ASGarD: Adaptive Sparse Grid Discretization

David L. Green[1], **Mark Cianciosa**[1], **Ed D'Azevedo**[1], **Wael Elwasif**[1], **Steven E. Hahn**[1], **Coleman J. Kendrick**[1], **Hao Lau**[1], **M. Graham Lopez**[1], **Adam McDaniel**[1,4], **B.Tyler McDaniel**[1,2], **Lin Mu**[1,3], **Timothy Younkin**[1], **Stefan Schanke**[1], **Hugo Brunie**[5,6,7], and **Nestor Demeure**[5,6]

1 Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA **2** University of Tennessee Knoxville, Knoxville, Tennessee, USA **3** University of Georgia, Athens, Georgia, USA **4** South Doyle High School, Knoxville, Tennessee, USA **5** National Energy Research Scientific Computing Center, Berkeley, California, USA **6** Lawrence Berkley Laboratory, Berkeley, California, USA **7** National Institute for Research in Digital Science and Technology, France

Summary

Many areas of science exhibit physical[1] processes which are described by high dimensional partial differential equations (PDEs), e.g., the 4D (Dorf et al., 2013), 5D (Candy et al., 2009) and 6D models (Juno et al., 2018) describing magnetized fusion plasmas, models describing quantum chemistry, or derivatives pricing (Bandrauk et al., 2007). In such problems, the so called "curse of dimensionality" whereby the number of degrees of freedom (or unknowns) required to be solved for scales as N^D where N is the number of grid points in any given dimension D. A simple, albeit naive, 6D example is demonstrated in the left panel of Figure 1. With $N = 1000$ grid points in each dimension, the memory required just to store the solution vector, not to mention forming the matrix required to advance such a system in time, would exceed an exabyte - and also the available memory on the largest of supercomputers available today. The right panel of Figure 1 demonstrates potential savings for a range of problem dimensionalities and grid resolution. While there are methods to simulate such high-dimensional systems, they are mostly based on Monte-Carlo methods (E et al., 2020) which rely on a statistical sampling such that the resulting solutions include noise. Since the noise in such methods can only be reduced at a rate proportional to $\sqrt{N_p}$ where N_p is the number of Monte-Carlo samples, there is a need for continuum, or grid / mesh based methods for high-dimensional problems which both do not suffer from noise and bypass the curse of dimensionality. Here we present a simulation framework which provides such a method using adaptive sparse grids (Pflüger et al., 2010).

The Adaptive Sparse Grid Discretization (ASGarD) code is a framework specifically targeted at solving high-dimensional PDEs using a combination of a Discontinuous-Galerkin Finite Element solver implemented atop an adaptive sparse grid basis. The adaptivity aspect allows for the sparsity of the basis to be adapted to the properties of the problem of interest, which facilitates retaining the advantages of sparse grids in cases where the standard sparse grid selection rule is not the best match. A prototype of the non-adaptive sparse-grid implementation was used to produce the results of (D'Azevedo et al., 2020) for 3D time-domain Maxwell's equations. The implementation utilizes both CPU and GPU resources, as well as being single and multi-node capable. Performance portability is achieved by casting the computational kernels as linear algebra operations and relying on vendor provided BLAS libraries. Several test problems are provided, including advection up to 6D with either explicit or implicit timestepping.

---

[1]This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The publisher acknowledges the US government license to provide public access under the DOE Public Access Plan (https://energy.gov/downloads/doe-public-access-plan).
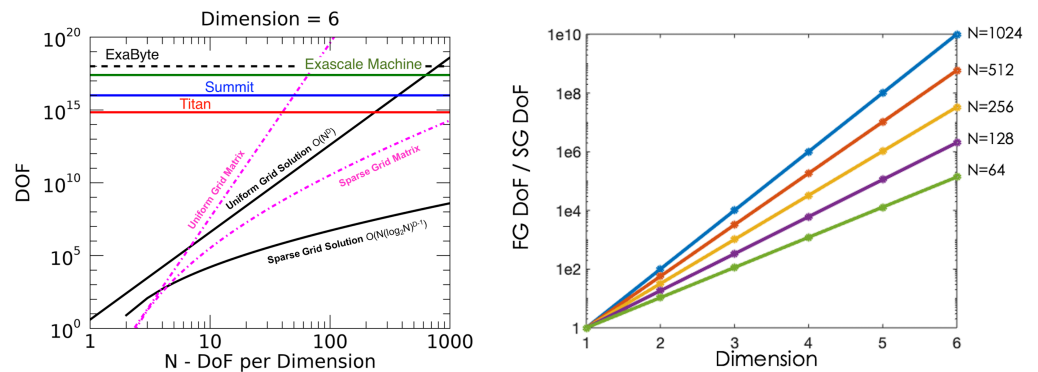
Figure 1: (left) Illustration of the curse of dimensionality in the context of solving a 6 dimensional PDE (e.g., those at the heart of magnetically confined fusion plasma physics) on modern supercomputers, and how the memory required to store the solution vector (solid black curves) and the matrix (magenta curves) in both naive and Sparse Grid based discretizations as the resolution of the simulation domain is varied. Memory limits of the Titan and Summit supercomputers at Oak Ridge National Laboratory, in addition to an approximate value for an ExaScale supercomputer, are overlaid for context; (right) Potential memory savings of a Sparse Grid based solver represented as the ratio of the naive tensor product full-grid (FG) degrees of freedom (DoF) to the sparse-grid (SG) DoF.

Acknowledgements

References

Bandrauk, A. D., Delfour, M. C., & Le Bris, C. (2007). *High-dimensional partial differential equations in science and engineering* (Vol. 41). American Mathematical Soc.

Candy, J., Holland, C., Waltz, R. E., Fahey, M. R., & Belli, E. (2009). Tokamak profile prediction using direct gyrokinetic and neoclassical simulation. *Physics of Plasmas*, *16*(6), 060704. https://doi.org/10.1063/1.3167820

D'Azevedo, E., Green, D. L., & Mu, L. (2020). Discontinuous galerkin sparse grids methods for time domain maxwell's equations. *Computer Physics Communications*, *256*, 107412. https://doi.org/https://doi.org/10.1016/j.cpc.2020.107412

Dorf, M. A., Cohen, R. H., Dorr, M., Rognlien, T., Hittinger, J., Compton, J., Colella, P., Martin, D., & McCorquodale, P. (2013). Simulation of neoclassical transport with the continuum gyrokinetic code COGENT. *Physics of Plasmas*, *20*(1), 012513. https://doi.org/10.1063/1.4776712

E, W., Han, J., & Jentzen, A. (2020). *Algorithms for solving high dimensional PDEs: From nonlinear monte carlo to machine learning*. http://arxiv.org/abs/2008.13333

Juno, J., Hakim, A., TenBarge, J., Shi, E., & Dorland, W. (2018). Discontinuous galerkin algorithms for fully kinetic plasmas. *Journal of Computational Physics*, *353*, 110–147. https://doi.org/10.1016/j.jcp.2017.10.009

66 Pflüger, D., Peherstorfer, B., & Bungartz, H.-J. (2010). Spatially adaptive sparse grids
67   for high-dimensional data-driven problems. *Journal of Complexity*, *26*(5), 508–522.
68   https://doi.org/https://doi.org/10.1016/j.jco.2010.04.001