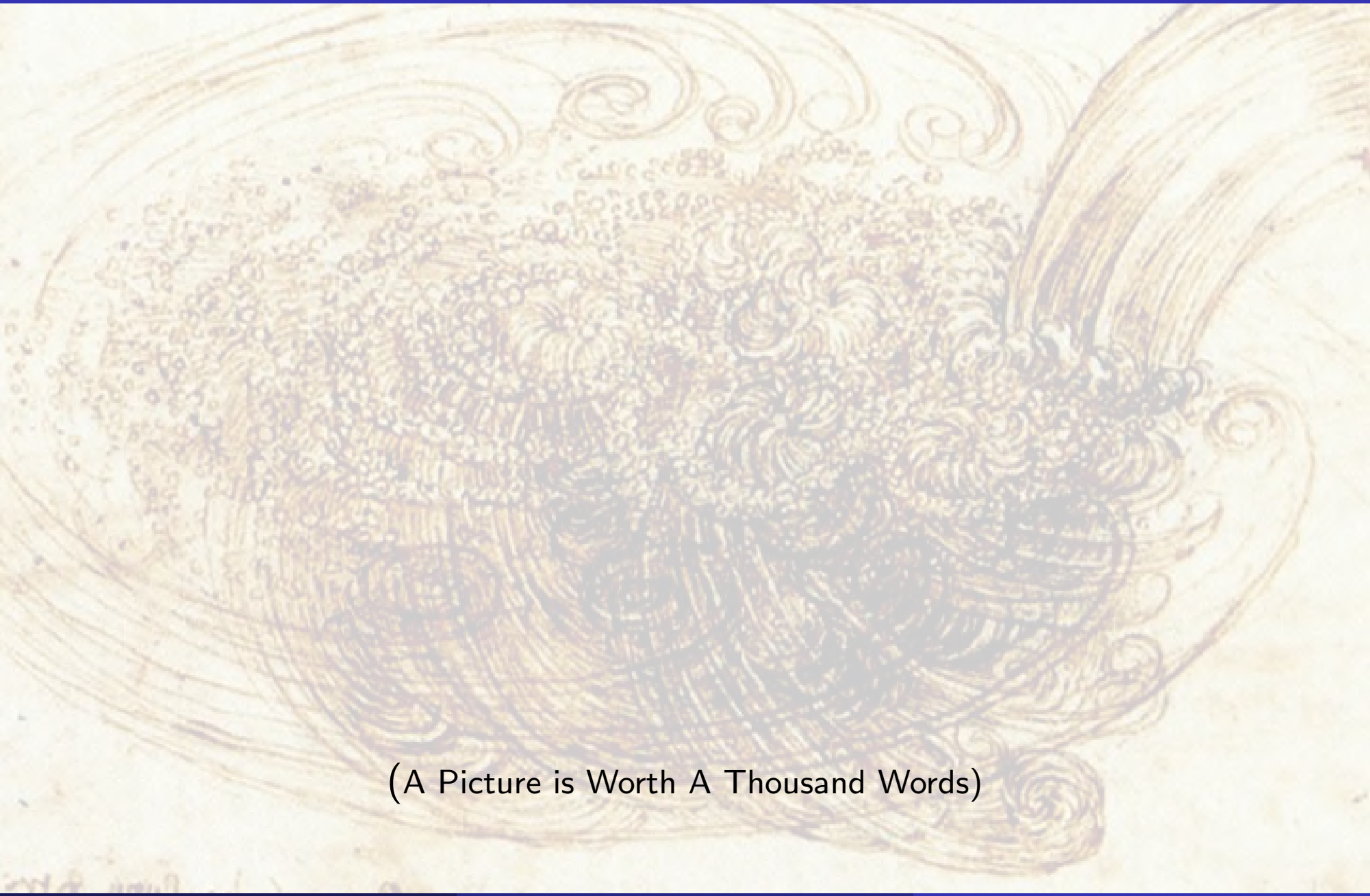
Resolution Effects and Local Averaging in Turbulence Simulations up to 4 Trillion Grid Points

P.K. Yeung (PI)
Schools of AE and ME, Georgia Tech
E-mail: pk.yeung@ae.gatech.edu

NSF:	PRAC (1036170, 1640771) and Fluid Dynamics Programs
BW Team, Cray:	Scaling, Reservations, Help Requests, Storage, Visualization
Collaborators:	T. Gotoh, S.B. Pope, B.L. Sawford, K.R. Sreenivasan
PhD Students:	K.P. Iyer (2014), D. Buaria (2016), M.P. Clay (2017), X.M. Zhai ($\sim$ 2018)

Blue Waters Symposium, June 4-7, 2018

Turbulence? Since the days of Leonardo da Vinci



(A Picture is Worth A Thousand Words)

Turbulence and High-Performance Computing

Disorderly fluctuations over a wide range of scales

- Pervasive in many branches of science and engineering
- Reynolds number: a measure of the range of scales
- Numerical simulation often best source for detailed information

A Grand Challenge problem in computing

- Flow is 3D: domain decomposition, and communication-intensive
- Every step-up in problem size: 8X in number of grid points

Some notable references in the field:

- Kaneda *et al.* PoF 2003: 4096^3, on Earth Simulator
- Yeung, Zhai & Sreenivasan PNAS 2015: 8192^3, on Blue Waters
- Ishihara *et al.* PRF 2016: 12288^3, on K Computer

Some Fundamental Questions

- Intermittency at high Reynolds number
  - How large/extreme fluctuations can be, and how likely?
  - How high a Reynolds number is high enough? For what purpose?
  - In both space and time
- Turbulent mixing and dispersion
  - How do things far apart come together? Or spread around?
  - Does molecular diffusivity matter, and to what extent?
- Turbulence in broader contexts, both natural and man-made:
  - Combustion: how does mixing interact with the chemistry?
  - The oceans: how are heat and salinity coupled to the flow?
  - etc., etc.

Some understanding of these questions is necessary in the design of improved engineering devices and responses to natural disasters, say.

What Blue Waters Has Enabled (Not Over Yet!)

Forced isotropic turbulence, R_λ up to 1300; various resolutions

- Largest production run at 8192^3, using 262,144 parallel processes
- Some shorter (yet arduous) runs at 12288^3 and 16384^3 (4 trillion)
- Hundreds of millions of core hours, 2.5 PB Nearline storage

Topics and Publications (to date):

- Extreme events (Y, Zhai & Sreenivasan PNAS 2015)
- Velocity increments and similarity (Iyer, S & Y, PRE 2015, 2017)
- Nested OpenMP for low-diffusivity mixing (Clay, *et al.* CPC 2017)
- Highly scalable particle tracking (Buaria & Y, CPC 2017)
- Resolution and extreme events (Y, S & Pope, PRF 2018)

The Question of Accuracy

Sources of errors and uncertainties in DNS:

- Resolution in space: requires $\Delta x/\eta \leq 1$?
- Resolution in time: is Courant number < 1 small enough?
- Aliasing errors due to nonlinear terms in Navier-Stokes equations
- Statistical: average over independent trials when possible

Can a simulation at higher resolution resolve any doubts?

- Higher moments, extreme fluctuations inherently more sensitive
- An expensive proposition, may have to be short
- Small scales most sensitive, but they have short time scales
  — hence shorter test runs may well suffice

Spatial Resolution in Pseudo-Spectral DNS

$$k_{max} = \sqrt{2}N/3; \qquad \Delta x/\eta \approx 2.96/(k_{max}\eta)$$

- $k_{max}\eta \approx 1.5$ may be fine for low-order velocity statistics
- But better resolution needed for small-scale statistics, especially at high Reynolds numbers (Yakhot & Sreenivasan 2005)
- Refine the grid, run again at same Re, and compare (e.g. acceleration statistics, Yeung *et al.* PoF 2006)
- For a given snapshot, what features may be less reliable?
  - Take best-resolved velocity field, truncate at some wavenumber k_c. Compute various statistics again, compare, and repeat
  - Large differences would indicate insufficient accuracy
  - A post-processing task, not a large new simulation — but, no information on global error after N-S time evolution

Temporal Resolution

- Courant number constraint for numerical stability:

$$C = \left[\frac{|u|\Delta t}{\Delta x} + \frac{|v|\Delta t}{\Delta y} + \frac{|w|\Delta t}{\Delta z} \right]_{max} \sim \alpha \frac{u' \Delta t}{\Delta x}$$

- Classical scaling and our experience in the simulations (take $\alpha = 12$)

$$\Delta t/\tau_\eta \approx (C/12)(15)^{1/4}(\Delta x/\eta)\ R_\lambda{}^{-1/2}$$

- $\Delta t/\tau_\eta$ often well under 1%: much less than $\Delta x/\eta$ (which is $O(1)$)
- But there are time scales smaller than τ_η in the simulations
  - η/u': advection of small scales by large scales (Tennekes 1975)
  - $\Delta x/u'$: advection over 1 grid spacing (re: Courant number)
- Short tests (approx 10 τ_η, say) at $C = 0.6$, 0.3, 0.15 can help.
  - three spatial resolutions; R_λ 390, 650 (easier than for R_λ 1300 ...)

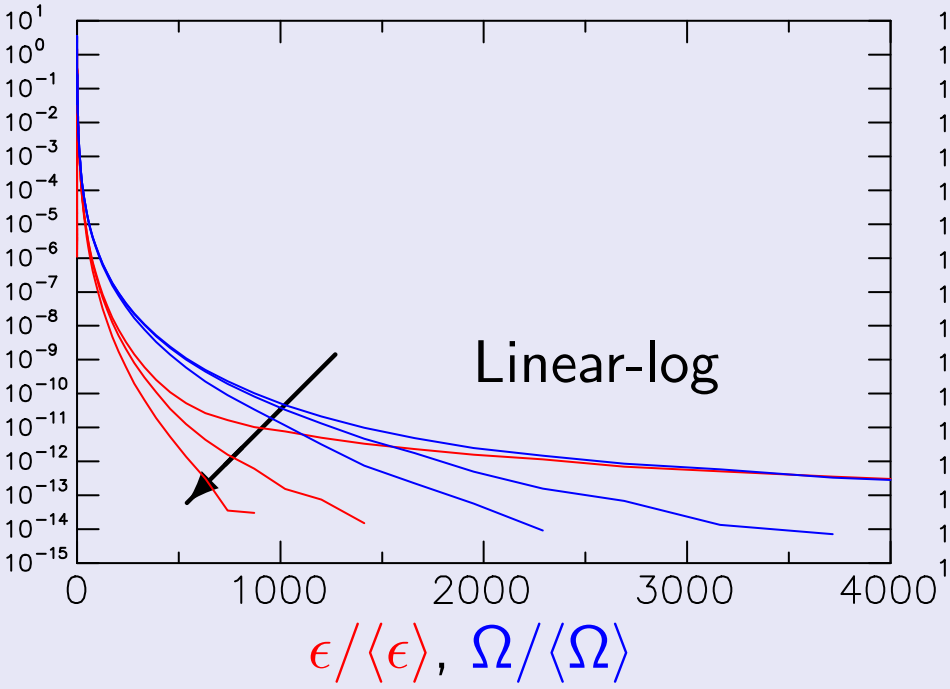
Fluctuations of Dissipation and Enstrophy

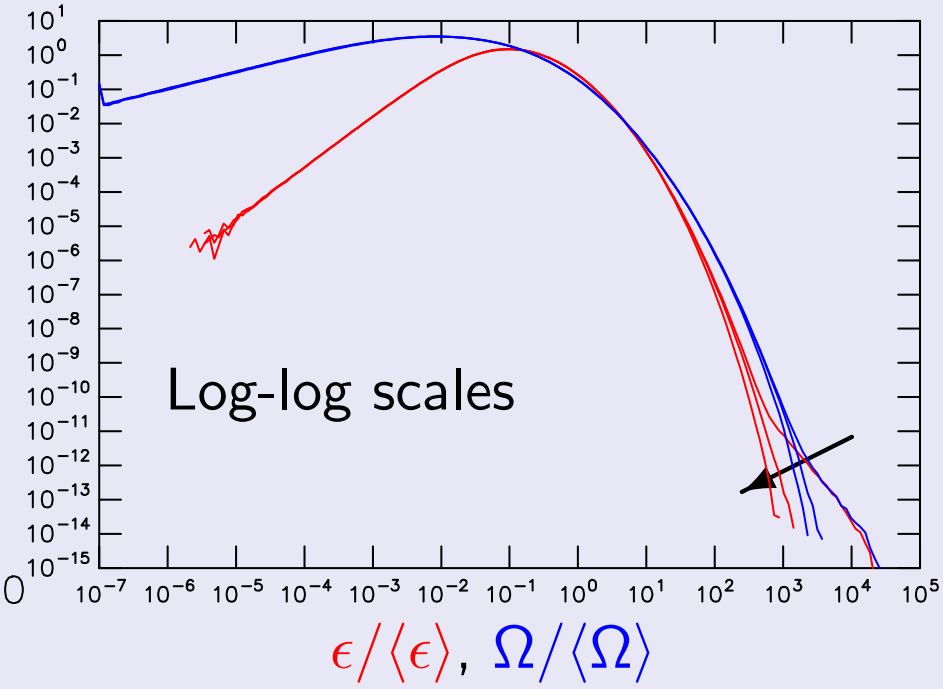
$$\epsilon \equiv 2\nu s_{ij} s_{ij} \ ; \quad \Omega \equiv \omega_i \omega_i$$

- Quadratic measures of local rates of strain and rotation of fluid elements subjected to disorder in turbulence
  - symmetric and anti-symmetric parts of velocity gradient tensor
- High strain rate: dispersion, breakup of flame surfaces
- High rotation rate: vortex filaments, preferential concentrations
- Fluctuations of ϵ vital in theories of intermittency (Kolmogorov 1962)
  - properties of local averages over 3D region of linear size r, with r varied through dissipative and inertial ranges
  - larger fluctuations expected at higher Reynolds numbers
- Homogeneous turbulence: $\langle \epsilon \rangle = \nu \langle \Omega \rangle$, but higher moments differ

Dissipation and Enstrophy: PDFs and Spatial Filtering

- Enstrophy is more intermittent. But dissipation is more sensitive
- R_λ 650, $k_{max}\eta = 2.8$ (4096^3); cutoff at $k_c/k_{max} = 1$, 0.75, 0.5





- Far (power-law like) tails due to high k modes, affected by aliasing
- Better spatial resolution: similar issue, pushed to larger amplitudes

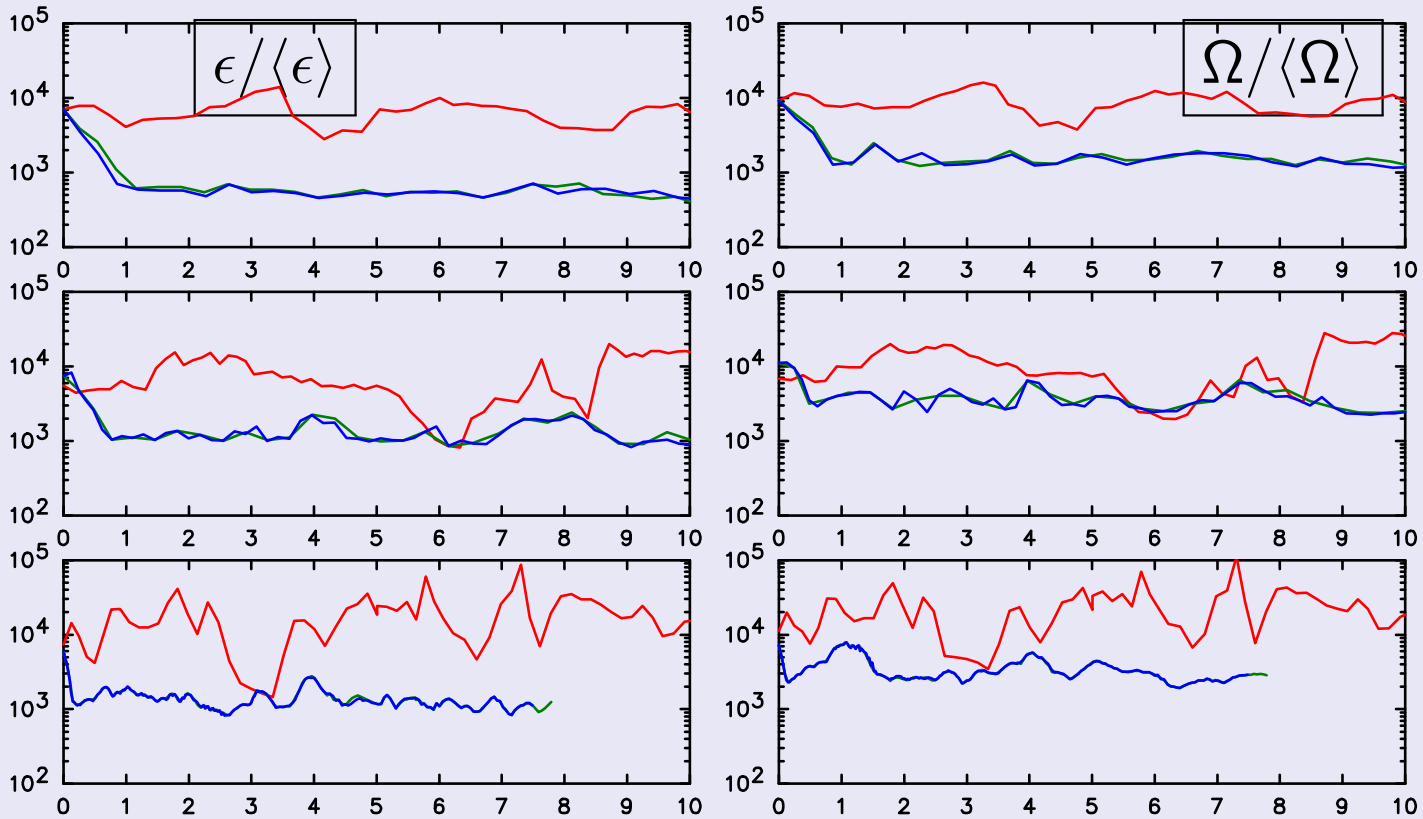
Resolution in time: Peak dissipation and enstrophy

$R_\lambda \sim 650$, $C = 0.6, 0.3, 0.15$; $\epsilon/\langle\epsilon\rangle$ and $\Omega/\langle\Omega\rangle$ vs. t/τ_η

$k_{max}\eta \approx 1.3$

$k_{max}\eta \approx 2.7$

$k_{max}\eta \approx 5.4$



- $C = 0.6$ gives spuriously large peaks; 0.3 and 0.15 almost the same
- Sensitivity greater for $\epsilon/\langle\epsilon\rangle$, consistent with effects of filtering

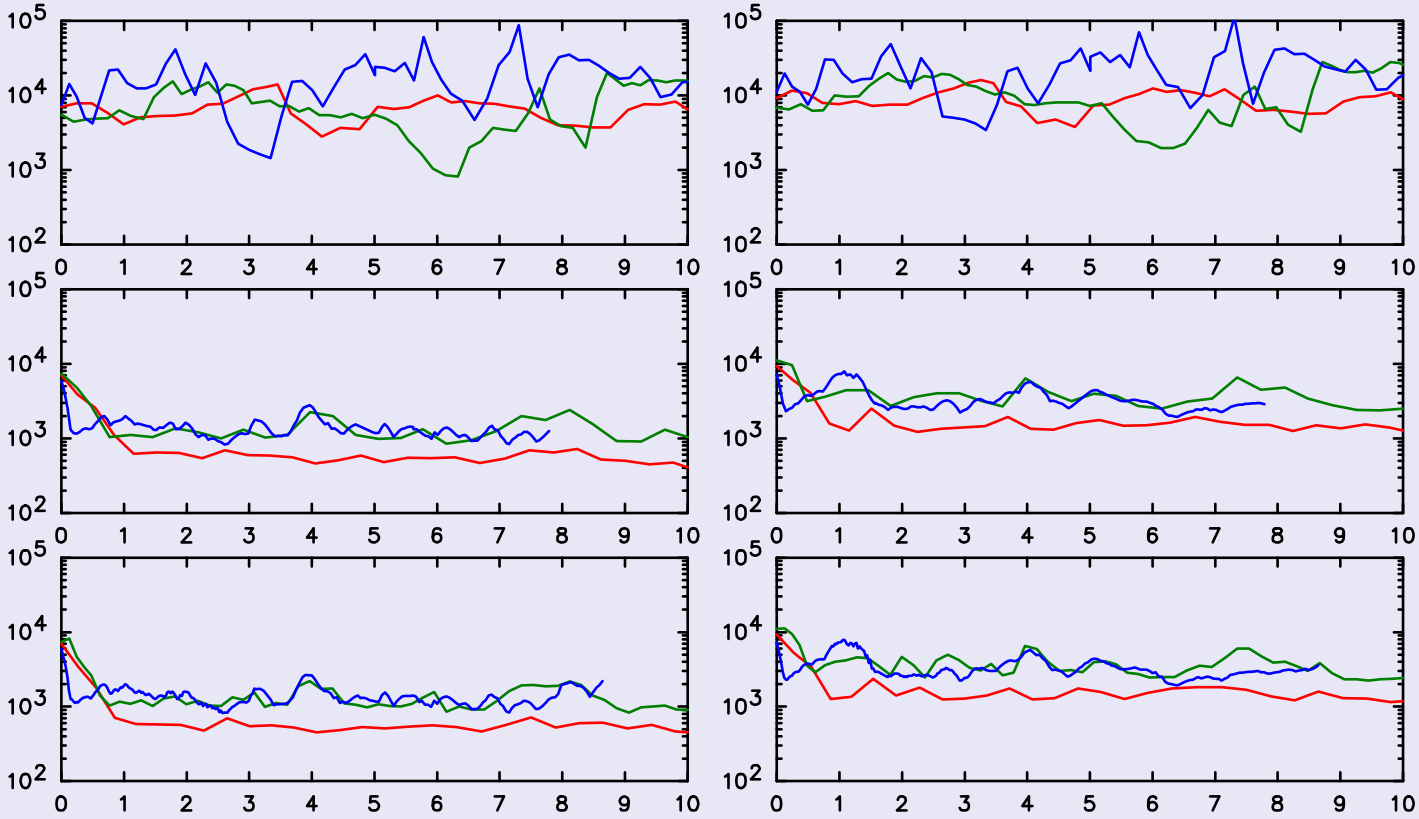
Resolution in space: Peak dissipation and enstrophy

$R_\lambda \sim 650$, $k_{max}\eta = 1.3, 2.7, 5.4$; $\epsilon/\langle\epsilon\rangle$ and $\Omega/\langle\Omega\rangle$ vs. t/τ_η

$C = 0.6$

$C = 0.3$

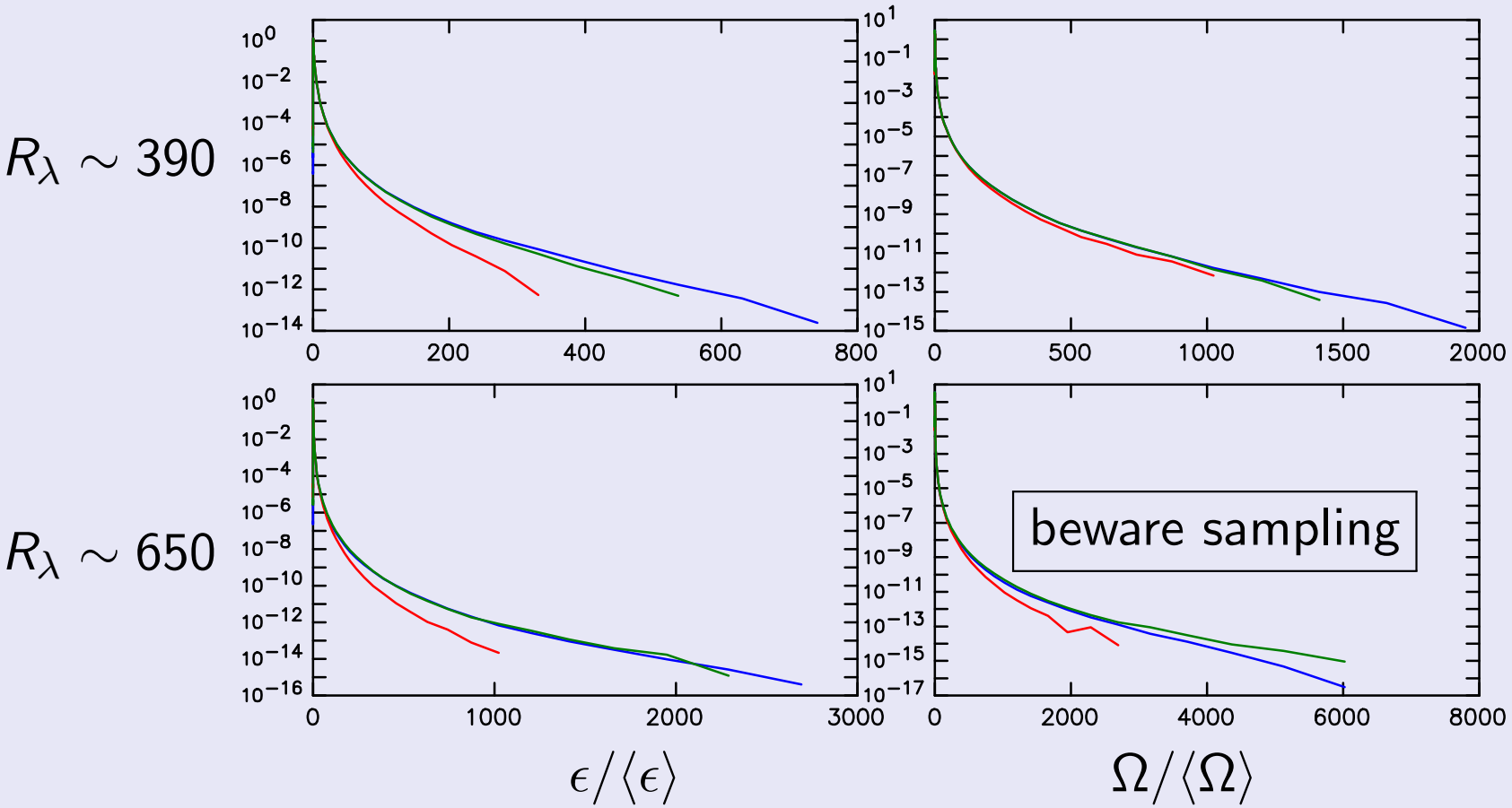
$C = 0.15$



- At $C = 0.6$: impact of using higher $k_{max}\eta$ is somewhat erratic
- $C = 0.3$ or lower: $k_{max}\eta$ at 5.4 captures larger gradients

Best results available for dissipation and enstrophy PDFs

$C = 0.15$, $k_{max}\eta \approx$ 1.3 2.7 5.4: Convergence apparently achieved



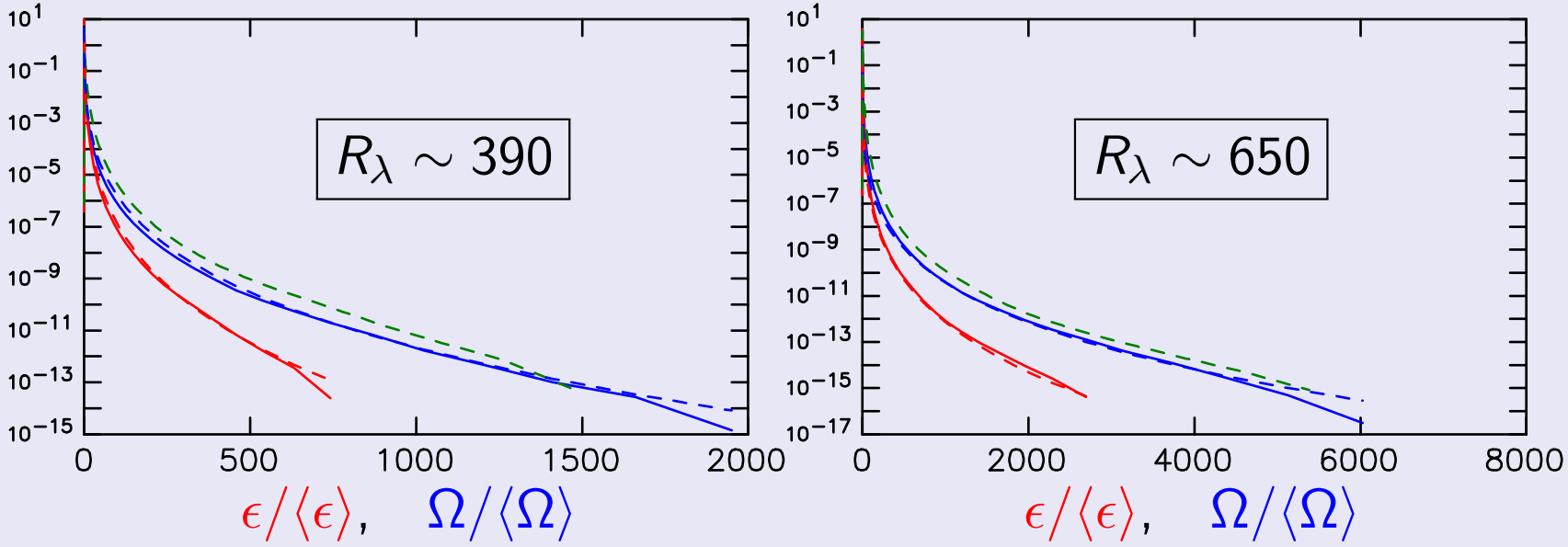
Tails stretch out further: as R_λ increases, and Ω relative to ϵ

Compare dissipation and enstrophy PDFs again

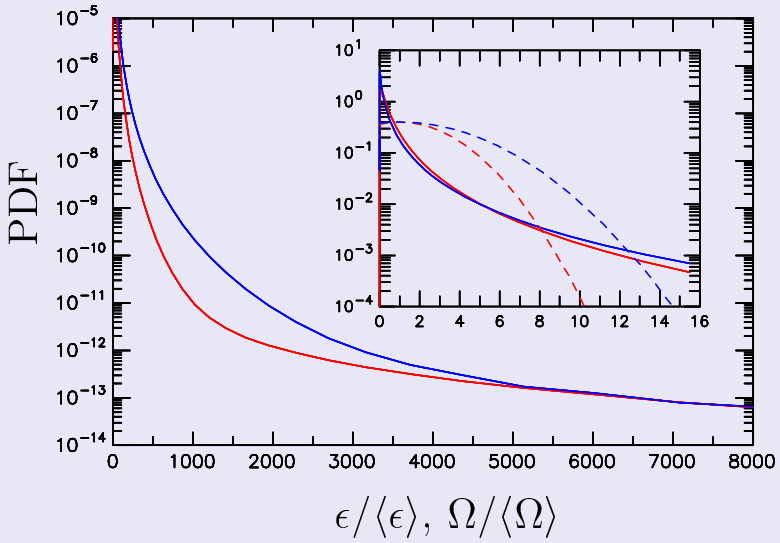
Best data at $C = 0.15$ (RK2), $k_{max}\eta \approx 5.4$: tails do not coincide, but the two PDFs have a strong similarity in shape ("stretched exponentials")

$$f_\epsilon(\epsilon/\langle\epsilon\rangle) \sim \exp[-b_\epsilon(\epsilon/\langle\epsilon\rangle)^{\gamma_\epsilon}] \; ; \qquad f_\Omega(\Omega/\langle\Omega\rangle) \sim \exp[-b_\Omega(\Omega/\langle\Omega\rangle)^{\gamma_\Omega}]$$

- Very good fit with $\gamma_\epsilon = \gamma_\Omega$ (dashed lines); $b_\epsilon > b_\Omega$
- Green dashed line is PDF of $2\epsilon/\langle\epsilon\rangle$
- Explanations may be possible using multi-fractal theory

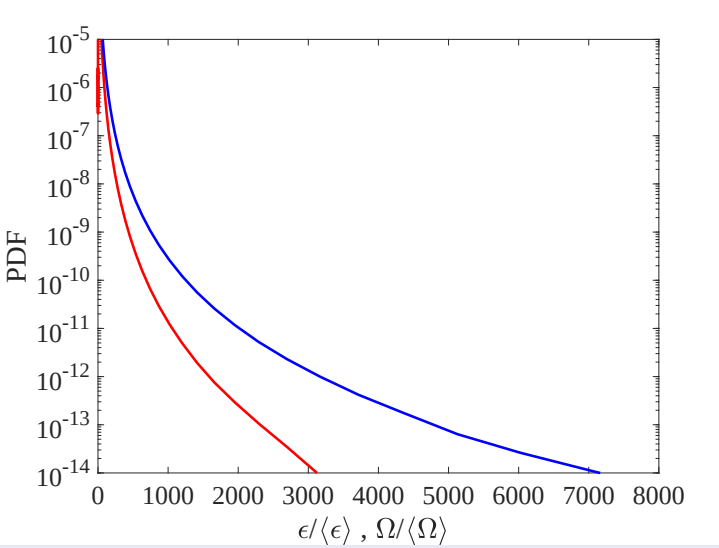


$R_\lambda \approx 1300$: Need to go beyond 8192^3



8192^3, $C = 0.6$, $\Delta x/\eta \approx 1.5$

(YZS, PNAS 2015)



12288^3, $C = 0.3$, $\Delta x/\eta \approx 1$

(D. Buaria, 95% holiday disc. on BW)

- Insufficient accuracy in time can lead to over-estimation of likelihood and intensity of extreme events (affected by aliasing errors)
- What about inertial range statistics, averaged over regions of linear size $\eta \ll r \ll L$ (expecting some smoothing in space)?

Local Averaging and Intermittency

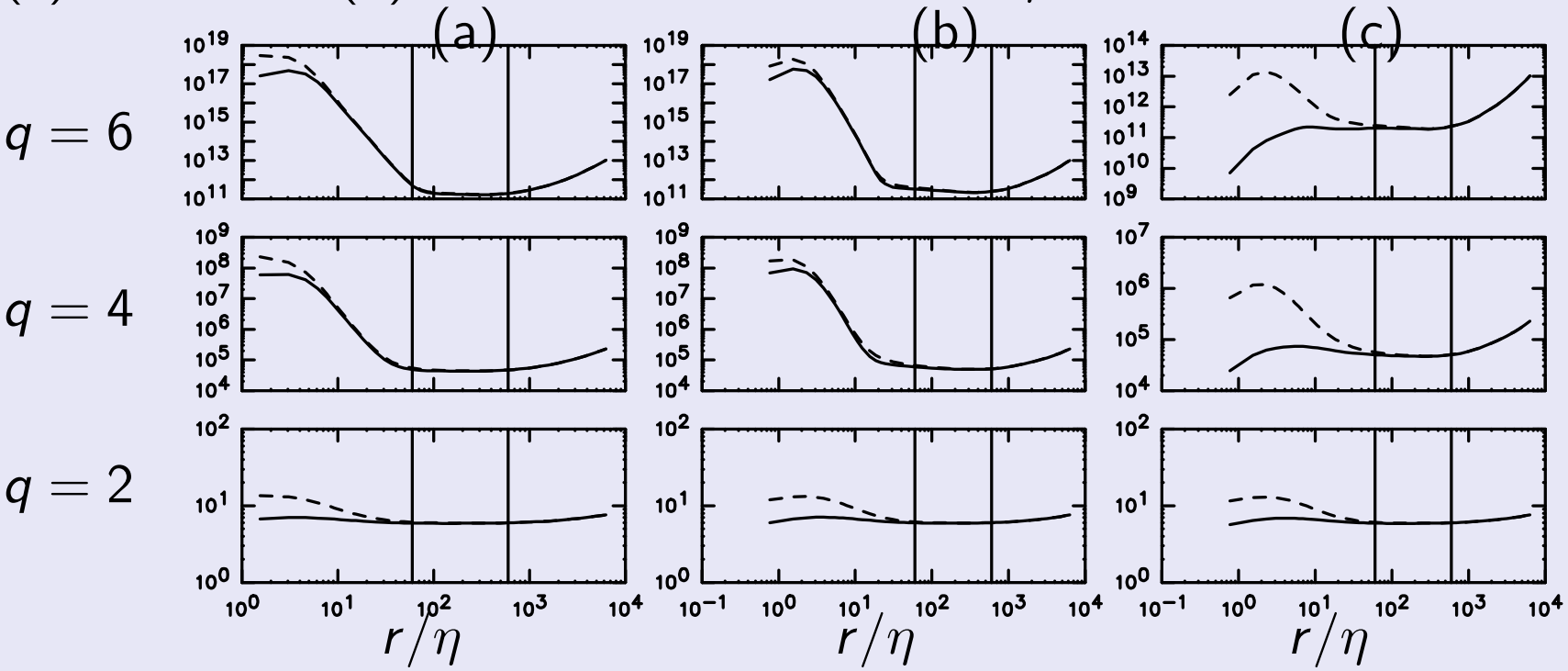
- Kolmogorov Refined Similarity (1962): replace $\langle\epsilon\rangle$ in inertial range formulas at scale size r by a local volume average

$$\epsilon_r(\mathbf{x}, t) = \frac{1}{Vol} \int_{Vol} \epsilon(\mathbf{x} + \mathbf{r}', t)\ d\mathbf{r}'$$

- 3D averages are important, but not often reported:
  - averaging along a line (1D) is *much* easier
  - 1D surrogate $(\partial u/\partial x)^2$ often used in experiments
  - DNS: also, nontrivial, only recently available (Iyer 2014)
- Inertial Range (intermediate scales):
  - $\langle\epsilon_r^q\rangle/\langle\epsilon\rangle^q \propto (r/\eta)^{-\tau_q}$ for orders $q = 2, 4, 6...$
  - Find τ_q: look for best fit of flat region for $\langle\epsilon_r^q\rangle/\langle\epsilon\rangle^q (r/\eta)^{\tau_q}$.
  - Exponents provide useful test of intermittency theories

Scaling Exponents

Compensated plots using same exponents as found for
(a) 8192^3 ensemble-averaged scaling of 3D averages, $k_{max}\eta \approx 2$
(b) single 16384^3 snapshot, by grid refinement, $k_{max}\eta \approx 3.8$
(c) filter from (b) to 8192^3 resolution, $k_{max}\eta \approx 1.9$



- Great variability at small r (dominated by extreme events)
- But relatively robust scaling in inertial range

Circulation

A measure of rotation in fluid flow

- Line integral of velocity vector along a closed path C, equivalent to integral of vorticity ($\omega \equiv \nabla \times \mathbf{u}$) around the enclosed area

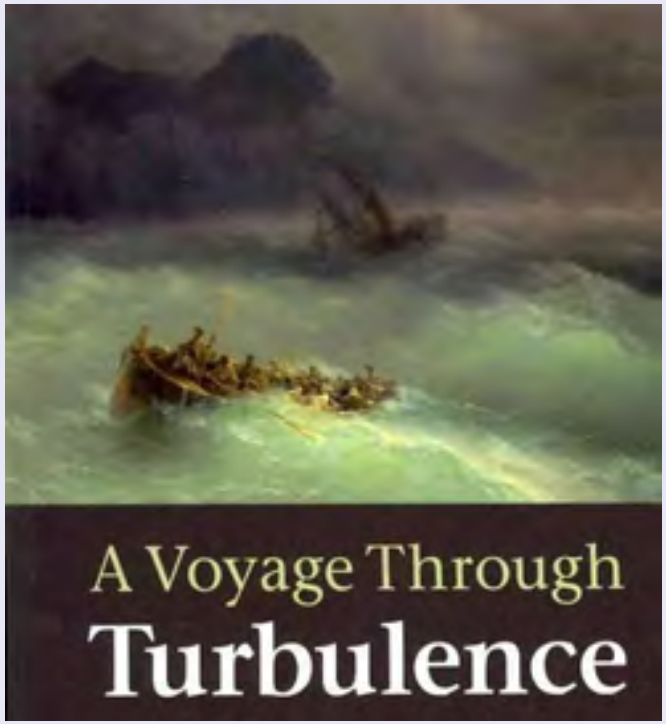
$$\Gamma = \oint_C \mathbf{u} \cdot d\mathbf{s} = \iint_A \omega \cdot \mathbf{n} \; dA$$

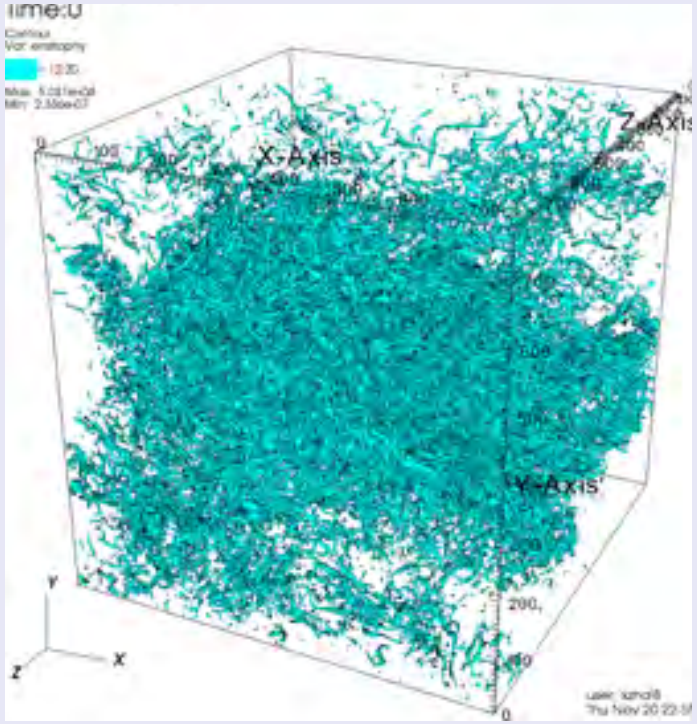
- Aerodynamics: circulation proportional to lift force
- Turbulence: spatial structure of vorticity field

What do we know about circulation in turbulence?

- Not so much — not a directly measurable quantity
- Theory — perhaps shape of closed path plays only a minor role?
- Vorticity is highly intermittent, but integration weakens this property
- DNS results show only very weak intermittency at intermediate scales (and a low sensitivity to numerical errors)

Concluding Remarks I: BW and Turbulence





An exciting journey, some eventful moments, but very rewarding:

- High-resolution simulations allowed us to address difficult questions
- Learned some lessons, but perhaps that is how science is done (?)
- A serious ramp-up in publication activity is under way

Concluding Remarks II: The Past and the Future

NSF infrastructure investments have had huge impact:

- Computational science as pursuit of scientific inquiry using computational methods and resources not imaginable before
- The meaning of "massive parallelism" changing over time, while researchers were challenged, and helped, to dream on ...
- Massive datasets, and gains in human resource development

As we look towards the "next Blue Waters":

- Recent architectural trends demand a change in paradigm for many (turbulence or not), and willingness to take risks
- The need for computational resources knows no boundaries in regard to discipline, nor the fundamental or applied debate
- Greater support needed for data and software repositories, as well as algorithmic development (often involving much risk!)