Advanced Modeling and Simulation
in Engineering Sciences
a SpringerOpen Journal

RESEARCH ARTICLE

Open Access

CrossMark

Stabilization in relation to wavenumber in HDG methods

Jay Gopalakrishnan[1], Stéphane Lanteri[2], Nicole Olivares[1*] and Ronan Perrussel[3]

*Correspondence: nmo@ pdx.edu
[1] Portland State University, PO Box 751, Portland, OR 97207-0751, USA
Full list of author information is available at the end of the article

Abstract

Background: Simulation of wave propagation through complex media relies on proper understanding of the properties of numerical methods when the wavenumber is real and complex.

Methods: Numerical methods of the Hybrid Discontinuous Galerkin (HDG) type are considered for simulating waves that satisfy the Helmholtz and Maxwell equations. It is shown that these methods, when wrongly used, give rise to singular systems for complex wavenumbers.

Results: A sufficient condition on the HDG stabilization parameter for guaranteeing unique solvability of the numerical HDG system, both for Helmholtz and Maxwell systems, is obtained for complex wavenumbers. For real wavenumbers, results from a dispersion analysis are presented. An asymptotic expansion of the dispersion relation, as the number of mesh elements per wave increase, reveal that some choices of the stabilization parameter are better than others.

Conclusions: To summarize the findings, there are values of the HDG stabilization parameter that will cause the HDG method to fail for complex wavenumbers. However, this failure is remedied if the real part of the stabilization parameter has the opposite sign of the imaginary part of the wavenumber. When the wavenumber is real, values of the stabilization parameter that asymptotically minimize the HDG wavenumber errors are found on the imaginary axis. Finally, a dispersion analysis of the mixed hybrid Raviart–Thomas method showed that its wavenumber errors are an order smaller than those of the HDG method.

Keywords: HDG, Raviart–Thomas, Dispersion, Dissipation , Absorbing material, Complex, Wave speed, Optimal , Stabilization, Helmholtz, Maxwell, Unisolvency

Background

Wave propagation through complex structures, composed of both propagating and absorbing media, are routinely simulated using numerical methods. Among the various numerical methods used, the Hybrid Discontinuous Galerkin (HDG) method has emerged as an attractive choice for such simulations. The easy passage to high order using interface unknowns, condensation of all interior variables, availability of error estimators and adaptive algorithms, are some of the reasons for the adoption of HDG methods.

It is important to design numerical methods that remain stable as the wavenumber varies in the complex plane. For example, in applications like computational lithography, one finds absorbing materials with complex refractive index in parts of the domain of simulation. Other examples are furnished by meta-materials. A separate and important

Springer

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 2 of 24

reason for requiring such stability emerges in the computation of resonances by iterative searches in the complex plane. It is common for such iterative algorithms to solve a source problem with a complex wavenumber as its current iterate. Within such algorithms, if the HDG method is used for discretizing the source problem, it is imperative that the method remains stable for all complex wavenumbers.

One focus of this study is on complex wavenumber cases in acoustics and electromagnetics, motivated by the above-mentioned examples. Ever since the invention of the HDG method in Ref. [1], it has been further developed and extended to other problems in many works (so many so that it is now impractical to list all references on the subject here). Of particular interest to us are works that applied HDG ideas to wave propagation problems such as [2–8]. We will make detailed comparisons with some of these works in a later section. However, none of these references address the stability issues for complex wavenumber cases. While the choice of the HDG stabilization parameter in the real wave number case can be safely modeled after the well-known choices for elliptic problems [9], the complex wave number case is essentially different. This will be clear right away from a few elementary calculations in the next section, which show that the standard prescriptions of stabilization parameters are not always appropriate for the complex wave number case. This then raises further questions on how the HDG stabilization parameter should be chosen in relation to the wavenumber, which are addressed in later sections.

Another focus of this study is on the difference in speeds of the computed and the exact wave, in the case of real wavenumbers. By means of a dispersion analysis, one can compute the discrete wavenumber of a wave-like solution computed by the HDG method, for any given exact wavenumber. An extensive bibliography on dispersion analyses for the standard finite element method can be obtained from Refs. [10, 11]. For nonstandard finite element methods however, dispersion analysis is not so common [12], and for the HDG method, it does not yet exist. We will show that useful insights into the HDG method can be obtained by a dispersion analysis. In multiple dimensions, the discrete wavenumber depends on the propagation angle. Analytic computation of the dispersion relation is feasible in the lowest order case. We are thus able to study the influence of the stabilization parameter on the discrete wavenumber and offer recommendations on choosing good stabilization parameters. The optimal stabilization parameter values are found not to depend on the wavenumber. In the higher order case, since analytic calculations pose difficulties, we conduct a dispersion analysis numerically.

We begin, in the next section, by describing the HDG methods. We set the stage for this study by showing that the commonly chosen HDG stabilization parameter values for elliptic problems are not appropriate for all complex wavenumbers. In the subsequent section, we discover a constraint on the stabilization parameter, dependent on the wavenumber, that guarantees unique solvability of both the global and the local HDG problems. Afterward, we perform a dispersion analysis for both the HDG method and a mixed method and discuss the results.

Methods of the HDG type

We borrow the basic methodology for constructing HDG methods from Ref. [1] and apply it to the time-harmonic Helmholtz and Maxwell equations (written as first order systems). While doing so, we set up the notations used throughout, compare the formulation we use

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 3 of 24

with other existing works, and show that for complex wavenumbers there are stabilization parameters that will cause the HDG method to fail.

Undesirable HDG stabilization parameters for the Helmholtz system

We begin by considering the lowest order HDG system for Helmholtz equation. Let k be a complex number. Consider the Helmholtz system on $\Omega \subset \mathbb{R}^2$ with homogeneous Dirichlet boundary conditions,

$$\hat{i}k\vec{\mathcal{U}} + \vec{\nabla}\Phi = \vec{0}, \qquad \text{in} \quad \Omega, \tag{1a}$$

$$\hat{i}k\Phi + \vec{\nabla}\cdot\vec{\mathcal{U}} = f, \qquad \text{in} \quad \Omega, \tag{1b}$$

$$\Phi = 0, \qquad \text{on} \quad \partial\Omega, \tag{1c}$$

where $f \in L^2(\Omega)$. Note that the second order form of the Helmholtz equation, $-\Delta\Phi - k^2\Phi = \hat{i}kf$ can be recovered by eliminating the $\vec{\mathcal{U}}$ variable. Also, although it is straightforward to define the method for spatially varying k, in this paper all our results will be stated only for constant k.

Let $\mathcal{T}_h$ denote a square or triangular mesh of disjoint elements K, so $\overline{\Omega} = \cup_{K \in \mathcal{T}_h}\overline{K}$, and let $\mathcal{F}_h$ denote the collection of edges. The HDG method produces an approximation $(\vec{u}, \phi, \hat{\phi})$ to the exact solution $(\vec{\mathcal{U}}, \Phi, \hat{\Phi})$, where $\hat{\Phi}$ denotes the trace of Φ on the collection of element boundaries $\partial\mathcal{T}_h$. The HDG solution $(\vec{u}, \phi, \hat{\phi})$ is in the finite dimensional space $V_h \times W_h \times M_h$ defined by

$$V_h = \left\{\vec{v} \in (L^2(\Omega))^2 \; : \; \vec{v}|_K \in V(K), \, \forall K \in \mathcal{T}_h\right\}$$

$$W_h = \left\{\psi \in L^2(\Omega) \; : \; \psi|_K \in W(K), \, \forall K \in \mathcal{T}_h\right\}$$

$$M_h = \left\{\hat{\psi} \in L^2\left(\bigcup_{F \in \mathcal{F}_h} F\right) \; : \; \hat{\psi}|_F \in M(F), \, \forall F \in \mathcal{F}_h \text{ and } \hat{\psi}|_{\partial\Omega} = 0\right\},$$

with polynomial spaces $V(K)$, $W(K)$, and $M(F)$ specified differently depending on element type:

Triangles	Squares
$V(K) = (\mathcal{P}_p(K))^2$	$V(K) = (\mathcal{Q}_p(K))^2$
$W(K) = \mathcal{P}_p(K)$	$W(K) = \mathcal{Q}_p(K)$
$M(F) = \mathcal{P}_p(F)$	$M(F) = \mathcal{P}_p(F).$

Here, for a given domain D, $\mathcal{P}_p(D)$ denotes polynomials of degree at most p, and $\mathcal{Q}_p(D)$ denotes polynomials of degree at most p in each variable.

The HDG solution satisfies

$$\sum_{K \in \mathcal{T}_h} \hat{i}k(\vec{u}, \vec{v})_K - (\phi, \vec{\nabla}\cdot\vec{v})_K + \langle\hat{\phi}, \vec{v}\cdot\vec{n}\rangle_{\partial K} = 0, \tag{2a}$$

$$\sum_{K \in \mathcal{T}_h} -(\vec{\nabla}\cdot\vec{u}, \psi)_K + \langle\tau\hat{\phi}, \psi\rangle_{\partial K} - \langle\tau\phi, \psi\rangle_{\partial K} - \hat{i}k(\phi, \psi)_K = -(f, \psi)_\Omega, \tag{2b}$$

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 4 of 24

$$\sum_{K \in \mathcal{T}_h} \langle \vec{u} \cdot \vec{n} + \tau(\phi - \hat{\phi}), \hat{\psi} \rangle_{\partial K} = 0, \tag{2c}$$

for all $\vec{v} \in V_h$, $\psi \in W_h$, and $\hat{\psi} \in M_h$. The last equation enforces the conservativity of the numerical flux

$$\hat{u} \cdot \vec{n} = \vec{u} \cdot \vec{n} + \tau(\phi - \hat{\phi}). \tag{3}$$

The stabilization parameter τ is assumed to be constant on each ∂K. We are interested in how the choice of τ in relation to k affects the method, especially when k is complex valued. Comparisons of this formulation with other HDG formulations for Helmholtz equations in the literature are summarized in Table 1.

One of the main reasons to use an HDG method is that all interior unknowns $(\vec{u}, \phi)$ can be eliminated to get a global system for solely the interface unknowns $(\hat{\phi})$. This is possible whenever the local system

$$\hat{\imath} k (\vec{u}, \vec{v})_K - (\phi, \vec{\nabla} \cdot \vec{v})_K = -\langle \hat{\phi}, \vec{v} \cdot \vec{n} \rangle_{\partial K}, \quad \forall \vec{v} \in V(K), \tag{4a}$$

$$\begin{aligned} -(\vec{\nabla} \cdot \vec{u}, \psi)_K - \langle \tau \phi, \psi \rangle_{\partial K} - \hat{\imath} k(\phi, \psi)_K &= -\langle \tau \hat{\phi}, \psi \rangle_{\partial K} \\ &\quad - (f, \psi)_K, \quad \forall \psi \in W(K), \end{aligned} \tag{4b}$$

is uniquely solvable. (For details on this elimination and other perspectives on HDG methods, see [1].) In the lowest order ($p = 0$) case, on a square element K of side length h, if we use a basis in the following order

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \phi_1 = 1, \qquad \text{on} \quad K,$$

then the element matrix for the system (4a, 4b) is

$$M = \begin{bmatrix} \hat{\imath} k\, h^2 & 0 & 0 \\ & \hat{\imath} k\, h^2 & 0 \\ 0 & 0 & -4\, h\tau - \hat{\imath} k\, h^2 \end{bmatrix}.$$

Table 1 Comparison with some HDG formulations in other papers

Reference	Their notations and equations	Connection to our formulation
Helmholtz case [2]	$\vec{q}_{[2]} + \vec{\nabla} u_{[2]} = \vec{0}$ $\vec{\nabla} \cdot \vec{q}_{[2]} - k^2 u_{[2]} = 0$ $\hat{q}_{[2]} \cdot \vec{n} = \vec{q}_{[2]} \cdot \vec{n} + \hat{\imath}\tau_{[2]}(u_{[2]} - \hat{u}_{[2]})$	$\tau_{[2]} = k\,\tau$ $\hat{\imath} k u_{[2]} = \phi$ $\vec{q}_{[2]} = \vec{u}$
Helmholtz case [4]	$\hat{\imath} k \vec{q}_{[4]} + \vec{\nabla} u_{[4]} = \vec{0}$ $\hat{\imath} k u_{[4]} + \vec{\nabla} \cdot \vec{q}_{[4]} = 0$ $\hat{q}_{[4]} \cdot \vec{n} = \vec{q}_{[4]} \cdot \vec{n} + \tau_{[4]}(u_{[4]} - \hat{u}_{[4]})$	$\tau_{[4]} = \tau$ $u_{[4]} = \phi$ $\vec{q}_{[4]} = \vec{u}$
2D Maxwell case [6]	$\hat{\imath}\omega_{[6]}\varepsilon_r E_{[6]} - \nabla \times \vec{H}_{[6]} = 0$ $\hat{\imath}\omega_{[6]}\mu_r \vec{H}_{[6]} + \vec{\nabla} \times E_{[6]} = \vec{0}$ $\hat{H}_{[6]} = \vec{H}_{[6]} + \tau_{[6]}(E_{[6]} - \hat{E}_{[6]})\vec{t}$	$\tau_{[6]} = \sqrt{\frac{\varepsilon_r}{\mu_r}}\tau$ $\omega_{[6]} = \omega\sqrt{\varepsilon_0\mu_0}$ $E_{[6]} = \frac{1}{\sqrt{\varepsilon_r}}E$, $\vec{H}_{[6]} = \frac{1}{\sqrt{\mu_r}}\vec{H}$
Maxwell case [8]	$\mu\vec{w}_{[8]} - \vec{\nabla} \times \vec{u}_{[8]} = \vec{0}$ $\vec{\nabla} \times \vec{w}_{[8]} - \varepsilon\omega^2\vec{u}_{[8]} = \vec{0}$ $\hat{w}_{[8]} = \vec{w}_{[8]} + \tau_{[8]}(\vec{u}_{[8]} - \hat{u}_{[8]}) \times \vec{n}$	$\tau_{[8]} = \hat{\imath}\sqrt{\frac{\varepsilon\omega^2}{\mu}}\tau$ $\mu\vec{w}_{[8]} = -\hat{\imath} k\vec{H}$, with $k = \omega\sqrt{\mu\varepsilon}$, $\vec{u}_{[8]} = \vec{E}$

Notations in the indicated external references are used after subscripting them by the reference number. Notations without subscripts are those defined in this paper.

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 5 of 24

This shows that if

$$4\tau = -\hat{\imath}kh, \tag{5}$$

then M is singular, and so the HDG method will fail. *The usual recipe of choosing $\tau = 1$ is therefore inappropriate when k is complex valued.* Indeed, if $\tau = 1$ and a complex wave number k happens to be very near $4\hat{\imath}/h$, the element matrices will be so close to being singular that interior variables cannot be reliably condensed out.

Intermediate case of the 2D Maxwell system

It is an interesting exercise to consider the 2D Maxwell system before going to the full 3D case. In fact, the HDG method for the 2D Maxwell system can be determined from the HDG method for the 2D Helmholtz system. The 2D Maxwell system is

$$\hat{\imath}k\mathcal{E} - \nabla \times \vec{\mathcal{H}} = -J, \tag{6a}$$

$$\hat{\imath}k\vec{\mathcal{H}} + \vec{\nabla} \times \mathcal{E} = 0, \tag{6b}$$

where $J \in L^2(\Omega)$, and the scalar curl $\nabla \times \cdot$ and the vector curl $\vec{\nabla} \times \cdot$ are defined by

$$\nabla \times \vec{\mathcal{H}} = \partial_1 \mathcal{H}_2 - \partial_2 \mathcal{H}_1 = \vec{\nabla} \cdot R(\vec{\mathcal{H}}), \qquad \vec{\nabla} \times \mathcal{E} = (\partial_2 \mathcal{E}, -\partial_1 \mathcal{E}) = R(\vec{\nabla}\mathcal{E}).$$

Here $R(v_1, v_2) = (v_2, -v_1)$ is the operator that rotates vectors counterclockwise by $+\pi/2$ in the plane. Clearly, if we set $\vec{r} = -R\vec{\mathcal{H}}$, then (6a, 6b) becomes

$$\hat{\imath}k\mathcal{E} + \vec{\nabla} \cdot \vec{r} = -J,$$
$$-\hat{\imath}k\vec{r} + RR\vec{\nabla}\mathcal{E} = 0,$$

which, since $RR\vec{v} = -\vec{v}$ (rotation by π), coincides with (1a, 1b, 1c) with $\Phi = \mathcal{E}$, $\vec{\mathcal{U}} = \vec{r}$, and $f = -J$. This also shows that the HDG method for Helmholtz equation should yield an HDG method for the 2D Maxwell system. We thus conclude that there exist stabilization parameters that will cause the HDG system for 2D Maxwell system to fail.

To examine this 2D HDG method, if we let $\vec{H}$ and E denote the HDG approximations for $R\vec{r}$ and $\mathcal{E}$, respectively, then the HDG system (2a, 2b, 2c) with $\vec{u}$ and ϕ replaced by $-R\vec{H}$ and E, respectively, gives

$$\sum_{K \in \mathcal{T}_h} -(E, \vec{\nabla} \times \vec{w})_K + \langle \hat{E}, \vec{n} \times \vec{w} \rangle_{\partial K} - \hat{\imath}k(\vec{H}, \vec{w})_K = 0,$$

$$\sum_{K \in \mathcal{T}_h} \hat{\imath}k(E, \psi)_K - (\vec{\nabla} \times \vec{H}, \psi)_K + \langle \tau(E - \hat{E}), \psi \rangle_{\partial K} = -(\vec{J}, \psi)_\Omega,$$

$$\sum_{K \in \mathcal{T}_h} \langle \widehat{R\vec{H}} \cdot \vec{n}, \hat{\psi} \rangle_{\partial K} = 0,$$

for all $\vec{w} \in R(V_h), \psi \in W_h$ and $\hat{\psi} \in M_h$. We have used the fact that $-(R\vec{H}) \cdot \vec{n} = \vec{H} \cdot \vec{t}$, where $\vec{t} = R\vec{n}$ the tangent vector, and we have used the 2D cross product defined by $\vec{v} \times \vec{n} = \vec{v} \cdot \vec{t}$. In particular, the numerical flux prescription (3) implies

$$-\widehat{R\vec{H}} \cdot \vec{n} = -R\vec{H} \cdot \vec{n} + \tau(E - \hat{E}),$$

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 6 of 24

where $\widehat{R\vec{H}}$ denotes the numerical trace of $R\vec{H}$. We rewrite this in terms of $\vec{H}$ and E, to obtain

$$\hat{H} \cdot \vec{t} = \vec{H} \cdot \vec{t} + \tau (E - \hat{E}).$$

One may rewrite this again, as

$$\hat{H} \times \vec{n} = \vec{H} \times \vec{n} + \tau (E - \hat{E}). \tag{7}$$

This expression is notable because it will help us consistently transition the numerical flux prescription from the Helmholtz to the full 3D Maxwell case discussed next. A comparison of this formula with those in the existing literature is included in Table 1.

The 3D Maxwell system

Consider the 3D Maxwell system on $\Omega \subset \mathbb{R}^3$ with a perfect electrically conducting boundary condition,

$$\hat{\imath} k \vec{\mathcal{E}} - \vec{\nabla} \times \vec{\mathcal{H}} = -\vec{J}, \quad \text{in} \quad \Omega, \tag{8a}$$

$$\hat{\imath} k \vec{\mathcal{H}} + \vec{\nabla} \times \vec{\mathcal{E}} = \vec{0}, \qquad \text{in} \quad \Omega, \tag{8b}$$

$$\vec{n} \times \vec{\mathcal{E}} = \vec{0}, \qquad \text{on} \quad \partial\Omega, \tag{8c}$$

where $\vec{J} \in (L^2(\Omega))^3$. For this problem, $\mathcal{T}_h$ denotes a cubic or tetrahedral mesh, and $\mathcal{F}_h$ denotes the collection of mesh faces. The HDG method approximates the exact solution $(\vec{\mathcal{E}}, \vec{\mathcal{H}}, \hat{\mathcal{E}})$ by the discrete solution $(\vec{E}, \vec{H}, \hat{E}) \in Y_h \times Y_h \times N_h$. The discrete spaces are defined by

$$Y_h = \left\{ \vec{v} \in (L^2(\Omega))^3 : \vec{v}|_K \in Y(K), \ \forall K \in \mathcal{T}_h \right\},$$

$$N_h = \left\{ \hat{\eta} \in (L^2(\mathcal{F}_h))^3 : \hat{\eta}|_F \in N(F), \ \forall F \in \mathcal{F}_h \text{ and } \hat{\eta}|_{\partial\Omega} = \vec{0} \right\},$$

with polynomial spaces $Y(K)$ and $N(F)$ specified by:

Tetrahedra	Cubes
$Y(K) = (\mathcal{P}_p(K))^3$	$Y(K) = (\mathcal{Q}_p(K))^3$
$N(F) = \left\{ \hat{\eta} \in (\mathcal{P}_p(F))^3 \ : \ \hat{\eta} \cdot \vec{n} = 0 \right\}$	$N(F) = \left\{ \hat{\eta} \in (\mathcal{Q}_p(F))^3 \ : \ \hat{\eta} \cdot \vec{n} = 0 \right\}$

Our HDG method for (8a, 8b, 8c) is

$$\sum_{K \in \mathcal{T}_h} \hat{\imath} k (\vec{E}, \vec{v})_K - (\vec{\nabla} \times \vec{H}, \vec{v})_K + \langle (\hat{H} - H) \times \vec{n}, \ \vec{v} \rangle_{\partial K} = -(\vec{J}, \vec{v})_\Omega, \quad \forall \vec{v} \in Y_h,$$

$$\sum_{K \in \mathcal{T}_h} -(\vec{E}, \vec{\nabla} \times \vec{w})_K + \langle \hat{E}, \vec{n} \times \vec{w} \rangle_{\partial K} - \hat{\imath} k (\vec{H}, \vec{w})_K = 0, \quad \forall \vec{w} \in Y_h,$$

$$\sum_{K \in \mathcal{T}_h} \langle \hat{H} \times \vec{n}, \ \hat{w} \rangle_{\partial K} = 0, \quad \forall \hat{w} \in N_h,$$

where, in analogy with (7), we now set numerical flux by

$$\hat{H} \times \vec{n} = \vec{H} \times \vec{n} + \tau (\vec{E} - \hat{E})_t, \tag{9}$$

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 7 of 24

where $(\vec{E} - \hat{E})_t$ denotes the tangential component, or equivalently

$$\hat{H} \times \vec{n} = \vec{H} \times \vec{n} + \tau(\vec{n} \times (\vec{E} - \hat{E})) \times \vec{n}.$$

Note that the 2D system (6a, 6b) is obtained from the 3D Maxwell system (8a, 8b, 8c) by assuming symmetry in x_3-direction. Hence, for consistency between 2D and 3D formulations, we should have the same form for the numerical flux prescriptions in 2D and 3D.

The HDG method is then equivalently written as

$$\sum_{K \in \mathcal{T}_h} \hat{\imath} k(\vec{E}, \vec{v})_K - (\vec{\nabla} \times \vec{H}, \vec{v})_K + \langle \tau(\vec{E} - \hat{E}) \times \vec{n}, \ \vec{v} \times \vec{n} \rangle_{\partial K} = -(\vec{J}, \vec{v})_\Omega, \tag{10a}$$

$$\sum_{K \in \mathcal{T}_h} -(\vec{E}, \vec{\nabla} \times \vec{w})_K + \langle \hat{E}, \vec{n} \times \vec{w} \rangle_{\partial K} - \hat{\imath} k(\vec{H}, \vec{w})_K = 0, \tag{10b}$$

$$\sum_{K \in \mathcal{T}_h} \langle \vec{H} + \tau \, \vec{n} \times (\vec{E} - \hat{E}), \ \hat{w} \times \vec{n} \rangle_{\partial K} = 0, \tag{10c}$$

for all $\vec{v}, \vec{w} \in Y_h$, and $\hat{w} \in N_h$. For comparison with other existing formulations, see Table 1.

Again, let us look at the solvability of the *local element problem*

$$\hat{\imath} k(\vec{E}, \vec{v})_K - (\vec{\nabla} \times \vec{H}, \vec{v})_K + \langle \tau \vec{E} \times \vec{n}, \ \vec{v} \times \vec{n} \rangle_{\partial K} = \langle \tau \hat{E} \times \vec{n}, \ \vec{v} \times \vec{n} \rangle_{\partial K} - (\vec{J}, \vec{v})_K, \tag{11a}$$

$$-(\vec{E}, \vec{\nabla} \times \vec{w})_K - \hat{\imath} k(\vec{H}, \vec{w})_K = -\langle \hat{E}, \vec{n} \times \vec{w} \rangle_{\partial K}, \tag{11b}$$

for all $\vec{v}, \vec{w} \in Y(K)$. In the lowest order ($p = 0$) case, on a cube element K of side length h, if we use a basis in the following order

$$\vec{E}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{E}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{E}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{H}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{H}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{H}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \tag{12}$$

then the 6×6 element matrix for the system (11a, 11b) is

$$M = \begin{bmatrix} (4h^2\tau + \hat{\imath} k h^3) I_3 & 0 \\ 0 & -(\hat{\imath} k h^3) I_3 \end{bmatrix},$$

where I_3 denotes the 3×3 identity matrix.

Again, exactly as in the Helmholtz case—cf. (5)—we find that if

$$4\tau = -\hat{\imath} k h, \tag{13}$$

then the local static condensation required in the HDG method will fail in the Maxwell case also.

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 8 of 24

Behavior on tetrahedral meshes

For the lowest order ($p = 0$) case on a tetrahedral element, just as for the cube element described above, there are bad stabilization parameter values. Consider, for example, the tetrahedral element of size h defined by

$$K = \left\{ \vec{x} \in \mathbb{R}^3 \ : \ x_j \geq 0 \ \forall j, \ x_1 + x_2 + x_3 \leq h \right\}, \tag{14}$$

with a basis ordered as in (12). The element matrix for the system (11a, 11b) is then

$$M = \frac{1}{6} \begin{bmatrix} (2\sqrt{3}+6)h^2\tau + \hat{i}kh^3 & -\sqrt{3}h^2\tau & -\sqrt{3}h^2\tau & 0 & 0 & 0 \\ -\sqrt{3}h^2\tau & (2\sqrt{3}+6)h^2\tau + \hat{i}kh^3 & -\sqrt{3}h^2\tau & 0 & 0 & 0 \\ -\sqrt{3}h^2\tau & -\sqrt{3}h^2\tau & 4h^2\tau + \hat{i}kh^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\hat{i}kh^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\hat{i}kh^3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\hat{i}kh^3 \end{bmatrix}.$$

We immediately see that the rows become linearly dependent if

$$(3\sqrt{3}+6)\tau = -\hat{i}kh.$$

Hence, for $\tau = -\hat{i}kh/(3\sqrt{3}+6)$, the HDG method will fail on tetrahedral meshes.

For orders $p \geq 1$, the element matrices are too complex to find bad parameter values so simply. Instead, we experiment numerically. Setting $\tau = -\hat{i}$, which is equivalent to the choice made in Ref. [8] (see Table 1), we compute the smallest singular value of the element matrix M [the matrix of the left hand side of (11a, 11b) with K set by (14)] for a range of normalized wavenumbers kh. Figure 1a, b show that, for orders $p = 1$ and $p = 2$, there are values of kh for which $\tau = -\hat{i}$ results in a singular value very close to zero. Taking a closer look at the first nonzero local minimum in Figure 1a, we find that the local matrix corresponding to normalized wavenumber $kh \approx 7.49$ has an estimated condition number exceeding 3.9×10^{15}, i.e., for all practical purposes, the element matrix is singular. To illustrate how a different choice of stabilization parameter τ can affect the conditioning of the element matrix, Figure 1c, d show the smallest singular values for the same range of kh, but with $\tau = 1$. Clearly the latter choice of τ is better than the former. In other unreported experiments we observed similar behavior for orders up to $p = 5$.

From another perspective, Figure 1e shows the smallest singular value of the element matrix as τ is varied in the complex plane, while fixing kh to 1. Figure 1f is similar except that we fixed kh to the value discussed above, approximately 7.49. In both cases, we find that the values of τ that yielded the smallest singular values are along the imaginary axis. Finally, in Figure 1g, h, we see the effects of multiplying these real values of kh by $1 + \hat{i}$. The region of the complex plane where such values of τ are found changes significantly when kh is complex.

Results on unisolvent stabilization

We now turn to the question of how we can choose a value for the stabilization parameter τ that will guarantee that the local matrices are not singular. The answer, given by a condition on τ, surprisingly also guarantees that the global condensed HDG matrix is
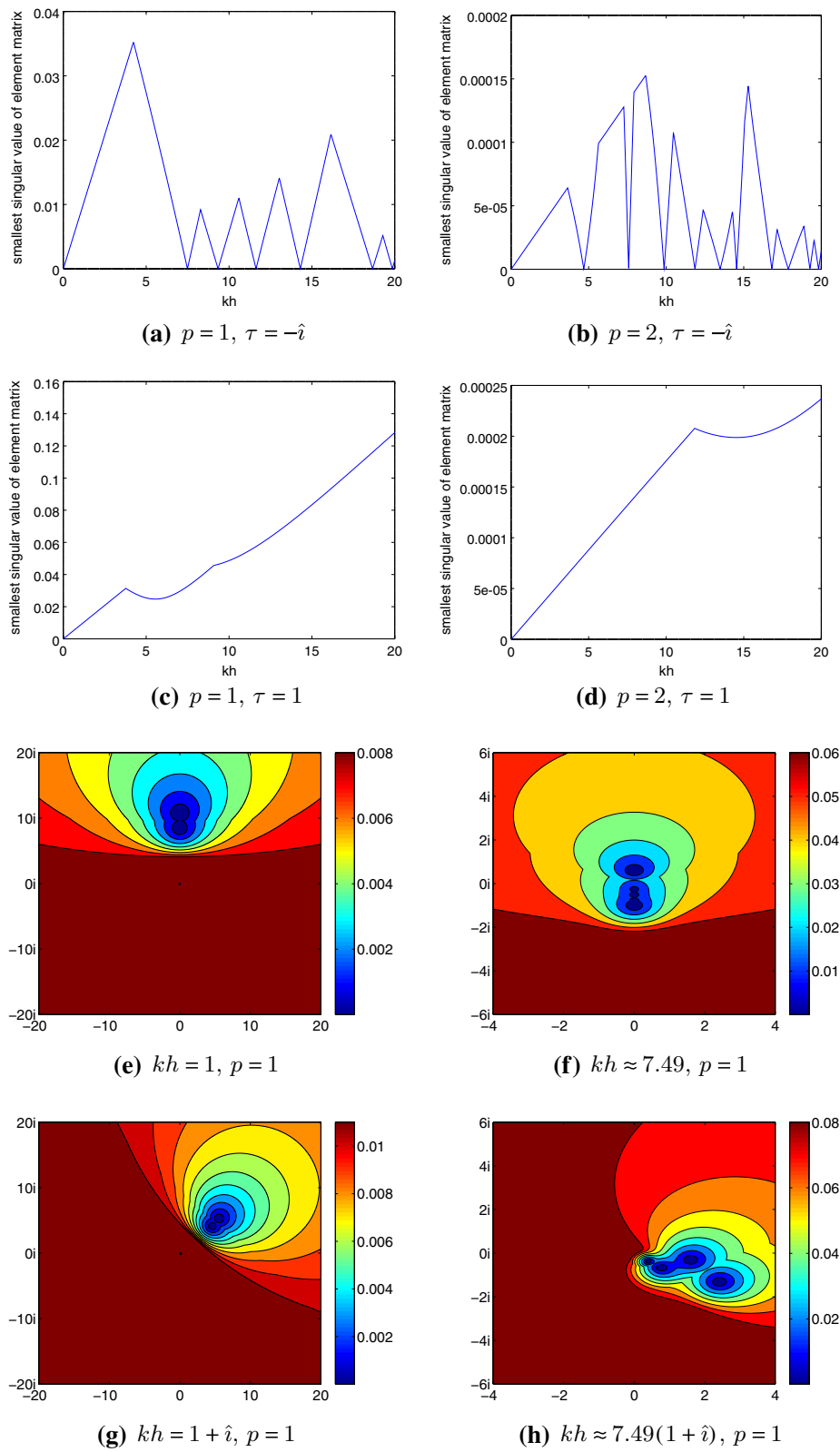
Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 9 of 24

(a) $p = 1$, $\tau = -\hat{\imath}$

(b) $p = 2$, $\tau = -\hat{\imath}$

(c) $p = 1$, $\tau = 1$

(d) $p = 2$, $\tau = 1$

(e) $kh = 1$, $p = 1$

(f) $kh \approx 7.49$, $p = 1$

(g) $kh = 1 + \hat{\imath}$, $p = 1$

(h) $kh \approx 7.49(1 + \hat{\imath})$, $p = 1$

Figure 1 The smallest singular values of a tetrahedral HDG element matrix.

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 10 of 24

nonsingular. These results are based on a tenuous stability inherited from the fact nonzero polynomials are never waves, stated precisely in the ensuing lemma. Then we give the condition on τ that guarantees unisolvency, and before concluding the section, present some caveats on relying solely on this tenuous stability.

As is standard in all HDG methods, the unique solvability of the element problem allows the formulation of a condensed global problem that involves only the interface unknowns. We introduce the following notation to describe the condensed systems. First, for Maxwell's equations, for any $\eta \in N_h$, let $(\vec{E}^\eta, \vec{H}^\eta) \in Y_h \times Y_h$ denote the fields such that, for each $K \in \mathcal{T}_h$, the pair $(\vec{E}^\eta|_K, \vec{H}^\eta|_K)$ satisfies the local problem (11a, 11b) with data $\eta|_{\partial K}$. That is,

$$\hat{\imath}k(\vec{E}^\eta, \vec{v})_K - (\vec{\nabla} \times \vec{H}^\eta, \vec{v})_K + \langle \tau \vec{E}^\eta \times \vec{n},\, \vec{v} \times \vec{n} \rangle_{\partial K} = \langle \tau \eta \times \vec{n},\, \vec{v} \times \vec{n} \rangle_{\partial K}, \tag{15a}$$

$$-(\vec{E}^\eta, \vec{\nabla} \times \vec{w})_K - \hat{\imath}k(\vec{H}^\eta, \vec{w})_K = -\langle \eta, \vec{n} \times \vec{w} \rangle_{\partial K}, \tag{15b}$$

for all $\vec{v} \in Y(K)$, $\vec{w} \in Y(K)$. If all the sources in (10a, 10b, 10c) vanish, then the *condensed global problem* for $\hat{E} \in N_h$ takes the form

$$a(\hat{E}, \eta) = 0, \quad \forall \eta \in N_h, \tag{16}$$

where

$$a(\Lambda, \eta) = \sum_{K \in \mathcal{T}_h} \langle \hat{H}^\Lambda \times \vec{n}, \eta \rangle_{\partial K}.$$

By following a standard procedure [1] we can express $a(\cdot, \cdot)$ explicitly as follows:

$$\begin{aligned}
a(\Lambda, \eta) &= \sum_{K \in \mathcal{T}_h} \langle \vec{H}^\Lambda \times \vec{n}, \eta \rangle_{\partial K} + \langle (\hat{H}^\Lambda - \vec{H}^\Lambda) \times \vec{n}, \eta \rangle_{\partial K} \\
&= \sum_{K \in \mathcal{T}_h} \hat{\imath}\bar{k}(\vec{H}^\Lambda, \vec{H}^\eta)_K - (\vec{\nabla} \times \vec{H}^\Lambda, \vec{E}^\eta)_K + \langle \tau \vec{n} \times (\vec{n} \times (\Lambda - \vec{E}^\Lambda)), \eta \rangle_{\partial K} \\
&= \sum_{K \in \mathcal{T}_h} \hat{\imath}\bar{k}(\vec{H}^\Lambda, \vec{H}^\eta)_K - \hat{\imath}k(\vec{E}^\Lambda, \vec{E}^\eta)_K + \langle \tau \vec{n} \times (\Lambda - \vec{E}^\Lambda), \vec{n} \times \vec{E}^\eta \rangle_{\partial K} \\
&\quad - \langle \tau \vec{n} \times (\Lambda - \vec{E}^\Lambda), \vec{n} \times \eta \rangle_{\partial K} \\
&= \sum_{K \in \mathcal{T}_h} \hat{\imath}\bar{k}(\vec{H}^\Lambda, \vec{H}^\eta)_K - \hat{\imath}k(\vec{E}^\Lambda, \vec{E}^\eta)_K - \tau \langle \vec{n} \times (\Lambda - \vec{E}^\Lambda), \vec{n} \times (\eta - \vec{E}^\eta) \rangle_{\partial K}.
\end{aligned}$$

Here we have used the complex conjugate of (15b) with $\vec{w} = \vec{H}^\Lambda$, along with the definition of $\hat{H}^\Lambda$, and then used (15a).

Similarly, for the Helmholtz equation, let $(\vec{u}^\eta, \phi^\eta) \in V_h \times W_h$ denote the fields such that, for all $K \in \mathcal{T}_h$, the functions $(\vec{u}^\eta|_K, \phi^\eta|_K)$ solve the element problem (4a, 4b) for given data $\hat{\phi} = \eta$. If the sources in (2a, 2b, 2c) vanish, then the *condensed global problem* for $\hat{\phi} \in M_h$ is written as

$$b(\hat{\phi}, \eta) = 0, \quad \forall \eta \in M_h, \tag{17}$$

where the form is found, as before, by the standard procedure:

$$b(\Lambda, \eta) = \sum_{K \in \mathcal{T}_h} \langle \hat{u}^\Lambda \cdot \vec{n}, \eta \rangle_{\partial K}$$
$$= \sum_{K \in \mathcal{T}_h} \hat{i}\bar{k}(\vec{u}^\Lambda, \vec{u}^\eta)_K - \hat{i}k(\phi^\Lambda, \phi^\eta)_K - \tau\langle \Lambda - \phi^\Lambda, \eta - \phi^\eta \rangle_{\partial K}.$$

The sesquilinear forms $a(\cdot, \cdot)$ and $b(\cdot, \cdot)$ are used in the main result, which gives sufficient conditions for the solvability of the local problems (11a, 11b), (4a, 4b) and the global problems (16), (17).

Before proceeding to the main result, we give a simple lemma, which roughly speaking, says that nontrivial harmonic *waves are not polynomials.*

Lemma 1   *Let $p \geq 0$ be an integer, $0 \neq k \in \mathbb{C}$, and D an open set. Then, there is no nontrivial $\vec{E} \in (\mathcal{P}_p(D))^3$ satisfying*

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) - k^2\vec{E} = 0$$

and there is no nontrivial $\phi \in \mathcal{P}_p(D)$ satisfying

$$\Delta\phi + k^2\phi = 0.$$

Proof   We use a contradiction argument. If $E \not\equiv \vec{0}$, then we may assume without loss of generality that at least one of the components of $\vec{E}$ is a polynomial of degree exactly p. But this contradicts $k^2\vec{E} = \vec{\nabla} \times (\vec{\nabla} \times \vec{E})$ because all components of $\vec{\nabla} \times (\vec{\nabla} \times \vec{E})$ are polynomials of degree at most $p - 2$. Hence $\vec{E} \equiv \vec{0}$. An analogous argument can be used for the Helmholtz case as well.   □

Theorem 1   *Suppose*

$$\text{Re}(\tau) \neq 0, \quad \text{whenever} \quad \text{Im}(k) = 0, \text{ and} \tag{18a}$$

$$\text{Im}(k)\text{Re}(\tau) \leq 0, \quad \text{whenever} \quad \text{Im}(k) \neq 0. \tag{18b}$$

Then, in the Maxwell case, the local element problem (11a, 11b) and the condensed global problem (16) are both unisolvent. Under the same condition, in the Helmholtz case, the local element problem (4a, 4b) and the condensed global problem (17) are also unisolvent.

Proof   We first prove the theorem for the local problem for Maxwell's equations. Assume (18a, 18b) holds and set $\hat{E} = \vec{0}$ in the local problem (11a, 11b). Unisolvency will follow by showing that $\vec{E}$ and $\vec{H}$ must equal $\vec{0}$. Choosing $\vec{v} = \vec{E}$, and $\vec{w} = \vec{H}$, then subtracting (11b) from (11a), we get

$$\hat{i}k\left(||\vec{E}||_K^2 + ||\vec{H}||_K^2\right) + 2\hat{i}\text{Im}(\vec{E}, \vec{\nabla} \times \vec{H})_K + \tau||\vec{E} \times \vec{n}||_{\partial K}^2 = 0,$$

whose real part is

$$-\text{Im}(k)\left(||\vec{E}||_K^2 + ||\vec{H}||_K^2\right) + \text{Re}(\tau)||\vec{E} \times \vec{n}||_{\partial K}^2 = 0.$$

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 12 of 24

Under condition (18b), we immediately have that the fields $\vec{E}$ and $\vec{H}$ are zero on K. Otherwise, (18a) implies $\vec{E} \times \vec{n}|_{\partial K} = 0$, and then (11a, 11b) gives

$$\hat{i}k\vec{E} - \vec{\nabla} \times \vec{H} = 0,$$
$$\hat{i}k\vec{H} + \vec{\nabla} \times \vec{E} = 0,$$

implying

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = k^2 \vec{E}.$$

By Lemma 1 this equation has no nontrivial solutions in the space $Y(K)$. Thus, the element problem for Maxwell's equations is unisolvent.

We prove that the global problem for Maxwell's equations is unisolvent by showing that $\hat{E} = \vec{0}$ is the unique solution of Eq. (16). This is done in a manner almost identical to what was done above for the local problem: First, set $\eta = \hat{E}$ in Eq. (16) and take the real part to get

$$\sum_{K \in \mathcal{T}_h} \mathrm{Im}(k) \left(||\vec{H}||_K^2 + ||\vec{E}||_K^2 \right) - \mathrm{Re}(\tau) ||\vec{n} \times (\hat{E} - \vec{E})||_{\partial K}^2 = 0. \tag{19}$$

This immediately shows that if condition (18b) holds, then the fields $\vec{E}$ and $\vec{H}$ are zero on $\Omega \subset \mathbb{R}^3$ and the proof is finished. In the case of condition (18a), we have $\vec{n} \times (\hat{E} - \vec{E}|_{\partial K}) = \vec{0}$ for all K. Using Eqs. (10a, 10b, 10c), this yields
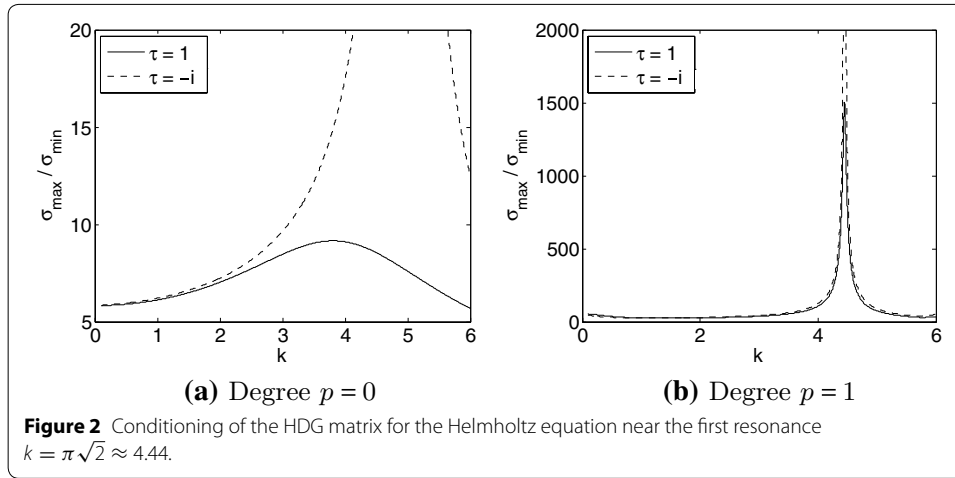
$$|_K = k^2 \vec{E}|_K,$$

so Lemma 1 proves that the fields on element interiors are zero, which in turn implies $\hat{E} = \vec{0}$ also. Thus, the theorem holds for the Maxwell case.

The proof for the Helmholtz case is entirely analogous. □

Note that even with Dirichlet boundary conditions and real k, the theorem asserts the existence of a unique solution for the Helmholtz equation. However, the exact Helmholtz problem (1a, 1b, 1c) is well-known to be *not* uniquely solvable when k is set to one of an infinite sequence of real resonance values. The fact that the discrete system is uniquely solvable even when the exact system is not, suggests the presence of artificial dissipation in HDG methods. We will investigate this issue more thoroughly in the next section.

However, we do not advocate relying on this discrete unisolvency near a resonance where the original boundary value problem is not uniquely solvable. The discrete matrix, although invertible, can be poorly conditioned near these resonances. Consider, for example, the Helmholtz equation on the unit square with Dirichlet boundary conditions. The first resonance occurs at $k = \pi\sqrt{2}$. In Figure 2, we plot the condition number $\sigma_{\max}/\sigma_{\min}$ of the condensed HDG matrix for a range of wavenumbers near the resonance $k = \pi\sqrt{2}$, using a small fixed mesh of mesh size $h = 1/4$, and a value of $\tau = 1$ that satisfies (18a, 18b). We observe that although the condition number remains finite, as predicted by the theorem, it peaks near the resonance for both the $p = 0$ and the $p = 1$ cases. We also observe that a parameter setting of $\tau = -\hat{i}$ that does not satisfy the

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 13 of 24

Figure 2 Conditioning of the HDG matrix for the Helmholtz equation near the first resonance $k = \pi\sqrt{2} \approx 4.44$.

conditions of the theorem produce much larger condition numbers, e.g., the condition numbers that are orders of magnitude greater than 10^{10} (off axis limits of Figure 2b) for k near the resonance were obtained for $p = 1$ and $\tau = -\hat{\imath}$. To summarize the caveat, *even though the condition number is always bounded for values of τ that satisfy* (18a, 18b), *it may still be practically infeasible to solve a source problem near a resonance by the HDG method.* (Of course, the eigenvalue problem can be discretized by the HDG method [13] and used to approximate the resonant eigenfunctions.)

Results of dispersion analysis for real wavenumbers

When the wavenumber k is *complex*, we have seen that it is important to choose the stabilization parameter τ such that (18b) holds. We have also seen that when k is real, the stability obtained by (18a) is so tenuous that it is of negligible practical value. For real wavenumbers, it is safer to rely on stability of the (un-discretized) boundary value problem, rather than the stability obtained by a choice of τ.

The focus of this section is on *real k* and the Helmholtz equation (1a, 1b, 1c). In this case, having already separated the issue of stability from the choice of τ, we are now free to optimize the choice of τ for other goals. By means of a dispersion analysis, we now proceed to show that some values of τ are better than others for minimizing discrepancies in wavespeed. Since dispersion analyses are limited to the study of propagation of plane waves (that solve the Helmholtz equation), we will not explicitly consider the Maxwell HDG system in this section. However, since we have written the Helmholtz and Maxwell system consistently with respect to the stabilization parameter [see the transition from (3) to (9) via (7)], we anticipate our results for the 2D Helmholtz case to be useful for the Maxwell case also.

The dispersion relation in the one-dimensional case

Consider the HDG method (2a, 2b, 2c) in the lowest order ($p = 0$) case in one dimension (1D)—after appropriately interpreting the boundary terms in (2a, 2b, 2c). We follow the techniques of [10] for performing a dispersion analysis. Using a basis on a segment of size

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 14 of 24

h in this order $u_1 = 1$, $\quad \phi_1 = 1$, $\quad \hat{\phi}_1 = 1$, $\quad \hat{\phi}_2 = 1$, the HDG element matrix takes the form $M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$ where

$$M_{11} = \begin{bmatrix} \hat{\imath}kh & 0 \\ 0 & -\hat{\imath}kh - 2\tau \end{bmatrix} \quad M_{12} = \begin{bmatrix} -1 & +1 \\ \tau & \tau \end{bmatrix}$$

$$M_{21} = M_{12}^t \qquad\qquad M_{22} = \begin{bmatrix} -\tau & 0 \\ 0 & -\tau \end{bmatrix}.$$

The Schur complement for the two endpoint basis functions $\{\hat{\phi}_1, \hat{\phi}_2\}$ is then

$$S = - \begin{bmatrix} \dfrac{1}{\hat{\imath}kh} - \dfrac{\tau^2}{\hat{\imath}kh + 2\tau} + \tau & -\dfrac{1}{\hat{\imath}kh} - \dfrac{\tau^2}{\hat{\imath}kh + 2\tau} \\ -\dfrac{1}{\hat{\imath}kh} - \dfrac{\tau^2}{\hat{\imath}kh + 2\tau} & \dfrac{1}{\hat{\imath}kh} - \dfrac{\tau^2}{\hat{\imath}kh + 2\tau} + \tau \end{bmatrix}.$$

Applying this matrix on an infinite uniform grid (of elements of size h), we obtain the stencil at an arbitrary point. If $\hat{\psi}_j$ denotes the solution (trace) value at the jth point ($j \in \mathbb{Z}$), then the jth equation reads

$$2\hat{\psi}_j \left(\frac{1}{\hat{\imath}kh} - \frac{\tau^2}{\hat{\imath}kh + 2\tau} + \tau \right) + (\hat{\psi}_{j-1} + \hat{\psi}_{j+1}) \left(-\frac{1}{\hat{\imath}kh} - \frac{\tau^2}{\hat{\imath}kh + 2\tau} \right) = 0.$$

In a dispersion analysis, we are interested in how this equation propagates plane waves on the infinite uniform grid. Hence, substituting $\hat{\psi}_j = \exp(\hat{\imath}k^h jh)$, we get the following dispersion relation for the unknown discrete wavenumber k^h:

$$\cos(k^h h) \left(\frac{1}{\hat{\imath}kh} + \frac{\tau^2}{\hat{\imath}kh + 2\tau} \right) = \left(\frac{1}{\hat{\imath}kh} - \frac{\tau^2}{\hat{\imath}kh + 2\tau} + \tau \right)$$

Simplifying,

$$k^h h = \cos^{-1} \left(1 - \frac{(kh)^2}{2 + \hat{\imath}kh(\tau + \tau^{-1})} \right). \tag{20}$$

This is the *dispersion relation* for the HDG method in the lowest order case in one dimension. Even when τ and k are real, the argument of the arccosine is not. Hence

$$\operatorname{Im}(k^h) \neq 0, \tag{21}$$

in general, indicating the *presence of artificial dissipation in HDG methods.* Note however that if τ is purely imaginary and kh is sufficiently small, (20) implies that $\operatorname{Im}(k^h) = 0$.

Let us now study the case of small kh (i.e., large number of elements per wavelength). As $kh \to 0$, using the approximation $\cos^{-1}(1 - x^2/2) \approx x + x^3/24 + \cdots$ valid for small x, and simplifying (20), we obtain

$$k^h h - kh = -\frac{(\tau^2 + 1)\hat{\imath}}{4\tau}(kh)^2 + O((kh)^3). \tag{22}$$

Comparing this with the discrete dispersion relation of the standard finite element method in one space dimension (see [10]), namely $k^h h - kh \approx O((kh)^3)$, we find that wavespeed discrepancies from the HDG method can be larger depending on the value

of τ. In particular, we conclude that *if we choose* $\tau = \pm \hat{\imath}$, *then the error* $k^h h - kh$ *in both methods are of the same order* $O((kh)^3)$.

Before concluding this discussion of the one-dimensional case, we note an alternate form of the dispersion relation suitable for comparison with later formulas. Using the half-angle formula, Eq. (20) can be rewritten as
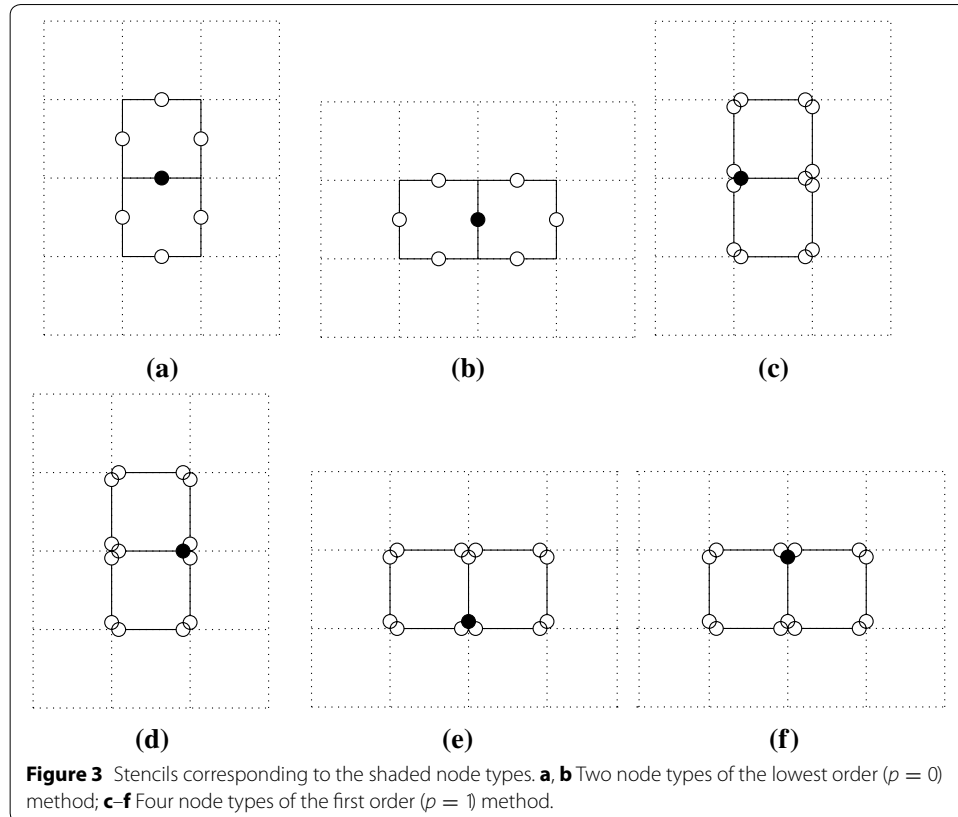
$$c^2 = 1 - \left(\frac{(kh)^2}{2} \right) \left(\frac{\tau}{\hat{\imath} kh(\tau^2 + 1) + 2\tau} \right), \tag{23}$$

where $c = \cos(k^h h/2)$.

Lowest order two-dimensional case

In the 2D case, we use an infinite grid of square elements of side length h. The HDG element matrix associated to the lowest order ($p = 0$) case of (2a, 2b, 2c) is now larger, but the Schur complement obtained after condensing out all interior degrees of freedom is only 4×4 because there is one degree of freedom per edge. Note that horizontal and vertical edges represent two distinct types of degrees of freedom, as shown in Figure 3a, b. Hence there are two types of stencils.

For conducting dispersion analysis with multiple stencils, we follow the approach in Ref. [11] (described more generally in the next subsection). Accordingly, let C_1 and C_2 denote the infinite set of stencil centers for the two types of stencils present in our case. Then, we get an infinite system of equations for the unknown solution (numerical trace)

Figure 3 Stencils corresponding to the shaded node types. **a**, **b** Two node types of the lowest order ($p = 0$) method; **c**–**f** Four node types of the first order ($p = 1$) method.

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 16 of 24

values $\hat{\psi}_{1,\vec{p}_1}$ and $\hat{\psi}_{2,\vec{p}_2}$ at all $\vec{p}_1 \in C_1$ and $\vec{p}_2 \in C_2$, respectively. We are interested in how this infinite system propagates plane wave solutions in every angle θ. Therefore, with the ansatz $\hat{\psi}_{j,\vec{p}_j} = a_j \exp(\hat{\imath}\vec{\kappa}_h \cdot \vec{p}_j)$ for constants a_j ($j = 1$ and 2), where the discrete wave vector is given by

$$\vec{\kappa}_h = k^h \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

we proceed to find the relation between the discrete wavenumber k^h and the exact wavenumber k.

Substituting the ansatz into the infinite system of equations and simplifying, we obtain a 2×2 system $F \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = 0$ where

$$F = \begin{bmatrix} 2\,kh\tau^2 c_1\,c_2 & d_1\left(4\,\tau + \hat{\imath}kh\right) + 2\,kh\tau^2 c_1{}^2 \\ d_2\left(4\,\tau + \hat{\imath}kh\right) + 2\,kh\tau^2 c_2{}^2 & 2\,kh\tau^2 c_1 c_2 \end{bmatrix}$$

and, for $j = 1, 2$,

$$c_j = \cos\left(\frac{1}{2}hk_j^h\right), \quad d_j = 2\hat{\imath}\left(1 - c_j^2\right) - \tau kh, \quad k_1^h = k^h\cos\theta, \quad k_2^h = k^h\sin\theta. \quad (24)$$

Hence the *2D dispersion relation* relating k^h to k in the HDG method is

$$\det(F) = 0. \tag{25}$$

To formally compare this to the 1D dispersion relation, consider these two sufficient conditions for $\det(F) = 0$ to hold:

$$2(kh)^2\tau^2 c_j^2 + d_j\left(2\tau kh + \hat{\imath}(k_jh)^2\right) = 0, \qquad \text{for } j = 1, 2, \tag{26}$$

where $k_1 = k\cos\theta$ and $k_2 = k\sin\theta$. (Indeed, multiplying (26)$_j$ by d_{j+1} ($j = 1$) or d_{j-1} ($j = 2$) and summing over $j = 1, 2$, one obtains a multiple of $\det(F)$.) The Equations in (26) can be simplified to

$$c_j^2 = 1 - \frac{(k_jh)^2}{2\hat{\imath}}\left(\frac{kh\,\tau}{(k_jh)^2 + (kh)^2\,\tau^2 - 2\,\hat{\imath}\,kh\,\tau}\right), \qquad j = 1, 2, \tag{27}$$

which are relations that have a form similar to the 1D relation (23). Hence we use asymptotic expansions of arccosine for small kh, similar to the ones used in the 1D case, to obtain an expansion for k_j^h, for $j = 1, 2$.

The final step in the calculation is the use of the simple identity

$$k^h = \left(\left(k_1^h\right)^2 + \left(k_2^h\right)^2\right)^{1/2}. \tag{28}$$

Simplifying the above-mentioned expansions for each term on the right hand side above, we obtain

$$k^h h - kh = \frac{\hat{\imath}(\cos(4\,\theta) + 3 + 4\,\tau^2)}{16\,\tau}(kh)^2 + O((kh)^3) \tag{29}$$

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 17 of 24

as $kh \to 0$. Thus, we conclude that if we want dispersion errors to be $O((kh)^3)$, then we must choose

$$\tau = \pm \frac{1}{2} \hat{\imath} \sqrt{\cos(4\theta) + 3}, \tag{30}$$

a prescription that is not very useful in practice because it depends on the propagation angle θ. However, we can obtain a more practically useful condition by setting τ to be the constant value that best approximates $\pm \frac{1}{2} \hat{\imath} \sqrt{\cos(4\theta) + 3}$ for all $0 \le \theta \le \pi/2$, namely

$$\tau = \pm \hat{\imath} \frac{\sqrt{3}}{2}. \tag{31}$$

These values of τ *asymptotically minimize errors in discrete wavenumber over all angles for the lowest order 2D HDG method.* Note that for any purely imaginary τ, (27) implies that k_j^h is real if kh is sufficiently small, so

$$\mathrm{Im}(k^h) = 0, \tag{32}$$

thus eliminating artificial dissipation.

We now report results of numerical computation of $k^h = k^h(\theta)$ by directly applying a nonlinear solver to the 2D dispersion relation (25) (for a set of propagation angles θ). The obtained values of the real part $\mathrm{Re}\,k^h(\theta)$ are plotted in Figure 3a, for a few fixed values of τ. The discrepancy between the exact and discrete curves quantifies the difference in the wave speeds for the computed and the exact wave. Next, analyzing the computed $k^h(\theta)$ for values of τ on a uniform grid in the complex plane, we found that the values of τ that minimize $|kh - k^h(\theta)h|$ are purely imaginary. As shown in Figure 4, these τ-values approach the asymptotic values determined analytically in Eq. (30). A second validation of our analysis is performed by considering the maximum error over all θ for each value of τ and then determining the practically optimal value of τ. The results, given in Table 2, show that the optimal τ values do approach the analytically determined value [see (31)] of $\pm \hat{\imath} \frac{\sqrt{3}}{2} \approx \pm 0.866 \hat{\imath}$. Further numerical results for the $p = 0$ case are presented together with a higher order case in the next subsection.
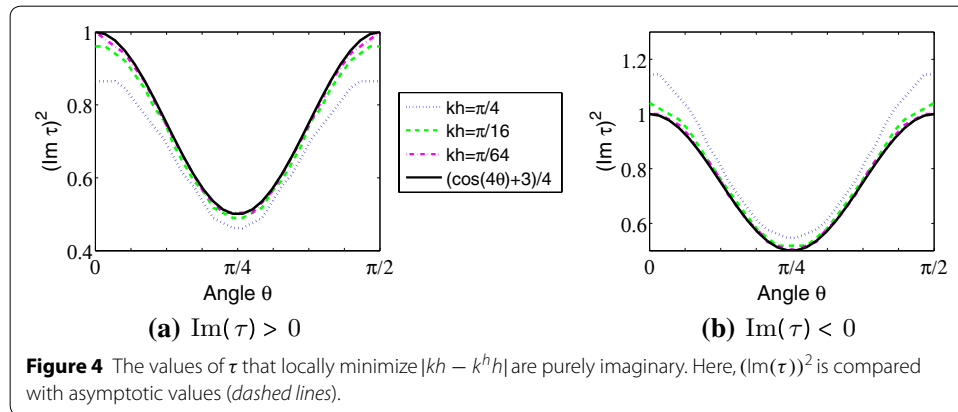

Figure 4 The values of τ that locally minimize $|kh - k^h h|$ are purely imaginary. Here, $(\mathrm{Im}(\tau))^2$ is compared with asymptotic values (*dashed lines*).

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 18 of 24

Table 2 Numerically found values of τ that minimize $|kh - k^h(\theta)h|$ for all θ in the $p = 0$ case

kh	Optimal τ, $\mathrm{Im}(\tau) > 0$	Optimal τ, $\mathrm{Im}(\tau) < 0$
$\pi/4$	$0.807\hat{\imath}$	$-0.931\hat{\imath}$
$\pi/8$	$0.837\hat{\imath}$	$-0.898\hat{\imath}$
$\pi/16$	$0.851\hat{\imath}$	$-0.882\hat{\imath}$
$\pi/32$	$0.859\hat{\imath}$	$-0.874\hat{\imath}$
$\pi/64$	$0.863\hat{\imath}$	$-0.871\hat{\imath}$
$\pi/128$	$0.865\hat{\imath}$	$-0.868\hat{\imath}$
$\pi/256$	$0.866\hat{\imath}$	$-0.867\hat{\imath}$

Higher order case

To go beyond the $p = 0$ case, we extend a technique of [11] (as in [12]). Using a higher order HDG stencil, we want to obtain an analogue of (25), which can be numerically solved for the discrete wavenumber $k^h = k^h(\theta)$. The accompanying dispersive, dissipative, and total errors are defined respectively by

$$\epsilon_{\text{disp}} = \max_\theta |\mathrm{Re}(k^h(\theta)) - k|, \qquad \epsilon_{\text{dissip}} = \max_\theta |\mathrm{Im}(k^h(\theta))|,$$
$$\epsilon_{\text{total}} = \max_\theta |k^h(\theta) - k|. \tag{33}$$

Again, we consider an infinite lattice of $h \times h$ square elements with the ansatz that the HDG degrees of freedom interpolate a plane wave traveling in the θ direction with wavenumber k^h. The lowest order and next higher order HDG stencils are compared in Figure 3. Note that the figure only shows the interactions of the degrees of freedom corresponding to the $\hat{\phi}$ variable—the only degrees of freedom involved after elimination of the $\vec{u}$ and ϕ degrees of freedom via static condensation. The lowest order method has two node types (shown in Figure 3a, b), while the first order method has four node types (shown in Figure 3c–f). For a method with S distinct node types, denote the solution value at a node of the sth type, $1 \le s \le S$, located at $\vec{l}h \in \mathbb{R}^2$, by $\psi_{s,\vec{l}}$. With our ansatz that these solution values interpolate a plane wave, we have

$$\psi_{s,\vec{l}} = a_s e^{\hat{\imath}\vec{k}^h \cdot \vec{l}h},$$

for some constants a_s.

Now, to develop notation to express each stencil's equation, we fix a stencil within the lattice. Suppose that it corresponds to a node of the tth type, $1 \le t \le S$, that is located at $\vec{j}h$. For $1 \le s \le S$, define $J_{t,s} = \{\vec{l} \in \mathbb{R}^2 : \text{a node of type } s \text{ is located at } (\vec{j} + \vec{l})h\}$ and, for $\vec{l} \in J_{t,s}$, denote the stencil coefficient of the node at location $(\vec{j} + \vec{l})h$ by $D_{t,s,\vec{l}}$. The stencil coefficient is the linear combination of the condensed local matrix entries that would likewise appear in the global matrix of Eq. (17). Both it and the set $J_{t,s}$ are translation invariant, i.e., independent of $\vec{j}$. Since plane waves are exact solutions to the Helmholtz equation with zero sources, the stencil's equation is

$$\sum_{s=1}^{S} \sum_{l \in J_{t,s}} D_{t,s,\vec{l}} \, \psi_{s,\vec{j}+\vec{l}} = 0.$$

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 19 of 24

Finally, we remove all dependence on $\vec{j}$ in this equation by dividing by $e^{i\vec{k}^h \cdot \vec{j}h}$, so there are S equations in total, with the tth equation given by

$$\sum_{s=1}^{S} a_s \sum_{l \in J_{t,s}} D_{t,s,\vec{l}} \, e^{i\vec{k}^h \cdot \vec{l}h} = 0. \tag{34}$$

Defining the $S \times S$ matrix $F(k^h)$ by

$$_{t,s} = \sum_{l \in J_{t,s}} D_{t,s,\vec{l}} \, e^{ik^h[\cos\theta, \sin\theta] \cdot \vec{l}h},$$

we observe that non-trivial coefficients $\{a_s\}$ exist if and only if k^h is such that
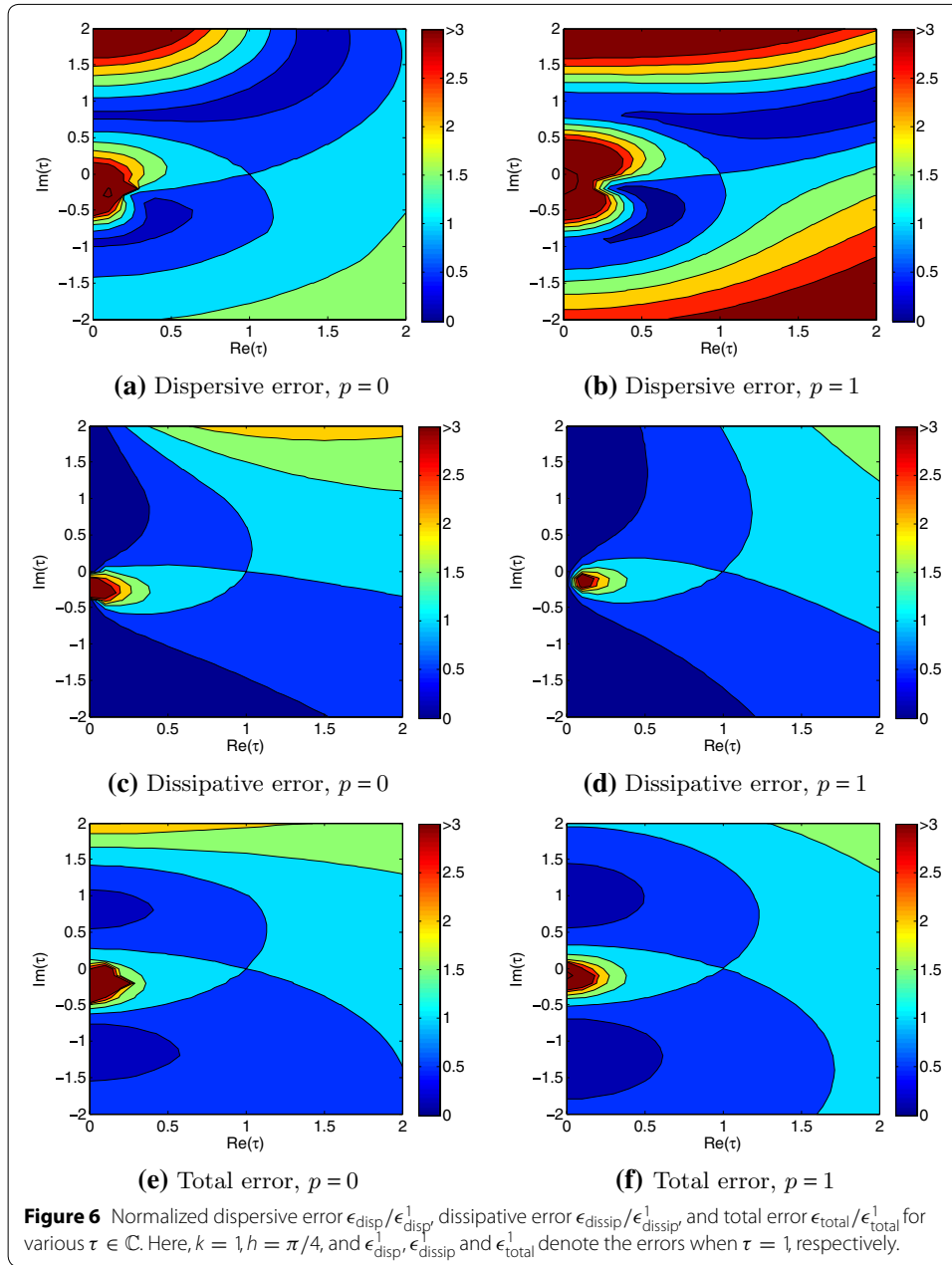
$$\det(F(k^h)) = 0. \tag{35}$$

This is the equation that we solve to determine k^h for a given θ for any order (cf. 25).

Results of the dispersion analysis are shown in Figures 5 and 6. These figures combine the results from previously discussed $p = 0$ case and the $p = 1$ cases to facilitate comparison. Here, we set $k = 1$ and $h = \pi/4$, i.e., 8 elements per wavelength. Figure 6 shows the dispersive, dissipative, and total errors for various values of $\tau \in \mathbb{C}$. For both the lowest order and first order cases, although the dispersive error is minimized at a value of τ having nonzero real part, the total error is minimized at a purely imaginary value of τ. This is attributed to the small dissipative errors for such τ. Specifically, the total error is minimized when $\tau = 0.87\hat{i}$ in the $p = 1$ case. This is close to the optimal value of τ found (both analytically and numerically) for $p = 0$. This value of τ reduces the total wavenumber error by 90% in the $p = 1$ case, relative to the total error when using $\tau = 1$.

Comparison with dispersion relation for the Hybrid Raviart–Thomas method

The HRT (Hybrid Raviart–Thomas) method is a classical mixed method [14–16] which has a similar stencil pattern, but uses different spaces. Namely, the HRT method for the Helmholtz equation is defined by exactly the same equations as (2a, 2b, 2c) with τ set to zero, but with these choices of spaces on square elements: $V(K) = \mathcal{Q}_{p+1,p}(K) \times \mathcal{Q}_{p,p+1}(K)$, $W(K) = \mathcal{Q}_p(K)$, and $M(F) = \mathcal{P}_p(F)$. Here $\mathcal{Q}_{l,m}(K)$ denotes the space of polynomials

Figure 5 Numerical wave speed $\mathrm{Re}(\vec{k}^h(\theta))$ as a function of θ for various choices of τ. Here, $k = 1$ and $h = \pi/4$.

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 20 of 24

Figure 6 Normalized dispersive error $\epsilon_{disp}/\epsilon_{disp}^1$, dissipative error $\epsilon_{dissip}/\epsilon_{dissip}^1$, and total error $\epsilon_{total}/\epsilon_{total}^1$ for various $\tau \in \mathbb{C}$. Here, $k = 1$, $h = \pi/4$, and ϵ_{disp}^1, ϵ_{dissip}^1 and ϵ_{total}^1 denote the errors when $\tau = 1$, respectively.

which are of degree at most l in the first coordinate and of degree at most m in the second coordinate. The general method of dispersion analysis described in the previous subsection can be applied for the HRT method. We proceed to describe our new findings, which in the lowest order case includes an exact dispersion relation for the HRT method.

In the $p = 0$ case, after statically condensing the element matrices and following the procedure leading to (25), we find that the discrete wavenumber k^h for the HRT method satisfies the 2D dispersion relation

$$\left(c_1^2 + c_2^2\right)\left(2(hk)^2 - 12\right) + c_1^2 c_2^2\left(4(hk)^2 + 48\right) + (hk)^2 - 24 = 0, \tag{36}$$

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 21 of 24

where c_j, as defined in (24), depends on k_j^h, which in turn depends on k^h. Similar to the HDG case, we now observe that the two equations

$$\left(2(hk_j)^2 + 12\right)c_j^2 + (hk_j)^2 - 12 = 0, \qquad j = 1, 2, \tag{37}$$

are sufficient conditions for (36) to hold. Indeed, if l_j is the left hand side above, then $l_1(2c_2^2 + 1) + l_2(2c_1^2 + 1)$ equals the left hand side of (36). The equations of (37) can immediately be solved:

$$hk_j^h = 2\cos^{-1}\left(\frac{12 - (hk_j)^2}{2\,(hk_j)^2 + 12}\right)^{1/2}$$

Hence, using (28) and simplifying using the same type of asymptotic expansions as the ones we previously used, we obtain

$$k^h h - kh = -\frac{\cos(4\,\theta) + 3}{96}\,(kh)^3 + O((kh)^5) \tag{38}$$

as $kh \to 0$. Comparing with (29), we find that in the lowest order case, *the HRT method has an error in wavenumber that is asymptotically one order smaller than the HDG method* for any propagation angle, irrespective of the value of τ.

To conclude this discussion, we report the results from numerically solving the non-linear solution (36) for $k^h(\theta)$ for an equidistributed set of propagation angles θ. We have also calculated the analogue of (36) for the $p = 1$ case (following the procedure described in the previous subsection). Recall the dispersive, dissipative, and total errors in the wavenumbers, as defined in Eq. (33). After scaling by the mesh size h, these errors for both the HDG and the HRT methods are graphed in Figure 7 for $p = 0$ and $p = 1$. We find that the dispersive errors decrease at the same order for the HRT method and the HDG method with $\tau = 1$. While (38) suggests that the dissipative errors for the HRT method should be of higher order, our numerical results found them to be zero (up to machine accuracy). The dissipative errors also quickly fell to machine zero for the HDG method with the previously discussed "best" value of $\tau = \hat{\imath}\sqrt{3}/2$, as seen from Figure 7.

Conclusions

These are the findings in this paper:

1. There are values of stabilization parameters τ that will cause the HDG method to fail in time-harmonic electromagnetic and acoustic simulations using complex wavenumbers. [See Eq. (5) et seq.]
2. If the wavenumber k is complex, then choosing τ so that $\mathrm{Re}(\tau)\mathrm{Im}(k) \le 0$ guarantees that the HDG method is uniquely solvable. (See Theorem 1.)
3. If the wavenumber k is real, then even when the exact wave problem is not well-posed (such as at a resonance), the HDG method remains uniquely solvable when $\mathrm{Re}(\tau) \ne 0$. However, in such cases, we found the discrete stability to be tenuous. (See Figure 2 and accompanying discussion.)

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 22 of 24

(a) Dispersive error, $p = 0$

(b) Dispersive error, $p = 1$

(c) Dissipative error, $p = 0$

(d) Dissipative error, $p = 1$

(e) Total error, $p = 0$

(f) Total error, $p = 1$

HDG with $\tau = 1$

HDG with $\tau = 0.866\,i$

HRT

Order $(kh)^2$ reference

Order $(kh)^3$ reference

Order $(kh)^4$ reference

Order $(kh)^5$ reference

Figure 7 Convergence rates as $kh \to 0$.
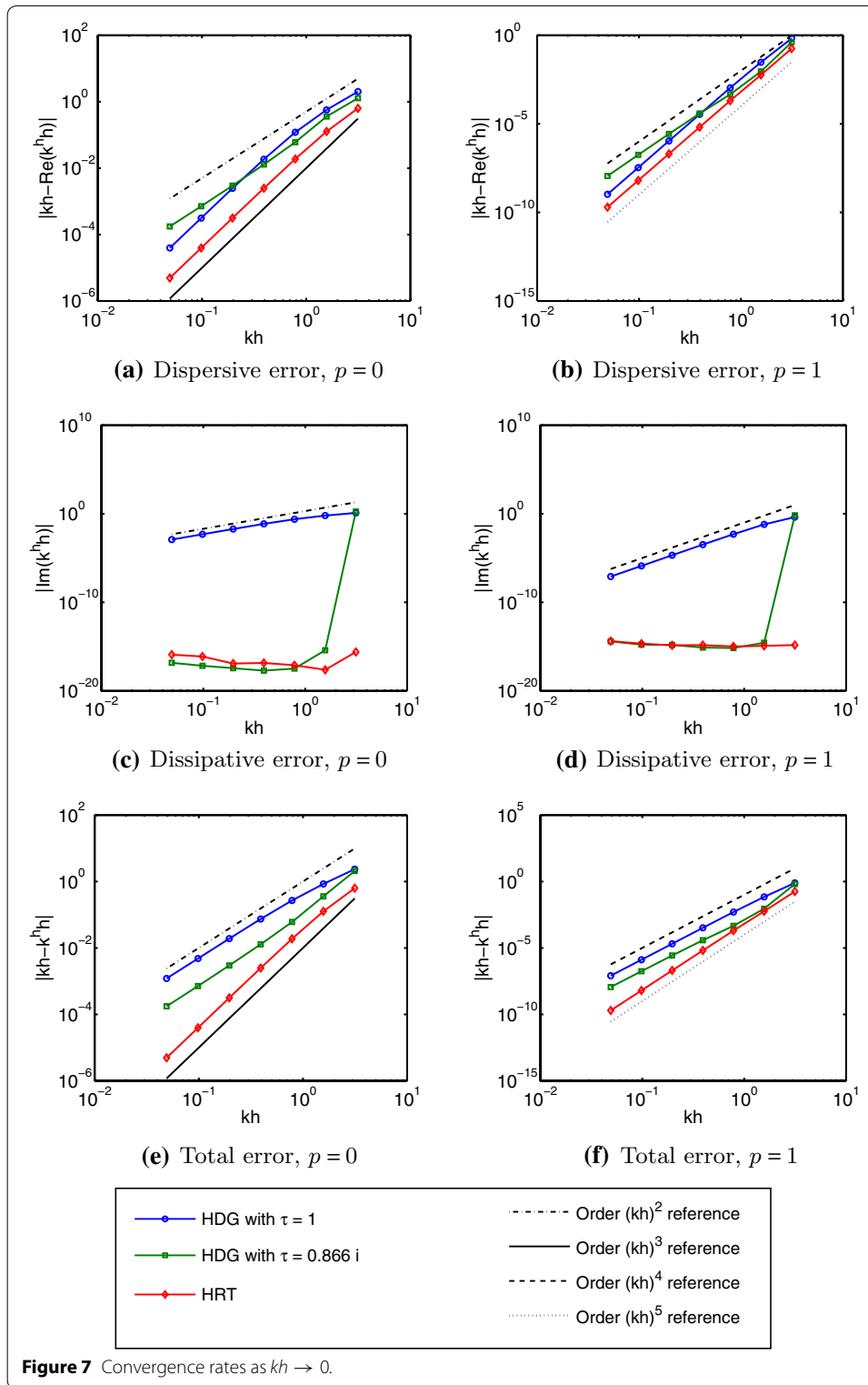
4. For real wavenumbers k, we found that the HDG method introduces small amounts of artificial dissipation [see Eq. (21)] in general. The artificial dissipation is eliminated [see Eq. (32)] when $\mathrm{Re}(\tau) = 0$ and kh is sufficiently small, but note that in this case,

Theorem 1 no longer guarantees unique solvability. In 1D, the optimal values of τ that asymptotically minimize the total error in the wavenumber (that quantifies dissipative and dispersive errors together) are $\tau = \pm\hat{\imath}$ [see Eq. (22)].

5. In 2D, for real wavenumbers k, the best values of τ are dependent on the propagation angle. Overall, values of τ that asymptotically minimize the error in the discrete wavenumber (considering all angles) is $\tau = \pm\hat{\imath}\sqrt{3}/2$ [per Eq. (31)]. While dispersive errors dominate the total error for $\tau = \hat{\imath}\sqrt{3}/2$, dissipative errors dominate when $\tau = 1$ (see Figure 7).

6. The HRT method, in both the numerical results and the theoretical asymptotic expansions, gave a total error in the discrete wavenumber that is asymptotically one order smaller than the HDG method. [See (38) and Figure 7.]

Abbreviations

HDG: hybrid (or hybridized) discontinuous Galerkin; 1D: one dimension(al); 2D: two dimension(al); 3D: three dimension(al); HRT: hybrid (or hybridized) Raviart–Thomas.

Authors' contributions

The contributions of all authors are equal. All authors read and approved the final manuscript.

Author details

[1] Portland State University, PO Box 751, Portland, OR 97207-0751, USA. [2] INRIA Sophia Antipolis Méditerranée, 2004 Route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex, France. [3] LAPLACE (LAboratoire PLasma et Conversion d'Energie), Université de Toulouse, CNRS/INPT/UPS, Toulouse, France.

Compliance with ethical guidelines

Competing interests

The authors declare that they have no competing interests.

References

1. Cockburn B, Gopalakrishnan J, Lazarov R (2009) Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. SIAM J Numer Anal 47(2):1319–1365. doi:10.1137/070706616
2. Cui J, Zhang W (2014) An analysis of HDG methods for the Helmholtz equation. IMA J Numer Anal 34(1):279–295. doi:10.1093/imanum/drt005
3. Giorgiani G, Fernández-Méndez S, Huerta A (2013) Hybridizable discontinuous Galerkin p-adaptivity for wave propagation problems. Int J Numer Methods Fluids 72:1244–1262. doi:10.1002/fld.3784
4. Griesmaier R, Monk P (2011) Error analysis for a hybridizable discontinuous Galerkin method for the Helmholtz equation. J Sci Comput 49(3):291–310. doi:10.1007/s10915-011-9460-z
5. Huerta A, Roca X, Aleksandar A, Peraire J (2012) Are high-order and hybridizable discontinuous Galerkin methods competitive? In: Oberwolfach Reports. Abstracts from the workshop held February 12–18, 2012, organized by Olivier Allix, Carsten Carstensen, Jörg Schröder and Peter Wriggers, vol 9. Oberwolfach, Blackforest, Germany, pp. 485–487. doi:10.4171/OWR/2012/09
6. Li L, Lanteri S, Perrussel R (2013) Numerical investigation of a high order hybridizable discontinuous Galerkin method for 2d time-harmonic Maxwell's equations. COMPEL 32(3):1112–1138. doi:10.1108/03321641311306196
7. Li L, Lanteri S, Perrussel R (2014) A hybridizable discontinuous Galerkin method combined to a Schwarz algorithm for the solution of 3d time-harmonic Maxwell's equation. J Comput Phys 256:563–581. doi:10.1016/j.jcp.2013.09.003
8. Nguyen NC, Peraire J, Cockburn B (2011) Hybridizable discontinuous Galerkin methods for the time-harmonic Maxwell's equations. J Comput Phys 230(19):7151–7175. doi:10.1016/j.jcp.2011.05.018
9. Cockburn B, Gopalakrishnan J, Sayas F-J (2010) A projection-based error analysis of HDG methods. Math Comp 79:1351–1367
10. Ainsworth M (2004) Discrete dispersion relation for hp-version finite element approximation at high wave number. SIAM J Numer Anal 42(2):553–575. doi:10.1137/S0036142903423460

Gopalakrishnan *et al. Adv. Model. and Simul. in Eng. Sci.* (2015) 2:13

Page 24 of 24

11. Deraemaeker A, Babuška I, Bouillard P (1999) Dispersion and pollution of the FEM solution for the Helmholtz equation in one, two and three dimensions. Int J Numer Methods Eng 46:471–499
12. Gopalakrishnan J, Muga I, Olivares N (2014) Dispersive and dissipative errors in the DPG method with scaled norms for the Helmholtz equation. SIAM J Sci Comput 36(1):20–39
13. Gopalakrishnan J, Li F, Nguyen N-C, Peraire J (2014) Spectral approximations by the HDG method. Math Comp 84(293):1037–1059
14. Arnold DN, Brezzi F (1985) Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. RAIRO Modél Math Anal Numér 19(1):7–32
15. Cockburn B, Gopalakrishnan J (2004) A characterization of hybridized mixed methods for the Dirichlet problem. SIAM J Numer Anal 42(1):283–301
16. Raviart PA, Thomas JM (1977) A mixed finite element method for 2nd order elliptic problems. In: Mathematical Aspects of Finite Element Methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975). Springer, Berlin, pp 292–315606