ARTICLE

**Yu-Kai The · Jacqueline Fernandez · M. Oana Popa
Holger Lerche · Jens Timmer**

Estimating rate constants from single ion channel currents when the initial distribution is known

Abstract Single ion channel currents can be analysed by hidden or aggregated Markov models. A classical result from Fredkin et al. (*Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer, vol I*, pp 269–289, 1985) states that the maximum number of identifiable parameters is bounded by $2n_o n_c$, where n_o and n_c denote the number of open and closed states, respectively. We show that this bound can be overcome when the probabilities of the initial distribution are known and the data consist of several sweeps.

Keywords Hidden Markov models · Aggregated Markov models · Identifiability · Maximum likelihood estimation · Sodium channel

Introduction

Ion channels are large proteins in the cell membrane regulating the ion concentration in the cell. The proteins can switch between different configurations called states, some of which conduct ions, while others do not. From the gating scheme, i.e. the number of closed and open states and the allowed transitions between these states, it is possible to gain insight into the functioning of the channel.

Y.-K. The (✉) · J. Timmer
Institut für Physik, Albert-Ludwigs-Universität,
Hermann-Herder-Str.3, 79104 Freiburg, Germany
E-mail: yuki@fdm.uni-freiburg.de
Tel.: +49-761-2035764
Fax: +49-761-2035967

Y.-K. The · J. Timmer
Freiburger Zentrum für Datenanalyse und Modellbildung,
Albert-Ludwigs-Universität,
Eckerstr. 3, 79104 Freiburg, Germany

J. Fernandez · M. O. Popa · H. Lerche
Abteilung für Angewandte Physiologie
und Neurologische Klinik, Universität Ulm,
Helmholtzstr. 8/1, 89081 Ulm, Germany

With the development of the patch clamp technique it has become possible to measure the current through single ion channels. Owing to the aggregation of states into closed and open conformations, the transitions between the various states cannot be directly observed. But information about these transitions can be accessed from the dynamical features of the measured current. Therefore, aggregated or hidden Markov models are widely used to analyse single channel data (Fredkin and Rice 1992; Chung et al. 1990). Generalisations of hidden Markov models to cope with coloured noise and filtered data have been developed (Venkataramanan and coworkers 1998a, 1998b, 2000; Qin et al. 2000; Michalek et al. 2000; Fredkin and Rice 2001).

The aggregation of states into closed and open conformations leads to the issues of model equivalence and parameter identifiability. Kienker (1989) gave the necessary and sufficient conditions for two gating schemes to be equivalent. Ito et al. (1992) obtained similar results in a more general context. In the seminal work of Fredkin et al. (1985) it was proven that the number of identifiable parameters cannot exceed $2n_o n_c$, where n_o and n_c denote the number of open and closed states, respectively. When deriving this result, it was assumed that the measurements are performed under stationary conditions so that the initial distribution is not explicitly incorporated into the model but follows directly from the rate constants.

Often the data are measured in several sweeps under the same, non-stationary conditions (Colquhoun et al. 1997), e.g. for voltage-gated channels where every sweep starts with a depolarising voltage pulse (Horn and Lange 1983; Horn et al. 1984; Vandenberg and Horn 1984; Vandenberg and Bezanilla 1991; Böhle and Benndorf 1995a, 1995b; Michalek et al. 1999). In this case the initial distribution does not follow from the rate constants but it can be estimated from the data.

The aim of this paper is to show that it is possible to estimate the parameters of models with more than $2n_o n_c$ rate constants consistently if some elements of the initial probability distribution are fixed to their true value and

the data are measured in several sweeps. Thus, models that otherwise are non-identifiable can be studied. In general, the initial distribution is not known. But in the context of single channel measurements the biological knowledge that certain states are inaccessible at the start of the measurements can be incorporated into a model. In these cases our result applies. After an introduction to the modelling of ion channels and a proof of our result in the "Theory" section, we demonstrate it by a simulation study in the section "Simulation study". In the section "Application" the method is applied to measured data.

Single channel measurements and data analysis

Single channel data were recorded from human embryonic kidney cells (tsA201) transfected with a mutant α-subunit of the adult human skeletal muscle Na^+ channel ($Na_V1.4$). Depolarising pulses were applied from -120 to -20 mV for a duration of 40 ms. The data were low-pass filtered at a frequency of 10 kHz with a four-pole Bessel filter. Seven hundred and fifty sweeps were recorded at a sampling rate of 50 kHz using an Axopatch 200B amplifier and pClamp 8.02 data acquisition. When no overlapping openings were observed, we concluded that there was a single channel in the patch (see also the section "Application" and Horn 1991).

We fitted different gating schemes to the data. To cope with the correlations in the data owing to the anti-aliasing filter the parameters were estimated by an approximate likelihood estimator for autoregressive-moving average (ARMA) filtered hidden Markov models (see Michalek et al. 2000 for details). The standard errors were estimated by the inverse of a numerical approximation to the Hessian matrix of the likelihood (subroutine e04xaf of The Numerical Algorithms Group 1999).

Theory

Modelling of ion channels by hidden Markov models

Background process

Hidden Markov models consist of two stochastic processes. When modelling ion channels the unobserved switching between the different states is described by a finite-state, continuous-time Markov chain X_t. The so-called generator matrix $\mathbf{Q}$ determines the time evolution of this Markov process

$$\dot{P}_i(t) = \sum_j P_j(t)Q_{ji},$$

where $P_i(t)$ denotes the probability of being in state i at time t. The rows of the generator matrix sum to zero and if a transition from state i to state j is not allowed the corresponding entry Q_{ij} of the generator matrix equals zero. The transition probability matrix $\mathbf{A}$ for a time interval Δt is related to the generator matrix $\mathbf{Q}$ by

$$\mathbf{A} = \exp{(\mathbf{Q}\Delta t)}.$$

Observation process

The measured current is described by the measurement or observation process Y_t. For a given realisation x_t of the Markov chain the measured current Y_t is determined by a conditional distribution $f(y_t|x_t)$. This distribution is the same for all closed and open states, respectively.

In the case of aggregated Markov models the observation process Y_t is taken to be discrete with outcomes of open and closed. If the channel is in an open configuration the measurement process Y_t equals 1, with probability 1 indicating that a current is flowing. If the channel is in a closed state $Y_t = 0$ with probability 1.

Compared with aggregated Markov models, hidden Markov models additionally take measurement noise into account which is usually assumed to be white and Gaussian. Thus, $f(y_t|x_t)$ is a Gaussian density with mean value and variance depending on whether x_t is in an open or a closed aggregate.

In the following we will not distinguish between hidden and aggregated Markov models and generally use the term hidden Markov model.

The likelihood L is calculated by

$$L = \sum_{x_1,\ldots,x_N=1}^m \pi_{x_1} f(y_1|x_1) \prod_{t=2}^N A_{x_{t-1}x_t} f(y_t|x_t),$$

where m denotes the number of states and N the number of data points. The vector π describes the initial probability of being in one of the background states. If the measurements are performed with several sweeps such expressions have to be calculated for every sweep. The complete likelihood results by summing up all these terms. Efficient methods for calculating and optimising the likelihood are described, for example, by Rabiner (1989).

Fixing the initial distribution

In this section we sketch why it is possible to use information known about the initial distribution to gain information on the rate constants. A detailed mathematical description is given in the "Appendix".

Our result is based mainly on the equivalence of models worked out by Kienker (1989). In that paper it is proven that two hidden Markov models $(\mathbf{Q},\pi)$ and $(\mathbf{Q},\pi')$ are equivalent if and only if there exists a Matrix $\mathbf{S}$ such that $\mathbf{Q}' = \mathbf{S}^{-1}\mathbf{Q}\mathbf{S}$ and $\pi' = \pi\mathbf{S}$. Two models are called equivalent if they generate the same observable distributions.

Consider the case of a continuous-time hidden Markov model with generator matrix $\mathbf{Q}$ that has one parameter too many to be identifiable. The generator matrix $\mathbf{Q}$ and the initial distribution π are parameterised by two vectors θ_Q and θ_π, respectively. The non-identifiability of the generator matrix implies that for every parameter vector of rate constants θ_Q there is another parameter θ'_Q such that the two generator matrices obtained from θ_Q and θ'_Q yield two equivalent hidden Markov models. If the generator matrix has only one parameter too many to be identifiable there exists a one-dimensional curve in the space of the parameter θ_Q along which all corresponding models are equivalent. Such a curve is related to a family of similarity transformations $\mathbf{S}$ as introduced earlier which also acts on the initial distribution. Thus, there exists a corresponding curve in the space of the parameter of the initial distribution θ_π. Fixing one component of the initial distribution constrains these two curves and the hidden Markov model becomes identifiable.

Results

Simulation study

To illustrate our result numerically we examined the three-state model sketched in Fig. 1. The rate constants are given in Hertz and the sampling frequency was chosen as 25 kHz. The model has more than $2n_o n_c$ parameters even if one rate constant is determined by the principle of detailed balance. For gating schemes containing loops detailed balance is preserved if the product of the rate constants in a clockwise direction around the loop equals the product of the rate constants counterclockwise (Colquhorn and Hawkes 1982).

It can be shown that the assumptions from the "Theory" section are fulfilled and the dimension of the vector ε is 1.

We simulated data sets consisting of 256, 512, 1024 and 2048 sweeps. The initial distribution for every sweep was chosen as $\pi = (0.8, 0.0, 0.2)$. The equilibrium distribution $\mathbf{p}$ of this model can be calculated from the $\mathbf{Q}$ matrix by $\mathbf{pQ} = \mathbf{0}$. This yields $\mathbf{p} = (0.35354, 0.40404, 0.24242)$.
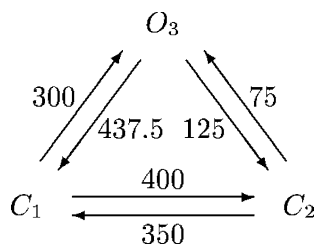

Fig. 1 The simulated model. We denote an open and a closed state by O and C, respectively. The model obeys the law of detailed balance. The rate constants are given in Hertz

For the data set with 256 sweeps each sweep had a length of 8,192 data points. For each of the 512 sweeps we generated 4,096 data points. The 1,024 and 2,048 sweeps consisted of 2,048 and 1,024 data points, respectively. Thus, every data set had a total number of 2^{21} data points.

From these data we calculated the maximum likelihood estimators of the parameters of the generator matrix under the constraint that detailed balance is fulfilled. The first component of the initial distribution was also estimated from the data by maximising the likelihood, the second component was set to zero and the third component was determined by normalisation to 1. Note that Eq. 2 from the theorem is not fulfilled if the last component of the initial distribution is fixed (see the "Discussion"). The maximisation of the likelihood was performed directly by a quasi-Newton method (subroutine e04ucf of The Numerical Algorithms Group 1999).

From 500 replications we calculated the mean values and estimated the standard deviation of the maximum likelihood estimator of the rate constants. The results are summarised in Table 1. As confidence intervals of the mean values the estimated standard deviations divided by $\sqrt{500}$ are given. Histograms of the estimated rate constants for 256 and 2,048 sweeps are shown in Fig. 2.

The standard deviation decreases when the number of sweeps increases, although the total number of data points remains constant. For the data set consisting of 2,048 sweeps all parameters are estimated without bias. The second parameter is upwardly biased in the case of 256, 512 and 1,024 sweeps, whereas the third parameter is underestimated. In the case of 256 and 512 sweeps the first parameter shows a bias downwards. The magnitude of all biases decreases as the number of sweeps increases.

In Fig. 3 we plot the 500 estimated values of the rate constants Q_{12} and Q_{13} for 256 and 2,048 sweeps. The estimates for the 2,048 sweeps are less scattered around the true value than the values for the 256 sweeps. Moreover, the distribution of the estimates in the lower plot looks approximately Gaussian, while the other does not.

In Fig. 4 we investigate the estimated condition number of the Hessian matrix of the maximum likelihood point. The condition number of the Hessian matrix is a measure for the non-identifiability of the parameters (Wagner et al. 1999). It decreases as the number of sweeps increases. The plot with logarithmic axes shows evidence that the decrease is inversely proportional to the number of sweeps used.

Application

In this section we apply our method to real Na^+-channel data. The sodium channel is a protein that consists of four domains (I–IV), each of which contains six transmembrane segments (S1–S6). The S4 segments contain

Table 1 Mean values and estimated standard deviations (*SD*) of the maximum likelihood estimators of the rate constants

True values	256 sweeps		512 sweeps		1,024 sweeps		2,048 sweeps	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
$Q_{12} = 400$	388.5 ± 5.1	113.5	393.2 ± 4.5	100.0	396.0 ± 4.1	90.6	403.8 ± 3.7	83.8
$Q_{21} = 350$	376.4 ± 7.2	161.0	369.0 ± 6.0	133.7	362.8 ± 4.6	103.4	353.7 ± 3.2	71.5
$Q_{23} = 75$	66.8 ± 1.7	37.2	68.9 ± 1.4	30.2	71.3 ± 1.0	22.0	73.8 ± 0.7	15.3
$Q_{32} = 125$	125.8 ± 3.8	86.9	123.7 ± 3.2	71.2	123.8 ± 2.4	54.1	125.9 ± 1.8	39.7
$Q_{13} = 300$	303.2 ± 1.8	39.3	302.3 ± 1.4	31.8	299.9 ± 1.0	23.2	300.9 ± 0.8	18.2

The parameter Q_{31} is determined by the principle of detailed balance. The confidence intervals for the mean values are given by the SD divided by $\sqrt{500}$
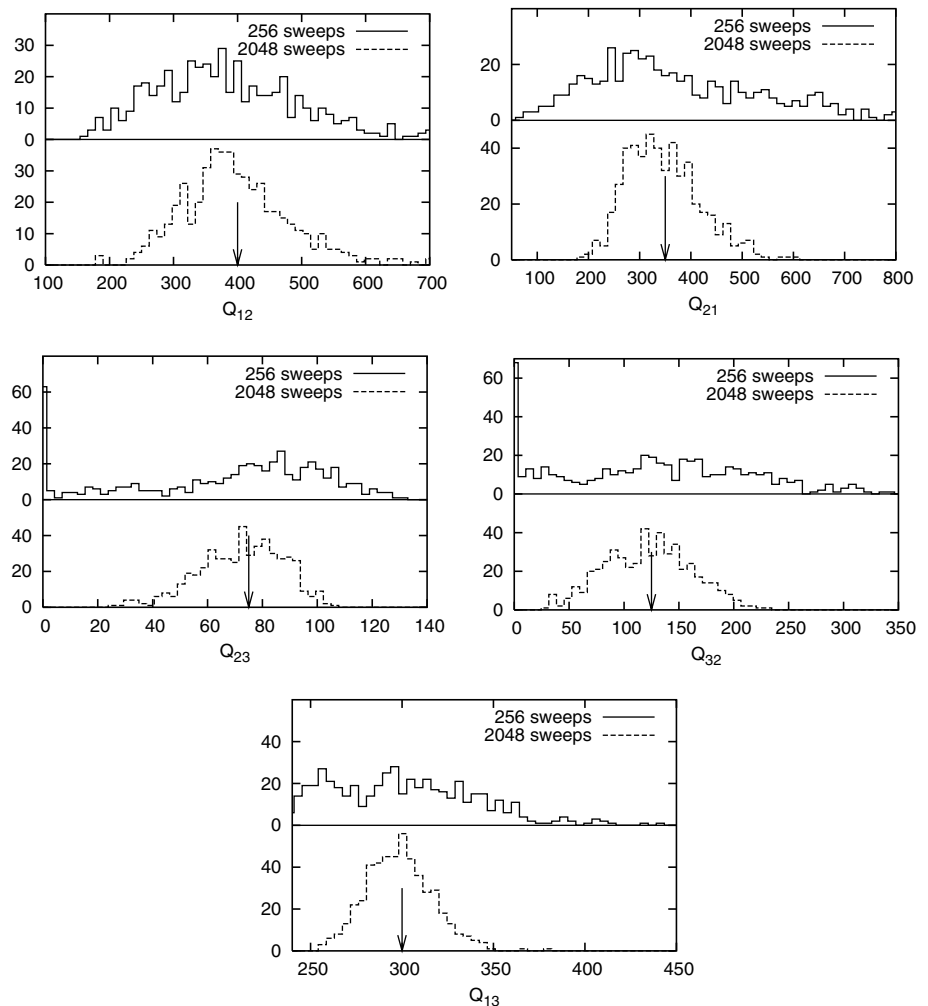
positively charged amino acids constituting the so-called voltage sensors that can move when a voltage is applied. The movement of the voltage sensors causes the channel to open (Catterall 2000).

Single channel records show that during a maintained depolarisation the wild-type sodium channel after some brief initial openings eventually closes and remains closed. A part of the protein located between domains III and IV on the intracellular side of the membrane is responsible for this so-called inactivation. It is believed that there is a hydrophobic particle (IFM) blocking the pore (West et al. 1992). This particle contains the amino acid sequence isoleucine–phenylalanine–methionine, which is commonly abbreviated by IFM.

A simple gating scheme for the wild-type sodium channel derived from these considerations is sketched in Fig. 5. It is believed that at the start of the depolarisation the channel is in the left-most state and all four voltage sensors are in their resting position. The transition between two closed states corresponds to the

Fig. 2 The histograms of the estimated rate constants for 256 sweeps (*top*) and 2,048 sweeps (*bottom*). The *arrows* mark the true values
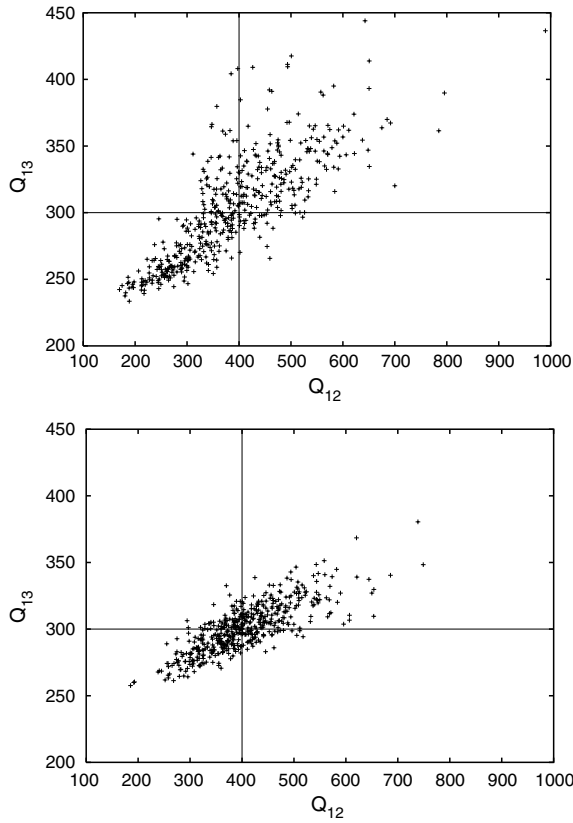
Fig. 3 The estimated parameters Q_{12} and Q_{13} for 256 (*top*) and 2,048 (*bottom*) sweeps. The true values are marked by the *solid lines*

movement of one of the voltage sensors. When all voltage sensors have moved, the channel is open. Transitions from both the closed and the open conformations to the inactivated state are possible owing to the hydrophobic particle (IFM) blocking the pore from inside the cell. Similar models have been found by Horn and Vandenberg (1984) and Michalek et al. (1999).
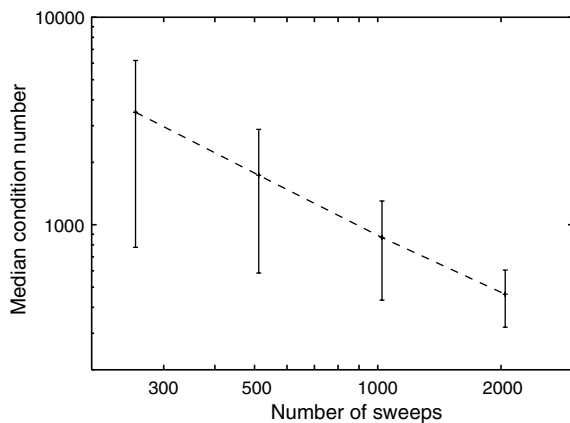

Fig. 4 The median condition number of the estimated correlation matrix with logarithmic axes. The *error bars* indicate the median absolute deviation of the condition number from the median condition number
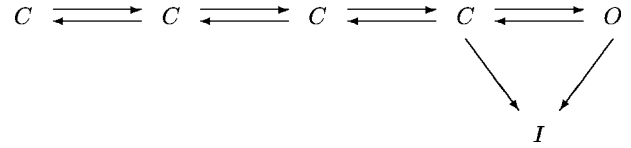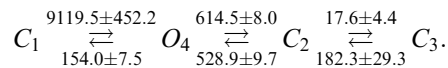

Fig. 5 A simple model for the wild-type channel. Closed, open and inactivated states are denoted by C, O and I, respectively

Since the four closed states cannot be distinguished from the magnitude of the current and are mainly occupied at the beginning of the voltage pulse, it cannot be expected that, in practice, all four states can be resolved from measurements. Most published models for the sodium channel fitted to real data had fewer than four closed states that are related to the movement of the voltage sensor.

The mutant we used in this study contains two different amino acids in segments S4–S5 of III and IV than the wild-type channel. We replaced isoleucine with cysteine in position 1160 and leucine with alanine in position 1482. This mutation causes the channel not to inactivate properly (Popa et al. 2002). One typical trace of raw data is shown in Fig. 6. A preliminary analysis of the amplitude histogram (not shown) shows that the probability of an opening is approximately $P_{open} \approx 0.44$. If these openings were due to two independent, identical channels a lower bound for the probability that both channels open simultaneously can be estimated as 0.04. This should be observable in the data and we concluded that the patch contained only one channel.

Different physiologically plausible models with at most $2n_o n_c$ parameters have been fitted to the data. A detailed discussion of the models and the model selection process will be presented elsewhere. The resulting model was

$$C_1 \underset{154.0 \pm 7.5}{\overset{9119.5 \pm 452.2}{\rightleftarrows}} O_4 \underset{528.9 \pm 9.7}{\overset{614.5 \pm 8.0}{\rightleftarrows}} C_2 \underset{182.3 \pm 29.3}{\overset{17.6 \pm 4.4}{\rightleftarrows}} C_3.$$
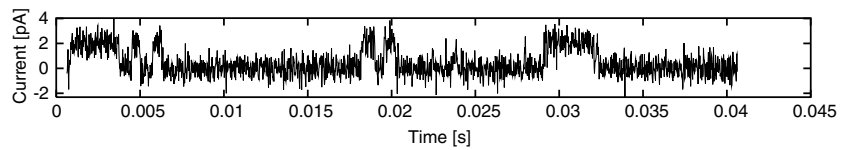
The vector of the initial probability distribution was estimated as follows:

$$\pi = (\pi_c, \pi_o) = (0.709, 0.255, 0.0, 0.036).$$

The high initial probability of state C_1 suggests that this state corresponds to the four closed states of the wild-type sodium channel in Fig. 5. The inactivated state of the wild-type channel transforms into the closed states C_2 and C_3. If this was true, one would expect direct transitions between states C_1 and C_2 corresponding to the transition from the closed aggregate to the inactivated state of the wild-type channel. Such a model is not identifiable. But since in this case the channel should occupy state C_1 at the start of the pulse the initial probabilities of states C_2 and C_3 can be constrained to zero and the estimation of the generator matrix is possible. The results are summarised in Fig. 7.

The comparison of this result with the estimation of the C–O–C–C model discussed earlier shows that the estimated standard deviations are of the same order of

Fig. 6 A trace of raw data. We show the section of 40 ms during which the depolarising pulse is applied

magnitude for both models. Also, the rate constants Q_{23} and Q_{32} are equal to the corresponding constants in the C–O–C–C model. The other rate constants are of the same order of magnitude as their respective counterparts.

We also fitted a model without the constraint of detailed balance but omitted it here since a likelihood ratio test does not reject the null hypothesis that the principle of detailed balance is fulfilled. Again, the standard errors of the parameter estimates are similar to the errors of the C–O–C–C model, which indicates that the parameter is identifiable.

Discussion

Parameter identifiability is a major issue when analysing ion channel currents by hidden Markov models. For this task Fredkin et al. (1985) gave an easily checkable upper bound on the number of identifiable parameters and a more refined one was given by Fredkin and Rice (1986). Generally, the exact number of identifiable parameters is not yet known and there are models with non-identifiable parameters where this bound is not achieved (Wagner et al. 1999).

In this paper we show under which conditions it is possible to estimate parameters of a hidden Markov model even though the number of parameters exceeds the bound $2n_o n_c$. To this end, biological knowledge about the initial distribution of the channel has to be available.

The estimation procedure is based on the fact that equivalent models are related by a similarity transformation. For a parameter that is not identifiable there exists a connected subset of the parameter space on

which the likelihood is constant. A family of similarity transformations generates this subset. Constraining the initial distribution also fixes this family of transformations if Eqs. 1 and 2 are fulfilled.

If in our simulation study we had fixed the last entry of the initial distribution, the family of similarity transformations would not have been fixed and Eq. 2 would not have been fulfilled. This can be seen directly if one considers the form of the similarity transformations. Since the model has only one open state, the sub-matrix S_o of a valid similarity transformation also has only one entry. Owing to row normalisation this entry has to be 1. Thus, constraining the initial distribution of the open state does not fix the family of similarity transformations.

The simulation study shows that the error bounds of the parameter estimates become smaller as the number of sweeps is increased (Table 1, Figs. 2, 3). Furthermore, the parameter estimates are biased if the number of sweeps does not suffice. When the data are measured under non-stationary conditions, the transient decays exponentially and the channel reaches its steady state, which contains no more information about the initial distribution. Since the rate constants are related to the initial distribution via a similarity transformation, the accuracy of the parameter estimates increases with the number of sweeps measured even if the total number of data points remains the same. This dependence is illustrated by Fig. 4, which suggests that the condition number of the correlation matrix is inversely proportional to the number of sweeps measured.

Note that the information about the rate constants is entirely due to the fact that the initial distribution is known and not to the fact that the data are measured under non-stationary conditions. The estimation also works when the data are generated under stationary conditions, but in this case it is not to be expected that the initial distribution is known.

Our application indicates that it can be helpful to incorporate biological knowledge into a model and to constrain the initial distribution to fixed values. The comparison of both fitted models presented in the "Simulation study" section shows that all respective rate constants do not differ much, providing evidence that the model in Fig. 7 fits the data well. Moreover, the standard errors are moderate, which demonstrates that the parameters are identifiable.

Thus, when single channel measurements are performed in several sweeps it is possible to fit models to the data that otherwise are not identifiable.
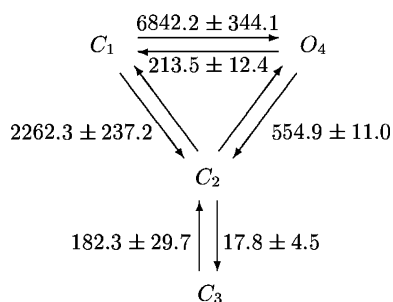

Fig. 7 The estimated rate constants when the initial probabilities of C_2 and C_3 are constrained to zero. One of the missing rate constants is estimated as $Q_{24} = 483.5 \pm 10.7$. The other parameter, $Q_{21} = 61.5 \pm 8.6$, is determined by the principle of detailed balance

Appendix: proof of the main theorem

Consider a continuous-time hidden Markov model $(\mathbf{Q},\pi)$ with n_c closed and n_o open states. The states are numbered such that states $1, ..., n_c$ are the closed states and $n_c + 1, ..., n_c + n_o = m$ are the open states. Let $\mathbf{Q}$, the generator matrix of the hidden Markov model, be partitioned into

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{cc} & \mathbf{Q}_{co} \\ \mathbf{Q}_{oc} & \mathbf{Q}_{oo} \end{pmatrix}.$$

The initial distribution is denoted by the row vector $\pi = (\pi_c, \pi_o)$.

The generator matrix $\mathbf{Q}$ and the initial distribution π are parameterised by vectors θ_Q and θ_π, respectively, and $\theta = (\theta_Q, \theta_\pi) \in \Theta \subset \mathbf{R}_+^k \times [0, 1]^{n-1}$. Let the mapping $\mathscr{F} : \theta_Q \to \mathbf{Q}(\theta_Q)$ denote the parameterisation of the generator matrix $\mathbf{Q}$. We require $\mathscr{F}$ to be injective; thus, the parameter vector θ_Q can be determined from the rate constant. The parameter θ_π contains $n-1$ elements of the initial distribution itself.

Now we give some technical conditions for the kind of non-identifiability we investigate. Consider an arbitrary but fixed parameter θ. We assume that there exists a continuous family of similarity transformations as introduced by Kienker (1989), $\mathbf{S} : \varepsilon \to \mathbf{S}(\varepsilon) \in \mathbf{R}^{n \times n}$, $\varepsilon \in U \subset \mathbf{R}^M$. Here U is a neighbourhood around zero that is sufficiently small. The similarity transformations fulfil the following conditions:

1. $\mathbf{S}(0) = 1$.
2. $\mathbf{S}$ has the form

$$\mathbf{S} = \left(\begin{array}{c|c} \mathbf{S}_c & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{S}_o \end{array} \right).$$

3. The rows of $\mathbf{S}$ sum to 1.
4. For all θ' in a neighbourhood of θ such that the hidden Markov model $[\mathbf{Q}(\theta'_Q), \pi(\theta'_\pi)]$ is equivalent to $[\mathbf{Q}(\theta_Q), \pi(\theta_\pi)]$ there exists an $\epsilon \in U$ such that $\mathbf{Q}(\theta'_Q) = \mathbf{S}^{-1}(\varepsilon)\mathbf{Q}(\theta_Q)\mathbf{S}(\varepsilon)$ and $\pi(\theta'_\pi) = \pi(\theta_\pi)\mathbf{S}(\varepsilon)$.

Conditions 1–3 ensure that the matrices $\mathbf{Q}' = \mathbf{S}^{-1}(\varepsilon)\mathbf{Q}\mathbf{S}(\varepsilon)$ are equivalent to the matrix $\mathbf{Q}$. The corresponding initial distributions are given by $\pi' = \pi\mathbf{S}$. If the parameter point θ is locally non-identifiable, there exists a parameter θ' in every neighbourhood of θ such that the hidden Markov model $[\mathbf{Q}(\theta'_Q), \pi(\theta'_\pi)]$ is equivalent to $[\mathbf{Q}(\theta_Q), \pi(\theta_\pi)]$. Condition 4 ensures that all equivalent parameters in a neighbourhood of θ are covered by the family of similarity transformations.

Loosely speaking, the hidden Markov model $[\mathbf{Q}(\theta), \pi(\theta)]$ has at most M parameters too many to be identifiable. This is, for example, the case when the dimension of the parameter vector θ_Q is greater than $2n_o n_c$.

Now, consider the sub-model that results from fixing M components of the vector θ_π. We investigate the conditions under which such a sub-model is identifiable. Let $\mathbf{a} = (\mathbf{a}_c, \mathbf{a}_o) \in \mathbf{R}^{M_c} \times \mathbf{R}^{M_o} = \mathbf{R}^M$ denote the vector of the fixed values of the initial distribution. Without loss of generality the initial distribution of the sub-model can be written as $\pi = (\mathbf{a}_c, \tilde{\pi}_c, \mathbf{a}_o, \tilde{\pi}_o)$. With $\mathbf{J}(\varepsilon) = \mathbf{S}(\varepsilon) - 1$ and a natural partitioning of $\mathbf{J}$ into $\mathbf{J}_c$ and $\mathbf{J}_o$ we are able to state our main result.

Theorem 1 *Let $[\mathbf{Q}(\theta), \pi(\theta)]$ be a parameterised hidden Markov model with parameter vector θ that fulfils conditions 1–4. It holds for all $\varepsilon \in U \setminus \{0\}$ that*

$$\sum_{i=1}^{n} (\boldsymbol{\pi}_c)_i [\mathbf{J}_c(\boldsymbol{\varepsilon})]_{ij} \neq 0 \tag{1}$$

for all $j = 1, ..., M_c$ and

$$\sum_{i=1}^{n} (\boldsymbol{\pi}_o)_i [\mathbf{J}_o(\boldsymbol{\varepsilon})]_{ij} \neq 0 \tag{2}$$

for all $j = 1, ..., M_o$.

Then the parameter vector $\tilde{\boldsymbol{\theta}} = (\boldsymbol{\theta}'_Q, \tilde{\boldsymbol{\pi}}'_c, \tilde{\boldsymbol{\pi}}'_c)$ of the sub-model with $\pi = (\mathbf{a}_c, \tilde{\pi}_c, \mathbf{a}_o, \tilde{\pi}_o)$ is identifiable.

Proof Assume that there exists a parameter $\tilde{\boldsymbol{\theta}}' = (\boldsymbol{\theta}'_Q, \tilde{\boldsymbol{\pi}}'_c, \tilde{\boldsymbol{\pi}}'_o)$ for the sub-model that is equivalent to the parameter $\tilde{\theta}$. Following condition 4, there exists an $\epsilon \in U$ with $\mathbf{Q}' = \mathbf{Q}(\theta'_Q) = \mathbf{S}^{-1}\mathbf{Q}(\theta_Q)\mathbf{S}$. The corresponding initial probability is given by $\pi' = (\pi_c\mathbf{S}_c, \pi_o\mathbf{S}_o)$. But this initial probability has to satisfy the constraint $\boldsymbol{\pi}' = (\mathbf{a}_c, \tilde{\boldsymbol{\pi}}'_c, \mathbf{a}_o, \tilde{\boldsymbol{\pi}}'_o)$. This can be rewritten as

$$\sum_{i=1}^{n} (\boldsymbol{\pi}_c)_i [\mathbf{J}_c(\boldsymbol{\varepsilon})]_{ij} = 0$$

for all $j = 1, ..., M_c$ and

$$\sum_{i=1}^{n} (\boldsymbol{\pi}_o)_i [\mathbf{J}_o(\boldsymbol{\varepsilon})]_{ij} = 0$$

for all $j = 1, ..., M_o$, which is a contradiction to the assumptions. $\square$

References

Böhle T, Benndorf K (1995a) Multimodal action of single Na^+ channels in myocardial mouse cells. Biophys J 68:121–130

Böhle T, Benndorf K (1995b) Voltage-dependent properties of three different gating modes in single cardiac Na^+ channels. Biophys J 69:873–882

Catterall WA (2000) From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channel. Neuron 26(1):13–25

Chung S-H, Moore JB, Xia L, Premkumar LS, Gage PW (1990) Characterization of single channel currents using digital signal processing techniques based on hidden Markov models. Philos Trans R Soc Lond B 329:265–285

Colquhoun D, Hawkes AG (1982) On the stochastic properties of bursts of single ion channel openings and of clusters of bursts. Philos Trans R Soc Lond B 300:1–59

Colquhoun D, Hawkes AG, Merlushkin A, Edmonds B (1997) Properties of single ion channel currents elicited by a pulse of agonist concentration or voltage. Philos Trans R Soc Lond A 355:1743–1786

Fredkin DR, Rice JA (1986) On aggregated Markov processes. J Appl Prob 23:208–214

Fredkin DR, Rice JA (1992) Maximum likelihood estimation and identification directly from single-channel recordings. Proc R Soc Lond B 249:125–132

Fredkin DR, Rice JA (2001) Fast evaluation of the likelihood of an HMM: ion channel currents with filtering and colored noise. IEEE Trans Signal Process 49(3):625–633

Fredkin DR, Montal M, Rice JA (1985) Identification of aggregated Markovian models: application to the nicotinic acetylcholine receptor. In: Le Cam LM, Olshen RA (eds) Proceedings of the Berkeley conference in Honor of Jerzy Neyman and Jack Kiefer, volume I, pp 269–289

Horn R (1991) Estimating the number of channels in patch recordings. Biophys J 60:433–439

Horn R, Lange K (1983) Estimating kinetic constants from single channel data. Biophys J 43:207–223

Horn R, Vandenberg CA (1984) Statistical properties of single sodium channels. J Gen Physiol 84:505–534

Horn R, Vandenberg C, Lange K (1984) Statistical analysis of single sodium channels. Biophys J 45:323–335

Ito H, Amari SI, Kobayashi K (1992) Identifiability of hidden Markov information sources and their minimum degree of freedom. IEEE Trans Inf Theory 38(2):324–333

Kienker P (1989) Equivalence of aggregated Markov models of ion-channel gating. Proc R Soc Lond B 236:269–309

Michalek S, Lerche H, Wagner M, Mitrovic N, Schiebe M, Lehmann-Horn F, Timmer J (1999) On identification of Na^+-channel gating schemes using moving-average filtered hidden Markov models. Eur Biophys J 28(7):605–609

Michalek S, Wagner M, Timmer J (2000) A new approximate likelihood estimator for ARMA-filtered hidden Markov models. IEEE Trans Signal Process 48(6):1537–1547

Popa MO, Alekov AK, Maljevic S, Mitrovic N, Lehmann-Horn F, Lerche H (2002) Cooperative effect on fast inactivation of S4-S5 loops of domains III and IV of the Na^+ channel ($hNa_{V\ 1.4}$). Biophys J 82:90a–91a

Qin F, Auerbach A, Sachs F (2000) Hidden Markov modeling for single channel kinetics with filtering and correlated noise. Biophys J 79:1928–1944

Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 77:257–285

The Numerical Algorithms Group Ltd. (1999) Fortran Library Mark 19. The Numerical Algorithms Group Ltd., Oxford

Vandenberg CA, Bezanilla F (1991) Single-channel, macroscopic, and gating currents from sodium channels in the squid giant axon. Biophys J 60:1499–1510

Vandenberg CA, Horn R (1984) Inactivation viewed through single sodium channels. J Gen Physiol 84:535–564

Venkataramanan L, Kuc R, Sigworth F (1998) Identification of hidden Markov models for ion channel currents—Part II: state-dependent excess noise. IEEE Trans Signal Process 46(7):1916–1929

Venkataramanan L, Walsh JL, Kuc R, Sigworth F (1998) Identification of hidden Markov models for ion channel currents—Part I: colored background noise. IEEE Trans Signal Process 46(7):1901–1915

Venkataramanan L, Kuc R, Sigworth FJ (2000) Identification of hidden Markov models for ion channel currents—Part III: bandlimited, sampled data. IEEE Trans Signal Process 48(2):376–385

Wagner M, Michalek S, Timmer J (1999) Estimating transition rates in aggregated Markov models of ion-channel gating with loops and with nearly equal dwell times. Proc R Soc Lond B 266:1919–1926

West JW, Patton DE, Scheuer T, Wang Y, Goldin AL, Catterall WA (1992) A cluster of hydrophobic amino acid residues required for fast Na^+-channel inactivation. Proc Natl Acad Sci USA 89(22):10910–10914