A Spacetime DPG Method for the Wave Equation in Multiple Dimensions

Jay Gopalakrishnan
*Portland State University*, gjay@pdx.edu

Paulina Sepulveda
*Basque Center for Applied Mathematics*

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/pics_pub](https://pdxscholar.library.pdx.edu/pics_pub)

Let us know how access to this document benefits you.

Citation Details
[https://pdxscholar.library.pdx.edu/pics_pub/14](https://pdxscholar.library.pdx.edu/pics_pub/14)

This Post-Print is brought to you for free and open access. It has been accepted for inclusion in Portland Institute for Computational Science Publications by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.
A SPACETIME DPG METHOD FOR THE WAVE EQUATION IN MULTIPLE DIMENSIONS

J. GOPALAKRISHNAN AND P. SEPÚLVEDA

ABSTRACT. A spacetime discontinuous Petrov-Galerkin (DPG) method for the linear wave equation is presented. This method is based on a weak formulation that uses a broken graph space. The wellposedness of this formulation is established using a previously presented abstract framework. One of the main tasks in the verification of the conditions of this framework is proving a density result. This is done in detail for a simple domain in arbitrary dimensions. The DPG method based on the weak formulation is then studied theoretically and numerically. Error estimates and numerical results are presented for triangular, rectangular, tetrahedral, and hexahedral meshes of the spacetime domain. The potential for using the built-in error estimator of the DPG method for an adaptivity mesh refinement strategy in two and three dimensions is also presented.

1. Introduction

This is a study on the feasibility of the discontinuous Petrov Galerkin (DPG) method [5, 6] for the spacetime wave equation. We follow the approach laid out in our earlier study of the DPG method for the spacetime Schrödinger equation [7].

Currently, the most widely used numerical techniques for transient problems are time-stepping schemes (based on the method of lines approach). However, there has been increasing interest recently in direct spacetime discretizations (where time is viewed as just another coordinate). Some reasons for investigating these approaches include their potential for performing natural spacetime adaptivity, possibility to obtain convergence even under limited spacetime regularity, exploitation of parallelism without causality constraints, and treatment of moving boundaries (see e.g. [7, 15, 16, 17, 18]). The analysis and implementation of 4D finite element discretizations is already underway [13, 18], hence our interest in obtaining a wellposed formulation in arbitrary dimensions.

Since the DPG method has a built-in error estimator and exhibits good pre-asymptotic mesh-independent stability properties, it is natural to consider its extension to spacetime problems. Applications of the DPG method for spacetime problems have already been computationally studied in [10] for the transient parabolic partial differential equations and [9] for the time-dependent convection-diffusion equation. We also note that a scheme that combines DPG spatial discretization with backward Euler time stepping for the heat equation has been analyzed in [11].

This work was partly supported by AFOSR grant FA9550-17-1-0090. Numerical studies were facilitated by the Portland Institute of Sciences (PICS) established under NSF grant DMS-1624776. Paulina Sepúlveda has received funding from the Spanish Ministry of Economy and Competitiveness with reference MTM2016-76329-R (AEI/FEDER, EU) and BCAM “Severo Ochoa” accreditation of excellence SEV-2013-0323 and SEV-2017-0718.
In contrast to these works, here we consider the transient acoustic wave system in arbitrary dimensions. One contribution of this work is a proof of the wellposedness of the ultraweak DPG formulation for the spacetime wave problem in a non-standard Hilbert space, without developing a trace theory for this function space. By using the abstract theory developed in [7], the proof reduces to verification of some conditions. This verification proceeds by proving a density result. The presented proof only applies for a multi-dimensional hyper-rectangle.

We also present, both practically and theoretically, how the built-in DPG error estimator is useful for spacetime adaptive refinement in two and three dimensions using conforming meshes of simplices. We also show that depending on how the interfacial variables are treated, one may end up with a discrete DPG system that has a nontrivial kernel for some alignment of mesh facets, a difficulty that we have not previously encountered in any other DPG example. We then provide practical solutions for solving for the DPG wave approximations despite the null space. The solutions computed using these techniques were observed to converge at the optimal rate.

In Section 2 we introduce the model wave problem and put it into the abstract variational setting of [7]. In Section 3 we introduce a broken weak formulation (upon which the DPG method is based) and prove its wellposedness subject to a density condition. In Section 4 we give a proof of the density condition for a simple multi-dimensional domain. In Section 5 we prove error estimates for the ideal DPG method for solutions with enough regularity. Finally, numerical experiments and implementation techniques are presented in Section 6.

2. The transient wave problem

Let \( \Omega_0 \) be a spatial domain in \( \mathbb{R}^d \), with Lipschitz boundary \( \partial \Omega_0 \), and let \( \Omega = \Omega_0 \times (0, T) \) be the spacetime domain, where \( T > 0 \) represents the final time. We consider the first order system for the wave equation given by

\[
\begin{align*}
\partial_t q - c \, \text{grad}_x \mu &= g, \\
\partial_t \mu - c \, \text{div}_x q &= f,
\end{align*}
\]

where \( f \in L^2(\Omega) \) and \( g \in L^2(\Omega)^d \) and \( c > 0 \) is the constant wave speed. Here the differential operators \( \text{div}_x \) and \( \text{grad}_x \) represent the (distributional) divergence and gradient operators that differentiate only along the spatial components \( (x) \). We add homogeneous initial and boundary conditions:

\[
\mu|_{t=0} = 0, \quad q|_{t=0} = 0, \quad \mu|_{\partial \Omega_0 \times (0,T)} = 0.
\]

Here, \( q \) represents the velocity and \( \mu \) the pressure. We now cast this problem in the framework of the abstract setting in [7, Appendix A].

2.1. The formal wave operator. Formally, the wave operator generated by the above system may be considered as a first order distributional derivative operator. Namely, set \( A : L^2(\Omega)^{d+1} \rightarrow \mathcal{D}'(\Omega)^{d+1} \) by

\[
Au = \begin{bmatrix} \partial_t u_q - c \, \text{grad}_x u_\mu \\ \partial_t u_\mu - c \, \text{div}_x u_q \end{bmatrix}
\]
where \( u \) in \( L^2(\Omega)^{d+1} \) is block partitioned into
\[
u = \begin{bmatrix} u_q \\ u_\mu \end{bmatrix}, \quad u_q \in L^2(\Omega)^d, \quad u_\mu \in L^2(\Omega).
\]

Next, we introduce the space
\[
W(\Omega) = \{ u \in L^2(\Omega)^{d+1} : Au \in L^2(\Omega)^{d+1} \}.
\]
By \( W(K) \) we mean the similarly defined space on an open subset \( K \) of \( \Omega \), but when considering this space with domain \( \Omega \), we abbreviate \( W(\Omega) = W \). Hereon, we denote by \( \langle \cdot, \cdot \rangle \) and \( \| \cdot \| \) the \( L^2(\Omega)^{d+1} \) inner product and norm, respectively, and \( D(\Omega)^{d+1} \) and \( D'(\Omega)^{d+1} \) is the space of infinitely differentiable vector functions with compact support in \( \Omega \) and its dual space, respectively. It is well known that the space \( W(\Omega) \) is a Hilbert space when endowed with the graph norm \( \| u \|_W = (\| u \|^2 + \| Au \|^2)^{1/2} \) (see [7, Lemma A.1]). The formal adjoint of \( A \) is the distributional differentiation operator \(-A\) and it satisfies
\[
\langle Aw, \tilde{w} \rangle = -\langle w, A\tilde{w} \rangle \quad \text{for all } w, \tilde{w} \in D(\Omega)^{d+1}.
\]
Define the operator \( D : W \to W' \) by
\[
\langle Du, v \rangle_W = \langle Au, v \rangle_\Omega + \langle u, Av \rangle_\Omega \quad \text{for all } u, v \in W.
\]
Here \( W' \) is the dual space of \( W \), and \( \langle \cdot, \cdot \rangle_W \) represents the duality pairing of a functional in \( W' \) with an element of \( W \). For smooth functions \( u, v \in D(\Omega)^{d+1} \), integration by parts shows that
\[
\langle Du, v \rangle_W = \int_{\partial \Omega} u_q \cdot (n_t v_q - c n_x v_\mu) + u_\mu (n_t v_\mu - c n_x v_q).
\]
Here and throughout, \( n = (n^x_t, n^\tau_t)^T \) represents the unit outward normal component to \( \Omega \) in \( \mathbb{R}^{d+1} \) and functions in \( L^2(\Omega)^{d+1} \), like the \( u \) and \( v \) above, are block partitioned as in (2.3).

2.2. The unbounded wave operator. In order to consider the boundary and initial conditions, we now proceed as suggested in [7, Appendix A], to define an unbounded operator with a domain that takes these conditions into account. Below, by abusing the notation, we shall denote this unbounded operator also by \( A \).

First, let us partition the spacetime boundary \( \partial \Omega \) into
\[
\Gamma_0 = \Omega_0 \times \{0\}, \quad \Gamma_T = \Omega_0 \times \{T\}, \quad \Gamma_b = \partial \Omega_0 \times [0, T].
\]
We define the following sets of smooth functions:
\[
\mathcal{V} = \{ u \in D(\Omega)^{d+1} : u|_{\Gamma_0} = 0, u_\mu|_{\Gamma_b} = 0 \},
\]
\[
\mathcal{V}^* = \{ v \in D(\Omega)^{d+1} : v|_{\Gamma_T} = 0, v_\mu|_{\Gamma_b} = 0 \}.
\]
Next, let \( A : \text{dom}(A) \subset L^2(\Omega)^{d+1} \to L^2(\Omega)^{d+1} \) be the unbounded operator in \( L^2(\Omega)^{d+1} \) defined by the right hand side of (2.2) with
\[
\text{dom}(A) = \{ u \in W : \langle Du, v \rangle_W = 0 \text{ for all } v \in \mathcal{V}^* \}.
\]
From (2.5), we see that the set of smooth functions \( D(\Omega)^{d+1} \) is contained in \( \text{dom}(A) \). Hence, \( A \) is a densely defined operator in \( L^2(\Omega)^{d+1} \). Therefore, it has a uniquely defined
adjoint $A^*$, which is again an unbounded operator. The adjoint $A^*$ equals the distributional derivative operator $-A$ when applied to $\text{dom}(A^*)$. This domain is prescribed as in standard functional analysis [2] by
\[
\text{dom}(A^*) = \{ \tilde{w} \in L^2(\Omega)^{d+1} : \exists \ell \in L^2(\Omega)^{d+1} \text{ such that } (Au, \tilde{w}) = (u, \ell) \text{ for all } u \in \text{dom}(A) \}.
\]

By definition, $\text{dom}(A)$ is a subset of $W(\Omega)$. When this subset is given the topology of $W(\Omega)$, we obtain a closed subset of $W(\Omega)$, which we call $V$, i.e., $V$ and $\text{dom}(A)$ coincide as sets or vector spaces, but not as topological spaces. Note that $\text{dom}(A^*)$ is also a subset of $V$, since for any $\tilde{w} \in \text{dom}(A^*)$, the distribution $-A\tilde{w}$ is in $L^2(\Omega)^{d+1}$. When $\text{dom}(A^*)$ is given the topology of $W$, it will be denoted by $V^*$. Observe that since $V$ is closed, $A$ is a closed operator. For any $S \subset W$ subspace, the right annihilator of $S$, denoted by $\perp S$, is defined by
\[
\perp S = \{ w \in W : \langle s', w \rangle_W = 0 \text{ for all } s' \in S \}.
\]
The definition of $\text{dom}(A^*)$, when written in terms of $D$ reveals that
\[
V^* = \perp D(V).
\]

Thus $V^*$ is also a closed subset of $W$.

The next observation is that from the definitions of $V$ and the operator $D$ (namely (2.8) and (2.4)) it immediately follows that $\mathcal{V} \subset V$. Note also that if $v^* \in \mathcal{V}^*$, then $\langle Au, v^* \rangle = -(u, Av^*) + \langle Du, v^* \rangle_W = -(u, Av^*)$ for all $u \in V$, since $\langle Du, v^* \rangle_W = 0$ by the definition of $V$. Therefore $v^*$ is in $V^*$. To summarize these observations, we have introduced $\mathcal{V}, \mathcal{V}^*, V$ and $V^*$, satisfying
\[
\mathcal{V} \subset V \quad \text{and} \quad \mathcal{V}^* \subset V^*.
\]

These are the abstract ingredients in the framework of [7, Appendix A] applied to the wave problem.

3. The Broken Weak Formulation

Following the settings of [3] and [7, Appendix], we partition the spacetime Lipschitz domain $\Omega$ into a mesh $\Omega_h$ of finitely many open elements $K$, (e.g. $(d + 1)$-simplices or $(d + 1)$-hyperrectangles) such that $\bar{\Omega} = \bigcup_{K \in \Omega_h} \bar{K}$ where $h = \max_{K \in \Omega_h} \text{diam}(K)$. The DPG method is based on a variational formulation in a “broken” analogue of $W$, which we call $W_h$, defined below.

We let $A_h$ be the wave operator applied element by element, i.e.,
\[
(A_h w)|_K = A(w|_K), \quad \forall w \in W(K), \quad K \in \Omega_h.
\]
Let $W_h = \{ w \in L^2(\Omega)^{d+1} : A_h w \in L^2(\Omega)^{d+1} \}$. The operator $D_h : W_h \rightarrow W_h'$ is defined by
\[
\langle D_h w, v \rangle_{W_h} = (A_h w, v)_\Omega + \langle w, A_h v \rangle_\Omega
\]
for all $w, v \in W_h$, where $\langle \cdot, \cdot \rangle_{W_h}$ denotes the duality pairing in $W_h$ in accordance with our previous notation. Below we abbreviate $\langle \cdot, \cdot \rangle_{W_h}$ to $\langle \cdot, \cdot \rangle_h$. Let $D_{h,V} : V \rightarrow W_h'$ denote
the restriction of $D_h$ to $V$, i.e., $D_{h,V} = D_h|_V$. The range of $D_{h,V}$, denoted by $Q$, is made into a complete space by the quotient norm

$$
\|\rho\|_Q = \inf_{v \in D_{h,V}^{-1}(\{\rho\})} \|v\|_W, \quad \rho \in Q \equiv \text{ran}(D_{h,V}).
$$

(3.1)

Define the bilinear form on $(L^2(\Omega)^{d+1} \times Q) \times W_h$ by

$$
b((v, \rho), w) = -\langle v, A_h w \rangle_{\Omega} + \langle \rho, w \rangle_h.
$$

The “broken” variational formulation for the wave problem now reads as follows. Given any $F$ in the dual space $W'_h$, find $u \in L^2(\Omega)^{d+1}$ and $\lambda \in Q$ such that

$$
b((u, \lambda), w) = F(w) \quad \text{for all } w \in W_h.
$$

(3.2)

Critical to the success of any numerical approximation of this formulation, in particular, the DPG approximation, is its wellposedness. By [7, Theorem A.5], this formulation is well-posed, provided we verify

$$
V = \bot D(V^*),
$$

(3.3)

$$
A : V \rightarrow L^2(\Omega)^{d+1} \text{ is a bijection.}
$$

(3.4)

Therefore our next focus is on proving (3.3) and (3.4). Recall from (2.11) that $V$ and $V^*$ are subspaces of smooth functions within $V$ and $V^*$. We now show that (3.3) and (3.4) follow if these are dense subspaces.

**Theorem 3.1.** Suppose

$$
\mathcal{V} \text{ is dense in } V \text{ and } \mathcal{V}^* \text{ is dense in } V^*.
$$

(3.5)

Then (3.3) and (3.4) holds. Consequently, the broken weak formulation (3.2) is well posed.

**Proof.** In view of the continuity of $D$, (2.11), and the assumption that $\mathcal{V}^*$ is dense in $V^*$, the condition (3.3) now immediately follows.

Next, we will prove that

$$
\|u\| \leq 2T\|Au\|, \quad \text{for all } u \in V,
$$

(3.6a)

$$
\|v\| \leq 2T\|A^*v\|, \quad \text{for all } v \in V^*.
$$

(3.6b)

These inequalities follow by well-known energy arguments, as shown in [8, Lemma 3]. We briefly include the proof for completeness. Let $v \in V^*$. Then

$$
\|v\|^2 = \int_0^T \left( \int_{\Omega_b} |v(x,t)|^2 \, dx \right) dt = 2 \int_0^T \int_0^T \int_{\Omega_b} \partial_s v(x, s) \cdot v(x, s) \, dx \, ds \, dt
$$

$$
= 2 \int_0^T \int_{\Omega_b} v(x, s) \cdot A^*v(x, s) \, dx \, ds \, dt - 2c \int_0^T \int_{\partial \Omega_b} (v_q \cdot n_x) v_\mu d\sigma \, ds \, dt
$$

$$
\leq 2T\|v\| \|A^*v\|.
$$

The inequality for $V$ is similarly proved by using its boundary conditions instead of those of $\mathcal{V}^*$.

Using the density assumptions, we conclude that (3.6) implies

$$
\|u\| \leq 2T\|Au\| \quad \text{for all } u \in V \quad \text{and}
$$

$$
\|v\| \leq 2T\|A^*v\| \quad \text{for all } v \in V^*.
$$

(3.7a)

(3.7b)
The inequality (3.7a) and the closed range theorem for closed operators imply that \( \text{dom}(A) = V \to L^2(\Omega)^{d+1} \) is injective and has closed range. Moreover, its adjoint \( A^* \) is injective (on its domain) by (3.7b), so the range of \( A \) must be all of \( L^2(\Omega)^{d+1} \) (see e.g., [2, Corollary 2.18]). Hence \( A \) is a bijection, i.e., condition (3.4) holds. Finally, since we have verified both (3.3) and (3.4), applying [7, Theorem A.5], the wellposedness follows.

Note that the wellposedness result of Theorem 3.1, in particular, implies that

\[
\beta = \inf_{0 \neq (v, \rho) \in L^2(\Omega)^{d+1} \times Q} \sup_{0 \neq w \in W_h} \frac{b(v, \rho, w)}{\|(v, \rho)\|_{L^2(\Omega)^{d+1} \times Q} \|w\|_{W_h}} > 0. \tag{3.8}
\]

4. Verification of the density condition

In this section, we verify (3.5) for a simple domain, namely a hyperrectangle (or an orthotope). Accordingly, throughout this section, we fix \( \Omega = \Omega_0 \times (0, T) \) and

\[
\Omega_0 = \prod_{i=1}^{d} (0, a_i),
\]

for some \( a_i > 0 \). While density of smooth functions in general graph spaces can be proved by standard Sobolev space techniques [1], to obtain the density of smooth functions with boundary conditions (like those in \( V \)) we need more arguments. The following proof has some similarities with the proof of [7, Theorem 3.1], an analogous density result for the one-dimensional Schrödinger operator. The main differences from [7] in the proof below include the consideration of multiple spatial dimensions and the construction of extension operators for vector functions in the wave graph space by combining even and odd reflections appropriately.

**Theorem 4.1.** On the above \( \Omega \), \( V^* \) is dense in \( V^* \) and \( V \) is dense in \( V \).

**Proof.** We shall only prove that \( V \) is dense in \( V \), since the proof of the density of \( V^* \) in \( V^* \) is similar. We divide the proof into three main steps.

**Step 1. Extension:** In this step, we will extend a function in \( V \) using spatial reflections to a domain which has larger spatial extent than \( \Omega \) (see Figure 1).

Let \( e_i \) denote the standard unit basis vectors in \( \mathbb{R}^{d+1} \) and \( y \in \mathbb{R}^{d+1} \) arbitrary. The following operations

\[
R_{i,-} y = y - 2y_i e_i, \quad R_{i,+} y = y + 2(a_i - y_i) e_i
\]

perform reflections of the coordinate vector \( y \) about \( y_i = 0 \) and \( y_i = a_i \) for \( i = 1, \ldots, d \). We set \( Q_0 \equiv \overline{\Omega} \) and then define extended domains \( Q_i \) in a recursive way, starting from \( i = 1 \) through \( i = d \) as follows.

\[
Q_{i,-} = R_{i,-}^{-1} Q_{i-1}, \quad Q_{i,+} = R_{i,+}^{-1} Q_{i-1}, \quad Q_i = Q_{i,-} \cup Q_{i-1} \cup Q_{i,+}.
\]

The final extended domain is \( Q \equiv Q_d \).

Next, we introduce even and odd extensions (in the \( x_i \)-direction) of scalar functions. Namely, let \( G_{i,e}, G_{i,o}: L^2(Q_{i-1}) \to L^2(Q_i) \) be defined by

\[
G_{i,e} f(x, t) = \begin{cases} f(R_{i,-}(x, t)) & \text{if } (x, t) \in Q_{i,-}, \\ f(R_{i,+}(x, t)) & \text{if } (x, t) \in Q_{i,+}, \\ f(x, t) & \text{if } (x, t) \in Q_{i-1}, \end{cases}
\]
and

\[ G_{i,o}f(x,t) = \begin{cases} 
-f(R_{i,-}(x,t)) & \text{if } (x,t) \in Q_{i,-}, \\
-f(R_{i,+}(x,t)) & \text{if } (x,t) \in Q_{i,+}, \\
f(x,t) & \text{if } (x,t) \in Q_{i,0}.
\end{cases} \]  

(4.2)

In the case of a vector function \( v \in L^2(Q_{i-1})^{d+1} \), we define \( G_i v(x,t) \) to be the extended vector function obtained by extending (in the \( i \)th direction) all the components of \( v \) using the odd scalar extension, except the \( i \)th component, which is extended using the even scalar extension. In other words, for any \( i = 1, \ldots, d \), we define \( G_i : L^2(Q_{i-1})^{d+1} \rightarrow L^2(Q_i)^{d+1} \) by

\[ G_i v = (G_{i,e}v_i)e_i + \sum_{j \neq i} (G_{i,o}v_j)e_j \]  

(4.3)

where the sum runs over all \( j = 1, \ldots, d+1 \) except \( i \). Let \( E_k = G_k \circ G_{k-1} \circ \ldots \circ G_1 \). The cumulative extension over all spatial directions is thus obtained using \( E = E_d \). It extends functions in \( \overline{\Omega} \) to \( Q \).

By change of variable formula for integration, we obtain

\[ (G_{i,o}f,g)_{Q_i} = (f,G_{i,o}g)_{Q_{i-1}}, \quad \text{for all } f \in L^2(Q_{i-1}), \ g \in L^2(Q_i), \]

\[ (G_{i,e}f,g)_{Q_i} = (f,G_{i,e}g)_{Q_{i-1}}, \quad \text{for all } f \in L^2(Q_{i-1}), \ g \in L^2(Q_i), \]

where the “folding” operators \( G'_{i,e/o} : L^2(Q_i) \rightarrow L^2(Q_{i-1}) \), that go the reverse direction of the extension operators, are defined by

\[ G'_{i,e}g(x,t) = g(x,t) - g(R_{i,-}^{-1}(x,t)) - g(R_{i,+}^{-1}(x,t)), \]

\[ G'_{i,o}g(x,t) = g(x,t) + g(R_{i,-}^{-1}(x,t)) + g(R_{i,+}^{-1}(x,t)). \]  

(4.4)

(4.5)

These scalar folding operators combine to form an analogue for vector functions as in (4.3), namely

\[ G_i'w = (G_{i,e}'w_i)e_i + \sum_{j \neq i} (G_{i,o}'w_j)e_j. \]

It satisfies \( (G_i v, w)_{Q_i} = (v, G_i'w)_{Q_{i-1}} \) for all \( v \in L^2(Q_{i-1})^{d+1}, \ w \in L^2(Q_i)^{d+1} \), and for each \( i \) from 1 to \( d \). Let \( E'_k = G'_k \circ G'_{k-1} \circ \ldots \circ G'_d \). Then \( E' = E'_1 \) folds functions in \( Q \) to \( \overline{\Omega} \) and is the adjoint of the extension \( E \) in the following sense.

\[ (Eu, w)_Q = (u, E'w)_{\overline{\Omega}}, \quad \text{for all } u \in L^2(\Omega)^{d+1}, \ w \in L^2(Q)^{d+1}. \]  

(4.6)

We want to prove that \( Ev \) is in \( W(Q) \) for any \( v \in V \). Note that if \( w \in L^2(\Omega)^{d+1} \), then \( Ew \) in \( L^2(Q)^{d+1} \), because each \( G_i \) maps \( L^2 \) functions into \( L^2 \) per (4.3). Therefore, in order to prove \( Ev \) is in \( W(Q) \), it only remains to prove that \( A(Ev) \) is in \( L^2(Q)^{d+1} \). Let \( \varphi \in D(Q)^{d+1} \) (where we abuse the notation and write \( D(Q) \) for \( D(Q^0) \) whenever \( Q^0 \) is the interior of \( Q \)). Using (4.6), the action of the distribution \( AEv \) on \( \varphi \) equals

\[ (AEv, \varphi)_{D(Q)^{d+1}} = -(Ev, A\varphi)_Q = -(v, E'A\varphi)_{\overline{\Omega}}. \]  

(4.7)

To analyze the last term, first observe that by the chain rule applied to a smooth scalar function \( \phi \) on \( Q_1 \),

\[ \partial_i(G_{i,o}' \phi) = G'_{i,o}(\partial_i \phi), \quad \partial_i(G_{i,e}' \phi) = G'_{i,e}(\partial_i \phi), \quad \partial_j(G_{i,o}' \phi) = G'_{i,o}(\partial_j \phi), \]

\[ \partial_j(G'_{i,e} \phi) = G'_{i,e}(\partial_j \phi), \quad \partial_i(G'_{i,e} \phi) = G'_{i,e}(\partial_i \phi), \quad \partial_j(G'_{i,o} \phi) = G'_{i,o}(\partial_j \phi). \]
for all \( j \neq i \). Combining these appropriately for smooth vector function \( \psi \) on \( Q_1 \), we find that

\[
\partial_t (G'_t \psi) = G'_t \partial_t \psi, \quad c \nabla_x G'_{i,o} \psi_\mu = G'_t (c \nabla_x \psi_\mu), \quad c \operatorname{div}_x G'_t \psi_q = G'_t (c \operatorname{div}_x \psi_q).
\]

Thus, for any \( \varphi \in \mathcal{D}(Q)^{d+1} \) we have \( E'_i A \varphi = A E'_i \varphi \) for all \( i = 1, \cdots, d \), and in particular

\[
E' A \varphi = A E' \varphi. \tag{4.8}
\]

Returning to (4.7) and using (4.8) and (2.4),

\[
\langle AE v, \varphi \rangle_{\mathcal{D}(Q)^{d+1}} = (Av, E' \varphi)_\Omega - \langle Dv, E' \varphi \rangle_{W(\Omega)}. \tag{4.9}
\]

We shall now show that the last term above vanishes. Since \( v \) is in \( V \), the last term will vanish by the definition of \( V \), provided \( E' \varphi \) is in \( V^* \). To prove that \( E' \varphi \) is in \( V^* \), we only need to verify that \( E' \varphi \) satisfies the boundary conditions in (2.7). Since \( \varphi \) is compactly supported in \( Q \), we obviously have \( (E' \varphi)|_{\Gamma_T} = 0 \) as \( E' \) only involves spatial folding.

Next, we claim that \( [E' \varphi]_{\mu}|_{\Gamma_b} = 0 \) also. To see this, let \( \Gamma^j \) denote the two facets of \( \partial Q_j \) where \( x_j \) is constant and \( \gamma^j \) denote the two facets of \( \partial Q_{j-1} \) where \( x_j \) is constant. The value of \( G'_{d,o} \varphi_\mu(x, t) \) for any \((x, t)\) in \( \gamma^{d-1} \) is the sum of the three terms in (4.4), two of which cancel each other, and one of which vanishes because \( \varphi_\mu|_{\Gamma^d} = 0 \). Thus \( \varphi_\mu|_{\Gamma^d} = 0 \implies (G'_{d,o} \varphi_\mu)|_{\partial Q_{d-1}} = 0 \) (where we have also used the fact that \( \varphi_\mu \) vanishes on the remainder \( \partial Q_{d-1} \setminus \gamma^{d-1} \)). The same argument can now be repeated to get that \( (G'_{d,o} \varphi_\mu)|_{\Gamma^d} = 0 \implies (G_{d-1,o} (G'_{d,o} \varphi_\mu))|_{\partial Q_{d-2}} = 0 \). Continuing similarly, we obtain that \( [E' \varphi]_\mu = G'_{1,o} \circ G'_{2,o} \circ \cdots \circ G'_{d,o} \varphi_\mu \) vanishes on \( \partial Q = \Gamma_b \). Thus, the last term in (4.9) is zero and by (4.6) we conclude that

\[
\langle AE v, \varphi \rangle_{\mathcal{D}(Q)^{d+1}} = (AE v, \varphi)_Q \tag{4.10}
\]

for all \( \varphi \) in \( \mathcal{D}(Q)^{d+1} \).

By virtue of (4.10), we have proved that for any \( v \in V \), \( AE v \) is in \( L^2(Q)^{d+1} \), \( AE v \) coincides with \( EA v \), and \( E v \) is in \( W(\Omega) \).

**Step 2. Translation:** In this step, we will translate up the previously obtained extension in time coordinate. This will give us room to mollify in the next step. Such a translation step is standard in many density proofs (see e.g., [1]).

---

**Figure 1.** Left: Extended domains \( Q_1 \) and \( Q_2 \) when \( \Omega \subseteq \mathbb{R}^3 \). Right: Translation by \( \delta \) in the \( t \) direction.
Let \( v \in V \) and let \( \tilde{E}v \) denote the extension of \( Ev \) to \( \mathbb{R}^{d+1} \), i.e., \( \tilde{E}v \) equals \( Ev \) in \( Q \) and it is zero elsewhere. Denote by \( \tau_\delta \) the translation operator in the \( t \)-direction by \( \delta > 0 \); i.e., \( (\tau_\delta w)(x,t) = w(x,t-\delta) \) for scalar or vector functions \( w \). It is well known \cite{2} that

\[
\lim_{\delta \to 0} \| \tau_\delta g - g \|_{L^2(\mathbb{R}^{d+1})} = 0, \quad \forall g \in L^2(\mathbb{R}^{d+1}). \tag{4.11}
\]

Let \( Q_\delta = \prod_{i=1}^{d+1} (-a_i, 2a_i) \times (-\delta, T+\delta) \) and let \( H_\delta \) be the restriction from \( \mathbb{R}^{d+1} \) to \( Q_\delta \).

We will now show that

\[
AH_\delta \tau_\delta \tilde{E}v = H_\delta \tau_\delta \tilde{E}Av. \tag{4.12}
\]

By a change of variable,

\[
(\tau_\delta \tilde{E}w, \tilde{w})_{Q_\delta} = (Ew, \tau_{-\delta} \tilde{w})_Q \tag{4.13}
\]

for all \( w \in L^2(\Omega)^{d+1} \) and \( \tilde{w} \in L^2(Q_\delta)^{d+1} \). Note that \( \tau_\delta \tilde{E}v \big|_\Omega \in L^2(\Omega)^{d+1} \). The distribution \( AH_\delta \tau_\delta \tilde{E}v \) applied to a smooth function \( \varphi \in D(Q_\delta)^{d+1} \) equals

\[
\langle AH_\delta \tau_\delta \tilde{E}v, \varphi \rangle_{D(Q_\delta)^{d+1}} = -\langle \tau_\delta \tilde{E}v, A\varphi \rangle_{Q_\delta} = -(Ev, A\tau_{-\delta} \varphi)_Q
\]

due to (4.13) and the fact that \( \tau_{-\delta} A \varphi = A \tau_{-\delta} \varphi \). Using also (4.6) and (4.8),

\[
\langle AH_\delta \tau_\delta \tilde{E}v, \varphi \rangle_{D(Q_\delta)^{d+1}} = -(v, E' A \tau_{-\delta} \varphi)_\Omega = -(v, A E' \tau_{-\delta} \varphi)_\Omega = (Av, E' \tau_{-\delta} \varphi)_\Omega - (Dv, E' \tau_{-\delta} \varphi)_\Omega.
\]

Note that \( E' \tau_{-\delta} \varphi \) satisfies all the boundary conditions required for it to be in \( \mathcal{V}^* \). Hence the last term in the above display is zero. We therefore conclude that

\[
\langle AH_\delta \tau_\delta \tilde{E}v, \varphi \rangle_{D(Q_\delta)^{d+1}} = (\tau_\delta EAv, \varphi)_{Q_\delta},
\]

which proves (4.12). In particular, \( H_\delta \tau_\delta \tilde{E}v \in W(Q_\delta) \) whenever \( v \in V \).

**Step 3. Mollification:** In this step we finish the proof by considering a \( v \in V \) and mollifying the time-translated extension \( \tau_\delta \tilde{E}v \) constructed above.

To recall the standard symmetric mollifier, let \( \rho_\varepsilon \in D(\mathbb{R}^{d+1}) \), for each \( \varepsilon > 0 \) be defined by

\[
\rho_\varepsilon(x,t) = \varepsilon^{-(d+1)} \rho_1(\varepsilon^{-1}x,\varepsilon^{-1}t),
\]

where

\[
\rho_1(x,t) = \begin{cases} 
  k \exp\left( -\frac{1}{2} \frac{1}{1 - |x|^2 + t^2} \right) & \text{if } |x|^2 + t^2 < 1, \\
  0 & \text{if } |x|^2 + t^2 \geq 1,
\end{cases}
\]

and \( k \) is a constant chosen so that \( \int_{\mathbb{R}^{d+1}} \rho_1 = 1 \). Here \( | \cdot | \) is the euclidean norm in \( \mathbb{R}^d \). Let \( \rho_\varepsilon*v \) denote the function obtained by component-wise convolution, i.e., \( [\rho_\varepsilon*v]_j = [v]_j * \rho_\varepsilon \) for all \( j \)-components. Then \( \rho_\varepsilon*v \) is a infinitely smooth vector function that satisfies

\[
\lim_{\varepsilon \to 0} \| v - \rho_\varepsilon*v \|_{\mathbb{R}^{d+1}} = 0, \quad \forall v \in L^2(\mathbb{R}^{d+1})^{d+1}. \tag{4.14}
\]

Consider any

\[
0 < \delta < \min_{1 \leq i \leq d} \left( a_i/2, T/2 \right),
\]

and define two functions \( v_\varepsilon = \rho_\varepsilon * \tau_\delta \tilde{E}v \) and \( a_\varepsilon = \rho_\varepsilon * \tau_\delta \tilde{E}Av \). Note that the \( Av_\varepsilon = a_\varepsilon \) on \( \Omega \) whenever \( \varepsilon < \delta/2 \), thanks to (4.12).
We now proceed to show that
\[ \lim_{\varepsilon \to 0} \| v_\varepsilon \|_\Omega - v \|_W = 0. \] (4.15)

Set \( \delta = 3\varepsilon \) and let \( \varepsilon < \min_{1 \leq i \leq d}(a_i/2, T/2)/3 \) go to zero. Note that
\[
\| A v_\varepsilon - Av \| \leq \| a_\varepsilon - Av \| + \| \rho_\varepsilon * \tau_\delta \tilde{E} Av - Av \|_\Omega \\
\leq \| \rho_\varepsilon * \tau_\delta \tilde{E} Av - \tilde{E} Av \|_{\mathbb{R}^{d+1}} \\
\leq \| \rho_\varepsilon * \tau_\delta \tilde{E} Av - \tau_\delta \tilde{E} Av \|_{\mathbb{R}^{d+1}} + \| \tau_\delta \tilde{E} Av - \tilde{E} Av \|_{\mathbb{R}^{d+1}}, \\
\| v_\varepsilon - v \| \leq \| \rho_\varepsilon * \tau_\delta \tilde{E} v - \tau_\delta \tilde{E} v \|_\Omega + \| \tau_\delta \tilde{E} v - v \|_\Omega \\
\leq \| \rho_\varepsilon * \tau_\delta \tilde{E} v - \tau_\delta \tilde{E} v \|_{\mathbb{R}^{d+1}} + \| \tau_\delta \tilde{E} v - \tilde{E} v \|_{\mathbb{R}^{d+1}}.
\]

Using (4.11) and (4.14) it now immediately follows that (4.15) holds.

To conclude, it suffices to prove that \( v_\varepsilon \|_\Omega \) is in \( \mathcal{V} \). Clearly, \( \tau_\delta \tilde{E} v \) is identically zero in a neighborhood of \( \Gamma_0 \). Hence we conclude that \( v_\varepsilon = \rho_\varepsilon * \tau_\delta \tilde{E} v \) vanishes on \( \Gamma_0 \) for small enough \( \varepsilon \). Next, let us examine the value of \( [v_\varepsilon]_\mu \) at points \((x, t)\) on \( \Gamma_b \), namely
\[
[v_\varepsilon]_\mu(x, t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_\varepsilon(x - x', t - t') [\tau_\delta \tilde{E} \rho](x', t') \, dx' \, dt'.
\]

Note that \( \rho_\varepsilon(x - x', t - t') \) is a symmetric function of \( x' \) about \( x \). The other term in the integrand, namely \( [\tau_\delta \tilde{E} \rho]_\mu(x', t') \), is odd about every facet of \( \Gamma_b \). Hence the integral of their product vanishes whenever \((x, t)\) is in \( \Gamma_b \). Thus, \( [v_\varepsilon]_\mu|_{\Gamma_b} = 0 \) and \( v_\varepsilon \in \mathcal{V} \).

5. The method and its error estimates

In this section, we present the approximation of the previously described broken weak formulation by the (ideal) DPG method and provide \( a \) priori and \( a \) posteriori error estimates.

5.1. The DPG method. The ideal DPG method [5] seeks \( u_h \) and \( \lambda_h \) in finite-dimensional subspaces \( U_h \subset L^2(\Omega)^{d+1} \) and \( Q_h \subset Q \), respectively, satisfying
\[ b((u_h, \lambda_h), w_h) = F(w_h), \quad \text{for all } w_h \in T(U_h \times Q_h), \] (5.1)

where \( T : L^2(\Omega)^{d+1} \times Q \to W_h \) is such that \( (T(v, \rho), w)_{W_h} = b((v, \rho), w) \) for all \( w \in W_h \) and any \( (v, \rho) \in L^2(\Omega)^{d+1} \times Q \). Hereon we denote \( T \) to be \( L^2(\Omega)^{d+1} \) and abbreviate the \( W_h \) inner product \( \langle \cdot, \cdot \rangle_{W_h} \) to simply \( \langle \cdot, \cdot \rangle_h \).

It is well known [6] that there is a mixed method that is equivalent to the above Petrov-Galerkin method (5.1). One of the variables in this mixed method is the error representation function \( \varepsilon_h \in W_h \) defined by
\[ (\varepsilon_h, w)_h = (f, w) - b((u_h, \lambda_h), w), \quad \text{for all } w \in W_h. \] (5.2)

One of the two equations in the mixed formulation given below is a restatement of this defining equation for \( \varepsilon_h \). The mixed formulation seeks \( \varepsilon_h \in W_h \) and \((u_h, \lambda_h) \in (U_h \times Q_h)\) such that
\[ (\varepsilon_h, w)_h + b((u_h, \lambda_h), w) = F(w) \quad \text{for all } w \in W_h, \]
\[ b((v, \rho), \varepsilon_h) = 0 \quad \text{for all } (v, \rho) \in U_h \times Q_h. \] (5.3)
We think of
\[ \eta = \| \varepsilon_h \|_{W_h} \equiv \left( \sum_{K \in \Omega_h} \| \varepsilon_h \|_{W(K)}^2 \right)^{1/2} \]
as an \textit{a posteriori} error estimator because \( \varepsilon_h \) can be computed from (5.2) after \( u_h \) and \( \lambda_h \) has been computed. Alternately, one can view \( \varepsilon_h \) as one of the unknowns together with \( u_h \) and \( \lambda_h \) as in (5.3). Note that (5.2) implies
\[ \eta = \sup_{w \in W_h} \frac{b((u - u_h, \lambda - \lambda_h), w)}{\|w\|_{W_h}}, \]
so it immediately follows that the estimator is globally reliable and efficient, namely
\[ \beta \| (u - u_h, \lambda - \lambda_h) \|_{U \times Q} \leq \eta \leq \|b\| \| (u - u_h, \lambda - \lambda_h) \|_{U \times Q} \]
where \( \beta \) is as in (3.8). In practice, we use element-wise norms of \( \varepsilon_h \) as \textit{a posteriori} element error indicator.

To give an \textit{a priori} error estimate with convergence rates, we need to specify all the approximation subspaces. We choose the space \( Q_h \subset Q \) by first selecting a finite element space \( V_h \subset V \) and then applying \( D_h \) to all functions in it, namely
\[ Q_h = D_h V_h. \]
This way we guarantee that \( Q_h \) is a subspace of \( Q \). The definition of \( V_h \) and the finite element subspaces of \( U \) are based on the type of elements in \( \Omega_h \). We consider two cases:

Case A: \( \Omega_h \) is a geometrically conforming mesh of \((d + 1)\)-simplices:
\begin{align*}
V_h &= \{ u \in V \cap C(\Omega)^{d+1} : u|_K \in P_{p+1}(K)^{d+1} \text{ for all } K \in \Omega_h \} \quad (5.4a) \\
U_h &= \{ u \in L^2(\Omega)^{d+1} : u|_K \in P_p(K)^{d+1} \text{ for all } K \in \Omega_h \}, \quad (5.4b)
\end{align*}
where \( P_p(K) \) is the space of polynomials of total degree \( \leq p \) on \( K \).

Case B: \( \Omega_h \) is a geometrically conforming mesh of hyperrectangles.
\begin{align*}
V_h &= \{ u \in V \cap C(\Omega)^{d+1} : u|_K \in Q_{p+1}(K)^{d+1} \text{ for all } K \in \Omega_h \} \quad (5.5a) \\
U_h &= \{ u \in L^2(\Omega)^{d+1} : u|_K \in Q_p(K)^{d+1} \text{ for all } K \in \Omega_h \}, \quad (5.5b)
\end{align*}
where \( Q_p(K) \) is the space of polynomials on \( K \) that are of degree at most \( p \) in each variable.

Since the wave operator \( A \) is a first order differential operator, \( H^1(\Omega)^{d+1} \subset W(\Omega) \). Hence, the Lagrange finite element space \( V_h \) is contained in \( W \). The space \( V_h \) has a nodal interpolation operator \( I_h : H^{s+1}(\Omega)^{d+1} \rightarrow V_h \) which is bounded for \( s + 1 > (d + 1)/2 \), which we shall use in the proof below. We will use \( C \) to denote a generic mesh-independent constant whose value at different occurrences may differ. Note that in the estimate of the theorem below, \( h \) is the discretization parameter in both space and time.

**Theorem 5.1.** Suppose \( u \in V \cap H^{s+1}(\Omega)^{d+1} \) and \( \lambda = D_h u \) solve (3.2). Suppose also that \( U_h \) and \( V_h \) are set as in (5.4) or (5.5) depending on the mesh type, and \( Q_h = D_h V_h \). Then, there exists a constant \( C > 0 \) independent of \( h \) such that the discrete solution \( u_h \in U_h \) and \( \lambda_h \in Q_h \) solving (5.1) satisfies
\[ \|u - u_h\| + \|\lambda - \lambda_h\|_Q \leq Ch^s |u|_{H^{s+1}(\Omega)^{d+1}} \quad (5.6) \]
for \((d - 1)/2 < s \leq p + 1\).

Proof. The ideal DPG method is quasioptimal, i.e., by [5, Theorem 2.2],
\[
\| (u, \lambda) - (u_h, \lambda_h) \|^2_{U \times Q} \leq C \inf_{(u_h, \lambda_h) \in U_h \times Q_h} \| (u, \lambda) - (u_h, \lambda_h) \|^2_{U \times Q} \\
\leq C \inf_{(u_h, \lambda_h) \in U_h \times Q_h} \| u - v_h \|^2 + \| \lambda - \rho_h \|^2_Q.
\]
The well-known best approximation estimates for \(U_h\) imply
\[
\inf_{v_h \in U_h} \| u - v_h \| \leq C h^s \| u \|_{H^s(\Omega)^{d+1}}, \quad \text{for all } 0 < s \leq p + 1.
\]
To estimate the remaining term, choose \(\rho_h = D_h I_h u\). Then, since \(\lambda = D_h u\), by the definition of the \(Q\)-norm in (3.1) and the Bramble-Hilbert lemma,
\[
\inf_{\rho_h \in Q_h} \| \lambda - \rho_h \|_Q \leq \| u - I_h u \|_W \\
\leq C \| u - I_h u \|_{H^1(\Omega)^{d+1}} \leq C h^s \| u \|_{H^{s+1}(\Omega)^{d+1}}
\]
for any \(u \in H^{s+1}(\Omega)\), for \((d - 1)/2 < s \leq p + 1\). Thus, from (5.7) and (5.8), we have that (5.6) holds. \(\square\)

6. Implementation and Numerical Results

We implemented the DPG discretization in the form (5.3) with the following change. Since \(W_h\) is infinite-dimensional, in order to get a practical method, we must replace \(W_h\) by a sufficiently rich finite-dimensional space \(Y^m_h\). A full theoretical analysis of this practical realization of the ideal DPG method is currently open, but we will provide numerical studies showing its efficacy in this section. For some non-negative integer \(m\), set \(Y^m_h\) as follows.

- In Case A (see (5.4)) we set \(Y^m_h = \{ w \in W_h(\Omega) : w|_K \in P_m(K)^{d+1} \}\),
- In Case B (see (5.5)) we set \(Y^m_h = \{ w \in W_h(\Omega) : w|_K \in Q_m(K)^{d+1} \}\).

Then, we compute \(e_h \in Y^m_h\), \(u_h \in U_h\) and \(\lambda_h \in Q_h\) satisfying
\[
(e_h, w)_h + b(u_h, \lambda_h, w) = F(w) \quad \text{for all } w \in Y^m_h, \\
b(v, \rho), e_h) = 0 \quad \text{for all } (v, \rho) \in U_h \times Q_h.
\]
In our numerical experience, the choice \(m = p + d + 1\) gave optimal convergence rates (as reported in detail below). This choice is motivated by the study in [14]. The choice \(m = p + d\) did not give optimal convergence rates for \(p > 2\) and \(d = 1\). A brief report of the performance of an adaptive algorithm is also included in the \(d = 1\) case. Here again, we observed marked deterioration of adaptivity if \(m = p + d\) is used instead of \(m = p + d + 1\) for higher degrees. Beyond these comments, we shall not describe these negative results further, but will henceforth focus solely on the \(m = p + d + 1\) case. All the numerical results have been implemented using the NGSolve [19] finite element software and the codes used for the experiments below are available in [20].

6.1. A null space. In order to implement (6.1), one strategy is to set \(\lambda_h = D_h z_h\) for some \(z_h \in V_h\) and solve
\[
(e_h, w)_h + b(u_h, D_h z_h, w) = F(w) \quad \text{for all } w \in Y^m_h, \\
b(v, D_h r), e_h) = 0 \quad \text{for all } v \in U_h, r \in V_h.
\]
We can decompose $V_h$ into interior “bubbles” in $V_h^0 = \{ z \in V_h : z|_{\partial K} = 0 \text{ for all } K \in \Omega_h \}$, and a remainder $V_h^1 = V_h/V_h^0$. Since $b((v, D_h V_h^0), w) = 0$, we may replace $V_h$ by $V_h^1$ in (6.2) (and compute a $z_h \in V_h^1$). Let $\{ y_k \}, \{ u_i \},$ and $\{ z_j \}$ denote a local finite element basis for $Y_h^m, U_h,$ and $V_h^1$, respectively. Using this basis, the system (6.2) with $V_h$ replaced by $V_h^1$ yields a matrix equation of the following form

$$
\begin{bmatrix}
A & B \\
B^T & 0
\end{bmatrix}
\begin{bmatrix}
e \\
x
\end{bmatrix}
=
\begin{bmatrix}
f \\
0
\end{bmatrix},
$$

(6.3)

where $e$ and $x$ are the vectors of coefficients in the basis expansion of $e_h \in Y_h^m$ and $(u_h, z_h) \in U_h \times V_h$, respectively; $A_{kl} = (y_l, y_k)_h$, $[B_0]_{ki} = b((u_i, 0), y_k)$, $[B_1]_{kj} = b((0, D_h z_j), y_k)$, and $B = [B_0, B_1]$. In all our numerical experiments, for the above-mentioned choice of $m = p + d + 1$, we observed that the matrices $A$ and $B_0$ have trivial null spaces.

However, we caution that $B_1$ may have a null space. This runs contrary to our experience with DPG methods on non-spacetime problems, so we expand on it. Note that (cf. (2.5))

$$
[B_1]_{kj} = b((0, D_h z_j), y_k) = \sum_{K \in \Omega_h} \int_{\partial K} D_{x, t} z_j \cdot y_k
$$

where

$$
D_{x, t} = \begin{bmatrix}
n_t I_d & -cn_x \\
-cn_x^T & nt_t
\end{bmatrix}
$$

and $I_d$ is the $d \times d$ identity matrix. It is immediate that on mesh facets with certain combinations of $n_x$ and $n_t$, the matrix $D_{x, t}$ is singular. Then $B_1$ will have a nontrivial kernel.

As an example, in Figure 2, we show one of the $z_j$ that is in the null space of $B_1$ on a triangular mesh for $p = 1$ and $c = 1$. In fact, on the mesh shown, there are 8 basis functions of $V^1_h$ that are in the null space of $B_1$, two for each diagonal edge. Recall that the wave speed is 1 for our model wave problem, so these edges align with the light cone for $d = 1$. In the case of $d = 2$ space dimensions, we continued to find a nontrivial null space for $B_1$ on analogous meshes.

![Figure 2. Example of a spacetime shape function $z_j$ in the kernel](image)

This null space problem occurs because the interface variable $\lambda_h$ is set indirectly by applying the singular operator $D_h$ on $V_h$. If one could directly construct a basis for $Q_h = D_h V_h$, then one can directly implement (6.1) (instead of (6.2)). However, we do not know how to construct such a basis easily on general simplicial meshes. Hence we continue on to describe how to solve (6.2) despite its kernel.
6.2. **Techniques to solve despite the null space.** Despite the above-mentioned problem, one may solve the DPG system using one of the following approaches.

6.2.1. **Technique 1: Remaining orthogonal to null space in conjugate gradients.** The matrix system (6.3) can be solved by reducing it to its Schur complement

\[ B^T A^{-1} B x = B^T A^{-1} f \]  \hspace{1cm} (6.4)

first. Let \( C = B^T A^{-1} B \) and \( g = B^T A^{-1} f \). The matrix \( C \) is symmetric and positive semi-definite. Its easy to see that

\[ \ker C = \ker B. \]

Thus solutions of (6.4) are defined only up to this kernel. Note however that since \( \ker B = \ker B_1 \) and \( B_0 \) has only the trivial kernel, the \( U_h \)-component of the DPG solution is uniquely defined independently of \( \ker B_1 \).

One may obtain one solution of (6.4) using the conjugate gradient method, which computes its \( n \)th iterate \( x_n \) in the Krylov space

\[ K_n(C, r_0) = \text{span}\{ C^k r_0 : k = 0, \ldots, n - 1 \} \]

where \( r_0 = g - A x_0 \) is the initial residual. This iteration will converge if \( K_n(C, r_0) \) remains \( (\ell^2) \) orthogonal to \( \ker(C) \) for all \( n \). A simple prescription to guarantee this orthogonality is to choose the initial iterate \( x_0 = 0 \). Indeed, if \( x_0 = 0 \), then \( r_0 = g = B^T A^{-1} f \) is in the range of \( B^T \) which equals the orthogonal complement of \( \ker B = \ker C \). Then for all \( n \geq 1 \), it’s obvious that \( C^n r_0 \) is also orthogonal to \( \ker C \). Thus \( K_n(C, r_0) \) is orthogonal to \( \ker C \).

To summarize this technique, we use the conjugate gradient algorithm to compute one solution orthogonal to \( \ker(C) \) and extract the unique \( U_h \)-component from that solution for reporting the errors.

6.2.2. **Technique 2: Regularization of the linear system.** Another technique to solve the singular system (6.4) approximately is regularization. First, we rewrite (6.4) in block form as

\[ \begin{bmatrix} B_0^T A^{-1} B_0 & B_0^T A^{-1} B_1 \\ B_1^T A^{-1} B_0 & B_1^T A^{-1} B_1 \end{bmatrix} x = g. \]

Since only \( B_1 \) may have a nontrivial kernel in \( V_h \), we can convert this to an invertible system by adding a small positive-definite term in \( V_h \). Namely, let \( M \) be the mass matrix \( M_{ij} = (z_i, z_j) \). Instead of solving (6.4), we solve for

\[ \begin{bmatrix} B_0^T A^{-1} B_0 & B_0^T A^{-1} B_1 \\ B_1^T A^{-1} B_0 & B_1^T A^{-1} B_1 + \alpha M \end{bmatrix} x = g. \]  \hspace{1cm} (6.5)

where \( \alpha \) is a positive regularization parameter, usually set much smaller than the order of the expected discretization errors. In all our reported experiments it was set to \( 10^{-9} \). The regularized system (6.5) is invertible and can be solved using any direct or iterative methods.
6.3. Convergence rates in two-dimensional spacetime. Let $\Omega = (0,1)^2$. We consider a problem with homogeneous boundary and initial conditions where the exact solution to the second order wave equation is given by $\phi(x,t) = \sin(\pi x)\sin^2(\pi t)$. Then, the exact solution for the first order system is

$$u = \begin{bmatrix} c\pi \cos(\pi x) \sin^2(\pi t) \\ \pi \sin(\pi x) \sin(2\pi t) \end{bmatrix}$$

and the corresponding source terms are

$$g = 0, \quad f = \pi^2 \sin(\pi x)(2 \cos(2\pi t) + c^2 \sin^2(\pi t)).$$

In each experiment, a (non-uniform) coarse triangular mesh of $\Omega$ was constructed, with element diameters not exceeding a reported mesh size $h$ and consider $c = 1$. Successive refinements of the mesh were obtained by connecting the mid points of the edges.

We observe in Table 1 that the order of convergence for $u_h$ in the $L^2$ norm is $O(h^{p+1})$ in accordance with Theorem 5.1. Similarly in Table 2, we observe the same convergence rates for rectangular meshes. All results in both tables were obtained using Technique 1.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$p = 0$</th>
<th>Order</th>
<th>$p = 1$</th>
<th>Order</th>
<th>$p = 2$</th>
<th>Order</th>
<th>$p = 3$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/4</td>
<td>1.2849e+00</td>
<td>–</td>
<td>1.5371e-01</td>
<td>–</td>
<td>2.0385e-02</td>
<td>–</td>
<td>1.2619e-03</td>
<td>–</td>
</tr>
<tr>
<td>1/8</td>
<td>5.6379e-01</td>
<td>1.19</td>
<td>5.6172e-02</td>
<td>1.45</td>
<td>4.7540e-03</td>
<td>2.10</td>
<td>1.5370e-04</td>
<td>+3.04</td>
</tr>
<tr>
<td>1/16</td>
<td>2.2067e-01</td>
<td>1.35</td>
<td>1.2472e-02</td>
<td>2.17</td>
<td>5.4897e-04</td>
<td>3.11</td>
<td>7.8519e-06</td>
<td>+4.29</td>
</tr>
<tr>
<td>1/32</td>
<td>1.0214e-01</td>
<td>1.11</td>
<td>3.0308e-03</td>
<td>2.04</td>
<td>6.6955e-05</td>
<td>3.00</td>
<td>4.7863e-07</td>
<td>+4.04</td>
</tr>
</tbody>
</table>

Table 1. Convergence rates for $\|u - u_h\|$ on triangular meshes using Technique 1.

6.4. Adaptivity. Let $\Omega = (0,1)^2$. We consider the same model problem (2.1) but now with zero sources $f = g = 0$ and the non-zero initial condition

$$\mu|_{t=0} = -\phi_0, \quad q|_{t=0} = \phi_0$$

in place of (2.1c), where $\phi_0 = \exp(-1000((x - 0.5)^2))$. The boundary condition $\mu = 0$ continues to remain the same. This simulates a beam reflecting off the Dirichlet boundary.

In Figure 3, we display a few iterates from the standard adaptive refinement algorithm using $p = 3$ and the DPG error estimator. We started with the extremely coarse mesh shown in Figure 3(a), used the element-wise norms of $e_h$ to compute the DPG element error indicator, marked elements with more than 50% of the total indicated error, refined the marked elements (and more for conformity) by bisection, and repeated this adaptivity loop. The few iterates from the adaptivity loop shown in Figure 3 show the potential of the spacetime DPG method to easily capture localized features in spacetime.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$p = 0$</th>
<th>Order</th>
<th>$p = 1$</th>
<th>Order</th>
<th>$p = 2$</th>
<th>Order</th>
<th>$p = 3$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/4</td>
<td>9.7226e-01</td>
<td>–</td>
<td>1.6834e-01</td>
<td>–</td>
<td>6.6722e-03</td>
<td>–</td>
<td>2.0910e-03</td>
<td>–</td>
</tr>
<tr>
<td>1/8</td>
<td>4.7357e-01</td>
<td>1.04</td>
<td>4.2869e-02</td>
<td>1.97</td>
<td>8.5059e-04</td>
<td>2.97</td>
<td>1.3308e-04</td>
<td>3.97</td>
</tr>
<tr>
<td>1/16</td>
<td>2.3291e-01</td>
<td>1.35</td>
<td>1.0763e-02</td>
<td>1.99</td>
<td>1.0707e-04</td>
<td>2.99</td>
<td>8.3773e-06</td>
<td>3.99</td>
</tr>
<tr>
<td>1/32</td>
<td>1.1587e-01</td>
<td>1.11</td>
<td>2.6935e-03</td>
<td>2.00</td>
<td>1.3409e-05</td>
<td>3.00</td>
<td>5.2613e-07</td>
<td>3.99</td>
</tr>
</tbody>
</table>

Table 2. Convergence rates for $\|u - u_h\|$ on rectangular meshes using Technique 1.
6.5. **Adaptivity with inhomogeneous materials.** Consider the case when the domain consists of two regions, namely $\Omega_l = (0, 0.5) \times (0, 1.4)$, and $\Omega_r = (0.5, 1) \times (0, 1.4)$, and a more general first order wave equation

$$\begin{bmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{bmatrix} \partial_t u - \begin{bmatrix} 0 & c \\ c & 0 \end{bmatrix} \partial_x u = 0,$$

where

$$\kappa_1 = \begin{cases} 2, & 0 < x < 1/2 \\ 1/2, & 1/2 < x < 1 \end{cases}, \quad \kappa_2 = \begin{cases} 2, & 0 < x < 1/2 \\ 1/2, & 1/2 < x < 1 \end{cases},$$

as in [12], we set $c = 1$. Here, $\kappa_1, \kappa_2$ are material parameters. The wave speed is given by $c/\sqrt{\kappa_1 \kappa_2}$, and jumps between 0.5 to 2. The impedance, given by $\kappa_1/\kappa_2$, is the same in both regions, therefore we expect no reflections between the regions. We set vanishing Dirichlet boundary conditions as the previous example and $f = g = 0$, $u_q(x, 0) = e^{-5000(x-0.2)^2}$, and $u_{\mu}(x, 0) = -e^{-5000(x-0.2)^2}$.

We can observe the results of adaptive algorithm in Figure 4.

6.6. **Convergence rates in three-dimensional spacetime.** On $\Omega = (0, 1)^3$, we consider the problem where the exact solution to the second order wave equation is given by $\phi(x, t) = \sin(\pi x) \sin(\pi y) t^2$. This corresponds to

$$u = \begin{bmatrix} \pi \cos(\pi x) \sin(\pi y) t^2 \\ \pi \cos(\pi y) \sin(\pi x) t^2 \\ 2 \sin(\pi x) \sin(\pi y) t \end{bmatrix},$$

Figure 3. Iterates from the adaptive algorithm. Numerical pressure $\mu$ is shown for $p = 3$. Time axis is vertical.
Figure 4. Iterative from the adaptive algorithm. Numerical pressure $\mu$ is shown for $p = 1$.

\[ f = \sin(\pi x) \sin(\pi y)(2 + 2\pi^2 t^2) \] and $g = 0$.

In Table 3, we show the convergence rates of $u_h$ for successively refined tetrahedral meshes, obtained using Technique 2 for $p = 0, 1, 2, 3$. Table 4 shows analogous results obtained for successively refined hexahedral meshes using Technique 1. In all these cases, we observe $O(h^{p+1})$ convergence rates for $u_h$.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$p = 0$</th>
<th>Order</th>
<th>$p = 1$</th>
<th>Order</th>
<th>$p = 2$</th>
<th>Order</th>
<th>$p = 3$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>6.0557e-01</td>
<td>0.58</td>
<td>1.3924e-01</td>
<td>1.78</td>
<td>1.3912e-02</td>
<td>3.35</td>
<td>3.2845e-03</td>
<td>3.74</td>
</tr>
<tr>
<td>1/4</td>
<td>3.3896e-01</td>
<td>0.84</td>
<td>3.3508e-02</td>
<td>2.05</td>
<td>1.4769e-03</td>
<td>3.24</td>
<td>1.6490e-04</td>
<td>4.32</td>
</tr>
<tr>
<td>1/8</td>
<td>1.5469e-01</td>
<td>1.13</td>
<td>8.9554e-03</td>
<td>1.90</td>
<td>1.7210e-04</td>
<td>3.10</td>
<td>9.9691e-06</td>
<td>4.05</td>
</tr>
</tbody>
</table>

Table 3. Convergence rates for $\|u - u_h\|$ on tetrahedral meshes obtained using Technique 2.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$p = 0$</th>
<th>Order</th>
<th>$p = 1$</th>
<th>Order</th>
<th>$p = 2$</th>
<th>Order</th>
<th>$p = 3$</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1149e+00</td>
<td>–</td>
<td>6.0068e-01</td>
<td>–</td>
<td>2.8828e-02</td>
<td>–</td>
<td>3.3262e-02</td>
<td>–</td>
</tr>
<tr>
<td>1/2</td>
<td>7.5769e-01</td>
<td>0.56</td>
<td>1.5124e-01</td>
<td>1.99</td>
<td>2.8264e-03</td>
<td>3.35</td>
<td>2.0540e-03</td>
<td>4.02</td>
</tr>
<tr>
<td>1/4</td>
<td>4.2035e-01</td>
<td>0.85</td>
<td>3.8592e-02</td>
<td>1.97</td>
<td>3.5256e-04</td>
<td>3.00</td>
<td>1.3234e-04</td>
<td>3.96</td>
</tr>
<tr>
<td>1/8</td>
<td>2.1338e-01</td>
<td>0.98</td>
<td>9.6918e-03</td>
<td>1.99</td>
<td>3.8023e-05</td>
<td>3.21</td>
<td>9.3766e-06</td>
<td>3.82</td>
</tr>
</tbody>
</table>

Table 4. Convergence rates for $\|u - u_h\|$ on hexahedral meshes using Technique 1.
6.7. **Adaptivity in 3D.** Consider $\Omega = (0, 1)^3$, and the problem where the exact solution is given by

$$u(x, y, t) = e^{-200((x-x_0-ct)^2 + (y-y_0-ct)^2)} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Here we have chosen $x_0 = y_0 = 0.2$. This corresponds to set $f = 0$, and

$$g = 400c e^{-200((x-x_0-ct)^2 + (y-y_0-ct)^2)} \begin{bmatrix} y - y_0 - ct \\ x - x_0 - ct \end{bmatrix}.$$ 

After setting $c = 1/2$ and homogeneous Dirichlet boundary conditions, we observe that the adaptive scheme captures with precision the behavior of the wave propagation in Figure 5.

![Figure 5. Adaptivity example in three dimensions after 10 iterations, $u_\mu$ component is shown, and $p = 1$.](image)

6.8. **Adaptivity with variant wave speed.** Consider $\Omega = (-4, 4)^2 \times (0, 8)$ and the exact solution

$$u(x, y, t) = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} e^{-20((x-c\cos(\frac{\pi}{2} t))\sin(\frac{\pi}{2} t) + (y+c\sin(\frac{\pi}{2} t))\cos(\frac{\pi}{2} t))}.$$ 

After setting $c = 1$, the component $u_\mu$ corresponds to a pulse propagating from the coordinates $(1, 0, 0)$ (at time $t = 0$), to $(1, 0, 8)$ (at time $t = 8$), rotating in time with respect to the $t$-axis with distance equals $1$. We have chosen the solution so we observe two complete rotations from $t = 0$ to $t = 8$. (see Figure 6).

Consider the initial condition $u(x, y, 0) = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} e^{-20((x-1)^2 + y^2)}$, with homogeneous Dirichlet boundary conditions, and set

$$g = 40c \left[ x\left(\frac{\pi}{2} \sin\left(\frac{\pi}{2} t\right) + 1\right) + \cos\left(\frac{\pi}{2} t\left(\frac{\pi}{2} y - c\right)\right)\right] u_\mu(x, y, t),$$

$$f = -40c \left[ x\left(\frac{\pi}{2} \sin\left(\frac{\pi}{2} t\right) + 1\right) + y\left(\frac{\pi}{2} \cos\left(\frac{\pi}{2} t\right) + 1\right) - \sqrt{2} c \sin\left(\frac{\pi}{2} + \frac{1}{4}\right)\right] u_\mu(x, y, t).$$
(a) Spacetime contour of the solution $u_\mu$.

(b) Slides of the solution $u_\mu$ at particular time steps.

Figure 6. Iterate from the adaptive algorithm. Numerical pressure $\mu$ is shown for $p = 1$.

REFERENCES


Portland State University, Po Box 751, Portland, OR 97207-0751, USA.
E-mail address: gjay@pdx.edu

Basque Center for Applied Mathematics, Mazarredo 14, Bilbao, E48009, Spain.
E-mail address: psepulveda@bcamath.org