An inner-outer subcycling algorithm for parallel cardiac electrophysiology simulations

Sebastian Laudenschlager | Xiao-Chuan Cai

Abstract
This paper explores cardiac electrophysiological simulations of the monodomain equations and introduces a novel subcycling time integration algorithm to exploit the structure of the ionic model. The aim of this work is to improve upon the efficiency of parallel cardiac monodomain simulations by using our subcycling algorithm in the computation of the ionic model to handle the local sharp changes of the solution. This will reduce the turnaround time for the simulation of basic cardiac electrical function on both idealized and patient-specific geometry. Numerical experiments show that the proposed approach is accurate and also has close to linear parallel scalability on a computer with more than 1000 processor cores. Ultimately, the reduction in simulation time can be beneficial in clinical applications, where multiple simulations are often required to tune a model to match clinical measurements.

Keywords
cardiac electrophysiology, finite element on unstructured meshes, parallel processing, patient-specific cardiac geometry, time integration with subcycling

1 | INTRODUCTION
The prevalence of cardiovascular disease gives cardiac research a sense of great importance, because advancements in cardiac research could ultimately lead to a reduction in deaths related to cardiovascular disease. One crucial area of cardiac research is cardiac electrophysiology, which investigates the electrical propagation that drives a heartbeat in the human heart. In a clinical setting, this can be studied, for example, with magnetic resonance imaging (MRI) and electrocardiograms (ECG) in order to monitor and detect irregularities in the electrical activity. A different approach is to simulate the electrical activity in the heart using cardiac models. Cardiac electrophysiology simulations at the tissue level can often be modeled by either the bidomain or the monodomain equations, and incorporate a cellular model to deal with the intracellular currents. In this paper, we utilize the monodomain equations. This is a tradeoff between complexity and accuracy, as the bidomain equations more accurately describe cardiac electrical activity, but are also more expensive to solve computationally.

The monodomain equations are solved via the finite element method (FEM) on a variety of meshes, ranging from simple slabs for benchmarking to meshes generated from actual MRI data. In order to generate patient-specific meshes from MRI data, we developed a semi-automatic mesh generation pipeline, which utilizes user-guided segmentation and automatic model creation from segmentation. This pipeline uses a variety of open source software, and allows for flexibility in mesh generation, including mesh size and quality.
With the semi-implicit discretization, a large sparse linear system of equations needs to be solved at every time step, which is the most time consuming part of the simulation. We solve the system with a conjugate gradient method preconditioned by a parallel algebraic multigrid. The actual software is a modified version of the FEM variant of the Cardioid package. It uses the FEM library MFEM and the solver/preconditioning library Hypre. Using MFEM allows us to abstract the details of FEM, and instead focus on the higher level details of the problem. Furthermore, the Hypre library gives us access to various high performance solvers and preconditioners, giving us flexibility to test various parameters to find the optimal combination for our problem. In addition to MFEM and Hypre, a variety of Python scripts were developed for post-processing simulation data, as well as for visualization and data analysis.

An important aspect of our solver is parallelization, meaning simply that we can use more processors in parallel to solve the problem much more efficiently. This is crucial because it reduces the turnaround time for cardiac electrophysiology simulations, which typically require very fine resolutions, thus allowing more simulations to be run in a shorter time frame. Many strategies and algorithms have been explored in parallel cardiac electrophysiology simulations to improve the efficiency of solving the bidomain or monodomain equations, including techniques like parallel preconditioning and fully implicit solvers. The parallelization used for our work shares many aspects of existing parallel cardiac electrophysiology codes such as Chaste, e.g. using a parallel FEM solver with preconditioning. We tested the parallel scalability of the Cardioid code on both simplified and realistic geometries, and found excellent scalability up to 1000 processors, similar to existing approaches. Beyond the scalability of the code, we also performed a verification study of our simulation results. To that end, we compared activation times with those of the Niederer benchmark across a variety of mesh sizes and timesteps. The results indicate good spatial and temporal convergence, and compare well with other implementations of monodomain solvers.

Due to the inherent multiscale nature of cardiac electrophysiology simulations, numerous previous studies have been conducted to improve the computational efficiency of simulations through spatial and temporal adaptivity. Spatial adaptivity through Adaptive Mesh Refinement (AMR) has been used on both the bidomain and monodomain equations, with modest observed speedups as a result of balancing the speedup with the cost of performing the mesh refinement. Temporal adaptivity has also been studied in the context of cardiac electrophysiology simulations, including simple global timestep adaptivity, applying full adaptivity to both the reactive and diffusive terms, and even uniformization methods based on Markov chains. These time adaptive methods are globally applied, meaning a residual estimator is needed to determine when to refine the timestep and how much the refinement should be.

In this paper, we instead focus on locally refining the timestep based on the underlying dynamics of ionic model, similar to the locality found in spatial adaptivity. Thus, a residual estimator is not necessary since we know in advance that, for the human heart, the solution changes sharply during the depolarization phase. The main goal of this paper is to improve the turnaround time for a simulation by introducing a novel subcycling algorithm for timesteping in the ionic model. Instead of using the existing timestepping method, where a global fixed timestep is used for both the nonlinear parabolic partial differential equation (PDE) for the electrical conduction and the system of ordinary differential equations (ODE) for the cellular ionic currents, we use several iterations of smaller timesteps in the ionic model for the sharp depolarization phase. This allows us to use a larger timestep for the PDE, as well as for the ionic model ODEs during the repolarization phase. For different ionic components, the length of the sharp change periods are slightly different and our selection is based on the longest component. Through the use of subcycling methods, we can accurately utilize timesteps of 0.1 ms, instead of the small timestep sizes of 0.01–0.05 ms used in the benchmark tests.

The rest of this paper is organized as follows. In Section 2, we present the continuous and discretized models for tissue-level electrical propagation and the cellular ionic model, as well as the mesh generation pipeline. In Section 3, we present the subcycling method in the ionic model computation, followed by numerical results and associated discussion in Section 4. Finally, in Section 5, we wrap up with some concluding remarks.

## 2 MATHEMATICAL MODELS

The cardiac electrical propagation described by the monodomain equations, based on the pioneering Hodgkin-Huxley model, consists of two main components, the PDEs for electrical conduction and the ODEs for cellular ionic currents. We first present the monodomain equations, including the discretized form using FEM, followed by the ionic model.
2.1 | Monodomain equations

The monodomain equations describing the propagation of transmembrane potential \( V(x, t) \) on \( \Omega \subset \mathbb{R}^3 \) and \( t \in [0, T] \) are given by

\[
\chi C_m \left( \frac{\partial V}{\partial t} + I_{\text{ion}}(V) - I_{\text{stim}} \right) = \nabla \times (\sigma \nabla V)
\]

\[I_{\text{ion}}(V) = \sum_{i \in \Gamma} I_i(V)\]

\[(\sigma \nabla V) \cdot n = 0 \text{ on } \partial \Omega \]

\[V(x, 0) = V_{\text{init}},\]

where \( \chi \) is the surface area to volume ratio, \( C_m \) is the cell membrane capacitance, and \( \sigma \) is the conductivity tensor. The \( I_{\text{ion}} \) term represents the current due to ionic flow through cell membrane channels, while \( I_{\text{stim}} \) is an externally applied stimulus. The boundary condition given in this formulation is an implicit no-flux condition, while the initial voltage \( V_{\text{init}} \) is set to \(-83 \text{ mV}\). Note that \( \Gamma \) here represents the set of different subsystem currents that contribute to \( I_{\text{ion}} \).

The ionic model used in this paper is the Ten Tusscher 2006 model, \(^{31}\) which is commonly used in cardiac electrophysiology studies.\(^{32,33}\) This model consists of 12 subsystem currents \( I_i \) where \( i \in \Gamma \), which is given by the set

\[\Gamma = \{ \text{Na}, \text{K1}, \text{to}, \text{Kr}, \text{Ks}, \text{CaL}, \text{NaCa}, \text{NaK}, \text{pCa}, \text{pK}, \text{bCa}, \text{bNa} \}.\]

These currents, in turn, are dependent on a set of 12 gating variables and six non-gating variables. The (independent) gating variables \( g^i \) are all given by Hodgkin-Huxley type ODEs of the form

\[
\frac{dg^i}{dt} = \frac{g^i_{\infty} - g^i}{\tau_{g^i}} \text{ for } i \in G,
\]

where \( G \) is the set of gating variables. Here \( g^i_{\infty} \) and \( \tau_{g^i} \) are exponential functions of \( V \). The non-gating variables, on the other hand, are not independent. They satisfy the following ODEs:

\[
\frac{d\text{Na}}{dt} = -I_{\text{Na}} + I_{\text{bNa}} + 3I_{\text{NaK}} + 3I_{\text{NaCa}}
\]

\[
\frac{d\text{Ca}}{dt} = \text{Ca}_{\text{bufc}} \left[ V_{\text{c}} \left( I_{\text{leak}} - I_{\text{up}} \right) + I_{\text{fex}} - \frac{I_{\text{bCa}} + I_{\text{pCa}} - 2I_{\text{NaCa}}}{2V_c F} \right]
\]

\[
\frac{d\text{Ca}_{\text{ss}}}{dt} = \text{Ca}_{\text{bufs}} \left( I_{\text{up}} - I_{\text{leak}} - I_{\text{rel}} \right)
\]

\[
\frac{d\text{Ca}_{\text{ss}}}{dt} = \text{Ca}_{\text{bufss}} \left( -\frac{I_{\text{Cal}}}{2V_{\text{ss}} F} + \frac{V_{\text{sf}}}{V_{\text{ss}}} I_{\text{rel}} - \frac{V_{\text{c}} I_{\text{fex}}}{V_{\text{ss}}} \right)
\]

\[
\frac{dK}{dt} = -I_{\text{K1}} + I_{\text{to}} + I_{\text{Kr}} + I_{\text{Ks}} - 2I_{\text{NaK}} + I_{\text{pK}} + I_{\text{stim}}
\]

\[
\frac{d\mathcal{R}}{dt} = -k_2 \text{Ca}_{\text{ss}} \mathcal{R} + k_4 (1 - \mathcal{R}).
\]

The right side of the equations in (3) are dependent on a variety of ionic currents and gating variables. We can write the non-gating ODEs more generally as

\[
\frac{dw^j}{dt} = F^j(w, g, V) \text{ for } j \in \text{NG},
\]

where NG is the set of non-gating variables, and \( F^j \) represents the right-hand side of the \( j \)th non-gating variable ODE.
The complexity of the ionic model stems from the sharp spikes found in the ionic model variables. In particular, they can be attributed to the gating variables, which can change rapidly during the depolarization phase. This, in turn, induces a rapid change in the ionic currents, which depend on the gating variables. For example, the fast sodium current $I_{\text{Na}}$, which depends on three gates, and the transient outward current $I_{\text{o}}$, which depends on two gates, both undergo a rapid change over a small time period, as shown in Figure 1. As we shall see later, these rapid changes limit the timestep sizes that can be used for the numerical solution of the non-gating variables, as they depend on these rapidly changing quantities.

### 2.2 Spatial and temporal discretization

In order to discretize the PDE component that is the first equation in (1), in time, we will use the second-order implicit Crank-Nicolson\textsuperscript{34} method. Using a timestep size of $\Delta t$, this yields

$$-\frac{\Delta t}{2\chi C_m} \nabla \cdot (\sigma \nabla V_n) + V_n = \frac{\Delta t}{2\chi C_m} \nabla \cdot (\sigma \nabla V_{n-1}) + V_{n-1} + \Delta t I_{\text{total}}, \tag{5}$$

where $V_n$ and $V_{n-1}$ are the approximations to $V$ at timesteps $n$ and $n-1$, respectively, and $I_{\text{total}} = I_{\text{stim}} - I_{\text{ion}}(V_{n-1})$ groups the ionic current terms together, and the use of $V_{n-1}$ in $I_{\text{ion}}$ linearizes the problem.

For the spatial discretization, we discretize (5) in space using a finite element method with continuous piecewise linear tetrahedral elements. The resulting linear system of equations can then be written as

$$
\begin{pmatrix} M + \frac{\Delta tK}{2\chi C_m} \end{pmatrix} V^h_n = \begin{pmatrix} M - \frac{\Delta tK}{2\chi C_m} \end{pmatrix} V^h_{n-1} + B_{n-1}, \tag{6}
$$

where $M$ and $K$ are the mass and stiffness matrices, respectively, and $B$ is the vector representing the contribution of the ionic model. Note that the notation $V^h_n$ indicates the finite element approximation to $V$ at timestep $n$. Thus, (6) is a linear system of equations, to be constructed and solved at every timestep. The coefficients are time-dependent and therefore the matrix changes from timestep to timestep, but the matrix on the left-hand side of (6) is symmetric and positive definite, and ill-conditioned when $\Delta t$ is large. Moreover, the system becomes more difficult to solve as the geometry of $\Omega$ becomes more complex.

### 2.3 Mesh generation

For generation of patient-specific cardiac meshes, we developed a new mesh-generation pipeline, which utilizes a mix of existing software and software developed in-house. The pipeline begins with the segmentation process, where

**FIGURE 1** Fast sodium ($I_{\text{Na}}$) and transient outward ($I_{\text{o}}$) currents exhibit rapid changes over a small time period.
sequences of Digital Imaging and Communications In Medicine (DICOM) images are segmented in a semi-automatic fashion in order to identify the left ventricle (LV), right ventricle (RV), and the epicardium. This process uses Slicer3D, an open source medical image processing software package,\textsuperscript{35} and an extension called SlicerIGT.\textsuperscript{36} The user selects a set of fiducials for each structure and for each image slice, and the resulting points are automatically interpolated to form a surface. The inter and intra-slice interpolation is achieved using Delaunay alpha triangulation,\textsuperscript{37} which is useful over standard Delaunay triangulation because of the large variance in shape and size between structures, especially the RV due to its non-convex, crescent shape. This allows for flexible segmentation of the different cardiac structures of interest. In addition to identifying the three cardiac shapes, we identify another set of fiducials representing a cutting plane (on one image slice only) in order to denote the top of our region of interest. No interpolation is performed for the cutting plane. The whole first stage produces three VTK\textsuperscript{38} files (one for each structure), and four fiducial files containing point information for the three structures and the cutting plane. Figure 2 shows sample interpolated deliniations during the segmentation process.

Once we have a surface mesh for each cardiac structure, we need to combine them into a single surface mesh, and cut it at the base of the heart. To that end, we developed a Python script to automatically handle this. It works by computing distance functions upon a regular grid, and using VTK’s implicit cutting planes to carve away the non-heart regions. The distance functions are used to first remove everything above the cutting plane, then to remove the interior of the LV and RV, all while interpolating across the boundary of the cut. The effect of this procedure is to stitch together the three models and to hollow out the LV and RV cavities. See Figure 2 for a visualization of the three surfaces before the clipping process, and the resulting surface after.

The surface mesh after the clipping procedure is rather messy because of the simple clipping procedure. An example of this is shown in Figure 3, where the surface mesh on the left contains numerous poorly shaped triangular elements. In order to improve the surface mesh quality, we run the mesh through an open source software package called Approximate Centroidal Voronoi Diagrams (ACVD).\textsuperscript{40} This package utilizes an optimized clustering algorithm in order to create a more efficient mesh.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Top: The segmentation process via Slicer3D, showing the deliniations for LV (green), RV (red), and Epicardium (cyan) on a single image slice. Magnetic resonance imaging data is from the Sunnybrook Cardiac Dataset.\textsuperscript{39} Bottom: The three superimposed surfaces (left) before clipping (LV in green, RV in red, Epicardium in cyan), and the combined surface after clipping (right).}
\end{figure}
to produce a higher quality surface mesh. The resulting surface mesh in Figure 3 on the right shows that, after running through ACVD, we get a much higher quality mesh.

The final step in the mesh generation pipeline is to convert the surface mesh into a 3D unstructured tetrahedral mesh. To this end, we utilize TetGen, a versatile tetrahedral mesh generation package using constrained Delaunay tetrahedralizations. TetGen allows us to specify a myriad of mesh quality parameters, such as the radius-edge ratio or the volume constraint of a tetrahedron, in order to create a mesh of arbitrary quality or size. In addition, TetGen can output resulting meshes in several formats, including ones readable by VTK and Medit, providing flexibility for multiple workflows.

2.4 | Solvers

In order to compute the ionic model contribution, we need to solve the gating and non-gating ODEs in (2) and (4). To that end, we use the Rush-Larsen method, which solves the gating variables via an exponential integrator method and the non-gating variables via the Forward Euler method. The Rush-Larsen method is commonly used in cardiac electrophysiology simulations due to the stiffness of the ionic model ODEs, so we use this method as a starting point.

Using the Rush-Larsen method, the solution of the gating variable ODEs at timestep \( n \) are given by

\[
g_{n}^g = g_{n-1}^g + \left(1 - e^{-\frac{\Delta t}{\tau_{g}}}\right) r_{g} \left(\frac{g_{\infty}^g - g_{n-1}^g}{r_{g}}\right),
\]

where \( \tau_{g} \) and \( g_{\infty}^g \) are functions depending only on \( V \). The solution of the non-gating variable ODEs is more complicated due to the lack of variable independency. Because of this, a single step of the Forward Euler method is used, which then gives the solution at timestep \( n \) as

\[
w_{n}^{j} = w_{n-1}^{j} + \Delta t f_{j}^{j}(w_{n-1}^{j}, g_{n-1}, V_{n-1}).
\]

The order in which the ionic model is solved for at a single timestep is: (1) compute the ionic currents and form \( I_{\text{ion}} \), (2) update the non-gating variables, and (3) update the gating variables.

Due to the rapidly changing nature of the ionic variables during the depolarization process, we are limited to using a small timestep \( \Delta t \), usually between 0.005 and 0.05 ms. Larger values of \( \Delta t \) are unable to accurately resolve the sharp spikes in the ionic model variables, which results in a significant delay in activation time.

Once the ionic model contribution is computed, we incorporate it into (6) and use an iterative method to solve the subsequent linear system. Because the system (6) is symmetric and positive definite, we choose a preconditioned conjugate gradient (CG) method as our linear solver, with a relative stopping tolerance of \( 10^{-6} \). The choice of preconditioner is the algebraic multigrid (AMG) method (V-cycle), specifically the BoomerAMG preconditioner in the
Hypre library, which has been shown to perform well in cardiac electrophysiology simulations.\textsuperscript{16} The AMG parameters chosen include the $l_1$ hybrid symmetric Gauss–Seidel smoother ($l_1$-GS),\textsuperscript{46} the Hybrid Modified Independent Set (H MIS) coarsening algorithm,\textsuperscript{47} a distance-two modified extended interpolation algorithm,\textsuperscript{48} and a strength threshold of $\theta = 0.25$.

3 \hspace{1em} AN INNER-OUTER ITERATIVE METHOD WITH SUBCYCLING

The goal of this section is to describe a temporal subcycling algorithm in the computation of the ionic model, which is developed to allow the use of a larger timestep, thus reducing the computational overhead of a cardiac electrophysiology simulation. As described in Section 2.4, the current method of solving the ODE system of the ionic model is applying an exponential integrator to the gating variables, while the non-gating variables are solved with a single step of the Forward Euler method. Our focus will be on improving the performance of the solution of the non-gating variables.

A simple approach to such a performance improvement is to use a larger timestep, thereby requiring fewer timesteps for a given simulation. When using the Forward Euler method for the solution of the non-gating variable ODEs, however, the timestep size is limited because of accuracy and stability concerns. Thus, a global increase to the timestep size is not feasible. Since the outer (PDE) and inner (ODE) portions are solved separately, we can use different timestep sizes for each portion. Due to the outer portion being solved with a Crank–Nicolson method, we can increase the timestep size for the outer portion, while keeping the smaller timestep (with more iterations) for the inner portion. As we show in Section 4.3, we are able to increase the timestep size by a factor of 10, while only incurring a slight increase in cost associated with the computation of the ionic model, and retaining very similar accuracy to a reference solution.

In order to implement this subcycling idea, we developed a simple subcycling algorithm for the Forward Euler method, herafter known as Subcycling Forward Euler (SFE). This algorithm works by introducing a subcycling timestep, $\Delta t_{\zeta}$, which is smaller than the original timestep by some factor $\zeta \in \mathbb{N}$ that is, $\Delta t_{\zeta} = \frac{\Delta t}{\zeta}$. We then run the Forward Euler scheme a total of $\zeta$ times to match the outer PDE timestep. For this SFE algorithm, subcycling takes place over the entirety of the temporal domain. This ends up being unnecessary because of the short timeframe of the rapid depolarization process. Since the voltage is relatively smooth outside of the depolarization, we can optimize this algorithm to apply subcycling only to the timeframe in which the depolarization occurs, and use a globally larger timestep for the rest of the time.

Therefore, we implemented a hybrid subcycling version using Forward Euler (referred to as HFE hereafter), where we use a time threshold $t_{\text{thresh}}$ to determine whether or not to apply subcycling. The intuition behind this method is that more accuracy is needed during the depolarization phase (with its associated sharp jumps), while the repolarization phase is smoother (see Section 4.2 for a more detailed discussion). For the HFE method, we use the SFE method for $t < t_{\text{thresh}}$, and a single Forward Euler step (as before) for $t \geq t_{\text{thresh}}$. Additionally, we implemented two more algorithms which utilize a combination of a first-order (SFE) and either a second-order or fourth-order Runge–Kutta method, referred to as HRK2 and HRK4, respectively. Specifically, we retain the SFE method for $t < t_{\text{thresh}}$, then switch to one step of either the second-order or fourth-order Runge Kutta for $t \geq t_{\text{thresh}}$. The subcycling algorithm for a single timestep is summarized in Algorithm 1.

In contrast to fully adaptive timestepping methods, where error estimates are computed at each timestep to possibly adjust the timestep size, our subcycling algorithm uses only a temporal threshold to turn off subcycling for the repolarization phase. Thus, we utilize problem-specific information on the structure of a cardiac electrophysiology cycle to inform the subcycling algorithm. This represents a simpler, more efficient solver for a given ionic model, but at the expense of flexibility, since the subcycling threshold might change based on stimulus patterns and cardiac geometry.

4 \hspace{1em} NUMERICAL EXPERIMENTS AND OBSERVATIONS

In order to test the subcycling algorithm, we consider two domains of different shapes and several meshes for each domain. The first domain is a rectangular slab, conforming to the specifications in the Niederer benchmark. This domain is used for verification purposes, by comparing activation times to other cardiac electrophysiology codes at a predefined set of points in the domain. The second domain is a patient-specific heart with the left and right ventricles, and the mesh is generated via the mesh generation pipeline described in Section 2.3.
4.1 Experimental results

The first test will be to establish a reference solution to compare against the Niederer benchmark. To that end, we ran the non-subcycling version of the algorithm with a variety of temporal and spatial resolutions, and measured the activation times at a set of predefined points, shown in Figure 4.

Here, we define activation at a point as the time at which the voltage crosses the 0 mV threshold. For all tests on this rectangular geometry, we used a single 1.5 mm stimulus with a 2 ms duration in the corner at P1, as is used in the Niederer benchmark. The results of this test, shown in Table 1, indicate good spatial and temporal convergence, exhibiting only small variations across resolutions. Furthermore, if we plot the activation times for the $\Delta t = 0.01$ ms test across a diagonal line from P1 to P8, we can see in Figure 5 that the results are robust across the three spatial resolutions.

The next test utilizes biventricular patient-specific geometry. Specifically, we generate three meshes of the same geometry but with different resolutions, and compare activation times at a set of predefined points, similar to the Niederer benchmark. We chose a set of nine points, four are located in the RV (P1–P4) and four in the LV (P5–P8),
respectively, with another point near the center of the septum (P9). These points were chosen such that each part of the domain is roughly represented in the activation time analysis. For these tests, we choose a sufficiently small timestep size of $\Delta t = 0.01$ ms so that there are no accuracy concerns attributed to using a larger timestep. Unlike the single stimulus used with the rectangular geometry, we used a seven stimulus protocol for the patient-specific geometry, four in the LV and three in the RV, based on existing data. Each of these seven stimuli has the same strength and duration as those in the Niederer benchmark, the only difference being they are spherical stimuli with a 3 mm radius.

The results in Table 2 show the activation times (in ms) at the aforementioned points for all three meshes. As with the Niederer benchmark, we see little difference in activation times between the three mesh resolutions, indicating good spatial convergence. It should be noted that because of the use of an unstructured grid, the points at which activation times are measured are not exactly the same between the different meshes. Instead, after choosing the reference points, we use the closest possible points relative to the reference points and measure the activation times there.

In order to test the various subcycling algorithms’ accuracy against the existing method, we created a set of tests on the same geometry. For these tests, we use the 0.1 mm resolution rectangular mesh from the Niederer benchmark. The reference solution is computed with $\Delta t = 0.01$ ms, while the SFE, HFE, HRK2, and HRK4 methods use $\Delta t = 0.1$ ms.

<table>
<thead>
<tr>
<th>$\Delta x$, $\Delta t$</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1, 0.005</td>
<td>1.27</td>
<td>33.53</td>
<td>7.95</td>
<td>34.56</td>
<td>24.28</td>
<td>41.32</td>
<td>25.48</td>
<td>42.14</td>
<td>20.36</td>
</tr>
<tr>
<td>0.1, 0.01</td>
<td>1.27</td>
<td>33.88</td>
<td>8.14</td>
<td>34.90</td>
<td>24.52</td>
<td>41.69</td>
<td>25.71</td>
<td>42.51</td>
<td>20.38</td>
</tr>
<tr>
<td>0.1, 0.05</td>
<td>1.37</td>
<td>36.70</td>
<td>8.76</td>
<td>37.81</td>
<td>26.48</td>
<td>45.17</td>
<td>27.78</td>
<td>46.06</td>
<td>22.01</td>
</tr>
<tr>
<td>0.2, 0.005</td>
<td>1.26</td>
<td>33.09</td>
<td>7.63</td>
<td>33.98</td>
<td>22.65</td>
<td>39.90</td>
<td>23.37</td>
<td>40.70</td>
<td>19.89</td>
</tr>
<tr>
<td>0.2, 0.01</td>
<td>1.27</td>
<td>33.40</td>
<td>7.70</td>
<td>34.30</td>
<td>22.83</td>
<td>40.27</td>
<td>23.57</td>
<td>41.07</td>
<td>20.07</td>
</tr>
<tr>
<td>0.2, 0.05</td>
<td>1.37</td>
<td>35.91</td>
<td>8.20</td>
<td>36.82</td>
<td>24.29</td>
<td>43.16</td>
<td>25.10</td>
<td>43.95</td>
<td>21.43</td>
</tr>
<tr>
<td>0.5, 0.005</td>
<td>1.25</td>
<td>35.68</td>
<td>7.63</td>
<td>36.80</td>
<td>23.03</td>
<td>43.98</td>
<td>25.37</td>
<td>44.89</td>
<td>21.97</td>
</tr>
<tr>
<td>0.5, 0.01</td>
<td>1.26</td>
<td>35.87</td>
<td>7367</td>
<td>36.99</td>
<td>23.12</td>
<td>44.18</td>
<td>25.47</td>
<td>45.09</td>
<td>22.07</td>
</tr>
<tr>
<td>0.5, 0.05</td>
<td>1.36</td>
<td>37.40</td>
<td>7.99</td>
<td>38.53</td>
<td>23.90</td>
<td>45.79</td>
<td>26.26</td>
<td>46.67</td>
<td>22.85</td>
</tr>
</tbody>
</table>

Note: Spatial data is measured in mm, while temporal data is measured in ms.

FIGURE 5  Activation times across a diagonal line of the Niederer benchmark domain
with a subcycling factor of $\zeta = 10$. Additionally, we include a test for $\Delta t = 0.1$ ms without any subcycling. Before running these tests, we need to determine what value to use for $t_{\text{thresh}}$ for the HFE, HRK2, and HRK4 methods. Since we are using the Niederer mesh, we know that the whole domain undergoes electrical activation by the 50 ms mark, so we will use $t_{\text{thresh}} = 50$ ms. The importance of choosing $t_{\text{thresh}}$ correctly will be discussed further in Section 4.2.

Since the algorithmic changes were only applied to the computation of the non-gating variables, it is natural to look at the solutions for these variables for all of the aforementioned algorithms, shown in Figures 6 and 7. In these figures, the solutions of the non-gating variables are plotted over time, for a single point in the same 24 processor partitioning of the Niederer mesh. From these figures, it can be seen that without any subcycling, simply setting $\Delta t = 0.1$ ms causes an inaccuracy in the form of an activation delay, causing the solutions to be shifted away from the reference solution until we reach the plateau phase where the solutions converge. Furthermore, we can see that our subcycling algorithms can use a timestep of $\Delta t = 0.1$ ms without significant loss of accuracy, i.e. the non-gating solutions closely match those of the reference solution. We can also notice that the solutions for the various subcycling methods are almost indistinguishable until we zoom in closely.

Using the same mesh, we can also compute the activation times (using the HFE algorithm) at points P1–P9, as well as the center of the domain C, and compare them to two reference solutions (which use a globally small timestep). We chose the HFE algorithm for this test because it is the fastest, and the other subcycling algorithms use the same pre-threshold subcycling method, so their activation times do not differ from HFE, as the threshold is set to occur after entire domain activation. These activation times, shown in Table 3, are very similar to the reference activation times, which is not surprising given the accuracy retention in the non-gating variables. When compared to the $\Delta x = 0.1$ mm, $\Delta t = 0.005$ ms reference solution, our HFE algorithm has, on average over all points in the domain, a difference of only 0.56% in the reference activation times, while for the $\Delta t = 0.01$ ms solution, the average difference is approximately 0.86%. This, in turn, means our subcycling algorithm retains roughly the same conduction velocities as our reference solutions, which are 0.59 m/s longitudinally and 0.12 m/s transversely.

### 4.2 Subcycling time threshold

Since the goal of our subcycling algorithms is to better resolve the sharp jumps in the solutions of the ionic model variables with a larger timestep, it is important to perform subcycling until most of the domain has undergone electrical activation. To see why, we performed a test on the Niederer geometry using our HFE algorithm, but setting $t_{\text{thresh}} = 25$ ms instead of 50 ms. We then compare the overall solution (voltage) to all previously mentioned algorithms for a point which has not yet undergone electrical activation prior to the subcycling change. The results can be seen in Figure 8, where we can see that when setting $t_{\text{thresh}}$ to 25 ms for the Niederer geometry, points which have not undergone electrical activation have a subsequent delay in activation, similar to the poor accuracy observed when using a global timestep of $\Delta t = 0.1$ ms without subcycling. This delay is not seen when $t_{\text{thresh}}$ is set to a time after entire domain activation. Thus, when using a larger timestep with subcycling, the selection of the subcycling threshold is critical to retaining the accuracy of a smaller timestep. By choosing the threshold to turn off subcycling only after whole domain activation, we can achieve similar accuracy with a larger timestep.

### 4.3 Performance and parallel scalability

Now that the accuracy of our subcycling algorithms has been established, we can look at their performance, in terms of the computational time. The whole point of using a larger timestep is to reduce the compute time taken to complete a
given simulation. Using a larger timestep means fewer timesteps are needed to complete a cardiac cycle. For instance, consider our tests in Section 4.1, where we compared our use of subcycling with $\Delta t = 0.1$ ms to the Niederer benchmark, which uses $\Delta t = 0.01$ ms. Since our timestep is 10 times larger, we need 10 times fewer timesteps to simulate the same time period. There is a bit of a tradeoff here, because when $t < t_{\text{thresh}}$, a single timestep using subcycling is more
expensive than one without subcycling. However, for our problem, it is still much more expensive to solve a single PDE timestep than a single ODE timestep, so the tradeoff becomes favorable. This is made even more favorable by the hybrid algorithms, because electrical activation occurs relatively early in the cardiac cycle, which means the subcycling threshold occurs very early as well. Thus, for most of a cardiac cycle, no subcycling is being performed.

FIGURE 7 Solution comparison over time for CaSR (top), CaSS (middle), and R' (bottom) for various ionic model timestepping algorithms.
In terms of testing raw timings of the various subcycling algorithms, we use the same test setup as in Section 4.1, namely the 0.1 mm Niederer geometry. We run all tests until with 24 cores until \( t = 600 \) ms to get to the plateau phase, the results of which can be seen in Table 4. As expected, the subcycling algorithms outperform the reference method, in this case by 85%–90%, due to computing 10 times fewer timesteps. It can also be seen that the average time taken (per timestep) to solve the ionic model and perform the PCG solve increase when using the subcycling methods. This, however, is more than offset by the reduction in the number of timesteps. It should be noted that for all component timings, like those in Table 4, we exclude the time taken for the assembly of the matrices and vectors associated with the finite element method, as those are intrinsic to external packages. These assembly operations constitute a majority of the time for a simulation (especially for a large mesh), but they are not the focus of this paper.

In order to assess the parallel scalability of the code, we first perform a strong scalability study on the Niederer benchmark without the use of subcycling to establish reference scalability data. In particular, we created a finer mesh on the same rectangular geometry, with approximately 20 M points. We then set the timestep size to \( \Delta t = 0.02 \) ms and ran our code until \( t = 50 \) ms with a number of different processor counts in order to determine the speedup from using more processors. A slightly larger timestep of 0.02 ms was used here to allow for quicker results without affecting the solution accuracy significantly. To complement this scalability study, we also performed the same strong scalability tests on patient-specific geometry, using a finer version of the mesh from Section 4.1 with approximately 36 M points and using the same seven stimulus activation protocol. The only difference is that we ran the scalability test on the realistic geometry until \( t = 20 \) ms due to wall-time constraints for small core counts.

The result of these scalability studies, shown in Figure 9, shows good scalability performance up to 1000 cores, with parallel efficiencies of approximately 90% at 1000 processors, and an almost linear speedup. The drop in efficiency is likely due to the communication overhead of using a large number of processors relative to the size of the mesh. In other words, the amount of work per processor becomes too low, and communication costs start to dominate the timings.
Furthermore, in Figure 10 we can see the different components of the solver, and how each one scales with the number of processors. We can see that the evaluation of the ionic model scales very well even to 1000 processors, which is to be expected due to minimal communication costs involved in solving the ionic model. Additionally, the PCG setup and solve phases scale well, with slight performance degradations at 1000 processors due to communication overhead.

The scalability of our subcycling algorithms is not expected to change, since the only change occurs in the ionic model which scales almost perfectly. To compare the scalability using subcycling against the reference scalability, we run the same tests (both on rectangular and patient-specific geometry), except using the HFE algorithm. Additionally, we run this test until \( t = 100 \text{ ms} \) with \( t_{\text{thresh}} = 50 \text{ ms} \) to allow for equal time with and without subcycling. Table 5 shows overall timings for pre and post threshold, as well as solver components as the processor count increases. We can see that 0–50 ms timeframe (with subcycling) takes significantly longer to compute than the subsequent 50–100 ms timeframe, again due to the need to resolve the sharpness in the depolarization phase. Once the initial 50 ms have been resolved via subcycling, the entire rest of the cardiac cycle can be computed using a larger global timestep, yielding significant speedup.

As is the case for the reference simulation, scalability using the HFE algorithm is good up to 1000 processors, with parallel efficiency at approximately 90%–95% at 1000 processors (see Figure 11). Similar scalability results can be expected from the other subcycling methods (SFE, HRK2, and HRK4), since they are all variations of the same algorithm.

### 5 | CONCLUDING REMARKS

Given the predictive potential of cardiac simulations in clinical applications such as drug discovery and treatment planning, it is important to be able to rapidly simulate cardiac electrical propagation on patient-specific geometry.

---

**TABLE 4** Timing comparison of subcycling algorithms and the reference solution on the Niederer geometry

<table>
<thead>
<tr>
<th>Method</th>
<th>Total</th>
<th>Ionic model</th>
<th>Ionic model (avg)</th>
<th>PCG</th>
<th>PCG (avg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
<td>13152.9</td>
<td>765.92</td>
<td>0.0128</td>
<td>1116.32</td>
<td>0.0186</td>
</tr>
<tr>
<td>SFE</td>
<td>2071.21</td>
<td>758.28</td>
<td>0.1264</td>
<td>161.38</td>
<td>0.0269</td>
</tr>
<tr>
<td>HFE</td>
<td>1447.89</td>
<td>135.76</td>
<td>0.0226</td>
<td>159.48</td>
<td>0.0266</td>
</tr>
<tr>
<td>HRK2</td>
<td>1474.41</td>
<td>153.29</td>
<td>0.0255</td>
<td>169.04</td>
<td>0.0282</td>
</tr>
<tr>
<td>HRK4</td>
<td>1505.88</td>
<td>185.83</td>
<td>0.0310</td>
<td>165.31</td>
<td>0.0276</td>
</tr>
</tbody>
</table>

Note: All timings are given in seconds, and averages are per timestep.

**FIGURE 9** Speedup (left) and parallel efficiency (right) for both the rectangular geometry from the Niederer benchmark and the patient-specific geometry.
Our base cardiac electrophysiology pipeline, adapted from the Cardioid package from LLNL, compares well to other approaches via the Niederer benchmark, exhibits both spatial and temporal convergence, and scales well up to 1000 cores, including on realistic patient geometry. In order to improve the turnaround time for a cardiac electrophysiology simulation, we proposed a novel subcycling time integration algorithm for the non-gating variables in the TT06 ionic model. This subcycling method allows a larger overall timestep to be used, while utilizing

<table>
<thead>
<tr>
<th>Cores</th>
<th>Total</th>
<th>( t &lt; t_{\text{thresh}} )</th>
<th>( t \geq t_{\text{thresh}} )</th>
<th>Ionic</th>
<th>PCG Setup</th>
<th>PCG Solve</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>20045.2</td>
<td>12419.5</td>
<td>7625.74</td>
<td>4102.15</td>
<td>7.25</td>
<td>4600.16</td>
</tr>
<tr>
<td>48</td>
<td>9563.4</td>
<td>5911.58</td>
<td>3651.82</td>
<td>2055.91</td>
<td>3.5</td>
<td>1911.63</td>
</tr>
<tr>
<td>96</td>
<td>4882.16</td>
<td>2984.82</td>
<td>1897.34</td>
<td>1045.32</td>
<td>1.49</td>
<td>1013.85</td>
</tr>
<tr>
<td>192</td>
<td>2441.48</td>
<td>1490.6</td>
<td>950.88</td>
<td>536.04</td>
<td>0.77</td>
<td>528.8</td>
</tr>
<tr>
<td>384</td>
<td>1298.29</td>
<td>801.09</td>
<td>497.2</td>
<td>271.15</td>
<td>1.25</td>
<td>322.97</td>
</tr>
<tr>
<td>768</td>
<td>662.75</td>
<td>402.27</td>
<td>260.48</td>
<td>138.51</td>
<td>0.79</td>
<td>154.61</td>
</tr>
<tr>
<td>1008</td>
<td>519.05</td>
<td>316.66</td>
<td>202.39</td>
<td>107.52</td>
<td>1.02</td>
<td>134.73</td>
</tr>
</tbody>
</table>

Note: Here, \( \Delta t = 0.1 \) with \( \zeta = 10 \).
subcycling (with a smaller timestep) for the rapid depolarization phase. The accuracy of this method is shown through various tests, including the Niederer benchmark and a seven stimuli activation pattern on patient-specific geometry, to be almost the same as when using a smaller global timestep. Furthermore, because the subcycling method only modifies the ODE solver in the ionic model, scalability remains excellent up to 1000 cores. When compared to the small timestep reference solutions, the subcycling methods show significant improvement to the overall completion time of a simulation, being 85%–90% faster than the original Cardioid FEM code for the Niederer benchmark.

The subcycling time integration algorithm presented has been shown to be promising for improving cardiac monodomain simulation time, but can still be improved in future work. One such improvement could be an automatic selection of the subcycling threshold, based on the underlying ionic model. This would further reduce simulation turnaround time, but also allow the subcycling algorithm to be seamlessly used on other potential ionic models, including the more recent ORd model. However, further testing is required to determine the applicability of our subcycling algorithm to other ionic models.

ACKNOWLEDGMENTS

This work utilized resources from the University of Colorado Boulder Research Computing Group, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Xiao-Chuan Cai https://orcid.org/0000-0003-0296-8640

REFERENCES


50.とした．

**How to cite this article:** Laudenschlager S, Cai X-C. An inner-outer subcycling algorithm for parallel cardiac electrophysiology simulations. *Int J Numer Meth Biomed Engng*. 2023;39(3):e3677. doi:10.1002/cnm.3677