Efficient Adaptive Solvers for Doubly Degenerate Elliptic Equations using A Posteriori Error Estimates

Elyes Ahmed, Saber Amdouni

Efficient Adaptive Solvers for Doubly Degenerate Elliptic Equations using A Posteriori Error Estimates

Elyes Ahmed* Saber Amdouni†

February 19, 2024

,

Abstract

In this work, we derive a posteriori error estimates for a class of doubly nonlinear and degenerate elliptic equations, including the Stefan problem and fast and slow diffusion in porous media. Our approach employs equilibrated flux reconstructions, providing guaranteed and fully computable upper bounds on an energy-type norm and local efficiency. These bounds remain independent of the strength of nonlinearity and degeneracy rates. These estimators drive an adaptive solver, dynamically switching between nonlinear solvers to achieve optimal iterations. The adaptive algorithm accounts for discretization, regularization, quadrature, and linearization error components. When Newton's method encounters challenges in achieving convergence, the adaptive algorithm transitions to the L-scheme solver. This solver optimally precomputes the stabilization (or tuning) parameter $L > 0$ during an offline phase, mirroring the behavior of the Jacobian. The adaptive algorithm is exemplified through four prototypical examples, showcasing its effective error control and notable computational savings.

Key words: degenerate elliptic equations, fast and slow diffusion, a posteriori error analysis, local efficiency, adaptive linearization, switching criteria

1 Introduction

In this article, we present an adaptive solver designed to approximate doubly degenerate elliptic equations of the Stefan and porous media forms, characterized by the expression $\beta(u) - \Delta\varphi(u) = f + \nabla\cdot\boldsymbol{\xi}$. Our paper focuses specifically on addressing scenarios where both φ and β lack uniform Lipschitz continuity. This is particularly relevant in cases such as the fast diffusion model [25], which is non-Lipschitz at infinity, or when these functions are not Lipschitz-continuous at specific points, as observed in slow diffusion. We also allow for the possibility of plateaus in these functions. These degenerate elliptic equations hold significant relevance in various fields, particularly in porous media applications such as density-dependent groundwater flow modeling [17] and immiscible two-phase flow problems [39, 40]. Additionally, they find crucial applications in biology [29, 41] chemistry [55], and material sciences [7]. It is of utmost importance to highlight that the solution operator of this model not only showcases nonlinearity and degeneracy but also may displays a striking transition between two distinct degeneration rates, which is contingent on the specific choices of φ and β. Consequently, the diffusion process becomes singularly perturbed, potentially resulting in collapsing effects. To effectively address this challenge and ensure numerical stability, common remedies include exploring a regularized version of the problem [23] or reformulating the system with a different form of nonlinearity [9]. These approaches are employed, possibly in combination, to address degenerate cases, especially those featuring double degeneracy with potentially distinct rates, and effectively manage the resulting singular effects.

The purpose of numerical methods for Partial Differential Equations (PDEs) is to approximate the solution, as precisely as possible, to the underlying continuous solution. Accurately approximating complex

*SINTEF Digital Oslo Norway. elyes.ahmed@sintef.no

†Laboratory For Mathematical And Numerical Modeling In Engineering Science, LAMSIN, LR99ES20, University of Tunis El Manar, National Engineering School of Tunis, Tunis, Tunisia. saber.amdouni@enit.utm.tn

PDEs can be challenging, especially when errors are intertwined, arising from various sources such as discretization, regularization, and linearization errors. To effectively address these challenges, the use of a posteriori error estimates becomes crucial. A key property of a reliable a posteriori estimate [30, 34, 35, 32, 33] is its ability to quantify the error in the numerical approximation accurately, making it suitable for guiding adaptive algorithms. Adaptive algorithms enable efficient approximation of localized features in the solution even with intricate features. Considerable progress has been made in recent years regarding a posteriori analysis for linear elliptic equations [31, 52, 16]. However, rigorous a posteriori error estimates for nonlinear elliptic problems, particularly in degenerate cases, remain a less-explored area [48, 44]. Explorations into degenerate parabolic equations have also been undertaken [11, 21], typically under the assumption that nonlinearities are globally Lipschitz-continuous.

Ongoing efforts are dedicated to designing robust and efficient linearization schemes tailored for degenerate elliptic and parabolic problems [51, 6, 43, 50]. While Newton's method [26] is commonly the default choice due to its potential for quadratic convergence, it encounters challenges when the initial solution is not sufficiently close. In order to control degeneracy, Newton's method often involves the introduction of regularized versions of nonlinear functions. Another strategy is to explore globally convergent solvers such as the L-scheme, Picard, or damped Newton's methods; however, this comes at the trade-off of slower convergence. For instance, the L-scheme has demonstrated unconditional convergence for many model problems, but it's slow and only linearly convergent [47]. The quest for a balance between robustness and speed has spurred the development of hybrid solvers, as explored in [6, 45]. In this context, multiple iterations of a slower scheme are employed to generate an initial approximation for Newton's method. This strategy was further refined in [53], where a switching criteria between the L-scheme and Newton's method for the Richards equation was guided by an a posteriori error estimator.

Addressing doubly degenerate elliptic problems presents an additional challenge for nonlinear solvers, demanding robustness in handling scenarios involving singular and double degeneracy [48]. Specifically, φ and β do not necessarily have to be Lipschitz-continuous or single-valued. This entails developing or combining solvers capable of adaptive switching and termination, all while selecting the appropriate regularization parameters and/or formulation [4, 20, 9]. Figuring out when to stop or switch solvers depends on a key connection to the error in how the problem is discretized. A posteriori error analysis stands as a powerful tool for addressing these tasks [28, 1, 2]. It not only guides the entire nonlinear and regularization process but also certifies errors [15]. A comprehensive review of controlling both linear and nonlinear solvers through adaptive stopping criteria can be found in [42, 38, 5] and the references therein. When evaluating the effectiveness of a posteriori error estimators, key attributes to consider include reliability (upper bound for error) and efficiency (lower bound for error). Ideally, the constants in the upper bounds should be explicit, independent of PDE data, and finite-dimensional approximation parameters. The requirement of having a constant-free upper bound, coupled with a lower bound featuring constants ideally independent of model parameters, is particularly appealing for adaptively balancing error sources.

Building upon prior contributions, we extend our focus to adaptive solvers tailored for addressing nonlinear doubly degenerate elliptic equations within conforming methods. The proposed solver dynamically alternates between two linearization strategies: the linearly and globally L-scheme linearization and the quadratic and locally convergent Newton's method. Notably, in instances where Newton's method struggles with non-convergence, the L-scheme is enlisted to initiate the linearization process until an adequately precise starting point for Newton's method is attained. This decision is contingent on achieving a balance between spatial and nonlinear error components. Specifically, the L-scheme persists until the linearization error estimator drops below the spatial error estimator or becomes sufficiently small in comparison. Once this criterion is met, the solver transitions to the Newton method. Furthermore, the choice of the regularization parameter, final stopping criteria, and adaptive grid refinement/coarsening is informed by a posteriori indicators derived from local flux reconstructions and saturation post-processing.

This paper is structured as follows: In Section 2, we introduce essential notation, outline the finite element spaces employed, and describe the nonlinear degenerate elliptic model. Section 3 is focused on presenting the discrete linearized scheme. Therein, we discuss various linearization methods and introduce our adaptive solver, built through the formulation of a posteriori error estimators and the design of balancing criteria. Section 4 is dedicated to the construction of the a posteriori error estimators, leveraging two error measure based on the dual norm of the residual. We first introduce the residual and energy norms used to evaluate the distance between the approximate solution and the exact solution and demonstrate equivalence between the error and dual norm of the residual. Subsequently, we define the local a posteriori estimators,

which rely on the calculated equilibrated flux and post-processed saturation. We also furnish proof of the guaranteed upper bound, local efficiency, and robustness (with respect to the strength of the nonlinearities) of our a posteriori estimate. Finally, in Section 5, we validate our adaptive solver through numerical experiments involving four illustrative examples.

2 Model problem

2.1 Sobolev spaces and notations

Let $\Omega \subset \mathbb{R}^d$, with $d \in \{2,3\}$, represent an open, bounded, and connected domain. We define $H^1(\Omega)$ as the standard Sobolev space of $L^2(\Omega)$ functions with weak gradients in $\left[L^2(\Omega)\right]^d$, and $V := H^1_0(\Omega)$ as its subspace consisting of functions with zero trace. The space $\mathbf{H}(\mathrm{div})$ denotes the set of $\left[L^2(\Omega)\right]^d$ functions with weak divergences in $L^2(\Omega)$. The symbols ∇ and $\nabla\cdot$ (or div) respectively represent the weak gradient and divergence operators in Ω. For any subdomain ω of Ω, we use $(\cdot,\cdot)_\omega$ to denote the L^2-inner product, $||\cdot||_\omega$ to indicate the corresponding norm (with the index omitted when $\omega := \Omega$), and $|\omega|$ to represent the Lebesgue measure on ω. Furthermore, we define $H^{-1}(\omega)$ as the topological dual space of $H^1_0(\omega)$, and we express the duality bracket between these spaces as $\left\langle v, w\right\rangle_{-1,1,\omega}$, which reduces to the L^2-inner product if v, $w \in L^2(\omega)$. When $\omega = \Omega$, the duality bracket is denoted simply as $\left\langle \cdot\ ,\ \cdot\right\rangle_{-1,1}$.

2.2 Partition of Ω and finite element spaces

Let $\mathcal{T}_h$ be a family of closed simplices, such that $\overline{\Omega} = \bigcup_{K\in\mathcal{T}_h} K$. The partition is assumed to be conforming, meaning that if K and L are in $\mathcal{T}_h$, with $K \neq L$, then $K \cap L$ is either an empty set, a common face, edge, or vertex of both K and L. We denote the set of all faces of $\mathcal{T}_h$ by $\mathcal{F}_h$, which is composed of $\mathcal{F}_h^{\mathrm{int}} \cup \mathcal{F}_h^{\mathrm{ext}}$, where $\mathcal{F}_h^{\mathrm{int}}$ and $\mathcal{F}_h^{\mathrm{ext}}$ represent the set of internal and external (boundary) faces, respectively. The faces of an element K are collected in the set $\mathcal{F}_K$. The diameter of an element K is denoted by h_K, and the diameter of a face $\sigma \in \mathcal{F}_h$ is denoted by h_σ. We set $h = \max_{K\in\mathcal{T}_h} h_K$. The set of vertices of the mesh $\mathcal{T}_h$ is denoted by $\mathcal{V}_h$, with $\mathcal{V}_h^{\mathrm{int}}$ representing the set of interior vertices and $\mathcal{V}_h^{\mathrm{ext}}$ for the set of boundary vertices. For each $\mathbf{a} \in \mathcal{V}_h$, we use $\mathcal{T}_h^{\mathbf{a}}$ to denote the patch of the vertex $\mathbf{a}$, i.e., all the elements $K \in \mathcal{T}_h$ that share the vertex $\mathbf{a}$. We denote by $\omega_{\mathbf{a}}$ the corresponding open subset of Ω with diameter $h_{\omega_{\mathbf{a}}}$. Additionally, we define $\mathcal{F}_{\mathbf{a}} \subset \mathcal{F}_h$ as the set of all faces in $\mathcal{T}_h^{\mathbf{a}}$, with $\mathcal{F}_{\mathbf{a}}^{\mathrm{int}}$ and $\mathcal{F}_{\mathbf{a}}^{\mathrm{ext}}$ representing its interior and boundary faces, respectively. we define $\mathcal{F}_{\mathbf{a}}^{\partial\Omega} = \mathcal{F}_{\mathbf{a}} \bigcap \mathcal{F}_h^{\mathrm{ext}}$. For each element $K \in \mathcal{T}_h$, we collect in $\mathcal{V}_K$ the set of vertices belonging to K from $\mathcal{V}_h$. Furthermore, for an element $K \in \mathcal{T}_h$, and for any edge $\sigma \in \partial K$, we define $\mathbf{n}_\sigma$ as the unit normal vector on σ pointing outside K.

Let $\mathbb{P}_p(K)$ represent the space of polynomials on $K \subset \mathcal{T}_h$ of total degree less than or equal to p. The set $\mathbb{P}_p(\mathcal{T}_h)$ denotes the corresponding space of piecewise p-degree polynomials on $\mathcal{T}_h$, meaning $\mathbb{P}_p(\mathcal{T}_h) = \{p_h \in L^2(\Omega);\, p_h|_K \in \mathbb{P}_p(K),\, \forall K \in \mathcal{T}_h\}$. Additionally, we consider the piecewise Raviart–Thomas–Nédélec space $\mathbf{RTN}_p(\mathcal{T}_h) \subset L^2(\Omega)$, defined by $\mathbf{RTN}p(\mathcal{T}_h) = \{v_h \in \mathbf{H}(\mathrm{div});\, v_h|_K \in \mathbf{RTN}_p(K),\, \forall K \in \mathcal{T}_h\}$, where $\mathbf{RTN}_p(K) = [\mathbb{P}_p(K)]^d + \mathbb{P}_p(K)\mathbf{x}$. Let $\Pi_h : L^2(\Omega) \to \mathbb{P}_p(\mathcal{T}_h)$ denote the L^2-orthogonal projection operator onto $\mathbb{P}_p(\mathcal{T}_h)$ or $(\mathbb{P}_p(\mathcal{T}_h;\mathbb{R}^d))$. We will also need $\Pi_h^{\mathrm{RTN}} : L^2(\Omega) \to \mathbf{RTN}_p(\mathcal{T}_h)$ denote the L^2-orthogonal projection operator onto $\mathbf{RTN}_p(\mathcal{T}_h)$. For a vertex $\mathbf{a} \in \mathcal{V}_h$, we introduce the hat operator $\psi_{\mathbf{a}} \in \mathbb{P}_1(\mathcal{T}_h) \cap H^1(\Omega)$ such that $\psi_{\mathbf{a}}(\mathbf{a}) = 1$ and is zero otherwise (at other vertices), thus, $\mathrm{supp}(\psi_{\mathbf{a}}) = \mathcal{T}_h^{\mathbf{a}}$. For each $\mathbf{a} \in \mathcal{V}_h$, $\mathbf{RTN}_p(\mathcal{T}_h^{\mathbf{a}})$ and $\mathbb{P}_p(\mathcal{T}_h^{\mathbf{a}})$ denote the restriction of $\mathbf{RTN}_p(\mathcal{T}_h)$ and $\mathbb{P}_p(\mathcal{T}_h)$ to the patch $\mathcal{T}_h^{\mathbf{a}}$, respectively. We also introduce $V_h := \mathbb{P}_p(\mathcal{T}_h) \cap H^1_0(\Omega)$ and let $Q_h := \mathbb{P}_p(\mathcal{T}_h)$ and $\mathbf{W}_h := \mathbf{RTN}_p(\mathcal{T}_h)$.

2.3 Model problem and variational formulation

We consider in this paper the following nonlinear and doubly degenerate elliptic equation as the basis of our discussion [13, 24]: find u such that

$$\beta(u) - \Delta\varphi(u) = f + \nabla\cdot\boldsymbol{\xi}, \quad \text{in } \Omega, \tag{2.1a}$$

$$u(\cdot,0) = u_0, \quad \text{in } \Omega, \tag{2.1b}$$

$$\varphi(u) = 0, \quad \text{on } \partial\Omega. \tag{2.1c}$$

As a preliminary step towards establishing the weak formulation of (2.1), we address the necessary assumptions that must hold:

(**A1**) $\varphi : \mathbb{R} \to \mathbb{R}$ is continuous and non-decreasing, further $\varphi(0) = 0$, and there exist M_0, $M_1 > 0$ such that $|\varphi(s)| \geq M_0|s| - M_1$, for all $s \in \mathbb{R}$.

(**A2**) $\beta : \mathbb{R} \to \mathbb{R}$ is continuous non-decreasing, and $\beta(0) = 0$, and there exist K_0, $K_1 > 0$ such that $|\beta(s)| \leq K_0|s| + K_1$, for all $s \in \mathbb{R}$.

(**A3**) $\varphi + \beta : \mathbb{R} \to \mathbb{R}$ is strictly increasing.

(**A4**) The source terms are such that $f \in L^2(\Omega)$ and $\boldsymbol{\xi} \in (L^2(\Omega))^d$.

The superlinearity of $\varphi(u)$ in (**A1**) ensures that u belongs to $L^2(\Omega)$ (at least). Thus, the sub-linearity of β assumed by (**A2**) implies that $\beta(u) \in L^2(\Omega)$ as well. That, assumption (**A4**) implies that the resulting source term $f + \nabla\cdot\boldsymbol{\xi}$ should be understood in the sense of distribution [19, 31]. Note that further generalizations are possible, bringing more technicalities. Assume (**A1**)–(**A4**) hold true, we introduce the weak solution to problem (2.1): find $u \in L^2(\Omega)$ with $\varphi(u) \in V$ such that

$$(\beta(u), v) + (\nabla\varphi(u), \nabla v) = (f, v) - (\boldsymbol{\xi}, \nabla v)\,, \quad \forall v \in V. \tag{2.2}$$

The theoretical aspects of (2.2) can be found in the seminal paper [14]. The existence and uniqueness of the weak solution to (2.2) are established in [24], where detailed regularity properties are also presented (see also [27]). Still following [25], one can opt for a reformulation of 2.2 using a new variable, say w, to serve as the primary unknown in the new system, in which Lipschitz-continuous counterparts of φ and β are obtained. Under Assumption (**A1**)–(**A4**), we observe that $\beta + \varphi : \mathbb{R} \to \mathbb{R}$ is a bijective function. This allows us to define $\mu(s) = \zeta\left((\beta + \zeta)^{-1}(s)\right)$ and $\rho(t) = t - \mu(t) = \beta\left((\beta + \zeta)^{-1}(t)\right)$. These functions are non-decreasing and have 1-Lipschitz continuity (see [24]). Problem 2.2 can be equivalently expressed as finding $w \in L^2(\Omega)$ with $\rho(w) \in V$ such that

$$(\mu(w), v) + (\nabla\rho(w), \nabla v) = (f, v) - (\boldsymbol{\xi}, \nabla v)\,, \quad \forall v \in V. \tag{2.3}$$

This reformulation brings advantages for a priori error analysis in the context of conforming methods. This is mainly because it provides more flexibility in using the chain rule and choosing suitable test functions compared to 2.2. Specifically, the Lipshitz continuity of μ and ρ makes formulation 2.2 more suitable for conforming methods. However, it's worth noting that computing the inverse transform $w \longrightarrow u$ can be computationally demanding, and w doesn't have a direct physical interpretation. The equivalence between the formulations comes from the uniqueness of solutions, making both applicable for our a posteriori error analysis. For a detailed exploration of the strengths and weaknesses in these two formulations, along with the numerical approaches applied to (2.1), we refer the reader to [24]. This work introduces the Gradient Discretisation Method (GDM), a generic numerical analysis framework enabling a unified convergence analysis of different discretisations [25].

3 Discrete setting and adaptive algorithm

In this section, we provide the discretized, linearized, and regularized version of the variational formulation (2.2). We also detail the methodology and algorithm utilized to solve this formulation.

3.1 Discretization

To illustrate our framework, we present a conforming Galerkin method applied to (2.2). We aim to find $u_h \in V_h$ satisfying the following discrete formulation:

$$(\beta(u_h), v) + (\nabla\varphi(u_h), \nabla v) \quad = (f, v) - (\boldsymbol{\xi}, \nabla v)\,, \quad \forall v \in V_h. \tag{3.1}$$

Yet, due to the necessity of mass-lumping to ensure a monotone scheme [37, 49], there is limited interest in employing high-order methods for 3.1. However, one can still rely on the Gradient Discretisation Method

(GDM) framework corresponding to MLFE to ensure higher-order convergence, as detailed in [26]. Furthermore, to derive any form of a posteriori error estimate from these conforming schemes, we are constrained to assume $\varphi(u_h) \in V_h$, a condition that cannot be guaranteed by the above formulation due to the lack of global Lipschitz continuity in φ. To address this, we may use the discrete counterpart of (2.3), which provide the necessary $V-$conformity for $\rho(w_h)$ and, by equivalence, for $\varphi(u_h)$. We shall later comment on the implications of each formulation on the a posteriori error analysis. Nevertheless, we maintain the generality of our presentation and analysis. Our framework covers conforming schemes where the discrete solution $u_h \in L^2(\Omega)$ is such that $\varphi(u_h) \in V$. Examples of methods fitting into our analysis include conforming Finite Element methods (FE), including the Mass-Lumped version (MLFE) and Vertex Approximate Gradient (VAG) methods [25].

3.2 Linearization

The primary challenge in addressing the non-linear and doubly degenerate problem (3.1) lies in devising a robust linearization scheme that converges, particularly in cases where $\varphi'(\cdot)$ and/or $\beta'(\cdot)$ approach(es) zero. Equally significant is the scenario where $\varphi'(\cdot)$ and/or $\beta'(\cdot)$ become(es) unbounded. This is of important interest due to its potential impact on the convergence properties of the solution method. Let us assume a linearization process is adopted to deal with nonlinearities within (3.1): At the iteration $k \geq 1$, find $u_h^k \in V_h$ such that

$$\left(\beta^{k-1}(u_h^k), v\right) + \left(\nabla\varphi^{k-1}(u_h^k), \nabla v\right) = (f, v) - (\boldsymbol{\xi}, \nabla v), \quad \forall v \in V_h, \tag{3.2}$$

where the functions φ^{k-1} and β^{k-1} are respectively linear approximations of $\varphi(u_h)$ and $\beta(u_h)$. If Newton–Raphson's method is used, these approximations are of the form:

$$\zeta^{k-1}(u_h^k) := \zeta'(u_h^{k-1})\left(u_h^k - u_h^{k-1}\right) + \zeta(u_h^{k-1}), \quad k \geq 1. \tag{3.3a}$$

Although, Newton–Raphson's method is quadratically convergent, it is only locally convergent and involves the computation of derivatives; when $\zeta(\cdot)$ degenerates, the method might fail to converge (algebraic residuals are either stagnating, diverging or oscillating) and a remedy based on relaxation techniques is needed. We might also need smoothing functions of $\varphi(\cdot)$ and $\beta(\cdot)$.

An alternative to Newton-Raphson's method is to consider a quasi-Newton version that replaces the exact derivative ζ' with an approximation. For instance, we adopt the so-called L-scheme, a quasi-Newton method which lies between fixed-point and Newton solvers [50, 51, 58], in which the approximations in (3.2) are simply given by

$$\zeta^{k-1}(u_h^k) := L\left(u_h^k - u_h^{k-1}\right) + \zeta(u_h^{k-1}), \quad k \geq 1, \tag{3.3b}$$

with $L > 0$ is a constant mimicking the derivative of ζ. The L-scheme is linearly convergent assuming Lipschitz continuity of $\zeta(\cdot)$. The only requirement for the convergence of the L-scheme is that the parameter L should satisfy $L \geq L_\zeta := \sup_s |\zeta'(s)|$. It has been shown that the convergence can also be obtained for smaller values of L, e.g. for $L = L_\zeta/2$ convergence is guaranteed (see [47]) while the contraction proof can not be ensured in this case. Also, one can choose $L = \max_s |\zeta'(s)|)$ and that's the classical Picard iterations (see [45] for details). Therefore, the L-scheme is robust when $\zeta'(\cdot) \approx 0$, however, when $\zeta'(\cdot)$ becomes unbounded (ζ is not Lipschitz), a regularization procedure is also needed. We refer to [57] for general sub-classes of quasi-Newton methods and particularly to [46, 6, 43] for combinations of Newton and L-scheme or Picard iterations.

3.3 Regularization

The regularization approach involves introducing a smoothing function, denoted as ζ_ϵ, for a parameter $\epsilon > 0$. This smoothing function $\zeta_\epsilon(\cdot) : \mathbb{R} \to \mathbb{R}$ is monotone and Lipschitz continuous,

$$0 \leq |\zeta_\epsilon'(s)| \leq L_{\zeta_\epsilon} < \infty, \ \forall s \in \mathbb{R}. \tag{3.4}$$

Let u_h^ϵ be the regularized solution, defined by the recursive sequence $u_h^{\epsilon,k} \in V_h$, $\forall k \geq 1$, such that

$$\left(\beta_\epsilon^{k-1}(u_h^{\epsilon,k}), v\right) + \left(\nabla\varphi_\epsilon^{k-1}(u_h^{\epsilon,k}), \nabla v\right) = (f, v) - (\boldsymbol{\xi}, \nabla v)\,, \quad \forall v \in V_h. \tag{3.5}$$

Although this regularization procedure could prove effective for an L-scheme, strict monotonicity might be a necessary condition for the application of Newton's method. Furthermore, if the L-scheme is employed for the alternative formulation 2.3, regularization may not be necessary at all. It is evident that developing an efficient solver that maintains a balance between robustness and efficiency while addressing all nonlinear aspects of the problem necessitates combining nonlinear solvers and/or formulations. This can be achieved through the so-called adaptive *solver/variable switching* [53] or parameterization [9], enabling us to leverage the strengths of each method or formulation. In our context, we will focus exclusively on adaptive solver switching to optimize the performance of the standard nonlinear solver tailored for our specific problem.

3.4 Adaptive algorithm

The purpose of this section is to reduce as much as possible the computational effort to solve (3.5). The estimation and localization of distinct error components lead to the development of an adaptive nonlinear solver, enabling optimal balancing of different error contributions. In doing so, one in particular can supervise Newton iterations with an adaptive switch back-and-forth to a slower but safer algorithm, if a convergence failure is detected.

3.4.1 Balancing criteria

Let $\eta_{\text{disc}}^{k,\epsilon}$, $\eta_{\text{reg}}^{k,\epsilon}$, $\eta_{\text{qd}}^{k,\epsilon}$, $\eta_{\text{osc}}^{k,\epsilon}$ and $\eta_{\text{lin}}^{k,\epsilon}$, be respectively fully computable a posteriori estimators of *discretization error*, *regularization error*, *quadrature error*, *data-oscillation error*, and *linearization error* at the k-th iteration, such that

$$\textit{error}(u, u_h^{\epsilon,k}) \leq \eta_{\text{disc}}^{k,\epsilon} + \eta_{\text{reg}}^{k,\epsilon} + \eta_{\text{qd}}^{k,\epsilon} + \eta_{\text{osc}}^{k,\epsilon} + \eta_{\text{lin}}^{k,\epsilon}, \tag{3.6}$$

where $\textit{error}(u, u_h^{\epsilon,k})$ represents an energy norm, and its definition will be provided in a posteriori error analysis in Section 4.1. To minimize efficiently the right-hand side of (3.6), we adopt suitable strategy for each estimator so that each of will looks forward the smallest possible cost to himself, and such that the ideal estimate is obtained. The strategy we adopt for $\eta_{\text{lin}}^{k,\epsilon}$ is to use an *adaptive switching* from the L-scheme to Newton-Raphson's method, i.e., we switch from the L-scheme to Newton iteration when the following criteria is satisfied

$$\eta_{\text{lin}}^{k,\epsilon} \leq \Gamma_{\text{sw}} \left(\eta_{\text{disc}}^{k,\epsilon} + \eta_{\text{reg}}^{k,\epsilon} + \eta_{\text{qd}}^{k,\epsilon} + \eta_{\text{osc}}^{k,\epsilon}\right), \tag{3.7}$$

where $\Gamma_{\text{sw}} < 1$ is a user-given weight to be chosen adaptively. The use of (3.7) implies that the L-scheme is pursued until the L-scheme error approaches enough from the total error, which hopefully is sufficiently good (but perhaps not!) to (re)-launch the Newton solver. This later is then continued until the following stopping criteria

$$\eta_{\text{lin}}^{k,\epsilon} \leq \Gamma_{\text{lin}} \left(\eta_{\text{disc}}^{k,\epsilon} + \eta_{\text{reg}}^{k,\epsilon} + \eta_{\text{qd}}^{k,\epsilon} + \eta_{\text{osc}}^{k,\epsilon}\right), \tag{3.8}$$

is thrown. The parameter Γ_{lin} is to be given in $(0, \Gamma_{\text{sw}})$. If the L-scheme fails to deliver a good initial guess for Newton, the L-scheme will be pursued either with smaller parameters Γ_{sw} or until the stopping criteria (3.8) is satisfied. The strategy we adopt for minimizing $\eta_{\text{reg}}^{k,\epsilon}$ is based on reducing adaptively the regularization parameter ϵ so that

$$\eta_{\text{reg}}^{k,\epsilon} \leq \Gamma_{\text{reg}} \left(\eta_{\text{disc}}^{k,\epsilon} + \eta_{\text{qd}}^{k,\epsilon} + \eta_{\text{osc}}^{k,\epsilon}\right), \tag{3.9}$$

with a user-given parameter $\Gamma_{\text{reg}} \in (0, 1)$. The strategy we adopt for $\eta_{\text{disc}}^{k,\epsilon}$ is to adapt the space mesh; refinement/coarsening so that the local discretization error estimators are distributed equally: for all K_1, $K_2 \in \mathcal{T}_h$,

$$\eta_{\text{disc},K_1}^{k,\epsilon} \approx \eta_{\text{disc},K_2}^{k,\epsilon}. \tag{3.10}$$

To address restrictions in computing resources, we establish a fixed refinement threshold $h_{\min} > 0$ for the mesh size. This involves ensuring that

$$\min_{K \in \mathcal{T}_h} h_K \geq h_{\min}.$$

We also assume that the adaptive mesh is shape regular, i.e., that there exists a constant κ_T such that, for all $K \in \mathcal{T}_h$,

$$\max_{K\in\mathcal{T}_h} \frac{h_K}{\rho_K} \leq \kappa_T,$$

where h_K is the diameter of K, and ρ_K is the radius of the largest inscribed ball of K.

3.4.2 Algorithm

Recall that $u_h^{\epsilon,k}$ is an approximation of the solution u obtained after k linearization iterations using a regularization parameter ϵ. Our algorithm, incorporating the aforementioned balancing criteria, is as follows:

Algorithm 3.1: The adaptive algorithm

---

(1) Choose an initial approximation $u_h^{\epsilon,-1} \in V_h$ of u_h, switching parameters $(\Gamma_{\text{sw}}, \Gamma_{\text{lin}}, \Gamma_{\text{reg}})$, and a regularization parameter $\epsilon > 0$. Set $k = -1$.

(2) **Do** {**Spatial adaptivity**}

(I) **Do** {**Regularization**}

(A) Set $k = -1$.

(B) **Do** {**Newton-Raphson**}

(i) Increase $k = k + 1$.

(ii) Compute $u_h^{\epsilon,k}$ using (3.3a) and (3.5).

(iii) Compute the estimators $(\eta_{\text{disc}}^{k,\epsilon},\ \eta_{\text{lin}}^{k,\epsilon},\ \eta_{\text{reg}}^{k,\epsilon},\ \eta_{\text{qd}}^{k,\epsilon},\ \eta_{\text{osc}}^{k,\epsilon})$.

(iv) **if** $\eta_{\text{lin}}^{k,\epsilon}$ increases, stagnates, or oscillates. `/*Switching solver*/`

(a) Reset $k = -1$.

(b) **Do** {**Quasi-Newton**}

(1) Increase $k = k + 1$.

(2) Compute $u_h^{\epsilon,k}$ using (3.3b) and (3.5).

(3) Compute the estimators $(\eta_{\text{disc}}^{k,\epsilon},\ \eta_{\text{lin}}^{k,\epsilon},\ \eta_{\text{reg}}^{k,\epsilon},\ \eta_{\text{qd}}^{k,\epsilon},\ \eta_{\text{osc}}^{k,\epsilon})$.

While (3.7) or (3.8) is not satisfied.

End

While (3.8) is not satisfied. `/*Stopping solver*/`

(C) **if** (3.9) is not satisfied.

(i) Replace ϵ by $\epsilon/2$.

End

While (3.9) is not satisfied.

(II) Refine or coarsen the cells $K \in \mathcal{T}_h$ until (3.10) is satisfied or $h_K < h_{\min}$.

While (3.10) is not satisfied.

Remark 3.1 (Offline L calculation)**.** *Given that the classical stabilization parameters, such as L_φ or $\frac{L_\varphi}{2}$, may not always yield optimal results, an enhancement to the adaptive algorithm 3.1 can be achieved through a 'brute optimization' process (see [54]). This process aims to determine the optimal parameters $(L_{\beta_\epsilon}^{opt}, L_{\varphi_\epsilon}^{opt})$ by solving the optimization problem:*

$$(L_{\beta_\epsilon}^{opt}, L_{\varphi_\epsilon}^{opt}) := \underset{a\leq L_\zeta \leq b, \zeta\in\{\varphi,\beta\}}{arg\,min} \eta_{lin}^{k,\epsilon}(L_\beta, L_\varphi; \mathcal{T}_H, \epsilon) \tag{3.11}$$

This optimization is performed over a coarse grid $\mathcal{T}_H$ and for a sufficiently small regularization parameter ϵ. The advantages lie in the discretization-error-independent nature of $\eta_{lin}^{k,\epsilon}$ and the almost regularization-independent performance of the L-scheme.

Remark 3.2 (Localized balancing criteria)**.** *We can express localized versions of the criteria on an element-by-element basis using local estimators, ensuring their validity for all $K \in \mathcal{T}_h$. The switching criteria* (3.7) *can also be tailored to balance linearization and spatial errors:*

$$\gamma_{\mathrm{sw}} \left(\sum_{\bullet} \eta_{\bullet}^{k,\epsilon} \right) \leq \eta_{\mathrm{lin}}^{k,\epsilon} \leq \Gamma_{\mathrm{sw}} \left(\sum_{\bullet} \eta_{\bullet}^{k,\epsilon} \right), \tag{3.12}$$

where Γ_{sw} and γ_{sw} are set close to one. Balancing linearization and spatial errors, especially with a small $\eta_{\mathrm{reg}}^{k,\epsilon}$ through (3.9)*, prevents specific errors from dominating. This delivers a smoother starting point for Newton, avoiding oscillations and promoting stability.*

Remark 3.3 (Adaptive switching)**.** *The switching parameter Γ_{sw} is to be chosen smaller if quasi-Newton's method fails to provide a good initialization for Newton's method.*

Remark 3.4 (Descent methods)**.** *The adaptive algorithm is valid if we replace the L-scheme with damped Newton's methods or any other descent solver with explicit formula of the step size [43]. The damping parameter can be computed in the offline phase. For optimal descent methods, we must constitute a line search rule to infer the ideal step size at each iteration.*

4 Energy-type a posteriori error estimates

In the present work, we adopt energy-type a posteriori error estimates [1, 52, 32]. Classically, the derivation of energy a posteriori error estimates is based on the definition the residual (and its dual norm) and the mathematical link between the energy norm of the error and the dual norm of the residual.

4.1 Residuals and energy norms

Following [32], we introduce the residual operator $\mathcal{R}_u(u_h) \in H^{-1}(\Omega)$, defined by

$$\langle \mathcal{R}_u(u_h), v \rangle := (\beta(u) - \beta(u_h), v) + \big(\nabla\big(\varphi(u) - \varphi(u_h)\big), \nabla v\big), \quad \forall v \in V. \tag{4.1}$$

Next, we let the parameter $\delta_\star$ be equal to 0 or 1, depending on the norm in which the error is measured. We then equip the V space with the energy norm

$$\|\psi\|_\star := \left(\|\nabla\psi\|^2 + \delta_\star \|\psi\|^2\right)^{\frac{1}{2}}, \tag{4.2}$$

where clearly the natural choice being $\|\psi\|_\star := \|\nabla\psi\|$ with $\delta_\star = 0$. The dual norm of the residual is then given through

$$|||\mathcal{R}_u(u_h)|||_\star := \sup_{\substack{\psi \in H_0^1(\Omega) \\ \|\psi\|_\star = 1}} \langle \mathcal{R}_u(u_h), \psi \rangle, \tag{4.3}$$

which expresses by how much the trial function u_h fails to satisfy (2.2). We use simply $|||\mathcal{R}(u_h)|||$ for $\delta_\star = 0$. Note that, in the case of $\varphi(u) := u$ and $\beta(u) := u + g$, with given $g \in L^2(\Omega)$, that is, when (2.2) is linear and strongly elliptic, the norm $\|\cdot\|_\star$ with $\delta_\star = 1$ is a natural energy norm on $H_0^1(\Omega)$. Remark also that the chosen $H_0^1(\Omega)$-norm delivers different continuity estimates for the equation (2.2). The sharpness of the continuity estimate is in fact of particular interest for designing goal-oriented error estimation. For energy-error estimation, the choice of the $H_0^1(\Omega)$-norm affects the dual norm of the residual and leads to different error measures. Evidently, the dual norm of the residual $|||\mathcal{R}_u(u_h)|||_\star$ vanishes for $u_h = u$ (see Corollary (4.4) for details), and this is when $\delta_\star = 0$ or 1, but its actual size is unknown and its value is incomputable. Thus, the link between these dual norms and the underlying error energy norms is necessary. To show this link, we introduce the energy norm

$$||u - u_h||_{\#}^2 = ||\beta(u) - \beta(u_h)||_{H^{-1}(\Omega)}^2 + ||\nabla\left(\varphi(u) - \varphi(u_h)\right)||^2. \tag{4.4}$$

We also introduce the weaker distance

$$\mathcal{J}_u(u_h) := 2\left(\beta(u) - \beta(u_h),\, \varphi(u) - \varphi(u_h)\right). \tag{4.5}$$

We have this result.

Proposition 4.1 (Residual-Energy norm)**.** *Let u be the weak formulation of* (2.2) *with V equipped with the energy-norm* (4.2) *and let $u_h \in L^2(\Omega)$ be an arbitrary function such that $\varphi(u_h) \in V$. Then, if $\delta_\star = 0$, there holds*

$$\mathcal{J}_u(u_h) + ||u - u_h||_{\#}^2 = |||\mathcal{R}_u(u_h)|||^2. \tag{4.6}$$

For $\delta_\star = 1$, define the metric

$$||u - u_h||_{\Sigma,\star}^2 := ||u - u_h||_{\#}^2 + ||\varphi(u) - \varphi(u_h)||_{H^{-1}(\Omega)}^2 - ||\varphi(u) - \varphi(u_h)||^2. \tag{4.7}$$

Then, there holds,

$$||u - u_h||_{\Sigma,\star}^2 = |||\mathcal{R}_u(u_h)|||_\star^2. \tag{4.8}$$

Proof. To obtain the above relations, we use duality techniques as in [32, 10]. Utilizing Assumption (**A2**), which ensures $\beta(u)$ and $\beta(u_h)$ are in $L^2(\Omega)$, we proceed by letting $\psi \in V$ be the solution to the auxiliary problem:

$$(\nabla\psi, \nabla v) = (\beta(u) - \beta(u_h), v)\,, \quad \forall v \in V. \tag{4.9}$$

Clearly, we have

$$||\nabla\psi|| = ||\beta(u) - \beta(u_h)||_{H^{-1}(\Omega)}. \tag{4.10}$$

We replace (4.9) in (4.1) to get

$$\langle \mathcal{R}_u(u_h), v\rangle = (\nabla\left(\psi + \varphi(u) - \varphi(u_h)\right), \nabla v)\,, \quad \forall v \in V. \tag{4.11}$$

so that from (4.3),

$$|||\mathcal{R}_u(u_h)||| = ||\nabla\left(\psi + \varphi(u) - \varphi(u_h)\right)||. \tag{4.12}$$

By expanding the square, we obtain

$$|||\mathcal{R}_u(u_h)|||^2 = ||\nabla\psi||^2 + ||\nabla\left(\varphi(u) - \varphi(u_h)\right)||^2 + 2(\nabla\psi, \nabla\left(\varphi(u) - \varphi(u_h)\right)). \tag{4.13}$$

We replace the first term using (4.10) as well as we make use of (4.9) to get (4.6).
For the second identity (4.8), we let $\psi \in V$ be the solution of the auxiliary problem

$$(\nabla\psi, \nabla v) + (\psi, v) = (\beta(u) - \beta(u_h) - \varphi(u) + \varphi(u_h), v)\,, \quad \forall v \in V. \tag{4.14}$$

We have first

$$\|\psi\|_\star = ||\beta(u) - \beta(u_h) - \varphi(u) + \varphi(u_h)||_{H^{-1}(\Omega)}. \tag{4.15}$$

Then it holds

$$\langle \mathcal{R}_u(u_h), v\rangle = (\psi + \varphi(u) - \varphi(u_h), v) + (\nabla\left(\psi + \varphi(u) - \varphi(u_h)\right), \nabla v)\,, \quad \forall v \in V.$$

Thus, also

$$|||\mathcal{R}_u(u_h)|||_\star^2 = \|\psi + \varphi(u) - \varphi(u_h)\|_\star^2.$$

Expanding the terms, utilizing the definitions of ψ and d_Σ, we finally obtain

$$\begin{aligned}
|||\mathcal{R}_u(u_h)|||_\star^2 &= \|\psi\|_\star^2 + \|\varphi(u) - \varphi(u_h)\|_\star^2 + 2\left(\psi, \varphi(u) - \varphi(u_h)\right) + 2\left(\nabla\psi, \nabla\left(\varphi(u) - \varphi(u_h)\right)\right) \\
&= \|\psi\|_\star^2 + \|\varphi(u) - \varphi(u_h)\|_\star^2 + 2\left(\beta(u) - \beta(u_h), \varphi(u) - \varphi(u_h)\right) - 2\,\|\varphi(u) - \varphi(u_h)\|^2 \\
&= ||u - u_h||_{\Sigma,\star}^2.
\end{aligned}$$

This completes the proof. □

Remark 4.2 (About $\mathcal{J}_u(u_h)$)**.** *Note that $\mathcal{J}_u(u_h)$ is positive due to the monotonicity of β and φ. As Arbogast [3] points out, the nonlinear form $\mathcal{J}_u(u_h)$ tells us something about the error of the scheme [56, 10]. Particularly, if $\beta = \mathbf{I}$ and φ is Lipschitz-continuous with Lipschitz constant L_φ, then,*

$$\frac{1}{L_\varphi}||\varphi(u) - \varphi(u_h)||^2 \leq \mathcal{J}_u(u_h). \tag{4.16}$$

Remark 4.3 (On norm(u, u_h))**.** *Note that in the linear case, as indicated by [52], we have:*

$$||u - u_h||^2 + ||\nabla\,(u - u_h)\,||^2 = |||\mathcal{R}_u(u_h)|||_\star^2, \quad \forall u,\, u_h \in V. \tag{4.17}$$

For the general nonlinear case, there holds

$$||u - u_h||_{\Sigma,\star}^2 \leq ||u - u_h||_\#^2 \leq ||\beta(u) - \beta(u_h)||^2 + ||\nabla\,(\varphi(u) - \varphi(u_h))\,||^2, \quad \forall u,\, u_h \in V. \tag{4.18}$$

Corollary 4.4 (On the weak solution)**.** *There exists one solution to problem* (2.2)*. Furthermore, there holds*

$$\delta_\star||u||_{\Sigma,\star} + (1 - \delta_\star)\left\{\ \left\{\mathcal{J}_0(u) + ||u||_\#^2\right\}^{\frac{1}{2}}\ \right\} = \sup_{\substack{v \in H_0^1(\Omega) \\ \|v\|_\star = 1}} \{(f, v) - (\boldsymbol{\xi}, \nabla v)\}. \tag{4.19}$$

Proof. The existence of a solution is a consequence of [24, Theorem 2.9]. To show the uniqueness of the solution under Assumption (**A1**)–(**A4**), we will see that if $\mathcal{R}(u_h) = 0$, we then have $u_h = u$. Suppose that $\mathcal{R}(u_h) = 0$, then from (4.6) we have immediately that $\varphi(u) - \varphi(u_h) = 0$ and $\beta(u) - \beta(u_h) = 0$. The use of Assumption (**A3**) implies that $u_h = u$ and the uniqueness of the solution. To infer the uniqueness of the solution from the weaker identity (4.8), we consider u, u_h as two solutions of problem (2.2). We then subtract their respective equations and let $v = \varphi(u) - \varphi(u_h)$, to get $\mathcal{J}_u(u_h)/2 + ||\nabla(\varphi(u) - \varphi(u_h))||^2 = 0$. Due to the positivity of $\mathcal{J}_u(u_h)$, it follows that $\varphi(u) - \varphi(u_h) = 0$. The weaker error measure (4.8) also results in $\beta(u) - \beta(u_h) = 0$. Again, the application of Assumption (**A3**) implies that $u_h = u$. Finally, obtaining the identity 4.19 is straightforward, given that $\varphi(0) = \beta(u) = 0$ (and zero is an admissible element of V). We achieve this by substituting (4.3) into (4.6) and (4.8), respectively.

□

4.2 Equilibrated flux $\boldsymbol{\sigma}_h^{\epsilon,k}$ and post-processed saturation $\Phi_h^{\epsilon,k}$

The constructions introduced in this section follows closely [35], in generalization of [52]. Let us introduce for each $\mathbf{a} \in \mathcal{T}_h^{\mathbf{a}}$, the local spatial mixed finite element spaces posed on the patch domain $\omega_{\mathbf{a}}$:

$$Q_h^{\mathbf{a}} := \begin{cases} \mathbb{P}_p(\mathcal{T}_h^{\mathbf{a}}) & \text{if } \delta_\star > 0 \text{ or } \mathbf{a} \in \mathcal{V}_h^{\text{ext}}, \\ \{q \in \mathbb{P}_p(\mathcal{T}_h^{\mathbf{a}}),\ (q,1)_{\omega_{\mathbf{a}}} = 0\} & \text{if } \delta_\star = 0 \text{ and } \mathbf{a} \in \mathcal{V}_h^{\text{int}}, \end{cases} \tag{4.20a}$$

$$\mathbf{W}_h^{\mathbf{a}} := \begin{cases} \{\mathbf{v}_h \in \mathbf{RTN}_p(\mathcal{T}_h^{\mathbf{a}}) \cap \mathbf{H}(\text{div}, \omega_{\mathbf{a}}),\ \mathbf{v}_h \cdot \mathbf{n}_K = 0 \ \text{ on } \partial\omega_{\mathbf{a}} \setminus \partial\Omega\} & \text{if } \mathbf{a} \in \mathcal{V}_h^{\text{ext}}, \\ \{\mathbf{v}_h \in \mathbf{RTN}_p(\mathcal{T}_h^{\mathbf{a}}) \cap \mathbf{H}(\text{div}, \omega_{\mathbf{a}}),\ \mathbf{v}_h \cdot \mathbf{n}_K = 0 \ \text{ on } \partial\omega_{\mathbf{a}}\} & \text{if } \mathbf{a} \in \mathcal{V}_h^{\text{int}}. \end{cases} \tag{4.20b}$$

Definition 4.5 (Notations)**.** *We denote by $\mathbf{I}_h^{\epsilon,k}$ the linearized approximate flux $\nabla\widetilde{\varphi}_\epsilon^{k-1}(u_h^{\epsilon,k})$ within* (3.5)*.*

The constructions given next are based on independent local mixed finite element approximations of residual problems from (3.5) over patches of elements around mesh vertices (See Section (3.2) for notations). As used in the above discrete spaces, the reconstructions will depend on the choice of the error measure. To simplify the notation, we introduce the residual function $\mathbf{r}_h^{\epsilon,k} \in \mathbb{P}_p(\mathcal{T}_h)$, $\mathbf{r}_h^{\epsilon,k}|_{\partial\Omega} = 0$, represented by

$$(\mathbf{r}_h^{\epsilon,k}, v) := (f, v) - (\widetilde{\beta}_\epsilon(u_h^{\epsilon,k}), v) - ((\boldsymbol{\xi} + \nabla\widetilde{\varphi}_\epsilon(u_h^{\epsilon,k})), \nabla v), \quad \forall v \in V_h. \tag{4.21}$$

Then, for a given parameter $\delta_\star = 1$ or 0, the equilibrated flux $\boldsymbol{\sigma}_h^{\epsilon,k}$ and postprocessed saturation $\Phi_h^{\epsilon,k}$ are obtained using the following reconstructions (See Appendix A. for the local minimization problems):

Reconstruction 1 ($(\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, \Phi_{\mathbf{a}}^{\epsilon,k})$)**.** *At each iteration $k \geq 1$, and for each patch $\omega_{\mathbf{a}}$, $\mathbf{a} \in \mathcal{V}_h$, we construct $(\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, r_{\mathbf{a}}^{\epsilon,k}) := (\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, r_{\mathbf{a}}^{\epsilon,k})_{\text{disc}} + (\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, r_{\mathbf{a}}^{\epsilon,k})_{\text{lin}} \in \mathbf{W}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$ such that:*

$$(\boldsymbol{\sigma}_{\mathbf{a},\text{disc}}^{\epsilon,k} + \boldsymbol{\sigma}_{\mathbf{a},\text{lin}}^{\epsilon,k}, \mathbf{v})_{\omega_{\mathbf{a}}} - (\nabla\cdot\mathbf{v}, r_{\mathbf{a}}^{\epsilon,k})_{\omega_{\mathbf{a}}} = (\boldsymbol{\lambda}_{\mathbf{a}}, \mathbf{v})_{\omega_{\mathbf{a}}}, \qquad \forall \mathbf{v} \in \mathbf{W}_h^{\mathbf{a}}, \tag{4.22a}$$

$$(\nabla\cdot(\boldsymbol{\sigma}_{\mathbf{a},\text{disc}}^{\epsilon,k} + \boldsymbol{\sigma}_{\mathbf{a},\text{lin}}^{\epsilon,k}), q)_{\omega_{\mathbf{a}}} + \delta_\star(r_{\mathbf{a}}^{\epsilon,k}, q)_{\omega_{\mathbf{a}}} = (g_{\mathbf{a}}, q)_{\omega_{\mathbf{a}}}, \quad \forall q \in Q_h^{\mathbf{a}}, \tag{4.22b}$$

with

$$g_{\mathbf{a}} := \Pi_h\Big(\psi_{\mathbf{a}}\big(f - \widetilde{\beta}_\epsilon^{k-1}(u_h^{\epsilon,k})\big)\Big)|_{\omega_{\mathbf{a}}} - \nabla\psi_{\mathbf{a}}\cdot(\Pi_h\boldsymbol{\xi} + \mathbf{I}_h^{\epsilon,k})|_{\omega_{\mathbf{a}}}, \text{ and } \boldsymbol{\lambda}_{\mathbf{a}} := -(\psi_{\mathbf{a}}\boldsymbol{\xi} + \psi_{\mathbf{a}}\mathbf{I}_h^{\epsilon,k})|_{\omega_{\mathbf{a}}}.$$

Then, we define the postprocessed saturation by $\Phi_{\mathbf{a}}^{\epsilon,k} = \Phi_{\mathbf{a},\text{disc}}^{\epsilon,k} + \Phi_{\mathbf{a},\text{lin}}^{\epsilon,k} := \delta_\star r_{\mathbf{a}}^{\epsilon,k} + \Pi_h(\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon^{k-1}(u_h^{\epsilon,k}))$ and we extend each $(\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, \Phi_{\mathbf{a}}^{\epsilon,k})$ by zero in $\Omega \setminus \omega_{\mathbf{a}}$, $(\boldsymbol{\sigma}_h^{\epsilon,k}, \Phi_h^{\epsilon,k}) \in \mathbf{RTN}_p(\mathcal{T}_h) \times \mathbb{P}_p(\mathcal{T}_h)$ is given by

$$(\boldsymbol{\sigma}_h^{\epsilon,k}, \Phi_h^{\epsilon,k}) := \sum_{\mathbf{a}\in\mathcal{V}_h} (\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, \Phi_{\mathbf{a}}^{\epsilon,k}). \tag{4.23}$$

Reconstruction 2 ($(\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, \Phi_{\mathbf{a}}^{\epsilon,k})_{\text{disc}}$)**.** *At each iteration $k \geq 1$, and for each patch $\omega_{\mathbf{a}}$, $\mathbf{a} \in \mathcal{V}_h$, construct $(\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, r_{\mathbf{a}}^{\epsilon,k})_{\text{disc}} \in \mathbf{W}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$ such that:*

$$(\boldsymbol{\sigma}_{\mathbf{a},\text{disc}}^{\epsilon,k}, \mathbf{v})_{\omega_{\mathbf{a}}} - (\nabla\cdot\mathbf{v}, r_{\mathbf{a},\text{disc}}^{\epsilon,k})_{\omega_{\mathbf{a}}} = (\boldsymbol{\lambda}_{\mathbf{a}}^{\epsilon,k}, \mathbf{v})_{\omega_{\mathbf{a}}}, \qquad \forall \mathbf{v} \in \mathbf{W}_h^{\mathbf{a}}, \tag{4.24a}$$

$$(\nabla\cdot\boldsymbol{\sigma}_{\mathbf{a},\text{disc}}^{\epsilon,k}, q)_{\omega_{\mathbf{a}}} + \delta_\star(r_{\mathbf{a},\text{disc}}^{\epsilon,k}, q)_{\omega_{\mathbf{a}}} = (g_{\mathbf{a}}^{\epsilon,k} - \psi_{\mathbf{a}}\mathbf{r}_h^{\epsilon,k}, q)_{\omega_{\mathbf{a}}}, \quad \forall q \in Q_h^{\mathbf{a}}, \tag{4.24b}$$

with

$$g_{\mathbf{a}}^{\epsilon,k} := \Pi_h\Big(\psi_{\mathbf{a}}\big(f - \widetilde{\beta}_\epsilon(u_h^{\epsilon,k})\big)\Big)|_{\omega_{\mathbf{a}}} - \nabla\psi_{\mathbf{a}}\cdot\big(\Pi_h\boldsymbol{\xi} + \nabla\widetilde{\varphi}_\epsilon(u_h^{\epsilon,k})\big)|_{\omega_{\mathbf{a}}}, \qquad \boldsymbol{\lambda}_{\mathbf{a}}^{\epsilon,k} := -\Big(\psi_{\mathbf{a}}\boldsymbol{\xi} + \psi_{\mathbf{a}}\nabla\widetilde{\varphi}_\epsilon(u_h^{\epsilon,k})\Big)|_{\omega_{\mathbf{a}}}.$$

Then, we define the postprocessed saturation by $\Phi_{\mathbf{a},\text{disc}}^{\epsilon,k} := \delta_\star r_{\mathbf{a},\text{disc}}^{\epsilon,k} + \Pi_h(\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon(u_h^{\epsilon,k}))$ and we extend each $(\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, \Phi_{\mathbf{a}}^{\epsilon,k})_{\text{disc}}$ by zero in $\Omega \setminus \omega_{\mathbf{a}}$, $(\boldsymbol{\sigma}_h^{\epsilon,k}, \Phi_h^{\epsilon,k})_{\text{disc}} \in \mathbf{RTN}_p(\mathcal{T}_h) \times \mathbb{P}_p(\mathcal{T}_h)$ is given by

$$(\boldsymbol{\sigma}_h^{\epsilon,k}, \Phi_h^{\epsilon,k})_{\text{disc}} := \sum_{\mathbf{a}\in\mathcal{V}_h} (\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, \Phi_{\mathbf{a}}^{\epsilon,k})_{\text{disc}}. \tag{4.25}$$

The key properties of the above reconstructions with $\delta_\star \in \{0, 1\}$ are summarized in the following result.

Lemma 4.6 (Properties of $(\boldsymbol{\sigma}_h^{\epsilon,k}, \Phi_h^{\epsilon,k})$)**.** *Let $(\boldsymbol{\sigma}_h^{\epsilon,k}, \Phi_h^{\epsilon,k})$ be given by Reconstruction 1 or 2. Then it hold*

1. $(\boldsymbol{\sigma}_h^{\epsilon,k}, \Phi_h^{\epsilon,k}) \in \mathbf{H}(\text{div}, \Omega) \times L^2(\Omega)$,
2. $\nabla\cdot\boldsymbol{\sigma}_h^{\epsilon,k} + \Phi_h^{\epsilon,k} = \Pi_h f$, *with*
   - 2.a. $\nabla\cdot\boldsymbol{\sigma}_{h,\text{disc}}^{\epsilon,k} + \Phi_{h,\text{disc}}^{\epsilon,k} = \Pi_h f - \mathbf{r}_h^{\epsilon,k}$,
   - 2.b. $\nabla\cdot\boldsymbol{\sigma}_{h,\text{lin}}^{\epsilon,k} + \Phi_{h,\text{lin}}^{\epsilon,k} = \mathbf{r}_h^{\epsilon,k}$,
3. $(\nabla\cdot\boldsymbol{\sigma}_{\mathbf{a},\text{disc}}^{\epsilon,k}, 1)_{\omega_{\mathbf{a}}} = (\nabla\cdot\boldsymbol{\sigma}_{\mathbf{a},\text{lin}}^{\epsilon,k}, 1)_{\omega_{\mathbf{a}}} = 0$, $\forall \mathbf{a} \in \mathcal{V}_h^{\text{int}}$,
4. $(\Phi_{\mathbf{a},\text{disc}}^{\epsilon,k}, 1)_{\omega_{\mathbf{a}}} = (\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon(u_h^{\epsilon,k}), 1)_{\omega_{\mathbf{a}}}$ *and* $(\Phi_{\mathbf{a},\text{lin}}^{\epsilon,k}, 1)_{\omega_{\mathbf{a}}} = (\psi_{\mathbf{a}}(\widetilde{\beta}_\epsilon^{k-1}(u_h^{\epsilon,k}) - \widetilde{\beta}_\epsilon(u_h^{\epsilon,k})), 1)_{\omega_{\mathbf{a}}}$, $\forall \mathbf{a} \in \mathcal{V}_h^{\text{int}}$.

Proof. Its straightforward to see that for any $\mathbf{a} \in \mathcal{V}_h$ the zero extension of $\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}$ is $\mathbf{H}(\text{div}, \Omega)$-conforming due to the Neumann boundary condition specified in $\mathbf{W}_h^{\mathbf{a}}$. The properties 2.a.-2.b. use the fact that $\sum_{\mathbf{a}\in\mathcal{V}_K} \psi_{\mathbf{a}}|_K = 1$ and $\sum_{\mathbf{a}\in\mathcal{V}_K} \nabla\psi_{\mathbf{a}}|_K = 0$. To get 2.a. (2.b. uses the same arguments), we have from (4.24b), $\nabla\cdot\boldsymbol{\sigma}_{h,\text{disc}}^{\epsilon,k} + \Phi_{h,\text{disc}}^{\epsilon,k} = \sum_{\mathbf{a}\in\mathcal{V}_h} \nabla\cdot\boldsymbol{\sigma}_{\mathbf{a},\text{disc}}^{\epsilon,k} + \Phi_{\mathbf{a},\text{disc}}^{\epsilon,k} = \sum_{\mathbf{a}\in\mathcal{V}_h} \nabla\cdot\boldsymbol{\sigma}_{\mathbf{a},\text{disc}}^{\epsilon,k} + \delta_\star r_{\mathbf{a},\text{disc}}^{\epsilon,k} + \Pi_h(\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon(u_h^{\epsilon,k})) = \sum_{\mathbf{a}\in\mathcal{V}_h}\{\Pi_h(\psi_{\mathbf{a}} f - \psi_{\mathbf{a}}\widetilde{\beta}_\epsilon(u_h^{\epsilon,k}) + \psi_{\mathbf{a}}\widetilde{\beta}_\epsilon(u_h^{\epsilon,k})) - \psi_{\mathbf{a}}\mathbf{r}_h^{\epsilon,k} - \nabla\psi_{\mathbf{a}}\cdot\big(\Pi_h\boldsymbol{\xi} + \nabla\widetilde{\varphi}_\epsilon(u_h^{\epsilon,k})\big)\} = \Pi_h f - \mathbf{r}_h^{\epsilon,k}$. To show 3, we have for $\mathbf{a} \in \mathcal{V}_h^{\text{int}}$,

for $(\nabla\cdot\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\star},1)_{\omega_\mathbf{a}} = (\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\star}\cdot\mathbf{n}_{\omega_\mathbf{a}},1)_{\partial\omega_\mathbf{a}} = 0$ because of the zero Neumann boundary condition in $\mathbf{W}^\mathbf{a}_h$. The property 4 is trivial when $\delta_\star = 0$, and for $\delta_\star = 1$, we can remark that the source term $g^{\epsilon,k}_\mathbf{a}$ in (4.24b) satisfies

$$(g^{\epsilon,k}_\mathbf{a} - \psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h, 1)_{\omega_\mathbf{a}} = 0, \quad \forall \mathbf{a}\in\mathcal{V}^{\rm int}_h. \tag{4.26}$$

As a result, we can use (4.24b) and property 3 to see that $(r^{\epsilon,k}_{\mathbf{a},\rm disc},1)_{\omega_\mathbf{a}} = 0$ as well, and thus $(\Phi^{\epsilon,k}_{\mathbf{a},\rm disc},1)_{\omega_\mathbf{a}} = (\psi_\mathbf{a}\widetilde{\beta}_\epsilon(u^{\epsilon,k}_h),1)_{\omega_\mathbf{a}}$. We repeat the same steps to get that $(\Phi^{\epsilon,k}_{\mathbf{a},\rm lin},1)_{\omega_\mathbf{a}} = \left(\psi_\mathbf{a}\big(\widetilde{\beta}^{k-1}_\epsilon(u^{\epsilon,k}_h) - \widetilde{\beta}_\epsilon(u^{\epsilon,k}_h)\big),1\right)_{\omega_\mathbf{a}}$. □

Remark 4.7 (Cheaper flux and saturation reconstruction)**.** *In accordance with the approach outlined in [12], it is possible to perform a reconstruction with one degree lower in the space* $\mathbf{RTN}_{p-1}(\mathcal{T}_h)\times\mathbb{P}_{p-1}(\mathcal{T}_h)$*. While this method can be satisfactory in practical applications, it is important to note that the theoretical polynomial degree robustness (as expressed in Eq.* (4.49)*) may be compromised.*

Remark 4.8 (On the $(\rho-\mu)$ formulation)**.** *The above reconstruction and the analysis linking the dual norm of the residual with the underlying error energy norms in Proposition 4.1 can be adapted for the* $(\rho-\mu)$ *formulation. This adjustment simply requires replacing the nonlinear functions with their counterparts from* (2.3)*.*

4.3 Guaranteed reliability

We recall here two basic inequalities necessary in order to obtain our a posteriori error estimates. Let K be an element of $\mathcal{T}_h$, we recall the Poincaré–Friedrichs inequality,

$$||v-\Pi_h v||_K \le C_{\mathrm{P},K}||\nabla v||_K, \quad \text{for any } v\in H^1(K). \tag{4.27}$$

Since simplices are convex, there holds $C_{\mathrm{P},K} = \frac{h_K}{\pi}$. The generalized Friedrichs inequality states that

$$||v||_K \le \frac{h_\Omega}{\pi}||\nabla v||_K, \quad \text{for any } v\in H^1(K). \tag{4.28}$$

Theorem 4.9 (A basic upper bound)**.** *Let* u *be the weak solution of* (2.2) *and let* $u^{\epsilon,k}_h\in H^1_0(\Omega)$ *be the solution to* (3.5) *at the iteration* $k\ge 1$*. Let* $\boldsymbol{\sigma}^{\epsilon,k}_h$ *and* $\Phi^{\epsilon,k}_h$ *be, respectively, equilibrated flux and saturation reconstructions given in Section 4.2. Then, with* $\kappa_\star = (1-\delta_\star)\frac{h_\Omega}{\pi}+\delta_\star$*, for* $\delta_\star\in\{0,1\}$*, we have*

$$\begin{aligned}\delta_\star||u-u^{\epsilon,k}_h||^2_{\Sigma,\star} + (1-\delta_\star)\left\{\mathcal{J}_u(u^{\epsilon,k}_h) + ||u-u^{\epsilon,k}_h||^2_{\#}\right\}\\ \le \sum_{K\in\mathcal{T}_h}\left[\frac{h_K}{\pi}||f-\Pi_h f||_K + \kappa_\star||\Phi^{\epsilon,k}_h - \beta(u^{\epsilon,k}_h)||_K + ||\boldsymbol{\sigma}^{\epsilon,k}_h + \boldsymbol{\xi} + \nabla\varphi(u^{\epsilon,k}_h)||_K\right]^2.\end{aligned} \tag{4.29}$$

Proof. Recalling the flux equilibration identity 2, we get

$$\begin{aligned}\langle\mathcal{R}_u(u^{\epsilon,k}_h),v\rangle &= (f,v) - (\boldsymbol{\xi},\nabla v) - \left(\beta(u^{\epsilon,k}_h),v\right) - \left(\nabla\varphi(u^{\epsilon,k}_h),\nabla v\right),\\ &= (f-\Pi_h f,v) + \left(\Phi^{\epsilon,k}_h - \beta(u^{\epsilon,k}_h),v\right) - \left(\boldsymbol{\sigma}^{\epsilon,k}_h + \boldsymbol{\xi} + \nabla\varphi(u^{\epsilon,k}_h),\nabla v\right),\end{aligned} \tag{4.30}$$

where we inserted $(\boldsymbol{\sigma}^{\epsilon,k}_h,\nabla v) + (\nabla\cdot\boldsymbol{\sigma}^{\epsilon,k}_h,v) = 0$. The next step is to estimate each term in (4.30) elementwise. For each element $K\in\mathcal{T}_h$, we use the identity $(g-\Pi_h g,v)_K = (g-\Pi_h g, v-\Pi_h v)_K$, for all $g\in L^2(\Omega)$, to obtain the following bound using Poincaré–Friedrichs inequality (4.27),

$$\left|\left(f-\Pi_h f,v\right)_K\right| \le ||f-\Pi_h f||_K\frac{h_K}{\pi}\|\nabla v\|_K. \tag{4.31}$$

The upper bound of the second term will depend on the considered error measure. If $\delta_\star = 0$, we can apply the Cauchy-Schwarz, and the Poincaré–Friedrichs inequality (4.28) to infer

$$\left(\Phi^{\epsilon,k}_h - \beta(u^{\epsilon,k}_h),v\right)_K \le ||\Phi^{\epsilon,k}_h - \beta(u^{\epsilon,k}_h)||_K\frac{h_\Omega}{\pi}||\nabla v||_K. \tag{4.32}$$

For $\delta_\star = 1$, we can simply use the Cauchy-Schwarz inequality,

$$\left(\Phi_h^{\epsilon,k} - \beta(u_h^{\epsilon,k}), v\right)_K \leq ||\Phi_h^{\epsilon,k} - \beta(u_h^{\epsilon,k})||_K ||v||_K. \tag{4.33}$$

For the last term of (4.30), one can apply the Cauchy–Schwarz inequality to get

$$\left(\boldsymbol{\sigma}_h^{\epsilon,k} + \boldsymbol{\xi} + \nabla\varphi(u_h^{\epsilon,k}), \nabla v\right)_K \leq ||\boldsymbol{\sigma}_h^{\epsilon,k} + \boldsymbol{\xi} + \nabla\varphi(u_h^{\epsilon,k})||_K ||\nabla v||_K. \tag{4.34}$$

Therefore, we deduce from (4.30) and the above inequalities that

$$\langle \mathcal{R}_u(u_h^{\epsilon,k}), v\rangle \leq \sum_{K\in\mathcal{T}_h} \left[\frac{h_K}{\pi}||f - \Pi_h f||_K + \kappa_\star||\Phi_h^{\epsilon,k} - \beta(u_h^{\epsilon,k})||_K + ||\boldsymbol{\sigma}_h^{\epsilon,k} + \boldsymbol{\xi} + \nabla\varphi(u_h^{\epsilon,k})||_K\right] ||v||_{\star,K}. \tag{4.35}$$

By applying the Cauchy–Schwarz inequality, the definitions of the dual norm of the residual (4.3) and (4.6) and (4.8) together with the fact that$\sum_{K\in\mathcal{T}} ||v||_{\star,K} := \sum_{K\in\mathcal{T}} ||\nabla v||_K^2 + \delta_\star||v||_K^2 = \|v\|_\star^2$, to get (4.29). □

We based our developments of Algorithm 3.1 on an energy-type a posteriori error estimate distinguishing the different error components at each nonlinear iteration $k \geq 1$. To prove this estimate, let us introduce, for an element $K \in \mathcal{T}_h$, the local discretization, quadrature, regularization, linearization, and data oscillation estimators as follows:

$$\eta_{\mathrm{disc},K}^{\epsilon,k} := ||\boldsymbol{\sigma}_{h,\mathrm{disc}}^{\epsilon,k} + \Pi_h^{\mathrm{RTN}}\boldsymbol{\xi} + \nabla\widetilde{\varphi}_\epsilon(u_h^{\epsilon,k})||_K + \delta_\star||\Phi_{h,\mathrm{disc}}^{\epsilon,k} - \widetilde{\beta}_\epsilon(u_h^{\epsilon,k})||_K \tag{4.36a}$$

$$\eta_{\mathrm{qd},K}^{\epsilon,k} := ||\nabla\widetilde{\varphi}(u_h^{\epsilon,k}) - \nabla\varphi(u_h^{\epsilon,k})||_K + \frac{h_K}{\pi}||\widetilde{\beta}(u_h^{\epsilon,k}) - \beta(u_h^{\epsilon,k})||_K, \tag{4.36b}$$

$$\eta_{\mathrm{reg},K}^{\epsilon,k} := ||\nabla\widetilde{\varphi}(u_h^{\epsilon,k}) - \nabla\widetilde{\varphi}_\epsilon(u_h^{\epsilon,k})||_K + \kappa_\star||\widetilde{\beta}(u_h^{\epsilon,k}) - \widetilde{\beta}_\epsilon(u_h^{\epsilon,k})||_K, \tag{4.36c}$$

$$\eta_{\mathrm{lin},K}^{\epsilon,k} := ||\boldsymbol{\sigma}_{h,\mathrm{lin}}^{\epsilon,k}||_K + \kappa_\star||\Phi_{h,\mathrm{lin}}^{\epsilon,k}||_K, \tag{4.36d}$$

$$\eta_{\mathrm{osc},K}^{\epsilon,k} := \frac{h_K}{\pi}||f - \Pi_h f||_K + ||\boldsymbol{\xi} - \Pi_h^{\mathrm{RTN}}\boldsymbol{\xi}||_K. \tag{4.36e}$$

The global versions of the above estimators are given by

$$\eta_\bullet^{\epsilon,k} := \left\{\sum_{K\in\mathcal{T}_h}\left[\eta_{\bullet,K}^{\epsilon,k}\right]^2\right\}^{\frac{1}{2}}. \tag{4.37}$$

Then we have:

Theorem 4.10 (A posteriori estimate distinguishing the different error components)**.** *Let u be the weak solution of* (2.2) *and let $u_h^{\epsilon,k} \in H_0^1(\Omega)$ be the solution to* (3.5) *at the iteration $k \geq 1$. Let $\boldsymbol{\sigma}_h^{\epsilon,k}$ and $\Phi_h^{\epsilon,k}$ be, respectively, equilibrated flux and saturation reconstructions given in Section 4.2. Then, with $\kappa_\star = (1-\delta_\star)\frac{h_\Omega}{\pi} + \delta_\star$, for $\delta_\star \in \{0,1\}$, we have*

$$\delta_\star||u - u_h^{\epsilon,k}||_{\Sigma,\star} + (1-\delta_\star)\left[\left\{\mathcal{J}_u(u_h^{\epsilon,k}) + ||u - u_h^{\epsilon,k}||_\#^2\right\}^{\frac{1}{2}}\right] \leq \eta_{\mathrm{disc}}^{\epsilon,k} + \eta_{\mathrm{lin}}^{\epsilon,k} + \eta_{\mathrm{reg}}^{\epsilon,k} + \eta_{\mathrm{osc}}^{\epsilon,k}. \tag{4.38}$$

Proof. Similar to Theorem 4.9, we use the equilibration property (4.30) to get

$$\begin{aligned}\langle \mathcal{R}_u(u_h^{\epsilon,k}), v\rangle = (f - \Pi_h f, v) + \left(\boldsymbol{\xi} - \Pi_h^{\mathrm{RTN}}\boldsymbol{\xi}, \nabla v\right) + \left(\widetilde{\beta}(u_h^{\epsilon,k}) - \beta(u_h^{\epsilon,k}), v\right) + \left(\Phi_h^{\epsilon,k} - \widetilde{\beta}(u_h^{\epsilon,k}), v\right)\\ + \left(\nabla\widetilde{\varphi}(u_h^{\epsilon,k}) - \nabla\varphi(u_h^{\epsilon,k}), \nabla v\right) - \left(\boldsymbol{\sigma}_h^{\epsilon,k} + \Pi_h^{\mathrm{RTN}}\boldsymbol{\xi} + \nabla\widetilde{\varphi}(u_h^{\epsilon,k}), \nabla v\right).\end{aligned} \tag{4.39}$$

where we added and subtracting $\widetilde{\varphi}(u_h^{\epsilon,k})$ and $\widetilde{\beta}(u_h^{\epsilon,k})$. To estimate each term, we proceed as in Theorem 4.9 to first get

$$\begin{aligned}\langle \mathcal{R}_u(u_h^{\epsilon,k}), v\rangle \leq \sum_{K\in\mathcal{T}_h}\Bigg[\frac{h_K}{\pi}||f - \Pi_h f||_K + ||\boldsymbol{\xi} - \Pi_h^{\mathrm{RTN}}\boldsymbol{\xi}||_K + \frac{h_K}{\pi}||\widetilde{\beta}(u_h^{\epsilon,k}) - \beta(u_h^{\epsilon,k})||_K + ||\nabla\widetilde{\varphi}(u_h^{\epsilon,k}) - \nabla\varphi(u_h^{\epsilon,k})||_K\\ + \kappa_\star||\Phi_h^{\epsilon,k} - \widetilde{\beta}(u_h^{\epsilon,k})||_K + ||\boldsymbol{\sigma}_h^{\epsilon,k} + \Pi_h^{\mathrm{RTN}}\boldsymbol{\xi} + \nabla\widetilde{\varphi}(u_h^{\epsilon,k})||_K\Bigg] ||v||_{\star,K}.\end{aligned} \tag{4.40}$$

By adding and subtracting $\widetilde{\varphi}_\epsilon(u_h^{\epsilon,k})$ and $\widetilde{\beta}_\epsilon(u_h^{\epsilon,k})$ to the last two terms respectively, applying the triangular inequality, and using the fact that $\Phi_h^{\epsilon,k} = \Phi_{h,\mathrm{disc}}^{\epsilon,k} + \Phi_{h,\mathrm{lin}}^{\epsilon,k}$ and $\boldsymbol{\sigma}_h^{\epsilon,k} = \boldsymbol{\sigma}_{h,\mathrm{disc}}^{\epsilon,k} + \boldsymbol{\sigma}_{h,\mathrm{lin}}^{\epsilon,k}$, we obtain

$$\langle \mathcal{R}_u(u_h^{\epsilon,k}), v\rangle \leq \sum_{K\in\mathcal{T}_h} \left[\eta_{\mathrm{disc},K}^{\epsilon,k} + \eta_{\mathrm{lin},K}^{\epsilon,k} + \eta_{\mathrm{reg},K}^{\epsilon,k} + \eta_{\mathrm{osc},K}^{\epsilon,k}\right] ||v||_{\star,K},$$

after assembling the terms using (4.36)-(4.37). We conclude the estimate using Cauchy–Schwarz inequality after recalling the identity (4.6) for $\delta_\star = 0$ and (4.8) for $\delta_\star = 0$, while we use the H^1-energy norm (4.7) so that $\sum_{K\in\mathcal{T}} ||\nabla v||_K^2 + \delta_\star ||v||_K^2 = \|v\|_\star^2$. □

4.4 Local efficiency and robustness

This section proves the local and therefrom the global efficiency of our a posteriori error estimate. To accomplish this, we must introduce the Sobolev spaces defined on a patchwise basis

$$H_+^1(\omega_{\mathbf{a}}) := \begin{cases} \{H^1(\omega_{\mathbf{a}}); \quad v = 0 \text{ on } \partial\omega_{\mathbf{a}} \cap \partial\Omega\} \text{ if } \mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}, \\ \{H^1(\omega_{\mathbf{a}}); \quad (v,1)_{\omega_{\mathbf{a}}} = 0\} \text{ if } \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}, \end{cases} \tag{4.41}$$

that have mean-value zero on the patch subdomain ω_a if $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$, or that vanish on $\partial\omega_{\mathbf{a}} \cap \partial\Omega$ if $\mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}$ is a boundary vertex. Next, we recall that there is a constant depending only on the mesh shape-regularity such that for $v \in H_+^1(\omega_{\mathbf{a}})$

$$||\nabla(\psi_{\mathbf{a}} v)||_{\omega_{\mathbf{a}}} \lesssim ||\nabla v||_{\omega_{\mathbf{a}}}, \quad \text{for all } \mathbf{a} \in \mathcal{V}. \tag{4.42}$$

Therefore, utilizing the $||\cdot||_\star$ norm defined by (4.2), one can readily verify that

$$||\psi_{\mathbf{a}} v||_{\star,\omega_{\mathbf{a}}} \lesssim ||v||_{\star,\omega_{\mathbf{a}}}, \quad \text{for all } \mathbf{a} \in \mathcal{V}. \tag{4.43}$$

We also recall the Poincaré–Friedrichs inequality for functions in $H_+^1(\omega_{\mathbf{a}})$:

$$||v||_{\omega_{\mathbf{a}}} \lesssim h_{\omega_{\mathbf{a}}} ||\nabla v||_{\omega_{\mathbf{a}}}, \quad \text{for all } \mathbf{a} \in \mathcal{V}. \tag{4.44}$$

Additionally, we need to introduce the patchwise counterpart of the oscillation estimator (4.36e):

$$\tilde{\eta}_{\mathrm{osc},\omega_{\mathbf{a}}}^{\epsilon,k} := \sum_{K\in\mathcal{T}_h^{\mathbf{a}}} \frac{h_K}{\pi} ||\Pi_h(\psi_{\mathbf{a}} f) - \psi_{\mathbf{a}} f||_K + ||\boldsymbol{\xi} - \Pi_h \boldsymbol{\xi}||_K + ||\psi_{\mathbf{a}}\boldsymbol{\xi} - \Pi_h^{\mathrm{RTN}}(\psi_{\mathbf{a}}\boldsymbol{\xi})||_K, \tag{4.45}$$

and let

$$\left[\tilde{\eta}_{\mathrm{osc}}^{\epsilon,k}\right]^2 := \sum_{\mathbf{a}\in\mathcal{V}_h} \left\{\tilde{\eta}_{\mathrm{osc},\omega_{\mathbf{a}}}^{\epsilon,k}\right\}^2. \tag{4.46}$$

Referring to (4.36e), it is evident that for $K \in \mathcal{T}_h$, we have $\eta_{\mathrm{osc},K}^{\epsilon,k} \leq \tilde{\eta}_{\mathrm{osc},\omega_{\mathbf{a}}}^{\epsilon,k}$ due to the partition of unity. Consequently, we have $\eta_{\mathrm{osc}}^{\epsilon,k} \leq \tilde{\eta}_{\mathrm{osc}}^{\epsilon,k}$. We obtain the following stability result by following the procedure in [31, 47].

Proposition 4.11 (Stability of discretization-flux equilibration)**.** *Let $f \in L^2(\Omega)$ and $\boldsymbol{\xi} \in L^2(\Omega)$. Consider u as the weak solution of (2.2). Let $u_h^{\epsilon,k} \in L^2(\Omega)$ be the solution to (3.5) at iteration $k \geq 1$ such that $\varphi(u_h^{\epsilon,k}) \in H_0^1(\Omega)$, and $(\boldsymbol{\sigma}_{\mathbf{a},\mathrm{disc}}^{\epsilon,k}, \Phi_{\mathbf{a},\mathrm{disc}}^{\epsilon,k})$ be the equilibrated flux and saturation reconstructions given by Reconstruction 2 in accordance with Definition A.2. Let $\mathbf{a} \in \mathcal{V}_h$ be fixed, and define $\boldsymbol{\tau}_{\mathbf{a},\mathrm{disc}}^{\epsilon,k} := \Pi_h^{\mathrm{RTN}}(\psi_{\mathbf{a}}\boldsymbol{\xi}) + \psi_{\mathbf{a}} \nabla \widetilde{\varphi}_\epsilon(u_h^{\epsilon,k}) \in \mathbf{RTN}_p(\mathcal{T}_h^{\mathbf{a}})$. Then, for $\delta_\star \in \{0,1\}$, there holds*

$$\begin{aligned} \left\{ ||\boldsymbol{\sigma}_{\mathbf{a},\mathrm{disc}}^{\epsilon,k} + \boldsymbol{\tau}_{\mathbf{a},\mathrm{disc}}^{\epsilon,k}||_{\omega_{\mathbf{a}}}^2 + \delta_\star \left\| r_{\mathbf{a},\mathrm{disc}}^{\varepsilon,k} \right\|_{\omega_{\mathbf{a}}}^2 \right\}^{\frac{1}{2}} &\lesssim \min_{\substack{\mathbf{v}_h \in \mathbf{W}_h^{\mathbf{a}}, \\ \nabla\cdot\mathbf{v}_h = g_{\mathbf{a}}^{\epsilon,k} - \Pi_h(\psi_{\mathbf{a}} \mathbf{r}_h^{\epsilon,k})}} ||\mathbf{v}_h + \boldsymbol{\tau}_{\mathbf{a},\mathrm{disc}}^{\epsilon,k}||_{\omega_{\mathbf{a}}} \\ &\lesssim \max_{\substack{|\nabla v| = 1 \\ v \in H_+^1(\omega_{\mathbf{a}}) \setminus \{0\}}} \left(g_{\mathbf{a}}^{\epsilon,k} - \psi_{\mathbf{a}} \mathbf{r}_h^{\epsilon,k}, v\right)_{\omega_{\mathbf{a}}} - \left(\boldsymbol{\tau}_{\mathbf{a},\mathrm{disc}}^{\epsilon,k}, \nabla v\right)_{\omega_{\mathbf{a}}}, \end{aligned} \tag{4.47}$$

where the constant in $\lesssim$ depends only on the space dimension d, the shape-regularity constant of $\mathcal{T}_h$, and the polynomial degree p.

Theorem 4.12 (Local efficiency)**.** *Let the assumptions of Proposition 4.11 be satisfied. For a fixed* $\mathbf{a} \in \mathcal{V}_h$*, the following patchwise efficiency bound holds*

$$\left\{ ||\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}||^2_{\omega_{\mathbf{a}}} + \delta_\star \left\| r^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}} \right\|^2_{\omega_{\mathbf{a}}} \right\}^{\frac{1}{2}} \lesssim \max_{v \in H^1_0(\omega_{\mathbf{a}})\setminus\{0\}} \frac{\langle \mathcal{R}_u(u^{\epsilon,k}_h), \psi_{\mathbf{a}} v\rangle}{||\psi_{\mathbf{a}} v||_\star} + \eta^{\epsilon,k}_{\mathrm{lin},\omega_{\mathbf{a}}} + \eta^{\epsilon,k}_{\mathrm{reg},\omega_{\mathbf{a}}} + \eta^{\epsilon,k}_{\mathrm{qd},\omega_{\mathbf{a}}} + \tilde{\eta}^{\epsilon,k}_{\mathrm{osc},\omega_{\mathbf{a}}}. \tag{4.48}$$

Assuming that the adaptive criteria given by (3.8)–(3.10) *are satisfied. For all* $K \in \mathcal{T}_h$*, let* $\mathfrak{T}_K = \cup_{\mathbf{a}\in\mathcal{V}_K} \mathcal{T}^{\mathbf{a}}_h$*, and* ω_K *be the corresponding open subdomain. Then, we obtain the following local efficiency:*

$$\eta^{\epsilon,k}_{\mathrm{disc},K} + \eta^{\epsilon,k}_{\mathrm{lin},K} + \eta^{\epsilon,k}_{\mathrm{reg},K} \lesssim \max_{w \in H^1_0(\omega_K)\setminus\{0\}} \frac{\langle \mathcal{R}_u(u^{\epsilon,k}_h), \psi_{\mathbf{a}} w\rangle}{||\psi_{\mathbf{a}} w||_\star} + \eta^{\epsilon,k}_{\mathrm{qd},\omega_K} + \tilde{\eta}^{\epsilon,k}_{\mathrm{osc},\omega_K}, \tag{4.49}$$

where the constant in $\lesssim$ *depends only on the balancing parameters* Γ_{lin} *and* Γ_{reg}*, the space dimension* d*, the shape-regularity constant of* $\mathcal{T}_h$*, and the polynomial degree* p*.*

Lemma 4.13 (Global efficiency)**.** *Let the assumptions of Proposition 4.11 be satisfied. Then, for* $\delta_\star \in \{0, 1\}$*, the global efficiency can be summarized as*

$$\eta^{\epsilon,k}_{\mathrm{disc}} + \eta^{\epsilon,k}_{\mathrm{lin}} + \eta^{\epsilon,k}_{\mathrm{reg}} \lesssim \delta_\star ||u - u^{\epsilon,k}_h||_{\Sigma,\star} + (1-\delta_\star)\left\{ \mathcal{J}_u(u^{\epsilon,k}_h) + ||u - u^{\epsilon,k}_h||^2_{\#} \right\}^{\frac{1}{2}} + \eta^{\epsilon,k}_{\mathrm{qd}} + \tilde{\eta}^{\epsilon,k}_{\mathrm{osc}}. \tag{4.50}$$

Proof. For each $K \in \mathcal{T}_h$, we recall that $\eta^{\epsilon,k}_{\mathrm{disc},K} = ||\boldsymbol{\sigma}^{\epsilon,k}_{h,\mathrm{disc}} + \Pi^{\mathrm{RTN}}_h \boldsymbol{\xi} + \nabla\widetilde{\varphi}_\epsilon(u^{\epsilon,k}_h)||_K + \delta_\star ||r^{\epsilon,k}_{h,\mathrm{disc}}||_K$. Noting that K has $(d+1)$ vertices collected in the set $\mathcal{V}_K$ for the simplicial mesh we consider, so that using the partition of unity we have

$$\begin{aligned}
\sum_{K\in\mathcal{T}_h} \left[\eta^{\epsilon,k}_{\mathrm{disc},K}\right]^2 &\lesssim \sum_{K\in\mathcal{T}_h} ||\boldsymbol{\sigma}^{\epsilon,k}_{h,\mathrm{disc}} + \Pi^{\mathrm{RTN}}_h \boldsymbol{\xi} + \nabla\widetilde{\varphi}_\epsilon(u^{\epsilon,k}_h)||^2_K + \delta_\star ||r^{\epsilon,k}_{h,\mathrm{disc}}||^2_K, \\
&\leq (d+1) \sum_{K\in\mathcal{T}_h} \sum_{\mathbf{a}\in\mathcal{V}_K} ||\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}||^2_K + \delta_\star \left\| r^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}} \right\|^2_K, \\
&= (d+1) \sum_{\mathbf{a}\in\mathcal{V}_h} ||\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}||^2_{\omega_{\mathbf{a}}} + \delta_\star \left\| r^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}} \right\|^2_{\omega_{\mathbf{a}}}.
\end{aligned}$$

We now rely on (4.48) to infer

$$\left\{ \sum_{K\in\mathcal{T}_h} \left[\eta^{\epsilon,k}_{\mathrm{disc},K}\right]^2 \right\}^{\frac{1}{2}} \lesssim \sum_{\mathbf{a}\in\mathcal{V}_h} \max_{v\in H^1_0(\omega_{\mathbf{a}})\setminus\{0\}} \frac{\langle \mathcal{R}_u(u^{\epsilon,k}_h), \psi_{\mathbf{a}} v\rangle}{||\psi_{\mathbf{a}} v||_\star} + \eta^{\epsilon,k}_{\mathrm{lin},\omega_{\mathbf{a}}} + \eta^{\epsilon,k}_{\mathrm{reg},\omega_{\mathbf{a}}} + \eta^{\epsilon,k}_{\mathrm{qd},\omega_{\mathbf{a}}} + \tilde{\eta}^{\epsilon,k}_{\mathrm{osc},\omega_{\mathbf{a}}}.$$

Therefrom, we use the stopping criteria (3.8) and (3.9) with sufficiently small Γ_{lin} and Γ_{reg} to bound the error estimators with the discretization estimator:

$$\begin{aligned}
\eta^{\epsilon,k}_{\mathrm{disc}} + \eta^{\epsilon,k}_{\mathrm{lin}} + \eta^{\epsilon,k}_{\mathrm{reg}} &\lesssim \sum_{\mathbf{a}\in\mathcal{V}_h} \sup_{v\in H^1_0(\omega_{\mathbf{a}})\setminus\{0\}} \frac{\langle \mathcal{R}_u(u^{\epsilon,k}_h), \psi_{\mathbf{a}} v\rangle}{||\psi_{\mathbf{a}} v||_\star} + \eta^{\epsilon,k}_{\mathrm{qd}} + \eta^{\epsilon,k}_{\mathrm{osc}} \\
&\lesssim \delta_\star ||u - u^{\epsilon,k}_h||_{\Sigma,\star} + (1-\delta_\star)\left\{ \mathcal{J}_u(u^{\epsilon,k}_h) + ||u - u^{\epsilon,k}_h||^2_{\#} \right\}^{\frac{1}{2}} \\
&\quad + \eta^{\epsilon,k}_{\mathrm{qd}} + \tilde{\eta}^{\epsilon,k}_{\mathrm{osc}},
\end{aligned} \tag{4.51}$$

Therein, we use the the localization of dual norms with $\psi_{\mathbf{a}}$-Galerkin orthogonality as in [18], to verify $\sum_{\mathbf{a}\in\mathcal{V}_h} \sup_{v\in H^1_0(\omega_{\mathbf{a}})\setminus\{0\}} \frac{\langle \mathcal{R}_u(u^{\epsilon,k}_h), \psi_{\mathbf{a}} v\rangle}{||\psi_{\mathbf{a}} v||_\star} \lesssim \left|\left|\left| \mathcal{R}_u(u^{\epsilon,k}_h) \right|\right|\right|_\star$. The proof is complete by employing (4.6) and (4.8).

□

5 Numerical results

In this section, we illustrate our theoretical results through a series of numerical experiments using the finite element method. The discretization is conducted using Lagrange finite elements. Our focus in this paper is on the linearization and regularization processes, with no detailed reporting on quadrature rules, associated errors, or oscillation errors. It's important to note that, in the literature, standard finite elements for (3.2) may require specific quadrature rules or mass-lumped versions of high-order methods to ensure convergence properties.

Given that we exclusively deal with analytical solutions, we compute the exact error components for discretization and linearization as follows: $\text{error}^k_{\text{lin}} = ||u_h - u_h^k||_{\#}$ and $\text{error}^k_{\text{disc}} = ||u - u_h||_{\#}$. Specifically, we assume $||\varphi(u) - \varphi(u_h)||_{H^{-1}(\Omega)} = ||\varphi(u) - \varphi(u_h)||$, which, in fact, results in a slightly overestimation of the exact error (see Remark 4.3). All presented results are derived using Reconstruction 1. Furthermore, for approximations with orders $p \geq 2$, cost-effective error estimations can still be achieved by utilizing a reconstruction of one degree lower.

5.1 Test case 1: An illustrative example

5.1.1 Offline optimization of the L-scheme

We set $\Omega = [0,1]^2$ and let $\beta(u) = u$ and $\varphi(u) = \log(u+1)$, $\boldsymbol{\xi} = 0$, and f complying to the solution $u(x) = x(1-x)y(1-y)$ on Ω. In particular, due to the Lipschitz continuity of φ, there is no need for regularization or variable transformation. The discretization is achieved using first-order Lagrange conforming finite elements, for which we conducted tests with both first-order and one-degree-lower reconstruction (cheaper), both yielding comparable results. In this initial test case, we utilize developed error estimates to compare the L-scheme and Newton's method. Prior to conducting the comparison, we capitalize on the advantage of splitting error sources to optimize the L-scheme parameter.

To achieve this optimization, we implement a "brute force" iterative method that utilizes the linearization estimator to find the optimal parameter L during an offline phase, as detailed in Remark 3.1. This process involves exploring various values of L to determine the one that leads to the minimum iteration number, effectively reducing the overall computational cost. The offline computation of L is carried out on coarse grids, which further contributes to its negligible impact on total computational costs, thanks to the error decomposition technique employed. Overall, this approach allows us to fine-tune the L-scheme, leading to improved efficiency and convergence in the subsequent calculations.

Figure 1 (left) illustrates the L-scheme iteration as a function of the parameter L, comparing both standard and adaptive (with $\Gamma_{\text{lin}} \in \{0.02, 0.2\}$) stopping criteria. The corresponding error estimators over the L-scheme iterations are depicted in Figure 1 (right) for $\Gamma_{\text{lin}} = 0.02$ and various values of L (with final iterations utilizing the standard stopping criteria). The optimal value for L is determined to be approximately 0.975, which is then used in Algorithm 3.1 to perform computations on the desired fine grids. Note that this value is close but not equal to the Lipschitz constant, which is 1 for this case. Employing the adaptive stopping criteria leads to significant savings in the number of iterations, and it also helps mitigate the impact of non-optimal values of L. The effectiveness of our balancing strategy is evident from the results shown in Table 1, where the exact and estimated errors exhibit an effectivity index close to 1. To further compare Newton's method with the L-scheme, we utilize Algorithm 3.1 and plot the various

$1/h$	error estimate	exact error	effectivity ind.
8	0.03184	0.0296325	1.074
16	0.01578	0.0149003	1.058
32	0.00785	0.0074607	1.053
64	0.00394	0.0037316	1.051

Table 1: Comparison of exact and estimated errors and resulting effectivity index relative to mesh size h. Results are obtained with the L-scheme with $\Gamma_{\text{lin}} = 0.02$.

error components for both methods in Figure 2. It is important to note that the discretization estimators are equal in both methods, while the linearization estimator η^k_{lin} differs. Therefore, the results presented in

Table 1 are equally applicable when using Newton's method. This observation further supports the error splitting techniques employed. In comparing the exact and estimated linearization errors with the two methods, we iterate each of the nonlinear solvers until $\eta_{\text{lin}}^k \leq 10^{-10}$, yielding an almost converged solution. Figure 3 showcases the exact and estimated error components after applying our adaptive stopping criteria, demonstrating a strong agreement between the two. For both methods, the linearization estimator becomes negligible, contributing minimally to the total error.

In conclusion, the above results reaffirm that the optimization of the L-parameter in combination with our adaptive stopping criteria makes the L-scheme an effective alternative to Newton's method.

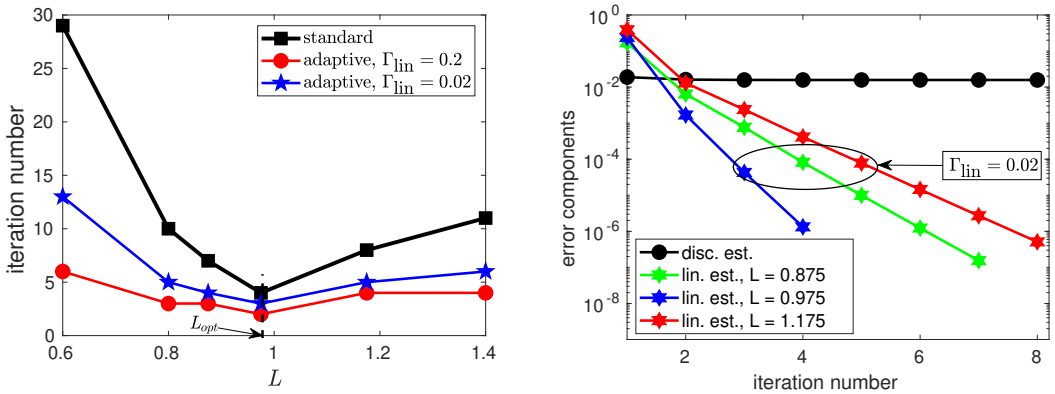

Figure 1: Offline computation of optimal L (left) alongside estimated error components (right). The error estimates are derived for $h = 1/16$ using the L-scheme, where the discretization estimator aligns with either the Newton or L-scheme.

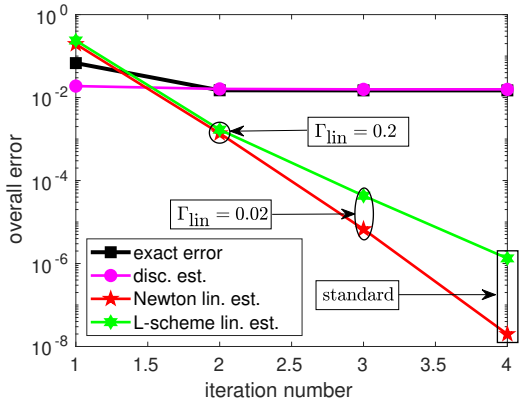

Figure 2: Comparing Newton and optimized L-scheme for various parameters Γ_{lin}. Both exact and estimated discretization errors coincides with either the Newton or L-scheme.

5.1.2 Mesh adaptivity for an L-shaped domain

In this scenario, we examine an L-shaped domain $\Omega = (-1,1) \times (-1,1) \setminus [0,1] \times [-1,0]$ and reverse the choices for β and φ from the previous case. Specifically, we set $\varphi(u) = u$ and $\beta(u) = \log(u+1)$, with $\boldsymbol{\xi} = 0$. The exact solution is expressed in polar coordinates:

$$u(r,\theta) = r^{\frac{2}{3}} \sin\left(\frac{2\theta}{3}\right), \quad \forall \theta \in (0, \frac{3\pi}{2}).$$

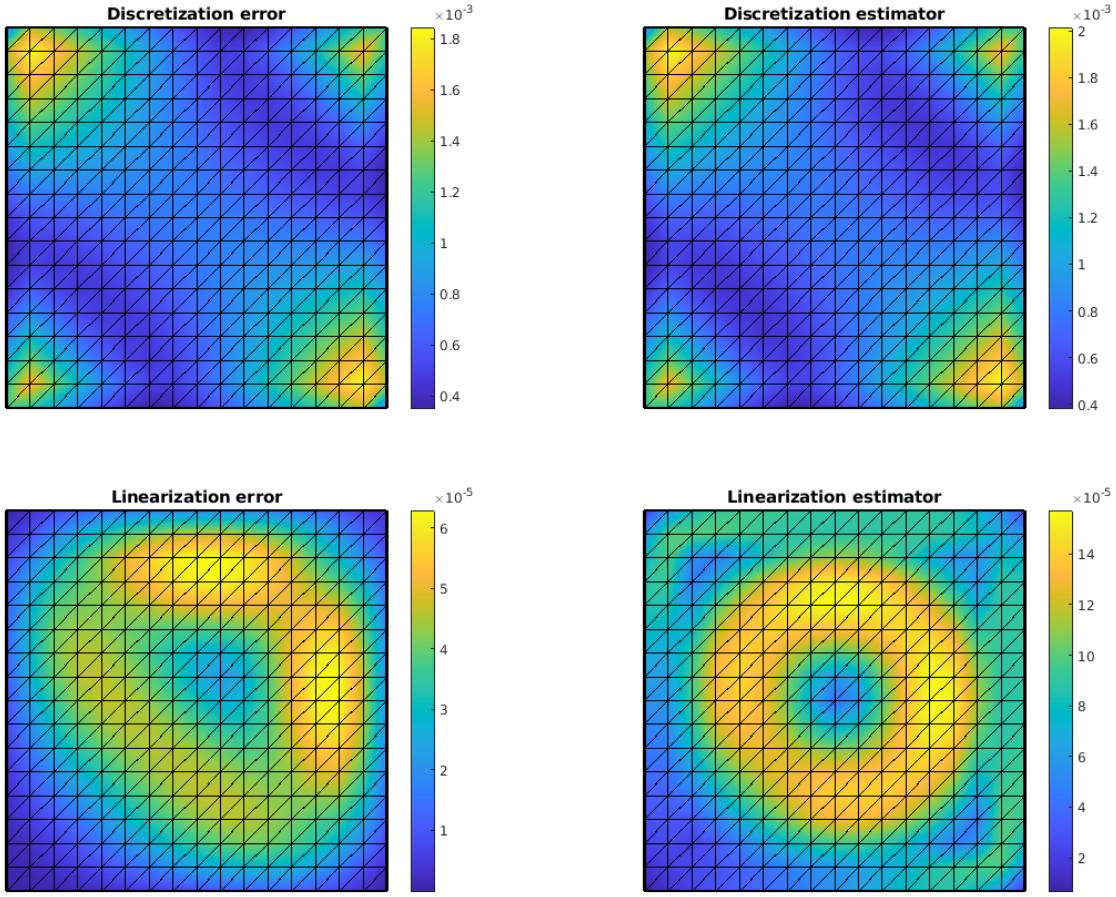

Figure 3: Comparison of discretization errors (top left) and estimators (top right), along with linearization errors (bottom left) and estimators (bottom right) using the L-scheme. The discretization errors remain identical when employing the Newton method.

We observe that the exact solution features a singularity at the origin, with $u \in H^{\frac{5}{3}-\varepsilon}(\Omega)$ for arbitrarily small $\varepsilon > 0$. The corresponding source term is $f(u) = \log(1+u)$. The singularity at the origin penalizes the convergence rate for uniformly refined grids [22].

For our comparisons, we set the parameter $\Gamma_{\text{lin}} = 0.1$ and assess the effectiveness of adaptive stopping criteria against standard ones. Subsequently, we examine the performance of adaptive mesh refinement in contrast to standard uniform mesh refinement. Here, regularization is find unnecessary due to the Lipschitz continuity of β and the non-singularity of the derivative. Our adaptive strategy aims to achieve equi-distribution of discretization errors, as outlined in (3.10).

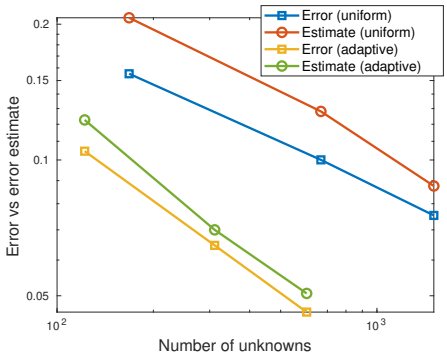

Figure 4: Comparison of errors and estimators for uniform and adaptive mesh refinement.

Method	Standard	Adaptive
Newton	4	2
optimized L-scheme	6	3

Table 2: Comparison of standard and adaptive linearizations

In Table 2, we present the total iteration numbers for the L-scheme and Newton methods with and without adaptive stopping criteria. Although this case is not particularly demanding in terms of nonlinear iterations, our stopping criteria lead to a 50% reduction in iterations. Figure 4 illustrates the errors and estimators for uniform and adaptive mesh refinements. A strong agreement is observed between the estimated and exact errors, with adaptive grids showcasing improved performance. With fewer elements, we achieve a superior convergence rate and smaller discretization errors compared to all runs with uniform refinement. This emphasizes the effectiveness of our localized estimators in controlling and reducing singular effects, consequently enhancing convergence rates for such scenarios. In Figure 5, we illustrate the distribution of the exact and estimated errors following the application of our adaptive stopping criteria. An excellent agreement is observed between the estimated and exact error components. In the non-adaptive case, localized and significant discretization errors around the corner are evident. However, these errors are effectively redistributed evenly over the domain by refining and coarsening the grids, in accordance with our discretization criteria.

As a result, the discretization error becomes equidistributed, and the maximum error improves by a factor of 2.5 after grid adaptivity. The linearization errors and estimators for both methods exhibit a sufficiently small contribution from the linearization into the overall error, as required by our adaptive stopping criteria. Clearly, the distribution of the error differs between the two methods, with the Newton method displaying a smoother distribution compared to the L-scheme. The differences in error distribution between the two methods can be attributed to their inherent characteristics. The Newton method, with its second-order convergence and full Jacobian, facilitates a more gradual and smoother adjustment to the correct solution. In contrast, the L-scheme's stabilization amendment may introduce oscillations and localized features, resulting in a less smooth error distribution.

5.2 Test case 2: A stefan-type model

In this case, the nonlinearity is represented by the Stefan-like function [49], posing a greater challenge for the standard Newton's method:

$$\varphi(u) = \begin{cases} u & \text{if } u < 0, \\ 0 & \text{if } 0 \le u \le 1, \\ u - 1 & \text{if } u \ge 1, \end{cases} \tag{5.1}$$

This increased complexity arises from the potential for the solution u to exhibit discontinuities, as φ contains plateaux. This test case, taken from [26], introduces the interface given by $s(x, y) = \gamma$ where $s(x, y) = \frac{x+y}{\sqrt{2}}$ with $\gamma = \frac{1}{3}$, so that the exact solution

$$u(x, y) = \cosh(s(x, y) - \gamma), \qquad \text{if } s(x, y) \ge \gamma, \tag{5.2}$$

$$u(x, y) = 0, \qquad \text{if } s(x, y) < \gamma, \tag{5.3}$$

is discontinuous along γ. Note that $\varphi(u)$ is, in fact, continuous since the discontinuity of u coincides with values where φ remains constant. Note that Assumptions (**A1**) to (**A4**) are satisfied for this case. We use a regularized function $\varphi_\epsilon(u)$ so that its strictly increasing, where the regularization parameter will be set up

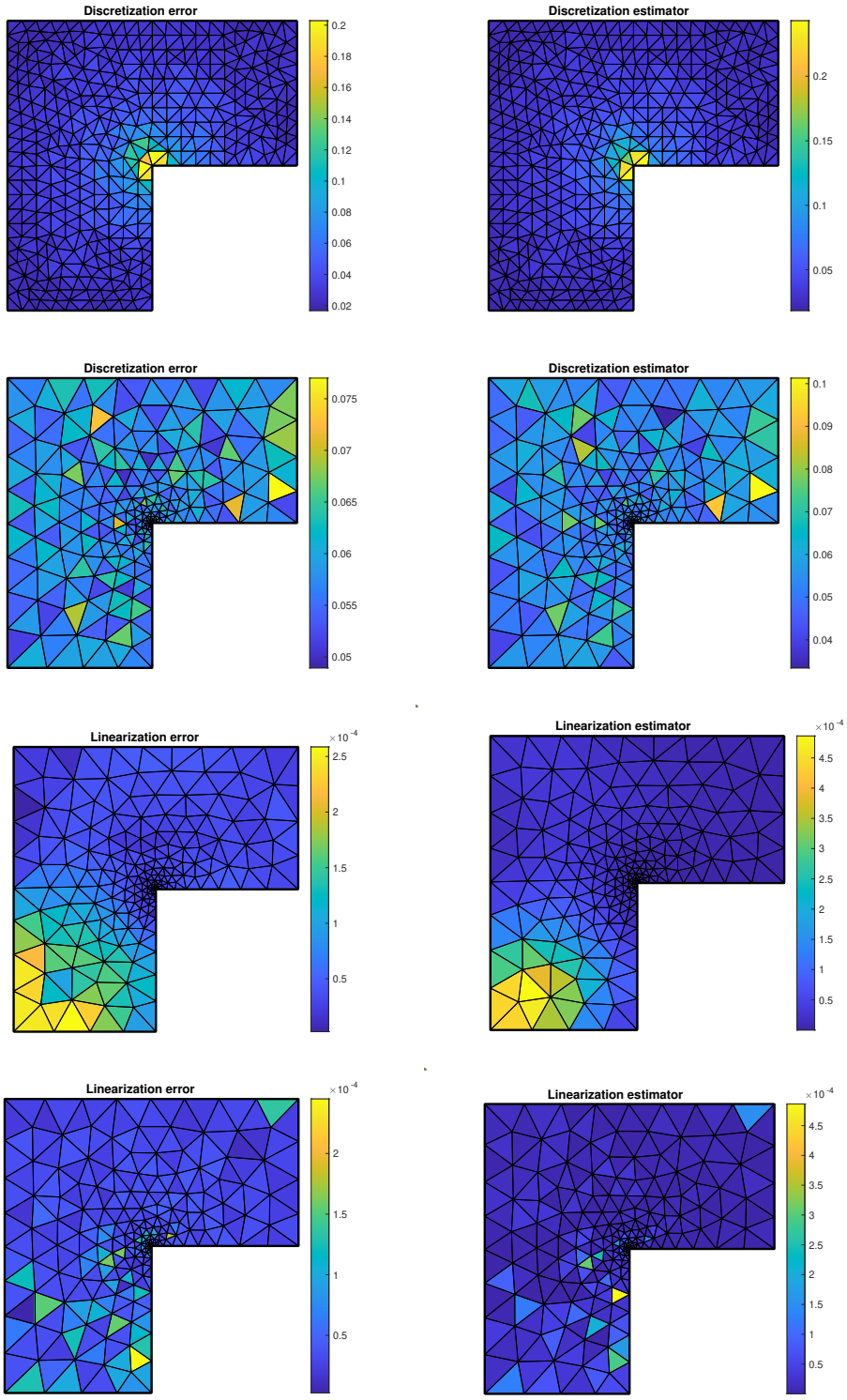

Figure 5: Spatial distribution of errors (left) and corresponding estimators (right). The top four figures illustrate the discretization error component, showcasing the use of a uniform mesh (top) and an adaptive mesh (bottom). The remaining figures depict the linearization error, illustrating its evaluation through Newton's method (top) and the L-scheme (bottom).

adaptively with the solver. We also set $\beta(u) = u$, leading to a zero source term in (2.2). In Figure 6 (left), we depict the functions φ and its regularized counterpart:

$$\varphi_\epsilon(u) = \begin{cases} u & \text{if } u < -\epsilon, \\ \dfrac{\epsilon\,(2\,u-1)}{1+2\,\epsilon} & \text{if } -\epsilon \le u \le 1+\epsilon, \\ u-1 & \text{if } u \ge 1+\epsilon, \end{cases} \tag{5.4}$$

5.2.1 Overall algorithm assessment

We used Algorithm 3.1 for $p = 1$ while setting $\Gamma_{\text{sw}} = 0.9$, $\Gamma_{\text{lin}} = 0.01$ and $\Gamma_{\text{reg}} = 0.1$. The offline computation of the optimal L-scheme parameter resulted in $L_\varphi^{\text{opt}} = 0.6225$, as illustrated in Figure 6 (right). We also set a fixed regularization parameter $\epsilon = 0.05$. Figure 7 (left) depicts the evolution of the spatial, nonlinear, and regularization error estimators as a function of the number of iterations for a fixed mesh and fixed regularization parameter ϵ. As anticipated, the linearization estimator consistently diminishes, whereas the other components remain stagnant, underscoring the distinct error component separation feature. By employing a switching criterion with $\Gamma_{\text{sw}} = 0.9$, we leverage improved initialization for Newton after completing 9 iterations with the L-scheme. The L-scheme iterations in our adaptive solver are stopped when the linearization estimator falls below the discretization estimator. Although the L-scheme may appear slow initially, it serves as a beneficial initialization for Newton, which exhibits rapid convergence towards the end. Coupled with the adaptive criterion, our solver necessitated a total of 13 iterations with $\Gamma_{\text{lin}} = 0.01$ and 11 iterations with $\Gamma_{\text{lin}} = 0.1$. It is crucial to emphasize that a standard Newton solver may experience oscillations or fail to converge when not provided with a sufficiently close starting point, especially for refined grids. Even with a close initialization, our adaptive Newton method significantly reduces the number of unnecessary iterations compared to a conventional criterion based on a fixed threshold of 10^{-6}. Given the slower convergence of the L-scheme compared to the faster convergence of Newton, especially towards the end, the selection of Γ_{sw} plays a crucial role. Opting for a larger or smaller Γ_{sw} can introduce numerous additional iterations, including oscillating ones for Newton. In Figure 7 (right), analogous findings are depicted regarding the criterion for the selection of the regularization parameter ϵ. As ϵ decreases, the regularization error estimator decreases as expected, while the spatial estimator plateaus. The regularization criterion with $\Gamma_{\text{reg}} = 0.1$ results in stopping the iterations after the fourth step, and with $\Gamma_{\text{reg}} = 0.01$, it stops after six steps.

5.2.2 Mesh adaptivity

The subsequent set of experiments is designed to evaluate the spatial balancing criterion (mesh adaptivity) by demonstrating its influence on the estimated error and effectivity index in relation to the total number of unknowns, and comparing it to fixed and uniform meshes. We commenced the experiment with a fixed mesh, utilizing consistent parameters for the adaptive, switching, and regularization criteria, as done previously. In Figure 8, we present the distribution of discretization (top left), regularization (top right), and linearization (bottom left) estimators using the adaptive algorithm. The spatial balancing criterion has resulted in a more evenly distributed error, with the maximum error concentrated at the end-points of the interface γ. This distribution is attributed to the still-discontinuous behavior of the boundary conditions initially computed from the exact and non-regularized solution.

Moreover, it is clear that both regularization and linearization errors are sufficiently small to have a negligible influence on the total error. The regularization error predominantly traces the interface γ, where the discontinuity is prominent. To evaluate the overall convergence of the adaptive solver, we present in Figure 8 (bottom right) the error and estimates as a function of the total number of unknowns in both the fully adaptive case and with uniform meshes (with $p = 2$). Even in the uniform case, Algorithm 3.1 is employed to adaptively control the linearization and regularization processes. Clearly, the adaptive approach outperforms the uniform one, delivering enhanced convergence rates with reduced computational costs. Note that the convergence rates in the uniform case was observed to consistently maintain at or slightly exceed $O(h^{0.8})$ for $p = 1$, only slightly declining to around $O(h^{1.5})$ for $p = 2$. The convergence rates were clearly enhanced through mesh adaptivity, although the specific rate improvement has not been calculated.

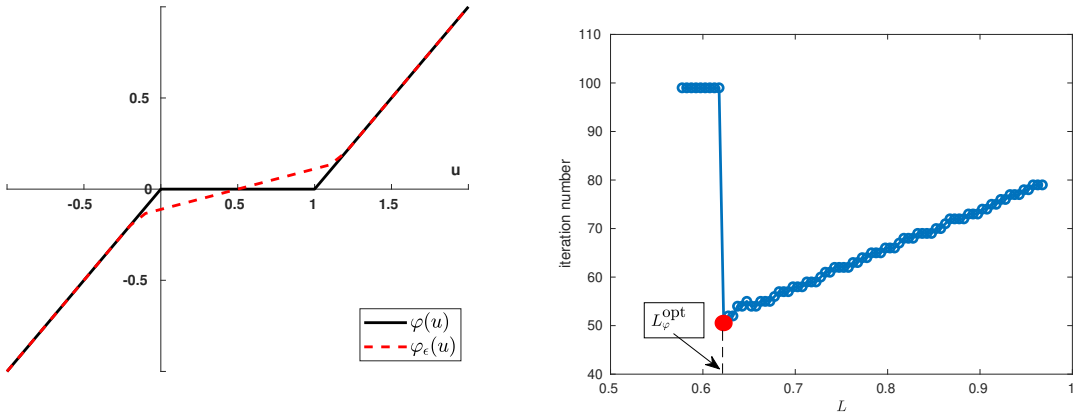

Figure 6: Illustration of a function φ and its regularization φ_ϵ (left), along with the offline computation of the optimal L-scheme parameter. The optimal L is determined for a small regularization parameter $\epsilon = 10^{-3}$ and a uniform coarse mesh with $h = 1/8$.

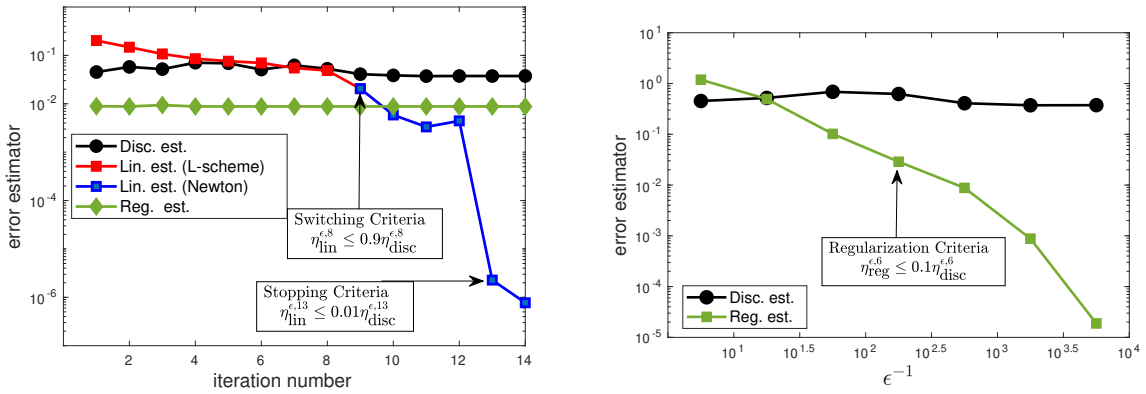

Figure 7: Evolution of the error estimators as a function of nonlinear iterations with adaptive switching and stopping criteria (left). Evolution of the regularization and discretization estimators with respect to the regularization parameter ϵ (right).

5.3 Test case 3: A porous media model

We consider now porous media models by setting $\beta(u) = u$ and $\varphi(u) = |u|^{m-1}u$, where $m > 0$. This particular selection of β and φ leads to a degenerate elliptic equation with intriguing properties depending on the value of m. We set f to comply with the solution $u(x) = \sin(\pi x)\sin(\pi y)$ on $\Omega = [0,1]^2$. To investigate the behavior of Algorithm 3.1 in handling the different choices of m, we conduct extensive testing on the model problem. This testing involves numerical simulations with various values of m, spanning both the slow and fast diffusion regimes. By employing a posteriori error estimation, adaptive schemes, and regularization techniques provided by Algorithm 3.1, we aim to obtain accurate and efficient approximations for the diverse cases of interest.

5.3.1 Models for slow regime

When $m > 1$, the model represents the slow diffusion case, which finds application in gas diffusion within porous media, often referred to as the slow diffusion model. In this scenario, the equation exhibits degeneracy at unknown points where $u = 0$. This means that the diffusion process becomes slow or even stagnant in regions where the solution approaches zero, leading to spatially localized behavior with potential concentration gradients near these points.

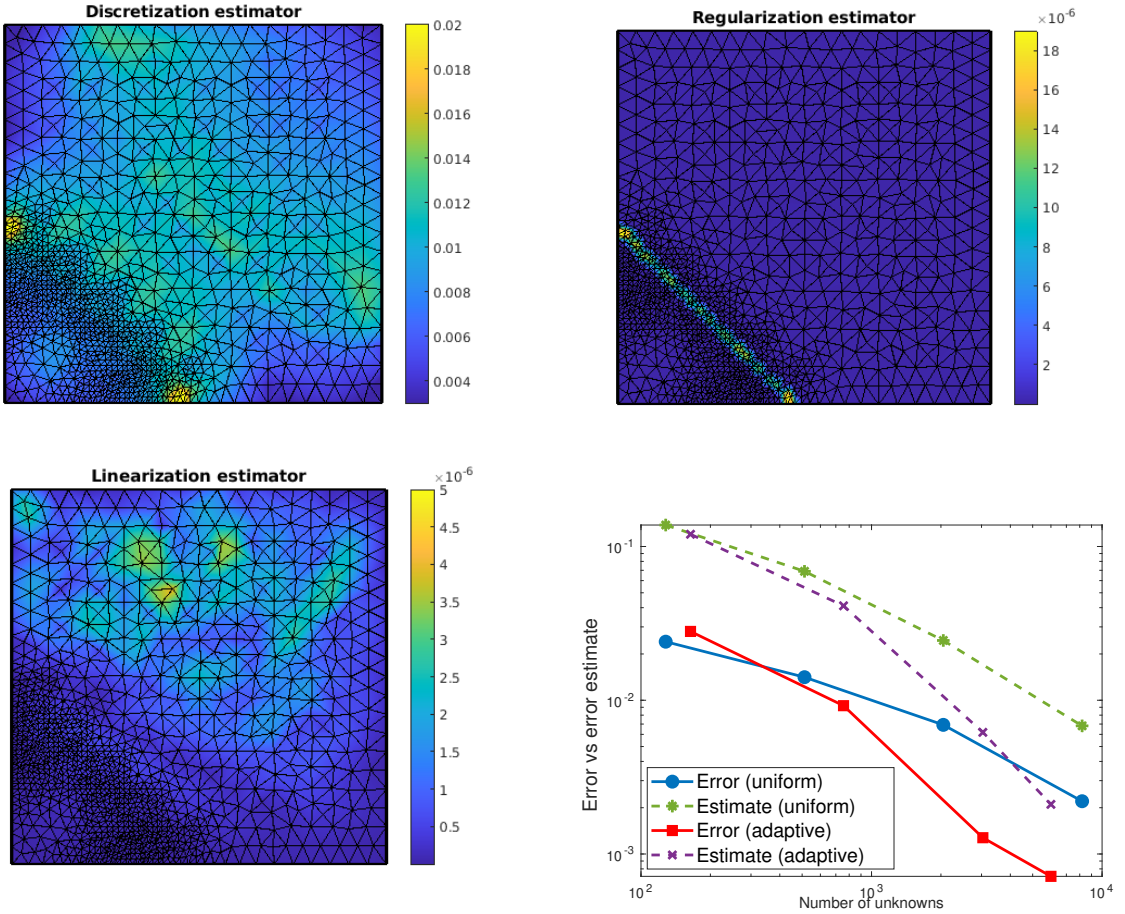

Figure 8: Spatial distribution of error estimators at the end of simulation: discretization (top left), regularization (top right), and linearization (bottom left). Convergence results comparing uniform and adaptive mesh refinements for $p = 1$ (bottom right).

Table 3 offers a comprehensive comparison of the L-scheme, Newton, and L-scheme-Newton solvers across various parameter values of m and different switching criteria. The obtained effectivity indices underscore the high quality and robustness of our estimators. Particularly, the estimators demonstrate their efficacy even in cases of reaction dominance (larger m). The consistent trend observed in both our numerically estimated errors and their resulting effectivity indices is an increase as m also increases. The near-singular regions near the boundary do not influence the estimators, highlighting their reliability. We observe that, in almost all cases, the L-scheme-Newton approach outperforms both the individual L-scheme and Newton methods in terms of the total number of iterations. Particularly for fine meshes and $m \geq 3$, Newton fails to converge, and the L-scheme alone requires a significant number of iterations. It is noteworthy that the optimal L parameter is found to be close to $L_\varphi^{\text{opt}} = L_\varphi/2 = m/2$. As an example, we plot in Figure 9 the total iterations number in function of Γ_{sw} for the case of $m = 4$ and $h = 1/16$. The results reveal that choosing Γ_{sw} larger or smaller than 0.25 leads to an increase in the number of nonlinear iterations. However, when opting for a small Γ_{sw}, the solver tends to exhibit enhanced stability, effectively mitigating oscillations in the Newton method (left figure). In conclusion, the search for the optimal switching parameter Γ_{sw} alongside the appropriate choice of L significantly enhances the efficiency and stability of the L-scheme-Newton solver.

5.3.2 Models with fast regime

When $0 < m < 1$, the model aligns with the fast diffusion case, a prevalent scenario in diverse fields such as plasma physics, the kinetic theory of gases, or fluid transportation in porous media [27]. In this regime,

m	$1/h$	Γ_{sw}	Γ_{lin}	Γ_{reg}	Adaptive Solver	Iterations	Estimator	Error	Eff. Idx
2	16	–	0.01	0.01	L-scheme	31			
		–	0.01	0.01	Newton	10	2.59e-01	2.16e-01	1.19
		0.95	0.01	0.01	L-scheme-Newton	6 (2+4)			
	32	–	0.01	0.01	L-scheme	60			
		–	0.01	0.01	Newton	10	1.24e-01	1.04e-01	1.18
		0.75	0.01	0.01	L-scheme-Newton	9 (6+3)			
3	16	–	0.01	0.01	L-scheme	84			
		–	0.01	0.01	Newton	12	3.29e-01	2.43e-01	1.35
		0.5	0.01	0.01	L-scheme-Newton	14 (4+10)			
	32	–	0.01	0.01	L-scheme	121			
		–	0.01	0.01	Newton	NO	1.54e-01	1.22e-01	1.26
		0.35	0.01	0.01	L-scheme-Newton	25 (16+9)			
4	16	–	0.01	0.01	L-scheme	112			
		–	0.01	0.01	Newton	NO	4.03e-01	2.79e-01	1.44
		0.25	0.01	0.01	L-scheme-Newton	22 (7+15)			
	32	–	0.01	0.01	L-scheme	144			
		–	0.01	0.01	Newton	NO	2.34e-01	1.72e-01	1.36
		0.15	0.01	0.01	L-scheme-Newton	32 (20+12)			

Table 3: Comprehensive comparison of nonlinear iterations, error estimates, exact errors, and effectivity indices across different values of m and mesh size h. The results for the L-scheme are obtained with an optimized L-parameter computed offline.

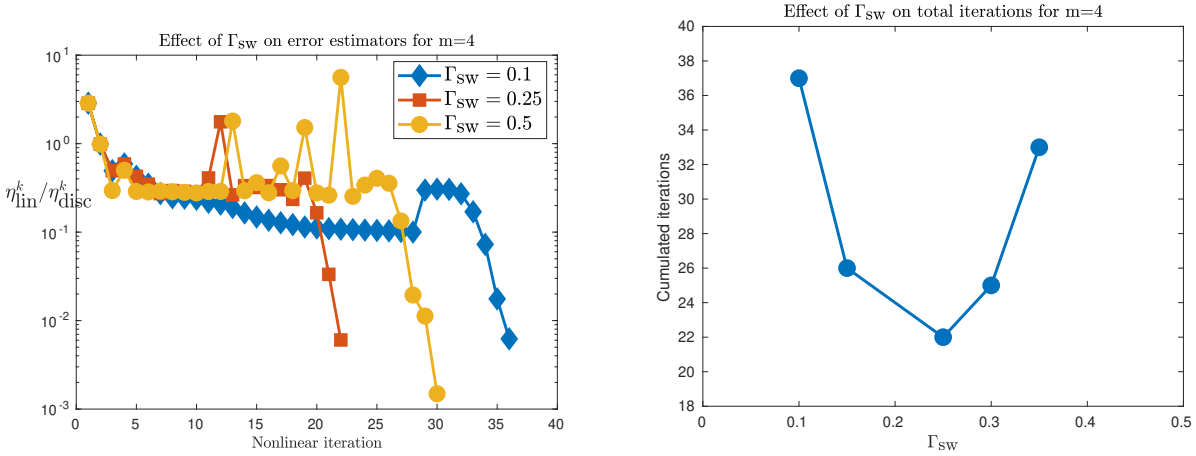

Figure 9: The ratio $\eta^k_{\text{lin}}/\eta^k_{\text{disc}}$ as a function of nonlinear iterations (left) for $\Gamma_{\text{sw}} = 0.1, 0.25,$ and 0.5. The results highlight the impact of Γ_{sw} on cumulative iterations (right).

the equation becomes singular as $|u|^{m-1}$ becomes unbounded whenever u approaches zero. This behavior indicates that the diffusion process undergoes rapid spreading, causing solutions to diffuse swiftly across the domain and potentially leading to sharp transitions and boundary layer effects. Here, as φ is not Lipschitz for arguments approaching 0, we contemplate regularization through $\varepsilon > 0$ and the introduction of the functions $\varphi_\varepsilon : \mathbb{R} \to \mathbb{R}$. Specifically, $\varphi_\varepsilon(u) = \varepsilon^{m-1}u$ if u lies in the interval $(0, \varepsilon)$ and $\varphi_\varepsilon(u) = \varphi(u)$ elsewhere. The Lipschitz constant is $L_{\varphi_\varepsilon} = \varepsilon^{m-1}$, and even φ'_ε is Lipschitz and $L_{b'_\varepsilon} = m(m-1)|\varepsilon|^{m-2}$. We can employ a transformation by defining $w = |u|^{m-1}u$ and addressing the problem for w using linear diffusion and a nonlinear reaction term represented by $|w|^{1/m-1}w$. This poses no challenges since $1/m > 1$. For this formulation, we apply the L-scheme without regularization.

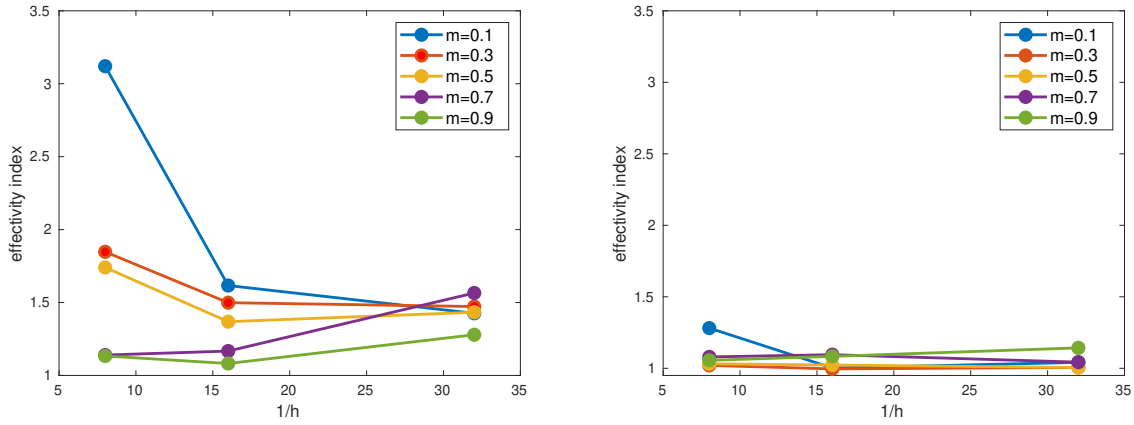

Figure 10: Effectivity index as a function of model parameter m and mesh size h for the original formulation (left) and the transformed formulation (right). These results were obtained utilizing the L-scheme with an optimized L-parameter computed offline.

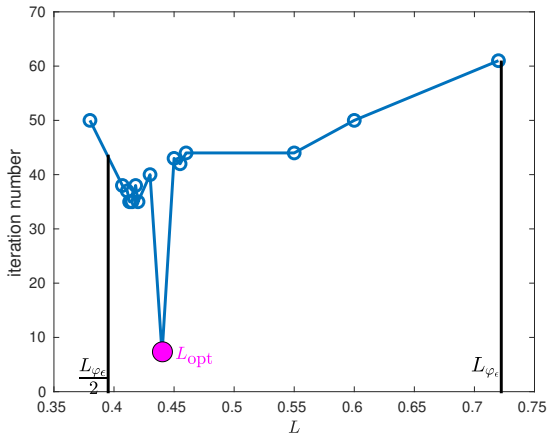

Figure 11: Offline computation of the optimal L-scheme parameter for the original formulation for $m = 0.5$. The results reaffirm the correlation between the performance of the L-scheme and the stabilization parameter L. Similar observations were made for $0 < m < 1$, confirming that L_{opt} falls within the range $[\frac{L_{\varphi_\varepsilon}}{2}, L_{\varphi_\varepsilon}]$.

Figure 10 presents a comparative analysis of effectivity indices for both formulations, considering varying values of m and mesh sizes. In the standard formulation (left figure), the estimator tends to overestimate the exact error, especially for small m values, causing the effectivity index to surpass the optimal value of 1. However, as m approaches 1 (linear case), the overestimation diminishes, and the effectivity index progressively converges towards 1. Overall, the effectivity index for the standard formulation remains within a reasonable range close to the optimal value of 1. In contrast, the transformed problem exhibits excellent results (right figure), maintaining an effectivity index consistently close to 1, irrespective of the m value. This highlights the enhanced properties of the transformed problem. Importantly, the transformed formulation demonstrates efficiency across various m values, requiring only a modest number of iterations—typically averaging 2 to 3 iterations. On the other hand, the original formulation demands a higher iteration count, ranging from 5 to 6 iterations when optimizing the stability parameter L_β. Without optimizing the L-scheme, the iteration count could significantly increase, as depicted in Figure 11. Figure 12 showcases the spatial distribution of discretization and linearization errors, along with estimators, for the original formulations with $m = 0.5$ after convergence. The results reveal a remarkable agreement between estimated and exact errors. This robust agreement underscores the effectiveness of employing these estimators to control nonlinearity and certify the accuracy of numerical simulations, especially in scenarios involving

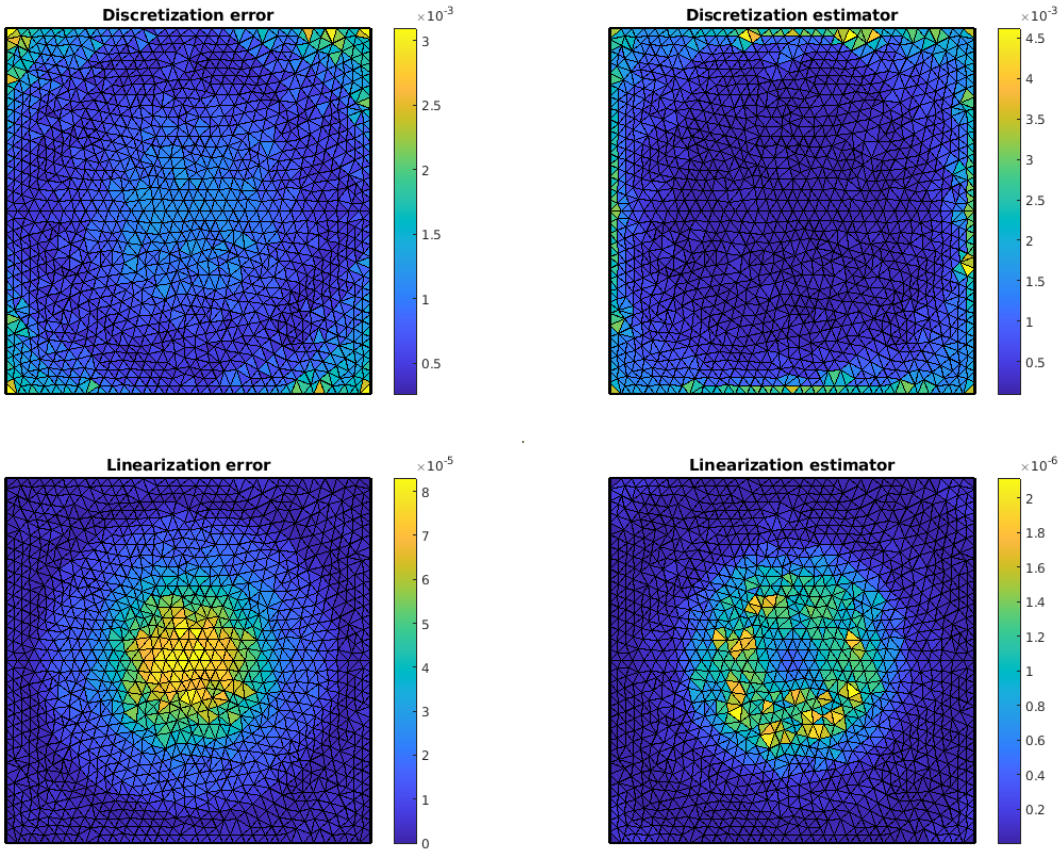

Figure 12: Comparison of discretization errors (top left) and estimators (top right), along with linearization errors (bottom left) and estimators (bottom right) using the L-scheme for $m = 0.5$. The discretization errors remain identical when employing the Newton method (results not shown).

degeneracy.

6 Conclusions

Our work presents an adaptive solver for a class of doubly nonlinear and degenerate elliptic equations, leveraging the strengths of the L-scheme and Newton methods. To this end, we developed a posteriori error estimators providing a guaranteed global upper bound on the error. We also proved the efficiency of the estimates and showed their robustness with respect to mesh size and nonlinearity strengths, and how to evaluate separately the various error components, including regularization and linearization errors.

In our numerical experiments focused on the Stefan problem and scenarios involving fast and slow diffusion in porous media, the adaptive L-scheme–Newton solver consistently outperforms the L-scheme and demonstrates greater stability compared to standard Newton methods, particularly in challenging convergence scenarios. The L-scheme's role as a robust initialization for Newton is crucial in mitigating oscillations and unnecessary iterations. The adaptive solver stands out by providing users with accurate error estimations for different error components at each nonlinear iteration.

Appendix A. Patchwise Euler-Lagrange Equations for Flux Equilibration

The a posteriori error estimates presented in this paper rely on a locally computable equilibrated flux, denoted as $\boldsymbol{\sigma}_h^{\epsilon,k}$, and the post-processed saturation, denoted as $\Phi_h^{\epsilon,k}$. For any $\mathbf{a} \in \mathcal{V}_h$, the local equilibrated flux $\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}$ and the local post-processed saturation $\Phi_{\mathbf{a}}^{\epsilon,k}$ are defined over vertex patches. These reconstructions correspond to the Euler-Lagrange equations resulting from the constrained minimization within vertex patches.

Definition A.1 (Reconstruction of $(\boldsymbol{\sigma}_{\mathbf{a}}^{\varepsilon,k}, \Phi_{\mathbf{a}}^{\varepsilon,k})$)**.** *For each $\mathbf{a} \in \mathcal{V}_h$, let $(\boldsymbol{\sigma}_{\mathbf{a}}^{\varepsilon,k}, \Phi_{\mathbf{a}}^{\varepsilon,k}) \in \mathbf{W}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$ be the minimizers of the following minimization problem:*

$$(\boldsymbol{\sigma}_{\mathbf{a}}^{\varepsilon,k}, \Phi_{\mathbf{a}}^{\varepsilon,k}) := \arg \min_{\substack{(\boldsymbol{\sigma}_h,\Phi_h)\in \mathbf{W}_h^{\mathbf{a}}\times Q_h^{\mathbf{a}} \\ \nabla\cdot\boldsymbol{\sigma}_h+\Phi_h=b}} \overbrace{\|\boldsymbol{\sigma}_h + \boldsymbol{\lambda}_{\mathbf{a}}\|_{\omega_{\mathbf{a}}}^2 + \delta_\star \left\|\Phi_h - \Pi_h(\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon^{k-1}(u_h^{\epsilon,k}))\right\|_{\omega_{\mathbf{a}}}^2}^{H(\boldsymbol{\sigma}_h,\Phi_h)}, \tag{A.1}$$

where $b := g_{\mathbf{a}} + \Pi_h(\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon^{k-1}(u_h^{\epsilon,k})) = \Pi_h(\psi_{\mathbf{a}} f) - \nabla\psi_{\mathbf{a}} \cdot (\Pi_h\boldsymbol{\xi} + \mathbf{I}_h^{\epsilon,k})$.

To solve the minimization problem A.1, we employ the method of Lagrange multipliers associated with the Lagrangian function:

$$\mathcal{L}(\boldsymbol{\sigma}_h, \Phi_h, r) = H(\boldsymbol{\sigma}_h, \Phi_h) - (\nabla\cdot\boldsymbol{\sigma}_h, r)_{\omega_{\mathbf{a}}} - (\Phi_h, r)_{\omega_{\mathbf{a}}} + (b, r)_{\omega_a} \quad \forall(\boldsymbol{\sigma}, \Phi, r) \in \mathbf{W}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}, \tag{A.2}$$

where r is the Lagrange multiplier used to impose the needed constraint on $(\boldsymbol{\sigma}_h, \Phi_h)$. Denote by $(\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, \Phi_{\mathbf{a}}^{\epsilon,k}, r_{\mathbf{a}}^{\epsilon,k})$ the saddle point of the Lagrangian function $\mathcal{L}$. Then, for all $(\mathbf{v}, r, q) \in \mathbf{W}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$, we have

$$\begin{aligned}
(\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, \mathbf{v})_{\omega_{\mathbf{a}}} - (\nabla\cdot\mathbf{v}, r_{\mathbf{a}}^{\epsilon,k})_{\omega_{\mathbf{a}}} &= (-\psi_{\mathbf{a}}(\boldsymbol{\xi} + \mathbf{I}_h^{\epsilon,k}), \mathbf{v})_{\omega_{\mathbf{a}}}, \\
(\Phi_{\mathbf{a}}^{\epsilon,k}, r)_{\omega_{\mathbf{a}}} &= \left(\Pi_h(\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon^{k-1}(u_h^{\epsilon,k})), r\right)_{\omega_{\mathbf{a}}} + \delta_\star(r_{\mathbf{a}}^{\epsilon,k}, r)_{\omega_{\mathbf{a}}}, \\
(\nabla\cdot\boldsymbol{\sigma}_{\mathbf{a}}^{\epsilon,k}, q)_{\omega_{\mathbf{a}}} + (\Phi_{\mathbf{a}}^{\epsilon,k}, q)_{\omega_{\mathbf{a}}} &= (\Pi_h(\psi_{\mathbf{a}} f), q)_{\omega_{\mathbf{a}}} - \left(\nabla\psi_{\mathbf{a}} \cdot (\Pi_h\boldsymbol{\xi} + \mathbf{I}_h^{\epsilon,k}), q\right)_{\omega_{\mathbf{a}}}.
\end{aligned}$$

The Euler-Lagrange equations can then be simplified to two coupled equations as given by (4.22), followed by a post-processing step:

$$\Phi_{\mathbf{a}}^{\epsilon,k} = \Pi_h(\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon^{k-1}(u_h^{\epsilon,k})) + \delta_\star r_{\mathbf{a}}^{\epsilon,k}, \quad \text{for } \delta_\star \in \{0, 1\}.$$

The global reconstructions are constructed afterward using the partition of unity property of the hat functions through (4.23).

Definition A.2 (Reconstruction of $(\boldsymbol{\sigma}_{\mathbf{a},\mathrm{disc}}^{\varepsilon,k}, \Phi_{\mathbf{a},\mathrm{disc}}^{\varepsilon,k})$)**.** *For each $\mathbf{a} \in \mathcal{V}_h$, let $(\boldsymbol{\sigma}_{\mathbf{a},\mathrm{disc}}^{\varepsilon,k}, \Phi_{\mathbf{a},\mathrm{disc}}^{\varepsilon,k}) \in W_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$ be the minimizers of the following minimization problem:*

$$(\boldsymbol{\sigma}_{\mathbf{a},\mathrm{disc}}^{\varepsilon,k}, \Phi_{\mathbf{a},\mathrm{disc}}^{\varepsilon,k}) := \arg \min_{\substack{(\boldsymbol{\sigma}_h,\Phi_h)\in \mathbf{W}_h^{\mathbf{a}}\times Q_h^{\mathbf{a}} \\ \nabla\cdot\boldsymbol{\sigma}_h+\Phi_h=b}} \overbrace{\left\|\boldsymbol{\sigma}_h + \boldsymbol{\lambda}_{\mathbf{a}}^{k,\epsilon}\right\|_{\omega_{\mathbf{a}}}^2 + \delta_\star \left\|\Phi_h - \Pi_h(\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon(u_h^{\epsilon,k}))\right\|_{\omega_{\mathbf{a}}}^2}^{H_{\mathrm{disc}}(\boldsymbol{\sigma}_h,\Phi_h)}, \tag{A.3}$$

where $b := g_{\mathbf{a}}^{\epsilon,k} + \Pi_h(\psi_{\mathbf{a}}\widetilde{\beta}_\epsilon(u_h^{\epsilon,k})) - \psi_{\mathbf{a}}\mathbf{r}_h^{\epsilon,k} = \Pi_h(\psi_{\mathbf{a}} f) - \nabla\psi_{\mathbf{a}} \cdot (\Pi_h\boldsymbol{\xi} + \nabla\widetilde{\varphi}_\epsilon(u_h^{\epsilon,k})) - \Pi_h(\psi_{\mathbf{a}}\mathbf{r}_h^{\epsilon,k})$.

Repeating the same steps as for (A.1), one can obtain the Euler-Lagrange equations (4.24). Therefore, the global constructions are obtained through (4.25).

Appendix B. Proof of Proposition 4.11

We establish the stability result (4.47) using the approach presented in [52] and, more specifically, in [31]. These references utilize stability bounds and quasi-minimizers for the patchwise discretization-flux equilibration problem (4.24).

Case with $\delta_\star = 0$: For this case, we recall that $\eta^{\epsilon,k}_{\mathrm{disc},K} = ||\boldsymbol{\sigma}^{\epsilon,k}_{h,\mathrm{disc}} + \Pi^{\mathrm{RTN}}_h \boldsymbol{\xi} + \nabla\widetilde{\varphi}_\epsilon(u^{\epsilon,k}_h)||_K$, which corresponds to an equilibration problem based on a pure diffusion problem. We also note that $\Phi^{\epsilon,k}_{h,\mathrm{disc}}$ is undefined by (4.22) and is not needed in the estimates.

By orthogonality of the L^2-projection operator Π^{RTN}_h onto $\mathbf{RTN}_p(\mathcal{T}^{\mathbf{a}}_h)$ together with (4.24a), we can see that $(\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}, \mathbf{v})_{\omega_\mathbf{a}} - (\nabla\cdot\mathbf{v}, r^{\epsilon,k}_{\mathbf{a},\mathrm{disc}})_{\omega_\mathbf{a}} = (\boldsymbol{\lambda}^{\epsilon,k}_\mathbf{a}, \mathbf{v})_{\omega_\mathbf{a}} = -(\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}, \mathbf{v})_{\omega_\mathbf{a}}$, for all $\mathbf{v} \in \mathbf{W}^{\mathbf{a}}_h$ and from (A.3) it's inferred that, for each $\mathbf{a} \in \mathcal{V}_h$,

$$\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} = \arg \min_{\substack{\mathbf{v}_h \in \mathbf{W}^{\mathbf{a}}_h \\ \nabla\cdot\mathbf{v}_h = g^{\epsilon,k}_\mathbf{a} - \Pi_h(\psi_\mathbf{a} \mathbf{r}^{\epsilon,k}_h)}} ||\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \mathbf{v}_h||_{\omega_\mathbf{a}}. \tag{B.1}$$

Recall that $(\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}, g^{\epsilon,k}_\mathbf{a} - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h)) \in \mathbf{RTN}_p(\mathcal{T}^{\mathbf{a}}_h) \times \mathbb{P}_p(\mathcal{T}^{\mathbf{a}}_h)$, and the Neumann compatibility condition (4.26) holds true for all interior vertices. Then, using [8, Thm 7] in the case of two space dimensions and [36, Thm 2.3] in the case of three space dimensions, we obtain, using (B.1) that

$$\begin{aligned} ||\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}||_{\omega_\mathbf{a}} &= \min_{\substack{\mathbf{v}_h \in \mathbf{W}^{\mathbf{a}}_h, \\ \nabla\cdot\mathbf{v}_h = g^{\epsilon,k}_\mathbf{a} - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h)}} ||\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \mathbf{v}_h||_{\omega_\mathbf{a}} \\ &\lesssim \max_{\substack{|\nabla v| = 1 \\ v \in H^1_+(\omega_\mathbf{a})\setminus\{0\}}} \left(g^{\epsilon,k}_\mathbf{a} - \psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h, v\right)_{\omega_\mathbf{a}} - \left(\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}, \nabla v\right)_{\omega_\mathbf{a}}. \end{aligned} \tag{B.2}$$

Case with $\delta_\star = 1$: We leverage the results obtained for $\delta_\star = 0$ along with techniques from [52] to verify the estimate (4.47) for this case. The idea relies on the fact that $(\boldsymbol{\sigma}^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}}, \Phi^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}})$ are minimizers of the problem A.3:

$$(\boldsymbol{\sigma}^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}}, \Phi^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}}) = \arg \min_{\substack{(\mathbf{v}_h, \Phi_h) \in \mathbf{W}^{\mathbf{a}}_h \times Q^{\mathbf{a}}_h \\ \nabla\cdot\mathbf{v}_h + \Phi_h \\ = g^{\epsilon,k}_\mathbf{a} + \Pi_h(\psi_\mathbf{a}\widetilde{\beta}_\epsilon(u^{\epsilon,k}_h)) - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h)}} \underbrace{\left\|\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \mathbf{v}_h\right\|^2_{\omega_\mathbf{a}} + \left\|\Phi_h - \Pi_h\big(\psi_\mathbf{a}\widetilde{\beta}_\epsilon(u^{\epsilon,k}_h)\big)\right\|^2_{\omega_\mathbf{a}}}_{H_{\mathrm{disc}}(\mathbf{v}_h, \Phi_h)}. \tag{B.3}$$

The functional $H_{\mathrm{disc}}(\mathbf{v}_h, \Phi_h) : \mathbf{W}^{\mathbf{a}}_h \times Q^{\mathbf{a}}_h \to \mathbb{R}$ is convex with respect to both variables. Therefore, by fixing $\Phi^{**}_h := \Pi_h\big(\psi_\mathbf{a}\widetilde{\beta}_\epsilon(u^{\epsilon,k}_h)\big)$, which is an admissible element of $Q^{\mathbf{a}}_h$, we obtain:

$$\begin{aligned} \min_{\mathbf{v}_h, \Phi_h} H_{\mathrm{disc}}(\mathbf{v}_h, \Phi_h) &\le \min_{\mathbf{v}_h} H_{\mathrm{disc}}(\mathbf{v}_h, \Phi^{**}_h), \\ &\overset{\text{(B.3)}}{=} \min_{\substack{\mathbf{v}_h \in \mathbf{W}^{\mathbf{a}}_h \\ \nabla\cdot\mathbf{v}_h = g^{\epsilon,k}_\mathbf{a} - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h)}} \left\|\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \mathbf{v}_h\right\|^2_{\omega_\mathbf{a}}. \end{aligned} \tag{B.4}$$

Here, the constraint is derived from $\nabla\cdot\mathbf{v}_h = g^{\epsilon,k}_\mathbf{a} + \Pi_h(\psi_\mathbf{a}\widetilde{\beta}_\epsilon(u^{\epsilon,k}_h)) - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h) - \Phi^{**}_h = g^{\epsilon,k}_\mathbf{a} - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h)$ (see Definition A.3). Let us define

$$\mathbf{v}^{**}_h = \arg \min_{\substack{\mathbf{v}_h \in \mathbf{W}^{\mathbf{a}}_h \\ \nabla\cdot\mathbf{v}_h = g^{\epsilon,k}_\mathbf{a} - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h)}} ||\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \mathbf{v}_h||_{\omega_\mathbf{a}}, \quad (\neq \boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}) \tag{B.5}$$

Here again, we observe that $(\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}, g^{\epsilon,k}_\mathbf{a} - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h)) \in \mathbf{RTN}_p(\mathcal{T}^{\mathbf{a}}_h) \times \mathbb{P}_p(\mathcal{T}^{\mathbf{a}}_h)$, and the Neumann compatibility condition $(g^{\epsilon,k}_\mathbf{a} - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h), 1)_{\omega_\mathbf{a}}$ for all interior vertices. Notably, we can verify that the constraint $\nabla\cdot\mathbf{v}^{**}_h + \Phi^{**}_h = g^{\epsilon,k}_\mathbf{a} + \Pi_h(\psi_\mathbf{a}\widetilde{\beta}_\epsilon(u^{\epsilon,k}_h)) - \Pi_h(\psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h)$ of problem (B.3) is met. The arguments used for (B.2) togethet with (B.5) implies

$$\begin{aligned} &\left\|\mathbf{v}^{**}_h + \boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}\right\|^2_{\omega_\mathbf{a}} + \overbrace{\left\|\Phi^{**} - \Pi_h\big(\psi_\mathbf{a}\widetilde{\beta}_\epsilon(u^{\epsilon,k}_h)\big)\right\|^2_{\omega_\mathbf{a}}}^{=0} \\ &\qquad \overset{\text{(B.5)}}{\lesssim} \left[\max_{\substack{|\nabla v| = 1 \\ v \in H^1_+(\omega_\mathbf{a})\setminus\{0\}}} \left(g^{\epsilon,k}_\mathbf{a} - \psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h, v\right)_{\omega_\mathbf{a}} - \left(\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}, \nabla v\right)_{\omega_\mathbf{a}}\right]^2. \end{aligned} \tag{B.6}$$

Since $(\boldsymbol{\sigma}^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}}, \Phi^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}})$ are minimizers of problem A.3, we conclude first that above estimate holds true as well for $(\boldsymbol{\sigma}^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}}, \Phi^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}})$. The proof is complete.

Appendix C. Proof of Theorem 4.12

We choose $\mathbf{a} \in \mathcal{V}_K$ and consider $v \in H^1_+(\omega_\mathbf{a})$ then we express the right-hand side of (4.47) as $\Big(g^{\epsilon,k}_\mathbf{a} - \psi_\mathbf{a}\mathbf{r}^{\epsilon,k}_h, v\Big)_{\omega_\mathbf{a}} - \Big(\boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}, \nabla v\Big)_{\omega_\mathbf{a}} = \sum_{i=1}^5 T_i$ where

$$
\begin{aligned}
T_1 &:= (f, \psi_\mathbf{a} v)_{\omega_\mathbf{a}} - (\widetilde{\beta}_\epsilon(u^{\epsilon,k}_h), \psi_\mathbf{a} v)_{\omega_\mathbf{a}} - (\boldsymbol{\xi} + \nabla\widetilde{\varphi}_\epsilon(u^{\epsilon,k}_h), \nabla(\psi_\mathbf{a} v))_{\omega_\mathbf{a}},\\
T_2 &:= (\Pi_h(\psi_\mathbf{a} f) - \psi_\mathbf{a} f, v)_{\omega_\mathbf{a}},\\
T_3 &:= (\boldsymbol{\xi} - \Pi_h\boldsymbol{\xi}, v\,\nabla\psi_\mathbf{a})_{\omega_\mathbf{a}},\\
T_4 &:= (\psi_\mathbf{a}\boldsymbol{\xi} - \Pi^{\mathrm{RTN}}_h(\psi_\mathbf{a}\boldsymbol{\xi}), \nabla v)_{\omega_\mathbf{a}},\\
T_5 &:= -\Big(\mathbf{r}^{\epsilon,k}_h, \psi_\mathbf{a} v\Big)_{\omega_\mathbf{a}}.
\end{aligned}
$$

Next, we can easily verify that $|T_2| = \sum_{K\in\mathcal{T}^\mathbf{a}_h}(\Pi_h(\psi_\mathbf{a} f) - \psi_\mathbf{a} f, v - \Pi_h v)_K$ using the orthogonality of the L^2-projection so that $|T_2| \leq \sum_{K\in\mathcal{T}^\mathbf{a}_h} \dfrac{h_K}{\pi}||\Pi_h(\psi_\mathbf{a} f) - \psi_\mathbf{a} f||_K||\nabla v||_K$ by using the Cauchy-Schwarz and Poincaré–Friedrichs inequalities, as demonstrated in the proof of Theorem 4.10. Similarly, we can prove that $|T_3| \leq ||\boldsymbol{\xi} - \Pi_h\boldsymbol{\xi}||_{\omega_\mathbf{a}}||\nabla v||_{\omega_\mathbf{a}}$ and $|T_4| \leq ||\psi_\mathbf{a}\boldsymbol{\xi} - \Pi_h(\psi_\mathbf{a}\boldsymbol{\xi})||_{\omega_\mathbf{a}}||\nabla v||_{\omega_\mathbf{a}}$, so that using (4.45)

$$
|T_2| + |T_3| + |T_4| \lesssim \overbrace{\left\{\sum_{K\in\mathcal{T}^\mathbf{a}_h} \frac{h_K}{\pi}||\Pi_h(\psi_\mathbf{a} f) - \psi_\mathbf{a} f||_K + ||\boldsymbol{\xi} - \Pi_h\boldsymbol{\xi}||_K + ||\psi_\mathbf{a}\boldsymbol{\xi} - \Pi_h(\psi_\mathbf{a}\boldsymbol{\xi})||_K\right\}}^{\tilde{\eta}^{\epsilon,k}_{\mathrm{osc},\omega_\mathbf{a}}} ||\nabla v||_{\omega_\mathbf{a}} \quad \text{(C.1)}
$$

To bound T_5, we use the Cauchy-Schwarz inequality and Poincaré–Friedrichs inequality to first get $|T_5| \lesssim h_{\omega_\mathbf{a}}||\mathbf{r}^{\epsilon,k}_h||_{\omega_\mathbf{a}}||\nabla v||_{\omega_\mathbf{a}}$. Now, we use property (2.b.), the triangle inequality, followed by an inverse inequality to get

$$
\begin{aligned}
||\mathbf{r}^{\epsilon,k}_h||_K &= ||\nabla\cdot\boldsymbol{\sigma}^{\epsilon,k}_{h,\mathrm{lin}} + \Phi^{\epsilon,k}_{h,\mathrm{lin}}||_K,\\
&\leq \sum_{K\in\mathcal{T}^\mathbf{a}_h} ||\nabla\cdot\boldsymbol{\sigma}^{\epsilon,k}_{h,\mathrm{lin}}||_K + ||\Phi^{\epsilon,k}_{h,\mathrm{lin}}||_K,\\
&\lesssim \sum_{K\in\mathcal{T}^\mathbf{a}_h} h_K^{-1}||\boldsymbol{\sigma}^{\epsilon,k}_{h,\mathrm{lin}}||_K + ||\Phi^{\epsilon,k}_{h,\mathrm{lin}}||_K. \qquad \text{(C.2)}
\end{aligned}
$$

That is we can conclude that

$$
\begin{aligned}
T_5 &\lesssim \left\{\sum_{K\in\mathcal{T}^\mathbf{a}_h} ||\boldsymbol{\sigma}^{\epsilon,k}_{h,\mathrm{lin}}||_K + \frac{h_K}{\pi}||\Phi^{\epsilon,k}_{h,\mathrm{lin}}||_K\right\}||\nabla v||_{\omega_\mathbf{a}}\\
&\lesssim \underbrace{\left\{\sum_{K\in\mathcal{T}^\mathbf{a}_h} ||\boldsymbol{\sigma}^{\epsilon,k}_{h,\mathrm{lin}}||_K + \kappa_\star||\Phi^{\epsilon,k}_{h,\mathrm{lin}}||_K\right\}}_{\eta^{\epsilon,k}_{\mathrm{lin},\omega_\mathbf{a}}} ||\nabla v||_{\omega_\mathbf{a}}, \qquad \text{(C.3)}
\end{aligned}
$$

where the constant depends only on the mesh regularity. Now we are left with the bound of T_1. As $\psi_\mathbf{a} v \in H^1_0(\Omega)$, we obtain first from (2.2) that

$$
T_1 = \Big(\beta(u) - \widetilde{\beta}_\epsilon(u^{\epsilon,k}_h), \psi_\mathbf{a} v\Big)_{\omega_\mathbf{a}} + \Big(\nabla\big(\varphi(u) - \widetilde{\varphi}_\epsilon(u^{\epsilon,k}_h)\big), \nabla(\psi_\mathbf{a} v)\Big)_{\omega_\mathbf{a}}.
$$

Now, we proceed as in the proof of Theorem 4.10 by (1) using the triangle inequality and (4.1) to incorporate the regularization and quadrature estimators (2) applying the Cauchy-Schwarz and Poincaré–Friedrichs inequalities to bound each term (3) employing the mesh shape-regularity, yielding $h_{\omega_{\mathbf{a}}} \approx h_K$ and the fact that $||\nabla(\psi_{\mathbf{a}} v)||_{\omega_{\mathbf{a}}} \lesssim ||\nabla v||_{\omega_{\mathbf{a}}}$:

$$T_1 \lesssim \langle \mathcal{R}_u(u_h), \psi_{\mathbf{a}} v \rangle + \left[\eta^{\epsilon,k}_{\mathrm{reg},\omega_{\mathbf{a}}} + \eta^{\epsilon,k}_{\mathrm{qd},\omega_{\mathbf{a}}} \right] ||\nabla v||_{\omega_{\mathbf{a}}}.$$

By combining the above inequalities in (B.2), we find out that

$$\begin{aligned}
\left\{ ||\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}||^2_{\omega_{\mathbf{a}}} + \delta_\star \left\| r^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}} \right\|^2_{\omega_{\mathbf{a}}} \right\}^{\frac{1}{2}} &\lesssim \max_{v \in H^1_0(\omega_{\mathbf{a}}) \setminus \{0\}} \frac{\langle \mathcal{R}_u(u_h), \psi_{\mathbf{a}} v \rangle}{|\nabla v|} + \left[\eta^{\epsilon,k}_{\mathrm{lin},\omega_{\mathbf{a}}} + \eta^{\epsilon,k}_{\mathrm{reg},\omega_{\mathbf{a}}} + \eta^{\epsilon,k}_{\mathrm{qd},\omega_{\mathbf{a}}} + \eta^{\epsilon,k}_{\mathrm{osc},\omega_{\mathbf{a}}} \right], \\
&\lesssim \max_{v \in H^1_+(\omega_{\mathbf{a}}) \setminus \{0\}} \left[\frac{\langle \mathcal{R}_u(u_h), \psi_{\mathbf{a}} v \rangle}{||\psi_{\mathbf{a}} v||_{\star,\omega_{\mathbf{a}}}} \times \frac{||\psi_{\mathbf{a}} v||_{\star,\omega_{\mathbf{a}}}}{||\nabla v||_{\omega_{\mathbf{a}}}} \right] + \sum_{\bullet} \eta^{\epsilon,k}_{\bullet,\omega_{\mathbf{a}}}, \\
&\lesssim \max_{v \in H^1_0(\omega_{\mathbf{a}}) \setminus \{0\}} \frac{\langle \mathcal{R}_u(u_h), \psi_{\mathbf{a}} v \rangle}{||\psi_{\mathbf{a}} v||_{\star,\omega_{\mathbf{a}}}} + \sum_{\bullet} \eta^{\epsilon,k}_{\bullet,\omega_{\mathbf{a}}}, \qquad \text{(C.4)}
\end{aligned}$$

where we used simply (4.42) for the last inequality for the case of $\delta_\star = 0$ and the fact that the $||\psi_{\mathbf{a}} v||_{\star,\omega_{\mathbf{a}}} \lesssim (1+h_{\omega_{\mathbf{a}}})||\nabla v||_{\omega_{\mathbf{a}}}$ by combining (4.43)-(4.44) for $\delta_\star = 1$. The last inequality corresponds to (4.48) in which the constant in $\lesssim$ depends only the space dimension d, the shape-regularity constant of $\mathcal{T}_h$, and the polynomial degree p. Now, to prove the local efficiency (4.49), we use the partition of unity to get

$$\begin{aligned}
\left[\eta^{\epsilon,k}_{\mathrm{disc},K} \right]^2 &\lesssim \sum_{\mathbf{a} \in \mathcal{V}_K} ||\boldsymbol{\sigma}^{\epsilon,k}_{h,\mathrm{disc}} + \Pi^{\mathrm{RTN}}_h(\psi_{\mathbf{a}} \boldsymbol{\xi}) + \psi_{\mathbf{a}} \nabla \widetilde{\varphi}_\epsilon(u^{\epsilon,k}_h)||^2_K + \delta_\star ||r^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}||^2_K, \\
&\leq \sum_{\mathbf{a} \in \mathcal{V}_K} ||\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}||^2_K + \delta_\star \left\| r^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}} \right\|^2_K, \\
&\leq \sum_{\mathbf{a} \in \mathcal{V}_K} ||\boldsymbol{\sigma}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}} + \boldsymbol{\tau}^{\epsilon,k}_{\mathbf{a},\mathrm{disc}}||^2_{\omega_{\mathbf{a}}} + \delta_\star \left\| r^{\varepsilon,k}_{\mathbf{a},\mathrm{disc}} \right\|^2_{\omega_{\mathbf{a}}}.
\end{aligned}$$

W replace the above inequality in the patchwise efficiency estimate (4.48), and proceeding by using the stopping criteria (3.8) and (3.9) with small enough user-parameters $(\Gamma_{\mathrm{lin}}, \Gamma_{\mathrm{reg}})$ to bound $\eta^{\epsilon,k}_{\mathrm{lin},\tilde{K}} + \eta^{\epsilon,k}_{\mathrm{reg},\tilde{K}}$ with $\eta^{\epsilon,k}_{\mathrm{disc},\tilde{K}}$ for all $\tilde{K} \in \mathcal{T}^{\mathbf{a}}_h$. We conclude the estimate by employing the balancing criteria (3.10) so that $\eta^{\epsilon,k}_{\mathrm{disc},\tilde{K}} \approx \eta^{\epsilon,k}_{\mathrm{disc},K}$ for all $\tilde{K} \in \mathcal{T}^{\mathbf{a}}_h$. The proof is complete.

References

[1] E. Ahmed, S. A. Hassan, C. Japhet, M. Kern, and M. V. k, *A posteriori error estimates and stopping criteria for space-time domain decomposition for two-phase flow between different rock types*, The SMAI journal of computational mathematics, 5 (2019), pp. 195–227, https://doi.org/10.5802/smai-jcm.47, https://doi.org/10.5802/smai-jcm.47.

[2] E. Ahmed, J. M. Nordbotten, and F. A. Radu, *Adaptive asynchronous time-stepping, stopping criteria, and a posteriori error estimates for fixed-stress iterative schemes for coupled poromechanics problems*, Journal of Computational and Applied Mathematics, 364 (2020), p. 112312, https://doi.org/10.1016/j.cam.2019.06.028, http://www.sciencedirect.com/science/article/pii/S0377042719303097.

[3] T. Arbogast and M. F. Wheeler, *A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media*, SIAM Journal on Numerical Analysis, 33 (1996), pp. 1669–1687, https://doi.org/10.1137/s0036142994266728, https://doi.org/10.1137/s0036142994266728.

[4] M. Arioli, D. Loghin, and A. Wathen, *Stopping criteria for iterations in finite element methods*, Numerische Mathematik, 99 (2005), pp. 381–410, https://doi.org/10.1007/s00211-004-0568-z, https://doi.org/10.1007/s00211-004-0568-z.

[5] R. Becker, M. Brunner, M. Innerberger, J. M. Melenk, and D. Praetorius, *Cost-optimal adaptive iterative linearized fem for semilinear elliptic pdes*, ESAIM: Mathematical Modelling and Numerical Analysis, 57 (2023), pp. 2193–2225, https://doi.org/10.1051/m2an/2023036, https://doi.org/10.1051/m2an/2023036.

[6] L. Bergamaschi and M. Putti, *Mixed finite elements and newton-type linearizations for the solution of richard's equation*, International Journal for Numerical Methods in Engineering, 45 (1999), pp. 1025–1046, https://doi.org/10.1002/(SICI)1097-0207(19990720)45:8<1025::AID-NME615>3.0.CO;2-G, https://onlinelibrary.wiley.com/doi/abs/10.1002, https://arxiv.org/abs/https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291097-0207%2819990720%2945%3A8%3C1025%3A%3AAID-NME615%3E3.0.CO%3B2-G.

[7] E. Bonetti and G. Schimperna, *Local existence for frémond's model of damage in elastic materials*, Continuum Mechanics and Thermodynamics, 16 (2004), pp. 319–335, https://doi.org/https://doi.org/10.1007/s00161-003-0152-2.

[8] D. Braess, V. Pillwein, and J. Schöberl, *Equilibrated residual error estimates are p-robust*, Computer Methods in Applied Mechanics and Engineering, 198 (2009), pp. 1189–1197, https://doi.org/https://doi.org/10.1016/j.cma.2008.12.010, https://www.sciencedirect.com/science/article/pii/S0045782508004441. HOFEM07.

[9] K. Brenner and C. Cancès, *Improving newton's method performance by parametrization: The case of the richards equation*, SIAM Journal on Numerical Analysis, 55 (2017), pp. 1760–1785, https://doi.org/10.1137/16M1083414, https://doi.org/10.1137/16M1083414, https://arxiv.org/abs/https://doi.org/10.1137/16M1083414.

[10] M. K. Brun, E. Ahmed, J. M. Nordbotten, and N. ChrStenseth, *Modeling the process of speciation using a multiscale framework including a posteriori error estimates*, SIAM Journal on Applied Mathematics, 82 (2022), pp. 450–475, https://doi.org/10.1137/21M1405228, https://doi.org/10.1137/21M1405228, https://arxiv.org/abs/https://doi.org/10.1137/21M1405228.

[11] C. Cancès, F. Nabet, and M. Vohralík, *Convergence and a posteriori error analysis for energy-stable finite element approximations of degenerate parabolic equations*, Mathematics of Computation, 90 (2021), pp. 517–563, https://doi.org/10.1090/mcom/3577, https://doi.org/10.1090/mcom/3577.

[12] E. Cancès, G. Dusson, Y. Maday, B. Stamm, and M. Vohralík, *Guaranteed and robust a posteriori bounds for laplace eigenvalues and eigenvectors: Conforming approximations*, SIAM Journal on Numerical Analysis, 55 (2017), pp. 2228–2254, https://doi.org/10.1137/15M1038633, https://doi.org/10.1137/15M1038633, https://arxiv.org/abs/https://doi.org/10.1137/15M1038633.

[13] J. Carrillo, *Entropy solutions for nonlinear degenerate problems*, Archive for Rational Mechanics and Analysis, 147 (1999), pp. 269–361, https://doi.org/10.1007/s002050050152, https://doi.org/10.1007/s002050050152.

[14] J. Carrillo, *Entropy solutions for nonlinear degenerate problems*, Archive for rational mechanics and analysis, 147 (1999), pp. 269–361, https://doi.org/10.1007/s002050050152, https://doi.org/10.1007/s002050050152.

[15] L. Chamoin and F. Legoll, *A pedagogical review on a posteriori error estimation in finite element computations*, 2021, https://arxiv.org/abs/2110.02160.

[16] I. Cheddadi, R. Fűi k, M. I. Prieto, and M. Vohralí k, *Guaranteed and robust a posteriori error estimates for singularly perturbed reaction-diffusion problems*, ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique, 43 (2009), pp. 867–888, https://doi.org/10.1051/m2an/2009012, www.numdam.org/item/M2AN_2009__43_5_867_0/.

[17] Z. Chen and R. E. Ewing, *Fully discrete finite element analysis of multiphase flow in groundwater hydrology*, SIAM Journal on Numerical Analysis, 34 (1997), p. 2228–2253, https://doi.org/10.1137/S0036142995290063, https://doi.org/10.1137/S0036142995290063, https://arxiv.org/abs/https://doi.org/10.1137/S0036142995290063.

[18] P. Ciarlet and M. Vohralík, *Localization of global norms and robust a posteriori error control for transmission problems with sign-changing coefficients*, ESAIM: Mathematical Modelling and Numerical Analysis, 52 (2018), pp. 2037–2064, https://doi.org/10.1051/m2an/2018034, https://doi.org/10.1051/m2an/2018034.

[19] A. Cohen, R. DeVore, and R. H. Nochetto, *Convergence rates of AFEM with h -1 data*, Foundations of Computational Mathematics, 12 (2012), pp. 671–718, https://doi.org/10.1007/s10208-012-9120-1, https://doi.org/10.1007/s10208-012-9120-1.

[20] D. A. Di Pietro, M. Vohralík, and S. Yousef, *Adaptive regularization, linearization, and discretization and a posteriori error control for the two-phase Stefan problem*, Math. Comp., 84 (2015), pp. 153–186, https://doi.org/10.1090/S0025-5718-2014-02854-8, https://doi.org/10.1090/S0025-5718-2014-02854-8.

[21] V. Dolejší and H.-G. Shin, *A posteriori error estimate and mesh adaptation for the numerical solution of the richards equation*, in Spectral and High Order Methods for Partial Differential Equations ICOSAHOM 2020+1, J. M. Melenk, I. Perugia, J. Schöberl, and C. Schwab, eds., Cham, 2023, Springer International Publishing, pp. 209–223.

[22] V. Dolejší, A. Ern, and M. Vohralík, *hp-adaptation driven by polynomial-degree-robust a posteriori error estimates for elliptic problems*, SIAM Journal on Scientific Computing, 38 (2016), pp. A3220–A3246, https://doi.org/10.1137/15M1026687, https://doi.org/10.1137/15M1026687, https://arxiv.org/abs/https://doi.org/10.1137/15M1026687.

[23] J. Droniou and R. Eymard, *Uniform-in-time convergence of numerical methods for non-linear degenerate parabolic equations*, Numerische Mathematik, 132 (2016), pp. 721–766, https://doi.org/10.1007/s00211-015-0733-6, https://doi.org/10.1007/s00211-015-0733-6.

[24] J. Droniou and R. Eymard, *High-order mass-lumped schemes for nonlinear degenerate elliptic equations*, SIAM Journal on Numerical Analysis, 58 (2020), pp. 153–188, https://doi.org/10.1137/19m1244500, https://doi.org/10.1137/19m1244500.

[25] J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin, *The gradient discretisation method*, vol. 82, Springer, 2018.

[26] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin, *Non-conforming Finite Elements on Polytopal Meshes*, Springer International Publishing, Cham, 2021, pp. 1–35, https://doi.org/10.1007/978-3-030-69363-3_1, https://doi.org/10.1007/978-3-030-69363-3_1.

[27] J. Droniou and K.-N. Le, *The gradient discretization method for slow and fast diffusion porous media equations*, SIAM Journal on Numerical Analysis, 58 (2020), pp. 1965–1992, https://doi.org/10.1137/19m1260165, https://doi.org/10.1137/19m1260165.

[28] L. El Alaoui, A. Ern, and M. Vohralík, *Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems*, Computer Methods in Applied Mechanics and Engineering, 200 (2011), pp. 2782–2795, https://doi.org/https://doi.org/10.1016/j.cma.2010.03.024, https://www.sciencedirect.com/science/article/pii/S0045782510001015. Special Issue on Modeling Error Estimation and Adaptive Modeling.

[29] C. L. Epstein and R. Mazzeo, *Degenerate Diffusion Operators Arising in Population Biology (AM-185)*, vol. 185, Princeton University Press, 2013.

[30] A. Ern and J.-L. Guermond, *Eléments finis: théorie, applications, mise en oeuvre*, vol. 36, Springer Science & Business Media, 2002.

[31] A. Ern, I. Smears, and M. Vohralík, *Discrete p-robust* $\mathbf{H}(\mathrm{div})$*-liftings and a posteriori estimates for elliptic problems with* H^{-1} *source terms*, Calcolo, 54 (2017), pp. 1009–1025, https://doi.org/10.1007/s10092-017-0217-4, https://doi.org/10.1007/s10092-017-0217-4.

[32] A. Ern, I. Smears, and M. Vohralík, *Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems*, SIAM Journal on Numerical Analysis, 55 (2017), pp. 2811–2834, https://doi.org/10.1137/16m1097626, https://doi.org/10.1137/16m1097626.

[33] A. Ern, A. F. Stephansen, and M. Vohralík, *Guaranteed and robust discontinuous galerkin a posteriori error estimates for convection–diffusion–reaction problems*, Journal of Computational and Applied Mathematics, 234 (2010), pp. 114–130, https://doi.org/10.1016/j.cam.2009.12.009, http://www.sciencedirect.com/science/article/pii/S0377042709008176.

[34] A. Ern and M. Vohralík, *A posteriori error estimation based on potential and flux reconstruction for the heat equation*, SIAM J. Numer. Anal., 48 (2010), pp. 198–223, https://doi.org/10.1137/090759008, https://doi.org/10.1137/090759008.

[35] A. Ern and M. Vohralík, *Adaptive inexact newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs*, SIAM Journal on Scientific Computing, 35 (2013), pp. A1761–A1791, https://doi.org/10.1137/120896918, https://doi.org/10.1137/120896918.

[36] A. Ern and M. Vohralík, *Stable broken h1 and h (div) polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions.*, Math. Comput., 89 (2020), pp. 551–594.

[37] R. Eymard, P. Féron, T. Gallouët, R. Herbin, and C. Guichard, *Gradient schemes for the stefan problem*, International Journal On Finite Volumes, (2013), pp. Volume–10.

[38] F. Févotte, A. Rappaport, and M. Vohralík, *Adaptive regularization, discretization, and linearization for nonsmooth problems based on primal–dual gap estimators*, Computer Methods in Applied Mechanics and Engineering, 418 (2024), p. 116558, https://doi.org/https://doi.org/10.1016/j.cma.2023.116558, https://www.sciencedirect.com/science/article/pii/S0045782523006825.

[39] C. Galusinski and M. Saad, *On a degenerate parabolic system for compressible, immiscible, two-phase flows in porous media*, Advances in Differential Equations, 9 (2004), pp. 1235 – 1278, https://doi.org/10.57262/ade/1355867902, https://doi.org/10.57262/ade/1355867902.

[40] V. Girault, B. Riviere, and L. Cappanera, *A finite element method for degenerate two-phase flow in porous media. part i: Well-posedness*, Journal of Numerical Mathematics, 29 (2021), pp. 81–101, https://doi.org/doi:10.1515/jnma-2020-0004, https://doi.org/10.1515/jnma-2020-0004.

[41] M. E. Gurtin and R. C. MacCamy, *On the diffusion of biological populations*, Mathematical Biosciences, 33 (1977), pp. 35–49, https://doi.org/https://doi.org/10.1016/0025-5564(77)90062-1, https://www.sciencedirect.com/science/article/pii/0025556477900621.

[42] A. Haberl, D. Praetorius, S. Schimanko, and M. Vohralík, *Convergence and quasi-optimal cost of adaptive algorithms for nonlinear operators including iterative linearization and algebraic solver*, Numerische Mathematik, 147 (2021), pp. 679–725, https://doi.org/10.1007/s00211-021-01176-w, https://doi.org/10.1007/s00211-021-01176-w.

[43] W. Hao, *A gradient descent method for solving a system of nonlinear equations*, Applied Mathematics Letters, 112 (2021), p. 106739, https://doi.org/10.1016/j.aml.2020.106739, https://doi.org/10.1016/j.aml.2020.106739.

[44] A. Harnist, K. Mitra, A. Rappaport, and M. Vohralík, *Robust energy a posteriori estimates for nonlinear elliptic problems.* working paper or preprint, Dec. 2023, https://hal.science/hal-04033438.

[45] D. Illiano, I. S. Pop, and F. A. Radu, *Iterative schemes for surfactant transport in porous media*, Computational Geosciences, (2020), https://doi.org/10.1007/s10596-020-09949-2, https://doi.org/10.1007/s10596-020-09949-2.

[46] F. List and F. A. Radu, *A study on iterative methods for solving richards' equation*, Computational Geosciences, 20 (2016), pp. 341–353, https://doi.org/10.1007/s10596-016-9566-3, https://doi.org/10.1007/s10596-016-9566-3.

[47] K. Mitra and I. Pop, *A modified l-scheme to solve nonlinear diffusion problems*, Computers & Mathematics with Applications, 77 (2019), pp. 1722–1738, https://doi.org/10.1016/j.camwa.2018.09.042, http://www.sciencedirect.com/science/article/pii/S0898122118305546. 7th International Conference on Advanced Computational Methods in Engineering (ACOMEN 2017).

[48] K. Mitra and M. Vohralík, *Guaranteed, locally efficient, and robust a posteriori estimates for nonlinear elliptic problems in iteration-dependent norms. An orthogonal decomposition result based on iterative linearization*. working paper or preprint, July 2023, https://inria.hal.science/hal-04156711.

[49] R. Nochetto, M. Paolini, and C. Verdi, *An adaptive finite element method for two-phase stefan problems in two space dimensions. i. stability and error estimates*, Mathematics of Computation, 57 (1991), pp. 73–108, https://doi.org/10.1090/S0025-5718-1991-1079028-X, https://doi.org/10.1090/S0025-5718-1991-1079028-X.

[50] I. Pop, F. Radu, and P. Knabner, *Mixed finite elements for the richards' equation: linearization procedure*, Journal of Computational and Applied Mathematics, 168 (2004), pp. 365–373, https://doi.org/10.1016/j.cam.2003.04.008, https://doi.org/10.1016/j.cam.2003.04.008.

[51] M. Slodicka, *A robust and efficient linearization scheme for doubly nonlinear and degenerate parabolic problems arising in flow in porous media*, SIAM Journal on Scientific Computing, 23 (2002), pp. 1593–1614, https://doi.org/10.1137/s1064827500381860, https://doi.org/10.1137/s1064827500381860.

[52] Smears, Iain and Vohralík, Martin, *Simple and robust equilibrated flux a posteriori estimates for singularly perturbed reaction-diffusion problems*, ESAIM: M2AN, 54 (2020), pp. 1951–1973, https://doi.org/10.1051/m2an/2020034, https://doi.org/10.1051/m2an/2020034.

[53] J. S. Stokke, K. Mitra, E. Storvik, J. W. Both, and F. A. Radu, *An adaptive solution strategy for richards' equation*, 2023, https://arxiv.org/abs/2301.02055.

[54] E. Storvik, J. W. Both, K. Kumar, J. M. Nordbotten, and F. A. Radu, *On the optimization of the fixed-stress splitting for biot's equations*, International Journal for Numerical Methods in Engineering, 120 (2019), pp. 179–194, https://doi.org/https://doi.org/10.1002/nme.6130, https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.6130, https://arxiv.org/abs/https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.6130.

[55] G. A. Truskey, F. Yuan, and D. F. Katz, *Transport phenomena in biological systems*, Pearson/Prentice Hall Upper Saddle River, NJ, USA:, 2004.

[56] C. S. Woodward and C. N. Dawson, *Analysis of expanded mixed finite element methods for a nonlinear parabolic equation modeling flow into variably saturated porous media*, SIAM Journal on Numerical Analysis, 37 (2000), pp. 701–724.

[57] X.-S. Yang, *Nature-inspired optimization algorithms*, Academic Press, 2020.

[58] W.-A. Yong and I. S. Pop, *A numerical approach to porous medium equations*, Citeseer, 1996.