TIME-DISCRETE HIGHER ORDER ALE FORMULATIONS: A PRIORI ERROR ANALYSIS

ANDREA BONITO, IRENE KYZA, AND RICARDO H. NOCHETTO

Abstract. We derive optimal a priori error estimates for discontinuous Galerkin (dG) time discrete schemes of any order applied to an advection-diffusion model defined on moving domains and written in the Arbitrary Lagrangian Eulerian (ALE) framework. Our estimates hold without any restrictions on the time steps for dG with exact integration or Reynolds' quadrature. They involve a mild restriction on the time steps for the practical Runge-Kutta-Radau (RKR) methods of any order. The key ingredients are the stability results shown earlier in [6] along with a novel ALE projection. Numerical experiments illustrate and complement our theoretical results.

1. Introduction

Problems governed by partial differential equations (PDEs) on deformable domains $\Omega_t \subset \mathbb{R}^d$, which change in time $0 \leq t \leq T < \infty$, are of fundamental importance in science and engineering, especially for space dimensions $d \geq 2$. The boundary $\partial\Omega_t$ of Ω_t may move according to a law given a priori (moving boundary) or a law we need to solve for (free boundary). The latter are of course more common and much more challenging to study theoretically and solve numerically (e.g. fluid-structure interactions). The *Arbitrary Lagrangian Eulerian (ALE) approach* was introduced in [10, 19, 20] to prevent excessive mesh distortion within the Lagrangian approach. The mesh boundary is deformed according to the prescribed boundary velocity $\mathbf{w}$, but an *arbitrary*, yet adequate, extension is used to perform the bulk deformation. This extension of $\mathbf{w}$ from $\partial\Omega_t$ to Ω_t can be performed using various techniques such as solving for a suitable boundary value problem with Dirichlet boundary condition $\mathbf{w}$; see [13, 24, 16, 23] and the references therein. This extension induces a map $\mathcal{A}_t : \Omega_0 \to \Omega_t$, the so-called *ALE map*, with the key property that

$$\mathbf{w}(\mathbf{x}, t) = \partial_t \mathcal{A}_t(\mathbf{y}), \qquad \mathbf{x} = \mathcal{A}_t(\mathbf{y}).$$

The ALE velocity $\mathbf{w}$ is unrelated to the fluid velocity $\mathbf{b}$ and dictated mostly by the geometric principle of preserving mesh regularity. This fact should be reflected in the error analysis.

The ALE approach has been introduced in [19] for finite difference numerical schemes and in [10, 20] for finite element schemes. Second order in time finite element schemes, such as Crank-Nicolson and BDF finite element schemes have been proposed by Formaggia and Nobile, [15, 14, 22], Boffi and Gastaldi, [4, 16] and by Badia and Codina [2]. They considered the following time-dependent diffusion-advection model problem defined on moving domains:

$$\begin{cases} \partial_t u + \nabla_{\mathbf{x}} \cdot (\mathbf{b}u) - \mu \Delta_{\mathbf{x}} u = f & \mathbf{x} \in \Omega_t, \ t \in [0, T] \\ u(\mathbf{x}, t) = 0 & \mathbf{x} \in \partial\Omega_t, \ t \in [0, T] \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}) & \mathbf{x} \in \Omega_0; \end{cases} \tag{1.1}$$

Date: September 9, 2012.

Key words and phrases. ALE formulations, moving domains, domain velocity, material derivative, dG methods in time, RKR methods, a priori error analysis, ALE projection, Reynolds' quadrature .

A.B. was partially supported by NSF Grant DMS-0914977 and by Award No. KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST). I.K. was partially supported by NSF Grants DMS-0807811 and DMS-0807815. R.H.N. was partially supported by NSF Grants DMS-0807811 and DMS-1109325.

$\mathbf{b}$ is a (divergence-free) convective velocity, $\mu > 0$ is a constant diffusion parameter, f is a forcing term and u_0 is the initial condition. In order to write (1.1) in the ALE framework, we consider for $t = 0, \Omega_0 \subseteq \mathbb{R}^d$ as the reference domain with Lipschitz boundary $\partial\Omega_0$, and let $\Omega_t \subseteq \mathbb{R}^d$ be a moving domain at time $t \in (0, T]$. We assume that the family of maps $\{\mathcal{A}_t\}_{t\in[0,T]}$ is bi-Lipschitz, that $\mathcal{A}_0 := I_d$ is the identity map, and that $\Omega_t = \mathcal{A}_t(\Omega_0)$. Therefore, each point $\mathbf{y}$ from the reference domain Ω_0 is mapped through $\mathcal{A}_t$ to the corresponding point $\mathbf{x} \in \Omega_t$, i.e., the map $\mathcal{A}_t$ is given by

$$\mathcal{A}_t : \Omega_0 \subseteq \mathbb{R}^d \to \Omega_t \subseteq \mathbb{R}^d, \quad \mathbf{x}(\mathbf{y}, t) = \mathcal{A}_t(\mathbf{y}).$$

We frequently regard $\mathcal{A}_t$ as a space-time function $\mathcal{A}(\mathbf{y}, t) := \mathcal{A}_t(\mathbf{y})$, and we refer to $\mathbf{y} \in \Omega_0$ as the ALE coordinate and $\mathbf{x} = \mathbf{x}(\mathbf{y}, t)$ as the spatial or Eulerian coordinate. The space-time domain is

$$\mathcal{Q}_T := \big\{(\mathbf{x}, t) \in \mathbb{R} : t \in [0, T], \mathbf{x} = \mathcal{A}_t(\mathbf{y}), \mathbf{y} \in \Omega_0\big\}.$$

The *domain velocity* $\widehat{\mathbf{w}} : \Omega_0 \times [0, T] \to \mathbb{R}^d$ in the ALE frame is given by

$$\widehat{\mathbf{w}}(\mathbf{y}, t) := \partial_t \mathbf{x}(\mathbf{y}, t),$$

and we indicate by $\mathbf{w} : \mathcal{Q}_T \to \mathbb{R}^d$ the corresponding function on the Eulerian frame. We always denote by ∂_t the usual weak partial derivative in time, namely keeping the space variable constant. Given a function $g : \mathcal{Q}_T \to \mathbb{R}$, $D_t g$ denotes the *material or ALE time-derivative* [6, Lemma 2.1] :

$$(D_t g)(\mathbf{x}, t) := (\partial_t g)(\mathcal{A}_t(\mathbf{y}), t) = \partial_t g(\mathbf{x}, t) + \mathbf{w}(\mathbf{x}, t) \cdot \nabla_{\mathbf{x}} g(\mathbf{x}, t).$$

If $\nabla_{\mathbf{x}} \cdot \mathbf{b} = 0$, then we can now rewrite problem (1.1) in the ALE framework, equivalently, as follows

$$\begin{cases} D_t u + (\mathbf{b} - \mathbf{w}) \cdot \nabla_{\mathbf{x}} u - \mu \Delta_{\mathbf{x}} u = f & \text{in } \mathcal{Q}_T, \\ u = 0 & \text{on } \partial\mathcal{Q}_T, \\ u(\cdot, 0) = u_0 & \text{in } \Omega_0, \end{cases} \tag{1.2}$$

in *non-conservative* form. For $\tau, t \in [0, T]$ with $\tau < t$, the variational formulation of (1.2) reads

$$\int_\tau^t \langle D_s u, v\rangle_{\Omega_s}\, ds + \int_\tau^t \langle (\mathbf{b} - \mathbf{w}) \cdot \nabla_{\mathbf{x}} u, v\rangle_{\Omega_s}\, ds + \mu \int_\tau^t \langle \nabla_{\mathbf{x}} u, \nabla_{\mathbf{x}} v\rangle_{\Omega_s}\, ds = \int_\tau^t \langle f, v\rangle_{\Omega_s}\, ds. \tag{1.3}$$

At this point we emphasize that we use test functions that do not have vanishing material derivative. We refer to [6] for a discussion of existence, uniqueness and regularity of the solution of (1.1).

Despite the important fact that for realistic simulations involving fluids in 3D the ALE method should to be at least second order in time, such methods are very limited in the literature. The aforementioned methods are known to be unconditionally stable (no restriction in the time steps) and lead to optimal order a priori error bounds only when applied to the counterpart of (1.1) on time-independent domains. However, this is not the case for moving domains for which only conditional stability and error bounds are proven. The restriction on the time steps is related to the nature of the ALE map $\mathcal{A}_t$, thereby being more serious for oscillatory ALE maps, even though $\mathbf{w}$ is absent in (1.1) and plays just a geometric role. These results were observed even for numerical schemes that satisfy the so called *geometric conservation law (GCL)*, namely schemes that reproduce exactly a constant solution [4, 12, 15, 14, 16, 17, 18, 22].

Since the main obstruction to higher order accuracy is the time discretization, we continue here our analysis of time-discrete discontinuous Galerkin methods (dG) with exact integration and suitable quadrature started in [6]. Exploiting the variational structure of dG, we have been able to prove that the scheme with exact integration is ALE-free stable (stability constants independent of the ALE map) and unconditionally stable, as well as the equivalence between conservative and non-conservative ALE formulations. This is also true for dG with Reynolds' quadrature provided the ALE map is a piecewise polynomial in time. The key ingredient is the *Reynolds' identity*

$$\frac{d}{dt} \int_{\Omega_t} v\, d\mathbf{x} = \int_{\Omega_t} \big(D_t v + v \nabla_{\mathbf{x}} \cdot \mathbf{w}\big)\, d\mathbf{x}, \tag{1.4}$$

whose discrete satisfaction entails essential cancellations. Reynolds' quadrature enforces (1.4), generalizes the concept of GCL, and is responsible for ALE-free stability [6]. We have also been able to examine the practical Runge-Kutta-Radau (RKR) methods, which are dG methods with Radau quadrature, and show their conditional stability for any order, under a mild ALE-dependent restriction on the time steps. This is due to violation of (1.4) which, in turn, leads to a practical scheme with minimal complexity. Before moving on, we note that Reynolds' methods are more appropriate in cases of highly oscillatory ALE maps because they lead to ALE-free stable schemes. In contrast, the minimal complexity of RKR methods makes them more appropriate in cases of non-oscillatory maps when the time-step requirement is less restrictive".

The aim in this paper is to provide optimal order a priori error estimates for the dG methods proposed in [6]. Our results hinge on the following two main ingredients:

- *Stability* of the dG schemes developed in [6], and minor variations in Section 2.4 to account for quadrature;
- The novel *ALE projection* of Section 3 which is designed not only to provide optimal approximability but also to preserve a key property related to Reynolds' formula.

The error estimates follow from Lax's axiom that stability plus consistency gives convergence. We resort to the usual implementation of this principle within variational methods [26]: we exploit the stability of dG for the discrete function $U - Pu$, with U being the dG solution and Pu the ALE projection of u. This is carried out in Sections 4 and 6. The main contributions of this paper are:

- *Optimal order a priori error estimates* for dG of degree $q \geq 0$ with exact integration and Reynolds' quadrature, the latter for piecewise polynomial ALE maps $\mathcal{A}_t$, and for dG with Radau quadrature. The first a priori error bounds for dG applied to (1.1) were derived by Jamet [21] but without quadrature and so for a non-practical method. More recently, Chrysafinos and Walkington [8] have examined dG for an advection-dominated diffusion problem of the form (1.1) with emphasis on the hyperbolic regime and assuming exact integration.
- *Geometric role of* $\mathbf{w}$: The ALE velocity $\mathbf{w}$ is designed to preserve mesh regularity in the ALE framework, and is absent in the physical problem (1.1). Hence, its role in the stability and a priori error analyses is different from the pure advection $\mathbf{b}$. This is a distinctive feature of our estimates, which are a consequence of [6] and contain constants depending on $\mathcal{A}_t$ of order 1 rather than exponential.
- *Time step restrictions*: Our results are valid without time step constraints for all $q \geq 0$ and dG with exact intergration and Reynolds' quadrature. This is because we exploit the discrete satisfaction of Reynolds' identity (1.4) in the error analysis. With the exception of Euler method ($q = 0$), there are no methods for $q \geq 1$ known with this property. Since Radau quadrature violates (1.4), the ensuing Runge-Kutta-Radau methods for $q \geq 0$ possess error estimates with a mild time step restriction depending on μ and $\mathcal{A}_t$.
- *ALE projection*: Our study of this projection, which extends the usual dG projection, reveals that space and time are tangled together and cannot be separated without further regularity assumption of $\mathcal{A}_t$ in space beyond Lipschitz continuity. We explore this feature with a counterexample and give a complete error analysis provided $\mathcal{A}_t \in W^2_\infty$ for a.e. $t \in (0, T)$.
- *Piecewise polynomial approximation of the ALE map*: We discuss a mathematical model that mimics practical ALE formulations, for which only the ALE velocity $\mathbf{w}$ is given on the boundary $\partial\Omega_t$ rather than the entire ALE map $\mathcal{A}_t$ in Ω_0. We reconstruct $\mathcal{A}_t$ by time integration and suitable extension inside Ω_t and evaluate the error committed in this process. We deduce that we need piecewise polynomial interpolation of $\mathbf{w}$ of degree q to match the dG accuracy.

This paper is organized as follows. In Section 2 we recall the methods introduced in [6] along with stability bounds that are necessary for the error analysis; in particular, in Subsection 2.4 we discuss an alternative stability bound useful to tackle the effect of quadrature. We devote Section 3 to the construction of the ALE projection Pu and proofs of its approximation properties in time and space. In Section 4 we derive *unconditional* (no time step constraints) optimal order a priori error estimates for dG methods of any order $q \geq 0$ with exact integration and with Reynolds' quadrature, the latter provided the ALE map is piecewise polynomial in time. In Section 5 we examine the practical issue of approximating in time the ALE velocity $\mathbf{w}$ and extending it inside the domain, and also prove error estimates induced by this process. The required mild time-step restriction in this section is needed to guarantee that the discrete-time ALE map indeed satisfies the one-to-one requirement (2.1) below. Finally, we conclude in Section 6 with *conditional* optimal order a priori error bounds for the RKR methods, which are practical dG methods. The time step constraint is rather mild and depends on the diffusion μ and the domain velocity $\mathbf{w}$. We corroborate the sharpness of our error estimates with several numerical experiments.

2. The dG Method

In this section we recall the discontinuous Galerkin methods (dG) discussed in [6] along with their stability estimates. The latter will be needed for the error analyses of Sections 4 and 6.

2.1. **The ALE Framework.** For any domain D of $\mathbb{R}^m$, $m = d$ or $d+1$, we denote by $W^\ell_\kappa(D)$ the standard Sobolev spaces with integrability $1 \leq \kappa \leq \infty$ and differentiability $0 \leq \ell < \infty$. We use the notation $L^\kappa(D)$ when $\ell = 0$ and $H^\ell(D)$ when $r = 2$ and $\ell \geq 1$. With $H^1_0(D)$ we denote the subspace of $H^1(D)$ consisting of functions with vanishing trace and equipped with the norm $\|\nabla_{\mathbf{x}} v\|_{L^2(D)}$; we denote its dual by $H^{-1}(D)$. We indicate with $\langle \cdot, \cdot \rangle_D$ both the $H^1_0 - H^{-1}$ duality pairing and the L^2-inner product in D, depending on the context. Spaces of vector-valued functions are written in bold-face. For $Y = W^\ell_\kappa$, $\ell \geq 0, 1 \leq \kappa \leq \infty$, H^1_0, or H^{-1}, we define the spaces

$$L^2(Y; \mathcal{Q}_T) := \Big\{ v : \mathcal{Q}_T \to \mathbb{R} : \int_0^T \|v(t)\|^2_{Y(\Omega_t)}\, dt < \infty \Big\}.$$

We define accordingly the spaces $C(Y; \mathcal{Q}_T)$ of continuous functions with values in Y, and set

$$L^\infty(\mathrm{div}; \mathcal{Q}_T) := \{ \mathbf{c} : \mathcal{Q}_T \to \mathbb{R}^d : \operatorname{ess\,sup}_{t \in (0,T)} \big(\|\mathbf{c}(t)\|_{\mathbf{L}^\infty(\Omega_t)} + \|\nabla_{\mathbf{x}} \cdot \mathbf{c}(t)\|_{L^\infty(\Omega_t)} \big) < \infty \}.$$

To simplify the notation we omit writing the dependency in $\mathcal{Q}_T$ when there is no confusion.

We say that $\{\mathcal{A}_t\}_{t \in [0,T]}$ is a family of *ALE maps* if the following two conditions are satisfied [6]

- Regularity: $\mathcal{A}(\cdot,\cdot) \in \mathbf{W}^1_\infty\big((0,T); \mathbf{W}^1_\infty(\Omega_0)\big)$;
- Injectivity: there exists a constant $\lambda > 0$ such that for all $t \in [0,T]$,

$$\|\mathcal{A}_t(\mathbf{y}_1) - \mathcal{A}_t(\mathbf{y}_2)\| \geq \lambda \|\mathbf{y}_1 - \mathbf{y}_2\|, \qquad \forall\, \mathbf{y}_1,\ \mathbf{y}_2 \in \Omega_0, \tag{2.1}$$

for some norm $\|\cdot\|$ in $\mathbb{R}^d$. Note that the one-to-one requirement implies that $\mathcal{A}_t : \Omega_0 \to \Omega_t$ is invertible with Lipschitz inverse, i.e., $\mathcal{A}_t$ is bi-Lipschitz and so a homeomorphism. This, in particular, implies that $\widehat{v} \in H^1_0(\Omega_0)$ if and only if $v = \widehat{v} \circ \mathcal{A}_t^{-1} \in H^1_0(\Omega_t)$; cf. [15, Proposition 1]. Moreover, since Ω_0 is Lipschitz so is Ω_t and $\Omega_t \subset \Omega$ for some bounded domain Ω, for all $t \in [0,T]$. Hence, the Poincaré inequality in Ω implies the existence of an absolute constant C_Ω, independent of t, so that

$$\|v\|_{L^2(\Omega_t)} \leq C_\Omega \|\nabla_{\mathbf{x}} v\|_{\mathbf{L}^2(\Omega_t)}, \quad \forall\, v \in H^1_0(\Omega_t). \tag{2.2}$$

In addition, since the Jacobian matrix of $\mathcal{A}_t$, $\mathbf{J}_{\mathcal{A}_t} := \nabla_{\mathbf{y}} \mathcal{A}_t$, is Lipschitz in time with $\mathcal{A}_0 = I_d$ and invertible for $t \in [0,T]$, we deduce that

$$\frac{d}{dt} \det \mathbf{J}_{\mathcal{A}_t}(\mathbf{y},t) = \nabla_{\mathbf{x}} \cdot \mathbf{w}(\mathcal{A}_t(\mathbf{y}),t) \det \mathbf{J}_{\mathcal{A}_t}(\mathbf{y},t) \quad \Rightarrow \quad \det \mathbf{J}_{\mathcal{A}_t}(\mathbf{y},t) = \mathrm{e}^{\int_0^t \nabla_{\mathbf{x}} \cdot \mathbf{w}(\mathcal{A}_s(\mathbf{y}),s)\, ds}. \tag{2.3}$$

As a consequence, $\det \mathbf{J}_{\mathcal{A}_t}$ is positive and bounded away from 0 and ∞ uniformly for $t \in [0,T]$. Additional regularity on the ALE map needed for the error analysis will be specified later.

Sometimes later it will be more convenient to use Ω_τ, $\tau \in (0,T]$ as reference domain rather than Ω_0. In such a case, the letter $\mathbf{y} \in \Omega_\tau$ will still indicate points in the reference domain and the letter $\mathbf{x} \in \Omega_t$ indicate points in any other domain Ω_t, $t \in [0,T] \setminus \{\tau\}$. Moreover, for $\tau, s \in [0,T]$, we denote by $\mathcal{A}_{\tau\to s} : \Omega_\tau \to \Omega_s$ the map $\mathcal{A}_{\tau\to s} := \mathcal{A}_s \circ \mathcal{A}_\tau^{-1}$, whence $\mathcal{A}_s = \mathcal{A}_{0\to s}$. Taking Ω_τ, $\tau \in [0,T]$ as the reference domain, to every function $g : \mathcal{Q}_T \to \mathbb{R}$ we associate the function $\widehat{g} : \Omega_\tau \times [0,T] \to \mathbb{R}$ defined by $\widehat{g}(\mathbf{y},t) := g\big(\mathcal{A}_{\tau\to t}(\mathbf{y}),t\big)$.

To discretize we let $0 =: t_0 < t_1 < \cdots < t_N := T$ be a partition of $[0,T]$ and for $n = 0,1,\ldots,N-1$, let $I_n := (t_n, t_{n+1}]$, $k_n := t_{n+1} - t_n$ be the variable time-steps, $k := \max\limits_{0\le n\le N-1} k_n$, and

$$\mathcal{Q}_n := \{(x,t) \in \mathcal{Q}_T : \, t \in I_n\}.$$

For $q \ge 0$, the discrete space of order $q+1$ for dG in time on moving domains, is defined as follows:

$$\mathcal{V}_q := \{V : \mathcal{Q}_T \to \mathbb{R} : \, V|_{I_n} = \sum_{j=0}^{q} t^j \varphi_j \text{ where } \varphi_j \in L^2(H_0^1) \text{ with } D_t\varphi_j = 0, \; j = 0,1,\ldots,q\}.$$

This definition of $\mathcal{V}_q$ ensures that whenever $V \in \mathcal{V}_q$, $\widehat{V}(\mathbf{y},t) = V(\mathbf{x},t)$ is defined on a reference domain and is a piecewise polynomial of degree at most q in t with coefficients in H_0^1. We also define

$$\mathcal{V}_q(I_n) := \{V : \mathcal{Q}_n \to \mathbb{R} : \, V = W|_{\mathcal{Q}_n}, \, W \in \mathcal{V}_q\},$$

for $n = 0,1,\ldots,N-1$, consisting of restrictions to $\mathcal{Q}_n$ of functions in $\mathcal{V}_q$. In the forthcoming analysis there will be constants depending explicitly on $\nabla_{\mathbf{y}}\mathcal{A}_t$, the space differential of the ALE map $\mathcal{A}_t$, and its time derivatives. They may change at each appearance and be multiplied by other constants depending on the polynomial degree q and the space dimension d. To simplify the notation, and make it clear that the constants are explicit we now introduce these characteristic constants:

$$\begin{aligned} A_n &:= \|\nabla_{\mathbf{y}}\mathcal{A}_{t_m\to t}\|^{\kappa}_{\mathbf{L}^\infty(I_n;\mathbf{L}^\infty(\Omega_{t_m}))} \|(\nabla_{\mathbf{y}}\mathcal{A}_{t_m\to t})^{-1}\|^{\ell}_{\mathbf{L}^\infty(I_n;\mathbf{L}^\infty(\Omega_{t_m}))}, \\ B_{n,j} &:= \|\nabla_{\mathbf{y}}\mathcal{A}_{t_n\to t}\|_{\mathbf{W}^j_\infty(I_n;\mathbf{L}^\infty(\Omega_{t_n}))}, \quad G_{n,j} := A_n B_{n,j}, \quad 0 \le j \le q, \end{aligned} \tag{2.4}$$

where $m = n, n+1$ and the powers $\kappa, \ell \ge 0$ with $\kappa + \ell \ge 1$ will not be specified, but they can be equal to $0, 1, d, d+2$ depending on the context . We do not specify the norm used in (2.4) for the finite dimensional space $\mathbb{R}^{d\times d}$ due to their equivalence. Since $\mathcal{A}_{t_m\to t} = \mathcal{A}_t \circ \mathcal{A}_{t_m}^{-1}$, we realize that

$$\lim_{t\to t_m} \|\nabla_{\mathbf{y}}\mathcal{A}_{t_m\to t}\|_{\mathbf{L}^\infty(\Omega_{t_m})} = \|I_d\|_{\mathbf{L}^\infty(\Omega_{t_m})} = 1,$$

whence the constants $A_n, B_{n,j} = O(1)$ are local and do not involve exponentials of either geometric quantities or T; see [6] for more details. We also emphasize that constants $B_{n,j}$ and $G_{n,j}$, which require additional regularity in time for the ALE map, are used in the a priori error analysis to guarantee that quadrature errors are indeed of optimal order of accuracy.

We will often use the notation $\lesssim$ to indicate absolute constants depending only on the polynomial degree q, the space dimension d and the constant C_Ω in (2.2).

Finally, we consider semidiscrete schemes with discretization only in time. Thus, for $n = 0,1,\ldots,N-1$, $\mathcal{V}_q(I_n)$ is not a finite-dimensional space. This follows a similar approach to semidiscrete dG for time-independent domains [26, Chapter 12].

Remark 2.1 (Finite Dimensionality)**.** Any function $V \in \mathcal{V}_q(I_n)$ is a polynomial in time of degree at most q when viewed on a reference domain Ω_{t_n}. Specifically, the quantity

$$I_n \ni t \mapsto \|\widehat{V}(t)\|^2_{\mathcal{H}(\Omega_{t_n})} \tag{2.5}$$

is a polynomial in time of degree at most $2q$ and the Hilbert space $\mathcal{H}(\Omega_{t_n})$ is either $L^2(\Omega_{t_n})$ or $H_0^1(\Omega_{t_n})$. Therefore, finite-dimensional arguments such as inverse inequalities ([7, Chapter 4, Lemma 4.5.3]) and the equivalence of norms in finite-dimensional spaces of polynomials, can be applied to (2.5). Quantities of the form (2.5) appear often in our subsequent analysis.

2.2. **dG with Exact Integration.** The dG approximation U to u with respect to the *non-conservative* ALE formulation (1.3) is defined as follows: we seek a $U \in \mathcal{V}_q$ such that

$$U(\cdot,0) = u_0 \quad \text{in} \quad \Omega_0, \tag{2.6}$$

and for $n = 0, 1, \ldots, N-1$,

$$\begin{aligned} &\int_{I_n} \langle D_t U, V\rangle_{\Omega_t}\, dt + \langle U(t_n^+) - U(t_n), V(t_n^+)\rangle_{\Omega_{t_n}} + \int_{I_n} \langle (\mathbf{b}-\mathbf{w})\cdot\nabla_{\mathbf{x}} U, V\rangle_{\Omega_t}\, dt \\ &\qquad + \mu \int_{I_n} \langle \nabla_{\mathbf{x}} U, \nabla_{\mathbf{x}} V\rangle_{\Omega_t}\, dt = \int_{I_n} \langle f, V\rangle_{\Omega_t}\, dt, \quad \forall\, V \in \mathcal{V}_q(I_n). \end{aligned} \tag{2.7}$$

The key relation used in [6] for the proof of the well-posedness of U and of the unconditional stability of the scheme (2.7) is the following consequence of the Reynolds' identity (1.4)

$$\int_{I_n} \Big(\langle D_t V, V\rangle_{\Omega_t} - \langle \mathbf{w}\cdot\nabla_{\mathbf{x}} V, V\rangle_{\Omega_t} \Big) dt = \frac{1}{2}\|V(t_{n+1})\|^2_{L^2(\Omega_{t_{n+1}})} - \frac{1}{2}\|V(t_n^+)\|^2_{L^2(\Omega_{t_n})}, \tag{2.8}$$

which is valid for all $V \in \mathcal{V}_q(I_n)$, $n = 0, 1, \ldots, N-1$. Identity (2.8) leads also to the equivalence of non-conservative and conservative formulations, as well as the following stability bound [6].

Lemma 2.1 (stability of dG with exact integration)**.** *The solution $U \in \mathcal{V}_q$ of (2.7) satisfies*

$$\begin{aligned} &\|U(t_n)\|^2_{L^2(\Omega_{t_n})} + \sum_{m=0}^{n-1} \|U(t_m^+) - U(t_m)\|^2_{L^2(\Omega_{t_m})} + \mu \int_0^{t_n} \|\nabla_{\mathbf{x}} U(t)\|^2_{L^2(\Omega_t)}\, dt \\ &\qquad \le \|U(0)\|^2_{L^2(\Omega_0)} + \frac{1}{\mu}\int_0^{t_n} \|f(t)\|^2_{H^{-1}(\Omega_t)}\, dt, \quad \forall\, n = 1, \cdots, N. \end{aligned} \tag{2.9}$$

The variational structure of (2.7) is convenient from the theoretical point of view, but (2.7) is not a practical method. We now recall the use of quadrature, explored in [6], for moving domains.

2.3. **dG with Quadrature.** Given $r \ge q$, we let $Q(g) := \sum_{j=0}^r \omega_j g(\tau_j)$ denote a quadrature on $[0,1]$ with positive weights ω_j and nodes $\tau_j \in (0,1]$. The induced quadrature over $I_n = (t_n, t_{n+1}]$ is denoted by $Q_n(g) := \sum_{j=0}^r \omega_{n,j} g(t_{n,j})$ with weights $\{\omega_{n,j}\}_{j=0}^r$ and nodes $\{t_{n,j}\}_{j=0}^r$ in I_n given by

$$\omega_{n,j} = k_n \omega_j, \quad t_{n,j} = t_n + k_n \tau_j, \quad j = 0, 1, \ldots, r,$$

for all $n = 0, 1, \ldots, N-1$. The non-conservative dG-scheme (2.7) with quadrature reads

$$\begin{aligned} &Q_n(\langle D_t U, V\rangle_{\Omega_t}) + \langle U(t_n^+) - U(t_n), V(t_n^+)\rangle_{\Omega_{t_n}} + Q_n(\langle (\mathbf{b}-\mathbf{w})\cdot\nabla_{\mathbf{x}} U, V\rangle_{\Omega_t}) \\ &\qquad + \mu Q_n(\langle \nabla_{\mathbf{x}} U, \nabla_{\mathbf{x}} V\rangle_{\Omega_t}) = Q_n(\langle f, V\rangle_{\Omega_t}),\ \forall\, V \in \mathcal{V}_q(I_n),\ n = 0, 1, \ldots, N-1. \end{aligned} \tag{2.10}$$

We first assume that $\mathcal{A}_t$ is a continuous piecewise polynomial ALE map of degree q' in time. The so-called *Reynolds' quadratures* are quadratures exact for polynomials of degree p with

$$p := 2q + \max\{dq' - 1, 0\}, \tag{2.11}$$

and generalize the geometric conservation law (GCL) to any polynomial degree q. For the error analysis we need Reynolds' quadratures satisfying

$$q \le r \le p - q; \tag{2.12}$$

cf. Lemma 2.2 below. Reynolds' quadratures satisfying (2.12) do exist. We could use $r+1$ Radau quadrature points with $r = q + [\frac{dq'}{2}]$, where $[\cdot]$ denotes the integer part. Consequently, $q \le r \le p - q = q + \max\{dq' - 1, 0\}$, and this quadrature is exact for polynomials of degree less or equal than $2r = 2q + 2[\frac{dq'}{2}] \ge 2q + \max\{dq' - 1, 0\} = p$. On the other hand, we could use $r+1$ Gauss points with r as before. This is because such a quadrature is exact for polynomials of degree less or equal than $2r + 1 \ge p$.

Our choice of Reynolds' quadrature hinges on the fact that it retains the Reynolds' identity (2.8)

$$Q_n\big(\langle D_t V - \mathbf{w}\cdot\nabla_{\mathbf{x}} V, V\rangle_{\Omega_t}\big) = \frac{1}{2}\|V(t_{n+1})\|^2_{L^2(\Omega_{t_{n+1}})} - \frac{1}{2}\|V(t_n^+)\|^2_{L^2(\Omega_{t_n})}, \qquad \forall V \in \mathcal{V}_q(I_n); \tag{2.13}$$

see [6] for a proof. This leads to a stability bound similar to (2.9).

Our second choice is Radau quadrature with $q+1$ nodes, namely $r = q$, which is exact for polynomials of degree $2q$ and is the minimal one that preserves stability properties of dG for non-moving domains. The resulting scheme (2.10) is the Runge-Kutta-Radau (RKR) method of order q on moving domains. Unfortunately, Reynolds' identity (2.13) is no longer true, which leads to a stability bound similar to (2.9) but with the mild time step constraint

$$A_n(1 + B_{n,2})k_n \lesssim \mu, \qquad \forall\, 1 \le n \le N, \tag{2.14}$$

provided $\mathcal{A}_{t_n\to t} \in \mathbf{W}^2_\infty\big(I_n; \mathbf{W}^1_\infty(\Omega_{t_n})\big)$; this compensates for the ensuing quadrature error [6]. RKR methods are important for several reasons. First the number of Radau nodes is $q+1$, independently of the dimension d, whereas the number of nodes for Reynolds' quadrature is $r+1$ with $r \ge q$ depending on d. This shows that Radau quadrature is cheaper than Reynolds' quadrature, but at the expense of (2.14), which appears to be more significant for smaller diffusion μ. Secondly, Radau quadrature with $q+1$ nodes is the most natural for dG on non-moving domains, so our results extend previous ones to moving domains in the ALE framework. The last and perhaps most important reason is that Radau quadrature does not require the ALE map to be piecewise polynomial as Reynolds' quadrature does, thereby avoiding the approximation of the ALE map discussed in Section 5.

2.4. **The Effect of Quadrature.** We start by recalling the following equivalence of norms

$$\frac{1}{A_n} Q_n\big(\|\nabla_{\mathbf{x}} V(t)\|^2_{\mathbf{L}^2(\Omega_t)}\big) \le \int_{I_n} \|\nabla_{\mathbf{x}} V(t)\|^2_{\mathbf{L}^2(\Omega_t)}\, dt \le A_n Q_n\big(\|\nabla_{\mathbf{x}} V(t)\|^2_{\mathbf{L}^2(\Omega_t)}\big) \tag{2.15}$$

for all $V \in \mathcal{V}_q(I_n)$, where the constant A_n is defined in (2.4) [6].

It turns out that to account for the quadrature error in the error analysis below we need a slightly more general form than $Q_n(\langle f, V\rangle_{\Omega_t})$ for the right-hand side in (2.10). We now explore the consequences to stability after adding to $Q_n(\langle f, V\rangle_{\Omega_t})$ a term like

$$E_n\big(\langle \mathbf{g}, \nabla_{\mathbf{x}} V\rangle_{\Omega_t}\big) := \int_{I_n} \langle \mathbf{g}, \nabla_{\mathbf{x}} V\rangle_{\Omega_t}\, dt - Q_n\big(\langle \mathbf{g}, \nabla_{\mathbf{x}} V\rangle_{\Omega_t}\big), \qquad V \in \mathcal{V}_q(I_n),$$

with $\widehat{\mathbf{g}} \in \mathbf{H}^{j+1}\big(I_n; \mathbf{L}^2(\Omega_{t_n})\big)$ for $j = 0, 1, \ldots, q$. Since functions in $\mathbf{H}^{j+1}\big(I_n; \mathbf{L}^2(\Omega_{t_n})\big)$ are at least Hölder $1/2$ in time for all $j \ge 0$, with values in $\mathbf{L}^2(\Omega_{t_n})$, the quantity $Q_n(\langle \mathbf{g}, \nabla_{\mathbf{x}} V\rangle_{\Omega_t})$ is well defined. We have the following error estimate.

Lemma 2.2 (effect of quadrature)**.** *Let $q \ge 0$ and Q_n be a any quadrature exact for polynomials of degree $q + r$ with $r \ge q$. If $\mathcal{A}_{t_n\to t} \in \mathbf{W}^{j+1}_\infty\big(I_n; \mathbf{W}^1_\infty(\Omega_{t_n})\big)$ and $\widehat{\mathbf{g}} \in \mathbf{H}^{j+1}\big(I_n; \mathbf{L}^2(\Omega_{t_n})\big)$ then, for all $j = 0, 1, \ldots, q$, $V \in \mathcal{V}_q(I_n)$ and all $\varepsilon > 0$, we have*

$$\Big|E_n\big(\langle \mathbf{g}, \nabla_{\mathbf{x}} V\rangle_{\Omega_t}\big)\Big| \le \frac{G_{n,j+1}}{\varepsilon\mu} k_n^{2(j+1)} \sum_{i=0}^{j+1} \int_{I_n} \|D_t^i \mathbf{g}(t)\|^2_{L^2(\Omega_t)} dt + \varepsilon\mu Q_n\big(\|\nabla_{\mathbf{x}} V(t)\|^2_{L^2(\Omega_t)}\big). \tag{2.16}$$

Proof. We denote by $\widehat{\mathbf{I}}_r$ the vector-valued Lagrange interpolation operator at the quadrature nodes of I_n and use the fact that Q_n is exact for polynomials of degree $q + r$ to write $Q_n(\langle \mathbf{g}, \nabla_{\mathbf{x}} V\rangle_{\Omega_t})$ on the reference domain Ω_{t_n} as follows:

$$Q_n\big(\langle \mathbf{g}, \nabla_{\mathbf{x}} V\rangle_{\Omega_t}\big) = \sum_{j=0}^{r} \omega_{n,j} \langle \widehat{\mathbf{h}}(t_{n,j}), \nabla_{\mathbf{y}} \widehat{V}(t_{n,j})\rangle_{\Omega_{t_n}} = \int_{I_n} \langle \widehat{\mathbf{I}}_r \widehat{\mathbf{h}}, \nabla_{\mathbf{y}} \widehat{V}\rangle_{\Omega_{t_n}}\, dt,$$

with $\widehat{\mathbf{h}}(t) = \det \mathbf{J}_{\mathcal{A}_{t_n\to t}} \mathbf{J}^{-1}_{\mathcal{A}_{t_n\to t}} \widehat{\mathbf{g}}(t)$. Since $r \geq q \geq j$, applying polynomial interpolation theory we deduce

$$\Big|E_n\big(\langle \mathbf{g}, \nabla_{\mathbf{x}} V\rangle_{\Omega_t}\big)\Big| \leq \int_{I_n} \Big|\langle \widehat{\mathbf{h}} - \widehat{\mathbf{I}}_r \widehat{\mathbf{h}}, \nabla_{\mathbf{y}} \widehat{V}\rangle_{\Omega_{t_n}}\Big| \, dt \lesssim k_n^{j+1} \int_{I_n} \|\partial_t^{j+1} \widehat{\mathbf{h}}(t)\|_{\mathbf{L}^2(\Omega_{t_n})} \|\nabla_{\mathbf{y}} \widehat{V}(t)\|_{\mathbf{L}^2(\Omega_{t_n})} dt.$$

The desired estimate (2.16) follows from using the Leibniz rule for differentiating $\widehat{\mathbf{h}}$, invoking the stated regularity for $\mathbf{g}$ and the ALE maps $\mathcal{A}_{t_n\to t}$, as well as (2.15). □

We conclude this section with a stability estimate that accounts for $E_n\big(\langle g, V\rangle_{\Omega_t}\big)$.

Lemma 2.3 (stability of dG with quadrature)**.** *Let $U \in \mathcal{V}_q$ be the dG solution of* (2.10)*, with $f \in C(H^{-1}; \mathcal{Q}_T) \cap L^2(\mathcal{Q}_T)$ and $\widehat{\mathbf{g}} \in \mathbf{H}^{j+1}\big(I_n; \mathbf{L}^2(\Omega_{t_n})\big)$. Then, for all $j = 0, 1, \ldots, q$ and $0 < n \leq N$, U satisfies*

$$\begin{aligned} &\|U(t_n)\|^2_{L^2(\Omega_{t_n})} + \sum_{m=0}^{n-1} \|U(t_m^+) - U(t_m)\|^2_{L^2(\Omega_{t_m})} + \frac{\mu}{2} \sum_{m=0}^{n-1} Q_j(\|\nabla_{\mathbf{x}} U(t)\|^2_{\mathbf{L}^2(\Omega_t)}) \\ &\leq \|U(0)\|^2_{L^2(\Omega_0)} + \frac{2}{\mu} \sum_{m=0}^{n-1} Q_j(\|f(t)\|^2_{H^{-1}(\Omega_t)}) + \frac{G_{n,j+1}}{\mu} k_n^{2(j+1)} \sum_{i=0}^{j+1} \int_{I_n} \|D_t^i \mathbf{g}(t)\|^2_{L^2(\Omega_t)} dt, \end{aligned} \tag{2.17}$$

for any time step k_n if Q_n is Reynolds' quadrature and for k_n subject to (2.14) *if Q_n is Radau quadrature with $q+1$ nodes.*

Proof. This is a minor modification of Theorem 4.1 (Reynolds' quadrature) and Theorem 5.1 (Radau quadrature) of [6]. Upon taking $V = U$ we realize that the only new term is $\varepsilon \mu Q_n(\|\nabla_{\mathbf{x}} V(t)\|^2_{\mathbf{L}^2(\Omega_t)})$ of (2.16), which can be absorbed on the left hand side of (2.17) with $\varepsilon = 1/4$. □

3. The ALE Projection

Since u does not, in general, belong to the subspace $\mathcal{V}_q$, in order to prove a priori error estimates, we introduce the *ALE projection* operator $P : C(H_0^1; \mathcal{Q}_T) \to \mathcal{V}_q$ which is defined as follows:

$$Pu(\cdot, 0) = u(0), \quad \text{in} \ \ \Omega_0, \tag{3.1}$$

and for $n = 0, 1, \ldots, N-1$,

$$Pu(\cdot, t_{n+1}) = u(\cdot, t_{n+1}), \quad \text{in} \ \ \Omega_{t_{n+1}} \tag{3.2}$$

and

$$\int_{I_n} \langle Pu - u, V\rangle_{\Omega_t} \, dt = 0, \quad \forall\, V \in \mathcal{V}_{q-1}(I_n). \tag{3.3}$$

Note that for $q = 0$ the condition (3.3) is void, and $\|Pu\|_{L^2(\Omega_t)}$ is discontinuous at $t = t_n$ for $q \geq 0$ and $0 \leq n < N$. We prove below that P is well defined by (3.2)-(3.3) and that has good approximation properties (cf. (3.9) and (3.10)). However, the present proofs are much more technical than those in non-moving domains, because integration in time and space in (3.3) do not commute.

Before embarking on the study of the ALE projection P, we present a simple but fundamental consequence of its definition, which makes it quite useful in the subsequent error analysis.

Lemma 3.1 (key property of P)**.** *For every $0 \leq n < N$, the error $\rho := u - Pu$ satisfies*

$$\int_{I_n} \langle D_t \rho, V\rangle_{\Omega_t} \, dt + \langle \rho(t_n^+) - \rho(t_n), V(t_n^+)\rangle_{\Omega_{t_n}} + \int_{I_n} \langle \rho \nabla_{\mathbf{x}} \cdot \mathbf{w}, V\rangle_{\Omega_t} \, dt = 0, \quad \forall\, V \in \mathcal{V}_q(I_n). \tag{3.4}$$

Proof. Given $V \in \mathcal{V}_q(I_n)$, Reynolds' identity (1.4) yields

$$\int_{I_n} \langle D_t \rho, V\rangle_{\Omega_t}\, dt = -\int_{I_n} \langle \rho, D_t V + V\nabla_{\mathbf{x}} \cdot \mathbf{w}\rangle_{\Omega_t}\, dt + \int_{I_n} \frac{d}{dt}\langle \rho, V\rangle_{\Omega_t}\, dt,$$

whence, using that $\rho(\cdot, t_{n+1}) = 0$ in $\Omega_{t_{n+1}}$ because of (3.2),

$$\int_{I_n} \langle D_t \rho, V\rangle_{\Omega_t}\, dt = -\int_{I_n} \langle \rho, D_t V + V\nabla_{\mathbf{x}} \cdot \mathbf{w}\rangle_{\Omega_t}\, dt - \langle \rho(t_n^+), V(t_n^+)\rangle_{\Omega_{t_n}}.$$

Since $D_t V \in \mathcal{V}_{q-1}(I_n)$, (3.3) implies $\int_{I_n} \langle \rho, D_t V\rangle_{\Omega_t}\, dt = 0$, which in turn gives (3.4) because $\rho(\cdot, t_n) = 0$ in Ω_{t_n}. $\square$

We are now ready to examine P. We start with a rather elementary proof of *uniqueness*: if $V_1, V_2 \in \mathcal{V}_q$ are two functions satisfying (3.2)-(3.3), then $V_1 - V_2 = (t_{n+1} - t)V$ with $V \in \mathcal{V}_{q-1}(I_n)$, whence $V = 0$ from (3.3). This would suffice for existence if $\mathcal{V}_q$ was finite dimensional.

To tackle existence of the ALE projection, we first dwell on the regularity issue $Pu \in \mathcal{V}_q$, namely that $Pu(t) \in H_0^1(\Omega_t)$ for all $t \in I_n$, provided $u \in C(H_0^1; \mathcal{Q}_n)$. This is automatic for time independent domains, even though (3.2)-(3.3) is just an L^2-type projection. Indeed, the commutativity between time integration and space differentiation implies $\nabla_{\mathbf{x}} Pu = P\nabla_{\mathbf{x}} u \in \mathbf{C}(\mathbf{L}^2; \mathcal{Q}_n)$ [26]. However, the ALE projection is not a pure time-projection if the ALE map $\mathcal{A}_t$ is not the identity, which leads to tangling of time and space. The question discussed in the following remark is whether W^1_∞-regularity of $\mathcal{A}_t$ would suffice.

Remark 3.1 (Counterexample)**.** We show that if the ALE map is just a W^1_∞-function in space, then the ALE projection may not be stable in $C(H_0^1; \mathcal{Q}_n)$. We consider the one-dimensional case $d = 1$, the fixed domain $\Omega = (-1, 1)$ and $T = 1$. We assume that

$$\det \mathbf{J}_{\mathcal{A}_t}(y) = \frac{\partial}{\partial y}\mathcal{A}_t(y) := \varphi(y)t + 1, \qquad \varphi(y) = \alpha|y|^\gamma,$$

with $0 < \alpha < 1$ and $0 < \gamma < \frac{1}{2}$, and that the polynomial degree is $q = 1$. Note that in this setting the domain is not changing in time. Let $\widehat{g}(y,t) = (1 - y^2)(1 - t)^2$ which satisfies $g \in C\big((0,1); H_0^1(\Omega)\big)$. In order to find the ALE projection Pg of g, we rewrite (3.3) on the reference domain Ω as $\int_0^1 \int_\Omega \big(\widehat{g} - \widehat{Pg}\big)(y,t)v(y) \det \mathbf{J}_{\mathcal{A}_t}(y)\, dt\, dy = 0$ for all $v \in H_0^1(\Omega)$, whence

$$\int_0^1 (\widehat{g}(y,t) - \widehat{Pg}(y,t)) \det \mathbf{J}_{\mathcal{A}_t}(y)\, dt = 0, \quad \forall y \in (-1, 1).$$

If $\widehat{Pu}(y,t) = \beta(y)(1-t)$, then tedious algebraic calculations lead to $\beta(y) = \frac{1}{2}(1-y^2)\big(1 + \frac{1}{\varphi(y)+3}\big)$, and $\beta'(y) = -y - \big[2y\big(\varphi(y)+3\big) + (1+y^2)\varphi'(y)\big]\big[2\big(\varphi(y)+3\big)^2\big]^{-1}$. In particular, we obtain

$$\big|\beta'(y)\big|^2 = G(y) + \frac{\big(\varphi'(y)\big)^2}{4\big(\varphi(y)+3\big)^4}, \tag{3.5}$$

for a continuous function G on $[-1, 1]$. Since $\varphi(y) + 3$ is continuous and positive, it is thus bounded away from zero in [-1,1]. However, $\varphi'(y)$ is not not square integrable in $[-1, 1]$, and neither is $\beta'(y)$. This shows that the ALE projection $\widehat{Pu} \notin H_0^1(\Omega)$ for any $t \in (0, 1]$.

The above remark implies that additional regularity assumption on the ALE map beyond Lipschitz continuity in space is needed to guarantee that $Pu(t) \in H_0^1(\Omega_t)$. We will assume the sufficient (and convenient) condition

$$\mathcal{A}_{t_{n+1}\to t} \in \mathbf{L}^\infty\big(I_n; \mathbf{W}^2_\infty(\Omega_{t_{n+1}})\big), \qquad n = 0, 1, \dots N-1. \tag{3.6}$$

This might seem to rule out finite element maps, which are just piecewise W^2_∞ in space. We postpone until Remark 5.1 a discussion on how to overcome this obstacle.

Proposition 3.1 (existence of the ALE projection)**.** *If the family of ALE maps satisfies* (3.6)*, then there exists a unique function $Pu \in \mathcal{V}_q$ such that* (3.1)*-*(3.3) *is valid for all $u \in C(H_0^1; \mathcal{Q}_T)$.*

Proof. We split the proof into four steps.
Step 1: Formulation in reference domain. We fix $n = 0, 1, \dots, N-1$, and use $\Omega_{t_{n+1}}$ as a reference domain. We recall that for $t \in I_n$, $\mathbf{y} \in \Omega_{t_{n+1}}$, we have defined

$$\widehat{v}(\mathbf{y}, t) = v\big(\mathcal{A}_{t_{n+1}\to t}(\mathbf{y}), t\big), \quad \forall v \in L^2(\mathcal{Q}_n).$$

If $\widehat{\mathcal{V}}_q(I_n)$ is the space of functions $\widehat{V}$ so that $V \in \mathcal{V}_q(I_n)$, then (3.2)-(3.3) is equivalent to finding $\widehat{U}_n \in \widehat{\mathcal{V}}_q(I_n)$ so that

$$\widehat{U}_n(\mathbf{y}, t_{n+1}) = \widehat{u}(\mathbf{y}, t_{n+1}), \quad \mathbf{y} \in \Omega_{t_{n+1}}$$

$$\int_{I_n}\int_{\Omega_{t_{n+1}}} \big(\widehat{U}_n(\mathbf{y}, t) - \widehat{u}(\mathbf{y}, t)\big)\widehat{V}(\mathbf{y}, t)\det(\mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}})(\mathbf{y}, t)\,d\mathbf{y}\,dt = 0, \quad \forall \widehat{V} \in \widehat{\mathcal{V}}_{q-1}(I_n). \tag{3.7}$$

Step 2: Existence and L^2-regularity. Let $\widehat{\mathcal{W}}_q(I_n)$ be defined as $\widehat{\mathcal{V}}_q(I_n)$ with the difference that H_0^1 is replaced by L^2. Then $\widehat{\mathcal{W}}_q(I_n)$ is a Hilbert space with respect to the inner product $(\cdot,\cdot)_n$ defined as

$$(\widehat{W}, \widehat{V})_n := \int_{I_n}\int_{\Omega_{t_{n+1}}} (t_{n+1} - t)\widehat{W}\widehat{V}\det\mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}\,d\mathbf{y}\,dt, \quad \forall \widehat{W}, \widehat{V} \in \widehat{\mathcal{W}}_q(I_n).$$

By Riesz representation Theorem, for every $\widehat{v}_n \in L^2\big(I_n; L^2(\Omega_{t_{n+1}})\big)$, there exists a unique $\widehat{V}_n \in \widehat{\mathcal{W}}_{q-1}(I_n)$ such that

$$(\widehat{V}_n, \widehat{W})_n = \int_{I_n} \langle \widehat{v}_n \det\mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}, \widehat{W}\rangle_{\Omega_{t_{n+1}}}, \quad \forall \widehat{W} \in \widehat{\mathcal{W}}_{q-1}(I_n).$$

Consequently, if $\widehat{v}_n := -\widehat{u}(\mathbf{y}, t_{n+1}) + \widehat{u}|_{I_n} \in L^2\big(I_n; L^2(\Omega_{t_{n+1}})\big)$, then

$$\widehat{U}_n := \widehat{u}(\mathbf{y}, t_{n+1}) + (t_{n+1} - t)\widehat{V}_n$$

satisfies (3.7) for all $\widehat{V} \in \widehat{\mathcal{W}}_{q-1}$. It remains to show that $\widehat{U}_n \in \widehat{\mathcal{V}}_q$.
Step 3: Representation of $\widehat{V}_n$. We can write $\widehat{V}_n = \sum_{j=0}^{q-1}\widehat{V}_{n,j}(t_{n+1} - t)^j$ with $\widehat{V}_{n,j} \in L^2(\Omega_{t_{n+1}})$, $j = 0, \dots, q-1$, because $\widehat{V}_n \in \widehat{\mathcal{W}}_{q-1}$. We intend to derive a matrix equation for the vector $\widehat{\mathbf{V}}_n := (\widehat{V}_{n,j})_{j=0}^{q-1}$. To this end, we rewrite (3.7) as follows: for all $i = 0, 1, \dots, q-1$:

$$\int_{\Omega_{t_{n+1}}} v \int_{I_n} \big((t_{n+1} - t)\widehat{V}_n - \widehat{v}_n\big)(t_{n+1} - t)^i \det\mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}} dt d\mathbf{y} = 0, \ \forall v \in L^2(\Omega_{t_{n+1}}).$$

This in turn can be equivalently reformulated in matrix form, for a.e. $\mathbf{y} \in \Omega_{t_{n+1}}$:

$$\mathbf{A}_n(\mathbf{y})\widehat{\mathbf{V}}_n(\mathbf{y}) = \widehat{\mathbf{v}}_n(\mathbf{y}), \tag{3.8}$$

with

$$\mathbf{A}_n(\mathbf{y})_{i,j} := \int_{I_n} (t_{n+1} - t)^{(i-1)+(j-1)+1} \det\mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}(\mathbf{y}, t)\,dt$$

and

$$\widehat{\mathbf{v}}_n(\mathbf{y})_i := \int_{I_n} \widehat{v}_n(\mathbf{y}, t)(t_{n+1} - t)^{i-1} \det\mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}(\mathbf{y}, t)\,dt,$$

for $i, j = 1, \cdots, q$.
Step 4: H_0^1-regularity. We first observe that $\mathcal{A}_{t_{n+1}\to t} \in \mathbf{L}^\infty\big(I_n; \mathbf{W}_\infty^2(\Omega_{t_{n+1}}\big)$ yields $\det\mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}} \in L^\infty\big(I_n; W_\infty^1(\Omega_{t_{n+1}})\big)$, whence $\widehat{\mathbf{v}}_n \in \mathbf{H}_0^1(\Omega_{t_{n+1}})$ because $\widehat{v}_n \in L^2(I_n; H_0^1(\Omega_{t_{n+1}}))$. Moreover, we also deduce $\mathbf{A}_n \in [W_\infty^1(\Omega_{t_{n+1}})]^{q\times q}$; in particular $\mathbf{A}_n$ is Lipschitz continuous and thus well defined for all $\mathbf{y} \in \overline{\Omega}_{t_{n+1}}$. We now verify that $\mathbf{A}_n$ is invertible and $\mathbf{A}_n^{-1}$ is also Lipschitz.

It suffices to prove that, for fixed $\mathbf{y} \in \overline{\Omega}_{t_{n+1}}$, the linear system

$$\mathbf{A}_n(\mathbf{y})\mathbf{X} = \mathbf{0}, \text{ with } \mathbf{X} = (X_i)_{i=0}^q$$

has only the trivial solution. In fact, we have $\mathbf{A}_n(\mathbf{y})\mathbf{X} = \mathbf{0}$ is equivalent to

$$\sum_{j=1}^q X_j \int_{I_n} (t_{n+1} - t)^{(j-1)+(i-1)+1} \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}(\mathbf{y}, t)\, dt = 0, \quad \forall\, i = 1, \dots, q.$$

So,

$$\int_{I_n} \Big(\sum_{j=1}^q X_j (t_{n+1} - t)^{j-1}\Big)\Big(\sum_{i=1}^q X_i (t_{n+1} - t)^{i-1}\Big)(t_{n+1} - t) \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}(\mathbf{y}, t)\, dt = 0,$$

whence we immediately arrive at $\sum_{j=1}^q X_j (t_{n+1} - t)^{j-1} = 0$, and thus $X_1 = \cdots = X_q = 0$. We infer that $\mathbf{A}_n(\mathbf{y})$ is invertible for every $\mathbf{y} \in \overline{\Omega}_{t_{n+1}}$, and that (3.8) is equivalently written as

$$\widehat{\mathbf{V}}_n(\mathbf{y}) = \mathbf{A}_n^{-1}(\mathbf{y})\widehat{\mathbf{v}}_n(\mathbf{y}).$$

Since $\det \mathbf{A}_n > 0$ is Lipschitz and $\overline{\Omega}_{t_{n+1}}$ is compact, $\det \mathbf{A}_n > 0$ is uniformly bounded away from 0. Therefore

$$\mathbf{A}_n^{-1} = \frac{1}{\det \mathbf{A}_n} \operatorname{cof}(\mathbf{A}_n) \in [W^1_\infty(\Omega_{t_{n+1}})]^{q\times q}$$

and $\widehat{\mathbf{V}}_n = \mathbf{A}_n^{-1}\widehat{\mathbf{v}}_n \in \mathbf{H}^1_0(\Omega_{t_{n+1}})$, as asserted. This concludes the proof. □

We next prove that the ALE projection Pu satisfies suitable approximation properties. As in Proposition 3.1, we will need the extra space regularity (3.6) and the corresponding local constants:

$$M_n := \|\mathcal{A}_{t_n\to t}\|_{\mathbf{L}^\infty\left(I_n;\mathbf{W}^2_\infty(\Omega_{t_n})\right)}, \quad n = 0, 1, \dots, N-1.$$

Proposition 3.2 (approximation properties of the ALE projection)**.** *If the family of ALE maps satisfies* (3.6)*, then the following approximation properties of the ALE projection P are valid for $j = 0, 1, \dots, q$:*

$$\|(u - Pu)(t)\|^2_{L^2(\Omega_t)} \le C_n k_n^{2j+1} \int_{I_n} \|D_t^{j+1} u(t)\|^2_{L^2(\Omega_t)}\, dt, \tag{3.9}$$

$$\|\nabla_{\mathbf{x}}(u - Pu)(t)\|^2_{\mathbf{L}^2(\Omega_t)} \le D_n k_n^{2j+1} \int_{I_n} \Big(\|D_t^{j+1} u(t)\|^2_{L^2(\Omega_t)} + \|\nabla_{\mathbf{x}} D_t^{j+1} u(t)\|^2_{\mathbf{L}^2(\Omega_t)}\Big)\, dt, \tag{3.10}$$

where the constant C_n is proportional to $A_n^3 + A_n$ and D_n is proportional to $(1+M_n^2)A_n^6 + M_n^2 A_n^3 + A_n^2$.

Proof. We compare Pu with the standard dG projection [1, 26]. For $0 \le n < N$, we consider $\Omega_{t_{n+1}}$ as a reference domain, and let the dG projection $\widehat{\Psi}_u \in \widehat{\mathcal{V}}_q(I_n)$ of $\widehat{u}$ be defined by

$$\widehat{\Psi}_u(\mathbf{y}, t_{n+1}) = \widehat{u}(\mathbf{y}, t_{n+1}), \quad \mathbf{y} \in \Omega_{t_{n+1}}, \tag{3.11}$$

$$\int_{I_n} \langle \widehat{u} - \widehat{\Psi}_u, \widehat{V}\rangle_{\Omega_{t_{n+1}}}\, dt = 0, \quad \forall \widehat{V} \in \widehat{\mathcal{V}}_{q-1}(I_n); \tag{3.12}$$

the latter condition is void for $q = 0$. It is well known that $\widehat{\Psi}_u$ is well defined and satisfies the following approximation properties for $j = 0, 1, \dots, q$ (cf. [26, p. 207–208] or [1]):

$$\|(\widehat{u} - \widehat{\Psi}_u)(t)\|^2_{L^2(\Omega_{t_{n+1}})} \lesssim k_n^{2j+1} \int_{I_n} \|\partial_t^{j+1}\widehat{u}(s)\|^2_{L^2(\Omega_{t_{n+1}})}\, ds, \quad t \in I_n, \tag{3.13}$$

$$\|\nabla_{\mathbf{y}}(\widehat{u} - \widehat{\Psi}_u)(t)\|^2_{\mathbf{L}^2(\Omega_{t_{n+1}})} \lesssim k_n^{2j+1} \int_{I_n} \|\nabla_{\mathbf{y}}\partial_t^{j+1}\widehat{u}(s)\|^2_{\mathbf{L}^2(\Omega_{t_{n+1}})}\, ds, \quad t \in I_n. \tag{3.14}$$

It is then obvious that $\Psi_u \in \mathcal{V}_q(I_n)$ with

$$\Psi_u(\cdot, t_{n+1}) = Pu(\cdot, t_{n+1}) \quad \text{in} \quad \Omega_{t_{n+1}},$$

whence $\Psi_u - Pu = (t_{n+1} - t)W$ for some $W \in \mathcal{V}_{q-1}(I_n)$, and that (3.3) implies

$$\int_{I_n} \langle \Psi_u - Pu, V\rangle_{\Omega_t}\, dt = \int_{I_n} \langle \Psi_u - u, V\rangle_{\Omega_t}\, dt, \quad \forall\, V \in \mathcal{V}_{q-1}(I_n).$$

Upon taking $V = W$, we thus deduce

$$\begin{aligned}\int_{I_n} \|(\Psi_u - Pu)(t)\|^2_{L^2(\Omega_t)}\, dt &= \int_{I_n} \langle \Psi_u - Pu, (t_{n+1} - t)W\rangle\, dt \le k_n \int_{I_n} \langle \Psi_u - Pu, W\rangle_{\Omega_t}\, dt \\ &= k_n \int_{I_n} \langle \Psi_u - u, W\rangle_{\Omega_t}\, dt \le k_n \Big(\int_{I_n} \|(u - \Psi_u)(t)\|^2_{L^2(\Omega_t)}\, dt\Big)^{\frac{1}{2}} \Big(\int_{I_n} \|W(t)\|^2_{L^2(\Omega_t)}\, dt\Big)^{\frac{1}{2}}.\end{aligned} \tag{3.15}$$

On the other hand, we note that

$$\begin{aligned}k_n^2 \int_{I_n} \|V(t)\|^2_{L^2(\Omega_t)}\, dt &= k_n^2 \int_{I_n} \int_{\Omega_{t_{n+1}}} \widehat{V}^2 \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}} \\ &\le k_n^2 \sup_{t\in I_n} \|\det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}\|_{L^\infty(\Omega_{t_{n+1}})} \int_{I_n} \|\widehat{V}(t)\|^2_{L^2(\Omega_{t_{n+1}})}\, dt\end{aligned}$$

Therefore, the equivalence of norms in the finite dimensional space of polynomials of degree $\le 2q-2$ (see Remark 2.1), ensures that for all $V \in \mathcal{V}_{q-1}(I_n)$,

$$\begin{aligned}k_n^2 \int_{I_n} \|\widehat{V}(t)\|^2_{L^2(\Omega_{t_{n+1}})}\, dt &\lesssim \int_{I_n} (t_{n+1} - t)^2 \|\widehat{V}(t)\|^2_{L^2(\Omega_{t_{n+1}})}\, dt \\ &\lesssim \sup_{t\in I_n} \|(\det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}})^{-1}\|_{L^\infty(\Omega_{t_{n+1}})} \int_{I_n} (t_{n+1} - t)^2 \|V(t)\|^2_{L^2(\Omega_t)}\, dt\end{aligned}$$

Hence,

$$k_n^2 \int_{I_n} \|V(t)\|^2_{L^2(\Omega_t)}\, dt \lesssim A_n \int_{I_n} (t_{n+1} - t)^2 \|V(t)\|^2_{L^2(\Omega_t)}\, dt, \tag{3.16}$$

where the constant A_n is defined in (2.4). Combining estimates (3.15), (3.16), and using again the equivalence of norms in the space of polynomials of degree $\le q$, we get

$$\|(\Psi_u - Pu)(t)\|^2_{L^2(\Omega_t)} \lesssim \frac{A_n}{k_n} \int_{I_n} \|(\Psi_u - Pu)(t)\|^2_{L^2(\Omega_t)}\, dt \lesssim \frac{A_n^2}{k_n} \int_{I_n} \|(u - \Psi_u)(t)\|^2_{L^2(\Omega_t)}\, dt. \tag{3.17}$$

This, in conjunction with (3.13) and $(u - Pu)^2 \le 2(u - \Psi_u)^2 + 2(\Psi_u - Pu)^2$, gives the desired (3.9).

For the proof of (3.10) we face once again the lack of commutativity between space and time integration. We thus proceed as in Proposition 3.1 by first writing (3.3) in $\Omega_{t_{n+1}}$ with $\widehat{U}_n = \widehat{Pu}$

$$\int_{I_n} \langle (\widehat{u} - \widehat{U}_n) \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}, \widehat{V}\rangle_{\Omega_{t_{n+1}}}\, dt = 0, \quad \forall\, \widehat{V} \in \widehat{\mathcal{V}}_{q-1}(I_n),$$

which is equivalent to

$$\int_{I_n} (\widehat{u} - \widehat{U}_n) t^j \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}\, dt = 0, \quad j = 0, 1, \dots, q-1,$$

and differentiating in space each term on the left-hand side, both being in $H^1_0(\Omega_{t_{n+1}})$,

$$\int_{I_n} \nabla_{\mathbf{y}} \big((\widehat{u} - \widehat{U}_n) \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}\big) t^j\, dt = \mathbf{0}, \quad j = 0, 1, \dots, q-1.$$

This implies

$$\int_{I_n} \langle \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}} \nabla_{\mathbf{y}}(\widehat{u}-\widehat{U}_n) + (\widehat{u}-\widehat{U}_n)\nabla_{\mathbf{y}} \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}, \nabla_{\mathbf{y}}\widehat{V}\rangle_{\Omega_{t_{n+1}}}\, dt = 0, \quad \forall \widehat{V} \in \widehat{\mathcal{V}}_{q-1}(I_n),$$

whence,

$$\begin{aligned}\int_{I_n} \langle \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}} \nabla_{\mathbf{y}}(\widehat{\Psi}_u-\widehat{U}_n), \nabla_{\mathbf{y}}\widehat{V}\rangle_{\Omega_{t_{n+1}}}\, dt &= \int_{I_n} \langle \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}} \nabla_{\mathbf{y}}(\widehat{\Psi}_u-\widehat{u}), \nabla_{\mathbf{y}}\widehat{V}\rangle_{\Omega_{t_{n+1}}}\, dt \\ &+ \int_{I_n} \langle (\widehat{U}_n-\widehat{u})\nabla_{\mathbf{y}} \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}}, \nabla_{\mathbf{y}}\widehat{V}\rangle_{\Omega_{t_{n+1}}}\, dt,\end{aligned}$$

for all $\widehat{V} \in \widehat{\mathcal{V}}_{q-1}(I_n)$. Furthermore, since $\nabla_{\mathbf{y}}\widehat{v}(\mathbf{y}) = \mathbf{J}^T_{\mathcal{A}_{t_{n+1}\to t}}\nabla_{\mathbf{x}} v(\mathbf{x})$ for any $v \in H^1_0(\Omega_t)$ and $\widehat{\Psi}_u - \widehat{U}_n = (t_{n+1}-t)\widehat{W}$, we deduce

$$\begin{aligned}\int_{I_n} \|\nabla_{\mathbf{x}}(\Psi_u - U_n)(t)\|^2_{\mathbf{L}^2(\Omega_t)}\, dt \lesssim & k_n \|(\nabla_{\mathbf{y}}\mathcal{A}_{t_{n+1}\to t})^{-1}\|^2_{\mathbf{L}^\infty\left(I_n;\mathbf{L}^\infty(\Omega_{t_{n+1}})\right)} \times \\ & \int_{I_n} \langle \det \mathbf{J}_{\mathcal{A}_{t_{n+1}\to t}} \nabla_{\mathbf{y}}(\widehat{\Psi}_u-\widehat{U}_n), \nabla_{\mathbf{y}}\widehat{W}\rangle_{\Omega_{t_{n+1}}}\, dt.\end{aligned}$$

Taking $\widehat{V} = \widehat{W}$, invoking the relation above, and proceeding as in the proof of (3.9), we obtain

$$\int_{I_n} \|\nabla_{\mathbf{x}}(\Psi_u - U_n)\|^2_{\mathbf{L}^2(\Omega_t)}\, dt \lesssim A_n^4 \int_{I_n} \|\nabla_{\mathbf{x}}(u-\Psi_u)(t)\|^2_{\mathbf{L}^2(\Omega_t)}\, dt + M_n^2 A_n^2 \int_{I_n} \|(u-U_n)(t)\|^2_{L^2(\Omega_t)}\, dt.$$

Using again the equivalence of norms (3.17) for gradients, namely

$$\|\nabla_{\mathbf{x}}(\Psi_u - Pu)(t)\|^2_{\mathbf{L}^2(\Omega_t)} \lesssim \frac{A_n^2}{k_n}\int_{I_n} \|\nabla_{\mathbf{x}}(\Psi_u - Pu)(t)\|^2_{\mathbf{L}^2(\Omega_t)}\, dt,$$

along with the triangle and Young's inequalities, we deduce (3.10). This concludes the proof. □

Notice that Chrysafinos and Walkington introduce a projection P similar to ours in a Lagrangian framework and for a fully discrete scheme [8]. Existence of P is equivalent to uniqueness within the space-time finite-dimensional setting, but the H^1_0-error estimate (related to our (3.10)) requires a restriction between time step and meshsize, namely a CFL type condition. This approach does not apply to our semidiscrete setting.

We conclude this section with a simple stability result for time derivatives $D^j Pu$ of the ALE projection Pu. We start with following auxiliary lemma.

Lemma 3.2 (stability of time derivatives of the dG projection)**.** *If $\widehat{\Psi}_u$ is the dG projection of $\widehat{u}$ defined in* (3.11)-(3.12)*, then the following estimate holds true for $j = 1, \dots, q$,*

$$\int_{I_n} \|\partial_t^j \widehat{\Psi}_u(t)\|^2_{L^2(\Omega_{t_{n+1}})}\, dt \lesssim \int_{I_n} \|\partial_t^j \widehat{u}(t)\|^2_{L^2(\Omega_{t_{n+1}})}\, dt. \tag{3.18}$$

Proof. Let $\widehat{I}_q\widehat{u}$ be an average Taylor polynomial in time of degree q of $\widehat{u}$ (cf., e.g., [7, Chapter 4, p. 93–105]). Applying an inverse estimate, as discussed in Remark 2.1, yields

$$\int_{I_n} \|\partial_t^j(\widehat{\Psi}_u - \widehat{I}_q\widehat{u})(t)\|^2_{L^2(\Omega_{t_{n+1}})} \lesssim k_n^{-2j} \int_{I_n} \|(\widehat{\Psi}_u - \widehat{I}_q\widehat{u})(t)\|^2_{L^2(\Omega_{t_{n+1}})}\, dt.$$

Adding and subtracting $\widehat{u}$ to the right-hand side, and invoking the approximation properties of the average Taylor polynomial together with (3.13), we deduce

$$\int_{I_n} \|(\widehat{\Psi}_u - \widehat{I}_q\widehat{u})(t)\|^2_{L^2(\Omega_{t_{n+1}})}\, dt \lesssim k_n^{2j} \int_{I_n} \|\partial_t^j \widehat{u}(t)\|^2_{L^2(\Omega_{t_{n+1}})}\, dt,$$

which combined with the previous estimate and the stability bound $\int_{I_n} \|\partial_t^j \widehat{I}_q \widehat{u}(t)\|^2_{L^2(\Omega_{t_{n+1}})}\,dt \lesssim \int_{I_n} \|\partial_t^j \widehat{u}(t)\|^2_{L^2(\Omega_{t_{n+1}})}\,dt$ yields the asserted estimate. □

We use now (3.18) to derive the desired stability bound for $D^j Pu$.

Proposition 3.3 (stability of time derivatives of the ALE projection)**.** *Let Pu be the ALE projection defined in* (3.2)-(3.3)*. Then, for all $n = 0, \cdots, N-1$ and $j = 0, 1, \cdots, q$, we have*

$$\int_{I_n} \|D_t^j Pu(t)\|^2_{L^2(\Omega_t)}\,dt \lesssim C_n \int_{I_n} \|D_t^j u(t)\|^2_{L^2(\Omega_t)}\,dt. \tag{3.19}$$

Proof. After adding and subtracting $D_t^j \Psi_u$ to the left-hand side of (3.19) and changing variables to the reference domain $\Omega_{t_{n+1}}$, we realize that it suffices to estimate $\int_{I_n} \|\partial_t^j (\widehat{Pu} - \widehat{\Psi}_u)(t)\|^2_{L^2(\Omega_{t_{n+1}})}\,dt$. We now proceed as in Lemma 3.2, first using an inverse estimate and then resorting to (3.17), (3.13) and (3.18) to derive (3.19). □

4. Unconditional a priori error estimates

We are now ready to derive a priori error estimates of optimal order for the numerical schemes (2.7) and (2.10) with Reynolds' quadrature. It turns out that there are no constraints on the time steps (unconditional error estimates). We postpone to Section 6 the study of Radau quadrature.

As it is customary in the error analysis of dG, we split the error $u - U = \rho + \Theta$ with

$$\rho := u - Pu, \qquad \Theta := Pu - U.$$

In view of (3.9)-(3.10), it suffices to prove an a priori error bound for Θ. Since $\Theta \in \mathcal{V}_q$ is discrete, this will be achieved by finding first the equation satisfied by Θ and next applying the discrete *stability* estimates of Lemmas 2.1 and 2.3.

4.1. dG with Exact Integration.

We now prove the following consequence of Lemma 2.1.

Theorem 4.1 (a priori error estimate with exact integration)**.** *Let the ALE map satisfy* (3.6)*. Let u be the solution of* (1.2) *and U be the dG solution of* (2.7)*. Then the following estimate holds*

$$\begin{aligned}
\max_{0 \le n \le N} \|(u-U)(t_n)\|^2_{L^2(\Omega_{t_n})} &+ \mu \int_0^T \|\nabla_{\mathbf{x}}(u-U)(t)\|^2_{\mathbf{L}^2(\Omega_t)}\,dt \\
&\le \frac{1}{\mu} \sum_{n=0}^{N-1} C_n k_n^{2j+2} \sup_{t \in I_n} \|(\mathbf{b}-\mathbf{w})(t)\|^2_{\mathbf{L}^\infty(\Omega_t)} \int_{I_n} \|D_t^{j+1} u(t)\|^2_{L^2(\Omega_t)}\,dt \\
&+ \mu \sum_{n=0}^{N-1} D_n k_n^{2j+2} \int_{I_n} \Big(\|D_t^{j+1} u(t)\|^2_{L^2(\Omega_t)} + \|\nabla_{\mathbf{x}} D_t^{j+1} u(t)\|^2_{\mathbf{L}^2(\Omega_t)} \Big)\,dt,
\end{aligned}$$

for all $j = 0, \cdots, q$, where C_n, D_n, $n = 0, 1, \ldots, N-1$, are constants proportional to those in (3.9) *and* (3.10)*, respectively.*

Proof. Taking $\tau = t_n$ and $t = t_{n+1}$ in (1.3), subtracting (2.7) and splitting $u - U = \rho + \Theta$, we obtain

$$\begin{aligned}
\int_{I_n} \langle D_t \Theta, V \rangle_{\Omega_t}\,dt + \langle \Theta(t_n^+) - \Theta(t_n), V(t_n^+) \rangle_{\Omega_{t_n}} + \int_{I_n} \langle (\mathbf{b}-\mathbf{w}) \cdot \nabla_{\mathbf{x}} \Theta, V \rangle_{\Omega_t}\,dt \\
+ \mu \int_{I_n} \langle \nabla_{\mathbf{x}} \Theta, \nabla_{\mathbf{x}} V \rangle_{\Omega_t}\,dt = \int_{I_n} \langle \rho(\mathbf{b}-\mathbf{w}) - \mu \nabla_{\mathbf{x}} \rho, \nabla_{\mathbf{x}} V \rangle_{\Omega_t}\,dt =: \int_{I_n} \langle g, V \rangle_{\Omega_t}\,dt,
\end{aligned}$$

where the last is a duality pairing $H^{-1}-H_0^1$. Since this equation is similar to (2.7) with f replaced by the distribution $g \in H^{-1}(\Omega_t)$, to apply Lemma 2.1 we just need to estimate $\|g(t)\|_{H^{-1}(\Omega_t)}$:

$$\|g(t)\|_{H^{-1}(\Omega_t)} = \sup_{v\in H_0^1(\Omega_t)} \frac{\langle g, v\rangle_{\Omega_t}}{\|\nabla_{\mathbf{x}} v\|_{\mathbf{L}^2(\Omega_t)}} \le \|(\mathbf{b}-\mathbf{w})(t)\|_{\mathbf{L}^\infty(\Omega_t)}\|\rho(t)\|_{L^2(\Omega_t)} + \mu\|\nabla_{\mathbf{x}}\rho(t)\|_{\mathbf{L}^2(\Omega_t)}.$$

Since $\Theta(0)=0$, the asserted estimate finally follows from Lemma 2.1 and Proposition 3.2. □

4.2. **dG with Reynolds' Quadrature.** In this subsection we prove a priori error estimates for the numerical scheme (2.10) with Reynolds' quadrature, for which we assume that the ALE map is a piecewise polynomial of degree q'.

Our first task is to find the equation satisfied by Θ. To this end, we add and subtract the ALE projection Pu to U in (2.10) and observe that the first term and third term involving $\mathbf{w}$ in (2.10) integrate exactly by design of Reynolds' quadrature. Invoking Lemma 3.1, and manipulating the resulting right-hand side properly, it is rather elementary but tedious to deduce

$$\begin{aligned}
&Q_n(\langle D_t\Theta, V\rangle_{\Omega_t}) + \langle \Theta(t_n^+)-\Theta(t_n), V(t_n^+)\rangle_{\Omega_{t_n}} + Q_n(\langle(\mathbf{b}-\mathbf{w})\cdot\nabla_{\mathbf{x}}\Theta, V\rangle_{\Omega_t})\\
&\qquad + \mu Q_n(\langle\nabla_{\mathbf{x}}\Theta, \nabla_{\mathbf{x}}V\rangle_{\Omega_t}) = Q_n(\langle(\mathbf{b}-\mathbf{w})\rho, \nabla_{\mathbf{x}}V\rangle_{\Omega_t}) - \mu Q_n(\langle\nabla_{\mathbf{x}}\rho, \nabla_{\mathbf{x}}V\rangle_{\Omega_t})\\
&\qquad + E_n(\langle(\mathbf{b}-\mathbf{w})u, \nabla_{\mathbf{x}}V\rangle_{\Omega_t}) - \mu E_n(\langle\nabla_{\mathbf{x}}u, \nabla_{\mathbf{x}}V\rangle_{\Omega_t}) + E_n(\langle f, V\rangle_{\Omega_t}),
\end{aligned} \tag{4.1}$$

for all $V \in \mathcal{V}_q(I_n)$. As expected, this equation is similar to that derived in Theorem 4.1, except that $\int_{I_n}$ is replaced by Q_n and the last three terms on the right-hand side reflect the effect of quadrature.

Theorem 4.2 (a priori error estimate with Reynolds' quadrature)**.** *Let the ALE map $\mathcal{A}_t$ be a piecewise polynomial in time of degree q' and satisfy* (3.6). *Let u be the solution of* (1.2) *and U be the dG solution of* (2.10) *with Reynolds' quadrature. Then the following a priori error estimate holds*

$$\begin{aligned}
\max_{0\le n\le N} &\|(u-U)(t_n)\|^2_{L^2(\Omega_{t_n})} + \frac{\mu}{2}\sum_{n=0}^{N-1} Q_n(\|\nabla_{\mathbf{x}}(u-U)(t)\|^2_{\mathbf{L}^2(\Omega_t)})\\
&\le \frac{1}{\mu}\sum_{n=0}^{N-1} C_n k_n^{2j+2}\sup_{t\in I_n}\|(\mathbf{b}-\mathbf{w})(t)\|^2_{\mathbf{L}^\infty(\Omega_t)}\int_{I_n}\|D_t^{j+1}u(t)\|^2_{L^2(\Omega_t)}\,dt\\
&+ \frac{\mu}{2}\sum_{n=0}^{N-1} D_n k_n^{2j+2}\int_{I_n}\Big(\|D_t^{j+1}u(t)\|^2_{L^2(\Omega_t)} + \|\nabla_{\mathbf{x}}D_t^{j+1}u(t)\|^2_{\mathbf{L}^2(\Omega_t)}\Big)\,dt\\
&+ \frac{1}{\mu}\sum_{n=0}^{N-1} G_{n,j+1}k_n^{2j+2}\sum_{i=0}^{j+1}\int_{I_n}\|D_t^i\Big((\mathbf{b}-\mathbf{w})u\Big)(t)\|^2_{\mathbf{L}^2(\Omega_t)}\,dt\\
&+ \mu\sum_{n=0}^{N-1} G_{n,j+1}k_n^{2j+2}\sum_{i=0}^{j+1}\int_{I_n}\|\nabla_{\mathbf{x}}D_t^i u(t)\|^2_{\mathbf{L}^2(\Omega_t)}\,dt\\
&+ \frac{1}{\mu}\sum_{n=0}^{N-1} G_{n,j+1}k_n^{2j+2}\sum_{i=0}^{j+1}\int_{I_n}\|D_t^i f(t)\|^2_{H^{-1}(\Omega_t)}\,dt,
\end{aligned} \tag{4.2}$$

for all $j=0,\cdots,q$, where the constants C_n, D_n, $n=0,1,\ldots,N-1$, are proportional to those in (3.9) *and* (3.10)*, respectively. Furthermore, a similar estimate can be obtained for the continuous $L^2(H^1)-$norm $\int_0^T \|\nabla_{\mathbf{x}}(u-U)(t)\|^2_{\mathbf{L}^2(\Omega_t)}\,dt$.*

Proof. We just recall (2.17) and simply estimate the right-hand side of (4.1). We handle the first two terms as we did in Theorem 4.1, thereby resorting to Proposition 3.2,

$$\frac{Q_n(\langle(\mathbf{b}-\mathbf{w})\rho,\nabla_{\mathbf{x}}V\rangle_{\Omega_t})}{Q_n(\|\nabla_{\mathbf{x}}V\|^2_{\mathbf{L}^2(\Omega_t)})^{\frac12}} \le k_n^{\frac12}\sup_{0\le j\le r}\|(\mathbf{b}-\mathbf{w})(t_{n,j})\|_{\mathbf{L}^\infty(\Omega_{t_{n,j}})}\|\rho(t_{n,j})\|_{L^2(\Omega_{t_{n,j}})}$$

$$\le C_n^{\frac12}k_n^{q+1}\sup_{t\in I_n}\|(\mathbf{b}-\mathbf{w})(t)\|_{\mathbf{L}^\infty(\Omega_t)}\Big(\int_{I_n}\|D_t^{q+1}u(t)\|^2_{L^2(\Omega_t)}\,dt\Big)^{\frac12}$$

and

$$\frac{Q_n(\langle\nabla_{\mathbf{x}}\rho,\nabla_{\mathbf{x}}V\rangle_{\Omega_t})}{Q_n(\|\nabla_{\mathbf{x}}V\|^2_{\mathbf{L}^2(\Omega_t)})^{\frac12}} \le k_n^{\frac12}\sup_{0\le j\le r}\|\nabla_{\mathbf{x}}\rho(t_{n,j})\|_{\mathbf{L}^2(\Omega_{t_{n,j}})}$$

$$\le D_n^{\frac12}k_n^{q+1}\Big(\int_{I_n}\|D_t^{q+1}u(t)\|^2_{L^2(\Omega_t)}+\|\nabla_{\mathbf{x}}D_t^{q+1}u(t)\|^2_{\mathbf{L}^2(\Omega_t)}\,dt\Big)^{\frac12}.$$

For the next two terms in (4.1) we readily see that $\mathbf{g}_1=(\mathbf{b}-\mathbf{w})u$ and $\mathbf{g}_2=-\mu\nabla_{\mathbf{x}}u$. For the last term, instead, we can always write $f=\nabla_{\mathbf{x}}\cdot\mathbf{F}$ with $\mathbf{F}\in\mathbf{L}^2(\Omega_t)$, integrate by parts to get $E_n(\langle\mathbf{F},\nabla_{\mathbf{x}}V\rangle_{\Omega_t})$, apply Lemma 2.2 with $\mathbf{g}_3=\mathbf{F}$, and minimize over $\|\mathbf{F}\|_{L^2(\Omega_t)}$ to find $\|f\|_{H^{-1}(\Omega_t)}$. Collecting these estimates, and employing (2.17) for Θ, we end up with (4.2). □

5. Approximating the ALE map

In Section 4 we have proven unconditional optimal a priori error estimates, namely without any restriction on the time steps, for the methods (2.7) and (2.10). Method (2.7) requires exact integration in time, and it is thus not practical, whereas results for (2.10) with Reynolds' quadrature are valid under the assumption that the ALE map is a piecewise polynomial in time. Despite the fact that the motion of the boundary is given, it is certainly not true that we can always describe the ALE map by a piecewise polynomial. In this section we intent to overcome this obstacle by discretizing a general ALE map in time and invoking a perturbation PDE argument. We emphasize that only the regularity of the continuous ALE map enters in the a priori error estimates. We prove a priori error estimates for the exact solution but not for the domain displacement.

5.1. **A Perturbation Argument and Error Estimates.** In this subsection we approximate the ALE map by a continuous piecewise polynomial in time of order $q'=q+1$ and we prove optimal order a priori error estimates for general ALE maps through a perturbation PDE argument. As we shall see, the only mild restriction on the time steps will be to ensure that the perturbed ALE map satisfies the one-to-one requirement with respect to $\mathbf{y}$ for $t\in(0,T]$.

To this end, we consider the same partition into subintervals I_n, $n=1,\dots,N$, as in Section 4. To reflect computational practice most closely, we approximate the domain velocity $\widehat{\mathbf{w}}$ by $\widehat{\mathbf{W}}$ using a piecewise polynomial in time of order q in the ALE frame. In turn, the approximation of the ALE map $\widetilde{\mathcal{A}}_t$ is defined by $\widetilde{\mathcal{A}}_0=I_d$ and for $n=0,\cdots,N-1$,

$$\widetilde{\mathcal{A}}_t(\mathbf{y})=\widetilde{\mathcal{A}}_{t_n}(\mathbf{y})+\int_{t_n}^t\widehat{\mathbf{W}}(\mathbf{y},s)\,ds,\qquad \widetilde{\mathbf{x}}(\mathbf{y},t):=\widetilde{\mathcal{A}}_t(\mathbf{y}). \tag{5.1}$$

To avoid any confusion, we use the symbol $\widetilde{v}$ to indicate $\widetilde{v}(\widetilde{\mathbf{x}},\cdot):=\widehat{v}\big(\widetilde{\mathcal{A}}_t^{-1}(\widetilde{\mathbf{x}}),\cdot\big)$, namely functions defined on the perturbed domain $\widetilde{\Omega}_t=\widetilde{\mathcal{A}}_t(\Omega_0)$. In this vein, $\widetilde{u},\widetilde{U}$ stand for the exact and dG solutions written over $\widetilde{\Omega}_t$.

Note that approximating the ALE map in time using (5.1) means that $\widetilde{\mathcal{A}}_{t_n}\neq\mathcal{A}_{t_n}$, and thus $\widetilde{\Omega}_{t_n}\neq\Omega_{t_n}$, in general. For this reason, we never assume that the perturbed ALE map coincides with the actual one at the nodes, but only that $\widetilde{\Omega}_{t_n}$ is "close enough" to Ω_{t_n}. We quantify this next.

In view of (5.1), we can write $\mathcal{A}_t(\mathbf{y}) = \mathbf{y} + \int_0^t \widehat{\mathbf{w}}(y,s)\,ds$ and $\widetilde{\mathcal{A}}_t(\mathbf{y}) = \mathbf{y} + \int_0^t \widehat{\mathbf{W}}(\mathbf{y},s)\,ds$, whence

$$\nabla_{\mathbf{y}}\mathcal{A}_t(\mathbf{y}) - \nabla_{\mathbf{y}}\widetilde{\mathcal{A}}_t(\mathbf{y}) = \int_0^t \nabla_{\mathbf{y}}(\widehat{\mathbf{w}} - \widehat{\mathbf{W}})(\mathbf{y},s)\,ds, \qquad \forall\, \mathbf{y} \in \Omega_0. \tag{5.2}$$

If we interpolate $\widehat{\mathbf{w}}$ with a piecewise polynomial of degree q to produce $\widehat{\mathbf{W}}$, then we can assume

$$\|\widehat{\mathbf{w}} - \widehat{\mathbf{W}}\|_{\mathbf{L}^\infty(I_n;\mathbf{W}^1_\infty(\Omega_0))} \lesssim k_n^{j+1}\|\partial_t^{j+1}\widehat{\mathbf{w}}\|_{\mathbf{L}^\infty(I_n;\mathbf{W}^1_\infty(\Omega_0))}, \tag{5.3}$$

for all $j = 0, 1, \ldots, q$. Hence, the map $\widetilde{\mathcal{A}}_t$ is piecewise polynomial of degree $q+1$. Combining (5.2) and (5.3), we get

$$\|\nabla_{\mathbf{y}}\mathcal{A}_t - \nabla_{\mathbf{y}}\widetilde{\mathcal{A}}_t\|_{\mathbf{W}^1_\infty\big((0,T);\mathbf{L}^\infty(\Omega_0)\big)} \lesssim k^{j+1}T\|\partial_t^{j+1}\widehat{\mathbf{w}}\|_{\mathbf{L}^\infty\big((0,T);\mathbf{W}^1_\infty(\Omega_0)\big)}. \tag{5.4}$$

For $\widetilde{\mathcal{A}}_t$ to be an ALE map, it remains to show that $\widetilde{\mathcal{A}}_t$ is one to one. For all $\mathbf{y}_1, \mathbf{y}_2 \in \Omega_0$ we have

$$\|\widetilde{\mathcal{A}}_t(\mathbf{y}_1) - \widetilde{\mathcal{A}}_t(\mathbf{y}_2)\| \geq \|\mathcal{A}_t(\mathbf{y}_1) - \mathcal{A}_t(\mathbf{y}_2)\| - \|(\widetilde{\mathcal{A}}_t - \mathcal{A}_t)(\mathbf{y}_1) - (\widetilde{\mathcal{A}}_t - \mathcal{A}_t)(\mathbf{y}_2)\|.$$

Since Ω_0 is Lipschitz, (5.4) implies that the difference map $\widetilde{\mathcal{A}}_t - \mathcal{A}_t$ is uniformly Lipschitz and

$$\|(\widetilde{\mathcal{A}}_t - \mathcal{A}_t)(\mathbf{y}_1) - (\widetilde{\mathcal{A}}_t - \mathcal{A}_t)(\mathbf{y}_2)\| \leq C_{\Omega_0}k^{j+1}T\|\partial_t^{j+1}\widehat{\mathbf{w}}\|_{\mathbf{L}^\infty\big((0,T);\mathbf{W}^1_\infty(\Omega_0)\big)}\|\mathbf{y}_1 - \mathbf{y}_2\|,$$

where C_{Ω_0} depends on the reference domain Ω_0 [11, p. 131–132],[25]. Consequently, if we choose the time steps k_n so that

$$C_{\Omega_0}k^{j+1}T\|\partial_t^{j+1}\widehat{\mathbf{w}}\|_{\mathbf{L}^\infty((0,T);\mathbf{W}^1_\infty(\Omega_0))} < \frac{\lambda}{2}, \tag{5.5}$$

then

$$\|\widetilde{\mathcal{A}}_t(\mathbf{y}_1) - \widetilde{\mathcal{A}}_t(\mathbf{y}_2)\| \geq \frac{\lambda}{2}\|\mathbf{y}_1 - \mathbf{y}_2\| \qquad \forall\, \mathbf{y}_1, \mathbf{y}_2 \in \Omega_0,$$

where $\lambda > 0$ is the one-to-one constant of $\mathcal{A}_t$; cf. (2.1). This does not only imply that $\widetilde{\mathcal{A}}$ is an ALE map, and so invertible, but also gives the bound

$$\|(\nabla_{\mathbf{y}}\widetilde{\mathcal{A}}_t)^{-1}\|_{\mathbf{L}^\infty((0,T);\mathbf{L}^\infty(\Omega_0))} \leq \frac{2}{\lambda}. \tag{5.6}$$

Finally, for $j = 0, 1, \ldots, q$, we introduce two global constants that will appear in the sequel:

$$\Gamma_j := A e^{k^{j+1}cTA\|\partial_t^{j+1}\widehat{\mathbf{w}}\|_{\mathbf{L}^\infty((0,T);\mathbf{W}^1_\infty(\Omega_0))}} \quad \text{with} \quad A := \frac{2T}{\lambda}\Big(1 + \frac{T}{\lambda}\|\nabla_{\mathbf{y}}\widehat{\mathbf{w}}\|_{\mathbf{L}^\infty(\Omega_0\times[0,T])}\Big), \tag{5.7}$$

where c is the constant hidden in (5.4), and $\Lambda_j = 1 + \mathcal{O}(k^{2j+2})$ given by

$$\begin{aligned}\Lambda_j := 1 &+ \Gamma_j^2T^2\|\partial_t^{j+1}\widehat{\mathbf{w}}\|^2_{\mathbf{L}^\infty((0,T);\mathbf{W}^1_\infty(\Omega_0))}k^{2j+2} \\ &+ (\lambda^{-4} + \Gamma_j^4)T^4\|\partial_t^{j+1}\widehat{\mathbf{w}}\|^2_{\mathbf{L}^\infty((0,T);\mathbf{W}^1_\infty(\Omega_0))}k^{4j+4}.\end{aligned} \tag{5.8}$$

Quantifying the effect of domain perturbation on the solution is a natural question, or equivalently

if $\widetilde{\phi}$ is the solution of (1.1) *with respect to $\widetilde{\Omega}_t$, how "large" is the error $\widetilde{u} - \widetilde{\phi}$?*

This simple question gives rise to the approximation $\widetilde{\Phi}$ of $\widetilde{\phi}$ defined as follows. Let Q_n be a Reynolds' quadrature on I_n as presented in Section 2.3. According to (2.10), the dG approximation $\widetilde{\Phi} \in \widetilde{\mathcal{V}}_q$ of $\widetilde{\phi}$ is given by $\widetilde{\Phi}(\cdot, 0) = u_0$ in Ω_0 and for $n = 0, 1, \ldots, N-1$,

$$\begin{aligned} Q_n(\langle D_t\widetilde{\Phi}, \widetilde{V}\rangle_{\widetilde{\Omega}_t}) + \langle \widetilde{\Phi}(t_n^+) - \widetilde{\Phi}(t_n), \widetilde{V}(t_n^+)\rangle_{\widetilde{\Omega}_{t_n}} + Q_n\big(\langle(\widetilde{\mathbf{b}} - \widetilde{\mathbf{W}})\cdot\nabla_{\widetilde{\mathbf{x}}}\widetilde{\Phi}, \widetilde{V}\rangle_{\widetilde{\Omega}_t}\big) \\ + \mu Q_n(\langle\nabla_{\widetilde{\mathbf{x}}}\widetilde{\Phi}, \nabla_{\widetilde{\mathbf{x}}}\widetilde{V}\rangle_{\widetilde{\Omega}_t}) = Q_n(\langle\widetilde{f}, \widetilde{V}\rangle_{\widetilde{\Omega}_t}), \quad \forall \widetilde{V} \in \widetilde{\mathcal{V}}_q(I_n). \end{aligned} \tag{5.9}$$

The dG method is now defined directly on the domain $\widetilde{\Omega}_t$ described by a piecewise polynomial in time. This is desirable because the method should (i) be computationally practical and (ii) take into account the discretization of the domain displacement in time. It is to be emphasized that the approximation $\widetilde{\Phi}$ is not the same as $\widetilde{U} = U \circ \mathcal{A}_t \circ \widetilde{\mathcal{A}}_t^{-1}$; in fact, $\widetilde{U} \notin \mathcal{V}_q$, generically.

One viable approach to estimate the error $\widetilde{u} - \widetilde{\Phi}$ would be to split $\widetilde{u} - \widetilde{\Phi} = (\widetilde{u} - \widetilde{\phi}) + (\widetilde{\phi} - \widetilde{\Phi})$. The error $\widetilde{u} - \widetilde{\phi}$ would account for the domain perturbation whereas $\widetilde{\phi} - \widetilde{\Phi}$ could be estimated using Theorem 4.2. We stress though that such an approach would lead to upper bounds depending on $\widetilde{\phi}$ (and its derivatives), instead of $\widetilde{u}$. To avoid this difficulty, we compare directly $\widetilde{u}$ and $\widetilde{\Phi}$.

We start by rewriting (1.3) in the perturbed domain $\widetilde{\Omega}_t$, $t \in [0, T]$.

Lemma 5.1. (from Ω_t to $\widetilde{\Omega}_t$.) *If $\widetilde{u}, \widetilde{\mathbf{w}}$ stand for $u, \mathbf{w}$ written in $\widetilde{\Omega}_t$, then* (1.3) *is equivalent to*

$$\int_\tau^t \langle z D_s \widetilde{u}, \widetilde{v}\rangle_{\widetilde{\Omega}_s}\, ds + \int_\tau^t \langle (\mathbf{b} - \widetilde{\mathbf{w}}) \cdot zK\nabla_{\widetilde{\mathbf{x}}} \widetilde{u}, \widetilde{v}\rangle_{\widetilde{\Omega}_s}\, ds + \mu \int_\tau^t \langle zK^{\mathrm{T}} K \nabla_{\widetilde{\mathbf{x}}} \widetilde{u}, \nabla_{\widetilde{\mathbf{x}}} \widetilde{v}\rangle_{\widetilde{\Omega}_s}\, ds = \int_\tau^t \langle z\widetilde{f}, \widetilde{v}\rangle_{\widetilde{\Omega}_s}\, ds, \tag{5.10}$$

for all $\widetilde{v} \in L^2(H_0^1; \widetilde{\mathcal{Q}}_T)$ and $\tau, t \in [0, T]$ with $\tau < t$, where $z := \det \mathbf{J}_{\mathcal{A}_t \circ \widetilde{\mathcal{A}}_t^{-1}}$ and $K := \mathbf{J}^{\mathrm{T}}_{\widetilde{\mathcal{A}}_t \circ \mathcal{A}_t^{-1}}$.

Proof. We just note that $\widetilde{\mathbf{x}} = \widetilde{\mathcal{A}}_t \circ \mathcal{A}_t^{-1}(\mathbf{x})$ and we apply the chain rule. □

Elementary algebraic manipulations of (5.10) lead to

$$\begin{aligned} &\int_\tau^t \langle D_s \widetilde{u}, \widetilde{v}\rangle_{\widetilde{\Omega}_s}\, ds + \int_\tau^t \langle (\widetilde{\mathbf{b}} - \widetilde{\mathbf{W}}) \cdot \nabla_{\widetilde{\mathbf{x}}} \widetilde{u}, \widetilde{v}\rangle_{\widetilde{\Omega}_s}\, ds + \mu \int_\tau^t \langle \nabla_{\widetilde{\mathbf{x}}} \widetilde{u}, \nabla_{\widetilde{\mathbf{x}}} \widetilde{v}\rangle_{\widetilde{\Omega}_s}\, ds \\ &\qquad = \int_\tau^t \langle \widetilde{f}, \widetilde{v}\rangle_{\widetilde{\Omega}_s}\, ds + \int_\tau^t \langle \widetilde{R}_1, \widetilde{v}\rangle_{\widetilde{\Omega}_s}\, ds + \mu \int_\tau^t \langle \widetilde{\mathrm{R}}_2, \nabla_{\widetilde{\mathbf{x}}} \widetilde{v}\rangle_{\widetilde{\Omega}_t}, \quad \forall \widetilde{v} \in L^2(H_0^1; \widetilde{\mathcal{Q}}_T), \end{aligned} \tag{5.11}$$

where

$$\widetilde{R}_1 := (1 - z) D_t \widetilde{u} + (\widetilde{\mathbf{b}} - \widetilde{\mathbf{w}}) \cdot (I_d - zK) \nabla_{\widetilde{\mathbf{x}}} \widetilde{u} + (\widetilde{\mathbf{w}} - \widetilde{\mathbf{W}}) \cdot \nabla_{\widetilde{\mathbf{x}}} \widetilde{u} + (z - 1) \widetilde{f}, \tag{5.12}$$

and

$$\widetilde{\mathrm{R}}_2 := (I_d - zK^{\mathrm{T}} K) \nabla_{\widetilde{\mathbf{x}}} \widetilde{u}. \tag{5.13}$$

Let $\widetilde{e} = \widetilde{u} - \widetilde{\Phi}$ be the error and let $\widetilde{P}\widetilde{u}$ denote the ALE projection of $\widetilde{u}$ with respect to $\widetilde{\mathcal{A}}_t$. Then $\widetilde{\rho} = \widetilde{u} - \widetilde{P}\widetilde{u}$ satisfies approximation properties similar to (3.9)-(3.10) with respect to $\widetilde{\Omega}_t$. To prove a priori error estimates for $\widetilde{e}$ we follow the same steps as in Subsection 4.2, namely, we first split the error $\widetilde{e}$ as $\widetilde{e} = \widetilde{\rho} + \widetilde{\Theta}$ with $\widetilde{\Theta} = \widetilde{P}\widetilde{u} - \widetilde{\Phi}$. Having at hand the expression (5.11) and proceeding as in the proof of (4.1) we can show that $\widetilde{\Theta}$ satisfies the following equation:

$$\begin{aligned} Q_n\big(\langle D_t \widetilde{\Theta}, \widetilde{V}\rangle_{\widetilde{\Omega}_t}\big) &+ \langle \widetilde{\Theta}(t_n^+) - \widetilde{\Theta}(t_n), \widetilde{V}(t_n^+)\rangle_{\widetilde{\Omega}_{t_n}} + Q_n\big(\langle (\widetilde{\mathbf{b}} - \widetilde{\mathbf{W}}) \cdot \nabla_{\widetilde{\mathbf{x}}} \widetilde{\Theta}, \widetilde{V}\rangle_{\widetilde{\Omega}_t}\big) \\ &+ \mu Q_n\big(\langle \nabla_{\widetilde{\mathbf{x}}} \widetilde{\Theta}, \nabla_{\widetilde{\mathbf{x}}} \widetilde{V}\rangle_{\widetilde{\Omega}_t}\big) = Q_n\big(\langle (\widetilde{\mathbf{b}} - \widetilde{\mathbf{W}}) \widetilde{\rho}, \nabla_{\widetilde{\mathbf{x}}} \widetilde{V}\rangle_{\widetilde{\Omega}_t}\big) - \mu Q_n\big(\langle \nabla_{\widetilde{\mathbf{x}}} \widetilde{\rho}, \nabla_{\widetilde{\mathbf{x}}} \widetilde{V}\rangle_{\widetilde{\Omega}_t}\big) \\ &+ E_n\big(\langle \nabla_{\widetilde{\mathbf{x}}} (\widetilde{\mathbf{b}} - \widetilde{\mathbf{W}}) \widetilde{u}, \widetilde{V}\rangle_{\widetilde{\Omega}_t}\big) - \mu E_n\big(\langle \nabla_{\widetilde{\mathbf{x}}} \widetilde{u}, \nabla_{\widetilde{\mathbf{x}}} \widetilde{V}\rangle_{\widetilde{\Omega}_t}\big) + E_n\big(\langle \widetilde{f}, \widetilde{V}\rangle_{\widetilde{\Omega}_t}\big) \\ &+ \int_{I_n} \langle \widetilde{R}_1, \widetilde{V}\rangle_{\widetilde{\Omega}_t}\, dt + \mu \int_{I_n} \langle \widetilde{\mathrm{R}}_2, \nabla_{\widetilde{\mathbf{x}}} \widetilde{V}\rangle_{\widetilde{\Omega}_t}\, dt \qquad \forall \widetilde{V} \in \widetilde{\mathcal{V}}_q(I_n), \end{aligned} \tag{5.14}$$

where $\widetilde{R}_1$ and $\widetilde{\mathrm{R}}_2$ are given by (5.12) and (5.13), respectively. In view of (5.14) and the analysis of Subsection 4.2, it is clear that we can obtain optimal order a priori error estimates as long as we are able to prove that the residuals $\int_0^T \|\widetilde{R}_1(t)\|^2_{H^{-1}(\widetilde{\Omega}_t)}\, dt$ and $\int_0^T \|\widetilde{\mathrm{R}}_2(t)\|^2_{L^2(\widetilde{\Omega}_t)}\, dt$ are of optimal order.

The first step in this direction is to examine the quantities z and K. Using (2.3), we see that

$$1 - z = 1 - \mathrm{e}^{\int_0^t \left(\nabla_{\mathbf{x}} \cdot \mathbf{w}\left(\mathcal{A}_s(\mathbf{y}), s\right) - \nabla_{\widetilde{\mathbf{x}}} \cdot \widetilde{\mathbf{W}}\left(\widetilde{\mathcal{A}}_s(\mathbf{y}), s\right)\right) ds} \qquad \text{a.e. } \mathbf{y} \in \Omega_0, t \in I_n. \tag{5.15}$$

Since

$$\nabla_{\mathbf{x}} \cdot \mathbf{w} - \nabla_{\widetilde{\mathbf{x}}} \cdot \widetilde{\mathbf{W}} = (\mathbf{J}_{\mathcal{A}_t}^{-T} \nabla_{\mathbf{y}}) \cdot (\widehat{\mathbf{w}} - \widehat{\mathbf{W}}) + \big(\mathbf{J}_{\mathcal{A}_t}^{-T} (\mathbf{J}_{\widetilde{\mathcal{A}}_t}^{T} - \mathbf{J}_{\mathcal{A}_t}^{T}) \mathbf{J}_{\widetilde{\mathcal{A}}_t}^{T} \nabla_{\mathbf{y}}\big) \cdot \widehat{\mathbf{W}},$$

using in (5.15) Taylor's expansion and employing (5.4) together with (5.6), yields

$$\|1 - z\|_{L^\infty(\Omega_0 \times I_n)} \lesssim k^{j+1} T \Gamma_j \|\partial_t^{j+1} \widehat{\mathbf{w}}\|_{\mathbf{L}^\infty((0,T);\mathbf{W}^1_\infty(\Omega_0))}. \tag{5.16}$$

Moreover, (5.4) also gives

$$\|I_d - K\|_{L^\infty(\Omega_0 \times I_n)} \lesssim k^{j+1} \lambda^{-1} T \|\partial_t^{j+1} \widehat{\mathbf{w}}\|_{\mathbf{L}^\infty((0,T);\mathbf{W}^1_\infty(\Omega_0))}. \tag{5.17}$$

The second step, to estimate $\int_0^T \|\widetilde{R}_1(t)\|^2_{H^{-1}(\widetilde{\Omega}_t)}\, dt$, consists of writing $I_d - zK = (1-z)I_d + z(I_d - K)$ and using Poincare's inequality in conjunction with (5.16) and (5.17) to obtain

$$\int_0^T \|\widetilde{R}_1(t)\|^2_{L^2(\Omega_t)}\, dt \lesssim k^{2j+2} C(u, f, \mathbf{w}, \Gamma_j, T) \|\partial_t^{j+1} \widehat{\mathbf{w}}\|^2_{\mathbf{L}^\infty((0,T);\mathbf{W}^1_\infty(\Omega_0))}, \tag{5.18}$$

for all $j = 0, 1, \dots, q$, where $C(u, f, \mathbf{w}, \Gamma, T)$ is a constant that depends on $u, f, \mathbf{w}, \Gamma_j, T$. Similarly, writing $I_d - zK^{\mathrm{T}} K = (I_d - zK^{\mathrm{T}}) + zK^{\mathrm{T}}(I_d - K)$, we deduce that

$$\int_0^T \|\widetilde{\mathrm{R}}_2(t)\|^2_{L^2(\widetilde{\Omega}_t)}\, dt \le k^{2j+2} (\Gamma_j^2 + \Lambda_j) C(u) \|\partial_t^{j+1} \widehat{\mathbf{w}}\|^2_{\mathbf{L}^\infty((0,T);\mathbf{W}^1_\infty(\Omega_0))}, \tag{5.19}$$

for all $j = 0, 1, \dots, q$, where the constant $C(u)$ depends only on u.

We finish this subsection by stating the third a priori error estimate which corresponds to a general ALE map.

Theorem 5.1 (a priori error estimate with Reynolds' quadrature: general ALE map)**.** *Let the perturbed ALE map $\widetilde{\mathcal{A}}_t$ be a piecewise polynomial of degree $q+1$ with $\widetilde{\mathcal{A}}_{t_{n+1}\to t} \in \mathbf{L}^\infty\big(I_n; \mathbf{W}^2_\infty(\Omega_{t_{n+1}})\big)$ and so that* (5.3) *is satisfied. Let also $\widetilde{\Phi}$ be the approximation of u defined in* (5.9) *with a Reynolds' quadrature. If the time steps are chosen so that* (5.5) *is satisfied, then the following a priori error estimate holds true, for $j = 0, 1, \dots, q$,*

$$\begin{aligned} \max_{0\le n\le N} \|(\widetilde{u} - \widetilde{\Phi})(t_n)\|^2_{L^2(\widetilde{\Omega}_{t_n})} + \frac{\mu}{2} \sum_{n=0}^{N-1} Q_n(\|\nabla_{\widetilde{\mathbf{x}}}(\widetilde{u} - \widetilde{\Phi})(t)\|^2_{\mathbf{L}^2(\widetilde{\Omega}_t)}) \le \widetilde{\mathcal{E}}(u, f, \widetilde{\mathcal{A}}_t, \mathbf{b}; j) \\ + k^{2j+2} \Big(\frac{1}{\mu} C(u, f, \mathbf{w}, \Gamma_j, T) + \mu(\Gamma_j^2 + \Lambda_j) C(u)\Big) \|\partial_t^{j+1} \widehat{\mathbf{w}}\|^2_{\mathbf{L}^\infty((0,T);\mathbf{W}^1_\infty(\Omega_0))}, \end{aligned} \tag{5.20}$$

where $\widetilde{\mathcal{E}}(u, f, \widetilde{\mathcal{A}}_t, \mathbf{b}; j)$ denotes the right-hand-side of (4.2) *in Theorem 4.2, now with respect to the discrete ALE map $\widetilde{\mathcal{A}}_t$ and proportional constants, $C(u, f, \mathbf{w}, \Gamma_j, T)$ and $C(u)$ are constants proportional to those in* (5.18) *and* (5.19)*, respectively and Γ_j, Λ_j are defined in* (5.7) *and* (5.8)*. Furthermore, a similar estimate to* (5.20) *is valid for the continuous $L^2(H^1)-$norm $\int_0^T \|\nabla_{\widetilde{\mathbf{x}}}(\widetilde{u} - \widetilde{\Phi})(t)\|^2_{L^2(\widetilde{\Omega}_t)}\, dt$, as well.* □

Remark 5.1 (perturbation argument for finite element discretization)**.** It is clear that the perturbation argument of this section can also be applied when we approximate the ALE map in space using finite elements. The ALE projection will then be defined with respect to the continuous in space ALE map and *not* to the finite element map which is just W^1_∞ in space. Imposing the assumption (3.6) on the continuous in space ALE map is not unrealistic. In that respect, our analysis and theory is not restrictive.

In the next subsection we explain briefly how we can approximate the domain velocity only at the boundary and still the residuals be of optimal order of accuracy.

5.2. **An extension argument.** In the ALE framework, what we usually know in practice is the velocity of the boundary $\partial\Omega_t$ at every time $t \in [0,T]$ as discussed in the Introduction. In other words, only the value of the ALE map is imposed at the boundary and we are free to choose any motion inside. This is consistent with the fact that the solution u of (1.1) only depends on the shape of Ω_t, $t \in [0,T]$ and not on the artificial ALE motion inside which is usually designed to facilitate the numerical computations. We briefly discuss in this section how to extend estimate (5.20) in this context. Since this is closely related to the numerical implementation of the ALE method and approach, we discuss it in detail in a forthcoming computational paper [5].

We assume that the original ALE map $\mathcal{A}_t$ (and thus its velocity $\widehat{\mathbf{w}}$) is only given at the boundary, i.e., for $t \in [0,T]$ we assume the existence of a function $\mathcal{A}_t$,

$$\mathcal{A}_t : \partial\Omega_0 \to \partial\Omega_t.$$

We choose to keep the same notation for the functions defined at the boundary as well, to avoid making the notation heavier.

Next, we interpolate the domain velocity $\widehat{\mathbf{w}}$ by a piecewise polynomial $\widehat{\mathbf{W}}$ of degree q such that, for $n = 0, 1, \ldots, N-1$ and $j = 0, 1, \ldots, q$,

$$\|\widehat{\mathbf{w}} - \widehat{\mathbf{W}}\|_{\mathbf{L}^\infty(I_n;\mathbf{W}^1_\infty(\partial\Omega_0))} \lesssim k_n^{j+1} \|\partial_t^{j+1}\widehat{\mathbf{w}}\|_{\mathbf{L}^\infty(I_n;\mathbf{W}^1_\infty(\partial\Omega_0))}; \tag{5.21}$$

compare with (5.3). We compute the perturbed ALE map $\widetilde{\mathcal{A}}_t$ on the boundary using (5.1) and extend it inside $\widetilde{\Omega}_t$ with the most convenient technique to our purposes (see for instance [13, 24, 15, 16, 23]). Then the continuous ALE map $\mathcal{A}_t$ inside Ω_t is chosen to be the map that it is obtained by following the *exact same procedure* as for the extension of the discrete ALE map inside $\widetilde{\Omega}_t$, and to be such that for $n = 0, 1, \ldots, N-1$,

$$\|\widehat{\mathbf{w}} - \widehat{\mathbf{W}}\|_{\mathbf{L}^\infty(I_n;\mathbf{W}^1_\infty(\Omega_0))} \lesssim \|\widehat{\mathbf{w}} - \widehat{\mathbf{W}}\|_{\mathbf{L}^\infty(I_n;\mathbf{W}^1_\infty(\partial\Omega_0))}. \tag{5.22}$$

Combining (5.21) and (5.22) we end up with

$$\|\widehat{\mathbf{w}} - \widehat{\mathbf{W}}\|_{\mathbf{L}^\infty(I_n;\mathbf{W}^1_\infty(\Omega_0))} \lesssim k_n^{j+1} \|\partial_t^{j+1}\widehat{\mathbf{w}}\|_{\mathbf{L}^\infty(I_n;\mathbf{W}^1_\infty(\partial\Omega_0))}, \tag{5.23}$$

for $n = 0, 1, \ldots, N-1$ and $j = 0, 1, \ldots, q$. In view of (5.23) it is clear that the a priori error estimate (5.20) still holds true. It is to be emphasized though once more that the estimate is now obtained by approximating the domain velocity $\widehat{\mathbf{w}}$ *only at the boundary* and not inside the domain.

5.3. **Computational Asymptotic Accuracy.** We conclude this section by presenting numerical experiments illustrating the asymptotic precision of the method. We consider the square $\Omega_0 := [-1,1] \times [-1,1]$ deformed according to the ALE map $\mathcal{A}_t(\mathbf{y}) := \mathbf{y}(1 + \frac{1}{2}\, T_{11}(t))$ for $t \in (0, 0.99)$, where $T_n(t) := \cos(n \arccos(t))$ is the nth Chebychev polynomial of first kind. We set $\mu = 1.0$ and the exact solution is manufactured to be $u(\mathbf{x}, t) := \exp(x_1\ t)\sin(x_2\ t)$ with $\mathbf{x} := (x_1, x_2) \in \Omega_t$. To obtain a continuous piecewise polynomial approximation $\widetilde{\mathcal{A}}_t \in \widetilde{\mathcal{V}}_{q+1}$, we take $\widetilde{\mathbf{W}} \in \widetilde{\mathcal{V}}_q$ to be the L^2-projection in time of the ALE velocity $\mathbf{w}$ onto $\mathcal{V}_q$ and deduce $\widetilde{\mathcal{A}}_t \in \widetilde{\mathcal{V}}_{q+1}$ according to (5.1). Finally, a Reynolds' quadrature based on $q + 1 + [d(q+1))/2]$ Radau points is used for the integration over each interval I_n; cf. (2.12). We recall that the dG approximation $\widetilde{\Phi} \in \widetilde{\mathcal{V}}_q$ is defined using $\widetilde{\mathcal{A}}_t$ through (5.9) and $\widetilde{u} := u \circ \mathcal{A}_t \circ \widetilde{\mathcal{A}}_t^{-1}$. The error $\max_{0 \le n \le N} \|(\widetilde{u} - \widetilde{\Phi})(t_n)\|^2_{L^2(\widetilde{\Omega}_{t_n})}$ against the number of uniform time steps used for the computation is depicted in Fig. 1 for different schemes $(q = 0, 1, 2, 3)$. The optimal rate of convergence $q + 1$ is observed in each case.

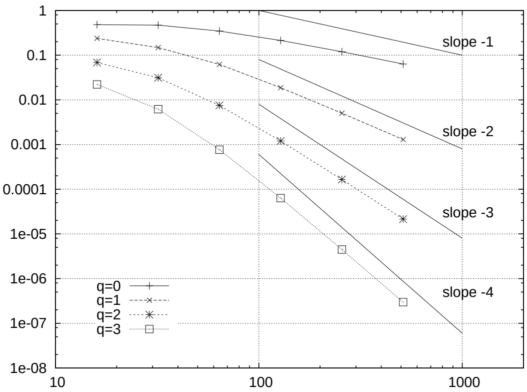

FIGURE 1. The error $\max_{0\le n\le N} \|(\widetilde{u}-\widetilde{\Phi})(t_n)\|^2_{L^2(\widetilde{\Omega}_{t_n})}$ against the number N of uniform timesteps is depicted for $q = 0, 1, 2, 3$. Over each interval I_n, a Reynolds' quadrature based on $q+1+[d(q+1))/2]$ Radau points is used, cf. (2.12). The discretization in space is chosen sufficiently fine, not to influence the time discretization error. As predicted by estimate (5.20), all schemes exhibit the optimal $\mathcal{O}(N^{-(q+1)})$ order of convergence.

6. Runge-Kutta-Radau Methods: Conditional Error Estimates

In this section we develop an error analysis for RKR methods, i.e., for (2.10) using $q+1$ Radau quadrature points. We recall that this is the minimal quadrature for dG giving unconditional optimal order a priori error estimates on time independent domains. In contrast to (2.10) with Reynolds' quadrature, the ALE map *does not* need to be piecewise polynomial in what follows.

6.1. A Priori Error Estimates.

The error $\Theta = Pu - U$ satisfies an equation similar to (4.1) with the extra quadrature errors $-E_n(\langle D_t Pu, V\rangle_{\Omega_t})$ and $-E_n\big(\langle \nabla_{\mathbf{x}}\cdot \mathbf{w}Pu, V\rangle_{\Omega_t}\big)$ on the right-hand side which do not vanish for Radau quadrature. We thus have the following conditional error estimate.

Theorem 6.1 (error estimate for the RKR methods). *Let $U \in \mathcal{V}_q$ be the solution of* (2.10) *with $q+1$ Radau quadrature points, and let $\mathcal{A}_t$ satisfy* (3.6). *If the time steps satisfy* (2.14)*, then the following error estimate for $u-U$ is valid, for all $j = 0, \cdots, q$,*

$$\begin{aligned}\max_{0\le n\le N} \|(u-U)(t_n)\|^2_{L^2(\Omega_{t_n})} &+ \frac{\mu}{2}\sum_{n=0}^{N-1} Q_n\big(\|\nabla_{\mathbf{x}}(u-U)(t)\|^2_{\mathbf{L}^2(\Omega_t)}\big) \le \mathcal{E}(u,f,\mathcal{A}_t,\mathbf{b};j)\\ &+\frac{1}{\mu}\sum_{n=0}^{N-1} C_n G_{n,j+1} k_n^{2j+2}\sum_{i=0}^{j+1}\int_{I_n}\Big(1+\|D_t^{j+1-i}\nabla_{\mathbf{x}}\mathbf{w}(t)\|^2_{\mathbf{L}^\infty(\Omega_t)}\Big)\|D_t^i u(t)\|^2_{L^2(\Omega_t)}\,dt\end{aligned} \tag{6.1}$$

where $\mathcal{E}(u,f,\mathcal{A}_t,\mathbf{b};j)$ denotes the right-hand side of (4.2) *(with proportional constants), Pu is the ALE projection defined in* (3.2)-(3.3) *and $C_n, G_{n,j+1}$ are as in* (3.9) *and* (2.4)*, respectively. Moreover, a similar estimate to* (6.1) *is valid for the continuous $L^2(H^1)-$norm $\int_0^T \|\nabla_{\mathbf{x}}(u-U)(t)\|^2_{\mathbf{L}^2(\Omega_t)}\,dt$.*

Proof. This is a consequence of Lemma 2.3 for Radau quadrature. In addition to Theorem 4.2, we now need to estimate $E_n(\langle D_t Pu, V\rangle_{\Omega_t})$ and $E_n\big(\langle\nabla_{\mathbf{x}}\cdot\mathbf{w}Pu, V\rangle_{\Omega_t}\big)$, which we do again employing Lemma 2.2. This leads to derivatives $D_t^i Pu$ for $i = 0, \cdots, j+1$, which can in turn be estimated by the stability bound (3.19). This concludes the proof. □

We point out that the analysis of this section has conceptual similarities with that of Ciarlet and Raviart on the effect of quadrature for the finite element method for elliptic equations [9, Chapter 4, p. 174–201]. The V_h-ellipticity of [9] is replaced by the stability result of Lemma 2.3. However, stability does not automatically lead to an optimal order a priori error bound since we need to prove the stability estimates of Proposition 3.3 for the ALE projection in order to close the argument.

6.2. **Computational Asymptotic Accuracy.** We now provide numerical experiments reflecting the asymptotic precision for the RKR methods. The test case considered is the same as described in Section 5.3 except that only $q+1$ Radau points are used for the time integration and that the ALE map is not projected into $\mathcal{V}_{q+1}$ but used directly in the computations. The error $\max_{0\le n\le N}\|(u-U)(t_n)\|^2_{L^2(\Omega_{t_n})}$ versus the number of uniform timesteps used for the computation is displayed in Fig. 1 for $q=0,1,2,3$. The optimal rate of convergence $q+1$ is observed in each case.

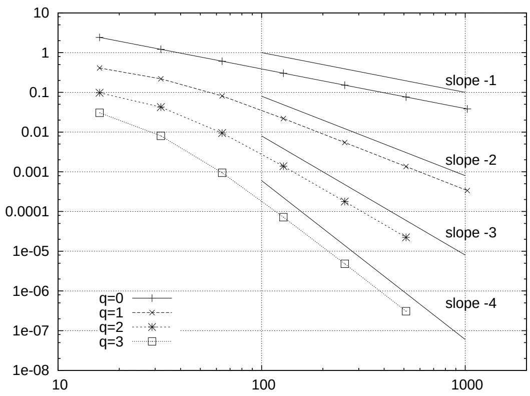

FIGURE 2. The error $\max_{0\le n\le N}\|(u-U)(t_n)\|^2_{L^2(\Omega_{t_n})}$ versus the number of uniform timesteps N is depicted for $q=0,1,2,3$. Over each interval I_n, a quadrature based on $q+1$ Radau points is used. The discretization in space is chosen not to influence the time discretization error. As predicted by estimate (6.1), all schemes exhibit the optimal $\mathcal{O}(N^{-(q+1)})$ order of convergence.

Acknowledgments

The authors would like to express their gratitude to W. Bangerth and G. Kanschat for participating in several discussions regarding the implementation of dG in time with deal.II [3].

References

[1] G. Akrivis and Ch. Makridakis. Galerkin time-stepping methods for nonlinear parabolic equations. *M2AN Math. Model. Numer. Anal.*, 38(2):261–289, 2004.

[2] S. Badia and R. Codina. Analysis of a stabilized finite element approximation of the transient convection-diffusion equation using an ALE framework. *SIAM J. Numer. Anal.*, 44(5):2159–2197 (electronic), 2006.

[3] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II—a general-purpose object-oriented finite element library. *ACM Trans. Math. Software*, 33(4):Art. 24, 27, 2007.

[4] D. Boffi and L. Gastaldi. Stability and geometric conservation laws for ALE formulations. *Comput. Methods Appl. Mech. Engrg.*, 193(42-44):4717–4739, 2004.

[5] A. Bonito, I. Kyza, and R.H. Nochetto. Time-discrete higher order ALE formulations: Computational investigation. In preparation.

[6] A. Bonito, I. Kyza, and R.H. Nochetto. Time-discrete higher order ALE formulations: Stability. Submitted.

[7] S.C. Brenner and L.R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[8] K. Chrysafinos and N.J. Walkington. Lagrangian and moving mesh methods for the convection diffusion equation. *M2AN Math. Model. Numer. Anal.*, 42(1):25–55, 2008.

[9] Ph.G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.

[10] J. Donéa, S. Giuliani, and J.P. Halleux. An arbitrary lagrangian-eulerian finite element method for transient dynamic fluid-structure interactions. *Comput. Methods Appl. Mech. Engrg.*, 33(1-3):689 – 723, 1982.
[11] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL,, 1992.
[12] Ch. Farhat, Ph. Geuzaine, and C. Grandmont. The discrete geometric conservation law and the nonlinear stability of ALE schemes for the solution of flow problems on moving grids. *J. Comput. Phys.*, 174(2):669–694, 2001.
[13] Ch. Farhat, M. Lesoinne, and N. Maman. Mixed explicit/implicit time integration of coupled aeroelastic problems: three-field formulation, geometric conservation and distributed solution. *Internat. J. Numer. Methods Fluids*, 21(10):807–835, 1995. Finite element methods in large-scale computational fluid dynamics (Tokyo, 1994).
[14] L. Formaggia and F. Nobile. A stability analysis for the arbitrary Lagrangian Eulerian formulation with finite elements. *East-West J. Numer. Math.*, 7(2):105–131, 1999.
[15] L. Formaggia and F. Nobile. Stability analysis of second-order time accurate schemes for ALE-FEM. *Comput. Methods Appl. Mech. Engrg.*, 193(39-41):4097–4116, 2004.
[16] L. Gastaldi. A priori error estimates for the arbitrary Lagrangian Eulerian formulation with finite elements. *East-West J. Numer. Math.*, 9(2):123–156, 2001.
[17] Ph. Geuzaine, C. Grandmont, and Ch. Farhat. Design and analysis of ale schemes with provable second-order time-accuracy for inviscid and viscous flow simulations. *J. Comp. Phys.*, 191(1):206–227, 2003.
[18] H. Guillard and Ch. Farhat. On the significance of the geometric conservation law for flow computations on moving meshes. *Comput. Methods Appl. Mech. Engrg.*, 190(11-12):1467–1482, 2000.
[19] C. W. Hirt, A. A. Amsden, and J. L. Cook. An arbitrary Lagrangian-Eulerian computing method for all flow speeds [J. Comput. Phys. **14** (1974), no. 3, 227–253]. *J. Comput. Phys.*, 135(2):198–216, 1997. With an introduction by L. G. Margolin, Commemoration of the 30th anniversary {of J. Comput. Phys.}.
[20] T.J.R. Hughes, W.K. Liu, and T.K. Zimmermann. Lagrangian-Eulerian finite element formulation for incompressible viscous flows. *Comput. Methods Appl. Mech. Engrg.*, 29(3):329–349, 1981.
[21] P. Jamet. Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain. *SIAM J. Numer. Anal.*, 15(5):912–928, 1978.
[22] F. Nobile. *Numerical approximation on fluid-structure interaction problems with application to haemodynamics.* PhD thesis, École Polytechnique Fédérale de Lausanne, 2001.
[23] L.A. Ortegaa and G. Scovazzi. A geometrically-conservative, synchronized, flux-corrected remap for arbitrary lagrangian-eulerian computations with nodal finite elements. Submitted.
[24] A. Quarteroni, M. Tuveri, and A. Veneziani. Computational vascular fluid dynamics: problems, models and methods. *Comput. Visual. Sci.*, 2(4):163–197, 2000.
[25] E.M.. Stein. *Singular integrals and differentiability properties of functions*, volume 30 of *Princeton Mathematical Series.* Princeton University Press, Princeton, N.J., 1970.
[26] V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 25 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, second edition, 2006.

(Andrea Bonito) Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA
E-mail address: bonito@math.tamu.edu

(Irene Kyza) Department of Mathematics, University of Maryland, College Park, MD 20742-4015, USA
E-mail address: kyza@math.umd.edu

(Ricardo H. Nochetto) Department of Mathematics and Institute of Physical Science and Technology, University of Maryland, College Park, MD 20742-4015, USA
E-mail address: rhn@math.umd.edu