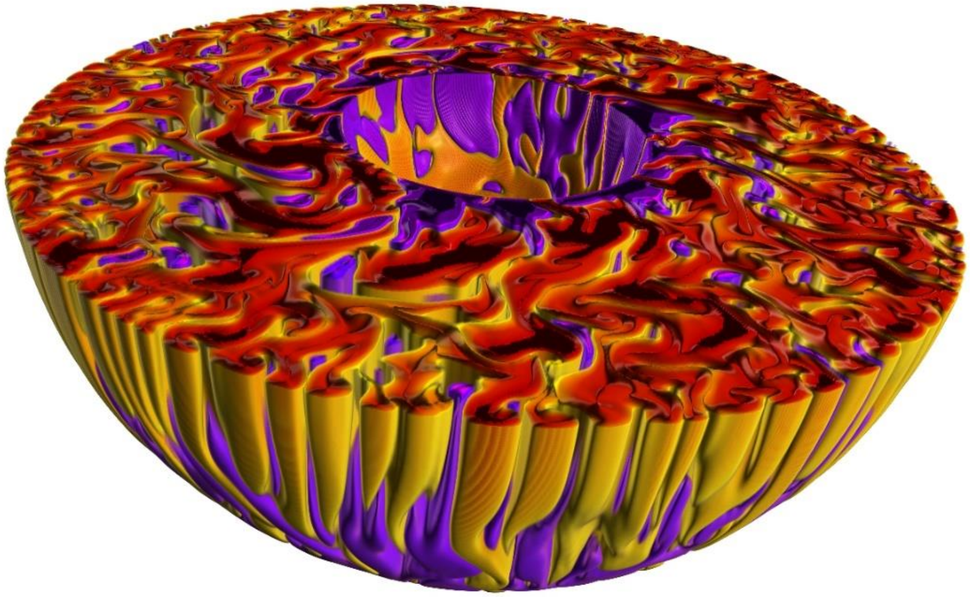
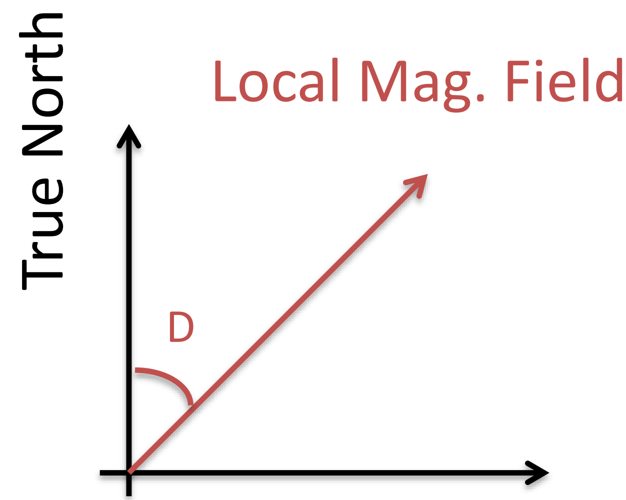
Rayleigh: An Open-Source, Pseudo-Spectral MHD Code for $O(10^5)$ cores … and beyond



CIG

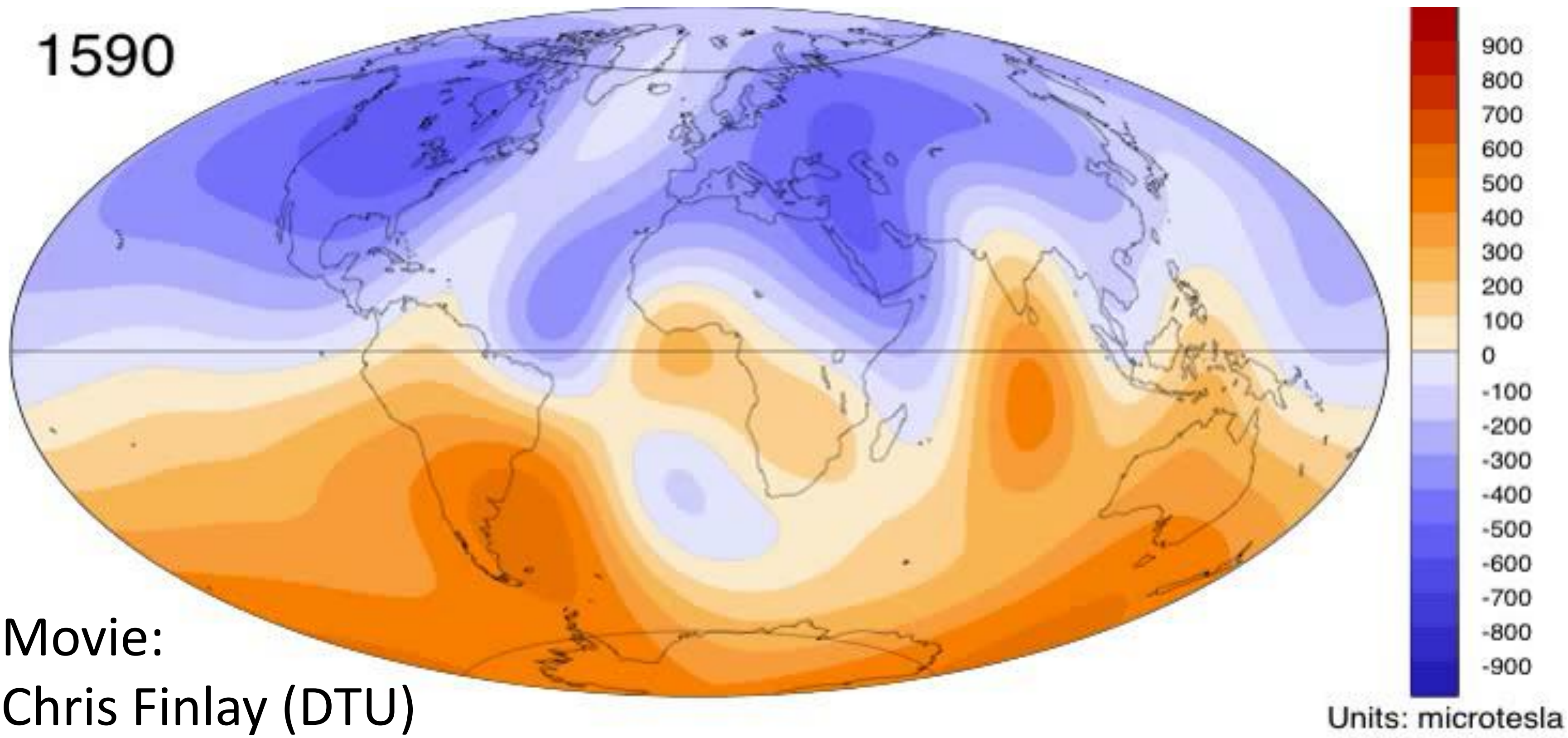
Nick Featherstone
Applied Math, CU Boulder / CIG

Geomagnetic Declination in 1701



[Edmond Haley, 1701]

History of Earth's Magnetic Field

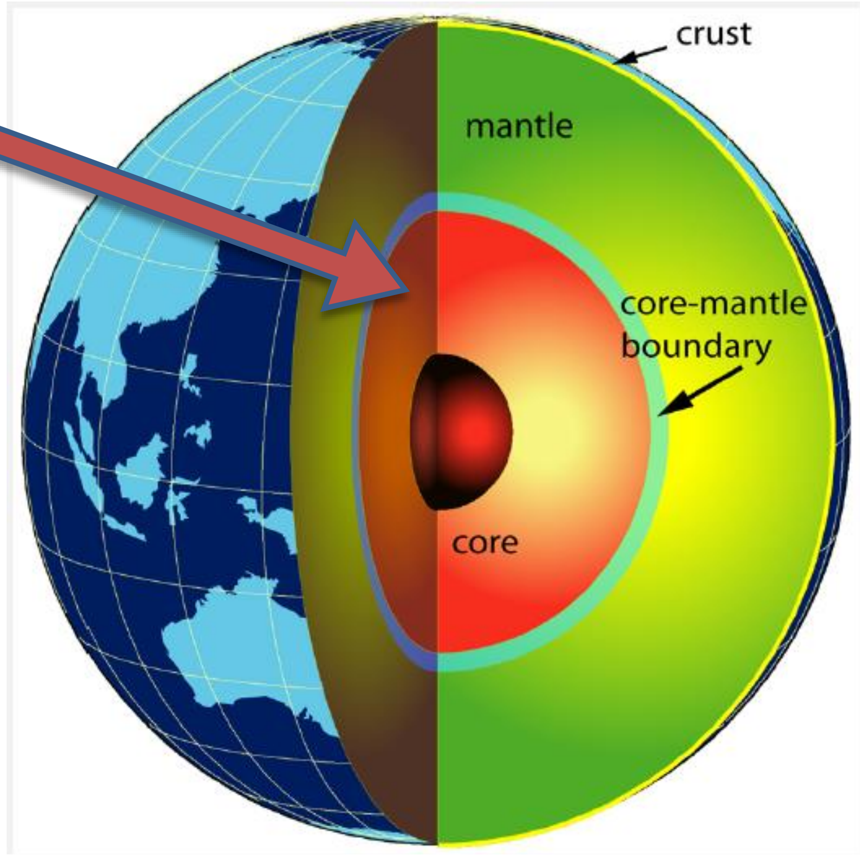


Movie:
Chris Finlay (DTU)

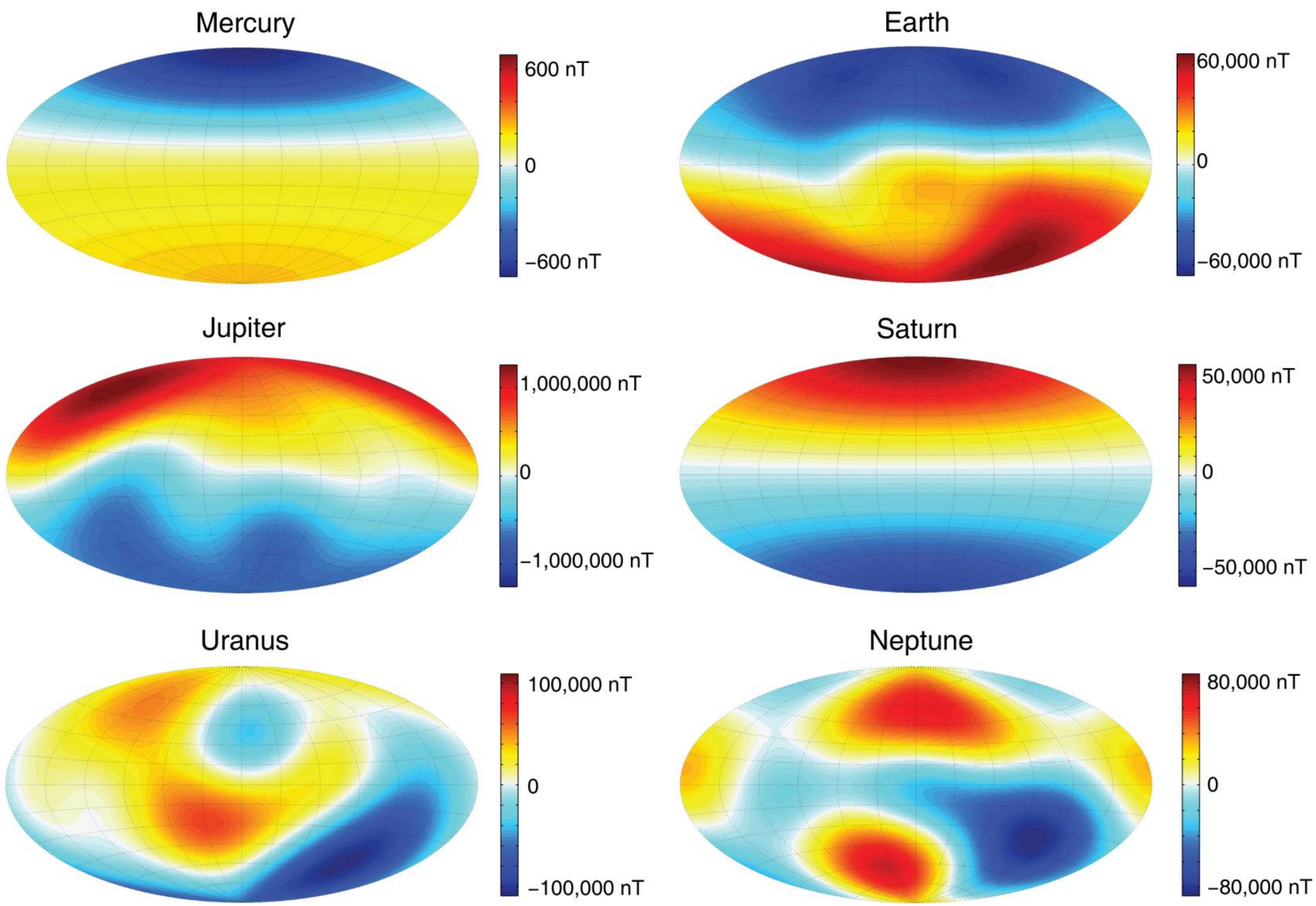
Geomagnetism is <u>Dynamic</u>
Something inside the Earth is causing this variation

Planetary Dynamo Schematic: The Geodynamo

- Liquid iron core:
  - Convection + Induction
  - Spherical geometry
- Thermal or compositional forcing:
  - Latent heat release
  - Light element release
- Difficult to observe directly:
  - Remote
  - Mantle-filtering



Most Planets Possess Magnetic Fields

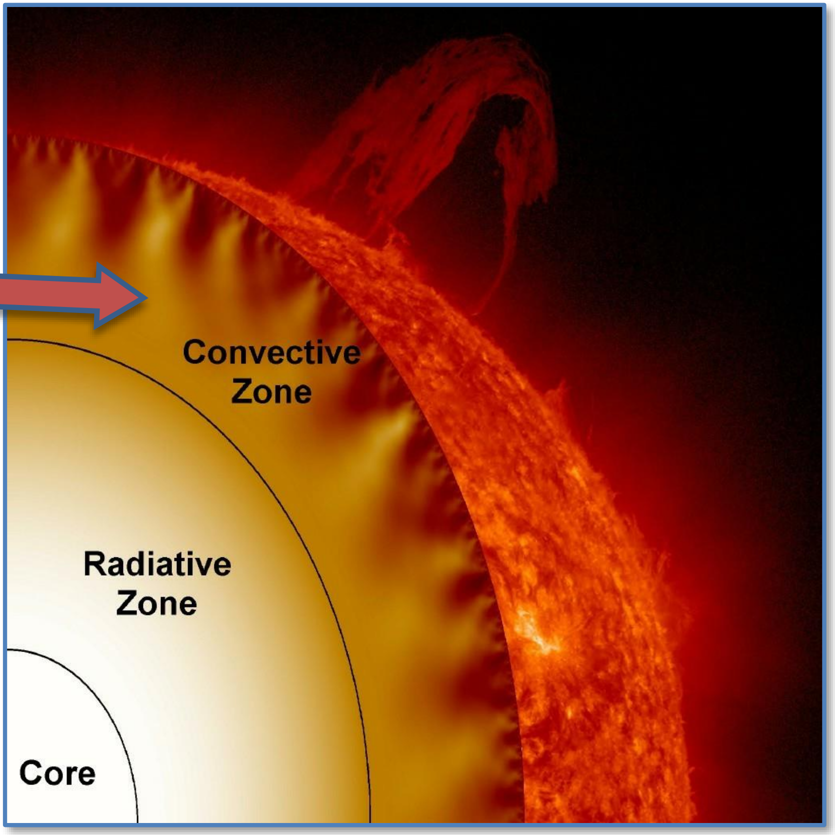


[Cao 2014]

…and of course the Sun too…

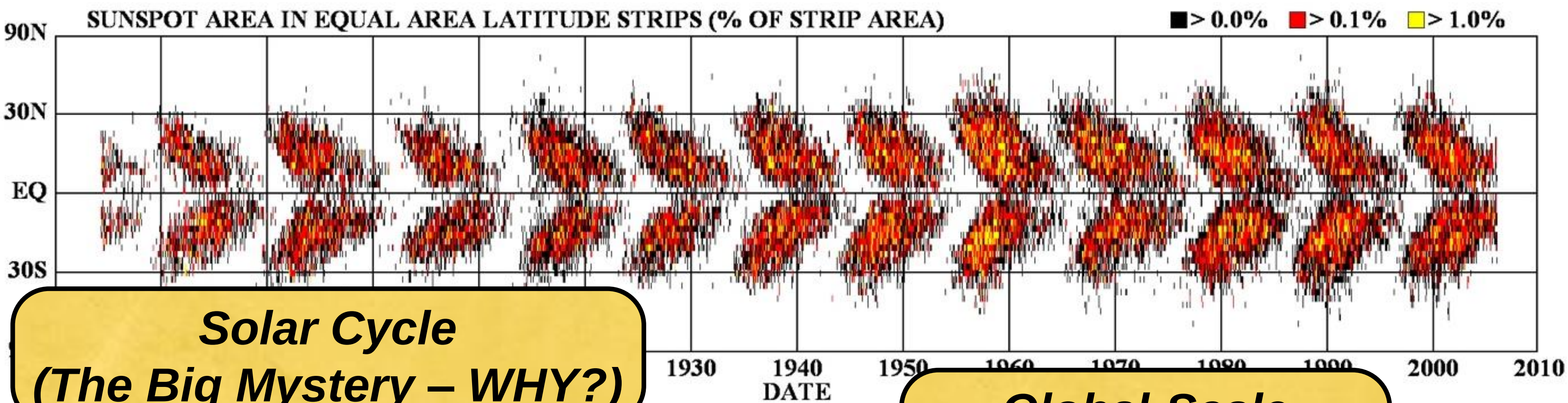
Stellar Dynamo Schematic: The Sun

- Dense plasma throughout:
  - Convection + Induction
  - Spherical geometry
- Thermal or compositional forcing:
  - Core fusion
- Difficult to observe below suface:
  - Helioseismology
- Magnetism is EVERYWHERE



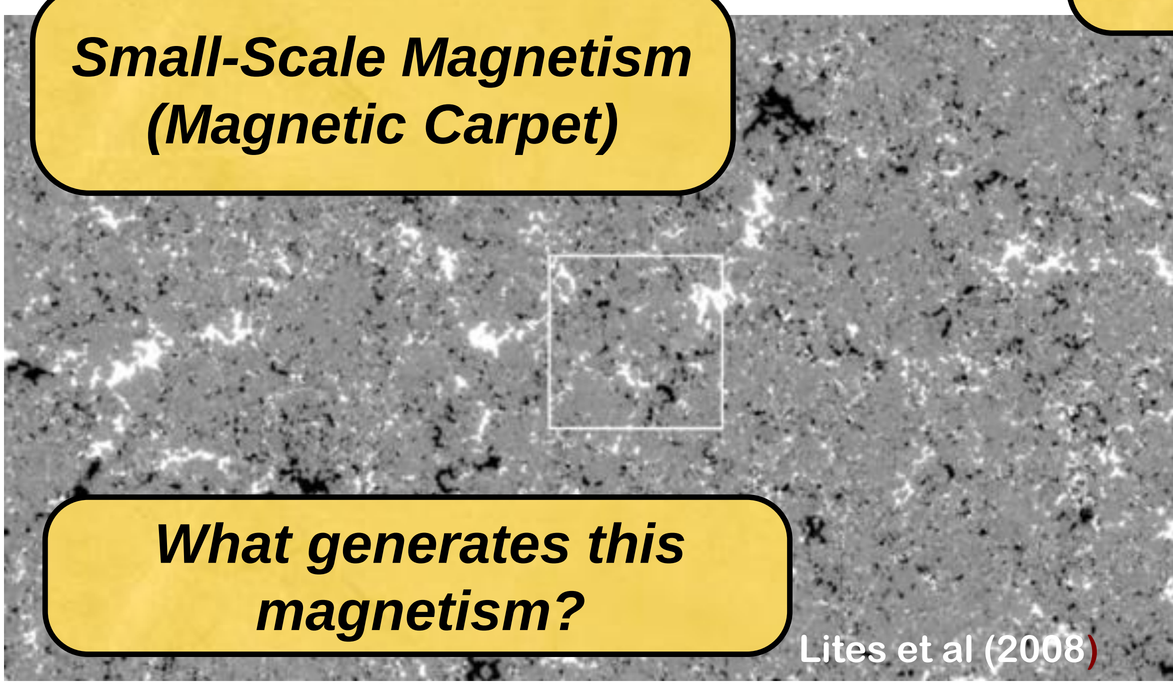
The Magnetic Sun

D. Hathaway (NASA MSFC)



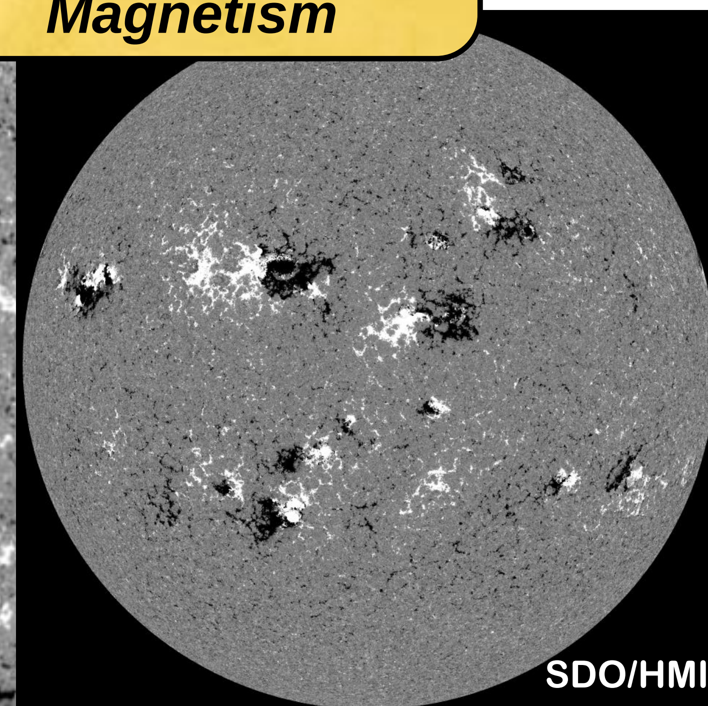
Solar Cycle
(The Big Mystery – WHY?)

Global-Scale Magnetism



Small-Scale Magnetism
(Magnetic Carpet)

What generates this magnetism?



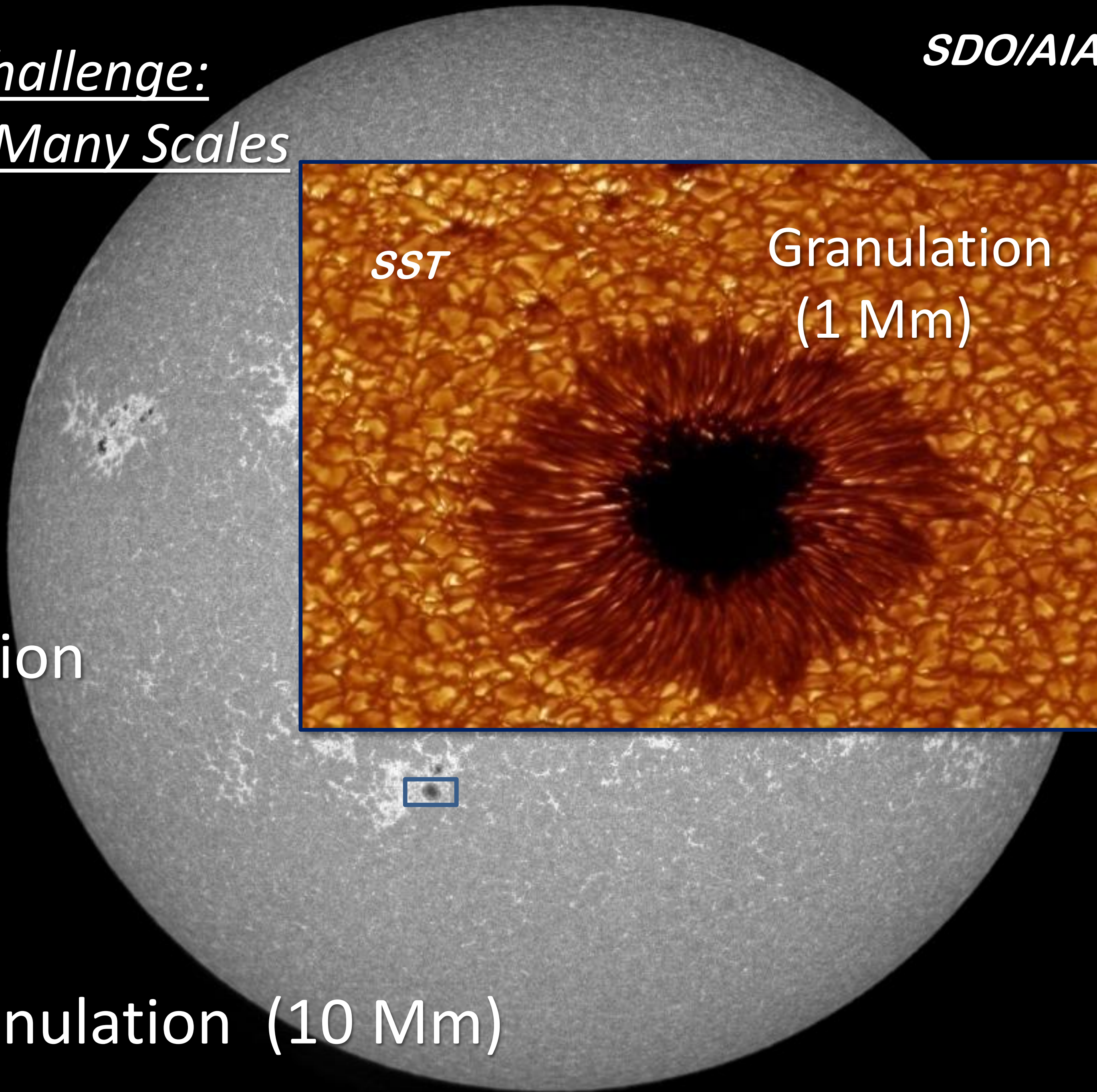
The Big Question:

How do any of these rotating bodies generate a magnetic field?

The Challenge:

- ALL of these examples possess a WIDE range of spatial scales of convection.
- We have to resolve the big stuff (spherical-scale)
- We also have to resolve the small stuff (application dependent)

The Solar Challenge: Convection on Many Scales

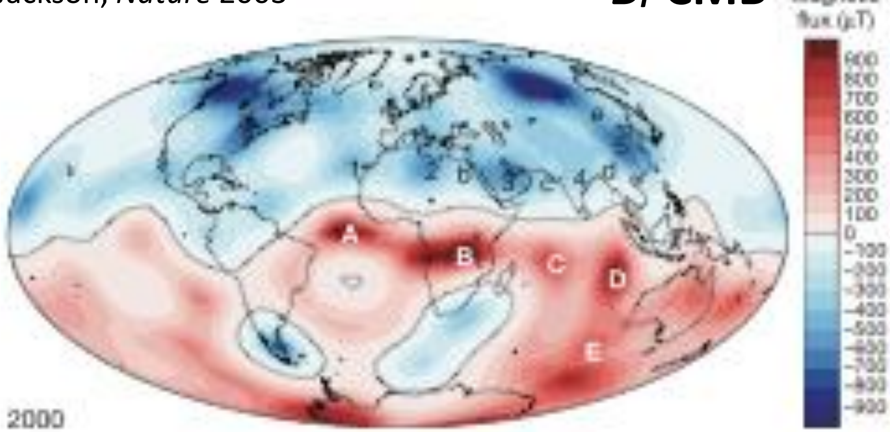


Observations

Models

Jackson, *Nature* 2003

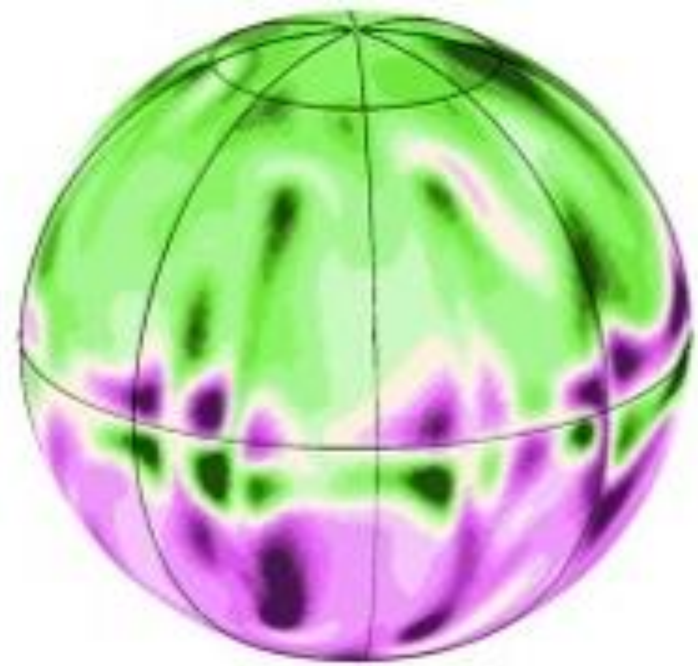
B_r CMB



L_{max} ~ 13

Soderlund et al. *EPSL* 2012

B_r CMB



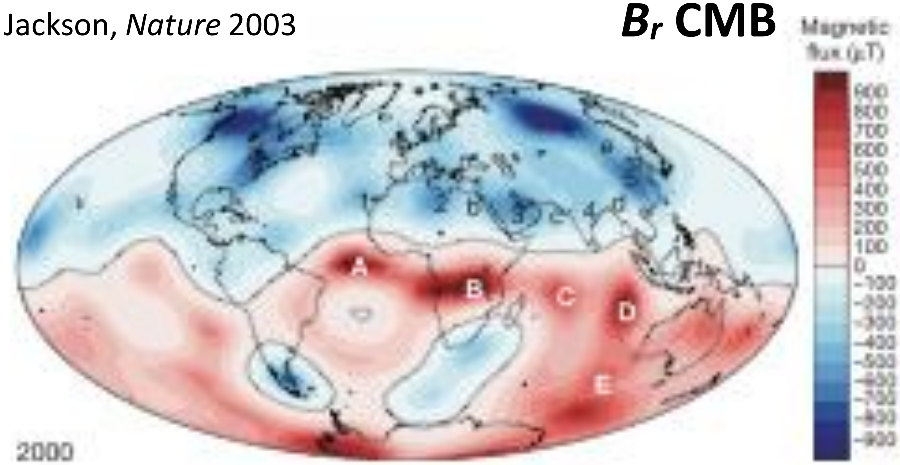
On the surface, things look pretty good…

Observations

Models

Jackson, *Nature* 2003

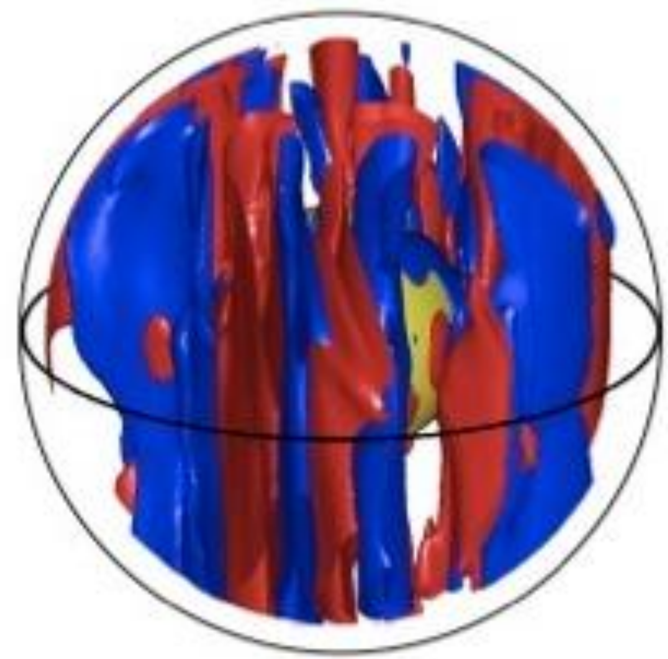
B_r **CMB**



L_{max} **~ 13**

Soderlund et al. *EPSL* 2012

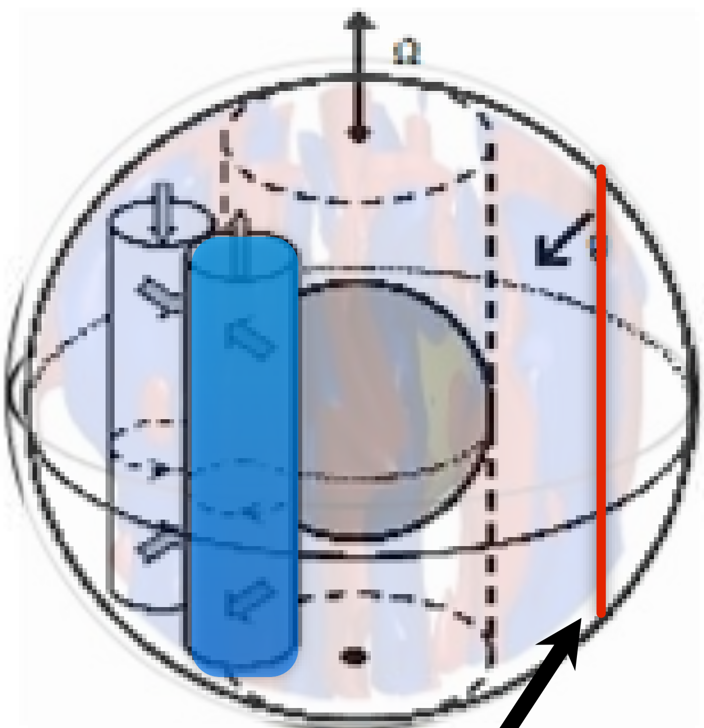
***z*-vorticity**



Beneath the surface …

… probably unphysical

Rotating Convection Columns: column size set by Ekman number E



10^3 too wide

$$E = \frac{\tau_{rotation}}{\tau_{viscous}} = \frac{\nu}{\Omega L^2}$$

Models:

E ~ 1e-4; l_c ~ 0.1

Earth's Core:

E ~ 1e-15; l_c ~ 1e-5

(i.e., 10^4 x smaller than scale of flux patches)

What is Rayleigh? Not just a number…

Rotating, anelastic MHD convection

Spectral: Spherical Harmonics
Chebyshev or FD

Scalable: 2048^3 on 132,000 cores

Open Source: Spring/Summer 2015

COMPUTATIONAL INFRASTRUCTURE for GEODYNAMICS

Geodynamo Working Group

Jon Aurnou, Ben Brown, Bruce Buffet,
Nick Featherstone, Gary Glatzmaier,
Moritz Heimpel, Lorraine Hwang, Louise Kellog,
Hiro Matsui, Peter Olson, Sabine Stanley

Ingredients of the Solar Convection Zone

- Density & Temperature Stratification
  11 density scale heights across the layer…
- Radiative heating
- Photospheric cooling (radiative transfer)
- Rotation
- Magnetism
- Sphericity
- Stable radiative zone/Convective overshoot
- Tachocline (ingredient or result?)
- Extreme values of non-dimensional parameters:
  Reynolds Number $\approx O(10^{12} - 10^{14})$
  Rayleigh Number $\approx O(10^{22} - 10^{24})$

This is tough to model … even badly

Ingredients of a *Model* Solar Convection Zone

- Density & Temperature Stratification

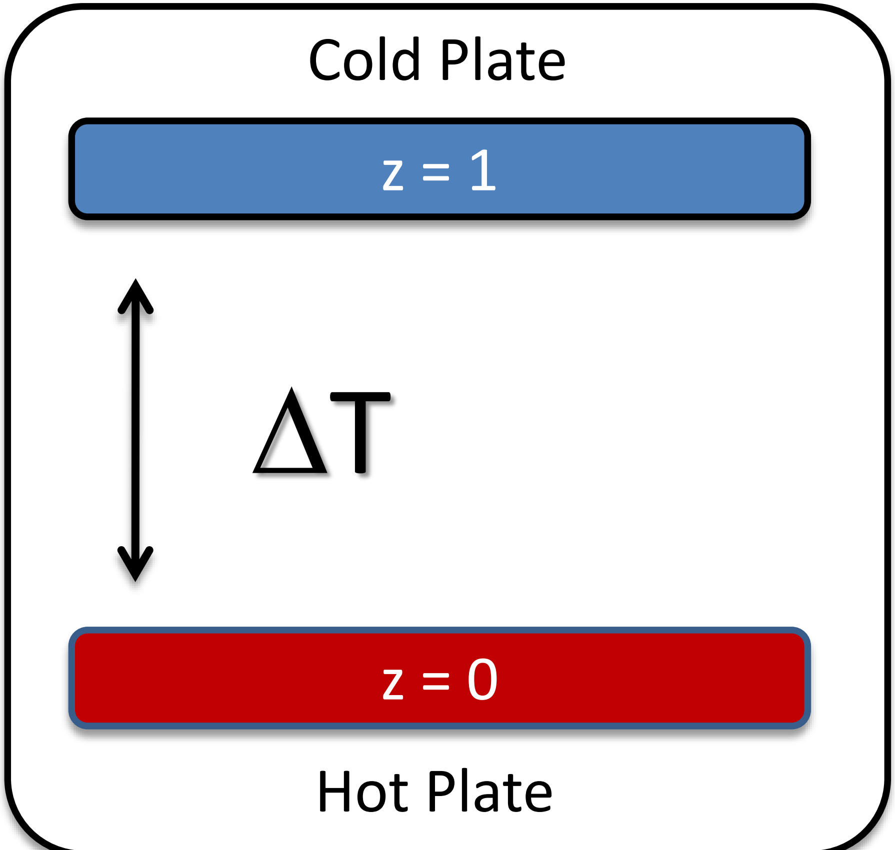
  ~~11~~ 4 or 5 density scale heights across the layer…
- Radiative heating
- ~~Photospheric cooling (radiative transfer)~~
- Rotation
- Magnetism
- Sphericity
- ~~Stable radiative zone/Convective overshoot~~
- ~~Tachocline (ingredient or result?)~~
- *Not so* extreme values of non-dimensional parameters:

  $Re \approx O(10^3)$

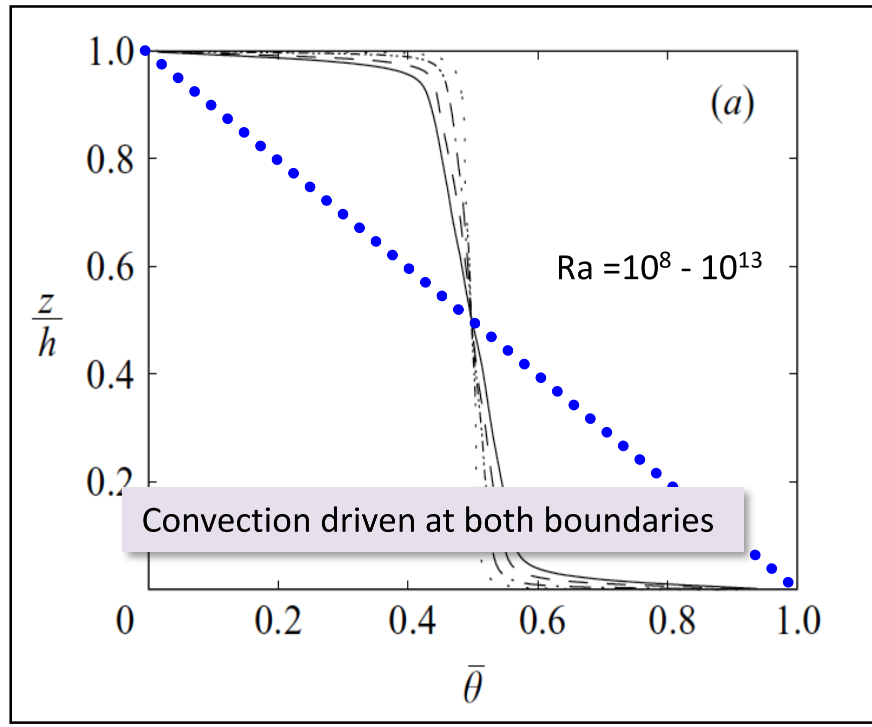
  $Ra \approx O(10^6 - 10^8)$

How do we model *this?* Start simply, and build up…

Modeling Solar Convection I: Start with Rayleigh-Bénard



Temperature Profiles



Verzicco & Sreenivasan (2008)

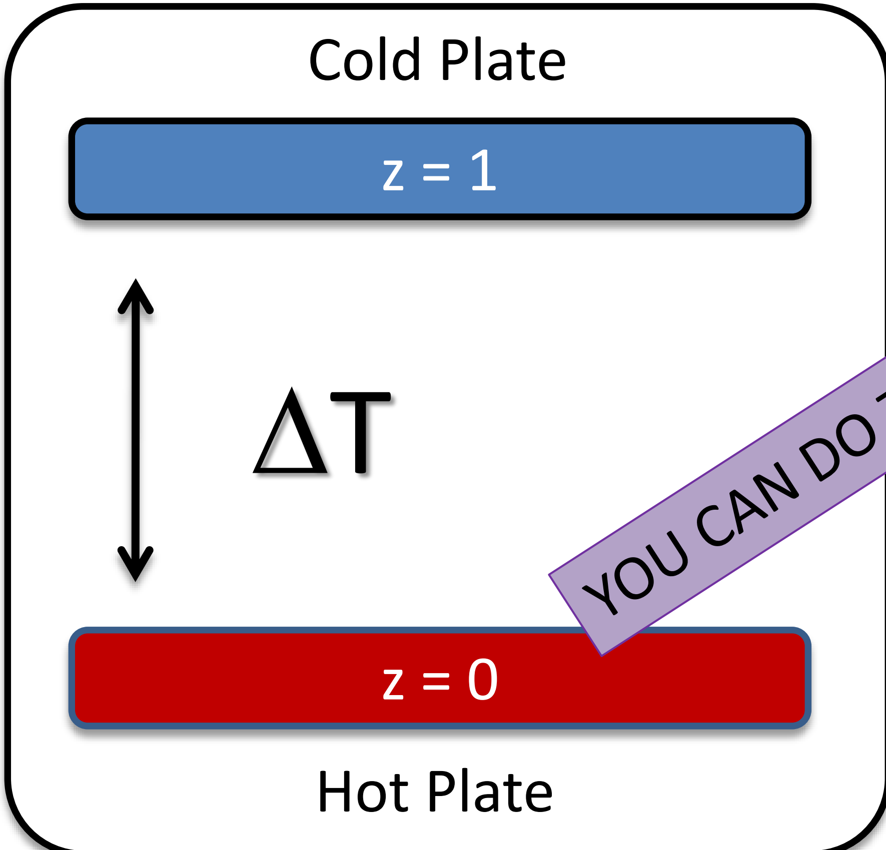
Energetics

Input: *Conduction*

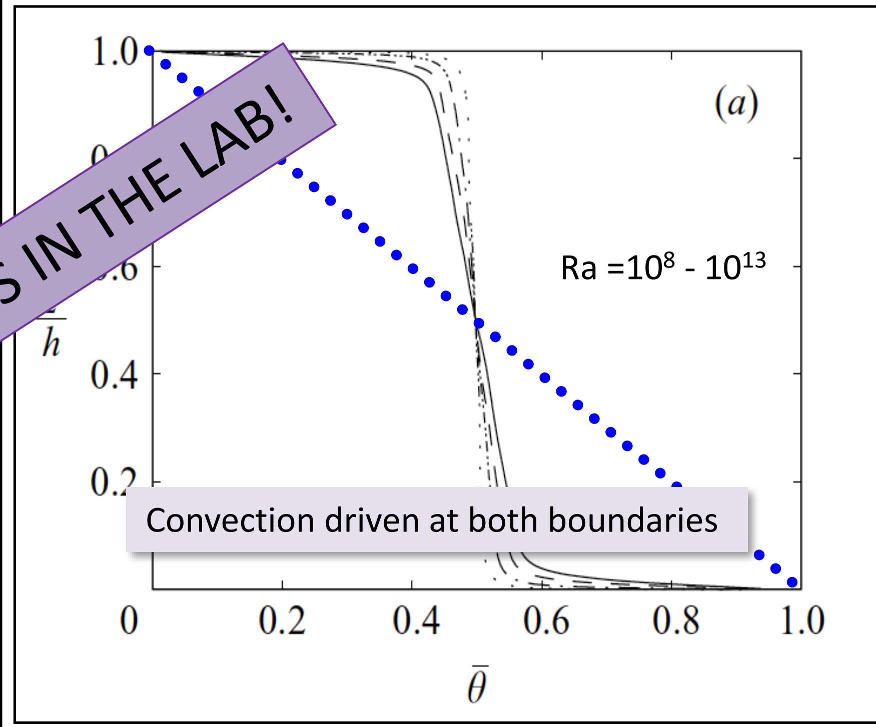
Throughput: *Rayleigh Number*

Output: *Conduction*

Modeling Solar Convection I: Start with Rayleigh-Bénard



YOU CAN DO THIS IN THE LAB!



Verzicco & Sreenivasan (2008)

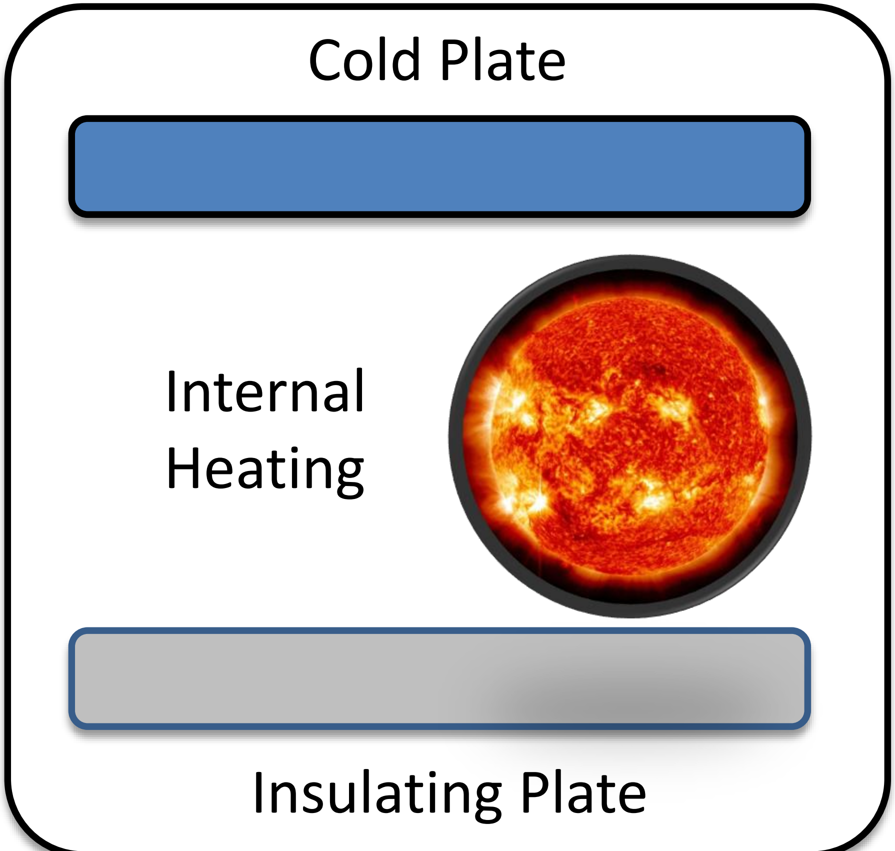
Energetics

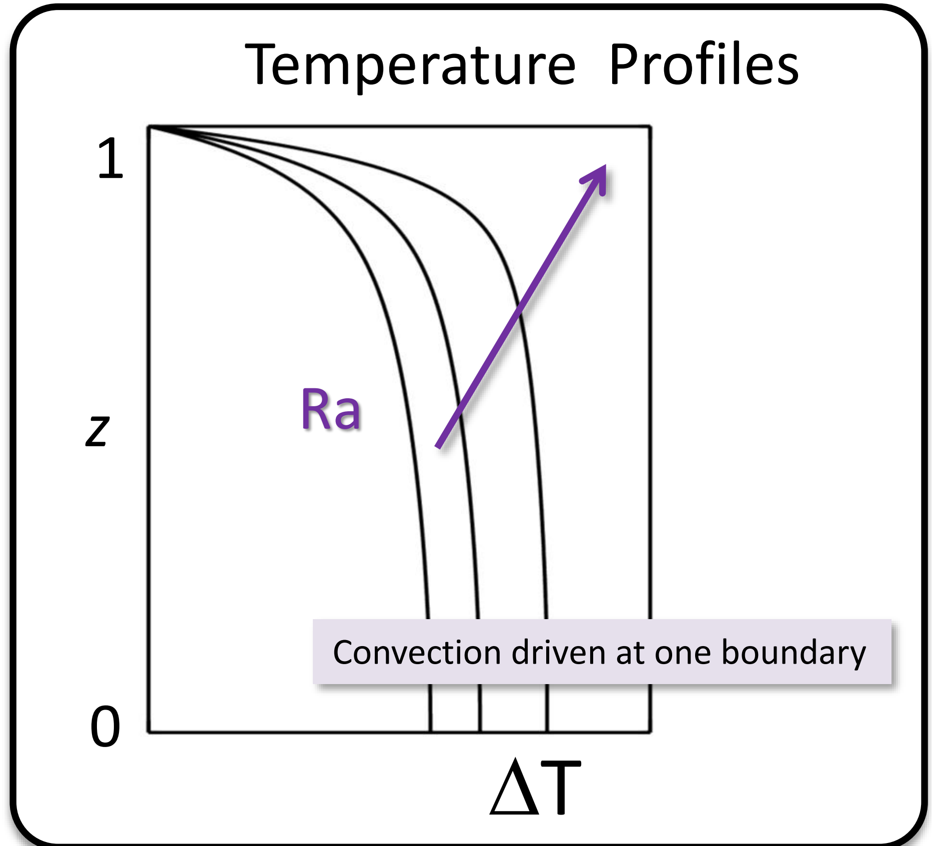
Input: *Conduction*

Throughput: *Rayleigh Number*

Output: *Conduction*

Modeling Solar Convection II: Internal Heating





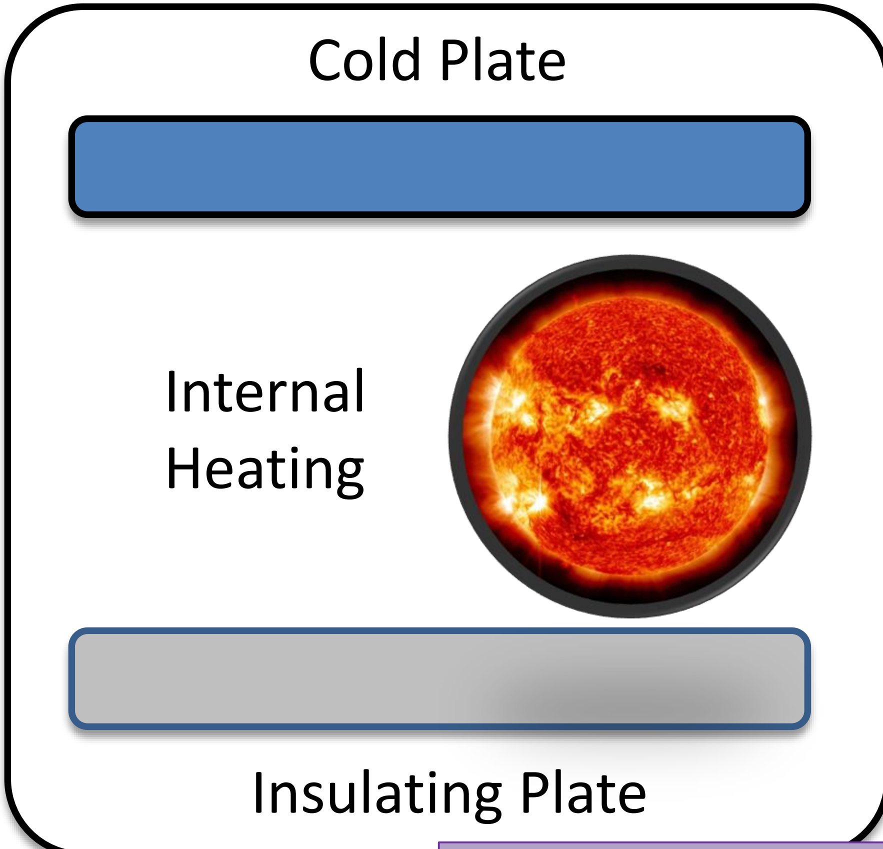
Energetics

Throughput: *Solar Luminosity*

Input: *Radiative Heating*

Output: *Conduction*

Modeling Solar Convection II: Internal Heating



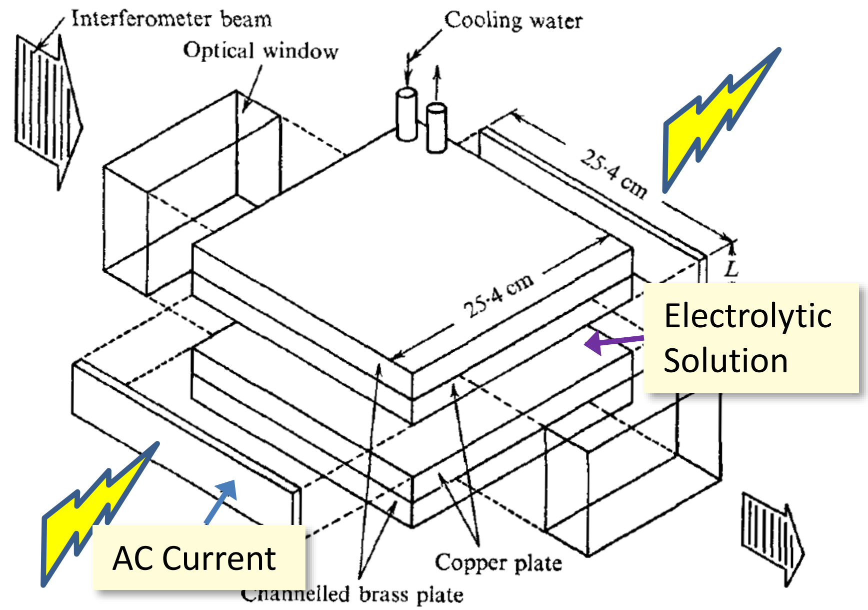
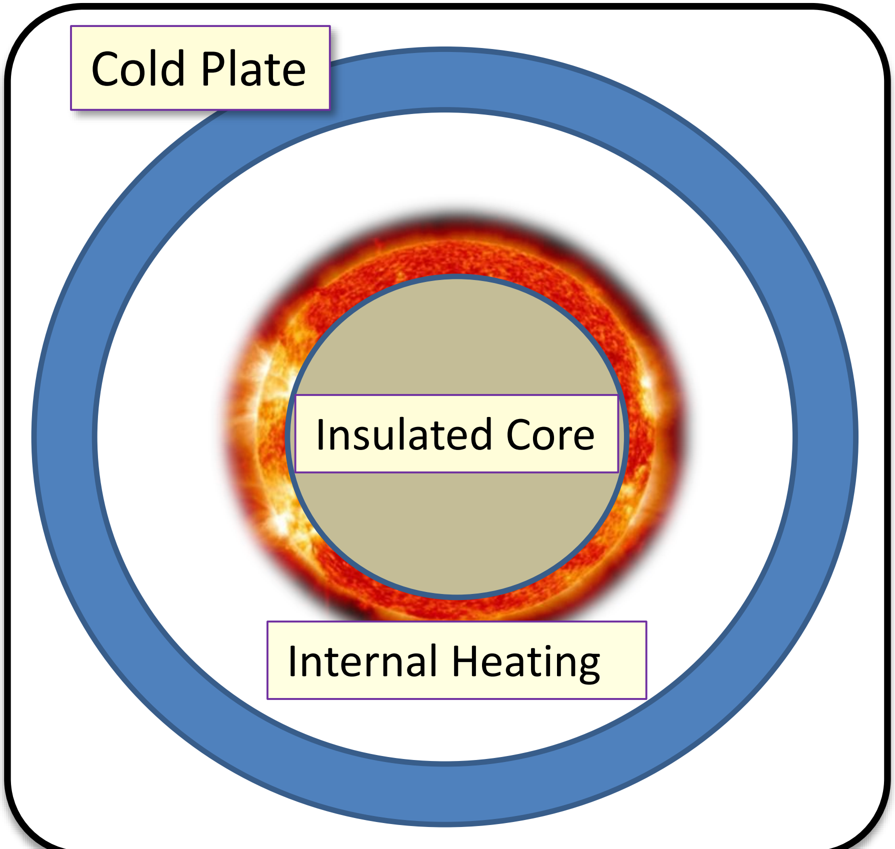


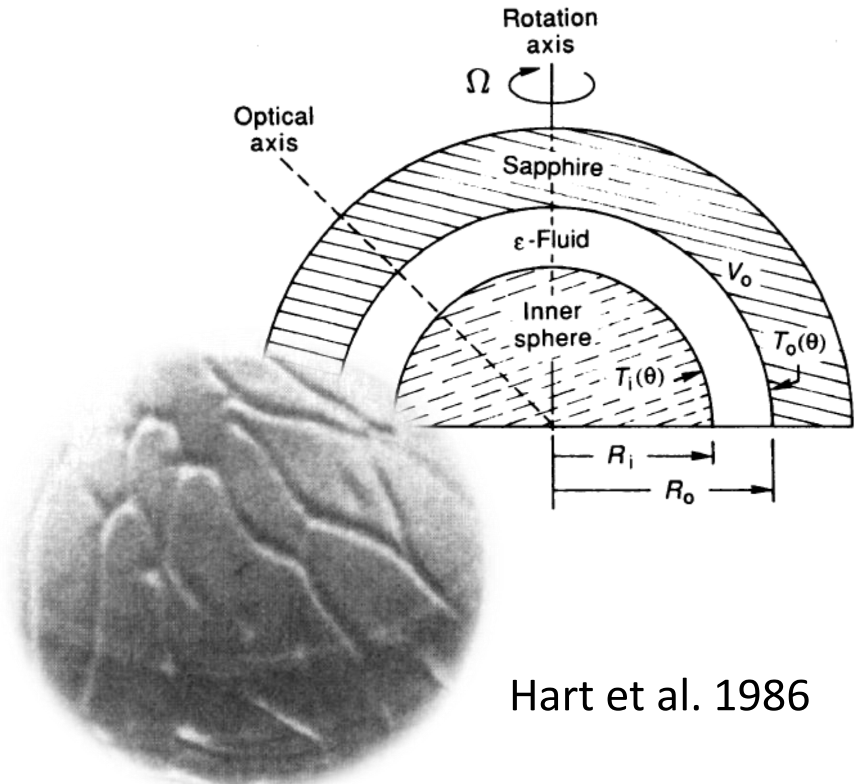
FIGURE 1. Convection chamber. $0{\cdot}127 \text{ cm} \leqslant L \leqslant 6{\cdot}352 \text{ cm}$.

Kulacki & Goldstein 1972

YOU CAN ALSO DO THIS IN THE LAB!

Modeling Solar Convection III: Sphericity





Hart et al. 1986

Can actually do this in the lab.
(space shuttle required…)

Density stratification,
magnetism,
still tough in the lab…

Most Models: Anelastic MHD Formulation

$$\frac{D\boldsymbol{v}}{Dt} = -2\Omega\hat{\boldsymbol{z}} \times \boldsymbol{v} - \boldsymbol{\nabla}\frac{P}{\bar{\rho}} + \boldsymbol{g}\frac{S}{c_p} + \frac{1}{4\pi\bar{\rho}}(\boldsymbol{\nabla} \times \boldsymbol{B}) \times \boldsymbol{B} + \frac{1}{\bar{\rho}}\boldsymbol{\nabla} \cdot \boldsymbol{D}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) - \nabla \times (\eta \boldsymbol{\nabla} \times \boldsymbol{B})$$

$$\boldsymbol{\nabla} \cdot \bar{\rho}\boldsymbol{v} = \boldsymbol{0}$$

$$\bar{\rho}\bar{T}\frac{DS}{Dt} = \nabla \cdot (\bar{\rho}\bar{T}\kappa\nabla S) + Qi + Qo + Qv$$

$$\mathsf{D}_{ij} \equiv 2\bar{\rho}\nu\left(e_{ij} - \frac{1}{3}(\nabla \cdot v)\delta_{ij}\right)$$

$$Q_i \equiv Internal\ Heating$$

$$Q_o \equiv \frac{1}{4\pi}(\boldsymbol{\nabla} \times \boldsymbol{B})^2$$

$$Q_v \equiv 2\bar{\rho}\nu\left(e_{ij}e_{ij} - \frac{1}{3}(\nabla \cdot v)^2\right)$$

Assumes perturbations about background thermal state are small

$$\frac{P}{\overline{P}} = \frac{\rho}{\overline{\rho}} + \frac{T}{\overline{T}}$$

fully compressible codes used as well

Serial Numerics

Horizontal "Discretization"

Spherical Harmonics: FFTW + DGEMM (Legendre)

Radial "Discretization"

Chebyshev Polynomials: Colocation scheme

Time-Stepping: Hybrid Crank-Nicolson/Adams-Bashforth

Direct Matrix Solve: LAPack LU Decomposition routines

Recalculate matrices when Δt changes

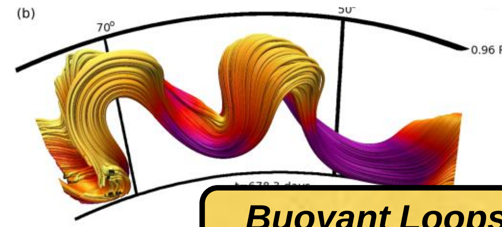
Brief Background

Gary (Dynamo)

1995

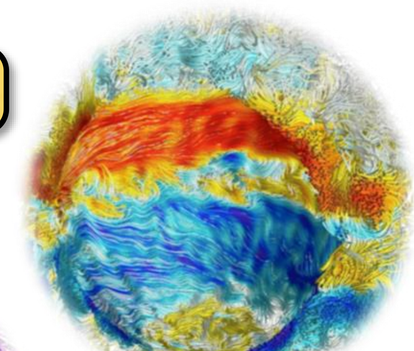
Juri Toomre
Tom Clune
Mark Miesch

ASH



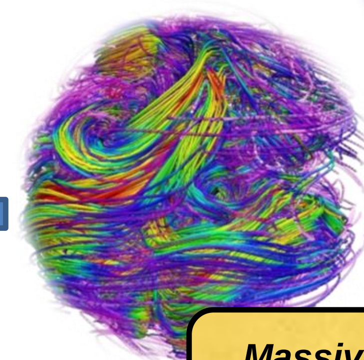
Buoyant Loops

Nelson et al. 2012



Coherent Magnetic Structures

Brown et al. 2010



Massive Stars

Featherstone et al. 2009

2010

ASH-FD

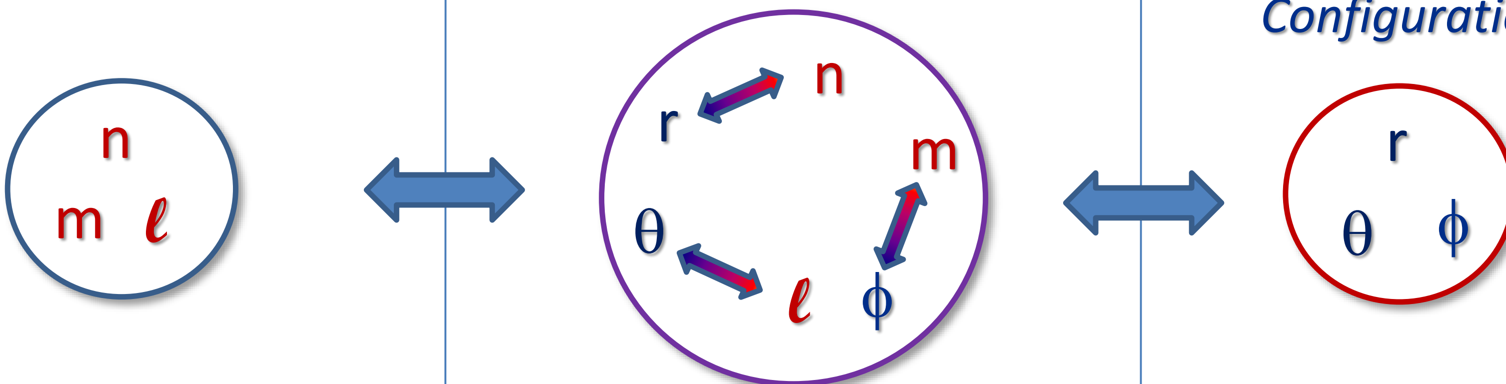
2012

COMPUTATIONAL INFRASTRUCTURE for GEODYNAMICS

~~PseudoCode~~
~~ACDC~~
Rayleigh

Conceptual View of a Pseudo-Spectral Approach

Spectral Configuration | *Hybrid Configuration* | *Physical Space Configuration*



- Linear terms

- Transforms
- Derivatives

- Nonlinear Terms

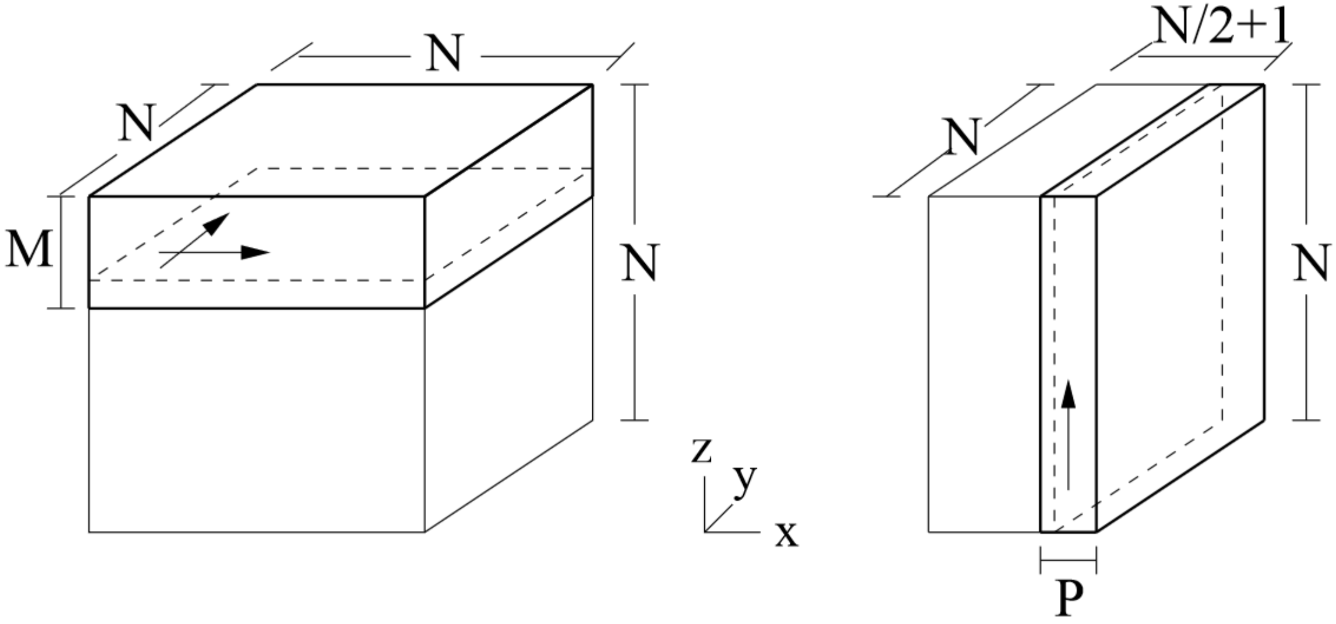
Movement between configurations requires:

Transforms: O(N^2) and O(N log N) … expensive but accurate

Transposes: All-to-Alls … limit scalability

Transpose: GLOBAL rearrangement of data

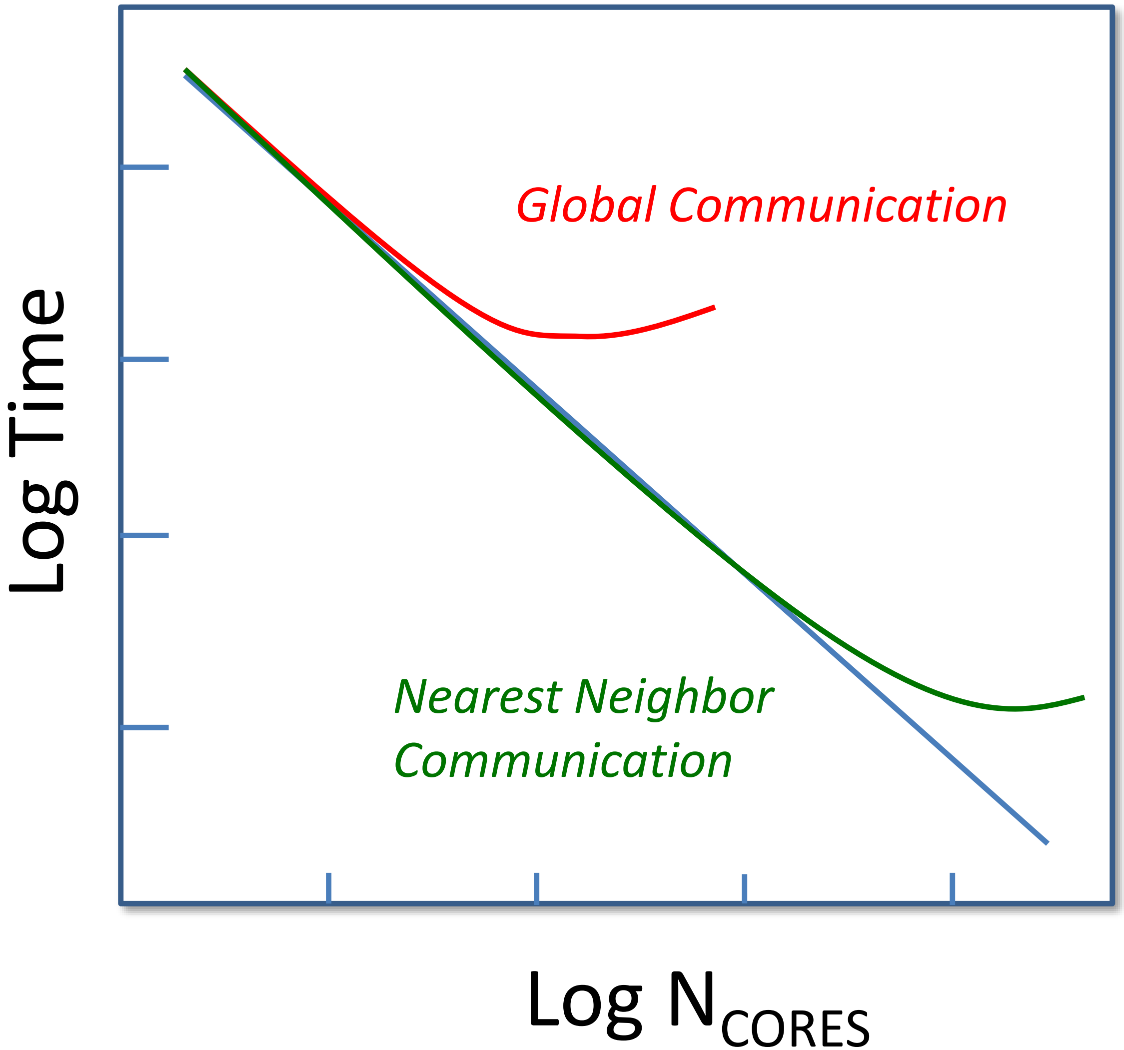
Why a Transpose?



Think of a 3-D FFT

Adapted from D.O. Go'mez et al. 2005

The Problem with Spectral Methods



Strong Scalability

Ideal Scaling:

$$\text{Time} \propto \frac{1}{N_{\text{CORES}}}$$

The game:
mitigation

How Long Should an All-to-All Take?

Time = *Initiation Time* + *Transmission Time*

Local Problem Size:	P
Number of MPI Ranks:	N
Single Message Initiation Time:	I
Bandwidth:	B

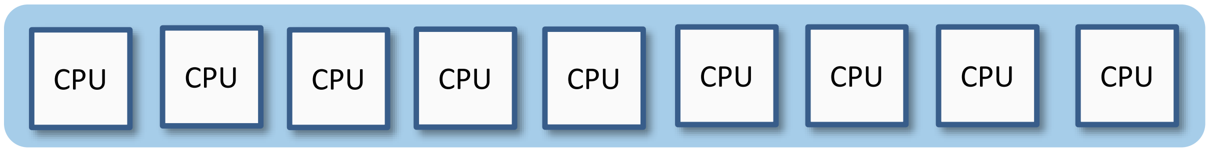
Transmission Time $= N \times \overbrace{(P/N)}^{\text{Message Size}} / B = P/B$ … *constant*

Initiation Time $= N \times \mathrm{I}$ … *growing*

Try to limit message count

Approach #1: 2-D Domain Decomposition

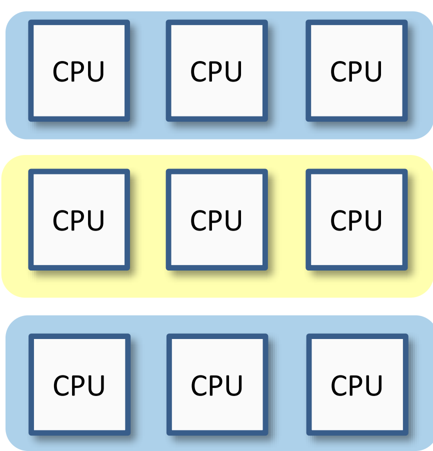
1-D Domain Decomposition



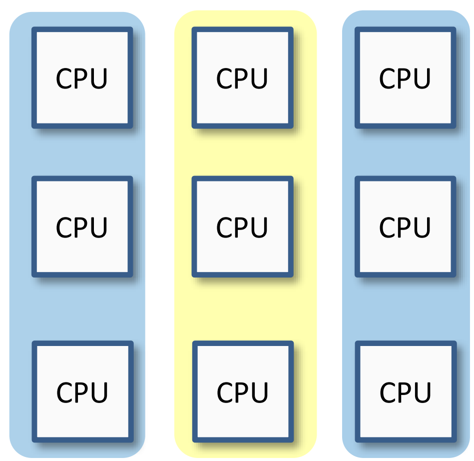
$$\mathrm{Time} = \frac{\mathrm{P}}{\mathrm{B}} + \mathrm{NI}$$

One Large All-to-All

2-D Domain decomposition



2 Passes



$$\mathrm{Time} = \frac{2\mathrm{P}}{\mathrm{B}} + 2\mathrm{I}\sqrt{\mathrm{N}}$$

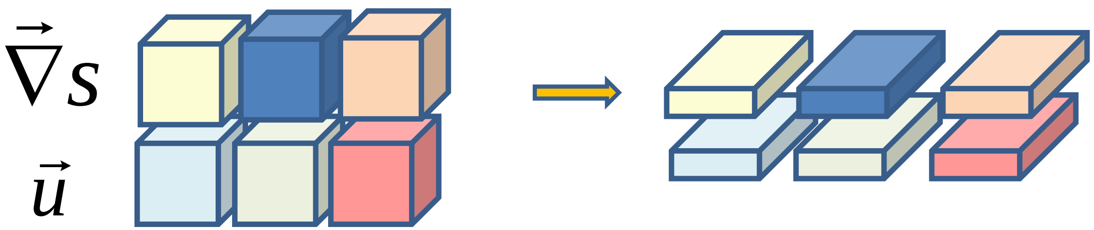
Higher Max N_{CORES}

Approach #2: Collect the Collectives

Example: Entropy Advection

Obvious approach: *transpose each field*

Six Transposes



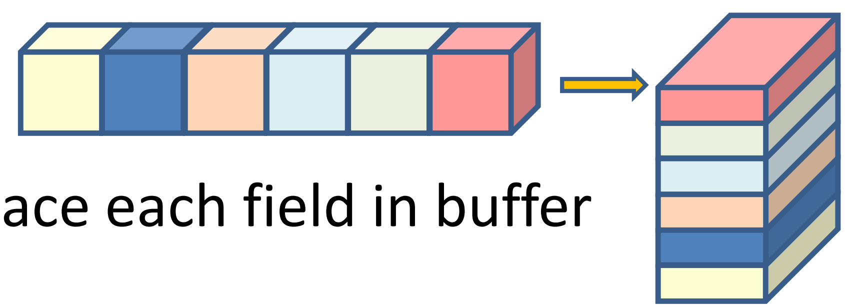
Spectral space

Physical space

$$T = \frac{6P}{B} + 6NI$$

Alternative approach: *transpose a buffer*

Single Transpose



Place each field in buffer

Transpose the buffer

$$T = \frac{6P}{B} + \underbrace{NI}_{\textit{Lower init time...}}$$

Rayleigh Parallelization: Load Balancing

Triangular Truncation:

$$0 \le \ell \le \ell_{\max} \qquad 0 \le m \le \ell$$

Uniform resolution on the Sphere.

Natural Load Balancing :

Keep all ℓ's for each m in processor.

Pair high and low m modes.

Max Angular NCPUs = $\dfrac{\ell_{\max}+1}{2}$

Example
Mode Distribution

$$\ell_{\max} = 5$$

Distributing m's is awkward

Fold the triangle

m = 0	m = 1	m = 2	m = 3	m = 4	m = 5
ℓ	ℓ	ℓ	ℓ	ℓ	ℓ
0	1	2	3	4	5
1	2	3	4	5	
2	3	4	5		
3	4	5			
4	5				
5					

Rayleigh Parallelization: Load Balancing

Triangular Truncation:

$$0 \le \ell \le \ell_{\max} \qquad 0 \le m \le \ell$$

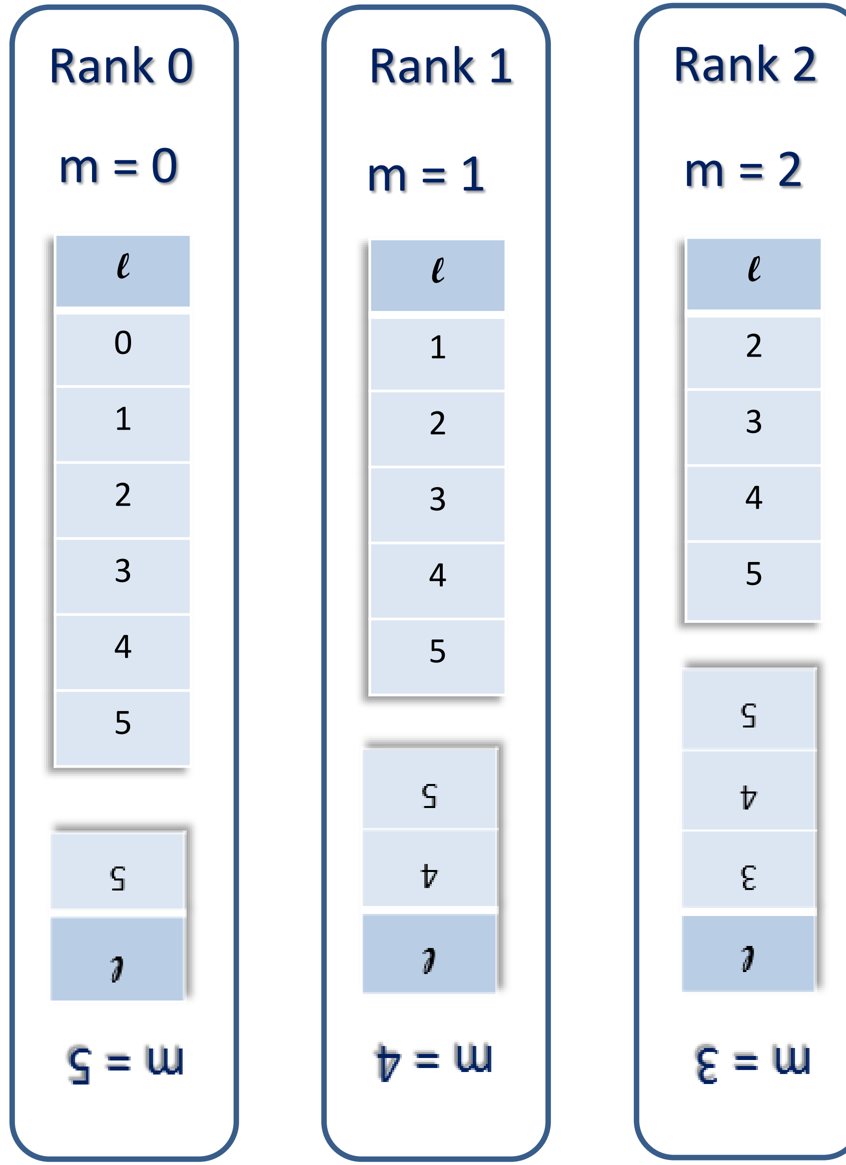
Uniform resolution on the Sphere.

Example
Mode Distribution

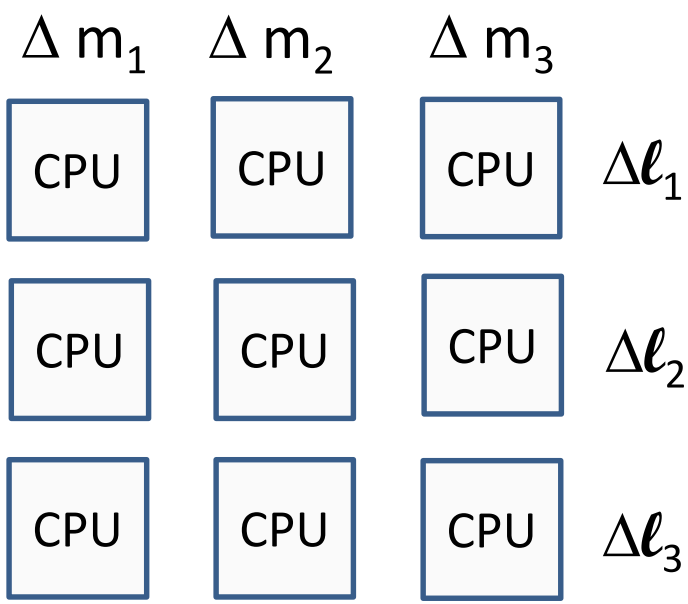
$$\ell_{\max} = 5$$

Distribute pairs of m-values

Each process gets same number of ℓ-values



Rayleigh Parallelization



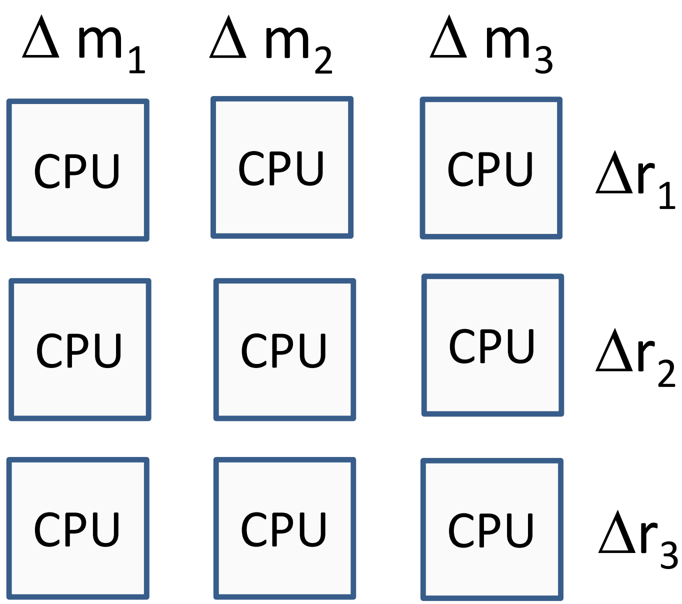
2-D Domain Decomposition

- Pure MPI
- Processes placed in columns and rows
- Three Configurations

Configuration 1

- ℓ-values distributed across rows
- m-values distributed across columns
- Radius in-processor
- Chebyshev Transforms
- Linear Solves

Rayleigh Parallelization



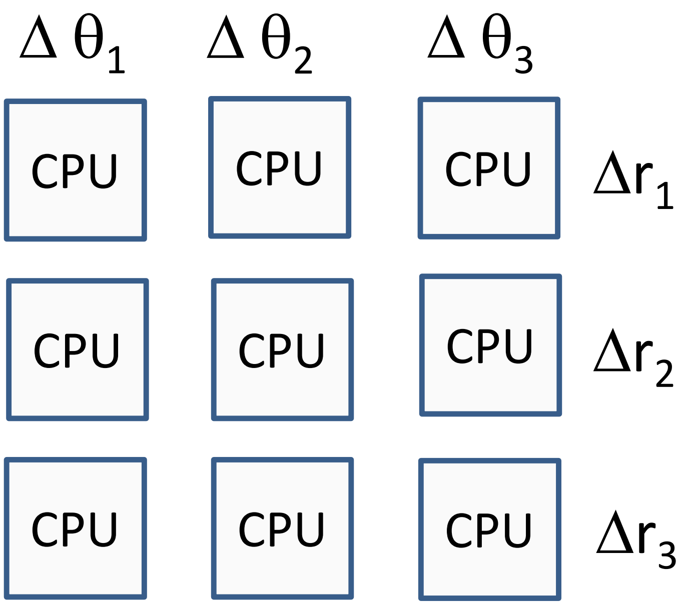
2-D Domain Decomposition

- Pure MPI
- Processes placed in columns and rows
- Three Configurations

Configuration 2

- Radial levels distributed across rows
- m-values distributed across columns
- ℓ / θ in-processor
- Legendre Transforms

Rayleigh Parallelization



2-D Domain Decomposition

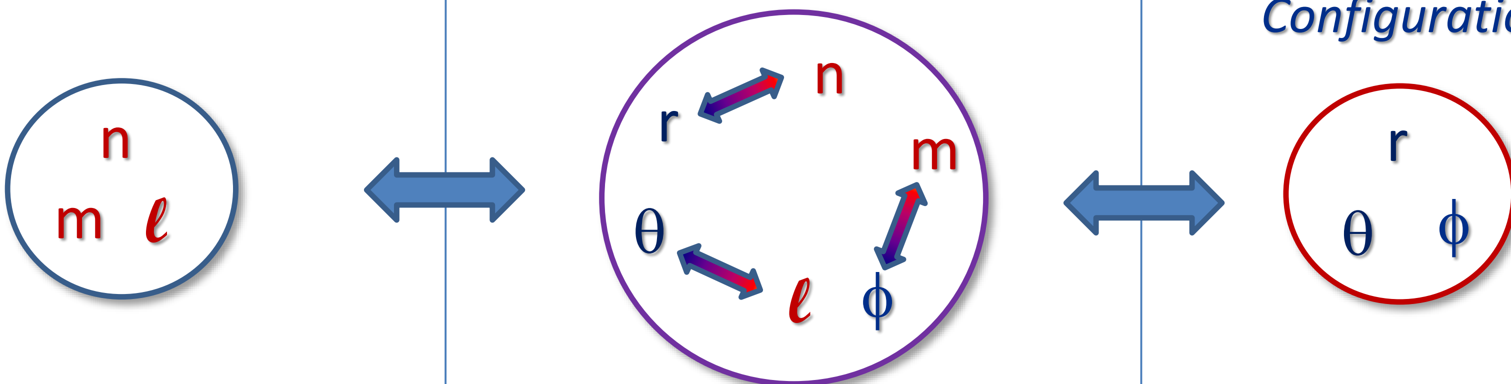
- Pure MPI
- Processes placed in columns and rows
- Three Configurations

Configuration 3

- Radial levels distributed across rows
- θ -values distributed across columns
- m/ϕ in-processor
- Fourier Transforms

Conceptual View of a Pseudo-Spectral Approach

Spectral Configuration | *Hybrid Configuration* | *Physical Space Configuration*



- Linear terms

- Transforms
- Derivatives

- Nonlinear Terms

Movement between configurations requires:

Transforms: $O(N^2)$ and $O(N \log N)$ … expensive but accurate

Transposes: All-to-Alls … limit scalability

The Flow of Rayleigh

Spectral Space

r in-processor

(ℓ, m distributed)

1. Time-stepping
2. Radial Derivatives

Transpose

Hybrid Space

ℓ in-processor

(r, m distributed)

1. θ-derivatives
2. Legendre Transforms

Transpose

Physical Space

ϕ in-processor

(r, θ distributed)

1. ϕ-derivatives
2. Fourier Transforms
3. Nonlinear Terms

Transpose

Hybrid Space

ℓ in-processor

(r, m distributed)

1. θ-derivatives
2. Legendre Transforms

Transpose

1 iteration

4 Calls to All-to-All

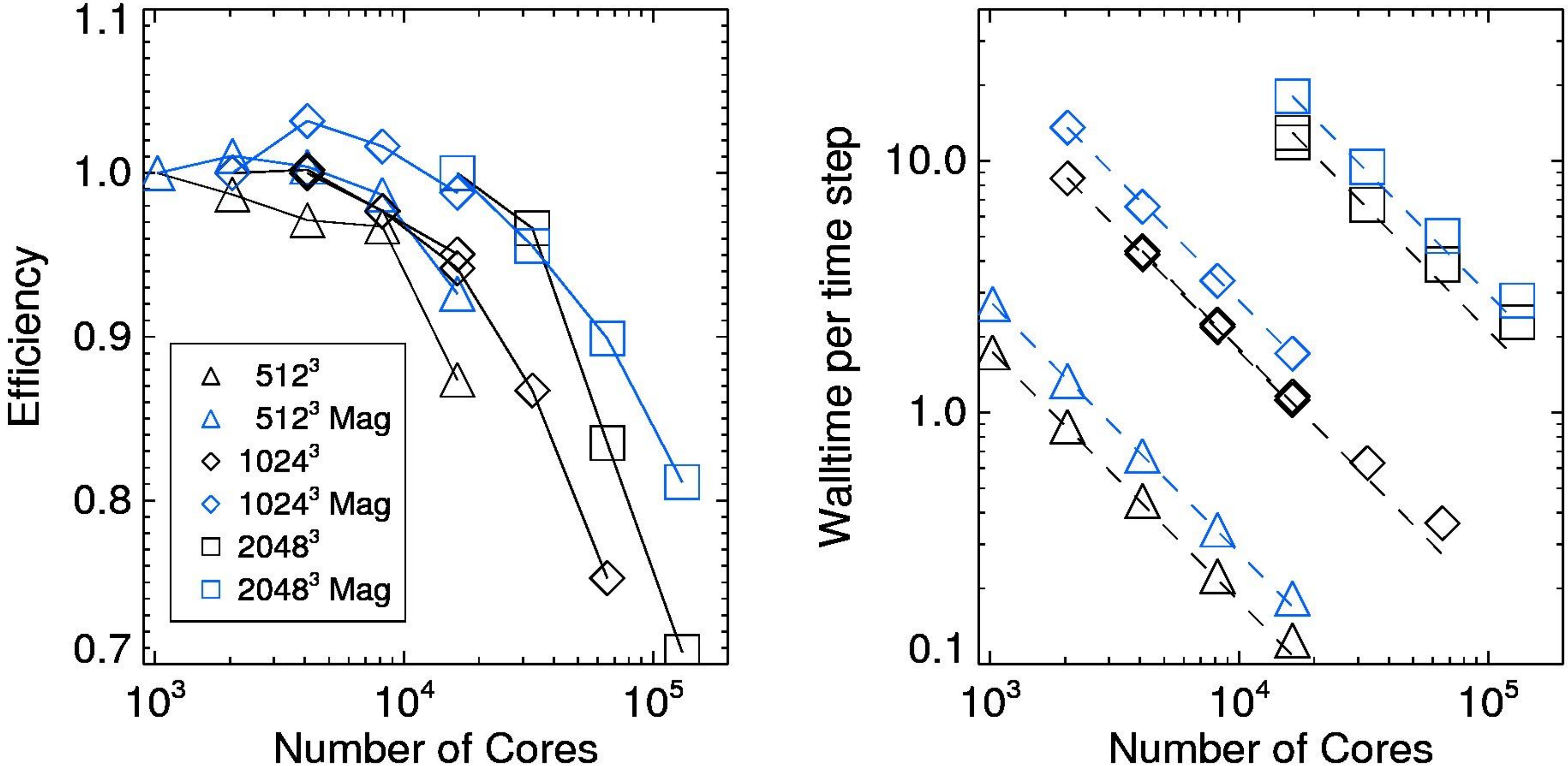
What about Memory?

Assumes $N_r = ½_{\text{Ntheta}}$

N_{theta}	Max CPUs	10% Max CPUs
256	50 kB	500 kB
512	100 kB	1 MB
1024	200 kB	2 MB

Largest Buffer holds 20 Fields
Often fits into cache…

Rayleigh Performance

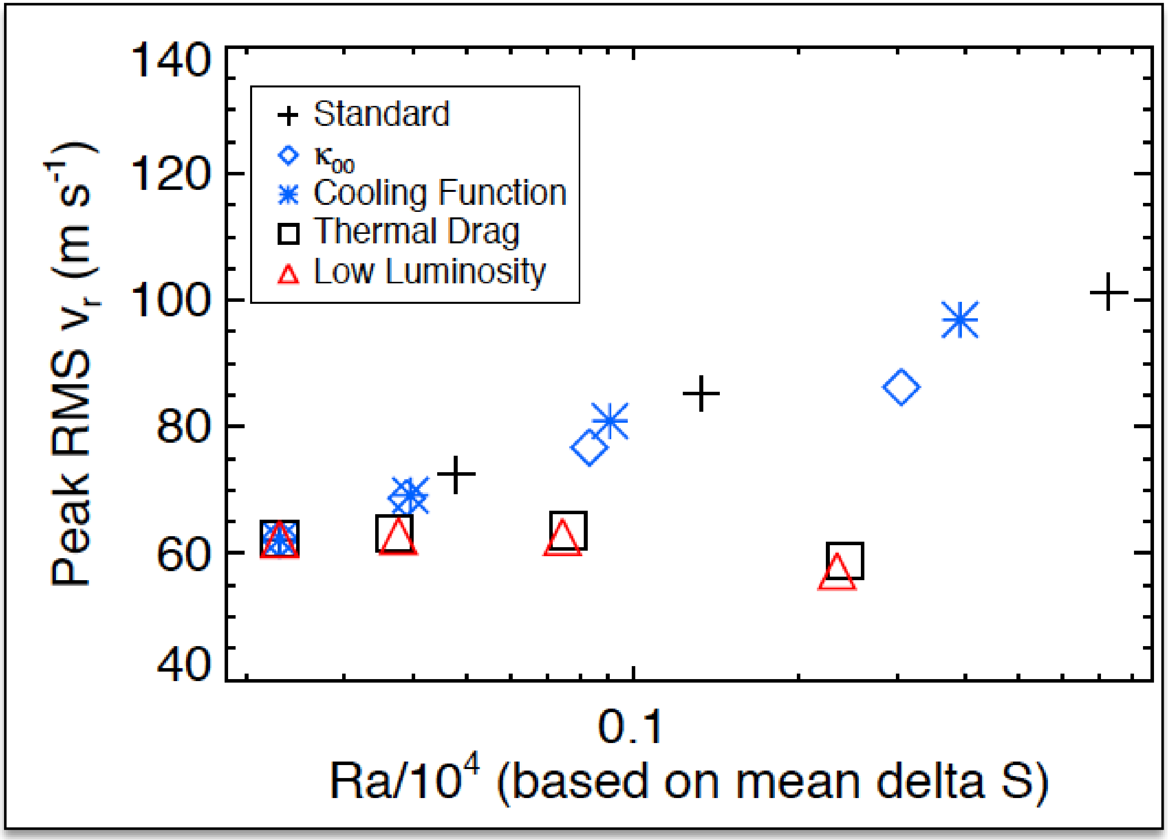


Mira (IBM Blue Gene/Q Argonne)

What can you do with one really scalable spectral code?

- Write computing time proposals…
- Quick parameter studies
- Fast vis runs
- Really big things (see #1)

Large Parameter Space Study
1 Week (Winter Break)



17 simulations

256^3 to 1024^3 grid points

300,000 to

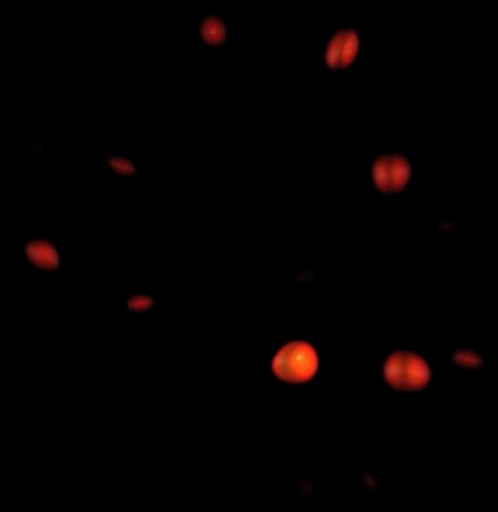
1 million time steps

Warm plumes

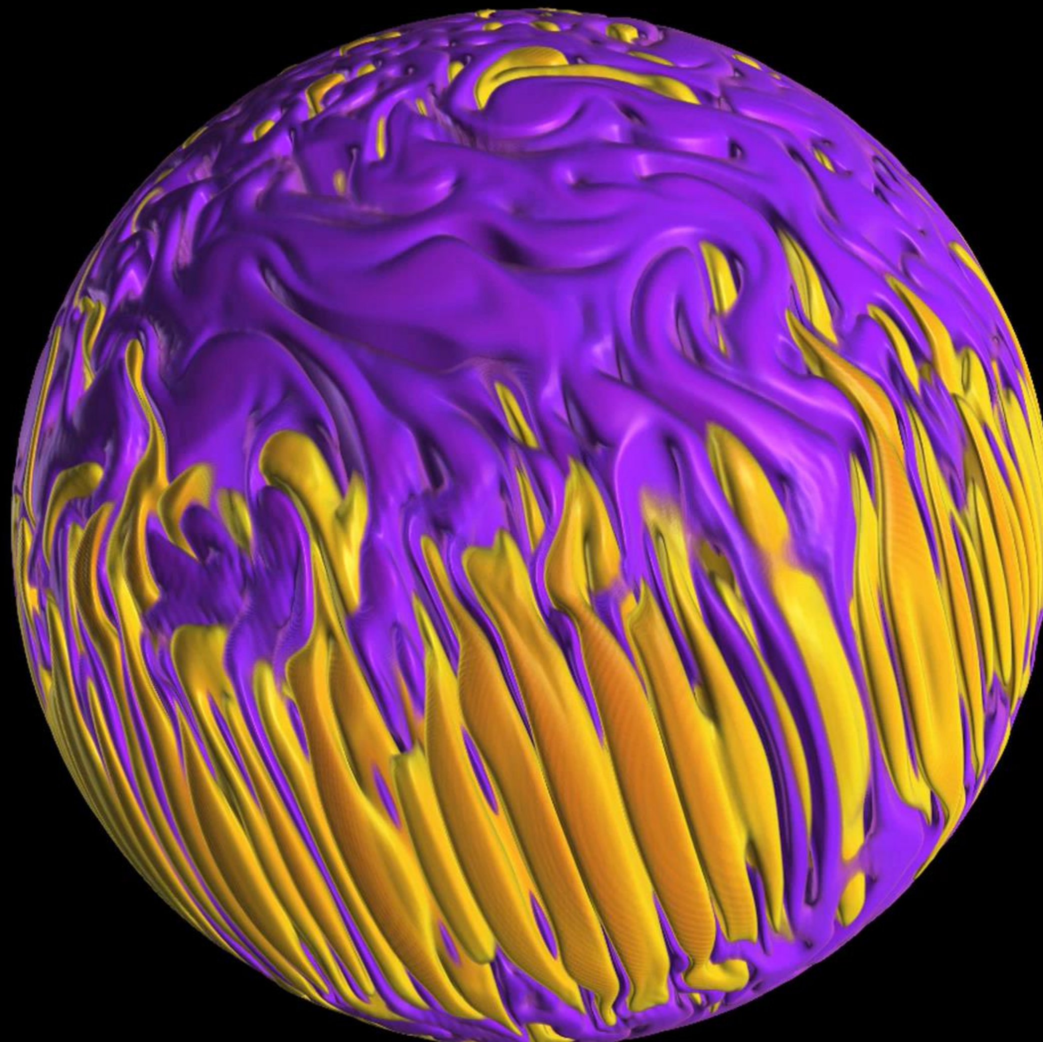
Earth-like geometry

Pr= 1
Ra = 10^7

4,000 cores
4 hours
NASA Pleiades



Rotating
(Exterior)

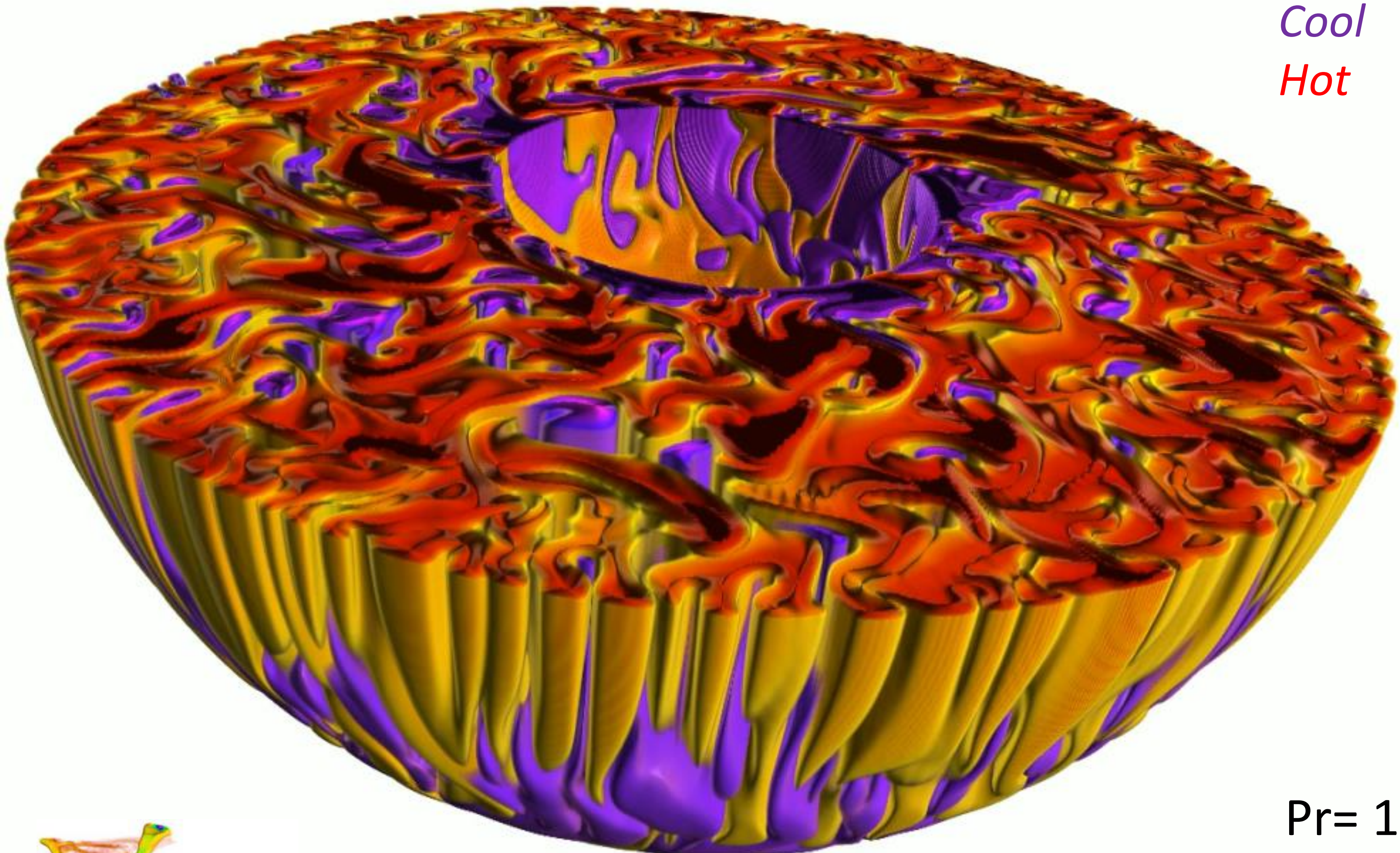


Rayleigh Code

4,000 cores
6 hours walltime
NASA Pleiades

Pr= 1
Ek = 10^{-5}
Ra = $2.5\text{x}10^{8}$

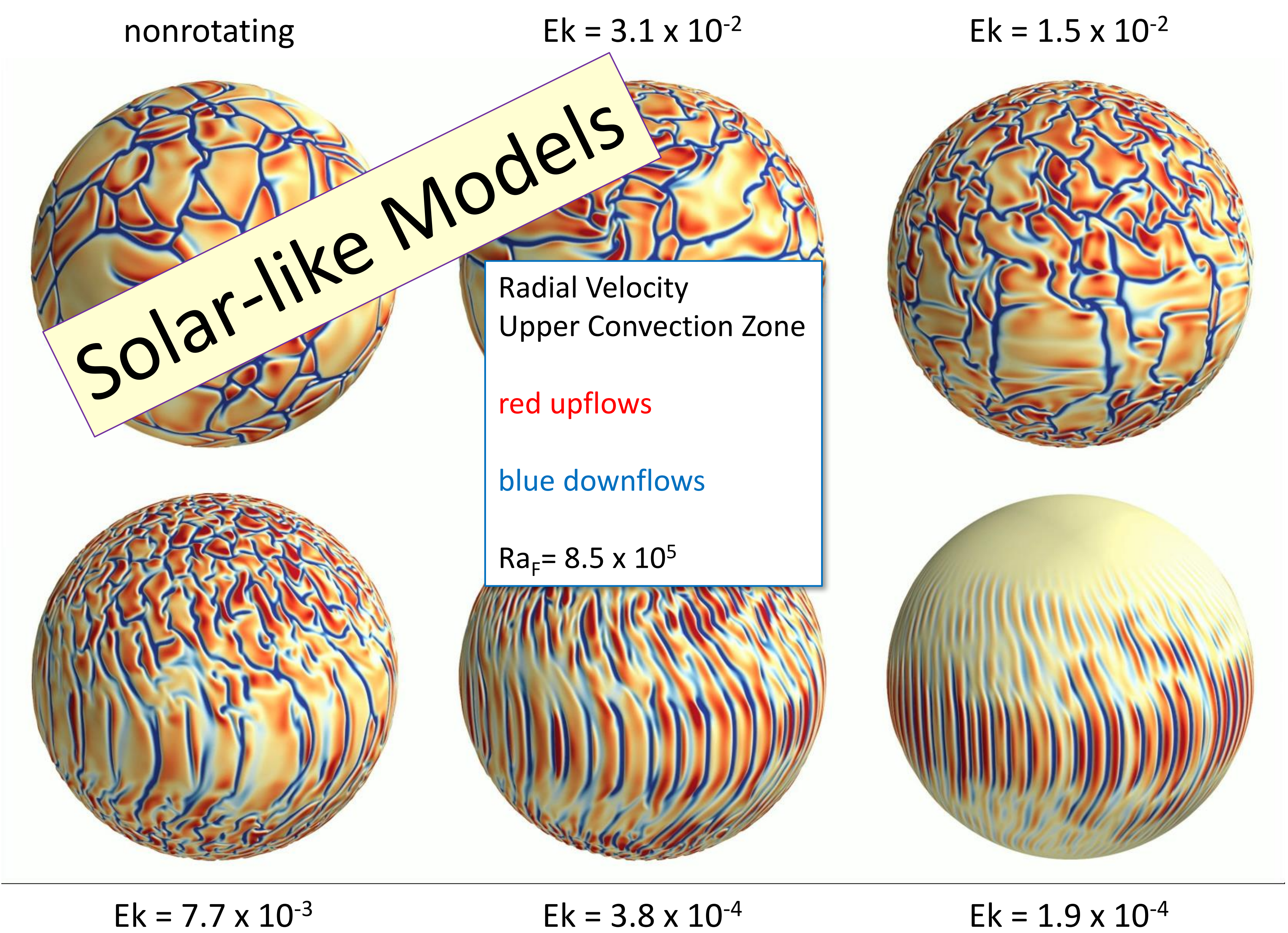
Cool
Hot



Pr= 1
Ek = 10^{-5}
Ra = 2.5×10^{8}

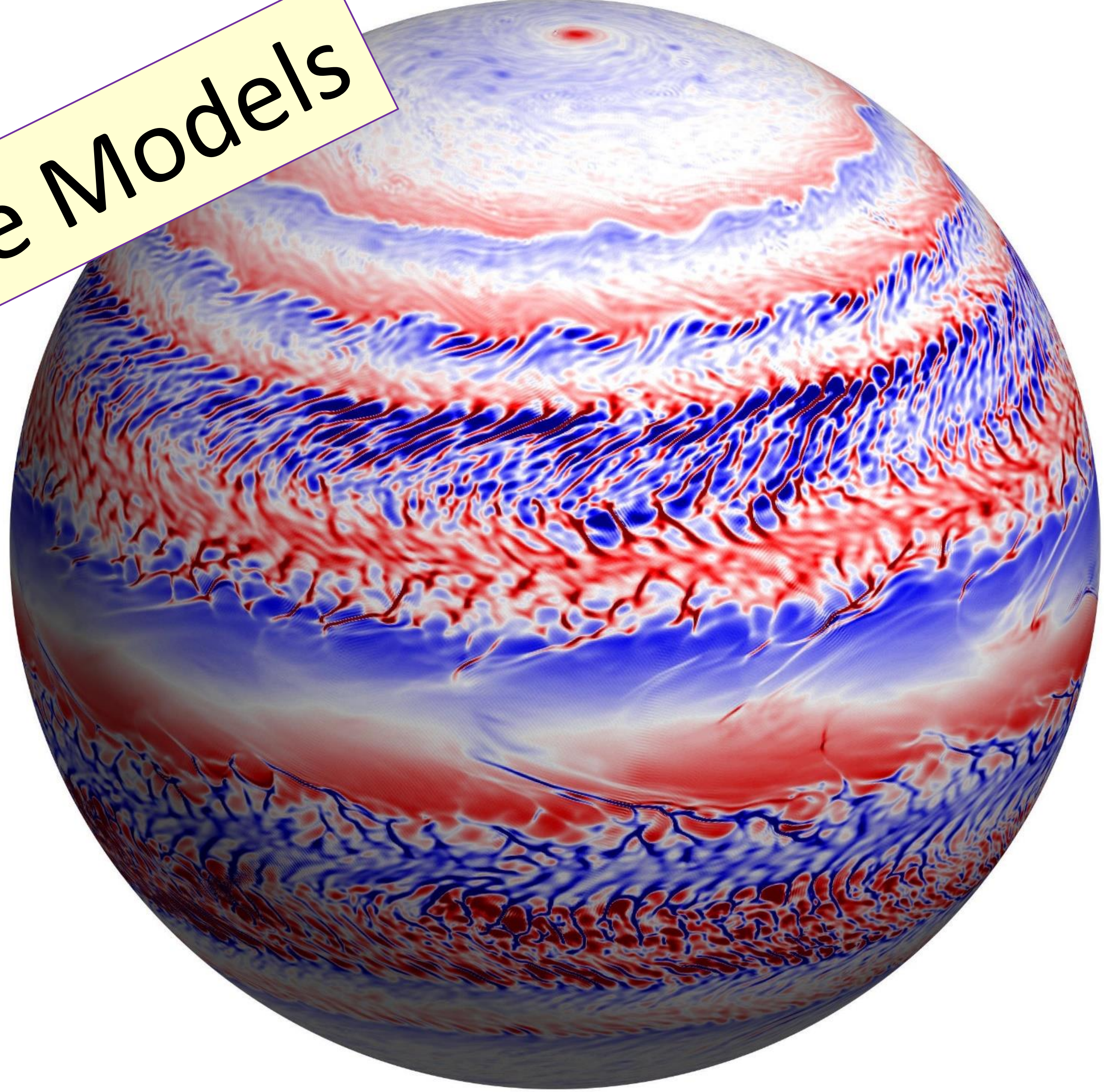
VAPOR
Visualization & Analysis Platform
NCAR

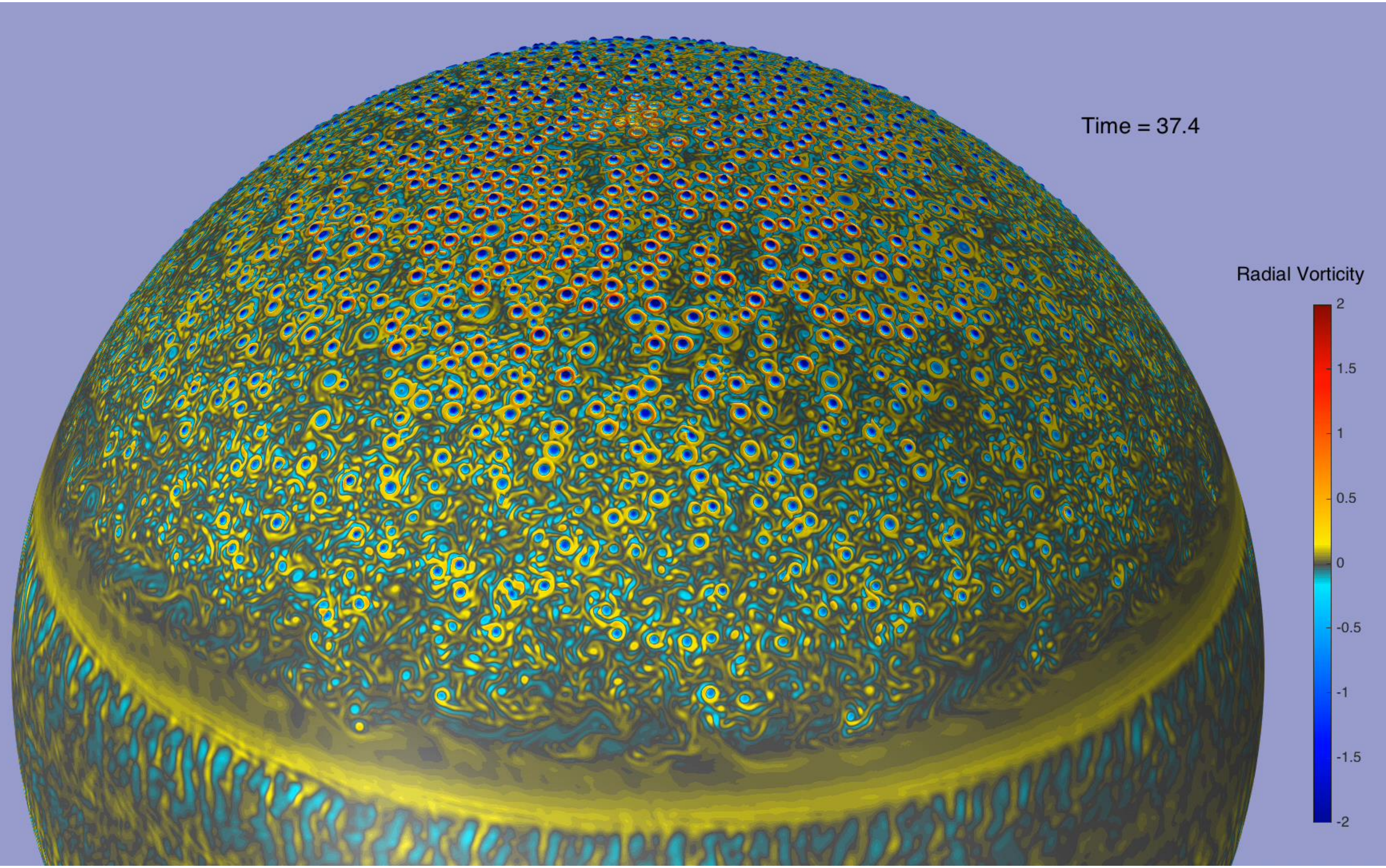
4,000 cores
6 hours walltime
NASA Pleiades



Solar-like Models
nonrotating
Ek = 3.1 x 10^-2
Ek = 1.5 x 10^-2
Radial Velocity
Upper Convection Zone
red upflows
blue downflows
Ra_F= 8.5 x 10^5
Ek = 7.7 x 10^-3
Ek = 3.8 x 10^-4
Ek = 1.9 x 10^-4

Jupiter-like Models




Time = 37.4
Radial Vorticity
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2

The Need for Community Effort

"BIG" calculations aren't cheap….

Resources for 1 million time steps

problem size	N_ℓ	N_r	N_θ	N_ϕ	Core Hours (Million)
512^3	512	256	768	1536	0.78
1024^3	1024	512	1536	3072	7.65
2048^3	2048	1024	3072	6144	97.65[1]

Conservative disk space Estimate

Output Type	Output Size	Output Frequency	Simulation Total
Checkpoints	413 GB	10,000[1]	8,260 GB
3-D Snapshots	930 GB	50,000	18,600 GB
2-D Slices	21.1 GB	1000	21,100 GB
Diagnostics	3.6 GB	1000	3,600 GB
Total	–	–	51,560 GB

2048^3