Reducing run-times of excitable cell models by replacing computationally expensive functions with splines

Michael Clerx and Pieter Collins

Abstract—Numerical simulation of muscle cells and tissue is an established tool in cardiac electrophysiology, where the electrical behavior of excitable heart muscle cells is commonly modeled as a stiff, non-linear system of ordinary differential equations. A common feature of this system’s right-hand side is the heavy use of computationally expensive univariate functions of the membrane potential. In this article, we investigate the performance benefits of replacing these functions with cubic spline approximations in an automated model simplification process. Clear performance gains were found when evaluating the right-hand side in isolation and when performing multi-cellular simulations using a simple forward Euler method. Single cell simulations run with an adaptive method saw smaller gains due to a higher overhead from the solver. A parallel multi-cellular simulation was also investigated, but the overhead of the implementation overshadowed the evaluation time of the right-hand side.

Index Terms—Biological Systems; Numerical and Symbolic; Systems Biology

I. INTRODUCTION

Numerical simulations based on models of ion channels and transporters are an established tool in cellular cardiac electrophysiology [1, 2]. This field, which studies the electrical behavior of heart muscle cells (myocytes) and its relation to muscle contraction, can deepen our knowledge of the fundamental processes underlying our heart beats and provides insights into life-threatening conditions of arrhythmia. The main components of these models: ion concentrations, channels, transporters and fluxes can be measured in isolation and described using ordinary differential equations (ODEs). Combining these components into cell models allows the effects of molecular differences to be investigated on a cellular level. For example, the effects of genetic mutations on channel proteins can be measured in vitro, leading to an updated ion channel model that can then be incorporated into existing models of the cell. By coupling adjacent cells through a current invoked by the difference in their potentials, multi-cellular models of muscle fibers or patches of tissue can be created.

Since their introduction in the last half of the 20th century, ODE models of myocytes have become increasingly complex. From a computational efficiency point of view modern models contain many expensive-to-evaluate functions, especially univariate functions of the membrane potential $V$ making heavy use of the exponential function [3, 4, 5, 6]. As more is learned about the complexity of the mammalian heart, the number of equations used to model a single cell is growing. This conflicts with the need to run simulations fast and the desire to create personalized models of the whole human heart, which contains roughly $5 \cdot 10^9$ of such cells [7].

In this paper we investigated the performance gains to be had from replacing computationally expensive parts of the right-hand-side (RHS) of the differential equations with easier-to-evaluate approximations. Numerical experiments were run to show the effects of simplifying with cubic spline approximations on the RHS evaluation times. We examined the influence of using a simplified RHS in three scenarios: single cell simulations using an advanced numerical solver, multi-cell simulations without parallelization, and multi-cell simulations run in parallel on a graphics processing unit (GPU).

To the best of our knowledge, no similar attempts have been reported. Although Mirin et al. [8] report using rational function approximates for the same purpose as part of a larger effort to simplify the model by ten Tusscher et al. [9], our approach differs in two ways: Firstly, in our approach we use splines, which are more stable and flexible when approximating functions with strong localized variations and may evaluate faster, and secondly, instead of simplifying one specific model, we have described and implemented a generic approach.

II. PRELIMINARY CONSIDERATIONS

A cell model’s RHS can be broken down into two parts:

\[
\frac{dV}{dt} = -\frac{1}{C} [I_{\text{ion}}(V, u) + I_{\text{stim}}]
\]

\[
\frac{du}{dt} = f(V, u)
\]

(see also [10]). Here, $V$ is the cell’s membrane potential, $C$ is the membrane capacitance and the vector $u$ is a collection of channel protein states, ion concentrations, and sometimes other variables such as fractions of bound enzymes. From
these variables a set of ionic currents is calculated as $I_i = g_i(V, u) (V - E_i)$ where $E_i$ is an equilibrium potential for the ion type and $g_i(V, u)$ is a model for the ion channel’s conductance. An external stimulus $I_{\text{stim}}$ is periodically applied to drive the system. It should be noted that $f(V, u)$ is non-linear and typically stiff. In electrophysiological terms, it contains slow and fast currents. In all models, the fast sodium current rises and falls in a time span of $< 5$ ms, whereas other currents, for example the slow delayed rectifier potassium current take several seconds to reach their peak. For a classic example of a cardiac myocyte model and the accompanying equations see [11] or [12], a good example of the complexity of modern models is found in the detailed appendix to [3]. An example of a reduced cell model is given in this paper as Table I. These equations describe the Morris-Lecar model [13], a reduced form of the original neuronal model by Hodgkin and Huxley [14] which contains only two state variables. In this example, the external input forcing the system is given as the dimensionless value pacing.

A. Single cell simulations

For detailed modern single-cell models the RHS is a costly function to evaluate and the system of equations is stiff. As a result, higher order adaptive schemes that require multiple RHS evaluations or suffer from reduced stability fail to produce any performance benefits. Implicit methods offer greater stability and therefore bigger step sizes, but the need to approximate the next solution often involves more RHS evaluations which can counteract this benefit. As a result, the explicit forward Euler method has remained a competitive choice. A difficult to implement, but much faster approach is to use an implicit adaptive multi-step method, e.g. CVODE [15]. The multi-step approach means only a limited number of RHS evaluations per time step are required; in most cases a single evaluation per step is sufficient (see Figure 2, part D). Myocyte models are typically paced (forced) periodically with a block wave stimulus, which introduces two discontinuities per cycle. At these points, it is necessary to reset the simulation routine, leading to a higher number of steps around those points.

B. Multi-cell simulations

Single cell models can be used for multi-cellular simulations by simply duplicating the single cell state vector. Without any interaction between the cells, the system’s Jacobian takes on a diagonal block structure. Connections between the cells are then introduced in the form of a diffusion current. For each cell $i$:

$$I_{\text{diff},i} = \sum_{j} g_{ij} (V_i - V_j)$$

Where the sum is over all neighboring cells and the cell-to-cell conduction $g_{ij}$ is assumed constant. This diffusion current connects all membrane potentials to that of their neighbors, but leaves the term $\frac{dV_i}{dt} = f(V, u)$ untouched. The system now becomes:

$$\frac{dV_i}{dt} = -\frac{1}{C} [I_{\text{ion}} (V_i, u_i) + I_{\text{stim},i} + I_{\text{diff},i}]$$

$$\frac{du_i}{dt} = f (V_i, u_i)$$

The most common strategy for running multi-cell simulations is the explicit forward Euler method. There are a number of reasons for this. First of all, in a multi-cell simulation, the evaluation time of the duplicated RHS becomes extremely high, so that methods requiring multiple evaluations all suffer from poor performance. Secondly, in the single cell case the system’s fast dynamics are highly localized in time (about 5ms per 1000ms for a heart rate of 60bmp) and occur at the onset of excitation. In a multi-cell scenario each cell’s excitation triggers its neighbors’ excitation with a small delay, causing the fast dynamics (and the need for a small step size) to be spread out over a much larger time. In extreme cases (which are often of high scientific interest) some part of the tissue is undergoing rapid changes at any given time. This limits the utility of adaptive methods. Finally, when performing parallelized multi-cellular simulations on GPU devices the cost of memory access and synchronization is such that a simple forward Euler scheme is

<table>
<thead>
<tr>
<th>Table I</th>
<th>THE REDUCED MORRIS-LECAR MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$dV \over dt = -\frac{1}{C} [I_{\text{ion}} + I_{\text{stim}}]$</td>
<td>$dw \over dt = 0.04 \frac{w_{\text{inf}} - w}{w_{\text{tau}}}$</td>
</tr>
<tr>
<td>$I_{\text{ion}} = I_K + I_{\text{Ca}} + I_{\text{leak}}$</td>
<td>$I_{\text{Ca}} = g_{\text{Ca}} (V - E_{\text{Ca}})$</td>
</tr>
<tr>
<td>$E_K = -84$</td>
<td>$E_{\text{Ca}} = 120$</td>
</tr>
<tr>
<td>$g_K = 8w$</td>
<td>$g_{\text{Ca}} = 4.4m_{\text{inf}}$</td>
</tr>
<tr>
<td>$w_{\text{inf}} = \frac{1}{2} [1 + \tanh \left( \frac{V - 2}{\theta} \right) ]$</td>
<td>$m_{\text{inf}} = \frac{1}{2} [1 + \tanh \left( \frac{V + 1.2}{18} \right) ]$</td>
</tr>
<tr>
<td>$w_{\text{tau}} = 1/ \cosh \left( \frac{V - 2}{\theta} \right)$</td>
<td>$I_{\text{stim}} = -80 \times \text{pacing}$</td>
</tr>
<tr>
<td>$W(t = 0) = 0.015$</td>
<td>$V(t = 0) = 60.86$</td>
</tr>
</tbody>
</table>

An important feature of these models is that they contain many computationally expensive functions depending only on the membrane potential $V$. We automatically identify these functions and replace them by spline approximations that are faster to evaluate. Most of these functions are found in ion channel models, which are a phenomenological class of models so that no valuable information is lost by altering the form of the expression. Furthermore, the precision of cell models is limited, and, as we illustrate in the next two sections, the stability of the ODE integrator takes precedence over the accuracy of the evaluation of the RHS.

Footnotes:

1In fact, this is Ohm’s law using conductance $g = 1/R$ instead of resistance $R$.

2The Jacobian of a cell model is typically not available analytically. For some good examples see [3, 4, 5].

3This is a simplification, as the diffusion current must be carried by one or more species of ion. Taking this into account, any ionic concentrations in the term $u$ would also be affected by the $I_{\text{diff}}$.  

---

---
almost always faster than a more delicate approach requiring the storage of multiple state or derivative vectors.

Garcia et al. [16] report experimenting with different adaptive Runge-Kutta pairs before selecting the simplest tested method: an explicit Euler / Heun pair. They report a 25% speed-up over fixed size forward Euler. Unfortunately, little details are given and we were unable to reproduce this result.

C. Software & model implementations

All simulations and benchmarks were carried out using the Myokit modeling framework.\(^4\) Models used were either implemented following the equations published in source files and appendices or imported from the CellML model repository [17]. Myokit is implemented in Python, but uses code generation and on-the-fly compilation of C-extensions to achieve high performance. Simulations are run in three stages: first a model file is read and parsed into a structured, symbolic representation of the model’s equations. This symbolic model is then passed on to a simulation class that uses it to generate, compile and link to a custom simulation object. Finally, the experiment is run and a pointer to the results is passed back to the caller. An advantage of this method is that it allows symbolic manipulation of the model’s equations before simulation, without the drawbacks of actually running simulations using the symbolic types.

III. RESULTS

A. Spline fitting procedure

A cubic spline function \( g(V) \) was fit to each selected function \( f(V) \) on the interval \( I = [-100\text{mV}, 150\text{mV}] \). Spline coefficients were set by imposing \( \frac{df}{dV} = \frac{dg}{dV} \) at the endpoints of \( I \), setting \( g(x_i) = f(x_i) \) at each knot \( x_i \) and then solving the resulting linear system. The absolute and relative error in the fit were estimated by sampling both functions at 1000 evenly spaced points on the selected interval and calculating

\[
\epsilon_{\text{abs}} = \frac{\text{max} |f - g|}{\epsilon_{\text{abs}}}
\]

\[
\epsilon_{\text{rel}} = \frac{\epsilon_{\text{abs}}}{\text{max} f - \text{min} f}
\]

over the full sample. The tolerance was set to \( \epsilon_{\text{abs}} \leq 10^{-3} \) and \( \epsilon_{\text{rel}} \leq 10^{-5} \). Whenever these limits were exceeded an extra knot was added at the point with the highest error. If a set maximum number of pieces (100) was reached the procedure was halted and an error returned, marking the function as unfit for approximation.

To show the performance advantage of spline approximations, cubic spline approximations using 32-piece\(^5\) splines were created for the following functions:

\[
\begin{align*}
    f_0(V) & = \exp(V/100) \\
    f_1(V) & = \exp(0.01V) \\
    f_2(V) & = 1/[1 + \exp((V + 40)/-10)] \\
    f_3(V) & = 1/[7 \exp((V + 12)/35) + 9 \exp(-(V + 77)/6)] \\
    f_4(V) & = 1/[1 + \exp(-0.1(V + 40))] \\
    f_5(V) & = 1/[7 \exp(0.03(V + 12)) + 9 \exp(-0.2(V + 77))]
\end{align*}
\]

The first function is a minimal example of an \( \exp \) function, scaled to the range \( I \). The second function shows the performance benefit of multiplication over division, compared to the run-time of an \( \exp \) evaluation. The third and fourth equations are adapted from the fast sodium channel formulation of [4] and are typical examples of the kind of equation we hoped to simplify: The first, \( f_2 \) represents the steady-state value of one of the state variables, \( m \), and defines a “hat” or “bell” shape over \( I \). The final two equations show the effect of replacing division operations with multiplication in \( f_2 \) and \( f_3 \).

The performance gains are shown in Table II. As can be seen, the performance of the splines is invariant with respect to the run-time of the original function, leading to large speed-ups for complex expressions. This illustrates the potential advantage of spline approximations in the RHS.

B. Isolating expressions for approximation

A symbolic version of the model’s equations was obtained and scanned for univariate functions of \( V \). Care was taken to include functions both directly dependent only on \( V \) and directly dependent on other functions themselves a function of only \( V \). For the selected functions, a spline was calculated using the methods described earlier. A piecewise polynomial with the calculated coefficients was then inserted into the symbolic model representation and finally the model was exported to standard C or OpenCL.\(^6\)

To estimate the maximum speed gain from this method, the model’s RHS was evaluated either in full or without the selected equations. The ratio between run-times, \( T_{\text{null}} = T_{\text{RHS, null}} / T_{\text{RHS}} \), is an estimate of fraction of the RHS evaluation time unaffected by the selected expressions. This estimate serves as a baseline value to compare the

<table>
<thead>
<tr>
<th>Table II: PERFORMANCE GAINS FOR SINGLE FUNCTION APPROXIMATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>( f_0 )</td>
</tr>
<tr>
<td>( f_1 )</td>
</tr>
<tr>
<td>( f_2 )</td>
</tr>
<tr>
<td>( f_3 )</td>
</tr>
<tr>
<td>( f_4 )</td>
</tr>
<tr>
<td>( f_5 )</td>
</tr>
</tbody>
</table>

\(^4\)Myokit, or the Maastricht Myocyte Toolkit, is an open-source Python-based software package designed to simplify the use of numerical models in the analysis of cardiac myocytes developed at Maastricht University. Source files and documentation are available from http://myokit.org

\(^5\)Splines with 32 pieces provided accurate fits in each of these examples. As can be seen in part B of Figure 1, the evaluation time of a full spline depends only weakly on the number of pieces used and using, for example, a 64-piece spline gives roughly the same results.

\(^6\)OpenCL™ (Open Computing Language) is a standard for parallel programming available from http://www.khronos.org/opencl/
actual performance $F_{\text{splines}}$ against. The results are shown in Table III.

In this table, $N_s$ and $N_c$ are the number of state variables and equations in each model respectively, while $N_r$ is the number of those equations that could be simplified. $F_{\text{null}}$ is an estimate of the best performance and $F_{\text{splines}}$ is the actual performance of the RHS evaluation time reduction measured as $F_{\text{splines}} = T_{\text{RHS,splines}} / T_{\text{RHS}}$. The run time of an unoptimized one-second single cell simulation is given as $T_0$ and the performance using splines is given as $F_0 = T_{0,\text{splines}} / T_0$. The number of RHS evaluations made by the adaptive method in these simulations is given as $M$ for the unaltered RHS and $M_{\text{splines}}$ for the version using splines. Finally, $F_1 = T_{1,\text{splines}} / T_1$ represents the speed-up gained in an unparallelized fixed step size cable simulation. Note that the smaller the $F$-value, the larger the performance gain using splines. The time taken to perform the spline approximations was not included in any of the benchmarks.

C. Accuracy & performance

Experiments were run to asses the trade-off between spline accuracy and run-time using the reduced Morris-Lecar model for excitable cells (see Table I). The expressions for $w_{\text{inf}}$, $w_{\text{tau}}$ and $m_{\text{inf}}$ were each selected for simplification: spline approximations were generated for each with an increasing number of segments (from 2 to 200). For each generated spline, a single cell simulation was run using CVODE. All simulations were started from a common set of initial values.\footnote{A similar test was run where models were “pre-paced” for $10^5$ periods before the steps taken were measured, but no significant difference with the pre-pacing free case was found.} Results are shown in Figure 1.

As can be seen, increasing the number of pieces reduces the error of fit almost without affecting the RHS evaluation time, which shows a reduction to about 13% of the evaluation time of non-optimized case. The difference between state variable $w$ with and without the approximation was tracked and seen to stabilize around 10 pieces. Similarly, the number of steps taken by the adaptive solving scheme and the total number of RHS evaluations performed in a simulation reach a stable average at about 5 pieces.

From these results we concluded that (1) Increasing the number of pieces in a spline below the minimum number required to achieve acceptable accuracy doesn’t dramatically affect RHS evaluation time. As a result, no special strategy to keep the number of pieces in a spline to a minimum is required. (2) Using spline approximations increases the number of steps taken by the adaptive solver, but the increase does not seem affected by the number of pieces. (3) The accuracy of the produced result increases dramatically as spline quality increases, but then stabilizes once the maximum relative error of fit has gone below $10^{-3}$.

We speculate that the unusually high and seemingly random variation in the number of RHS evaluations and integrator steps was caused by the higher order discontinuities at the spline knots. This is briefly discussed in Section III-E.

D. RHS Evaluation times

Right-hand-side evaluation times were measured using the following procedure: A simulation was run and the state vector was saved for every position. Next, a benchmarking tool revisited each state multiple times and measured the average execution time per call. The results show that employing spline approximation decreased the RHS evaluation time in each case. However, the utility of function approximation is seen to decrease with increasing model complexity. This can be attributed to two factors: the inclusion of more multivariate equations (for example signaling or intracellular diffusion processes) and the inclusion of more fast-to-evaluate equations (for example Markov formulations of ion channels).

For small models, the performance of the spline-based RHS sometimes exceeds our calculated estimate. This can be attributed to the difficulty of measuring small run times and the many optimizations employed by modern compilers and hardware, which make it difficult to get consistent benchmarking results when comparing small run times.

E. Single cell simulation

Single cell simulations were run using a simulation built on Myokit’s Simulation class, which creates, compiles and executes C code using CVODE to integrate the ODE.
results show a smaller speed-up than expected on the basis of
the RHS performance alone. We initially hypothesized that
this may be due to the higher order discontinuities created at
the knots, but this would suggest a strong relation between
the number of pieces used and the number of steps taken.
As can be seen in Table III no such relation was found.
Tentative experiments using splines with higher orders of
continuity failed to show a consistent performance gain in
terms of number of steps. By comparing the number of
RHS evaluations taken in simulations with and without spline
approximations (given in Table III as $M$ and $M_{Splines}$) it
can be seen that even with a reduced number of evaluations
the performance gain is not proportional to that in the RHS
evaluation times. This suggests overhead of the method is
a bigger factor in these simulations and the impact of RHS
run-time reduction is reduced.

To gain insight into the adaptive behavior of the CVODE
solver, we implemented a tracking mode that logs the number
of evaluations taken and the system time at each step of
a simulation. Figure 2 shows the results for a 1000ms
simulation using the model by Decker et al. [19]. As can be
seen from the top half of the figure, most of the work done
during a simulation occurs during the action potential (AP),
the period when the membrane potential is elevated from its
usual resting state of around $-80mV$. This example, which
is typical for ventricular myocyte models, shows how most of
the system’s fast dynamics are localized around the start
and finish of the AP. The lower half of the figure clearly shows
why this method is efficient in terms of RHS evaluations: the
most common number of evaluations needed at each step is 1
with occasional short burst of around 50 evaluations, leading
to an average of 3.1 evaluations per step throughout the
simulation. This average ranges from 1.6 to 6.0 evaluations
per step for the models described in Table III.

F. Multi-cell simulation

Non-parallel multi-cellular simulations were run using
Myokit’s Simulation1D class, which uses an explicit forward
Euler method to integrate the ODE. As described in Sec-
tion II-B, this is often the best approach for multi-cellular
simulations. Without a step-size choosing algorithm or other
complicating factors, it can be expected that the performance
boost in this scenario matches that seen in single RHS
evaluation times. A small number of tests were run which
clearly confirmed this notion so non-parallel data was not
included in Table III.

An OpenCL-based parallel version of this simulation was
implemented, but was unable to produce any speed-ups on
a GPU. Indeed, we found that eliminating the simplified
calculations from the RHS entirely only lead to marginal
speed-ups, indicating that the RHS evaluation time is of
little importance in our parallel implementation. We suspect
that the overhead of memory access and synchronization in
this case is a bottleneck in our calculations. Further work
is needed to reduce this overhead before the benefits of
reduced RHS evaluation time for parallel cell simulation can
be judged.

IV. Conclusions & Further Work

Splines are a natural candidate for providing simplified
approximations to univariate functions occurring in the right-

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|}
\hline
Model & $N_x$ & $N_y$ & $N_z$ & $F_{null}$ & $F_{Splines}$ & $T_0$ & $F_0$ & $M$ & $M_{Splines}$ & $F_1$ \\
\hline
2009, Tran [18] & 4 & 24 & 11 & 28% & 17% & 0.5ms & 69% & 359 & 350 & - \\
2009, Stewart [6] & 20 & 106 & 38 & 50% & 48% & 3.7ms & 80% & 942 & 991 & 51% \\
2011, O’Hara [4] & 41 & 249 & 67 & 53% & 64% & 1.2ms & 98% & 1483 & 2586 & - \\
2009, Decker [19] & 48 & 204 & 48 & 74% & 72% & 1.3ms & 88% & 1646 & 1594 & 73% \\
2010, Sampson [20] & 81 & 299 & 107 & 39% & 85% & 38ms & 97% & 3277 & 3209 & 78% \\
2011, Heijman [3] & 145 & 514 & 61 & 80% & 88% & 8.3ms & 98% & 4165 & 4071 & 89% \\
\hline
\end{tabular}
\caption{Model run times and performance}
\end{table}
hand side of a differential equation model of excitable cells. In this paper, we showed that the use of cubic spline approximations can reduce the time needed to evaluate the RHS of an excitable cell model’s ODE. Further, the non-smoothness of the approximation did not appear to negatively affect the number of RHS evaluations needed by the sophisticated implicit multi-step method used by the package CVODE.

The time benefit of this optimization was seen directly in unparallelized multi-cell simulations, but failed to materialize fully in single cell simulations using CVODE, and had no effect on parallel multi-cell simulations using the forward Euler method run on the GPU. We suggest that the RHS evaluation time is overshadowed by overhead resulting from the solver in the former case, and from overhead of memory access due to the hardware architecture in the latter.

Given the speedups obtained in the evaluation of the RHS, we believe that this method may still be useful in situations where RHS evaluation time is a dominant factor, either in the field of cellular electrophysiology or in other domains. Simplification of multivariate functions using splines is possible, but significantly more complicated. Further work is needed to optimize our parallel solver to the point that the benefits of RHS simplification become visible. Work is in progress on computing optimal spline approximations of various degrees in the uniform norm.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Enno de Lange for our valuable discussions on cell models, model implementations and writing for the GPU.

REFERENCES


