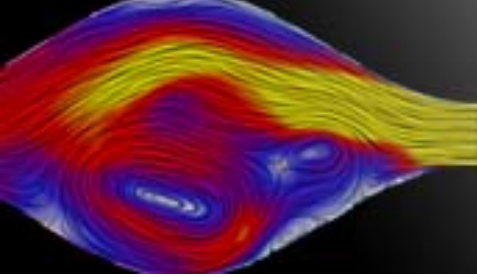


Sir James Lighthill Distinguished Lectureship in Mathematical Sciences

Tallahassee, Florida

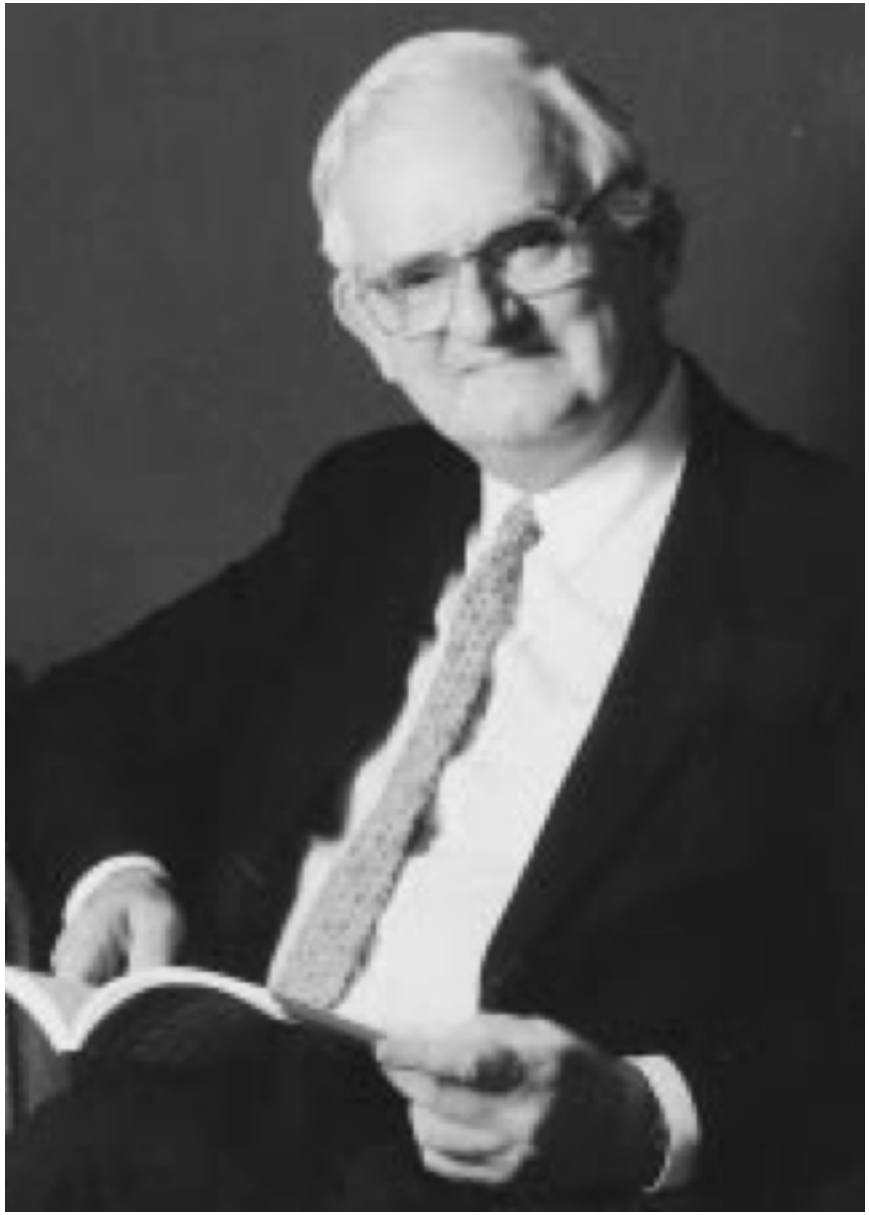
Applications that break techniques

David Keyes

Towards Optimal Petascale Simulations (TOPS), SciDAC Program, U.S. DOE

Mathematical and Computer Sciences & Engineering, KAUST
Applied Physics & Applied Mathematics, Columbia University

In Memoriam



23 Jan 1924 – 17 July 1998

"James Lighthill was acknowledged throughout the world as one of the great mathematical scientists of this century. He was the prototypical applied mathematician, immersing himself thoroughly in the essence and even the detail of every engineering, physical, or biological problem he was seeking to illuminate with mathematical description, formulating a sequence of clear mathematical problems and attacking them with a formidable range of techniques completely mastered, or adapted to the particular need, or newly created for the purpose, and then finally returning to the original problem with understanding, predictions, and advice for action."

(from the David Crighton memorial in *AMS Notices*)

Plan of series

- **Theme: role of mathematics in Computational Science & Engineering, specifically large-scale simulation**
- **Our philosophy will be to look at the scientific opportunity of large-scale simulation from three perspectives, concentrating one lecture on each**
  - **Applications, Architectures, Algorithms**
- **FSU Lighthill lectures are presumed neither cumulative nor exclusive**
  - **Individuals may attend any *one* without prerequisite**
  - **Individuals invited to attend all *three* (Engineering, Mathematics, Public)**
- **This requires a modicum of audience patience for either**
  - **Delegation (individual lectures not completely self-contained)**
  - **Repetition (lectures have some overlap)**

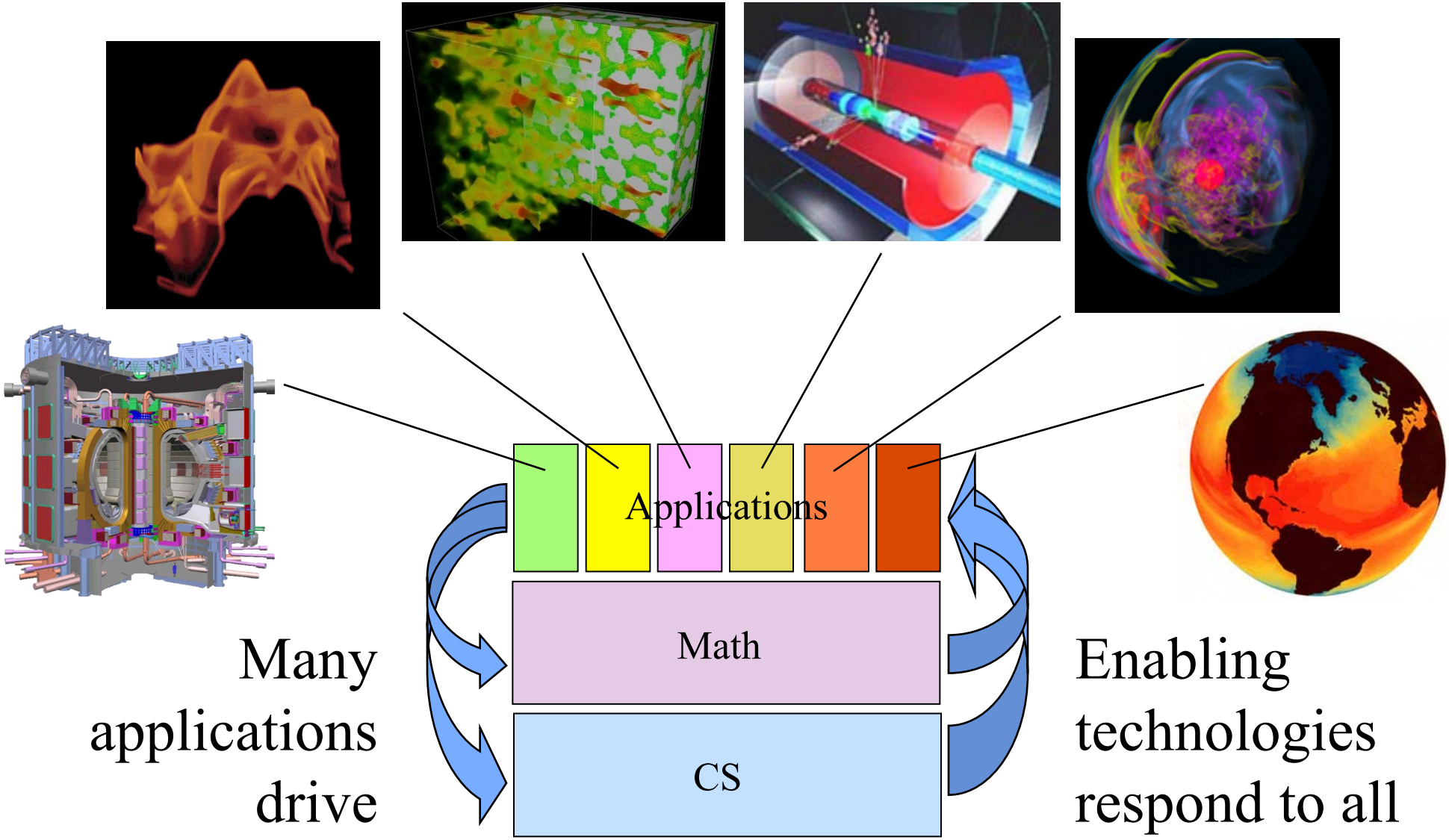
Purpose of the Engineering presentation

- **Expose the structure of a large multidisciplinary CS&E initiative**
  - **SciDAC**
  - **in its tenth year, hopefully to be continued by the 112th Congress**
- **Convey some of the fun of multidisciplinary collaborations between applications and mathematics**
- **Signal some specific topics for further discussion during the week**

Outline of the Engineering presentation

- Applied and computational mathematics in the U.S. Scientific Discovery through Advanced Computing program (SciDAC)
- The cornerstone of many large-scale simulations: the linear solver
- Applications that have "broken" standard solvers, and led to some advances (all remain "in progress")
  - Fusion (Off. of Fusion Energy Sciences)
  - Ice sheet fracture (Off. of Biological and Environmental Sciences)
  - Quantum chromodynamics (Off. of High Energy and Nuclear Physics)
  - Phase separation (Off. of Basic Energy Sciences)
- Summary and audience interaction

SciDAC philosophy: common cyberinfrastructure



Required cyberinfrastructure

Model-related

- Geometric modelers
- Meshers
- Discretizers
- Partitioners
- Solvers / integrators
- Adaptivity systems
- Random no. generators
- Subgridscale physics
- Uncertainty quantification
- Dynamic load balancing
- Graphs and combinatorial algs.
- Compression

Development-related

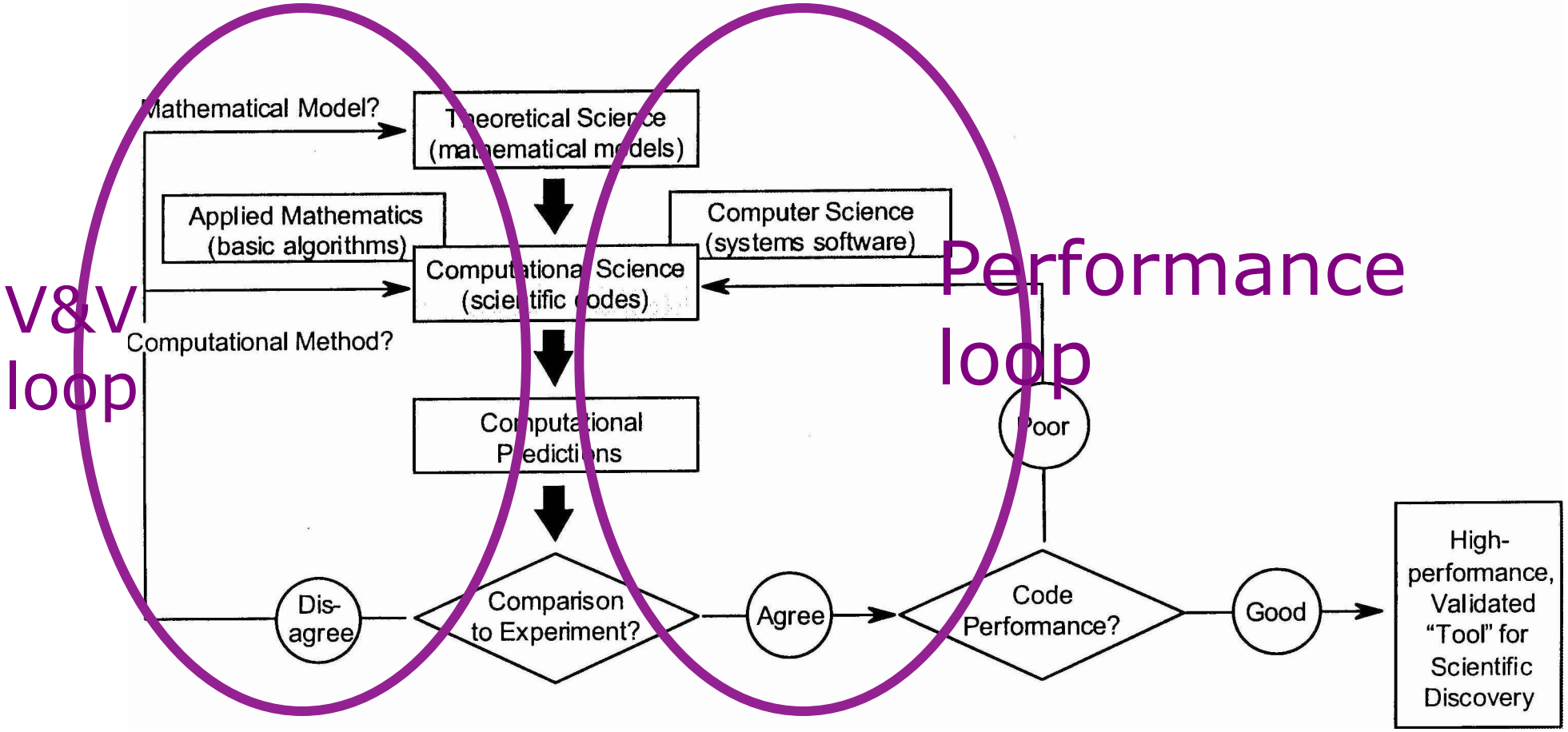
- Configuration systems
- Source-to-source translators
- Compilers
- Simulators
- Messaging systems
- Debuggers
- Profilers

Production-related

- Dynamic resource management
- Dynamic performance optimization
- Authenticators
- I/O systems
- Visualization systems
- Workflow controllers
- Frameworks
- Data miners
- Fault monitoring, reporting, and recovery

High-end computers come with little of this stuff. Most has to be contributed by the user community

Designing a simulation code – the diagram that launched the SciDAC program



c/o T. Dunning, 2000

SciDAC's four computational math centers

- **Interoperable Tools for Advanced Petascale Simulations (ITAPS)**

  PI: ***L. Freitag-Diachin, LLNL***

  For complex domain geometry

- **Algorithmic and Software Framework for Partial Differential Equations (APDEC)**

  PI: ***P. Colella, LBNL***

  For solution adaptivity

- **Combinatorial Scientific Computing and Petascale Simulation (CSCAPES)**

  PI: ***A. Pothen, Purdue U***

  For partitioning and ordering

- **Towards Optimal Petascale Simulations (TOPS)**

  PI: ***D. Keyes, Columbia U (since 2009: E. Ng, LBNL)***

  For scalable solution

See: www.scidac.gov/math/math.html

The TOPS center spans 4 labs and 5 universities

Our mission: Enable scientists and engineers to take full advantage of petascale hardware by overcoming the scalability bottlenecks traditional solvers impose, and assist them to move beyond "one-off" simulations to validation and optimization (~$32M/10 years)

Argonne NATIONAL LABORATORY

Lawrence Livermore National Laboratory

TOPS

Towards Optimal Petascale Simulations

BERKELEY LAB

Sandia National Laboratories

Columbia University

University of Colorado

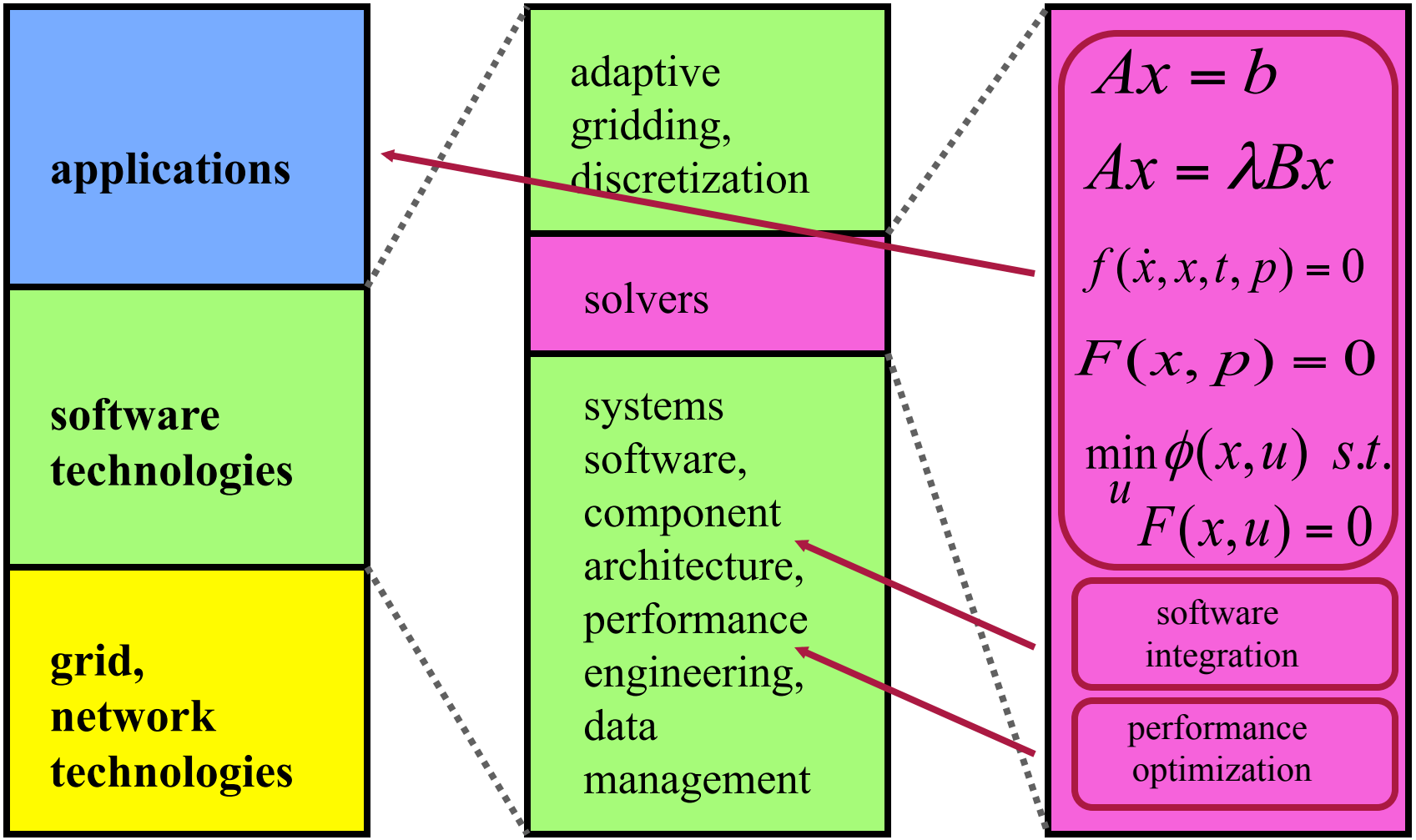
Berkeley University of California

The Institute for Computational Engineering and Sciences ICES

University of Texas

Southern Methodist University

SciDAC
Scientific Discovery through Advanced Computing



TOPS has built a toolchain of solver components that (increasingly) interoperate

- SciDAC project TOPS features these trusted packages, whose principal functions are keyed to the chart at the right:

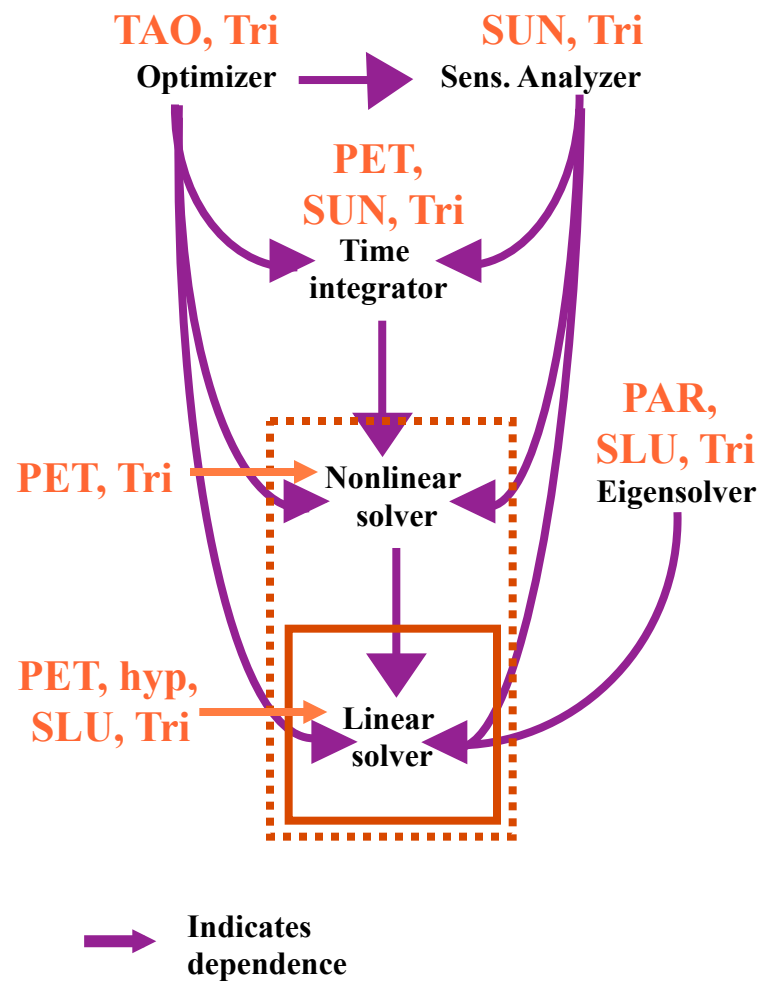
  hypre, PETSc,

  SUNDIALS, SuperLU,

  TAO, Trilinos, [PARPACK]

  These are in use and actively debugged in dozens of high-performance computing environments, in dozens of applications domains, by thousands of user groups around the world

- TOPS maintains about *half* of the software presented at DOE's ACTS toolkit tutorials



TOPS usage outside of SciDAC proper

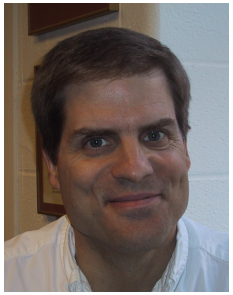
In articles, proceedings, theses:

- Astronomy
- Biomechanics
- Chemistry
- Climate
- Cognitive Sciences
- Combustion
- Economics
- Electrical Engineering
- Finance
- Geosciences
- Hydrodynamics
- Materials Science
- Mechanics
- Medical
- Micromechanics/Nanotechnology
- Numerical Analysis
- Optics
- Porous Media
- Shape Optimization

In widely distributed software:

- Cray LibSci®
- deal.II (2007 Wilkinson Prize)
- Dspice
- EMSolve
- FEMLAB®
- FIDAP®
- GlobalArrays
- HP Mathematical Library®
- IMSL ®
- libMesh
- Magpar
- Mathematica®
- NAG ®
- NIKE
- Prometheus
- SCIRun
- SciPy
- SLEPc
- Snark

Thousands of groups around the world use TOPS software without directly collaborating

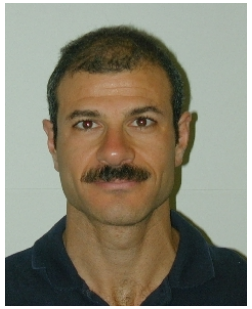

Adams


Baker


Cai


Demmel


Falgout


Ghattas


Heroux


Hu


Kaushik


Keyes


Knepley


Li


Manteuffel


McCormick


McInnes


Moré


Munson


Ng


Reynolds


Rouson


Salinger


Smith


Woodward


C. Yang


U. Yang

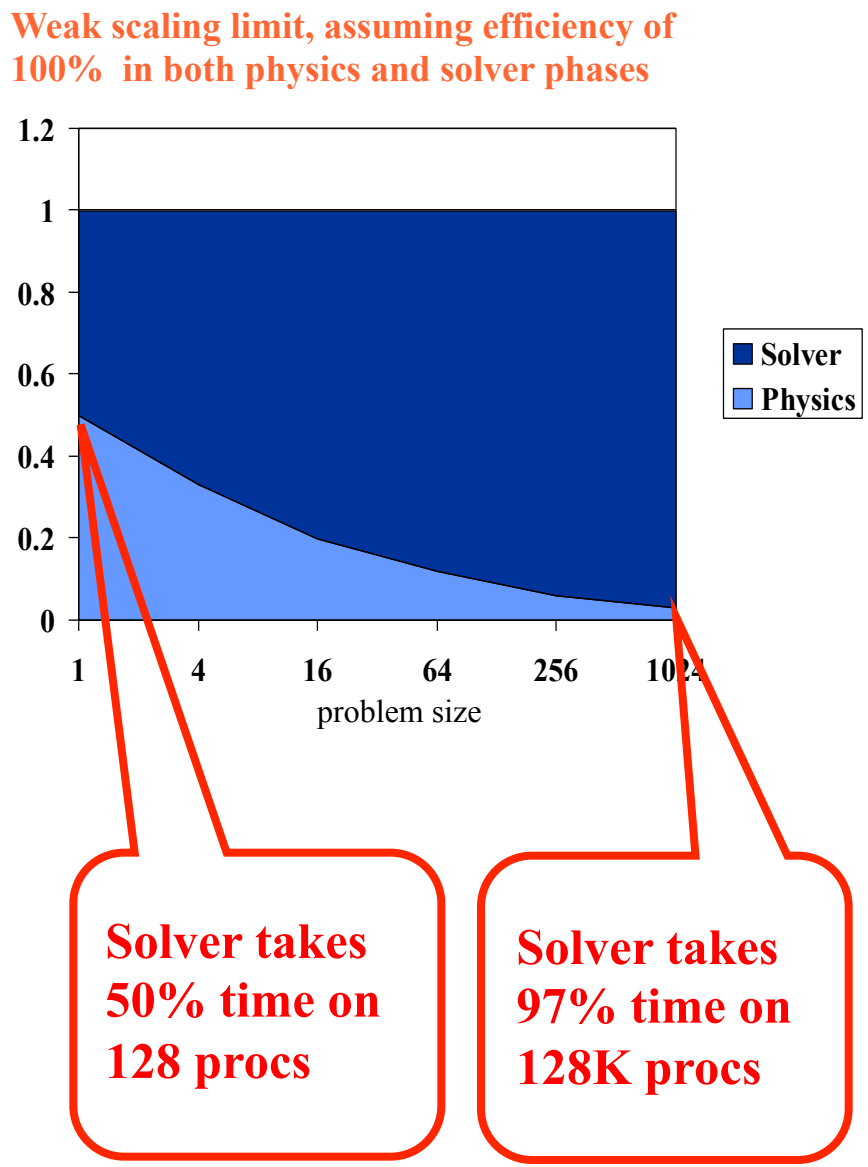

Zhang

Faces of TOPS – the coauthors of this presentation ☺

TOPS
Terascale Optimal PDE Simulations

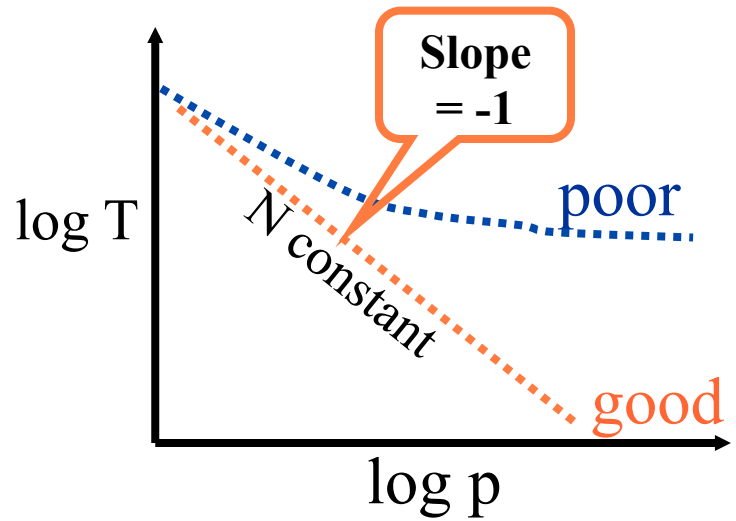
It's *all* about solvers at large scale

- **Given, for example:**
  - **a "physics" phase that scales as $O(N)$**
  - **a "solver" phase that scales as $O(N^{3/2})$**
  - **computation is almost all solver after several doublings**
- **Most applications groups have not yet "felt" the impact of this curve in their gut**
  - **as users actually get into queues with more than 4K processors, this will change**

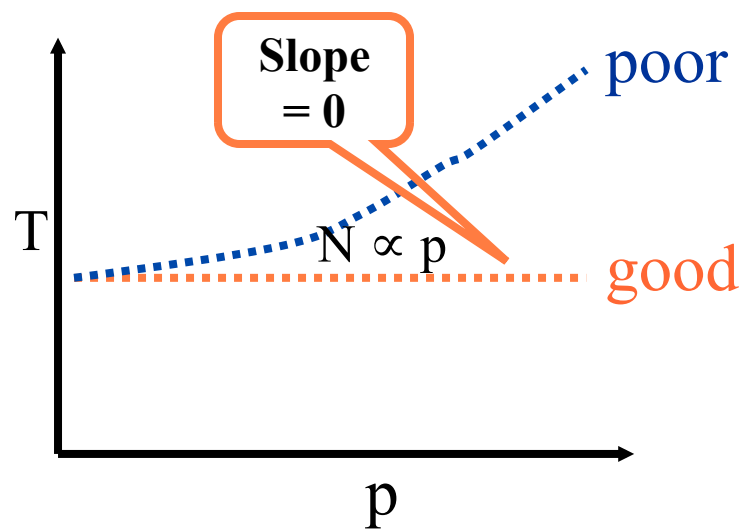


Review: two definitions of scalability

- **"Strong scaling"**
  - **execution time (*T*) decreases in inverse proportion to the number of processors (*p*)**
  - ***fixed size problem* (*N*) *overall***
  - **often instead graphed as reciprocal, "speedup"**



- **"Weak scaling" (memory bound)**
  - **execution time remains constant, as problem size and processor number are increased in proportion**
  - ***fixed size problem per processor***
  - **also known as "Gustafson scaling"**



Solvers *are* scaling: *hypre*'s algebraic multigrid (AMG) on BlueGene

- **Algebraic multigrid a key algorithmic technology**
  - Discrete operator defined for finest grid by the application, itself, *and* for many recursively derived levels with successively fewer degrees of freedom, for solver purposes only
  - Unlike geometric multigrid, AMG not restricted to problems with "natural" coarsenings derived from grid alone
- **Optimality (cost per cycle) intimately tied to the ability to coarsen aggressively**
- **Convergence scalability (number of cycles) and parallel efficiency also sensitive to rate of coarsening**
- **While much research and development remains, multigrid is practical at extreme concurrency**

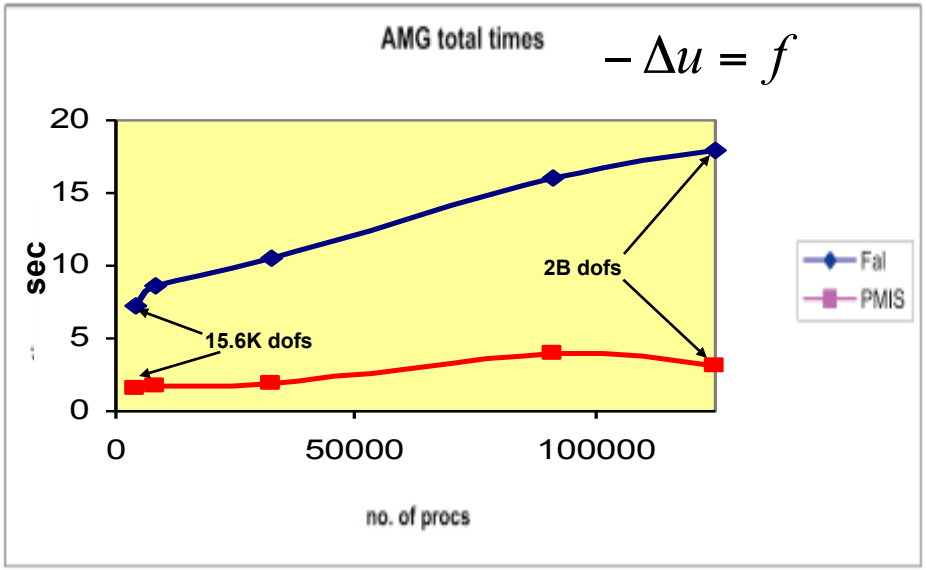


Figure shows weak scaling result for AMG out to 120K processors, with one 25×25×25block per processor (up to ~2B DOFs)

c/o U. M. Yang, LLNL

7-pt Laplacian, total execution time, AMG-CG, total problem size ~2 billion

Iterative correction: a generator of scalable algorithms

- **The most basic idea in iterative methods for** $Ax = b$

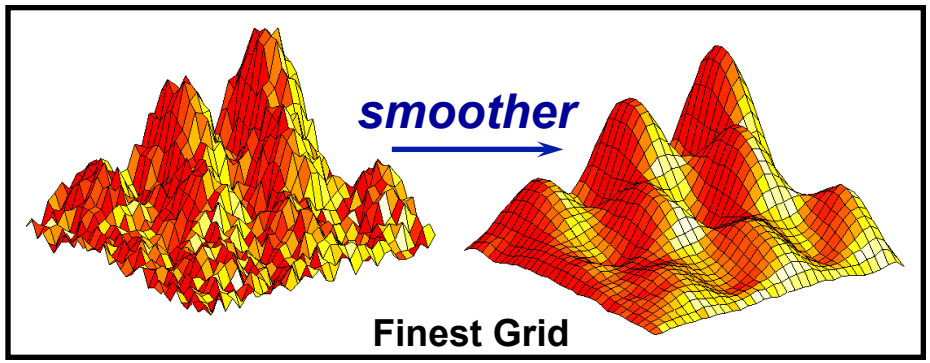
$$x \leftarrow x + B^{-1}(b - Ax)$$

- **Evaluate residual accurately, but solve approximately, where** B^{-1} **is an approximate inverse to** A
- **A sequence of complementary solves can be used, e.g., with** B_1 **first and then** B_2 **one has**

$$x \leftarrow x + [B_1^{-1} + B_2^{-1} - B_2^{-1} A B_1^{-1}](b - Ax)$$

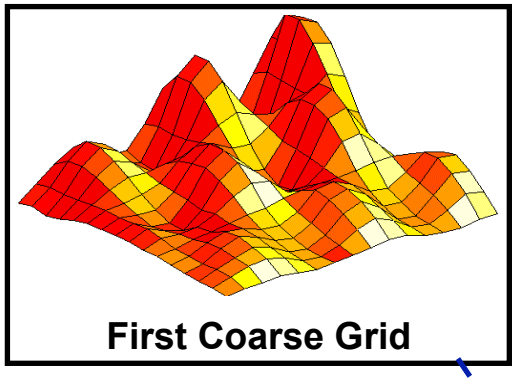
- **Scale recurrence, e.g., with** $B_2^{-1} = R^T (RAR^T)^{-1} R$ **, leads to *multilevel methods***
- **Characteristic choices of** R **lead to *domain decomposition***
- **Optimal polynomials of** $(B^{-1}A)$ **lead to various *preconditioned Krylov methods***

Multigrid treats each error component in an appropriate subspace



A Multigrid V-cycle

Restriction
transfer from fine to coarse grid



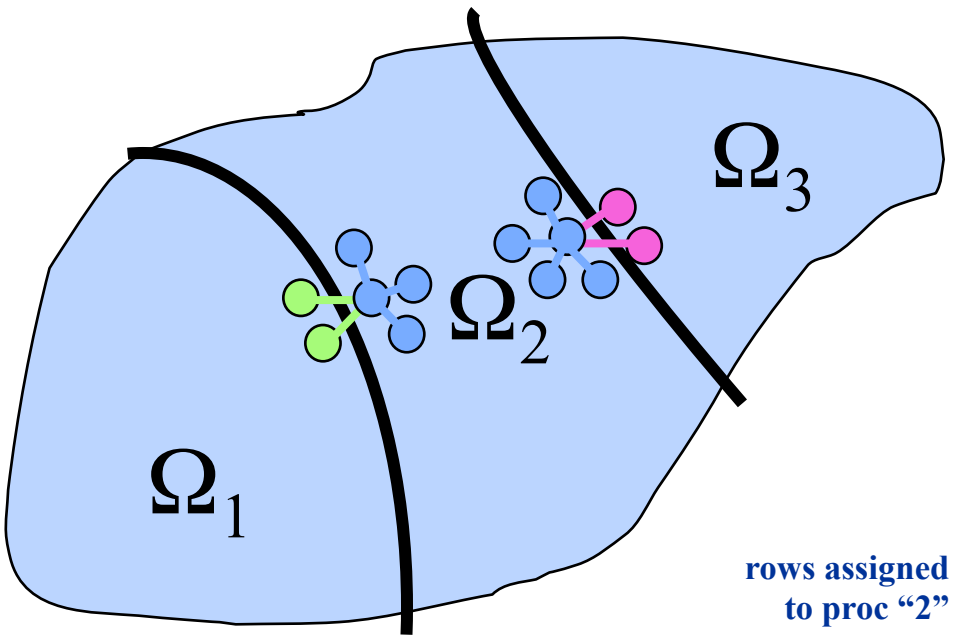
coarser grid has fewer cells (less work & storage)

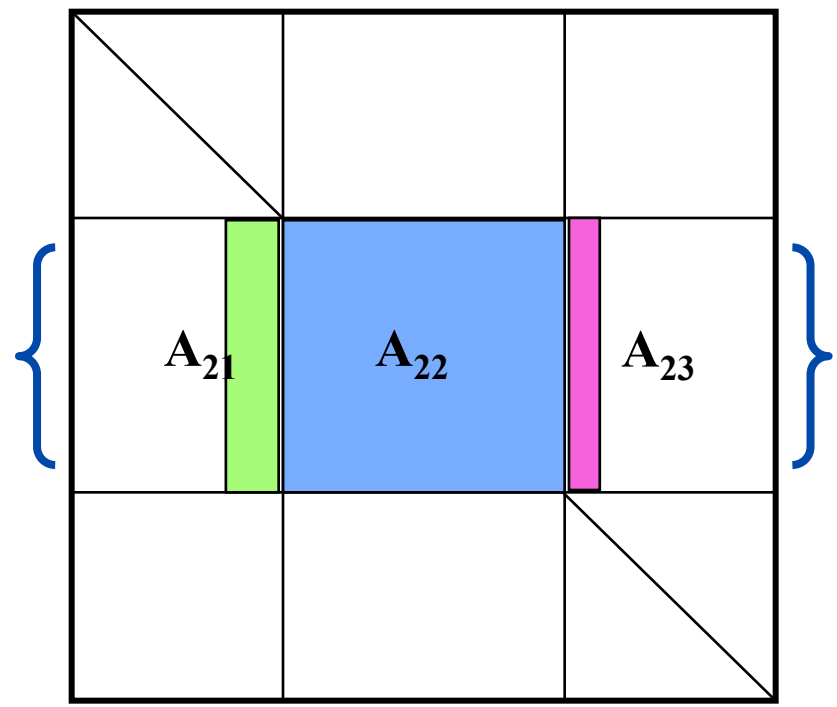
Prolongation
transfer from coarse to fine grid

Recursively apply this idea until we have an easy problem to solve

c/o R. Falgout, LLNL

Domain decomposition puts off limitation of Amdahl's Law in weak scaling



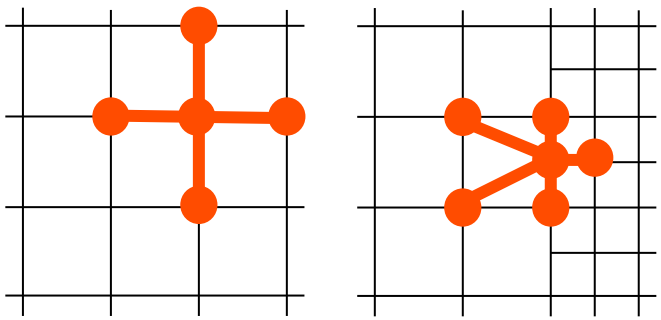


Partitioning of the grid induces block structure on the system matrix (Jacobian)

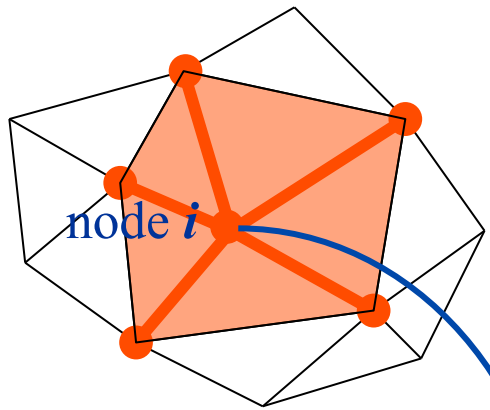
Computation scales with *area*; communication scales with *perimeter*; ratio *fixed* in weak scaling

DD relevant to any local stencil formulation

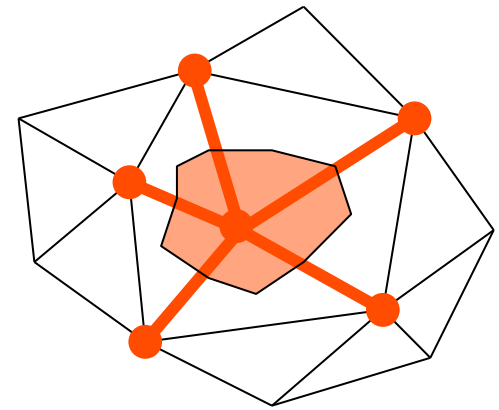
finite differences

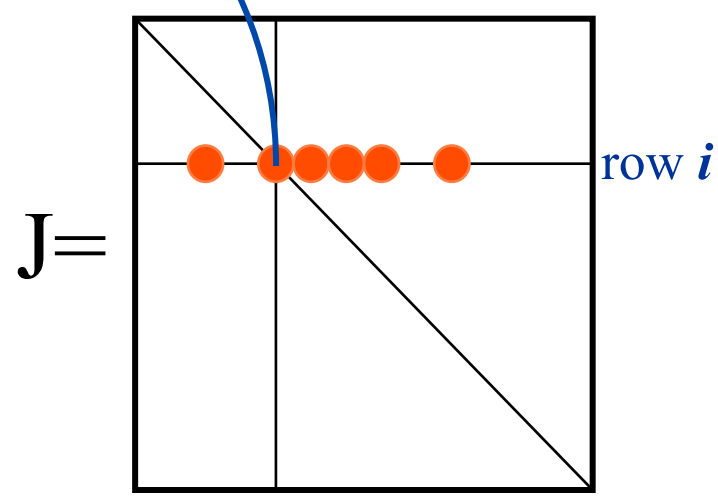


finite elements



finite volumes





- **lead to sparse Jacobian matrices**
- **however, the inverses are generally dense; even the factors suffer unacceptable fill-in in 3D**
- **want to solve in subdomains only, and use to precondition full sparse problem**

Krylov-Schwarz: a linear solver "workhorse"

$$Ax = b \qquad\qquad B^{-1}Ax = B^{-1}b$$

$$x = \underset{v \in V \equiv \{b, Ab, A^2 b, \cdots\}}{\arg\min} \{Av - b\} \qquad B^{-1} = \sum_i R_i^T (R_i A R_i^T)^{-1} R_i$$

Krylov

accelerator

spectrally adaptive

Schwarz

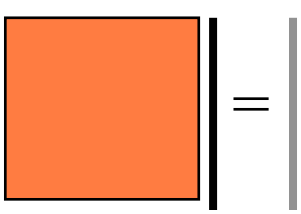
preconditioner

parallelizable

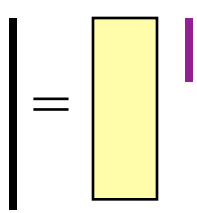
Krylov bases for sparse systems

- E.g., conjugate gradients (CG) for symmetric, positive definite systems, and generalized minimal residual (GMRES) for nonsymmetry or indefiniteness
- Krylov iteration is an algebraic projection method for converting a high-dimensional linear system into a lower-dimensional linear system

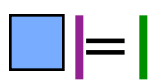
$$Ax = b$$



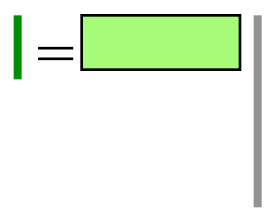
$$x = Vy$$



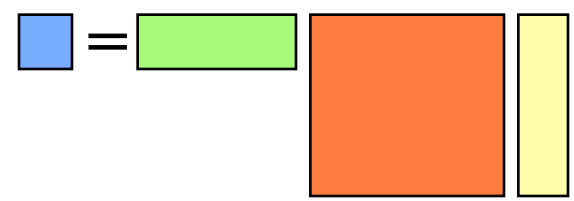
$$Hy = g$$



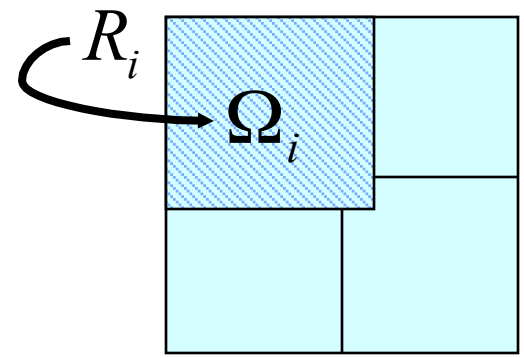
$$g = W^T b$$



$$H \equiv W^T A V$$



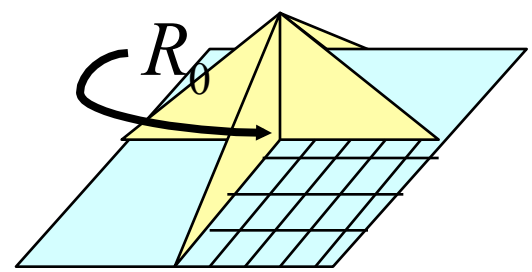
Schwarz domain decomposition method



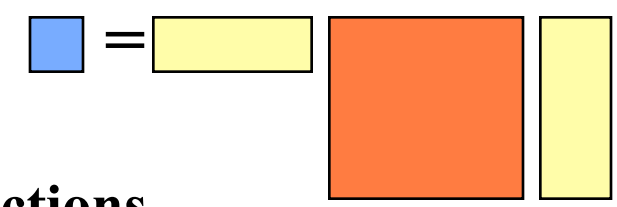
- **Consider restriction and extension operators for subdomains,** R_i, R_i^T, **and for possible coarse grid,** R_0, R_0^T
- **Replace discretized** $Au = f$ **with**

$$B^{-1}Au = B^{-1}f$$

$$B^{-1} = R_0^T A_0^{-1} R_0 + \sum_i R_i^T A_i^{-1} R_i$$



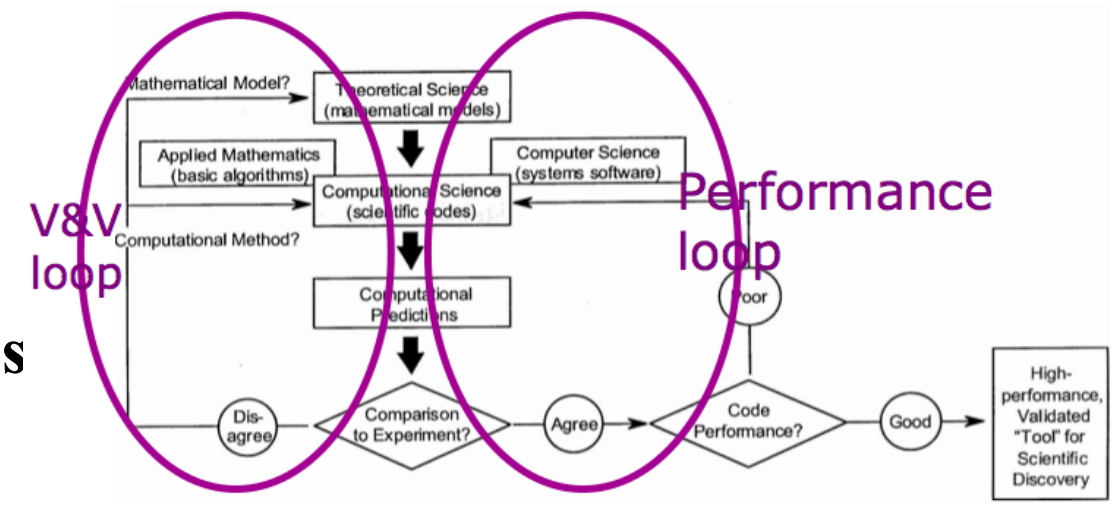
$$A_i = R_i A R_i^T$$



- **Solve by a Krylov method**
- **Matrix-vector multiplies with**
  - **parallelism on each subdomain**
  - **nearest-neighbor exchanges, global reductions**
  - **possible small global system (not needed for parabolic case)**

Remainder of the presentation

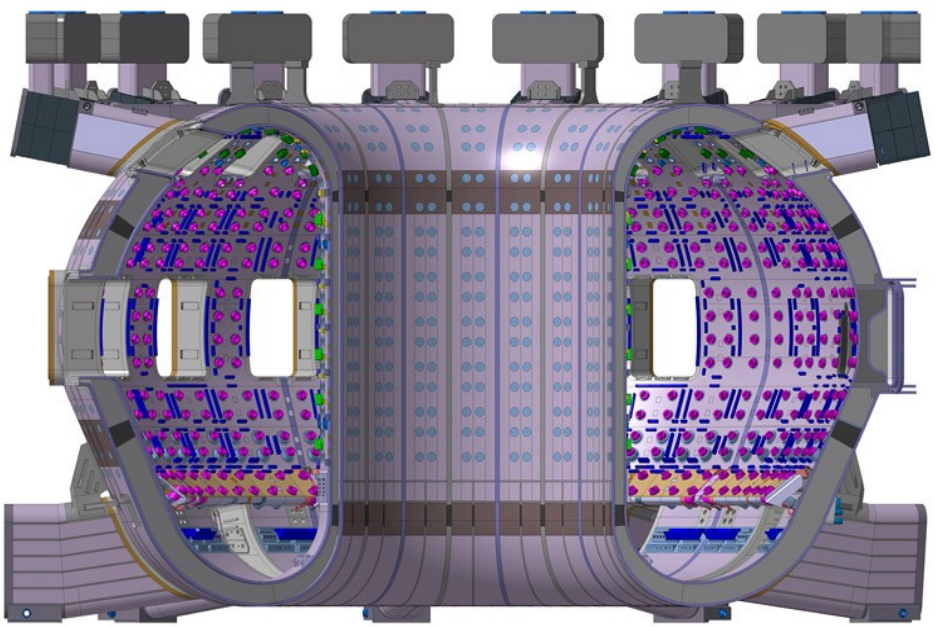
- **Four vignettes**
  - **Fusion**
  - **Ice sheet fracture**
  - **Quantum chromodynamics**
  - **Phase separation**



- **For each, a simple story**
  - **Application encounters limitations with existing solver**
  - **Interaction ensues between application scientist and TOPS computational mathematics group**
  - **Solutions are proposed, sometimes off-the-shelf, but usually in prolonged co-development**
  - **Application advances to next hurdle**
  - **Solutions get added to the infrastructure**

Application #1: MHD models of magnetically confined fusion

ITER, an $11B multinational project currently under construction in Cadaraches, France, aims to demonstrate magnetically confined fusion by 2020; photo at right shows tokamak pit at the far end of the construction site.

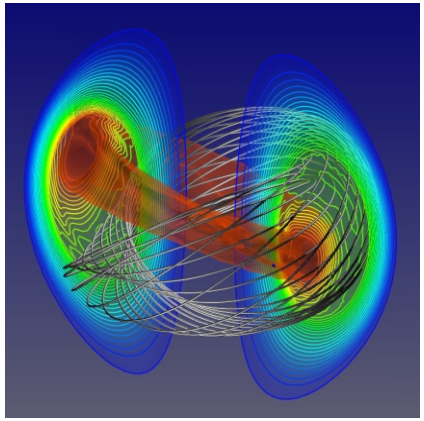


Contract for Vacuum Vessel signed 14 Oct 2010

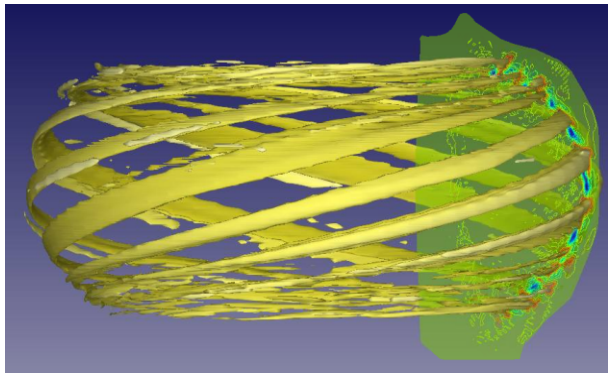
Top-to-bottom exascale computation is believed essential for efficient design and operation of large-scale experiments

- **Typical ITER discharge is estimated at $1M**
- **US will get so many "shots" per month**
- **Chief goal is to understand disruptions that could plague a practical power generating device**

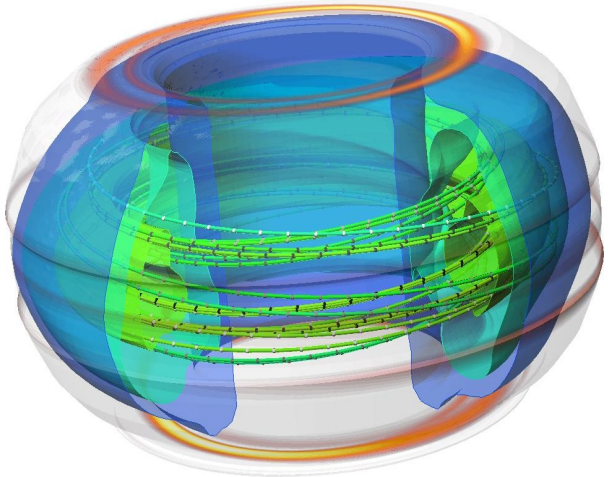
MHD codes predict onset of instabilities critical to ITER, and explore control scenarios



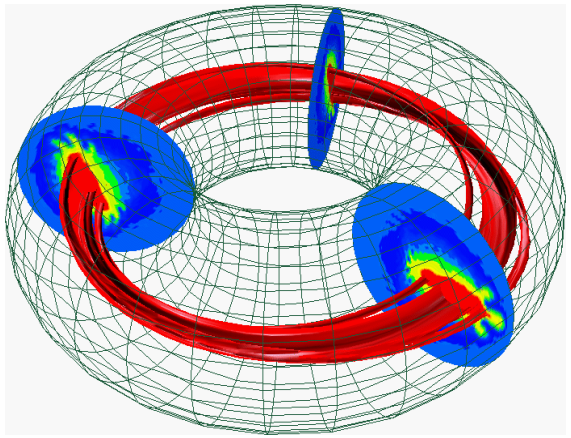
"sawtooth oscillations"



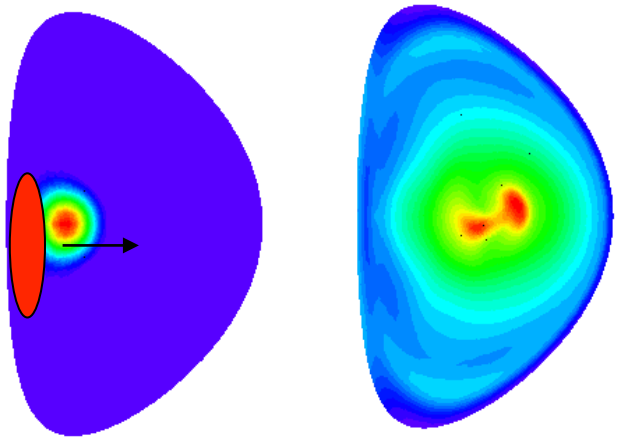
Edge Localized Modes



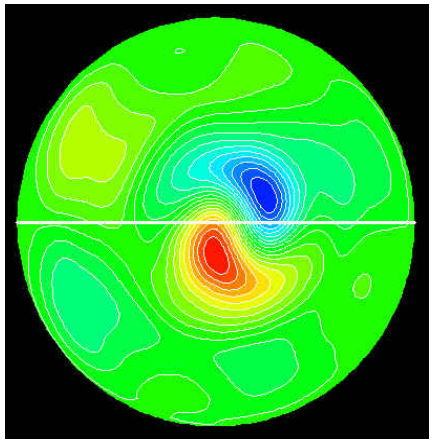
Heat loads during disruption



Disruptions caused by short wave-length modes interacting with helical structures.



Mass redistribution after pellet injection



Interaction of high-energy density particles with global modes

c/o A. Koniges, LBNL

WARNING: The following two slides are rated

E

for explicit equations.

No audience may look unless accompanied by a mathematician or engineer.

MHD: Maxwell coupled to Navier-Stokes

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} + \kappa$$

$$\mathbf{E} = -\mathbf{V} \times \mathbf{B} + \eta \mathbf{J}$$

$$\mu_0 \mathbf{J} = \nabla \times \mathbf{B}$$

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{V}) = 0$$

$$\rho\left(\frac{\partial \mathbf{V}}{\partial t} + \mathbf{V} \cdot \nabla \mathbf{V}\right) = \mathbf{J} \times \mathbf{B} - \nabla p + \nabla \cdot \nu \rho \nabla \mathbf{V}$$

$$\frac{n}{\gamma - 1}\left(\frac{\partial T}{\partial t} + \mathbf{V} \cdot \nabla T\right) = -p \nabla \cdot \mathbf{V} + \nabla \cdot n\left[\left(\chi_{\parallel} - \chi_{\perp}\right)\hat{\mathbf{b}}\hat{\mathbf{b}} + \chi_{\perp}\mathbf{I}\right] \cdot \nabla T + Q$$

Vector potential, **A**, gives the expansions

$$\mathbf{B} = \nabla \times \mathbf{A},$$

$$\mathbf{E} = -\nabla \varphi - \frac{\partial \mathbf{A}}{\partial t}$$

Scalar potentials, U, f, χ, and ψ, give the expansions (in (R, Φ, Z) coordinates):

$$\mathbf{A} = \Psi \nabla \phi + R \nabla f \times \nabla \phi$$

$$\mathbf{V} = R^2 \nabla U \times \nabla \phi + \nabla_{\perp} \chi + v_{\varphi} R \nabla \phi$$

MHD in scalar potential form

$$\frac{\partial Z}{\partial t} = -I\Delta^* \underline{I} - \Delta^* \underline{p} + \frac{\mu}{\rho}\nabla^2 \underline{Z} \ldots$$

$$\frac{\partial I}{\partial t} = -I\underline{Z} + \eta\Delta^* \underline{I} \ldots$$

$$\frac{\partial p}{\partial t} = -\gamma p \underline{Z} \ldots$$

$$\frac{\partial C}{\partial t} = \eta\Delta^* \underline{C} + \ldots$$

$$\frac{\partial W}{\partial t} = \frac{\mu}{\rho}\nabla^2 \underline{W} + \ldots$$

$$\frac{\partial v_\varphi}{\partial t} = \frac{\mu}{\rho}\nabla^2 \underline{v_\varphi} \ldots$$

$$\frac{\partial d}{\partial t} = \ldots$$

$$\Delta^* \chi = Z$$

$$\Delta^\dagger U = W$$

$$\nabla_\perp^2 \Phi = \ldots$$

$$\nabla_\perp^2 f = -I/R$$

$$\Delta^* \psi = C$$

Each Time Step:

3 coupled implicit time advance equations

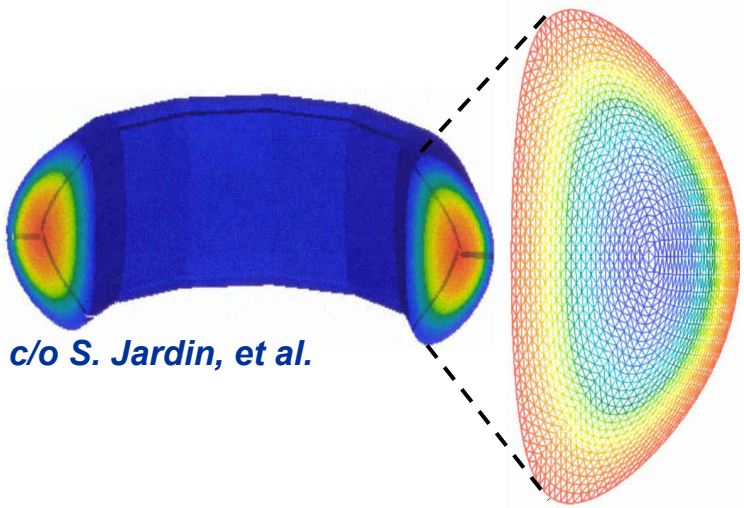
3 uncoupled implicit time advance equations

1 explicit time advance

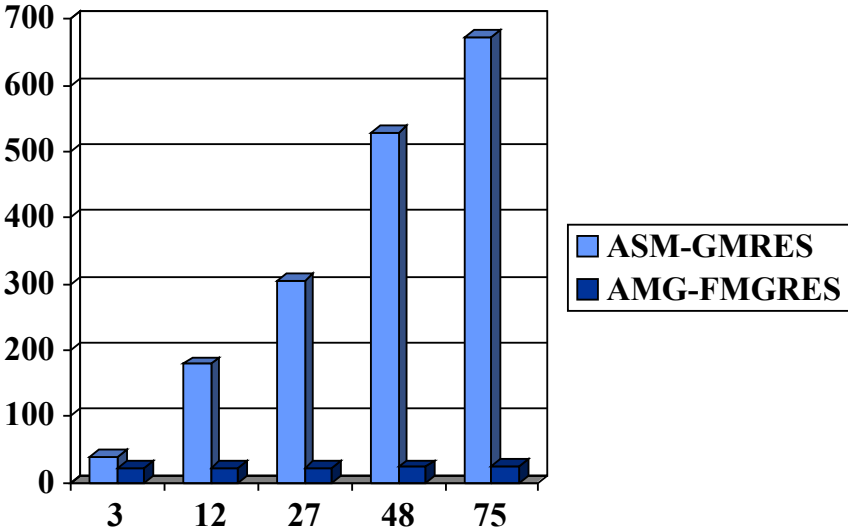
5 elliptic solves..but all 2D

M3D-C^0: multigrid for optimality

- **M3D code**
  - unstructured mesh, hybrid FE/FD discretization with C^0 elements in each poloidal crossplane
  - linear systems (>90% exe. time)



- **TOPS collaboration**
  - Replaced generic additive Schwarz (ASM) preconditioner with three different solvers tuned to coefficient structure, including algebraic multigrid (AMG) from hypre
  - achieved mesh-independent convergence rate
  - ~5× improvement in execution time



M3D-C^1 code development

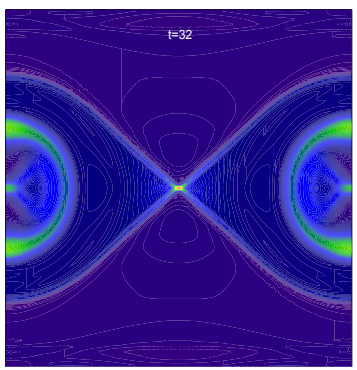
Existing M3D code needed to be upgraded	Numerous discussions: Math-CS-physics	Required new work in all areas
• low order accuracy not sufficient to resolve multi-scale phenomena (current sheets) • grid construction and adaptation not optimal for current sheets • time-step restriction too severe for slowly growing modes	• 3D C^1 high-order FE allows fully implicit compact system • make use of ITAPS mesh and adaptation libraries and data structure • use TOPS block-Jacobi preconditioner that recognizes tight point-block and looser interplane couplings	• Physics team completely recodes M3D using new higher-continuity, higher-order elements • ITAPS makes many extensions as needed for periodic torus • TOPS team adds new capabilities as required for preconditioner from SuperLU

c/o S. Jardin, PPPL

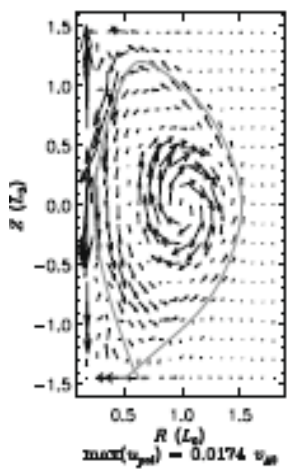
M3D-C^1 benchmarking

Development went through four stages of increasing complexity where useful results and critical benchmarking were performed at each step

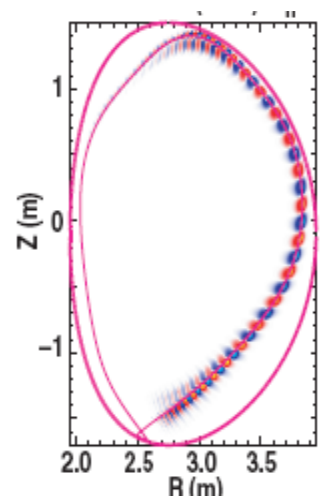
2D slab → 2D torus → 3D L torus → 3D NL torus



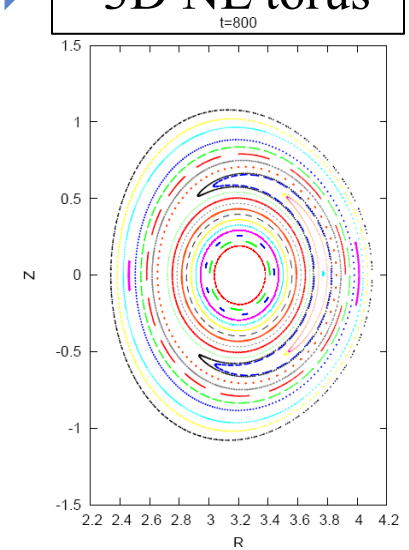
Two-fluid magnetic reconnection in model geometry



Two-fluid tokamak equilibrium with flow included

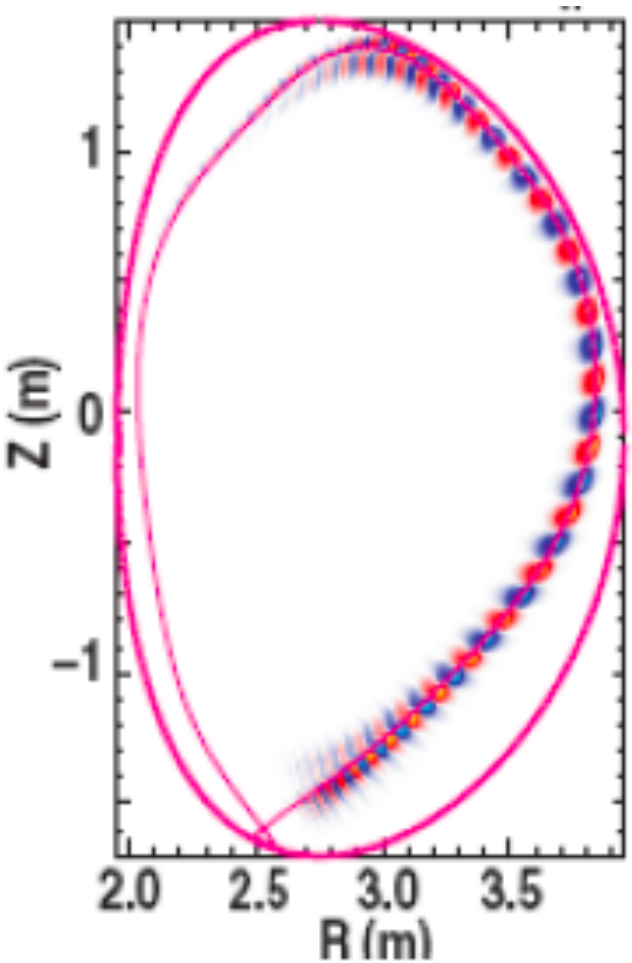


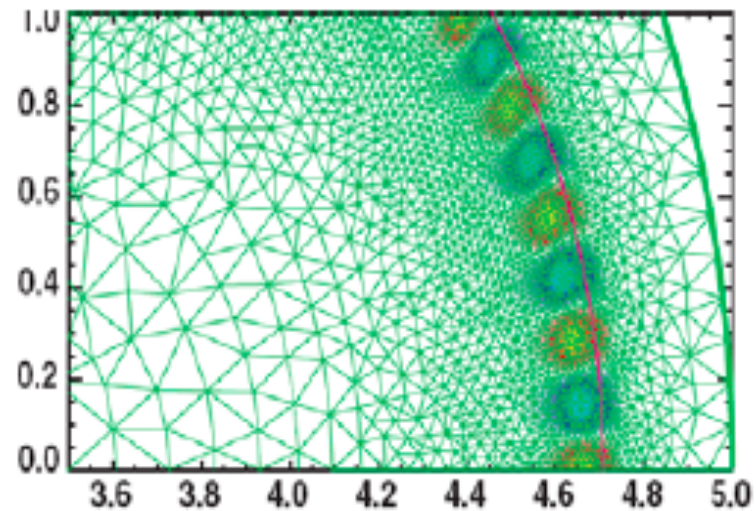
Linear study of edge localized modes in tokamak



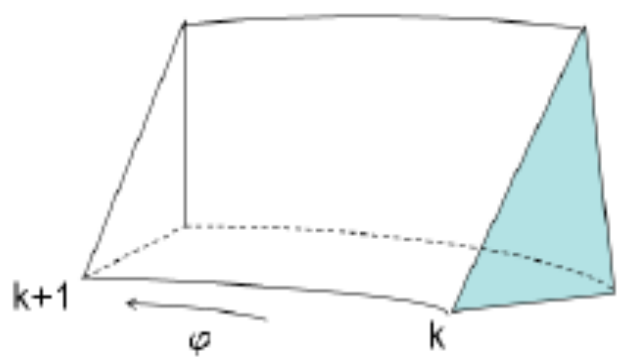
Nonlinear internal reconnection event in tokamak

Close-up of the resolution of edge-localized modes





Typical 3D C^1 wedge element obtained by tensoring 2D basis with Hermites in the toroidal direction.

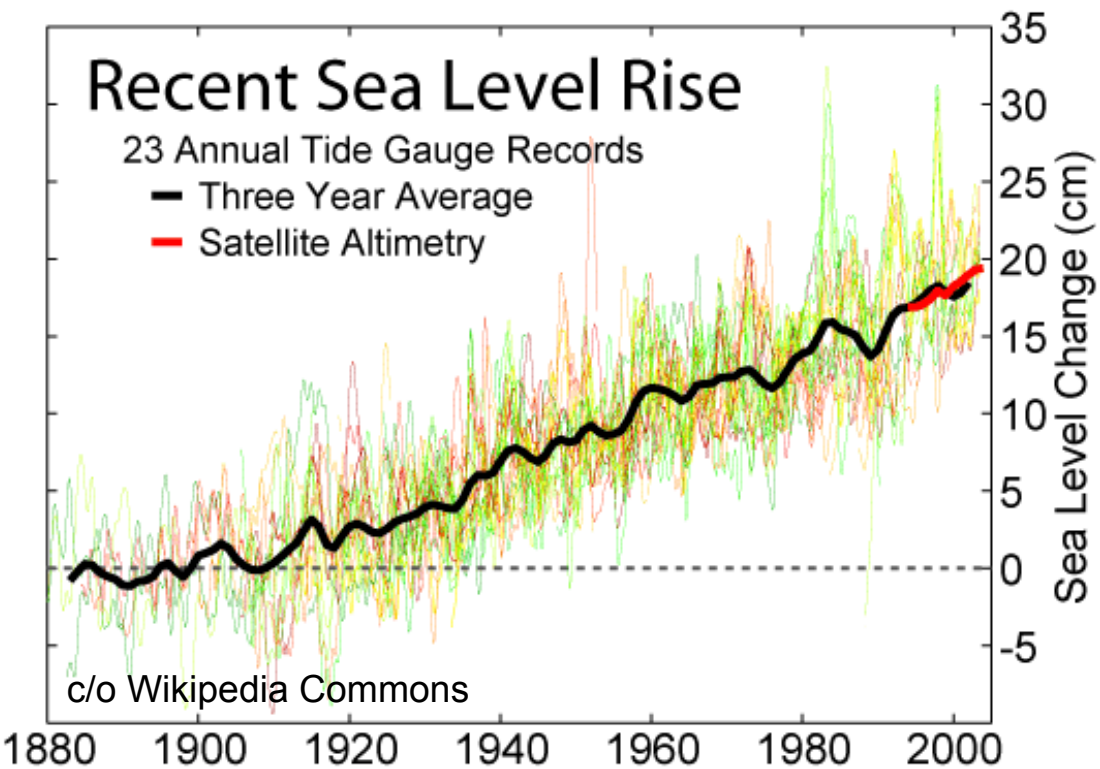


Fusion simulation: current status

- **Fusion group has a toolkit of linear solvers to call dynamically from the command line (PETSc)**
  - field-by-field scalar elliptic and time-implicit solvers
  - point-blocked solvers for tighter coupling of fields
  - direct sparse solver with fill-minimizing ordering (SuperLU) for 2D poloidal planes and other aggregates
  - algebraic multigrid solvers
  - additive Schwarz extensions to precondition 3D problems
  - Krylov accelerators
- **Current solver allows physicists to move about on a spectrum from robustness to optimality, with orders of magnitude runtime improvements over the robust default of direct sparse solves, at relevant contemporary granularities**

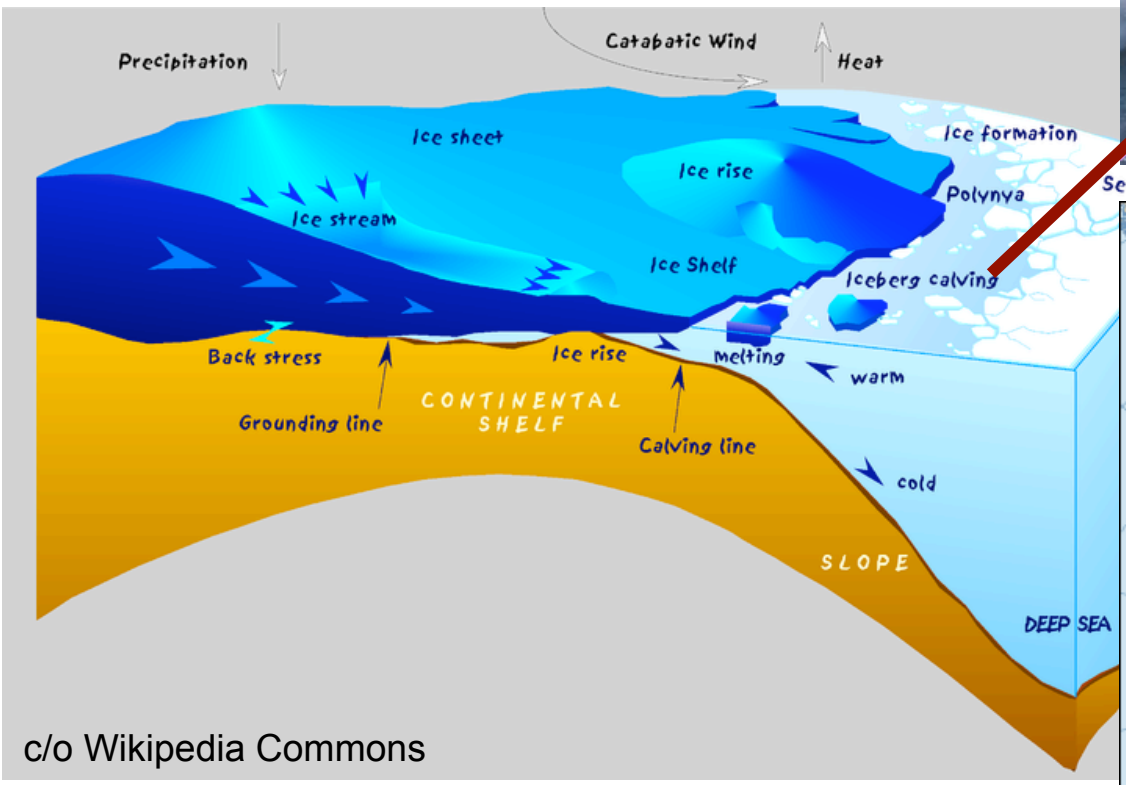
Application #2: fracture in ice sheets

- Ice sheets sitting on Greenland and Antarctica keep 77% of the world's freshwater locked up "high and dry"
- Average thickness 2.1 kilometers; now cover 10% of Earth's land area
- If all the fresh water land-locked in ice sheets and glaciers were to melt, it would cause a sea level rise of nearly 80 meters

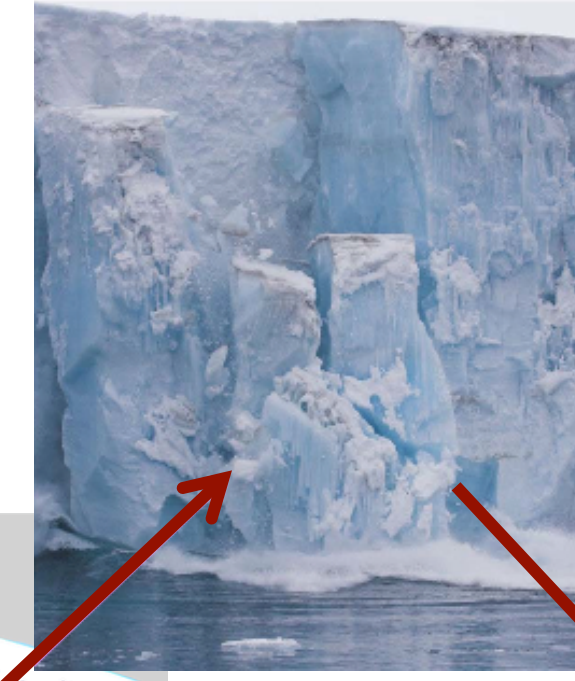


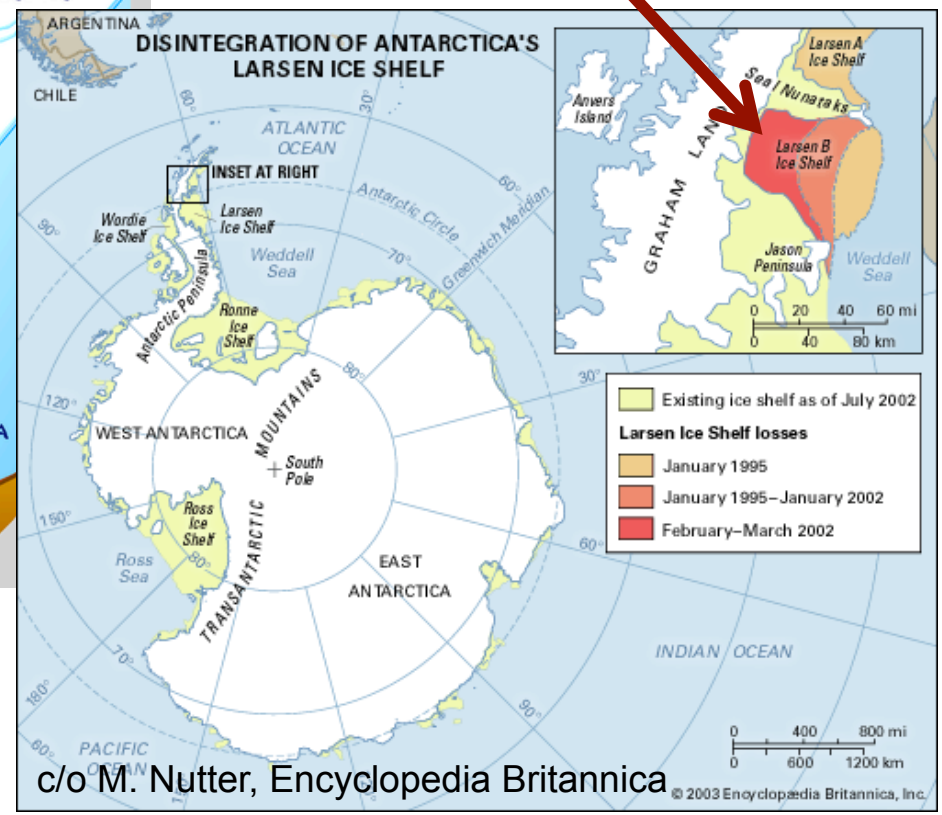
- In the last century, sea levels have risen about 0.2 meters
- On 6 Aug 2010, a piece of the Greenland ice sheet 4X the size of Manhattan fell into the sea
- Primary mechanism for losses:
  - sliding off land
  - calving at overhanging shelves
  - accelerated by fracture
- Climate models currently lack these dynamics

Ice *sheet* (on land)
vs.
ice *shelf* (over sea)



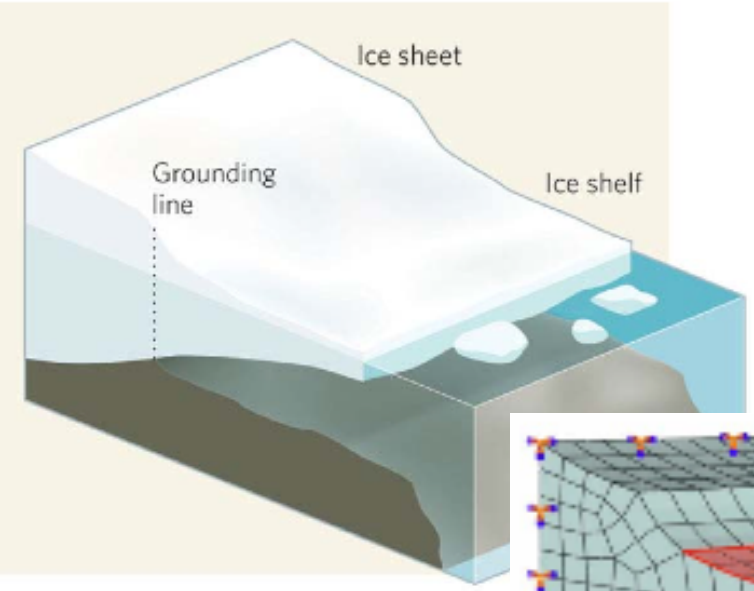
c/o Wikipedia Commons



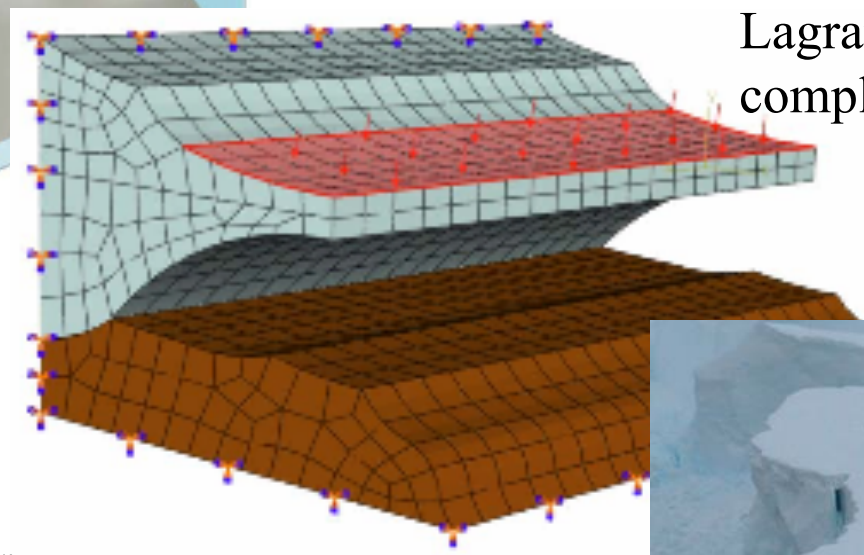


c/o M. Nutter, Encyclopedia Britannica

Ice sheet/shelf modeling: start with linear elasticity



- Cracks can be homogenized into the stress-strain constitutive relationship with a "damage" assumption, or they can be explicitly treated
- However, explicit transient Lagrangian remeshing can be complex



- Better to add new degrees of freedom rather than new mesh points?
- Have just received some ice sheet geometry data from the field; this talk is preliminary to the real application and is limited to 2D

c/o H. Waisman, Columbia University

Components of our computational model

- **Extended finite elements (XFEM)**
  - **XFEM developed in 1999 by Belytschko *et al.* at Northwestern to extend finite elements to problems with cracks (or other discontinuities) without slavish remeshing**
  - **Here, it is applied to brittle fracture**
- **Algebraic multigrid (AMG) solvers based on smoothed aggregation prolongators**
  - **SA-AMG developed in 1996 by Vanek *et al.* at Denver to build operator information into the coarsening strategy**
- **Domain decomposition (DD) to isolate the extra DOFs of XFEM in a small problem**
  - **Of the three main reasons for DD: (1) isolate different physics in different computational regimes, (2) achieving near-optimal sequential computational complexity, (3) scaling implicit finite element solvers to massively parallel computers, we are most closely related to the first, with a twist: isolating different discretizations**

Computational modeling of fracture

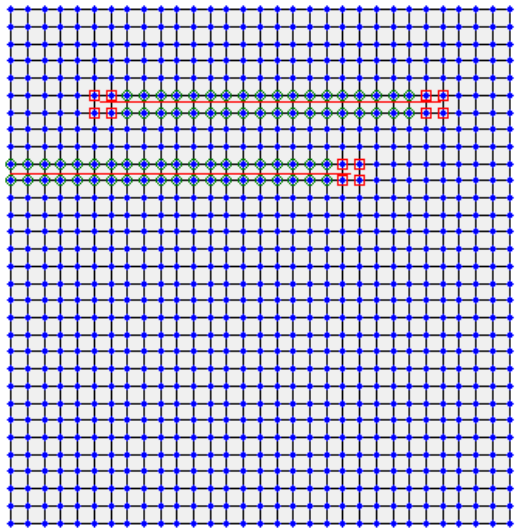
Classical FEM approach to fracture:

- Mesh conforms to crack boundaries
- Crack propagation requires remeshing at each step
  - Requires double-nodes for crack opening and fine mesh for tip singularities

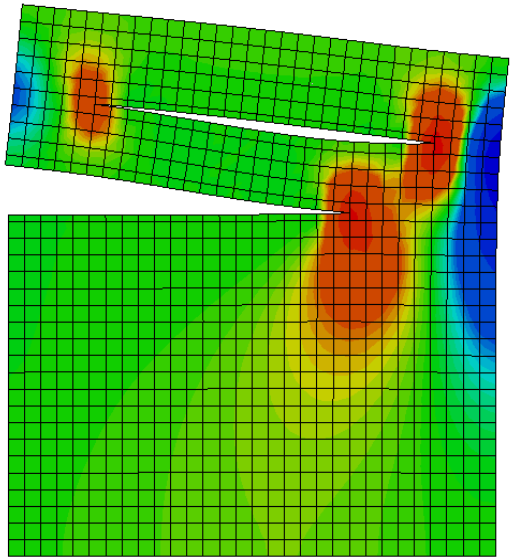
XFEM approach:

- Base mesh independent of crack geometry
- Crack propagation requires adding "enriched" DOF with special basis functions to existing nodes
  - Crack geometry defined through intersections of two levelset functions (for each crack), normal and tangential
  - Discontinuities and singularities captured through special basis functions (enrichments)
  - Enrichments have local support

XFEM mesh

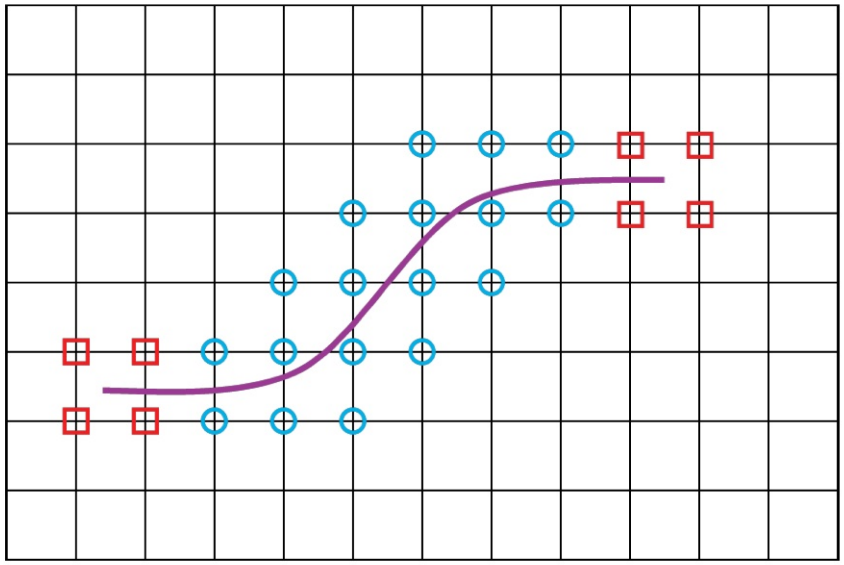


Stresses in *y* direction when bottom edge fixed and uniform traction applied on top edge in *y* direction



XFEM: employ regular elements and add degrees of freedom to parameterize the crack(s)

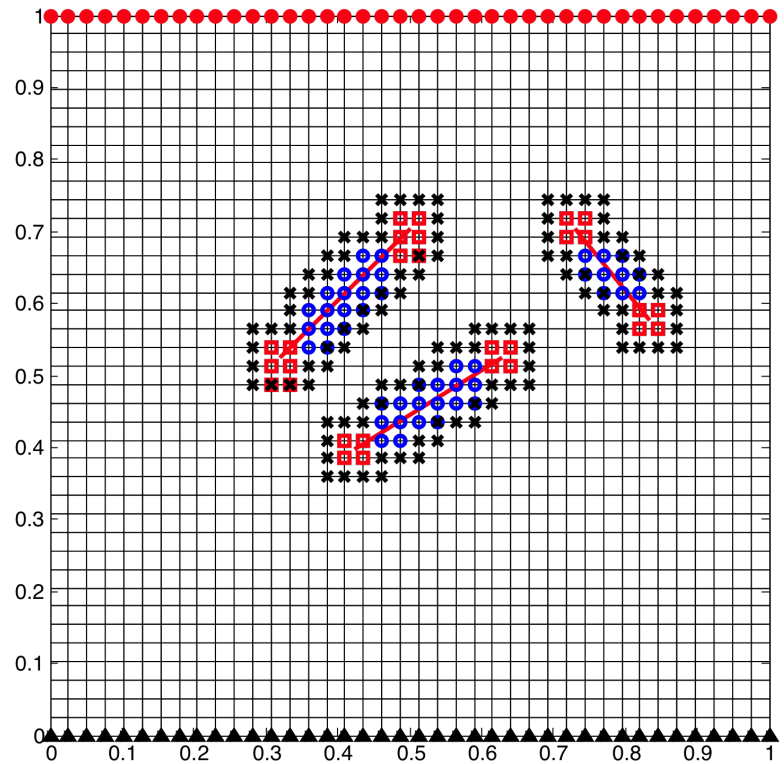
Single crack in 2D, zoom



○ Crack DOFs:
Heaviside functions

□ Crack-tip DOFs:
analytical singularities

Multiple cracks in 2D



The number of extended DOFs should be relatively “small” but can still be algebraically significant, worse in 3D

XFEM for fracture

XFEM Discrete model:
(Belytschko et al. 1999)

$$u^h(\mathbf{x}) = \overbrace{\sum_{I=1}^{n} N_I(\mathbf{x}) u_I}^{\text{Reg DOFs}} + \overbrace{\sum_{I=1}^{n_h} N_I(\mathbf{x}) H(\mathbf{x}) a_I}^{\text{Line DOFs}} + \overbrace{\sum_{I=1}^{n_f} N_I(\mathbf{x}) \sum_{J=1}^{n_J} F_J(\mathbf{x}) b_{IJ}}^{\text{Tip DOFs}}$$

Enrichment functions:

$$H(\mathbf{x}) = \begin{cases} 1 & \text{above } \Gamma_{c+} \\ -1 & \text{below } \Gamma_{c-} \end{cases}.$$

The $F_J(\mathbf{x})$ are given in local polar coordinates (r, θ) as

$$F_J(r,\theta) = \left\{ \overbrace{\sqrt{r}\sin\left(\frac{\theta}{2}\right)}^{J=1}, \overbrace{\sqrt{r}\cos\left(\frac{\theta}{2}\right)}^{J=2}, \overbrace{\sqrt{r}\sin\left(\frac{\theta}{2}\right)\sin(\theta)}^{J=3}, \overbrace{\sqrt{r}\cos\left(\frac{\theta}{2}\right)\sin(\theta)}^{J=4} \right\}$$

XFEM element stiffness matrix:

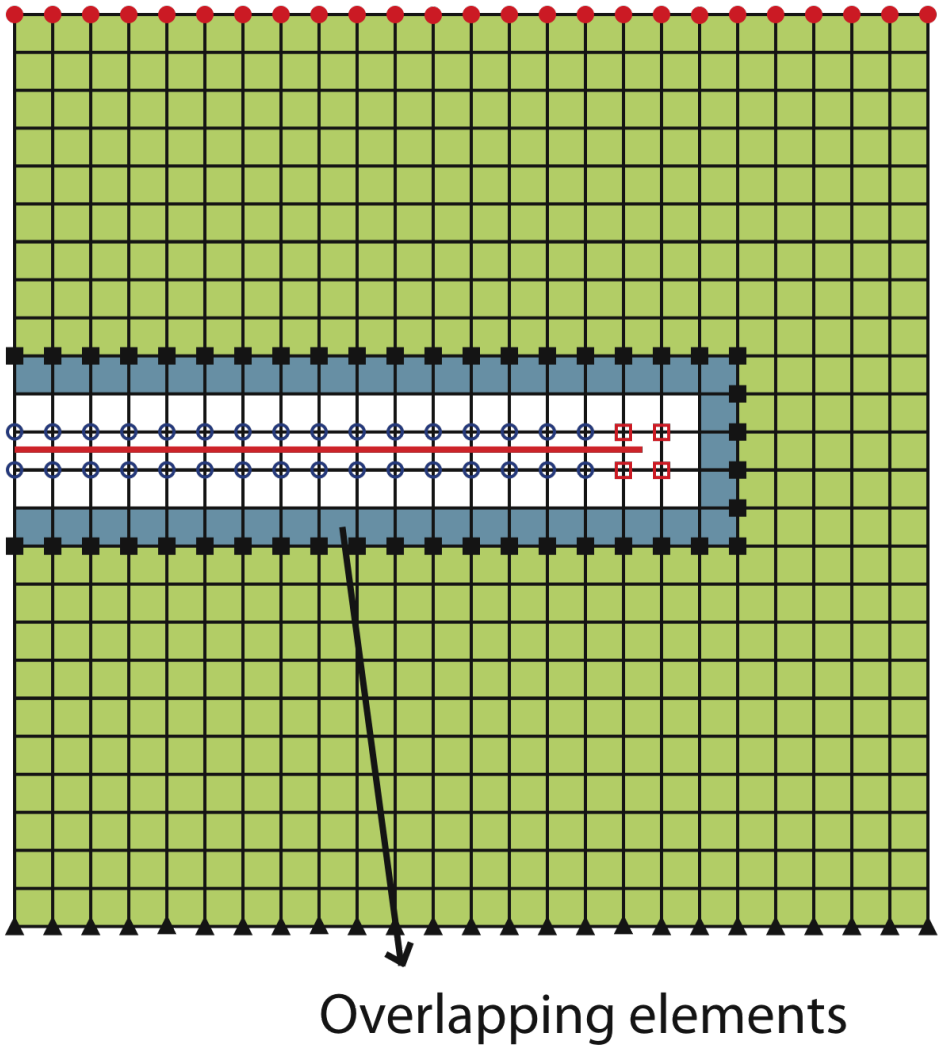
$$\mathbf{A}_e = \int_{\Omega_e} (\mathbf{B}^e_{enr})^T \mathbf{D} \mathbf{B}^e_{enr} d\Omega_e$$

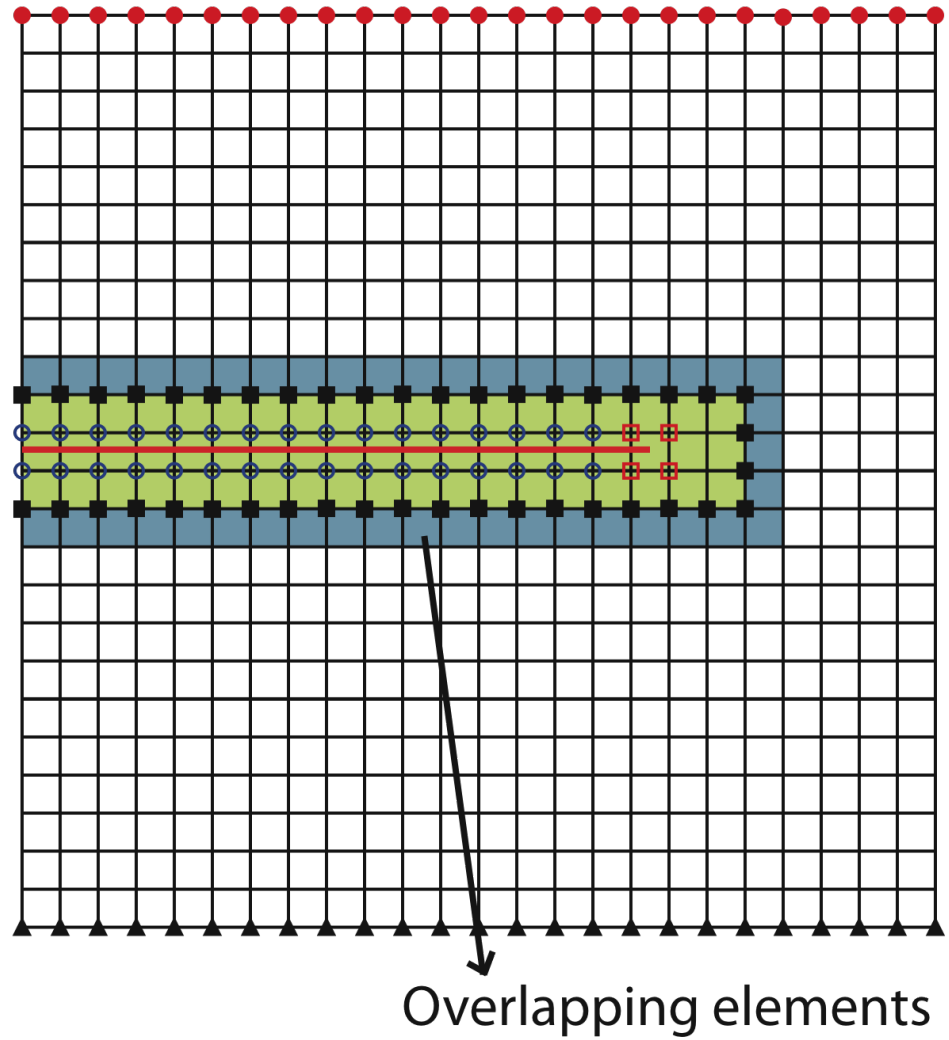
XFEM linear system after assembly:

$$\begin{bmatrix} A_{rr} & A_{rx} \\ A_{xr} & A_{xx} \end{bmatrix} \begin{bmatrix} u_r \\ u_x \end{bmatrix} = \begin{bmatrix} f_r \\ f_x \end{bmatrix}$$

Good methods exist for the red block, e.g., AMG

Schwarz approach builds preconditioner out of ambient "healthy" piece and local crack pieces





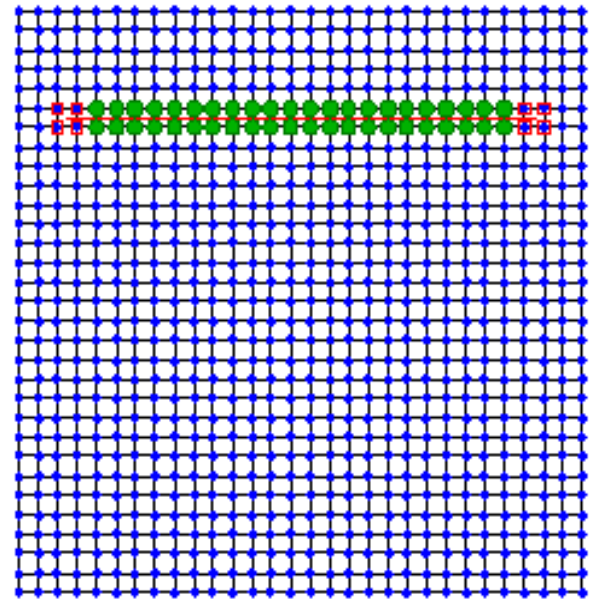
Structure of the A_{xx} block

XFEM Linear System:

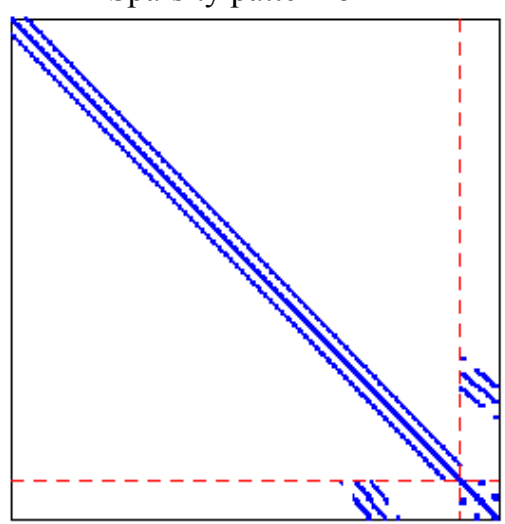
$$\begin{bmatrix} A_{rr} & A_{rx} \\ A_{xr} & A_{xx} \end{bmatrix} \begin{bmatrix} u_r \\ u_x \end{bmatrix} = \begin{bmatrix} \tilde{f}_r \\ \tilde{f}_x \end{bmatrix}$$

- Enriched DOF grouped together at the end in u_x
- A_{xx} small compared to A_{rr} for relatively small number of cracks
- Dense blocks in A_{xx} correspond to tip functions

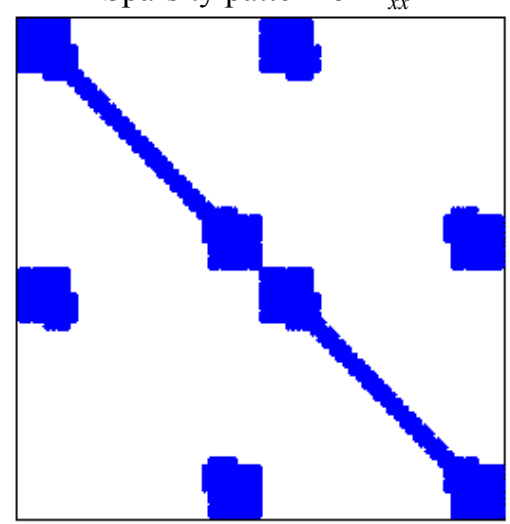
XFEM mesh



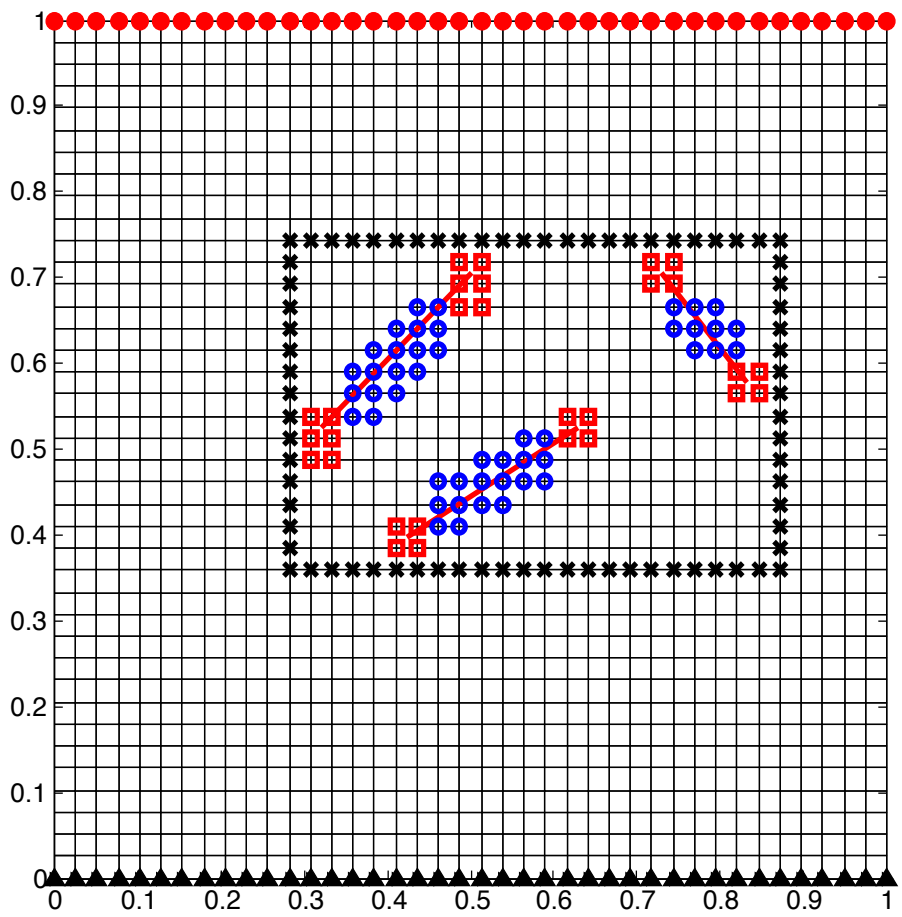
Sparsity pattern of A

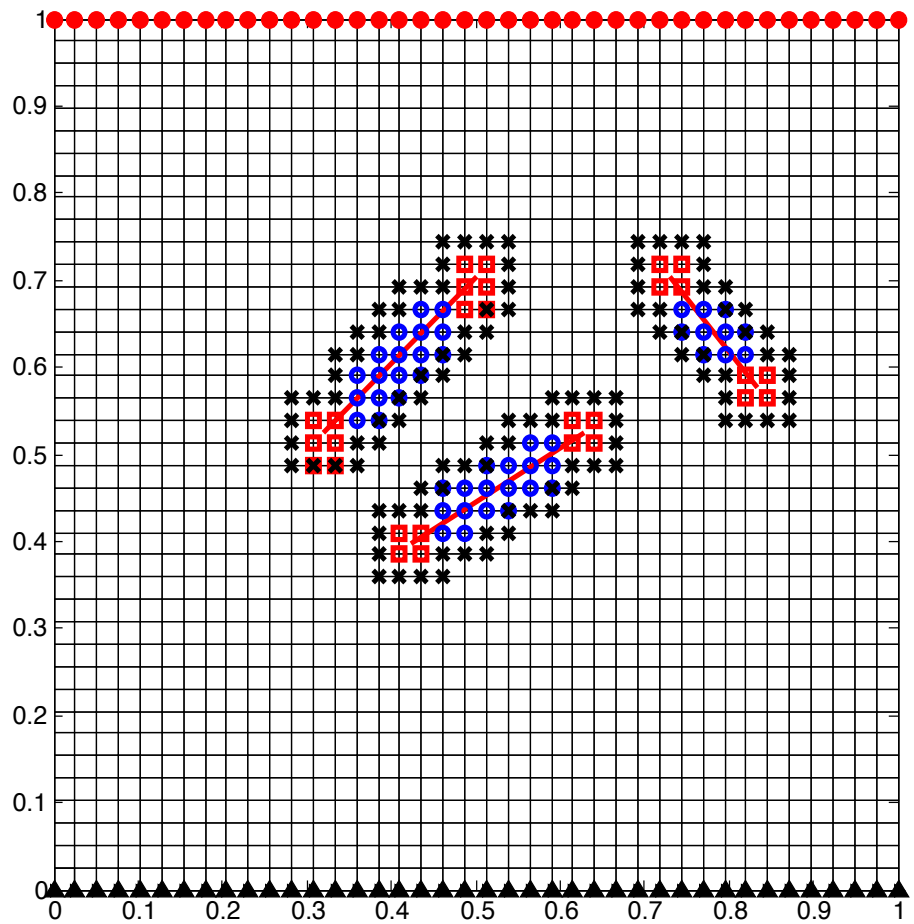


Sparsity pattern of A_{xx}

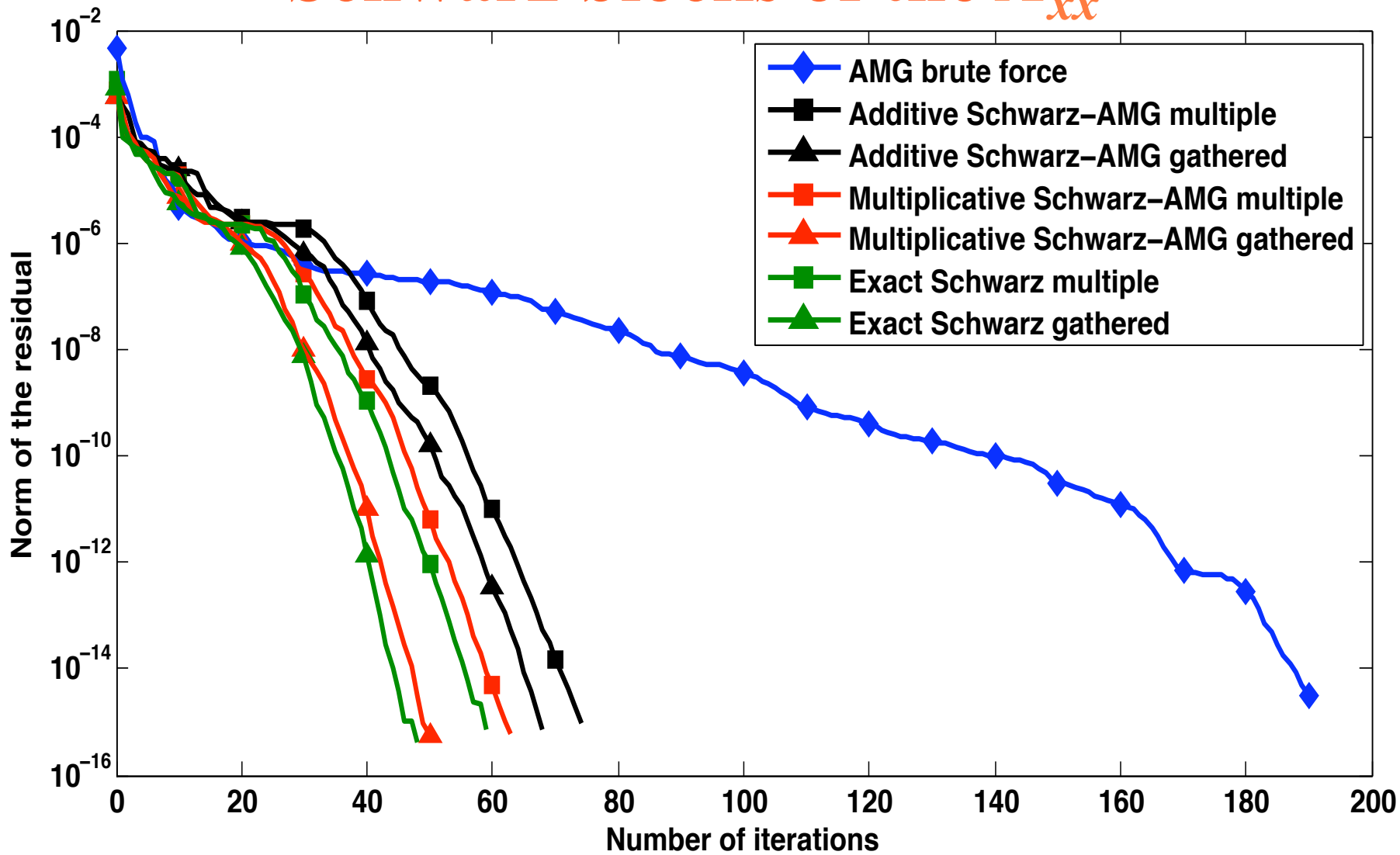


Cracks embedded with random angle





Brute-force AMG poor on the overall system but handles the A_{rr} piece, with LU/ILU on the Schwarz blocks of the A_{xx}

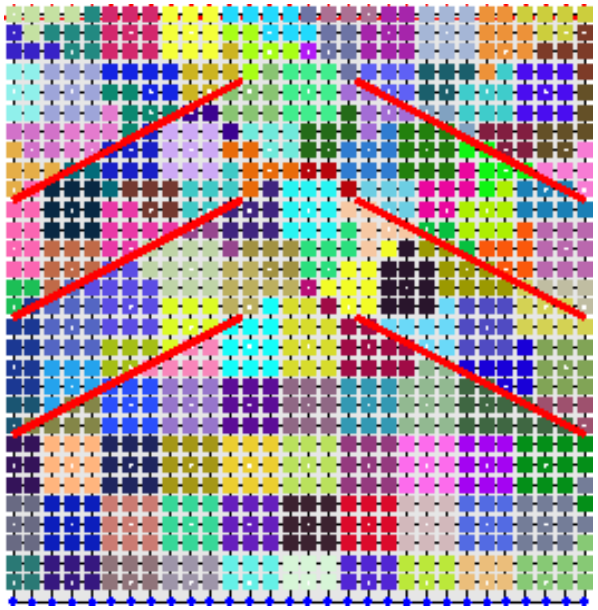


c/o L. Berger-Vergiat, Columbia University

Quasi-AMG

- Aggregates form the "coarse" nodes on the next level
- Aggregates should respect crack boundaries; otherwise, coarsening couples across cracks
- Break the graph edges in A corresponding to couplings across the crack interfaces using levelset information
  - Results in aggregates that respect crack boundaries

Before



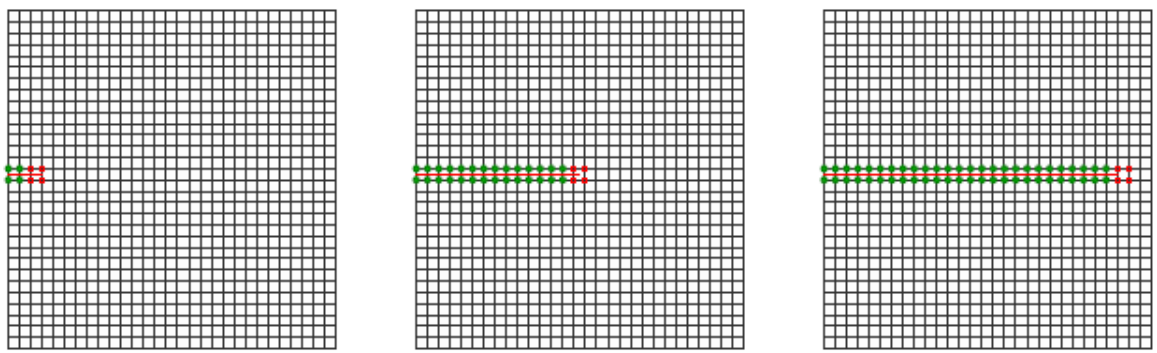
After

Each color represents one aggregate at the coarse level

Test cases

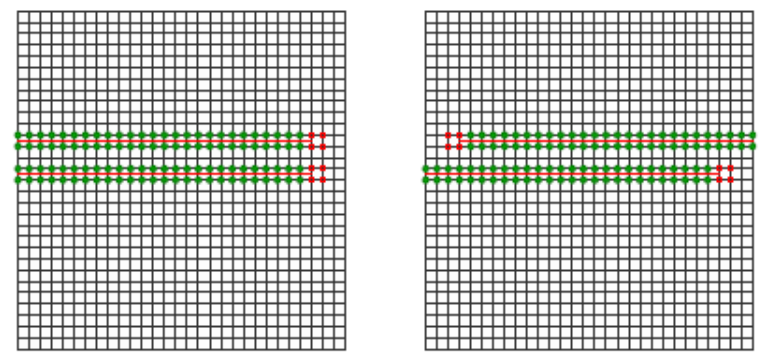
- **Both edge cracks and interior cracks are considered**
- **For each crack-configuration, following mesh densities are considered**
  - **30 x 30**
  - **60 x 60**
  - **90 x 90**
  - **120 x 120**

Single Propagating Crack



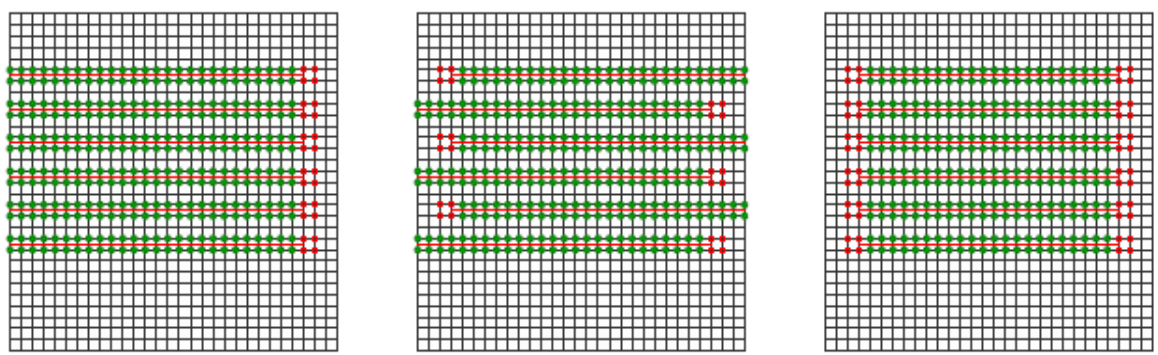
(a) Case 1a (b) Case 1b (c) Case 1c

Two Cracks



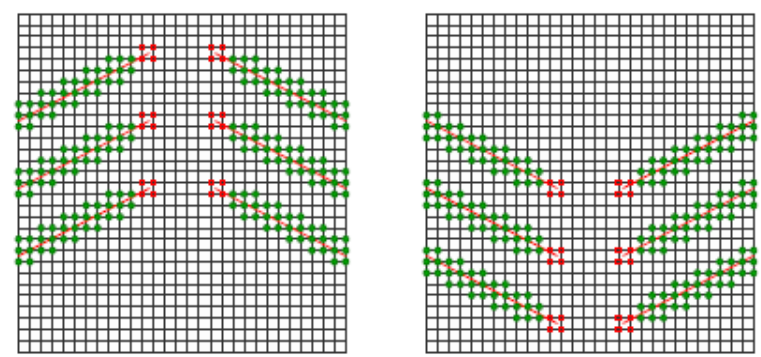
(d) Case 2a (e) Case 2b

Six Cracks



(f) Case 3a (g) Case 3b (h) Case 4

Inclined Cracks



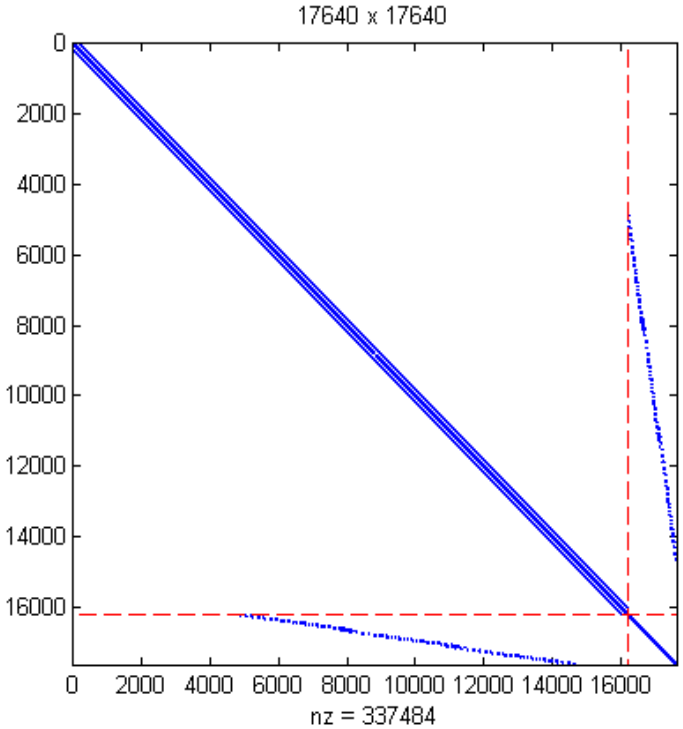
(i) Case 5a (j) Case 5b

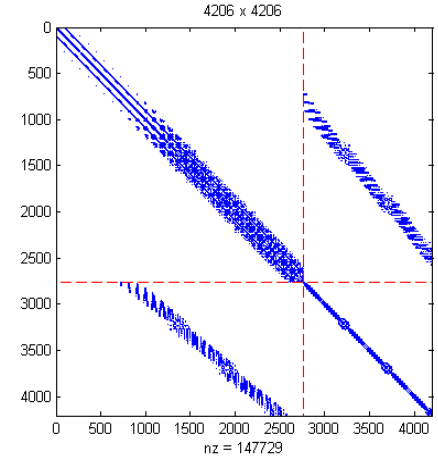
Numerical results

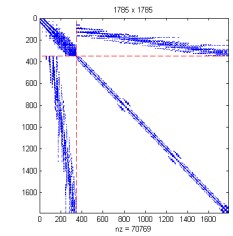
Case	VBlk AMG	Hybrid Standard AMG	Quasi AMG	Mesh	Case	VBlk AMG	Hybrid Standard AMG	Quasi AMG
1a	28	13	11	30^2	3a	154	-	16
	29	15	10	60^2		127	-	14
	37	17	12	90^2		-	-	25
	37	19	12	120^2		-	-	21
1b	24	22	11	30^2	3b	-	-	18
	24	29	12	60^2		-	-	21
	36	35	14	90^2		-	-	28
	35	41	13	120^2		-	-	22
1c	31	31	13	30^2	4	116	107	15
	32	43	14	60^2		102	154	21
	47	53	16	90^2		142	190	23
	45	61	15	120^2		151	-	22
2a	64	57	15	30^2	5a	80	76	12
	52	80	14	60^2		91	107	13
	87	98	20	90^2		124	131	15
	92	113	18	120^2		140	151	15
2b	73	59	16	30^2	5b	89	81	16
	72	81	17	60^2		103	116	15
	97	104	21	90^2		134	143	17
	95	122	19	120^2		151	165	16

c/o R. Tuminaro, Sandia

Remark on coarsening







Not to scale

$$A^{[k]} \longrightarrow \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix}^T \begin{bmatrix} A_{rr} & A_{rx} \\ A_{xr} & A_{xx} \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \longrightarrow A^{[k+1]}$$

Currently, only regular DOFs are currently coarsened. This works well when there are few cracks, but limits scalability for problems with many cracks. We are currently considering special coarsening of some enrichment functions.

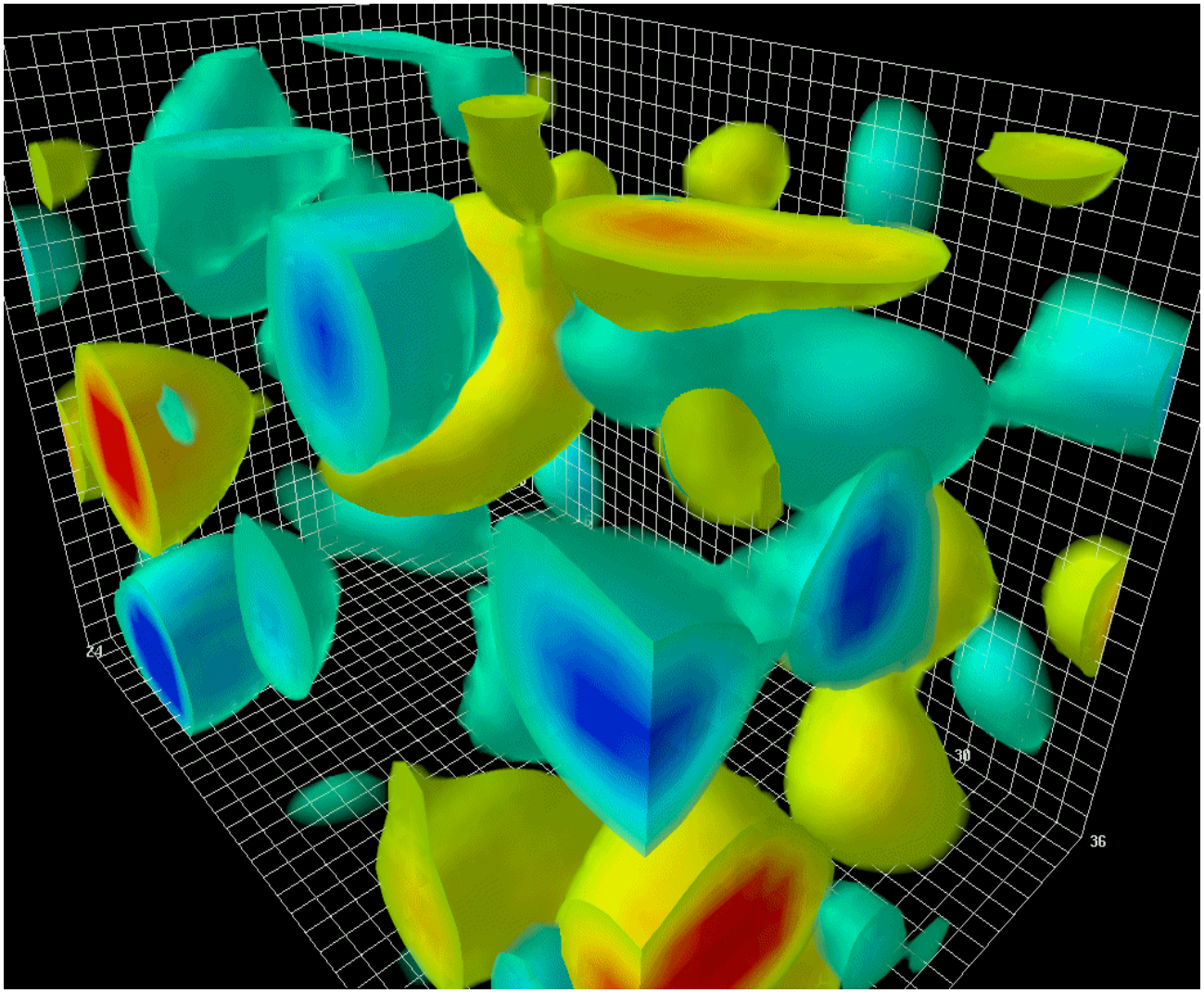
Ice-sheet modeling: current status

- Straightforward Schwarz approach to employ AMG in XFEM by avoiding the irregular blocks of the enriched degrees of freedom
- Uses only blackbox solvers
- Crack-sensitive prolongator for AMG is yet further superior
  - Current AMG solver coarses "healthy" domains, but preserves extended degrees of freedom, resulting in increasing relative size of extended DOFs at coarser levels
- Challenge looms in 3D, where extended DOFs live on 2D crack surfaces

Application #3: Quantum Chromodynamics

QCD is the theory of strong forces in the Standard Model of particle physics. It Describes the structure of nucleons, which are made up of quarks interacting in a gluon field.

The lattice is typically 4- or 5-dimensional, and the fields consist of 12 components at each node. The gluon field mixes the components through unitary matrix coefficients on each edge.



The regularity of the mesh and the high resolution required has long suggested multigrid to physicists, but prior to SciDAC, achievements were limited.

SUPERCOMPUTER
COMPUTATIONS
RESEARCH INSTITUTE

PROJECTIVE MULTIGRID FOR WILSON FERMIONS

by

Richard C. Brower, Robert G. Edwards,
Claudio Rebbi, and Ettore Vicari

FSU-SCRI-91-54

R. C. Brower, R. Edwards, C.Rebbi,and E. Vicari,
"Projective multigrid for Wilson fermions", Nucl. Phys.B366 (1991) 689
(a.k.a. “Spectral AMG”, Tim Chartier, 2000)

The Dirac PDE (for Quarks)

$$\sum_{\mu=1}^{4} \gamma_\mu^{ij}[\frac{\partial}{\partial x_\mu} - iA_\mu^{ab}(x)]\psi_{jb}(x) + m\psi_{ia}(x) = b_{ia}(x)$$

3x3 color gauge matrices

4x4 sparse spin matrices:
4 non-zero entries 1,-1, i, -i

$x_\mu = (x_1,x_2,x_3,x_4)$
(space,time)

On a Hypercubic Lattice **(x_μ = integer, a = lattice spacing):**

$$\sum_{\pm\mu} \frac{\pm\gamma_\mu - 1}{2} U(x, x\pm\mu)\psi(x\pm\mu) + (am+4)\psi(x) = b(x)$$

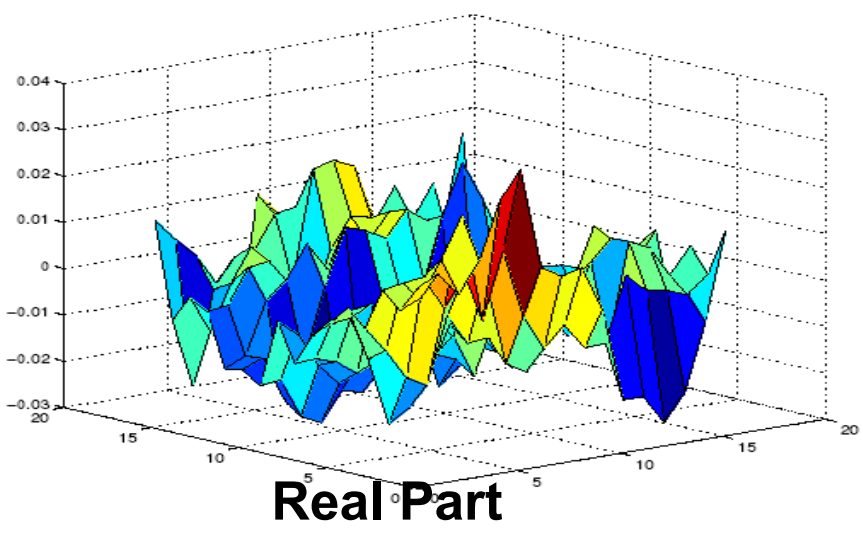
3x3 Unitary : $U(x,x+\mu) = \exp[i\, a\, A_\mu(x)]$ and $U(x,x-\mu) = U^*(x-\mu,x)$

c/o R. Brower, Boston University

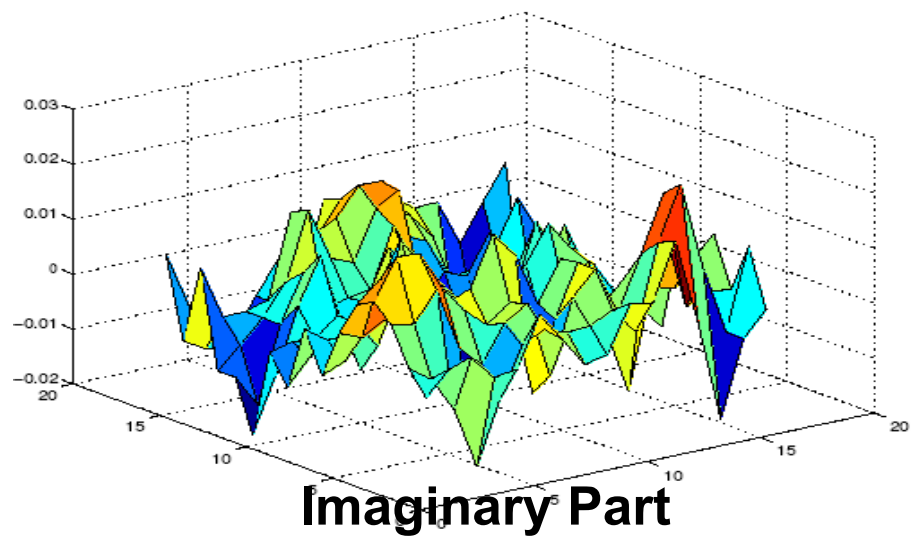
Scalable solvers for the Dirac equations in QCD have been elusive until now

- **Challenges in solving the Dirac equations**
  - System is complex and indefinite
  - System can be extremely ill-conditioned
  - Near null space is unknown and oscillatory

Two-dimensional model problem for the a scalar complex field, showing an "instanton", or a null-space mode of the Dirac operator



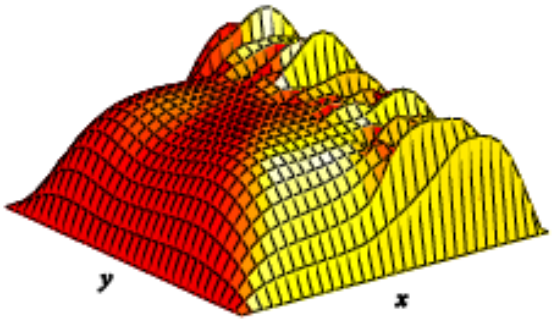
Real Part



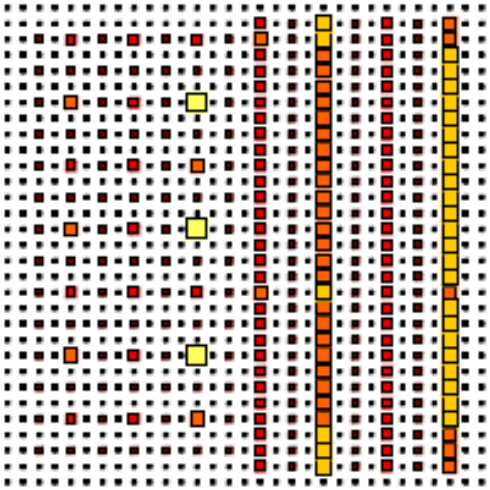
Imaginary Part

The importance of the near null space

- **Multigrid (MG) methods are based on knowledge of the near null space**
  - **For the Laplacian, these are the smoothest modes and are well approximated geometrically**
  - **Generally, for indefinite (oscillatory) problems, smooth modes are irrelevant**
- **Algebraic Multigrid (AMG) uses matrix coefficients**
  - **Automatically coarsens "grids"**
- **Error left by relaxation is *algebraically smooth***
- **Coarsening *must interpolate small eigenmodes well***



Error after 7 GS sweeps

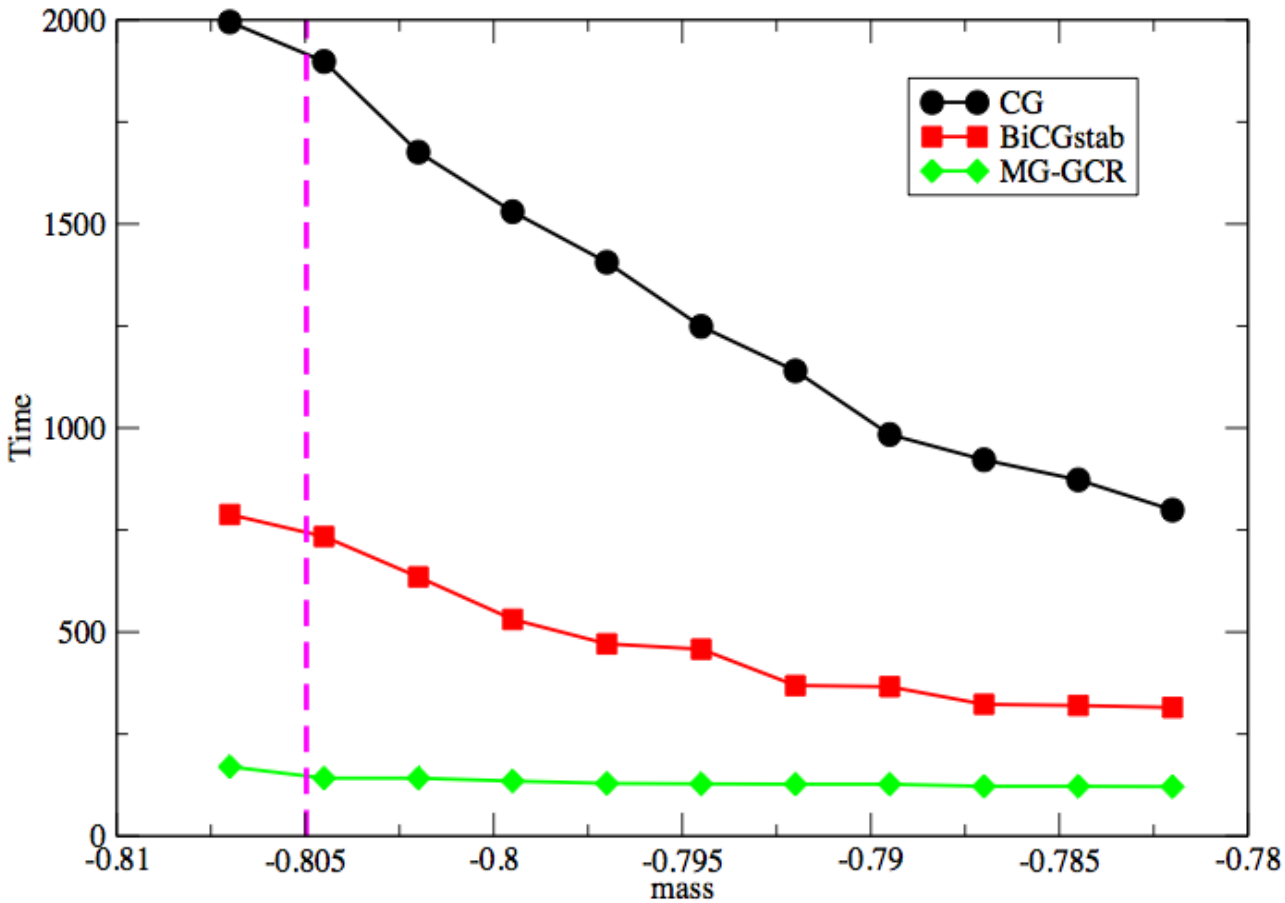


AMG coarsens grids in the direction of geometric smoothness

Adaptive smoothed aggregation (αSA) automatically builds the coarse space

- **Generate the basis one vector at a time**
  - **Start with relaxation on $Au=0 \rightarrow u_1 \rightarrow \alpha SA(u_1)$**
  - **Use $\alpha SA(u_1)$ on $Au=0 \rightarrow u_2 \rightarrow \alpha SA(u_1,u_2)$**
  - **Iterate until we have a good coarse basis**
- **Setup is expensive, but is amortized over many right-hand sides**
- **Published in 2004 by SciDAC TOPS team**
  - **Brezina, Falgout, MacLachlan, Manteuffel, McCormick, and Ruge, "*Adaptive smoothed aggregation (αSA),*" SIAM J. Sci. Comput. (2004)**
- **Demonstrated in 2D QED in 2005**
  - **Brannick, Brezina, Keyes, Livne, Livshits, MacLachlan, Manteuffel, McCormick, Ruge, and Zikatanov, "*Adaptive smoothed aggregation in lattice QCD,*" Springer (2006)**
- **Subsequently migrated to realistic models**

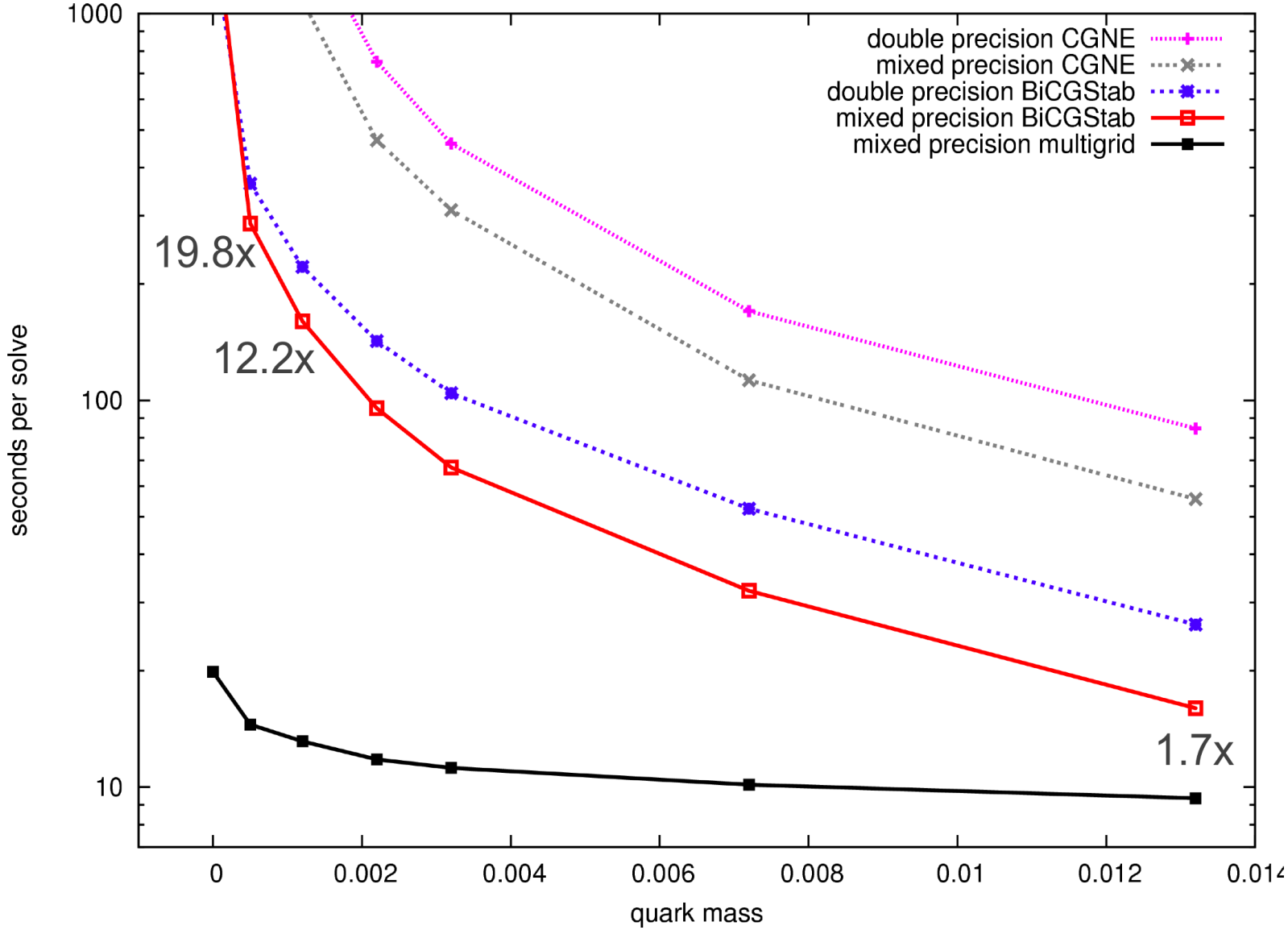
4D Wilson-Dirac Results
αSA-MG shows no slowing down



- **Parameters: $N=16^3x32$, $\beta=6.0$, $m_{crit} = -0.8049$**
- **MG Parameters: 4^4x3x2 blocking, 3 levels, W(2,2,4) cycle, $N_v = 20$, setup run at m_{crit}**

c/o R. Brower, Boston University

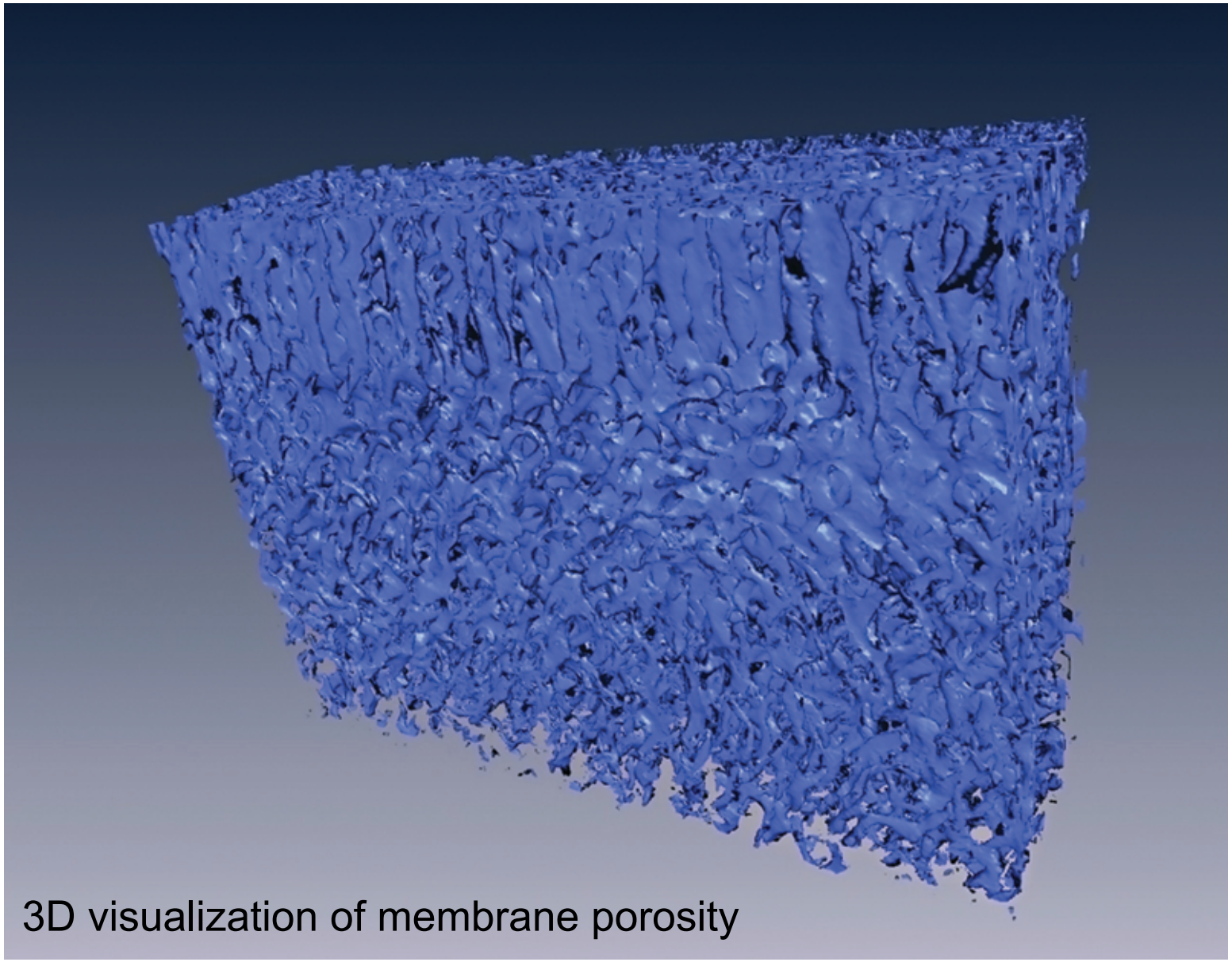
4D Wilson-Dirac outpaces competition at all quark masses



QCD Dirac solves: current status

- **Adaptive Smoothed Aggregation Algebraic Multigrid is expensive to use just once due to set up costs, but in QCD applications, where the same gluon coefficient field is used tens to hundreds of times on different right-hand sides, it is a paradigm shifter**
- **The superiority of the method increases with the resolution of the problem and the approach of quark mass to physically realistic values**
- **Complicated algorithm, with data-dependent sparsity is challenging to encode as efficiently as brute-force methods, but this is now underway – even on GPUs**

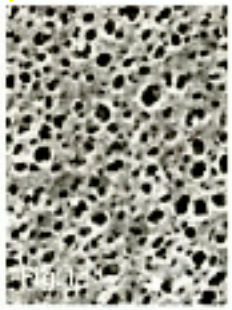
Application #4: Modeling phase separation with Cahn-Hilliard

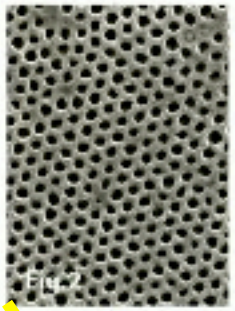


c/o K.-V. Peinemann, KAUST

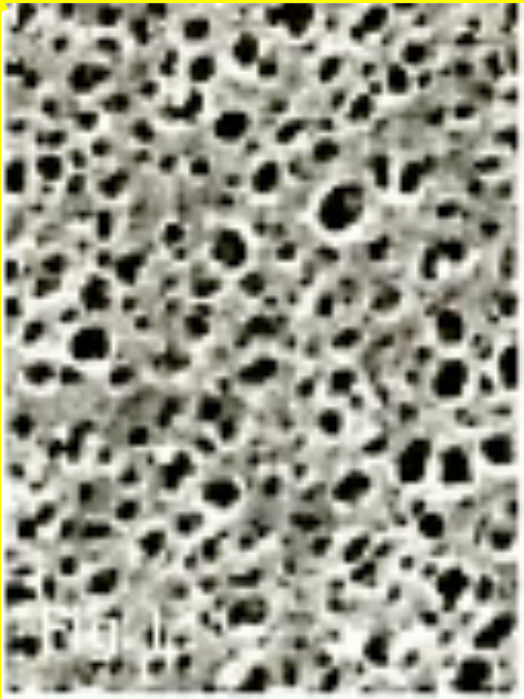
High density, high regulari s

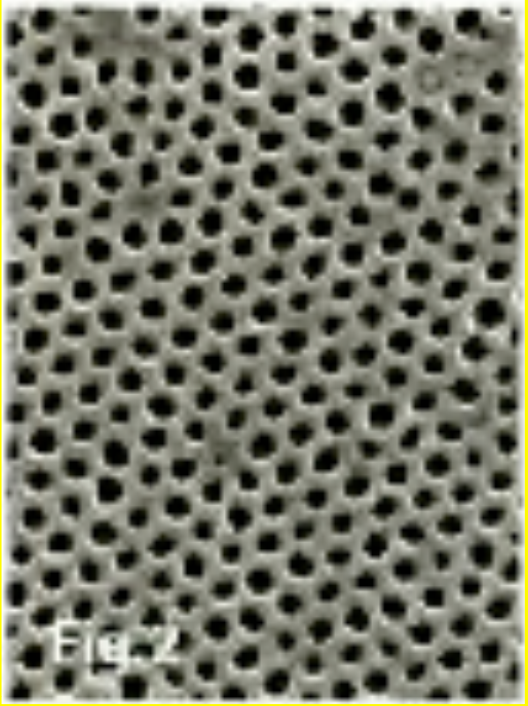
RIGHT FIG.1: Conventional polyacrylonitrile (PAN) membrane
RIGHT FIG.2: Isoporous membrane surface prepared by the self-assembly of block-copolymer micelles





CO-AUTHORS
RIGHT): Ali R
Bobby Hoogha
Sougrat and D
Absent from p
Neckalanda P
Pinnau and Ul



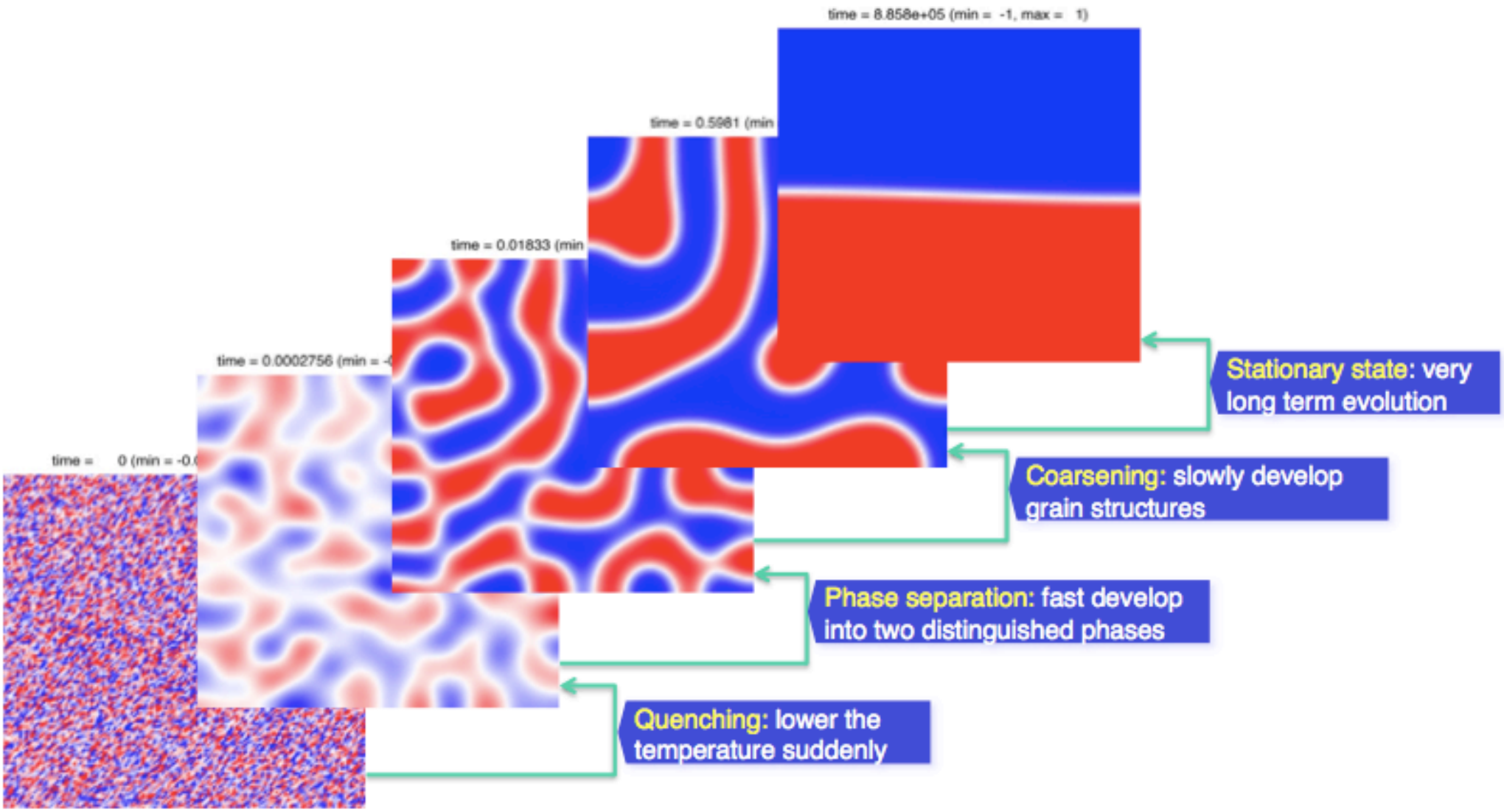


Klaus-Viktor, a veteran of the membrane technology industry, realized that "phase inversion" (abrupt phase separation by immersion in a nonsolvent bath), a well-defined process in the manufacture of commercial membranes, might be the key to the commercial viability of their findings. Using a very simple and easily scalable casting procedure followed by an immediate water bath quenching (for "phase inversion"), the pore size and distribution in the m
stabilized and set.

With a patent app
ess and a small-se
manufacture small
fi n currently under
KA ST researchers h
comp y to manufa
membra es. Meanwh
lab, they will contin
science and chniqu
films, moving toward
creating membra es
pores."

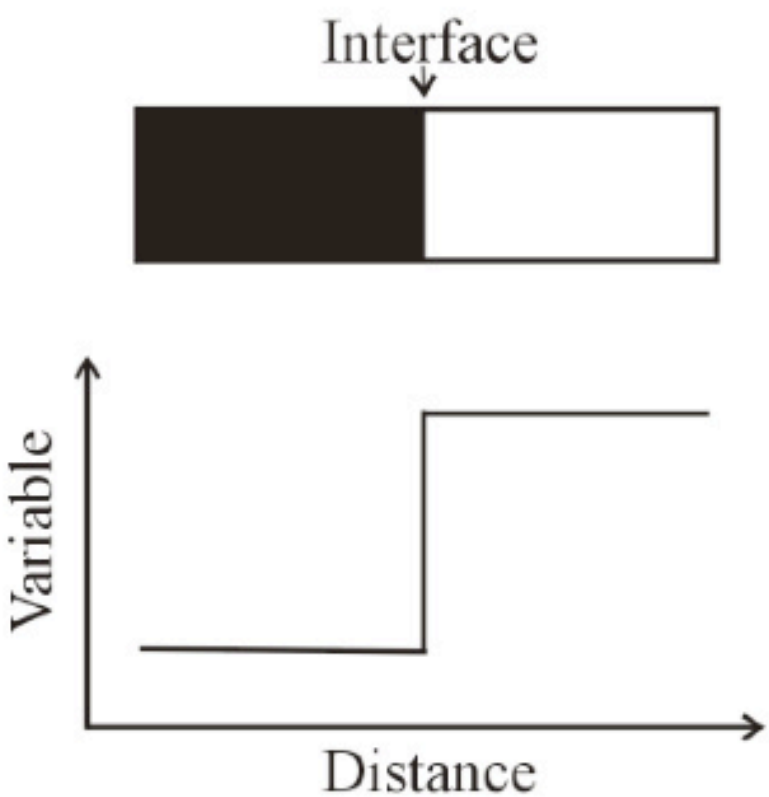
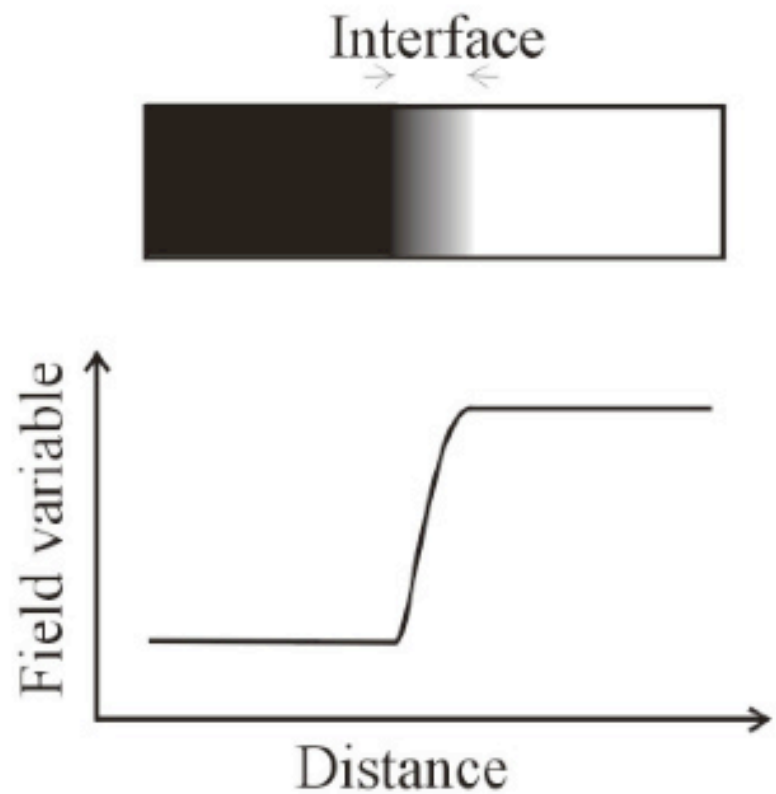
uzana and
example of
ory condi-
duction of
hough their
ST Research
kanda told
nce of high
ST allowed
whether the
ot." In this
potential to
millions of

Spinodal decomposition
(Cahn & Hilliard, J. Chem. Phys. 1958)



c/o Chao Yang, CAS

Sharp (Lagrangian) vs. diffuse (Eulerian) models

Sharp-interface models	Diffuse-interface models
	
• Discontinuous • Explicit interface-tracking • Simplified grain morphologies	• Continuous • Interfaces implicitly included • Complex grain morphologies

General form of Cahn-Hilliard model

$$\boxed{\frac{\partial u}{\partial t} - \nabla \cdot M(u) \nabla \frac{\delta E(u)}{\delta u} = 0}$$

- Domain: $\mathbf{x} \in [0,1]^d, d = 1,2,3,\ t \geq 0$
- $u(\mathbf{x}, t) \in [-1, 1]$: concentration difference of a binary mixture
- $M(u) \geq 0$: mobility
- $E(u) = E^c(u) + E^i(u)$: Ginzburg-Landau free energy
  - $E^c(u) = \int_\Omega W(u) d\mathbf{x}$: chemical energy (or bulk energy)
  - $E^i(u) = \int_\Omega \frac{\epsilon^2}{2} |\nabla u|^2 d\mathbf{x}$: interfacial energy
  - $0 < \epsilon << 1$: interfacial sharpness parameter

The C-H equation: $\boxed{\frac{\partial u}{\partial t} + \epsilon^2 \nabla \cdot M(u) \nabla \Delta u - \nabla \cdot M(u) \nabla W'(u) = 0}$

- Boundary conditions
  - Periodic
  - Neumann: $\frac{\partial u}{\partial \nu} = \frac{\partial \Delta u}{\partial \nu} = 0$

Parameterizations

- **Case 1: ideal**
  - Quartic chemical potential (double well):

$$W(u) = \frac{1}{4}(1 - u^2)^2$$

  - Constant mobility:

$$M(u) = 1$$

- **Case 2: realistic**
  - Logarithmic chemical potential:

$$W(u) = \frac{1}{2}\left((1 + u)\ln(1 + u) + (1 - u)\ln(1 - u) - \theta u^2\right)$$

The constant $\theta > 1$: quench ratio $(T_{critical}/T_{absolute})$

  - Thermodynamically consistent mobility:

$$M(u) = \frac{1}{4}(1 - u^2)$$

Time discretization (1)

Discretize $u_{i,j}(t)$ with $u_{i,j}^n = u_{i,j}(t_n)$ and put in a vector U^n.

The C-H equation: $\frac{\partial u}{\partial t} + \epsilon^2 \nabla \cdot M(u) \nabla \Delta u - \nabla \cdot M(u) \nabla W'(u) = 0.$

Denote A and $\tilde{A}^n$ as the discrete operators of $-\Delta$ and $-\nabla \cdot M(U^n)\nabla$.

- Forward Euler: suffers severely from stability limit

$$\frac{U^{n+1} - U^n}{\Delta t} + \epsilon^2 \tilde{A}^n A U^n + \tilde{A}^n W'(U^n) = 0. \qquad \text{(Euler)}$$

- Semi-implicit: also suffers from stability limit

$$\frac{U^{n+1} - U^n}{\Delta t} + \epsilon^2 \tilde{A}^{n+1} A U^{n+1} + \tilde{A}^n W'(U^n) = 0. \qquad \text{(SI)}$$

- Fully implicit backward Euler: typically $\Delta t \leq \mathcal{O}(\epsilon^2)$, Copetti92'NM

$$\frac{U^{n+1} - U^n}{\Delta t} + \epsilon^2 \tilde{A}^{n+1} A U^{n+1} + \tilde{A}^{n+1} W'(U^{n+1}) = 0. \qquad \text{(B-Euler)}$$

Stability through energy splitting

- **The numerical scheme should discretely obey energy decay**

$$\mathcal{E}(U^{n+1}) \leq \mathcal{E}(U^n), \quad n = 1, 2, 3, \ldots$$

- **Energy splitting (Eyre'98, He'07, Shen'10) breaks energy into a convex and a concave part**
  - **Keep convex on the LHS (implicit), where it enhances definiteness**
  - **Keep concave on the RHS (explicit), lest it oppose definiteness**

Time discretization (2)

- **Case 2: realistic (variable mobility)**

  The C-H eq.

$$\frac{\partial u}{\partial t} + \epsilon^2 \nabla \cdot M(u) \nabla \Delta u + \theta \nabla \cdot M(u) \nabla u - (1/4) \Delta u = 0$$

  The free energy

$$E = \frac{1}{2} \int_\Omega \left(\epsilon^2 |\nabla u|^2 + (1+u) \ln(1+u) + (1-u) \ln(1-u) - \theta u^2 \right) d\mathbf{x}$$

  Nonlinearly stabilized (NS) scheme

  - The splitting

$$E_2 = -\frac{\theta}{2} \int_\Omega u^2 d\mathbf{x}, \quad E_1 = E - E_2$$

  - The NS scheme

$$\frac{U^{n+1} - U^n}{\Delta t} + \epsilon^2 \tilde{A}^{n+1} A U^{n+1} - (1/4) A U^{n+1} + \theta \tilde{A}^n U^n = 0$$

Time discretization (3)

- **The rate of evolution of the interface varies enormously, becoming extremely slow near minimum surface area equilibrium**
- **A main interest is to get stably to long-term configuration**
- **Use "Switched Evolution-Relaxation" (SER) of Mulder & Van Leer (1985)**
  - **build up time step in inverse proportion to some fractional power of steady-state residual decrease**
  - **subject to maximum increase and minimum decrease ratio**
- **Initially, $\Delta t \approx O(\varepsilon^2)$**
- **Robustification feature**
  - **if Newton diverges, recursively halve the timestep**
  - **happens less than 10% of the time, despite sometimes rapid increases in Δt**

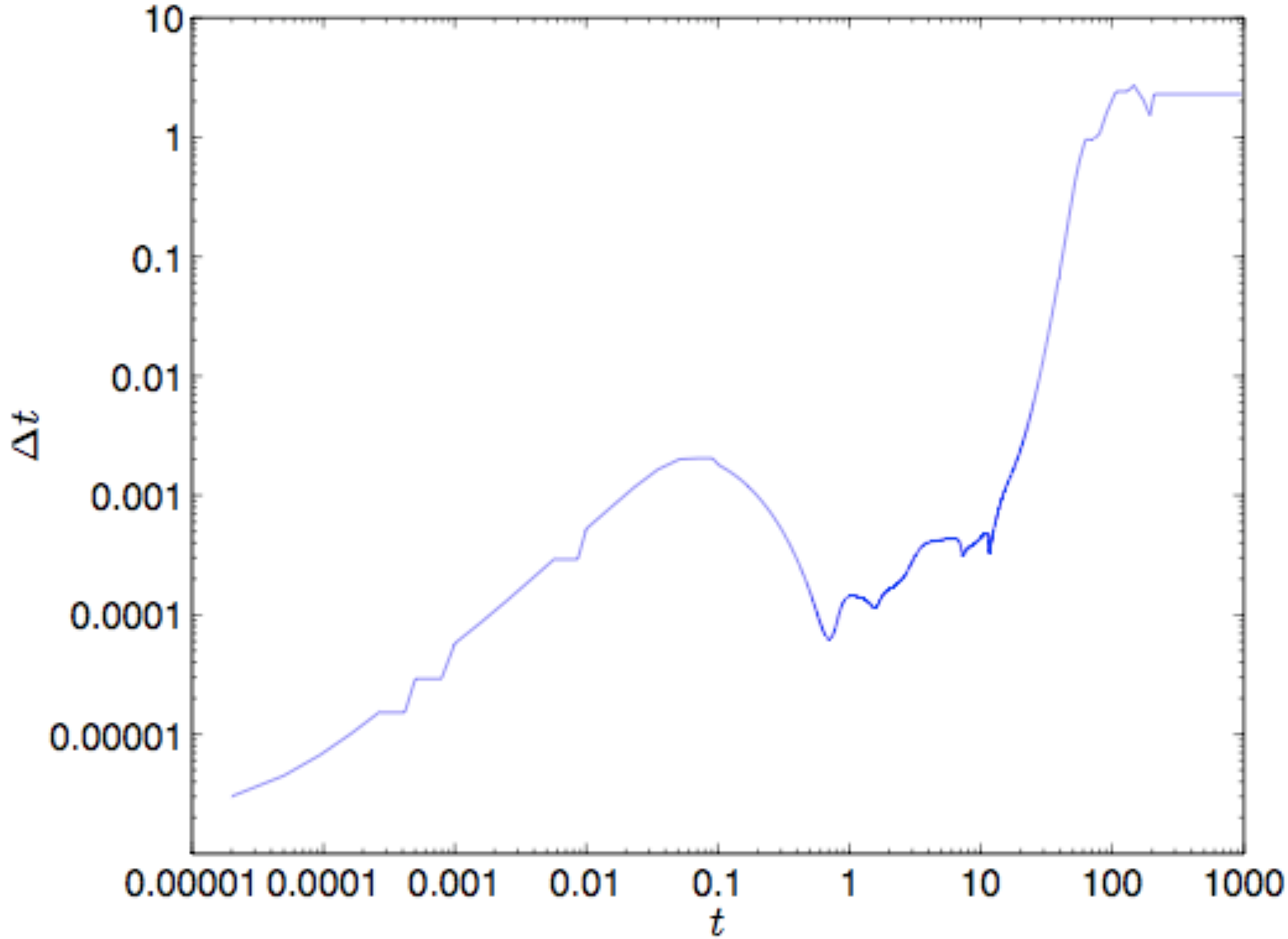
Test case: 3D, variable mobility

- The C-H equation

$$\frac{\partial u}{\partial t} + \epsilon^2 \nabla \cdot M(u) \nabla \Delta u + \theta \nabla \cdot M(u) \nabla u - (1/4) \Delta u = 0$$

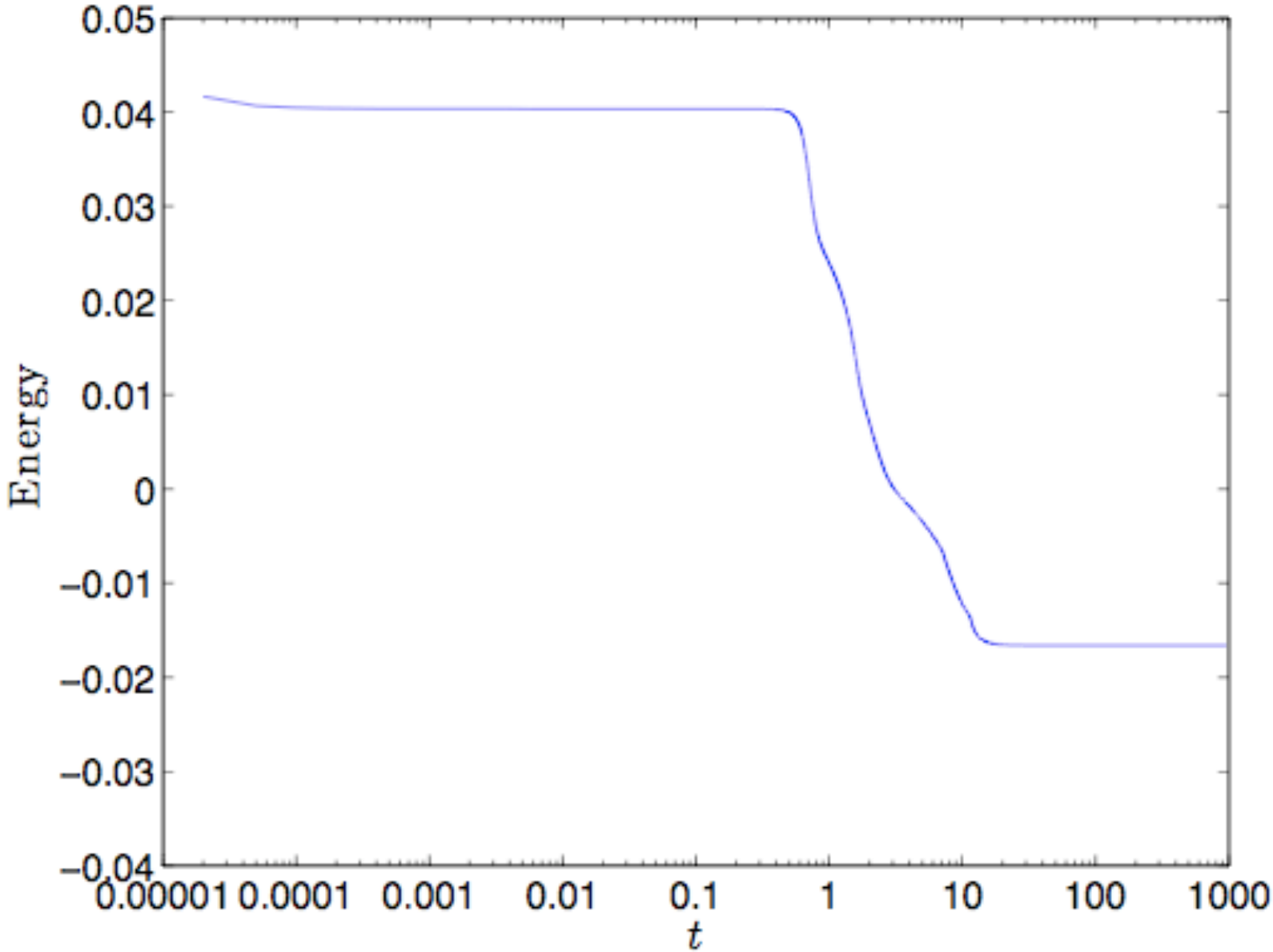
- Coefficients: $\theta = 3/2$, $\epsilon^2 = 1/800$
- Initial conditions: $u_0 = 0.26 + 0.05 * \text{rand}(-1, 1)$
- Boundary conditions: periodic
- Time integration: Backward Euler and NS schemes
- Time step size $\Delta t_0 = 8 \times 10^{-5}$ with adaptive Δt
- c.f. Gomez, Calo, Bazilevs and Hughes 2008 CMAME

Timestep build-up



c/o Chao Yang, CAS

Energy asymptotics



c/o Chao Yang, CAS

Strong scalability, BG/L

- LU subdomain solves, RAS with $\delta = 2$, $\Delta t = 8 \times 10^{-5}$

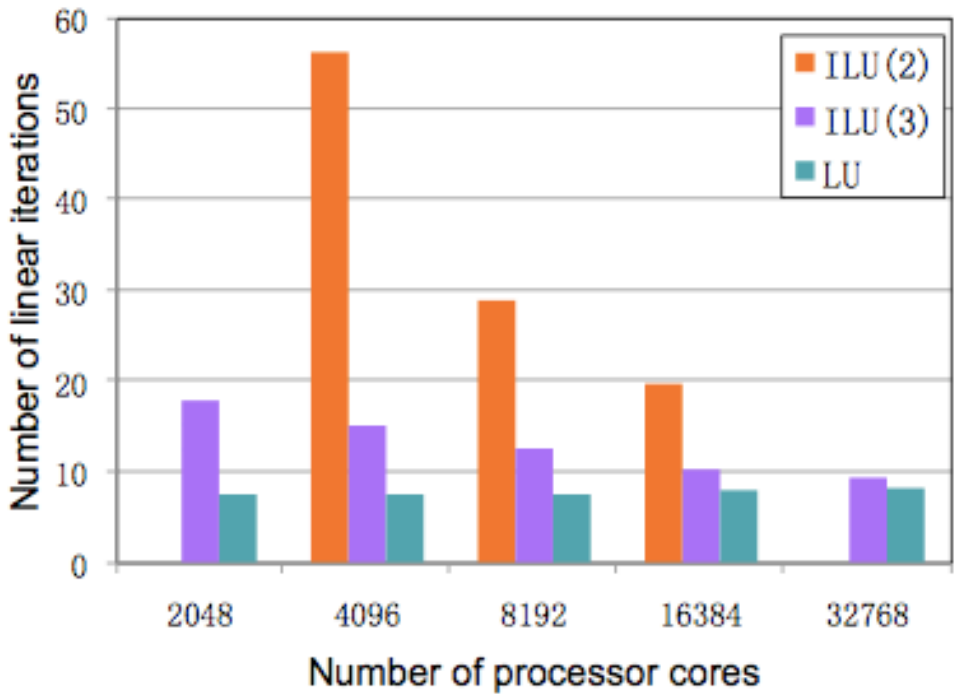
Mesh	np	$\frac{\text{Newton}}{\Delta t}$	$\frac{\text{GMRES}}{\text{Newton} \cdot \Delta t}$	Time (sec)
64^3	256	2.1	3.6	223.6
64^3	512	2.1	3.7	86.5
64^3	1024	2.1	3.8	37.8
Mesh	np	$\frac{\text{Newton}}{\Delta t}$	$\frac{\text{GMRES}}{\text{Newton} \cdot \Delta t}$	Time (sec)
128^3	1024	2.1	6.6	656.4
128^3	2048	2.1	6.7	247.6
128^3	4096	2.1	7.3	95.6

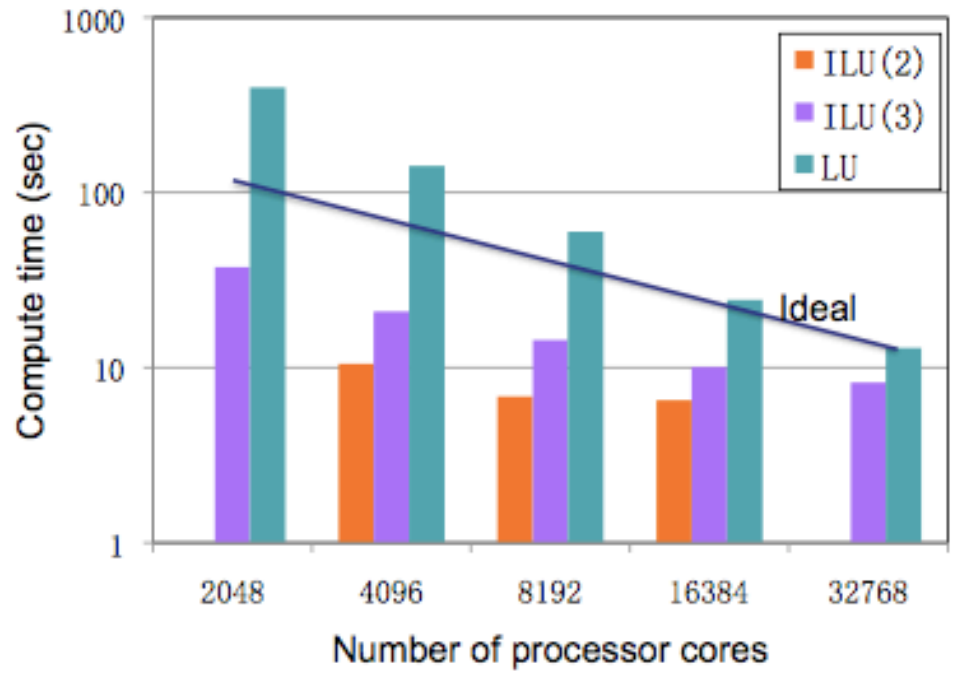
- ILU(3) subdomain solves, RAS with $\delta = 2$, $\Delta t = 1.25 \times 10^{-5}$

Mesh	np	$\frac{\text{Newton}}{\Delta t}$	$\frac{\text{GMRES}}{\text{Newton} \cdot \Delta t}$	Time (sec)
256^3	8192	2.1	13.4	36.51

c/o Chao Yang, CAS

Strong scalability, BG/P





- The one-level RAS scales well up to 32K processors in terms of iteration numbers
- LU subdomain solver provides superlinear speedup, though the compute time is large
- ILU subdomain solver are faster but the speedup is no longer superlinear
- A trade-off between LU and ILU, or other sequential solver, is needed in real applications

(now using BoomerAMG from hypre on the subdomains, for log-linear complexity subdomain solves)

c/o Xiao-Chuan Cai, CU-Boulder

Phase field crystal (PFC) model

- Free energy

$$E(u) = \int_\Omega \left(\frac{1}{4}u^4 + \frac{1-\eta}{2}u^2 - |\nabla u|^2 + \frac{1}{2}(\Delta u)^2 \right)$$

- Conserved dynamics: phase field crystal (PFC) equation

$$\frac{\partial u}{\partial t} - \Delta \frac{\delta E(u)}{\delta u} = 0 \longrightarrow \boxed{\frac{\partial u}{\partial t} - \Delta^3 u - 2\Delta^2 u - \Delta[u^3 + (1-\eta)u] = 0}$$

“Movie” (1): Cahn-Hilliard (4th order)

Movie (2): Phase field crystal (6th order)

Phase separation, current status

- **Within an adaptive nonlinearly stabilized timestepping scheme, standard Additive Schwarz leads to efficient massive parallelism, enabling phase separation problems to run to asymptotically long times**
- **Better preconditioners are sought prior to turning the Cahn-Hilliard simulator loose in a vast physical parameter space to determine optimal designs of self-assembling membranes**

Teams have become a "way to go"

The Increasing Dominance of Teams in Production of Knowledge

Stefan Wuchty,[1]* Benjamin F. Jones,[2]* Brian Uzzi[1,2]*†

We have used 19.9 million papers over 5 decades and 2.1 million patents to demonstrate that teams increasingly dominate solo authors in the production of knowledge. Research is increasingly done in teams across nearly all fields. Teams typically produce more frequently cited research than individuals do, and this advantage has been increasing over time. Teams now also produce the exceptionally high-impact research, even where that distinction was once the domain of solo authors. These results are detailed for sciences and engineering, social sciences, arts and humanities, and patents, suggesting that the process of knowledge creation has fundamentally changed.

An acclaimed tradition in the history and sociology of science emphasizes the role of the individual genius in scientific discovery (*1*, *2*). This tradition focuses on guiding contributions of solitary authors, such as Newton and Einstein, and can be seen broadly in the tendency to equate great ideas with particular names, such as the Heisenberg uncertainty principle, Euclidean geometry, Nash equilibrium, and Kantian ethics. The role of individual contributions is also celebrated through science's award-granting institutions, like the Nobel Prize Foundation (*3*).

Several studies, however, have explored an apparent shift in science from this individual-based model of scientific advance to a teamwork model. Building on classic work by Zuckerman and Merton, many authors have established a rising propensity for teamwork in samples of research fields, with some studies going back a century (*4*–*7*). For example, de Solla Price examined the change in team size in chemistry from 1910 to 1960, forecasting that in 1980 zero percent of the papers would be written by solo authors (*8*). Recently, Adams *et al.* established that over time, teamwork had increased across broader sets of fields among elite U.S. research universities (*9*). Nevertheless, the breadth and depth of this projected shift in manpower remains indefinite, particularly in fields where the size of experiments and capital investments remain small, raising the question as to whether the projected growth in teams is universal or cloistered in specialized fields.

A shift toward teams also raises new questions of whether teams produce better science. Teams may bring greater collective knowledge and effort, but they are known to experience social network and coordination losses that make them underperform individuals even in highly complex tasks (*10*–*12*), as F. Scott Fitzgerald concisely observed when he stated that "no grand idea was ever born in a conference" (*13*). From this viewpoint, a shift to teamwork may be a costly phenomenon or one that promotes low-impact science, whereas the highest-impact ideas remain the domain of great minds working alone.

We studied 19.9 million research articles in the Institute for Scientific Information (ISI) Web of Science database and an additional 2.1 million patent records. The Web of Science data covers research publications in science and engineering since 1955, social sciences since 1956, and arts and humanities since 1975. The patent data cover all U.S. registered patents since 1975 (*14*). A team was defined as having more than one listed author (publications) or inventor (patents). Following the ISI classification system, the universe of scientific publications is divided into three main branches and their constituent subfields: science and engineering (with 171 subfields), social sciences (with 54 subfields), and arts and humanities (with 27 subfields). The universe of U.S. patents was treated as a separate category (with 36 subfields). See the Supporting Online Material (SOM) text for details on these classifications.

For science and engineering, social sciences, and patents, there has been a substantial shift toward collective research. In the sciences, team size has grown steadily each year and nearly

Table 1. Patterns by subfield. For the three broad ISI categories and for patents, we counted the number (*N*) and percentage (%) of subfields that show (i) larger team sizes in the last 5 years compared to the first 5 years and (ii) RTI measures larger than 1 in the last 5 years. We show RTI measures both with and without self-citations removed in calculating the citations received. Dash entries indicate data not applicable.

		Increasing team size		RTI > 1 (with self-citations)		RTI > 1 (no self-citations)	
	N_{fields}	N_{fields}	%	N_{fields}	%	N_{fields}	%
Science and engineering	171	170	99.4	167	97.7	159	92.4
Social sciences	54	54	100.0	54	100.0	51	94.4
Arts and humanities	27	24	88.9	23	85.2	18	66.7
Patents	36	36	100.0	32	88.9	–	–

[1]Northwestern Institute on Complexity (NICO), Northwestern University, Evanston, IL 60208, USA. [2]Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: uzzi@northwestern.edu

Study of 20 million ISI papers since 1955 and 2 million patents:

- ✓ **Papers by multiple authors more than twice as likely to be cited as solo**
- ✓ **Teams are six times more likely to author "home run papers" (> 1000 citations)**
- ✓ **Average team size grows by 20% per decade (over past five decades)**

Stories from the Audience?

EOF