Development of a Scalable Parallel Solver for Macroscopic Extended MHD Modeling

A. H. Glasser, PSI Center, University of Washington
V. S. Lukin, Naval Research Laboratory

Presented at the 2009 APS/DPP and CEMM Meetings
Atlanta, Georgia, November 1 & 3, 2009

Scalable Parallel Solver Development Strategy

- Physics-based preconditioning
  - Divide and conquer, reduces size and improves diagonal dominance of matrices to be solve.
  - Similar to split time step, but wrapped inside a full nonlinear Newton-Krylov solve. Convergence requires accurate preconditioning.
- Need scalable method for solving reduced matrices.
  - FETI-DP: proven scalability, natural preconditioner, but limited to SPD matrices. No longer under development.
  - Static condensation, GMRES, additive Schwarz: more general and robust, scales up to moderate size.
  - Algebraic multigrid: remains to be investigated.
- General framework developed and tested. Requires problem-specific Schur complement in flux-source form.
- Sequence of increasingly complete model problems developed and tested.
  - Linear ideal MHD traveling waves in 2D.
  - Nonlinear, dissipative, traveling and standing MHD waves in 2D.
  - 1D cylindrical magnetic confinement, theta pinch, radial compression, nonlinear, dissipative.
  - 2D, FRC, numerical initial conditions.

Physics-Based Preconditioning

Factorization and Schur Complement

Linear System

$$\mathsf{L}\mathbf{u} = \mathbf{r}, \quad \mathsf{L} \equiv \begin{pmatrix} \mathsf{L}_{11} & \mathsf{L}_{12} \\ \mathsf{L}_{21} & \mathsf{L}_{22} \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}, \quad \mathbf{r} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix}$$

Factorization

$$\mathsf{L} \equiv \begin{pmatrix} \mathsf{L}_{11} & \mathsf{L}_{12} \\ \mathsf{L}_{21} & \mathsf{L}_{22} \end{pmatrix} = \begin{pmatrix} \mathsf{I} & \mathsf{0} \\ \mathsf{L}_{21}\mathsf{L}_{11}^{-1} & \mathsf{I} \end{pmatrix} \begin{pmatrix} \mathsf{L}_{11} & \mathsf{0} \\ \mathsf{0} & \mathsf{S} \end{pmatrix} \begin{pmatrix} \mathsf{I} & \mathsf{L}_{11}^{-1}\mathsf{L}_{12} \\ \mathsf{0} & \mathsf{I} \end{pmatrix}$$

Schur Complement

$$\mathsf{S} \equiv \mathsf{L}_{22} - \mathsf{L}_{21}\mathsf{L}_{11}^{-1}\mathsf{L}_{12}$$

Exact and Approximate Inverse

Preconditioned Krylov Iteration

Inverse

$$\mathbf{L}^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{L}_{11}^{-1}\mathbf{L}_{12} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{L}_{21}\mathbf{L}_{11}^{-1} & \mathbf{I} \end{pmatrix}$$

Exact Solution

$$\mathbf{s}_1 = \mathbf{L}_{11}^{-1}\mathbf{r}_1, \quad \mathbf{s}_2 = \mathbf{r}_2 - \mathbf{L}_{21}\mathbf{s}_1$$

$$\mathbf{u}_2 = \mathbf{S}^{-1}\mathbf{s}_2, \quad \mathbf{u}_1 = \mathbf{s}_1 - \mathbf{L}_{11}^{-1}\mathbf{L}_{12}\mathbf{u}_2$$

Preconditioned Krylov Iteration

$$\mathbf{P} \approx \mathbf{L}^{-1}, \quad (\mathbf{L}\mathbf{P})\left(\mathbf{P}^{-1}\mathbf{u}\right) = \mathbf{r}$$

Outer iteration preserves full nonlinear accuracy.
Need approximate Schur complement **S**
and scalable solution procedure for $\mathbf{L}_{11}$ and **S**.

Ideal MHD Waves

Linearized, Normalized Equations

$$\frac{\partial p}{\partial t} + \gamma \nabla \cdot \mathbf{v} = 0, \quad \frac{\partial \mathbf{b}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B})$$

$$\frac{\partial \mathbf{v}}{\partial t} + \nabla \cdot \mathsf{T} = 0, \quad \mathsf{T} = (\beta p + \mathbf{B} \cdot \mathbf{b})\,\mathsf{I} - \mathbf{B}\mathbf{b} - \mathbf{b}\mathbf{B}$$

Approximate Schur Complement

$$\mathsf{S}\mathbf{v} = \mathbf{v} + \nabla \cdot \mathsf{T},$$

$$\mathsf{T} \equiv h^2\theta^2 \left\{ [\mathbf{B} \cdot \nabla \times (\mathbf{v} \times \mathbf{B}) - \gamma\beta\nabla \cdot \mathbf{v}]\,\mathsf{I} - \mathbf{B}\nabla \times (\mathbf{v} \times \mathbf{B}) - \nabla \times (\mathbf{v} \times \mathbf{B})\,\mathbf{B} \right\}$$

Ideal MHD Schur Complement, 1

Evaluation of T_{kl}

$$T_{kl} = T_{lk} = \nabla x_k \cdot \mathsf{T} \cdot \nabla x_l = \nabla x_k \cdot \{[\mathbf{B} \cdot \nabla \times (\mathbf{v} \times \mathbf{B}) - \gamma p \nabla \cdot \mathbf{v}] \mathsf{I}$$
$$- \mathbf{B} \nabla \times (\mathbf{v} \times \mathbf{B}) - \nabla \times (\mathbf{v} \times \mathbf{B}) \mathbf{B}\} \cdot \nabla x_l = \partial_i v_j S_{ijkl} + v_j R_{jkl}$$

Evaluation of S_{ijkl}

$$\begin{aligned}
S_{ijkl} =& S_{ijlk} = \frac{\partial T_{kl}}{\partial(\partial_i v_j)} \\
=& \nabla x_k \cdot \{[\mathbf{B} \cdot \nabla x_i \times (\nabla x_j \times \mathbf{B}) - \gamma p \delta_{ij}] \mathsf{I} \\
& - \mathbf{B} \nabla x_i \times (\nabla x_j \times \mathbf{B}) - \nabla x_i \times (\nabla x_j \times \mathbf{B}) \mathbf{B}\} \cdot \nabla x_l \\
=& \nabla x_k \cdot \left\{ \left[n_i n_j - (B^2 + \gamma p) \delta_{ij} \right] \mathsf{I} \right. \\
& \left. + 2\mathbf{B}\mathbf{B} \delta_{ij} - n_i (\mathbf{B} \nabla x_j + \nabla x_j \mathbf{B}) \right\} \cdot \nabla x_l \\
=& \left[n_i n_j - (B^2 + \gamma p) \delta_{ij} \right] \delta_{kl} + 2 n_k n_l \delta_{ij} - n_i (n_k \delta_{jl} + n_l \delta_{jk})
\end{aligned}$$

Ideal MHD Schur Complement, 2

Pressure Gradient Schur Terms

$$\mathbf{T}_P = -\mathsf{I}\mathbf{v} \cdot \nabla P$$

$$\mathbf{T}_{Pij} = -\delta_{ij} v_k \partial_k P$$

Current Schur Terms

$$\mathbf{S}_J \equiv \mathbf{J} \times \frac{\partial \mathbf{b}}{\partial t} = \mathbf{J} \times [\nabla \times (\mathbf{v} \times \mathbf{B})]$$

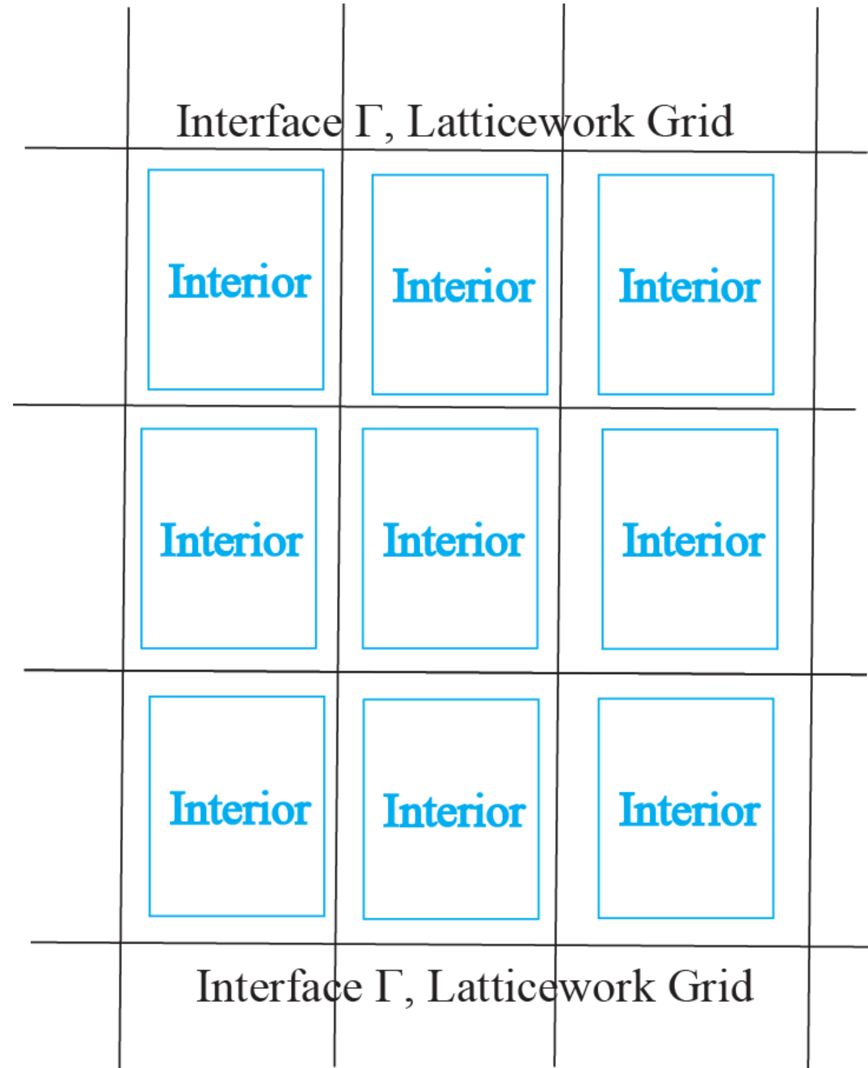
$$S_{Ji} = \mathbf{S}_J \cdot \nabla x_i = \partial_j v_k \Sigma_{ijk}$$

$$\begin{aligned}\Sigma_{ijk} =& J_l B_m \nabla x_l \times [\nabla x_j \times (\nabla x_k \times \nabla x_m)] \cdot \nabla x_i \\ =& J_l B_m (\nabla x_i \times \nabla x_l) \cdot [\nabla x_j \times (\nabla x_k \times \nabla x_m)] \\ =& J_l B_m \left(\delta_{ij} \epsilon_{lkm} - \delta_{lj} \epsilon_{ikm}\right)\end{aligned}$$

$$\partial_j v_k = \frac{\partial_j (\rho v_k)}{\rho} - \frac{(\partial_j \rho)(\rho v_k)}{\rho^2}$$

Static Condensation

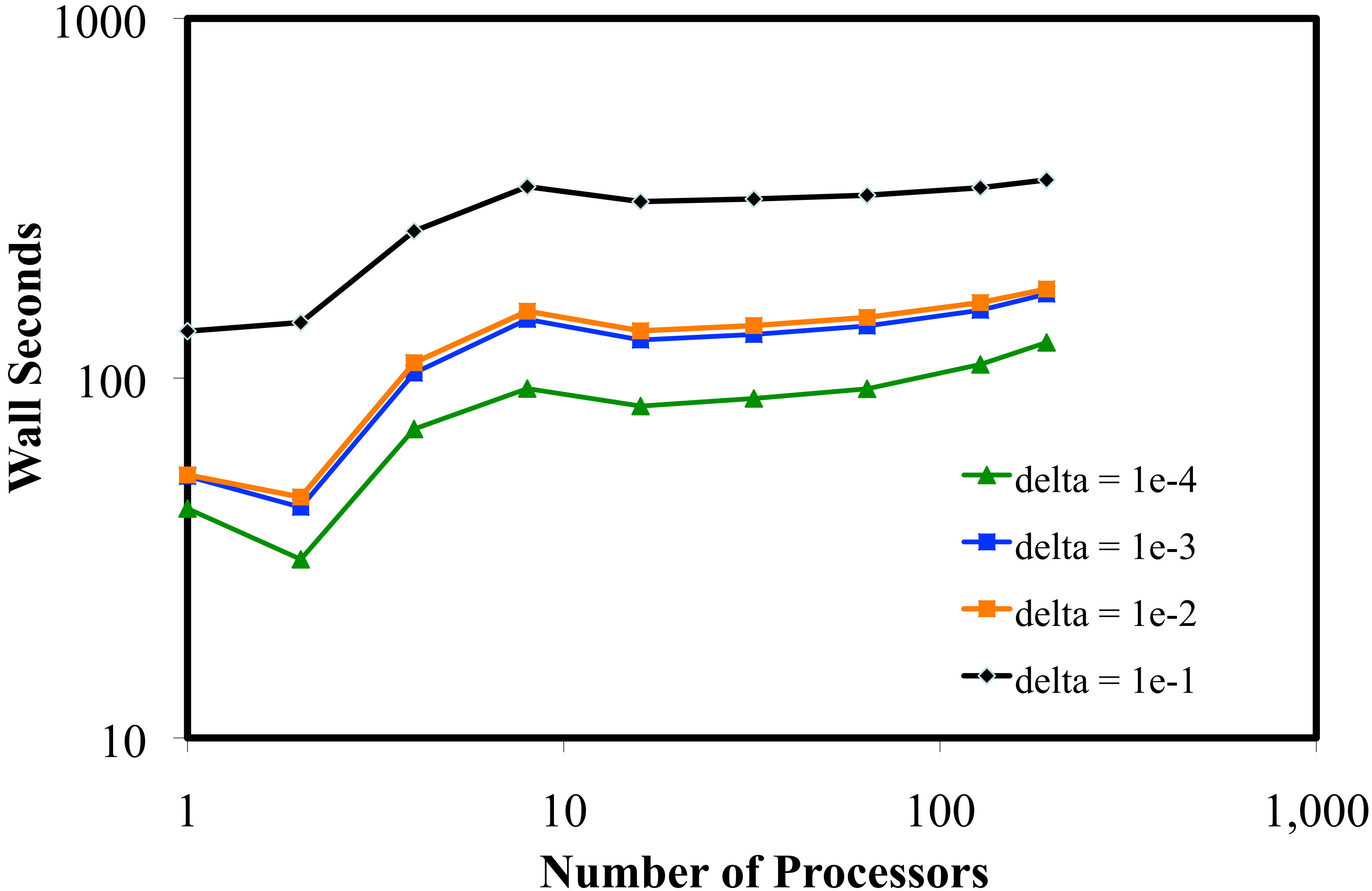
- Implicit time step requires linear system solution: **L u = r**.
- Direct solution time grows as n^3.
- Break up large matrix into smaller pieces: Interiors + Interface.
- Small direct solves for interior.
- Interface solve by CG or GMRES, precoditioned with LU or ILU(k) on each processor, with Schwarz overlap between processors.
- Substantially reduces solution time, condition number.



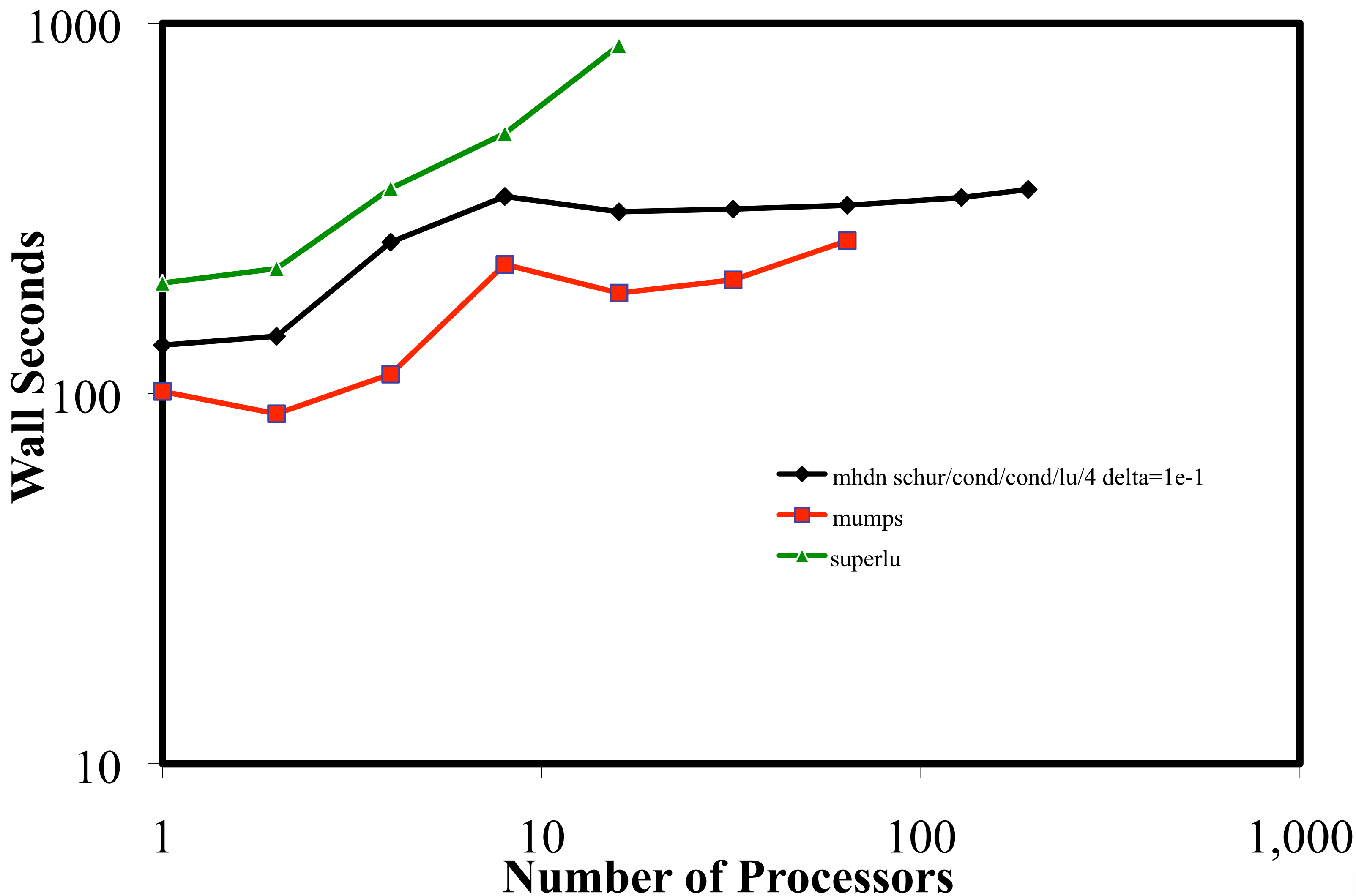
Nonlinear, Dissipative Wave Test Problem

- Nonlinear, dissipative, standing or traveling MHD waves in a doubly periodic uniform plane.
- 2D **k** vector in computational plane, 3D **B** vector specified by spherical angles about normal to plane. Continuous control of angle θ between **k** and **B.**
- Initialize to pure linear eigenvector: fast, shear, or slow wave.
- Unit cell: 1 full wavelength in each direction, nx = ny = 8, np = 6, nqty = 8.
- Weak scaling test case: each processor has one unit cell.
  Nonlinear amplitude delta, resistivity, viscosity, thermal conductivity to damp nonlinear coupling to high frequency, short-wavelength modes.
- 1 – 192 processors on PSI Center Ice cluster.
- Largest test problem size:
  - 128x96 unit cells, 192 processors. 2-6 minutes of wall time.
  - 3.8M dependent variables, 64 large time steps. For large delta, multiple Jacobian evaluations.

Wall Time to Solution, Schur Solve
Nonlinear, Dissipative MHD Slow Traveling Wave



Wall Time to Solution
Comparison to Direct Solvers



Comments on Nonlinear, Dissipative Wave Test Problem

- Erratic scaling up to 16 processors, then smoothly scales up to 192.
- Increasing nonlinear wave amplitude requires substantial increase in effort due to larger number of Jacobian evaluations, but no degradation in scaling.
- Deviation from perfect scaling: (wall time) = (nproc) $^{\gamma}$
  Perfect: $\gamma = 0$. Actual: $\gamma = 0.13$. Disclaimer: limited to nproc = 192.
- Memory requirement primarily due to computed and stored sparse matrix, very small, scales up linearly, requires much less than available.
- Capable of treating entire 3D problem.
- Direct solvers: SuperLU and MUMPS, condensed matrix, worse time scaling, run out of memory. Only capable of preconditioning 3D solve.

1D Magnetic Confinement Model, Theta Pinch

$$\begin{aligned}\frac{\partial B_z}{\partial r} &= -\alpha r\left(r^2-r_1^2\right)\left(r^2-r_2^2\right)\\ &= -\alpha\left[r^5-\left(r_1^2+r_2^2\right)r^3+r_1^2r_2^2r\right]\end{aligned}$$

$$B_z(r) = 1-\frac{\alpha r^2}{12}\left[2r^4-3\left(r_1^2+r_2^2\right)r^2+6r_1^2r_2^2\right]$$

$$\begin{aligned}\psi(r) &= -\int_0^r B_z(r')r'dr' = -\frac{1}{2}\int_0^{r^2} B_z(x)dx, \quad x\equiv r'^2\\ &= -\frac{r^2}{2}\left\{1-\frac{\alpha r^2}{24}\left[r^4-2\left(r_1^2+r_2^2\right)r^2+6r_1^2r_2^2\right]\right\}\end{aligned}$$

$$p(r)+\frac{1}{2}B_z^2(r) = \frac{1}{2}(1+\beta_0)$$

$$\beta(r) = 2p(r) = \beta_0+1-B_z^2(r)$$

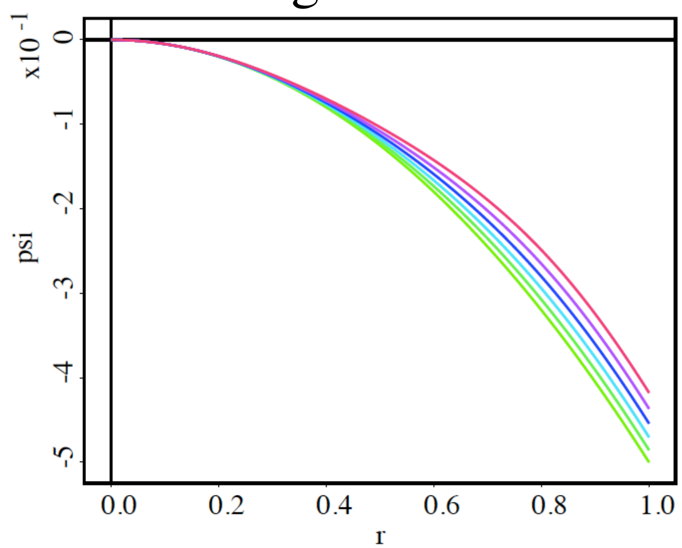
$$\Delta\beta \equiv \beta(r_1)-\beta_0 = 1-B_z^2(r_1)$$

$$r_1 = \frac{1}{\sqrt{3}}, \quad r2 = 1, \quad \alpha = \frac{81}{5}\left[1-(1-\Delta\beta)^{1/2}\right]$$

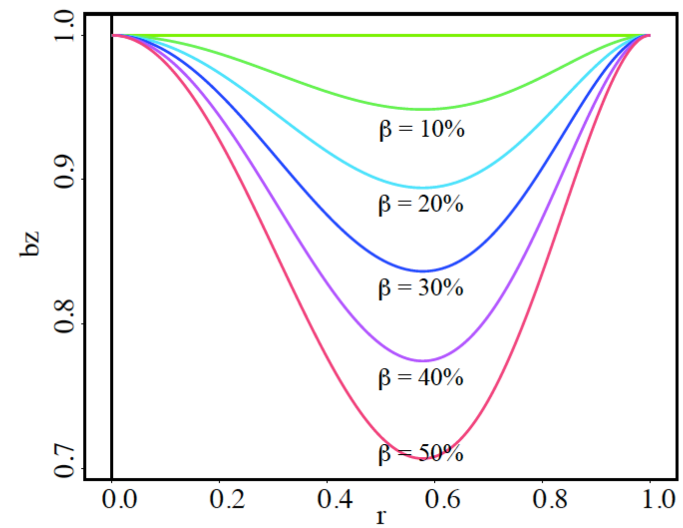
Time dependence due to resistive decay of magnetic field or radial compression.

1D Magnetic Confinement Model, Theta Pinch

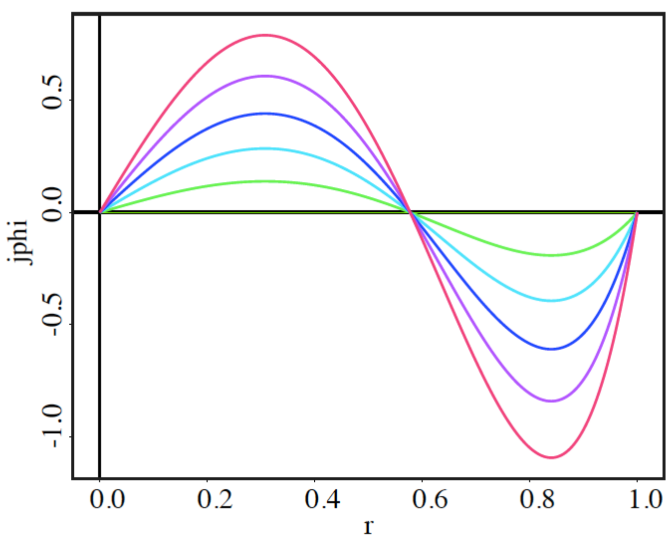
Magnetic Flux



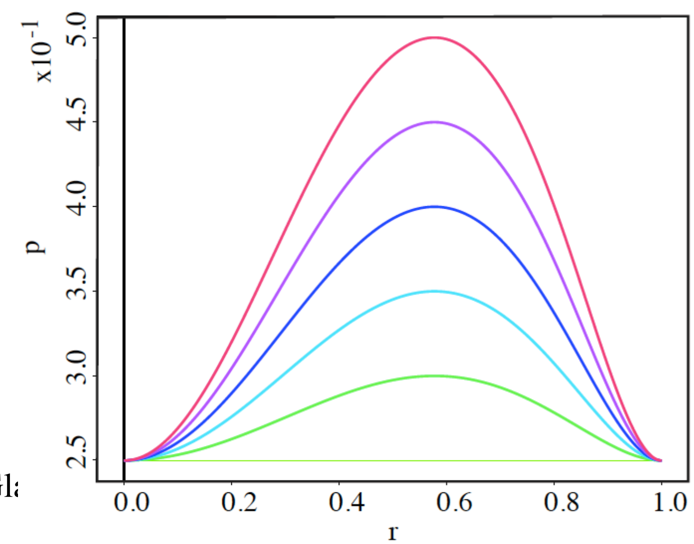
Axial Magnetic Field



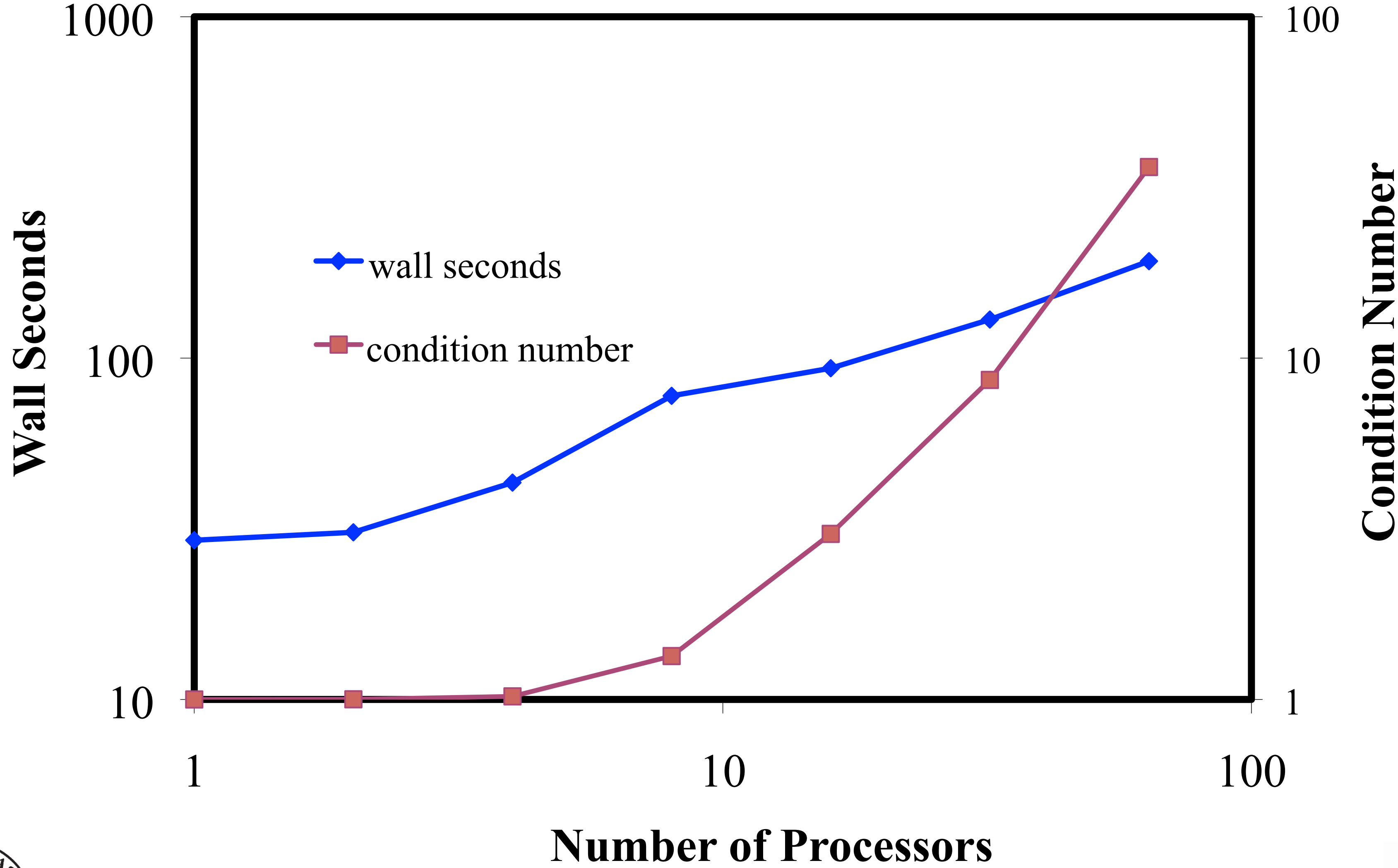
Azimuthal Current



Pressure



Wall Time to Solution
1D Magnetic Confinement



Comments on 1D Magnetic Confinement Test

- Deviation from perfect scaling: (wall time) = (nproc) $^{\gamma}$
  Perfect: $\gamma = 0$. Actual: $\gamma = 0.57$. Much worse than 2D wave test.
- Cause of poor scaling: increasing condition number of Schur complement, GMRES iterations. Not scalable.
- Why does this show up for this test but not for 2D wave tests?
  - 1D initial conditions, 1D scaling of grid, only ny scales up, by factor of 64.
  - 2D initial conditions, 2D scaling of grid, nx and ny scale up by factors of 16 and 12.
  - Condition number of Schur complement scales as $\omega_{fast}^2 = (k_r^2+k_z^2)*(\omega_A^2+\omega_S^2)$.
    Gets much larger in 1D scaling test.
  - Slow time scale also scales up in wave test, but not in 1D confinement test.
- 1D case not of practical interest, 2D and 3D cases scale up to reasonable values.
- FETI-DP scalable but only for SPD matrices. GMRES not scalable. Geometric multigrid scalable but not applicable to spectral elements. Algebraic multigrid?

Additional Tests and Future Plans

- 2D numerical initial conditions, FRC, Grad-Shafranov solution, George Marklin. Schur complement solution procedure works correctly, but with excessive Newton iterations, indicating inaccurate Schur complement, not large condition number. Not yet diagnosed.
- Algebraic multigrid will be investigated for improved scalability in cases where condition number is an issue.
  BoomerAMG, Hypre, PETSc.
- 3D: HiFi and other codes. Since physics-based preconditioning involves physical rather than geometric decomposition, and doesn't require large memory, extension to 3D should be straightforward.