Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2018-0235

Vol. **27**, No. 2, pp. 513-545
February 2020

A GPU-Accelerated Hybridizable Discontinuous Galerkin Method for Linear Elasticity

Maurice S. Fabien*

Department of Computational and Applied Mathematics, Rice University, 6100 Main MS-134, Houston, TX 77005, USA.

Received 13 September 2018; Accepted (in revised version) 6 January 2019

Abstract. We design and analyze an efficient GPU-accelerated hybridizable discontinuous Galerkin method for linear elasticity. Performance analysis of the method is done using the state-of-the-art Time-Accuracy-Size (TAS) spectrum. TAS is a new performance measure which takes into account the accuracy of the solution. Standard performance measures, like floating point operations or run-time, are not completely appropriate for gauging the performance of approximations of continuum mechanics problems, as they neglect the solutions accuracy. A standard roofline model demonstrates that our method is utilizing computational resources efficiently, and as such, significant speed ups over a serial implementation are obtained. By combining traditional performance measures and the novel time-accuracy measures [7] into our performance model, we are able to draw more complete conclusions about which discretizations are best suited for an application. Several numerical experiments validate and verify our numerical scheme.

AMS subject classifications: 65Y05, 74S05, 65N30, 65N55

Key words: GPU-acceleration, discontinuous Galerkin, hybridization, multigrid, performance analysis.

1 Introduction

Computer architecture is becoming more sophisticated at the node level, where individual core clock rates are reduced, but more cores are packed into a chipset; to the point that floating point performance is greatly eclipsing memory operations. Efficiently utilizing this type of computational hardware has already been shown to require different programming models and parallel computing paradigms. This trend has consequences for the planning and designing of next-generation high-performance computing software [19, 45]. With this in mind, we design and analyze a GPU-accelerated hybridizable discontinuous Galerkin (HDG) method for linear elasticity.

*Corresponding author. *Email address:* fabien@rice.edu (M. S. Fabien)

HDG methods have several attractive properties, especially for problems where the solution of linear systems is required. As the HDG method is a discontinuous finite element method, it retains many of the features of discontinuous Galerkin (DG) methods that are celebrated, like local conservation, arbitrary order approximations, rigorous mathematical foundation, *hp*-adaptivity, admits unstructured meshes, and so on [12]. One of the important properties of the HDG method that differentiates it from classical DG methods is its ability to use static condensation. This well-known technique introduces additional unknowns on the mesh skeleton, and due to a judicious choice of numerical trace and flux, the original unknowns can be eliminated in an element-by-element fashion [13]. The total number of unknowns for the HDG method is thus reduced by a large amount for higher orders when compared to standard DG methods. As a result, drastic savings for memory storage and computational time are feasible [33].

For the linear elasticity equations, several HDG methods have been proposed, e.g., [15, 17, 17, 25, 53, 57, 58]. These discretizations have variations in how the primal and mixed variables are treated. We follow the discretization outlined in [43], as it is locking-free (for any $k \geq 0$), and is easily extendible to other equations, like Stokes and nonlinear elasticity. Related discretizations, like face centred finite volumes [56], the weak Galerkin method [9, 62], and the hybrid high order method [14, 20] have also been studied. Relevant applications can be found in [16, 34, 35].

The HDG method for linear elasticity presented in [43] has all approximate variables (displacement, gradient of displacement, and hydrostatic pressure) converge at the optimal rate of $k+1$ in the L^2-norm for polynomials of degree $k \geq 0$. There are a few interesting consequences of the gradient of displacement converging at the optimal rate. Quantities of engineering interest, like vorticity, stress, and strain also all converge at the rate of $k+1$ in the L^2-norm [43]. In addition, there exist local postprocessing schemes for the displacement variable which result in a new displacement approximation that superconverges at the rate of $k+2$ in the L^2-norm [44]. As the postprocessing is performed in an element-by-element manner, it is much cheaper than solving the full system at one polynomial order higher.

GPU-accelerated numerical methods for partial differential equations (PDEs) have received a great deal of attention. In particular, lattice Boltzmann and discontinuous Galerkin finite element methods have been demonstrated to perform well for linear wave problems and hyperbolic conservation laws [10, 66, 67]. However, few many-core or even GPU-accelerated DG methods are considered for partial differential equations that are inherently implicit in their nature, like PDEs that have elliptic or parabolic characteristics. The dominant difficulty for this class of PDEs is that they require the solution of large sparse linear systems. Efficient sparse linear solvers for high order DG methods are in general bandwidth bound, and more sophisticated [24, 29]. Preconditioners that map well to many-core architectures (e.g. [1, 3]) may not provide the best performance due to their poor convergence rates [22].

A GPU-accelerated HDG method for the 2D Poisson problem is proposed in [37], which analyzes in detail an efficient HDG assembly based on batch processing. The lin-