Learning Cycle-Linear Hybrid Automata for Excitable Cells

R. Grosu[1], S. Mitra[2], P. Ye[1], E. Entcheva[3],
I.V. Ramakrishnan[1], and S.A. Smolka[1]

[1] Department of Computer Science, Stony Brook University
[2] Computer Science and AI Lab, Massachusetts Institute of Technology
[3] Department of Biomedical Engineering, Stony Brook University

Abstract. We show how to automatically learn the class of Hybrid Automata called Cycle-Linear Hybrid Automata (CLHA) in order to model the behavior of excitable cells. Such cells, whose main purpose is to amplify and propagate an electrical signal known as the action potential (AP), serve as the "biologic transistors" of living organisms. The learning algorithm we propose comprises the following three phases: (1) Geometric analysis of the APs in the training set is used to identify, for each AP, the modes and switching logic of the corresponding Linear Hybrid Automata. (2) For each mode, the modified Prony's method is used to learn the coefficients of the associated linear flows. (3) The modified Prony's method is used again to learn the functions that adjust, on a per-cycle basis, the mode dynamics and switching logic of the Linear Hybrid Automata obtained in the first two phases. Our results show that the learned CLHA is able to successfully capture AP morphology and other important excitable-cell properties, such as refractoriness and restitution, up to a prescribed approximation error. Our approach is fully implemented in MATLAB and, to the best of our knowledge, provides the most accurate approximation model for ECs to date.

1 Introduction

Hybrid automata [2,19] are an increasingly popular modeling formalism for systems that exhibit both continuous and discrete behavior. Intuitively, Hybrid automata (HA) are extended finite-state automata whose discrete states correspond to the various modes of continuous dynamics a system may exhibit, and whose transitions express the switching logic between these modes.

Traditionally, HA have been used to model embedded systems, including automated highway systems, air traffic management, embedded automotive controllers, robotics, and real-time circuits. More recently, they have been used to model and analyze biological systems, such as cellular cycles and immune response [3], bio-molecular networks [1], gene-regulatory networks [7,17,23], protein-signaling pathways[11], and metabolic processes [4]. The hybrid-system metaphor has also been used to develop algorithms for large-scale simulation of

A. Bemporad, A. Bicchi, and G. Buttazzo (Eds.): HSCC 2007, LNCS 4416, pp. 245–258, 2007.

biological systems [15]. Biological systems are intrinsically hybrid in nature: biochemical concentrations may vary continuously, yet discrete transitions between distinct states are also possible.

Excitable cells (ECs) are a typical example of biological systems exhibiting hybrid behavior: transmembrane ion fluxes and voltages may vary continuously but the transition from the resting state to the excited state is generally considered an all-or-nothing discrete response. Traditionally, however, the preferred modeling approach for ECs uses large sets of coupled nonlinear differential equations. Although an invaluable asset for integrating genomics and proteomics data to reveal local interactions, such models are not typically amenable to control-theoretic techniques developed for linear systems, and render large-scale simulation impractical.

In previous work [26,27] we showed that it is possible to construct a conceptually simpler HA model for ECs that approximates with reasonable accuracy their electrical properties. We called these HA *Cycle-Linear HA* (CLHA) to highlight their cyclic structure and the fact that, while in each cycle they exhibit linear dynamics, the coefficients of the corresponding linear equations and mode-transition guards may vary in interesting ways from cycle to cycle.

The manual construction of CLHA, however, proved to be a tedious task, and the CLHA so derived were tied to a particular type of EC and to a particular species. Moreover, since recent advances in measuring *in vitro* the electrical activity of ECs have resulted in the availability of extensive data sets, it was natural to turn our attention to the following question: *Given a training set of electrical measurements of an EC, is it possible to automatically learn a CLHA that approximates the behavior of the EC up to a required error margin?*

In this paper we address this question, by presenting a methodology for automatically learning CLHA models for two types of ECs: the squid giant axon and the guinea pig ventricular cell. To the best of our knowledge, these are the most accurate approximation models (with per-AP-cycle linear dynamics), developed for these ECs to date, both in terms of electrical-signal morphology and typical excitable-cell characteristics such as refractoriness and restitution.

To simplify the process of obtaining training sets, we used *virtual* measurements obtained by applying the so-called *S1S2*-protocol to existing nonlinear models of these ECs. Extending the method outlined here to *in vitro* data obtained in the laboratory of the fourth author is a direction for future work.

The learning technique we have developed for CLHA is also of independent interest, as we learn all aspects of excitable-cell CLHA models up to a given error margin, including the number of modes; for each mode, the dimension of the state space and the coefficients of its linear time-invariant dynamics; and all aspects of the mode switching logic, including the jump conditions, thresholds and resets. To do so, we use the modified Prony's method to obtain an exponential fit for the continuous per-mode linear dynamics. Moreover, in learning the CLHA, we make no *a priori* assumptions about the dimension of the state space of the nonlinear system we are targetting, nor the degree of its input and output.

We also learn the functions that adjust a CLHA's mode dynamics and switching logic on a per-cycle basis. This aspect of our technique is critical in the case of excitable cells, which exhibit the following *restitution property*: the longer the recovery time for an EC, the longer in duration its subsequent *action potential*.

Organization. The rest of the paper is organized as follows. Section 2 defines Cycle Linear Hybrid Automata (CLHA). Section 3 describes the requisite biology of excitable cells. Our proposed methodology for learning a CLHA from a given training set is presented in Section 4. In this section, we also present our results for the training/testing set generated from the highly nonlinear Luo-Rudi dynamic model previously developed for Guinea Pig ventricular cells. Section 5 discusses related work. Section 6 contains our concluding remarks and directions for future research.

2 Hybrid Automata

We define *Cycle-Linear Hybrid Automata* as a restricted class of *Structured Hybrid Automata* [20]. SHA, which are derived from *Timed Input/Output Automata* [16], can model a general class of hybrid systems for which the input/output distinction intrinsic to the IOA methodology is ignored. We use a variable structure to specify states of an automaton. Let X be a set of variables. A *valuation* $\mathbf{x}$ for X is a function that associates with each $x \in X$ a value in its type. The set of all valuations of X is denoted by $val(X)$. For $\mathbf{x} \in val(X)$, let $\mathbf{x}.x$ denote the value of the variable x within the state $\mathbf{x}$. A *trajectory* $\tau : J \rightarrow val(X)$ specifies the values of all variables in X over a real time interval J. The *limit time* of a trajectory τ, written as $\tau.ltime$, is the supremum of the domain of τ. A *state model* for X is a collection F of differential and algebraic equations involving the continuous variables in X such that for every $\mathbf{x} \in val(X)$, there exists solution trajectory τ of F that starts from $\mathbf{x}$.

Definition 1. *A* Structured Hybrid Automaton *(SHA) with mode set $\mathcal{M}$ is a tuple $\mathcal{A} = (X, Q, \Theta, A, \mathcal{D}, P)$, where X is a set of* variables*, including a special discrete variable called* mode *of type $\mathcal{M}$; $Q \subseteq val(X)$ is the set of states; $\Theta \subseteq Q$ is a nonempty set of* start states*; A is a set of* actions, *$\mathcal{D} \subseteq Q \times A \times Q$ is a set of* discrete transitions*; and P is an indexed family F_i, $i \in \mathcal{M}$, of state models.*

As usual, we will specify the set of transitions in $\mathcal{D}$ corresponding to an action $a \in A$ by a guard predicate E_a and a reset map $R_a : X \rightarrow X$. A transition $(\mathbf{x}, a, \mathbf{x}') \in \mathcal{D}$ is called a *mode switch* if $\mathbf{x}.mode \neq \mathbf{x}'.mode$. The set $\mathcal{T}_{\mathcal{A}}$ of trajectories of an SHA $\mathcal{A}$ is defined as follows: a trajectory τ for X is in $\mathcal{T}_{\mathcal{A}}$ if the restriction of τ to the set of continuous variables in X satisfies the state model $F_{\tau(0).mode}$, and the restriction of τ to the discrete variables in X remain constant over the domain of τ. An *execution fragment* captures a particular run of $\mathcal{A}$; it is defined as an alternating sequence of actions and trajectories $\alpha = \tau_0 a_1 \tau_1 a_2 \ldots$, where each $\tau_i \in \mathcal{T}$, and if τ_i is not the last trajectory then $(\tau_i(ltime), a_{i+1}, \tau_{i+1}(0)) \in \mathcal{D}$.

The SHA model represents one end of the spectrum of hybrid-system models where one can specify systems with very general dynamics, and discrete mechanics. Linear and rectangular hybrid automata [2], at the other end of the spectrum, enable powerful analysis techniques by restricting the state models to linear differential equations. The Cycle-Linear Hybrid Automaton (CLHA) model was proposed in [25] for describing and analyzing highly nonlinear but periodic systems, and provides the mathematical basis for the remainder of this paper. Definition 2 gives a precise semantic definition of this model as restricted SHA. Informally, a CLHA captures a class of hybrid system where the state models, reset maps, and guards are all linear but with coefficients that are functions of a discrete state variable called *epoch*. The *epoch* variable is reset only when a particular mode is entered.

Definition 2. *A* Cycle-Linear Hybrid Automaton *(CLHA) with state space $Q \subseteq val(X)$, mode set $\mathcal{E} \times \mathcal{P}$, and snapshot map $S : Q \rightarrow \mathcal{E}$, is an SHA with mode set $\mathcal{M}$ satisfying:*

1. *Variable mode has type $\mathcal{M} = \mathcal{E} \times \mathcal{P}$, and its first component is a discrete variable of type $\mathcal{E}$ referred to as* epoch. *There is a unique $\zeta \in \mathcal{P}$ that is visited infinitely many times in any execution with an infinite number of mode switches.*
2. *For each $(\epsilon, p) \in \mathcal{M}$, $F_{\epsilon,p}$ is a linear state model. For each action $a \in A$, the guard E_a (reset map R_a) can be expressed as a linear predicate (resp. function) on X, with coefficients that are functions of epoch.*
3. *Suppose $(\mathbf{x}, a, \mathbf{x}')$ is a mode switch with $\mathbf{x}.mode = (\epsilon_1, p_1)$ and $\mathbf{x}'.mode = (\epsilon_2, p_2)$, for some $\epsilon_1, \epsilon_2 \in \mathcal{E}$, $p_1, p_2 \in \mathcal{P}$. If $p_2 = \zeta$ then $\epsilon_2 = S(\mathbf{x})$; otherwise, $\epsilon_2 = \epsilon_1$. A mode switch of the first type is called is called an* epoch transition.

3 Excitable Cells (ECs)

Action potentials. Excitable cells (ECs) can be viewed as active electrical circuits with nonlinear behavior, capable of amplifying and propagating electrical signals. ECs include neurons, cardiac cells, skeletal and smooth muscle cells. In cardiac cells for example, on each heart beat, an electrical control signal is generated by the sinoatrial node, the heart's internal pacemaking region. This signal is amplified and propagated as an electrical wave along a prescribed path, exciting cells in the main chambers of the heart (atria and ventricles) and assuring synchronous contractions. At the cellular level, the electrical signal is a change in the potential across the cell membrane caused by different ion currents flowing through the cell membrane. This electrical signal for each excitation event is known as an *action potential* (AP). A typical AP for Guinea-Pig ventricular cells is shown in Figure 1(a). Its morphology is usually defined as a sequence of six phases: *stimulated* (S), *upstroke* (U), *early repolarization* (E), *plateau* (P), *final repolarization* (F), and *resting* (R).

For non-pacemaking ECs, APs are externally triggered events: a cell fires an AP as an all-or-nothing response to a supra-threshold stimulus, and each

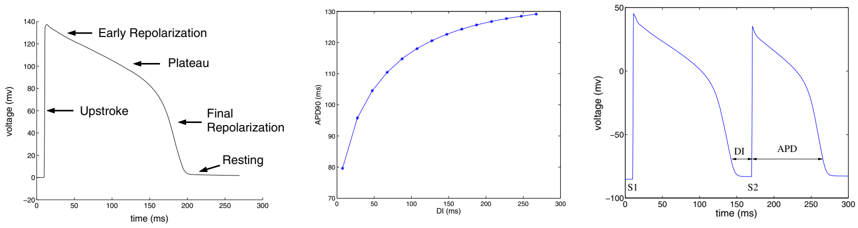

Fig. 1. (a) AP and its phases. (b) Restitution curve. (c) APD, DI and S1S2 protocol.

AP follows the same sequence of phases and maintains approximately the same magnitude regardless of the applied stimulus. After an initial step-like increase in the membrane potential, an AP lasts for a couple milliseconds to hundreds milliseconds in most mammals. During phases U, E, P and the first part of F, generally no re-excitation can occur. This portion of an AP is therefore known as the *absolute refractory period* (ARP). Starting with the second part of phase F, an altered secondary excitation event is possible if the stimulation strength or duration is raised. This portion of the AP is therefore known as the *relative refractory period* (RRP).

Restitution function. When an EC is subjected to repeated stimuli, two important time intervals can be identified: the *action potential duration* (APD), the period when the AP is above some prescribed percentage (e.g. 10%) of its maximum height, and the *diastolic interval* (DI), the period from the end of the APD to the end of the cycle, i.e. the end of phase R. Figure 1(c) illustrates the two intervals.

The function relating APD to DI as the cell is subjected to different stimulation frequencies is called the APD *restitution function.* As shown in Figure 1(b), the function is nonlinear and captures the phenomenon that a longer recovery time is followed by a longer APD. A physiological explanation of a cell's restitution is rooted in the ion-channel kinetics as a limiting factor in the cell's response to multiple stimuli over time. The sum of the APD and DI is called the *Basic Cycle Length* (BCL).

The *S1S2 protocol* is often used to determine the restitution function of an excitable cell. In this protocol, a cell is driven into a stable mode, in which a stable APD may be observed, by first subjecting it to a train of so-called S1 stimuli at a fixed BCL. Immediately thereafter, a single S2 stimulus, having a different (i.e. shorter) BCL is delivered. As such, one can associate a DI-APD pair with each running of the protocol, viz. the DI preceding the S2-induced APD. By repeating this procedure and varying the DIs before S2, one gradually constructs the graph of the restitution curve. Figure 1(c) illustrates the placement of the last S1 stimulus followed by the S2 stimulus.

Mathematical models of excitation. Modeling of the ionic processes that underlie cell excitation dates back to 1952, when Hodgkin and Huxley formulated their model of the squid giant axon [13]. Intuitively, the HH model is that of a nonlinear resistor-capacitor (RC) circuit with current sources, defining AP in terms of a stimulation current and three ionic currents: (fast) inward sodium, (slow) outward potassium, and a time-independent linear (leak) current. The ionic currents depend themselves on the AP via a gating mechanism (a time-varying conductance). The corresponding nonlinear system of equations is given below, where: V, m, n and h are continuous state variables; V is the AP, m, n and h are the ion-channel gates; $\overline{g}_{\mathrm{Na}}$,$\overline{g}_{\mathrm{K}}$,$\overline{g}_{\mathrm{L}}$ are the constants which represent the maximum channel conductances for the sodium, potassium and leakage channel, respectively; E_{Na},E_{K},E_{L} are the constants for reversal potentials for these channels; m_{∞}, h_{∞} and n_{∞} are the ion-channel gates' steady-state values, and τ_m, τ_h and τ_n are their time-constant values; C is the constant cell capacitance and I_{st} is the stimulation current.

$$C\dot{V} = -\overline{g}_{\mathrm{Na}} m^3 h(V - E_{\mathrm{Na}}) - \overline{g}_{\mathrm{K}} n^4 (V - E_{\mathrm{K}}) - \overline{g}_{\mathrm{L}}(V - E_{\mathrm{L}}) + I_{\mathrm{st}}$$

$$\begin{array}{lll}
\tau_m\ \dot{m} = m - m_{\infty} & \tau_m\ = 1/(\alpha_m + \beta_m) & m_{\infty}\ = \alpha_m/(\alpha_m + \beta_m) \\
\tau_h\ \dot{h}\ = h - h_{\infty} & \tau_h\ = 1/(\alpha_h + \beta_h) & h_{\infty}\ = \alpha_h/(\alpha_h + \beta_h) \\
\tau_n\ \dot{n}\ = n - n_{\infty} & \tau_n\ = 1/(\alpha_n + \beta_n) & n_{\infty}\ = \alpha_n/(\alpha_n + \beta_n)
\end{array}$$

$$\begin{array}{lll}
\alpha_m(V) = \frac{2.5-0.1V}{e^{2.5-0.1V}-1} & \alpha_h(V) = 0.07e^{-\frac{V}{20}} & \alpha_n(V) = \frac{0.1-0.01V}{e^{1-0.1V}-1} \\
\beta_m(V) = 4e^{-\frac{V}{18}} & \beta_h(V) = \frac{1}{e^{3-0.1V}+1} & \beta_n(V) = 0.125e^{-\frac{V}{80}}
\end{array}$$

The HH model with its 3 membrane currents, 4 state variables, and 12 fitted parameters laid the foundation for subsequent models of excitable cells of increasing complexity. All of these models use multiple continuous state variables (voltage, ion-channel gates, ion concentrations) to describe action potential in different cell types. One of the most popular cardiac-cell models is the dynamic Luo-Rudy model [18]. The LRd model uses 11 different membrane currents, more than 20 state variables and over 150 fitted parameters to describe the AP. Due to space constraints, the full structure of the LRd model is not listed here.

4 Learning the CLHA of an Excitable Cell

Given a training set of APs generated by applying the *S1S2*-protocol to an excitable cell of a particular species, our methodology for learning the CLHA that approximates the cell's behavior up to a given error margin consists of two phases. In the first phase, we obtain for each AP a linear Hybrid automaton (LHA) whose output is within the specified error bound. This involves identifying the segments of the APs that correspond to the modes of the LHA, deriving the flows for each mode, and the guards and reset maps for each transition. In each mode, we use the modified Prony's method (MPM) [21] to approximate the AP with a (normalized) linear dynamics, i.e., with a sum of exponentials.

In the second phase, we derive a CLHA that combines the behavior of all the LHAs and therefore captures all the APs in the training set. We exploit the fact that the coefficients defining the flows, guards, and reset maps of the CLHA are functions of the *epoch* variable which is updated during an *epoch transition*. We

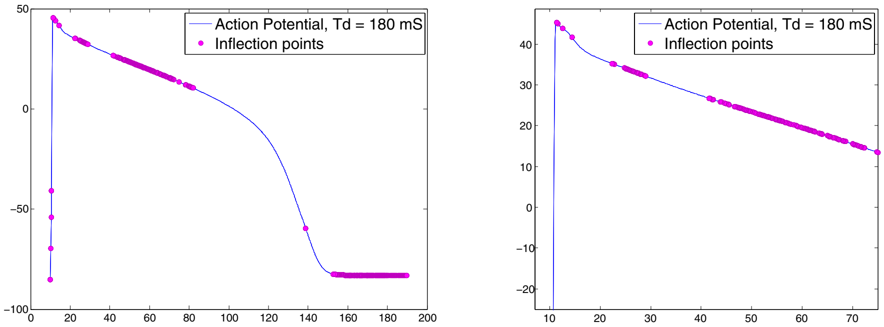

Fig. 2. Null/inflection points in the LRd APs

choose the variable to be a voltage-valued variable called v_0 and epoch transitions to be those that are brought about by the occurrence of a stimulus. In finding the snapshot map which sets the value of the epoch variable in the post-state of epoch transitions, we once again use MPM. Specifically, we estimate the voltage-dependent coefficients of the CLHA as an exponential regression of the constant coefficients in the LHAs obtained in the first phase.

Assumptions. Our goal is to derive a CLHA, the output of which is within ± 2 mv of the output of the Luo-Rudy model, under the following class of stimuli: each stimulation is a step of amplitude -80 $\mu\mathtt{A}/\mathtt{cm}^2$, duration 0.6 msecs, and BCL between 160 and 400 msecs. The set of 25 APs sampled every 0.2 msecs, corresponding to BCL 160 to 400 msecs, in 10 msecs intervals, serves as the *training set* for deriving the CLHA. The performance of the learned CLHA is evaluated on the *test set* consisting of APs with BCL from 165 to 405 msecs, in 10 msecs intervals, sampled at the same frequency.

Identifying Modes. To discover the points in the APs that correspond to mode transitions in the target LHAs, we computed the null points (zeros of the first-order derivative) and the inflection points (zeros of the second-order derivatives) of the APs. This approach worked very well for the HH model, and the sections of APs between successive null or inflection points were identified as the modes of the LHAs.

When directly applied to the LRd model, this approach yielded far too many modes. In particular, there exist trains of inflection points in the P and R phases of the APs (see Figure 2). This was somewhat surprising because the AP of these phases appears as rather smooth line segments corresponding to "stretched" inflection points. The higher-order nonlinearity of the LRd model seems to have dealt with such segments by generating trains of points whose tangent (first-order derivative) difference was smaller (in absolute value) than 10^{-5}. Based on this observation, we designed our own parameterized filter to eliminate such long sequences of closely-spaced inflexion points. The filter parameter enable us to increase or decrease the number of segments and thereby achieve the desired accuracy of the CLHA.

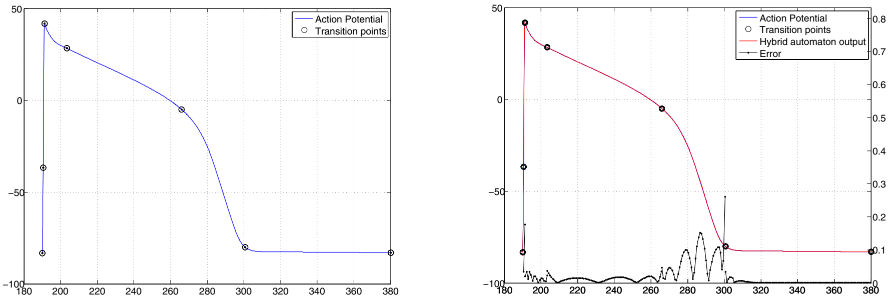

Fig. 3. (a) Inflection points after filtering. (b) Hybrid-automaton output.

Using the MPM described below, we were able to approximate each segment with two exponentials and the entire AP to within the desired accuracy. Since, however, this approach seemed to split each of the E and F phases in two, we decided to eliminate one inflection point in each. In doing so, we were not able to maintain the desired accuracy, unless we moved down the end-point of phase P and up the starting-point of phase R. The correctness of both transformations was confirmed by analyzing higher-resolution APs, where these points were indeed very close to their inferred position. The final seven points chosen are shown in Figure 3(a).

Using Modified Prony's Method to Obtain LHA. The null/inflection points partition the AP into sections defining six modes of the LHA: S, U, E, P, F and R. We denote the set of modes by $\mathcal{P}$. Since these modes are always visited in order, the voltages of the six inflection points define the guards (thresholds) for the corresponding transition edges. We denote the transition voltages by V_p, where $p \in \mathcal{P}$, is the mode in the post-state of the transition. For example, in the AP of Figure 3(a), the transition from U to E occurs at $V_{\mathsf{E}} = 45.32$ mv. To completely define the LHA, it remains to define the flows and the reset maps; for this we use the modified Prony's method [21].

The modified Prony's method is a technique for fitting exponential or sinusoidal functions to time-series data. For fixed n, MPM minimizes the $L2$ distance between time-series data and any function y that solves a differential equation with constant coefficients:

$$\sum_{i=1}^{n+1} c_i \frac{d^{i-1}y}{dt^i} = 0. \tag{1}$$

Depending on the coefficients c_i, the function estimating the solution of Equation 1 may be a complex or a real exponential, damped or undamped sinusoids. Furthermore, the input to the algorithm can be noisy periodic samples from the actual solution. Because of these attractive features, MPM has found many practical applications.

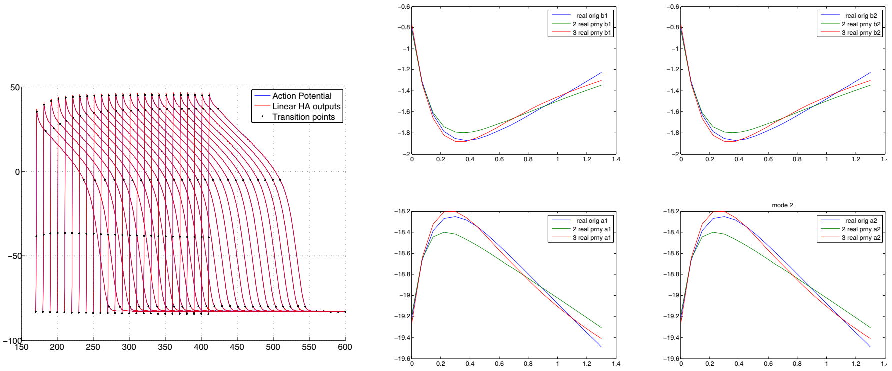

Fig. 4. (a) Original APs and superposed LHA outputs for training set. (b) Sums of 2 and 3 exponentials for estimating a_1, a_2, b_1 and b_2 for mode U.

Suppose the voltage in mode $p \in \mathcal{P}$ of the AP can be approximated as a sum of exponential functions:

$$v(t) = \sum_{i=1}^{n} a_{ip} e^{b_{ip} t} \tag{2}$$

Then, we can specify the flows in each mode as :

$$\forall\ i \in \{1, \ldots, n\}, \quad \dot{x}_i = b_{ip} x_i \text{ and } x_i(0) = a_{ip} \tag{3}$$
$$v = \sum_{i=1}^{n} x_i,$$

where the x_i's are the state variables. The initial condition on the state variables is set by the reset map of the transition from the previous mode. The accuracy of the above approximation is a function of n, that is, the number of state variables used. Using the MPM with $n = 2$, we obtained, approximations that were within the acceptable error bounds for all modes. The output of the resulting LHA, the original AP, and the error between the two, are plotted in Figure 3(b). We apply this procedure to obtain an LHA for each AP in the training set. The output of these automata, superimposed on the original APs, are shown in Figure 4(a).

Linear to Cycle-Linear HA. From the first phase, we obtain for each AP in the training set and for each mode $p \in \mathcal{P}$, the transition voltage V_p for the guards, and the coefficients b_{1p}, b_{2p}, and a_{1p}, a_{2p} corresponding to the the differential equations and initial values for the state variables x_1 and x_2. In other words, we obtain one linear hybrid automaton approximating each of the APs in the training set.

In the second phase, we combine these LHAs into a single CLHA by using the transition to mode S (stimulus arrival) as the epoch transition, setting the value of the epoch variable v_0. We call the value of v_0 the *epoch voltage*. For each mode, we find a function mapping v_0 of each LHA to transition voltages and coefficients; this function implicitly defines the snapshot map. We once again

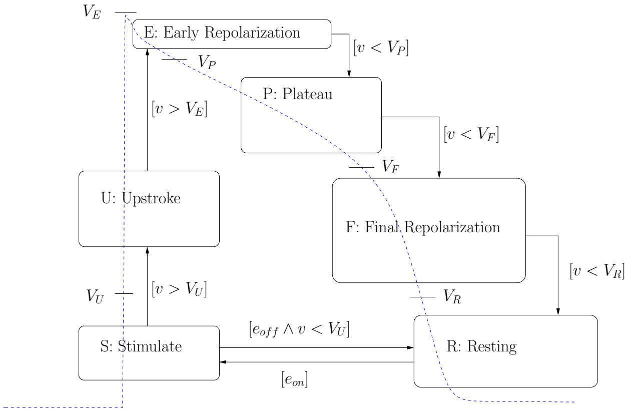

Fig. 5. Structure of the CLHA

use sums of two exponentials for these functions and obtain their coefficients by applying MPM. These functions are defined below, where $p \in \mathcal{P}$ and $i \in [1..2]$:

$$\begin{aligned} V_p(v_0) &= \vartheta_p \ e^{\theta_p v_0} \ + \vartheta'_p \ e^{\theta'_p v_0} \\ a_{ip}(v_0) &= \alpha_{ip} \ e^{\lambda_{ip} v_0} + \alpha'_{ip} \ e^{\lambda'_{ip} v_0} \\ b_{ip}(v_0) &= \beta_{ip} \ e^{\gamma_{ip} v_0} \ + \beta'_{ip} \ e^{\gamma'_{ip} v_0} \end{aligned}$$

Thus, a_{ip}, b_{ip} and V_p in the CLHA depend on the AP value stored in variable v_0 on the epoch transition between modes R and S. The way MPM approximates a_{1U}, a_{2U}, b_{1U} and b_{2U} with sums of two or three exponentials is shown in Figure 4(b). The structure of the CLHA thus obtained is given in Figure 5. For simplicity, the figure does not show the actions on the transitions and the flows within the modes.

While the above equations give the general pattern for the transition voltages and coefficients, a few observations are in place. First, by construction, V_F and V_R are constant in all LHAs and therefore no exponential fitting is necessary for the CLHA. Secondly, the a_i and b_i coefficients of modes F and R are up to a very small variation the same in all LHAs. Although we expressed them as functions in the CLHA, we are confident that using constants instead would have still satisfied the required accuracy. Thirdly, for the rest of the modes, the a_i and b_i obtained for the LHAs are complex values. We therefore separately fitted their real and imaginary parts. The constant coefficients ϑ_p, ϑ'_p, θ_p, θ'_p, α_{ip}, α'_{ip}, λ_{ip}, λ'_{ip}, β_{ip}, β'_{ip}, γ_{ip} and γ'_{ip} are complex too. Finally, due to space restrictions, we defer including a table with all voltage and coefficient values to the full version of the paper. We are happy, however, to provide them upon request.

Simulation results. We have implemented the above-described learning technique in MATLAB, and applied it to both the HH and LRd models. The accuracy of the resulting CLHA was analyzed on both the training and test sets. Due to space constraints, the results on the simpler HH model are omitted.

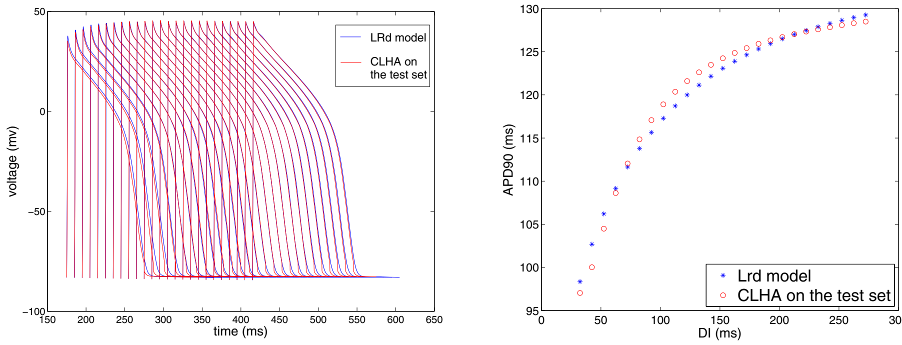

Fig. 6. (a) Comparison of AP. (b) Comparison of restitution curve.

The output of the CLHA on the LRd test set is shown superposed on the original APs in Figure 6(a). As can be observed, the morphology of the output, as well as the required accuracy, is maintained on this set. In Figure 6(b), the restitution curve obtained from the CLHA by running it on the v_0's specified in the test set is compared to the restitution curve obtained from the APs in the test set. Although not perfect, the results are very satisfactory. To our knowledge, these are the best results among the LRd-approximation models proposed so far.

5 Related Work

We have developed a learning/identification technique for cycle-linear hybrid automata (CLHA), and applied it to a classical, highly nonlinear model of ventricular cardiac myocytes. The technique of hybrid-automaton identification has been previously used in a number of communication and control applications, including interplanetary life-support systems [12], dynamic power management [8], autonomous systems and intelligent robots [14,10], and figure tracking [22]. To the best of our knowledge, our application of this technique in the area of systems biology, in general, and excitable cells, in particular, is the first of its kind.

Our approach to hybrid-automaton identification is further distinguished from prior work in the area by the novelty of the identification technique itself. Specific contributions in this regard include the following: (1) Our approach is applicable to continuous-time nonlinear systems that exhibit some level of periodicity and adaptation. Given such a system, the CLHA we learn are also continuous-time, specifically, linear time-invariant (LTI). In contrast, the techniques of [24,5] target discrete-time PWARX (piecewise-affine auto-regressive exogenous) models. Furthermore, in contrast to these approaches, when learning the CLHA for a system S, we make no *a priori* assumptions about the dimension of S's state space nor the degree of its input and output.

(2) Our technique learns all aspects of a hybrid automaton, including the number of modes; for each mode, the dimension of the state space and the coefficients of its LTI dynamics; and all aspects of the mode switching logic, including the jump conditions, thresholds and resets. To do so, we use a modified

Prony method to obtain an exponential fit for the continuous per-mode linear dynamics. Cf. [24], where polynomial fitting is used for the case of discrete-time PWARX systems.

(3) We also learn the functions that adjust a CLHA's mode dynamics and switching logic on a per-cycle basis. This aspect of our technique is critical in the case of excitable cells because of their *restitutional* nature (see Section 3). In this case, the coefficients of the mode dynamics and the voltage thresholds are functions of V_0, the cell's initial transmembrane voltage for the current cycle.

Other approximate models for cardiac-tissue excitability have been proposed in the literature, including the piecewise-linear model of Biktashev [6] and the nonlinear model of Fenton and Karma [9]. The CLHA models of excitable cells learned by our technique retain the simplicity of Biktashev's model without sacrificing the expressiveness of Fenton-Karma.

6 Conclusions

We have presented a method for automatically learning CLHA that approximate, up to a prescribed error margin, the complex, nonlinear processes of amplification and propagation of electrical signals (APs) in excitable cells. Our method, implemented in MATLAB, combines geometric analysis with exponential regression (using the modified Prony's method) to derive a CLHA that covers in a cycle-linear manner the input/output behavior of the original nonlinear system. Moreover, it provides, to the best of our knowledge, the most accurate approximation of extant nonlinear excitable-cell models, such as HH and LRd.

A source of complexity in the HH and LRd models is the coupling between state variables, which seemingly occurs continuously throughout an AP cycle. In contrast, the coupling between state variables in the CLHA model is markedly reduced: (i) the membrane voltage v_0 at the time a stimulus arrives determines the coefficients of the flows of the "gated voltages" x_1 and x_2 for a complete cycle; (ii) x_1 and x_2's flows determine the membrane voltage within a mode; and (iii) the transition voltages V_p, $p \in \mathcal{P}$, determine the mode-switching logic within a cycle. This decoupling of state variable within the CLHA model may provide additional insight into essential properties of ECs, such as refractoriness and restitution. The derivatives of x_1 and x_2 approximate in each mode of the CLHA the inward and the outward currents, respectively.

It was therefore no surprise that the AP phases that were most difficult to linearly approximate were upstroke (U) and early repolarization (E): it is in these phases when disparate time constants coexist. For smaller error margins, we will most likely require three exponentials in the MPM approximation of these phases, and therefore three state variables in the corresponding modes. The third variable presumably will distinguish between the K and Ca currents contributing to repolarization.

As future work, we plan to more carefully consider APs at higher stimulation frequencies, i.e. in the 130 to 160 msecs range; accurately capturing their behavior also seems to require a third state variable. We also plan to develop additional training sets based on protocols incorporating stimuli of varying shapes and intensity. Furthermore, we intend to study AP propagation within an array of CLHA. Our preliminary data [26], showed that a 400-x-400 cell array was able to produce spirals (arrhythmia-related phenomena). We expect that the increased accuracy of the learned CLHA will better match observed behavior.

Finally, we would like to better understand the class of nonlinear systems whose behavior CLHA can successfully approximate. Intuitively, they exhibit periodic, but nonetheless adaptive, behavior with respect to the input stimuli. For narrow ranges of the epoch voltage v_0, CLHA naturally provide a linear approximation, and well-established techniques for reachability, stability, observability and controllability analysis can be readily applied. It would be interesting, however, to investigate whether the additional structure provided by CLHA (its parametrization on v_0) can be exploited to extend the reach of such techniques.

References

1. R. Alur, C. Belta, F. Ivancic, V. Kumar, M. Mintz, G. Pappas, H. Rubin, and J. Schug. Hybrid modeling and simulation of biomolecular networks. In *Hybrid Systems: Computation and Control, HSCC 2001*, volume 2034 of *LNCS*, pages 1932, 2001.
2. R. Alur, C. Coucoubetis, N. Halbwachs, T.A. Henzinger, P.H. Ho, X. Nicolin, A. Olivero, J. Sifakis, and S. Yovine. The altorithmic analysis of hybrid systems. *Theoretical Computer Siences*, 138:334, 1995.
3. P. Barbano, M. Spivak, J. Feng, M. Antoniotti, and B. Misra. A coherent framework for multi-resolution analysis of biological networks with memory: Ras pathway, cell cycle and immune system. In *Proc. National Academy of Science*, volume 102, pages 62456250, 2005.
4. C. Belta, P. Finin, L. Habets, A. Halasz, M. Imielinski, V. Kumar, and H. Rubin. Understanding the bacterial stringent response using reachability analysis of hybrid systems. In *Hybrid Systems:Computation and Control, HSCC 2004*, volume 2993 of *LNCS*, pages 111126, 2004.
5. A. Bemporad, A. Garulli, S. Paoletti, and A. Vicino. A bounded-error approach to piecewise affine system identification. *IEEE Transactions on Automatic Control*, 50(10):14731634, 2005.
6. V. N. Biktashev. A simplified model of propagation and dissipation of excitation fronts. *International Journal of Bifurcation and Chaos*, 13:36053619, 2003.
7. H. de Jong, J.-L. Gouz, C. Hernandez, M. Page, T. Sari, and J. Geiselmann. Hybrid modeling and simulation of genetic regulatory networks: A qualitative approach. In *Hybrid Systems: Computation and Control, HSCC 2003*, volume 2623 of *LNCS*, pages 267282, 2003.
8. T. Erbes. Stochastic learning feedback hybrid automata for dynamic power management in embedded systems, 2004.
9. F. Fenton and A. Karma. Vortex dynamics in 3d continous myocardium with fiber rotation: Filament instability and fibrillation. CHAOS, 8:2047, 1998.
10. G.C. Rigatos, I.P Vlahavas, and C.D. Spyropoulos. Fuzzy stochastic automata for reactive learning and hybrid control. In *Methods and Applications of Artificial Intelligence*, pages 366377. LNCS, Vol. 2308, Springer-Verlag, 2002.
11. R. Ghosh and C.J. Tomlin. Symbolic reachable set computation of piecewise affine hybrid automata and its application to biological modeling: Delta-notch protein signaling. *IEEE Transactions on Systems Biology*, 1(1):170183, 2004.

12. M. M. Henry. Model-based estimation of probabilistic hybrid automata, 2002.
13. A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane currents and its application to conduction and excitation in nerve. *J Physiol*, 117:500544, 1952.
14. M. Huber and R.A. Grupen. A hybrid architecture for learning robot control tasks. *Robotics Today*, 13(4), 2000.
15. K. Joshi, N. Neogi, and W. Sanders. Dynamic partitioning of large discrete event biological systems for hybrid simulation and analysis. In *Hybrid Systems: Computation and Control, HSCC 2004*, volume 2993 of *LNCS*, 2004.
16. D.K. Kaynar, N. Lynch, R. Segala, and F. Vaandrager. *The Theory of Timed I/O Automata*. Synthesis Lectures on Computer Science. Morgan Claypool, November 2005. Also available as Technical Report MIT-LCS-TR-917.
17. P. Lincoln and A. Tiwari. Symbolic systems biology: Hybrid modeling and analysis of biological networks. In *Hybrid Systems: Computation and Control, HSCC 2004*, volume 2993 of *LNCS*, pages 660672, 2004.
18. C. H. Luo and Y. Rudy. A dymanic model of the cardiac ventricular action potential: I. simulations of ionic currents and concentration changes. *Circ Res*, 74:1071 1096, 1994.
19. N. Lynch, R. Segala, and F. Vaandrager. Hybrid I/O automata.*Information and Computation*, 185(1):103157, 2003.
20. S. Mitra, D. Liberzon, and N. Lynch. Verifying average dwell time by solving optimization problems. In Ashish Tiwari and Joao P. Hespanha, editors,*Hybrid Systems: Computation and Control (HSCC 06)*, LNCS, Santa Barbara, CA, March 2006. Springer.
21. M.R. Osborne and G.K. Smyth. A modified Prony algorithm for exponential function fitting.*SIAM J. Sci. Comput.*, 16(1):119138, 1995.
22. V. Pavlovic, J.M. Rehg, T.-J. Cham, and K.P. Murphy. A dynamic bayesian network approach to figure tracking using learned dynamic models. In*Proc. of 7th IEEE Int. Conf. on Computer Vision*, 1999.
23. A. Singh and J. Hespanha. Models for generegulatory networks using polynomial stochastic hybrid systems. In*CDC05*, 2005.
24. R. Vidal, S. Soatto, Y. Ma, and S. Sastry. An algebraic geometric approach to the identification of a class of linear hybrid systems. In *Proc. of 42nd IEEE Conf. on Decision and Control*, 2003.
25. P. Ye, E. Entcheva, R. Grosu, and S. A. Smolka. Efficient modeling of excitable cells using hybrid automata. In Computational Methods in Systems Biology, 2005.
26. P. Ye, E. Entcheva, R. Grosu, and S.A. Smolka. Efficient modeling of excitable cells using hybrid automata. In CMSB05, *Computational Methods in Systems Biology Workshop*, Edinburgh, UK, April 2005.
27. P. Ye, E. Entcheva, S.A. Smolka, M.R. True, and R. Grosu. A cycle-linear approach to modeling action potentials. In *EMBC06, the IEEE International Conference of the Engineering in Medicine and Biology Society*, IEEE, New York City, NY, September 2006. IEEE.