

AUTOMATED ADJOINTS OF COUPLED PDE-ODE SYSTEMS*

P. E. FARRELL[†], J. E. HAKE[‡], S. W. FUNKE[§], AND M. E. ROGNES[§]

Abstract. Mathematical models that couple partial differential equations (PDEs) and spatially distributed ordinary differential equations (ODEs) arise in biology, medicine, chemistry, and many other fields. In this paper we discuss an extension to the FEniCS finite element software for expressing and efficiently solving such coupled systems. Given an ODE described using an augmentation of the Unified Form Language (UFL) and a discretization described by an arbitrary Butcher tableau, efficient code is automatically generated for the parallel solution of the ODE. The high-level description of the solution algorithm also facilitates the automatic derivation of the adjoint and tangent linearization of coupled PDE-ODE solvers. We demonstrate the capabilities of the approach on examples from cardiac electrophysiology and mitochondrial swelling.

Key words. finite element methods, coupled PDE-ODE, adjoints, FEniCS, dolfin-adjoint, code generation

AMS subject classifications. 65L06, 65M60, 65M32, 35Q92

DOI. 10.1137/17M1144532

1. Introduction. In this paper we discuss solvers for systems involving spatially dependent ordinary differential equations (ODEs) of the following general form: given an initial condition $y_0(x)$, find $y = y(x, t)$ such that

$$(1) \quad y_t(x, t) = f(y, x, t), \quad y(x, t_0) = y_0(x)$$

for all x in a point set $X \subset \Omega \subseteq \mathbb{R}^d$, where d is the spatial dimension. The subscript t refers to differentiation in time. The right-hand side f cannot depend on spatial derivatives of y ; the ODE is decoupled at different points. We also discuss the automated derivation of the adjoint and tangent linearization of such models: these can be used to identify the sensitivity of the solution to model parameters, solve inverse problems for unknown parameters, and characterize the stability of trajectories.

Problems of the form (1) may arise in a number of different contexts. One important example, of particular interest here, is as a subproblem in an operator splitting algorithm for solving a system of coupled time-dependent PDEs. Examples of application areas include *computational physiology*, e.g., simulating cardiac electrophysiology [13, 27, 36], mitochondrial swelling [10], or pulmonary gas transport [7, 40]; *computational geology*, e.g., studying groundwater flow and contamination [39]; or *computational chemistry*, e.g., modeling plasma-enhanced chemical vapor deposition [14].

For motivational purposes and concreteness, we begin this exposition by introducing two distinct examples of coupled PDE-ODE systems where problems of the

*Submitted to the journal's Software and High-Performance Computing section August 22, 2017; accepted for publication (in revised form) March 26, 2019; published electronically June 6, 2019.

<http://www.siam.org/journals/sisc/41-3/M114453.html>

Funding: This work was supported by a Center of Excellence grant awarded to the Center for Biomedical Computing at Simula Research Laboratory from the Research Council of Norway, by EPSRC grants EP/K030930/1 and EP/M011151/1, a NOTUR grant NN9316K, and the generous support of Sir Michael Moritz and Harriet Heyman.

[†]Mathematical Institute, University of Oxford, OX2 6GG, UK (patrick.farrell@maths.ox.ac.uk).

[‡]Ski Videregaende skole, Ski, Norway (johan.hake@gmail.com).

[§]Simula Research Laboratory, P.O. Box 134, 1325 Lysaker, Norway (simon@simula.no, meg@simula.no).

form (1) arise. We note that both of our motivating examples originate from computational physiology but emphasize that the abstractions, algorithms, and framework presented here targets the general setting. We return to numerical results for these examples in section 6.

Example 1 (the bidomain equations). As our first motivating example, we will consider the bidomain equations for the propagation of an electrical signal in a nondeforming domain Ω [36]: find the transmembrane potential $v = v(x, t)$, the extracellular potential $u_e(x, t)$, and additional state variables $s = s(x, t)$ such that for $t \in (0, T]$:

$$(2a) \quad v_t - \operatorname{div}(M_i \operatorname{grad} v + M_i \operatorname{grad} u_e) = -I_{\text{ion}}(v, s) + I_s \quad \text{in } \Omega,$$

$$(2b) \quad -\operatorname{div}(M_i \operatorname{grad} v + (M_i + M_e) \operatorname{grad} u_e) = 0 \quad \text{in } \Omega,$$

$$(2c) \quad s_t = F(v, s) \quad \forall x \in \Omega.$$

In (2), I_{ion} is a given nonlinear function describing ionic currents and F defines a system of nonlinear functions, while M_i and M_e are the intracellular and extracellular conductivity tensors, respectively, and I_s is a given stimulus current. The function F cannot depend on spatial derivatives of v and s ; it defines a pointwise system of ODEs. The specific form of I_{ion} and F are typically prescribed by a given *cardiac cell model* and may vary greatly in complexity: from involving a single state variable s such as the FitzHugh–Nagumo model [13] to models with, e.g., 41 state variables such as the model of O’Hara et al. [28]. The system (2) is closed with appropriate initial and boundary conditions. After the application of operator splitting, the ODE step is decoupled from the PDE step and is a system of the form of (1).

Example 2 (mitochondrial swelling). As a second example, we consider a model proposed in [10] to describe the swelling of mitochondria. Mitochondrial swelling plays a key role in the process of programmed cell death (*apoptosis*) and thus for the life cycle of cells. Mathematically, we consider the following model [10, p. 26]: find the calcium concentration $u = u(x, t)$ and the densities of mitochondria states $N_i = N_i(x, t)$ for $i = 1, 2, 3$ such that

$$(3a) \quad u_t = d_1 \Delta(|u|^{q-2}u) + d_2 g(u) N_2 \quad \text{in } \Omega,$$

$$(3b) \quad N_{1t} = -f(u) N_1 \quad \forall x \in \Omega,$$

$$(3c) \quad N_{2t} = f(u) N_1 - g(u) N_2 \quad \forall x \in \Omega,$$

$$(3d) \quad N_{3t} = -g(u) N_2 \quad \forall x \in \Omega,$$

where $d_1 \geq 0$ is a diffusion coefficient, $d_2 \geq 0$ is a feedback parameter, $q \geq 2 \in \mathbb{N}$ determines the nonlinearity of the diffusion-type operator, and g and f are prescribed functions. The functions N_1 , N_2 , and N_3 describe the densities of unswollen, swelling, and completely swollen mitochondria, respectively. Equation (3a) is a spatially coupled PDE, while equations (3b)–(3d) define pointwise ODEs. The system is closed with initial conditions and Dirichlet boundary conditions for u . Again, after operator splitting, a subproblem of the form (1) results.

Over the last decade, there has been a growing interest in computational frameworks for the rapid development of numerical solvers for PDEs such as the FEniCS Project [21], the Firedrake Project [30], and Feel++ [29]. The rapid development of solvers is achieved by offering high-level abstractions for the expression of such problems. Each of these projects provides a domain-specific language for specifying finite element variational formulations of PDEs and associated software for their efficient solution. Moreover, such domain-specific languages typically allow for automation

of high-level algorithms. For instance, the dolfin-adjoint project [12] automatically derives discrete adjoint and tangent linear models from a FEniCS forward model. These auxiliary models allow for the efficient computation of functional gradients and Hessian-vector products and are essential ingredients in stability analyses, parameter identification, and inverse problems. We refer the reader to [1, 23] and [12] for more detailed descriptions of the existing functionality provided by these packages. As the adjoint discretization is derived automatically from the forward discretization, the resulting derivatives are discretely consistent [18, sect. 2.9].¹ Similar computer systems also exist for solving ODEs and their associated adjoint problems [24, 33, 41].

However, none of these computer systems do currently extend to allow for *efficient* implementations of coupled PDE-ODE systems. While monolithic finite element discretizations of coupled PDE-ODE systems such as (2) or (3) may easily be specified and solved using the current software features in FEniCS, this approach will involve an unreasonably large computational expense, as the monolithic system is highly nonlinear and the solver cannot exploit the fact that the ODEs are spatially decoupled.

Instead, operator splitting is a popular method of choice for solving coupled PDE-ODE systems. Operator splitting decouples a PDE-ODE system into a spatially coupled system of PDEs and a spatially decoupled collection of ODE systems. We refer the reader to [15, 25, 34] for in-depth descriptions of operator splitting as applied to general time-dependent PDE systems. This collection of ODE systems then typically takes the form (1), which can be solved using well-known temporal discretization methods. For instance, in the context of cardiac electrophysiology, operator splitting is considered the most efficient approach to solving the bidomain equations (2) [36] and is implemented as standard in many hand-written codes; see, e.g., [26, 38]. Heretofore it has not been possible to specify spatially decoupled ODE systems of the form (1) in high-level PDE frameworks such as the FEniCS or Firedrake projects.

This work addresses the gap in available abstractions, algorithms, and software for arbitrary PDE-ODE systems. The methods presented in this paper are implemented in the FEniCS Project and dolfin-adjoint. We introduce high-level domain-specific language constructs for specifying collections of ODE systems and for specifying multistage ODE schemes via Butcher tableaux [8]. This has several major benefits. By automatically generating the solver from a high-level description of the problem, practitioners can flexibly explore a range of mathematical models. Another advantage is that a high-level description facilitates the automated derivation of the associated tangent linear and adjoint models. This represents a significant saving in the time taken to investigate, e.g., application-oriented research questions.

The main new contributions of this paper are (i) the extension of the FEniCS finite element system to enable efficient the large-scale forward solution of coupled, time-dependent PDE-ODE systems via operator splitting, and (ii) the extension of dolfin-adjoint to automatically derive and solve the associated tangent linear and adjoint models, enabling efficient automated computation of functional gradients for use in, e.g., optimization or adjoint-based sensitivity analysis.

This paper is organized as follows. In section 2, we summarize the general operator splitting setting, the resulting separate PDE and ODE systems, and natural discretizations of these. We continue in section 3 by briefly describing the

¹Note that for some applications it can be more appropriate to discretize the adjoint differently; for example, higher-order adjoint discretizations can be used to compute more accurate goal-based error estimates [5, sect. 5]. This latter approach is not considered here.

well-established ODE schemes that we consider in this work: the multistage and Rush–Larsen families. We discuss the adjoint and tangent linear discretizations of a general operator splitting scheme and derive the adjoint and tangent linear models for the multistage and Rush–Larsen schemes in section 4. Key features of our new implementation are described in section 5. We present numerical results for two different application examples in section 6 and use these to evaluate the performance of our implementation. In section 7, we provide some concluding remarks and discuss current limitations and possible future extensions.

2. Operator splitting for coupled PDE-ODE systems.

2.1. Operator splitting. Operator splitting is a classical approach to the numerical solution of differential equations in general and coupled PDE-ODE systems in particular [15, 25, 34]. For coupled PDE-ODE systems, the algorithm approximately solves the full system of equations by alternating between the solution of a system of PDEs and a system of ODEs defined over each time interval. The advantage of this approach is that it decouples the solution of the (spatially decoupled) ODEs from the spatially coupled PDEs at each time step, typically allowing for more specialized solution algorithms for the separate subproblems. For instance, for the bidomain system (2) where the ODEs are typically highly nonlinear while the PDEs are linear, operator splitting can offer significant computational advantages [36].

To present a typical operator splitting algorithm and make ideas more concrete, we consider (2). For an in-depth description of this application case, we refer the reader to [36]. Applying a variable order operator split to (2), the resulting scheme reads: given initial conditions v^0, s^0 , an order parameter $\theta \in [0, 1]$, and time points $\{t_0, \dots, t_N\}$ with an associated timestep $\kappa_n = t_{n+1} - t_n$, then for each time step $n = 0, 1, \dots, N - 1$:

1. Compute v^* and s^* by solving the ODE system

$$(4a) \quad v_t = -I_{\text{ion}}(v, s),$$

$$(4b) \quad s_t = F(v, s)$$

over $\Omega \times [t_n, t_n + \theta\kappa_n]$ with initial conditions v^n, s^n .

2. Compute v^\dagger and u_e^{n+1} by solving the PDE system

$$(5a) \quad v_t - \text{div}(M_i \text{grad } v + M_i \text{grad } u_e) = I_s,$$

$$(5b) \quad -\text{div}(M_i \text{grad } v + (M_i + M_e) \text{grad } u_e) = 0$$

over $\Omega \times [t_n, t_{n+1}]$ with initial condition v^* .

3. If $\theta < 1$, compute solutions v^{n+1} and s^{n+1} solving (4) over $\Omega \times [t_n + \theta\kappa_n, t_{n+1}]$ with initial conditions v^\dagger and s^* .

The split scheme relies on the repeated solution of a nonlinear system of ODEs and (in this case) a linear system of PDEs. The main advantage of the approach is that it allows for the separate discretization and solution of the ODEs and PDEs, with the respective solutions as feedback into the other system. For $\theta = 1/2$, the resulting Strang splitting scheme is second-order accurate; for other values of θ the resulting scheme is first-order accurate [36].

2.2. Discretization of the separate PDE and ODE systems. Suppose now that the relevant PDE system (e.g., (5)) is discretized in space by a finite element method defined over a mesh \mathcal{T}_h of the domain Ω and in time by some suitable

temporal discretization. Efficient solution algorithms for such discretizations are well-established (see, e.g., [21]) and will not be detailed further here.

At each iteration the solution of the PDE system relies on the solution of the ODE system (e.g., (4)), or at least the ODE solution evaluated at some finite set of points

$$X = \{x_i\}_{i=1}^{|X|}$$

in space. For instance, the set of points X may be taken as the nodal locations of the finite element degrees of freedom or the quadrature points of the mesh. As a consequence of this and the spatial locality of ODEs, a natural approach to discretizing the system of ODEs is to step the ODE system forward in time at this set of points X . Typically, $|X|$ is very large and the efficient repeated solution of these systems of ODEs is key. This is the setting that we focus on next.

3. Solution schemes for the ODE systems. The initial value problem (1) decouples in space, and henceforth we consider its solution for a fixed x . With a minor abuse of notation, let $y = y(x)$ and $f(y, t) = f(y, x, t)$ so that (1) reads as the classical ODE problem: find $y \in C^1([T_0, T_1]; \mathbb{R}^m)$ for a certain $m \in \mathbb{N}$ such that

$$(6) \quad y_t(t) = f(y, t), \quad y(T_0) = y_0$$

for $t \in [T_0, T_1]$. There exists a wide variety of solution schemes for (6) including but not limited to multistage, multistep, and implicit-explicit schemes; see, e.g., [8]. In this work, we focus on two classes of schemes: multistage schemes and so-called Rush–Larsen schemes. The Rush–Larsen schemes are commonly used in cardiac electrophysiology in general and for discretizations of the bidomain equations (2) in particular. These two classes of schemes are detailed below. Keeping the general iterative operator splitting setting in mind, we present the schemes on a single time step $[T_0, T_1]$ with $\kappa = T_1 - T_0$ for brevity of notation.

3.1. Multistage schemes. An s -stage multistage scheme for (6) is defined by a set of coefficients a_{ij} , b_i , and c_i for $i, j = 1, \dots, s$, commonly listed in a so-called *Butcher tableau*; for more details, see, e.g., [8]. Given y^0 at T_0 , the scheme finds the stage variables k_i for $i = 1, \dots, s$ satisfying

$$(7) \quad k_i = f \left(y^0 + \kappa \sum_{j=1}^s a_{ij} k_j, T_0 + c_i \kappa \right),$$

and subsequently sets the solution y^1 at T_1 via

$$(8) \quad y^1 = y^0 + \kappa \sum_{i=1}^s b_i k_i.$$

Note that (7) defines a system of (nonlinear) equations to solve for the stage variables k_i ($i = 1, \dots, s$) if $a_{ij} \neq 0$ for any $j \geq i$. Our implementation demands that $a_{ij} = 0$ for $j > i$, i.e., does not allow the computation of earlier stages to depend on the values of later stages.

3.2. The Rush–Larsen scheme and its generalization. Recall that $y = \{y_i\}_{i=1}^m$ and $f(y) = \{f_i(y)\}_{i=1}^m$. The original Rush–Larsen scheme employs an exponential integration scheme for all linear terms for each component f_i (for $i = 1, 2, \dots, M$), and a forward Euler step for all nonlinear terms [32]. Let $J_i = \frac{\partial f_i}{\partial y_i}(y^0, T_0)$

be the i th diagonal component of the Jacobian at time T_0 for $i = 1, \dots, M$, and assume that J_i is nonzero. For brevity, denote $f_i(y^0, T_0) = f_i^0$. The Rush–Larsen scheme then computes y_i^1 at T_1 by

$$(9) \quad y_i^1 = y_i^0 + \begin{cases} J_i^{-1} f_i^0 (e^{\kappa J_i} - 1) & \text{if } f_i \text{ is linear in } y_i, \\ \kappa f_i^0 & \text{if } f_i \text{ is not linear in } y_i. \end{cases}$$

For stiff systems of ODEs, the forward Euler step in (9) can become unstable for large κ . This motivates a generalized version of the Rush–Larsen scheme [35]. In this generalization, the exponential integration step is used for all components of y reducing the scheme to

$$(10) \quad y_i^1 = y_i^0 + J_i^{-1} f_i^0 (e^{\kappa J_i} - 1).$$

Both (9) and (10) are first-order accurate in time [32, 35].

Both the original and the generalized Rush–Larsen schemes can be developed into second order schemes by repeated use as follows. The solution

$$y^{\frac{1}{2}} \text{ at } T_{\frac{1}{2}} = T_0 + \frac{\kappa}{2}$$

is computed in the first step and is used in f and its linearization J to compute y^1 in the second step. Let

$$f_i^{\frac{1}{2}} = f_i(y^{\frac{1}{2}}, T_{\frac{1}{2}}),$$

and let

$$J_{i, \frac{1}{2}} = \frac{\partial f_i}{\partial y_i}(y^{\frac{1}{2}}, T_{\frac{1}{2}})$$

be the diagonal component of the Jacobian at time $T_{\frac{1}{2}}$ for $i = 1, \dots, M$. More precisely, the second order version of (9) and (10) is then given by

$$(11a) \quad y_i^{\frac{1}{2}} = y_i^0 + \begin{cases} J_i^{-1} f_i^0 (e^{\frac{\kappa}{2} J_i} - 1) & \text{if } f_i \text{ is linear in } y_i, \\ \frac{\kappa}{2} f_i^0 & \text{if } f_i \text{ is not linear in } y_i, \end{cases}$$

$$(11b) \quad y_i^1 = y_i^0 + \begin{cases} J_{i, \frac{1}{2}}^{-1} f_i^{\frac{1}{2}} (e^{\kappa J_{i, \frac{1}{2}}} - 1) & \text{if } f_i \text{ is linear in } y_i, \\ \kappa f_i^{\frac{1}{2}} & \text{if } f_i \text{ is not linear in } y_i \end{cases}$$

for the original Rush–Larsen scheme and

$$(12a) \quad y_i^{\frac{1}{2}} = y_i^0 + J_i^{-1} f_i^0 (e^{\frac{\kappa}{2} J_i} - 1),$$

$$(12b) \quad y_i^1 = y_i^0 + J_{i, \frac{1}{2}}^{-1} f_i^{\frac{1}{2}} (e^{\kappa J_{i, \frac{1}{2}}} - 1)$$

for the generalized Rush–Larsen scheme. Here, (12b) is simplified compared to the scheme in [35]: instead of evaluating f and J at $\bar{y}^{(i)}$ (cf. Eq. (7) in [35]) we use $y_i^{\frac{1}{2}}$.

4. Adjoints and tangent linearizations of coupled PDE-ODE systems.

Our aim is to automatically derive the adjoint and tangent linear equations of operator splitting schemes in a manner that allows for efficient solution of the resulting systems. We proceed as follows: we first state the adjoint and tangent linear system for general problems, then consider the special case of mixed PDE-ODE systems, and finally discuss the specific adjoint and tangent linear versions of multistage and Rush–Larsen schemes.

We begin by considering a general system of discretized equations in the form: find $y \in Y \subset \mathbb{R}^M$ such that

$$(13) \quad F(y) = 0.$$

The adjoint equation of (13) associated with a real-valued functional of interest $J : Y \rightarrow \mathbb{R}$ is: find the adjoint solution $\bar{y} \in \mathbb{R}^M$ such that

$$(14) \quad \frac{\partial F^*}{\partial y}(y) \bar{y} = \frac{\partial J}{\partial y}(y),$$

where the superscript $*$ denotes the adjoint operator. The associated tangent linear equation with respect to some auxiliary control parameter m is given by: find the tangent linear solution \dot{y} such that

$$(15) \quad \frac{\partial F}{\partial y}(y) \dot{y} = \frac{\partial F}{\partial m}(y).$$

We now turn our attention to the case where F is a coupled PDE-ODE system. This system might consist of multiple timesteps and stages within each timestep. For a guiding example, we consider a single timestep of the bidomain equations (4)–(5) ($N = 1$). This does not cause a loss of generality, as the central feature of the dolfin-adjoint software is to combine the adjoint of each individual step in a transient simulation together to derive the adjoint of the entire system via graph manipulation; for more details we refer the reader to [12]. For brevity, we denote the ODE operator (4) as O and the PDE operator (5) as P . We can then rewrite the scheme in the form (13):

$$(16a) \quad v_0 - v^0 = 0, \quad (\text{Set initial condition})$$

$$(16b) \quad s_0 - s^0 = 0, \quad (\text{Set initial condition})$$

$$(16c) \quad O(v^*, s^*; v^0, s^0) = 0, \quad (\text{Solve bidomain ODEs (4)})$$

$$(16d) \quad P(v^\dagger, u_e^1; v^*) = 0, \quad (\text{Solve bidomain PDEs (5)})$$

$$(16e) \quad O(v^1, s^1; v^\dagger, s^*) = 0. \quad (\text{Solve bidomain ODEs for final state})$$

The first pair of equations in (16) set the initial conditions, the following equations represents a collection of nonlinear ODEs, the third equation is a system of PDEs, and the last equation is again a collection of ODEs.

From (14), the adjoint system for (16) derives as

$$(17) \quad \begin{pmatrix} I & \frac{\partial O^*}{\partial(v^0, s^0)} & 0 & 0 \\ 0 & \frac{\partial O^*}{\partial(v^*, s^*)} & \frac{\partial P^*}{\partial(v^0, s^0)} & \frac{\partial O^*}{\partial s^*} \\ 0 & 0 & \frac{\partial P^*}{\partial(v^\dagger, u_e^1)} & \frac{\partial O^*}{\partial v^\dagger} \\ 0 & 0 & 0 & \frac{\partial O^*}{\partial(v^1, s^1)} \end{pmatrix} \begin{pmatrix} \bar{v}^0 \\ \bar{s}^0 \\ \bar{v}^* \\ \bar{s}^* \\ \bar{v}^\dagger \\ \bar{u}_e^1 \\ \bar{v}^1 \\ \bar{s}^1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \frac{\partial J}{\partial v^1} \\ \frac{\partial J}{\partial s^1} \end{pmatrix}.$$

For brevity, the left-hand side matrix combines 2×2 blocks, and the functional of interest is assumed to only depend on the final state variables v^1, s^1 . The adjoint system (17) is a linear coupled PDE-ODE system with upper-triangular block structure, which can efficiently be solved by backwards substitution. The last block row represents a linear adjoint ODE system, preceded by an adjoint PDE system, and preceded by another adjoint ODE system, and finalized by a variable assignment in the first block row which represents the adjoint version of setting the initial conditions.

The derivation of adjoint and tangent linear equations for finite element discretizations of PDEs, such as the third block row in (17), is well established; see, e.g., [12]. However, the automated derivation of adjoint and tangent linear systems

for the multistage and Rush–Larsen discretizations of the generic ODE system (6), such as the second and fourth block row in (17), is less so, and these derivations are presented below.

4.1. Adjoints and tangent linearizations of multistage schemes. Consider a multistage discretization of (6) as described in section 3.1 with a given Butcher tableau a_{ij}, b_j, c_i , $i, j = 1, \dots, s$. Taking $s = 3$ for illustrative purposes, we can write (7)–(8) in form (13) as:

$$\begin{pmatrix} y^0 \\ k_1 \\ k_2 \\ k_3 \\ y^1 \end{pmatrix} - \begin{pmatrix} y_0 \\ f(w_1, T_0 + c_1 \kappa) \\ f(w_2, T_0 + c_2 \kappa) \\ f(w_3, T_0 + c_3 \kappa) \\ y^0 + \kappa \sum_{i=1}^3 b_i k_i \end{pmatrix} = 0,$$

with

$$w_i = y^0 + \kappa \sum_{j=1}^s a_{ij} k_j.$$

Assuming that the functional of interest J is independent of the internal stage values k_i , we can derive the adjoint problem as

$$\begin{pmatrix} \bar{y}^0 \\ \bar{k}_1 \\ \bar{k}_2 \\ \bar{k}_3 \\ \bar{y}^1 \end{pmatrix} - \begin{pmatrix} 0 & \frac{\partial f_1}{\partial w_1}^* & \frac{\partial f_2}{\partial w_2}^* & \frac{\partial f_3}{\partial w_3}^* & I \\ 0 & \kappa a_{11} \frac{\partial f_1}{\partial w_1}^* & \kappa a_{21} \frac{\partial f_2}{\partial w_2}^* & \kappa a_{31} \frac{\partial f_3}{\partial w_3}^* & \kappa b_1 \\ 0 & 0 & \kappa a_{22} \frac{\partial f_2}{\partial w_2}^* & \kappa a_{32} \frac{\partial f_3}{\partial w_3}^* & \kappa b_2 \\ 0 & 0 & 0 & \kappa a_{33} \frac{\partial f_3}{\partial w_3}^* & \kappa b_3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{y}^0 \\ \bar{k}_1 \\ \bar{k}_2 \\ \bar{k}_3 \\ \bar{y}^1 \end{pmatrix} = \begin{pmatrix} \frac{\partial J}{\partial y^0} \\ 0 \\ 0 \\ 0 \\ \bar{y}_0 \end{pmatrix},$$

where \bar{y}_0 is the terminal condition for the adjoint solution at T_1 .

Generalizing to s stages, we see that we first solve for the adjoint stage values and then compute the adjoint solution at time T_0 via

$$(18a) \quad \left(I - \kappa a_{ii} \frac{\partial f_i}{\partial w_i}^* \right) \bar{k}_i = \kappa b_i \bar{y}^1 + \sum_{j=i+1}^s \kappa a_{ji} \frac{\partial f_j}{\partial w_j}^* \bar{k}_j,$$

$$(18b) \quad \bar{y}^0 = \bar{y}^1 + \sum_{i=1}^s \frac{\partial f_i}{\partial w_i}^* \bar{k}_i.$$

Following a similar calculation, the tangent linearization of the multistage scheme with s stages is: given y_0 at T_0 compute \dot{k}_i for $i = 1, \dots, s$ and then \dot{y}^1 at T_1 via

$$(19a) \quad \left(I - \kappa a_{ii} \frac{\partial f_i}{\partial w_i} \right) \dot{k}_i = \frac{\partial f_i}{\partial w_i} \dot{y}^n + \sum_{j=1}^{i-1} \kappa a_{ij} \frac{\partial f_i}{\partial w_i} \dot{k}_j + \frac{\partial f_i}{\partial m},$$

$$(19b) \quad \dot{y}^{n+1} = \dot{y}^n + \sum_{i=1}^s \kappa b_i \dot{k}_i.$$

4.2. Adjoints and tangent linearizations of Rush–Larsen schemes. We now turn to consider adjoints and tangent linearizations of the (generalized) Rush–Larsen schemes. We here derive the equations for the first order generalized Rush–Larsen scheme given by (10):

$$y_i^1 = y_i^0 + J_i^{-1} f_i^0 (e^{\kappa J_i} - 1)$$

for each component $i = 1, \dots, M$ of the state variable.

We can write (10) in form (13) as

$$\begin{pmatrix} y^0 - y_0 \\ y^1 - L(y^0) - y^0 \end{pmatrix} = 0,$$

where y_0 is the given initial condition at T_0 and $L(y^0) = \{L_i(y^0)\}_{i=1}^M$ with

$$L_i(y^0) = J_i^{-1} f_i^0 (e^{\kappa J_i} - 1) \equiv \frac{\partial f_i}{\partial y_i}(y^0, T_0)^{-1} f_i(y_0, T_0) (e^{\kappa \frac{\partial f_i}{\partial y_i}(y^0, T_0)} - 1).$$

The adjoint system of the first-order generalized Rush–Larsen scheme is then given by

$$(20) \quad \begin{pmatrix} \bar{y}^0 \\ \bar{y}^1 \end{pmatrix} - \begin{pmatrix} 0 & I + \frac{\partial L}{\partial y^0}^* \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{y}^0 \\ \bar{y}^1 \end{pmatrix} = \begin{pmatrix} \frac{\partial J}{\partial y^0} \\ \bar{y}_0 \end{pmatrix},$$

where \bar{y}_0 is a given terminal condition for the adjoint at T_1 . Similarly, the tangent linear system (15) with respect to an auxiliary parameter m is given by

$$(21) \quad \begin{pmatrix} \dot{y}^0 \\ \dot{y}^1 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ I + \frac{\partial L}{\partial y^0} & 0 \end{pmatrix} \begin{pmatrix} \dot{y}^0 \\ \dot{y}^1 \end{pmatrix} = \begin{pmatrix} \dot{y}_0 \\ \frac{\partial L(y^0)}{\partial m} \end{pmatrix},$$

where \dot{y}_0 is a given initial condition at T_0 .

The adjoint and tangent linear equations for the other Rush–Larsen schemes are derived following the same steps, and we therefore omit the details here.

5. FEniCS and dolfin-adjoint abstractions and algorithms. The FEniCS Project defines a collection of software components targeting the automated solution of differential equations via finite element methods [21]. The components include the Unified Form Language (UFL) [3], the FEniCS Form Compiler (FFC) [22], and the finite element library DOLFIN [23]. The separate dolfin-adjoint project and software automatically derives the discrete adjoint and tangent linear models from a forward model written in the Python interface to DOLFIN [12]. This section presents our extensions of the FEniCS form language and the FFC, and other new FEniCS and dolfin-adjoint software features targeting coupled PDE-ODE systems in general and collections of systems of ODE in particular.

5.1. Variational formulation of collections of ODE systems. Consider a spatial domain $\Omega \subset \mathbb{R}^d$ tessellated by a mesh \mathcal{T}_h and a collection of general initial problems as defined by (1) over a set of points X defined relative to \mathcal{T}_h . For $x_i \in X$, denote by δ_{x_i} the Dirac delta function centered at $x_i \in \mathbb{R}^d$ such that

$$f(x_i) = \int_{\mathbb{R}^d} f \delta_{x_i} \, dx$$

for all continuous, compactly supported functions f . We will also write

$$(22) \quad \sum_{x_i \in X} \int_{\mathbb{R}^d} f \delta_{x_i} \, dx \equiv \sum_{x_i \in X} \int f \, dP(x_i) \equiv \sum_{x_i \in X} \langle f, 1 \rangle_{x_i} \equiv \langle f, 1 \rangle_X.$$

Drawing inspiration from our context of finite element variational formulations, we can then write (1) as: find $y(\cdot, t)$ such that

$$(23) \quad \langle y_t(t), \psi_i \rangle_X = \langle f(y, t), \psi_i \rangle_X$$

for all (basis) functions (or distributions) ψ_i such that $\psi_i(x_i) = 1$ and $\psi_j(x_i) = 0$ for $j \neq i$.

The remainder of this section describes the extensions of the FEniCS and dolfin-adjoint systems to allow for in particular abstract representation and efficient forward and reverse solution of systems of the form (23).

5.2. Extending UFL with vertex integrals. UFL is an expressive domain-specific language for abstractly representing (finite element) variational formulations of differential equations. In particular, the language defines syntax for integration over various domains. Consider a mesh \mathcal{T}_h of geometric dimension d with cells $\{\mathcal{T}\}$, interior facets $\{\mathcal{F}^i\}$, and boundary facets $\{\mathcal{F}\}^b$. Interior facets are defined as the $d-1$ dimensional intersections between two cells and thus for any interior facet \mathcal{F}^i we can write $\mathcal{F}^i = \mathcal{T}^+ \cap \mathcal{T}^-$. UFL defines the sum of integrals over cells, sum of integrals over boundary facets, and sum of integrals over interior facets by the \mathbf{dx} , \mathbf{ds} , and \mathbf{dS} measures, respectively. For example, the following linear variational form defined in terms of a piecewise polynomial u defined relative over \mathcal{T}_h :

$$L(u) = \sum_{\{\mathcal{T}\}} \int_{\mathcal{T}} u \, \mathbf{dx} + \sum_{\{\mathcal{F}^i\}} \int_{\mathcal{F}^i} u|_{\mathcal{T}^+} \, \mathbf{ds} + \sum_{\{\mathcal{F}^b\}} \int_{\mathcal{F}^b} u \, \mathbf{ds}$$

is naturally expressed in UFL as the following.

```
L = u*dx + u('++')*dS + u*ds
```

To allow for point evaluation over the vertices of a mesh (cf. (22)), we have introduced a new *vertex integral* (or *vertex measure* in UFL terms) type with default instantiation \mathbf{dP} .

```
from ufl import Measure
dP = Measure('vertex')
L = f*dP
```

In agreement with (22), the vertex integral is defined by:

$$(24) \quad \mathbf{f} * \mathbf{dP} \equiv \langle f, 1 \rangle_{\mathcal{V}(\mathcal{T}_h)} = \sum_{x_i \in \mathcal{V}(\mathcal{T}_h)} f(x_i),$$

where $\mathcal{V}(\mathcal{T}_h)$ is the set of all vertices of the mesh \mathcal{T}_h . Vertex measures restricted to a subset of vertices can be defined as for all other UFL measure types.

This basic extension of UFL crucially allows for the abstract specification of collections of ODEs such as (23) in a manner that is consistent and compatible with abstract specification of finite element variational formulations of PDEs. It also allows for the native specification of, e.g., point sources in finite element formulations of PDEs in FEniCS.

5.3. Vertex integrals in the FFC. The FFC generates specialized C++ code [2] from the symbolic UFL representation of variational forms and finite element spaces [22]. To accommodate the new vertex integral type, we have extended the UFC interface with a class `vertex_integral` defining the interface for the tabulation of the element matrix corresponding to the evaluation of an expression at a given vertex.

The FFC code generation pipeline has been correspondingly extended to allow for the generation of optimized code from the UFL representation of variational forms involving vertex integrals. This allows for subsequent automated assembly of variational forms involving single vertex integrals (as illustrated above) and also in combination with other (cell, interior facet, exterior facet) integrals.

5.4. DOLFIN features for solving collections of ODE systems.

5.4.1. Assembly of vertex integrals. We have extended DOLFIN with support for automated assembly of variational forms that include vertex integrals. The

support is currently limited to forms defined over test and trial spaces with vertex-based degrees of freedom only. The assembly algorithm follows the standard finite element assembly pattern by iterating over the vertices of the mesh, computing the map from local to global degrees of freedom, evaluating the local element tensor based on generated code, and adding the contributions to the global tensor. The vertex integral assembler runs natively in parallel via message passing interface (MPI).

5.4.2. Specification and generation of multistage schemes. Since version 2016.1, DOLFIN supports the specification of multistage schemes of the form (7)–(8) for the solution of collections of ODE systems of the form (1), via their Butcher tableaux. The DOLFIN class `ButcherMultiStageScheme` takes as input the right-hand side expression f , the solution y , the Butcher tableau specified via a , b , and c , and secondary variables such as a time variable, the (integer) order of the scheme. From this specification, DOLFIN automatically generates a variational formulation of each of the separate stages in the multistage scheme, with each stage in accordance with (7)–(8). DOLFIN can also automatically generate the variational formulations corresponding to the adjoint scheme (18) and to the tangent linear scheme (19) on demand. A set of common multistage schemes are predefined including Crank–Nicolson, explicit Euler, implicit Euler, Fourth-order explicit Runge–Kutta, ESDIRK3, ESDIRK4 [8, 20].

Similar features have also been implemented for easy specification of Rush–Larsen schemes; cf. (9)–(12) and the automated generation of the corresponding adjoint and tangent linear schemes; cf., e.g., (20) and (21) through the class `RushLarsenScheme`. Both `ButcherMultiStageScheme` and `RushLarsenScheme` subclass the `MultiStageScheme` class.

5.4.3. PointIntegralSolvers. To allow for the efficient solution of the multistage schemes for collections of systems of ODEs, we have introduced a targeted `PointIntegralSolver` class in DOLFIN 2016.1. This solver class takes a `MultiStageScheme` as input, and its main functionality is to compute the solutions over a single time step.

The solver iterates over all vertices of the mesh and solves for the relevant (stage and/or final) variables at each vertex. The solution algorithm for each stage depends on whether the stage is explicit or implicit. For implicit stages, a simplified Newton algorithm is used; i.e., the Jacobian is calculated once at the beginning of the solve and used for every subsequent Newton step [9, sect. 2.1.2]. As the resulting linear systems are typically small and dense, a direct (LU) algorithm is used for the inner solves. For explicit stages, a simple vector update is performed.

The point integral solver runs natively in parallel via MPI. For a cell-partitioned mesh distributed between N processes, each cell is owned by one process and that process performs the solve for each vertex on the cell (once). As little communication is required between processes, the total solve is expected to scale linearly in N .

5.5. Extensions to dolfin-adjoint. Dolfin-adjoint has been extended to support the new features including automatically deriving and computing adjoint and tangent linear solutions for `PointIntervalSolver`. The symbolic derivation of the variational formulation for the adjoint and tangent linear equations for the `MultiStageScheme` are based on (18) and (19) for the Butcher tableau defined schemes and (20), (21), and its analogies for the Rush–Larsen schemes.

Since the adjoint problem depends on the forward solution, dolfin-adjoint must implement a strategy for storing the forward solution at each timestep. By default, dolfin-adjoint stores every forward solution in computer memory. For multistage ODE schemes, this would require storing all intermediate stage solutions. To avoid excessive

memory usage in these cases, dolfin-adjoint only stores the initial stage solution at each time step and recomputes the intermediate stages when needed. Optionally, dolfin-adjoint can employ binomial checkpointing to reduce the memory requirements of solving the adjoint problem at the cost of computational overhead; see [12, 17] for more details. Dolfin-adjoint has been extended to support binomial checkpointing also for coupled ODE-PDE problems. Initial testing yielded the expected results as well as computational overhead, but checkpointing methods will not be further discussed in this paper.

The overall result is that operator splitting algorithms as described in section 2.1 can be fully specified and efficiently solved within FEniCS. Moreover, the corresponding adjoint and tangent linear models and first and second order functional derivatives may be efficiently computed using dolfin-adjoint with only a few lines of additional code.

6. Applications. In this section, we demonstrate the applicability and performance of our implementation by considering the two examples presented in the introduction, originating from the computational modeling of mitochondrial swelling and cardiac electrophysiology, respectively. The complete supplementary code is openly available; see [11].

6.1. Application: Mitochondrial swelling. We consider the mathematical model defined by (3). Inspired by [10], we let $x = (x_0, x_1) \in \Omega = [0, 1]^2$ and $t \in [0, T]$ with $T = 35$, and consider the initial conditions $N_{1,0} = 1$, $N_{2,0} = 0$, $N_{3,0} = 0$, and

$$(25) \quad u_0(x) = \frac{30}{\int_{\Omega} M(x) \, dx} M(x), \quad M(x_0, x_1) = \frac{1}{2\pi} e^{-0.5(X_0^2 + X_1^2)},$$

with $X_0 = (\alpha - \beta)x_0 + \beta$, $X_1 = (\alpha - \beta)x_1 + \beta$ taking $\alpha = 3$ and $\beta = -1$. We consider

$$(26) \quad f(s) = \begin{cases} 0, & s < C^-, \\ f^*, & s > C^+, \\ \frac{f^*}{2} \left(1 - \cos \left(\frac{s - C^-}{C^+ - C^-} \pi \right) \right), & \text{otherwise,} \end{cases}$$

and

$$(27) \quad g(s) = \begin{cases} g^*, & s > C^+, \\ \frac{g^*}{2} \left(1 - \cos \left(\frac{s}{C^+} \pi \right) \right), & \text{otherwise} \end{cases}$$

with $C^- = 20$, $C^+ = 200$, $f^* = 1$, $g^* = 0.1$, $d_1 = 2 \times 10^{-6}$, and $d_2 = 30$.

We are interested in the total amount of completely swollen mitochondria (whose density is given by N_3) at the final time T and its sensitivity to the initial condition for u (u_0). Thus, our functional of interest J is given by

$$(28) \quad J = \int_{\Omega} N_3(T),$$

and our goal is to compute the derivative of J with respect to u_0 .

We use a second order Strang splitting scheme for (3), a Crank–Nicolson discretization in time and continuous piecewise linear finite elements in space. The scheme then reads as: given initial conditions u_0 , $N_{1,0}$, $N_{2,0}$, $N_{3,0}$, and time points $\{t_0, t_1, \dots, t_N\}$ with time step $\kappa_n = t_{n+1} - t_n$, then for each $n = 0, 1, \dots, N - 1$:

1. Compute u^* and $N_{i,0}^*$ $i = 1, 2, 3$ solving

(29)

$$u_t = d_2 g(u) N_2, \quad N_{1t} = -f(u) N_1, \quad N_{2t} = f(u) N_1 - g(u) N_2, \quad N_{3t} = g(u) N_2,$$

over $\Omega \times [t_n, t_n + \frac{1}{2}\kappa_n]$ with initial conditions u^n, N_i^n for $i = 1, 2, 3$.

2. Compute a solution u^\dagger of: find $u_h \in V_h$ such that

$$(30) \quad \int_{\Omega} (u_h - u^*) v + \kappa_n \operatorname{grad} \mathcal{A}(\langle u_h \rangle) \cdot \operatorname{grad} v \, dx = 0$$

for all $v \in V_h$ with $\langle u \rangle = \frac{1}{2}(u + u^*)$ and where $\mathcal{A}(u) = u$ or $\mathcal{A}(u) = |u|^{q-2}u$.

3. Compute solutions u^{n+1} and N_i^{n+1} , $i = 1, 2, 3$, solving (29) over $\Omega \times [t_n + \frac{1}{2}\kappa_n, t_{n+1}]$ with initial conditions u^\dagger and N_i^* , $i = 1, 2, 3$.

We choose to discretize (29) via the ESDIRK4 method in time and set a tolerance of 10^{-10} for the inner Newton solves. Further, we take $q = 3$, take $\kappa_n = 0.5$, and let \mathcal{T}_h be a uniform tessellation of Ω with $N_x \times N_x \times 2$ triangles.

The initial and final solutions are presented in Figure 1 while the L^2 -gradient of the objective functional J given by (28) with respect to the initial calcium concentration u_0 is presented in Figure 2.

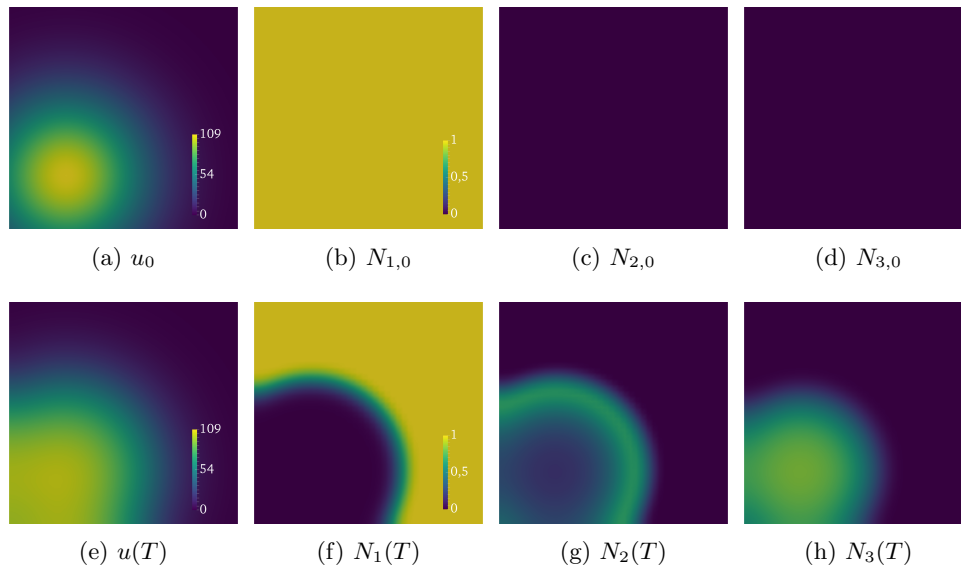


FIG. 1. Initial conditions and solutions $t = T = 35$ for the mitochondrial swelling example with $N_x = 40$. The scale is common for A and E and ranges from 0 (blue) to 108.3 (yellow). The scale is common for B–D and F–G and ranges from 0 (blue) to 1 (yellow).

To verify that the computed gradient is correct, we performed a Taylor test at a given spatial and temporal resolution with a given perturbation seed. The results are listed in Table 1 and demonstrate the expected orders of convergence, indicating a correctly computed gradient.

In a multistage scheme, a number of intermediate stage solutions are computed during the solution process. Since dolfin-adjoint recomputes the intermediate stage solutions for each time step, the minimal ratio of adjoint runtime to forward runtime for multistage ODE solves using this strategy is approximately 2. For a linear PDE

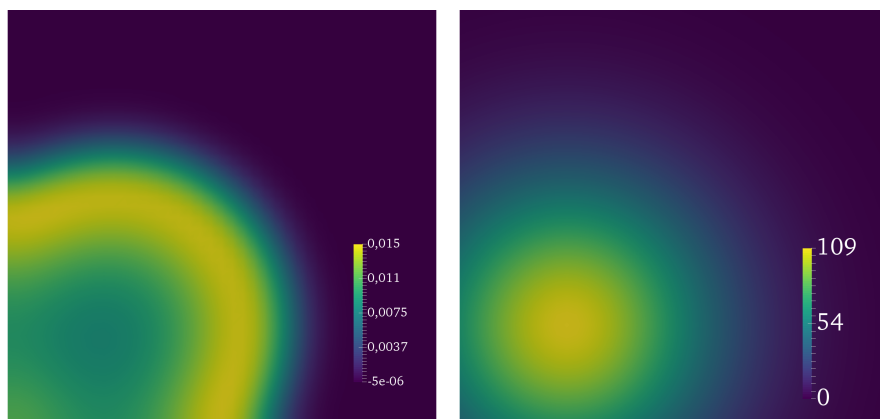


FIG. 2. Sensitivity of J , the total amount of completely swollen mitochondria at T , L^2 -gradient of J (left) with respect to the initial condition u_0 (right).

TABLE 1

Taylor remainders $R_0 = |J(u_0 + \delta u_0) - J(u_0)|$ and $R_1 = |J(u_0 + \delta u_0) - J(u_0) - \nabla J(u_0)\delta u_0|$ for the mitochondria example with functional given by (28) (to three significant digits). Computations performed with $T = 35$, $\kappa_n = 0.5$, and $N_x = 40$ and the ESDIRK4 scheme.

δu_0	$R_0(\delta u_0)$	Order	$R_1(\delta u_0)$	Order
0.5	1.22×10^{-3}		1.25×10^{-6}	
0.25	6.14×10^{-4}	1.00	3.11×10^{-7}	2.00
0.125	3.07×10^{-4}	1.00	7.77×10^{-8}	2.00
0.0625	1.53×10^{-4}	1.00	1.94×10^{-8}	2.00
0.03125	7.67×10^{-5}	1.00	4.85×10^{-9}	2.00

solve, the optimal ratio of adjoint run time to forward run time, assuming all linear solves take the same amount of time, is 1. For a nonlinear PDE solve via a Newton iteration, the optimal ratio of adjoint run time to forward run time, assuming all linear solves take the same amount of time, is $1/N$ where N is the number of average Newton iterations in the forward solve. These optimal ratios for linear and nonlinear PDE solves have been observed for typical dolfin-adjoint usage [12]. For this example, two Newton iterations were required on average to solve the PDEs to within a tolerance of 10^{-10} both for $N_x = 40$ and 80, and assembly dominated the PDE solve runtime. Thus for this example with two multistage ODE solves and a nonlinear PDE solve in each time step, the optimal ratio of adjoint runtime to forward runtime is expected to be in the range $0.5-2$, and closer to the upper bound than the lower bound depending on the distribution of computational cost between the PDEs and ODEs.

Experimentally observed timings for this example are listed in Tables 2–3 for $N_x = 40, 80, 160$. From Table 2, we observe that the adjoint-to-forward runtime ratio for the total solve is in the range $1.38-1.95$ and decreasing with increasing problem size. The corresponding gradient-to-forward runtime is in the range $1.44-2.28$, again decreasing with increasing problem size as expected. From Table 3, we observe that adjoint-to-forward runtime ratio for the ODE solves is in the range $2.35-2.41$ for this set of problem sizes, also decreasing with increasing problem size. We also note that the ODE solve runtime, both for the forward and adjoint solves, appears to scale linearly with the problem size $M = N_x^2$, as is optimal. The ODE solves account for 37–40% of the total forward runtime for this example.

TABLE 2

Run times to compute the forward solution t_F and to compute the adjoint t_A and the functional gradient t_G of the functional defined by (28) for the mitochondria example with increasing spatial resolution for $T = 35$, $\kappa_n = 0.5$ with the ESDIRK4 multistage scheme.

N_x	Forward t_F (s)	Adjoint t_A (s)	Ratio $\frac{t_A}{t_F}$	Gradient t_G (s)	Ratio $\frac{t_G}{t_F}$
40	9.30	18.1	1.95	21.2	2.28
80	34.0	52.2	1.53	56.1	1.65
160	135	186	1.38	193	1.44

TABLE 3

Run times for solving the ODE systems when computing the forward solution $t_{F,O}$ and when computing the adjoint $t_{A,O}$ of the functional defined by (28) for $T = 35$, $\kappa_n = 0.5$ with the ESDIRK4 multistage scheme.

N_x	Forward ODEs (s)	(Of total)	Adjoint ODEs (s)	Ratio
	$t_{F,O}$		$t_{A,O}$	$\frac{t_{A,O}}{t_{F,O}}$
40	3.48	(37 %)	8.40	2.41
80	13.7	(40 %)	32.6	2.37
160	53.6	(40 %)	126	2.35

We also conducted numerical experiments varying the multistage scheme used in the ODE solves, including the 1-stage Backward Euler (BDF1), the 2-stage Crank–Nicolson (CN2), the 3-stage ESDIRK3 scheme in addition to the previously reported 4-stage ESDIRK4 scheme. The results (using $N_x = 40$) are reported in Table 4. We observe that the adjoint-to-forward runtime ratio ranges from 2.04 (BDF1) to 2.41 (ESDIRK3, ESDIRK4) and the general trend is that this ratio increases with the number of stages as expected but stabilizes around 2.3–2.4.

TABLE 4

Run times computing the forward solution t_F , for solving the ODE systems when computing the forward solution $t_{F,O}$ and when computing the adjoint $t_{A,O}$ of the functional defined by (28) for $T = 35$, $\kappa_n = 0.5$, $N_x = 40$ for some common multistage schemes.

Scheme	Forward	Forward ODEs (s)	Adjoint ODEs (s)	Ratio
	t_F	$t_{F,O}$	$t_{A,O}$	$\frac{t_{A,O}}{t_{F,O}}$
BDF1	6.56	0.76	1.55	2.04
CN2	6.67	0.94	2.14	2.28
ESDIRK3	8.30	2.54	6.02	2.37
ESDIRK4	9.30	3.48	8.40	2.41

6.2. Application: Cardiac electrophysiology (two dimensions). In this example, we consider the bidomain equations (2) over a two-dimensional rectangular domain $\Omega = [0, 50] \times [0, 50]$ (mm) with coordinates (x_0, x_1) . We will consider a set of different cardiac cell models of increasing complexity: a reparametrized FitzHugh–Nagumo (FHN) model with 1 ODE state variable [13], the Beeler–Reuter (BR) model with 7 ODE state variables [6], the ten Tusscher and Panfilov (TTP) epicardial cell model with 18 ODE state variables [37], and the Grandi, Pasqualini, and Bers (GPB) cell model with 38 ODE state variables [16]. The description of each of these cell models is available via the CellML repository, and our implementation of the ionic current I_{ion} and F arising in (2) was automatically generated from the CellML models for the BR, TTP, and GPB models. We refer to the supplementary code [11] for the precise description of the models including our choice of FHN parameters.

Our parameter setup is otherwise as follows. We let $\chi = 140$ (mm^{-1}), $C_m = 0.01$ ($\mu\text{F}/\text{mm}^2$), and let $M_i = \text{diag}(g_{if}, g_{if})$ and $M_e = \text{diag}(g_{ef}, g_{es})$, where $g_{ef} = 0.625/(\chi C_m)$, $g_{es} = 0.236/(\chi C_m)$, and $g_{if} = 0.174/(\chi C_m)$. We let $I_s = 0$, and set the initial condition

$$(31) \quad v_0 = 10 \left(\frac{x_0}{50} \right)^2 + 10$$

for v . The other state variables are initialized to the default initial conditions given by the respective CellML models.

We use a second-order Strang splitting scheme (i.e., $\theta = 0.5$ for (4)–(5)), a Crank–Nicolson discretization in time and continuous piecewise linear finite elements in space for the PDE system, and different Rush–Larsen type discretizations (Rush–Larsen (RL)1, generalized Rush–Larsen (GRL)1, RL2, GRL2) for the time discretization of the ODE systems.² The coupled linear systems that arise at each PDE step (5) were solved with a block-preconditioned generalized minimal residual method scheme with a relative tolerance of 10^{-10} . The complete details of the solver are detailed in the supplementary code [11].

We take $\kappa_n = 0.1$ (ms) and consider a uniform mesh of the computational domain with $N_x \times N_x \times 2$ triangles. The coarsest mesh considered used $N_x = 160$. For this case, simulations converged using any of the cell models and schemes and gave qualitatively correct results, except the Grandi cell model discretized by the RL1 scheme for which the numerical solution scheme failed to converge.

We are interested in computing the gradient with respect to the extracellular fiber conductivity g_{ef} of the following objective functional

$$(32) \quad J(v) = \sum_{i=1}^5 \int_{\Omega} (v(t_i)^2 + s(t_i)^2) \, dx$$

for a equidistributed set of time points t_i . The model initial condition and solution for the transmembrane potential at $T = 100$, together with the L^2 -gradient of the objective functional with respect to the fiber conductivity are illustrated in Figure 3.

The following numerical experiments were performed on the Abel supercomputer at the University of Oslo. All runtime experiments were repeated three times, of which the minimum timing values are reported here.

6.2.1. Verifying the correctness of the discrete gradient. To verify that the computed functional gradient is correct, we performed a Taylor test in a random perturbation direction at given spatial and temporal resolutions. The results with the FHN cell model and the GRL1 scheme are listed in Table 5 and demonstrate the expected orders of convergence without a computed gradient (order 1) and when using the computed discrete gradient (order 2) in the Taylor expansion. We also performed similar experiments for the other cell models (BR, TTP, and GPB) and other Rush–Larsen solution schemes (RL1, GRL2, RL2). We used the same settings as in Table 5, except for the Grandi cell model with any other Rush–Larsen scheme than GRL1, where the time step had to be reduced to 0.01 in order for the forward solver to converge. We obtained the expected convergence order for all combination of cell models and schemes tested. These results indicate that the automatically derived and computed adjoints and gradient are correct for all the Rush–Larsen schemes implemented.

²We observed spurious wave propagation results when using a first-order splitting scheme and implicit Euler for the BR model.

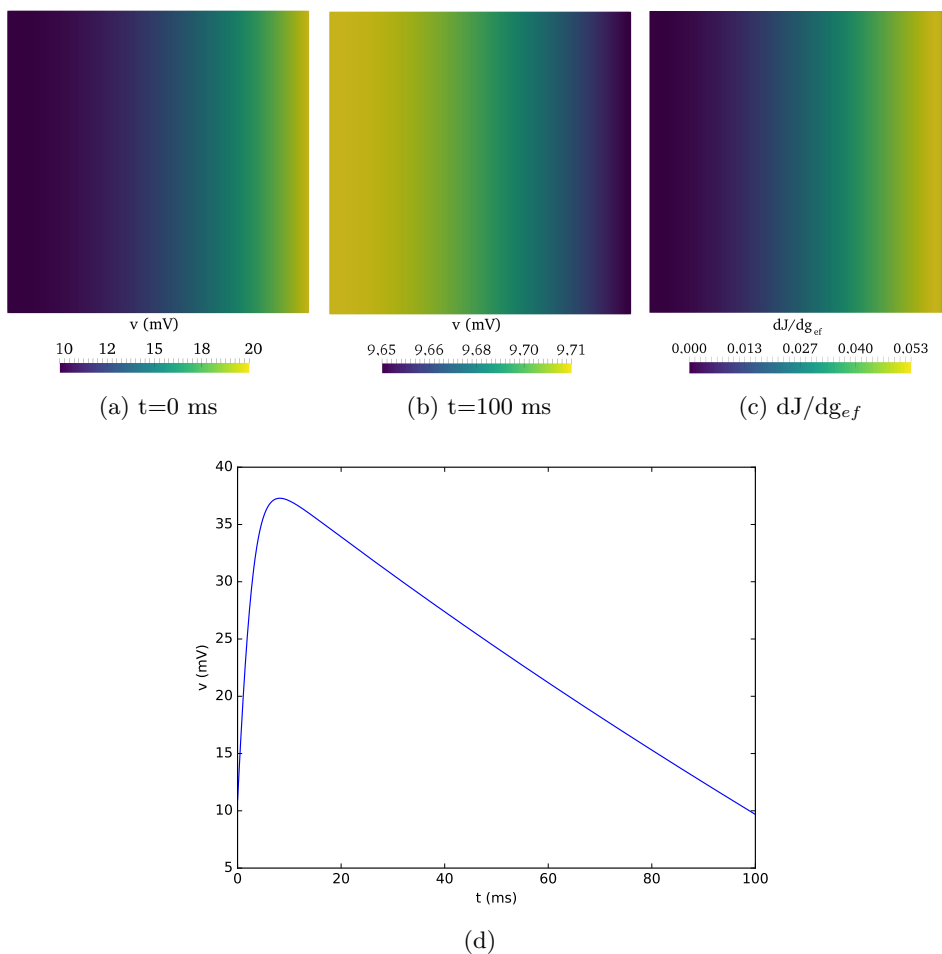


FIG. 3. (a) Transmembrane potential v at $t = 0$ ms and (b) at $t = 100$ ms using the FHN model, computed with $\kappa_n = 0.1$ and $N_x = 160$. (c) L^2 -gradient of the functional with respect to the extracellular fiber conductivity g_{ef} . (d) Timeplot of transmembrane potential v at $x = y = 25$ mm.

TABLE 5

Taylor convergence results for the cardiac electrophysiology two-dimensional test case. Taylor remainders $R_0 = |J(g_{ef} + s\delta)|$ and $R_1 = |J(g_{ef} + s\delta) - J(g_{ef}) - \nabla J(g_{ef})s\delta|$ for decreasing step sizes s and a random direction δ with functional given by (32). Computations performed with $T = 10.0$, $\kappa_n = 0.1$, and $N_x = 160$ with the FHN cell model and the GRL1 scheme. We observe that the remainders converge at first order without gradient information and at second order with gradient information, as expected.

s	$R_0(s)$	Order	$R_1(s)$	Order
0.005	15.7	0.96	0.541	
0.0025	7.97	0.98	0.145	1.90
0.00125	4.02	0.99	0.0382	1.92
0.000625	2.02	0.99	0.00983	1.96

6.2.2. Adjoint runtime performance. As the bidomain PDEs (5) are linear, a rough theoretical estimate of the optimal PDE adjoint to PDE forward runtime

ratio is 1, assuming that the solution times in the adjoint and forward PDE solves are equal and dominated by the finite element assembly and linear solvers. Similarly, the ODEs are solved with explicit Rush–Larsen schemes with a theoretically optimal adjoint-to-forward ratio of 1 as well.

In Table 6, we list the total runtime and the runtime components corresponding to only the ODE solves, only the PDE solves, and an intermediate variable update (merge) step for both the forward and the adjoint solution computation for a test case with $T = 10.0$, $\kappa_n = 0.1$, $N_x = 160$, the GRL1 scheme, and the four different cardiac cell models (FHN, BR, TTP, GPB).

TABLE 6

Adjoint-to-forward runtime performance for a cardiac electrophysiology two-dimensional test case. Breakdown of forward and adjoint runtimes and their ratio for different cell models. Each row shows total runtime (s), runtime for the ODE, PDE, and merge steps (s), and percentage of the total runtime not accounted for by these three steps (Other). The computations were performed with $T = 10$, $\kappa_n = 0.1$, $N = 160$, GRL1 as the ODE scheme, and (32) as the adjoint functional.

Model		Total	ODEs	PDEs	Merge	Other
FHN	Forward	38.8	3.20	34.1	0.466	3 %
	Adjoint	44.4	8.27	28.1	1.75	14 %
	Ratio	1.15	2.58	0.82	3.76	
BR	Forward	48.3	10.9	34.5	0.979	4 %
	Adjoint	62.1	20.4	28.2	4.66	14 %
	Ratio	1.29	1.88	0.81	4.76	
TTP	Forward	67.3	24.4	35.5	1.97	8 %
	Adjoint	99.5	47.8	28.2	9.98	14 %
	Ratio	1.48	1.96	0.79	5.07	
GPB	Forward	91.8	40.8	34.7	3.63	14 %
	Adjoint	152	81.7	28.2	19.6	15 %
	Ratio	1.66	2.00	0.81	5.40	

From Table 6, for the forward solves, we observe that the total runtime increases with cell model complexity. The forward PDE step runtime is essentially independent of the cell model, while the forward ODE step runtime increases with the cell model complexity as expected: the ODE runtime corresponds to approximately 10% of the PDE runtime for the simplest FHN model, and approximately 114% of the PDE runtime for the most complex Grandi model. The runtime of the forward merge step is insignificant in comparison to the ODE and PDE steps but increases with cell model complexity.

We observe that also the adjoint PDE runtimes are comparable for all cell models. The resulting adjoint-to-forward PDE step ratios are in the range 0.79–0.82. The decrease in compute time for the adjoint PDE solves compared to the forward PDE solves are attributable to the Krylov solvers: fewer Krylov solver iterations were necessary for convergence for the adjoint solves compared to the forward solves. For the ODE solves, we observe that the adjoint-to-forward ratio for the ODE solves is in the range 1.88–2.58 for the different cell models. Additionally, we observe that the runtime of the adjoint merge step is significantly larger than the corresponding time for the forward merge step and with an increasing adjoint-to-forward ratio with increasing cell model complexity 3.76–5.40. We suggest that the reason for the increased and increasing adjoint runtime is that the adjoint merge step involves additional assembly of variational forms of complexity comparable to that of the cell model.

Overall, we observe that the adjoint-to-forward total runtime ratios are in the range 1.15–1.66 for the different cell models considered, with increasing adjoint-to-forward ratio with increasing cell model complexity. Since the PDE runtime dominates the total runtime for the simpler cell models, the total adjoint-to-forward ratio is closer to 1 for these models. On the other hand, for the more complicated cell models, the ODE runtimes dominate, and so we observe adjoint-to-forward ratios closer to 2 for these.

In the last column of Table 6, we observe that the combined ODE and PDE solve time contributes to between 97% (for FHN) and 86% (for Grandi) of the total forward runtime for the cell models considered. The remaining forward runtime cost is primarily caused by the initialization routines such as loading the computational mesh and creating the required function spaces. For the corresponding adjoint runtimes, the combined ODE and PDE solve times contribute to between 86% (for FHN) and 85% (for Grandi) of the total run time, and the percentage contribution decreases with the complexity of the cell model. The remaining runtime cost can primarily be attributed to recording the forward states during the adjoint solve (which is included in the reported total run time but not in the ODE or PDE runtimes here). We also timed the computation of the discrete gradient and noted that the cost of additionally computing the gradient was negligible (less than 0.1% of the adjoint solve runtime).

6.2.3. Forward and adjoint parallel scalability. In this section, we discuss the weak and strong parallel scalability of the forward and adjoint ODE solvers and the scalability of the point integral solvers in particular applied to this two-dimensional electrophysiology test case. We focus on the scalability of the point integral solver in particular as this is the primary new feature described in this paper. Previous studies have discussed the parallel scalability of FEniCS in general [1, 12, 31].

We ran a series of weak scaling experiments on Abel, hosted by the Norwegian national computing infrastructure, for the previously considered series of cell models (FHN, BR, TTP, GPB) on an increasing number of CPU cores (1, 4, 16, 64) and with an increasing mesh resolution. We timed the total runtime of the point integral solves, both for the forward solves and the adjoint solves. For each cell model and resolution, we ran three experiments and extracted the minimal run times. The results are listed in Table 7. For both the forward and adjoint run times, we computed the parallel efficiency (PE) as the ratio of the 1-CPU core runtime to the n -CPU core run time for $n = 4, 16, 64$, also listed in Table 7. The optimal PE would be 1. The last column of the table lists the computed the adjoint-to-forward ratio for these solves.

For the forward runtimes, we observe that the parallel efficiencies range from 0.66 to 0.95 for all cell models and resolutions, with efficiencies higher than 0.8 for all but the least complex cell model. In general, we observe that the parallel efficiency increases with increasing cell model complexity and that it decreases moderately with the number of CPU cores. Some of the loss of parallel efficiency (from 1) is attributable to only a near-perfect distribution of mesh vertices across CPU cores; we observed approximately 10% variation between CPU cores in the number of local vertices.

For the adjoint runtimes, we observe PEs ranging from 0.75 (FHN on 64 CPU cores) to 0.95, again with efficiencies higher than 0.8 for all but the least complex cell model. In general, we note the same trends as for the forward run times: increasing parallel efficiency with increasing cell model complexity and moderately decreasing parallel efficiency with increasing number of CPU cores. We emphasize that we thus

TABLE 7

Weak scaling of the point integral solver (PIS) for a cardiac electrophysiology two-dimensional test case. Forward and adjoint PIS runtimes for increasing N_x and simultaneously increasing number of CPU cores and for different cell models. Each row shows number of CPU cores, the mesh resolution, the total PIS runtime (s) for the forward solves, the PE for the point integral forward solves, the total PIS runtime (s) for the adjoint solves, the PE for the point integral adjoint solves, and the adjoint-to-forward runtime ratio. The computations were performed with $T = 1.0$, $\kappa_n = 0.1$, and GRL1 as the ODE scheme and (32) as the adjoint functional.

	No. CPU cores	N_x	Forward PIS (s)	PE	Adjoint PIS (s)	PE	Ratio
FHN	1	160	0.185		0.251		1.36
	4	320	0.253	0.73	0.311	0.81	1.23
	16	640	0.263	0.70	0.330	0.76	1.25
	64	1280	0.281	0.66	0.333	0.75	1.19
BR	1	160	0.740		0.925		1.25
	4	320	0.795	0.93	0.998	0.93	1.26
	16	640	0.868	0.85	1.076	0.86	1.24
	64	1280	0.899	0.82	1.143	0.81	1.27
TTP	1	160	1.657		2.516		1.52
	4	320	1.736	0.95	2.640	0.95	1.52
	16	640	1.886	0.88	2.939	0.86	1.56
	64	1280	1.991	0.83	3.072	0.82	1.54
GPB	1	160	2.618		3.895		1.49
	4	320	2.771	0.94	4.131	0.94	1.49
	16	640	3.059	0.86	4.553	0.86	1.49
	64	1280	3.125	0.84	4.632	0.84	1.48

achieve the *same* PE for the adjoint point integral solves as for the forward point integral solves.

Comparing the forward and adjoint runtimes, we see that the adjoint-to-forward ratio for the point integral solvers range from 1.19 to 1.56. Comparing these numbers with the adjoint-to-forward ratios of the overall ODE solves as discussed in the previous paragraphs, we observe that the adjoint-to-forward point integral solver performance is better. The adjoint-to-forward ratio is moderately increasing with cell model complexity, but is stable with respect to the number of CPU cores. This latter point again illustrates that the adjoint point integral solvers demonstrate the same PE as the forward point integral solves.

We also ran a series of strong scaling experiments on Abel for the previously considered series of cell models (FHN, BR, TTP, GPB) on an increasing number of CPU cores (1, 4, 16, 64) and with a fixed mesh resolution. We timed the total runtime of the point integral solver steps for both the forward and adjoint solves. Again, for each cell model and resolution, we ran three experiments and extracted the minimal run times. The results are listed in Table 8. For both the forward and adjoint runtimes, we computed the (strong scaling) PE as the ratio of the 1-CPU core runtime to n -times the n -CPU core runtime for $n = 4, 16, 64$, also listed in table 8. The optimal (strong scaling) PE would be 1. The last column of the table lists the computed the adjoint-to-forward ratio for these solves.

We observe that the PEs are comparable to the results from the weak scaling test case with efficiencies in the range from 0.60 to 0.94. We note that the adjoint strong PE is comparable to the forward strong PE. The adjoint-to-forward ratios for this strong scaling test are also in the same range as for the weak scaling test.

TABLE 8

Strong scaling of the point integral solver (PIS) for a cardiac electrophysiology two-dimensional test case. Forward and adjoint PIS runtimes for a fixed (fine) mesh resolution $N_x = 1280$ with an increasing number of CPU cores, for different cell models (FHN, BR, TTP, GPB). Each row show the number of CPU cores, the total PIS runtime for the forward solves, the corresponding PE, the total PIS runtime for the adjoint solves, the corresponding PE, and the adjoint-to-forward runtime ratio. The computations were performed with $T = 1.0$, $\kappa_n = 0.1$, and GRL1 as the ODE scheme and (32) as the adjoint functional.

Model	No. CPU cores	Forward PIS	PE	Adjoint PIS	PE	Ratio
FHN	1	12.874		17.039		1.32
	4	3.503	0.92	4.826	0.88	1.38
	16	1.159	0.69	1.373	0.78	1.18
	64	0.281	0.72	0.333	0.80	1.19
BR	1	48.286		59.673		1.24
	4	13.760	0.88	16.455	0.91	1.20
	16	3.525	0.86	4.307	0.87	1.22
	64	0.899	0.84	1.143	0.82	1.27
TTP	1	105.451		161.443		1.53
	4	28.003	0.94	44.613	0.90	1.59
	16	8.323	0.79	11.451	0.88	1.38
	64	1.991	0.83	3.072	0.82	1.54
GPB	1	166.582		249.980		1.50
	4	45.811	0.91	68.084	0.92	1.49
	16	17.426	0.60	25.787	0.61	1.48
	64	3.125	0.83	4.632	0.84	1.48

6.2.4. Verifying the correctness of the forward model. To verify the correctness of the forward solver, in particular the new DOLFIN features and the operator splitting implementation, we performed a convergence study based on an analytical solution. Specifically, we consider the bidomain equation, defined by (2), with $I_{\text{ion}}(v, s) = -s$, $F(v, s) = v$, $M_i = M_e = 1$, $\Omega = [0, 1] \times [0, 1]$, and $t \in [0, 1]$. For this setup, the following analytical solution exists:

$$\begin{aligned}
 (33) \quad & v(x, y) = g(x, y) \sin(t), \\
 & u(x, y) = -v(x, y)/2, \\
 & s(x, y) = -g(x, y) \cos(t), \\
 & I_s(x, y) = -2s(x, y) + \pi^2 v(x, y),
 \end{aligned}$$

with $g(x, y) = \cos(\pi x) \cos(\pi y)$. If we solve this problem numerically with the Strang splitting scheme, Crank–Nicolson discretization in time for the PDE, an at least second order scheme for the ODE, and continuous piecewise linear finite elements in space, then we expect a second-order convergence in space and time [36]. To test the spatial convergence, we chose a sufficiently small time step and solved the problem for increasingly fine mesh resolutions. The resulting errors and the convergence rates are listed in Table 9. A second-order convergence rate is observed, indicating that the spatial discretization of the forward model is implemented correctly. Similarly, the temporal convergence was tested by choosing a sufficiently small mesh resolution and solving the problem for increasingly fine time steps. The results, listed in Table 10, show the expected convergence rate, indicating that also the temporal discretization is implemented correctly.

TABLE 9

Spatial convergence test for the cardiac electrophysiology two-dimensional test case with analytical solution (33). The table shows the numerical errors $\mathcal{E}(u_h) = |u(T) - u_h(T)|_{L^2(\Omega)}$ and $\mathcal{E}(v_h) = |v(T) - v_h(T)|_{L^2(\Omega)}$ for different mesh resolutions and the associated spatial convergence rates for the manufactured solution (MMS) example. Computations performed with $T = 1$ and $\kappa_n = 0.01$ and the ESDIRK4 scheme.

N_n	$\mathcal{E}(u)$	Order	$\mathcal{E}(v)$	Order
5	1.89×10^{-2}		4.11×10^{-2}	
10	5.07×10^{-3}	1.90	1.09×10^{-2}	1.91
20	1.30×10^{-3}	1.97	2.77×10^{-3}	1.98
40	3.26×10^{-4}	1.99	6.91×10^{-4}	2.00

TABLE 10

Temporal convergence test for the cardiac electrophysiology two-dimensional test case with analytical solution (33). The table shows the numerical errors $\mathcal{E}(u_h) = |u(T) - u_h(T)|_{L^2(\Omega)}$ for different time steps and the associated temporal convergence rates for the MMS example. Computations performed with $T = 1$ and $N_x = 300$ and the ESDIRK4 scheme.

κ_n	$\mathcal{E}(u)$	Order	$\mathcal{E}(v)$	Order
1.0	5.17×10^{-3}		9.65×10^{-2}	
0.5	1.49×10^{-3}	1.80	2.56×10^{-2}	1.91
0.25	2.76×10^{-4}	2.43	6.15×10^{-3}	2.05
0.125	6.66×10^{-5}	2.05	1.52×10^{-3}	2.01

6.3. Application: Biventricular cardiac electrophysiology (three-dimensional). In this example, we aim to compute the sensitivity of the squared L^2 -norm of the transmembrane potential with respect to its initial condition, in order to illustrate the features discussed in this paper for a more complex test case. This example also aims to illustrate how different representations may seamlessly be used for the control variables; this is useful, for instance, when computing sensitivities with respect to spatially constant, regionally defined or highly resolved control variables.

Inspired by [4], we consider the monodomain variation of (2) over a mesh of a three-dimensional biventricular domain Ω with the TTP epicardial cell model [37] as detailed in section 6.2. For the monodomain equations, we replace the bidomain equations (2a)–(2b) by: find v such that

$$(34) \quad v_t - \operatorname{div} M \operatorname{grad} v = -I_{\text{ion}}(v, s) \quad \text{in } \Omega,$$

with homogeneous Neumann boundary conditions and with $v(t = 0) = v_0$. As before, we let $\chi = 140$ (mm^{-1}), $C_m = 0.01$ ($\mu\text{F}/\text{mm}^2$), and for simplicity (ignoring realistic, spatially varying fiber directions), let $M = \operatorname{diag}(g_f, g_s, g_n)$ where

$$(35) \quad g_f = 0.255, \quad g_n = 0.0775, \quad g_s = 0.0775.$$

To solve the coupled equations (34) and (4), we use a second-order Strang splitting scheme ($\theta = 0.5$), a Crank–Nicolson discretization in time and continuous piecewise linear finite elements in space for the resulting PDEs, and the first-order GRL1 scheme for the resulting ODEs. We take $\kappa_n = 0.05$ (ms). The moderately coarse mesh has mesh cell diameters in the range $[0.61, 2.70]$ (mm) and bounding box $[0.103, 40.2] \times [-15.8, 36.8] \times [-30, 13.9]$ (mm^3), with 45, 625 vertices and 236, 816 cells, and is illustrated in Figure 4.

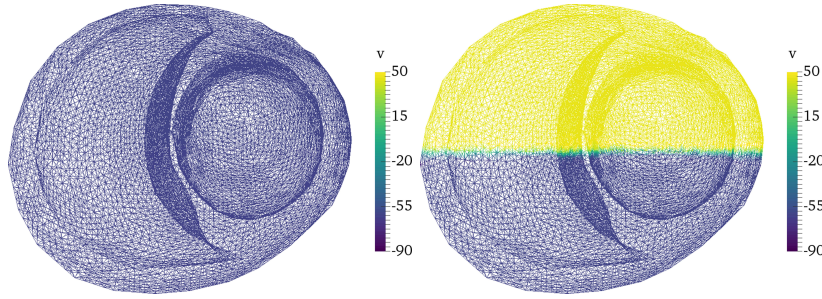


FIG. 4. Transmembrane potentials: wireframe view (along positive x -axis) of the computed transmembrane potential $v(t)$ at $t = 0$ (left) and $t = 6.0$ ms (right).

To quantify the transmembrane potential exceeding a zero threshold, we consider the following objective functional:

$$(36) \quad J(v) = \int_{\Omega_v} v(T)^2 dx$$

at a fixed time $T = 6.0$ (ms) and where $\Omega_v = \{x \in \Omega \mid v(x) \geq 0\}$. We are interested in computing the gradient of J with respect to the initial condition u_0 .

We consider two different representations of the initial condition to illustrate computing the gradient with respect to low and high resolution fields. For both cases, we let

$$(37) \quad v_0(x) = \begin{cases} -61.1, & x_2 > z_{\text{mid}}, \\ -61.3, & x_2 \leq z_{\text{mid}}, \end{cases}$$

with $z_{\text{mid}} = -8.07$, but consider (i) v_0 as represented by two constants (c_0, c_1) , i.e., the two-dimensional space \mathbb{R}^2 and (ii) v_0 represented by a continuous piecewise linear defined relative to the mesh, i.e., as a n_{verts} -dimensional space V where n_{verts} denotes the number of mesh vertices.

The initial condition and computed solution at $t = 6.0$ (ms) are illustrated in Figure 4. (With reference to the chosen perspective in this figure, we denote the subdomain where $x_2 > z_{\text{mid}}$ as the *top* part of the domain and the subdomain $x_2 \leq z_{\text{mid}}$ as the *bottom* part.)

For the case where the initial condition is represented at a low resolution by $v_0 = (c_0, c_1) \in \mathbb{R}^2$, we considered a series of experiments starting with different initial conditions $v_0^\alpha = (c_0 + \alpha, c_1 + \alpha) \in \mathbb{R}^2$ for $\alpha \in \{-0.1, -0.05, 0.0, 0.05, 0.1\}$. The resulting gradients (in \mathbb{R}^2) with respect to $v_0 \in \mathbb{R}^2$ are illustrated in Figure 5 (left). We observe that as we increase α , the gradient with respect to the top initial condition c_0 goes from large and positive to small and negative, while the gradient with respect to the bottom initial condition c_1 varies from moderately negative to close-to-zero and rapidly to large and positive. The rapid changes in the gradients illustrate the nonlinear nature of the problem at hand.

The sensitivity of the objective functional J defined by (36) with respect to the initial condition $v_0 \in V$ is illustrated in Figure 5 (right). We observe that the gradient in the upper part of the domain is large and negative, that the gradient in a boundary zone is large and positive, and the gradient in the lower part of the domain is close to zero. This indicates that infinitesimal increases/decreases in the upper part of

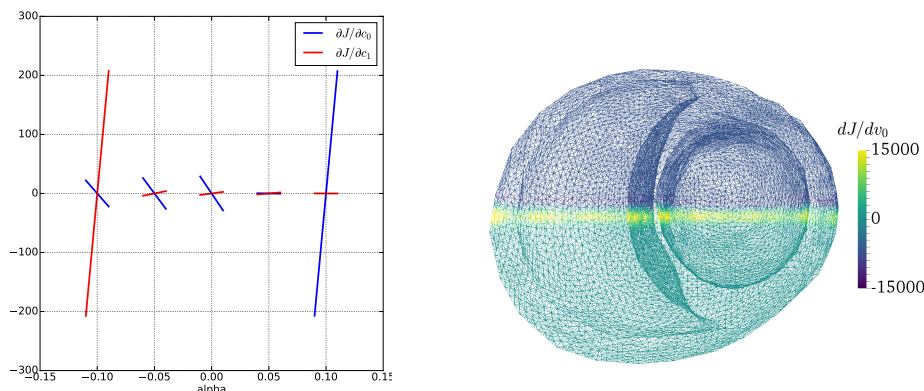


FIG. 5. Sensitivities computed with respect to low and high resolution representations of an initial transmembrane potential. Left: sensitivities $\frac{\partial J}{\partial v_0} = (\frac{\partial J}{\partial c_0}, \frac{\partial J}{\partial c_1})$ of the objective functional J defined by (36) with respect to a low resolution initial condition $v_0 = (c_0, c_1) \in \mathbb{R}^2$ for different initial conditions $v_0 = v_0^\alpha = (-61.1 + \alpha, -63.3 + \alpha)$ for $\alpha = \{-0.1, -0.05, 0.0, 0.05, 0.1\}$. For each point α and $i = 0, 1$, each plotted line segment (spanning from $\alpha - 0.01$ to $\alpha + 0.01$) has slope $\frac{\partial J}{\partial c_i}$. Right: Sensitivity $\frac{\partial J}{\partial v_0}$ of the objective functional J defined by (36) with respect to high resolution initial condition $v_0 \in V$.

the domain will lead to a large decrease/increase in the functional value, infinitesimal increases/decreases in the boundary zone domain will lead to a large increase/decrease in the functional value, and that infinitesimal increases/decreases in the lower part of the domain will have relatively small effect on the functional value.

7. Concluding remarks. We have presented a high-level framework that allows for the specification of coupled PDE-ODE systems and their efficient forward and adjoint solution with operator splitting schemes. We have illustrated the features of the framework with a series of examples originating from cell modeling and computational cardiac electrophysiology. Our numerical results indicate adjoint runtime performance near optimal (measured in terms of adjoint-to-forward runtimes), and parallel efficiency indices of 84–94% for the more realistic cardiac cell models.

The main advantages of the framework are the speed of developing solvers for new systems and the ability to automatically derive their adjoints from the high-level specification. This allows for flexible exploration of models when the relevant governing equations are uncertain and for the identification of unknown parameters via the solution of inverse problems as demonstrated, e.g., in [19]. Both of these capabilities are of significant importance in numerous areas of scientific computing and in particular, in computational physiology, chemistry, and geology. Future work will focus on further improving the parallel scalability of the solvers and its application to problems in personalized medicine.

REFERENCES

- [1] M. S. ALNÆS, J. BLECHTA, J. HAKE, A. JOHANSSON, B. KEHLET, A. LOGG, C. RICHARDSON, J. RING, M. E. ROGNES, AND G. N. WELLS, *The fenics project version 1.5*, Arch. Numer. Softw., 3 (2015), <https://doi.org/10.11588/ans.2015.100.20553>.
- [2] M. S. ALNÆS, A. LOGG, K.-A. MARDAL, O. SKAVHAUG, AND H. P. LANGTANGEN, *Unified framework for finite element assembly*, Int. J. Comput. Sci. Engrg., 4 (2009), pp. 231–244, <https://doi.org/10.1504/IJCSE.2009.029160>.

- [3] M. S. ALNÆS, A. LOGG, K. B. ØLGAARD, M. E. ROGNES, AND G. N. WELLS, *Unified form language: A domain-specific language for weak formulations of partial differential equations*, ACM Trans. Math. Softw., 40 (2014), <https://doi.org/10.1145/2566630>.
- [4] H. J. AREVALO, F. VADAKKUMPADAN, E. GUALLAR, A. JEBB, P. MALAMAS, K. C. WU, AND N. A. TRAYANOVA, *Arrhythmia risk stratification of patients after myocardial infarction using personalized heart models*, Nature Commun., 7 (2016), 11437.
- [5] R. BECKER AND R. RANNACHER, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numerica, 10 (2001), pp. 1–102, <https://doi.org/10.1017/S0962492901000010>.
- [6] G. W. BEELER AND H. REUTER, *Reconstruction of the action potential of ventricular myocardial fibres*, J. Physiology, 268 (1977), pp. 177–210.
- [7] G. J. BIGNELL AND P. R. JOHNSTON, *Split operator finite element method for modelling pulmonary gas exchange*, in Proceedings of the 13th Biennial Computational Techniques and Applications Conference, CTAC-2006, W. Read and A. J. Roberts, eds., vol. 48 of ANZIAM J., Aug. 2007, pp. C364–C380. <http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/125>.
- [8] J. C. BUTCHER, *Numerical Methods for Ordinary Differential Equations*, John Wiley & Sons, New York, 2008.
- [9] P. DEUFLHARD, *Newton Methods for Nonlinear Problems*, Springer Series in Computational Mathematics 35, Springer-Verlag, New York, 2011.
- [10] S. EISENHOFER, *A coupled system of ordinary and partial differential equations modeling the swelling of mitochondria*, PhD thesis, Technical University of Munich, 2013.
- [11] P. E. FARRELL, J. E. HAKE, S. W. FUNKE, AND M. E. ROGNES, *Supplementary code for “Automated adjoints of coupled PDE-ODE systems”*, Aug. 2017, <https://doi.org/10.5281/zenodo.843495>.
- [12] P. E. FARRELL, D. A. HAM, S. W. FUNKE, AND M. E. ROGNES, *Automated derivation of the adjoint of high-level transient finite element programs*, SIAM J. Sci. Comput., 35 (2013), pp. C369–C393, <https://doi.org/10.1137/120873558>.
- [13] R. FITZHUGH, *Impulses and physiological states in theoretical models of nerve membrane*, Biophys. J., 1 (1961), pp. 445–466.
- [14] J. GEISER, *Iterative operator splitting method for coupled problems: Transport and electric fields*, J. Informatics Math. Sci., 3 (2011), pp. 107–125.
- [15] R. GLOWINSKI, S. J. OSHER, AND W. YIN, *Splitting Methods in Communication, Imaging, Science, and Engineering*, Springer, New York, 2017.
- [16] E. GRANDI, F. S. PASQUALINI, AND D. M. BERS, *A novel computational model of the human ventricular action potential and ca transient*, J. Molecular Cellular Cardiology, 48 (2010), pp. 112–121.
- [17] A. GRIEWANK AND A. WALTHER, *Algorithm 799: Revolve: An implementation of checkpointing for the reverse or adjoint mode of computational differentiation*, ACM Trans. Math. Softw., 26 (2000), pp. 19–45.
- [18] M. D. GUNZBURGER, *Perspectives in Flow Control and Optimization*, Advances in Design and Control, SIAM, 2003.
- [19] S. KALLHOVD, M. M. MALECKAR, AND M. E. ROGNES, *Inverse estimation of cardiac ACTIVATION times via gradient-based optimisation*, Int. J. Numer. Methods Biomed. Engrg., 34 (2018), e2919, <https://doi.org/10.1002/cnm.2919>.
- [20] A. KVÆRNØ, *Singly diagonally implicit Runge–Kutta methods with an explicit first stage*, BIT, 44 (2004), pp. 489–502.
- [21] A. LOGG, K.-A. MARDAL, G. N. WELLS, ET AL., *Automated Solution of Differential Equations by the Finite Element Method*, Springer, New York, 2012, <https://doi.org/10.1007/978-3-642-23099-8>.
- [22] A. LOGG, K. B. ØLGAARD, M. E. ROGNES, AND G. N. WELLS, *FFC: the FEniCS Form Compiler*, Springer, New York, 2012.
- [23] A. LOGG AND G. N. WELLS, *Dolfin: Automated finite element computing*, ACM Trans. Math. Softw., 37 (2010), <https://doi.org/10.1145/1731022.1731030>.
- [24] J. LOTZ, U. NAUMANN, R. HANNEMANN-TAMÁS, T. PLOCH, AND A. MITSOS, *Higher-order discrete adjoint ODE solver in C++ for dynamic optimization*, Procedia Comput. Sci., 51 (2015), pp. 256–265.
- [25] R. I. MCLACHLAN AND G. R. W. QUISPTEL, *Splitting methods*, Acta Numerica, 11 (2002), pp. 341–434.
- [26] G. R. MIRAMS, C. J. ARTHURS, M. O. BERNABEU, R. BORDAS, J. COOPER, A. CORRIAS, Y. DAVIT, S.-J. DUNN, A. G. FLETCHER, D. G. HARVEY, ET AL., *Chaste: An open source C++ library for computational physiology and biology*, PLoS Comput. Biol., 9 (2013), e1002970.

- [27] S. NIEDERER, L. MITCHELL, N. SMITH, AND G. PLANK, *Simulating human cardiac electrophysiology on clinical time-scales*, *Frontiers Physiology*, 2 (2011), 14, <https://doi.org/10.3389/fphys.2011.00014>.
- [28] T. O'HARA, L. VIRÁG, A. VARRÓ, AND Y. RUDY, *Simulation of the undiseased human cardiac ventricular action potential: model formulation and experimental validation*, *PLoS Comput. Biol.*, 7 (2011), e1002061.
- [29] C. PRUD'HOMME, V. CHABANNES, V. DOYEUX, M. ISMAIL, A. SAMAKE, AND G. PENA, *Feel++: A Computational Framework for Galerkin Methods and Advanced Numerical Methods*, *ESAIM: Proceedings*, 38 (2012), pp. 429–455, <https://doi.org/10.1051/proc/201238024>.
- [30] F. RATHGEBER, D. A. HAM, L. MITCHELL, M. LANGE, F. LUPORINI, A. T. T. McRAE, G.-T. BERCEA, G. R. MARKALL, AND P. H. J. KELLY, *Firedrake: Automating the finite element method by composing abstractions*, preprint, arXiv:1501.01809, 2015, <http://arxiv.org/abs/1501.01809>.
- [31] C. N. RICHARDSON AND G. N. WELLS, *Parallel scaling of DOLFIN on ARCHER* Retrieved: 11 35, May 29, 2017 (GMT), 2015, <https://doi.org/10.6084/m9.figshare.1304537.v1>.
- [32] S. RUSH AND H. LARSEN, *A practical algorithm for solving dynamic membrane equations*, *IEEE Trans. Biomed. Eng.*, 25 (1978), pp. 389–392, <https://doi.org/10.1109/TBME.1978.326270>.
- [33] R. SERBAN AND A. C. HINDMARSH, *CVODES: The sensitivity-enabled ode solver in sundials*, in *ASME 2005 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers, Long Beach, CA, 2005, pp. 257–269.
- [34] G. STRANG, *On the construction and comparison of difference schemes*, *SIAM J. Numer. Anal.*, 5 (1968), pp. 506–517.
- [35] J. SUNDNES, R. ARTEBRANT, O. SKAVHAUG, AND A. TVEITO, *A second-order algorithm for solving dynamic cell membrane equations.*, *IEEE Trans. Biomed. Eng.*, 56 (2009), pp. 2546–2548, <https://doi.org/10.1109/TBME.2009.2014739>.
- [36] J. SUNDNES, G. T. LINES, X. CAI, B. F. NIELSEN, K.-A. MARDAL, AND A. TVEITO, *Computing the Electrical Activity in the Heart*, Springer-Verlag, New York, 2006.
- [37] K. TEN TUSSCHER AND A. PANFILOV, *Cell model for efficient simulation of wave propagation in human ventricular tissue under normal and pathological conditions*, *Phys. Med. Biol.*, 51 (2006), pp. 6141–6156.
- [38] E. VIGMOND, R. W. DOS SANTOS, A. PRASSL, M. DEO, AND G. PLANK, *Solvers for the cardiac bidomain equations*, *Progress Biophys. Molecular Biol.*, 96 (2008), pp. 3–18.
- [39] F. WANG, J. BRIGHT, AND J. HADFIELD, *Simulating nitrate transport in an alluvial aquifer: a three dimensional n-dynamics model*, *J. Hydrology*, 42 (2003), pp. 145–162.
- [40] P. J. WHITELEY, J. D. GAVAGHAN, AND E. C. HAHN, *Mathematical modelling of pulmonary gas transport*, *J. Math. Biol.*, 47 (2003), pp. 79–99, <https://doi.org/10.1007/s00285-003-0196-8>.
- [41] H. ZHANG AND A. SANDU, *FATODE: A library for forward, adjoint, and tangent linear integration of ODEs*, *SIAM J. Sci. Comput.*, 36 (2014), pp. C504–C523.