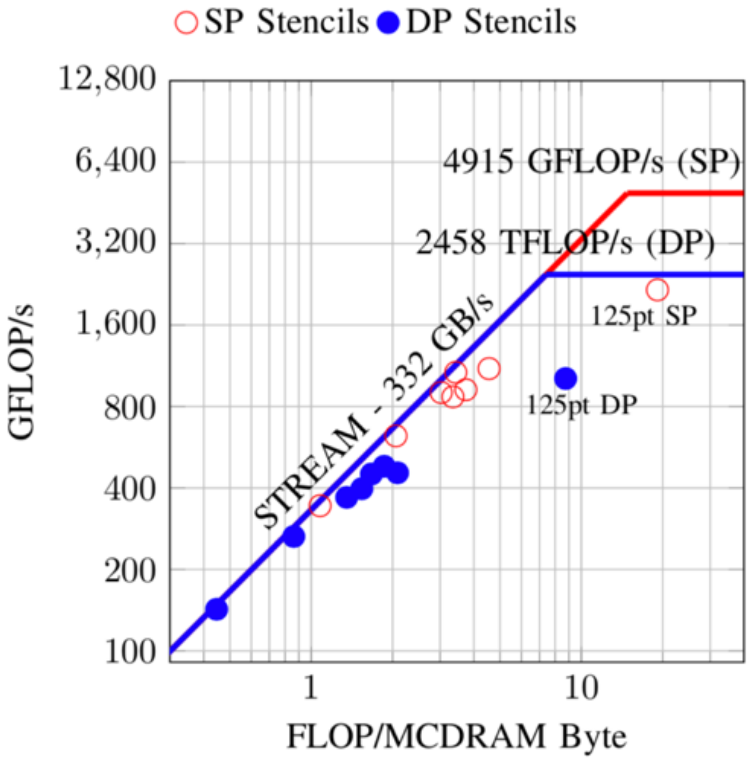
Energy-conserving numerics and consistency

1. What is the role for the 2nd law in the formulation of energy-consistent subgrid physics and physics-dynamics coupling?
2. Do we separate physics from dynamics for good reasons? Is it an obstacle to some "better approaches"? Conversely are there good arguments in favor of not separating?
3. Should all physics be written as PDEs? Would that exclude certain approaches to parameterizing certain processes (e.g. deep convection)?
4. What needs to be specified in order to clarify which energetics we are talking about?
   → Total energy, thermodynamic potentials, dissipation rates ...
5. Suppose we find a way to do everything right, and it is not affordable.
   How do we minimize the errors induced by inevitable compromises? Monitor errors?
6. What to expect / demand in terms of accuracy / convergence?
7. What approaches could we learn from other fields?

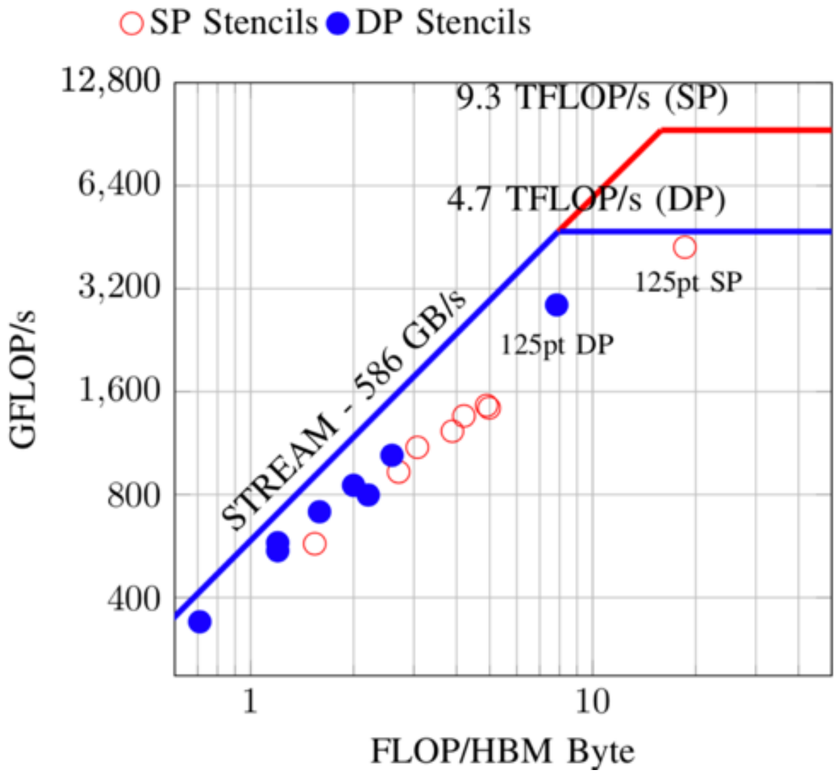
5. Correctness vs. Cost

Correctness	**Current state-of-the-art:**
• Have we split the problem correctly?	→ Non-hydrostatic with good asymptotics?
• Correctly solving the mathematical problem?	→ Conservation, energy, entropy?
Accuracy:	
• Is higher-order worth it?	→ Higher-order HEVI schemes?
• Is implicit worth it?	→ Numerical convergence in space/time?
Feasibility:	
• Efficient calculation (time, resources)?	→ AMR + block-structured efficiency?
• “Grey zone” confidence?	→ More rigorous tests?

Block-structured stencils, “brick” code-generation



Intel KNL 7250



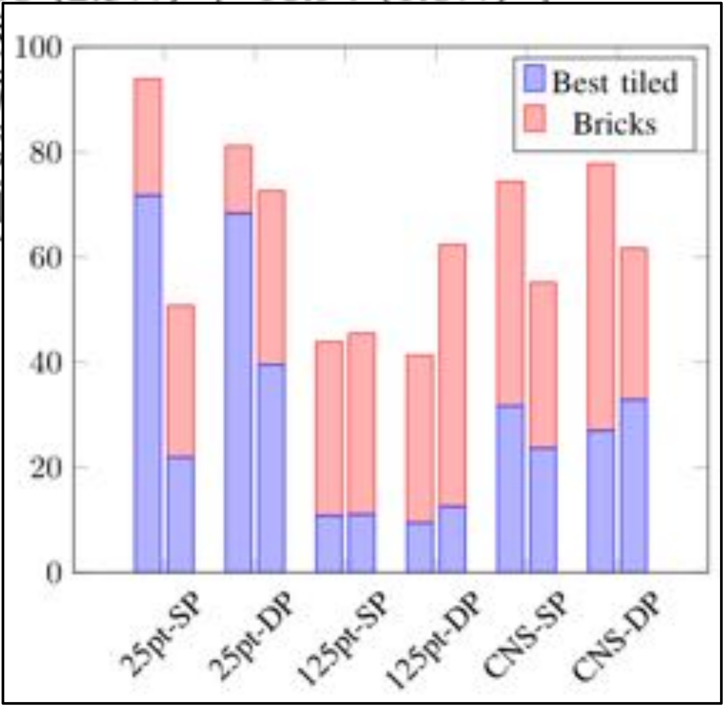
NVIDIA P100

Bricks can be faster for higher-order, coupled stencils

Name	Brick GStencil/s & Speedup			
	KNL SP	KNL DP	P100 SP	P100 DP
7pt	26.46 (0.9×)	10.96 (0.7×)	41.04 (1.4×)	24.25 (1.1×)
13pt	24.93 (1.0×)	10.59 (0.8×)	35.79 (1.8×)	21.06 (1.4×)
19pt	24.35 (1.2×)	9.98 (0.9×)	32.18 (2.1×)	18.84 (1.7×)
25pt	21.83 (1.3×)	9.20 (1.2×)	29.08 (2.3×)	16.94 (1.8×)
27pt	20.84 (1.0×)	8.59 (0.9×)	26.5	
125pt	8.67 (4.1×)	4.08 (4.4×)	16.7	
iso	14.24 (1.1×)	6.52 (1.0×)	19.9	
CNS	1.98 (2.3×)	1.03 (2.9×)	3.1	

Speed-up depends on a lot of factors:

- Aggressive 6D tiling w/ auto-tuning baseline
- Very low HBM (& L1) data movement compared to tiling larger array sizes
- Greatly reduces # of streams, cache misses



Higher is better

Nonlinear column solves use batched

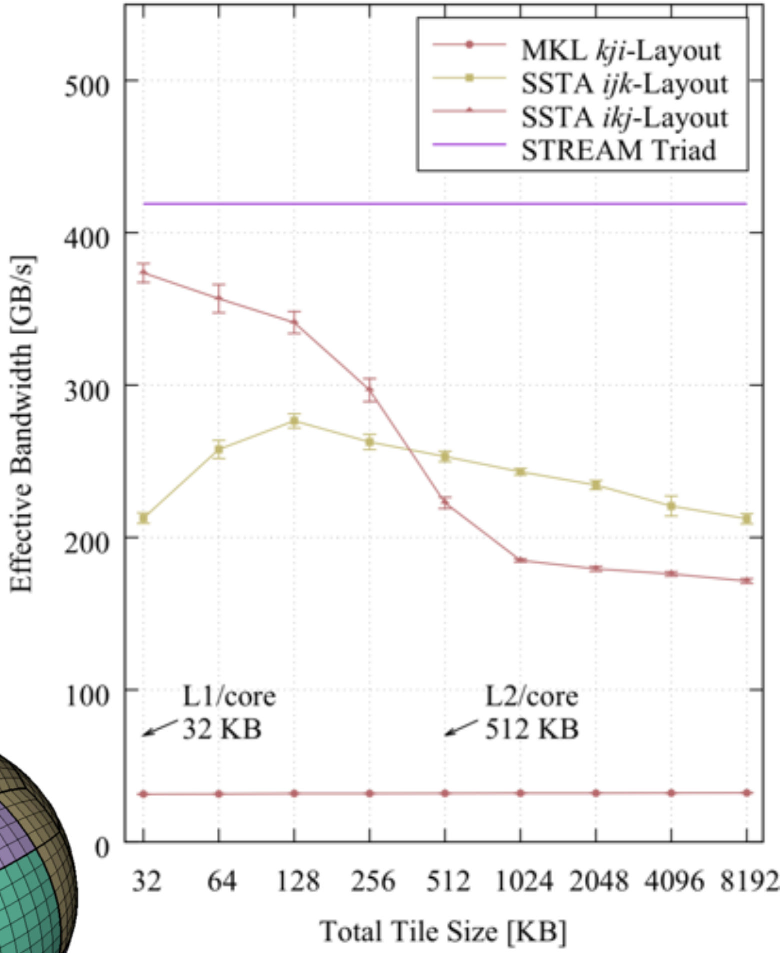
Column solvers are an issue

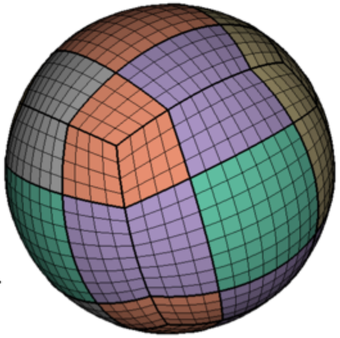
- Non-linear for numerical accuracy, asymptotics
- Typically 1-3 iterations
- k-major or i-major?

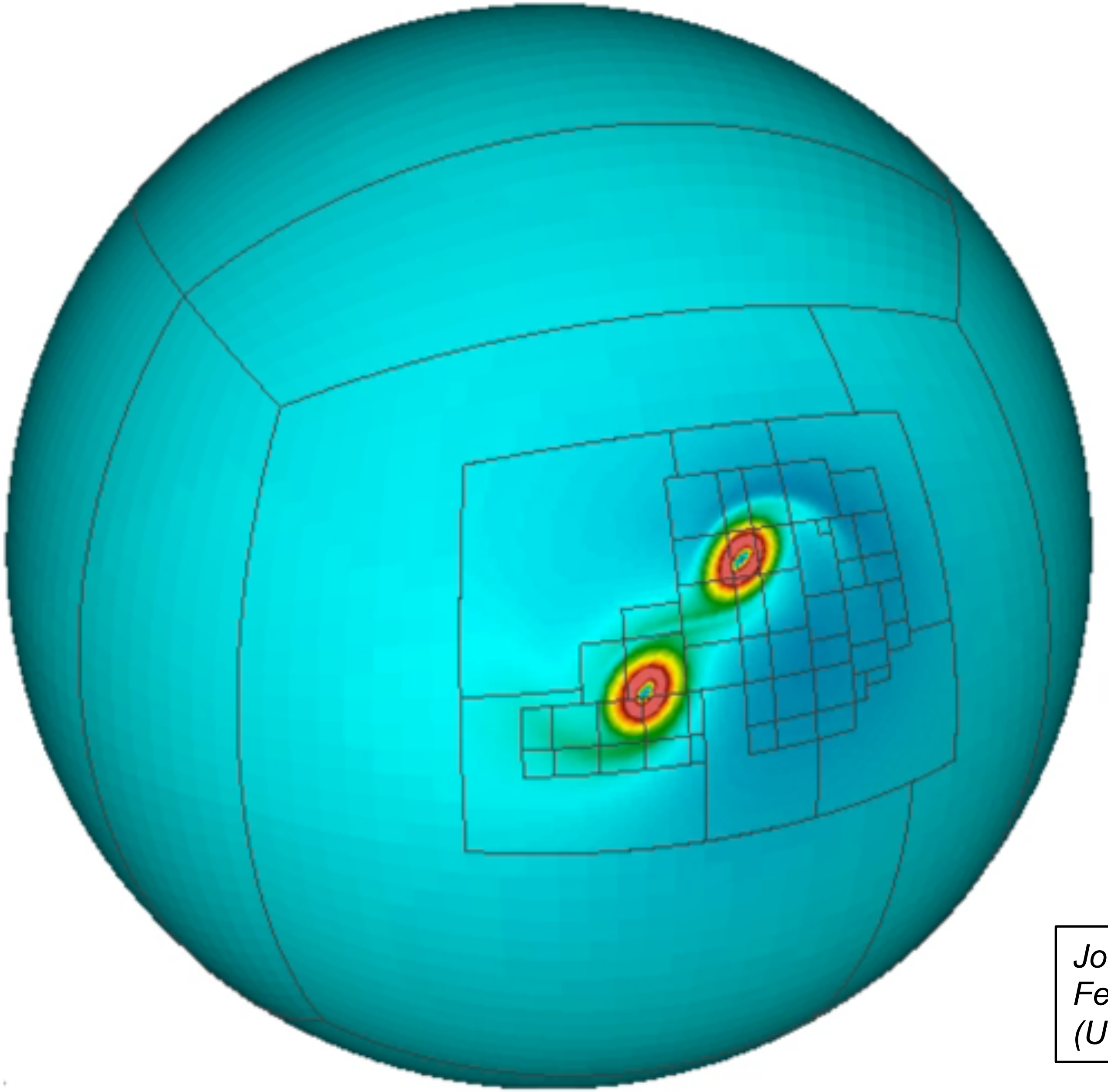
Banded solvers are bw-bound

- Despite lots of flops, also requires same order of data
- Fast divides/reciprocal on KNL mean memory bw-bound
- Even batched MKL library calls have overhead, generalities

Porting to GPU and MPI optimizations in progress …

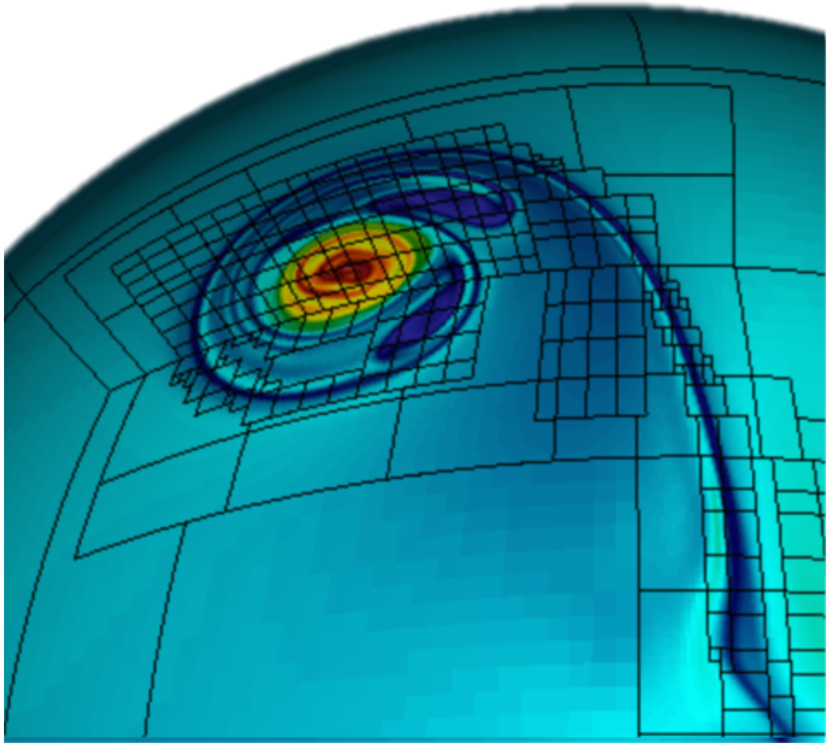




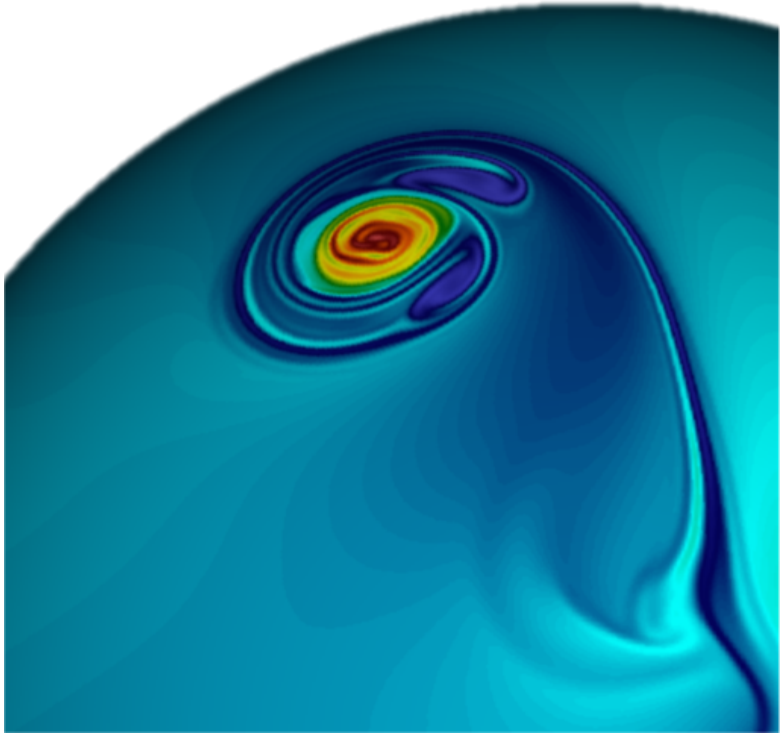


Joint work with
Ferguson, Jablonowski
(U. Michigan)

Example: Vorticity dynamics, AMR vs. uniform



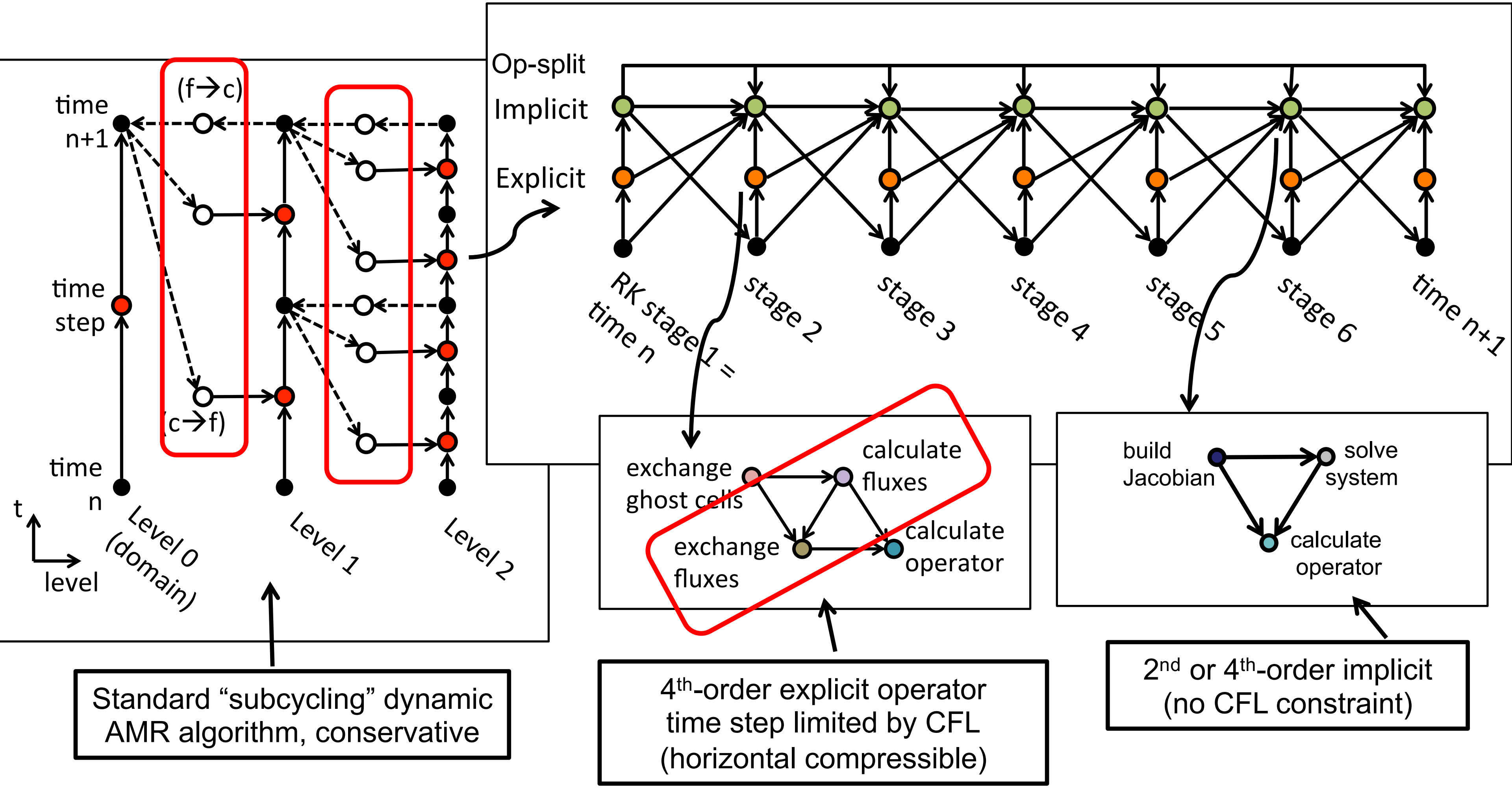
3x AMR C1024



Uniform C1024

Joint work with Ferguson, Jablonowski (U. Michigan)

AMR time stepping: RK ImEx, but also SDC, R-W, QSS, …

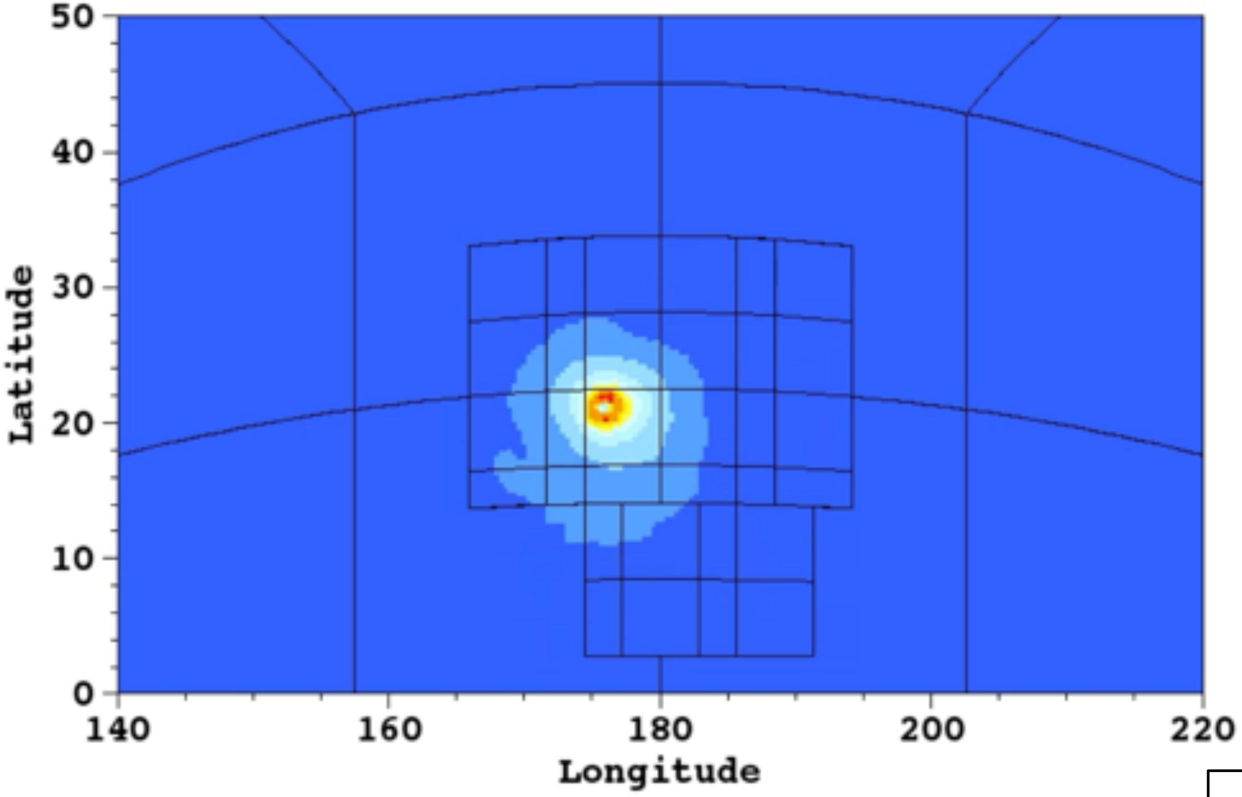


Standard “subcycling” dynamic AMR algorithm, conservative

4th-order explicit operator time step limited by CFL (horizontal compressible)

2nd or 4th-order implicit (no CFL constraint)

3D AMR: Idealized Tropical Cyclone



c64/c256 (~160/40 km)

Joint work with Ferguson, Jablonowski (U. Michigan)

AMR Time accuracy across refinement boundaries?

- AMR in “grey zone:” hydrostatic (> 20km, large aspect ratios) vs. non-hydrostatic (< 10 km)?
- Aspect ratios depend on horizontal resolution, orography, kinked/stretched vertical mappings
- Implicit solver should asymptote to Richardson’s eqn / hydrostatic vertical velocity

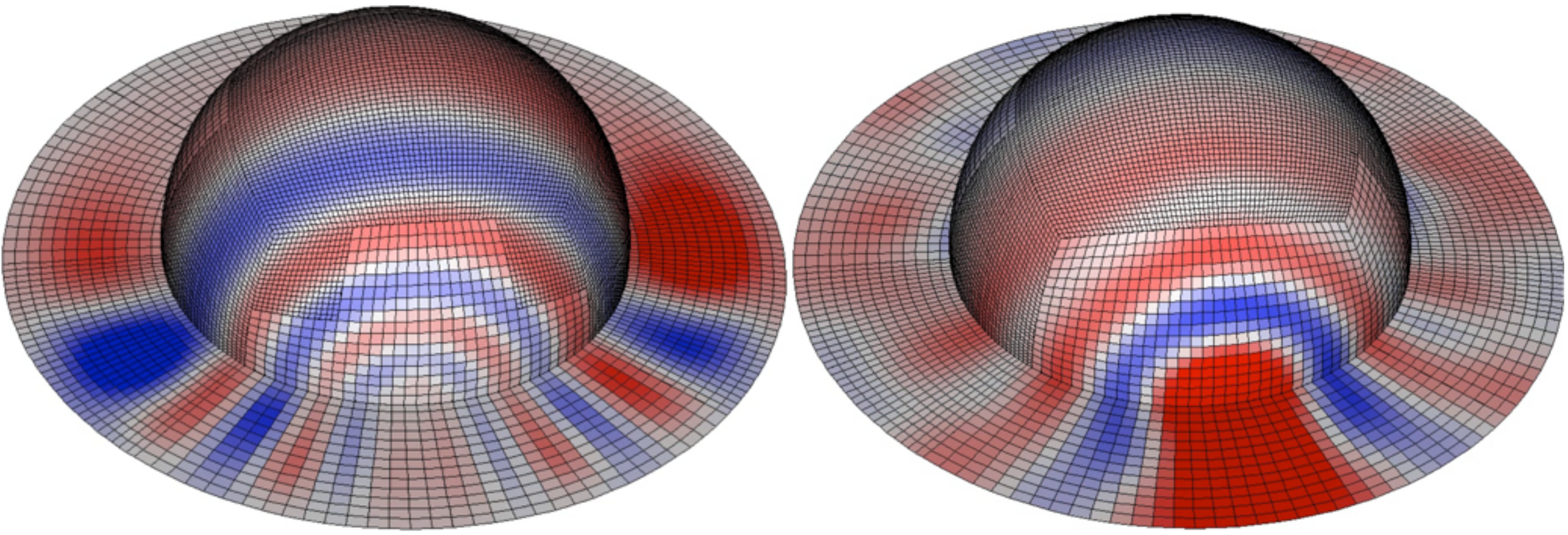
$$\frac{\partial p^H}{\partial t} = w\rho g - g\int_r^\infty \nabla^\perp \cdot (\rho \boldsymbol{u}^\perp)dr'$$

$$\boxed{\frac{\partial w}{\partial t} + \frac{1}{\rho}\frac{\partial \pi}{\partial r}} = -\boldsymbol{u}^\perp \cdot \nabla^\perp w - w\frac{\partial w}{\partial r},$$

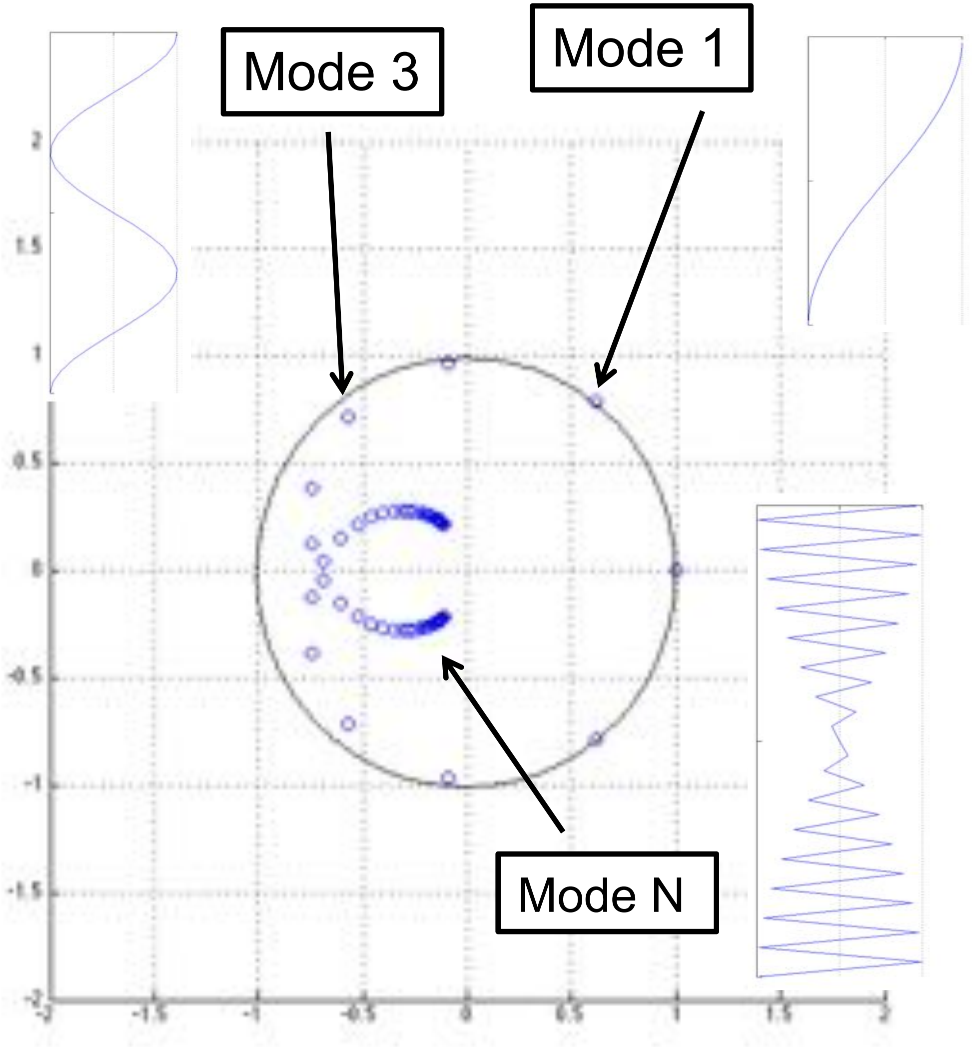
$$\boxed{\frac{\partial \pi}{\partial t} + \rho c^2\frac{\partial w}{\partial r}} = g\int_r^\infty \nabla^\perp \cdot (\rho \boldsymbol{u}^\perp)dr' - \rho c^2 \nabla^\perp \cdot \boldsymbol{u}^\perp - \boldsymbol{u}^\perp \cdot \nabla^\perp p^H \boxed{- \boldsymbol{u} \cdot \nabla \pi}$$

$$\Rightarrow \quad \rho c^2\frac{\partial w}{\partial r} = g\int_r^\infty \nabla^\perp \cdot (\rho \boldsymbol{u}^\perp)dr' - \rho c^2 \nabla^\perp \cdot \boldsymbol{u}^\perp - \boldsymbol{u}^\perp \cdot \nabla^\perp p^H$$

Nonhydrostatic Gravity vs. Acoustic Waves



Vertical AMR Challenges


Mode 3

Mode 1

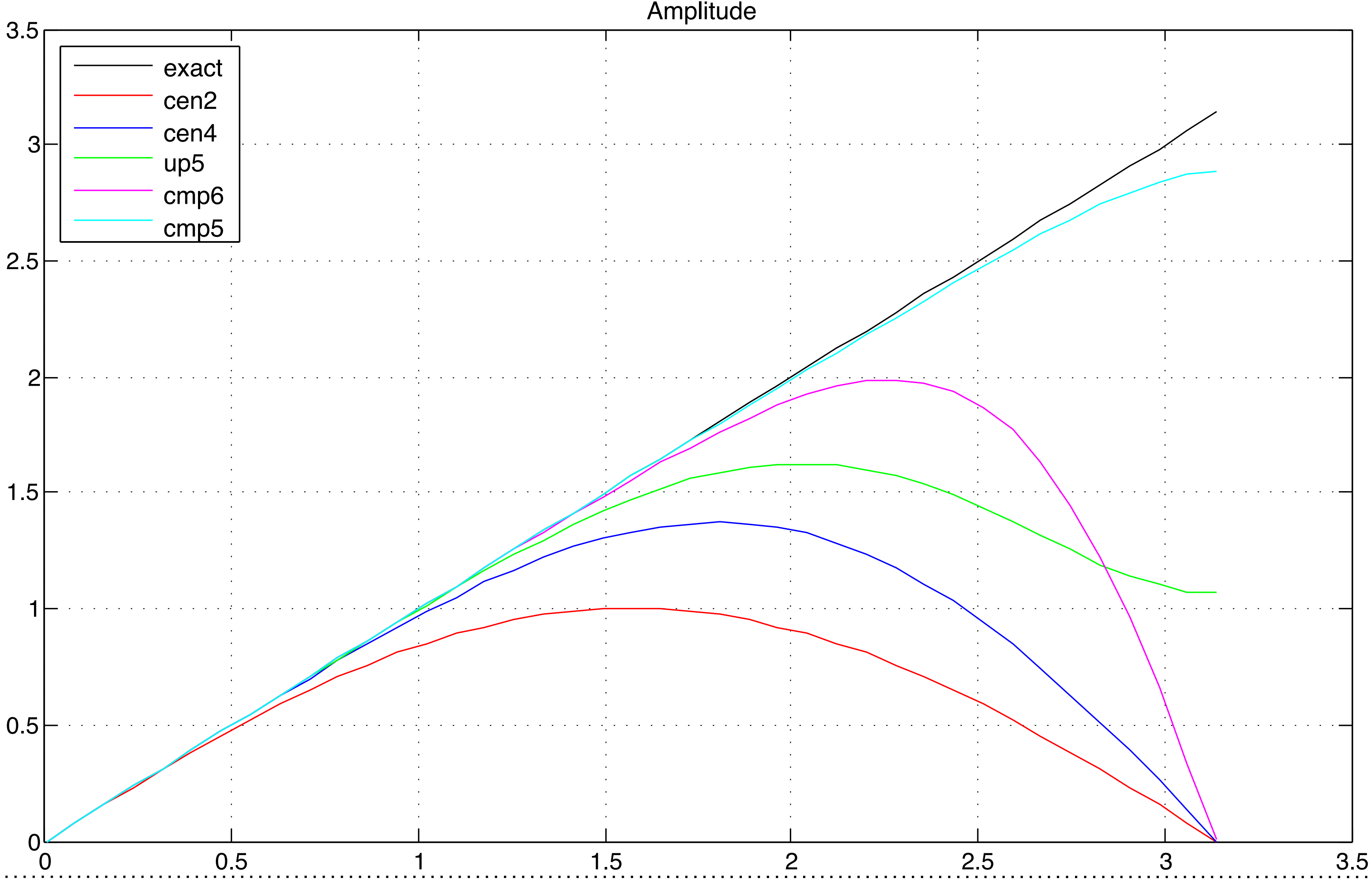
Mode N

- Coupling implicit wave equation, “volume discrepancy,” and physics
  → *as in complex combustion codes*
- At coarse resolutions, system only good for 1-2 modes at 10x CFL
- Fast modes damp rapidly, but oscillate

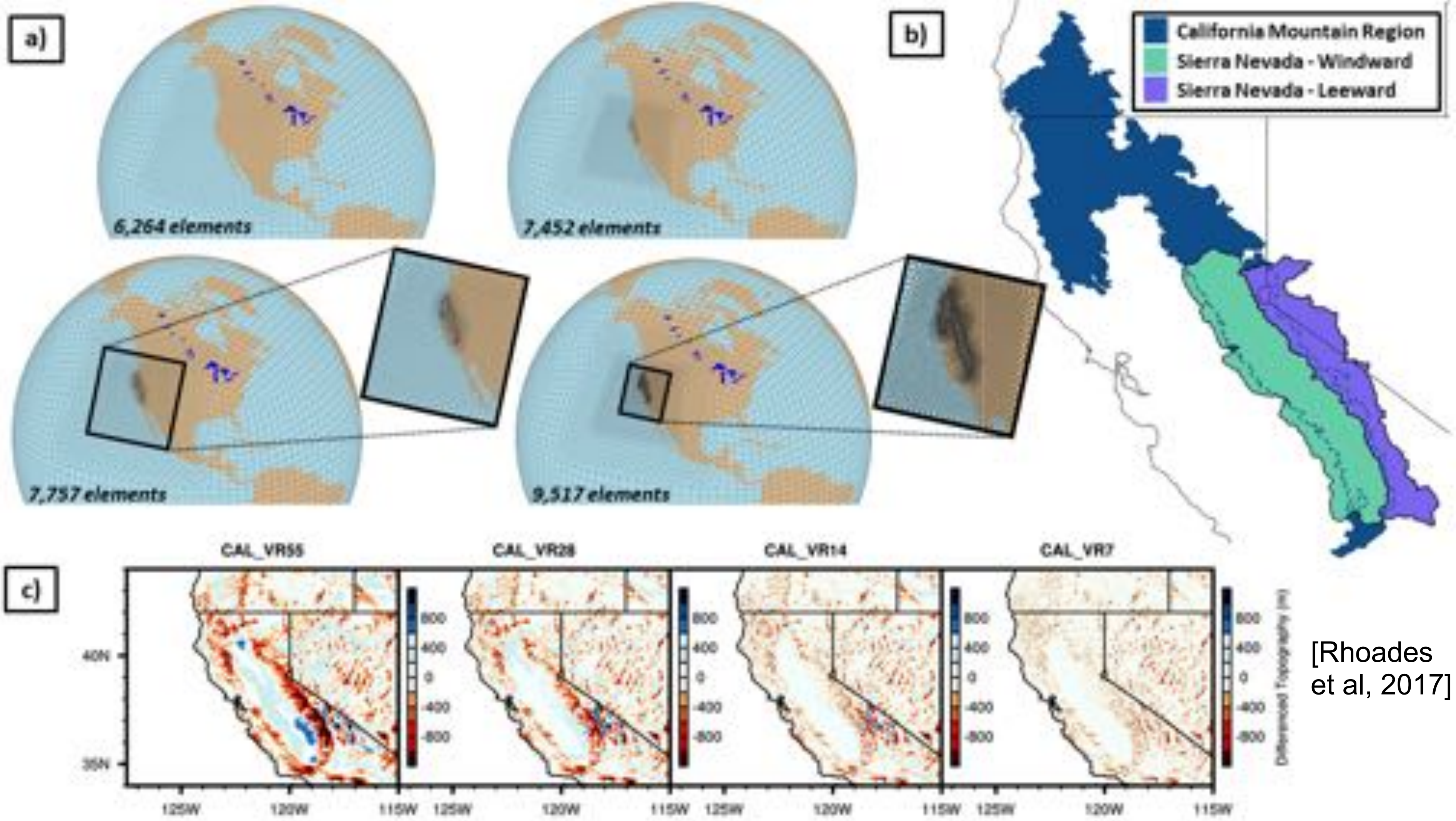
→ Very under-resolved, phase error flips signs every solve

- Kinked vertical mappings exacerbate with eigenmode shapes
- AMR complications – space-time convergence vs. vertical refinement?

Space-time accuracy, spectral resolution matters



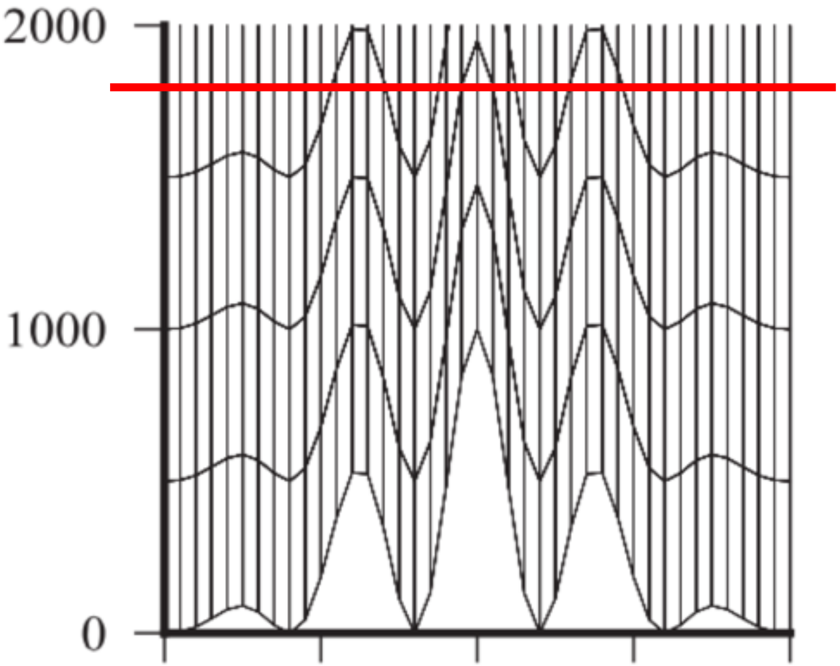
Global simulations at coarse resolutions?

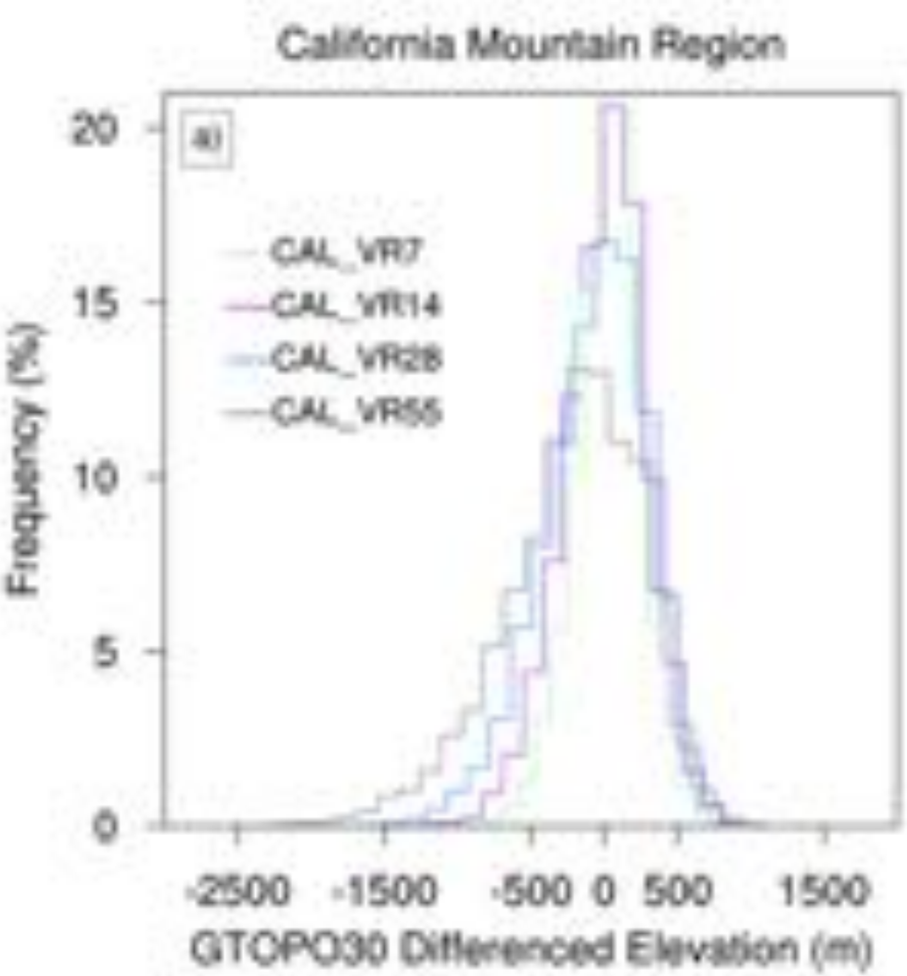


[Rhoades et al, 2017]

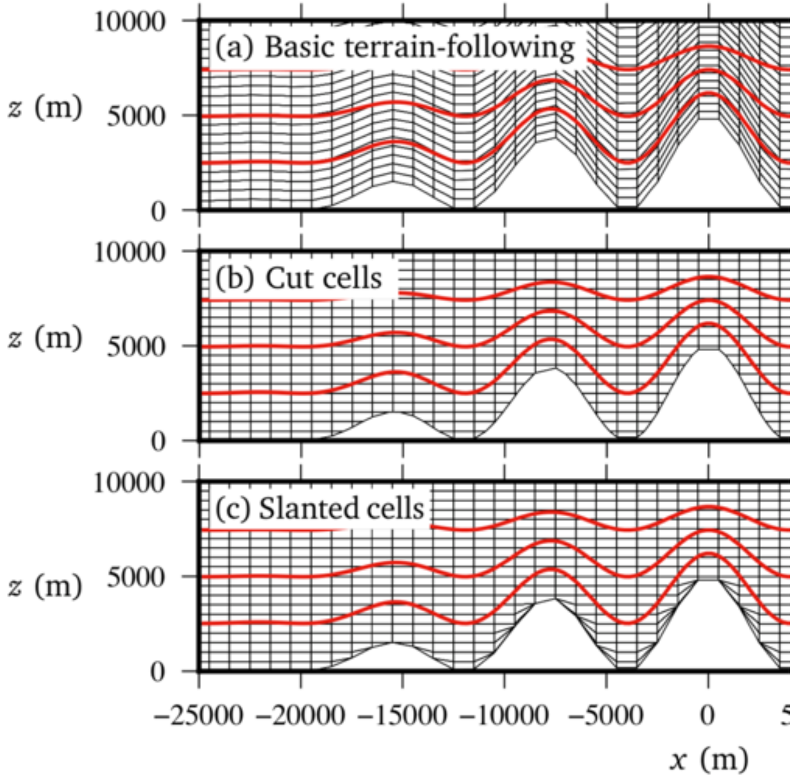
Complex orography ←→ non-smooth mappings?

Distorted mesh into PBL?

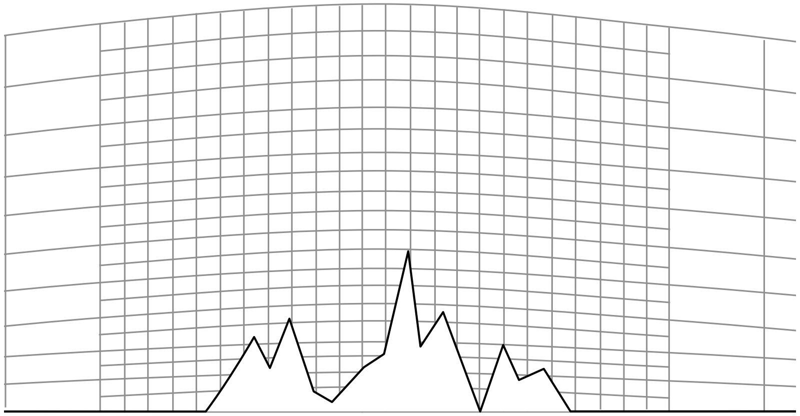




Complex orography with cut cells, smooth mappings?

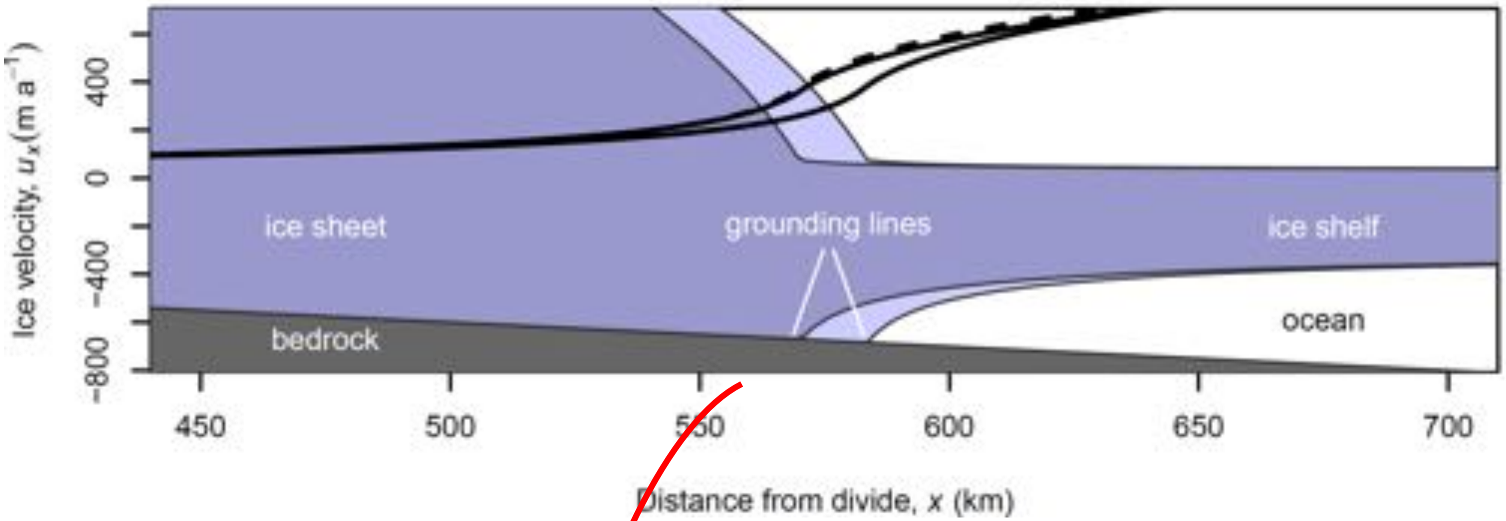


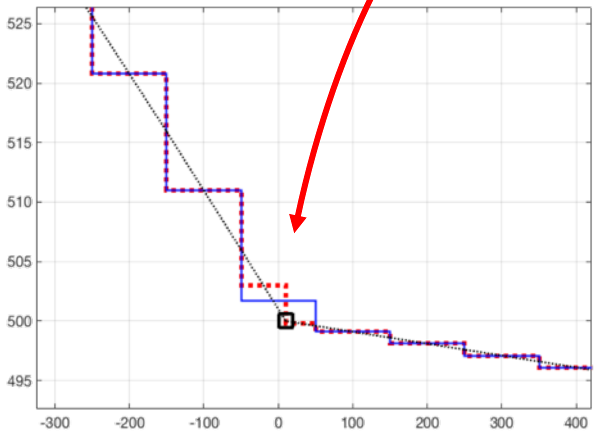
[Shaw, Weller 2016]

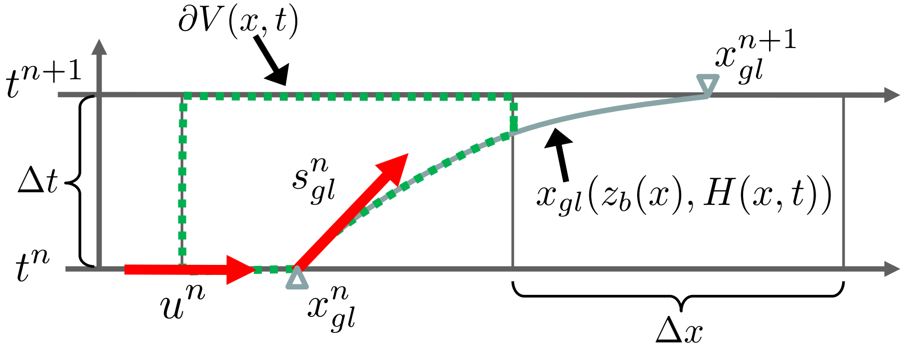


AMR + cut cells have the potential to correctly represent topography at all resolutions, *without* smoothing terrain or distorting mapping.

Example: Grounding line as a Multifluid Interface







Energy-conserving numerics and consistency

1. What is the role for the 2nd law in the formulation of energy-consistent subgrid physics and physics-dynamics coupling?
2. Do we separate physics from dynamics for good reasons? Is it an obstacle to some "better approaches"? Conversely are there good arguments in favor of not separating?
3. Should all physics be written as PDEs? Would that exclude certain approaches to parameterizing certain processes (e.g. deep convection)?
4. What needs to be specified in order to clarify which energetics we are talking about?
   → Total energy, thermodynamic potentials, dissipation rates ...
5. Suppose we find a way to do everything right, and it is not affordable.
   How do we minimize the errors induced by inevitable compromises? Monitor errors?
6. What to expect / demand in terms of accuracy / convergence?
7. What approaches could we learn from other fields?