

## Identification of individual channel kinetics from recordings containing many identical channels<sup>☆</sup>

Graham Pulford<sup>a</sup>, Rodney A. Kennedy<sup>b</sup>, Shin-Ho Chung<sup>a,\*</sup>

<sup>a</sup>*Protein Dynamics Unit, Department of Chemistry, The Australian National University, Canberra, A.C.T. 0200, Australia*

<sup>b</sup>*Department of Engineering, Faculty of Engineering, The Australian National University, Canberra, A.C.T. 0200, Australia*

Received 11 October 1993; revised 13 January 1994, 27 May 1994 and 14 November 1994

---

### Abstract

Given a discrete-time signal consisting of  $N$  identical, independent, binary Markov chains observed in white noise, we consider the problem of estimating the non-zero state level, the number of chains and the elementary transition probability matrix. We derive formulae for the central moments, first- and second-order auto-correlation functions and the power spectrum of a first-order, discrete-time Markov chain. We show that the mean, variance, third central moment and power spectrum provide sufficient information for the estimation of the parameters of the signal in question. We demonstrate the estimation procedure with numerical examples for both simulated and real biological data, and describe a method for estimating the non-unity eigenvalue of the transition matrix as well as the noise variance from the power spectrum of the noisy signal.

### Zusammenfassung

Gegeben sei ein zeitdiskretes Signal, daß aus  $N$  gleichartigen, unabhängigen, binären Markov-Ketten besteht und in weißes Rauschen eingebettet ist. Wir betrachten das Problem der Schätzung des Niveaus des Nicht-Nullzustandes, der Anzahl der Ketten und der elementaren Matrix der Übergangswahrscheinlichkeiten. Wir leiten Formeln her für die zentralen Momente, die Autokorrelationsfunktionen erster und zweiter Ordnung und das Leistungsspektrum einer zeitdiskreten Markov-Kette erster Ordnung. Wir zeigen, daß der Mittelwert, die Varianz, das dritte Zentralmoment und das Leistungsspektrum genügend Information zur Schätzung der Parameter des fraglichen Signals bereitstellen. Wir demonstrieren die Schätzprozedur anhand von numerischen Beispielen mit sowohl simulierten als auch realen biologischen Daten und beschreiben eine Methode zur Schätzung des Nicht-Einheitseigenwertes der Übergangsmatrix sowie der Rauschvarianz aus dem Leistungsspektrum des verrauschten Signals.

### Résumé

Pour un signal à temps discret donné composé de  $N$  chaînes de Markov binaires indépendantes et identiques observé dans du bruit blanc, nous considérons le problème de l'estimation du niveau d'état non nul, du nombre de chaînes et de la matrice de probabilités de transition. Nous dérivons des formules pour les moments centrés, pour les fonctions d'auto-corrélation du premier et du second ordre, et pour le spectre de puissance d'une chaîne de Markov à temps discret

---

<sup>☆</sup>This work was supported in part by grants from the Ramaciotti Foundation and NH&MRC of Australia.

\*Corresponding author. Fax: (06)247 2792.

du premier ordre. Nous montrons que la moyenne, la variance, le moment centré d'ordre trois, et le spectre de puissance fournissent une information suffisante pour l'estimation des paramètres du signal en question. Nous illustrons la procédure d'estimation avec des exemples numériques provenant de données simulées et de données biologiques réelles, et décrivons une méthode d'estimation de la valeur propre non unitaire de la matrice de transition et de la variance du bruit à partir du spectre de puissance du signal bruité.

**Keywords:** Markov chain; Parameter estimation; Ionic channel; Whole-cell recording

---

## 1. Introduction

Since the advent of the patch-clamp technique for monitoring drug-induced activity in cell membrane channels, there has been considerable interest in identifying probabilistic models of the kinetics of these systems [1, 2, 11, 22, 29, 32]. In the biological sciences, these parameter estimation techniques are commonly referred to as noise analysis or fluctuation analysis. Patch clamping allows the current due to a *single* ion channel to be observed. In contrast, our work is motivated by the possibility of measuring the membrane current due to a *whole cell* comprising many identical channels and hence estimating the parameters of a Markov model for the individual channels. Moreover, the present approach leads to an identification procedure that is computationally efficient and practically viable.

Our approach is to regard the net channel recording as a linear superposition of an unknown, possibly large number of identical and independent elementary channels, or pores, observed in white noise. Each elementary channel under the influence of an introduced drug (chemical transmitter) is modelled as a discrete-time, binary, homogeneous Markov chain. The effect of the drug is to tend to open the pores but the action is in fact random in the sense that pores open and close in a non-deterministic manner. This gives justification to the use of a Markov chain model. We assume initially that the variance of the noise can be determined a priori from a control data segment and also that the closed current level due to each channel is zero. We then go on to derive a practicable statistical procedure which identifies the individual open channel current, the matrix of four transition probabilities and the number of channels present in the record.

We later relax the assumption of a known noise variance, showing how this can be estimated from the power spectrum of the noisy data. We point out that although techniques such as hidden Markov modelling [8, 9] could in principle be applied to this multi-channel estimation problem, in practice the computational burden would be excessive when the assumed number of channels is large.

Fluctuation analysis has been fruitfully exploited in biophysics. Utilizing the signal mean and variance, and the corner frequency of the power spectrum, it has been possible to estimate the conductance, the mean open duration and the number of single channels contributing to the record [1, 32]. Previous estimation procedures, however, relied on the assumption that the probability of a channel being in the open state was small. The 'low concentration limit' used by Anderson and Stevens [1] is such an example. This assumption may not be realistic in some circumstances, and consequently we have extended the estimation technique so that it can be applied without the need to make unnecessary assumptions on the behaviour of the system. Further discussions concerning these and other modelling assumptions may be found in [10].

We deal in Section 2 with the computation of statistics of finite-state, discrete-time Markov chains. For further details of this theory, the reader is referred to [4, 12, 17, 19]. We obtain expressions for the moments, second- and third-order correlation functions and cumulants, and the power density spectrum. The computations are quite straightforward but have been included since the treatment of the discrete-time case does not appear to be well documented in the literature. The computation of power spectra of Markov chains has been addressed in the communications literature

on coded modulation [3, 5, 6], but the emphasis is on Markov chains arising from shift register processes, and these tend to have a special structure. We present our results in a more accessible form. Our calculations for the moments and spectra extend easily to the case of memoryless functions of a Markov chain [14, 15] so that the theory may be applied to chains with aggregated states such as discrete-time Colquhoun and Hawkes models, although in such cases it is not clear how to solve the parameter estimation problem. We show how to modify the results when the observed process consists of a sum of independent and identical chains with additive noise.

Section 3 is concerned with the problem of identifying the parameters of a superposition of binary Markov chains. We solve this problem for the case of additive white noise of known variance using the theory developed in Section 2. In Section 4 we discuss the estimation of moment statistics and spectra, indicating how to estimate the non-unity eigenvalue of the transition matrix and the noise variance from measurements of the power spectrum. Numerical examples are presented in Section 5 for simulated as well as real biological channel data, followed by a discussion of some possible extensions of the work.

## 2. Mathematical preliminaries

We will concern ourselves in this part with the computation of statistics of finite-state, discrete-time Markov chains, namely, the central moments of arbitrary order, the second- and third-order correlations and cumulants and the power spectrum. We treat the case of a single, homogeneous, regular chain with  $n$  states first, then specialise to the 2-state case. We present results which are most relevant to solving the inverse multi-channel problem of Section 3. We go on to consider superpositions of  $N$  identical, independent, 2-state chains observed in white noise. We demonstrate that only the first three moments give non-redundant information regarding the process parameters, and that these moments are most economically expressed in terms of the steady-state distribution.

### 2.1. Moments of finite-state Markov chains

We consider an  $n$ -state, homogeneous, discrete-time, scalar Markov chain  $\{x_k\}_{k \geq 0}$  with state space  $\mathcal{S} = \{s_1, \dots, s_n\} \subset \mathbb{R}^n$ . The stochastic process  $\{x_k\}$  is fully specified once we know the initial probability distribution

$$\pi_0 = [\Pr(x_0 = s_1), \dots, \Pr(x_0 = s_n)]$$

and the transition probability matrix

$$\mathbf{P} = (p_{ij}), \quad i, j = 1, \dots, n,$$

where the transition probabilities are defined as

$$p_{ij} = \Pr(x_{k+1} = s_j | x_k = s_i).$$

In what follows, we will assume that  $\{x_k\}$  is a *regular* Markov chain in the sense that a unique stationary distribution (or steady state distribution)  $\pi$  with components  $\pi_i$  exists which satisfies

$$\pi = \pi \mathbf{P} \tag{1}$$

subject to the constraint  $\sum_{i=1}^n \pi_i = 1$ . This guarantees that the dominant eigenvalue of  $\mathbf{P}$  is unity and that all other eigenvalues lie inside the unit circle in the complex plane.

Note that a sufficient condition for this to hold is that all the entries of  $\mathbf{P}$  be positive. Equivalently, we may write [17, p. 89]

$$\pi = \text{any row of } \bar{\mathbf{P}}, \quad \bar{\mathbf{P}} = \lim_{k \rightarrow \infty} \mathbf{P}^k.$$

Furthermore, we will only be interested in obtaining the steady-state statistical properties of  $\{x_k\}$ , and so we ignore the initial distribution  $\pi_0$ . Thus we may consider  $\{x_k\}$  to be defined by its transition probability matrix  $\mathbf{P}$  and state space  $\mathcal{S}$ . We turn to the calculation of the moments.

The  $q$ th order moment of a stationary stochastic process  $\{x_k\}$  is defined as

$$m_q(x) = E\{x_k^q\}, \quad q = 1, 2, \dots,$$

and the  $q$ th order central moment as

$$\begin{aligned} \mu_q(x) &= E\{(x_k - E\{x_k\})^q\}, \quad q = 2, 3, \dots \\ &= \sum_{i=0}^q \binom{q}{i} (-m_1(x))^{q-i} m_i(x). \end{aligned} \tag{2}$$

The last line follows from the binomial theorem, with  $m_0 = 1$ . So that, generally,  $\mu_q(x)$  is expressible in terms of  $\{m_i(x), i = 1, \dots, q\}$ . When  $\{x_k\}$  is a discrete-time Markov chain it follows from the definition that, in the steady state,

$$\begin{aligned} m_q(x) &= \sum_{i=1}^n s_i^q \Pr(x_k = s_i) \\ &= \sum_{i=1}^n s_i^q \pi_i = \boldsymbol{\pi} \mathbf{S}^q \mathbf{1}, \end{aligned} \quad (3)$$

where  $\boldsymbol{\pi}$  is given by (1),  $\mathbf{S} = \text{diag}[s_1, \dots, s_n]$  and  $\mathbf{1} = [1, \dots, 1]^T \in \mathbb{R}^n$  ( $\mathbf{T}$  denotes transposition).

It is clear from (3) that the moments of all orders of the Markov chain  $\{x_k\}$ , and hence its first-order probability distribution or, equivalently, its histogram, are specified once the state levels  $\mathcal{S}$  and stationary distribution  $\boldsymbol{\pi}$  are known. In other words, any two chains sharing the same state space and steady-state probabilities have the same probability distributions, although they may have differing correlations. Since  $\boldsymbol{\pi}$  has  $n - 1$  independent entries and there are  $n$  state levels  $s_i$ , from the point of view of identification this means that at most  $2n - 1$  independent equations involving these parameters may be obtained from the set of moment equations of the process. This fact has important consequences when we come to solving the multi-pore problem in Section 3.

We explained that an  $n$ -state Markov chain can be regarded as specified by the  $n^2 + n$  quantities comprising  $\mathbf{P}$  and  $\mathcal{S}$ . Now, since  $\mathbf{P}$  is a stochastic matrix, only  $n(n - 1) + n = n^2$  of these parameters may be independent. Thus, in general, whenever there are two or more process states, there is an excess of  $n^2 - (2n - 1) = (n - 1)^2$  independent quantities which are not fixed by specifying the moments alone. For this reason, we will need to consider ‘correlation-type’ information, and we present this next.

## 2.2. Correlations and cumulants of Markov chains

In this section we derive expressions for the second- and third-order correlations and cumulants of the  $n$ -state Markov chain introduced in the last section. The calculations for the correlation

functions are roughly analogous to the continuous-time case treated by Fredkin and Rice [15] whose matrix notation we have borrowed.

The second-order correlation function of a homogeneous Markov chain  $\{x_k\}$  is defined as

$$R_x(k_1) = E\{x_k x_{k+k_1}\} \quad (4)$$

and the third-order correlation by

$$R_x(k_1, k_2) = E\{x_k x_{k+k_1} x_{k+k_2}\}, \quad (5)$$

whereas the second- and third-order cumulants are defined, respectively, as

$$C_x(k_1) = E\{(x_k - m_1(x))(x_{k+k_1} - m_1(x))\},$$

$$\begin{aligned} C_x(k_1, k_2) &= E\{(x_k - m_1(x))(x_{k+k_1} - m_1(x)) \\ &\quad \times (x_{k+k_2} - m_1(x))\}, \end{aligned}$$

where  $m_1(x)$  is the mean of  $\{x_k\}$ . The second-order correlation is known as the auto-correlation and the second-order cumulant as the autocovariance. We look at the case of  $k_2 \geq k_1 \geq 0$  first. We have, with the same notation as before,

$$R_x(k_1) = \boldsymbol{\pi} \mathbf{S}^{\mathbf{P}^{k_1}} \mathbf{S} \mathbf{1}, \quad k_1 \geq 0, \quad (6)$$

$$R_x(k_1, k_2) = \boldsymbol{\pi} \mathbf{S}^{\mathbf{P}^{k_1}} \mathbf{S}^{\mathbf{P}^{k_2 - k_1}} \mathbf{S} \mathbf{1}, \quad k_2 \geq k_1 \geq 0, \quad (7)$$

which can be derived by using the fact that the  $k$ -step transition probabilities  $p_{ij}^{(k)}$  are the entries of  $\mathbf{P}^k$ . By considering all six possible orderings of the quantities  $0, k_1, k_2$ , it is straightforward to show that the correlations for arbitrary lags are given by

$$R_x(k_1) = \boldsymbol{\pi} \mathbf{S}^{\mathbf{P}^{|\kappa_1|}} \mathbf{S} \mathbf{1}, \quad (8)$$

$$R_x(k_1, k_2) = \boldsymbol{\pi} \mathbf{S}^{\mathbf{P}^{\kappa_1}} \mathbf{S}^{\mathbf{P}^{\kappa_2}} \mathbf{S} \mathbf{1}, \quad (9)$$

where the non-negative lags  $\kappa_i$  are given by

$$\kappa_1 = \text{med}\{0, k_1, k_2\} - \min\{0, k_1, k_2\}, \quad (10)$$

$$\kappa_2 = \max\{0, k_1, k_2\} - \text{med}\{0, k_1, k_2\}, \quad (11)$$

and  $\text{med}$  is the statistical median.

To proceed further, we need to find a spectral representation for  $\mathbf{P}$ . If  $\mathbf{P}$  has distinct eigenvalues, this is furnished by

$$\mathbf{P} = \sum_{i=1}^n \lambda_i \mathbf{A}_i, \quad (12)$$

in which  $\{\lambda_i\}$  are the eigenvalues of  $\mathbf{P}$  and the projector matrices  $\mathbf{A}_i$  are generated according to

$$\mathbf{A}_i = \mathbf{m}_i \mathbf{n}_i, \quad (13)$$

where  $\mathbf{m}_i$  and  $\mathbf{n}_i$  are, respectively, the right and left eigenvectors of  $\mathbf{P}$  corresponding to  $\lambda_i$ . Note that the  $\mathbf{A}_i$  matrices satisfy

$$\mathbf{A}_i \mathbf{A}_j = \mathbf{A}_i \delta_{ij}, \quad \sum_{i=1}^n \mathbf{A}_i = \mathbf{I}_n, \quad (14)$$

where  $\delta_{ij}$  is the Kronecker delta and  $\mathbf{I}_n$  is the unit matrix of order  $n$ . It follows that

$$\mathbf{P}^k = \sum_{i=1}^n \lambda_i^k \mathbf{A}_i, \quad k \geq 0. \quad (15)$$

If the eigenvalues of  $\mathbf{P}$  are not distinct, we must use a Jordan representation for  $\mathbf{P}$  and the computations become more involved [6, 13, 17]. We do not treat the repeated eigenvalue case in this paper. Assuming then that (15) is valid, (8) and (9) simplify to

$$R_x(k_1) = \sum_{i=1}^n r_i \lambda_i^{k_1}, \quad (16)$$

$$R_x(k_1, k_2) = \sum_{i=1}^n \sum_{j=1}^n r_{ij} \lambda_i^{k_1} \lambda_j^{k_2},$$

where we have defined

$$r_i = \boldsymbol{\pi} \mathbf{S} \mathbf{A}_i \mathbf{S} \mathbf{1}, \quad i = 1, \dots, n, \quad (17)$$

$$r_{ij} = \boldsymbol{\pi} \mathbf{S} \mathbf{A}_i \mathbf{S} \mathbf{A}_j \mathbf{S} \mathbf{1}, \quad i, j = 1, \dots, n.$$

We can simplify these expressions even further by noting that, due to the assumption of irreducibility of  $\mathbf{P}$ , the unity eigenvalue, say  $\lambda_1$ , is unique and has associated projector  $\mathbf{A}_1 = \mathbf{1}\boldsymbol{\pi}$ . Therefore, we have

$$r_1 = \boldsymbol{\pi} \mathbf{S} \mathbf{A}_1 \mathbf{S} \mathbf{1} = (\boldsymbol{\pi} \mathbf{S} \mathbf{1})^2 = m_1^2(x) \quad (18)$$

and also

$$r_{1j} = \boldsymbol{\pi} \mathbf{S} \mathbf{1} \boldsymbol{\pi} \mathbf{S} \mathbf{A}_j \mathbf{S} \mathbf{1} = m_1(x) r_j, \quad (19)$$

$$r_{i1} = \boldsymbol{\pi} \mathbf{S} \mathbf{A}_i \mathbf{S} \mathbf{1} \boldsymbol{\pi} \mathbf{S} \mathbf{1} = m_1(x) r_i = r_{1i},$$

as obtained by Fredkin and Rice [15] in the continuous-time case. The autocovariance is now expressible as

$$C_x(k_1) = \sum_{i=2}^n r_i \lambda_i^{k_1}, \quad (20)$$

where  $\lambda_2, \dots, \lambda_n$  are the distinct, non-unity eigenvalues of the transition matrix. Furthermore, the third-order cumulant is given by

$$C_x(k_1, k_2) = R_x(k_1, k_2) - m_1(x) \{R_x(k_1) + R_x(k_2) + R_x(k_1 - k_2)\} + 2m_1^3(x). \quad (21)$$

In particular, the variance and third-order central moment are given, respectively, by

$$\mu_2(x) = C_x(0) = \boldsymbol{\pi} \mathbf{S} (\mathbf{I} - \mathbf{A}_1) \mathbf{S} \mathbf{1}, \quad (22)$$

$$\mu_3(x) = C_x(0, 0) = \boldsymbol{\pi} \mathbf{S} (\mathbf{S} - 3\mathbf{A}_1 \mathbf{S} + 2\mathbf{A}_1 \mathbf{S} \mathbf{A}_1) \mathbf{S} \mathbf{1}, \quad (23)$$

where  $\mathbf{A}_1 = \mathbf{1}\boldsymbol{\pi}$  and  $m_1(x) = \boldsymbol{\pi} \mathbf{S} \mathbf{1}$  is the process mean. These formulae will be of use in Section 3.

### 2.3. Power spectrum of a Markov chain

The power spectrum computation for the continuous-time case and discrete-time case may be found, respectively, in [32] and [6, 21]. However, especially in the communications literature, the calculations focus on transition matrices with special structure, corresponding to, for instance, coded signals. This makes it hard to interpret the results when arbitrary transition matrices are involved.

The power spectrum  $S_x(\omega)$  is the Fourier transform of the autocovariance, or, equivalently, its Z-transform evaluated on the unit circle [25], viz,

$$S_x(\omega) = Z\{C_x(k)\}_{z=e^{j\omega T}}. \quad (24)$$

Here,  $T$  represents the sampling interval. From (20) we have

$$S_x(\omega) = \sum_{i=2}^n r_i Z\{\lambda_i^{k_1}\}_{z=e^{j\omega T}}, \quad (25)$$

which suggests that we need to compute the following Z-transform:

$$\begin{aligned} Z\{\lambda^{k_1}\} &= \sum_{k=-\infty}^{\infty} z^{-k} \lambda^{|k|} \\ &= \frac{1 - \lambda^2}{(1 - \lambda z^{-1})(1 - \lambda z)}, \end{aligned} \quad (26)$$

which converges whenever  $|\lambda| < |z| < 1/|\lambda|$  and  $|\lambda| < 1$ . Noting that, since  $\mathbf{P}$  is the transition matrix

of a regular Markov chain,  $|\lambda_i| < 1$  for  $i \neq 1$  in (25), we see that

$$S_x(\omega) = \sum_{i=2}^n \frac{(1 - \lambda_i^2)r_i}{(1 - \lambda_i e^{-j\omega T})(1 - \lambda_i e^{j\omega T})}, \quad (27)$$

where  $r_i$  was defined in (17). We note that  $S_x(\omega)$  could also be generated by passing white noise through a linear auto-regressive moving-average (ARMA) filter with appropriately chosen coefficients. Let us briefly examine the structure of (27). Clearly  $S_x(\omega)$  is a rational function of  $\cos(\omega T)$  (as is the case for discrete-time linear systems corresponding to ARMA processes). Suppose we partition the set of eigenvalues of  $\mathbf{P}$  in the following way:

$$\lambda_1 = 1, \quad \lambda_2, \dots, \lambda_r \in \mathbb{R}, \quad (28)$$

$$\lambda_{r+i}, \lambda_{r+i}^*, \dots, \lambda_{r+c}, \lambda_{r+c}^* \in \mathbb{C},$$

in which  $\lambda^*$  is the complex conjugate of  $\lambda$  and  $r + 2c = n$ . Then we may write (27) as

$$S_x(\omega) = \sum_{i=2}^r \frac{(1 - \lambda_i^2)r_i}{1 + \lambda^2 - 2\lambda_i \cos \omega T} + 2 \sum_{i=1}^c \operatorname{Re} \left\{ \frac{(1 - \lambda_{r+i}^2)r_{r+i}}{(1 - \lambda_{r+i} e^{-j\omega T})(1 - \lambda_{r+i} e^{j\omega T})} \right\}, \quad (29)$$

in which  $\operatorname{Re}(\lambda)$  denotes the real part of  $\lambda$ . This form makes apparent the components of the spectrum.

In the biological literature, the power spectrum of binary Markov chains observed in white noise is fitted with a Lorentzian distribution. The terms corresponding to the real eigenvalues,

$$\frac{(1 - \lambda_i^2)r_i}{1 + \lambda_i^2 - 2\lambda_i \cos \omega T},$$

can be compared to their Lorentzian counterparts for the continuous-time case, namely

$$\frac{2r_i \lambda_i}{\lambda_i + (\omega T)^2}.$$

The two components have similar behaviour for  $\omega T \ll 1$ , but differ as  $\omega$  approaches the Nyquist frequency  $1/2T$ .

As an example which will be important later, we consider a regular Markov chain which is binary

i.e., it has only two possible states  $s_1$  and  $s_2$ . Let the transition probability matrix be parametrised as

$$\mathbf{P} = \begin{bmatrix} \zeta & \bar{\zeta} \\ \bar{\rho} & \rho \end{bmatrix}, \quad (30)$$

where  $\bar{\zeta} = 1 - \zeta$  and  $\bar{\rho} = 1 - \rho$ , and  $\zeta$  and  $\rho$  are not both zero. The eigenvalues of  $\mathbf{P}$  are

$$\lambda_1 = 1, \quad \lambda_2 \triangleq \lambda = \zeta + \rho - 1 = 1 - \bar{\zeta} - \bar{\rho}. \quad (31)$$

Note that  $-1 < \lambda < 1$ . The steady-state distribution is easily shown to have components

$$\pi_1 = \frac{\bar{\rho}}{\bar{\rho} + \bar{\zeta}}, \quad \pi_2 = \frac{\bar{\zeta}}{\bar{\rho} + \bar{\zeta}}. \quad (32)$$

The spectral expansion of  $\mathbf{P}$  is given by

$$\mathbf{P} = \mathbf{A}_1 + \lambda \mathbf{A}, \quad (33)$$

where

$$\mathbf{A}_1 = \mathbf{1}\boldsymbol{\pi} = \begin{bmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \pi_2 & -\pi_2 \\ -\pi_1 & \pi_1 \end{bmatrix}.$$

This holds for any  $2 \times 2$  stochastic  $\mathbf{P}$  since a repeated eigenvalue can only occur if  $\mathbf{P} = \mathbf{I}$ .

Suppose now that one state level, e.g.,  $s_1$ , in the binary chain may be taken as zero. This can be arranged by shifting the origin in the state space. In this case, the mean, variance and third-central moment can be written as

$$m_1(x) = \pi_2 s_2, \quad \mu_2(x) = \pi_1 \pi_2 s_2^2, \quad (34)$$

$$\mu_3(x) = \pi_1 \pi_2 (\pi_1 - \pi_2) s_2^3,$$

while the spectrum may be expressed as

$$S_x(\omega) = \frac{(1 - \lambda^2)\mu_2(x)}{1 + \lambda^2 - 2\lambda \cos \omega T}, \quad (35)$$

where  $\lambda$  is the real, non-unity eigenvalue of  $\mathbf{P}$ . It is possible to deduce a simple expression for the  $q$ th-order central moment in this case, and this may be found in Appendix A.

#### 2.4. Superpositions of independent binary chains

We now indicate how to compute the central moments and power spectrum of a process consisting

of  $N$  independent, binary Markov chains. We make the simplifying assumption that all chains are identical, so that they possess the same state space and transition matrix. For some results concerning superpositions of non-identical Markov chains, see [16, 33]. Specifically, the process is represented by

$$x_k = \sum_{i=1}^N x_k^{(i)}, \quad (36)$$

where the  $\{x_k^{(i)}\}$  are independent and identically distributed (iid), binary Markov chains (as treated in the preceding section). Since the chains are identical, we cannot distinguish states of  $\{x_k\}$  in which equal numbers of elementary chains are in (say) the open state. Therefore,  $\{x_k\}$  has a state space with  $N + 1$  elements called *aggregated states*. The Markov property is preserved by this aggregation because all the chains  $\{x_k^{(i)}\}$  are *identical*.

The transition probability matrix  $\mathbf{P}_{\text{agg}}$  of  $\{x_k\}$  can be computed as

$$\mathbf{P}_{\text{agg}} = \mathbf{L} \underbrace{(\mathbf{P} \otimes \mathbf{P} \otimes \cdots \otimes \mathbf{P})}_{N} \mathbf{R}, \quad (37)$$

where  $\mathbf{L}$  and  $\mathbf{R}$  are aggregation matrices [18, 20] and  $\otimes$  is the Kronecker product. We could compute the moments and power spectrum directly using this  $(N + 1) \times (N + 1)$  transition matrix. This is unnecessary, however, as the moments and spectrum of the process are clearly  $N$  times their values for a single chain, when the chains are independent, as described in Sections 2.2 and 2.3.

Finally in our model we assume that we only have access to noisy observations  $\{y_k\}$  of the process (36),

$$y_k = x_k + n_k, \quad (38)$$

where  $\{n_k\}$  is a sequence of white Gaussian random variables with zero mean and variance  $\sigma^2$  independent of the signal process  $\{x_k\}$ . The mean and variance of  $\{y_k\}$  are given in terms of the mean and variance of  $\{x_k\}$  by

$$m_1(y) = m_1(x), \quad \mu_2(y) = \mu_2(x) + \sigma^2. \quad (39)$$

Moreover, and less obviously, the third-order central moment, like the mean, is unaffected by the noise and we have  $\mu_3(y) = \mu_3(x)$ . This holds whenever the noise density is symmetric about its mean

value since all odd-order cumulants of  $\{y_k\}$  are the same as those of  $\{x_k\}$  [24]. We point out that the same cannot be said for odd-order moments about the origin, e.g.,  $m_3(y) \neq m_3(x)$  in general. This simplification justifies the use of central moments in the calculations. The spectrum of  $\{y_k\}$  is given by

$$S_y(\omega) = S_x(\omega) + \sigma^2, \quad (40)$$

since the noise is white and has a flat spectrum.

### 3. Multi-channel inverse problem

Up to this point we have characterised, in terms of various statistical quantities, the superposition of a set of independent binary Markov chains. This model is a close approximation of transmembrane currents which arise from the collective behaviour of many single channels that open and close intermittently. Such situations can occur when membrane currents are measured with intracellular electrodes, from whole-cell configurations or excised patches containing multiple channels.

We can now state the following problem, called the *multi-pore problem* (MPP):

**Problem 3.1.** Given observations of a process  $\{y_k\}$  consisting of an addition of iid, binary, homogeneous Markov chains, observed in additive white Gaussian noise, determine the number of binary chains, the two elementary state levels and the transition probability matrix of the individual chains in terms of statistically measurable quantities.

Initially, we will not attempt to solve the MPP in this generality, instead, we make the following additional assumptions:

- (A1) The variance of the noise is known.
- (A2) One state level,  $s_1$ , of the identical binary chains is zero.

Assumption (A2) may be made without loss of generality. We show in Section 4 how to estimate the noise variance and non-unity eigenvalue of the single-chain transition matrix using the power spectrum of the noisy data, so that assumption (A1) should not be an impediment in practice. We stress here that, although the general problem can be solved in principle using hidden Markov model

techniques, it is impractical to do so when the number of individual chains is large. In addition, the transition matrix of an individual chain  $\mathbf{P}$  must be obtained from (37), and this involves the solution of an overdetermined system of non-linear equations. We now proceed to solve the MPP subject to assumptions (A1) and (A2), using the theory developed in Section 2.

Let the observable process be written in the usual notation as

$$y_k = \sum_{i=1}^N x_k^{(i)} + n_k. \quad (41)$$

From Sections 2.1–2.3 it follows that the mean, variance, third-order central moment and power spectrum of  $\{y_k\}$  are given, respectively, by

$$m_1(y) = Nm_1(x^{(i)}) = N\pi_2 s_2, \quad (42)$$

$$\mu_2(y) = N\mu_2(x^{(i)}) + \sigma^2 = N\pi_1\pi_2 s_2^2 + \sigma^2, \quad (43)$$

$$\mu_3(y) = N\mu_3(x^{(i)}) = N\pi_1\pi_2(\pi_1 - \pi_2)s_2^3 \quad (44)$$

and

$$S_y(\omega) = NS_{x^{(i)}}(\omega) + \sigma^2 \\ = \frac{N\pi_1\pi_2 s_2^2(1 - \lambda^2)}{1 + \lambda^2 - 2\lambda \cos \omega T} + \sigma^2. \quad (45)$$

The fourth central moment of  $\{y_k\}$  may be shown to be

$$\mu_4(y) = N\pi_1\pi_2(\pi_1^3 + \pi_2^3)s_2^4 + 6N\pi_1\pi_2\sigma^2 s_2^2 + 3\sigma^4.$$

As before, if  $\mathbf{P}$  is the transition matrix of  $\{x_k^{(i)}\}$ ,  $\pi_1$ ,  $\pi_2$  are the components of the steady-state distribution of  $\mathbf{P}$ ,  $\lambda$  is the non-unity eigenvalue and  $s_2$  is the non-zero state level. Note that, in the absence of noise,

$$\mu_2(y) = s_2 m_1(y) - m_1^2(y)/N,$$

a similar result holds for Bernoulli processes for which  $\zeta = 1 - \rho$  in (30), as obtained by Sigworth [31].

The unknowns in this problem are  $s_2$ ,  $\lambda$ ,  $\pi_1$ ,  $\pi_2$ ,  $N$  and the transition matrix itself, which may be determined from the spectral expansion (33) once  $\lambda$ ,  $\pi_1$ ,  $\pi_2$  are known. Of course,  $\pi_1 + \pi_2 = 1$ , so we can reduce the number of unknown parameters to four. We emphasize again here that no further information relevant to determining the transition

probabilities can be gained by considering central moments of order higher than 3. (This can be established using the results of Appendix A.) As mentioned in Section 2, the moments provide only  $2 \times 2 - 1 = 3$  independent equations, so that correlation-type information must be considered (or equivalently spectra).

With Eqs. (42)–(45) in hand, it is a straightforward matter to identify the desired parameters. Since the noise variance is assumed to be known, we can replace  $\mu_2(y)$  by

$$\mu_2(x) = \mu_2(y) - \sigma^2$$

and subtract the noise power from the spectrum to obtain the noiseless signal spectrum

$$S_x(\omega) = S_y(\omega) - \sigma^2.$$

It follows simply from Eqs. (42)–(44) that the unique solution for  $(\pi, s_2, N)$  is

$$\pi_1 = (2 - \gamma)^{-1}, \quad \pi_2 = 1 - \pi_1, \quad (46)$$

$$s_2 = \frac{\mu_2(x)}{m_1(y)\pi_1}, \quad N = \frac{\pi_1 m_1^2(y)}{\pi_2 \mu_2(x)},$$

where the dimensionless quantity  $\gamma$  is defined by

$$\gamma \triangleq \frac{m_1(y)\mu_3(y)}{(\mu_2(x))^2}. \quad (47)$$

Note that in practice  $N$  would need to be rounded to an integer value. Since  $\pi_1 \in (0, 1)$  for a Markov chain with no absorbing state, it follows that  $\gamma < 1$ . In practice, outliers in the data may need to be removed to ensure that the latter condition holds for the estimated moments.

Furthermore, we can obtain  $\lambda$  and hence the transition matrix from the spectrum. An analytical solution is given by

$$\lambda - 1 - 2(1 + v)^{-1}, \quad v \triangleq \left( \frac{S_x(0)}{S_x(\pi)} \right)^{1/2}, \quad (48)$$

although, from a numerical point of view, it is better to fit the analytical expression (45) to the spectrum estimate rather than relying on information at only two frequency points. The transition probabilities are determined via

$$\zeta = \pi_1 + \pi_2 \lambda, \quad \rho = \pi_2 + \pi_1 \lambda,$$

$$\bar{\zeta} = 1 - \zeta, \quad \bar{\rho} = 1 - \rho.$$

**Remarks.** (1) The above results indicate the dependence of the parameters on the moments and power spectrum of the process (41).

(2) Satisfactory performance of the algorithm depends on the implementation of the statistical moment and power spectrum estimators.

#### 4. Moment and spectrum estimators

We consider how to obtain estimates of the required statistics from a finite-length sample of noisy data  $\{y_k\}_{k=1}^K$  from the process (41). We assume that the noise has an effectively flat spectrum over the frequency range of interest  $(0, 1/2T)$ , where  $T$  is the sampling period, and that an initial estimate of its variance may be obtained prior to estimation of the parameters. With this initial estimate, a signal variance estimate can be calculated by subtracting the noise variance from the variance of the noisy signal. The noise variance may also be estimated independently from the power spectrum. We also assume implicitly that the process is stationary over the observation period.

The sample mean, and second and third central moments of the data are computed using standard formulae. It is well known that the standard estimators for the variance and third central moment can be de-biased by the appropriate scaling, but this bias is inconsequential for large  $N$ . Further remarks concerning the consistency of standard moment estimators may be found in, e.g., [26]. We mention that the skew of the process (41), defined as

$$\zeta \triangleq \frac{E\{(y_k - E\{y_k\})^3\}}{(E\{(y_k - E\{y_k\})^2\})^{3/2}}, \quad (49)$$

is expressible as

$$\zeta = \frac{1}{\sqrt{N}} \frac{\pi_1 - \pi_2}{\sqrt{\pi_1 \pi_2}} \quad (50)$$

in the absence of noise. Moreover, the skew is small when the components of the steady-state distribution are approximately equal, or  $\pi_1 \approx \pi_2$ . This has implications in terms of the number of data points needed to obtain good estimates of the third central moment, which becomes more difficult to estimate accurately as the number of chains  $N$  gets large, or

for chains with roughly equal diagonal transition probabilities. We later give an example for which both these conditions apply.

Power spectrum estimation is well covered in the literature, and details may be found in [7, 23, 27]. For the purposes of demonstrating the algorithm, segmental averaging of windowed periodograms of the data was found to be adequate and we describe this briefly. The data are divided into non-overlapping blocks, typically of size 1024 points. The data in each block are de-biased and windowed. The fast Fourier transform of the block is then computed and the periodogram formed. The running average of the periodogram over blocks is taken and yields the power spectrum estimate  $\hat{S}_y(\omega_i)$  for the  $N/2$  frequencies  $\omega_i, i = 0, \dots, N/2 - 1$ . This estimate is then normalised using the computed data variance.

The power spectrum estimate of the raw data should be examined graphically to ascertain initial estimates for the signal parameters. Let  $S_{\max}$  and  $S_{\min}$  be the maximum and minimum values of the power spectrum obtained by visual examination. Furthermore, let  $\omega_0$  represent the  $-3\text{dB}$  frequency or half-power point obtained by locating the frequency at which the power spectrum drops to half its maximum value, measured with respect to its minimum value. It then follows simply from (45) that  $\omega_0$  and the non-unity eigenvalue  $\lambda$  are related by

$$\cos(\omega_0) = \frac{2\lambda}{1 + \lambda^2}, \quad (51)$$

so that an estimate of  $\lambda$  may be obtained as

$$\lambda = \frac{1 - \sin(\omega_0)}{\cos(\omega_0)}. \quad (52)$$

Estimates of the signal variance  $V$  and noise variance  $\sigma^2$  can then be obtained using

$$V = \frac{1 - \lambda^2}{4\lambda} (S_{\max} - S_{\min}), \quad (53)$$

$$\sigma^2 = \frac{(1 + \lambda)^2 S_{\min} - (1 - \lambda)^2 S_{\max}}{4\lambda}. \quad (54)$$

These variance estimates should sum to give the variance of the noisy signal. The initial parameter estimates can be refined by fitting a function of the

form (45) using least squares or a recursive variant thereof. Further details may be found in [10]. Alternatively, the power spectrum fitting procedure could be replaced by a recursive estimator for the linear auto-regressive system having the same power spectrum (for white noise input) as the Markov chain signal. Finally, we point out that all of the operations involved in the estimation procedure can be performed sequentially over blocks of data.

## 5. Numerical examples

We first give some numerical simulation examples demonstrating the effectiveness of the multi-pore identification technique. We assume that the noise variance is known, and so only use power spectrum fitting to estimate the non-unity eigenvalue of the process. We chose the following parameters for the multi-pore signal: transition probabilities  $\zeta = 0.99$ ,  $\rho = 0.97$ ; non-zero state level  $s_2 = -5$ ; noise variance  $\sigma^2 = 25$ ; number of samples  $= 10^6$ . The data were rounded to three significant figures to simulate quantisation noise. The number of points was taken to be large so that the initial Markov chain transient did not bias the estimation. The cases we have listed correspond to channels with  $N = 10$  and  $N = 100$  pores. The exact statistics of the net signal resulting from the sum of the elementary pore signals, computed from (42)–(44), are shown in Table 1. Note that the non-unity eigenvalue of the elementary transition matrix is 0.96. The first-, second- and third-order statistics of each of the three data records were computed as described in Section 4, and these are displayed in Table 2.

Fig. 1(a) shows a 1000 point segment of noise-free data from the 10-chain simulation, with a sample of

Table 2  
Estimated statistics of multi-pore data for one run of length  $10^6$  points

$N$	Mean	Noise + signal variance	Third-order central moment	Eigenvalue
10	-12.5127	72.2969	-115.994	0.9612
100	-124.938	469.027	-1035.875	0.9578

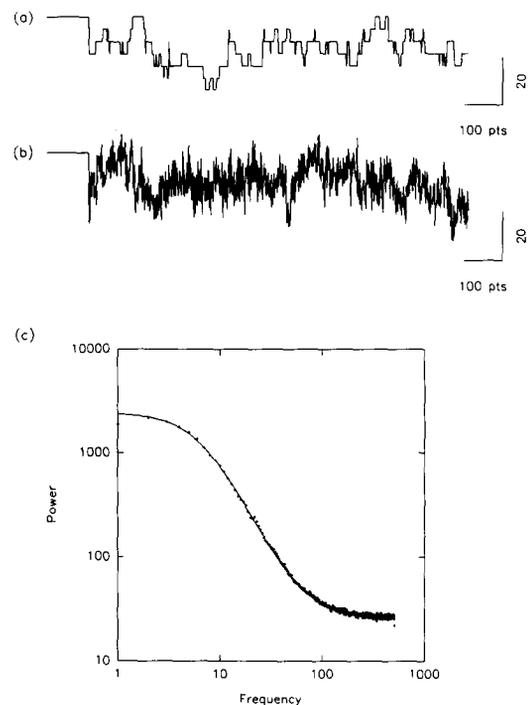


Fig. 1. Parameter identification from a record containing 10 chains. A segment of the noise-free signal and a segment with additive white noise are shown in (a) and (b). The estimated power spectrum, and corresponding fitted curve, are shown in (c).

Table 1  
Exact statistics of multi-pore data

$N$	Mean	Variance	Noise + signal variance	Third-order central moment
10	-12.5	46.875	71.875	-117.1875
100	-125	468.75	493.75	-1171.875

the noisy data displayed in Fig. 1(b). The 512-point power spectrum, obtained as described in Section 4, is displayed in Fig. 1(c). We used non-overlapping blocks of 1024 points with a Hanning window. Also shown in Fig. 1(c) is the fitted spectrum curve which was obtained using the method in [10]. Since the noise variance was assumed known, only the eigenvalue  $\lambda$  was required in the fit and the results are

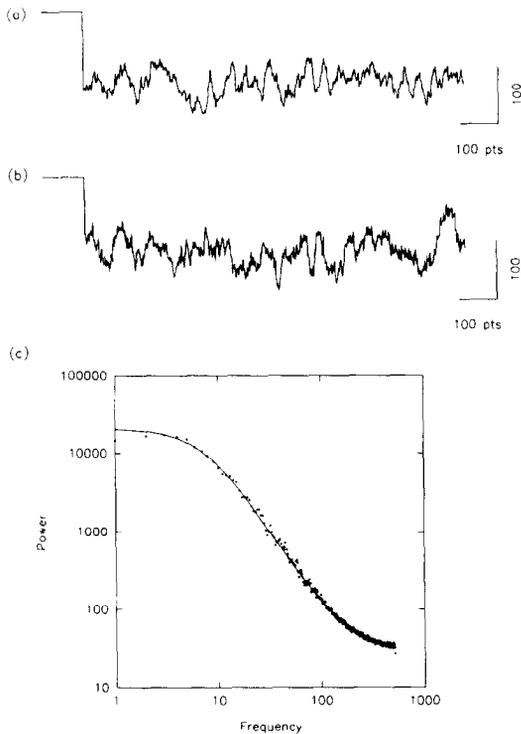


Fig. 2. Simulation for a large number of chains. A segment of the noise-free signal sequence, containing 100 identical chains is displayed in (a) and an observation sequence contaminated by white noise is shown in (b). The estimated power spectrum and fitted curve are shown in (c).

listed in Table 2. Similarly, a segment of the noiseless data for the 100-chain example is shown in Fig. 2(a), with a 1000-point sample of the noisy data in Fig. 2(b). The corresponding computed and fitted power spectra are given in Fig. 2(c).

Using the central moment estimates together with the estimated non-unity eigenvalue and the known noise variance, the elementary chain parameters were calculated from Eqs. (46) and the

values are given in Table 3. The quantities  $\hat{\pi}_i$  and  $\hat{p}_{ii}$  are the elements of the estimated steady-state distribution and elementary transition matrix. It is clear that these estimates are in good agreement with their true values. Note that the eigenvalue estimate  $\hat{\lambda}$  is only needed to estimate the transition probabilities, the other quantities being determined from the moment statistics and noise variance.

In order to demonstrate the consistency of the estimation algorithm we have included a set of Monte Carlo simulations. We chose the following parameters for the simulations: transition probabilities  $\zeta = 0.98$ ,  $\rho = 0.97$ ; non-zero state level  $s_2 = -25$ ; noise variance  $\sigma^2 = 625$ ; number of chains  $N = 100$ . In this example the skew (50) of the noiseless process is only  $O(10^{-3})$ , making the estimation of the third moment more difficult than for typical cases involving fewer chains or processes with more widely differing diagonal transition probabilities.

The experiment consisted of 100 Monte Carlo runs each for simulation lengths varying between 20000 and 500000 points. Note that theoretical values for the process statistics are in this case: mean  $-1000$ ; variance (signal plus noise) 15625; third central moment  $-75000$ ; non-unity eigenvalue 0.95. For each simulation, the parameters were computed from the estimated moments and power spectrum (assuming a known noise variance). A simple outlier detection strategy was adopted to improve the accuracy of the third moment estimator.

The mean parameter values and their standard deviations (depicted as error bars extending one standard deviation above and below the line) are plotted in Figs. 3(a)–(e) against the number of points in the simulations. The convergence of the parameter estimates is evident in these graphs. For this example, adequate performance was obtained

Table 3  
Estimated signal parameters from 10 and 100 pore simulations, assuming a known noise variance

	$\hat{N}$	$\hat{s}_2$	$\hat{\pi}_1$	$\hat{\pi}_2$	$\hat{\zeta}$	$\hat{\rho}$
True value	—	-5	0.75	0.25	0.99	0.97
$N = 10$	9	-5.11	0.74	0.26	0.9900	0.9713
$N = 100$	102	-4.77	0.7443	0.2557	0.9892	0.9686

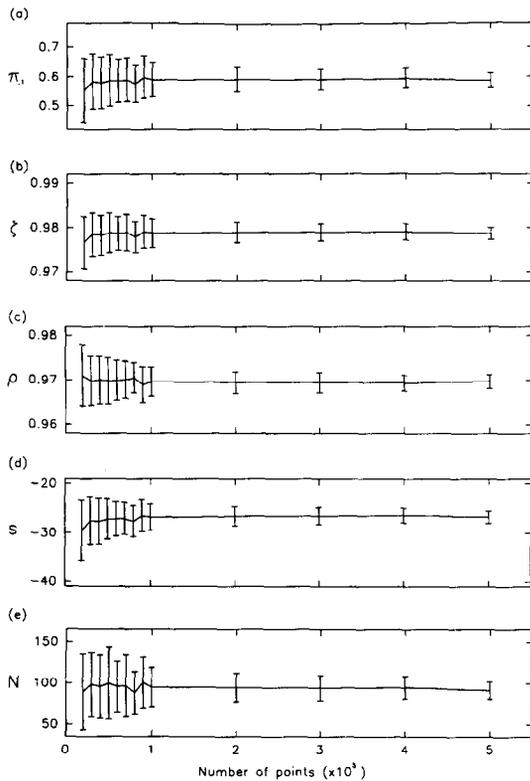


Fig. 3. Estimator performance as a function of data length. The mean value of the parameter estimates for steady-state distribution (a)  $\pi_1$  and transition probabilities (b)  $\zeta$  and (c)  $\rho$ , non-zero state level (d)  $s_2$  and number of chains (e)  $N$  are shown. The error bars have length equal to twice the standard deviation of the estimate.

for around 50 000 points, although for processes with larger values of skew, the data length can be reduced substantially. In typical biophysical experiments, membrane currents are recorded over a period of between 10 and 30 min and the entire record is digitised at 10 kHz. At this sampling rate, 50 000 points of data could be acquired in 5 s real time.

Lastly, we provide examples of the performance of the algorithm when applied to real data. Inward currents due to sodium ions, activated by a brief application of 100  $\mu\text{M}$  N-methyl-D-aspartate (NMDA) to a cultured hippocampal neuron, were recorded with a whole-cell patch clamp technique. Fig. 4 shows the current trace recorded immediately before, during and after the application of

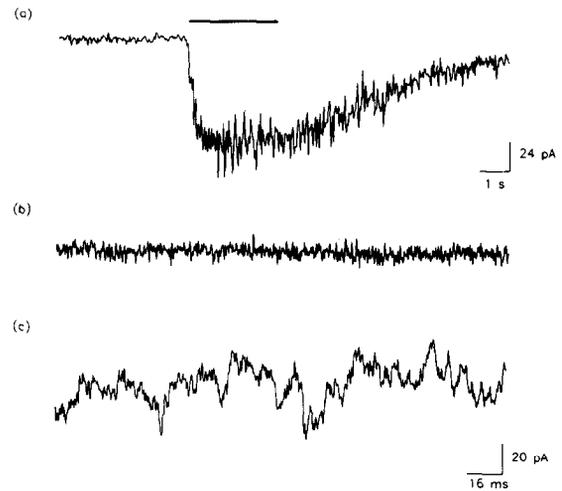


Fig. 4. Parameter estimation for real data. The current trace of a whole-cell patch clamp experiment is shown in (a). The 75 000-point record was passed through a median filter and then compressed into 500 points. The period during which 100  $\mu\text{M}$  N-methyl-D-aspartate (NMDA) was applied to the cell surface is indicated by the bar above the current trace. Short segments corresponding to activity prior to, and during the application of the drug (NMDA) are shown in (b) and (c) respectively on a larger time scale. The record was filtered at 2 kHz and sampled at 5 kHz.

NMDA. An unknown number of channels opened and closed intermittently while NMDA was being applied to the cell surface, as indicated by a bar just above Fig. 4(a). When the application of the channel agonist ceased, the magnitude of inward currents slowly decayed to the baseline (the zero-current level), reflecting a slow diffusion of agonist molecules from the cell surface. Figs. 4(b) and (c) show the current before and during the application of NMDA on an expanded time scale. Although the process is clearly non-stationary during the period in which the system relaxes back to the equilibrium, it may be considered *stationary* during the application of NMDA. Using 10 000-points of the record taken during the stationary period, we estimated the relevant parameters of the single NMDA channels. The estimated number of channels activated during the peak of inward currents was 77 with an estimated conductance of 51.2 pS (corresponding to an amplitude of  $-2.56$  pA). The estimated diagonal transition probabilities were  $\zeta = 0.9876$  and

$\rho = 0.985$ . These results are in accordance with generally accepted experimental values.

## 6. Discussion

We have presented a technique for the identification of parameters of a signal consisting of a superposition of identical, independent, binary Markov chains observed in white noise. This was based on the computation of the mean, variance, third-order central moment and power spectrum of the process data. The technique is simpler than other methods based on hidden Markov models or histograms and for this reason is better suited to the estimation of single channel characteristics given data arising from a large number of identical channels. In addition, the technique is still useful even when the Markov property fails to hold due to aggregation of model states or if the independence assumption is violated by weak coupling of the channels [10]. From a practical standpoint, we only require one data segment for analysis.

We provided numerical examples demonstrating the performance of the technique. These tests showed the accuracy of estimation given a large enough data set, and also the general dependence of errors on the data length. The estimate of the third central moment can be quite noisy, especially for processes with small values of skew (as assumed in some of the examples), and it is advisable to select data points that lie within some confidence interval around the mean in order to lessen the variance of this estimate. The length of data required for satisfactory estimation of the third moment is typically much less for processes with larger skew.

Although we assumed that the noise variance was known a priori for simplicity in the simulation, it is feasible to obtain an estimate from the power spectrum, using Eq. (54). The interest in biological applications is in obtaining a second, independent, estimate of the noise variance as this is thought to increase with the onset of channel activity due to the shot-noise effect [28, 30]. If the contribution of shot-noise to the power spectrum can be considered constant

over the bandwidth of interest, then our approach can be applied to yield an estimate of its variance. All that is required is a comparison of the variance of the control data segment, containing no channel activity, and the noise variance obtained from the power spectrum of the multi-channel record.

Two possible refinements of the technique concern the effects of filtering on the measurements and the violation of the independence assumption of the individual chains. Regarding filtered observations, it is a relatively easy matter to correct the statistics and power spectrum assuming the transfer function of the filter is known. We have derived compensation formulae for the case of multi-channel data that have been passed through a first-order auto-regressive filter [10]. For typical values of membrane resistance and capacitance encountered in physiological preparations, the correction factors relating the actual and filtered signal statistics are close to unity. When the individual pores are identical but no longer independent, the net channel current process is still a Markov chain, but Eq. (37), for instance, is no longer valid. A simple description of such a coupled Markov process has been formulated in [18] and essentially involves only one further parameter to describe the amount of coupling. The extension to the case of non-identical chains seems difficult, although some results for the continuous-time case may be found in Fredkin and Rice [16].

## Acknowledgements

The authors gratefully acknowledge Mr Andrew Slater for his assistance in the programming and numerical simulations.

## Appendix A

*Formula for the  $q$ th central moment of a binary Markov chain*

In the notation of Section 2.3 we have

$$\begin{aligned}
& E\{(x - m_1(x))^q\} \\
&= E\left\{\sum_{j=0}^{q-1} \binom{q}{j} x_k^{q-j} (-m_1(x))^j\right\} + (-m_1(x))^q \\
&= \sum_{j=0}^{q-1} \binom{q}{j} (-m_1(x))^j E\{x_k^{q-j}\} + (-m_1(x))^q \\
&= \pi_2 \sum_{j=0}^{q-1} \binom{q}{j} (-m_1(x))^j s_2^{q-j} + (-m_1(x))^q
\end{aligned}$$

and  $E\{x_k^q\} = \pi_2 s_2^q$ . Continuing, we have

$$\begin{aligned}
& E\{(x - m_1(x))^q\} \\
&= \pi_2 (s_2 - m_1(x))^q - \pi_2 (-m_1(x))^q \\
&\quad + (-m_1(x))^q \\
&= \pi_2 (s_2 - m_1(x))^q + \pi_1 (-m_1(x))^q \\
&= \pi_1 \pi_2 s_2^q [\pi_1^{q-1} + (-1)^q \pi_2^{q-1}]
\end{aligned}$$

since  $m_1(x) = \pi_2 s_2$  when  $s_1 = 0$ .

## References

- [1] C.R. Anderson and C.F. Stevens, "Voltage clamp analysis of acetylcholine produced end-plate current fluctuations at frog neuromuscular junction", *J. Physiol. London*, Vol. 235, 1973, pp. 655–691.
- [2] F. Ball and M.S.P. Sansom, "Ion-channel gating mechanisms: model identification and parameter estimation from single channel recordings", *Proc. Roy. Soc. London B*, Vol. 236, 1989, pp. 385–416.
- [3] W.R. Bennett, "Statistics of regenerative digital transmission", *Bell Systems Technical J.*, Vol. 37, 1958, pp. 1501–1542.
- [4] A.T. Bharucha-Reid, *Elements of the Theory of Markov Processes and Their Applications*, McGraw-Hill, New York, 1960.
- [5] B.S. Bosik, "The special density of a coded digital signal", *Bell Systems Technical J.*, Vol. 51, 1972, pp. 921–932.
- [6] G.L. Cariolaro and G.P. Tronca, "Spectra of block coded digital signals", *IEEE Trans. Comm.*, Vol. COM-22, 1974, pp. 1555–1564.
- [7] D.G. Childers, ed., *Modern Spectrum Analysis*, IEEE Press, New York, 1978.
- [8] 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. London B*, Vol. 329, 1990, pp. 265–285.
- [9] 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. London B*, Vol. 334, 1991, pp. 357–384.
- [10] S.H. Chung and G.W. Pulford, "Fluctuation analysis of patch-clamp or whole-cell recordings containing many single channels", *J. Neurosci. Meth.*, Vol. 50, 1993, pp. 369–384.
- [11] D. Colquhoun and A.G. Hawkes, "Relaxation of membrane currents that flow through drug-operated channels", *Proc. Roy. Soc. London B*, Vol. 199, 1977, pp. 231–262.
- [12] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 2, Wiley, New York, 1966, p. 470.
- [13] J.S. Frame, "Matrix functions and applications – Part IV", *IEEE Spectrum*, Vol. 1, 1964, p. 123–131.
- [14] D.R. Fredkin and J.A. Rice, "On aggregated Markov processes", *J. Appl. Probab.*, Vol. 23, 1986, pp. 208–214.
- [15] D.R. Fredkin and J.A. Rice, "Correlation functions of a finite-state Markov process with applications to channel kinetics", *Math. Biosciences*, Vol. 87, 1987, pp. 161–172.
- [16] D.R. Fredkin and J.A. Rice, "On the superposition of currents from ion channels", *Phil. Trans. Roy. Soc. London B*, Vol. 334, 1991, pp. 347–356.
- [17] F.R. Gantmacher, *The Theory of Matrices*, Vol. 2, Chelsea, New York, 1974.
- [18] R.A. Kennedy and S.H. Chung, "Identification of coupled Markov chain models with application", *Proc. 31st Conf. on Decision and Control*, Tucson, Arizona, 16–18 December 1992, pp. 3529–3534.
- [19] J. Keilson, *Markov Chain Models – Rarity and Exponentiality*, Springer, New York, 1979.
- [20] J.G. Kemeny and J.L. Snell, *Finite-State Markov Chains*, Van Nostrand, Princeton, NJ, 1960.
- [21] E.A. Lee and D.G. Messerschmitt, *Digital Communication*, Kluwer, Boston, MA, 1988.
- [22] L.S. Liebovitch, J. Fishbarg and J.P. Koniarek, "Optical correlation functions applied to the random telegraph signal: how to analyze patch clamp data without measuring the open and closed times", *Math. Biosciences*, Vol. 78, 1986, pp. 203–215.
- [23] S.L. Marple Jr., *Digital Spectral Analysis with Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1987.
- [24] J.M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications", *Proc. IEEE*, Vol. 79, 1991, pp. 278–305.
- [25] A.V. Oppenheim and R.W. Schaffer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.
- [26] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, NY, 1991.
- [27] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, *Numerical Recipes – The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1989.
- [28] S.O. Rice, "Mathematical analysis of random noise", *Bell Systems Technical J.*, Vol. 23, 1944, pp. 282–332; *ibid*, Vol. 24, 1945, pp. 46–156.

- [29] M.S.P. Sansom, F.G. Ball, C.J. Kerry, R. McGee, R.L. Ramsey and P.N.R. Usherwood, “Markov, fractal, diffusion, and related models of ion channel gating – a comparison with experimental data from two ion channels”, *Biophys. J.*, Vol. 56, 1989, pp. 1229–1243.
- [30] E. Siebenga, A.W.A. Meyer and A.A. Verveen, “Membrane shot-noise in electrically depolarized nodes of Ranvier”, *Pflügers Archiv.*, Vol. 341, 1973, pp. 87–96.
- [31] F.J. Sigworth, “The variance of sodium current fluctuations at the node of Ranvier”, *J. Physiol. London*, Vol. 307, pp. 97–129.
- [32] C.F. Stevens, “Inferences about membrane properties from electrical noise measurements”, *Biophys. J.*, Vol. 307, pp. 97–129.
- [33] D.M. Titterton, A.F.M. Smith and V.E. Makov, *Statistical Analysis of Finite Mixture Distributions*, Wiley, New York, 1985.