FAST COMMUNICATION

CRANK-NICOLSON FINITE ELEMENT METHODS USING SYMMETRIC STABILIZATION WITH AN APPLICATION TO OPTIMAL CONTROL PROBLEMS SUBJECT TO TRANSIENT ADVECTION–DIFFUSION EQUATIONS[*]

ERIK BURMAN[†]

Abstract. We consider a finite element method with symmetric stabilization for transient advection-diffusion-reaction problems. The Crank-Nicolson finite difference scheme is used for discretization in time. We prove stability of the numerical method both for implicit and explicit treatment of the stabilization operator. The resulting convergence results are given and the results are illustrated by a numerical experiment. We then consider a model problem for pde-constrained optimization. Using discrete adjoint consistency of our stabilized method we show that both the implicit and semi-implicit methods proposed yield optimal convergence for the control and the state variable.

Key words. Transient advection–diffusion, stabilized finite element methods, Crank-Nicolson, optimal control.

AMS subject classifications. 65M12, 65M38, 49M25, 49M29.

1. Introduction

In a previous work [4] we studied finite element methods with symmetric stabilization for transient advection-diffusion-reaction problems. For the time-discretization we proposed to use either the θ-method or the second order backward differentiation formula. We also proposed to evaluate the stabilization operator at the previous time-step, resulting in a semi-implicit scheme, in order to reduce computational cost. The procedure was proven to be stable and optimally convergent in cases where the time-discretization produced some numerical dissipation. This excluded the Crank-Nicolson scheme and we claimed that the explicit treatment was not viable for this method. Herein we will show that this claim was premature. Indeed it turns out that the same stability/convergence analysis that we performed in [4] is valid for the Crank-Nicolson method with explicit treatment of the stabilization term. This is somewhat surprising since the method has no dissipation that can counter the destabilizing effect of shifting the stabilization. Symmetric stabilization methods have some advantages compared to Petrov-Galerkin type stabilized methods; in particular the resulting adjoint consistency is appealing for the design of methods for optimal control problems and a posteriori error estimates. In the second part of this note we exploit the symmetry property to show how the numerical scheme may be applied for the solution of transient optimal control problems or inverse problems, provided that the stabilization is treated in the same fashion for both the state and the adjoint equation. Note that in this case, where repeated solves are necessary to find the optimal solution, the possibility of treating the stabilization explicitly is a definite advantage.

[†]Department of Mathematics, University of Sussex, Brighton, BN1 9RF United Kingdom, (E.N.Burman@sussex.ac.uk).

2. Problem setting

Let Ω be an open domain in $\mathbb{R}^d$ ($d=1,2$ or 3), with a polyhedral boundary $\partial\Omega$, and let $I=]0,T[$ with $T>0$. We consider the following problem: find $u:\Omega\times I\mapsto\mathbb{R}$ such that

$$\begin{cases} \partial_t u + \boldsymbol{\beta}\cdot\nabla u + \sigma u - \mu\Delta u = f & \text{in} \quad \Omega\times I, \\ \qquad\qquad\qquad\qquad u = 0 & \text{on} \quad \partial\Omega\times I, \\ \qquad\qquad\quad u(\cdot,0) = u_0 & \text{in} \quad \Omega, \end{cases} \tag{2.1}$$

where $\boldsymbol{\beta}:\Omega\mapsto\mathbb{R}^d$ is a given, Lipschitz continuous, velocity field satisfying $\nabla\cdot\boldsymbol{\beta}=0$, $f\in L^2(\Omega)$ is a source function, $u_0\in L^2(\Omega)$ the initial data and $\sigma\geq 0$, $\mu\geq\mu_0>0$ are given bounded functions. Formally the case $\mu=0$ (hyperbolic regime) is also covered by the theory below, however the boundary conditions (2.1) and the variational framework must be modified accordingly. In that case we impose $u=0$ on the inflow boundary $\Gamma_{\text{in}}\overset{\text{def}}{=}\{\boldsymbol{x}\in\partial\Omega:(\boldsymbol{\beta}\cdot\boldsymbol{n})(\boldsymbol{x})<0\}$ only, where $\boldsymbol{n}$ stands for the outward unit normal of the domain Ω.

Let us introduce some standard notation. For a given domain $\omega\subset\mathbb{R}^d$, the space of functions whose distributional derivatives of order up to $m\geq 0$ belong to $L^2(\omega)$ is denoted by $H^m(\omega)$. The subspace of $H^1(\omega)$ consisting of functions vanishing on the boundary is denoted by $H^1_0(\omega)$. The norm of $H^m(\omega)$ is denoted by $\|\cdot\|_{m,\omega}$. The L^2-norm is denoted by $\|\cdot\|_{0,\omega}$ and its inner product by $(\cdot,\cdot)_\omega$, the latter being simplified in the case $\omega=\Omega$ as $(\cdot,\cdot)\overset{\text{def}}{=}(\cdot,\cdot)_\Omega$. We will use $\|\cdot\|_\infty$ to denote the $L^\infty(I;L^\infty(\Omega))$ norm and $\|\cdot\|_{1,\infty}$ to denote the $L^\infty(I;W^{1,\infty}(\Omega))$ norm.

In order to introduce a variational setting for (2.1) we consider the space $V\overset{\text{def}}{=}H^1_0(\Omega)$ and we define the bilinear form $a(u,v)=(\boldsymbol{\beta}\cdot\nabla u+\sigma u,v)+(\mu\nabla u,\nabla v)$, $\forall u,v\in V$. The above problem can then be cast in weak form as follows:

$$\begin{cases} \text{For a.e. } t\in I, \text{ find } u(t)\in V \text{ such that} \\ (\partial_t u, v) + a(u,v) = (f,v) \quad \forall v\in V, \\ u(0) = u_0. \end{cases} \tag{2.2}$$

3. Space discretization

Let $\{\mathcal{T}_h\}_{0<h\leq 1}$ denote a family of shape regular and quasi-uniform triangulations of the domain Ω. The elements of $\mathcal{T}_h$ will be denoted by K. To each element we associate an outward unit normal $\boldsymbol{n}_K$. The faces of the triangulation will be denoted F, $\mathcal{F}(K)$ stands for the set of faces in the boundary of $K\in\mathcal{T}_h$ and $\mathcal{F}_{\text{in}}(K)$ for the set of inner faces in the boundary of K, i.e. such that $\text{int}(F)\cap\partial\Omega=\emptyset$.

For each triangulation $\mathcal{T}_h$, the subscript $h\in(0,1]$ refers to the level of refinement of the triangulation, which is defined by $h\overset{\text{def}}{=}\max_{K\in\mathcal{T}_h}h_K$, $\quad h_K\overset{\text{def}}{=}\max_{F\subset\mathcal{F}(K)}h_F$, with h_F the diameter of the face F. We let V_h denote the standard finite element space of continuous, piecewise polynomial functions of degree $k\geq 1$, defined on $\mathcal{T}_h$. Dirichlet boundary conditions will be enforced in a weak sense, using Nitsche's method [8] (see also [9, Page 24]). Note that this weak treatment of the boundary conditions allows us to include both the hyperbolic and the parabolic problems in the same framework.

Then the stabilized finite element approximation of (2.2) is defined as follows:

$$\begin{cases} \text{For } t\in I, \text{ find } u_h(t)\in V_h \text{ such that} \\ (\partial_t u_h, v_h) + a_h(u_h,v_h) + s_h(u_h,v_h) = (f,v_h) \quad \forall v_h\in V_h, \\ (u_h(0),v_h) = (u_0,v_h) \quad \forall v_h\in V_h, \end{cases} \tag{3.1}$$

where

$$a_h(u_h, v_h) \overset{\text{def}}{=} a(u_h, v_h) - (\mu \nabla u_h \cdot \boldsymbol{n}, v_h)_{\partial \Omega} - (u_h, \mu \nabla v_h \cdot \boldsymbol{n})_{\partial \Omega}$$
$$+ (\gamma_{\text{bc}} \mu h^{-1} u_h, v_h)_{\partial \Omega} - (\boldsymbol{\beta} \cdot \boldsymbol{n} u_h, v_h)_{\Gamma_{\text{in}}}. \tag{3.2}$$

Here, $\gamma_{\text{bc}} > 0$ is Nitsche's penalty parameter. Moreover, $s_h(\cdot, \cdot)$ represents an abstract stabilization term — a symmetric, positive semi-definite, bilinear form. The matrix pattern corresponding to $s_h(\cdot, \cdot)$ does not necessarily coincide with that of the standard Galerkin method.

For the convergence analysis below, we introduce the following norms

$$\|v_h\|^2 \overset{\text{def}}{=} \|\sigma^{\frac{1}{2}} v_h\|_{0,\Omega}^2 + \frac{1}{2} \|\mu^{\frac{1}{2}} \nabla v_h\|_{0,\Omega}^2 + \frac{1}{2} \|(\gamma_{\text{bc}} \mu / h)^{\frac{1}{2}} v_h\|_{0,\partial\Omega}^2 + \frac{1}{2} \||\boldsymbol{\beta} \cdot \boldsymbol{n}|^{\frac{1}{2}} v_h\|_{0,\partial\Omega}^2$$

and $\|\|v_h\|\|^2 \overset{\text{def}}{=} \|v_h\|^2 + s_h(v_h, v_h)$ for all $v_h \in V_h$. Note that for $\sigma = \mu = 0$ this is only a semi-norm. It is well known that for γ_{bc} large enough there exists $c_a > 0$ such that for all $u_h \in V_h$

$$c_a \|u_h\|^2 \leq a_h(u_h, u_h) \text{ and } c_a \|\|u_h\|\|^2 \leq a_h(u_h, u_h) + s_h(u_h, u_h). \tag{3.3}$$

In what follows, the symbol $\lesssim$ indicates an inequality up to a multiplicative constant independent of the discretization and physical parameters. Note, in particular, that the constant is independent of T.

3.1. Symmetric stabilization methods. Herein we only assume that the stabilization operator is symmetric (i.e. that $s_h(u_h, v_h) = s_h(v_h, u_h)$) and that the following inequalities hold:

$$0 \leq s_h(u_h, u_h) \leq C_S h^{-1} \|u_h\|_{0,\Omega}^2, \quad \forall u_h \in V_h. \tag{3.4}$$

An example of a stabilization entering the framework is the continuous interior penalty method (CIP), obtained by the addition of the following penalty term on the normal derivative jump over element faces:

$$s_h(u_h, v_h) \overset{\text{def}}{=} \gamma \sum_{K \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_{\text{in}}(K)} \int_F h_F^2 |\boldsymbol{\beta}_h \cdot \boldsymbol{n}_K| [\![\partial_{\boldsymbol{n}_K} v_h]\!][\![\partial_{\boldsymbol{n}_K} v_h]\!], \tag{3.5}$$

with $\gamma > 0$ the stabilization parameter and where $\boldsymbol{\beta}_h$ is some interpolant of $\boldsymbol{\beta}$ in the space of piecewise affine functions. Moreover $[\![\partial_{\boldsymbol{n}_K} v_h]\!] \overset{\text{def}}{=} \partial_{\boldsymbol{n}_K} v_h^- - \partial_{\boldsymbol{n}_K} v_h^+$, where

$$\partial_{\boldsymbol{n}_K} v_h^\pm(\boldsymbol{x}) \overset{\text{def}}{=} \boldsymbol{n}_K(\boldsymbol{x}) \cdot \lim_{\epsilon \to 0^\pm} \nabla v_h(\boldsymbol{x} + \epsilon \boldsymbol{n}_K) \quad \forall \boldsymbol{x} \in \bigcup_{F \in \mathcal{F}_{\text{in}}(K)} \text{int}(F)$$

denotes the jump of the normal derivative across the boundary of the simplex $K \in \mathcal{T}_h$. For work on other symmetric stabilization methods for transient transport problems we refer to [7, 5].

4. Time discretization

Let $N > 0$ be a given positive integer. In what follows, we consider a uniform partition of I with time step size $\tau \overset{\text{def}}{=} T/N$. In addition, $u_h^n \in V_h$ stands for an approximation of $u(t^n)$ in V_h, with $t^n \overset{\text{def}}{=} n\tau$ and $0 \leq n \leq N$.

We propose a Crank-Nicolson scheme where the stabilization term is evaluated either at the midpoint of each time interval in the standard way or in a completely implicit or completely explicit fashion; indeed we introduce a parameter λ and apply the stabilization operator to the solution $\lambda u_h^{n+1} + (1-\lambda)u_h^n$, with $\lambda \in \{0, 1/2, 1\}$. Note that for $\lambda = 0$ the treatment of the stabilization term is fully explicit. For $\lambda \in \{0, 1/2, 1\}$ the time discretization of (2.2) reads as follows:

$$\begin{cases} \text{For } 0 \leq n \leq N-1, \text{ find } u_h^{n+1} \in V_h \text{ such that} \\ (D_\tau u_h^{n+1}, v_h) + a_h(u_h^{n+\frac{1}{2}}, v_h) + s_h(u_h^{n+\lambda}, v_h) = (f^{n+\frac{1}{2}}, v_h) \quad \forall v_h \in V_h, \\ u_h^0 = \pi_h u_0, \end{cases} \quad (4.1)$$

with the notations $D_\tau u_h^{n+1} \stackrel{\text{def}}{=} (u_h^{n+1} - u_h^n)/\tau$, $u_h^{n+\frac{1}{2}} \stackrel{\text{def}}{=} \frac{1}{2}(u_h^{n+1} + u_h^n)$, $f^{n+\frac{1}{2}} \stackrel{\text{def}}{=} f(t^{n+\frac{1}{2}})$, and $u_h^{n+\lambda} = \lambda u_h^{n+1} + (1-\lambda)u_h^n$.

At each time-level $0 \leq n \leq N-1$ the problem (4.1) involves the solution of a linear system with a non-standard Galerkin matrix pattern, except in the case $\lambda = 0$ when the stabilization term is fully explicit.

4.1. Stability.   We first prove an elementary lemma on the properties of the stabilization operator.

LEMMA 4.1. *For* $\lambda \in \{0, \frac{1}{2}, 1\}$ *there holds*

$$\tau \sum_{n=0}^{N-1} s_h(u_h^{n+\lambda}, u_h^{n+\frac{1}{2}}) = \tau \frac{1}{2}\left(\lambda - \frac{1}{2}\right) s_h(u_h^N, u_h^N) + \tau \sum_{n=0}^{N-1} s_h(u_h^{n+\frac{1}{2}}, u_h^{n+\frac{1}{2}})$$
$$- \tau \frac{1}{2}\left(\lambda - \frac{1}{2}\right) s_h(u_h^0, u_h^0). \quad (4.2)$$

Proof.   For $\lambda = \frac{1}{2}$ the statement is trivial. If $\lambda = 0$ we have $s_h(u_h^n, u_h^{n+\frac{1}{2}}) = \frac{1}{2}s_h(u_n^n, u_h^{n+1}) + \frac{1}{2}s_h(u_n^n, u_h^n)$. By the symmetry of $s_h(\cdot, \cdot)$ it follows that $s_h(u_h^n, u_h^{n+\frac{1}{2}}) = s_h(u_h^{n+\frac{1}{2}}, u_h^{n+\frac{1}{2}}) - \frac{1}{4}s_h(u_h^{n+1}, u_h^{n+1}) + \frac{1}{4}s_h(u_h^n, u_h^n)$ and we conclude by summing over $0 \leq n \leq N-1$. The case $\lambda = 1$ is similar.   □

The following result states the stability properties of the fully discrete scheme (4.1).

LEMMA 4.2 (L^2-*stability, Crank-Nicolson*). *Let* $\{u_h^n\}$ *be the solution of* (4.1). *Then*

- *for* $\lambda \in \{\frac{1}{2}, 1\}$ *there holds*

$$\frac{1}{2}\|u_h^N\|_{0,\Omega}^2 + \tau \frac{1}{2}\left(\lambda - \frac{1}{2}\right) s_h(u_h^N, u_h^N) + \tau c_a \sum_{n=0}^{N-1} \||u_h^{n+\frac{1}{2}}\||^2$$
$$\leq \tau \sum_{n=0}^{N-1} (f^{n+\frac{1}{2}}, u_h^{n+\frac{1}{2}}) + \frac{1}{2}\|u_h^0\|_{0,\Omega}^2 + \tau\left(\lambda - \frac{1}{2}\right) s_h(u_h^0, u_h^0),$$
$$(4.3)$$

- *for* $\lambda = 0$ *and under the CFL-condition* $\tau \leq (C_S)^{-1}h$ *there holds*

$$\frac{1}{4}\|u_h^N\|_{0,\Omega}^2 + \tau c_a \sum_{n=0}^{N-1} \||u_h^{n+\frac{1}{2}}\||^2 \leq \tau \sum_{n=0}^{N-1} (f^{n+\frac{1}{2}}, u_h^{n+\frac{1}{2}}) + \frac{1}{2}\|u_h^0\|_{0,\Omega}^2. \quad (4.4)$$

Proof. Take $v_h = u_h^{n+\frac{1}{2}}$ in (4.1) and multiply by τ. By definition there holds

$$\tau(D_\tau u_h^{n+1}, u_h^{n+\frac{1}{2}}) = \frac{1}{2}\|u_h^{n+1}\|_{0,\Omega}^2 - \frac{1}{2}\|u_h^n\|_{0,\Omega}^2$$

and by (3.3) we have $c_a\|u_h^{n+\frac{1}{2}}\|^2 \leq a_h(u_h^{n+\frac{1}{2}}, u_h^{n+\frac{1}{2}})$. Summing over n, using Lemma 4.1 and the right-hand inequality of (3.3) we obtain

$$\frac{1}{2}\|u_h^N\|_{0,\Omega}^2 + \tau\frac{1}{2}\left(\lambda - \frac{1}{2}\right)s_h(u_h^N, u_h^N) + c_a\tau\sum_{n=0}^{N-1}\left\|\|u_h^{n+\frac{1}{2}}\|\right\|^2$$
$$\leq \tau\sum_{n=0}^{N-1}(f^{n+\frac{1}{2}}, u_h^{n+\frac{1}{2}}) + \frac{1}{2}\|u_h^0\|_{0,\Omega}^2 + \tau\frac{1}{2}\left(\lambda - \frac{1}{2}\right)s_h(u_h^0, u_h^0). \qquad (4.5)$$

Since both the second term in the left hand side and the last term in the right hand side are non-negative for $\lambda \geq \frac{1}{2}$, this proves the case $\lambda \in \{\frac{1}{2}, 1\}$.

For $\lambda = 0$ note that the last term in the right hand side is negative and can therefore be omitted. Then note that under the CFL-condition there holds

$$\frac{1}{2}\|u_h^N\|_{0,\Omega}^2 - \tau\frac{1}{4}s_h(u_h^N, u_h^N) \geq \frac{1}{2}\|u_h^N\|_{0,\Omega}^2 - \tau\frac{1}{4}C_S h^{-1}\|u_h^N\|_{0,\Omega}^2 \geq \frac{1}{4}\|u_h^N\|_{0,\Omega}^2.$$

$\square$

COROLLARY 4.3. *Let $\{u_h^n\}$ be the solution of (4.1). Then*
- *for $\lambda \in \{\frac{1}{2}, 1\}$ there holds*

$$\|u_h^N\|_{0,\Omega}^2 + \tau\left(\lambda - \frac{1}{2}\right)s_h(u_h^N, u_h^N) + \tau c_a\sum_{n=0}^{N-1}\left\|\|u_h^{n+\frac{1}{2}}\|\right\|^2$$
$$\lesssim T\tau\sum_{n=0}^{N-1}\|f^{n+\frac{1}{2}}\|_{0,\Omega}^2 + \frac{1}{2}\tau\left(\lambda - \frac{1}{2}\right)s_h(u_h^0, u_h^0) + \|u_h^0\|_{0,\Omega}^2, \quad (4.6)$$

- *for $\lambda = 0$ and under the CFL-condition $\tau \leq (C_S)^{-1}h$ there holds*

$$\|u_h^N\|_{0,\Omega}^2 + \tau c_a\sum_{n=0}^{N-1}\left\|\|u_h^{n+\frac{1}{2}}\|\right\|^2 \lesssim T\tau\sum_{n=0}^{N-1}\|f^{n+\frac{1}{2}}\|_{0,\Omega}^2 + \|u_h^0\|_{0,\Omega}^2. \quad (4.7)$$

Proof. An immediate consequence of the inverse inequality (3.4), Proposition 4.2 and the discrete Gronwall's inequality. $\square$

REMARK 4.4. It follows that the fully implicit method, $\lambda = 1$ must satisfy the CFL-condition for initial conditions with minimal regularity, i.e. only if $u_0 \in L^2(\Omega)$; since

$$\tau s_h(u_h^0, u_h^0) \leq C_S\frac{\tau}{h}\|u_h^0\|_{0,\Omega}^2.$$

This stability condition can be lifted if $u_0 \in H^1(\Omega)$ ($s_h(u_h^0, u_h^0) \lesssim h\|\beta\|_\infty\|\nabla u_h^0\|_{0,\Omega}^2$), or by using the following projection for the definition of u_h^0: find $u_h^0 \in V_h$ such that

$$(u_h^0, v_h) + \tau s_h(u_h^0, v_h) = (u^0, v_h), \quad \forall v_h \in V_h.$$

This projection leads to optimal estimates in the convection-dominated regime, or under the CFL-condition.

4.2. Convergence.   The following convergence result is readily obtained from Lemma 4.2 by using arguments in the spirit of [4].

THEOREM 4.5 (Convergence of the Crank-Nicolson scheme (4.1)). *Let u be the solution of (2.1) and $\{u_h^n\}_{n=1}^N$ be the solution of (4.1) with $s_h(\cdot,\cdot)$ defined by (3.5). Assume that the hypothesis of Lemma 4.2 are satisfied and let $e_h^n \overset{\text{def}}{=} \pi_h u(t^n) - u_h^n$, where π_h denotes the standard L^2-projection onto V_h. Then for $N \geq 1$ there holds*

$$\|e_h^n\|_{0,\Omega}^2 + \sum_{m=0}^{n-1} \tau \|\!|e_h^{m+\frac{1}{2}}\|\!|^2 \lesssim T\mathcal{H}^2(T,\mu,\boldsymbol{\beta},\sigma)\|u\|_{C^0(I;H^{k+1}(\Omega))}^2$$
$$+ T\tau^4 \|\partial_{ttt}u\|_{L^2(I;L^2(\Omega))}^2, \quad 1 \leq n \leq N,$$

with $\mathcal{H}(T,\mu,\boldsymbol{\beta},\sigma) \overset{\text{def}}{=} h^k\left(\|\boldsymbol{\beta}\|_\infty^{\frac{1}{2}}h^{\frac{1}{2}} + T^{\frac{1}{2}}\|\boldsymbol{\beta}\|_{1,\infty,\Omega}h + \|\sigma\|_\infty^{\frac{1}{2}}h + \|\mu\|_\infty^{\frac{1}{2}}\right).$

Note that this result implies (quasi-) optimal convergence for smooth solutions independent of the diffusion parameter.

5. A numerical example

We consider as example a pure transport problem in two dimensions: a rotating Gaussian on the unit disc. Hence, in (2.1), we take

$$\boldsymbol{\beta} = (-y,x)^{\mathrm{T}}, \quad \sigma = \mu = f = 0, \quad \text{and} \quad \Omega := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\},$$

with the Gaussian initial condition centered at $(0.3, 0.3)$ given by

$$u_0(x,y) = e^{-10[(x-0.3)^2 + (y-0.3)^2]}.$$

We compute on three different meshes with $\tau = h$. In Fig. 5.1 we show the convergence plots of the L^2-error for the computation of the solution using the CIP stabilized method and $\lambda = \frac{1}{2}$ (dotted), $\lambda = 1$ (dotted) and $\lambda = 0$ (dashed). The error is very similar for explicit and implicit treatments of the stabilization term. In fact the graphs of the two implicit methods can not be distinguished, whereas the semi-implicit method has a slightly larger error. Similar observations were made for non-smooth solutions. It should be noted however that the stabilization parameter was chosen ten times larger in the semi-implicit case than in the implict case for optimal results. For reference we also plot the suboptimal convergence of the standard Galerkin method (dash/dot) and the optimal second order convergence (full line).

6. Application to optimal control or inverse problems

Several authors have discussed the advantages of symmetric stabilization for optimal control problems in the stationary case; see [1, 2, 10]. The design of numerical methods for the solution of optimization problems subject to *transient* advection–diffusion–reaction equations, on the other hand, has received relatively little attention in the literature; we cite the work of Fu and Rui [6], where a characterisitic finite element method of first order is analyzed and implemented. We will here show how the above stabilized Crank-Nicolson method can be applied in a framework similar to the one of [6], leading to optimal estimates for smooth solutions of the control problem independent of the diffusion coefficient, and also for high order discretization methods. Consider the following distributed optimal control problem on the space-time domain $Q \overset{\text{def}}{=} \Omega \times I$:

$$\begin{cases} \min_{u \in L^2(Q)} J(y(u),u), \\ \partial_t y + \boldsymbol{\beta} \cdot \nabla y - \mu \Delta y = f + u, \quad \text{in } Q, \\ y = 0 \quad \text{on} \quad \partial\Omega \times I \text{ and } y(\cdot,0) = 0, \text{ in } \Omega, \end{cases} \qquad (6.1)$$
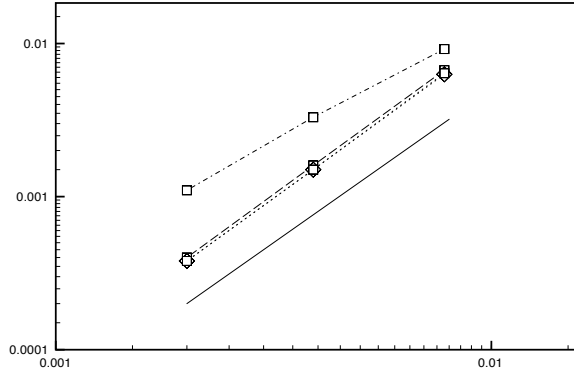
FIG. 5.1. *Convergence in the L^2-norm plotted against τ, of the stabilized (dashed line for $\lambda = 0$ and dotted line for $\lambda \in \{1/2, 1\}$ in the stabilization operator) and unstabilized (dash-dot) Crank-Nicolson method. The full line shows optimal second order convergence.*

where $\mu \geq \mu_0 > 0$, $J(u) \stackrel{\text{def}}{=} \frac{1}{2}\|y(u) - y_d\|_{0,Q}^2 + \alpha\|u\|_{0,Q}^2$, with $\alpha \in \mathbb{R}$, $\alpha > 0$, and $y_d \in L^2(Q)$. The weak formulation of the constraint equation is given by (2.2). It is straightforward to show that the minimizer u^* is a solution to the following system:

$$\begin{cases} (\partial_t y, v) + a(y, v) = (f + u^*, v) & \forall v \in V, \quad a.e.\, t \in I, \quad \text{and with } y(x,0) = 0, \\ (-\partial_t p, w) + a(w, p) = (y - y_d, w) & \forall w \in V, \quad a.e.\, t \in I, \quad \text{and with } p(x,T) = 0, \quad (6.2) \\ \int_0^T (\alpha u^* + p, x - u^*)\, dt = 0 & \forall x \in L^2(Q). \end{cases}$$

6.1. Discrete optimization problem and analysis. We obtain the discrete optimization problem by discretizing the state and adjoint equations using the formulation (4.1). Here we simply consider the standard finite element space V_h for all the unknowns. In particular we discretize the control u by $u_h \in U_h \stackrel{\text{def}}{=} [V_h]^N$. The control at a given time level t^n will be denoted $u^n \stackrel{\text{def}}{=} u(t^n)$. We will look for the control $u_h(t)$ such that $u_h|_{I_n} = u_h^{n+\frac{1}{2}} \approx u(t^{n+\frac{1}{2}})$, where $I_n = (t^n, t^{n+1}]$. Note the slight abuse of notation: contrary to $y_h^{n+\frac{1}{2}}$ and $p_h^{n+\frac{1}{2}}$, $u_h^{n+\frac{1}{2}}$ is *not* defined by $(u_h^{n+1} + u_h^n)/2$ but is an element of V_h. We may then state the discrete formulation of the optimality problem (6.2):

$$\begin{cases} \text{For } 0 \leq n \leq N-1, \text{ find } u_h^{*,n+\frac{1}{2}}, y_h^{n+1}, \text{ and } p_h^n \in V_h \text{ such that } y_h^0 = 0, \, p_h^N = 0, \text{ and} \\ (D_\tau y_h^{n+1}, v_h) + a_h(y_h^{n+\frac{1}{2}}, v_h) + s_h(y_h^{n+\lambda}, v_h) = (f^{n+\frac{1}{2}} + u_h^{*,n+\frac{1}{2}}, v_h) \quad \forall v_h \in V_h, \\ (-D_\tau p_h^{n+1}, w_h) + a_h(w_h, p_h^{n+\frac{1}{2}}) + s_h(p_h^{n+1-\lambda}, w_h) = (y_h^{n+\frac{1}{2}} - y_d^{n+\frac{1}{2}}, w_h) \quad \forall w_h \in V_h, \\ \tau \sum_{m=0}^{N-1} (\alpha u_h^{*,m+\frac{1}{2}} + p_h^{m+\frac{1}{2}}, x_h - u_h^{*,m+\frac{1}{2}}) = 0 \quad \forall x_h \in U_h. \end{cases}$$

$$(6.3)$$

Note that the stabilization is of the same type, either explicit or implicit, both in the forward and the backward equations.

LEMMA 6.1. *For the system* (6.3) *there holds*

$$\sum_{n=0}^{N-1}(f^{n+\frac{1}{2}}+u_h^{*,n+\frac{1}{2}},p_h^{n+\frac{1}{2}})=\sum_{n=0}^{N-1}(y_h^{n+\frac{1}{2}}-y_d^{n+\frac{1}{2}},y_h^{n+\frac{1}{2}}).\qquad(6.4)$$

Proof. Immediately by the formulation we have

$$\sum_{n=0}^{N-1}(f^{n+\frac{1}{2}}+u_h^{*,n+\frac{1}{2}},p_h^{n+\frac{1}{2}})$$

$$=\sum_{n=0}^{N-1}\left((D_\tau y_h^{n+1},p_h^{n+\frac{1}{2}})+a_h(y_h^{n+\frac{1}{2}},p_h^{n+\frac{1}{2}})+s_h(y_h^{n+\lambda},p_h^{n+\frac{1}{2}})\right)$$

$$=\sum_{n=0}^{N-1}\left((-D_\tau p_h^{n+1},y_h^{n+\frac{1}{2}})+a_h(y_h^{n+\frac{1}{2}},p_h^{n+\frac{1}{2}})+s_h(y_h^{n+\lambda},p_h^{n+\frac{1}{2}})\right)$$

$$=\sum_{n=0}^{N-1}(y_h^{n+\frac{1}{2}}-y_d^{n+\frac{1}{2}},y_h^{n+\frac{1}{2}})+\sum_{n=0}^{N-1}\left(s_h(y_h^{n+\lambda},p_h^{n+\frac{1}{2}})-s_h(y_h^{n+\frac{1}{2}},p_h^{n+1-\lambda})\right).$$

It remains to show that the last sum in the right hand side is zero. For $\lambda=1/2$ the result is trivially true. Consider now the case $\lambda=0$. Using the fact that $y_h^0=p_h^T=0$ we have

$$\sum_{n=0}^{N-1}s_h(y_h^n,p_h^{n+\frac{1}{2}})=\frac{1}{2}\sum_{n=0}^{N-1}(s_h(y_h^n,p_h^n)+s_h(y_h^n,p_h^{n+1}))$$

$$=\frac{1}{2}\left(\sum_{n=0}^{N-2}s_h(y_h^{n+1},p_h^{n+1})+\sum_{n=0}^{N-2}s_h(y_h^n,p_h^{n+1})\right)=\sum_{n=0}^{N-1}s_h(y_h^{n+\frac{1}{2}},p_h^{n+1}).$$

The case $\lambda=1$ is similar. □

We also introduce the auxiliary equations:

$$\begin{cases} \text{For } 0\le n\le N-1, \text{ with } \tilde{y}_h^0=0, \tilde{p}_h^N=0, \rho_h^N=0 \text{ and find } \tilde{y}_h^{n+1}, p_h^n, \rho_h^n\in V_h \text{ such that} \\ (D_\tau \tilde{y}_h^{n+1},v_h)+a_h(\tilde{y}_h^{n+\frac{1}{2}},v_h)+s_h(\tilde{y}_h^{n+\lambda},v_h)=(f^{n+\frac{1}{2}}+u^{*,n+\frac{1}{2}},v_h) \quad \forall v_h\in V_h, \\ (-D_\tau \tilde{p}_h^{n+1},v_h)+a_h(v_h,\tilde{p}_h^{n+\frac{1}{2}})+s_h(\tilde{p}_h^{n+1-\lambda},v_h)=(\tilde{y}_h^{n+\frac{1}{2}}-y_d^{n+\frac{1}{2}},v_h) \quad \forall v_h\in V_h, \\ (-D_\tau \rho_h^{n+1},v_h)+a_h(v_h,\rho_h^{n+\frac{1}{2}})+s_h(\rho_h^{n+1-\lambda},v_h)=(y^{n+\frac{1}{2}}-y_d^{n+\frac{1}{2}},v_h) \quad \forall v_h\in V_h, \end{cases}$$
$$(6.5)$$

where u^*,y are the control and state functions satisfying (6.2). Recall that $u^{*,n+\frac{1}{2}}\stackrel{\text{def}}{=}u^*(t^{n+\frac{1}{2}})$ and similarly for $y_d^{n+\frac{1}{2}}$ and $y^{n+\frac{1}{2}}$.

THEOREM 6.1. *Let the triplet* $\{u_h^{*,n+\frac{1}{2}},y_h^{n+1},p_h^n\}_{n=0}^{N-1}$ *be the solution of the discrete optimality system* (6.3), *and* $u^*\in C^0(I;L^2(\Omega))$, $y\in C^0(I;H^1(\Omega))$, $p\in C^0(I;H^1(\Omega))$ *the solution of the continuous optimality system* (6.2), *then there holds*

$$\tau\sum_{n=0}^{N-1}\alpha\|u^{*,n+\frac{1}{2}}-u_h^{*,n+\frac{1}{2}}\|_{0,\Omega}^2\lesssim\tau\sum_{n=0}^{N-1}\alpha^{-1}(\|\rho_h^{n+\frac{1}{2}}-p^{n+\frac{1}{2}}\|_{0,\Omega}^2+T^2\|\tilde{y}_h^{n+\frac{1}{2}}-y^{n+\frac{1}{2}}\|_{0,\Omega}^2),$$
$$(6.6)$$

and for $n = 1, \dots, N$ we have

$$\|y_h^n - y^n\|_{0,\Omega}^2 \lesssim \|\tilde{y}_h^n - y^n\|_{0,\Omega}^2 + \tau T \sum_{m=0}^{N-1} \|u_h^{*,m+\frac{1}{2}} - u^{*,m+\frac{1}{2}}\|_{0,\Omega}^2. \tag{6.7}$$

Proof. By adding and subtracting $\alpha^{-1} p^{n+\frac{1}{2}}$, $\alpha^{-1} p_h^{n+\frac{1}{2}}$, and $\alpha^{-1} \tilde{p}_h^{n+\frac{1}{2}}$ in the left slot of the scalar product of the left hand side of (6.6) we have

$$\alpha\tau \sum_{n=0}^{N-1} \|u^{*,n+\frac{1}{2}} - u_h^{*,n+\frac{1}{2}}\|_{0,\Omega}^2 = \tau \sum_{n=0}^{N-1} (\alpha u(t^{n+\frac{1}{2}}) + p^{n+\frac{1}{2}}, u^{*,n+\frac{1}{2}} - u_h^{*,n+\frac{1}{2}})$$

$$-\tau \sum_{n=0}^{N-1} (\alpha u_h^{*,n+\frac{1}{2}} + p_h^{n+\frac{1}{2}}, \pi_h u^{*,n+\frac{1}{2}} - u_h^{*,n+\frac{1}{2}}) + \tau \sum_{n=0}^{N-1} (p_h^{n+\frac{1}{2}} - \tilde{p}_h^{n+\frac{1}{2}}, u^{*,n+\frac{1}{2}} - u_h^{*,n+\frac{1}{2}})$$

$$+\tau \sum_{n=0}^{N-1} (\tilde{p}_h^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}, u^{*,n+\frac{1}{2}} - u_h^{*,n+\frac{1}{2}}) = I_1 + I_2 + I_3 + I_4.$$

Under the regularity assumptions, $I_1 \equiv 0$ by the continuous optimality condition (6.2) and $I_2 \equiv 0$ by the discrete optimality condition (6.3). For I_4 we immediately have

$$I_4 \leq \alpha^{-1} \frac{1}{2} \tau \sum_{n=0}^{N-1} \|\tilde{p}_h^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\|_{0,\Omega}^2 + \frac{1}{2} \alpha\tau \sum_{n=0}^{N-1} \|u^{*,n+\frac{1}{2}} - u_h^{*,n+\frac{1}{2}}\|_{0,\Omega}^2.$$

It only remains to estimate term I_3. To this aim we introduce the perturbation equation for $\theta_h^n = \tilde{y}_h^n - y_h^n$ and $\zeta_h^n = \tilde{p}_h^n - p_h^n$ obtained by taking the difference of the first two equations of (6.5) and the first two equations of (6.3):

$$\begin{cases} \text{For } 0 \leq n \leq N, \ \theta_h^n \text{ and } \zeta_h^n \in V_h \text{ satisfy } \theta_h^0 = 0, \ \zeta_h^N = 0, \text{ and} \\ (D_\tau \theta_h^{n+1}, v_h) + a_h(\theta_h^{n+\frac{1}{2}}, v_h) + s_h(\theta_h^{n+\lambda}, v_h) = (u^{*,n+\frac{1}{2}} - u_h^{*,n+\frac{1}{2}}, v_h) \quad \forall v_h \in V_h, \\ (-D_\tau \zeta_h^{n+1}, w_h) + a_h(w_h, \zeta_h^{n+\frac{1}{2}}) + s_h(w_h, \zeta_h^{n+1-\lambda}) = (\theta_h^{n+\frac{1}{2}}, w_h) \quad \forall w_h \in V_h. \end{cases} \tag{6.8}$$

It follows from the definition of the perturbation equations and the discrete adjoint consistency of Lemma 6.1 that, since $I_3 = -\tau \sum_{n=0}^{N-1} (\zeta_h^{n+\frac{1}{2}}, u^{*,n+\frac{1}{2}} - u_h^{*,n+\frac{1}{2}})$,

$$I_3 = -\tau \sum_{n=0}^{N-1} \left[(D_\tau \theta_h^{n+1}, \zeta_h^{n+\frac{1}{2}}) + a_h(\theta_h^{n+\frac{1}{2}}, \zeta_h^{n+\frac{1}{2}}) + s_h(\theta_h^{n+\lambda}, \zeta_h^{n+\frac{1}{2}}) \right]$$

$$= -\tau \sum_{n=0}^{N-1} \left[-(\theta_h^{n+\frac{1}{2}}, D_\tau \zeta_h^{n+1}) + a_h(\theta_h^{n+\frac{1}{2}}, \zeta_h^{n+\frac{1}{2}}) + s_h(\zeta_h^{n+1-\lambda}, \theta_h^{n+\frac{1}{2}}) \right]$$

$$= -\tau \sum_{n=0}^{N-1} \|\theta_h^{n+\frac{1}{2}}\|_{0,\Omega}^2.$$

We deduce that $\alpha\tau \sum_{n=0}^{N-1} \|u_h^{*,n+\frac{1}{2}} - u^{*,n+\frac{1}{2}}\|_{0,\Omega}^2 \leq \alpha^{-1} \tau \sum_{n=0}^{N-1} \|\tilde{p}_h^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\|_{0,\Omega}^2$. To obtain the inequality (6.6) we add and subtract $\rho_h^{n+\frac{1}{2}}$ in the right hand side of the last equation and apply the triangle inequality:

$$\tau \sum_{n=0}^{N-1} \|\tilde{p}_h^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\|_{0,\Omega}^2 \leq 2\tau \sum_{n=0}^{N-1} \|\tilde{p}_h^{n+1} - \rho_h^{n+1}\|_{0,\Omega}^2 + 2\tau \sum_{n=0}^{N-1} \|\rho_h^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\|_{0,\Omega}^2.$$

Then, to control the first term in the right hand side, apply the stability estimate of Corollary 4.3 to the perturbation equation for $\tilde{p}_h^n - \rho_h^n$, obtained by taking the difference of the two last equations of (6.5):

$$\tau \sum_{n=0}^{N-1} \|\tilde{p}_h^{n+1} - \rho_h^{n+1}\|_{0,\Omega}^2 \lesssim T^2 \tau \sum_{n=0}^{N-1} \|\tilde{y}_h^{n+\frac{1}{2}} - y^{n+\frac{1}{2}}\|_{0,\Omega}^2.$$

Note that Corollary 4.3 is applied to all the time levels n separately and then summed up, resulting in an additional factor T. For (6.7) we proceed in a similar fashion, first using the triangular inequality $\|y_h^n - y^n\|_{0,\Omega}^2 \le 2\|\theta_h^n\|_{0,\Omega}^2 + 2\|\tilde{y}_h^n - y^n\|_{0,\Omega}^2$ and then the stability estimate of Corollary 4.3 applied to $\|\theta_h^n\|_{0,\Omega}^2$, yielding $\|\theta_h^n\|_{0,\Omega}^2 \lesssim T\tau \sum_{m=0}^{n-1} \|u_h^{*,n+\frac{1}{2}} - u^{*,n+\frac{1}{2}}\|_{0,\Omega}^2.$ $\qquad\square$

THEOREM 6.2. *Let* $y,p \in C^3(I;L^2(\Omega)) \cap C^0(I;H^{k+1}(\Omega))$ *be the exact solution of* (2.2) *and* u_h^*, y_h, p_h *the solution of* (6.3) *in the convection dominated regime* ($\mu < |\boldsymbol{\beta}|h$) *with* $s_h(\cdot,\cdot)$ *defined by* (3.5). *Then*

$$\left(\tau \sum_{n=0}^{N-1} \|u_h^{*,n+\frac{1}{2}} - u^{*,n+\frac{1}{2}}\|_{0,\Omega}^2\right)^{\frac{1}{2}} \le C(\boldsymbol{\beta},\alpha,u,T)(h^{k+\frac{1}{2}} + \tau^2)$$

and $\|y_h^m - y^m\|_{0,\Omega} \le C(\boldsymbol{\beta},\alpha,u,T)(h^{k+\frac{1}{2}} + \tau^2), \quad 0 \le m \le N.$

Proof. By Equation (6.7) of Theorem 6.1 it is enough to prove the first inequality. Applying Theorem 4.5 to the errors in ρ_h^n and $\tilde{y}_h^n$ leads to

$$\|\rho_h^n - p^n\|_{0,\Omega} \le C(\boldsymbol{\beta},u,T)(h^{k+\frac{1}{2}} + \tau^2) \text{ and } \|\tilde{y}_h^n - y^n\|_{0,\Omega} \le C(\boldsymbol{\beta},u,T)(h^{k+\frac{1}{2}} + \tau^2). \tag{6.9}$$

By considering the right hand side of (6.6) we have

$$\tau \sum_{n=0}^{N-1} \|\rho_h^{n+\frac{1}{2}} - p^{n+\frac{1}{2}}\|_{0,\Omega}^2 \lesssim \tau \sum_{n=0}^{N} \|\rho_h^n - p^n\|_{0,\Omega}^2 + \tau \sum_{n=0}^{N} \|p^{n+\frac{1}{2}} - (p^{n+1} + p^n)/2\|_{0,\Omega}^2.$$

The bound for the first sum is obtained by applying (6.9) termwise, and the bound for the second term is obtained by a standard Taylor development. The second contribution of (6.6) is treated similarly, using the right inequality of (6.9). $\qquad\square$

7. Conclusion

We have designed stabilized Crank-Nicolson methods for transient advection–diffusion–reaction equations where the stabilization may be treated either implicitly or fully explicitly (under a hyperbolic CFL-condition). We have also shown that the proposed scheme may be used for the solution of optimal control problems or inverse problems, yielding optimal error estimates that are robust with respect to the transport regime (local Peclet number) for smooth solutions.

REFERENCES

[1] R. Becker and B. Vexler, *Optimal control of the convection-diffusion equation using stabilized finite element methods*, Numer. Math., 106(3), 349–367, 2007.

[2] M. Braack, *Optimal control in fluid mechanics by finite elements with symmetric stabilization*, SIAM J. Control Optim., 48(2), 672–687, 2009.

[3] E. Burman and P. Hansbo, *Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems*, Comput. Methods Appl. Mech. Engrg., 193(15-16), 1437–1453, 2004.

[4] E. Burman and M. Fernández, *Finite element methods with symmetric stabilization for the transient convection-diffusion-reaction equation*, Comput. Methods Appl. Mech. Engrg., 198(33-36), 2508–2519, 2009.

[5] R. Codina and J. Blasco, *Analysis of a stabilized finite element approximation of the transient convection-diffusion-reaction equation using orthogonal subscales*, Comput. Vis. Sci., 4(3), 167–174, 2002.

[6] H. Fu and H. Rui, *A priori error estimates for optimal control problems governed by transient advection-diffusion equations*, J. Sci. Comput., 38(3), 290–315, 2009.

[7] J-L. Guermond, *Subgrid stabilization of Galerkin approximations of linear contraction semi-groups of class C^0 in Hilbert spaces*, Numer. Meth. Part. Differ. Equs., 17(1), 1–25, 2001.

[8] J. Nitsche, *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilrumen, die keinen Randbedingungen unterworfen sind*, (German) Collection of Articles Dedicated to Lothar Collatz on His Sixtieth Birthday. Abh. Math. Sem. Univ. Hamburg, 36, 9–15, 1971.

[9] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 25, 1997.

[10] N. Yan and Z. Zhou, *A priori and a posteriori error analysis of edge stabilization Galerkin method for the optimal control problem governed by convection-dominated diffusion equation*, J. Comput. Appl. Math., 223(1), 198–217, 2009.