Comput. Methods Appl. Mech. Engrg. 267 (2013) 1–22

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

A posteriori analysis of an iterative multi-discretization method for reaction–diffusion systems ☆

J.H. Chaudhry [a], D. Estep [b,*], V. Ginting [c], S. Tavener [a]

[a] Department of Mathematics, Colorado State University, Fort Collins, CO 80523, United States
[b] Department of Statistics, Colorado State University, Fort Collins, CO 80523, United States
[c] Department of Mathematics, University of Wyoming, Laramie, WY 82071, United States

ARTICLE INFO

Article history:
Received 2 July 2012
Received in revised form 10 February 2013
Accepted 15 August 2013
Available online 23 August 2013

Keywords:
Reaction–diffusion
A posteriori estimates
Discontinuous Galerkin method
Multirate method
Multi-scale discretization
Operator decomposition

ABSTRACT

This paper is concerned with the accurate computational error estimation of numerical solutions of multi-scale, multi-physics systems of reaction–diffusion equations. Such systems can present significantly different temporal and spatial scales within the components of the model, indicating the use of independent discretizations for different components. However, multi-discretization can have significant effects on accuracy and stability. We perform an adjoint-based analysis to derive asymptotically accurate *a posteriori* error estimates for a user-defined quantity of interest. These estimates account for leading order contributions to the error arising from numerical solution of each component, an error due to incomplete iteration, an error due to linearization, and for errors arising due to the projection of solution components between different spatial meshes. Several numerical examples with various settings are given to demonstrate the performance of the error estimators.

1. Introduction

This paper is concerned with the accurate computational error estimation of numerical solutions of multi-scale, multi-physics systems of reaction–diffusion equations. The components of solutions of such multi-scale, multi-physics models typically exhibit spatial and temporal behavior occurring over a significant range of scales. For example, consider the well-known Brusselator model for chemical dynamics [25,1]. This is a system of reaction–diffusion equations whose separate components can behave over different spatial and temporal scales for particular choices of parameters. The model is

$$\begin{aligned}
&\dot{u}_1 - \epsilon_1 \Delta u_1 = \alpha - (\beta+1)u_1 + u_1^2 u_2, && \boldsymbol{x} \in \Omega \subset \mathbb{R}^2,\ t > 0,\\
&c\dot{u}_2 - \epsilon_2 \Delta u_2 = \beta u_1 - u_1^2 u_2, && \boldsymbol{x} \in \Omega,\ t > 0,\\
&u_1(\boldsymbol{x},t) = \alpha,\ u_2(\boldsymbol{x},t) = \beta/\alpha, && \boldsymbol{x} \in \partial\Omega,\ t > 0,\\
&u_1(\boldsymbol{x},0) = u_{1,0}(\boldsymbol{x}),\ u_2(\boldsymbol{x},0) = u_{2,0}(\boldsymbol{x}), && \boldsymbol{x} \in \Omega,
\end{aligned} \tag{1}$$

☆ J.H. Chaudhry's work is supported in part by the Department of Energy (DE-SC0005304). D. Estep's work is supported in part by the Defense Threat Reduction Agency (HDTRA1-09-1-0036), Department of Energy (DE-FG02-04ER25620, DE-FG02-05ER25699, DE-FC02-07ER54909, DE-SC0001724, DE-SC0005304, INL00120133, DE0000000SC9279), Idaho National Laboratory (00069249, 00115474), Lawrence Livermore National Laboratory (B573139, B584647, B590495), National Science Foundation (DMS-0107832, DMS-0715135, DGE-0221595003, MSPA-CSE-0434354, ECCS-0700559, DMS-1065046, DMS-1016268, DMS-FRG-1065046, DMS-1228206), National Institutes of Health (#R01GM096192). V. Ginting's work is supported in part by the National Science Foundation (DMS-1016283), the Department of Energy (DE-SC0004982).

* Corresponding author. Tel.: +1 970 491 6722; fax: +1 970 491 7895.
E-mail address: estep@stat.colostate.edu (D. Estep).

0045-7825/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cma.2013.08.007

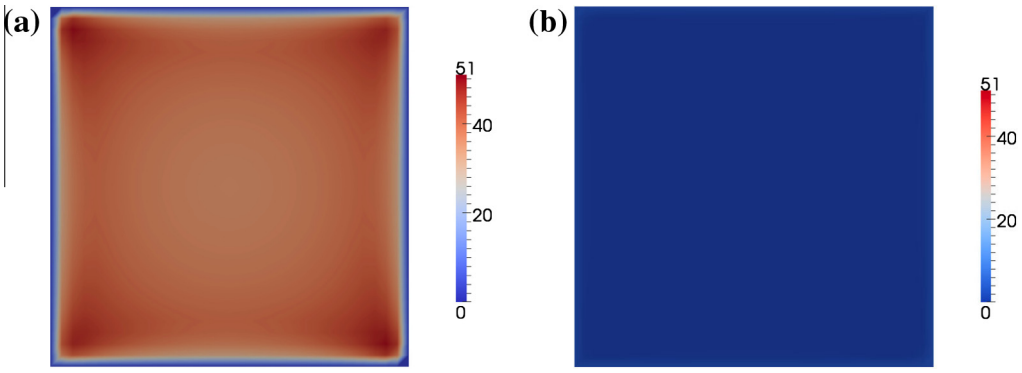

Fig. 1. Brusselator: Color contour plots of the solution at $T = 1.0$. (a) $u_1(\boldsymbol{x})$. (b) $u_2(\boldsymbol{x})$.

where u_1 and u_2 are concentrations of species 1 and 2, respectively. We assume that $0 < \epsilon_0 < \epsilon_i$, $i = 1, 2$, for some positive constant ϵ_0. The solutions are multi-scale in time and space for a wide range of parameter values. In Fig. 1 we show the solution at $t = 1.0$ corresponding to $\alpha = 2$, $\beta = 5.45$, $\epsilon_1 = 0.008$, $\epsilon_2 = 0.08$, $c = 20$ and initial conditions $u_{1,0}(\boldsymbol{x}) = \alpha + 0.1\sin(\pi x_1)\sin(\pi x_2)$ and $u_{2,0}(\boldsymbol{x}) = \beta/\alpha + 0.1\sin(\pi x_1)\sin(\pi x_2)$. We plot a cross section of the numerical solution at $x_2 = 0.25$ in Fig. 2. There are sharp spatial gradients for the component u_1, while u_2 shows relatively less spatial variation, suggesting that we might use an relatively finer mesh to resolve u_1. The time evolution of the solution at the point $\boldsymbol{x} = (0.25, 0.25)$ is also shown in Fig. 2 and indicates the multirate nature of the solutions in which u_1 is a faster component than u_2 and requires relatively fine time steps for accurate resolution.

In practical situations, the error of approximate solutions of multi-scale, multi-physics evolution models is always significant. Simply providing an *a priori* analysis of convergence and an assertion that the error is small for sufficiently refined discretizations that cannot be achieved in practice is inadequate for scientific purposes. Hence, application of numerical solution to predictive science and engineering applications requires accurate estimation of information computed from numerical solution as part of the overall uncertainty quantification critical to scientific and engineering needs.

For multi-scale problems, the demands of computational efficiency (or simple necessity) suggests a multi-discretization approach that involves solving the distinct components of a multi-physics model using independent meshes and time steps chosen to resolve behavior on the pertinent scales. A multi-discretization strategy often has significant effects on the accuracy and stability of the numerical solution. Indeed, such multi-discretization methods fall into the general class of multi-scale operator decomposition methods [11], that typically employ some form of projection to link solutions computed on different spatial and temporal meshes and necessarily "synchronize" solutions that have been decoupled during an iterative solution process. Since these practices can have a complex effect on accuracy and stability, there has been a steady development of *a posteriori* error estimates for a wide range of multi-scale operator decomposition methods in recent years [13,12,16,18,6,7,22,21,23] extending earlier work on *a posteriori* error analysis employing computable residuals and adjoint problems, see e.g. [10,8,9,15,19,5,3,4]. While the primary purpose of such estimates is to quantify the contributions of various sources of discretization error on accuracy and stability, the estimates can also provide guidance as to the choice of numerical parameters needed to obtain a desired accuracy.

The analysis of multi-discretization numerical methods for multi-scale systems of partial differential equations in this paper extends earlier results for multi-rate time integration schemes for initial value problems for ordinary differential equations in [14]. For simplicity, we consider a system comprised of two reaction–diffusion equations: Find $u = (u_1 u_2)^\top$ that satisfies

$$\begin{aligned}
&\dot{u}_1 - \nabla\cdot(\epsilon_1\nabla u_1) = f_1(u_1,u_2), && (\boldsymbol{x},t)\in\Omega\times(0,T],\\
&\dot{u}_2 - \nabla\cdot(\epsilon_2\nabla u_2) = f_2(u_1,u_2), && (\boldsymbol{x},t)\in\Omega\times(0,T],\\
&u_i(\boldsymbol{x},t) = 0, && (\boldsymbol{x},t)\in\partial\Omega\times(0,T],\ i=1,2,\\
&u_i(\boldsymbol{x},0) = g_i(\boldsymbol{x}), && \boldsymbol{x}\in\Omega,\ i=1,2,
\end{aligned} \tag{2}$$

where Ω is a convex polygonal domain with boundary $\partial\Omega$, $\{f_i\}$ are differentiable functions of their arguments, $\{\epsilon_i\}$ and $\{g_i\}$ are smooth functions in Ω, and there is a constant $\epsilon_0 > 0$ such that $\epsilon_i \geqslant \epsilon_0 > 0$ on Ω. Finally, we also assume that

$$f_i(0) = 0, \quad i = 1, 2. \tag{3}$$

The latter assumption is used to define the adjoint problems employed for the *a posteriori* error analysis carried out in Section 4. The ideas and results extend to systems consisting of more than two equations in a straightforward way. Condition (3)

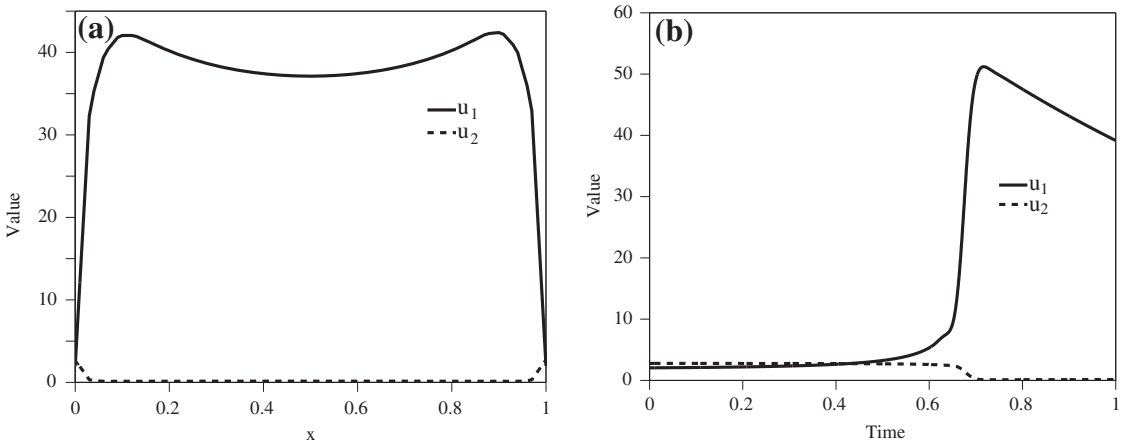

Fig. 2. Brusselator. (a) Spatial cross section of the solution at $x_2 = 0.25$ and $T = 1.0$. (b) Temporal cross section of the solution at $\boldsymbol{x} = (0.25, 0.25)$.

can also be generalized, see Section 4. Finally, neglecting the vastly more difficult questions of existence, uniqueness, and regularity for the problem, the analysis also extends to problems with nonlinear diffusion constants, and we show the formal result in Section 7.

Whenever appropriate, we write the differential equations in a compact form

$$\dot{u} - \nabla \cdot (\epsilon \nabla u) = f(u),$$

where $\epsilon = \mathrm{diag}(\epsilon_1, \epsilon_2)$, $\nabla u = [\nabla u_1 \quad \nabla u_2]^\top$, and $f(u) = [f_1(u) \quad f_2(u)]^\top$. The diffusion coefficients, ϵ_1 and ϵ_2, and reaction terms f_1 and f_2 may induce different spatial and temporal properties for u_1 and u_2. We adopt a multi-discretization approach in which each component model is solved on its own scale. In order to facilitate this approach, we compute the solution using a common iterative approach in which each component model is solved while fixing the other component solutions. The individual component solves are synchronized by exchanging information at designated "synchronization" times. At each synchronization time, component exchanges are iterated a specified number of times before the solution proceeds to the next synchronization time.

In this paper, we derive accurate *a posteriori* error estimates for a quantity of interest obtained from a numerical solution computed using the iterative multi-discretization scheme. The estimates account for *leading order* contributions to the error arising from numerical solution of each component, multi-discretization, and iterative solution. The estimates quantify the relative size of the various contributions to the error. We demonstrate the accuracy of the estimates on a variety of examples.

The rest of the paper is organized as follows. In Section 2, we formulate an iterative multi-discretization Galerkin finite element method for (2). In Section 3, we formulate an *analytic* version of (2) that we use for the purpose of analysis. We present the first results of an analysis for the multi-discretization solution method in Section 4 followed by numerical examples in Section 5. In Section 6, we expand the analysis to include the effects of using different space meshes for the two components. We also give numerical results for the Brusselator problem in this section. Finally, in Section 7 we consider the analysis for systems in which the diffusion coefficient may depend on the solution.

2. An iterative multi-discretization Galerkin finite element method

In Algorithm 1 we formulate the iterative multi-discretization Galerkin finite element method for (2). We first discretize $[0, T]$ into $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$ with time steps $\{\Delta t_n = t_n - t_{n-1}\}_{n=1}^N$, $\Delta t = \max_{1 \leqslant n \leqslant N} \{\Delta t_n\}$ and time intervals $I_n = [t_{n-1}, t_n]$. We think of $\{t_n\}$ as synchronization times during which information between the two component solves interior to the nodes is exchanged iteratively. To each t_n, we assign a positive integer M_n which is the number of iterations to be used when synchronizing the fast and slow components.

To solve the components over each synchronization interval, we divide the intervals $\{I_n\}$ into a number of smaller time steps. We let $L_{i,n}$, $i = 1, 2$ be two positive integers, where $L_{1,n}$ denotes the number of time steps used to solve the subsystem 1 and $L_{2,n}$ the number of steps used for subsystem 2 on each synchronization interval. Without loss of generality, we assume $L_{1,n} = d_n L_{2,n}$ for some positive integer d_n, i.e., $L_{1,n}$ is divisible by $L_{2,n}$. We denote time steps for each component in the Galerkin formulation by $\Delta s_{i,n} = \Delta t_n / L_{i,n}$, with $\Delta s_i = \max_{1 \leqslant n \leqslant N} \{\Delta s_{i,n}\}$. We use an extension of the discontinuous Galerkin method [15]. The method naturally extends to the continuous Galerkin method [15].

To construct the finite dimensional spaces, we first discretize Ω into triangulations $\mathcal{T}_{h_i}$, where h_i denotes the maximum diameter of the elements of $\mathcal{T}_{h_i}$, $i = 1, 2$, i.e., each equation has different triangulation. Each of these triangulations is arranged in such a way that the union of the elements of $\mathcal{T}_{h_i}$ is Ω, and the intersection of any two elements is either a common edge, node, or is empty.

The approximations are polynomials in time and continuous piecewise polynomials in space on each space–time slab $S_{l,n} = \Omega \times I_{l,n}$, for $l = 1, \ldots, L_{1,n}$ and $S_{k,n} = \Omega \times I_{k,n}$, for $k = 1, \ldots, L_{2,n}$. Here $I_{l,n} = [t_{n-1} + (l-1)\Delta s_{1,n}, t_{n-1} + l\Delta s_{1,n}]$ and $I_{k,n} = [t_{n-1} + (k-1)\Delta s_{2,n}, t_{n-1} + k\Delta s_{2,n}]$ are the smaller time intervals. In space, we let $V_{h_i} \subset H_0^1(\Omega)$ denote the space of continuous piecewise polynomial functions $v(\boldsymbol{x}) \in \mathbb{R}$ defined on $\mathcal{T}_{h_i}$. (For simplicity we confine our attention to problems with homogeneous Dirichlet boundary conditions). On each slab, we define

$$W_{l,n}^{q_1} = \left\{ w(\boldsymbol{x}, t) : w(\boldsymbol{x}, t) = \sum_{j=0}^{q_1} t^j v_j(\boldsymbol{x}),\ v_j \in V_{h_1},\ (\boldsymbol{x}, t) \in S_{l,n} \right\},$$

$$W_{k,n}^{q_2} = \left\{ w(\boldsymbol{x}, t) : w(\boldsymbol{x}, t) = \sum_{j=0}^{q_2} t^j v_j(\boldsymbol{x}),\ v_j \in V_{h_2},\ (\boldsymbol{x}, t) \in S_{k,n} \right\}.$$

We denote the jump across t_n by $[w]_n = w_n^+ - w_n^-$, where $w_n^{\pm} = \lim_{s \to t_n^{\pm}} w(s)$. We let $\Pi_{1\to 2} : W_{l,n}^{q_1} \to W_{k,n}^{q_2}$, $\Pi_{2\to 1} : W_{k,n}^{q_2} \to W_{l,n}^{q_1}$ denote projections between the two spaces. The iterative discontinuous Galerkin dG (q) finite element approximation is written down in Algorithm 1. In the algorithm, $U^{(m)} = [U_1^{(m)}, U_2^{(m)}]^\top \in W_{l,n}^{q_1} \times W_{k,n}^{q_2}$ are the finite element solutions, defined locally on time intervals $I_{l,n}$ and $I_{k,n}$. The notation (a, b) denotes the L^2 inner product, or simply the spatial integral, $\int_\Omega a\, b\, d\boldsymbol{x}$.

Algorithm 1. Iterative multi-discretization Galerkin finite element method

Set $U^{(M_0)}(\cdot, t_0^-) = u(\cdot, t_0)$
for $n = 1$ to N **do**
  Set $U_2^{(0)} = U_2^{(M_{n-1})}(\cdot, t_{n-1})$
  for $m = 1$ to M_n **do**
    Set $U^{(m)}(\cdot, t_{n-1}^-) = U^{(M_{n-1})}(\cdot, t_{n-1}^-)$
    for $l = 1$ to $L_{1,n}$ **do**
      Compute $U_1^{(m)} \in W_{l,n}^{q_1}$ satisfying

$$\int_{I_{l,n}} \left(\dot{U}_1^{(m)} - f_1(U_1^{(m)}, \Pi_{2\to 1} U_2^{(m-1)}), V \right) dt + \int_{I_{l,n}} \left(\epsilon_1 \nabla U_1^{(m)}, \nabla V \right) dt + \left([U_1^{(m)}]_{l-1,n}, V_{l-1}^+ \right) = 0 \tag{4}$$

        for all $V \in W_{l,n}^{q_1}$
    end for
    for $k = 1$ to $L_{2,n}$ **do**
      Compute $U_2^{(m)} \in W_{k,n}^{q_2}$ satisfying

$$\int_{I_{k,n}} \left(\dot{U}_2^{(m)} - f_2(\Pi_{1\to 2} U_1^{(m)}, U_2^{(m)}), Z \right) dt + \int_{I_{k,n}} \left(\epsilon_2 \nabla U_2^{(m)}, \nabla Z \right) dt + \left([U_2^{(m)}]_{k-1,n}, Z_{k-1}^+ \right) = 0 \tag{5}$$

        for all $Z \in W_{k,n}^{q_2}$
    end for
  end for
end for

3. An analytic iterative method

The approach to the *a posteriori* analysis of the multi-discretization finite element approximation in Algorithm 1 we use in this paper starts with defining an iterative method to determine an *analytic* solution of (2) obtained via a sequence of functions $\{u_i^{(m)}(t)\}$ that map the time intervals to the Banach space $X = L^2(\Omega)$, i.e., $u_i^{(m)}(t) : [t_{n-1}, t_n] \times X \to X$ for $i = 1, 2$. The iterative method defining $\{u_i^{(m)}\}$ is given in Algorithm 2.

Algorithm 2. Analytic iterative method

for $n = 1$ to N **do**
  Set $u_2^{(0)} = u_2^{(M_{n-1})}(\cdot, t_{n-1})$
  for $m = 1$ to M_n **do**
    Compute $u_1^{(m)}(\boldsymbol{x}, t)$ in $\Omega \times I_n$ satisfying

$$\begin{cases} \dot{u}_1^{(m)} - \nabla \cdot \left(\epsilon_1 \nabla u_1^{(m)}\right) = f_1(u_1^{(m)}, u_2^{(m-1)}), & (\boldsymbol{x}, t) \in \Omega \times I_n, \\ u_1^{(m)}(\boldsymbol{x}, t) = 0, & (\boldsymbol{x}, t) \in \partial\Omega \times I_n, \\ u_1^{(m)}(\boldsymbol{x}, t_{n-1}) = u_1^{(M_{n-1})}(\boldsymbol{x}, t_{n-1}), & \boldsymbol{x} \in \Omega. \end{cases} \tag{6}$$

    Compute $u_2^{(m)}(\boldsymbol{x}, t)$ in $\Omega \times I_n$ satisfying

$$\begin{cases} \dot{u}_2^{(m)} - \nabla \cdot \left(\epsilon_2 \nabla u_2^{(m)}\right) = f_2\left(u_1^{(m)}, u_2^{(m)}\right), & (\boldsymbol{x}, t) \in \Omega \times I_n, \\ u_2^{(m)}(\boldsymbol{x}, t) = 0, (\boldsymbol{x}, t) \in \partial\Omega \times I_n, & \\ u_2^{(m)}(\boldsymbol{x}, t_{n-1}) = u_2^{(M_{n-1})}(\boldsymbol{x}, t_{n-1}), & \boldsymbol{x} \in \Omega. \end{cases} \tag{7}$$

  end for
end for

The accuracy of the computational error estimate derived below assumes that the analytic iteration has converged to a sufficient extent and the discretization error is sufficiently small. The following assumptions provide sufficient general conditions to guarantee convergence of $u_i^{(m)}$ to u_i, $i = 1, 2$:

Assumption A.1. Assume that $f(t, u) : [t_{n-1}, t_n] \times X \times X \to X \times X$ is uniformly Lipschitz continuous with constant L, i.e.

$$\|f(t, u) - f(t, v)\|_{X\times X} \leqslant L\|u - v\|_{X\times X} \quad \forall t \geqslant 0. \tag{8}$$

Similarly, we assume that $f'(u)$ is uniformly Lipschitz continuous with constant L'.'

Assumption A.2. Let M be the bound on the semigroup G associated with (2) (defined in Appendix). We assume that the time steps Δt_n satisfy the inequality,

$$ML\Delta t_n \exp(ML\Delta t_n) < 1. \tag{9}$$

The convergence proof is given in the Appendix. We note that these are sufficient conditions to guarantee convergence of the iteration. They are not necessary and the iteration may converge in specific cases without satisfying these assumptions. Our *a posteriori* analysis assumes the iteration is convergent and employs the Lipschitz assumptions, but does not specifically depend on the bound on the semigroup.

The motivation for introducing the analytic iterative solution method is the realization that the iterative multi-discretization Galerkin finite element method in Algorithm 1 is a consistent finite element space–time discretization of Algorithm 2. In particular, in (4) and (5) we have chosen piecewise space–time polynomials that solve the weak or variational formulation of (6) and (7) respectively. The variational formulation is obtained by multiplying each (6) and (7) by appropriate test functions, integrating over space and time, and using Green's formula on the elliptic part. In practice, we evaluate the finite element function using quadrature to approximate the associated integral, which results in a set of discrete equations.

4. *A posteriori* analysis of the iterative multi-discretization Galerkin finite element method

We derive computational *a posteriori* error estimates based on variational analysis, residuals of the finite element approximation, and the generalized Green's function solving the adjoint problem [8,10,9,15,19,5,3,11,4]. We first develop the analysis assuming the same spatial meshes for both components. We relax this restriction in Section 6 where we include the effect of projection between different spatial meshes.

A key feature of the analysis is the realization that the iterative multi-discretization approximation is naturally associated with a different adjoint operator than that for the original problem. For this reason, we use a different linearization than commonly employed for nonlinear problems [13]. We assume that the operators for the original problem and the analytic operator decomposition version share a common solution, and use that as a linearization point for determining the stability properties of solutions in the neighborhood of the linearization point. The simplest example is to assume a common steady-

state solution such as 0, which is guaranteed by the homogeneity assumption (3), i.e., $f(0) = 0$. This assumption is employed in many standard analyses of the model (2) and it is satisfied in a great many cases. The condition can be generalized (see [13]), e.g. to other steady state solutions or to a given function of time. We give an example of a system (Brusselator) that uses an alternative condition in Section 6 [13]. We let

$$\overline{f'_{ij}(u)} = \int_0^1 \frac{\partial f_i}{\partial u_j}(su)\,ds, \quad i,j = 1,2 \tag{10}$$

and $\overline{f'(u)}$ denotes the square matrix whose entries are (10). Then $f(u) = \overline{f'(u)}u$. Associated with this linearized form, we denote by φ, the generalized Green's function satisfying the following adjoint problem:

$$\begin{cases} -\dot{\varphi} - \nabla\cdot(\epsilon\nabla\varphi) = \overline{f'(u)}^{\top}\varphi, & (\boldsymbol{x},t) \in \Omega\times(T,0], \\ \varphi(\boldsymbol{x},t) = 0, & (\boldsymbol{x},t) \in \partial\Omega\times(T,0], \quad \epsilon\nabla\varphi = \begin{pmatrix} \epsilon_1\nabla\varphi_1 \\ \epsilon_2\nabla\varphi_2 \end{pmatrix}. \\ \varphi(\boldsymbol{x},T) = \psi(\boldsymbol{x}), & \boldsymbol{x}\in\Omega, \end{cases} \tag{11}$$

On subinterval $I_n = (t_{n-1}, t_n)$, we define the solution operators Φ_n associated with the Green's function,

$$\varphi(x,t) = \Phi_n(t)\psi_n(x)$$

for $t_n > t \geqslant t_{n-1}$ and some initial data ψ_n. To get solution representation using the Green's functions, we multiply u with (11), integrate in time and space, resulting in

$$(u_n,\psi_n) = (u_{n-1},\varphi_{n-1}) + \int_{I_n}\left(\dot{u} - \nabla\cdot(\epsilon\nabla u) - \overline{f'(u)}u, \varphi\right)dt = (u_{n-1},\varphi_{n-1}) + \int_{I_n}(\dot{u} - \nabla\cdot(\epsilon\nabla u) - f(u),\varphi)\,dt. \tag{12}$$

Because u solves (2), this last equality gives

$$(u_n,\psi_n) = (u_{n-1},\Phi_n\psi_n). \tag{13}$$

4.1. Analysis of the analytic iterative method

To simplify presentation, we express the analytic iterative method in Algorithm 2 in a more compact format. In particular, for any iteration index m, we write (6) and (7) as

$$\dot{u}^{(m)} - \nabla\cdot\left(\epsilon\nabla u^{(m)}\right) = f\left(u^{(m)}\right) + \delta_R^{(m)}, \quad \delta_R^{(m)} = \begin{pmatrix} -\left(f_1\left(u_1^{(m)},u_2^{(m)}\right) - f_1\left(u_1^{(m)},u_2^{(m-1)}\right)\right) \\ 0 \end{pmatrix}. \tag{14}$$

The vector $\delta_R^{(m)}$ can be interpreted as residuals at the iteration level m.

To define an adjoint for the approximation in Algorithm 2, we let φ_i denote the generalized Green's function that satisfies an adjoint problem on time interval I_n as given in Algorithm 3. Here K_n refers to the number of iterations to be used when synchronizing the two components of the adjoint.

Algorithm 3. Adjoint for the analytic iterative method

Set $\varphi_1^{(0)} = \psi_{1,n}$
for $k = 1$ to K_n **do**
Compute $\varphi_2^{(k)}(\boldsymbol{x},t)$ in $\Omega\times(t_n,t_{n-1}]$, satisfying

$$\begin{cases} -\dot{\varphi}_2^{(k)} - \nabla\cdot(\epsilon_2\cdot\nabla\varphi_2^{(k)}) = \overline{f'_{22}(u^{(m)})}\varphi_2^{(k)} + \overline{f'_{12}(u^{(m)})}\varphi_1^{(k-1)}, & (\boldsymbol{x},t)\in\Omega\times(t_n,t_{n-1}], \\ \varphi_2^{(k)}(\boldsymbol{x},t) = 0, & (\boldsymbol{x},t)\in\partial\Omega\times(t_n,t_{n-1}], \\ \varphi_2^{(k)}(\boldsymbol{x},t_n) = \psi_{2,n}(\boldsymbol{x}), & \boldsymbol{x}\in\Omega. \end{cases} \tag{15}$$

Compute $\varphi_1^{(k)}(\boldsymbol{x},t)$ in $\Omega\times(t_n,t_{n-1}]$, satisfying

$$\begin{cases} -\dot{\varphi}_1^{(k)} - \nabla\cdot\left(\epsilon_1\nabla\varphi_1^{(k)}\right) = \overline{f'_{11}(u^{(m)})}\varphi_1^{(k)} + \overline{f'_{21}(u^{(m)})}\varphi_2^{(k)}, & (\boldsymbol{x},t)\in\Omega\times(t_n,t_{n-1}], \\ \varphi_1^{(k)}(\boldsymbol{x},t) = 0, & (\boldsymbol{x},t)\in\partial\Omega\times(t_n,t_{n-1}], \\ \varphi_1^{(k)}(\boldsymbol{x},t_n) = \psi_{1,n}(\boldsymbol{x}), & \boldsymbol{x}\in\Omega. \end{cases} \tag{16}$$

end for

Notice that the adjoint problems are solved backward in time and in the reverse order to that of the forward problem, starting with φ_2 followed by φ_1. These generalized Green's functions are an iterative approximation of (11). We note that the coefficients $\overline{f'_{ij}(u^{(m)})}$ are linearized around $u^{(m)}$. As in the forward problem, we can also rewrite this last algorithm into a compact form

$$-\dot{\varphi}^{(k)} - \nabla\cdot(\epsilon\nabla\varphi^{(k)}) = \overline{f'(u^{(m)})}^{\top}\varphi^{(k)} + \xi_R^{(k)}, \quad \xi_R^{(k)} = \begin{pmatrix} 0 \\ -\left(\overline{f'_{12}(u^{(m)})}\left(\varphi_1^{(k)} - \varphi_1^{(k-1)}\right)\right) \end{pmatrix} \tag{17}$$

for adjoint iteration level k. Here, $\xi_R^{(k)}$ is the residual of the adjoint at iteration level k. We also introduce the solution operators $\Phi_n^{(k)}$, with $\varphi^{(k)}(\boldsymbol{x},t) = \Phi_n^{(k)}(t)\psi_n(\boldsymbol{x})$, for $t_n > t \geqslant t_{n-1}$. To get a representation of the iterative solution, we follow a similar derivation for the fully coupled problem (see (12)). Multiplying Eq. (17) by $u^{(m)}$, integrating each over $\Omega\times I_n$ and applying integration by parts in time, and Green's Theorem in space and using (14), we obtain the solution representation of the analytic iterative method

$$(u_n^{(m)},\psi_n) = (u_{n-1}^{(m)},\Phi_n^{(k)}\psi_n) + \int_{I_n}\left(\delta_R^{(m)},\varphi^{(k)}\right)dt - \int_{I_n}\left(u^{(m)},\xi_R^{(k)}\right)dt. \tag{18}$$

We note that this representation is not in the standard format (in which the solution at the current time level solely depends on the previous time level values). It contains remnants arising from the iterative procedure used to compute both forward and backward problems. The second term can be interpreted as the weighted average of the forward problem residual over a time step. The third term, on the other hand, is the weighted average of the backward problem residual over a time step. Thus, the iterative nature of solution procedure is reflected in this representation. Once convergence is reached both on forward and backward problems, then the standard convention of solution representation using the adjoint technique is recovered.

We are now able to express the error representation of the iterative implicit method. Let $\hat{e}_n^{(m)} = u_n - u_n^{(m)}$. Now, we state a lemma concerning an error equation over one time step.

Lemma 4.1. *The analytic iterative method satisfies the following error equation over one time step:*

$$(\hat{e}_n^{(m)},\psi_n) = (u_n - u_n^{(m)},\psi_n) = \left(\hat{e}_{n-1}^{(m)},\Phi_n^{(k)}\psi_n\right) + \left(\hat{e}_{n-1}^{(m)},\Delta\Phi_n\psi_n\right) + \left(u_{n-1}^{(m)},\Delta\Phi_n\psi_n\right) - \int_{I_n}\left(\delta_R^{(m)},\varphi^{(k)}\right)dt + \int_{I_n}\left(u^{(m)},\xi_R^{(k)}\right)dt,$$

where $\Delta\Phi_n = (\Phi_n - \Phi_n^{(k)})$.

Proof. Subtracting (18) from (13), and adding and subtracting $(u_{n-1}^{(m)},\Phi_n\psi_n)$,

$$\begin{aligned}(\hat{e}_n^{(m)},\psi_n) &= (u_n - u_n^{(m)},\psi_n)\\ &= (u_{n-1},\Phi_n\psi_n) - \left(u_{n-1}^{(m)},\Phi_n\psi_n\right) + \left(u_{n-1}^{(m)},\Phi_n\psi_n\right) - \left(u_{n-1}^{(m)},\Phi_n^{(k)}\psi_n\right) - \int_{I_n}\left(\delta_R^{(m)},\varphi^{(k)}\right)dt + \int_{I_n}\left(u^{(m)},\xi_R^{(k)}\right)dt\\ &= \left(u_{n-1} - u_{n-1}^{(m)},\Phi_n\psi_n\right) + \left(u_{n-1}^{(m)},\Delta\Phi_n\psi_n\right) - \int_{I_n}\left(\delta_R^{(m)},\varphi^{(k)}\right)dt + \int_{I_n}\left(u^{(m)},\xi_R^{(k)}\right)dt.\end{aligned}$$

Adding and subtracting $\left(\hat{e}_{n-1}^{(m)},\Phi_n^{(k)}\psi_n\right)$ to above equation completes the proof. □

4.2. Analysis of the iterative multi-discretization Galerkin finite element method

To construct the adjoint, let $z^{(m)} = su^{(m)} + (1-s)U^{(m)}$, with $s\in[0,1]$. Then let $\overline{f'(z^{(m)})}$ be a matrix whose entries are

$$\overline{f'(z^{(m)})}_{ij} = \int_0^1 \frac{\partial f_i}{\partial u_j}(z^{(m)})\,ds.$$

Consequently, $f(u^{(m)}) - f(U^{(m)}) = \overline{f'(z^{(m)})}(u^{(m)} - U^{(m)})$.

Associated with the finite element solution, we denote by ϑ the generalized Green's function that satisfies the adjoint problem in Algorithm 4.

Algorithm 4. Adjoint for the iterative multi-discretization Galerkin finite element method

Set $\vartheta_1^{(0)} = \psi_{1,n}$

for $k = 1$ to K_n **do**

Compute $\vartheta_2^{(k)}(\boldsymbol{x}, t)$ in $\Omega \times (t_n, t_{n-1}]$ satisfying

$$\begin{cases} -\dot{\vartheta}_2^{(k)} - \nabla \cdot \left(\epsilon_2 \nabla \vartheta_2^{(k)}\right) = \overline{f'_{22}(z^{(m)})}\vartheta_2^{(k)} + \overline{f'_{12}(z^{(m)})}\vartheta_1^{(k-1)}, & (\boldsymbol{x}, t) \in \Omega \times (t_n, t_{n-1}], \\ \vartheta_2^{(k)}(\boldsymbol{x}, t) = 0, & (\boldsymbol{x}, t) \in \partial\Omega \times (t_n, t_{n-1}], \\ \vartheta_2^{(k)}(\boldsymbol{x}, t_n) = \psi_{2,n}(\boldsymbol{x}), & \boldsymbol{x} \in \Omega. \end{cases} \tag{19}$$

Compute $\vartheta_1^{(k)}(\boldsymbol{x}, t)$ in $\Omega \times (t_n, t_{n-1}]$ satisfying

$$\begin{cases} -\dot{\vartheta}_1^{(k)} - \nabla \cdot \left(\epsilon_1 \nabla \vartheta_1^{(k)}\right) = \overline{f'_{11}(z^{(m)})}\vartheta_1^{(k)} + \overline{f'_{21}(z^{(m)})}\vartheta_2^{(k)}, & (\boldsymbol{x}, t) \in \Omega \times (t_n, t_{n-1}], \\ \vartheta_1^{(k)}(\boldsymbol{x}, t) = 0, & (\boldsymbol{x}, t) \in \partial\Omega \times (t_n, t_{n-1}], \\ \vartheta_1^{(k)}(\boldsymbol{x}, t_n) = \psi_{1,n}(\boldsymbol{x}), & \boldsymbol{x} \in \Omega. \end{cases} \tag{20}$$

end for

As was the case in the adjoint formulation associated with analytic iterative method, this algorithm can be expressed as a compact form

$$-\dot{\vartheta}^{(k)} - \nabla \cdot (\epsilon \nabla \vartheta^{(k)}) = \overline{f'(z^{(m)})}^{\top} \vartheta^{(k)} + \eta_R^{(k)}, \quad \eta_R^{(k)} = \begin{pmatrix} 0 \\ -\left(\overline{f'_{12}(z^{(m)})}\left(\vartheta_1^{(k)} - \vartheta_1^{(k-1)}\right)\right) \end{pmatrix}. \tag{21}$$

Here, $\eta_R^{(k)}$ is the residual of the adjoint at iteration level k.

At this stage, we are in position to derive an error equation associated with the iterative multi-discretization Galerkin finite element method. Let $\tilde{e}^{(m)} = u^{(m)} - U^{(m)}$. First notice that using integration by parts,

$$\left(\epsilon \nabla u^{(m)} - \epsilon \nabla U^{(m)}, \nabla \vartheta^{(k)}\right) = \left(\epsilon \nabla \tilde{e}^{(m)}, \nabla \vartheta^{(k)}\right) = \left(\tilde{e}^{(m)}, -\nabla \cdot \left(\epsilon \nabla \vartheta^{(k)}\right)\right).$$

Similarly,

$$\left(f\left(u^{(m)}\right) - f\left(U^{(m)}\right), \vartheta^{(k)}\right) = \left(\overline{f'(z^{(m)})}\tilde{e}^{(m)}, \vartheta^{(k)}\right) = \left(\tilde{e}^{(m)}, \overline{f'(z^{(m)})}^{\top}\vartheta^{(k)}\right).$$

Furthermore, using continuity of $u^{(m)}$,

$$\tilde{e}_{l-1,n}^{(m)+} = u_{l-1,n}^{(m)+} - U_{l-1,n}^{(m)+} = \left(u_{l-1,n}^{(m)-} - U_{l-1,n}^{(m)-}\right) - \left(U_{l-1,n}^{(m)+} - U_{l-1,n}^{(m)-}\right) = \tilde{e}_{l-1,n}^{(m)-} - \left[U^{(m)}\right]_{l-1,n}.$$

We use these three expressions on time interval $I_{l,n}$, $l = 1, 2, \ldots, L_{1,n}$, to obtain

$$\begin{aligned} 0 &= \int_{I_{l,n}} \left(\tilde{e}^{(m)}, \dot{\vartheta}^{(k)} + \nabla \cdot \left(\epsilon \nabla \vartheta^{(k)}\right) + \overline{f'(z^{(m)})}^{\top}\vartheta^{(k)} + \eta_R^{(k)}\right) dt \\ &= \left(\tilde{e}_{l,n}^{(m)-}, \vartheta_{l,n}^{(k)}\right) - \left(\tilde{e}_{l-1,n}^{(m)+}, \vartheta_{l-1,n}^{(k)}\right) + \int_{I_{l,n}} \left(-\dot{\tilde{e}}^{(m)} + f\left(u^{(m)}\right) - f\left(U^{(m)}\right), \vartheta^{(k)}\right) dt + \int_{I_{l,n}} \left(\epsilon \nabla U^{(m)} - \epsilon \nabla u^{(m)}, \nabla \vartheta^{(k)}\right) dt \\ &\quad + \int_{I_{l,n}} \left(\tilde{e}^{(m)}, \eta_R^{(k)}\right) dt. \end{aligned}$$

Hence,

$$\begin{aligned} 0 &= \left(\tilde{e}_{l,n}^{(m)-}, \vartheta_{l,n}^{(k)}\right) - \left(\tilde{e}_{l-1,n}^{(m)-} - \left[U^{(m)}\right]_{l-1,n}, \vartheta_{l-1,n}^{(k)}\right) + \int_{I_{l,n}} \left(\dot{U}^{(m)} - f\left(U^{(m)}\right), \vartheta^{(k)}\right) dt + \int_{I_{l,n}} \left(\epsilon \nabla U^{(m)}, \nabla \vartheta^{(k)}\right) dt \\ &\quad + \int_{I_{l,n}} \left(\tilde{e}^{(m)}, \eta_R^{(k)}\right) dt + \int_{I_{l,n}} \left(-\dot{u}^{(m)} + \nabla \cdot \left(\epsilon \nabla u^{(m)}\right) + f\left(u^{(m)}\right), \vartheta^{(k)}\right) dt. \end{aligned} \tag{22}$$

Rearranging the terms in (22) and using (14) we obtain a recursive relation

$$\left(\tilde{e}_{l,n}^{(m)-},\vartheta_{l,n}^{(k)}\right)=\left(\tilde{e}_{l-1,n}^{(m)-},\vartheta_{l-1,n}^{(k)}\right)-\left(\left[U^{(m)}\right]_{l-1,n},\vartheta_{l-1,n}^{(k)}\right)-\int_{I_{l,n}}\left[\left(\dot{U}^{(m)}-f\left(U^{(m)}\right),\vartheta^{(k)}\right)+\left(\epsilon\nabla U^{(m)},\nabla\vartheta^{(k)}\right)\right]dt$$
$$+\int_{I_{l,n}}\left(\delta_R^{(m)},\vartheta^{(k)}\right)dt-\int_{I_{l,n}}\left(\tilde{e}^{(m)},\eta_R^{(k)}\right)dt. \tag{23}$$

This is the basis for the equation for the error at time t_n stated in the following lemma.

Lemma 4.2. *The iterative multi-discretization finite element solution satisfies an error equation over one time step:*

$$(\tilde{e}_n^{(m)-},\psi_n)=\left(\tilde{e}_{n-1}^{(m)-},\vartheta_{n-1}^{(k)}\right)+\hat{Q}_{1,n}+\hat{Q}_{2,n}-\sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\left(\tilde{e}^{(m)},\eta_R^{(k)}\right)dt+\sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\left(\delta_R\left(u^{(m)}\right)-\delta_R\left(U^{(m)}\right),\vartheta^{(k)}\right)dt$$

where

$$\hat{Q}_{1,n}=\sum_{l=1}^{L_{1,n}}\left\{\int_{I_{l,n}}\left[\left(-\dot{U}_1^{(m)}+f_1\left(U_1^{(m)},U_2^{(m-1)}\right),\vartheta_1^{(k)}\right)-\left(\epsilon_1\nabla U_1^{(m)},\nabla\vartheta_1^{(k)}\right)\right]dt-\left(\left[U_1^{(m)}\right]_{l-1,n},\vartheta_{1,l-1,n}^{(k)}\right)\right\}, \tag{24}$$

$$\hat{Q}_{2,n}=\sum_{l=1}^{L_{1,n}}\left\{\int_{I_{l,n}}\left[\left(-\dot{U}_2^{(m)}+f_2\left(U_1^{(m)},U_2^{(m)}\right),\vartheta_2^{(k)}\right)-\left(\epsilon_2\nabla U_2^{(m)},\nabla\vartheta_2^{(k)}\right)\right]dt-\left(\left[U_2^{(m)}\right]_{l-1,n},\vartheta_{2,l-1,n}^{(k)}\right)\right\}. \tag{25}$$

Proof. This is obtained by using the recursive relation (23) and applying integration by parts. □

We note that this equation reflects the error arising from the consistent finite element numerical discretization of the analytical iterative method. Similar to Lemma 4.1, this error contains the iteration residuals weighted by the adjoint $\vartheta^{(k)}$. The last term cannot be approximated easily since it involves the error $\tilde{e}^{(m)}$ weighted by the iteration residual in the adjoint computation. However, provided that an *a priori* estimate on $\tilde{e}^{(m)}$ is available, we can control this term to be relatively small due the fact that the residual can be made as small as needed when the adjoint computation is driven to convergence.

We now collect all the results above and obtain an error representation of the finite element multi-scale iterative implicit method by setting $e^{(m)}=u-U^{(m)}=(u-u^{(m)})+(u^{(m)}-U^{(m)})=\hat{e}^{(m)}+\tilde{e}^{(m)}$.

Theorem 4.1. *Set $\psi_N=\psi$ and $\psi_{n-1}=\vartheta_{n-1}^{(K_n)}$ in Algorithm 4 and $\psi_{n-1}=\varphi_{n-1}^{(K_n)}$ in Algorithm 3, for $n=N,N-1,\ldots,1$. Then the error of iterative multi-discretization finite element solution at final time $t_N=T$ can be expressed as*

$$\left(e_N^{(M_n)-},\psi_N\right)=\left(u_N-U_N^{(M_N)-},\psi\right)=\sum_{n=1}^{N}\left(\hat{Q}_{1,n}+\hat{Q}_{2,n}+\hat{Q}_{3,n}+\hat{Q}_{4,n}+\hat{Q}_{5,n}+\hat{Q}_{6,n}\right), \tag{26}$$

$\hat{Q}_{1,n}$ and $\hat{Q}_{2,n}$ are given in Lemma 4.2 with $m=M_n$ and

$$\hat{Q}_{3,n}=\sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\left(-\delta_R(U^{(M_n)}),\vartheta^{(K_n)}\right)dt$$
$$\hat{Q}_{4,n}=\sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\left(\delta_R(u^{(M_n)}),\vartheta^{(K_n)}-\varphi^{(K_n)}\right)dt$$
$$\hat{Q}_{5,n}=\left(u_{n-1}^{(M_n)},\Delta\Phi_n\psi_n\right)+\int_{I_n}\left(u^{(M_n)},\xi_R^{(K_n)}\right)dt$$
$$\hat{Q}_{6,n}=\left(\hat{e}_{n-1}^{(M_n)},\Delta\Phi_n\psi_n\right)-\sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\left(\tilde{e}^{(M_n)},\eta_R^{(K_n)}\right)dt,$$

Proof. First we estimate the error over one time step. Combining Lemma 4.2 and with Lemma 4.1 we get,

$$(e_n^{(M_n)-},\psi_n)=\left(\tilde{e}_{n-1}^{(M_n)-},\vartheta_{n-1}^{(K_n)}\right)+\left(\hat{e}_{n-1}^{(M_n)},\phi_{n-1}^{(K_n)}\right)+\hat{Q}_{1,n}+\hat{Q}_{2,n}+\hat{Q}_{3,n}+\hat{Q}_{4,n}+\hat{Q}_{5,n}+\hat{Q}_{6,n}. \tag{27}$$

We note that since $U_{n-1}^{(M_n)-}=U_{n-1}^{(M_{n-1})-}$ and $u_{n-1}^{(M_n)}=u_{n-1}^{(M_{n-1})}$ (see Algorithm 1), we have $\tilde{e}_{n-1}^{(M_n)-}=\tilde{e}_{n-1}^{(M_{n-1})-}$ and $\hat{e}_{n-1}^{(M_n)}=\hat{e}_{n-1}^{(M_{n-1})}$. This yields a recursive relation in terms of $\tilde{e}^{(M_n)-}$ and $\hat{e}^{(M_n)}$ for the total error over one time step. The error at the final time is obtained from undoing this relation and assuming $\tilde{e}_0^{M_0-}=\hat{e}_0^{M_0}=0$. □

The terms $\hat{Q}_{5,n}$ and $\hat{Q}_{6,n}$ are not easy to approximate. However, provided the discretization error and the iteration error are sufficiently small, $\hat{Q}_{5,n}$ and $\hat{Q}_{6,n}$ are asymptotically small compared with $\hat{Q}_{1,n},\ldots,\hat{Q}_{4,n}$.

Theorem 4.2. *The terms $\sum_{n=1}^{N}\hat{Q}_{5,n}$ and $\sum_{n=1}^{N}\hat{Q}_{6,n}$ are asymptotically small compared with $\sum_{n=1}^{N}\hat{Q}_{1,n},\ldots,\sum_{n=1}^{N}\hat{Q}_{4,n}$ in the limit of iteration errors $\|u^{(M_n)}-u\|_{L^\infty(I_n;L^2(\Omega))}$ and $\|\varphi^{(K_n)}-\varphi\|_{L^\infty(I_n;L^2(\Omega))}$ tending to zero for all n.*

Proof. Of the two, $\hat{Q}_{5,n}$ is more difficult to estimate. Let $\hat{\varphi}$ be the solution of

$$\begin{cases} -\dot{\hat{\varphi}}-\nabla\cdot(\epsilon\nabla\hat{\varphi})=\overline{f'(u^{(M_n)})}^{\top}\hat{\varphi}, & (\boldsymbol{x},t)\in\Omega\times(t_n,t_{n-1}],\\ \hat{\varphi}(\boldsymbol{x},t)=0, & (\boldsymbol{x},t)\in\partial\Omega\times(t_n,t_{n-1}],\\ \hat{\varphi}(\boldsymbol{x},t_n)=\psi_n(\boldsymbol{x}), & \boldsymbol{x}\in\Omega. \end{cases} \tag{28}$$

Notice that (28) and (11) differ only in terms of linearization point for $\overline{f'}$. Now we write

$$\varphi-\varphi^{(K_n)}=(\varphi-\hat{\varphi})+\left(\hat{\varphi}-\varphi^{(K_n)}\right)=\alpha+\beta,$$

where α and β satisfy, respectively

$$-\dot{\alpha}-\nabla\cdot(\epsilon\nabla\alpha)=\overline{f'(u^{(M_n)})}^{\top}\alpha+\delta_{f'}^{\top}\varphi, \tag{29}$$

$$-\dot{\beta}-\nabla\cdot(\epsilon\nabla\beta)=\overline{f'(u^{(M_n)})}^{\top}\beta-\xi_R^{(K_n)}, \tag{30}$$

with zero initial and boundary conditions, and we designate the 2×2 matrix $\delta_{f'}=\overline{f'(u)}-\overline{f'(u^{(M_n)})}$. Multiplying (29) by α, and following by integration over $(t,t_n)\times\Omega$ yields

$$\begin{aligned}\|\alpha(t)\|^2 &\leqslant \|\alpha(t)\|^2+2\int_t^{t_n}(\epsilon\nabla\alpha,\nabla\alpha)\,d\tau=2\int_t^{t_n}\left(\alpha,\overline{f'(u^{(M_n)})}\alpha\right)d\tau+2\int_t^{t_n}(\varphi,\delta_{f'}\alpha)\,d\tau\\ &\leqslant 2L\int_t^{t_n}\|\alpha\|^2\,d\tau+2\int_t^{t_n}\int_\Omega|\varphi||\delta_{f'}||\alpha|\,d\tau,\end{aligned} \tag{31}$$

where $\|\cdot\|$ is the norm in $L^2(\Omega)\times L^2(\Omega)$, and $|\cdot|$ is understood as the usual Euclidean vector norm for φ and α, or its corresponding matrix norm for $\delta_{f'}$. There is a constant $C_\varphi<\infty$ such that $\|\varphi_i\|_{L^\infty(t,t_n,L^2(\Omega))}<C_\varphi$, see for example [20]. We apply the Cauchy–Schwarz and arithmetic–geometric mean inequalities to the last term on the right hand side of (31) to get

$$\|\alpha(t)\|^2\leqslant 2L\int_t^{t_n}\|\alpha\|^2\,d\tau+\int_t^{t_n}\|\delta_{f'}\|^2\,d\tau+C_\varphi^2\int_t^{t_n}\|\alpha\|^2\,d\tau=\int_t^{t_n}\|\delta_{f'}\|^2\,d\tau+(2L+C_\varphi^2)\int_t^{t_n}\|\alpha\|^2\,d\tau. \tag{32}$$

Gronwall's inequality then implies

$$\|\alpha(t)\|^2\leqslant\exp\left((2L+C_\varphi^2)(t_n-t)\right)\int_t^{t_n}\|\delta_{f'}\|^2\,d\tau. \tag{33}$$

Similarly, we get

$$\|\beta(t)\|^2\leqslant\exp\left((2L)(t_n-t)\right)\int_t^{t_n}\|\xi_R^{(K_n)}\|^2\,d\tau. \tag{34}$$

Next, we multiply $u^{(M_n)}$ to (29) and (30), respectively, integrate each of them over I_n, and apply integrations by parts, and use (14) to get

$$\begin{aligned}\left(u_{n-1}^{(M_n)},\alpha_{n-1}\right)-\int_{I_n}\left(u^{(M_n)},\delta_{f'}^{\top}\varphi\right)dt&=\int_{I_n}\left(\alpha,\delta_R^{(M_n)}\right)dt,\\ \left(u_{n-1}^{(M_n)},\beta_{n-1}\right)+\int_{I_n}\left(u^{(M_n)},\xi_R^{(K_n)}\right)dt&=\int_{I_n}\left(\beta,\delta_R^{(M_n)}\right)dt\end{aligned} \tag{35}$$

and thus

$$\begin{aligned}\hat{Q}_{5,n}&=\int_{I_n}\left(u^{(M_n)},\delta_{f'}^{\top}\varphi\right)dt+\int_{I_n}\left(\alpha,\delta_R^{(M_n)}\right)dt+\int_{I_n}\left(\beta,\delta_R^{(M_n)}\right)dt\\ &\leqslant\int_{I_n}\left(\delta_{f'}u^{(M_n)},\varphi\right)dt+\frac{1}{2}\int_{I_n}(\|\alpha\|^2+\|\beta\|^2)\,dt+\frac{1}{2}\int_{I_n}\|\delta_R^{(M_n)}\|^2\,dt\\ &\leqslant\int_{I_n}\left(\overline{f'(u)}u^{(M_n)}-f(u^{(M_n)}),\varphi\right)dt+\frac{1}{2}\int_{I_n}\exp\left(C(t_n-t)\right)\int_t^{t_n}\left(\|\delta_{f'}\|^2+\|\xi_R^{(K_n)}\|^2\right)d\tau dt+\frac{1}{2}\int_{I_n}\|\delta_R^{(M_n)}\|^2\,dt.\end{aligned} \tag{36}$$

Notice that the second term and third terms in (36) involve integration of the square of residuals. Thus these terms are asymptotically small compared to $\hat{Q}_{j,n}$, $j=1,\ldots,4$ as $u^{(M_n)}$ converges to u. Moreover, the first integral in (36) dominates the other terms. We show this term is asymptotically small compared to $\hat{Q}_{3,n}$ as $u^{(M_n)}$ converges to u.

First we bound the term $\hat{Q}_{3,n}$. From Assumption A.1 we have,

$$\hat{Q}_{3,n}=\sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\left(f_1\left(U_1^{(M_n)},U_2^{(M_n)}\right)-f_1\left(U_1^{(M_n)},U_2^{(M_n-1)}\right),\vartheta_1^{(K_n)}\right)dt\leqslant L\sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\|U_2^{(M_n)}-U_2^{(M_n-1)}\|\,\|\vartheta_1^{(K_n)}\|\,dt. \tag{37}$$

Now,

$$\sum_{l=1}^{L_{1,n}} \int_{I_{l,n}} \|U_2^{(M_n)} - U_2^{(M_n-1)}\|\,dt \leqslant \sum_{l=1}^{L_{1,n}} \int_{I_{l,n}} \|U_2^{(M_n)} - u_2^{(M_n)}\|dt + \sum_{l=1}^{L_{1,n}} \int_{I_n} \|u_2^{(M_n)} - u_2^{(M_n-1)}\|dt + \sum_{l=1}^{L_{1,n}} \int_{I_{l,n}} \|u_2^{(M_n-1)} - U_2^{(M_n-1)}\|dt. \quad (38)$$

Now, the first and the third terms are discretization errors, and depend on the order of the numerical method, say $\rho(\Delta t, h)$ for some homogeneous function ρ. From (68) in the proof of Theorem A.1, $\|u_2^{(M_n)} - u_2^{(M_n-1)}\| = O(\tau^{M_n})$, for some $\tau < 1$, and

$$\int_{I_n} \|u_2^{(M_n)} - u_2^{(M_n-1)}\|\,dt = O(\tau^{M_n+1}). \quad (39)$$

Combining this with (37) and (38), we have,

$$\hat{Q}_{3,n} = O(\rho(\Delta t, h) + \tau^{M_n+1}). \quad (40)$$

We now return to estimate the first term in (36). Noting that $\overline{f'(u^{(M_n)})}u^{(M_n)} = f(u^{(M_n)})$ and by the assumption that f' is Lipschitz continuous with constant L', we have,

$$\int_{I_n} \left(\overline{f'(u)}u^{(M_n)} - f(u^{M_n}), \varphi\right) dt = \int_{I_n} \left(\left(\overline{f'(u)} - \overline{f'(u^{(M_n)})}\right)u^{(M_n)}, \varphi\right) dt \leqslant \int_{I_n} L'\|u - u^{(M_n)}\|\,\|\phi\|\,dt. \quad (41)$$

An analysis of the semigroup associated with the problem similar to that used in the Appendix to derive (68) yields $\|u - u^{(M_n)}\| = O(\tau^{M_n+1})$. Combining this with (41) and using appropriate scaling we have,

$$\int_{I_n} \left(\overline{f'(u)}u^{(M_n)} - f\left(u^{M_n}\right), \varphi\right) dt = O(\tau^{M_n+2}). \quad (42)$$

Hence, $\hat{Q}_{5,n}$ is asymptotically smaller than $\hat{Q}_{3,n}$ as $u^{(M_n)}$ converges to u.

Turning to $\hat{Q}_{6,n}$, we note that it is a sum of two terms. The first term is a product of iteration errors for the forward and adjoint problems, and is straightforward to bound as smaller than $\hat{Q}_{j,n},\ j = 1, \ldots, 4$ as the iterations converge. The second term in $\hat{Q}_{j,n},\ j = 1, \ldots, 4$ is a product of discretization error and iteration residual in the adjoint. This is bounded smaller than $\hat{Q}_{j,n},\ j = 1, o, 4$ by an argument similar to that used for analogous expressions in $\hat{Q}_{5,n}$. □

4.3. A computational error estimate

The error representation in Theorem 4.1 contains terms involving the true continuum solution $u^{(M_n)}$ as well as the true adjoint solutions $\varphi^{(K_n)}$ and $\vartheta^{(K_n)}$. We form a computational error estimate by approximating the adjoint solutions, $\varphi^{(K_n),h}$ and $\vartheta^{(K_n),h}$, in a finite dimensional space. These adjoint problems are approximated by substituting the finite element solution $U^{(M_n)}$ for $u^{(M_n)}$, as is common in adjoint based error estimation literature. Further, the term $\hat{Q}_{4,n}$ is expressed as,

$$\hat{Q}_{4,n} = \sum_{l=1}^{L_{1,n}} \int_{I_{l,n}} \left(\delta_R\left(U^{(M_n)}\right), \vartheta^{(K_n)} - \varphi^{(K_n)}\right) dt + \left(\delta_R(u^{(M_n)}) - \delta_R\left(U^{(M_n)}\right), \vartheta^{(K_n)} - \varphi^{(K_n)}\right) dt \quad (43)$$

Here the term $\left(\delta_R(u^{(M_n)}) - \delta_R(U^{(M_n)}), \vartheta^{(K_n)} - \varphi^{(K_n)}\right)$ is a product of difference of two residuals, and hence we drop it in the computational error estimate. This leads to the following computational error estimate.

Theorem 4.3. *The error of the iterative multi-discretization finite element solution at final time $t_N = T$ can be approximated as,*

$$\left(e_N^{(M_n)-}, \psi_N\right) = \left(u_N - U_N^{(M_N)-}, \psi\right) \approx \sum_{n=1}^{N} (Q_{1,n} + Q_{2,n} + Q_{3,n} + Q_{4,n}), \quad (44)$$

where,

$$Q_{1,n} = \sum_{l=1}^{L_{1,n}} \left\{ \int_{I_{l,n}} \left[\left(-\dot{U}_1^{(m)} + f_1(U_1^{(m)}, U_2^{(m-1)}), \vartheta_1^{(K_N),h}\right) - \left(\epsilon_1 \nabla U_1^{(m)}, \nabla \vartheta_1^{(K_n),h}\right)\right] dt - \left([U_1^{(m)}]_{l-1,n}, \vartheta_{1,l-1,n}^{(K_n),h}\right) \right\},$$

$$Q_{2,n} = \sum_{l=1}^{L_{1,n}} \left\{ \int_{I_{l,n}} \left[\left(-\dot{U}_2^{(m)} + f_2(U_1^{(m)}, U_2^{(m)}), \vartheta_2^{(K_n),h}\right) - \left(\epsilon_2 \nabla U_2^{(m)}, \nabla \vartheta_2^{(K_n),h}\right)\right] dt - \left([U_2^{(m)}]_{l-1,n}, \vartheta_{2,l-1,n}^{(K_n),h}\right) \right\},$$

$$Q_{3,n} = \sum_{l=1}^{L_{1,n}} \int_{I_{l,n}} \left(-\delta_R(U^{(M_n)}), \vartheta^{(K_n),h}\right) dt$$

$$Q_{4,n} = \sum_{l=1}^{L_{1,n}} \int_{I_{l,n}} \left(\delta_R(U^{(M_n)}), \vartheta^{(K_n),h} - \varphi^{(K_n),h}\right) dt$$

Table 1
Error contributions and their interpretations.

Notation	Contribution
Q_1	Discretization error in component U_1
Q_2	Discretization error in component U_2
Q_3	Iteration error for the numerical solution
Q_4	Error due to linearization in the computed adjoint problem

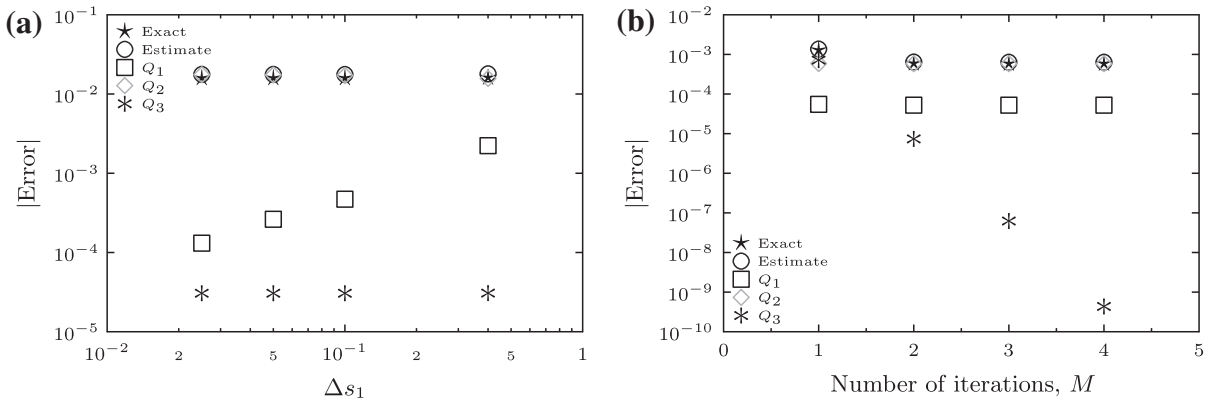

Fig. 3. Example 5.2: $T = 0.2$. (a) $\Delta t = \Delta s_2 = 0.4,\ M = 2$. Error contributions as Δs_1 is varied. (b) $\Delta t = 0.04,\ \Delta s_1 = \Delta t/32,\ \Delta s_2 = \Delta t/16$. Error contributions as M is varied.

We present interpretations of the computational error contributions in Table 1. Note that we have dropped $\hat{Q}_{5,n}$ and $\hat{Q}_{6,n}$ to obtain (44). As explained, this is reasonable provided the iteration has converged to a sufficient degree and the discretization is sufficiently refined. The examples below demonstrate the estimate (44) provides a reasonably accurate approximation of the true error.

Remark 4.1. We note that computing the error estimate (44) involves the cost of solving the adjoint problem in addition to computing the original approximation. The computational cost depends on how the numerical adjoint problem is solved, however the adjoint problem is at least linear, and hence often involves less iteration than solving the original problem.

On this issue, it is important to note that if the practical application requires an accurate error estimate to accompany a numerical solution, then the issue of cost of the error estimate has to be related to the cost of alternative approaches to error estimation. There are other ways to treat numerical solutions of coupled systems involving iteration, e.g. [18,17,16,7]. Some of these approaches provide for direct estimation of the effect of finite iteration on accuracy, at the cost of greatly increasing the number of adjoint problems that must be solved. The estimate in Theorem 4.1 is thus relatively inexpensive at the cost of assuming that the iteration has converged to a sufficient degree.

Remark 4.2. Standard adaptive error control strategies based on the Principle of Equidistribution applied to "dual-weighted" *a posteriori* estimates, [8,9,5,19,3], can be extended to (44) in a straightforward way to balance all sources of error. For example, if the component Q_1 is large, then refining the spatial and temporal mesh for the first component may lead to a more accurate solutions. A similar conclusion follows for Q_2. The terms Q_3 and Q_4 reflect errors incurred due to finite iterations, and these errors may be reduced by increasing the number of iterations. However, we note that many application codes for multi-physics problems eschew adaptive computation.

5. Numerical experiments using equal spatial meshes

In this section, we present numerical examples to illustrate the performance of the error estimates. For various problems, we show plots of the error estimate and true error accompanied by plots of the individual contributions to the error estimate, $Q_{1,n},\ Q_{2,n},\ Q_{3,n},\ Q_{4,n}$ as defined in Lemma 4.2 and Theorem 4.3. A comparison of the relative sizes of the different contributions to the error is often illuminating.

All forward problems are solved using continuous piecewise linear functions in space and using the piecewise constant discontinuous Galerkin method in time. The piecewise constant discontinuous Galerkin method, or dG (0), is

Table 2
Effectivity ratios for the experiment in Fig. 3. (a) Effectivity ratios as Δs_1 is varied. (b) Effectivity ratios as M is varied.

Δs_1	Effectivity ratio
(a) Effectivity ratios as Δs_1 is varied	
0.4	1.12
0.1	1.11
0.05	1.12
0.025	1.12
M	Effectivity ratio
(b) Effectivity ratios as M is varied	
1	1.06
2	1.09
3	1.09
4	1.09

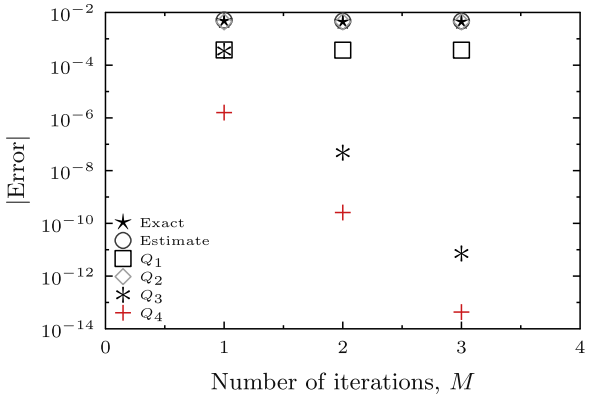

Fig. 4. Example 5.3: $T = 0.2,\ \Delta t = 0.04,\ \Delta s_1 = \Delta t/16,\ \Delta s_2 = \Delta t/2$. Error contributions as the number of iterations M is varied.

equivalent to the backward Euler scheme. The nonlinear equations are solved using Newton's Method. The adjoint solutions are approximated using continuous piecewise quadratic functions in space and piecewise linear continuous Galerkin method in time. The piecewise linear continuous Galerkin method, or cG (1), is equivalent to the second order Crank–Nicholson scheme. All problems are posed on the unit square, i.e., on $\Omega = [0,1] \times [0,1]$ and solved using a uniform mesh containing $(20 \times 20 \times 2)$ triangular elements. The initial conditions at time $t = 0$ are $u = (\sin(\pi x_1)\sin(\pi x_2),\ \sin(\pi x_1)\sin(\pi x_2))^\top$.

The quantity of interest in all cases is given by the globally supported function $\psi = (\sin(\pi x_1)\sin(\pi x_2), \sin(\pi x_1)\sin(\pi x_2))^\top$. We compare the performance of estimators using either the analytical solution when available. Otherwise we use a "reference solution" using a higher order spatial discretization and a finer time step. In our numerical results, we plot the different error components and tabulate the effectivity ratio of the estimator. The effectivity ratio is defined as the ratio of the estimated error to the true error in the quantity of interest, provided the true error is not zero. An accurate error estimator has effectivity ratio close to one.

5.1. An equal rate one-way coupled linear system

We consider the system,

$$\begin{cases} \dot{u}_1 - \Delta u_1 = \pi^2 u_1, \\ \dot{u}_2 - \Delta u_2 = \pi^2(0.5u_2 + u_1). \end{cases}$$

Notice that this is a one-way coupled system in which the variable of subsystem 1, u_1, is coupled to the variable of subsystem 2, but u_1 can be solved independently of u_2. The exact solution is $u_1 = e^{-\pi^2 t}\sin(\pi x_1)\sin(\pi x_2)$ and $u_2 = 2e^{-\pi^2 t}\sin(\pi x_1)\sin(\pi x_2)$, hence there is not a significant difference in spatial or temporal scales. Since the system is only coupled in one direction there is no need to iterate to solve the system and there is no iteration error, i.e., $Q_3 = 0$. Moreover, for linearly coupled systems $\phi = \vartheta$, and hence $Q_4 = 0$. The system is solved until $T = 0.2$ with $\Delta t = \Delta s_1 = \Delta s_2 = 0.02$. The error estimate was -0.0177161, as compared to the true error of -0.0169774 for an effectivity ratio of 1.04.

Table 3
Effectivity ratios for the experiment in Fig. 4.

M	Effectivity ratio
1	1.08
2	1.09
3	1.09

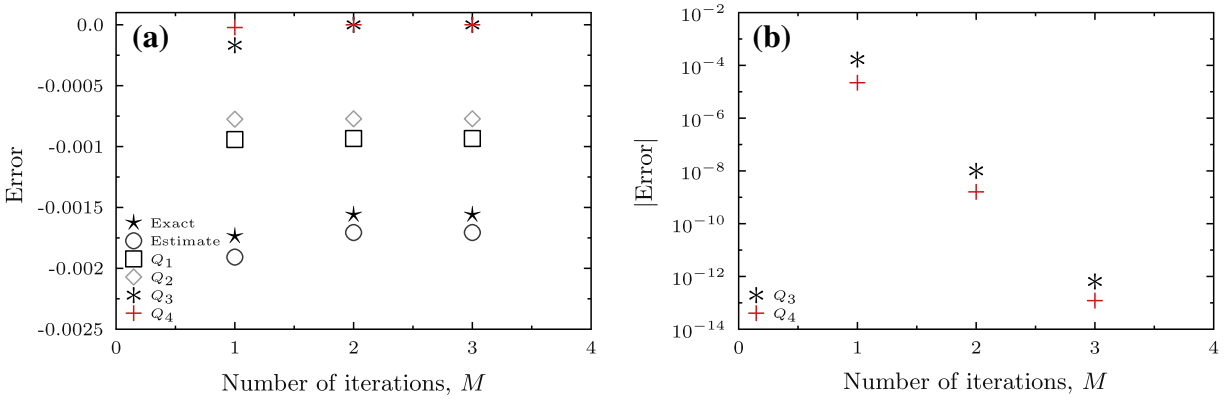

Fig. 5. Example 5.4: $T = 0.2,\ \Delta t = 0.01,\ \Delta s_1 = \Delta s_2 = \Delta t/2$. Error contributions as the number of iterations M is varied. (a) True and estimated errors. (b) Q_3 and Q_4 only.

Table 4
Effectivity ratios for the experiment in Fig. 5(a).

M	Effectivity ratio
1	1.10
2	1.09
3	1.09

5.2. A multirate coupled linear system

We consider the system,

$$\begin{cases} \dot{u}_1 - \Delta u_1 = -1000u_1 + u_2, \\ \dot{u}_2 - \Delta u_2 = 999u_1 - 2u_2. \end{cases} \tag{45}$$

Here, u_1 is a fast variable and u_2 is a slow variable. We solve until $T = 0.2$ and plot the error components as a function of Δs_1 in Fig. 3(a) while fixing $\Delta t = \Delta s_2 = 0.4$. We use two iterations at each of the time steps Δt. As expected, the error in the component Q_1 decreases as Δs_1 is reduced. In Fig. 3(b) we plot the effect of employing different number of iterations to solve the system at each time step Δt. In this case, we use $\Delta t = 0.04,\ \Delta s_1 = \Delta t/32$ and $\Delta s_2 = \Delta t/16$. The iteration error decreases as the number of iterations is increased. Except in the extreme case of just one iteration, the contribution to the error from iteration is relatively small. In all cases, the error estimator provided an accurate prediction of the exact error. We recall that for linear systems, $\phi = \vartheta$, and hence $Q_4 = 0$. The accuracy of the estimator is also illustrated in Tables 2, which show effectivity ratios close to the ideal value of 1.0.

5.3. A coupled nonlinear system using different time steps

We consider the system,

$$\begin{cases} \dot{u}_1 - \Delta u_1 = u_1^4 + u_2^2, \\ \dot{u}_2 - \Delta u_2 = u_1 - u_2^3. \end{cases} \tag{46}$$

The system is solved until $T = 0.2$, with $\Delta t = 0.04,\ \Delta s_1 = \Delta t/16$ and $\Delta s_2 = \Delta t/2$. In Fig. 4 the result of increasing the number of iterations is demonstrated. The component Q_3 is initially large, but decays to a small value after two iterations. The component Q_4 is nonzero for this problem, since the adjoints ϑ and ϕ differ from one another. However, it is quite small compared to other components. Again, we obtained very accurate error estimates. Once again, the effectivity ratios, shown in Table 3 are close to the ideal value of 1.0.

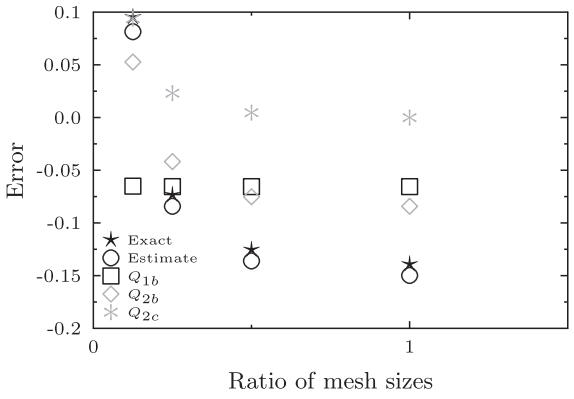

Fig. 6. Example 6.1: $T = 0.2,\ \Delta t = \Delta s_1 = \Delta s_2 = 0.01$. Error contributions versus ratio of mesh sizes.

5.4. A coupled nonlinear system using equal time steps

We consider the system,

$$\begin{cases} \dot{u}_1 - \Delta u_1 = \exp(u_1) + \exp(u_2) - 2, \\ \dot{u}_2 - \Delta u_2 = -\exp(u_1) - \exp(u_2) + 2. \end{cases} \tag{47}$$

The system is solved until $T = 0.2$, with $\Delta t = 0.01,\ \Delta s_1 = \Delta s_2 = \Delta t/2$. The effect of increasing the number of iterations is shown in Fig. 5. The component Q_3 is large after just one iteration, but contributes relatively little after two iterations. The component Q_4 is nonzero for this problem since the adjoints ϑ and ϕ differ from one another. The effectivity ratios for this experiment are shown in Table 4. The effectivity ratios are quite close to 1.0, indicating the accuracy of our estimator.

6. *A posteriori* analysis of the iterative multi-discretization Galerkin finite element method for different spatial meshes

In this section, we derive an estimate for the case in which the two subsystems in Algorithm 2 are solved on different space meshes. For such systems, we can further decompose the error components to reflect the projection errors. Solution of (4) involves the projection of $U_2^{(m-1)}$, denoted as $\Pi_{2\to1}U_2^{(m-1)}$, from $W_{l,n}^{q_1}$ to $W_{k,n}^{q_2}$. If the number of time steps are the same for the two subsystems, then $\Pi_{2\to1}$ is the projection of functions from the mesh for subsystem 1 to functions on the mesh for subsystem 2. Similarly, solution to (5) involves the projection, $\Pi_{1\to2}U_1^{(m)}$, of $U_1^{(m)}$ on the space of functions on the mesh of subsystem 2. With these projections we have the following error representation.

Theorem 6.1. *Set $\psi_N = \psi$ and $\psi_{n-1} = \vartheta_{n-1}^{K_n}$ for $n = N, N-1, \ldots, 1$. Then, with Assumptions A.1 and A.2, the error of iterative multi-discretization finite element solution at final time $t_N = T$ can be expressed as*

$$\left(u_N - U_N^{(M_N)-}, \psi\right) = \sum_{n=1}^{N} \left(\hat{Q}_{1b,n} + \hat{Q}_{1c,n} + \hat{Q}_{2b,n} + \hat{Q}_{2c,n} + \hat{Q}_{3,n} + \hat{Q}_{4,n} + \hat{Q}_{5,n} + \hat{Q}_{6,n}\right), \tag{48}$$

where $\hat{Q}_{3,n},\ \hat{Q}_{4,n},\ \hat{Q}_{5,n},\ \hat{Q}_{6,n}$ are as given in Theorem 4.3 and

$$\hat{Q}_{1b,n} = \sum_{l=1}^{L_{1,n}} \left\{ \int_{I_{l,n}} \left[\left(-\dot{U}_1^{(m)} + f_1(U_1^{(m)}, \Pi_{2\to1}U_2^{(m-1)}), \vartheta_1^{(k)} \right) - \left(\epsilon_1 \nabla U_1^{(m)}, \nabla \vartheta_1^{(k)} \right) \right] dt + \left([U_1^{(m)}]_{l-1,n}, \vartheta_{1,l-1,n}^{(k)} \right) \right\},$$

$$\hat{Q}_{1c,n} = \left(f_1(U_1^{(m)}, U_2^{(m-1)}) - f_1(U_1^{(m)}, \Pi_{2\to1}U_2^{(m-1)}), \vartheta_1^{(k)} \right)$$

$$\hat{Q}_{2b,n} = \sum_{l=1}^{L_{1,n}} \left\{ \int_{I_{l,n}} \left[\left(-\dot{U}_2^{(m)} + f_2(\Pi_{1\to2}U_1^{(m)}, U_2^{(m)}), \vartheta_2^{(k)} \right) - \left(\epsilon_2 \nabla U_1^{(m)}, \nabla \vartheta_1^{(k)} \right) \right] dt + \left([U_2^{(m)}]_{l-1,n}, \vartheta_{2,l-1,n}^{(k)} \right) \right\},$$

$$\hat{Q}_{2c,n} = \left(f_2(U_1^{(m)}, U_2^{(m)}) - f_2(\Pi_{1\to2}U_1^{(m)}, U_2^{(m)}), \vartheta_2^{(k)} \right).$$

Proof. Adding and subtracting $\left(f_1(U_1^{(m)}, \Pi_{2\to1}U_2^{(m-1)}), \vartheta_1^{(k)} \right)$ to (24),

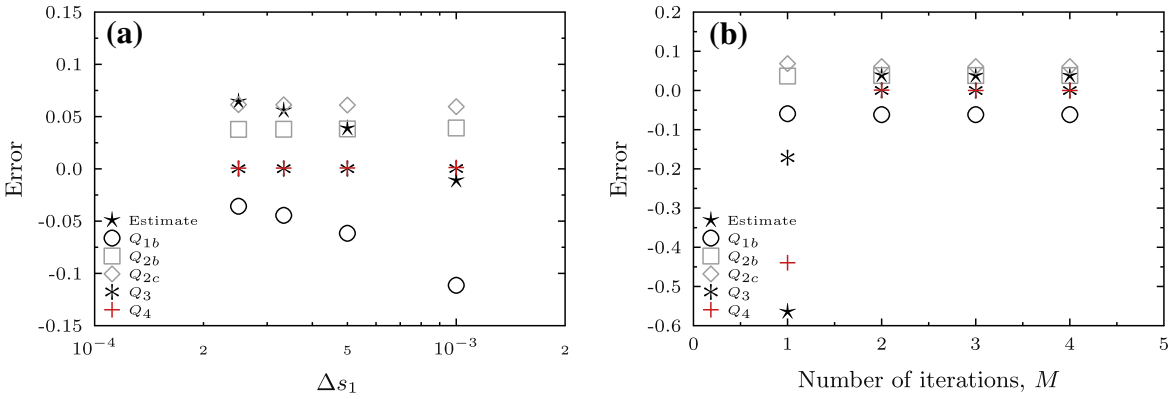

Fig. 7. Brusselator: $T = 0.7,\ \Delta t = \Delta s_2 = 0.001$. (a) $M = 2$. Error contributions as Δs_1 is varied. (b) $\Delta s_1 = 0.001$. Error contributions as M is varied.

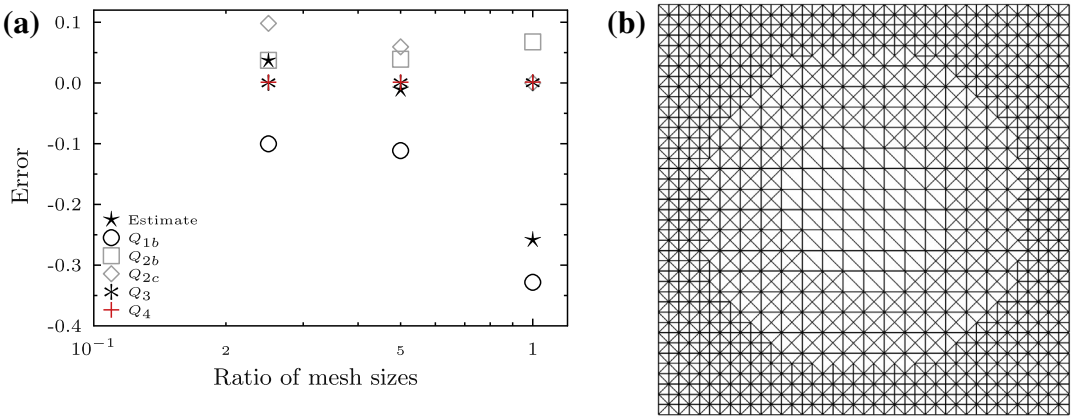

Fig. 8. Brusselator: $T = 0.7,\ \Delta t = \Delta s_1 = \Delta s_2 = 0.001,\ M = 2$. (a) Error contributions versus ratio of mesh sizes. (b) Refined mesh used to produce error contributions provided in row 3 of Table 5.

Table 5
Brusselator: $T = 0.7,\ \Delta t = \Delta s_2 = 0.001,\ \Delta s_1 = \Delta t/4 = 0.00025,\ M = 2$. Error components for two uniform and one non-uniform mesh 1. Here Dof_i refers to degrees-of-freedom for the component u_i.

Mesh	Elements	Dof_1	Dof_2	Estimate	Q_{1b}	Q_{2b}	Q_{2c}	Q_3	Q_4
Coarse uniform	800	441	441	−0.1509	−0.2139	0.0618	0.0000	0.0003	0.0008
Fine uniform	3200	1681	441	0.0644	−0.0358	0.0377	0.0614	0.0002	0.0007
Non-uniform	2320	1241	441	0.0600	−0.0511	0.0398	0.0704	0.0003	0.0007

$$\hat{Q}_{1,n} = \sum_{l=1}^{L_{1,n}} \left\{ \int_{I_{l,n}} \left[\left(-\dot{U}_1^{(m)} + f_1\left(U_1^{(m)}, \Pi_{2\to 1} U_2^{(m-1)} \right), \vartheta_1^{(k)} \right) - \left(\epsilon_1 \nabla U_1^{(m)}, \nabla \vartheta_1^{(k)} \right) \right] dt + \left(\left[U_1^{(m)} \right]_{l-1,n}, \vartheta_{1,l-1,n}^{(k)} \right) + \left(f_1\left(U_1^{(m)}, U_2^{(m-1)} \right) - f_1\left(U_1^{(m)}, \Pi_{2\to 1} U_2^{(m-1)} \right), \vartheta_1^{(k)} \right) \right.$$

Similarly, adding and subtracting $\left(f_1\left(\Pi_{1\to 2} U_1^{(m)}, U_2^{(m)} \right), \vartheta_1^{(k)} \right)$ to (25) leads to,

$$\hat{Q}_{2,n} = \hat{Q}_{2b,n} + \hat{Q}_{2c,n}.$$

Combining these with (26) leads to (48). □

For simplicity in our examples, one mesh will always be a refinement of the other mesh. Nodal projection for the space meshes is employed for the operators $\Pi_{1\to 2}$ and $\Pi_{2\to 1}$. Further, we form a computational error estimate in the manner outlined in Section 4.3, representing the approximations of the terms $\hat{Q}_{i,n}$ as $Q_{i,n}$. We recall Table 1 that describes the contributions to the error.

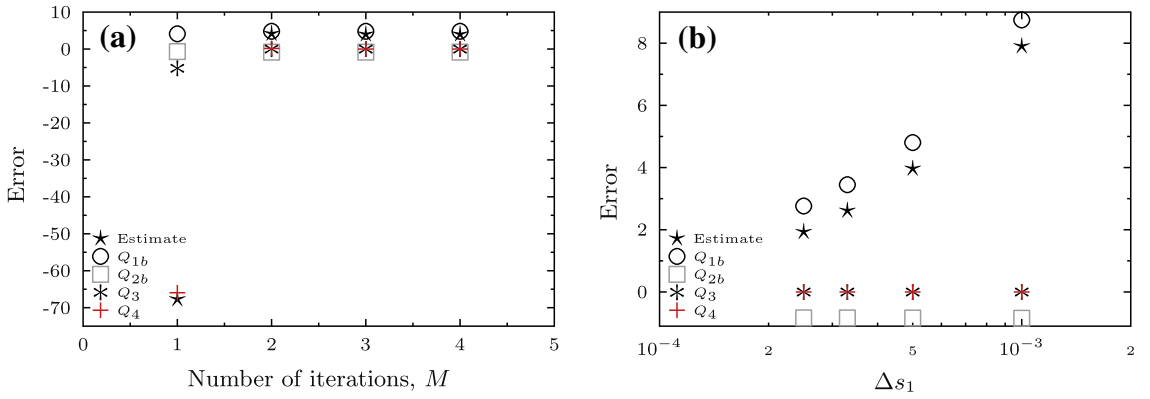

Fig. 9. Brusselator: Time derivative at $t_D = 0.7$ (final time is $T = 0.8$). (a) Effect as the number of iterations M varies. (b) Effect as Δs_1 varies given $M = 3$.

6.1. A linear system

In this section, we consider the system,

$$\begin{cases} \dot{u}_1 - \Delta u_1 = u_1, & (\boldsymbol{x}, t) \in \Omega \times (0, T], \\ \dot{u}_2 - \Delta u_2 = \mathbf{b} \cdot \nabla u_1 + u_2, & (\boldsymbol{x}, t) \in \Omega \times (0, T], \\ u_1 = 5x_1^2(1-x_1)^2(e^{10x_1^2} - 1)x_2^2(1-x_2)^2(e^{10x_2^2} - 1), & (\boldsymbol{x}, t) \in \Omega \times \{0\}, \\ u_2 = \sin(\pi x_1)\sin(\pi x_2), & (\boldsymbol{x}, t) \in \Omega \times \{0\}, \end{cases} \tag{49}$$

where $\mathbf{b} = (1000,\ 1000)^\top$. The quantity of interest is taken to be $\psi = (0,\ 100x_1^2(1-x_1)^2x_2^2(1-x_2)^2)^\top$. Note that due to the presence of the term $\mathbf{b} \cdot \nabla u_1$, the term $\overline{f'_{21}(u^{(m)})}\phi_2^{(k)}$ in Algorithm 3 is interpreted as $-\nabla \cdot (\mathbf{b}\phi_2^{(k)})$. The term $\overline{f'_{21}(z^{(m)})}\vartheta_2^{(k)}$ in Algorithm 4 is treated in a similar fashion.

In the numerical experiments, subsystem 1 is solved on a uniform mesh comprising $(40 \times 40 \times 2)$ triangular elements. The mesh for subsystem 2 is varied through $(5 \times 5 \times 2), (10 \times 10 \times 2), (20 \times 20 \times 2)$ and $(40 \times 40 \times 2)$ triangular elements and the system is solved with $\Delta t = \Delta s_1 = \Delta s_2 = 0.01$. We plot the error components as a function of the ratio of mesh sizes in Fig. 6. The figure indicates that the projection error Q_{2c} dominates the total error when there is a large difference between the mesh sizes, and goes to 0 as the two meshes have the same size.

6.2. The Brusselator

We recall the Brusselator problem (1) in Section 1. The values of different parameters are the same as in Section 1. The system as posed does not satisfy $f(\tilde{u}) = 0$. However, we use a change of variable to accomplish this. The new variables are defined as,

$$\begin{cases} u_1 = \tilde{u}_1 - \alpha, \\ u_2 = \tilde{u}_2 - \beta/\alpha. \end{cases}$$

With these new variables, $u = (u_1, u_2)^\top$, the new set of equations satisfy the requirement that $f(u) = 0$. We experiment with two different quantities of interest; a spatial quantity of interest at the final time, and a time based quantity of interest approximating the temporal derivative at a certain time.

6.2.1. A spatial quantity of interest at the final time

For this experiment, we take the quantity of interest to be

$$\psi = \begin{pmatrix} x_1^2(1-x_1)^2(\exp(6x_1^2) - 1)x_2^2(1-x_2)^2(\exp(6x_2^2) - 1) \\ 0 \end{pmatrix}$$

evaluated at final time $T = 0.7$. This quantity of interest is adapted from Ch. 8 in [2]. Mesh 1 and mesh 2 were chosen to be uniform with $(40 \times 40 \times 2)$ and $(20 \times 20 \times 2)$ triangular elements respectively, $\Delta t = \Delta s_2 = 0.001$ and $M = 2$. In Fig. 7(a) the effect of decreasing Δs_1 on the error components is evident and the error component Q_{1b} decreases as Δs_1 is reduced as expected. Note that the total error increases as Δs_1 is reduced, due to cancellation of errors with opposite sign. In Fig. 7(b) the effect of varying the number of iterations M is shown. The rest of the parameters are the same as for Fig. 7(a) except that Δs_1 is fixed at 0.001. For $M = 1$ there are significant errors in the components Q_3 and Q_4, but these errors decrease as the number of iterations is increased.

6.2.2. Competing effects of discretization and projection

Separate refinement of either of the spatial meshes may result in a reduction of discretization errors for the solution component(s) computed on that mesh, but may also increase projection errors. For this experiment, mesh 2 was held fixed with $(20 \times 20 \times 2)$ triangular elements while mesh 1 was varied having $(20 \times 20 \times 2), (40 \times 40 \times 2)$ and $(80 \times 80 \times 2)$ triangular elements. Here $\Delta t = \Delta s_1 = \Delta s_2 = 0.001$ and $M = 2$. In Fig. 8(a) the error components are plotted for this series of different discretization levels for mesh 1. Note that discretization error Q_{1b} decreased as the mesh ratio decreased (as mesh 1 was refined), but that the projection error Q_{2c} *increased*. While the magnitude of reduction of Q_{1b} exceeded the magnitude of the increase in Q_{2c}, the *total* error *increased* as mesh 1 was refined due to cancellation of errors with opposite sign.

In Table 5 we tabulate the error contributions for three different choices of mesh 1, two uniform and one non-uniform. Fig. 8(b) where the mesh is refined in regions of rapid variation of component u_1. Mesh 2 was uniform with $(20 \times 20 \times 2)$ triangular elements for all three cases. Here $\Delta t = \Delta s_2 = 0.001,\ \Delta s_1 = \Delta t/4 = 0.00025$, and $M = 2$.

When mesh 1 has $(20 \times 20 \times 2)$ uniform triangular elements, the first row of Table 5 indicates that the dominant error contribution is Q_{1b}, the discretization error on mesh 1. Halving each element on mesh 1 produces a situation in which the discretization errors on both meshes and the projection error are roughly of the same magnitude (row 2 of Table 5). Non-uniform refinement of mesh 1 such that it has a finer mesh in regions of sharp variation produces a similar distribution of error with 2/3 of the number of elements (row 3 of Table 5).

6.2.3. A temporal derivative as the quantity of interest

For this experiment, we approximate the time derivative $\frac{\partial}{\partial t}\int_\Omega u_1\,dx$ of the average value of u_1 at some $t = t_D$ using a central difference. We approximate the temporal derivative of a function v by,

$$\left.\frac{\partial v}{\partial t}\right|_{t=t_D} \approx \frac{v(t_D + 0.5\Delta t) - v(t_D - 0.5\Delta t)}{\Delta t}. \tag{50}$$

In practice we approximate the point value using a local average. That is, $v(\tau) \approx \overline{v(\tau)} = \int_{\tau-r}^{\tau+r} v(t)\,dt$. As $r \to 0,\ \overline{v(\tau)} \to v(\tau)$. The adjoint solution required a finer (time) discretization near t_D to accurately resolve the adjoint solution. Near $t = t_D$, we used a time discretization that was 100 times finer than that used for the forward problem. That is, in this region the time step is $\Delta t/100$, where Δt is the time step for the forward problem. Moreover, we chose $r = \Delta t/10$.

In Fig. 9(a) we investigate the effect of the number of iterations. We use $\Delta t = \Delta s_2 = 0.001,\ \Delta s_1 = 0.0005$ and the same uniform mesh with $(40 \times 40 \times 2)$ triangular elements for both components. For $M = 1$, the estimate is dominated by the term Q_3 and then by Q_4, which measure the effect of the number of iterations.

In Fig. 9(b) we show the effect of varying Δs_1 for fixed M. We use $\Delta t = \Delta s_2 = 0.001,\ M = 3$ and the same uniform mesh with $(40 \times 40 \times 2)$ triangular elements for both components. We see that the error in the component Q_{1b} decreases as Δs_1 is reduced, as expected. The component Q_{1b} also dominates the total error, so refining the time steps for this fast component leads to significant reduction of total error as well.

7. *A posteriori* analysis for systems with nonlinear diffusion coefficients

To explain how the *a posteriori* analysis can be extended to fully nonlinear coupled systems, we provide a formal derivation of an *a posteriori* error estimates for systems of parabolic initial boundary value problems having nonlinear diffusion coefficients. We consider the problem of finding $u = (u_1,\ u_2)^\top$ for systems in which the diffusion coefficient, $\epsilon = \epsilon(u)$ is a function of u,

$$\begin{aligned}
&\dot{u}_1 - \nabla\cdot(\epsilon_1(u_1,u_2)\nabla u_1) = f_1(u_1,u_2), && (\boldsymbol{x},t) \in \Omega\times(0,T],\\
&\dot{u}_2 - \nabla\cdot(\epsilon_2(u_1,u_2)\nabla u_2) = f_2(u_1,u_2), && (\boldsymbol{x},t) \in \Omega\times(0,T],\\
&u_i(\boldsymbol{x},t) = 0, && (\boldsymbol{x},t) \in \partial\Omega\times(0,T],\ i = 1,2\\
&u_i(\boldsymbol{x},0) = g_i(\boldsymbol{x}), && \boldsymbol{x} \in \Omega,\ i = 1,2
\end{aligned} \tag{51}$$

or in a compact form,

$$\dot{u} - \nabla\cdot(\epsilon(u)\nabla u) = f(u),$$

where $\epsilon(u) = \mathrm{diag}(\epsilon_1(u) \quad \epsilon_2(u))$. Meaningful analysis of general parabolic systems with nonlinear diffusion coefficients is very challenging, see for example [20]. Generally, analytic results can be greatly improved by employing the special properties of particular systems. We assume that the *a priori* analysis is in place and proceed to focus on the *a posteriori* analysis.

We again use Algorithm 1 to solve this system, with the obvious modification that $\epsilon = \epsilon(u)$ and for simplicity we consider a scenario of having the same spatial discretization for u_1 and u_2. The adjoint problem similar to (11) is modified to account for the dependence of ϵ on u,

$$-\dot{\varphi} - \nabla\cdot\left(\overline{\epsilon(u)}\nabla\varphi\right) + \overline{\epsilon'(u)}^\top\cdot\nabla\varphi = \overline{f'(u)}^\top\varphi, \tag{52}$$

where $\overline{\epsilon'(u)}$ denotes a square matrix and $\overline{\epsilon(u)}$ is a diagonal matrix whose entries, respectively, are

$$\overline{\epsilon'_{ij}(u)} = \left[\int_0^1 \frac{\partial \epsilon_i}{\partial u_j}(su)\nabla(u_i s)\,ds\right] \quad \text{and} \quad \overline{\epsilon_i(u)} = \int_0^1 \epsilon_i(us)ds. \tag{53}$$

In compact form, the adjoint problems similar to (17) is also modified as

$$-\dot{\varphi}^{(k)} - \nabla\cdot(\overline{\epsilon(u^{(m)})}\nabla\varphi^{(k)}) + \overline{\epsilon'(u^{(m)})}^{\top}\cdot\nabla\varphi^{(k)} = \overline{f'(u^{(m)})}^{\top}\varphi^{(k)} + \xi_R^{(k)} + \xi_D^{(k)}, \tag{54}$$

where $\xi_R^{(k)}$ is as defined in (17), $\overline{\epsilon'(u^{(m)})}$ and $\overline{\epsilon(u^{(m)})}$ are similarly defined as in (53), and

$$\xi_D^{(k)} = \left[0\overline{\epsilon'_{12}(u^{(m)})}^{\top}\cdot\nabla\left(\varphi_1^{(k)} - \varphi_1^{(k-1)}\right)\right]^{\top}.$$

Similarly, the finite element adjoint problem (21) is modified as,

$$-\dot{\vartheta}^{(k)} - \nabla\cdot\left(\overline{\epsilon(z^{(m)})}\nabla\vartheta^{(k)}\right) + \overline{\epsilon'(z^{(m)})}^{\top}\cdot\nabla\vartheta^{(k)} = \overline{f'(z^{(m)})}^{\top}\vartheta^{(k)} + \eta_R^{(k)} + \eta_D^{(k)}, \tag{55}$$

where $z^{(m)} = su^{(m)} + (1-s)U^{(m)}$, $\overline{\epsilon'(z^{(m)})}$ is a square matrix and $\overline{\epsilon(z^{(m)})}$ is a diagonal matrix whose entries, respectively, are

$$\overline{\epsilon'_{ij}(z^{(m)})} = \int_0^1 \frac{\partial \epsilon_i}{\partial u_j}(z_i{}^{(m)})\nabla z^{(m)}ds \quad \text{and} \quad \overline{\epsilon_i(z^{(m)})} = \int_0^1 \epsilon_i(z^{(m)})ds. \tag{56}$$

The residual $\eta_R^{(k)}$ is as defined in (21) and

$$\eta_D^{(k)} = \left[0\overline{\epsilon'_{12}(z^{(m)})}\cdot\nabla\left(\vartheta_1^{(k)} - \vartheta_1^{(k-1)}\right)\right]^{\top}.$$

Analysis of this system leads to the following error representation.

Theorem 7.1. *Set $\psi_N = \psi$ and $\psi_{n-1} = \vartheta_{n-1}^{K_n}$ for $n = N, N-1, \ldots, 1$. Then the error of iterative multi-discretization finite element solution of (51) at final time $t_N = T$ can be expressed as*

$$\left(u_N - U_N^{(M_N)-}, \psi\right) = \sum_{n=1}^{N}\left(\hat{Q}_{1,n} + \hat{Q}_{2,n} + \hat{Q}_{3,n} + \hat{Q}_{4,n} + \hat{Q}_{5b,n} + \hat{Q}_{6b,n} + \hat{Q}_{5c,n} + \hat{Q}_{6c,n}\right), \tag{57}$$

where $\hat{Q}_{1,n}$, $\hat{Q}_{2,n}$, $\hat{Q}_{3,n}$, $\hat{Q}_{4,n}$ are as given in Theorem 4.3, and

$$\hat{Q}_{5b,n} = \left(u_{n-1}^{(M_n)}, \Delta\Phi_n\psi_n\right) + \int_{I_n}\left(u^{(M_n)}, \xi_R^{(K_n)} + \xi_D^{(K_n)}\right)dt$$

$$\hat{Q}_{6b,n} = (\hat{e}_{n-1}, \Delta\Phi_n\psi_n) - \sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\left(\tilde{e}^{(M_n)}, \eta_R^{(K_n)} + \eta_D^{(K_n)}\right)dt$$

$$\hat{Q}_{5c,n} = \sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\left(\delta\epsilon_1(U^{(M_n)})\nabla U_1^{(M_n)}, \nabla\vartheta_1^{(K_n)}\right)dt$$

$$\hat{Q}_{6c,n} = \sum_{l=1}^{L_{1,n}}\int_{I_{l,n}}\left(\delta\epsilon_1(u^{(M_n)})\nabla u_1^{(M_n)}, \nabla\vartheta_1^{(K_n)} - \nabla\varphi_1^{(K_n)}\right)dt,$$

with $\delta\epsilon_1(u^{(m)}) = \epsilon_1\left(u_1^{(m)}, u_2^{(m)}\right) - \epsilon_1(u_1^{(m)}, u_2^{(m-1)})$, and similarly for $\delta\epsilon_1(U^{(m)})$.

A proof similar to that of Lemma 4.2 is beyond the scope of this paper. With the appropriate *a priori* analysis in place, we expect that the terms $\hat{Q}_{5b,n}$, $\hat{Q}_{5c,n}$, $\hat{Q}_{6b,n}$ and $\hat{Q}_{6c,n}$ are small compared to $\hat{Q}_{1,n}\ldots\hat{Q}_{4,n}$.

8. Conclusions

In this paper we formulate and analyze an iterative multi-discretization Galerkin finite element method for multi-scale reaction–diffusion equations. Subsystems in such reaction–diffusion equations may exhibit significantly different spatial and temporal scales, motivating a multi-discretization numerical method. We employ adjoint operators and variational analysis to form computational error estimates for a quantity of interest calculated from the multi-discretization finite element method. A key insight in analyzing the multi-discretization method is the realization that the adjoint operator associated with the iterative multi-discretization approximation is different from that of the original problem. Hence, our analysis utilizes two adjoint operators. One of the operators utilizes a different linearization than the one commonly used for nonlinear problems. The other adjoint is based on the property that our iterative multi-discretization Galerkin finite element method is a consistent discretization of the analytic iterative method.

We derive *a posteriori* error estimates to quantify various sources of error in a quantity of interest computed from our iterative finite element method. We first derive estimates for the case when the different components of the system are solved on the same spatial mesh, and then extend the analysis to include distinct meshes. The error estimator has terms indicating errors arising from discretization of each component, finite iteration, differences between the two different adjoints

and projection. We demonstrate the accuracy of our method through a variety of numerical examples, starting from simple linear problems and ending with the non-linear multi-scale Brusselator problem. We demonstrate how refining one or both meshes or increasing the number of iterations can decrease the specific error components arising from a specific source. Hence our error estimates are useful not only for computing the total error in a quantity of interest, but also applicable in guiding an adaptive refinement strategy.

Appendix A

We prove the convergence of the iterative scheme in Algorithm 2. We consider $u_i(t)$ as functions from the interval to a Banach space X, $u_i(t) : [t_{n-1}, t_n] \times X \to X$, where $X = L^2(\Omega)$, for $i = 1, 2$. Let $f(u) : [t_{n-1}, t_n] \times X \times X \to X \times X$ be uniformly Lipschitz continuous with constant L, i.e. $\|f(u) - f(v)\|_{X\times X} \leqslant L\|u - v\|_{X\times X}\ \forall t$. Let $-A_i = -\nabla \cdot \epsilon_i \nabla$ be the infinitesimal generator of the C_0 semigroup $G_i(t),\ t \geqslant 0$, on X. For simplicity of notation, we denote $A_1 = A_2 = A$ and $G_1 = G_2 = G$. Then, based on the theory of semigroups [24], (6) and (7) on an interval $[t_{n-1}, t_n]$ are recast as,

$$u_1^{(m)}(t) = G(t - t_{n-1})u_1^{(M_{n-1})} + \int_{t_{n-1}}^{t} G(t - s)f_1(u_1^{(m)}, u_2^{(m-1)})\, ds, \tag{58}$$

$$u_2^{(m)}(t) = G(t - t_{n-1})u_2^{(M_{n-1})} + \int_{t_{n-1}}^{t} G(t - s)f_2(u_1^{(m)}, u_2^{(m)})\, ds. \tag{59}$$

Let M denote the bound on $\|G(t)\|$ on $[0, T]$. We then have,

Lemma A.1. *With Assumptions A.1 and A.2, the integral equation*

$$\xi(t) = G(t - t_{n-1})\alpha + \int_{t_{n-1}}^{t} G(t - s)f_i(\xi, \beta)\, ds \tag{60}$$

admits a unique solution (ξ, β).

Proof. The proof follows arguments used for ordinary differential equations [14] and employs techniques from [24]. Set $\xi^{(0)} = \alpha$ and compute

$$\xi^{(j)} = G(t - t_{n-1})\alpha + \int_{t_{n-1}}^{t} G(t - s)f_i\left(\xi^{(j-1)}, \beta\right) ds \tag{61}$$

for $j = 1, 2, \ldots$. For $j = 1$ we have,

$$\begin{aligned}\|\xi(t) - G(t - t_{n-1})\alpha\| &= \left\| \int_{t_{n-1}}^{t} G(t - s)f_i(\alpha, \beta)\, ds \right\|, = \left\| \int_{t_{n-1}}^{t} G(t - s)(f_i(\alpha, \beta) - f_i(0, 0))\, ds \right\|, \quad (\text{since} \quad f_i(0) = 0) \\ &\leqslant \Delta t_n ML\|(\alpha, \beta)\|.\end{aligned}$$

Moreover, using a semigroup property (cf. page 5 in [24]),

$$\|G(t - t_{n-1})\alpha - \alpha\| = \left\| \int_0^{t-t_{t_{n-1}}} G(s)A\alpha\, ds \right\| \leqslant \Delta t_n M\|A\alpha\|.$$

Using the above results and the triangle inequality,

$$\|\xi^{(1)}(t) - \alpha\| \leqslant \|\xi^{(1)}(t) - G(t - t_{n-1})\alpha\| + \|G(t - t_{n-1})\alpha - \alpha\| \leqslant \Delta t_n MLc_1,$$

where $c_1 = \|(\alpha, \beta)\| + L^{-1}\|A\alpha\|$. Now we use induction argument, where our induction hypothesis is

$$\|\xi^{(j-1)}(t) - \xi^{(j-2)}(t)\| \leqslant c_1(ML\Delta t_n)^{(j-1)}. \tag{62}$$

Then, using the Lipschitz continuity of f and our induction hypothesis (62) we have,

$$\|\xi^{(j)}(t) - \xi^{(j-1)}(t)\| = \int_{t_{n-1}}^{t} G(t - s)(f_i(\xi^{(j-1)}, \beta) - f_i(\xi^{(j-2)}, \beta))\, ds \leqslant \Delta t_n ML\|\xi^{(j-1)}(t) - \xi^{(j-2)}(t)\| \leqslant c_1(ML\Delta t_n)^j$$

Now, if $ML\Delta t_n < 1$, then for $l > k > N$,

$$\|\xi^{(l)}(t) - \xi^{(k)}(t)\| \leqslant \sum_{j=k+1}^{l} \|\xi^{(j)}(t) - \xi^{(j-1)}(t)\| \leqslant \frac{c_1(ML\Delta t_n)^N}{1 - ML\Delta t_n}. \tag{63}$$

Thus, $\|\xi^{(l)}(t) - \xi^{(k)}(t)\| \to 0$ as $N \to \infty$. Hence, $\xi^{(l)}(t)$ is a Cauchy sequence in the Banach space X, and hence converges to an element in X. We pass to the limit in (61), so that this limit satisfies (60). □

Now we use this lemma to prove the convergence of Algorithm 2.

Theorem A.1. *With Assumptions A.1 and A.2, there exists $t_n > t_{n-1}$ such that the sequence of functions $\left\{u_1^{(m)}\right\}$ and $\left\{u_2^{(m)}\right\}$ as defined in Algorithm 2 converges to the exact solution of (2) on the time interval $I_n = [t_{n-1}, t_n]$.*

Proof. The existence of the sequences $\left\{u_1^{(m)}\right\}$ and $\left\{u_2^{(m)}\right\}$ are established by repeated application of Lemma A.1. For $m = 1$, we set $\alpha = u_1(t_{n-1})$ and $\beta = u_2^{(0)}(t_{n-1})$. Then, by Lemma A.1, there exists a solution $(u_1^{(1)}, u_2^{(0)})$ to the integral equation governing $u_1^{(1)}$. We obtain a similar result for $u_2^{(1)}$ by setting $\alpha = u_2(t_{n-1})$ and $\beta = u_1^{(1)}(t_{n-1})$. Hence, repeated application of this lemma shows the existence of the sequences $\left\{u_1^{(m)}\right\}$ and $\left\{u_2^{(m)}\right\}$. Moreover, from the proof of Lemma A.1 we have,

$$\|u_2^{(1)}(t) - u_2^{(0)}(t)\| = \|u_2^{(1)}(t) - u_2(t_{n-1})\| \leqslant c_1 ML\Delta t_n.$$

Thus,

$$\begin{aligned}\|u_1^{(2)}(t) - u_1^{(1)}(t)\| &\leqslant \int_{t_{n-1}}^{t} \|G(t-s)(f_1(u_1^{(2)}, u_2^{(1)}) - f_1(u_1^{(1)}, u_2^{(1)}))\| + \|G(t-s)(f_1(u_1^{(1)}, u_2^{(1)}) - f_1(u_1^{(1)}, u_2^{(0)}))\|\, ds \\ &\leqslant ML \int_{t_{n-1}}^{t} \|u_1^{(2)} - u_1^{(1)}\|\, ds + c_1 ML(t - t_{n-1})^2.\end{aligned}$$

Setting $\tau_n = ML\Delta t_n \exp(ML\Delta t_n)$, we apply Gronwall's inequality,

$$\|u_1^{(2)}(t) - u_1^{(1)}(t)\| \leqslant c_1 ML(t - t_{n-1})^2 \exp(ML\Delta t_n) \leqslant \frac{c_1 \tau_n^2}{ML \exp(ML\Delta t_n)}. \tag{64}$$

Similarly,

$$\|u_2^{(2)}(t) - u_2^{(1)}(t)\| \leqslant \frac{c_1 \tau_n^2}{ML \exp(ML\Delta t_n)}. \tag{65}$$

Now we use induction, where our induction hypothesis is,

$$\|u_1^{(m-1)}(t) - u_1^{(m-2)}(t)\| \leqslant c_n \tau_n^{m-1} \tag{66}$$

and

$$\|u_2^{(m-1)}(t) - u_2^{(m-2)}(t)\| \leqslant c_n \tau_n^{m-1}, \tag{67}$$

where $c_n = \frac{c_1}{ML \exp(ML\Delta t_n)}$. We have, by our induction hypothesis and Gronwall's inequality,

$$\begin{aligned}\|u_1^{(m)}(t) - u_1^{(m-1)}(t)\| &\leqslant \int_{t_{n-1}}^{t} \|G(t-s)(f_1(u_1^{(m)}, u_2^{(m-1)}) - f_1(u_1^{(m-1)}, u_2^{(m-1)}))\| \quad + \|G(t-s)(f_1(u_1^{(m-1)}, u_2^{(m-1)}) \\ &\quad - f_1(u_1^{(m-1)}, u_2^{(m-2)}))\|\, ds \\ &\leqslant ML \int_{t_{n-1}}^{t} \|u_1^{(m)} - u_1^{(m-1)}\|\, ds + ML\Delta t_n \|u_2^{(m-1)}(t) - u_2^{(m-2)}(t)\| \\ &\leqslant ML \int_{t_{n-1}}^{t} \|u_1^{(m)} - u_1^{(m-1)}\|\, ds + c_n ML\Delta t_n \tau_n^{m-1} \leqslant ML\Delta t_n c_n \tau_n^{m-1} \exp(ML\Delta t_n) = c_n \tau_n^m.\end{aligned} \tag{68}$$

For $\tau_n < 1$, and $l > k > N$,

$$\|u_1^{(l)}(t) - u_1^{(k)}(t)\| \leqslant \sum_{m=k+1}^{l} \|u_1^{(m)}(t) - u_1^{(m-1)}(t)\| \tag{69}$$

$$\leqslant \sum_{m=N}^{\infty} \|u_1^{(m)}(t) - u_1^{(m-1)}(t)\| \tag{70}$$

$$\leqslant \frac{c_n \tau_n^N}{1 - \tau_n} \tag{71}$$

By enforcing $\tau_n < 1$, we get that $u_1^{(m)}$ is a Cauchy sequence that converges to an element in X. This is also true for $u_2^{(m)}$. We pass to the limit in (59), so that it converges to the solution of the implicit equation. □

References

[1] G. Adomian, The diffusion-Brusselator equation, Comput. Math. Appl. 29 (1995) 1–3.
[2] M. Ainsworth, T. Oden, *A Posteriori* Error Estimation in Finite Element Analysis, John Wiley-Teubner, 2000.

[3] W. Bangerth, R. Rannacher, Adaptive Finite Element Methods for Differential Equations, Birkhauser Verlag, 2003.
[4] T.J. Barth, A-Posteriori Error Estimation and Mesh Adaptivity for Finite Volume and Finite Element Methods, Lecture Notes in Computational Science and Engineering, vol. 41, Springer, New York, 2004.
[5] R. Becker, R. Rannacher, An optimal control approach to a posteriori error estimation in finite element methods, Acta Numer. (2001) 1–102.
[6] V. Carey, D. Estep, S. Tavener, *A posteriori* analysis and adaptive error control for operator decomposition solution of elliptic systems. I: Triangular systems, SIAM J. Numer. Anal. 47 (2009) 740–761.
[7] V. Carey, D. Estep, S. Tavener, *A posteriori* analysis and adaptive error control for multiscale operator decomposition solution of elliptic systems II: Fully coupled systems, Int. J. Numer. Methods Engrg., 2012 (in revision).
[8] K. Eriksson, D. Estep, P. Hansbo, C. Johnson, Introduction to adaptive methods for differential equations, Acta Numerica, vol. 1995, Cambridge Univ. Press, Cambridge, 1995, pp. 105–158.
[9] K. Eriksson, D. Estep, P. Hansbo, C. Johnson, Computational Differential Equations, Cambridge University Press, Cambridge, 1996.
[10] D. Estep, A posteriori error bounds and global error control for approximation of ordinary differential equations, SIAM J. Numer. Anal. 32 (1995) 1–48.
[11] D. Estep, Error estimation for multiscale operator decomposition for multiphysics problems, in: J. Fish (Ed.), Bridging the Scales in Science and Engineering, Oxford University Press, 2008.
[12] D. Estep, V. Carey, V. Ginting, S. Tavener, T. Wildey, A posteriori error analysis of multiscale operator decomposition methods for multiphysics models, J. Phys. Conf. Ser. 125 (2008) 012075.
[13] D. Estep, V. Ginting, D. Ropp, J. Shadid, S. Tavener, An a posteriori-a priori analysis of multiscale operator splitting, SIAM J. Numer. Anal. 46 (2008) 1116–1146.
[14] D. Estep, V. Ginting, S. Tavener, A posteriori analysis of a multirate numerical method for ordinary differential equations, CMAME 223–224 (2012) 10–227.
[15] D. Estep, M.G. Larson, R.D. Williams, Estimating the error of numerical solutions of systems of reaction–diffusion equations, Mem. Am. Math. Soc. 146 (2000). viii+109.
[16] D. Estep, S. Tavener, T. Wildey, *A posteriori* analysis and improved accuracy for an operator decomposition solution of a conjugate heat transfer problem, SIAM J. Numer. Anal. 46 (2008) 2068–2089.
[17] D. Estep, S. Tavener, T. Wildey, *A posteriori* error analysis for a transient conjugate heat transfer problems, Finite Elem. Anal. Des. 45 (2009) 263–271.
[18] D. Estep, S. Tavener, T. Wildey, *A posteriori* error estimation and adaptive mesh refinement for a multi-discretization operator decomposition approach to fluid–solid heat transfer, J. Comput. Phys. 229 (2010) 4143–4158.
[19] M. Giles, E. Süli, Adjoint methods for PDEs: *a posteriori* error analysis and postprocessing by duality, Acta Numer. (2002) 145–236.
[20] O.A. Ladyženskaja, V.A. Solonnikov, N.N. Ural'ceva, Linear and quasilinear equations of parabolic type, in: Translations of Mathematical Monographs, vol. 23, American Mathematical Society, Providence, RI, 1968 (Translated from the Russian by S. Smith).
[21] M.G. Larson, F. Bengzon, Adaptive finite element approximation of multiphysics problems, Commun. Numer. Methods Engrg. 24 (2008) 505–521.
[22] M.G. Larson, A. Maylqvist, Goal oriented adaptivity for coupled flow and transport problems with applications in oil reservoir simulations, Comput. Methods Appl. Mech. Engrg. 196 (2007) 3546–3561.
[23] A. Logg, Multi-adaptive time integration, Appl. Numer. Math. 48 (2004) 339–354.
[24] A. Pazy, Semigroups of linear operators and applications to partial differential equations, Applied Mathematical Sciences, Springer-Verlag, 1983.
[25] I. Prigogine, R. Lefever, Symmetry breaking instabilities in dissipative systems, J. Chem. Phys. 48 (1968) 1695–1700.