Turbulence on Petascale Computers:

What have we learned, and
What we hope to learn

P.K. Yeung

Schools of AE, CSE and ME
Georgia Inst of Tech, USA

pk.yeung@ae.gatech.edu

KITP Turbulence Program, UC Santa Barbara

May 26, 2011

Ack.: NSF and NSF/DoE Supercomputing Centers

Outline of This Talk

- **Turbulence and Petascale Computers:**
  - *some general remarks*
- **What have we learned (examples):**
  - *intermittency, mixing, dispersion*
- **What we hope to learn (challenges):**
  - *in both science and computing*

Turbulence and Computing

- Turbulence: disorderly fluctuations over a wide range of scales in time and 3D space, with diverse applications
  - efficient mixing (of heat, substances and momentum) is key to combustion, aerodynamic drag, pollutant dispersion, etc.
- Direct numerical simulation: compute all the scales, based on exact governing equations
  - for physical understanding and model development
  - CPU intensive (repeat: wide range of scales)
- Petascale: 10^{15} operations/sec, or bytes of data:
  - exponential increase in CPU power over at least 25 years, world's fastest currently at 2.4 Pflop/s (theoretical peak)

A brief history of DNS

— (selected major markers) —

- Orszag 1969-1971: Spectral and pseudo-spectral methods
- Riley & Patterson 1972: particle tracking (32^3)
- (Large-eddy simulation: Leonard, 1974)
- Rogallo 1981: homogeneous turbulence (128^3)
- Kim, Moin & Moser 1987: channel flow (Chebyshev)
- Various authors: $\sim 512^3$, early to late 1990s
- Kaneda *et al.* 2002: 4096^3 on Earth Simulator, Japan

A brief history of DNS

— (selected major markers) —

- Orszag 1969-1971: Spectral and pseudo-spectral methods
- Riley & Patterson 1972: particle tracking (32^3)
- (Large-eddy simulation: Leonard, 1974)
- Rogallo 1981: homogeneous turbulence (128^3)
- Kim, Moin & Moser 1987: channel flow (Chebyshev)
- Various authors: $\sim 512^3$, early to late 1990s
- Kaneda *et al.* 2002: 4096^3 on Earth Simulator, Japan

Future: Turbulence at 12288^3, RK4, 10000 time steps in 40 hours is an acceptance test criterion for 10-Pflop *Blue Waters*, 2012

Uses of Massive Computing Power

- A wider range of scales (in space and/or time)
  - higher Reynolds number (always!)
  - high Schmidt number ($Sc = \nu/D$): smaller scales
  - *very* low Schmidt number: growh of large scales
- Improved accuracy at the small scales
  - fine-scale intermittency, thin reaction zones
- Longer simulations, e.g. to provide better sampling
  - amount of data IS a challenge
- More complex physics
  - e.g. stratification, rotation, MHD
- More complex boundary conditions
  - channel, boundary layer, mixing layer (still canonical)

More thoughts about Computers

- Good access to a top-of-the-line machine would let us:
  - compute faster, bigger, longer; analyze deeper
  - compute better too? (hopefully)
- But to get the best benefit is not trivial
  - massive parallelism (up to $O(10^5)$ CPU cores)
  - Cyber: how to use/re-use, maintain, and share data
  - intense competition for CPU resources vs. other fields
  - new programming models to be investigated
- Good science gets done only if:
  - good questions are being asked (needs collaborators)
  - humans and computers working together well

Simulation Approach

- Forced, stationary isotropic turbulence on a periodic domain, using Fourier pseudo-spectral method (Rogallo 1981)
- Resolution: in most simulations pushing the Reynolds number $k_{max}\eta \approx 1.5$ ($\Delta x/\eta \approx 2$, with $k_{max} \equiv \sqrt{2}N/3$)
  - effects on intermittency examined in Donzis *et al.* PoF 2008
- Passive scalar fluctuations driven by a uniform mean gradient: ($\nabla\Phi = (1, 0, 0)$: allows tests of local isotropy)

$$\partial\phi/\partial t + \mathbf{u}\cdot\nabla\phi = -\mathbf{u}\cdot\nabla\Phi + D_\phi\nabla^2\phi$$

- Size of smallest scale for each scalar depends on Sc — unequal accuracy for multiple scalars in a given simulation
- Massively parallel code, in principle up to N^2 processors (Donzis, Yeung & Pekurovsky; TeraGrid Conf. 2008)

Simulation database

R_λ	N	$k_{max}\eta$	Sc					
140	256	1.38	0.125	1				
140	512	2.74	0.125	1	4			
140	1024	5.48		1	4			
140	2048	11.2			4			64
240	512	1.41	0.125	1				
240	2048	5.14		1		8		
240	4096	~ 11					32	
390	1024	1.4	0.125	1				
650	2048	1.4	0.125	1				
650	4096	2.8		1	4			
1000	4096	1.4						

(Also recent runs on larger domains, and very low Sc)

What have we learned:

1. Intermittency and extreme events

Dissipation and Enstrophy

- Dissipation: $\epsilon = 2\nu s_{ij} s_{ij}$ (strain rates squared)
- Enstrophy: $\Omega = (\nu)\omega_i \omega_i$ (rotation rates squared)
- Same mean values in homogeneous turbulence, but moments and PDFs can be different
- Both represent small scales, but most data sources suggest enstrophy is more intermittent, contrary to expectation at high Reynolds no. (Nelkin 1999)
- In relative dispersion, straining pulls particle pairs apart but rotation makes them move around together
- Difficulties in resolution and sampling, nature of infrequent but extreme events

PDFs of Dissipation and Enstrophy

- Stretched-exponential fits:

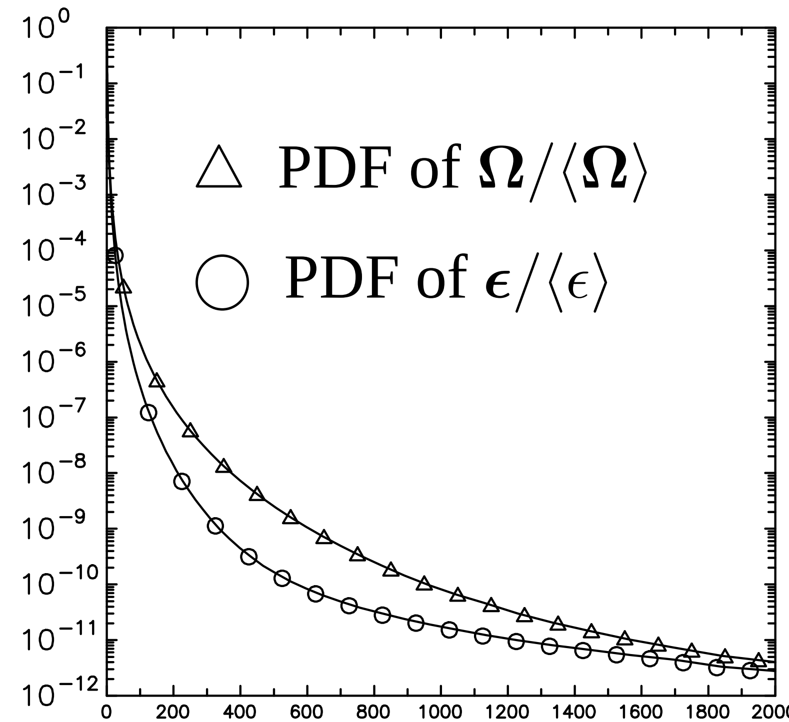
$$f_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}') \sim \exp[-b_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}')^{c_{\boldsymbol{\epsilon}}}]$$

- Donzis *et al.* PoF 2008:

  PDFs of $\boldsymbol{\epsilon}/\langle\epsilon\rangle$ and $\boldsymbol{\Omega}/\langle\boldsymbol{\Omega}\rangle$ coincide at extreme tails (only at high Reynolds no.)

- Similar results observed in two 4096^3 simulations:
  - higher Reynolds no.
  - higher resolution

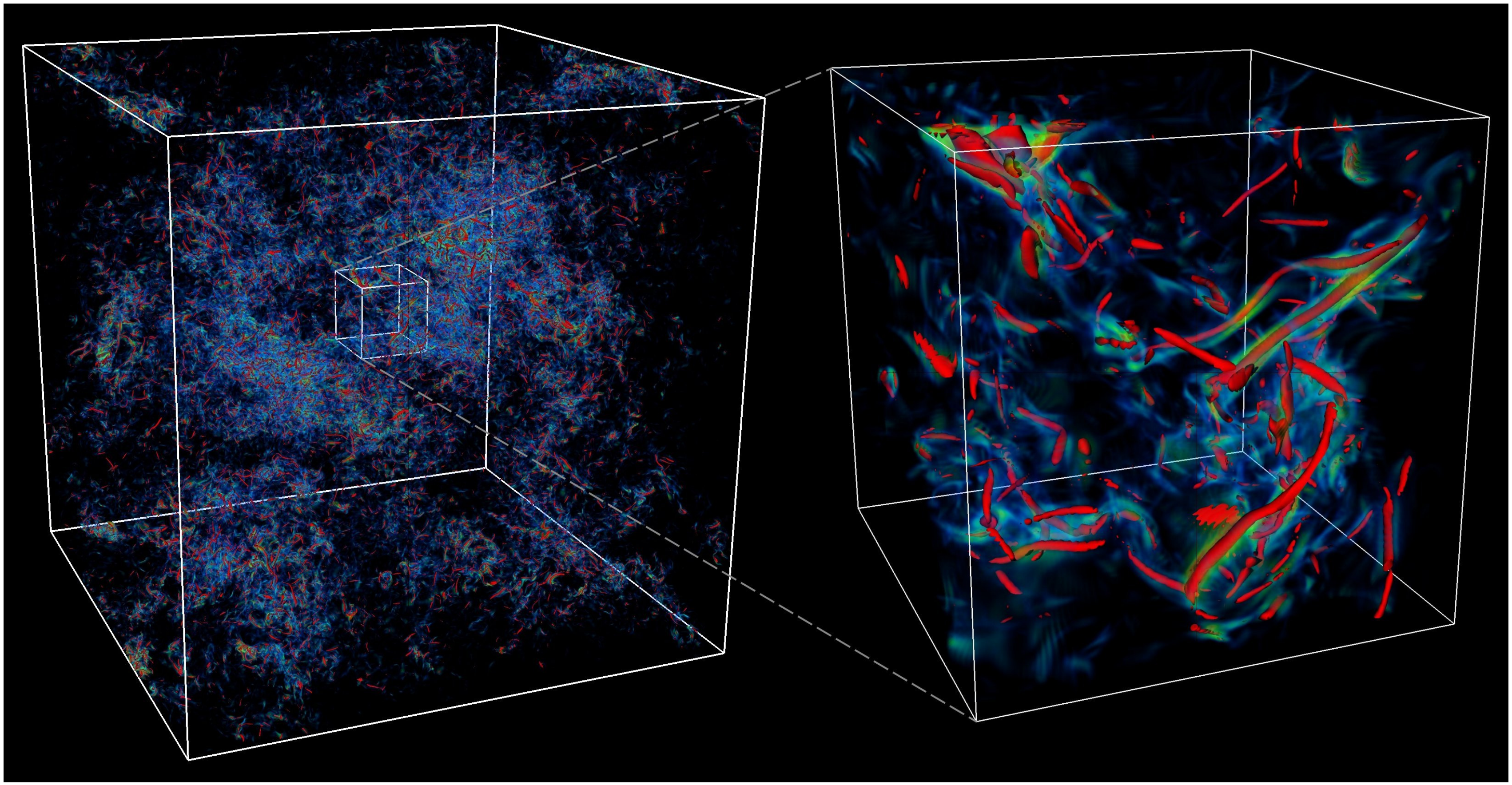
$R_\lambda \approx 1000,\ 4096^3$



$\boldsymbol{\epsilon}' \equiv \boldsymbol{\epsilon}/\langle\epsilon\rangle;\ \boldsymbol{\Omega}' \equiv \boldsymbol{\Omega}/\langle\boldsymbol{\Omega}\rangle$

Extreme $\boldsymbol{\epsilon}$ usually accompanied by large $\boldsymbol{\Omega}$, but extreme $\boldsymbol{\Omega}$ may occur with moderate $\boldsymbol{\epsilon}$

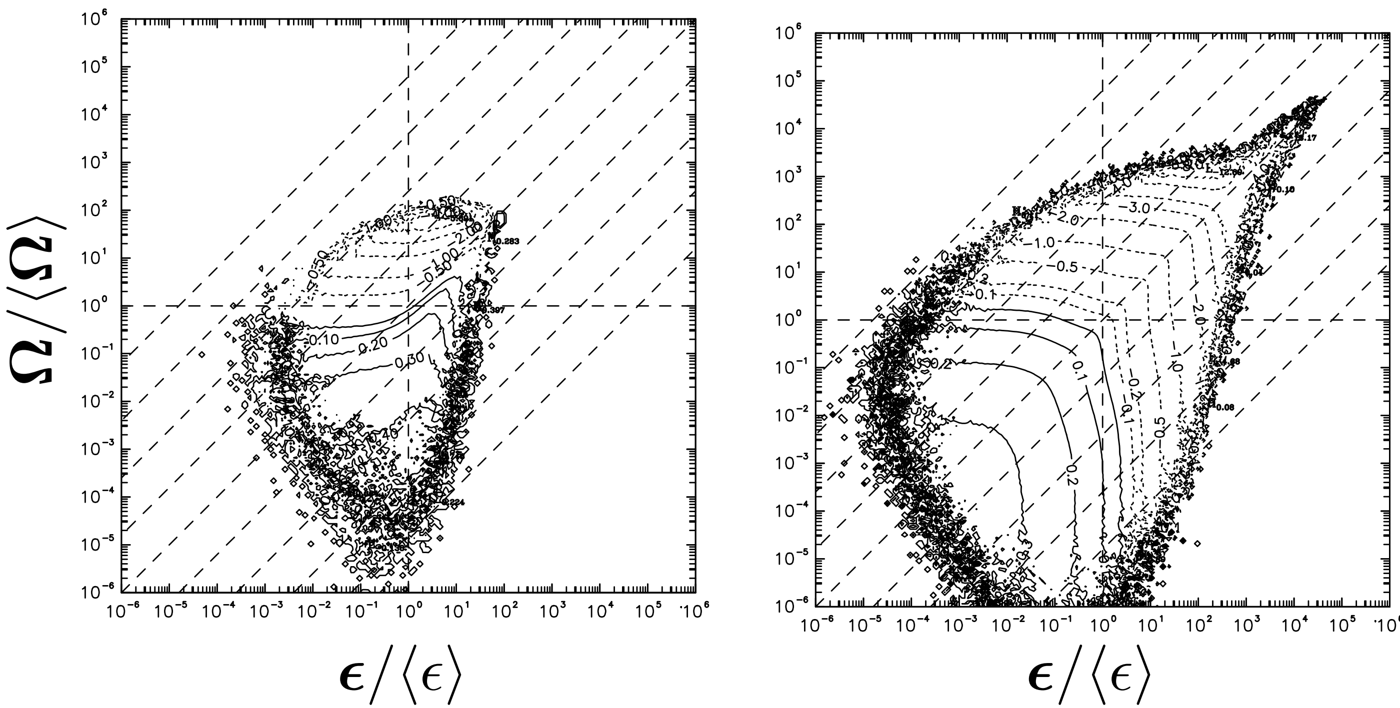
3D Visualization



[TACC visualization staff] 2048^3, $\boldsymbol{R_\lambda} \approx 650$: intense enstrophy (red) has worm-like structure, while dissipation (blue) is more diffuse

Flow variables conditioned on ϵ and Ω

From $\nabla^2(p/\rho) = \frac{1}{2}(\Omega - \epsilon/\nu)$, an indirect connection to pressure field?



At high Re, $\langle p|\epsilon, \mathbf{\Omega}\rangle$ is almost symmetric across the diagonal line
— both high ϵ and high $\mathbf{\Omega}$ lead to negative pressure fluctuations
— but nonlocal nature of pressure adds some complexity

What have we learned:

2. Turbulent Mixing (Passive Scalars)

Turbulent Mixing: Similarity theory

- (Besides Re) the Schmidt number is also an important parameter
- Sc varies: $O(0.01)$ in liquid metals, $O(1)$ for gaseous combustion, ~ 7 for heat in water, $O(1000)$ for salinity in oceans
- Smallest scales thought to be

$$\begin{array}{rrcll} \text{Obukhov-Corrsin:} & \eta_{OC} & = & \eta Sc^{-3/4} & \text{for } Sc \lesssim 1 \\ \text{Batchelor:} & \eta_B & = & \eta Sc^{-1/2} & \text{for } Sc \gg 1 \end{array}$$

- Different scaling regimes for $Sc \lesssim 1$, $\gg 1$ and $\ll 1$, but data less available in latter two
- Local isotropy: do the small scales remain isotropic in response to a mean gradient?
- Intermittency of scalar gradients and scalar dissipation: what is the effect of the Schmidt number?

Sc $\lesssim$ 1: Obukhov-Corrsin scaling

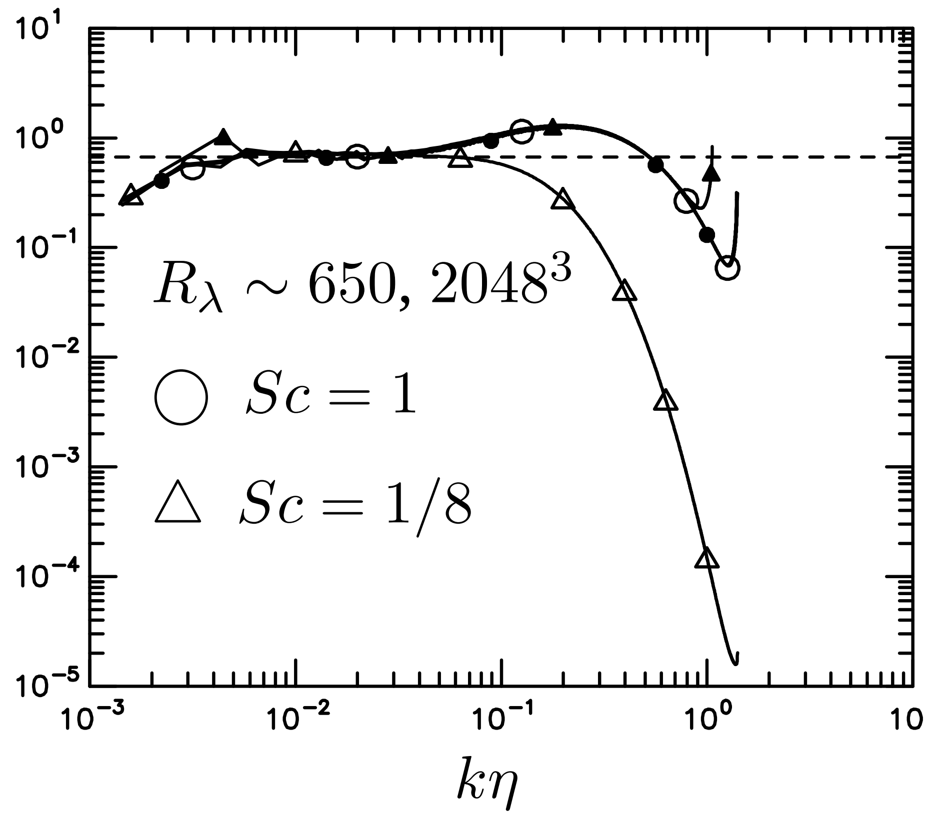
- Inertial-convective:

$$E_\phi(k) \sim \langle\chi\rangle\langle\epsilon\rangle^{-1/3}k^{-5/3}$$

(for $1/L \ll k \ll 1/\eta_{OC}$)

- Yeung *et al.* PoF 2005:
  - $C_{OC} \approx 0.67$ in 3D spectrum, consistent with survey of experiments (Sreenivasan PoF 1996)
  - bottleneck apparent for $Sc = 1$ (precursor to k^{-1} for $Sc > 1$?)

Compensated spectra

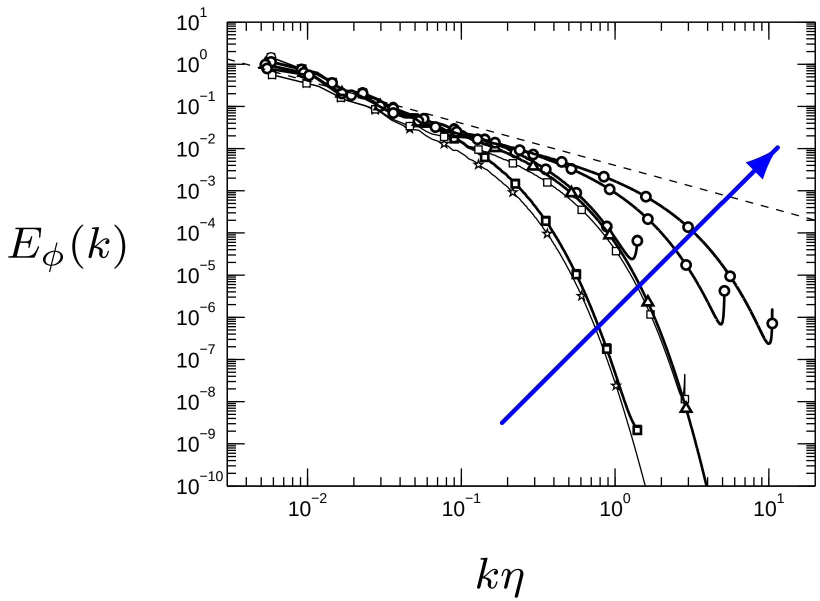


Consistent with isotropic random forcing of scalars (Watanabe & Gotoh 2004, 2007; ▲, •)

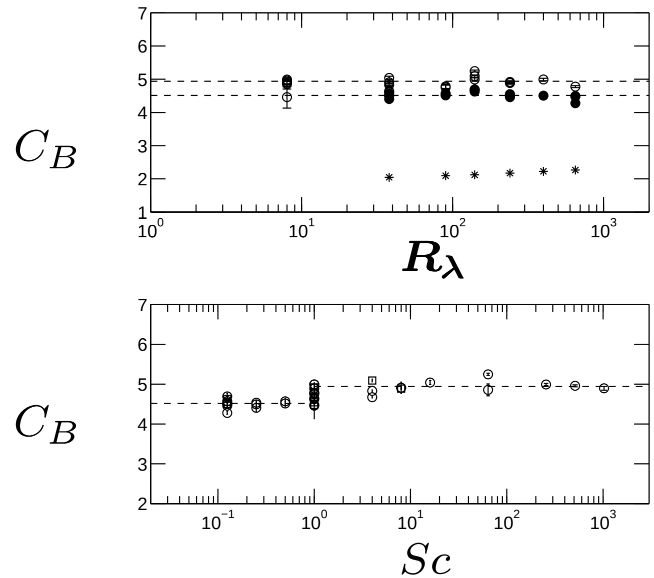
$Sc \gg 1$: Batchelor's spectrum

$\sim$ Donzis, Sreenivasan & Yeung (FTC 2010)

$E_\phi(k) \sim C_B \langle \chi \rangle \tau_\eta k^{-1}$



Value of C_B is less certain



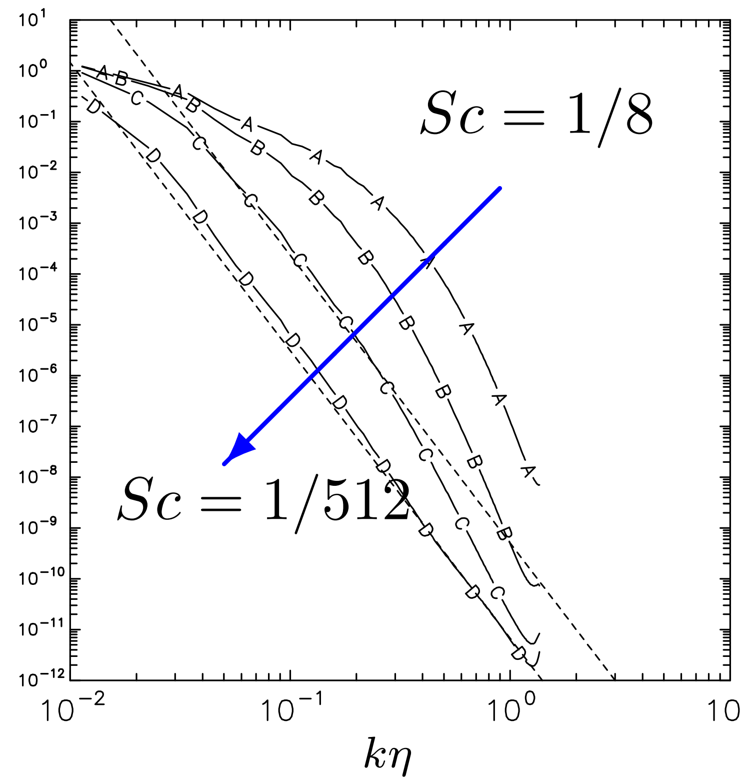
- Viscous-convective: $1/\eta \ll k \ll 1/\eta_B$
- R_λ 240, $Sc = 1/8, 1, 8, 32$
- Sustained trend towards k^{-1}

- Batchelor (1959): 2.0, but DNS close to 5
- PDF of most compressive principal strain rate

Sc ≪ 1: Batchelor *et al.* JFM 1959

"Inertial-diffusive": $E_\phi(k) \sim \langle\chi\rangle\langle\epsilon\rangle^{2/3} D^{-3} k^{-17/3}$

- Few data available: $Sc \ll 1$ in liquid metals and astrophysics
- Needs larger domain for larger length scales — while keeping Re high!
- Spectral cascade not the same, since velocity is now at "intermediate" scale
- Preliminary data: $Sc = 1/8$, 1/32, 1/128, 1/512

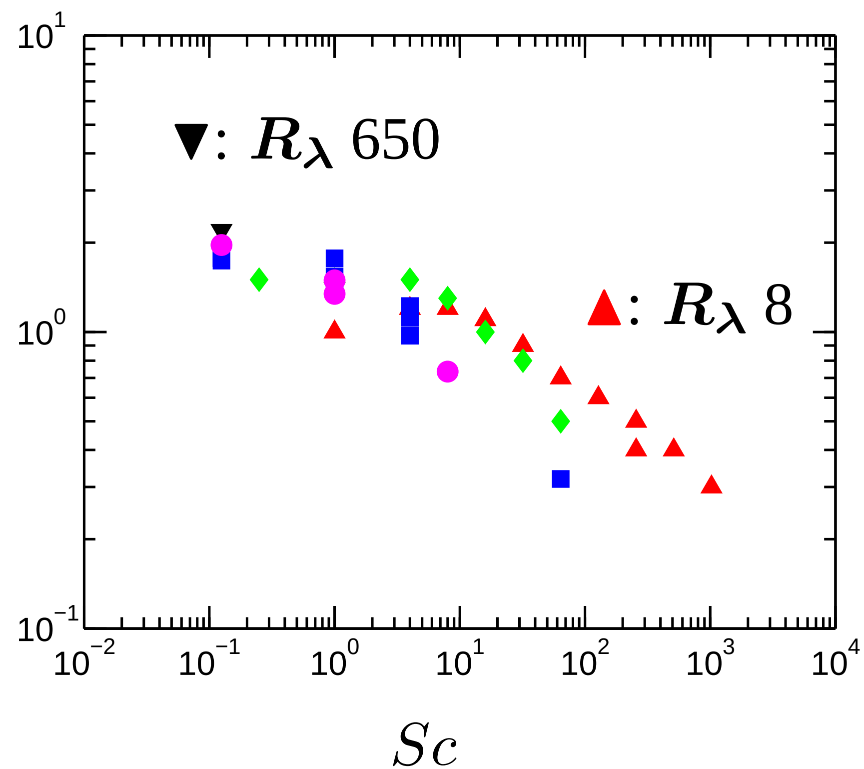


Tentative support for $k^{-17/3}$ (quality of data to be improved)

Local (An)isotropy

- Most lab. and DNS data indicate $\nabla_{\|}\phi$ is skewed, in conflict with notion of local isotropy at high Re
- Beyond $Sc \sim 4$, skewness drops with increasing Sc (faster if Re is higher)
- A return to isotropy at high Re may have been masked by finite resolution (Donzis & Yeung FTC 2010)

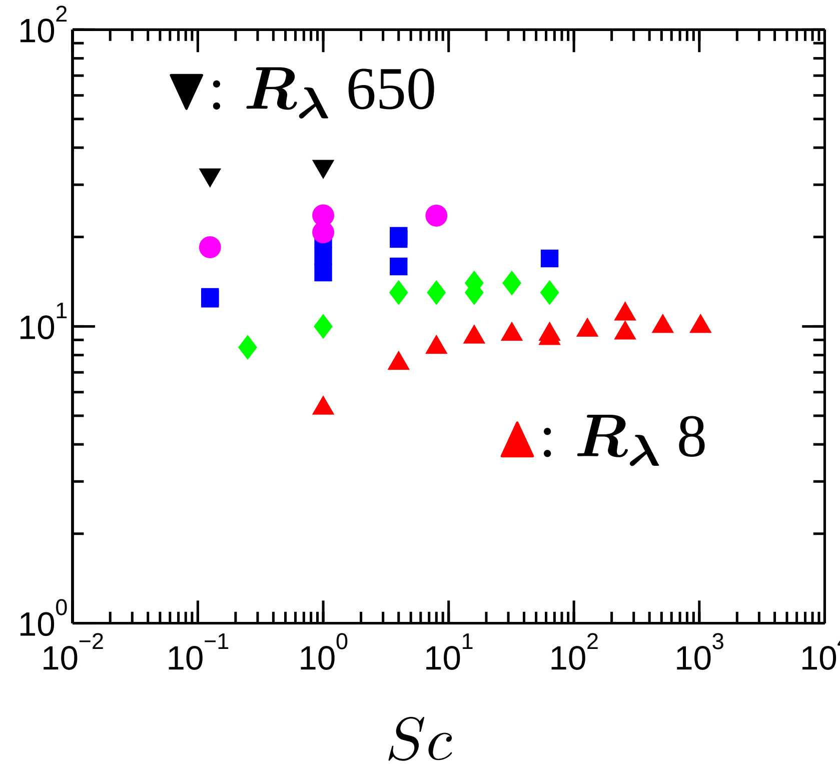
Gradient skewness at various $\boldsymbol{R_\lambda}$, Sc, and resolutions:



Intermittency of scalar gradients

- Scalar gradients are highly non-Gaussian, with $\nabla_{\|}\phi$ about 10% higher flatness than $\nabla_{\perp}\phi$
- Strong increase with $\boldsymbol{R_\lambda}$ at low Sc
- High Sc: a trend of saturation (but flatness for highest Sc in simulation may be underestimated)
- Sc needed for saturation is lower if Re is high

Flatness of $\nabla_{\|}\phi$



What have we learned:

3. Turbulent Dispersion (Lagrangian view)

Lagrangian Approach and DNS

- Motion of fluid particles is fundamental to turbulent dispersion and multiphase flows
  - G.I. Taylor 1921, Proc. Lond. Math. Soc.: "Diffusion by continuous movements"
  - L.F. Richardson 1926: particle pairs moving apart
  - multiparticle clusters also give useful info. on flow structure
- DNS is a powerful source of Lagrangian data
  - particle tracking algorithm based on cubic-spline interpolation (Yeung & Pope 1988): $d\mathbf{x}^+/dt = \mathbf{u}^+$; $\mathbf{u}^+(t) = \mathbf{u}(\mathbf{x}^+(t), t)$
  - velocity gradients sampled along particle trajectories
  - enormous detail under controlled conditions, for modeling

Lagrangian Kolmogorov Similarity

Lagrangian structure function:

$$D_2^L(\tau) \equiv \langle [u^+(t+\tau) - u^+(t)]^2 \rangle$$

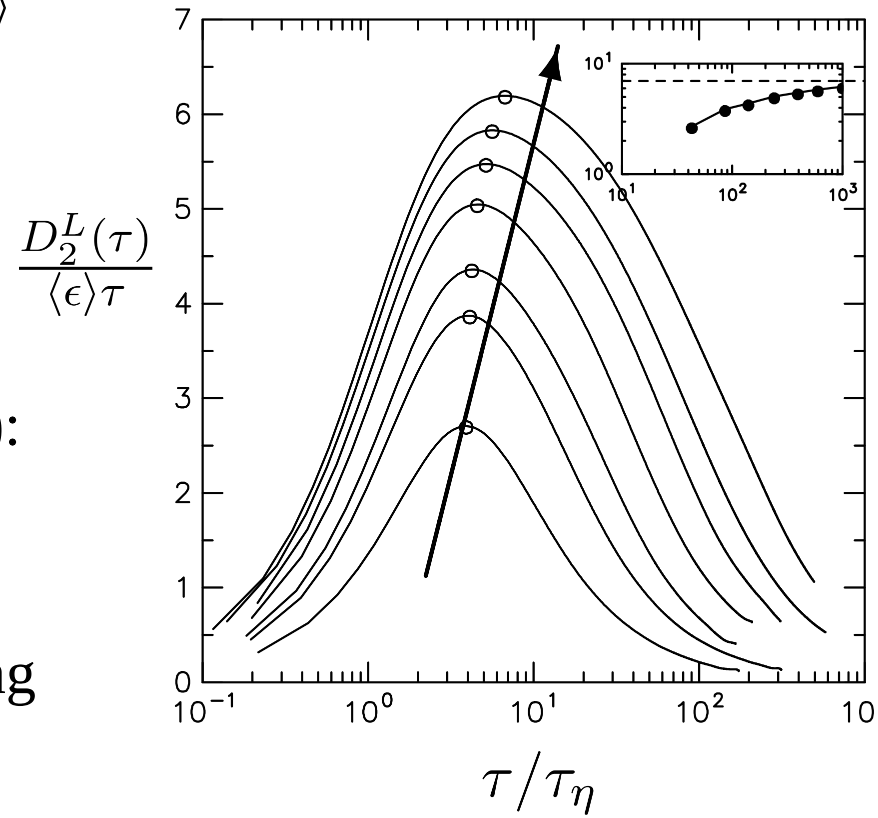
- Range of time scales from τ_η to integral time scale, $T_L \equiv \int_0^\infty \rho_L(\tau)\, d\tau$
- For $\tau_\eta \ll \tau \ll T_L$ ("inertial"):

$$D_2^L(\tau) = C_0 \epsilon \tau$$

  C_0 used in stochastic modeling

- R_λ up to ≈ 1000 in DNS (similar to expts) with $T_L/\tau_\eta \approx 80$: $C_0 \to O(7)$

DNS database, 64^3 to 4096^3

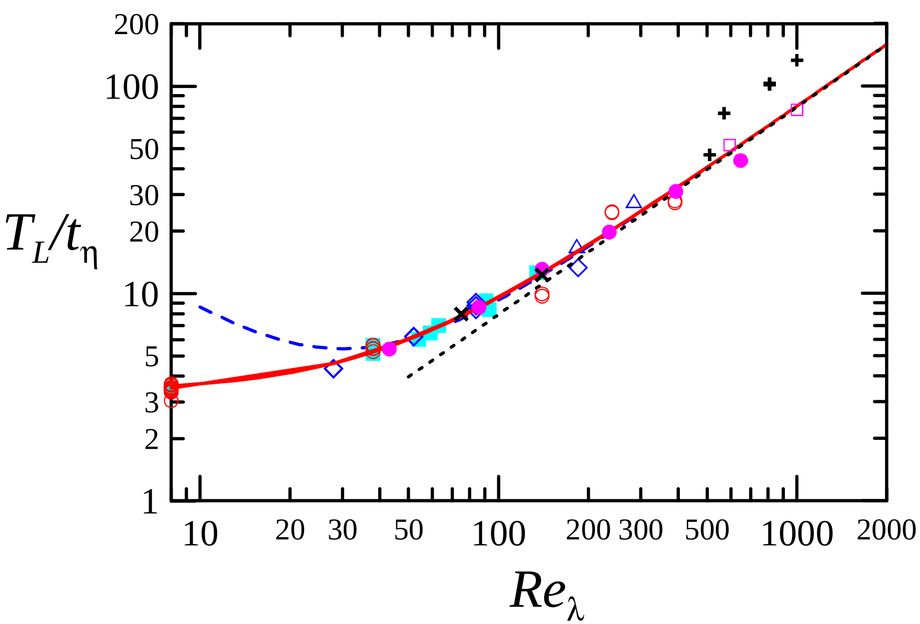


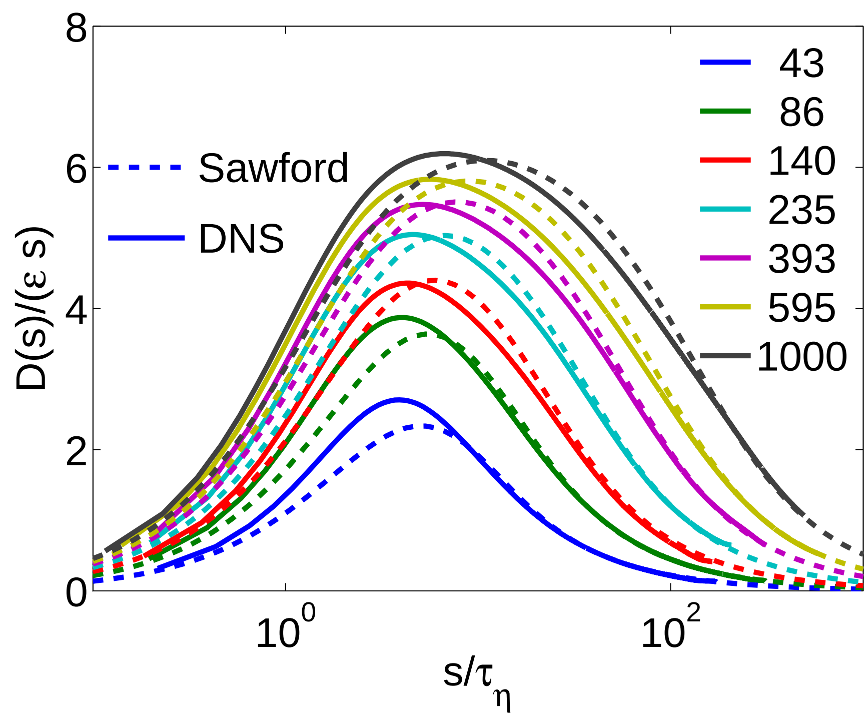
[Model of Sawford (1991) predicts $D_2^L(\tau)$ well, but not higher orders]

Dispersion Modeling

Stochastic modeling with drift and randomness terms: need T_L/τ_η as function of Reynolds number; and value of C_0

Sawford: multiple data sources



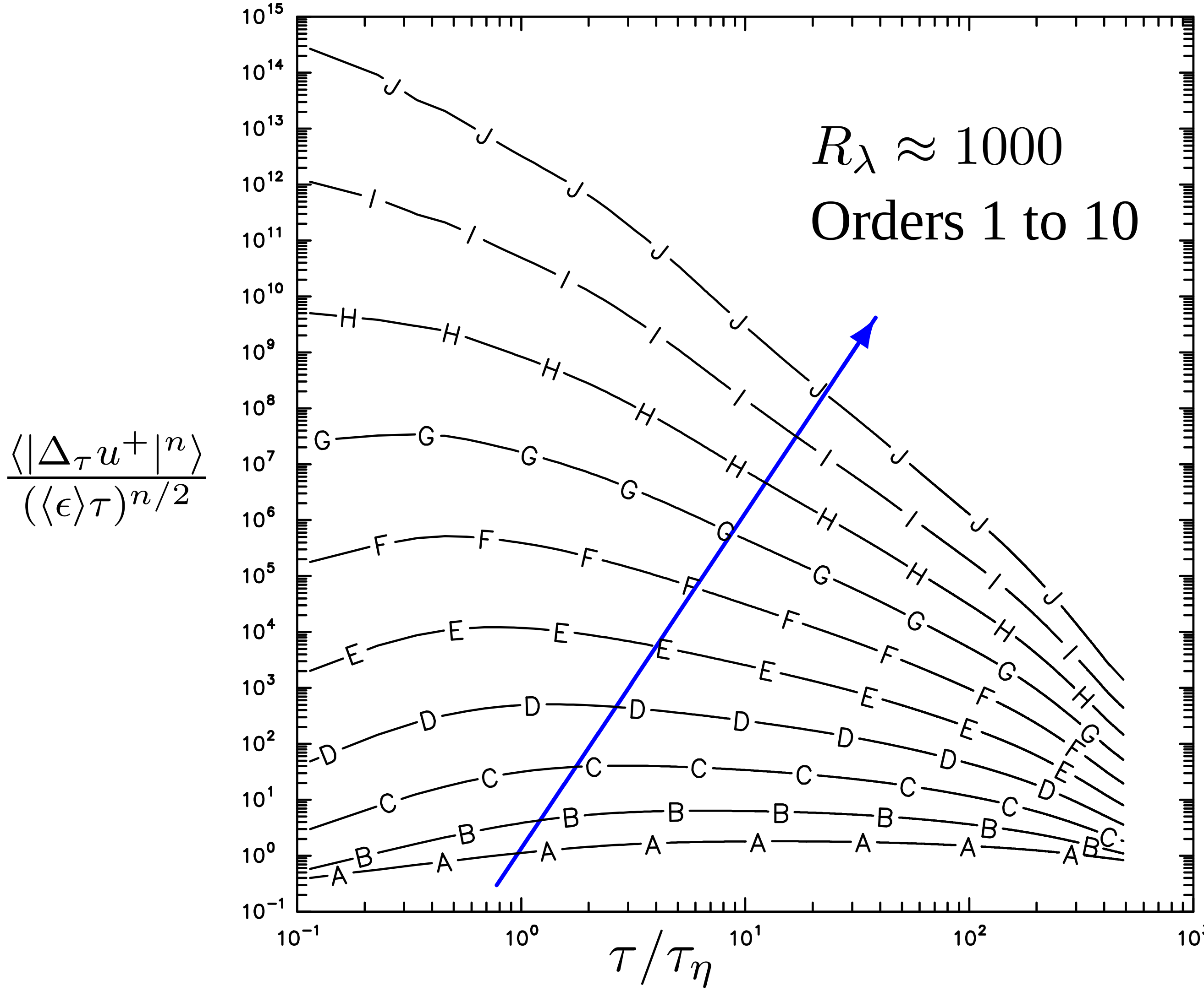


DNS at 4096^3, $\boldsymbol{R_\lambda} \approx 1000$:

$T_L/\tau_\eta \approx 80$

Pope 2009: based on acceleration model by Sawford (PoF 1991)

Higher-order structure functions



Much more difficult. Resolution issues due to intermittency (small τ)

Local flow structure

Fluid particles moving in regions of large fluctuating velocity gradients will experience a rapid change in velocity, i.e., a large acceleration

- Local straining, rotation, or combination of effects
  - dissipation: $\boldsymbol{\epsilon} \equiv 2\nu s_{ij} s_{ij}$
  - enstrophy: $\boldsymbol{\Omega} \equiv \omega_i \omega_i$
  - pseudo-dissipation: $\boldsymbol{\varphi} \equiv \nu(\partial u_i/\partial u_j)(\partial u_i/\partial u_j)$
- Strain-dominated vs rotation-dominated regions
  - strain is very important in dispersion of particle pairs
  - rotation can cause frequent changes in direction
- Need to know statistics and time scales of ϵ, ζ and $\boldsymbol{\varphi}$, along fluid particle trajectories, as function of Reynolds no.

Lagrangian conditional statistics

- Conditional sampling based on $\boldsymbol{\epsilon}$, $\boldsymbol{\Omega}$, or $\boldsymbol{\varphi}$ along particle paths, e.g.:

$$\rho_u(\tau|Z) \equiv \frac{\langle u^+(t)u^+(t+\tau)|Z^+(t)=Z\rangle}{\langle\{u^+(t)\}^2|Z^+(t)=Z\rangle}$$

  with $Z = \boldsymbol{\epsilon}$, $\boldsymbol{\Omega}$ or $\boldsymbol{\varphi}$ in logarithmic intervals

  - Lagrangian time series of $\boldsymbol{\epsilon}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\varphi}$ can be obtained by high-order interpolation in DNS
  - larger acceleration and more rapid-decorrelation expected in regions of large velocity gradients
  - dependence expected to last for time lags comparable to integral time scales of $\boldsymbol{\epsilon}^+(t)$, $\boldsymbol{\Omega}^+(t)$, $\boldsymbol{\varphi}^+(t)$

- A promising tool for introducing effects of fine-scale intermittency into stochastic modeling (Lamogese *et al.* JFM 2007)

Conditional structure functions

- Stochastic modeling at different levels of complexity:
  - Given $\mathbf{u}^+(t)$, "predict" increment $\Delta_\tau \mathbf{u}^+ = \mathbf{u}^+(t+\tau) - \mathbf{u}^+(t)$ (then integrate to recover displacement)
  - Given $\mathbf{u}^+(t)$ and $\mathbf{a}^+(t)$, "predict" $\Delta_\tau \mathbf{a}^+ = \mathbf{a}^+(t+\tau) - \mathbf{a}^+(t)$ (then integrate to recover velocity, then displacement)
  - Given $\mathbf{u}^+(t)$ and $\epsilon^+(t)$: incorporate fine-scale intermittency
- Dissipation (strain), enstrophy (vorticity), or pseudo-dissipation (all velocity gradients)
- Conditional flatness factor: (with $X = \epsilon$, $\mathbf{\Omega}$ or $\boldsymbol{\varphi}$)

$$\mu_4(\tau|X) = \langle(\Delta_\tau u^+)^4|X\rangle / \langle(\Delta_\tau u^+)^2|X\rangle^2$$

"Conditional Gaussianity" is closest approximation for acceleration given pseudo-dissipation (Yeung *et al.* PoF 2006)

Local slopes and extended self-similarity

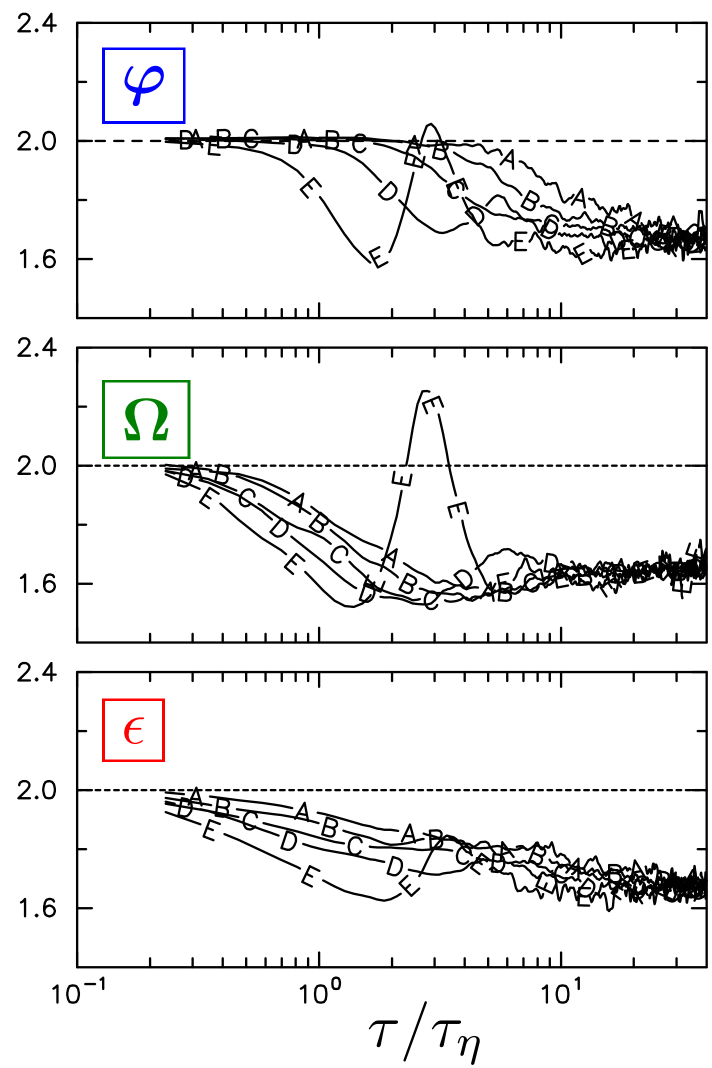
Let $D_m^L(\tau) \propto \tau^{\zeta_m}$

Lines A-E: increasing ϵ, Ω or φ

- K41 gives $\zeta_m = m/2$, but affected by intermittency
- ESS: Consider "local slope"

$$\zeta_4'(\tau) = \frac{d\log[D_4^L(\tau)]}{d\log[D_2^L(\tau)]}$$

- Biferale *et al.* PoF 2008: suggests "dip" in $\zeta_4'(\tau)$ for $\tau/\tau_\eta \approx 2$ due to vortices
- Can intense strain rate cause a similar observation?



Multiparticle clusters

- Motions of 2,3,4 particles considered together can be related to 2nd, 3rd and 4th moments of concentration fluctuations — triangles and tetrads also carry info on shape distortion
- Richardson (1926) for mean-square separation:

$$\langle r^2 \rangle = g \langle \epsilon \rangle t$$

inertial range conditions and independent of initial separation

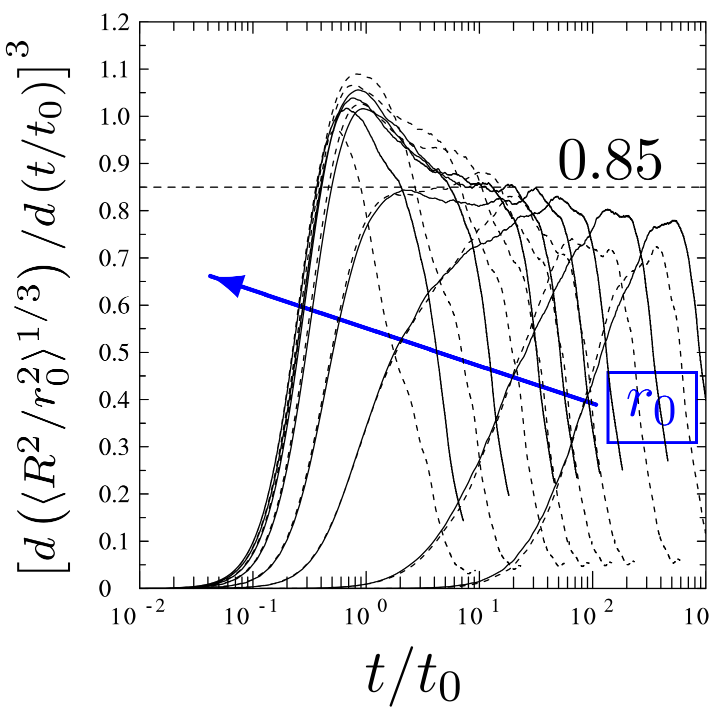
- Focus on tetrads (Pumir *et al.* 2000):
  - Size measured by volume or (better) gyration radius:
    $V = \frac{1}{6} \left| \left(\mathbf{X}^{(2)} - \mathbf{X}^{(1)} \right) \cdot \left[\left(\mathbf{X}^{(3)} - \mathbf{X}^{(1)} \right) \times \left(\mathbf{X}^{(4)} - \mathbf{X}^{(1)} \right) \right] \right|$
    $R^2 = \frac{1}{2n} \sum_{l,m=1}^{n} |\mathbf{X}^{(l)} - \mathbf{X}^{(m)}|^2 \quad (n = 4)$
  - Shape: $0 \leq \Lambda = V^{2/3}/R^2 \leq 3^{-5/3}$, or ratios of eigenvalues of a moment-of-inertia tensor

Tetrads: Evolution of Size and Shape

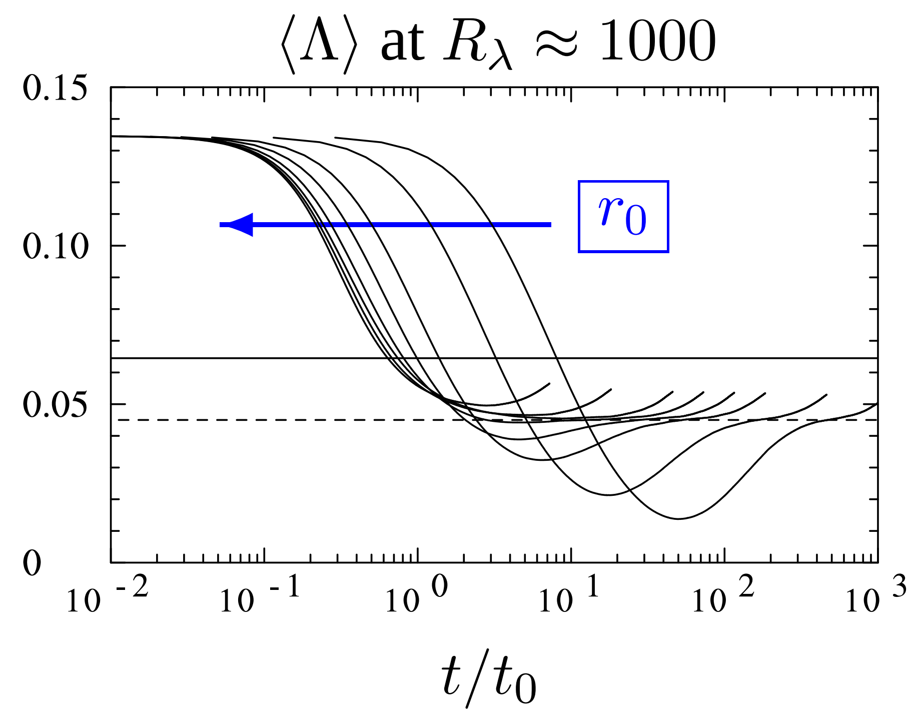
Data from $\boldsymbol{R_\lambda} \approx 1000$, 4096^3 and $\boldsymbol{R_\lambda} \approx 650$, 2048^3 runs used to test

$$\langle R^2 \rangle / r_0^2 = (3g/2)(t/t_0)^3$$

(where $t_0 = (r_0^2/\langle\epsilon\rangle)^{1/3}$)



Shape shows more scaling:



- inertial and diffusive regimes: 0.45 (DNS) and 0.645 (MC)
- more details in Hackl *et al.* PoF May 2011

What do we hope (and need) to learn:

A. Unresolved issues in turbulence

B. Cyber challenges and opportunities

Issues in Turbulence and DNS

Would like to better understand (e.g.) the following:

Issues in Turbulence and DNS

Would like to better understand (e.g.) the following:

- Connections among pressure, dissipation, and enstrophy
  - knowledge of ϵ and Ω determines $\nabla^2 p$, but not p itself

Issues in Turbulence and DNS

Would like to better understand (e.g.) the following:

- Connections among pressure, dissipation, and enstrophy
  - knowledge of ϵ and Ω determines $\nabla^2 p$, but not p itself
- A unified view of Sc-dependence in turbulent mixing
  - spectral transfer, local isotropy, intermittency
  - low or high Sc, at reasonably high Re

Issues in Turbulence and DNS

Would like to better understand (e.g.) the following:

- Connections among pressure, dissipation, and enstrophy
  - knowledge of ϵ and Ω determines $\nabla^2 p$, but not p itself
- A unified view of Sc-dependence in turbulent mixing
  - spectral transfer, local isotropy, intermittency
  - low or high Sc, at reasonably high Re
- Lagrangian intermittency in turbulent dispersion
  - refined similarity, info. for stochastic modeling

Issues in Turbulence and DNS

Would like to better understand (e.g.) the following:

- Connections among pressure, dissipation, and enstrophy
  - knowledge of ϵ and Ω determines $\nabla^2 p$, but not p itself
- A unified view of Sc-dependence in turbulent mixing
  - spectral transfer, local isotropy, intermittency
  - low or high Sc, at reasonably high Re
- Lagrangian intermittency in turbulent dispersion
  - refined similarity, info. for stochastic modeling
- DNS: effects of resolution, domain size, simulation time span

Issues in Turbulence and DNS

Would like to better understand (e.g.) the following:

- Connections among pressure, dissipation, and enstrophy
  - knowledge of ϵ and Ω determines $\nabla^2 p$, but not p itself
- A unified view of Sc-dependence in turbulent mixing
  - spectral transfer, local isotropy, intermittency
  - low or high Sc, at reasonably high Re
- Lagrangian intermittency in turbulent dispersion
  - refined similarity, info. for stochastic modeling
- DNS: effects of resolution, domain size, simulation time span
- Effects of more complex physics on everything above:
  - stratification, rotation, MHD turbulence

Cyber Challenges

How to keep going, and to the next level: (*Exascale* by 2018):

- How to scale our codes, efficiently, to $O(10^5)$ CPU cores? (currently, largest system in the world has $\approx 220,000$ cores)

Cyber Challenges

How to keep going, and to the next level: (*Exascale* by 2018):

- How to scale our codes, efficiently, to $O(10^5)$ CPU cores? (currently, largest system in the world has $\approx 220,000$ cores)
- New *programming models* and new types of architectures:
  - hybrid shared- and distributed memory
  - graphical processor units, etc.

  Can require massive effort, not necessarily rewarding!

Cyber Challenges

How to keep going, and to the next level: (*Exascale* by 2018):

- How to scale our codes, efficiently, to $O(10^5)$ CPU cores? (currently, largest system in the world has $\approx 220,000$ cores)
- New *programming models* and new types of architectures:
  - hybrid shared- and distributed memory
  - graphical processor units, etc.

  Can require massive effort, not necessarily rewarding!
- How to better compete with other fields for computer time?

Cyber Challenges

How to keep going, and to the next level: (*Exascale* by 2018):

- How to scale our codes, efficiently, to $O(10^5)$ CPU cores? (currently, largest system in the world has $\approx 220,000$ cores)
- New *programming models* and new types of architectures:
  - hybrid shared- and distributed memory
  - graphical processor units, etc.

  Can require massive effort, not necessarily rewarding!
- How to better compete with other fields for computer time?
- How to work with *Big Data* and engage the community?
  - can we be more collaborative, a little more like other fields with trusted community codes?

Cyber Challenges

How to keep going, and to the next level: (*Exascale* by 2018):

- How to scale our codes, efficiently, to $O(10^5)$ CPU cores? (currently, largest system in the world has $\approx 220,000$ cores)
- New *programming models* and new types of architectures:
  - hybrid shared- and distributed memory
  - graphical processor units, etc.

  Can require massive effort, not necessarily rewarding!
- How to better compete with other fields for computer time?
- How to work with *Big Data* and engage the community?
  - can we be more collaborative, a little more like other fields with trusted community codes?

Future Opportunities

- **Towards 8192^3 or (equivalent):**
  - NSF-funded *Blue Waters*, predicted 1 Pflop/s sustained execution speed, over 300,000 CPU cores
  - while 4096^3 becomes easier?
  - code development and choice of simulation parameters
  - analyses spanning years...

Future Opportunities

- **Towards 8192^3 or (equivalent):**
  - NSF-funded *Blue Waters*, predicted 1 Pflop/s sustained execution speed, over 300,000 CPU cores
  - while 4096^3 becomes easier?
  - code development and choice of simulation parameters
  - analyses spanning years...
- **What will we be doing in 2018?**