Computing and Visualization in Science manuscript No.
(will be inserted by the editor)

Randolph E. Bank · **Maximilian S. Metti**

An Error Analysis of Some Higher Order Space-Time Moving Finite Elements

Received: May 27, 2014 / Accepted: date

Abstract This is a study of certain finite element methods designed for convection-dominated, time-dependent partial differential equations. Specifically, we analyze high order space-time tensor product finite element discretizations, used in a method of lines approach coupled with mesh modification to solve linear partial differential equations. Mesh modification can be both continuous (moving meshes) and discrete (static rezone). These methods can lead to significant savings in computation costs for problems having solutions that develop steep moving fronts or other localized time-dependent features of interest. Our main result is a symmetric a priori error estimate for the finite element solution computed in this setting.

Keywords Moving Finite Elements, Error Analysis, Symmetric Error Estimate, Convection-Dominated

Mathematics Subject Classification (2000) 65M15, 65M50, 65M60

1 Introduction

Computing accurate solutions to convection-dominated partial differential equations using standard finite element methods can be computationally expensive, and sometimes prohibitively so. Consequently, the use of adaptive methods can lead to great savings in computation time and maintain accuracy of the computed solution [1,6]. It is often the case that regions in which the solution to a partial differential equation is rough or rapidly changing are relatively small compared to the overall domain and adaptive methods leverage this fact by focusing more computational effort by placing a higher concentration of the degrees of freedom in these regions and avoiding "over-solving" where the solution is smooth [8]. Effectively, adaptive methods are designed to automate the process of finding a finite element space that is well-suited to solving a given differential equation. In using adaptive methods, one inherently employs nonuniform meshes. When analyzing finite element methods that use nonuniform meshes, it is necessary to make certain shape regularity assumptions on the mesh to ensure reasonable approximation properties of the finite element space as well as the well-posedness of discrete problem. Furthermore, existing software packages, like PLTMG [2], have the capability to employ nonuniform meshes as well as measure the mesh quality of a given nonuniform grid.

For time-dependent problems, these critical regions can move throughout the domain, as in the case of steep moving fronts, and short time steps may be required to maintain a desired level of accuracy in these regions [14,18]. In order to avoid these short time steps, a moving mesh can be used to continuously track this moving region.

Moving finite elements were initially proposed by Miller and Miller in [18,19] and have been analyzed and implemented in the context of linear and nonlinear problems. One of the main themes in moving finite element methods has been to devise strategies for effectively moving the mesh efficiently and accurately solve a given problem. In this paper, however, we focus on providing an error analysis for a space-time moving finite element method that allows for a variety of mesh motion schemes.

The first error analysis of moving finite element methods is given by Dupont in [12], where a symmetric error bound of the form

$$|||u-u_h||| \leq C \inf_{v\in V_h} |||u-v|||, \tag{1}$$

was proved in the semi-discrete case (continuous in time, discrete in space). Here u and u_h are the true solution and the finite element solution to the differential equation, respectively, V_h is the finite element space, and $|||\cdot|||$ is a specially defined mesh-dependent energy norm related to the differential equation. Symmetric error bounds are proven for linear moving finite elements by Bank and Santos in [7,21] and

Bank: The work of this author was supported by the National Science Foundation under contract DMS-1318480.

Bank: Department of Mathematics, University of California, San Diego, La Jolla, California 92093-0112. Email: rbank@ucsd.edu.
Metti: Department of Mathematics, The Pennsylvania State University, University Park, Pennsylvania 16802. Email: msm37@psu.edu.

by Dupont and Mogultay [13]. Some symmetric error estimates for mixed methods that use moving meshes are proven in [16].

Here we prove a symmetric error bound like (1) for finite element spaces of arbitrary order, and more general time integration schemes. An important extension in this work to previous research is the use of multistage time integration techniques that lead to higher-order convergence. To do this, we first describe an underlying space-time tensor-product finite element space that allows the spatial discretization to evolve continuously in time, except at discrete time steps where the mesh is allowed to reconfigure in a discontinuous fashion. Since this analysis employs higher order finite element spaces, say order p, the spatial nodes are permitted to follow piecewise polynomial trajectories of degree p in time, allowing smoother and more dynamic mesh motion and, ultimately, longer time steps than linear mesh motion or static meshes. The analysis of this paper indicates that reasonable mesh motion schemes align with the convection velocity of a given problem, though the aim of this paper is to provide an error analysis for a class of general moving mesh methods, rather than prescribe the details of a particular implementation of these methods.

In §2, the time-dependent linear convection-diffusion equation is introduced and the construction of a space-time tensor-product finite element space is described. In §3, the notation for the analysis is established, a new space-time shape regularity constraint is proposed for the moving finite elements, and some preliminary results are given. A space-time moving finite element method is proposed and analyzed in §4, and a symmetric error estimate is proved for finite element spaces of arbitrary order. We conclude in §5 with some remarks on this error analysis and an error analysis for which more general time integration schemes can be employed.

2 Notation and Definitions

The spatial domain Ω is assumed to be a compact subset of $\mathbb{R}^d$, where $d = 1, 2$, or 3, with boundary $\partial\Omega$, and the time domain is a finite interval, $(0,T]$. Let a, b, c, and f be smooth and bounded functions defined on $\Omega \times (0,T]$ such that $a \geq \bar{a} > 0$ and $c \geq 0$, and g is integrable on $\partial\Omega$. Let u_0 be a given initial condition for the solution on Ω and let n denote the outward unit normal vector to the boundary $\partial\Omega$.

The solution to the differential equation, denoted by u, is the function that satisfies

$$u_t - \nabla\cdot(a\nabla u) + b\cdot\nabla u + cu = f, \qquad \text{in } \Omega\times(0,T], \quad (2)$$

$$a\nabla u\cdot n = g, \qquad \text{on } \partial\Omega\times(0,T], \quad (3)$$

$$u(x,0) = u_0(x), \qquad \text{for } x \text{ in } \Omega.$$

For convenience, a Neumann boundary condition (3) is assumed; our results still hold with minor changes when other boundary conditions are imposed.

When the convection velocity, b in (2), is large, the solution of the equation may develop steep shock layers that sweep through the spatial domain. When computing solutions numerically, such moving structures can be difficult to track accurately and require small time steps for non-moving finite elements. To offset the effects of a potentially large convection velocity, we replace the time derivative in (2) with space-time directional derivative, as in the method of characteristics. Define a time-dependent parameterization of the spatial variable, $x : \Omega\times[0,T] \to \Omega$, that is invertible in space and differentiable in time. This leads to the *characteristic* derivative:

$$\begin{aligned} u_\tau(x(y,t),t) &\equiv \tfrac{\partial}{\partial t}u(x(y,t),t) \\ &= u_t(x(y,t),t) + x_t(y,t)\cdot\nabla u(x(y,t),t). \end{aligned}$$

To simplify notation, we use the spatial inverse $y = y(x,t)$ to write the velocity field $x_t(t) = x_t(y(x,t),t)$ in terms of x, and typically choose $x_t(t)$ to approximate the convection velocity $b(x,t)$

We replace the time derivative with the characteristic derivative in (2) and propose a weak form of the differential equation that can reduce the presence of the convection velocity: find u with $u(t) \in H^1(\Omega)$ and $u_t(t) \in L_2(\Omega)$ such that for all χ in $H^1(\Omega)$ and $0 < t \leq T$,

$$\big(u_\tau(\cdot,t),\chi\big) + A_\tau(t;u,\chi) = \big(f(\cdot,t),\chi\big) + \big\langle g(\cdot,t),\chi\big\rangle, \qquad (4)$$

and when $t = 0$

$$\big(u(\cdot,0),\chi\big) = \big(u_0,\chi\big).$$

The inner-products are given by

$$(f,\chi) = \int_\Omega f(x)\chi(x)\,dx,$$

$$\langle g,\chi\rangle = \int_{\partial\Omega} g(s)\chi(s)\,ds,$$

and define the time-dependent bilinear form

$$\begin{aligned} A_\tau(t;u,\chi) \equiv \int_\Omega \Big[& a(x,t)\nabla u(x,t)\cdot\nabla\chi(x) \\ & + (b(x,t) - x_t(t))\cdot\nabla u(x,t)\;\chi(x) + c(x,t)u(x,t)\chi(x)\Big]\,dx. \end{aligned}$$

Notice that parameterizing the spatial variable so that $x_t \approx b$ leads to a formulation where the convection velocity is much less prominent, as it is "absorbed" into the characteristic derivative.

2.1 A space-time tensor-product moving finite element space

To compute a solution to the differential equation, we restrict the trial and test spaces of equation (4) to a finite element space, V_h^p, where $p \geq 1$ is the maximum order of the piecewise polynomials that generate the space. The finite element spaces described in this paper are tensor-products of finite element spaces on Ω with a finite element discretization of

the time domain. These are akin to the finite element spaces described in [7,13,21], except that these previous works restricted attention to the case of linear elements ($p = 1$).

Partition the time domain into m disjoint intervals, where the endpoints of the partitions are given by $\{t_i\}$ and satisfy

$$0 = t_0 < t_1 < \ldots < t_m = T,$$

with $\Delta t_i \equiv t_i - t_{i-1}$ for $i = 1,\ldots,m$. Within each time partition, we define a triangulation of the spatial domain Ω with nodes $\{x_k(t)\}_{k=0}^{n_i}$ such that each node is a polynomial of degree p on $(t_{i-1},t_i]$. To avoid mesh degeneration, the mesh topology is required not to change in time — when $d = 1$, this corresponds to $x_{k-1}(t) < x_k(t)$ for $t_{i-1} \le t \le t_i$ and $k = 1,\ldots,n_i$. The resulting mesh is a series of time partitions on which we have a triangulation of the mesh that evolves continuously in time. An example time partition, with quadratic mesh motion ($p = 2$) and $d = 1$, is depicted in figure 1.

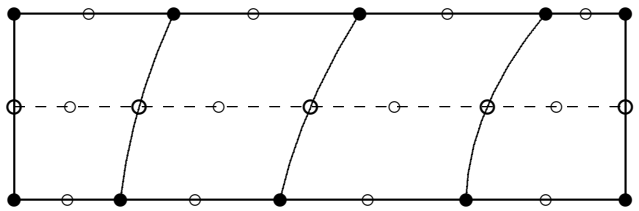

Fig. 1 An example space-time mesh partition with $d = 1$ and $p = 2$. The filled circles represent the space-time "hat" basis functions; hollow circles correspond to basis functions that are the product of a "bump" function with a "hat" or "bump" function.

As described above, a continuously moving mesh is defined on each time partition. Between time partitions, however, discontinuous changes in the mesh are permitted. These discontinuities provide for the periodic addition, deletion, and relocation of the spatial nodes to track structures that may develop or vanish over time. Another important benefit of these mesh reconfigurations is that they provide a means to avoid the nodes from entangling or colliding [9, 10,15].

For $d = 1$, the elements in the mesh are curvilinear trapezoids with flat parallel edges corresponding to the beginning and end of the time partitions, and curved edges representing the moving space nodes — recall that these edges are polynomial curves of degree p. The associated space-time reference element is the Cartesian product of the reference element in space with the reference element in time. Moreover, the space-time basis functions are given by the tensor products of the degree p polynomial spatial basis functions on $[0,1]$ and the degree p polynomial temporal basis functions on $[0,1]$. This implies that there are $(p+1)^2$ degrees of freedom associated with the reference element. Since a tensor product is used to define the space-time basis functions, the degrees of freedom on the reference element are aligned into time slices. The degrees of freedom are represented by the filled and empty circles in figure 1 and their alignment into time slices is emphasized by the dashed line.

When $d = 2$, the elements are curvilinear triangular prisms with flat triangular parallel edges corresponding to the beginning and end of a time partition, and the curved edges corresponding to the space nodes moving along polynomial trajectories of degree p. The reference element is the Cartesian product of the unit triangle and the unit interval, giving a wedge-like shape, and has $(p+1)\times(p+1)(p+2)/2$ degrees of freedom. The space-time basis functions are the tensor product of the degree p polynomial basis functions in space, on the unit triangle, and the degree p polynomial basis functions in time on the unit interval.

For $d = 3$, the space-time reference element is the Cartesian product of the unit tetrahedron with the unit interval reference element for time and has $(p+1)\times(p+1)(p+2)(p+3)/6$ degrees of freedom.

Isoparametric maps are used to map the degrees of freedom from the reference element to elements in the mesh. We start with the case $d = 1$ and then generalize to higher dimensions. Let e be an element in the i^{th} time partition $(t_{i-1},t_i]$ given by $e(t) = [x_{k-1}(t),x_k(t)]$, where $x_{k-1}(t)$ and $x_k(t)$ are polynomials of degree p satisfying $x_{k-1}(t) < x_k(t)$. The time component of the isoparametric map is given by the affine map

$$t = (1-\hat{t})t_{i-1} + \hat{t}\,t_i = t_{i-1} + \hat{t}\Delta t_i,$$

for $0 \le \hat{t} \le 1$. Taking $\Delta x_k(t) = x_k(t) - x_{k-1}(t)$, the time dependent isoparametric map for the element $e(t)$ is

$$x(t) = (1-\hat{x})x_{k-1}(t) + \hat{x}x_k(t) = x_{k-1}(t) + \hat{x}\Delta x_k(t),$$

with $0 \le \hat{x} \le 1$.

The inverse of the isoparametric map for element e is

$$\hat{t} = \frac{t - t_{i-1}}{\Delta t_i}$$

and

$$\hat{x}(t) = \frac{x - x_{k-1}(t)}{\Delta x_k(t)},$$

and exists whenever $\Delta x_k(t) > 0$ for t in $[t_{i-1},t_i]$ and x in $e(t) = [x_{k-1}(t),x_k(t)]$. Since the inverse of the spatial component is an affine transformation in space, it holds that the finite element space at any fixed time t is a standard finite element space of continuous piecewise polynomials of degree p defined over Ω. We denote this slice of the finite element space by $V_h^p(t)$. This property of the finite element space justifies analyzing finite element functions on a time slice — for example, $\phi(t)$ in $V_h^p(t)$ — in a way that is consistent with the study of finite element methods for autonomous problems. Note that $\Delta t_i > 0$ and $\Delta x_k(t) > 0$ are always required for the finite element space to be well-defined.

When $d = 2$ or 3, the time component of the isoparametric map is unaffected, but the spatial component is now given by a vector mapping that is affine in space and a polynomial

of degree p in time. Taking the derivative of the isoparametric map, one obtains the block triangular Jacobian matrix

$$J_e(t) = \begin{bmatrix} J_e(t) & \Delta t_i x_t(t)^T \\ 0 & \Delta t_i \end{bmatrix},$$

where $J_e(t)$ represents the $d \times d$ Jacobian matrix of isoparametric map of element $e(t)$, and $x_t(t)$ is a d-vector that is affine in space and a polynomial of degree $p-1$ along the node trajectories $x_k(t)$. The trajectories traced out by $x(t)$ are called the *characteristic* trajectories of the mesh and the vector x_t describes the mesh motion of the finite element space. These characteristic trajectories can be chosen to offset the convection velocity, by setting $x_t \approx b$, in hopes of attaining greater flexibility in the length of permissible time steps. As in the case of one spatial dimension, we require $\Delta t_i > 0$ and for the spatial mesh

$$D_e(t) \equiv |\det(J_e(t))| > 0.$$

Note that $D_e(t)$ is proportional to the size of the spatial element $e(t)$, so this constraint prevents the element from degenerating in time.

As mentioned above, each time slice of the finite element space is a standard finite element discretization on Ω. Thus, for ϕ in V_h^p we have $\phi(t) \in V_h^p(t) \subset H^1(\Omega)$. Along the characteristic trajectories, the finite element function ϕ is a piecewise polynomial of degree p, defined on the partition given by $\{t_i\}_{i=0}^m$. That is, the function $\phi(x(y,t),t)$ is a polynomial of degree p for $t_{i-1} < t \le t_i$ and fixed y in Ω.

Recall that the finite element functions can have discontinuities between the mesh partitions; define the jump of ϕ by

$$[\phi](t_i) = \lim_{\delta \to 0^+} \phi(t_i + \delta) - \phi(t_i - \delta) \equiv \phi(t_{i^+}) - \phi(t_{i^-}).$$

To uniquely define the finite element functions at these discontinuities, we follow Dupont [12] by requiring the jump to be orthogonal to the finite element space at the beginning of the new partition:

$$\big([\phi](t_i), \chi\big) = 0, \tag{5}$$

for all χ in $V_h^p(t_{i^+})$.

3 Preliminary Results

Multi-index notation is used to represent spatial derivatives, but time and characteristic derivatives do not follow this convention. The $H^k(\Omega)$ semi-norm and norm follow conventional notation and we write

$$|v|_k = \left(\sum_{|\alpha|=k} (D_\alpha v, D_\alpha v) \right)^{1/2}$$

and

$$\|v\|_k = \left(\sum_{|\alpha|\le k} (D_\alpha v, D_\alpha v) \right)^{1/2}.$$

Following Dupont [12], a mesh-dependent semi-norm is defined that allows us to prove our symmetric error estimate,

$$\|v\|_{(-1,V_h^p(t))} = \sup_{\substack{\chi \in V_h^p(t) \\ \chi \neq 0}} \frac{|(v,\chi)|}{\|\chi\|_1}.$$

We also use the infinity norm, $\|v\|_\infty = \max_{x\in\Omega} |v(x)|$.

Let Q_{ref} represent a *reference* quadrature rule, which is defined on the time reference element $[0,1]$ as the interpolatory quadrature rule with knots at a given set of time collocation nodes $\{\hat{t}_j\}_{j=1}^p$. It is required that the knots of the Q_{ref} satisfy

$$0 < \hat{t}_1 < \ldots < \hat{t}_p \le 1, \tag{6}$$

and that the weights $\{w_j\}$ are all positive for $j = 1, \ldots, p$. For convenience, let $\hat{t}_0 = 0$ be coincident with a degree of freedom on the reference element. Using the isoparametric maps, this quadrature rule can be applied to the time partitions:

$$Q_i(v) \equiv \sum_{j=1}^p w_j v(t_{i-1} + \hat{t}_j \Delta t_i) \approx \frac{1}{\Delta t_i} \int_{t_{i-1}}^{t_i} v(t)\, dt.$$

We now introduce a space-time shape regularity constraint for the moving finite elements that controls the time evolution of the spatial elements and prevents degenerate elements. Fix e to be an element in the time partition with $t_{i-1} \le t \le t_i$. Then, the Jacobian matrix at time t can be represented as

$$J_e(t) = \big(R_e(t) + \Delta t_i H_e(t)\big) J_e(t_{i-1^+}), \tag{7}$$

for some orthogonal *rotation* matrix, $R_e(t)$, and *evolution* matrix, $H_e(t)$. The matrix $R_e + \Delta t_i H_e$ is constrained to have polynomial entries of degree at most p throughout the time partition. As the name suggests, the matrix $R_e(t)$ describes the element rotation in time, and the evolution matrix describes the deformation of the shape of the element. Since the trajectories of the spatial nodes are restricted to polynomial paths of degree p, elements do not rotate perfectly and more of a twisting action is observed; the evolution matrix necessarily reflects these deformations. If an element is merely translated in time, without rotation or changing shape, then the Jacobian matrix, $J_e(t)$, remains unchanged.

Let $\rho(\cdot)$ represent the spectral radius for $d \times d$ matrices. For space-time regularity, it is assumed that the evolution matrix H_e has a uniformly bounded spectral radius throughout the time step; namely, there exists some positive constant μ that does not depend on e or t such that

$$\rho\big(H_e(t)\big) \le \mu. \tag{8}$$

This can be interpreted as bounding the relative change in shape and size of the element over time.

Assuming a non-degenerate finite element space and the space-time shape regularity bound (8), it follows that

$$\rho\left(J_e(t)J_e^{-1}(t_{i-1^+})\right) = \rho\left(R_e(t) + \Delta t_i H_e(t)\right) \leq 1 + \mu\Delta t_i \quad (9)$$

and, for $\tilde{c}_{\mu,d} = [(1+\mu\Delta t_i)^d - 1]/\Delta t_i = O(1)$ and $\Delta t_i \leq 1/2\tilde{c}_{\mu,d}$,

$$\begin{aligned}(1-\tilde{c}_{\mu,d}\Delta t_i) \leq (1-\mu\Delta t_i)^d &\leq \frac{D_e(t)}{D_e(t_{i-1^+})} \\ = \det\left(J_e(t)J_e^{-1}(t_{i-1^+})\right) &\leq (1+\mu\Delta t_i)^d \leq (1+\tilde{c}_{\mu,d}\Delta t_i).\end{aligned} \quad (10)$$

Let ϕ be a function in the finite element space $V_h^p(t)$ for some t in the time partition $(t_{i-1}, t_i]$. We *shift* ϕ onto the mesh of $V_h^p(t_{i-1^+})$, at the beginning of the time partition by replacing the basis functions of $V_h^p(t)$ with with their corresponding basis functions in $V_h^p(t_{i-1^+})$, while preserving the basis coefficients. Formally, this operation can be defined by an element-wise composition of the inverse of the affine spatial isoparametric maps for the elements in the mesh at time t, which is well-defined for non-degenerate meshes, with the affine spatial isoparametric maps for the elements at the beginning of the time step t_{i-1^+}. The following lemma establishes the relationship between the space-time shape regularity constraint (8) and the continuity of this shift operation.

Lemma 1 (Shift Lemma) *Let $\phi, \chi \in V_h^p(t)$ and $\tilde{\phi}, \tilde{\chi} \in V_h^p(t_{i-1^+})$ represent a pair of finite element functions and their shifts, respectively, on a non-degenerate time partition of the mesh that satisfies (8) on each element. If $\Delta t_i \leq 1/2\tilde{c}_{\mu,d}$, as defined in (10), then there exists a positive constant $C_{\mu,d}$ such that*

$$\left|(\phi,\chi) - (\tilde{\phi},\tilde{\chi})\right| \leq C_{\mu,d}\Delta t_i \frac{\|\tilde{\phi}\|_0^2 + \|\tilde{\chi}\|_0^2}{2}, \quad (11)$$

$$\left|\|\phi\|_0^2 - \|\tilde{\phi}\|_0^2\right| \leq C_{\mu,d}\Delta t_i \|\tilde{\phi}\|_0^2, \quad (12)$$

$$\left|\|\phi\|_1^2 - \|\tilde{\phi}\|_1^2\right| \leq C_{\mu,d}\Delta t_i \|\tilde{\phi}\|_1^2. \quad (13)$$

Proof The proof follows from an element-wise change of variables and using the space-time shape regularity constraint (8) to uniformly bound the relative change in element size by bound (10). For bounding the difference of a function and its shift in the H^1-norm, the Jacobian matrices of the element-wise transformations multiply the finite element function, following from the chain rule. These Jacobian matrices are bounded by (9). See [17] for a more detailed proof. □

We conclude this section with the following discrete Grönwall inequality.

Lemma 2 (Discrete Grönwall Inequality) *Let $\Delta t_i > 0$ and $\alpha_i, \gamma_i, \theta_i, q_i \geq 0$, for $1 \leq i \leq m$, with $\theta_i\Delta t_i \leq \frac{1}{2}$ and $\theta = \max_i \theta_i$. Then, if*

$$\frac{q_i - q_{i-1}}{\Delta t_i} + \gamma_i \leq \alpha_i + \theta_i(q_i + q_{i-1}),$$

there exists a positive constant C_θ such that

$$\max_{1\leq i\leq m} q_i + \sum_{i=1}^m \gamma_i\Delta t_i \leq C_\theta\left\{q_0 + \sum_{i=1}^m \alpha_i\Delta t_i\right\}.$$

This theorem comes directly from [21] and the proof can be found therein. Its argument primarily follows the proof of the standard discrete Grönwall lemma with minor modifications.

4 A space-time moving finite element method

We now discretize the weak formulation of the differential equation (4). Given the mesh velocity, x_t, and collocation nodes, $\{t_{i,j}\}$, find u_h in V_h^p such that for each $t_{i,j}$ and all χ in $V_h^p(t_{i,j})$, the finite element solution satisfies

$$\left(\partial_\tau u_h(t_{i,j}),\chi\right) + A_\tau\left(t_{i,j}; u_h, \chi\right) = \left(f(t_{i,j}),\chi\right) + \left\langle g(t_{i,j}),\chi\right\rangle \quad (14)$$

for $i = 1,\ldots,m$ and $j = 1,\ldots,p$, and when $t = 0$,

$$\left(u_h(\cdot,0),\chi\right) = \left(u_0,\chi\right).$$

Each time partition is coupled to the previous time partition by the jump orthogonality condition (5), which states that the finite element solution must satisfy

$$(u_h(t_{i^+}),\chi) = (u_h(t_{i^-}),\chi),$$

for all χ in $V_h^p(t_{i^+})$, at each time step $i = 1,\ldots,m-1$. We emphasize that using L_2-projection in practice is typically a difficult and expensive task, especially in higher dimensions. To this effect, interpolation is recommended for practical purposes, though L_2-projection facilitates the theoretical analysis of the method. However, should one wish to leverage the stability properties of projection, it is worth noting that the discontinuities should be minor between each time step, only adding and removing a few nodes in most realistic cases, which can simplify the task of computing the inner-products of the basis functions on these two meshes.

This formulation effectively solves for the solution on each time partition sequentially. The mesh motion is assumed to be pre-computed for the time partition and no specific mesh motion is prescribed. As a result, this analysis encompasses many mesh moving strategies including the method of characteristics ($x_t = b$) and non-moving meshes ($x_t \equiv 0$). Since the mesh motion is assumed to be known before computing the solution, however, strategies for moving the mesh cannot depend on the values of the computed solution, unless values from previous time partitions or predictor-corrector schemes are used.

An important feature of this method is that the time collocation nodes, where (14) is imposed, need not coincide with the time basis nodes. Here the collocation nodes will be the p nodes of quadrature formula Q of order at least $2p-1$. For any polynomial $v \in \mathbb{R}^{2p-1}$, we assume that Q satisfies

$$\int_0^1 v\,dx = Q(v) - C_p v^{(2p-1)} \tag{15}$$

where $C_p \geq 0$. The constant $C_p = 0$ when Q is classical Gaussian quadrature, since its order is $2p$, while $C_p > 0$ for Gauss-Radau quadrature [20]. C_p is also positive for the one parameter family of quadrature rules of order $2p-1$ that interpolates between the Gauss and Gauss-Radau rules. In the special case $p = 1$, the Gauss rule corresponds to an integration scheme analogous to the Crank-Nicolson method; the Gauss-Radau rule is analogous to the first backward difference formula. The family of rules that interpolate between the two corresponds to the family of so-called θ-methods.

Let $\{t_{i,j}\}$ represent the collocation nodes and $\{\zeta_{i,j}\}$ represent the time basis nodes, and let $\phi \in V_h^p$. At the collocation node $t = t_{i,j}$, let $\chi \in V_h^p(t_{i,j})$ and $\tilde{\phi}(t) \in V_h^p(t_{i,j})$ be the shift of $\phi(t)$ onto the mesh at time $t_{i,j}$. Then, the finite element solution is determined by

$$\sum_{k=0}^{p} \Big\{ \frac{1}{\Delta t_i} \beta_k'(\hat{t}_j)(\tilde{\phi}(\zeta_{i,k}),\chi) + \beta_k(\hat{t}_j)\big[(a\nabla\tilde{\phi}(\zeta_{i,k}),\nabla\chi) + ((b-x_t)\cdot\nabla\tilde{\phi}(\zeta_{i,k}),\chi) + (c\tilde{\phi}(\zeta_{i,k}),\chi)\big]\Big\} = (f,\chi) + \langle g,\chi\rangle,$$

where β_j represents the time basis functions on the reference element and $\hat{t}_j$ represents the collocation nodes on the reference element, $j = 1,\ldots,p$. To ensure the existence and uniqueness of the finite element solution for sufficiently small Δt_i, we show that the following necessary condition is satisfied.

Lemma 3 *Let $\{\beta_j\}_{j=0,\ldots,p}$ be the Lagrange basis for polynomials of degree p with nodes $0 = \hat{\zeta}_0 < \hat{\zeta}_1 < \cdots < \hat{\zeta}_p \leq 1$. Then, the matrix defined by $B \equiv [\beta_k'(\hat{t}_j)]_{j,k}$ is invertible, where $\hat{t}_j$ is a strictly ordered partition of $(0,1]$ and $1 \leq j,k \leq p$. Furthermore $\|B\|$ and $\|B^{-1}\|$ are independent of the time steps, space discretization, and details of the pde.*

Proof Let $v \in \mathbb{R}^p$ be chosen to satisfy $Bv = 0$. Then, the polynomial of degree p defined by

$$v(\hat{t}) \equiv \sum_{j=1}^{p} \beta_j(\hat{t}) v_j$$

satisfies $v'(\hat{t}_j) = 0$ for $j = 1,\ldots,p$. Accordingly, the derivative $v'(\hat{t})$ is identically zero as it has p distinct roots. Thus, the polynomial $v(\hat{t})$ is constant and $\beta_j(0) = 0$ for $j = 1,\ldots,p$, which implies that $v(\hat{t}) \equiv 0$. Equivalently, the vector v must be trivial. Since all calculations take place on the reference interval, $\|B\|$ and $\|B^{-1}\|$ can only depend on the quadrature rule Q, the basis functions on the reference element, and the choice of norm. □

From lemma 3, the linear system corresponding to the finite element formulation is nonsingular, which proves the existence and uniqueness of the finite element solution exists and is unique when Δt_i is sufficiently small.

We now present a local Grönwall lemma that will be used to bound the maximum error of the finite element solution in the L_2-norm over each time partition.

Lemma 4 (Local Grönwall Inequality) *Let Q_i be the interpolatory quadrature rule with positive weights and p distinct collocation nodes $\{t_{i,j}\}$ on the time partition $(t_{i-1},t_i]$. Suppose the space-time mesh on the time partition satisfies the regularity constraint (8) at the collocation nodes, where $1 \leq i \leq m$ and that there exists a positive constant κ such that*

$$\|b - x_t\|_\infty \leq \kappa. \tag{16}$$

If $\Delta t_i \leq 1/2\tilde{c}_{\mu,d}$, as defined in (10), and functions ϕ in V_h^p and η in the solution space satisfy

$$\big(\phi_\tau(t_{i,j}),\chi\big) + A_\tau\big(t_{i,j};\phi,\chi\big) = \big(\eta_\tau(t_{i,j}),\chi\big) + A_\tau\big(t_{i,j};\eta,\chi\big) \tag{17}$$

for all χ in $V_h^p(t_{i,j})$ at time each collocation node $t_{i,j}$, then, there exists a constant such that

$$\max_{1\leq j\leq p} \big\|\phi(t_{i,j})\big\|_0^2 \leq C\Big\{ \big\|\phi(t_{i-1^+})\big\|_0^2 + \Delta t_i Q_i\left(\|\eta_\tau\|^2_{(-1,V_h^p(\cdot))} + \|\eta\|_1^2 + \|\phi\|_1^2\right)\Big\},$$

where C depends on κ,μ,d,p, and the differential equation.

Proof To simplify notation, the time partition index, i, is assumed to be fixed and we let $\phi_j \equiv \phi(t_{i,j})$. Define $\tilde{\phi}_\ell^{(j)}$ in $V_h^p(t_{i,j})$ to be the finite element function with the same basis coefficients as ϕ_ℓ multiplying the basis functions for time $t_{i,j}$. The function $\tilde{\phi}_\ell^{(j)}$ is, therefore, the shift of $\phi(t_{i,\ell})$ onto the mesh at time $t_{i,j}$, and $\tilde{\phi}_j^{(j)} = \phi_j$.
Let k index the collocation node where ϕ attains its maximal L_2 norm: $\|\phi_k\|_0 = \max_{1\leq j\leq p}\|\phi_j\|_0$. Choose $\chi = \tilde{\phi}_k^{(j)}$ in $V_h^p(t_{i,j})$ for equation (17), $j = 1,\ldots,p$. Use the time basis expansion of the characteristic derivative to get

$$\frac{1}{\Delta t_i}\sum_{\ell=0}^{p} \beta_\ell'(\hat{t}_j)\big(\tilde{\phi}_\ell^{(j)},\tilde{\phi}_k^{(j)}\big) + A_\tau\big(t_{i,j};\phi_j,\tilde{\phi}_k^{(j)}\big) = \big(\eta_{j,\tau},\tilde{\phi}_k^{(j)}\big) + A_\tau\big(t_{i,j};\eta_j,\tilde{\phi}_k^{(j)}\big),$$

for $j = 1,\ldots,p$, where β_j and $\hat{t}_j$ represents the time basis functions and the collocation nodes on the reference

element. Equivalently,

$$\sum_{\ell=1}^{p}\beta_\ell'(\hat{t}_j)\left(\tilde{\phi}_\ell^{(j)},\tilde{\phi}_k^{(j)}\right)=-\beta_0'(\hat{t}_j)\left(\tilde{\phi}_0^{(j)},\tilde{\phi}_k^{(j)}\right)+\Delta t_i\left\{\left(\eta_{j,\tau},\tilde{\phi}_k^{(j)}\right)+A_\tau\left(t_{i,j};\eta_j-\phi_j,\tilde{\phi}_k^{(j)}\right)\right\}. \quad (18)$$

By lemma 3, there exists a linear combination of the derivatives of the time basis functions such that

$$\sum_{j=1}^{p}\alpha_j\sum_{\ell=1}^{p}\beta_\ell'(\hat{t}_j)v_\ell=v_k,$$

for $1\le k\le p$ and any $v=[v_j]_j\in\mathbb{R}^p$. Thus, we take this linear combination of the equation (17) so that the left side simplifies and can be bounded using lemma 1:

$$\begin{aligned}&\sum_{j=1}^{p}\alpha_j\sum_{\ell=1}^{p}\beta_\ell'(\hat{t}_j)\left(\tilde{\phi}_\ell^{(j)},\tilde{\phi}_k^{(j)}\right)\\&\ge\sum_{j=1}^{p}\alpha_j\sum_{\ell=1}^{p}\beta_\ell'(\hat{t}_j)\left(\tilde{\phi}_\ell^{(0)},\tilde{\phi}_k^{(0)}\right)\\&\quad-\frac{1}{2}C_\mu\Delta t_i\sum_{j=1}^{p}|\alpha_j|\sum_{\ell=1}^{p}|\beta_\ell'(\hat{t}_j)|\left(\|\tilde{\phi}_\ell^{(0)}\|_0^2+\|\tilde{\phi}_k^{(0)}\|_0^2\right)\\&\ge\|\tilde{\phi}_k^{(0)}\|_0^2-\hat{C}_{\mu,d,p}\Delta t_i\sum_{j=1}^{p}\|\tilde{\phi}_j^{(0)}\|_0^2\\&\ge\|\tilde{\phi}_k^{(0)}\|_0^2-\hat{C}'_{\mu,d,p}\Delta t_i\sum_{j=1}^{p}\|\phi_j\|_0^2. \quad (19)\end{aligned}$$

To bound the right side, choose $\delta>0$ to be sufficiently small, say $\delta<1/2$, and use the shift lemma (lemma 1) to get

$$-\sum_{j=1}^{p}\alpha_j\beta_0'(\hat{t}_j)\left(\tilde{\phi}_0^{(j)},\tilde{\phi}_k^{(j)}\right)\le\hat{C}''_{\mu,d,p}\|\phi_0\|_0^2+\delta\|\tilde{\phi}_k^{(0)}\|_0^2. \quad (20)$$

Furthermore, we use the mesh dependent negative norm and the shift lemma to bound

$$\left(\eta_{j,\tau},\tilde{\phi}_k^{(j)}\right)\le\frac{1}{2}\left(\|\eta_{j,\tau}\|^2_{(-1,V_h^p(t_{i,j}))}+C_{\mu,d,p}\|\phi_k\|_1^2\right) \quad (21)$$

and

$$\begin{aligned}A_\tau\left(t_{i,j};\eta_j-\phi_j,\tilde{\phi}_k^{(j)}\right)&\le C_\kappa\left(\|\eta_j\|_1^2+\|\phi_j\|_1^2+\|\tilde{\phi}_k^{(j)}\|_1^2\right)\\&\le C_{\kappa,\mu,d,p}\left(\|\eta_j\|_1^2+\|\phi_j\|_1^2+\|\phi_k\|_1^2\right).\end{aligned} \quad (22)$$

From (19)–(22), we get

$$\|\tilde{\phi}_k^{(0)}\|_0^2\le C_{\kappa,\mu,d,p}\Big\{\|\phi(t_{i-1^+})\|_0^2+\sum_{j=1}^{p}\left(\|\eta_{j,\tau}\|^2_{(-1,V_h^p(t_{i,j}))}+\|\eta_j\|_1^2+\|\phi_j\|_1^2\right)\Big\}. \quad (23)$$

Since the quadrature weights are positive, there is a constant $C_Q>0$ such that

$$\begin{aligned}&\sum_{j=1}^{p}\left(\|\eta_{j,\tau}\|^2_{(-1,V_h^p(t_{i,j}))}+\|\eta_j\|_1^2+\|\phi_j\|_1^2\right)\\&\quad\le C_Q Q_i\left(\|\eta_\tau(\cdot)\|^2_{(-1,V_h^p(\cdot))}+\|\eta(\cdot)\|_1^2+\|\phi(\cdot)\|_1^2\right). \quad (24)\end{aligned}$$

By the shift lemma, $\|\tilde{\phi}_k^{(0)}\|\ge C_{\mu,d,p}\|\phi_k\|_0=C_{\mu,d,p}\max_{1\le j\le p}\|\phi_j\|_0$, and combining the bounds (23)–(24) yields the desired result. □

The mesh-dependent energy norm used for our symmetric error estimate is given by

$$\begin{aligned}|||u|||^2&=\max_{\substack{1\le i\le m\\1\le j\le p}}\|u(t_{i,j})\|_0^2\\&\quad+\sum_{i=1}^{m}\Delta t_i Q_i\left(\|u_\tau(\cdot)\|^2_{(-1,V_h^p(\cdot))}+\|u(\cdot)\|_1^2\right). \quad (25)\end{aligned}$$

This energy norm is a discrete approximation of a norm associated with Grönwall estimates applied to continuous parabolic equations,

$$|||u|||^2\approx\max_{0\le t\le T}\|u(t)\|_0^2+\int_0^T\|u_\tau(t)\|_{-1}^2+\|u(t)\|_1^2\,dt,$$

which measures the spatial and characteristic derivatives in space-time, as well as the maximal L_2-norm of u. We note that refinement of the time partition leads to a better approximation of the time integral by the composite quadrature rule in (25).

The main result of this paper is the quasi-optimal error bound for the finite element solution determined by (14).

Theorem 1 *Suppose that V_h^p is a finite element space with a non-degenerate mesh and time collocation nodes that satisfy (15). Furthermore, assume that there exist positive constants μ and κ such that at each collocation node*

$$\rho\left(H_e(t_{i,j})\right)\le\mu, \quad (26)$$

and

$$\|b-x_t\|_\infty\le\kappa. \quad (27)$$

Then, if $\Delta t=\max_{1\le i\le m}\Delta t_i$ is sufficiently small, there exists a positive constant C such that the finite element solution satisfies

$$|||u-u_h|||\le C\inf_{\chi\in V_h^p}|||u-\chi|||, \quad (28)$$

where C depends on μ,κ,d,p, and the differential equation.

Proof From the Galerkin orthogonalities, at each collocation node

$$\begin{aligned}\big(\partial_\tau u_h(t_{i,j}),\chi\big)+A_\tau\big(t_{i,j};u_h,\chi\big)\\ =\big(\partial_\tau u(t_{i,j}),\chi\big)+A_\tau\big(t_{i,j};u,\chi\big),\end{aligned}$$

for all χ in $V_h^p(t_{i,j})$ with $i=1,\dots,m$ and $j=1,\dots,p$. Let $\psi\in V_h^p$ and define $\phi\equiv u_h-\psi$ in V_h^p and $\eta\equiv u-\psi$. Re-write the statement of the Galerkin orthogonality as

$$\begin{aligned}(\partial_\tau\phi(t_{i,j}),\chi)+A_\tau\big(t_{i,j};\phi,\chi\big)\\ =(\partial_\tau\eta(t_{i,j}),\chi)+A_\tau\big(t_{i,j};\eta,\chi\big),\end{aligned}\tag{29}$$

for all χ in $V_h^p(t_{i,j})$. Using equation (29) we will show

$$|||\phi|||\le C|||\eta|||$$

and use the triangle inequality to obtain the sought after bound (28).
Fix i and j and choose $\chi=\phi(t_{i,j})$ so that (29) gives

$$(\phi_\tau,\phi)+A_\tau(\phi,\phi)=(\eta_\tau,\phi)+A_\tau(\eta,\phi),\tag{30}$$

at time $t=t_{i,j}$. The bound at this collocation node comes from bounding the terms in (30) individually. For the first term on the left, the shift lemma gives the bound

$$(\phi_\tau,\phi)=\frac{1}{2}\partial_\tau\|\phi\|_0^2\ge\frac{1}{2}\frac{d}{dt}\|\tilde\phi\|_0^2-C_{\mu,d,p}\max_{1\le k\le p}\|\phi(t_{i,k})\|_0^2,\tag{31}$$

where $\tilde\phi$ is the shift onto the initial mesh of the time partition. Since the mesh motion satisfies (27),

$$\begin{aligned}A_\tau(\phi,\phi)&\ge\bar a|\phi|_1^2-\kappa|\phi|_1\|\phi\|_0+\bar c\|\phi\|_0^2\\ &\ge C_A\|\phi\|_1^2-C_{A,\kappa}\|\phi\|_0^2.\end{aligned}\tag{32}$$

Now, choose $\varepsilon>0$ to be sufficiently small and bound

$$(\eta_\tau,\phi)\le C\|\eta_\tau\|^2_{(-1,V_h^p(t_{i,j}))}+\varepsilon\|\phi\|_1^2\tag{33}$$

and, since (27) holds,

$$A_\tau(\eta,\phi)\le C_{A,\kappa}\|\eta\|_1^2+\varepsilon\|\phi\|_1^2.\tag{34}$$

Hence, from (30)–(34), it is true at $t=t_{i,j}$ that

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|\tilde\phi\|_0^2+C_A\|\phi\|_1^2\\ &\le C_{A,\kappa,\mu,d,p}\left\{\|\eta_\tau\|^2_{(-1,V_h^p(t_{i,j}))}+\|\eta\|_1^2+\max_{1\le k\le p}\|\phi(t_{i,k})\|_0^2\right\}.\end{aligned}\tag{35}$$

Using the local quadrature rule Q_i, with positive weights, to aggregate (35) over the time partition, we recover

$$\begin{aligned}&Q_i\Big(\frac{d}{dt}\|\tilde\phi\|_0^2\Big)+C_AQ_i\big(\|\phi\|_1^2\big)\\ &\le C_{A,\kappa,\mu,d,p}\Big\{Q_i\big(\|\eta_\tau\|^2_{(-1,V_h^p(t_{i,j}))}+\|\eta\|_1^2\big)\\ &\qquad+\max_{1\le k\le p}\|\phi(t_{i,k})\|_0^2\Big\}.\end{aligned}\tag{36}$$

The collocation nodes satisfy (15), which implies that

$$\begin{aligned}Q_i\Big(\frac{d}{dt}\|\tilde\phi\|_0^2\Big)&\ge\frac{1}{\Delta t_i}\int_{t_{i-1}}^{t_i}\frac{d}{dt}\|\tilde\phi\|_0^2\,dt\\ &=\frac{\|\tilde\phi(t_{i^-})\|_0^2-\|\tilde\phi(t_{i-1^+})\|_0^2}{\Delta t_i}.\end{aligned}\tag{37}$$

From bound (37), the shift lemma, and the local Grönwall inequality, we may write

$$\begin{aligned}&\frac{\|\phi(t_{i^-})\|_0^2-\|\phi(t_{i-1^+})\|_0^2}{\Delta t_i}+C_AQ_i\big(\|\phi\|_1^2\big)\\ &\le C_{A,\kappa,\mu,d,p}\Big\{Q_i\big(\|\eta_\tau\|^2_{(-1,V_h^p(t_{i,j}))}+\|\eta\|_1^2\big)\\ &\qquad+\|\phi(t_{i^-})\|_0^2+\|\phi(t_{i-1^+})\|_0^2\Big\}.\end{aligned}\tag{38}$$

We now use the discrete Grönwall inequality to show

$$\max_{0\le i\le m}\|\phi(t_{i^-})\|_0^2+\sum_{i=1}^m\Delta t_iQ_i(\|\phi\|_1^2)\le C\left\{\|\phi(0)\|_0^2+|||\eta|||^2\right\}.\tag{39}$$

Since an L_2-projection is used for the initial condition,

$$\|\phi(0)\|_0\le\|\eta(0)\|_0\le|||\eta|||.\tag{40}$$

Furthermore, at each collocation node, $t_{i,j}$,

$$(\partial_\tau\phi,\chi)=(\partial_\tau\eta,\chi)+A_\tau(\eta,\chi)-A_\tau(\phi,\chi)$$

for all χ in $V_h^p(t_{i,j})$, which implies that

$$\begin{aligned}&\|\partial_\tau\phi\|_{(-1,V_h^p(t_{i,j}))}\\ &\qquad\le C_{A,\kappa}\left\{\|\partial_\tau\eta\|_{(-1,V_h^p(t_{i,j}))}+\|\eta\|_1+\|\phi\|_1\right\}.\end{aligned}\tag{41}$$

From (39)–(41),

$$\max_{0\le i\le m}\|\phi(t_{i^-})\|_0^2+\sum_{i=1}^m\Delta t_iQ\big(\|\phi\|_1^2+\|\phi\|^2_{(-1,V_h^p(\cdot))}\big)\le C|||\eta|||^2.$$

All that is needed to conclude the proof is the local Grönwall inequality once more to show

$$\max_{1\le j\le p}\|\phi(t_{i,j})\|_0^2 \le C\Big\{\|\phi(t_{i-1^+})\|_0^2 + \Delta t_i Q_i\left(\|\eta_\tau\|^2_{(-1,V_h^p(\cdot))}+\|\eta\|_1^2+\|\phi\|_1^2\right)\Big\} \le C|||\eta|||^2 \quad (42)$$

for $i=1,\dots,m$. Thus, we have $|||\phi||| \le C|||\eta|||$, as desired. □

Theorem 1 states that the finite element solution computed by the presented method provides a quasi-optimal approximation to the solution of the variational problem relative to the finite element space. Furthermore, the application of Grönwall inequalities indicate exponential dependence on the alignment of the convection and mesh velocities in the bounding coefficient. This highlights the benefits of the additional flexibility of using higher order mesh motion. We now turn to a result that establishes the approximation properties of these higher order moving finite element spaces.

Lemma 5 *Suppose V_h^p is a finite element space with a non-degenerate mesh that is shape regular at each collocation node and (8) holds with $\Delta t \le 1/2\tilde{c}_{\mu,d}$ and $\Delta x/\Delta t_i \le \gamma$, for some positive constant γ, $i=1,\dots,m$. Then, the error of the interpolant, u_I, using the collocation nodes of the quadrature rule Q is bounded by*

$$|||u-u_I||| \le C(\Delta x^p+\Delta t^p)\Big\{\max_{0<t\le T}|u(t)|^2_{p+1} + \int_0^T \|\partial_\tau^{p+1}u\|_0^2\,dt + \sum_{i=1}^m \max_{0\le k\le n_i}\int_{t_{i-1}}^{t_i}\Big[\Big(\frac{d}{dt}\Big)^{p+1}\tilde{u}(x_k,t)\Big]^2 dt + \sum_{i=1}^m\sum_{j=0}^p \Delta t_i\big(|u(t_{i,j})|^2_{p+1}+|u_\tau(t_{i,j})|^2_{p+1}\big)\Big\}^{1/2},$$

for some positive constant $C=C_{\gamma,\mu,d,p}$ which also depends on the shape regularity and the collocation scheme.

The proof for this result is contained in the appendix, though we remark upon the condition $\Delta x/\Delta t_i \le \gamma$. This condition ensures that the spatial discretization is not too coarse relative to the time discretization so that the projection error between time partitions does not accumulate over many short time steps. In most practical cases, the mesh discontinuities will be limited to a relatively few nodes and small regions of the domain and this constraint plays a less significant role. Immediately from Theorem 1 and Lemma 5, we have the following bound on the error of the finite element solution.

Theorem 2 *Suppose V_h^p is a non-degenerate finite element space such that the conditions of Theorem 1 and Lemma 5 hold. Then,*

$$|||u-u_h||| \le C(\Delta x^p+\Delta t^p)\Big\{\max_{0<t\le T}|u(t)|^2_{p+1} + \int_0^T \|\partial_\tau^{p+1}u\|_0^2\,dt + \sum_{i=1}^m \max_{0\le k\le n_i}\int_{t_{i-1}}^{t_i}\Big[\Big(\frac{d}{dt}\Big)^{p+1}\tilde{u}(x_k,t)\Big]^2 dt + \sum_{i=1}^m\sum_{j=0}^p \Delta t_i\big(|u(t_{i,j})|^2_{p+1}+|u_\tau(t_{i,j})|^2_{p+1}\big)\Big\}^{1/2}.$$

5 Discussion

In the case $p=1$, Theorem 1 improves on the results given in [7] in several small but significant ways. First, the energy norm $|||\cdot|||$ employs the characteristic space-time derivative rather than the time derivative, an improvement first used in [16]. Theorem 1 also unifies proofs for Crank-Nicolson and first backward difference approaches under one umbrella, and extends the theory to the family of first order θ-methods. For the case $p>1$, we believe our result to be new.

The family of methods covered by Theorem 1 corresponds to a family of p-stage fully implicit Runge-Kutta methods. Except for the case $p=1$, these methods are rarely used in practice. Being fully implicit, all degrees of freedom from all stages become coupled, resulting in a system of equations of order Np to be solved, where N is the number of degrees of freedom in the space dimensions. When $p>1$, one normally moves to diagonally implicit Runge-Kutta methods, where degrees of freedom associated with stage r depend only on degrees of freedom from stages $0,1,\dots,r-1$. This gives the system of equations a block triangular shape, yielding p similar systems of order N to be solved in each time step, rather than one system of order Np. Indeed, the example that inspired this work is the TR-BDF2 method [3,4,11,22], which in this context is a second order two stage diagonally implicit Runge-Kutta method. More informally, it consists of a first half-step using the trapezoid rule (Crank-Nicolson) followed by a half-step using the second backward difference formula. Unfortunately, thus far we have been unable to prove a symmetric error estimate like Theorem 1 for the general case of p-stage diagonally implicit Runge-Kutta methods. The main issue is that we have not been able to cast such methods in a fully Galerkin finite element framework. We have been successful in analyzing the TR-BDF2 method [5,17], but were able to achieve only a partially symmetric error estimate; in particular, some extra time-truncation terms appear on the right hand side of the analogue of (28). This is reminiscent of similar terms that appeared in the analysis of time discretization schemes in [12].

6 Appendix

We prove the interpolation error bound of Lemma 5.

Proof This proof uses the definition of the energy semi-norm (25) to split the approximation error into several terms that can be bounded independently. We have

$$|||u-u_I|||^2 = \max_{\substack{1\le i\le m\\ 1\le j\le p}} \left\|\left(u-u_I\right)(t_{i,j})\right\|_0^2 + Q\left(\left\|\partial_\tau\left(u-u_I\right)(\cdot)\right\|^2_{(-1,V_h^p(\cdot))} + \left\|\left(u-u_I\right)(\cdot)\right\|_1^2\right). \quad (43)$$

For any collocation node $t_{i,j}$ with $j>0$, the shape regularity assumptions can be used to bound

$$\left\|\left(u-u_I\right)(t_{i,j})\right\|_k \le C\Delta x^{p+1-k}|u(t_{i,j})|_{p+1}, \quad (44)$$

for $k=0$ and 1, where C depends on d, p, and the shape regularity of the mesh.
Bounding the terms involving the characteristic derivatives is more intricate due to the moving nodes in the mesh. To circumvent the complications that arise from the characteristic derivative, we use the results proven in section 3 to shift the functions onto a stationary mesh. The characteristic derivative corresponds to a time derivative upon shifting u and u_I onto a stationary mesh:

$$\left\|\partial_\tau\left(u-u_I\right)(t_{i,j})\right\|_{(-1,V_h^p(t_{i,j}))} \le (1+C_{\mu,d}\Delta t_i)\left\|\tilde{u}_t(t_{i,j}) - \frac{d}{dt}\tilde{u}_{\tilde{I}}(t_{i,j})\right\|_0, \quad (45)$$

where $\tilde{u}_{\tilde{I}}$ represents the interpolant of shifted function $\tilde{u}$ on the stationary mesh. Since the finite element interpolant must satisfy (5), we introduce an auxiliary interpolant $\tilde{u}_{\tilde{K}}$ that interpolates $\tilde{u}$ at on the mesh at time t_{i-1^+} so that we have

$$\begin{aligned}
&\left\|\tilde{u}_t(t_{i,j}) - \frac{d}{dt}\tilde{u}_{\tilde{I}}(t_{i,j})\right\|_0 \\
&\le \left\|\tilde{u}_t(t_{i,j}) - \frac{d}{dt}\tilde{u}_{\tilde{K}}(t_{i,j})\right\|_0 + \left\|\frac{d}{dt}\tilde{u}_{\tilde{K}}(t_{i,j}) - \frac{d}{dt}\tilde{u}_{\tilde{I}}(t_{i,j})\right\|_0 \\
&\le \left\|\tilde{u}_t(t_{i,j}) - \frac{d}{dt}\tilde{u}_{\tilde{K}}(t_{i,j})\right\|_0 \\
&\qquad + \frac{1}{\Delta t_i}|\hat{\beta}'_{i,0}(0)|\left\|\tilde{u}_{\tilde{K}}(t_{i-1}) - \tilde{u}_{\tilde{I}}(t_{i-1})\right\|_0. \quad (46)
\end{aligned}$$

Since $\tilde{u}_{\tilde{K}}(t_{i-1})$ and $\tilde{u}_{\tilde{I}}(t_{i-1})$ are interpolations of $\tilde{u}(t_{i-1})$ on two (potentially) distinct meshes, we use (44) and $\Delta x \le \gamma\Delta t_i$ to show

$$\begin{aligned}
\frac{1}{\Delta t_i}\left\|\tilde{u}_{\tilde{K}}(t_{i-1}) - \tilde{u}_{\tilde{I}}(t_{i-1})\right\|_0 &\le 2\gamma C\Delta x^p|\tilde{u}(t_{i-1})|_{p+1} \\
&\le 2\gamma(1+C_{\mu,d}\Delta t_i)C\Delta x^p|u(t_{i-1})|_{p+1}.
\end{aligned}$$

We further split the first term in the upper bound and use (44):

$$\begin{aligned}
&\left\|\tilde{u}_t(t_{i,j}) - \frac{d}{dt}\tilde{u}_{\tilde{K}}(t_{i,j})\right\|_0 \\
&\le \left\|\tilde{u}_t(t_{i,j}) - (\tilde{u}_t(t_{i,j}))_{\tilde{K}}\right\|_0 + \left\|(\tilde{u}_t(t_{i,j}))_{\tilde{K}} - \frac{d}{dt}\tilde{u}_{\tilde{K}}(t_{i,j})\right\|_0 \\
&\le C(1+C_{\mu,d}\Delta t_i)\Delta x^{p+1}|u_\tau(t_{i,j})|_{p+1} \\
&\qquad + \left\|(\tilde{u}_t(t_{i,j}))_{\tilde{K}} - \frac{d}{dt}\tilde{u}_{\tilde{K}}(t_{i,j})\right\|_0.
\end{aligned}$$

The only remaining term to bound is

$$\begin{aligned}
&\left\|(\tilde{u}_t(t_{i,j}))_{\tilde{K}} - \frac{d}{dt}\tilde{u}_{\tilde{K}}(t_{i,j})\right\|_0 \\
&= \left\|\sum_{k=0}^{n}\left(\tilde{u}_t(x_k,t_{i,j}) - \frac{d}{dt}\tilde{u}_{\tilde{K}}(x_k,t_{i,j})\right)\phi_\ell(\cdot)\right\|_0 \\
&\le C|\Omega|\Delta t_i^p \max_{0\le k\le n_i}\left\{\int_{t_{i-1}}^{t_i}\left[\left(\frac{d}{dt}\right)^{p+1}\tilde{u}(x_k,t)\right]^2 dt\right\}^{1/2}, \quad (47)
\end{aligned}$$

where the Peano-Kernel theorem is used to attain the upper bound.
Thus, we have bounded each term in (43). So, combining (43)—(47) gives the desired result.

□

References

1. I. Babuška and M. R. Dorr, *Error estimates for the combined h and p versions of the finite element method*, Numer. Math., 37 (1981), pp. 257–277.
2. R. E. Bank, *PLTMG, a Software Package for Solving Elliptic Partial Differential Equations: Users' Guide 11.0*, vol. 5, Department of Mathematics, University of California, San Diego, 2012.
3. R. E. Bank, W. M. Coughran, W. Fichtner, E. H. Grosse, and R. K. Smith, *Transient simulation of silicon devices and circuits*, IEEE Trans. on Electron Devices, ED-32 (1985), pp. 1992–2005.
4. ———, *Transient simulation of silicon devices and circuits*, IEEE Trans. CAD, CAD-4 (1985), pp. 436–451.
5. R. E. Bank and M. S. Metti, *Generalized time integration schemes for space-time moving finite elements*, (in preparation).
6. R. E. Bank and H. Nguyen, *hp adaptive finite elements based on derivative recovery and superconvergence*, Computing and Visualization in Science, 14 (2012), pp. 287–299. Original Article.
7. R. E. Bank and R. F. Santos, *Analysis of some moving space-time finite element methods*, SIAM J. Numer. Anal., 30 (1993), pp. 1–18.
8. R. E. Bank, A. H. Sherman, and A. Weiser, *Refinement algorithms and data structures for regular local mesh refinement*, in Scientific Computing (Applications of Mathematics and Computing to the Physical Sciences) (R. S. Stepleman, ed.), North Holland, 1983, pp. 3–17.
9. N. N. Carlson and K. Miller, *Design and application of a gradient-weighted moving finite element code. I. In one dimension*, SIAM J. Sci. Comput., 19 (1998), pp. 728–765.

10. ———, *Design and application of a gradient-weighted moving finite element code. II. In two dimensions*, SIAM J. Sci. Comput., 19 (1998), pp. 766–798.
11. S. DHARMARAJA, Y. WANG, AND G. STRANG, *Optimal stability for trapezoidal–backward difference split-steps*, IMA J. Numer. Anal., 30 (2010), pp. 141–148.
12. T. DUPONT, *Mesh modification for evolution equations*, Math. Comp., 39 (1982), pp. 85–107.
13. T. F. DUPONT AND I. MOGULTAY, *A symmetric error estimate for Galerkin approximations of time-dependent Navier-Stokes equations in two dimensions*, Math. Comp., 78 (2009), pp. 1919–1927.
14. W. HUNDSDORFER AND J. VERWER, *Numerical solution of time-dependent advection-diffusion-reaction equations*, vol. 33 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2003.
15. A. P. KUPRAT, *Creation and Annihilation of Nodes for the Moving Finite Element Method*, PhD thesis, Department of Mathematics, University of California, Berkeley, CA, 1992.
16. Y. LIU, R. E. BANK, T. F. DUPONT, S. GARCIA, AND R. F. SANTOS, *Symmetric error estimates for moving mesh mixed methods for advection-diffusion equations*, SIAM J. Numer. Anal., 40 (2002), pp. 2270–2291 (electronic) (2003).
17. M. S. METTI, *Analysis of Some Higher Order Space-Time Moving Finite Element Methods*, PhD thesis, Department of Mathematics, University of California, San Diego, CA, 2013.
18. K. MILLER, *Moving finite elements.* II, SIAM J. Numer. Anal., 18 (1981), pp. 1033–1057.
19. K. MILLER AND R. N. MILLER, *Moving finite elements.* I, SIAM J. Numer. Anal., 18 (1981), pp. 1019–1032.
20. S. E. NOTARIS, *The error norm of Gauss-Radau quadrature formulae for Chebyshev weight functions*, BIT, 50 (2010), pp. 123–147.
21. R. F. SANTOS, *Moving Space-Time Finite Element Methods for Convection-Diffusion Problems*, PhD thesis, Department of Mathematics, University of California, San Diego, CA, 1991.
22. L. F. SHAMPINE, I. GLADWELL, AND S. THOMPSON, *Solving ODEs with MATLAB*, Cambridge University Press, Cambridge, 2003.