Statistical Analysis of Ion Channel Data Using Hidden Markov Models With Correlated State-Dependent Noise and Filtering

M. C. M. DE GUNST, H. R. KÜNSCH, and J. G. SCHOUTEN

A hidden Markov model that describes ion channel data, including correlated, state-dependent noise and filter characteristics, and a Markov chain Monte Carlo algorithm that enables Bayesian inference under this model are presented. The method provides parameter estimates and an estimate of the noiseless signal. It was tested on simulated data and applied to real recordings to estimate the model parameters. Modeling the noise and filter correctly turned out to be crucial for the analyzed data sets. The assumption of white noise is too simple, and negligence of the smoothing effect of the low-pass filter leads to errors in the detection of rapid transitions. The hidden Markov model that we propose treats these effects simultaneously.

KEY WORDS: Bayesian inference; Ion channel recording; Kalman filter; Markov chain Monte Carlo; Signal reconstruction.

1. INTRODUCTION

Ion channels are proteins in the membranes of both plant and animal cells. These proteins conduct selected ions in and out of the cells. An ion channel is not permanently open for conduction of ions. Its kinetic behavior, the changes in being closed or open, has a random nature. Certain gating mechanisms, related to the configuration of the protein, govern the kinetic behavior of the ion channel. Ion channels play an important role in, for example, the transduction of signals in neural and muscular systems. Models that describe ion channel recordings are developed with the goal of better understanding the physiological states of the protein, its kinetic behavior, and the factors that influence its activation or inactivation.

The conduction of ions over the membrane corresponds to small electrical currents, which can be measured by using the patch clamp technique (Sakmann and Neher 1995). Recordings of a single ion channel then consist of a closed and an open conductance level. Because the currents are very small, the signal is usually masked by noise from various sources. Before the recordings are sampled and stored, they are low-pass filtered to suppress aliasing of high frequencies.

Our goal is to analyze ion channel recordings that were made at the biology department of the Free University by Sake Vogelzang; see Section 5. More precisely, our objectives are the selection of a model that gives a satisfactory description of the ion channel gating mechanism and the estimation of the unknown parameters of this model. In this article, we consider the problem of parameter estimation; the topic of model selection is treated by De Gunst and Schouten (2000).

The kinetic behavior of single ion channels often is modeled by a continuous-time, time-homogeneous, irreducible Markov chain, and this is our approach. The state space S is finite, and the states correspond to the physiological states of the gating mechanism of a single ion channel. The state space is the union of two disjoint sets, $S = S_c \bigcup S_o$, in which the ion channel is observed to be closed or open, respectively. For an overview of the history of modeling ion channel currents, see Ball and Rice (1992).

The best one can do is to observe whether the ion channel is open. If one is able to reconstruct the noiseless signal, inferences can be made directly from the reconstruction; see, for example, Colquhoun and Sigworth (1995) or Qin, Auerbach, and Sachs (1997). In most recordings, the variance of the noise is large relative to the underlying signal, which makes a noiseless reconstruction very difficult or impossible. Hidden Markov models incorporating the noise process are used to overcome this problem. Chung et al. (1990) and Chung, Krishnamurthy, and Moore (1991) introduced hidden Markov models containing white background noise. They show that the hidden Markov models are reasonably robust with respect to the first order Markov model assumption. However, under such models, correlated noise and filtering may lead to errors in event detection.

From preliminary analyses of our data, it followed that we had to incorporate both autocorrelations and state-dependent variances in our noise model. Also, the blurring effect of the low-pass filter was visible in our data. Ignoring it is likely to cause problems when rapid transitions between closed and open states occur frequently, as is the case in most of our experimental recordings. For these reasons, we found it necessary to incorporate both the correlated and state-dependent noise and the filter effects in our hidden Markov model. This is explained in detail in Section 2. Additional features, like baseline drift and other deterministic interferences, do occur, but we found that these could be removed from the signal in a preprocessing step.

Most methods in the literature deal with some but not all of these complications. Fredkin and Rice (1992a, 1997) used hidden Markov models with filtering and correlated but state-independent noise. Venkataramanan et al. (1998) developed a modified Baum–Welch algorithm for hidden Markov models with correlated, state-independent noise but no filtering. This was extended by Venkataramanan, Kuc, and Sigworth (1998) to hidden Markov models with correlated, state-independent

M. C. M. de Gunst is Professor of Statistics, Free University, Amsterdam, The Netherlands (E-mail: *degunst@cs.vu.nl*). H. R. Künsch is Professor of Statistics, ETH-Zentrum, Zürich, Switzerland (E-mail: *kuensch@stat.math.ethz.ch*). J. G. Schouten is a trainee at Statistics Netherlands, Voorburg, The Netherlands (E-mail: *bstn@cbs.nl*). The authors thank Sake Vogelzang for providing the close-to-ideal ion channel recordings and for his helpful advice and remarks. The authors thank the associate editor and both referees for their useful comments and many helpful suggestions.

and white, state-dependent noise. Likelihood approximations were developed by Michalek, Wagner, and Timmer (1998) in a model that contains all three features. Their formulas are, however, quite involved. Autoregressive noise can be handled with the idea of metastates; see Venkataramanan et al. (1998). Complications occur if moving average parts are present, and the addition of a state-dependent white noise to an autoregressive noise introduces such parts.

We adopt here the Bayesian point of view because there is usually prior information about the noise parameters and transition rates and because this allows us to take into account parameter uncertainty in the reconstruction of the signal. The analysis relies heavily on Markov chain Monte Carlo (MCMC) methods. A key element is that we are able to perform updates of the unobserved Markov chain in a single step. This is explained in Sections 3 and 4. The results are shown in Section 5.

The Bayesian approach also has been used by Ball et al. (1997) and Hodgson (1999). The difference is that we do not generate the jump times of the Markov chain, but we generate its states at the sample time points. Also, Ball et al. (1999) do not consider filter effects or correlated noise, and Hodgson (1999) deals only with state-independent noise.

2. THE HIDDEN MARKOV MODEL

As stated in the Introduction, it is assumed that the observed signal is a hidden Markov model, where the Markov chain is related to the gating mechanism. The aggregation of states that correspond to a closed and open ion channel, correlated and state-dependent background noise, and low-pass filtering make the Markov model hidden.

2.1 Markov Models for Ion Channel Kinetics

The kinetics of the channel are described by a continuous-time, time-homogeneous irreducible Markov chain $\{X(t) : t \geq 0\}$. The state space S is finite and consists of the union of S_c, the set of closed states, and S_o, the set of open states. One cannot observe the exact state of the Markov chain, but in the ideal noiseless case, one would observe the process $\{\mu(X(t)) : t \geq 0\}$, where

$$\mu(X(t)) = \begin{cases} \mu_c & \text{if } X(t) \in S_c, \\ \mu_o & \text{if } X(t) \in S_o. \end{cases} \tag{1}$$

Candidate models specifying the number of open and closed states and the allowed transitions between them arise naturally from the physiological background of the data, that is, from the conjectured configuration of the ion channel protein. In our examples in Section 5, we consider two different gating mechanisms, but other models have been discussed and can be treated in the same way; see, e.g., Schouten (2000).

The probability distribution of $\{X(t) : t \geq 0\}$ is determined by an infinitesimal matrix Q, containing the transition rates between all pairs of states, and the initial distribution at $t = 0$. We assume that the latter is the stationary distribution π_x, satisfying $\pi_x Q = 0$ and $\sum_{i \in S} \pi_x(i) = 1$. This means that the experiment is in equilibrium.

The continuous-time Markov chain is sampled with sample step Δ_t. The sampled chain $X_t = X(t\Delta_t)$, $t = 0, 1, \ldots, T$, is then again Markov, and the one-step transition matrix $P = P_{\Delta_t}$ is given by the matrix exponential $P = \exp(\Delta_t Q)$.

2.2 Modeling the Noise

In the recordings, some background noise is always present, whether the ion channel is active or not, because the measured currents are very small (in the order of pA). This state-independent background noise stems from different sources, e.g., Johnson or thermal noise from conductors, shot noise from potential barriers, dielectric noise from capacitors, and so forth. Noise analyses (Schouten 2000) of the baseline current, i.e., when all ion channels are closed, indicate that this noise is not white, but AR(p) processes may suffice to describe the correlations. We thus will assume that the correlated state-independent noise $\{C_t : 0 \leq t \leq T\}$ is a stationary causal AR(p) process. So for $p \leq t \leq T$,

$$C_t = \sum_{i=1}^{p} \phi_i C_{t-i} + \epsilon_t, \tag{2}$$

where the innovations ϵ_t are independent and $\mathrm{N}(0, \sigma_\epsilon^2)$ distributed.

When the ion channel opens, the amplitude of the current increases and the effects of some of the sources of background noise mentioned earlier become stronger. For the model, this means that the variance of the noise increases when the channel opens. We assume that the correlation of the noise is not state dependent and that the increase in noise variance can be described by an additional noise process $\{W_t : 0 \leq t \leq T\}$, where

$$W_t = \sigma(X_t)\delta_t = \begin{cases} \sigma_c \delta_t & \text{if } X_t \in S_c, \\ \sigma_o \delta_t & \text{if } X_t \in S_o \end{cases} \tag{3}$$

and $\{\delta_t : 0 \leq t \leq T\}$ is white standard Gaussian noise.

2.3 Filter Effects

Analog low-pass Bessel or Butterworth filters are used for antialiasing of high frequencies when an analog signal is sampled. A drawback of using a filter is that some short-time memory is introduced, and the filter tends to smooth rapid transitions between the closed and open current levels. As a result, the current does not always reach the closed and open current levels before returning to the original current level; see Figures 3(a) and 4(a). Because of this, fast transitions are easily missed, and estimates for some transition rates are poor.

Therefore, a discrete approximation to the analog filter is incorporated in the model by introducing weights $\gamma_{-r}, \gamma_{-r+1}, \ldots, \gamma_r$ to describe the smoothing effect of the filter. It is assumed that the noiseless current $\{I_t : 0 \leq t \leq T\}$ is a weighted sum, i.e.,

$$I_t = \sum_{k=-r}^{r} \gamma_k \mu(X_{t-k}) \tag{4}$$

with μ as defined by (1).

The weights should be chosen according to the filter characteristics. Analog filters may of course depend only on past values, so theoretically the weight coefficients $\gamma_{-1}, \ldots, \gamma_{-r}$ should all be equal to 0. However, for approximation purposes it is convenient to introduce also dependence on future values.

2.4 The Complete Model

With the assumptions made in the previous sections, we give the full hidden Markov model. Let $\{Y_t : 0 \le t \le T\}$ be the recording of the current that is observed. Then for $(r \vee p) \le t \le T - r$,

$$Y_t = I_t + C_t + W_t = \sum_{k=-r}^{r} \gamma_k \mu(X_{t-k}) + C_t + \sigma(X_t)\delta_t. \qquad (5)$$

In addition to the means μ_c and μ_o, the noise variances σ_c^2, σ_o^2 and σ_ϵ^2 and the autoregressive (AR) parameters $\phi = (\phi_1, \phi_2, \ldots, \phi_p)'$, there are unknown parameters in the generator matrix Q. But some entries in Q are zero, corresponding to infinitesimal transitions that are not allowed within the model, and between the positive entries the model often postulates linear dependencies; see Section 5. Thus, Q depends on m free parameters $q = (q_1, q_2, \ldots, q_m)'$, where m is usually small. The vector θ of unknown parameters in the model (5) is therefore

$$\theta = \left(\mu_c, \mu_o, \sigma_c^2, \sigma_o^2, \sigma_\epsilon^2, \phi', q'\right)'. \qquad (6)$$

Venkataramanan and Sigworth (1998) used the same model for the noise but omitted the effect of the filter. The order p of the AR process and the weights $\gamma = (\gamma_{-r}, \gamma_{-r+1}, \ldots, \gamma_r)'$ are assumed to be known and are fixed in advance. This is reasonable in our case, because we can obtain p independently from noise analysis of stretches of recordings when the channel was closed, and the weights follow from the known characteristics of the applied filter.

We remark that the pair $(\{Y_t : 0 \le t \le T\}, \{(X_{t-r}, \ldots, X_{t+r}, C_{t-p+1}, \ldots, C_t) : 0 \le t \le T\})$ is a discrete-time hidden Markov model governed by a continuous-time Markov chain. Because we are interested in the gating mechanism of the ion channel, we parametrize $\{X_t : 0 \le t \le T\}$ by the unknown parameters q in Q, which appear in $P = \exp(\Delta_t Q)$. Hence, the model takes transitions between sample time points into account. This differs from the model used by Venkataramanan et al. (1998a,b), who did not parametrize the one-step transition matrix P by the transition rates but used the entries of P themselves as parameters in their model.

Variations of our model are possible. For example, we could assume that the innovations variance σ_ϵ^2 or the autoregression coefficients $\phi_1, \ldots, \phi_p$ of the correlated noise process is also state dependent. The computations then would be similar. However, our noise analyses gave no indication that the model (5) is not adequate.

Finally, we introduce some notation that we use in the following. When needed, we distinguish a random variable and a realization of it by writing a capital and a lowercase letter, respectively. For any sequence $z_1, z_2, \ldots, z_T$, we write z_s^{s+t} for the stretch $(z_s, z_{s+1}, \ldots, z_{s+t})$ $(1 \le s \le s+t \le T)$. We use $\rho = r \vee p$ for the starting point of the observations (y_t) and write $y = y_\rho^{T-r}$, $x = x_{\rho-r}^{T}$, and $c = c_{\rho-p}^{T-r}$.

3. BAYESIAN INFERENCE FOR THE HIDDEN MARKOV MODEL

Our goal is to estimate the state sequence x and the parameter vector θ based on observations y in model (5) and to make inferences about x and θ. We adopt a Bayesian viewpoint and treat θ as a random variable. This gives us a unified computational framework; we can take uncertainty about θ into account in the inference on x and we can incorporate prior information about θ. Such prior information is often available. Different recordings obtained under approximately the same experimental conditions should have similar background noise characteristics. For the transition rates, physiological bounds can be given. Bayesian inference is based on the posterior densities $p_{\Theta|Y}$ and $p_{X|Y}$. In the following, the indices for the probability densities will be omitted, and we will use p everywhere.

3.1 An MCMC Algorithm With Multiple Updates

Expressions for the posterior densities of our interest, $p(x|y)$ and $p(\theta|y)$, cannot be derived analytically. For this reason, we resort to MCMC methods. For a thorough treatment of MCMC methods see, for example, Gilks, Richardson, and Spiegelhalter (1996) or Robert and Casella (1999). Information about $p(x|y)$ and $p(\theta|y)$ can then be obtained from the sequences $\{x^i\}_{1<i\le N}$ and $\{\theta^i\}_{1<i\le N}$, respectively, that are generated by the MCMC algorithm. However, the straightforward Gibbs sampler is not useful here. First, because the noise is a superposition of C_t and W_t, the conditional distribution $p(y|x)$ and thus also the full conditionals $(x_t|\{x_s\}_{s\ne t}, \theta, y)$ are complicated and cannot be expressed easily. This can be overcome by adding the noise c to the Markov chain; that is, we sample the variables (c, x, θ) given y. The second problem is that with single updates of x_t or c_t given all other variables, the Markov chain typically is very slow; see, for example, Carter and Kohn (1994). Therefore, as Carter and Kohn (1994) suggested, we want to update $(x|\theta, y)$ all at the same time. As we show, these updates are computationally tractable. Thus, our Gibbs sampler has the following form:

1. Construct starting values x^0 and θ^0 for x and θ. Let $i = 0$.
2. Sample the background noise c^{i+1} from $p(c|y, x^i, \theta^i)$.
3. Sample the ion channel states x^{i+1} from $p(x|y, c^{i+1}, \theta^i)$.
4. Sample θ^{i+1} from $p(\theta|y, c^{i+1}, x^{i+1})$. Return to step 2 with $i := i + 1$.

Because it is easier to construct starting values for x and θ than for c, we update c first. Steps 3 and 4 of the algorithm can be interchanged. The separate steps are described in detail in Section 4.

The results presented in Section 5 use quadratic loss, i.e., we take the expectation of the posterior distribution, to obtain point estimates for θ. Our estimates for $1_{S_o}(x)$ are based on the total number of misclassified states as loss function, which we refer to as indicator loss. As point estimator of $1_{S_o}(x_t)$, we use 1 if $P[X_t \in S_o|y] > 0.5$ and 0 otherwise. This is optimal for the indicator loss function that counts the number of misclassified states. It is more useful than the estimates based on quadratic loss, which are between 0 and 1. See, for example, Berger (1980) for a general reference on Bayesian concepts. Analytic expressions for the expectations of $(\theta_i|y)$ are not available, but they can be estimated by arithmetic means from the Gibbs sampler.

3.2 Convergence and Geometric Ergodicity

Here we merely state that it can be shown that the transition probability kernel of the Gibbs sampler defined by steps 1–4 converges to the invariant distribution $\pi = p(x, c, \theta|y)$ in total variation norm except for a π-null set, and that this Gibbs sampler generates a π-irreducible, aperiodic, and recurrent Markov chain.

Because we have a special interest in the transition rates $q_1, \ldots, q_m$ of x, we also computed confidence intervals for $\mathbb{E}(q_i|y)$ for $1 \le i \le m$. Indeed, by using the general theory in Meyn and Tweedie (1993) and arguments like those in Robert, Celeux, and Diebolt (1993), it follows that both $\{x^i|y\}_{1\le i\le N}$ and $\{\theta^i|y\}_{1\le i\le N}$ are geometrically ergodic Markov chains. This means that we can use the functional central limit theorem for Markov chains to construct such confidence intervals. We have that an approximate $100(1-\alpha)\%$ confidence interval for $\mathbb{E}(q_i|y)$ is

$$\frac{1}{N-n}\sum_{j=n+1}^{N} q_i^j \pm \xi_{1-\alpha)2}\frac{\sigma_{q_i}}{\sqrt{N-n}}, \tag{7}$$

where N is the number of updates, n is the burn-in, ξ_α is the α-quantile of the standard normal distribution, and $\sigma_{q_i}^2 = \mathrm{Var}(q_i^n|y) + 2\sum_{j=1}^{\infty}\mathrm{Cov}(q_i^n, q_i^{n+j}|y)$, which is estimated from the generated Markov chain.

4. THE FOUR STEPS OF THE GIBBS SAMPLER

4.1 Starting Values

The choice of starting values can be important to have good convergence to the stationary distribution, i.e., have a small burn-in n. In choosing our starting values, we make use of a rough reconstruction of the noiseless recording. Here a reconstruction means that for every time point, it is decided whether the ion channel was open or closed. Because we found the so-called Hinkley detector (Schultze and Draber 1993) to perform better than the standard threshold methods and because it is easy to implement, we employ this method to obtain the reconstruction. It makes use of a threshold that is based on the signal-to-noise ratio of a recording. Rough estimates of the current levels and noise variances are needed to compute this threshold.

In the updating step of c in the Gibbs sampler, we do not need the exact ion channel states x but only θ and $1_{S_o}(x)$. Therefore, it suffices to have starting values for θ and $1_{S_o}(x)$; for the latter we use the open-closed reconstruction. If we ignore that short closures and openings between the sample time points fail to be detected, a phenomenon called time-interval omission, then this reconstructed signal is a sequence of closed and open holding times. The probability density of such a sequence can be computed by following Ball and Sansom (1989), and maximum likelihood estimates for q can then be approximated numerically. We use these numerical approximations as starting values for q. Correction for time-interval omission may be applied to obtain better estimates (Ball et al. 1993), but since we use the estimates only as starting values, there is no need for such correction.

Starting values for the correlated noise variables are found by using long closed periods in the reconstructed signal. Estimates for the state-dependent noise variables are made directly from long closed and open periods.

4.2 Simulation of Correlated Noise

The model (5) given $X_{\rho-r}^T = x$ and θ can be written as a linear Gaussian state space model. Namely, if we define the state vector $S_t = (C_t, C_{t-1}, \ldots, C_{t-p+1})'$ and the observations $Z_t = Y_t - \sum_{k=-r}^{r}\gamma_k\mu(x_{t-k})$, then we have the dynamics

$$\begin{cases} S_{t+1} = FS_t + u_{t+1}, \\ Z_{t+1} = HS_{t+1} + v_{t+1} \end{cases} \tag{8}$$

for $\rho \le t \le T - r$. Here, the noise variables are $u_t = (\epsilon_t, 0, \ldots, 0)'$ and $v_t = \sigma(x_t)\delta_t$. The form of the matrices F and H is obvious from (2) and (5).

To simulate the correlated noise C given x and y, we therefore use the forward Kalman filtering–backward sampling procedures in linear Gaussian state space models, as introduced by Carter and Kohn (1994) and Frühwirth-Schnatter (1994). Because this is by now fairly standard and has been used frequently in the literature (Shephard 1994; West 1995; Carter and Kohn 1996), we omit the details. We just mention that we use the stationary distribution for $S_{\rho-1}$ to initialize the Kalman filter. The mean is 0, and the covariances follow from the Yule–Walker equations.

4.3 Simulation of the States of the Ion Channel

We discuss here how to generate a sample of $X_{\rho-r}^T$ from the conditional distribution given y, c, and θ. Note that, given c and θ, the joint distribution of $(X, Y) = (X_{\rho-r}^T, Y_\rho^{T-r})$ is that of a hidden Markov model with noninstantaneous dependence of the observations Y_t on X. That is, conditional on X, the Y_t's are independent with

$$\begin{aligned}(Y_t|x, c, \theta) &= \left(Y_t|x_{t-r}^{t+r}, c, \theta\right) \\ &\stackrel{d}{=} N\left(c_t + \sum_{k=-r}^{r}\gamma_k\mu(x_{t-k}), \sigma^2(x_t)\right).\end{aligned} \tag{9}$$

We show that, conditional on (y, c, θ), X is a Markov chain of order $2r$ and that we can compute the backward transition probabilities and the backward starting distribution by a forward algorithm similar to the forward procedure in the standard hidden Markov model (Rabiner and Juang 1986). We thus will again use a forward filtering–backward sampling algorithm, as in Section 4.2.

To formulate our main result, define

$$h_{\rho+r-1}\left(x_{\rho-r}^{\rho+r-1}\right) = \pi_x(x_{\rho-r}|q)\prod_{i=\rho-r+1}^{\rho+r-1} p(x_i|x_{i-1}, q), \tag{10}$$

and for $\rho + r \le t \le T$

$$\begin{aligned}h_t\left(x_{t-2r}^t\right) &= p(x_t|x_{t-1}, q)\frac{1}{\sigma(x_{t-r})} \\ &\quad\times \varphi\left(\frac{y_{t-r} - c_{t-r} - \sum_{k=-r}^{r}\gamma_k\mu(x_{t-r-k})}{\sigma(x_{t-r})}\right).\end{aligned} \tag{11}$$

Then we have the following theorem.

Theorem 1. Conditional on (y, c, θ), X is a backward Markov chain of order $2r$ with starting distribution $p(x_{T-2r+1}^T|y, c, \theta) = \alpha_T(x_{T-2r+1}^T)$ and transition probabilities

$$p\left(x_{t-2r}|x_{t-2r+1}^t, y, c, \theta\right) = \frac{h_t\left(x_{t-2r}^t\right)\alpha_{t-1}\left(x_{t-2r}^{t-1}\right)}{\sum_{x_{t-2r}\in S} h_t\left(x_{t-2r}^t\right)\alpha_{t-1}\left(x_{t-2r}^{t-1}\right)},$$

where $\alpha_t(x_{t-2r+1}^t) = p(x_{t-2r+1}^t|y_\rho^{t-r}, c, \theta)$ is computed recursively as

$$\alpha_t(x_{t-2r+1}^t) = \frac{\sum_{x_{t-2r}\in S} h_t(x_{t-2r}^t)\alpha_{t-1}(x_{t-2r}^{t-1})}{\sum_{x_{t-2r}^t\in S^{2r+1}} h_t(x_{t-2r}^t)\alpha_{t-1}(x_{t-2r}^{t-1})},$$

starting with

$$\alpha_{\rho+r-1}\left(x_{\rho-r}^{\rho+r-1}\right) = h_{\rho+r-1}\left(x_{\rho-r}^{\rho+r-1}\right).$$

Proof. The joint density of (X, Y) is equal to

$$\begin{aligned} &p(x, y|c, \theta) \\ &\quad = \pi(x_{\rho-r}|q) \prod_{i=\rho-r+1}^{T} p(x_i|x_{i-1}, q) \\ &\qquad \times \prod_{t=\rho}^{T-r} \frac{1}{\sigma(x_t)}\varphi\left(\frac{y_t - c_t - \sum_{k=-r}^{r}\gamma_k\mu(x_{t-k})}{\sigma(x_t)}\right). \end{aligned} \tag{12}$$

The conditional density $p(x|y, c, \theta)$ is proportional to this. Therefore, it is sufficient to show that, for some term $K(y, c, \theta)$ not depending on x,

$$\begin{aligned} p(x, y|c, \theta) &= \alpha_T\left(x_{T-2r+1}^T\right) \\ &\quad \times \prod_{t=\rho+r}^{T} p\left(x_{t-2r}|x_{t-2r+1}^t, y, c, \theta\right)K(y, c, \theta). \end{aligned}$$

From (10)–(12) and the definition of the α_t's, it follows that

$$\begin{aligned} p(x, y|c, \theta) &= h_{\rho+r-1}\left(x_{\rho-r}^{\rho+r-1}\right) \prod_{t=\rho+r}^{T} h_t\left(x_{t-2r}^t\right) \\ &= \alpha_T\left(x_{T-2r+1}^T\right) \prod_{t=\rho+r}^{T} \frac{h_t\left(x_{t-2r}^t\right)\alpha_{t-1}\left(x_{t-2r}^{t-1}\right)}{\alpha_t\left(x_{t-2r+1}^t\right)} \\ &= \alpha_T\left(x_{T-2r+1}^T\right) \prod_{t=\rho+r}^{T} p\left(x_{t-2r}|x_{t-2r+1}^t, y, c, \theta\right) \\ &\quad \times \prod_{t=\rho+r}^{T} \sum_{\tilde{x}_{t-2r}^t\in S^{2r+1}} h_t\left(\tilde{x}_{t-2r}^t\right)\alpha_{t-1}\left(\tilde{x}_{t-2r}^{t-1}\right). \end{aligned}$$

Because the terms in the last product are constants with respect to x, this completes the proof. □

In practice, a lot of memory must be allocated to store all the α's that are necessary for the backward step. This can be overcome partially by storing only every kth iteration and repeating the forward step for pieces of length $k-1$ while performing the backward step. The computation time will then increase. Hence, there is a trade-off between computation time and memory allocation. It should be noted, however, that we are not interested in the α's, so that we only need to allocate memory for one iteration.

4.4 Simulation of the Parameters

In the last step of our Gibbs sampler, θ is updated by taking a sample from $p(\theta|y, c^{i+1}, x^{i+1})$. For this we only need to know $p(\theta|y, c, x)$ up to a constant. We have

$$\begin{aligned} p(\theta|y, x, c) \propto\ & p\left(y|x, c, \mu_c, \mu_o, \sigma_c^2, \sigma_o^2\right) \\ & \times p(x|q)p\left(c|\phi, \sigma_\epsilon^2\right)p(\theta), \end{aligned} \tag{13}$$

and the conditional distributions on the right follow from the definition of the model. From (13), we see that if we choose a prior distribution $p(\theta)$ with independent components, then we can do the updates independently. Moreover, if we choose normal priors for μ_c, μ_o, and ϕ and inverse gamma priors for σ_c^2, σ_o^2, and σ_ϵ^2, then the posteriors for these parameters follow by conjugacy. The details can be found in Schouten (2000).

For the transition rates $q_i, 1 \le i \le m$, the situation is somewhat more complex, because we cannot construct priors for q that are conjugate, and we do not have explicit expressions for the one-step probabilities $p(j|i, q)$ in terms of q. To overcome this problem, we use gamma priors for q and a Metropolis–Hastings algorithm. We update all q_i in one go by using the $N_m(q_{ML}(x), \Sigma_{ML}(x))$ distribution as proposal distribution. Here, $q_{ML}(x)$ is the maximum likelihood estimate of q given x, which is easy to compute, and $\Sigma_{ML}(x) = \Sigma_{\text{small}} + 1)100\, q_{ML}(x)' I_m q_{ML}(x)$. The matrix Σ_{small} is diagonal, and its entries are fixed in advance and are chosen small relative to the starting values for q. The term Σ_{small} is added to prevent the sampled Markov chain from sticking to small samples for q and, therefore, to improve the mixing conditions of the Markov chain.

5. RESULTS FROM SIMULATED AND RECORDED SAMPLES

The algorithms described in the previous sections were written in C codes and Matlab codes and finally linked to Matlab. Several recordings were simulated to test the algorithms for varying models, transition rates, and noise characteristics, and several real recordings were analyzed to investigate the workings of the proposed method on real data sets. The simulated recordings were created according to (5); i.e., the states of the Markov process $\{X_t : 0 \le t \le T\}$ were sampled with transition matrix $P = \exp(\Delta_t Q)$, which allows for transitions between the sample time points. In this section, we present four typical examples, two simulated and two real recorded. Parts of the four examples are shown in Figures 1–4. Data and software that were used for the examples are available on request.

The two Markov models for the gating mechanism used in the examples are the following:

- Model 1:

$$C_1 \underset{q_2}{\overset{q_1}{\rightleftharpoons}} O \underset{q_4}{\overset{q_3}{\rightleftharpoons}} C_2$$

- Model 2:

$$C_1 \underset{q_2}{\overset{4q_1}{\rightleftharpoons}} C_2 \underset{2q_2}{\overset{3q_1}{\rightleftharpoons}} C_3 \underset{3q_2}{\overset{2q_1}{\rightleftharpoons}} C_4 \underset{4q_2}{\overset{q_1}{\rightleftharpoons}} O \underset{q_4}{\overset{q_3}{\rightleftharpoons}} C_5$$

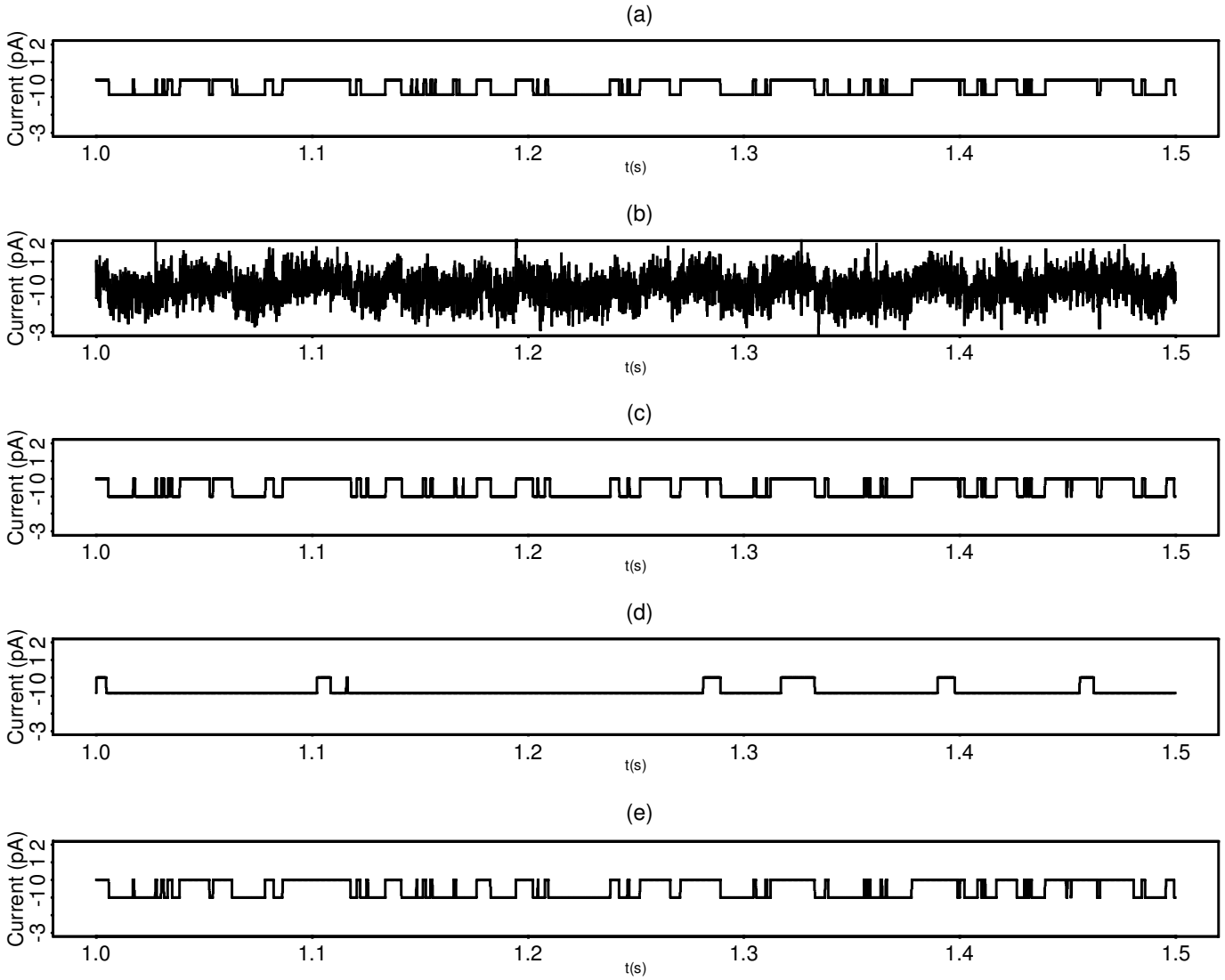

Figure 1. Part of Simulation 1 (.5 sec). (a) The simulated jumps; (b) the simulated recording; (c) the outcomes of the Gibbs sampler with filtering for indicator loss; (d) the reconstruction from the Hinkley detector; and (e) the outcomes of the Gibbs sampler without filtering for indicator loss.

Simulation 1 is from model 1 and simulation 2 from model 2. Both recordings are analyzed under model 2. Model 1 is frequently used in the literature and model 2 arises from biological indications for the presence of four identical slow gates plus one fast gate. In model 2, the closed state C_i, for $1 \le i \le 4$, corresponds to $5 - i$ open identical slow gates, and C_5 represents a closed fast gate.

The parameter values used for the simulated data are given in Table 1. The signal-to-noise ratio is low for simulation 1 and high for simulation 2. For simulation 2, the parameters were chosen to be comparable to the values of the estimates of recording 2 to mimic the biological reality. The simulated samples have a length of 10^5 and 2×10^5 points, respectively.

The recorded samples are a cell-attached-patch and an inside-out-patch measurement of a single potassium outward rectifier in a leaf protoplast of barley. The data were sampled at fixed holding potentials [$V_p = -96$(mV) and $V_p = -150$(mV)], and an eight-pole Bessel filter was applied with a cutoff frequency of $f_c = 5$ kHz and a roll-off of 48 dB/octave. They were corrected for trend by fitting and subtracting polynomials to the baseline noise. From the recordings, 2×10^5 points were taken for the analysis. Recording 1 is an example of a high noise recording, whereas in recording 2, signal and noise are easily distinguished by eye because of a low noise level. Parts of recordings 1 and 2 are given in Figures 3(a) and 4(a), respectively. See Vogelzang and Prins (1995) for experimental details and comparable recordings.

The simulations and the recording were sampled at $f_s = 10$ kHz. The smoothing coefficients $\gamma_{-r}, \ldots, \gamma_r$ were chosen according to the digital Gauss filter with $r = 1$ (Colquhoun and Sigworth 1995), which approximates very well the characteristics of the Bessel filter used in the experiments. It should be noted that in case $f_s) f_c < 4.75$, it suffices to let $r = 1$. For larger values of $f_s) f_c$, r increases linearly with this ratio.

Starting values for the Gibbs sampler were taken as outlined in Section 4.1. The computation of starting values for q took a couple of hours on a Sun workstation (128 MB RAM, Ultrasparc8 processor), depending on the complexity

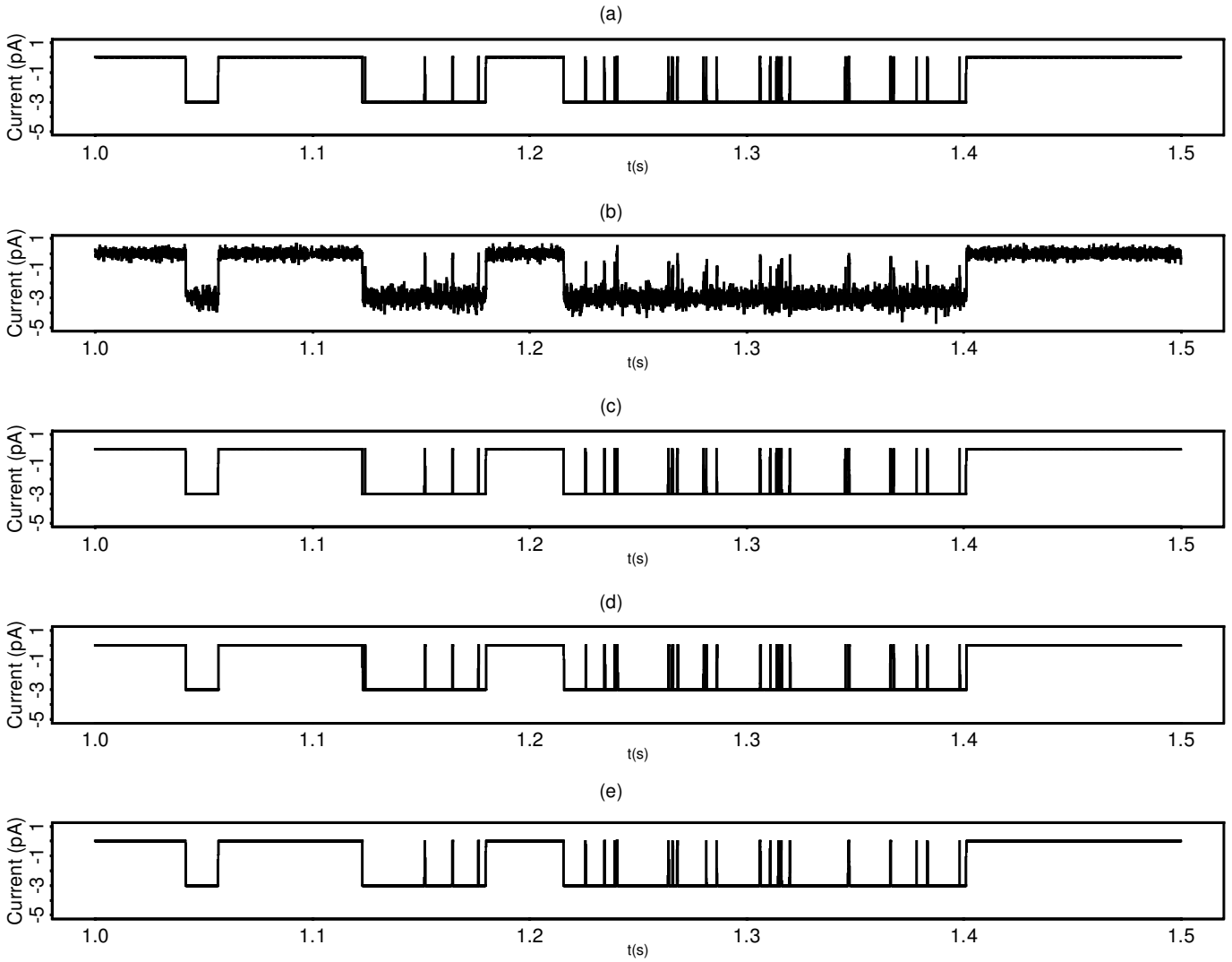

Figure 2. Part of Simulation 2 (.5 sec). (a) The simulated jumps; (b) the simulated recording; (c) the outcomes of the Gibbs sampler for the model with filtering for indicator loss; (d) the reconstruction from the Hinkley detector; and (e) the outcomes of the Gibbs sampler for the model without filtering.

of the model and the number of transition rates to be estimated. The Gibbs sampler performed 50,000 and 5,000 iterations for the simulations; for the recordings the number of iterations was 4,000 and 7,000. Although the approximation technique proposed by Fredkin and Rice (1997) was implemented, one run of the Gibbs sampler took several days.

In general, a more complex gating mechanism, as with model 2, does not lead to more iterations that are necessary to have good approximations. The burn-in period n was chosen by eye, between 1,000 and 2,500 in these examples. Quantile plots, the gibbsit software (Raftery and Lewis 1992), and modified CUSUM path plots (Brooks 1998) were used to diagnose convergence of the sampler. Except for the variance parameter σ_c^2, these diagnostics gave satisfactory results.

As a typical example, the resulting MCMC chains for the parameters of simulation 1 are shown in Figure 5. It can be seen from these that the method is not very sensitive to the starting values and that the Gibbs sampler gives a satisfying picture for all parameters, except for σ_c^2. Indeed, when σ_c^2 is taken very small compared to σ_o^2 and σ_ϵ^2, the sampler mixes slowly with respect to this parameter. After 2,500 iterations, the generated MCMC chains for all parameters except that for σ_c^2 fluctuated around the values that were used in the simulation, and the averages of these chains were almost constant. The starting values, especially those for q and $1_{S_o}(x)$, however, deviated considerably from the simulation values. We found the same for other simulations. Keeping σ_c^2 fixed at different small values did not change the results of the computations for the other parameters. The fact that σ_c^2 cannot be recovered too well is of no practical importance precisely because it is small. Because we are not interested in σ_c^2 and because the convergence of the other parameters does not seem to be affected by that of σ_c^2, we restricted the number of iterations for the other simulation and both recordings.

In Figures 1–4, the results of our Gibbs sampler for the depicted parts of the examples are shown. The estimates of the jump process based on indicator loss and the reconstructions via the Hinkley detector are shown. In the case of the simulated examples, the estimates of the jump process when filtering is neglected are also shown. From Figures 1 and 2 we see

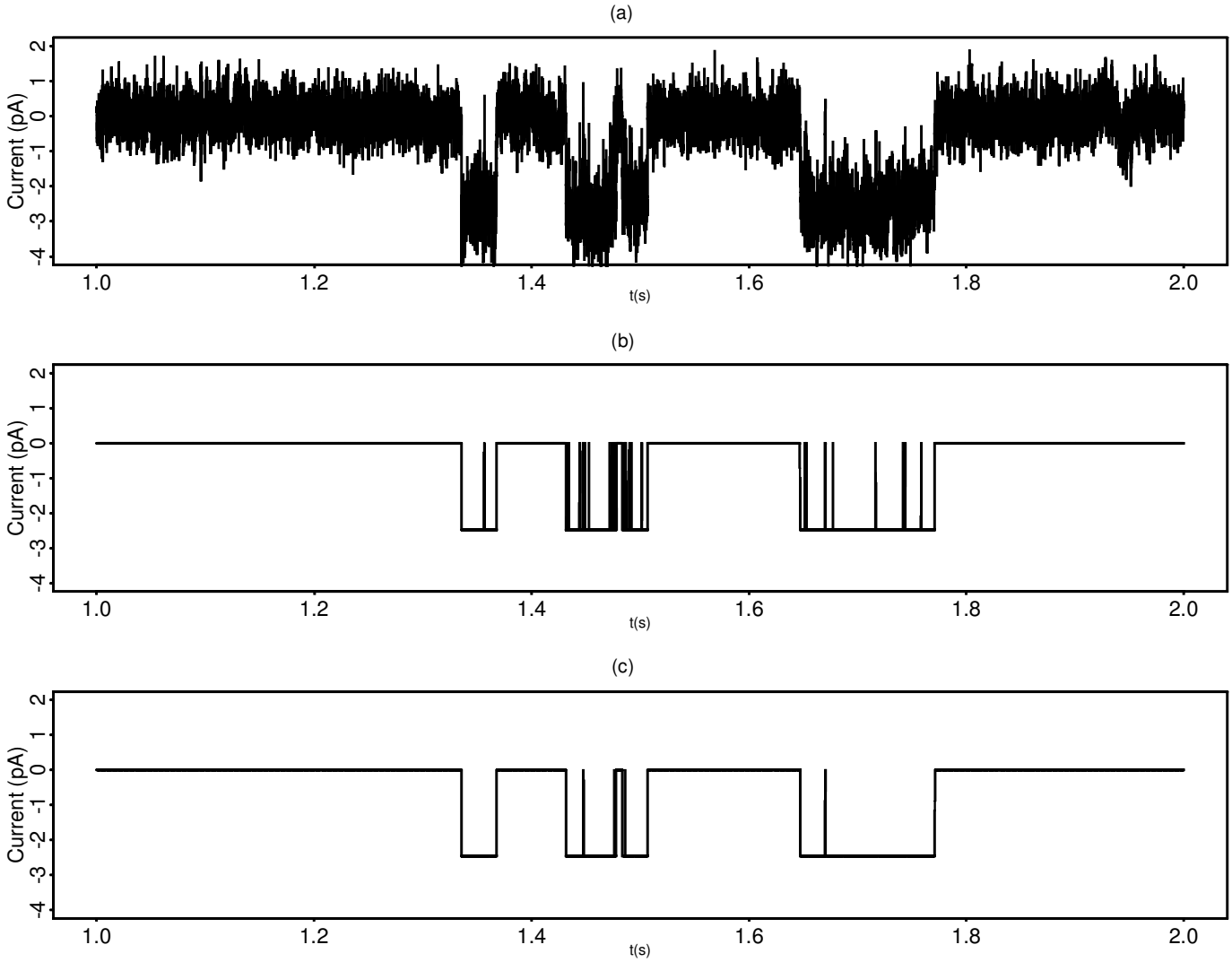

Figure 3. (a) Part of Recording 1 (1.0 sec); (b) the Outcomes of the Gibbs Sampler for Indicator Loss; and (c) the Reconstruction From the Hinkley Detector.

that the underlying jump process is recovered quite well by our Gibbs sampler. Most rapid transitions between the closed and open current levels were recovered in the simulated and, seemingly, also in recorded samples. Table 2 gives the percentage of misclassified states for the Gibbs sampler, without and with filtering, and the Hinkley detector. We see that as the signal-to-noise ratio decreases, the Hinkley detector is likely to miss more transitions. This affects the starting values. For the part of recording 2 shown in Figure 2, the Hinkley detector gave the same jump process as did our estimates, but for the whole sample approximately 50 of 475 jump points were assigned differently. As mentioned, short closures and openings are smoothed by the filter and are not always recognized as jumps when the smoothing effect of the filter is not included in the model; see the last pictures in Figures 1 and 2. In case filtering is neglected, the percentage of misclassified states is larger, as can be seen from Table 2.

In Table 1, the point estimates for the θ_i computed by taking the sample means of the respective generated marginal MCMC runs are given for all examples. We see that in general for the simulated samples they approximate the true values reasonably well. In particular, the current levels were recovered rather well. The point estimates for the transition rates in the recorded samples give biologically plausible values. The 95% highest posterior density credible regions are given in Table 3 and contain the true values in most cases. Clearly, incorporating the filter makes a big difference. Without the filter, the estimates of the variance parameters are larger, and the difference between the current levels $|\mu_o - \mu_c|$ is slightly smaller, because some jumps are missed. But the biggest difference is in the estimated transition rates q. Without the filter, these estimates can be off by an order of magnitude. Only the estimate for q_4 in simulation 2 is worse with filtering than without. But this can be explained by the fact that in this simulation the maximum likelihood estimate for q_4 based on the simulated jumps, i.e., the noiseless process, is 8,100.

Finally, because the transition rates turned out to be the most difficult to recover, we checked whether we ran the Gibbs sampler long enough by computing for the transition rates the 95% confidence intervals for the transition rates, as given in formula (7). Here, we estimated $\sigma^2_{\theta_i}$ by using a truncated window estimator based on the autocovariance function.

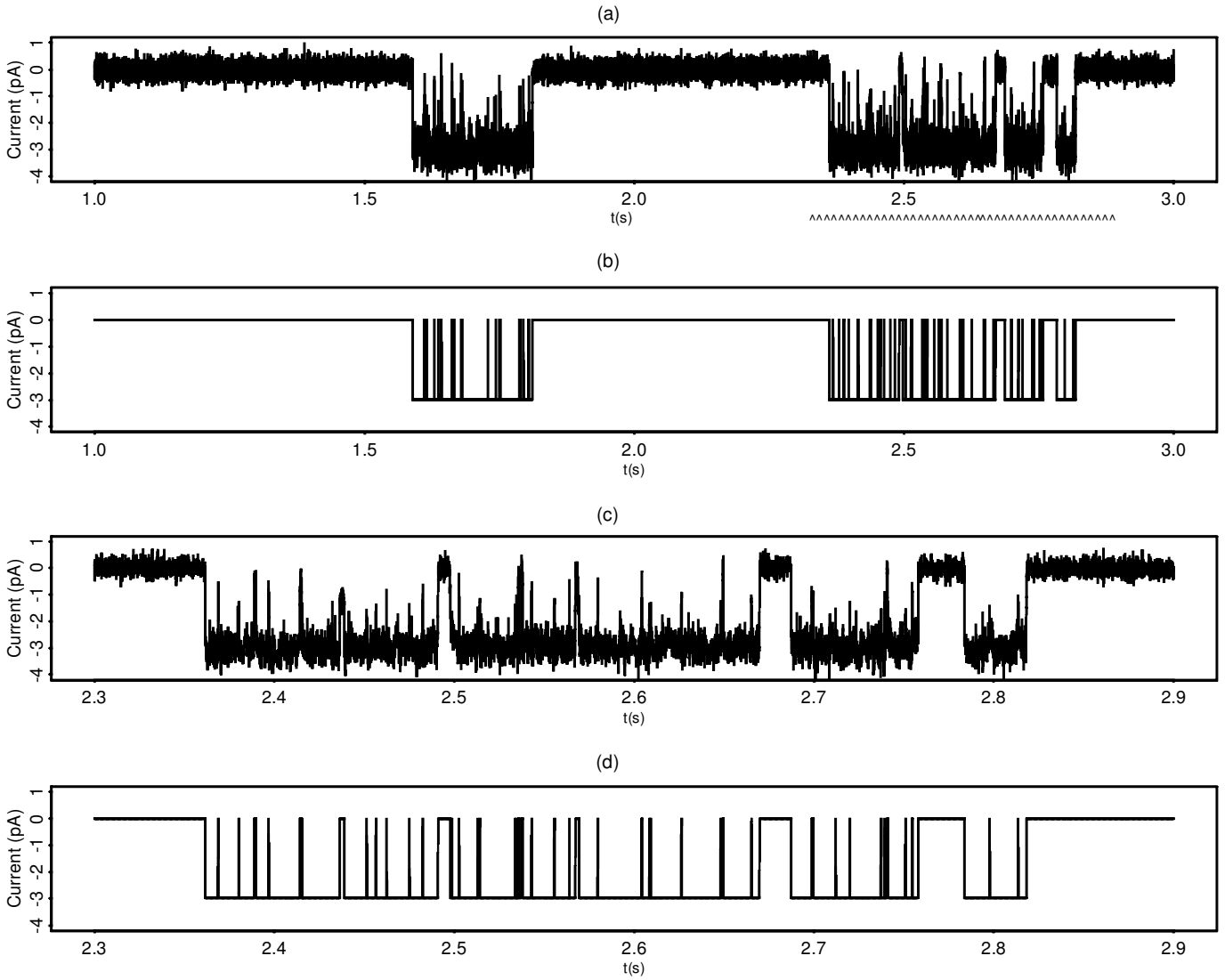

Figure 4. (a) Part of Recording 2 (2.0 sec); (b) the Outcomes of the Gibbs Sampler for Quadratic and Indicator Loss (the Same); (c) Part of Recording 2 Marked in (a) (0.6 sec); and (d) the Outcomes of the Gibbs Sampler for Quadratic and Indicator Loss (the same).

Table 1. Values of θ Used in Simulation, and Estimates Without and With Filtering

	Simulation 1			Simulation 2			Recording 1	Recording 2
Variable	True	Estimate no filter	Estimate filter	True	Estimate no filter	Estimate filter	Estimate	Estimate
q_1	100	989	104	3.6	3.59	3.64	8.91	3.57
q_2	100	108	103	2.8	3.54	3.67	8.07	2.79
q_3	100	965	109	130	114	145	811	132
q_4	1000	989	1021	7000	6794	7840	17914	7041
ϕ_1	−.3	−.31	−.31	−.28	−.27	−.27	3×10^{-4}	−.28
ϕ_2	−.1	−.10	−.10	−.10	−.10	−.10	-3×10^{-6}	−.06
ϕ_3	−.05	−.05	−.05	−.05	−.05	−.05	-7×10^{-5}	.002
$\sigma_\epsilon^2(\times10^{-1})$	4	3.87	3.97	.48	.49	.49	1×10^{-6}	.48
$\mu_c(\times10^{-3})$	0	.41	2.03	0	−.13	.20	−1.24	−.83
μ_o	−.85	−.84	−.85	−3	−2.99	−3.00	−2.47	−2.98
$\sigma_c^2(\times10^{-3})$	5	10.3	.95	1	.15	.32	291	.89
$\sigma_o^2(\times10^{-1})$	.5	.60	.54	1.1	1.24	1.15	4.37	1.14

NOTE: The transition rates are given in s^{-1}, the current levels in pA, and the noise variances in $(pA)^2$.

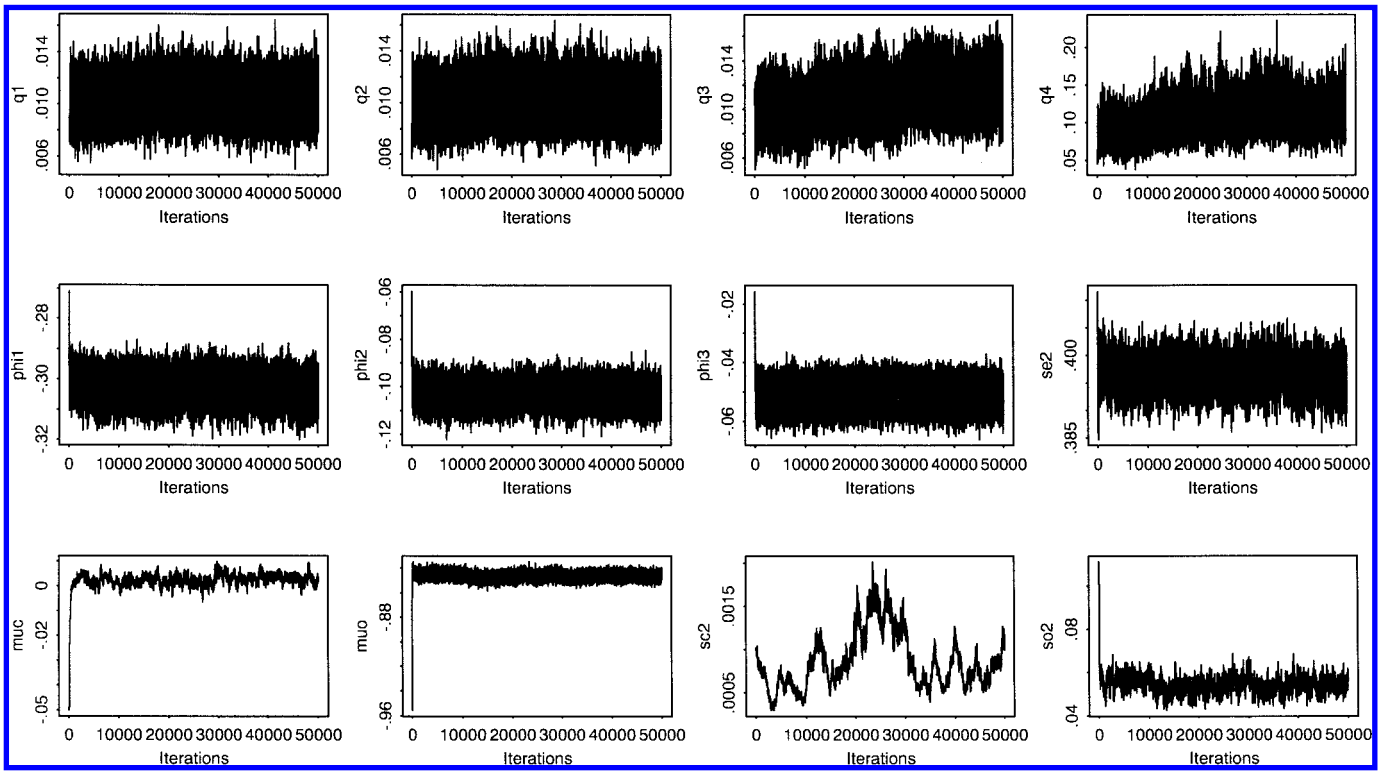

Figure 5. First 50,000 Iterations of Gibbs Sampler for $\Delta_t q$, ϕ, σ_ϵ^2, μ_c,μ_o,σ_c^2, and σ_o^2 (simulation 1). The samples of μ_c and μ_o are given in pA, those of σ_ϵ^2,σ_c^2, and σ_o^2 in (pA)^2.

These intervals are given in Table 4. The results in Table 4 indicate that the sample mean in general seems to be a good estimator of the posterior expectation. Note that these intervals describe only the uncertainty due to simulating from the posterior. It should not be confused with the uncertainty due to insufficient information in the data. So, for example, for q_4 in simulation 1, taking more iterations in the Gibbs sampler may lead to a better estimate of $\mathbb{E}(q_4|y)$, but in view of the rather broad empirical posterior density, this gain is hardly worth the computational burden.

6. CONCLUDING REMARKS

From the results it is clear that the approach outlined in Sections 2–4 is successful in dealing with those aspects that we found most important in our preliminary analysis of the data. In particular, most rapid transitions between the closed and open current levels are recovered satisfactorily. When filtering is neglected, such transitions are sometimes missed, which leads to bias in the estimates. Also, in most cases, the noise parameters and transition rates in the simulations could be estimated well. Simple AR(p) processes fit the background noise reasonably well, which makes the hidden Markov model rather straightforward. More general noise models, including ARMA(p, q) background noise and state-dependent correlations, can in principle be treated in the same way. Computations would become more difficult, however.

Table 2. Percentages of Misclassified States for Gibbs Sampler Without and With Filtering and Hinkley Detector

	% misclassified		
Data	*Gibbs sampler without filter (%)*	*Gibbs sampler with filter (%)*	*Hinkley detector (%)*
Simulation 1	2.9	2.2	20.3
Simulation 2	0.021	0.009	0.018

A drawback of the approach outlined here is the fact that computation times are sometimes rather large. Especially, the updating step of the ion channel states x may be time consuming. When results are needed within a short time, or whenever the algorithm needs to be repeated several times, this may not be acceptable. More efficient programming, faster machines, and additional approximation techniques like those by Fredkin and Rice (1997) can further shorten computation times.

In all the examples we analyzed, it was never found that the Hinkley detector produced the same jump process as did the estimates based on the Gibbs sampler. In the case when a recording has a low noise level, it may, however, not be worth the computational burden of running the proposed Gibbs sampler. Instead, one may base the estimates on the reconstructed, noiseless recording. In the case when the noise level is not low, so that many errors are made while subtracting the noise, or, as is discussed in De Gunst and Schouten (2000), if a model needs to be selected from a class of candidates, the proposed Gibbs sampler is a useful tool.

The MCMC methods discussed by Ball et al. (1999) have smaller computation times than do ours, although we need fewer iterations. We note that our method is especially designed for recordings that exhibit filtering effects and have many rapid transitions between the closed and open states.

Table 3. 95% Highest Posterior Density Credible Regions for All Samples

Variable	Simulation 1	Simulation 2	Recording 1	Recording 2
q_1	[76.3, 129]	[2.04, 5.15]	[6.08, 12.1]	[2.28, 4.87]
q_2	[71.6, 129]	[2.08, 5.64]	[5.26, 11.9]	[1.56, 4.01]
q_3	[74.9, 138]	[116, 176]	[514, 1117]	[103, 158]
q_4	[653, 1479]	[6330, 9560]	[12276, 23935]	[5289, 8583]
ϕ_1	[−.31, −.30]	[−.28, −.27]	[−.013, .014]	[−.29, −.28]
ϕ_2	[−.11, −.09]	[−.10, −.09]	[−.016, .013]	[−.071, −.059]
ϕ_3	[−.060, −.044]	[−.056, −.045]	[−.016, .014]	[−.003, .008]
$\sigma_\epsilon^2(\times 10^{-1})$	[3.91, 4.02]	[.486, .493]	$[1.13\times 10^{-6}, 1.15\times 10^{-6}]$	[.477, .488]
$\mu_c(\times 10^{-3})$	[−2.35, 6.57]	[−.53, 1.00]	[−3.95, 1.30]	[−1.67, −.58]
μ_o	[−.85, −.84]	[−3.01, −2.99]	[−2.49, −2.46]	[−2.99, −2.98]
$\sigma_c^2(\times 10^{-3})$	[.29, 1.58]	[.16, .49]	[289, 293]	[.49, 1.20]
$\sigma_o^2(\times 10^{-1})$	[.45, .62]	[1.10, 1.19]	[4.23, 4.52]	[1.11, 1.17]

Table 4. 95% Confidence Intervals for Posterior Expectations of Transition Rates

Variable	Simulation 1	Simulation 2	Recording 1	Recording 2
q_1	[103, 104]	[3.54, 3.74]	[8.72, 9.11]	[3.52, 3.63]
q_2	[100, 102]	[3.54, 3.80]	[7.87, 8.27]	[2.74, 2.84]
q_3	[103, 106]	[143, 146]	[799, 823]	[132, 133]
q_4	[992, 1050]	[7790, 7890]	[17712, 18116]	[7004, 7079]

For such data, Ball et al.'s method seems less suitable, because there may be many transitions between the sample time points.

Finally, we remark that our approach can in principle be extended to recordings with subconductance levels and recordings of multiple ion channels in a straightforward way. However, more efficient programming may be needed as the size of the state space will be larger and the computation times longer.

[Received January 1999. Revised January 2001.]

REFERENCES

Ball, F. G., Cai, Y., Kadane, J. B., and O'Hagan, A. (1999), "Bayesian Inference for Ion Channel Gating Mechanisms Directly from Single Channel Recordings, Using Markov Chain Monte Carlo," *Proceedings of the Royal Society of London A*, 455, 2879–2932.

Ball, F. G., and Rice, J. A. (1992), "Stochastic Models for Ion Channels: Introduction and Bibliography," *Mathematical Biosciences*, 112, 189–206.

Ball, F. G., and Sansom, M. S. P. (1989), "Ion-Channel Gating Mechanisms: Model Identification and Parameter Estimation from Single Channel Recordings," *Proceedings of the Royal Society of London B*, 236, 385–416.

Ball, F. G., Yeo, G. F., Milne, R. K., Edeson, R. O., Madsen, B. W., and Sansom, M. S. P. (1993), "Single Ion Channel Models Incorporating Aggregation and Time Interval Omission," *Biophysical Journal*, 64, 357–374.

Berger, J. O. (1980), *Statistical Decision Theory. Foundations, Concepts and Methods*, New York: Springer Verlag.

Brooks, S. P. (1998), "Quantitative Convergence Assessment for MCMC via CUSUMS," *Statistics and Computing*, 8, 267–274.

Carter, C. K., and Kohn, R. (1994), "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541–553.

——— (1996), "Markov Chain Monte Carlo in Conditionally Gaussian State Space Models," *Biometrika*, 83, 589–601.

Chung, S. H., Moore, J. B., Xia, L., Premkumar, L. S., and Gage, P. W. (1990), "Characterization of Single Channel Currents Using Digital Signal Processing Techniques Based on Hidden Markov Models," *Philosophical Transactions of the Royal Society of London B*, 329, 265–285.

Chung, S. H., Krishnamurthy, V., and Moore, J. B. (1991), "Adaptive Processing Techniques Based on Hidden Markov Models for Characterizing Very Small Currents Buried in Noise and Deterministic Interferences," *Philosophical Transactions of the Royal Society of London B*, 334, 357–384.

Colquhoun, D., and Sigworth, F. J. (1995), "Fitting and Statistical Analysis of Single Channel Records," in *Single-Channel Recording* (2nd ed.), eds. B. Sakmann and F. Neher, New York: Plenum Press.

De Gunst, M. C. M., and Schouten, J. G. (2000), "The Selection of a Markov Model for the Gating Mechanism of an Ion Channel Using Reversible Jump MCMC," Technical Report WS-543, Free University of Amsterdam, Dept. of Exact Sciences.

Fredkin, R. F., and Rice, J. A. (1992a), "Bayesian Restoration of Single Channel Patch Clamp Recordings," *Biometrics*, 48, 427–448.

——— (1992b), "Maximum Likelihood Estimation and Identification Directly from Single-Channel Recordings," *Proceedings of the Royal Society of London B*, 249, 125–132.

——— (1997), "Fast Evaluation of the Likelihood of an HMM: Ion Channel Currents With Filtering and Colored Noise," preprint.

Frühwirth-Schnatter, S. (1994), "Data Augmentation and Dynamic Linear Models," *Journal of Time Series Analysis*, 15, 183–202.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996), *Markov Chain Monte Carlo Methods in Practice*, London: Chapman & Hall.

Hodgson, M. E. A. (1999), "A Bayesian Restoration of an Ion Channel Signal," *Journal of the Royal Statistical Society B*, 61, 95–114.

Michalek, S., Wagner, M., and Timmer, J. (1998), "A New Approximate Likelihood Estimator for ARMA-Filtered Hidden Markov Models," preprint.

Meyn, S. P., and Tweedie, R. L. (1993), *Markov Chains and Stochastic Stability*, London: Springer-Verlag.

Qin, F., Auerbach, A., and Sachs, F. (1997), "Maximum Likelihood Estimation of Aggregated Markov Processes," *Proceedings of the Royal Society of London B*, 264, 375–383.

Rabiner, L. R., and Juang, B. H. (1986), "An Introduction to Hidden Markov Models," *IEEE Acoustics, Speech and Signal Processing Magazine*, Jan., 4–16.

Raftery, A. E., and Lewis, S. M. (1992), "How Many Iterations in the Gibbs Sampler?," in *Bayesian Statistics 4*, eds. J. M. Bernardo, A. F. M. Smith, A. P. Dawid, and J. O. Berger, Oxford: Oxford University Press.

Robert, C. P., and Casella, G. (1999), *Monte Carlo Statistical Methods*, New York: Springer-Verlag.

Robert, P. R., Celeux, G., and Diebolt, J. (1993), "Bayesian Estimation of Hidden Markov Chains: A Stochastic Implementation," *Statistics and Probability Letters*, 16, 77–83.

Sakmann, B., and Neher, E. (1995), *Single-Channel Recording* (2nd ed.), New York: Plenum Press.

Schouten, J. G. (2000), "Stochastic Modeling of Ion Channel Kinetics," Ph.D. dissertation, Free University of Amsterdam, Dept. of Exact Sciences.

Schultze, R., and Draber, S. (1993), "A Nonlinear Filter Algorithm for the Detection of Jumps in Patch-Clamp Data," *Journal of Membrane Biology*, 132, 41–52.

Shephard, N. (1994), "Partial Non-Gaussian State Space," *Biometrika*, 81, 115–131.

Venkataramanan, L., Walsh, J. L., Kuc, R., and Sigworth, F. J. (1998), "Identification of Hidden Markov Models for Ion Channel Currents, Part I: Colored Background Noise," *IEEE Transactions on Signal Processing*, 46, 1901–1915.

Venkataramanan, L., Kuc, R., and Sigworth, F. J. (1998), "Identification of Hidden Markov Models for Ion Channel Currents, Part II: State-Dependent Excess Noise," *IEEE Transactions on Signal Processing*, 46, 1916–1929.

Vogelzang, S. A., and Prins, H. B. A. (1995), "Kinetic Analysis of Two Simultaneously Activated K^+ Currents in Root Cell Protoplasts of Plantago Media L.," *Journal of Membrane Biology*, 146, 59–71.

West, M. (1995), "Bayesian Inference in Cyclical Component Dynamic Linear Models," *Journal of the American Statistical Association*, 90, 1301–1312.

This article has been cited by:

1. Elan Gin, Larry E. Wagner, David I. Yule, James Sneyd. 2009. Inositol trisphosphate receptor and ion channel models based on single-channel data. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **19**:3, 037104. [CrossRef]
2. M. C. M. Gunst, O. Shcherbakova. 2008. Asymptotic behavior of Bayes estimators for hidden Markov models with application to ion channels. *Mathematical Methods of Statistics* **17**:4, 342-356. [CrossRef]
3. Brenton Clarke, Peter McKinnon. 2005. Robust inference and modelling for the single ion channel. *Journal of Statistical Computation and Simulation* **75**:7, 513-529. [CrossRef]
4. M.C.M. de Gunst, J.G. Schouten. 2005. Model selection and parameter estimation for ion channel recordings with an application to the K+ outward-rectifier in barley leaf. *Journal of Mathematical Biology* **50**:3, 233-256. [CrossRef]
5. R. Nazim Khan, Boris Martinac, Barry W. Madsen, Robin K. Milne, Geoffrey F. Yeo, Robert O. Edeson. 2005. Hidden Markov analysis of mechanosensitive ion channel gating. *Mathematical Biosciences* **193**:2, 139-158. [CrossRef]
6. Iain L. MacDonald, Walter ZucchiniMarkov Models, Hidden . [CrossRef]