RESEARCH STATEMENT

JED BROWN

Never has there been more intense demand for algorithmically optimal algorithms and software for extreme scale multiphysics simulation. This greater emphasis on algorithmic optimality is a consequence of increased hardware capability and correspondingly more ambitious simulations, for which suboptimal methods suffer greatly or fail completely. Multilevel methods, including multigrid and the fast multipole method, are an essential ingredient in any modern solution algorithm, both for algorithmic scalability and to limit and structure the necessary parallel communication. My research features the development of hardware-adapted, algorithmically optimal, multiphysics-capable, composable algorithms and numerical software.

In this statement, I elaborate on the common theme of high-performance integrative multiscale methods, explain my philosophy regarding software and reproducibility, and introduce several applications on which I am currently working and that are suitable for driving further targeted research.

1. Pervasive multiscale methods and fused analysis

Today's computational methods for analysis of systems governed by partial differential equations are founded on a series of abstractions that have been developed relatively independently. For example, a spatial discretization takes a PDE and produces an assembled operator, a linear solver takes an assembled sparse matrix and produces a solution, a nonlinear solver takes a nonlinear expression of a problem and transforms to a sequence of linear problems, an implicit time discretization takes a semi-discrete problem and yields a sequence of fully discrete problems, an optimization method views the PDE solver as a black box, most methods for uncertainty quantification do not look inside the solver, and adaptation of these methods to modern parallel computing architectures is done at a stage with little ability to modify the underlying algorithm. For each abstraction there are necessary and/or sufficient conditions for convergence, sometimes with quantitative prediction of convergence *rate* or at least asymptotic robustness to parameters. These interfaces were necessary to facilitate a division of labor, in terms of both theorems and software implementation, that has served the community well over the past few decades. On the other hand, these interfaces are *lossy* in the sense that they destroy structure that could greatly improve the efficiency and robustness of implementations. As a community, we have only recently developed a level of maturity where the entire stack can be reasoned about in a principled way. It's time to redesign the interfaces to emphasize matrix-free [Bro10] and nonlinear multilevel methods adapted to modern computing architectures, with in-situ uncertainty quantification, optimization, and stability analysis. In doing so, we must

Date: Fall 2012.

emphasize composability [BKM+12] so that the methods can reliably be combined for coupled multiphysics simulation [KMW+12].

2. Accessibility and open source software

As an early signatory of the Science Code Manifesto [1], I am deeply committed to responsible scientific software practices. The greatest impediment to widespread adoption of recently-developed methods is *accessibility*. A method presented only in journal papers is accessible only when it is conceptually simple and easily adaptable to different circumstances. In contrast, a method implemented as a library need not be internally simple, but must have robust, predictable features. Underlying my research is the principle that developers of computational methods should strive to create user-friendly methods complete with library implementations, and to work with applications to put the methods into production. Since all methods "break" eventually, the stress testing provided by production science and engineering applications is invaluable to truly understand the utility of methods and to identify the most important future research directions. Additionally, since the vast majority of algorithms cannot be completely defined within the space constraints of papers, it is crucial for reproducibility to release a reference implementation. Releasing code requires extra effort, but results in a higher quality product, fosters collaboration, and ultimately increases the impact of the research.

3. Selected projects and applications

3.1. **Low-communication multigrid.** I have been working with Mark Adams (Columbia University) to extend Brandt and Diskin's work [BD94] in eliminating communication in multigrid algorithms. These methods are unique in their ability to remove *all* communication from fine levels without significantly affecting the multigrid convergence rate. This remarkable result is a result of re-examining the coarse grid equation in the Full Approximation Scheme. Conventional multigrid cycles start and end on the fine grid, using the coarse grid for accelerating convergence of long-wavelength components of the error. A formally equivalent perspective interprets a multigrid cycle as starting and ending on the coarse grid, with fine grids "visited" only to provide local feedback in the form of the "τ correction" which quantifies the local difference between the coarse and fine grid operators. Side-effects of the τ formulation include compressed (and locally decompressible) checkpoints which are important for computation of transient adjoints, resilience in case of hardware errors, migration for dynamic load balancing, and in-situ visualization, as well as the ability to evaluate functionals of a solution using only polylogarithmic memory (no explicit storage of solution). This formulation exposes many opportunities in parallel implementations and on-node efficiency (e.g., fused fine-level visits operating entirely within cache) with potentially deep co-design consequences (e.g., fine levels can reside exclusively on accelerator devices, communicating with the host only via a coarse level). On the mathematical front, there is a fresh demand for sharp convergence analysis of compatible relaxation [Liv04] variants as well as a rigorous convergence theory for the τ formulation without full regularity assumptions. When combined with compatible Monte-Carlo [BR03], this

[1] http://sciencecodemanifesto.org

framework is sufficiently general to express many existing paradigms for multiscale models with stochastic micro/mesoscales (e.g., [WEL+07, Bra08, BL11, LBL12]).

3.2. **User-, solver-, and analysis-friendly discretizations.** An ideal discretization would be robust and generally applicable, highly accurate, entail minimal communication and memory movement, admit simple and rapidly-converging optimal solvers (like multigrid), inherit continuum structure such as local conservation and maximum principles/positivity, and commute with the adjoint operation (so that continuum and discrete adjoints coincide). Such a wonder-discretization is currently only available for the simplest model problems (and some non-existence theorems exist), but there is a nearly inexhaustible amount of research remaining to be done in extending the most attractive families to efficiently support the entire suite of desired capabilities. This research is most effectively done in collaboration with an application where the improvements can have an immediate impact.

3.3. **Scalable interactive in-situ eigenanalysis.** The objective of this recently-funded project is to develop methods and software with similar capability to EigTool [WT02], with ability to interactively explore the spectrum/pseudospectrum of an operator and plot selected eigenvalues even in case of large 3D multiphysics models, all transparently accessible within existing parallel applications that use PETSc for solvers. In addition to the usual scientific applications of eigenanalysis, convenient targeted eigenanalysis will be invaluable for designing multilevel and domain-decomposed solvers. For example, eigenanalysis applied to a filtered operator on a level of multigrid immediately identifies deficiencies in the smoother and grid transfer operators. In addition to scientific and methods research, I expect this tool to be of significant benefit for teaching about solver convergence.

3.4. **Lithosphere and mantle dynamics.** More so than in many scientific applications, solver performance is a bottleneck that limits inquiry into the process of subduction. Specifically, the solution of highly heterogeneous Stokes-like systems with strong material nonlinearities due to plasticity and phase change. Solution of these problems represents a critical bottleneck to high-resolution simulation of lithosphere dynamics. In collaboration with Dave May (ETH Zurich), Matt Knepley (University of Chicago), and Mark Adams (Columbia University), I am exploring robust adaptive coarsening and highly-parallel smoothers, with attention to fast updating due to evolving nonlinearities.

3.5. **Modular software for automotive CFD and heat transfer.** In collaboration with Zoltan Horvath (SZE, Hungary) and Matt Knepley (University of Chicago), I am designing open source software for use in air and air-oil cooling of DC motors and shape optimization of air intakes and wheel wells. Our objectives are threefold, (a) to create a test-bed for evaluating discretizations (including finite volume, spectral difference, and discontinuous Galerkin), structure-preserving IMEX time integration, and multigrid, (b) to make a practical engineering tool in which components like turbulence models, kinetic effects, and chemistry can be developed by engineers with limited software expertise, and (c) to be portable to emerging architectures and scale to very large problem sizes and computers. For efficiency without sacrificing modularity, we are developing a JIT for dynamically-configured models, and including the ability to write model components in Julia[2].

---

[2] `http://julialang.org`

We are also developing multiscale time integrators for quasi-time-periodic systems to efficiently handle time integration of motors over thousands of revolutions (necessary to ensure that a design can maintain safe operating temperatures).

3.6. **Solvers for PDE problems augmented with low-rank coupling.** What do lightweight inflatable robots and gyrokinetics have in common? They both consist of elliptic PDE systems augmented with low-rank coupling blocks. With robots, air pressure in various chambers couples the (sparse) fabric equations (collaboration with Geoffrey Irving at Otherlab, Inc.) while in gyrokinetics, flux surface averaging (arising from electron drift relative to ions) couples the poloidal potential equations (collaboration with Mark Adams (Columbia University) and the XGC1 team at Princeton). The low-rank blocks typically couple several thousand degrees of freedom, rendering explicit assembly unattractive. I am developing multilevel methods and software to solve problems of this type efficiently independent of the number and size of the low-rank blocks.

3.7. **IMEX time integration.** In collaboration with Emil Constantinescu, I designed and implemented an extensible suite of adaptive IMEX time integrators from the Rosenbrock-W and additive Runge-Kutta families within PETSc's `TS` package. These methods were formulated to work with the general nonlinear IMEX form $g(t, x, \dot{x}) = f(t, x)$ where g represents the stiff part requiring implicit treatment and f is non-stiff. Compared to the representations appearing in the literature, this form is more solver-friendly and more convenient for users with nonlinear transient terms such as arise with moving meshes. Schemes up to order 5 have been implemented and new methods can be defined by providing a coefficient tableau or design parameters within a known family. Most methods come with embedded error estimates, dense output formulas, and "hot starts" for iterative solvers. A unique attribute of my implementation is the extensible adaptive controller which allows applications with additional knowledge, especially regarding locations of model nonsmoothness, to control the time steps. Current research directions are in online operator-dependent method design and in exposing parallelism by simultaneously solving all stages of an implicit Runge-Kutta method within a single multigrid cycle.

References

[BD94] A. Brandt and B. Diskin, *Multigrid solvers on decomposed domains*, Contemporary Mathematics **157** (1994), 135–155.

[BKM+12] J. Brown, M. G. Knepley, D. A. May, L. C. McInnes, and B. F. Smith, *Composable linear solvers for multiphysics*, Proceeedings of the 11th International Symposium on Parallel and Distributed Computing (ISPDC 2012), IEEE Computer Society, 2012, pp. 55–62.

[BL11] I. Babuska and R. Lipton, *Optimal local approximation spaces for generalized finite element methods with application to multiscale problems*, Multiscale Modeling & Simulation **9** (2011), 373.

[BR03] A. Brandt and D. Ron, *Multigrid solvers and multilevel optimization strategies*, pp. 1–69, Kluwer Academic Publishers, 2003.

[Bra08] A. Brandt, *Principles of systematic upscaling*, Bridging the Scales in Science and Engineering (2008), 193–215.

[Bro10] Jed Brown, *Efficient nonlinear solvers for nodal high-order finite elements in 3D*, Journal of Scientific Computing **45** (2010), 48–63, 10.1007/s10915-010-9396-8.

[KMW+12] David E. Keyes, Lois Curfman McInnes, Carol Woodward, William Gropp, Eric Myra, Michael Pernice, John Bell, Jed Brown, Alain Clo, Jeffrey Connors, Emil Constantinescu, Don Estep, Kate Evans, Charbel Farhat, Ammar Hakim, Glenn Hammond, Glen Hansen, Judith Hill, Tobin Isaac, Xiangmin Jiao, Kirk Jordan, Dinesh Kaushik, Efthimios Kaxiras, Alice Koniges, Kihwan Lee, Aaron Lott, Qiming Lu, John Magerlein, Reed Maxwell, Michael McCourt, Miriam Mehl, Roger Pawlowski, Amanda Peters Randles, Daniel Reynolds, Beatrice Rivière, Ulrich Rüde, Tim Scheibe, John Shadid, Brendan Sheehan, Mark Shephard, Andrew Siegel, Barry Smith, Xianzhu Tang, Cian Wilson, and Barbara Wohlmuth, *Multiphysics Simulations: Challenges and Opportunities*, Tech. Report ANL/MCS-TM-321, Revision 1.1, Argonne National Laboratory, Oct 2012, To appear as a special issue of the *International Journal of High Performance Computing Applications*.

[LBL12] C. Le Bris and T. Lelièvre, *Micro-macro models for viscoelastic fluids: modelling, mathematics and numerics*, Science China Mathematics **55** (2012), no. 2, 353–384.

[Liv04] O.E. Livne, *Coarsening by compatible relaxation*, Numerical Linear Algebra with Applications **11** (2004), no. 2-3, 205–227.

[WEL+07] E. Weinan, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden, *Heterogeneous multiscale methods: a review*, Commun Comput Phys **2** (2007), no. 3, 367–450.

[WT02] T.G. Wright and LN Trefethen, *EigTool*, Software available at `http://www.comlab.ox.ac.uk/pseudospectra/eigtool`, 2002.