



## **Space-time CutFEM on overlapping meshes I: simple continuous mesh motion**

Downloaded from: <https://research.chalmers.se>, 2024-06-30 16:29 UTC

Citation for the original published paper (version of record):

Larson, M., Logg, A., Lundholm, C. (2024). Space-time CutFEM on overlapping meshes I: simple continuous mesh motion. *Numerische Mathematik*, 156(3): 1015-1054.

<http://dx.doi.org/10.1007/s00211-024-01417-8>

N.B. When citing this work, cite the original published paper.



# Space-time CutFEM on overlapping meshes I: simple continuous mesh motion

Mats G. Larson<sup>1</sup> · Anders Logg<sup>2</sup> · Carl Lundholm<sup>1</sup>

Received: 24 February 2023 / Revised: 11 March 2024 / Accepted: 23 April 2024 /

Published online: 25 May 2024

© The Author(s) 2024

## Abstract

We present a cut finite element method for the heat equation on two overlapping meshes: a stationary background mesh and an overlapping mesh that moves around inside/“on top” of it. Here the overlapping mesh is prescribed by a simple continuous motion, meaning that its location as a function of time is *continuous* and *piecewise linear*. For the discrete function space, we use continuous Galerkin in space and discontinuous Galerkin in time, with the addition of a discontinuity on the boundary between the two meshes. The finite element formulation is based on Nitsche’s method and also includes an integral term over the space-time boundary between the two meshes that mimics the standard discontinuous Galerkin time-jump term. The simple continuous mesh motion results in a space-time discretization for which standard analysis methodologies either fail or are unsuitable. We therefore employ what seems to be a relatively uncommon energy analysis framework for finite element methods for parabolic problems that is general and robust enough to be applicable to the current setting. The energy analysis consists of a stability estimate that is slightly stronger than the standard basic one and an a priori error estimate that is of optimal order with respect to both time step and mesh size. We also present numerical results for a problem in one spatial dimension that verify the analytic error convergence orders.

**Mathematics Subject Classification** 65M12 · 65M15 · 65M60 · 65M85

## 1 Introduction

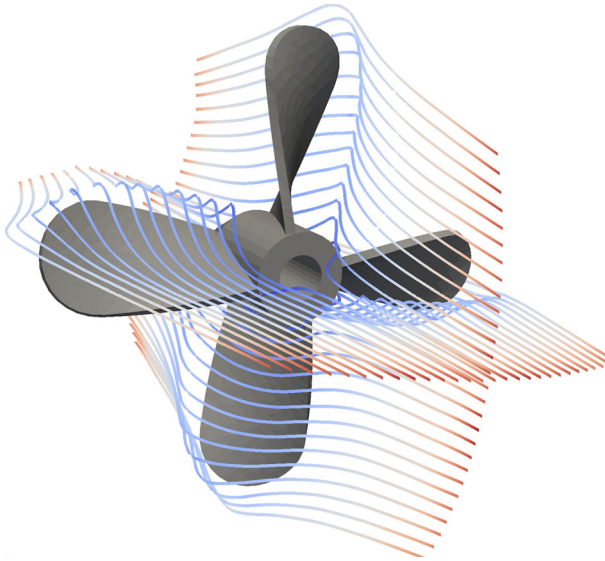
**Issue—Cost of mesh generation:** Generating computational meshes for numerically solving differential equations can be a computationally costly procedure. In practical

---

✉ Carl Lundholm  
carl.lundholm@umu.se

<sup>1</sup> Department of Mathematics and Mathematical Statistics, Umeå University, 901 87 Umeå, Sweden

<sup>2</sup> Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 412 96 Gothenburg, Sweden

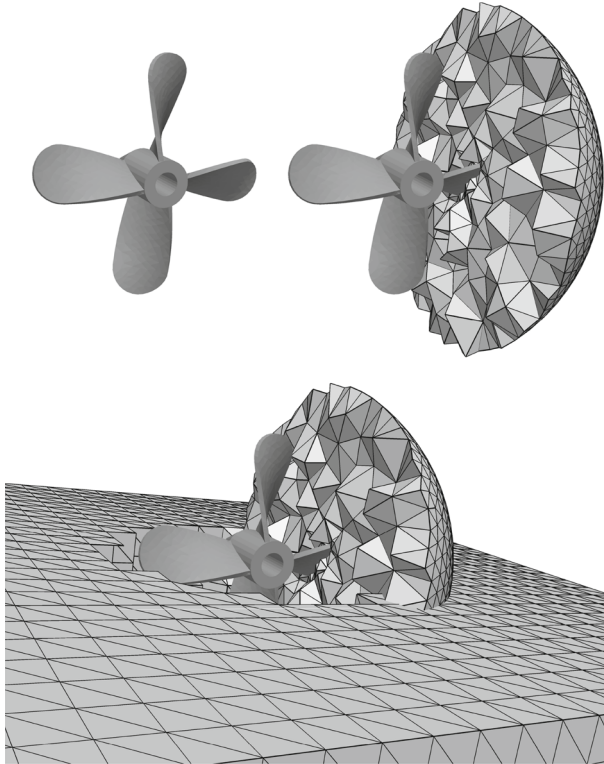


**Fig. 1** Computed streamlines around a propeller. Image by Anders Logg is licensed under CC BY 4.0

applications the mesh generation can often represent a substantial amount of the total computation time. This is especially true for problems where the solution domain changes during the solution process, e.g., evolving geometry and shape optimization. With standard methods the mesh then has to be constantly checked for degeneracy and updated if needed, meaning a persisting meshing cost for the entire solution process.

**Remedy—CutFEM:** Cut finite element methods (CutFEMs) provide a way of decoupling the computational mesh from the problem geometry. This means that the same discretization can be used for a changing solution domain. CutFEMs can thus make remeshing redundant for problems with changing geometry but also for other applications involving meshing such as adaptive mesh refinement. The cost of CutFEMs is treating the mesh cells that are arbitrarily cut by the independent problem geometry.

**CutFEM on overlapping meshes:** A common type of problem with changing geometry is one where there is a moving object in the solution domain, e.g., see Fig. 1. A straightforward CutFEM-approach to this problem would be to consider CutFEM for the interface problem, i.e., to use a background mesh of the empty solution domain together with an interface that represents the object. However, a more advantageous and sophisticated approach is to consider CutFEM on *overlapping meshes*, meaning two or more meshes ordered in a mesh hierarchy. This is also called composite grids/meshes and multimesh in the literature. The idea is to use a background mesh of the empty solution domain, just as for the interface problem, but instead to encapsulate the object in a second mesh. The mesh containing the object is then placed “on top” of the background mesh, creating a mesh hierarchy. The motion of the object will thus also cause its encapsulating mesh to move. There are some advantages of using a second overlapping mesh instead of an interface. Firstly, an overlapping mesh can incorporate boundary layers close to the object. Something an interface cannot. Sec-



**Fig. 2** Overlapping meshes for a problem with a rotating propeller. Image by Anders Logg is licensed under CC BY 4.0

only, if the object has a complicated geometry, representing it with an interface can lead to tricky cut situations and thus a higher computational cost. By instead using an object-encapsulating mesh with a simply-shaped exterior boundary, the cut situations can be made less tricky, see Fig. 2. A way to further sophisticate this is to allow the moving object to deform the interior of the overlapping mesh while initially keeping its exterior boundary fixed. Only when the deformations have become too large is the overlapping mesh “snapped” into place to avoid degeneracy. Such a snapping feature provides a choice between computing cut situations or computing deformations, thus allowing the cheapest option for the situation at hand to be chosen. A drawback of using a second overlapping mesh instead of an interface is that overlapping meshes require collision computations between the cells of the meshes, something that can be computationally expensive.

CutFEM on overlapping meshes can also be used as an alternative to adaptive mesh refinement by keeping a smaller finer mesh in regions requiring higher accuracy. Yet another application example is to use a composition of simpler structured meshes to represent a complicated domain.

**Literary background:** Over the past two decades, a theoretical foundation for the formulation of stabilized CutFEM has been developed by extending the ideas of Nitsche, presented in [1], to a general weak formulation of the interface conditions,

thereby removing the need for domain-fitted meshes. The foundations of CutFEM were presented in [2] and then extended to overlapping meshes in [3]. The CutFEM methodology has since been developed and applied to a number of important multi-physics problems [4–7]. For overlapping meshes in particular, see for example [8–11]. As already touched upon, CutFEM is especially relevant for applications with changing geometry such as time-dependent problems and the methodology has been employed partially or fully in several works on unfitted FEM for such problems. These include moving interfaces [12, 13], moving domains [14–18], and evolving surfaces for which the methodology can take the form of TraceFEM [19–22]. However, when it comes to CutFEM on overlapping meshes, only methods for *stationary* problems have been developed and analysed to a satisfactory degree, thus leaving analogous work for *time-dependent* problems to be desired.

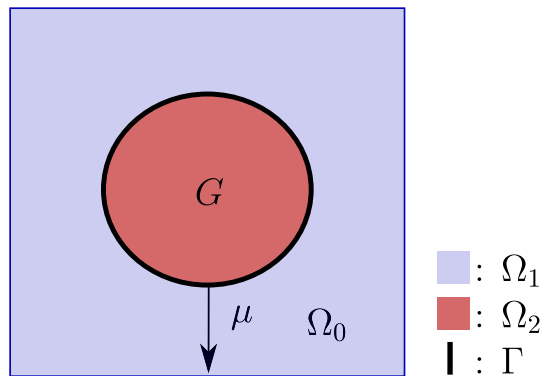
**This work:** The work presented here is together with [23–25] intended to be an initial part of developing and analysing CutFEMs on overlapping meshes for time-dependent problems. We consider a CutFEM for the heat equation on two overlapping meshes: one stationary background mesh and one moving overlapping mesh with no object. Depending on how the mesh motion is represented discretely, quite different space-time discretizations may arise, allowing for different types of analyses to be applied. Generally the mesh motion may either be continuous or discontinuous, which might also affect the suitability for different applications. For instance, the information in a prismatic space-time method flows along the space-time trajectories of the underlying spatial mesh. This means that the flow of information of the overlapping mesh is more well-connected in the continuous case, whereas in the discontinuous one, the information “leaks” out of the overlapping mesh. This could suggest that continuous mesh motion is more suitable when the overlapping mesh represents something *physical*, and that discontinuous mesh motion is more suitable when it is a *computational* feature that should not be “seen”, e.g., alternative to adaptive mesh refinement. More discussion of a comparison of the two motions can be found in [24]. We have considered the simplest case of both of these two types of mesh motion, which we refer to as *simple continuous* and *simple discontinuous* mesh motion. Simple continuous mesh motion means that the location of the overlapping mesh as a function of time is *continuous* and *piecewise linear*, and simple discontinuous mesh motion means that it is *discontinuous* and *piecewise constant*. The first study on this topic, presented in the MSc thesis [23], considered simple continuous mesh motion together with an  $L^2$ -analysis (error in spatial  $L^2$ -norm at the final time). Partially due to  $L^2$ -analysis’s demanding stability requirements, error bounds were unfortunately not obtained and the analysis was left incomplete. However, the numerical results indicate that the superconvergence with respect to the time step is lost with continuous mesh motion, but that the other error convergence orders are preserved. After the first study, simplifications were made in two directions: less demanding analysis and less complicated mesh motion. This resulted in two new studies with complete analyses: energy analysis for continuous mesh motion and  $L^2$ -analysis for discontinuous. The former is presented in this work and the latter in [25]. They are also part of the Phd thesis [24] as long and technical manuscripts, referred to as Paper I and Paper II, respectively. There, detailed discussions and comparisons of all *three* studies are also presented.

**Analysis:** The simple continuous mesh motion results in a space-time discretization with skewed space-time nodal trajectories and cut prismatic space-time cells. This discretization lacks a slabwise product structure between space and time. Standard analysis methodology relying on such a structure therefore either fail or require too restrictive assumptions here. The reason for this is that standard analysis methodology typically use spatial operators that map to the momentaneous finite element space, such as the discrete Laplacian (used in the aforementioned  $L^2$ -analysis) and the solution operator used to define the  $H^{-1}$ -norm on  $L^2$  (used in standard energy analysis). If the spatial discretization changes within slabs these operators get an intrinsic time dependence that standard methodologies fail to incorporate. We therefore employ what seems to be a relatively uncommon analysis framework for finite element methods for parabolic problems that avoids the use of operators of the aforementioned type which thus makes it general and robust enough to be applicable to the current discretization. The framework has its roots in analysis of the streamline diffusion method, first presented in [26] and first analyzed in [27], where certain analytic components later were used to obtain improved and optimal order error bounds for the discontinuous Galerkin method in [28]. The full abstract formulation of the analysis framework can be found in [29]. For time-dependent problems it has been applied, e.g., in [30] for general polytopic spatial meshes and in [15] for an unfitted FEM for moving domains. The analysis framework is of an energy type, where space-time energy norms are used to derive and obtain a stability estimate that is slightly stronger than the standard basic one and an a priori error estimate that is of optimal order with respect to both time step and mesh size. The main steps of the energy analysis are:

0. Handling of the time derivative: This is the initial step that characterizes and sets the course for the whole analysis. Instead of the  $H^{-1}$ -norm, the  $L^2$ -norm scaled with the time step is used to include the time-derivative term in a space-time energy norm. This may intuitively be viewed as treating the time derivative as temporal advection. An alternative intuition for the handling is as a discrete version of the  $H^{-1}$ -norm.
1. Analytic preliminaries: A “perturbed coercivity” is proved which is used to show an inf-sup condition. These results become slightly different compared with corresponding standard ones due to the handling of the time derivative.
2. Stability analysis: The “perturbed coercivity” is used to derive a stability estimate that is somewhat stronger than the standard basic one obtained by testing with the discrete solution.
3. Error analysis: Just as in a standard energy analysis, a Cea’s lemma type argument is followed by using the inf-sup condition, Galerkin orthogonality, and continuity. A difference here is that the continuity comes with a twist, namely temporal integration by parts, which is needed because of the slightly different inf-sup condition. Finally, together with an interpolation estimate, an optimal order a priori error estimate may be proved.

**Paper overview:** The rest of the paper is organized as follows. We start by formulating the model problem in Section 2, followed by a corresponding CutFEM in Section 3. Then we present and prove analytic tools in Section 4 and a discrete stability estimate in Section 5. In Section 6, the main analytic result which is an optimal order a priori

**Fig. 3** Partition of  $\Omega_0$  into  $\Omega_1$  (blue) and  $\Omega_2$  (red) for  $d = 2$  with  $G$  moving with velocity  $\mu$



error estimate is presented and proved. Numerical results for a problem in one spatial dimension that verify the analytic convergence orders are presented in Section 7. The last part of the paper is an appendix that contains technical estimates and interpolation results used in the analysis.

## 2 Problem

For  $d = 1, 2$ , or  $3$ , let  $\Omega_0 \subset \mathbb{R}^d$  be a bounded convex domain with polygonal boundary  $\partial\Omega_0$ . Let  $T > 0$  be a given final time. Let  $G \subset \Omega_0 \subset \mathbb{R}^d$  be another bounded domain with polygonal boundary  $\partial G$ . We let the *location* of  $G$  be time-dependent by prescribing for  $G$  a velocity  $\mu : [0, T] \rightarrow \mathbb{R}^d$ . We point out that this makes the size and shape of  $G$  remain the same for all times. That  $\mu$  does not depend on space slightly simplifies some analytic technicalities, especially the proofs of Lemma A.8 and Lemma A.10. Using  $\Omega_0$  and  $G$ , we define the following two domains:

$$\Omega_1 := \Omega_0 \setminus (G \cup \partial G) \quad (2.1)$$

$$\Omega_2 := \Omega_0 \cap G \quad (2.2)$$

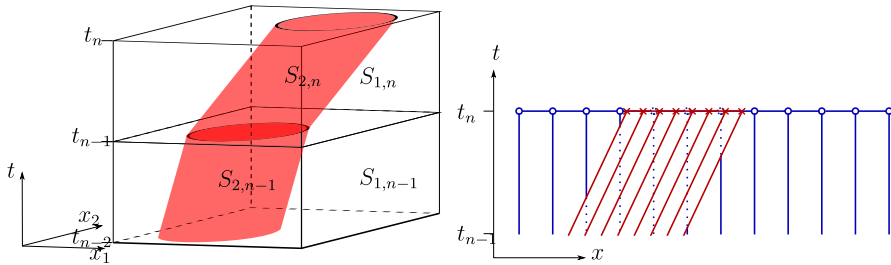
with boundaries  $\partial\Omega_1$  and  $\partial\Omega_2$ , respectively. Let their common boundary be

$$\Gamma := \partial\Omega_1 \cap \partial\Omega_2 \quad (2.3)$$

For  $t \in [0, T]$ , we have the partition

$$\Omega_0 = \Omega_1(t) \cup \Gamma(t) \cup \Omega_2(t) \quad (2.4)$$

See Fig. 3 for an illustration. We consider the heat equation in  $\Omega_0 \times (0, T]$  with source  $f \in L^2((0, T], \Omega_0)$ , homogeneous Dirichlet boundary conditions, and initial data



**Fig. 4** Left: Space-time slabs with simple continuous mesh motion. Right: Space-time discretization for  $S_{0,n}$  for  $d = 1$  when  $\mu > 0$ . At time  $t = t_n$ , the nodes of the blue background mesh  $\mathcal{T}_0$  are marked with circles and the nodes of the red moving mesh  $\mathcal{T}_G$  with crosses. The blue vertical lines are thus the nodal trajectories of  $\mathcal{T}_0$  and the red skewed vertical lines those of  $\mathcal{T}_G$

$$u_0 \in H^2(\Omega_0) \cap H_0^1(\Omega_0):$$

$$\begin{cases} \partial_t u - \Delta u = f & \text{in } \Omega_0 \times (0, T] \\ u = 0 & \text{on } \partial\Omega_0 \times (0, T] \\ u = u_0 & \text{in } \Omega_0 \times \{0\} \end{cases} \tag{2.5}$$

We stress that although we have the partition (2.4), the problem (2.5) is itself a one-domain problem for ease of analysis.

### 3 Method

#### 3.1 Preliminaries

Let  $\mathcal{T}_0$  and  $\mathcal{T}_G$  be quasi-uniform simplicial meshes of  $\Omega_0$  and  $G$ , respectively. We denote by  $h_K$  the diameter of a simplex  $K$ . We partition the time interval  $(0, T]$  quasi-uniformly into  $N$  subintervals  $I_n = (t_{n-1}, t_n]$  of length  $k_n = t_n - t_{n-1}$ , where  $0 = t_0 < t_1 < \dots < t_N = T$  and  $n = 1, \dots, N$ . We assume the following space-time quasi-uniformity: For  $h = \max_{K \in \mathcal{T}_0 \cup \mathcal{T}_G} \{h_K\}$ , and  $k = \max_{1 \leq n \leq N} \{k_n\}$ ,

$$h^2 \lesssim k_{\min} \quad k \lesssim h_{\min} \tag{3.1}$$

where  $k_{\min} = \min_{1 \leq n \leq N} \{k_n\}$ , and  $h_{\min} = \min_{K \in \mathcal{T}_0 \cup \mathcal{T}_G} \{h_K\}$ . We next define the following slabwise space-time domains:

$$S_{0,n} := \Omega_0 \times I_n \tag{3.2}$$

$$S_{i,n} := \{(x, t) \in S_{0,n} : x \in \Omega_i(t)\} \tag{3.3}$$

$$\bar{\Gamma}_n := \{(s, t) \in S_{0,n} : s \in \Gamma(t)\} \tag{3.4}$$

In general we will use bar, i.e.,  $\bar{\cdot}$ , to denote something related to space-time, such as domains and variables. In addition to the domains  $\Omega_1(t)$  and  $\Omega_2(t)$ , we also consider



the “covered” overlap domain  $\Omega_O(t)$ . To define it we will use the set of simplices  $\mathcal{T}_{0,\bar{\Gamma}_n} := \{K \in \mathcal{T}_0 : \exists t \in I_n \text{ such that } K \cap \Gamma(t) \neq \emptyset\}$ , i.e., all simplices in  $\mathcal{T}_0$  that are cut by  $\bar{\Gamma}_n$ . We define the overlap domain  $\Omega_O(t)$  for a time  $t \in I_n$  by

$$\Omega_O(t) := \bigcup_{K \in \mathcal{T}_{0,\bar{\Gamma}_n}} K \cap \Omega_2(t) \tag{3.5}$$

As a discrete counterpart to the motion of the domain  $G$ , we prescribe a simple continuous motion for the overlapping mesh  $\mathcal{T}_G$ . By this we mean that the location of the overlapping mesh  $\mathcal{T}_G$  is a function with respect to time that is *continuous* on  $[0, T]$  and *linear* on each  $I_n$ . This means that the discrete velocity we prescribe for  $\mathcal{T}_G$  is constant on each  $I_n$ . Henceforth, we let  $\mu$  denote this discrete velocity. Letting  $\mu_{\text{cont}}$  denote the velocity prescribed for  $G$ , we take the discrete velocity to be  $\mu|_{I_n} = k_n^{-1} \int_{I_n} \mu_{\text{cont}}(t) dt$ , for  $n = 1, \dots, N$ , i.e., the slabwise average. An illustration of the slabwise space-time domains  $S_{i,n}$  defined by (3.3) is shown in Fig. 4 (Left). Figure 4 (Right) shows a slabwise space-time discretization that has both straight and skewed space-time trajectories as a result of the simple continuous mesh motion. In a standard setting with only straight space-time trajectories, the time-derivative operator  $\partial_t$  is naturally also a derivative operator in the direction of the trajectories. This is convenient and we would like have an analogous operator for our setting. We start by defining the domain-dependent velocity  $\mu_i = \mu_i(t)$  by

$$\mu_i(t) := \begin{cases} \mathbf{0} & i = 1 \\ \mu(t) & i = 2 \end{cases} \tag{3.6}$$

We use this velocity to define the domain-dependent differential operator  $D_t = D_{t,i}$  by

$$D_{t,i}\{\cdot\} := \partial_t\{\cdot\} + \mu_i \cdot \nabla\{\cdot\} \tag{3.7}$$

The operator  $D_t$  is a *scaled* derivative operator in the direction of the space-time trajectories. To see this, consider the space-time vector  $\bar{\mu}_i = (\mu_i, 1)$  and the space-time gradient  $\bar{\nabla} = (\nabla, \partial_t)$ . The *unscaled* derivative operator in the direction of the space-time trajectories is

$$D_{s,i} = \frac{\bar{\mu}_i}{|\bar{\mu}_i|} \cdot \bar{\nabla} = \frac{1}{|\bar{\mu}_i|} (\mu_i \cdot \nabla + \partial_t) = \frac{1}{|\bar{\mu}_i|} D_{t,i} \tag{3.8}$$

We thus have  $D_{t,i} = |\bar{\mu}_i| D_{s,i}$ . Let  $\bar{\tau} = \bar{\tau}(t)$  denote a space-time trajectory that is uncut on the time interval  $(t_a, t_b)$ , and  $v$  be a function of sufficient regularity. The intrinsic scaling of  $D_t$  gives the convenient integral identity

$$\int_{\bar{\tau}(t_a)}^{\bar{\tau}(t_b)} D_s v ds = \int_{t_a}^{t_b} D_t v dt \tag{3.9}$$

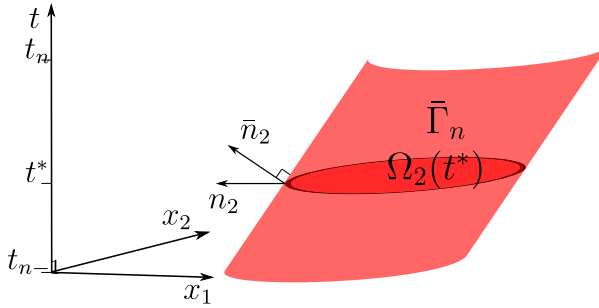


Fig. 5 Space-time normal vector  $\bar{n}_2$  to  $\bar{\Gamma}_n$  (red) in relation to the spatial normal vector  $n_2$  to  $\partial\Omega_2$

Next we introduce some normal vectors. Let the spatial vector  $n = n_i$  denote the outward pointing unit normal vector to  $\partial\Omega_i$ . Let the space-time vector  $\bar{n} = \bar{n}_i = (\bar{n}_i^x, \bar{n}_i^t)$  denote the outward pointing unit normal vector to  $\partial S_{i,n}$ , where  $\bar{n}_i^x$  and  $\bar{n}_i^t$  denote the spatial and temporal component(s), respectively. On a purely spatial subset, the space-time unit normal vector is purely temporal, i.e.,  $\bar{n}_i = (0, \pm 1)$ , and vice versa, i.e.,  $\bar{n}_i = (n_i, 0)$ . The remaining case is a mixed space-time subset and the only such set is  $\bar{\Gamma}_n$ . See Fig. 5 for an illustration. We define the space-time unit normal vector to  $\bar{\Gamma}_n$  by

$$\bar{n}_i|_{\bar{\Gamma}_n} = (\bar{n}_i^x, \bar{n}_i^t)|_{\bar{\Gamma}_n} := \frac{1}{\sqrt{(n_i \cdot \mu)^2 + 1}}(n_i, -(n_i \cdot \mu)) \tag{3.10}$$

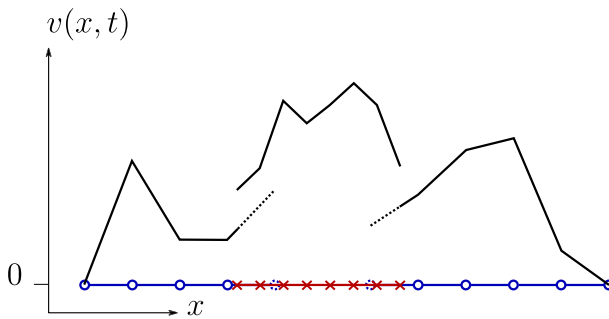
### 3.2 Finite element spaces

We define the discrete spatial finite element spaces  $V_{h,0}$  and  $V_{h,G}$  as the spaces of continuous piecewise polynomials of degree  $\leq p$  on  $\mathcal{T}_0$  and  $\mathcal{T}_G$ , respectively. We also let the functions in  $V_{h,0}$  be zero on  $\partial\Omega_0$ . For  $t \in [0, T]$ , we use these two spaces to define the broken finite element space  $V_h(t)$  by

$$V_h(t) := \{v : v|_{\Omega_1(t)} = v_0|_{\Omega_1(t)} \text{ for some } v_0 \in V_{h,0} \text{ and } v|_{\Omega_2(t)} = v_G|_{\Omega_2(t)} \text{ for some } v_G \in V_{h,G}\} \tag{3.11}$$

See Fig. 6 for an illustration of a function  $v \in V_h(t)$ . For  $n = 1, \dots, N$ , we define the discrete space-time finite element spaces  $V_{h,0}^n$  and  $V_{h,G}^n$  as the spaces of functions that for a  $t \in I_n$  lie in  $V_{h,0}$  and  $V_{h,G}$ , respectively, and in time are polynomials of degree  $\leq q$  along the trajectories of  $\mathcal{T}_0$  and  $\mathcal{T}_G$  for  $t \in I_n$ , respectively. For  $n = 1, \dots, N$ , we use these two spaces to define the broken finite element space  $V_h^n$  by:

$$V_h^n := \{v : v|_{S_{1,n}} = v_0^n|_{S_{1,n}} \text{ for some } v_0^n \in V_{h,0}^n \text{ and } v|_{S_{2,n}} = v_G^n|_{S_{2,n}} \text{ for some } v_G^n \in V_{h,G}^n\} \tag{3.12}$$



**Fig. 6** Example of  $v(\cdot, t) \in V_h(t)$  for  $d = 1$  and  $p = 1$ , where  $\mathcal{T}_0$  is blue,  $\mathcal{T}_G$  red, and the overlap parts are dotted

We define the global space-time finite element space  $V_h$  by:

$$V_h := \{v : v|_{S_{0,n}} \in V_h^n, n = 1, \dots, N\} \tag{3.13}$$

### 3.3 Finite element formulation

We may now formulate the space-time cut finite element formulation for the problem described in Sect. 2 as follows: Find  $u_h \in V_h$  such that

$$B_h(u_h, v) = \int_0^T (f, v)_{\Omega_0} dt + (u_0, v_0^+)_{\Omega_0} \quad \forall v \in V_h \tag{3.14}$$

The non-symmetric bilinear form  $B_h$  is defined by

$$B_h(w, v) := \sum_{i=1}^2 \sum_{n=1}^N \int_{I_n} (\partial_t w, v)_{\Omega_i(t)} dt + \sum_{n=1}^N \int_{I_n} A_{h,t}(w, v) dt + \sum_{n=1}^{N-1} ([w]_n, v_n^+)_{\Omega_0} + (w_0^+, v_0^+)_{\Omega_0} + \sum_{n=1}^N \int_{\bar{\Gamma}_n} -\bar{n}^t [w] v_\sigma d\bar{s} \tag{3.15}$$

where  $(\cdot, \cdot)_{\Omega}$  is the  $L^2(\Omega)$ -inner product,  $[v]_n$  is the jump in  $v$  at time  $t_n$ , i.e.,  $[v]_n = v_n^+ - v_n^-$ ,  $v_n^\pm = \lim_{\varepsilon \rightarrow 0^+} v(x, t_n \pm \varepsilon)$ . The last term in  $B_h$  mimics the standard dG-time-jump term, but over  $\bar{\Gamma}_n$ . Here,  $\bar{n}$  is the space-time normal vector to  $\bar{\Gamma}_n$  defined by (3.10),  $[v]$  is the jump in  $v$  over  $\bar{\Gamma}_n$ , i.e.,  $[v] = v_1 - v_2$ ,  $v_i = \lim_{\varepsilon \rightarrow 0^+} v(\bar{s} - \varepsilon \bar{n}_i)$ ,  $\bar{s} = (s, t)$ . If  $\bar{n} = \bar{n}_1$ , we take  $\sigma = \frac{1}{2}(3 + \text{sgn}(\bar{n}^t))$  and if  $\bar{n} = \bar{n}_2$ , we take  $\sigma = \frac{1}{2}(3 - \text{sgn}(\bar{n}^t))$ , where  $\text{sgn}$  is the sign function. These choices make it so that  $\sigma$  always picks the limit on the positive (in time) side of  $\bar{\Gamma}_n$ . The symmetric bilinear form  $A_{h,t}$  is defined by

$$\begin{aligned}
 A_{h,t}(w, v) := & \sum_{i=1}^2 (\nabla w, \nabla v)_{\Omega_i(t)} - |\bar{\mu}| \langle (\partial_{\bar{n}^x} w), [v] \rangle_{\Gamma(t)} - |\bar{\mu}| \langle (\partial_{\bar{n}^x} v), [w] \rangle_{\Gamma(t)} \\
 & + |\bar{\mu}| (\gamma h_K^{-1} [w], [v])_{\Gamma(t)} + ([\nabla w], [\nabla v])_{\Omega_O(t)}
 \end{aligned} \tag{3.16}$$

where  $|\bar{\mu}| = \sqrt{|\mu|^2 + 1}$ ,  $\langle v \rangle$  is a convex-weighted average of  $v$  on  $\Gamma$ , i.e.,  $\langle v \rangle = \omega_1 v_1 + \omega_2 v_2$ , where  $\omega_1, \omega_2 \in [0, 1]$  and  $\omega_1 + \omega_2 = 1$ ,  $\partial_{\bar{n}^x} v = \bar{n}^x \cdot \nabla v$ ,  $\gamma \geq 0$  is a stabilization parameter,  $h_K = h_K(x) = h_{K_0}$  for  $x \in K_0$ , where  $h_{K_0}$  is the diameter of simplex  $K_0 \in \mathcal{T}_0$ , and  $\Omega_O(t)$  is the overlap domain defined by (3.5). The reason for including the factor  $|\bar{\mu}|$  in the  $\Gamma(t)$  terms is that when considering spacetime, these terms should be on  $\bar{\Gamma}_n$ . Since  $|\bar{\mu}|$  is the skewed temporal scaling, we have that

$$\int_{I_n} |\bar{\mu}| (w, v)_{\Gamma(t)} dt = \int_{\bar{\Gamma}_n} w v d\bar{s} \tag{3.17}$$

**Remark** The method presented here is formulated with a discrete space  $V_h$  of arbitrary polynomial degree  $q$  in time. However, the main analytic results Lemma 5.1 and Theorem 6.1 are only presented for the cases  $q = 0, 1$ . This is because in the proofs of the underlying technical estimates Lemma A.10 and Lemma A.11, terms involving  $D_t^2 v$  for  $v \in V_h$  show up which we make vanish by simply assuming  $q \leq 1$ . To handle these terms for  $q > 1$  requires adding stabilization to the mass form. Here we choose not to do that in order to keep things simple for this first study and since we think that the method for  $q \leq 1$  is relevant and provides value.

### 4 Analytic preliminaries

#### 4.1 The bilinear form $A_{h,t}$

The space of  $A_{h,t}$  is  $H^{3/2+\varepsilon}(\cup_i \Omega_i(t))$  where  $\varepsilon > 0$  may be arbitrarily small. Let  $\Gamma_K(t) := K \cap \Gamma(t)$ . We define the following two mesh-dependent norms:

$$\|w\|_{1/2,h,\Gamma(t)}^2 := \sum_{K \in \mathcal{T}_0,\Gamma(t)} h_K^{-1} \|w\|_{\Gamma_K(t)}^2 \quad \|w\|_{-1/2,h,\Gamma(t)}^2 := \sum_{K \in \mathcal{T}_0,\Gamma(t)} h_K \|w\|_{\Gamma_K(t)}^2 \tag{4.1}$$

Note that

$$\|w\|_{\Gamma(t)}^2 \leq h \|w\|_{1/2,h,\Gamma(t)}^2 \quad (w, v)_{\Gamma(t)} \leq \|w\|_{-1/2,h,\Gamma(t)} \|v\|_{1/2,h,\Gamma(t)} \tag{4.2}$$

We define the time-dependent spatial energy norm  $\|\cdot\|_{A_{h,t}}$  by

$$\|w\|_{A_{h,t}}^2 := \sum_{i=1}^2 \|\nabla w\|_{\Omega_i(t)}^2 + |\bar{\mu}| \| \langle \partial_{\bar{n}^x} w \rangle \|_{-1/2,h,\Gamma(t)}^2$$

$$+|\bar{\mu}|\| [w] \|_{1/2,h,\Gamma(t)}^2 + \| [\nabla w] \|_{\Omega_O(t)}^2 \tag{4.3}$$

Continuity of  $A_{h,t}$  follows from using (4.2) in (3.16). Next we consider the coercivity:

**Lemma 4.1** [Discrete coercivity of  $A_{h,t}$ ] *Let the bilinear form  $A_{h,t}$  and the energy norm  $\| \cdot \|_{A_{h,t}}$  be defined by (3.16) and (4.3), respectively. Then, for  $t \in [0, T]$  and  $\gamma$  sufficiently large,*

$$A_{h,t}(v, v) \gtrsim \| v \|_{A_{h,t}}^2 \quad \forall v \in V_h(t) \tag{4.4}$$

**Proof** Following the proof of the coercivity in [2], we consider

$$\begin{aligned} 2|\bar{\mu}|(\langle \partial_{\bar{n}^x} v \rangle, [v])_{\Gamma(t)} &\leq \frac{|\bar{\mu}|}{\varepsilon} \| \langle \partial_{\bar{n}^x} v \rangle \|_{-1/2,h,\Gamma(t)}^2 + \varepsilon |\bar{\mu}| \| [v] \|_{1/2,h,\Gamma(t)}^2 \\ &\leq \frac{2|\bar{\mu}|}{\varepsilon} C_I \left( \sum_{i=1}^2 \| \nabla v \|_{\Omega_i(t)}^2 + \| [\nabla v] \|_{\Omega_O(t)}^2 \right) \\ &\quad - \frac{|\bar{\mu}|}{\varepsilon} \| \langle \partial_{\bar{n}^x} v \rangle \|_{-1/2,h,\Gamma(t)}^2 + \varepsilon |\bar{\mu}| \| [v] \|_{1/2,h,\Gamma(t)}^2 \end{aligned} \tag{4.5}$$

where we have used Lemma A.5 and denoted its constant by  $C_I$ . We use (4.5) in

$$\begin{aligned} A_{h,t}(v, v) &= \sum_{i=1}^2 \| \nabla v \|_{\Omega_i(t)}^2 - 2|\bar{\mu}|(\langle \partial_{\bar{n}^x} v \rangle, [v])_{\Gamma(t)} \\ &\quad + \gamma |\bar{\mu}| \| [v] \|_{1/2,h,\Gamma(t)}^2 + \| [\nabla v] \|_{\Omega_O(t)}^2 \\ &\geq \left( 1 - \frac{2|\bar{\mu}|C_I}{\varepsilon} \right) \sum_{i=1}^2 \| \nabla v \|_{\Omega_i(t)}^2 + \frac{|\bar{\mu}|}{\varepsilon} \| \langle \partial_{\bar{n}^x} v \rangle \|_{-1/2,h,\Gamma(t)}^2 \\ &\quad + (\gamma - \varepsilon) |\bar{\mu}| \| [v] \|_{1/2,h,\Gamma(t)}^2 + \left( 1 - \frac{2|\bar{\mu}|C_I}{\varepsilon} \right) \| [\nabla v] \|_{\Omega_O(t)}^2 \end{aligned} \tag{4.6}$$

By taking  $\varepsilon > 2|\bar{\mu}|C_I$ , and  $\gamma > \varepsilon$  we may obtain (4.4) from (4.6). □

### 4.2 The bilinear form $B_h$

The bilinear form  $B_h$  can be expressed differently, as noted in the following lemma:

**Lemma 4.2** [Alternative form of  $B_h$ ] *Let  $\zeta = \frac{1}{2}(3 - \text{sgn}(\bar{n}^t))$ . The bilinear form  $B_h$ , defined by (3.15), can be written as*

$$\begin{aligned} B_h(w, v) &= \sum_{i=1}^2 \sum_{n=1}^N \int_{I_n} (w, -\partial_t v)_{\Omega_i(t)} dt + \sum_{n=1}^N \int_{I_n} A_{h,t}(w, v) dt \\ &\quad + \sum_{n=1}^{N-1} (w_n^-, -[v]_n)_{\Omega_0} + (w_N^-, v_N^-)_{\Omega_0} + \sum_{n=1}^N \int_{\bar{\Gamma}_n} \bar{n}^t w_\zeta [v] d\bar{s} \end{aligned} \tag{4.7}$$

**Proof** The proof is analogous to the standard case. The first term in (3.15) is integrated by parts in time via  $\int_{S_{i,n}} (\nabla, \partial_t) \cdot (\mathbf{0}, wv) \, d\bar{x}$  and the result is combined with the last three terms in (3.15). The combination of purely time-jump-related terms is exactly as in the standard case. For the  $\bar{\Gamma}_n$ -integral terms, we let  $\zeta = \frac{1}{2}(3 - \text{sgn}(\bar{n}'))$ , if  $\sigma = \frac{1}{2}(3 + \text{sgn}(\bar{n}'))$  and  $\bar{n} = \bar{n}_1$ . This makes  $\zeta, \sigma \in \{1, 2\}$  and  $\zeta \neq \sigma$ .  $\square$

An important result for the analysis is obtained by first taking the same function as both arguments of  $B_h$ . We present this result as a coercivity of  $B_h$  with the following space-time energy norm:

$$\begin{aligned} \|v\|_{B_h}^2 &:= \sum_{n=1}^N \int_{I_n} \|v\|_{A_{h,t}}^2 \, dt \\ &+ \sum_{n=1}^{N-1} \|[v]_n\|_{\Omega_0}^2 + \|v_N^-\|_{\Omega_0}^2 + \|v_0^+\|_{\Omega_0}^2 + \sum_{n=1}^N \|\bar{n}'|^{1/2}[v]\|_{\bar{\Gamma}_n}^2 \end{aligned} \tag{4.8}$$

**Lemma 4.3** [Discrete coercivity of  $B_h$ ] *Let the bilinear form  $B_h$  and the energy norm  $\|\cdot\|_{B_h}$  be defined by (3.15) and (4.8), respectively. Then, for  $\gamma$  sufficiently large, we have that*

$$B_h(v, v) \gtrsim \|v\|_{B_h}^2 \quad \forall v \in V_h \tag{4.9}$$

**Proof** The proof is analogous to the standard case. First the same function  $v$  is taken as both arguments of  $B_h$ . Then the first term in (3.15) is integrated in time via  $\int_{S_{i,n}} (\nabla, \partial_t) \cdot (\mathbf{0}, v^2) \, d\bar{x}$  and the result is combined with the last three terms in (3.15). The combination of purely time-jump-related terms is exactly as in the standard case. For the  $\bar{\Gamma}_n$ -integral terms, we note from the interdependence of  $\sigma$  and  $\bar{n}$  that the combined integrand may be written as  $\bar{n}' \text{sgn}(\bar{n}') [v]^2$ . Also using Lemma 4.1 then shows the desired estimate.  $\square$

For the continued analysis, we define three space-time energy norms by

$$\|v\|_X^2 := \sum_{i=1}^2 \sum_{n=1}^N \int_{I_n} k_n \|D_t v\|_{\Omega_i(t)}^2 \, dt + \|v\|_{B_h}^2 \tag{4.10}$$

$$\|v\|_{Y_+}^2 := \sum_{n=1}^N \left( \int_{I_n} \frac{1}{k_n} \|v\|_{\Omega_0}^2 \, dt + \int_{I_n} \|v\|_{A_{h,t}}^2 \, dt + \|v_{n-1}^+\|_{\Omega_0}^2 \right) \tag{4.11}$$

$$\|v\|_{Y_-}^2 := \sum_{n=1}^N \left( \int_{I_n} \frac{1}{k_n} \|v\|_{\Omega_0}^2 \, dt + \int_{I_n} \|v\|_{A_{h,t}}^2 \, dt + \|v_n^-\|_{\Omega_0}^2 \right) \tag{4.12}$$

The  $X$ -norm is the main norm, meaning that it is in this norm that we obtain stability and error estimates. The  $Y$ -norms are auxiliary norms. We use the  $X$ -norm and  $Y$ -norms to obtain continuity of  $B_h$  which comes in two variants depending on the starting point, i.e., the standard form of  $B_h$  (3.15) or the alternative (4.7).

**Lemma 4.4** (Continuity of  $B_h$ ) *Let the bilinear form  $B_h$  be defined by (3.15) and the norms  $\|\cdot\|_X$ ,  $\|\cdot\|_{Y_+}$ , and  $\|\cdot\|_{Y_-}$  by (4.10), (4.11), and (4.12), respectively. Then for any functions  $w$  and  $v$  with sufficient spatial and temporal regularity we have that*

$$B_h(w, v) \lesssim \|w\|_X \|v\|_{Y_+} \tag{4.13}$$

$$B_h(w, v) \lesssim \|w\|_{Y_-} \|v\|_X \tag{4.14}$$

**Proof** The proofs of (4.13) and (4.14) are analogous so we only consider the latter since it gives the continuity result needed in the error analysis. The starting point is the alternative form of  $B_h$  (4.7). Applying the Cauchy–Schwarz inequality to all the terms (several times and different versions for some), (3.7) to split the first term followed by Corollary A.1 for the  $w$ -factor in the resulting  $\mu_i \cdot \nabla$ -part, the continuity of  $A_{h,t}$  in the treatment of the second term, and Lemma A.3 in the treatment of the fifth, we get product terms, where one factor may be estimated by  $\|w\|_{Y_-}$  and the other by  $\|v\|_X$ .  $\square$

Next, we present an estimate involving the bilinear form  $B_h$  and the  $X$ -norm that may be viewed as a counterpart to such a coercivity. Due to the appearance of the estimate, we call it “perturbed coercivity”. The estimate is a cornerstone of the energy analysis. It is fundamental to the stability analysis and also the starting point for deriving an inf-sup condition that in turn is essential for the error analysis. Key technical results used in the proof of the perturbed coercivity are Lemma A.8 and Lemma A.10.

**Lemma 4.5** [Discrete perturbed coercivity of  $B_h$ ] *Let the bilinear form  $B_h$  and the norm  $\|\cdot\|_X$  be defined by (3.15) and (4.10), respectively. Then, for  $q = 0, 1$ , and  $\gamma$  sufficiently large, there exists a constant  $\delta > 0$  such that*

$$B_h(v, v + \delta k_n D_t v) \gtrsim \|v\|_X^2 \quad \forall v \in V_h \tag{4.15}$$

**Proof** Using Lemma 4.3 with constant  $\beta > 0$ , the left-hand side of (4.15) is

$$B_h(v, v + \delta k_n D_t v) \geq \beta \|v\|_{B_h}^2 + B_h(v, \delta k_n D_t v) \tag{4.16}$$

The second term on the right-hand side is

$$\begin{aligned} B_h(v, \delta k_n D_t v) &= \sum_{i=1}^2 \sum_{n=1}^N \int_{I_n} (\partial_t v, \delta k_n D_t v)_{\Omega_i(t)} + \sum_{n=1}^N \int_{I_n} A_{h,t}(v, \delta k_n D_t v) \, dt \\ &+ \sum_{i=1}^2 \sum_{n=1}^{N-1} ([v]_n, (\delta k_n D_t v)_n^+)_{\Omega_{i,n}} + \sum_{i=1}^2 (v_0^+, (\delta k_n D_t v)_0^+)_{\Omega_{i,0}} \\ &+ \sum_{n=1}^N \int_{\bar{\Gamma}_n} -\bar{n}^t [v] (\delta k_n D_t v)_\sigma \, d\bar{s} \end{aligned} \tag{4.17}$$

The treatment of most of the terms involve the Cauchy–Schwarz inequality and for some also an  $\varepsilon$ -weighted Young’s inequality. The first term in (4.17) is split using

(3.7), where the  $D_t$ -part is good, and we use standard estimates for the  $\mu_i \cdot \nabla$ -part. For the second term in (4.17), we use the continuity of  $A_{h,t}$  followed by Lemma A.8. The third and fourth term in (4.17) are estimated by Lemma A.10. For the fifth and final term in (4.17), we use Lemma A.3 and Lemma A.8. Collecting all the estimates and using the result in (4.16), we may obtain

$$\begin{aligned}
 B_h(v, v + \delta k_n D_t v) &\geq \delta \left(1 - \frac{\delta}{\varepsilon} C\right) \sum_{i=1}^2 \sum_{n=1}^N \int_{I_n} k_n \|D_t v\|_{\Omega_i(t)}^2 dt \\
 &\quad + \left(\beta - \left(\varepsilon + \delta + \frac{\delta^2}{\varepsilon}\right) C\right) \|v\|_{B_h}^2
 \end{aligned}
 \tag{4.18}$$

where  $C > 0$  denote various constants. First taking  $\varepsilon > 0$  sufficiently small and then taking  $\delta > 0$  sufficiently small gives the desired estimate.  $\square$

Using Lemma 4.5 and Lemma A.11, we may obtain the discrete inf-sup condition:

**Corollary 4.1** (A discrete inf-sup condition for  $B_h$ ) *Let the bilinear form  $B_h$  and the norm  $\|\cdot\|_X$  be defined by (3.15) and (4.10), respectively. Then, for  $q = 0, 1$ , and  $\gamma$  sufficiently large, we have that*

$$\|w\|_X \lesssim \sup_{v \in V_h \setminus \{0\}} \frac{B_h(w, v)}{\|v\|_X} \quad \forall w \in V_h
 \tag{4.19}$$

To show Galerkin orthogonality, we need the following lemma on consistency:

**Lemma 4.6** (Consistency) *The solution  $u$  to problem (2.5) also solves (3.14).*

**Proof** First insert  $u$  in place of  $u_h$  on the left-hand side of (3.14) and use the regularity of  $u$ . Then integrate by parts in space via  $\int_{S_{i,n}} (\nabla, \partial_t) \cdot (\nabla u v, 0) d\bar{x}$  to get interior and boundary terms. The exterior boundary terms vanish because of the boundary conditions imposed on  $v$  thus leaving the  $\Gamma$ -terms which are combined. Applying Lemma A.1 and the regularity of  $u$  only leaves terms which from (2.5) equals the right-hand side of (3.14).  $\square$

From Lemma 4.6, we may obtain the Galerkin orthogonality:

**Corollary 4.2** (Galerkin orthogonality) *Let the bilinear form  $B_h$  be defined by (3.15), and let  $u$  and  $u_h$  be the solutions of (2.5) and (3.14), respectively. Then*

$$B_h(u - u_h, v) = 0 \quad \forall v \in V_h
 \tag{4.20}$$

### 5 Stability analysis

In this section we present and prove a stability estimate for the solution  $u_h$  to (3.14). The key component in the proof is Lemma 4.5, i.e., the perturbed coercivity of  $B_h$  on  $V_h$ .



**Lemma 5.1** (A stability estimate in  $\|\cdot\|_X$ ) *Let  $u_h$  be the solution of (3.14). Let  $u_0$  and  $f$  be the initial data and source in (2.5), respectively. Then, for  $q = 0, 1$ , and  $\gamma$  sufficiently large, we have that*

$$\|u_h\|_X \lesssim \|u_0\|_{\Omega_0} + \|f\|_{L^2((0,T];L^2(\Omega_0))} \tag{5.1}$$

**Proof** By taking  $v = u_h \in V_h$  in Lemma 4.5 and  $v = u_h + \delta k_n D_t u_h \in V_h$  in (3.14), we have

$$\begin{aligned} \|u_h\|_X^2 &\lesssim B_h(u_h, u_h + \delta k_n D_t u_h) \\ &= (u_0, u_{h,0}^+)_{\Omega_0} + (u_0, \delta k_1 (D_t u_h)_0^+)_{\Omega_0} \\ &\quad + \sum_{n=1}^N \int_{I_n} (f, u_h)_{\Omega_0} dt + \sum_{n=1}^N \int_{I_n} (f, \delta k_n D_t u_h)_{\Omega_0} dt \end{aligned} \tag{5.2}$$

Applying the Cauchy–Schwarz inequality to all the terms (several times and different versions for some), Lemma A.10 in the treatment of the second term, and Corollary A.1 in the treatment of the third, we get product terms, where one factor is  $\|u_0\|_{\Omega_0}$  or  $\|f\|_{L^2((0,T];L^2(\Omega_0))}$  and the other may be estimated by  $\|u_h\|_X$ . Dividing both sides by  $\|u_h\|_X$  thus gives (5.1). □

### 6 A priori error analysis

**Theorem 6.1** (An optimal order a priori error estimate in  $\|\cdot\|_X$ ) *Let  $\|\cdot\|_X$  be defined by (4.10), let  $u$  be the solution of (2.5) and let  $u_h$  be the finite element solution defined by (3.14). Then, for  $q = 0, 1$ , and  $\gamma$  sufficiently large, we have that*

$$\|u - u_h\|_X^2 \lesssim k^{2q+1} F_k^2(u) + h^{2p} \left( F_h^2(u) + E_{h,1}^2(u) \right) \tag{6.1}$$

where  $F_k, F_h$ , and  $E_{h,1}$  are defined by (B.25), (B.26), and (B.23), respectively.

**Proof** We use the interpolant  $\bar{I}_h u \in V_h$ , where  $\bar{I}_h$  is the space-time interpolation operator defined by (B.19), to split the error  $e = u - u_h$  into  $\rho = u - \bar{I}_h u$  and  $\theta = \bar{I}_h u - u_h$ . Thus

$$\|e\|_X \leq \|\rho\|_X + \|\theta\|_X \tag{6.2}$$

where we focus on the  $\theta$ -part first. From Corollary 4.2, i.e., Galerkin orthogonality, we have for any  $v \in V_h$  that

$$B_h(\theta, v) = -B_h(\rho, v) \tag{6.3}$$

We note that  $\theta \in V_h$  and use Corollary 4.1, i.e., a discrete inf-sup condition for  $B_h$ , the Galerkin orthogonality result (6.3), and Lemma 4.4, i.e., continuity of  $B_h$ , to estimate

the  $\theta$ -part by

$$\begin{aligned} \|\theta\|_X &\lesssim \sup_{v \in V_h \setminus \{0\}} \frac{B_h(\theta, v)}{\|v\|_X} = \sup_{v \in V_h \setminus \{0\}} \frac{-B_h(\rho, v)}{\|v\|_X} \\ &\lesssim \sup_{v \in V_h \setminus \{0\}} \frac{\|\rho\|_{Y_-} \|v\|_X}{\|v\|_X} = \|\rho\|_{Y_-} \end{aligned} \tag{6.4}$$

Using (6.4) in (6.2), we estimate the approximation error by

$$\begin{aligned} \|e\|_X^2 &\lesssim \|\rho\|_X^2 + \|\rho\|_{Y_-}^2 \\ &\lesssim \sum_{i,n} \left( k_n \int_{I_n} \|D_t \rho\|_{\Omega_i(t)}^2 dt + \frac{1}{k_n} \int_{I_n} \|\rho\|_{\Omega_i(t)}^2 dt \right) + \|\rho\|_{B_h}^2 + \sum_{n=1}^N \|\rho_n^-\|_{\Omega_0}^2 \end{aligned} \tag{6.5}$$

By applying various interpolation error estimates: Lemma B.4 and using (3.1) for the first term, Lemma B.5 for the second, and Corollary B.1 for the third, we get results that may be estimated by the right-hand side of (6.1). □

### 7 Numerical results

The implementation used to obtain the numerical results is freely available online at <https://github.com/Carl-Lundholm/STCutFEMOverlapMesh>.

Here we present numerical results for a problem in one spatial dimension on the unit interval with exact solution  $u(x, t) = \sin^2(\pi x)e^{-t/2}$ . We compute  $u_h$  for  $p = 1$  and  $q = 0, 1$ . For dG(1) in time, *some* of the left-hand side integrals involving time have been approximated locally by quadrature. For integrals over cut space-time prisms, *composite three-point* Lobatto quadrature has been used in time. By this we mean one quadrature rule for *each* temporal part of a cut space-time prism where the polynomial degree of the integrand is unchanged. For integrals over intraprismatic segments of the space-time boundary  $\bar{\Gamma}_n$ , *three-point* Lobatto quadrature has been used. Both of these choices of quadrature result in a quadrature error =  $O(k^4)$ . The right-hand side integrals have been approximated locally by quadrature over the space-time prisms: first *composite* quadrature in time, then quadrature in space. In space, the trapezoidal rule has been used, thus resulting in a quadrature error =  $O(h^2)$ . For dG(0) in time, the *composite* midpoint rule has been used, thus resulting in a quadrature error =  $O(k^2)$ . For dG(1) in time, *composite three-point* Lobatto quadrature has been used, thus resulting in a quadrature error =  $O(k^4)$ . For simplicity, the velocity  $\mu$  of the overlapping mesh is set to be constant at the value  $\mu(t_n)$  on every subinterval  $I_n = (t_{n-1}, t_n]$ . The stabilization parameter  $\gamma = 10$ .

For the error convergence study, both  $\mathcal{T}_0$  and  $\mathcal{T}_G$  are uniform meshes, with mesh sizes  $h_0$  and  $h_G$ , respectively. The temporal discretization is also uniform with time step  $k$  for each instance. The final time is set to  $T = 1$ , the length of  $\mathcal{T}_G$  is 0.25, and

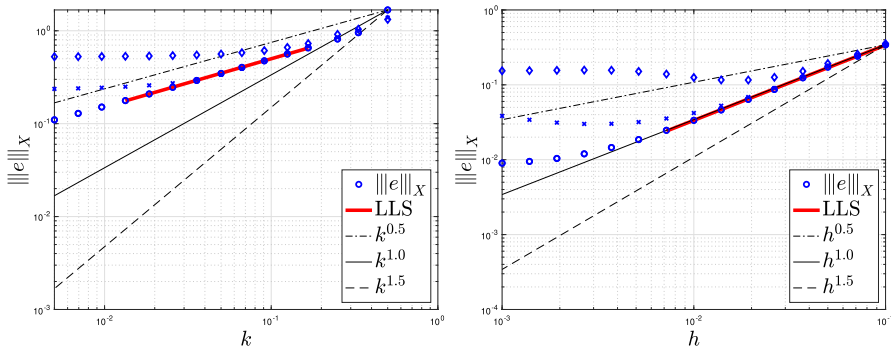


Fig. 7 Error convergence for dG(0) with  $\mu = 0.6$

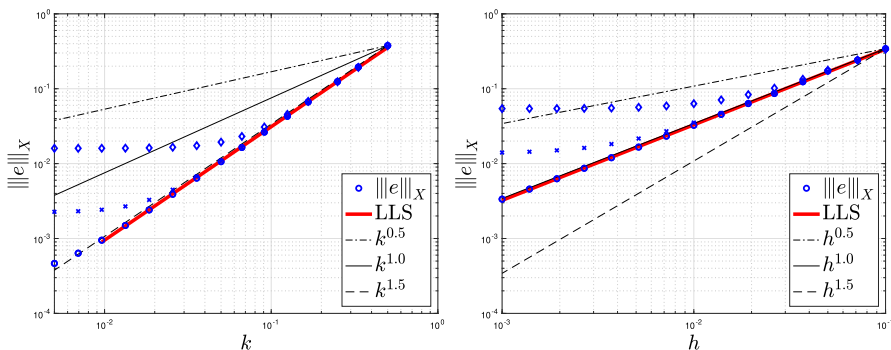


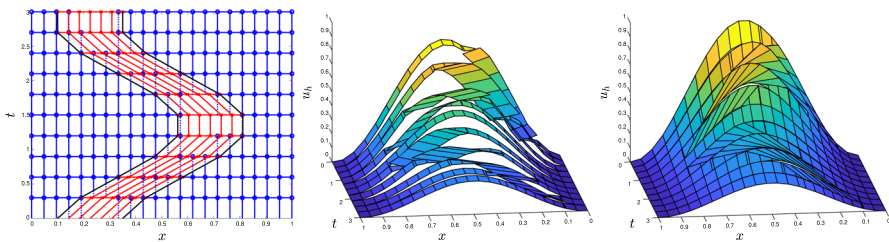
Fig. 8 Error convergence for dG(1) with  $\mu = 0.6$

the initial position of  $\mathcal{T}_G$  is the spatial interval  $[0.125, 0.125 + 0.25]$ . The error is  $\|e\|_X = \|u - u_h\|_X$ . All time, space, and space-time integrals involving  $u$  in the  $X$ -norm have been approximated locally by *three-point* Gauss-Legendre quadrature: first *composite* quadrature in time, then quadrature in space where applicable. This results in a quadrature error  $= O((k^6 + h^6)^{1/2})$ . In the  $k$ -convergence study, the mesh sizes have been fixed at  $h = 1.5 \cdot 10^{-1}, 7 \cdot 10^{-2}, 10^{-3}$  for dG(0) and  $h = 5 \cdot 10^{-3}, 7 \cdot 10^{-4}, 10^{-4}$  for dG(1). Analogously, in the  $h$ -convergence study, the time step has been fixed at  $k = 10^{-2}, 10^{-3}, 10^{-4}$  for dG(0) and  $k = 1.5 \cdot 10^{-1}, 6 \cdot 10^{-2}, 10^{-2}$  for dG(1). Figures 7 and 8 display error convergence plots for dG(0) and dG(1) in time with  $\mu = 0.6$ . The left plots show the error versus  $k$  and the right plots versus  $h = h_0 \geq h_G$ . Besides the computed error for different fixed values of  $h$  or  $k$ , each plot contains a line segment that has been computed with the linear least squares method to fit the error data for the smallest fixed value of  $h$  or  $k$ . This line segment is referred to as the LLS of the error. Reference slopes are also included. In Table 1 we summarize the slope of the LLS of the error for different values of  $\mu$ .

The numerical solutions presented in Fig. 9 have been computed for an equidistant space-time discretization: 22 nodes for  $\mathcal{T}_0$ , 7 nodes for  $\mathcal{T}_G$  for all times, and 10 time steps on the interval  $(0, 3]$ . The length of  $\mathcal{T}_G$  has again been 0.25 and the velocity  $\mu$  has for simplicity been slabwise constant at  $\mu|_{I_n} = \frac{1}{2} \sin(\frac{2\pi I_n}{3})$ .

**Table 1** The slope of the LLS of the *error* versus *k* and *h* for different values of  $\mu$

$\mu$	dG(0) in time		dG(1) in time	
	Versus <i>k</i> (points)	Versus <i>h</i> (points)	Versus <i>k</i> (points)	Versus <i>h</i> (points)
0	0.5058 (4–12)	1.0190 (1–9)	1.4506 (1–5)	1.0137 (1–15)
0.1	0.5026 (4–12)	1.0205 (1–9)	1.4893 (1–7)	1.0155 (1–15)
0.2	0.4985 (4–12)	1.0221 (1–9)	1.4947 (1–8)	1.0162 (1–15)
0.4	0.5163 (4–12)	1.0179 (1–9)	1.5031 (1–11)	1.0147 (1–15)
0.6	0.5179 (4–12)	1.0047 (1–9)	1.5151 (1–13)	1.0091 (1–15)



**Fig. 9** Space-time discretization (left) with resulting dG(0)cG(1)-solution (middle) and dG(1)cG(1)-solution (right)

### 8 Conclusions

We have presented a cut finite element method for a parabolic model problem on an overlapping mesh situation: one stationary background mesh and one *continuously* moving overlapping mesh. We have applied what we believe to be a relatively uncommon analysis framework for finite element methods for parabolic problems. This analysis framework may arguably be considered more robust and natural than standard ones, since it is the only one that we have been able to successfully apply to our overlapping mesh situation. The analysis is of an energy type and the main results are a basic stability estimate and an optimal order a priori error estimate. We have also presented numerical results for a parabolic problem in one spatial dimension that verify the analytic error convergence orders.

### A Analytic tools

**Lemma A.1** (A jump identity) *Let  $\omega_+, \omega_- \in \mathbb{R}$  and  $\omega_+ + \omega_- = 1$ , let  $[A] := A_+ - A_-$ , and  $\langle A \rangle := \omega_+ A_+ + \omega_- A_-$ . We then have*

$$[AB] = [A]\langle B \rangle + \langle A \rangle[B] + (\omega_- - \omega_+)[A][B] \tag{A.1}$$

**Proof** Using the definitions and evaluating both sides shows the identity. □

### A.1 Spatial estimates

**Lemma A.2** [A Poincaré inequality for  $H_0^1(\cup_i \Omega_i(t))$ ] For  $t \in [0, T]$  we have that

$$\|v\|_{\Omega_0} \lesssim \|\nabla v\|_{\Omega_1(t) \cup \Omega_2(t)} + \|[v]\|_{\Gamma(t)} \quad \forall v \in H_0^1(\cup_i \Omega_i(t)) \tag{A.2}$$

**Proof** To lighten the notation we omit the time dependence, which has no importance here anyways. For  $v \in H_0^1(\Omega_1 \cup \Omega_2)$ , we consider the dual problem: Find  $\phi \in H^2(\Omega_0) \cap H_0^1(\Omega_0)$  such that  $-\Delta\phi = v$  in  $\Omega_0$ . By using the dual problem, partial integration, that  $v|_{\partial\Omega_0} = 0$ , Lemma A.1, and the regularity of  $\phi$  ( $[\partial_n\phi]|_{\Gamma} = 0$  in  $L^2(\Gamma)$ ), we have

$$\|v\|_{\Omega_0}^2 = \sum_{i=1}^2 (-\Delta\phi, v)_{\Omega_i} = \sum_{i=1}^2 (\nabla\phi, \nabla v)_{\Omega_i} - ((\partial_n\phi), [v])_{\Gamma} \tag{A.3}$$

Using a standard trace inequality for  $\nabla\phi|_{\Omega_i} \in H^1(\Omega_i)$ , elliptic regularity on  $H^2(\Omega_0) \cap H_0^1(\Omega_0)$  for  $\phi$ , and the dual problem, the first argument to the last inner product may be estimated by

$$\|(\partial_n\phi)\|_{\Gamma} \leq \sum_{i=1}^2 \|\nabla\phi_i\|_{\partial\Omega_i} \lesssim \sum_{i=1}^2 \|\nabla\phi\|_{1,\Omega_i} \lesssim \|\phi\|_{2,\Omega_0} \lesssim \|\Delta\phi\|_{\Omega_0} = \|v\|_{\Omega_0} \tag{A.4}$$

We note that this also gives an estimate for the first argument to the penultimate inner product. Thus using (A.4) in (A.3) followed by cancellation of a factor  $\|v\|_{\Omega_0}$  on both sides gives (A.2). □

By squaring both sides of (A.2), using Young’s inequality, and (4.2), we may estimate the resulting right-hand side by  $\|\cdot\|_{A_{h,t}}$ :

**Corollary A.1** (An energy Poincaré inequality for  $H^{3/2+\varepsilon}(\cup_i \Omega_i(t)) \cap H_0^1(\cup_i \Omega_i(t))$ ) Let the time-dependent spatial energy norm  $\|\cdot\|_{A_{h,t}}$  be defined by (4.3). Then, for  $t \in [0, T]$ , we have that

$$\|v\|_{\Omega_0} \lesssim \|v\|_{A_{h,t}} \quad \forall v \in H^{3/2+\varepsilon}(\cup_i \Omega_i(t)) \cap H_0^1(\cup_i \Omega_i(t)) \tag{A.5}$$

**Lemma A.3** (A spatial continuity result for  $\Gamma(t)$ ) Let the space-time vector  $\bar{n} = (\bar{n}^x, \bar{n}^t)$  and the time-dependent spatial energy norm  $\|\cdot\|_{A_{h,t}}$  be defined by (3.10) and (4.3), respectively. Let  $\sigma$  change arbitrarily along  $\Gamma(t)$  between the values 1 and 2 and let  $|\mu|_{[0,T]} = \max_{t \in [0,T]} \{|\mu|\}$ . Then, for  $t \in [0, T]$ , we have that

$$\begin{aligned} (\bar{n}^t[w], v_{\sigma})_{\Gamma(t)} &\lesssim |\mu|_{[0,T]} h^{1/2} \|w\|_{A_{h,t}} \|v\|_{A_{h,t}} \\ \forall w, v \in H^{3/2+\varepsilon}(\cup_i \Omega_i(t)) \cap H_0^1(\cup_i \Omega_i(t)) \end{aligned} \tag{A.6}$$

**Proof** To lighten the notation we omit the time dependence, which has no importance here anyways. Using  $|\bar{n}^t| \leq |\mu|$ , which follows from (3.10), the left-hand side of (A.6) is

$$(\bar{n}^t[w], v_\sigma)_\Gamma \leq |\mu|_{[0,T]} \| [w] \|_\Gamma \| v_\sigma \|_\Gamma \tag{A.7}$$

Using (4.2) and (4.3), the  $w$ -factor may be estimated by  $h^{1/2} \| \| w \| \|_{A_{h,t}}$ . Applying the standard trace inequality for  $H^1(\Omega_i)$ , Corollary A.1, and (4.3), the  $v$ -factor may be estimated by  $\| \| v \| \|_{A_{h,t}}$ . This shows (A.6).  $\square$

**Lemma A.4** (A scaled trace inequality for domain-partitioning manifolds of codimension 1) *For  $d = 1, 2$ , or  $3$ , let  $\Omega \subset \mathbb{R}^d$  be a bounded domain with diameter  $L$ , i.e.,  $L = \text{diam}(\Omega) = \sup_{x,y \in \Omega} |x - y|$ . Let  $\Gamma \subset \Omega$  be a continuous manifold of codimension 1 that partitions  $\Omega$  into  $N$  subdomains. Then*

$$\| v \|_\Gamma^2 \lesssim L^{-1} \| v \|_\Omega^2 + L \| \nabla v \|_\Omega^2 \quad \forall v \in H^1(\Omega) \tag{A.8}$$

**Proof** If (A.8) holds for the case  $N = 2$ , then that result may be applied repeatedly to show (A.8) for  $N > 2$ . We thus assume that  $\Gamma$  partitions  $\Omega$  into two subdomains denoted  $\Omega_1$  and  $\Omega_2$  with diameters  $L_1$  and  $L_2$ , respectively. From the regularity assumptions on  $v$ , we have for  $i = 1, 2$ , that  $v \in H^1(\Omega_i)$  and thus

$$\| v \|_\Gamma^2 \leq \| v \|_{\partial\Omega_i}^2 \lesssim L_i^{-1} \| v \|_{\Omega_i}^2 + L_i \| \nabla v \|_{\Omega_i}^2 \tag{A.9}$$

where we have used a standard scaled trace inequality. Using the triangle type inequality  $L \leq L_1 + L_2$  and (A.9), the left-hand side of (A.8) is

$$\| v \|_\Gamma^2 \leq \sum_{i=1}^2 \frac{L_i}{L} \| v \|_\Gamma^2 \lesssim \sum_{i=1}^2 \left( L^{-1} \| v \|_{\Omega_i}^2 + L \| \nabla v \|_{\Omega_i}^2 \right) \lesssim L^{-1} \| v \|_\Omega^2 + L \| \nabla v \|_\Omega^2 \tag{A.10}$$

which shows (A.8).  $\square$

Let  $\Gamma_K = \Gamma_K(t) = K \cap \Gamma(t)$ . For  $t \in [0, T]$ ,  $j \in \{0, G\}$ , a simplex  $K \in \mathcal{T}_{j,\Gamma(t)} = \{K \in \mathcal{T}_j : K \cap \Gamma(t) \neq \emptyset\}$ , and  $v \in H^1(K)$ , we have from Lemma A.4 that

$$\| v \|_{\Gamma_K}^2 \lesssim h_K^{-1} \| v \|_K^2 + h_K \| \nabla v \|_K^2 \tag{A.11}$$

where  $h_K$  is the diameter of  $K$ . For  $v \in \mathcal{P}(K)$ , we have the standard inverse estimate

$$\| D_x^k v \|_K^2 \lesssim h_K^{-2} \| D_x^{k-1} v \|_K^2 \quad \text{for } k \geq 1 \tag{A.12}$$

Using (A.12) in (A.11), we get the following corollary:

**Corollary A.2** (A discrete spatial local inverse inequality for  $\Gamma_K(t)$ ) For  $t \in [0, T]$ ,  $j \in \{0, G\}$ ,  $K \in \mathcal{T}_{j,\Gamma(t)}$  with diameter  $h_K$ , let  $\Gamma_K(t) = K \cap \Gamma(t)$ . Then, for  $k \geq 0$ , we have that

$$\|D_x^k v\|_{\Gamma_K(t)}^2 \lesssim h_K^{-1} \|D_x^k v\|_K^2 \quad \forall v \in V_h(t) \tag{A.13}$$

**Lemma A.5** (A discrete spatial inverse inequality for  $\Gamma(t)$ ) Let the mesh-dependent norm  $\|\cdot\|_{-1/2,h,\Gamma(t)}$  be defined by (4.1). Then, for  $t \in [0, T]$ , we have that

$$\|\langle \partial_{\bar{n}^x} v \rangle\|_{-1/2,h,\Gamma(t)}^2 \lesssim \sum_{i=1}^2 \|\nabla v\|_{\Omega_i(t)}^2 + \|[\nabla v]\|_{\Omega_O(t)}^2 \quad \forall v \in V_h(t) \tag{A.14}$$

**Proof** To lighten the notation we omit the time dependence, which has no importance here anyways. We follow the proof of the corresponding inequality in [2] with some modifications. We use index  $j \in \{0, G\}$ , such that, if  $j = 0$ , then  $i = 1$  and if  $j = G$ , then  $i = 2$ , and let  $\Gamma_{K_j} = K_j \cap \Gamma$  and  $\mathcal{T}_{j,\Gamma} = \{K_j \in \mathcal{T}_j : K_j \cap \Gamma \neq \emptyset\}$ . Note that for  $i = 1, 2$ ,

$$\sum_{K_0 \in \mathcal{T}_{0,\Gamma}} h_{K_0} \|v_i\|_{\Gamma_{K_0}}^2 \lesssim \sum_{K_G \in \mathcal{T}_{G,\Gamma}} h_{K_G} \|v_i\|_{\Gamma_{K_G}}^2 \tag{A.15}$$

which follows from  $\cup_{K_0 \in \mathcal{T}_{0,\Gamma}} \Gamma_{K_0} = \Gamma = \cup_{K_G \in \mathcal{T}_{G,\Gamma}} \Gamma_{K_G}$  and the inter-quasi-uniformity of the meshes. Since  $\partial_{\bar{n}^x} v = \bar{n}^x \cdot \nabla v$  and  $|\omega_i| |\bar{n}^x| \leq 1$ , we have  $\|\omega_i (\partial_{\bar{n}^x} v)_i\|_{\Gamma_{K_j}}^2 \leq \|(\nabla v)_i\|_{\Gamma_{K_j}}^2$ . Using this after (A.15), and followed by Corollary A.2, the left-hand side of (A.14) is

$$\begin{aligned} \|\langle \partial_{\bar{n}^x} v \rangle\|_{-1/2,h,\Gamma}^2 &\lesssim \sum_{i=1}^2 \sum_{K_j \in \mathcal{T}_{j,\Gamma}} h_{K_j} \|(\nabla v)_i\|_{\Gamma_{K_j}}^2 \lesssim \sum_{i=1}^2 \sum_{K_j \in \mathcal{T}_{j,\Gamma}} \|(\nabla v)_i\|_{K_j}^2 \\ &= \sum_{K_0 \in \mathcal{T}_{0,\Gamma}} \left( \|\nabla v\|_{K_0 \cap \Omega_1}^2 + \|(\nabla v)_1\|_{K_0 \cap \Omega_2}^2 \right) + \sum_{K_G \in \mathcal{T}_{G,\Gamma}} \|\nabla v\|_{K_G}^2 \end{aligned} \tag{A.16}$$

The resulting terms may be estimated by the right-hand side of (A.14). □

### A.2 Temporal estimates

Recall the domain-dependent velocity  $\mu_i$ , defined by (3.6). For a time  $t^* \in I_n$ , a point  $x \in \Omega_i(t^*)$  and a point  $s \in \Gamma(t^*)$ , approached from  $\Omega_i(t^*)$ , we define the spatial components  $\hat{x}(t)$  and  $\hat{s}_i(t)$  of the slabwise space-time trajectory through  $x$  and that through  $s$ , respectively, by

$$\hat{x}(t) := x + \int_{t^*}^t \mu_i(\tau) \, d\tau \quad \forall t \in I_n \tag{A.17}$$

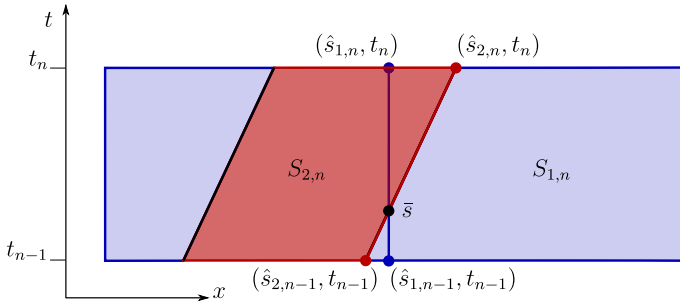


Fig. 10 Slabwise space-time trajectories through a point  $\bar{s} \in \bar{\Gamma}_n$  for  $d = 1$

$$\hat{s}_i(t) := s + \int_{t^*}^t \mu_i(\tau) \, d\tau \quad \forall t \in I_n \tag{A.18}$$

For  $i = 1$ , we get a straight space-time trajectory parallel to the time axis. For  $i = 2$ , we simply follow a point along the space-time surface  $\bar{\Gamma}_n$ . See Fig. 10 for an illustration. To lighten the notation, we omit the index  $i$  and the time dependence when there is no risk of confusion. Thus  $(\hat{s}, t) = (\hat{s}_i(t), t)$  and  $\hat{s}_k = \hat{s}_{i,k} = \hat{s}_i(t_k)$ , if not explicitly stated otherwise.

**Lemma A.6** (Discrete temporal inverse estimates in  $\|\cdot\|_{\Omega(t)}$ ) *Let  $k_n$  be the length of interval  $I_n$  and the scaled differential operator  $D_t$  be defined by (3.7). For  $v \in V_h^n$ , let  $w = w_r = D_x^r v$ , where  $0 \leq r \leq p$ . Then, for any  $v \in V_h^n$ , we have that*

$$\int_{I_n} k_n^2 \|D_t w\|_{\Omega_1(t) \cup \Omega_0(t)}^2 \, dt \lesssim \int_{I_n} \|w\|_{\Omega_1(t) \cup \Omega_0(t)}^2 \, dt \tag{A.19}$$

$$\int_{I_n} k_n^2 \|D_t w\|_{\Omega_2(t)}^2 \, dt \lesssim \int_{I_n} \|w\|_{\Omega_2(t)}^2 \, dt \tag{A.20}$$

**Proof** The estimates follow from applying a standard one-dimensional inverse estimate for polynomials along the space-time trajectories. The presence of  $D_t$  in the  $I_n$ -integrals gives the correct scaling for going to the space-time trajectories and back.  $\square$

**Lemma A.7** (An inequality for  $W^{1,1}((a, b))$ ) *For an open interval  $(a, b)$ , a point  $c \in (a, b)$ , and for any function  $w \in W^{1,1}((a, b))$  it holds that*

$$(b - a)w(c) \leq \int_a^b w(x) \, dx + (b - a) \int_a^b |w'(x)| \, dx \tag{A.21}$$

**Proof** Consider an open interval  $(\alpha, \beta) \subset (a, b)$ . For an arbitrary point  $y \in (\alpha, \beta)$ , we use integration by parts to get

$$(\beta - y)w(\beta^-) + (y - \alpha)w(\alpha^+) \leq \int_\alpha^\beta w(x) \, dx + (\beta - \alpha) \int_\alpha^\beta |w'(x)| \, dx \tag{A.22}$$



The left-hand side of (A.21) is

$$\begin{aligned}
 (b - a)w(c) &= (b - c)w(c) + (c - a)w(c) \\
 &\leq \int_c^b w(x) \, dx + (b - c) \int_c^b |w'(x)| \, dx \\
 &\quad + \int_a^c w(x) \, dx + (c - a) \int_a^c |w'(x)| \, dx \\
 &\leq \int_a^b w(x) \, dx + (b - a) \int_a^b |w'(x)| \, dx
 \end{aligned}
 \tag{A.23}$$

where we have used (A.22) with  $y = \beta^- = \beta = b$  and  $\alpha = c$ , and (A.22) with  $\beta = c$  and  $y = \alpha^+ = \alpha = a$  to obtain the first inequality. This concludes the proof.  $\square$

**Lemma A.8** (A discrete temporal inverse estimate in  $\|\cdot\|_{A_{h,t}}$ ) *Let  $\|\cdot\|_{A_{h,t}}$  be defined by (4.3),  $k_n$  be the length of interval  $I_n$ , and the scaled differential operator  $D_t$  be defined by (3.7). Then we have that*

$$\int_{I_n} \|k_n D_t v\|_{A_{h,t}}^2 \, dt \lesssim \int_{I_n} \|v\|_{A_{h,t}}^2 \, dt \quad \forall v \in V_h^n \tag{A.24}$$

**Proof** We expand the left-hand side of (A.24) by using (4.3)

$$\begin{aligned}
 \int_{I_n} \|k_n D_t v\|_{A_{h,t}}^2 \, dt &= \underbrace{\sum_{i=1}^2 \int_{I_n} k_n^2 \|\nabla D_t v\|_{\Omega_i(t)}^2 \, dt}_{=I} + \underbrace{\int_{I_n} k_n^2 |\bar{\mu}| \|\langle \partial_{\bar{n}x} D_t v \rangle\|_{-1/2,h,\Gamma(t)}^2 \, dt}_{=II} \\
 &\quad + \underbrace{\int_{I_n} k_n^2 |\bar{\mu}| \| [D_t v] \|_{1/2,h,\Gamma(t)}^2 \, dt}_{=III} + \underbrace{\int_{I_n} k_n^2 \| [\nabla D_t v] \|_{\Omega_O(t)}^2 \, dt}_{=IV}
 \end{aligned}
 \tag{A.25}$$

We treat the terms separately, starting with the first. Using that  $\nabla D_t v = D_t \nabla v$  and Lemma A.6, the first term in (A.25) is

$$\begin{aligned}
 I &= \sum_{i=1}^2 \int_{I_n} k_n^2 \|D_t \nabla v\|_{\Omega_i(t)}^2 \, dt \lesssim \int_{I_n} \sum_{i=1}^2 \|\nabla v\|_{\Omega_i(t)}^2 + \|[\nabla v]\|_{\Omega_O(t)}^2 \, dt \\
 &\leq \int_{I_n} \|v\|_{A_{h,t}}^2 \, dt
 \end{aligned}
 \tag{A.26}$$

The second term in (A.25) receives the same treatment after first using Lemma A.5, thus

$$II \lesssim \int_{I_n} k_n^2 |\bar{\mu}| \left( \sum_{i=1}^2 \|\nabla D_t v\|_{\Omega_i(t)}^2 + \|[\nabla D_t v]\|_{\Omega_O(t)}^2 \right) \, dt \lesssim \int_{I_n} \|v\|_{A_{h,t}}^2 \, dt \tag{A.27}$$

The third term in (A.25) requires some more work than the others. Recall the slabwise space-time trajectories through a point  $\bar{s} \in \bar{\Gamma}_n$ , whose spatial components  $\hat{s} = \hat{s}_i(t)$  are defined by (A.18). Let  $S_n^q$  denote the set of points in  $I_n$  corresponding to temporal degrees of freedom for  $V_h^n$ . Thus  $S_n^0 = \{t_n\}$  and  $S_n^1 = \{t_n, t_{n-1}^+\}$ . For  $q > 1$ , interior points of  $I_n$  are also included in  $S_n^q$ . We consider the temporal basis functions  $\lambda_k \in \mathcal{P}^q(I_n)$ , where every  $\lambda_k$  corresponds to a point  $t_k \in S_n^q$ . Writing  $\hat{x}_k = \hat{x}(t_k)$ , where  $\hat{x}$  is defined by (A.17), and using a somewhat relaxed notation, any  $v \in V_h^n$  may be represented as  $v(x, t) = \sum_{t_k \in S_n^q} v(\hat{x}_k, t_k) \lambda_k(t)$ . With simple continuous mesh motion,  $\mu$  is constant along every slabwise space-time trajectory, which means that  $D_t v(\hat{x}_k, t_k) = 0$ . Using this together with the somewhat relaxed representation, we have that

$$D_t v(x, t) = \sum_{t_k \in S_n^q} v(\hat{x}_k, t_k) D_t \lambda_k(t) = \sum_{t_k \in S_n^q} v(\hat{x}_k, t_k) \lambda'_k(t) \tag{A.28}$$

With (A.28), the third term in (A.25) is

$$\begin{aligned} \text{III} &\leq \frac{k_n^2}{h_{\min}} \int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} |(D_t v(s, t))_1 - (D_t v(s, t))_2|^2 ds dt \\ &\leq \frac{k_n^2}{h_{\min}} \int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} \left( \sum_{t_k \in S_n^q} |v_1(\hat{s}_{1,k}, t_k) - v_2(\hat{s}_{2,k}, t_k)| \underbrace{|\lambda'_k(t)|}_{\leq C(q)/k_n} \right)^2 ds dt \tag{A.29} \\ &\lesssim h_{\min}^{-1} \sum_{t_k \in S_n^q} \underbrace{\int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} |v_1(\hat{s}_{1,k}, t_k) - v_2(\hat{s}_{2,k}, t_k)|^2 ds dt}_{=\text{III}.k} \end{aligned}$$

We split III.k by

$$\begin{aligned} \text{III}.k &\lesssim \underbrace{\int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} |v_1(\hat{s}_{1,k}, t_k) - v_1(\hat{s}_{2,k}, t_k)|^2 ds dt}_{=\text{III}.k.1} \\ &\quad + \underbrace{\int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} |v_1(\hat{s}_{2,k}, t_k) - v_2(\hat{s}_{2,k}, t_k)|^2 ds dt}_{=\text{III}.k.2} \tag{A.30} \end{aligned}$$

For the first term in (A.30), we consider the spatial plane curve resulting from projecting  $s(\tau) \in \Gamma(\tau)$ , for all  $\tau$  between  $t_k$  and  $t$ , onto the spatial plane at time  $t_k$ . By applying the fundamental theorem of calculus for line integrals to this curve, we have

that

$$\begin{aligned}
 \text{III}.k.1 &= \int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} |v_1(s(t), t_k) - v_1(s(t_k), t_k)|^2 \, ds \, dt \\
 &= \int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} \left| \int_{t_k}^t \mu(\tau) \cdot \nabla v_1(s(\tau), t_k) \, d\tau \right|^2 \, ds \, dt \\
 &\leq |\mu|_{I_n}^2 k_n \int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} \int_{I_n} |\nabla v_1(s(\tau), t_k)|^2 \, d\tau \, ds \, dt \\
 &= |\mu|_{I_n}^2 k_n^2 |\bar{\mu}| \int_{\Gamma(t_k)} \int_{I_n} |\nabla v_1(s(\tau), t_k)|^2 \, d\tau \, ds \tag{A.31} \\
 &\lesssim |\mu|_{I_n}^2 k_n^2 |\bar{\mu}| \int_{\Omega_1(t_k) \cup \Omega_O(t_k)} |\nabla v_1(x, t_k)|^2 \, dx \\
 &\lesssim |\mu|_{I_n}^2 k_n \int_{I_n} \|\nabla v_1\|_{\Omega_1(t) \cup \Omega_O(t)}^2 \, dt \\
 &\lesssim |\mu|_{I_n}^2 k_n \int_{I_n} \|v\|_{A_{h,t}}^2 \, dt
 \end{aligned}$$

where, in the fifth step, we have taken possible multiples of the same line integrals into account and expanded the domain of integration. In the sixth step, we have used a standard inverse inequality for polynomials. For the second term in (A.30), we use Lemma A.7, thus

$$\begin{aligned}
 \text{III}.k.2 &= k_n |\bar{\mu}| \int_{\Gamma(t_k)} |v_1(s, t_k) - v_2(s, t_k)|^2 \, ds = \int_{\Gamma(t_k)} \left( k_n |\bar{\mu}| [v]^2(s, t_k) \right) \, ds \\
 &\lesssim \int_{\Gamma(t_k)} \left( \int_{I_n} |\bar{\mu}| [v]^2(\hat{s}_2, t) \, dt + k_n |\bar{\mu}| \int_{I_n} |D_{t,2}[v]^2(\hat{s}_2, t)| \, dt \right) \, ds \\
 &\lesssim \int_{\bar{\Gamma}_n} |[v]|^2 \, d\bar{s} + k_n \int_{\bar{\Gamma}_n} |[v]| |D_{t,2}[v]| \, d\bar{s} \tag{A.32} \\
 &\leq \|[v]\|_{\bar{\Gamma}_n}^2 + k_n \|[v]\|_{\bar{\Gamma}_n} \|D_{t,2}[v]\|_{\bar{\Gamma}_n} \\
 &\leq \underbrace{\left(1 + \frac{1}{\varepsilon}\right) \|[v]\|_{\bar{\Gamma}_n}^2}_{=\text{III}.k.2.1} + \underbrace{\varepsilon k_n^2 \|D_{t,2}[v]\|_{\bar{\Gamma}_n}^2}_{=\text{III}.k.2.2}
 \end{aligned}$$

Using (4.2), the first term is

$$\begin{aligned}
 \text{III}.k.2.1 &= \left(1 + \frac{1}{\varepsilon}\right) \int_{I_n} |\bar{\mu}| \|[v]\|_{\Gamma(t)}^2 \, dt \lesssim h \int_{I_n} |\bar{\mu}| \|[v]\|_{1/2,h,\Gamma(t)}^2 \, dt \tag{A.33} \\
 &\leq h \int_{I_n} \|v\|_{A_{h,t}}^2 \, dt
 \end{aligned}$$

Using again (4.2), the second term is

$$\begin{aligned}
 \text{III.k.2.2} &= \varepsilon k_n^2 \|D_{t,2}v_1 - D_{t,2}v_2\|_{\Gamma_n}^2 = \varepsilon k_n^2 \|D_t v_1 + \mu \cdot \nabla v_1 - D_t v_2\|_{\Gamma_n}^2 \\
 &\lesssim \varepsilon k_n^2 \| [D_t v] \|_{\Gamma_n}^2 + \varepsilon k_n^2 \| \mu \cdot \nabla v_1 \|_{\Gamma_n}^2 \\
 &\leq \varepsilon k_n^2 \int_{I_n} |\bar{\mu}| \| [D_t v] \|_{\Gamma(t)}^2 dt + |\mu|_{I_n}^2 k_n^2 \| \nabla v_1 \|_{\Gamma_n}^2 \\
 &\leq \varepsilon h \underbrace{\int_{I_n} k_n^2 |\bar{\mu}| \| [D_t v] \|_{1/2,h,\Gamma(t)}^2 dt}_{=\text{III}} + \underbrace{|\mu|_{I_n}^2 k_n^2 \| \nabla v_1 \|_{\Gamma_n}^2}_{=\text{III.k.2.2.2}}
 \end{aligned}
 \tag{A.34}$$

where the first term is done. For the second term, we use Corollary A.2, thus

$$\begin{aligned}
 \text{III.k.2.2.2} &= |\mu|_{I_n}^2 k_n^2 \int_{I_n} |\bar{\mu}| \sum_{K \in \mathcal{T}_{0,\Gamma(t)}} \| \nabla v_1 \|_{\Gamma_K}^2 dt \\
 &\lesssim \frac{|\mu|_{I_n}^2 k_n^2}{h_{\min}} \int_{I_n} \sum_{K \in \mathcal{T}_{0,\Gamma(t)}} \| \nabla v_1 \|_K^2 dt \\
 &\leq \frac{|\mu|_{I_n}^2 k_n^2}{h_{\min}} \int_{I_n} \| \nabla v_1 \|_{\Omega_1(t) \cup \Omega_0(t)}^2 dt \lesssim \frac{|\mu|_{I_n}^2 k_n^2}{h_{\min}} \int_{I_n} \| \|v\| \|_{A_{h,t}}^2 dt
 \end{aligned}
 \tag{A.35}$$

This concludes the separate treatment of all the terms unfolding in the estimation of the third term in (A.25). Collecting all the estimates and using (3.1) gives us

$$\begin{aligned}
 \text{III} &\lesssim h_{\min}^{-1} \sum_{t_k \in S_n^g} \left( |\mu|_{I_n}^2 k_n \int_{I_n} \| \|v\| \|_{A_{h,t}}^2 dt + h \int_{I_n} \| \|v\| \|_{A_{h,t}}^2 dt + \text{III.k.2.2} \right) \\
 &\lesssim h_{\min}^{-1} \left( \left( |\mu|_{I_n}^2 k_n + h + \frac{|\mu|_{I_n}^2 k_n^2}{h_{\min}} \right) \int_{I_n} \| \|v\| \|_{A_{h,t}}^2 dt + \varepsilon h (\text{III}) \right) \\
 &\lesssim \int_{I_n} \| \|v\| \|_{A_{h,t}}^2 dt + \varepsilon (\text{III})
 \end{aligned}
 \tag{A.36}$$

By kicking back the  $\varepsilon$ -term and taking  $\varepsilon$  sufficiently small, we may estimate the third term in (A.25) by the first term on the right-hand side of A.36. The fourth term in (A.25) receives the same treatment as the first, thus

$$\text{IV} \lesssim \int_{I_n} k_n^2 \left( \| D_t \nabla v \|_{\Omega_1(t) \cup \Omega_0(t)}^2 + \| D_t \nabla v \|_{\Omega_2(t)}^2 \right) dt \lesssim \int_{I_n} \| \|v\| \|_{A_{h,t}}^2 dt
 \tag{A.37}$$

The treatment of all the terms in (A.25) is done. This shows (A.24). □

**Lemma A.9** (An inverse inequality for  $\mathcal{P}((a, c), (c, b))$ ) *For an open interval  $(a, b)$ , a point  $c \in (a, b)$ , and for  $w \in \mathcal{P}((a, c), (c, b))$ , i.e.,  $w$  is a polynomial on  $(a, c)$ ,*

possibly another polynomial on  $(c, b)$ , and possibly discontinuous at  $c$ , there exists a positive constant depending on the polynomial degree such that

$$(b - a)|w(a^+)|^2 \lesssim \int_a^c |w(x)|^2 dx + \int_c^b |w(x)|^2 dx + (b - c)|[w](c)|^2 + (b - c)(c - a) \int_a^c |w'(x)|^2 dx \tag{A.38}$$

**Proof** Using a standard inverse inequality for polynomials with a positive constant that depends on the polynomial degree, the left-hand side of (A.38) is

$$(b - a)|w(a^+)|^2 \lesssim \int_a^c |w(x)|^2 dx + (b - c)|w(a^+)|^2 \tag{A.39}$$

Adding and subtracting  $w(c^-)$  and  $w(c^+)$  within the absolute value, followed by using standard estimates, the second term is

$$(b - c)|w(a^+)|^2 = (b - c) \left| - \int_a^c w'(x) dx - [w](c) + w(c^+) \right|^2 \lesssim (b - c)(c - a) \int_a^c |w'(x)|^2 dx + (b - c)|[w](c)|^2 + \int_c^b |w(x)|^2 dx \tag{A.40}$$

□

**Lemma A.10** (Discrete temporal inverse inequalities for  $V_h^n$ ) *Let  $k_n$  be the length of interval  $I_n$ , the scaled differential operator  $D_t$  be defined by (3.7),  $|\mu|_{I_n} = \max_{t \in I_n} \{|\mu(t)|\}$ , and  $\|\cdot\|_{A_{h,t}}$  be defined by (4.3). Then, for  $q = 0, 1$ , we have that*

$$\sum_{i=1}^2 k_n^2 \|(D_t v)_{n-1}^+\|_{\Omega_{i,n-1}}^2 \lesssim \sum_{i=1}^2 \int_{I_n} k_n \|D_t v\|_{\Omega_i(t)}^2 dt + |\mu|_{I_n} \int_{I_n} \|v\|_{A_{h,t}}^2 dt \quad \forall v \in V_h^n \tag{A.41}$$

$$\sum_{i=1}^2 k_n^2 \|(D_t v)_n^-\|_{\Omega_{i,n}}^2 \lesssim \sum_{i=1}^2 \int_{I_n} k_n \|D_t v\|_{\Omega_i(t)}^2 dt + |\mu|_{I_n} \int_{I_n} \|v\|_{A_{h,t}}^2 dt \quad \forall v \in V_h^n \tag{A.42}$$

**Proof** We only prove (A.41), since the proof of (A.42) is analogous. Recall  $\hat{x}(t)$  defined by (A.17). We denote by  $\hat{x}_i^n$  the slabwise space-time trajectory through a point  $x_{i,n-1} \in \Omega_{i,n-1}$ . We define the set of points in  $\Omega_{1,n-1}$  with cut and uncut space-time

trajectories by

$$\Omega_{1,n-1}^\Gamma := \{x \in \Omega_{1,n-1} : \tilde{x}_1^n \cap \bar{\Gamma}_n \neq \emptyset\} \tag{A.43}$$

$$\Omega_{1,n-1}^n := \{x \in \Omega_{1,n-1} : \tilde{x}_1^n \cap \bar{\Gamma}_n = \emptyset\} \tag{A.44}$$

The idea to prove (A.41) is that if a point’s space-time trajectory is *uncut*, we use a standard inverse inequality, and if it is *cut*, we use Lemma A.9. Using that  $\Omega_{1,n-1}^\Gamma$  and  $\Omega_{1,n-1}^n$  form a partition of  $\Omega_{1,n-1}$ , the left-hand side of (A.41) is

$$\begin{aligned} \sum_{i=1}^2 k_n^2 \|(D_t v)_{n-1}^+\|_{\Omega_{i,n-1}}^2 &= \underbrace{k_n^2 \|(D_t v)_{n-1}^+\|_{\Omega_{1,n-1}^n}^2}_{=I} + \underbrace{k_n^2 \|(D_t v)_{n-1}^+\|_{\Omega_{2,n-1}}^2}_{=II} \\ &\quad + \underbrace{k_n^2 \|(D_t v)_{n-1}^+\|_{\Omega_{1,n-1}^\Gamma}^2}_{=III} \end{aligned} \tag{A.45}$$

We treat the terms separately. See Fig. 11 for an illustration of the proof idea. Using a standard inverse inequality, the first and second term in (A.45) are

$$I \lesssim \int_{I_n} k_n \|D_t v\|_{\Omega_1(t)}^2 dt \quad II \lesssim \int_{I_n} k_n \|D_t v\|_{\Omega_2(t)}^2 dt \tag{A.46}$$

For the third term in (A.45), we recall  $\hat{s}$  defined by (A.18). We consider a space-time curve that starts at  $x \in \Omega_{1,n-1}^\Gamma$ , goes straight up in time until it hits  $\bar{\Gamma}_n$ , which occurs at time  $t_\Gamma$ , then travels on  $\bar{\Gamma}_n$  along  $(\hat{s}(t), t)$  up to  $t_n$ . We will apply Lemma A.9 to the function that is  $(D_t v)_1$  up until  $t_\Gamma$  along this space-time curve, and  $(D_t v)_2$  afterwards. Here the corresponding derivative term on the right-hand side of (A.38) vanishes since  $D_t^2 v(x, t) = 0$  for  $q \leq 1$ . Thus

$$\begin{aligned} III &= k_n \int_{\Omega_{1,n-1}^\Gamma} \left( k_n |D_t v(x, t_{n-1}^+)|^2 dx \right) \\ &\lesssim k_n \int_{\Omega_{1,n-1}^\Gamma} \left( \int_{t_{n-1}}^{t_\Gamma} |D_t v(x, t)|^2 dt + \int_{t_\Gamma}^{t_n} |\bar{\mu}| |(D_t v)_2(\hat{s}(t), t)|^2 dt \right. \\ &\quad \left. + (t_n - t_\Gamma) |[D_t v](x, t_\Gamma)|^2 + (t_n - t_\Gamma)(t_\Gamma - t_{n-1}) \int_{t_{n-1}}^{t_\Gamma} |D_t^2 v(x, t)|^2 dt \right) dx \\ &\lesssim \underbrace{k_n \int_{\Omega_{1,n-1}^\Gamma} \int_{t_{n-1}}^{t_\Gamma} |D_t v(x, t)|^2 dt dx}_{=III.1} + \underbrace{k_n \int_{\Omega_{1,n-1}^\Gamma} \int_{t_\Gamma}^{t_n} |\bar{\mu}| |(D_t v)_2(\hat{s}(t), t)|^2 dt dx}_{=III.2} \\ &\quad + \underbrace{k_n \int_{\Omega_{1,n-1}^\Gamma} (t_n - t_\Gamma) |[D_t v](x, t_\Gamma)|^2 dx}_{=III.3} \end{aligned} \tag{A.47}$$

Simply expanding the domain of integration, the first term is

$$\text{III.1} \lesssim \int_{I_n} k_n \|D_t v\|_{\Omega_1(t)}^2 dt \tag{A.48}$$

For the second and third term in (A.47), we want to change the domain of integration from  $\Omega_{1,n-1}^\Gamma$  to its temporal projection onto  $\bar{\Gamma}_n$ . To do this, we note that  $dx \lesssim |\mu|_{I_n} d\bar{s}$ , where  $dx$  and  $d\bar{s}$  are the integration differentials for  $\Omega_{1,n-1}^\Gamma$  and  $\bar{\Gamma}_n$ , respectively. Using this, a standard trace inequality for  $H^1(\Omega_2)$ , that  $\nabla D_t v = D_t \nabla v$ , and Lemma A.6, the second term is

$$\begin{aligned} \text{III.2} &\lesssim k_n \int_{\Omega_{1,n-1}^\Gamma} \int_{I_n} |\bar{\mu}| |(D_t v)_2(\hat{s}(t), t)|^2 dt dx \\ &\lesssim k_n \int_{\bar{\Gamma}_n} \int_{I_n} |\bar{\mu}| |(D_t v)_2(\bar{s})|^2 dt |\mu|_{I_n} d\bar{s} \lesssim |\mu|_{I_n} k_n^2 \int_{I_n} |\bar{\mu}| \|(D_t v)_2\|_{\bar{\Gamma}(t)}^2 dt \\ &\lesssim |\mu|_{I_n} k_n \int_{I_n} k_n \|D_t v\|_{\Omega_2(t)}^2 dt + |\mu|_{I_n} \int_{I_n} k_n^2 \|D_t \nabla v\|_{\Omega_2(t)}^2 dt \\ &\lesssim \int_{I_n} k_n \|D_t v\|_{\Omega_2(t)}^2 dt + |\mu|_{I_n} \int_{I_n} \|v\|_{A_{h,t}}^2 dt \end{aligned} \tag{A.49}$$

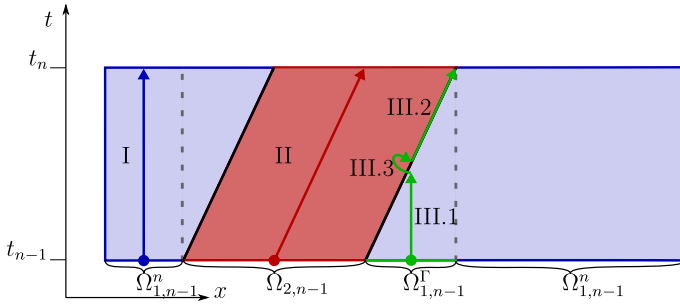
Using the relation between the integration differentials, the estimate (4.2), and Lemma A.8, the third term in (A.47) is

$$\begin{aligned} \text{III.3} &\lesssim k_n^2 \int_{\Omega_{1,n-1}^\Gamma} |[D_t v](x, t_\Gamma)|^2 dx \lesssim k_n^2 \int_{\bar{\Gamma}_n} |[D_t v](\bar{s})|^2 |\mu|_{I_n} d\bar{s} \\ &\lesssim |\mu|_{I_n} k_n^2 \int_{I_n} |\bar{\mu}| \|[D_t v]\|_{\bar{\Gamma}(t)}^2 dt \lesssim |\mu|_{I_n} h \int_{I_n} \|k_n D_t v\|_{A_{h,t}}^2 dt \\ &\lesssim |\mu|_{I_n} \int_{I_n} \|v\|_{A_{h,t}}^2 dt \end{aligned} \tag{A.50}$$

The treatment of all the terms in (A.45) is done. This shows (A.41). □

**Lemma A.11** (A discrete temporal inverse estimate in  $\|\cdot\|_X$ ) *Let the norm  $\|\cdot\|_X$  be defined by (4.10),  $k_n$  be the length of time interval  $I_n$ , and the scaled differential operator  $D_t$  be defined by (3.7). Then, for  $q = 0, 1$ , we have that*

$$\|k_n D_t v\|_X \lesssim \|v\|_X \quad \forall v \in V_h \tag{A.51}$$



**Fig. 11** The starting domains  $\Omega_{1,n-1}^n$ ,  $\Omega_{2,n-1}$ , and  $\Omega_{1,n-1}^\Gamma$ . The arrows represent the treatment of the corresponding right-hand side terms

**Proof** The square of the left-hand side of (A.51) is

$$\begin{aligned}
 \| \|k_n D_t v\| \|_X^2 &= \sum_{i=1}^2 \sum_{n=1}^N \int_{I_n} k_n \|D_t(k_n D_t v)\|_{\Omega_i(t)}^2 dt \\
 &+ \sum_{n=1}^N \left( \int_{I_n} \| \|k_n D_t v\| \|_{A_{h,t}}^2 dt + \| |\bar{n}^t|^{1/2} [k_n D_t v] \|_{\bar{\Gamma}_n}^2 \right) \\
 &+ \sum_{i=1}^2 \sum_{n=1}^{N-1} \| [k_n D_t v]_n \|_{\Omega_{i,n}}^2 \\
 &+ \sum_{i=1}^2 \left( \| (k_n D_t v)_N^- \|_{\Omega_{i,N}}^2 + \| (k_n D_t v)_0^+ \|_{\Omega_{i,0}}^2 \right)
 \end{aligned}
 \tag{A.52}$$

The first term vanishes since  $D_t^2 v(x, t) = 0$  for  $q \leq 1$ . The  $\bar{\Gamma}_n$ -norm term is estimated by the  $A_{h,t}$ -norm term by using (4.2). Applying Lemma A.8 to the  $A_{h,t}$ -norm term and Lemma A.10 to all the terms in the last two rows, we get terms which may be estimated by  $\| \|v\| \|_X^2$ . □

### B Interpolation

Let  $\cdot^\circ$  denote the interior of a set, e.g.,  $I_n^\circ = (t_{n-1}, t_n)$ . Also let  $C_b(\cup_n I_n^\circ)$  denote the space of functions that are continuous and bounded on every  $I_n^\circ$ . In this section, the space-time interpolation operator  $\bar{I}_h : C_b(\cup_n I_n^\circ; L^1(\Omega_0)) \rightarrow V_h$  is successively constructed by first defining spatial interpolation operators, then temporal ones, and finally combining them. Interpolation error estimates are also presented.



### B.1 Slabwise operators and local estimates

**Definition B.1** (Spatial interpolation operators) We define the spatial interpolation operators  $\pi_{h,0} : L^1(\Omega_0) \rightarrow V_{h,0}$  and  $\pi_{h,G} : L^1(G) \rightarrow V_{h,G}$  to be the Scott-Zhang interpolation operators for the spaces  $V_{h,0}$  and  $V_{h,G}$ , respectively, where the defining integrals are taken over entire simplices.

Note that  $\pi_{h,G}$  is time-dependent but to lighten the notation we omit this. The temporal interpolation operators will interpolate along the space-time trajectories of the domains  $\Omega_0$  and  $G$ . For  $n = 1, \dots, N$ , we define the slabwise space-time trajectory for a point  $x \in \Omega_0$  and that of a point  $x_n \in G(t_n)$  by

$$\tilde{x}_0^n := \{(\hat{x}(t), t) : \hat{x}(t) = x, t \in I_n\} \tag{B.1}$$

$$\tilde{x}_G^n := \{(\hat{x}(t), t) : \hat{x}(t) = x_n - \int_t^{t_n} \mu(\tau) \, d\tau, t \in I_n\} \tag{B.2}$$

Note that (A.17) can be used to obtain all trajectories defined by (B.2) but *not* all defined by (B.1) because some  $\tilde{x}_0^n$  may lie completely in  $S_{2,n}$ . Let  $S_n^q$  denote the set of temporal interpolation points for interpolation to  $\mathcal{P}^q(I_n)$ . We take  $S_n^0 = \{t_n^-\}$  and  $S_n^1 = \{t_n^-, t_{n-1}^+\}$ . For  $q > 1$ , we include interior points of  $I_n$  in some suitable fashion.

**Definition B.2** (Temporal interpolation operators) For each time subinterval  $I_n$ , where  $n = 1, \dots, N$ , we define the temporal interpolation operators  $\pi_0^n : C_b(\tilde{x}_0^n) \rightarrow \mathcal{P}^q(\tilde{x}_0^n)$  and  $\pi_G^n : C_b(\tilde{x}_G^n) \rightarrow \mathcal{P}^q(\tilde{x}_G^n)$  to be the nodal interpolation operators that use the points in  $S_n^q$  as nodal interpolation points.

Note that  $\pi_0^n$  and  $\pi_G^n$  are spatially dependent but to lighten the notation we omit this. We combine the spatial and temporal interpolation operators to define space-time ones.

**Definition B.3** (Slabwise space-time interpolation operators) For  $n = 1, \dots, N$ , we define the slabwise space-time interpolation operators  $\bar{I}_{h,0}^n : C_b(I_n^\circ; L^1(\Omega_0)) \rightarrow V_{h,0}^n$  and  $\bar{I}_{h,G}^n : C_b(I_n^\circ; L^1(G)) \rightarrow V_{h,G}^n$  by

$$\bar{I}_{h,0}^n := \pi_0^n \pi_{h,0} \quad \bar{I}_{h,G}^n := \pi_G^n \pi_{h,G} \tag{B.3}$$

Recall the interdependent indices  $i \in \{1, 2\}$  and  $j \in \{0, G\}$  where  $j = 0$  for  $i = 1$  and  $j = G$  for  $i = 2$ . Let  $\bar{K}_n := \{(x, t) : x \in K = K(t), t \in I_n\}$  denote an arbitrary space-time prism, where  $K = K_j \in \mathcal{T}_j$ . Let  $\mathcal{N}(K)$  denote the neighborhood of a simplex  $K$ , i.e., the set of all adjacent simplices to and including  $K$ . We also use the notation  $\|w\|_{K, I_n} = \max_{t \in I_n} \{\|w(\cdot, t)\|_{K(t)}\}$ .

**Lemma B.1** (Local space-time interpolation error estimates for  $\bar{K}_n$ ) *Let  $\bar{I}_{h,j}^n$  be defined by (B.3), where  $j \in \{0, G\}$ , and let  $D_t$  be defined by (3.7). Then, for a function  $v$  with sufficient spatial and temporal regularity, we have for  $0 \leq s \leq q+1$  and  $0 \leq r \leq p+1$  that*

$$\|D_t^s(v - \bar{I}_{h,j}^n v)\|_{\bar{K}_n} \lesssim k_n^{q+1-s} \|D_t^{q+1} v\|_{\bar{K}_n} + h^{p+1} k_n^{1/2} \|D_x^{p+1} D_t^s v\|_{\mathcal{N}(K), I_n} \tag{B.4}$$

$$\|D_x^r(v - \bar{I}_{h,j}^n v)\|_{\bar{K}_n} \lesssim k_n^{q+1} \|D_t^{q+1} D_x^r v\|_{\bar{K}_n} + h^{p+1-r} k_n^{1/2} \|D_x^{p+1} v\|_{\mathcal{N}(K), I_n} \tag{B.5}$$

**Proof** We show the two estimates separately, starting with the first. Using that  $\bar{I}_{h,j}^n = \pi_j^n \pi_{h,j} = \pi_{h,j} \pi_j^n$ , that  $D_t^s \pi_{h,j} = \pi_{h,j} D_t^s$ , stability of  $\pi_{h,j}$ , and a trivial estimate, the left-hand side of (B.4) is

$$\begin{aligned} \|D_t^s(v - \bar{I}_{h,j}^n v)\|_{\bar{K}_n} &\leq \|D_t^s \pi_{h,j} (\mathbb{1} - \pi_j^n) v\|_{\bar{K}_n} + \|D_t^s(v - \pi_{h,j} v)\|_{\bar{K}_n} \\ &\lesssim \|D_t^s (\mathbb{1} - \pi_j^n) v\|_{\bar{K}_n} + k_n^{1/2} \|D_t^s v - \pi_{h,j} D_t^s v\|_{K, I_n} \end{aligned} \tag{B.6}$$

Applying standard estimates for  $\pi_j^n$  and  $\pi_{h,j}$  shows the first estimate and we move on to the second. We are going to use the expansion of interpolants of  $\pi_j^n$  into a sum over the temporal interpolation points  $t_k \in S_n^q$  with  $\lambda_k \in \mathcal{P}^q(I_n)$  denoting the corresponding shape function. For a function  $w$  of sufficient regularity, we have that

$$\begin{aligned} \|\pi_j^n w\|_{\bar{K}_n}^2 &= \int_{I_n} \int_{K(t)} \left| \sum_{t_k \in S_n^q} w(\hat{x}(t_k), t_k) \lambda_k(t) \right|^2 dx dt \\ &\leq (q+1) \sum_{t_k \in S_n^q} \int_{I_n} \int_{K(t)} |w(\hat{x}(t_k), t_k)|^2 dx dt \lesssim k_n \|w\|_{K, I_n}^2 \end{aligned} \tag{B.7}$$

Using this after using that  $D_x^r \pi_j^n = \pi_j^n D_x^r$ , the left-hand side of (B.5) is

$$\begin{aligned} \|D_x^r(v - \bar{I}_{h,j}^n v)\|_{\bar{K}_n} &\leq \|D_x^r(v - \pi_j^n v)\|_{\bar{K}_n} + \|D_x^r \pi_j^n (\mathbb{1} - \pi_{h,j}) v\|_{\bar{K}_n} \\ &\lesssim \|D_x^r v - \pi_j^n D_x^r v\|_{\bar{K}_n} + k_n^{1/2} \|D_x^r (\mathbb{1} - \pi_{h,j}) v\|_{K, I_n} \end{aligned} \tag{B.8}$$

Applying standard estimates for  $\pi_j^n$  and  $\pi_{h,j}$  shows the second estimate. □

Recall  $\hat{s}$  defined by (A.18). Let  $\mathcal{T}_{j,D} = \{K \in \mathcal{T}_j : K \cap D \neq \emptyset\}$ , where  $D$  is a possibly time-dependent subset of  $\mathbb{R}^{d+1}$ .

**Lemma B.2** (Slabwise space-time interpolation error estimates for  $\bar{\Gamma}_n$ ) *Let  $\bar{I}_{h,j}^n$  be defined by (B.3), where  $j \in \{0, G\}$ , and let  $D_t$  be defined by (3.7). Then, for any function  $v$  with sufficient spatial and temporal regularity, we have that*

$$\|(v - \bar{I}_{h,j}^n v)_i\|_{\bar{\Gamma}_n}^2 \lesssim k_n^{2q+2} \|D_t^{q+1} v\|_{L^2(\bar{\Gamma}_n, L^\infty(I_n))}^2 + h^{2p+1} \sum_{K \in \mathcal{T}_{j, \bar{\Gamma}_n}} k_n \|D_x^{p+1} v\|_{\mathcal{N}(K), I_n}^2 \tag{B.9}$$

**Proof** The general proof idea is the same for all  $q \geq 0$ . What varies is how a temporal difference is treated. We show how to treat if for  $q = 1$  from which it should be relatively straightforward how to handle the other cases. Using the shape functions

$\lambda_k \in \mathcal{P}^q(I_n)$ , corresponding to interpolation points  $t_k \in S_n^q$ , the argument of the norm on the left-hand side of (B.9) is

$$\begin{aligned} (v - \bar{I}_{h,j}^n v)_i|_{\bar{\Gamma}_n} &= v(s, t) - \sum_{t_k \in S_n^q} \pi_{h,j} v(\hat{s}_k, t_k) \lambda_k(t) \\ &= \underbrace{\sum_{t_k \in S_n^q} (v(s, t) - v(\hat{s}_k, t_k)) \lambda_k(t)}_{=A} + \underbrace{\sum_{t_k \in S_n^q} (v(\hat{s}_k, t_k) - \pi_{h,j} v(\hat{s}_k, t_k)) \lambda_k(t)}_{=B} \end{aligned} \tag{B.10}$$

The left-hand side of (B.9) may thus be split by  $\|(v - \bar{I}_{h,j}^n v)_i\|_{\bar{\Gamma}_n}^2 \lesssim \|A\|_{\bar{\Gamma}_n}^2 + \|B\|_{\bar{\Gamma}_n}^2$  where we consider the terms separately, starting with the first. We proceed with some further treatment of  $A$  for which we restrict ourselves to the case  $q = 1$ . From this case it should however be relatively straightforward how to treat  $A$  for  $q \neq 1$ . Using the mean value theorem along the space-time trajectories, the explicit expressions for the shape functions  $\lambda_{n-1}$  and  $\lambda_n$  for  $q = 1$ , and the fundamental theorem of calculus, we have

$$A = \sum_{t_k \in S_n^q} D_t v(\hat{s}, c_k)(t - t_k) \lambda_k(t) = \frac{(t - t_n)(t - t_{n-1})}{k_n} \int_{c_{n-1}}^{c_n} D_t^2 v(\hat{s}, \tau) \, d\tau \tag{B.11}$$

Using (B.11), we have that

$$\begin{aligned} \|A\|_{\bar{\Gamma}_n}^2 &\leq \int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} k_n^2 (c_n - c_{n-1}) \int_{c_{n-1}}^{c_n} |D_t^2 v(\hat{s}, \tau)|^2 \, d\tau \, ds \, dt \\ &\leq k_n^4 \|D_t^2 v\|_{L^2(\bar{\Gamma}_n, L^\infty(I_n))}^2 \end{aligned} \tag{B.12}$$

Writing  $B_k = v(\hat{s}_k, t_k) - \pi_{h,j} v(\hat{s}_k, t_k)$ , using (A.11), and standard estimates for  $\pi_{h,j}$ , we have that

$$\begin{aligned} \|B\|_{\bar{\Gamma}_n}^2 &\lesssim \int_{I_n} |\bar{\mu}| \int_{\Gamma(t)} \sum_{t_k \in S_n^q} |v(\hat{s}_k, t_k) - \pi_{h,j} v(\hat{s}_k, t_k)|^2 \, ds \, dt \\ &\lesssim \sum_{t_k \in S_n^q} \sum_{K \in \mathcal{T}_{j, \bar{\Gamma}_n}} k_n \|B_k\|_{\bar{\Gamma}_K}^2 \lesssim \sum_{t_k \in S_n^q} \sum_{K \in \mathcal{T}_{j, \bar{\Gamma}_n}} k_n \left( h_K^{-1} \|B_k\|_K^2 + h_K \|D_x B_k\|_K^2 \right) \\ &\lesssim h^{2p+1} \sum_{K \in \mathcal{T}_{j, \bar{\Gamma}_n}} k_n \|D_x^{p+1} v\|_{\mathcal{N}(K), I_n}^2 \end{aligned} \tag{B.13}$$

□

**Lemma B.3** (Local spatial interpolation error estimates for temporal endpoints) *Let  $\bar{I}_{h,j}^n$  be defined by (B.3), where  $j \in \{0, G\}$ , let  $t_k^\tau \in \{t_{n-1}^+, t_n^-\}$ , and let  $D_t$  be defined by (3.7). Then, for any function  $v$  with sufficient spatial and temporal regularity, we have for  $q > 0$  that*

$$\|(v - \bar{I}_{h,j}^n v)_k^\tau\|_K \lesssim h^{p+1} \|D_x^{p+1} v(\cdot, t_k^\tau)\|_{\mathcal{N}(K)} \tag{B.14}$$

and for  $q = 0$  that

$$\|(v - \bar{I}_{h,j}^n v)_n^-\|_K \lesssim h^{p+1} \|D_x^{p+1} v(\cdot, t_n^-)\|_{\mathcal{N}(K)} \tag{B.15}$$

$$\|(v - \bar{I}_{h,j}^n v)_{n-1}^+\|_K \lesssim k_n^{1/2} \|D_t v\|_{\bar{K}_n} + h^{p+1} \|D_x^{p+1} v(\cdot, t_n^-)\|_{\mathcal{N}(K)} \tag{B.16}$$

**Proof** Estimates (B.14) and (B.15) follow from simply using that  $t_k^\tau$  is an interpolation point of  $\pi_j^n$  and then a standard estimate for  $\pi_{h,j}$ . This does not work for (B.16), since  $t_{n-1}^+$  is not an interpolation point for  $q = 0$ . Instead, we integrate along the slabwise space-time trajectory of an element  $x \in K$  to obtain

$$\begin{aligned} \|v(\cdot, t_{n-1}^+) - v(\cdot, t_n^-)\|_K^2 &= \int_K \left| v(x(t_{n-1}^+), t_{n-1}^+) - v(x(t_n^-), t_n^-) \right|^2 dx \\ &= \int_K \left| \int_{I_n} D_t v(x(t), t) dt \right|^2 dx \leq k_n \|D_t v\|_{\bar{K}_n}^2 \end{aligned} \tag{B.17}$$

Using the definition of  $\pi_j^n$  for  $q = 0$ , the left-hand side of (B.16) is

$$\begin{aligned} \|(v - \bar{I}_{h,j}^n v)_{n-1}^+\|_K &\leq \|(v - \pi_j^n v)(\cdot, t_{n-1}^+)\|_K + \|(\pi_j^n v - \pi_j^n \pi_{h,j} v)(\cdot, t_{n-1}^+)\|_K \\ &= \|v(\cdot, t_{n-1}^+) - v(\cdot, t_n^-)\|_K + \|(v - \pi_{h,j} v)(\cdot, t_n^-)\|_K \end{aligned} \tag{B.18}$$

Applying (B.17) and a standard estimate for  $\pi_{h,j}$  shows (B.16). □

### B.2 Global operator and estimates

**Definition B.4** (Main space-time interpolation operator) We define the main space-time interpolation operator  $\bar{I}_h : C_b(\cup_n I_n^o; L^1(\Omega_0)) \rightarrow V_h$  by, for  $n = 1, \dots, N$ ,

$$\bar{I}_h v|_{S_{1,n}} := \bar{I}_{h,0}^n v|_{S_{1,n}} \quad \bar{I}_h v|_{S_{2,n}} := \bar{I}_{h,G}^n v|_{S_{2,n}} \tag{B.19}$$

**Lemma B.4** (Global space-time interpolation error estimates for  $\Omega_0 \times (0, T]$ ) *Let  $\bar{I}_h$  be defined by (B.19) and  $D_t$  by (3.7). Then, for any function  $v$  with sufficient spatial and temporal regularity, we have for  $0 \leq s \leq q + 1$  and  $0 \leq r \leq p + 1$  that*

$$\sum_{i=1}^2 \sum_{n=1}^N \int_{I_n} \|D_t^s (v - \bar{I}_h v)\|_{\Omega_i(t)}^2 dt \lesssim k^{2(q+1-s)} E_{k,0}^2(v) + h^{2(p+1)} E_{h,s}^2(v) \tag{B.20}$$

$$\sum_{i=1}^2 \sum_{n=1}^N \int_{I_n} \|D_x^r(v - \bar{I}_h v)\|_{\Omega_i(t)}^2 dt \lesssim k^{2(q+1)} E_{k,r}^2(v) + h^{2(p+1-r)} E_{h,0}^2(v) \quad (B.21)$$

where

$$E_{k,r}^2(v) = \sum_{i=1}^2 \sum_{n=1}^N \sum_{K \in \mathcal{T}_{j,S_{i,n}}} \|D_t^{q+1} D_x^r v\|_{\bar{K}_n}^2 \quad (B.22)$$

$$E_{h,s}^2(v) = \sum_{i=1}^2 \sum_{n=1}^N \sum_{K \in \mathcal{T}_{j,S_{i,n}}} k_n \|D_x^{p+1} D_t^s v\|_{\mathcal{N}(K), I_n}^2 \quad (B.23)$$

**Proof** Both estimates follow by applying Lemma B.1. □

**Lemma B.5** (An interpolation error estimate in  $\|\cdot\|_{B_h}$ ) *Let  $\|\cdot\|_{B_h}$ ,  $\bar{I}_h$ , and  $D_t$  be defined by (4.8), (B.19), and (3.7), respectively. Then, for any function  $v$  with sufficient spatial and temporal regularity, we have that*

$$\|v - \bar{I}_h v\|_{B_h}^2 \lesssim k^{2q+1} F_k^2(v) + h^{2p} F_h^2(v) \quad (B.24)$$

where

$$F_k^2(v) = \sum_{i=1}^2 \sum_{n=1}^N \sum_{K \in \mathcal{T}_{j,S_{i,n}}} \left( \|D_t^{q+1} v\|_{\bar{K}_n}^2 + \|D_t^{q+1} \nabla v\|_{\bar{K}_n}^2 + \|D_t^{q+1} D_x^2 v\|_{\bar{K}_n}^2 \right) + \sum_{i=1}^2 \sum_{n=1}^N \|D_t^{q+1} v\|_{L^2(\bar{\Gamma}_n, L^\infty(I_n))}^2 \quad (B.25)$$

$$F_h^2(v) = \sum_{i=1}^2 \sum_{n=1}^N \sum_{K \in \mathcal{T}_{j,S_{i,n}}} k_n \|D_x^{p+1} v\|_{\mathcal{N}(K), I_n}^2 \quad (B.26)$$

**Proof** Letting  $w = v - \bar{I}_h v$ , the left-hand side of (B.24) is

$$\begin{aligned} \|w\|_{B_h}^2 &= \sum_{n=1}^N \underbrace{\int_{I_n} \|w\|_{A_{h,t}}^2 dt}_{=I} + \sum_{n=1}^N \underbrace{\|\bar{n}^t\|^{1/2} [w]\|_{\bar{\Gamma}_n}^2}_{=II} + \sum_{i=1}^2 \sum_{n=1}^{N-1} \underbrace{\|[w]_n\|_{\Omega_{i,n}}^2}_{=III} \\ &\quad + \sum_{i=1}^2 \underbrace{\|w_N^-\|_{\Omega_{i,N}}^2}_{=IV} + \sum_{i=1}^2 \underbrace{\|w_0^+\|_{\Omega_{i,0}}^2}_{=V} \end{aligned} \quad (B.27)$$

We consider the terms separately, starting with first.

$$\begin{aligned}
 \text{I} &= \underbrace{\sum_{i=1}^2 \int_{I_n} \|\nabla w\|_{\Omega_i(t)}^2 dt}_{= \text{I.i}} + \underbrace{\int_{I_n} |\bar{\mu}| \|\langle \partial_{\bar{n}^x} w \rangle\|_{-1/2,h,\Gamma(t)}^2 dt}_{= \text{I.ii}} \\
 &+ \underbrace{\int_{I_n} |\bar{\mu}| \|[w]\|_{1/2,h,\Gamma(t)}^2 dt}_{= \text{I.iii}} + \underbrace{\int_{I_n} \|[\nabla w]\|_{\Omega_O(t)}^2 dt}_{= \text{I.iv}}
 \end{aligned}
 \tag{B.28}$$

Letting  $w_j^n = v - \bar{I}_{h,j}^n v$ , we treat each term in (B.28) separately, starting with the first.

$$\text{I.i} \leq \int_{I_n} \sum_{K \in \mathcal{T}_j, \Omega_i(t)} \|\nabla w_j^n\|_K^2 dt \leq \sum_{K \in \mathcal{T}_j, S_{i,n}} \|\nabla w_j^n\|_{\bar{K}_n}^2
 \tag{B.29}$$

By using standard estimates, (A.15), and (A.11), the second term is

$$\begin{aligned}
 \text{I.ii} &\lesssim \int_{I_n} \sum_{i=1}^2 \sum_{K_j \in \mathcal{T}_j, \Gamma(t)} h_{K_j} \|(\nabla w)_i\|_{\Gamma_{K_j}}^2 dt \\
 &\lesssim \sum_{i=1}^2 \sum_{K \in \mathcal{T}_j, \bar{\Gamma}_n} \left( \|\nabla w_j^n\|_{\bar{K}_n}^2 + h_K^2 \|D_x^2 w_j^n\|_{\bar{K}_n}^2 \right)
 \end{aligned}
 \tag{B.30}$$

For the third term we use the same standard estimates and again (A.15), thus

$$\text{I.iii} \lesssim \int_{I_n} |\bar{\mu}| \sum_{i=1}^2 \sum_{K_j \in \mathcal{T}_j, \Gamma(t)} h_{K_0}^{-1} \|w_i\|_{\Gamma_{K_j}}^2 dt \leq h_{\min}^{-1} \sum_{i=1}^2 \|(w_j^n)_i\|_{\bar{\Gamma}_n}^2
 \tag{B.31}$$

The fourth term is

$$\text{I.iv} \lesssim \int_{I_n} \sum_{i=1}^2 \|(\nabla w)_i\|_{\Omega_O(t)}^2 dt \leq \sum_{i=1}^2 \sum_{K \in \mathcal{T}_j, \bar{\Gamma}_n} \|\nabla w_j^n\|_{\bar{K}_n}^2
 \tag{B.32}$$

We are done with the separate treatments of all the terms in (B.28) and move on to the second term in (B.27). For this term, using that  $|\bar{n}^t| \leq |\mu|$  and (4.2) results in a factor that is I.iii which may simply be estimated by (B.31), thus

$$\text{II} \leq |\mu|_{(0,T)} h \int_{I_n} |\bar{\mu}| \|[w]\|_{1/2,h,\Gamma(t)}^2 dt \lesssim |\mu|_{(0,T)} h h_{\min}^{-1} \sum_{i=1}^2 \|(w_j^n)_i\|_{\bar{\Gamma}_n}^2
 \tag{B.33}$$

Combining the third, fourth and fifth term in (B.27), we have

$$\begin{aligned}
 \text{III} + \text{IV} + \text{V} &\lesssim \sum_{n=1}^N \left( \|w_n^-\|_{\Omega_{i,n}}^2 + \|w_{n-1}^+\|_{\Omega_{i,n-1}}^2 \right) \\
 &\leq \sum_{n=1}^N \left( \sum_{K \in \mathcal{T}_{j,\Omega_{i,n}}} \|(w_j^n)_n^-\|_K^2 + \sum_{K \in \mathcal{T}_{j,\Omega_{i,n-1}}} \|(w_j^n)_{n-1}^+\|_K^2 \right)
 \end{aligned}
 \tag{B.34}$$

The separate treatments of all the terms in (B.27) are done. The obtained estimates give

$$\begin{aligned}
 \|w\|_{B_h}^2 &\lesssim \sum_{n=1}^N \sum_{i=1}^2 \left( \sum_{K \in \mathcal{T}_{j,S_{i,n}}} \underbrace{\|\nabla w_j^n\|_{\tilde{K}_n}^2}_{=A} + \sum_{K \in \mathcal{T}_{j,\tilde{\Gamma}_n}} \underbrace{h_K^2 \|D_x^2 w_j^n\|_{\tilde{K}_n}^2}_{=B} \right. \\
 &\quad \left. + \underbrace{h_{\min}^{-1} \|(w_j^n)_i\|_{\tilde{\Gamma}_n}^2}_{=C} + \sum_{K \in \mathcal{T}_{j,\Omega_{i,n}}} \underbrace{\|(w_j^n)_n^-\|_K^2}_{=D} + \sum_{K \in \mathcal{T}_{j,\Omega_{i,n-1}}} \underbrace{\|(w_j^n)_{n-1}^+\|_K^2}_{=E} \right)
 \end{aligned}
 \tag{B.35}$$

We proceed by considering the five different types of terms separately. For term A we use Lemma B.1 with  $r = 1$ :

$$A = \|\nabla(v - \tilde{I}_{h,j}^n v)\|_{\tilde{K}_n}^2 \lesssim k_n^{2(q+1)} \|D_t^{q+1} \nabla v\|_{\tilde{K}_n}^2 + h^{2p} k_n \|D_x^{p+1} v\|_{\mathcal{N}(K), I_n}^2
 \tag{B.36}$$

For term B we apply Lemma B.1 with  $r = 2$ :

$$B = h_K^2 \|D_x^2(v - \tilde{I}_{h,j}^n v)\|_{\tilde{K}_n}^2 \lesssim k_n^{2(q+1)} \|D_t^{q+1} D_x^2 v\|_{\tilde{K}_n}^2 + h^{2p} k_n \|D_x^{p+1} v\|_{\mathcal{N}(K), I_n}^2
 \tag{B.37}$$

For term C we use Lemma B.2 and (3.1):

$$\begin{aligned}
 C &= h_{\min}^{-1} \|(v - \tilde{I}_{h,j}^n v)_i\|_{\tilde{\Gamma}_n}^2 \\
 &\lesssim k_n^{2q+1} \|D_t^{q+1} v\|_{L^2(\tilde{\Gamma}_n, L^\infty(I_n))}^2 + h^{2p} \sum_{K \in \mathcal{T}_{j,\tilde{\Gamma}_n}} k_n \|D_x^{p+1} v\|_{\mathcal{N}(K), I_n}^2
 \end{aligned}
 \tag{B.38}$$

By applying Lemma B.3 to term D and using (3.1), we get

$$D = \|(v - \tilde{I}_{h,j}^n v)_n^-\|_K^2 \lesssim h^{2p} k_n \|D_x^{p+1} v\|_{\mathcal{N}(K), I_n}^2
 \tag{B.39}$$

Using Lemma B.3 and (3.1) for term E, we get

$$E = \|(v - \tilde{I}_{h,j}^n v)_{n-1}^+\|_K^2 \lesssim k_n^{2q+1} \|D_t^{q+1} v\|_{\tilde{K}_n}^2 + h^{2p} k_n \|D_x^{p+1} v\|_{\mathcal{N}(K), I_n}^2
 \tag{B.40}$$

Using these local estimates in (B.35) gives (B.24).  $\square$

By applying Lemma B.3 and using (3.1), we get the estimate:

**Corollary B.1** (A global spatial interpolation error estimate for temporal endpoints) *Let  $\bar{I}_h$  and  $F_h$  be defined by (B.19) and (B.26), respectively. Then, for any function  $v$  with sufficient spatial and temporal regularity, we have that*

$$\sum_{n=1}^N \|(v - \bar{I}_h v)_n^-\|_{\Omega_0}^2 \lesssim h^{2p} F_h^2(v) \quad (\text{B.41})$$

**Funding** Open access funding provided by Umea University.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Nitsche, J.: Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. In: *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 36, pp. 9–15. Springer (1971)
2. Hansbo, A., Hansbo, P.: An unfitted finite element method, based on Nitsche's method, for elliptic interface problems. *Comp. Methods Appl. Mech. Engrg.* **191**(47), 5537–5552 (2002)
3. Hansbo, A., Hansbo, P., Larson, M.G.: A finite element method on composite grids based on Nitsche's method. *ESAIM Math. Model. Numer. Anal.* **37**(03), 495–514 (2003)
4. Burman, E., Hansbo, P.: A unified stabilized method for Stokes' and Darcy's equations. *J. Comput. Appl. Math.* **198**(1), 35–51 (2007)
5. Burman, E., Fernández, M.A.: Stabilized explicit coupling for fluid-structure interaction using Nitsche's method. *C. R. Math. Acad. Sci. Paris* **345**(8), 467–472 (2007)
6. Becker, R., Burman, E., Hansbo, P.: A Nitsche extended finite element method for incompressible elasticity with discontinuous modulus of elasticity. *Comp. Methods Appl. Mech. Eng.* **198**(41), 3352–3360 (2009)
7. Massing, A., Larson, M.G., Logg, A., Rognes, M.E.: A stabilized Nitsche fictitious domain method for the stokes problem. *J. Sci. Comput.* **61**(3), 604–628 (2014)
8. Massing, A., Larson, M.G., Logg, A., Rognes, M.E.: A stabilized Nitsche overlapping mesh method for the Stokes problem. *Numer. Math.* **128**(1), 73–101 (2014)
9. Johansson, A., Larson, M.G., Logg, A.: High order cut finite element methods for the Stokes problem. *Adv. Model. Simul. Eng. Sci.* **2**(1), 1–23 (2015). <https://doi.org/10.1186/s40323-015-0043-7>
10. Dokken, J.S., Funke, S.W., Johansson, A., Schmidt, S.: Shape optimization using the finite element method on multiple meshes with Nitsche coupling. *SIAM J. Sci. Comput.* **41**(3), A1923–A1948 (2019)
11. Johansson, A., Kehlet, B., Larson, M.G., Logg, A.: Multimesh finite element methods: solving PDEs on multiple intersecting meshes. *Comput. Methods Appl. Mech. Eng.* (2019)
12. Lehrenfeld, C., Reusken, A.: Analysis of a Nitsche XFEM-DG discretization for a class of two-phase mass transport problems. *SIAM J. Numer. Anal.* **51**(2), 958–983 (2013). <https://doi.org/10.1137/120875260>



13. Voulis, I., Reusken, A.: A time dependent Stokes interface problem: well-posedness and space-time finite element discretization. *ESAIM: Math. Model. Numer. Anal.* **52**(6), 2187–2213 (2018)
14. Lehrenfeld, C., Olshanskii, M.: An Eulerian finite element method for PDEs in time-dependent domains. *ESAIM: Math. Model. Numer. Anal.* **53**(2), 585–614 (2019)
15. Preuß, J.: Higher order unfitted isoparametric space-time FEM on moving domains, Master's thesis, GRO.data, University of Göttingen (2021). <https://data.goettingen-research-online.de/dataset.xhtml?persistentId=https://doi.org/10.25625/UACWXS>
16. von Wahl, H., Richter, T., Lehrenfeld, C.: An unfitted Eulerian finite element method for the time-dependent Stokes problem on moving domains. *IMA J. Numer. Anal.* **42**(3), 2505–2544 (2022)
17. Heimann, F., Lehrenfeld, C., Preuß, J.: Geometrically higher order unfitted space-time methods for PDEs on moving domains. *SIAM J. Sci. Comput.* **45**(2), B139–B165 (2023). <https://doi.org/10.1137/22M1476034>
18. Badia, S., Dilip, H., Verdugo, F.: Space-time unfitted finite element methods for time-dependent problems on moving domains. *Comput. Math. Appl.* **135**, 60–76 (2023)
19. Olshanskii, M.A., Reusken, A.: Trace Finite Element Methods for PDEs on Surfaces. In: Bordas, S.P.A., Burman, E., Larson, M.G., Olshanskii, M.A. (eds.) *Geometrically Unfitted Finite Element Methods and Applications*, vol. 121, pp. 211–258. Series Title: *Lecture Notes in Computational Science and Engineering*. Springer International Publishing, Cham (2017). [https://doi.org/10.1007/978-3-319-71431-8\\_7](https://doi.org/10.1007/978-3-319-71431-8_7)
20. Sass, H.: Space-time trace finite element methods for partial differential equations on evolving surfaces, Ph.D. dissertation, RWTH Aachen University, 2022, number: RWTH-2022-09895. <https://publications.rwth-aachen.de/record/854968>
21. Olshanskii, M.A., Reusken, A., Zhiliakov, A.: Tangential Navier–Stokes equations on evolving surfaces: analysis and simulations. *Math. Models Methods Appl. Sci.* **32**(14), 2817–2852 (2022). <https://doi.org/10.1142/S0218202522500658>
22. Olshanskii, M., Reusken, A., Schwering, P.: An Eulerian finite element method for tangential Navier–Stokes equations on evolving surfaces. *Math. Comput.* (2023). <https://www.ams.org/mcom/0000-0000/S0025-5718-2023-03931-X/>
23. Lundholm, C.: A space-time cut finite element method for a time-dependent parabolic model problem, Master's thesis, Chalmers University of Technology and University of Gothenburg (2015). <https://odr.chalmers.se/items/c8f07fea-5b84-44c3-8409-d1bd69845b97>
24. Lundholm, C.: *Cut Finite Element Methods on Overlapping Meshes: Analysis and Applications*, Ph.D. dissertation, Chalmers University of Technology and University of Gothenburg (2021). <https://research.chalmers.se/en/publication/524200>
25. Larson, M.G., Lundholm, C.: Space-time CutFEM on overlapping meshes II: simple discontinuous mesh evolution. *Numerische Mathematik* (2024)
26. Hughes, T., Brooks, A.: A multi-dimensional upwind scheme with no crosswind diffusion. In: Hughes, T.J.R. (ed.) *Finite Element Methods for Convection Dominated Flows*, ASME Winter Annual Meeting, New York, USA., vol. 34, pp. 19–35 (1979)
27. Johnson, C., Nävert, U.: An analysis of some finite element methods for advection-diffusion problems. In: Axelsson, O., Frank, L., Van Der Sluis, A. (eds.) *Analytical and Numerical Approaches to Asymptotic Problems in Analysis*, Ser. *North-Holland Mathematics Studies*, . North-Holland, vol. 47, pp. 99–116 (1981). <https://www.sciencedirect.com/science/article/pii/S0304020808711046>
28. Johnson, C., Pitkäranta, J.: An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. *Math. Comput.* **46**(173), 1–26 (1986)
29. Di Pietro, D.A., Ern, A.: *Mathematical Aspects of Discontinuous Galerkin Methods*, Ser. *Mathématiques et Applications*, vol. 69. Springer, Berlin, Heidelberg (2012). <https://doi.org/10.1007/978-3-642-22980-0>
30. Cangiani, A., Dong, Z., Georgoulis, E.H.:  $\mathbb{P}_k$ -version space-time discontinuous Galerkin methods for parabolic problems on prismatic meshes. *SIAM J. Sci. Comput.* **39**(4), A1251–A1279 (2017). <https://doi.org/10.1137/16M1073285>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.