A-POSTERIORI ERROR ESTIMATION AND ADAPTIVITY FOR NONLINEAR PARABOLIC EQUATIONS USING IMEX-GALERKIN DISCRETIZATION OF PRIMAL AND DUAL EQUATIONS [*]

X. WU[†], K. G. VAN DER ZEE[‡], G. SIMSEK[†], AND E. H. VAN BRUMMELEN[†]

Abstract. While many methods exist to discretize nonlinear time-dependent partial differential equations (PDEs), the rigorous estimation and adaptive control of their discretization errors remains challenging. In this paper, we present a methodology for duality-based a posteriori error estimation for nonlinear parabolic PDEs, where the full discretization of the PDE relies on the use of an implicit-explicit (IMEX) time-stepping scheme and the finite element method in space. The main result in our work is a decomposition of the error estimate that allows to separate the effects of spatial and temporal discretization error, and which can be used to drive adaptive mesh refinement and adaptive time-step selection. The decomposition hinges on a specially-tailored IMEX discretization of the dual problem. The performance of the error estimates and the proposed adaptive algorithm is demonstrated on two canonical applications: the elementary heat equation and the nonlinear Allen–Cahn phase-field model.

Key words. A posteriori error estimate, Duality-based error estimate, IMEX scheme, Implicit-explicit schemes, Space-time error, Adaptivity, Parabolic PDE

AMS subject classifications. 65M15, 65M20, 65M50

1. Introduction. Nonlinear parabolic PDEs are ubiquitous in science, however, their efficient numerical solution remains challenging. Implicit-explicit (IMEX) methods have been widely used for the time integration of complex time-dependent PDEs with terms of different type [1, 9]. Recently, a number of IMEX time-stepping schemes, paired with spatial Galerkin finite-element discretizations, have been proposed for phase-field models [40, 20, 42, 35, 39], which are currently a much-studied class of nonlinear parabolic problems [24, 29, 22, 33, 21]. When the PDE solution displays alternating fast and slow variations, the numerical discretization can, obviously, benefit significantly from adaptivity in both space and time.

This paper is devoted to the development of a posteriori error estimates and corresponding adaptive algorithms for these popular discretizations. In particular, we consider dual-based error estimates that assess the discretization error with respect to user specified quantities of interest describing the goal of the analyses. The quantities of interest might, for instance, be physical quantities or some appropriate norms of the error of the solution (e.g. energy norm, L^2 norm). To efficiently drive adaptive mesh refinement and adaptive time-step selection, the error estimates need to address the temporal and the spatial discretization errors separately.

There have been several studies on goal-oriented adaptive techniques for parabolic equations during the last decades, but mostly in the context of space-time (discontinuous) Galerkin finite element discretization, see for instance Eriksson and Johnson [15, 16, 17, 18], Schmich and Vexler [32], Carey et al. [12], Bermejo and Carpio [7], Braack et al. [10], Besier and Rannacher [8], and Asner et al. [2]. Alternative a posteriori techniques for parabolic problems use energy arguments based on elliptic

[†]Multiscale Engineering Fluid Dynamics, Eindhoven University of Technology, P. O. Box 513, 5600 MB Eindhoven, Netherlands (x.wu@tue.nl, g.simsek@tue.nl, e.h.v.brummelen@tue.nl).

[‡] School of Mathematical Sciences, University of Nottingham, University Park, NG7 2RD Nottingham, United Kingdom (kg.vanderzee@nottingham.ac.uk).

reconstruction, see e.g. Makridakis and Nochetto [28], or sharp stability estimates for the dual problem; see Lakkis et al. [27] for an overview of both these techniques. Applications of these techniques to nonlinear parabolic problems including phase-field models and blow-up phenomena have been studied by, e.g., Kessler et al. [23], Bartels and Muller [5] and Cangiani et al. [11].

The use of IMEX time-stepping schemes for parabolic equations has been the focus of recent developments: Chaudhry et al. [13, 14] proposed a posteriori error estimates for various IMEX schemes, based on an equivalence relation between IMEX schemes and *time-Galerkin* finite element methods. They rewrite the time-Galerkin method using special numerical quadrature rules and carry out a standard duality-based analysis for the resultant approximations. The splitting of the temporal and the spatial error contributions in these error estimates are commonly achieved by inserting and subtracting suitable projections of the dual solution.

The objective of this paper is to present an alternative approach to duality-based a posteriori error estimates for fully discretized semi-linear parabolic PDEs using conforming finite elements in space and first-order IMEX schemes in time. Contrary to Chaudhry et al, in our approach we *directly* obtain a posteriori error estimates without resorting to an interpretation of IMEX as a Galerkin-in-time method. This paper is a follow-up to our recent paper [34], where we only considered errors due to spatial discretization. The focus of this work is on the *total* discretization error which contains both the spatial and temporal parts.

The starting point of our analysis is the exact duality-based error representation, which is a duality pairing of the global space-time residual with the solution of the mean-value-linearized (backward-in-time) dual problem. This error representation can be decomposed into various distinct residuals weighted by the same dual solution. A fundamental framework for successfully decomposing the residuals for (non)linear parabolic PDEs, discretized by a classical A-stable θ-scheme in time, has been developed by Verfürth [38] in the context of *energy*-based a posteriori error analysis.

By extending Verfürth's framework to IMEX schemes and a *duality*-based error analysis, we will decompose the error representation into three contributions which can be associated to the temporal and spatial discretization error, and additionally data oscillation. This novel decomposition hinges on a special nonstandard IMEX discretization of the dual problem. We then propose a general space-time adaptive algorithm for an efficient distribution of the discretization parameters: a set of time steps and the refined mesh at each time step.

This work is structured as follows. In Section 2, we introduce the abstract setting for a general (non)linear parabolic PDE and its IMEX-Galerkin discretization. Section 3 is devoted to the methodology for a space-time decomposition of a duality-based a posteriori error estimate. After having established computable error estimates in Section 4, we propose the associated adaptive algorithm in Section 5. The application to the elementary heat equation and the nonlinear Allen–Cahn equation (an elementary phase-field model), together with numerical results are presented in Section 6, after which we present our conclusions.

2. Abstract setting. In this section, we start by introducing an abstract setting of nonlinear parabolic PDEs and the corresponding dual problem in weak formulations. Then we present the discretization of the primal problem using IMEX time-stepping schemes and conforming finite elements in space.

2.1. Weak formulation and error representation. As a model problem we consider a general semi-linear parabolic equation in a bounded space-time domain $\Omega \times (0, T]$ with $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, having natural boundary conditions. To provide a setting for the weak formulation, we denote by $\mathcal{V}$ a suitable Hilbert space and by $\mathcal{V}^*$ its dual space, such that $\mathcal{V} \subset L^2(\Omega) \subset \mathcal{V}^*$ with continuous embeddings. We denote the inner product in $L^2(\Omega)$ by $(\cdot, \cdot)$, and duality pairings between $\mathcal{V}^*$ and $\mathcal{V}$ by $\langle \cdot, \cdot \rangle$. By defining $\mathcal{W} := \{ v \in L^2(0, T; \mathcal{V}), \partial_t v \in L^2(0, T; \mathcal{V}^*) \}$ as a suitable space for u, the weak form reads: find $u \in \mathcal{W}_{u^0} := \{ v \in \mathcal{W} : v(0) = u^0 \}$ such that $\forall v \in L^2(0, T; \mathcal{V})$

$$
(1) \qquad \int_0^T \left(\langle \partial_t u, v \rangle + \mathcal{B}(u, v) + \mathcal{N}(u; v) \right) \mathrm{d}t = \int_0^T \langle f, v \rangle \, \mathrm{d}t
$$

where $f \in L^2(0, T; \mathcal{V}^*)$, $u^0 \in L^2(\Omega)$, the semi-linear form $\mathcal{N}(\cdot; \cdot)$ of a sufficiently smooth nonlinear operator represents the nonlinear components which is linear with respect to arguments on the right of the semicolon, and $\mathcal{B}(\cdot, \cdot)$ is the bilinear form of a elliptic self-adjoint operator. A prime example of the abstract setting is the Allen–Cahn equation $\partial_t u - \Delta u + \frac{1}{\varepsilon^2} \psi'(u) = 0$, which will be discussed later in Subsection 6.1.

Given the solution u, we consider the quantity

$$
(2) \qquad \mathcal{Q}(u) := (\bar{q}, u(T)) + \int_0^T (q, u) \, \mathrm{d}t,
$$

with $\bar{q} \in L^2(\Omega)$ and $q \in L^2(0, T; L^2(\Omega))$ so that $\mathcal{Q} : \mathcal{W} \to \mathbb{R}$ is a continuous linear functional.[1] One example of $\mathcal{Q}(u)$ would be the value of the solution at the final time $t = T$ at a critical area of the domain centered at $\mathbf{x}_0$, $\mathcal{Q}(u) = \int_\Omega \rho_\epsilon(\mathbf{x}_0 - \mathbf{x}) u(\mathbf{x}, T) \, \mathrm{d}\mathbf{x}$, where $\rho_\epsilon \in C^\infty$ is a kernel function with radius and center of ϵ and $\mathbf{x}_0$. Alternatively, one might wish to estimate the error in the L^2 norm at the final time T. To achieve this, we set $\bar{q} = u(T) - \hat{u}(T)$ and $q = 0$ where $\hat{u}$ is an approximation of the solution u. Then we have $\mathcal{Q}(u) - \mathcal{Q}(\hat{u}) = \| u(T) - \hat{u}(T) \|_{L^2(\Omega)}^2$.

For any $u, \hat{u} \in \mathcal{V}$, we denote by $\mathcal{N}^s(u, \hat{u}; \cdot, \cdot)$ the mean-value linearization of $\mathcal{N}(\cdot; \cdot)$ performed at a value in between u and $\hat{u}$, namely,

$$
(3) \qquad \mathcal{N}^s(u, \hat{u}; w, v) = \int_0^1 \mathcal{N}' \big(su + (1 - s)\hat{u} \big)(w, v) \mathrm{d}s, \qquad \forall w, v \in \mathcal{V}
$$

where $\mathcal{N}'$ is the Gâteaux derivative of $\mathcal{N}$, i.e.

$$
\mathcal{N}'(\hat{w})(w, v) = \lim_{s \to 0} \frac{\mathcal{N}(\hat{w} + s\,w; v) - \mathcal{N}(\hat{w}; v)}{s} \mathrm{d}s, \qquad \forall w, v \in \mathcal{V}.
$$

Note that if we set $w = u - \hat{u}$, the chain rule gives

$$
(4) \qquad \mathcal{N}^s(u, \hat{u}; u - \hat{u}, v) = \mathcal{N}(u; v) - \mathcal{N}(\hat{u}; v).
$$

The mean-value-linearized (backward-in-time) dual problem takes the form: find $z \in \mathcal{W}^{\bar{q}} := \{ v \in \mathcal{W} : v(T) = \bar{q} \}$ such that $\forall w \in L^2(0, T; \mathcal{V})$

$$
(5) \qquad \int_0^T \left(\langle -\partial_t z, w \rangle + \mathcal{B}(z, w) + \mathcal{N}^s(u, \hat{u}; w, z) \right) \mathrm{d}t = \int_0^T (q, w) \, \mathrm{d}t
$$

---

[1]One can more generally consider $q \in L^2(\tau_1, T; \mathcal{V}^*) \cup L^2(0, T; L^2(\Omega))$ where $\tau_1 > 0$ is the size of the first time step. For technical reasons later on (i.e., $\langle q_0, u_0 \rangle$ must be well-defined, with q_0 defined in (21)), we can not take $\tau_1 = 0$.

Let $\hat{u} \in \mathcal{W}$ denote any approximation of the solution u in (1). We define the residual of the primal PDE, $\mathcal{R}_{\mathrm{PDE}}$, and the residual of the initial condition, $\mathcal{R}_0$, as

(6) $\qquad \mathcal{R}_{\mathrm{PDE}}(\hat{u}(t); v) := \langle f(t), v \rangle - \langle \partial_t \hat{u}(t), v \rangle - \mathcal{B}(\hat{u}(t), v) - \mathcal{N}(\hat{u}(t); v)$

(7) $\qquad \mathcal{R}_0(\hat{u}(0); w) := \left(u^0 - \hat{u}(0), w\right)$

for all $v \in \mathcal{V}$ and $w \in L^2(\Omega)$. Following the general framework of goal-oriented error analysis (see, e.g. [6, 30]), we obtain an exact error representation assessing the error in $\mathcal{Q}$, which can generally be represented by a global space-time residual weighted by the solution of the dual problem (5).

THEOREM 2.1 (Global space-time error representation). *Given any approximation $\hat{u} \in \mathcal{W}$ of the solution u of the primal problem (1), we have the following a posteriori error representation:*

(8) $\qquad \mathcal{Q}(u) - \mathcal{Q}(\hat{u}) = \mathcal{R}_0\left(\hat{u}(0); z(0)\right) + \int_0^T \mathcal{R}_{\mathrm{PDE}}\left(\hat{u}(t); z(t)\right) \, \mathrm{d}t$

where $z \in \mathcal{W}^{\bar{q}}$ is the solution of the dual problem (5).

Proof. The proof is standard, see, e.g. [37] or [34, Theorem 2.A]. $\qquad\square$

Note that errors in norm are also included in Theorem 2.1 by suitably changing $\bar{q}$ and q, e.g., as in the example above.

2.2. IMEX - FEM Discretization. We next describe a full discretization of problem (1) by partitioning $[0, T]$ as $0 = t_0 < t_1 < t_2 < \cdots < t_k < \cdots < t_N = T$ into N subintervals $\mathcal{I}_{k+1} = [t_k, t_{k+1}]$ of length $\tau_{k+1} = t_{k+1} - t_k$, $k = 0, 1, \cdots, N - 1$. Because of the nonlinearity in the system (1), one has to be careful in choosing a time discretization to avoid prohibitive stability restrictions and high computational complexity. In this paper, we focus on first-order IMEX time-stepping schemes, which employ a splitting of the nonlinear term $\mathcal{N}$ according to

$$\mathcal{N}(u; v) = \mathcal{N}_c(u; v) - \mathcal{N}_e(u; v),$$

The notation $\mathcal{N}_c$ and $\mathcal{N}_e$ comes from the phase-field modeling community, and refers to the contractive and expansive part, respectively, which can also refer to the stiff and non-stiff term. The fundamental idea is to treat the contractive part *implicitly* and the expansive part *explicitly*. Such a time scheme for problem (1) is defined recursively by: find $u_{k+1} \in \mathcal{V}$ such that $\forall v \in \mathcal{V}$

(9) $\qquad \left(\dfrac{u_{k+1} - u_k}{\tau_{k+1}}, v\right) + \mathcal{B}(u_{k+1}, v) + \mathcal{N}_c(u_{k+1}; v) - \mathcal{N}_e(u_k; v) = (f_{k+1}, v)$

for $k = 0, 1, \cdots, N - 1$, where the initial condition is

(10) $\qquad (u_0, v) = (u^0, v) \qquad \forall v \in L^2(\Omega).$

Here, $f_{k+1} = f(\cdot, t_{k+1})$, which is well-defined upon assuming that the function f is sufficiently regular, e.g., $f \in C^0((0, T]; L^2(\Omega))$. We remark that instead of the time approximation f_{k+1}, a time-averaged approximation $\bar{f} = \frac{1}{\tau_{k+1}} \int_{t_k}^{t_{k+1}} f(\cdot, t) \, \mathrm{d}t$ can be used, provided that $f \in L^2(0, T; L^2(\Omega))$. We also implicitly assume in (9) that $\mathcal{N}_e(u_0; v)$ is bounded for $u_0 \in L^2(\Omega)$. If $\mathcal{N}_e(u_0; v)$ is not well-defined, one can remove

4

this term from (9) for the first time step. For simplicity, we continue our analysis assuming that $\mathcal{N}_e(\cdot;\cdot)$ is bounded on $L^2(\Omega) \times L^2(\Omega)$.

To fully discretize the primal problem (1), we consider a standard shape-regular mesh $\mathcal{K}_k$ of Ω and an associated conforming finite element space $\mathcal{S}_k^{h,p}$ defined by

$$\mathcal{S}_k^{h,p} := \{v \in \mathcal{V} : v(x)|_K \in \mathbb{P}^p(K), \forall K \in \mathcal{K}_k\}$$

for $k = 0, 1, \ldots, N$, where $\mathbb{P}^p(K)$ is the space of polynomials up to order p on element K and h denotes the mesh parameter. The fully discrete approximation is then formulated as: find $u_{k+1}^h \in \mathcal{S}_{k+1}^{h,p}$ such that $\forall v^h \in \mathcal{S}_{k+1}^{h,p}$

$$(11) \quad \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, v^h\right) + \mathcal{B}(u_{k+1}^h, v^h) + \mathcal{N}_c(u_{k+1}^h; v^h) - \mathcal{N}_e(u_k^h; v^h) = (f_{k+1}, v^h)$$

for $k = 0, 1, \cdots, N-1$, where the initial condition is

$$(12) \qquad\qquad (u_0^h, v^h) = (u^0, v^h) \qquad \forall v^h \in \mathcal{S}_0^{h,p}.$$

We assume that the solutions $u_\tau := \{u_k\}_{k=0}^N$ and $u_{\tau h} := \{u_k^h\}_{k=0}^N$ exist for the time-discrete primal problem (9) and the fully-discrete primal problem (11), respectively.

3. Space-time decomposed a posteriori error estimate. Space-time adaptivity is heavily dependent on an appropriate decomposition of error estimates, which will be derived in this section. Our approach to isolate error contributions from different sources is inspired by the work of Verfürth in [38, Chapter 6], which contains a general framework for deriving residual-based a posteriori error estimates for nonlinear parabolic problems with the θ-scheme. In the following Lemma, we adapt Verfürth's residual decomposition to our fully discrete primal problem (11).

LEMMA 3.1 (Residual decomposition). *Let $u_{\tau h} := \{u_k^h\}_{k=0}^N$ denote the solution of the fully discrete problem (11), and $Iu_{\tau h}$ denote the piecewise-linear time reconstruction of $u_{\tau h}$ on time intervals $[t_k, t_{k+1}]$, $k = 0, 1, \ldots, N-1$, i.e.,*

$$(13) \qquad Iu_{\tau h}(t) = \frac{t_{k+1} - t}{\tau_{k+1}} u_k^h + \frac{t - t_k}{\tau_{k+1}} u_{k+1}^h \qquad t \in [t_k, t_{k+1}].$$

Let the spatial residual r_h^k, the temporal residual r_τ^k and the data-oscillation contribution r_f^k be defined, for each $k = 0, 1, \cdots, N-1$, by

(14)
$$\left\langle r_h^{k+1}, v \right\rangle := (f_{k+1}, v) - \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, v\right) - \mathcal{B}(u_{k+1}^h, v) - \mathcal{N}_c(u_{k+1}^h; v) + \mathcal{N}_e(u_k^h; v)$$

(15)
$$\left\langle r_\tau^{k+1}(t), v \right\rangle := \mathcal{B}(u_{k+1}^h, v) + \mathcal{N}_c(u_{k+1}^h; v) - \mathcal{N}_e(u_k^h; v) - \mathcal{B}(Iu_{\tau h}(t), v) - \mathcal{N}(Iu_{\tau h}(t); v)$$

(16)
$$\left\langle r_f^{k+1}(t), v \right\rangle := \langle f(t) - f_{k+1}, v \rangle .$$

for all $v \in \mathcal{V}$ and $t \in (t_k, t_{k+1}]$. Then, for each $k = 0, 1, \cdots, N-1$, the following decomposition of the space-time residual (6) holds:

$$(17) \quad \mathcal{R}_{\mathrm{PDE}}(Iu_{\tau h}(t); v) = \langle f(t), v \rangle - \langle \partial_t Iu_{\tau h}(t), v \rangle - \mathcal{B}(Iu_{\tau h}(t), v) - \mathcal{N}(Iu_{\tau h}(t); v)$$

$$(18) \qquad\qquad = \left\langle r_h^{k+1}, v \right\rangle + \left\langle r_\tau^{k+1}(t), v \right\rangle + \left\langle r_f^{k+1}(t), v \right\rangle$$

5

where $t \in (t_k, t_{k+1}]$.

Proof. Since $\partial_t I u_{\tau h} = \frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}$ on $(t_k, t_{k+1}]$, the identities in (18) follow from a straightforward substitution in (17) using the definition (14), (15) and (16). □

Remark 3.2. We note that the spatial residuals (14) are independent of time, and due to Galerkin orthogonality, the spatial residuals will be equal to zero if $v \in \mathcal{S}_{k+1}^{h,p}$. Furthermore, upon convergence $u_k^h \to u_k$ as $h \to 0^+$, for all k, we also have $r_h^{k+1} \to 0$ (see (9)). Similarly, assuming sufficient smoothness in time, then $u_k^h, u_{k+1}^h \to I u_{\tau h}(t)$ for $t \in [t_k, t_{k+1}]$ as $\tau_{k+1} \to 0^+$, which implies $r_\tau^{k+1}(t) \to 0$ as $\tau_{k+1} \to 0^+$. This is the motivation for calling r_τ^{k+1} and r_h^{k+1} the *temporal residual* and the *spatial residual*, respectively.

3.1. Time-discrete error representation. The first step toward a decomposition of duality-based error estimates is to introduce a time-discrete error representation identifying only the spatial discretization error. To this end, we introduce a novel and specially-tailored IMEX time-discrete dual problem. This time-discrete problem is driven by the following discrete representation of $\mathcal{Q}$.

Let us rewrite the piecewise-linear time reconstruction $I w_\tau \in \mathcal{W}$ of any sequence $w_\tau := \{w_k\}_{k=0}^N$, $w_k \in \mathcal{V}$, as

$$(19) \qquad I w_\tau(\mathbf{x}, t) = \sum_{k=0}^N w_k(\mathbf{x}) \mathbf{N}_k(t)$$

where

$$\mathbf{N}_k(t) := \begin{cases} \dfrac{t_{k+1} - t}{\tau_{k+1}} & \text{if } t \in \mathcal{I}_{k+1}, k \le N - 1 \\[2mm] \dfrac{t - t_{k-1}}{\tau_k} & \text{if } t \in \mathcal{I}_k, k \ge 1 \\[2mm] 0 & \text{otherwise} \end{cases}$$

We consider the following discrete representation of $\mathcal{Q} : \mathcal{W} \to \mathbb{R}$ when applied to $I w_\tau$.

LEMMA 3.3. *Let us define*

$$(20) \qquad q_k = \frac{1}{\tau_k} \int_0^T q \mathbf{N}_k(t) \mathrm{d}t \qquad \text{for } k = 1, 2, \ldots, N$$

and

$$(21) \qquad q_0 = \frac{1}{\tau_1} \int_0^T q \mathbf{N}_0(t) \mathrm{d}t.$$

Then, the following time-discrete representation of $\mathcal{Q} : \mathcal{W} \to \mathbb{R}$ holds

$$(22) \qquad \mathcal{Q}(I w_\tau) = \tau_1(q_0, w_0) + \sum_{k=1}^N \tau_k(q_k, w_k) + (\bar{q}, w_N).$$

Proof. For $I w_\tau$ defined in (19), we observe that, according to (2),

$$\mathcal{Q}(I w_\tau) = (\bar{q}, w_N) + \int_0^T \left(q, \sum_{k=0}^N w_k \mathbf{N}_k(t) \right) \mathrm{d}t = (\bar{q}, w_N) + \sum_{k=0}^N \left(\int_0^T q \, \mathbf{N}_k(t) \, \mathrm{d}t, \, w_k \right).$$

By virtue of

$$\sum_{k=0}^{N}\left(\int_0^T q\,\mathbf{N}_k(t)\,\mathrm{d}t,\ w_k\right) = \left(\int_0^T q\mathbf{N}_0(t),w_0\right) + \sum_{k=1}^{N}\left(\int_0^T q\,\mathbf{N}_k(t)\,\mathrm{d}t,\ w_k\right),$$

we obtain (22) by substituting the definition (20) and (21). $\qquad\square$

We now state the novel IMEX time-stepping scheme to discretize the dual problem backwards in time: Find $z_k \in \mathcal{V}$, $k = 0, 1, \ldots, N$, such that

$$(23)\qquad -\left(\frac{z_1 - z_0}{\tau_1}, w\right) - \mathcal{N}_e^s(u_0, u_0^h; w, z_1) = (q_0, w) \qquad \forall w \in \mathcal{V}$$

and for $k = 1, 2, \ldots, N - 1$:

$$(24)\qquad -\left(\frac{z_{k+1} - z_k}{\tau_k}, w\right) + \mathcal{B}(z_k, w) + \mathcal{N}_c^s(u_k, u_k^h; w, z_k)$$
$$-\frac{\tau_{k+1}}{\tau_k}\mathcal{N}_e^s(u_k, u_k^h; w, z_{k+1}) = (q_k, w) \qquad \forall w \in \mathcal{V}$$

where the terminal condition is

$$(25)\quad (z_N, w) + \tau_N \mathcal{B}(z_N, w) + \tau_N \mathcal{N}_c^s(u_N, u_N^h; w, z_N) = \tau_N(q_N, w) + (\bar{q}, w) \qquad \forall w \in \mathcal{V}.$$

The time-discrete dual (23)-(25) has been defined so as to provide an exact error representation for $\mathcal{Q}(Iu_{\tau h})$ with respect to $\mathcal{Q}(Iu_\tau)$.

THEOREM 3.4 (Time-discrete error representation). *Let $u_\tau = \{u_k\}_{k=0}^N$ denote the solution of the time discrete system (9), and $u_{\tau h} = \{u_k^h\}_{k=0}^N$ denote the solution of the fully discrete system (11). Let $z_\tau = \{z_k\}_{k=0}^N$ denote the time discrete approximation of the dual problem obtained from (23)-(25). Then the following error representation holds:*

$$(26)\qquad \mathcal{Q}(Iu_\tau) - \mathcal{Q}(Iu_{\tau h}) = \left(u^0 - u_0^h, z_0 - v_0^h\right) + \sum_{k=1}^N \tau_k \left\langle r_h^k, z_k - v_k^h \right\rangle,$$

for any $v_k^h \in \mathcal{S}_k^{h,p}$, $k = 0, 1, \ldots, N$.

Proof. From (22), it follows that $\mathcal{Q}(Iu_\tau) - \mathcal{Q}(Iu_{\tau h})$ can be formulated as

$$(27)\qquad \mathcal{Q}(Iu_\tau) - \mathcal{Q}(Iu_{\tau h}) = \tau_1(q_0, u_0 - u_0^h) + (\bar{q}, u_N - u_N^h) + \sum_{k=1}^N \tau_k(q_k, u_k - u_k^h)$$

Substituting the time-discrete dual problem (23)–(25) into (27), we get

$$
\mathcal{Q}(Iu_\tau) - \mathcal{Q}(Iu_{\tau h})
$$

$$
= \tau_1 \left\{ -\left(\frac{z_1 - z_0}{\tau_1}, u_0 - u_0^h \right) - \mathcal{N}_e^s(u_0, u_0^h; u_0 - u_0^h, z_1) \right\}
$$

$$
+ \tau_N \left\{ \left(\frac{z_N}{\tau_N}, u_N - u_N^h \right) + \mathcal{B}(z_N, u_N - u_N^h) + \mathcal{N}_c^s(u_N, u_N^h; u_N - u_N^h, z_N) \right\}
$$

$$
+ \sum_{k=1}^{N-1} \tau_k \left\{ -\left(\frac{z_{k+1} - z_k}{\tau_k}, u_k - u_k^h \right) + \mathcal{B}(z_k, u_k - u_k^h) \right.
$$

$$
\left. + \mathcal{N}_c^s(u_k, u_k^h; u_k - u_k^h, z_k) - \frac{\tau_{k+1}}{\tau_k} \mathcal{N}_e^s(u_k, u_k^h; u_k - u_k^h, z_{k+1}) \right\}
$$

After applying summation by parts on $(z_{k+1} - z_k, u_k - u_k^h)$, i.e.,

$$
\sum_{k=1}^{N-1} \left(u_k - u_k^h, z_{k+1} - z_k \right) =
$$

$$
\left(u_N - u_N^h, z_N \right) - \left(u_1 - u_1^h, z_1 \right) - \sum_{k=1}^{N-1} \left(z_{k+1}, \left(u_{k+1} - u_{k+1}^h \right) - \left(u_k - u_k^h \right) \right)
$$

it follows that

$$
\mathcal{Q}(Iu_\tau) - \mathcal{Q}(Iu_{\tau h})
$$

$$
= \left(u_1 - u_1^h, z_1 \right) + \tau_1 \left\{ -\left(\frac{z_1 - z_0}{\tau_1}, u_0 - u_0^h \right) - \mathcal{N}_e^s(u_0, u_0^h; u_0 - u_0^h, z_1) \right\}
$$

$$
+ \tau_N \left\{ \mathcal{B}(z_N, u_N - u_N^h) + \mathcal{N}_c^s(u_N, u_N^h; u_N - u_N^h, z_N) \right\}
$$

$$
+ \sum_{k=1}^{N-1} \tau_k \left\{ \left(\frac{u_{k+1} - u_k}{\tau_k}, z_{k+1} \right) - \left(\frac{u_{k+1}^h - u_k^h}{\tau_k}, z_{k+1} \right) + \mathcal{B}(z_k, u_k - u_k^h) \right.
$$

$$
\left. + \mathcal{N}_c^s(u_k, u_k^h; u_k - u_k^h, z_k) - \frac{\tau_{k+1}}{\tau_k} \mathcal{N}_e^s(u_k, u_k^h; u_k - u_k^h, z_{k+1}) \right\}
$$

Then, by shifting the indices of the arguments of $\mathcal{B}$ and $\mathcal{N}_c^s$:

$$
\tau_N \mathcal{B}(z_N, u_N - u_N^h) + \sum_{k=1}^{N-1} \tau_k \mathcal{B}(z_k, u_k - u_k^h) = \sum_{k=0}^{N-1} \tau_{k+1} \mathcal{B}(z_{k+1}, u_{k+1} - u_{k+1}^h)
$$

$$
\tau_N \mathcal{N}_c^s(u_N, u_N^h; u_N - u_N^h, z_N) + \sum_{k=1}^{N-1} \tau_k \mathcal{N}_c^s(u_k, u_k^h; u_k - u_k^h, z_k)
$$

$$
= \sum_{k=0}^{N-1} \tau_{k+1} \mathcal{N}_c^s(u_{k+1}, u_{k+1}^h; u_{k+1} - u_{k+1}^h, z_{k+1})
$$

8

and employing the mean-value linearization property (4) on $\mathcal{N}_c^s$ and $\mathcal{N}_e^s$, we arrive at

$$
\mathcal{Q}(Iu_\tau) - \mathcal{Q}(Iu_{\tau h}) = \left(u_0 - u_0^h, z_0\right)
$$
$$
+ \sum_{k=0}^{N-1} \tau_{k+1} \Bigg\{ \left(\frac{u_{k+1} - u_k}{\tau_{k+1}}, z_{k+1}\right) + \mathcal{B}(u_{k+1}, z_{k+1}) + \mathcal{N}_c(u_{k+1}; z_{k+1}) - \mathcal{N}_e(u_k; z_{k+1})
$$
$$
- \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, z_{k+1}\right) - \mathcal{B}(u_{k+1}^h, z_{k+1}) - \mathcal{N}_c(u_{k+1}^h; z_{k+1}) + \mathcal{N}_e(u_k^h; z_{k+1}) \Bigg\}
$$

After substituting the time-discrete primal problem (9) weighted by dual solution z_{k+1}, we finally obtain

$$
\mathcal{Q}(Iu_\tau) - \mathcal{Q}(Iu_{\tau h}) = \left(u^0 - u_0^h, z_0\right) + \sum_{k=0}^{N-1} \tau_{k+1} \Bigg\{ (f_{k+1}, z_{k+1}) - \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, z_{k+1}\right)
$$
$$
- \mathcal{B}(u_{k+1}^h, z_{k+1}) - \mathcal{N}_c(u_{k+1}^h; z_{k+1}) + \mathcal{N}_e(u_k^h; z_{k+1}) \Bigg\}.
$$

This is (26) by the definition in (14), and noting that $\forall v_{k+1}^h \in \mathcal{S}_{k+1}^{h,p}$ it holds that $\langle r_{k+1}^h, v_{k+1}^h \rangle = 0$ by (11) for $k = 0, 1, \ldots, N-1$ and $\left(u^0 - u_0^h, v_0^h\right) = 0$ $\forall v_0^h \in \mathcal{S}_0^{h,p}$ by (12). □

3.2. Spatial and temporal error representation. Building on Verfürth's residual decomposition (18) and the time-discrete error representation (26), we are now ready to state our main result: A suitable decomposition of the dual-weighted residual (8).

THEOREM 3.5 (Decomposed error representation). *Let the assumptions of Theorem 3.4 hold. Let u denote the solution of the primal problem (1) and z denote the solution of the dual problem (5). Then the following error representation holds:*

$$
(28) \qquad \mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h}) = \mathcal{R}_{\mathrm{s}}(u_{\tau h}; z_\tau - v_\tau^h) + \mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau) + \mathrm{Osc}
$$

for any $v_\tau^h := \{v_k^h\}_{k=0}^N$, $v_k^h \in \mathcal{S}_k^{h,p}$, where $\mathcal{R}_{\mathrm{s}}(u_{\tau h}; z_\tau - v_\tau^h)$ is the spatial error representation

$$
(29) \qquad \mathcal{R}_{\mathrm{s}}(u_{\tau h}; z_\tau - v_\tau^h) := \left(u^0 - u_0^h, z_0 - v_0^h\right) + \sum_{k=1}^N \tau_k \left\langle r_h^k, z_k - v_k^h \right\rangle,
$$

$\mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau)$ is the temporal error representation

$$
(30) \quad \mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau) := \left(u^0 - u_0^h, z(0) - z_0\right)
$$
$$
+ \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \left\{ \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); z(t) - z_{k+1}\right) + \left\langle r_\tau^{k+1}(t), z_{k+1} \right\rangle \right\} \mathrm{d}t
$$

and Osc denotes the data-oscillation contribution

$$
(31) \qquad \mathrm{Osc} := \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \left\langle r_f^{k+1}(t), z_{k+1} \right\rangle \mathrm{d}t.
$$

Remark 3.6. By virtue of Galerkin orthogonality, $\mathcal{R}_{\mathrm{s}}(u_{\tau h}; z_\tau)$ will vanish if $z_k \in \mathcal{S}_k^{h,p}$, $k = 0, 1, \ldots, N$. In addition, as $h \to 0^+$ it holds that $u_k^h \to u_k$ for all k and, accordingly, $\mathcal{R}_{\mathrm{s}}(u_{\tau h}; \cdot) \to 0$ (see (9) and (10)). Similarly, assuming sufficient smoothness in time, then $z_k, z_{k+1} \to z(t)$ for $t \in [t_k, t_{k+1}]$ as $\tau_{k+1} \to 0^+$, which implies $(u^0 - u_0^h, z(0) - z_0) \to 0$ and $\mathcal{R}_{\mathrm{PDE}}(Iu_{\tau h}(t); z(t) - z_{k+1}) \to 0$ as $\tau_{k+1} \to 0^+$. And since $r_\tau^{k+1}(t) \to 0$ as $\tau_{k+1} \to 0^+$ (see Remark 3.2), we conclude that $\mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau) \to 0$ as $\tau_{k+1} \to 0^+$. This is the motivation for calling $\mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau)$ and $\mathcal{R}_{\mathrm{s}}(u_{\tau h}; z_\tau)$ the *temporal error representation* and the *spatial error representation*, respectively.

Remark 3.7. If we choose $f_{k+1} = \frac{1}{\tau_{k+1}} \int_{t_k}^{t_{k+1}} f(t)\,\mathrm{d}t$, then the data-oscillation contribution (31) will vanish.

Proof. (of Theorem 3.5) The global space-time error representation is Theorem 2.1 with $\hat{u} = Iu_{\tau h}$:

$$(32) \qquad \mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h}) = \mathcal{R}_0\left(u_0^h; z(0)\right) + \int_0^T \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); z(t)\right)\,\mathrm{d}t.$$

The spatial error representation (29) satisfies the representation in Theorem 3.4.

The temporal error representation is obtained by subtracting the spatial error representation (29) and the data-oscillation contribution (31) from the space-time error representation (32), i.e.,

$$\mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau) = \mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h}) - \mathcal{R}_{\mathrm{s}}(u_{\tau h}; z_\tau) - \mathrm{Osc}$$

$$= \mathcal{R}_0\left(u_0^h; z(0)\right) - \left(z_0, u^0 - u_0^h\right) - \mathrm{Osc}$$

$$+ \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \left\{ \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); z(t)\right) - \left\langle r_h^{k+1}, z_{k+1} \right\rangle \right\}\,\mathrm{d}t.$$

Adding and subtracting $\mathcal{R}_{\mathrm{PDE}}(Iu_{\tau h}; z_{k+1})$ yields

$$\mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau) = \mathcal{R}_0\left(u_0^h; z(0)\right) - \left(z_0, u^0 - u_0^h\right) - \mathrm{Osc}$$

$$+ \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \left\{ \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); z(t) - z_{k+1}\right) + \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); z_{k+1}\right) - \left\langle r_h^{k+1}, z_{k+1} \right\rangle \right\}\,\mathrm{d}t.$$

Since $\partial_t Iu_{\tau h} = \left(u_{k+1}^h - u_k^h\right)/\tau_{k+1}$ on $(t_k, t_{k+1}]$ according to the definition of $Iu_{\tau h}$ (19), we employ the definition of the residuals in (6), (7) and (14), and obtain

$$\mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau)$$

$$= \left(u^0 - u_0^h, z(0) - z_0\right) - \mathrm{Osc} + \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \left\{ \langle f(t), z_{k+1} \rangle - (f_{k+1}, z_{k+1}) \right\}\,\mathrm{d}t$$

$$+ \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \left\{ \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); z(t) - z_{k+1}\right) - \mathcal{B}\left(Iu_{\tau h}(t) - u_{k+1}^h, z_{k+1}\right) \right.$$

$$\left. - \mathcal{N}\left(Iu_{\tau h}(t); z_{k+1}\right) + \mathcal{N}_c\left(u_{k+1}^h; z_{k+1}\right) - \mathcal{N}_e\left(u_k^h; z_{k+1}\right) \right\}\,\mathrm{d}t.$$

Finally, substituting the definitions in (15), (16) and (31) gives the result (30). □

A useful interpretation of the spatial and the temporal error representation can be obtained by writing the global space-time error as:

$$\mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h}) = \mathcal{Q}(u) - \mathcal{Q}(Iu_\tau) + \mathcal{Q}(Iu_\tau) - \mathcal{Q}(Iu_{\tau h}) \tag{33}$$

The following Corollary holds:

COROLLARY 3.8. *Under the assumptions of* Theorem 3.4 *and* Theorem 3.5*, we have*

$$\mathcal{Q}(u) - \mathcal{Q}(Iu_\tau) = \mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau) + \mathrm{Osc} \tag{34}$$

and

$$\mathcal{Q}(Iu_\tau) - \mathcal{Q}(Iu_{\tau h}) = \mathcal{R}_{\mathrm{s}}(u_{\tau h}; z_\tau - v_\tau^h) \tag{35}$$

for any $v_\tau^h := \{v_k^h\}_{k=0}^N$, $v_k^h \in \mathcal{S}_k^{h,p}$.

Proof. The identity in (35) is a direct consequence of (26) and (29). Equation (34) then follows from (28) and (29). □

4. Computable error estimate. There are two approximations commonly involved in evaluating the exact error representations (32), (29) and (30):

- one for approximating the exact primal solutions u and u_τ in the mean-value-linearized dual problem (5) and its time-discrete system (23)-(25),
- the other for approximating the exact dual solutions z and z_τ in the error representation formulas (32), (29) and (30).

The resulting error estimate can only be accurate if the approximations are sufficiently close to the true solutions. We refer to [6, Section 6] for techniques on controlling the linearization error due to the first approximation.

Here, to obtain computable and asymptotically effective error estimates, we consider a hierarchical two-level methodology developed in [34] where the estimate is directly evaluated with an enriched dual approximation that is computed with help of an additional primal approximation at an enriched discretization level for the mean-value-linearization. Since the focus of [34] is on the spatial discretization error, we now extend this methodology to our problem: We need two additional discretization levels: one which is spatially-enriched for evaluating the spatial error representation (29) and the other which is space-time enriched for evaluating (30) and (32).

We first introduce the following notations:

- $\mathcal{S}_k^{h,p}$: the original FE space with spatial mesh of size $h = h_k$ at time t_k for $k = 0, 1, \ldots, N$.

- $\mathcal{S}_k^{h/2,p}$: an *enriched* FE space with finer spatial mesh of size $\dfrac{h_k}{2}$ at time t_k for $k = 0, 1, \ldots, N$. $\mathcal{S}_k^{h/2,p}$ is obtained by global refinement of all the element in $\mathcal{S}_k^{h,p}$.

- $\mathcal{S}_{k+1/2}^{h/2,p}$: an enriched FE space with finer spatial mesh of size $\dfrac{h_k}{2}$ at the intermediate time level $t_{k+1/2} = \dfrac{t_{k+1} + t_k}{2}$ for $k = 0, 1, \ldots, N-1$. We set $\mathcal{S}_{k+1/2}^{h/2,p} = \mathcal{S}_{k+1}^{h/2,p}$.

- $u_{\tau,h} = \{u_k^h\}_{k=0}^N$: the solution of (11) and (12) using time-step sizes $\{\tau_k\}_{k=1}^N$ and FE spaces $\{\mathcal{S}_k^{h,p}\}_{k=0}^N$.

- $u_{\tau,h/2} = \{u_k^{h/2}\}_{k=0}^N$: the solution of (11) and (12) using time-step sizes $\{\tau_k\}_{k=1}^N$ and enriched FE spaces $\{\mathcal{S}_k^{h/2,p}\}_{k=0}^N$; $u_{\tau,h/2}$ represents an approximation of the time-discrete primal solution u_τ.

11

- $u_{\tau/2,h/2} = \{\tilde{u}_l^{h/2}\}_{l=0,1/2,\ldots,N}$: the solution of (11) and (12) using *half* time-step sizes $\left\{\dfrac{\tau_1}{2}, \dfrac{\tau_1}{2}, \dfrac{\tau_2}{2}, \dfrac{\tau_2}{2}, \ldots, \dfrac{\tau_N}{2}, \dfrac{\tau_N}{2}\right\}$ and enriched FE spaces $\{\mathcal{S}_l^{h/2,p}\}_{l=0,1/2,\ldots,N}$; $u_{\tau/2,h/2}$ represents an approximation of the exact primal solution u.

- $z_{\tau,h/2} = \{z_k^{h/2}\}_{k=0}^N$: the solution of the approximate dual problem obtained by replacing u_τ with $u_{\tau,h/2}$ in (23)-(25), using time-step sizes $\{\tau_k\}_{k=1}^N$ and enriched FE spaces $\{\mathcal{S}_k^{h/2,p}\}_{k=0}^N$; $z_{\tau,h/2}$ represents an approximation of the time-discrete dual z_τ.

- $z_{\tau/2,h/2} = \{\tilde{z}_l^{h/2}\}_{l=0,1/2,\ldots,N}$: the solution of the approximate dual problem obtained by replacing u_τ with $u_{\tau/2,h/2}$ in (23)-(25), using half time-step sizes $\left\{\dfrac{\tau_1}{2}, \dfrac{\tau_1}{2}, \dfrac{\tau_2}{2}, \dfrac{\tau_2}{2}, \ldots, \dfrac{\tau_N}{2}, \dfrac{\tau_N}{2}\right\}$ and enriched FE spaces $\{\mathcal{S}_l^{h/2,p}\}_{l=0,1/2,\ldots,N}$; $z_{\tau/2,h/2}$ represents the approximation of the exact dual solution z.

The strategy for computing the primal and dual solutions is illustrated in Figure 1. For evaluating the error representations (32), (29) and (30), we compute two enriched dual solutions $z_{\tau,h/2}$ and $z_{\tau/2,h/2}$ solved backwards in time to approximate z_τ and z, respectively. In order to make $z_{\tau,h/2}$ computable, an additional primal approximation $u_{\tau,h/2}$ is computed forwards in time using FE spaces $\{\mathcal{S}_k^{h/2,p}\}_{k=0}^N$ to approximate the mean-value-linearization of the dual problem (23)-(25). Similarly, another additional primal approximation $u_{\tau/2,h/2}$ is computed using spaces $\{\mathcal{S}_l^{h/2,p}\}_{l=0,1/2,\ldots,N}$ for obtaining $z_{\tau/2,h/2}$.
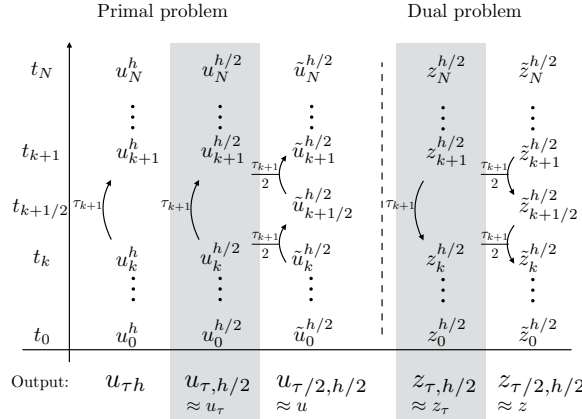

Fig. 1: Approximations of the primal and dual solutions. The primal approximations $u_{\tau h}$, $u_{\tau,h/2}$ and $u_{\tau/2,h/2}$ are computed forwards in time and the dual approximations $z_{\tau,h}$, $z_{\tau,h/2}$ and $z_{\tau/2,h/2}$ are computed backwards in time with the corresponding spatial meshes and time steps. The computational cost of the algorithm can be reduced by discarding the approximations in the grey columns; see Remark 4.1.

Now let us denote by $\hat{z} \in \mathcal{V}$ a time-reconstruction of the dual solution $z_{\tau/2,h/2}$ (e.g. a piecewise-constant time-reconstruction will be used in numerical applications; see Section 6). By replacing z with $\hat{z}$ in (32), the estimate of the space-time error in $\mathcal{Q}$ can then be computed as:

$$(36) \qquad \mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h}) \approx \mathcal{E}_{\mathrm{st}} := \mathcal{R}_0\left(u_0^h; \hat{z}(0)\right) + \int_0^T \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); \hat{z}(t)\right) \mathrm{d}t.$$

Replacing z_τ with the computable $z_{\tau,h/2}$ in (29), we compute the spatial error estimate $\mathcal{E}_s$ as:

$$\text{(37)} \qquad \mathcal{R}_s(u_{\tau h}; z_\tau) \approx \mathcal{E}_s := \left(u_0^{h/2} - u_0^h, z_0^{h/2} \right) + \sum_{k=1}^N \tau_k \left\langle r_h^k, z_k^{h/2} \right\rangle.$$

Finally, replacing z and z_τ with the computable $\hat{z}$ and $z_{\tau,h/2}$ in (30), respectively, we compute the temporal error estimate $\mathcal{E}_t$ as:

$$\text{(38)} \quad \mathcal{R}_t(u_{\tau h}, u_\tau, z, z_\tau) \approx \mathcal{E}_t := \left(u_0^{h/2} - u_0^h, \hat{z}(0) - z_0^{h/2} \right)$$
$$+ \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \left\{ \mathcal{R}_{\text{PDE}} \left(I u_{\tau h}(t); \hat{z}(t) - z_{k+1}^{h/2} \right) + \left\langle r_\tau^{k+1}(t), z_{k+1}^{h/2} \right\rangle \right\} \mathrm{d}t.$$

Remark 4.1 (Reduced-cost implementation). If one wants to reduce the number of distinct approximations in the error estimates (36)–(38), the most straightforward strategy is to simply take $u_{\tau,h/2} = u_{\tau/2,h/2}$ and $z_{\tau,h/2} = z_{\tau/2,h/2}$ at concurrent time steps (i.e. $u_k^{h/2} = \tilde{u}_k^{h/2}$ and $z_k^{h/2} = \tilde{z}_k^{h/2}$ for $k = 0, 1, \ldots, N$). In this manner, one only needs to compute $u_{\tau h}$, $u_{\tau/2,h/2}$ and $z_{\tau/2,h/2}$, without the gray columns in Figure 1. More detailed description and examples are given in numerical applications; see Section 6.

There are various cheaper alternatives than $u_{\tau/2,h/2}$ and $z_{\tau/2,h/2}$, for example, reconstructing a higher-order approximation of a low-order approximation. A discussion and a comparison of techniques can be found in Section 4.1 of the book by Bangerth and Rannacher [4].

Significant computational savings can be further obtained by using so-called block-wise adaptivity and coarse-scale dual approximations. These techniques use distinct spatial meshes only between blocks of time-intervals, and compute very cheap dual approximations based on coarse-scale representations of the primal solution. For details we refer to Carey et al. [12].

If an improved approximation $u_{\tau/2,h/2}$ is available, this could be used to directly estimate the error using $Q(u_{\tau/2,h/2}) - Q(u_{\tau,h})$. Note however that this expression is not helpful for determining where to refine in space–time (for example, if $Q(v) = \int_\Omega v(T, x)\, \mathrm{d}x$ then it is clearly not localizable in space–time). One indeed requires a dual solution that contains the sensitivity to errors accumulated at earlier times.

5. Adaptive algorithm. Our goal is now to design an adaptive algorithm to iteratively increase the accuracy of the numerical solution by using the error estimates. In this section, we first derive error indicators of local contributions that serve as the basis to control adaptive mesh refinement and adaptive time-step selection, and then present the space-time adaptive algorithm.

5.1. Error indicators. To drive space-time adaptivity, the information of the global error estimates has to be localized to time-intervals and spatially-local contributions. To this end, we rewrite the computable error estimates $\mathcal{E}_{st}$, $\mathcal{E}_s$ and $\mathcal{E}_t$ in (36)–(38) as a sum of their local contributions on each time intervals $[t_k, t_{k+1}]$, $k = 0, 1, \ldots, N - 1$, respectively. The absolute values of these local contributions are identified as the local indicators, which can directly be used for adaptive time-step selection. For adaptive mesh refinement, the local contributions associated to the spatial discretization error have to be localized further in space. We summarize the result in the following propositions.

PROPOSITION 5.1. *The error estimate $\mathcal{E}_{\mathrm{st}}$, $\mathcal{E}_{\mathrm{t}}$ and $\mathcal{E}_{\mathrm{s}}$ can be bounded from above by*

$$|\mathcal{E}_{\mathrm{st}}| \le \mathcal{E}_{h\tau}^0 + \sum_{k=0}^{N-1} \mathcal{E}_{h\tau}^{k+1} \qquad |\mathcal{E}_{\mathrm{t}}| \le \mathcal{E}_{\tau}^0 + \sum_{k=0}^{N-1} \mathcal{E}_{\tau}^{k+1} \qquad |\mathcal{E}_{\mathrm{s}}| \le \mathcal{E}_{h}^0 + \sum_{k=0}^{N-1} \mathcal{E}_{h}^{k+1}$$

where the local space-time error indicators $\mathcal{E}_{h\tau}^0$ and $\mathcal{E}_{h\tau}^{k+1}$ are defined by

$$(39) \qquad \mathcal{E}_{h\tau}^0 := \left| \mathcal{R}_0\left(u_0^h; \hat{z}(0)\right) \right| \qquad \mathcal{E}_{h\tau}^{k+1} := \left| \int_{t_k}^{t_{k+1}} \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); \hat{z}(t)\right) \mathrm{d}t \right|,$$

the temporal error indicators $\mathcal{E}_{\tau}^0$ and $\mathcal{E}_{\tau}^{k+1}$ are defined by

$$(40) \qquad \mathcal{E}_{\tau}^0 := \left| \left(u_0^{h/2} - u_0^h, \hat{z}(0) - z_0^{h/2} \right) \right|$$

$$(41) \qquad \mathcal{E}_{\tau}^{k+1} := \left| \int_{t_k}^{t_{k+1}} \left\{ \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); \hat{z}(t) - z_{k+1}^{h/2}\right) + \left\langle r_{\tau}^{k+1}(t), z_{k+1}^{h/2} \right\rangle \right\} \mathrm{d}t \right|$$

and the spatial error indicators $\mathcal{E}_{h}^0$ and $\mathcal{E}_{h}^{k+1}$ are defined by

$$(42) \qquad \mathcal{E}_{h}^0 := \left| \left(u_0^{h/2} - u_0^h, z_0^{h/2} \right) \right| \qquad \mathcal{E}_{h}^{k+1} := \left| \tau_{k+1} \langle r_h^{k+1}, z_{k+1}^{h/2} \rangle \right|.$$

Proof. We split the error estimator $\mathcal{E}_{\mathrm{st}}$ (36) into local space-time error indicators (39) by

$$\mathcal{E}_{\mathrm{st}} = \mathcal{R}_0\left(u_0^h; \hat{z}(0)\right) + \int_0^T \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); \hat{z}(t)\right) \mathrm{d}t$$

$$= \mathcal{R}_0\left(u_0^h; \hat{z}(0)\right) + \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); \hat{z}(t)\right) \mathrm{d}t$$

$$\le \left| \mathcal{R}_0\left(u_0^h; \hat{z}(0)\right) \right| + \sum_{k=0}^{N-1} \left| \int_{t_k}^{t_{k+1}} \mathcal{R}_{\mathrm{PDE}}\left(Iu_{\tau h}(t); \hat{z}(t)\right) \mathrm{d}t \right|$$

Following the same procedure as above, the temporal error estimate $\mathcal{E}_{\mathrm{t}}$ (38) is localized to the temporal error indicators (41) on each time intervals $[t_k, t_{k+1}]$, $k = 0, 1, \ldots, N-1$, and the spatial error estimate $\mathcal{E}_{\mathrm{s}}$ (37) is localized in time to the spatial error indicators (42). □

For spatial adaptivity, we consider the following mesh-refinement indicators as in [34, Section 4.2]. These indicators are associated with the support of dual basis functions instead of the support of elements.

PROPOSITION 5.2. *Let $\mathcal{S}_k^{h/2,p} = \mathrm{span}\{\varphi_k^i, i = 1, 2, \ldots, M_k\}$ be the span of finite element shape functions φ_k^i for the enriched dual space. For $k = 0, 1, \ldots, N$ let the dual solution $z_k^{h/2} \in \mathcal{S}_k^{h/2,p}$ of problem (23)-(25) have the expansion:*

$$(43) \qquad z_k^{h/2}(\mathbf{x}) = \sum_{i=1}^{M_k} \bar{z}_k^i \varphi_k^i(\mathbf{x}).$$

Then, the spatial error indicators of Proposition 5.1 satisfy,

$$(44) \qquad \mathcal{E}_h^0 \le \sum_{i=1}^{M_0} \mathcal{E}_{h,i}^0, \qquad and \qquad \mathcal{E}_h^{k+1} \le \sum_{i=1}^{M_k} \mathcal{E}_{h,i}^{k+1}, \qquad k = 0, 1, \ldots, N-1$$

14

where the local mesh-refinement indicators are defined by:

(45a) $$\mathcal{E}_{h,i}^0 = \left| \bar{z}_0^i \left(u_0^{h/2} - u_0^h, \varphi_0^i \right) \right| \qquad \forall i = 1, 2, \ldots, M_0,$$

and for $k = 0, 1, 2, \ldots, N - 1,$

(45b) $$\mathcal{E}_{h,i}^{k+1} = \left| \tau_{k+1} \bar{z}_{k+1}^i \left\langle r_h^{k+1}, \varphi_{k+1}^i \right\rangle \right| \qquad \forall i = 1, 2, \ldots, M_{k+1}.$$

Proof. Simply insert the ansatz (43) into (42) and use the triangle inequality. □

Remark 5.3. Since the above local refinement indicators are associated to supports of basis functions, they are particularly convenient in hierarchical refinement procedures based on those supports instead of a traditional element-wise marking and refinement strategy. We have employed the hierarchical strategy as in Kuru et al. [26] (see also [31, 36]). In this strategy, once supports are marked for refinement, basis functions with support within the marked set are replaced by a set of globally-refined basis functions whose supports are also within the marked set; see [26] for further details.

Remark 5.4. In general in (42), one needs to subtract an interpolant of the dual solution, $\Pi z_{\tau h/2}$ from $z_{\tau,h/2}$ to get a sharp spatial indicator. However, this is not needed for the indicators in (45); see [34, Remark 4.4] for more details.

5.2. The space-time adaptive algorithm. In Algorithm 1, we propose a global space-time adaptive procedure using the above duality-based indicators. The pseudocode consists of three parts: (1) The computation of the primal and dual approximations (comprised of lines 3-8, 9-11 and 13-14), (2) the evaluation of the error estimates (given in lines 12, 15, 18-19 and 23), and (3) the error control (comprised of the remaining lines of Algorithm 1).

Within the adaptive procedure, the error control is built on a two-step approach. First, in lines 16-17 of the pseudocode we apply the maximum marking strategy (following Babuška and Vogelius [3]) with fraction $\theta \in [0, 1]$ on space-time error indicators $\{\mathcal{E}_{h\tau}^k\}_{k=0}^N$ to globally select time steps $\{k'\}$ (i.e. $|\mathcal{E}_{h\tau}^{k'}| \geq \theta \max\{|\mathcal{E}_{h\tau}^0|, |\mathcal{E}_{h\tau}^1|, \ldots, |\mathcal{E}_{h\tau}^N|\}$), which contain the largest error contributions throughout the time period. Second, in line 20 we locally check the leading causes of the error at the targeted time steps $\{k'\}$, whether from the spatial error or from the temporal error. An example of this is shown and explained in Figure 2. The figure on the left indicates the time steps $\{k'\}$, where the major error contributions are located. Then, if the spatial indicator $\mathcal{E}_h^{k'}$ is larger than the temporal indicator $\mathcal{E}_\tau^{k'}$, the spatial mesh is targeted for refinement according to the mesh indicators $\mathcal{E}_i^{k'}$; see Figure 2 (center). Otherwise, the time step size $\tau_{k'}$ is marked and reduced by half ; see Figure 2 (right).

The adaptive spatial mesh refinement is also based on a maximum marking strategy. The nodes $\{i'\}$ are marked for which their mesh-refinement indicators are at least a fraction $\lambda \in [0, 1]$ of the maximal mesh indicator (i.e. $|\mathcal{E}_{i'}^k| \geq \lambda \max\{|\mathcal{E}_0^k|, \ldots, |\mathcal{E}_M^k|\}$). The addition of the basis function on selected nodes is performed using hierarchical refinement for finite element methods [25, 26, 31, 36]. Moreover, instead of projection, we introduce a common refinement to transfer the solution from one mesh to another without loss of accuracy in any quadrature approximations.

Remark 5.5. A standard adaptive algorithm for time-dependent problems starts with an initial coarse mesh, and proceeds *sequentially*. Based on the mesh for the current time step, a new space mesh is generated for each new time step. Such a

Algorithm 1 Duality-based space-time adaptive algorithm

Given: Choose a coarse spatial mesh $\mathcal{K}_0$ and a coarse time step size τ

1: Initialize a list of spatial mesh $\{\mathcal{K}_k\}_{k=1}^N$ ($\mathcal{K}_k = \mathcal{K}_0$) for time steps $\{\tau_k\}_{k=1}^N$ ($\tau_k = \tau$)
2: **while** the maximal error estimate $Max > \text{tol}$ **do**
3:     **for** $k \in \{1, 2, \ldots, N\}$ **do**
4:         Compute $u_{\tau h}$ in $\mathcal{K}_k$ with τ_k
5:         Compute $u_{\tau, h/2}$ in $\mathcal{K}_k^{h/2}$ with τ_k
6:         Compute $u_{\tau/2, h/2}$ in $\mathcal{K}_k^{h/2}$ with $\tau_k/2$
7:         $t = t + \tau_k$
8:     **end for**
9:     **for** $k \in \{N, N-1, \ldots, 1\}$ **do**
10:        Compute $z_{\tau, h/2}$ in $\mathcal{K}_{k-1}^{h/2}$ with τ_k
11:        Compute $z_{\tau/2, h/2}$ in $\mathcal{K}_{k-1}^{h/2}$ with $\tau_k/2$
12:        Estimate the error contribution $\mathcal{E}_{h\tau}^k$
13:        $t = t - \tau_k$
14:     **end for**
15:     Estimate the initial error contribution $\mathcal{E}_{h\tau}^0$
16:     Compute the maximal error contribution for the whole time period $Max = \max\{\mathcal{E}_{h\tau}^0, \ldots, \mathcal{E}_{h\tau}^N\}$
17:     **while** $|\mathcal{E}_{h\tau}^k| > \theta\,|Max|$ **do**
18:        Estimate the local temporal error indicator $\mathcal{E}_\tau^k$
19:        Estimate the local spatial error indicator $\mathcal{E}_h^k$
20:        **if** $\mathcal{E}_\tau^k \geq \mathcal{E}_h^k$ **then**
21:           Refine the time step τ_k by half
22:        **else**
23:           Estimate $\mathcal{E}_i^k$ for the mesh $\mathcal{K}_k$
24:           Refine the mesh $\mathcal{K}_k$ by using hierarchical refinement strategy and maximum strategy with parameter λ
25:        **end if**
26:     **end while**
27: **end while**

---

sequential procedure commonly uses residual-based error estimates in space which only contain information at the current time step. Duality-based error estimates, on the contrary, contain the entire evolution history of the error dependence implicitly via the dual solution.
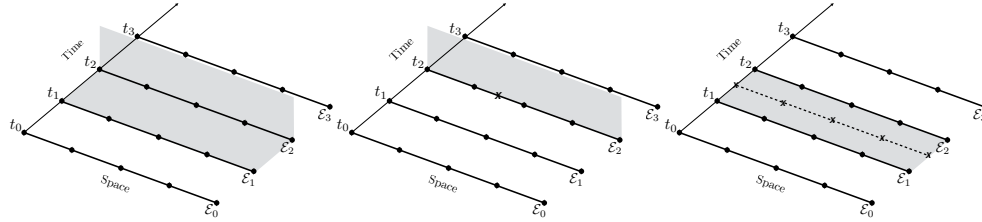

Fig. 2: Adaptive mesh refinement. Left: assume that $\mathcal{E}_{h\tau}^2 = \max\{\mathcal{E}_{h\tau}^0, \mathcal{E}_{h\tau}^1, \mathcal{E}_{h\tau}^2, \mathcal{E}_{h\tau}^3\}$, i.e., the second time step $[t_1, t_2]$ is targeted as the largest error contribution throughout the time period. Middle: If the spatial indicator $\mathcal{E}_h^2$ is larger than the temporal indicator $\mathcal{E}_\tau^2$, the space mesh is targeted to refine according to the mesh indicators $\mathcal{E}_i^2$. Right: If the spatial indicator is smaller than the temporal indicator, the time step size is cut into half.

6. Applications. In this section, we give two examples of problems which fit into the abstract framework introduced in section 2: the nonlinear Allen–Cahn equation and the linear heat equation (as a special case of the Allen–Cahn equation). We numerically investigate the performance of the duality-based error estimates and the proposed adaptive algorithm.

Let us point out that the abstract framework easily accommodates other applications, for example, systems of parabolic equations; see [41, Section 6.3] for the application to a phase-field tumor-growth system.

6.1. Allen–Cahn equation. We subject the (forced) Allen–Cahn equation, $\partial_t u - \Delta u + \epsilon^{-2} \psi'(u) = f(t)$, to homogeneous Neumann boundary conditions. We choose the function spaces as $\mathcal{V} = H^1(\Omega)$, $\mathcal{V}^* = [H^1(\Omega)]^*$ and set $\mathcal{B}(u,v) = (\nabla u, \nabla v)$, $\mathcal{N}(u;v) = \frac{1}{\epsilon^2}(\psi'(u),v)$ in (1), where ϵ is a parameter that controls the thickness of the diffuse interface (typical in phase-field models), and the nonlinear double-well function $\psi(u)$ is defined as (a standard truncated quartic polynomial)[2]

$$(46) \qquad \psi(u) := \begin{cases} (u+1)^2 & u < -1 \\ \frac{1}{4}(u^2 - 1)^2 & u \in [-1,1] \\ (u-1)^2 & u > 1. \end{cases}$$

Then, we obtain the weak form of the Allen–Cahn equation is: Find $u \in \mathcal{W}_{u^0} := \left\{ v \in L^2(0,T;\mathcal{V}), \partial_t v \in L^2(0,T;\mathcal{V}^*) : v(0) = u^0 \right\}$ such that $\forall v \in L^2(0,T;H^1(\Omega))$

$$(47) \qquad \int_0^T \left(\langle \partial_t u, v \rangle + (\nabla u, \nabla v) + \frac{1}{\epsilon^2}(\psi'(u),v) \right) \mathrm{d}t = \int_0^T \langle f, v \rangle \, \mathrm{d}t.$$

In our setting the IMEX scheme for (47) leads to the energy-stable time-stepping scheme introduced in [19]: find $u_{k+1} \in H^1(\Omega)$ such that $\forall v \in H^1(\Omega)$

$$(48) \quad \left(\frac{u_{k+1} - u_k}{\tau_{k+1}}, v \right) + (\nabla u_{k+1}, \nabla v) + \frac{1}{\epsilon^2}(\psi_c'(u_{k+1}),v) - \frac{1}{\epsilon^2}(\psi_e'(u_k),v) = (\bar{f}_{k+1}, v)$$

for $k = 0, 1, \ldots, N-1$, where the initial condition is $(u_0,v) = (u^0, v)$, $\forall v \in L^2(\Omega)$, and where we choose $\bar{f}_{k+1} = \frac{1}{\tau_{k+1}} \int_{t_k}^{t_{k+1}} f(\cdot, t) \, \mathrm{d}t$. In particular, for a splitting of ψ with a quadratic convex part, the resulting system is linear, for example:

$$\psi = \psi_c - \psi_e = \begin{cases} \left(u^2 + \frac{1}{4}\right) - \left(-2u - \frac{3}{4}\right) & u < -1 \\ \left(u^2 + \frac{1}{4}\right) - \left(\frac{3}{2}u^2 - \frac{1}{4}u^4\right) & u \in [-1,1] \\ \left(u^2 + \frac{1}{4}\right) - \left(2u - \frac{3}{4}\right) & u > 1. \end{cases}$$

We then have the full discretization: find $u_{k+1}^h \in \mathcal{S}_{k+1}^{h,1}$ such that $\forall v^h \in \mathcal{S}_{k+1}^{h,1}$

$$(49) \quad \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, v^h \right) + (\nabla u_{k+1}^h, \nabla v^h)$$

$$+ \frac{1}{\epsilon^2}\left(\psi_c'(u_{k+1}^h), v^h\right) - \frac{1}{\epsilon^2}\left(\psi_e'(u_k^h), v^h\right) = \left(\bar{f}_{k+1}, v^h\right)$$

---

[2]Note that by choosing $\psi(u) = 0$, one obtains the linear heat equation.

for $k = 0, 1, \ldots, N - 1$, where the initial condition is $(u_0^h, v^h) = (u^0, v^h)$, $\forall v^h \in \mathcal{S}_0^{h,1}$.

According to the definition of $\mathcal{N}^s$ in (3), we can explicitly write the mean-value linearization of $\psi'(u)$ in terms of u and $\hat{u}$:

$$\psi'^s(u, \hat{u}) = \int_0^1 \psi''(su + (1 - s)\hat{u}) \mathrm{d}s,$$

although, because of its piecewise definition (46), this is an elaborate expression. For example, for $u, \hat{u} > 1$ or $u, \hat{u} < -1$, we have $\psi'^s(u, \hat{u}) = 2$, and for $u, \hat{u} \in [-1, 1]$, we have $\psi'^s(u, \hat{u}) = u^2 + \hat{u}^2 + u\hat{u} - 1$.

Then, by setting $\mathcal{N}^s(u, \hat{u}; w, z) = \frac{1}{\epsilon^2}\left(\psi'^s(u, \hat{u})z, w\right)$ in (5), the dual problem reads: find $z \in \mathcal{W}^{\bar{q}} := \left\{v \in L^2(0, T; \mathcal{V}), \partial_t v \in L^2(0, T; \mathcal{V}^*) : v(T) = \bar{q}\right\}$ such that $\forall w \in L^2(0, T; H^1(\Omega))$

$$(50) \qquad \int_0^T \left(\langle -\partial_t z, w \rangle + (\nabla z, \nabla w) + \frac{1}{\epsilon^2}\left(\psi'^s(u, \hat{u})z, w\right)\right) \mathrm{d}t = \int_0^T (q, w)\, \mathrm{d}t,$$

And the IMEX time-discrete dual problem, based on (23)-(25), is defined by: find $z_k \in H^1(\Omega)$, $k = 0, 1, \ldots, N$, such that

$$(51) \qquad -\left(\frac{z_1 - z_0}{\tau_1}, w\right) - \frac{1}{\epsilon^2}\left(\psi_e'^s(u_0, u_0^h)z_1, w\right) = (q_0, w) \quad \forall w \in H^1(\Omega)$$

and for $k = 1, 2, \ldots, N - 1$:

$$(52) \qquad -\left(\frac{z_{k+1} - z_k}{\tau_k}, w\right) + (\nabla z_k, \nabla w) + \frac{1}{\epsilon^2}\left(\psi_c'^s(u_k, u_k^h)z_k, w\right)$$
$$- \frac{\tau_{k+1}}{\tau_k}\frac{1}{\epsilon^2}\left(\psi_e'^s(u_k, u_k^h)z_{k+1}, w\right) = (q_k, w) \qquad \forall w \in H^1(\Omega)$$

where the terminal condition is

$$(53) \qquad (z_N, w) + \tau_N(\nabla z_N, \nabla w) + \tau_N\frac{1}{\epsilon^2}\left(\psi_c'^s(u_N, u_N^h)z_N, w\right)$$
$$= (\bar{q}, w) + \tau_N(q_N, w) \qquad \forall w \in H^1(\Omega)$$

with q_k, $k = 0, 1, \ldots, N$ defined in Lemma 3.3. Note that for our choice of ψ_c, the derivative $\psi_c'^s$ reduces to a constant. The main results of Section 3 hold, as shown in the following corollary.

COROLLARY 6.1 (Decomposed error representation for Allen–Cahn equation). *The following error representation holds*

$$\mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h}) = \left(z(0), u^0 - u_0^h\right) + \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \Big\{ \langle f(t), z(t) \rangle - \left(\partial_t Iu_{\tau h}(t), z(t)\right)$$
$$- \left(\nabla Iu_{\tau h}(t), \nabla z(t)\right) - \frac{1}{\epsilon^2}\left(\psi'(Iu_{\tau h}(t)), z(t)\right) \Big\} \mathrm{d}t$$
$$= \mathcal{R}_\mathrm{s}(u_{\tau h}; z_\tau) + \mathcal{R}_\mathrm{t}(u_{\tau h}, u_\tau, z, z_\tau),$$

where the spatial error representation reduces to

$$\mathcal{R}_\mathrm{s}(u_{\tau h}; z_\tau) = (z_0, u^0 - u_0^h) + \sum_{k=0}^{N-1} \tau_{k+1} \Big\{ \left(\bar{f}_{k+1}, z_{k+1}\right) - \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, z_{k+1}\right)$$
$$- (\nabla z_{k+1}, \nabla u_{k+1}^h) - \frac{1}{\epsilon^2}\left(\psi_c'(u_{k+1}^h), z_{k+1}\right) + \frac{1}{\epsilon^2}\left(\psi_e'(u_k^h), z_{k+1}\right) \Big\}$$

and the temporal error representation reduces to

$$\mathcal{R}_{\mathrm{t}}(u_{\tau h}, u_\tau, z, z_\tau) = \left(z(0) - z_0, u^0 - u_0^h\right) + \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \bigg\{ \langle f(t), z(t) - z_{k+1} \rangle$$

$$- \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, z(t) - z_{k+1}\right) - (\nabla I u_{\tau h}(t), \nabla z(t)) + (\nabla u_{k+1}^h, \nabla z_{k+1})$$

$$- \frac{1}{\epsilon^2}\left(\psi'(I u_{\tau h}(t)), z(t)\right) + \frac{1}{\epsilon^2}\left(\psi_c'(u_{k+1}^h), z_{k+1}\right) - \frac{1}{\epsilon^2}\left(\psi_e'(u_k^h), z_{k+1}\right) \bigg\}\,\mathrm{d}t.$$

Proof. The result simply follows from Theorem 3.5 applied to (47)–(53). □

6.2. Computable error indicator. Let $u_{\tau h} = \{u_k^{h/2}\}_{k=0}^N$ denote the solution of the fully-discrete primal problem (49) using time-step sizes $\{\tau_k\}_{k=1}^N$ and FE spaces $\{\mathcal{S}_k^{h,1}\}_{k=0}^N$, and let $u_{\tau,h/2} = \{u_k^{h/2}\}_{k=0}^N$ denote the solution of (49) using time-step sizes $\{\tau_k\}_{k=1}^N$ and enriched piecewise-linear FE spaces $\{\mathcal{S}_l^{h/2,1}\}_{l=0,1/2,\ldots,N}$. Replacing u_τ with the computable $u_{\tau,h/2}$ in (51)-(53), we obtain the full discretization of the dual problem using enriched FE spaces: find $z_k^{h/2} \in \mathcal{S}_k^{h/2,1}$, $k = 0, 1, \ldots, N$, such that
(54)

$$-\left(\frac{z_1^{h/2} - z_0^{h/2}}{\tau_1}, w^{h/2}\right) - \frac{1}{\epsilon^2}\left(\psi_e'^s(u_0^{h/2}, u_0^h)z_1^{h/2}, w^{h/2}\right) = (q_0, w^{h/2}) \quad \forall w^{h/2} \in \mathcal{S}_0^{h/2,1}$$

and for $k = 1, 2, \ldots, N-1$:

$$(55) \quad -\left(\frac{z_{k+1}^{h/2} - z_k^{h/2}}{\tau_k}, w^{h/2}\right) + (\nabla z_k^{h/2}, \nabla w^{h/2}) + \frac{1}{\epsilon^2}\left(\psi_c'^s(u_k^{h/2}, u_k^h)z_k^{h/2}, w^{h/2}\right)$$

$$-\frac{\tau_{k+1}}{\tau_k}\frac{1}{\epsilon^2}\left(\psi_e'^s(u_k^{h/2}, u_k^h)z_{k+1}, w^{h/2}\right) = (q_k, w^{h/2}) \qquad \forall w^{h/2} \in \mathcal{S}_k^{h/2,1}$$

where the terminal condition is

$$(56) \quad \left(z_N^{h/2}, w^{h/2}\right) + \tau_N(\nabla z_{N-1}^{h/2}, \nabla w^{h/2}) + \frac{\tau_N}{\epsilon^2}\left(\psi_c'^s(u_N^{h/2}, u_N^h)z_N, w^{h/2}\right)$$

$$= \tau_N(q_N, w^{h/2}) + (\bar{q}, w^{h/2}) \qquad \forall w^{h/2} \in \mathcal{S}_N^{h/2,1}$$

We denote by $z_{\tau,h/2} = \{z_k^{h/2}\}_{k=0}^N$ the solution of (54)-(56). To get $z_{\tau/2,h/2}$, we first compute the space-time enriched approximation $u_{\tau/2,h/2} = \{\tilde{u}_l^{h/2}\}_{l=0,1/2,\ldots,N}$ of the primal problem using half time-step sizes $\{\tau_1/2, \tau_1/2, \tau_2/2, \tau_2/2, \ldots, \tau_N/2, \tau_N/2\}$ and enriched FE spaces $\{\mathcal{S}_l^{h/2,1}\}_{l=0,1/2,\ldots,N}$. Then, replacing u_τ with $u_{\tau/2,h/2}$ in (51)-(53), we compute the space-time enriched approximation $z_{\tau/2,h/2} = \{\tilde{z}_l^{h/2}\}_{l=0,1/2,\ldots,N}$ using the same time-step sizes and FE spaces as $u_{\tau/2,h/2}$.

In the numerical examples for the Allen–Cahn equation in Subsection 6.3 we compute $u_{\tau,h/2}$ and $z_{\tau,h/2}$ directly by taking $u_{\tau,h/2} = u_{\tau/2,h/2}$ and $z_{\tau,h/2} = z_{\tau/2,h/2}$ at concurrent time steps as in Remark 4.1, and consider a piecewise-constant time-reconstruction $\hat{z}$ of $z_{\tau/2,h/2}$ for each time interval $[t_k, t_{k+1})$, $k = 0, 1, \ldots, N-1$, i.e.

$$\hat{z}(t) = \tilde{z}_k^{h/2} \quad \text{for } t \in [t_k, t_{k+1/2}), \qquad \hat{z}(t) = \tilde{z}_{k+1/2}^{h/2} \quad \text{for } t \in [t_{k+1/2}, t_{k+1}).$$

According to (36) and (39)-(42), we then get the following global space-time error estimate:

$$(57) \quad \mathcal{E}_{\mathrm{st}} = \left(\hat{z}(0), u_0^{h/2} - u_0^h\right) + \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \left\{ \langle f(t), \hat{z}(t) \rangle - \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, \hat{z}(t) \right) \right.$$
$$\left. - \left(\nabla I u_{\tau h}(t), \nabla \hat{z}(t)\right) - \frac{1}{\epsilon^2} \left(\psi'(I u_{\tau h}(t)), \hat{z}(t)\right) \right\} \, \mathrm{d}t$$

and the local error indicators

$$\mathcal{E}_{h\tau}^{k+1} = \left| \int_{t_k}^{t_{k+1}} \left\{ \langle f(t), \hat{z}(t) \rangle - \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, \hat{z}(t) \right) - \left(\nabla I u_{\tau h}(t), \nabla \hat{z}(t)\right) \right. \right.$$
$$\left. \left. - \frac{1}{\epsilon^2} \left(\psi'(I u_{\tau h}(t)), \hat{z}(t)\right) \right\} \, \mathrm{d}t \right|$$

$$\mathcal{E}_{\tau}^{k+1} = \left| \int_{t_k}^{t_{k+1}} \left\{ \left\langle f(t), \hat{z}(t) - \tilde{z}_{k+1}^{h/2} \right\rangle \right. \right.$$
$$- \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, \hat{z}(t) - \tilde{z}_{k+1}^{h/2} \right) - \left(\nabla I u_{\tau h}(t), \nabla \hat{z}(t)\right) + (\nabla u_{k+1}^h, \nabla \tilde{z}_{k+1}^{h/2})$$
$$\left. \left. - \frac{1}{\epsilon^2} \left(\psi'(I u_{\tau h}(t)), \hat{z}(t)\right) + \frac{1}{\epsilon^2} \left(\psi_c'(u_{k+1}^h), \tilde{z}_{k+1}^{h/2}\right) - \frac{1}{\epsilon^2} \left(\psi_e'(u_k^h), \tilde{z}_{k+1}^{h/2}\right) \right\} \, \mathrm{d}t \right|$$

$$\mathcal{E}_h^{k+1} = \left| \tau_{k+1} \left\{ \left(\bar{f}_{k+1}, \tilde{z}_{k+1}^{h/2} \right) - \left(\frac{u_{k+1}^h - u_k^h}{\tau_{k+1}}, \tilde{z}_{k+1}^{h/2} \right) - (\nabla \tilde{z}_{k+1}^{h/2}, \nabla u_{k+1}^h) \right. \right.$$
$$\left. \left. - \frac{1}{\epsilon^2} \left(\psi_c'(u_{k+1}^h), \tilde{z}_{k+1}^{h/2}\right) + \frac{1}{\epsilon^2} \left(\psi_e'(u_k^h), \tilde{z}_{k+1}^{h/2}\right) \right\} \right|$$

for $k = 0, 1, \ldots, N-1$ and $i = 0, 1, \ldots, M_k$. And since $z_0^{h/2} = \hat{z}(0) = \tilde{z}_0^{h/2}$,

$$\mathcal{E}_{h\tau}^0 = \mathcal{E}_h^0 = \left(u_0^{h/2} - u_0^h, \tilde{z}_0^{h/2} \right)$$
$$\mathcal{E}_{\tau}^0 = 0.$$

Remark 6.2. For the linear heat equation, $\partial_t u - \Delta u = f$, note that the dual problem and the computable error indicators are easily obtained from the above results by neglecting the nonlinear terms ψ_e' and ψ_c'.

6.3. Numerical results. In the following numerical experiments, we investigate the efficiency of the duality-based error estimates and the performance of the proposed adaptive algorithm. The results will be demonstrated in three parts. In the first part, we illustrate the consistency of the dual time scheme (51)–(53) since we introduced a non-standard IMEX time-discrete dual problem (23)–(25) which contains a nonstandard coefficient τ_{k+1}/τ_k. The second part is on the convergence of error estimate $\mathcal{E}_{\mathrm{st}}$ under uniform refinements for the nonlinear Allen–Cahn equation. In the third part, we apply the proposed duality-based adaptive algorithm to the linear heat equation and the nonlinear Allen–Cahn equation. We compare our adaptive result
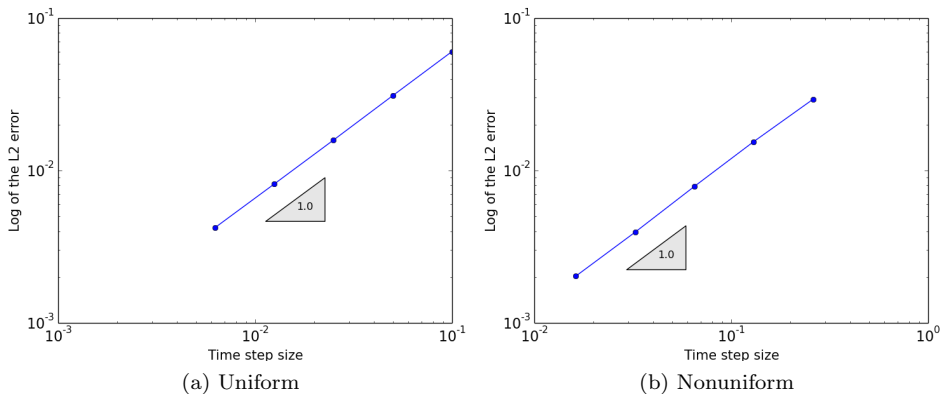
Fig. 3: Accuracy test for dual IMEX scheme: (a) log of the L^2 norm of the error at time T versus time step size (uniform initial time steps); (b) log of the L^2 norm of the error at time T versus maximum time step size (nonuniform initial time steps)

for the heat equation with the sequential-in-time adaptive algorithm (specifically for the heat equation) of Verfürth stated in [38, Sec. 6.8]. For the Allen–Cahn equation, we compare our adaptive results with uniform space-time refinements.

Consistency test. We begin with verifying numerically that the special IMEX time scheme of the dual problem (23)-(25) is first-order accurate in time with respect to refinements of uniform initial time steps (i.e. $\tau_{k+1}/\tau_k = 1$) and, in particular, nonuniform initial time steps (i.e. $\tau_{k+1}/\tau_k \neq 1$). Here, the Allen–Cahn equation is considered in 1D on the domain $\Omega = (-3, 3)$ with parameter $\epsilon = 1$. The spatial mesh is composed of 256 elements (of polynomial degree 1) along the axis. We consider a manufactured solution which oscillates in time:

$$z(\mathbf{x}, t) = e^{-10x^2 + \sin t}$$

The convergence results are presented on a double logarithmic scale in Figure 3. Figure 3a is the convergence of the error in L^2 norm at time T based on time-step refinements using uniform initial time steps $\{\tau_k\}_{k=1}^5 = \{0.1, 0.1, 0.1, 0.1, 0.1\}$, and Figure 3b uses nonuniform initial time steps $\{\tau_k\}_{k=1}^5 = \{2 \cdot 10^{-2}(3k + 1)\}_{k=4}^0 = \{0.26, 0.2, 0.14, 0.08, 0.02\}$. For both uniform and nonuniform initial time steps, the observed rates are close to 1, which demonstrates that (23)-(25) is a first-order time-accurate scheme.

Effectivity test. In this numerical experiment, we consider the Allen–Cahn equation with an exact solution

$$u(\mathbf{x}, t) = \sin(\pi x) \sin(\pi y) e^{-t}$$

on the domain $\Omega = (0, 1) \times (0, 1)$ where $\epsilon = 1$. We suppose that we are interested in the error at final time T, i.e., in (2) we take $q = 0$ and $\bar{q} = u(T) - I u_{\tau h}(T)$ (which is approximated in the computations by $\bar{q} \approx \tilde{u}_N^{h/2} - u_N^h$, see Section 4). The error of interest is thus $\mathcal{Q}(u) - \mathcal{Q}(I u_{\tau h}) = \|u(T) - I u_{\tau h}(T)\|_{L^2(\Omega)}^2 = \|u(T) - u_N^h\|_{L^2(\Omega)}^2$.

To investigate the error estimate with respect to the spatial discretization, temporal discretization and space-time discretization, we compute $\mathcal{E}_{\mathrm{st}}$ according to (57)

under uniform spatial refinement, uniform temporal refinement and uniform space-time refinement, respectively. In view of the nonstandard time discretization scheme for the dual problem, we also investigate the accuracy of the error estimate $\mathcal{E}_{st}$ with respect to uniform and nonuniform initial time step sizes. Table 1 presents the convergence of $\mathcal{E}_{st}$ under uniform spatial refinement for a sufficiently small time step size $(\tau_k = 1e - 4)$. For uniform initial time steps $\{\tau_k\}_{k=1}^4 = \{0.05, 0.05, 0.05, 0.05\}$, the left side of Table 2 shows the convergence of $\mathcal{E}_{st}$ under uniform temporal refinement for a sufficiently fine spatial mesh (128×128 elements) and the left side of Table 3 shows the convergence under uniform space-time refinements. For nonuniform initial time steps $\{\tau_k\}_{k=1}^4 = \{0.08, 0.06, 0.04, 0.02\}$, the convergence result for uniform temporal refinement with a sufficiently fine spatial mesh is presented on the right side of Table 2, and the convergence result for uniform space-time refinements is presented on the right side of Table 3. The effectivity results for different refinements are also presented in the two tables.

M	$\mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h})$	$\mathcal{E}_{st}$	Effectivity
16	0.0003046	0.0002871	0.943
64	1.711e-05	1.606e-05	0.939
256	1.041e-06	9.756e-07	0.938
1024	6.458e-08	6.054e-08	0.937
4096	4.028e-09	3.776e-09	0.937

Table 1: Effectivity of error estimate (57) under spatial refinement

	Uniform initial time steps			Nonuniform initial time steps		
N	$\mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h})$	$\mathcal{E}_{st}$	Eff	$\mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h})$	$\mathcal{E}_{st}$	Eff
4	0.0001378	5.893e-05	0.428	0.0001058	4.773e-05	0.451
8	3.584e-05	1.659e-05	0.463	2.612e-05	1.253e-05	0.480
16	8.918e-06	4.272e-06	0.479	6.247e-06	3.054e-06	0.489
32	2.130e-06	1.032e-06	0.484	1.444e-06	7.063e-07	0.489
64	4.778e-07	2.303e-07	0.482	3.122e-07	1.507e-07	0.483

Table 2: Effectivity of error estimate (57) under temporal refinement with uniform initial time steps and nonuniform initial time steps

	Uniform initial time steps			Nonuniform initial time steps		
$M \times N$	$\mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h})$	$\mathcal{E}_{st}$	Eff	$\mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h})$	$\mathcal{E}_{st}$	Eff
256	6.061e-05	2.583e-05	0.426	4.184e-05	2.013e-05	0.481
2048	2.471e-05	1.086e-05	0.440	1.651e-05	7.581e-06	0.459
16384	7.926e-06	3.717e-06	0.469	5.295e-06	2.534e-06	0.479
131072	2.233e-06	1.082e-06	0.485	1.491e-06	7.299e-07	0.489

Table 3: Effectivity of error estimate (57) under space-time refinement with uniform initial time steps and nonuniform initial time steps

From all the results of Table 1–Table 3, we observe that the error and the estimate converge with the same order under various refinements. The effectivity indices are always between 0.4 and 1, and seem to converge to a constant. This indicates the asymptotic effectivity of the error estimate.

Adaptivity test for the linear heat equation: Clockwise moving source. Let $\Omega = (0,1) \times (0,2)$. We take the example from Asner, Tavener and Kay [2] by

choosing the right hand side and the initial and the boundary condition so that the exact solution is given by

$$u(\mathbf{x},t) = \exp(-100(x - 0.5 - 0.25\sin(\pi t))^2 - 100(y - 1 - 0.5\cos(\pi t))^2)$$

which is an exponential peak moving clockwise inside the domain, see Figure 4. We set up the final time $T = 0.5$ and $\theta = \lambda = 0.8$, and start the adaptive procedure with a uniform mesh containing 4×8 elements and 10 equally distributed time steps (i.e. $\tau = 0.05$). We aim to minimize the error at final time T, i.e., as before, $\mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h}) = \|u(T) - Iu_{\tau h}(T)\|_{L^2(\Omega)}^2 = \|u(T) - u_N^h\|_{L^2(\Omega)}^2$. For the sequential adaptive algorithm, we choose an initial coarse mesh with 4×8 elements and an initial $\tau = 0.01$.
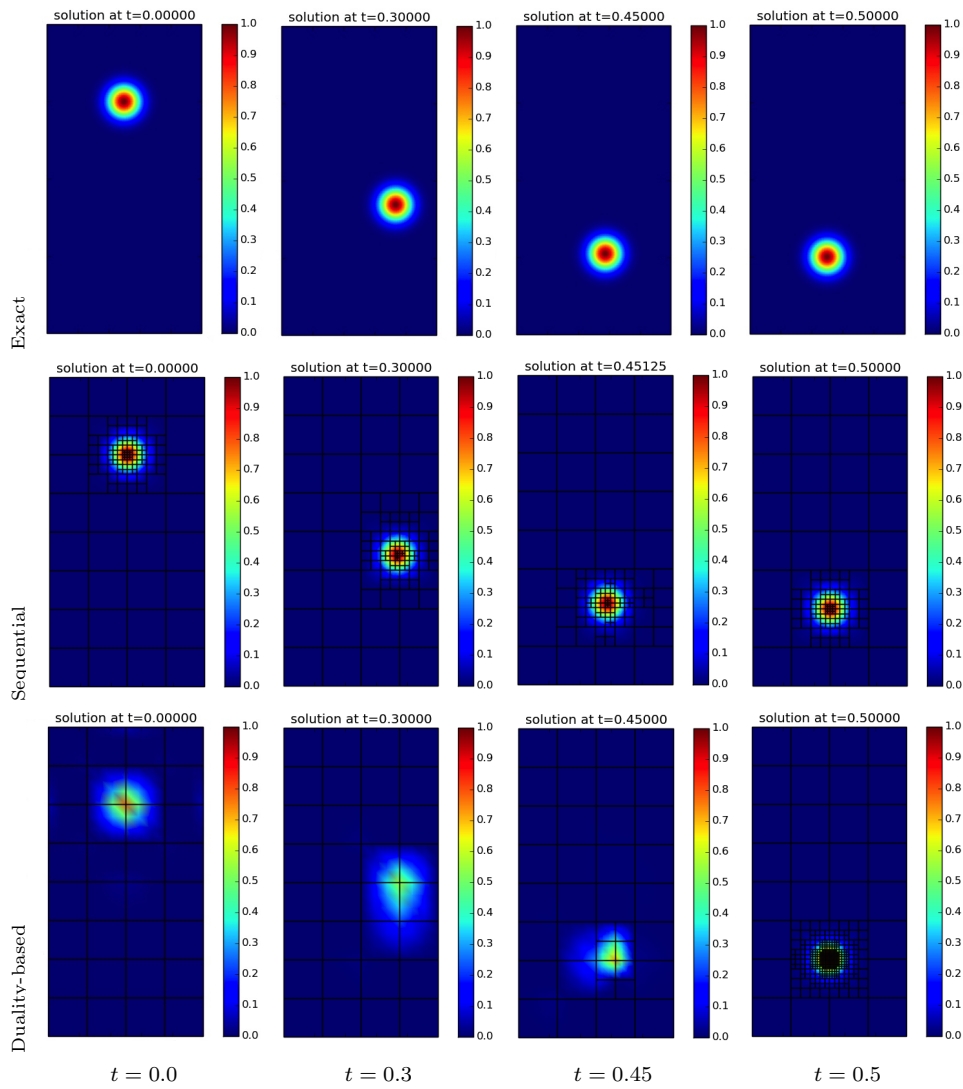
Fig. 4: Snapshot of the exact solution (first row), the sequential adaptive solution with the computational mesh (second row) and the duality-based adaptive solution with the computational mesh (third row) at several time points

23

In Figure 4, we show a comparison of the results obtained by the sequential adaptive algorithm and our duality-based adaptive algorithm. The first row corresponds to the exact solution. The second and third row are the snapshots of adaptively refined meshes with corresponding approximations by using the two adaptive algorithms. Figure 5 illustrates the various time steps used over time for the two adaptive algorithms. From these results, it can be seen that the spatial and temporal refinements of the sequential adaptive algorithm focus on tracing the movement of the Gaussian shaped peak of the solution in the space-time domain. In contrast, since our duality-based adaptive algorithm targets the L^2 norm of the final error, the refinements of our algorithm are only concentrated at the final moment of the space-time domain. Away from this final moment, where residuals contribute much less to the final error, the original mesh and time step already provide sufficient resolution and need not be refined. This leads to fewer degrees of freedom and time steps for reaching the same accuracy of the final solution than for the sequential adaptive algorithm.
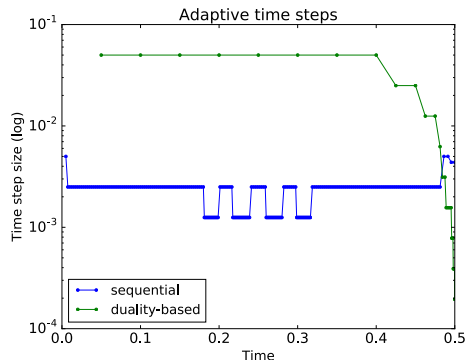

Fig. 5: Adaptive time step refinement for the sequential adaptive algorithm and the non-sequential duality-based adaptive algorithm
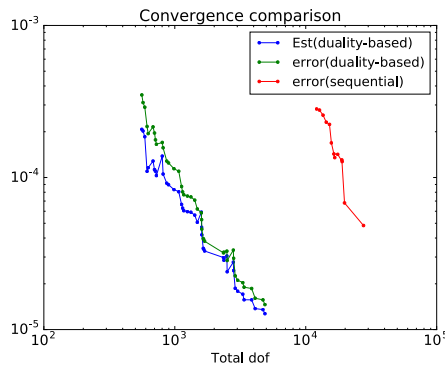
Fig. 6: Convergence comparison between the sequential adaptive algorithm and the non-sequential duality-based adaptive algorithm

In Figure 6, we show the convergence of the dual-based adaptive algorithm versus the 'Total dof', which refers to total number of degrees of freedom, $\sum_{k=0}^{N} M_k$, where $M_k = \dim S_k^{h,1}$ is the total number of spatial degrees of freedom in the primal (linear) FE space at time t_k. Since the error estimate $\mathcal{E}_{st}$ is close to the error, this highlights the accuracy of the error estimate under adaptive refinement. In the same figure, we have also plotted the error of the sequential adaptive algorithm versus the total number of degrees of freedom used, for different final tolerances set in the sequential algorithm (see Algorithm 6.61 in [38, Sec. 6.8]). As expected, the sequential algorithm requires many more total degrees of freedom to attain a particular error tolerance, compared to the dual-based adaptive algorithm. This may seem as an unfair comparison, since the dual-based algorithm is much more expensive than the sequential one, however, the gap between the error curves in Figure 6 can be made arbitrarily large depending on the problem set-up and in particular the corresponding final time T required. This argument shows that for these situations, dual-based adaptive algorithms will certainly outperform sequential ones.

Adaptivity test for the Allen–Cahn equation: Shrinking ring.
The Allen–Cahn dynamics of this test case is a shrinking ring (with diffuse interfaces)

24

in the middle of the domain $\Omega = (-1, 1)^2$. The initial condition is set as

$$(58) \qquad u(0) = -\tanh\left(\frac{\sqrt{x^2 + y^2} - 0.6}{\sqrt{2}\epsilon}\right) + \tanh\left(\frac{\sqrt{x^2 + y^2} - 0.15}{\sqrt{2}\epsilon}\right) - 1$$

where $\epsilon = 0.0625$, see Figure 7 (top-left) for an illustration of the initial condition. The inner circle has a small radius of 0.15 which is expected to vanish much earlier than the outer circle. We are interested in the final error (i.e. $\mathcal{Q}(u) - \mathcal{Q}(Iu_{\tau h}) = \|u - Iu_{\tau h}\|_{L^2}^2(T)$) when the inner circle has disappeared, i.e., the final time $T = 0.02$. To have a reference value for the error, we compute an approximation to (49) on a uniform mesh with 512^2 elements and a uniform time step size $\tau = 1e - 5$. For the adaptive algorithm, we take a coarse initial time-step size $\tau_k = 5e - 3$ and a coarse initial mesh with 16^2 elements. The fractions in the adaptive algorithm are selected as $\theta = \lambda = 0.8$.

Snapshots of the results are presented in Figure 7. The first row shows snapshots of the reference solution, while the second and fourth row show the primal approximation and the corresponding adaptive mesh obtained by the proposed adaptive algorithm. The computed dual solution z is displayed in the third row. It can be observed that the dual solution grows as time progresses. This growth is localized at the interface of the outer ring. Let us mention that, for better visualization, in the plots of the dual solution the range of the color bars is adapted to each plot. We note that the picture of the dual problem at $t = 0.02$ does not display numerical instabilities, but oscillations originating from the final-time condition for the dual solution z. Indeed, in accordance with the quantity of interest, we set the final-time condition $z(T) = u_{N_{\tau/2}}^{h/2} - u_{N_\tau}^h$, where $u_{N_{\tau/2}}^{h/2}$ is the final-time approximation of the primal problem using $\tau/2$ and $h/2$, and $u_{N_\tau}^h$ is the approximation using τ and h. The dual thus exhibits element-wise oscillations due to the difference $u_{N_{\tau/2}}^{h/2} - u_N^h$. Since the dual problem runs backwards, the element-wise oscillations gradually dissipate.

Figure 8 displays the various time steps over time for the shrinking ring. As expected, smaller time steps are need at the vanishing moment of the inner ring and towards the final time T. Furthermore, the interfaces in the solution are well-resolved throughout the simulation time, with a significant increase of resolution towards the final time T.

The convergence of the error estimate and the error for the duality-based adaptive algorithm in comparison to uniform space-time refinement is shown in Figure 9 where 'Total dof' refers again to the total number of primal degrees of freedom $\sum_{k=0}^N M_k$. Note that the error exhibits a plateau for the most refined approximations because the accuracy of the adaptively-refined approximations surpasses that of the reference approximation.

7. Conclusion. In this work we carried out a comprehensive study of duality-based a posteriori error estimates for semi-linear parabolic problems, with a special focus on discretizations using the finite element method in space combined with IMEX time stepping. We introduced a decomposition of the error estimates to identify the separate error contributions due to temporal and spatial approximation. The key idea is to adapt the residual decomposition by Verfürth to our duality-based error representation and propose a specially-tailored time-discrete dual problem. The resultant error indicators quantify the spatial and temporal discretization errors and provide information to drive adaptive mesh refinement and adaptive time-step selection.
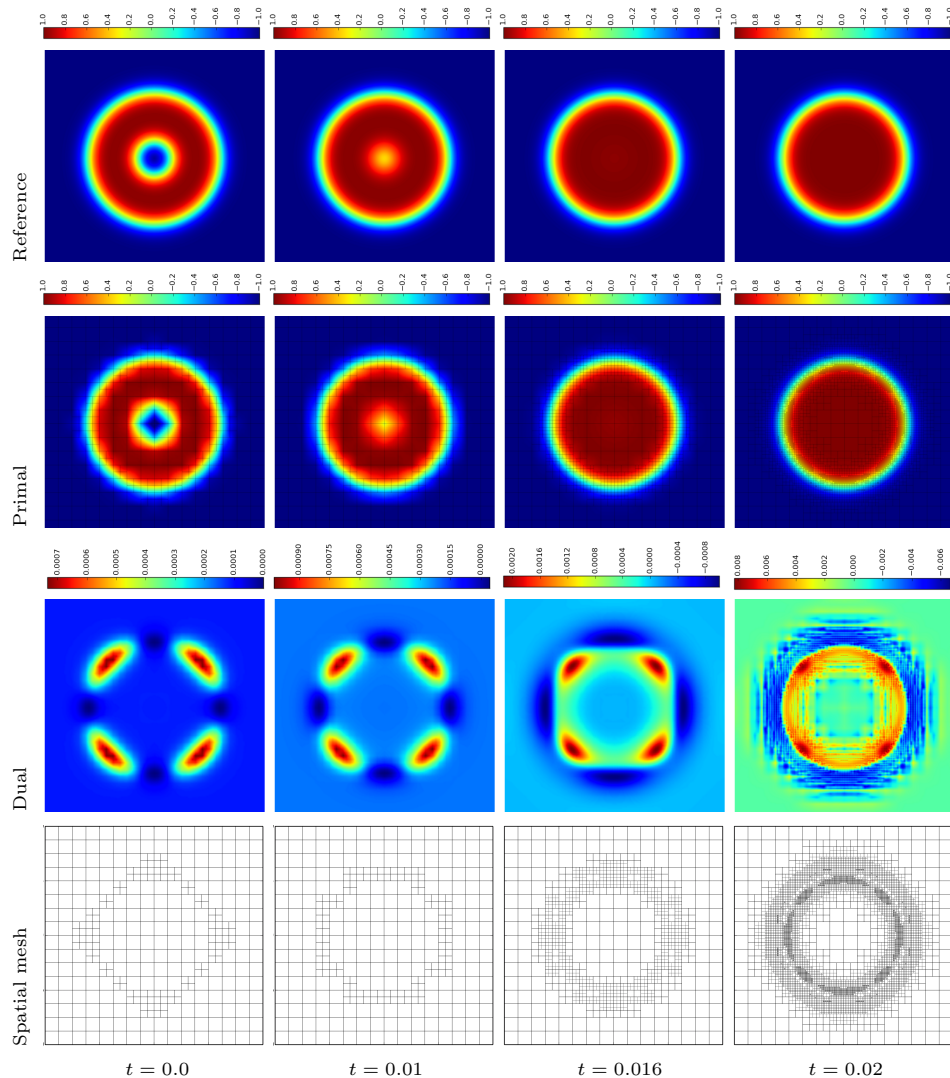
Fig. 7: Shrinking ring: snapshot of the reference solution (first row), the adaptive primal solution (second row), the adaptive dual solution (third row) and the computational mesh (fourth row) at several time points.

To illustrate the performance of the duality-based error estimates and the proposed adaptive algorithm, we presented numerical experiments for the heat equation and Allen–Cahn equation. We refer to [41, Section 6.3] for the application to systems. The numerical results verified the accuracy and the effectivity of the error estimate in test problems. We also observed the overall good quality of the adaptive algorithm.

The proposed methodology can be further extended to other finite difference time-stepping schemes which do not fit in our considered abstract setting, e.g. higher-order multi-stage Runge-Kutta schemes. The key challenge in any such extension is the derivation of a specially-tailored time-discrete dual problem. Our analysis indicates that these dual problems can be derived systematically by means of "backwards"
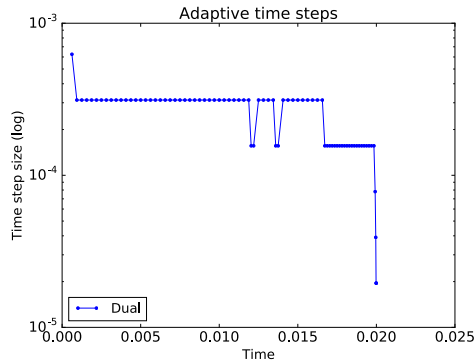
Fig. 8: Adaptive time step refinement for the duality-based adaptive algorithm
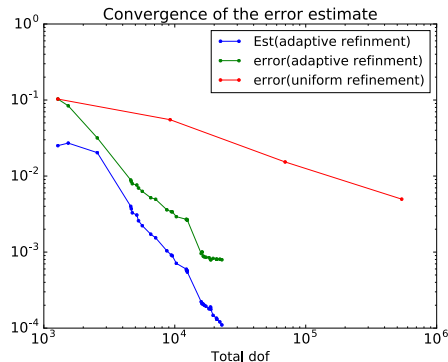

Fig. 9: Convergence of the error estimate for the duality-based adaptive algorithm

summation-by-parts on the time-discrete system.

REFERENCES

[1] U. M. Ascher, *Numerical methods for evolutionary differential equations*, vol. 5, Siam, 2008.

[2] L. Asner, S. Tavener, and D. Kay, *Adjoint-based a posteriori error estimation for coupled time-dependent systems*, SIAM Journal on Scientific Computing, 34 (2012), pp. A2394–A2419.

[3] I. Babuška and M. Vogelius, *Feedback and adaptive finite element solution of one-dimensional boundary value problems*, Numerische Mathematik, 44 (1984), pp. 75–102.

[4] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Lectures in Mathematics, ETH Zürich, Birkhäuser, 2003.

[5] S. Bartels and R. Müller, *A posteriori error controlled local resolution of evolving interfaces for generalized Cahn–Hilliard equations*, Interfaces Free Bound., 12 (2010), pp. 45–73.

[6] R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numer., 10 (2001), pp. 1–102.

[7] R. Bermejo and J. Carpio, *A space–time adaptive finite element algorithm based on dual weighted residual methodology for parabolic equations*, SIAM J. Sci. Comput., 31 (2009), pp. 3324–3355.

[8] M. Besier and R. Rannacher, *Goal-oriented space–time adaptivity in the finite element galerkin method for the computation of nonstationary incompressible flow*, International Journal for Numerical Methods in Fluids, 70 (2012), pp. 1139–1166.

[9] S. Boscarino, F. Filbet, and G. Russo, *High order semi-implicit schemes for time dependent partial differential equations*, Journal of Scientific Computing, (2016), pp. 1–27.

[10] M. Braack, E. Burman, and N. Taschenberger, *Duality based a posteriori error estimation for quasi-periodic solutions using time averages*, SIAM J. Sci. Comput., 33 (2011), pp. 2199–2216.

[11] A. Cangiani, E. H. Georgoulis, I. Kyza, and S. Metcalfe, *Adaptivity and blow-up detection for nonlinear evolution problems*, SIAM Journal on Scientific Computing, 38 (2016), pp. A3833–A3856.

[12] V. Carey, D. Estep, A. Johansson, M. Larson, and S. Taverner, *Blockwise adaptivity for time dependent problems based on coarse scale adjoint solutions*, SIAM J. Sci. Comput., 32 (2010), pp. 2121–2145.

[13] J. H. Chaudhry, J. Collins, and J. N. Shadid, *Error estimation for multi-stage runge-kutta imex schemes*, arXiv preprint arXiv:1509.08576, (2015).

[14] J. H. Chaudhry, D. Estep, V. Ginting, J. N. Shadid, and S. Tavener, *A posteriori error analysis of imex multi-step time integration methods for advection–diffusion–reaction equations*, Computer Methods in Applied Mechanics and Engineering, 285 (2015), pp. 730–751.

[15] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43–77.

[16] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems II:*

Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$, SIAM J. Numer. Anal., 32 (1993), pp. 706–740.

[17] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems IV: Nonlinear problems*, SIAM J. Numer. Anal., 32 (1995), pp. 1729–1749.

[18] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems V: Long-time integration*, SIAM J. Numer. Anal., 32 (1995), pp. 1750–1763.

[19] D. EYRE, *Unconditionally gradient stable time marching the Cahn–Hilliard equation*, in Computational and Mathematical Models of Microstructural Evolution, J. W. Bullard, L.-Q. Chen, R. K. Kalia, and A. M. Stoneham, eds., vol. 529 of Mater. Res. Soc. Symp. Proc., Warrendale, Pennsylvania, 1998, Materials Research Society, pp. 39–46.

[20] H. GOMEZ AND T. J. R. HUGHES, *Provably unconditionally stable, second-order time-accurate, mixed variational methods for phase-field models*, J. Comput. Phys., 230 (2011), pp. 5310–5327.

[21] H. GOMEZ AND K. G. VAN DER ZEE, *Computational phase-field modeling*, in Encyclopedia of Computational Mechanics, Second Edition, E. Stein, R. de Borst, and T. J. R. Hughes, eds., Wiley, 2017. Part 1 Fluids.

[22] D. HILHORST, J. KAMPMANN, T.-N. NGUYEN, AND K. G. VAN DER ZEE, *Formal asymptotic limit of a diffuse-interface tumor-growth model*, Math. Models Methods Appl. Sci., 25 (2015), pp. 1011–1043.

[23] D. KESSLER, R. H. NOCHETTO, AND A. SCHMIDT, *A posteriori error control for the allen–cahn problem: circumventing gronwall's inequality*, ESAIM: Mathematical Modelling and Numerical Analysis, 38 (2004), pp. 129–142.

[24] J. KIM AND J. LOWENGRUB, *Phase field modeling and simulation of three-phase flows*, Interfaces Free Bound., 7 (2005), pp. 435–466.

[25] P. KRYSL, E. GRINSPUN, AND P. SCHRÖDER, *Natural hierarchical refinement for finite element methods*, International Journal for Numerical Methods in Engineering, 56 (2003), pp. 1109–1124.

[26] G. KURU, C. V. VERHOOSEL, K. G. VAN DER ZEE, AND E. H. VAN BRUMMELEN, *Goal-adaptive isogeometric analysis with hierarchical splines*, Comput. Methods Appl. Mech. Engrg., 270 (2014), pp. 270–292.

[27] O. LAKKIS, C. MAKRIDAKIS, AND T. PRYER, *A comparison of duality and energy a posteriori estimates for $L_\infty(0, T; L_2(\Omega))$ in parabolic problems*, Math. Comp., 84 (2014), pp. 1537–1569.

[28] C. MAKRIDAKIS AND R. H. NOCHETTO, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 1585–1594.

[29] J. T. ODEN, A. HAWKINS, AND S. PRUDHOMME, *General diffuse-interface theories and an approach to predictive tumor growth modeling*, Math. Models Methods Appl. Sci., 20 (2010), pp. 477–517.

[30] J. T. ODEN AND S. PRUDHOMME, *Goal-oriented error estimation and adaptivity for the finite element method*, Comput. Math. Appl., 41 (2001), pp. 735–756.

[31] T. RICHTER AND T. WICK, *Variational localizations of the dual-weighted residual estimator*, J. Comp. Appl. Math., 279 (2015), pp. 192–208.

[32] M. SCHMICH AND B. VEXLER, *Adaptivity with dynamic meshes for space–time finite element discretization of parabolic equations*, SIAM J. Sci. Comput., 30 (2008), pp. 369–393.

[33] J. SHEN AND X. YANG, *Decoupled, energy stable schemes for phase-field models of two-phase incompressible flows*, SIAM J. Numer. Anal., 53 (2015), pp. 279–296.

[34] G. ŞIMŞEK, X. WU, K. VAN DER ZEE, AND E. VAN BRUMMELEN, *Duality-based two-level error estimation for time-dependent pdes: Application to linear and nonlinear parabolic equations*, Computer Methods in Applied Mechanics and Engineering, 288 (2015), pp. 83–109.

[35] G. TIERRA AND F. GUILLÉN-GONZÁLEZ, *Numerical methods for solving the Cahn–Hilliard equation and its applicability to related energy-based models*, Arch. Comput. Methods Eng., 22 (2015), pp. 269–289.

[36] E. VAN BRUMMELEN, S. ZHUK, AND G. VAN ZWIETEN, *Worst-case multi-objective error estimation and adaptivity*, Comput. Methods Appl. Mech. Engrg., 313 (2017), pp. 723–743.

[37] K. G. VAN DER ZEE, J. T. ODEN, S. PRUDHOMME, AND A. HAWKINS-DAARUD, *Goal-oriented error estimation for Cahn–Hilliard models of binary phase transition*, Numerical Methods for Partial Differential Equations, 27 (2011), pp. 160–196.

[38] R. VERFÜRTH, *A posteriori error estimation techniques for finite element methods*, Oxford University Press, 2013.

[39] P. VIGNAL, L. DALCIN, D. L. BROWN, N. COLLIER, AND V. M. CALO, *An energy-stable convex splitting for the phase-field crystal equation*, Comput. & Structures, 158 (2015), pp. 355–368.

[40] S. M. WISE, C. WANG, AND J. S. LOWENGRUB, *An energy-stable and convergent finite-difference scheme for the phase field crystal equation*, SIAM J. Numer. Anal., 47 (2009), pp. 2269–2288.

[41] X. WU, *Space-Time Adaptive Methods for Phase-Field Models*, PhD thesis, Technische Universiteit Eindhoven, The Netherlands, June 2017. Available at http://repository.tue.nl/866034.

[42] X. WU, G. J. VAN ZWIETEN, AND K. G. VAN DER ZEE, *Stabilized second-order convex splitting schemes for Cahn–Hilliard models with application to diffuse-interface tumor-growth models*, Int. J. Numer. Meth. Biomed. Engng., 30 (2014), pp. 180–203.