

# Unstructured Discontinuous Galerkin for Seismic Inversion

*S. Scott Collis, Curtis C. Ober, and Bart G. van Bloemen Waanders, Sandia National Laboratories*

## SUMMARY

This abstract explores the potential advantages of discontinuous Galerkin (DG) methods for the time-domain inversion of media parameters within the earth's interior. In particular, DG methods enable local polynomial refinement to better capture localized geological features within an area of interest while also allowing the use of unstructured meshes that can accurately capture discontinuous material interfaces. This abstract describes our initial findings when using DG methods combined with Runge-Kutta time integration and adjoint-based optimization algorithms for full-waveform inversion. Our initial results suggest that DG methods allow great flexibility in matching the media characteristics (faults, ocean bottom and salt structures) while also providing higher fidelity representations in target regions.

## INTRODUCTION

Inversion of earth media parameters is of primary importance to exploration geophysics with the specific goal of constructing accurate subsurface characterizations in a computationally efficient manner. Full-waveform inversion (FWI) is one of several techniques that are being pursued to produce higher quality earth models. Recently, Krebs et al. (2009) have demonstrated the viability of this approach by combining FWI with phase encoding in order to efficiently produce accurate earth models. Although their synthetic experiments show accurate reconstructions when using time-domain finite-difference methods on structured grids, in practice complex geological structures and material properties may hinder the ability of these methods to accurately invert for these features. In this paper, we build on the FWI and phase-encoding ideas but make use of an unstructured, yet high-order accurate, numerical approach based upon the DG method. We explore the flexibility engendered by DG methods to improve the characterization of complex subsurface features through unstructured meshes and localized polynomial refinement.

Discontinuous Galerkin methods have been developed and utilized in many fields since their inception in the early 1970's (see *e.g.*, Reed and Hill 1973; Cockburn 1999; Cockburn et al. 2000; Hesthaven and Warburton 2008). In the past several years, DG methods have been applied to seismic modeling by Käser, Dumbser, and co-workers who have investigated several aspects of DG methods in relationship to forward seismic-wave propagation on unstructured meshes including: elastic wave propagation (Käser and Dumbser, 2006), numerical properties (Käser et al., 2008), viscoelastic attenuation (Käser et al., 2007), topography representation (Käser et al., 2008; de la Puente, 2008; Park and Antin, 2004), and  $p$ -adaptivity (Dumbser et al., 2007). DG methods in conjunction with inversion have also recently been explored in the frequency domain (Brossier et al., 2009). However, one feature, common to all

these prior studies is that they are limited to piecewise constant representations of medium parameters within each element. This significantly reduces their ability to accurately represent complex earth models. For example, even a simple linear variation of properties with depth must be represented with a staircase-like, elementwise, constant-layered approximation.

In this study, we investigate time-domain acoustic inversion with DG methods that are *not* artificially constrained to use piecewise constant material models. In this context, we investigate the advantages of both unstructured meshes and higher-order polynomial refinement. The large-scale nature of the earth-model inversion problem requires efficient algorithms and to meet those needs, we use adjoint-based gradient decent algorithms, simultaneous source encoding, and fully parallel, distributed-memory implementations. Results are presented for synthetic numerical experiments based on the Marmousi2 model (Martin et al., 2006).

## FORMULATION AND METHODOLOGY

The acoustic FWI problem can be formulated as a constrained optimization problem with a least-squares objective function

$$\min_{\beta \in \mathcal{B}} \frac{1}{2} \sum_{r=1}^{N_r} \int_0^T \int_{\Omega} \xi_r(\mathbf{x}) \left( p(\mathbf{x}, t) - \sum_{s=1}^{N_s} \omega_s \bar{p}(\mathbf{x}, t) \right)^2 d\mathbf{x} dt \quad (1)$$

subject to the acoustic wave equation

$$\beta \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{v} = \beta \phi \quad \text{in } \Omega \times (0, T] \quad (2a)$$

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \nabla p = \mathbf{0} \quad \text{in } \Omega \times (0, T] \quad (2b)$$

$$p(\mathbf{x}, 0) = 0 \quad \text{for } \mathbf{x} \in \Omega \quad (2c)$$

$$\mathbf{v}(\mathbf{x}, 0) = \mathbf{0} \quad \text{for } \mathbf{x} \in \Omega \quad (2d)$$

which is solved subject to appropriate boundary conditions. In these expressions,  $\beta = 1/(\rho c^2)$  is the compressibility;  $\mathcal{B}$  is the space of admissible media;  $\rho$  is the mass-density;  $c$  is the wave speed;  $\Omega$  is the computational domain;  $T$  is the time horizon;  $N_r$  is the number of receivers;  $N_s$  is the number of sources;  $\omega_s$  is the random phase encoding for source  $s$ ;  $\phi$  is encoded sum of all explosive pressure sources;  $\bar{p}$  is the measured pressure data; and  $\xi_r(\mathbf{x})$  is the spatial kernel for receiver  $r$ . The forms for  $\phi$ ,  $\omega_s$  and  $\xi_r$  are described in more detail below. The state variables are pressure,  $p(\mathbf{x}, t)$  and particle velocity,  $\mathbf{v}(\mathbf{x}, t)$  defined for  $(\mathbf{x}, t) \in \Omega \times [0, T]$ .

We have written the acoustic wave equations in first-order, flux form such that the material properties,  $\rho$  and  $\beta$  appear on the time derivative. In this way, the system can be readily seen to be a conservative, hyperbolic system where the conservation variables are the products  $\beta p$  and  $\rho \mathbf{v}$ . This allows the direct application of a standard conservative discontinuous Galerkin method with no artificial restrictions on the variation of media parameters.

## Unstructured DG for Seismic Inversion

The adjoint-state method (see, *e.g.*, Tarantola 2005) is used to solve the inversion problem, by forming a Lagrangian functional that combines the objective function and the constraint (state equations) multiplied by adjoint variables. The optimality conditions are derived by taking variations with respect to the state, adjoint and inversion variables and setting the resulting equations to zero. This produces the original state equations, adjoint equations, and a nonlinear gradient equation along with adjoint boundary and end conditions. The adjoint equation is driven by a source term that comes from the linearized objective function but is otherwise identical in form to the original state equations for this self-adjoint system.

Our numerical solution strategy consists of a sequential approach in which the state equation is first solved and then the adjoint equation is integrated backwards starting from the adjoint end-condition. The adjoint solution is then used to solve the gradient equation and a gradient descent algorithm along with a line search is used to update the inversion parameters.

In principle our numerical implementation mimics the solution strategy described above but there are some differences that are worth noting. Instead of deriving the optimization conditions, then discretizing, and finally linearizing, we instead discretize, linearize and derive the optimality conditions. The latter approach is better suited to accommodate a DG discretization. We have verified our gradients against directional finite differences (both second and fourth order) using random direction vectors with agreement to double-precision machine accuracy.

Our DG spatial discretization is based on the work of Collis and co-workers (Collis, 2002a,b; Collis and Ghayour, 2003; Chen, 2004; Chen and Collis, 2004; Ramakrishnan and Collis, 2004; Ramakrishnan, 2005; Chen and Collis, 2008) and is a modal DG implementation in which numerical quadrature is used to accurately evaluate integrals in the resulting weak-form. This is to be contrasted with the work of Dumbser, Käser and co-workers who use a quadrature-free approach that is particularly attractive for linear, constant coefficient systems where exact integration can be used to improve computationally efficiency. Unfortunately, the quadrature-free approach, which is well-known in the the CFD community (Atkins and Shu, 1997), comes with several important restrictions: only simplicial meshes are allowed, curved elements are not supported, elements must have constant medium properties, and nonlinearities lead to aliasing or the need for spatial filters. While these constraints may at first seem daunting, there are important problems (linear wave propagation through element-wise homogeneous materials) for which these methods are quite useful. Nevertheless, it is the opinion of the authors that such an approach is not viable for seismic inversion where sub-element level variations in material properties are required – especially for the rather large element sizes that are ideally used in high-order DG for the wavefield variables.

Our DG implementation removes these restrictions thereby allowing hybrid meshes of quadrilateral and triangular elements, curved boundaries to more accurately capture topology, and high-order polynomial variations of material properties within each element. A companion submitted abstract (T. M. Smith, S. S. Collis, C. C. Ober, J. R. Overfelt, and H. F. Schwaiger,

personal communication, 2010) describes our variable media DG implementation in more detail in the context of isotropic, linear elasticity. For the acoustic equations presented here, the formulation is analogous except that differentiability of the medium properties within each element is not required since the equations are solved in conservation form. We recently reported simple verification studies for our DG implementation using manufactured solutions of the acoustic wave equation (Ober et al., 2009) and, subsequently, the method has been validated in both 2D and 3D against reference solutions and time-domain finite-difference codes for both acoustic and elastic physics. Due to space limitations, both the detailed formulation and verification studies are not reported here, but will be summarized in the associated presentation.

Our DG spatial discretization is used in conjunction with explicit time-stepping using a standard fourth-order Runge-Kutta algorithm for both the state and adjoint equations where care is taken to use a method that is self-adjoint. It should be noted that the optimality conditions of this material inversion problem are nonlinear and therefore a gradient-descent method coupled with a line-search algorithm (Brent, 1973) is used. A random phase-encoding scheme is implemented, equivalent to the one developed by Krebs et al. (2009) in which re-encoding is used on every iteration. Under these conditions, it is not clear whether one can reliably use a conjugate-gradient algorithm which requires a two-step recurrence. Therefore, we use a simple steepest-descent approach here and future work will explore the use of more advanced optimization algorithms. Despite the possibly unattractive convergence properties of steepest descent, the benefits of phase encoding (and thereby simultaneously inverting multiple shots), far outweighs the benefit of sequentially inverting for each source.

In this study, an explosion source is used of the form

$$\phi(\mathbf{x}, t) = \sum_{s=1}^{N_s} \omega_s w(t) \xi_s(\mathbf{x}) \quad (3)$$

where  $\omega_s \in \{-1, 1\}$  is the random encoding scalar for source  $s$  and  $w(t)$  is a Ricker wavelet defined by

$$w(t) = (1 - 2\pi^2 f_p^2 (t - t_0)) \exp(-\pi^2 f_p^2 (t - t_0)^2) \quad (4)$$

and using  $f_p = 5$  Hz and  $t_0 = 0.3$  s. The spatial kernel is given by a Gaussian ball centered at  $\mathbf{x}_s$  with standard deviation,  $\sigma$ ,

$$\xi_s(\mathbf{x}) = \left( \frac{1}{\sigma\sqrt{2\pi}} \right)^N \exp\left( -\frac{|\mathbf{x} - \mathbf{x}_s|^2}{2\sigma^2} \right) \quad (5)$$

where  $N$  is the number of space-dimensions ( $N = 2$  for the studies presented here). Receivers are also modeled using the Gaussian kernel for  $\xi_r(\mathbf{x})$  and  $\sigma = 30$  m for the cases presented here. Note that we could also have used a kernel of the form  $\xi_s(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_s)$  which would closely match the formulation used by Krebs et al. (2009). However, Dirac delta functions introduce a fundamentally unresolvable feature within any discretized numerical implementation and we chose to make sure that our problem setup was fully resolvable for our initial studies. Future studies will explore point sources and receivers.

## Unstructured DG for Seismic Inversion

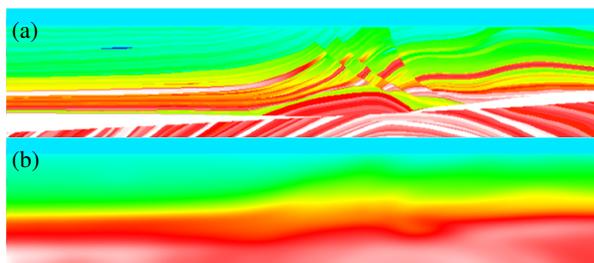


Figure 1: Marmousi2 earth model: (a) the true model with 20 m sampling and (b) the smoothed initial model.

## RESULTS

### Structured Mesh Inversion

We consider an inversion problem based on the Marmousi2 model (Martin et al., 2006), down sampled to 20 m over the region  $0 \leq x \leq 16000$  m, and  $0 \leq y \leq 3500$  m as shown in Figure 1(a). The model has been padded on the bottom by 500 m, not shown, so that a sponge-type (Grosch and Orszag, 1977) non-reflecting boundary condition does not affect the lower portion of the model. This padded region simply duplicates values of the bottom of the model and produces a total model depth of 4000 m. The sources are uniformly spaced at  $(x_s = s * 1000$  m, 300 m) where  $1 \leq s \leq (N_s = 15)$ , with simultaneous random phase-encoding. The receivers are uniformly spaced at  $(x_r = r * 200 + 500$  m, 100 m) where  $0 \leq r \leq (N_r = 75)$ .

First consider a structured mesh composed of quadrilaterals mimicking a standard finite-difference grid but with cell sizes of 200 m. With 1,600 elements ( $80 \times 20$ ) and a polynomial order of five, a total of  $80 \times 20 \times (5 + 1)^2 = 57,600$  degrees of freedom are used for each field variable, relative to  $800 \times 200 = 160,000$  for finite-difference at 20 m resolution. The greater resolving power of DG affords a significant reduction in degrees of freedom, both in the wavefield and in model parameters. A free-surface boundary condition is enforced on the ocean surface and a sponge-type, non-reflecting boundary treatment along with first-order characteristic-based non-reflecting boundary conditions are imposed on the sides and bottom of the computational domain.

To obtain an initial model, the true model, shown in Figure 1(a), was smoothed using a damped least-squares method where the smoothing operator at the top surface was 500 m (vertically) by 1000 m (horizontally) and linearly increased by a factor of four towards the bottom of the model. The initial model is shown in Figure 1(b) and is similar to that used by Krebs et al. (2009).

One method to reduce the computational costs is to start inversion with a short-time horizon,  $T$ , to initially invert for the shallow layers. This shortened time horizon reduces the cost of the forward modeling and subsequently the inversion algorithm. Once the shallow layers are fairly well resolved (*i.e.*, the reduction in the model-fit slows), one can increase  $T$  to include additional data and then continue the inversion to obtain additional details at depth.

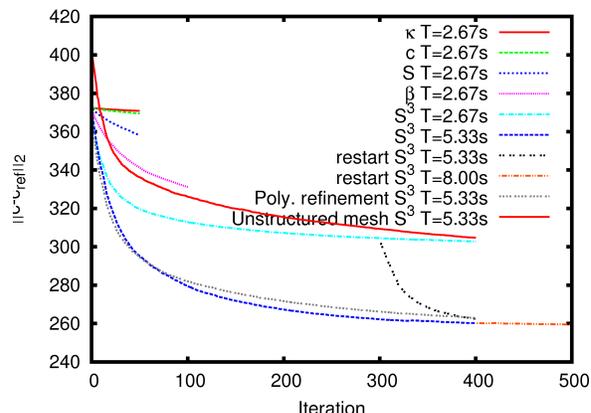


Figure 2: Influence of parameterization on model-fit convergence in terms of  $c$  in units of m/s (first five curves). Model-fit for different time horizons using inversion parameter  $S^3$  (fifth through eighth curves). Comparison of the polynomial refinement and the unstructured mesh (last two curves).

We start with a time horizon of  $T = 2.67$  s and begin by investigating the influence of parameterization on the rate of convergence. One such set of parameterizations are powers of the wave speed,  $c$ . As shown in Figure 2, there is a clear trend that negative powers of  $c$  lead to faster convergence in the model-fit. The slowest convergence is observed when using bulk modulus,  $\kappa = \rho c^2$ . Inverting directly for  $c$  leads to a slight improvement in convergence. But, successive negative powers of  $c$ , moving from slowness,  $S = 1/c$ , to compressibility,  $\beta = 1/(\rho c^2)$  to  $S^3 = 1/c^3$  each result in increased rates of convergence in model-fits. While this suggests that  $S^3$  may be a particularly effective parameterization for acoustic inversion of this model, additional investigations are required to determine the generality of this result.

Figure 2 also shows the effect of increasing the time-horizon from  $T = 2.67$  s to  $T = 5.33$  s. With more trace data, the  $T = 5.33$  s case is better able to reduce the model-fit, however this is with additional cost. At 300 iterations, the inverted model from the  $T = 2.67$  s time horizon was used as the initial model to restart the  $T = 5.33$  s time horizon run. As seen in Figure 2, this restart method quickly approaches the original  $T = 5.33$  s time-horizon curve but at a computational savings of 23%. Further savings could be achieved if additional restarts had been employed with intermediate time-horizons. Inversions were also performed using  $T = 8$  s but no additional improvement in model-fit was observed (see Figure 2).

The projected and predicted models are shown in Figure 3(a.1) and (a.2), respectively, and are in good agreement. It should be noted that the projected model is the best one could expect as it shows what can be represented by the DG method using a polynomial order of five on 200 m elements.

### Inversion with Local Polynomial Refinement

In Figure 3(a) there are two areas of interest denoted with black arrows. One is near a target region in the upper left quadrant and the other is in the anticline region. Because of the thin

## Unstructured DG for Seismic Inversion

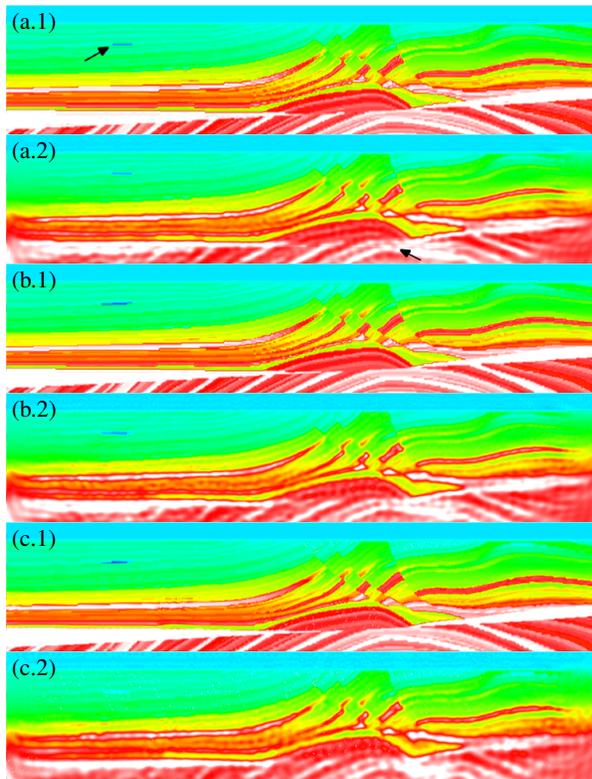


Figure 3: Acoustic wave-speed for both (1) projected and (2) predicted models using  $S^3$  inversion parameter at 400 iterations and  $T = 5.33$  s: (a) uniform structured mesh; (b) local  $p$ -refinement; (c) unstructured mesh at 215 iterations.

structure of the target, the media is not well represented on the uniform mesh with 200 m elements and  $p = 5$ . Likewise, in the anticline region, inter-element jump can be seen in the predicted model (*i.e.*, a slight discontinuous horizontal behavior). These types of DG jumps at mesh boundaries can be used as error-indicators that point to regions requiring mesh refinement. Our future work will explore automated solution adaptive inversion using jumps as error indicators. Here, we do a proof-of-principle study by locally increasing the polynomial order from  $p = 5$  to 8 in these two areas and re-running the inversion. Figure 3(b.1) shows that the projection of the truth solution is indeed improved in these two regions (*e.g.*, the magnitude and size of the target region). Figure 3(b.2) shows the predicted model on the  $p$ -refined mesh with improved results around the target region and removal of the inter-element jumps near the anticline. The model-fit with local polynomial refinement is shown in Figure 2 and is similar to that of the structured-mesh inversion confirming that a global  $L_2$  measure is insensitive to these local improvements.

### Unstructured Mesh Inversion

Discontinuous-material interfaces, such as ocean bottom, salt structures, and faults, occur routinely in surveyed regions. Traditionally, these interfaces are smoothed to allow for meshing as well as compatibility with high-order finite-difference

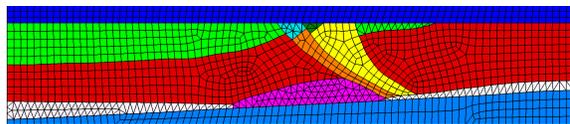


Figure 4: Unstructured mesh aligning with faults, salt flanks and selected layers.

methods. Using unstructured DG methods, these interfaces can be captured with the mesh to produce more accurate results (see Figure 4). We used the Cubit (Clark, 2010) meshing package to discretize the Marmousi2 model with a hybrid mesh of triangles and quadrilaterals. The layers, faults and salt structure were obtained through the original Marmousi2 model specification. In a real problem, this information would be iteratively determined as part of the inversion, but here it serves as a proof-of-principle for mesh-adaptive inversion.

Model-fit convergence is shown in Figure 2 with a trend similar to the structured mesh cases. However, the magnitude of the initial model-fit is higher than the structured mesh results. Since the unstructured mesh exactly captures several high-contrast geological features, it does a relatively better job of representing the truth model, Fig. 1(a), then the structured mesh while the smooth initial model, Fig. 1(b), is well-represented on all meshes thereby leading to a higher initial model miss-fit. The better representation of the projected truth model is clear when comparing Figure 3(c.1) to (b.1). Figure 3(c.2) shows the predicted model (after 215 iterations) for the unstructured mesh which is similar in quality to the structured mesh in most regions of the model. The target region is not as well represented since the mesh is relatively coarse there (see Figure 4). A careful examination of the fault regions suggest that the unstructured mesh is better able to represent these features for inversion but additional work is needed to fully exploit the potential of unstructured meshing.

## CONCLUSIONS

Time-domain inversion using discontinuous Galerkin on unstructured meshes and with local polynomial refinement is shown to better capture localized geological features and accurately capture discontinuous-material interfaces. These approaches provide the ability to surgically refine representations in order to improve predicted models for specific geological features. Our future work will entail automated extensions to directly incorporate local refinement and adaptive unstructured meshes within the inversion process.

## ACKNOWLEDGMENTS

This work has been supported by a Cooperative Research and Development Agreement (CRADA) between Sandia and an industrial partner. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

## Unstructured DG for Seismic Inversion

### REFERENCES

- Atkins, H. L., and C.-W. Shu, 1997, Quadrature-free implementation of discontinuous Galerkin methods for hyperbolic equations: *AIAA Journal*, **36**, 775–782.
- Brent, R., 1973, Algorithms for minimization without derivatives: Englewood Cliffs, N.J.
- Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: *Geophysics*, **74**, WCC105–WCC118.
- Chen, G., 2004, Multimodel simulation for optimal control of aeroacoustics: PhD thesis, Rice University, Mechanical Engineering and Materials Science.
- Chen, G., and S. S. Collis, 2004, Toward optimal control of aeroacoustic flows using discontinuous Galerkin discretizations: *AIAA Paper* 2004-0364.
- , 2008, Discontinuous Galerkin multi-model methods for optimal control of aeroacoustics: *AIAA Journal*, **46**, 1890–1899.
- Clark, B. W., 2010, Cubit website: URL <http://cubit.sandia.gov>.
- Cockburn, B., ed., 1999, Discontinuous Galerkin methods for convection-dominated problems, *in* High-order methods for computational applications, Lecture Notes in Computational Science and Engineering: Springer, 69–224.
- Cockburn, B., G. Karniadakis, and C.-W. Shu, eds., 2000, Discontinuous Galerkin methods: Theory, computation, and applications: Springer.
- Collis, S. S., 2002a, The DG/VMS method for unified turbulence simulation: *AIAA Paper* 2002-3124.
- , 2002b, Discontinuous Galerkin methods for turbulence simulation: Proceedings of the 2002 Center for Turbulence Research Summer Program, 155–167.
- Collis, S. S., and K. Ghayour, 2003, Discontinuous Galerkin for compressible DNS: ASME Paper Number FEDSM2003-45632.
- de la Puente, J., 2008, Seismic wave simulation for complex rheologies on unstructured meshes: PhD thesis, Ludwig-Maximilians-Universität.
- Dumbser, M., M. Käser, and E. F. Toro, 2007, An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes -V. Local time stepping and p-adaptivity: *Geophysical Journal International*, **171**, 695–717.
- Grosch, C. E., and S. A. Orszag, 1977, Numerical solution of problems in unboundary regions: *Journal of Computational Physics*, **25**, 273–296.
- Hesthaven, J. S., and T. Warburton, 2008, Nodal discontinuous Galerkin methods: Algorithms, analysis and applications: Springer, volume **54** of Texts in Applied Mathematics.
- Käser, M., and M. Dumbser, 2006, An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes -I. The two-dimensional isotropic case with external source terms: *Geophysical Journal International*, **166**, 855–877.
- Käser, M., M. Dumbser, J. de la Puente, and H. Igel, 2007, An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes-III. Viscoelastic attenuation: *Geophysical Journal International*, **168**, 224–242.
- Käser, M., V. Hermann, and J. de la Puente, 2008, Quantitative accuracy analysis of the discontinuous Galerkin method for seismic wave propagation: *Geophysical Journal International*, **173**, 990–999.
- Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, 2009, Fast full wave seismic inversion using encoded sources: *Geophysics*, **74**, 177–188.
- Martin, G. S., R. Wiley, and K. J. Marfurt, 2006, Marmousi2: An Elastic Upgrade for Marmousi: *The Leading Edge*, **25**, 156–166.
- Ober, C. C., S. S. Collis, B. van Bloemen Waanders, and C. Marcinkovich, 2009, Method of manufactured solutions for the acoustic wave equation: SEG Technical Program Expanded Abstracts, **28**, 3615–3619.
- Park, S.-H., and N. Antin, 2004, A discontinuous Galerkin method for seismic soil-structure interaction analysis in the time domain: *Earthquake Engineering and Structural Dynamics*, **33**, 285–293.
- Ramakrishnan, S., 2005, Local variational multiscale method for turbulence simulation: PhD thesis, Rice University.
- Ramakrishnan, S., and S. S. Collis, 2004, Multiscale modeling for turbulence simulation in complex geometries: *AIAA Paper* 2004-0241.
- Reed, W. H., and T. R. Hill, 1973, Triangular mesh methods for the neutron transport equation: Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory.
- Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: SIAM.