School of Computing

UNIVERSITY OF LEEDS

"Parallel, adaptive, multilevel solution of nonlinear systems arising in phase-field problems"

Peter Jimack

School of Computing, University of Leeds

Joint with: Peter Bollada, Chris Goodyer and Andrew Mullis

Recent Advances in Parallel and High Performance Computing Techniques and Applications – NUS, IMS, Jan 12-16, 2015

Introduction

• Overview of talk

• This talk focuses on a specific phase-field problem:

  ➢ Modelling the solidification of a dilute binary alloy, involving multiple length and time scales.

• However this requires us to consider some overarching issues that apply equally to other phase-field systems:

  ➢ Implicit time-stepping;

  ➢ Multilevel solvers (at each time step);

  ➢ 3D mesh adaptivity (considered at each time step);

  ➢ Combining adaptivity, multigrid and parallel solution...

A Thermal-Solute Phase-Field Model

- J.C. Ramirez, C. Beckermann, A. Karma, et al, Phys. Rev. E, 69 (2004) 051607.

We consider a 3-d extension of one particular (2-d) P-F model for the solidification of a dilute binary alloy:

- A phase equation (nonlinear)
- A chemical concentration diffusion equation (nonlinear)
- A heat diffusion equation (linear)

Lewis number (Le) = (thermal diffusivity)/(chemical diffusivity)

=> *stiffness*

These are solved to obtain three (time-dependent) fields...

- φ (phase), u (concentration) and ϑ (temperature)

A Thermal-Solute Phase-Field Model

- Phase equation shown in 2-d

$$A(\psi)^2\left[\frac{1}{Le}+Mc_\infty[1+(1-k)U]\right]\frac{\partial\phi}{\partial t}=A(\psi)^2\nabla^2\phi+2A(\psi)A'(\psi)\left[\frac{\partial\psi}{\partial x}\frac{\partial\phi}{\partial x}+\frac{\partial\psi}{\partial y}\frac{\partial\phi}{\partial y}\right]$$
$$-\frac{\partial}{\partial x}\left(A(\psi)A'(\psi)\frac{\partial\phi}{\partial y}\right)+\frac{\partial}{\partial y}\left(A(\psi)A'(\psi)\frac{\partial\phi}{\partial x}\right)+$$
$$\phi-\phi^3-\lambda(1-\phi^2)^2(\theta+Mc_\infty U)$$

Phase Equation

Properties:

- highly nonlinear
- noise introduced by anisotropy function A(Ψ)
- where $\psi=\arctan(\phi_y/\phi_x)$

A Thermal-Solute Phase-Field Model

- Concentration and temperature equations shown in 2-d

$$\left(\frac{1+k}{2}-\frac{1-k}{2}\phi\right)\frac{\partial U}{\partial t}=D\left(-\frac{1}{2}\left[\frac{\partial \phi}{\partial x}\frac{\partial U}{\partial x}+\frac{\partial \phi}{\partial y}\frac{\partial U}{\partial y}\right]+\frac{1-\phi}{2}\nabla^2 U\right)+$$

$$\frac{1}{2\sqrt{2}}\left(\{1+(1-k)U\}\left(\frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial t}\frac{\phi_x}{|\nabla \phi|}\right)+\frac{\partial}{\partial y}\left(\frac{\partial \phi}{\partial t}\frac{\phi_y}{|\nabla \phi|}\right)\right)\right.$$

$$\left.+(1-k)\left(\frac{\partial U}{\partial x}\left(\frac{\partial \phi}{\partial t}\frac{\phi_x}{|\nabla \phi|}\right)+\frac{\partial U}{\partial y}\left(\frac{\partial \phi}{\partial t}\frac{\phi_y}{|\nabla \phi|}\right)\right)\right)$$

$$+\frac{1}{2}\left((1+(1-k)U)\frac{\partial \phi}{\partial t}\right)$$

Concentration Equation

Temperature Equation

$$\frac{\partial \theta}{\partial t}=\alpha\nabla^2\theta+\frac{1}{2}\frac{\partial \phi}{\partial t}$$

A Thermal-Solute Phase-Field Model

- Concentration and temperature equations shown in 2-d

$$\left(\frac{1+k}{2}-\frac{1-k}{2}\phi\right)\frac{\partial U}{\partial t}=D\left(-\frac{1}{2}\left[\frac{\partial\phi}{\partial x}\frac{\partial U}{\partial x}+\frac{\partial\phi}{\partial y}\frac{\partial U}{\partial y}\right]+\frac{1-\phi}{2}\nabla^{2}U\right)+$$

$$\frac{1}{2\sqrt{2}}\left(\{1+(1-k)U\}\left(\frac{\partial}{\partial x}\left(\frac{\partial\phi}{\partial t}\frac{\phi_x}{|\nabla\phi|}\right)+\frac{\partial}{\partial y}\left(\frac{\partial\phi}{\partial t}\frac{\phi_y}{|\nabla\phi|}\right)\right)\right.$$

$$\left.+(1-k)\left(\frac{\partial U}{\partial x}\left(\frac{\partial\phi}{\partial t}\frac{\phi_x}{|\nabla\phi|}\right)+\frac{\partial U}{\partial y}\left(\frac{\partial\phi}{\partial t}\frac{\phi_y}{|\nabla\phi|}\right)\right)\right)$$

$$+\frac{1}{2}\left((1+(1-k)U)\frac{\partial\phi}{\partial t}\right)$$

D

Concentration Equation

α

Temperature Equation

$$\frac{\partial\theta}{\partial t}=\alpha\nabla^{2}\theta+\frac{1}{2}\frac{\partial\phi}{\partial t}$$

- Le = α / D

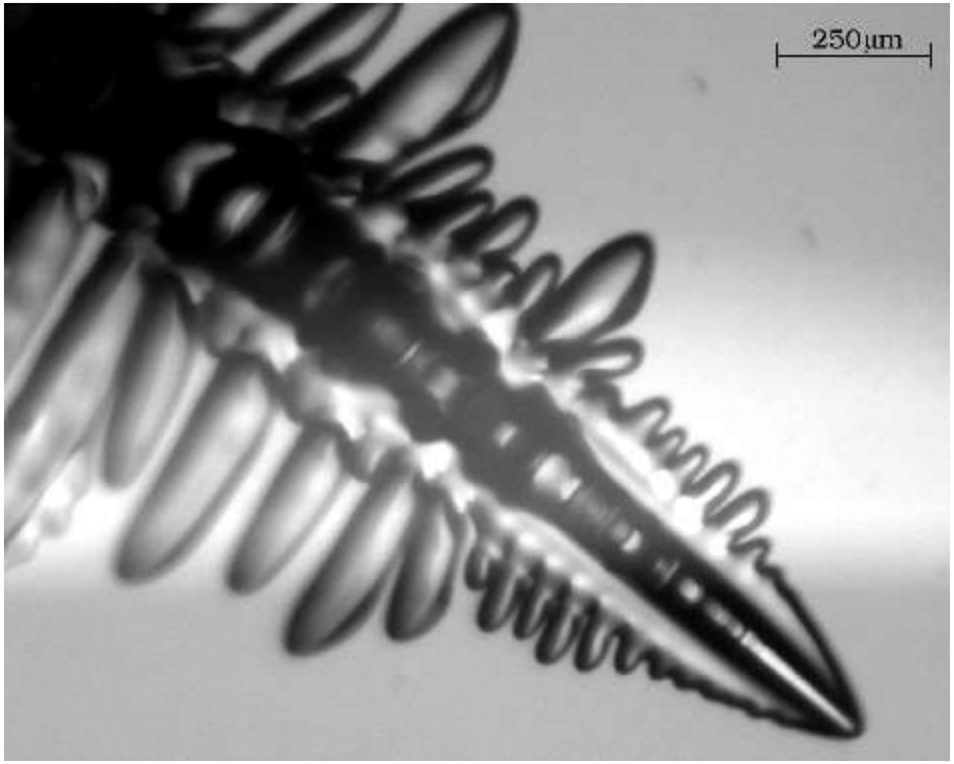
Dendritic Solutions in Three Dimensions

• Illustration of a real dendrite

• This is a snapshot of the dendritic crystal structure that can arise when rapid solidification occurs.

• This is a xenon crystal (Singer & Bilgram 2004 *Europhys. Lett.* **68** 240).

• Similar structures occur in metallic alloys...



Dendritic Solutions in Three Dimensions

• Illustration of a time-dependent simulation

• Animation of a typical solution.

•Plots of the $\varphi = 0$ isosurface at different time intervals.

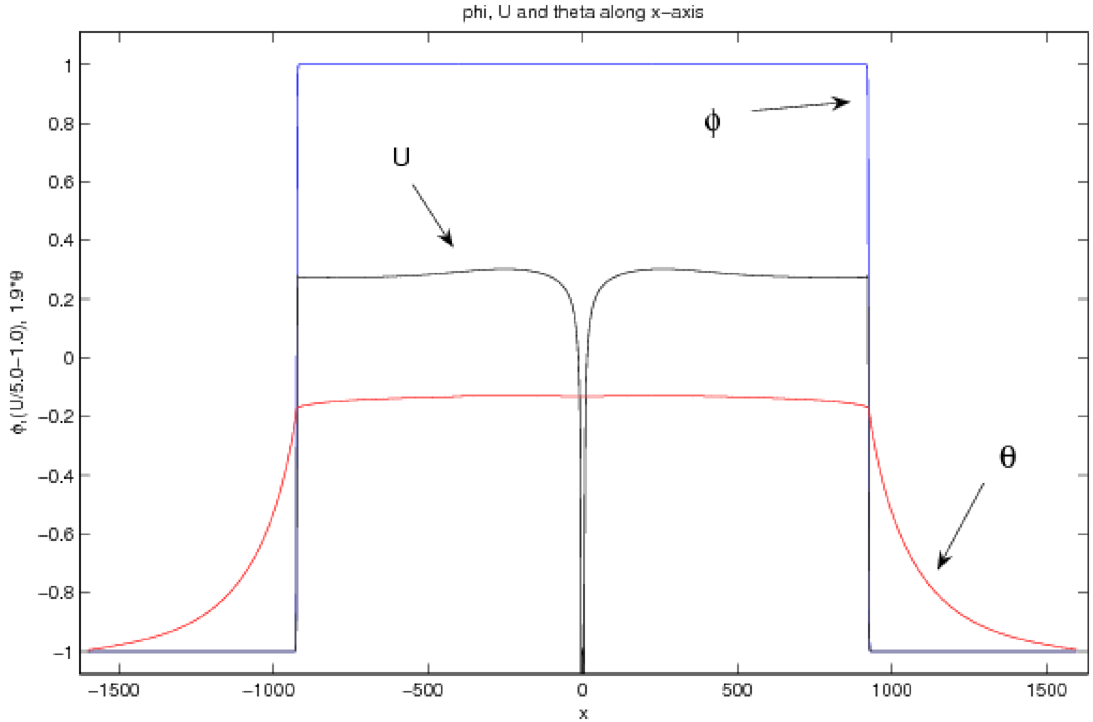
•Begins with a small solid seed.

•Preferred growth directions through our choice of anisotropy function...

Dendritic Solutions in Three Dimensions

• Cross-section of a typical solution

Cross-section of typical solution



Large values of the Lewis number lead to a multiscale problem that is highly stiff

Summary of Computational Challenges

• All of these issues are significant for quantitative predictions

• Should include both thermal and chemical diffusion for models of metallic alloys:

  - ➢ leads to a stiff system making explicit solution impossible

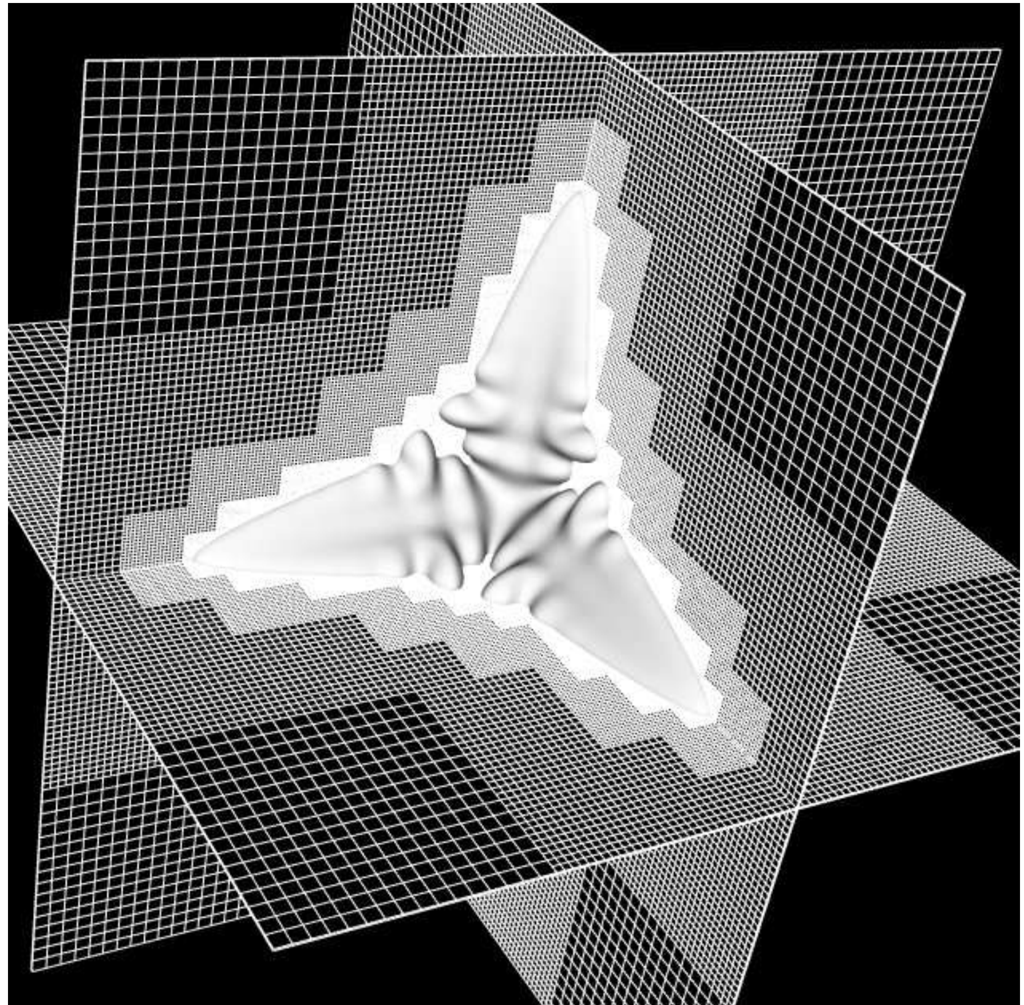
• Results are interface-width dependent unless this is sufficiently small: need to be able to resolve this interface.

  - ➢ Very fine mesh required at moving interface
  - ➢ But need very large domain for the thermal field ahead of the interface

• Everything needs to be done in three space dimensions of course...

Spatial Adaptivity in Three Dimensions

• Snapshot of adaptive mesh refinement

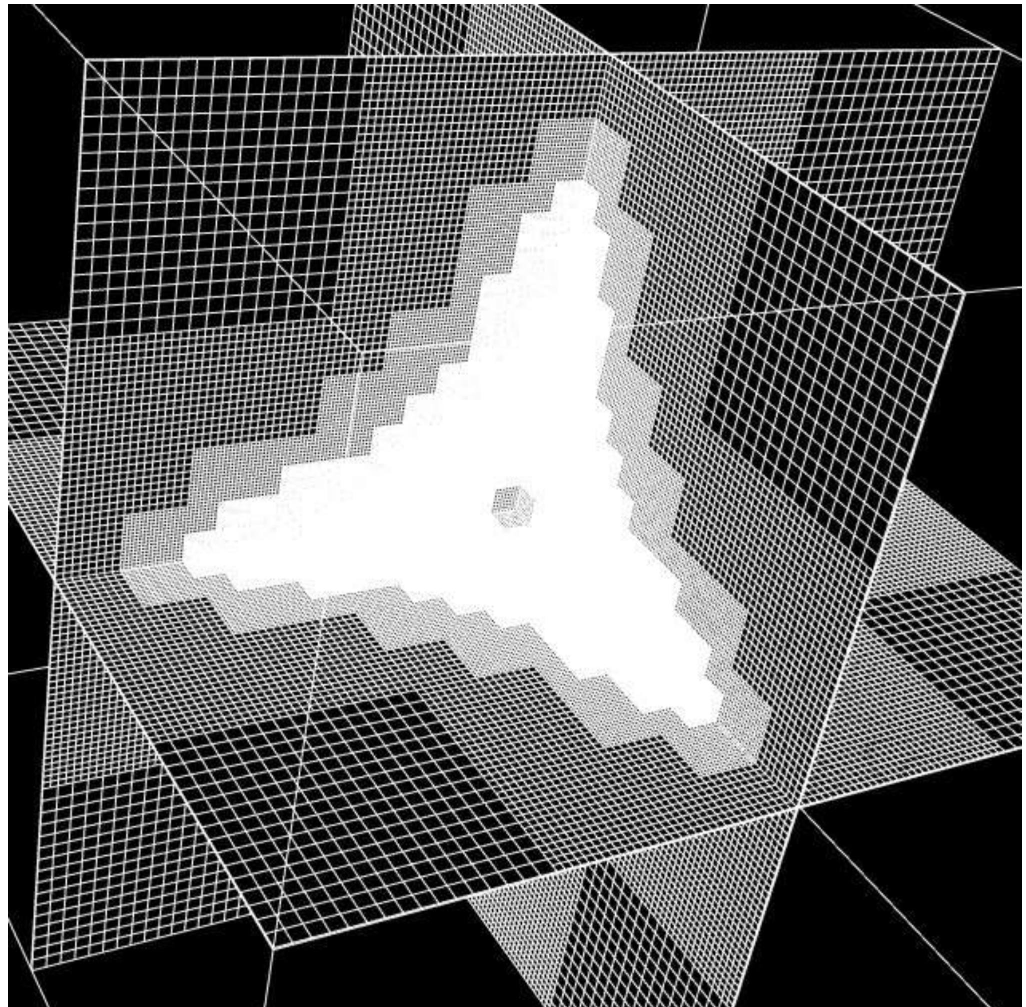


- Mesh adaptivity is clearly essential.
- Here we see local mesh refinement around the $\varphi = 0$ isosurface.
- Implementation is based upon the Open Source PARAMESH library (MacNeice & Olsen).

Spatial Adaptivity in Three Dimensions

- Snapshot of adaptive mesh refinement

- Here we see only the adapted mesh – which includes coarsening behind the interface.
- Based upon oct-tree of blocks (16x16x16 in this case).
- Use just one ghost layer (even for 19-pt stencil)



Implicit Time-Stepping and Nonlinear Multigrid

• Application of the nonlinear multigrid solver for adaptive meshes

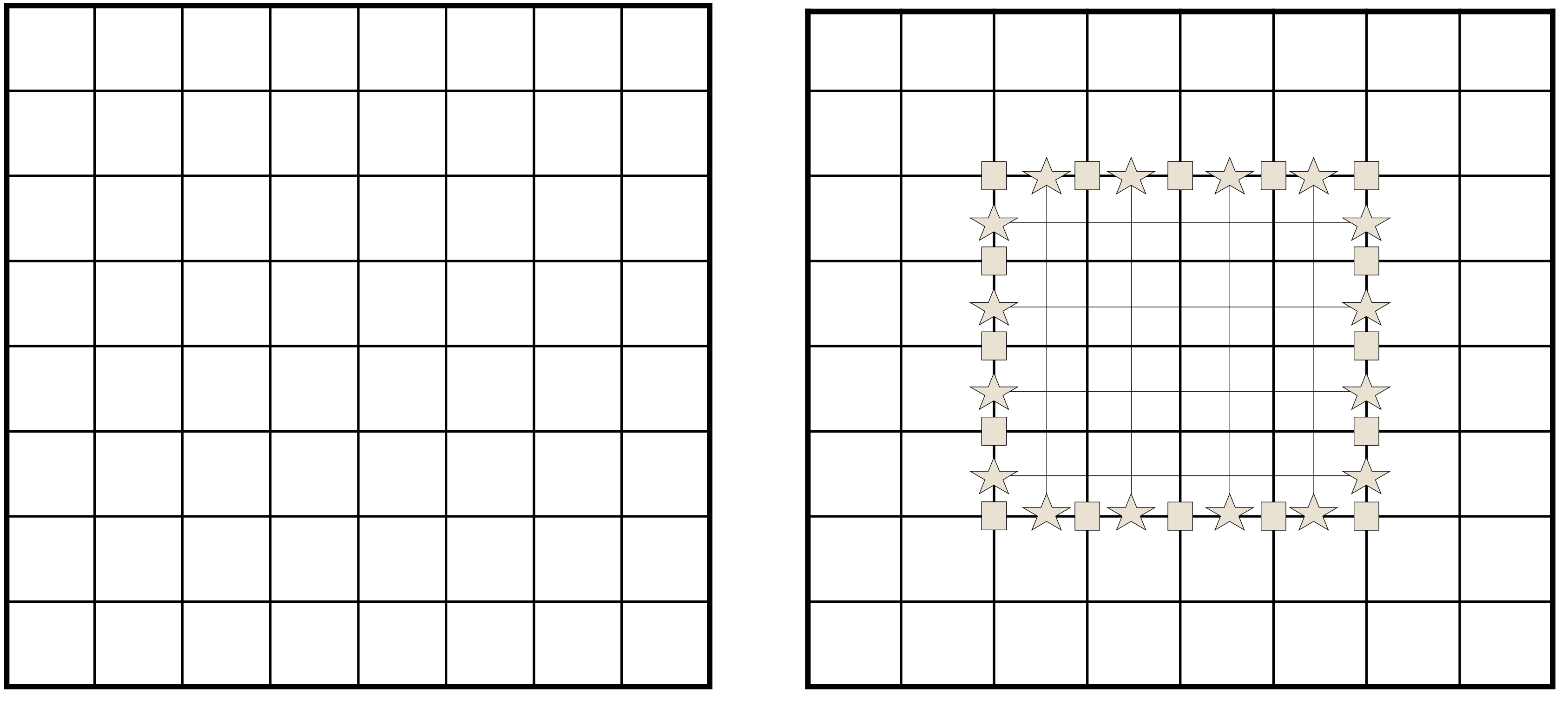
At each time step (we use BDF2) a large nonlinear algebraic system of equations is solved for the new values: ϕ_{ijk}^{n+1}, U_{ijk}^{n+1} and θ_{ijk}^{n+1}.

- A f*ully coupled* nonlinear Multigrid solver is used to for this:
  - ➢ based upon the FAS (full approximation scheme) approach to resolve the non-linearity;
  - ➢ and the MultiLevel AdapTive (MLAT) scheme of Brandt to handle the adaptivity;
  - ➢ a weighted nonlinear block Jacobi iterative scheme is seen to be an adequate smoother.
- Optimal, *h*-independent, convergence results are obtained.

Adaptivity and Multigrid

- The MLAT scheme

For the MLAT scheme the nodes at the interface between refinement levels are treated as a Dirichlet boundary by the smoother…



Parallel Implementation Issues

• Adaptivity and multigrid within a parallel solver

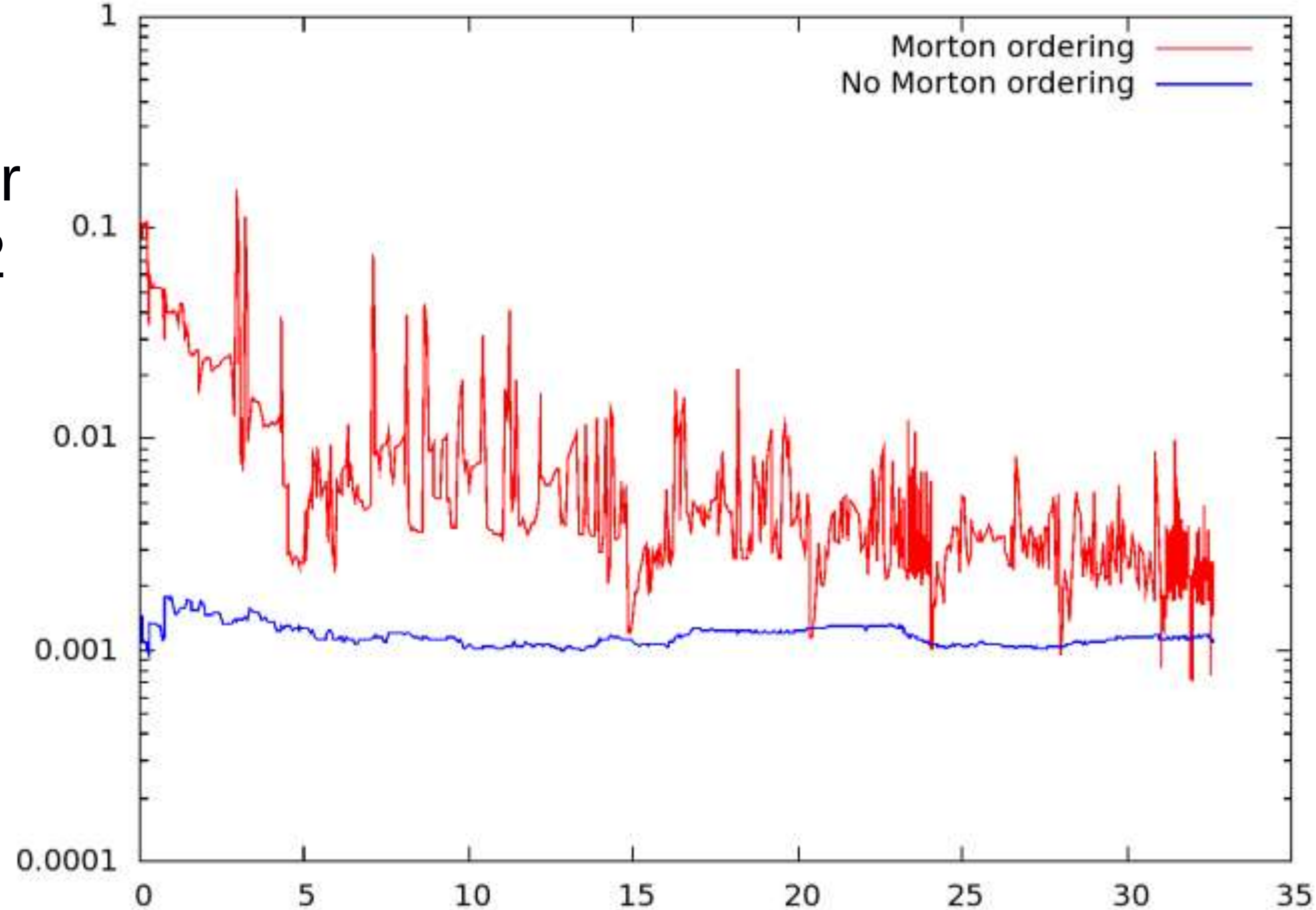
- Dynamic load-balancing when re-meshing occurs:
  - ➢ Nested hierarchy of hexahedral blocks (e.g. 8x8x8);
  - ➢ Each block has a single ghost layer at each edge (so 10x10x10).
- Parallel multigrid implementation:
  - ➢ Geometric MG visits one mesh level at a time;
  - ➢ Load-balancing and grid-transfer operations must reflect this.
- Coarse grid problem is expensive to solve in FAS:
  - ➢ Must make coarsest level as coarse as possible.
  - ➢ Typically this will imply idle cores at coarsest level(s).

Parallel Implementation Issues

• Improved dynamic load-balancing strategy compared to PARAMESH default

• This plot shows the computation time *per block* per time step for a 32 core run.

• Blue shows the benefits of a partitioning strategy that ensures a good load-balance at *each* mesh level.

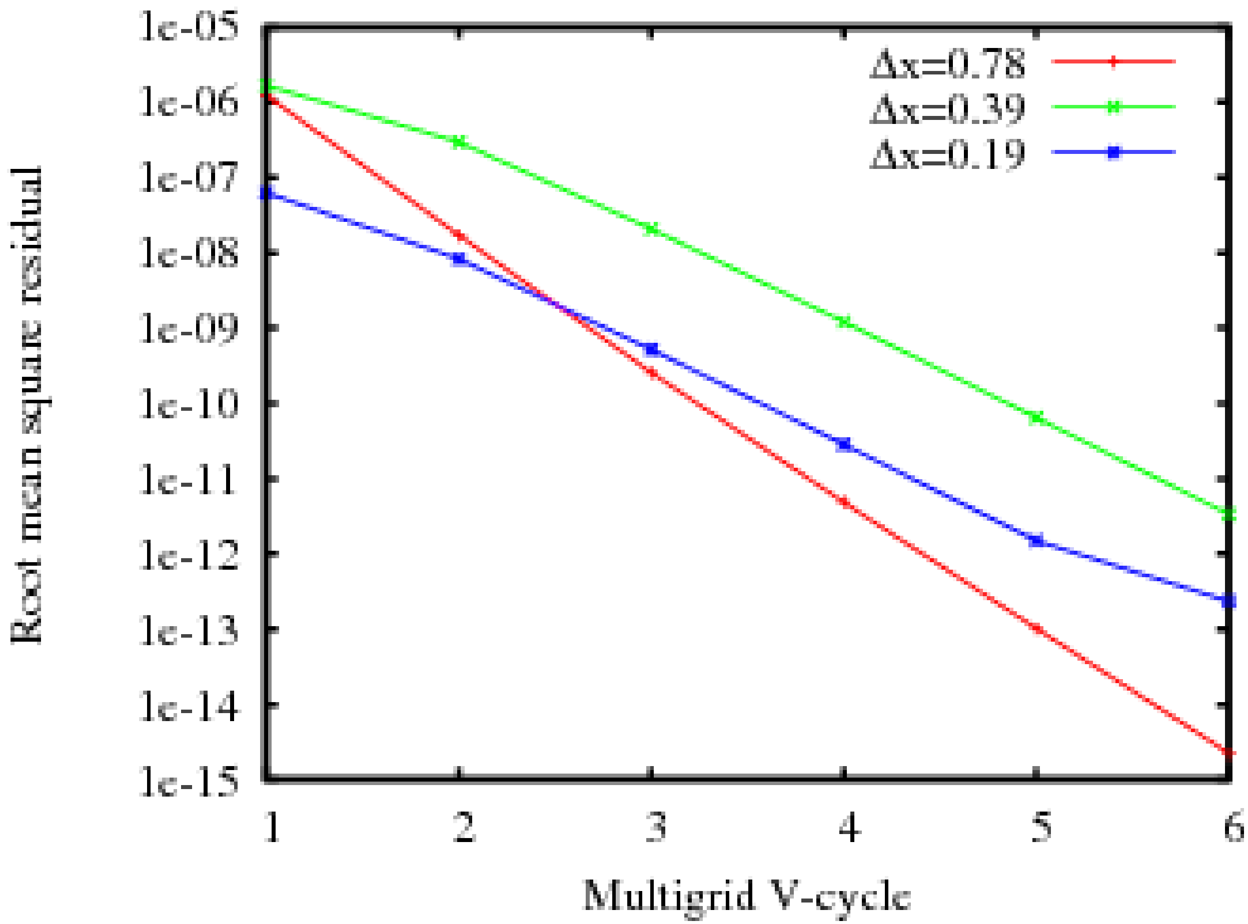


UNIVERSITY OF LEEDS

Typical Simulation Results

- Optimal multigrid convergence is obtained for adaptive meshes in 3-d

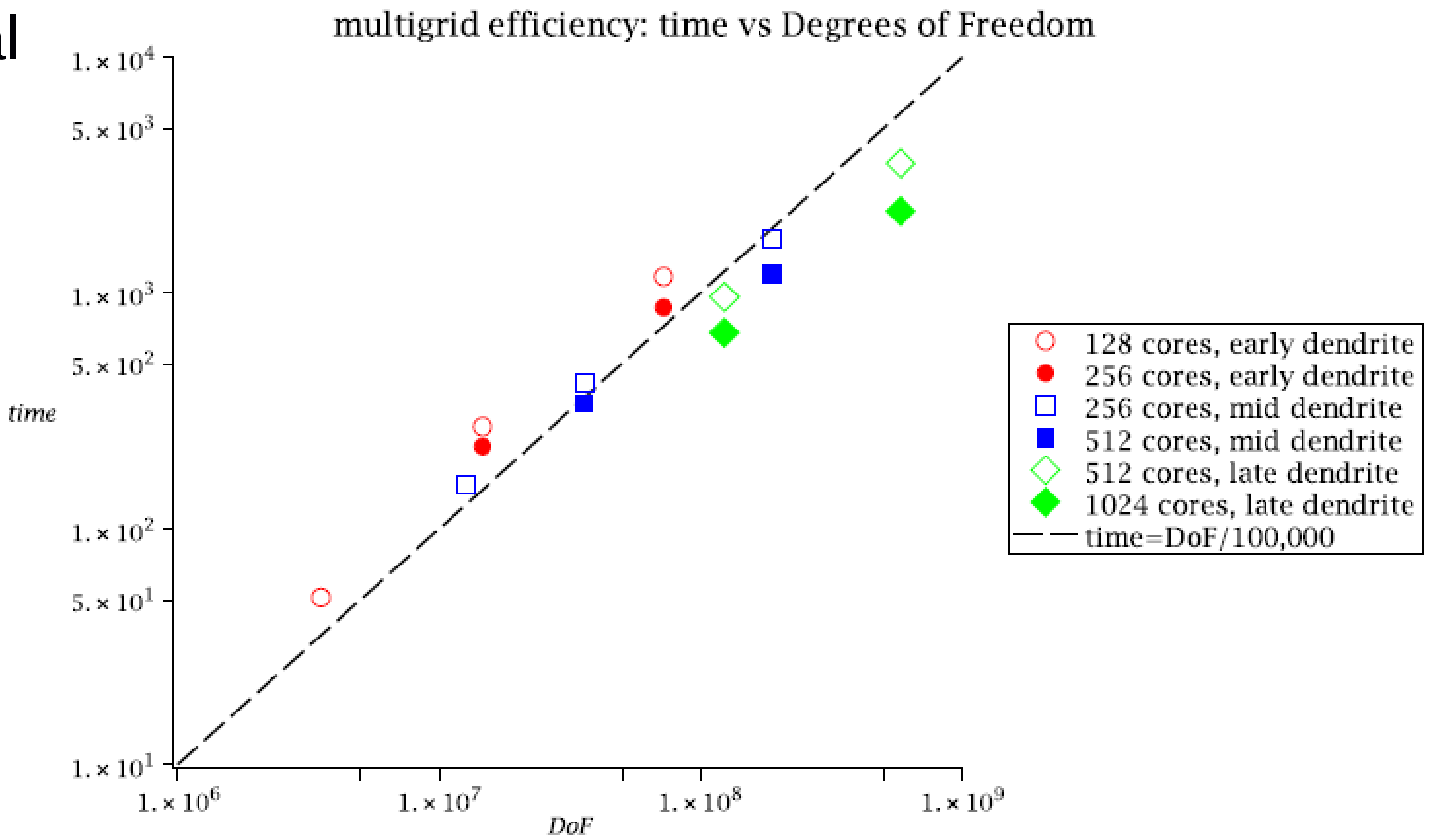
- Here we see close to an optimal convergence rate for the nonlinear multigrid solver applied to the isothermal problem in 3-d .



Typical Simulation Results

- Optimal multigrid convergence is obtained for adaptive meshes in 3-d

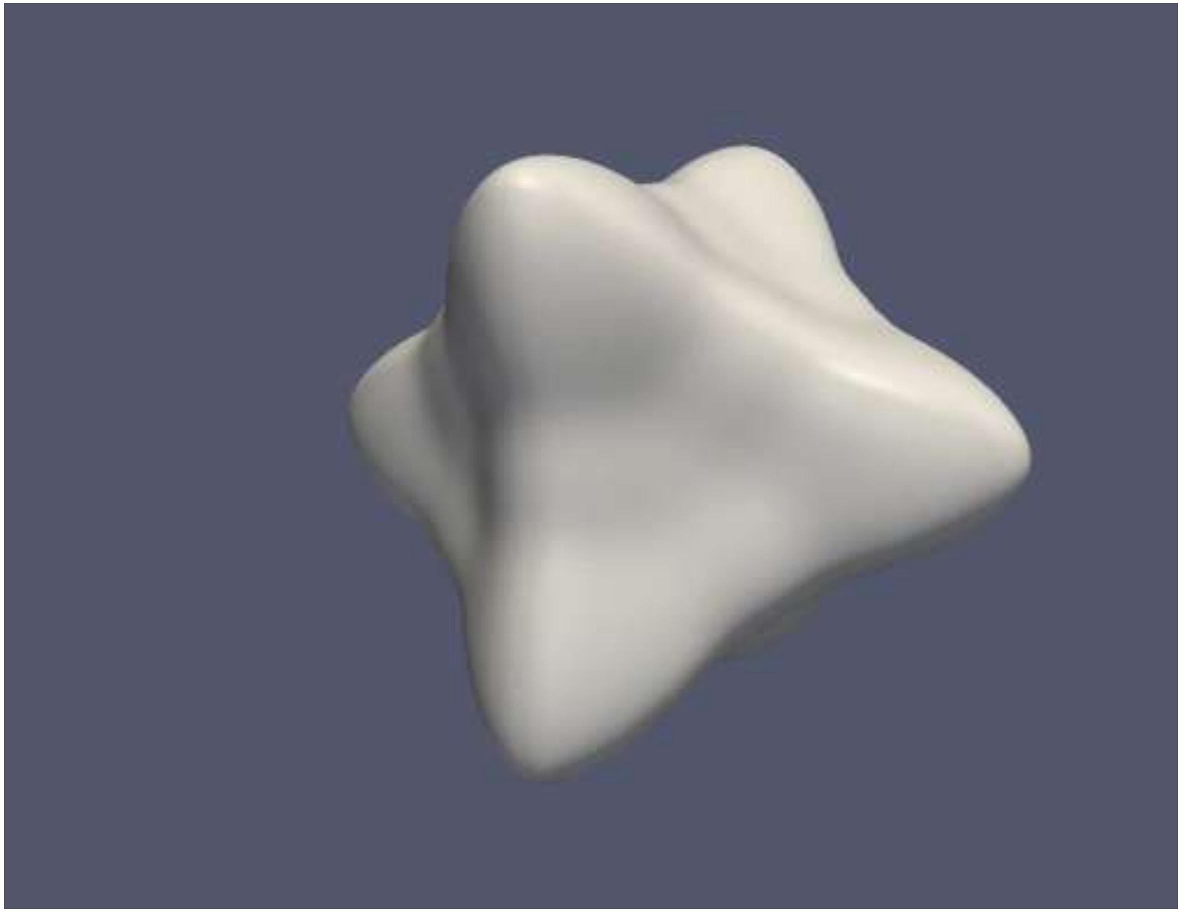
- And optimal timings for the full non-isothermal problem in 3-d, using different numbers of cores as a dendrite grows...



Typical Simulation Results

- Snapshots of dendrite formation – Le=40 , Δ=0.525

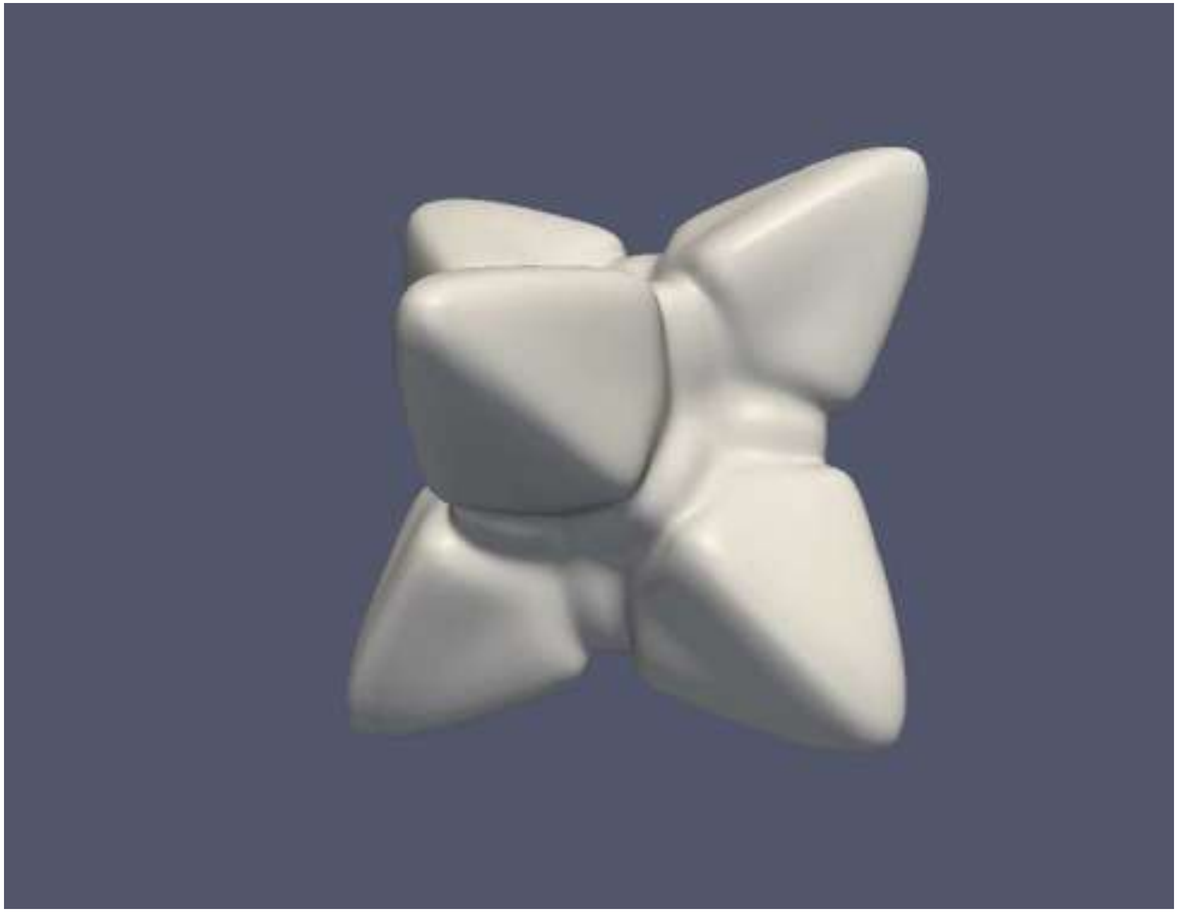
These images show snapshots of this particular run: early...



Typical Simulation Results

- Snapshots of dendrite formation – Le=40 , Δ=0.525

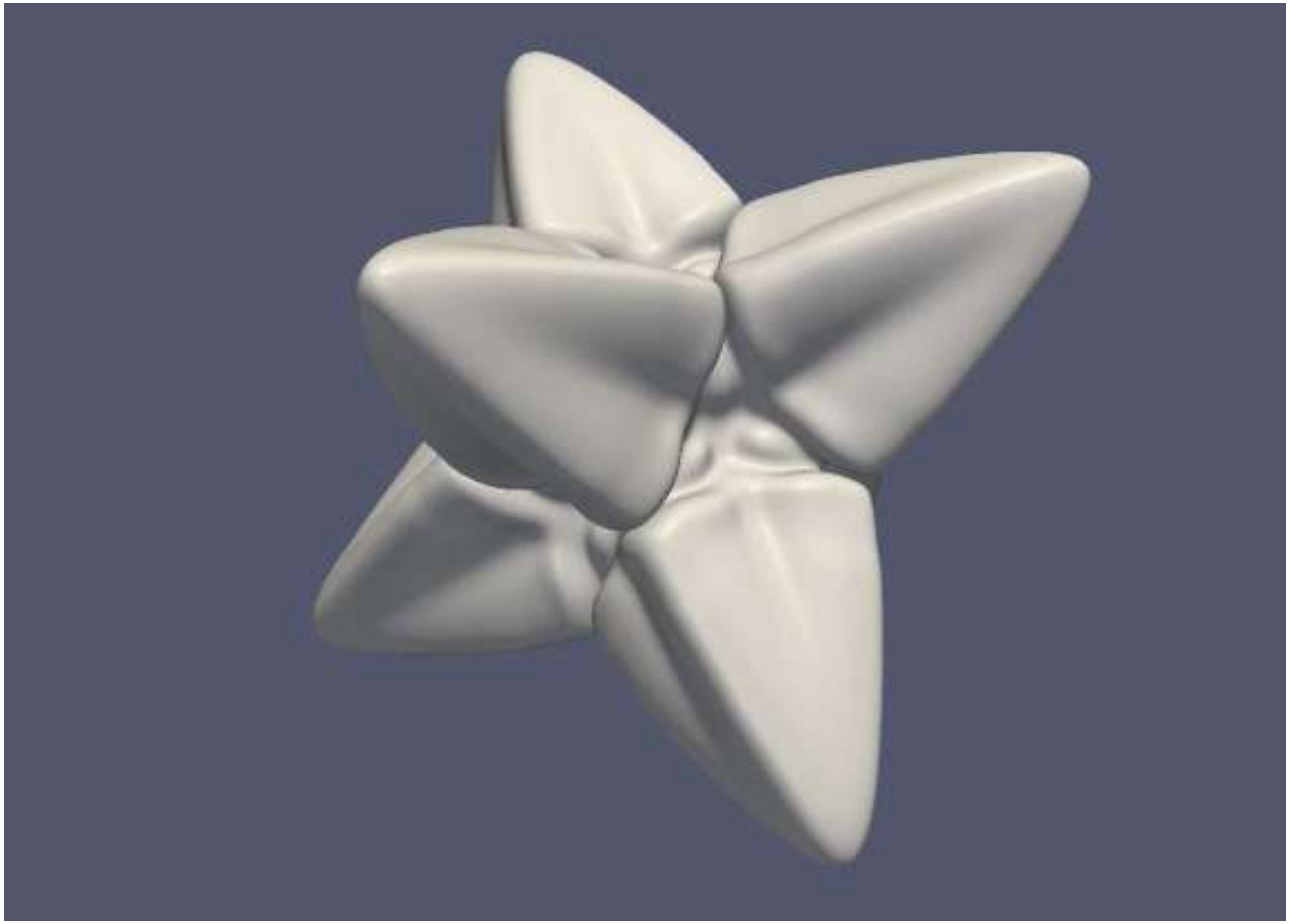
These images show snapshots of this particular run: mid...



Typical Simulation Results

- Snapshots of dendrite formation – Le=40 , Δ=0.525

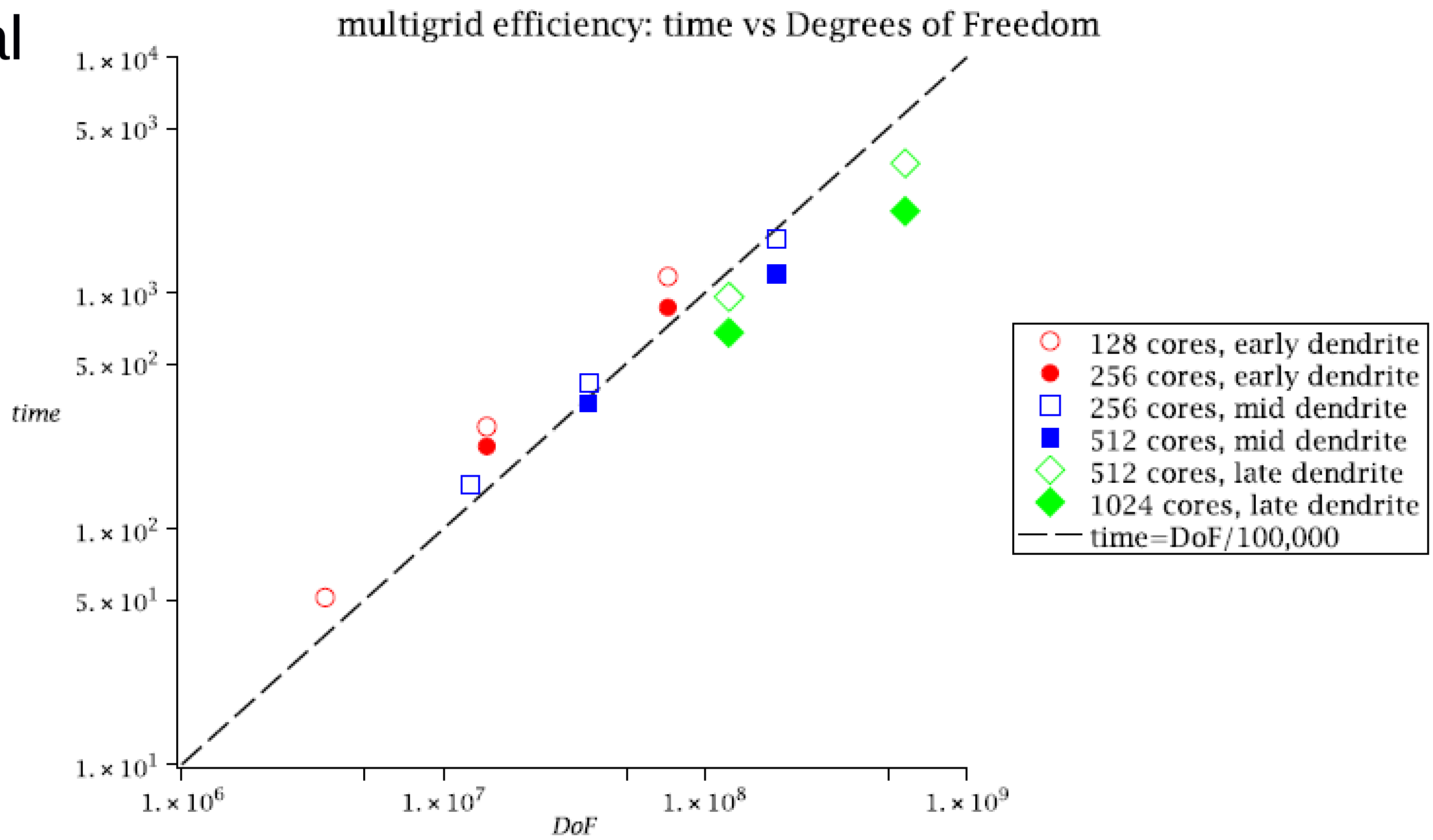
These images show snapshots of this particular run: late...



Typical Simulation Results

- Optimal multigrid convergence is obtained for adaptive meshes in 3-d

- And optimal timings for the full non-isothermal problem in 3-d, using different numbers of cores as a dendrite grows...



Parallel Performance

- Parallel performance: Le = 40, Δ=0.325

All runs based upon 32-core nodes (1 GB/core)

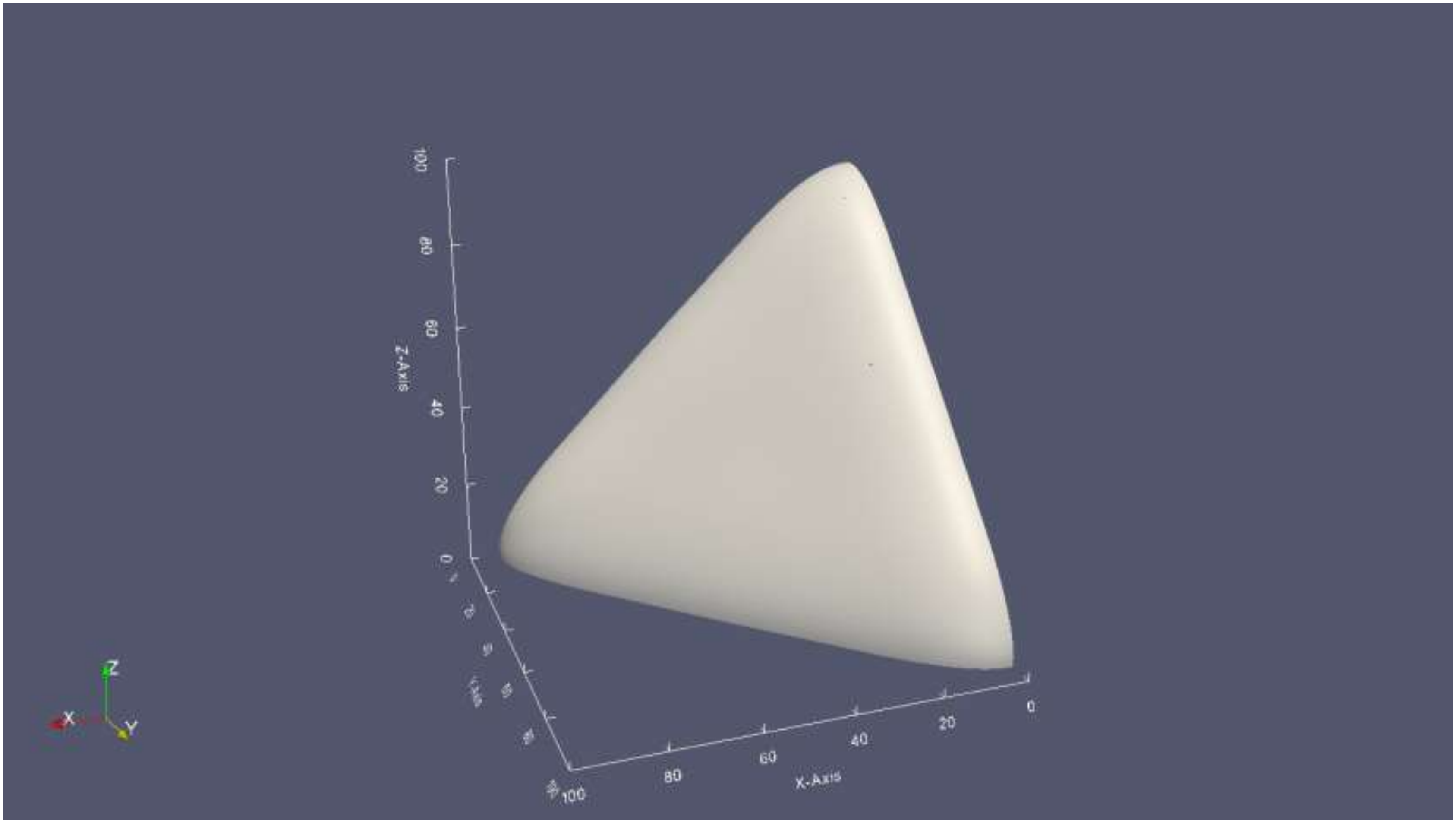
- Up to 100M DoFs per time step;
- Timings for 10 time steps – after steps 4000/7000...

Core #	Run times including (without) remeshing					
	dx=0.78	dx=0.78	dx=0.39	dx=0.39	dx=0.195	dx=0.195
32	120 (120)	456 (450)	735 (715)	-	-	-
64	73 (73)	264 (259)	420 (404)	896 (884)	-	-
128	57 (57)	181 (176)	298 (267)	546 (522)	1292 (1234)	-
256		171 (164)	246 (228)	412 (401)	955 (868)	1864 (1792)
512				335 (331)	707 (675)	1331 (1228)
1024					479 (388)	900 (701)

Typical Simulation Results

- Snapshots of dendrite formation – Le=40, Δ=0.325

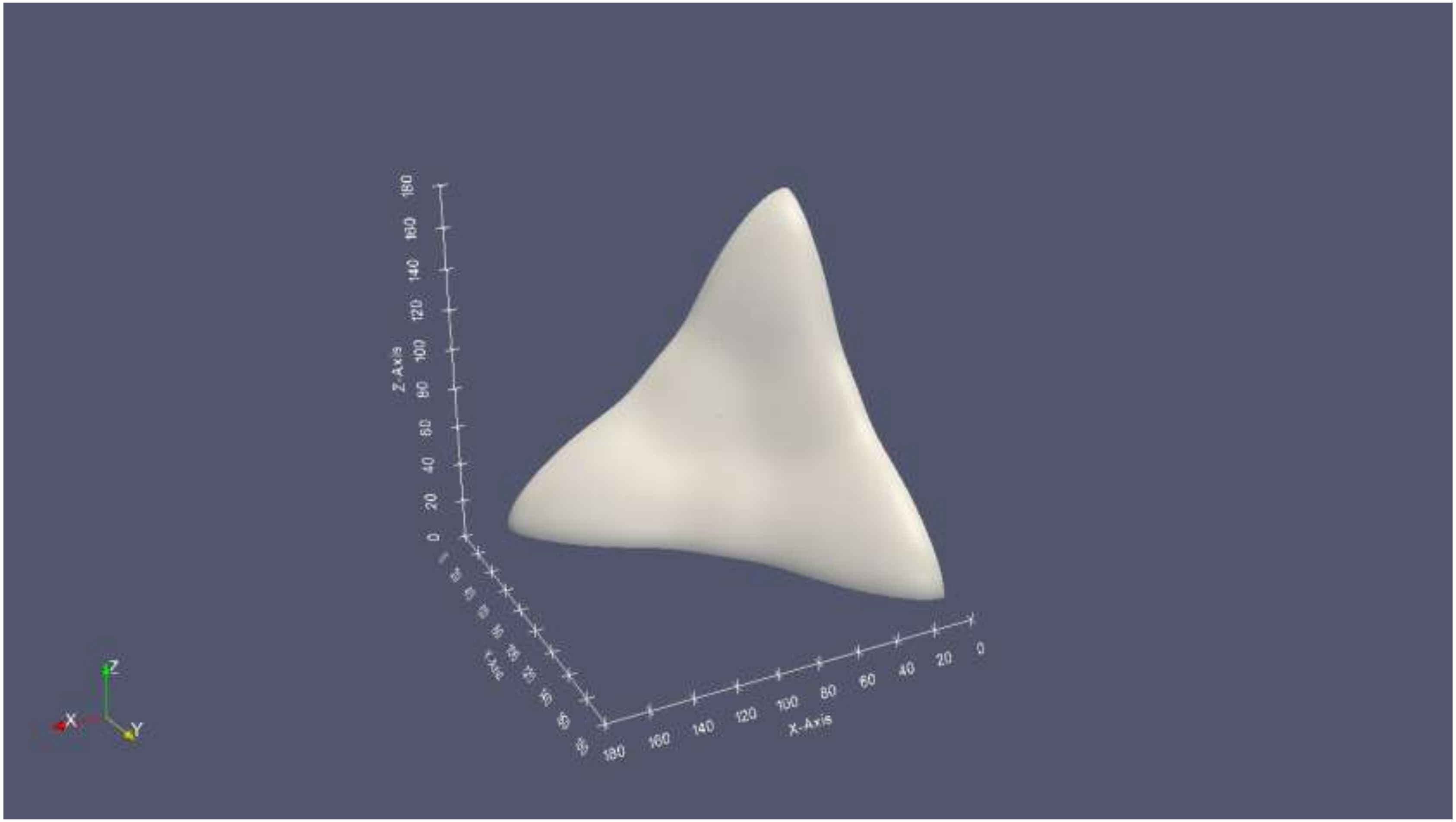
These images show snapshots of this particular run: step 4000...



Typical Simulation Results

- Snapshots of dendrite formation – Le=40 , Δ=0.325

These images show snapshots of this particular run: step 7000...



UNIVERSITY OF LEEDS

Parallel Performance

- Parallel performance: Le = 40, Δ=0.325

All runs based upon 32-core nodes (1 GB/core)

- Up to 100M DoFs per time step;
- Timings for 10 time steps – after steps 4000/7000...

Core #	**Run times including (without) remeshing**					
	dx=0.78	dx=0.78	dx=0.39	dx=0.39	dx=0.195	dx=0.195
32	120 (120)	456 (450)	735 (715)	-	-	-
64	73 (73)	264 (259)	420 (404)	896 (884)	-	-
128	57 (57)	181 (176)	298 (267)	546 (522)	1292 (1234)	-
256		171 (164)	246 (228)	412 (401)	955 (868)	1864 (1792)
512				335 (331)	707 (675)	1331 (1228)
1024					479 (388)	900 (701)

Parallel Performance

• Loss of efficiency

There are a number of causes of the loss of efficiency:

- Overhead with undertaking mesh adaptation and the resulting dynamic load-balancing.
- Choice of 8x8x8 block size amplifies this.
- Grid transfer operations have a high communication to computation ratio.
- Coarsest grid solver also has high communication to computation ratio (and idle cores).

Nevertheless, the problem would be intractable without the Combination of adaptivity, multigrid and parallel solution!

Discussion

Numerical methods implemented in parallel:

1. Second order (19-point) finite differences for the spatial discretization of the highly nonlinear coupled system of parabolic PDEs.
2. Hierarchical *adaptivity* to refine and coarsen the spatial mesh as the solution evolves in time.
3. Fully implicit *second order* BDF time integration for the stiff ODE systems that arise after spatial discretization (essential for stiff problems).
4. Fully coupled nonlinear Multigrid solver for the nonlinear algebraic systems that occur at each time step: *optimal complexity*.
5. *Adaptive time step* selection based upon local error estimation and/or convergence rate of MG solver.
6. Implemented within a general-purpose software framework

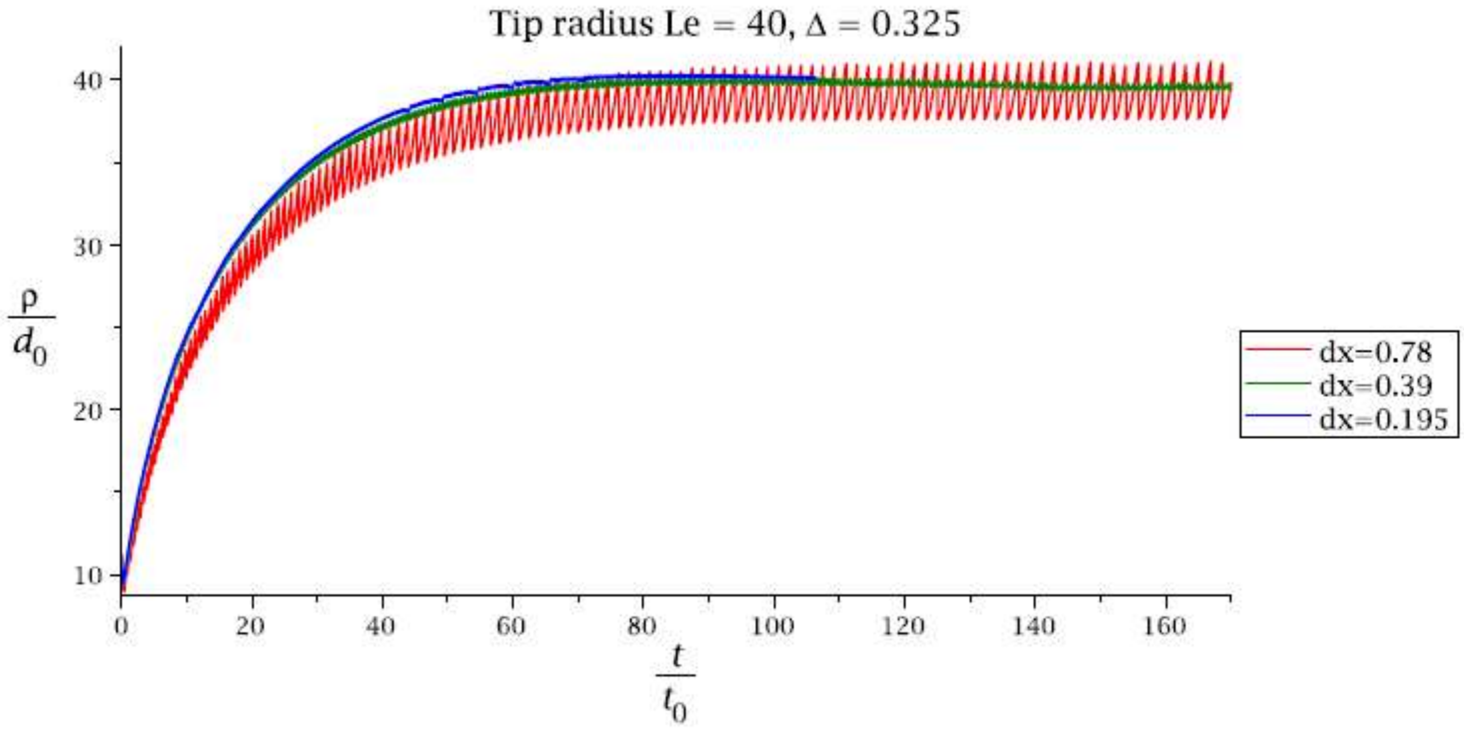
Discussion

Results obtained:

1. Parallelism and adaptivity provide a capability in terms of memory use (high resolution in 3-d on up 1TB of RAM).
2. Parallelism and multigrid provide a capability for implicit solution of very large stiff systems (run times in hours)
3. Strong scaling is difficult due to achieve due to high overheads of adaptivity and nonlinear multigrid.
4. Weak scaling is obtainable up to a certain level – provided the work per core is sufficient.
5. Able to obtain new results for the quantitative description of the solidification of metal alloys, using physically realistic parameter values!
6. Similar outcome for completely different P-F model of 3D Tumor Growth (e.g. Wise, Lowengrub & Christini (2011)).

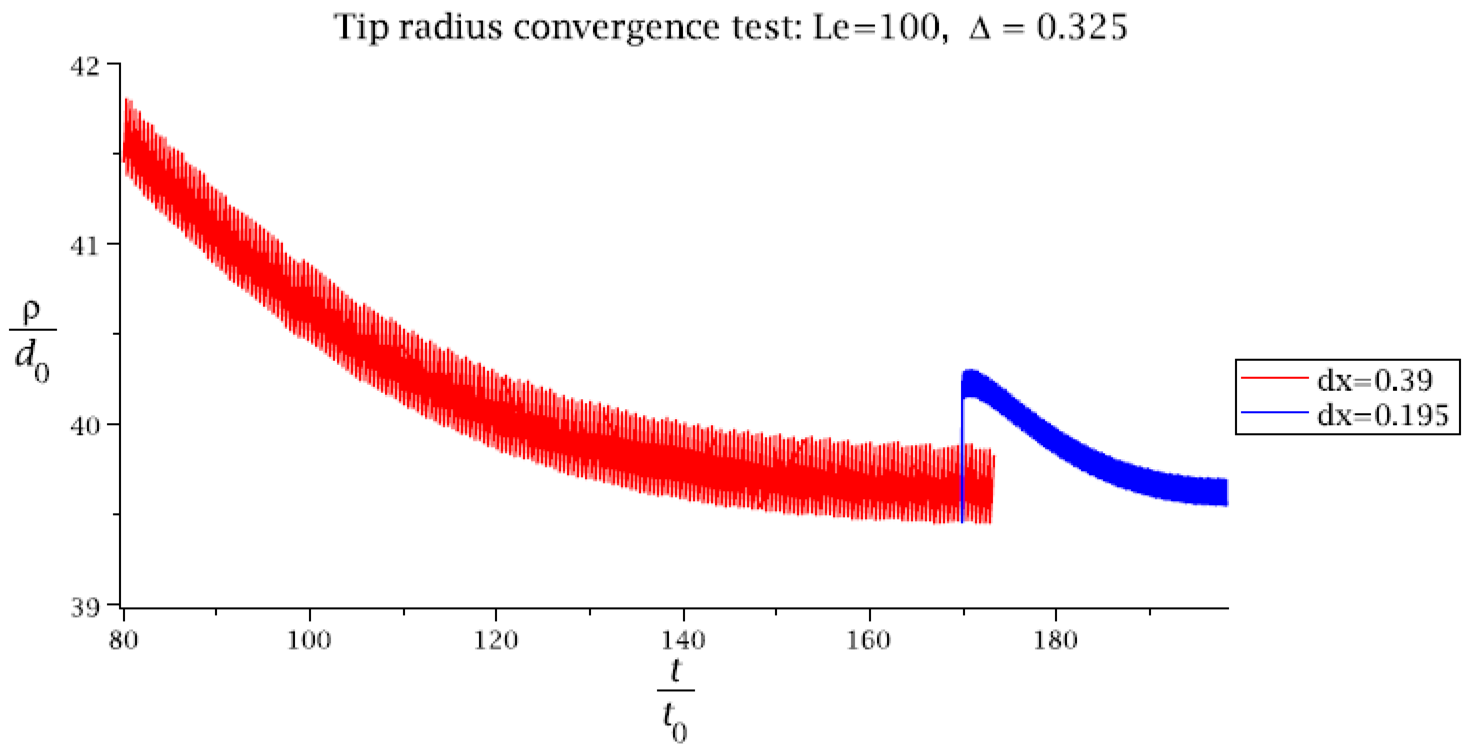
Examples of New Quantitative Results

Mesh independent prediction of dendrite tip radius…



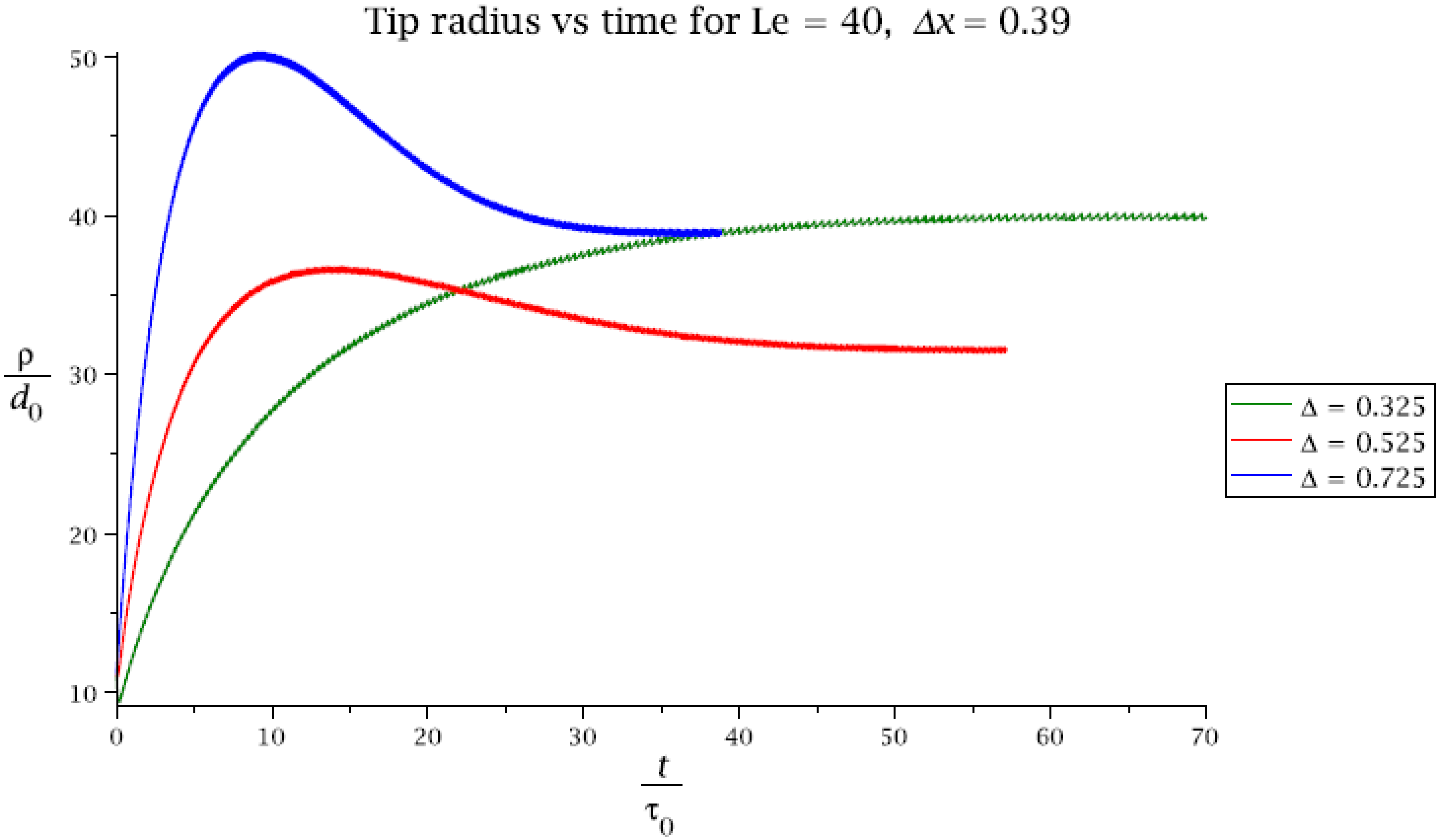
Examples of New Quantitative Results

Mesh independent prediction of dendrite tip radius



Examples of New Quantitative Results

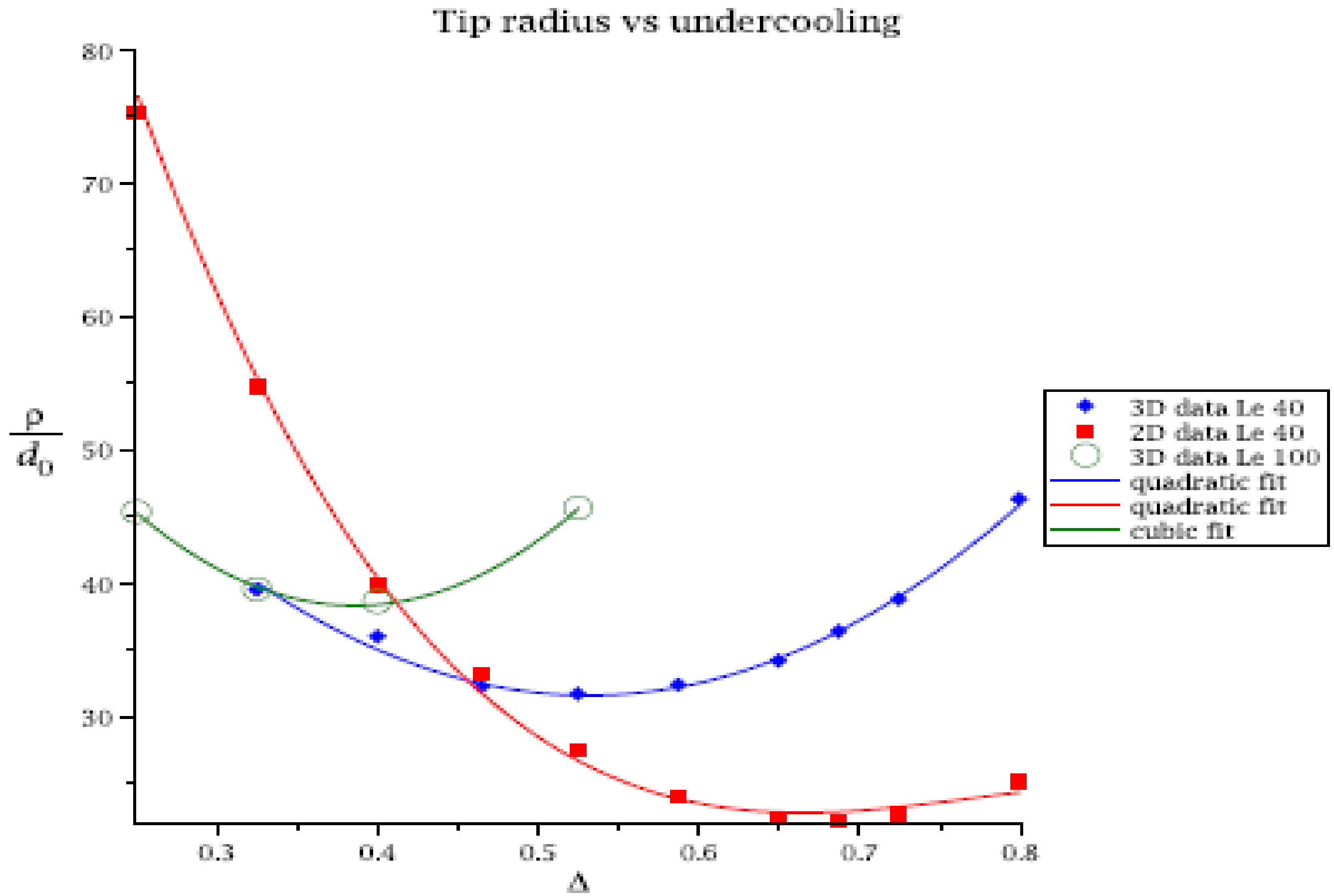
Can therefore predict quantitative dendrite features…



UNIVERSITY OF LEEDS

Examples of New Quantitative Results

And see the importance of three-dimensional simulations…



Can we Scale to >10000 Cores?

Two algorithmic priorities:

1. Improve dynamic load-balancing
   - Enhance the data locality between parent and child blocks
   - Enhance the data remapping algorithms to improve their efficiency
2. Replace nonlinear MG solver (FAS+MLAT) with a matrix-free Newton-Krylov solver with linear MG preconditioner based upon FAC (fast adaptive composite grid) approach:
   - Preliminary analysis shows lower constant of optimality in the MG cost and slightly faster convergence
   - Only requires a (non-exact) linear solve at the coarsest level which will allow a finer coarsest grid and better parallel performance

Thank you…

To the organisers for the invitation to attend this workshop…

Any questions?