ERROR ESTIMATES FOR THE DISCONTINUOUS GALERKIN METHODS FOR IMPLICIT PARABOLIC EQUATIONS

K. CHRYSAFINOS[†] AND NOEL. J. WALKINGTON[†]

Abstract. We analyze the classical discontinuous Galerkin method for "implicit" parabolic equations. Symmetric error estimates for schemes of arbitrary order are presented. The ideas we develop allow us to relax many assumptions frequently required in previous work. For example, we allow different discrete spaces to be used at each time step and do not require the spatial operator to be self adjoint or independent of time. Our error estimates are posed in terms of projections of the exact solution onto the discrete spaces and are valid under the minimal regularity guaranteed by the natural energy estimate. These projections are local and enjoy optimal approximation properties when the solution is sufficiently regular.

1. Introduction. We consider implicit parabolic partial differential equations of the form,

$$(M(t)u)_t + A(t)u = F(t), \qquad u(0) = u_0. \tag{1.1}$$

The operators act on Hilbert spaces related through the standard pivot construction, $U \hookrightarrow H \simeq H' \hookrightarrow U'$, where each embedding is continuous and dense. Then, $A(.) : U \to U'$ is a linear map and $F(.) \in U'$. We assume that $M(.) : H \to H$ is a self adjoint positive definite operator.

Conservation laws for systems undergoing diffusion may take the form of (1.1) when the capacity changes with time; for example, in a porous medium flow the porosity could change as the medium collapses due to oil being extracted from the reservoir. Classical parabolic equations (i.e. equations with M the identity) also take the form of (1.1) under a time dependent change of coordinates; common examples being diffusion on surfaces (more generally manifolds) which are in motion, and the Lagrange (or characteristic) Galerkin formulation of the convection diffucsion equation [9, 18].

Here we analyze the classical discontinuous Galerkin (DG) scheme for approximating solutions of (1.1) and derive fully-discrete error estimates under minimal regularity assumptions. The class of DG schemes we consider are classical in the sense that the discrete solutions may be discontinuous in time but are conforming in space, i.e. are in (a subspace of) U at each time. Our analysis extends the ideas introduced in [5] and addresses the following issues which have not yet been adequately considered in the literature.

- A systematic treatment of DG approximations of implicit parabolic eqations of the form (1.1) have not been considered in the past.

*Mathematics subject classification: 65M60

[†]Department of Mathematics, Carnegie Mellon University, Pittsburgh, PA 15213. Supported in part by National Science Foundation Grants DMS–0208586, and ITR 0086093. This work was also supported by the NSF through the Center for Nonlinear Analysis.

Example: Diffusion on Manifolds As an illustrative example, consider diffusion on a cell membrane, $\mathcal{S}(t) \subset \mathbb{R}^3$, which is being transported in an ambient fluid with velocity $\mathbf{V} = \mathbf{V}(t, x)$. Since the membrane is diffeomorphic to a sphere, a numerical scheme may triangulate the sphere S^2, and at each time t construct a mapping $x(t, .) : S^2 \rightarrow \mathcal{S} \subset \mathbb{R}^3$. If the sphere is locally parameterized by coordinates $X \in U \subset \mathbb{R}^2$ the diffusion equation takes the form

$$u_t - (1/\sqrt{J}) div_X \left(\sqrt{J} (F^T F)^{-1} \nabla_X u \right) = f,$$

where F is the 3×2 matrix with components $F_{i\alpha} = \partial x_i / \partial X_\alpha$, and $J = det(F^T F)$ is the determinant of the first fundamental form. The determinant J satisfies $J_t = 2J\,(I - \mathbf{n} \otimes \mathbf{n}) \cdot (\nabla_x \mathbf{V}) = 2J \sum_{ij} (\delta_{ij} - n_i n_j)(\partial V_i / \partial x_j)$, where $\mathbf{n} = \mathbf{n}(t, X)$ is the normal to $\mathcal{S}(.)$. It follows that the diffusion equation is of the form (1.1) with $M(.)u = \sqrt{J} u$ and

$$A(.)u = -(I - \mathbf{n} \otimes \mathbf{n}) \cdot (\nabla_x \mathbf{V}) u \sqrt{J} - div_X \left(\sqrt{J} (F^T F)^{-1} \nabla_X u \right).$$

As $\mathcal{S}(.)$ evolves the matrix $F^T F$ will become ill conditioned and the sphere S^2 will need to be retriangulated so different discrete subspaces will be required on different intervals $(t^{n-1}, t^n]$ of $[0, T]$.

1.1. Related Results. The discontinuous Galerkin method was first introduced to model and simulate neutron transport by Lasaint and Raviart in [16]. There is an abundant literature concerning applications of the DG scheme in hyperbolic problems, see e.g. [4, 15, 26] and references within. The DG method for ordinary differential equations was considered by Delfour, Hager and Trochu in [6]. They showed that the DG scheme was super convergent at the partition points (order $2k+2$ for polynomials of degree k).

In the context of parabolic equations DG schemes were first analyzed for linear parabolic problems by Jamet in [14] where $\mathcal{O}(\tau^k)$ results were proved and then by Eriksson, Johnson and Thomée [12] where $\mathcal{O}(\tau^{2k-1})$ estimates are established at the partition points for smooth smooth solutions. An excellent exposition of their results and, more generally, the DG method for parabolic equations, can be found in Thomée's book [25]. In [25] nodal and interior estimates are presented in various norms. One may also consult [19] for the analysis of a related formulation based on the backward Euler scheme. The relation between the DG scheme and adaptive techniques was studied in [10] and [11]. Finally, some results concerning the analysis of parabolic integro-differential equations by discontinuous Galerkin method are presented in [17] (see also references therein).

In [9] DuPont and Liu introduce the concept of "symmetric error estimates" for parabolic problems. They define such an error estimate to be one of the form

$$|||u - u_h||| \leq C \inf_{w_h \in \mathcal{U}_h} |||u - w_h|||,$$

where u and u_h are the exact and approximate solutions respectively, $|||.|||$ is an appropriate norm, and $\mathcal{U}_h$ is the discrete subspace in which approximation solutions are sought. While estimates of this form are standard for elliptic problems, this is not the case for evolution problems. For example, error estimates for evolution problems approximated by the implicit Euler scheme frequently involve terms of the form $\|u_{tt}\|_{L^2(\Omega)}$.

error estimates which take the form of the sum of (i) the "local truncation error", (ii) projection errors between different subspaces, and (iii) errors in the initial data.

One technical distinction between the error estimate developed for the classical parabolic problem in [5, Theorem 3.1] and Theorem 5.1 of Section 5 is that the latter assumes the existence of an inverse hypothesis of the form $\|u_h\|_{U(t)} \leq C_{inv}(h)\|u_h\|_{H(t)}$ for u_h in the discrete subspaces of U. In Theorem 5.1 the product $\tau C_{inv}(h)$ enters into the error estimate where τ is the time step size. For classical second order parabolic problems this term will be of order $\mathcal{O}(1)$ if $\tau \sim h$ and quasi-uniform finite element meshes with no small angles are used.

1.3. Notation. We deal with spaces $H(t) = (H, \|.\|_{H(t)})$ and $U(t) = (U, \|.\|_{U(t)})$ whose norms depend upon time. The pivot spaces $H(t)$ have inner product $(u,v)_{H(t)} = (M(t)u, v)_H$, and we often denote the norms on the these spaces by $|.|_{H(t)} \equiv \|.\|_{H(t)}$. Motivated by various examples, we assume that $\|u\|^2_{U(t)} = |u|^2_{U(t)} + |u|^2_{H(t)}$ where $|u|_{U(t)}$ is a semi norm on $U(t)$ (the principle part). In particular $U(t) \hookrightarrow H(t)$ with embedding constant independent of time. Each norm and inner product will be explicitly subscripted; while this is rather cumbersome it helps to minimize confusion due the plethora of spaces and projections. Notation of the form $L^2[0,T;U(.)]$, $H^1[0,T;U'(.)]$ etc. is used to indicate the temporal regularity of functions with values in $U(.)$, $U'(.)$ etc.

Approximations of (1.1) will be constructed on a partition $0 = t^0 < t^1 < \ldots < t^N = T$ of $[0,T]$. On each interval of the form $(t^{n-1}, t^n]$ a subspace U_h^n of U is specified, and the approximate solutions will lie in the space

$$\mathcal{U}_h = \{u_h \in L^2[0,T;U(.)] \mid u_h|_{(t^{n-1},t^n]} \in \mathcal{P}_k(t^{n-1},t^n;U_h^n)\}.$$

Here $\mathcal{P}_k(t^{n-1},t^n;U_h^n)$ is the space of polynomials of degree k or less having values in U_h^n. Notice that, by convention, we have chosen functions in $\mathcal{U}_h$ to be left continuous with right limits. We will write u^n for $u_h(t^n) = u_h(t^n_-)$, and let u^n_+ denote $u(t^n_+)$. This notation will is also used with functions like the error $e = u - u_h$. We always assume the exact solution, u, is in $C[0,T;H(.)]$ so that the jump in the error at t^n, denoted by $[e^n]$, is equal to $[u^n] = u^n_+ - u^n$.

2. Implicit Parabolic Equations. In this section we introduce structural assumptions required for our analysis of the implicit parabolic problem

$$(M(t)u)_t + A(t)u = F(t), \qquad u(0) = u_0. \tag{2.1}$$

To characterize the time dependence of $A(.)$ we introduce equivalent norms on U of the form $\|u\|^2_{U(t)} = |u|^2_{H(t)} + |u|^2_{U(t)}$ where $|.|_{U(t)}$ is a seminorm on U and $|.|_{H(t)}$ is the norm on H with Riesz map (the symmetric positive operator) $M(t)$. We denote by $a(.;u,v)$ the natural bilinear forms associated with $A(.)$ and will assume that the time dependent spaces $U(t), H(t)$ satisfy $U(t) \hookrightarrow H(t) \hookrightarrow U'(t)$, where each embedding is dense and continuous.

2.2. Properties of $H(t)$. The smoothness assumption 2.1 guarantees that the norms on the pivot spaces $H(t)$ vary continuously with t. This following lemma quantifies this and will be used ubiquitously below.

LEMMA 2.2. *Let w, $z \in H$ and $s \leq t$, then* $e^{C_\mu(s-t)} \leq |z|^2_{H(t)}/|z|^2_{H(s)} \leq e^{C_\mu(t-s)}$ *and*

$$|(w,z)_{H(t)} - (w,z)_{H(s)}| \leq (t-s)C_\mu e^{C_\mu(t-s)}|w|_{H(\xi_1)}|z|_{H(\xi_2)}, \qquad \xi_1,\ \xi_2 \in [s,t].$$

Proof. The differentiability of $(.,.)_{H(.)}$ implies

$$(w,z)_{H(t)} - (w,z)_{H(s)} = \int_s^t \frac{d}{d\xi}(w,z)_{H(\xi)}\, d\xi = \int_s^t \mu(\xi,w,z).$$

Putting $w = z$ gives

$$|z|^2_{H(t)} = |z|^2_{H(s)} + \int_s^t \mu(\xi,z,z)\, d\xi \leq |z|^2_{H(s)} + C_\mu \int_s^t |z|^2_{H(\xi)}\, d\xi.$$

Gronwall's inequality then shows $|z|^2_{H(t)} \leq |z|^2_{H(s)} e^{C_\mu(t-s)}$. Since this argument is symmetric in s and t the first inequality follows.

The second inequality now follows from the intermediate value theorem:

$$|(w,z)_{H(t)} - (w,z)_{H(s)}| \leq C_\mu \int_s^t |w|_{H(\xi)}|z|_{H(\xi)}\, d\xi = C_\mu(t-s)|w|_{H(\hat{\xi})}|z|_{H(\hat{\xi})},$$

for some $\hat{\xi} \in [s,t]$. Upon introducing a factor of $e^{C_\mu(t-s)}$ we may replace each instance of $\hat{\xi}$ on the right by any $\xi \in [s,t]$. □

3. Discrete Spaces. Discrete subspaces $\mathcal{U}_h$ of $L^2[0,T;U(.)]$ are constructed from a partition $0 = t^0 < t^1 < \ldots < t^N = T$ and a sequence of subspaces $\{U_h^n\}_{n=1}^N$ of U as

$$\mathcal{U}_h = \{u_h \in L^2[0,T;U(.)] \ \mid \ u_h|_{(t^{n-1},t^n]} \in \mathcal{P}_k(t^{n-1},t^n;U_h^n)\}.$$

Projections of H onto the subspaces U_h^n with respect to the norms $H(t)$ appear in our analysis; the following notation is used to denote these.

NOTATION 1. *$P_n(t)$ is the projection $P_n(t) : H(t) \to U_h^n$ characterized by $P_n(t)u \in U_h^n$, $(P_n(t)u, v_h)_{H(t)} = (u, v_h)_{H(t)}$ for all $v_h \in U_h^n$.*

3.1. Discrete Characteristic Functions. Estimates for the solution $u(t)$ of an evolution equation are frequently obtained by multiplying the equation by $\chi_{[0,t)}u$. This choice of test functions is not available in the discrete context unless the terminal time is one of the partition points. To estimate the solution at times $t \in [t^{n-1}, t^n)$ we first recall the discrete discrete characteristic functions introduced in [5, Section 2.3].

The discrete characteristic functions is invariant under translation so it is convenient to work on the interval $[0,\tau)$ with $\tau = t^n - t^{n-1}$. We begin by considering polynomials $p \in$

The proof is essentially the same as in [5, Lemma 2.10]; the only difference concerns the scaling argument required to show that C_k can be chosen to be independent of time. To do this Lemma 2.2 is used to estimate $\|u-\tilde{u}\|_{L^2[0,\tau,H(.)]}$ by $\|u-\tilde{u}\|_{L^2[0,\tau,H(\tau/2)]}$ to remove the implicit time dependence through $H(t)$.

To bound $\tilde{u}$ in $L^2[0,\tau;U(.)]$ we will compare it with the algebraic projection $\hat{u}$ and invoke an inverse hypothesis.

LEMMA 3.3. *Let $u \in \mathcal{P}_k(0,\tau,U_h^n)$ and $\tilde{u}$ be the projections defined in (3.2). If $\hat{u}$ is the* algebraic *projection characterized by (3.1), then*

$$\|\hat{u}-\tilde{u}\|_{L^2[0,\tau;H(.)]} \le C(k,\mu)\tau\|u\|_{L^2[0,\tau;H(.)]},$$

where $C(\mu,k)$ is a constant depending on k and μ through C_μ and C_k, the constant in Lemma 3.2.

Proof. In this proof $C(k,\mu)$ denotes a constant depending only on C_k and C_μ which may change from step to step. Recall that $\hat{u} \in \mathcal{P}_k(0,\tau;U_h)$ satisfies $\hat{u}(0)=u(0)=\tilde{u}(0)$,

$$\int_0^\tau (\hat{u},w)_{H(0)} = \int_0^t (u,w)_{H(0)}, \qquad w \in \mathcal{P}_{k-1}(0,\tau;U_h),$$

and $\|\hat{u}\|_{L^2[0,\tau,H(0)]} \le C_k\|u\|_{L^2[0,\tau;H(0)]}$. If $w \in \mathcal{P}_{k-1}[0,\tau;U_h]$ then

$$\begin{aligned}\int_0^\tau (\tilde{u}-\hat{u},w)_{H(.)} &= \int_0^t (u,w)_{H(.)} - \int_0^\tau (\hat{u},w)_{H(.)},\\ &= \int_0^t \Big((u,w)_{H(.)}-(u,w)_{H(0)}\Big) - \int_0^\tau \Big((\hat{u},w)_{H(.)}-(\hat{u},w)_{H(0)}\Big).\end{aligned}$$

Since $(\hat{u}-\tilde{u})(0)=0$ it follows that $(\hat{u}-\tilde{u})(s)=s\bar{u}(s)$ where $\bar{u}\in\mathcal{P}_{k-1}(0,\tau;U_h)$. Putting $w=\bar{u}$ and using Lemma 2.2 to estimate the right hand side gives

$$\int_0^\tau s|\bar{u}|^2_{H(s)}\,ds \le \int_0^\tau sC(k,\mu)\big(|u(s)|_{H(0)}+|\hat{u}(s)|_{H(0)}\big)|\bar{u}(s)|_{H(s)}\,ds.$$

An application of the Cauchy Schwarz inequality then shows

$$\int_0^\tau s|\bar{u}(s)|^2_{H(s)}\,ds \le C(k,\mu)\int_0^\tau s\big(|u(s)|^2_{H(0)}+|\hat{u}(s)|^2_{H(0)}\big)\,ds \le C(k,\mu)\tau\int_0^\tau |u|^2_{H(0)}.$$

Using Lemma 2.2 to compare $|.|_{H(s)}$ with the fixed norm $|.|_{H(0)}$ we obtain

$$\int_0^\tau s|\bar{u}|^2_{H(0)}\,ds \le C(k,\mu)\tau\int_0^\tau |u|^2_{H(0)} \le C(k,\mu)\tau\int_0^\tau |u|^2_{H(.)}.$$

Using the equivalence of norms on $\mathcal{P}_{k-1}(0,\tau)$ and Lemma 2.2 once again to compare $|.|_{H(0)}$ with $|.|_{H(.)}$ we obtain

$$\int_0^\tau |\tilde{u}-\hat{u}|^2_{H(.)} = \int_0^\tau s^2|\bar{u}|^2_{H(s)}\,ds \le C(k,\mu)\tau^2\int_0^\tau |u|^2_{H(.)}.$$

Proof. Galerkin orthogonality shows that the error $e = u - u_h$ satisfies

$$(e^n, v^n)_{H(t^n)} - \int_{t^{n-1}}^{t^n} (e_h, v_{ht})_{H(t)} - (e^{n-1}, v_+^{n-1})_{H(t^{n-1})} = 0. \tag{4.3}$$

Using the definition of $\hat{e} = \hat{u} - u_h$ we have

$$\begin{aligned}(\hat{e}^n, v^n)_{H(t^n)} &- \int_{t^{n-1}}^{t^n} (e_h, v_{ht})_{H(t)} - (\hat{e}^{n-1}, v_+^{n-1})_{H(t^{n-1})} \\ &= \left(P_n(t^{n-1})(I - P_{n-1}(t^{n-1}))u(t^{n-1}), v_+^{n-1}\right)_{H(t^{n-1})}.\end{aligned}$$

Setting $v_h(t) = \hat{e}^n$ to be independent of time gives

$$|\hat{e}^n|^2_{H(t^n)} = (\hat{e}^{n-1}, \hat{e}^n)_{H(t^{n-1})} + \left(P_n(t^{n-1})(I - P_{n-1}(t^{n-1}))u(t^{n-1}), \hat{e}^n\right)_{H(t^{n-1})},$$

so that

$$|\hat{e}^n|^2_{H(t^n)} \le \left(|\hat{e}^{n-1}|_{H(t^{n-1})} + |P_n(t^{n-1})(I - P_{n-1}(t^{n-1}))u(t^{n-1})|_{H(t^{n-1})}\right)|\hat{e}^n|_{H(t^{n-1})}.$$

Lemma 2.2 shows that $|\hat{e}^n|_{H(t^{n-1})} \le \mathrm{e}^{C_\mu \tau_n}|\hat{e}^n|_{H(t^n)}$ where $\tau_n = t^n - t^{n-1}$, which gives

$$|\hat{e}^n|_{H(t^n)} \le \mathrm{e}^{C_\mu \tau_n}\left(|\hat{e}^{n-1}|_{H(t^{n-1})} + |P_n(t^{n-1})(I - P_{n-1}(t^{n-1}))u(t^{n-1})|_{H(t^{n-1})}\right)$$

The statement of the theorem now follows from the discrete Gronwall inequality. □

The next definition characterizes the local truncation error in the present context.

DEFINITION 4.2. *(1) The projection $\mathbb{P}_n^{loc} : C[t^{n-1}, t^n; H(.)] \to \mathcal{P}_k(t^{n-1}, t^n; U_h^n)$ satisfies $(\mathbb{P}_n^{loc}u)^n = P_n(t^n)u(t^n)$, and*

$$\int_{t^{n-1}}^{t^n} (u - \mathbb{P}_n^{loc}u, v_h)_{H(.)} = 0, \qquad \forall v_h \in \mathcal{P}_{k-1}(t^{n-1}, t^n; U_h^n).$$

Here we have used the convention $(\mathbb{P}_n^{loc}u)^n \equiv (\mathbb{P}_n^{loc}u)(t^n)$.

(2) The projection $\mathbb{P}_h^{loc} : C[0, T; H(.)] \to \mathcal{U}_h$ satisfies

$$\mathbb{P}_h^{loc}u \in \mathcal{U}_h \qquad \text{and} \qquad (\mathbb{P}_h^{loc}u)|_{(t^{n-1},t^n]} = \mathbb{P}_n^{loc}(u|_{[t^{n-1},t^n]}).$$

(3) $\mathbb{P}_h : \{u \in C[0, T; H(.)] \mid (Mu) \in H^1[0, T; U'(.)]\} \to \mathcal{U}_h$ is the discontinuous Galerkin solution of (4.1) with $f = u_t$.

REMARK 1. *Notice that $\mathbb{P}_n^{loc}u$ is the solution of the DG approximation of $(M(.)u)' = f$ on $(t^{n-1}, t^n]$ with $u(t^{n-1})$ specified as the initial data. It follows that $\sup_{t^{n-1} \le t \le t^n} |u - \mathbb{P}_n^{loc}u|_{H(.)}$ (or related norms) measure the local truncation error of the scheme.*

The following theorem compares the (global) solution, $\mathbb{P}_h u$, of the DG scheme with the local solutions, $\mathbb{P}_n^{loc}u$.

5.1. DG Scheme. To approximate the solution of the weak formulation (2.2) we introduce a partition $0 = t^0 < t^1 < \dots < t^N = T$ of $[0,T]$ and on each partition construct a closed subspace $U_h^n \subset U$. The discontinuous Galerkin approximates the solution of (2.2) on $(t^{n-1}, t^n]$ by $u_h \in \mathcal{P}_k(t^{n-1}, t^n; U_h^n)$ satisfying

$$\begin{aligned}(u^n, v^n)_{H(t^n)} &+ \int_{t^{n-1}}^{t^n} \Big(-(u_h, v_{ht})_{H(t)} + a(.; u_h, v_h) \Big) \qquad (5.1)\\ &- (u^{n-1}, v_+^{n-1})_{H(t^{n-1})} = \int_{t^{n-1}}^{t^n} \langle F, v_h \rangle \quad \forall v_h \in \mathcal{P}_k(t^{n-1}, t^n; U_h^n).\end{aligned}$$

The stability and error estimates are established using very similar arguments; for this reason we will just focus on the error estimate. The Galerkin orthogonality condition shows that the error $e = u - u_h$ satisfies

$$(e^n, v^n)_{H(t^n)} + \int_{t^{n-1}}^{t^n} \Big(-(e, v_{ht})_{H(t)} + a(.; e, v_h) \Big) - (e^{n-1}, v_+^{n-1})_{H(t^{n-1})} = 0, \qquad (5.2)$$

for all $v_h \in \mathcal{P}_k(t^{n-1}, t^n; U_h^n)$. We decompose the error as $e = e_p + e_h \equiv (u - \mathbb{P}_h u) + (\mathbb{P}_h u - u_h)$, where $\mathbb{P}_h : \{u \in C[0,T; H(.)] \mid M(.)u \in H^1[0,T; U'(.)]\} \to \mathcal{U}_h$ is the projection introduced in Definition 4.2. The orthogonality condition (5.2) becomes

$$\begin{aligned}(e_h^n, v^n)_{H(t^n)} &+ \int_{t^{n-1}}^{t^n} \Big(-(e_h, v_{ht})_{H(.)} + a(.; e_h, v_h) \Big) - (e_h^{n-1}, v_+^{n-1})_{H(t^{n-1})}\\ &= -(e_p^n, v^n)_{H(t^n)} + \int_{t^{n-1}}^{t^n} (e_p, v_{ht})_{H(.)} + (e_p^{n-1}, v_+^{n-1})_{H(t^{n-1})}\\ &\quad - \int_{t^{n-1}}^{t^n} a(.; e_p, v_h).\end{aligned}$$

By construction, $\mathbb{P}_h u$ is the discontinuous Galerkin approximation of the (4.1), so e_p satisfies the orthogonality condition (4.3). It follows that the first three terms of right hand side vanish so that

$$(e_h^n, v^n)_{H(t^n)} + \int_{t^{n-1}}^{t^n} (-(e_h, v_{ht})_{H(.)} + a(.; e_h, v_h)) - (e_h^{n-1}, v_+^{n-1})_{H(t^{n-1})} = -\int_{t^{n-1}}^{t^n} a(\cdot; e_p, v_h). \qquad (5.3)$$

A preliminary estimate satisfied by the error at the partition points is obtained by setting $v_h = e_h$ and using the coercivity of $a(.;.,.)$

$$\begin{aligned}\frac{1}{2}|e_h^n|^2_{H(t^n)} &+ c_\alpha \int_{t^{n-1}}^{t^n} |e_h|^2_{U(.)} + \frac{1}{2}|e_{h+}^{n-1} - e_h^{n-1}|^2_{H(t^{n-1})} \qquad (5.4)\\ &\leq \frac{1}{2}|e_h^{n-1}|^2_{H(t^{n-1})} + \int_{t^{n-1}}^{t^n} \left(\mu(\cdot; e_h, e_h) - a(.; e_p, e_h) + C_\alpha |e_h|^2_{H(.)} \right)\end{aligned}$$

Using Assumptions 2 on the continuity of $a(.;.,.)$ and $\mu(.;.,.)$ we obtain

$$\begin{aligned}|e_h^n|^2_{H(t^n)} &+ c_\alpha \int_{t^{n-1}}^{t^n} |e_h|^2_{U(.)} + |e_h^{n-1} - e_{h+}^{n-1}|^2_{H(t^{n-1})} \leq |e_h^{n-1}|^2_{H(t^{n-1})} \qquad (5.5)\\ &+ \int_{t^{n-1}}^{t^n} \Big((1 + c_a/c_\alpha) \Big(c_a |e_p|^2_{U(.)} + C_a |e_p|^2_{H(.)} \Big) + (2C_\alpha + 2C_\mu + C_a) |e_h|^2_{H(.)} \Big).\end{aligned}$$

Recalling that $(e_{ht}, e_h)_{H(.)} = (1/2)(d/dt)|e_h|^2_{H(.)} - (1/2)\mu(.; e_h, e_h)$ we obtain

$$\frac{1}{2}|e_h(t)|^2_{H(t)} + \frac{1}{2}|e_h^{n-1} - e_{h+}^{n-1}|^2_{H(t^{n-1})} - \frac{1}{2}|e_h^{n-1}|^2_{H(t^{n-1})}$$
$$= \int_{t^{n-1}}^{t} \frac{1}{2}\mu(.; e_h, e_h) - \int_{t^{n-1}}^{t^n} \Big(a(.; e_p, \tilde{e}_h) + a(.; e_h, \tilde{e}_h) + \mu(., e_h, \tilde{e}_h)\Big)$$

Estimating the right hand using Lemma 3.2 and Corollary 3.4, as in the derivation of equation (5.5), we obtain

$$\begin{aligned} |e_h(t)|^2_{H(t^n)} + |e_h^{n-1} - e_{h+}^{n-1}|^2_{H(t^{n-1})} &\leq |e_h^{n-1}|^2_{H(t^{n-1})} \\ &+ C(k,\mu)\int_{t^{n-1}}^{t^n} \Big((1 + c_a/c_\alpha)(c_a|e_p|^2_{U(.)} + C_a|e_p|^2_{H(.)}) \\ &+ c_\alpha C\big(C_u, c_u, c_a/c_\alpha\big)|e_h|^2_{U(.)} + C(...)|e_h|^2_{H(.)}\Big). \end{aligned} \tag{5.6}$$

Here $C(...)$ is a constant depending upon $C\big(C_a, C_\alpha, C_\mu, \sqrt{c_a}\tau C_{inv(h)}, c_a/c_\alpha\big)$. The remainder of the proof parallels the the proof of [5, Theorem 3.1]. Specifically we construct the convex combination of $(1-\lambda)$ equation (5.5) and λ of equation (5.6) and choose the coefficient, λ, so that the term involving $|e_h(t)|^2_{U(.)}$ on the right hand side of (5.6) is dominated by the corresponding term on the left of (5.5). Setting

$$\lambda C\big(C_u, c_u, c_a/c_\alpha\big) = (1/2)(1-\lambda), \qquad \text{or} \qquad \lambda = \frac{1}{1 + 2C\big(C_u, c_u, c_a/c_\alpha\big)}.$$

leads to an estimate of the form:

$$(1-\lambda)|e_h^n|^2_{H(t^n)} + \lambda|e_h(t)|^2_{H(.)} + (1-\lambda)\frac{c_\alpha}{2}\int_{t^{n-1}}^{t^n} \|e_h\|^2_{U(.)} + |e_h^{n-1} - e_{h+}^{n-1}|^2_{H(t^{n-1})} \tag{5.7}$$
$$\leq |e_h^{n-1}|^2_{H(t^{n-1})} + C(...)\lambda\int_{t^{n-1}}^{t^n} \Big(c_a\|e_p\|^2_{U(.)} + C_a|e_p|^2_{H(.)} + |e_h|^2_{H(.)}\Big)$$

Bound the first and last terms on the right by

$$|e_h^{n-1}|^2_{H(t^n)} \leq (1-\lambda)|e_h^{n-1}|^2_{H(t^{n-1})} + \lambda \sup_{t^{n-2} < s \leq t^{n-1}} |e_h(s)|^2_{H(s)}$$

and

$$\int_{t^{n-1}}^{t^n} |e_h|^2_{H(.)} \leq \tau^n \sup_{t^{n-1} < s \leq t^n} |e_h(s)|^2_{H(s)}, \qquad \tau^n \equiv t^n - t^{n-1},$$

respectively, and select the time t on the left so that $|e_h(t)|_{H(s)} = \sup_{t^{n-1} < s \leq t^n} |e_h(s)|_{H(s)}$ to get

$$\begin{aligned} (1-\lambda)|e_h^n|^2_{H(t^n)} + \lambda(1 - C(...)\tau^n) &\sup_{t^{n-1} < s \leq t^n} |e_h(t)|^2_{H(s)} \\ &+(1-\lambda)\frac{c_\alpha}{2}\int_{t^{n-1}}^{t^n} \|e_h\|^2_{H(.)} + |e_h^{n-1} - e_{h+}^{n-1}|^2_{H(t^{n-1})} \\ \leq (1-\lambda)|e_h^{n-1}|^2_{H(t^{n-1})} &+ \lambda \sup_{t^{n-2} < s \leq t^{n-1}} |e_h(s)|^2_{H(s)} \\ &+C(...)\lambda\int_{t^{n-1}}^{t^n} (c_a\|e_p\|^2_{U(.)} + C_a|e_p|^2_{H(.)}). \end{aligned}$$

[7] J. DOUGLAS JR. AND T. F. RUSSEL, Numerical methods for convection dominated difussion problems based on combining the method of characteristics with finite element or finite difference procedures, *SIAM J. Numer. Anal.*, **19**, 1982, pp 871-885.

[8] T. DUPONT, Mesh modification for evolutionary equations, *Math. Comp.*, **39**, 1982, pp 85-107.

[9] T. F. DUPONT AND YINGJIE LIU, Symmetric error estimates for moving mesh Galerkin methods for advection-diffusion equations, *SIAM J. Numer. Anal.* **40**, 2002, pp 914-927.

[10] K. ERIKSSON AND C. JOHNSON, Adaptive finite element methods for parabolic problems. I. A linear model problem, *SIAM J. Numer. Anal.*, **28**, 1991, pp. 43-77.

[11] K. ERIKSSON AND C. JOHNSON, Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty(L^2)$ and $L_\infty(L_\infty)$, *SIAM J. Numer. Anal.*, **32**, 1995, pp. 706-740.

[12] K. ERIKSSON, C. JOHNSON AND V. THOMÉE, Time discretization of parabolic problems by the discontinuous Galerkin method, *RAIRO Modél. Math. Anal. Numér.*, **29**, 1985, pp 611-643.

[13] D. ESTEP AND S. LARSSON, The discontinuous Galerkin method for semilinear parabolic equations, *RAIRO Modél. Math. Anal. Numér.*, **27**, 1993, pp 35-54.

[14] P. JAMET, Galerkin-type approximations which are discontinuous in time for parabolic equations in a varible domain, *SIAM J. Numer. Anal.* **15**, 1978, pp 912-928.

[15] C. JOHNSON, Numerical Solution of Partial Differential Equations by the Finite Element Method, Cambridge, 1987.

[16] P. LASAINT AND P.-A. RAVIART, On a finite element method for solving the neutron transport equation, in Mathematical aspects of finite elements in partial differential equations, C. de Boor, ed., Academic Press, New York, 1974, pp. 89-123.

[17] S. LARSSON, V. THOMÉE AND L.B. WALHBIN, Numerical solution of parabolic intregro-differential equations by the discontinuous Galerkin method, *Math. Comp.* **67**, 1998, pp 45-71.

[18] YINGJIE LIU, R.E. BANK, T.F DUPONT, S. GARCIA AND R.P. SANTOS, Symmetric error estimates for moving mesh mixed methods for advection-diffusion equations, *SIAM J. Numer. Anal.* **40**, 2003, pp 2270-2291.

[19] M. LUSKIN AND R. RANNACHER, On the smoothing property of the Galerkin method for parabolic equations, *SIAM J. Numer. Anal.* **19**, 1982, pp 93-113.

[20] K. MILLER, Moving finite elements II, *SIAM J. Numer. Anal.* **18**, 1981, pp 1033-1057.

[21] K. MILLER AND R.N. MILLER, Moving finite elements I, *SIAM J. Numer. Anal.* **18**, 1981, pp 1019-1032.

[22] K.W. MORTON, A. PRIESTLEY AND E. SÜLI, Stability of the Lagrange-Galerkin method with nonexact integration, *RAIRO Modél. Math. Anal. Numér.*, **22**, 1988, pp 625-653.

[23] SHOWALTER, R. E., *Hilbert Space Methods for Partial Differential Equations*, Pitman, 1979. Available at `http://ejde.math.txstate.edu/`

[24] SHOWALTER, R. E., *Monotone operators in Banach space and nonlinear partial differential equations*, American Mathematical Society, Providence, RI, 1997.

[25] V. THOMÉE, Galerkin finite element methods for parabolic problems, Spinger-Verlag, Berlin, 1997

[26] N. J. WALKINGTON, Convergence of the discontinuous Galerkin method for discontinuous solutions, *submitted* SINUM 1/03.