Journal of Mathematics and Statistics Studies
ISSN: 2709-4200
DOI: 10.32996/jmss
Journal Homepage: www.al-kindipublisher.com/index.php/jmss

JMSS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Adaptive Space–Time Numerical Methods for Nonlinear Parabolic Equations with Moving Interfaces

Abbas Allawi Haboub[1]✉ and Ali Hussein Nuaima[2]
[1]*Directorate General of Wasit Education, Al-Hikma Boys' Intermediate School, Wasit, Iraq*
[2]*Directorate General of Wasit Education, Al-Mustafa Boys' Preparatory School, Wasit, Iraq*
Corresponding Author: Abbas Allawi Haboub, **E-mail**: lawyb007@gmail.com

| ABSTRACT
We study a space–time Galerkin method for nonlinear parabolic problems whose coefficients jump across an interface that moves in time. In space–time, the interface is a hypersurface, and the mesh is built to fit it. As a result, no mesh velocity is needed, and the time step does not have to follow the front. Inside a time slab, the discrete functions are continuous in space and time, and slabs are coupled by an upwind jump term. We derive a residual error estimator with four parts (space, time, interface geometry and nonlinear solver) and give proof sketches for its reliability and local efficiency. We also show that, in the exact norm, the thin strip between the true and the discrete interface causes an error of order $\delta^{1/2}$, where δ is the distance between the two. An adaptive algorithm with curvature-aware anisotropic refinement and a Newton stopping rule tied to the estimator is described. A prototype for one spatial dimension and one time slab was used for the tests. The estimator bounded the error in every run, with an effectivity index near 15 for equal diffusion coefficients. The geometry indicator decayed at the predicted rate, and adaptive refinement needed about a third as many unknowns as uniform refinement. The effectivity index grew sharply with the diffusion contrast, so the estimator is not robust in that respect.

| KEYWORDS
space–time finite elements, a posteriori error estimation, moving interfaces, nonlinear parabolic equations, anisotropic mesh adaptation, discontinuous Galerkin in time.

| ARTICLE INFORMATION
ACCEPTED: 22 October 2026 **PUBLISHED:** 16 November 2026 **DOI:** 10.32996/jmss.2026.7.6.17

1. Introduction

Diffusion problems with a moving internal interface occur in phase-change and Stefan-type models, in multiphase flow with a transported scalar, in tumour and biofilm growth, in reaction fronts and in the thermal processing of composite materials. The diffusion coefficients, and often the nonlinearity, differ on the two sides of a front $\Gamma(t)$. The solution is continuous across the front but its gradient jumps, so it is only piecewise smooth. Two difficulties arise together. The first is geometric, since the front moves through the computational domain, sometimes quickly. The second is that the solution has features at several scales in space and time.

Moving-mesh methods in ALE form (Bonito et al., 2013) resolve the interface accurately but need mesh-velocity terms and careful control of mesh quality. The idea of treating moving boundaries by meshing space–time goes back at least to Tezduyar et al. (1992). Unfitted Eulerian methods such as Nitsche-XFEM and CutFEM (Burman et al., 2015; Hansbo & Hansbo, 2002; Lehrenfeld & Olshanskii, 2019) avoid remeshing, but they need ghost-penalty terms, special quadrature on cut elements, and, when the time step is fixed, a limit on how far the interface may travel in one step. Adaptive methods for parabolic problems are well developed on time-stepping meshes (Eriksson & Johnson, 1991; Makridakis & Nochetto, 2003; Verfürth, 2013)، estimators that account for inexact linearisation exist for nonlinear diffusion (Ern & Vohralík, 2013)، and space–time discretisations are an active field (Langer & Steinbach, 2019; Steinbach, 2015). What we could not find is an a posteriori

Copyright: © 2026 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (https://creativecommons.org/licenses/by/4.0/). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

framework in which the interface geometry is an error source with its own indicator, balanced in one adaptive loop against the space, time and solver errors.

The paper combines the following four elements.

C1. The moving interface is treated as a fixed hypersurface $\Sigma \subset \mathbb{R}^{d+1}$, and the mesh is a simplicial mesh of space–time fitted to a polynomial approximation Σ_h. Inside a slab, the discrete space is continuous in space and time, which removes any mesh-velocity term. Slabs are coupled by an upwind jump and may carry different meshes. The interface motion enters only through the geometry of the mesh, so it is separated from the choice of the time step.

C2. We derive a computable estimator $\eta^2 = \eta_S^2 + \eta_T^2 + \eta_J^2 + \eta_\Gamma^2 + \eta_N^2$. Its geometry part η_Γ measures the mismatch between the exact and the discrete interface, both in the coefficients and in the interface source term.

C3. In the exact energy norm, the strip S_h between Σ and Σ_h contributes an error of order $|S_h|^{1/2} \sim \delta^{1/2}$ even if the discrete solution is the best possible. A degree-p_Σ interface therefore limits the order to $(p_\Sigma + 1)/2$ in the worst case, and η_Γ measures exactly this contribution. A sharper rate, $h^{p_\Sigma+1/2}$, holds only in a norm that leaves the strip out (Proposition 4.5).

C4. The refinement metric combines a Hessian-based metric for the solution with an interface metric built from the principal curvatures of Σ in space–time. These curvatures contain both the spatial curvature of $\Gamma(t)$ and the acceleration of the front. The Newton iteration stops when its estimator η_N falls below a fraction of the discretisation estimators.

2. Model problem

Let $\Omega \subset \mathbb{R}^d$ $(d = 1,2,3)$ be a bounded polytopal domain, $T > 0$, and $Q = \Omega \times (0,T)$. Let $\varphi \in C^2(\overline{Q})$ be a level-set function with $|\nabla_x \varphi| \geq c_\varphi > 0$ in a neighbourhood of $\Sigma := \{\varphi = 0\}$. Define

$$\Omega_1(t) = \{\varphi(\cdot,t) < 0\}, \quad \Omega_2(t) = \{\varphi(\cdot,t) > 0\}, \quad \Gamma(t) = \{\varphi(\cdot,t) = 0\},$$

and $Q_i = \cup_t\, \Omega_i(t) \times \{t\}$. We assume $\Gamma(t) \Subset \Omega$ for all t and that $\Gamma(t)$ undergoes no topological change on $[0,T]$ (see Section 7). Let $\nu = (\nu_x, \nu_t) \in \mathbb{R}^{d+1}$ be the unit normal of Σ pointing from Q_1 to Q_2. The spatial unit normal is $n = \nu_x/|\nu_x|$, and the surface measures are related by $\mathrm{d}s\,\mathrm{d}t = |\nu_x|\,\mathrm{d}\Sigma$.

Problem. Find $u: Q \to \mathbb{R}$ such that

$$\partial_t u - \nabla \cdot \sigma_i(x,t,\nabla u) = f(x,t,u) \qquad \text{in } Q_i,\ i = 1,2, \tag{1}$$

$$[[u]] = 0, \qquad [[\sigma(x,t,\nabla u)\cdot n]] = g(x,t,u) \qquad \text{on } \Gamma(t), \tag{2}$$

$$u = 0 \qquad \text{on } \partial\Omega \times (0,T), \tag{3}$$

$$u(\cdot\,,0) = u_0 \qquad \text{in } \Omega, \tag{4}$$

where $[[w]] = w|_{Q_1} - w|_{Q_2}$ on Σ. A typical flux law is $\sigma_i(x,t,\xi) = a_i(x,t,|\xi|)\,\xi$ with a power-law or Carreau-type dependence on $|\xi|$.

Assumption 1 (Structure**)** *There are $0 < \alpha \leq \beta$ and $L_f, L_g \geq 0$ such that for a.e. $(x,t) \in Q_i$, all $\xi, \eta \in \mathbb{R}^d$ and $w, z \in \mathbb{R}$:*

$$\begin{aligned} &(\sigma_i(x,t,\xi) - \sigma_i(x,t,\eta))\cdot(\xi-\eta) \geq \alpha|\xi-\eta|^2, \qquad && |\sigma_i(x,t,\xi) - \sigma_i(x,t,\eta)| \leq \beta|\xi-\eta|, \\ &|f(x,t,w) - f(x,t,z)| \leq L_f|w-z|, && |g(x,t,w) - g(x,t,z)| \leq L_g|w-z|. \end{aligned}$$

Moreover, $\sigma_i(\cdot,\cdot\,,0)$ is bounded, $f(\cdot,\cdot\,,0) \in L^2(Q)$, $g(\cdot,\cdot\,,0) \in L^\infty(Q)$, $u_0 \in L^2(\Omega)$. We write $\kappa_\sigma := \beta/\alpha$ for the ratio between the largest and the smallest slope of the flux. For linear fluxes, it is the diffusion contrast.

The weak form reads: find $u \in L^2(0,T;H_0^1(\Omega))$ with $\partial_t u \in L^2(0,T;H^{-1}(\Omega))$, $u(0) = u_0$, such that for a.e. t and all $v \in H_0^1(\Omega)$

$$\langle \partial_t u, v\rangle + \int_\Omega \sigma(x,t,\nabla u)\cdot\nabla v\,\mathrm{d}x = \int_\Omega f(x,t,u)\,v\,\mathrm{d}x + \int_{\Gamma(t)} g(x,t,u)\,v\,\mathrm{d}s, \tag{5}$$

where $\sigma = \sigma_i$ on $\Omega_i(t)$.

Theorem 2.1 (Well-posedness**)** *Under Assumption 1, problem (5) has a unique solution.*

Proof (sketch). Because φ is C^2, the trace inequality $\| v \|^2_{L^2(\Gamma(t))} \leq \epsilon \| \nabla v \|^2_{L^2(\Omega)} + C_\epsilon \| v \|^2_{L^2(\Omega)}$ holds with constants that do not depend on t. With it, the operator $v \mapsto -\nabla\cdot\sigma(\cdot,t,\nabla v) - f(\cdot,t,v) - g(\cdot,t,v)\,\delta_{\Gamma(t)}$ is hemicontinuous and bounded for almost every t, and it satisfies a Gårding inequality. Measurability in t follows from the regularity of φ. The Lions theory for monotone parabolic operators with a Gårding-type perturbation then gives existence. For uniqueness, one tests the difference of two solutions with itself and applies Gronwall's lemma.

Regularity. For the a priori rates, we assume $u|_{Q_i} \in H^{p+1}(Q_i)$; the a posteriori analysis needs no such assumption. The solution is in general not in $H^2(Q)$, because ∇u jumps across Σ. A mesh fitted to Σ places this kink along element faces, so that it does not have to be approximated by smooth polynomials inside elements.

3. An interface-fitted space–time DG-in-time method

3.1 Meshes

Let $0 = t_0 < \cdots < t_N = T$, $I_n = (t_{n-1}, t_n)$, $\tau_n = t_n - t_{n-1}$, and $Q_n = \Omega \times I_n$. For each n let $\mathcal{T}_h^n$ be a conforming simplicial mesh of $\overline{Q_n}$ in $\mathbb{R}^{d+1}$ (no tensor-product structure is required) such that Σ_h^n, a degree-p_Σ polynomial approximation of $\Sigma \cap Q_n$, is a union of (possibly curved) faces of $\mathcal{T}_h^n$. The meshes $\mathcal{T}_h^n$ and $\mathcal{T}_h^{n+1}$ may be *non-matching* at $t = t_n$. Write $Q_{i,h}^n$ for the resulting discrete subdomains and $S_h := (Q_1 \triangle Q_{1,h})$ for the *mismatch strip*. Faces are collected in $\mathcal{F}_h^n$; for $F \in \mathcal{F}_h^n$ the normal is $\nu_F = (\nu_{x,F}, \nu_{t,F})$, and *horizontal* faces (on $t = \text{const}$) have $\nu_{x,F} = 0$.

Assumption 2 (Mesh regularity) *(i) $\mathcal{T}_h^n$ is shape-regular in $\mathbb{R}^{d+1}$ in a fixed metric M_K (so anisotropic elements are allowed after the linear map $M_K^{1/2}$); (ii) the spatial slices $K \cap \{t = s\}$ are uniformly shape-regular up to a metric-dependent factor in s ("slice regularity"); (iii) the distance between Σ and Σ_h satisfies $\delta := supdist(\Sigma_h, \Sigma) \lesssim h^{p_\Sigma+1}$ locally.*

3.2 Scheme

Let $V_n = \{v \in C^0(\overline{Q_n})\colon\ v|_K \in \mathbb{P}_p(K)\ \ \forall K \in \mathcal{T}_h^n,\ \ v = 0 \text{on} \partial\Omega \times I_n\}$. Denote by $v^{n-1,+}$ and $v^{n,-}$ the traces of $v \in V_n$ at t_{n-1} and t_n, and by $u_h^{n-1,-}$ the trace of the previous slab solution (with $u_h^{0,-} := \Pi_h u_0$, an L^2-projection). For slab n, find $u_h \in V_n$ such that for all $v \in V_n$

$$\int_{Q_n} (\partial_t u_h\, v + \sigma_h(x,t,\nabla u_h) \cdot \nabla v - f(x,t,u_h)\, v)\, \mathrm{d}x\, \mathrm{d}t + \int_\Omega [[u_h]]^{n-1}\, v^{n-1,+}\, \mathrm{d}x$$

$$= \int_{\Sigma_h^n} g(x,t,u_h)\, |\nu_x|\, v\, \mathrm{d}\Sigma_h, \tag{6}$$

with the temporal jump $[[u_h]]^{n-1} := u_h^{n-1,+} - u_h^{n-1,-}$ (the integral over Ω is understood on the common refinement of the two meshes), and $\sigma_h = \sigma_i$ on $Q_{i,h}^n$.

Remark 3.1 (Role of the interface) *Since Σ_h consists of faces of the mesh, the flux condition $[[\sigma \cdot n]] = g$ is imposed weakly and conservatively by the single integral on the right-hand side of (6), and the time component of the normal does not appear. The speed of the front, $V_n = -\varphi_t/|\nabla_x \varphi|$, influences the method only through the shape of the mesh, namely the tilt of the faces of Σ_h, and through the weight $|\nu_x| = (1 + V_n^2)^{-1/2}$. For this reason, the scheme contains no condition that ties τ_n to the number of cells the front crosses in one step. The experiments in Section 6 do not test this property.*

Theorem 3.2 (Discrete well-posedness and stability) *Let Assumption 1 hold. There is $\tau_0 = \tau_0(\alpha, L_f, L_g, \Omega, \varphi) > 0$ such that for $\tau_n \le \tau_0$ the scheme (6) has a unique solution, and*

$$\| u_h^{N,-} \|^2 + \sum_{n=1}^N \| [[u_h]]^{n-1} \|^2 + \alpha \| \nabla u_h \|_{L^2(Q)}^2 \ \le\ C(\| u_0 \|^2 + \| f(\cdot,\cdot,0) \|_{L^2(Q)}^2 + \| g(\cdot,\cdot,0) \|_{L^2(\Sigma)}^2), \tag{7}$$

with C depending on α, L_f, L_g, T but *not* on the mesh, the number of slabs or the interface speed.

Proof (sketch). Test (6) with $v = u_h$. Because u_h is continuous in Q_n,

$$\int_{Q_n} \partial_t u_h\, u_h\, \mathrm{d}x\, \mathrm{d}t + \int_\Omega [[u_h]]^{n-1} u_h^{n-1,+}\, \mathrm{d}x = \tfrac{1}{2} \| u_h^{n,-} \|^2 - \tfrac{1}{2} \| u_h^{n-1,-} \|^2 + \tfrac{1}{2} \| [[u_h]]^{n-1} \|^2 ,$$

using $a(a-b) = \frac{1}{2}(a^2 - b^2 + (a-b)^2)$. Strong monotonicity yields the term $\alpha \| \nabla u_h \|^2$, while the Lipschitz continuity of f and g and the uniform trace inequality control the right-hand side. Summing over n and applying a discrete Gronwall lemma give (7). The same computation for the difference of two solutions shows that the slab operator is strongly monotone on V_n when $\tau_n \le \tau_0$. Existence then follows from the Browder–Minty theorem on the finite-dimensional space V_n.

4. A posteriori and a priori analysis

4.1 Residuals and indicators

Let $\tilde{u}_h$ be the *computed* (inexact) approximation: it lies in V_n on each slab but satisfies (6) only up to an algebraic residual. Let $e = u - \tilde{u}_h$. For $K \in \mathcal{T}_h^n$ and $F \in \mathcal{F}_h^n$ define the element and face residuals

$$r_K := f(\cdot,\cdot,\tilde{u}_h) - \partial_t \tilde{u}_h + \nabla\cdot\sigma_h(\cdot,\cdot,\nabla\tilde{u}_h)|_K,$$
$$j_F := \begin{cases} [[\sigma_h(\nabla\tilde{u}_h)\cdot\nu_{x,F}]]_F & F \not\subset \Sigma_h^n,\ F \not\subset \partial\Omega\times I_n, \\ [[\sigma_h(\nabla\tilde{u}_h)\cdot\nu_{x,F}]]_F - |\nu_{x,F}|\, g(\cdot,\cdot,\tilde{u}_h) & F \subset \Sigma_h^n, \\ 0 & \text{otherwise.} \end{cases}$$

Let $h_{x,K}$ be the spatial diameter of K and $h_{x,F}$ that of F (its projection onto $\mathbb{R}^d$). The indicators are

$$\eta_{S,K}^2 := \frac{h_{x,K}^2}{\alpha} \| r_K \|_{L^2(K)}^2 + \frac{1}{\alpha}\sum_{F\subset\partial K} h_{x,F} \| j_F \|_{L^2(F)}^2 \quad (\text{space, incl. intra} - \text{slab time resolution}), \tag{8}$$

$$\eta_{T,n}^2 := \frac{\beta^2}{\alpha}\,\tau_n \| \nabla[[\tilde{u}_h]]^{n-1} \|_{L^2(\Omega)}^2 \quad (\text{slab} - \text{wise temporal}), \qquad \eta_{J,n}^2 := \| [[\tilde{u}_h]]^{n-1} \|_{L^2(\Omega)}^2 \quad (\text{time jump}), \tag{9}$$

$$\eta_{\Gamma,K}^2 := \frac{1}{\alpha} \| (\sigma_1 - \sigma_2)(\cdot,\cdot,\nabla\tilde{u}_h) \|_{L^2(S_K)}^2 + |S_K|\,\theta_K^2 \quad (\text{interface geometry}), \tag{10}$$

where $S_K := S_h \cap K$, $|S_K|$ is its $(d+1)$-dimensional measure, and $\theta_K := \| g(\cdot,\cdot,\tilde{u}_h) \|_{W^{1,\infty}(K\cap\Sigma_h)}$ (set $\eta_{\Gamma,K} = 0$ if K does not touch Σ_h). The algebraic indicator is

$$\eta_{N,n} := \sup_{0\neq v\in V_n} \frac{\mathcal{R}_n(\tilde{u}_h; v)}{|||v|||_n}, \qquad |||v|||_n^2 := \| v^{n,-} \|^2 + \| v^{n-1,+} \|^2 + \alpha \| \nabla v \|_{L^2(Q_n)}^2, \tag{11}$$

where $\mathcal{R}_n(\tilde{u}_h;\cdot)$ is the residual of (6) (the left-hand side minus the right-hand side). In practice $\eta_{N,n}$ is obtained from the Riesz representative of the algebraic residual, for example with one multigrid cycle. The global estimator is

$$\eta^2 := \eta_S^2 + \eta_T^2 + \eta_J^2 + \eta_\Gamma^2 + \eta_N^2, \qquad \eta_\bullet^2 = \sum \eta_{\bullet,\cdot}^2. \tag{12}$$

Lemma 4.1 (Geometric consistency) *Let Assumptions 1–2 hold. For every $v \in L^2(0,T;H_0^1(\Omega))$ and every $w_h \in V_n$ on slab n,*

$$\left| \int_{Q_n} (\sigma_h - \sigma)(\cdot,\cdot,\nabla w_h)\cdot\nabla v \right| \le \| (\sigma_1 - \sigma_2)(\cdot,\cdot,\nabla w_h) \|_{L^2(S_h\cap Q_n)} \| \nabla v \|_{L^2(Q_n)}, \tag{13}$$

$$\left| \int_\Sigma g\,|\nu_x|\, v\,\mathrm{d}\Sigma - \int_{\Sigma_h} g\,|\nu_x|\, v\,\mathrm{d}\Sigma_h \right| \le C_\Gamma \Big(\sum_K |S_K| \| g \|_{W^{1,\infty}(K\cap\Sigma_h)}^2\Big)^{1/2} \| v \|_{L^2(0,T;H^1(\Omega))}. \tag{14}$$

Here $g = g(\cdot,\cdot,w_h)$ and C_Γ depends only on φ and the shape regularity.

Proof (sketch). (13): $\sigma_h - \sigma$ vanishes outside S_h and equals $\pm(\sigma_1 - \sigma_2)$ on S_h; apply Cauchy–Schwarz. (14): take a slice at fixed t so that only *spatial* derivatives of v are needed. The curves/surfaces $\Gamma(t)$ and $\Gamma_h(t)$ bound a thin region $S_h(t)$; the divergence theorem on $S_h(t)$ applied to $g\,v\,\bar{n}$ ($\bar{n}$ a smooth extension of the unit normal) gives $\lesssim \int_{S_h(t)} (|\nabla(gv)| + |gv|\,|\nabla\cdot\bar{n}|)\,\mathrm{d}x$; Cauchy–Schwarz in x and then in t yields (14). The change of surface measure $\mathrm{d}\Sigma \leftrightarrow \mathrm{d}s\,\mathrm{d}t$ is absorbed through $|\nu_x|$ and is Lipschitz in the geometry.

4.2 Reliability and efficiency

Let $|||e|||^2 := \sup_t \| e(t) \|_{L^2(\Omega)}^2 + \alpha \| \nabla e \|_{L^2(Q)}^2$.

Theorem 4.2 (Reliability) *Under Assumptions 1–2 and with $\tau_n \le \tau_0$ there is a constant C_{rel} depending on κ_σ, L_f, L_g, T, p, shape regularity and the slice-regularity constant, but independent of h, τ_n and the interface speed, such that*

$$|||u - \tilde{u}_h||| \;\le\; C_{\mathrm{rel}}\,\eta. \tag{15}$$

Proof (sketch). Step 1 (error equation). For $v \in H_0^1$ and a.e. t in a slab, $\langle \partial_t e, v\rangle + (\sigma(\nabla u) - \sigma(\nabla\tilde{u}_h), \nabla v) = (f(u) - f(\tilde{u}_h), v) + (g(u) - g(\tilde{u}_h), v)_\Gamma + \mathcal{R}(v)$, where $\mathcal{R}$ is the residual functional collecting (a) elementwise integration by parts of $\sigma_h(\nabla\tilde{u}_h)$, which produces r_K and j_F (horizontal faces do not contribute because $\nu_x = 0$ there); (b) the geometric consistency terms of Lemma 4.1; and (c) the algebraic defect. *Step 2 (time reconstruction).* Following Makridakis–Nochetto (2003) and the DG-in-time analysis of Schötzau–Schwab (2000), one introduces on each slab a continuous-in-time reconstruction $\hat{u}_h = \tilde{u}_h + (\text{a linear} - \text{in} - \text{time correction built from} [[\tilde{u}_h]]^{n-1})$. The difference $\hat{u}_h - \tilde{u}_h$ is bounded in $L^\infty(L^2)$ by η_J and in $L^2(H^1)$ (weighted by $\beta/\sqrt{\alpha}$) by η_T; this is the origin of (9). *Step 3 (parabolic stability for $u - \hat{u}_h$).* Testing with $u - \hat{u}_h$ and using strong monotonicity,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} \| u - \hat{u}_h \|^2 + \alpha \| \nabla(u - \hat{u}_h) \|^2 \le \| \mathcal{R} \|_{H^{-1}} \| \nabla(u - \hat{u}_h) \| + (L_f + C_\epsilon L_g) \| u - \hat{u}_h \|^2 + \epsilon \| \nabla(u - \hat{u}_h) \|^2 + \cdots.$$

Step 4 (bounding $\| \mathcal{R} \|_{H^{-1}}$*).* Apply, for each t, a Scott–Zhang quasi-interpolation operator on the spatial slice mesh (shape-regular by Assumption 2(ii)); Galerkin orthogonality up to η_N yields $\| \mathcal{R}(t) \|_{H^{-1}} \lesssim \alpha^{1/2} (\sum_K \eta_{S,K}^2(t))^{1/2}$ + geometry + algebra. *Step 5 (Gronwall).* Integrate in time, absorb the ϵ-terms, and apply Gronwall's lemma. The constant $e^{C(L_f, L_g)T}$ is the only place where T enters. A complete proof also has to track how Step 4 depends on slice regularity near the slab boundaries. We leave out these long but routine estimates.

Remark 4.3 (Sharpness of η_Γ) *The bound (13) uses only* $\| \nabla e \|_{L^2(S_h)} \leq \| \nabla e \|_{L^2(Q)}$*, and* $\eta_\Gamma^2 \approx \alpha^{-1} \| \sigma_1 - \sigma_2 \|_{L^2(S_h)}^2 \sim \delta$*. This is not pessimistic. In the strip, the exact solution carries the gradient of the other phase, whereas the discrete solution essentially continues the gradient of the phase of its element. Hence* $\| \nabla(u - u_h) \|_{L^2(S_h)} \sim \| \nabla u_1 - \nabla u_2 \|_{L^2(S_h)} \sim \delta^{1/2}$*, as stated in Proposition 4.5.*

Theorem 4.4 (Local efficiency) *Let* ω_K *be the patch of* K *and let* $\| \mathcal{R} \|_{*,\omega_K}$ *be the dual norm of the residual functional on* $H_0^1(\omega_K \cap \{t\})$*, integrated in time. Then for every* $K \in \mathcal{T}_h^n$ *not touching* Σ_h*,*

$$\eta_{S,K} \;\leq\; C_{eff}(\alpha^{-1/2} \| \mathcal{R} \|_{*,\omega_K} + \mathrm{osc}_K) \;\leq\; C'_{eff}(\| \partial_t e \|_{L^2(H^{-1}(\omega_K))} + \beta \alpha^{-1/2} \| \nabla e \|_{L^2(\omega_K)} + L_f \| e \|_{L^2(\omega_K)}) + C_{eff}\, \mathrm{osc}_K, \tag{16}$$

where $\mathrm{osc}_K = \alpha^{-1/2} h_{x,K} \| r_K - \bar{r}_K \|_{L^2(K)}$ + (face analogue) and $\bar{r}_K$ is a polynomial approximation. For elements near Σ_h the bound holds with an additional term, the oscillation of g on the faces of Σ_h weighted by θ_K. The temporal indicators are efficient only globally in time: $\eta_T^2 + \eta_J^2 \lesssim |||e|||_*^2 + \mathrm{osc}$, where $||| \cdot |||_*$ additionally contains $\| \partial_t e \|_{L^2(H^{-1})}$.

Proof (sketch). Use standard Verfürth bubble-function arguments: test the residual with $v = b_K r_K$ (element bubble) and $v = b_F j_F$ (face bubble extended to the patch)، use the equivalence of norms on polynomials, inverse estimates in the spatial slices (again Assumption 2(ii)), and the error equation of Step 1. The estimate is local in time because the bubbles are supported in K or in the patch ω_F of a face.

4.3 A priori analysis: the cost of geometry

Let u^{ext} denote the function equal to the smooth extension of $u|_{Q_i}$ on $Q_{i,h}$, and let $||| \cdot |||_{\mathrm{br}}$ be the energy norm measured against u^{ext} (the strip is discounted) ; $||| \cdot |||$ is the norm against the exact u.

Proposition 4.5 (A priori error and geometry) *Let* $u|_{Q_i} \in H^{p+1}(Q_i)$ *be extended smoothly,* $\tau_n \leq \tau_0$*,* $g \in W^{1,\infty}$*, and let* Σ_h *have degree* p_Σ *with* $\delta \lesssim h^{p_\Sigma + 1}$*. Then:*

$$|||u^{ext} - u_h|||_{\mathrm{br}} \leq C(h^p + h^{p_\Sigma + 1/2}), \tag{17}$$

$$|||u - u_h||| \leq C(h^p + h^{p_\Sigma + 1/2}) + \| \nabla(u - u^{ext}) \|_{L^2(S_h)} \;\lesssim\; h^p + h^{(p_\Sigma + 1)/2}. \tag{18}$$

In general, the last term cannot be improved: if ∇u jumps by a non-vanishing amount across Σ, $\| \nabla(u - u^{ext}) \|_{L^2(S_h)} \simeq |S_h|^{1/2} \sim \delta^{1/2}$.

Proof (sketch). Strang's lemma splits the error into the approximation error of u^{ext} in the fitted space V_h, which is of order h^p, and a consistency error. The part of the consistency error that comes from the coefficient mismatch is bounded by $\| \nabla u \|_{L^2(S_h)} \| \nabla v_h \|_{L^2(S_h)} \lesssim \delta^{1/2} (\delta / h)^{1/2} = \delta h^{-1/2}$ using an inverse inequality on cut-adjacent elements; the interface-source part is treated via (14). This gives (17). For (18), the triangle inequality adds $u^{ext} - u$, which is supported in S_h.

Corollary 4.6 (Interface degree) *In the exact-domain norm, order* h^p *requires* $\delta \lesssim h^{2p}$*; if* $\delta \sim h^{p_\Sigma + 1}$ *in the worst case, this means* $p_\Sigma \geq 2p - 1$ *(*$p_\Sigma = 1$ *for* $p = 1$*,* $p_\Sigma = 3$ *for* $p = 2$*). If* δ *is smaller than the worst case, e.g. because the vertices are placed exactly on* Σ*, a lower* p_Σ *may suffice (Section 6).*

5. Adaptive anisotropic space–time algorithm

5.1 Metric-based refinement rule

We equip space–time with a Riemannian metric $\mathcal{M} : \overline{Q} \to \mathrm{SPD}(d+1)$ and aim for elements that are (quasi-)equilateral in $\mathcal{M}$. The metric is the intersection (maximum in the Löwner sense, (Loseille & Alauzet, 2011)) $\mathcal{M} = \mathcal{M}_u \cap \mathcal{M}_\Sigma$ of

$$\mathcal{M}_u = \frac{1}{\varepsilon_u} |\mathbf{H}(\tilde{u}_h)| \quad (\text{absolute value of a recovered space} - -\text{time Hessian, one} - \text{sided on each } \mathrm{Q_i}), \tag{19}$$

$$\mathcal{M}_\Sigma(z) = \sum_{j=1}^{d} \frac{|\kappa_j(z)|}{2\, \varepsilon_\Sigma} \mathbf{t}_j \mathbf{t}_j^{\top} + \frac{1}{h_{\max}^2} \nu \nu^{\top}, \qquad z = \pi_\Sigma(\cdot), \tag{20}$$

where κ_j and $\mathbf{t}_j$ are the principal curvatures and directions of $\Sigma \subset \mathbb{R}^{d+1}$, π_Σ the closest-point projection, and ε_Σ a target for the strip width. Since for a linear interface approximation the width of the strip on a segment of tangential length ℓ_j is $\delta \approx \kappa_j \ell_j^2/8$, (20) prescribes $\ell_j \approx (8\varepsilon_\Sigma/|\kappa_j|)^{1/2}$ along each principal direction. The curvature κ_j in space–time contains *both* the spatial curvature of $\Gamma(t)$ and the interface acceleration $\partial_t V_n$; a front that accelerates is thus automatically resolved more finely in time-like tangent directions. The targets $\varepsilon_u, \varepsilon_\Sigma$ are tied to the estimator: ε_Σ is reduced whenever η_Γ dominates and ε_u whenever η_S dominates (see the marking step).

5.2 The loop

Let $\theta \in (0,1)$ be the Dörfler parameter (Dörfler, 1996), $\gamma_N \in (0,1)$ the algebraic balance parameter, and $\gamma_T > 0$ the slab-temporal balance parameter.

Algorithm 1. ST-AFEM for nonlinear parabolic problems with a moving interface

1. **Input:** $\varphi, \sigma_i, f, g, u_0$, tolerance TOL, parameters $\theta, \gamma_N, \gamma_T$, initial slab length τ_1.
2. **for** $n = 1, \dots, N$ **do**
3. Build $\mathcal{T}_h^n$ fitted to Σ_h^n (vertices on Σ obtained by projecting onto $\{\varphi = 0\}$).
4. **repeat**
5. Solve: inexact Newton for (6) until $\eta_{N,n} \le \gamma_N(\eta_{S,n}^2 + \eta_{\Gamma,n}^2 + \eta_{T,n}^2 + \eta_{J,n}^2)^{1/2}$.
6. Estimate: compute $\eta_{S,K}, \eta_{\Gamma,K}$ for $K \in \mathcal{T}_h^n$ and $\eta_{T,n}, \eta_{J,n}$.
7. **if** $\eta_{T,n}^2 + \eta_{J,n}^2 > \gamma_T\,(\eta_{S,n}^2 + \eta_{\Gamma,n}^2)$ **then**
8. Slab control: replace $\tau_n \leftarrow \tau_n/2$, restart slab (reuse $\mathcal{T}_h^n$ spatial data).
9. **else if** $\eta_n > \mathrm{TOL}_n$ **then**
10. Mark: minimal $\mathcal{M}_n \subset \mathcal{T}_h^n$ s.t. $\sum_{K\in\mathcal{M}_n}(\eta_{S,K}^2 + \eta_{\Gamma,K}^2) \ge \theta \sum_K (\eta_{S,K}^2 + \eta_{\Gamma,K}^2)$.
11. Metric update: update ε_u (resp. ε_Σ) by factor $1/2$ on marked elements dominated by η_S (resp. η_Γ); regenerate $\mathcal{T}_h^n$ by local metric-based refinement, edge flips, node snapping onto Σ, and slice-regularity check.
12. **end if**
13. **until** $\eta_n \le \mathrm{TOL}_n$
14. Coarsen (optional) using the inverse Dörfler criterion; set τ_{n+1} from the last accepted τ_n and the ratio $\eta_{T,n}^2/\eta_n^2$.
15. **end for**

Remark 5.1 (Local time resolution)**.** *The temporal resolution has two parts: inside a slab, it is controlled element-wise by $\eta_{S,K}$ (since $\partial_t \tilde{u}_h$ is part of r_K and the metric $\mathcal{M}$ can refine elements in the time direction only where needed) , whereas $\eta_{T,n}, \eta_{J,n}$ control the slab-to-slab coupling. A slab can therefore be long, for example to suit a parallel space–time solve, and the slab-control test in line 7 keeps the slab length from becoming the limiting factor.*

Remark 5.2 (Estimator reduction: what is proved and what is not) *With this marking, one expects an estimator-reduction property for $\eta_S + \eta_\Gamma$ if refinement reduces the local metric size by a fixed factor and the data are piecewise polynomial. We do not prove it. Contraction of the error and quasi-optimal complexity are open for a nonlinear space–time DG method with moving geometry.*

5.3 Complexity

A uniform isotropic space–time mesh in $d+1$ dimensions with spacing h has $\mathcal{O}(h^{-(d+1)})$ unknowns. Suppose the solution has a feature of width $\epsilon \ll h$ normal to Σ and is smooth along it. An anisotropic mesh with normal size $\sim \epsilon$ and tangential size $\sim h$ then needs only $\mathcal{O}(h^{-d})$ elements, provided a bounded number of element layers resolves the feature, while an isotropic mesh needs $\mathcal{O}(\epsilon^{-1}h^{-d})$. This heuristic motivates (20). We do not prove a complexity result here.

6. Numerical Experiments

6.1 Setting

All results in this section come from a prototype written in Python with NGSolve 6.2. The code and the raw result files are supplied together with this manuscript. The prototype implements the scheme (6) for one space dimension, so that Q is a planar domain triangulated by a mesh that fits the curved interface Σ. Refinement is isotropic bisection, and new vertices on Σ_h are moved onto Σ. Several parts of the method are not covered by these experiments. The anisotropic metric (19) – (20) is not implemented, and neither is time marching over several slabs with non-matching meshes. The whole interval $(0, T)$ is a single

slab, so η_J and η_T only measure the mismatch with the initial data, and this mismatch is tiny in every run. The prototype also does not handle $d = 2{,}3$ and has not been compared with ALE or unfitted methods. The statements in contributions C1 and C4 about several slabs, anisotropy and independence of the time step from the front speed are therefore not backed by data here.

Test problem. We take $\Omega = (0{,}1)$ and $T = 1$. The front is an arc of a conic in the (x,t)-plane from $(0.2{,}0)$ to $(0.7{,}1)$ with initial speed 0.3 and final speed 0.9, so it accelerates. Netgen represents this arc as a rational quadratic spline; we derived its implicit equation and checked it against the mesh vertices, which lie on it up to 10^{-15}. Let ρ be a level-set function for the front ($\rho = 0$ on Σ, $\rho < 0$ in Q_1) and let $\epsilon = 0.05$. The manufactured solution is

$$u_i = \sin(\pi x)[(1 + \frac{t}{2})(1 + \kappa_i \rho) + 0.3\,(1 + t)\, e^{-\rho^2/\epsilon^2}], \qquad \kappa_1 = 2, \qquad \kappa_2 = -1.$$

It is continuous, its x-derivative jumps across Σ, and a narrow Gaussian layer travels with the front. The flux is $\sigma_i(\xi) = a_i\, \xi\, (1 + \xi^2)^{0.1}$, a growth law with exponent $q = 2.2$. The data f and g are computed symbolically so that u solves the problem exactly. The standard case is $a_1 = a_2 = 1$. Figure 1 shows an adapted mesh and the discrete solution.

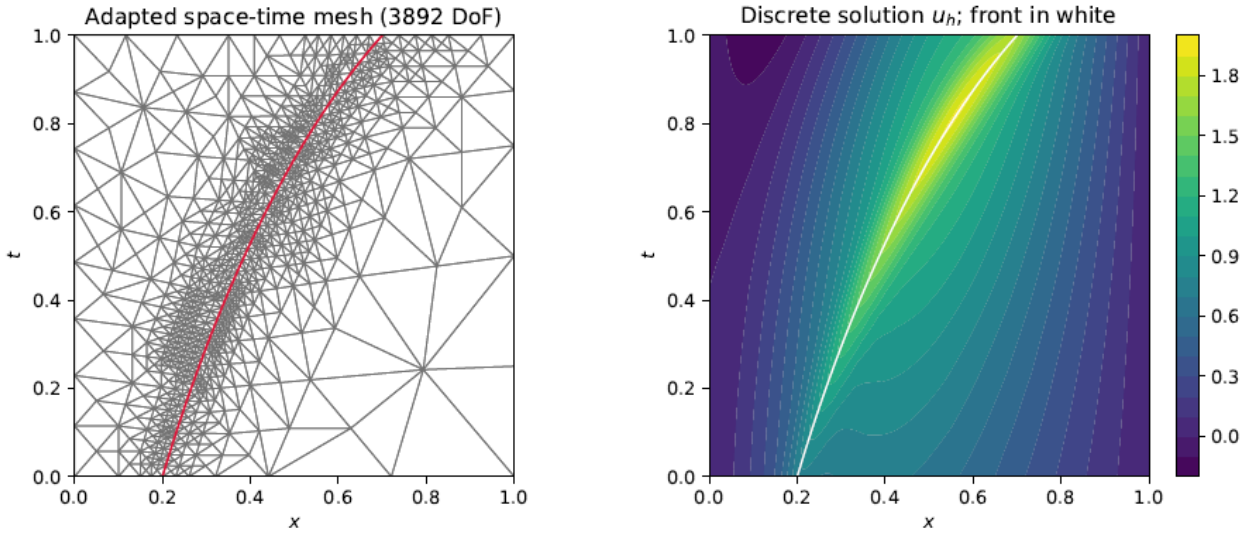


Fig. 1. Left: space–time mesh after eight adaptive iterations ($p = p_\Sigma = 2$, 3892 degrees of freedom). The interface Σ_h is drawn in red; the mesh follows it as a line of tilted elements. Right: the discrete solution, with the front in white.

Error measure and implementation notes. The discrete problem is solved by Newton's method, and the stopping test uses the algebraic estimator η_N of (11), computed from the Riesz representative of the residual in the norm $|||\cdot|||_n$. The supremum over time in $|||\cdot|||$ was not implemented. Instead, we report

$$\mathfrak{E}^2 = \| e(T) \|^2_{L^2(\Omega)} + \alpha \, \| \partial_x e \|^2_{L^2(Q)}, \qquad \alpha = \min(a_1, a_2),$$

which is bounded by the norm of Theorem 4.2. The effectivity index $\mathrm{eff} = \eta/\mathfrak{E}$ is therefore a lower bound for $\eta/|||e|||$. The mismatch strip S_h is integrated to first order across its thickness, with $|d| \approx |\rho|/|\rho_x|$. The broken error $\mathfrak{E}_{\mathrm{br}}$ is the same quantity with the strip left out. Two simplifications were made in the indicators. In (8), the face size $h_{x,F}$ is replaced by the spatial diameter $h_{x,K}$ of the element, which is an upper bound. In (10), the norm $\| g \|^2_{W^{1,\infty}}$ is replaced by the pointwise value $g^2 + |\nabla g|^2$.

6.2 Uniform refinement

Table 1 lists results for $(p, p_\Sigma) = (1{,}1), (2{,}1), (2{,}2)$ and Figure 2 (left) plots the error against the number of unknowns. The estimator was above the error in every run, and the smallest effectivity index observed was 9.2. For $p = 2$ the error falls at about h^2. For $p = 1$ it falls at about h once the mesh resolves the moving layer; the first levels are slower because the layer has width $\epsilon = 0.05$.

Table 1. Uniform refinement with $a_1 = a_2 = 1$. $\mathfrak{E}$ is the error measure defined above, EOC is the observed order in h (one level is one halving of h), η_Γ the geometry indicator, and $\mathrm{eff} = \eta/\mathfrak{E}$.

Level	DoF	$\mathfrak{E}$	EOC	η_S	η_Γ	EOC $_\Gamma$	Eff
$p = 1,\ p_\Sigma = 1$							
0	27	2.91×10^0	–	3.48×10^1	6.53×10^0	–	12.2
1	89	1.36×10^0	1.10	1.95×10^1	3.94×10^0	0.73	14.7

2	321	8.69×10^{-1}	0.64	1.05×10^{1}	1.96×10^{0}	1.01	12.3
3	1217	6.33×10^{-1}	0.46	5.88×10^{0}	9.80×10^{-1}	1.00	9.5
4	4737	3.41×10^{-1}	0.89	3.09×10^{0}	4.90×10^{-1}	1.00	9.2
5	18689	1.71×10^{-1}	1.00	1.57×10^{0}	2.45×10^{-1}	1.00	9.4
$p = 2,\ p_\Sigma = 1$							
0	89	2.19×10^{0}	–	3.69×10^{1}	6.53×10^{0}	–	17.1
1	321	9.60×10^{-1}	1.19	2.00×10^{1}	3.94×10^{0}	0.73	21.3
2	1217	5.45×10^{-1}	0.82	8.60×10^{0}	1.96×10^{0}	1.01	16.2
3	4737	1.67×10^{-1}	1.71	2.55×10^{0}	9.80×10^{-1}	1.00	16.4
4	18689	3.93×10^{-2}	2.09	5.79×10^{-1}	4.90×10^{-1}	1.00	19.3
5	74241	1.12×10^{-2}	1.81	1.47×10^{-1}	2.45×10^{-1}	1.00	25.5
$p = 2,\ p_\Sigma = 2$							
0	89	2.03×10^{0}	–	3.64×10^{1}	9.75×10^{-2}	–	17.9
1	321	9.40×10^{-1}	1.11	2.00×10^{1}	2.35×10^{-2}	2.05	21.3
2	1217	5.41×10^{-1}	0.80	8.59×10^{0}	5.85×10^{-3}	2.01	15.9
3	4737	1.65×10^{-1}	1.71	2.55×10^{0}	1.46×10^{-3}	2.00	15.5
4	18689	3.75×10^{-2}	2.14	5.79×10^{-1}	3.65×10^{-4}	2.00	15.5
5	74241	9.60×10^{-3}	1.97	1.47×10^{-1}	9.14×10^{-5}	2.00	15.4

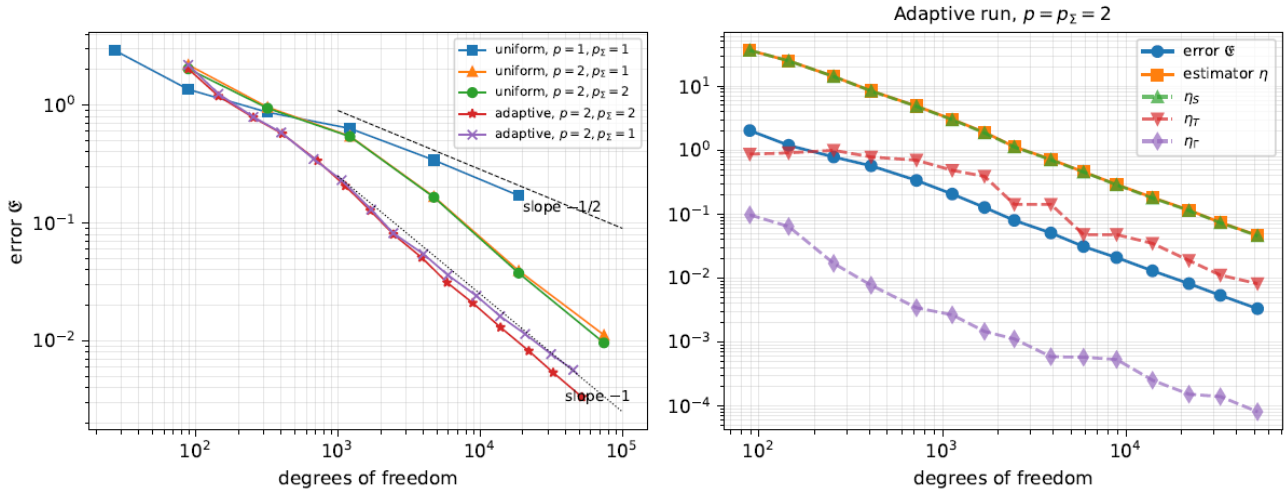


Fig. 2. Left: error against degrees of freedom for uniform and adaptive refinement. Right: error, total estimator and its parts along the adaptive run with $p = p_\Sigma = 2$. The estimator follows the error at a nearly constant factor; η_S dominates, and η_Γ is smaller by orders of magnitude.

6.3 The role of the interface geometry

The geometry indicator behaves as the analysis suggests. It falls like h for $p_\Sigma = 1$ and like h^2 for $p_\Sigma = 2$ (EOC $_\Gamma = 1.00$ and 2.00 on the finest levels). Since $\eta_\Gamma \sim \delta^{1/2}$, this means $\delta \sim h^2$ and $\delta \sim h^4$ respectively. The second value is better than the worst case $\delta \sim h^{p_\Sigma+1}$ used in Corollary 4.6. We think this comes from the mesh vertices lying exactly on Σ and the edge midpoints being projected onto it, and we do not expect it to hold for every front. For $(p, p_\Sigma) = (2,1)$ the geometry term starts to matter at the finest level. There η_Γ is larger than η_S, the exact-domain error departs from the h^2 trend (EOC drops from 2.09 to 1.81), and eff rises from 16 to 25. The broken error, which ignores the strip, keeps a rate close to h^2 (Table 2 and Figure 3). This matches Proposition 4.5: in the exact norm the strip is a real source of error. With $p_\Sigma = 2$ no such effect is visible up to $7.4 \cdot 10^4$ unknowns.

Table 2. Effect of the interface degree for $p = 2$, levels 3 to 5: error in the exact-domain norm, broken error (strip left out), and geometry indicator.

p_Σ	DoF	$\mathfrak{E}$	$\mathfrak{E}_{br}$	η_Γ
1	4737	1.67×10^{-1}	1.65×10^{-1}	9.80×10^{-1}
1	18689	3.93×10^{-2}	3.75×10^{-2}	4.90×10^{-1}

1	74241	1.12×10^{-2}	9.60×10^{-3}	2.45×10^{-1}
2	4737	1.65×10^{-1}	1.65×10^{-1}	1.46×10^{-3}
2	18689	3.75×10^{-2}	3.75×10^{-2}	3.65×10^{-4}
2	74241	9.60×10^{-3}	9.60×10^{-3}	9.14×10^{-5}

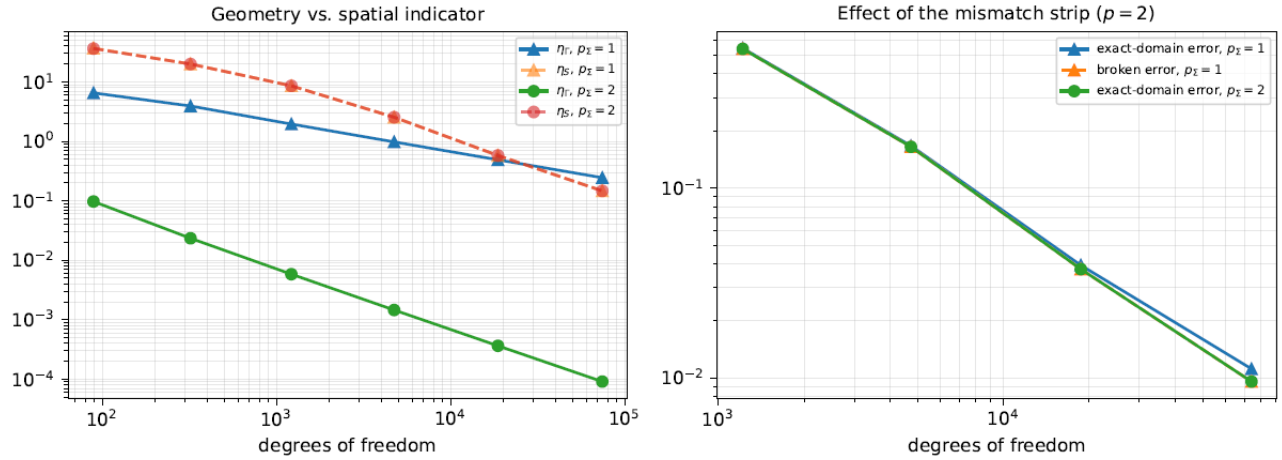


Fig. 3. Left: geometry and spatial indicators for $p = 2$ with a linear and a quadratic interface. With $p_\Sigma = 1$ the geometry indicator falls only like h and overtakes η_S on the finest level. Right: error in the exact-domain norm and in the broken norm.

6.4 Diffusion contrast

For the contrast study, we set $a_1 = 1$, $a_2 = 1/c$ and $p = p_\Sigma = 2$. The estimator stays above the error and the error still falls at about h^2 (EOC between 2.1 and 2.3). The effectivity index, however, grows quickly with the contrast: it is 15, $1.2 \cdot 10^3$ and $6.2 \cdot 10^4$ for $c = 1, 10^2, 10^4$ (Table 3 and Figure 4). This agrees with the dependence of C_{rel} on κ_σ in Theorem 4.2. Our estimator is scaled by the global constant $\alpha = \min a_i$ and is not robust with respect to the contrast. A variant with weights and a quasi-monotonicity assumption on the coefficients would be needed.

Table 3. Contrast study with $p = p_\Sigma = 2$, levels 3 and 4 (4737 and 18689 DoF).

$c = a_1/a_2$	$\mathfrak{E}$ (level 3)	$\mathfrak{E}$ (level 4)	EOC	η (level 4)	eff
1	1.65×10^{-1}	3.75×10^{-2}	2.14	5.81×10^{-1}	15
100	1.86×10^{-2}	4.10×10^{-3}	2.18	4.95×10^{0}	1206
10^4	4.06×10^{-3}	8.00×10^{-4}	2.34	4.95×10^{1}	61883

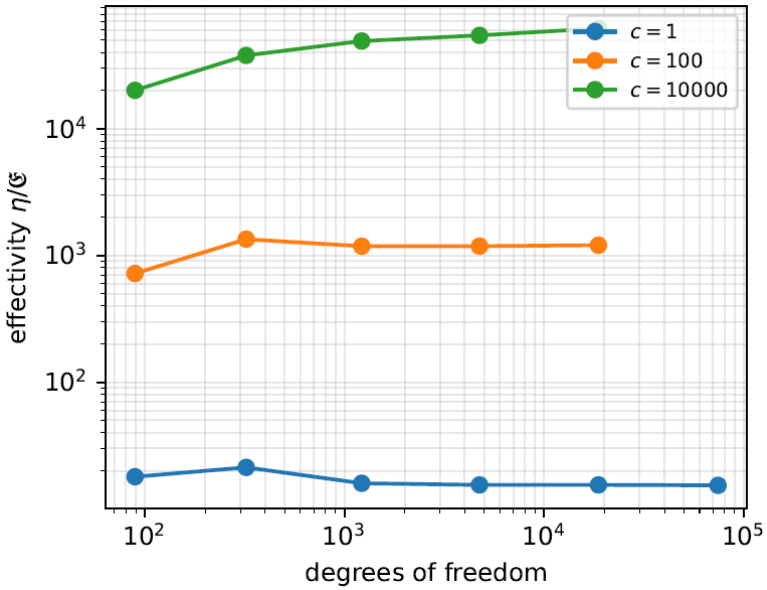


Fig. 4. Effectivity index $\eta/\mathfrak{E}$ for three diffusion contrasts $c = a_1/a_2$ ($p = p_\Sigma = 2$, uniform refinement).

6.5 Adaptive refinement

We ran Algorithm 5.2 without slab control, with $\theta = 0.5$, marking on $\eta_S^2 + \eta_\Gamma^2$ and $p = p_\Sigma = 2$. Between iterations 4 and 14 the effectivity index stays between 13.8 and 15.0 (Table 4). The adaptive error of $8.2 \cdot 10^{-3}$ at $22\,019$ unknowns is a little below the uniform error of $9.6 \cdot 10^{-3}$ at $74\,241$ unknowns, so for equal accuracy adaptivity needs about 3.4 times as few unknowns in this example. Both strategies converge at about the same rate in terms of unknowns (Figure 2), so the gain is a constant factor and not a better rate. With $p_\Sigma = 1$ the adaptive run also converges, reaching an error of $5.7 \cdot 10^{-3}$ at $45\,194$ unknowns with eff between 15 and 16.

Table 4. Adaptive and uniform refinement for $p = p_\Sigma = 2$ and $a_1 = a_2 = 1$.

Strategy	DoF	$\mathfrak{E}$	eff
adaptive	721	3.37×10^{-1}	14.5
adaptive	1693	1.27×10^{-1}	15.0
adaptive	3892	5.08×10^{-2}	14.2
adaptive	8878	2.09×10^{-2}	13.9
adaptive	22019	8.20×10^{-3}	14.1
adaptive	52047	3.33×10^{-3}	14.0
uniform	4737	1.65×10^{-1}	15.5
uniform	18689	3.75×10^{-2}	15.5
uniform	74241	9.60×10^{-3}	15.4

6.6 Stopping the Newton iteration

We compared two stopping rules in the adaptive runs. The first iterates until $\eta_N \le 10^{-12}$. The second stops when $\eta_N \le \gamma\, \eta_{disc}$ with $\gamma = 0.1$, where η_{disc} is the discretisation estimator from the previous adaptive iteration (Table 5). The second rule cuts the total number of Newton steps from 67 to 18, which is one step per iteration after the first. The price is accuracy. At the last iteration the algebraic estimator is $5.9 \cdot 10^{-3}$, about as large as the error itself ($5.5 \cdot 10^{-3}$), and the error at a similar problem size is about 1.6 times that of the tight run. In this problem eff ≈ 14, so $\gamma = 0.1$ is too loose compared with the true error. A smaller γ, for example 0.01, or a factor $1/\text{eff}$ in the threshold should be used.

Table 5. Newton stopping rules in the adaptive runs ($p = p_\Sigma = 2$), last iteration.

Newton stopping	DoF	$\mathfrak{E}$	total steps	η_N	eff
tight, $\eta_N \le 10^{-12}$	52047	3.33×10^{-3}	67	2.39×10^{-14}	14.0
balanced, $\eta_N \le 0.1\, \eta_{disc}$	56726	5.47×10^{-3}	18	5.95×10^{-3}	9.1

6.7 Limits of these experiments

The study covers one space dimension, one slab, one family of manufactured solutions and one front geometry. There is no comparison with other methods, refinement is isotropic, and the error measure uses $\| e(T) \|$ in place of the supremum over time. The proofs in Section 4 are sketches. The numbers agree with the theory, but they do not prove it.

7. Conclusions and outlook

We have described a space–time method with upwind coupling between slabs for nonlinear parabolic problems with a moving interface, built on meshes that fit the interface. We derived an error estimator with an explicit geometry part, sketched the proofs of its reliability and local efficiency, and showed that the strip between the exact and the discrete interface limits the accuracy in the exact norm. A one-dimensional prototype with a single slab confirmed that the estimator is reliable, that the geometry indicator decays as predicted, that adaptive refinement reduces the number of unknowns by a constant factor compared with uniform refinement, and that the estimator is not robust against a large diffusion contrast. It also showed that a Newton tolerance of 0.1 times the discretisation estimator saves work but leaves too large an algebraic error.

The most urgent open points are on the implementation side. The anisotropic metric, time marching over several slabs with non-matching meshes, and the extension to two and three space dimensions have to be implemented and tested, together with a comparison with ALE and unfitted methods. On the theory side, there are five directions. The first is a coupling with a Stefan-type law, where the front speed follows from $[[\sigma \cdot n]]$ and the mesh becomes part of the unknowns, so a front-tracking indicator is needed. The second is topological change, such as merging or splitting fronts, which probably needs local unfitted patches inside a fitted mesh. The third is hp-adaptivity for smooth fronts. The fourth is a convergence proof and optimality for the adaptive loop. The fifth is a contrast-robust estimator; Table 3 shows that it is needed, and we expect it to require a quasi-monotonicity assumption and weighted quasi-interpolation.

Data and code. The prototype, the run scripts and the raw results behind all tables and figures are supplied with the manuscript.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
ORCID iD: 0009-0008-4190-6323
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers.

References

[1] Bonito, A., Kyza, I., & Nochetto, R. H. (2013). Time-discrete higher-order ALE formulations: Stability. SIAM Journal on Numerical Analysis, 51(1), 577–604. https://doi.org/10.1137/120862715
[2] Burman, E., Claus, S., Hansbo, P., Larson, M. G., & Massing, A. (2015). CutFEM: Discretizing geometry and partial differential equations. International Journal for Numerical Methods in Engineering, 104(7), 472–501. https://doi.org/10.1002/nme.4823
[3] Dörfler, W. (1996). A convergent adaptive algorithm for Poisson's equation. SIAM Journal on Numerical Analysis, 33(3), 1106–1124.
[4] Eriksson, K., & Johnson, C. (1991). Adaptive finite element methods for parabolic problems I: A linear model problem. SIAM Journal on Numerical Analysis, 28(1), 43–77. https://doi.org/10.1137/0728003
[5] Ern, A., & Vohralík, M. (2013). Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. SIAM Journal on Scientific Computing, 35(4), A1761–A1791. https://doi.org/10.1137/120896918
[6] Hansbo, A., & Hansbo, P. (2002). An unfitted finite element method, based on Nitsche's method, for elliptic interface problems. Computer Methods in Applied Mechanics and Engineering, 191(47–48), 5537–5552. https://doi.org/10.1016/S0045-7825(02)00524-8
[7] Langer, U., & Steinbach, O. (Eds.). (2019). Space-time methods: Applications and techniques. De Gruyter.
[8] Lehrenfeld, C., & Olshanskii, M. A. (2019). An Eulerian finite element method for PDEs in time-dependent domains. ESAIM: Mathematical Modelling and Numerical Analysis, 53, 585–614.
[9] Loseille, A., & Alauzet, F. (2011). Continuous mesh framework. Part I: Well-posed continuous interpolation error. SIAM Journal on Numerical Analysis, 49(1), 38–60.
[10] Makridakis, C., & Nochetto, R. H. (2003). Elliptic reconstruction and a posteriori error estimates for parabolic problems. SIAM Journal on Numerical Analysis, 41(4), 1585–1594. https://doi.org/10.1137/S0036142902406314
[11] Schötzau, D., & Schwab, C. (2000). Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method. SIAM Journal on Numerical Analysis, 38(3), 837–875. https://doi.org/10.1137/S0036142999352394
[12] Steinbach, O. (2015). Space-time finite element methods for parabolic problems. Computational Methods in Applied Mathematics, 15, 551–566.
[13] Tezduyar, T. E., Behr, M., & Liou, J. (1992). A new strategy for finite element computations involving moving boundaries and interfaces—The deforming-spatial-domain/space-time procedure: I. The concept and the preliminary numerical tests. Computer Methods in Applied Mechanics and Engineering, 94(3), 339–351. https://doi.org/10.1016/0045-7825(92)90059-S
[14] Verfürth, R. (2013). A posteriori error estimation technique for finite element methods. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679423.001.0001