

# Space–time discontinuous Galerkin approximation of acoustic waves with point singularities

P. Bansal and A. Moiola and I. Perugia and Ch. Schwab

Research Report No. 2020-10

February 2020

Latest revision: August 2020

Seminar für Angewandte Mathematik  
Eidgenössische Technische Hochschule  
CH-8092 Zürich  
Switzerland

---

# Space–time discontinuous Galerkin approximation of acoustic waves with point singularities

Pratyuksh Bansal\*, Andrea Moiola†, Ilaria Perugia‡, Christoph Schwab§

August 5, 2020

## Abstract

We develop a convergence theory of space–time discretizations for the linear, 2nd-order wave equation in polygonal domains  $\Omega \subset \mathbb{R}^2$ , possibly occupied by piecewise homogeneous media with different propagation speeds. Building on an unconditionally stable space–time DG formulation developed in [33], we (a) prove optimal convergence rates for the space–time scheme with local isotropic corner mesh refinement on the spatial domain, and (b) demonstrate numerically optimal convergence rates of a suitable *sparse* space–time version of the DG scheme. The latter scheme is based on the so-called *combination formula*, in conjunction with a family of anisotropic space–time DG-discretizations. It results in optimal-order convergent schemes, also in domains with corners, with a number of degrees of freedom that scales essentially like the DG solution of one stationary elliptic problem in  $\Omega$  on the finest spatial grid. Numerical experiments for both smooth and singular solutions support convergence rate optimality on spatially refined meshes of the full and sparse space–time DG schemes.

## 1 Introduction

In recent years, substantial interest and progress has been made on so-called *space–time discretizations of evolution equations*. This development has been driven by several factors. We mention only the need to compute the evolution of the solution over a finite time interval  $(0, T)$  rather than a sequence of spatial solutions resulting in the solution only at the final time horizon  $T$ . This issue arises for example in optimal control of parabolic and hyperbolic PDEs, where the feedback depends on the solution of an adjoint PDE which is driven by a functional of the forward solution. Similar issues and needs arise in space–time optimization of such PDEs, and in particular in space–time adaptivity. Here, traditional time-marching schemes encounter several difficulties which can be easily overcome by combined space–time, “monolithic” discretizations. This is purchased, however, with a number of new issues: increased memory requirement for the simultaneous storage of the entire solution history over  $(0, T)$ , efficient solvers, etc. In the present paper, we address a class of space–time discretizations of the linear, acoustic wave equation in two space dimensions, in polygonal domains which are occupied by possibly heterogeneous media with corresponding variable wave speeds. Acoustic (and other types of) waves exhibit diffraction at conical singularities such as corners or multi-material interfaces (e.g. [6, 26, 30]). It has been clarified in the past decades (e.g. [27–29] and the references there) that these singularities are, in a sense, the hyperbolic counterpart of elliptic conical singularities, and that the solutions of linear, second order hyperbolic equations admit regularity results in scales of corner-weighted spaces. Based on the regularity results in [27–29, 31], it was shown by some of the authors of the present manuscript in [36] that high-order, time-marching discretizations of linear, acoustic wave equations with FEM discretizations in the spatial domain can recover the maximal convergence rate afforded by the elemental polynomial degree, even in the presence of conical singularities.

### 1.1 Previous results

The necessity of some form of stabilization has been recognized early on in the development of space–time discretizations. We refer to the classical stabilized spacetime FEM [23, 24], where least squares, mesh-dependent stabilization has been introduced. For general first-order systems, Monk–Richter [34] introduced a DG space–time scheme based on numerical fluxes which upwind in the transport direction.

---

\*SAM ETH Zürich, Rämistrasse 101, CH-8092 Zürich (pratyuksh.bansal@sam.math.ethz.ch)

†Department of Mathematics, University of Pavia, Via Ferrata 5, I-27100 Pavia (andrea.moiola@unipv.it)

‡Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna (ilaria.perugia@univie.ac.at)

§SAM ETH Zürich Rämistrasse 101, CH-8092 Zürich, (christoph.schwab@sam.math.ethz.ch)

More recently, Trefftz DG methods, i.e. schemes employing locally exact solutions of the wave equation, have been considered by Muiola–Perugia [33], following the DG framework of [34], and Banjai–Georgoulis–Lijoka [3], in an interior penalty (IP) setting. Most of these error analyses have been developed under the assumption that the exact solution is sufficiently smooth, as a function of space and time. In polygonal spatial domains or in domains composed of different homogeneous materials, high regularity of solutions in standard Sobolev scales is known to fail, due to diffractive solution components with point singularities. We refer to results obtained in recent years on the regularity of solutions, also in nonsmooth domains, which are phrased in terms of *corner-weighted Sobolev spaces of Kondrat’ev type*, see [29] and the references there, and the more recent Luong–Tung [31]. The solutions of linear, second-order wave equations in polygons has been shown in these references to consist of a smooth part plus a finite, time-dependent linear combination of singular “corner” solutions, with coefficients which depend regularly on the time-variable. The canonical corner solutions correspond to certain solutions of certain conical diffraction problems in infinite wedges given by tangent cones to the polygon with vertices in the corners.

A-priori estimates of solutions in corner-weighted spaces were used in Müller–Schwab [37], and in the PhD thesis [35], to develop convergence rate bounds for time-marching spatial DG-FEM approximations of time-domain wave propagation in polygonal domains under realistic regularity assumptions on the solution.

Concurrent (full-tensor, in our terminology) space–time discretizations for linear wave equations were already proposed by [23, 24]. Error bounds for smooth solutions were obtained and numerical examples in one spatial dimension were reported. For related recent numerical analysis results of space–time discretization schemes for evolution equations, we mention the work Cangiani–Dong–Georgoulis [9] where space–time DG schemes for linear, parabolic PDEs were investigated on prismatic meshes. Maximal regularity of the solution in the spatial and the temporal variable was assumed in the analysis in [9]. Steinbach–Zank [43] proposed a stabilization for a conforming space–time finite element method on Cartesian meshes, in order to overcome time-step restrictions in meshes that are locally refined in space. Ernesti–Wieners [15] consider a (full-tensor) *Petrov-Galerkin (DPG) discretization* of linear acoustic wave equations, and prove a-priori error bounds assuming maximal solution regularity in the spatial and temporal variable in standard Sobolev spaces, i.e. precluding point singularities admitted in the present work. The work [13] focuses again on space–time discretization of the heat equation, and addresses the solver complexity. A space–time DG discretization for the parabolic Navier-Stokes equation was proposed in [42].

For the linear, acoustic wave equation, the CFL constraint encountered in explicit time-stepping is exacerbated by local spatial mesh-refinement near corners. One remedy with a space–time discretization flavour is here the so-called *local time-stepping*. We refer to Diaz–Grote [12] for details and to [41] for a method-of-lines based, related approach to circumvent the CFL-constraint in explicit time-marching.

The recent work on explicit, marching-type space–time schemes by Gopalakrishnan–Schöberl–Wintersteiger [19] is rooted in Falk–Richter [16]. These schemes are based on so-called *tent-pitched space-time meshes*, where the PDE can be explicitly evolved in the causal direction, i.e. from the “bottom” to the “top” of the space–time cylinder element by element, thereby avoiding global CFL constraints. This evolution is performed, after mapping each element into a space–time cylinder, by applying a Runge-Kutta or a Taylor time-stepping; see [19] and [18], respectively. An alternative is to combine tent-pitching with Trefftz basis functions [40].

## 1.2 Contributions

In the present paper, we obtain the following novel contributions. Firstly, we extend the consistency analysis of the space–time DG discretization proposed in [33] to general (namely non-Trefftz) discrete spaces and to the (realistic) setting of solutions which exhibit spatial point singularities situated at corners of the spatial domain  $\Omega$  or, for transmission problems, at multimaterial interface points. These points generate, in two spatial dimensions, diffraction terms which propagate radially with the acoustic speed of sound afforded by the medium. In this case, convergence rate bounds which are based on maximal spatial regularity of solutions on quasi-uniform meshes of the spatial domain are moot, as higher order spatial regularity can only be expressed in scales of *corner-weighted Sobolev spaces* of Kondrat’ev type. Our present error analysis and convergence rate bounds (see Proposition 7.4) also cover *interior singularities* of solutions as arise, for example, due to multi-material interfaces in  $\Omega$ , based on corresponding regularity in weighted Sobolev scales; see [36, 37] and the references there.

Furthermore, in this (realistic, i.e. including corner-singularities) setting, we propose a *novel, sparse space-time DG discretization*. It allows to build an approximate solution of the evolution equation in error vs. work which corresponds, asymptotically, to that of one elliptic solve at the highest spatial resolution in  $\Omega$ , with the corresponding number of degrees of freedom. This is comparable with the error vs. work

offered, for example, by integral equation based methods combined with convolution quadrature (see, e.g., [5] and the references there), without access to explicit fundamental solutions.

The *consistency error bounds* in mesh-dependent norm we provide for the full space–time DG scheme holds true without any time-step size constraint, also in the presence of corner singularities and spatially refined meshes. This indicates that its sparse–tensor version may feature superior error vs. work performance, as observed subsequently in detailed numerical experiments. Unlike integral-equation-based methods, which require an explicit fundamental solution, the presently proposed results and space–time DG discretization does *not* mandate homogeneous media: the presently considered *DG formulation* can accommodate also more general, inhomogeneous but (piecewise) sufficiently regular coefficients (in particular the wave speed). Here we focus on the piecewise-constant materials case and leave the extension of the analysis to smooth coefficients to future work.

Some further comments on the novelty and the generality of our results are in order. Part of the abstract analysis of the DG scheme follows that of [33, 34], but it differs in that the techniques used in these references do not allow the use of tensor-product discrete spaces, see Remark 5.4 for details. The presently developed error analysis is a high-order “*h*-convergence” analysis; it does not establish “*p*-version” convergence for increasing polynomial degrees on a fixed space–time mesh. Finally, we only consider two-dimensional spatial domains as the regularity and corner singularity theory for the wave equation in polygons, in weighted spaces of finite order is nowadays quite complete, as described above. The corresponding theory for solutions in polyhedra (with diffraction from both, vertex and edge singularities) will mandate *anisotropic mesh refinement in the vicinity of edges* which, in turn, requires significant technical modifications in all parts of the present DG error analysis and whose development is beyond the scope of the present paper.

### 1.3 Outline

The structure of this paper is as follows. In §2 we state initial boundary value problems for the acoustic wave equation for a homogeneous medium in both first- and second-order formulations; we also provide sufficient conditions for their well-posedness. §3 reviews relevant regularity results on the exact solution from [35, 37], based on [29, 31]. §4 prepares the (somewhat involved) notation for the ensuing development of the space–time DG discretization. §4.5 in particular has some polynomial inverse inequalities. §5 then develops the space–time DG formulation, with corresponding existence and uniqueness results of the discrete solution, and abstract, “quasi-optimality-like” a-priori error bounds. §6 then provides the convergence rate bounds for the full-tensor space–time DG scheme, subject to either smooth solutions or solutions belonging to corner-weighted spaces. §7 discusses the generation and properties of families of corner-refined triangulations of the spatial domain and their complexity, leading to convergence rate bounds for the space–time DG scheme posed on such meshes. §8 describes the outcomes of several numerical examples involving smooth and singular solutions in homogeneous and heterogeneous media, quasi-uniform and locally refined meshes. §9 finally introduces the sparse space–time DG discretization, compares its computational complexity against the full-tensor approach, and provides numerical results for the sparse version of the scheme applied to the same examples of the previous section.

## 2 Initial boundary value problem

In this section, we introduce the model problem given by linear, acoustic wave propagation in a bounded, polygonal domain  $\Omega$ . We discuss its well-posedness.

### 2.1 Problem statement

We consider an initial boundary value problem (IBVP) for the linear, acoustic wave-equation posed on a space–time domain  $Q = \Omega \times I$ , where  $\Omega \subset \mathbb{R}^2$  is an open, bounded, Lipschitz polygon with straight sides and with outward unit normal  $\mathbf{n}_\Omega^x$ , and where  $I = (0, T)$ ,  $T > 0$ , denotes a time-interval with finite time horizon.

The boundary of  $\Omega$  is divided in two parts, with mutually disjoint interiors, denoted  $\Gamma_D$  and  $\Gamma_N$  corresponding to Dirichlet and Neumann boundary conditions, respectively; one of them may be empty.

The first-order acoustic wave IBVP reads as

$$\begin{cases} \nabla v + \frac{\partial \boldsymbol{\sigma}}{\partial t} = \mathbf{0} & \text{in } Q, \\ \nabla \cdot \boldsymbol{\sigma} + c^{-2} \frac{\partial v}{\partial t} = f & \text{in } Q, \\ v(\cdot, 0) = v_0, \quad \boldsymbol{\sigma}(\cdot, 0) = \boldsymbol{\sigma}_0 & \text{on } \Omega, \\ v = g_D, & \text{on } \Gamma_D \times [0, T], \\ \boldsymbol{\sigma} \cdot \mathbf{n}_\Omega^x = g_N, & \text{on } \Gamma_N \times [0, T]. \end{cases} \quad (1)$$

Here  $f, v_0, \boldsymbol{\sigma}_0, g_D, g_N$  are the problem data;  $c(\mathbf{x}) \geq c_0 > 0$  is the wave speed, which is assumed to be piecewise constant on a fixed, finite polygonal partition  $\{\Omega_i\}$  of  $\Omega$  which is independent of  $t$ . We denote by  $\mathcal{S} := \{\mathbf{c}_i, i = 1 \dots, M\}$  the set of all vertices of the polygons  $\Omega_j \subset \Omega$  on which  $c$  is constant. We also assume that  $\overline{\Omega_j} \cap \partial\Omega$  is contained in either  $\overline{\Gamma_D}$  or in  $\overline{\Gamma_N}$ . This IBVP is a special case of [33, (1)] and [34, §6], where also impedance boundary conditions were allowed.

The corresponding IBVP for the second-order scalar wave equation is:

$$\begin{cases} -\Delta u + c^{-2} \frac{\partial^2 u}{\partial t^2} = f & \text{in } Q, \\ \frac{\partial u}{\partial t}(\cdot, 0) = v_0, \quad u(\cdot, 0) = u_0 & \text{on } \Omega, \\ \frac{\partial u}{\partial t} = g_D, & \text{on } \Gamma_D \times [0, T], \\ -\mathbf{n}_\Omega^x \cdot \nabla u = g_N, & \text{on } \Gamma_N \times [0, T]. \end{cases} \quad (2)$$

Formally, given a smooth solution  $u$  of (2), its first-order derivatives  $(v, \boldsymbol{\sigma}) = (\frac{\partial u}{\partial t}, -\nabla u)$  constitute a solution of (1) with  $\boldsymbol{\sigma}_0 = -\nabla u_0$ . Vice versa, if  $(v, \boldsymbol{\sigma})$  is a smooth solution of (1) with  $\boldsymbol{\sigma}_0 = -\nabla u_0$ , then  $u(\cdot, t) = u_0 + \int_0^t v(\cdot, s) ds$  is solution of (2).

The IBVP (2) can be written with the Dirichlet condition in the more commonly encountered form  $u = G_D$  on  $\Gamma_D \times [0, T]$  simply by taking  $G_D(\mathbf{x}, t) := u_0(\mathbf{x}) + \int_0^t g_D(\mathbf{x}, s) ds$  for  $(\mathbf{x}, t) \in \Gamma_D \times [0, T]$ , for, e.g., continuous  $u_0$  and  $g_D$ . An IBVP with boundary condition  $u = G_D$  on  $\Gamma_D \times [0, T]$  satisfying  $G_D(\mathbf{x}, 0) = u_0(\mathbf{x})$  can be written as (2) simply by taking  $g_D = \frac{\partial G_D}{\partial t}$ .

## 2.2 Variational solutions

In order to give conditions ensuring the well-posedness of the first-order IBVP (1), we firstly consider the analogous problem for the second-order wave equation (2). In this section we only consider the homogeneous Dirichlet case, namely with  $\Gamma_N = \emptyset$  and  $g_D = 0$ .

A classical result on variational methods for evolution equations, see [11, pp. 581–582, Chapt. XVIII, §6.1], states that, if

$$f \in L^2(Q), \quad u_0 \in H_0^1(\Omega), \quad v_0 \in L^2(\Omega), \quad \Gamma_N = \emptyset, \quad g_D = 0,$$

then (2) admits a unique variational solution

$$u \in C^0([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega)) \cap H^2(0, T; H^{-1}(\Omega)), \quad (3)$$

where we use the Bochner-space notation as in, e.g., [11]. The fact that  $\frac{\partial^2 u}{\partial t^2} \in L^2(0, T; H^{-1}(\Omega))$  follows from the wave equation, the space regularity of  $u$  and  $f$ , and the mapping  $\Delta : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$ .

We state some simple properties of negative-regularity Sobolev spaces of vector fields. We define  $H_0(\text{div}; \Omega)^*$  as the dual space of the classical space  $H_0(\text{div}; \Omega) = \{\boldsymbol{\tau} \in L^2(\Omega)^2, \nabla \cdot \boldsymbol{\tau} \in L^2(\Omega), \boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x = 0 \text{ in } H^{-\frac{1}{2}}(\partial\Omega)\}$ . Then, for all  $w \in L^2(\Omega)$ , we have that  $\nabla w \in H_0(\text{div}; \Omega)^*$ : this is because the duality  $\langle \nabla w, \mathbf{z} \rangle_{H_0(\text{div}; \Omega)^* \times H_0(\text{div}; \Omega)} = -\int_\Omega w \nabla \cdot \mathbf{z} dx$  for  $\mathbf{z} \in H_0(\text{div}; \Omega)$  is an extension of the  $(L^2(\Omega))^2$  scalar product (equivalently,  $\nabla : L^2(\Omega) \rightarrow H_0(\text{div}; \Omega)^*$  as it is the adjoint of  $-\nabla \cdot : H_0(\text{div}; \Omega) \rightarrow L^2(\Omega)$ ). Moreover, the embedding  $H_0(\text{div}; \Omega)^* \subset (H^{-1}(\Omega))^2$  holds, since the inclusion  $(H_0^1(\Omega))^2 \subset H_0(\text{div}; \Omega)$  is dense and continuous and  $H^{-1}(\Omega) = H_0^1(\Omega)^*$ .

We now make use of the second-order equation in the context of the first-order system (1). If

$$f \in L^2(Q), \quad \boldsymbol{\sigma}_0 \in L^2(\Omega)^2, \quad v_0 \in L^2(\Omega), \quad \Gamma_N = \emptyset, \quad g_D = 0, \quad (4)$$

we define  $u_0 \in H_0^1(\Omega)$  to be the solution of the Poisson equation  $-\Delta u_0 = \nabla \cdot \boldsymbol{\sigma}_0 \in H^{-1}(\Omega)$  in  $\Omega$ . Then we denote by  $u$  the unique variational solution of (2) with such initial condition  $u_0$ , and define

$$(v, \boldsymbol{\sigma}) := \left( \frac{\partial u}{\partial t}, -\nabla u + \nabla u_0 + \boldsymbol{\sigma}_0 \right) \in \left( C^0([0, T], L^2(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)) \right) \times \left( C^0([0, T], (L^2(\Omega))^2) \cap C^1([0, T]; H_0(\text{div}; \Omega)^*) \right). \quad (5)$$

Then,  $(v, \boldsymbol{\sigma})$  is solution of the first-order IBVP (1). Note that, in the definition of  $\boldsymbol{\sigma}$ , the time-independent term  $\nabla u_0 + \boldsymbol{\sigma}_0$  is needed in order to deal with initial conditions  $\boldsymbol{\sigma}_0$  that are not gradients, see [33, Rem. 1].

Thus, for all data satisfying (4), we have defined a solution field  $(v, \boldsymbol{\sigma})$ . In which sense does  $(v, \boldsymbol{\sigma})$  solve the IBVP (1)? If it is smooth, e.g.  $(v, \boldsymbol{\sigma}) \in (C^1(Q) \cap C^0(\overline{Q})) \times (C^1(Q) \cap C^0(\overline{Q}))^2$ , then (1) holds pointwise in the classical sense. More generally, we deduce from (5) that the two PDEs in (1) hold in the following spaces:

$$\nabla v + \frac{\partial \boldsymbol{\sigma}}{\partial t} = \mathbf{0} \text{ in } C^0([0, T]; H_0(\text{div}; (\Omega))^*), \quad \nabla \cdot \boldsymbol{\sigma} + c^{-2} \frac{\partial v}{\partial t} = f \text{ in } L^2(0, T; H^{-1}(\Omega))$$

and the initial conditions  $v = v_0$ ,  $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0$  hold in  $L^2(\Omega)$ . The reason why the scalar equation is valid in  $L^2(0, T; H^{-1}(\Omega))$  only and not in  $C^0([0, T]; H^{-1}(\Omega))$  is that  $f \in L^2(Q)$ . The spaces in (5) do not allow to take traces of  $(v, \boldsymbol{\sigma})$  on  $\partial\Omega \times (0, T)$ , so the boundary condition  $v = g_D = 0$  on  $\partial\Omega \times [0, T]$  is satisfied only weakly. If  $(v, \boldsymbol{\sigma})$  is a smooth solution, by testing the PDEs of (1) against a test field  $(w, \boldsymbol{\tau})$  and integrating by parts, we obtain

$$\begin{aligned} \mathcal{A}((v, \boldsymbol{\sigma}), (w, \boldsymbol{\tau})) &= - \int_Q \left( v \left( \nabla \cdot \boldsymbol{\tau} + c^{-2} \frac{\partial w}{\partial t} \right) + \boldsymbol{\sigma} \cdot \left( \nabla w + \frac{\partial \boldsymbol{\tau}}{\partial t} \right) \right) dV + \int_{\Omega \times \{T\}} (\boldsymbol{\sigma} \cdot \boldsymbol{\tau} + c^{-2} v w) dx \\ &= \int_Q f w dV + \int_{\Omega \times \{0\}} (\boldsymbol{\sigma}_0 \cdot \boldsymbol{\tau} + c^{-2} v_0 w) dx \\ \forall (w, \boldsymbol{\tau}) &\in \left( C^0([0, T]; H_0^1(\Omega)) \cap H^1(0, T; L^2(\Omega)) \right) \times \left( C^0([0, T]; H(\text{div}; \Omega)) \cap C^1([0, T]; (L^2(\Omega))^2) \right). \end{aligned} \quad (6)$$

Here and in the following, our notation for differentials within integrals is as follows: we use  $d\mathbf{x} = dx_1 dx_2$  within integrals over (two-dimensional) spatial regions,  $dV = d\mathbf{x} dt$  within volume integrals over (three-dimensional) space-time domains, and  $dS$  within surface integrals over general two-dimensional surfaces in space-time.

Note that, of the two boundary terms expected from integration by parts,  $\int_{\partial\Omega \times (0, T)} \boldsymbol{\sigma} \cdot \mathbf{n}_\Omega^x w dS$  is not present in  $\mathcal{A}(\cdot, \cdot)$  as  $w \in C^0([0, T]; H_0^1(\Omega))$ , while the absence of the term  $\int_{\partial\Omega \times (0, T)} v \boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x dS$  corresponds to weakly imposing the homogeneous Dirichlet boundary condition. By density<sup>1</sup>, the variational identity (6) is satisfied also by rougher  $(v, \boldsymbol{\sigma})$  with regularity (5): this makes precise the sense in which the boundary conditions are weakly satisfied by  $(v, \boldsymbol{\sigma})$ .

In order to show uniqueness of the variational solution of (1), let  $(v, \boldsymbol{\sigma})$  be a solution of (6), with  $f = 0$ ,  $v_0 = 0$ , and  $\boldsymbol{\sigma}_0 = \mathbf{0}$ , that belongs to the space indicated in (5). Define

$$u(\cdot, t) := \int_0^t v(\cdot, s) ds,$$

which, from the regularity (5) of  $(v, \boldsymbol{\sigma})$ , satisfies  $u \in C^1([0, T], L^2(\Omega)) \cap H^2(0, T; H^{-1}(\Omega))$ . Note that the continuity in time of  $v$  implies that  $u(\cdot, 0) = 0$  in  $L^2(\Omega)$ . As discussed above, in this setting,  $(v, \boldsymbol{\sigma})$  satisfies

$$\nabla v + \frac{\partial \boldsymbol{\sigma}}{\partial t} = \mathbf{0} \text{ in } C^0([0, T]; H_0(\text{div}; (\Omega))^*), \quad \nabla \cdot \boldsymbol{\sigma} + c^{-2} \frac{\partial v}{\partial t} = 0 \text{ in } C^0(0, T; H^{-1}(\Omega)). \quad (7)$$

<sup>1</sup>We sketch here the density argument. Given  $(v, \boldsymbol{\sigma})$  a variational solution as in (5), there is  $u$ , solution of the second order problem (2), with regularity (3). By density of smooth functions in Sobolev spaces, there exists a sequence  $u_j \in C^\infty(\overline{Q})$  such that  $u_j \rightarrow u$  in  $C^0([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega)) \cap H^2(0, T; H^{-1}(\Omega))$  norm. Defining  $(v_j, \boldsymbol{\sigma}_j) := (\frac{\partial u_j}{\partial t}, -\nabla u_j)$ , we have  $(v_j, \boldsymbol{\sigma}_j) \rightarrow (v, \boldsymbol{\sigma})$  in the norm of (5), and  $\nabla v_j + \frac{\partial \boldsymbol{\sigma}_j}{\partial t} = \mathbf{0}$ ,  $\nabla \cdot \boldsymbol{\sigma}_j + c^{-2} \frac{\partial v_j}{\partial t} = f_j$ , where  $f_j \rightarrow f$  in  $L^2(0, T; H^{-1}(\Omega))$ . Moreover  $(v_{0,j}, \boldsymbol{\sigma}_{0,j}) := (v_j, \boldsymbol{\sigma}_j)|_{t=0} \rightarrow (v, \boldsymbol{\sigma})$  in  $L^2(\Omega)^{1+2}$  because of the definition of the space in (5). From the continuity of the bilinear form  $\mathcal{A}$  in the topology in which we have convergence, we deduce that, for all  $(w, \boldsymbol{\tau})$  as in (6),

$$\mathcal{A}((v, \boldsymbol{\sigma}), (w, \boldsymbol{\tau})) = \lim_{j \rightarrow \infty} \mathcal{A}((v_j, \boldsymbol{\sigma}_j), (w, \boldsymbol{\tau})) = \lim_{j \rightarrow \infty} \int_Q f_j w + \int_{\Omega \times \{0\}} (\boldsymbol{\sigma}_{0,j} \cdot \boldsymbol{\tau} + c^{-2} v_{0,j} w) = \int_Q f w + \int_{\Omega \times \{0\}} (\boldsymbol{\sigma}_0 \cdot \boldsymbol{\tau} + c^{-2} v_0 w).$$



By integrating the first equation in (7) in time, taking into account that  $\boldsymbol{\sigma}_0 = \mathbf{0}$ , we obtain

$$\mathbf{0} = \int_0^t \nabla v(\cdot, s) \, ds + \int_0^t \frac{\partial \boldsymbol{\sigma}}{\partial t}(\cdot, s) \, ds = \nabla \int_0^t v(\cdot, s) \, ds + \boldsymbol{\sigma}(\cdot, t) = \nabla u(\cdot, t) + \boldsymbol{\sigma}(\cdot, t),$$

namely,

$$\nabla u(\cdot, t) = -\boldsymbol{\sigma}(\cdot, t) \text{ in } C^0([0, T], (L^2(\Omega))^2) \cap C^1([0, T]; H_0(\text{div}; \Omega)^*),$$

thus  $u$  also belongs to  $C^0([0, T], H^1(\Omega))$ .

The fact that  $u$  has zero Dirichlet trace on  $\partial\Omega$  follows from the homogeneous weak Dirichlet boundary condition on  $v$ , namely  $\int_{\partial\Omega \times (0, T)} v \boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x \, dS = 0$  for all  $\boldsymbol{\tau} \in C^0([0, T]; H(\text{div}; \Omega)) \cap C^1([0, T]; (L^2(\Omega))^2)$ . In fact, taking  $\boldsymbol{\tau}$  independent of time, replacing  $v$  by  $\frac{\partial u}{\partial t}$ , and integrating by parts in time immediately give that the trace of  $u(\cdot, T)$  on  $\partial\Omega$  is zero in the sense of  $H^{1/2}(\partial\Omega)$ , thus  $u(\cdot, T) \in H_0^1(\Omega)$ . Therefore, by integrating by parts in time, taking into account that the initial and final traces of  $u$  on  $\partial\Omega$  vanish, we have

$$0 = \int_{\partial\Omega \times (0, T)} v \boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x \, dS = \int_{\partial\Omega \times (0, T)} \frac{\partial u}{\partial t} \boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x \, dS = - \int_{\partial\Omega \times (0, T)} u \frac{\partial \boldsymbol{\tau}}{\partial t} \cdot \mathbf{n}_\Omega^x \, dS$$

for all  $\frac{\partial \boldsymbol{\tau}}{\partial t} \cdot \mathbf{n}_\Omega \in C^0([0, T]; H^{-1/2}(\partial\Omega))$ , from which we deduce that the trace of  $u$  on  $\partial\Omega \times (0, T)$  is zero in  $C^0([0, T]; H^{1/2}(\partial\Omega))$ . Consequently,  $u$  also belongs to  $C^0([0, T], H_0^1(\Omega))$ , and we conclude that  $u$  has the regularity (3).

Finally, inserting  $v = \frac{\partial u}{\partial t}$  and  $\boldsymbol{\sigma} = -\nabla u$  into the second equation in (7) gives  $-\Delta u + c^{-2} \frac{\partial^2 u}{\partial t^2} = 0$  in  $C^0(0, T; H^{-1}(\Omega))$ , together with  $u(\cdot, 0) = 0$  and  $\frac{\partial u}{\partial t}(\cdot, 0) = 0$ .

We have proven that  $u$  is actually variational solution of (2). Therefore, by invoking uniqueness of the variational solution of (2), we deduce uniqueness of the variational solution of (1). We summarize the results above in the following proposition.

**Proposition 2.1.** *Under assumptions (4), there exists a unique variational solution  $(v, \boldsymbol{\sigma})$ , in the sense of (6), of the IBVP (1), with regularity as in (5).*

Alternatively, when  $\Gamma_N = \emptyset$ ,  $f = 0$ ,  $g_D = 0$ ,  $v_0 \in H_0^1(\Omega)$ ,  $\boldsymbol{\sigma}_0 \in H(\text{div}; \Omega)$ , the well-posedness of IBVP (1) could be derived from the Lumer–Phillips theorem as in [21, pp. 238–239], without using the second-order problem.

### 3 Solution regularity

For the convergence rate analysis of high-order, full-tensor and sparse-tensor DG discretizations in the presence of geometric singularities due to corners and multimaterial interfaces, we require regularity results in corner-weighted spaces in  $\Omega$ , which we now recapitulate, based on [31, 35–38].

We collect the weight exponents  $\delta_i \in [0, 1)$  assigned to each  $\mathbf{c}_i \in \mathcal{S}$ , in  $\boldsymbol{\delta} = \{\delta_i\}_{i=1}^M \in [0, 1)^M$  (see Remarks 3.1, 3.2) and define the weight function  $\Phi_{\boldsymbol{\delta}}(\mathbf{x}) := \prod_{i=1}^M |\mathbf{x} - \mathbf{c}_i|^{\delta_i}$ . We express the regularity of the solution in the spatial domain  $\Omega$  in terms of weighted spaces  $H_{\boldsymbol{\delta}}^{k, \ell}(\Omega)$ , which are defined as the completions of  $C^\infty(\bar{\Omega})$  with respect to the weighted Sobolev norms  $\|\cdot\|_{H_{\boldsymbol{\delta}}^{k, \ell}(\Omega)}$  which, for  $1 \leq \ell \leq k \in \mathbb{N}$ , are given by

$$\|u\|_{H_{\boldsymbol{\delta}}^{k, \ell}(\Omega)}^2 := \|u\|_{H^{\ell-1}(\Omega)}^2 + |u|_{H_{\boldsymbol{\delta}}^{k, \ell}(\Omega)}^2, \quad |u|_{H_{\boldsymbol{\delta}}^{k, \ell}(\Omega)}^2 := \sum_{m=\ell}^k \int_{\Omega} \left( \Phi_{\boldsymbol{\delta}+m-\ell}^2 \sum_{\substack{\boldsymbol{\alpha} \in \mathbb{N}_0^2 \\ \alpha_1 + \alpha_2 = m}} |D^{\boldsymbol{\alpha}} u|^2 \right) \, dx.$$

We will use these spaces only for  $\ell = 1$  and  $\ell = 2$ ; in particular, we will mostly use the following norms and seminorms

$$\begin{aligned} \|u\|_{H_{\boldsymbol{\delta}}^{1, 1}(\Omega)}^2 &:= \|u\|_{L^2(\Omega)}^2 + |u|_{H_{\boldsymbol{\delta}}^{1, 1}(\Omega)}^2, & |u|_{H_{\boldsymbol{\delta}}^{1, 1}(\Omega)}^2 &:= \|\Phi_{\boldsymbol{\delta}} \nabla u\|_{L^2(\Omega)}^2, \\ \|u\|_{H_{\boldsymbol{\delta}}^{2, 2}(\Omega)}^2 &:= \|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2 + |u|_{H_{\boldsymbol{\delta}}^{2, 2}(\Omega)}^2, & |u|_{H_{\boldsymbol{\delta}}^{2, 2}(\Omega)}^2 &:= \|\Phi_{\boldsymbol{\delta}} D^2 u\|_{L^2(\Omega)^{2 \times 2}}^2, \end{aligned}$$

where  $D^2 u$  denotes the Hessian of  $u$ .

**Remark 3.1.** *For  $\delta_i = 0$ ,  $1 \leq i \leq M$ , we have  $\Phi_{\boldsymbol{\delta}} \equiv 1$ . In this case, some of the weighted seminorms reduce to standard ones, i.e.  $|\cdot|_{H_{\boldsymbol{\delta}}^{k, k}(\Omega)} = |\cdot|_{H^k(\Omega)}$  for  $\boldsymbol{\delta} = (0, \dots, 0)$  and  $k \in \mathbb{N}$ , and  $\|u\|_{H_{\boldsymbol{\delta}}^{2, 2}(\Omega)} = \|u\|_{H^2(\Omega)}$ .*

**Remark 3.2.** For a general polygon  $\Omega$ , admissible values of the parameters  $\delta_i \in [0, 1)$  depend on the coefficients of the elliptic spatial operator, on the boundary conditions at the sides of  $\Omega$  meeting at corner  $\mathbf{c}_i \in \mathcal{S}$  and on the interior opening angle  $\omega_i \in (0, 2\pi)$  at  $\mathbf{c}_i$ . For a homogeneous, isotropic material at a corner  $\mathbf{c}_i$ , with either homogeneous Dirichlet or homogeneous Neumann BCs at either side meeting at  $\mathbf{c}_i$ , we can take  $\delta_i > 1 - \pi/\omega_i$ . At convex corners  $\omega_i < \pi$ , so that  $\delta_i = 0$  is admissible. At corners  $\mathbf{c}_i$  where BCs change type  $\delta_i > 1 - \pi/(2\omega_i)$ . Similar conditions are valid for transmission problems with multi-material interface points  $\mathbf{c} \in \Omega$ . At such points, however,  $\delta_i$  may take values close to 1. We refer to [17, Section 3.2] for an example; a numerical example is provided in Section 8.4 ahead.

The regularity of the solution  $(v, \boldsymbol{\sigma})$  of the IBVP (1) follows from the regularity of the solution  $u$  of the second-order IBVP (2), by taking time and space derivatives. For example, if

$$v_0, u_0 \in C_0^\infty(\Omega), \quad \boldsymbol{\sigma}_0 = -\nabla u_0, \quad f \in C_0^\infty(Q), \quad g_D = g_N = 0,$$

then, by [35, Cor. 2.6.6] (see also [36, Prop. 2.2] for the case with constant  $c$ ), there exists  $\boldsymbol{\delta} \in [0, 1)^M$  such that, for all  $k_t, k_x \in \mathbb{N}$ , it holds that

$$(v, \boldsymbol{\sigma}) = \left( \frac{\partial u}{\partial t}, -\nabla u \right) \in C^{k_t-1}([0, T]; H_{\boldsymbol{\delta}}^{k_x+1, 2}(\Omega)) \times C^{k_t}([0, T]; H_{\boldsymbol{\delta}}^{k_x, 1}(\Omega)^2). \quad (8)$$

## 4 Space–time DG discretization: Notation

We prepare notation and conventions which shall be used in the ensuing space–time DG discretization.

### 4.1 Temporal, spatial, and space–time meshes

Let us introduce a partition  $\mathcal{T}_{h_t}^t$  of the time domain  $(0, T)$  into  $N \in \mathbb{N}$  intervals  $I_n$ ,  $1 \leq n \leq N$ , with

$$0 =: t_0 < t_1 < \dots, t_N := T, \quad I_n := (t_{n-1}, t_n), \quad h_n := |I_n|, \quad h_t := \max_{1 \leq n \leq N} h_n,$$

and introduce the following notation for the time slabs and the partial cylinders, respectively:

$$D_n := \Omega \times I_n, \quad Q_n := \Omega \times (0, t_n), \quad \text{for } n = 1, \dots, N.$$

For each  $1 \leq n \leq N$ , we introduce a polygonal finite element mesh  $\mathcal{T}_{h_{\mathbf{x}, n}}^{\mathbf{x}} = \{K_{\mathbf{x}}\}$  of the spatial domain  $\Omega$ , possibly with hanging nodes, with

$$h_{\mathbf{x}, n} := \max_{K_{\mathbf{x}} \in \mathcal{T}_{h_{\mathbf{x}, n}}^{\mathbf{x}}} h_{K_{\mathbf{x}}}, \quad h_{K_{\mathbf{x}}} := \text{diam } K_{\mathbf{x}}, \quad h_{\mathbf{x}} := \max_{1 \leq n \leq N} h_{\mathbf{x}, n}.$$

For each spatial mesh  $\mathcal{T}_{h_{\mathbf{x}, n}}^{\mathbf{x}}$ , we assume *i*) shape-regularity, *ii*) non-degeneracy of faces, namely all element- and face-sizes are locally comparable, *iii*) alignment with the fixed partition  $\{\Omega_i\}$  on which the wave speed  $c$  is piecewise constant, and *iv*) that for each  $K_{\mathbf{x}} \in \mathcal{T}_{h_{\mathbf{x}, n}}^{\mathbf{x}}$ ,  $\partial K_{\mathbf{x}}$  contains at most one element of  $\mathcal{S}$ , i.e. one vertex of the partition  $\{\Omega_i\}$ .

We partition the space–time domain  $Q = \Omega \times (0, T)$  with a finite element mesh  $\mathcal{T}_h$  given by

$$\mathcal{T}_h := \mathcal{T}_h(Q) := \{K = K_{\mathbf{x}} \times I_n : K_{\mathbf{x}} \in \mathcal{T}_{h_{\mathbf{x}, n}}^{\mathbf{x}}, 1 \leq n \leq N\}.$$

Note that  $\mathcal{T}_h$  is a tensor product mesh whenever  $\mathcal{T}_{h_{\mathbf{x}, n}}^{\mathbf{x}} = \mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}}$  for all  $1 \leq n \leq N$ , for a given spatial mesh  $\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}}$ , that is  $\mathcal{T}_h = \mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} \times \mathcal{T}_{h_t}^t$ .

As all spatial meshes  $\mathcal{T}_{h_{\mathbf{x}, n}}^{\mathbf{x}}$ ,  $1 \leq n \leq N$ , are aligned with the partition  $\{\Omega_i\}$ , then the wave speed  $c$  is constant in each element  $K$  of the space–time mesh  $\mathcal{T}_h$ , and we set  $c_K := c|_K$ . Moreover, if  $K = K_{\mathbf{x}} \times I_n$  is such that  $\partial K_{\mathbf{x}} \cap \mathcal{S} = \{\mathbf{c}_i\}$  is non-empty, we denote by  $\delta_K = \delta_{\mathbf{c}_i} \in [0, 1)$  the exponent of the space in (8).

For any  $1 \leq n \leq N$ , we define the time–truncated mesh

$$\mathcal{T}_h(Q_n) := \{K \in \mathcal{T}_h, K \subset Q_n\}.$$

We partition  $\mathcal{T}_h$  into elements abutting at a corner in  $\mathcal{S}$  and the remaining elements, i.e.

$$\mathcal{T}_h^{\angle} := \{K = K_{\mathbf{x}} \times I \in \mathcal{T}_h, \overline{K_{\mathbf{x}}} \cap \mathcal{S} \text{ nonempty}\}, \quad \mathcal{T}_h^{\circ} := \mathcal{T}_h \setminus \mathcal{T}_h^{\angle}.$$



## 4.2 Mesh faces

Each internal face  $F = \partial K_1 \cap \partial K_2$ , for  $K_1, K_2 \in \mathcal{T}_h$ , with positive 2-dimensional measure, is a subset of a hyperplane:

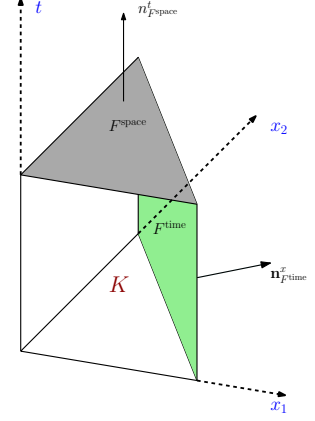
$$F \subset \Pi_F := \{(\mathbf{x}, t) \in \mathbb{R}^{2+1} : \mathbf{x} \cdot \mathbf{n}_F^x + t n_F^t = a_F\},$$

where  $(\mathbf{n}_F^x, n_F^t)$  is a unit vector in  $\mathbb{R}^{2+1}$  and  $a_F \in \mathbb{R}$ . We assume that all internal faces  $F$  are either “space-like”, i.e. with  $\mathbf{n}_F^x = \mathbf{0}$ , or “time-like”, i.e. with  $n_F^t = 0$ . On space-like faces, by convention, we choose  $n_F^t > 0$ , i.e. the unit normal vector  $(\mathbf{n}_F^x, n_F^t)$  points forward in time. Note that all time-like faces are rectangles of the form  $F = F_x \times F_t$  with  $h_{F_x} = |F_x|$  and  $h_{F_t} = |F_t|$ ; we recall that  $F_t = I_n$  for some  $1 \leq n \leq N$ . Finally, we denote the outward-pointing unit normal vector on  $\partial K$  by  $(\mathbf{n}_K^x, n_K^t)$ .

We use the following notation for unions and sets of faces:

$$\begin{aligned} \mathcal{F}_h &:= \bigcup_{K \in \mathcal{T}_h} \partial K \quad (\text{the mesh skeleton}), \\ \mathcal{F}_h^{\text{space}} &:= \text{the union of the internal space-like faces}, \\ \mathcal{F}_h^{\text{time}} &:= \text{the union of the internal time-like faces}, \\ \mathcal{F}_h^0 &:= \Omega \times \{0\}, & \mathcal{F}_h^T &:= \Omega \times \{T\}, \\ \mathcal{F}_h^D &:= \Gamma_D \times [0, T], & \mathcal{F}_h^N &:= \Gamma_N \times [0, T], \\ \mathbb{F} &:= \{\text{faces of elements of } \mathcal{T}_h\}, \\ \mathbb{F}^\star &:= \{F \in \mathbb{F}, F \subset \mathcal{F}_h^\star\} && \text{for } \star \in \{\text{time, D, N}\}, \\ \mathbb{F}_\angle^\star &:= \{F \in \mathbb{F}^\star, F \subset \partial K \text{ for } K \in \mathcal{T}_h^\angle\} && \text{for } \star \in \{\text{time, D, N}\}, \\ \mathbb{F}_\circ^\star &:= \mathbb{F}^\star \setminus \mathbb{F}_\angle^\star && \text{for } \star \in \{\text{time, D, N}\}, \\ \mathcal{F}_h^\angle &:= \text{the union of the faces } F \in \mathbb{F}_\angle^{\text{time}} \cup \mathbb{F}_\angle^D \cup \mathbb{F}_\angle^N, \\ \mathcal{F}_h^\circ &:= \text{the union of the faces } F \in \mathbb{F}_\circ^{\text{time}} \cup \mathbb{F}_\circ^D \cup \mathbb{F}_\circ^N, \\ \mathbb{F}_\bullet^\star(\Upsilon) &:= \{F \in \mathbb{F}_\bullet^\star : F \subset \overline{\Upsilon}\}, \end{aligned}$$

for all subsets  $\Upsilon \subset Q$  and for all indices  $\star, \bullet$  allowed by the previous definitions; in particular we will use  $\mathbb{F}_\bullet^\star(Q_n)$  and  $\mathbb{F}_\bullet^\star(K)$ ,  $K \in \mathcal{T}_h$ . The wave speed  $c$  is assumed independent of time, so that it may jump only across time-like faces in  $\mathcal{F}_h^{\text{time}}$ .



Space-time prism element  $K$ , with time-like face  $F^{\text{time}}$  and space-like face  $F^{\text{space}}$ .

## 4.3 Averages and jumps

In the formulation and error analysis of the DG discretization, we require the “usual” notation for interelement averages and jumps. We follow the notation in [33] in order to refer to results on space-time DG formulations which were established there.

For piecewise-continuous scalar ( $w$ ) and vector ( $\boldsymbol{\tau}$ ) fields, we define averages  $\{\!\{ \cdot \}\!\}$ , space normal jumps  $\llbracket \cdot \rrbracket_{\mathbf{N}}$ , and time (full) jumps  $\llbracket \cdot \rrbracket_t$  on internal mesh faces in the standard DG notation: on  $F = \partial K_1 \cap \partial K_2$ ,  $K_1, K_2 \in \mathcal{T}_h$ ,

$$\begin{aligned} \{\!\{w\}\!\} &:= \frac{w|_{K_1} + w|_{K_2}}{2}, & \{\!\{\boldsymbol{\tau}\}\!\} &:= \frac{\boldsymbol{\tau}|_{K_1} + \boldsymbol{\tau}|_{K_2}}{2}, \\ \llbracket w \rrbracket_{\mathbf{N}} &:= w|_{K_1} \mathbf{n}_{K_1}^x + w|_{K_2} \mathbf{n}_{K_2}^x, & \llbracket \boldsymbol{\tau} \rrbracket_{\mathbf{N}} &:= \boldsymbol{\tau}|_{K_1} \cdot \mathbf{n}_{K_1}^x + \boldsymbol{\tau}|_{K_2} \cdot \mathbf{n}_{K_2}^x, \\ \llbracket w \rrbracket_t &:= w|_{K_1} n_{K_1}^t + w|_{K_2} n_{K_2}^t = (w^- - w^+) n_F^t, & \llbracket \boldsymbol{\tau} \rrbracket_t &:= \boldsymbol{\tau}|_{K_1} n_{K_1}^t + \boldsymbol{\tau}|_{K_2} n_{K_2}^t = (\boldsymbol{\tau}^- - \boldsymbol{\tau}^+) n_F^t. \end{aligned}$$

Here we have denoted by  $w^-$  and  $w^+$  the traces of the function  $w$  from the adjacent elements at lower and higher times, respectively (and similarly for  $\boldsymbol{\tau}^\pm$ ). The temporal jumps  $\llbracket \cdot \rrbracket_t$  are different from zero on space-like faces only.

## 4.4 Some properties of $H_\delta^{1,1}(\Omega)$ functions

We state in our notation some results from [45] that are needed in the error analysis of the space-time discretization.

**Lemma 4.1.** *If  $u \in H_\delta^{1,1}(\Omega)$  then*

$$u \in L^1(F) \quad \text{and} \quad \llbracket u \rrbracket_{\mathbf{N}} = 0 \quad \text{a.e. on } F \quad (9)$$

where  $F$  is any segment in  $\Omega$ . In general,  $u \in H_\delta^{1,1}(\Omega)$  does not guarantee that  $u \in L^2(F)$ .

Moreover, the embedding  $H_{\delta}^{1,1}(K_{\mathbf{x}}) \subset L^2(K_{\mathbf{x}})$  is compact:

$$H_{\delta}^{1,1}(K_{\mathbf{x}}) \subset\subset L^2(K_{\mathbf{x}}) \quad (10)$$

*Proof.* See [45, Lemma 1.3.2 and Remark 1.3.3], [45, Lemma 1.3.4], and [45, Proposition A.2.5], respectively.  $\square$

If the IBVP solution  $(v, \boldsymbol{\sigma})$  satisfies the regularity condition (8), then  $\boldsymbol{\sigma}$  has poor regularity in space, so we expect its traces on mesh edges close to the domain corners at a fixed time to be in  $L^1$  but not in  $L^2$ . This is relevant, e.g., to define the DG skeleton seminorms in §5.3. On the other hand, in the analysis we make use of the vanishing of the jumps of  $\boldsymbol{\sigma}$  on these edges.

## 4.5 Inverse inequalities

As the ensuing error analysis will be an “ $h$ -version” analysis, where the polynomial degrees are assumed fixed and convergence is achieved by suitable mesh refinement, we will require inverse inequalities in order to provide bounds on some face terms.

For any time-like face  $F = F_{\mathbf{x}} \times F_t \in \mathbb{F}^{\text{time}} \cup \mathbb{F}^{\text{D}}$ ,

$$\|\varphi\|_{L^1(F_{\mathbf{x}})} \leq h_{F_{\mathbf{x}}}^{1/2} \|\varphi\|_{L^2(F_{\mathbf{x}})} \quad \forall \varphi \in L^2(F_{\mathbf{x}}), \quad \|\varphi\|_{L^2(F_{\mathbf{x}})} \leq h_{F_{\mathbf{x}}}^{1/2} \|\varphi\|_{L^\infty(F_{\mathbf{x}})} \quad \forall \varphi \in L^\infty(F_{\mathbf{x}}). \quad (11)$$

For polynomial functions, the following inverse inequalities hold:

$$\|P\|_{L^2(F_{\mathbf{x}})} \leq C_{2|1} h_{F_{\mathbf{x}}}^{-1/2} \|P\|_{L^1(F_{\mathbf{x}})}, \quad \|P\|_{L^\infty(F_{\mathbf{x}})} \leq C_{\infty|2} h_{F_{\mathbf{x}}}^{-1/2} \|P\|_{L^2(F_{\mathbf{x}})} \quad \forall P \in \mathbb{P}^p(F_{\mathbf{x}}), \quad (12)$$

where  $\mathbb{P}^p(F_{\mathbf{x}})$  denotes the space of polynomials of degree at most  $p \in \mathbb{N}_0$  on  $F_{\mathbf{x}}$ , and the constants  $C_{2|1}, C_{\infty|2}$  depend on  $p$ . The inequalities in (11) are obvious, while the inverse inequalities in (12) follow from the equivalence of norms in finite dimensional spaces and scaling; note that  $C_{\infty|2} \geq 1$ . The constants in (12) depend linearly on the polynomial degree. Denote by  $P_p$  the Legendre polynomials and by  $Q_p = \sqrt{p + \frac{1}{2}} P_p$  their scalings which are *orthonormal* in  $L^2(-1, 1)$  and which satisfy  $\|Q_p\|_{L^\infty(-1,1)} = \sqrt{p + \frac{1}{2}}$ . Then, for any polynomial  $v = \sum_{j=0}^p a_j Q_j \in \mathbb{P}^p(-1, 1)$ ,

$$\begin{aligned} \|v\|_{L^\infty(-1,1)} &\leq \sum_{j=0}^p |a_j| \|Q_j\|_{L^\infty(-1,1)} \leq \left( \sum_{j=0}^p |a_j|^2 \right)^{\frac{1}{2}} \left( \sum_{j=0}^p (j + 1/2) \right)^{\frac{1}{2}} = \|v\|_{L^2(-1,1)} \frac{p+1}{\sqrt{2}}, \\ \|v\|_{L^2(-1,1)}^2 &= \sum_{j=0}^p |a_j|^2 = \sum_{j=0}^p \left( \int_{-1}^1 v Q_j \right)^2 \leq \|v\|_{L^1(-1,1)}^2 \sum_{j=0}^p \|Q_j\|_{L^\infty(-1,1)}^2 = \|v\|_{L^1(-1,1)}^2 \frac{(p+1)^2}{2}. \end{aligned}$$

By scaling from  $(-1, 1)$  to a general interval we obtain (12) with

$$C_{2|1} \leq p + 1, \quad C_{\infty|2} \leq p + 1 \quad p \in \mathbb{N}_0. \quad (13)$$

The inverse inequalities (12) involve spaces of single-variable polynomials defined on segments; in the error analysis carried out in the next sections, we do not use inverse inequalities for spaces defined on polygons or polyhedra, so the shapes of the mesh elements do not play any role in this respect.

## 5 Space–time DG discretization: Formulation

With the notation introduced in the preceding section, we now derive the space–time DG discretization of the IBVP (1). We prove existence and uniqueness of discrete solutions, as well as abstract error estimates for these.

Throughout this section we assume the following regularity for the data and for the solutions  $(v, \boldsymbol{\sigma})$  of the IBVP (1):

$$\begin{aligned} f &\in L^2(Q), \quad v_0 \in H_{\delta}^{2,2}(\Omega), \quad \boldsymbol{\sigma}_0 \in H_{\delta}^{1,1}(\Omega)^2, \quad g_{\text{D}} \in L^2(0, T; H^{\frac{1}{2}}(\Gamma_{\text{D}})), \quad g_{\text{N}} \in L^2(0, T; L^2(\Gamma_{\text{N}})), \\ (v, \boldsymbol{\sigma}) &\in H^1([0, T]; H_{\delta}^{2,2}(\Omega)) \times H^1([0, T]; H_{\delta}^{1,1}(\Omega)^2). \end{aligned}$$

## 5.1 Derivation of the space–time discrete formulation

We rewrite the formulation of [33] in the non-Trefftz case. This is a special case of the formulation of [34] for general hyperbolic first-order systems.

To derive the DG formulation, we multiply the firstly two equations of (1) with test fields  $\boldsymbol{\tau}$  and  $w$  and integrate by parts on a single mesh element  $K \in \mathcal{T}_h$ :

$$\begin{aligned} & - \int_K \left( v \left( \nabla \cdot \boldsymbol{\tau} + c^{-2} \frac{\partial w}{\partial t} \right) + \boldsymbol{\sigma} \cdot \left( \nabla w + \frac{\partial \boldsymbol{\tau}}{\partial t} \right) \right) dV \\ & + \int_{\partial K} \left( (v \boldsymbol{\tau} + \boldsymbol{\sigma} w) \cdot \mathbf{n}_K^x + \left( \boldsymbol{\sigma} \cdot \boldsymbol{\tau} + c^{-2} v w \right) n_K^t \right) dS = \int_K f w dV; \end{aligned} \quad (14)$$

we recall that the wave speed  $c$  is assumed to be constant in each  $K \in \mathcal{T}_h$ . We seek a discrete solution  $(v_h, \boldsymbol{\sigma}_h)$  approximating  $(v, \boldsymbol{\sigma})$  in a finite-dimensional (arbitrary, at this stage) space  $\mathbf{V}_p(\mathcal{T}_h)$ . We take the test field  $(w, \boldsymbol{\tau})$  in the same space  $\mathbf{V}_p(\mathcal{T}_h)$ . The traces of  $v_h$  and  $\boldsymbol{\sigma}_h$  on the mesh skeleton are approximated by the (single-valued) *numerical fluxes*  $\hat{v}_h$  and  $\hat{\boldsymbol{\sigma}}_h$ , so that (14) is rewritten as:

$$\begin{aligned} & - \int_K \left( v_h \left( \nabla \cdot \boldsymbol{\tau}_h + c^{-2} \frac{\partial w_h}{\partial t} \right) + \boldsymbol{\sigma}_h \cdot \left( \nabla w_h + \frac{\partial \boldsymbol{\tau}_h}{\partial t} \right) \right) dV \\ & + \int_{\partial K} \left( \hat{v}_h \left( \boldsymbol{\tau}_h \cdot \mathbf{n}_K^x + \frac{w_h}{c^2} n_K^t \right) + \hat{\boldsymbol{\sigma}}_h \cdot \left( w_h \mathbf{n}_K^x + \boldsymbol{\tau}_h n_K^t \right) \right) dS = \int_K f w dV. \end{aligned} \quad (15)$$

We choose to define the numerical fluxes as:

$$\hat{v}_h := \begin{cases} v_h^- & \text{on } \mathcal{F}_h^{\text{space}}, \\ v_h & \text{on } \mathcal{F}_h^T, \\ v_0 & \text{on } \mathcal{F}_h^0, \\ \{\!\{v_h}\!\} + \beta[\![\boldsymbol{\sigma}_h]\!]_{\mathbf{N}} & \text{on } \mathcal{F}_h^{\text{time}}, \\ g_{\text{D}} & \text{on } \mathcal{F}_h^{\text{D}}, \\ v_h + \beta(\boldsymbol{\sigma}_h \cdot \mathbf{n}_{\Omega}^x - g_{\text{N}}) & \text{on } \mathcal{F}_h^{\text{N}}. \end{cases} \quad \hat{\boldsymbol{\sigma}}_h := \begin{cases} \boldsymbol{\sigma}_h^- & \text{on } \mathcal{F}_h^{\text{space}}, \\ \boldsymbol{\sigma}_h & \text{on } \mathcal{F}_h^T, \\ \boldsymbol{\sigma}_0 & \text{on } \mathcal{F}_h^0, \\ \{\!\{\boldsymbol{\sigma}_h}\!\} + \alpha[\![v_h]\!]_{\mathbf{N}} & \text{on } \mathcal{F}_h^{\text{time}}, \\ \boldsymbol{\sigma}_h + \alpha(v_h - g_{\text{D}}) \mathbf{n}_{\Omega}^x & \text{on } \mathcal{F}_h^{\text{D}}, \\ g_{\text{N}} \mathbf{n}_{\Omega}^x & \text{on } \mathcal{F}_h^{\text{N}}. \end{cases}$$

The stabilization parameters  $\alpha \in L^\infty(\mathcal{F}_h^{\text{time}} \cup \mathcal{F}_h^{\text{D}})$  and  $\beta \in L^\infty(\mathcal{F}_h^{\text{time}} \cup \mathcal{F}_h^{\text{N}})$  will be chosen depending on the mesh, but constant on each time-like face; see Corollary 6.6 and §7 below.

These fluxes are *consistent*, in the sense that they coincide with the traces of the exact solution  $(v, \boldsymbol{\sigma})$  of the IBVP (1) if they are applied to  $(v, \boldsymbol{\sigma})$  itself, which satisfies the boundary conditions and has no jumps across mesh faces by the  $C^0([0, T]; H_\delta^{1,1}(\Omega))$  regularity of all components of the solution and the property (9) (see [45, Lemma 1.3.4]). The numerical fluxes can be understood as upwind fluxes in the usual sense on the space-like faces, and as standard DG-elliptic fluxes with jump penalisation on the time-like ones.

By summing the elemental DG equation (15) over the elements  $K \in \mathcal{T}_h$ , with the fluxes defined above, we obtain the space–time DG variational formulation:

$$\begin{aligned} & \text{Find } (v_h, \boldsymbol{\sigma}_h) \in \mathbf{V}_p(\mathcal{T}_h) \text{ such that } \mathcal{A}_{\text{DG}}(v_h, \boldsymbol{\sigma}_h; w, \boldsymbol{\tau}) = \ell(w, \boldsymbol{\tau}) \quad \forall (w, \boldsymbol{\tau}) \in \mathbf{V}_p(\mathcal{T}_h), \text{ where} \\ \mathcal{A}_{\text{DG}}(v, \boldsymbol{\sigma}; w, \boldsymbol{\tau}) & := - \sum_{K \in \mathcal{T}_h} \int_K \left( v \left( \nabla \cdot \boldsymbol{\tau} + c^{-2} \frac{\partial w}{\partial t} \right) + \boldsymbol{\sigma} \cdot \left( \nabla w + \frac{\partial \boldsymbol{\tau}}{\partial t} \right) \right) dV \\ & + \int_{\mathcal{F}_h^{\text{space}}} (c^{-2} v^- [w]_t + \boldsymbol{\sigma}^- \cdot [\boldsymbol{\tau}]_t) dx + \int_{\mathcal{F}_h^T} (c^{-2} v w + \boldsymbol{\sigma} \cdot \boldsymbol{\tau}) dx \\ & + \int_{\mathcal{F}_h^{\text{time}}} (\{\!\{v}\!\} [\boldsymbol{\tau}]_{\mathbf{N}} + \{\!\{\boldsymbol{\sigma}\}\!\} \cdot [w]_{\mathbf{N}} + \alpha [v]_{\mathbf{N}} \cdot [w]_{\mathbf{N}} + \beta [\boldsymbol{\sigma}]_{\mathbf{N}} [\boldsymbol{\tau}]_{\mathbf{N}}) dS \\ & + \int_{\mathcal{F}_h^{\text{D}}} (\boldsymbol{\sigma} \cdot \mathbf{n}_{\Omega}^x w + \alpha v w) dS + \int_{\mathcal{F}_h^{\text{N}}} (v (\boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x) + \beta (\boldsymbol{\sigma} \cdot \mathbf{n}_{\Omega}^x) (\boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x)) dS, \\ \ell(w, \boldsymbol{\tau}) & := \int_Q f w dV + \int_{\mathcal{F}_h^0} (c^{-2} v_0 w + \boldsymbol{\sigma}_0 \cdot \boldsymbol{\tau}) dx \\ & + \int_{\mathcal{F}_h^{\text{D}}} g_{\text{D}} (\alpha w - \boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x) dS + \int_{\mathcal{F}_h^{\text{N}}} g_{\text{N}} (\beta \boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x - w) dS. \end{aligned} \quad (16)$$

## 5.2 Existence and uniqueness of the discrete solution

We prove existence and uniqueness of the DG solution under the assumption that all  $(w_h, \boldsymbol{\tau}_h) \in \mathbf{V}_\mathbf{p}(\mathcal{T}_h)$  are elementwise polynomials (at this stage elementwise analytic suffices), and that we have

$$\left( \nabla \cdot \boldsymbol{\tau}_h + c^{-2} \frac{\partial w_h}{\partial t}, \nabla w_h + \frac{\partial \boldsymbol{\tau}_h}{\partial t} \right) \in \mathbf{V}_\mathbf{p}(\mathcal{T}_h) \quad \forall (w_h, \boldsymbol{\tau}_h) \in \mathbf{V}_\mathbf{p}(\mathcal{T}_h); \quad (17)$$

recall that  $c$  is piecewise constant in  $\mathcal{T}_h$ . We define the following DG seminorm:

$$\begin{aligned} |(w, \boldsymbol{\tau})|_{\text{DG}}^2 := & \frac{1}{2} \|c^{-1} \llbracket w \rrbracket_t\|_{L^2(\mathcal{F}_h^{\text{space}})}^2 + \frac{1}{2} \|\llbracket \boldsymbol{\tau} \rrbracket_t\|_{L^2(\mathcal{F}_h^{\text{space}})}^2 + \frac{1}{2} \|c^{-1} w\|_{L^2(\mathcal{F}_h^0 \cup \mathcal{F}_h^T)}^2 + \frac{1}{2} \|\boldsymbol{\tau}\|_{L^2(\mathcal{F}_h^0 \cup \mathcal{F}_h^T)}^2 \\ & + \|\alpha^{1/2} \llbracket w \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\text{time}})}^2 + \|\beta^{1/2} \llbracket \boldsymbol{\tau} \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\text{time}})}^2 + \|\alpha^{1/2} w\|_{L^2(\mathcal{F}_h^{\text{D}})}^2 + \|\beta^{1/2} \boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x\|_{L^2(\mathcal{F}_h^{\text{N}})}^2. \end{aligned} \quad (18)$$

**Proposition 5.1.** *Under assumption (17), the solution of the space–time DG formulation (16) exists and is unique.*

*Proof.* We specialise the proof of [34, Lemma 4.4] to our setting. We start by rewriting  $\mathcal{A}_{\text{DG}}$  in (16) in the following equivalent form, which is obtained by integration by parts in each element:

$$\begin{aligned} \mathcal{A}_{\text{DG}}(v, \boldsymbol{\sigma}; w, \boldsymbol{\tau}) = & \sum_{K \in \mathcal{T}_h} \int_K \left( \left( \nabla \cdot \boldsymbol{\sigma} + c^{-2} \frac{\partial v}{\partial t} \right) w + \left( \nabla v + \frac{\partial \boldsymbol{\sigma}}{\partial t} \right) \cdot \boldsymbol{\tau} \right) dV \\ & - \int_{\mathcal{F}_h^{\text{space}}} (c^{-2} \llbracket v \rrbracket_t w^+ + \llbracket \boldsymbol{\sigma} \rrbracket_t \cdot \boldsymbol{\tau}^+) d\mathbf{x} + \int_{\mathcal{F}_h^0} (c^{-2} v w + \boldsymbol{\sigma} \cdot \boldsymbol{\tau}) d\mathbf{x} \\ & + \int_{\mathcal{F}_h^{\text{time}}} (-\llbracket v \rrbracket_{\mathbf{N}} \cdot \llbracket \boldsymbol{\tau} \rrbracket - \llbracket \boldsymbol{\sigma} \rrbracket_{\mathbf{N}} \llbracket w \rrbracket + \alpha \llbracket v \rrbracket_{\mathbf{N}} \cdot \llbracket w \rrbracket_{\mathbf{N}} + \beta \llbracket \boldsymbol{\sigma} \rrbracket_{\mathbf{N}} \llbracket \boldsymbol{\tau} \rrbracket_{\mathbf{N}}) dS \\ & + \int_{\mathcal{F}_h^{\text{D}}} (-v \boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x + \alpha v w) dS + \int_{\mathcal{F}_h^{\text{N}}} (-\boldsymbol{\sigma} \cdot \mathbf{n}_\Omega^x w + \beta (\boldsymbol{\sigma} \cdot \mathbf{n}_\Omega^x) (\boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x)) dS. \end{aligned} \quad (19)$$

By taking  $w = v$  and  $\boldsymbol{\tau} = \boldsymbol{\sigma}$ , and summing the two expressions for  $\mathcal{A}_{\text{DG}}$  given in (16) and (19), we obtain

$$\mathcal{A}_{\text{DG}}(v, \boldsymbol{\sigma}; v, \boldsymbol{\sigma}) = |(v, \boldsymbol{\sigma})|_{\text{DG}}^2, \quad (20)$$

for the DG seminorm defined in (18).

In order to prove uniqueness of the discrete solution of (16), due to the linearity of the problem, it is enough to consider the data  $f = u_0 = u_1 = g_{\text{D}} = g_{\text{N}} = 0$  so that  $\ell(w_h, \boldsymbol{\tau}_h) = 0$  for all  $(w_h, \boldsymbol{\tau}_h) \in \mathbf{V}_\mathbf{p}(\mathcal{T}_h)$ , and prove that  $(v_h, \boldsymbol{\sigma}_h)$  is zero in  $Q$ .

Choosing  $(w_h, \boldsymbol{\tau}_h) = (v_h, \boldsymbol{\sigma}_h)$ , from (20) we deduce  $|(v_h, \boldsymbol{\sigma}_h)|_{\text{DG}} = 0$ , i.e. all jumps and traces present in the definition of the DG seminorm vanish. Then (19) becomes

$$\begin{aligned} \mathcal{A}_{\text{DG}}(v_h, \boldsymbol{\sigma}_h; w_h, \boldsymbol{\tau}_h) = & \sum_{K \in \mathcal{T}_h} \int_K \left( \left( \nabla \cdot \boldsymbol{\sigma}_h + c^{-2} \frac{\partial v_h}{\partial t} \right) w_h + \left( \nabla v_h + \frac{\partial \boldsymbol{\sigma}_h}{\partial t} \right) \cdot \boldsymbol{\tau}_h \right) dV = 0 \\ & \forall (w_h, \boldsymbol{\tau}_h) \in \mathbf{V}_\mathbf{p}(\mathcal{T}_h). \end{aligned}$$

Choosing  $(w_h, \boldsymbol{\tau}_h) = (\nabla \cdot \boldsymbol{\sigma}_h + c^{-2} \frac{\partial v_h}{\partial t}, \nabla v_h + \frac{\partial \boldsymbol{\sigma}_h}{\partial t}) \in \mathbf{V}_\mathbf{p}(\mathcal{T}_h)$ , which is possible by (17), we deduce that for all  $K \in \mathcal{T}_h$

$$\nabla \cdot \boldsymbol{\sigma}_h + c^{-2} \frac{\partial v_h}{\partial t} = 0, \quad \nabla v_h + \frac{\partial \boldsymbol{\sigma}_h}{\partial t} = \mathbf{0}, \quad (21)$$

i.e.  $(v_h, \boldsymbol{\sigma}_h)$  is solution of the homogeneous PDE in each element.

In each  $K \subset D_1$ , the first time-slab,  $(v_h, \boldsymbol{\sigma}_h)$  is a polynomial solution (or an analytical solution) of (21) with homogeneous initial conditions, so it vanishes in  $K$ . Iterating over the time slabs and using that the jumps across space-like faces vanish, we deduce that  $(v_h, \boldsymbol{\sigma}_h)$  is zero in the whole space–time cylinder  $Q$ .

Existence of the solution follows from linearity and finite dimensionality.  $\square$

If the wave speed  $c$  were a general (piecewise) smooth function of  $\mathbf{x}$ , the condition (17) would not be satisfied. Condition (17) is used in the proof of Proposition 5.1 to show (21), i.e. to ensure that the solution of the homogeneous variational problem are elementwise solution of the PDE. This prevents the immediate extension of the current analysis to the case of general coefficients.

**Remark 5.2.** Whenever the problem is driven by initial conditions only (namely,  $f = 0$  and the boundary conditions on  $\mathcal{F}_h^D$  and  $\mathcal{F}_h^N$ , if non empty, are zero), the solution  $(v_h, \boldsymbol{\sigma}_h)$  of the space–time DG formulation (16) satisfies

$$|(v_h, \boldsymbol{\sigma}_h)|_{\text{DG}}^2 = \mathcal{A}_{\text{DG}}(v_h, \boldsymbol{\sigma}_h; v_h, \boldsymbol{\sigma}_h) = \ell(v_h, \boldsymbol{\sigma}_h) \leq \left( 2 \|c^{-1}v_0\|_{L^2(\mathcal{F}_h^0)}^2 + 2 \|\boldsymbol{\sigma}_0\|_{L^2(\mathcal{F}_h^0)^2}^2 \right)^{1/2} |(v_h, \boldsymbol{\sigma}_h)|_{\text{DG}},$$

and therefore

$$|(v_h, \boldsymbol{\sigma}_h)|_{\text{DG}} \leq \left( 2 \|c^{-1}v_0\|_{L^2(\mathcal{F}_h^0)}^2 + 2 \|\boldsymbol{\sigma}_0\|_{L^2(\mathcal{F}_h^0)^2}^2 \right)^{1/2}.$$

**Remark 5.3.** We stress that  $|\cdot|_{\text{DG}}$  defined in (18) is only a seminorm in usual broken Sobolev spaces and DG discrete spaces. In Trefftz discretization spaces,  $|\cdot|_{\text{DG}}$  is actually a norm. Therefore, in the analysis of the Trefftz-DG method in [33], well-posedness and quasi-optimality were straightforward consequences of the Lax–Milgram theorem.

### 5.3 Abstract error estimates at discrete times

We firstly prove error bounds in the seminorm  $|\cdot|_{\text{DG}}$  restricted to the partial cylinder  $Q_n$ . We denote this truncated seminorm  $|(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)} := |(w\chi_{Q_n}, \boldsymbol{\tau}\chi_{Q_n})|_{\text{DG}}$ , where  $\chi_{Q_n}$  is the characteristic function of  $Q_n$ . Similarly we define the truncated bilinear form  $\mathcal{A}_{\text{DG}(Q_n)}(v, \boldsymbol{\sigma}; w, \boldsymbol{\tau}) := \mathcal{A}(v\chi_{Q_n}, \boldsymbol{\sigma}\chi_{Q_n}; w\chi_{Q_n}, \boldsymbol{\tau}\chi_{Q_n})$ .

As  $|(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)}^2$  contains the terms  $\|c^{-1}w\|_{L^2(\Omega \times \{t_n\})}^2$  and  $\|\boldsymbol{\tau}\|_{L^2(\Omega \times \{t_n\})^2}^2$ , error bounds in the seminorm  $|\cdot|_{\text{DG}(Q_n)}$  imply bounds on the  $L^2$  norm of the error of the trace (as opposed to the time-jumps only) on the space-like interfaces  $\mathcal{F}_h^{t_n} := \Omega \times \{t_n\}$ , namely a control of spatial volume integrals of the error in  $v$  and in  $\boldsymbol{\sigma}$  at each discrete time  $t_n$ .

Let  $(v, \boldsymbol{\sigma}), (v_h, \boldsymbol{\sigma}_h), (w_h, \boldsymbol{\tau}_h)$  be the continuous solution, the discrete solution, and an arbitrary discrete test field, respectively. By (20) and Galerkin orthogonality

$$\begin{aligned} |(w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)}^2 &= \mathcal{A}_{\text{DG}(Q_n)}((w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h); (w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)) \\ &= \mathcal{A}_{\text{DG}(Q_n)}((w_h, \boldsymbol{\tau}_h) - (v, \boldsymbol{\sigma}); (w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)). \end{aligned} \quad (22)$$

We now want to prove that the right-hand side is bounded by  $|(w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)}$ .

If we choose  $(w_h, \boldsymbol{\tau}_h) \in \mathbf{V}_p$  so that

$$\int_K \left( (w_h - v) \left( \nabla \cdot \boldsymbol{\tau} + c^{-2} \frac{\partial w}{\partial t} \right) + (\boldsymbol{\tau}_h - \boldsymbol{\sigma}) \cdot \left( \nabla w + \frac{\partial \boldsymbol{\tau}}{\partial t} \right) \right) dV = 0 \quad \forall (w, \boldsymbol{\tau}) \in \mathbf{V}_p(K), \forall K \subset Q_n, \quad (23)$$

the volume terms in  $\mathcal{A}_{\text{DG}(Q_n)}((w_h, \boldsymbol{\tau}_h) - (v, \boldsymbol{\sigma}); (w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h))$  expressed as in (16) vanish. If assumption (17) is satisfied, taking  $(w_h, \boldsymbol{\tau}_h)$  as the orthogonal  $L^2$  projection of  $(v, \boldsymbol{\sigma})$  on  $\mathbf{V}_p(K)$  ensures (23).

**Remark 5.4.** A straightforward application of the analysis of [34, §5] is not possible here. In [34], total degree space–time polynomial spaces  $\mathbb{P}^p(K)$  were used, and the first-order wave operator maps  $\mathbb{P}^p(K)^{d+1}$  into  $\mathbb{P}^{p-1}(K)^{d+1}$ . Then, a projection that is orthogonal to  $\mathbb{P}^{p-1}(K)$  and  $\mathbb{P}^p(F)$ , where  $F$  is a face of  $K$ , was employed in order to define  $(w_h, \boldsymbol{\tau}_h)$  satisfying (23). In our case, where space–time tensor product polynomials are used, the first-order wave operator does not map to lower-degree polynomial spaces. Consequently, the  $L^2$  projection into the orthogonal to the image space does not “save” sufficiently many degrees of freedom to be able to use them to cancel some trace. Therefore, we define  $(w_h, \boldsymbol{\tau}_h)$  by projecting against the whole discrete space, and we exploit the coercivity in seminorm property in the first line of (22), which was not used in [34]; see §5.3.1 and §5.3.2 below.

For the sake of clarity, we firstly deduce a quasi-optimality bound under the assumption that  $(v, \boldsymbol{\sigma})$  does not present corner singularities (§5.3.1), and then in the more realistic case involving weighted spaces (§5.3.2). These “quasi-optimality” bounds (i.e. (26) and (30)) differ from classical ones (from Céa lemma) in that they allow to control the Galerkin error in terms of the  $L^2$ -projection error, as opposed to the best approximation error in the energy norm.

#### 5.3.1 The smooth solution case

If the IBVP data are artificially chosen so that the solution  $(v, \boldsymbol{\sigma}) \in C^0([0, T]; H^1(\Omega))^3$ , then all traces on the mesh faces possess  $L^2$  summability. Therefore, if  $(w_h, \boldsymbol{\tau}_h)$  satisfies (23) (e.g. because assumption (17)

is verified and  $(w_h, \boldsymbol{\tau}_h)$  is the orthogonal  $L^2$  projection of  $(v, \boldsymbol{\sigma})$  on  $\mathbf{V}_\mathbf{p}(\mathcal{T}_h)$ , by using the Cauchy-Schwarz inequality, one can see that the bilinear form in (16) admits the following upper bound:

$$\begin{aligned} \mathcal{A}_{\text{DG}(Q_n)}((w_h, \boldsymbol{\tau}_h) - (v, \boldsymbol{\sigma}); (w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)) \\ \leq 2 |(w_h, \boldsymbol{\tau}_h) - (v, \boldsymbol{\sigma})|_{\text{DG}(Q_n)^+} |(w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)}, \end{aligned} \quad (24)$$

where the  $|\cdot|_{\text{DG}(Q_n)^+}$  seminorm is defined by

$$\begin{aligned} |(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)^+}^2 &:= |(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)}^2 + 2 \|c^{-1}w^-\|_{L^2(\mathcal{F}_h^{\text{space}} \cap Q_n)}^2 + 2 \|\boldsymbol{\tau}^-\|_{L^2(\mathcal{F}_h^{\text{space}} \cap Q_n)}^2 \\ &\quad + \|\beta^{-1/2} \llbracket w \rrbracket\|_{L^2(\mathcal{F}_h^{\text{time}} \cap Q_n)}^2 + \|\beta^{-1/2} w\|_{L^2(\mathcal{F}_h^{\text{N}} \cap Q_n)}^2 \\ &\quad + \|\alpha^{-1/2} \llbracket \boldsymbol{\tau} \rrbracket \cdot \mathbf{n}_F^x\|_{L^2(\mathcal{F}_h^{\text{time}} \cap Q_n)}^2 + \|\alpha^{-1/2} \boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x\|_{L^2(\mathcal{F}_h^{\text{D}} \cap Q_n)}^2. \end{aligned}$$

Therefore, from (22) and (24) we get

$$|(w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \leq 2 |(v, \boldsymbol{\sigma}) - (w_h, \boldsymbol{\tau}_h)|_{\text{DG}(Q_n)^+}. \quad (25)$$

In the following proposition, we prove that the  $|\cdot|_{\text{DG}(Q_n)}$  seminorm of the error (thus the  $L^2$  norm in space of the error at each discrete time  $t_n$ ) is bounded by the  $\text{DG}^+$  seminorm on the partial cylinder  $Q_n$  of the  $L^2$  projection error of the solution into the discrete space.

**Proposition 5.5.** *Let  $(v, \boldsymbol{\sigma})$  and  $(v_h, \boldsymbol{\sigma}_h)$  be the solutions of (1) and (16), respectively, and assume that  $(v, \boldsymbol{\sigma}) \in C^0([0, T]; H^1(\Omega))^3$ . Let  $(w_h, \boldsymbol{\tau}_h)$  be the orthogonal  $L^2$  projection of  $(v, \boldsymbol{\sigma})$  into  $\mathbf{V}_\mathbf{p}(\mathcal{T}_h)$ . Assume that  $\mathbf{V}_\mathbf{p}(\mathcal{T}_h)$  is a piecewise polynomial space and that (17) holds true. Then, for each discrete time  $t_n$ , we have the error bounds*

$$\begin{aligned} \frac{1}{2} \|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times \{t_n\})} &\leq |(v, \boldsymbol{\sigma}) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \\ &\leq 3 |(v, \boldsymbol{\sigma}) - (w_h, \boldsymbol{\tau}_h)|_{\text{DG}(Q_n)^+}. \end{aligned} \quad (26)$$

*Proof.* The first inequality comes from the definition of the  $|\cdot|_{\text{DG}(Q_n)}$  seminorm, which contains the  $L^2$  norm of the traces on  $\Omega \times \{t_n\}$  (as opposed to the full  $|\cdot|_{\text{DG}}$  norm, which contains the  $L^2$  norm of the time-jumps only on  $\Omega \times \{t_n\}$ ), and the bound  $\frac{1}{2}(A + B) \leq (\frac{1}{2}(A^2 + B^2))^{1/2}$  for all  $A, B \in \mathbb{R}$ . Assumption (17) ensures identity (23) and thus bounds (24)–(25). The second inequality then follows from the triangle inequality, the bound (25), and  $|\cdot|_{\text{DG}(Q_n)} \leq |\cdot|_{\text{DG}(Q_n)^+}$ .  $\square$

### 5.3.2 The general case

We now allow for corner singularities and assume that  $(v, \boldsymbol{\sigma})$  satisfies (8) for some  $k_t, k_x \in \mathbb{N}$ .

In order to control the inter-element terms on the right-hand side of (22), we cannot use directly the Cauchy-Schwarz inequality as in the derivation of (24), because for  $\boldsymbol{\tau} \in C^0((0, T); H_\delta^{1,1}(\Omega)^2)$  we have  $\boldsymbol{\tau} \in L^1(F)^2$  for  $F \in \mathbb{F}_\Sigma^{\text{time}}$  thanks to (9), but not necessarily  $\boldsymbol{\tau} \in L^2(F)^2$ . So we opt to apply the Hölder inequality in  $L^1$ – $L^\infty$  [45, Appendix]:

$$\begin{aligned} \left| \int_F \llbracket \boldsymbol{\tau} \rrbracket \cdot \llbracket w \rrbracket_{\mathbf{N}} \, dS \right| &\leq \|\llbracket \boldsymbol{\tau} \rrbracket\|_{L^2(F_t; L^1(F_x))} \|\llbracket w \rrbracket_{\mathbf{N}}\|_{L^2(F_t; L^\infty(F_x))} \\ &\leq C_{\infty|2} h_{F_x}^{-1/2} \|\llbracket \boldsymbol{\tau} \rrbracket\|_{L^2(F_t; L^1(F_x))} \|\llbracket w \rrbracket_{\mathbf{N}}\|_{L^2(F_t; L^2(F_x))} \end{aligned}$$

for all  $F \in \mathbb{F}_\Sigma^{\text{time}}$ ,  $\boldsymbol{\tau} \in L^2((0, T); H_\delta^{1,1}(\Omega)^2)$  and elementwise-polynomial  $w$ . Here  $C_{\infty|2}$  stems from the inverse inequality (12) for the polynomial space, thus it depends on the maximal polynomial degree in space admitted for  $w$  in the corner elements. The term  $\int_{\mathcal{F}_h^{\text{D}}} \boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x w \, dS$  is treated similarly. The other terms appearing in  $\mathcal{A}_{\text{DG}}(\cdot, \cdot)$  can be controlled by the Cauchy-Schwarz inequality, as  $\llbracket \boldsymbol{\sigma} \rrbracket_{\mathbf{N}} = 0$  on  $\mathcal{F}_h^{\text{time}}$ ,  $\boldsymbol{\sigma} \cdot \mathbf{n}_\Omega^x = 0$  on  $\mathcal{F}_h^{\text{N}}$ , and all other traces are in  $L^2(F)$  for the corresponding face  $F$ .

Proceeding as in §5.3.1, we control the bilinear form in (16) as

$$\begin{aligned} \mathcal{A}_{\text{DG}(Q_n)}((w_h, \boldsymbol{\tau}_h) - (v, \boldsymbol{\sigma}); (w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)) \\ \leq 2C_{\infty|2} |(w_h, \boldsymbol{\tau}_h) - (v, \boldsymbol{\sigma})|_{\text{DG}(Q_n)^+} |(w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \end{aligned} \quad (27)$$



where now

$$\begin{aligned}
|(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)^+}^2 &:= |(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)}^2 + 2 \|c^{-1}w^-\|_{L^2(\mathcal{F}_h^{\text{space}} \cap Q_n)}^2 + 2 \|\boldsymbol{\tau}^-\|_{L^2(\mathcal{F}_h^{\text{space}} \cap Q_n)}^2 \\
&+ \|\beta^{-1/2}\{\!\!\{w\}\!\!\}\|_{L^2(\mathcal{F}_h^{\text{time}} \cap Q_n)}^2 + \|\beta^{-1/2}w\|_{L^2(\mathcal{F}_h^{\text{N}} \cap Q_n)}^2 \\
&+ \sum_{F \in \mathbb{F}_Z^{\text{time}}(Q_n)} h_{F_x}^{-1} \|\alpha^{-1/2}\{\!\!\{\boldsymbol{\tau}\}\!\!\} \cdot \mathbf{n}_F^x\|_{L^2(F_t; L^1(F_x))}^2 + \|\alpha^{-1/2}\{\!\!\{\boldsymbol{\tau}\}\!\!\} \cdot \mathbf{n}_F^x\|_{L^2(\mathcal{F}_h^{\text{time}} \cap \mathcal{F}_h^{\circ} \cap Q_n)}^2 \\
&+ \sum_{F \in \mathbb{F}_Z^{\text{D}}(Q_n)} h_{F_x}^{-1} \|\alpha^{-1/2}\boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x\|_{L^2(F_t; L^1(F_x))}^2 + \|\alpha^{-1/2}\boldsymbol{\tau} \cdot \mathbf{n}_\Omega^x\|_{L^2(\mathcal{F}_h^{\text{D}} \cap \mathcal{F}_h^{\circ} \cap Q_n)}^2.
\end{aligned} \tag{28}$$

Therefore, from (22) and (27) we get

$$|(w_h, \boldsymbol{\tau}_h) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \leq 2C_{\infty|2} |(v, \boldsymbol{\sigma}) - (w_h, \boldsymbol{\tau}_h)|_{\text{DG}(Q_n)^+}. \tag{29}$$

The following result is proved exactly as Proposition 5.5, using (29) instead of (25) and the corresponding modified  $|\cdot|_{\text{DG}(Q_n)^+}$  seminorm.

**Proposition 5.6.** *Let  $(v, \boldsymbol{\sigma})$  and  $(v_h, \boldsymbol{\sigma}_h)$  be the solutions of (1) and (16), respectively. Assume that  $(v, \boldsymbol{\sigma})$  satisfies the weighted regularity condition (8) for some  $k_t, k_x \in \mathbb{N}$ . Let  $(w_h, \boldsymbol{\tau}_h)$  be the orthogonal  $L^2$  projection of  $(v, \boldsymbol{\sigma})$  into  $\mathbf{V}_p(\mathcal{T}_h)$ . Assume that  $\mathbf{V}_p(\mathcal{T}_h)$  is a piecewise polynomial space and that (17) holds true. Then, for each discrete time  $t_n$ , we have the error bounds*

$$\begin{aligned}
\frac{1}{2} \|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times \{t_n\})} &\leq |(v, \boldsymbol{\sigma}) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \\
&\leq (2C_{\infty|2} + 1) |(v, \boldsymbol{\sigma}) - (w_h, \boldsymbol{\tau}_h)|_{\text{DG}(Q_n)^+}.
\end{aligned} \tag{30}$$

Recall from (13) that  $C_{\infty|2} \leq p + 1$ ,  $p$  being the maximal polynomial degree in the spatial variable of the elements of  $\mathbf{V}_p(\mathcal{T}_h^{\angle})$ , thus the abstract error bound (30) can be made fully explicit, and the bounding constant depends linearly on  $p$ . Note that when a local mesh refinement strategy is used, since corner elements are taken to be small,  $p$  can be chosen to be 0 or 1.

**Remark 5.7** (Energy bounds). *We briefly discuss the dissipation properties of the proposed scheme. We define the energy of a pair  $(w, \boldsymbol{\tau})$  at time  $t \in [0, T]$  as*

$$\mathcal{E}(t; w, \boldsymbol{\tau}) := \frac{1}{2} \int_{\Omega} (c^{-2}w^2(\mathbf{x}, t) + |\boldsymbol{\tau}(\mathbf{x}, t)|^2) \, d\mathbf{x}.$$

*Then, if  $(v, \boldsymbol{\sigma})$  is a solution of the IBVP (1) with  $f = 0$ , we have the identity  $\mathcal{E}(t; v, \boldsymbol{\sigma}) = \mathcal{E}(0; v, \boldsymbol{\sigma}) - \int_{\partial\Omega \times (0, t)} v\boldsymbol{\sigma} \cdot \mathbf{n}_\Omega^x \, dS$  (e.g. [33, equation (15)]). If, moreover, the boundary conditions are homogeneous (i.e.,  $g_{\text{D}} = g_{\text{N}} = 0$ ) the total energy is preserved,  $\mathcal{E}(t; v, \boldsymbol{\sigma}) = \mathcal{E}(0; v, \boldsymbol{\sigma})$ . Proceeding exactly as in [33, (16)–(17)] we see that the DG method is dissipative and we quantify the energy dissipated in  $Q_n$ : for  $(v_h, \boldsymbol{\sigma}_h)$  the solution of (16) and  $1 \leq n \leq N$ ,*

$$\begin{aligned}
\mathcal{E}(t_n; v_h, \boldsymbol{\sigma}_h) &= \mathcal{E}(0; v_0, \boldsymbol{\sigma}_0) - \frac{1}{2} \|c^{-1}\llbracket v_h \rrbracket_t\|_{L^2(\mathcal{F}_h^{\text{space}} \cap Q_n)}^2 - \|\alpha^{1/2}\llbracket v_h \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\text{time}} \cap Q_n)}^2 - \|\alpha^{1/2}v_h\|_{L^2(\mathcal{F}_h^{\text{D}} \cap Q_n)}^2 \\
&- \frac{1}{2} \|\llbracket \boldsymbol{\sigma} \rrbracket_t\|_{L^2(\mathcal{F}_h^{\text{space}} \cap Q_n)}^2 - \|\beta^{1/2}\llbracket \boldsymbol{\sigma} \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\text{time}} \cap Q_n)}^2 - \|\beta^{1/2}\boldsymbol{\sigma} \cdot \mathbf{n}_\Omega^x\|_{L^2(\mathcal{F}_h^{\text{N}} \cap Q_n)}^2.
\end{aligned}$$

*This means that the energy dissipated by the discrete solution is proportional to the jumps of the solution on the mesh skeleton and to the mismatch with the weakly enforced (homogeneous) boundary conditions.*

## 6 Projection and Galerkin error estimates

Given a vector of elemental polynomial degrees  $\mathbf{p} = ((p_{x,K}^v, p_{t,K}^v, p_{x,K}^\sigma, p_{t,K}^\sigma) \in \mathbb{N}_0^4, K \in \mathcal{T}_h)$ , we choose as trial and test space the piecewise-polynomial space (with  $\otimes$  denoting the algebraic tensor product)

$$\mathbf{V}_p(\mathcal{T}_h) = \prod_{K=K_x \times I_n \in \mathcal{T}_h} \left( \mathbb{P}^{p_{x,K}^v, K}(\mathbf{K}_x) \otimes \mathbb{P}^{p_{t,K}^v, K}(I_n) \right) \times \left( \mathbb{P}^{p_{x,K}^\sigma, K}(\mathbf{K}_x) \otimes \mathbb{P}^{p_{t,K}^\sigma, K}(I_n) \right)^2. \tag{31}$$

Here and in what follows, the space  $\mathbb{P}^{p_{\bullet}, \kappa}(K_{\mathbf{x}})$ , for  $\bullet \in \{v, \boldsymbol{\sigma}\}$ , of polynomials of degree at most  $p_{\bullet, K}$  in two variables can be replaced by the space  $\mathbb{Q}^{p_{\bullet}, \kappa}(K_{\mathbf{x}})$  of polynomials of the same degree in each of the two variables.

Bound (30) states that we can control the DG seminorm of the Galerkin error by the  $\text{DG}^+$  seminorm of the error of its elementwise  $L^2$  projection in  $\mathbf{V}_{\mathbf{p}}(\mathcal{T}_h)$ . So we need to derive error estimates for the *traces* of the *volume*  $L^2$  projection. The  $p$ -version of these estimates was derived in [22] and [10] for cubes and simplices, respectively; here we aim at the  $h$ -version, for which simpler scaling argument are sufficient. We partly follow [36, §5].

## 6.1 Bounds of DG seminorms in terms of volume norms

As this section contains auxiliary results, we consider directly the general case of solutions admitting corner singularities. If  $(v, \boldsymbol{\sigma})$  is smooth, in particular, if  $\boldsymbol{\sigma} \in C^0([0, T]; H^1(\Omega)^2)$ , then in (33) and (39) below, the summations over faces in  $\mathcal{F}_h^{\angle}$  or elements in  $\mathcal{T}_h^{\angle}$  can be dropped, if simultaneously the terms over (subsets of)  $\mathcal{F}_h^{\circ}$  and  $\mathcal{T}_h^{\circ}$  are extended to (analogous subsets of) the whole of  $\mathcal{F}_h$  and  $\mathcal{T}_h$ .

We define the space containing both the exact solution and the discrete functions

$$V_+ := \left( C^0([0, T]; H_{\delta}^{2,2}(\Omega)) \times C^0([0, T]; H_{\delta, N}^{1,1}(\Omega)^2) \right) + \mathbf{V}_{\mathbf{p}}(\mathcal{T}_h), \quad \text{where}$$

$$H_{\delta, N}^{1,1}(\Omega)^2 := \{ \boldsymbol{\tau} \in H_{\delta}^{1,1}(\Omega)^2, \boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x = 0 \text{ on } \mathcal{F}_h^N \}.$$

Here we want to derive a bound of  $|(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)_+}$  in terms of elementwise sums of traces, tracking the dependence on spatial and temporal meshsize for all  $(w, \boldsymbol{\tau}) \in V_+$ . For all  $(w, \boldsymbol{\tau}) \in V_+$  there exists a (not necessarily unique) decomposition in continuous + discrete components:

$$(w, \boldsymbol{\tau}) = (\tilde{w}, \tilde{\boldsymbol{\tau}}) + (w_h, \boldsymbol{\tau}_h), \quad (\tilde{w}, \tilde{\boldsymbol{\tau}}) \in C^0([0, T]; H_{\delta}^{2,2}(\Omega)) \times C^0([0, T]; H_{\delta, N}^{1,1}(\Omega)^2), \quad (32)$$

$$(w_h, \boldsymbol{\tau}_h) \in \mathbf{V}_{\mathbf{p}}(\mathcal{T}_h).$$

Thus for  $(w, \boldsymbol{\tau}) \in V_+$ , taking into account the definition (28), we get

$$|(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)_+}^2 \lesssim \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} \left[ \|c^{-1}w\|_{L^2(K_{\mathbf{x}} \times \{t_{n'-1}, t_{n'}\})}^2 + \|\boldsymbol{\tau}\|_{L^2(K_{\mathbf{x}} \times \{t_{n'-1}, t_{n'}\})}^2 \right. \quad (33)$$

$$+ \sum_{F \in \mathbb{F}^{\text{time}}(K) \cup \mathbb{F}^{\text{D}}(K)} \left\| \alpha^{1/2} w \right\|_{L^2(F)}^2 + \sum_{F \in \mathbb{F}^{\text{time}}(K) \cup \mathbb{F}^{\text{N}}(K)} \left\| \beta^{-1/2} w \right\|_{L^2(F)}^2$$

$$+ \sum_{F \in \mathbb{F}^{\text{time}}(K) \cup \mathbb{F}^{\text{D}}(K)} \left\| \alpha^{-1/2} \boldsymbol{\tau} \cdot \mathbf{n}_F^x \right\|_{L^2(F)}^2 + \sum_{F \in \mathbb{F}^{\text{time}}(K) \cup \mathbb{F}^{\text{D}}(K)} h_{F_{\mathbf{x}}}^{-1} \left\| \alpha^{-1/2} \boldsymbol{\tau} \cdot \mathbf{n}_F^x \right\|_{L^2(F_t; L^1(F_{\mathbf{x}}))}^2 \left. \right]$$

$$+ \left\| \beta^{1/2} \llbracket \boldsymbol{\tau} \rrbracket_{\mathbf{N}} \right\|_{L^2(\mathcal{F}_h^{\text{time}} \cap Q_n)}^2 + \left\| \beta^{1/2} \boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(\mathcal{F}_h^N \cap \overline{Q_n})}^2,$$

where all traces are taken from the element  $K$ . Here we used the  $L^1$ - $L^2$  inequality (11) to treat the term containing  $\llbracket \boldsymbol{\tau} \rrbracket$  on some of the faces  $F \in \mathbb{F}^{\text{time}}$ ; in particular, on the faces in the form  $F = \partial K_{\angle} \cap \partial K_{\circ}$  for  $K_{\angle} \subset \mathcal{T}_h^{\angle}$  and  $K_{\circ} \subset \mathcal{T}_h^{\circ}$ , the trace of  $\boldsymbol{\tau}|_{K_{\circ}}$  is lifted from the  $L^2(F_t; L^1(F_{\mathbf{x}}))$  norm to the  $L^2(F)$  norm.

The two terms containing  $\beta^{1/2} \boldsymbol{\tau}$  are non zero only for the discrete component  $\boldsymbol{\tau}_h$  of  $\boldsymbol{\tau}$ , recall (32). Since this is a polynomial of given degree, we control these terms using the inverse inequality (12). The terms on the Neumann boundary can be controlled as follows:

$$\left\| \beta^{1/2} \boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(\mathcal{F}_h^N \cap \overline{Q_n})}^2 \stackrel{\tilde{\boldsymbol{\tau}} \cdot \mathbf{n}_{\Omega}^x = 0}{=} \left\| \beta^{1/2} \boldsymbol{\tau}_h \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(\mathcal{F}_h^N \cap \overline{Q_n})}^2$$

$$= \sum_{F \in \mathbb{F}_{\circ}^N(Q_n)} \left\| \beta^{1/2} \boldsymbol{\tau}_h \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(F)}^2 + \sum_{F \in \mathbb{F}_{\angle}^N(Q_n)} \left\| \beta^{1/2} \boldsymbol{\tau}_h \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(F)}^2$$

$$\stackrel{(12)}{\leq} \sum_{F \in \mathbb{F}_{\circ}^N(Q_n)} \left\| \beta^{1/2} \boldsymbol{\tau}_h \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(F)}^2 + C_{2|1} \sum_{F \in \mathbb{F}_{\angle}^N(Q_n)} h_{F_{\mathbf{x}}}^{-1} \left\| \beta^{1/2} \boldsymbol{\tau}_h \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(F_t; L^1(F_{\mathbf{x}}))}^2$$

$$\stackrel{\tilde{\boldsymbol{\tau}} \cdot \mathbf{n}_{\Omega}^x = 0}{=} \sum_{F \in \mathbb{F}_{\circ}^N(Q_n)} \left\| \beta^{1/2} \boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(F)}^2 + C_{2|1} \sum_{F \in \mathbb{F}_{\angle}^N(Q_n)} h_{F_{\mathbf{x}}}^{-1} \left\| \beta^{1/2} \boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(F_t; L^1(F_{\mathbf{x}}))}^2$$

$$= \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} \left[ \sum_{F \in \mathbb{F}_{\odot}^{\mathbb{N}}(K)} \left\| \beta^{1/2} \boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(F)}^2 + C_{2|1} \sum_{F \in \mathbb{F}_{\angle}^{\mathbb{N}}(K)} h_{F_{\mathbf{x}}}^{-1} \left\| \beta^{1/2} \boldsymbol{\tau} \cdot \mathbf{n}_{\Omega}^x \right\|_{L^2(F_t; L^1(F_{\mathbf{x}}))}^2 \right].$$

The jump term is controlled similarly:

$$\begin{aligned} & \left\| \beta^{1/2} \llbracket \boldsymbol{\tau} \rrbracket_{\mathbf{N}} \right\|_{L^2(\mathcal{F}_h^{\text{time}} \cap Q_n)}^2 \stackrel{\llbracket \tilde{\boldsymbol{\tau}} \rrbracket_{\mathbf{N}}=0}{=} \left\| \beta^{1/2} \llbracket \boldsymbol{\tau}_h \rrbracket_{\mathbf{N}} \right\|_{L^2(\mathcal{F}_h^{\text{time}} \cap Q_n)}^2 \\ &= \sum_{F \in \mathbb{F}_{\odot}^{\text{time}}(Q_n)} \left\| \beta^{1/2} \llbracket \boldsymbol{\tau}_h \rrbracket_{\mathbf{N}} \right\|_{L^2(F)}^2 + \sum_{F \in \mathbb{F}_{\angle}^{\text{time}}(Q_n)} \left\| \beta^{1/2} \llbracket \boldsymbol{\tau}_h \rrbracket_{\mathbf{N}} \right\|_{L^2(F)}^2 \\ &\stackrel{(12)}{\leq} \sum_{F \in \mathbb{F}_{\odot}^{\text{time}}(Q_n)} \left\| \beta^{1/2} \llbracket \boldsymbol{\tau}_h \rrbracket_{\mathbf{N}} \right\|_{L^2(F)}^2 + C_{2|1} \sum_{F \in \mathbb{F}_{\angle}^{\text{time}}(Q_n)} h_{F_{\mathbf{x}}}^{-1} \left\| \beta^{1/2} \llbracket \boldsymbol{\tau}_h \rrbracket_{\mathbf{N}} \right\|_{L^2(F_t; L^1(F_{\mathbf{x}}))}^2 \\ &\stackrel{\llbracket \tilde{\boldsymbol{\tau}} \rrbracket_{\mathbf{N}}=0}{=} \sum_{F \in \mathbb{F}_{\odot}^{\text{time}}(Q_n)} \left\| \beta^{1/2} \llbracket \boldsymbol{\tau} \rrbracket_{\mathbf{N}} \right\|_{L^2(F)}^2 + C_{2|1} \sum_{F \in \mathbb{F}_{\angle}^{\text{time}}(Q_n)} h_{F_{\mathbf{x}}}^{-1} \left\| \beta^{1/2} \llbracket \boldsymbol{\tau} \rrbracket_{\mathbf{N}} \right\|_{L^2(F_t; L^1(F_{\mathbf{x}}))}^2 \\ &\stackrel{(11)}{\leq} (1 + C_{2|1}) \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^{\odot}(Q_n)} \sum_{F \in \mathbb{F}^{\text{time}}(K)} \left\| \beta^{1/2} \boldsymbol{\tau} \cdot \mathbf{n}_K^x \right\|_{L^2(F)}^2 \\ &\quad + C_{2|1} \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^{\angle}(Q_n)} \sum_{F \in \mathbb{F}^{\text{time}}(K)} h_{F_{\mathbf{x}}}^{-1} \left\| \beta^{1/2} \boldsymbol{\tau} \cdot \mathbf{n}_K^x \right\|_{L^2(F_t; L^1(F_{\mathbf{x}}))}^2, \end{aligned}$$

where in the last step we used again the  $L^1$ - $L^2$  inequality (11) to treat the contributions from  $\mathcal{T}_h^{\odot}$  to  $\llbracket \boldsymbol{\tau} \rrbracket_{\mathbf{N}}$  on the faces of the form  $F = \partial K_{\angle} \cap \partial K_{\odot}$  for  $K_{\angle} \subset \mathcal{T}_h^{\angle}$  and  $K_{\odot} \subset \mathcal{T}_h^{\odot}$ .

For all elements  $K = K_{\mathbf{x}} \times I_n \in \mathcal{T}_h$ , the standard weighted trace inequality, applied in the time and space directions independently, reads

$$\|\varphi\|_{L^2(K_{\mathbf{x}} \times \{t_{n-1}, t_n\})}^2 \lesssim h_n^{-1} \|\varphi\|_{L^2(K)}^2 + h_n |\varphi|_{H^1(I_n; L^2(K_{\mathbf{x}}))}^2 \quad \forall \varphi \in H^1(I_n; L^2(K_{\mathbf{x}})), \quad (34)$$

$$\|\varphi\|_{L^2(\partial K_{\mathbf{x}} \times I_n)}^2 \lesssim h_{K_{\mathbf{x}}}^{-1} \|\varphi\|_{L^2(K)}^2 + h_{K_{\mathbf{x}}} |\varphi|_{L^2(I_n; H^1(K_{\mathbf{x}}))}^2 \quad \forall \varphi \in L^2(I_n; H^1(K_{\mathbf{x}})), \quad (35)$$

where the hidden constant in (35) only depends on the shape-regularity parameter of the space mesh (see e.g. [7, 1.6.6] or [33, Lemma 2] for star-shaped elements). The corresponding result for weighted spaces follows from Lemma 4.1 ([45, Lemma 1.3.2c]):

$$\|\varphi\|_{L^2(I_n; L^1(\partial K_{\mathbf{x}}))}^2 \lesssim \|\varphi\|_{L^2(K)}^2 + h_{K_{\mathbf{x}}}^{2-2\delta_K} |\varphi|_{L^2(I_n; H_{\delta}^{1,1}(K_{\mathbf{x}}))}^2 \quad \forall \varphi \in L^2(I_n; H_{\delta}^{1,1}(K_{\mathbf{x}})). \quad (36)$$

We recall that the numerical flux parameters  $\alpha, \beta$  are assumed to be constant on each element face on which they are defined. For each element  $K \in \mathcal{T}_h$ , we write

$$\begin{aligned} \alpha_{\min}^K &:= \min_{F \in \mathbb{F}^{\text{time}}(K) \cup \mathbb{F}^{\text{D}}(K)} \{\alpha|_F\}, & \beta_{\min}^K &:= \min_{F \in \mathbb{F}^{\text{time}}(K) \cup \mathbb{F}^{\text{N}}(K)} \{\beta|_F\}, \\ \alpha_{\max}^K &:= \max_{F \in \mathbb{F}^{\text{time}}(K) \cup \mathbb{F}^{\text{D}}(K)} \{\alpha|_F\}, & \beta_{\max}^K &:= \max_{F \in \mathbb{F}^{\text{time}}(K) \cup \mathbb{F}^{\text{N}}(K)} \{\beta|_F\}, \\ \mathbf{a}_K &:= \max\{(\alpha_{\max}^K)^{1/2}, (\beta_{\min}^K)^{-1/2}\}, & \mathbf{b}_K &:= \max\{(\beta_{\max}^K)^{1/2}, (\alpha_{\min}^K)^{-1/2}\}. \end{aligned} \quad (37)$$

As the spatial elements are assumed not to have degenerating faces, there holds

$$h_{F_{\mathbf{x}}} \approx h_{K_{\mathbf{x}}}, \quad \text{for } F \in \mathbb{F}^{\text{time}}(K) \cup \mathbb{F}^{\text{D}}(K) \cup \mathbb{F}^{\text{N}}(K), \quad K \in \mathcal{T}_h. \quad (38)$$

We have now all the ingredients to derive an upper bound on the  $|\cdot|_{\text{DG}}$  seminorm in terms of standard space-time (weighted, Bochner) Sobolev norms of its argument over the mesh elements. Note that we cannot expect to obtain a uniform bound in term of the spatial and temporal meshsize (e.g. for  $w = 1$ ,  $\boldsymbol{\tau} = \mathbf{0}$ , and constant  $\alpha$  and  $\beta$ , the seminorm  $|(w, \boldsymbol{\tau})|_{\text{DG}^+}^2$  is proportional to the 2-dimensional Hausdorff measure of  $\mathcal{F}_h$ , which is not bounded for  $h_{\mathbf{x}}, h_t \searrow 0$ ).

**Proposition 6.1.** *For all  $(w, \boldsymbol{\tau}) \in V_+$ , the following bound holds true:*

$$\begin{aligned} & |(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)^+}^2 \quad (39) \\ & \lesssim \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} \left[ h_{n'}^{-1} \left( \|c^{-1}w\|_{L^2(K)}^2 + \|\boldsymbol{\tau}\|_{L^2(K)^2}^2 \right) + h_{n'} \left( |c^{-1}w|_{H^1(I_{n'}; L^2(K_{\mathbf{x}}))}^2 + \|\boldsymbol{\tau}\|_{H^1(I_{n'}; L^2(K_{\mathbf{x}})^2)}^2 \right) \right] \end{aligned}$$

$$\begin{aligned}
& + h_{K_{\mathbf{x}}}^{-1} \|\mathbf{a}_K w\|_{L^2(K)}^2 + h_{K_{\mathbf{x}}} |\mathbf{a}_K w|_{L^2(I_{n'}; H^1(K_{\mathbf{x}}))}^2 + \underbrace{h_{K_{\mathbf{x}}}^{-1}}_{h_{K_{\mathbf{x}}} \lesssim h_{F_{\mathbf{x}}}} \|\mathbf{b}_K \boldsymbol{\tau}\|_{L^2(K)^2}^2 \\
& + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\circ(Q_n)} h_{K_{\mathbf{x}}} |\mathbf{b}_K \boldsymbol{\tau}|_{L^2(I_{n'}; H^1(K_{\mathbf{x}})^2)}^2 + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\angle(Q_n)} \underbrace{h_{K_{\mathbf{x}}}^{1-2\delta_K}}_{h_{K_{\mathbf{x}}} \lesssim h_{F_{\mathbf{x}}}} |\mathbf{b}_K \boldsymbol{\tau}|_{L^2(I_{n'}; H_\delta^{1,1}(K_{\mathbf{x}})^2)}^2,
\end{aligned}$$

where the hidden constant depends linearly on  $\max\{p_{x,K}^\sigma, K \in \mathcal{T}_h^\angle\}$  through the inverse inequality constant  $C_{2|1}$  (recall (13)).

*Proof.* From the bound (33), after treating the terms with  $\beta^{1/2} \boldsymbol{\tau}$  as described, by using the trace inequalities (34), (35) and (36), notation (37), and assumption (38), for all  $(w, \boldsymbol{\tau}) \in V_+$ , we obtain (39).  $\square$

We note that the only polynomial degree affecting the bounding constant in (39) is the degree of the space associated with the corner elements, which is typically taken to be 0 or 1. So in this case the bounding constant only depends on the element shapes.

## 6.2 $L^2$ projection error estimates for a scalar function

In this section we prove error bounds in Bochner norms for the  $L^2$  projection on polynomial spaces.

### 6.2.1 The smooth case

For functions without corner singularities we have the following result.

**Proposition 6.2.** *Let  $K = K_{\mathbf{x}} \times I_n \in \mathcal{T}_h$  be a prismatic space-time element, i.e.  $K_{\mathbf{x}}$  is a shape-regular polygon, with  $h_{K_{\mathbf{x}}} = \text{diam}(K_{\mathbf{x}})$  and  $h_n = |I_n|$ . Denote by  $\Pi_{\mathbf{p}}$ ,  $\mathbf{p} = (p_x, p_t) \in \mathbb{N}_0^2$ , the  $L^2(K)$ -orthogonal projection:  $L^2(K) \rightarrow \mathbb{P}^{\mathbf{p}}(K) = \mathbb{P}^{p_x}(K_{\mathbf{x}}) \otimes \mathbb{P}^{p_t}(I_n)$ .*

*Let  $\varphi \in H^{k_t+1}(I_n; L^2(K_{\mathbf{x}})) \cap L^2(I_n; H^{k_x+1}(K_{\mathbf{x}}))$ ,  $k_x, k_t \in \mathbb{N}_0$ . Then*

$$\begin{aligned}
& \|\varphi - \Pi_{\mathbf{p}} \varphi\|_{L^2(I_n; L^2(K_{\mathbf{x}}))} + h_n |\varphi - \Pi_{\mathbf{p}} \varphi|_{H^1(I_n; L^2(K_{\mathbf{x}}))} + h_{K_{\mathbf{x}}} |\varphi - \Pi_{\mathbf{p}} \varphi|_{L^2(I_n; H^1(K_{\mathbf{x}}))} \\
& \lesssim h_n^{s_t+1} |\varphi|_{H^{s_t+1}(I_n; L^2(K_{\mathbf{x}}))} + h_{K_{\mathbf{x}}}^{s_x+1} |\varphi|_{L^2(I_n; H^{s_x+1}(K_{\mathbf{x}}))}, \tag{40}
\end{aligned}$$

where  $s_t := \min\{k_t, p_t\}$ ,  $s_x := \min\{k_x, p_x\}$ , and the hidden constants are independent of  $h_{K_{\mathbf{x}}}$  and  $h_n$ , but depend on the polynomial degrees  $p_t$  and  $p_x$ , and on the shape-regularity constant of  $K_{\mathbf{x}}$ .

*Proof.* Consider a reference element  $\widehat{K} = \widehat{K}_{\widehat{\mathbf{x}}} \times \widehat{I}$ , and denote by  $\widehat{\Pi}_{\mathbf{p}}$ ,  $\mathbf{p} \in \mathbb{N}_0^2$ , the  $L^2(\widehat{K})$ -orthogonal projection:  $L^2(\widehat{K}) \rightarrow \mathbb{P}^{\mathbf{p}}(\widehat{K}) = \mathbb{P}^{p_x}(\widehat{K}_{\widehat{\mathbf{x}}}) \otimes \mathbb{P}^{p_t}(\widehat{I})$ . Let  $\widehat{\varphi}$  be in  $H^{k_t+1}(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}})) \cap L^2(\widehat{I}; H^{k_x+1}(\widehat{K}_{\widehat{\mathbf{x}}}))$ . We prove that

$$\begin{aligned}
& \left\| \widehat{\varphi} - \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} \right\|_{L^2(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}}))} + \left| \widehat{\varphi} - \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} \right|_{H^1(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}}))} + \left| \widehat{\varphi} - \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} \right|_{L^2(\widehat{I}; H^1(\widehat{K}_{\widehat{\mathbf{x}}}))} \\
& \lesssim |\widehat{\varphi}|_{H^{s_t+1}(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}}))} + |\widehat{\varphi}|_{L^2(\widehat{I}; H^{s_x+1}(\widehat{K}_{\widehat{\mathbf{x}}}))}. \tag{41}
\end{aligned}$$

The assertion (40) then follows from a scaling argument.

We recall that the Peetre–Tartar lemma, as formulated in [14, Lemma A.38], states that if  $X, Y, Z$  are Banach spaces,  $A : X \rightarrow Y$  is injective,  $T : X \rightarrow Z$  is compact,  $\|x\|_X \lesssim \|Ax\|_Y + \|Tx\|_Z$  for all  $x \in X$ , then  $\|x\|_X \lesssim \|Ax\|_Y$  for all  $x \in X$ . Within this setting, we fix

$$\begin{aligned}
X &= H^{s_t+1}(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}})) \cap L^2(\widehat{I}; H^{s_x+1}(\widehat{K}_{\widehat{\mathbf{x}}})), \\
Y &= \left( \mathbb{P}^{s_x}(\widehat{K}_{\widehat{\mathbf{x}}}) \otimes \mathbb{P}^{s_t}(\widehat{I}) \right) \times L^2(\widehat{K}) \times L^2(\widehat{K})^{s_x+2}, \\
Z &= L^2(\widehat{K}), \\
A : \widehat{\psi} &\mapsto \left( \widehat{\Pi}_{(s_x, s_t)} \widehat{\psi}, \quad \partial_{\widehat{I}}^{s_t+1} \widehat{\psi}, \quad D_{\widehat{\mathbf{x}}}^{s_x+1} \widehat{\psi} \right)
\end{aligned}$$

where  $D_{\widehat{\mathbf{x}}}^{s_x+1} \widehat{\psi}$  denotes the collection of the  $s_x + 2$  space partial derivatives of order exactly  $s_x + 1$  of  $\widehat{\psi}$ , and  $T : X \rightarrow Z$  is the natural embedding.

The intersection  $L^2(\widehat{I}; H^1(\widehat{K}_{\widehat{\mathbf{x}}})) \cap H^1(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}}))$  is isomorphic to  $H^1(\widehat{K})$ , thus it is compactly<sup>2</sup> embedded in  $L^2(\widehat{K})$ . Since  $X \subset L^2(\widehat{I}; H^1(\widehat{K}_{\widehat{\mathbf{x}}})) \cap H^1(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}}))$ , the embedding operator  $T : X \rightarrow Z$ , is compact.

<sup>2</sup>Recall that if  $U \subset V$  is a compact inclusion of Banach spaces, the embedding of the corresponding Bochner spaces  $L^2(\widehat{I}; U) \subset L^2(\widehat{I}; V)$  is not compact; see a counterexample in [39, Appendix A].

If  $A\widehat{\psi} = 0$  then  $\widehat{\psi}$  has vanishing  $(s_t + 1)$ th time derivative and  $(s_x + 1)$ th space derivatives, so it is polynomial of degree at most  $s_t$  in time and  $s_x$  in space; furthermore  $A\widehat{\psi} = 0$  implies that  $\widehat{\psi}$  has vanishing projection on the polynomial space  $\mathbb{P}^{s_x}(\widehat{K}_{\widehat{\mathbf{x}}}) \otimes \mathbb{P}^{s_t}(\widehat{I})$ , so  $\widehat{\psi}$  vanishes identically in  $\widehat{K}$  and  $A$  is injective.

By the corollary on page 21 of [32] (see also the definitions on page 7),

$$\begin{aligned} \left\| \widehat{\psi} \right\|_X &= \left\| \widehat{\psi} \right\|_{H^{s_t+1}(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}}))} + \left\| \widehat{\psi} \right\|_{L^2(\widehat{I}; H^{s_x+1}(\widehat{K}_{\widehat{\mathbf{x}}}))} \lesssim \left\| \widehat{\psi} \right\|_{L^2(\widehat{K})} + \left\| \partial_t^{s_t+1} \widehat{\psi} \right\|_{L^2(\widehat{K})} + \left\| D_{\widehat{\mathbf{x}}}^{s_x+1} \widehat{\psi} \right\|_{L^2(\widehat{K})^{s_x+2}} \\ &\leq \left\| A\widehat{\psi} \right\|_Y + \left\| T\widehat{\psi} \right\|_Z, \end{aligned}$$

therefore, the Peetre–Tartar lemma gives

$$\left\| \widehat{\psi} \right\|_X = \left\| \widehat{\psi} \right\|_{H^{s_t+1}(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}}))} + \left\| \widehat{\psi} \right\|_{L^2(\widehat{I}; H^{s_x+1}(\widehat{K}_{\widehat{\mathbf{x}}}))} \lesssim \left\| A\widehat{\psi} \right\|_Y, \quad \forall \widehat{\psi} \in X. \quad (42)$$

When  $s_t = p_t$  and  $s_x = p_x$ , then  $\partial_t^{s_t+1} \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} = 0$  and  $D_{\widehat{\mathbf{x}}}^{s_x+1} \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} = 0$ . Otherwise, if  $s_t < p_t$ , in general  $\partial_t^{s_t+1} \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} \neq 0$ . However, denoting  $\widehat{\pi}_x^{s_x} : L^2(\widehat{K}_{\widehat{\mathbf{x}}}) \rightarrow \mathbb{P}^{s_x}(\widehat{K}_{\widehat{\mathbf{x}}})$  and  $\widehat{\pi}_t^{s_t} : L^2(\widehat{I}) \rightarrow \mathbb{P}^{s_t}(\widehat{I})$  the  $L^2$ -orthogonal projections, we split the projector  $\widehat{\Pi}_{\mathbf{p}} := (I \otimes \widehat{\pi}_x^{s_x}) \circ (\widehat{\pi}_t^{s_t} \otimes I)$ , and obtain

$$\begin{aligned} \left\| \partial_t^{s_t+1} \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} \right\|_{L^2(\widehat{K})} &= \left\| \widehat{\Pi}_x^{s_x} \partial_t^{s_t+1} \widehat{\Pi}_t^{s_t} \widehat{\varphi} \right\|_{L^2(\widehat{K})} \leq \left\| \partial_t^{s_t+1} \widehat{\Pi}_t^{s_t} \widehat{\varphi} \right\|_{L^2(\widehat{K})} \\ &\leq \left\| \partial_t^{s_t+1} \widehat{\varphi} \right\|_{L^2(\widehat{K})} + \left\| \partial_t^{s_t+1} (\widehat{\varphi} - \widehat{\Pi}_t^{s_t} \widehat{\varphi}) \right\|_{L^2(\widehat{K})} \lesssim \left\| \partial_t^{s_t+1} \widehat{\varphi} \right\|_{L^2(\widehat{K})} \end{aligned}$$

by the Bramble–Hilbert lemma (we have used the  $L^2$  continuity of the  $L^2$  projection in the first inequality). Similarly, if  $s_x < p_x$ , then

$$\left\| D_{\widehat{\mathbf{x}}}^{s_x+1} \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} \right\|_{L^2(\widehat{K})} \lesssim \left\| D_{\widehat{\mathbf{x}}}^{s_x+1} \widehat{\varphi} \right\|_{L^2(\widehat{K})}.$$

In all cases, bound (42) applied to  $\widehat{\psi} = \widehat{\varphi} - \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi}$ , together with  $\widehat{\Pi}_{\mathbf{p}}(\widehat{\varphi} - \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi}) = 0$ , gives

$$\begin{aligned} \left\| \widehat{\varphi} - \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} \right\|_{H^{s_t+1}(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}}))} + \left\| \widehat{\varphi} - \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} \right\|_{L^2(\widehat{I}; H^{s_x+1}(\widehat{K}_{\widehat{\mathbf{x}}}))} &= \left\| \widehat{\varphi} - \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi} \right\|_X \lesssim \left\| A(\widehat{\varphi} - \widehat{\Pi}_{\mathbf{p}} \widehat{\varphi}) \right\|_Y \\ &\lesssim \left\| \partial_t^{s_t+1} \widehat{\varphi} \right\|_{L^2(\widehat{K})} + \left\| D_{\widehat{\mathbf{x}}}^{s_x+1} \widehat{\varphi} \right\|_{L^2(\widehat{K})^{s_x+2}} = |\widehat{\varphi}|_{H^{s_t+1}(\widehat{I}; L^2(\widehat{K}_{\widehat{\mathbf{x}}}))} + |\widehat{\varphi}|_{L^2(\widehat{I}; H^{s_x+1}(\widehat{K}_{\widehat{\mathbf{x}}}))}, \end{aligned} \quad (43)$$

which admits (41) as a particular case.  $\square$

**Remark 6.3.** *The proof of Proposition 6.2 actually provides a slightly stronger bound under the same assumptions. Indeed, applying a scaling argument to bound (43) gives*

$$\begin{aligned} \sum_{z_t=0}^{s_t} h_n^{z_t} |\varphi - \Pi_{\mathbf{p}} \varphi|_{H^{z_t}(I_n; L^2(K_{\mathbf{x}}))} + \sum_{z_x=1}^{s_x} h_{K_{\mathbf{x}}}^{z_x} |\varphi - \Pi_{\mathbf{p}} \varphi|_{L^2(I_n; H^{z_x}(K_{\mathbf{x}}))} \\ \lesssim h_n^{s_t+1} |\varphi|_{H^{s_t+1}(I_n; L^2(K_{\mathbf{x}}))} + h_{K_{\mathbf{x}}}^{s_x+1} |\varphi|_{L^2(I_n; H^{s_x+1}(K_{\mathbf{x}}))}. \end{aligned}$$

## 6.2.2 The general case

The next proposition is concerned with functions with  $H_{\delta}^{1,1}$  regularity in space.

**Proposition 6.4.** *Let  $K = K_{\mathbf{x}} \times I_n \in \mathcal{T}_h^{\angle}$ , with  $h_{K_{\mathbf{x}}} = \text{diam}(K_{\mathbf{x}})$  and  $h_n = |I_n|$ . Denote by  $\Pi_{(0,p_t)}$ ,  $p_t \in \mathbb{N}_0$ , the  $L^2(K)$ -orthogonal projection:  $L^2(K) \rightarrow \mathbb{P}^0(K_{\mathbf{x}}) \otimes \mathbb{P}^{p_t}(I_n)$ .*

*Let  $\varphi \in H^{k_t+1}(I_n; L^2(K_{\mathbf{x}})) \cap L^2(I_n; H_{\delta}^{1,1}(K_{\mathbf{x}}))$ ,  $k_t \in \mathbb{N}_0$ . Then*

$$\begin{aligned} \left\| \varphi - \Pi_{(0,p_t)} \varphi \right\|_{L^2(K_{\mathbf{x}})} + h_n \left\| \varphi - \Pi_{(0,p_t)} \varphi \right\|_{H^1(I_n; L^2(K_{\mathbf{x}}))} + h_{K_{\mathbf{x}}}^{1-\delta_K} \left\| \varphi - \Pi_{(0,p_t)} \varphi \right\|_{L^2(I_n; H_{\delta}^{1,1}(K_{\mathbf{x}}))} \\ \lesssim h_n^{s_t+1} |\varphi|_{H^{s_t+1}(I_n; L^2(K_{\mathbf{x}}))} + h_{K_{\mathbf{x}}}^{1-\delta_K} |\varphi|_{L^2(I_n; H_{\delta}^{1,1}(K_{\mathbf{x}}))}, \end{aligned} \quad (44)$$

where  $s_t := \min\{k_t, p_t\}$ , and the hidden constants are independent of  $h_{K_{\mathbf{x}}}$  and  $h_n$ , but depend on the polynomial degree  $p_t$  and on the shape of  $K_{\mathbf{x}}$ .

*Proof.* It suffices to repeat the reasoning in the proof of Proposition 6.2, using the compactness of the embedding  $H_{\delta}^{1,1}(K_{\mathbf{x}}) \subset L^2(K_{\mathbf{x}})$  recalled in (10) (see [45, Proposition A.2.5]).  $\square$

### 6.3 Galerkin error estimates

As in §5.3 and §6.2, we start by considering the case where the problem data  $f, v_0, \boldsymbol{\sigma}_0, g_D, g_N$  are such that the solution  $(v, \boldsymbol{\sigma})$  does not present singularities due to the domain corners, then we consider the general case.

#### 6.3.1 The smooth solution case

**Proposition 6.5.** *Let  $(v, \boldsymbol{\sigma})$  and  $(v_h, \boldsymbol{\sigma}_h) \in \mathbf{V}_\mathbf{p}(\mathcal{T}_h)$  be the solutions of (1) and (16), respectively. Assume that*

$$\begin{aligned} v|_K &\in H^{k_{t,K}^v+1}(I_{n'}; H^2(K_\mathbf{x})) \cap L^2(I_{n'}; H^{k_{x,K}^v+1}(K_\mathbf{x})) \quad \forall K = K_\mathbf{x} \times I_{n'} \in \mathcal{T}_h, \\ \boldsymbol{\sigma}|_K &\in H^{k_{t,K}^\sigma+1}(I_{n'}; L^2(K_\mathbf{x})^2) \cap L^2(I_{n'}; H^{k_{x,K}^\sigma+1}(K_\mathbf{x})^2) \quad \forall K = K_\mathbf{x} \times I_{n'} \in \mathcal{T}_h, \end{aligned}$$

and  $\mathbf{V}_\mathbf{p}(\mathcal{T}_h)$  as in (31), with

$$k_{t,K}^v, k_{x,K}^v, k_{t,K}^\sigma, k_{x,K}^\sigma, p_{t,K}^v, p_{x,K}^v, p_{t,K}^\sigma, p_{x,K}^\sigma \in \mathbb{N}_0, \quad |p_{x,K}^\sigma - p_{x,K}^v| \leq 1, \quad p_{t,K}^\sigma = p_{t,K}^v.$$

Define

$$s_{t,K}^v := \min\{k_{t,K}^v, p_{t,K}^v\}, \quad s_{x,K}^v := \min\{k_{x,K}^v, p_{x,K}^v\}, \quad s_{t,K}^\sigma := \min\{k_{t,K}^\sigma, p_{t,K}^\sigma\}, \quad s_{x,K}^\sigma := \min\{k_{x,K}^\sigma, p_{x,K}^\sigma\}.$$

Then,

$$\begin{aligned} &\frac{1}{2} \|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times \{t_n\})^2} \leq |(v, \boldsymbol{\sigma}) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \\ &\lesssim \sum_{\substack{K=K_\mathbf{x} \times I_{n'} \\ \in \mathcal{T}_h(Q_n)}} \left[ (h_{n'}^{-1/2} c_K^{-1} + h_{K_\mathbf{x}}^{-1/2} \mathbf{a}_K) \left( h_{n'}^{s_{t,K}^v+1} h_{K_\mathbf{x}} |v|_{H^{s_{t,K}^v+1}(I_{n'}; H^2(K_\mathbf{x}))} + h_{K_\mathbf{x}}^{s_{x,K}^v+1} |v|_{L^2(I_{n'}; H^{s_{x,K}^v+1}(K_\mathbf{x}))} \right) \right. \\ &\quad \left. + (h_{n'}^{-1/2} + h_{K_\mathbf{x}}^{-1/2} \mathbf{b}_K) \left( h_{n'}^{s_{t,K}^\sigma+1} |\boldsymbol{\sigma}|_{H^{s_{t,K}^\sigma+1}(I_{n'}; L^2(K_\mathbf{x})^2)} + h_{K_\mathbf{x}}^{s_{x,K}^\sigma+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{x,K}^\sigma+1}(K_\mathbf{x})^2)} \right) \right]. \end{aligned}$$

The hidden constant depends on the spatial and temporal polynomial degrees, the regularity exponents and the spatial element shapes.

*Proof.* The DG scheme (16) is well-posed because the conditions on the polynomial degrees defining  $\mathbf{V}_\mathbf{p}(\mathcal{T}_h)$  ensure that condition (17) is satisfied, namely that the (piecewise) first-order wave operator maps the discrete space into itself. We use the quasi-optimality result of Proposition 5.5, the bound (39) on the  $|\cdot|_{\text{DG}(Q_n)_+}$  seminorm of the  $L^2$  projection error  $(w, \boldsymbol{\tau})|_K := (v, \boldsymbol{\sigma})|_K - (\Pi_{\mathbf{p}_K^v} v|_K, \Pi_{\mathbf{p}_K^\sigma} \boldsymbol{\sigma}|_K)$ ,  $\mathbf{p}_K^v = (p_{x,K}^v, p_{t,K}^v)$  and  $\mathbf{p}_K^\sigma = (p_{x,K}^\sigma, p_{t,K}^\sigma)$  (neglecting the summation over elements in  $\mathcal{T}_h^\triangleleft(Q_n)$ , and replacing  $\mathcal{T}_h^\circ(Q_n)$  by  $\mathcal{T}_h(Q_n)$ ), and the projection estimate (40) applied to each component of  $(w, \boldsymbol{\tau})$

$$\begin{aligned} &\frac{1}{2} \|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times \{t_n\})^2} \leq |(v, \boldsymbol{\sigma}) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \stackrel{(26)}{\lesssim} |(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)_+} \\ &\stackrel{(39)}{\lesssim} \left( \sum_{K=K_\mathbf{x} \times I_{n'} \in \mathcal{T}_h(Q_n)} \left[ h_{n'}^{-1} \left( \|c_K^{-1}(v - \Pi_{\mathbf{p}_K^v} v)\|_{L^2(K)}^2 + \|\boldsymbol{\sigma} - \Pi_{\mathbf{p}_K^\sigma} \boldsymbol{\sigma}\|_{L^2(K)^2} \right) \right. \right. \\ &\quad \left. \left. + h_{n'} \left( \|c_K^{-1}(v - \Pi_{\mathbf{p}_K^v} v)\|_{H^1(I_{n'}; L^2(K_\mathbf{x}))}^2 + \|\boldsymbol{\sigma} - \Pi_{\mathbf{p}_K^\sigma} \boldsymbol{\sigma}\|_{H^1(I_{n'}; L^2(K_\mathbf{x})^2)}^2 \right) \right. \right. \\ &\quad \left. \left. + h_{K_\mathbf{x}}^{-1} \|\mathbf{a}_K(v - \Pi_{\mathbf{p}_K^v} v)\|_{L^2(K)}^2 + h_{K_\mathbf{x}} \|\mathbf{a}_K(v - \Pi_{\mathbf{p}_K^v} v)\|_{L^2(I_{n'}; H^1(K_\mathbf{x}))}^2 + h_{K_\mathbf{x}}^{-1} \|\mathbf{b}_K(\boldsymbol{\sigma} - \Pi_{\mathbf{p}_K^\sigma} \boldsymbol{\sigma})\|_{L^2(K)^2} \right] \right. \\ &\quad \left. + \sum_{K=K_\mathbf{x} \times I_{n'} \in \mathcal{T}_h(Q_n)} h_{K_\mathbf{x}} \|\mathbf{b}_K(\boldsymbol{\sigma} - \Pi_{\mathbf{p}_K^\sigma} \boldsymbol{\sigma})\|_{L^2(I_{n'}; H^1(K_\mathbf{x})^2)}^2 \right)^{1/2} \\ &\stackrel{(40)}{\lesssim} \left( \sum_{K=K_\mathbf{x} \times I_{n'} \in \mathcal{T}_h(Q_n)} (h_{n'}^{-1} c_K^{-2} + h_{K_\mathbf{x}}^{-1} \mathbf{a}_K^2) \left( h_{n'}^{s_{t,K}^v+1} |v|_{H^{s_{t,K}^v+1}(I_{n'}; L^2(K_\mathbf{x}))} + h_{K_\mathbf{x}}^{s_{x,K}^v+1} |v|_{L^2(I_{n'}; H^{s_{x,K}^v+1}(K_\mathbf{x}))} \right) \right)^2 \\ &\quad \left. + \sum_{K=K_\mathbf{x} \times I_{n'} \in \mathcal{T}_h(Q_n)} (h_{n'}^{-1} + h_{K_\mathbf{x}}^{-1} \mathbf{b}_K^2) \left( h_{n'}^{s_{t,K}^\sigma+1} |\boldsymbol{\sigma}|_{H^{s_{t,K}^\sigma+1}(I_{n'}; L^2(K_\mathbf{x})^2)} + h_{K_\mathbf{x}}^{s_{x,K}^\sigma+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{x,K}^\sigma+1}(K_\mathbf{x})^2)} \right) \right)^2 \end{aligned}$$



$$\lesssim \left( \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} (h_{n'}^{-1} c_K^{-2} + h_{K_{\mathbf{x}}}^{-1} \mathbf{a}_K^2) \left( h_{n'}^{s_{t,K}^v+1} h_{K_{\mathbf{x}}} |v|_{H^{s_{t,K}^v+1}(I_{n'}; L^\infty(K_{\mathbf{x}}))} + h_{K_{\mathbf{x}}}^{s_{x,K}^v+1} |v|_{L^2(I_{n'}; H^{s_{x,K}^v+1}(K_{\mathbf{x}}))} \right)^2 \right. \\ \left. + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} (h_{n'}^{-1} + h_{K_{\mathbf{x}}}^{-1} \mathbf{b}_K^2) \left( h_{n'}^{s_{t,K}^\sigma+1} |\boldsymbol{\sigma}|_{H^{s_{t,K}^\sigma+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)} + h_{K_{\mathbf{x}}}^{s_{x,K}^\sigma+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{x,K}^\sigma+1}(K_{\mathbf{x}})^2)} \right)^2 \right)^{1/2},$$

where in the last inequality we have used the inclusion  $H^2(K_{\mathbf{x}}) \subset C^0(\overline{K_{\mathbf{x}}})$ , and  $|v|_{H^{s_{t,K}^v+1}(I_{n'}; L^2(K_{\mathbf{x}}))} \lesssim h_{K_{\mathbf{x}}} |v|_{H^{s_{t,K}^v+1}(I_{n'}; L^\infty(K_{\mathbf{x}}))}$ . The assertion follows by using again the inclusion  $H^2(K_{\mathbf{x}}) \subset C^0(\overline{K_{\mathbf{x}}})$ .  $\square$

Proposition 6.5 provides the same error convergence rates in both the  $L^2(\Omega)$  norm at a discrete time  $t_n$  and in the  $|\cdot|_{\text{DG}(Q_n)}$  seminorm. However, one would expect higher rates in the former norm: when the mesh is refined uniformly, in the computation of the  $|\cdot|_{\text{DG}(Q_n)}$  seminorm the error is integrated on larger and larger domains (the mesh skeletons). This suggests that the convergence rates with respect to the mesh size for the  $L^2(\Omega \times \{t_n\})$  norm of the error are suboptimal. The same reasoning applies to the bounds in Propositions 6.7 and 7.3.

Without using the information on the elementwise regularity of  $(v, \boldsymbol{\sigma})$ , but considering only the global regularity, the statement of Proposition 6.5 simplifies as in the next corollary.

**Corollary 6.6.** *Assume that, for the IBVP solution  $(v, \boldsymbol{\sigma})$  we have*

$$(v, \boldsymbol{\sigma}) \in C^{k_t-1}([0, T]; H^{k_x+1}(\Omega)) \times C^{k_t}([0, T]; H^{k_x}(\Omega)^2),$$

for some  $k_x, k_t \in \mathbb{N}$  and  $k_t \geq 2$ . Assume also that the space  $\mathbf{V}_{\mathbf{p}}(\mathcal{T}_h)$  and the local polynomial degrees are defined as in Proposition 6.5. Then the assertion of Proposition 6.5 holds true with

$$k_{t,K}^v = k_t - 2, \quad k_{x,K}^v = k_x, \quad k_{t,K}^\sigma = k_t - 1, \quad k_{x,K}^\sigma = k_x - 1 \quad \forall K \in \mathcal{T}_h.$$

Assuming in particular the velocity  $c$  to be constant in  $Q$ , and that the mesh satisfies the quasi-uniformity condition

$$h := \max \left\{ \max_{K \in \mathcal{T}_h(Q_n)} h_{K_{\mathbf{x}}}, \max_{n'=1, \dots, n} ch_{n'} \right\} \leq \rho \min \left\{ \min_{K \in \mathcal{T}_h(Q_n)} h_{K_{\mathbf{x}}}, \min_{n'=1, \dots, n} ch_{n'} \right\}$$

for some  $\rho \geq 1$ , that the polynomial degrees are uniform in space and time, i.e.  $p_{t,K}^v = p_{x,K}^v = p_{t,K}^\sigma = p_{x,K}^\sigma = p$ , with the numerical flux parameters chosen as  $\alpha^{-1} = \beta = c$ , then

$$\frac{1}{2} \|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times \{t_n\})^2} \leq |(v, \boldsymbol{\sigma}) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \\ \lesssim C(v, \boldsymbol{\sigma}) h^{\min\{k_t - \frac{1}{2}, k_x - \frac{1}{2}, p + \frac{1}{2}\}},$$

where  $C(v, \boldsymbol{\sigma})$  depends on the mesh  $\mathcal{T}_h$  only through the quasi-uniformity parameter  $\rho$  and the shape of the elements.

### 6.3.2 The general case

**Proposition 6.7.** *Let  $(v, \boldsymbol{\sigma})$  and  $(v_h, \boldsymbol{\sigma}_h) \in \mathbf{V}_{\mathbf{p}}(\mathcal{T}_h)$  be the solutions of (1) and (16), respectively. Assume that*

$$\begin{aligned} v|_K &\in H^{k_{t,K}^v+1}(I_{n'}; H_{\delta}^{2,2}(K_{\mathbf{x}})) \cap L^2(I_{n'}; H^{k_{x,K}^v+1}(K_{\mathbf{x}})) & \forall K = K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h, \\ \boldsymbol{\sigma}|_K &\in H^{k_{t,K}^\sigma+1}(I_{n'}; L^2(K_{\mathbf{x}})^2) \cap L^2(I_{n'}; H^{k_{x,K}^\sigma+1}(K_{\mathbf{x}})^2) & \forall K = K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\circ, \\ \boldsymbol{\sigma}|_K &\in H^{k_{t,K}^\sigma+1}(I_{n'}; L^2(K_{\mathbf{x}})^2) \cap L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2) & \forall K = K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\angle, \end{aligned}$$

with  $\mathbf{V}_{\mathbf{p}}(\mathcal{T}_h)$  as in (31) and the indices  $k_{b,K}^a, p_{b,K}^a, s_{b,K}^a$  (for  $a \in \{v, \boldsymbol{\sigma}\}$ ,  $b \in \{t, x\}$ ,  $K \in \mathcal{T}_h(Q_n)$ ) as in Proposition 6.5. Then

$$\frac{1}{2} \|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times \{t_n\})^2} \leq |(v, \boldsymbol{\sigma}) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \\ \lesssim \sum_{\substack{K=K_{\mathbf{x}} \times I_{n'} \\ \in \mathcal{T}_h(Q_n)}} (h_{n'}^{-1/2} c_K^{-1} + h_{K_{\mathbf{x}}}^{-1/2} \mathbf{a}_K) \left( h_{n'}^{s_{t,K}^v+1} h_{K_{\mathbf{x}}} |v|_{H^{s_{t,K}^v+1}(I_{n'}; H_{\delta}^{2,2}(K_{\mathbf{x}}))} + h_{K_{\mathbf{x}}}^{s_{x,K}^v+1} |v|_{L^2(I_{n'}; H^{s_{x,K}^v+1}(K_{\mathbf{x}}))} \right)$$

$$\begin{aligned}
& + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\circ(Q_n)} (h_{n'}^{-1/2} + h_{K_{\mathbf{x}}}^{-1/2} \mathbf{b}_K) \left( h_{n'}^{s_{t,K}^{\sigma}+1} |\boldsymbol{\sigma}|_{H^{s_{t,K}^{\sigma}+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)} + h_{K_{\mathbf{x}}}^{s_{x,K}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{x,K}^{\sigma}+1}(K_{\mathbf{x}})^2)} \right) \\
& + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\angle(Q_n)} (h_{n'}^{-1/2} + h_{K_{\mathbf{x}}}^{-1/2} \mathbf{b}_K) \left( h_{n'}^{s_{t,K}^{\sigma}+1} |\boldsymbol{\sigma}|_{H^{s_{t,K}^{\sigma}+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)} + h_{K_{\mathbf{x}}}^{1-\delta_K} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)} \right).
\end{aligned}$$

As in Proposition 6.5, the hidden constant depends on the spatial and temporal polynomial degrees, the solution regularity indices and on the shape-regularity constant of the spatial elements.

*Proof.* We proceed as in the proof of Proposition 6.5. By using the quasi-optimality result of Proposition 5.6, the bound (39), and the projection estimates (40) and (44), we get

$$\begin{aligned}
& \frac{1}{2} \|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times \{t_n\})^2} \leq |(v, \boldsymbol{\sigma}) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \stackrel{(30)}{\lesssim} |(w, \boldsymbol{\tau})|_{\text{DG}(Q_n)} + \\
& \stackrel{(39)}{\lesssim} \left( \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} \left[ h_{n'}^{-1} \left( \|c_K^{-1}(v - \Pi_{\mathbf{P}_K^v} v)\|_{L^2(K)}^2 + \|\boldsymbol{\sigma} - \Pi_{\mathbf{P}_K^{\sigma}} \boldsymbol{\sigma}\|_{L^2(K)}^2 \right) \right. \right. \\
& \quad + h_{n'} \left( \|c_K^{-1}(v - \Pi_{\mathbf{P}_K^v} v)\|_{H^1(I_{n'}; L^2(K_{\mathbf{x}}))}^2 + \|\boldsymbol{\sigma} - \Pi_{\mathbf{P}_K^{\sigma}} \boldsymbol{\sigma}\|_{H^1(I_{n'}; L^2(K_{\mathbf{x}})^2)}^2 \right) \\
& \quad \left. \left. + h_{K_{\mathbf{x}}}^{-1} \|\mathbf{a}_K(v - \Pi_{\mathbf{P}_K^v} v)\|_{L^2(K)}^2 + h_{K_{\mathbf{x}}} |\mathbf{a}_K(v - \Pi_{\mathbf{P}_K^v} v)|_{L^2(I_{n'}; H^1(K_{\mathbf{x}}))}^2 + h_{K_{\mathbf{x}}}^{-1} \|\mathbf{b}_K(\boldsymbol{\sigma} - \Pi_{\mathbf{P}_K^{\sigma}} \boldsymbol{\sigma})\|_{L^2(K)}^2 \right] \right. \\
& \quad + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\circ(Q_n)} h_{K_{\mathbf{x}}} |\mathbf{b}_K(\boldsymbol{\sigma} - \Pi_{\mathbf{P}_K^{\sigma}} \boldsymbol{\sigma})|_{L^2(I_{n'}; H^1(K_{\mathbf{x}})^2)}^2 \\
& \quad \left. + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\angle(Q_n)} h_{K_{\mathbf{x}}}^{1-2\delta_K} |\mathbf{b}_K(\boldsymbol{\sigma} - \Pi_{(0,p_{t,K}^{\sigma})} \boldsymbol{\sigma})|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)}^2 \right)^{1/2} \\
& \stackrel{(40),(44)}{\lesssim} \left( \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} (h_{n'}^{-1} c_K^{-2} + h_{K_{\mathbf{x}}}^{-1} \mathbf{a}_K^2) \left( h_{n'}^{s_{t,K}^v+1} |v|_{H^{s_{t,K}^v+1}(I_{n'}; L^2(K_{\mathbf{x}}))} + h_{K_{\mathbf{x}}}^{s_{x,K}^v+1} |v|_{L^2(I_{n'}; H^{s_{x,K}^v+1}(K_{\mathbf{x}}))} \right) \right)^2 \\
& \quad + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\circ(Q_n)} (h_{n'}^{-1} + h_{K_{\mathbf{x}}}^{-1} \mathbf{b}_K^2) \left( h_{n'}^{s_{t,K}^{\sigma}+1} |\boldsymbol{\sigma}|_{H^{s_{t,K}^{\sigma}+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)} + h_{K_{\mathbf{x}}}^{s_{x,K}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{x,K}^{\sigma}+1}(K_{\mathbf{x}})^2)} \right)^2 \\
& \quad + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\angle(Q_n)} (h_{n'}^{-1} + h_{K_{\mathbf{x}}}^{-1} \mathbf{b}_K^2) \left( h_{n'}^{s_{t,K}^{\sigma}+1} |\boldsymbol{\sigma}|_{H^{s_{t,K}^{\sigma}+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)} + h_{K_{\mathbf{x}}}^{1-\delta_K} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)} \right)^2 \Big)^{1/2} \\
& \lesssim \left( \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} (h_{n'}^{-1} c_K^{-2} + h_{K_{\mathbf{x}}}^{-1} \mathbf{a}_K^2) \left( h_{n'}^{s_{t,K}^v+1} h_{K_{\mathbf{x}}} |v|_{H^{s_{t,K}^v+1}(I_{n'}; L^\infty(K_{\mathbf{x}}))} + h_{K_{\mathbf{x}}}^{s_{x,K}^v+1} |v|_{L^2(I_{n'}; H^{s_{x,K}^v+1}(K_{\mathbf{x}}))} \right) \right)^2 \\
& \quad + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\circ(Q_n)} (h_{n'}^{-1} + h_{K_{\mathbf{x}}}^{-1} \mathbf{b}_K^2) \left( h_{n'}^{s_{t,K}^{\sigma}+1} |\boldsymbol{\sigma}|_{H^{s_{t,K}^{\sigma}+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)} + h_{K_{\mathbf{x}}}^{s_{x,K}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{x,K}^{\sigma}+1}(K_{\mathbf{x}})^2)} \right)^2 \\
& \quad + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\angle(Q_n)} (h_{n'}^{-1} + h_{K_{\mathbf{x}}}^{-1} \mathbf{b}_K^2) \left( h_{n'}^{s_{t,K}^{\sigma}+1} |\boldsymbol{\sigma}|_{H^{s_{t,K}^{\sigma}+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)} + h_{K_{\mathbf{x}}}^{1-\delta_K} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)} \right)^2 \Big)^{1/2},
\end{aligned}$$

where in the last inequality we have used the inclusion  $H_{\delta}^{2,2}(K_{\mathbf{x}}) \subset C^0(\overline{K_{\mathbf{x}}})$  (see [2, Formula (2.2)], recalled in [45, Remark 1.2.2 b)], and  $|v|_{H^{s_{t,K}^v+1}(I_{n'}; L^2(K_{\mathbf{x}}))} \lesssim h_{K_{\mathbf{x}}} |v|_{H^{s_{t,K}^v+1}(I_{n'}; L^\infty(K_{\mathbf{x}}))}$ . The assertion follows by using again the inclusion  $H_{\delta}^{2,2}(K_{\mathbf{x}}) \subset C^0(\overline{K_{\mathbf{x}}})$ .  $\square$

When the solution exhibits point singularities, either at corners of  $\Omega$  or at multimaterial interfaces in the interior of  $\Omega$ , the convergence rate of discretizations on quasi-uniform families of partitions of  $\Omega$  is well known to degrade, due to the limited solution regularity in scales of standard Sobolev spaces. Local mesh refinement, with the local refinement taking place in space only, can restore the largest possible convergence rates. We describe suitable mesh families in the next section, and derive the corresponding orders of convergence in Proposition 7.4. We remark that, if the numerical flux coefficients  $\alpha^{-1}$  and  $\beta$  are chosen to be proportional to the ratio between global and local space mesh size, the estimate in Proposition 6.7 guarantees that, under local spatial mesh refinement, the solution remains bounded. On the other hand, numerical results in §8.2 below seem to suggest that this condition is actually not necessary.

## 7 Error bounds on locally refined meshes

For each corner  $\mathbf{c} \in \mathcal{S}$ , define the distance

$$R_{\mathbf{c}} := \frac{1}{2} \min_{\mathbf{c}' \in \mathcal{S} \setminus \{\mathbf{c}\}} |\mathbf{c} - \mathbf{c}'| \quad (45)$$

and let  $\text{dist}(K_{\mathbf{x}}, \mathbf{c})$  be the spatial distance of the element  $K_{\mathbf{x}}$  from the corner  $\mathbf{c}$ . Assume  $p_{x,K}^{\sigma} = p_x^{\sigma} \in \mathbb{N}_0$  for all elements  $K$  of any given mesh.

We use the algorithm proposed by Gaspoz and Morin [17] to generate the meshes with appropriate grading towards the corners of a polygon:

**Definition 7.1** (Corner-refined, regular simplicial triangulation of  $\Omega$ ). *In a polygon  $\Omega$  with  $\text{diam}(\Omega) \leq 1$ , a corner-refined, regular simplicial triangulation  $\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}}$  of meshwidth  $h_{\mathbf{x}} \in (0, 1]$  for spatial polynomial degree  $p_{\mathbf{x}} \geq 1$  and corner weights  $\{\delta_{\mathbf{c}} : \mathbf{c} \in \mathcal{S}\} \subset [0, 1)^M$  as in §3 is a regular partition obtained by the following procedure. Let  $J \in \mathbb{N}_0$  be such that for each  $\mathbf{c} \in \mathcal{S}$*

$$2^{-\frac{(1-\delta_{\mathbf{c}})(J+1)}{p_{\mathbf{x}}^{\sigma}+1}} \leq h_{\mathbf{x}} \lesssim 2^{-\frac{(1-\delta_{\mathbf{c}})J}{p_{\mathbf{x}}^{\sigma}+1}}. \quad (46)$$

with the hidden constant being independent of  $h_{\mathbf{x}}$ . Corner-refined, regular simplicial triangulations of  $\Omega$  are constructed by regular refinement of the regular, simplicial initial triangulation  $\mathcal{T}_0^{\mathbf{x}}$  using Algorithms 1 and 2, and newest vertex bisection [35, Section 3.3.1].

---

**Algorithm 1:** `bisection_tree_refinement`( $\mathcal{T}_0^{\mathbf{x}}, h_{\mathbf{x}}, \delta, p_x^{\sigma}, R_{\mathbf{c}}, \text{bool\_conforming}$ )

---

**Data:**  $\mathcal{T}_0^{\mathbf{x}}, h_{\mathbf{x}} > 0, \delta \in [0, 1)^M, p_x^{\sigma} \in \mathbb{N}_0, \{R_{\mathbf{c}}\}_{\mathbf{c} \in \mathcal{S}}, \text{bool\_conforming}$   
**Result:** Mesh  $\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}}$  with local refinement towards corners  
*// Step 1: refine  $\mathcal{T}_0^{\mathbf{x}}$  until  $h_{K_{\mathbf{x}}} < h_{\mathbf{x}} \forall K_{\mathbf{x}} \in \mathcal{T}_0^{\mathbf{x}}$*   
 $\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} := \mathcal{T}_0^{\mathbf{x}}$   
 $\mathcal{M} := \{K_{\mathbf{x}} \in \mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} \mid h_{K_{\mathbf{x}}} > h_{\mathbf{x}}\}$   
**while**  $\mathcal{M} \neq \emptyset$  **do**  
   $\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} := \text{bisect\_marked}(\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}}, \mathcal{M}, \text{bool\_conforming})$   
   $\mathcal{M} := \{K_{\mathbf{x}} \in \mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} \mid h_{K_{\mathbf{x}}} > h_{\mathbf{x}}\}$   
**end**  
*// Step 2: local refinement in nested discs*  
**for**  $\mathbf{c} \in \mathcal{S}$  **do**  
   $J := \lceil -\log_2(h_{\mathbf{x}}) \frac{p_x^{\sigma}+1}{1-\delta_{\mathbf{c}}} - 1 \rceil$   
  **for**  $j = 0, 1, \dots, 2J + 1$  **do**  
     $\mathcal{M} := \{K_{\mathbf{x}} \in \mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} \mid \text{dist}(K_{\mathbf{x}}, \mathbf{c}) \leq 2^{-j/2} R_{\mathbf{c}} \wedge h_{K_{\mathbf{x}}} > h_{\mathbf{x}} 2^{-j(p_x^{\sigma} + \delta_{\mathbf{c}})/(2(p_x^{\sigma} + 1))}\}$   
     $\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} := \text{bisect\_marked}(\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}}, \mathcal{M}, \text{bool\_conforming})$   
  **end**  
**end**

---

**Algorithm 2:** `bisect_marked`( $\mathcal{T}_0^{\mathbf{x}}, \mathcal{M}, \text{bool\_conforming}$ )

---

**Data:**  $\mathcal{T}_0^{\mathbf{x}}, \mathcal{M} \subset \mathcal{T}_0^{\mathbf{x}}, \text{bool\_conforming}$   
**Result:**  $\mathcal{T}_j^{\mathbf{x}}$   
**if** `bool_conforming` **then**  
   $j := 0$   
   $\mathcal{M}_0 := \mathcal{M}$   
  **while**  $\mathcal{M}_j \neq \emptyset$  **do**  
    **for**  $K_{\mathbf{x}} \in \mathcal{M}_j$  **do**  
       $\mathcal{T}_{j+1}^{\mathbf{x}} := \text{newest\_vertex\_bisection}(K_{\mathbf{x}}, \mathcal{T}_j^{\mathbf{x}})$   
    **end**  
     $\mathcal{M}_{j+1} := \{K_{\mathbf{x}} \in \mathcal{T}_{j+1}^{\mathbf{x}} \mid K_{\mathbf{x}} \text{ has a hanging node}\}$   
     $j := j + 1$   
  **end**  
  Set  $J := j$   
**else**  
  **for**  $K_{\mathbf{x}} \in \mathcal{M}$  **do**  
     $\mathcal{T}_j^{\mathbf{x}} := \text{newest\_vertex\_bisection}(K_{\mathbf{x}}, \mathcal{T}_0^{\mathbf{x}})$   
  **end**  
**end**

---

**Remark 7.2.** Algorithms 1 and 2 construct a spatial mesh of meshwidth at most  $h_{\mathbf{x}}$  obtained as bisection tree refinement of a regular initial partition  $\mathcal{T}_0^{\mathbf{x}}$  of  $\Omega$ . For  $j \leq 2J + 1$  and  $J = \lceil -\log_2(h_{\mathbf{x}}) \frac{p_x^\sigma + 1}{1 - \delta_c} - 1 \rceil$ , the elements at distance not exceeding  $2^{-j/2} R_{\mathbf{c}}$  (cf. (45)) from the closest vertex  $\mathbf{c}$  have diameter at most  $h_{\mathbf{x}} 2^{-\frac{j(p_x^\sigma + \delta_c)}{2(p_x^\sigma + 1)}}$ . Note that at fixed target meshwidth  $h_{\mathbf{x}}$ ,  $J \uparrow \infty$  as  $\delta_c \uparrow 1$  (cf. Remark 3.2). The flag “bool\_conforming” in Algorithm 1 allows to choose between meshes with or without hanging nodes.

If one knows a priori that, in a neighbourhood of a corner  $\mathbf{c}_i \in \mathcal{S}$ , the IBVP solution  $(v, \boldsymbol{\sigma})$  has regularity better than that associated to the exponent  $\delta_i$ , then one can choose a smaller value of  $\delta_i$  in (46) and in the choice of  $J$  in Algorithm 1. This does not alter the asymptotic complexity of the algorithm, but allows to use fewer local refinement levels.

**Proposition 7.3.** Let  $p_x^\sigma \in \mathbb{N}$ ,  $h_{\mathbf{x}} > 0$ ,  $\boldsymbol{\delta} \in [0, 1)^M$  and  $\mathcal{T}_0^{\mathbf{x}}$  be an initial regular triangulation of the domain  $\Omega$ . For the mesh  $\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} := \text{bisection\_tree\_refinement}(\mathcal{T}_0^{\mathbf{x}}, h_{\mathbf{x}}, \boldsymbol{\delta}, p_x^\sigma, R_{\mathbf{c}}, \cdot)$ , there holds the following bound:

$$\#\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} - \#\mathcal{T}_0^{\mathbf{x}} \leq C h_{\mathbf{x}}^{-2}, \quad (47)$$

where the constant  $C$  is independent of  $h_{\mathbf{x}}$  but depends on  $\boldsymbol{\delta}$ . Furthermore, the constant depends on  $\mathcal{T}_0^{\mathbf{x}}$  and increases unboundedly if  $\delta_c \rightarrow 1$  (see Remark 3.2) or if  $p_x^\sigma \rightarrow \infty$ .

*Proof.* This complexity estimate has been proved in [17, Lemma 4.9] for conforming meshes. The estimate for non-conforming meshes follows as these have, asymptotically for  $h_{\mathbf{x}} \rightarrow 0$ , the same number of elements.  $\square$

For solutions with point singularities and for locally refined triangulations of  $\Omega$  spatial meshes constructed with Algorithms 1 and 2, we now prove error bounds with the same dependence on the meshwidth  $h$  as the bounds proved in Corollary 6.6 for smooth solutions.

**Proposition 7.4.** Assume that the IBVP solution  $(v, \boldsymbol{\sigma})$  admits the regularity in (8) for some smoothness orders  $k_x, k_t \in \mathbb{N}$ ,  $k_t \geq 2$  and  $\boldsymbol{\delta} \in [0, 1)^M$ . Furthermore, assume that the space  $\mathbf{V}_{\mathbf{p}}(\mathcal{T}_h)$  and the local polynomial degrees are defined as in Proposition 6.5. Given a spatial meshwidth bound  $h_{\mathbf{x}} > 0$  and an initial regular triangulation  $\mathcal{T}_0^{\mathbf{x}}$  of the domain  $\Omega$ , the locally refined mesh  $\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}}$  is generated using Algorithm 1.

In particular, assume that the velocity  $c$  is constant in  $Q$ , the polynomial degrees are  $p_{t,K}^v = p_{x,K}^v = p_{t,K}^\sigma = p_{x,K}^\sigma = p$ , the numerical flux parameters are chosen as

$$\alpha|_F = c^{-1} \frac{h_{\mathbf{x}}}{h_{F_{\mathbf{x}}}} \quad \forall F \in \mathcal{F}_h^{\text{time}} \cap \mathcal{F}_h^{\text{D}}, \quad \beta|_F = c \frac{h_{F_{\mathbf{x}}}}{h_{\mathbf{x}}} \quad \forall F \in \mathcal{F}_h^{\text{time}} \cap \mathcal{F}_h^{\text{N}}$$

We assume that the space-time mesh satisfies the condition

$$h := \max \left\{ h_{\mathbf{x}}, \max_{n'=1, \dots, n} ch_{n'} \right\} \leq \rho \min \left\{ h_{\mathbf{x}}, \min_{n'=1, \dots, n} ch_{n'} \right\}, \quad (48)$$

for each discrete time  $t_n$  and some  $\rho > 1$ . Define  $h_t := \max_{n'=1, \dots, n} h_{n'}$ . Then

$$\begin{aligned} \frac{1}{2} \|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times \{t_n\})} &\leq |(v, \boldsymbol{\sigma}) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}(Q_n)} \\ &\leq C(c, v, \boldsymbol{\sigma}) h^{\min\{k_t - \frac{1}{2}, k_x - \frac{1}{2}, p + \frac{1}{2}\}}, \end{aligned} \quad (49)$$

where  $C(c, v, \boldsymbol{\sigma})$  depends on the space-time mesh  $\mathcal{T}_h$  only through the shape-regularity constant of its spatial elements.

*Proof.* We adapt the proof of Proposition 3.3.9 in [35] to our setting. For  $K_{\mathbf{x}} \in \mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}}$  and  $\mathbf{c} \in \mathcal{S}$ , define

$$r_{\mathbf{c}}(\mathbf{x}) := |\mathbf{x} - \mathbf{c}|, \quad r_{K_{\mathbf{x}}; \mathbf{c}} := \inf_{\mathbf{x} \in K_{\mathbf{x}}} r_{\mathbf{c}}(\mathbf{x}).$$

Without loss of generality, assume  $R_{\mathbf{c}} = 1$  (recall (45)). Given  $\mathbf{c} \in \mathcal{S}$ , partition the mesh  $\mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}}$  into sets

$$\begin{aligned} D_{j; \mathbf{c}} &:= \left\{ K_{\mathbf{x}} \in \mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} : r_{K_{\mathbf{x}}; \mathbf{c}} \in \left( 2^{-\frac{j+1}{2}}, 2^{-\frac{j}{2}} \right] \right\}, \quad j = 0, 1, \dots, 2J + 1, \\ D_{2(J+1); \mathbf{c}} &:= \left\{ K_{\mathbf{x}} \in \mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} : r_{K_{\mathbf{x}}; \mathbf{c}} \in [0, 2^{-(J+1)}] \right\}, \\ \mathcal{K} &:= \mathcal{T}_{h_{\mathbf{x}}}^{\mathbf{x}} \setminus \left( \bigcup_{\mathbf{c} \in \mathcal{S}} \bigcup_{j=0}^{2(J+1)} D_{j; \mathbf{c}} \right). \end{aligned}$$

By the construction of Algorithm 1, all the elements  $K_{\mathbf{x}} \in \{\mathcal{K}, \{D_{j;\mathbf{c}}\}_{j=0,\dots,2J+1}, D_{2(J+1);\mathbf{c}}\}$  satisfy a quasi-uniformity condition within their respective sets, that is for some  $\hat{\rho} > 1$ ,

$$\hat{\rho}^{-1} \widehat{C}_{K_{\mathbf{x}}} h_{\mathbf{x}} \leq h_{K_{\mathbf{x}}} \leq \widehat{C}_{K_{\mathbf{x}}} h_{\mathbf{x}}, \quad \text{where} \quad (50)$$

$$\widehat{C}_{K_{\mathbf{x}}} = \begin{cases} 1 & \forall K_{\mathbf{x}} \in \mathcal{K}, \\ C(j, p, \delta_{\mathbf{c}}) & \forall K_{\mathbf{x}} \in D_{j;\mathbf{c}}, j = 0, \dots, 2J+1, \\ C(2J+1, p, \delta_{\mathbf{c}}) & \forall K_{\mathbf{x}} \in D_{2(J+1);\mathbf{c}}, \end{cases} \quad \text{and} \quad C(j, p, \delta_{\mathbf{c}}) := 2^{-\frac{j(p+\delta_{\mathbf{c}})}{2(p+1)}}.$$

Under the assumptions of the current proposition, recalling the definition of  $\mathbf{a}_K$  and  $\mathbf{b}_K$  in (37), we have

$$c_K = c, \quad \mathbf{a}_K^{-1} = \mathbf{b}_K = \left( \frac{c \max h_{F_{\mathbf{x}}}}{h_{\mathbf{x}}} \right)^{1/2} \stackrel{(38)}{\approx} \left( \frac{ch_{K_{\mathbf{x}}}}{h_{\mathbf{x}}} \right)^{1/2} \quad \forall K \in \mathcal{T}_h,$$

where the maximum is taken over the time-like faces of  $K$ ,  $\mathbf{a}_K$  and  $\mathbf{b}_K$  are defined in (37). Moreover, due to the definition of  $h_t$ , Step 1 of Algorithm 1 (see also (50)), and assumption (48), we have

$$h_{n'} \leq h_t, \quad h_{K_{\mathbf{x}}} \leq h_{\mathbf{x}}, \quad h_{n'}^{-1} \leq \rho ch^{-1}, \quad h_{\mathbf{x}}^{-1} \leq \rho h^{-1}. \quad (51)$$

Additionally, Corollary 6.6 implies that  $\forall K \in \mathcal{T}_h$ ,

$$s_t^v := s_{t,K}^v = \min\{k_t - 2, p\}, \quad s_x^v = s_{x,K}^v = \min\{k_x, p\},$$

$$s_t^\sigma := s_{t,K}^\sigma = \min\{k_t - 1, p\}, \quad s_x^\sigma := s_{x,K}^\sigma = \min\{k_x - 1, p\}.$$

Proceeding as in the proof of Proposition 6.7 and using the information presented above,

$$\begin{aligned} & \frac{1}{2} \|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\sigma - \sigma_h\|_{L^2(\Omega \times \{t_n\})^2} \leq |(v, \sigma) - (v_h, \sigma_h)|_{\text{DG}(Q_n)} \\ & \lesssim \underbrace{\left( \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} c^{-1} h_{\mathbf{x}}^2 (h_{n'}^{-1} c^{-1} + h_{\mathbf{x}}^{-1}) \left( h_t^{s_t^v+1} |v|_{H^{s_t^v+1}(I_{n'}; H_{\delta}^{2,2}(K_{\mathbf{x}}))} + h_{\mathbf{x}}^{s_x^v} |v|_{L^2(I_{n'}; H^{s_x^v+1}(K_{\mathbf{x}}))} \right)^2 \right)}_{(*)_1} \\ & + \underbrace{\sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\circ(Q_n)} c (h_{n'}^{-1} c^{-1} + h_{\mathbf{x}}^{-1}) \left( h_t^{s_t^\sigma+1} |\sigma|_{H^{s_t^\sigma+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)} + h_{K_{\mathbf{x}}}^{s_x^\sigma+1} |\sigma|_{L^2(I_{n'}; H^{s_x^\sigma+1}(K_{\mathbf{x}})^2)} \right)^2}_{(*)_2} \\ & + \underbrace{\sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\angle(Q_n)} c (h_{n'}^{-1} c^{-1} + h_{\mathbf{x}}^{-1}) \left( h_t^{s_t^\sigma+1} |\sigma|_{H^{s_t^\sigma+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)} + h_{K_{\mathbf{x}}}^{1-\delta_K} |\sigma|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)} \right)^2}_{(*)_3} \end{aligned} \quad (52)$$

From (51), we have

$$h_{n'}^{-1} c^{-1} + h_{\mathbf{x}}^{-1} \leq 2\rho h^{-1}.$$

For  $\vec{a} = (a_i), \vec{b} = (b_i) \in \ell_2(\mathbb{N})$ , the triangle inequality in  $\ell_2(\mathbb{N})$  implies  $\sum_i (a_i + b_i)^2 = \|\vec{a} + \vec{b}\|_2^2 \leq (\|\vec{a}\|_2 + \|\vec{b}\|_2)^2$ . Using this inequality, we find

$$\begin{aligned} (*)_1 & = \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} c^{-1} h_{\mathbf{x}}^2 (h_{n'}^{-1} c^{-1} + h_{\mathbf{x}}^{-1}) \left( h_t^{s_t^v+1} |v|_{H^{s_t^v+1}(I_{n'}; H_{\delta}^{2,2}(K_{\mathbf{x}}))} + h_{\mathbf{x}}^{s_x^v} |v|_{L^2(I_{n'}; H^{s_x^v+1}(K_{\mathbf{x}}))} \right)^2 \\ & \lesssim c^{-1} h \left( h_t^{s_t^v+1} |v|_{H^{s_t^v+1}((0,t_n); H_{\delta}^{2,2}(K_{\mathbf{x}}))} + h_{\mathbf{x}}^{s_x^v} |v|_{L^2((0,t_n); H^{s_x^v+1}(\Omega))} \right)^2. \end{aligned}$$

Now, for all  $K = K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^\angle(Q_n)$ ,  $r_{K_{\mathbf{x}};\mathbf{c}} = 0$ , which implies that  $K_{\mathbf{x}}$  belongs to the set  $D_{2(J+1);\mathbf{c}}$ . On element  $K_{\mathbf{x}}$ , we have  $\delta_K = \delta_{\mathbf{c}}$  and, from (50),  $h_{K_{\mathbf{x}}} \leq h_{\mathbf{x}} C(2J+1, p, \delta_{\mathbf{c}})$ . Then, we write

$$\begin{aligned} h_{K_{\mathbf{x}}}^{1-\delta_{\mathbf{c}}} |\sigma|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)} & \stackrel{(50)}{\leq} h_{\mathbf{x}}^{1-\delta_{\mathbf{c}}} 2^{-\frac{(2J+1)(p+\delta_{\mathbf{c}})(1-\delta_{\mathbf{c}})}{2(p+1)}} |\sigma|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)} \\ & = h_{\mathbf{x}}^{1-\delta_{\mathbf{c}}} 2^{-\frac{(2J+1)(p+\delta_{\mathbf{c}})(1-\delta_{\mathbf{c}})}{2(p+1)}} 2^{+\frac{(p+\delta_{\mathbf{c}})(1-\delta_{\mathbf{c}})}{2(p+1)}} |\sigma|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)} \end{aligned}$$

$$\begin{aligned}
&\stackrel{(46)}{\leq} h_{\mathbf{x}}^{1-\delta_{\mathbf{c}}} h_{\mathbf{x}}^{p+\delta_{\mathbf{c}}} \underbrace{2^{\frac{(p+\delta_{\mathbf{c}})(1-\delta_{\mathbf{c}})}{2(p+1)}}}_{=: \tilde{C}(p, \delta_{\mathbf{c}})} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)} \\
&\leq \tilde{C}(p, \delta_{\mathbf{c}}) h_{\mathbf{x}}^{p+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)} \leq \tilde{C}(p, \delta_{\mathbf{c}}) h_{\mathbf{x}}^{p+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)} \\
&\leq \tilde{C}(p, \delta_{\mathbf{c}}) h_{\mathbf{x}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)}.
\end{aligned}$$

Elements  $K = K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^{\circ}(Q_n)$  may belong to either of the sets  $\mathcal{K}$ ,  $D_{2(J+1); \mathbf{c}}$  and  $D_{j; \mathbf{c}}$ ,  $j = 0, \dots, 2J+1$ . Thus, we compute estimates for each case separately.

For elements  $K_{\mathbf{x}} \in \mathcal{K}$ : As  $1 < r_{K_{\mathbf{x}}; \mathbf{c}} < r_{\mathbf{c}}$  and, from (50),  $h_{K_{\mathbf{x}}} \leq h_{\mathbf{x}}$ , we have

$$h_{K_{\mathbf{x}}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{\mathbf{x}}^{\sigma}+1}(K_{\mathbf{x}})^2)} \leq h_{\mathbf{x}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{\mathbf{x}}^{\sigma}+1}(K_{\mathbf{x}})^2)} \leq h_{\mathbf{x}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)}.$$

For elements  $K_{\mathbf{x}} \in D_{2(J+1); \mathbf{c}}$ ,  $r_{K_{\mathbf{x}}; \mathbf{c}} > 0$ : Multiplying and dividing  $|\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{\mathbf{x}}^{\sigma}+1}(K_{\mathbf{x}})^2)}$  by the factor  $r_{K_{\mathbf{x}}; \mathbf{c}}^{\delta_{\mathbf{c}}+s_{\mathbf{x}}^{\sigma}}$  we write it in weighted spaces, and further, using  $h_{K_{\mathbf{x}}} \lesssim r_{K_{\mathbf{x}}; \mathbf{c}}$  from [17, Lemma 4.6] we obtain

$$\begin{aligned}
h_{K_{\mathbf{x}}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{\mathbf{x}}^{\sigma}+1}(K_{\mathbf{x}})^2)} &\lesssim r_{K_{\mathbf{x}}; \mathbf{c}}^{s_{\mathbf{x}}^{\sigma}+1} r_{K_{\mathbf{x}}; \mathbf{c}}^{-\delta_{\mathbf{c}}-s_{\mathbf{x}}^{\sigma}} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)} \\
&\lesssim r_{K_{\mathbf{x}}; \mathbf{c}}^{1-\delta_{\mathbf{c}}} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)} \lesssim 2^{-(1-\delta_{\mathbf{c}})(J+1)} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)} \\
&\stackrel{(46)}{\lesssim} h_{\mathbf{x}}^{p+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)} \lesssim h_{\mathbf{x}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)}.
\end{aligned}$$

For elements  $K_{\mathbf{x}} \in D_{j; \mathbf{c}}$ ,  $j = 0, \dots, 2J+1$ : From (50), we have  $h_{K_{\mathbf{x}}} \leq h_{\mathbf{x}}$   $C(j, p, \delta_{\mathbf{c}})$ , which implies

$$\begin{aligned}
h_{K_{\mathbf{x}}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{\mathbf{x}}^{\sigma}+1}(K_{\mathbf{x}})^2)} &\leq h_{K_{\mathbf{x}}}^{s_{\mathbf{x}}^{\sigma}+1} r_{K_{\mathbf{x}}; \mathbf{c}}^{-\delta_{\mathbf{c}}-s_{\mathbf{x}}^{\sigma}} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)} \\
&\stackrel{(50)}{\leq} h_{\mathbf{x}}^{s_{\mathbf{x}}^{\sigma}+1} \underbrace{2^{-j \frac{(p+\delta_{\mathbf{c}})(s_{\mathbf{x}}^{\sigma}+1)}{2(p+1)}} 2^{j \frac{\delta_{\mathbf{c}}+s_{\mathbf{x}}^{\sigma}}{2}}}_{\leq 1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)} \\
&\leq h_{\mathbf{x}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(K_{\mathbf{x}})^2)}.
\end{aligned}$$

Putting together these estimates and proceeding similarly to  $(*)_1$  yields

$$\begin{aligned}
(*)_2 + (*)_3 &\lesssim ch^{-1} \left( h_t^{s_t^{\sigma}+1} \left( \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h(Q_n)} |\boldsymbol{\sigma}|_{H^{s_t^{\sigma}+1}(I_{n'}; L^2(K_{\mathbf{x}})^2)}^2 \right)^{1/2} \right. \\
&\quad + \left( \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^{\circ}(Q_n)} h_{K_{\mathbf{x}}}^{2(s_{\mathbf{x}}^{\sigma}+1)} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H^{s_{\mathbf{x}}^{\sigma}+1}(K_{\mathbf{x}})^2)}^2 \right. \\
&\quad \left. \left. + \sum_{K=K_{\mathbf{x}} \times I_{n'} \in \mathcal{T}_h^{\angle}(Q_n)} h_{K_{\mathbf{x}}}^{2(1-\delta_{\mathbf{c}})} |\boldsymbol{\sigma}|_{L^2(I_{n'}; H_{\delta}^{1,1}(K_{\mathbf{x}})^2)}^2 \right)^{1/2} \right)^2 \\
&\lesssim ch^{-1} \left( h_t^{s_t^{\sigma}+1} |\boldsymbol{\sigma}|_{H^{s_t^{\sigma}+1}((0, t_n); L^2(\Omega)^2)} + h_{\mathbf{x}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2((0, t_n); H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(\Omega)^2)} \right)^2.
\end{aligned}$$

Finally,

$$\begin{aligned}
&\|c^{-1}(v - v_h)\|_{L^2(\Omega \times \{t_n\})} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times \{t_n\})^2} \\
&\lesssim 2(c^{-1}h)^{1/2} \left( h_t^{s_t^v+1} |v|_{H^{s_t^v+1}((0, t_n); H_{\delta}^{2,2}(K_{\mathbf{x}}))} + h_{\mathbf{x}}^{s_{\mathbf{x}}^v} |v|_{L^2((0, t_n); H^{s_{\mathbf{x}}^v+1}(\Omega))} \right) \\
&\quad + 2(ch^{-1})^{1/2} \left( h_t^{s_t^{\sigma}+1} |\boldsymbol{\sigma}|_{H^{s_t^{\sigma}+1}((0, t_n); L^2(\Omega)^2)} + h_{\mathbf{x}}^{s_{\mathbf{x}}^{\sigma}+1} |\boldsymbol{\sigma}|_{L^2((0, t_n); H_{\delta}^{s_{\mathbf{x}}^{\sigma}+1,1}(\Omega)^2)} \right) \\
&\lesssim C(c, v, \boldsymbol{\sigma}) h^{\min\{k_t - \frac{1}{2}, k_x - \frac{1}{2}, p + \frac{1}{2}\}}.
\end{aligned}$$

□

The space-time discrete space considered in Proposition 7.4 has dimension

$$\dim(\mathbf{V}_{\mathbf{p}}(\mathcal{T}_h)) = (\#\mathcal{T}_h) 3 \frac{(p+1)(p+2)}{2} (p+1),$$



where  $\frac{(p+1)(p+2)}{2}$  is the dimension of the scalar elemental polynomial space in a spatial element,  $p + 1$  is the dimension of the elemental polynomial space in a time interval, 3 is the number of components (i.e.  $v, \sigma_1, \sigma_2$ ). Then, from Proposition 7.3 and the quasi-uniformity of the temporal partition, we have

$$\dim(\mathbf{V}_p(\mathcal{T}_h)) = O(p^3 h_x^{-2} h_t^{-1}) = O(p^3 h^{-3}).$$

## 8 Implementation aspects and numerical experiments

In all numerical experiments presented in this section the following choices have been adopted.

The partition  $\mathcal{T}_h$  for the space–time domain  $Q = \Omega \times I$ , is generated as a Cartesian product of meshes  $\mathcal{T}_{h_x}^x$  in the spatial domain  $\Omega$  and a partition  $\mathcal{T}_{h_t}^t$  of the time interval  $I = (0, T)$ :

$$\mathcal{T}_h := \mathcal{T}_{h_x}^x \times \mathcal{T}_{h_t}^t = \{K = K_x \times I_n, \text{ for } K_x \in \mathcal{T}_{h_x}^x, I_n \in \mathcal{T}_{h_t}^t\}.$$

The temporal mesh  $\mathcal{T}_{h_t}^t$  is uniform, i.e. it consists of  $N_t = 2^{l_t}$  time steps of equal size  $T2^{-l_t}$ , for any given temporal refinement level  $l_t \in \mathbb{N}$ . The spatial mesh  $\mathcal{T}_{h_x}^x$  is a regular mesh of triangles which are shape-regular with respect to the maximum meshwidth of  $\mathcal{T}_{h_x}^x$  given by

$$h_x := \max_{K_x \in \mathcal{T}_{h_x}^x} h_{K_x}.$$

The wave speed is fixed at  $c = 1$ , unless specified otherwise.

The Cartesian tensor-product structure of the space–time mesh allows for an easy implementation of the space–time DG solver as a space-only DG discretization with DG time-stepping. We implement this scheme in Python 3.7.6 using the libraries NumPy 1.16.4, [44], and SciPy 1.2.0, [25], and FEniCS, [1], for the space-only DG discretization. The resulting linear systems of equations are solved using the sparse direct solver `spsolve`, included in the SciPy submodule `scipy.sparse.linalg`. The library FEniCS is also used to generate the uniform meshes in Test 1.1, §8.1, and the local bisection-tree refined meshes in Test 1.2, §8.2.

In our tests, we estimate the order of convergence by measuring the relative error with respect to the exact solution in  $L^2(\Omega)$ -norm at a given time  $t$ , indicated by  $L^2(\Omega \times \{t\})$ -norm.

### 8.1 Test 1: smooth solution, domain $\Omega = (0, 1)^2$

Choose the space–time domain  $Q = (0, 1)^2 \times (0, 1)$ . Consider the exact smooth solution

$$u(\mathbf{x}, t) = \sin(\pi x_1) \sin(\pi x_2) \sin(\sqrt{2}\pi t), \quad \forall (\mathbf{x}, t) \in Q, \quad (52a)$$

$$v = \partial_t u, \quad \boldsymbol{\sigma} = -\nabla_{\mathbf{x}} u, \quad (52b)$$

to the IBVP (1), with homogeneous Dirichlet boundary conditions and source  $f = 0$ .

We use uniform meshes with  $h_x \approx h_t = 2^{-l}$ ,  $l \in \mathbb{N}$ . The convergence rates are given in the Tables 1 and 2 for different choices of the stabilization parameters.

The tables show that for the choice  $\alpha = \beta = 1$  for the DG stabilization parameters and for uniform polynomial degrees  $p = p_v^x = p_v^t = p_\sigma^t = p_\sigma^x$  we obtain  $\|v - v_h\|_{L^2(\Omega \times T)} \approx \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega \times T)^2} = O(h^{p+1})$ . (Recall that here  $c = 1$ : for constant  $c \gg 1$  or  $c \ll 1$  we expect that the parameters need to be scaled as  $\alpha^{-1} = \beta = c$ .) If we over-penalize the spatial jumps of  $v$  by choosing  $\alpha = h_{F_x}^{-1}$ , or under-penalize the normal jumps of  $\boldsymbol{\sigma}$  by choosing  $\beta = h_{F_x}$ , or we combine the two strategies, the error in  $v$  is not substantially affected, while the accuracy in  $\boldsymbol{\sigma}$  is reduced to  $O(h^p)$ . These penalization strategies have previously been proposed for the analogous DG formulation in time-harmonic regime, i.e. for the Helmholtz equation, both for Trefftz and non-Trefftz schemes, see e.g. [20, Table 1] and references therein.

If we reduce by one the spatial polynomial degree of the  $\boldsymbol{\sigma}$  component, namely we take  $p_\sigma^x = p - 1$ , reflecting the mismatch in regularities (e.g. (8)), we observe (Table 2) a degradation of the convergence order in the case  $\alpha = \beta = 1$ . With all other choices of parameters considered, the order of convergence for  $\boldsymbol{\sigma}$  is apparently not affected by this reduction of  $p_\sigma^x$ .

The last columns of Tables 1 and 2 show the values of the errors measured in  $|\cdot|_{\text{DG}}$  seminorm and the corresponding experimental convergence rates. For  $p_\sigma^x = p$ , we observe convergence rates between  $O(h^p)$  and  $O(h^{p+\frac{1}{2}})$ , depending on the choice of the numerical flux parameters, while for  $p_\sigma^x = p - 1$  we observe rates between  $O(h^{p-\frac{1}{2}})$  and  $O(h^p)$ . In particular, the result for  $\alpha = \beta = 1$  and  $p_\sigma^x = p$  demonstrates that the estimate  $|(v, \boldsymbol{\sigma}) - (v_h, \boldsymbol{\sigma}_h)|_{\text{DG}} \lesssim C(v, \boldsymbol{\sigma})h^{p+\frac{1}{2}}$  of Corollary 6.6 is sharp for this test case. In the same way as for the  $L^2(\Omega \times T)$  norms, the choice  $\alpha = \beta$  is preferable for  $p_\sigma^x = p$  but not for  $p_\sigma^x = p - 1$ .

$p$	DOFs	$\alpha$	$\beta$	$v_h$		$\sigma_h$		$ \cdot _{\text{DG}}$	
				Error	Rate	Error	Rate	Error	Rate
1	9437184	1	1	3.1420E-04	2.08	1.8280E-04	2.11	1.2097E-02	1.51
		$h_{F_x}^{-1}$	1	2.1087E-04	2.20	9.8420E-04	1.55	1.2232E-02	1.51
		1	$h_{F_x}$	5.3481E-04	2.13	2.3935E-03	1.11	2.8166E-02	1.17
		$h_{F_x}^{-1}$	$h_{F_x}$	1.1571E-03	2.01	3.8674E-03	1.09	4.1202E-02	1.06
2	28311552	1	1	3.1949E-06	2.97	1.2851E-06	3.02	1.0190E-04	2.48
		$h_{F_x}^{-1}$	1	4.3013E-06	2.78	5.4307E-05	1.90	3.9421E-04	1.97
		1	$h_{F_x}$	1.3603E-06	3.02	7.4015E-06	2.26	1.7307E-04	2.24
		$h_{F_x}^{-1}$	$h_{F_x}$	2.5958E-06	3.01	7.4462E-05	1.98	4.3786E-04	2.00
3	62914560	1	1	1.3258E-08	4.00	6.9321E-09	4.00	5.0930E-07	3.49
		$h_{F_x}^{-1}$	1	1.5043E-08	3.81	2.2632E-07	2.95	1.4669E-06	3.05
		1	$h_{F_x}$	7.2375E-09	4.01	6.2051E-08	3.13	1.2966E-06	3.14
		$h_{F_x}^{-1}$	$h_{F_x}$	1.0360E-08	4.01	3.3719E-07	2.99	2.2607E-06	3.01

Table 1: Empirical convergence rates of the full-tensor space–time DG scheme for *Test 1*, as described in §8.1, on uniform meshes with  $h_x \approx h_t = 2^{-l}$ ,  $l \in \mathbb{N}$ , and fix  $p = p_v^x = p_v^t = p_\sigma^t = p_\sigma^x$ . The column labelled “DOFs” shows the total number of degrees of freedom in space–time, and the columns labelled “Error” show the errors in  $L^2(\Omega \times \{T\})$ -norm and  $|\cdot|_{\text{DG}}$  seminorm, for the mesh level  $l = 6$ . The columns labelled “Rate” show the estimated orders of convergence computed using the mesh levels  $l = 4, 5, 6$ .

$p$	DOFs	$\alpha$	$\beta$	$v_h$		$\sigma_h$		$ \cdot _{\text{DG}}$	
				Error	Rate	Error	Rate	Error	Rate
1	5242880	1	1	8.0082E-03	0.82	4.8002E-02	0.96	3.2155E-01	0.53
		$h_{F_x}^{-1}$	1	7.4382E-03	0.79	4.7774E-02	0.96	3.2314E-01	0.53
		1	$h_{F_x}$	3.9547E-04	1.82	1.8237E-02	1.00	4.6940E-02	1.06
		$h_{F_x}^{-1}$	$h_{F_x}$	9.9308E-04	1.98	1.7012E-02	1.02	4.9905E-02	1.04
2	18874368	1	1	2.8767E-04	2.02	1.7415E-04	2.00	3.6441E-03	1.53
		$h_{F_x}^{-1}$	1	2.9474E-05	2.92	1.9826E-04	1.97	9.8431E-04	2.02
		1	$h_{F_x}$	1.2984E-04	2.06	1.6404E-04	2.00	2.3004E-03	1.63
		$h_{F_x}^{-1}$	$h_{F_x}$	4.0016E-06	3.06	1.6247E-04	2.00	8.9843E-04	2.00
3	46137344	1	1	3.6718E-06	2.99	1.1350E-06	3.00	3.8272E-05	2.49
		$h_{F_x}^{-1}$	1	3.2268E-06	2.99	1.1832E-06	2.99	3.2379E-05	2.51
		1	$h_{F_x}$	2.1222E-07	3.67	1.4085E-06	2.93	7.4792E-06	2.97
		$h_{F_x}^{-1}$	$h_{F_x}$	5.6900E-08	4.00	1.1043E-06	3.00	5.9181E-06	2.99

Table 2: Empirical convergence rates of the full-tensor space–time DG scheme for *Test 1*, as described in §8.1, on uniform meshes with  $h_x \approx h_t = 2^{-l}$ ,  $l \in \mathbb{N}$ , and fix  $p = p_v^x = p_v^t = p_\sigma^t$  and  $p_\sigma^x = p - 1$ . The column labelled “DOFs” shows the total number of degrees of freedom in space–time, and the columns labelled “Error” show the errors in  $L^2(\Omega \times \{T\})$ -norm and  $|\cdot|_{\text{DG}}$  seminorm, for the mesh level  $l = 6$ . The columns labelled “Rate” show the estimated orders of convergence computed using the mesh levels  $l = 4, 5, 6$ .

mesh level $l$	parameter $J$			
	$p_\sigma^x = 0$	$p_\sigma^x = 1$	$p_\sigma^x = 2$	$p_\sigma^x = 3$
1	1	2	4	5
2	2	5	8	11
3	4	8	13	17
4	5	11	17	23
5	7	14	22	29
6	8	17	26	35

Table 3: The parameters used in Algorithm 1 for generating meshes with local refinement for the  $\Gamma$ -shaped spatial domain  $\Omega := (-\frac{1}{2}, \frac{1}{2})^2 \setminus \{(0, \frac{1}{2}) \times (-\frac{1}{2}, 0)\}$  of *Test 2*. Local refinement is performed only at the corner  $\mathbf{x}_0 = (0, 0)^\top$ . The cut-off radius  $R_c = 0.245$  and refinement weight  $\delta = 1/3$  (see Remark 3.2). The number of local refinements, given by parameter  $J$ , are presented for mesh resolutions  $h_x = 2^{-l}$ ,  $l \in \{1, 2, \dots, 6\}$ , and polynomial degrees  $p_\sigma^x \in \{0, 1, 2, 3\}$ .

## 8.2 Test 2: singular solution, $\Gamma$ -shaped domain

Choose the  $\Gamma$ -shaped spatial domain  $\Omega := (-\frac{1}{2}, \frac{1}{2})^2 \setminus \{(0, \frac{1}{2}) \times (-\frac{1}{2}, 0)\}$  and the end time  $T = 1$ . The polar coordinates centered at the re-entrant corner are

$$r = \|\mathbf{x}\|_2, \quad \theta = \arctan(x_2/x_1), \quad \forall \mathbf{x} \in \Omega.$$

Consider the singular solution

$$u(\mathbf{x}, t) = r^\gamma \sin(\gamma\theta) \sin(\sqrt{2}\pi t), \quad \text{for } \gamma = \frac{2}{3}, \quad \forall (\mathbf{x}, t) \in Q, \quad (53a)$$

$$v = \partial_t u, \quad \boldsymbol{\sigma} = -\nabla_{\mathbf{x}} u, \quad (53b)$$

to the IBVP (1), with Neumann boundary conditions and with the prescribed source term

$$f(\mathbf{x}, t) = -2\pi^2 r^\gamma \sin(\gamma\theta) \sin(\sqrt{2}\pi t), \quad \forall (\mathbf{x}, t) \in Q.$$

The parameters used in Algorithm 1 for generating locally refined meshes in this experiment are presented in Table 3 and some examples of these meshes are shown in Figure 1. The convergence rates are given in Tables 4 (quasi-uniform meshes), 5 and 6 (locally refined meshes,  $p = p_v^t = p_\sigma^t = p_v^x = p_\sigma^x$  and  $p = p_v^t = p_\sigma^t = p_v^x = p_\sigma^x + 1$ , respectively), for various choices of the stabilization parameters.

The tables show that for the singular solution (53), the corner-refined meshes, as described in §7, allow the space-time DG-scheme to preserve the same orders of convergence as for the smooth-solution case of §8.1.

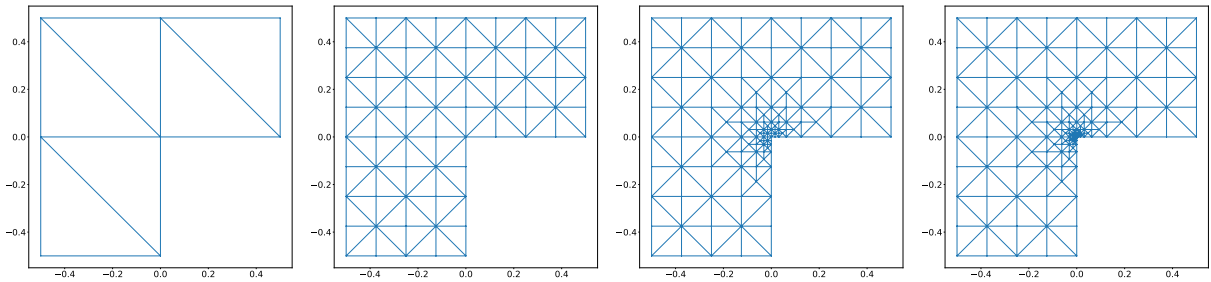


Figure 1:  $\Gamma$ -shaped spatial domain  $\Omega := (-\frac{1}{2}, \frac{1}{2})^2 \setminus \{(0, \frac{1}{2}) \times (-\frac{1}{2}, 0)\}$ . [Column 1] Initial mesh  $\mathcal{T}_0^{\mathbf{x}}$ . [Column 2] Uniform mesh for  $h_{\mathbf{x}} = 2^{-2}$ . Meshes generated from the initial mesh  $\mathcal{T}_0^{\mathbf{x}}$  using Algorithm 1 with parameters given in Table 3, only for the corner  $\mathbf{x}_0 = (0, 0)^{\top}$ : [column 3]  $p_x^{\sigma} = 1$ ,  $h_{\mathbf{x}} = 2^{-2}$ ; [column 4]  $p_x^{\sigma} = 2$ ,  $h_{\mathbf{x}} = 2^{-2}$ .

## 8.3 Test 3: smooth solution, heterogeneous medium

In this test case, we simulate the reflection and transmission of a wave at the interface of two different media. We follow the experimental setup provided in [40, §6.7] and use its results as reference for comparison.

Consider the space-time domain  $Q = (0, 2)^2 \times (0, 1)$ . The wavespeed is the piecewise constant function

$$c(\mathbf{x}) = \begin{cases} c_1 = 1, & x_1 \leq 1.2, \\ c_2 = 3, & x_1 > 1.2. \end{cases}$$

As the initial condition, we take a Gaussian wave given by

$$u_0 = \exp(-\|\mathbf{x} - \mathbf{x}_0\|^2 / \delta^2), \quad v_0 = 0, \quad \boldsymbol{\sigma}_0 = -\nabla_{\mathbf{x}} u,$$

where  $\mathbf{x}_0 = (1, 1)^{\top}$  and  $\delta = 0.01$ . We consider homogeneous Dirichlet boundary conditions. The computations are performed with polynomial degrees  $p_x^v = p_x^{\sigma} + 1 = 4$  and  $p_t^v = p_t^{\sigma} = 1$ . The quasi-uniform spatial mesh used in this experiment is shown in Figure 2 (center panel). The uniform temporal mesh has step size  $h_t = 2^{-6}$ . We use approximately  $7.35 \times 10^7$  degrees of freedom in space-time; the size of each linear system solved is approximately  $1.15 \times 10^6$ .

Snapshots of the solution are shown in Figure 3. First, the initial condition evolves in the left homogeneous medium. At time  $t = 0.2$ , the wave crosses over into the medium with higher wave speed. The

$p$	$v_h$		$\sigma_h$		$ \cdot _{\text{DG}}$	
	Error	Rate	Error	Rate	Error	Rate
1	6.5455E-02	0.66	5.2710E-02	0.60	3.9363E-02	0.69
2	4.8317E-02	0.67	3.9329E-02	0.60	2.8139E-02	0.63
3	6.8525E-02	0.66	5.3033E-02	0.59	3.6141E-02	0.61

Table 4: Empirical convergence rates of the full-tensor space–time DG scheme for *Test 2*, described in §8.2, on **uniform** meshes with  $h_{\mathbf{x}} \approx h_t = 2^{-l}$ ,  $l \in \mathbb{N}$ , and with fixed stabilization parameters  $\alpha = 1$  and  $\beta = 1$ , and polynomial degree  $p = p_v^x = p_v^t = p_\sigma^t = p_\sigma^x$ . The column labelled “Error” shows the estimated numerical error in  $L^2(\Omega \times \{T\})$ -norm and  $|\cdot|_{\text{DG}}$  seminorm: if  $p \in \{1, 2\}$ , it is given for the mesh level  $l = 6$ , else if  $p = 3$ , for  $l = 5$ . The column labelled “Rate” shows the estimated convergence rate: if  $p \in \{1, 2\}$ , it is computed with the mesh levels  $l = 4, 5, 6$ , else if  $p = 3$ , with  $l = 3, 4, 5$ .

$p$	DOFs	$\alpha$	$\beta$	$v_h$		$\sigma_h$		$ \cdot _{\text{DG}}$	
				Error	Rate	Error	Rate	Error	Rate
1	9780480	1	1	9.9964E-04	1.95	1.1227E-03	1.88	1.5032E-03	1.91
		$h_{F_{\mathbf{x}}}^{-1}$	1	1.0349E-03	1.94	4.0737E-03	1.11	4.2031E-03	1.27
		1	$h_{F_{\mathbf{x}}}$	9.9909E-04	1.95	1.3705E-03	1.75	1.6960E-03	1.83
		$h_{F_{\mathbf{x}}}^{-1}$	$h_{F_{\mathbf{x}}}$	1.0356E-03	1.94	4.2489E-03	1.13	4.3733E-03	1.27
2	35541504	1	1	1.8713E-05	3.00	2.4310E-05	2.85	3.0678E-05	2.91
		$h_{F_{\mathbf{x}}}^{-1}$	1	1.8715E-05	3.00	2.7236E-05	2.78	3.3047E-05	2.86
		1	$h_{F_{\mathbf{x}}}$	1.8714E-05	3.00	2.4686E-05	2.84	3.0978E-05	2.90
		$h_{F_{\mathbf{x}}}^{-1}$	$h_{F_{\mathbf{x}}}$	1.8718E-05	3.00	3.5567E-05	2.59	4.0191E-05	2.73
3	12441600	1	1	1.0517E-05	4.00	1.3890E-05	3.79	1.7422E-05	3.88
		$h_{F_{\mathbf{x}}}^{-1}$	1	1.0518E-05	4.00	1.4359E-05	3.77	1.7799E-05	3.87
		1	$h_{F_{\mathbf{x}}}$	1.0518E-05	4.00	1.3904E-05	3.79	1.7434E-05	3.88
		$h_{F_{\mathbf{x}}}^{-1}$	$h_{F_{\mathbf{x}}}$	1.0518E-05	4.00	1.4555E-05	3.76	1.7957E-05	3.86

Table 5: Empirical convergence rates of the full-tensor space–time DG scheme for *Test 2*, as described in §8.2, on **locally refined** meshes with  $h_{\mathbf{x}} \approx h_t = 2^{-l}$ ,  $l \in \mathbb{N}$ , and with fixed polynomial degree  $p = p_v^x = p_v^t = p_\sigma^t = p_\sigma^x$ . The column labelled “DOFs” shows the total number of degrees of freedom in space–time, and the column labelled “Error” shows the estimated numerical error in  $L^2(\Omega \times \{T\})$ -norm and  $|\cdot|_{\text{DG}}$  seminorm: if  $p \in \{1, 2\}$ , they are given for the mesh level  $l = 6$ , else if  $p = 3$ , for  $l = 5$ . The column labelled “Rate” shows the estimated convergence rate: if  $p \in \{1, 2\}$ , it is computed with the mesh levels  $l = 4, 5, 6$ , else if  $p = 3$ , with  $l = 3, 4, 5$ .

$p$	DOFs	$\alpha$	$\beta$	$v_h$		$\sigma_h$		$ \cdot _{\text{DG}}$	
				Error	Rate	Error	Rate	Error	Rate
1	4147200	1	1	2.4743E-02	0.92	2.6202E-02	0.83	3.6038E-02	0.87
		$h_{F_{\mathbf{x}}}^{-1}$	1	2.4702E-02	0.91	2.6246E-02	0.83	3.6042E-02	0.87
		1	$h_{F_{\mathbf{x}}}$	2.5666E-02	0.96	2.6355E-02	0.85	3.6788E-02	0.91
		$h_{F_{\mathbf{x}}}^{-1}$	$h_{F_{\mathbf{x}}}$	2.5740E-02	0.96	2.4506E-02	0.88	3.5540E-02	0.92
2	19560960	1	1	7.6086E-04	2.00	8.4941E-04	1.85	1.1404E-03	1.93
		$h_{F_{\mathbf{x}}}^{-1}$	1	7.5923E-04	2.00	8.4984E-04	1.85	1.1396E-03	1.93
		1	$h_{F_{\mathbf{x}}}$	7.5511E-04	2.00	8.4990E-04	1.85	1.1369E-03	1.93
		$h_{F_{\mathbf{x}}}^{-1}$	$h_{F_{\mathbf{x}}}$	7.5479E-04	2.00	8.3428E-04	1.86	1.1250E-03	1.93
3	57919488	1	1	1.6852E-05	3.00	2.1861E-05	2.84	2.7602E-05	2.91
		$h_{F_{\mathbf{x}}}^{-1}$	1	1.6821E-05	3.00	2.1854E-05	2.84	2.7579E-05	2.91
		1	$h_{F_{\mathbf{x}}}$	1.6732E-05	3.00	2.1896E-05	2.84	2.7557E-05	2.91
		$h_{F_{\mathbf{x}}}^{-1}$	$h_{F_{\mathbf{x}}}$	1.6726E-05	3.00	2.1606E-05	2.85	2.7324E-05	2.91

Table 6: Convergence rates of the full-tensor space–time DG scheme for *Test 2*, as described in §8.2, on **locally refined** meshes with  $h_{\mathbf{x}} \approx h_t = 2^{-l}$ ,  $l \in \mathbb{N}$ , and fix  $p = p_v^x = p_v^t = p_\sigma^t$  and  $p_\sigma^x = p - 1$ . The column labelled “DOFs” shows the total number of degrees of freedom in space–time, and the column labelled “Error” shows the estimated numerical error in  $L^2(\Omega \times \{T\})$ -norm and  $|\cdot|_{\text{DG}}$  seminorm, for the mesh level  $l = 6$ . The column labelled “Rate” shows the estimated order of convergence computed using the mesh levels  $l = 4, 5, 6$ .

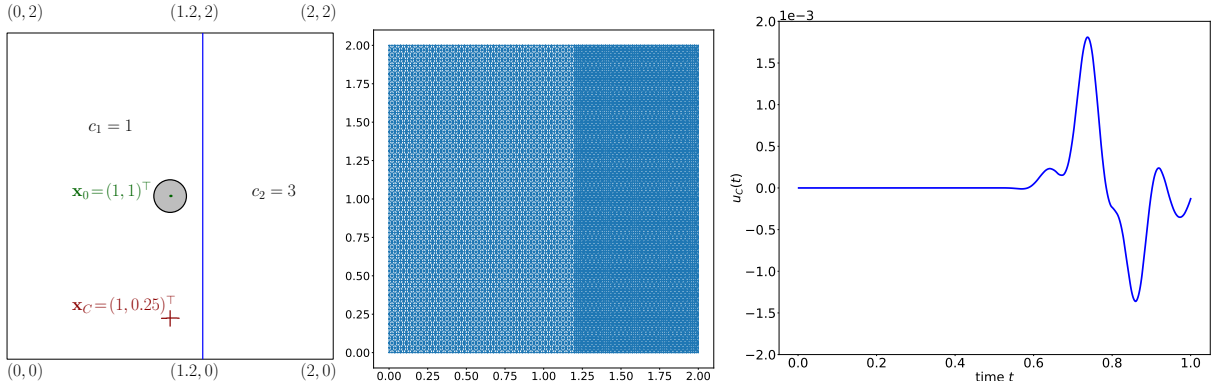


Figure 2: [left] Set-up of *Test 3*, §8.3: the grey region indicates the initial Gaussian wave with centre  $\mathbf{x}_0 = (1, 1)^\top$ , the material interface is at  $x_1 = 1.2$ , indicated by the blue line, with wave speed  $c_1$  to the left and  $c_2$  to the right of the interface, and the point  $\mathbf{x}_C = (1, 0.25)^\top$  in whose vicinity we measure the wave signal. [center] The quasi-uniform spatial mesh with meshwidth  $h_{\mathbf{x}} \approx 0.0365$  conforming to the material interface. [right] Solution signal  $u_C$  measured on the cell  $\Omega_C$  containing the point  $\mathbf{x}_C = (1, 0.25)^\top$ . The observed signal is in agreement with the measurements available for the same experiment in [40].

snapshot at  $t = 0.3$  shows that a part of the incident wave is transmitted across the interface with a higher speed and a shallow wavefront, and another part is reflected back. Finally, at  $t = 0.4$ , we also observe the weaker Huygens wave phenomenon.

Let  $\Omega_C$  be the mesh element that contains the point  $\mathbf{x}_C = (1, 0.25)$ . We measure  $v_C(t) = \int_{\Omega_C} v(\mathbf{x}, t) d\mathbf{x}$  as a time series signal and compute an approximation of  $u_C(t) = \int_{\Omega_C} u(\mathbf{x}, t) d\mathbf{x}$  by integrating the signal  $v_C(t)$  in time using the trapezoidal rule. Figure 2 (right panel) suggests that we numerically separate the three incoming waves: the very weak Huygens wave arrives first, followed by the initial wave and the reflected wave.

These observations are in agreement with the results for the same experimental set-up in [40, Fig. 11]. The signals are proportional but their amplitudes differ because the mesh elements  $\Omega_C$  on which they are integrated are different. They differ also in that here we plot the integral of the signal over  $\Omega_C$ , as opposed to its  $L^1(\Omega_C)$  norm, in order to capture the sign change of the reflected wave.

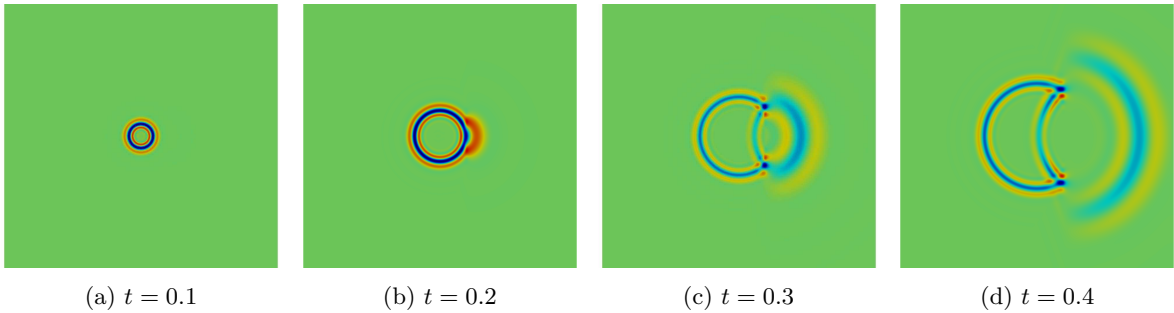


Figure 3: Wave propagation with the full-tensor DG scheme through a heterogeneous medium, the pressure  $v$  is shown at different times. The experimental set-up is described in *Test 3*, §8.3.

#### 8.4 Test 4: singular solution, heterogeneous medium

Consider the space–time domain  $Q = (0, 2)^2 \times (0, 0.3)$ . The wavespeed in  $\Omega = (0, 2)^2$  is the piecewise constant function

$$c(\mathbf{x}) = \begin{cases} c_1 = 3, & x_1 > 1.2, x_2 > 1, \\ c_2 = 1, & x_1 \leq 1.2, x_2 > 1, \\ c_3 = 3, & x_1 \leq 1.2, x_2 \leq 1, \\ c_4 = 1, & x_1 > 1.2, x_2 \leq 1. \end{cases}$$

As the initial condition, we take a Gaussian wave given by

$$u_0 = \exp(-\|\mathbf{x} - \mathbf{x}_0\|^2 / \lambda^2), \quad v_0 = 0, \quad \boldsymbol{\sigma}_0 = -\nabla_{\mathbf{x}} u,$$

where  $\mathbf{x}_0 = (1, 1.125)^\top$  and  $\lambda = 0.01$ . We consider homogeneous Dirichlet boundary conditions. The computations are performed with polynomial degrees  $p_x^v = p_x^\sigma = 2$  and  $p_t^v = p_t^\sigma = 1$ , and the uniform temporal mesh has step size  $h_t = 0.3 \cdot 2^{-4}$ .

In this experiment, we use the same quasi-uniform spatial mesh as in Test 3, Figure 2. Additionally, we use the locally refined spatial mesh shown in Figure 4. We have approximately 9.5 million and 16.5 million degrees of freedom in space–time for the quasi-uniform and locally refined meshes, respectively. The size of each linear system is approximately  $6 \times 10^5$  for the quasi-uniform case and  $10^6$  for the locally refined case. Snapshots of the solution at various times are shown in Figure 5.

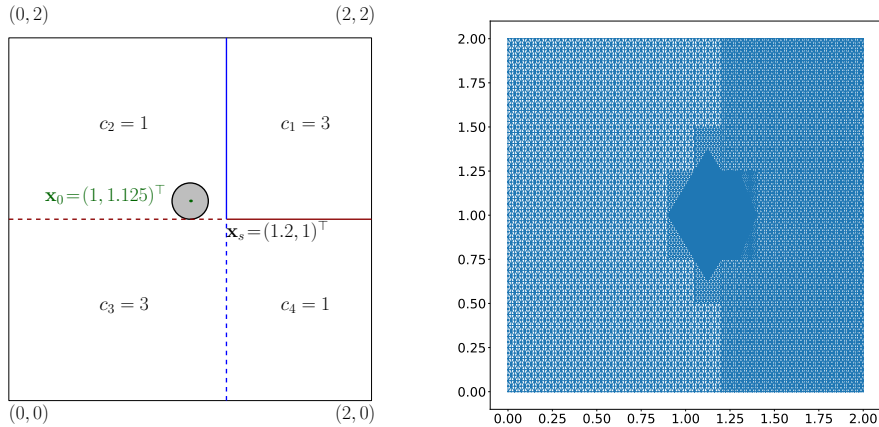


Figure 4: [left] Set-up of *Test 4*, §8.4: the grey region indicates the initial Gaussian wave with centre  $\mathbf{x}_0 = (1, 1.125)^\top$ , the material interfaces are indicated by the blue and brown lines. [right] The locally refined bisection-tree spatial mesh conforming to the material interfaces, generated using Algorithm 1 only for the material interface corner  $\mathbf{x}_s = (1.2, 1)^\top$ , with polynomial degree  $p_x^\sigma = 2$ , cut-off radius  $R_c = 0.392$ , refinement weight  $\delta = 0.4$  and mesh parameter  $h_x' = 0.0625$ , which yields the number of refinements  $J = 19$ . The resultant mesh has maximum meshwidth  $h_x \approx 0.0365$ .

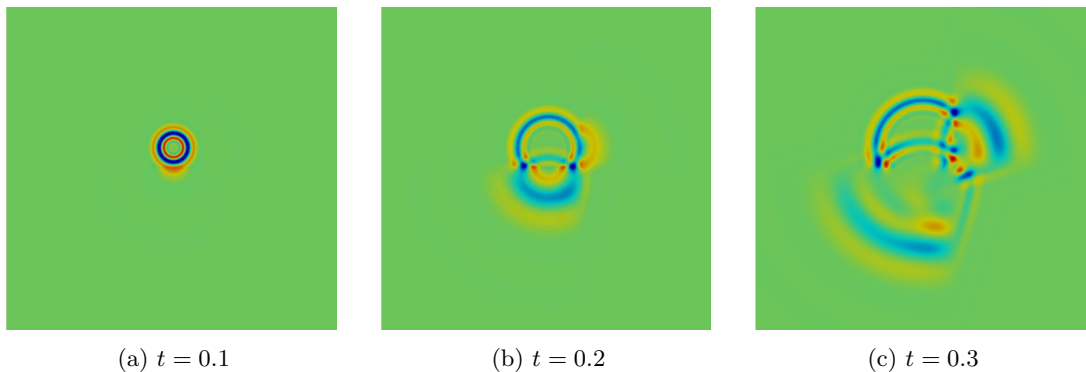


Figure 5: Wave propagation with the full-tensor DG scheme through a heterogeneous medium, the pressure  $v$  is shown at different times. The experimental set-up is described in *Test 4*, §8.4. The snapshots are computed with the locally refined mesh shown in Figure 4.

## 9 Sparse space–time discretization

Based on the work of Bungartz, Griebel et al., see [8] and the references therein, we develop now a *sparse-tensor space-time* (“sparse- $xt$ ” for short) approximation based on the above DG formulation and on *anisotropic* tensorized finite element spaces in the space–time cylinder  $Q$ . The sparse- $xt$  DG solution will be computed by (independent) numerical solution of several *anisotropic space-time* DG discretizations with different combinations of space and time steps, given by the so-called *combination formula*, equation (55). Subsequently, these anisotropic, full-tensor space–time numerical approximations will be combined into the sparse- $xt$  approximation. As we shall show, this approximation will afford the same asymptotic order of convergence as the full- $xt$  approximation, with the total number of degrees of freedom in the entire space–time domain proportional to that of one time step on the finest spatial grid.



## 9.1 Combination formula

To describe the sparse- $xt$  discretization, we shall require the following notation. We denote the discretization “levels” (corresponding, roughly, to generations in bisection-tree refinements in the spatial domain  $\Omega$  and the time-interval  $I$ ) by  $l_x$  and  $l_t \in \mathbb{N}_0$ , respectively. We tag the spatial and temporal discretization levels in the index pairs

$$\mathbf{l} := (l_x, l_t)^\top \in \mathbb{N}_0^2.$$

We introduce the total order  $|\mathbf{l}| = l_x + l_t$ . The indices  $l_x, l_t$  represent the resolution level in space and in time, respectively. We let  $L_x, L_t, L_{0,x}, L_{0,t} \in \mathbb{N}_0$  be such that

$$L_x - L_{0,x} = L_t - L_{0,t},$$

and define the vectors

$$\mathbf{L} = (L_x, L_t)^\top \in \mathbb{N}_0^2, \quad \mathbf{L}_0 = (L_{0,x}, L_{0,t})^\top \in \mathbb{N}_0^2,$$

and the set of admissible indices

$$\Xi_{(\mathbf{L}, \mathbf{L}_0)} := \{\mathbf{l} \in \mathbb{N}_0^2 : l_x \geq L_{0,x}, l_t \geq L_{0,t}, |\mathbf{l}| \in \{L_x + L_{0,t}, L_x + L_{0,t} - 1\}\}. \quad (54)$$

The indices  $L_x, L_t$  and  $L_{0,x}, L_{0,t}$  represent maximum and minimum resolution levels in space and in time, respectively.

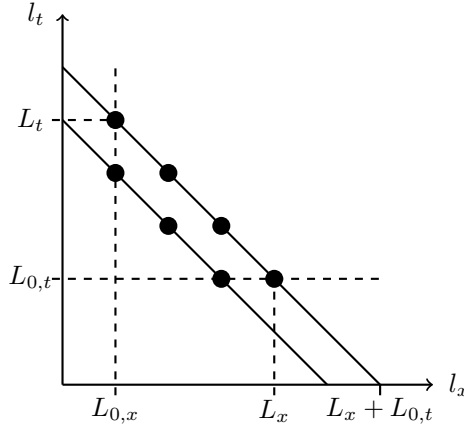


Figure 6: The indices in the set  $\Xi_{(\mathbf{L}, \mathbf{L}_0)}$  as in (54), with  $L_{0,x} = 1$ ,  $L_{0,t} = 2$ ,  $L_x = 4$ ,  $L_t = 5$ .

Let  $h_{l_x, x}, h_{l_t, t} \in \mathbb{R}^+$  and  $h_{0,x}, h_{0,t} \in \mathbb{R}^+$  be such that

$$h_{l_x, x} = 2^{-l_x} h_{0,x}, \quad h_{l_t, t} = 2^{-l_t} h_{0,t}.$$

A mesh  $\mathcal{T}_h^1$  has meshwidth bound  $h_{l_x, x}$  in the spatial variable and  $h_{l_t, t}$  with respect to the time variable. Assume that  $(\mathcal{T}_h^1)_{l_x=L_{0,x}}^{L_x}$  for a fixed  $h_{l_t, t}$ , and vice-versa  $(\mathcal{T}_h^1)_{l_t=L_{0,t}}^{L_t}$  for a fixed  $h_{l_x, x}$ , form a nested hierarchy of space–time meshes. For  $(v, \boldsymbol{\sigma}) \in L^2(I; H_\delta^{2,2}(\Omega)) \times L^2(I; H_\delta^{1,1}(\Omega)^2)$  and its discrete approximation  $(v_h^1, \boldsymbol{\sigma}_h^1) \in \mathbf{V}_P(\mathcal{T}_h^1)$ , define

$$\mathbf{w} := (v, \boldsymbol{\sigma}^\top)^\top, \quad \mathbf{w}_\mathbf{l} := (v_h^1, \boldsymbol{\sigma}_h^1)^\top.$$

Using the notation and definitions above, we can write a sparse- $xt$  approximation  $\hat{\mathbf{w}}_{(\mathbf{L}, \mathbf{L}_0)}$  of  $\mathbf{w}$  using the combination formula [8] in space and time dimensions as

$$\hat{\mathbf{w}}_{(\mathbf{L}, \mathbf{L}_0)} = \sum_{\mathbf{l} \in \Xi_{(\mathbf{L}, \mathbf{L}_0)}} c_\mathbf{l} \mathbf{w}_\mathbf{l} = (\hat{v}_{(\mathbf{L}, \mathbf{L}_0)}, \hat{\boldsymbol{\sigma}}_{(\mathbf{L}, \mathbf{L}_0)}^\top)^\top, \quad \text{where } c_\mathbf{l} = \begin{cases} +1, & \text{for } |\mathbf{l}| = L_x + L_{0,t} \\ -1, & \text{for } |\mathbf{l}| = L_x + L_{0,t} - 1, \end{cases} \quad (55)$$

are the so-called “combination coefficients”. We refer to  $\mathbf{w}_\mathbf{L}$  as *full- $xt$  DG numerical approximation*. Then, a *sparse- $xt$  DG numerical approximation* can be computed as follows:

- (a) Given  $\mathbf{L}, \mathbf{L}_0 \in \mathbb{N}_0^2$ , compute (in parallel) the space–time DG solution  $\mathbf{w}_\mathbf{l}$  for all  $\mathbf{l} \in \Xi_{(\mathbf{L}, \mathbf{L}_0)}$ .
- (b) Build the sparse- $xt$  numerical solution using the combination formula (55).

## 9.2 Error vs. complexity analysis

Given a maximal dyadic refinement level  $L \in \mathbb{N}$ , fix the indices  $L_{0,x} = L_{0,t} = 0$  and  $L_x = L_t = L$ . We omit the  $\mathbf{L}_0$  vector in the sparse- $xt$  approximation for ease of notation. Given  $h_{0,x} = h_{0,t} = 0.5$  and  $\mathbf{l} \in \mathbb{N}_0^2$ , consider a uniform time mesh with  $N_t = O((h_{l_t,t})^{-1})$  intervals of uniform (time) step size  $h_{l_t,t} = T/N_t$ . We shall further assume that  $h_{l_t,t} = O(2^{-l_t})$ . In the spatial domain  $\Omega$ , we consider a regular partition of  $\Omega$  into triangles of meshwidth  $h_{l_x,x} = O(2^{-l_x})$ . Then,  $h_{l_x,x} \simeq h_{l_t,t}$  holds for  $l_x = l_t = l$ , and in that case denote by  $h_l$  the common spatial and temporal meshwidth. Fix the polynomial degrees  $p_{x,K}^v = p_{x,K}^\sigma = p_x$  and  $p_{t,K}^v = p_{t,K}^\sigma = p_t$  for all  $K \in \mathcal{T}_h^1$ . Then,

$$M_h^1 := \dim(\mathbf{V}_p(\mathcal{T}_h^1)) = 6(p_t + 1) \frac{(p_x + 1)(p_x + 2)}{2} 2^{2l_x + l_t} = C(p_x, p_t) \cdot 2^{2l_x + l_t},$$

is the number of degrees of freedom of the finite-dimensional space  $\mathbf{V}_p(\mathcal{T}_h^1)$ . Let  $M_{\mathbf{L}}^f$  and  $M_{\mathbf{L}}^s$  denote the number of degrees of freedom for full- $xt$  and sparse- $xt$  approximations, respectively.

For a *full- $xt$  approximation*, with equal resolution levels  $L = l_x = l_t$  in space and time, we have, as  $L \rightarrow \infty$ ,

$$M_{\mathbf{L}}^f := M_h^{\mathbf{L}} = C(p_x, p_t) \cdot 2^{3L} = O(h_L^{-3}).$$

The overall number of degrees of freedom in a *sparse- $xt$  approximation* equals asymptotically, as  $L \rightarrow \infty$ ,

$$\begin{aligned} M_{\mathbf{L}}^s &= \sum_{\mathbf{l} \in \Xi_L} M_h^1 = C(p_x, p_t) \left( \sum_{l_x=0}^L 2^{2l_x + L - l_x} + \sum_{l_x=0}^{L-1} 2^{2l_x + L - l_x - 1} \right) \\ &= C(p_x, p_t) (2^L \cdot (2^{L+1} - 1) + 2^{L-1} \cdot (2^L - 1)) \\ &= C(p_x, p_t) \cdot \frac{(5 - 3 \cdot 2^{-L})}{2} \cdot 2^{2L} < \frac{5}{2} \cdot C(p_x, p_t) \cdot 2^{2L} = O(h_L^{-2}). \end{aligned}$$

We consider fixed, and equal spatial and temporal discretization orders  $p = \min(p_x, p_t) \geq 1$ . Then, under the solution regularity assumptions of Proposition 6.7, we have for the full-tensor  $xt$ -DG scheme, as  $L \rightarrow \infty$  (i.e.,  $h_L \rightarrow 0$ )

$$\begin{aligned} &\frac{1}{2} \|c^{-1}(v - v_h^{\mathbf{L}})\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\sigma - \sigma_h^{\mathbf{L}}\|_{L^2(\Omega \times \{t_n\})^2} \\ &\leq \mathcal{E}^{ft} := |(v, \sigma) - (v_h, \sigma_h)|_{\text{DG}(Q_n)} = O(h_L^{p+\frac{1}{2}}) = O([M_{\mathbf{L}}^f]^{-\frac{p+1/2}{3}}). \end{aligned}$$

On the other hand, if we consider the error bound for the sparse-tensor  $xt$ -DG scheme based on the combination formula (55), we expect the error to be bounded according to

$$\begin{aligned} &\frac{1}{2} \|c^{-1}(v - v_h^{\mathbf{L}})\|_{L^2(\Omega \times \{t_n\})} + \frac{1}{2} \|\sigma - \sigma_h^{\mathbf{L}}\|_{L^2(\Omega \times \{t_n\})^2} \\ &\leq \mathcal{E}^{st} := |(v, \sigma) - (v_h, \sigma_h)|_{\text{DG}(Q_n)} = O(h_L^{p+\frac{1}{2}}) = O([M_{\mathbf{L}}^s]^{-\frac{p+1/2}{2}}). \end{aligned}$$

These asymptotic relations are valid in both cases, for the quasi-uniform and the corner-refined meshes. This is demonstrated in the numerical experiments in §9.3.

The sparse-tensor discretization requires considerably *fewer degrees of freedom* than the standard, full-tensor scheme for comparable consistency order ( $M_{\mathbf{L}}^s = O(h_L^{-2})$  vs.  $M_{\mathbf{L}}^f = O(h_L^{-3})$ ), under realistic solution regularity (allowing, in particular, for conical singularities in the spatial variable).

A further advantage of the combination formula is provided by the *distribution of the sizes of the linear systems* to be solved, in parallel. We recall that the numbers of degrees of freedom  $M_{\mathbf{L}}^f$  and  $M_{\mathbf{L}}^s$  calculated above are the sums of the dimensions of all systems to be solved. The full-tensor  $xt$ -DG scheme requires the numerical solution of  $O(h_L^{-1}) = O(2^L)$  linear systems of size  $O(h_L^{-2}) = O(2^{2L})$  (i.e., one linear system solve per each time-step). On the other hand, the sparse  $xt$ -DG approximation requires the solution of  $\sum_{\mathbf{l} \in \Xi_L} O(2^{l_t}) = O(2^{L+1})$  many *linear systems of different sizes* given by  $O(2^{2l})$  for  $l = 0, \dots, L$ . That is, only  $O(1)$  systems of size  $O(2^{2L})$  ( $l_x = L, l_t = 0$ , i.e. finest mesh in space and coarsest in time), twice as many systems of a quarter size  $O(2^{2(L-1)})$  ( $l_x = L-1, l_t = 1$ ),  $\dots$ ,  $O(2^L)$  systems of size  $O(1)$  ( $l_x = 0, l_t = L$ , i.e. coarsest mesh in space and finest in time). Since  $O(2^{2L})$  is the size of the linear system for the DG discretization of a two-dimensional elliptic problem on the finest spatial mesh, the computational effort of the sparse  $xt$ -DG scheme amounts to that of this elliptic problem.

### 9.3 Numerical experiments

Our implementation to compute the sparse space–time solution is serial. However, it is important to note that the space–time combination formula (55) is *inherently parallel*. All the solutions  $\mathbf{w}_l$ , for  $l \in \Xi(\mathbf{L}, \mathbf{L}_0)$ , can be computed simultaneously using multiple processors. Computing a linear combination of these anisotropically discretized (so-called “detail”) solutions can be implemented as a reduction operation across processors.

We re-run the numerical experiments described in Test 1, §8.1, and Test 2, §8.2, for the sparse- $xt$  DG scheme. A comparison of the full- $xt$  and the sparse- $xt$  DG schemes, in terms of their accuracy versus the total number of degrees of freedom used, is shown in Figures 7 and 8.

In particular, for  $p = p_x^v = p_x^\sigma = p_t^v = p_t^\sigma \in \{1, 2\}$  and  $\alpha = \beta = 1$  we observe in the first rows of Figure 7 (smooth solution, quasi-uniform mesh) and Figure 8 (singular solution, locally refined mesh) that the errors decay as  $O([M_{\mathbf{L}}^f]^{-\frac{p+1}{3}})$  for the full-tensor product space, and as  $O([M_{\mathbf{L}}^s]^{-\frac{p+1}{2}})$  for the sparse tensor product space (both are slightly better than what is expected from the theory). Further experiments show that the errors for the two fields  $v$  and  $\sigma$  decay with the same rate (cf. the first two rows of Figures 7–10 in the first version of [4]). If  $p_x^\sigma$  is lowered to  $p-1$  and the flux parameters are scaled as  $\alpha^{-1} = \beta = h_{F_x}$ , then the errors for  $\sigma$  reduce to  $O([M_{\mathbf{L}}^f]^{-\frac{p}{2}})$  and  $O([M_{\mathbf{L}}^s]^{-\frac{p}{2}})$ , while those for  $v$  behave more erratically; error plots for this choice of the parameters are shown in the second rows of Figure 7 (smooth solution, quasi-uniform mesh) and Figure 8 (singular solution, locally refined mesh).

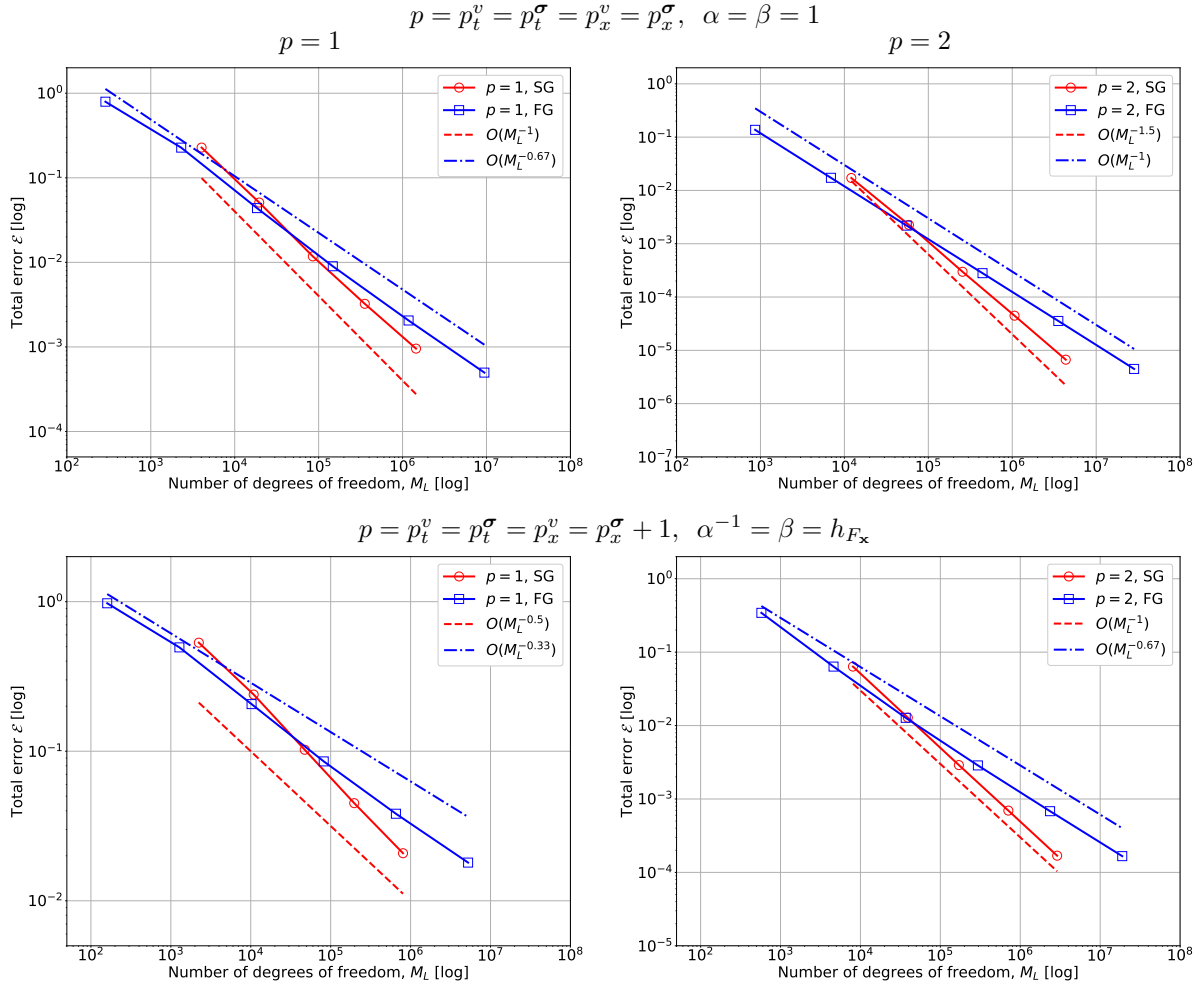


Figure 7: Numerical results for *Test 1*, as described in §8.1; FG: **full-tensor**  $xt$ -DG, SG: **sparse-tensor**  $xt$ -DG. The polynomial degrees and the stabilization parameters are chosen as  $p = p_t^v = p_t^\sigma = p_x^v = p_x^\sigma, \alpha = \beta = 1$  (first row) and  $p = p_t^v = p_t^\sigma = p_x^v = p_x^\sigma + 1, \alpha^{-1} = \beta = h_{F_x}$  (second row), with  $p = 1$  (left column) and  $p = 2$  (right column). The meshwidth for full-tensor  $xt$ -DG scheme is  $h_x \approx h_t = 2^{-l}$ , for  $l = 1, \dots, 6$ . The sparse-tensor  $xt$ -DG parameters are chosen as  $h_{0,x} \approx h_{0,t} = 0.5, L_{0,x} = 0, L_{0,t} = 1$  and  $L_x = L_t - 1 = L$ , for  $L = 1, \dots, 5$ .

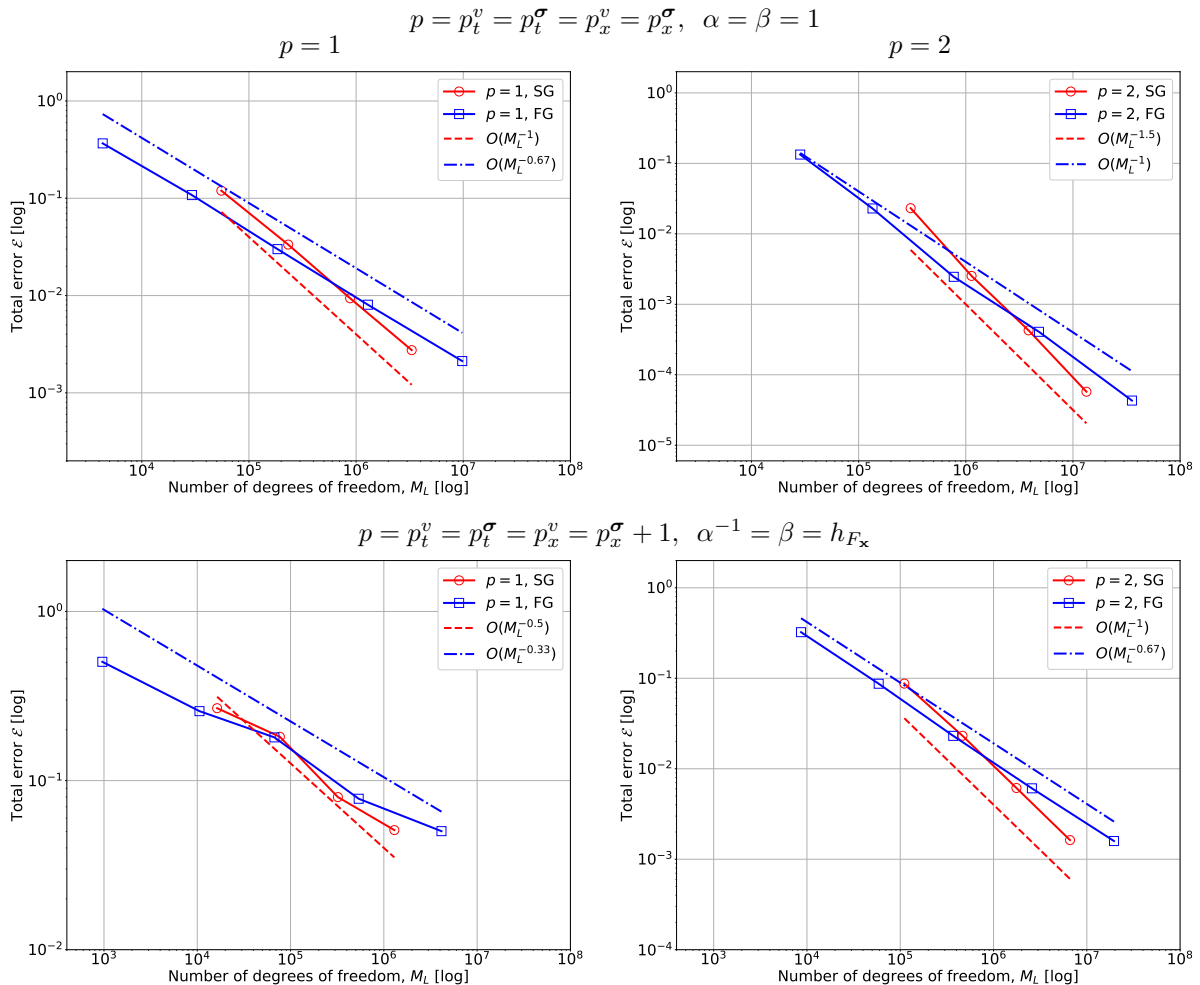


Figure 8: Numerical results for *Test 2*, as described in §8.2; FG: **full-tensor** *xt*-DG, SG: **sparse-tensor** *xt*-DG. The polynomial degrees and the stabilization parameters are chosen as  $p = p_t^v = p_t^\sigma = p_x^v = p_x^\sigma, \alpha = \beta = 1$  (first row) and  $p = p_t^v = p_t^\sigma = p_x^v = p_x^\sigma + 1, \alpha^{-1} = \beta = h_{F_x}$  (second row), with  $p = 1$  (left column) and  $p = 2$  (right column). The meshwidth for full-tensor *xt*-DG scheme is  $h_x \approx h_t = 2^{-l}$ , for  $l = 2, \dots, 6$ . The sparse-tensor *xt*-DG parameters are chosen as  $h_{0,x} \approx h_{0,t} = 0.25, L_{0,x} = 0, L_{0,t} = 1$  and  $L_x = L_t - 1 = L$ , for  $L = 1, \dots, 4$ .

## 10 Conclusions

We analyzed a space–time discontinuous Galerkin discretization based on the variational formulation from [33] for the linear, acoustic wave equation. We admitted, in particular, spatial mesh refinement in polygonal domains and for multi-material configurations where transmission conditions are imposed at material interfaces. The proposed DG discretization is unconditionally stable with respect to space and time-step sizes which arise in such space–time partitions. Using recent regularity results for the time-domain acoustic wave equation, we proved that spatial mesh grading towards corners and multi-material interface points can restore the maximal possible convergence rate. Numerical experiments for piecewise-constant coefficient in polygonal domains with manufactured exact solutions, which exhibit conical diffraction singularities, confirmed the convergence rate analysis. For piecewise-homogeneous materials occupying polygonal subdomains that are meshed exactly, the polynomial shape functions in the DG scheme give rise to cell contributions that can be evaluated exactly. For more general, (piecewise) smoothly varying coefficient functions, numerical integration has to be used in the setup of the proposed DG scheme. We expect that the presently developed stability bounds and the consistency analysis for the space–time DG scheme extend to variable coefficients subject to a quadrature consistency error analysis.

We also presented a novel, *sparse space–time DG discretization* admitting local mesh refinement towards point singularities in the spatial domain. It is based on the *combination* of several (possibly parallel) numerical solutions of the space–time DG scheme for a wide range of space and time-steps, including the mentioned spatial local mesh refinement towards conical singular points. The unconditional stability of the full tensor space–time DG scheme is essential here, as the combination technique will always access

CFL-violating combinations of space and time step sizes. As we showed in numerical experiments, this strategy furnishes the maximal asymptotic convergence order afforded by the discretization, even in the presence of conical singularities, and with overall work scaling as the solution of one elliptic spatial solve (i.e., one implicit time step) on the finest spatial grid.

In numerical experiments (§9.3) for model singular problems, the sparse space–time DG scheme (with polynomial degree  $p = 1, 2$ ) actually outperformed the corresponding full–tensor space–time DG scheme in terms of (relative) error versus total number of degrees of freedom, albeit only in the error regime below one percent. A consistency error analysis of the sparse space–time DG approximation is under investigation.

As mentioned, the unconditional stability of the considered space–time DG method is important in view of widely varying spatial step-sizes due to mesh refinement near point singularities in the spatial domain. This is also key in facilitating the sparse–tensor space–time DG algorithm: the terms which enter the combination formula involve arbitrary combinations of spatial and temporal step sizes, in particular, therefore, CFL-violating pairs. Furthermore, the global space–time nature of the discretization suggests that the presently proposed space–time DG method is very robust with respect to the control of *numerical dispersion*, which usually is significant in explicit time-marching schemes. A detailed account of this will be the subject of future investigations.

Recently, there has been considerable interest in time-explicit schemes for acoustic and seismic wave propagation. Strong corner mesh refinement in the spatial domain as considered here will, however, generally entail rather *anisotropic space–time slabs* which require prohibitively small explicit time-steps in elements situated close to corners. This could be addressed either by *implicit local time-stepping* or by multi-level, explicit local time-stepping, as recently advocated e.g. in [12]; an error and stability analysis for explicit, local time-stepping of the presently proposed, graded and bisection-tree meshes is yet to be developed.

The presently proposed space–time DG discretizations will extend also to other linear, second-order wave equations, in particular to elastic and electromagnetic waves, and to three-dimensional domains. These issues will be addressed in detail elsewhere.

## Acknowledgments

The Authors would like to acknowledge the kind hospitality of the Erwin Schrödinger International Institute for Mathematics and Physics (ESI), where part of this research was developed under the frame of the Thematic Programme *Numerical Analysis of Complex PDE Models in the Sciences*.

P. Bansal acknowledges support from the project *ModCompShock*, funded by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 642768. A. Moiola acknowledges support from GNCS-INDAM, from PRIN project “NA\_FROM-PDEs” and from MIUR through the “Dipartimenti di Eccellenza” Programme (2018-2022)–Dept. of Mathematics, University of Pavia. I. Perugia has been funded by the Austrian Science Fund (FWF) through the projects F 65 and P 29197-N32.

## References

- [1] M. S. ALNÆS, J. BLECHTA, J. HAKE, A. JOHANSSON, B. KEHLET, A. LOGG, C. RICHARDSON, J. RING, M. E. ROGNES, AND G. N. WELLS, *The FEniCS project version 1.5*, Archive of Numerical Software, 3 (2015).
- [2] I. BABUŠKA, R. B. KELLOGG, AND J. PITKÄRANTA, *Direct and inverse error estimates for finite elements with mesh refinements*, Numer. Math., 33 (1979), pp. 447–471.
- [3] L. BANJAI, E. H. GEORGOULIS, AND O. LIJOKA, *A Trefftz polynomial space-time discontinuous Galerkin method for the second order wave equation*, SIAM J. Numer. Anal., 55 (2017), pp. 63–86.
- [4] P. BANSAL, A. MOIOLA, I. PERUGIA, AND C. SCHWAB, *Space-time discontinuous Galerkin approximation of acoustic waves with point singularities*, arXiv:2002.11575 [math.NA] (2020).
- [5] L. BANJAI, C. LUBICH, AND J. M. MELENK, *Runge-Kutta convolution quadrature for operators arising in wave propagation*, Numer. Math., 119 (2011), pp. 1–20.
- [6] V. A. BOROVNIKOV AND B. Y. KINBER, *Geometrical theory of diffraction*, The Institution of Electrical Engineers, 1994.

- [7] S. C. BRENNER AND L. R. SCOTT, *Mathematical theory of finite element methods*, 3rd ed., Texts Appl. Math., Springer-Verlag, New York, 2007.
- [8] H.-J. BUNGARTZ AND M. GRIEBEL, *Sparse grids*, Acta Numer., 13 (2004), pp. 147–269.
- [9] A. CANGIANI, Z. DONG, AND E. H. GEORGIOULIS, *hp-Version space-time discontinuous Galerkin methods for parabolic problems on prismatic meshes*, SIAM Journal on Scientific Computing, 39 (2017), pp. A1251–A1279.
- [10] A. CHERNOV, *Optimal convergence estimates for the trace of the polynomial  $L^2$ -projection operator on a simplex*, Mathematics of Computation, 81 (2012), pp. 765–787.
- [11] R. DAUTRAY AND J.-L. LIONS, *Mathematical analysis and numerical methods for science and technology. Vol. 5*, Springer-Verlag, Berlin, 1992. Evolution problems. I, With the collaboration of Michel Artola, Michel Cessenat and Hélène Lanchon, Translated from the French by Alan Craig.
- [12] J. DIAZ AND M. J. GROTE, *Multi-level explicit local time-stepping methods for second-order wave equations*, Comput. Methods Appl. Mech. Engrg., 291 (2015), pp. 240–265.
- [13] S. DOHR, M. MERTA, G. OF, O. STEINBACH, AND J. ZAPLETAL, *A parallel solver for a pre-conditioned space-time boundary element method for the heat equation*, arXiv:1811.05224 [math.NA], (2018).
- [14] A. ERN AND J.-L. GUERMOND, *Theory and practice of finite elements*, vol. 159 of Applied Mathematical Sciences, Springer Science & Business Media, 2004.
- [15] J. ERNESTI AND C. WIENERS, *Space-time discontinuous Petrov-Galerkin methods for linear wave equations in heterogeneous media*, Comput. Methods Appl. Math., 19 (2019), pp. 465–481.
- [16] R. S. FALK AND G. R. RICHTER, *Explicit finite element methods for symmetric hyperbolic equations*, SIAM J. Numer. Anal., 36 (1999), pp. 935–952.
- [17] F. D. GASPOZ AND P. MORIN, *Convergence rates for adaptive finite elements*, IMA J. Numer. Anal., 29 (2009), pp. 917–936.
- [18] J. GOPALAKRISHNAN, M. HOCHSTEGER, J. SCHÖBERL, AND C. WINTERSTEIGER, *An explicit mapped tent pitching scheme for Maxwell equations*, 2019.
- [19] J. GOPALAKRISHNAN, J. SCHÖBERL, AND C. WINTERSTEIGER, *Mapped tent pitching schemes for hyperbolic systems*, SIAM Journal on Scientific Computing, 39 (2017), pp. B1043–B1063.
- [20] R. HIPTMAIR, A. MOIOLA, AND I. PERUGIA, *A Survey of Trefftz Methods for the Helmholtz Equation*, Lect. Notes Comput. Sci. Eng., Springer, 2016. pp. 237–278.
- [21] M. HOCHBRUCK, T. PAŽUR, A. SCHULZ, E. THAWINAN, AND C. WIENERS, *Efficient time integration for discontinuous Galerkin approximations of linear wave equations*, ZAMM Z. Angew. Math. Mech., 95 (2015), pp. 237–259.
- [22] P. HOUSTON, C. SCHWAB, AND E. SÜLI, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 39 (2002), pp. 2133–2163.
- [23] T. J. R. HUGHES AND G. M. HULBERT, *Space-time finite element methods for elastodynamics: formulations and error estimates*, Comput. Methods Appl. Mech. Engrg., 66 (1988), pp. 339–363.
- [24] G. M. HULBERT AND T. J. R. HUGHES, *Space-time finite element methods for second-order hyperbolic equations*, Comput. Methods Appl. Mech. Engrg., 84 (1990), pp. 327–348.
- [25] E. JONES, T. OLIPHANT, P. PETERSON, ET AL., *SciPy: Open source scientific tools for Python*, 2001–. [Online; accessed August 5, 2020].
- [26] J. KELLER, *Geometrical theory of diffraction*, Journal of the Optical Society of America, 52 (1962), pp. 116–130.
- [27] A. Y. KOKOTOV, P. NEITTAANMÄKI, AND B. A. PLAMENEVSKIĬ, *Problems of diffraction by a cone: asymptotic behavior of the solutions near the vertex*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI), 259 (1999), pp. 122–144, 297–298.

- [28] A. Y. KOKOTOV AND B. A. PLAMENEVSKIĬ, *A mixed problem for second-order hyperbolic systems in a wedge*, in *Nonlinear analysis and related problems (Russian)*, vol. 2 of Tr. Inst. Mat. (Minsk), Natl. Akad. Nauk Belarusi, Inst. Mat., Minsk, 1999, pp. 98–109.
- [29] A. Y. KOKOTOV AND B. A. PLAMENEVSKIĬ, *On the asymptotic behavior of solutions of the Neumann problem for hyperbolic systems in domains with conical points*, *Algebra i Analiz*, 16 (2004), pp. 56–98.
- [30] O. LAGHROUCHE, P. BETTES, AND R. J. ASTLEY, *Modelling of short wave diffraction problems using approximating systems of plane waves*, *Internat. J. Numer. Methods Engrg.*, 54 (2002), pp. 1501–1533.
- [31] V. T. LUONG AND N. T. TUNG, *The Dirichlet–Cauchy problem for nonlinear hyperbolic equations in a domain with edges*, *Nonlinear Analysis*, 125 (2015), pp. 457–467.
- [32] V. G. MAZ’YA, *Sobolev Spaces with Applications to Elliptic Partial Differential Equations*, Springer, 2nd ed., 2011.
- [33] A. MOIOLA AND I. PERUGIA, *A space–time Trefftz discontinuous Galerkin method for the acoustic wave equation in first-order formulation*, *Numerische Mathematik*, 138 (2018), pp. 389–435.
- [34] P. MONK AND G. R. RICHTER, *A discontinuous Galerkin method for linear symmetric hyperbolic systems in inhomogeneous media*, *J. Sci. Comput.*, 22/23 (2005), pp. 443–477.
- [35] F. MÜLLER, *Numerical analysis of finite element methods for second order wave equations in polygons*, PhD thesis, ETH Zurich, 2017. Available at <https://www.research-collection.ethz.ch/handle/20.500.11850/167502>.
- [36] F. MÜLLER, D. SCHÖTZAU, AND C. SCHWAB, *Discontinuous Galerkin methods for acoustic wave propagation in polygons*, *Journal of Scientific Computing*, (2018), pp. 1–27.
- [37] F. MÜLLER AND C. SCHWAB, *Finite elements with mesh refinement for wave equations in polygons*, *J. Comput. Appl. Math.*, 283 (2015), pp. 163–181.
- [38] F. MÜLLER AND C. SCHWAB, *Finite elements with mesh refinement for elastic wave propagation in polygons*, *Mathematical Methods in the Applied Sciences*, 39 (2016), pp. 5027–5042.
- [39] O. R. PEMBERY AND E. A. SPENCE, *The Helmholtz equation in random media: Well-posedness and a priori bounds*, *SIAM/ASA J. Uncertain. Quantif.*, 8 (2020), pp. 58–87.
- [40] I. PERUGIA, J. SCHÖBERL, P. STOCKER, AND C. WINTERSTEIGER, *Tent pitching and Trefftz-DG method for the acoustic wave equation*, *Computers and Mathematics with Applications*, (2020).
- [41] D. PETERSEIM AND M. SCHEDENSACK, *Relaxing the CFL condition for the wave equation on adaptive meshes*, *J. Sci. Comput.*, 72 (2017), pp. 1196–1213.
- [42] S. RHEBERGEN, B. COCKBURN, AND J. J. W. VAN DER VEGT, *A space-time discontinuous Galerkin method for the incompressible Navier-Stokes equations*, *J. Comput. Phys.*, 233 (2013), pp. 339–358.
- [43] O. STEINBACH AND M. ZANK, *A Stabilized Space–Time Finite Element Method for the Wave Equation*, Springer International Publishing, Cham, 2019, pp. 341–370.
- [44] S. V. D. WALT, S. C. COLBERT, AND G. VAROQUAUX, *The numpy array: A structure for efficient numerical computation*, *Computing in Science and Engg.*, 13 (2011), pp. 22–30.
- [45] T. P. WIHLER, *Discontinuous Galerkin FEM for Elliptic Problems in Polygonal Domains*, PhD thesis, Swiss Federal Institute of Technology Zurich, 2002. Available at <http://e-collection.library.ethz.ch/view/eth:26201>.