Scaling, structure functions and all that…

S. C. Chapman

Notes for MPAGS MM1 Time series Analysis

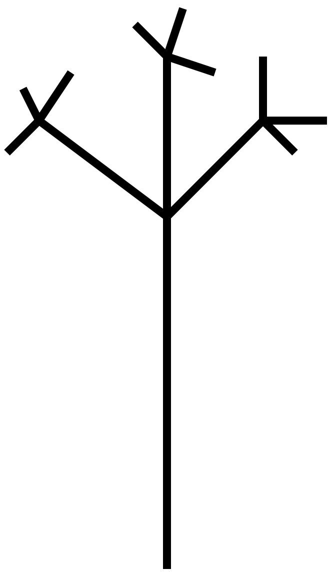
•SCALING: Some generic concepts: universality, turbulence, fractals and multifractals, stochastic models
•RESCALING PDFS AND STRUCTURE FUNCTIONS
•FINITE LENGTH TIMESERIES, UNCERTAINTIES, EXTREMES-'real data' examples

Scaling

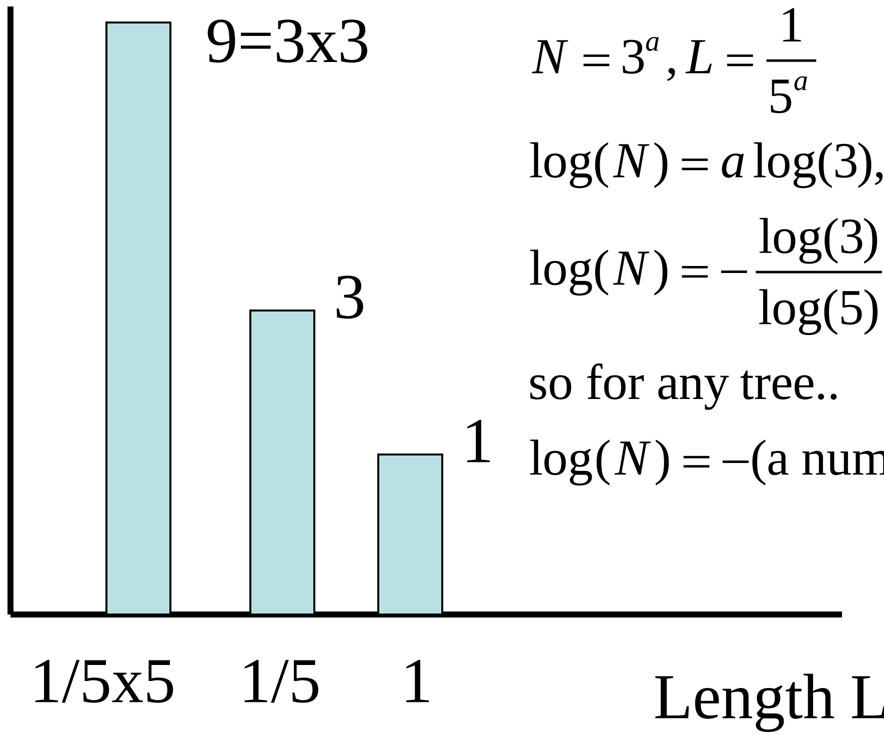
Some ideas and examples

Scaling and universality-Branches on a self-similar tree

Each branch grows 3 new branches, 1/5 as long as itself..



Number N of branches of length L



$$N = 3^a, L = \frac{1}{5^a}$$

$$\log(N) = a\log(3), \log(L) = -a\log(5)$$

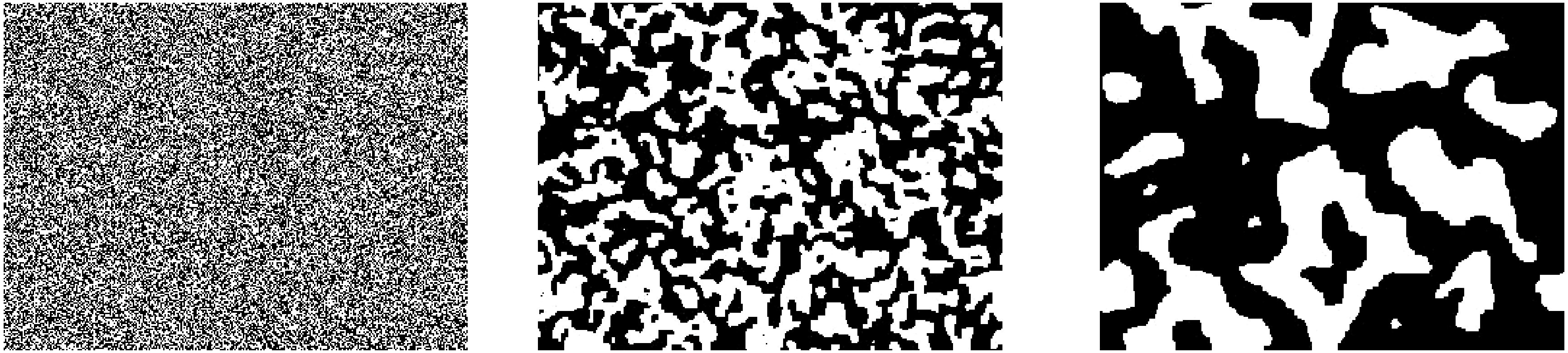
$$\log(N) = -\frac{\log(3)}{\log(5)}\log(L)$$

so for any tree..

$$\log(N) = -(\text{a number})\log(L)$$

Segregation/coarsening- a selfsimilar dynamics

Rules: each square changes to be like the majority of its neighbours
Coarsening, segregation, selfsimilarity

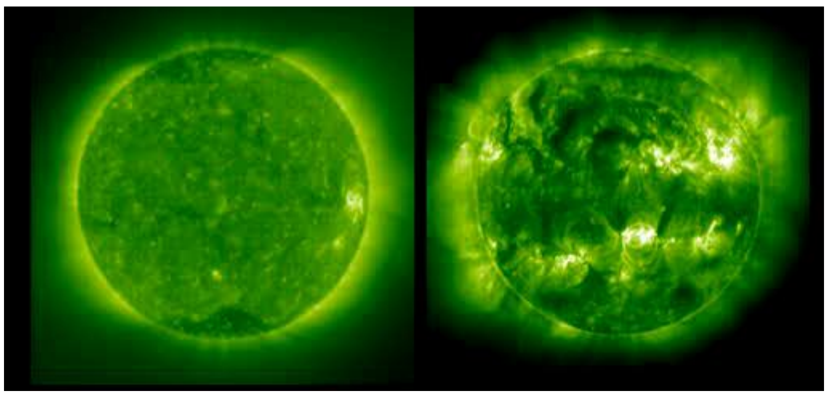


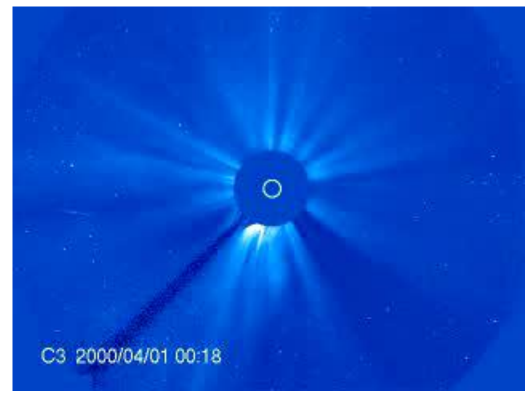
Courtesy P. Sethna

Solar corona over the solar cycle

SOHO-EIT image of the corona
at solar minimum and solar maximum
- Magnetic field structure

SOHO- LASCO image
of the outer corona
near solar maximum





The solar wind as a turbulence laboratory (will use as an example)

Solar wind at 1AU power spectra- suggests inertial range of (anisotropic MHD) turbulence. Multifractal scaling in velocity and magnetic field components.. AND something else in B magnitude..

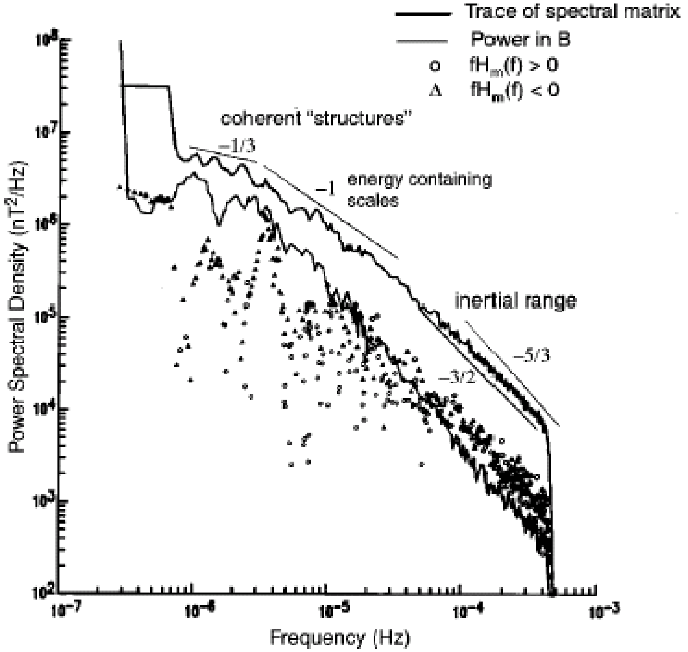


FIG. 1. A power spectrum of the solar wind magnetic field from a time series spanning more than a year. The upper curve is the trace of the power spectral matrix of the three components of **B**, the lower solid curve is the power in |**B**|, and the circles and triangles are the positive and negative values of the [reduced (Ref. 91)] magnetic helicity (Refs. 39 and 92) respectively.

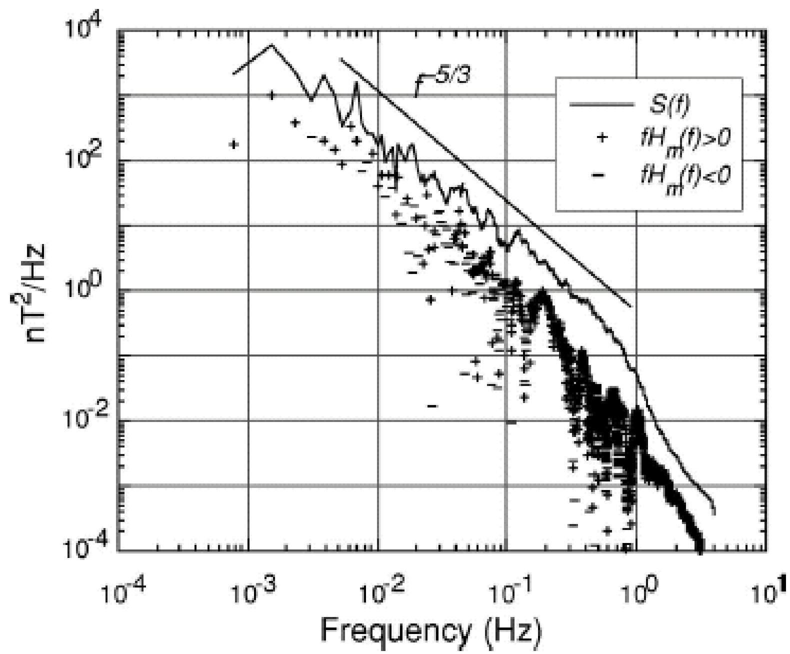


FIG. 2. A power spectrum of Mariner 10 data from 0.5 AU showing the dissipation range of magnetic fluctuations. Also shown are positive and negative values of $fH_m(f)$.

Goldstein and Roberts, POP 1999, See also Tu and Marsch, SSR, 1995

Quantifying scaling I

'Fractal' – self- affine scaling

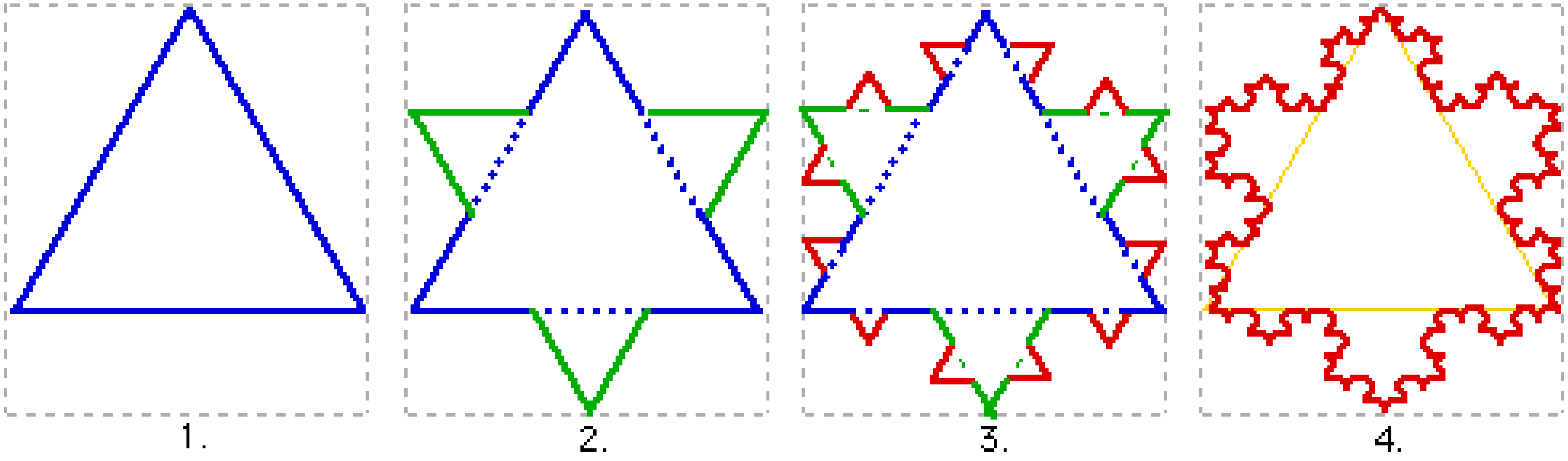
Uncertainties, finite size effects

Link to SDE models (self- affine processes)

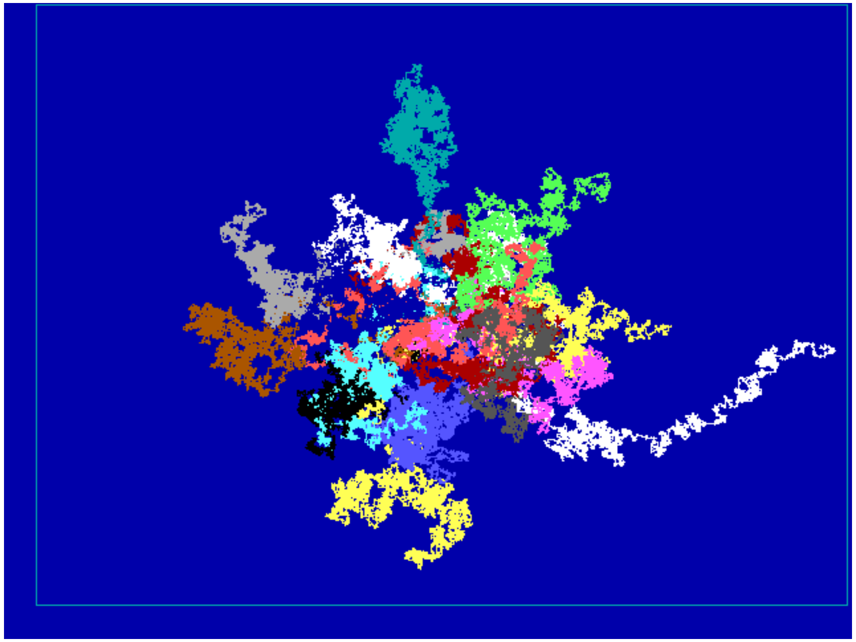
A regular fractal

Koch snowflake

line length $l \sim (4/3)^n$



A random fractal



16 particles- Brownian random walk

consider a random walk in 2D

$$\underline{r}(t_n) = \underline{r}_n = \underline{r}_{n-1} + \underline{l}$$

$$\underline{r}_n \bullet \underline{r}_n = r_{n-1}^2 + 2\underline{r}_{n-1} \bullet \underline{l} + l^2$$

$$\langle \underline{r}_n \bullet \underline{r}_n \rangle = \langle r_n^2 \rangle = \langle r_{n-1}^2 \rangle + l^2$$

$$\langle r_n^2 \rangle = nl^2$$

so if n steps take time t_n

$$\langle r_n^2 \rangle \sim t_n \text{ or } r \sim t^{1/2}$$

Data Renormalization

Consider a timeseries $x(t)$ sampled with precision Δ. We construct a *differenced* timeseries

$\delta x(t,\tau) = y(t,\tau) = x(t+\tau) - x(t)$ so

$x(t+\tau) = x(t) + y(t,\tau)$ and $y(t,\tau)$ is a random variable

then

$$x(t) = y(t_1,\Delta) + y(t_2,\Delta) + \ldots + y(t_k,\Delta) + y(t_{k+1},\Delta) + \ldots + y(t_N,\Delta)$$

$$= y^{(1)}(t_1, 2\Delta) + .. + y^{(1)}(t_k, 2\Delta) + .. + y^{(1)}(t_{N/2}, 2\Delta)$$

$$= y^{(n)}(t_1, 2^n\Delta) + \ldots + y^{(n)}(t_k, 2^n\Delta) + .. + y^{(n)}(t_{N/2^n}, 2^n\Delta)$$

we seek a self affine scaling

$y' = 2^\alpha y, \tau' = 2\tau, y^{(n)} = 2^{n\alpha} y,$ as $\tau = 2^n\Delta$

for arbitrary τ, normalize such that

$$y'(t,\tau) = \tau^\alpha y(t,\Delta)$$

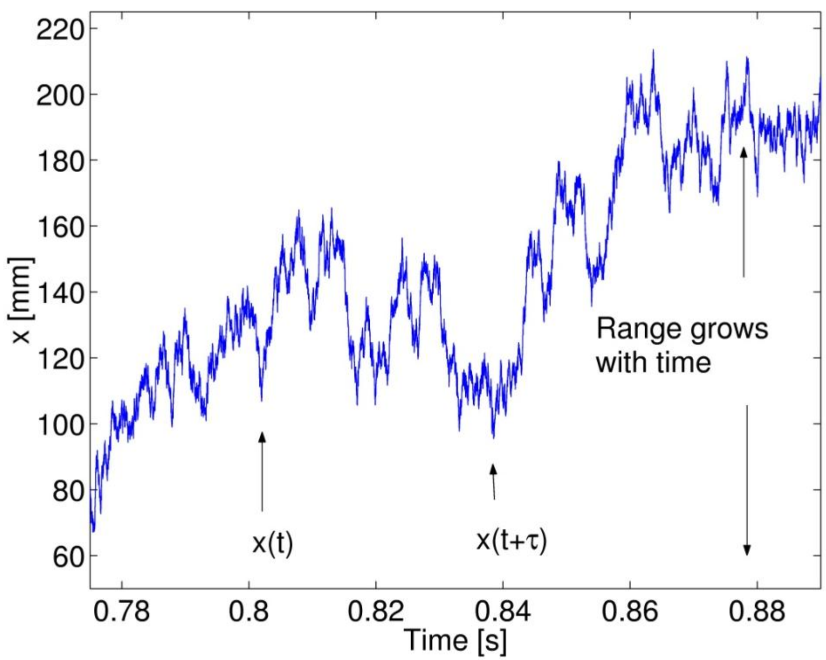
y is a random variable, so we have the same PDF under transformation:

$$P(y'\tau^{-\alpha})\tau^{-\alpha} = P(y)$$

the y are not Gaussian iid. We need to find α

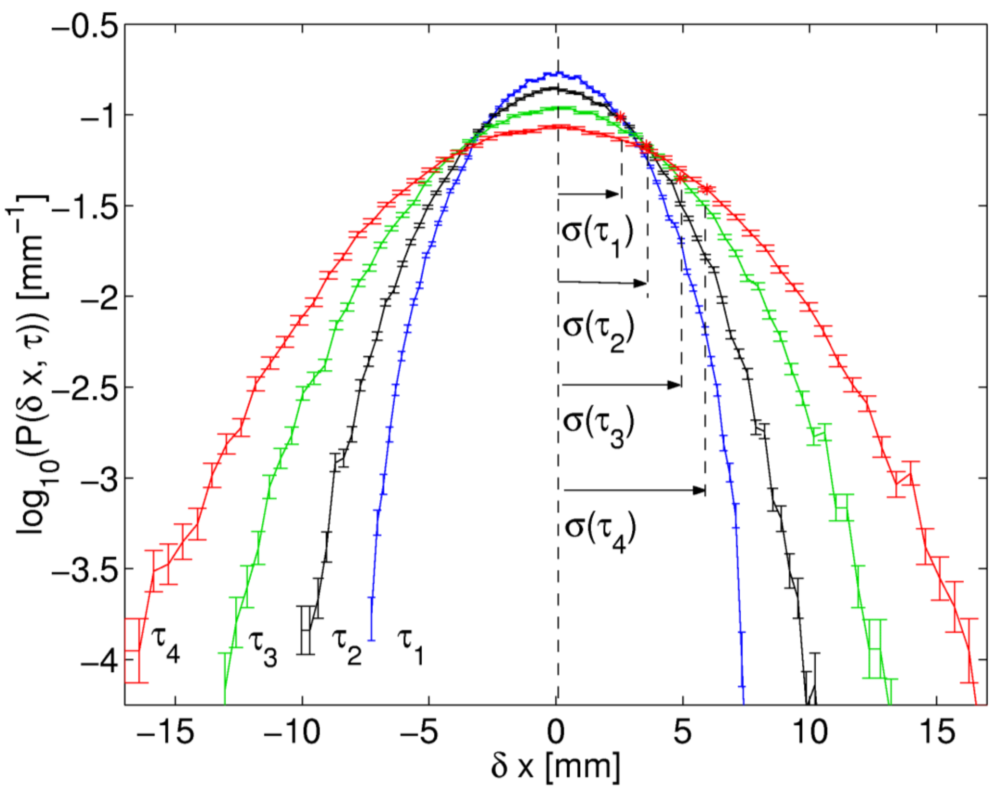
consider CLT case..

Self –affine ('fractal') scaling in timeseries



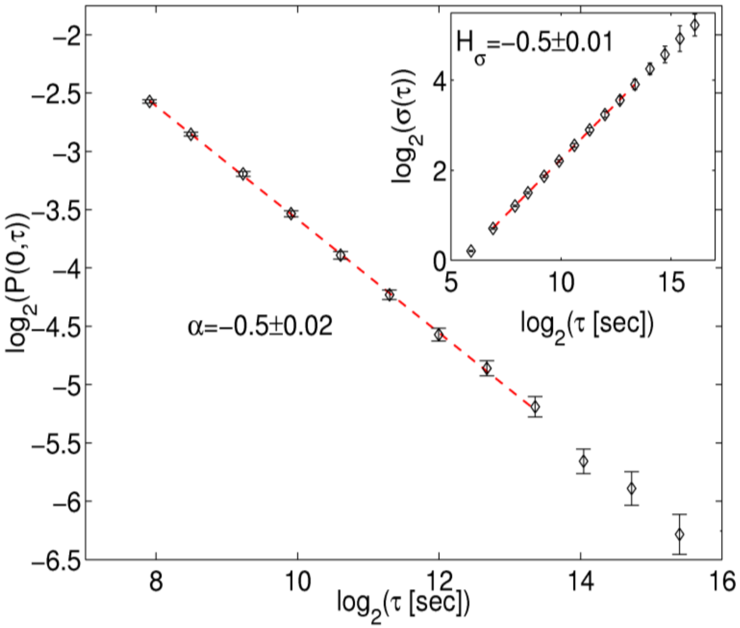
Example-Brownian walk

Fluctuations: $\delta x(t,\tau) = x(t+\tau) - x(t)$

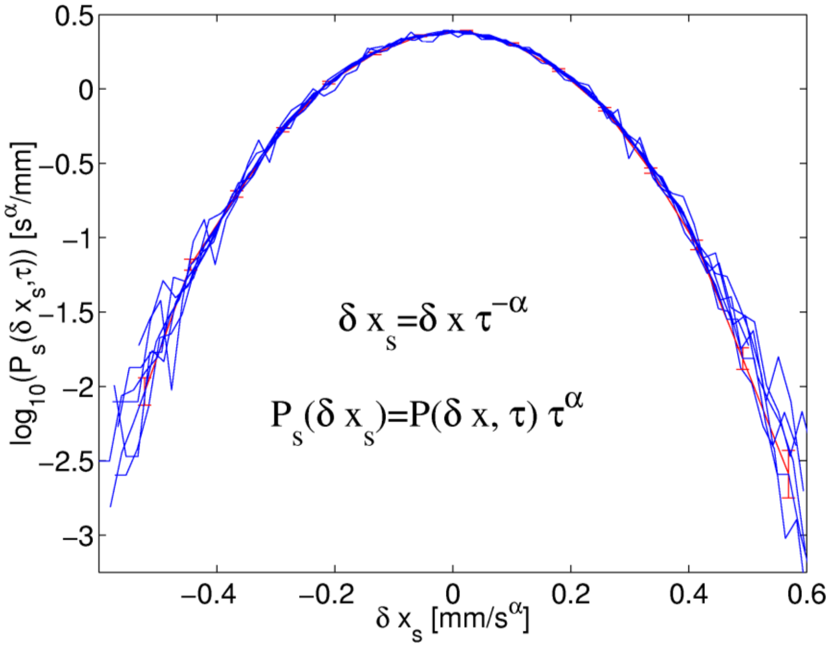


Probability of wandering different distances in a given time (Gaussian)

Rescale



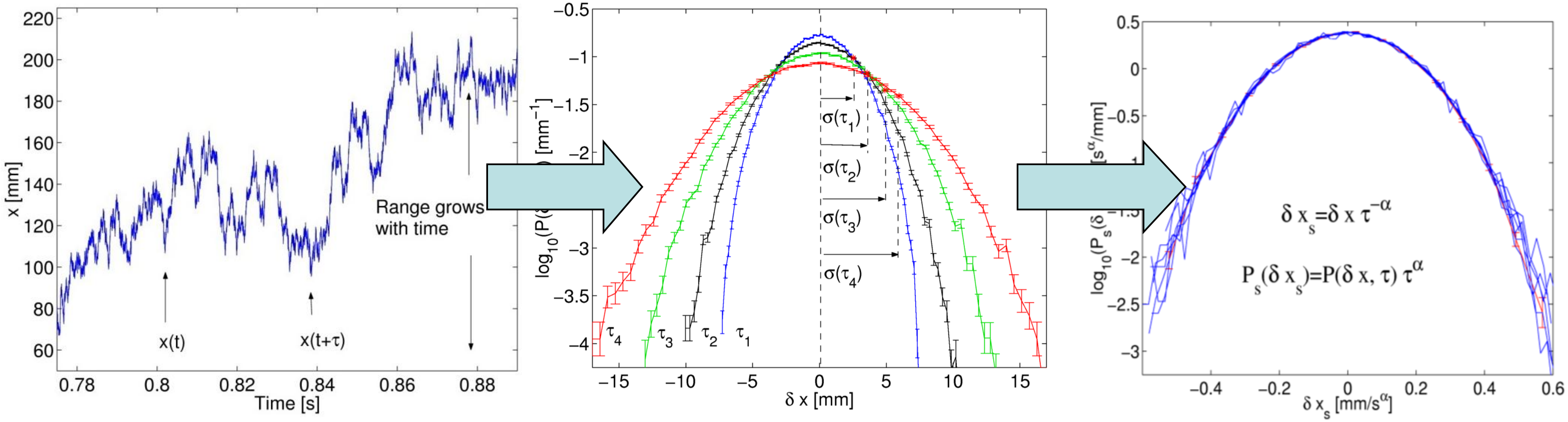
The height of the peaks is power law- a single factor rescales them



The same factor rescales all the curves- α=1/2

Self-similarity

Procedure to test for self affine scaling in a timeseries- Brownian walk (simple fractal)



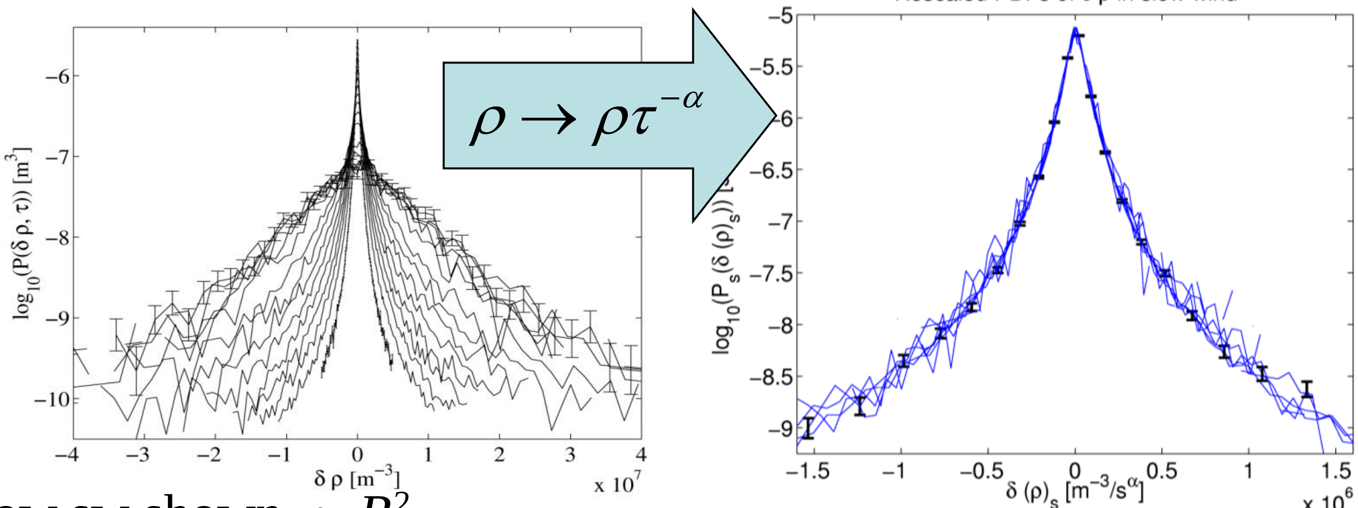
1) difference the timeseries $x(t)$ on timescale τ to obtain $y(t,\tau)=x(t+\tau)-x(t)$

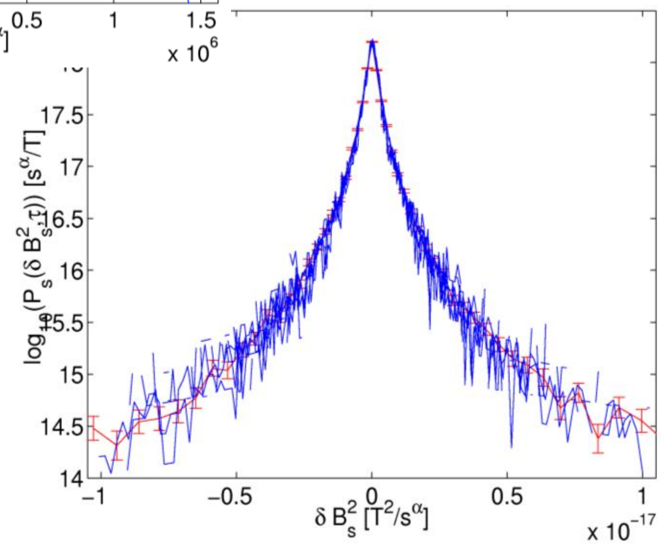
2) $P(y,\tau)$ are self- similar (fractal) - *if* same function under single parameter rescaling

3) rescaling parameter comes from the data eg $\sigma(\tau)\sim\tau^{\alpha}, \alpha = 1/2$ here

4) so moments of the PDF: $\left\langle y(t,\tau)^{p}\right\rangle_{t} \sim \tau^{\alpha p}$

example- ρ,B^2 in the solar wind





slow sw shown, ρ, B^2
selfsimilar scaling up to τ~few hrs
WIND 46/98s
Key Parameters '95-'98
Approx 10^6 samples
Verified with ACE
Hnat, SCC et al GRL,2002, POP 2004

Diffusion- random walk

Brownian random walk

$$\frac{dx}{dt} = \eta$$

η is stochastic iid

diffusion equation

$$\frac{\partial P(y,t)}{\partial t} = D\nabla^2 P(y,t)$$

$\Rightarrow P(y,t)$ is Gaussian

Note: $y(t)$ is distance travelled in interval $t = \tau$ –a differenced variable

Renormalization-scaling system looks the same under

$$t' = \frac{t}{\tau},\ y' = \frac{y}{\tau^{\alpha}} \text{ and } \alpha = \frac{1}{2} \text{..........which implies } P(y',t') = \tau^{\alpha} P(y,t)$$

$\Rightarrow P(y,t)$ is Gaussian, the fixed point under RG

Fokker- Planck models

(see also fractional kinetics and Lévy flights)

Langevin equation

$$\frac{dx}{dt} = \beta(x) + \gamma(x)\eta$$

η stochastic iid

Fokker- Planck equation

$$\frac{\partial P(y,t)}{\partial t} = \nabla(A(y)P(y,t) + B(y)\nabla P(y,t))$$

can solve for $P(y,t)$

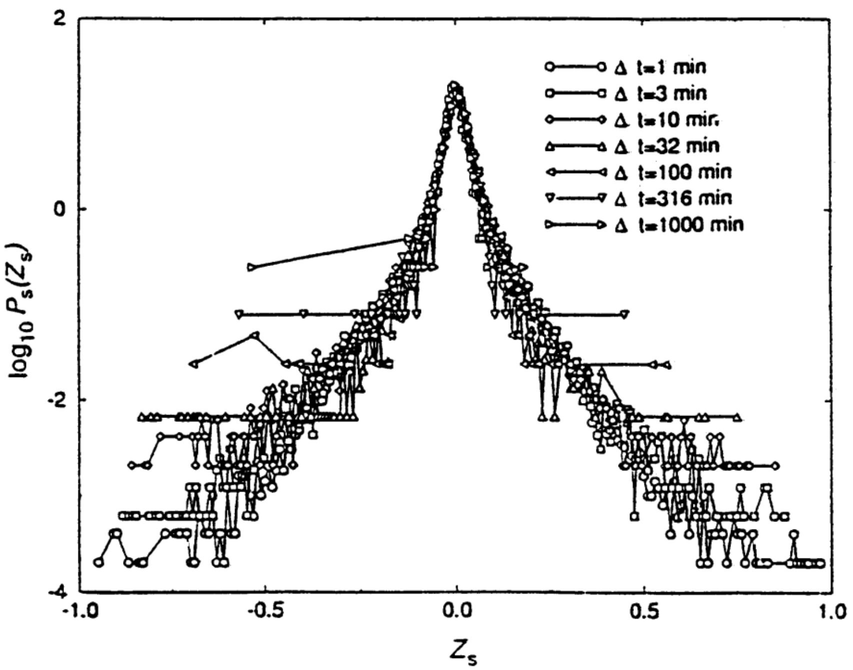
Note: $y(t)$ is distance travelled in interval $t = \tau$ –a differenced variable

Renormalization-scaling system looks the same under

$$t' = \frac{t}{\tau},\ y' = \frac{y}{\tau^{\alpha}} \text{ and } \alpha \neq \frac{1}{2} \text{..........which implies } P(y',t') = \tau^{\alpha} P(y,t)$$

A not so simple fractal timeseries- financial markets

- *Mantegna and Stanley- Nature,1995*
- S+P500 index
- ‘heavy tailed’ distributions
- Brownian walk in log(price) is the basis of Black Scholes (FP model for price dynamics)
- Non- Gaussian PDF, fractal scaling- Fractional Kinetics or non- linear FP



The efficient market

- ➢Efficient- arbitrageurs constantly trade to exploit differences in price
- ➢As a consequence any price differences are very short lived
- ➢The market is a ‘fair game’

Implies

- ▪Fluctuations are uncorrelated
- ▪Fluctuations aggregate many (*N*) trades, thus an equilibrium, large *N* model implies Gaussian statistics (CLT)
- ▪Change in price *S*, *dS* in *t*-*t*+*dt* governed by:

$$\frac{dS}{S} = \sigma dX + \mu dt$$

Black-Scholes and all that..

Anticipate a Diffusion equation for $\log(S)$ -since $\frac{dS}{S} = \sigma dX + \mu dt$

provided we have the self- similar scaling for

the stochastic variable dX

I $< dX^2 >\sim dt$

we can write an equation for price evolution

II $dS = A(S,t)dX + B(S,t)dt$

can then write a Taylor expansion for any $f(S)$ using I.

This leads to the B-S SDE for the price of options...

Riskless portfolio $\pi = f(S) + \beta S,\ f(S)$ is an option on stock S

key phenomenology is that of scaling

Nonlinear F-P model for self similar fluctuations- asymptotic result (alternative- fractional kinetics)

If the PDF of fluctuations $y = x(t+\tau) - x(t)$ on timescale τ is selfsimilar:

$P(y,\tau) = \tau^{-\alpha} P_s(y\tau^{-\alpha})$

P is then a solution of a Fokker- Planck equation:

$\frac{\partial P}{\partial \tau} = \nabla\left[AP + B\nabla P\right]$, where transport coefficients $A = A(y), B = B(y)$

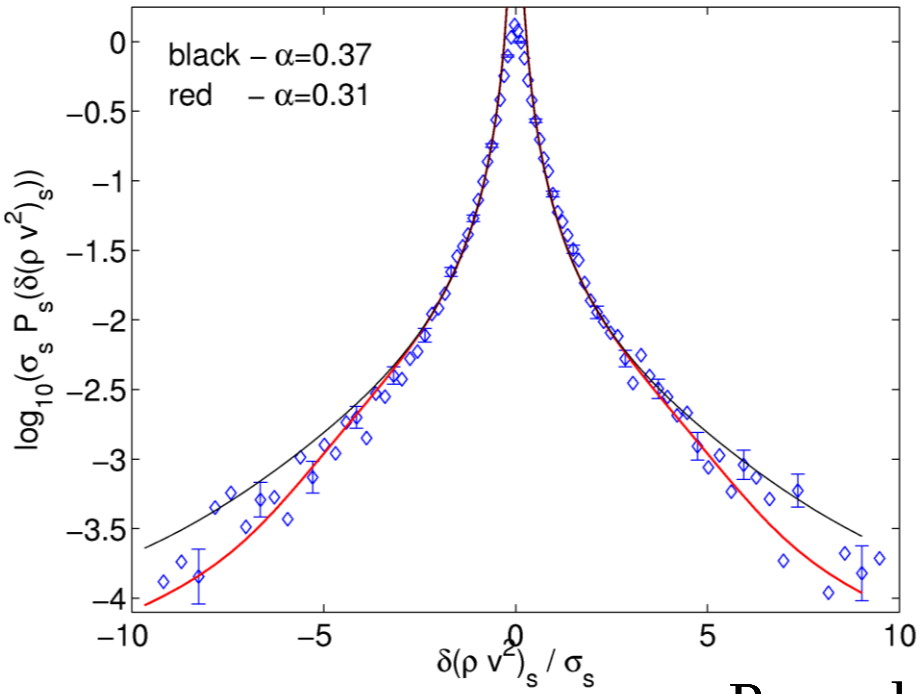
with $A \propto y^{1-1/\alpha}, B \propto y^{2-1/\alpha}$ we solve the Fokker- Planck for P_s

This corresponds to a Langevin equation: $\frac{dx}{dt} = \beta(x) + \gamma(x)\xi(t)$

and we can obtain β, γ via the Fokker- Planck coefficients

see Hnat, SCC et al. Phys. Rev. E (2003), Chapman et al, NPG (2005)

Fokker Planck fit to PDFs



Procedure:

1) Measure exponent
2) Solve FP for PDF functional form
3) Check this fits the observed PDF

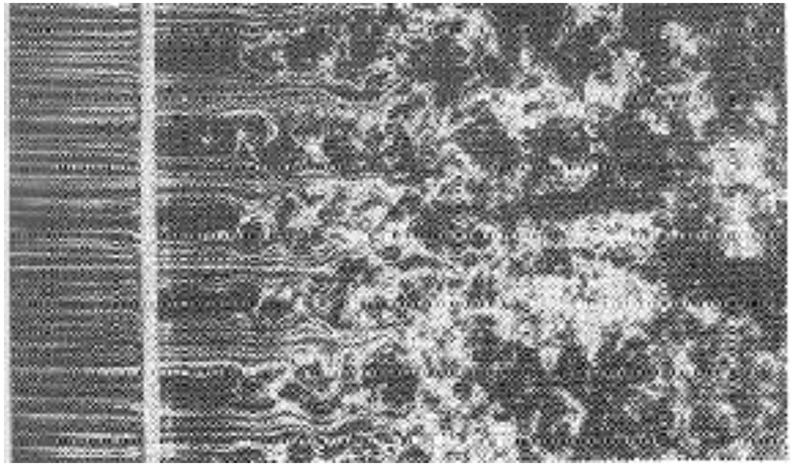
Turbulence

a la Komogorov, intermittency

beyond power spectra...

(NB we will introduce intermittency in the context of turbulence, but methods are quite general)

Turbulence



Dynamics are complex
Statistics are simple
Assume:
Isotropic
Stationary
Homogeneous

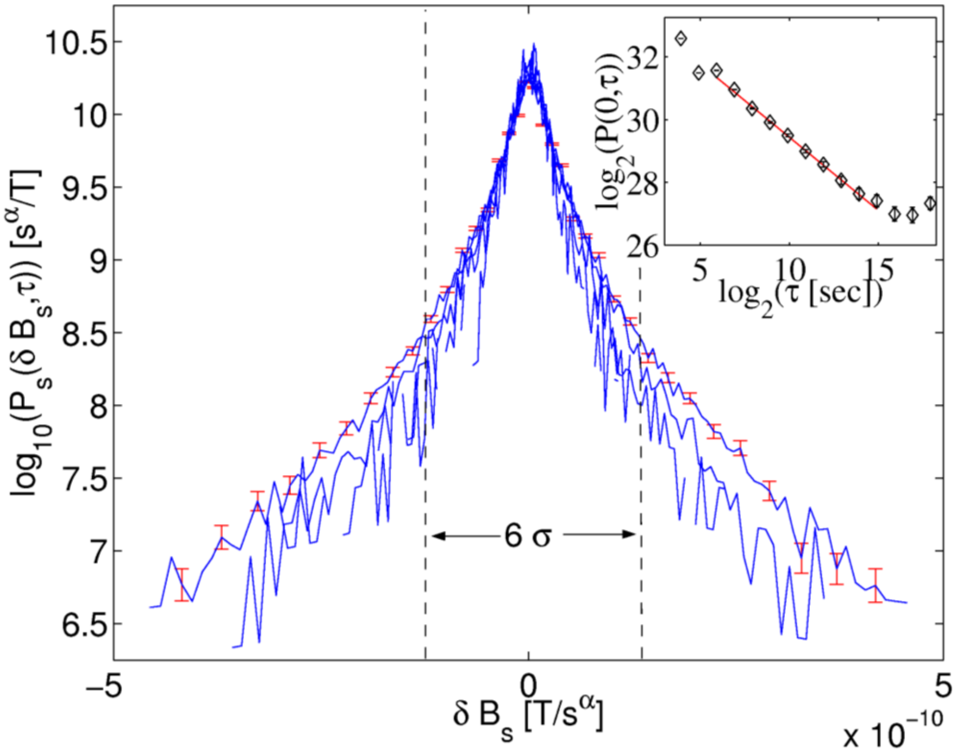
Example- strong multifractal

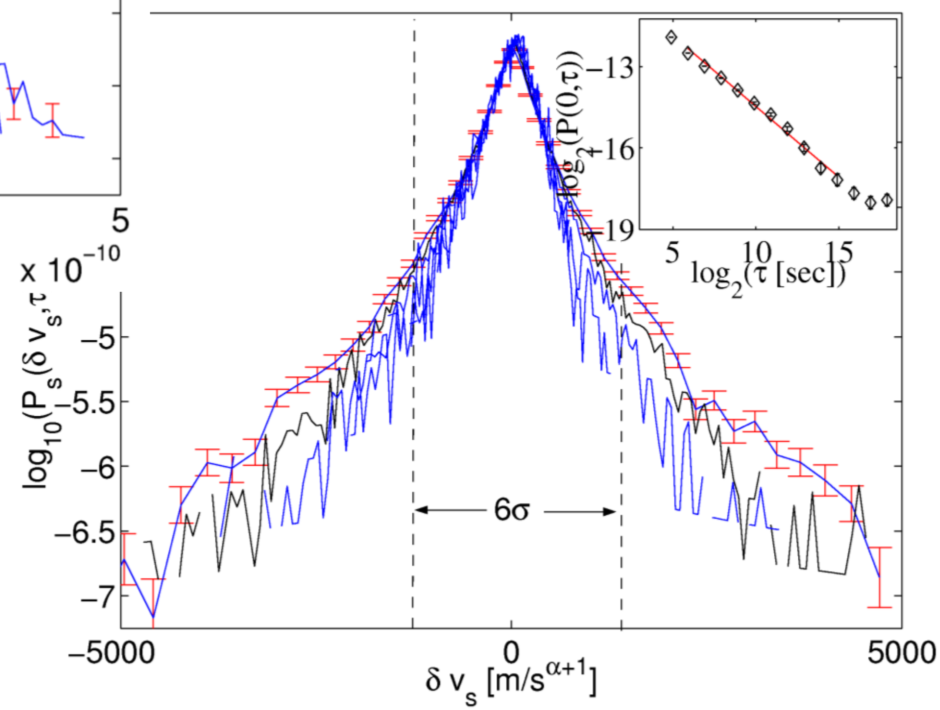
solar wind ***v,B***

moments

$$S^m = < \delta x^m > \sim \tau^{\zeta(m)}$$

$$\zeta(m) \text{ quadratic in } m$$





Intermittent turbulence-topology

Consider simple finite sized scaling system, scale lengths l_j

$\lambda = \left(l_{j-1}/l_j\right)^3, l = 1 \ldots N$ with N levels

from a smallest size $l_1 = \eta$ to the system size $l_N = L$, m_j patches on lengthscale l_j

Non space filling, intermittent patches: $< m_j^q > l_j^{\gamma(q)} = < m_{j-1}^q > l_{j-1}^{\gamma(q)} = < m_N^q > L^{\gamma(q)}$

Fractal support: $\frac{\varepsilon_{\mathrm{j}}^*}{l_j^\alpha} = \frac{\varepsilon_{\mathrm{j-1}}^*}{l_{j-1}^\alpha} = \frac{\varepsilon_{\mathrm{N}}^*}{L^\alpha}$ where $\varepsilon_{\mathrm{j}}^*$ is 'active quantity' per patch, lengthscale l_j

Conservation:

active quantity per lengthscale $\varepsilon_j = m_j \varepsilon_{\mathrm{j}}^*$

$< \varepsilon_j > = \varepsilon_0$ which fixes $\gamma(1) = \alpha$ or $\mu(1) = 0$

when these combine to give:

$$< \varepsilon_j^q > = (\varepsilon_N^*)^q < m_N^q > \left(\frac{l_j}{L}\right)^{[\alpha q - \gamma(q)]} = \varepsilon_0^q \left(\frac{l_j}{L}\right)^{-\mu(q)}$$

Intermittency-
as a deviation from a space filling cascade (Kolmogorov turbulence)

velocity difference across an eddy $d_r v = v(l+r) - v(l)$

eddy time $T(r)$ and energy transfer rate $\varepsilon_{\mathrm{r}} \propto \dfrac{d_r v^2}{T}$

have T as the eddy turnover time $T \propto r/d_r v$ so that $\varepsilon_{\mathrm{r}} \propto \dfrac{d_r v^3}{r}$

If the flow is non- intermittent $\langle \varepsilon_r{}^p \rangle = \bar{\varepsilon}^p, r$ independent for any p

$\Rightarrow \langle d_r v^p \rangle \propto r^{p/3} \bar{\varepsilon}^{p/3} \sim r^{\zeta(p)}$ -$\zeta(p) = \alpha p$ linear with p – *selfsimilar(fractal)* scaling

intermittency correction- r dependence $\langle \varepsilon_r{}^p \rangle \propto \bar{\varepsilon}^p \left(r/L \right)^{\tau(p)}$

$\Rightarrow \langle d_r v^p \rangle \propto r^{p/3} \bar{\varepsilon}^{p/3} \left(L/r \right)^{\tau\left(p/3\right)} \sim r^{\zeta(p)}$- $\zeta(p)$ quadratic in p

$\langle \varepsilon_r \rangle = \bar{\varepsilon}$ independent of r (steady state) so $\tau(1) = 0,$

$\Rightarrow \zeta(p)$ must monotonically increase (and $\zeta(p) > 1$ for some p)

in situ single point observations take $r \equiv t$: measure $\zeta(p)$ from $< d_t v^p > \sim t^{\zeta(p)}$

$p = 6$ needed to measure $\tau(2)$! predicted from phenomenology

Quantifying scaling II

Multifractal scaling and structure functions

Turbulence and scaling

structures on many length/timescales.

single spacecraft- time interval τ a proxy for space

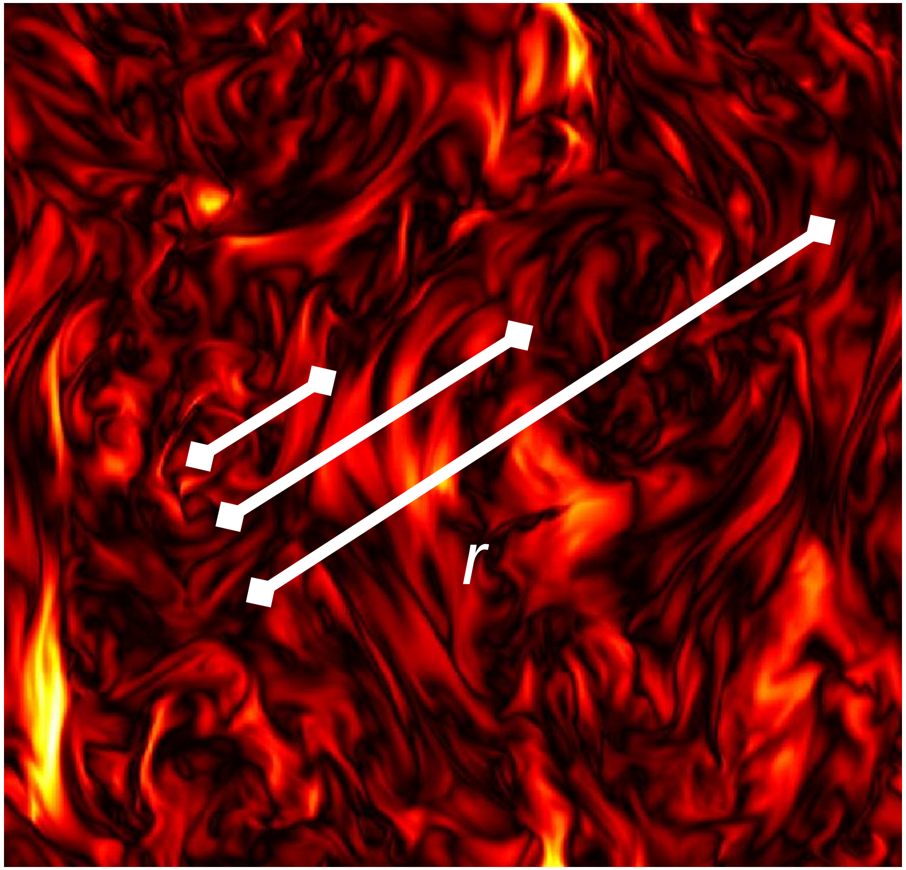
Reproducible, predictable in a *statistical* sense.

to focus on any particular scale r take a difference:

$y(l,r) = x(l+r) - x(l)$

look at the statistics of $y(l,r)$

power spectra- compare power in Fourier modes on different scales r



DNS of 2D compressible MHD turbulence
Merrifield, SCC et al, POP 2006,2007

Quantifying scaling

structures on many length/timescales.

Reproducible, predictable in a *statistical* sense.

look at (time-space) differences:

$y(r,l) = x(r+l) - x(r)$

$y(t,\tau) = x(t+\tau) - x(t)$

for all available t_k of the timeseries $x(t_k)$

test for statistical scaling i.e

structure functions $S_p(r) = <|y(r,l)|^p> \propto l^{\zeta(p)}$

or $S_p(\tau) = <|y(t,\tau)|^p> \propto \tau^{\zeta(p)}$

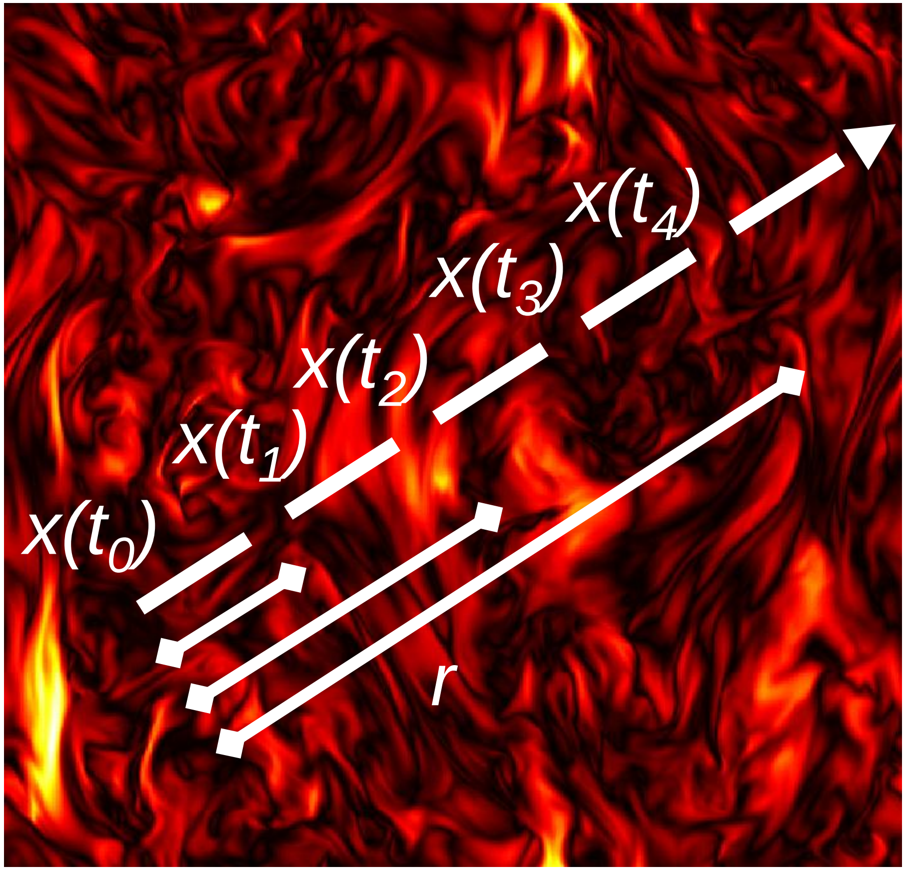
we want to measure the $\zeta(p)$

fractal (self- affine) $\zeta(p) \sim \alpha p$

multifractal $\zeta(p) \sim \alpha p - \beta p^2 + \ldots$

would like $<|y(r,l)|^p> = \int_{-\infty}^{\infty} |y|^p \, P(y,l) dy$

BUT finite system/data!



Theory-data comparisons- examples

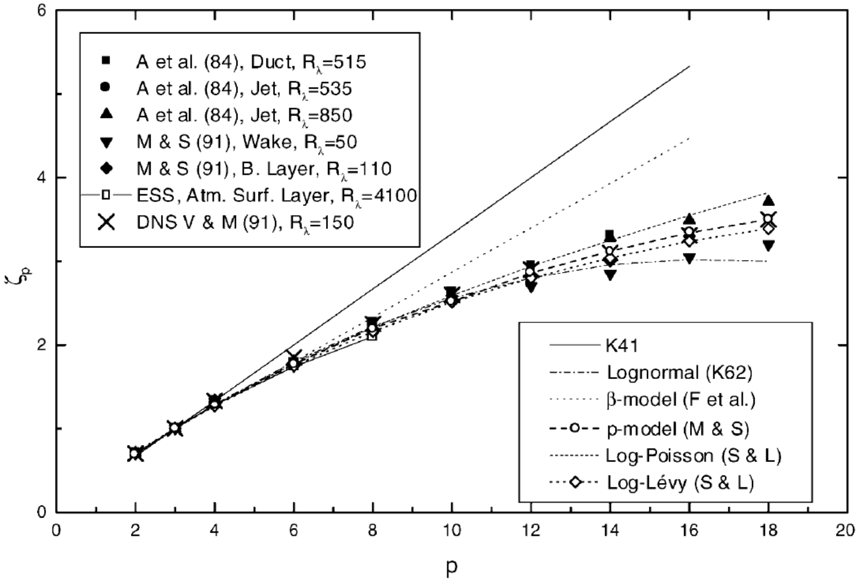


Fig. 11. Power-law exponents ζ_p of the structure functions as a function of the order p, together with the values predicted by K41 and the various intermittency models of Table 1.

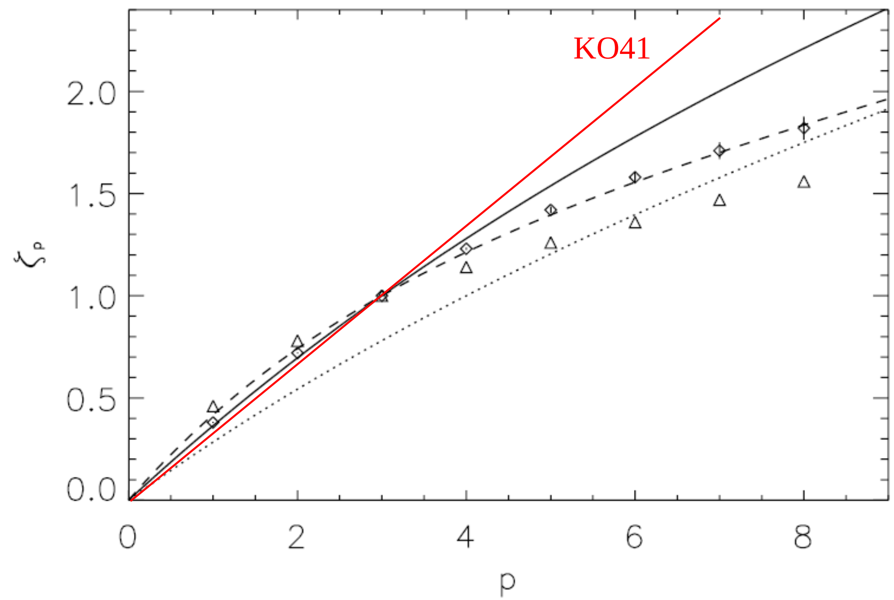


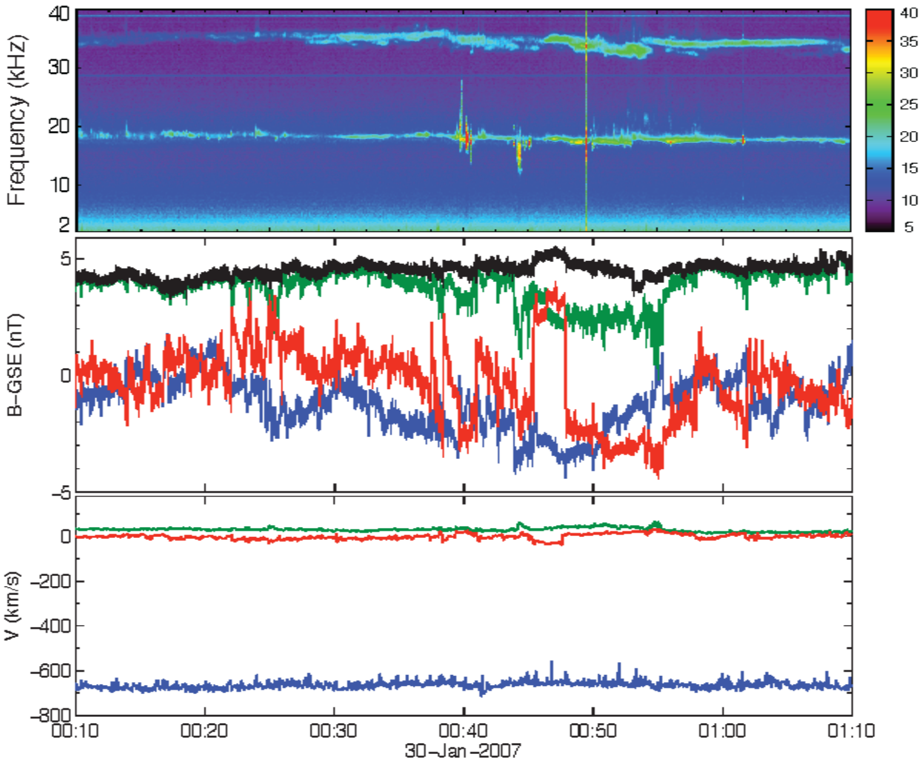
FIG. 4. Scaling exponents ζ_p^+ for 3D MHD turbulence (diamonds) and relative exponents ζ_p^+/ζ_3^+ for 2D MHD turbulence (triangles). The continuous curve is the She-Leveque model ζ_p^{SL}, the dashed curve the modified model ζ_p^{MHD} (7), and the dotted line the IK model ζ_p^{IK}.

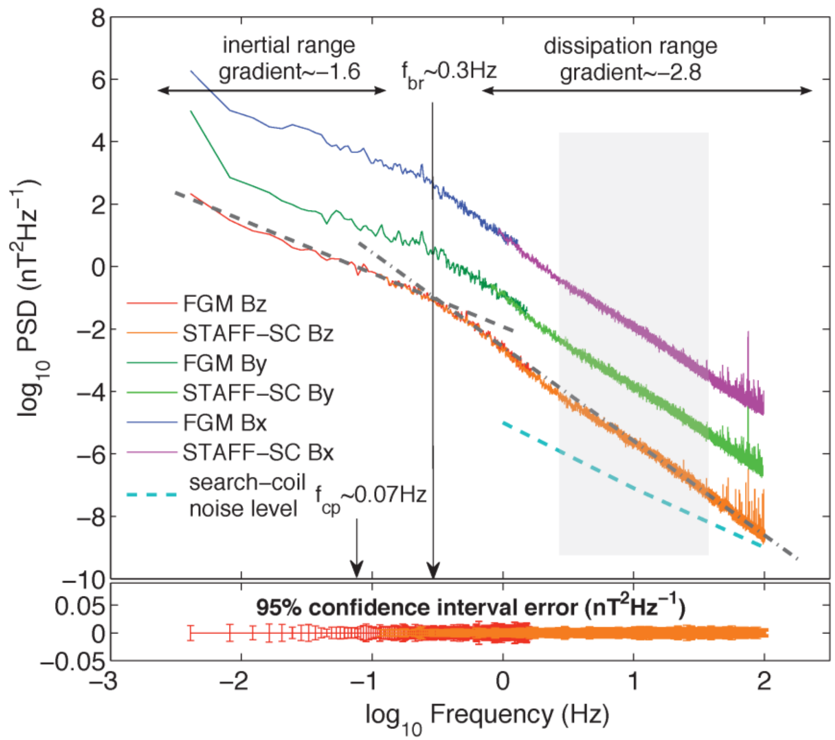
Fluid experiments,
Anselmet et al, PSS, 2001

2 and 3D MHD simulations
Muller & Biskamp PRL 2000

How large can we take p? *See eg Dudok De Wit, PRE, 2004*

A nice quiet fast interval of solar wind- CLUSTER high cadence B field spanning IR and dissipation range



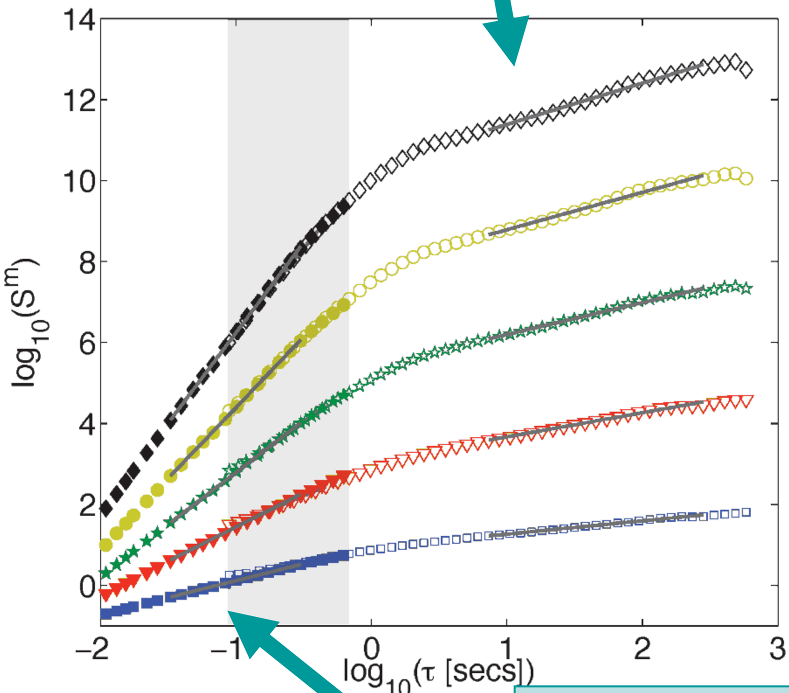


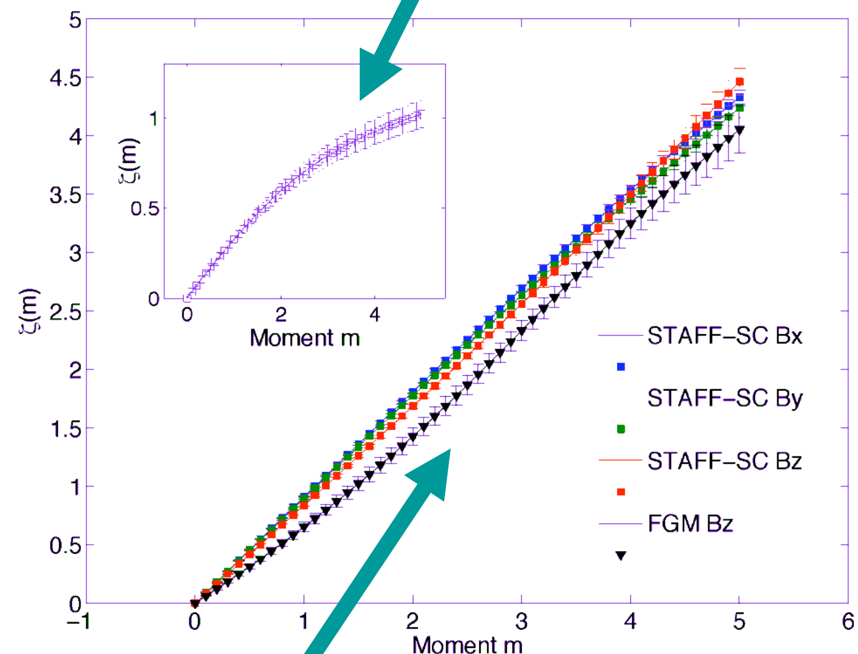
CLUSTER STAFF and FGM shown overlaid.

Kiyani, SCC et al PRL 2009

$$S_p = <|x(t+\tau) - x(t)|^p> \sim \tau^{\xi(p)}, \text{ plot } \log(S_p) \text{ vz. } \log(\tau) \text{ to obtain } \xi(p)$$

Inertial range- multifractal





Dissipation range- fractal

CLUSTER STAFF and FGM shown overlaid.

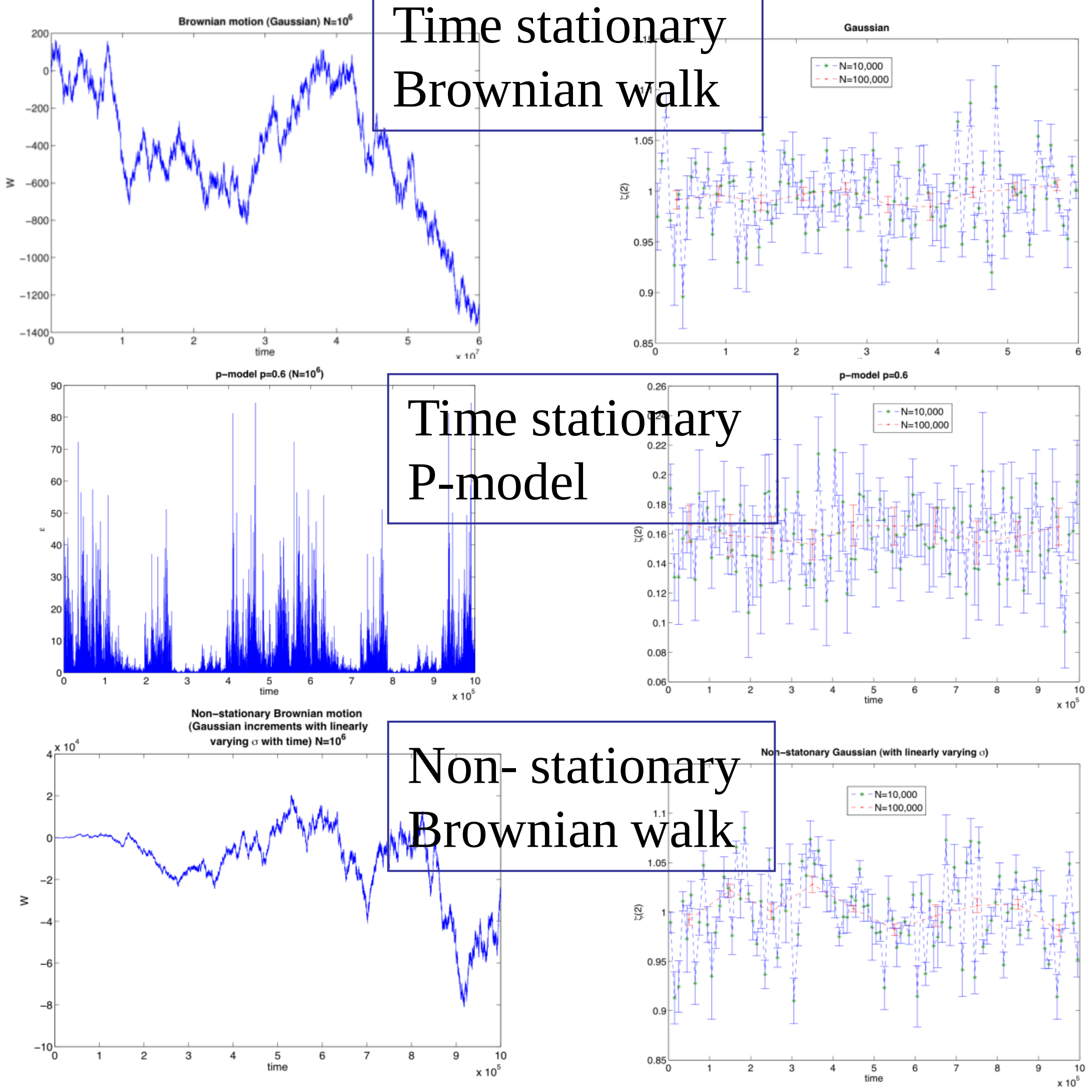
Kiyani, SCC et al PRL 2009,

Quantifying scaling III

Uncertainties, extreme events, finite size effects

Will discuss structure functions but remarks relate to other measures of scaling

Finite sample effect- error on exponent ζ(2) as a function of sample size N



Time stationary Brownian walk

Time stationary P-model

Non- stationary Brownian walk

Errors decrease in power law with N!

Kiyani, SCC et al, PRE (2009)

Structure functions-estimating the $\zeta(p)$ from data

Define structure function (generalized variogram) S_p for differenced timeseries:

$$y(t,\tau) = x(t+\tau) - x(t)$$

$$S_p(\tau) = <| y(t,\tau) |^p> \propto \tau^{\zeta(p)} \text{ if scaling}$$

We would like to calculate $S_p(\tau) = <| y(t,\tau) |^p> = \int_{-\infty}^{\infty} | y |^p \, P(y,\tau) dy$

then $S_p(\tau) = \tau^{\zeta(p)} \int_{-\infty}^{\infty} y_s^p P_s dy_s$

Conditioning- an estimate is:

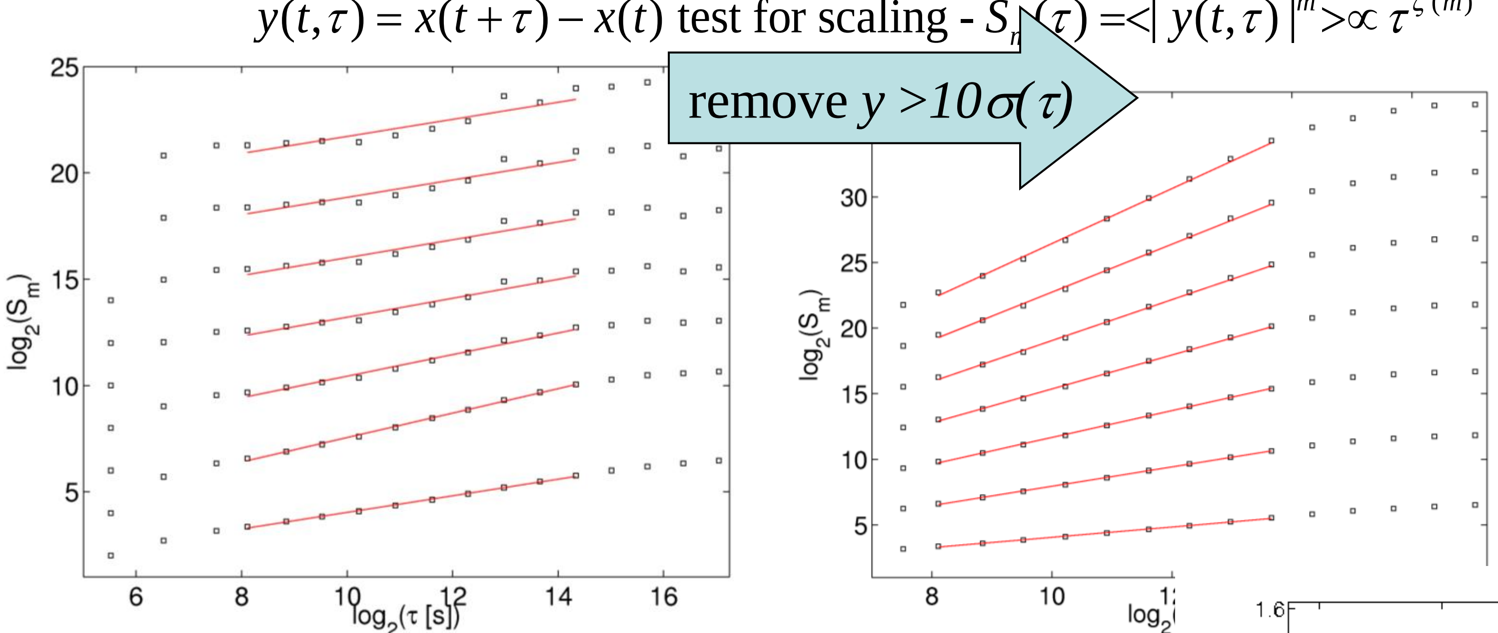
$<| y |^p> = \int_{-A}^{A} | y |^p \, P(y,\tau) dy$ where $A = [10-20]\sigma(\tau)$

strictly ok if selfsimilar: $y \to y_s \tau^{\alpha}, P \to P_s \tau^{-\alpha}$, $\zeta(p) = p\alpha$

if ξ(p) is quadratic in p (multifractal)- weaker estimate

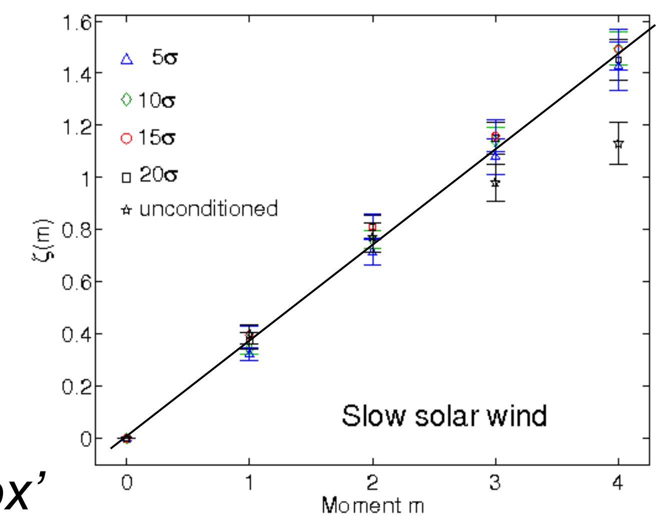
Structure functions- sensitive to undersampling of largest events (example - ρ in slow sw)

$y(t,\tau)=x(t+\tau)-x(t)$ test for scaling - $S_m(\tau)=\langle|y(t,\tau)|^m\rangle\propto\tau^{\zeta(m)}$



2 sources of uncertainty in exponent

1) Fitting error of lines (error bar estimates)
2) Outliers- Shown: removed < 1% of the data
   ACE 98-01 (4years)-10^6 samples.
   Threshold 450 km/sec.



fractal or multifractal?

fractal (self- affine) $\zeta(p)\sim\alpha p$

multifractal $\zeta(p)\sim\alpha p-\beta p^2+\ldots$

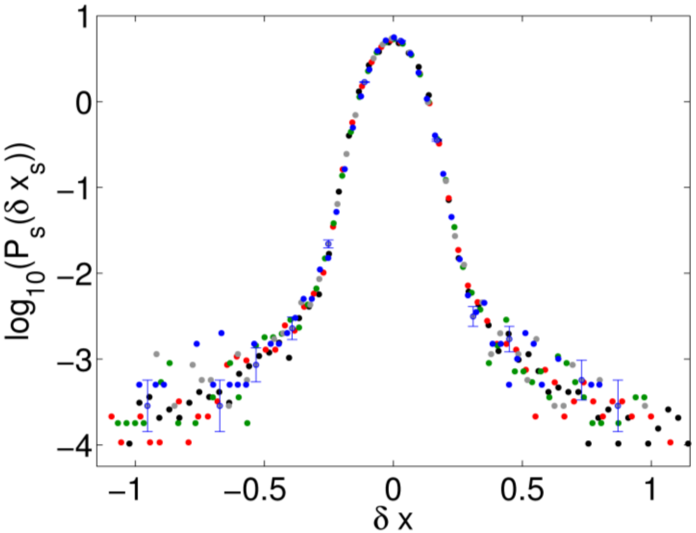
cf Fogedby et al PRE 'anomalous diffusion in a box'

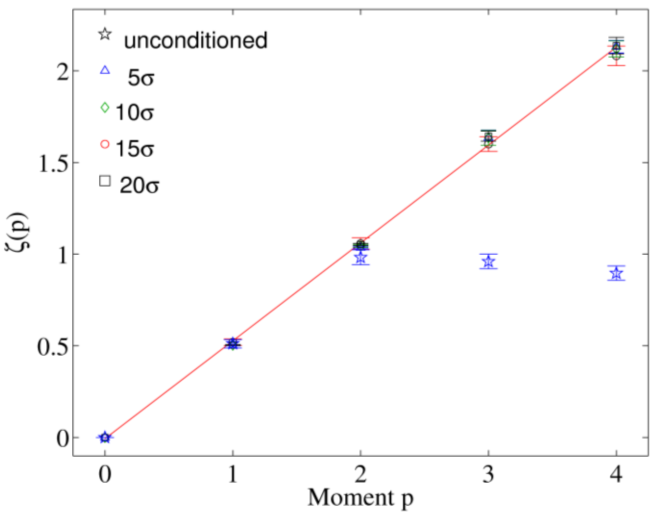
Outliers and a more precise test for fractality-example-Lévy flight ('fractal')

$P(x) \sim \frac{C}{x^{1+\mu}}, x \to \pm\infty, 1 < \mu < 2$ power law tails, self similar

for a finite length flight $(x - \langle x \rangle)^2 \sim t^{2/\mu}$

so $\mu = 2$ is Gaussian distributed, Brownian walk





PDF rescaling $x \to x_s \tau^{\alpha}, P \to P_s \tau^{-\alpha}$

Structure functions: $S_p(\tau) = \langle | x(t,\tau) |^p \rangle \propto \tau^{\zeta(p)}$

expect $\zeta(p) \sim \alpha p, \alpha = 1/\mu$

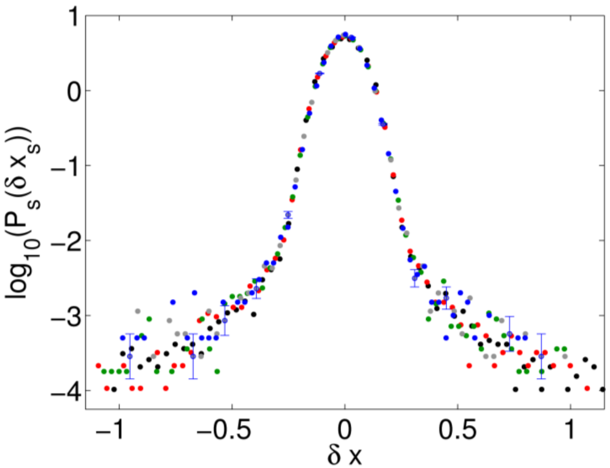
Chapman et al, NPG, 2005, Kiyani et al PRE, 2006

A more precise test for fractality-
outliers and convergence: example-Lèvy flight ('fractal')

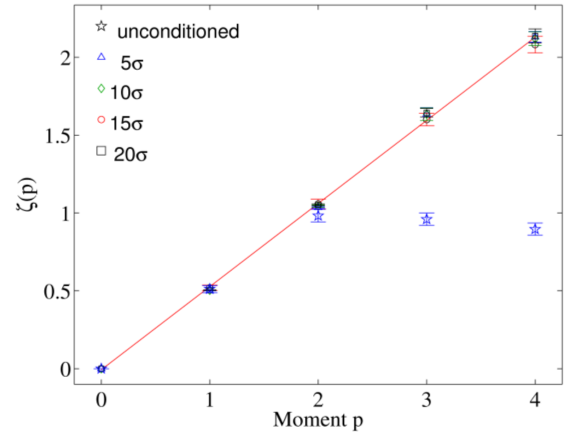
$P(x) \sim \frac{C}{x^{1+\mu}}, x \to \pm\infty, 1 < \mu < 2$ power law tails, self similar

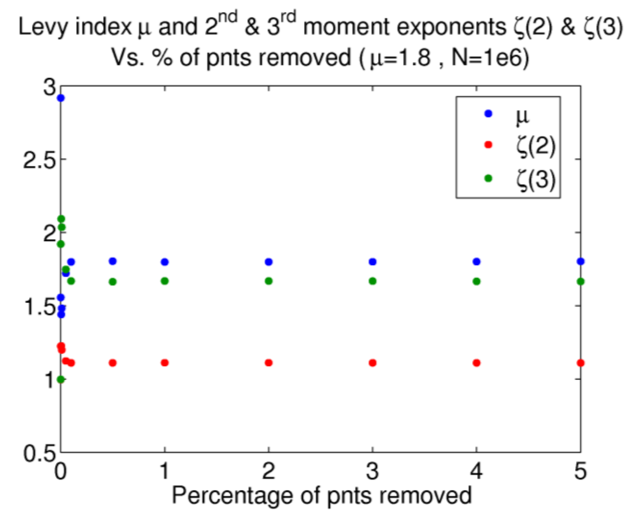
for a finite length flight $(x - \langle x \rangle)^2 \sim t^{2/\mu}$

so $\mu = 2$ is Gaussian distributed, Brownian walk



PDF rescaling $x \to x_s \tau^{\alpha}, P \to P_s \tau^{-\alpha}$





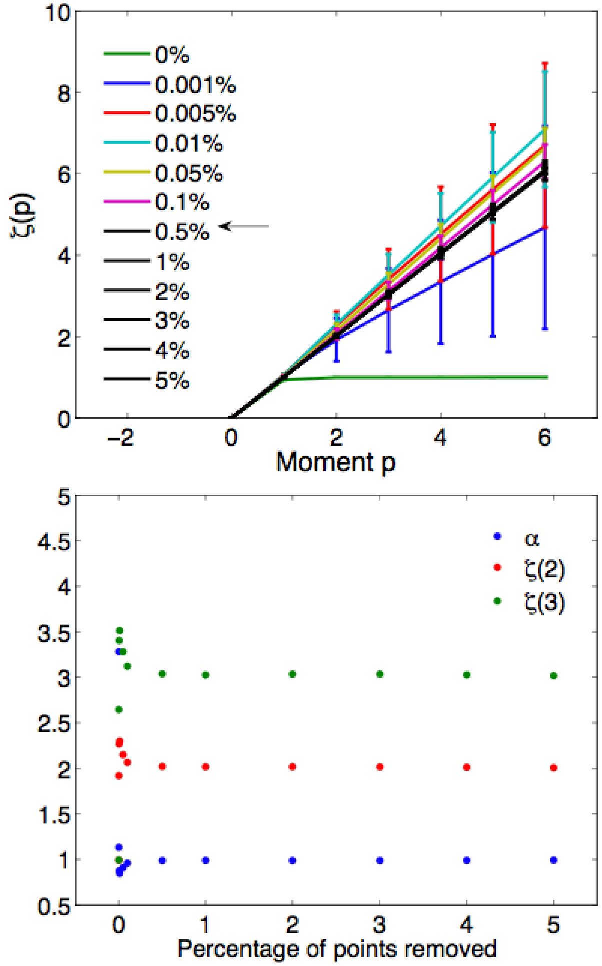
Structure functions: $S_p(\tau) = \langle |x(t,\tau)|^p \rangle \propto \tau^{\zeta(p)}$

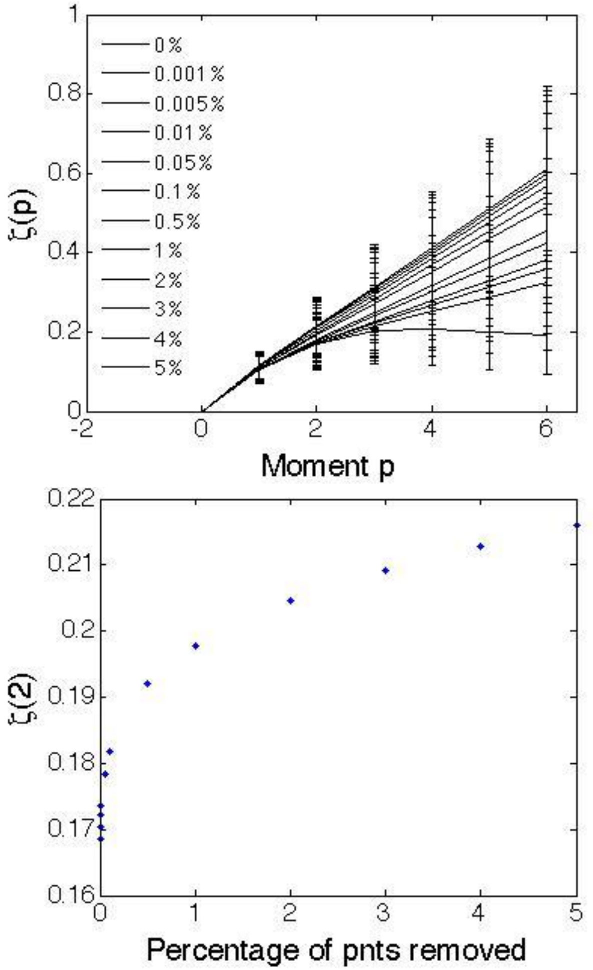
expect $\zeta(p) \sim \alpha p, \alpha = 1/\mu$

Chapman et al, NPG, 2005,Kiyani, SCC et al PRE (2006)

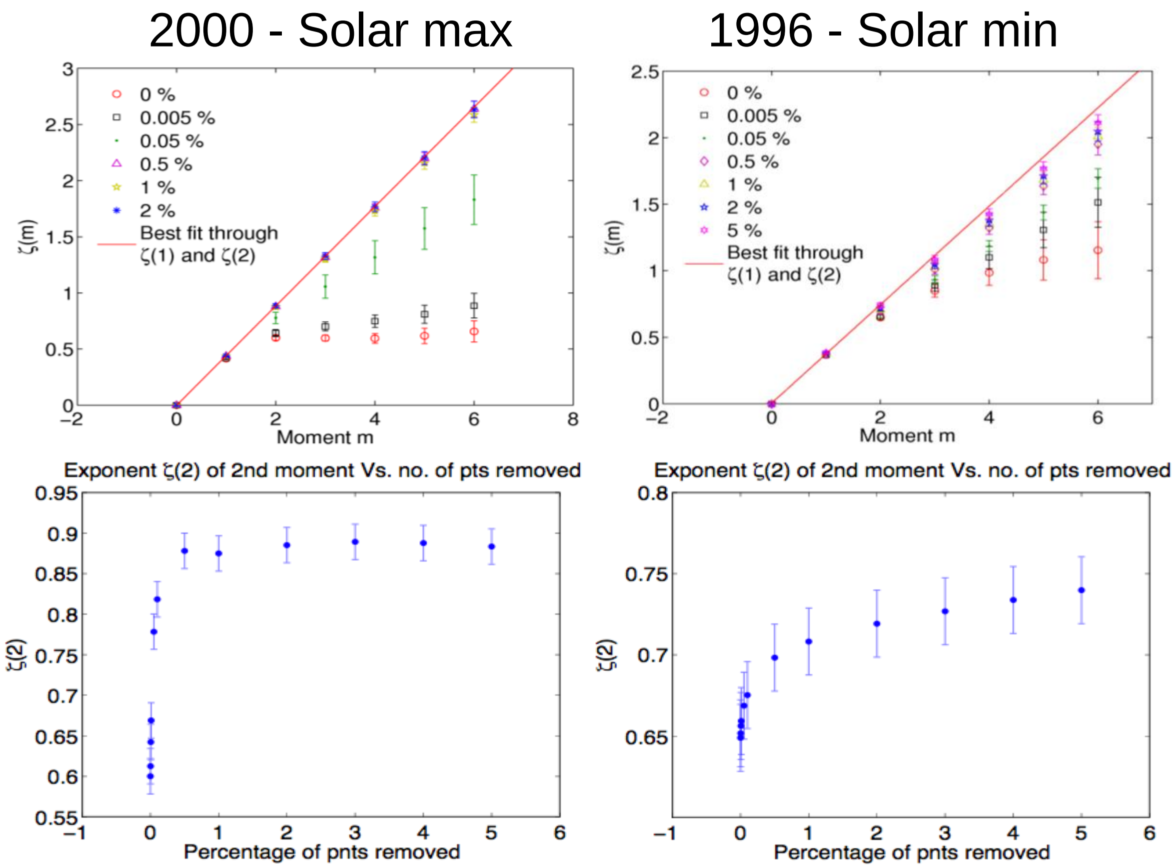
Distinguishing self- affinity (fractality) and multifractality

Lévy flight -fractal P-model -multifractal





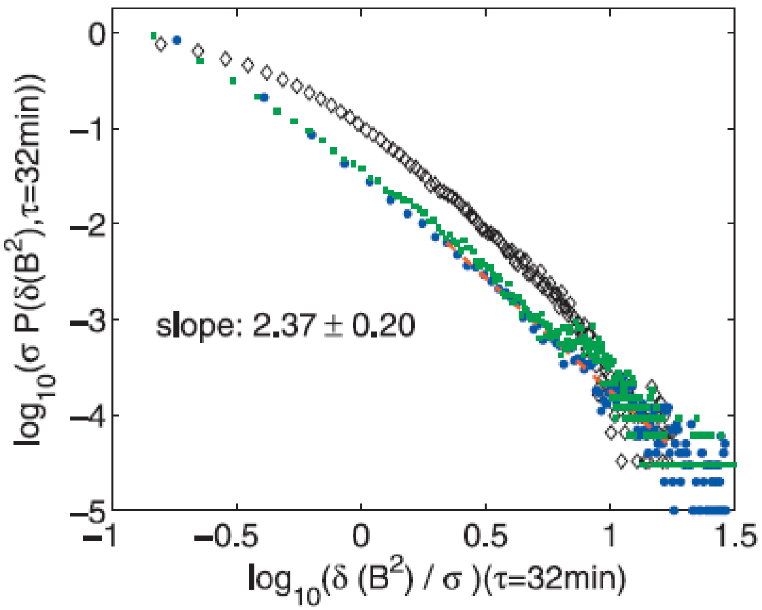
Solar cycle variation WIND Inertial Range- $|B|^2$

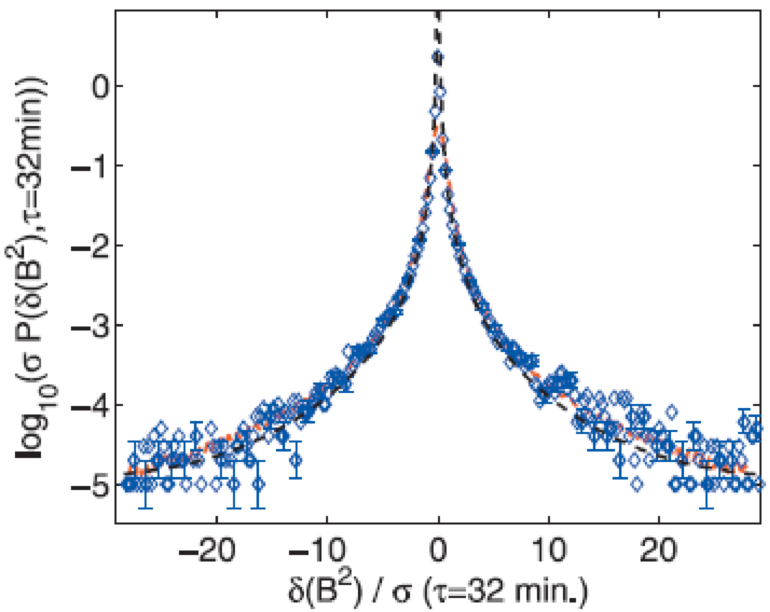


Kiyani et al, PRL 2007, Hnat et al, GRL 2007

Fractal signature 'embedded' in (multifractal) solar wind inertial range turbulence-coincident with complex coronal magnetic topology

Left: B² fluctuation PDF solar max and solar min
Right: solar max, FP and Lévy fit





WIND 1996 min (◊), 2000 max (◦), ACE 2000 max (□)

Hnat, SCC et al, GRL, (2007)

Quantifying scaling IV

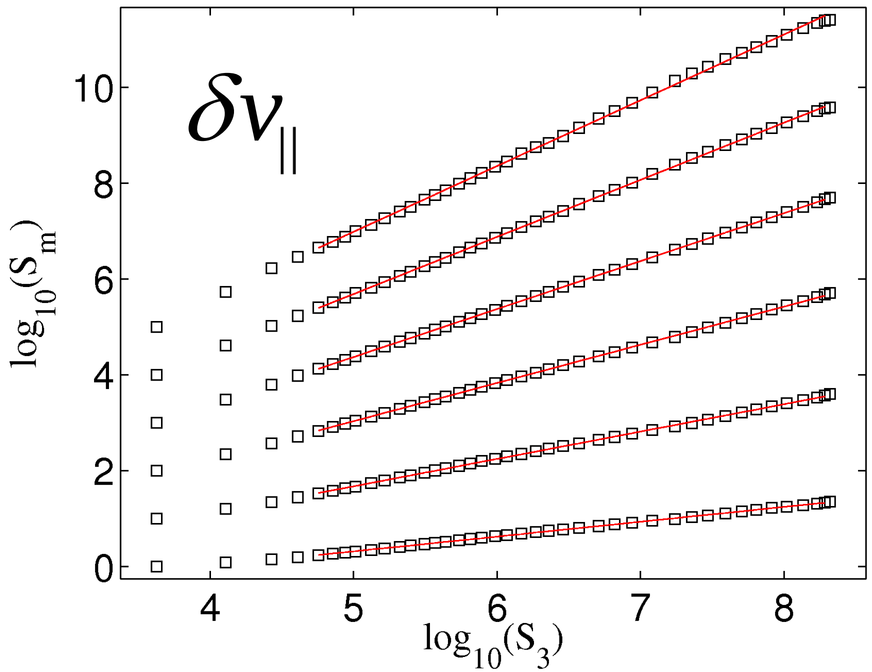
Extended self similarity

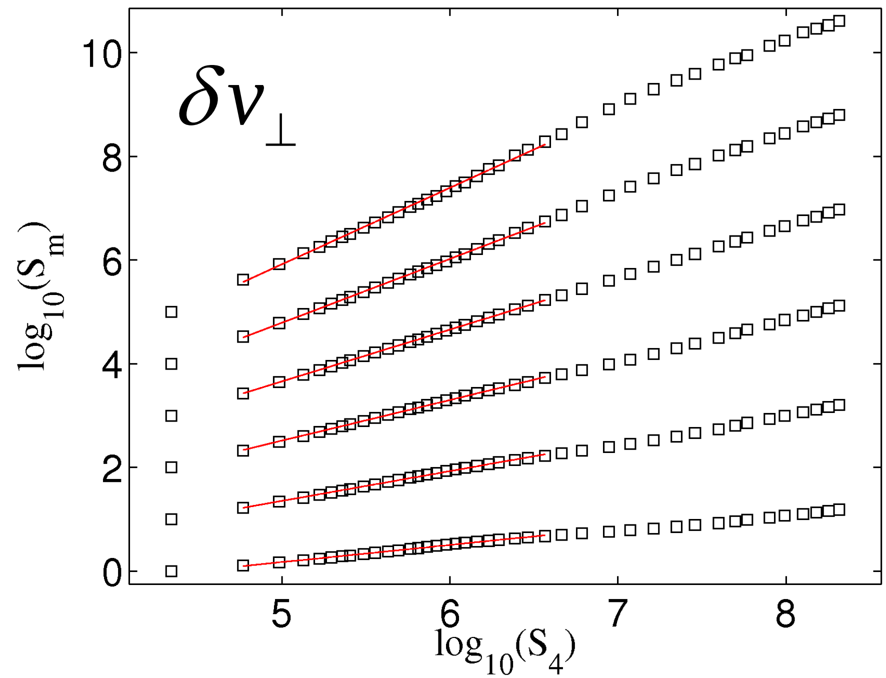
Generalized or extended self simlarity- ESS plots:

$S_p = \langle |\delta \mathbf{v} \bullet \hat{\mathbf{b}}|^p \rangle$ and its remainder versus S_3, S_4

ESS tests $S_p = S_q^{\zeta(p)/\zeta(q)}$ i.e. $S_p \sim G(\tau)^{\zeta(p)}$

gives exponents when e.g. $\zeta(3) \approx 1$ or $\zeta(4) \approx 1$





End

See the MPAGS web site for more reading…