Geophysical Journal International

Geophys. J. Int. (2014) **196,** 1123–1133
Advance Access publication 2013 December 3

doi: 10.1093/gji/ggt446

Time-stepping stability of continuous and discontinuous finite-element methods for 3-D wave propagation

W. A. Mulder,[1,2] E. Zhebel[1] and S. Minisini[1]

[1] *Shell Global Solutions International B.V., PO Box* 60, 2280 *AB Rijswijk, the Netherlands. E-mail: wim.mulder@shell.com*
[2] *Delft University of Technology, Faculty of Civil Engineering and Geosciences, PO Box* 5048, 2600 *GA Delft, the Netherlands*

Accepted 2013 November 1. Received 2013 October 29; in original form 2013 February 22

SUMMARY

We analyse the time-stepping stability for the 3-D acoustic wave equation, discretized on tetrahedral meshes. Two types of methods are considered: mass-lumped continuous finite elements and the symmetric interior-penalty discontinuous Galerkin method. Combining the spatial discretization with the leap-frog time-stepping scheme, which is second-order accurate and conditionally stable, leads to a fully explicit scheme. We provide estimates of its stability limit for simple cases, namely, the reference element with Neumann boundary conditions, its distorted version of arbitrary shape, the unit cube that can be partitioned into six tetrahedra with periodic boundary conditions and its distortions. The Courant–Friedrichs–Lewy stability limit contains an element diameter for which we considered different options. The one based on the sum of the eigenvalues of the spatial operator for the first-degree mass-lumped element gives the best results. It resembles the diameter of the inscribed sphere but is slightly easier to compute. The stability estimates show that the mass-lumped continuous and the discontinuous Galerkin finite elements of degree 2 have comparable stability conditions, whereas the mass-lumped elements of degree one and three allow for larger time steps.

Key words: Numerical solutions; Fourier analysis; Numerical approximations and analysis; Computational seismology; Wave propagation.

1 INTRODUCTION

Solving the 3-D wave equation remains a challenging problem in geophysical applications, especially for seismic imaging of the subsurface. Traditionally, the wave equation is discretized with finite-difference methods. However, their accuracy deteriorates in complex geological settings with sharp contrasts in the medium and with topography. Finite elements on tetrahedral meshes are more flexible and accurate if the mesh follows the geometry of the interfaces and of the topography.

We focus on two types of methods: the continuous mass-lumped (CML) finite elements and the symmetric interior-penalty discontinuous Galerkin (SIPDG) method, since both allow for explicit time stepping. Accurate estimates of the Courant–Friedrichs–Lewy (CFL) number are required for efficiency, in particular if local time stepping is used (Grote & Mitkova 2010; Minisini *et al.* 2013).

Because mass lumping of standard finite elements leads to a degradation of the accuracy, except for the standard linear element, polynomials of higher degree have to be added in the interior of the element. In 2-D, elements up to degree six have been constructed for triangles (Mulder 2013). For 3-D, tetrahedral elements of degree two (Mulder 1996) and three (Chin-Joe-Kong *et al.* 1999) are known. By borrowing nodes from higher-degree elements on the faces and in the interior, the mass matrix is made diagonal while preserving the accuracy of the spatial discretization.

The interior-penalty discontinuous Galerkin (IPDG) methods allow for the approximating functions to be discontinuous at the element interfaces. However, to impose continuity, additional terms called fluxes have to be included (Wheeler 1978; Darlow 1980; Arnold 1982; Oden *et al.* 1998; Rivière *et al.* 1999, 2002; Dawson *et al.* 2004; Grote *et al.* 2006; Epshteyn & Rivière 2007; Rivière 2008; Minisini *et al.* 2013). The discontinuous representation leads to a local mass matrix that can be easily inverted. Here, we will only consider the symmetric variant, SIPDG. Other versions are the incomplete (IIPDG) and non-symmetric (NIPDG) ones.

Combining the spatial discretization with the leap-frog time-integration scheme, which is second-order accurate and conditionally stable, leads to a fully explicit method. Accuracy was discussed elsewhere (Grote *et al.* 2006; Grote & Schötzau 2009; Minisini *et al.* 2012). Here, we will focus on the time-stepping stability for CML and SIPDG finite elements.

An analysis of the convergence properties of SIPDG has been presented by Epshteyn & Rivière (2007) and Grote & Schötzau (2009), while Ainsworth *et al.* (2006) and De Basabe & Sen (2010) investigated the dispersive properties and stability for 2-D problems on Cartesian meshes. The last authors also considered IIPDG and

NIPDG. Agut *et al.* (2011) carried out stability analysis of the SIPDG method on triangles.

Ideally, the stability condition is split into a problem-independent CFL-number (Courant *et al.* 1928) and properties that are related to the mesh and velocity model. We will examine several choices for the diameter of an element in an attempt to minimize the variation of the CFL-number with respect to variations in the mesh. In order to estimate the CFL-number for CML and SIPDG tetrahedral finite elements, we consider a set of simplified problems: the reference element with Neumann boundary conditions, its distorted version of arbitrary shape, the unit cube that can be partitioned into six tetrahedra with periodic boundary conditions, and its distortions. For SIPDG, we also consider the reference element and its distortions with zero values for the solution in neighbouring elements. The estimates for the CFL-number are obtained by numerical optimization for these examples. They represent a necessary but perhaps not sufficient condition for stability.

The paper is organized as follows. Section 2 reviews the finite-element discretization in space and time for the CML as well as the SIPDG finite-element method. Section 3 describes the Fourier stability analysis and suggests several choices for the element diameter. Section 4 contains the stability estimates for the reference element and the six tetrahedra with periodic boundary conditions including their distorted versions. We also consider the dependence on the penalty parameter for the SIPDG method. Section 5 summarizes the conclusions.

2 FINITE ELEMENTS DISCRETIZATION

This section reviews the discretizations of the acoustic wave equation with the CML finite elements and the IPDG methods.

The scalar wave equation for constant-density acoustics in a bounded 3-D domain, $\Omega \subset \mathbb{R}^3$ reads:

$$\begin{cases} \dfrac{1}{c^2}\dfrac{\partial^2 u}{\partial t^2} - \Delta u = s, & \text{in } \Omega, \\ u(\mathbf{x},0) = u_0, \quad \dfrac{\partial u}{\partial t}(\mathbf{x},0) = u_1, & \text{in } \Omega, \\ u = 0, & \text{on } \Gamma_{\mathrm{D}}, \\ \nabla u \cdot \mathbf{n} = 0, & \text{on } \Gamma_{\mathrm{N}}, \end{cases} \tag{1}$$

where $c(\mathbf{x})$ is the velocity, $u(\mathbf{x}, t)$ the pressure at time t and position $\mathbf{x} = (x, y, z) \in \Omega$ and $s(\mathbf{x}, t)$ the source term. The initial conditions are $u_0(\mathbf{x})$ and $u_1(\mathbf{x})$. The external boundary of the domain is denoted by Γ_e and consists of subsets Γ_{D} and Γ_{N} with Dirichlet or Neumann boundary conditions, respectively. We have $\Gamma_e = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}}$, $\Gamma_{\mathrm{D}} \cap \Gamma_{\mathrm{N}} = \emptyset$ and $\mathbf{n}$ is the unit outward vector normal to Γ_e. For simplicity, we have prescribed homogeneous boundary conditions, but the formulation can be extended to other types of boundary conditions, for instance periodic or absorbing, without major difficulties.

The weak formulation of the acoustic wave equation is obtained by multiplying eq. (1) by a test function, integrating over the domain and applying Gauss's divergence theorem. For the continuous finite elements, this leads to

$$\int_\Omega c^{-2} v \frac{\partial^2 u}{\partial t^2}\,\mathrm{d}\Omega + \int_\Omega \nabla u \cdot \nabla v\,\mathrm{d}\Omega - \int_{\Gamma_e} (\mathbf{n}\cdot\nabla u)\, v\,\mathrm{d}\Gamma = \int_\Omega v\, s\,\mathrm{d}\Omega,$$

$\forall v \in V$, with $V = \{v | v \in H^1(\Omega), v = 0 \text{ on } \Gamma_{\mathrm{D}}\}$, $t \in (0, T)$, $u : (0, T) \times \mathbb{R}^3(\Omega) \to H^1(\Omega)$. The exterior boundary of Ω is denoted by Γ_e.

Next, the domain is partitioned into tetrahedral elements E_m, such that $\Omega_h = \bigcup_m E_m$. We introduce a finite-dimensional subspace $V_h \subset V$ described by the basis $\{\phi_i\}_{i=1}^N$, N being the number of nodes in the discretization. We choose Lagrangian basis functions of polynomial degree M, such that the approximating polynomials are equal to 1 on one node of the discretization and 0 on the other nodes. To preserve the spatial accuracy after mass lumping, polynomials of higher degree M_f that vanish on the edges have to be added to the faces, increasing the number of nodes on the faces (Fried & Malkus 1975; Tordjman 1995; Cohen *et al.* 2001). Similarly, additional polynomials of higher degree M_i that vanish on faces need to be included in the interior, leading to a larger number of nodes in the interior. For details, we refer to papers by Mulder (1996) and Chin-Joe-Kong *et al.* (1999). We introduce the approximation

$$u(\mathbf{x}, t) = \sum_j u_j(t)\phi_j(\mathbf{x}), \quad v(\mathbf{x}) = \sum_i v_i \phi_i(\mathbf{x}),$$

where u_j is the value of the unknown in node j of the discretization. Substitution of this approximation for each u and v into the weak formulation and imposing stationarity in the v_i yields

$$\frac{\partial^2 \mathbf{u}}{\partial t^2} = \mathcal{M}^{-1}\mathbf{s} - \mathcal{L}\mathbf{u}, \tag{2}$$

with spatial operator $\mathcal{L} = \mathcal{M}^{-1}\mathcal{K}$, where

$$\mathcal{M}_{ij} = \int_\Omega c^{-2}\phi_i\phi_j\,\mathrm{d}\Omega, \quad \mathcal{K}_{ij} = \int_\Omega \nabla\phi_i \cdot \nabla\phi_j\,\mathrm{d}\Omega,$$

are the elements of the mass matrix and the stiffness matrix, respectively. These are assembled from their local counterparts per element that obey a similar expression. We assume that $c(\mathbf{x})$ is constant per element. The vector $\mathbf{u}$ contains the unknowns of the problem, the degrees of freedom, and $\mathbf{s}$ contains the elements of the discrete source term, $s_i = \int_\Omega s\,\phi_i\,\mathrm{d}\Omega$.

The mass matrix is replaced by a diagonal one using the mass-lumping technique. Eq. (2) is called the semi-discrete form because it has been discretized in space through the substitution of the basis functions but the second time derivative still remains.

The discontinuous Galerkin method is a generalization of the FEM that allows the approximating function to be discontinuous at the elements interfaces. The continuity is then enforced weakly by some extra terms. For the discontinuous Galerkin interior penalty scheme, the weak formulation is given by

$$\int_\Omega c^{-2} v \frac{\partial^2 u}{\partial t^2}\,\mathrm{d}\Omega + \int_\Omega \nabla u \cdot \nabla v\,\mathrm{d}\Omega - \int_\Gamma (\mathbf{n}\cdot\nabla u)\, v\,\mathrm{d}\Gamma = \int_\Omega v s\,\mathrm{d}\Omega \tag{3}$$

for all test functions v that are chosen as Lagrange polynomials up to degree M per element. Here, $\mathbf{n}$ denotes the outward normal and Γ consists of internal and external boundaries of the domain Ω. The term with the normal in (3), called flux term, is given by

$$-\int_\Gamma (\mathbf{n}\cdot\nabla u)\, v\,\mathrm{d}\Gamma = -\int_\Gamma \{\nabla u\} \cdot [v]\,\mathrm{d}\Gamma + \epsilon \int_\Gamma [u] \cdot \{\nabla v\}\,\mathrm{d}\Gamma + \gamma \int_\Gamma [u] \cdot [v]\,\mathrm{d}\Gamma. \tag{4}$$

If u^+ is the solution inside the element, E^+, and u^- lives on one of the neighbouring elements, E^-, then $[u] := u^+\mathbf{n}^+ + u^-\mathbf{n}^-$ denotes the jump across the element boundary and $\{\nabla u\} := \frac{1}{2}(\nabla u^+ + \nabla u^-)$ is the average. The same applies for the function v. The parameter ϵ assumes the values 1, 0 or -1, depending on the particular formulation of the IPDG (Arnold *et al.* 2002). In particular, $\epsilon = 1$ leads

to the non-symmetric variant (NIPDG; Rivière *et al.* 1999, 2002), $\epsilon = 0$ to the incomplete version (IIPDG; Dawson *et al.* 2004) and the symmetric variant (SIPDG) has $\epsilon = -1$ (Darlow 1980; Grote *et al.* 2006). The third integral in the expression (4) for the fluxes is introduced to enforce the coercivity of the spatial operator, which is required for existence and uniqueness of the solution; see the paper by Grote *et al.* (2006) for details. The penalty parameter γ is defined on each interior face as $\gamma = C/\eta_F$, where C is a positive parameter and η_F has different definitions in the literature. Whatever its definition, the coercivity is assured for $C \geq C_0 > 0$ (Grote *et al.* 2006; Epshteyn & Rivière 2007). Shahbazi (2005) chooses $\eta_F = \min(d_{i,E^+}, d_{i,E^-})$, where $d_{i,E}$ is the diameter of the inscribed sphere in the element E, while Grote *et al.* (2006) let $\eta_F = \min(h_{E^+}, h_{E^-})$, with h_E the diameter of the element. If the elements are squares, h is the length of the diagonal. The paper does not give its definition for triangles. Ainsworth *et al.* (2006) chooses η_F as length of the edge of a square for a 2-D problem and sets $C = \frac{1}{2}(M+1)(M+2)$, where M is the polynomial degree used for the discretization. De Basabe *et al.* (2008) use the same definition. Epshteyn & Rivière (2007) determine a lower bound for the penalty constant that guarantees coercivity. In the Appendix, we derive a sharper, element-wise estimate

$$\gamma_E \geq \gamma_{0,E} = \frac{M(M+2)}{d_{i,E}} \tag{5}$$

per element E. We will also consider the time-stepping stability limit as a function of $\alpha = \gamma_E/\gamma_{0,E}$.

The discretization of the weak formulation (3) is similar to that of the continuous finite elements. The spatial operator becomes

$$\mathcal{L} = \mathcal{M}^{-1}(\mathcal{K} + \mathcal{F}), \tag{6}$$

where $\mathcal{M}$ and $\mathcal{K}$ are the mass and stiffness matrix, respectively, and $\mathcal{F}$ takes the flux contributions into account.

To obtain a fully discrete problem, we use the standard second-order discretization for the second derivative in time, known as the leap-frog scheme. Then,

$$\mathbf{u}^{n+1} - 2\mathbf{u}^n + \mathbf{u}^{n-1} = \Delta t^2(\mathcal{M}^{-1}\mathbf{s}^n - \mathcal{L}\mathbf{u}^n), \tag{7}$$

where $\mathcal{L}$ is a spatial discretization operator. Here, $\mathbf{u}^n$ contains the degrees of freedom or unknowns on the nodes at time $t_n = t_0 + n\Delta t$ ($n = 0, 1, \ldots$).

3 METHOD

The stability analysis of the time-stepping scheme requires an estimate of the spectral radius of the spatial operator for a given problem. To get an indication of typical values, we consider either a single element or a cube filled with six tetrahedra and periodic boundary conditions. For the last, a spatial Fourier transform is used. In the analysis, the source term, $\mathbf{s}$, is set to zero and a unit velocity $c = 1$ is adopted. The corresponding spatial operator will be denoted by $\bar{\mathcal{L}}$, to distinguish it from the operator $\mathcal{L}$ with variable $c(\mathbf{x})$.

The classic van Neumann stability analysis (Charney *et al.* 1950) for finite-difference schemes employs Fourier transforms in time and space. Instead of applying a Fourier transform in time, we consider the second-order difference equation $u^{n+1} - 2u^n + u^{n-1} = -\Delta t^2 a\, u^n$ with given initial conditions u^{-1} and u^0, time step Δt and spatial operator replaced by scalar multiplication with some positive constant a. For $n > 1$, the solution obeys

$$u^n = u^0 U_n(\zeta) - u^{-1} U_{n-1}(\zeta), \quad \zeta = 1 - \tfrac{1}{2}\Delta t^2 a.$$

where $U_n(\zeta)$, $\zeta \in [-1, 1]$, denotes a Chebyshev polynomial of the second kind, with $U_{-1}(\zeta) = 0$. Since the extrema occur at the endpoints and equal $U_n(\pm 1) = (n+1)(\pm 1)^n$, we can have linear growth with n. This is formally allowed, as for instance $u = ct$ with constant c obeys $\mathrm{d}^2u/\mathrm{d}t^2 = 0$ for $a = 0$. For values of ζ outside the interval $[-1, 1]$, the Chebyshev polynomials have power-law growth with n, so stability requires $|\zeta| \leq 1$. Next, we generalize from a scalar a and u to a spatial operator $\mathcal{L}$ acting on a vector $\mathbf{u}$. Stability then leads to the requirement

$$\Delta t \leq \frac{2}{\sqrt{\rho(\mathcal{L})}} \quad \text{and} \quad \mathcal{L} \geq 0, \tag{8}$$

where $\rho(\mathcal{L})$ denotes the spectral radius of the operator $\mathcal{L}$. The second requirement, that $\mathcal{L}$ be non-negative with real eigenvalues, is guaranteed for CML by construction (Mulder 1996; Chin-Joe-Kong *et al.* 1999) and for SIPDG by the coercivity and the symmetry of $\mathcal{L}$.

For problems of realistic size, the computation of the spectral radius is too costly. One way to obtain an estimate is by Geršgorin's circle theorem (1931), which states that for a square matrix A, its eigenvalues λ obey $|\lambda - a_{k,k}| \leq \sum_{m \neq k} |a_{k,m}|$. From this, it follows that $|\lambda| \leq \max_k \sum_m |a_{k,m}|$, so the spectral radius $\rho(A)$ is bounded by the largest row sum of absolute values of A. Potentially sharper estimates can be obtained by considering special cases for which the spectral radius can be computed and that are sufficiently generic to assume a more general validity. We will take that approach by considering the reference element and its distortions as well as the unit cube with six tetrahedra and periodic boundary conditions, without and with distortions.

Since this spectral radius contains information about the mesh, the question is how to express the stability condition into a mesh-independent constant CFL and mesh-dependent properties. A scaling argument suggests an expression of the form

$$\Delta t \leq \min_m \left(\sigma_m \frac{d_m}{c_m}\right) \leq \frac{2}{\sqrt{\rho(\mathcal{L})}}, \tag{9}$$

involving the ratio of some diameter d_m and the velocity c_m per element and the minimum is taken over all elements $E_m \in \Omega_h$. Ideally, the number σ_m should be independent of m and be equal to a constant number CFL. This Courant–Friedrichs–Lewy number (Courant *et al.* 1928) can then be taken as $\mathrm{CFL} = \min_m \sigma_m$, leading to a stability bound $\Delta t \leq \mathrm{CFL} \min_m (d_m/c_m)$. Below, we will adopt $\mathrm{CFL} \leq 2/[d_{\max}\sqrt{\rho(\bar{\mathcal{L}})}]$, using the spectral radius ρ of the spatial operator $\bar{\mathcal{L}}$ for a constant unit velocity and the largest diameter $d_{\max}$ over all elements for a specific case.

We have not been able to find a choice for the diameter d in (9) that produces a value of σ that is independent of the mesh and the velocity model, but some choices are better than others. As an example, consider the single reference element with vertices (0, 0, 0), (1, 0, 0), (0, 1, 0) and (0, 0, 1). If this is discretized with a degree-1 CML finite element with homogeneous Neumann boundary conditions for a unit velocity, the stiffness matrix $\mathcal{K}$ and lumped mass matrix $\mathcal{M}$ become

$$\mathcal{K} = \frac{1}{6}\begin{pmatrix} 3 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{M} = \frac{1}{24}\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}.$$

Following Shewchuk (2002), we can consider a version of this element on an arbitrary tetrahedron and compute the eigenvalues of

$\tilde{\mathcal{L}} = \mathcal{M}^{-1}\mathcal{K}$ for $c = 1$. One of those is zero and the other three are the roots of

$$\lambda^3 - \beta\lambda^2 + \gamma\lambda - \delta = 0, \tag{10}$$

with

$$\beta = 4J_0^{-2}\sum_{k=1}^{4} A_k^2, \quad \gamma = 16J_0^{-2}\sum_{j=1}^{6} \ell_j^2, \quad \delta = 256J_0^{-2},$$

where A_k is the area of face F_k, ℓ_j the length of edge j, and $J_0 = 6V$ with V the volume of the tetrahedron. Note that $\beta = \sum_{i=1}^{3}\lambda_i$ is the sum and $\delta = \prod_{i=1}^{3}\lambda_i$ the product of the non-zero eigenvalues. By symmetry of the spatial operator, all eigenvalues are non-negative, so $0 \le \lambda_i \le \beta$. We will consider the following four choices for a diameter d:

(i) $d_i = 2J_0/\sum_{k=1}^{4} A_k$, the diameter of the inscribed sphere,
(ii) $d_l = \min_j \ell_j$, the length of the shortest edge,
(iii) $d_b = 2/\sqrt{\beta}$, based on the sum of the eigenvalues mentioned above, or
(iv) $d_e = 2/\sqrt{\max_i \lambda_i}$.

To see which of these choices produces a nearly constant σ, we first consider the same case as before, but let the reference element have a varying height. Then, the four vertices are (0, 0, 0), (1, 0, 0), (0, 1, 0) and (0, 0, h). In this case, $J_0 = h$ and

$$d_i = 2h\big/(1 + 2h + \sqrt{1 + 2h^2}), \qquad d_l = \min(1, h),$$

$$d_b = h\big/\sqrt{2 + 4h^2}, \quad d_e = h\big/\sqrt{2[2 + 3h^2 + \sqrt{4 - h^2(4 - 9h^2)}]}.$$

The coefficients in (10) are $\beta = 16 + 8/h^2$, $\gamma = 48 + 96/h^2$, and $\delta = 256/h^2$, leading to eigenvalues 4 and $2h^{-2}(2 + 3h^2 \pm \sqrt{4 - 4h^2 + 9h^4})$, apart from the zero eigenvalue. For stability, we can define

$$\mathrm{CFL} = \min_h \sigma(h), \quad \sigma(h) = \frac{2}{d(h)\sqrt{\lambda_{\max}}},$$

$$\lambda_{\max} = \frac{2}{h^2}\left(2 + 3h^2 + \sqrt{4 - 4h^2 + 9h^4}\right),$$

where $d(h)$ can be one of the four choices mentioned above. By definition, $\sigma(h) = 1$ for $d(h) = d_e$. Fig. 1 displays $\sigma(h)$ for the other three choices of $d(h)$. Clearly, d_b is the second best in the sense that it has the smallest variation with h, namely $\sqrt{3/2}$ or about 22 per cent.

The idea of the stability analysis described in this paper is to estimate the spectral radius of the spatial discretization operator, in our case the CML finite-element operator and the SIPDG operator. One way to obtain estimates is by considering a single reference element and distorted version of it. These can be obtained by the coordinate transform

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} 1 & p_1 & q_1 \\ 0 & p_2 & q_2 \\ 0 & 0 & q_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{11}$$

where $p_2 > 0$ and $q_3 > 0$. Other shapes follow by symmetry. To avoid too strong distortions, we require $d_o/d_i \le 100$, where d_o is the diameter of the circumscribed and d_i of the inscribed sphere ($d_o \ge 3d_i$ for any tetrahedron). Here we assume that a mesh generation code will be able to construct tetrahedra with this property. We then estimated the smallest and largest occurring value of

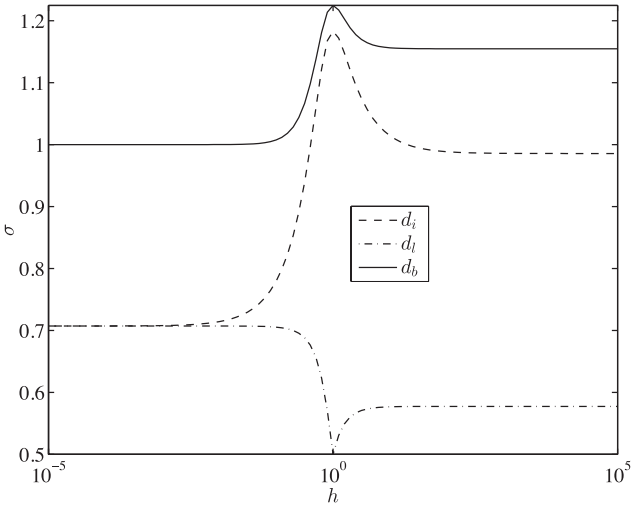

Figure 1. Graph of $\sigma(h)$ for three choices of $d(h)$, d_b, d_i and d_l. The result for d_e is not shown, since it equals 1 by definition.

$$\sigma(p_1, p_2, q_1, q_2, q_3) = \frac{2}{\sqrt{\lambda_{\max}}}\,\frac{1}{\max_m d(p_1, p_2, q_1, q_2, q_3)} \tag{12}$$

using $d = d_i$, d_b and d_e, respectively.

In case of the CML finite elements, the spectral radius of the spatial operator $\lambda_{\max}$ depends on five distortion parameters p_1, p_2, q_1, q_2, q_3. In case of the SIPDG operator, the spectral radius additionally depends on the interior penalty γ, enlarging the optimization problem (12) by one dimension.

Fourier or plane-wave analysis of the spatial operator is a natural approach in a finite-difference setting with a regular Cartesian grid. The method is also applicable to finite elements if a regular Cartesian grid is used, packed with tetrahedra. Here, we select a cube divided into six tetrahedra as shown on Fig. 2. Note that this cube also contains the reference element with a diameter of the inscribed sphere $1 - 1/\sqrt{3} = 0.423$. The diameters of the inscribed sphere for the six tetrahedra in the cube are either 0.360 or 0.423. As with the single tetrahedron, we can consider distorted versions by

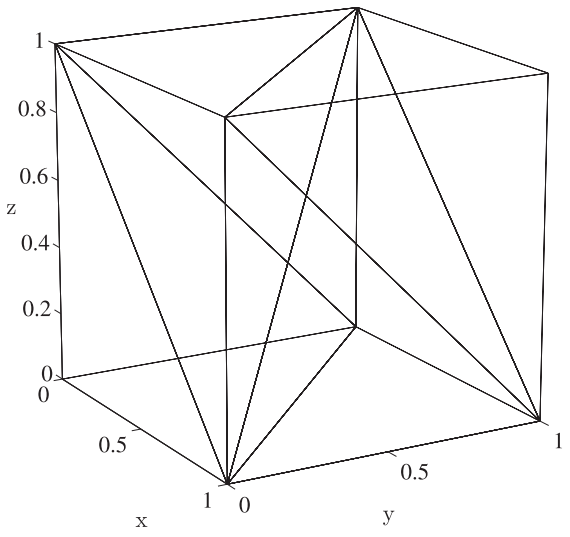

Figure 2. Unit cube packed with six tetrahedral elements, one of them being the reference element.

using the coordinate transform (11). The following cases will be considered:

(i) reference element with homogeneous Neumann boundary conditions (CML) or zero values in neighbouring elements (SIPDG);
(ii) distorted versions of the reference element;
(iii) unit cube containing six tetrahedra with periodic boundary conditions;
(iv) distorted versions of the latter.

For the single reference element and its distorted version, $\lambda_{\max} = \rho(\bar{\mathcal{L}})$ is the largest eigenvalue of the operator with unit velocity. With periodic boundaries, the spectral radius of the Fourier symbol $\hat{\bar{\mathcal{L}}}(\xi_1, \xi_2, \xi_3)$ of the spatial operator with unit velocity is

$$\lambda_{\max} = \max_{\xi_1,\xi_2,\xi_3} \rho\{\hat{\bar{\mathcal{L}}}(\xi_1, \xi_2, \xi_3)\},$$

where the ξ_k are normalized wavenumbers, specified later in Section 3.1, and $\rho\{\hat{\bar{\mathcal{L}}}(\xi_1, \xi_2, \xi_3)\}$ the spectral radius of the Fourier symbol of operator in the wavenumber domain for unit velocity, $c = 1$. We used the Nelder–Mead (1965) method for finding the extremum of the optimization problem (12). Since this method is not guaranteed to converge to the global minimum or maximum, several runs with random starting values were performed. In this way, we increased the chance of reaching the global extremum.

Next, we will review periodicity for the CML and the SIPDG elements separately.

3.1 Periodic mass-lumped continuous finite elements

We consider the four geometries described above. The construction of the matrices is straight-forward for the reference element and its distorted versions. Imposing periodic boundary conditions for the six tetrahedra requires additional calculations.

To build the matrix $\bar{\mathcal{L}} = \mathcal{M}^{-1}\mathcal{K}$ with unit velocity and periodic boundary conditions, we assembled the matrices $\mathcal{M}$ and $\mathcal{K}$ for a configuration of $3^3 = 27$ unit cubes on the domain $[-1, 2]^3$. Let $\mathcal{N}$ denote set of degrees of freedom in the central unit cube $[0, 1)^3$. We arrange the numerical solution of (7) with zero source term $\mathbf{s}$ on the $\mathcal{N}$ nodes in a single vector $\mathbf{u}$. The subset of the matrix $\bar{\mathcal{L}}$ that has $\mathcal{N}$ rows corresponding to the nodes inside $[0, 1)^3$ acts on this vector as well as on the degrees of freedom in neighbouring cubes.

To simplify the notation, we introduce a shift operator T. In the 1-D case with only x as coordinate, this shift operator is defined by

$$T^s u(x) = u(x + s\Delta x),$$

where integer s and Δx the spacing, in this case of unit length. The shift operator enables straightforward expression of the values on the interval $[-1, 2)$ into those inside $[0, 1)$ in the periodic case for a unit interval, as sketched in Fig. 3. Basically, T^{+1} is a shift to the right over a distance Δx and T^{-1} to the left over the same distance, whereas T^0 is the identity operator.

The Fourier symbol of T^s is

$$\hat{T}^s = \exp(\imath s k \Delta x),$$

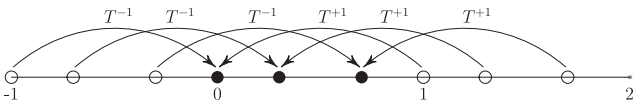

Figure 3. For the periodic 1-D case, the shift operators map nodes in the interval $[-1, 2)$ to $[0, 1)$.

where k is the wavenumber in the x-coordinate. We define $\xi = k\Delta x$ with $\xi \in (-\pi, \pi]$. Then,

$$\hat{T}^s = \exp(\imath s \xi).$$

To determine the spatial operator $\bar{\mathcal{L}}$ for the interval $[-1, 2]$, we can sort the solution vector $\mathbf{u}$ in a part $\mathbf{u}_L$ related to the interval $[-1, 0)$, a part $\mathbf{u}_C$ for the central interval $[0, 1)$ and a part $\mathbf{u}_R$ for the interval $[1, 2)$ on the right. We then can select the subset of rows from $\mathcal{L}$ that correspond to $[0, 1)$ and call that $\mathbf{L}$. Then, $\mathbf{Lu} = \mathbf{L}\mathbf{u}_L + \mathbf{L}\mathbf{u}_C + \mathbf{L}\mathbf{u}_R$ and we can set up $\mathbf{L}_L$ by taking the columns of $\mathbf{L}$ that correspond to $\mathbf{u}_L$, $\mathbf{L}_C$ by taking those corresponding to $\mathbf{u}_C$ and $\mathbf{L}_R$ as those corresponding to $\mathbf{u}_R$. This results in $\mathbf{Lu} = (\mathbf{L}_L T^{-1} + \mathbf{L}_C + \mathbf{L}_R T^{+1})\mathbf{u}$, *cf.* Fig. 3. The discrete Fourier symbol $\hat{\bar{\mathcal{L}}}$ of the periodic spatial operator $\bar{\mathcal{L}}$ becomes

$$\hat{\bar{\mathcal{L}}}(\xi) = \mathbf{L}_L \hat{T}^{-1} + \mathbf{L}_C + \mathbf{L}_R \hat{T}^{+1}.$$

The generalization to three dimensions is straightforward. For the analysis, we have assembled the global mass and stiffness matrix for the 27 cubes with 27×6 tetrahedra in distorted form and computed the spatial operator. To determine the shift operators, the distorted coordinates were mapped back to the unit cubes and its neighbours. We selected the rows of $\mathcal{L}$ that correspond to the central cube $[0, 1)^3$. For the columns, the node positions readily provided the shifts s_j for each of the coordinates $x_j, j = 1, 2, 3$. The component of $\mathcal{L}$ was then multiplied by $\hat{T}_j^{s_j}$ and added to the corresponding component for the node in the central cube to obtain $\hat{\bar{\mathcal{L}}}(\xi_1, \xi_2, \xi_3)$.

For mass-lumped elements of degree $M = 1$ and the unit cube, this approach produces

$$\begin{aligned}\hat{\bar{\mathcal{L}}} = \frac{1}{3}\Big[&20 - 2\left(\hat{T}_1^{-1} + \hat{T}_1\right) - 4\left(\hat{T}_2^{-1} + \hat{T}_2\right) - 4\left(\hat{T}_3^{-1} + \hat{T}_3\right)\\ &- \left(\hat{T}_1^{-1}\hat{T}_2 + \hat{T}_1\hat{T}_2^{-1}\right) - \left(\hat{T}_1^{-1}\hat{T}_3 + \hat{T}_1\hat{T}_3^{-1}\right)\\ &+ \left(\hat{T}_2^{-1}\hat{T}_3 + \hat{T}_2\hat{T}_3^{-1}\right) + \left(\hat{T}_1^{-1}\hat{T}_2\hat{T}_3 + \hat{T}_1\hat{T}_2^{-1}\hat{T}_3^{-1}\right)\Big]\\ = \frac{2}{3}\Big[&10 - 2\cos(\xi_1) - 4\cos(\xi_2) - 4\cos(\xi_3) - \cos(\xi_1 - \xi_2)\\ &- \cos(\xi_1 - \xi_3) + \cos(\xi_2 - \xi_3) + \cos(\xi_1 - \xi_2 - \xi_3)\Big].\end{aligned}$$

In this case, $\hat{\bar{\mathcal{L}}}(\xi_1, \xi_2, \xi_3)$ is a scalar. For higher degrees, it becomes a matrix. At the lowest degree $M = 1$, we obtain $0 \le \hat{\bar{\mathcal{L}}}(\xi_1, \xi_2, \xi_3) \le 6 + 3\sqrt{6} = 13.35$, where the minimum occurs at $(\xi_1, \xi_2, \xi_3) = 0$, among others, and the maximum at $\xi_1 = -2\arcsin\left(\frac{1}{2}(3/8)^{1/4}\right) = -0.8040$, $\xi_2 = \xi_3 = 2\arcsin\left(\sqrt{3/8 + \sqrt{3/8}}\right) = 2.916$, among others. For the reference element, which is one of the six tetrahedra in the unit cube, we have $d_i = 1 - 1/\sqrt{3}$, $d_b = 1/\sqrt{6}$, $d_e = 1/2$ and $d_l = \sqrt{2}$. The largest diameters over the six tetrahedra in the cube are: $d_i = 1 - 1/\sqrt{3}$, $d_b = 1/\sqrt{6}$, $d_e = 1/\sqrt{2 + \sqrt{2}}$ and $d_l = \sqrt{3}$ and the corresponding values of $\mathrm{CFL} = 2/[d_{\max}\sqrt{\rho(\hat{\bar{\mathcal{L}}})}]$ are $\mathrm{CFL}_i = \sqrt{(\sqrt{3} + 2)(\sqrt{6} - 2)} = 1.295$, $\mathrm{CFL}_b = 2\sqrt{\sqrt{6} - 2} = 1.341$, $\mathrm{CFL}_e = \sqrt{(2/3)(\sqrt{2} + 2)(\sqrt{6} - 2)} = 1.011$ and $\mathrm{CFL}_l = \sqrt{2/9(\sqrt{6} - 2)} = 0.316$.

The shortest edge length, d_l, resulted in extremely large variations of the estimates over all distortions, so we will only consider the other measures in the following.

3.2 Periodic discontinuous Galerkin finite elements

To estimate the spectral radius of the SIPDG spatial operator, we performed stability analysis on the reference element and its distorted versions as well as the unit cube with six tetrahedra and periodic boundary conditions, as shown on Fig. 4, and its distortions. As with the CML elements, the last two cases require additional computations for each set of wavenumbers.

For the unit cube with six tetrahedra, the stiffness matrix of the spatial operator including fluxes is given by

$$\begin{pmatrix} \mathcal{K}+\mathcal{F}^+ & & & & \mathcal{F}_2^- & \\ & \mathcal{K}+\mathcal{F}^+ & & \mathcal{F}_7^- & \mathcal{F}_4^- & \\ & & \mathcal{K}+\mathcal{F}^+ & \mathcal{F}_9^- & & \\ & \mathcal{F}_7^- & \mathcal{F}_9^- & \mathcal{K}+\mathcal{F}^+ & & \mathcal{F}_{12}^- \\ \mathcal{F}_2^- & \mathcal{F}_4^- & & & \mathcal{K}+\mathcal{F}^+ & \mathcal{F}_{15}^- \\ & & & \mathcal{F}_{12}^- & \mathcal{F}_{15}^- & \mathcal{K}+\mathcal{F}^+ \end{pmatrix},$$

where each line of the matrix represents an element and the subscript of the fluxes from neighbouring tetrahedra identifies the number of the face (Fig. 5). Empty entries correspond to blocks that contain zeros. The mass matrix is a block-diagonal matrix consisting of local mass matrices.

In order to add periodic boundary conditions, we use shift operators similar to those introduced for the CML elements. The components of the periodic spatial operator are then given by $\hat{T}_j^s \mathcal{F}_m^-$, where the index j denotes x_j-dimension ($j = 1, 2, 3$), $\mathcal{F}_m^-$ denotes a face adjacent to neighbouring elements outside the unit cube, s has the integer values of -1 and 1, meaning a shift to the left or to the right, respectively. The Fourier symbol of the periodic spatial operator is then computed as

$$\hat{\tilde{\mathcal{L}}} = \mathcal{M}^{-1} \begin{pmatrix} \mathcal{K}+\mathcal{F}^+ & \hat{T}_2^1\mathcal{F}_3^- & & \hat{T}_1^1\mathcal{F}_1^- & \mathcal{F}_2^- & \hat{T}_3^1\mathcal{F}_0^- \\ \hat{T}_2^{-1}\mathcal{F}_5^- & \mathcal{K}+\mathcal{F}^+ & \hat{T}_3^1\mathcal{F}_6^- & \mathcal{F}_7^- & \mathcal{F}_4^- & \\ & \hat{T}_3^{-1}\mathcal{F}_8^- & \mathcal{K}+\mathcal{F}^+ & \mathcal{F}_9^- & \hat{T}_1^{-1}\mathcal{F}_{10}^- & \hat{T}_2^{-1}\mathcal{F}_{11}^- \\ \hat{T}_1^{-1}\mathcal{F}_{13}^- & \mathcal{F}_7^- & \mathcal{F}_9^- & \mathcal{K}+\mathcal{F}^+ & & \mathcal{F}_{12}^- \\ \mathcal{F}_2^- & \mathcal{F}_4^- & \hat{T}_1^1\mathcal{F}_{14}^- & & \mathcal{K}+\mathcal{F}^+ & \mathcal{F}_{15}^- \\ \hat{T}_3^{-1}\mathcal{F}_{17}^- & & \hat{T}_2^1\mathcal{F}_{16}^- & \mathcal{F}_{12}^- & \mathcal{F}_{15}^- & \mathcal{K}+\mathcal{F}^+ \end{pmatrix}.$$

The spectral radius $\rho(\hat{\tilde{\mathcal{L}}})$ is the largest eigenvalue of the Fourier symbol of the spatial operator, $\hat{\tilde{\mathcal{L}}}$, over all sets of scaled wavenumbers $\{\xi_1, \xi_2, \xi_3\}$ and can again be found by a numerical maximization algorithm.

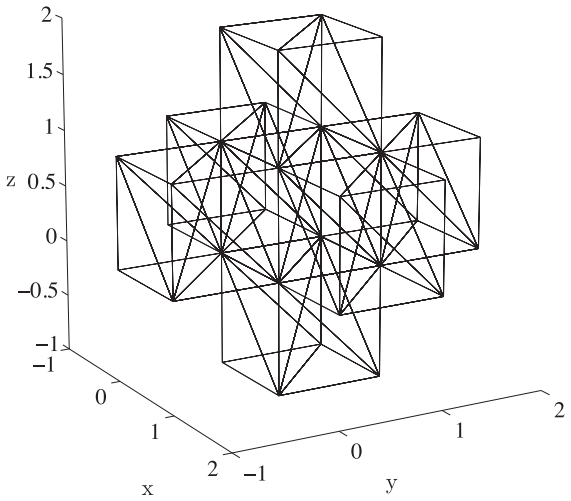

Figure 4. Unit cube with its neighbours for DG.

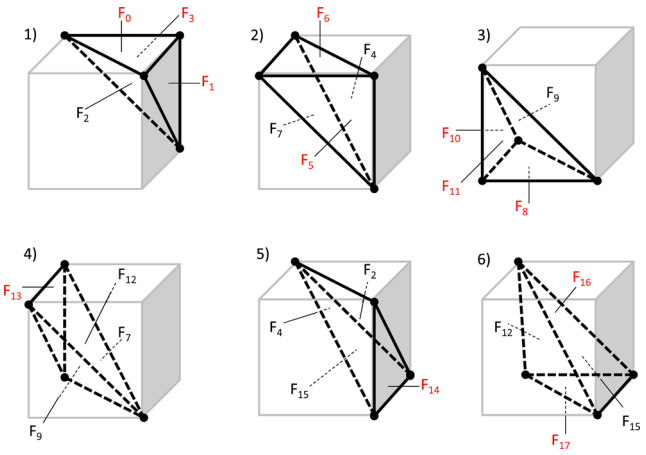

Figure 5. Unit cube with six tetrahedra. The internal and external faces are marked by black and red, respectively.

4 RESULTS

4.1 Estimated CFL numbers

Table 1 lists the values of σ for the CML reference element as well as the smallest and largest values obtained for a single distorted element for three choices of the diameter. For d_e, we observe no variation for the CML element of degree 1, by definition. For higher-order elements however, the smallest variation appears to be obtained for d_b. Note that the CML element of degree 3 has a more favourable stability limit for type 2 than for type 1. These types are as defined by Chin-Joe-Kong *et al.* (1999), who found two mass-lumped elements of degree 3 augmented with polynomials of degree 5 on the faces and degree 6 in the interior, leading to a total of 50 nodes per element. The two types differ in the positions of the nodes and the corresponding integration weights. The results for $M = 1$ can be found analytically. For d_i, the minimum value is $1/\sqrt{2}$ for a flat degenerate element, which will not be reached numerically because of the diameter constraint. We have $(3 + \sqrt{3})/4$ for the reference element and a maximum value of $\sqrt{3}$, reached for the equilateral tetrahedron. For d_b, the results are 1, $\sqrt{3/2}$, and $\sqrt{3}$. For d_e, $\sigma = 1$ by definition.

Table 2 shows similar results but for the cube containing six tetrahedra with periodic boundary conditions. The results for the unit cube lie between the smallest and largest values obtained for the distorted cases. An estimate of CFL can be obtained by choosing the smallest value of σ in these two tables. These are underlined.

For the SIPDG method, we first consider a single element with Neumann boundary conditions. In that case, all fluxes are zero. This can be considered as an extreme case and does not involve the penalty terms, so the same result would be obtained for IIPDG and NIPDG as well as the classic choice of finite elements that requires inversion of the full mass matrix. Table 3 lists the values of σ for a single distorted element for each of the three choices of the diameter. Note that for degree $M = 1$, the results are different

Table 1. Values of σ based on the spectral radius for a single reference element and its distorted versions with homogeneous Neumann boundary conditions for various CML elements of degree M. For each of the diameters d_i, d_b or d_e, the smallest and largest occurrence of σ is listed, with the value for the undistorted reference element in between. The ratio of largest to smallest is given in brackets.

M	Type	d_i				d_b				d_e			
1		0.717	1.18	1.73	(2.42)	1.00	1.22	1.73	(1.73)	1.00	1.00	1.00	(1.00)
2		0.0616	0.0900	0.118	(1.92)	0.0790	0.0932	0.118	(1.49)	0.0681	0.0761	0.0990	(1.45)
3	1	0.0422	0.0590	0.0624	(1.48)	0.0561	0.0611	0.0624	(1.11)	0.0360	0.0499	0.0615	(1.71)
3	2	0.0720	0.105	0.121	(1.68)	0.0960	0.109	0.121	(1.26)	0.0697	0.0887	0.112	(1.60)

Table 2. As Table 1, but for the distorted unit cube with periodic boundary conditions. The worst-case values are underlined and can serve as an estimate of CFL.

M	Type	d_i				d_b				d_e			
1		0.848	1.30	1.41	(1.66)	1.01	1.34	1.41	(1.41)	0.775	1.01	1.41	(1.82)
2		0.0705	0.104	0.114	(1.62)	0.0942	0.107	0.114	(1.21)	0.0625	0.0810	0.110	(1.76)
3	1	0.0367	0.0557	0.0624	(1.70)	0.0475	0.0576	0.0623	(1.31)	0.0339	0.0435	0.0601	(1.77)
3	2	0.0662	0.104	0.120	(1.82)	0.0877	0.107	0.120	(1.37)	0.0652	0.0811	0.105	(1.62)

from those in Table 1, since we use the full mass matrix instead of the lumped one. This property was exploited by Marfurt (1984) to reduce numerical dispersion by taking a weighted average of the consistent and lumped mass matrices.

Next, we consider SIPDG with $\gamma = \gamma_0$, as defined in eq. (5). We assumed that the solution was zero in neighbouring elements and only the fluxes from the interior, $\mathcal{F}^+$, were included in the computed spectral radii. For the reference case, the matrices were assembled on a mesh with five elements, the central one being the reference element and the other four its mirror image in each face. Its distorted versions were obtained by the coordinate transformation given earlier.

Table 4 shows the corresponding values for σ. Table 5 lists the results for the cube containing six tetrahedra and periodic boundary conditions, without and with distortions. The assumption of zero-valued neighbours in Table 5 provides values closer to the periodic result in Table 5 than the Neumann boundary conditions in Table 3. The worst-case results are underlined and can serve as an estimate for CFL.

Contrary to CML, the choice of diameter, d_b, provides a similar variation as d_i over all distortions considered, but both are consirable smaller that the range observed for d_e.

Compared to the CML elements, the value of CFL for the SIPDG method is smaller by about a quarter for degree 1. For degree 2, the value for SIPDG is about 40 per cent larger than that of SIPDG. The same is true for degree 3 and CML type 1, whereas CML degree 3 type 2 has CFL about two times larger than SIPDG. The degree 4 CML element is unknown at present.

The reader should bear in mind that, since the results in the tables were obtained by numerical optimization, there is chance that the global extrema were missed.

4.2 Dependence on the penalty parameter

So far, the penalty parameter γ was taken per element as $\gamma_0 = M(M + 2)/d_i$. We also considered σ as a function of $\alpha = \gamma/\gamma_0$.

Fig. 6 shows the smallest and largest values of $\sigma(\alpha)$, as well as those for the reference element, for SIPDG in case of a distorted cube with six tetrahedra and periodic boundary conditions. As before, CFL can be taken as the smallest value of σ for a given α. As mentioned before, the estimates for σ are based on the spectral radius, the largest eigenvalue, of the spatial operator. Its smallest eigenvalue should be zero, but in some cases becomes negative because the coercivity bound is violated for some $\alpha \leq 1$, meaning that the spatial discretization becomes unstable. In that case, the curves are truncated at the smallest value of α that still results in a smallest non-negative eigenvalue.

Table 3. Values of σ for SIPDG on the reference element and its distorted versions with Neumann boundary conditions. For each of the diameters d_i, d_b or d_e, the smallest and largest occurrence of σ is listed, with the value for the undistorted reference element in between. The ratio of largest to smallest is given in brackets.

M	d_i				d_b				d_e			
1	0.320	0.529	0.775	(2.42)	0.447	0.548	0.775	(1.73)	0.447	0.447	0.447	(1.00)
2	0.181	0.267	0.374	(2.07)	0.228	0.276	0.374	(1.64)	0.216	0.226	0.252	(1.17)
3	0.118	0.169	0.232	(1.97)	0.144	0.175	0.232	(1.61)	0.134	0.143	0.165	(1.23)
4	0.0832	0.118	0.163	(1.96)	0.100	0.122	0.163	(1.63)	0.0941	0.0995	0.116	(1.23)

Table 4. Values of σ for SIPDG on the reference element and its distorted versions, assuming that the neighbouring elements have zero values. For each of the diameters d_i, d_b or d_e, the smallest and largest occurrence of σ is listed, with the value for the undistorted reference element in between.

M	d_i				d_b				d_e			
1	0.226	0.359	0.359	(1.59)	0.260	0.372	0.431	(1.66)	0.172	0.304	0.430	(2.50)
2	0.108	0.175	0.175	(1.63)	0.117	0.181	0.199	(1.70)	0.0891	0.148	0.199	(2.23)
3	0.0635	0.106	0.106	(1.67)	0.0734	0.110	0.117	(1.59)	0.0513	0.0897	0.117	(2.27)
4	0.0424	0.0705	0.0705	(1.66)	0.0487	0.0729	0.0771	(1.58)	0.0342	0.0596	0.0770	(2.25)

Table 5. As Table 4, but for six tetrahedra in the unit cube with periodic boundaries and its distorted versions.

M	d_i				d_b				d_e			
1	0.179	0.251	0.269	(1.51)	0.234	0.259	0.356	(1.52)	0.148	0.196	0.351	(2.37)
2	0.0870	0.119	0.128	(1.47)	0.111	0.123	0.170	(1.52)	0.0702	0.0930	0.169	(2.40)
3	0.0524	0.0730	0.0791	(1.51)	0.0675	0.0756	0.100	(1.48)	0.0428	0.0567	0.0998	(2.33)
4	0.0346	0.0478	0.0517	(1.49)	0.0447	0.0495	0.0663	(1.49)	0.0283	0.0373	0.0664	(2.34)

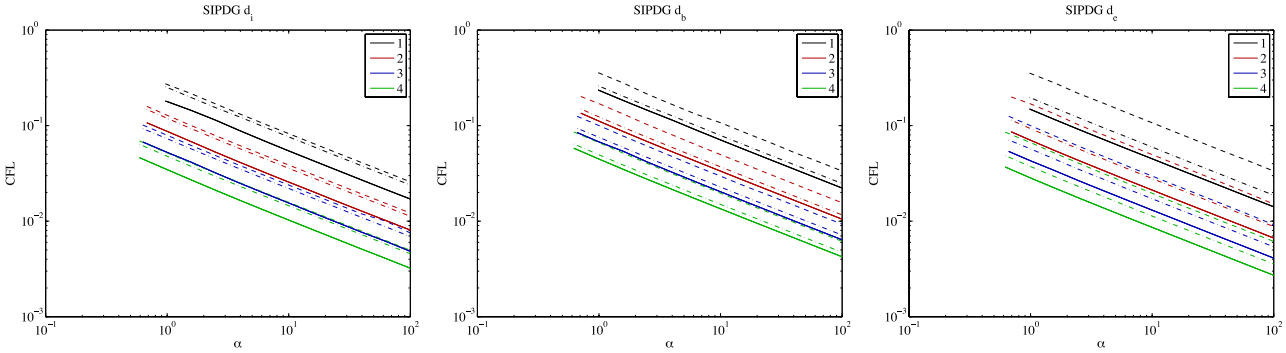

Figure 6. SIPDG stability limits for (a) d_i, (b) d_b and (c) d_e as a function of the penalty parameter α for the periodic case. The colours mark the degrees from 1 to 4. Drawn lines correspond to the smallest values of σ, which can be taken as an estimate for CFL, and dashed to the largest over all distortions considered. The smallest variation occurs for d_b. The dot–dashed lines refer to the undistorted case. The curves truncate at a smallest value of α where the smallest eigenvalues of the spatial operator becomes negative.

Table 6. Values of α_0 for SIPDG and various choices of diameter, assuming that $\sigma(\alpha) \sim \sqrt{\alpha_0/\alpha}$ for $\alpha \geq 1$. The last two rows provide data fits for the behaviour of α_0 with degree M from 1 to 4 as a function of N_{DG} and $M(M+2)$, respectively. The latter provides slightly more accurate fits.

M	d_i	d_b	d_e
1	0.030	0.051	0.020
2	0.0068	0.0011	0.0046
3	0.0025	0.0042	0.0017
4	0.00011	0.0018	0.00075
$\alpha_0 \simeq$	$\frac{0.234}{N_{\mathrm{DG}}^{1.53}}$	$\frac{0.392}{N_{\mathrm{DG}}^{1.52}}$	$\frac{0.156}{N_{\mathrm{DG}}^{1.51}}$
$\alpha_0 \simeq$	$\frac{0.172}{[M(M+2)]^{1.60}}$	$\frac{0.288}{[M(M+2)]^{1.58}}$	$\frac{0.115}{[M(M+2)]^{1.58}}$

The results in Fig. 6 are represented on a log–log scale. The lines are close to linear, meaning that σ follows a power law in α. When we determined linear fits to the logarithmic values, we found a power of -0.5 within less than 5 per cent. Turning this into the assumption that $\sigma(\alpha) \sim \sqrt{\alpha_0/\alpha}$, we then derived α_0 for the various cases. The results are listed in Table 6. The last row lists the result of fitting α_0 as function of $M(M+2)$ on a log–log scale. These result can be used for implementation in a simulation code, with an estimated error of around 5 per cent for $\alpha \geq 1$ and $1 \leq M \leq 4$.

4.3 Higher-order time stepping

Higher-order time stepping can be obtained by the Cauchy–Kovalewski or Lax–Wendroff procedure (Lax & Wendroff 1960) that substitutes higher time-derivatives with spatial derivatives using the partial differential equation. This approach is also known as the modified equation approach (Shubin & Bell 1987) or Dablain's scheme (Dablain 1986). The idea is to replace higher-order time derivatives in the Taylor series expansion of the time stepping scheme by spatial derivatives. For an even time-stepping order M_t,

Table 7. With time-stepping of order M_t, CFL has to be multiplied by $\sqrt{\sigma_t}$.

M_t	2	4	6	8	10	12	14	16
$\sqrt{\sigma_t}$	1	1.73	1.38	2.32	1.54	2.77	1.57	3.04

the truncated Taylor series in the absence of source terms is

$$\mathbf{u}^{n+1} - 2\mathbf{u}^n + \mathbf{u}^{n-1} = 2\sum_{m=1}^{M_t/2} \frac{(\Delta t)^{2m}}{(2m)!} \frac{\partial^{2m}\mathbf{u}^n}{\partial t^{2m}} = 2\sum_{m=1}^{M_t/2} \frac{(-\Delta t^2 \mathcal{L})^m}{(2m)!}\mathbf{u}^n, \tag{13}$$

which can be evaluated by repeated application of the spatial operator. The maximum allowable time step is increased by a factor $\sqrt{\sigma_t}$ relative to $M_t = 2$, which can be found in a paper by Gilbert & Joly (2008) and is listed for some choice of M_t in Table 7 for completeness, correcting errors in Chen (2009) and De Basabe & Sen (2010). Note that the alternating values of $\sqrt{\sigma_t}$ approach $\pi/2$ or π for large M_t.

5 CONCLUSIONS

Estimates of the time-stepping stability constant were considered for the continuous mass-lumped and symmetric interior-penalty DG finite elements. They were obtained by numerical optimization for a set of example problems and represent a necessary condition for stability. There may be extreme cases where smaller values of the CFL-number are required, but we have not encountered those so far in various numerical simulations. Different choices of the measure of element size were suggested such as the diameter of the inscribed sphere, the length of the shortest edge and measures based on the spectral properties of the spatial operator for the lowest-degree mass-lumped finite element on a single tetrahedron with natural boundary conditions. Based on each measure, the values of the CFL-number were computed for both methods on the reference element as well as a single tetrahedral element of arbitrary shape. Also, the unit cube packed with six tetrahedra and periodic boundary

conditions was considered. For the SIPDG method, we examined variations of the penalty parameter. We found that the smallest variation of CFL appears to be obtained for the measure based on the sum of the eigenvalues for the above-mentioned lowest-degree mass-lumped element, both for the higher-order CML elements as well as for the SIPDG elements. This measure resembles the expression for the diameter of the inscribed sphere but has a slightly lower computational cost.

The first-degree mass-lumped elements have a much larger CFL, allowing for a larger time step than required for SIPDG. The CFL for elements of second degree are comparable. Of the two variants of the third degree mass-lumped element, the second allows for a larger time step than the cubic SIPDG element. Since mass-lumped elements of degree 4 and higher are presently unknown, only stability results for SIPDG were given. The general trend for SIPDG is that CFL becomes smaller for higher-order elements.

REFERENCES

Agut, C., Bart, J.-M. & Diaz, J., 2011. Numerical study of the stability of the Interior Penalty Discontinuous Galerkin method for the wave equation with 2D triangulations, Tech. Rep. 7719, INRIA, Bordeaux.

Ainsworth, M., Monk, P. & Muniz, W., 2006. Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation, *J. Scient. Comput.,* **27**(1–3), 5–40.

Arnold, D.N., 1982. An interior penalty finite element method with discontinuous elements, *SIAM J. Numer. Anal.,* **19,** 742–760.

Arnold, D.N., Brezzi, F., Cockburn, B. & Marini, L.D., 2002. Unified analysis of discontinuous Galerkin methods for elliptic problems, *SIAM J. Numer. Anal.,* **39**(5), 1749–1779.

Charney, J.G., Fjörtoft, R. & von Neumann, J., 1950. Numerical integration of the barotropic vorticity equation, *Tellus,* **2,** 237–254.

Chen, J.-B., 2009. Lax-Wendroff and Nyström methods for seismic modelling, *Geophys. Prospect.,* **57**(6), 931–941.

Chin-Joe-Kong, M.J.S., Mulder, W.A. & van Veldhuizen, M., 1999. Higher-order triangular and tetrahedral finite elements with mass lumping for solving the wave equation, *J. Eng. Math.,* **35**(4), 405–426.

Cohen, G., Joly, P., Roberts, J.E. & Tordjman, N., 2001. Higher order triangular finite elements with mass lumping for the wave equation, *SIAM J. Numer. Anal.,* **38**(6), 2047–2078.

Courant, R., Friedrichs, K. & Lewy, H., 1928. Über die partiellen Differenzengleichungen der mathematischen Physik, *Math. Ann.,* **100**(1), 32–74.

Dablain, M.A., 1986. The application of high-order differencing to the scalar wave equation, *Geophysics,* **51**(1), 54–66.

Darlow, B.L., 1980. A penalty-Galerkin method for solving the miscible displacement problem, *PhD thesis,* Rice University, Houston, Texas.

Dawson, C., Sun, S. & Wheeler, M., 2004. Compatible algorithms for coupled flow and transport, *Comp. Methods Appl. Mech. Eng.,* **193**(23–26), 2565–2580.

De Basabe, J.D. & Sen, M.K., 2010. Stability of the high-order finite elements for acoustic or elastic wave propagation with high-order time stepping, *Geophys. J. Int.,* **181**(1), 577–590.

De Basabe, J.D., Sen, M.K. & Wheeler, M.F., 2008. The interior penalty discontinuous Galerkin method for elastic wave propagation: grid dispersion, *Geophys. J. Int.,* **175**(1), 83–93.

Epshteyn, Y. & Rivière, B., 2007. Estimation of penalty parameters for symmetric interior penalty Galerkin methods, *J. Comput. Appl. Math.,* **206**(2), 843–872.

Fried, I. & Malkus, D.S., 1975. Finite element mass matrix lumping by numerical integration with no convergence rate loss, *Int. J. Solids Struct.,* **11,** 461–466.

Geršgorin, S., 1931. Über die Abgrenzung der Eigenwerte einer Matrix, *Izvestiya Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk, Seriya VII,* **6,** 749–754.

Gilbert, J.C. & Joly, P., 2008. Higher order time stepping for second order hyperbolic problems and optimal CFL conditions, pp. 67–93, *Computational Methods in Applied Sciences*, Vol. 16, Springer.

Grote, M., Schneebeli, A. & Schotzau, D., 2006. Discontinuous Galerkin finite element method for the wave equation, *SIAM J. Numer. Anal.,* **44**(6), 2408–2431.

Grote, M.J. & Mitkova, T., 2010. Discontinuous Galerkin methods and local time stepping for wave propagation, in *Proceedings of ICNAAM 2010: International Conference of Numerical Analysis and Applied Mathematics, AIP Conference Proceedings,* Vol. **1281,** pp. 2119–2122.

Grote, M.J. & Schötzau, D., 2009. Optimal error estimates for the fully discrete interior penalty DG method for the wave equation, *J. Scient. Comput.,* **40**(1–3), 257–272.

Lax, P. & Wendroff, B., 1960. Systems of conservation laws, *Commun. Pure appl. Math.,* **31**(2), 217–237.

Marfurt, K., 1984. Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations, *Geophysics,* **49**(5), 533–549.

Minisini, S., Zhebel, E., Kononov, A. & Mulder, W.A., 2012. Efficiency comparisons for higher-order continuous mass-lumped and discontinuous Galerkin finite-element methods for the 3-D acoustic wave equation, in *Proceedings of the 74th EAGE Conference & Exhibition,* Extended Abstracts, A004, Copenhagen, Denmark.

Minisini, S., Zhebel, E., Kononov, A. & Mulder, W.A., 2013. Local time stepping with the discontinuous Galerkin method for wave propagation in 3-D heterogeneous media, *Geophysics,* **78**(3), T67–T77.

Mulder, W.A., 1996. A comparison between higher-order finite elements and finite differences for solving the wave equation, in *Proceedings of the Second ECCOMAS Conference on Numerical Methods in Engineering,* pp. 344–350, John Wiley & Sons.

Mulder, W.A., 2013. New triangular mass-lumped finite elements of degree six for wave propagation, *Prog. Electromag. Res.,* **141,** 671–692.

Nelder, J.A. & Mead, R., 1965. A simplex method for function minimization, *Comput. J.,* **7**(4), 308–313.

Oden, J., Babushka, I. & Baumann, C., 1998. A discontinuous *hp* finite element method for diffusion problems, *J. Comput. Phys.,* **146**(2), 491–519.

Rivière, B., 2008. *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation,* Vol. 35 of Frontiers in Mathematics, SIAM.

Rivière, B., Wheeler, M.F. & Girault, V., 1999. Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I, *Comput. Geosci.,* **3**(3–4), 337–360.

Rivière, B., Wheeler, M.F. & Girault, V., 2002. A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems, *SIAM J. Numer. Anal.,* **39**(3), 902–931.

Shahbazi, K., 2005. An explicit expression for the penalty parameter of the interior penalty method, *J. Comput. Phys.,* **205**(2), 401–407.

Shewchuk, J.R., 2002. What is a good linear finite element?—Interpolation, conditioning, anisotropy, and quality measures, in *Proceedings of the 11th International Meshing Roundtable*, Ithaca, New York, pp. 115–126, Sandia National Laboratories.

Shubin, G.R. & Bell, J.B., 1987. A modified equation approach to constructing fourth order methods for acoustic wave propagation, *SIAM J. Scient. Stat. Comput.,* **8**(2), 135–151.

Tordjman, N., 1995. Élements finis d'order élevé avec condensation de masse pour l'equation des ondes, *PhD thesis,* L'Université Paris IX Dauphine.

Warburton, T. & Hesthaven, J., 2003. On the constants in *hp*-finite element trace inverse inequalities, *Comput. Methods Appl. Mech. Eng.,* **192**(25), 2765–2773.

Wheeler, M.F., 1978. An elliptic collocation-finite element method with interior penalties, *SIAM J. Numer. Anal.,* **15**(1), 152–161.

APPENDIX: PENALTY PARAMETER

The penalty parameter, γ, should be large enough to ensure coercivity. If chosen too large, the stability limit for time stepping will become too small and make the method inefficient. Therefore,

a sharp estimate of a lower bound γ_0 is desirable. Epshteyn & Rivière (2007) provide a lower bound for SIPDG that guarantees coercivity. Here, we will derive alternative, sharper bounds, following their paper. The first one is based on dropping the estimate in eq. (56) of that paper and just use the expression on its left-hand side. The second one replaces the estimate per face by one per element. For simplicity, we will assume zero Neumann boundary conditions, which produce zero flux terms on the boundary of the domain.

We start with the first, based on eq. (56) of the paper mentioned above:

$$\frac{A}{V} \leq \frac{9}{2}\frac{d_l^{\max}|\cot\theta_m|}{A}, \tag{A1}$$

with V the volume of a given element and A the area of one of its faces, $d_l^{\max}$ the length of its longest edge and θ_m the dihedral angle that has the smallest value of $\sin\theta$ over all dihedral angles θ in the tetrahedron. The dihedral angle is defined as the angle between two faces and its cosine can be taken as the dot product of the face normals: $\cos\theta_{k,l} = -\mathbf{n}_k \cdot \mathbf{n}_l$ for two faces k and l. Then, $\sin\theta_{k,l} = \sqrt{1-\cos^2\theta_{k,l}}$ and θ_m has the smallest value of $\sin\theta_{k,l}$ among all six pairs of faces. Note that the evaluation of $|\cot\theta_m|$ does not require the explicit computation of θ_m. For an interior face $F(E^+, E^-)$ between two elements E^+ and E^-, Epshteyn & Rivière (2007) determine the penalty bound

$$\gamma_{0,F}^{(1)} = \lambda_\varepsilon^2 \frac{3}{2} M(M+2)\, A_F^{-1}\,(\psi_{E^+} + \psi_{E^-}), \quad \psi = d_l^{\max}|\cot\theta_m|. \tag{A2}$$

The area $A_{F(E^+,E^-)}$ of the common face is abbreviated to A_F. We have incorporated an additional factor λ_ε^2 for the generalization of SIPDG ($\varepsilon = -1$) to IIPDG ($\varepsilon = 0$) and NIPDG ($\varepsilon = 1$), to be discussed later on.

We can simply choose the left-hand side of (A1) instead of its right-hand side to obtain a sharper bound that is also easier to compute:

$$\gamma_{0,F}^{(2)} = \lambda_\varepsilon^2 \frac{1}{3} M(M+2)\, A_F \left(V_{E^+}^{-1} + V_{E^-}^{-1}\right). \tag{A3}$$

In either case, the local bounds can be made global. Then,

$$\gamma_0^{(1)} = \lambda_\varepsilon^2\, 3\; M(M+2)\frac{\max_E(d_{l,E}^{\max}|\cot\theta_{m,E}|)}{\min_E A_{F_E}}$$

and

$$\gamma_0^{(2)} = \lambda_\varepsilon^2 \frac{2}{3} M(M+2) \max_E\left(\frac{\max_{F_E} A_{F_E}}{V_E}\right)$$

at the expense of sharpness. Note that this an element-wise formulation. It is easy to see how to change this into a face-based approach.

The second estimate is based on the inequality in theorem 4 by Warburton & Hesthaven (2003) for polynomials basis functions v of degree M:

$$\|v\|_{\partial E}^2 \leq \kappa_E^{1/2}\|v\|_E^2, \quad \kappa = \frac{2(M'+1)(M'+3)}{d_i} \tag{A4}$$

We have used the expression for the diameter of the inscribed sphere, $d_i = 6\,V/\sum_{k=1}^4 A_k$, of a tetrahedral element E with volume V and boundary ∂E consisting of four faces with areas A_k. The inequality (A4) will be applied to ∇v instead of v, requiring $M' = M-1$. Epshteyn & Rivière (2007) employ this estimate for one face at the time, replacing the full boundary ∂E by only a single face, thereby sacrificing sharpness. To avoid this, we will use (A4) directly by going to an element-wise formulation of the fluxes in the discrete scheme rather than one per face.

Let the discretization of the stiffness term, $-\int_\Omega v\Delta u\,\mathrm{d}\Omega$, be given by

$$A(u,v) = \int_\Omega \nabla u \cdot \nabla v\,\mathrm{d}\Omega + \sum_F \left[\varepsilon J_F(v,u) - J_F(u,v) + \gamma_F G_F(u,v)\right].$$

Here,

$$J_{F(E^+,E^-)}(u,v) = \int_F [v]\cdot\{\nabla u\}\,\mathrm{d}\Gamma = \frac{1}{2}\int_F (v^+\mathbf{n}^+ + v^-\mathbf{n}^-)\cdot(\nabla u^+ + \nabla u^-)\,\mathrm{d}\Gamma,$$

$$G_{F(E^+,E^-)}(u,v) = \int_F [u]\cdot[v] = \gamma_F\int_F (u^+-u^-)(v^+-v^-)\,\mathrm{d}\Gamma,$$

and u and v are now expanded into piecewise polynomial basis functions of degree M. We earlier made the assumption of zero Neumann boundary conditions for simplicity. This allows us to ignore the faces on the boundary Γ_e of the domain. We would like to replace the summation over the individual faces by a summation over faces per element, taking the element E^+ as reference, for instance. If we set $\gamma_{F(E^+,E^-)} = \frac{1}{2}(\gamma_{E^+} + \gamma_{E^-})$, then

$$\sum_F \gamma_F G_F(u,v) = \sum_{F(E^+,E^-)} \frac{1}{2}\gamma_{E^+} G_{F(E^+,E^-)}(u,v) + \sum_{F(E^-,E^+)} \frac{1}{2}\gamma_{E^-} G_{F(E^-,E^+)}(u,v) = \sum_E \frac{1}{2}\gamma_E \sum_{k=1}^4 \int_{F_k} (u-u_k^-)(v-v_k^-)\,\mathrm{d}\Gamma.$$

In the last expression, the superscript $^+$ that denotes the element under consideration has been dropped whereas the superscript $^-$ referring to the neighbouring elements has been retained. The inner summation involves all four faces, F_k. Note that interior faces are visited twice in this way, from $E = E^+$ and from $E = E^-$. We can do the same with

$$J_F(u,v) = \int_F [v]\cdot\{\nabla u\}\,\mathrm{d}\Gamma = \frac{1}{2}\int_F (v^+ - v^-)\mathbf{n}^+\cdot\nabla u^+\,\mathrm{d}\Gamma + \frac{1}{2}\int_F (v^- - v^+)\mathbf{n}^-\cdot\nabla u^-\,\mathrm{d}\Gamma,$$

using $\mathbf{n}^+ = -\mathbf{n}^-$. Then,

$$\sum_F J_F(u,v) = \sum_E\sum_{k=1}^4 \frac{1}{2}\int_k (v-v^-)\mathbf{n}\cdot\nabla u\,\mathrm{d}\Gamma.$$

For coercivity, we need $A(v,v) \geq c_0\|v\|_\Omega^2$ with $c_0 > 0$. The sum of $J_F(v,v)$ over the four faces F_E of a given element E obeys

$$\sum_{k=1}^4 \frac{1}{2}\int_k (v-v^-)\mathbf{n}\cdot\nabla v\,\mathrm{d}\Gamma \leq \frac{1}{2}\|[v]\|_{\partial E}\|\mathbf{n}\cdot\nabla v\|_{\partial E}.$$

This can be further bounded by the inverse inequality (Warburton & Hesthaven 2003, theorem 4), quoted here in (A4) and applied to ∇v instead of v, requiring $M' = M-1$. Using Young's

inequality, $2\langle a, b\rangle \le \psi^{-1}\langle a, a\rangle + \psi\langle b, b\rangle$, with $a = \|\mathbf{n}.\nabla v\|_E$ and $b = \kappa_E^{1/2}\|[v]\|_{\partial E}$, we obtain

$$\sum_E \sum_{k=1}^{4} J_k^E(v, v) \le \sum_E \frac{1}{4}\psi^{-1}\|\nabla v||_E^2 + \sum_E \frac{1}{4}\psi\kappa_E\|[v]\|_{\partial E}^2.$$

From

$$A(v, v) = \sum_E \|\nabla v\|_E^2 + \sum_E \gamma_E \sum_{k=1}^{4} \frac{1}{2}\|v - v_k^-\|_k^2 - (1-\varepsilon)\sum_E \sum_{k=1}^{4} \frac{1}{2} J_k^E(v, v),$$

it follows that

$$A(v, v) \ge \left(1 - \frac{1-\varepsilon}{4\psi}\right)\|\nabla v\|_\Omega^2 + \sum_E \left[\frac{\gamma_E}{2} - \frac{\psi}{4}(1-\varepsilon)\kappa_E\right]\|[v]\|_{\partial E}^2.$$

Let $\psi = \frac{1}{4}(1-\varepsilon)$ and

$$\gamma_E > \frac{1}{2}\psi(1-\varepsilon)\kappa_E = \frac{1}{2}\left(\frac{1-\varepsilon}{2}\right)^2 \kappa_E = \left(\frac{1-\varepsilon}{2}\right)^2 \frac{M(M+2)}{d_i}.$$

We therefore define

$$\gamma_{0,E}^{(3)} = \lambda_\varepsilon^2 \frac{M(M+2)}{d_i}. \tag{A5}$$

According to the above, $\lambda_\varepsilon = (1-\varepsilon)/2$. This shows that for NIPDG with $\varepsilon = 1$, the penalty factor can be arbitrarily small. For accuracy, though, a finite value is needed, for which for instance the SIPDG value can be taken.

As an aside: in earlier work, we made the educated guess $\gamma_{0,E}^{(4)} = N_{\mathrm{DG}}/\min_E d_{i,E}$, with $N_{\mathrm{DG}} = (M+1)(M+2)(M+3)/6$ the number of nodes per element, a 3-D generalization of the 2-D estimate by Ainsworth *et al.* (2006) for SIPDG. The ratio of $M(M+2)$ to N_{DG} happens to be about 0.75 for the degrees M from 1 to 4 considered in this paper.