ESAIM: Modélisation mathématique et analyse numérique

Zhiming Chen
Ricardo H. Nochetto
Alfred Schmidt

Error control and adaptivity for a phase relaxation model

ESAIM: Modélisation mathématique et analyse numérique, tome 34, n° 4 (2000), p. 775-797
<http://www.numdam.org/item?id=M2AN_2000__34_4_775_0>

Mathematical Modelling and Numerical Analysis ESAIM: M2AN
Modélisation Mathématique et Analyse Numérique Vol. 34, N° 4, 2000, pp. 775–797

ERROR CONTROL AND ADAPTIVITY FOR A PHASE RELAXATION MODEL

ZHIMING CHEN[1], RICARDO H. NOCHETTO[2] AND ALFRED SCHMIDT[3]

Abstract. The phase relaxation model is a diffuse interface model with small parameter ε which consists of a parabolic PDE for temperature θ and an ODE with double obstacles for phase variable χ. To decouple the system a semi-explicit Euler method with variable step-size τ is used for time discretization, which requires the stability constraint $\tau \leq \varepsilon$. Conforming piecewise linear finite elements over highly graded simplicial meshes with parameter h are further employed for space discretization. *A posteriori* error estimates are derived for both unknowns θ and χ, which exhibit the correct asymptotic order in terms of ε, h and τ. This result circumvents the use of duality, which does not even apply in this context. Several numerical experiments illustrate the reliability of the estimators and document the excellent performance of the ensuing adaptive method.

Mathematics Subject Classification. 65N15, 65N30, 65N50, 80A22, 35M05, 35R35.

Received: July 2, 1999.

1. INTRODUCTION

Let $\Omega \subset \mathbb{R}^d (d = 1, 2, 3)$ be a bounded convex polyhedral domain, $\Gamma = \partial\Omega$, and $Q_T = \Omega \times (0, T)$ for some $0 < T < +\infty$. Given a small parameter ε, we consider the phase relaxation model of Visintin [18, 19]

$$\partial_t\theta + \partial_t\chi - \Delta\theta = f \qquad \text{in } Q_T, \tag{1.1}$$
$$\varepsilon\partial_t\chi + \Lambda(\chi) \ni \theta \qquad \text{in } Q_T, \tag{1.2}$$

subject to the parabolic boundary conditions

$$\theta = 0 \qquad \text{on } \Gamma \times (0, T), \tag{1.3}$$
$$\theta(\cdot, 0) = \theta_0(\cdot), \; \chi(\cdot, 0) = \chi_0(\cdot) \qquad \text{in } \Omega. \tag{1.4}$$

Here θ stands for the temperature of a substance that occupies the domain Ω and undergoes solidification, χ is the phase variable (or order parameter), $u = \theta + \chi$ is the enthalpy, and Λ is the inverse of the sign function,

Keywords and phrases. Phase relaxation, diffuse interface, subdifferential operator, finite elements, *a posteriori* estimates, adaptivity.

[1] Institute of Mathematics, Academia Sinica, Beijing 100080, PR China. The first author was partially supported by the National Natural Science Foundation of China under the grant No. 19771080 and China National Key Project "Large Scale Scientific and Engineering Computing".

[2] Department of Mathematics, University of Maryland, College Park, MD 20742, USA. Partially supported by NSF Grant DMS-9623394 and NSF SCREMS 9628467.

[3] Institut für Angewandte Mathematik, Universität Freiburg, 79104 Freiburg, Germany. Partially supported by DFG and EU Grant HCM "Phase Transitions and Surface Tension". e-mail: alfred@mathematik.uni-freiburg.de

namely, the following maximal monotone graph

$$\Lambda(s) = \begin{cases} [-\infty, 0] & \text{if } s = -1, \\ 0 & \text{if } -1 < s < 1, \\ [0, +\infty] & \text{if } s = 1. \end{cases} \tag{1.5}$$

This graph Λ forces χ to be within ± 1, the obstacles. Liquid and solid phases are determined by the coincidence sets where either $\chi = 1$ or $\chi = -1$, respectively. The remaining non-contact set is the so-called *transition region*, where the phase change takes place. The thickness of such a region is $O(\sqrt{\varepsilon})$ for smooth evolutions, as dictated by the behavior of 1d traveling waves [12]. This model allows for θ to exhibit values above the melting temperature in the solid phase (*super-heating*) and below in the liquid phase (*under-cooling*). This simple model however does not exhibit surface tension effects.

This paper introduces, analyzes and tests an adaptive finite element method for (1.1)-(1.4) based on *a posteriori error control.* The issue at stake is the strong nonlinearity hidden in Λ that prevents duality techniques to apply in the present context. Duality hinges on linearization, and has been systematically used to derive *a posteriori* error estimates for parabolic problems, both linear [5,7] and mildly nonlinear [6]. It has even been employed for the simplest model of liquid-solid phase change, namely the Stefan problem [15]. The analysis of [15], which is strikingly different from those of [5–7] already hints at the limitations of duality in dealing with strong nonlinearities. The new difficulty posed by the above model is the lack of Lipschitz regularity of Λ. This is an essential feature that cannot be overcome by regularization, since in that case χ must be updated in the entire Ω instead of a narrow transition region; this is indeed a major advantage of the double obstacle formulation. Moreover, the regularization error should be accounted for *a posteriori*, which is problematic.

A new technique has been recently introduced in [13,14] for *implicit* time stepping of nonlinear evolution equations for subdifferential operators which are not necessarily smooth. This is the case of our system, since $\Lambda = \partial\Phi$ is the subdifferential of the proper, lower semi-continuous, and convex functional $\Phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$

$$\Phi(s) = \begin{cases} 0 & \text{if } |s| \leq 1, \\ +\infty & \text{otherwise;} \end{cases} \tag{1.6}$$

see Section 4.2 for details. In this paper we study the *semi-explicit* method of [16,17] for time discretization, combined with conforming P^1 finite elements for space discretization, which are discussed in Sections 3.1 and 3.2 respectively. We establish *a posteriori* error estimators for both the semi-discrete scheme in Section 4.2 and the fully discrete scheme in Section 4.3, and state that the former are optimal with respect to both rate of convergence and regularity requirements in Section 4.2. Full proofs are given in Sections 5 and 6. We also present several numerical experiments in Section 7 that document the efficiency and robustness of the ensuing adaptive method, which handles topological changes automatically.

In contrast to [11], which uses *a priori* information for mesh design, the present paper is entirely based on *a posteriori* error estimation. Instrumental to this goal is an interpolation operator which extends a node-wise relation stemming from the discretization of (1.2) to the whole domain, but does *not* cause additional error contributions in the discrete coincidence region; similar ideas were previously used in [11]. The importance of this property is evident in practical applications of error control since unnecessary mesh refinement should always be avoided in the discrete coincidence region.

A new positivity preserving interpolation operator has been recently introduced in [1] for *a posteriori* error estimation of elliptic obstacle problems. This operator, in conjunction with ideas in this paper, constitutes the basis for mesh design and analysis of adaptive methods for diffuse interface models [2].

2. Setting

Let $\mathcal{V} = H_0^1(\Omega)$, $\mathcal{V}^* = H^{-1}(\Omega)$ be the dual space of $\mathcal{V}$, and $\mathcal{W} = L^2(\Omega)$. Let $\mathcal{K} = \{\, \phi \in \mathcal{W} : \ |\phi| \leq 1 \text{ a.e. on } \Omega\}$ be a (convex) set of admissible functions. The data satisfy the following hypotheses (weaker ones are needed in Sects. 4.2 and 5):

(H1) Initial temperature: $\|\, \theta_0 \,\|_{W_0^{1,\infty}(\Omega)} \leq C$.

(H2) Initial phase variable: $\chi_0 \in \mathcal{K} \cap C(\bar{\Omega})$.

(H3) Compatibility condition: $\|\theta_0 - z_0\|_{L^2(\Omega)} \leq A\sqrt{\varepsilon}$ with $z_0 \in \Lambda(\chi_0)$ given by

$$z_0 = \begin{cases} \theta_0 - \max(\theta_0, 0) & \text{if } \chi_0 = -1, \\ 0 & \text{if } -1 < \chi_0 < 1 \\ \theta_0 - \min(\theta_0, 0) & \text{if } \chi_0 = 1. \end{cases}$$

(H4) Source term: $f \in W^{1,1}(0,T; C(\bar{\Omega}))$.

The compatibility condition (H3) penalizes the lack of satisfaction of $\theta_0 \in \Lambda(\chi_0)$ as well as the thickness of the initial transition region $\mathcal{T}(0)$. If

$$\mathcal{T}(t) = \{x \in \Omega : |\chi(x,t)| < 1\}$$

denotes the *transition region*, its thickness is $O(\sqrt{\varepsilon})$ for smooth solutions as revealed by the analysis of traveling waves [12].

The weak formulation of the problem (1.1)-(1.4) reads as follows: *Find θ and χ such that*

$$\theta \in L^2(0,T; H_0^1(\Omega)) \cap H^1(0,T; H^{-1}(\Omega)), \quad \chi \in H^1(0,T; L^2(\Omega)),$$

$$\theta(\cdot,0) = \theta_0(\cdot), \quad \chi(\cdot,0) = \chi_0(\cdot),$$

and for a.e. $t \in (0,T)$ the following relations hold

$$\langle \partial_t u, \phi \rangle + \langle \nabla\theta, \nabla\phi \rangle = \langle f, \phi \rangle \qquad \forall\, \phi \in \mathcal{V}, \tag{2.1}$$

$$\langle \varepsilon\partial_t\chi, \chi - \varphi \rangle \leq \langle \theta, \chi - \varphi \rangle \qquad \forall\, \varphi \in \mathcal{K}. \tag{2.2}$$

Hereafter, $u := \theta + \chi$, and $\langle \cdot,\cdot \rangle$ stands for either the inner product in $L^2(\Omega)$ or the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

In view of (H1–4) the following *a priori* estimates are valid uniformly in ε:

$$\|\, \theta \,\|_{L^\infty(0,T;H_0^1(\Omega))} + \sqrt{\varepsilon}\|\, \partial_t\chi \,\|_{L^\infty(0,T;L^2(\Omega))} + \|\, \theta \,\|_{H^1(0,T;L^2(\Omega))} \leq C. \tag{2.3}$$

3. Discretization

In this section we discuss the time discretization, which decouples the variables θ and χ, as well as the space discretization by finite elements.

3.1. Time discretization

We now introduce the *semi-explicit* scheme of Verdi and Visintin for solving (2.1)-(2.2) [16, 17]. Let τ_n be the nth time step and set

$$t^n := \sum_{i=1}^{n} \tau_n, \quad \varphi(\cdot) := \varphi(\cdot, t^n)$$

for any function φ continuous in $(t^{n-1}, t^n]$. Let N be the total number of steps, that is $t^N \geq T$. The semi-discrete problem then reads as follows: *Given* $\Theta^0 = \theta_0$ *and* $X^0 = \chi_0$, *find* $\Theta^n \in \mathcal{V}, \mathrm{X}^n \in \mathcal{K}, U^n \in \mathcal{W}$ *for* $1 \leq n \leq N$ *such that*

$$\varepsilon\Big\langle \frac{\mathrm{X}^n - \mathrm{X}^{n-1}}{\tau_n}, \mathrm{X}^n - \varphi \Big\rangle \leq \langle \Theta^{n-1}, \mathrm{X}^n - \varphi \rangle \quad \forall\, \varphi \in \mathcal{K}, \tag{3.1}$$

$$\Big\langle \frac{U^n - U^{n-1}}{\tau_n}, \phi \Big\rangle + \langle \nabla\Theta^n, \nabla\phi \rangle = \langle f^n, \phi \rangle \quad \forall\, \phi \in \mathcal{V}, \tag{3.2}$$

where $U^n = \Theta^n + \mathrm{X}^n$. We notice that (3.1) is merely an algebraic correction for X^n in terms of Θ^{n-1} and X^{n-1}, which decouples from the *linear* elliptic PDE (3.2) for Θ^n. The price for this appealing feature is the stability constraint $\tau_n \leq \varepsilon$. We point out that (3.1) may be replaced by the equivalent equation

$$\varepsilon\Big\langle \frac{\mathrm{X}^n - \mathrm{X}^{n-1}}{\tau_n}, \varphi \Big\rangle + \langle \mathrm{Z}^n, \varphi \rangle = \langle \Theta^{n-1}, \varphi \rangle \; \forall \varphi \in L^2(\Omega). \tag{3.3}$$

Here $\mathrm{Z}^n \in \Lambda(\mathrm{X}^n)$ may be viewed as a Lagrange multiplier associated to the unilateral constraint $|\mathrm{X}^n| \leq 1$.

In contrast to (3.1), the *implicit* time discretization

$$\varepsilon\Big\langle \frac{\mathrm{X}^n - \mathrm{X}^{n-1}}{\tau_n}, \mathrm{X}^n - \varphi \Big\rangle \leq \langle \Theta^n, \mathrm{X}^n - \varphi \rangle \; \forall \varphi \in \mathcal{K}, \tag{3.4}$$

does not suffer from any stability constraint but leads to a coupled system. Both methods are studied in [10] for constant step-size. In Section 4.2 we state *a posteriori* error estimates for (3.1)-(3.4) along with their optimal asymptotic rate, and prove them in Section 5. In addition to providing computable error estimators, these results simplify and extend the error analysis of [10] to variable step-sizes.

3.2. Space discretization

We now combine the semi-explicit method of Section 3.1 with conforming piecewise linear finite elements in space.

We denote by $\mathcal{M}^n$ a uniformly regular partition of Ω into simplexes [3]. We use refining/coarsening procedures based on bisection, which lead to *compatible* consecutive meshes; this is a major difference with the method proposed in [11]. Given a triangle $S \in \mathcal{M}^n$, h_S stands for its diameter and ρ_S for its sphericity and they satisfy $h_S \leq 2\rho_S / \sin(\gamma_S/2)$, where γ_S is the minimum angle of S. Shape regularity of the family of triangulations is equivalent to $\gamma_S \geq \gamma > 0$, with γ independent of n. We denote by $\mathcal{B}^n$ the collection of interior inter-element boundaries or sides e of $\mathcal{M}^n$ in Ω; h_e stands for the size of $e \in \mathcal{B}^n$.

Let $\mathcal{V}^n$ indicate the usual space of piecewise linear finite elements over $\mathcal{M}^n$ and $\mathcal{V}_0^n = \mathcal{V}^n \cap \mathcal{V}$. Let $\{x_j^n\}_{j=1}^{K^n}$ denote the nodes of $\mathcal{M}^n$. Let $I^n : C(\bar{\Omega}) \to \mathcal{V}^n$ be the usual Lagrange interpolation operator, namely $(I^n\varphi)(x_j^n) = \varphi(x_j^n)$ for all $1 \leq j \leq K^n$. Finally, let the discrete inner products $\langle \cdot, \cdot \rangle^n$ be defined by

$$\langle \varphi, \psi \rangle^n = \sum_{S \in \mathcal{M}^n} \int_S I^n(\varphi\psi) \mathrm{d}x,$$

for any piecewise uniformly continuous functions φ, ψ. This vertex quadrature rule is easy to evaluate in practice [3]. It is known that for any $S \in \mathcal{M}^n$ [3, 15]

$$\Big| \int_S \varphi\psi \mathrm{d}x - \int_S I^n(\varphi\psi) \mathrm{d}x \Big| \leq \frac{1}{8} h_S^2 \| \nabla\varphi \|_{L^2(S)} \| \nabla\psi \|_{L^2(S)} \;\; \forall \varphi, \psi \in \mathcal{V}^n. \tag{3.5}$$

Let $\Theta^0 := I^0\theta_0$ and $X^0 := I^0\chi_0$, which makes sense in view of (H1) and (H2). Then the fully discrete finite element approximation reads as follows: *Given* $\Theta^{n-1} \in \mathcal{V}_0^{n-1}, X^{n-1} \in \mathcal{V}^{n-1}$ *and* $Z^{n-1} \in \mathcal{V}^{n-1}$, *then* $\mathcal{M}^{n-1}$ *and*

τ_{n-1} are modified as described below to give rise to $\mathcal{M}^n$ and τ_n and thereafter $\Theta^n \in \mathcal{V}_0^n, X^n \in \mathcal{V}^n$ and $Z^n \in \mathcal{V}^n$ computed according to the following prescription

$$\varepsilon\Big\langle \frac{X^n - I^n X^{n-1}}{\tau_n}, \varphi\Big\rangle^n + \langle Z^n, \varphi\rangle^n = \langle I^n \Theta^{n-1}, \varphi\rangle^n \quad \forall\, \varphi \in \mathcal{V}^n, \tag{3.6}$$

$$\Big\langle \frac{U^n - I^n U^{n-1}}{\tau_n}, \varphi\Big\rangle^n + \langle \nabla\Theta^n, \nabla\varphi\rangle = \langle I^n f^n, \varphi\rangle^n \quad \forall\, \varphi \in \mathcal{V}_0^n, \tag{3.7}$$

with $U^n := \Theta^n + X^n$ and $Z^n(x_j^n) \in \Lambda(X^n(x_j^n))$ for $1 \le j \le K^n$.

We note that (3.6) is a simple and inexpensive node-wise algebraic correction which computes X^n and Z^n in terms of Θ^{n-1} and X^{n-1}. In fact, (3.6) can be rewritten explicitly as

$$\frac{\varepsilon}{\tau_n}(X^n(x_j^n) - X^{n-1}(x_j^n)) + Z^n(x_j^n) - \Theta^{n-1}(x_j^n) = 0 \quad \forall\, 1 \le j \le K^n, \tag{3.8}$$

or equivalently as

$$X^n(x_j^n) = (I - \beta)\Big(\frac{\tau_n}{\varepsilon} I^n \Theta^{n-1} + I^n X^{n-1}\Big)(x_j^n) \quad \forall\, 1 \le j \le K^n, \tag{3.9}$$

where $\beta(s) = \min(s+1, 0) + \max(s-1, 0)$ and I is the identity. Once X^n is computed, then Z^n can be obtained node-wise from (3.6). On the other hand, (3.7) can be rewritten equivalently as

$$\langle \Theta^n, \varphi\rangle^n + \tau_n\langle \nabla\Theta^n, \nabla\varphi\rangle = \langle I^n\Theta^{n-1} - X^n + I^n X^{n-1} + \tau_n I^n f^n, \varphi\rangle^n \quad \forall\, \varphi \in \mathcal{V}_0^n. \tag{3.10}$$

This is a linear positive definite symmetric system in the unknown Θ^n, which can be solved efficiently by the conjugate gradient method with BPX preconditioner for instance.

It is crucial to extend the node-wise relation (3.8) to the entire domain Ω while preserving monotonicity properties. The key issue is to capture the different behavior of Θ^n and X^n within the *transition regions* $\mathcal{T}^n$ and outside, where

$$\mathcal{T}^n = \{x \in \Omega : \ |\mathrm{X}^n(x)| < 1\}. \tag{3.11}$$

Since $\mathrm{X}^n \in \mathcal{K}^n$ is piecewise linear, the closure of $\mathcal{T}^n$ is a union of simplices. We thus introduce the operator $P^n : \mathcal{V}^n \to L^\infty(\Omega)$ defined element-wise as follows: for any $\varphi \in \mathcal{V}^n$ and $S \in \mathcal{M}^n$ with vertices $\{x_{j(i)}^n\}_{i=1}^{d+1}$ and barycentric coordinates $\{\lambda_i\}_{i=1}^{d+1}$, define

$$P^n\varphi|_S = \begin{cases} \varphi|_S & \text{if } S \cap \mathcal{T}^n = \emptyset, \\ \displaystyle\sum_{i=1}^{d+1} \varphi(x_{j(i)}^n)\chi_i & \text{otherwise,} \end{cases} \tag{3.12}$$

where χ_i is the characteristic function of the polyhedral set

$$S_i := \{x \in S : \lambda_i(x) \ge \lambda_l(x), \ l \ne i\}.$$

It is now easy to see that $P^n Z^n \in \Lambda(P^n X^n)$ a.e. in Ω and that (3.8) is equivalent to the global relation

$$\varepsilon\, \frac{P^n X^n - P^n(I^n X^{n-1})}{\tau_n} + P^n Z^n - P^n(I^n\Theta^{n-1}) = 0 \ \text{ a.e. in } \Omega. \tag{3.13}$$

Moreover, there exists a constant C, depending only on the minimum angle of the mesh $\mathcal{M}^n$, such that

$$\| P^n\varphi \|_{L^2(\Omega)} \le C\| \varphi \|_{L^2(\Omega)} \quad \forall\, \varphi \in \mathcal{V}^n, \tag{3.14}$$

$$\| P^n\varphi - \varphi \|_{L^2(\Omega)} \le C\| h_n \nabla\varphi \|_{L^2(\mathcal{T}^n)} \quad \forall\, \varphi \in \mathcal{V}^n. \tag{3.15}$$

3.3. Residuals

We define the *interior residual*

$$R_\theta^n := \frac{U^n - I^n U^{n-1}}{\tau_n} - I^n f^n \in \mathcal{V}^n,$$

along with the *jump residual* across $e \in \mathcal{B}^n$

$$[\![\nabla\Theta^n]\!] := (\nabla\Theta^n_{|S_1} - \nabla\Theta^n_{|S_2}) \cdot \nu_e,$$

using the convention that the unit normal vector ν_e to e points from S_2 to S_1. We observe that integration by parts yields

$$\langle \nabla\Theta^n, \nabla v\rangle = -\sum_{e\in\mathcal{B}^n} \int_e [\![\nabla\Theta^n]\!] v \quad \forall\, v \in H_0^1(\Omega). \tag{3.16}$$

Moreover, if $\mathcal{A}$ is a union of elements in $\mathcal{M}^n$ and $\mathcal{B}^n(\mathcal{A})$ is the set of edges interior to $\mathcal{A}$, and $k \in \mathbb{R}$, we set

$$\| h_n^k\varphi \|_{L^2(\mathcal{A})} := \Big(\sum_{S\subset\mathcal{A}} h_S^{2k} \| \varphi \|^2_{L^2(S)} \Big)^{1/2} \quad \forall\, \varphi \in L^2(\Omega), \tag{3.17}$$

$$|\!|\!| h_n^k [\![\nabla\Theta^n]\!] |\!|\!|_{L^2(\mathcal{A})} := \Big(\sum_{e\in\mathcal{B}^n(\mathcal{A})} h_e^{2k} \| [\![\nabla\Theta^n]\!] \|^2_{L^2(e)} \Big)^{1/2}. \tag{3.18}$$

4. A posteriori error estimates

In this section we state the *a posteriori* error estimates for the schemes of Sections 3.1 and 3.2. Proofs are given in Sections 5 and 6.

4.1. Error representation formula

We start with a representation of the error for any approximation of $\{\theta, \chi, u\}$

$$\{\boldsymbol{\Theta}^n, \mathbf{X}^n, \mathbf{U}^n\} \in \mathcal{V} \times \mathcal{K} \times \mathcal{W}$$

at time $t = t^n$. The ensuing formula will play a fundamental role later.

We denote by $\ell(t)$ the piecewise linear function

$$\ell(t) := \frac{t^n - t}{\tau_n} \qquad \forall\, t_{n-1} < t \le t_n.$$

Given a sequence $\{W^n\}_{n=1}^N \in \mathcal{W}$ we indicate with W and $\bar{W}$ its piecewise linear and constant interpolants, namely, $W(\cdot,0) = \bar{W}(\cdot,0) := W^0(\cdot)$ and for all $t_{n-1} < t \le t_n$

$$W(\cdot,t) := \ell(t)\mathrm{W}^{n-1}(\cdot) + (1-\ell(t))\mathrm{W}^n(\cdot), \qquad \bar{W}(\cdot,t) = \mathrm{W}^n(\cdot). \tag{4.1}$$

Let $G : \mathcal{V}^* \to \mathcal{V}$ be the Green operator defined by

$$\langle \nabla G\phi, \nabla v \rangle = \langle \phi, v \rangle \quad \forall\, v \in \mathcal{V}, \phi \in \mathcal{V}^*. \tag{4.2}$$

We first rewrite (2.1) as

$$\langle \partial_t(u - \mathbf{U}), \phi \rangle + \langle \nabla(\theta - \bar{\mathbf{\Theta}}), \nabla\phi \rangle = \langle f - \partial_t \mathbf{U}, \phi \rangle - \langle \nabla \bar{\mathbf{\Theta}}, \nabla\phi \rangle \quad \forall\, \phi \in \mathcal{V}.$$

Upon taking $\phi = G(u - \mathbf{U}) \in \mathcal{V}$, we get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\| u - \mathbf{U} \|^2_{H^{-1}(\Omega)} + \langle \theta - \bar{\mathbf{\Theta}}, u - \mathbf{U} \rangle = \langle f - \partial_t \mathbf{U}, G(u - \mathbf{U}) \rangle - \langle \nabla\bar{\mathbf{\Theta}}, \nabla G(u - \mathbf{U}) \rangle. \tag{4.3}$$

Next we rewrite (2.2) as

$$\langle \varepsilon\partial_t(\chi - \mathbf{X}), \chi - \mathbf{X} \rangle - \langle \theta - \bar{\mathbf{\Theta}}, \chi - \mathbf{X} \rangle \le \langle \bar{\mathbf{\Theta}} - \varepsilon\partial_t \mathbf{X}, \chi - \mathbf{X} \rangle. \tag{4.4}$$

We introduce the error function

$$\delta(\theta; \Theta) := \frac{1}{2}\| \theta - \Theta \|^2_{L^2(\Omega)} + \frac{1}{2}\| \theta - \bar{\Theta} \|^2_{L^2(\Omega)}, \tag{4.5}$$

and note that a simple algebraic calculation yields

$$\langle \theta - \bar{\mathbf{\Theta}}, \theta - \mathbf{\Theta} \rangle = \delta(\theta; \mathbf{\Theta}) - \frac{1}{2}\| \bar{\mathbf{\Theta}} - \mathbf{\Theta} \|^2_{L^2(\Omega)}.$$

Upon adding (4.3) and (4.4), and integrating in time from 0 to $t^* \le T$, we easily obtain the desired error representation formula

$$\begin{aligned}
&\frac{1}{2}\| (u - \mathbf{U})(t^*) \|^2_{H^{-1}(\Omega)} + \int_0^{t^*} \delta(\theta; \mathbf{\Theta})\mathrm{d}t + \frac{\varepsilon}{2}\| (\chi - \mathbf{X})(t^*) \|^2_{L^2(\Omega)} \\
&\quad \le \frac{1}{2}\| u_0 - \mathbf{U}^0 \|^2_{H^{-1}(\Omega)} + \frac{\varepsilon}{2}\| \chi_0 - \mathbf{X}^0 \|^2_{L^2(\Omega)} \\
&\quad + \frac{1}{2}\int_0^{t^*} \| \mathbf{\Theta} - \bar{\mathbf{\Theta}} \|^2_{L^2(\Omega)}\mathrm{d}t + \int_0^{t^*} \langle \theta - \bar{\mathbf{\Theta}}, \mathbf{U} - (\mathbf{X} + \mathbf{\Theta}) \rangle \mathrm{d}t \\
&\quad + \int_0^{t^*} \Big(\langle f - \partial_t \mathbf{U}, G(u - \mathbf{U}) \rangle - \langle \nabla\bar{\mathbf{\Theta}}, \nabla G(u - \mathbf{U}) \rangle \Big)\mathrm{d}t \\
&\quad + \int_0^{t^*} \langle \bar{\mathbf{\Theta}} - \varepsilon\partial_t \mathbf{X}, \chi - \mathbf{X} \rangle \mathrm{d}t.
\end{aligned} \tag{4.6}$$

We conclude that evaluating the various terms in (4.6) in suitable Sobolev norms would yield a computable error estimate solely in terms of discrete quantities and data. The fourth term on the right-hand side of (4.6) *vanishes provided* $\mathbf{U} = \mathbf{\Theta} + \mathbf{X}$, *which is not assumed in this derivation.*

4.2. Time discretization

We first state *a posteriori* error estimates for the scheme of Section 3.1, and then their *optimal* rate of convergence. Proofs are given in Section 5.

If the Hilbert space $\mathcal{V}^* \times \mathcal{W}$ is endowed with the scalar product

$$\langle\!\langle (u, \chi), (v, \eta) \rangle\!\rangle = \langle u, v \rangle_{\mathcal{V}^*} + \varepsilon\langle \chi, \eta \rangle_{\mathcal{W}},$$

then (1.1)-(1.2) can be written equivalently as $\partial_t(u,\chi)+\mathfrak{F}(u,\chi)=(f,0)$ in terms of the (multi-valued) monotone operator $\mathfrak{F}:\mathcal{V}^*\times\mathcal{W}\to\mathcal{V}^*\times\mathcal{W}$ given by

$$\mathfrak{F}(u,\chi)=\Big(-\Delta(u-\chi),\frac{1}{\varepsilon}(\Lambda(\chi)-(u-\chi)\Big); \tag{4.7}$$

$\mathfrak{F}$ is the sub-gradient of the *convex* lower semi-continuous potential [9,18,19]

$$\Psi(u,\chi)=\int_\Omega\frac{1}{2}|u-\chi|^2+\Phi(\chi). \tag{4.8}$$

Therefore the abstract theory of *a posteriori* error analysis of [13,14] applies directly to the implicit method (3.4), thereby giving an upper bound in terms of the computable *monotonicity residual* (use that $U^n=\Theta^n+\mathrm{X}^n$)

$$\begin{aligned}\langle\!\langle\mathfrak{F}(U^n,\mathrm{X}^n)-\mathfrak{F}(U^{n-1},\mathrm{X}^{n-1}),(U^n,\mathrm{X}^n)-(U^{n-1},\mathrm{X}^{n-1})\rangle\!\rangle &= \|\Theta^n-\Theta^{n-1}\|^2_{L^2(\Omega)}\\ &\quad+\langle\Lambda(\mathrm{X}^n)-\Lambda(\mathrm{X}^{n-1}),\mathrm{X}^n-\mathrm{X}^{n-1}\rangle. \end{aligned}\tag{4.9}$$

Theorem 4.1. *The following a posteriori error estimates are valid*

$$\begin{aligned}&\|(u-U)(T)\|^2_{H^{-1}(\Omega)}+\varepsilon\|(\chi-X)(T)\|^2_{L^2(\Omega)}+4\int_0^T\delta(\theta;\Theta)\mathrm{d}t\\ &\le 2\sum_{n=1}^N\tau_n\Big(\|\Theta^n-\Theta^{n-1}\|^2_{L^2(\Omega)}+\langle \mathrm{Z}^n-\mathrm{Z}^{n-1},\mathrm{X}^n-\mathrm{X}^{n-1}\rangle\Big)\\ &\quad+\frac{4}{\varepsilon}\Big(\sum_{n=1}^N\tau_n\|\Theta^n-\Theta^{n-1}\|_{L^2(\Omega)}\Big)^2+4\Big(\sum_{n=1}^N\int_{t^{n-1}}^{t^n}\|f-f^n\|_{H^{-1}(\Omega)}\mathrm{d}t\Big)^2\end{aligned}\tag{4.10}$$

for the semi-explicit method (3.1)-(3.2). The third term on the right-hand side can be omitted for the method (3.4).

We now stress that the first two terms in (4.10) correspond to (4.9). In particular, the second term captures the essence of the variational inequality for phase variable. Since $\mathrm{Z}^n=0$ within $\mathcal{T}^n$, and $\mathcal{T}^n$ does not change too abruptly, such a term vanishes in most of the domain Ω. The third term on the right-hand side of (4.10) is due to the explicit treatment of temperature in (3.1).

The quantity $\delta(\theta,\Theta)$ arises from the coercivity of the first component of $\mathfrak{F}$ in $\mathcal{V}^*$, whereas the second component of $\mathfrak{F}$ is not coercive and thus contributes nothing to the error. This is consistent with [13,14].

Our next objective is to show that the *a posteriori* error estimators in Theorem 4.1 are optimal with respect to both order and regularity requirements. It is obvious that the last term at the right hand of (4.10) is of optimal order

$$\sum_{n=1}^N\int_{t^{n-1}}^{t^n}\|f-f^n\|_{H^{-1}(\Omega)}\mathrm{d}t\le\sum_{n=1}^N\tau_n\int_{t^{n-1}}^{t^n}\|\partial_t f(t)\|_{H^{-1}(\Omega)}\mathrm{d}t\le C\tau \tag{4.11}$$

provided $f\in H^1(0,T;H^{-1}(\Omega))$ and $\tau:=\max_{1\le n\le N}\tau_n$ is the largest time-step. It remains to estimate the first two terms on the right hand of (4.10), which is possible under weaker assumptions than (H1)-(H4).

Theorem 4.2. *In addition to (H3), let θ_0 and f satisfy*

$$\|\theta_0\|_{H^1_0(\Omega)}\le C^*,\quad \|f\|_{W^{1,1}(0,T;H^{-1}(\Omega))}\le C^*. \tag{4.12}$$

There exists a constant C depending on C^*, T and A in (H3), but independent of τ_n and ε, such that the following a priori estimates are valid provided the stability constraint $\tau_n \le \varepsilon$ is enforced for all $1 \le n \le N$

$$\sum_{n=1}^{N} \frac{1}{\tau_n}\Big(\| \Theta^n - \Theta^{n-1} \|^2_{L^2(\Omega)} + \langle \mathrm{Z}^n - \mathrm{Z}^{n-1}, \mathrm{X}^n - \mathrm{X}^{n-1} \rangle \Big) \le C. \tag{4.13}$$

The compatibility assumption (H3) is equivalent to requiring that (u_0, χ_0) belong to the domain of $\mathfrak{F}$ uniformly in ε. Such a regularity is also used in [13, 14] to prove optimal rates of convergence in an abstract setting.

Corollary 4.1. *If (H3) and (4.12) hold, then there is a constant C depending on C^*, T and A in (H3), but independent of τ_n and ε, such that*

$$\|(u-U)(T)\|_{H^{-1}(\Omega)} + \sqrt{\varepsilon}\|(\chi - X)(T)\|_{L^2(\Omega)} + \Big(\int_0^T \delta(\theta;\Theta)\mathrm{d}t \Big)^{1/2} \le C\frac{\tau}{\sqrt{\varepsilon}} \tag{4.14}$$

provided $\tau_n \le \varepsilon$ for all $1 \le n \le N$. The rate of convergence becomes $O(\tau)$ for (3.4).

Proof. To derive (4.14), it suffices to apply Theorem 4.2 together with the Cauchy-Schwarz inequality to estimate the last term in (4.10). □

The rate of convergence (4.14) is *optimal* according to the semi-discrete traveling wave solution of [12]. Similar rates were derived in [9] but the present argument is simpler and also more intrinsic.

4.3. Space discretization

We first define the error functions $e_u := u - U$, $e_\chi := \chi - PX$, where P is the interpolation operator of (3.12). Our goal is to establish a fully discrete version of Theorem 4.1.

Theorem 4.3. *Let $\tau_n \le C^{\#}\tau_{n-1}$ for all $1 \le n \le N$ with $C^{\#} > 0$ arbitrary. Then there exist constants $C > 0$ depending solely on the minimum angle of $\mathcal{M}^n$ and $C^{\#}$ such that the following a posteriori error estimate holds for all $t^m \in [0,T]$,*

$$\| e_u \|_{L^\infty(0,t^m;H^{-1}(\Omega))} + \Big(\int_0^{t^m} \delta(\theta;\Theta)\mathrm{d}t \Big)^{1/2} + \sqrt{\varepsilon}\| e_\chi \|_{L^\infty(0,t^m;L^2(\Omega))} \le C \sum_{i=0}^{7} \mathcal{E}_i,$$

where the error indicators $\mathcal{E}_i$ are given by

$$\mathcal{E}_0 := \| u_0 - U^0 \|_{H^{-1}(\Omega)} + \varepsilon^{1/2} \| \chi_0 - P^0 X^0 \|_{L^2(\Omega)} \qquad \textit{initial error,}$$

$$\mathcal{E}_1 := \Big(\sum_{n=1}^{m} \tau_n \| \Theta^n - I^n \Theta^{n-1} \|_{L^2(\Omega)}^2 \Big)^{1/2} + \varepsilon^{-1/2} \sum_{n=1}^{m} \tau_n \| \Theta^n - I^n \Theta^{n-1} \|_{L^2(\Omega)}$$
$$+ \Big(\sum_{n=1}^{m} \tau_n \langle P^n Z^n - P^{n-1} Z^{n-1}, P^n X^n - P^{n-1} X^{n-1} \rangle \Big)^{1/2} \qquad \textit{time residual,}$$

$$\mathcal{E}_2 := \Big(\sum_{n=1}^{m} \tau_n \| | h_n^{3/2} [\![\nabla \Theta^n]\!] \| |_{L^2(\Omega)}^2 \Big)^{1/2} + \varepsilon^{-1/2} \sum_{n=1}^{m} \tau_n \| | h_n^{3/2} [\![\nabla \Theta^n]\!] \| |_{L^2(\Omega)} \qquad \textit{jump residual,}$$

$$\mathcal{E}_3 := \Big(\sum_{n=1}^{m} \tau_n \| h_n^2 R_\theta^n \|_{L^2(\Omega)}^2 \Big)^{1/2} + \varepsilon^{-1/2} \sum_{n=1}^{m} \tau_n \| h_n^2 R_\theta^n \|_{L^2(\Omega)} \qquad \textit{interior residual,}$$

$$\mathcal{E}_4 := \sum_{n=1}^{m} \| U^{n-1} - I^n U^{n-1} \|_{H^{-1}(\Omega)} + \varepsilon^{1/2} \sum_{n=1}^{m} \| P^n (I^n X^{n-1}) - P^{n-1} X^{n-1} \|_{L^2(\Omega)}$$
$$+ \Big(\sum_{n=1}^{m} \tau_n \| \Theta^{n-1} - I^n \Theta^{n-1} \|_{L^2(\Omega)}^2 \Big)^{1/2} \qquad \textit{coarsening,}$$

$$\mathcal{E}_5 := \sum_{n=1}^{m} \tau_n \| h_n^2 \nabla R_\theta^n \|_{L^2(\Omega)} \qquad \textit{quadrature,}$$

$$\mathcal{E}_6 := \varepsilon^{-1/2} \sum_{n=1}^{m} \tau_n \| h_n \nabla I^n \Theta^{n-1} \|_{L^2(\mathcal{T}^n)} + \Big(\sum_{n=0}^{m} \tau_n \| h_n \nabla \mathrm{X}^n \|_{L^2(\mathcal{T}^n)}^2 \Big)^{1/2} \qquad \textit{transition layer,}$$

$$\mathcal{E}_7 := \sum_{n=1}^{m} \int_{t^{n-1}}^{t^n} \| f - I^n f^n \|_{H^{-1}(\Omega)} \mathrm{d}t. \qquad \textit{source discretization.}$$

Remark 4.1. The estimators $\mathcal{E}_1$ and $\mathcal{E}_7$ are consistent with those in (4.10). The remaining estimators reflect the effect of space discretization.

5. Proof of Theorem 4.1 and 4.2

5.1. Proof of Theorem 4.1

We use the crucial estimate (4.6). Let $\Theta^n, \mathrm{X}^n, U^n$ be the solutions of (3.1)-(3.2), and set $\boldsymbol{\Theta}^n = \Theta^n, \mathbf{X}^n = \mathrm{X}^n, \mathbf{U}^n = U^n$. Consequently, the fourth term on the right-hand side of (4.6) vanishes. From (3.2) we know that

$$\langle f - \partial_t U, G(u-U) \rangle - \langle \nabla \bar{\Theta}, \nabla G(u-U) \rangle = \langle f - f^n, G(u-U) \rangle.$$

Rewriting (3.3) as $Z^n = \Theta^{n-1} - \varepsilon \partial_t X$ we deduce for $t^{n-1} < t \le t^n$ that

$$\begin{aligned} \langle \bar{\Theta} - \varepsilon \partial_t X, \chi - X \rangle &= \langle \Theta^n - \Theta^{n-1}, \chi - X \rangle + \langle \mathrm{Z}^n, \chi - X \rangle \\ &= \langle \Theta^n - \Theta^{n-1}, \chi - X \rangle + \langle \mathrm{Z}^n, \chi - \mathrm{X}^n \rangle + \langle \mathrm{Z}^n, \mathrm{X}^n - X \rangle \\ &\le \langle \Theta^n - \Theta^{n-1}, \chi - X \rangle + \langle \mathrm{Z}^n, \mathrm{X}^n - X \rangle, \end{aligned}$$

where we have used the fact that $\mathrm{Z}^n \in \Lambda(\mathrm{X}^n)$ and $|\chi| \le 1$ to conclude that

$$\langle \mathrm{Z}^n, \chi - \mathrm{X}^n \rangle \le \Phi(\chi) - \Phi(\mathrm{X}^n) = 0,$$

Φ being the primitive of Λ defined in (1.6). Now inserting the above estimates into (4.6) for $t^* \in (t^{m-1}, t^m]$ with $1 \le m \le N$ and noticing that

$$\Theta^n - \Theta = \ell(t)(\Theta^n - \Theta^{n-1}), \quad \mathrm{X}^n - X = \ell(t)(\mathrm{X}^n - \mathrm{X}^{n-1}),$$

as well as that

$$\langle \mathrm{Z}^{n-1}, \mathrm{X}^n - \mathrm{X}^{n-1}\rangle \le \Phi(\mathrm{X}^n) - \Phi(\mathrm{X}^{n-1}) = 0,$$

we finally obtain with $t_*^n = \min(t^n, t^*)$

$$\begin{aligned}
&\frac{1}{2}\| (u-U)(t^*) \|_{H^{-1}(\Omega)}^2 + \frac{1}{2}\varepsilon \| (\chi - X)(t^*) \|_{L^2(\Omega)}^2 + \int_0^{t^*} \delta(\theta;\Theta)\mathrm{d}t \\
&\quad \le \frac{1}{2}\sum_{n=1}^m \tau_n \Big(\| \Theta^n - \Theta^{n-1} \|_{L^2(\Omega)}^2 + \langle \mathrm{Z}^n - \mathrm{Z}^{n-1}, \mathrm{X}^n - \mathrm{X}^{n-1}\rangle \Big) \\
&\quad + \frac{1}{4}\max_{0\le t\le t^*} \Big(\| u - U \|_{H^{-1}(\Omega)}^2 + \varepsilon \| \chi - X \|_{L^2(\Omega)}^2 \Big) \\
&\quad + \varepsilon^{-1}\Big(\sum_{n=1}^m \tau_n \| \Theta^n - \Theta^{n-1} \|_{L^2(\Omega)} \Big)^2 + \Big(\sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \| f - f^n \|_{H^{-1}(\Omega)} \mathrm{d}t \Big)^2 .
\end{aligned} \tag{5.1}$$

Let $t^* \in [0,T]$ be chosen so that it realizes the maximum of $\| u - U \|_{H^{-1}(\Omega)}^2 + \varepsilon \| \chi - X \|_{L^2(\Omega)}^2$. We then hide the second term on the right-hand side of (5.1) into its left. This immediately gives the asserted error estimate for u and χ. The remaining estimate for θ results from replacing on the right-hand side of (5.1) the one just obtained. This argument completes the proof of Theorem 4.1 for the semi-explicit method (3.1), as well as reveals that the third term on the right-hand side does not occur for the implicit method (3.4). □

5.2. Proof of Theorem 4.2

We first introduce the following notation of discrete derivatives

$$\delta U^n := \frac{U^n - U^{n-1}}{\tau_n}, \quad \delta \mathrm{X}^n := \frac{\mathrm{X}^n - \mathrm{X}^{n-1}}{\tau_n}.$$

We next rewrite (3.2) and (3.3) for $n-1$ in place of n with $n \ge 2$

$$\langle \delta U^{n-1}, \phi\rangle + \langle \nabla\Theta^{n-1}, \nabla\phi\rangle = \langle f^{n-1}, \phi\rangle \quad \forall\, \phi \in \mathcal{V}, \tag{5.2}$$

$$\varepsilon\langle \delta \mathrm{X}^{n-1}, \varphi\rangle + \langle \mathrm{Z}^{n-1}, \varphi\rangle = \langle \Theta^{n-2}, \varphi\rangle \quad \forall\, \varphi \in \mathcal{W}. \tag{5.3}$$

To extend these equations to $n = 1$ we need to define suitable quantities U^{-1}, Θ^{-1} and X^{-1}. To this end, we set $\Theta^{-1} := \Theta^0$ and $Z^0 := z_0$, and thereby define X^{-1} according to the equation

$$\varepsilon\frac{X^0 - X^{-1}}{\tau_0} + Z^0 = \Theta^{-1}. \tag{5.4}$$

We then define U^{-1} to satisfy

$$\delta U^0 = \frac{U^0 - U^{-1}}{\tau_0} = \Delta\Theta^0 + f^0. \tag{5.5}$$

We observe that such definitions do *not* enforce the constitutive relation $U^{-1} = \Theta^{-1} + X^{-1}$, but instead the satisfaction of (5.2)-(5.3) for $n = 1$ along with the following estimates based on (H3) and (4.12)

$$\varepsilon^{1/2}\| \delta X^0 \|_{L^2(\Omega)} = \varepsilon^{-1/2}\| \theta_0 - z_0 \|_{L^2(\Omega)} \le C, \tag{5.6}$$

$$\| \delta U^0 \|_{H^{-1}(\Omega)} = \| \Delta\Theta^0 \|_{H^{-1}(\Omega)} + \| f^0 \|_{H^{-1}(\Omega)} \le C. \tag{5.7}$$

We subtract (5.2) from (3.2), and use $\phi = G(\delta U^n) \in \mathcal{V}$ to obtain

$$\begin{aligned}\frac{1}{2}\| \delta U^n \|^2_{H^{-1}(\Omega)} - \frac{1}{2}\| \delta U^{n-1} \|^2_{H^{-1}(\Omega)} + \frac{1}{\tau_n}\| \Theta^n - \Theta^{n-1} \|^2_{L^2(\Omega)} + \frac{1}{\tau_n}\langle \Theta^n - \Theta^{n-1}, \mathrm{X}^n - \mathrm{X}^{n-1}\rangle \\ = \langle f^n - f^{n-1}, G(\delta U^n)\rangle.\end{aligned} \tag{5.8}$$

Similarly, we subtract (5.3) from (3.3) and take $\varphi = \delta \mathrm{X}^n \in \mathcal{W}$ to get

$$\begin{aligned}\frac{\varepsilon}{2}\| \delta \mathrm{X}^n \|^2_{L^2(\Omega)} - \frac{\varepsilon}{2}\| \delta \mathrm{X}^{n-1} \|^2_{L^2(\Omega)} + \frac{\varepsilon}{2}\| \delta \mathrm{X}^n - \delta \mathrm{X}^{n-1} \|^2_{L^2(\Omega)} + \frac{1}{\tau_n}\langle \mathrm{Z}^n - \mathrm{Z}^{n-1}, \mathrm{X}^n - \mathrm{X}^{n-1}\rangle \\ = \frac{1}{\tau_n}\langle \Theta^n - \Theta^{n-1}, \mathrm{X}^n - \mathrm{X}^{n-1}\rangle - \langle (\Theta^n - \Theta^{n-1}) - (\Theta^{n-1} - \Theta^{n-2}), \delta \mathrm{X}^n\rangle.\end{aligned} \tag{5.9}$$

By adding (5.8) and (5.9), and summing the resulting equality over n for $1 \le n \le m \le N$, we arrive at

$$\begin{aligned}&\frac{1}{2}\| \delta U^m \|^2_{H^{-1}(\Omega)} + \frac{\varepsilon}{2}\| \delta X^m \|^2_{L^2(\Omega)} + \frac{\varepsilon}{2}\sum_{n=1}^{m}\| \delta \mathrm{X}^n - \delta \mathrm{X}^{n-1} \|^2_{L^2(\Omega)} + \sum_{n=1}^{m}\frac{1}{\tau_n}\| \Theta^n - \Theta^{n-1} \|^2_{L^2(\Omega)} \\ &\quad + \sum_{n=1}^{m}\frac{1}{\tau_n}\langle \mathrm{Z}^n - \mathrm{Z}^{n-1}, \mathrm{X}^n - \mathrm{X}^{n-1}\rangle \\ &= \frac{1}{2}\| \delta U^0 \|^2_{H^{-1}(\Omega)} + \frac{\varepsilon}{2}\| \delta X^0 \|^2_{L^2(\Omega)} + \sum_{n=1}^{m}\langle f^n - f^{n-1}, G(\delta U^n)\rangle \\ &\quad - \sum_{n=1}^{m}\langle (\Theta^n - \Theta^{n-1}) - (\Theta^{n-1} - \Theta^{n-2}), \delta \mathrm{X}^n\rangle.\end{aligned} \tag{5.10}$$

In light of (5.6) and (5.7) we see that the first term of the right-hand side of (5.10) is bounded uniformly in ε. The second term can be handled *via* (4.11) as follows:

$$\sum_{n=1}^{m}\langle f^n - f^{n-1}, G(\delta U^n)\rangle \le C + \frac{1}{2}\max_{1\le n\le m}\| \delta U^n \|^2_{L^2(\Omega)}. \tag{5.11}$$

To bound the third term in (5.10) we do sum by parts, use the stability constraint $\tau_n \le \varepsilon$ together with $\Theta^{-1} = \Theta^0$, to get

$$\begin{aligned}\sum_{n=1}^{m}\langle (\Theta^n - \Theta^{n-1}) - (\Theta^{n-1} - \Theta^{n-2}), \delta \mathrm{X}^n\rangle &= -\langle \Theta^m - \Theta^{m-1}, \delta X^m\rangle + \sum_{n=1}^{m}\langle \Theta^{n-1} - \Theta^{n-2}, (\delta \mathrm{X}^n - \delta \mathrm{X}^{n-1})\rangle \\ &\le \frac{\varepsilon}{2}\| \delta X^m \|^2_{L^2(\Omega)} + \frac{1}{2}\sum_{n=1}^{m}\tau_n^{-1}\| \Theta^n - \Theta^{n-1} \|^2_{L^2(\Omega)} \\ &\quad + \frac{\varepsilon}{2}\sum_{n=1}^{m}\| \delta \mathrm{X}^n - \delta \mathrm{X}^{n-1} \|^2_{L^2(\Omega)}.\end{aligned} \tag{5.12}$$

Substituting (5.11)-(5.12) into (5.10) yields the asserted estimate (4.13), and completes the proof. □

Remark 5.1. Note that the stability constraint $\tau_n \le \varepsilon$ is only employed in (5.12), which is not present for the implicit method (3.4).

6. Proof of Theorem 4.3

Since we need to interpolate functions under minimal regularity assumptions, we resort to the Clément interpolation operator $\Pi^n : H_0^1(\Omega) \to \mathcal{V}_0^n$, which satisfies the following local approximation properties for all $\varphi \in H^k(N(A))$ and $k = 1, 2$

$$\|\varphi - \Pi^n\varphi\|_{L^2(S)} + h_S\|\nabla(\varphi - \Pi^n\varphi)\|_{L^2(S)} \le C^* h_S^k \|\phi\|_{H^k(N(S))}, \tag{6.1}$$

$$\|\varphi - \Pi^n\varphi\|_{L^2(e)} \le C^* h_e^{k-1/2}\|\phi\|_{H^k(N(e))}, \tag{6.2}$$

where $N(A)$ is the union of all elements of $\mathcal{M}^n$ surrounding the sets $A = S \in \mathcal{M}^n$ or $A = e \in \mathcal{B}^n$ [4]. The constant C^* depends solely on the minimum angle of the mesh $\mathcal{M}^n$, which is also responsible for a universal bound on the number of elements belonging to $N(S)$. This, in conjunction with (6.1) for $k = 1$, yields

$$\|\nabla\Pi^n\varphi\|_{L^2(\Omega)} \le C\|\nabla\varphi\|_{L^2(\Omega)}. \tag{6.3}$$

In contrast to Theorem 4.1, we now invoke the representation formula (4.6) with $\mathbf{\Theta}^n = \Theta^n$, $\mathbf{X}^n = P^nX^n$ and $\mathbf{U}^n = \Theta^n + \mathrm{X}^n$, which leads to

$$\mathbf{U}^n - (\mathbf{\Theta}^n + \mathbf{X}^n) = \mathrm{X}^n - P^n\mathrm{X}^n \ne 0.$$

We next estimate the last two terms on the right-hand side of (4.6). Using (3.7), we have for all $t \in (t^{n-1}, t^n]$

$$\begin{aligned}\langle f - \partial_t U, Ge_u\rangle - \langle\nabla\bar\Theta, \nabla Ge_u\rangle = {}& \tau_n^{-1}\langle U^{n-1} - I^nU^{n-1}, Ge_u\rangle + \langle R_\theta^n, \Pi^nGe_u - Ge_u\rangle\\ &+ \langle\nabla\Theta^n, \nabla(\Pi^nGe_u - Ge_u)\rangle + \langle R_\theta^n, \Pi^nGe_u\rangle^n - \langle R_\theta^n, \Pi^nGe_u\rangle\\ &+ \langle f - I^nf^n, Ge_u\rangle.\end{aligned} \tag{6.4}$$

Since $P^nZ^n \in \Lambda(P^nX^n)$, we see that

$$\begin{aligned}\langle P^nZ^n, \chi - PX\rangle &= \langle P^nZ^n, \chi - P^nX^n\rangle + \langle P^nZ^n, P^nX^n - PX\rangle\\ &\le \Phi(\chi) - \Phi(P^nZ^n) + \langle P^nZ^n, P^nX^n - PX\rangle\\ &= \ell(t)\langle P^nZ^n, P^nX^n - P^{n-1}\mathrm{X}^{n-1}\rangle.\end{aligned}$$

Therefore, we easily obtain from (4.6) that for all $t \in (t^{n-1}, t^n]$

$$\begin{aligned}\langle\bar\Theta - \varepsilon\partial_tPX, e_\chi\rangle \le {}& \ell(t)\langle P^nZ^n, P^nX^n - P^{n-1}X^{n-1}\rangle + \langle\Theta^n - P^n(I^n\Theta^{n-1}), e_\chi\rangle\\ &+ \frac{\varepsilon}{\tau_n}\langle P^{n-1}X^{n-1} - P^n(I^nX^{n-1}),\ e_\chi\rangle.\end{aligned} \tag{6.5}$$

Inserting these estimates into (4.6), we end up with the following concrete error representation formula, valid for any $t^* \in (t^{m-1}, t^m]$ and $1 \le m \le N$,

$$
\begin{aligned}
\frac{1}{2}\| e_u(t^*) \|_{H^{-1}(\Omega)}^2 + \int_0^{t^*} \delta(\theta;\Theta)\mathrm{d}t + \frac{\varepsilon}{2}\| e_\chi(t^*) \|_{L^2(\Omega)}^2 \le{}& \frac{1}{2}\| e_u(0) \|_{H^{-1}(\Omega)}^2 + \frac{\varepsilon}{2}\| e_\chi(0) \|_{L^2(\Omega)}^2 \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \frac{1}{2}\ell(t)^2 \| \Theta^n - \Theta^{n-1} \|_{L^2(\Omega)}^2 \mathrm{d}t \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} (\theta - \Theta^n, X - PX)\mathrm{d}t \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \tau_n^{-1}\langle U^{n-1} - I^n U^{n-1}, Ge_u\rangle \mathrm{d}t \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \langle R_\theta^n, \Pi^n Ge_u - Ge_u\rangle \mathrm{d}t \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \langle \nabla\Theta^n, \nabla(\Pi^n Ge_u - Ge_u)\rangle \mathrm{d}t \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \Big(\langle R_\theta^n, \Pi^n Ge_u\rangle^n - \langle R_\theta^n, \Pi^n Ge_u\rangle\Big)\mathrm{d}t \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \langle f - I^n f^n, Ge_u\rangle \mathrm{d}t \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \ell(t)\langle P^n Z^n, P^n X^n - P^{n-1}X^{n-1}\rangle \mathrm{d}t \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \langle \Theta^n - P^n(I^n\Theta^{n-1}), e_\chi\rangle \mathrm{d}t \\
&+ \sum_{n=1}^m \int_{t^{n-1}}^{t_*^n} \frac{\varepsilon}{\tau_n}\langle P^{n-1}X^{n-1} - P^n(I^n X^{n-1}), e_\chi\rangle \mathrm{d}t \\
=:{}& \mathrm{I}_0 + \cdots + \mathrm{I}_{10},
\end{aligned}
\tag{6.6}
$$

where $t_*^n = \min(t^n, t^*)$. We now estimate each term I_1 to I_{10} separately.

First, the triangle inequality yields

$$
\begin{aligned}
\mathrm{I}_1 &\le \frac{1}{3}\sum_{n=1}^m \tau_n \| \Theta^n - I^n\Theta^{n-1} \|_{L^2(\Omega)}^2 + \frac{1}{3}\sum_{n=1}^m \tau_n \| \Theta^{n-1} - I^n\Theta^{n-1} \|_{L^2(\Omega)}^2 \\
&\le C(\mathcal{E}_1^2 + \mathcal{E}_4^2).
\end{aligned}
$$

For I_2 we need an estimate on $\| X - PX \|_{L^2(\Omega)}$. In light of (4.1) we have

$$
X - PX = \ell(t)(\mathrm{X}^{n-1} - P^{n-1}X^{n-1}) + (1 - \ell(t))(\mathrm{X}^n - P^n X^n),
$$

for all $t^{n-1} < t \le t^n$. Since $\tau_n \le C^\# \tau_{n-1}$, we immediately get

$$\int_{t^{n-1}}^{t_*^n} \| X - PX \|_{L^2(\Omega)} \mathrm{d}t \le \frac{\tau_n}{2} \| \mathrm{X}^n - P^n X^n \|_{L^2(\Omega)} + \frac{C^\# \tau_{n-1}}{2} \| \mathrm{X}^{n-1} - P^{n-1} X^{n-1} \|_{L^2(\Omega)}, \tag{6.7}$$

whence, applying (3.15), we obtain

$$\mathrm{I}_2 \le \frac{1}{8} \int_0^{t^*} \delta(\theta, \Theta) \mathrm{d}t + C \sum_{n=0}^{m} \tau_n \| X^n - P^n X^n \|_{L^2(\Omega)}^2 \le \frac{1}{8} \int_0^{t^*} \delta(\theta, \Theta) \mathrm{d}t + C \mathcal{E}_6^2.$$

Since $\| G e_u \|_{H^1(\Omega)} \le \| e_u \|_{H^{-1}(\Omega)}$, according to the definition (4.2) of the Green operator $G : \mathcal{V}^* \to \mathcal{V}$, we thus have

$$\mathrm{I}_3 + \mathrm{I}_7 \le \frac{1}{8} \max_{0 \le t \le t^*} \| e_u(t) \|_{H^{-1}(\Omega)}^2 + C(\mathcal{E}_4^2 + \mathcal{E}_7^2).$$

Also, the regularity theory of second order elliptic operators such as G on non-smooth domains yields [8]

$$\begin{aligned} \| G e_u \|_{H^2(\Omega)} &\le C \| e_u \|_{L^2(\Omega)} \\ &\le C(\| e_\theta \|_{L^2(\Omega)} + \| e_\chi \|_{L^2(\Omega)} + \| X - PX \|_{L^2(\Omega)}), \end{aligned}$$

because Ω is convex. Upon using the estimates (3.15), (6.1), and (6.7), we get

$$\begin{aligned} \mathrm{I}_4 &\le C \sum_{n=1}^{m} \int_{t^{n-1}}^{t_*^n} \| h_n^2 R_\theta^n \|_{L^2(\Omega)} \| G e_u \|_{H^2(\Omega)} \mathrm{d}t \\ &\le \frac{1}{8} \int_0^{t^*} \delta(\theta; \Theta) \mathrm{d}t + \frac{\varepsilon}{16} \max_{0 \le t \le t^*} \| e_\chi(t) \|_{L^2(\Omega)}^2 + C(\mathcal{E}_3^2 + \mathcal{E}_6^2). \end{aligned}$$

Similarly, this time using (6.2) instead of (6.1), we have

$$\mathrm{I}_5 \le \frac{1}{8} \int_0^{t^*} \delta(\theta; \Theta) \mathrm{d}t + \frac{1}{16} \varepsilon \max_{0 \le t \le t^*} \| e_\chi(t) \|_{L^2(\Omega)}^2 + C(\mathcal{E}_2^2 + \mathcal{E}_6^2).$$

Applying (3.5) and (6.3) we arrive at

$$\begin{aligned} \mathrm{I}_6 &\le C \sum_{n=1}^{m} \int_{t^{n-1}}^{t_*^n} \| h_n^2 R_\theta^n \|_{L^2(\Omega)} \| \nabla \Pi^n G e_u \|_{L^2(\Omega)} \\ &\le \frac{1}{8} \max_{0 \le t \le t^*} \| e_u(t) \|_{H^{-1}(\Omega)}^2 + C \mathcal{E}_5^2. \end{aligned}$$

Since $P^n \mathrm{Z}^{n-1} \in \Lambda(P^{n-1} X^{n-1})$, invoking the convexity of Φ in (1.6) we realize that

$$\langle P^{n-1} \mathrm{Z}^{n-1}, P^n X^n - P^{n-1} X^{n-1} \rangle \le \Phi(P^n X^n) - \Phi(P^{n-1} X^{n-1}) = 0,$$

whence

$$\mathrm{I}_8 \le \frac{1}{2} \sum_{n=1}^{m} \tau_n \langle P^n \mathrm{Z}^n - P^{n-1} \mathrm{Z}^{n-1}, P^n X^n - P^{n-1} X^{n-1} \rangle \mathrm{d}t \le \frac{1}{2} \mathcal{E}_1^2.$$

Since $\Theta^n - P^n(I^n\Theta^{n-1}) = (\Theta^n - I^n\Theta^{n-1}) + (I^n\Theta^{n-1} - P^n(I^n\Theta^{n-1}))$, and P^n is the identity outside of the transition region $\mathcal{T}^n$, using (3.15) we can write

$$\| \Theta^n - P^n(I^n\Theta^{n-1}) \|_{L^2(\Omega)} \le \| \Theta^n - I^n\Theta^{n-1} \|_{L^2(\Omega)} + \| h_n \nabla I^n \Theta^{n-1} \|_{L^2(\mathcal{T}^n)},$$

which yields

$$\mathrm{I}_9 \le \frac{\varepsilon}{16} \max_{0\le t\le t^*} \| e_\chi(t) \|^2_{L^2(\Omega)} + C(\mathcal{E}_1^2 + \mathcal{E}_6^2).$$

Finally, we easily obtain

$$\mathrm{I}_{10} \le \frac{\varepsilon}{16} \max_{0\le t\le t^*} \| e_\chi(t) \|^2_{L^2(\Omega)} + C\mathcal{E}_4^2,$$

and realize that substituting the above estimates for $(\mathrm{I}_i)_{i=1}^{10}$ into (6.6) leads to the asserted estimate of Theorem 4.1. □

7. Simulations

After a brief discussion of implementation issues, we present two examples which illustrate the performance and efficiency of our adaptive finite element method with error control.

7.1. Implementation

Let $\mathbf{M}$ and $\mathbf{K}$ be the mass and stiffness matrices, namely,

$$\mathbf{M} := \left(\langle \phi_i, \phi_j \rangle_h\right)_{i,j=1}^J, \quad \mathbf{K} := \left(\langle \nabla\phi_i, \nabla\phi_j \rangle\right)_{i,j=1}^J. \tag{7.1}$$

The integrals can be computed *via* the vertex quadrature rule, which gives rise to a diagonal $\mathbf{M}$. Then the linear algebraic system for Θ^n becomes

$$\mathbf{M}\Theta^n + \tau\mathbf{K}\Theta^n = \mathbf{M}(\Theta^{n-1} + \mathrm{X}^{n-1} - \mathrm{X}^n), \tag{7.2}$$

or equivalently

$$\Theta^n + \mathbf{A}\Theta^n = \Theta^{n-1} + \mathrm{X}^{n-1} - \mathrm{X}^n, \tag{7.3}$$

where $\mathbf{A} := \tau(\mathbf{M})^{-1}\mathbf{K}$ is easy to compute. The matrix $\mathbf{Id} + \mathbf{A}$ of the linear system is symmetric and positive definite, thus the system can be solved by standard iterative solvers like SOR, (preconditioned) conjugate gradient or multigrid methods. Even though the phase variable X^n is mostly constant, its update is performed everywhere to detect the spontaneous appearance of a phase change (nucleation), which would not be easy to track otherwise. Nevertheless, the simple algebraic calculation of X^n is much cheaper than the solution (7.3).

We can now design an adaptive method to automatically generate meshes and time-steps for the total error *err* to be below a given tolerance *tol*. The *a posteriori* error estimate of Theorem 4.3 is first rewritten as follows:

$$\begin{aligned}\mathrm{err} :=& \| e_u \|_{L^\infty(0,t^m;H^{-1}(\Omega))} + \Big(\int_0^{t^m} \delta(\theta;\Theta)\mathrm{d}t\Big)^{1/2} + \sqrt{\varepsilon}\| e_\chi \|_{L^\infty(0,t^m;L^2(\Omega))}\\ \le& \Big(\sum_{S\in\mathcal{M}^0} (\eta_S^0)^2\Big)^{1/2} + \max_{n=1,\dots,m} \Big(\eta_\tau^n + \Big(\sum_{S\in\mathcal{M}^n} (\eta_S^n)^2\Big)^{1/2}\Big),\end{aligned}$$

where η_τ^n includes all error indicators of time discretization (from $\mathcal{E}_1, \mathcal{E}_7$) and η_S^n represents all local error indicators of space discretization on element S. The adaptive method adjusts time steps τ_n and adapts meshes $\mathcal{M}^n$ so that

$$\eta_S^0 \le \frac{\gamma_0 \,\mathrm{tol}}{\sqrt{\#\mathcal{M}^0}}, \quad \eta_\tau^n \le \gamma_\tau \,\mathrm{tol}, \quad \eta_S^n \le \frac{\gamma_h \,\mathrm{tol}}{\sqrt{\#\mathcal{M}^n}}, \tag{7.4}$$

where $\gamma_0 + \gamma_\tau + \gamma_h \le 1$ are given parameters. Mesh modification is done by refinement/coarsening *via* bisection. Elements violating (7.4) are refined, whereas elements with small local error indicators relative to the local tolerance may be coarsened. Coarsening is the inverse operation to a previous local refinement. This approach tends to equidistribute errors over space and time, which is standard for parabolic problems; see [5].

Our adaptive method is able to detect the presence, and spontaneous appearance, of transition layers and refine accordingly, and is insensitive to topological changes such as merging, extinction, and mush or singularity formation (see Ex. 7.3). The transition layer velocity need not be computed explicitly for mesh design, which is a major improvement with respect to [11].

7.2. Example: Modified traveling wave

This is an example with exact solution based on the traveling waves of [12]. Even though all functions exhibit radial symmetry, this is a true 2D test because symmetry is not exploited by the code. Let $\delta_{12}^+ := V/2$, $\delta_{12}^- := \sqrt{1/\varepsilon + V^2/4}$. A 1D traveling wave $\theta(x,t) = \hat\theta(x+Vt)$, $\chi(x,t) = \hat\chi(x+Vt)$ is given by

$$\hat\theta(x) := \begin{cases} \frac{\gamma}{V}(\exp(Vx) - 1), & x \le 0 \\ \frac{\gamma}{\delta_{12}^-}\exp(\delta_{12}^+ x)\sinh(\delta_{12}^- x), & 0 < x \le x_\varepsilon \\ \frac{\hat\theta'(x_\varepsilon)}{V}(\exp(V(x-x_\varepsilon)) - 1) + \frac{\gamma}{\delta_{12}^-}\exp(\delta_{12}^+ x_\varepsilon)\sinh(\delta_{12}^- x_\varepsilon), & x > x_\varepsilon, \end{cases}$$

$$\hat\chi(x) := \begin{cases} 0, & x \le 0 \\ \frac{\gamma}{2\delta_{12}^- V\varepsilon}\left(\frac{\exp((\delta_{12}^+ + \delta_{12}^-)x) - 1}{\delta_{12}^+ + \delta_{12}^-} - \frac{\exp((\delta_{12}^+ - \delta_{12}^-)x) - 1}{\delta_{12}^+ - \delta_{12}^-}\right), & 0 < x < x_\varepsilon \\ 1, & x \ge x_\varepsilon, \end{cases}$$

where γ is a parameter, V represents the traveling wave velocity, and x_ε the transition region thickness; they satisfy $\int_0^{x_\varepsilon} \hat\theta(s)\mathrm{d}s = V\varepsilon$ [12]. Note that $\hat\chi$ does not take values between -1 and 1 but rather between 0 and 1 as in [12]. We perform the experiments with $V = 1$ and $\varepsilon = 0.1, 0.01$.

Figures 1 and 2 show the behavior of the true errors $e_\chi = \chi - X$ and $e_\theta = \theta - \Theta$ for tolerances tol $= 4, 2, 1$ and $\varepsilon = 0.1, 0.01$: the errors decrease almost at a linear rate as predicted by the theory; the total error $||e_\chi||_{L^\infty(L^1(\Omega))} + \sqrt{\varepsilon}||e_\chi||_{L^\infty(L^2(\Omega))} + ||e_\theta||_{L^2(L^2(\Omega))}$ is depicted in Figure 4. The almost constant time-steps τ are displayed in Figure 3, which are consistent with a constant layer velocity V. The meshes are however highly refined near the outer boundary of the transition layer $\mathcal{T}$, where ∇X exhibits a jump discontinuity, as indicated by Figure 5. Figure 6 shows isolines of X, again concentrated near the outer boundary of $\mathcal{T}$, together with the rapid variation of X from 0 to 1 without oscillations.

7.3. Example: Oscillating source

This is a phase change in a container $\Omega = (-1,1)^2$, $T = 20.0$, with initial temperature $\Theta(x,0) = 0.1\,x_2$, prescribed temperature at three walls $\Theta(x,t) = 0.1\,x_2$ for $x_2 > -1$, a fourth insulated wall $\partial_\nu\Theta(x,t) = 0$ for $x_2 = -1$, and two circular oscillating heat sources driving the evolution,

$$f(x,t) = \cos(0.2\,t)\max(0.0,\ 3.125 - 50|x - (-0.2,-0.5)|^2) + \sin(0.2\,t)\max(0.0,\ 3.125 - 50|x - (-0.2,0.5)|^2);$$

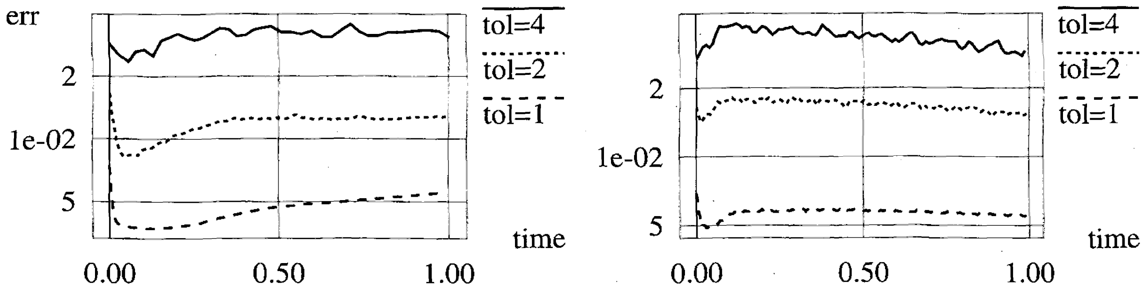

FIGURE 1. Example 5.1. $||e_\chi(t)||_{L^2(\Omega)}$ for $\varepsilon = 0.1,\ 0.01$.

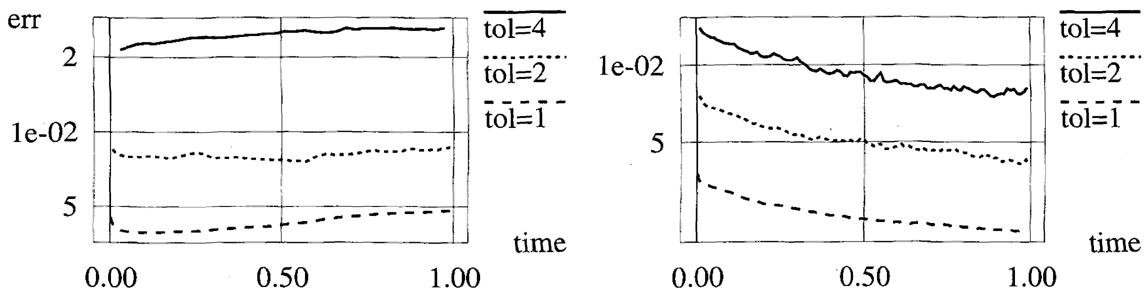

FIGURE 2. Example 5.1. $||e_\theta(t)||_{L^2(\Omega)}$ for $\varepsilon = 0.1,\ 0.01$.

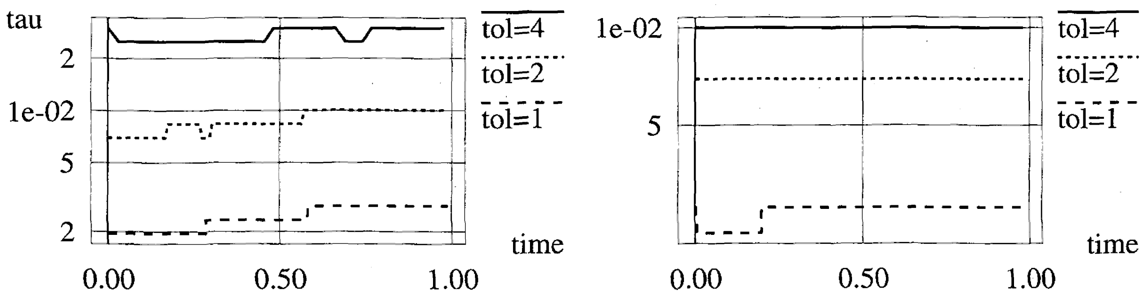

FIGURE 3. Example 5.1. Time step size for $\varepsilon = 0.1,\ 0.01$.

see Figure 12. The exact solution is unknown. Figures 7 and 8 show the effect of decreasing the relaxation parameter from $\varepsilon = 0.1$ to $\varepsilon = 0.01$ in terms of transition layer width. They also demonstrate the ability of the method to handle topological changes such as merging, extinction, and creation of transition layers; all these features take place in the proposed example. The meshes are highly refined to capture both layer location and presence of strong heat sources, and corresponding large variations of second derivatives. Figures 9 and 10 depict isolines of phase variable, confirm the shrinkage of layer width, and hint at the complicated topology of the phase change. Finally Figure 11 displays isothermal curves.

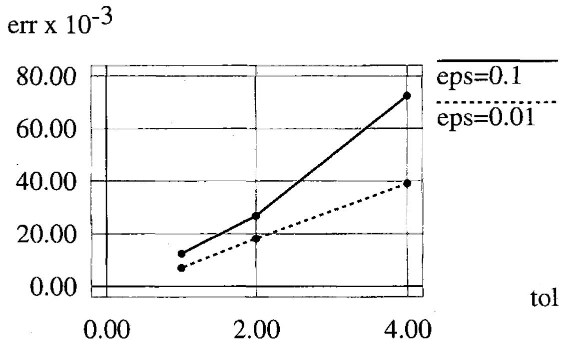

FIGURE 4. Example 5.1. True error for $\varepsilon = 0.1, 0.01$.

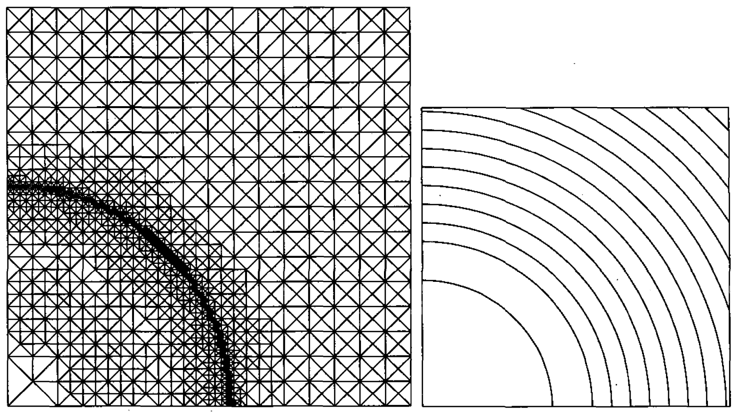

FIGURE 5. Example 5.1. Mesh and isolines $\Theta = 0.5 \cdot k (k \geq 0)$ at $t \approx 0.3$ with $\varepsilon = 0.1, \mathrm{tol} = 2$.

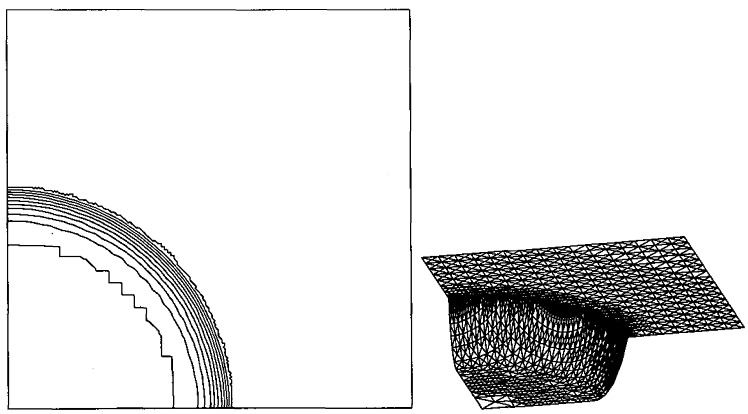

FIGURE 6. Example 5.1. Isolines $X = 0.1.k (0 \leq k \leq 10)$ and graph of X at $t \approx 0.3$ with $\varepsilon = 0.1, \mathrm{tol} = 2$.

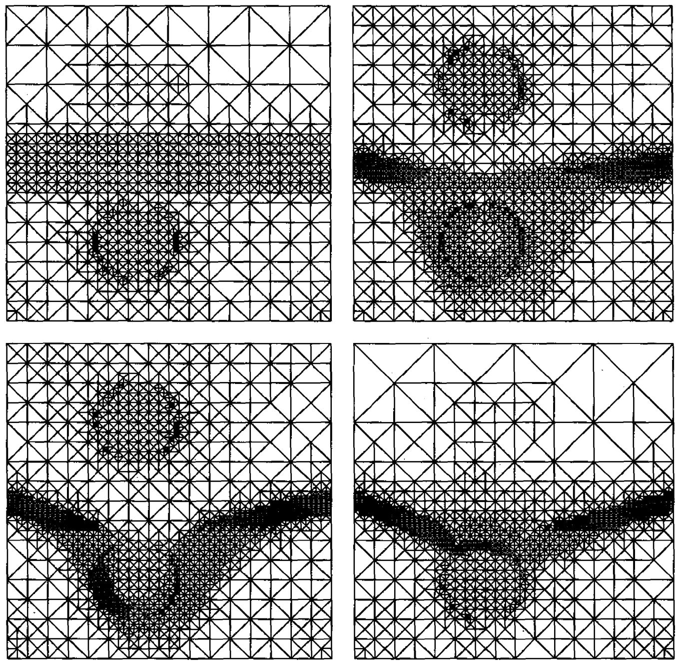

FIGURE 7. Example 5.2. Meshes for $\varepsilon = 0.1, \mathrm{tol} = 3$ at $t \approx 0.5, 5.5, 10.5, 15.5$.

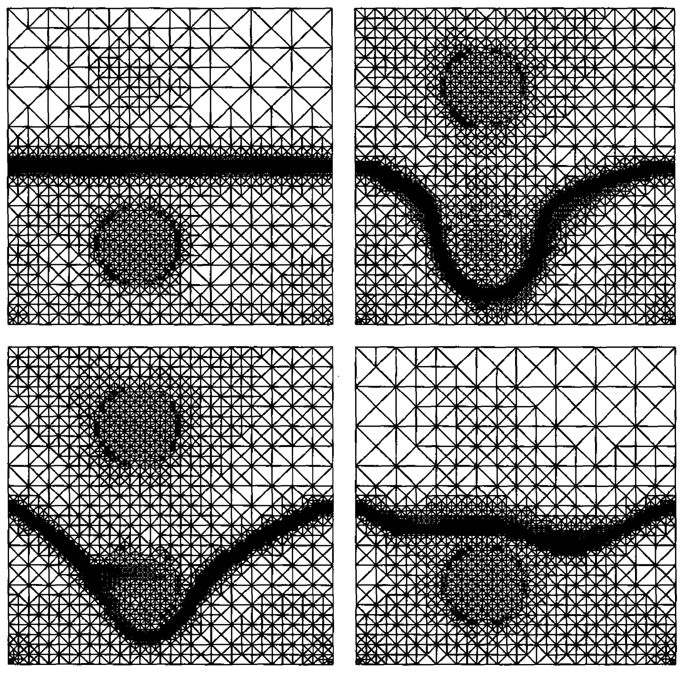

FIGURE 8. Example 5.2. Meshes for $\varepsilon = 0.01, \mathrm{tol} = 3$ at $t \approx 0.5, 5.5, 10.5, 15.5$.

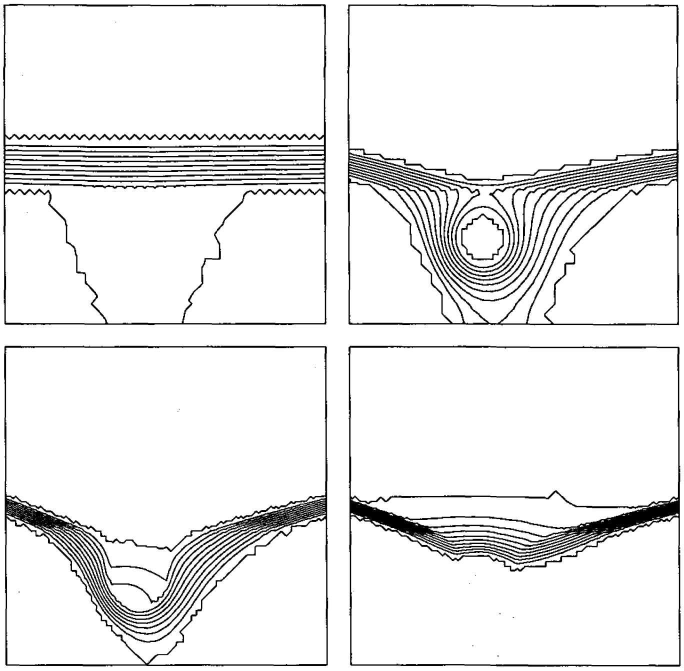

FIGURE 9. Example 5.2. Isolines $X = 0.2 \cdot k$ $(-5 \leq k \leq 5)$ for $\varepsilon = 0.1, \mathrm{tol} = 3$ at $t \approx 0.5, 5.5, 10.5, 15.5$.

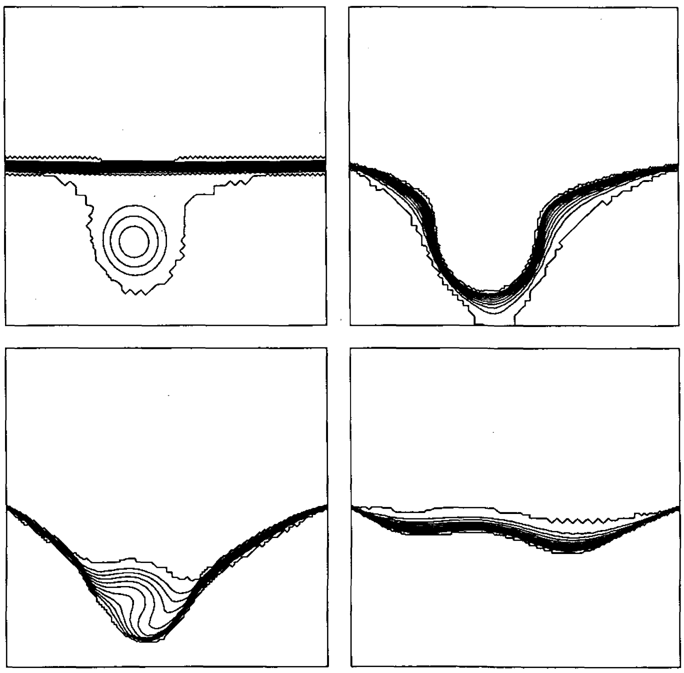

FIGURE 10. Example 5.2. Isolines $X = 0.2 \cdot k$ $(-5 \leq k \leq 5)$ for $\varepsilon = 0.01, \mathrm{tol} = 3$ at $t \approx 0.5, 5.5, 10.5, 15.5$.

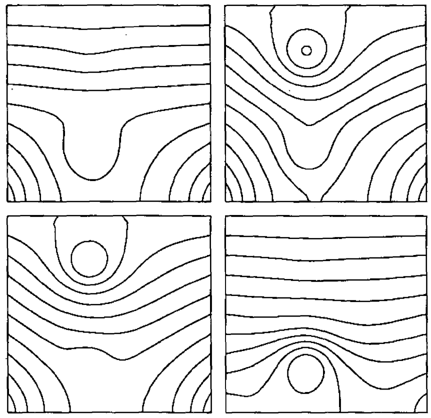

FIGURE 11. Example 5.2. Isolines $\Theta = 0.02 \cdot k(-5 \leq k \leq 5)$ for $\varepsilon = 0.01, \mathrm{tol} = 3$ at $t \approx 0.5, 5.5, 10.5, 15.5$.

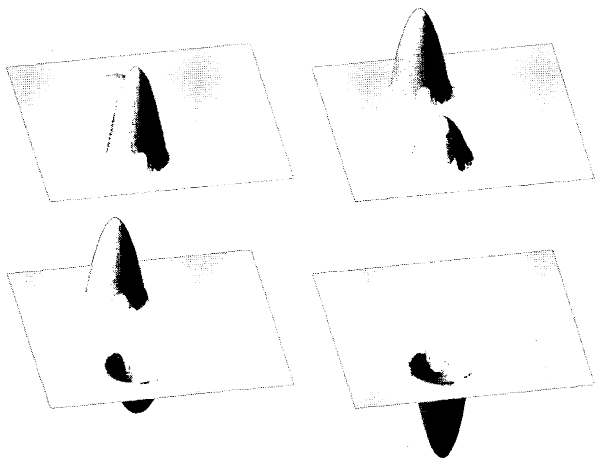

FIGURE 12. Example 5.2. Graphs of right-hand side f for $\varepsilon = 0.01, \mathrm{tol} = 3$ at $t \approx 0.5, 5.5, 10.5, 15.5$.

REFERENCES

[1] Z. Chen and R.H. Nochetto, Residual type *a posteriori* error estimates for elliptic obstacle problems. *Numer. Math.* **84** (2000) 527–548.

[2] Z. Chen, R.H. Nochetto and A. Schmidt, Adaptive finite element methods for diffuse interface models (in preparation).

[3] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems.* North-Holland, Amsterdam (1978).

[4] Ph. Clément, Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.* **9** (1975) 77–84.

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

[6] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. IV. Nonlinear problems. *SIAM J. Numer. Anal.* **32** (1995) 1729–1749.

[7] K. Eriksson, C. Johnson and S. Larsson, Adaptive finite element methods for parabolic problems. VI. Analytic semigroups. *SIAM J. Numer. Anal.* **35** (1998) 1315–1325.

[8] P. Grisvard, *Elliptic Problems on Non-smooth Domains.* Pitman, Boston (1985).

[9] X. Jiang and R.H. Nochetto, Optimal error estimates for semidiscrete phase relaxation models. *RAIRO Modél. Math. Anal. Numér.* **31** (1997) 91–120.

[10] X. Jiang and R.H. Nochetto, A $P^1 - P^1$ finite element method for a phase relaxation model. I. Quasi uniform mesh. *SIAM J. Numer. Anal.* **35** (1998) 1176–1190.

[11] X. Jiang, R.H. Nochetto and C. Verdi, A $P^1 - P^1$ finite element method for a phase relaxation model. II. Adaptively refined meshes. *SIAM J. Numér. Anal.* **36** (1999) 974–999.
[12] R.H. Nochetto, M. Paolini and C. Verdi, Continuous and semidiscrete traveling waves for a phase relaxation model. *European J. Appl. Math.* **5** (1994) 177–199.
[13] R.H. Nochetto, G. Savaré and C. Verdi, Error control for nonlinear evolution equations. *C.R. Acad. Sci. Paris Sér. I* **326** (1998) 1437–1442.
[14] R.H. Nochetto, G. Savaré and C. Verdi, *A posteriori* error estimates for variable time-step discretizations of nonlinear evolution equations. *Comm. Pure Appl. Math.* **53** (2000) 529–589.
[15] R.H. Nochetto, A. Schmidt and C. Verdi, *A posteriori* error estimation and adaptivity for degenerate parabolic problems. *Math. Comp.* **69** (2000) 1–24.
[16] C. Verdi and A. Visintin, Numerical analysis of the multidimensional Stefan problem with supercooling and superheating. *Boll. Un. Mat. Ital. B* **7** (1987) 795–814.
[17] C. Verdi and A. Visintin, Error estimates for a semi-explicit numerical scheme for Stefan-type problems. *Numer. Math.* **52** (1988) 165–185.
[18] A. Visintin, Stefan problem with phase relaxation. *IMA J. Appl. Math.* **34** (1985) 225–245.
[19] A. Visintin, Supercooling and superheating effects in phase transitions. *IMA J. Appl. Math.* **35** (1986) 233–256.