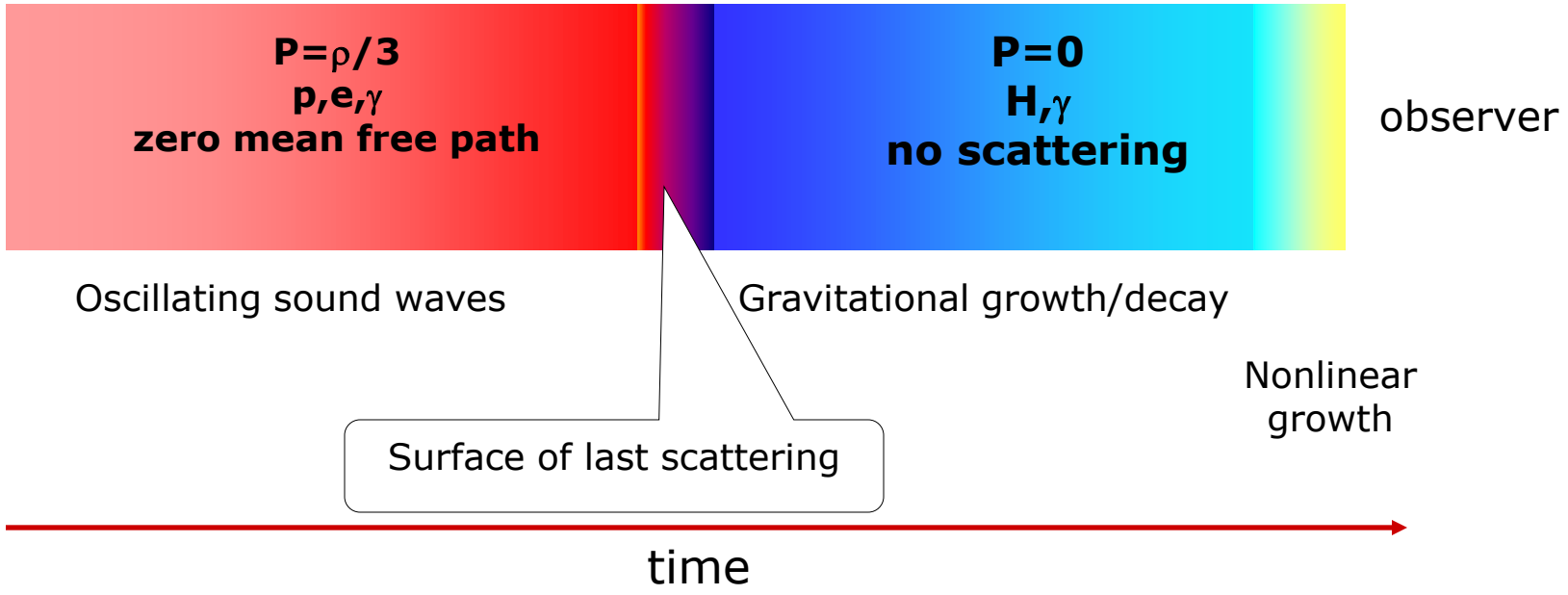
Redshift-Space Distortions of Baryon Acoustic Oscillations

Alex Szalay, JHU

The Growth of Fluctuations

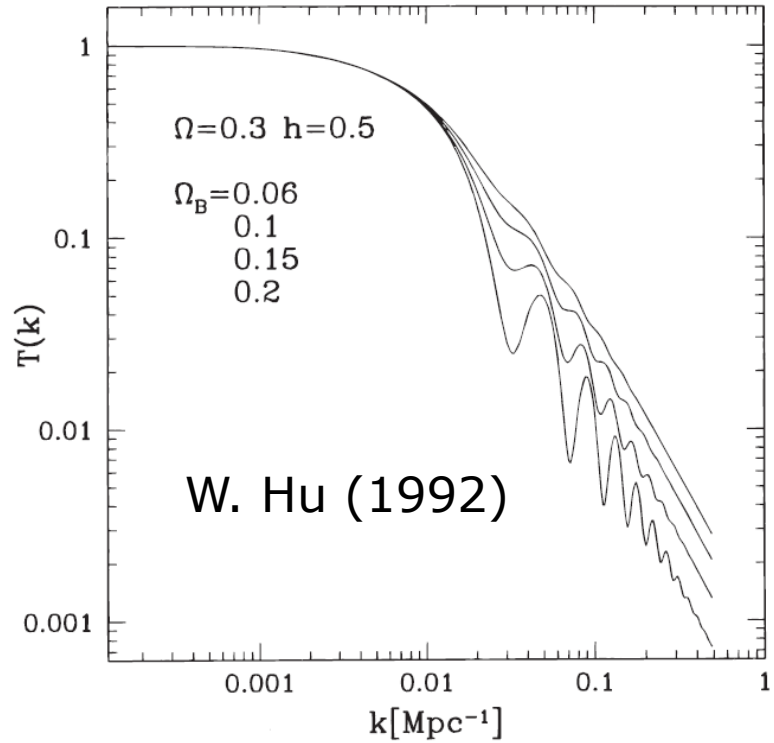
- Jeans (1902): gravitational instability in a P=0 fluid
- When the pressure is large, sound waves
- At around recombination:
  - Phase interference => modulation of fluctuations
  - Behaves like a resonant cavity



The Effect of Sound Waves

- The remnants of early sound waves show up as small bumps in the LSS P(k)
- They fully affect the CMB
- Wavelength given by the sound horizon (Sakharov)
- Amplitude of LSS bumps depends on

$$f_B = \Omega_B / \Omega_{tot}$$



- The position of the bumps determined from first principles: the speed of sound and the age of the universe at recombination
- An perfect ruler to measure cosmological parameters

How can we detect these?

Correlation Function

- Defined through the ensemble average

$$\xi(\mathbf{x}_1,\mathbf{x}_2)=\left\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\right\rangle$$

- Can be expressed through the Fourier transform

$$\xi(\mathbf{x}_1,\mathbf{x}_2)=\frac{1}{(2\pi)^6}\int d^3\mathbf{k}_1 d^3\mathbf{k}_2\; e^{i(\mathbf{k}_1\mathbf{x}_1-\mathbf{k}_2\mathbf{x}_2)}\left\langle \delta(\mathbf{k}_1)\delta^*(\mathbf{k}_2)\right\rangle$$

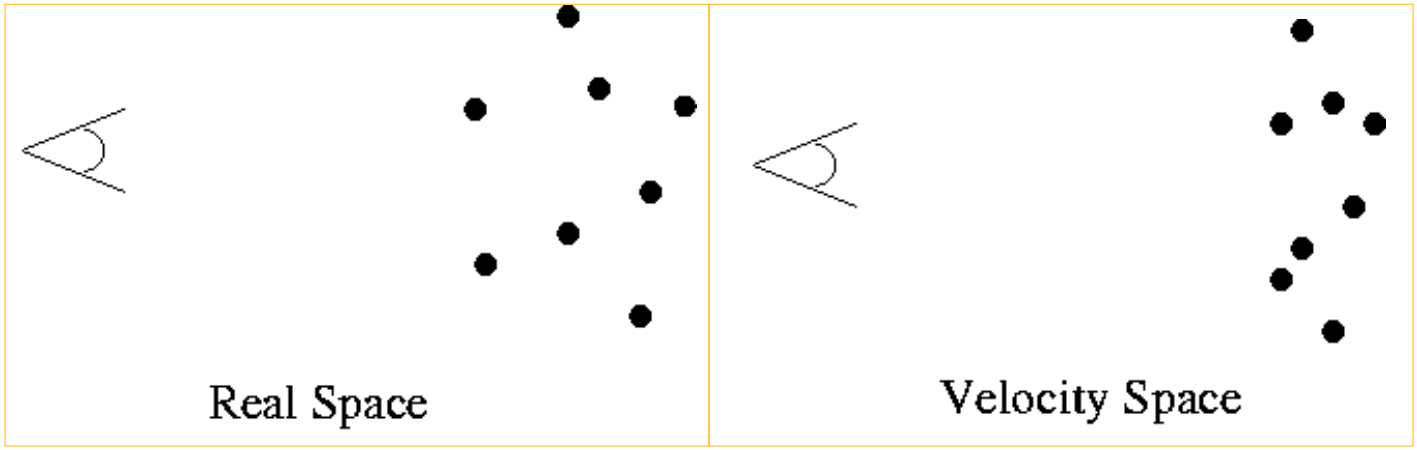
- The correlation function only depends on the distance

$$\xi(\mathbf{x}_1,\mathbf{x}_2)=\int d^3\mathbf{k}\; e^{i\mathbf{k}(\mathbf{x}_1-\mathbf{x}_2)}P(k)=\xi(|\,\mathbf{x}_1-\mathbf{x}_2\,|)=\xi(r)$$

- The correlation and the power spectrum form a Fourier transform pair

Distance from Redshift

- Redshift measured from Doppler shift
- Hubble's Law $v = H_0 r$
- Gives distance to zeroth order
- But, galaxies are not at rest in the comoving frame:
  - Distortions along the radial directions
  - Originally homogeneous isotropic random field, now anisotropic!

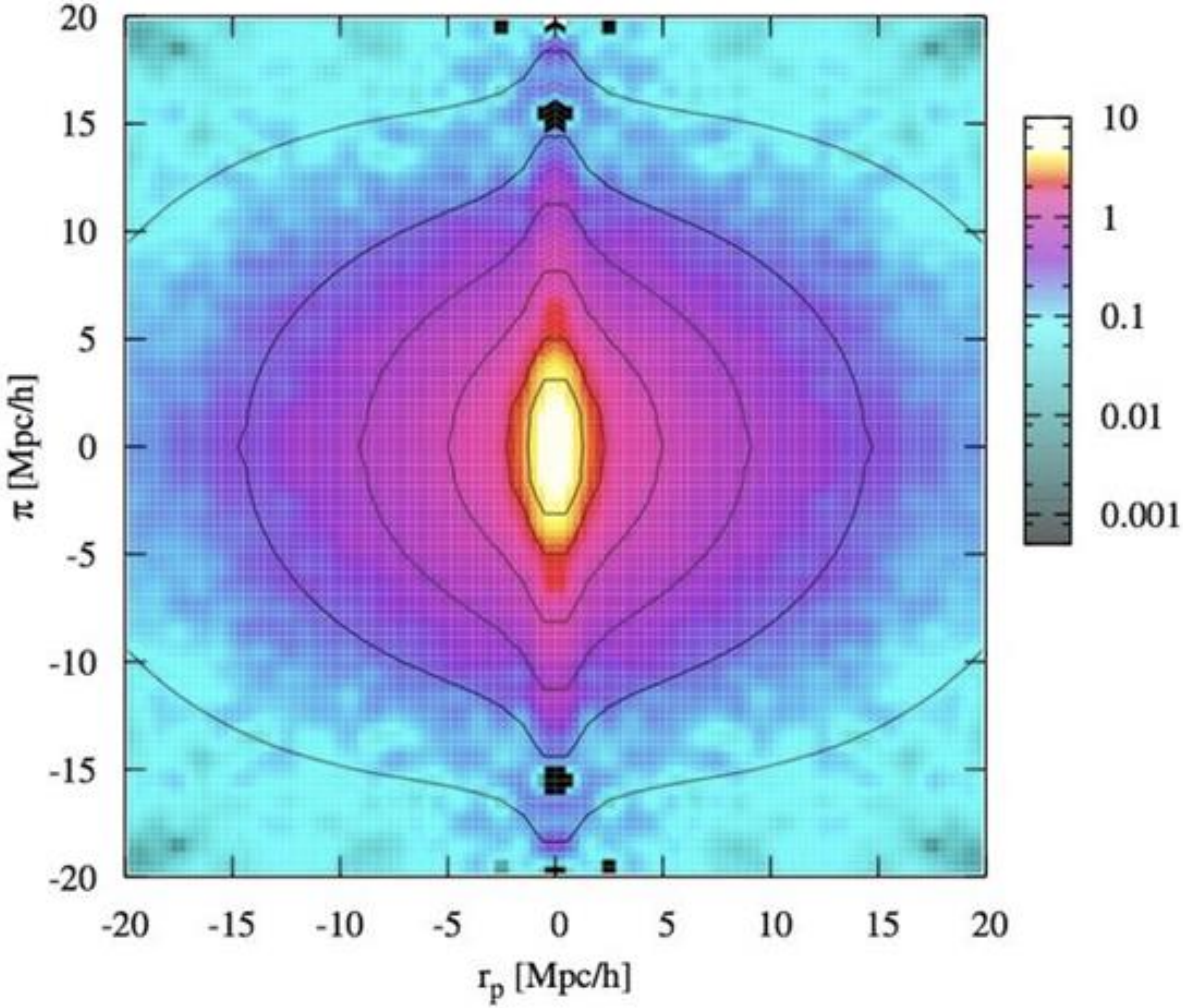


Redshift Space Distortions

Three different distortions

- Linear infall (large scales)
  - Flattening of the redshift space correlations
  - Kaiser 86
- Thermal motion (small scales)
  - 'Fingers of God'
  - Cuspy exponential $P(v_{12}) \propto e^{-|v_{12}|/\sigma}$
- Nonlinear infall (intermediate scales)

The 2dF Correlation Function



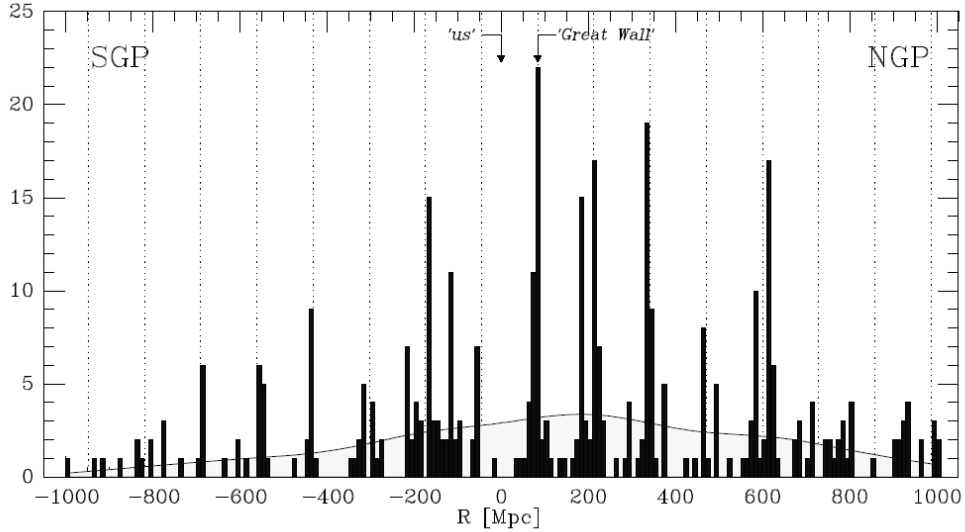
Redshift Space Power Spectrum

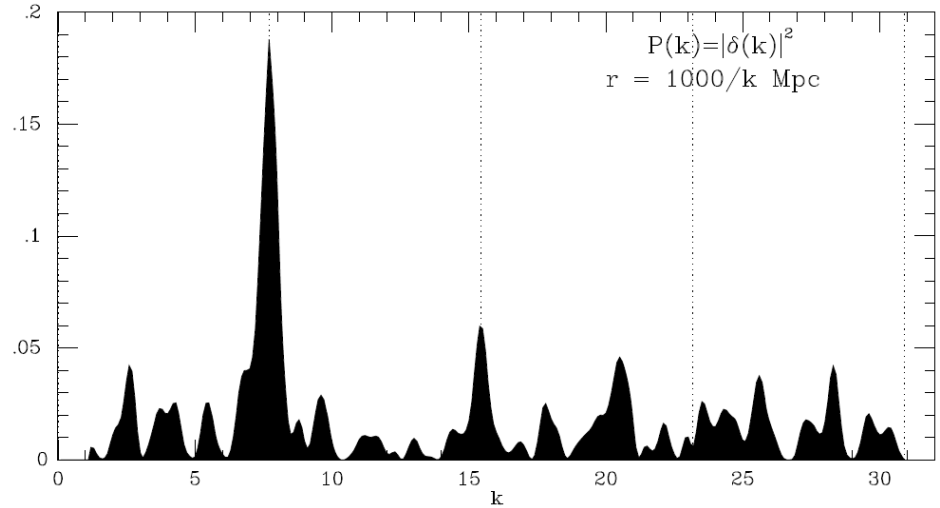
- Expressing the power spectrum

$$P_s(\mathbf{k}) = P_r(k)[1 + \beta\mu^2]^2,$$

  where μ is the angle between the line of sight and the wave vector
- $P(k)$ and $\xi(r)$ are duals:
  - P gets elongated
  - ξ gets compressed along the line of sight
- This is due to linear infall
- Fingers of God on the axis

BEKS 1990 – peaks at 128 h^{-1} Mpc





Broadhurst, Ellis, Koo and Szalay, Nature 1990

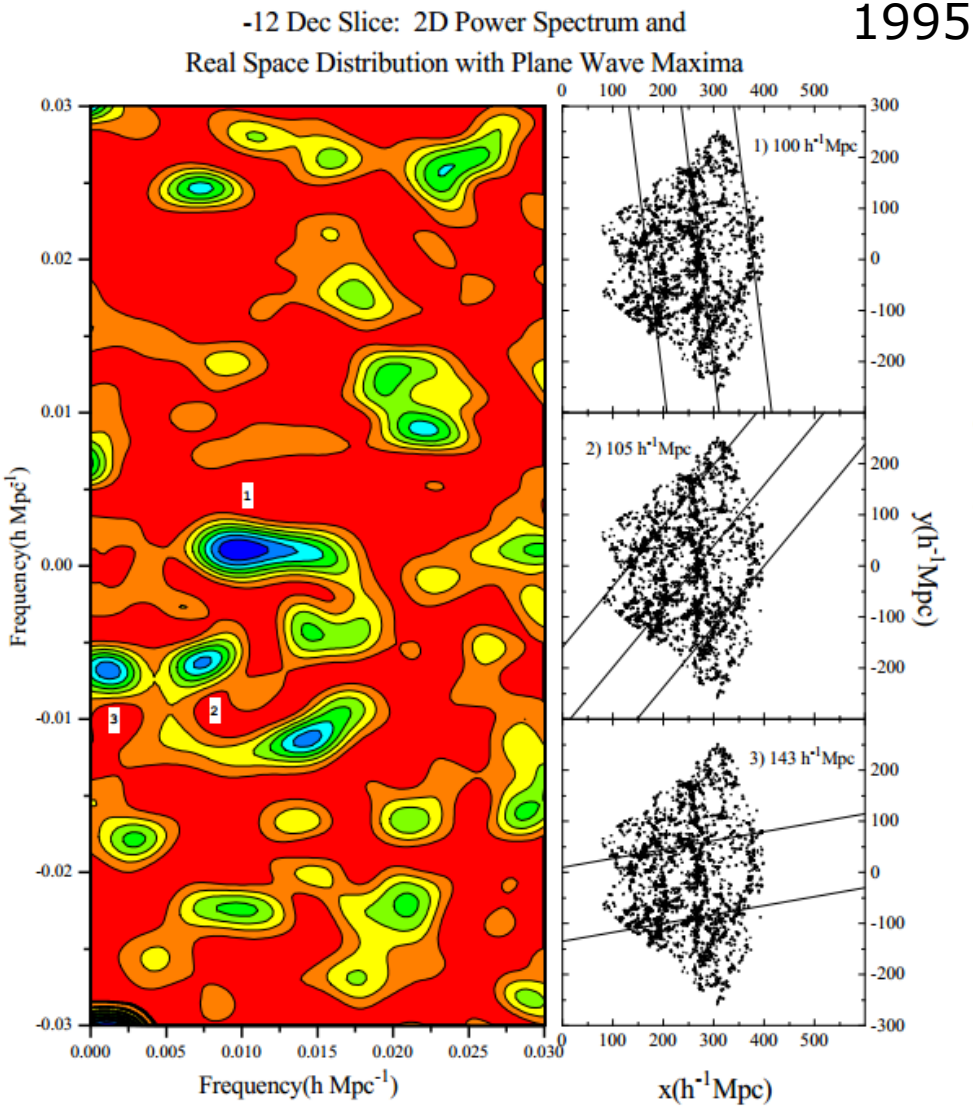
7 Summary

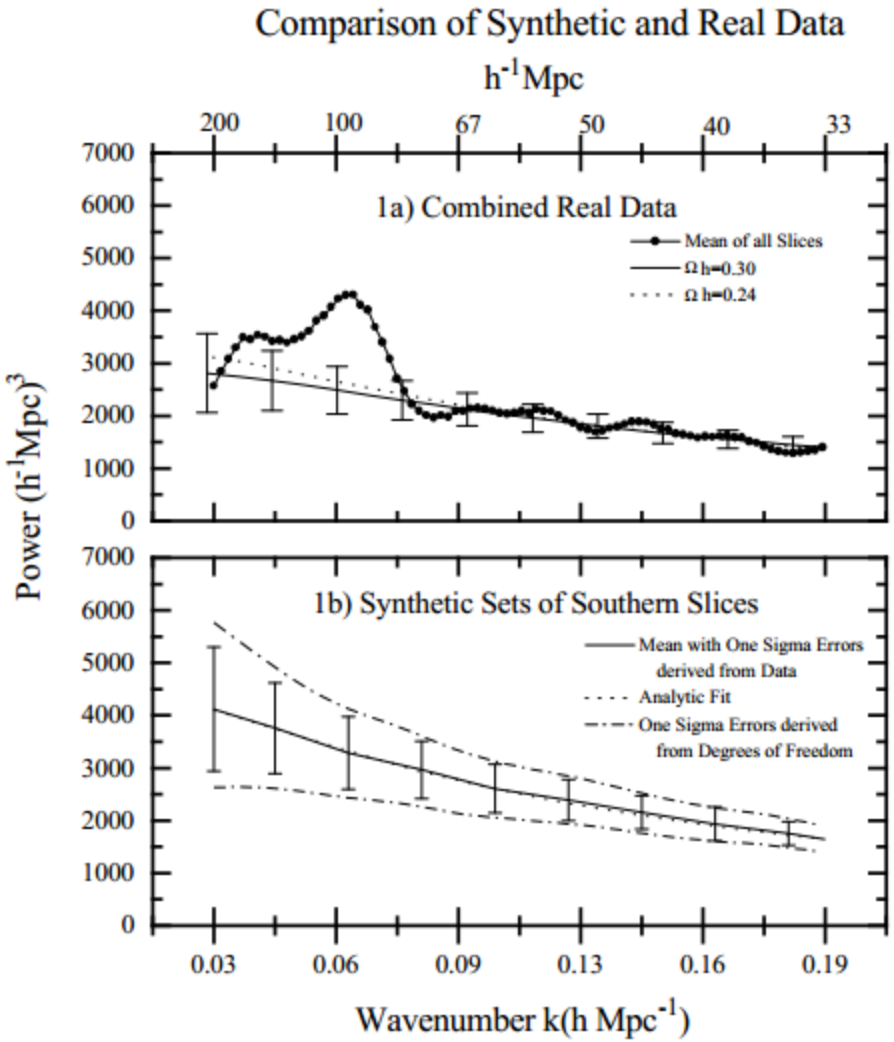
In summary, there are several observations pointing to excess power on 100-130 h^{-1} Mpc scales, which manifests itself in a small number of sharp spikes in Fourier space. These reflect the presence of walls and voids on similar scales. The emergence of this comoving scale at high redshift implies that this is imprinted on the fluctuations. Such a scale occurs naturally at recombination. The Sakharov oscillations, remnants of the sound waves at that epoch may provide an intriguing explanation. In such scenarios the baryon content of the Universe must be high, the Hubble constrant low, and the Universe open. This family of models deserves further investigation, and it is just barely possible that the 100 Mpc bumps may be the first preview of the elusive Doppler peaks — a fascinating preview of further connections between the galaxy distribution and the Cosmic Microwave Background.

AS, In Proc 18th Texas Symposium, 1996

Las Campanas Redshift Survey

Landy et al 1995



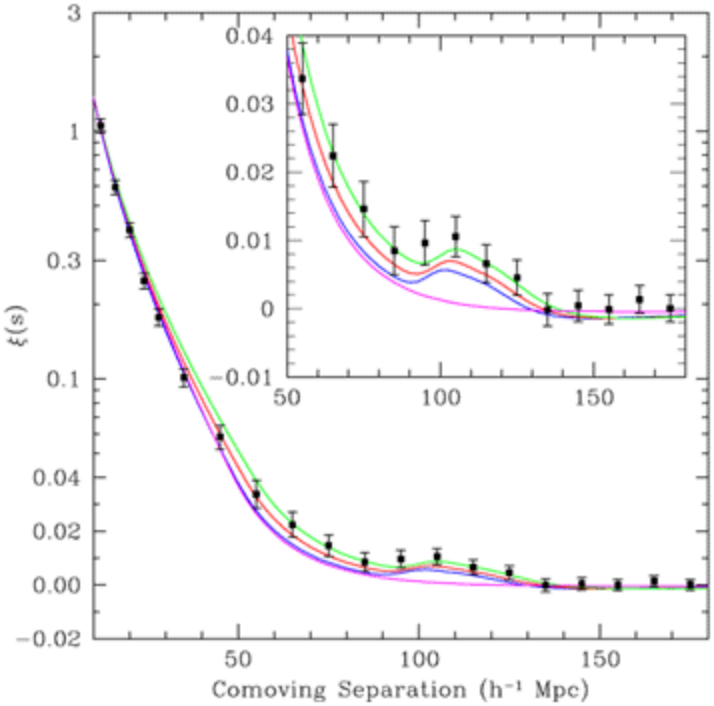


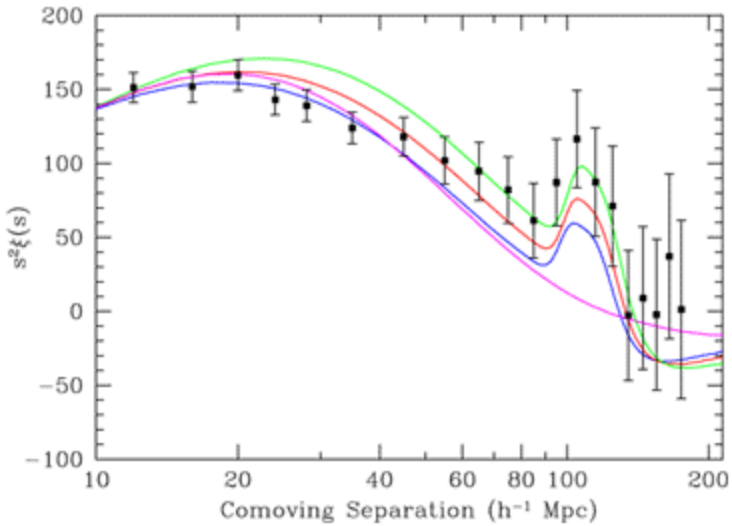
SDSS DR7 Redshift Samples

- Main Galaxies
  - 800K galaxies, high sampling density, but not too deep
- Luminous Red Galaxies
  - 100K galaxies, color and flux selected
  - $m_r < 19.5$, $0.15 < z < 0.45$, close to volume-limited
- Quasars
  - 20K QSOs, cover huge volume, but too sparse

Finding the BAO in DR4

- Eisenstein et al (2005) – LRG sample





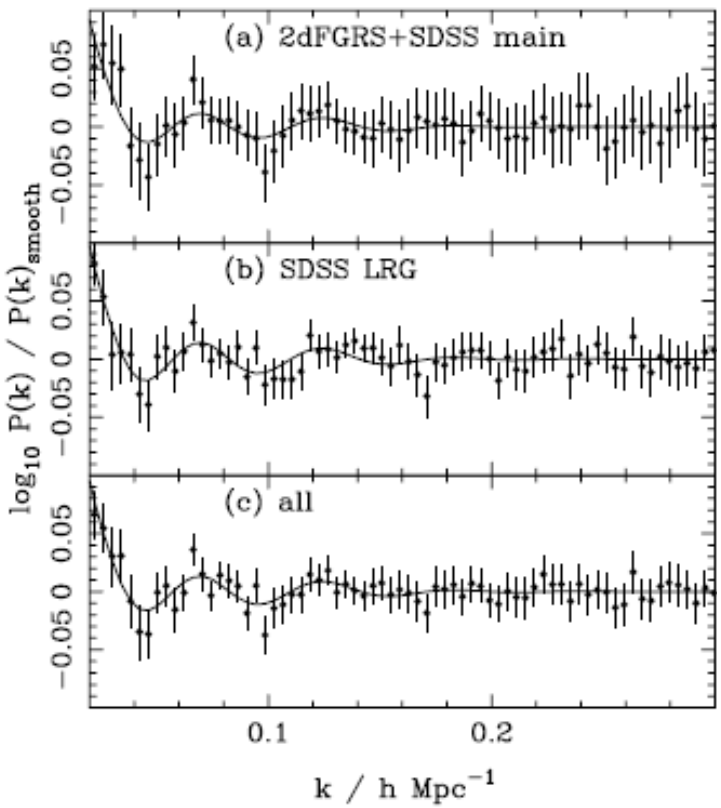
Correlation function

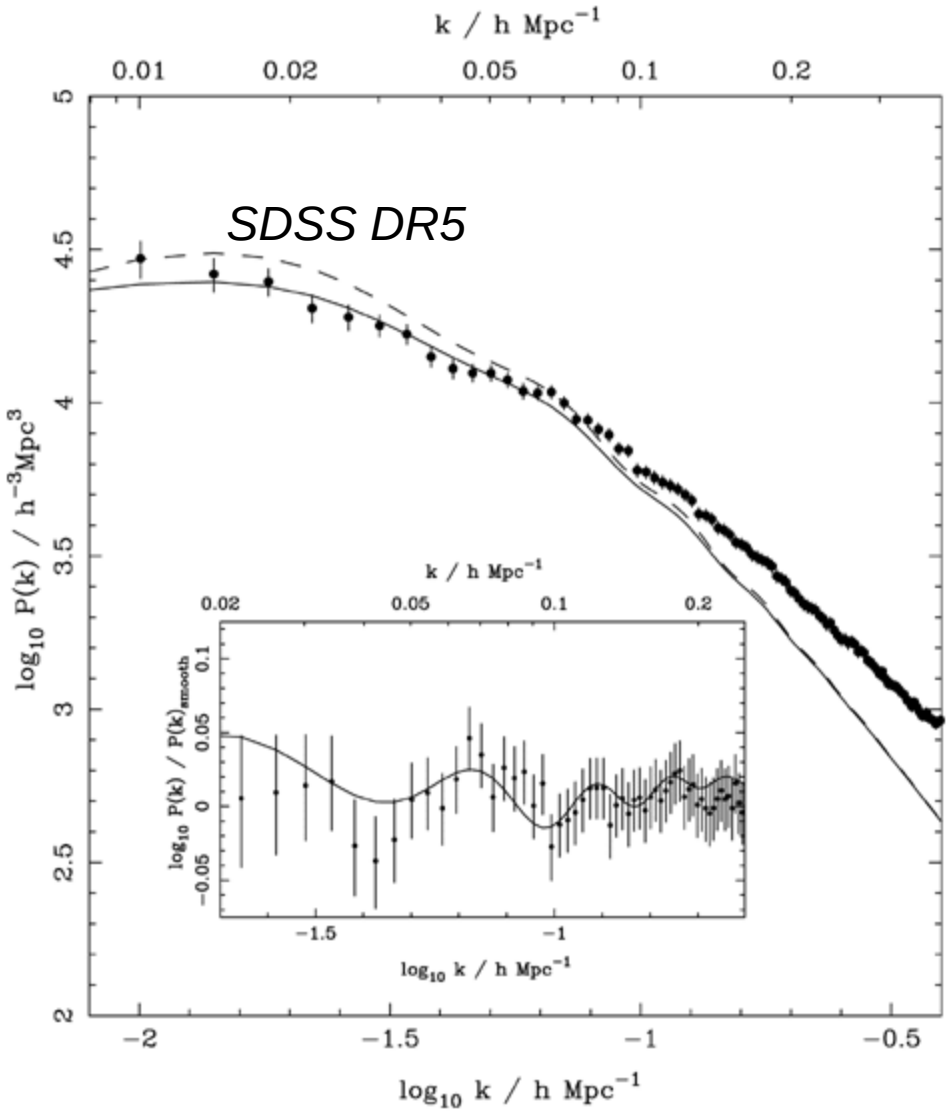
Primordial Sound Waves in SDSS

Power Spectrum

(Percival et al 2006, 2007)

SDSS DR6+2dF





800K galaxies

Redshift Space Correlation Fn

$$\xi^{(s)}(r,\gamma)=\left(1+\frac{2\beta}{3}+\frac{\beta^2}{5}\right)\xi_0(r)-\left(\frac{4\beta}{3}+\frac{4\beta^2}{7}\right)\xi_2(r)P_2(\cos\gamma)+\frac{8\beta^2}{35}\xi_4(r)P_4(\cos\gamma)$$

$$\xi_L^{(n)}(r)=\frac{1}{2\pi^2}\int dk\,k^2k^{-n}j_L(kr)P(k).$$

Szalay, Matsubara, Landy (1998)

- Azimuthal average wipes out the P_2 and P_4 terms, only the real space correlation function remains
- β terms contain a second derivative (Hamilton '92)

$$\frac{2\beta}{3}\left[j_0(kr)-2j_2(kr)\right]=-2\beta\left[j_0''(kr)\right]$$

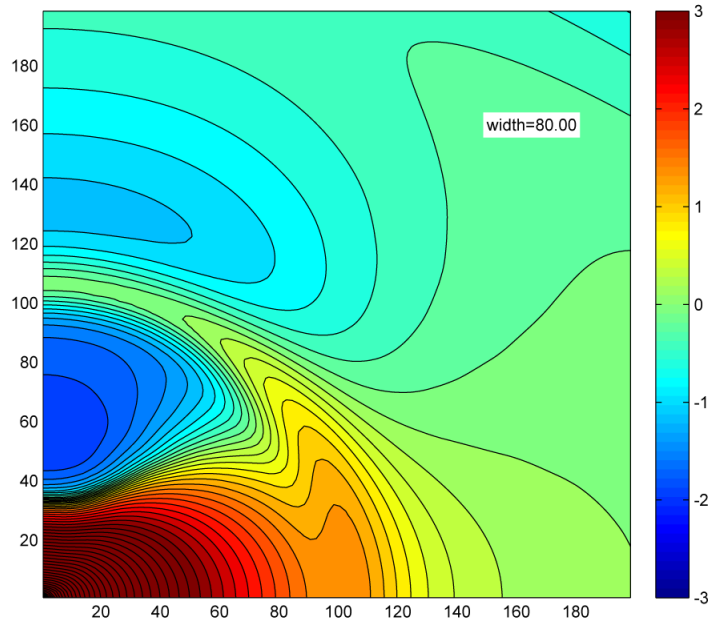
- Interesting effect on power spectra with sharp features: redshift-space distortions make features even sharper!
- Baryon Acoustic Oscillations !!!

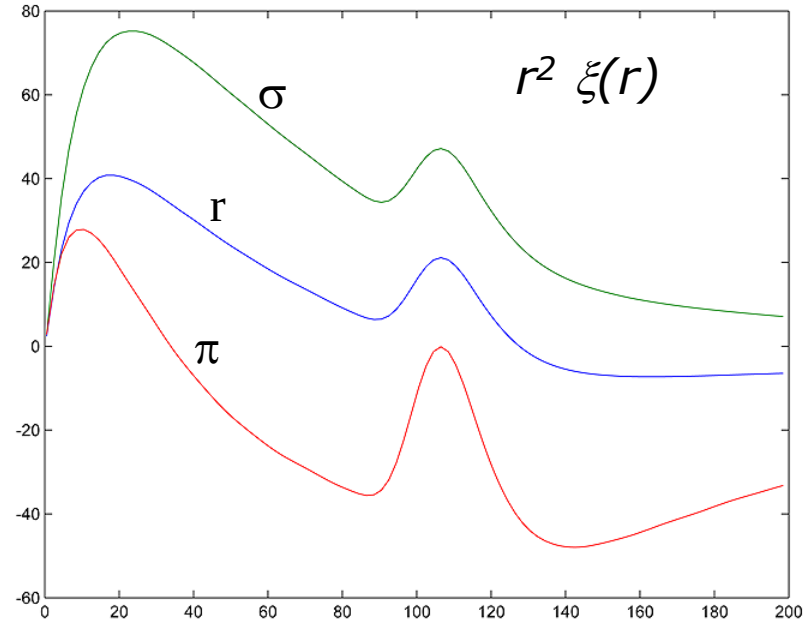
$\xi(r)$ from linear theory + BAO

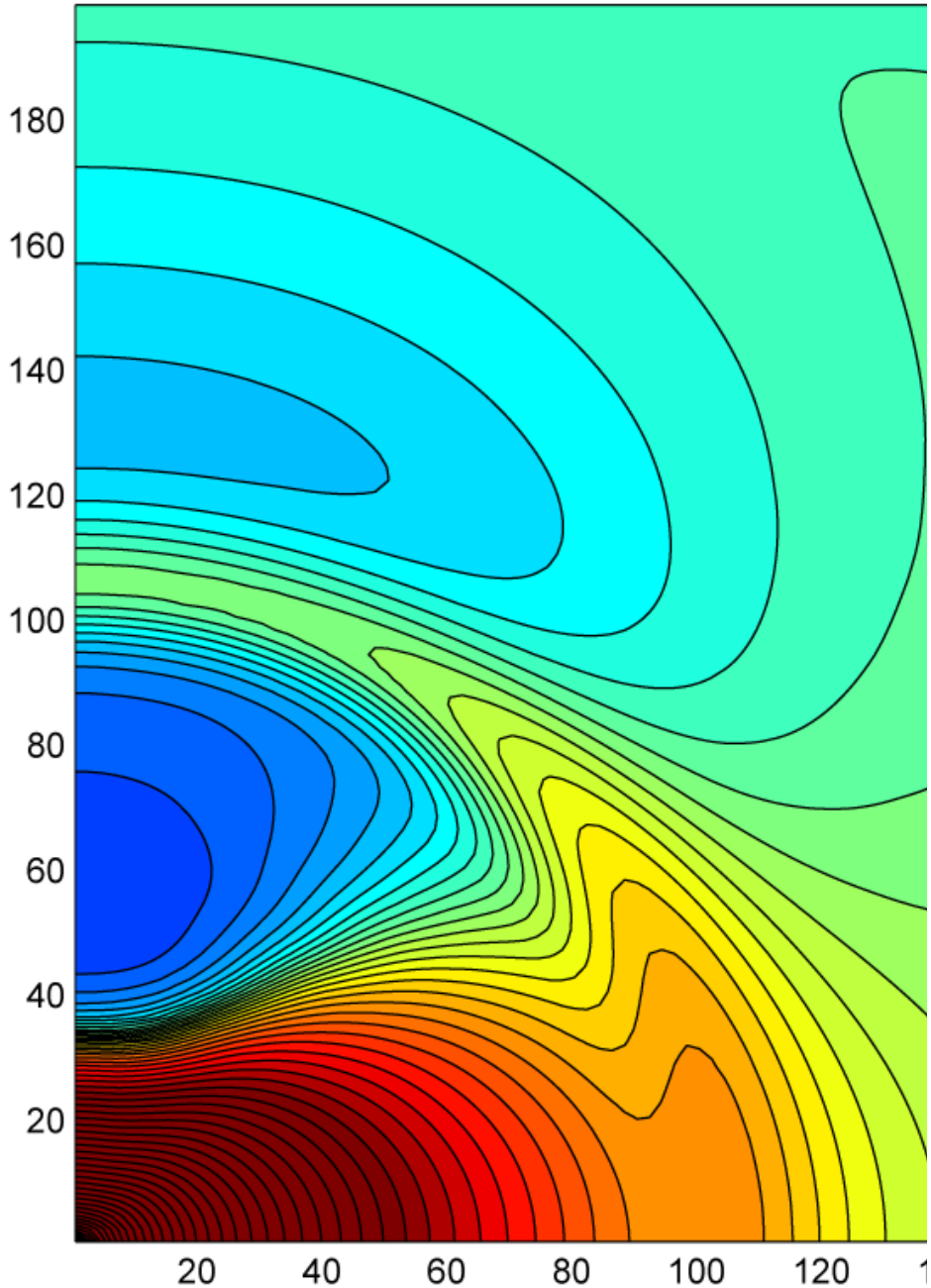
- Mixing of ξ_0, ξ_2 and ξ_4
  - Along the line of sight

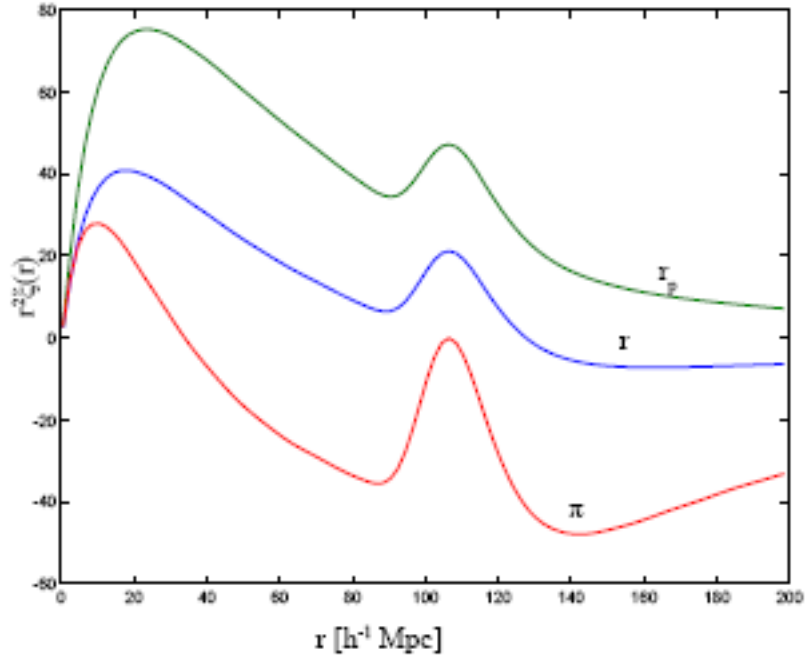
$$\xi_n(r) = \frac{1}{2\pi^2}\int_0^\infty dk\, k^2 j_n(kr) P(k)$$

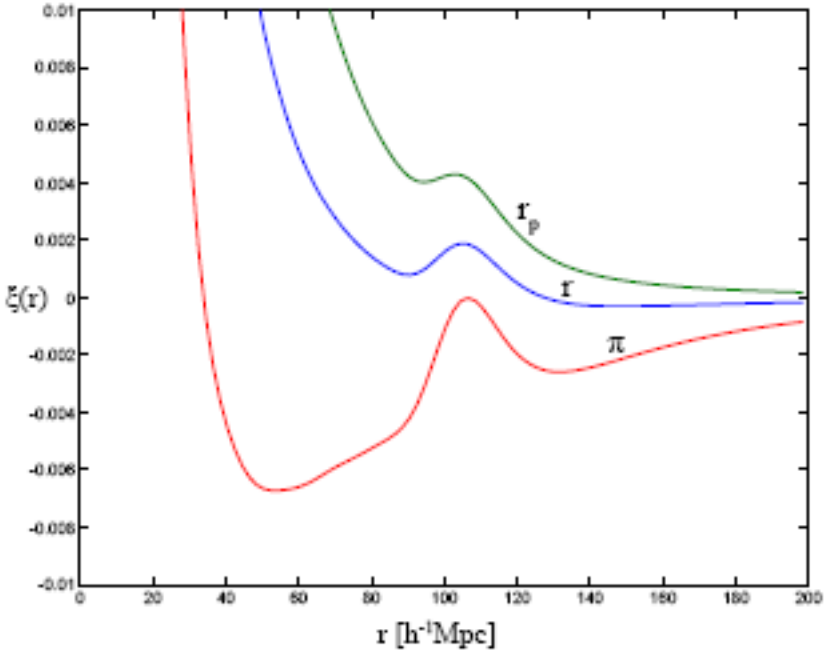
$$\xi^{(s)}(r) = \left(1 + \frac{2\beta}{3} + \frac{\beta^2}{5}\right)\xi_0(r) - \left(\frac{4\beta}{3} + \frac{4\beta^2}{7}\right)\xi_2(r) + \frac{8\beta^2}{35}\xi_4(r)$$






180
160
140
120
100
80
60
40
20
20
40
60
80
100
120


$r^2\xi(r)$
r_p
r
π
r [h^{-1} Mpc]


ξ(r)
r_p
r
π
r [h^{-1}Mpc]

2D Symmetry

- There is a planar symmetry:
  - Observer + 2 galaxies
- Thus 2D correlation of a slice is the same
- We usually average over $d\cos\theta = \sin\theta\, d\theta$

$$\left(1+\frac{2\beta}{3}+\frac{\beta^2}{5}\right)\xi_0(r)$$

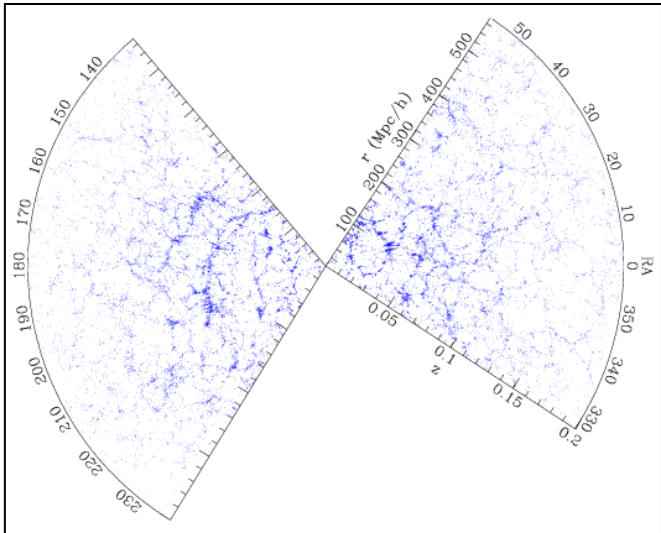
- Very little weight along the π axis:

$$\langle \xi(r,\theta) \rangle_\theta = \int_0^{\pi/2} d\theta \sin\theta\, \xi(r,\theta)$$

Sharp features at the line-of-sight go away with averaging!

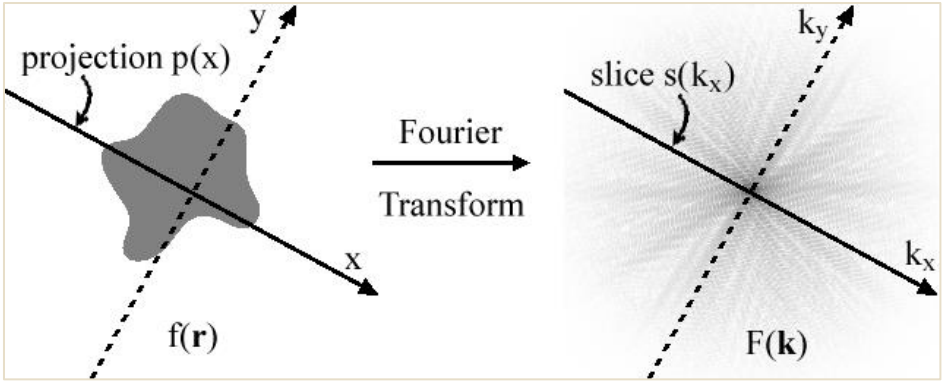
Lower Dimensional Subsets

- For a homogeneous isotropic process, the correlation function in a lower dimensional subset is identical
- Statement even true if the field is not isotropic, as long as subsets obey the same symmetry
- Projection-Slicing theorem



Projection and Slicing Theorem

$$F_m P_m = S_m F_N$$



$$p(x) = \int_{-\infty}^{\infty} f(x,y)\, dy.$$

$$F(k_x, k_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, e^{-2\pi i(xk_x + yk_y)}\, dxdy.$$

$$s(k_x) = F(k_x, 0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, e^{-2\pi i x k_x}\, dxdy$$

$$= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x,y)\, dy\right] e^{-2\pi i x k_x} dx$$

$$= \int_{-\infty}^{\infty} p(x)\, e^{-2\pi i x k_x} dx$$

The basis of CAT-SCAN / Radon transform

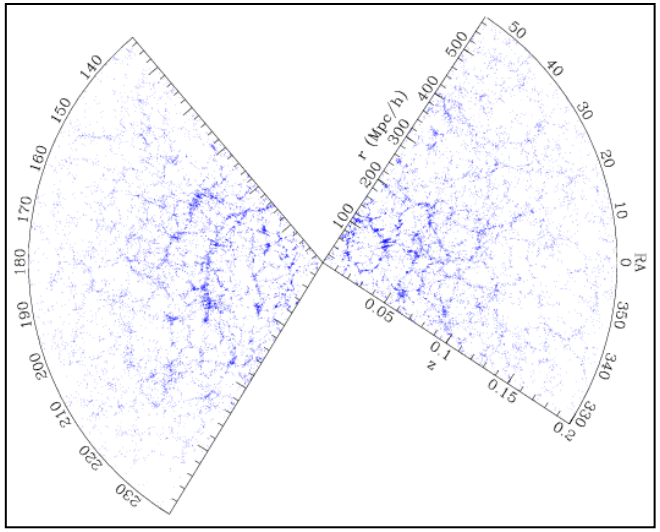
Slices and Pencilbeams

Slicing along a plane

- The isotropic correlation function is the same
- Corresponding 2D power spectrum is given by the projection onto the plane of the slice

Slicing along pencilbeams

- 1D power spectrum is the projection onto the axis of the pencilbeam



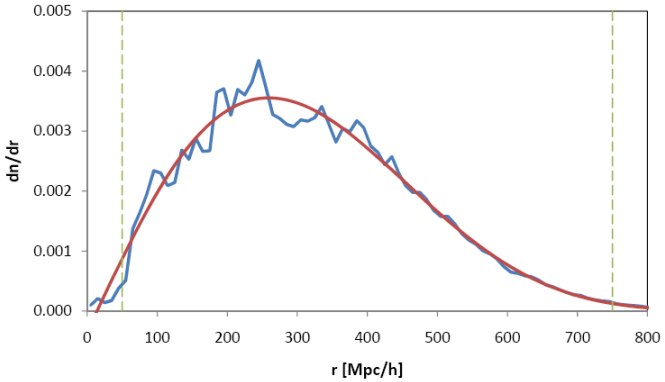
Tomography of SDSS

- SDSS DR7 Main Galaxy Sample
  - Limit distances to $100<r<750\ h^{-1}$ Mpc
- Compute 3D correlation function
- Also cut 3D data into thin angular slices
  - Project down to plane (only 2D info)
  - Different widths (2.5, 5, 10 deg)
  - Rotate slicing direction by 15 degrees
  - Ensemble of 661 slices to work with
- Analyze 2D correlation functions $\xi(\pi,\sigma)$

(Tian, Budavari, Neyrinck and AS 2011)

SDSS Sample

- SDSS DR7 MGS, Stripes 9 through 37, Northern Cap only
- 0.001<z<0.18, zConf>0.9, zErr<0.1
- Remove all the objects in the incomplete areas
- 527,362 objects
- 17M random galaxies
- Slices:

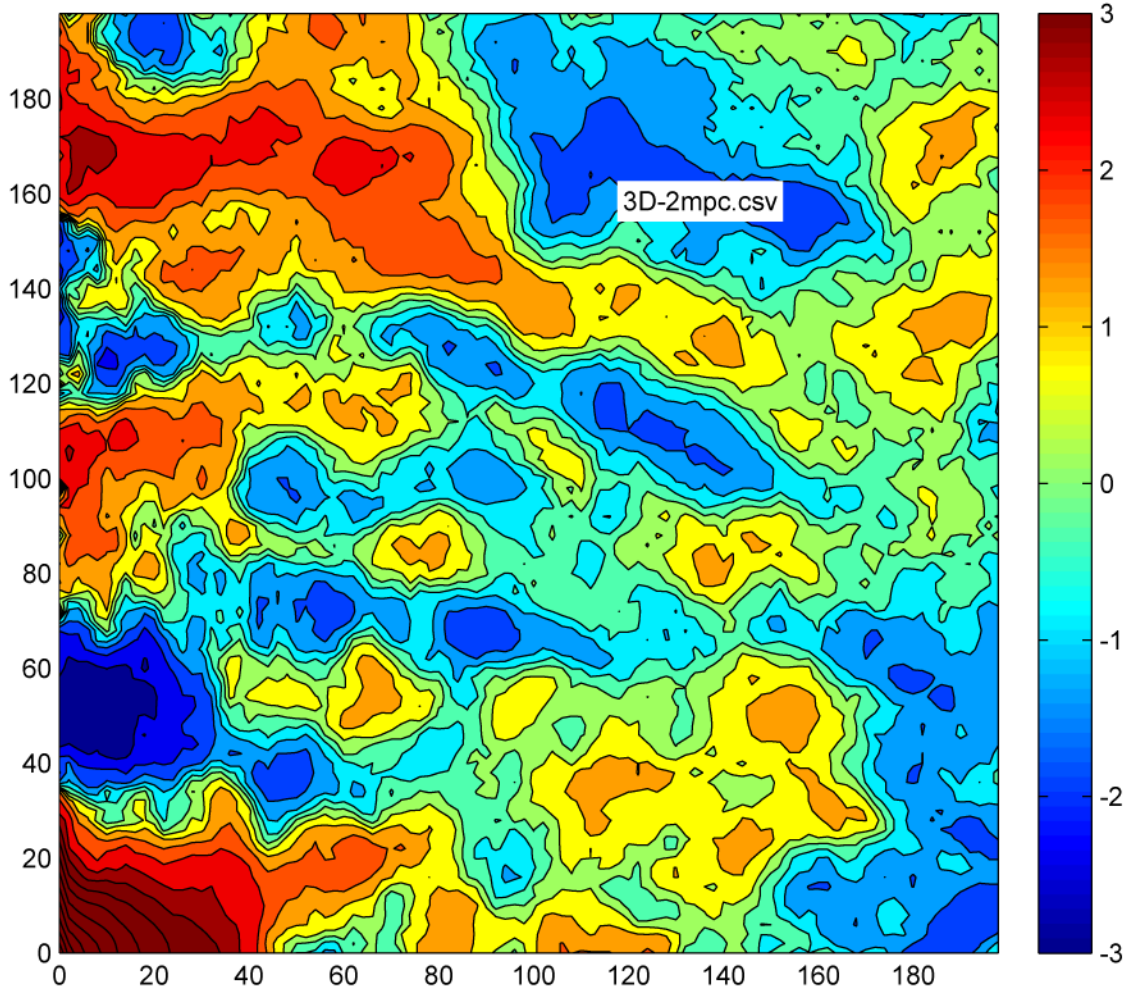


From 0 to 165°, 15° increments, 12 angular orientations, 2.5° thickness, 20°<width<80°, 661 slices total

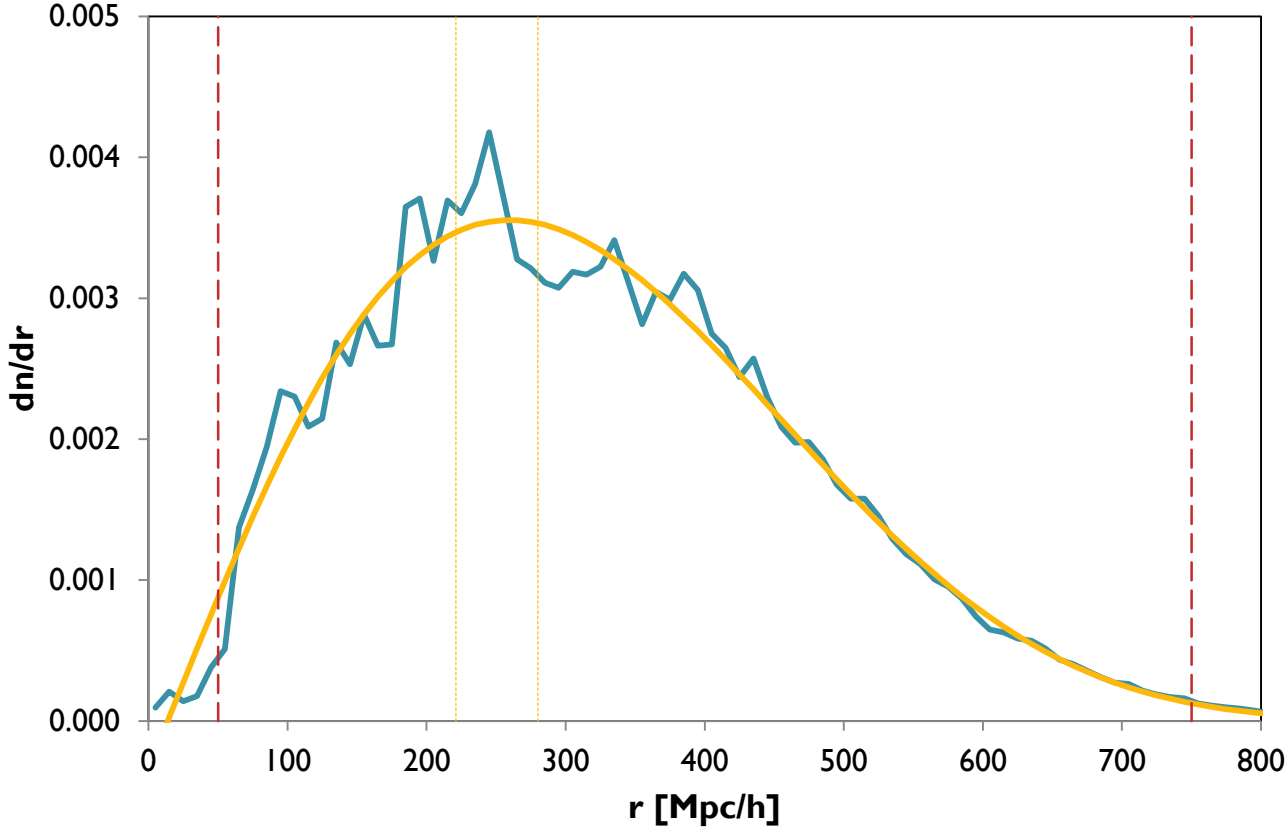
Computations on GPUs

- Generated 16M randoms with correct radial and angular selection for SDSS-N
- Done on an NVIDIA GeForce 260 card
- **600 trillion** galaxy/random pairs
- Brute force massively parallel code, much faster than tree-code for hi-res ξ
- All done inside the SDSS database
- Correlation function is now DB utility

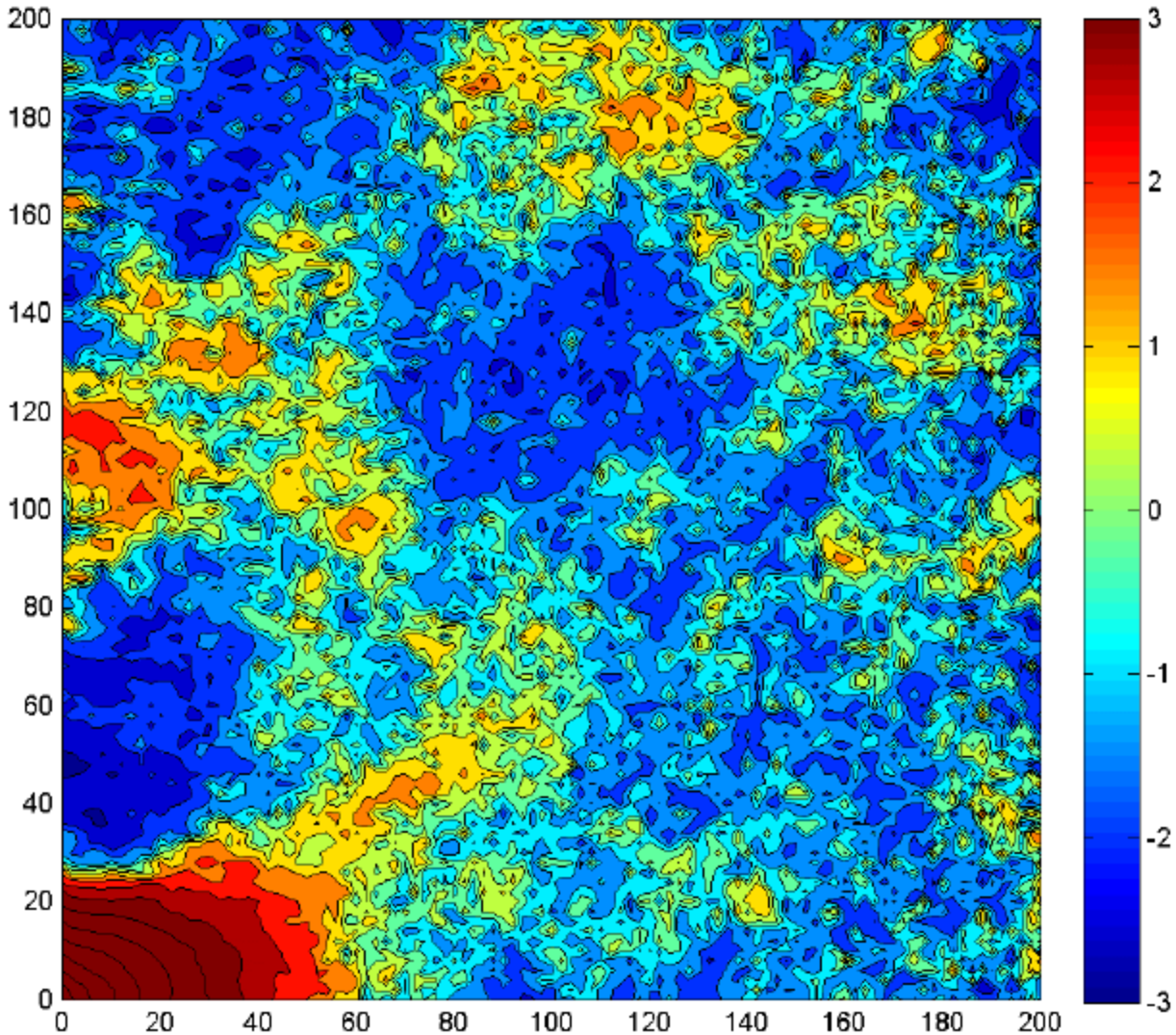
Full 3D correlation function



Radial Distribution



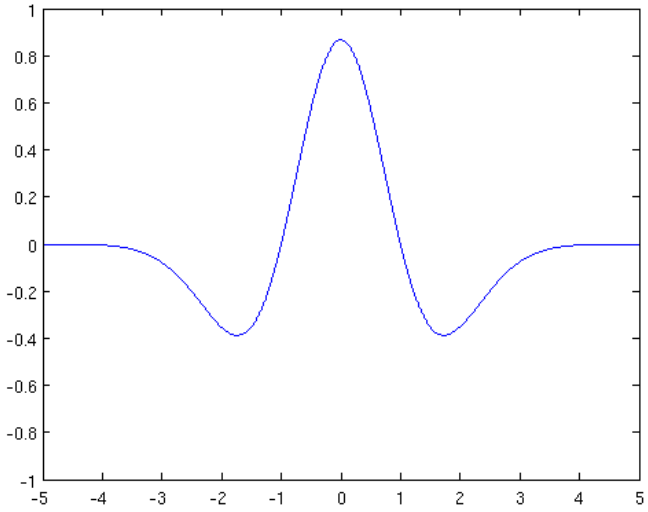
300-750 Mpc/h Cut



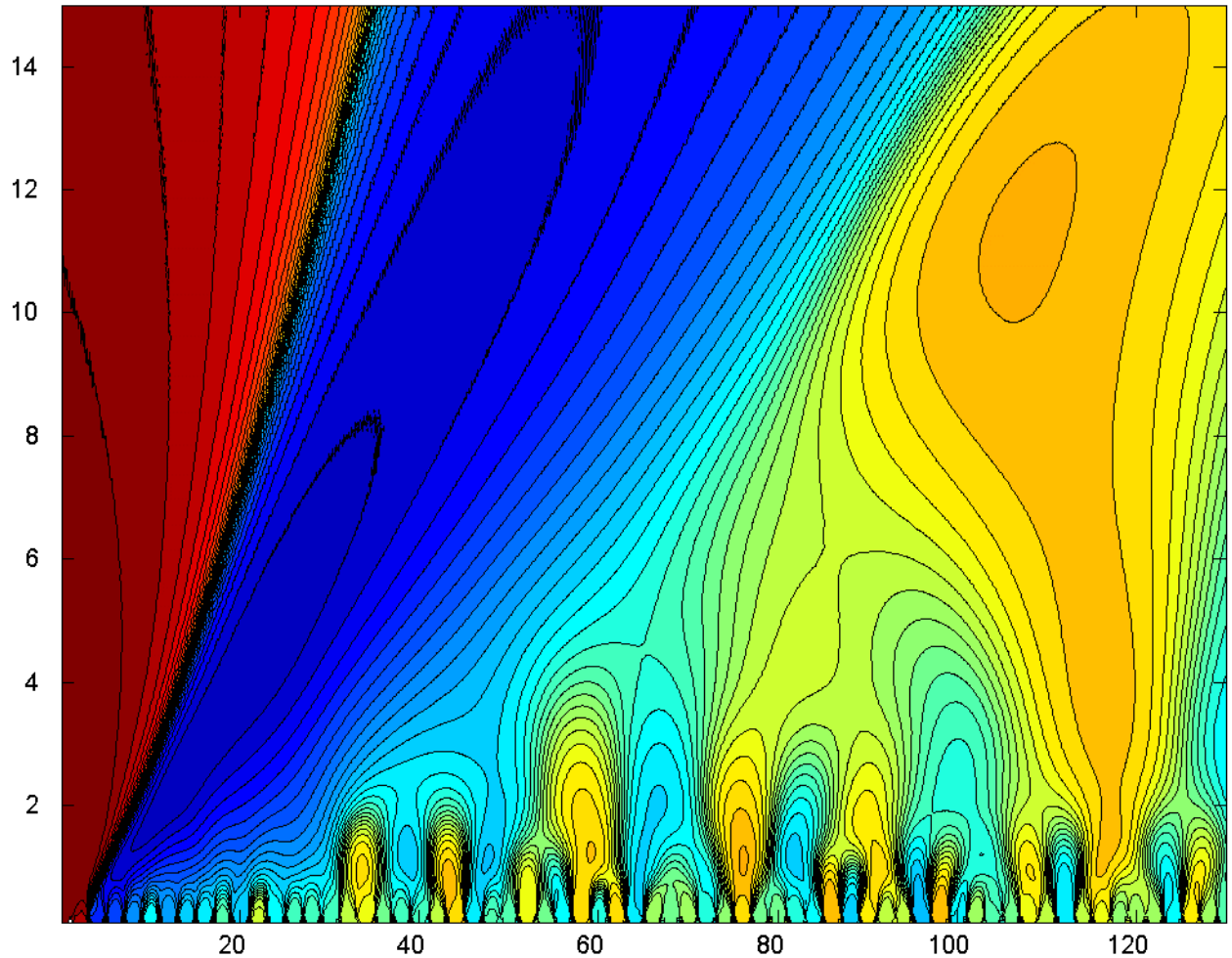
Lose 50% of galaxies, only 1/27th of volume

Wavelet Transform

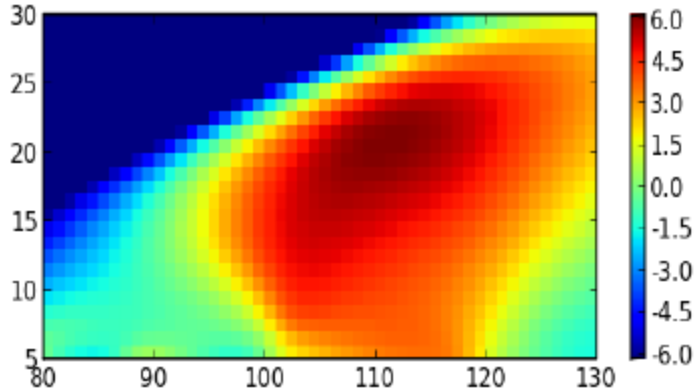
- Mexican hat wavelet transform
  - Compensated filter
  - Enhances localized "bump"
- Zero signal for constant background
- Decreases correlations among bins



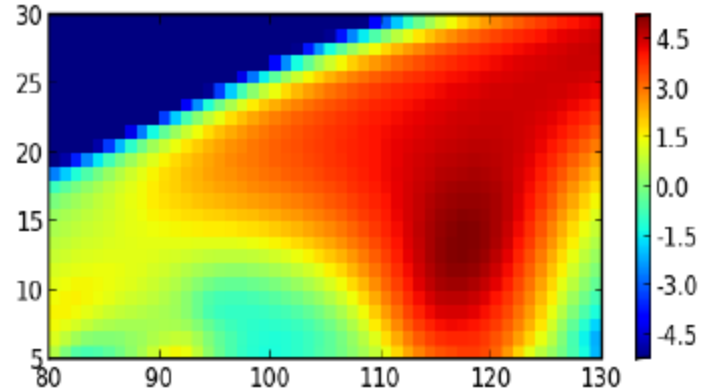
Wavelet Transform



S/N of the Wavelet Transform



Flat theta



LOS(6 deg)

Noise estimated from slices, divided by 16

Velocity effects

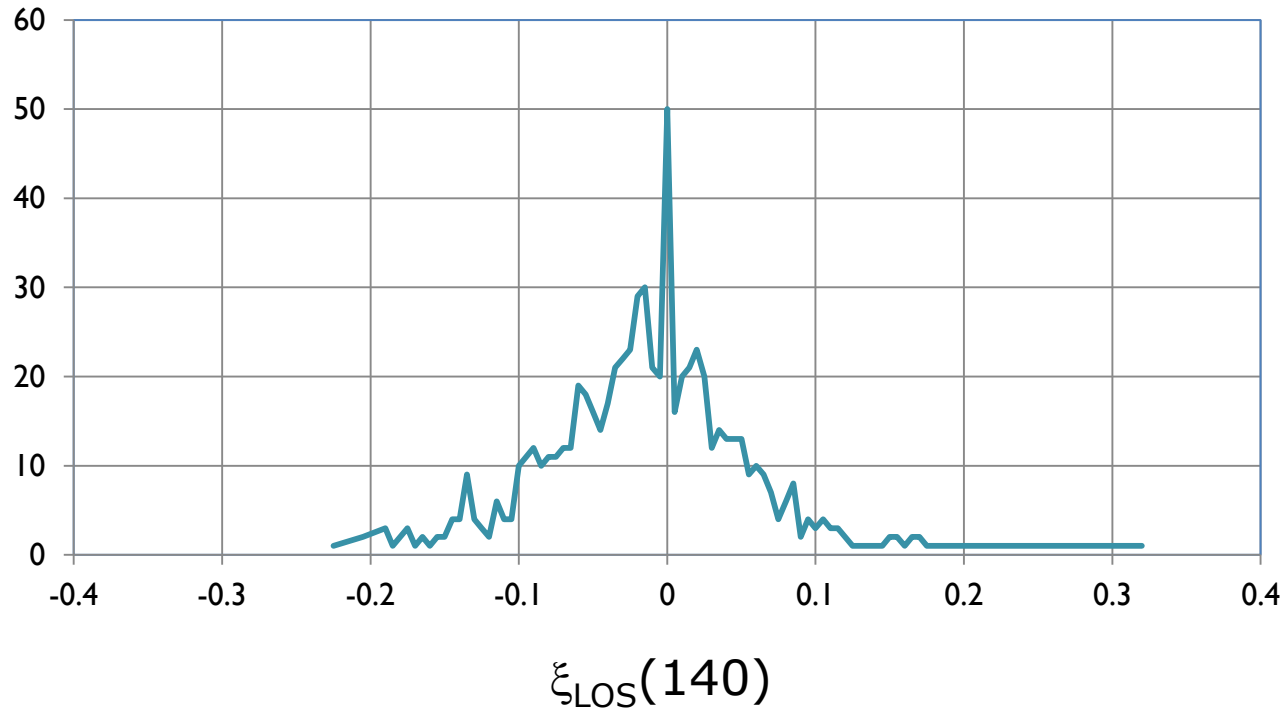
- Velocity correlations have longer range

$$\xi_{vv}(\mathbf{x}_1, \mathbf{x}_2) = (Haf)^2 \int d^3\mathbf{k}\, e^{i\mathbf{k}(\mathbf{x}_1 - \mathbf{x}_2)} \frac{1}{k^2} P(k)$$

- Nonlinear contributions in velocity correlations will extend much farther than in the density

Far Side Infall (140Mpc)

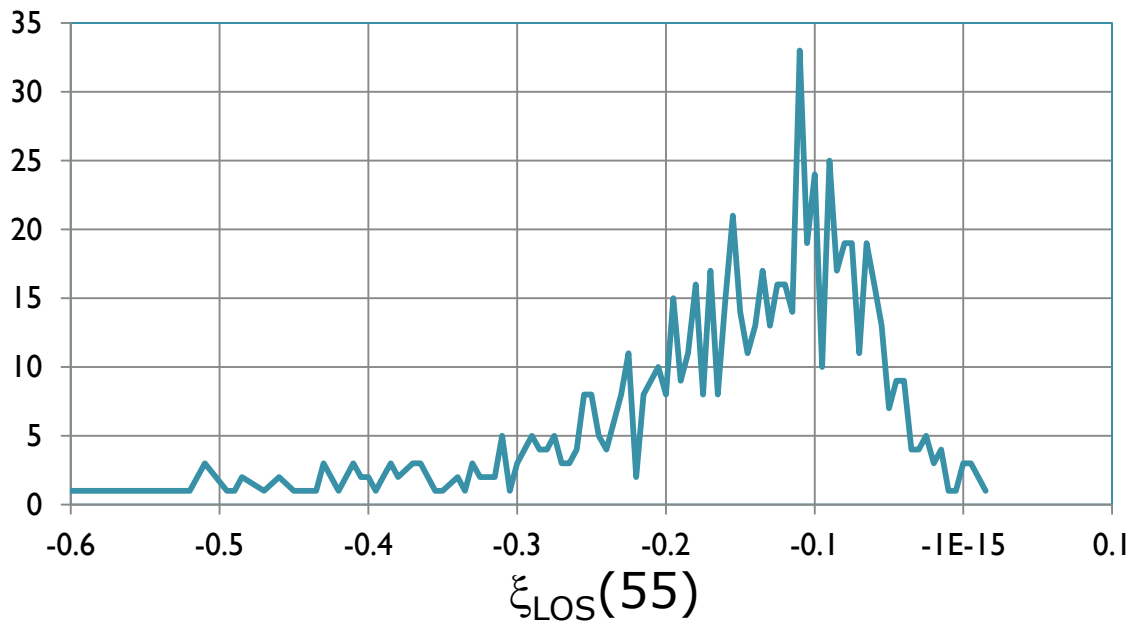
- Center: 140 h^{-1}Mpc, width: 25 h^{-1}Mpc
- Still shows some skewness



Strong Non-Linear Infall (55Mpc)

Distribution of LOS Mexican Hat wavelet coefficients over the 660 slices,

- Center at 55 h^{-1}Mpc, width 25 h^{-1}Mpc



BAO are Sharper in Redshift Space

- Near and far side infall onto the BAO bump
- Angular averaging wipes out most of this effect
- Evidence for BAO even in SDSS DR7 MGS at ~110 h^{-1}Mpc, potentially constraining the equation of state at low z
- Trough at 55 h^{-1}Mpc indicates effects of strong nonlinear infall on these scales
- Nonlinearities significant even on BAO scales!

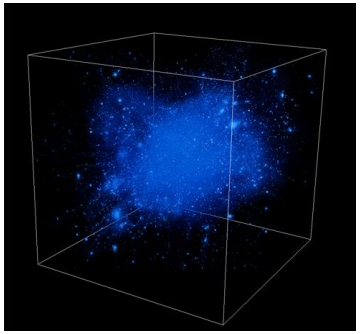
Science is Changing

THOUSAND YEARS AGO
science was empirical
describing natural phenomena

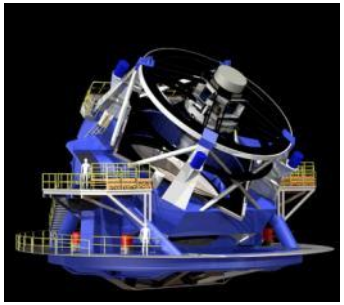
LAST FEW HUNDRED YEARS
theoretical branch using models, generalizations

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - \mathrm{K}\frac{c^2}{a^2}$$

LAST FEW DECADES
a computational branch simulating complex phenomena



TODAY
data intensive science, synthesizing theory, experiment and computation with statistics
- new way of thinking required!



Impact of Sky Surveys

Astronomy

Sloan Digital Sky Survey tops astronomy citation list

NASA's Sloan Digital Sky Survey (SDSS) is the most significant astronomical facility, according to an analysis of the 200 most cited papers in astronomy published in 2006. The survey, carried out by Juan Madrid from McMaster University in Canada and Duccio Macchetto from the Space Telescope Science Institute in Baltimore, puts NASA's Swift satellite in second place, with the Hubble Space Telescope in third (arXiv:0901.4552).

Madrid and Macchetto carried out their analysis by looking at the top 200 papers using NASA's Astrophysics Data System (ADS), which charts how many times each paper has been cited by other research papers. If a paper contains data taken only from one observatory or satellite, then that facility is awarded all the citations given to that article. However, if a paper is judged to contain data from different facilities – say half from SDSS and half from Swift – then both facilities are given 50% of the citations that paper received.

Top 10 telescopes

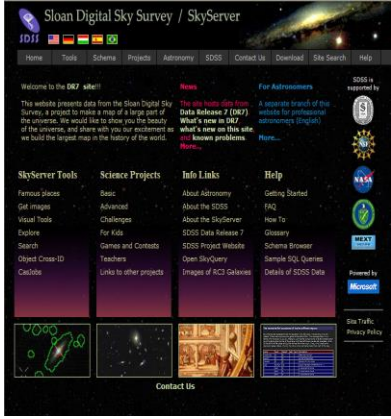
Rank	Telescope	Citations	Ranking in 2004
1	Sloan Digital Sky Survey	1892	1
2	Swift	1523	N/A
3	Hubble Space Telescope	1078	3
4	European Southern Observatory	813	2
5	Keck	572	5
6	Canada–France–Hawaii Telescope	521	N/A
7	Spitzer	469	N/A
8	Chandra	381	7
9	Boomerang	376	N/A
10	High Energy Stereoscopic System	297	N/A

The researchers then totted up all the citations and produced a top 10 ranking (see table). Way out in front with 1892 citations is the SDSS, which has been running since 2000 and uses the 2.5 m telescope at Apache Point in New Mexico to obtain images of more than a quarter of the sky. NASA's Swift satellite, which studies gamma-ray bursts, is second with 1523 citations, while the Hubble Space Telescope (1078 citations) is third.

Although the 200 most cited papers make up only 0.2% of the references indexed by the ADS for papers published in 2006, those 200 papers account for 9.5% of the citations. Madrid and Macchetto also ignored theory papers on the basis that they do not directly use any telescope data. A similar study of papers published in 2004 also puts SDSS top with 1843 citations. This time, though, the European Southern Observatory, which has telescopes in Chile, comes second with 1365 citations and the Hubble Space Telescope takes third spot with 1124 citations.

Michael Banks

Skyserver

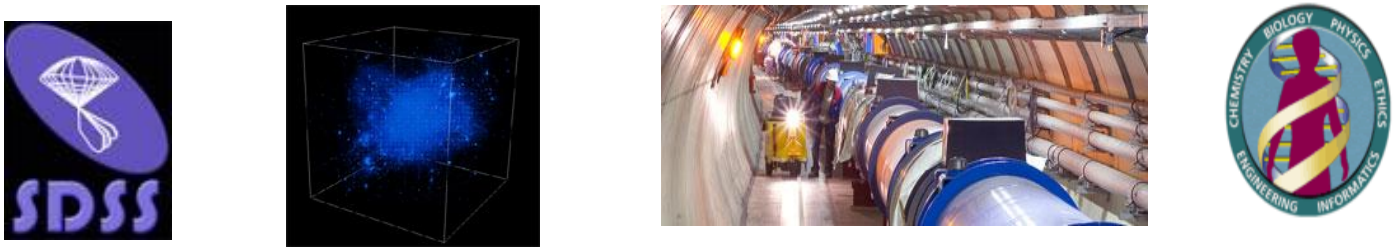


Prototype in 21st Century data access

- 1.2B web hits in 12 years
- 200M external SQL queries
- 4,000,000 distinct users vs. 15,000 astronomers
- The emergence of the "Internet Scientist"
- The world's most used astronomy facility today
- Collaborative server-side analysis done by 7K astronomers

Trends

- Broad sociological changes
  - Convergence of Physical and Life Sciences
  - Data collection in ever larger collaborations
  - Virtual Observatories: CERN, IVOA, NCBI, NEON, OOI,…
  - Analysis decoupled, off archived data by smaller groups
  - Emergence of the citizen/internet scientist (GalaxyZoo…)
- Need to start training the next generations
  - Π-shaped vs I- and T-shaped people
  - Early involvement in "Computational thinking"



- Science is increasingly driven by data (big and small)
- Surveys analyzed by individuals
- From hypothesis-driven to data-driven science
- New instruments: "microscopes" & "telescopes" for data
- A new, Fourth Paradigm of Science is emerging…
- SDSS has been at the cusp of this transition