INDUSTRIAL
MATHEMATICS
INSTITUTE

2003:09

Galerkin finite element method for a class of porous medium equations

K.B. Fadimba and R.C. Sharpley

IMI
Preprint Series

Department of Mathematics
University of South Carolina

Galerkin Finite Element Method for a Class of Porous Medium Equations [1]

by

Koffi B. Fadimba
Department of Mathematical Sciences
University of South Carolina at Aiken
Aiken, SC 29801

E-mail: KoffiF@usca.edu
FAX: (803) 641-3726

and

Robert C. Sharpley
Department of Mathematics
University of South Carolina
Columbia, SC 29208

E-mail: sharpley@math.sc.edu

Abstract

We study the numerical approximation of the Saturation Equation which arises in the formulation of two phase fluid flow through porous media, idealized as either a convex bounded polyhedral domain or a domain with smooth boundary. This equation is degenerate and the solutions are not guaranteed to be sufficiently smooth for direct numerical approximation. Through regularization, a family of approximate non-degenerate problems is considered along with their numerical approximations. Error estimates are established for appropriately transformed continuous Galerkin approximations, followed by corresponding error estimates for a fully discretized Galerkin method for this class of problems.

1 Introduction

In modelling immiscible two phase flow through porous media (see e.g. [15, 1, 9]), a class of saturation equations of the form

$$\frac{\partial}{\partial t}S + \nabla \cdot (f(S)\mathbf{u}) - \nabla \cdot (k(S)\nabla S) = Q(S) \qquad \text{on } \Omega \times (0,T]. \tag{1.1}$$

on a bounded domain Ω ($\Omega \subset \mathbf{R}^d$, $d \leq 3$), is derived which satisfies the boundary condition

$$(f(S)\mathbf{u} - k(S)\nabla S) \cdot \mathbf{n} = q \quad \text{on } \partial\Omega \times [0, T_0] \tag{1.2}$$

and has initial condition

$$S(x,0) = S^0(x) \qquad \text{on } \Omega \tag{1.3}$$

with $0 \leq S^0(x) \leq 1$, for all $x \in \Omega$. For simplicity we let $|\Omega| = 1$.

In these equations S is the saturation of the invading fluid (see [1, 9, 15]) and it follows from the general theory [4] that $0 \leq S \leq 1$. The diffusion coefficient $k = k(S)$ is the conductivity of the media, which is assumed here to depend only on the saturation S. The fractional flow function f governs the transport term $\nabla \cdot (f(S)\mathbf{u})$ where $\mathbf{u}$ is the total velocity of the two phase flow. We assume in this analysis that $\mathbf{u}$

[1]This work was supported in part by NSF Grant EHR-910-18772 and the Partnership In Computational Sciences Consortium of DOE through Martin Marietta Contracts 19X-5K965C and 19X-5K966V.

2000 Mathematics Subject Classification. 65J15, 65M12, 65M15, 65M60, 35J70, 35J99, 35Q35

Key words and phrases Porous media, immiscible flow, regularization, advection–diffusion, finite element, error analysis, degenerate elliptic equations, numerical analysis.

is sufficiently smooth and is provided, but in practice, it is obtained by solving another (possibly coupled) elliptic partial differential equation which models the total pressure of the two phases. The term $Q = Q(S)$ represents the source/sink terms and q denotes the boundary flux. We assume throughout that the domain Ω is sufficiently "nice" in order that the standard analysis for elliptic problems [2] be valid; in particular we require elliptic regularity and second order error estimates to hold (see the Appendix for details). This is the case, for example, when Ω is either a convex polyhedral domain satisfying a maximal interior angle condition [12] or has a smooth boundary [2].

For a given fractional flow f, we require that there be a constant C^*, such that

$$C^*|f(b) - f(a)|^2 \leq \big(K(b) - K(a)\big)(b - a). \tag{1.4}$$

Lemma 2.1 below shows that this requirement is reasonable.

We also suppose the diffusion coefficient k satisfies the growth condition

$$k(s) \geq \begin{cases} c_1|s|^{\mu_1} & 0 \leq s \leq \alpha_1 \\ c_2 & \alpha_1 \leq s \leq \alpha_2 \\ c_3|1-s|^{\mu_2} & \alpha_2 \leq s \leq 1 \end{cases} \tag{1.5}$$

where $0 < \alpha_1 < \frac{1}{2} < \alpha_2 < 1$ are given and assume $0 < \mu_1, \mu_2 \leq 2$. Define

$$\begin{aligned} \mu &:= \max(\mu_1, \mu_2) \\ \gamma &:= \frac{2+\mu}{1+\mu} \end{aligned} \tag{1.6}$$

and set

$$K(\xi) = \int_0^\xi k(\tau)d\tau.$$

Because of possible roughness of the solution to the degenerate problem (1.1)–(1.3), one often regularizes the conductivity to obtain a non-degenerate formulation

$$\frac{\partial}{\partial t}S_\beta + \nabla \cdot (f(S_\beta)\mathbf{u}) - \nabla \cdot (k_\beta(S_\beta)\nabla S_\beta) = Q(S_\beta) \qquad \text{on } \Omega \times (0, T_0] \tag{1.7}$$

$$f(S_\beta)\mathbf{u} \cdot \mathbf{n} - \frac{\partial}{\partial \mathbf{n}}K_\beta(S_\beta) = q \qquad \text{on } \partial\Omega \times [0, T_0] \tag{1.8}$$

$$S_\beta(x, 0) = S^0(x) \qquad \text{on } \Omega \tag{1.9}$$

where $K_\beta(\xi) = \int_0^\xi k_\beta(\tau)d\tau$, and $k_\beta \to k$ in an appropriate sense as the regularization parameter β converges to zero. We define $C_0(\beta)$ by

$$C_0(\beta) := \|K_\beta - K\|_{L^\infty}^\gamma \tag{1.10}$$

If $0 < \beta < \frac{1}{2}$ and $k(0) = k(1) = 0$, an example of an acceptable perturbation k_β of k is defined by

$$k_\beta(\xi) = k(\xi) \quad \text{for } k(\xi) > \delta; \qquad k_\beta(\xi) \text{ lies between } \delta \text{ and } \frac{1}{2}\delta, \text{ otherwise.} \tag{1.11}$$

where

$$\delta := \delta(\beta) = \min(k(\beta), k(1 - \beta)). \tag{1.12}$$

For this particular perturbation (see [10, 11]) there holds

$$C_0(\beta) \leq c(\beta\delta(\beta))^\gamma. \tag{1.13}$$

In an earlier paper [11], error estimates were established for any perturbation of k and some of those results which we require are summarized in Section 2. The equations (1.7)–(1.9) constitute the problem that we approximate numerically by a transformed Galerkin Finite Element procedure.

In Section 3 we first approximate the solution by a continuous time Galerkin approximation, i.e. the space variable is discretized. This variational method yields a solution S_h convergent to S_β in a controlled manner as $h \to 0^+$. Our previous knowledge of the rate of approximation of S by the solution S_β of the regularized equation provides then an estimate of the error $\|S - S_h\|$ in the desired function spaces.

In Section 4 we proceed to discretize in time and provide several error estimates for a fully discretized (backward in time) solution. We extend here the results of [16, 18]. M.E. Rose in [16] treats the one dimensional case of this problem and assumes a single and more regulated degeneracy for k. In that case the operator T can be expressed explicitly and the problem can be transformed into a purely parabolic problem. D.L. Smylie in [18] treats the multidimensional case for the parabolic equation. The paper of Nochetto and Verdi [14] also establishes error estimates for the same type of problem, using numerical integration, again for the case of one degeneracy: $k(s) = s^m$, $m > 1$. An example for our setting would be any k for which $k(s) \geq s^{\mu_1}(1-s)^{\mu_2}$, for $0 < \mu_1, \mu_2 \leq 2$. Compare Corollary 3 of [14] to our Corollary 4.1.

As previously mentioned, the total velocity, total pressure formulation of two phase flow in porous medium in several dimensions presents additional analytical and numerical difficulties for the saturation equation. The treatment of the transport term in [11], reproduced here as Lemma 2.1, is very helpful in our estimates. Our arguments in the proofs of Lemma 3.1 and Theorem 3.1 are somewhat different from [16], although we follow the same general line as in that paper.

In section 4, by manipulating the inequalities in a different manner, we are able to improve the convergence rate in the time step as given in [16]. We also give a proof for the existence of a solution of the fully discretized scheme, for the case that the total velocity $\mathbf{u}$ is not a function of the time variable t.

We next describe additional notation which will be used throughout the remainder of this paper. We define $(f,g) := (f,g)_\Omega := \int_\Omega fg dx$ when this has a meaning (extended when appropriate to the distributional sense), and in particular we set $f_\Omega := \frac{1}{|\Omega|}(f,1)_\Omega$. We drop the subscript Ω when there is no ambiguity. The notation $\|f\|_{L^p} := \|f\|_{L^p(\Omega)}$ is used for the standard Lebesgue norm of a measurable function, when this quantity is finite. Similarly, we denote by $\|f\|_{L^p(L^q)} := \|f\|_{L^p(0,T,L^q(\Omega))}$ the mixed Lebesgue norm for f, while $\|f\|_{L^p(H^q)} := \|f\|_{L^p(0,T,H^q(\Omega))}$ designates the mixed Sobolev-Lebesgue norm of a function. We use C, c, to denote constants which may change from line, but which are independent of the parameters β, h and Δt, unless explicitly specified.

Finally, we would like to thank the referee for valuable suggestions and pointing out appropriate references which improved the error estimates for numerical approximations by a logarithmic factor.

2 Regularization Results

We summarize here some results from [11, 10], which are required in the error analysis that follows.

Lemma 2.1 *If $f \in C^1([0,1])$ with $f'(0) = f'(1) = 0$, then there is a positive constant C^* so that for all $0 \leq a \leq b \leq 1$ we have*

$$C^* \, |f(b) - f(a)|^2 \leq (K(b) - K(a))\,(b-a). \tag{2.1}$$

This lemma will typically be applied in the integral form

$$C^* \, \|f(u) - f(v)\|^2_{L^2(\Omega)} \leq (K(u) - K(v),\, u - v)\,. \tag{2.2}$$

Conversely, if (2.1) is satisfied, then

$$|f'(\xi)| \leq C\sqrt{k(\xi)} \tag{2.3}$$

Lemma 3.1 in [11] imples that there is a positive constant C^{**} so that

$$C^{**} \, \|u - v\|^{2+\mu}_{L^{2+\mu}} \leq (K(u) - K(v),\, u - v) \tag{2.4}$$

is valid for all $u, v \in L^{2+\mu}$. Moreover, the inequalities above remain true if k is replaced by the regularized k_β (or K is replaced by K_β) with constants independent of β. Finally, we observe that when K is Lipschitz, we have:

$$\begin{aligned}(K(u) - K(v))^2_\Omega &\leq \|K(u) - K(v)\|^2_{L^2} \\ &\leq \|k\|_\infty \,(K(u) - K(v),\, u - v)\,. \end{aligned} \tag{2.5}$$

The next Theorem gives the error estimates for $S-S_\beta$, when the initial problem (1.1)–(1.3) is replaced by the regularized nondegenerate problem (1.7)–(1.9). We assume for the remaining of this paper, to simplify the analysis, that $Q \equiv 0$ and $q \equiv 0$.

Theorem 2.1 *(Theorem 4.1 of [11]) Assume that the coefficients f and k satisfy the conditions (1.4)-(1.5). Let S_β be the solution to regularized equations (1.7)–(1.9), and S the solution to (1.1)–(1.3), then*

$$\sup_{0\le t\le T_0} \|S_\beta - S\|^2_{(H^1)^*} + \eta \int_0^{T_0} \Big(K_\beta(S_\beta) - K_\beta(S),\, S_\beta - S\Big)(\tau)d\tau \;\le\; C_0(\beta) \tag{2.6}$$

$$\Big\|K_\beta(S_\beta) - K_\beta(S)\Big\|^2_{L^2(L^2)} \;\le\; C_0(\beta) \tag{2.7}$$

and

$$\|S_\beta - S\|^{2+\mu}_{L^{2+\mu}(L^{2+\mu})} \;\le\; C_0(\beta) \tag{2.8}$$

where $C_0(\beta)$ is defined by (1.10).

Another useful result is the following which establishes estimates for the solution to the regularized equation.

Lemma 2.2 *Assume the hypotheses of Theorem 2.1 hold. If S_β is the solution to (1.7)–(1.9), then there exist constants C, C_1, C_2 (independent of β) such that*

$$\left\|\frac{\partial S_\beta}{\partial t}\right\|_{L^\infty(0,T_0,L^1(\Omega))} \le C_1 + C_2\|S^0\|_{W_2^1(\Omega)}. \tag{2.9}$$

$$\|S_\beta\|^2_{L^\infty(L^2)} + \left\|\sqrt{k_\beta(S_\beta)}\nabla S_\beta\right\|^2_{L^2(L^2)} \le C\cdot T_0 + \|S_0\|^2_{L^2} \tag{2.10}$$

$$\left\|\sqrt{k_\beta(S_\beta)}(S_\beta)_t\right\|^2_{L^2(L^2)} + \|\nabla K_\beta(S_\beta)\|^2_{L^\infty(L^2)} \le C + \left\|\nabla K_\beta(S^0)\right\|^2_{L^2}. \tag{2.11}$$

This Lemma provides the elements for the proof of the following result (see Theorem 4.4 of [11]).

Theorem 2.2 *Assume the hypotheses of Theorem 2.1 holds and define*

$$m(\beta) := \inf_{0\le s\le 1} k_\beta(s). \tag{2.12}$$

If $\gamma = \frac{2+\mu}{1+\mu}$, then there is a constant C, independent of β, such that

$$\|(S_\beta)_t\|_{L^\gamma(L^\gamma)} \le C\, m(\beta)^{-\frac{1}{2+\mu}} \tag{2.13}$$

and hence

$$\left\|\frac{\partial S_\beta}{\partial t} + \nabla\cdot f(S_\beta)\mathbf{u}\right\|_{L^\gamma(L^\gamma)} \le C\, m(\beta)^{-\frac{1}{2+\mu}} \tag{2.14}$$

We note that the respective estimates (2.11) and (2.10), using the proof of Theorem 4.4 of [11], immediately imply the inequalities

$$\|(S_\beta)_t\|_{L^2(L^2)} \le C\, m(\beta)^{-\frac{1}{2}} \tag{2.15}$$

$$\|\Delta K_\beta(S_\beta)\|_{L^2(L^2)} \le C\, m(\beta)^{-\frac{1}{2}}. \tag{2.16}$$

If we choose the specific regularization where k_β is defined by

$$k_\beta(s) := \max(k(s), \beta^\mu) \tag{2.17}$$

and set

$$\delta(\beta) = \beta^{\min(1,\mu)},$$

then Theorem 4.6 of [11] gives the following estimates.

Theorem 2.3 *Assume the hypotheses of Theorem 2.1 hold. Then*

$$\sup_{0\le t\le T_0}\Big(K(S_\beta)-K(S),\,S_\beta-S\Big)+\eta\left\|\nabla(K(S_\beta)-K(S))\right\|^2_{L^2(L^2)}\le C\,\delta(\beta) \tag{2.18}$$

$$\|S_\beta-S\|^{2+\mu}_{L^\infty(L^{2+\mu})}\le C\,\delta(\beta) \tag{2.19}$$

$$\left\|K(S_\beta)-K(S)\right\|^2_{L^2(H^1,[0,T_0])}\le C\,\delta(\beta). \tag{2.20}$$

3 The Continuous Galerkin Method

3.1 The Finite Element Space

We give a brief description of the approximation subspaces which provide the finite element solutions for the Galerkin problems. For general background references of the methods used in this section, see [2] and [12, 19]. We let $\{M_h\}_{0<h<1}$ denote a family of finite dimensional spaces, with $M_h\subset H^1(\Omega)$ and assume that M_h has the approximation property:

$$\inf_{\chi\in M_h}\|f-\chi\|_{L^p(\Omega)}\le Ch^2|f|_{W^{2,p}}\qquad\qquad \text{for all } f\in W^{2,p}(\Omega). \tag{3.1}$$

We will also need the *inverse estimate* assumption on M_h (see, for example, Section 4.5 of [2]):

$$\|\chi\|_{H^1}\le Ch^{-1}\|\chi\|_{L^2}\qquad\qquad \text{for all } \chi\in M_h. \tag{3.2}$$

which by duality implies

$$\|\chi\|^2_{L^2}=(\chi,\chi)\le\|\chi\|_{H^1}\|\chi\|_{(H^1)^*}\le Ch^{-1}\|\chi\|_{L^2}\|\chi\|_{(H^1)^*}.$$

and consequently

$$\|\chi\|_{L^2(\Omega)}\le Ch^{-1}\|\chi\|_{(H^1)^*}\qquad\qquad \text{for all } \chi\in M_h. \tag{3.3}$$

An important case is that for Ω, a convex bounded polygonal domain in $\mathbf{R}^2$ with a triangulation $\mathcal{T}_h=\{\mathcal{K}\}$ where the parameter h $(0<h<1)$ is defined as follows: for a triangle $\mathcal{K}\in\mathcal{T}$, define $h_\mathcal{K}$ and $\rho_\mathcal{K}$ by

$$h_\mathcal{K}:=\operatorname{diam}(\mathcal{K}) \tag{3.4}$$

and

$$\rho_\mathcal{K}:=\sup\{\operatorname{diam}(C):\ \text{C is a circle inscribed in } \mathcal{K}\}, \tag{3.5}$$

then

$$h:=\max\{h_\mathcal{K}:\mathcal{K}\in\mathcal{T}_h\}. \tag{3.6}$$

The space of piecewise linear elements for this triangulation is defined by

$$M_h:=\{\phi\in C^0(\Omega):\ \phi/\mathcal{K}\text{ is linear for all }\mathcal{K}\in\mathcal{T}_h\}. \tag{3.7}$$

We assume, in addition, that $\mathcal{T}$ is a regular family, i.e. there exists a constant $d_1>0$ such that

$$\frac{\rho_\mathcal{K}}{h_\mathcal{K}}\ge d_1. \tag{3.8}$$

in which case [2, 12, 8, 5] there exists $\overline{p}=\overline{p}(\Omega)>2$ so that M_h satisfies the direct approximation estimate (3.1) for all $1<p<\overline{p}$ (see (A.13) of the Appendix for the dependence of $\overline{p}$ on Ω). If we assume $\mathcal{T}$ is a quasi–uniform triangulation, i.e. (3.8) is satisfied and there exists $d_2>0$ such that

$$h_\mathcal{K}\ge d_2h,\qquad\qquad \text{for all } \mathcal{K}\in\mathcal{T}_h, \tag{3.9}$$

then M_h defined by (3.7) also satisfies the inverse estimate (3.2) (cf. [2, 13]).

3.2 The Discretized Problem

Although analytically the saturation lies in the interval $[0,1]$, small numerical oscillations may occur and so we extend the domain of the functions f and k_β as follows:

$$k_\beta(\xi) = \begin{cases} k_\beta(1) & \text{if } \xi \geq 1 \\ k_\beta(-\xi) & \text{if } \xi \leq 0 \end{cases} \tag{3.10}$$

$$f(\xi) = \begin{cases} 0 & \text{if } \xi \leq 0 \\ f(1) & \text{if } \xi \geq 1 \end{cases} \tag{3.11}$$

We continue to define the primitive K_β by

$$K_\beta(\xi) = \int_0^\xi k_\beta(\tau)d\tau \tag{3.12}$$

and observe that it is a strictly increasing C^1 function on $\mathbf{R}$, since $k_\beta(\xi) > 0$ as long as $\beta > 0$. Hence K_β has a C^1 inverse function H_β:

$$H_\beta(\xi) = K_\beta^{-1}(\xi) \qquad\qquad \forall \xi \in \mathbf{R}. \tag{3.13}$$

We consider the ordinary differential equation (actually a coupled ODE system):

$$\left(\frac{\partial}{\partial t}H_\beta(V_h), \chi\right) - (f(H_\beta(V_h))\mathbf{u}, \nabla\chi) + (\nabla V_h, \nabla\chi) = 0 \tag{3.14}$$

required to hold for all $\chi \in M_h$, and $t \in (0, T_0]$. This system has the initial condition:

$$P_h H_\beta(V_h(0)) = P_h S^0 \tag{3.15}$$

where S^0 is as in (1.3), and P_h is the L^2 projection on M_h. We solve for V_h in M_h where V_h is the Galerkin approximation to $K_\beta(S_\beta)$. We then set $S_h = H_\beta(V_h)$, so that S_h approximates S_β. The operator $P_h \circ H_\beta$ is a continuous, nonlinear, coercive map from M_h into itself [16] and is thus bijective. Therefore by (3.15), $V_h(0)$ exists in M_h.

Suppose that $\{e_i\}_1^m$ is a basis for M_h, with $m = m(h) = \dim(M_h)$, so that for all $\chi \in M_h$, $\chi(x) = \sum_{i=1}^m \chi_i e_i(x)$. Then (3.14) is equivalent to the system of m coupled O.D.E:

$$\frac{d}{dt}(P_h(H_\beta(V_h)), e_i) - (f(H_\beta(V_h))\mathbf{u}, \nabla e_i) + (\nabla V_h, \nabla e_i) = 0 \qquad (1 \leq i \leq m) \tag{3.16}$$

This can be rewritten in a vector form as a Cauchy problem:

$$\begin{cases} \frac{d}{dt}P_h H_\beta(V_h) = F(P_h H_\beta(V_h)) \\ P_h H_\beta(V_h(0)) = P_h S^0 \end{cases} \tag{3.17}$$

With our assumption on f, k_β and $\mathbf{u}$, the function F is Lipschitz, so we are guaranteed the existence and uniqueness of the solution $\tilde{S}_h = P_h H_\beta(V_h)$ to (3.17). We have previously observed that $P_h H_\beta$ is bijective, so we have that $V_h = (P_h H_\beta)^{-1}\tilde{S}_h$ exists in M_h.

For convenience we define $S_h := H_\beta(V_h)$ and rewrite (3.14) as

$$((S_h)_t, \chi) - (f(S_h)\mathbf{u}, \nabla\chi) + (\nabla K_\beta(S_h), \nabla\chi) = 0, \qquad \forall \chi \in M_h. \tag{3.18}$$

By approximating $K_\beta(S_\beta)$ by $V_h \in M_h$, the approximation of S_β by $S_h = H_\beta(V_h)$ is shown in Theorem 3.1 below to have higher rate of convergence than approximating S_β directly by an element of M_h.

3.3 The main results

As stated above we give our main estimates for the error $\|S - S_h\|$ in two theorems. We need the following discrete version of inequality (2.11) with S_β replaced by S_h:

Lemma 3.1 *If V_h is the solution to (3.14)–(3.15), and if we set $S_h = H_\beta(V_h)$, then*

$$\int_0^{T_0} \Big((S_h)_t, K_\beta(S_h)_t\Big)\, dt + \eta \left\|\nabla K_\beta(S_h)\right\|^2_{L^\infty(L^2)} \le \tilde{C}\{\|\mathbf{u}\|^2_{L^\infty(L^2)} + \|\mathbf{u}_t\|^2_{L^2(L^2)}\} \tag{3.19}$$

where $\tilde{C} = \tilde{C}(T_0, \mathbf{u}) = \exp(C\{1 + \|\mathbf{u}\|^2_{L^\infty(L^\infty)}\})$, and η a positive constant.

Proof. In (3.18), let $\chi = V_{ht} = K_\beta(S_h)_t$. Then

$$((S_h)_t, K_\beta(S_h)_t) - (f(S_h)\mathbf{u}, \nabla K_\beta(S_h)_t) + (\nabla K_\beta(S_h), \nabla(K_\beta(S_h)_t)) = 0,$$

or equivalently

$$((S_h)_t, K_\beta(S_h)_t) + \frac{1}{2}\frac{d}{dt}\|\nabla K_\beta(S_h)\|^2_{L^2} = (f(S_h)\mathbf{u}, \nabla K_\beta(S_h)_t), \tag{3.20}$$

where we use the fact that Δ and $\frac{\partial}{\partial t}$ commute. Applying the product rule on the right-hand side of (3.20), and substituting, we obtain

$$((S_h)_t, K_\beta(S_h)_t) + \frac{1}{2}\frac{d}{dt}\|\nabla K_\beta(S_h)\|^2_{L^2} = \\ \frac{d}{dt}(f(S_h)\mathbf{u}, \nabla K_\beta(S_h)) - ((f(S_h)_t\mathbf{u}, \nabla K_\beta(S_h)) - (f(S_h)\mathbf{u}_t, \nabla K_\beta(S_h)) \,. \tag{3.21}$$

By (2.3)

$$\|f(S_h)_t\mathbf{u}\|^2_{L^2} = \|f'(S_h)(S_h)_t\mathbf{u}\|^2_{L^2} \le c\|\mathbf{u}\|^2_{L^\infty} \left\|\sqrt{k_\beta(S_h)}(S_h)_t\right\|^2_{L^2} \\ = c\|\mathbf{u}\|^2_{L^\infty}((S_h)_t, K_\beta(S_h)_t). \tag{3.22}$$

Thus (3.21) implies

$$((S_h)_t, K_\beta(S_h)_t) + \frac{1}{2}\frac{d}{dt}\|\nabla K_\beta(S_h)\|^2_{L^2} \\ \le \frac{d}{dt}(f(S_h)\mathbf{u}, \nabla K_\beta(S_h)) + \sigma_1\|f(S_h)_t\mathbf{u}\|^2_{L^2} + \frac{1}{\sigma_1}\|\nabla K_\beta(S_h)\|^2_{L^2} \\ + \frac{1}{2}\|f(S_h)\mathbf{u}_t\|^2_{L^2} + \frac{1}{2}\|\nabla K_\beta(S_h)\|^2_{L^2}, \tag{3.23}$$

where σ_1 is an arbitrary positive constant. This yields, by (3.22),

$$((S_h)_t, K_\beta(S_h)_t) + \frac{1}{2}\frac{d}{dt}\|\nabla K_\beta(S_h)\|^2_{L^2} \\ \le \frac{d}{dt}(f(S_h)\mathbf{u}, \nabla K_\beta(S_h)) + \sigma_1 c\|u\|_{L^\infty(L^\infty)}((S_h)_t, K_\beta(S_h)_t) \\ + \left(\frac{1}{\sigma_1} + \frac{1}{2}\right)\|\nabla K_\beta(S_h)\|^2_{L^2} + \frac{1}{2}\|f(S_h)\mathbf{u}_t\|^2_{L^2}. \tag{3.24}$$

Now, for σ_1 sufficiently small, we can hide the second term of the right-hand side of (3.24) in the left-hand side of (3.23). Also $\|f(S_h)\mathbf{u}_t\|_{L^2} \le c\|\mathbf{u}_t\|_{L^2}$, given the smoothness assumptions on f. So (3.24) becomes

$$\frac{1}{2}((S_h)_t, K_\beta(S_h)_t) + \frac{1}{2}\frac{d}{dt}\|\nabla K_\beta(S_h)\|^2_{L^2} \\ \le \frac{d}{dt}(f(S_h)\mathbf{u}, \nabla K_\beta(S_h)) + C(\mathbf{u})\|\nabla K_\beta(S_h)\|^2_{L^2} + C\|\mathbf{u}_t\|^2_{L^2(L^2)} \tag{3.25}$$

Finally using the Grönwall Lemma, we get

$$\int_0^{T_0} ((S_h)_t, K_\beta(S_h)_t)dt + \sigma\|\nabla K_\beta(S_h)\|^2_{L^\infty(L^2)}$$
$$\leq \tilde{C}(\mathbf{u})\{\sup_{0\leq t\leq T_0} |(f(S_h)\mathbf{u}, \nabla K_\beta(S_h))(t)| + C\|\mathbf{u}_t\|^2_{L^2(L^2)}\}. \tag{3.26}$$

with $\tilde{C}(u) = \exp(C(u))$. But

$$\sup_{0\leq t\leq T_0} |(f(S_h)\mathbf{u}, \nabla K_\beta(S_h))(t)| \leq C\|f(S_h)\mathbf{u}\|_{L^\infty(L^2)} + \frac{1}{2}\sigma\|\nabla K_\beta(S_h)\|^2_{L^\infty(L^2)}. \tag{3.27}$$

Combining these last two inequalities, we obtain the Lemma by taking $\eta = \frac{1}{2}\sigma$. □

Remark 3.1 *From Lemma 3.1 it follows that*

$$\left\|\sqrt{k_\beta(S_h)}(S_h)_t\right\|_{L^2(L^2)} \leq \tilde{C}(\mathbf{u})\{\|\mathbf{u}\|^2_{L^\infty(L^2)} + \|\mathbf{u}_t\|^2_{L^2(L^2)}\}^{\frac{1}{2}} \tag{3.28}$$

and

$$\|\nabla K_\beta(S_h)\|_{L^\infty(L^2)} \leq \tilde{C}(\mathbf{u})\{\|\mathbf{u}\|^2_{L^\infty(L^2)} + \|\mathbf{u}_t\|^2_{L^2(L^2)}\}, \tag{3.29}$$

where the constants are independent of β and h.

With this Lemma, we are now in position to formulate and prove the main theorem of this section. The theorem states that V_h converges to $K_\beta(S_\beta)$, and that S_h converges to S, the solution to (1.1)–(1.3). We note that the rate of convergence of V_h to $K_\beta(S_\beta)$ is higher than that of S_h to S_β, since the elements incorporate attributes of the diffusion coefficient.

Theorem 3.1 *Suppose μ and γ are given by (1.6), and $m(\beta), C_0(\beta)$ are as defined in (2.12) and (1.10), respectively. Furthermore, assume that $0 < \mu < \overline{p}(\Omega) - 2$, where $\overline{p}(\Omega)$ is defined by the relation (A.13). Let S be the solution to the degenerate equation (1.1)–(1.3) with $Q = 0$ and $q = 0$, and with coefficients f and k which satisfy conditions (1.4)-(1.5). If M_h satisfies the approximation (3.1) and inverse (3.2) properties, and V_h solves (3.14), then the approximate solution $S_h := K_\beta(V_h)$ satisfies the inequality*

$$\|S - S_h\|^2_{L^\infty((H^1)^*)} + \int_0^{T_0} \Big(K_\beta(S_\beta) - K_\beta(S_h), S_\beta - S_h\Big)$$
$$\leq ch^{2\gamma}m(\beta)^{-\frac{1}{\mu+1}} + C_0(\beta). \tag{3.30}$$

with constant c independent of β and h.

Remark 3.2 *The condition, $0 < \mu < \overline{p}(\Omega)$, on the degeneracy of the conductivity, is relatively mild. For example, if the maximal interior angle $\Theta(\Omega)$ of Ω is no larger than $\frac{2\pi}{3}$, then by its definition in (A.13) $\overline{p} - 2 = 2$ and there is no additional restriction on μ. If Θ increases to $\frac{7\pi}{8}$, then the maximum value of μ allowed is reduced to $\frac{1}{3}$.*

Proof of Theorem 3.1. The proof will be split into two main steps by writing

$$S - S_h = (S - S_\beta) + P_h(S_\beta - S_h) + (I - P_h)(S_\beta - S_h).$$

The two steps will estimate respectively, the second and third terms, while inequality (2.6) of Theorem 2.1 is used to estimate the first term by $C_0(\beta)$.

In order to obtain the desired results by applying $T(S_\beta - S_h)$ as a test function in the weak formulation of the regularized problem (1.7)-(1.9), we first observe that for all $t > 0$

$$(S_\beta - S_h)_\Omega = 0. \tag{3.31}$$

Indeed, equations (1.7)–(1.9) imply

$$((S_\beta)_t, \chi) - (f(S_\beta)\mathbf{u}, \nabla\chi) + (\nabla K_\beta(S_\beta), \nabla\chi) = 0 \tag{3.32}$$

for all $\chi \in M_h$. By subtracting (3.18) from this equation, we obtain

$$((S_\beta - S_h)_t, \chi) - ((f(S_\beta) - f(S_h))\mathbf{u}, \nabla\chi) + (\nabla(K_\beta(S_\beta) - K_\beta(S_h)), \nabla\chi) = 0 \tag{3.33}$$

If we set $\chi = 1 \in M_h$ in this equation, then $\frac{d}{dt}(S_\beta - S_h)_\Omega = 0$, in which case, $(S_\beta - S_h)_\Omega = (S^0 - S_h(0))_\Omega$. for all $t > 0$. But the initial condition (3.15) for the Galerkin solution implies

$$(S^0 - S_h(0))_\Omega = (S^0 - S_h(0), 1) = (P_h(S^0 - S_h(0)), 1) = 0,$$

which verifies that mean values are preserved.

Step 1: We derive the estimate:

$$\|P_h(S_\beta - S_h)\|^2_{L^\infty(H_h^{-1})} + \eta \int_0^{T_0} (K_\beta(S_\beta) - K_\beta(S_h), S_\beta - S_h)(t)dt \leq Ch^{2\gamma} m(\beta)^{-\frac{1}{1+\mu}} \tag{3.34}$$

with the definition of the norm $\|\cdot\|_{H_h^{-1}}$ given in the Appendix. We use as test function $\phi = T(S_\beta - S_h) \in H^1(\Omega)$ in the weak formulation of the regularized saturation equation for S_β and use $\phi = T_h(S_\beta - S_h) \in M_h$ for the Galerkin formulation (3.18) with solution S_h, where we recall that T_h is defined as $E_h \circ T$. Subtracting these two equations and rearranging, we obtain the reference equation for Step 1:

$$\begin{aligned}\Big((S_\beta - S_h)_t, T_h(S_\beta - S_h)\Big) + \Big(\nabla(K_\beta(S_\beta) - K_\beta(S_h)), \nabla T(S_\beta - S_h)\Big) = \\ - \Big(\nabla\cdot(f(S_\beta) - f(S_h))\mathbf{u}, T_h(S_\beta - S_h)\Big) \\ - \Big((S_\beta)_t + \nabla\cdot f(S_\beta)\mathbf{u}, (T - T_h)(S_\beta - S_h)\Big) \ .\end{aligned} \tag{3.35}$$

We have used here the fact that the additional term

$$\Big((I - E_h)K_\beta(S_h), S_\beta - S_h\Big)$$

vanishes since $K_\beta(S_h) = V_h \in M_h$ and so $(I - E_h)V_h = 0$.

For the first term on the left hand side of the reference equation (3.35), we use the identity

$$T_h f = T_h P_h f \qquad \forall f \in (H^1)^*, \tag{3.36}$$

which follows directly from the definitions of T and the projections, in order to write

$$((S_\beta - S_h)_t, T_h(S_\beta - S_h)) = \frac{1}{2}\frac{d}{dt}\|P_h(S_\beta - S_h)\|^2_{H^{-1}}. \tag{3.37}$$

For the second term on the left hand side of reference equation (3.35), we use the properties of the operator T and the fact that $S_\beta - S_h$ has vanishing mean in order to see that it reduces as

$$(\nabla(K_\beta(S_\beta) - K_\beta(S_h)), \nabla T(S_\beta - S_h)) = (K_\beta(S_\beta) - K_\beta(S_h), S_\beta - S_h). \tag{3.38}$$

To handle the first term on the right hand side of equation (3.35), we use Cauchy-Schwartz, the arithmetic-geometric mean inequality, relation (2.2), and the properties of T (see (3.36) and (A.27)), respectively, to obtain

$$\begin{aligned}&\Big|\Big(\nabla\cdot(f(S_\beta) - f(S_h))\mathbf{u}, T_h(S_\beta - S_h)\Big)\Big| \\ &\leq \|f(S_\beta) - f(S_h)\|_{L^2}\,\|\mathbf{u}\|_\infty\,\|\nabla T_h(S_\beta - S_h)\|_{L^2} \\ &\leq \frac{C^*}{4}\|f(S_\beta) - f(S_h)\|^2_{L^2} + c(\mathbf{u})\|\nabla T_h(S_\beta - S_h)\|^2_{L^2} \\ &\leq \frac{1}{4}\Big(K_\beta(S_\beta) - K_\beta(S_h), S_\beta - S_h\Big) + c(\mathbf{u})\|P_h(S_\beta - S_h)\|_{H_h^{-1}}\end{aligned} \tag{3.39}$$

which is of the desired form for employing the standard method of burying terms and using Grönwall's lemma. We must first prepare the remaining terms. For the second term on the right hand side of reference equation (3.35), we set for convenience the variable $W := (S_\beta)_t + \nabla \cdot f(S_\beta)\mathbf{u}$, use the fact that $T - T_h$ is a symmetric operator and follow in a similar manner to the estimates used in (3.39): here Hölder's equality with conjugate indices γ and $2+\mu$ replaces the Cauchy-Schwartz inequality, and inequality (2.4) replaces (2.2), which results in

$$\begin{aligned}\left|\Big(W\,, (T-T_h)[S_\beta - S_h]\Big)\right| &\le \|S_\beta - S_h\|_{L^{2+\mu}} \|(T-T_h)(W)\|_{L^\gamma} \\ &\le \frac{C^{**}}{4}\|S_\beta - S_h\|^{2+\mu}_{L^{2+\mu}} + C\|(T-T_h)(W)\|^\gamma_{L^\gamma} \\ &\le \frac{1}{4}\Big(K_\beta(S_\beta) - K_\beta(S_h), S_\beta - S_h\Big) + C\|(T-T_h)(W)\|^\gamma_{L^\gamma} \qquad (3.40)\end{aligned}$$

In this last expression, the rightmost term may be estimated by using the fact that $T_h = E_h T$, and then applying the finite element error estimates for the elliptic approximation, as outlined in (A.21) of the Appendix, in order to obtain

$$\begin{aligned}\|(T-T_h)(W)\|_{L^\gamma} &= \|(I-E_h)T(W)\|_{L^\gamma} \\ &\le ch^2\|T(W)\|_{W^{2,\gamma}} \\ &\le ch^2\|W\|_{L^\gamma} \qquad (3.41)\end{aligned}$$

where the last inequality follows from the mapping properties of the operator T (see (A.14) below) in the range $1 < \gamma' = 2 + \mu < \overline{p}(\Omega)$. This restriction on γ is equivalent to a restriction on μ which depends on the parameter $\Theta(\Omega)$ appearing in the definition of $\overline{p}(\Omega)$ given in (A.13).

If we substitute the identities (3.37) and (3.38) into the reference equation (3.35) for Step 1, and follow by using the inequalities (3.39), (3.40), and (3.41), we obtain

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|P_h(S_\beta - S_h)\|^2_{H^{-1}} + \frac{1}{2}(K_\beta(S_\beta) - K_\beta(S_h), S_\beta - S_h) \\ &\qquad = c(\mathbf{u})\|P_h(S_\beta - S_h)\|^2_{H^{-1}} + ch^{2\gamma}\|W\|^\gamma_{L^\gamma} \qquad (3.42)\end{aligned}$$

where $W = (S_\beta)_t + \nabla\cdot f(S_\beta)\mathbf{u}$. Applying Grönwall's inequality, and using our fact (2.14), we have established the desired estimate (3.34) for Step 1 with norm for H^{-1} in place of that of H_h^{-1}. But the inverse property of M_h implies (A.28) (in the Appendix) and therefore we have

$$C\|P_h(S_\beta - S_h)\|_{(H^1)^*} \le \|P_h(S_\beta - S_h)\|_{H_h^{-1}} \le \|P_h(S_\beta - S_h)\|_{(H^1)^*}. \qquad (3.43)$$

and the proof of Step 1 is complete.

Step 2: We show that

$$\|(I-P_h)(S_\beta - S_h)\|_{L^\infty((H^1)^*)} \le Ch^\gamma. \qquad (3.44)$$

Define parameter $0 < \epsilon < 1$ by

$$\epsilon := h^{\frac{1}{1+\mu}}, \qquad (3.45)$$

where we may assume that h is small enough so that $\epsilon \le \min\{\alpha_1, 1-\alpha_2\}$, where α_1 and α_2 were prescribed in (1.5). We introduce as in [16] and [18] a new dependent variable S^ϵ_β, defined as follows:

$$S^\epsilon_\beta := \begin{cases} \max(S_\beta, \epsilon) & \text{if } S_\beta \le \alpha_2 \\ \min(S_\beta, 1-\epsilon) & \text{if } S_\beta > \alpha_2 \end{cases} \qquad (3.46)$$

Obviously, with S^ϵ_β defined in this way, we have

$$|S^\epsilon_\beta - S_\beta| \le \epsilon \qquad (3.47)$$

and

$$k_\beta(S_\beta^\epsilon) \geq C\epsilon^\mu. \tag{3.48}$$

We define S_h^ϵ similarly and obtain the same estimates (3.47) and (3.48) where S_β and S_β^ϵ are replaced by S_h and S_h^ϵ, respectively. Using (3.48) we see

$$\|\nabla K_\beta(S_\beta^\epsilon)\|_{L^2} = \|k_\beta(S_\beta^\epsilon)\nabla S_\beta^\epsilon\|_{L^2} \geq C\epsilon^\mu \|\nabla S_\beta^\epsilon\|_{L^2} \tag{3.49}$$

which implies

$$\|\nabla S_\beta^\epsilon\|_{L^\infty(L^2)} \leq \epsilon^{-\mu}\|\nabla K_\beta(S_\beta^\epsilon)\|_{L^\infty(L^2)} \tag{3.50}$$

But in the weak sense $\nabla K_\beta(S_\beta^\epsilon) = \chi_{\{\epsilon \leq S_\beta \leq 1-\epsilon\}} \nabla K_\beta(S_\beta)$, so our earlier estimate (2.11) implies that

$$\|\nabla K_\beta(S_\beta^\epsilon)\|_{L^\infty(L^2)} \leq \|\nabla K_\beta(S_\beta)\|_{L^\infty(L^2)} \leq C \tag{3.51}$$

and thus combining with (3.50) we obtain

$$\|\nabla S_\beta^\epsilon\|_{L^\infty(L^2)} \leq C\epsilon^{-\mu} \; . \tag{3.52}$$

Similarly, if we use (3.29) this same proof shows that

$$\|\nabla S_h^\epsilon\|_{L^\infty(L^2)} \leq C\epsilon^{-\mu}. \tag{3.53}$$

Next recall the approximation property of the L^2–projection, [2, 12] which states for $j = -1, 0, 1$ that

$$\|(I - P_h)\phi\|_{H^j} \leq Ch\|\phi\|_{H^{j+1}} \qquad\qquad \forall \phi \in H^{j+1}(\Omega) \; . \tag{3.54}$$

It then follows directly that

$$\begin{aligned}\|(I - P_h)S_\beta\|_{L^\infty(L^2)} &\leq \|(I - P_h)(S_\beta - S_\beta^\epsilon)\|_{L^\infty(L^2)} + \|(I - P_h)S_\beta^\epsilon\|_{L^\infty(L^2)} \\ &\leq \; C\,\epsilon + C\,h\|S_\beta^\epsilon\|_{L^\infty(H^1)} \quad \leq \quad C(\epsilon + h\epsilon^{-\mu}).\end{aligned} \tag{3.55}$$

where we have made use of inequality (3.47) and the mapping properties of P_h, the approximation property (3.54) applied with $j = 0$, and the estimate (3.52). But from its definition, $\epsilon = h^{\frac{1}{1+\mu}}$ and so substituting into (3.55) above, we get

$$\|(I - P_h)S_\beta\|_{L^\infty(L^2)} \leq Ch^{\frac{1}{1+\mu}} \; . \tag{3.56}$$

Similarly $\|(I - P_h)S_h\|_{L^\infty(L^2)} \leq Ch^{\frac{1}{1+\mu}}$ and so we obtain

$$\|(I - P_h)(S_\beta - S_h)\|_{L^\infty(L^2)} \leq Ch^{\frac{1}{1+\mu}}. \tag{3.57}$$

Upon another application of the error estimate (3.54) with $j = -1$, we see that $\|(I-P_h)(S_\beta - S_h)\|_{L^\infty(H^{-1})} \leq Chh^{\frac{1}{1+\mu}}$ and the inequality (3.44) stated in Step 2 is verified.

Finally, by combining the inequalities established in Steps 1 and 2, together with the estimate for $S - S_\beta$ provided by Theorem 2.1, the proof of Theorem 3.1 is complete. □

We illustrate some immediate consequences of Theorem 3.1 and its proof by a particular choice of the perturbation

$$k_\beta(s) = \max(k(s), c_0\beta^\mu) \qquad 0 \leq s \leq 1 \; .$$

In this case, it is straightforward to estimate that $C(\beta) \leq C_0\beta^{2+\mu}$ and $m(\beta) \geq c_0\beta^\mu$. Finally, let $\beta = \beta(h)$ be given by

$$\beta = \beta_0 h^\lambda \qquad\qquad \text{with } \lambda = \tfrac{4+2\mu}{2+4\mu+\mu^2} \tag{3.58}$$

for a fixed positive constant β_0.

Corollary 3.1 *Suppose $0 < \mu < \overline{p}(\Omega) - 2$, with $\overline{p}(\Omega)$ defined as in (A.13). Let γ be given by (1.6) and S be the solution to the equation (1.1)–(1.3) with $Q = 0$ and $q = 0$, whose coefficients f and k satisfy conditions (1.4)-(1.5). If M_h satisfies the approximation (3.1) and inverse (3.2) properties, then the approximate solution $S_h := K_\beta(V_h)$, where V_h solves (3.14) with β as in (3.58), satisfies the estimates*

$$\|S - S_h\|_{L^\infty((H^1)^*)} \leq C\, h^{\frac{2+\mu}{2}\lambda} \tag{3.59}$$

$$\|K_\beta(S_h) - K_\beta(S_\beta)\|_{L^2(L^2)} \leq C\, h^{\frac{2+\mu}{2}\lambda} \tag{3.60}$$

$$\|S - S_h\|_{L^{2+\mu}(L^{2+\mu})} \leq C\, h^{\lambda} \tag{3.61}$$

$$\|S - S_h\|_{L^\infty(L^2)} \leq C\, h^{\frac{2+\mu}{2}\lambda - 1}. \tag{3.62}$$

Proof. The estimates (3.59)–(3.61) are by now clear. Inequality (3.59) follows directly from (3.30) with the designated assignment of β. Inequality (3.61) follows from (3.30) of Theorem 3.1 together with inequality (2.4). Inequality (3.60) follows similarly, but uses the simple pointwise inequality $|K_\beta(b) - K_\beta(a)|^2 \leq \|k\|_\infty (K_\beta(b) - K_\beta(a))(b-a)$ which is uniform in β. Finally, to establish the estimate (3.62) we notice that in the course of the proof of Theorem 3.1 (see (3.3) and (3.34) and (3.57), respectively) that we have

$$\|P_h(S_\beta - S_h)\|^2_{L^\infty(L^2)} \leq Ch^{-2}\|P_h(S_\beta - S_h)\|^2_{L^\infty((H^1)^*)} \leq C\, h^{2\gamma-2} m(\beta)^{-\frac{1}{1+\mu}} \tag{3.63}$$

$$\|(I - P_h)(S_\beta - S_h)\|_{L^\infty(L^2)} \leq Ch^{\frac{1}{1+\mu}}. \tag{3.64}$$

The proof is completed by combining this with the estimate (2.19) from Lemma 2.3. □

3.4 Additional Error Estimates

The following theorem was stated in [16] without proof in the one dimensional case where K has one degeneracy. In this subsection we give a multivariate proof, if K has two degeneracies, for the special case where the regularization k_β is defined by

$$k_\beta(s) := \max(k(s), \beta^\mu) \tag{3.65}$$

which, as we have seen, implies $C(\beta) \leq c\, \beta^{2+\mu}$, and $m(\beta) \geq \beta^\mu$. In order to establish this theorem and to provide the Galerkin error estimates in the next section, we assume for the remainder of this paper that $K_\beta(S_\beta)$ is sufficiently regular. In particular, we assume

$$\|K_\beta(S_\beta)\|_{W^{2,\gamma}} \leq C_\gamma \left(\|\Delta K_\beta(S_\beta)\|_{L^\gamma} + 1\right). \tag{3.66}$$

This inequality holds, for example, under a diffusive flux assumption

$$\frac{\partial K_\beta(S_\beta)}{\partial n} = 0 \quad \text{on } \partial\Omega. \tag{3.67}$$

Indeed, if this condition holds, then since $1 < \gamma < 2$, it follows that $\gamma < \overline{p}(\Omega)$ and so the elliptic regularity (see A.12) holds for the Neumann problem over domains Ω which satisfy our standing assumptions. Hence,

$$\|K_\beta(S_\beta) - K_\beta(S_\beta)_\Omega\|_{W^{2,\gamma}} \leq C\,(\|\Delta K_\beta(S_\beta)\|_{L^\gamma} + \|K_\beta(S_\beta)\|_{L^\gamma}). \tag{3.68}$$

In what follows we use C and c for constants which are independent of the parameters β and h, but may depend on the Darcy velocity $\mathbf{u}$.

Theorem 3.2 *Suppose the hypotheses of Corollary 3.1 hold and β, λ are given as in (3.58). Furthermore, suppose that either condition (3.66) or condition (3.67) holds, then*

$$\|S - S_h\|_{L^\infty(L^{2+\mu})} \leq Ch^{\frac{\lambda}{2+\mu}\min(1,\mu)} \tag{3.69}$$

$$\|K(S) - K_\beta(S_h)\|_{L^2(H^1)} \leq Ch^{\frac{\lambda}{2}\min(1,\mu)} \tag{3.70}$$

Proof. By Theorem 2.3, it suffices to establish the estimate

$$\sup_{0\le t\le T}\Big(K_\beta(S_\beta)-K_\beta(S_h),S_\beta-S_h\Big)+\Big\|\nabla(K_\beta(S_\beta)-K_\beta(S_h))\Big\|^2_{L^2(L^2)} \;\le\; C\,h^{\lambda\min(1,\mu)}.$$

Set $W_h=E_h(K_\beta(S_\beta))$, and let $\chi=W_h-K_\beta(S_h)\in M_h$. Then (3.33) becomes

$$\begin{aligned}((S_\beta)_t-(S_h)_t,W_h-K_\beta(S_h))-((f(S_\beta)-f(S_h))\mathbf{u},\nabla(W_h-K_\beta(S_h)))&\\ +(\nabla(K_\beta(S_\beta)-K_\beta(S_h)),\nabla(W_h-K_\beta(S_h)))=0&\end{aligned}\tag{3.71}$$

which can be rewritten as

$$\begin{aligned}\frac{d}{dt}(S_\beta-S_h,K_\beta(S_\beta)-K_\beta(S_h))+\|\nabla(K_\beta(S_\beta)-K_\beta(S_h))\|^2_{L^2}&\\ =\Big((f(S_\beta)-f(S_h))\mathbf{u},\nabla(W_h-K_\beta(S_h))\Big)&\\ +((S_\beta-S_h)_t,K_\beta(S_\beta)-W_h)&\\ +(S_\beta-S_h,(K_\beta(S_\beta)-K_\beta(S_h))_t)&\\ +(\nabla(K_\beta(S_\beta)-K_\beta(S_h)),\nabla(K_\beta(S_\beta)-W_h))&\end{aligned}\tag{3.72}$$

The first term on the right-hand side of (3.72) can be rewritten as

$$\begin{aligned}((f(S_\beta)-f(S_h))\mathbf{u},\nabla(W_h-K_\beta(S_h)))=((f(S_\beta)-f(S_h))\mathbf{u},\nabla(W_h-K_\beta(S_\beta)))&\\ +((f(S_\beta)-f(S_h))\mathbf{u},\nabla(K_\beta(S_\beta)-K_\beta(S_h)))&\end{aligned}\tag{3.73}$$

So we can bound this term by

$$\begin{aligned}\|(f(S_\beta)-f(S_h))\mathbf{u}\|_{L^2}\|\nabla(W_h-K_\beta(S_\beta))\|_{L^2}+C\|(f(S_\beta)-f(S_h))\mathbf{u}\|^2_{L^2}&\\ +\frac{1}{4}\|\nabla(K_\beta(S_\beta)-K_\beta(S_h))\|^2_{L^2}.&\end{aligned}\tag{3.74}$$

We can then hide the last term of (3.74) in the left hand side of (3.72). The first term of (3.74) is bounded as follows:

$$\begin{aligned}&\|(f(S_\beta)-f(S_h))\mathbf{u}\|_{L^2}\|\nabla(W_h-K_\beta(S_\beta))\|_{L^2}\\ &\qquad\le C\|f(S_\beta)-f(S_h)\|_{L^2}\|(I-E_h)K_\beta(S_\beta)\|_{H^1}\\ &\qquad\le C\|S_\beta-S_h\|_{L^2}\|(I-E_h)K_\beta(S_\beta)\|_{H^1}\end{aligned}\tag{3.75}$$

where we have used the fact that f is Lipschitz. Applying the elliptic approximation estimate for H^1 to this inequality, followed by the estimate (3.66), we then get

$$\begin{aligned}\Big|\Big((f(S_\beta)-f(S_h))\mathbf{u},\nabla(W_h-K_\beta(S_\beta))\Big)\Big| &\le C\,h\,\|S_\beta-S_h\|_{L^2}|K_\beta(S_\beta)|_{H^2}\\ &\le C\,h\,\|S_\beta-S_h\|_{L^2}(\|\Delta K_\beta(S_\beta)\|_{L^2}+1).\end{aligned}\tag{3.76}$$

The second term on the righthand of (3.72) is bounded similarly:

$$\begin{aligned}\Big|\Big((S_\beta-S_h)_t,W_h-K_\beta(S_\beta)\Big)\Big| &\le \|(S_\beta-S_h)_t\|_{L^2}\|(I-E_h)K_\beta(S_\beta)\|_{L^2}\\ &\le C\,h^2\,\|(S_\beta-S_h)_t\|_{L^2}(\|\Delta K_\beta(S_\beta)\|_{L^2}+1),\end{aligned}\tag{3.77}$$

while the third term is estimated by

$$\Big|\Big(S_\beta-S_h,(K_\beta(S_\beta)-K_\beta(S_h))_t\Big)\Big| \le \|S_\beta-S_h\|_{L^{2+\mu}}\|(K_\beta(S_\beta)-K_\beta(S_h))_t\|_{L^\gamma}.\tag{3.78}$$

Finally, the last term on the righthand side of (3.72) is bounded as follows:

$$
\begin{aligned}
&\Big|\Big(\nabla(K_\beta(S_\beta)-K_\beta(S_h)),\nabla(I-E_h)K_\beta(S_\beta)\Big)\Big| \\
&\qquad\qquad \le \frac{1}{4}\|\nabla(K_\beta(S_\beta)-K_\beta(S_h))\|^2_{L^2} + C\,|(I-E_h)K_\beta(S_\beta)|^2_{H^1} \\
&\qquad\qquad \le \frac{1}{4}\|\nabla(K_\beta(S_\beta)-K_\beta(S_h))\|^2_{L^2} + C\,h^2(\|\Delta K_\beta(S_\beta)\|_{L^2}+1). \qquad (3.79)
\end{aligned}
$$

We can hide the first term on the righthand side of (3.79) in the left hand side of (3.72). Now combining the inequalities (3.72), (3.76)-(3.79) together, we obtain

$$
\begin{aligned}
&\frac{d}{dt}(S_\beta-S_h,K_\beta(S_\beta)-K_\beta(S_h)) + \frac{1}{2}\|\nabla(K_\beta(S_\beta)-K_\beta(S_h))\|^2_{L^2} \\
&\quad \le \ C\left\{h\|S_\beta-S_h\|_{L^2} \ + \ h^2\ \|(S_\beta-S_h)_t\|_{L^2} + h^2\right\}(\|\Delta K_\beta(S_\beta)\|_{L^2}+1) \\
&\qquad + C\left\{\|S_\beta-S_h\|_{L^{2+\mu}}\|(K_\beta(S_\beta)-K_\beta(S_h))_t\|_{L^\gamma} + \|f(S_\beta)-f(S_h)\|_{L^2}\right\}. \qquad (3.80)
\end{aligned}
$$

Now integrate over the interval $[0,T_0]$ and use Hölder's inequality to get

$$
\begin{aligned}
&\max_{0\le t\le T_0}(S_\beta-S_h,K_\beta(S_\beta)-K_\beta(S_h))(t) + \frac{1}{2}\|\nabla(K_\beta(S_\beta)-K_\beta(S_h))\|^2_{L^2(L^2)} \\
&\quad \le \ C\left\{h\|S_\beta-S_h\|_{L^{2+\mu}(L^{2+\mu})} \ + \ h^2\|(S_\beta-S_h)_t\|_{L^2(L^2)} \ + \ h^2\right\}h^{-\frac{\lambda\mu}{2}} \\
&\qquad + C\left\{\|S_\beta-S_h\|_{L^{2+\mu}(L^{2+\mu})}\|(K_\beta(S_\beta)-K_\beta(S_h))_t\|_{L^\gamma(L^\gamma)} + \|(f(S_\beta)-f(S_h))\mathbf{u}\|_{L^2(L^2)}\right\} \\
&\qquad + \ \Big(S^0-S_h(0),K_\beta(S^0)-K_\beta(S_h(0))\Big) \qquad (3.81)
\end{aligned}
$$

where we used the fact from (2.16) that $\|\Delta K_\beta(S_\beta)\|_{L^2(L^2)} \le c\,m(\beta)^{-\frac{1}{2}}$. The last term of (3.81) is $\mathcal{O}(h^\lambda)$, by (3.59). Therefore, using (2.11), (3.61), (3.28), (2.15), (2.16) and (3.58), respectively, we get

$$
\begin{aligned}
&\sup_{0\le t\le T_0}(K_\beta(S_\beta)-K_\beta(S_h),S_\beta-S_h) + \frac{1}{2}\|\nabla(K_\beta(S_\beta)-K_\beta(S_h))\|^2_{L^2(L^2)} \\
&\qquad\qquad \le \ C\{h^{\lambda+1}h^{-\frac{\mu\lambda}{2}} + h^2h^{-\mu\lambda} + h^\lambda + h^2h^{-\mu\lambda} + h^\lambda\} \\
&\qquad\qquad \le \ Ch^\lambda \qquad (3.82)
\end{aligned}
$$

The proof is completed by using (2.4)-(2.5) and (2.19)-(2.20) together with the triangle inequality. □

4 Discrete Galerkin Method

In the previous section we have derived error estimates for a continuous Galerkin method (the time variable remaining continuous) applied to the regularized problem (1.7)–(1.9). In this section we discretize in time (backward scheme) and give corresponding error estimates. Again, here we follow M. E. Rose's analysis in [16], but are generalizing to the multidimensional case with two degeneracies, and making modifications of the analysis to establish a higher rate of convergence in the time step.

The finite difference scheme used here is implicit with $V_h^n = FV_h^{n+1}$, where F is a nonlinear function. Thus one must show that this function is invertible for Δt chosen sufficiently small. We show this is the case when $\mathbf{u}$ constant in time. Therefore at each step $V_h^n \mapsto V_h^{n+1}$ is well defined which leads to a sequence of nonlinear algebraic equations. Finally we derive the error estimates for $S - S_h^n$, where $S_h^n = H_\beta(V_h^n)$ and H_β is defined by (4.13).

4.1 The Discretized Problem and Existence of a Solution

We consider the backward–difference time discretization

$$
\left(\frac{H_\beta(V_h^{n+1})-H_\beta(V_h^n)}{\Delta t},\chi\right) - \big(f(H_\beta(V_h^{n+1}))\mathbf{u}^{n+1},\nabla\chi\big) + (\nabla V_h^{n+1},\nabla\chi) = 0 \qquad (4.1)
$$

for all $\chi \in M_h$, $n = 0, 1, ..., N-1$ with

$$P_h H_\beta V_h^0 = P_h S^0 \tag{4.2}$$

The operator $P_h H_\beta$ is bijective and the results of [16] guarantee a solution V_h^0. We show that for the time step sufficiently small, the mapping $V_h^n \to V_h^{n+1}$ given by 4.1 is well defined. We get from (4.1) that

$$V_h^n = F V_h^{n+1} \tag{4.3}$$

for some function $F : M_h \to M_h$. To show $V_h^n \to V_h^{n+1}$ is well defined it is enough to show the function F appearing in (4.3) is bijective. We show this in two ways, the first indicating why the nonlinear equations can be solved for small time step, and the second giving a quantitative estimate for Δt to guarantee invertibility.

Indirect argument for invertibility.

Let $(e_i)_1^m$ be a basis for M_h, where $m = m(h) = \dim(M_h)$. Set $\chi = e_i$ in (4.1). Then

$$(H_\beta(V_h^n), e_i) \;=\; (H_\beta(V_h^{n+1}), e_i) - \Delta t\{(f(H_\beta(V_h^{n+1}))\mathbf{u}^{n+1}, \nabla e_i) - (\nabla V_h^{n+1}, \nabla e_i)\} \tag{4.4}$$

for $0 \le i \le N-1$. This can be rewritten in the vector form

$$P_h H_\beta(V_h^n) = P_h H_\beta(V_h^{n+1}) + \Delta t \mathcal{F}(P_h H_\beta(V_h^{n+1})) \tag{4.5}$$

where $\mathcal{F}(P_h H_\beta(V_h^{n+1}))$ is the vector in M_h with components $(-f(H_\beta(V_h^{n+1})) + \nabla V_h^{n+1}, \nabla e_i)$. If we choose $\max \Delta t$ sufficiently small then we see by (4.5) that the mapping $P_h H_\beta(V_h^{n+1}) \mapsto P_h H_\beta(V_h^n)$ is bijective, and, since $P_h H_\beta$ is bijective, we deduce that $V_h^{n+1} \to V_h^n$ is bijective for small Δt.

Direct argument for invertibility.

Here we provide a sufficient condition on Δt to show that the nonlinear function F is bijective and that the nonlinear equations are invertible. We assume for this analysis that $\mathbf{u}$ is constant in time, or at least its variation in time is negligible.

Let V_h^n (resp. V_h^m) be the iterated solution at time $n\Delta t$ (resp. $m\Delta t$). Then by (4.1) we have

$$\begin{aligned}\left(\frac{P_h H_\beta(V_h^n) - P_h H_\beta(V_h^m)}{\Delta t}, \chi\right) &+ ((f(H_\beta(V_h^{n+1})) - f(H_\beta(V_h^{m+1})))\mathbf{u}, \nabla\chi) \\ &= \left(\frac{P_h H_\beta(V_h^{n+1}) - P_h H_\beta(V_h^{m+1})}{\Delta t}, \chi\right) + (\nabla(V_h^{n+1} - V_h^{m+1}), \nabla\chi).\end{aligned} \tag{4.6}$$

Set $\chi = V_h^{n+1} - V_h^{m+1}$, and use the relation $V_h^n = F V_h^{n+1}$ to get

$$\begin{aligned}&\left(\frac{P_h H_\beta(F V_h^{n+1}) - P_h H_\beta(F V_h^{m+1})}{\Delta t}, V_h^{n+1} - V_h^{m+1}\right) \\ &\quad + \quad ((f(H_\beta(V_h^{n+1})) - f(H_\beta(V_h^{m+1})))\mathbf{u}, \nabla(V_h^{n+1} - V_h^{m+1})) \\ &\quad = \quad \left(\frac{P_h H_\beta(V_h^{n+1}) - P_h H_\beta(V_h^{m+1})}{\Delta t}, V_h^{n+1} - V_h^{m+1}\right) + \|\nabla(V_h^{n+1} - V_h^{m+1})\|_{L^2(\Omega)}^2.\end{aligned} \tag{4.7}$$

We can bound the second term of the left hand side of (4.7) as follows:

$$\begin{aligned}&\left|((f(H_\beta(V^{n+1})) - f(H_\beta(V_h^{m+1})))\mathbf{u}, \nabla(V_h^{n+1} - V_h^{m+1}))\right| \\ &\qquad \le \frac{1}{2}\|(f(H_\beta(V_h^{n+1})) - f(H_\beta(V_h^{n+1}))\mathbf{u}\|_{L^2}^2 + \frac{1}{2}\|\nabla(V_h^{n+1} - V_h^{m+1})\|_{L^2}^2.\end{aligned} \tag{4.8}$$

Now using Lemma 2.1 applied to K_β (in place of K) we get

$$
\begin{aligned}
\Big|((f(H_\beta(V^{n+1})) - f(H_\beta(V_h^{m+1})))\mathbf{u}, \nabla(V_h^{n+1} - V_h^{m+1}))\Big| \\
&\le \frac{\|\mathbf{u}\|^2_{L^\infty(L^\infty)}}{2C^*}(H_\beta(V_h^{n+1}) - H_\beta(V_h^{m+1}), V_h^{n+1} - V_h^{m+1}) \\
&\quad + \frac{1}{2}\|\nabla(V_h^{n+1} - V_h^{m+1})\|^2_{L^2} \\
&= \frac{\|\mathbf{u}\|^2_{L^\infty(L^\infty)}}{2C^*}(P_hH_\beta(V_h^{n+1}) - P_hH_\beta(V_h^{m+1}), V_h^{n+1} - V_h^{m+1}) \\
&\quad + \frac{1}{2}\|\nabla(V_h^{n+1} - V_h^{m+1})\|^2_{L^2}.
\end{aligned}
\tag{4.9}
$$

So the identity (4.7) leads to

$$
\begin{aligned}
&\Big(P_hH_\beta(FV_h^{n+1}) - P_hH_\beta(FV_h^{m+1}), V_h^{n+1} - V_h^{m+1}\Big) \\
&\quad + \Delta t\frac{\|\mathbf{u}\|^2_{L^\infty(L^\infty)}}{2C^*}\Big(P_hH_\beta(V_h^{n+1}) - P_hH_\beta(V_h^{m+1}), V_h^{n+1} - V_h^{m+1}\Big) \\
&\qquad \ge (P_hH_\beta(V_h^{n+1}) - P_hH_\beta(V_h^{m+1}), V_h^{n+1} - V_h^{m+1}) \\
&\qquad\quad + \frac{1}{2}\Delta t\|\nabla(V_h^{n+1} - V_h^{m+1})\|^2_{L^2},
\end{aligned}
\tag{4.10}
$$

and so,

$$
\begin{aligned}
&(P_hH_\beta(FV_h^{n+1}) - P_hH_\beta(FV_h^{m+1}), V_h^{n+1} - V_h^{m+1}) \\
&\quad \ge \left(1 - \Delta t\frac{\|\mathbf{u}\|^2_{L^\infty(L^\infty)}}{2C^*}\right)\left(P_hH_\beta(V_h^{n+1}) - P_hH_\beta(V_h^{m+1}), V_h^{n+1} - V_h^{m+1}\right) \\
&\quad\quad + \frac{1}{2}\Delta t\|\nabla(V_h^{n+1} - V_h^{m+1})\|^2_{L^2}.
\end{aligned}
\tag{4.11}
$$

Now we use the fact that P_hH_β is coercive [16] to get from (4.11)

$$
\begin{aligned}
&(P_hH_\beta(FV_h^{n+1}) - P_hH_\beta(FV_h^{m+1}), V_h^{n+1} - V_h^{m+1}) \\
&\quad \ge \left(1 - \Delta t\frac{\|\mathbf{u}\|^2_{L^\infty(L^\infty)}}{2C^*}\right)\|V_h^{n+1} - V_h^{m+1}\|^2_{L^2} \\
&\quad\quad + \frac{1}{2}\Delta t\|\nabla(V_h^{n+1} - V_h^{m+1})\|^2_{L^2}.
\end{aligned}
\tag{4.12}
$$

Thus, if the standard type of existence condition on the time step

$$
\left(1 - \Delta t\frac{\|\mathbf{u}\|^2_{L^\infty(L^\infty)}}{2C^*}\right) > 0
\tag{4.13}
$$

for nonlinear equations is satisfied, then $P_hH_\beta F$ is bijective; since it is clearly continuous [16], [3] and P_hH_β is bijective [16], then F is also bijective. Thus by taking Δt to satisfy (4.13), we can perform the backward solve to produce the solutions $V_h^0, V_h^1, ..., V_h^n$ to the sequence of nonlinear algebraic equations which approximate the solution to the nonlinear differential equations (3.14)–(3.15).

4.2 Error Analysis

We are interested in estimating the error $S(t_n) - H_\beta(V_h^n)$ of the discrete time Galerkin approximation to the differential equations. We set $S_h^n := H_\beta(V_h^n)$, in which case the equations (4.1) and (4.2), respectively, become

$$\left(\frac{S_h^{n+1} - S_h^n}{\Delta t}, \chi\right) - (f(S_h^{n+1})\mathbf{u}^{n+1}, \nabla\chi) + (\nabla K_\beta(S_h^{n+1}), \nabla\chi) = 0 \qquad \text{for all } \chi \in M_h \tag{4.14}$$

$$P_h S_h^0 = P_h S^0 \tag{4.15}$$

where S^0 is the initial saturation given as in (1.3). We then have the following fully discretized version of Theorem 3.1.

Theorem 4.1 *Suppose μ, γ are defined by (1.6) and $m(\beta)$ is as defined in (2.12). Let S_β be the solution to the regularized equation (1.7)-(1.9) with $Q = 0$ and $q = 0$, and with coefficients f and k which satisfy conditions (1.4)-(1.5). Let $S_h^n = H_\beta(V_h^n)$, where $V_h^n \in M_h$, $n = 0, 1, ..., N-1$ solves (4.1)-(4.2). Then*

$$\max_{0\le n\le N} \|S_\beta^n - S_h^n\|^2_{(H^1)^*} + \eta \sum_0^{N-1} \Delta t \left(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1}), S_\beta^{n+1} - S_h^{n+1}\right)$$
$$\le C\{h^{2\gamma} m(\beta)^{-\frac{1}{1+\mu}} + \Delta t^{\frac{\gamma+2}{2}}\} \tag{4.16}$$

where $S_\beta^n := S_\beta(t_n)$.

Proof. Subtract (4.14) from (3.32) to get

$$\left(\frac{S_\beta^{n+1} - S_\beta^n}{\Delta t} - \frac{S_h^{n+1} - S_h^n}{\Delta t}, \chi\right) - ((f(S_\beta^{n+1}) - f(S_h^n))\mathbf{u}^{n+1}, \nabla\chi)$$
$$+(\nabla(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1})), \nabla\chi)$$
$$+\left(\frac{\partial S_\beta^{n+1}}{\partial t} - \frac{S_\beta^{n+1} - S_\beta^n}{\Delta t}, \chi\right) = 0 \tag{4.17}$$

for all $\chi \in M_h$. If we set

$$\partial^+\phi^n = \frac{\phi^{n+1} - \phi^n}{\Delta t},$$

then (4.17) can be rewritten as

$$\left(\partial^+(S_\beta - S_h)^n, \chi\right) - \left((f(S_\beta^{n+1}) - f(S_h^{n+1}))\mathbf{u}^{n+1}, \nabla\chi\right)$$
$$+\left(\frac{\partial S_\beta^{n+1}}{\partial t} - \frac{S_\beta^{n+1} - S_\beta^n}{\Delta t}, \chi\right) + (\nabla(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1})), \nabla\chi) = 0. \tag{4.18}$$

Next we set $\chi = T_h(S_\beta^{n+1} - S_h^{n+1})$ in (4.18) to get

$$\left(\partial^+ P_h(S_\beta - S_h)^n, T_h(S_h^{n+1} - S_h^{n+1})\right)$$
$$+\left(\nabla(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1})), \nabla T_h(S_\beta^{n+1} - S_h^{n+1})\right)$$
$$= \left((f(S_\beta^{n+1}) - f(S_h^{n+1}))\mathbf{u}^{n+1}, \nabla T_h(S_\beta^{n+1} - S_h^{n+1})\right)$$
$$-\left(\frac{\partial S_\beta^{n+1}}{\partial t} - \partial^+ S_\beta^n, T_h(S_\beta^{n+1} - S_h^{n+1})\right), \tag{4.19}$$

where T_h is defined by (A.24).

The identity

$$\frac{1}{2\Delta t}\{\|\phi^{n+1}\|^2_{H_h^{-1}} - \|\phi^n\|^2_{H_h^{-1}}\} + \frac{\Delta t}{2}\|\partial^+\phi^n\|^2_{H_h^{-1}} = \big(T_h(\partial^+\phi^n), \phi^{n+1}\big), \tag{4.20}$$

which is the discretized analogue of the fact $\big(T\frac{\partial\phi}{\partial t}, \phi\big) = \frac{1}{2}\frac{d}{dt}(T\phi, \phi)$, is established using the definition of E_h and T_h for $\|\phi\|^2_{H_h^{-1}} = (T_h\phi, \phi)$, and properties of $T_h = E_hT$. If we apply this identity to $\phi = P_h(S_\beta - S_h) \in M_h$, then we obtain the estimate

$$\begin{aligned}\frac{1}{2\Delta t}&\left\{\|P_h(S_\beta^{n+1} - S_h^{n+1})\|^2_{H_h^{-1}} - \|P_h(S_\beta^n - S_h^n)\|^2_{H_h^{-1}}\right\}\\ &\leq \frac{1}{\Delta t}\big(P_h(S_\beta^{n+1} - S_h^{n+1}) - P_h(S_\beta^n - S_h^n), T_h(S_\beta^{n+1} - S_h^{n+1})\big).\end{aligned} \tag{4.21}$$

Upon substituting this estimate into identity (4.19) we obtain

$$\begin{aligned}\frac{1}{2\Delta t}\|P_h(S_\beta^{n+1} - S_h^{n+1})\|^2_{H_h^{-1}} &- \frac{1}{2\Delta t}\|P_h(S_\beta^n - S_h^n)\|^2_{H_h^{-1}}\\ &+\big(\nabla(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1})), \nabla T_h(S_\beta^{n+1} - S_h^{n+1})\big)\\ \leq\ & \big((f(S_\beta^{n+1}) - f(S_h^{n+1}))\mathbf{u}^{n+1}, \nabla T_h(S_\beta^{n+1} - S_h^{n+1})\big)\\ &-\left(\frac{\partial S_\beta^{n+1}}{\partial t} - \frac{S_\beta^{n+1} - S_\beta^n}{\Delta t}, T_h(S_\beta^{n+1} - S_h^{n+1})\right).\end{aligned} \tag{4.22}$$

But, we get by definition of E_h (See (A.18), and because $T_h(S_\beta^{n+1} - S_h^{n+1}) \in M_h$,

$$\begin{aligned}&\big(\nabla(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1})), \nabla T_h(S_\beta^{n+1} - S_h^{n+1})\big)\\ &\qquad = \big(\nabla E_h(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1})), \nabla T_h(S_\beta^{n+1} - S_h^{n+1})\big).\end{aligned} \tag{4.23}$$

Next using the definition of T_h (see (A.22) and (A.24)), and the fact that $(S_\beta^{n+1} - S_h^{n+1})_\Omega = 0$ (set $\chi = 1$ in (4.14) and use (3.15)), we have

$$\begin{aligned}&\big(\nabla(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1})), \nabla T_h(S_\beta^{n+1} - S_h^{n+1})\big)\\ &\quad= \big(E_h(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1})), (S_\beta^{n+1} - S_h^{n+1})\big)\\ &\quad= \big(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1}), S_\beta^{n+1} - S_h^{n+1}\big)\\ &\qquad+\Big((E_h - I)K_\beta(S_\beta^{n+1}), S_\beta^{n+1} - S_h^{n+1}\Big)\end{aligned} \tag{4.24}$$

since $V_h^{n+1} = K_\beta(S_h^{n+1}) \in M_h$. We combine estimates (4.22) and (4.24) to yield

$$\begin{aligned}\frac{1}{2\Delta t}\|P_h(S_\beta^{n+1} - S_h^{n+1})\|^2_{(H^1)^*} &- \frac{1}{2\Delta t}\|P_h(S_\beta^n - S_h^n)\|^2_{(H^1)^*}\\ &+\big(K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1}), S_\beta^{n+1} - S_h^{n+1}\big)\\ \leq\ & \big((f(S_\beta^{n+1}) - f(S_h^{n+1}))\mathbf{u}^{n+1}, \nabla T_h(S_\beta^{n+1} - S_h^{n+1})\big)\\ &+\left|\Big((I - E_h)K_\beta(S_\beta^{n+1}), S_\beta^{n+1} - S_h^{n+1}\Big)\right|\\ &-\left(\frac{\partial S_\beta}{\partial t} - \frac{S_\beta^{n+1} - S_\beta^n}{\Delta t}, T_h(S_\beta^{n+1} - S_h^{n+1})\right).\end{aligned} \tag{4.25}$$

The first term on the righthand side of (4.25) can be bounded as follows:

$$\begin{aligned}&\big((f(S_\beta^{n+1}) - f(S_h^{n+1}))\mathbf{u}^{n+1}, \nabla T_h(S_\beta^{n+1} - S_h^{n+1})\big)\\ &\quad\leq \frac{1}{2}C^*\|f(S_\beta^{n+1}) - f(S_h^{n+1})\|^2_{L^2} + \frac{1}{2C^*}\|\nabla T_hP_h(S_\beta^{n+1} - S_h^{n+1})\|^2_{L^2}\\ &\quad\leq \frac{1}{2}C^*\|f(S_\beta^{n+1}) - f(S_h^{n+1})\|^2_{L^2} + \frac{1}{2C^*}\|P_h(S_\beta^{n+1} - S_h^{n+1})\|^2_{H_h^{-1}}\end{aligned} \tag{4.26}$$

where C^* is as in (2.2). By Lemma 2.1 the first term on the righthand side of (4.26) can be hidden in the left hand side of (4.25).

The second term on the righthand side of (4.25) is bounded as follows:

$$|((I-E_h)K_\beta(S_\beta^{n+1}), S_\beta^{n+1}-S_h^{n+1})| \leq C\|(I-E_h)K_\beta(S_\beta^{n+1})\|_{L^\gamma}^\gamma + \frac{C^{**}}{8}\|S_\beta^{n+1}-S_h^{n+1}\|_{L^{2+\mu}}^{2+\mu} \tag{4.27}$$

by the arithmetic–geometric mean inequality. We can again hide the second term on the righthand side of (4.27) in the left hand side of (4.25). Using the error estimate for the elliptic projection and the inequality (3.66), we have for the first term

$$\|(I-E_h)K_\beta(S_\beta^{n+1})\|_{L^\gamma}^\gamma \leq Ch^{2\gamma}\|\Delta K_\beta(S_\beta^{n+1})\|_{L^\gamma}^\gamma \tag{4.28}$$

It remains to deal with the last term on the righthand side of (4.25). For this, we have

$$\frac{\partial S_\beta^{n+1}}{\partial t} - \frac{S_\beta^{n+1}-S_\beta^n}{\Delta t} = \frac{1}{\Delta t}\int_{I_n} S_{\beta tt}(\tau)(t_{n+1}-\tau)d\tau \tag{4.29}$$

where we use the Taylor expansion of S_β about t_n and $I_n := [t_n, t_{n+1}]$. By (1.7) it follows that

$$S_{\beta tt} = (-\nabla\cdot f(S_\beta)\mathbf{u} + \Delta K_\beta(S_\beta))_t \tag{4.30}$$

Using (4.29) and (4.30) in the last term of (4.25) we get

$$\begin{aligned}&(\frac{\partial S_\beta^{n+1}}{\partial t} - \frac{S_\beta^{n+1}-S_\beta^n}{\Delta t}, T_h(S_\beta^{n+1}-S_h^{n+1}))\\ &= \left(\frac{1}{\Delta t}\int_{I_n}((-\nabla\cdot f(S_\beta)\mathbf{u})_t(\tau) + (\Delta K_\beta(S_\beta))_t(\tau))(t_{n+1}-\tau)d\tau, T_h(S_\beta^{n+1}-S_h^{n+1})\right)\end{aligned} \tag{4.31}$$

The divergence theorem together with (1.8) and (4.31) give

$$\begin{aligned}&\left(\frac{\partial S_\beta^{n+1}}{\partial t} - \frac{S_\beta^{n+1}-S_\beta^n}{\Delta t}, T_h(S_\beta^{n+1}-S_h^{n+1})\right)\\ &= (\frac{1}{\Delta t}\int_{I_n}(f(S_\beta)u)_t(\tau)(t_{n+1}-\tau)d\tau, \nabla T_h(S_\beta^{n+1}-S_h^{n+1}))\\ &\quad - \left(\frac{1}{\Delta t}\int_{I_n}(\nabla K_\beta(S_\beta))_t(t_{n+1}-\tau)d\tau, \nabla T_h(S_\beta^{n+1}-S_h^{n+1})\right)\end{aligned} \tag{4.32}$$

We estimate each of the terms of the righthand side of (4.32) separately.

The first term is bounded as follows:

$$\begin{aligned}&\left(\frac{1}{\Delta t}\int_{I_n}(f(S_\beta)\mathbf{u})_t(\tau)(t_{n+1}-\tau)d\tau, \nabla T_h(S_\beta^{n+1}-S_h^{n+1})\right)\\ &\leq \frac{1}{\Delta t}\int_{I_n}\|(f(S_\beta)\mathbf{u})_t(\tau)\|_{L^2}|t_{n+1}-\tau|d\tau\|\nabla T_h(S_\beta^{n+1}-S_h^{n+1})\|_{L^2}\end{aligned} \tag{4.33}$$

By the Cauchy–Schwartz inequality we have

$$\begin{aligned}&\frac{1}{\Delta t}\int_{I_n}\|(f(S_\beta)\mathbf{u})_t(\tau)\|_{L^2}|t_{n+1}-\tau|d\tau\\ &\leq \frac{1}{\Delta t}\|(f(S_\beta)\mathbf{u})_t\|_{L^2(L^2(\Omega),I_n)}\left(\int_{t_n}^{t_{n+1}}(t_{n+1}-\tau)^2d\tau\right)^{\frac{1}{2}}\\ &= (\Delta t)^{\frac{1}{2}}\|(f(S_\beta)\mathbf{u})_t\|_{L^2(L^2(\Omega),I_n)}\end{aligned} \tag{4.34}$$

So (4.33) becomes

$$\begin{aligned}
&\left(\frac{1}{\Delta t}\int_{I_n}(f(S_\beta)\mathbf{u})_t(\tau)(t_{n+1}-\tau)d\tau, \nabla T_h P_h(S_\beta^{n+1}-S_h^{n+1})\right)\\
&\qquad \leq \quad (\Delta t)^{\frac{1}{2}}\|(f(S_\beta)\mathbf{u})_t\|_{L^2(L^2(\Omega),I_n)}\|\nabla T_h P_h(S_\beta^{n+1}-S_h^{n+1})\|_{L^2}\\
&\qquad \leq \quad (\Delta t)/2\|(f(S_\beta)\mathbf{u})_t\|^2_{L^2(L^2(\Omega),I_n)}+\frac{1}{2}\|\nabla T_h P_h(S_\beta^{n+1}-S_h^{n+1})\|^2_{L^2}\\
&\qquad = \quad (\Delta t)/2\|(f(S_\beta)\mathbf{u})_t\|^2_{L^2(L^2(\Omega),I_n)}+\frac{1}{2}\|P_h(S_\beta^{n+1}-S_h^{n+1})\|^2_{H_h^{-1}} \qquad (4.35)
\end{aligned}$$

where we used the identity (A.27) for the last equality .

The second term on the righthand side of (4.32) can be rewritten as follows:

$$\begin{aligned}
&\left(\frac{1}{\Delta t}\int_{I_n}(\nabla K_\beta(S_\beta))_t(t_{n+1}-\tau)d\tau, \nabla T_h(S_\beta^{n+1}-S_h^{n+1})\right)\\
&\qquad = \left(\frac{1}{\Delta t}\int_{I_n}(\nabla E_h K_\beta(S_\beta))_t(t_{n+1}-\tau)d\tau, \nabla T_h(S_\beta^{n+1}-S_h^{n+1})\right)\\
&\qquad = \left(\frac{1}{\Delta t}\int_{I_n}(E_h K_\beta(S_\beta))_t(t_{n+1}-\tau)d\tau, S_\beta^{n+1}-S_h^{n+1}\right). \qquad (4.36)
\end{aligned}$$

Here we made use of the definition of E_h (see (A.18)) and the definition of T_h (see (A.22)–(A.24)). Using again the Cauchy–Schwartz inequality and (4.36) we get

$$\begin{aligned}
&|(\frac{1}{\Delta t}\int_{I_n}(\nabla K_\beta(S_\beta))_t(t_{n+1}-\tau)d\tau, \nabla T_h(S_\beta^{n+1}-S_h^{n+1}))|\\
&\qquad \leq \frac{1}{\Delta t}\int_{I_n}\|E_h K_\beta(S_\beta)_t\|_{L^2}\,(t_{n+1}-\tau)d\tau\,\|S_\beta^{n+1}-S_h^{n+1}\|_{L^2} \qquad (4.37)
\end{aligned}$$

Also

$$\begin{aligned}
&\int_{I_n}\|E_h(K_\beta(S_\beta))_t(\tau)\|\,(t_{n+1}-\tau)d\tau\\
&\qquad \leq \|E_h K_\beta(S_\beta)_t\|_{L^2(L^2(\Omega),I_n)}\left(\int_{t_n}^{t_{n+1}}(t_{n+1}-\tau)^2 d\tau\right)^{\frac{1}{2}}\\
&\qquad \leq \|E_h K_\beta(S_\beta)_t\|_{L^2(L^2(\Omega),I_n)}(\Delta t)^{\frac{3}{2}} \qquad (4.38)
\end{aligned}$$

So (5.37) becomes

$$\begin{aligned}
&|(\frac{1}{\Delta t}\int_{I_n}(\nabla K_\beta(S_\beta))_t(t_{n+1}-\tau)d\tau, \nabla T_h(S_\beta^{n+1}-S_h^{n+1}))|\\
&\qquad \leq (\Delta t)^{\frac{1}{2}}\|E_h K_\beta(S_\beta)_t\|_{L^2(L^2(\Omega),I_n)}\;\|S_\beta^{n+1}-S_h^{n+1}\|_{L^2(\Omega)} \qquad (4.39)
\end{aligned}$$

which then gives

$$\begin{aligned}
&|(\frac{1}{\Delta t}\int_{I_n}(\nabla K_\beta(S_\beta))_t(t_{n+1}-\tau)d\tau, \nabla T_h(S_\beta^{n+1}-S_h^{n+1}))|\\
&\qquad \leq \quad C(\Delta t)^{\frac{\gamma}{2}}\|E_h K_\beta(S_\beta)_t\|^{\gamma}_{L^2(L^2(\Omega),I_n)}\\
&\qquad\qquad +\frac{C^{**}}{8}\|S_\beta^{n+1}-S_h^{n+1}\|^{2+\mu}_{L^2(\Omega)}\\
&\qquad \leq \quad C(\Delta t)^{\frac{\gamma}{2}}\|E_h K_\beta(S_\beta)_t\|^{\gamma}_{L^2(L^2(\Omega),I_n)}\\
&\qquad\qquad +\frac{C^{**}}{8}\|S_\beta^{n+1}-S_h^{n+1}\|^{2+\mu}_{L^{2+\mu}(\Omega)}. \qquad (4.40)
\end{aligned}$$

Now using estimates (4.21), (4.26), (4.27), (4.28), (4.35) and (4.40), after hiding the appropriate terms, we get

$$\begin{aligned}
&\frac{1}{2\Delta t}\|P_h(S_\beta^{n+1}-S_h^{n+1})\|^2_{H_h^{-1}}-\frac{1}{2\Delta t}\|P_h(S_\beta^{n}-S_h^{n})\|^2_{H_h^{-1}}\\
&\qquad+\frac{1}{4}\big(K_\beta(S_\beta^{n+1})-K_\beta(S_h^{n+1}),S_\beta^{n+1}-S_h^{n+1}\big)\\
&\quad\le C\{\|P_h(S_\beta^{n+1}-S_h^{n+1})\|^2_{H_h^{-1}}+h^{2\gamma}\|\Delta K_\beta(S_\beta^{n+1})\|^\gamma_{L^\gamma}\\
&\qquad+\frac{\Delta t}{2}\|(f(S_\beta)\mathbf{u})_t\|^2_{L^2(L^2(\Omega),I_n)}\\
&\qquad+(\Delta t)^{\frac{\gamma}{2}}\|E_hK_\beta(S_\beta)_t\|^\gamma_{L^2(L^2(\Omega),I_n)}\}
\end{aligned}\tag{4.41}$$

Next, multiplying (4.41) by Δt, summing from $n=0$ to $n=m-1$, with $0<m\le N$, and using the fact that $P_h(S_\beta^0-S_h^0)=0$, we obtain

$$\begin{aligned}
&\frac{1}{2}\|P_h(S_\beta^{m}-S_h^{m})\|^2_{H_h^{-1}}+\frac{1}{4}\sum_0^{m-1}\Delta t\big(K_\beta(S_\beta^{n+1})-K_\beta(S_h^{n+1}),S_\beta^{n+1}-S_h^{n+1}\big)\\
&\quad\le C\left\{\sum_0^{m-1}\Delta t\|P_h(S_\beta^{n+1}-S_h^{n+1})\|^2_{H_h^{-1}}+h^{2\gamma}\sum_0^{m-1}\Delta t\|\Delta K_\beta(S_\beta^{n+1})\|^\gamma_{L^\gamma}\right.\\
&\qquad+\Delta t^2\sum_0^{m-1}\|(f(S_\beta)\mathbf{u})_t\|^2_{L^2(L^2(\Omega),I_n)}\\
&\qquad\left.+(\Delta t)^{\frac{\gamma+2}{2}}\sum_0^{m-1}\|E_hK_\beta(S_\beta)_t\|^\gamma_{L^2(L^2(\Omega),I_n)}\right\}.
\end{aligned}\tag{4.42}$$

Next using the discrete Grönwall Lemma (see [10],[6]) and the fact $1\le\gamma\le 2$, we get

$$\begin{aligned}
&\max_{0\le n\le N}\|P_h(S_\beta^{n}-S_h^{n})\|^2_{H_h^{-1}}+\eta\sum_0^{N-1}\Delta t\big(K_\beta(S_\beta^{n+1})-K_\beta(S_h^{n+1}),S_\beta^{n+1}-S_h^{n+1}\big)\\
&\quad\le C\left\{h^{2\gamma}\sum_0^{N-1}\Delta t\|\Delta K_\beta(S_\beta^{n+1})\|^\gamma_{L^\gamma}\right.\\
&\qquad\left.+(\Delta t)^{\frac{\gamma+2}{2}}(\|(f(S_\beta)\mathbf{u})_t\|^2_{L^2(L^2)}+\|E_hK_\beta(S_\beta)_t\|^\gamma_{L^2(L^2)})\right\}
\end{aligned}\tag{4.43}$$

Since

$$\sum_{0\le n\le N}\Delta t\|K_\beta(S_\beta^n)\|^\gamma_{W^{2,\gamma}}\to\|K_\beta(S_\beta)\|^\gamma_{L^\gamma(W^{2,\gamma})}$$

as $\Delta t\to 0$, we have

$$\sum_{0\le n\le N}\Delta t\|K_\beta(S_\beta^n)\|^\gamma_{W^{2,\gamma}}\le Cm(\beta)^{-\frac{1}{1+\mu}}\tag{4.44}$$

where $m(\beta)$ is defined by (2.12), and we have used the inequalities (3.66) and (2.14) together with the equation (1.7).

For the second term of the righthand side of (4.43), we observe by (2.11) and (2.3), and [18] that $\|f(S_\beta)_t\|_{L^2(L^2)}$ and $\|E_hK_\beta(S_\beta)_t\|_{L^2(L^2)}$ are bounded independently of β, and h. Thus (4.43) becomes

$$\begin{aligned}
&\max_{0\le n\le N}\|P_h(S_\beta^{n}-S_h^{n})\|^2_{H_h^{-1}}+\eta\sum_0^{N-1}\Delta t\big(K_\beta(S_\beta^{n+1})-K_\beta(S_h^{n+1}),S_\beta^{n+1}-S_h^{n+1}\big)\\
&\qquad\le C\{h^{2\gamma}m(\beta)^{-\frac{1}{1+\mu}}+(\Delta t)^{\frac{\gamma+2}{2}}\}
\end{aligned}\tag{4.45}$$

An immediate consequence of (4.45) is the following.

$$\sum_{0}^{N-1} \|K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1})\|_{L^2}^2 \Delta t \;\le\; C\{h^{2\gamma} m(\beta)^{-\frac{1}{1+\mu}} + (\Delta t)^{\frac{\gamma+2}{2}}\} \tag{4.46}$$

and

$$\sum_{0}^{N-1} \Delta t \|S_\beta^{n+1} - S_h^{n+1}\|_{L^{2+\mu}}^{2+\mu} \;\le\; C\{h^{2\gamma} m(\beta)^{-\frac{1}{1+\mu}} + (\Delta t)^{\frac{\gamma+2}{2}}\}. \tag{4.47}$$

To obtain the desired estimate for $S_\beta^n - S_h^n$, we will use (4.45) together with an estimate for $(I-P_h)(S_\beta^n - S_h^n)$. This estimate however requires the following Lemma.

Lemma 4.1 *For K_β and S_h^n as in Theorem 4.1, there is a positive constant C so that*

$$\max_{0\le n\le N} \|\nabla K_\beta(S_h^n)\|_{L^2} \le C. \tag{4.48}$$

Proof. In (4.14) set $\chi = K_\beta(S_h^{n+1}) - K_\beta(S_h^n) \in M_h$ to get

$$\begin{aligned}\left(\frac{S_h^{n+1} - S_h^n}{\Delta t}, K_\beta(S_h^{n+1}) - K_\beta(S_h^n)\right) - \left(f(S_h^{n+1})\mathbf{u}^{n+1}, \nabla(K_\beta(S_h^{n+1}) - K_\beta(S_h^n))\right) \\ + \left(\nabla K_\beta(S_h^{n+1}), \nabla(K_\beta(S_h^{n+1}) - K_\beta(S_h^n))\right) = 0\end{aligned} \tag{4.49}$$

This yields (Cauchy–Schwartz)

$$\begin{aligned}\left(\frac{S_h^{n+1} - S_h^n}{\Delta t}, K_\beta(S_h^{n+1}) - K_\beta(S_h^n)\right) + \frac{1}{2}\|\nabla K_\beta(S_h^{n+1})\|_{L^2}^2 - \frac{1}{2}\|\nabla K_\beta(S_h^n)\|_{L^2}^2 \\ \le \left(f(S_h^{n+1})\mathbf{u}^{n+1}, \nabla\big(K_\beta(S_h^{n+1}) - K_\beta(S_h^n)\big)\right)\end{aligned} \tag{4.50}$$

The righthand side of (4.50) can be rewritten as follow

$$\begin{aligned}&(f(S_h^{n+1})\mathbf{u}^{n+1}, \nabla(K_\beta(S_h^{n+1}) - K_\beta(S_h^n))) \\ &\quad= (f(S_h^{n+1})\mathbf{u}^{n+1}, \nabla K_\beta(S_h^{n+1})) - (f(S_h^n)\mathbf{u}^n, \nabla K_\beta(S_h^n)) \\ &\qquad +(f(S_h^n)\mathbf{u}^n - f(S_h^{n+1})\mathbf{u}^{n+1}, \nabla K_\beta(S_h^n)) \\ &\quad\le (f(S_h^{n+1})\mathbf{u}^{n+1}, \nabla K_\beta(S_h^{n+1})) - (f(S_h^n)\mathbf{u}^n, \nabla K_\beta(S_h^n)) \\ &\qquad +\frac{\tilde{C}}{2\Delta t}\|f(S_h^n)\mathbf{u}^n - f(S_h^{n+1})\mathbf{u}^n\|_{L^2}^2 + C\Delta t\|\nabla K_\beta(S_h^n)\|_{L^2}^2 \\ &\qquad +\frac{\tilde{C}}{2\Delta t}\|f(S_h^{n+1})(\mathbf{u}^{n+1} - \mathbf{u}^n)\|_{L^2}^2\end{aligned} \tag{4.51}$$

The last term on the righthand side of (4.51) is bounded by

$$\frac{1}{2}\tilde{C}\Delta t\|\mathbf{u}_t\|_{L^\infty}^2 \|f(S_h^{n+1})\|_{L^\infty}^2. \tag{4.52}$$

Now hide the second term of the right side of (4.51) in the left hand side of (4.50) by inequality (2.2), by making $\tilde{C}$ sufficiently small. Combine (4.50) and (4.51), sum over $0 \le n \le m$, and use the discrete Grönwall Lemma, to complete the proof of the Lemma. □

In order to complete the proof of Theorem 4.1, we use the inequality $\max_{0\le n\le N} \|\nabla K_\beta(S_\beta^n)\|_{L^2} \le C$ established by Lemma 3.1, and follow the analysis done in (3.44) through (3.57), in order to get

$$\max_{0\le n\le N-1} \|(I-P_h)(S_\beta^{n+1} - S_h^{n+1})\|_{(H^1)^*} \le Ch^\gamma \tag{4.53}$$

The proof of the theorem is completed by assuming the inverse estimates (3.2), and using estimates (3.43), (4.45), and (4.53). □

If, as in section 3, we take $\beta = \beta_0 h^\lambda$ with $\lambda = \frac{4+2\mu}{2+4\mu+\mu^2}$, and if we consider the specific perturbation defined by (2.17), then the conclusion of Theorem 4.1 becomes

$$\max_{0\le n\le N-1} \|S_\beta^{n+1} - S_h^{n+1}\|^2_{(H^1)^*} + \sum_0^{N-1} \Delta t (K_\beta(S_\beta^{n+1}) - K_\beta(S_h^{n+1}), S_\beta^{n+1} - S_h^{n+1}) \le C\{h^{(2+\mu)\lambda} + (\Delta t)^{\frac{\gamma+2}{2}}\} \quad (4.54)$$

Finally we have the following

Corollary 4.1 *Under the hypotheses of Theorem 4.1 we have*

$$\max_{0\le n\le N} \|S(t_n) - S_h^n\|^2_{(H^1)^*} \;\le\; C\{h^{2\gamma} m(\beta)^{-\frac{1}{1+\mu}} + C(\beta) + (\Delta t)^{\frac{\gamma+2}{2}}\} \quad (4.55)$$

and

$$\sum_0^{N-1} \Delta t \|S(t_{n+1}) - S_h^{n+1}\|^{2+\mu}_{L^{2+\mu}} \;\le\; C\{h^{2\gamma} m(\beta)^{-\frac{1}{1+\mu}} + C(\beta) + (\Delta t)^{\frac{\gamma+2}{2}}\} \quad (4.56)$$

where $C(\beta)$ is defined by (1.10), and $m(\beta)$ defined by (2.12).

Estimate (4.56) does not require the inverse estimate assumption (3.2). But we do need this assumption for estimate (4.55). Estimate (4.56) is a direct consequence of Theorem 2.1 and estimate (4.47). Estimate (4.55) is a direct consequence of Theorem 2.1 and Theorem 4.1.

A Appendix on Poisson Solutions: Regularity and Approximation

In [11], properties of the Poisson Solution Operator T were given which were needed in our development and are summarized here for convenience. In addition, we define the Mean-Value Preserving Elliptic Projection onto approximating subspaces, the corresponding discrete version of the operator T, and give some of their properties which are required for our analysis.

A.1 The Poisson Solution Operator

The elliptic boundary value problem

$$\begin{cases} -\Delta\omega &= f, \quad \text{in } \Omega \\ \frac{\partial\omega}{\partial n} &= 0, \quad \text{on } \partial\Omega \\ \omega_\Omega &= 0, \end{cases} \quad (A.1)$$

has a unique solution $\omega =: Sf$ with $\omega \in H^1$ when $f \in H^{-1}$ and f_Ω vanishes (see Sections 5.2-5.3 of [2]). Therefore for any $f \in H^{-1}$ the boundary value problem

$$\begin{cases} -\Delta u &= f - f_\Omega, \quad \text{in } \Omega \\ \frac{\partial u}{\partial n} &= 0, \quad \text{on } \partial\Omega \\ u_\Omega &= f_\Omega, \end{cases} \quad (A.2)$$

has a unique solution $u \in H^1$ given by $u := S(f - f_\Omega) + f_\Omega$, and we define the Mean-Value Preserving Elliptic Solution operator $T : (H^1)^* \to H^1$ by $T(f) = u$. A more convenient equivalent norm for H^1 is defined by (A.9) below and is closely related to T. The weak formulation of (A.2) is given by

$$\begin{aligned} (\nabla u, \nabla\phi) &= (f, \phi) - (f_\Omega, \phi) \\ &= (f, \phi) - f_\Omega \phi_\Omega \\ &= (f, \phi) - (Tf)_\Omega \phi_\Omega \end{aligned} \quad (A.3)$$

for all $\phi \in H^1$ and so

$$\Big(\nabla(Tf), \nabla\phi\Big) = (f, \phi) - f_\Omega \phi_\Omega. \tag{A.4}$$

In particular, if we take $\phi = Tf$, then we obtain

$$\|\nabla Tf\|_{L^2}^2 = (f, Tf) - (f_\Omega)^2 = (f, Tf) - (Tf)_\Omega^2. \tag{A.5}$$

The operator T is linear, symmetric and positive definite [18, 11]

$$(Tf,\, g) = (f,\, Tg) \text{ for all } f, g \in (H^1)^*. \tag{A.6}$$

and from (A.5) it follows that

$$(Tf,\, f) = \|\nabla Tf\|_{L^2}^2 + (Tf)_\Omega^2 = \|\nabla Tf\|_{L^2}^2 + (f_\Omega)^2. \tag{A.7}$$

With these properties in mind we can define on $(H^1)^*$ the norm:

$$|||f|||_{(H^1)^*} := (Tf, f)^{\frac{1}{2}} \tag{A.8}$$

which is the dual norm for H^1, when H^1 is equipped with the equivalent norm

$$|||u|||_{H^1} := (\|\nabla u\|_{L^2}^2 + (u_\Omega)^2)^{\frac{1}{2}}. \tag{A.9}$$

With these definitions we get the relationship

$$(Tf, f) = |||Tf|||_{H^1}^2 \simeq \|Tf\|_{H^1}^2. \tag{A.10}$$

where $\simeq$ means equivalent within fixed constants independent of f. The proof of the equivalence is a simple application of Poincaré's inequality (see e.g [7]) in one direction and Hölder inequality in the other.

Proposition A.1 *Suppose f belongs to $(H^1)^*$, then*

$$(Tf, f)^{\frac{1}{2}} = \|f\|_{(H^1)^*}, \tag{A.11}$$

in the sense of the norm (A.9).

The results of Chapter 4 of [12] further extend the properties of the operator T to more general Sobolev spaces over convex polygonal domains in R^2. In particular, elliptic *a priori* estimates of the form

$$\|u\|_{W^{2,p}(\Omega)} \leq c_\Omega \{\|\Delta u\|_{L^p} + \|u\|_{L^p}\}, \tag{A.12}$$

are established (see, inequality (4,1,2) of [12] and its proof using Theorem 4.3.2.4 and Remark 4.3.2.5), under the assumption that $1 < p < \overline{p}(\Omega)$, where

$$\overline{p}(\Omega) := \left(1 - \frac{\pi}{2\,\Theta(\Omega)}\right)^{-1} \tag{A.13}$$

and $\Theta(\Omega)$ is defined as the maximal interior angle of the polygon Ω. Hence in this range of p it follows that

$$\|Tf\|_{W^{2,p}(\Omega)} \leq c_\Omega \{\|f - f_\Omega\|_{L^p} + \|Tf\|_{L^p}\} \leq c\|f\|_{L^p}. \tag{A.14}$$

The last inequality on the right hand side of inequality (A.14) (i.e. the boundedness of T on L^p) follows by the Sobolev embedding theorem (since $d \leq 2$), the fact that the result holds in the case $p = 2$, and Hölder's inequality.

From [11], we also have the following result:

Proposition A.2 *For smooth f, the operators T and $\frac{\partial}{\partial t}$ commute*

$$\frac{\partial}{\partial t}(Tf) = T(\frac{\partial f}{\partial t}) \tag{A.15}$$

Furthermore, if $f \in H^1(\Omega)$ and

$$\frac{\partial f}{\partial n} = 0 \text{ on } \partial\Omega, \tag{A.16}$$

then

$$T(\Delta f) = \Delta(Tf). \tag{A.17}$$

A.2 The Mean-Value Preserving Projection

Let $\{M_h\}_{h>0}$ be a family of finite dimensional spaces (see, for example, Chapter 4 of [2]) such that $M_h \subset H^1(\Omega)$. Let $f \in H^1$, then

$$\begin{cases} (\nabla f_h, \nabla\chi) &= (\nabla f, \nabla\chi), \quad \forall\chi \in M_h \\ (f_h)_\Omega &= f_\Omega \end{cases} \tag{A.18}$$

has unique solution in M_h. We define the mean–value preserving operator $E_h : H^1(\Omega) \longrightarrow M_h$ by $E_h(f) := f_h$ where f_h is the unique solution to (A.18). By the definition of the projection E_h and orthogonality it follows that

$$0 \leq \|\nabla(f - E_h f)\|_{L^2}^2 + \|\nabla E_h f\|_{L^2}^2 = \|\nabla f\|_{L^2}^2$$

and so for $f \in H^1(\Omega)$, there holds

$$\|\nabla E_h f\|_{L^2} \leq \|\nabla f\|_{L^2}. \tag{A.19}$$

$$\|E_h f\|_{H^1} \leq \|f\|_{H^1}. \tag{A.20}$$

In fact, by using a duality argument (see Section 7.5 of [2]), the projection E_h can be shown to be bounded on $W^{1,p}(\Omega)$ for all $1 < p < \infty$, and therefore a corresponding Cea estimate holds for the elliptic projection. This and another lifting argument then provides a mean-preserving second order elliptic error estimate for E_h of the form

$$\|f - E_h f\|_{L^p} \leq ch^2\|f\|_{W^{2,p}(\Omega)} \tag{A.21}$$

if $1 < p' < \overline{p}(\Omega)$, where $\overline{p}$ is defined as in (A.13) and p' is the conjugate index to p.

A.3 The discrete analogue of the Solution Operator

We consider the discretized elliptic problem of finding for each $w_h \in M_h$ a solution $f \in H^1(\Omega)^*$ such that

$$\begin{cases} (\nabla w_h, \nabla\chi) &= (f - f_\Omega, \chi), \quad \forall\chi \in M_h \\ (w_h)_\Omega &= (f)_\Omega. \end{cases} \tag{A.22}$$

By the definition of the solution operator T to (A.2) there holds

$$\begin{cases} (\nabla Tf, \nabla\chi) &= (f - f_\Omega, \chi) \quad \forall\chi \in M_h \\ (Tf)_\Omega &= (f)_\Omega, \end{cases} \tag{A.23}$$

and so

$$T_h := E_h \circ T : H^1(\Omega)^* \to M_h \tag{A.24}$$

is the solution operator to the discrete problem (A.22), that is $w_h = E_h(Tf)$. It follows directly [18, 11] that the operator T_h is linear, symmetric in the sense

$$(T_h f,\, g) = (f,\, T_h g) \text{ for all } f, g \in (H^1)^*, \tag{A.25}$$

and is nonnegative since

$$(T_h f, f) = \|\nabla T_h f\|_{L^2}^2 + (f_\Omega)^2 \geq 0. \tag{A.26}$$

Although $\chi \to (T_h\chi, \chi)^{\frac{1}{2}}$ is only a semi–norm on $(H^1)^*$, it is a norm when restricted to M_h,

$$\|\chi\|_{H_h^{-1}} = (T_h\chi, \chi)^{\frac{1}{2}} = \left(\|\nabla T_h\chi\|_{L^2}^2 + (\chi_\Omega)^2\right)^{\frac{1}{2}}. \tag{A.27}$$

In fact, Lemma 4.4 of [17] established the following:

Lemma A.1 *If $\|\chi\|_{(H^1)^*}$ is defined by (A.11), then there is a positive constant C so that*

$$C\,\|\chi\|_{(H^1)^*} \leq \|\chi\|_{H_h^{-1}} \leq \|\chi\|_{(H^1)^*} \qquad \forall\chi \in M_h. \tag{A.28}$$

By combining the elliptic error estimates (A.21) with the elliptic regularity of T (A.12) the following lemma follows immediately.

Lemma A.2 *If the discrete solution operator T_h is defined by (A.24), then*

$$\|(T - T_h)(f)\|_{L^p} \leq ch^2 \|f\|_{L^p} \tag{A.29}$$

if $\max(p, p') < \overline{p}(\Omega)$, *where* $\overline{p}(\Omega)$ *is defined as in (A.13).*

References

[1] J. Bear and A. Verruijt. *Modeling Groundwater Flow and Pollution.* D. Reidel Publication Company, Dodreich, Holland, 1987.

[2] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods.* Texts in Applied Mathematics 15, Springer, Berlin, 1994.

[3] E. F. Browder. Nonlinear operators and nonlinear evolution equations in banach spaces. In *Proc. Pure Math.*, volume 18, Providence, Rhode Island, 1976. Amer. Math. Society.

[4] G. Chavent and J. Jaffre. *Mathematical Models and Finite Element for Reservoir Simulation: Single phase, multiphase and multicomponent flows through Porous Media.* North-Holland, New York, 1986.

[5] P. G. Ciarlet. *The Finite Element Methods for elliptic Problems.* North-Holland, 1978.

[6] H. K. Dahle. Adaptive characteristic operator splitting technique for convection dominated diffusion problems in one and two dimensions. Technical report, University of Bergen, Bergen, Norway, 1988.

[7] R. Dautray and J-L Lions. *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 2. Springer-Verlag, Berlin, 1988.

[8] T. Dupont and R. Scott. Polynomial approximation of functions in sobolev spaces. *Math. Comp.*, 34:441–463, 1980.

[9] R. E. Ewing. Problems arising in the modeling of processes for hydrocarbon recovery. In R. E. Ewing, editor, *The Mathematics of Reservoir Simulation*, pages 3–34, Philadelphia, 1983. S.I.A.M.

[10] K. B. Fadimba. *Regularization and Numerical Methods for a Class of Porous Medium Equations.* PhD thesis, University of South Carolina, Columbia, 1993.

[11] K. B. Fadimba and R. C. Sharpley. A priori estimates and regularization for a class of porous medium equations. *Nonlin. World*, 2:13–41, 1995.

[12] P. Grisvard. *Elliptic Problems in Nonsmooth Domains.* Monographs and Studies in Mathematics, Pitman, Boston, 1985.

[13] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method.* Cambridge University Press, Cambridge, 1987.

[14] R. H. Nochetto and C Verdi. Approximation of degenerate parabolic problems using numerical integration. *SIAM J. Numer. Anal.*, 25(4):784–814, 1988.

[15] D.W. Peaceman. *Fundamentals of Numerical Reservoir Simulation.* Elsevier Scientific Publishing Company, New York, 1977.

[16] M. E. Rose. Numerical methods for a general class of porous medium equations. Technical report, Argonne National Laboratory, Argonne, Il, 1980.

[17] M. E. Rose. Numerical methods for flow through porous media–I. *Math. Comp.*, 40:435–467, 1983.

[18] D. L. Smylie. *A Near Optimal Order Approximation to a Class of Two-sided Nonlinear Parabolic Partial differential Equations.* PhD thesis, University of Wyoming, Laramie, 1989.

[19] H. Treibel. *The Structure of Functions.* Birkhauser, Basel, 2001.