Dynamic Finite Element Methods for Second Order Parabolic Equations

Daoqi Yang*

June 16, 1997

Abstract

Dynamic finite element schemes are analyzed for second order parabolic problems. These schemes can employ different finite element spaces at different time levels in order to capture time-changing localized phenomena, such as moving sharp fronts or layers. The dynamically changing grids and interpolation polynomials are necessary and essential to many large-scale transient problems. Standard, characteristic, and mixed finite element methods with dynamic function spaces are considered for linear and nonlinear problems. The convergence results obtained in this paper are optimal and better than those published previously.

1 Introduction

Many time-dependent problems involve localized phenomena, such as sharp fronts, shocks, and layers, which also change with time. The numerical simulation of these problems using the finite element method requires capabilities for efficient, dynamic, and self-adaptive local grid refinement or unrefinement and interpolation polynomial modifications.

The object of this paper is to analyze a number of numerical schemes for parabolic problems which allow one to use different grids and different interpolation polynomials at different time levels when necessary. For many problems, such as oil reservoir and semiconductor simulation, the solution is rough in a very small region of the physical domain, but the region of roughness sweeps out a substantial part of the domain as time goes on. Thus a static (fixed with time) grid finite element method would require very fine grid over the entire domain and is often too expensive in practice. On the other hand, dynamic (changing with time) finite elements would provide great computational flexibility and efficiency, where local grid refinement and interpolation polynomial modification can be made dynamically with the changing location of singularities. With the popularity of the p and hp version finite element methods, the order of interpolation polynomials can also be adapted locally in space and dynamically in time according to the behavior of the solution. It is proved theoretically and shown experimentally that some singularities can be resolved not by just refining local grids, but by increasing the

*Institute for Mathematics and its Applications, University of Minnesota, 514 Vincent Hall, 206 SE Church Street, Minneapolis, MN 55455-0436. Email: yang@math.wayne.edu, yang@ima.umn.edu

order of approximation polynomials. A frequently encountered example is nonsmooth initial data parabolic problems. At the beginning the solution is not smooth, fine grids (i.e. the h version) and piecewise linear interpolation polynomials may be applied. After a while the solution becomes smooth, we may use coarse grid and higher-order basis functions (i.e. the p version). For general problems, hp version finite elements may be applied to improve efficiency and accuracy; see Babuska and Guo [2], Babuska and Suri [3], Schwab and Suri [39], and Suri [40]. If the grid and basis functions are chosen at each time level in accordance with the changing character of the solution at that time, then the dynamic finite element methods have the capability for accurately and efficiently resolving time-changing phenomena. For simplicity, however, we will consider only the h version in our analysis. The p and hp versions can be treated analogously.

Dynamic finite element schemes under the name of discontinuous Galerkin or space-time finite element methods have been discussed in Bank and Santos [6], Eriksson, Estep, Hansbo, and Johnson [18], Eriksson and Johnson [19, 20, 21, 22], Hulbert and Hughes [24], Johnson [28, 29], Jaffre, Johnson, and Szepessy [25] for model linear and nonlinear evolution problems. These methods can provide a-posteriori error estimates and adaptivity based on local grid refinement at different time levels. The error analysis obtained so far is not optimal (there is a logrithmic factor of the time step size contained in the error estimates; see Johnson [28] and Eriksson and Johnson [19]), depends on some strong stability estimates for the discrete dual problem, and imposes some restrictions on the time and space grids. For example, the error estimates in Hulbert and Hughes [24], and Johnson [29] are non-optimal in the sense that they contain the factor Δt^{-1}, where Δt is the time step size. The finite element spaces in [19] are required to satisfy that $S_n \subset S_{n-1}$, where S_n is the finite elment space at time level $t = t_n$, and the space and time grids are required to satisfy that $h_n^2 \leq C\Delta t_n$, where h_n and Δt_n are the space and time grid sizes at time $t = t_n$, respectively and C is a constant. Generalization of the estimates in Eriksson and Johnson [19, 20, 21, 22] to nonstandard finite elements (e.g. characteristic and mixed finite elements) has not been seen, except for Bank and Santos [6] where a space-time moving finite element method with discretization along characteristics was treated.

Moving finite element method (see Baines [4, 5], Lucier [33], Miller [34], and Miller and Miller [35]) is another class of such methods which provide dynamic change of grids according to the moving local phenomena. A unique feature of moving finite elements is the inclusion of grid point movement in weak forms or in the minimization of the residual of the differential equations. That is, the position of grid points and the approximate solution at these points are solved simultaneously for each time level in such a way that the weighted residual of the differential equation, possibly with a penalty term, is minimized. This method offers a good way of solving certain kind of problems, but employ essentially the same number of grid points at all time levels and has great difficulties tackling three dimensional problems.

A third class of dynamic finite element schemes was mainly analyzed at the theoretical level (see Dupont [16], Liang [30, 31], Liang and Chen [32], Rui [37], Yang [42, 43, 44, 45, 46, 47], and Yuan [50, 51]), although numerical experiments were given in Yang [47, 48] based on domain decomposition and finite element discretization at each time level. The idea is to follow the traditional finite differencing in time and finite elment discretization in space (see Douglas and Dupont [14] and Wheeler [41]). However, since we are applying different finite element spaces at different times, the finite differencing in time is achieved by first projecting the solution from the previous time level into the finite element space at the current time level, and then using it as initial value to compute the approximate solution at current time level. The projection is used for convergence analysis and may not need to be actually computed for some of the schemes in implementation. Relatively optimal convergence results were derived in most of the papers above. However, the error estimates contain as a factor the number of different finite element spaces applied up to the current time level, which could be as large as Δt^{-1}, where Δt is the time step size. Thus these error estimates are not quite optimal in the case of changing the finite element spaces frequently.

In this paper, we will consider some dynamic finite element schemes which may be categorized into the third class as defined above. We will derive optimal convergence estimates for general (variable and nonlinear coefficients with first order terms) parabolic problems and for general (standard, characteristic, mixed) finite element approximations. The number of different finite element spaces applied up to the current time level will disappear as a factor in the error estimates. Our error estimation consists of three parts: an optimal time finite difference discretization error, an optimal spatial finite element discretization error, and an optimal error term due to the projections of the approximated solution from old finite element spaces onto new finite element spaces. Note that the third error term due to projection in [30, 31, 32, 37, 42, 43, 44, 45, 46, 47, 50, 51] are not optimal.

We now introduce some notation which we will use throughout the paper. Let Ω denote a spatial domain in R^d with a piecewise uniformly Lipschitz boundary Γ. Here d is a positive integer. Denote by $H^m(\Omega) = W^{m,2}(\Omega)$ and $W^{m,p}(\Omega)$ the standard Sobolev spaces on Ω, with norms $\|\cdot\|_m$ and $\|\cdot\|_{m,p}$, respectively. Let $L^p(\Omega)$, $p = 2, \infty$, denote the standard Banach spaces, with $\|\cdot\|$ denoting the L^2 norm and $\|\cdot\|_\infty$ the L^∞ norm over Ω. However, for a positive function ϕ, we use $\|\cdot\|_\phi$ to denote the weighted L^2-norm with weight function ϕ. For a normed linear space Q with norm $\|\cdot\|_Q$ and a sufficient regular function $g : [t_1, t_2] \to Q$, we define

$$\|g\|_{L^p([t_1,t_2];Q)} = \left(\int_{t_1}^{t_2} \|g(\cdot,t)\|_Q^p dt\right)^{1/p}, \quad p = 1, 2, \infty,$$

with standard modification for $p = \infty$, where $[t_1, t_2] \subset [0, T]$ is a time interval, and T is a positive number. We omit $[t_1, t_2]$ from the notation when $[t_1, t_2] = [0, T]$, that is, we write $\|g\|_{L^p(Q)}$ instead of $\|g\|_{L^p([0,T];Q)}$.

We partition the time interval $[0, T]$ into $0 = t_0 < t_1 < \cdots < t_N = T$, and denote $\Delta t_n = t_n - t_{n-1}$. We will also use capital letter C, without subscripts, to denote a generic positive real constant, which may take on different values in different occurrences.

The organization of the paper is as follows. In §2 we give our approximation scheme and prove some convergence results for linear problems. In §3 we analyze the method and make error estimates for nonlinear problems. Then in §4, we consider the modified method of characteristics, and in §5, we treat mixed finite element methods. Finally in §6, we give some concluding remarks.

2 Linear Problems

Consider the following linear parabolic problem with Dirichlet boundary condition: find $u(x,t)$ satisfying

$$\phi(x)\frac{\partial u}{\partial t} - \nabla\cdot(a(x,t)\nabla u) + b(x,t)\cdot\nabla u + c(x,t)u = f(x,t), \quad x\in\Omega, \quad t\in(0,T], \tag{1}$$

$$u(x,t) = 0, \qquad x\in\Gamma, \quad t\in(0,T], \tag{2}$$

$$u(x,0) = g(x), \qquad x\in\Omega, \tag{3}$$

where f, g, a, b, c and ϕ are known real-valued functions. It is assumed that ϕ and a are bounded below and above by positive constants, and that b and its componentwise gradient are bounded from above by positive constants.

Our numerical method will allow us to apply different finite element spaces at different times in order to capture moving local phenomena. For $n = 0, 1, 2, \cdots, N$, let S_n be a finite element space of $H_0^1(\Omega)$ with grid parameter h_n, and interpolation polynomials of degree k_n. We assume that the following approximation property holds: for $n = 1, 2, \cdots, N$,

$$\inf_{w\in S_n}(\|v-w\| + h_n\|v-w\|_1) \le Ch_n^s\|v\|_s, \quad 0\le s\le k_n+1, \quad \forall v\in H_0^1(\Omega), \tag{4}$$

where we assume that C is a constant independent v, n, h_n. Since the element sizes can be very different, a better form for the right hand side of (4) would be to express it as a sum of contributions from individual elements, like (3.1) in Yang [46].

We first define the implicit Euler scheme. Suppose that $U_0 \in M_0$ is an initial approximation of $u(\cdot, 0)$, we define our first dynamic finite element scheme as follows:

Algorithm 2.1 *For $n = 1, 2, \cdots, N$, first compute the weighted L^2 projection $\widehat{U}_{n-1} \in S_n$ by solving*

$$(\phi(\widehat{U}_{n-1} - U_{n-1}), v) = 0, \qquad \forall v \in S_n; \tag{5}$$

then compute $U_n \in S_n$ by

$$(\phi\frac{U_n - \widehat{U}_{n-1}}{\Delta t_n}, v) + (a_n\nabla U_n, \nabla v) + (b_n\cdot\nabla U_n + c_nU_n, v) = (f_n, v), \;\; \forall v\in S_n, \tag{6}$$

where $(f,g) = \int_\Omega f\cdot g dx$, and $\xi_n = \xi(x, t_n)$ for any function ξ.

Some remarks about the scheme (5)-(6) are in order. Equation (5) gives a weighted L^2 projection $\widehat{U}_{n-1}$ of the previous approximate solution U_{n-1} into the current finite element space S_n when different finite element spaces are used at times $t = t_n$ and $t = t_{n-1}$. This projection is used in (6) as initial value to calculate U_n, the approximate solution at $t = t_n$. Note that when the finite element space remains unchanged for all time levels, the scheme (5)-(6) reduces to the standard one [14, 41]. Note that Algorithm 2.1 is very similar to the space-time finite element scheme in Eriksson and Johnson [19] with piecewise constant polynomials in time, and is the same as a scheme considered in Dupont [16] and Liang [30].

The Crank-Nicolson scheme can be defined in the standard way.

Algorithm 2.2 *For $n = 1, 2, \cdots, N$, first compute the weighted L^2 projection $\widehat{U}_{n-1} \in S_n$ by solving*

$$(\phi(\widehat{U}_{n-1} - U_{n-1}), v) = 0, \qquad \forall v \in S_n; \tag{7}$$

then compute $U_n \in S_n$ by

$$\begin{aligned}(\phi\frac{U_n - \widehat{U}_{n-1}}{\Delta t_n}, v) + (a_{n-1/2}\nabla\frac{U_n + \widehat{U}_{n-1}}{2}, \nabla v) + (b_{n-1/2}\nabla\frac{U_n + \widehat{U}_{n-1}}{2} + c_{n-1/2}\frac{U_n + \widehat{U}_{n-1}}{2}, v)\\ = (f_{n-1/2}, v), \ \ \forall v \in S_n,\end{aligned} \tag{8}$$

where $\xi_{n-1/2} = \xi(x, t_{n-1/2})$ for any function ξ, and $t_{n-1/2} = \dfrac{1}{2}(t_n + t_{n-1})$.

We now state and prove the following convergence estimates for the implicit Euler scheme.

Theorem 2.1 *Suppose that the solution u to problem (1)-(3) is sufficiently regular. Let U_n be the solution of Scheme (5)-(6). Then we have the error estimates for $m = 1, 2 \cdots, N$,*

$$\begin{aligned}&\max_{1\le n\le m} \|u_n - U_n\|\\ &\le C\Big\{\|u_0 - U_0\| + \sum_{n=1}^{m}\Big[h_n^{k_n+1}\|u_t\|_{L^1([t_{n-1},t_n];H^{k_n+1}(\Omega))} + \Delta t_n\|u_{tt}\|_{L^1([t_{n-1},t_n];L^2(\Omega))}\Big]\\ &+ \sum_{n=1}^{m}\Delta t_n^{1/2}h_n^{k_n+1}\|u(\cdot,t_n)\|_{k_n+1} + \max_{0\le n\le m} h_n^{k_n+1}\|u(\cdot,t_n)\|_{k_n+1}\\ &+ \sum_{n=1}^{m}\delta_n\Big[h_n^{k_n+1}\|u(\cdot,t_{n-1})\|_{k_n+1} + h_{n-1}^{k_{n-1}+1}\|u(\cdot,t_{n-1})\|_{k_{n-1}+1}\Big]\Big\},\end{aligned} \tag{9}$$

$$\begin{aligned}&\Big(\sum_{n=1}^{m}\Delta t_n[(a_n\nabla(u_n - U_n), \nabla(u_n - U_n)) + (c_n(u_n - U_n), u_n - U_n)]\Big)^{1/2}\\ &\le C\Big\{\|u_0 - U_0\| + \sum_{n=1}^{m}\Big[h_n^{k_n+1}\|u_t\|_{L^1([t_{n-1},t_n];H^{k_n+1}(\Omega))} + \Delta t_n\|u_{tt}\|_{L^1([t_{n-1},t_n];L^2(\Omega))}\Big]\\ &+ \sum_{n=1}^{m}\Delta t_n^{1/2}h_n^{k_n}\|u(\cdot,t_n)\|_{k_n+1}\\ &+ \sum_{n=1}^{m}\delta_n\Big[h_n^{k_n+1}\|u(\cdot,t_{n-1})\|_{k_n+1} + h_{n-1}^{k_{n-1}+1}\|u(\cdot,t_{n-1})\|_{k_{n-1}+1}\Big]\Big\},\end{aligned} \tag{10}$$

where $\delta_n = 0$ if $S_n = S_{n-1}$ and $\delta_n = 1$ otherwise.

Proof: We will make use of the elliptic projection $R_n u$ of u: find $R_n u(x,t) \in S_n$ for each $t \in [0, T]$ such that

$$(a_n\nabla(u - R_n u)(\cdot,t), \nabla v) + (c_n(u - R_n u)(\cdot,t), v) = 0, \qquad \forall v \in S_n. \tag{11}$$

Suppose that the triangulation is regular at each time level, and that all finite elements are affine, we can prove (see Ciarlet [11] and Brenner and Scott [7]) that there exists a constant C independent of h_n (and k_n, provided that polynomials of degree fewer than 20, say, are used) such that

$$\|u - R_n u\| + h_n\|u - R_n u\|_1 \le C h_n^{k_n+1}\|u\|_{k_n+1}, \qquad \forall t \in [0, T]. \tag{12}$$

We introduce the following notation.

$$\begin{aligned} e_n &= U_n - R_n u_n, & \hat{e}_{n-1} &= \widehat{U}_{n-1} - R_n u_{n-1}, \\ r_n &= u_n - R_n u_n, & \hat{r}_{n-1} &= u_{n-1} - R_n u_{n-1}. \end{aligned}$$

Notice that the exact solution u satisfies

$$\begin{aligned} &(\phi \frac{u_n - u_{n-1}}{\Delta t_n}, v) + (a_n \nabla u_n, \nabla v) + (b_n \nabla u_n + c_n u_n, v) \\ &= (f_n, v) + (\phi \rho_n, v), \qquad \forall v \in H_0^1(\Omega), \end{aligned} \tag{13}$$

where

$$\|\rho_n\| = \| \frac{u_n - u_{n-1}}{\Delta t_n} - \frac{\partial u}{\partial t}(t_n)\| \le \int_{t_{n-1}}^{t_n} \|u_{tt}\| dt. \tag{14}$$

Subtracting (13) from (6) and using (11) yield

$$\begin{aligned} &(\phi \frac{e_n - \hat{e}_{n-1}}{\Delta t_n}, v) + (a_n \nabla e_n, \nabla v) + (c_n e_n, v) \\ &= (\phi (\frac{r_n - \hat{r}_{n-1}}{\Delta t_n} - \rho_n), v) + (b_n \cdot \nabla (u_n - U_n), v), \qquad \forall v \in S_n. \end{aligned} \tag{15}$$

Letting $v = e_n$ in (15) we obtain the error equation

$$\begin{aligned} &(\phi \frac{e_n - \hat{e}_{n-1}}{\Delta t_n}, e_n) + (a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n) \\ &= (\phi (\frac{r_n - \hat{r}_{n-1}}{\Delta t_n} - \rho_n), e_n) + (b_n \cdot \nabla (u_n - U_n), e_n). \end{aligned} \tag{16}$$

It is easy to see that

$$(\phi \frac{e_n - \hat{e}_{n-1}}{\Delta t_n}, e_n) = \frac{\|e_n\|_\phi^2 - \|\hat{e}_{n-1}\|_\phi^2}{2\Delta t_n} + \frac{1}{2\Delta t_n} \|e_n - \hat{e}_{n-1}\|_\phi^2, \tag{17}$$

and

$$\begin{aligned} \|r_n - \hat{r}_{n-1}\| &= \|(I - R_n)(u_n - u_{n-1})\| \\ &\le C h_n^{k_n+1} \|u_n - u_{n-1}\|_{k_n+1} \\ &\le C h_n^{k_n+1} \| \int_{t_{n-1}}^{t_n} u_t d\tau \|_{k_n+1} \\ &\le C h_n^{k_n+1} \int_{t_{n-1}}^{t_n} \|u_t\|_{k_n+1} dt. \end{aligned} \tag{18}$$

Applying integration by parts and the ϵ-inequality, we have

$$\begin{aligned} &(b_n \cdot \nabla (u_n - U_n), e_n) = (b_n \cdot \nabla r_n, e_n) - (b_n \cdot \nabla e_n, e_n) \\ &= -(\nabla \cdot (e_n b_n), r_n) + \int_\Gamma e_n r_n b_n \cdot \nu ds - (b_n \cdot \nabla e_n, e_n) \\ &= -(\nabla \cdot (e_n b_n), r_n) - (b_n \cdot \nabla e_n, e_n) \\ &\le \frac{1}{2}(a_n \nabla e_n, \nabla e_n) + C(\|e_n\|_\phi^2 + \|r_n\|^2), \end{aligned} \tag{19}$$

where ν is the unit outward normal of Γ.

Combining (16), (17), (18), (19), and (14), we have the following error inequality

$$\begin{aligned}&\|e_n\|_\phi^2 - \|\hat{e}_{n-1}\|_\phi^2 + \Delta t_n[(a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n)] \\ &\le C[F_n \|e_n\|_\phi + \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2)],\end{aligned} \tag{20}$$

where

$$F_n = h_n^{k_n+1} \int_{t_{n-1}}^{t_n} \|u_t\|_{k_n+1} dt + \Delta t_n \int_{t_{n-1}}^{t_n} \|u_{tt}\| dt. \tag{21}$$

We now find the relationship between $\|e_n\|_\phi$ and $\|\hat{e}_n\|_\phi$. Equation (5) implies that

$$(\phi(\hat{e}_{n-1} - e_{n-1}), v) = (\phi(\hat{r}_{n-1} - r_{n-1}), v), \quad \forall v \in S_n.$$

Choosing $v = \hat{e}_{n-1}$ and using Schwarz's inequality, we have

$$\|\hat{e}_{n-1}\|_\phi^2 - \|e_{n-1}\|_\phi^2 \le 2\|\hat{e}_{n-1}\|_\phi \|\hat{r}_{n-1} - r_{n-1}\|_\phi \le \frac{1}{2}\|\hat{e}_{n-1}\|_\phi^2 + 2\|\hat{r}_{n-1} - r_{n-1}\|_\phi^2,$$

from which we derive that

$$\|\hat{e}_{n-1}\|_\phi^2 \le 2\|e_{n-1}\|_\phi^2 + 4\|\hat{r}_{n-1} - r_{n-1}\|_\phi^2, \tag{22}$$

and

$$\|\hat{e}_{n-1}\|_\phi^2 - \|e_{n-1}\|_\phi^2 \le 2\sqrt{2}\|e_{n-1}\|_\phi \, \|\hat{r}_{n-1} - r_{n-1}\|_\phi + 4\|\hat{r}_{n-1} - r_{n-1}\|_\phi^2. \tag{23}$$

Combining (20) and (23) we see that

$$\begin{aligned}&\|e_n\|_\phi^2 - \|e_{n-1}\|_\phi^2 + \Delta t_n[(a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n)] \\ &\le C[F_n \|e_n\|_\phi + \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2) + \|r_{n-1} - \hat{r}_{n-1}\|_\phi \, \|e_{n-1}\|_\phi + \|r_{n-1} - \hat{r}_{n-1}\|_\phi^2].\end{aligned} \tag{24}$$

Summing (24) from $n = 1$ to $n = m$ $(1 \le m \le N)$, we obtain

$$\begin{aligned}&\max_{1\le n\le m} \|e_n\|_\phi^2 - \|e_0\|_\phi^2 + \sum_{n=1}^m \Delta t_n[(a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n)] \\ &\le C \sum_{n=1}^m \Big(F_n \|e_n\|_\phi + \|r_{n-1} - \hat{r}_{n-1}\|_\phi \|e_{n-1}\|_\phi + \|r_{n-1} - \hat{r}_{n-1}\|_\phi^2 + \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2) \Big) \\ &\le C \Big\{ \max_{0\le n\le m} \|e_n\|_\phi \sum_{n=1}^m (F_n + \|r_{n-1} - \hat{r}_{n-1}\|_\phi) + \sum_{n=1}^m (\|r_{n-1} - \hat{r}_{n-1}\|_\phi^2 + \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2)) \Big\} \\ &\le \frac{1}{2} \max_{0\le n\le m} \|e_n\|_\phi^2 + C \Big\{ \Big[\sum_{n=1}^m F_n\Big]^2 + \Big[\sum_{n=1}^m \|r_{n-1} - \hat{r}_{n-1}\|_\phi\Big]^2 + \sum_{n=1}^m \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2) \Big\}.\end{aligned}$$

Then

$$\begin{aligned}&\frac{1}{2} \max_{0\le n\le m} \|e_n\|_\phi^2 - \|e_0\|_\phi^2 + \sum_{n=1}^m \Delta t_n[(a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n)] \\ &\le C \sum_{n=1}^m \Delta t_n \|e_n\|_\phi^2 + C \Big\{ \Big[\sum_{n=1}^m F_n\Big]^2 + \Big[\sum_{n=1}^m \|r_{n-1} - \hat{r}_{n-1}\|_\phi\Big]^2 + \sum_{n=1}^m \Delta t_n \|r_n\|^2 \Big\}.\end{aligned}$$

Applying the discrete Gronwall Lemma, triangular inequality, and (12) completes the proof of the theorem. ∎

Corollary 2.1 *When a static finite element space is used for $0 \le t \le T$, i. e., when $h_n = h, k_n = k$ for $n = 0, 1, \cdots, N$, we have the error estimates*

$$\|u_m - U_m\| = O(\Delta t + h^{k+1} + \|u_0 - U_0\|). \tag{25}$$

Thus the implicit Euler algorithm analyzed in Wheeler [41] is the $h_n = h$ and $k_n = k$ case of the scheme (5)-(6). ∎

These results are different from those obtained by Dupont [16] and Bank and Santos [6], in that our error estimates are given in standard norms independent of the finite element grids. Also, the finite element grids in our method are not required to change continuously in any fashion; just a minimum angle property is needed for grids at all time levels. Note that the error estimates in Dupont [16] and Bank and Santos [6] involve grid-dependent norms, are sub-optimal, and impose the assumption that the grid is locally quasi-uniform, while our estimates are in grid independent norms, optimal and require only regular grids. Similar sub-optimal results to [16] were obtained in Jamet [26]. A one-dimensional problem was analyzed in Jamet [27].

The results of Theorem 2.1 also improve those obtained previously [30, 31, 32, 37, 42, 43, 44, 45, 46, 47, 48, 50, 51], in that the number of different finite element spaces applied up to the current time level does not appear as a factor in the error estimates. Note that this number is of the order $O(\Delta t^{-1})$ if the finite element spaces are changed at every time step.

The Crank-Nicolson scheme (Algorithm 2.2) can be analyzed following the steps in the proof of Theorem 2.1. We omit the analysis here and consider a more general θ scheme for nonlinear problems in the next section. Note Dupont [16, page 92] claimed that an analysis of the Crank-Nicolson scheme was not possible in his framework. Thus our analysis here not only gives optimal error estimates, but also provides software implementors with more theoretically guaranteed convergent numerical schemes, on which previous theory had remained silent.

3 Nonlinear Problems

Consider the following quasilinear parabolic problem with Dirichlet boundary condition: find $u(x,t)$ satisfying

$$\phi(x)\frac{\partial u}{\partial t} - \nabla\cdot(a(x,u)\nabla u) + b(x,u)\cdot\nabla u = f(x,u), \quad x \in \Omega, \quad t \in (0,T], \tag{26}$$

$$u(x,t) = 0, \qquad x \in \Gamma, \quad t \in (0,T], \tag{27}$$

$$u(x,0) = g(x), \qquad x \in \Omega, \tag{28}$$

whose weak formulation can be put into the fashion: find a differentiable mapping $u : [0,T] \to H_0^1(\Omega)$ such that

$$(\phi u_t, v) + A(u; u, v) + B(u; u, v) = (f(u), v), \quad \forall v \in H_0^1(\Omega), \tag{29}$$

$$(u(\cdot, 0), v) = (g, v), \quad \forall v \in H_0^1(\Omega), \tag{30}$$

where $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega)$, $u_t = \dfrac{\partial u}{\partial t}$, $f(u) = f(x,u)$, and

$$A(w; u, v) = (a(x,w)\nabla u, \nabla v), \quad B(w; u, v) = (b(x,w)\cdot\nabla u, v). \tag{31}$$

Let S_n be a finite element space at time $t = t_n$ satisfying (4). We make some assumptions on the coefficients:

$$\text{For } (x,v) \in \Omega \times [0,T],\ C_1\|\eta\|^2 \le \sum_{i,j=1}^{d}(a_{ij}(x,v)\eta_i,\eta_j) \le C_2\|\eta\|^2,\ \max_{1\le i\le d}|b_i(x,v)| \le C_2, \tag{32}$$

$$f, b_i, a_{ij}, \text{are uniformly Lipschitz continuous with respect to their } (d+1)\text{-th variable}, \tag{33}$$

$$\text{For } 1 \le i \le d,\ \frac{\partial b_i(x,v)}{\partial x_i} \text{ exists and is bounded from above}, \tag{34}$$

$$f(\cdot,0) \in L^2(\Omega),\ g \in H^{k+1}(\Omega) \cap H_0^1(\Omega), \tag{35}$$

$$u \text{ is unique to (26)-(28), and } u, u_t \in L^\infty(H^{k+1}(\Omega)\cap H_0^1(\Omega)), u_{tt} \in L^1(L^2(\Omega)). \tag{36}$$

where $k = \max_{1\le n\le N} k_n$.

Suppose that $U_0 \in M_0$ is an initial approximation of $u(\cdot,0)$, we define our dynamic finite element scheme with a parameter θ as follows:

Algorithm 3.1 *For $n = 0,1,2,\cdots,N-1$, first compute the weighted L^2 projection $\widehat{U}_n \in S_{n+1}$ by solving*

$$(\phi(\widehat{U}_n - U_n), v) = 0, \qquad \forall v \in S_{n+1}; \tag{37}$$

then compute $U_{n+1} \in S_{n+1}$ by

$$(\phi\frac{U_{n+1}-\widehat{U}_n}{\Delta t_{n+1}}, v) + A(\widehat{U}_{n,\theta};\widehat{U}_{n,\theta},v) + B(\widehat{U}_{n,\theta};\widehat{U}_{n,\theta},v) = (f(\widehat{U}_{n,\theta}),v),\ \ \forall v \in S_{n+1}, \tag{38}$$

where $\widehat{U}_{n,\theta} = \frac{1}{2}(1+\theta)U_{n+1} + \frac{1}{2}(1-\theta)\widehat{U}_n,\ 0 \le \theta \le 1$.

Using Brouwer's fixed-point theorem, we can show that the scheme (37)-(38) has a solution for sufficiently small Δt_n (see Douglas and Dupont [14]). Note that $\theta = 0$ corresponds to the Crank-Nicolson scheme, while $\theta = 1$ corresponds to the implicit Euler scheme.

Theorem 3.1 *Suppose that the solution u to problem (26)-(28) is sufficiently regular. Let U_n be the solution of Scheme (37)-(38). Then we have the error estimates for $m = 1,2\cdots,N$,*

$$\begin{aligned}
&\max_{1\le n\le m}\|u_n - U_n\|^2 \\
&\le C\Big\{\|u_0 - U_0\| + \sum_{n=0}^{m-1}\Big[h_{n+1}^{k_{n+1}+1}\|u_t\|_{L^1([t_n,t_{n+1}];H^{k_{n+1}+1}(\Omega))} + \Delta t_{n+1}\|u_{tt}\|_{L^1([t_n,t_{n+1}];L^2(\Omega))}\Big] \\
&+ \sum_{n=0}^{m-1}\Delta t_{n+1}^{1/2}\Big(h_{n+1}^{k_{n+1}+1}\|u(\cdot,t_{n+1})\|_{k_{n+1}+1} + h_{n+1}^{k_{n+1}+1}\|u(\cdot,t_n)\|_{k_{n+1}+1}\Big) \\
&+ \max_{0\le n\le m-1} h_{n+1}^{k_{n+1}+1}\|u(\cdot,t_{n+1})\|_{k_{n+1}+1} + \sum_{n=0}^{m-1}\Delta t_{n+1}^{3/2}\Big(\|u\|_{L^\infty(H_0^1(\Omega))} + \|u_t\|_{L^\infty(H_0^1(\Omega))}\Big) \\
&+ \sum_{n=0}^{m-1}\delta_{n+1}\Big[h_{n+1}^{k_{n+1}+1}\|u(\cdot,t_n)\|_{k_{n+1}+1} + h_n^{k_n+1}\|u(\cdot,t_n)\|_{k_n+1}\Big]\Big\},
\end{aligned} \tag{39}$$

$$\Big(\sum_{n=0}^{m-1}\Delta t_{n+1}\|u_n - U_n\|_1^2\Big)^{1/2}$$

$$
\begin{aligned}
(40)\qquad &\leq C\Big\{\|u_0-U_0\| + \sum_{n=0}^{m-1}\Big[h_{n+1}^{k_{n+1}+1}\|u_t\|_{L^1([t_n,t_{n+1}];H^{k_{n+1}+1}(\Omega))} + \Delta t_{n+1}\|u_{tt}\|_{L^1([t_n,t_{n+1}];L^2(\Omega))}\Big]\\
&+\sum_{n=0}^{m-1}\Delta t_{n+1}^{1/2}\Big(h_{n+1}^{k_{n+1}+1}\|u(\cdot,t_{n+1})\|_{k_{n+1}+1} + h_{n+1}^{k_{n+1}+1}\|u(\cdot,t_n)\|_{k_{n+1}+1}\Big)\\
&+\sum_{n=0}^{m-1}\Delta t_{n+1}^{1/2}h_{n+1}^{k_{n+1}}\|u(\cdot,t_{n+1})\|_{k_{n+1}+1} + \sum_{n=0}^{m-1}\Delta t_{n+1}^{3/2}\Big(\|u\|_{L^\infty(H_0^1(\Omega))} + \|u_t\|_{L^\infty(H_0^1(\Omega))}\Big)\\
&+\sum_{n=0}^{m-1}\delta_{n+1}\Big[h_{n+1}^{k_{n+1}+1}\|u(\cdot,t_n)\|_{k_{n+1}+1} + h_n^{k_n+1}\|u(\cdot,t_n)\|_{k_n+1}\Big]\Big\},
\end{aligned}
$$

where $\delta_n = 0$ if $S_n = S_{n-1}$ and $\delta_n = 1$ otherwise.

Proof: We define the elliptic projection $R_n u$ of u: find $R_n u(x,t) \in S_n$ for each $t \in [0,T]$ such that

$$
(41)\qquad A(u(\cdot,t);(u-R_nu)(\cdot,t),v) = 0, \qquad \forall v \in S_n.
$$

Then we have

$$
\begin{aligned}
(42)\qquad & \|u-R_nu\| + h_n\|u-R_nu\|_1 \leq Ch_n^{k_n+1}\|u\|_{k_n+1}, \qquad \forall t\in[0,T],\\
(43)\qquad & \|(u-R_nu)_t\| + h_n\|(u-R_nu)_t\|_1 \leq Ch_n^{k_n+1}(\|u\|_{k_n+1}+\|u_t\|_{k_n+1}), \qquad \forall t\in[0,T],\\
(44)\qquad & \|(u-R_nu)_{tt}\|_1 \leq C(\|u\|_{L^\infty(H^1(\Omega))} + \|u_t\|_{L^\infty(H^1(\Omega))} + \|u_{tt}\|_{L^\infty(H^1(\Omega))}), \qquad \forall t\in[0,T].
\end{aligned}
$$

See Wheeler [41] for a proof. We introduce the following notation.

$$
\begin{aligned}
&e_n = U_n - R_nu_n, && \hat{e}_n = \hat{U}_n - R_{n+1}u_n,\\
&r_n = u_n - R_nu_n, && \hat{r}_n = u_n - R_{n+1}u_n,\\
&\hat{e}_{n,\theta} = \frac{1}{2}(1+\theta)e_{n+1} + \frac{1}{2}(1-\theta)\hat{e}_n, && \hat{r}_{n,\theta} = \frac{1}{2}(1+\theta)r_{n+1} + \frac{1}{2}(1-\theta)\hat{r}_n,\\
&t_{n,\theta} = \frac{1}{2}(1+\theta)t_{n+1} + \frac{1}{2}(1-\theta)t_n, && u_{n,\theta} = \frac{1}{2}(1+\theta)u_{n+1} + \frac{1}{2}(1-\theta)u_n.
\end{aligned}
$$

Using (41), the exact solution u of (29)-(30) satisfies

$$
\begin{aligned}
(45)\ &(\phi\frac{R_{n+1}u_{n+1}-R_{n+1}u_n}{\Delta t_{n+1}},v) + A(u_{n,\theta}+\beta_{n,\theta};R_{n+1}u_{n,\theta}+\alpha_{n,\theta},v) + B(u_{n,\theta}+\beta_{n,\theta};u_{n,\theta}+\beta_{n\theta},v)\\
&= -(\phi\frac{r_{n+1}-\hat{r}_n}{\Delta t_{n+1}}+\rho_{n,\theta},v) + (f(u_{n,\theta}+\beta_{n,\theta}),v), \qquad \forall v\in S_{n+1},
\end{aligned}
$$

where

$$
\begin{aligned}
(46)\qquad & \|\rho_{n,\theta}\| \leq \|u_{tt}\|_{L^1([t_n,t_{n+1}];L^2(\Omega))},\\
(47)\qquad & \|\beta_{n,\theta}\| \leq \Delta t_{n+1}\|u_t\|_{L^\infty(L^2(\Omega))},\\
(48)\qquad & \|\alpha_{n,\theta}\|_1 \leq \Delta t_{n+1}\|R_{n+1}u_t\|_{L^\infty(H_0^1(\Omega))}.
\end{aligned}
$$

Subtracting (45) from (38) and taking $v = \hat{e}_{n,\theta}$ we have the error equation

$$
\begin{aligned}
&(\phi \frac{e_{n+1} - \hat{e}_n}{\Delta t_{n+1}}, \hat{e}_{n,\theta}) + A(\widehat{U}_{n,\theta}; \hat{e}_{n,\theta}, \hat{e}_{n,\theta}) \\
&= (\phi \frac{r_{n+1} - \hat{r}_n}{\Delta t_{n+1}} + \rho_{n,\theta}, \hat{e}_{n,\theta}) + (f(\widehat{U}_{n,\theta}) - f(u_{n,\theta} + \beta_{n,\theta}), \hat{e}_{n,\theta}) \\
&+ [A(u_{n,\theta} + \beta_{n,\theta}; R_{n+1} u_{n,\theta} + \alpha_{n,\theta}, \hat{e}_{n,\theta}) - A(\widehat{U}_{n,\theta}; R_{n+1} u_{n,\theta}, \hat{e}_{n,\theta})] \\
&+ [B(u_{n,\theta} + \beta_{n,\theta}; u_{n,\theta} + \beta_{n,\theta}, \hat{e}_{n,\theta}) - B(\widehat{U}_{n,\theta}; \widehat{U}_{n,\theta}, \hat{e}_{n,\theta})] \\
&\equiv T_1 + T_2 + T_3 + T_4.
\end{aligned} \tag{49}
$$

Next, we estimate (49) term by term. Using the inequality $2ab \leq a^2 + b^2$ and the assumption (32), we see that the left hand side of (49) dominates

$$
\frac{1}{2\Delta t_{n+1}} (\|e_{n+1}\|_\phi^2 - \|\hat{e}_n\|_\phi^2) + C_1 \|\nabla \hat{e}_{n,\theta}\|^2. \tag{50}
$$

The right hand side of (49) can be estimated as follows. Applying (18) and Schwarz's inequality we have

$$
T_1 \leq \|\hat{e}_{n,\theta}\|_\phi (\|\rho_{n,\theta}\|_\phi + C \Delta t_{n+1}^{-1} h_{n+1}^{k_{n+1}+1} \int_{t_n}^{t_{n+1}} \|u_t\|_{k_{n+1}+1} dt). \tag{51}
$$

Applying the assumption (33) have obtain

$$
T_2 \leq C \|\hat{e}_{n,\theta}\|_\phi (\|\hat{e}_{n,\theta}\| + \|\hat{r}_{n,\theta}\| + \|\beta_{n,\theta}\|). \tag{52}
$$

In view of the assumptions (32) and (33), and the formulas (31) and (42), we see that

$$
\begin{aligned}
T_3 &= ((a(x, u_{n,\theta} + \beta_{n,\theta}) - a(x, \widehat{U}_{n,\theta})) \nabla (R_{n+1} u_{n,\theta}), \nabla \hat{e}_{n,\theta}) + (a(x, u_{n,\theta} + \beta_{n,\theta}) \nabla \alpha_{n,\theta}, \nabla \hat{e}_{n,\theta}) \\
&\leq \frac{1}{4} C_1 \|\nabla \hat{e}_{n,\theta}\|^2 + C(\|\hat{e}_{n,\theta}\|^2 + \|\hat{r}_{n,\theta}\|^2 + \|\beta_{n,\theta}\|^2 + \|\nabla \alpha_{n,\theta}\|^2),
\end{aligned} \tag{53}
$$

where C depends on $\|u\|_{L^\infty(H^{k+1}(\Omega))}$. In view of the assumptions (32), (33), (34), and the integration by parts technique (19), we derive that

$$
\begin{aligned}
T_4 &= (b(x, u_{n,\theta} + \beta_{n,\theta}) \cdot \nabla \hat{r}_{n,\theta}, \hat{e}_{n,\theta}) + ((b(x, u_{n,\theta} + \beta_{n,\theta}) - b(x, \widehat{U}_{n,\theta})) \cdot \nabla (R_{n+1} u_{n,\theta}), \hat{e}_{n,\theta}) \\
&\quad - (b(x, \widehat{U}_{n,\theta}) \cdot \nabla \hat{e}_{n,\theta}, \hat{e}_{n,\theta}) \\
&\leq \frac{1}{4} C_1 \|\nabla \hat{e}_{n,\theta}\|^2 + C(\|\hat{e}_{n,\theta}\|^2 + \|\hat{r}_{n,\theta}\|^2 + \|\beta_{n,\theta}\|^2).
\end{aligned} \tag{54}
$$

Combining (49), (50), (51), (52), (53), and (54) we have the following error inequality

$$
\begin{aligned}
&\|e_{n+1}\|_\phi^2 - \|\hat{e}_n\|_\phi^2 + C_1 \Delta t_{n+1} \|\nabla \hat{e}_{n,\theta}\|^2 \\
&\leq C \Big\{ \|\hat{e}_{n,\theta}\|_\phi E_n + \Delta t_{n+1} (\|\hat{e}_{n,\theta}\|_\phi^2 + \|\hat{r}_{n,\theta}\|^2 + \|\beta_{n,\theta}\|^2 + \|\nabla \alpha_{n,\theta}\|^2) \Big\},
\end{aligned} \tag{55}
$$

where

$$
E_n = h_{n+1}^{k_{n+1}+1} \int_{t_n}^{t_{n+1}} \|u_t\|_{k_{n+1}+1} dt + \Delta t_{n+1} \int_{t_n}^{t_{n+1}} \|u_{tt}\| dt. \tag{56}
$$

Combining (55), (22), and (23) we have

$$
\begin{aligned}
&\|e_{n+1}\|_\phi^2 - \|e_n\|_\phi^2 + C_1\Delta t_{n+1}\|\nabla \hat{e}_{n,\theta}\|^2 \\
&\le C\Big\{(\|e_{n+1}\|_\phi + \|e_n\|_\phi)E_n + \|e_n\|_\phi\|r_n - \hat{r}_n\|_\phi + \|r_n - \hat{r}_n\|_\phi E_n + \|r_n - \hat{r}_n\|_\phi^2 \\
&+\Delta t_{n+1}(\|e_{n+1}\|_\phi^2 + \|e_n\|_\phi^2 + \|\hat{r}_{n,\theta}\|^2 + \|\beta_{n,\theta}\|^2 + \|\nabla\alpha_{n,\theta}\|^2)\Big\},
\end{aligned} \tag{57}
$$

Employing the same technique as used in the proof of Theorem 2.1 leads to

$$
\begin{aligned}
&\max_{1\le n\le m} \|e_n\|_\phi^2 - \|e_0\|_\phi^2 + \sum_{n=0}^{m-1} \Delta t_{n+1}\|\nabla \hat{e}_{n,\theta}\|^2 \\
&\le C\Big\{\Big[\sum_{n=0}^{m-1} E_n\Big]^2 + \Big[\sum_{n=0}^{m-1} \|r_n - \hat{r}_n\|_\phi\Big]^2 + \sum_{n=0}^{m-1} \Delta t_{n+1}(\|\hat{r}_{n,\theta}\|^2 + \|\beta_{n,\theta}\|^2 + \|\nabla\alpha_{n,\theta}\|^2)\Big\}.
\end{aligned}
$$

Applying the triangular inequality, (42), and (43) completes the proof of this theorem. ∎

We have proved a first order accuracy in time for the the general scheme (37)-(38). When $\theta = 0$, a second order accuracy in time can be proved in the same way as in Theorem 3.1, except for (46)-(48), which should be changed respectively into

$$\|\rho_{n,\theta}\| \le \Delta t_{n+1}]\|u_{ttt}\|_{L^1([t_n,t_{n+1}];L^2(\Omega))}, \tag{58}$$

$$\|\beta_{n,\theta}\| \le \Delta t_{n+1}^2\|u_{tt}\|_{L^\infty(L^2(\Omega))}, \tag{59}$$

$$\|\alpha_{n,\theta}\|_1 \le \Delta t_{n+1}^2\|R_{n+1}u_{tt}\|_{L^\infty(H_0^1(\Omega))}. \tag{60}$$

Note the scheme (37)-(38) is nonlinear and requires some linearization technique such as Newton's iteration in computation. However, a first order linear scheme can be defined in a standard way.

Algorithm 3.2 *For $n = 0, 1, 2, \cdots, N-1$, first compute the weighted L^2 projection $\widehat{U}_n \in S_{n+1}$ by solving*

$$(\phi(\widehat{U}_n - U_n), v) = 0, \qquad \forall v \in S_{n+1}; \tag{61}$$

then compute $U_{n+1} \in S_{n+1}$ by

$$(\phi\frac{U_{n+1} - \widehat{U}_n}{\Delta t_{n+1}}, v) + A(\widehat{U}_n; \widehat{U}_{n,\theta}, v) + B(\widehat{U}_n; \widehat{U}_{n,\theta}, v) = (f(\widehat{U}_n), v), \ \ \forall v \in S_{n+1}. \tag{62}$$

A class of predictor-corrector schemes, which are second order in time for $\theta = 0$ and first order otherwise, can be defined as follows.

Algorithm 3.3 *For $n = 0, 1, 2, \cdots, N-1$, first compute the weighted L^2 projection $\widehat{U}_n \in S_{n+1}$ by solving*

$$(\phi(\widehat{U}_n - U_n), v) = 0, \qquad \forall v \in S_{n+1}; \tag{63}$$

then compute $W_{n+1} \in S_{n+1}$ *and* $U_{n+1} \in S_{n+1}$ *by*

$$(64) \qquad (\phi\frac{W_{n+1}-\widehat{U}_n}{\Delta t_{n+1}}, v) + A(\widehat{U}_n; \widehat{W}_{n,\theta}, v) + B(\widehat{U}_n; \widehat{W}_{n,\theta}, v) = (f(\widehat{U}_n), v), \quad \forall v \in S_{n+1},$$

$$(65) \qquad (\phi\frac{U_{n+1}-\widehat{U}_n}{\Delta t_{n+1}}, v) + A(\widehat{W}_{n,\theta}; \widehat{U}_{n,\theta}, v) + B(\widehat{W}_{n,\theta}; \widehat{U}_{n,\theta}, v) = (f(\widehat{W}_{n,\theta}), v), \quad \forall v \in S_{n+1},$$

where $\widehat{W}_{n,\theta} = \frac{1}{2}(1+\theta)W_{n+1} + \frac{1}{2}(1-\theta)\widehat{U}_n$.

An extrapolated scheme can be defined by the following algorithm.

Algorithm 3.4 *Given two initial approximations* U_0 *and* U_1*, for* $n = 1, 2, \cdots, N-1$*, first compute the weighted* L^2 *projection* $\widehat{U}_n \in S_{n+1}$ *by solving*

$$(66) \qquad (\phi(\widehat{U}_n - U_n), v) = 0, \qquad \forall v \in S_{n+1};$$

then compute $U_{n+1} \in S_{n+1}$ *by*

$$(67) \qquad (\phi\frac{U_{n+1}-\widehat{U}_n}{\Delta t_{n+1}}, v) + A(\tilde{U}_n; \widehat{U}_{n,\theta}, v) + B(\tilde{U}_n; \widehat{U}_{n,\theta}, v) = (f(\tilde{U}_n), v), \quad \forall v \in S_{n+1},$$

where $\tilde{U}_n = (1 + \frac{\Delta t_{n+1}}{2\Delta t_n})U_n - \frac{\Delta t_{n+1}}{2\Delta t_n}U_{n-1}$.

Compared to the second order in time space-time finite element schemes in Eriksson and Johnson [19], Algorithms 3.1, 3.3, 3.4 do not involve a coupled system like (1.1) of [19].

4 A Dynamic Characteristic Finite Element Scheme

For convection dominated diffusion problems, the modified method of characteristics may be preferred; see Douglas and Russell [15], Ewing, Russell, and Wheeler [23], Russell [38], Duran [17], Dawson, Russell, and Wheeler [12], Bank and Santos [6], and Yang [44, 49]. In this method, time discretization is made along or near the characteristic direction, instead of the t direction for standard finite difference methods. Consider the linear problem (1)-(3) as an example.

Define the characteristic direction $\tau(x)$ as

$$(68) \qquad \frac{\partial}{\partial \tau} = \frac{1}{\sqrt{\phi(x)^2 + |b(x,t)|^2}}(\phi\frac{\partial}{\partial t} + b(x,t)\cdot\nabla).$$

Thus, equation (1) can be rewritten in the form

$$(69) \qquad \sqrt{\phi^2 + |b|^2}\frac{\partial u}{\partial \tau} - \nabla\cdot(a\nabla u) + cu = f.$$

Since

$$(70) \qquad \bar{x} = x - \frac{b(x,t_n)}{\phi(x)}\Delta t_n$$

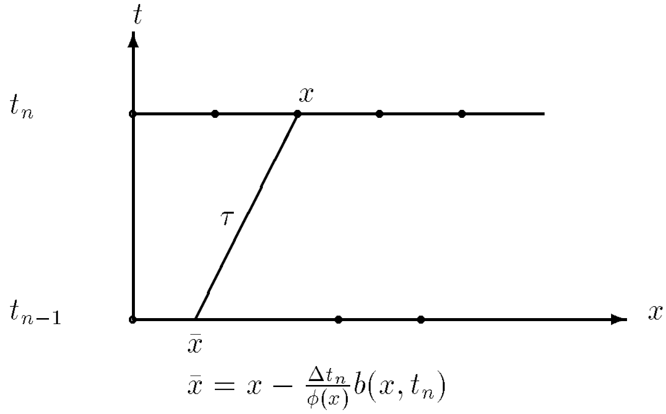

Figure 1: Discretization along the characteristic line, where x is a grid point while $\bar{x}$ is not.

approximates the characteristic through (x, t_n) by its tangent at (x, t_n), as in Figure 1, we have the following backward-difference approximation

$$
\begin{aligned}
&\sqrt{\phi^2(x) + |b(x,t_n)|^2}\, \frac{\partial u}{\partial \tau}(x, t_n) \\
(71) \qquad &\approx \sqrt{\phi^2(x) + |b(x,t_n)|^2}\, \frac{u(x,t_n) - u(\bar{x}, t_{n-1})}{[|x-\bar{x}|^2 + \Delta t_n^2]^{1/2}} \\
&= \phi(x) \frac{u(x,t_n) - u(\bar{x}, t_{n-1})}{\Delta t_n}.
\end{aligned}
$$

Then the implicit Euler scheme along characteristics reads:

Algorithm 4.1 *For $n = 1, 2, \cdots, N$, first compute the weighted L^2 projection $\widehat{U}_{n-1} \in S_n$ by solving*

$$
(72) \qquad (\phi(\widehat{U}_{n-1} - U_{n-1}), v) = 0, \qquad \forall v \in S_n;
$$

then compute $U_n \in S_n$ by

$$
(73) \qquad (\phi \frac{U_n - \widehat{\bar{U}}_{n-1}}{\Delta t_n}, v) + (a_n \nabla U_n, \nabla v) + (c_n U_n, v) = (f_n, v), \quad \forall v \in S_n,
$$

where $\bar{U}_{n-1} = U(\bar{x}, t_{n-1})$, $\widehat{\bar{U}}_{n-1} = \widehat{U}(\bar{x}, t_{n-1})$, $\bar{x}$ is defined by (70). Near the boundary of the domain, the characteristic curve may trace out of the domain. Thus periodicity of the solution function is required or some reflection principle needs to be used.

Theorem 4.1 *Suppose that the solution u to problem (1)-(3) is sufficiently regular. Let U_n be the solution of Algorithm 4.1. Then we have the error estimates for $m = 1, 2 \cdots, N$,*

$$
\begin{aligned}
&\max_{1 \le n \le m} \|u_n - U_n\| \\
(74) \qquad &\le C \Big\{ \|u_0 - U_0\| + \sum_{n=1}^{m} \Big[h_n^{k_n+1} \|u_t\|_{L^1([t_{n-1},t_n]; H^{k_n+1}(\Omega))} + \Delta t_n \|\frac{\partial^2 u}{\partial \tau^2}\|_{L^1([t_{n-1},t_n]; L^2(\Omega))} \Big] \\
&+ \sum_{n=1}^{m} \Delta t_n^{1/2} h_n^{k_n+1} \|u(\cdot, t_{n-1})\|_{k_{n-1}+1} + \max_{0 \le n \le m} h_n^{k_n+1} \|u(\cdot, t_n)\|_{k_n+1}
\end{aligned}
$$

$$+\sum_{n=1}^{m}\delta_n\Big[h_n^{k_n+1}\|u(\cdot,t_{n-1})\|_{k_n+1}+h_{n-1}^{k_{n-1}+1}\|u(\cdot,t_{n-1})\|_{k_{n-1}+1}\Big]\Big\},$$

$$\Big(\sum_{n=1}^{m}\Delta t_n[(a_n\nabla(u_n-U_n),\nabla(u_n-U_n))+(c_n(u_n-U_n),u_n-U_n)]\Big)^{1/2}$$

$$\le C\Big\{\|u_0-U_0\|+\sum_{n=1}^{m}\Big[h_n^{k_n+1}\|u_t\|_{L^1([t_{n-1},t_n];H^{k_n+1}(\Omega))}+\Delta t_n\|\frac{\partial^2 u}{\partial\tau^2}\|_{L^1([t_{n-1},t_n];L^2(\Omega))}\Big] \tag{75}$$

$$+\sum_{n=1}^{m}\Delta t_n^{1/2}h_n^{k_n}\|u(\cdot,t_n)\|_{k_n+1}$$

$$+\sum_{n=1}^{m}\delta_n\Big[h_n^{k_n+1}\|u(\cdot,t_{n-1})\|_{k_n+1}+h_{n-1}^{k_{n-1}+1}\|u(\cdot,t_{n-1})\|_{k_{n-1}+1}\Big]\Big\},$$

where $\delta_n=0$ if $S_n=S_{n-1}$ and $\delta_n=1$ otherwise.

Proof: We follow the proof of Theorem 2.1 and techniques developed in Douglas and Russell [16], Dawson, Russell, and Wheeler [12], and Duran [17]. Using the definition of the elliptic projection (11) and introducing the notation,

$$e_n=U_n-R_nu_n,\qquad \hat{\bar{e}}_{n-1}=\hat{\bar{U}}_{n-1}-R_n\bar{u}_{n-1},$$
$$r_n=u_n-R_nu_n,\qquad \hat{\bar{r}}_{n-1}=\bar{u}_{n-1}-R_n\bar{u}_{n-1},$$

we have the following error equation for Algorithm 4.1.

$$(\phi\frac{e_n-\hat{\bar{e}}_{n-1}}{\Delta t_n},e_n)+(a_n\nabla e_n,\nabla e_n)+(c_ne_n,e_n)$$
$$=(\phi\frac{r_n-\hat{\bar{r}}_{n-1}}{\Delta t_n},e_n)+(\sqrt{\phi^2+|b_n|^2}\frac{\partial u}{\partial\tau}(t_n)-\phi\frac{u_n-\bar{u}_{n-1}}{\Delta t_n},e_n). \tag{76}$$

Using a change of variable technique, we can easily obtain $\|\hat{\bar{e}}_n\|_\phi^2\le(1+C\Delta t_n)\|\hat{e}_n\|_\phi^2$. Thus the first term on the left hand side of (76) can be estimated as

$$(\phi\frac{e_n-\hat{\bar{e}}_{n-1}}{\Delta t_n},e_n)$$
$$\ge\frac{1}{2\Delta t_n}\Big(\|e_n\|_\phi^2-\|\hat{\bar{e}}_{n-1}\|_\phi^2\Big) \tag{77}$$
$$\ge\frac{1}{2\Delta t_n}\Big(\|e_n\|_\phi^2-\|\hat{e}_{n-1}\|_\phi^2\Big)-C\|\hat{e}_{n-1}\|_\phi^2.$$

Applying Lemma 1 in Douglas and Russell [15] and (18) to the first term on the right hand side of (76) we see that

$$(\phi\frac{r_n-\hat{\bar{r}}_{n-1}}{\Delta t_n},e_n)=(\phi\frac{r_n-\hat{r}_{n-1}}{\Delta t_n},e_n)+(\phi\frac{\hat{r}_{n-1}-\hat{\bar{r}}_{n-1}}{\Delta t_n},e_n)$$
$$\le\|e_n\|_\phi\|\frac{r_n-\hat{r}_{n-1}}{\Delta t_n}\|_\phi+\|e_n\|_1\|\frac{\hat{r}_{n-1}-\hat{\bar{r}}_{n-1}}{\Delta t_n}\|_{-1}$$
$$\le C\|e_n\|_\phi\Delta t_n^{-1}h_n^{k_n+1}\int_{t_{n-1}}^{t_n}\|u_t\|_{k_n+1}dt+\frac{1}{2}(a_n\nabla e_n,\nabla e_n)+C\|\hat{r}_{n-1}\|^2. \tag{78}$$

The second term on the right hand side of (76) can be estimated using Taylor's expansion along characteristics. Finally we obtain the error inequality

$$\|e_n\|_\phi^2 - \|\hat{e}_{n-1}\|_\phi^2 + \Delta t_n[(a_n\nabla e_n, \nabla e_n) + (c_n e_n, e_n)]$$
$$\leq C\Big[\|e_n\|_\phi\Big(h_n^{k_n+1}\int_{t_{n-1}}^{t_n}\|\frac{\partial u}{\partial t}\|_{k_n+1}dt + \Delta t_n\int_{t_{n-1}}^{t_n}\|\frac{\partial^2 u}{\partial \tau^2}\|dt\Big) + \Delta t_n(\|\hat{e}_{n-1}\|_\phi^2 + \|\hat{r}_{n-1}\|^2)\Big]. \tag{79}$$

Following the proof of Theorem 2.1, we can easily finish the rest of the proof. ∎

5 Dynamic Mixed Finite Element Schemes

Mixed finite element methods approximate the solution and its gradient simultaneously and transform the original second order problem into a system of first order equations. Define the flux $\sigma = -a\nabla u$ and the Sobolev space $H(\mathrm{div};\Omega) = \{v \in L^2(\Omega) : \nabla\cdot v \in L^2(\Omega)\}$, we have the mixed weak formulation for the problem (1)-(3): find $\{u,\sigma\} \in L^2(\Omega)\times H(\mathrm{div};\Omega)$ such that

$$(\phi\frac{\partial u}{\partial t}, v) + (\nabla\cdot\sigma, v) - (b\cdot(a^{-1}\sigma), v) + (cu, v) = (f, v), \quad v \in L^2(\Omega), \tag{80}$$

$$(a^{-1}\sigma, \psi) - (u, \nabla\cdot\psi) = 0, \quad \forall\psi \in H(\mathrm{div};\Omega), \tag{81}$$

$$(u(\cdot,0), v) = (g, v), \quad \forall v \in L^2(\Omega). \tag{82}$$

Let $M_n \times X_n \subset L^2(\Omega)\times H(\mathrm{div};\Omega)$ be a mixed finite element space (see Brezzi and Fortin [10] and Brenner and Scott [7]) at $t = t_n$. We assume that the following approximation property holds: for $n = 1, 2, \cdots, N$,

$$\inf_{z\in M_n}\|v - z\| \leq Ch_n^s\|v\|_s, \quad 0 \leq s \leq k_n + 1, \quad \forall v \in L^2(\Omega), \tag{83}$$

$$\inf_{w\in X_n}\|\psi - w\| \leq Ch_n^s\|\psi\|_s, \quad 0 \leq s \leq k_n + 1, \quad \forall\psi \in H(\mathrm{div};\Omega), \tag{84}$$

where we assume that C is a constant independent v, n, h_n. Discussions of the solvability and stability conditions of the mixed finite element method can also be found in Brezzi, Douglas, Durán, and Fortin [8], Brezzi, Douglas, Fortin, and Marini [9], and Raviart and Thomas [36]. Then a dynamic mixed finite element method reads:

Algorithm 5.1 *For $n = 1, 2, \cdots, N$, first compute the weighted L^2 projection $\widehat{U}_{n-1} \in M_n$ by solving*

$$(\phi(\widehat{U}_{n-1} - U_{n-1}), v) = 0, \qquad \forall v \in M_n; \tag{85}$$

then compute $\{U_n, W_n\} \in M_n \times X_n$ by

$$(\phi\frac{U_n - \widehat{U}_{n-1}}{\Delta t_n}, v) + (\nabla\cdot W_n, v) - (b_n\cdot(a_n^{-1}W_n), v) + (c_n U_n, v) = (f_n, v), \quad v \in M_n, \tag{86}$$

$$(a_n^{-1}W_n, w) - (U_n, \nabla\cdot w) = 0, \quad \forall w \in X_n. \tag{87}$$

Theorem 5.1 *Suppose that the solution u to problem (1)-(3) is sufficiently regular. Let U_n be the solution of Scheme (85)-(87). Then we have the error estimates for $m = 1, 2\cdots, N$,*

$$\max_{1\le n\le m}\|u_n-U_n\|+\Big(\sum_{n=1}^{m}\Delta t_n[(a_n^{-1}(\sigma_n-W_n),\sigma_n-W_n)+(c_n(u_n-U_n),u_n-U_n)]\Big)^{1/2}$$
$$\le C\Big\{\|u_0-U_0\|+\sum_{n=1}^{m}\Big[h_n^{k_n+1}\|u_t\|_{L^1([t_{n-1},t_n];H^{k_n+1}(\Omega))}+\Delta t_n\|u_{tt}\|_{L^1([t_{n-1},t_n];L^2(\Omega))}\Big] \tag{88}$$
$$+\sum_{n=1}^{m}\Delta t_n^{1/2}h_n^{k_n+1}\|\nabla u(\cdot,t_n)\|_{k_n+1}+\max_{0\le n\le m}h_n^{k_n+1}(\|u(\cdot,t_n)\|_{k_n+1}+\|\nabla u(\cdot,t_n)\|_{k_n+1})$$
$$+\sum_{n=1}^{m}\delta_n\Big[h_n^{k_n+1}(\|u(\cdot,t_{n-1})\|_{k_n+1}+\|\nabla u(\cdot,t_{n-1})\|_{k_n+1})$$
$$+h_{n-1}^{k_{n-1}+1}(\|u(\cdot,t_{n-1})\|_{k_{n-1}+1}+\|\nabla u(\cdot,t_{n-1})\|_{k_{n-1}+1})\Big]\Big\},$$

where $\delta_n = 0$ if $M_n\times X_n = M_{n-1}\times X_{n-1}$ and $\delta_n = 1$ otherwise.

Proof: Introduce the elliptic projection $\{R_nu, R_n\sigma\}$ of $\{u,\sigma\}$: find $\{R_nu(\cdot,t), R_n\sigma(\cdot,t)\}\in M_n\times X_n$ for each $t\in[0,T]$ such that

$$(\nabla\cdot(R_n\sigma-\sigma),v)+(c_n(R_nu-u),v)=0,\quad \forall v\in M_n, \tag{89}$$
$$(a_n^{-1}\cdot R_n\sigma,w)-(R_nu,\nabla\cdot w)=0,\quad \forall w\in X_n. \tag{90}$$

Then it is standard to show that

$$\|u-R_nu\|+\|\sigma-R_n\sigma\|$$
$$\le Ch_n^s(\|u\|_s+\|\nabla u\|_s),\ 0\le s\le k_n+1, t\in[0,T],\ n=0,1,\cdots,N. \tag{91}$$

Define

$$\begin{aligned} e_n &= U_n-R_nu_n, & \hat e_{n-1} &= \widehat U_{n-1}-R_nu_{n-1},\\ r_n &= u_n-R_nu_n, & \hat r_{n-1} &= u_{n-1}-R_nu_{n-1},\\ \epsilon_n &= W_n-R_n\sigma_n, & \eta_n &= \sigma_n-R_n\sigma_n. \end{aligned}$$

Combining (86), (87), (80), (81), (89), and (90), we have

$$(\phi\frac{e_n-\hat e_{n-1}}{\Delta t_n},v)+(\nabla\cdot\epsilon_n,v)+(c_ne_n,v)$$
$$=(\phi\frac{r_n-\hat r_{n-1}}{\Delta t_n},v)+(\phi\frac{\partial u}{\partial t}(t_n)-\phi\frac{u_n-u_{n-1}}{\Delta t_n},v)+(b_n\cdot a_n^{-1}(W_n-\sigma_n),v),\qquad \forall v\in M_n, \tag{92}$$
$$(a_n^{-1}\epsilon_n,w)-(e_n,\nabla\cdot w)=0,\ \ \forall w\in X_n. \tag{93}$$

Taking $v=e_n$ in (92) and $w=\epsilon_n$ in (93) and adding we obtain the error equation

$$(\phi\frac{e_n-\hat e_{n-1}}{\Delta t_n},e_n)+(a_n^{-1}\epsilon_n,\epsilon_n)+(c_ne_n,e_n)$$
$$=(\phi\frac{r_n-\hat r_{n-1}}{\Delta t_n},e_n)+(\phi\frac{\partial u}{\partial t}(t_n)-\phi\frac{u_n-u_{n-1}}{\Delta t_n},e_n) \tag{94}$$
$$+(b_n\cdot(a_n^{-1}\epsilon_n),e_n)-(b_n\cdot(a_n^{-1}\eta_n),e_n).$$

Applying (14), (17), (18), and the ϵ-inequality to (94) we have

$$\begin{aligned} &\|e_n\|_\phi^2 - \|\hat{e}_{n-1}\|_\phi^2 + \Delta t_n[(a_n^{-1}\epsilon_n, \epsilon_n) + (c_n e_n, e_n)] \\ &\leq C[(F_n + \Delta t_n\|\eta_n\|)\|e_n\|_\phi + \Delta t_n\|e_n\|_\phi^2], \end{aligned} \tag{95}$$

where F_n is defined in (21). Now the rest of the proof follows directly from Theorem 2.1. ∎

Crank-Nicolson and θ schemes can also be considered in a similar fashion. For nonlinear problems, our framework and error analysis presents no problems.

6 Concluding Remarks

We have analyzed a number of dynamic finite element methods for second order linear and nonlinear parabolic equations using a unified framework. This framework enables us to study different variants of the finite element schemes and obtain unified convergence results. In particular, when the finite element space changes from time step to time step, the modified method of characteristics and mixed finite element methods are all treated in the same way as standard finite element methods. The convergence results obtained in this paper are optimal and offer a clear picture on the propagation of error due to the change of the finite element space. Our analysis also proves the convergence of some schemes which were not guaranteed by previous theory.

This paper has emphasized on the convergence theory of dynamic finite element methods and paid little attention to implementation issues. For example, where to apply fine grids and how to make grid refinement are very import problems in practice. There is a large literature on grid refinement strategies and here we just mention a few. Large gradient areas are usually the places where the solution develops steep layers or fronts. Thus predicting large gradient areas from the solution obtained at previous time step and making local grid refinement is one strategy [13]. In this respect, mixed finite element methods provide a more accurate prediction of the gradient and thus may be a good choice. One popular method among the engineering community, though, is to postprocess the approximate solution to obtain more accurate representation of the gradient. A-posteriori error estimation is another way for doing adaptivity and local grid refinement [19, 1]. When an approximate solution is obtained, the error between the approximate solution and the true solution can be estimated based on the information about the coefficients of the given partial differential equation and the approximate solution, which can be evaluated elementwise. Elements with large error are then subdivided into finer grids. Explicit a-posteriori estimators can be computed directly from the finite element solution and the coefficients of the differential equation based on the residual equation, while implicit a-posteriori estimators require solving local boundary value problems approximating the residual equation satisfied by the error.

No matter how grid refinement and interpolation polynomial modification are made, our convergence theory states that the error between the exact solution and the approximate solution consists of three parts: a time finite difference discretization error, a spatial finite element discretization error, and an error term due to the projection of the approximated solution from old finite element spaces onto new finite element spaces. A good strategy to minimize the error would be that make grid refinement in a larger area to cover the local phenomena for several (maybe dozens of) time steps and that change the finite element space less frequently.

Numerical experiments have also shown that changing the grids at every time step or making grid refinement not according to the changing location of the local phenomena would affect the accuracy of the approximate solution. In Yang [48], the author combined grid refinement and domain decomposition techniques to capture moving local phenomena for a model parabolic problem. Grid refinement was made only in subdomains that contain the local layer and coarse grid was applied in other subdomains. When the local layer moves, the domain was decomposed dynamically in such a way that the local layer was always contained in some subdomains, minimizing its intersection with interdomain boundaries, to improve accuracy.

References

[1] M. Ainsworth and J. T. Oden, A Posteriori error estimation in finite element analysis, TICAM Report 96-19, Texas Institute for Computational and Applied Mathematics, University of Texas at Austin, May 1996.

[2] I. Babuska and B. Q. Guo, Approximation properties of the h-p version of the finite element method, Comput. Methods Appl. Mech. Engrg. 133 (1996), 319–346.

[3] I. Babuska and M. Suri, The p and h-p versions of the finite element method, basic principles and properties, SIAM Rev. 36 (1994) 578–632.

[4] M. J. Baines, An analysis of the moving finite element procedure, SIAM J. Numer. Anal. 28(1991) 1323-1349.

[5] M. J. Baines, Moving finite elements, Oxford University Press, New York, 1994.

[6] R. E. Bank and R. F. Santos, Analysis of some moving space-time finite element methods, SIAM J. Nume . Anal., 30(1993) 1-18.

[7] S. C. Brenner and L. R. Scott, The Mathematical Theory of Finite Element Methods, Springer-Verlag, New York, 1994.

[8] F. Brezzi, J. Douglas, Jr., R. Durán, and M. Fortin, Mixed finite elements for second order elliptic problems in three variables, Numer. Math. 51(1987) 237-250.

[9] F. Brezzi, J. Douglas, Jr., M. Fortin, and L. D. Marini, Efficient rectangular mixed finite elements in two and three space variables, RAIRO Modél. Math. Anal. Numér. 21(1987) 581-604.

[10] F. Brezzi, and M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, New York, 1991.

[11] P. G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

[12] C. N. Dawson, T. F. Russell, M. F. Wheeler, Some improved error estimates for the modified method of characteristics, SIAM J. Numer. Anal. 26 (1989) 1487–1512.

[13] J. C. Diaz, R. E. Ewing, A. C. McDonald, L. M. Uhler, and D. U. von Rosenberg, Self-adaptive local grid refinement for time-dependent two-dimensional simulations, in: R. H. Gallagher, D. Norrie, J. T. Oden, and O. C. Zienkiewicz, Eds., Finite Elements in Fluids, Vol. 6, Wiley, New York, 1985, pp. 279-290.

[14] J. Douglas, Jr. and T. F. Dupont, Galerkin methods for parabolic problems, SIAM J. Nume . Anal., 7(1970) 575-626.

[15] J. Douglas, Jr. and T. F. Russell, Numerical methods for convection–dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures, SIAM J. Numer. Anal. 19(1982) 871-885.

[16] T. F. Dupont, Mesh modification for evolution equations, Math. Comp., 39(1982) 85-108.

[17] R. Duran, On the approximation of miscible displacement in porous media by a method of characteristics combined with a mixed method, SIAM J. Numer. Anal. 25(1988) 989–1001.

[18] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, Introduction to adaptive methods for differential equations, Acta Numerica, 1995, 105–158,

[19] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems I: A linear model problem, SIAM J. Numer. Anal. 28 (1991) 43–77.

[20] 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) 706–740.

[21] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems IV: Nonlinear problems, SIAM J. Numer. Anal. 32 (1995) 1729–1749.

[22] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems V: Long-time integration, SIAM J. Numer. Anal. 32 (1995) 1750–1763.

[23] R. E. Ewing, T. F. Russell and M. F. Wheeler, Simulation of miscible displacement using mixed methods and a modified method of characteristics, SPE 12241, 7th SPE Symposium on Reservoir Simulation, San Francisco, November 15-18, 1983.

[24] G. M. Hulbert and T. J. R. Hughes, Space-time finite element methods for second-order hyperbolic equations, Comput. Meth. Appl. Mech. Engrg, 84(1990) 327-348.

[25] J. Jaffre, C. Johnson, and A. Szepessy, Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws, Math. Models Methods Appl. Sci. 5(1995) 367–386.

[26] P. Jamet, Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain, SIAM J. Numer. Anal. 15(1978) 912-928.

[27] P. Jamet, Stability and convergence of a generalized Crank-Nicolson scheme on a variable domain for the heat equation, SIAM J. Numer. Anal. 17(1980) 530-539.

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

[29] C. Johnson, Discontinuous Galerkin finite element methods for second order hyperbolic problems, Comput. Meth. Appl. Mech. Engrg, 107(1993) 117-129.

[30] G. P. Liang, A finite element method with moving grid, Math. Numer. Sinica 7 (1985) 377–384.

[31] G. P. Liang, A finite element method of semidiscretization with moving grid, J. Comput. Math. 4 (1986) 86–96.

[32] G. P. Liang and Z. M. Chen, A full-discretization moving FEM with optimal convergence rate, Chinese J. Numer. Math. Appl. 12(1990) 91–111.

[33] B. Lucier, A moving mesh numerical method for hyperbolic conservation laws, Math. Comp. 46(1986) 59-69.

[34] K. Miller, Moving finite elements II, SIAM J. Numer. Anal. 18(1981) 1033-1057.

[35] K. Miller and R. N. Miller, Moving finite elements I, SIAM J. Numer. Anal. 18(1981) 1019-1032.

[36] P.–A. Raviart and J.–M. Thomas, A mixed finite element method for second order elliptic problems, in: I. Galligani and E. Magenes, Eds., Mathematical Aspects of the Finite Element Method, Lecture Notes in Mathematics 606, Springer-Verlag, Berlin and New York, 1977, pp. 292-315.

[37] H. X. Rui, A finite element method with moving grid for solving quasilinear pseudohyperbolic equations, Numer. Math. J. Chinese Univ. 10 (1988) 318–331.

[38] T. F. Russell, Finite elements with characteristic finite element method for a miscible displacement problem, SPE 10500, Proc. 6th SPE Symposium on Reservoir Simulation, Dallas, TX, 1982, pp. 123-135.

[39] C. Schwab and M. Suri, The p and hp versions of the finite element method for problems with boundary layers, Math. Comp. 65 (1996) 1403–1429.

[40] M. Suri, Analytical and computational assessment of locking in the hp finite element method, Comput. Methods Appl. Mech. Engrg. 133 (1996) 347–371.

[41] M. F. Wheeler, A priori L^2 error estimates for Galerkin approximations to parabolic partial differential equations, SIAM J. Numer. Anal., 10(1973) 723-759.

[42] D. Q. Yang, The mixed finite element methods with moving grids for parabolic problems, Math. Numer. Sinica 10(1988) 266-271.

[43] D. Q. Yang, Mixed methods with dynamic finite-element spaces for miscible displacement in porous media, J. Comput. Appl. Math. 30(1990) 313-328.

[44] D. Q. Yang, A characteristic mixed method with dynamic finite element space for convection-dominated diffusion problems, J. Comput. Appl. Math. 43(1992) 343-353.

[45] D. Q. Yang, Grid modification for the wave equation with attenuation, Numer. Math. 67(1994) 391-401.

[46] D. Q. Yang, Grid modification for second order hyperbolic problems, Math. Comp. 64(1995) 1495-1509.

[47] D. Q. Yang, Different domain decompositions at different times for capturing moving local phenomena, J. Comput. Appl. Math. 59(1995) 39-48.

[48] D. Q. Yang, Dynamic domain decomposition and grid modification for parabolic problems, Computers and Mathematics with Applications, to appear.

[49] D. Q. Yang, Numerical simulation of miscible displacement in porous media using an iterative perturbation method combined with a modified method of characteristics. (in preparation)

[50] Y. R. Yuan, On finite element methods with moving mesh for 2-phase immiscible flow, Sci. Sinica, Ser. A, 29 (1986) 785–799.

[51] Y. R. Yuan, On characteristic finite element methods with moving mesh for convection diffusion problems, Numer. Math. J. Chinese Univ. 8(1986) 236-245.