ONE-STEP AHEAD PREDICTION FOR PARAMETER ESTIMATION IN PHYSIOLOGICAL HYBRID MODELS

Natal A.W. van Riel¹², Aleksandar Lj. Juloski² and Jorn op den Buijs²

¹Eindhoven University of Technology, Department of Biomedical Engineering, P.O. box 513, 5600 MB Eindhoven, The Netherlands. Tel: +31 40 247 5506, Fax: +31 40 243 4582, e-mail: N.A.W.v.Riel@tue.nl

Abstract: Identification of hybrid systems is complex because it is not a priori known which data is generated by which discrete mode of the system. A novel procedure is presented for the simultaneous determination of the model parameters and classification of the data. Initial parameter estimates are obtained based on a priori knowledge of the modes of the system. Next, the data points are subsequently classified and with each new classified data point the parameter estimates are refined. The proposed procedure is applied in the estimation of parameters in a hybrid description of calcium cycling in the intact heart. Copyright © 2005 IFAC

Keywords: hybrid models, identification algorithms, least-squares identification, parameter estimation, physiological models, prediction error methods.

1. INTRODUCTION

Besides in the system theory and control engineering community, hybrid models start to gain interest for applications in life sciences research, particularly in the rapidly emerging area of systems biology (e.g. Belta et al., 2001; Gosh et al., 2002). A hybrid system modeling paradigm is required to treat the combined discrete and continuous characteristics of physiological and biomolecular systems (McAdams and Shapiro, 2003). Processes that might operate on a fast continuous time-scale at a molecular level often appear as discrete events at the more macroscopic level of the experimental observations. Hybrid models can provide a computationally efficient approximation of the largely different time scales in biological systems. Hybrid models often arise naturally as will be illustrated in this paper by an example of ion currents in excitable cells. The purpose of mathematical models in life sciences research is often to test and generate hypotheses of the function of the system in health and disease. This requires a parametric model structure in which model variables and parameters have a physiologically meaningful interpretation. Often, a priori information is available that can be used to derive a suitable model structure. The possibility to extract (situation specific) parameter values from experimental data provides an important next step. Modern techniques enable experiments to generate real-time data for parameter estimation in dynamic grey-box hybrid models of certain biological systems. The complexity of identification of hybrid systems stems from the problem that it is not a priori known which data are generated by which discrete mode of the system. Furthermore, the determination of the model parameters of the continuous dynamics and data classification have to be accomplished simultaneously (Ferrari-Trecate et al., 2003).

The problem of identification of hybrid systems has only recently been considered (e.g. Ferrari-Trecate et al., 2003; Juloski et al., 2004a). These studies followed a black-box approach, using Piece-Wise Auto-Regressive eXogenous (PWARX) models. Grey-box approaches for hybrid descriptions of physiological systems have been previously employed using heuristic signal analysis and data preprocessing to first classify the data, irrespective of the subsequent parameter estimation problem (Op den Buijs et al., 2005; Van Riel and Van der Vusse 2002; Van Riel et al., 2003). Obtained estimates were
highly sensitive to the (erroneous) classification of the data.
In this paper we present a novel procedure for parameter estimation and simultaneous data classification in hybrid systems. The model class and identification problem are formally introduced in section 2. In section 3 the algorithm will be proposed, followed by a physiological example in section 4. In the final section the results are discussed and placed into perspective.

2. PROBLEM STATEMENT

Time-switched, nonlinear systems of the form:

\[ \dot{x}(t) = f_{\sigma(t)}(x(t)) \] (1)

are considered, with \( t \geq 0, \sigma : \mathbb{R}^+ \rightarrow \{1, \ldots, s\} \) a switching function, which is assumed to be arbitrary, but not known and \( \{f_1(\cdot, \theta_1), \ldots, f_s(\cdot, \theta_s)\} \) a known set of nonlinear functions, parameterized by \( \theta = [\theta_1, \ldots, \theta_s]^T \). The number of modes \( s \) is assumed to be known. Note that not all parameters \( \theta_i \) of the different modes are required to be of the same dimension.

The states \( x \) are measured at equidistant discrete time instants \( kT_s \) and \( N \) samples are collected according to the measurement equation:

\[ y(kT_s) = x(kT_s) + e(kT_s) \] (2)

where \( k = 1, \ldots, N \) and \( e(k) \) is additive noise. With the slight abuse of notation, in the remainder of the text we will use \( k \) instead of \( kT_s \) to indicate the sampled system. The identification problem consists of estimating the values of the unknown parameter vectors \( \theta_i \) \( (i = 1, \ldots, s) \), given the collected dataset \( \{y(1), \ldots, y(N)\} \).

Since the measurements are collected at discrete time instants we will discretize the switched dynamics (1) using an Euler approximation:

\[ \frac{x(k+1) - x(k)}{T_s} = f_{\sigma_{\chi_i}}(x(k), \theta_{\sigma_{\chi_i}}) \] (3)

Note that (3) is not an exact discretization of (1), since it does not account for the cases when the mode switches between the sampling instants. If it is assumed that the collected experimental data is generated by model (3), the index sets of the modes can be defined according to:

\[ \chi_i = \{k \mid \sigma(k) = i\} \] (4)

If it is known that \( \sigma(k) = i \), then the one-step ahead prediction can be computed on the basis of the measured output:

\[ \dot{x}(k+1) = y(k) + T_s f_i(y(k), \theta_i) \] (5)

The parameters can be estimated as the solution of the following nonlinear least squares problem:

\[ \hat{\theta} = \arg \min_\theta J(\theta) \] (6)

where the cost function \( J \) has the form:

\[ J = \sum_{i=1}^{s} \sum_{k \in \chi_i} \left( \hat{y}(k+1) - y(k+1) \right)^2 \] (7)

where \( \hat{y}(k+1) \) is computed as in (5). If it is assumed that the noise \( e(k) \) has a Gaussian distribution (zero mean white noise with a normal distribution \( N(0, \sigma) \)), then (6) and (7) yield a Maximum Likelihood Estimator (e.g. Ljung, 1999). The covariance matrix \( R_y \) of the estimate (6) can be computed from (J (7)):

\[ R_y = \frac{1}{N} \left( \frac{\partial^2 J}{\partial \theta^2} \right)^{-1} \] (8)

3. ALGORITHM

Since \( \sigma(k) \) is not known, the sets \( \chi_i \) are not a priori available and they have to be reconstructed by the identification procedure. The problem of assigning data points to sets \( \chi_i \) is the problem of data classification. If we attempt to construct the sets \( \chi_i \) directly, a combinatorial optimization problem would emerge. Hence we have to resort to suboptimal solutions and we propose to reconstruct the sets \( \chi_i \) iteratively. We will classify data points one at a time and simultaneously refine the estimate of \( \theta \). Usually some data points can be classified a priori to one of the modes of the system, as will be illustrated in the example in section 4. This key observation forms the basis of our procedure.

With \( \chi_i \) we denote the approximation to the set \( \chi_i \) in the \( i \)-th step of the algorithm. It is assumed that initial sets \( \chi_i^0 \) that contain several data points for each of the modes, are available and that the initial estimate \( \hat{\theta}^0 \) has been obtained using these sets and (6). In order to obtain \( \hat{\theta}^i \), theoretically no more than \( p_i + 1 \) data points are required in each \( \chi_i^0 \), where \( p_i \) is the number of parameters to be estimated for mode \( i \).

The main idea of the algorithm is to use the currently available estimate of \( \theta \) to determine which submodel \( f_i \) better predicts the following data point, to classify the data point to the respective set \( \chi_i \), and to reestimate the value of \( \theta \). The identification algorithm is formally stated as follows:

**Algorithm 3.1 Classification and parameter estimation**

1. Initialization: obtain \( \chi_1^0, \ldots, \chi_s^0, \hat{\theta}^0 \), set \( j = 0 \).
2. Choose a non-classified data point \( k \): choose \( k \) such that \( k \notin \bigcup_{i=1}^{s} \chi_i \); if all data points have been classified, exit the algorithm.
3. Compute one-step ahead prediction errors for all modes:

\[ J_i^k = (y(k+1) - (y(k) + T_i f_i(y(k), \hat{\theta}))^2, \quad i=1,...,s \]

4. Classify the data point to the mode that gives the smallest prediction error:

\[ \chi_i^{i+1} = \chi_i^i \cup \{k\} \quad \text{where} \quad i = \arg \min_i (J_i^k) \]

and \( \chi_m^{i+1} = \chi_m^i \) for \( m = 1,...,s \) and \( m \neq i \).

5. Estimate \( \hat{\theta}^{i+1} \), using (6) and \( \{\chi_i^{i+1}\}_{i=1}^s \).

6. Set \( j = j + 1 \); go to step 2.

Estimate \( \hat{\theta} \) is given by the value of \( \hat{\theta}^\star \) in the last step and the classified data points are in the sets \( \{\chi_i^\star\}_{i=1}^s \).

Using this algorithm, the model is optimized for one-step ahead prediction. The model predicts \( \hat{x}(k+1) \) from the measured value \( y(k) \). In contrast, in simulation \( \hat{x}(k+1) \) is predicted from the previously predicted \( \hat{x}(k) \). Hence, in simulation, errors between model and data will accumulate. To validate the estimated model parameters, a free run simulation of the model is performed in which the moments of switching are obtained from the classification procedure.

4. EXPERIMENTAL EXAMPLE FROM PHYSIOLOGY

To demonstrate the proposed parameter estimation procedure we applied it to a typical biological process that can be approximated as a hybrid system: ion currents in excitable cells. Examples of excitable cells are neurons and muscle cells (myocytes). Information is propagated along these cells as electrical signals through fast opening and closing of receptors and ion channels in the cell membrane, yielding a system with a switching, hybrid characteristic. In case of myocytes, the electrical excitation is followed by mechanical contraction through the intracellular release of calcium.

4.1 Physiological background

Failure of the regulation of intracellular free calcium \([\text{Ca}^{2+}]_i\) in the cardiac muscle cell has been shown to play an important role in deterioration of heart performance during e.g. diabetes related cardiomyopathy (Choi et al., 2002; Op den Buijs et al., 2005) and cardiac ischemia/reperfusion induced damage (Ivanics et al., 2001). A properly controlled cyclic alteration of \([\text{Ca}^{2+}]_i\), is a prerequisite for appropriate heart performance (Bers, 2002). During excitation-contraction coupling, calcium is released from the sarcoplasmic reticulum (SR) into the cytoplasm by the opening of ryanodine receptors (RyR). Opening of RyR is controlled by a small \([\text{Ca}^{2+}]_i\) current through the sarcolemmal L-Type channel, a mechanism known as \([\text{Ca}^{2+}]_i\)-induced \([\text{Ca}^{2+}]_i\)-release (CICR). The increase in \([\text{Ca}^{2+}]_i\), activates the contractile machinery, whereas sequestration of \([\text{Ca}^{2+}]_i\) back into the SR by the SR \([\text{Ca}^{2+}]_i\) ATP-ase (SERCA) induces a decline in \([\text{Ca}^{2+}]_i\) and enables relaxation of the contractile machinery.

Nowadays, high temporal resolution fluorescence measurement techniques are available to assess the \([\text{Ca}^{2+}]_i\) concentration in isolated beating hearts (e.g. Ivanics et al., 2001). Two problems arise in the quantitative analysis of \([\text{Ca}^{2+}]_i\) dynamics based on intact heart fluorescence measurements. First, the observed macroscopic signal is a lumped sum of release of \([\text{Ca}^{2+}]_i\) from the SR, reaction with various \([\text{Ca}^{2+}]_i\)-binding proteins and the indicator dye itself, and sequestration into the SR. This calls for a method to compute the kinetics of these underlying processes from the \([\text{Ca}^{2+}]_i\)-transient alone. Secondly, initiation and cessation of release of \([\text{Ca}^{2+}]_i\) are almost instantaneous. These switching moments cannot be easily determined from the noisy data.

4.2 Model components

A previously published simple, physiologically meaningful model of calcium handling was adopted (Op den Buijs et al., 2005). The main components of the mathematical model are: the release of \([\text{Ca}^{2+}]_i\) from the SR into the cytoplasm through RyR, sequestration of \([\text{Ca}^{2+}]_i\) back into the SR by SERCA, a small flux as a result of leakage of the SR and binding of \([\text{Ca}^{2+}]_i\) to \([\text{Ca}^{2+}]_i\) buffers in the SR and cytoplasm and the indicator dye. To determine \([\text{Ca}^{2+}]_i\), the following differential equations were solved for inactive \( f_{\text{act}}([\text{Ca}^{2+}]_i) \) and active \( f_{\text{act}}([\text{Ca}^{2+}]_i) \) \([\text{Ca}^{2+}]_i\) release:

\[
\frac{d}{dt}[\text{Ca}^{2+}]_i = \begin{cases} f_{\text{act}}([\text{Ca}^{2+}]_i), & \text{Ca}^{2+} \text{ release active} \quad (9) \\ f_{\text{act}}([\text{Ca}^{2+}]_i), & \text{Ca}^{2+} \text{ release inactive} \end{cases}
\]

\[
f_{\text{act}}([\text{Ca}^{2+}]_i) = \beta_{\text{cyt}} \cdot (J_{\text{pump}} + J_{\text{leak}}) \tag{10}
\]

\[
f_{\text{act}}([\text{Ca}^{2+}]_i) = \beta_{\text{cyt}} \cdot (J_{\text{pump}} + J_{\text{leak}} + J_{\text{rel}}) \tag{11}
\]

Descriptions for the different fluxes \( J_{\text{rel}}, J_{\text{pump}}, J_{\text{leak}} \) and the cytoplasmic buffer fraction \( \beta_{\text{cyt}} \) are outlined below, together with comments on choices for the parameter values that could be determined beforehand.

Switching behavior is introduced by \([\text{Ca}^{2+}]_i\) release from the SR \( J_{\text{rel}} \). During action-potential triggered CICR, the change in total intracellular \([\text{Ca}^{2+}]_i\) is considered to be dependent on the concentration gradient over the SR membrane:

\[
J_{\text{rel}} = k_{\text{rel}} \cdot ([\text{Ca}^{2+}]_i)_{\text{SR}} - [\text{Ca}^{2+}]_i \] \tag{12}

where \( k_{\text{rel}} \) is estimated by the algorithm. For the remaining time of the heart cycle \( J_{\text{rel}} \) equals zero. It is assumed that the total intracellular calcium is conserved and is only transferred between the
cytoplasm and the SR. Therefore, \([\text{Ca}^{2+}]_{\text{SR}}\) can be calculated from \([\text{Ca}^{2+}]_{\text{i}}\), according to:

\[
\frac{d}{dt}[\text{Ca}^{2+}]_{\text{SR}} = -\frac{V_{\text{cyt}}\beta_{\text{cyt}}}{V_{\text{SR}}\beta_{\text{SR}}} \frac{d}{dt}[\text{Ca}^{2+}]_{\text{i}}
\]  

A volume conversion factor \(V_{\text{cyt}}/V_{\text{SR}}\) of 18 L cytosol/L SR (Shannon et al., 2000) was used to determine \([\text{Ca}^{2+}]_{\text{SR}}\) from the total \([\text{Ca}^{2+}]_{\text{i}}\) flux out of the cytoplasm. The equations for the buffer ratio \(\beta_{\text{cyt}}/\beta_{\text{SR}}\) are given below. \([\text{Ca}^{2+}]_{\text{SR}}\) was obtained through integration of (13), treating the initial condition as unknown. This constant was estimated by the algorithm, whereas the physiologically more relevant maximal \([\text{Ca}^{2+}]_{\text{SR}}\) ([\text{Ca}^{2+}]_{\text{SR,max}}) is reported here.

Leakage of \([\text{Ca}^{2+}]\) from the SR \((J_{\text{leak}})\) was modeled as:

\[
J_{\text{leak}} = k_{\text{leak}} ([\text{Ca}^{2+}]_{\text{i}} - [\text{Ca}^{2+}]_{\text{cyt}})
\]  

with leak rate \(k_{\text{leak}}\) fixed at 0.00047 s\(^{-1}\) to maintain a gradient over the SR membrane in steady-state (Shannon et al., 2000).

\([\text{Ca}^{2+}]\) is continuously being sequestered back into the SR by SERCA. The turnover rate of SERCA \((J_{\text{pump}})\) was modeled by Hill kinetics:

\[
J_{\text{pump}} = \frac{v_{\text{max}}}{1 + (k_m/[\text{Ca}^{2+}])^h}
\]  

with \(v_{\text{max}}\) the maximal uptake rate of SERCA, \(k_m\) the half-saturation concentration and \(h\) the Hill coefficient. Parameters \(v_{\text{max}}\) and \(k_m\) were estimated by the algorithm, whereas \(h\), an index for the stoichiometry of SERCA, was fixed at 3.98 (Smith et al., 1998). The sarcolemmal Na\(^+\)/Ca\(^{2+}\) exchanger was omitted since its contribution is less than 10% of the total Ca\(^{2+}\) flux in rat cardiac myocytes (Bers, 2002).

Buffering of \([\text{Ca}^{2+}]\) by calmodulin, troponin C, several \([\text{Ca}^{2+}]\) binding proteins in the sarcolemma and SR membrane and the indicator dye Indo-1 was determined from 8 individual buffer reactions using the rapid buffer approximation (Wagner and Keizer, 1994):

\[
\beta_{\text{cyt}}^{-1} = 1 + \sum_{i=1}^{4} B_i \frac{k_{i,j}}{([\text{Ca}^{2+}]_{\text{i}} + k_{i,j})^2}
\]  

The values for the buffer concentrations \(B_i\) and dissociation constants \(k_{i,j}\) were fixed at values adopted from Shannon et al. (2000).

Calsequestrin proteins are the main \([\text{Ca}^{2+}]\) buffer in the SR:

\[
\beta_{\text{SR}}^{-1} = 1 + B_{\text{SR}} \frac{k_{\text{SR}}}{([\text{Ca}^{2+}]_{\text{i}} + k_{\text{SR}})^2}
\]  

The SR buffer concentration \(B_{\text{SR}}\) was fixed at 2.7 mM SR and dissociation constant \(k_{\text{SR}}\) at 0.630 mM SR (Shannon et al., 2000).

4.3 Methods and implementation

\([\text{Ca}^{2+}]\), in the isolated, beating heart was measured according to the method published by Ivancics et al. (2001). The signals were recorded with a sampling time \(T_s = 5\) ms (200 Hz) and low-pass filtered using a 5th-order Butterworth filter with cutoff frequency of 20 Hz. \([\text{Ca}^{2+}]\), was measured during control conditions and upon challenging the hearts with 0.5 nM isoproterenol, which acts as a \(\beta\)-adrenergic stimulator. \(N = 200\) samples were collected in each experiment.

To initialize Algorithm 3.1, initial sets \(\chi_{\text{on}}^0\) and \(\chi_{\text{off}}^0\) were obtained using the available \(a\ priori\) knowledge about the system. Data points where \([\text{Ca}^{2+}]\), increased compared to the previous sampling instant could be, with a degree of certainty, classified to \(\chi_{\text{on}}^0\) and conversely, data points where \([\text{Ca}^{2+}]\), decreased were classified to \(\chi_{\text{off}}^0\).

The model and algorithm were implemented in Matlab 6.5 R13 (The MathWorks, Inc., Natick, MA, USA), running on an IBM compatible PC (2.4 GHz) under MS Windows XP Pro. The nonlinear least squares problem was numerically solved using the Levenberg-Marquardt algorithm of the Optimization Toolbox (version 2.2, Matlab). The termination tolerance on the parameter values \(\theta\) was set to 1e-6 and the termination tolerance on the cost function \(J\) to 1e-6. Besides the estimated parameter values \(\hat{\theta}\), the algorithm returned \(J(\hat{\theta})\), the value of the squared 2-norm of the residuals at the solution and the Jacobian \(\partial J/\partial \hat{\theta}\), which were used to calculate the covariance matrix according to (8).

4.4 Results

Classification and identification The fluorescence measurements of \([\text{Ca}^{2+}]\), were clustered in sets of active \((\chi_{\text{on}})\) and inactive \((\chi_{\text{off}})\) \([\text{Ca}^{2+}]\) release using Algorithm 3.1 (Fig. 1). Furthermore, the one-step prediction and classification of the data in modes of \([\text{Ca}^{2+}]\) release ‘on’ and ‘off’ under control conditions.
ahead prediction of the data is shown using the final estimate $\hat{\theta}$.

Table 1 summarizes the values of the model parameters as identified by our method. Previously published values for $v_{\text{max}}$ of SERCA in unchallenged cardiomyocytes are $137 \pm 16 \, \mu M \, s^{-1}$ (Shannon et al., 2000), $\sim 144 \, \mu mol \, kg^{-1} \, s^{-1}$ (Choi et al., 2002), $210 \pm 4 \, \mu M \, s^{-1}$ (Balke et al., 1994) and $5 \, mM \, s^{-1}$ (Luo and Rudy, 1994). Corresponding values for $k_{m}$ are $260 \, nM$ (Shannon et al., 2000), $\sim 150 \, nM$ (Choi et al., 2002), $280 \pm 4 \, nM$ (Balke et al., 1994) and $920 \, nM$ (Luo and Rudy, 1994). Upon stimulation with 0.5 nM isoproterenol, we found a small increase in $v_{\text{max}}$, whereas the main effect on SERCA is a decrease of $k_{m}$ by nearly 20%, in line with biochemical findings (Odermatt et al., 1996). Because of the acceleration of SERCA, SR Ca$^{2+}$ load is elevated by 17% under 0.5 nM isoproterenol, similar to results reported before in ferret myocytes (Ginsburg et al., 1998). The model predicts that the applied dose of isoproterenol does not significantly change the channel conductance $k_{\text{rel}}$ of RyR. A previous reported value for $k_{\text{rel}}$ under unchallenged conditions is $1.8 \, s^{-1}$ (Jafri et al., 1998).

**Simulation** To validate the estimated model parameters, a free run simulation of the model was performed. The moments of switching were obtained from the classification procedure. Fig. 3 compares the measured and simulated [Ca$^{2+}$], and shows the simulated [Ca$^{2+}$]$_{\text{SR}}$. As could be expected, the fit was less good than with the one-step ahead prediction scheme, but remained acceptable.

### 5. DISCUSSION

A novel technique for the simultaneous estimation of model parameters and classification of data in hybrid systems was presented. It is aimed towards parameter estimation in switched, nonlinear systems and provides an approach to handle a priori knowledge in the identification of hybrid systems. The principle is somewhat similar to the recently developed Bayesian approach to identification of PWARX hybrid models (Juloski et al., 2004b), in the sense that a priori information is used that is iteratively refined as more data points are considered.

The method was applied to identify cardiac [Ca$^{2+}$]-cycling under physiologically relevant conditions in the intact beating rat heart. Analysis of control and isoproterenol-stimulated hearts resulted in parameter values within a physiologically reliable range. The method is more robust against measurement noise compared to when the data was classified a priori (i.e. separately from the estimation procedure), (Op den Buijs et al., 2005; Van Riel and Van der Vusse 2002). The one-step ahead method (5) requires that all model states are measured outputs (2). Although this was not the case for the calcium system, the procedure could be applied because the unobserved [Ca$^{2+}$]$_{\text{SR}}$ could be directly reconstructed from the measured [Ca$^{2+}$], (13). It was assumed that total intracellular calcium is conserved and is only transferred between the two model states. This particular mass conservation property does not generally hold for all (physiological) hybrid systems. Furthermore, the [Ca$^{2+}$] data were equidistantly sampled with a sufficiently high rate such that a single numerical Euler integration step was sufficient to predict the (discretized) state value at the next time instant. Especially in biological systems, these sampling

![Fig. 2. Validation (one-step ahead prediction). Algorithm 3.1 was only used to classify the data into the two modes.](image-url)

![Fig. 3. Simulation of [Ca$^{2+}$], and [Ca$^{2+}$]$_{\text{SR}}$ under control conditions.](image-url)
conditions can usually not be fulfilled. In the future, the proposed new identification procedure can be adapted to non-equidistantly sampled and undersampled systems.

ACKNOWLEDGMENT

The authors would like to thank Dr. Ivanics and Prof. Ligeti (Institute of Human Physiology and Clinical Experimental Research, Semmelweis University, Budapest, Hungary) for kindly providing the experimental data.

REFERENCES


Odermatt, A., K. Kurzydlowski and D.H. MacLennan (1996) The Vmax of the Ca2+-ATPase of cardiac sarcoplasmic reticulum (SERCA2a) is not altered by Ca2+/calmodulin-dependent phosphorylation or by interaction with phospholamban. J. Biol. Chem., 271, 14206—14213.


