Approximate Probabilistic Verification of Hybrid Systems

Benjamin M. Gyori[1], Bing Liu[2], Soumya Paul[3], R. Ramanathan[4], and P.S. Thiagarajan[1(✉)]

[1] Laboratory of Systems Pharmacology, Harvard Medical School, Boston, USA
benjamin_gyori@hms.harvard.edu, psthiagu@gmail.com
[2] Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, USA
[3] Institute de Recherche En Informatique de Toulouse, 31062 Toulouse, France
[4] Department of Computer Science, National University of Singapore, Singapore, Singapore

Abstract. Hybrid systems whose mode dynamics are governed by non-linear ordinary differential equations (ODEs) are often a natural model for biological processes. However such models are difficult to analyze. To address this, we develop a probabilistic analysis method by approximating the mode transitions as stochastic events. We assume that the probability of making a mode transition is proportional to the measure of the set of pairs of time points and value states at which the mode transition is enabled. To ensure a sound mathematical basis, we impose a natural continuity property on the non-linear ODEs. We also assume that the states of the system are observed at discrete time points but that the mode transitions may take place at any time between two successive discrete time points. This leads to a discrete time Markov chain as a probabilistic approximation of the hybrid system. We then show that for BLTL (bounded linear time temporal logic) specifications the hybrid system meets a specification iff its Markov chain approximation meets the same specification with probability 1. Based on this, we formulate a sequential hypothesis testing procedure for verifying–approximately–that the Markov chain meets a BLTL specification with high probability. Our case studies on cardiac cell dynamics and the circadian rhythm indicate that our scheme can be applied in a number of realistic settings.

Keywords: Hybrid systems · Markov chains · Dynamical systems · Statistical model checking

1 Introduction

Hybrid systems are often used to model biological processes [6,10,11]. The analysis of these models is difficult due to the high expressive power of the mixed

S. Paul—The author would like to thank ERC grant 269427 that supported a part of this work.

A. Abate and D. Šafránek (Eds.): HSB 2015, LNBI 9271, pp. 96–116, 2015.
DOI: 10.1007/978-3-319-26916-0_6

dynamics [22]. Various lines of work have explored ways to mitigate this problem with a common technique being to restrict the mode dynamics [2,3,14,18,20,23]. However, for many of the models arising in systems biology the mode dynamics will be governed by a system of non-linear ordinary differential equations (ODEs). To analyze such systems, we develop a scheme under which such systems can be approximated as a discrete time Markov chain.

A key difficulty in analyzing a hybrid system's behavior is that the time points and value states at which a trajectory meets a guard will depend on the solutions to the ODE systems associated with the modes. For high-dimensional systems these solutions will not be available in closed form. To get around this, we approximate the mode transitions as stochastic events by fixing the probability of a mode transition to be proportional to the measure of the set of value state and time point pairs at which this transition is enabled. More sophisticated hypotheses could be considered. For instance one could tie the mode transition probability to how long the guard has been continuously enabled or how deeply within a guard region the current state is. To bring out the main ideas we will postpone exploring such approximations to our future work.

To secure a sound mathematical basis for our approximation, we further assume: (i) The vector fields associated with the ODEs are Lipschitz continuous functions. (ii) The states of the hybrid system are observable only at discrete time points. (iii) The set of initial states and the guard sets are bounded open sets. (iv) The hybrid dynamics is strictly non-Zeno in the sense that there is a uniform upper bound on the number of transitions that can take place in a unit time interval. For technical convenience we in fact assume that time discretization is so chosen that at most one mode transition takes place between two successive discrete time points.

Under these assumptions, we show that the dynamics of the hybrid system H can be approximated as an infinite state Markov chain M. Given the application domain we have in mind, namely, biological pathway dynamics, we focus on the behavior of the hybrid system over a finite time horizon and BLTL (bounded linear time temporal logic) [15] to specify dynamic properties of interest. The maximum discrete time point we fix will be determined by the BLTL specification. Our probabilistic approximation is such that the set of trajectories satisfying a BLTL formula will correspond to a measurable set of paths of the Markov chain and hence can be assigned a probability value. We then show that H meets the specification ψ–i.e. every trajectory of H is a model of ψ–iff M meets the specification ψ with probability 1. This allows us to approximately verify interesting properties of the hybrid system using its Markov chain approximation. However, even a bounded portion of M can not be constructed effectively. This is because the transition probabilities of the Markov chain will depend on the solutions to the ODEs associated with the modes, which will not be available in a closed form. In addition, the structure of M itself will be unknown since the states of the chain will be those that can reached with non-zero probability from the initial mode and we can not determine which transitions have non-zero probabilities. To cope with this, we design a statistical model checking procedure

to approximately verify that the chain (and hence the hybrid system) almost certainly meets the specification. One just needs to ensure that the dynamics of the Markov chain is being sampled according to underlying probabilities. We achieve this by randomly generating trajectories of H through numerical simulations in a way that corresponds to randomly sampling the paths of the Markov chain according to its underlying structure and transition probabilities. We note that a simple minded Monte Carlo simulations based strategy consisting of sampling an initial state (according to the given initial distribution) followed by a random generation of trajectory will flounder on the issue of how one "randomly" picks a mode transition during the generation of a trajectory in the presence of the non-linear dynamics captured by the ODEs systems. Our approximation technique instead establishes a principled way of achieving this.

In establishing these results, we assume that the atomic propositions in the specification are interpreted over the modes of the hybrid system. Consequently one can specify patterns of mode visitations while quantitative properties can be inferred only indirectly and in a limited fashion. Our results however can be extended to handle quantitative atomic propositions ("the current concentration of protein X is greater than 2 μM"). Due to space limitations, we present the details of this extension in a technical report available online [4].

To demonstrate the applicability of our method, we first study the electrical activity of cardiac cells represented by a hybrid model. By varying parameters we analyze key dynamical properties on multiple cell types, in healthy and disease conditions, and under different input stimuli. We also analyze a hybrid model of the circadian rhythm, and find distinct roles of multiple feedback loops in maintaining oscillatory properties of the dynamics.

1.1 Related Work

Mode transitions have been approximated as random events in the literature. In [1] the dynamics of a hybrid system is approximated by substituting the guards with probabilistic barrier functions. Our transition probabilities are constructed using similar but simpler considerations. We have done so in order to be able to carry out temporal logic based verification based on simulations. An alternative approach to approximately verifying non-linear hybrid systems is one based on δ-reals [19]. Here one verifies bounded reachability properties that are robust under small perturbations of the numerical values mentioned in the specification. Since the approximation involved is of a very different kind, it is difficult to compare this line of work with ours. However, it may be fruitful to combine the two approaches to verify a richer set of reachability properties.

The present work may be viewed as an extension of [31] where a *single* system of ODEs is considered. This method however, breaks down in the multi-mode hybrid setting and one needs to construct–as we do here–an entirely new machinery. Finally, a wealth of literature is available on the analysis of stochastic automata [5,8,13,25]. It will be interesting to explore if these methods can be transported to our setting.

2 Hybrid Automata

We fix n real-valued variables $\{x_i\}_{i=1}^n$ viewed as functions of time $x_i(t)$ with $t \in \mathbb{R}_+$, the set of non-negative reals. A valuation of $\{x_i\}_{i=1}^n$ is $\mathbf{v} \in \mathbb{R}^n$ with $\mathbf{v}(i) \in \mathbb{R}$ representing the value of x_i. The language of *guards* is given by: (i) $a < x_i$ and $x_i < b$ are guards where a, b are rationals and $i \in \{1, 2, \ldots, n\}$. (ii) If g and g' are guards then so are $g \wedge g'$ and $g \vee g'$.

$\mathcal{G}$ denotes the set of guards. We define $\mathbf{v} \models g$ (i.e. $\mathbf{v}$ satisfies the guard g) via: $\mathbf{v} \models a < x_i$ iff $a < \mathbf{v}(i)$ and similarly for $x_i < b$. The clauses for conjunction and disjunction are standard. We let $\|g\| = \{\mathbf{v} \mid \mathbf{v} \models g\}$. We note that $\|g\|$ is an open subset of $\mathbb{R}^n$ for every guard g. We will abbreviate $\|g\|$ as g.

Definition 1. *A hybrid automaton is a tuple* $H=(Q, q_{in}, \{F_q(\boldsymbol{x})\}_{q \in Q}, \mathcal{G}, \rightarrow,$ INIT*), where*

- Q *is a finite set of* modes *and* $q_{in} \in Q$ *is the* initial *mode.*
- *For each* $q \in Q$, $d\boldsymbol{x}/dt = F_q(\boldsymbol{x})$ *is a system of ODEs, where* $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ *and* $F_q = (f_q^1(\boldsymbol{x}), f_q^2(\boldsymbol{x}), \ldots, f_q^n(\boldsymbol{x}))$*. Further,* f_q^i *is Lipschitz continuous for each* i*.*
- $\rightarrow \subseteq (Q, \mathcal{G}, Q)$ *is the mode transition relation. If* $(q, g, q') \in \rightarrow$ *we shall often write it as* $q \xrightarrow{g} q'$*.*
- INIT $= (L_1, U_1) \times (L_2, U_2) \ldots \times (L_n, U_n)$ *is the set of initial states where* $L_i < U_i$ *and* L_i, U_i *are rationals.*

We have not associated invariant conditions with the modes or reset conditions with the mode transitions. They can be introduced with some additional work.

Fixing a suitable unit time interval Δ, we discretize the time domain as $t = 0, \Delta, 2\Delta, \ldots$. We assume the states of the system are observed only at these discrete time points. Furthermore, we shall assume that only a bounded number of mode changes can take place between successive discrete time points. Both in engineered and biological processes this is a reasonable assumption. Given this, we shall in fact assume that Δ is such that at most one mode change takes place within a Δ time interval. We note that there can be multiple choices for Δ that meet this requirement and in practice one must choose this parameter carefully. (Our method can be extended to handle a bounded number of mode transitions in a unit time interval but this will entail notational complications that will obscure the main ideas.) In what follows, for technical convenience we also assume the time scale has been normalized so that $\Delta = 1$. As a result, the discretized set of time points will be $\{0, 1, 2, \ldots\}$.

2.1 Trajectories

We have assumed that for every mode q, the right hand side of the ODEs, $F_q(\mathbf{x})$, is Lipschitz continuous for each component. As a result, for each initial value $\mathbf{v} \in \mathbb{R}^n$ and in each mode q, the system of ODEs $d\mathbf{x}/dt = F_q(\mathbf{x})$ will have a

unique solution $Z_{q,\mathbf{v}}(t)$ [24]. We are also guaranteed that $Z_{q,\mathbf{v}}(t)$ is Lipschitz and hence measurable [24]. It will be convenient to work with two sets of functions derived from solutions to the ODE systems.

The (unit interval) *flow* $\Phi_q : (0,1) \times \mathbb{R}^n \to \mathbb{R}^n$ is given by $\Phi_q(t, \mathbf{v}) = Z_{q,\mathbf{v}}(t)$. Φ_q will also be Lipschitz. Next we define the parametrized family of functions $\Phi_{q,t} : \mathbb{R}^n \to \mathbb{R}^n$ given by $\Phi_{q,t}(\mathbf{v}) = \Phi_q(t, \mathbf{v})$. Applying once again the fact that the RHS of the ODEs are Lipschitz continuous functions, we can conclude that these parametrized functions $\Phi_{q,t}$ (which will be $1-1$) as well as their inverses will be Lipschitz.

A (finite) *trajectory* is a sequence $\tau = (q_0, \mathbf{v}_0)(q_1, \mathbf{v}_1) \ldots (q_k, \mathbf{v}_k)$ such that for $0 \leq j < k$ the following conditions are satisfied: (i) For $0 \leq j < k$, $q_j \xrightarrow{g_j} q_{j+1}$ for some guard g_j. (ii) there exists $t \in (0,1)$ such that $\Phi_{q_j,t}(\mathbf{v}_j) \in g$. Furthermore $\mathbf{v}_{j+1} = \Phi_{q_{j+1},1-t}(\Phi_{q_j,t}(\mathbf{v}_j))$.

We say that the trajectory τ as defined above *starts* from q_0 and *ends* in q_k. Further, its initial value state is $\mathbf{v}_0$ and its final value state is $\mathbf{v}_k$. We let TRJ denote the set of all finite trajectories that start from the initial mode q_{in} with an initial value state in INIT.

3 The Markov Chain Approximation

A (finite) path in H is a sequence $\rho = q_0 q_1 \ldots q_k$ such that for $0 \leq j < k$, there exists a guard g_j such that $q_j \xrightarrow{g_j} q_{j+1}$. We say that this path starts from q_0, ends at q_k and is of length $k+1$. We let paths_H denote the set of all finite paths that start from q_{in}.

In what follows μ will denote the standard Lebesgue measure over finite dimensional Euclidean spaces. We will construct $M_H = (\Upsilon, \Rightarrow)$, the Markov chain approximation of H inductively. Each state in Υ will be of the form $(\rho, X, \mathbf{P}_X)$ with $\rho \in \mathsf{paths}_H$, X an open subset of $\mathbb{R}^n$ of non-zero, finite measure and $\mathbf{P}_X$ a probability distribution over $SA(X)$, the Borel σ-algebra generated by X.

We start with $(q_{in}, \text{INIT}, \mathbf{P}_{\text{INIT}}) \in \Upsilon$. Clearly, INIT is an open set of non-zero, finite measure since $\mu(\text{INIT}) = \prod_i (U_i - L_i)$. For technical convenience we shall assume $\mathbf{P}_{\text{INIT}}$ to be the uniform probability distribution, but other probability distributions with respect to which the Lebesgue-measure is absolutely continuous could also be chosen. Assume inductively that $(\rho, X, \mathbf{P}_X)$ is in Υ with X an open subset of $\mathbb{R}^n$ of non-zero, finite measure and $\mathbf{P}_X$ a probability distribution over $SA(X)$. Suppose ρ ends in q and there are m outgoing transitions $q \xrightarrow{g_1} q_1, \ldots, q \xrightarrow{g_m} q_m$ from q in H (Fig. 1 illustrates this inductive step).

Then for $1 \leq j \leq m$ we define the triples $(\rho q_j, X_j, \mathbf{P}_{X_j})$ as follows. In doing so we will assume the required properties of the objects involved in this construction. We will then establish these properties and thus the soundness of the construction. For convenience, through the remaining parts of this section j will range over $\{1, 2, \ldots, m\}$.

For each $\mathbf{v} \in X$ and each j we first define the set of time points $\mathbb{T}_j(\mathbf{v}) \subseteq (0,1)$ via

$$\mathbb{T}_j(\mathbf{v}) = \{t \mid \Phi_q(t, \mathbf{v}) \in g_j\}. \tag{1}$$

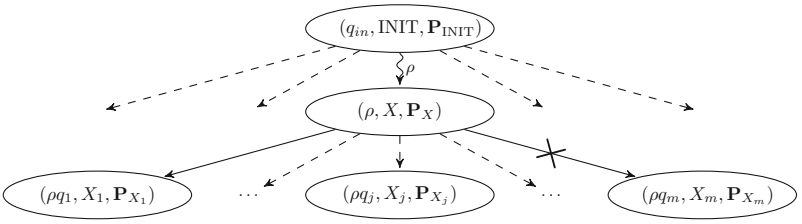


Fig. 1. The Markov chain construction. The edge from the state $(\rho, X, \mathbf{P}_X)$ to the state $(\rho q_m, X_m, \mathbf{P}_{X_m})$ marked with a '$\times$' represents the case where X_m has measure 0, and hence the probability of this transition is 0. Thus, $(\rho q_m, X_m, \mathbf{P}_{X_m})$ will not be a state of the Markov chain.

Thus $\mathbb{T}_j(\mathbf{v})$ is the set of time points in $(0, 1)$ at which the guard g_j is satisfied if the system starts from $\mathbf{v}$ in mode q and evolves according to dynamics of mode q up to time t. We next define X_j for each j as

$$X_j = \bigcup_{\mathbf{v} \in X} \{\Phi_{q_j}(1 - t, \Phi_q(t, \mathbf{v})) \mid t \in \mathbb{T}_j(\mathbf{v})\}. \tag{2}$$

Thus X_j is the set of all value states obtained by starting from some $\mathbf{v} \in X$ at time k, evolving up to $k+t$ according to the dynamics q, making an instantaneous mode switch to q_j at this time point, and evolving up to time $k + 1$ according to dynamics of mode q_j.

To complete the definition of the triples $(\rho q_j, X_j, \mathbf{P}_{X_j})$, we first denote by $\mathbf{P}_{\mathbb{T}_j(\mathbf{v})}$ a probability distribution over $\mathbb{T}_j(\mathbf{v})$. We shall choose this distribution to be uniform but it could be any other non-uniform probability distribution with respect to which the Lebesgue-measure is absolutely continuous. We now define the probability distributions $\mathbf{P}_{X_j}$ over $SA(X_j)$ as follows. Suppose Y is a measurable subset of X_j. Then

$$\mathbf{P}_{X_j}(Y) = \int_{\mathbf{v} \in X} \int_{t \in \mathbb{T}_j(\mathbf{v})} \mathbf{1}_{(\Phi_{q_j}(1-t, \Phi_q(t, \mathbf{v})) \cap Y)} d\mathbf{P}_{\mathbb{T}_j(\mathbf{v})} d\mathbf{P}_X. \tag{3}$$

As usual $\mathbf{1}_Z$ is the indicator function of the set Z while $d\mathbf{P}_{\mathbb{T}_j(\mathbf{v})}$ indicates that the inner integration over $\mathbb{T}_j(\mathbf{v})$ is w.r.t. the probability measure $\mathbf{P}_{\mathbb{T}_j(\mathbf{v})}$ and $d\mathbf{P}_X$ indicates that the outer integration over X is w.r.t. the probability measure $\mathbf{P}_X$. Thus $\mathbf{P}_{X_j}(Y)$ captures the probability that the value state $\Phi_{q_j}(1 - t, \Phi_q(t, \mathbf{v}))$ lands in $Y \subseteq X_j$ by taking the transition $q \xrightarrow{g_j} q_j$ at some time point in $\mathbb{T}_j(\mathbf{v})$ given that one started with some value state in X.

Next we define the triples $((\rho, X, \mathbf{P}_X), p_j, (\rho q_j, X_j, \mathbf{P}_{X_j}))$, where p_j is given by

$$p_j = \int_{\mathbf{v} \in X} \frac{\mu(\mathbb{T}_j(\mathbf{v}))}{\sum_{\ell=1}^{m} \mu(\mathbb{T}_\ell(\mathbf{v}))} d\mathbf{P}_X. \tag{4}$$

Thus p_j captures the probability of taking the mode transition $q \xrightarrow{g_j} q_j$ when starting from the value states in X and mode q. For every j we add

the state $(\rho q_j, X_j, \mathbf{P}_{X_j})$ to Υ and the triple $((\rho, X, \mathbf{P}_X), p_j, (\rho q_j, X_j, \mathbf{P}_{X_j}))$ to $\Rightarrow$ iff $\mu(X_j) > 0$.

Finally, $(q_{in}, \text{INIT}, \mathbf{P}_{\text{INIT}})$ is the initial state of M_H. We can summarize the key properties of our construction as follows (while assuming the associated terminology and notations).

Theorem 1. *1. $\mathbb{T}_j(\boldsymbol{v})$ is an open set of finite measure for each $\boldsymbol{v} \in X$ and each j.*
2. X_j is open and is of finite measure for each j.
3. If $(\rho q_j, X_j, \boldsymbol{P}_{X_j}) \in \Upsilon$ then $\mu(X_j) > 0$.
4. $\boldsymbol{P}_{X_j}$ is a probability distribution for each j.
5. $M_H = (\Upsilon, \Rightarrow)$ is an infinite state Markov chain whose underlying graph is a finitely branching tree.

Proof. To prove the first part, suppose $t \in \mathbb{T}_j(\mathbf{v})$. Then $\Phi_q(t, \mathbf{v}) = \mathbf{v}' \in g_j$ and g_j is open. Hence $\mathbf{v}'$ will be contained in an open neighborhood U contained in g_j. Since Φ_q is Lipschitz we can pick U such that $Y' = \Phi_q^{-1}(U)$ is an open set containing $(\mathbf{v}, t)$ with $Y' \subseteq (0,1) \times X$. Thus every element of $\mathbb{T}_j(\mathbf{v})$ is contained in an open neighborhood in $(0,1)$ and hence $\mathbb{T}_j(\mathbf{v})$ is open.

Using the definition of X_j, the fact that X and $\mathbb{T}_j(\mathbf{v})$ are open, and the continuity of the inverses of the flow functions it is easy to observe that X_j is open. To see that it is of finite measure, by the induction hypothesis, X is open and $\mu(X)$ is finite. Hence $((0,1) \times X)$ is open as well and $\mu((0,1) \times X)$ is finite. Since $\mathbb{R}^{n+1}$ is second-countable [33], there exists a countable family of disjoint open-intervals $\{I_i\}_{i \geq 1}$ in $\mathbb{R}^{n+1}$ such that $((0,1) \times X) = \bigcup_i I_i$. Clearly each I_i has a finite measure. By the Lipschitz continuity of Φ_q we know that there exists a constant c such that $\mu(\Phi_q(I_i)) < c \cdot \mu(I_i)$ for all i. We thus have

$$\begin{aligned} \mu(\Phi_q((0,1), X)) &\leq \sum_i \mu(\Phi_q(I_i)) \\ &< c \sum_i \mu(I_i) = c\mu((0,1) \times X) < \infty. \end{aligned} \tag{5}$$

Therefore $\Phi_q((0,1), X)$ has a finite measure. By a similar argument we can show that $\Phi_{q_j}((0,1), \Phi_q((0,1), X))$ has a finite measure as well. Since $X_j = \bigcup_t \Phi_{q_j, 1-t}(\Phi_{q,t}(X) \cap g) \subseteq \Phi_{q_j}((0,1), \Phi_q((0,1), X))$, it must have a finite measure.

The remaining parts of the theorem follow easily from the definitions and basic measure theory.

4 Relating the Behaviors of H and M_H

We shall use bounded linear-time temporal logic (BLTL) [15] to specify time bounded properties and use it to relate the behaviors of H and M_H. For convenience we shall write M instead of M_H from now on.

We assume a finite set of atomic propositions AP and a valuation function $Kr : Q \to 2^{AP}$. Formulas of BLTL are defined as: (i) Every atomic proposition as well as the constants *true*, *false* are formulas. (ii) If ψ, ψ' are formulas then $\neg\psi$ and $\psi \vee \psi'$ are formulas. (iii) If ψ, ψ' are formulas and ℓ is a positive integer then $\psi \mathbf{U}^{\leq \ell} \psi'$ is a formula. The derived operators $\mathbf{F}^{\leq \ell}$ and $\mathbf{G}^{\leq \ell}$ are defined as usual: $\mathbf{F}^{\leq \ell}\psi \equiv \mathit{true}\mathbf{U}^{\leq \ell}\psi$ and $\mathbf{G}^{\leq \ell}\psi \equiv \neg\mathbf{F}^{\leq \ell}\neg\psi$.

We shall assume through the rest of the paper that the behavior of the system is of interest only up to a maximum time point $K > 0$. This is guided by the fact that given a BLTL formula ψ there is a constant K_ψ that depends only on ψ so that it is enough to evaluate an execution trace of length at most K_ψ to determine whether ψ is satisfied [7]. Hence we assume that a sufficiently high K has been chosen to handle the specifications of interest. Having fixed K, we denote by TRJ^{K+1} the trajectories of length $K+1$, and view this set as representing the time bounded non-deterministic behavior of H of interest.

To develop the corresponding notion for M, we first define a finite path in M to be a sequence $\eta_0\eta_1 \ldots \eta_k$ such that $\eta_j \in \Upsilon$ for $0 \leq j \leq k$. Furthermore for $0 \leq j < k$ there exists $p_j \in (0, 1]$ such that $\eta_j \overset{p_j}{\Rightarrow} \eta_{j+1}$. Such a path is said to start from η_0 and its length is $k+1$. We define paths_M to be the set of finite paths that start from the initial state of M while paths_M^{K+1} is the set of paths in paths_M of length $K+1$.

The Trajectory Semantics: Let $\tau = (q_0, \mathbf{v}_0)\ (q_1, \mathbf{v}_1)\ \ldots\ (q_k, \mathbf{v}_k)$ be a finite trajectory, ψ a BLTL formula and $0 \leq j \leq K$. Then $\tau, j \models_H \psi$ is defined via:

- $\tau, j \models_H A$ iff $A \in Kr(q_j)$, where A is an atomic proposition.
- $\neg$ and $\vee$ are interpreted in the usual way.
- $\tau, j \models_H \psi \mathbf{U}^{\leq \ell}\psi'$ iff there exists j' such that $j' \leq \ell$ and $j + j' \leq k$ and $\tau, (j + j') \models_H \psi'$. Further, $\tau, (j + j'') \models_H \psi$ for every $0 \leq j'' < j'$.

We now define $\mathit{models}_H(\psi) \subseteq TRJ^{K+1}$ via: $\tau \in \mathit{models}_H(\psi)$ iff $\tau, 0 \models_H \psi$. We say that H *meets the specification* ψ, denoted $H \models \psi$, iff $\mathit{models}_H(\psi) = TRJ^{K+1}$.

The Markov chain semantics: Let $\pi = \eta_0\eta_1 \ldots \eta_k$ be a path in M with $\eta_j = (\rho q_j, X_j, \mathbf{P}_{X_j})$ for $0 \leq j \leq k$. Let ψ be a BLTL formula and $0 \leq j \leq k$. Then $\pi, j \models_M \psi$ is given by:

- $\pi, j \models_M A$ iff $A \in Kr(q_j)$, where A is an atomic proposition.
- The remaining clauses are defined just as in the case of $\models_H$.

Now we define $\mathit{models}_M(\psi) \subseteq \mathsf{paths}_M^{K+1}$ via: $\pi \in \mathit{models}_M(\psi)$ iff $\pi, 0 \models_M \psi$. We can now define the probability of satisfaction of a formula in M. Let $\pi = \eta_0\eta_1 \ldots \eta_K$ be in paths_M^{K+1}. Then $\Pr(\pi) = \prod_{0 \leq \ell < K} p_\ell$, where $\eta_\ell \overset{p_\ell}{\Rightarrow} \eta_{\ell+1}$ for $0 \leq \ell < K$. This leads to

$$\Pr(\mathit{models}_M(\psi)) = \sum_{\pi \in \mathit{models}_M(\psi)} \Pr(\pi).$$

We write $M \models \psi$ to denote $\Pr(models_M(\psi)) = 1$. For $p \in [0,1]$ we write as usual $\Pr_{\geq p}(\psi)$ instead of $\Pr(models_M(\psi)) \geq p$. We note that $Pr(\pi) > 0$ for every $\pi \in models_M(\psi)$. Furthermore $\sum_{\pi \in models_M(\psi)} \Pr(\pi) \leq 1$. Hence $\Pr_{\geq 1}(\psi)$ iff $models_M(\psi) = \mathsf{paths}_M^{K+1}$ iff $M \models \psi$.

4.1 The Correspondence Result

We wish to show that H meets the specification ψ iff $\Pr_{\geq 1}(\psi)$. To this end let $\pi = \eta_0\eta_1 \ldots \eta_k$ be a path in M with $\eta_j = (q_0 q_1 \ldots q_j, X_j, \mathbf{P}_{X_j})$ for $0 \leq j \leq k$ and let $\tau = (q'_0, \mathbf{v}_0)\ (q'_1, \mathbf{v}_1) \ldots (q'_{k'}, \mathbf{v}_{k'})$ be a trajectory. Then we say that π and τ are *compatible* iff $k = k'$ and $q_j = q'_j$ and $\mathbf{v}_j \in X_j$ for $0 \leq j \leq k$. The following three observations based on this notion will easily lead to the main result.

Lemma 1. *1. Suppose the path $\pi = \eta_0\eta_1 \ldots \eta_k$ in M and the trajectory $\tau = (q_0, \boldsymbol{v}_0)\ (q_1, \boldsymbol{v}_1) \ldots (q_k, \boldsymbol{v}_k)$ are compatible. Let $0 \leq j \leq k$ and ψ be a BLTL formula. Then $\pi, j \models_M \psi$ iff $\tau, j \models_H \psi$.*
2. Suppose π is a path in M. Then there exists a trajectory τ such that π and τ are compatible. Furthermore if $\pi \in \mathsf{paths}_M$ then $\tau \in TRJ$.
3. Suppose τ is a trajectory. Then there exists a path π in M such that τ and π are compatible. Furthermore if $\tau \in TRJ$ then $\pi \in \mathsf{paths}_M$.

Proof. The proof follows via a systematic application of the definitions. To prove the first part we note that if A is an atomic proposition then $\pi, j \models_M A$ iff $A \in Kr(q_j)$ iff $\tau, j \models_H A$. We next note that the suffix of length m of π will be compatible with the suffix of length m of τ whenever π and τ are compatible. The result now follows at once by structural induction on ψ.

To show the second part let $\pi = \eta_0\eta_1 \ldots \eta_k$ be a path in M with $\eta_j = (q_0 q_1 \ldots q_j, X_j, \mathbf{P}_{X_j})$ for $0 \leq j \leq k$. Clearly X_j is non-empty for $0 \leq j \leq k$ since $\eta_j \in \Upsilon$ implies $\mu(X_j) > 0$. We proceed by induction on k. If $k = 0$ then we can pick $\mathbf{v}_0 \in X_0$ and the trajectory $(q_0, \mathbf{v}_0)$ will be compatible with τ. So assume $k > 0$. Then by the induction hypothesis there exists a trajectory $(q_1, \mathbf{v}_1)(q_2, \mathbf{v}_2) \ldots (q_k, \mathbf{v}_k)$ which is compatible with the path $\eta_1\eta_2 \ldots \eta_k$. Let $q_0 \xrightarrow{g} q_1$. Since $\mathbf{v}_1 \in X_1$ there must exist $\mathbf{v}_0$ in X_0 and $t \in (0,1)$ such that $\Phi_{q_0,t}(\mathbf{v}_0) \in g$ and $\mathbf{v}_1 = \Phi_{q_1,1-t}(\Phi_{q_0,t}(\mathbf{v}_0))$. Clearly $\mathbf{v}_0\mathbf{v}_1 \ldots \mathbf{v}_k$ is a trajectory that is compatible with π. The fact that $\tau \in TRJ$ if $\pi \in \mathsf{paths}_M$ follows from the definition of compatibility.

To prove the third part let $\tau = (q_0, \mathbf{v}_0)\ (q_1, \mathbf{v}_1) \ldots (q_k, \mathbf{v}_k) \in TRJ$. Again we proceed by induction on k. Suppose $k = 0$. Then $(q_{in}, \mathrm{INIT}, \mathbf{P}_{\mathrm{INIT}})$ is in paths_M which is compatible with τ. So suppose $k > 0$. Then by the induction hypothesis there exits $\pi' = \eta_0\eta_1 \ldots \eta_{k-1}$ such that π' is compatible with $\tau' = (q_0, \mathbf{v}_0)(q_1, \mathbf{v}_1) \ldots (q_{k-1}, \mathbf{v}_{k-1})$. Let $q_{k-1} \xrightarrow{g} q_k$. Since X_{k-1} is open there exists an open neighborhood $Y \subseteq X_{k-1}$ that contains $\mathbf{v}_{k-1}$. But then both $\Phi^{-1}_{q_{k-1}}$ and $\Phi^{-1}_{q_k}$ are continuous. Thus $\Phi_{q_{k-1},t}(Y)$ is open and $\Phi_{q_{k-1},t}(Y) \cap g$ should be open and non-empty (since g is open and $(q_k, \mathbf{v}_k)$ is part of the trajectory). Hence $Y' = \bigcup_{t \in (0,1)} \Phi_{q_k,1-t}(\Phi_{q_{k-1},t}(Y) \cap g)$ is a non-empty open set with a positive measure. This means there will be a state of the form $\eta_k = (\rho_k, X_k, \mathbf{P}_{X_k})$ in Υ

with $Y' \subseteq X_k$ and $\eta_{k-1} \stackrel{p}{\Rightarrow} \eta_k$ for some $p \in (0,1]$. Clearly $\pi = \pi'\eta_k \in \mathsf{paths}_M$ and is compatible with τ. Again the fact that $\pi \in \mathsf{paths}_M$ if $\tau \in TRJ$ follows from the definition of compatibility.

Theorem 2. *$H \models \psi$ iff $M \models \psi$.*

Proof. Suppose H does not meet the specification ψ. Then there exists $\tau \in TRJ^{K+1}$ such that $\tau, 0 \not\models_H \psi$. By the third part of Lemma 1 there exists $\pi \in \mathsf{paths}_M^{K+1}$ which is compatible with τ. By the first part of Lemma 1 we then have $\pi \notin models_M(\psi)$ which leads to $Pr_{<1}(\psi)$.

Next suppose that $Pr_{<1}(\psi)$. Then there exists $\pi \in \mathsf{paths}^{K+1}$ such that $\pi, 0 \not\models_M \psi$. By the second part of Lemma 1 there exists $\tau \in TRJ^{K+1}$ which is compatible with π. By the first part of Lemma 1 this implies $\tau, 0 \not\models_H \psi$ and this in turn implies that H does not meet the specification ψ.

4.2 Quantitative Atomic Propositions

The above results can be extended to handle atomic propositions of the form $\langle x_i < c \rangle$ and $\langle x_i > c \rangle$ where c is a rational constant. We partition $\mathbb{R}^n$ into hypercubes according to the constants appearing in the given set of quantitative atomic propositions in AP_{qt}. We then blow up the state space of the Markov chain to record which hypercube the current values of the variables fall in. We restrict our attention to robust trajectories and show that every robust trajectory of H meets a BLTL specification iff its Markov chain approximation meets the same specification with probability 1. Informally a robust trajectory is one which has an open neighborhood of trajectories under a natural topology over the space of $K+1$-length trajectories. Under an associated measure the set of non-robust trajectories will have measure 0. The details can be found in [4].

5 The SMC Procedure

To verify whether H meets the specification ψ, we solve the equivalent problem whether $\Pr_{\geq 1}(\psi)$ on M. However, as discussed in Sect. 1, M cannot be constructed explicitly since both its structure and transition probabilities, defined in terms of the solutions to the ODEs, will not be available. Therefore we shall use randomly generated trajectories to sample the paths of M and formulate a sequential hypothesis test to decide with bounded error rate whether $\Pr_{\geq 1}(\psi)$ holds. Algorithm 1 describes our trajectory sampling procedure.

Clearly Algorithm 1 generates a trajectory in TRJ^{K+1}. We now relate these trajectories to paths in M.

The initial value $\mathbf{v}_0$ is sampled uniformly on INIT, and we start in mode q_{in}, consistent with the initial state (q_{in},INIT,$\mathbf{P}_{\text{INIT}}$) of M. Inductively, suppose $\eta = (\rho, X, \mathbf{P}_X)$ is a state of M with ρ ending in q. Suppose $\eta \stackrel{p_j}{\Rightarrow} \eta_j$ is a transition in M such that $\eta_j = (\rho q_j, X_j, \mathbf{P}_{X_j})$.

Algorithm 1. Trajectory simulation

Input: Hybrid automaton $H = (Q, q_{in}, \{F_q(\mathbf{x})\}_{q \in Q}, \mathcal{G}, \rightarrow, \text{INIT})$, maximum time step K.
Output: Trajectory τ

Sample $\mathbf{v}_0$ from INIT uniformly, set $q_0 := q_{in}$ and $\tau := (q_0, \mathbf{v}_0)$.
for $k := 1 \ldots K$ **do**
  Generate time points $T := \{t_1, \ldots, t_J\}$ uniformly in $(0, 1)$.
  Simulate $\mathbf{v}^j := \Phi_{q_{k-1}}(t_j, \mathbf{v}_{k-1})$, for $j \in \{1, \ldots, J\}$
  Let $\widehat{\mathbb{T}}_j := \{t \in T : \mathbf{v}^j \in g_j\}$ be the time points where g_j is enabled.
  Pick g_ℓ randomly according to probabilities $\{p_j := |\widehat{\mathbb{T}}_j| / \sum_{i=1}^m |\widehat{\mathbb{T}}_i|\}$.
  Pick t_ℓ uniformly at random from $\widehat{\mathbb{T}}_\ell$.
  Simulate $\mathbf{v}' := \Phi_{q'}(1 - t_\ell, \mathbf{v}^\ell)$, where q' is the target of g_ℓ.
  Set $q_k := q'$, $\mathbf{v}_k := \mathbf{v}'$, and extend $\tau := (q_0, \mathbf{v}_0) \ldots (q_k, \mathbf{v}_k)$.
end for
return τ

Proposition 1. *Suppose, we obtain a sample $\boldsymbol{v} \sim \boldsymbol{P}_X$. The probability of choosing guard g_j whose target mode is q_j in Algorithm 1 tends to p_j as $J \to \infty$.*

Proof. According to Algorithm 1, the probability of picking guard g_j for a trajectory starting at $\mathbf{v} \in X$ is defined as $|\widehat{\mathbb{T}}_j| / \sum_{i=1}^m |\widehat{\mathbb{T}}_i|$, which, by the law of large numbers tends to

$$p_j(\mathbf{v}) := \frac{\mu(\mathbb{T}_j(\mathbf{v}))}{\sum_{i=1}^m \mu(\mathbb{T}_i(\mathbf{v}))} \tag{6}$$

as J tends to ∞.

Now if $\mathbf{v}$ is randomly sampled according to $\mathbf{P}_X$, then the probability of picking guard j can be expressed as the expected value of $p_j(\mathbf{v})$ under $\mathbf{v} \sim \mathbf{P}_X$ as

$$\mathbb{E}_{\mathbf{v} \sim \mathbf{P}_X}[p_j(\mathbf{v})] = \int_{\mathbf{v} \in X} p_j(\mathbf{v}) d\mathbf{P}_X = \int_{\mathbf{v} \in X} \frac{\mu(\mathbb{T}_j(\mathbf{v}))}{\sum_{i=1}^m \mu(\mathbb{T}_i(\mathbf{v}))} d\mathbf{P}_X, \tag{7}$$

which by (4) is equal to p_j, the corresponding transition probability in the Markov chain.

Similarly, picking the transition time t from $\widehat{\mathbb{T}}_j$ will approximate sampling $t \sim \mathbf{P}_{\mathbb{T}_j(\mathbf{v})}$, for sufficiently high J. Next, assume that we have picked $q \xrightarrow{g_j} q_j$ as the transition to take. We sample $t \sim \mathbf{P}_{\mathbb{T}_j(\mathbf{v})}$, and obtain $\mathbf{v}'$ by numerical simulation via:

$$\mathbf{v}' = \Phi_{q_j}(1 - t, \Phi_q(t, \mathbf{v})). \tag{8}$$

Proposition 2. *$\boldsymbol{v}'$ is distributed according to $\boldsymbol{P}_{X_j}$.*

Proof. Clearly it suffices to show that for a measurable subset $Y \subseteq X_j$, $\Pr(\mathbf{v}' \in Y) = \mathbf{P}_{X_j}(Y)$. We start with

$$\Pr(\mathbf{v}' \in Y \mid \mathbf{v}) = \int_{t \in \mathbb{T}_j(\mathbf{v})} \mathbf{1}_{(\Phi_{q_j}(1-t, \Phi_q(t, \mathbf{v})) \cap Y)} d\mathbf{P}_{\mathbb{T}_j(\mathbf{v})}.$$

Integrating now over all possible choices of $\mathbf{v}$ with respect to $\mathbf{P}_X$ we have

$$\Pr(\mathbf{v}' \in Y) = \int_{\mathbf{v} \in X} \Pr(\mathbf{v}' \in Y \mid \mathbf{v}) d\mathbf{P}_X.$$

From (3) it follows that $\Pr(\mathbf{v}' \in Y) = \mathbf{P}_{X_j}(Y)$ with $\mathbf{v} \sim \mathbf{P}_X$ and $t \sim \mathbf{P}_{\mathbb{T}_j(\mathbf{v})}$.

Consequently, the trajectory being generated will now be in mode q_j with $\mathbf{v}' \in X_j$ and $\mathbf{v}'$ distributed according to $\mathbf{P}_{X_j}$, compatible with the state $\eta_j = (\rho q_j, X_j, \mathbf{P}_{X_j})$ of M. Inductively it is hence guaranteed that each subsequent iteration of Algorithm 1 will produce values compatible with a path of M.

Whether the generated trajectory of length $K+1$ (and hence the corresponding path of M) is a model of ψ can be determined using a standard BLTL model checker [15]. In fact this can be done on the fly which will often avoid generating the whole trajectory. Based on this, we can test whether $\Pr_{\geq 1}(\psi)$ on M by testing the following alternative pair of hypotheses: $H_0 : \Pr_{\geq 1}(\psi)$ and $H_1 : \Pr_{<1-\delta}(\psi)$, where $0 < \delta < 1$ is a parameter chosen by the user marking the interval $[1-\delta, 1)$ as an indifference region in which accepting either hypothesis is fine. In our setting, whenever we encounter a sample (i.e. a randomly generated trajectory) that does not satisfy ψ, we can reject H_0 and accept H_1. Thus we only have to deal with false positives (when H_0 is accepted while H_1 happens to be true).

This leads to Algorithm 2 that repeatedly generates a random trajectory (using Algorithm 1), and decides after a finite number of tries between H_0 and H_1. For doing so we also fix a user-defined false positive rate α.

Algorithm 2. Sequential hypothesis test

Input: Markov chain M, BLTL property ψ, indifference parameter δ, false positive bound α.
Output: H_0 or H_1.

```
Set N := ⌈log α / log(1 − δ)⌉
for i := 1 ... N do
    Generate a random trajectory τ using Algorithm 1
    if τ, 0, ⊨^H ψ then Continue
    else return H_1
end for
return H_0
```

The accuracy of Algorithm 2 is captured by the next result.

Theorem 3. *The probability of choosing H_1 when H_0 is true (false negative) is 0. Further, suppose $N \geq \log\alpha / \log(1-\delta)$. Then the probability of choosing H_0 when H_1 is true (false positive) is no more than α.*

Proof. As observed earlier the first part is obvious. To prove the second part, if H_1 is true, then we know that $\Pr_{<1-\delta}(\psi)$. The probability of N sampled trajectories all satisfying ψ (and thus returning H_0, a false positive) is at most $(1-\delta)^N$. Therefore we have $\alpha \leq (1-\delta)^N$, leading to $N \geq \log\alpha / \log(1-\delta)$.

Hence we use $N := \lceil \log\alpha / \log(1-\delta) \rceil$ to set the sample size. For example for $\delta = 0.01$ and $\alpha = 0.01$ we get $N = 459$ while for $\delta = 0.001$ and $\alpha = 0.01$ we get $N = 4603$.

6 Case Studies

We first evaluated our method on a model of the electrical dynamics of the cardiac cell [12]. We also applied our method on a model of circadian rhythm network [28]. The Δ time step parameter for the cardiac cell model and the

circadian rhythm model were both set to 0.1. The parameters used for the statistical model checking were $\delta = 0.01$ and $\alpha = 0.01$. We have implemented our method using MATLAB. The source code is available at http://github.com/bgyori/hybrid. The experiments were carried out on a PC with a 3.4 GHz Intel Core i7 processor with 8 GB RAM. Simulating one trajectory took, on average, 5.2 s for the circadian clock model and 18.3 s for the cardiac cell model. We note that when checking quantitative properties, the trajectories that hit corner points such as $u = 1.4$ will be non-robust and hence can be ignored. Our implementation exploits the parallelization enabled by statistical model checking, hence multiple trajectories can be simulated simultaneously. A summary of the results for the verification of all properties for both models along with the number of samples taken to complete the verification is given in Table 3 of the Appendix.

In our experiments, we used $J = 10$ as the number of intermediate time steps for choosing mode transitions. We investigated whether this choice is sufficient for accurate simulation. We simulated 1000 independent realizations of the cardiac cell system with $J = 10$ and $J = 100$, and compared the distributions of the modes that the system is in at a series of discrete time points. The Kolmogorov-Smirnov statistical test did not reject the hypothesis that the two distributions are the same (at confidence level 95 %). This indicates that using $J = 10$ is adequate.

6.1 Cardiac Cell Model

Heart rhythm depends on the organized opening and closing of gates–called ion channels–on the cell membrane, which govern the electrical activity of cardiac cells. Disordered electric wave propagation in heart muscle can cause cardiac abnormalities such as *tachycardia* and *fibrillation*. The dynamics of the electrical activity of a single human ventricular cell has been modeled as a hybrid automaton [12,21] shown in Fig. 2. The model contains 4 state variables and 26 parameters. Ventricular cells consist of three subtypes, namely epicardial, endocardial, and midmyocardial cells, which possess different dynamical characteristics. The cell-type-specific parameters of the model are summarized in Table 1.

An action potential (AP) is a change in the cell's transmembrane potential u, as a response to an external stimulus (current) ϵ. The flow of total currents is controlled by a fast channel gate v and two slow gates w and s.

In mode q_0, the "Resting mode", the cell is waiting for stimulation. We assume an external stimulus ϵ equal to 1 mV lasting for 1 millisecond. The stimulation causes u to increase which may trigger a mode transition to mode q_1. In mode q_1, gate v starts closing and the decay rate of u changes. The system will jump to mode q_2 if $u > \theta_w$. In mode q_2, gate w is also closing. When $u > \theta_v$, mode q_3 can be reached, which means a successful "AP initiation". In mode q_3, u reaches its peak due to the fast opening of a sodium channel. The cardiac muscle then contracts and u starts decreasing.

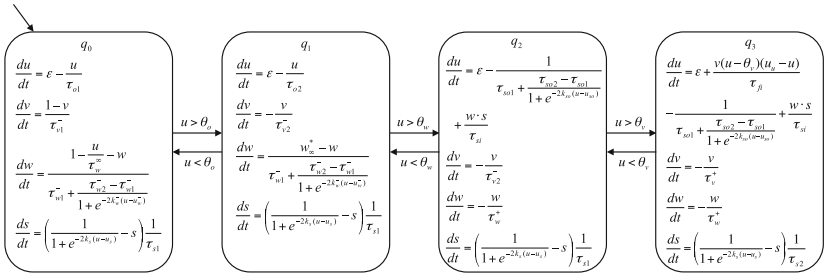


Fig. 2. The hybrid automaton model for the cardiac cell system [21].

Table 1. Parameter values of the cardiac model for epicardial (EPI), endocardial (ENDO), and midmyocardial (MID) cells under healthy condition.

Parameter	EPI	ENDO	MID	Parameter	EPI	ENDO	MID
θ_o	0.006	0.006	0.006	τ_{v1}^-	60	75	80
θ_w	0.13	0.13	0.13	τ_{v2}^-	1150	10	1.4506
θ_v	0.3	0.3	0.3	τ_{w1}^-	60	6	70
u_w^-	0.03	0.016	0.016	τ_{w2}^-	15	140	8
u_{so}	0.65	0.65	0.6	τ_{o1}	400	470	410
u_s	0.9087	0.9087	0.9087	τ_{o2}	6	6	7
u_u	1.55	1.56	1.61	τ_{so1}	30.0181	40	91
w_∞^*	0.94	0.78	0.5	τ_{so2}	0.9957	1.2	0.8
k_w^-	65	200	200	τ_{s1}	2.7342	2.7342	2.7342
k_{so}	2.0458	2	2.1	τ_{s2}	16	2	4
k_s	2.994	2.994	2.994	τ_{fi}	0.11	0.1	0.078
τ_v^+	1.4506	1.4506	1.4506	τ_{si}	1.8875	2.9013	3.3849
τ_w^+	200	280	280	$\tau_{w\infty}$	0.07	0.0273	0.01

Property C1. It is known that the cardiac cell can lose its excitability, which will lead to disorders such as ventricular tachycardia and fibrillation. We formulated the property for responding to stimulus by leaving the resting mode:

$$\mathbf{F}^{\leq 500}(\neg[\text{Resting mode}]).$$

The property was verified to be *true* for all three cell types under the healthy condition. However, under a disease condition (for example $\tau_{o1} = 0.004$ or $\tau_{o2} = 0.1$ [27]) the property was verified to be *false* no matter what stimulation value of ϵ was used. Consequently, a region of such unexcitable cells blocks the impulse conduction and can lead to cardiac disorders such as fibrillation. This is consistent with experimental results reported in [34].

Property C2. After successfully generating an AP (that is, reaching the "AP mode", q_3), the cardiac cell should return to a low transmembrane potential and

wait in "Resting mode" for the next stimulation. The corresponding formula is

$$\mathbf{F}^{\leq 500}([\text{AP mode}]) \wedge \mathbf{F}^{\leq 500}(\mathbf{G}^{\leq 100}([\text{Resting mode}])).$$

The above query was verified to be *true* for all three cell types under the healthy condition and transient stimulation. However, if we change the stimulation profile from transient to sustained, i.e. assuming ϵ lasts for 500 milliseconds, the property was verified to be *false*–the cell doesn't return to and settle at a low transmembrane potential resting state. In ventricular tissue the stimulus ϵ can be delivered from neighboring cells [12]. Thus, our results suggest that the transient activation of a single cardiac cell depends on the stimulation profile of its neighboring cells.

Property C3. It has been reported that epicardial, endocardial, and midmyocardial cells have different AP morphologies [16,29]. In particular, a crucial "spike-and-dome" (i.e. a sharp peak followed by a blunt peak) AP morphology can only be observed in epicardial cells but not endocardial and midmyocardial cells (Fig. 3).

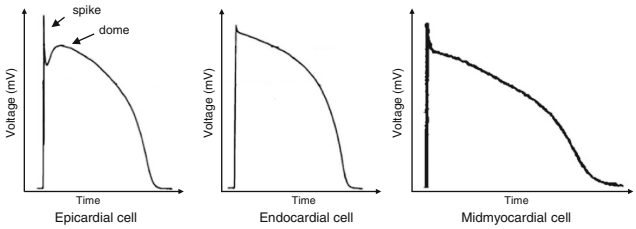


Fig. 3. The AP morphologies of epicardial [29], endocardial [29] and midmyocardial [16] cells.

We formulated the property for a spike-and-dome AP morphology as a quantitative property,

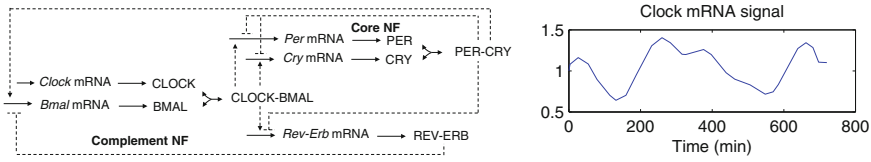
$$\mathbf{F}^{\leq 500}(\mathbf{G}^{\leq 1}([1.4 \leq u]) \wedge \mathbf{F}^{\leq 500}([0.8 \leq u] \wedge [u \leq 1.1] \wedge \mathbf{F}^{\leq 500}(\mathbf{G}^{\leq 50}([1.1 \leq u])))).$$

The property was verified to be *true* for epicardial cells and *false* for endocardial and midmyocardial cells under the healthy condition and transient stimulation. Among 26 model parameters, 20 of them have different values over different cell types. We then perturbed each epicardial parameter and checked if the above property still holds. Our results show that τ_{s2} is key to the AP morphology (i.e. the spike-and-dome AP morphology disappears when $\tau_{s2} = 2$), which highlights the importance of s gate to epicardial cells. This is consistent with [27] in that the model proposed in [17] (which does not include s gate) is unable to capture the dynamics of epicardial cells.

6.2 Circadian Rhythm Model

Mammalian cells follow a circadian rhythm with a 24 h period, which is generated and governed by a highly coupled transcription-translation network. The model diagram and the corresponding hybrid system dynamics proposed in [28,30] is described below.

The equations governing the dynamics of the circadian clock model are given in Fig. 4. The equations contain rate constants which are denoted k_1 to k_{28} and are set according to [30]. The combination of "mode indicator" binary variables θ_{CB} to $\theta_{RE}, \theta_{PC1}, \theta_{PC2}$ and θ_{PC3} define the mode of the dynamics, and each mode is defined by a unique value combination of the mode indicators. These value combinations are listed in Table 2. The guards associated with a source and target mode are constructed as follows. Each mode indicator corresponds to a guard component which is a threshold on a state variable. For instance, θ_{RE} has the corresponding guard component [REV-ERB]< 1.1. The guard to a target mode is enabled if all the mode indicators that are on in the mode are enabled according to their respective guard components. Finally, a transition between a source and a target mode only exists if there is only one difference in the combination of mode indicators. For instance, there is a transition from mode 1 to mode 2 but not from mode 1 to mode 9. The dynamics of the *Clock* mRNA is governed externally.



$$
\begin{aligned}
\mathrm{d}/\mathrm{d}t[\mathrm{Per}] &= -k_1 \cdot [\mathrm{Per}] + k_{13} \cdot \theta_{PC2} \cdot \theta_{CB} + k_{14} \\
\mathrm{d}/\mathrm{d}t[\mathrm{PER}] &= -k_2 \cdot [\mathrm{PER}] + k_{15} \cdot [\mathrm{Per}] - k_{16} \cdot [\mathrm{PER}] \cdot [\mathrm{CRY}] \\
\mathrm{d}/\mathrm{d}t[\mathrm{Cry}] &= -k_3 \cdot [\mathrm{Cry}] + k_{17} \cdot \theta_{PC2} \cdot \theta_{CB} + k_{18} \\
\mathrm{d}/\mathrm{d}t[\mathrm{CRY}] &= -k_4 \cdot [\mathrm{CRY}] + k_{19} \cdot [\mathrm{Cry}] - k_{16} \cdot [\mathrm{PER}] \cdot [\mathrm{CRY}] \\
\mathrm{d}/\mathrm{d}t[\mathrm{PER\text{-}CRY}] &= -k_5 \cdot [\mathrm{PER\text{-}CRY}] + k_{16} \cdot [\mathrm{PER}] \cdot [\mathrm{CRY}] \\
\mathrm{d}/\mathrm{d}t[\mathrm{Rev\text{-}Erb}] &= -k_6 \cdot [\mathrm{Rev\text{-}Erb}] + k_{20} \cdot \theta_{PC1} \cdot \theta_{CB} + k_{21} \\
\mathrm{d}/\mathrm{d}t[\mathrm{REV\text{-}ERB}] &= -k_7 \cdot [\mathrm{REV\text{-}ERB}] + k_{22} \cdot [\mathrm{Rev\text{-}Erb}] \\
\mathrm{d}/\mathrm{d}t[\mathrm{CLOCK}] &= -k_9 \cdot [\mathrm{CLOCK}] + k_{24} \cdot [\mathrm{Clock}] - k_{25} \cdot [\mathrm{CLOCK}] \cdot [\mathrm{BMAL}] \\
\mathrm{d}/\mathrm{d}t[\mathrm{Bmal}] &= -k_{10} \cdot [\mathrm{Bmal}] + k_{26} \cdot \theta_{PC3} \cdot \theta_{RE} + k_{27} \\
\mathrm{d}/\mathrm{d}t[\mathrm{BMAL}] &= -k_{11} \cdot [\mathrm{BMAL}] + k_{28} \cdot [\mathrm{Bmal}] - k_{25} \cdot [\mathrm{CLOCK}] \cdot [\mathrm{BMAL}] \\
\mathrm{d}/\mathrm{d}t[\mathrm{CLOCK\text{-}BMAL}] &= -k_{12} \cdot [\mathrm{CLOCK\text{-}BMAL}] + k_{25} \cdot [\mathrm{CLOCK}] \cdot [\mathrm{BMAL}]
\end{aligned}
$$

Fig. 4. The model diagram, the Clock mRNA signal and the equations governing the circadian clock model.

The system comprises 16 modes, each of which contains 12 state variables and 29 parameters. Each mode corresponds to a particular combination of ON or OFF transcriptional states of genes *Per*, *Cry*, *Rev-Erb*, *Clock*, and *Bmal*. The switches between modes are guarded by the threshold levels of protein complexes PER-CRY, CLOCK-BMAL and REV-REB. The mRNA levels of *Per* and *Cry*

Table 2. The 5 mode indicator variables and their associated guard components (top). The 16 modes of the circadian clock model with the corresponding combination of binary mode indicator variables (bottom).

Mode indicator	Guard component
θ_{RE}	[REV-ERB]< 1.1
θ_{CB}	[CLOCK-BMAL]> 1.0
θ_{PC1}	[PER-CRY]< 1.4
θ_{PC2}	1.4 <[PER-CRY]< 1.5
θ_{PC3}	2.2 <[PER-CRY]

Mode	1	2	3	4
$(\theta_{PC1}, \theta_{PC2}, \theta_{PC3}, \theta_{RE}, \theta_{CB})$	(1,1,0,1,0)	(1,1,0,1,1)	(1,1,0,0,0)	(1,1,0,0,1)
Mode	5	6	7	8
$(\theta_{PC1}, \theta_{PC2}, \theta_{PC3}, \theta_{RE}, \theta_{CB})$	(0,1,0,1,0)	(0,1,0,1,1)	(0,1,0,0,0)	(0,1,0,0,1)
Mode	9	10	11	12
$(\theta_{PC1}, \theta_{PC2}, \theta_{PC3}, \theta_{RE}, \theta_{CB})$	(0,0,0,1,0)	(0,0,0,1,1)	(0,0,0,0,0)	(0,0,0,0,1)
Mode	13	14	15	16
$(\theta_{PC1}, \theta_{PC2}, \theta_{PC3}, \theta_{RE}, \theta_{CB})$	(0,0,1,1,0)	(0,0,1,1,1)	(0,0,1,0,0)	(0,0,1,0,1)

are known to be oscillating due to the negative feedback loops in the network. Specifically, there are two major negative feedback (NF) loops: (i) the core NF formed by PER-CRY, CLOCK-BMAL, PER, and CRY and (ii) a complement NF formed by REV-ERB, BMAL, and CLOCK-BMAL. The time constants appearing in the properties are in minute units.

Property R1. Similar to *Per* and *Cry*, the expression level of *Bmal* gene is also oscillating [32]. We formulated this property as

$$\mathbf{F}^{\leq 500}([1.5 \leq Bmal] \wedge \mathbf{F}^{\leq 500}([Bmal \leq 0.8] \wedge \mathbf{F}^{\leq 500}([1.5 \leq Bmal] \wedge \mathbf{F}^{\leq 500}([Bmal \leq 0.8] \wedge \mathbf{F}^{\leq 500}([1.5 \leq Bmal])))))$$

The property was verified to be *true* under the wild type condition. It was verified to be *false* under *Cry* mutant condition but *true* in the *Rev-Erb* mutant condition, which is consistent with the experimental data in [26,32]. This suggests that the oscillatory behavior of *Bmal* mRNA is induced by the core negative feedback mediated by PER-CRY, instead of the complement negative feedback mediated by REV-ERB.

Property R2. It has been observed that the peaks of *Bmal* mRNA are always located between two successive *Per* or *Cry* mRNA peaks [26]. The corresponding formula is

$$\mathbf{F}^{\leq 500}([Bmal \leq 0.8] \wedge [2.0 \leq Per] \wedge [2.0 \leq Cry] \wedge \mathbf{F}^{\leq 500}([1.5 \leq Bmal] \wedge [Per \leq 0.8] \wedge [Cry \leq 0.8] \wedge \mathbf{F}^{\leq 500}([Bmal \leq 0.8] \wedge [2.0 \leq Per] \wedge [2.0 \leq Cry] \wedge \mathbf{F}^{\leq 500}([1.5 \leq Bmal] \wedge [Per \leq 0.8] \wedge [Cry \leq 0.8]))))$$

The above query was verified to be *true* under wild type condition. If we remove the dependence between *Bmal* transcription and PER-CRY concentration, the property R2 was verified to be *false*, while the property R1 was verified to *true* (i.e. oscillating). Thus, our results suggest that the complement negative feedback mediated by REV-ERB is responsible for maintaining the oscillatory behavior of *Bmal* mRNA level while PER-CRY plays a role in delaying the *Bmal* expression responses.

Table 3 is a summary of the performance of the SMC procedure for the hybrid systems for the two case studies presented above.

Table 3. Results summary of SMC for hybrid systems

Property	Condition	Decision	# samples
C1	Epicardial, Healthy	True	459
C1	Endocardial, Healthy	True	459
C1	Midmyocardial, Healthy	True	459
C1	Epicardial, Diseased	False	1
C1	Endocardial, Diseased	False	1
C1	Midmyocardial, Diseased	False	1
C2	Epicardial, Transient	True	459
C2	Endocardial, Transient	True	459
C2	Midmyocardial, Transient	True	459
C2	Epicardial, Sustained	False	1
C2	Endocardial, Sustained	False	1
C2	Midmyocardial, Sustained	False	1
C3	Epicardial, $\tau_{s2} = 16$	True	459
C3	Epicardial, $\tau_{s2} = 2$	False	1
C3	Endocardial	False	1
C3	Midmyocardial	False	1
R1	Wild type	True	459
R1	Cry mutant	False	1
R1	Rev-Erb mutant	True	459
R2	Wild type	True	459
R2	Without PER-CRY dependence	False	1
R1	Without PER-CRY dependence	True	459

7 Conclusion

We have presented an approximate probabilistic verification method for analyzing the dynamics of a hybrid system H in terms of a Markov chain M. For bounded time properties, we have shown a strong correspondence between the behaviors of H and M. We have also extended this result to handle quantitative atomic propositions and shown a similar correspondence result for the sub-dynamics consisting of robust trajectories. Thus the intractable verification problem for H can be solved approximately using its Markov chain approximation. Accordingly, we have devised a statistical model checking procedure to verify that M almost certainly meets a BLTL specification and then applied this

procedure to two examples to demonstrate the applicability of our approximation scheme. A hardware accelerated parallel implementation of the trajectory sampling procedure will considerably improve the performance and scalability of our method. Overall, we view our results as providing a mathematical basis for verifying if a hybrid system models satisfies a BLTL property with high probability.

As an extension, one could consider more sophisticated stochastic assumptions regarding the time points and value states at which the mode transitions take place. These assumptions will however have to be justified and motivated by the modeling problem at hand, especially in systems biology applications. Yet another valuable extension will be to study a network of hybrid systems. This will enable us to model the cross talk, feed-forward and feed-back loops involving multiple signaling pathways. A further discretization of the continuous component of the hybrid system could also be coupled with the proposed approach to reduce the complexity and increase the robustness of biological models [9].

References

1. Abate, A., Ames, A.D., Sastry, S.S.: Stochastic approximations of hybrid systems. In: ACC 2005. pp. 1557–1562 (2005)
2. Agrawal, M., Stephan, F., Thiagarajan, P.S., Yang, S.: Behavioural approximations for restricted linear differential hybrid automata. In: Hespanha, J.P., Tiwari, A. (eds.) HSCC 2006. LNCS, vol. 3927, pp. 4–18. Springer, Heidelberg (2006)
3. Alur, R., Henzinger, T.A., Lafferriere, G., Pappas, G.J.: Discrete abstractions of hybrid systems. Proc. IEEE **88**(7), 971–984 (2000)
4. Approximate probabilistic verification of hybrid systems (Technical report): http://arxiv.org/abs/1412.6953
5. Ballarini, P., Djafri, H., Duflot, M., Haddad, S., Pekergin, N.: COSMOS: a statistical model checker for the hybrid automata stochastic logic. In: QEST 2011. pp. 143–144 (2011)
6. Batt, G., Ropers, D., de Jong, H., Geiselmann, J., Page, M., Schneider, D.: Qualitative analysis and verification of hybrid models of genetic regulatory networks: nutritional stress response in *Escherichia coli*. In: Morari, M., Thiele, L. (eds.) Hybrid Systems: Computation and Control. LNCS, vol. 3414, pp. 134–150. Springer, Heidelberg (2005)
7. Biere, A., Cimatti, A., Clarke, E., Zhu, Y.: Symbolic model checking without BDDs. In: Cleaveland, W.R. (ed.) TACAS 1999. LNCS, vol. 1579, pp. 193–207. Springer, Heidelberg (1999)
8. Blom, H.A., Lygeros, J., Everdij, M., Loizou, S., Kyriakopoulos, K.: Stochastic Hybrid Systems: Theory and Safety Critical Applications. Springer, Heidelberg (2006)
9. Bortolussi, L., Policriti, A.: The importance of being (a little bit) discrete. ENTCS **229**(1), 75–92 (2009)
10. Bruce, D., Pathmanathan, P., Whiteley, J.P.: Modelling the effect of gap junctions on tissue-level cardiac electrophysiology. Bull. Math. Biol. **76**(2), 431–454 (2014)
11. Buckwar, E., Riedler, M.G.: An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. J. Math. Biol. **63**(6), 1051–1093 (2011)

12. Bueno-Orovio, A., Cherry, E.M., Fenton, F.H.: Minimal model for human ventricular action potentials in tissue. J. Theor. Biol. **253**, 544–560 (2008)
13. Cassandras, C.G., Lygeros, J.: Stochastic Hybrid Systems. CRC Press, Taylor & Francis Group (2006)
14. Clarke, E., Fehnker, A., Han, Z., Krogh, B.H., Stursberg, O., Theobald, M.: Verification of hybrid systems based on counterexample-guided abstraction refinement. In: Garavel, H., Hatcliff, J. (eds.) TACAS 2003. LNCS, vol. 2619, pp. 192–207. Springer, Heidelberg (2003)
15. Clarke, E.M., Grumberg, O., Peled, D.A.: Model Checking. MIT Press, Cambridge (1999)
16. Drouin, E., Charpentier, F., Gauthier, C., Laurent, K., Le Marec, H.: Electrophysiologic characteristics of cells spanning the left ventricular wall of human heart: evidence for presence of m cells. J Am Coll Cardiol **26**, 185–192 (1995)
17. Fenton, F., Karma, A.: Vortex dynamics in 3D continuous myocardium with fiber rotation: filament instability and fibrillation. Chaos **8**, 20–47 (1998)
18. Frehse, G.: PHAVer: algorithmic verification of hybrid systems past hytech. In: Morari, M., Thiele, L. (eds.) HSCC 2005. LNCS, vol. 3414, pp. 258–273. Springer, Heidelberg (2005)
19. Gao, S., Kong, S., Clarke, E.: Delta-complete reachability analysis (part i). In: Technical report, CMU SCS, CMU-CS-13-131 (2013)
20. Girard, A., Le Guernic, C., Maler, O.: Efficient computation of reachable sets of linear time-invariant systems with inputs. In: Hespanha, J.P., Tiwari, A. (eds.) HSCC 2006. LNCS, vol. 3927, pp. 257–271. Springer, Heidelberg (2006)
21. Grosu, R., Batt, G., Fenton, F.H., Glimm, J., Le Guernic, C., Smolka, S.A., Bartocci, E.: From cardiac cells to genetic regulatory networks. In: Gopalakrishnan, G., Qadeer, S. (eds.) CAV 2011. LNCS, vol. 6806, pp. 396–411. Springer, Heidelberg (2011)
22. Henzinger, T.: The theory of hybrid automata. In: LICS 1996. pp. 278–292 (1996)
23. Henzinger, T., Kopke, P.: Discrete-time control for rectangular hybrid automata. Theor. Comput. Sci. **221**(1), 369–392 (1999)
24. Hirsch, M., Smale, S., Devaney, R.: Differential Equations, Dynamical Systems, and an Introduction to Chaos. Academic Press, Waltham (2012)
25. Julius, A.A., Pappas, G.J.: Approximations of stochastic hybrid systems. IEEE Trans. Autom. Control **54**(6), 1193–1203 (2009)
26. Kim, J.K., Forger, D.B.: A mechanism for robust circadian timekeeping via stoichiometric balance. Mol Syst Biol **8**, 1–14 (2012)
27. Liu, B., Kong, S., Gao, S., Zuliani, P., Clarke, E.M.: Parameter synthesis for cardiac cell hybrid models using δ-decisions. In: Mendes, P., Dada, J.O., Smallbone, K. (eds.) CMSB 2014. LNCS, vol. 8859, pp. 99–113. Springer, Heidelberg (2014)
28. Matsuno, H., Inouye, S.T., Okitsu, Y., Fujii, Y., Miyano, S.: A new regulatory interaction suggested by simulations for circadian genetic control mechanism in mammals. J Bioinf. Comput Biol **4**(1), 139–153 (2006)
29. Nabauer, M., Beuckelmann, D.J., Uberfuhr, P., Steinbeck, G.: Regional differences in current density and rate-dependent properties of the transient outward current in subepicardial and subendocardial myocytes of human left ventricle. Circulation **93**, 169–177 (1996)
30. Nakamura, K., Yoshida, R., Nagasaki, M., Miyano, S., Higuchi, T.: Parameter estimation of in silico biological pathways with particle filtering towards a petascale computing. In: PSB 2009. pp. 227–238 (2009)

31. Palaniappan, S.K., Gyori, B.M., Liu, B., Hsu, D., Thiagarajan, P.S.: Statistical model checking based calibration and analysis of bio-pathway models. In: Gupta, A., Henzinger, T.A. (eds.) CMSB 2013. LNCS, vol. 8130, pp. 120–134. Springer, Heidelberg (2013)
32. Shearman, L., Sriram, S., Weaver, D., Maywood, E., Chaves, I., Zheng, B., Kume, K., Lee, C., van der Horst, G., Hastings, M., Reppert, S.: Interacting molecular loops in the mammalian circadian clock. Science **288**, 1013–1019 (2000)
33. Stephen, W.: General Topology. Dover Publications, Addison-Wesley Publishing Company, Reading, Massachusetts (1970)
34. Tanaka, K., Zlochiver, S., Vikstrom, K., Yamazaki, M., Moreno, J., Klos, M., Zaitsev, A., Vaidyanathan, R., Auerbach, D., Landas, S., Guiraudon, G., Jalife, J., Berenfeld, O., Kalifa, J.: Spatial distribution of fibrosis governs fibrillation wave dynamics in the posterior left atrium during heart failure. Circ. Res. **8**(101), 839–847 (2007)