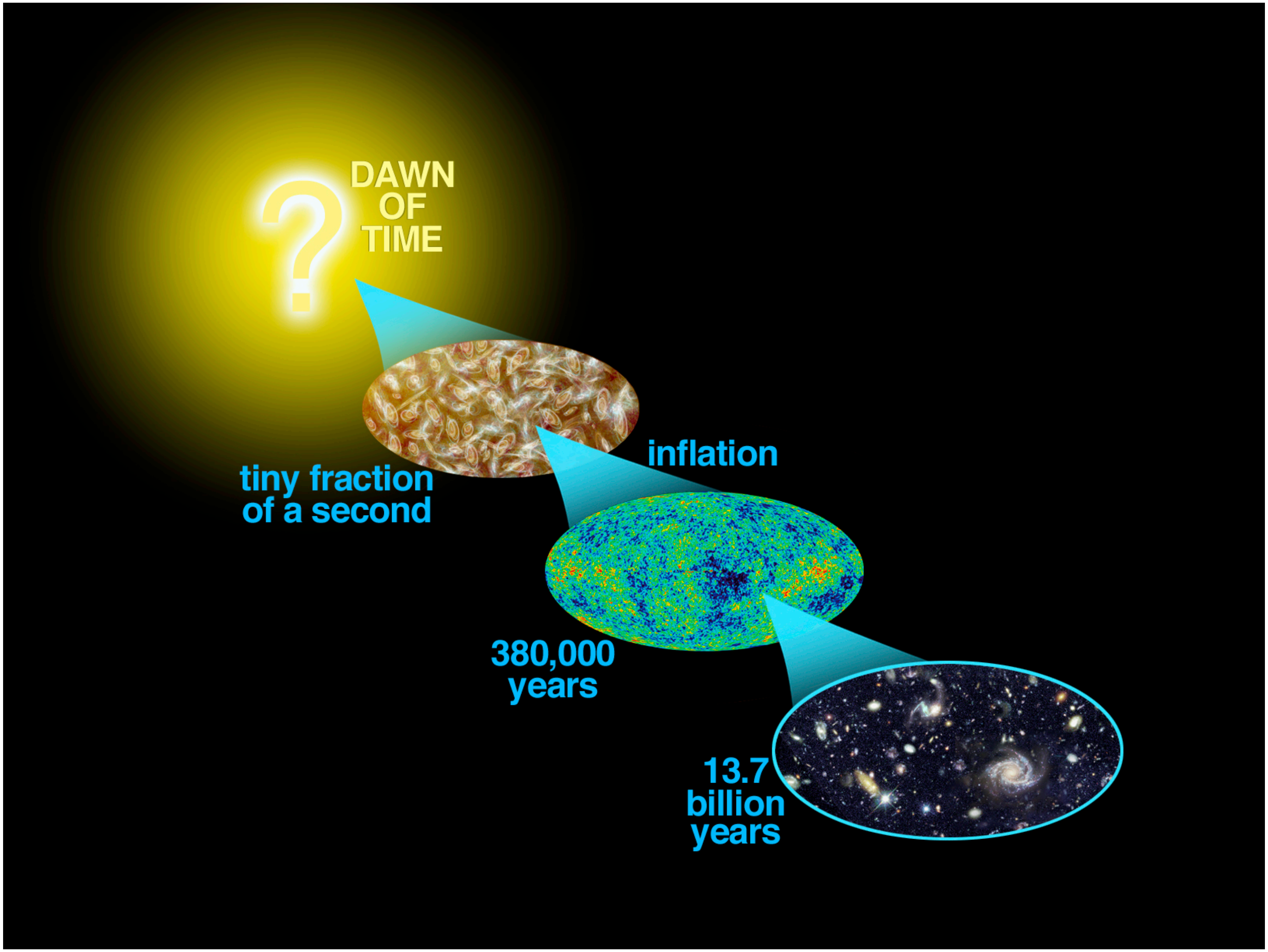
Questions?


DAWN
OF
TIME
inflation
tiny fraction
of a second
380,000
years
13.7
billion
years

What do we mean by “the large-scale structure” of the Universe?

What is a “massive redshift survey”?

What is galaxy biasing?

What can we learn from studying the spatial galaxy distribution?

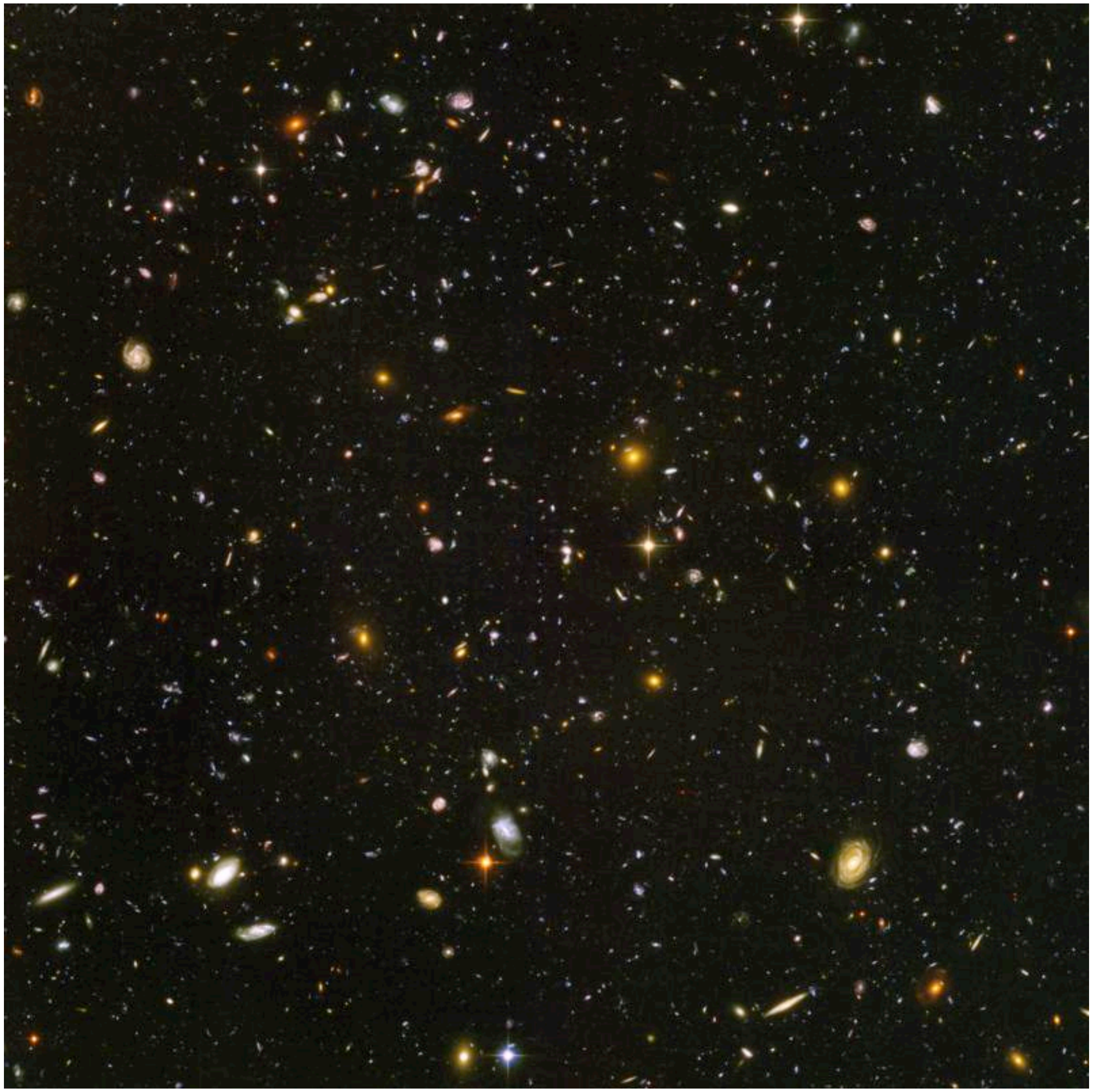
Courtesy of the HUDF team

The deepest image of the universe in the optical waveband

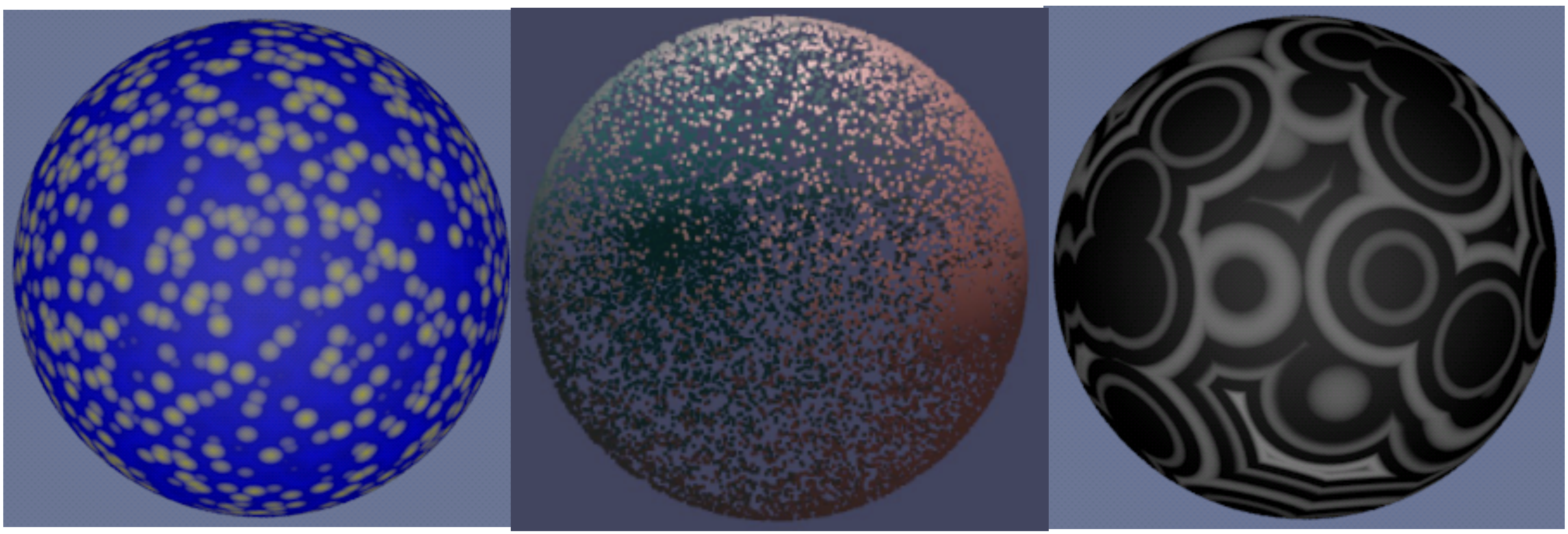
It contains ~ 10^4 galaxies in a solid angle corresponding to 1/50 of the Moon

The whole sky corresponds to 12.7 million times more solid angle

Galaxies are the fundamental building blocks of the distribution of luminous matter



How are galaxies distributed on the sky?



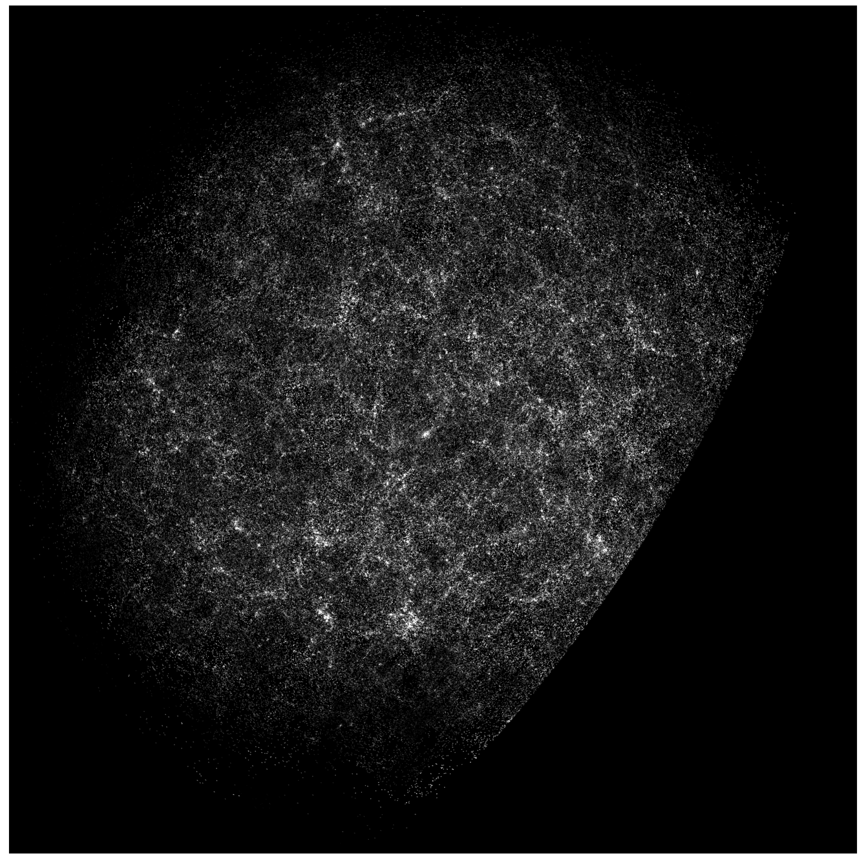
The Lick galaxy survey

The Lick survey conducted at the Lick observatory in Santa Cruz during the 1950s recorded the position of 10^6 galaxies in 10 years.

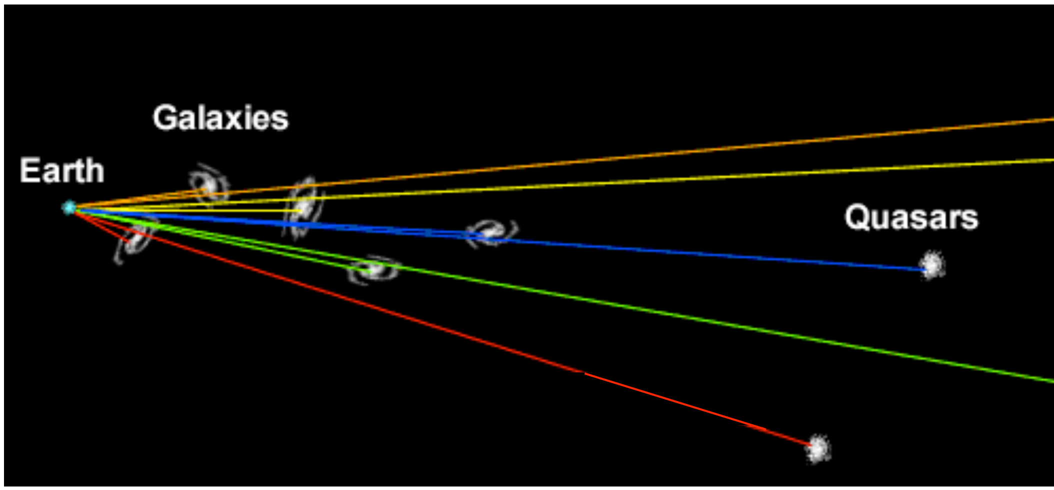
It provided the first evidence that galaxies are not distributed at random on the sky (Totsuji & Kihara 1969, Peebles & Hauser 1974, Groth & Peebles 1977)

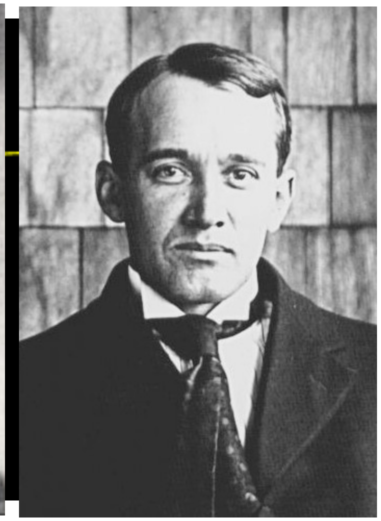
Projected filamentary structure is evident but what is the corresponding 3D distribution?

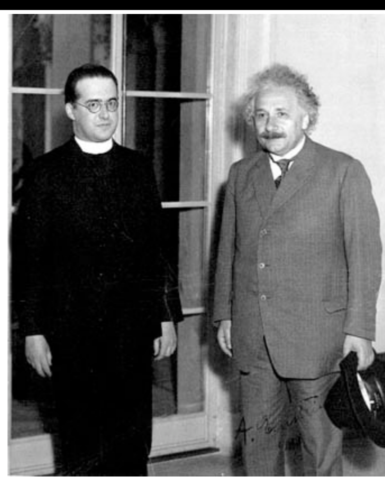
Shane & Wirtanen 1967

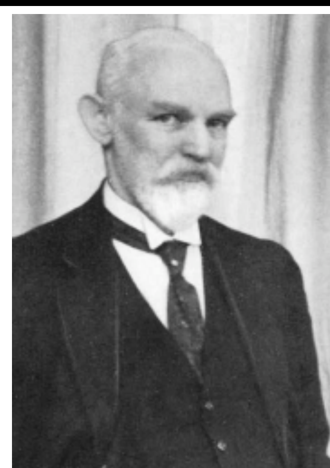


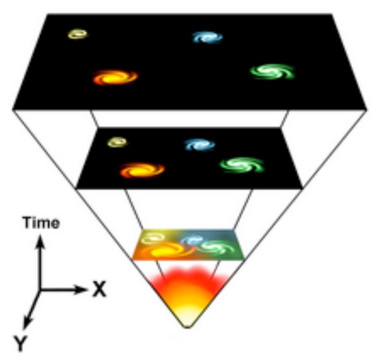
Adding the third dimension



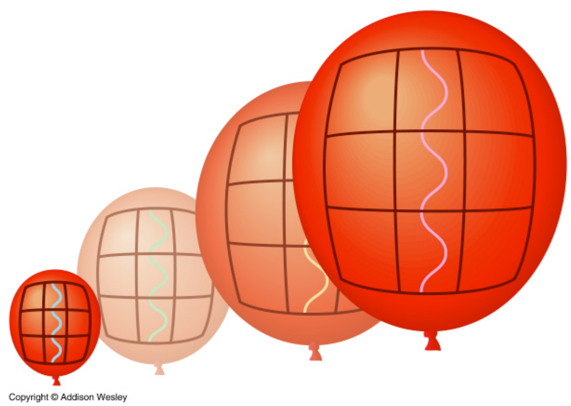


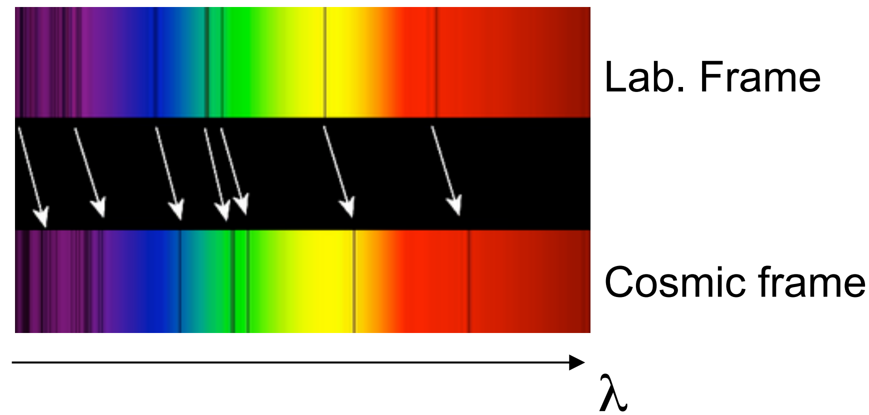






Adding the third dimension



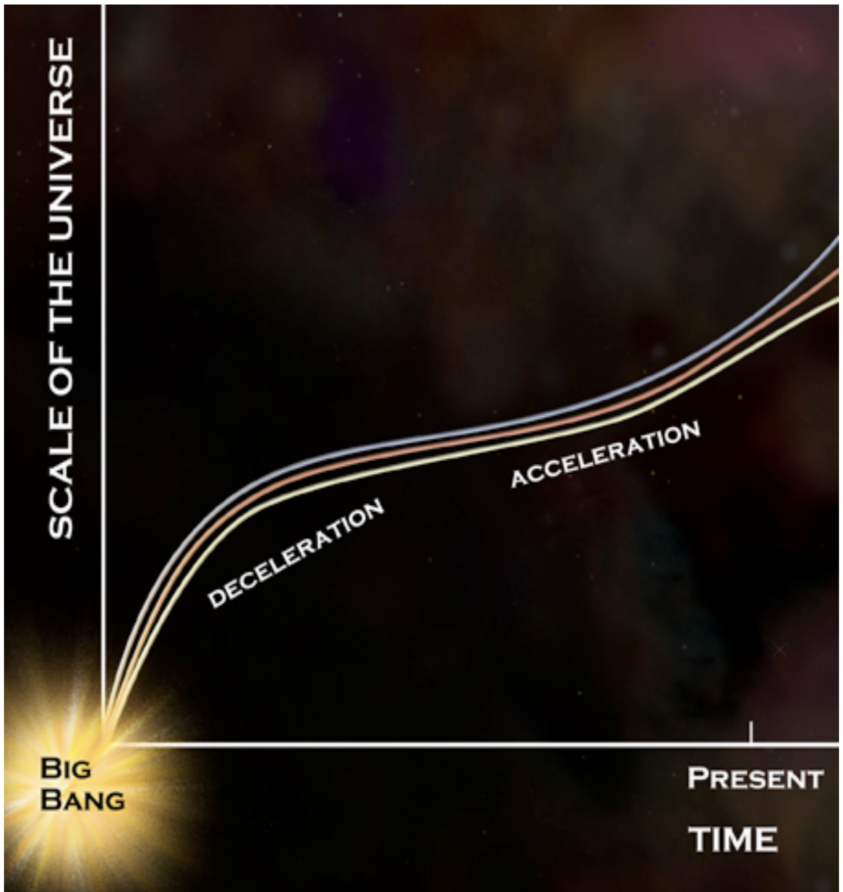


- Cosmological redshift: the wavelength of EM signals is stretched by the cosmic expansion.

- Spectra are shifted in frequency towards the red by a factor
  1+z = (present-day size of the universe)/(size at photon emission)=$a(t_0)/a(t_{em})$

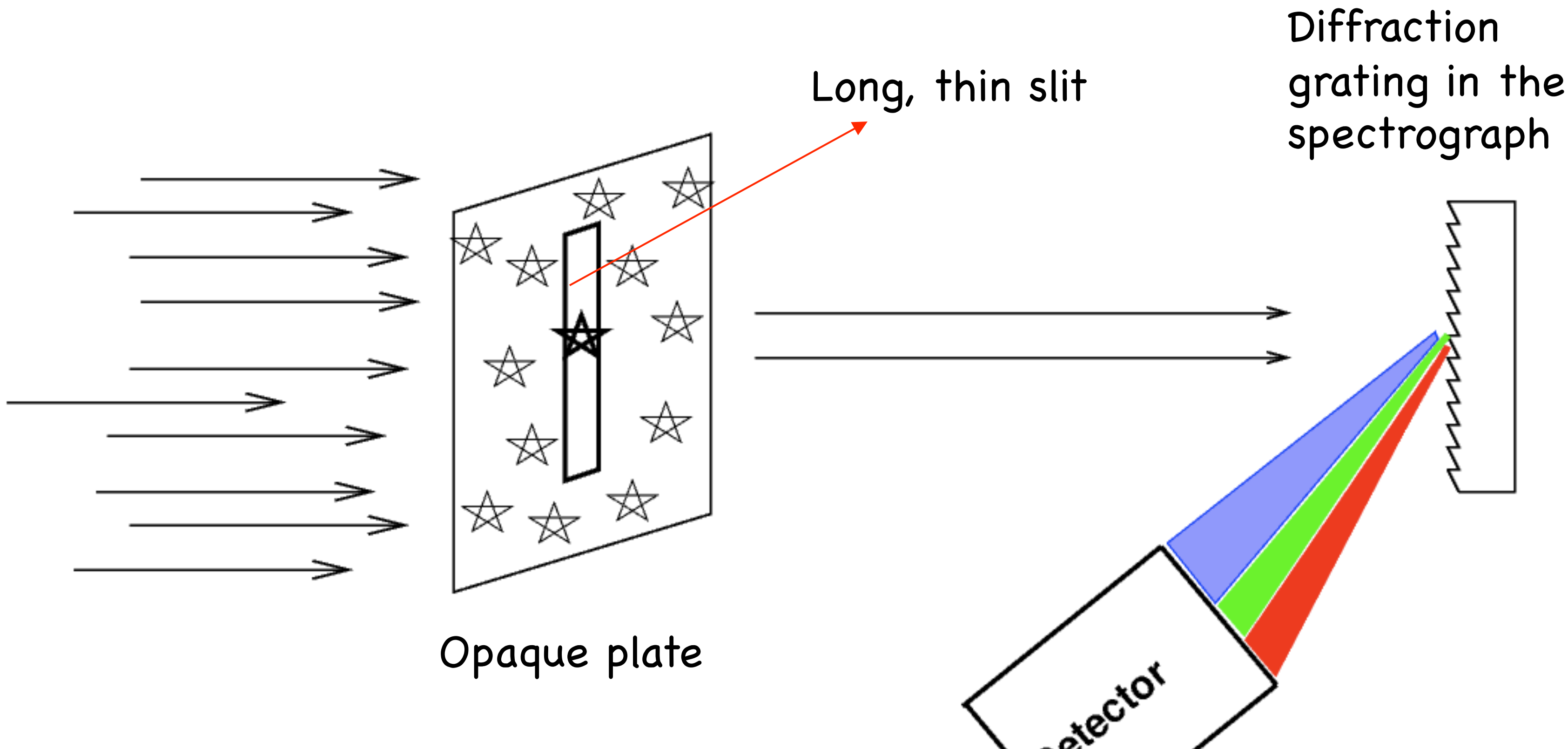
$$1+z=\lambda_{obs}/\lambda_{em}$$

Adding the third dimension - II



- Knowing the cosmic expansion history, a(t), would allow us to convert the ratio of scale factors into a look-back time: $t_{now} - t_{em}$
- Then, taking into account that EM signals propagate at constant velocity, c, one could derive the radial distance on our past-light cone corresponding to a given cosmological redshift.
- Measuring redshifts is then necessary to study the galaxy distribution in three dimensions

Slit spectroscopy



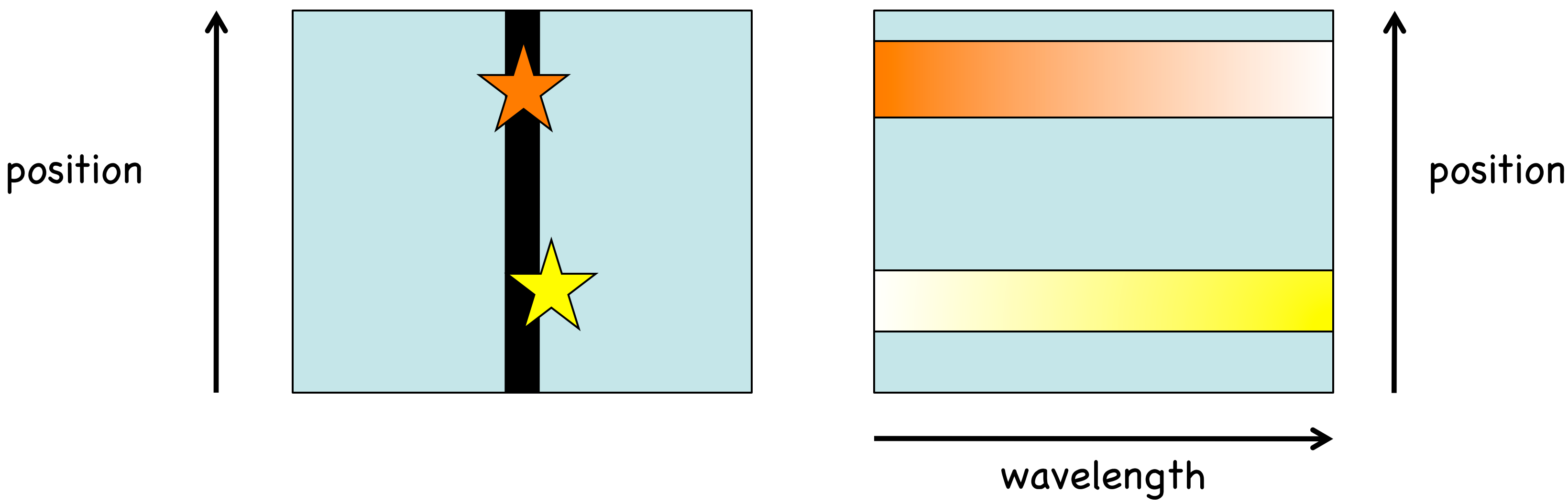
The image of a celestial body is focused on the slit. The diverging beam then reaches a collimator (either a lens or a mirror - not shown -). This produces parallel light which is then dispersed by a diffraction grating.

The dispersed light is imaged on a detector (typically a CCD)

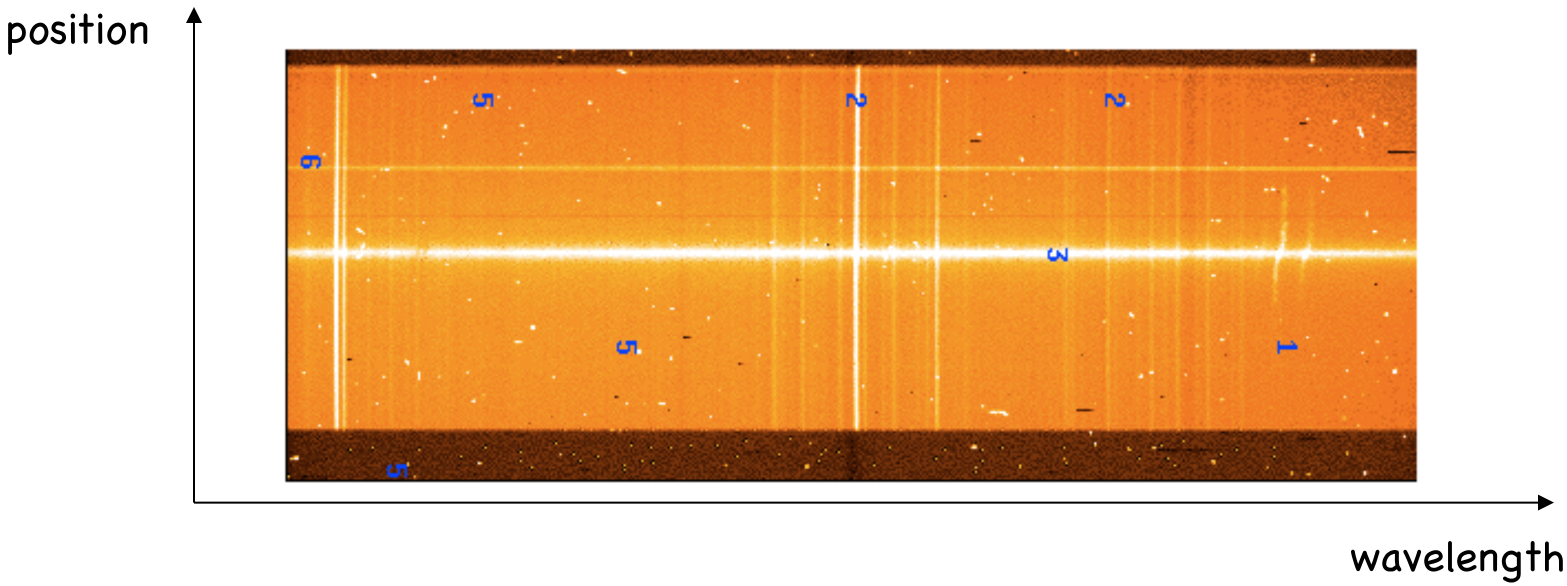
Slit spectroscopy

- Why using a slit? To keep out as much as background light as possible

- How does the output look like? 2D spectrum



A raw galaxy spectrum



- This is the typical output of a spectrograph mounted on a telescope.

- Can you identify the origin of the different features?

Answer

1. Emission lines from the observed galaxy (note that the galaxy is rotating)

2. Emission lines from the Earth's night sky

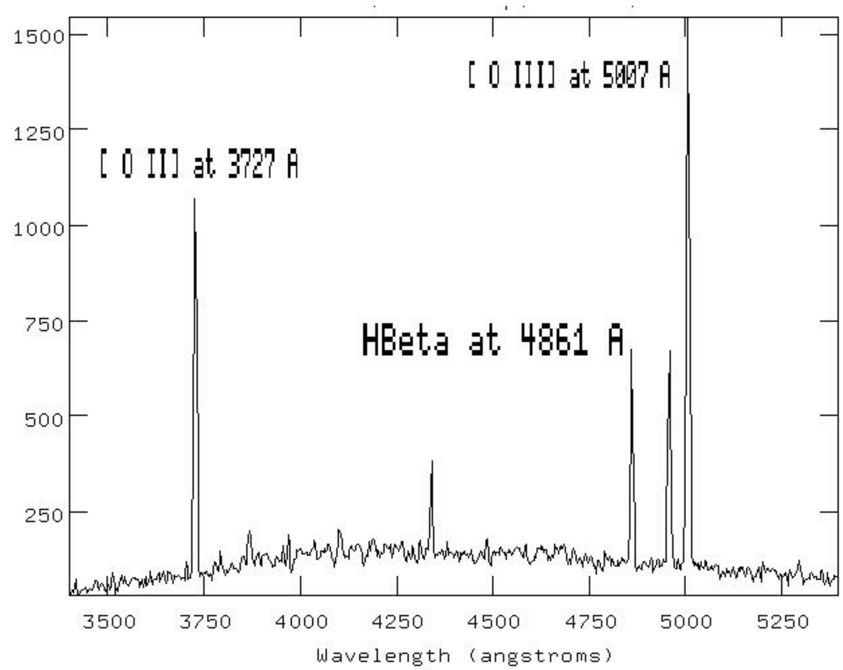
3. Continuum from the observed galaxy

4. -

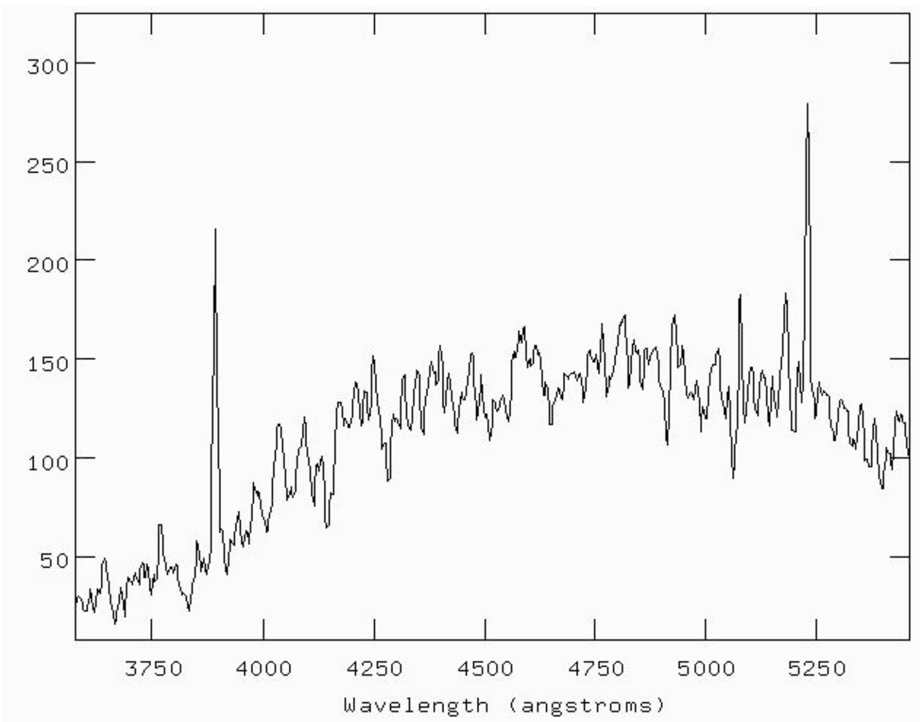
5. Cosmic rays

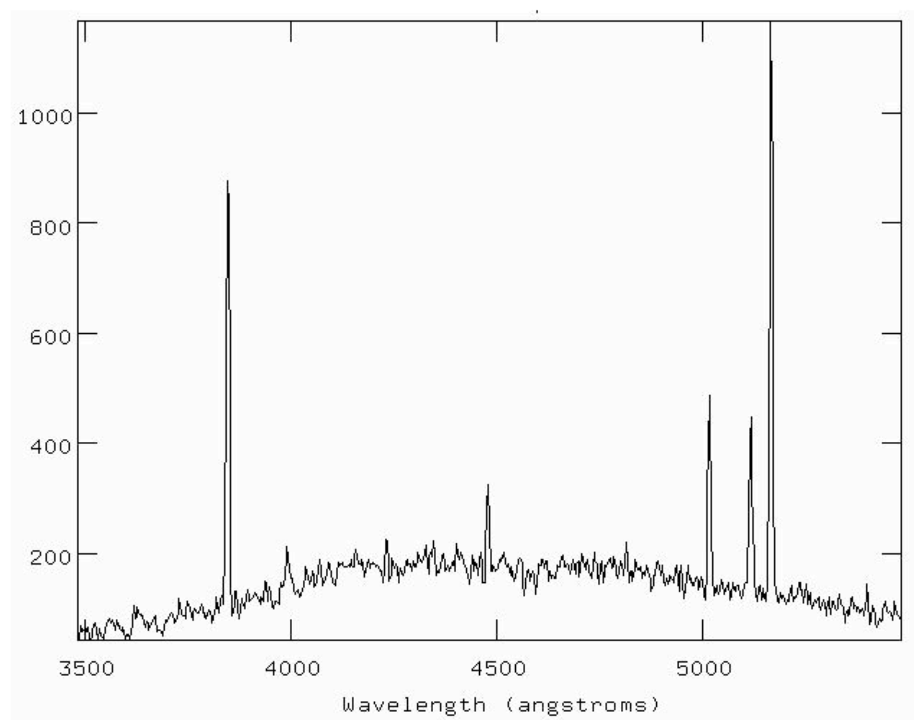
6. Continuum from a nearby star

Measuring galaxy redshifts



Template spectrum at z=0

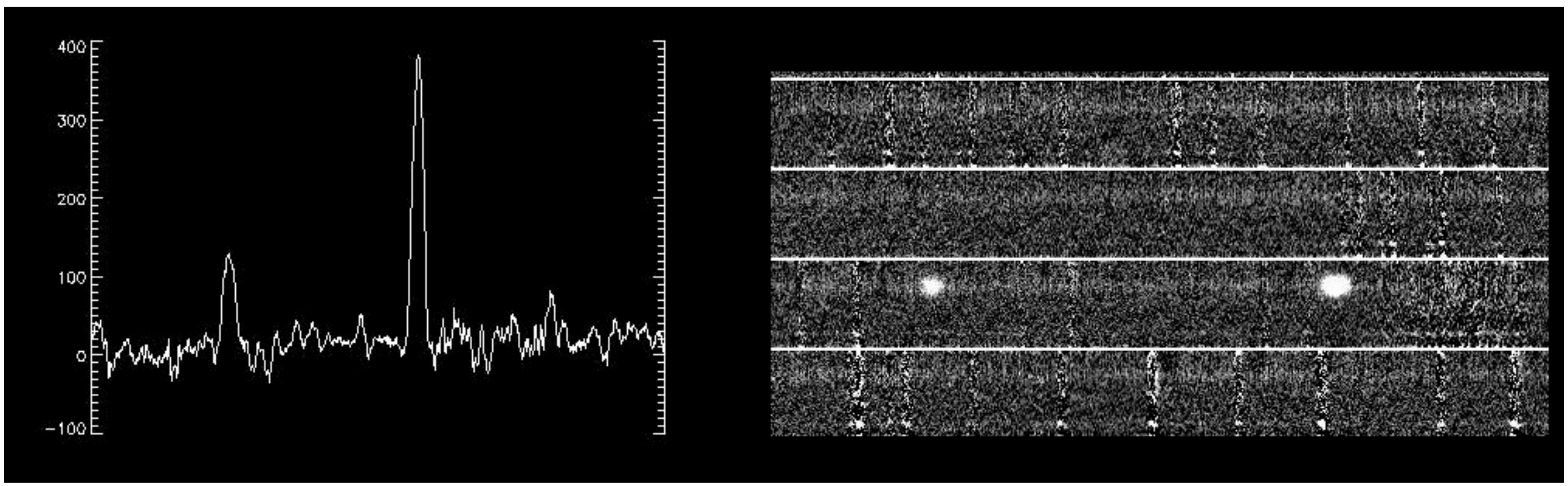




Spectrum of a galaxy with a faint continuum

1D spectrum

2D spectrum



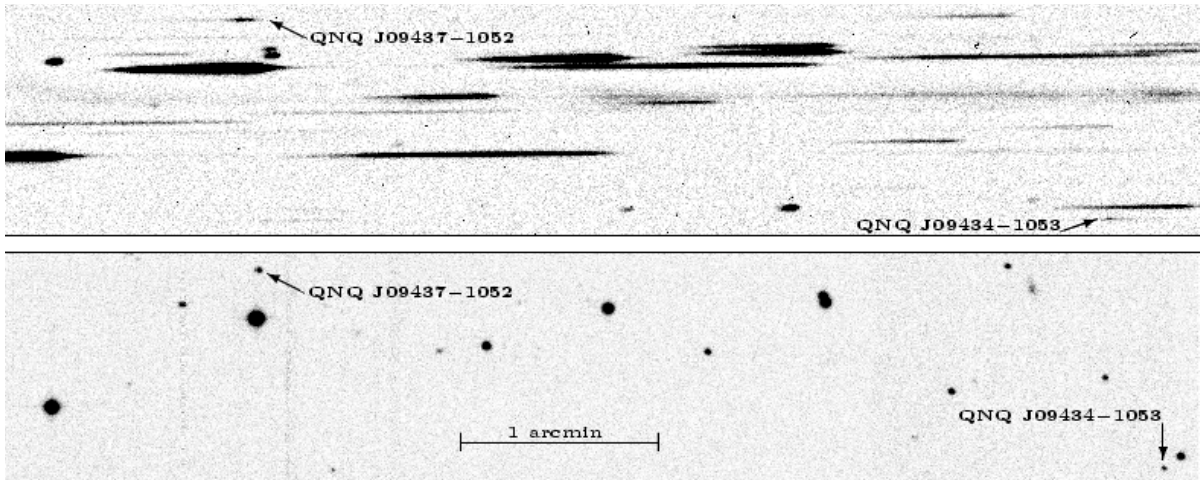
wavelength

Quality flags
an example from the zCOSMOS survey

Confidence Class	Description	Spectroscopic verification	ZEBRA photo-z consistency within $\Delta z=0.1(1+z)$
Class 4	Very secure redshift	>99.5%	97%
Class 3	Secure redshift	>99.5%	97%
Class 9	One line either Ha or [OII] 3727 (best guess)	95%	90%
Class 8	Unidentifed one line (best guess)		
Class 2	Probable redshift	92%	93%
Class 1	Insecure redshift	70%	72%

Confidence class	spectroscopic/photometric consistency
.5	Spectroscopic redshift consistent within 0.1(1+z) of the photometric redshift
.4	No photometric redshift available, includes all spectroscopic AGN and stars
.3	Special case for Class 9: Consistent with photo-z only after the redshift changed to the alternate redshift, a switch which is then applied
.1	Spectroscopic and photometric redshifts are not consistent at the level of 0.1(1+z)

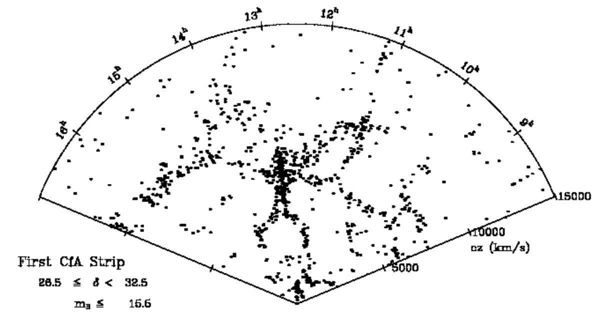
Slitless spectroscopy

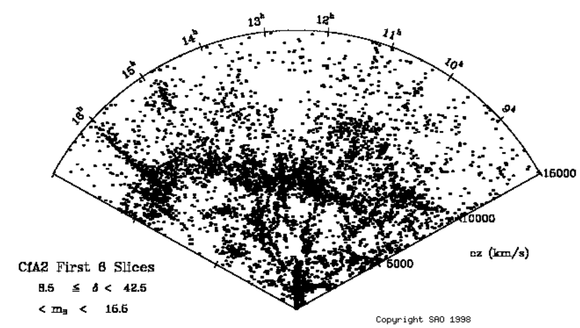


- It is possible to take spectra of all objects in the field of view that are brighter than a given threshold level (determined by the background)
- This is done using a dispersion element directly in combination with a telescope
- The disadvantage is that the background is high and that in crowded fields many spectra will overlap
- Filters are often used to isolate the spectral region of interest and reduce crowding and sky background intensity

Early redshift surveys: revealing the LSS

- Taking spectra of faint galaxies needs large telescopes and long integration times!
- The first redshift surveys were completed in the 1980s and 1990s (e.g. CfA, IRAS, Las Campanas)
- They measured 10^3 to 10^4 galaxy redshifts. Demonstrated the existence of LSS but not good enough for statistics
- No astronomer would have predicted the complexity of the galaxy distribution that was revealed!





Huchra, Davis, Geller et al. 1982-1998

Measuring galaxy redshifts

- Measuring galaxy redshifts in the optical band is not easy

- Using a telescope of diameter 1.5m (the standard in the 1970s), it takes nearly 5 minutes of integration time to measure the redshift of a bright emission line galaxy (from HII regions). Nearly 20% of the local galaxy population presents such strong emission features.

- For most galaxies, redshifts must be determined from absorption lines (from the integrated contribution of stellar atmospheres). This requires measuring the continuum and needs 15-60 minutes of integration on a 1.5m telescope

- In 1972 (≈60 years after the discovery of redshifts) there were only ≈1000 galaxy redshifts known (mostly in galaxy clusters)

John Huchra (1948-2010)

- "The universe is what it is, and we are trying to find out what it is"
- "The explorers of the new world weren't trying to prove theories, they were looking at what was out there"
- "I could make telescopes sit up and take data"

Mapmaker, Mapmaker Make Me a Map

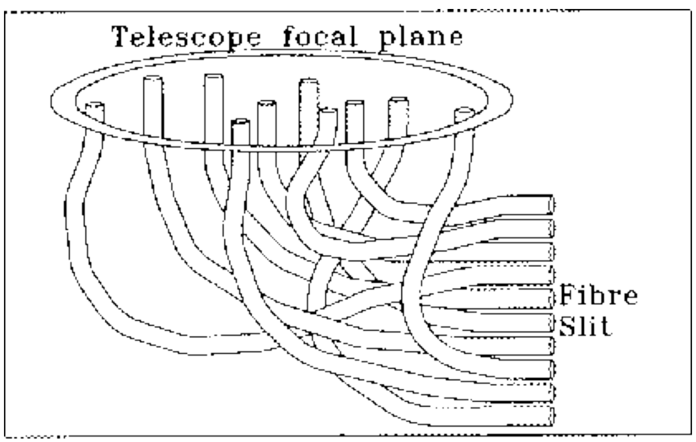
John P. Huchra

(Harvard-Smithsonian Center for Astrophysics)

I love being on mountaintops. Next best thing to being in space. I guess I also love counting things, whether the things are 4000 footers in New England, cards in games of chance, or galaxies on my observing list. There in, of course, lies the tale.

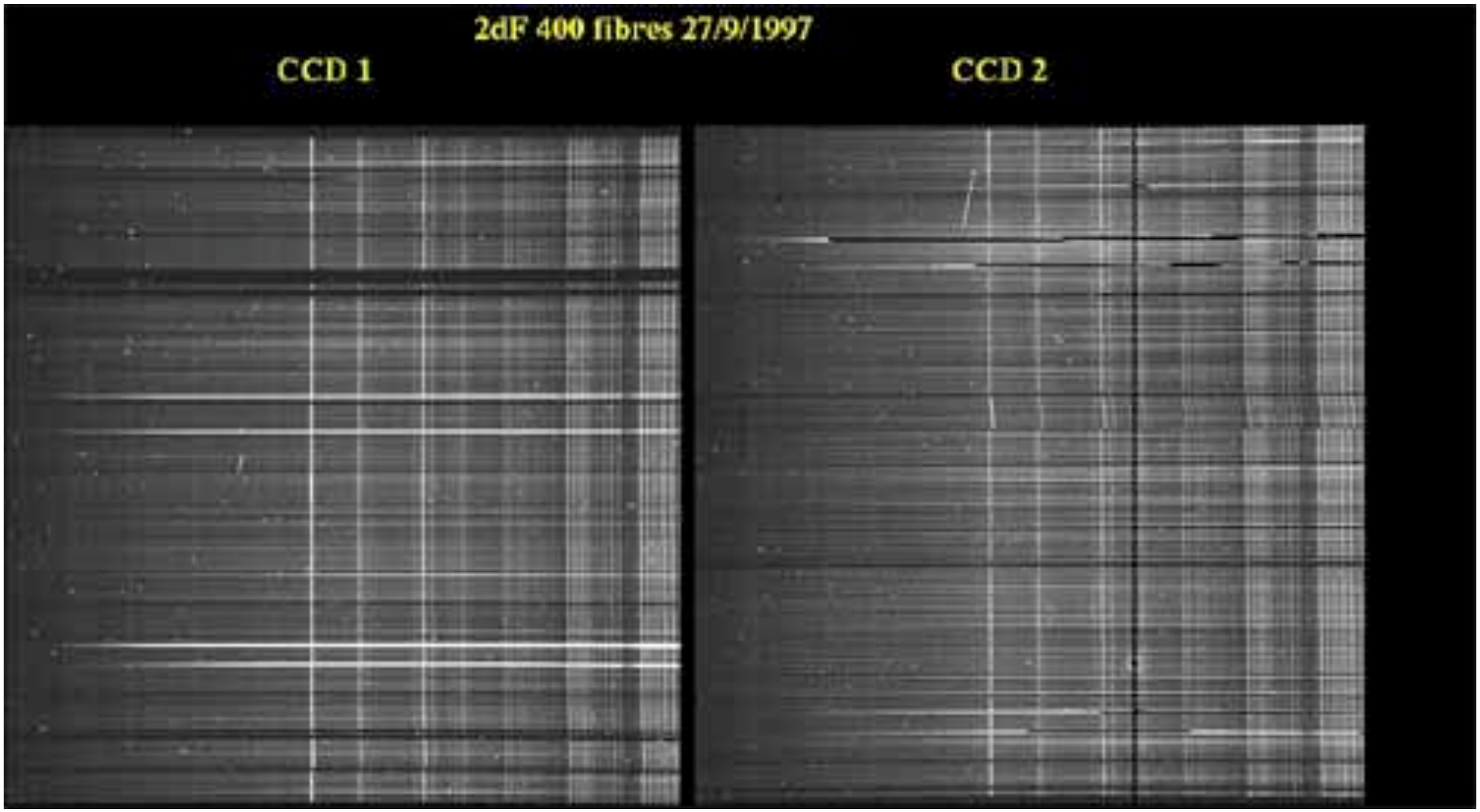
It all started because I was a little kid much more interested in reading than in sports.

Multi-object spectroscopy



- Single-slit spectroscopy does not make full use of the imaging capability of a telescope: several objects are imaged but only one is used
- Multi-slit and Fibre-fed spectrographs solve this problem.
- In the latter case, a set of optical fibers are positioned in the focal plane of the telescope so that each is illuminated by a target object. The fibers are then connected to a series of position in the spectrograph

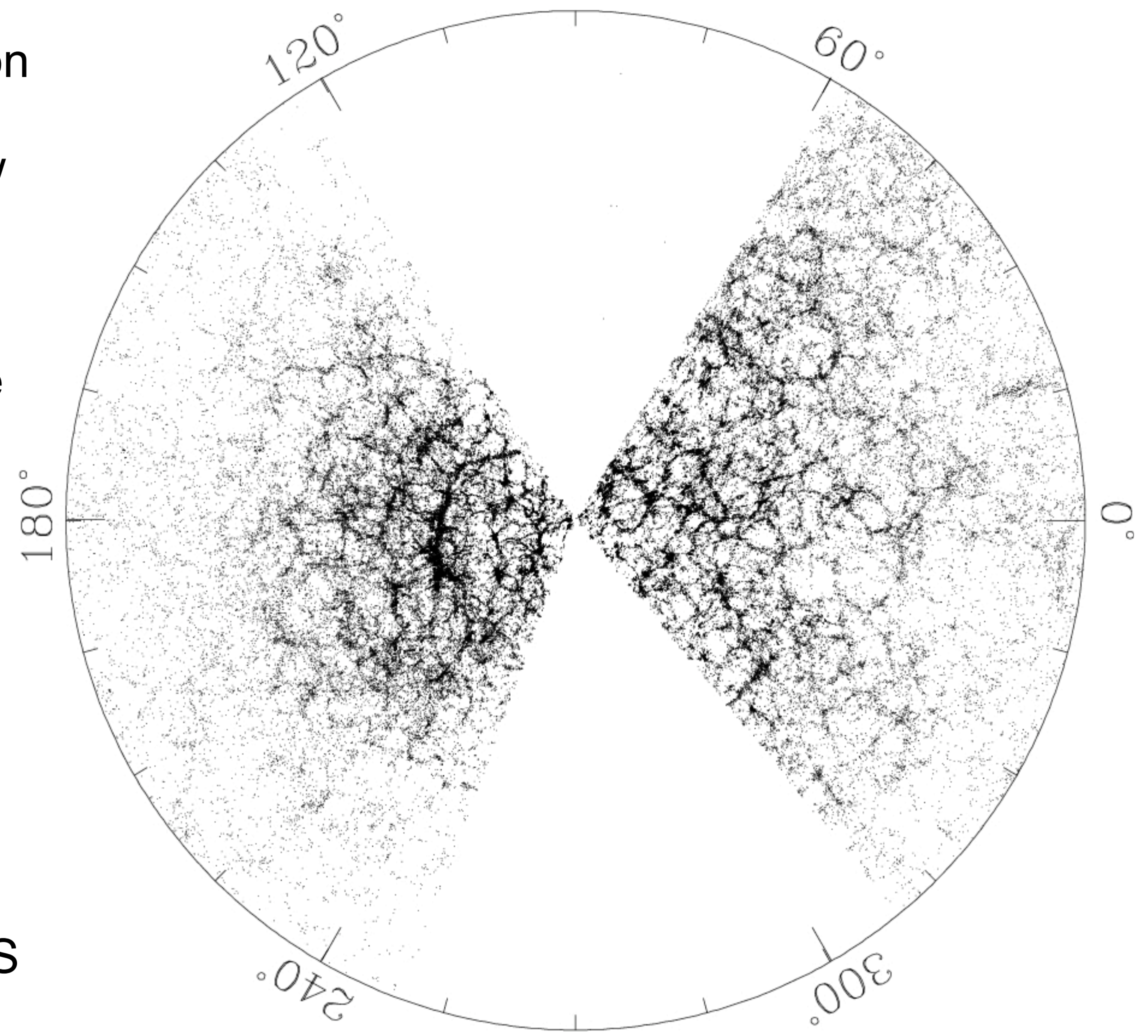
Multi-object spectra



Massive redshift surveys

- Multifibre technology, digitalization and multiobject spectrographs now allow us to measure redshift of millions of galaxies on a time scale of a few years.

- Recently completed or ongoing surveys: (local) 2dF, SDSS, 6dF (high-z) VVDS, DEEP2, zCOSMOS



The Sloan Digital Sky Survey

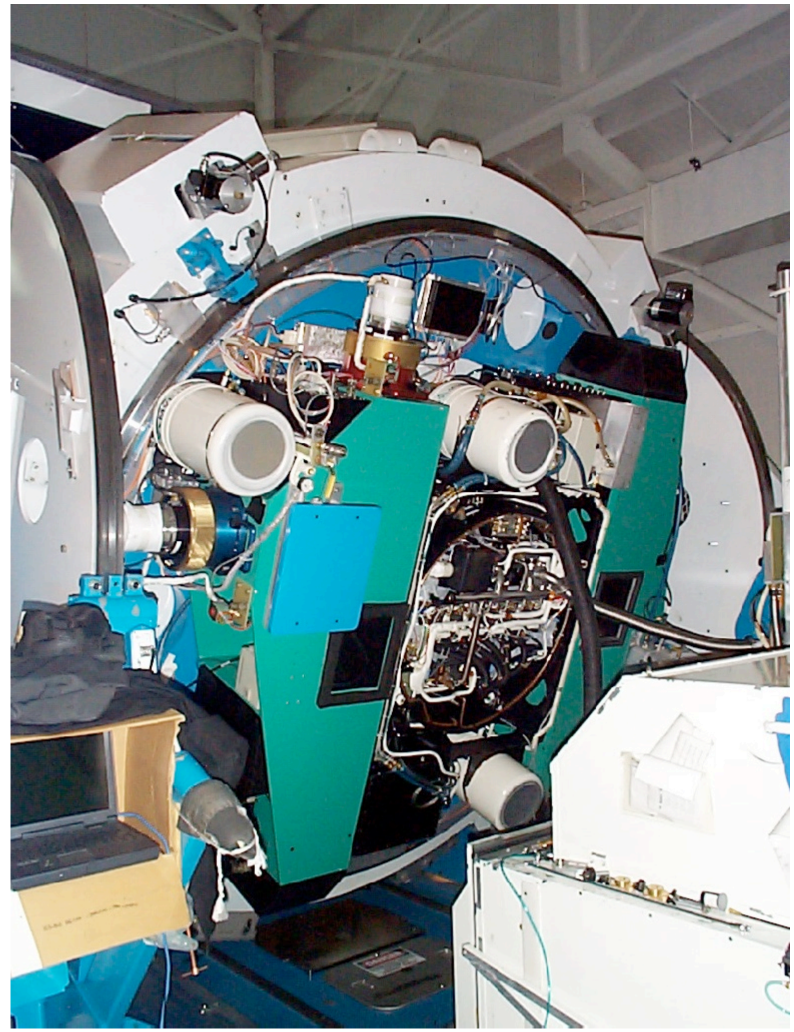
- Over eight years of operation (SDSS I, 2000-2005; SDSS II, 2005-2008; SDSS III, 2008-2014)

- It used a dedicated 2.5m telescope at Apache Point Observatory (New Mexico) equipped with 2 special purpose instruments: a 120 Mpixel camera imaging 1.5 sq. deg. of the sky at a time (8 times the area of the full moon); a pair of spectrographs fed by optical fibers (640 objects per pointing)

- It obtained deep multi-color images (u,g,r,i,z) covering more than a quarter of the sky (8,400 square degrees)

- Created 3D maps containing more than 930,000 galaxies and more than 120,000 quasars (in 5,700 square degrees)

The SDSS telescope and instruments



Groundbreaking technology

Photo: U. Montan

Charles Kuen Kao

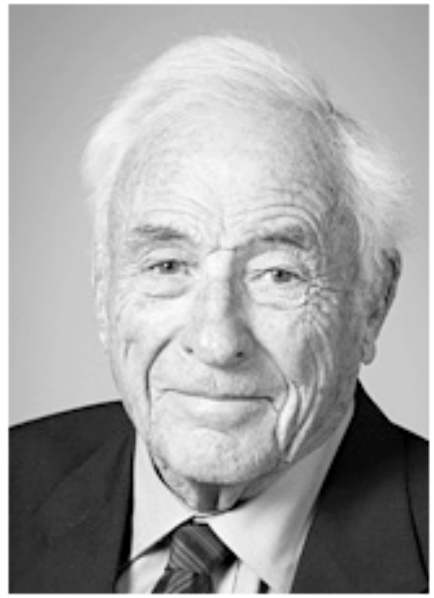


Photo: U. Montan

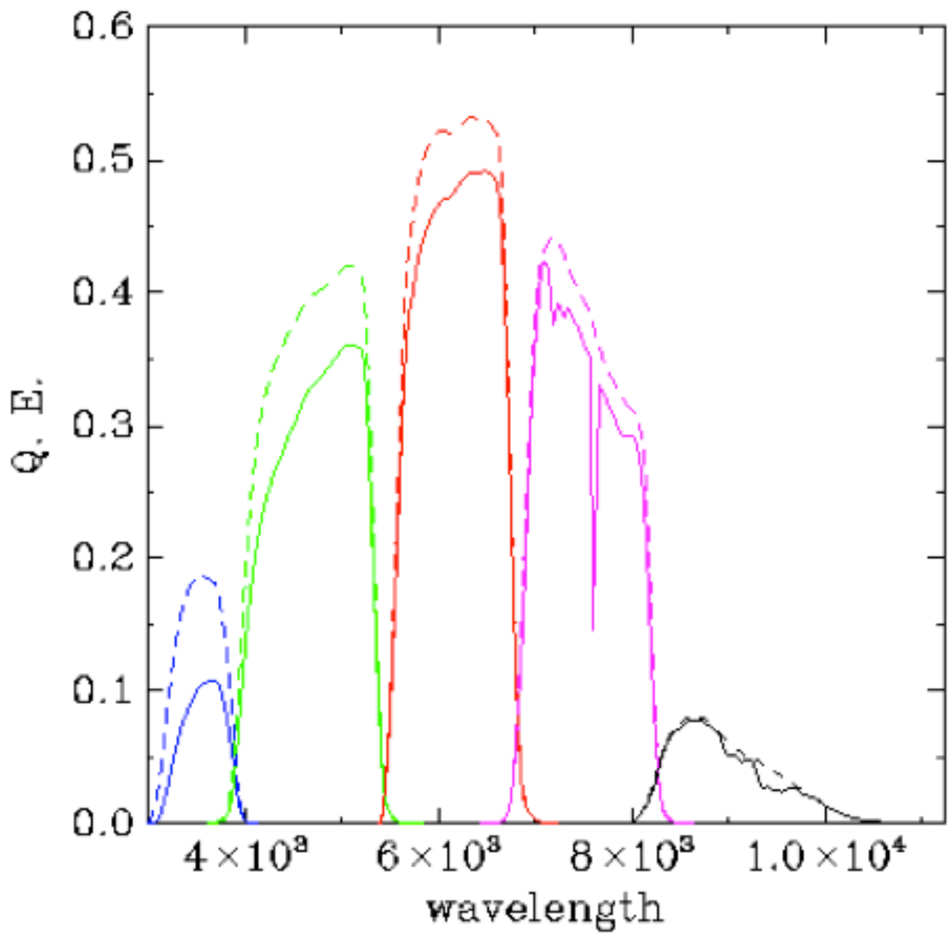
Willard S. Boyle

Photo: U. Montan

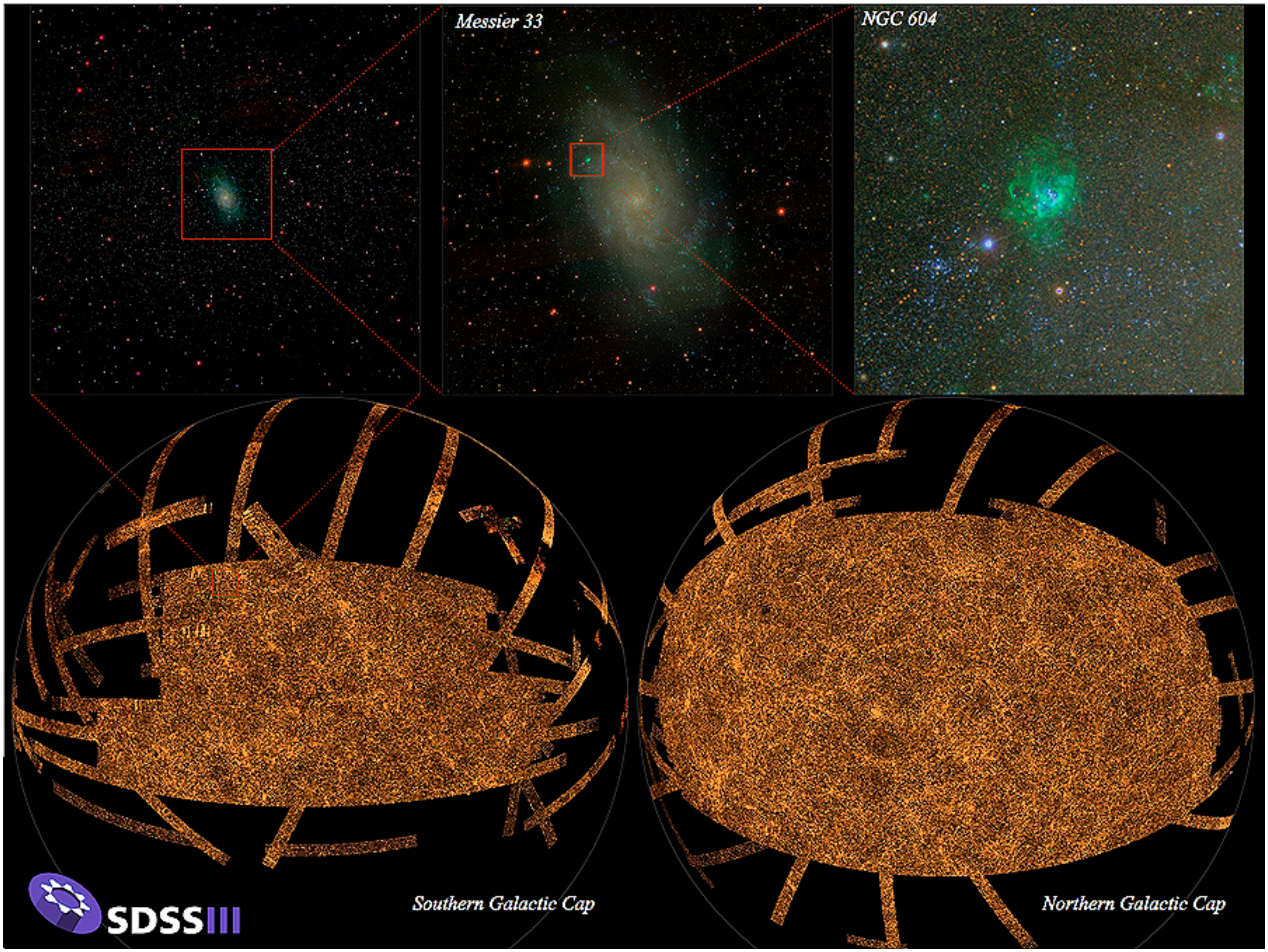
George E. Smith

The Nobel Prize in Physics 2009 was divided, one half awarded to Charles Kuen Kao *"for groundbreaking achievements concerning the transmission of light in fibers for optical communication"*,the other half jointly to Willard S. Boyle and George E. Smith *"for the invention of an imaging semiconductor circuit – the CCD sensor"*.

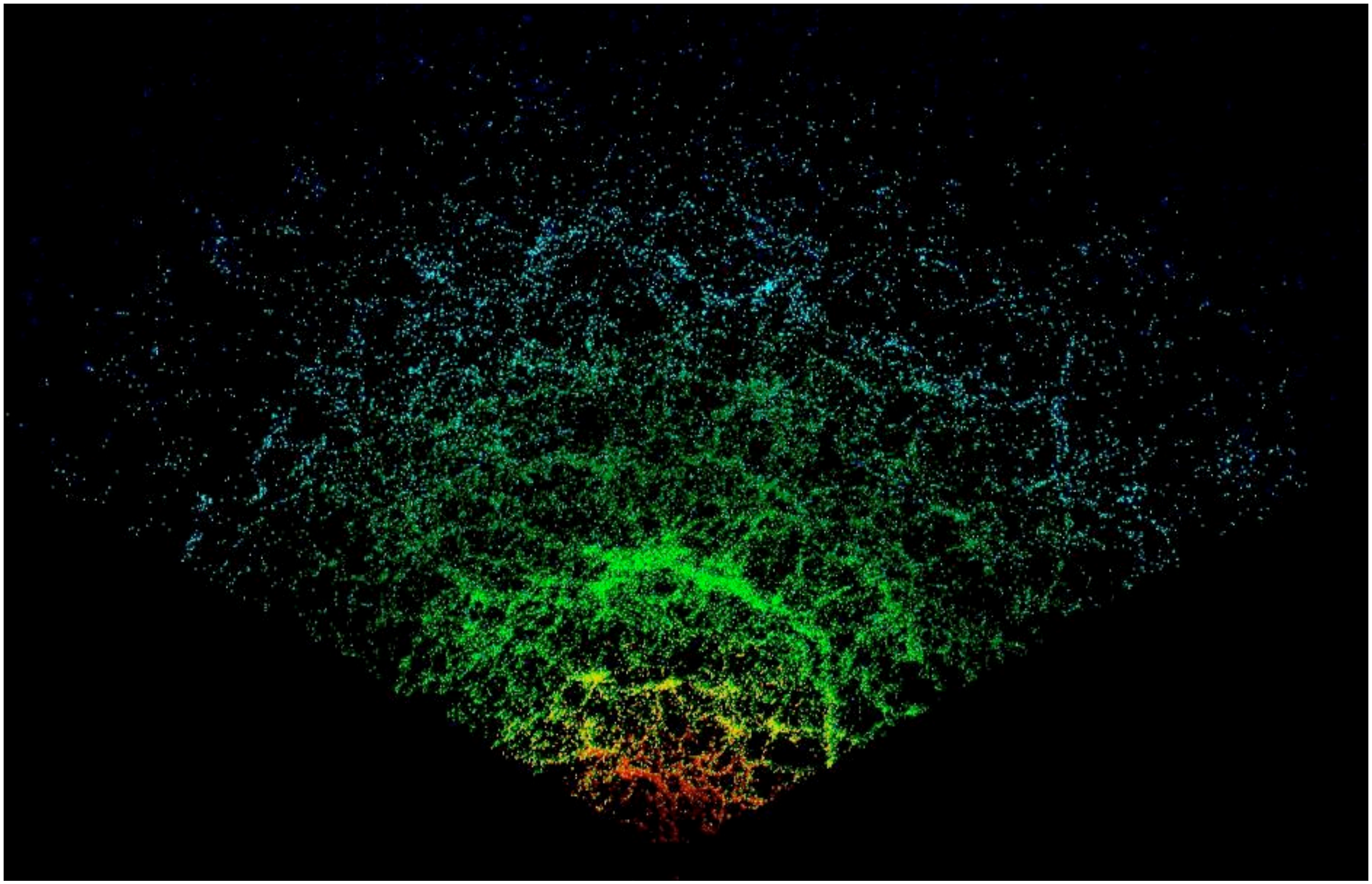
SDSS filter transmission curves

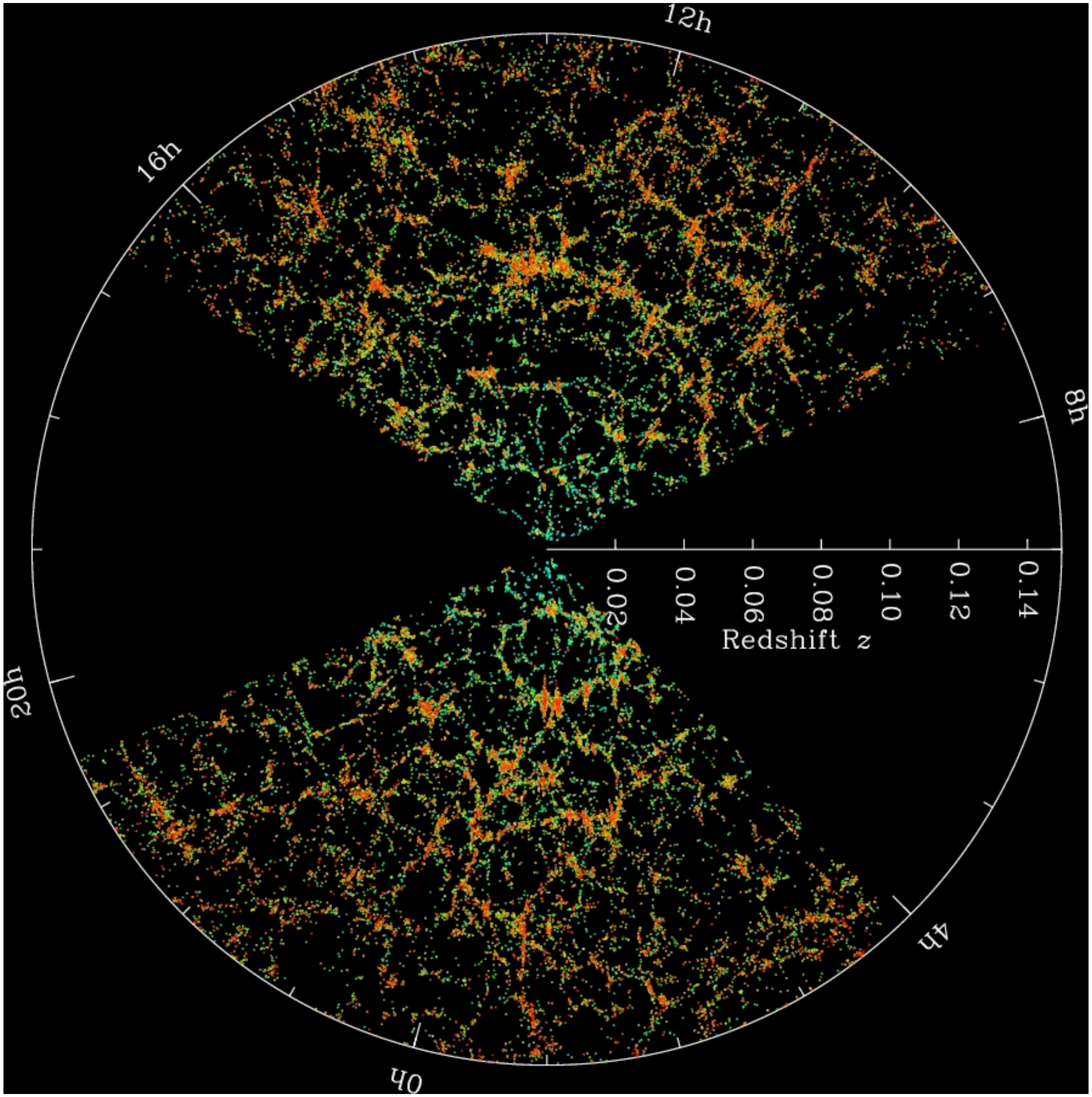


Filter	Average wavelength
u	3551 Å
v	4686 Å
r	6165 Å
i	7481 Å
z	8931 Å


Messier 33
NGC 604
SDSSIII
Southern Galactic Cap
Northern Galactic Cap

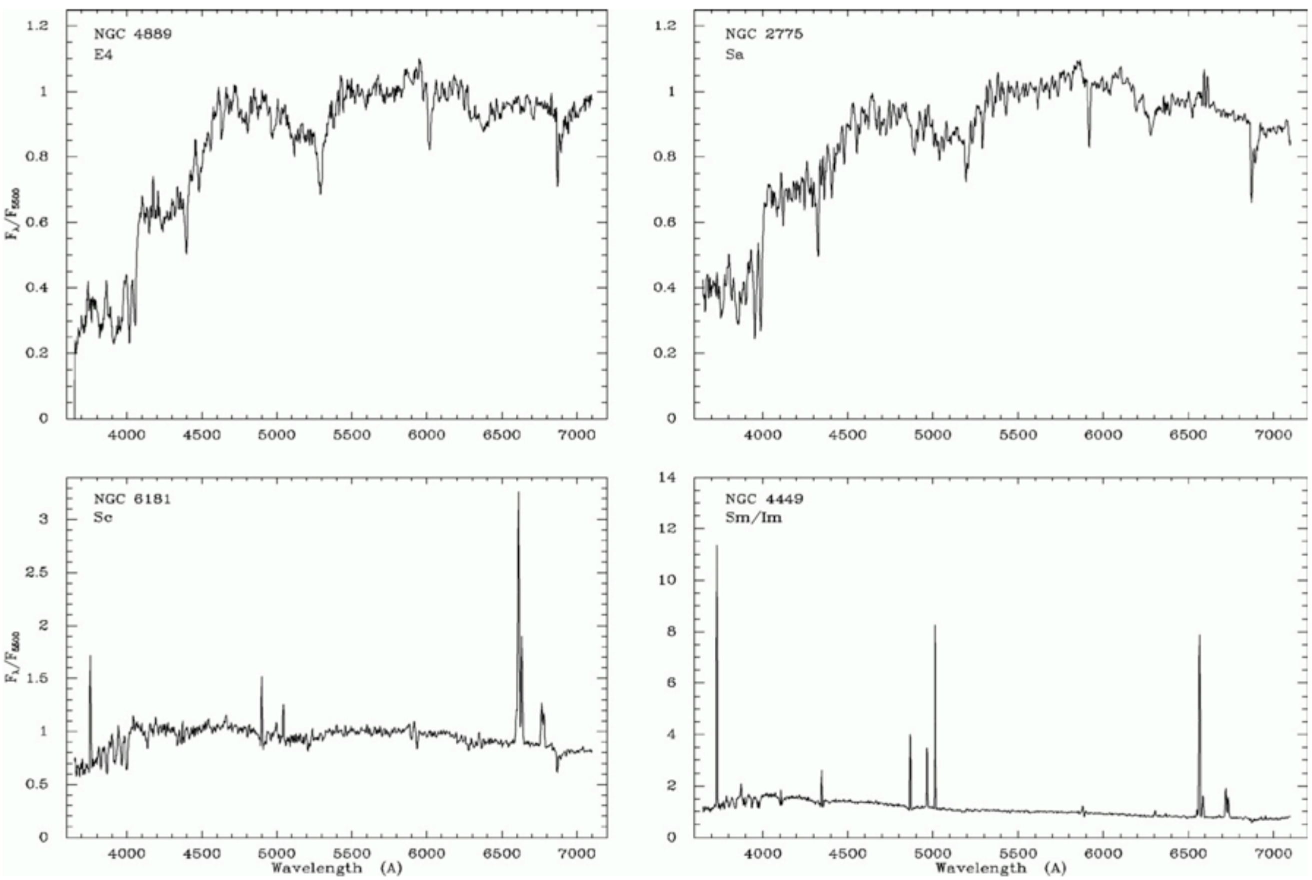
The Sloan Digital Sky Survey




12h
16h
8h
20h
4h
0h
0.02
0.04
0.06
0.08
0.10
0.12
0.14
Redshift z

Photometric redshifts

Galaxy spectra

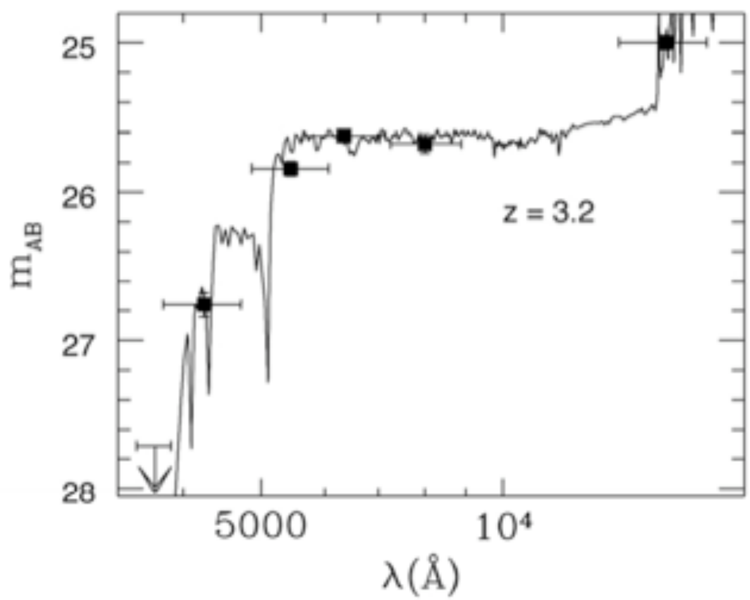


Galaxies show a variety of optical spectra which can be classified based on:

- strength of blue continuum
- composite stellar absorption features
- strength of nebular emission lines

Templates of the different classes can be easily built

Photometric redshifts

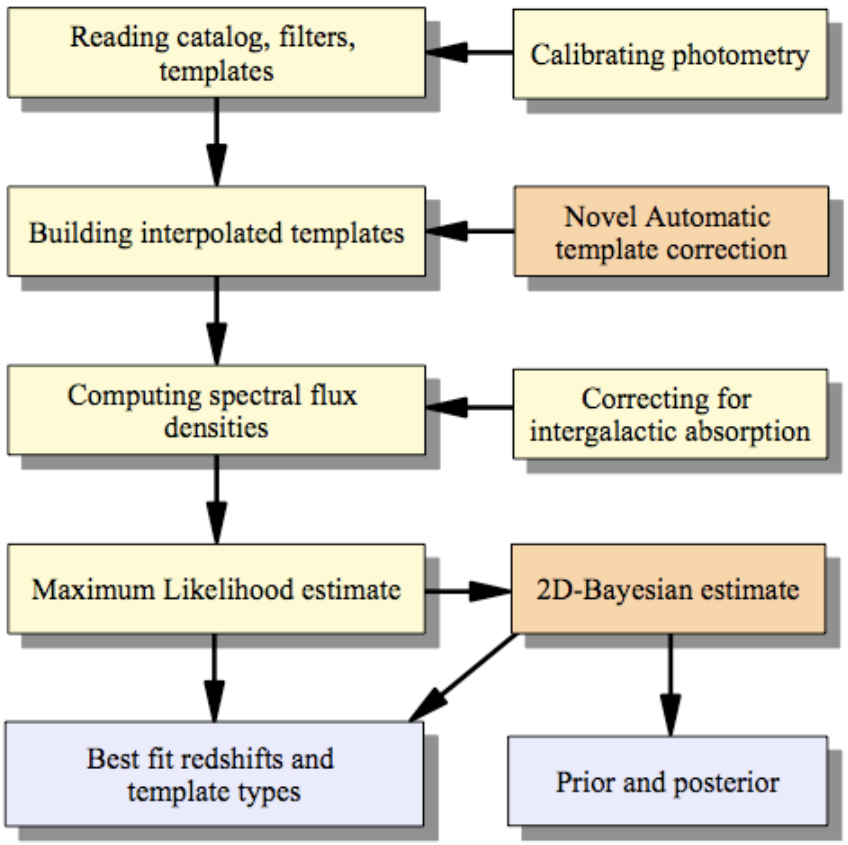


- Cheap estimate (in terms of observational time) of galaxy redshifts using multi-color, broad-band photometry instead of spectroscopy

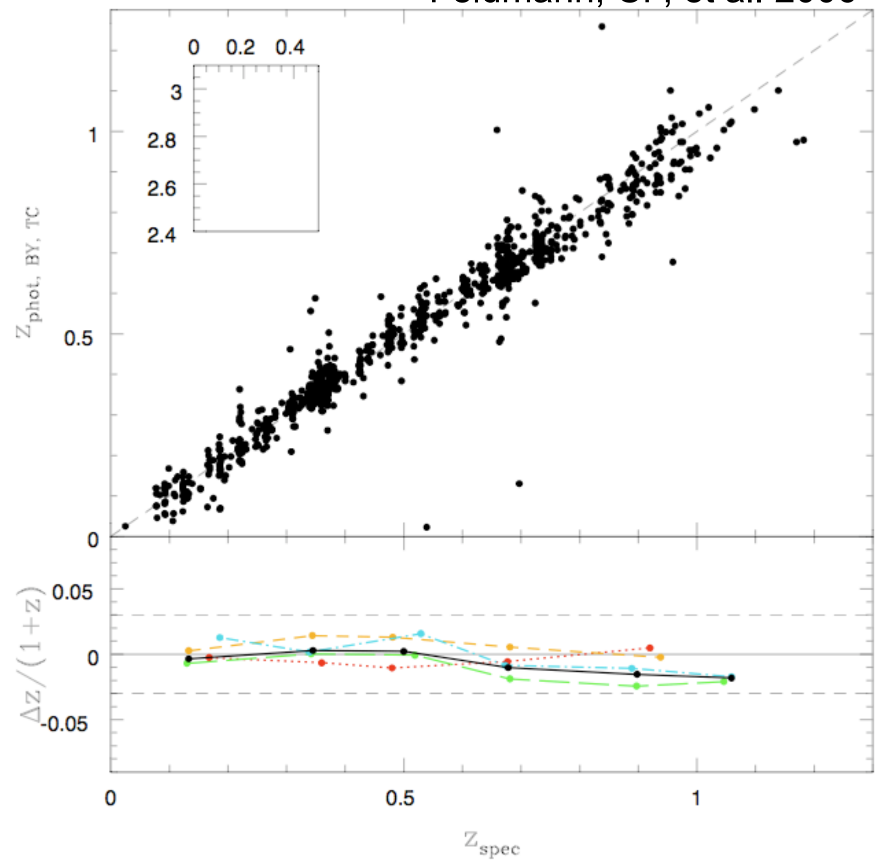
- It simply chooses the best-fitting redshifted spectrum out of a library of templates (either observationally or theoretically motivated)

- Rather than observing narrow spectral features of galaxy spectra (such as emission lines) this technique concentrate on broad features (such as spectral breaks) and the overall shape of a spectrum

Photometric redshifts



Feldmann, CP, et al. 2006



Characteristics of a galaxy survey

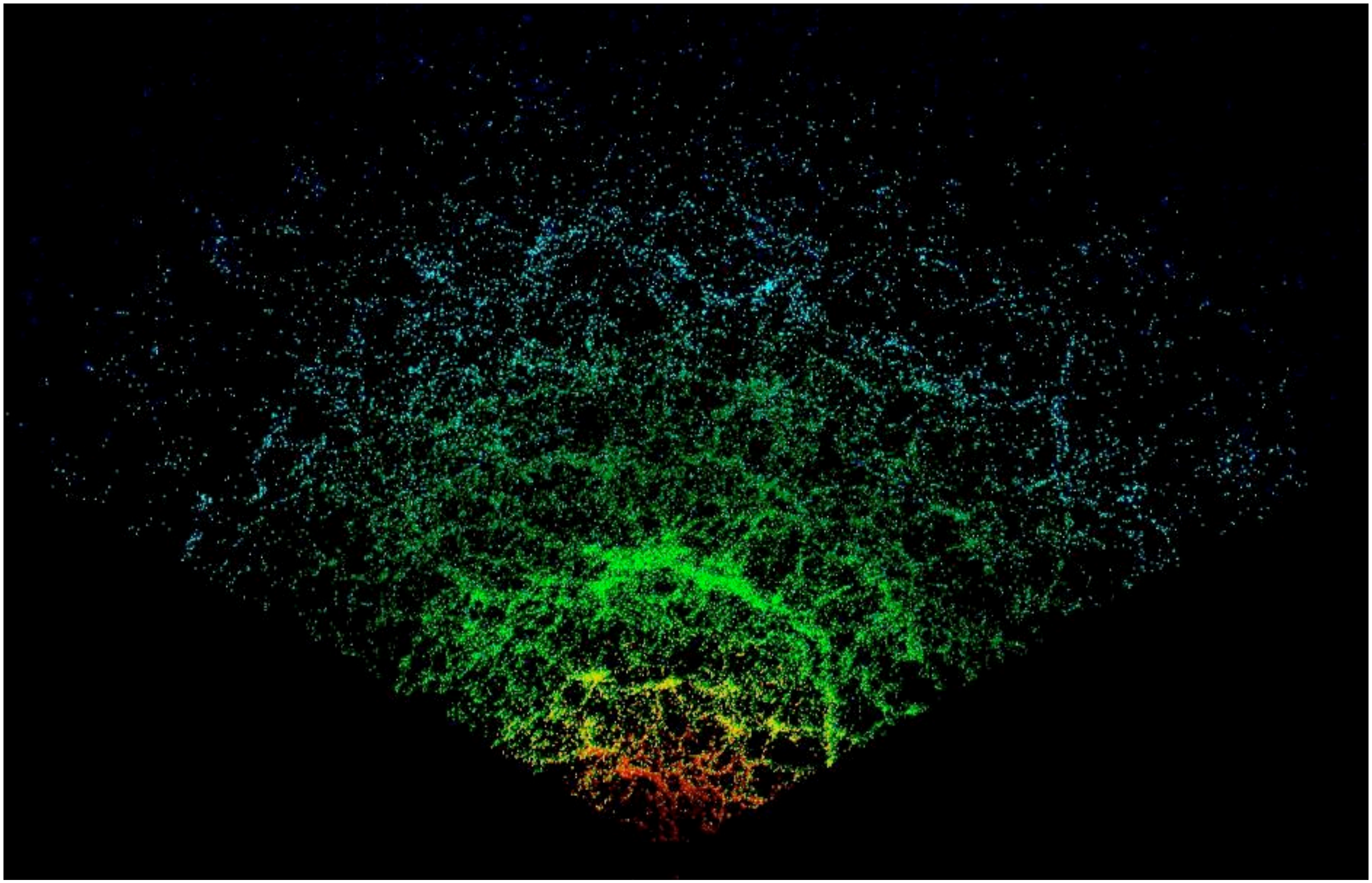
- Photometric (how many bands?) vs spectroscopic (What's the redshift completeness? What's the success rate?)

- All redshift surveys start from an parent (input) catalog. How are targets selected? Magnitude limited (in what band?) vs volume limited vs (pre) color selected

- How many square degrees? Down to what redshift? (wide vs deep) What geometry? (pencil beam vs wide angle)

- Before computing any statistic compute (or download) the selection function (density of galaxies as a function of redshift) and the completeness map (what fraction of objects are included in the survey as a function of position and redshift?). Otherwise you will fail miserably!

Science with galaxy surveys



What can you measure?

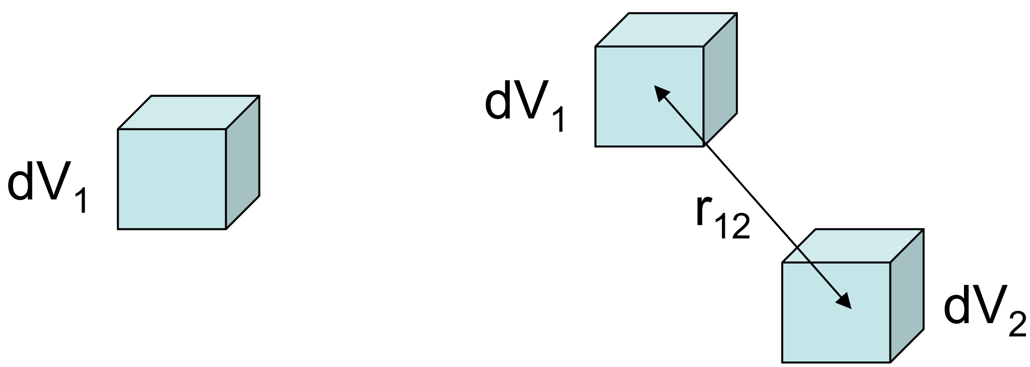
- Luminosity function and number densities
- Group and cluster catalogs (FoF, Voronoi, BCG)

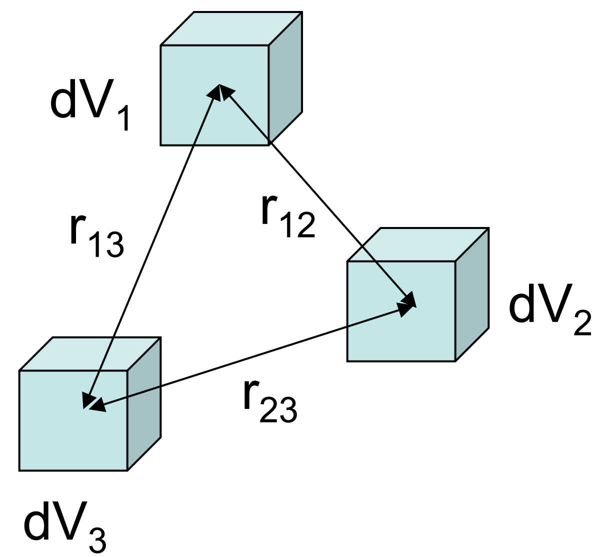
- The density field
- Reconstruct the linear density field (time machine)

- Counts in cells
- Measure 2-point, 3-point correlation function
- Measure power spectrum, bispectrum
- Topological invariants: Minkowski functionals (mean genus, void probability function)

Correlation functions

Consider a stationary point process with mean density n and write the probability of finding N points within N infinitesimal volume elements





$dP_1 = n\, dV_1$

$dP_{12} = n^2\,(1+\xi(r_{12}))\, dV_1\, dV_2$

$dP_{123} = n^3\,(1+\xi(r_{12})+\xi(r_{13})+\xi(r_{23})+\zeta(r_{12}, r_{13}, r_{23}))\, dV_1\, dV_2\, dV_3$

Power spectra

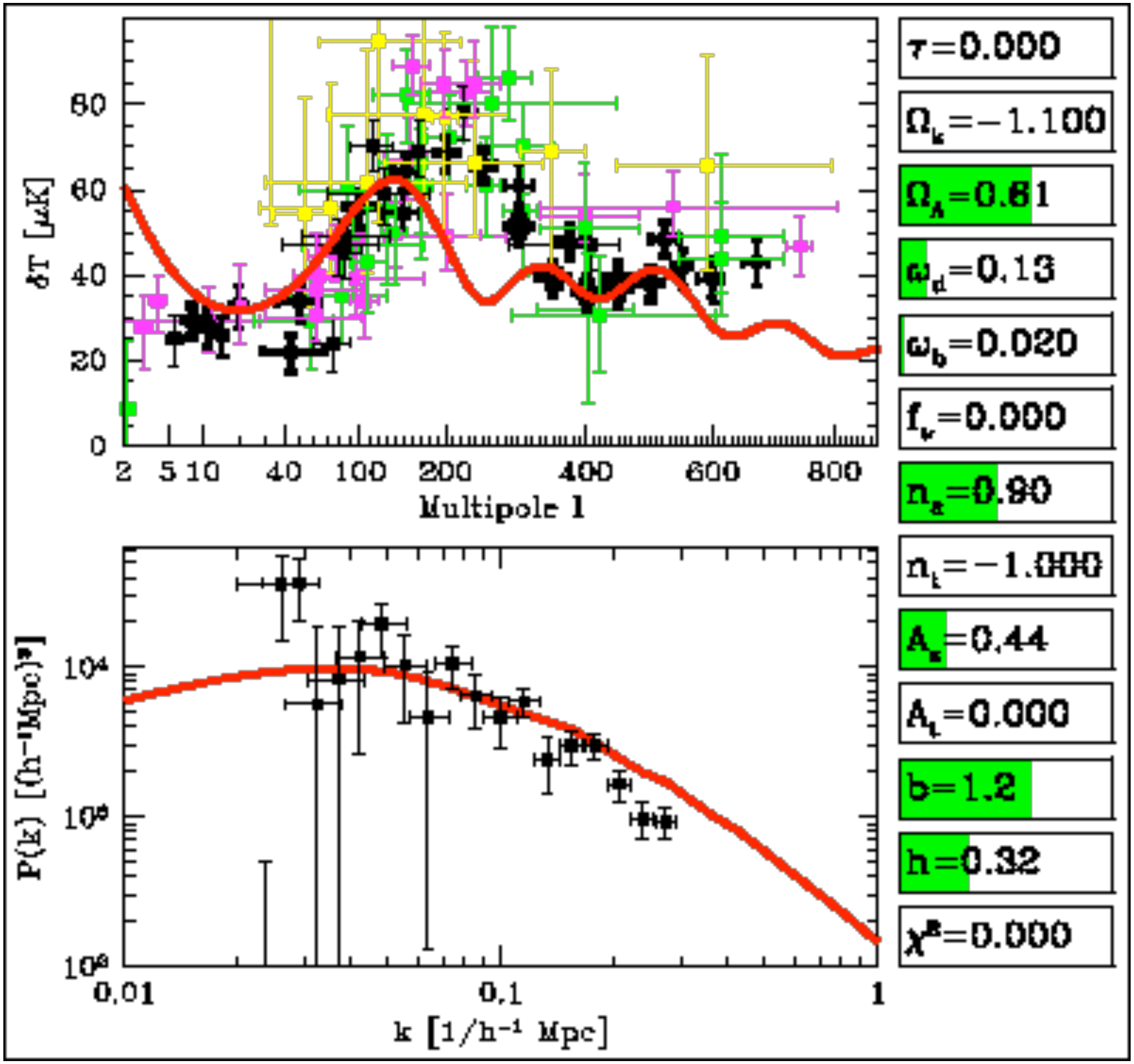
N-spectrum defined via the expectation value of the product of N+1 Fourier transforms of the overdensity field

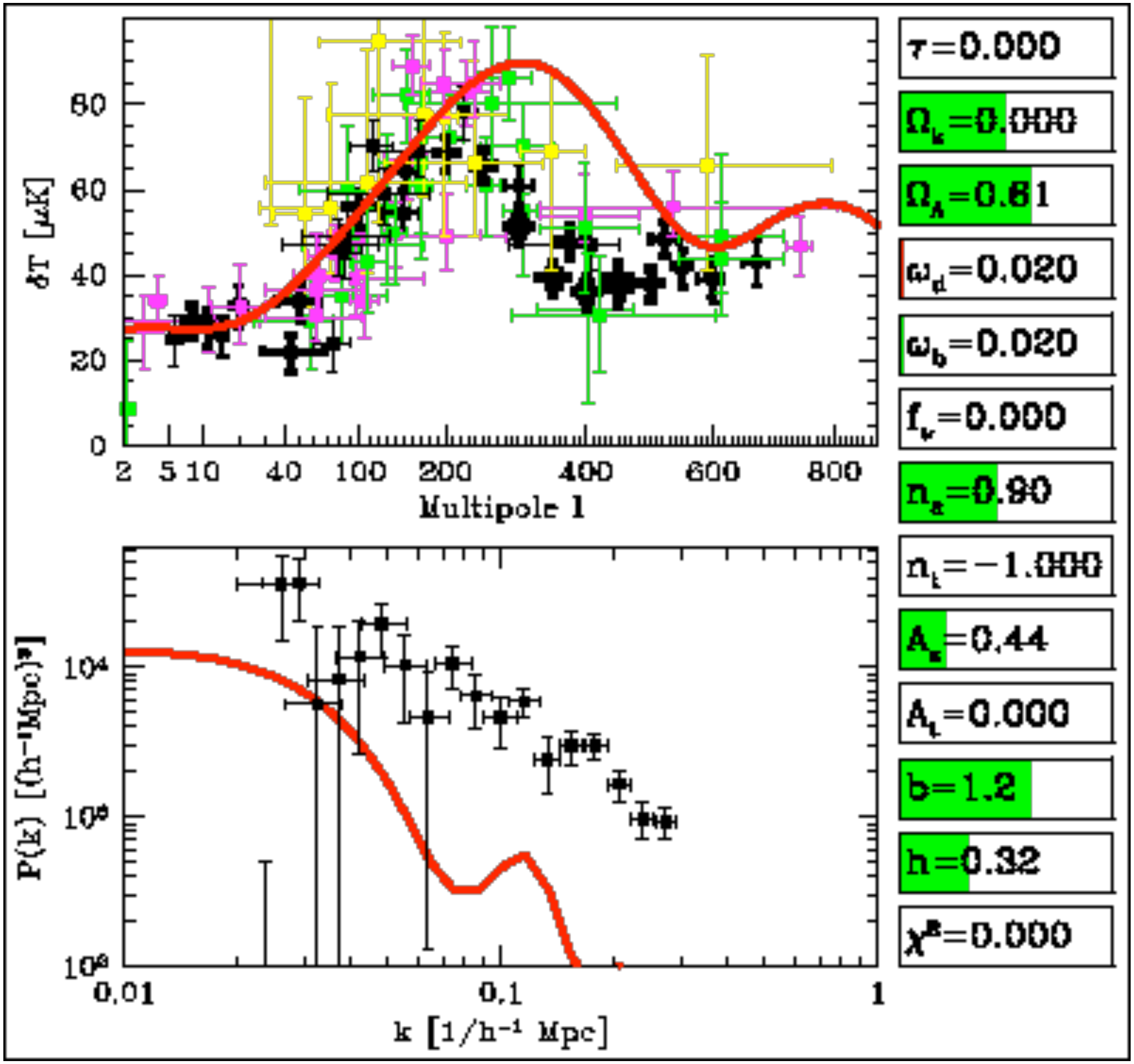
$$\langle \tilde{\delta}(\vec{k})\,\tilde{\delta}(\vec{q})\rangle = (2\pi)^3 P(k)\delta_D(\vec{k}+\vec{q})$$

$$\langle \tilde{\delta}(\vec{k})\,\tilde{\delta}(\vec{q})\,\tilde{\delta}(\vec{p})\rangle = (2\pi)^3 B(k,q,p)\delta_D(\vec{k}+\vec{q}+\vec{p})$$

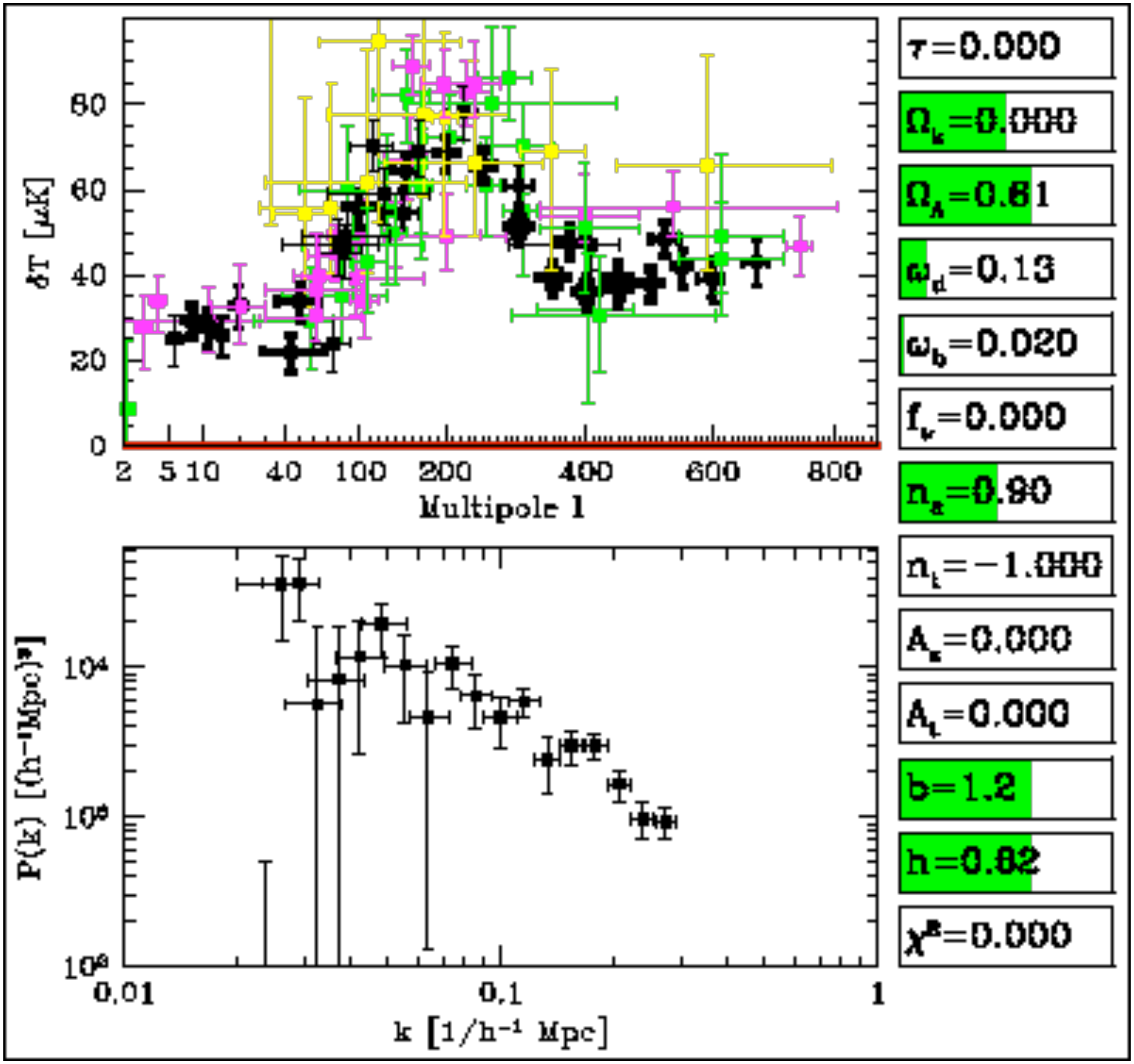
Wiener - Khintchine theorem:

$$\xi(r) = \frac{1}{2\pi^2}\int_0^\infty k^2 P(k)\, j_0(kr)\,dk$$

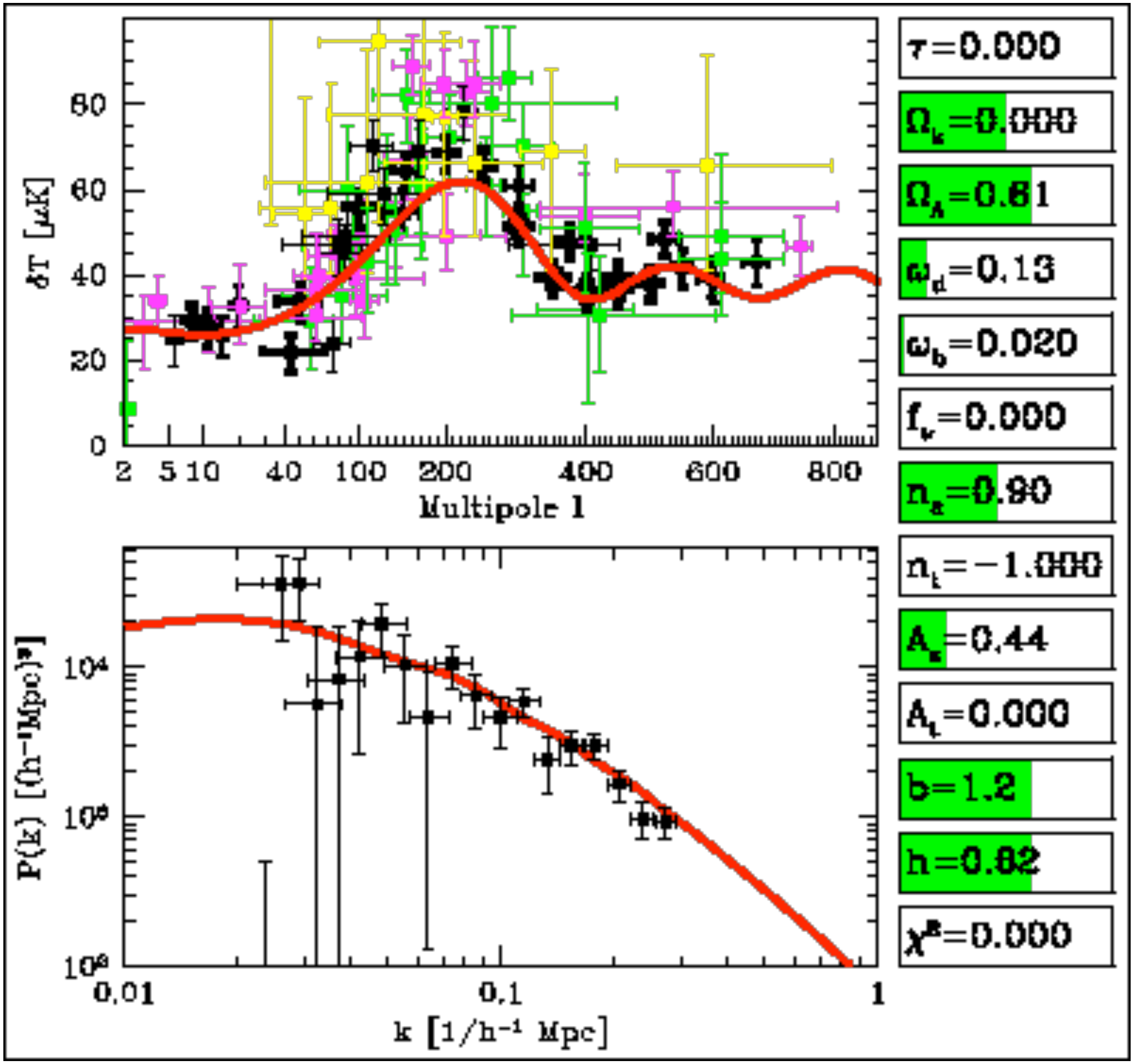

δT [μK]
Multipole l
$P(k)$ [$(h^{-1}\mathrm{Mpc})^3$]
k [1/h^{-1} Mpc]
τ=0.000
Ω_k=−1.100
Ω_Λ=0.61
ω_d=0.13
ω_b=0.020
f_ν=0.000
n_s=0.90
n_t=−1.000
A_s=0.44
A_t=0.000
b=1.2
h=0.32
χ^2=0.000



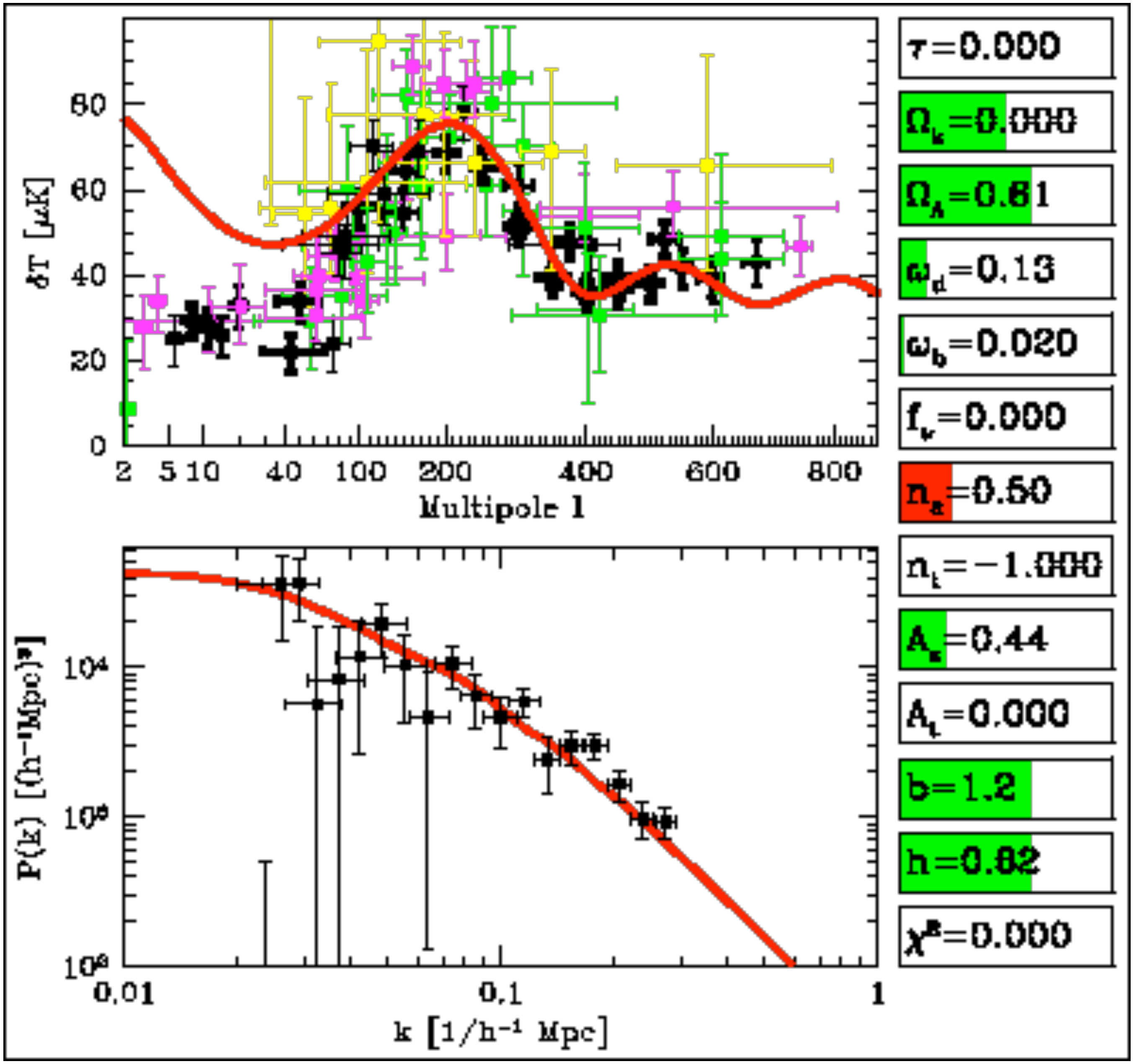
δT [μK]
Multipole l
P(k) [(h⁻¹Mpc)³]
k [1/h⁻¹ Mpc]
τ=0.000
Ωk=0.000
ΩΛ=0.61
ωd=0.020
ωb=0.020
fν=0.000
ns=0.90
nt=−1.000
As=0.44
At=0.000
b=1.2
h=0.32
χ²=0.000



δT [μK]
Multipole l
P(k) [(h⁻¹Mpc)³]
k [1/h⁻¹ Mpc]
τ=0.000
Ω_k=0.000
Ω_Λ=0.61
ω_d=0.13
ω_b=0.020
f_ν=0.000
n_s=0.90
n_t=−1.000
A_s=0.000
A_t=0.000
b=1.2
h=0.82
χ²=0.000



δT [μK]
Multipole l
P(k) [(h⁻¹Mpc)³]
k [1/h⁻¹ Mpc]
τ=0.000
Ω_k=0.000
Ω_Λ=0.81
ω_d=0.13
ω_b=0.020
f_ν=0.000
n_s=0.90
n_t=−1.000
A_s=0.44
A_t=0.000
b=1.2
h=0.82
χ²=0.000


δT [μK]
Multipole l
P(k) [(h⁻¹Mpc)³]
k [1/h⁻¹ Mpc]
τ=0.000
Ω_k=0.000
Ω_Λ=0.61
ω_d=0.13
ω_b=0.020
f_ν=0.000
n_s=0.50
n_t=−1.000
A_s=0.44
A_t=0.000
b=1.2
h=0.62
χ²=0.000

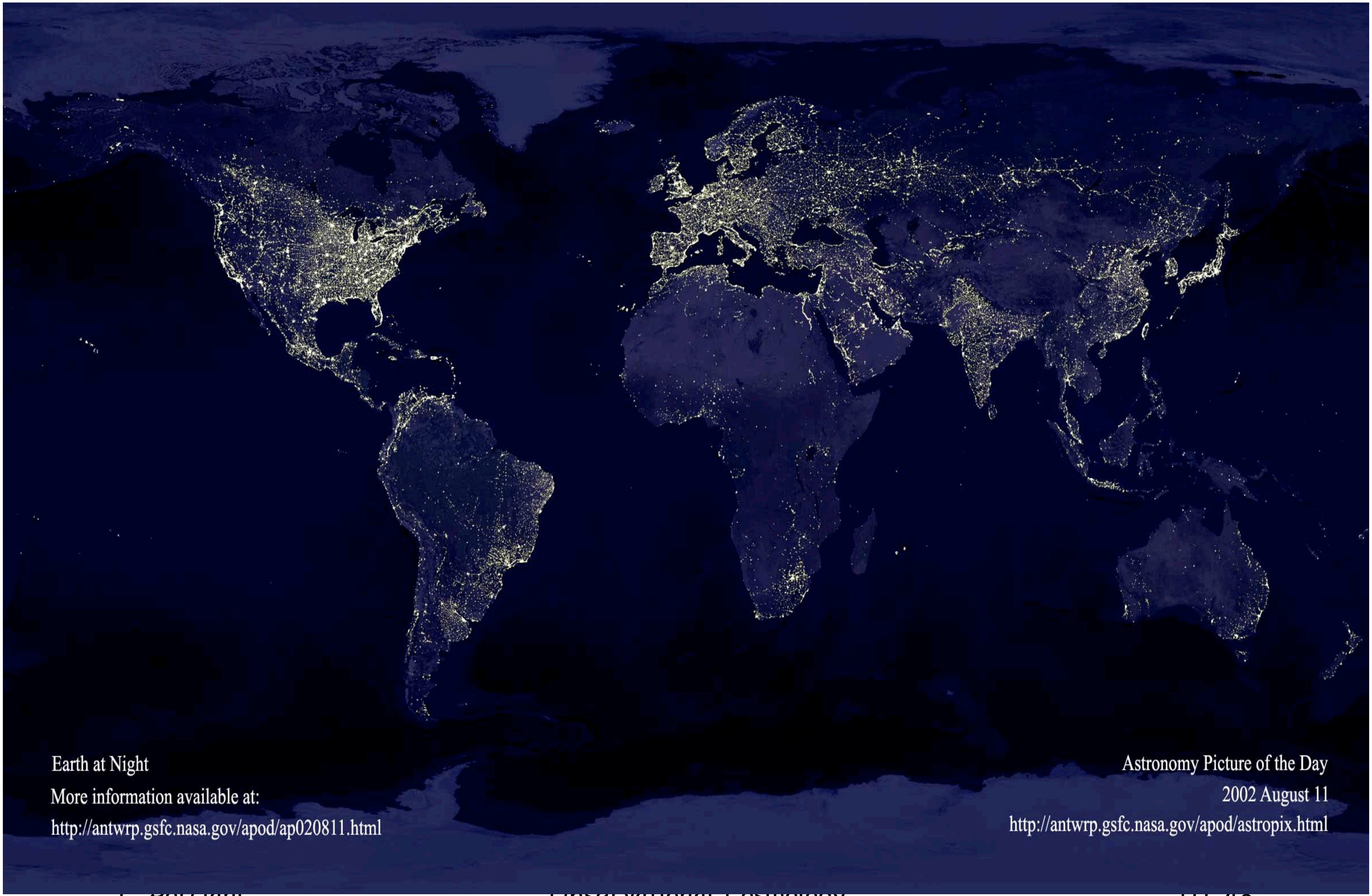
Complications…

I- Galaxy biasing

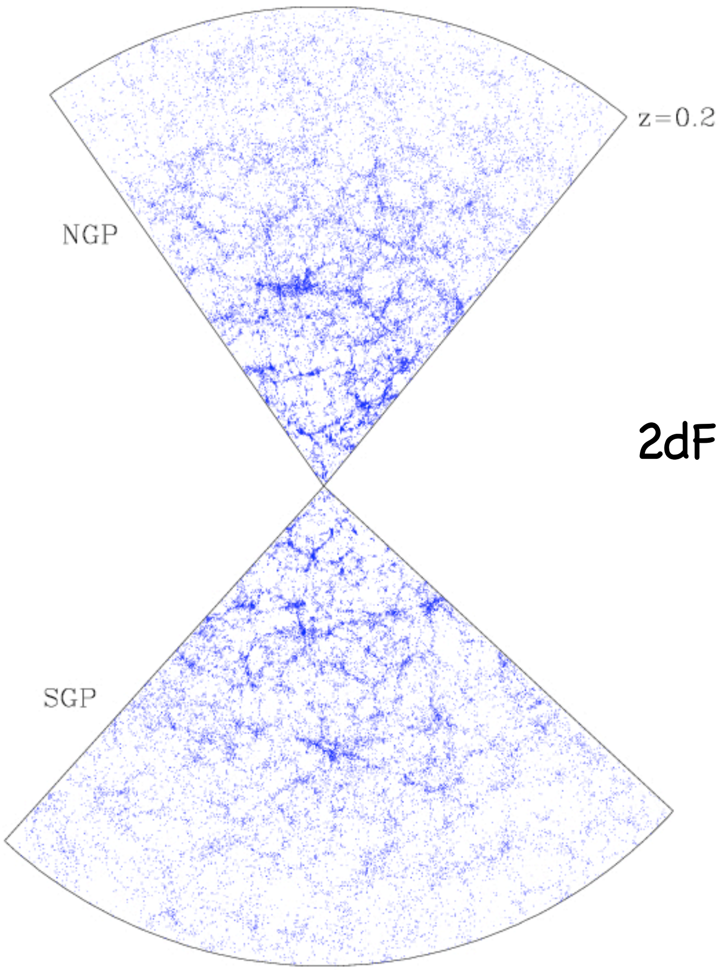
(where cosmology meets astrophysics)

Light does not trace mass

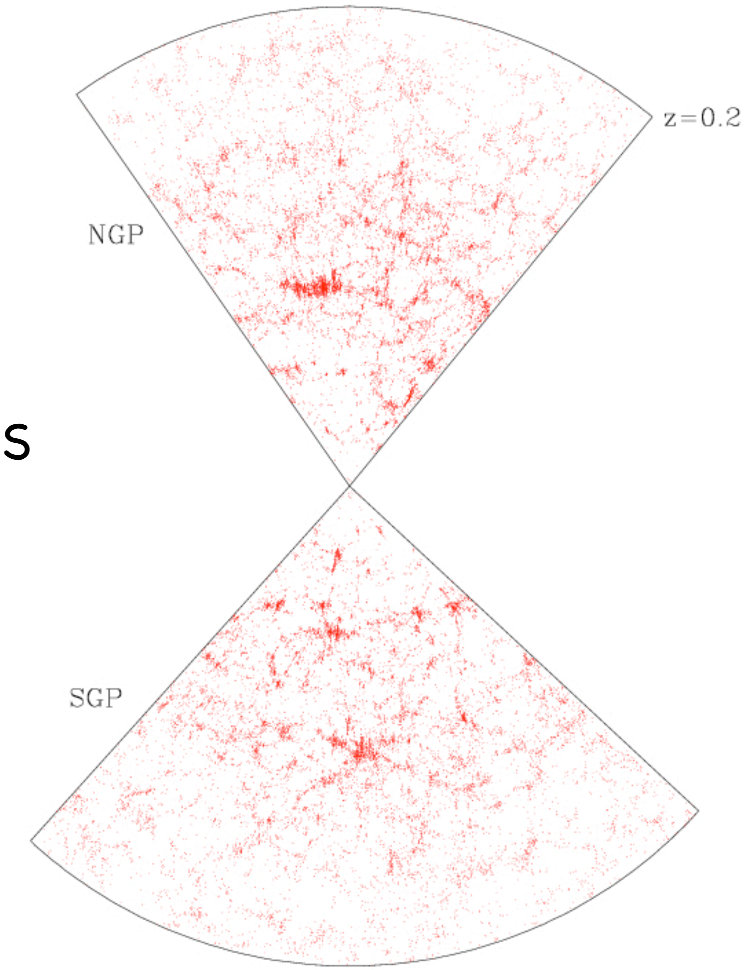
- We observe galaxies and use them to map the cosmic web
- Theory, however, predicts the mass distribution
- So far we have a limited understanding of the galaxy formation process (a complicated (g)astrophysical problem)
- It is clear, anyway, that galaxies form in special regions of the density field with different statistical properties


Earth at Night
More information available at:
http://antwrp.gsfc.nasa.gov/apod/ap020811.html
Astronomy Picture of the Day
2002 August 11
http://antwrp.gsfc.nasa.gov/apod/astropix.html

Galaxy biasing exists

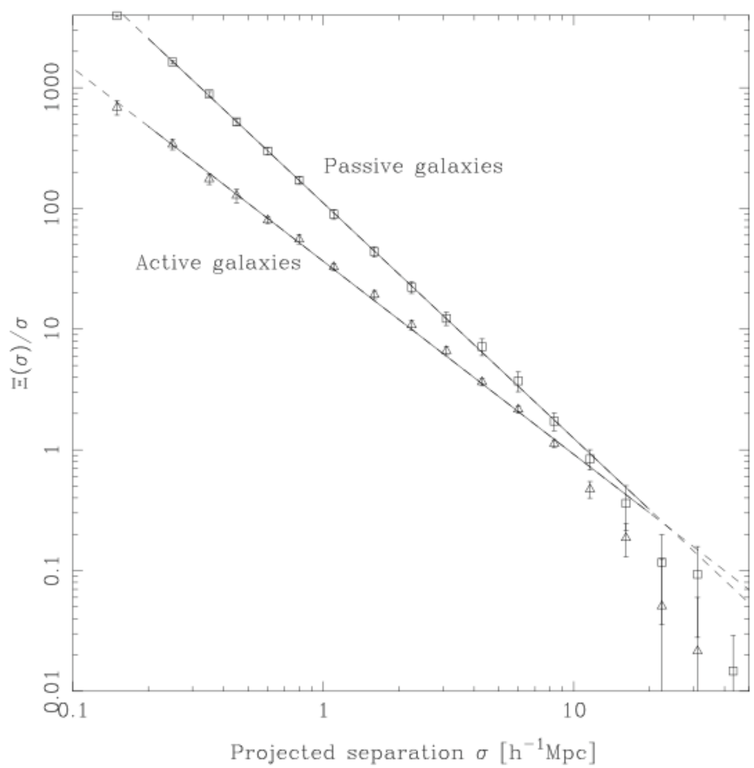


2dF GRS

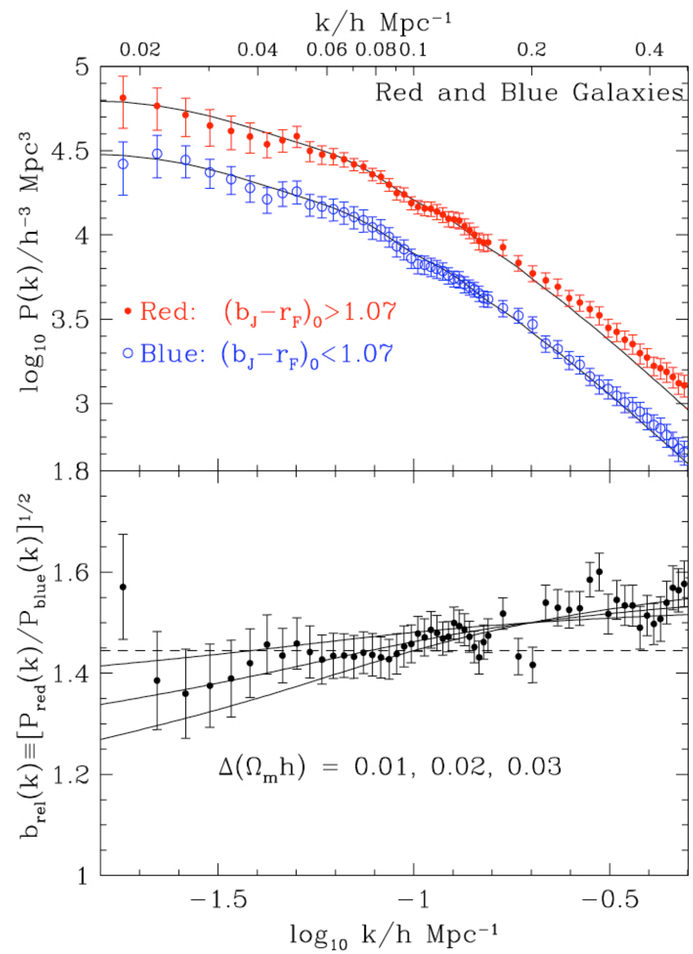


Galaxy biasing exists

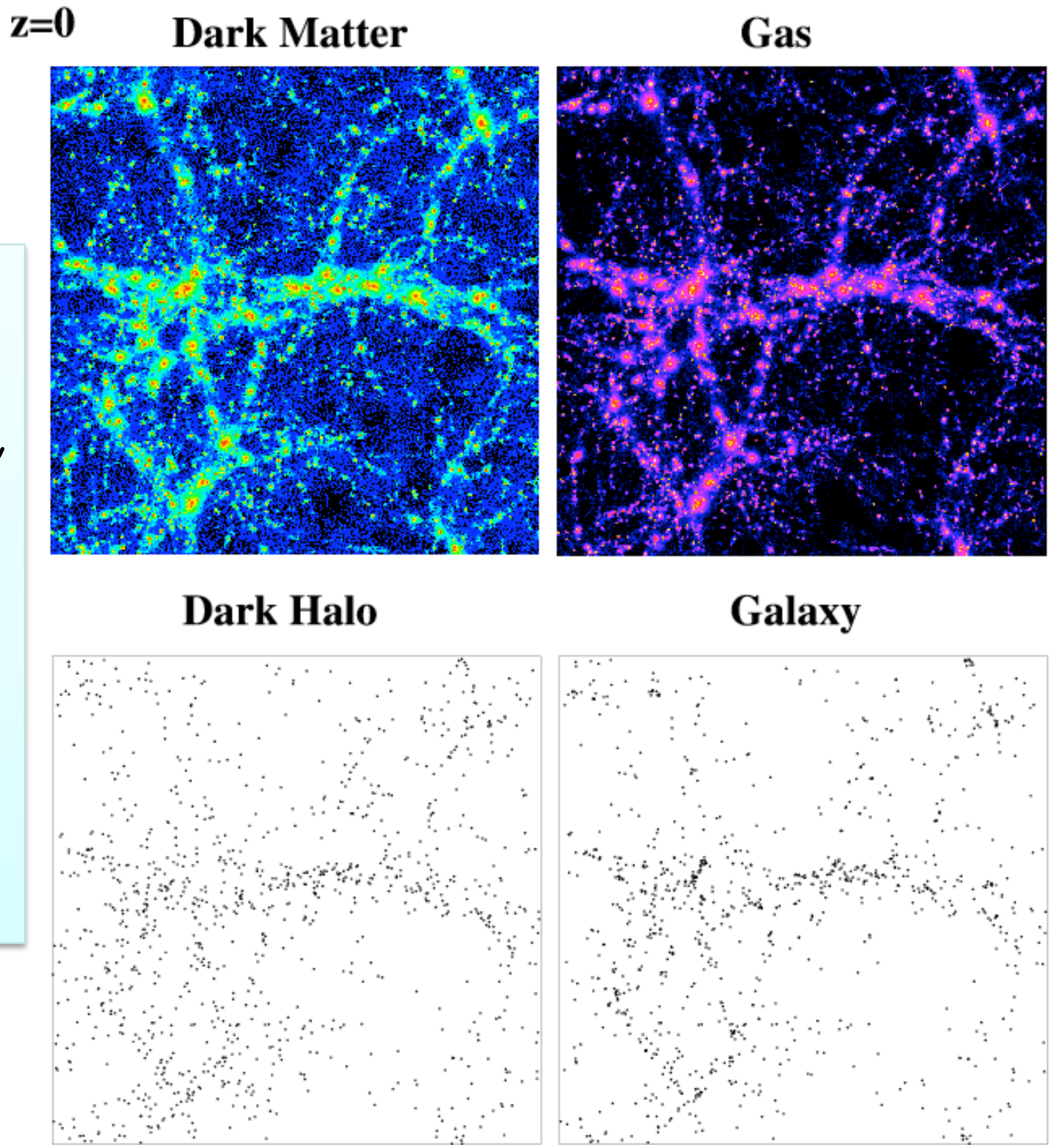
Madgwick et al. 2003



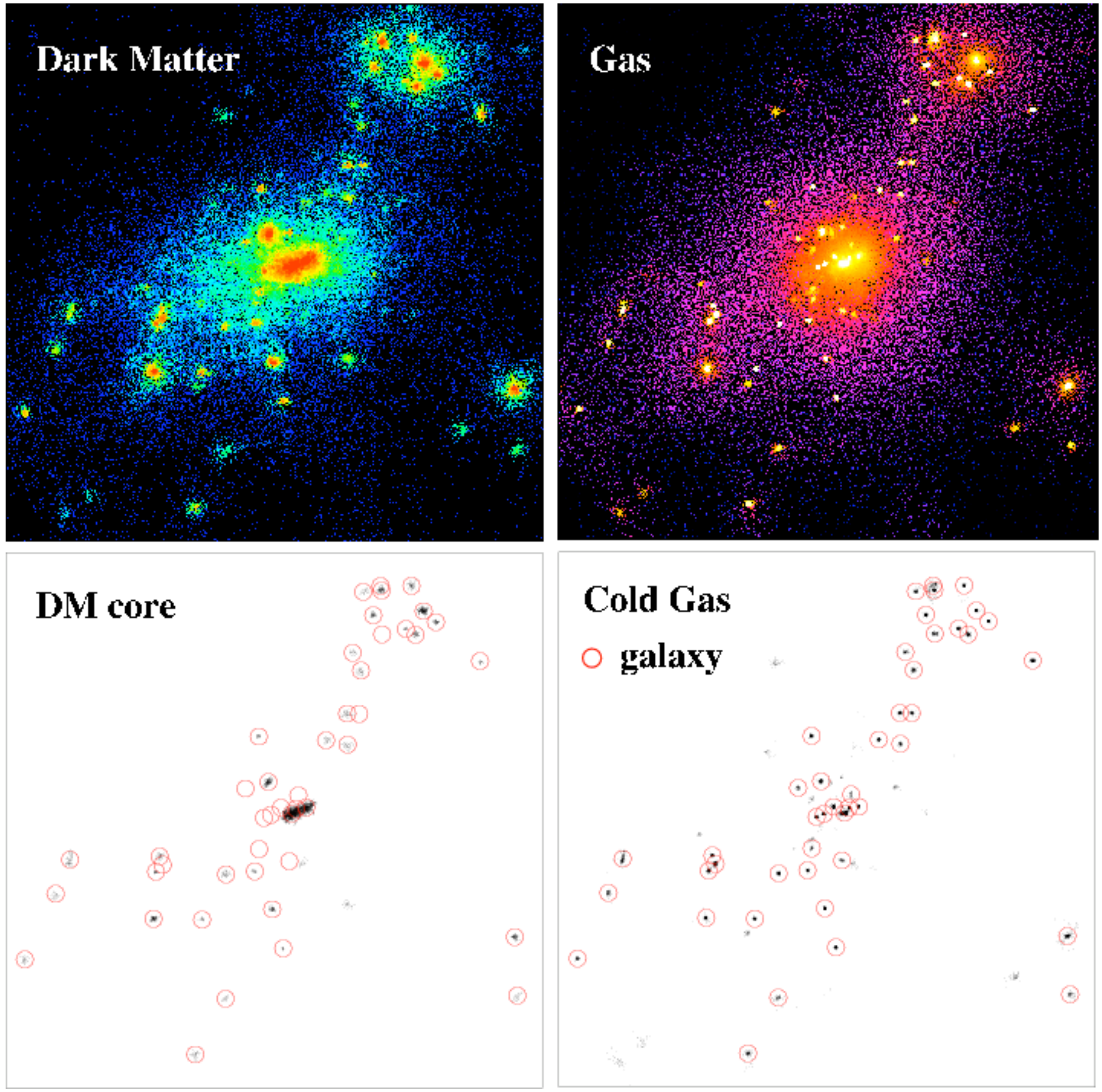
Cole et al. 2005



To better understand how to model galaxy biasing, we can compare the distribution of different tracers of the large-scale structure in a numerical simulation



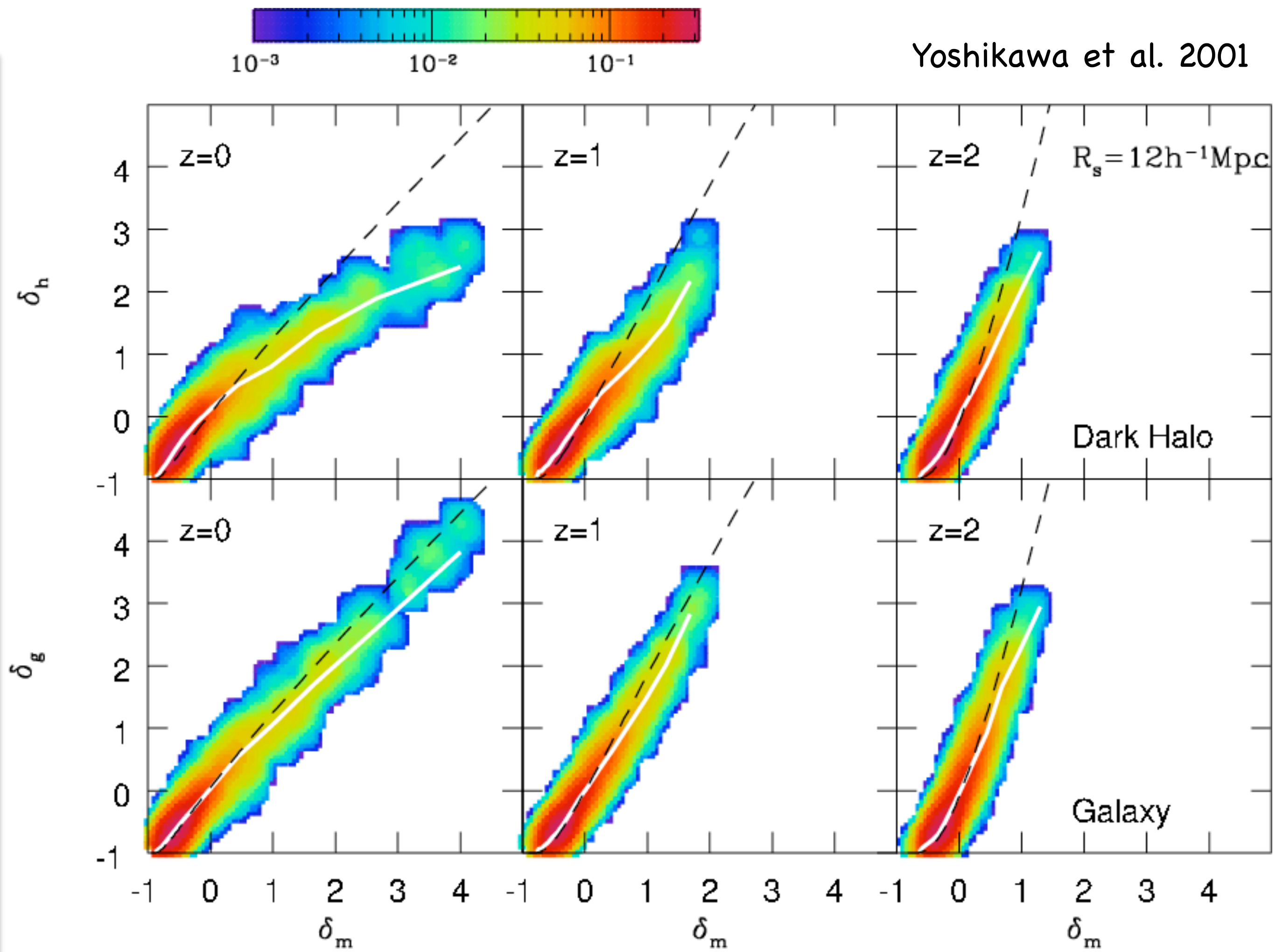
Zooming in a galaxy cluster



A local biasing scheme?

Smooth the density distributions of different tracers on the scale R_s and plot them against the mass density (also smoothed) at the same spatial location. Apart from some scatter there appears to be a deterministic relation.

Yoshikawa et al. 2001



A local biasing scheme

- Therefore, we can write that

Scatter, noise

$$\delta_g(x) = f[\delta_m(x)] + \varepsilon(x)$$

- And, for a large smoothing scale, for which δ_m<<1 we can Taylor expand the deterministic part and write (neglecting the scatter)

$$\delta_g(x) \approx b_0 + b_1\delta_m(x) + \frac{b_2}{2}[\delta_m(x)]^2 + \dots$$

- This implies that

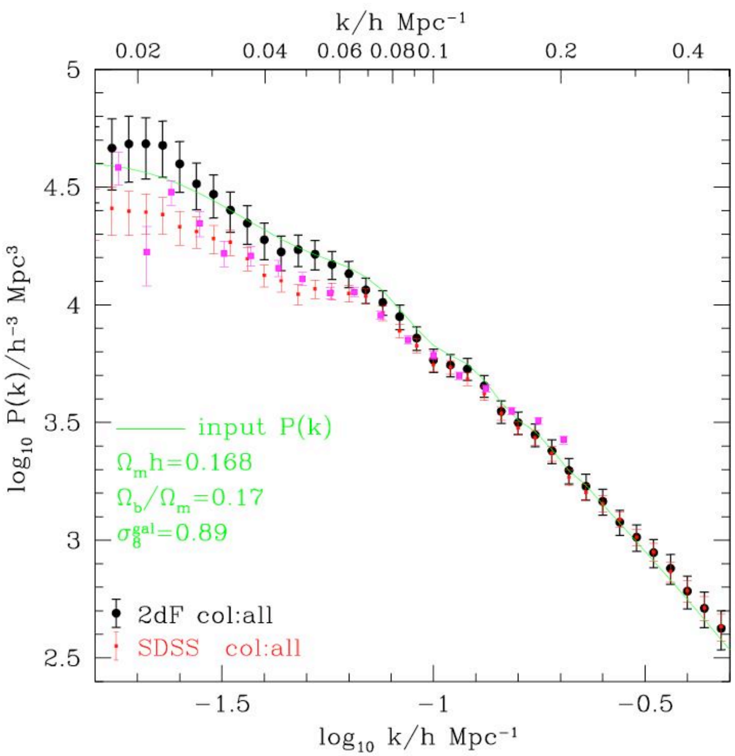
$$P_g(k) = b_1^2 P_m(k) + g[b_1, b_2, P_m(k)]$$

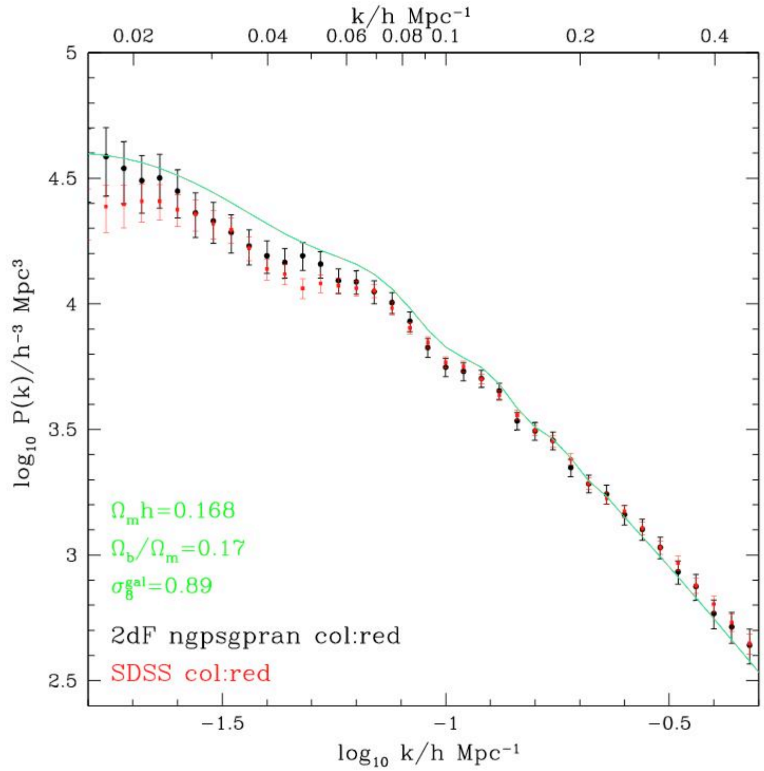
Linear biasing term: changes the amplitude

Non-linear biasing term: changes the shape

Power spectrum and galaxy selection 2dF GRS vs SDSS

Cole et al. 2005

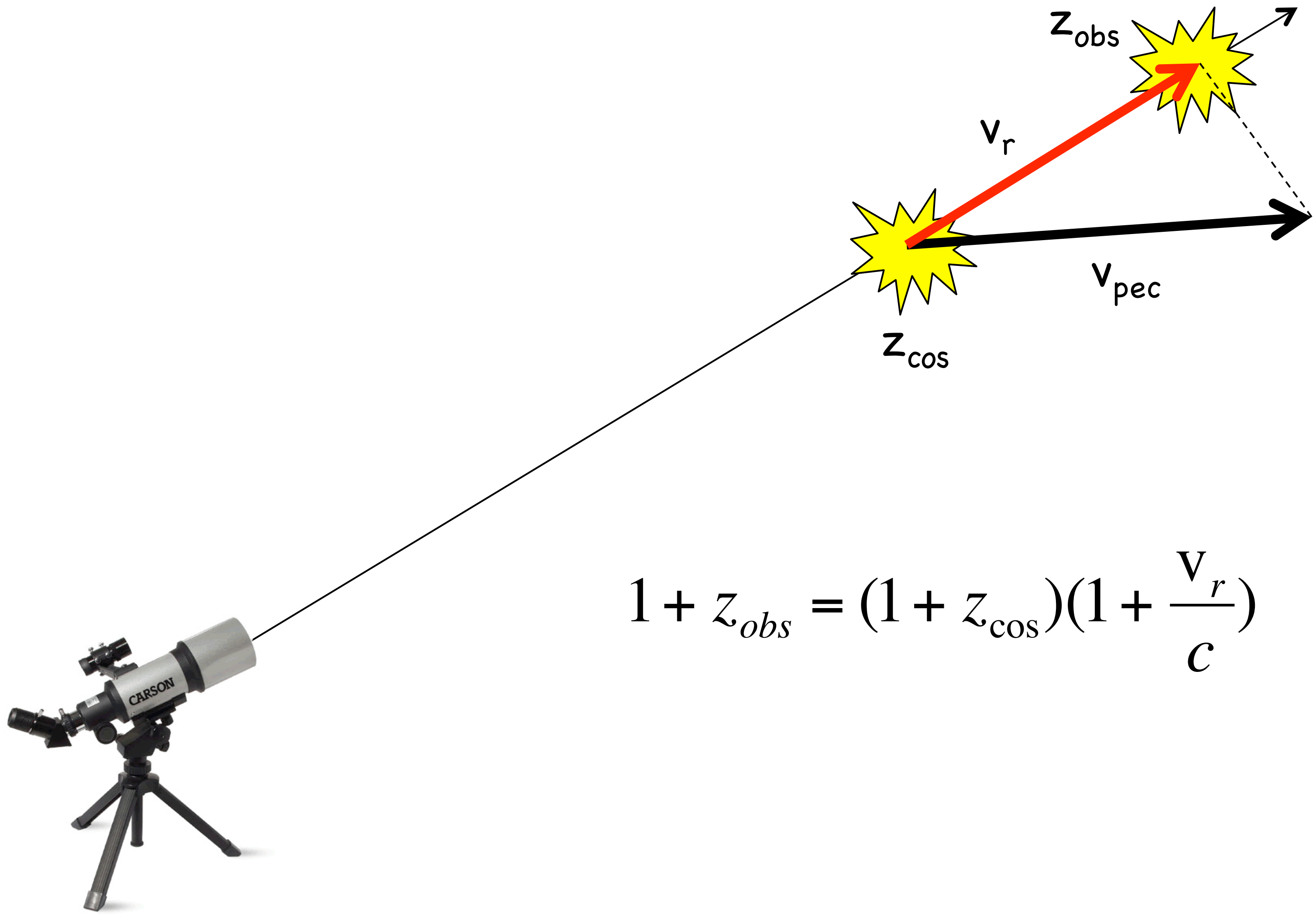




Complications…

II- Redshift-space distortions

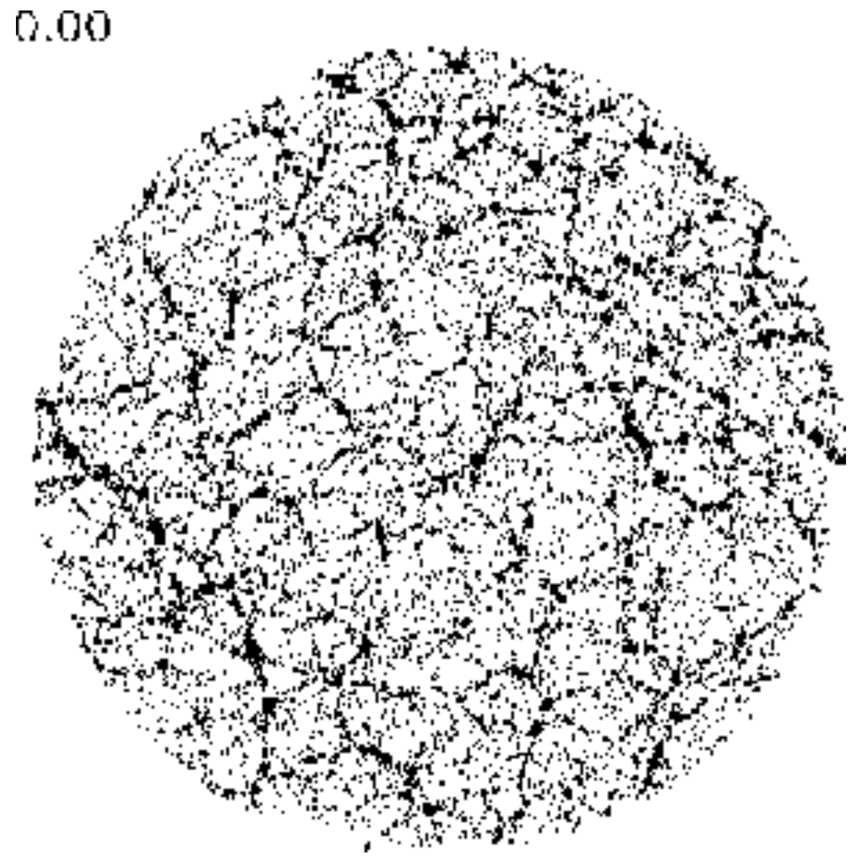
Redshift-space distortions



$$1 + z_{obs} = (1 + z_{\cos})(1 + \frac{v_r}{c})$$

Redshift-space distortions

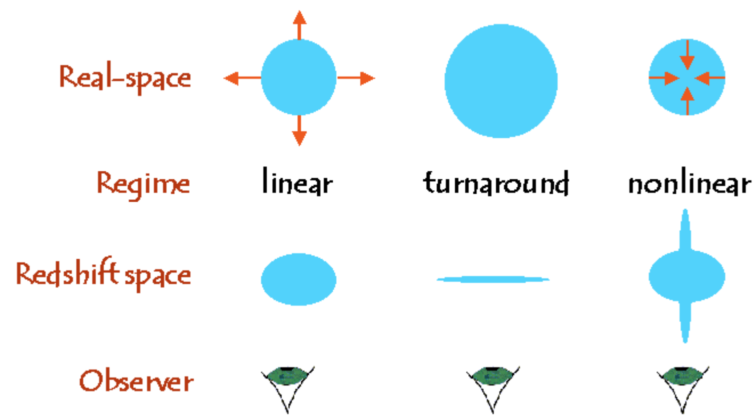
- Density fluctuations generate velocities on top of the global cosmic expansion
- The observed redshift of a galaxy includes a radial Doppler component: $1+z_{obs} = (1+z_{cos})\,(1+v_r/c)$
- Since we use the redshift to infer the distance to a galaxy, our 3D maps of the universe are "distorted".



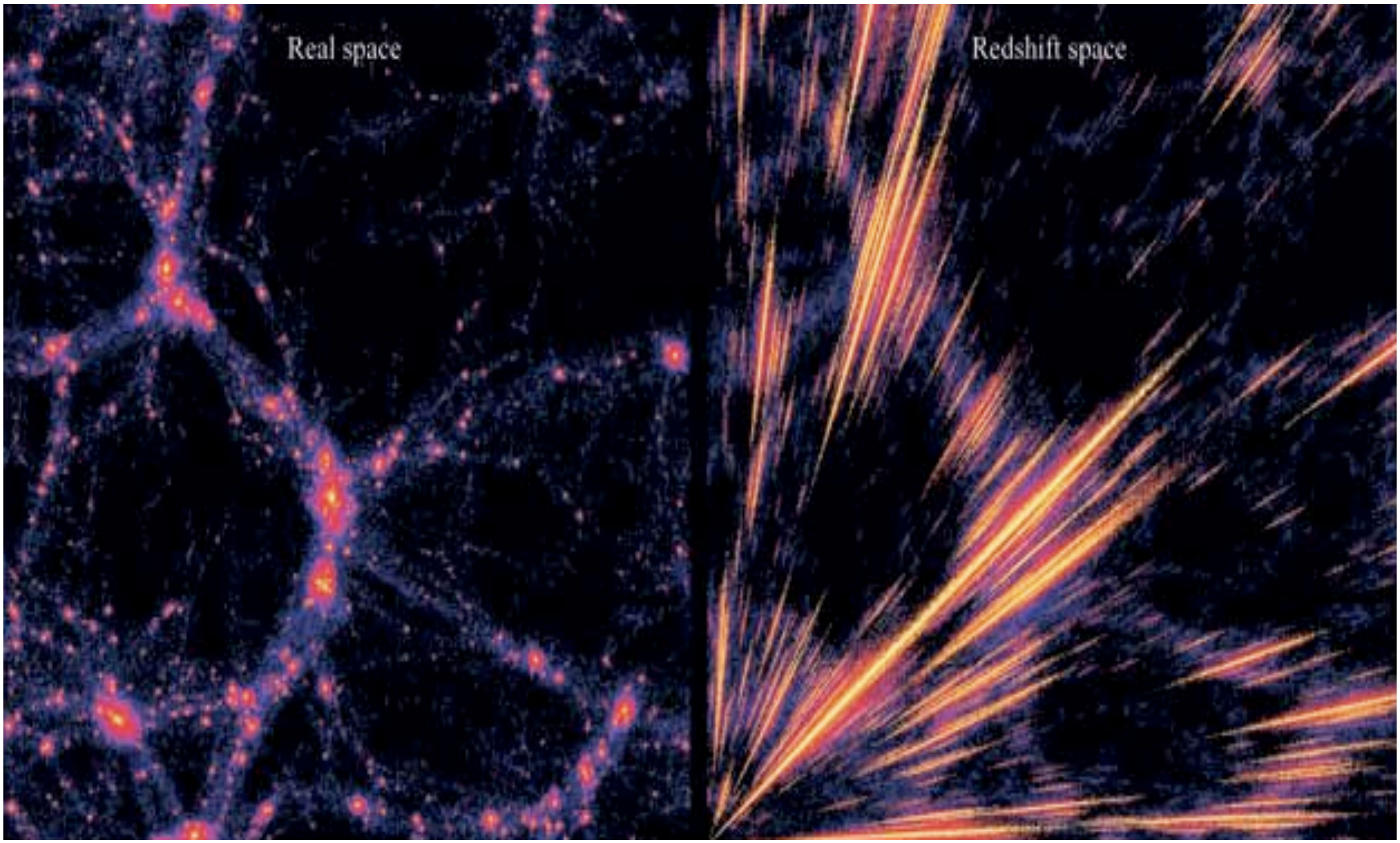
Redshift distortions

- Fingers of God: Radial stretching pointing towards the observer. They come about because of random velocities in clusters of galaxies

- Large overdensities lead to a coherent infall motion: walls appear denser and thicker, voids bigger and emptier



A closer look



Consequences of RS-distortions

- Bad news: we will never be able to measure the actual galaxy distribution

- Good news: the size of the distortions depends on cosmology. We can use them to learn something about the universe. Recall from cosmology class:

$$\nabla \cdot \mathrm{v} = -\frac{\partial \delta_m}{\partial t} \text{ (linearized continuity equation)}$$

$$\delta_{g,s}(\vec{k}) = \delta_{g,r}(\vec{k}) \left(1 + \beta\mu^2\right), \quad \beta \approx \frac{\Omega_m^{0.55}}{b_1}, \quad \mu = \cos(\theta_{\vec{r}\vec{k}})$$

The power spectrum in redshift space

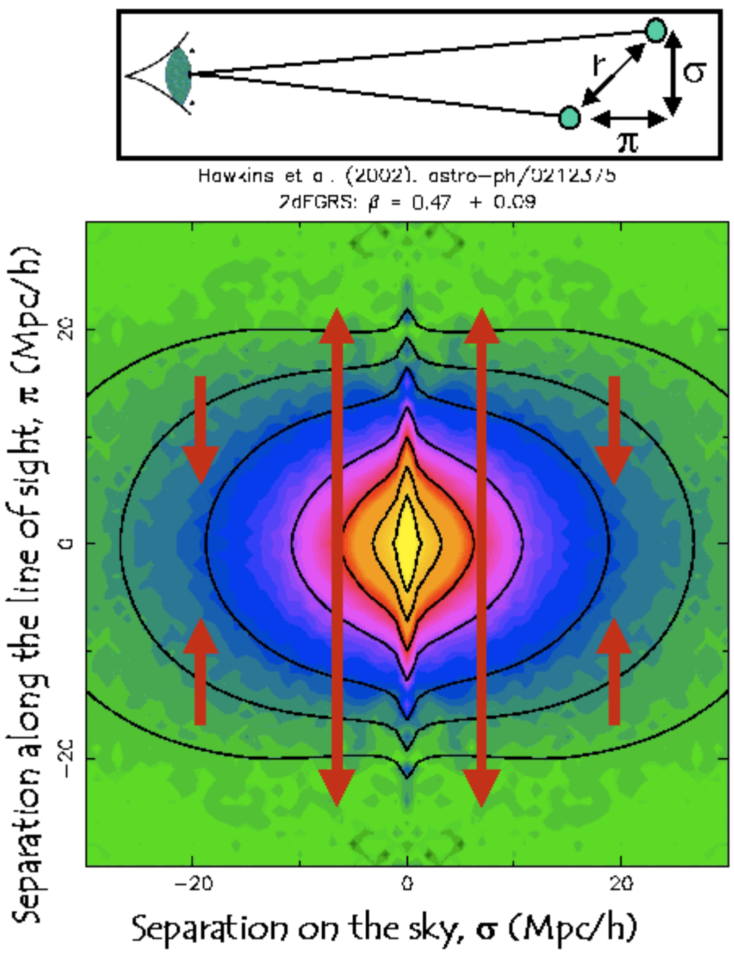
Boost in the average power + anisotropic terms

$$\frac{P_s(k)}{P(k)} = \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right) + \left(\frac{4}{3}\beta + \frac{4}{7}\beta^2\right) L_2(\mu) + \frac{8}{35}\beta^2 L_4(\mu)$$

where $L_i(x)$ denotes the Legendre polynomial of order i

The ratio of the quadrupole to monopole amplitudes is a monotonic function of β that rises from 0 at $\beta=0$ to just over unity at $\beta=1$. Redshift distortions can then be used to measure β and, if one already knows b_1, provide a measure of Ω_m

Anisotropic correlation function



- Redshift distortions also generate anisotropies in the 2-point correlation function
- The finger-of-god effect can be used to determine the velocity dispersion (and thus the typical mass) of the galaxy groups
- The squashing effect on large scales is equivalent to the quadrupole to monopole rato in the power spectrum and can be used to further constrain the cosmological model

Questions?

Complications…

III- Shot noise

Shot noise

- Galaxies are discrete objects
- For mathematical convenience, we describe their distribution with a continuous random field that it is sampled at random positions (note that there are 2 levels of randomness here)
- The effect of the sampling it is called shot noise and we need a model for it (there are infinite ways to do it). The most used is Poisson sampling (but never forget that it is just an approximation):

$$P(N\,|\,\delta) = \mathrm{Poisson}[(1+\delta_{gal})\bar{n}_{gal}V]$$

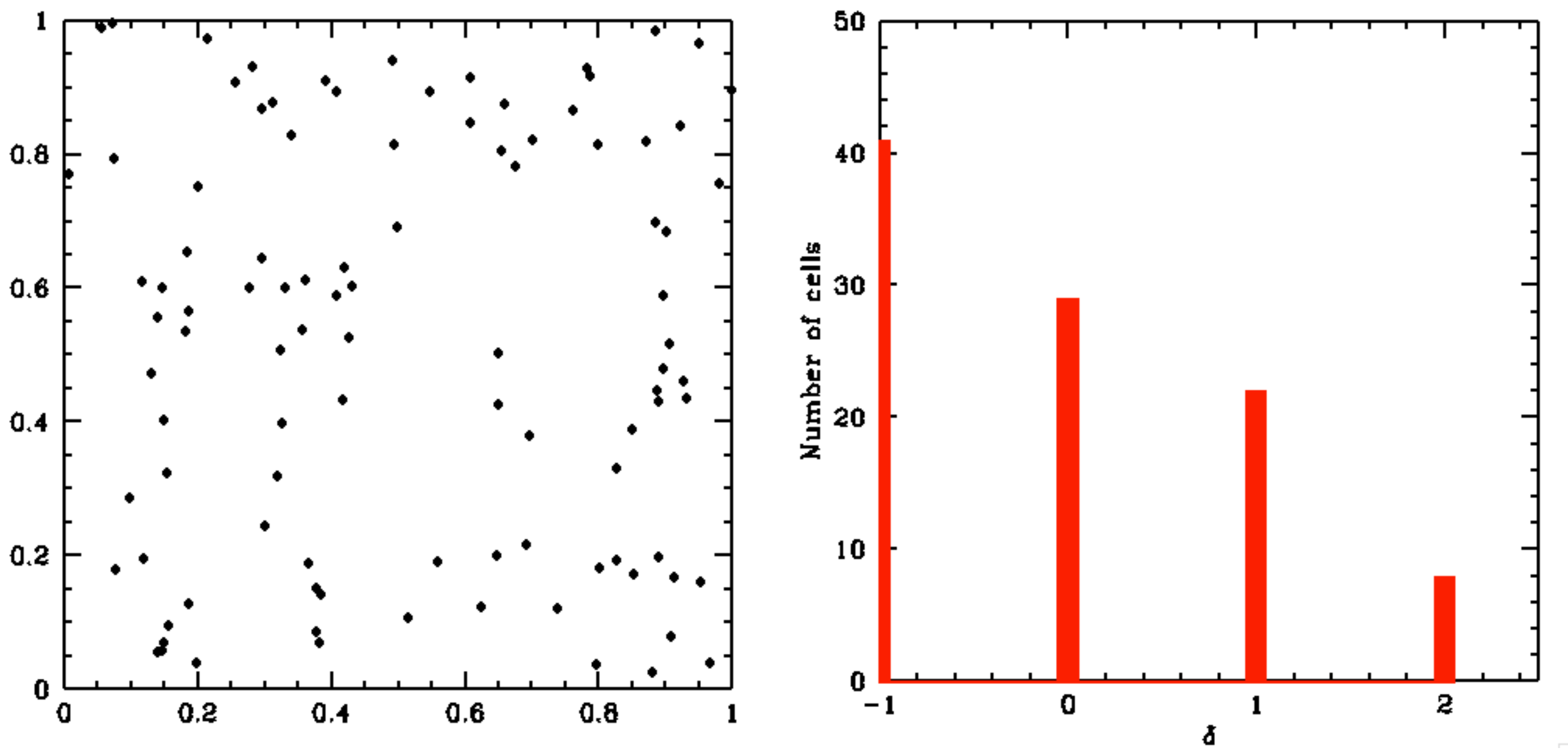
where P(N| δ) gives the probability of finding N galaxies in a volume V with underlying "continous" overdensity δ

Shot noise

- Shot noise also refers to the effect of self pairs (i.e. pairs made by a single objects) in N-point statistics

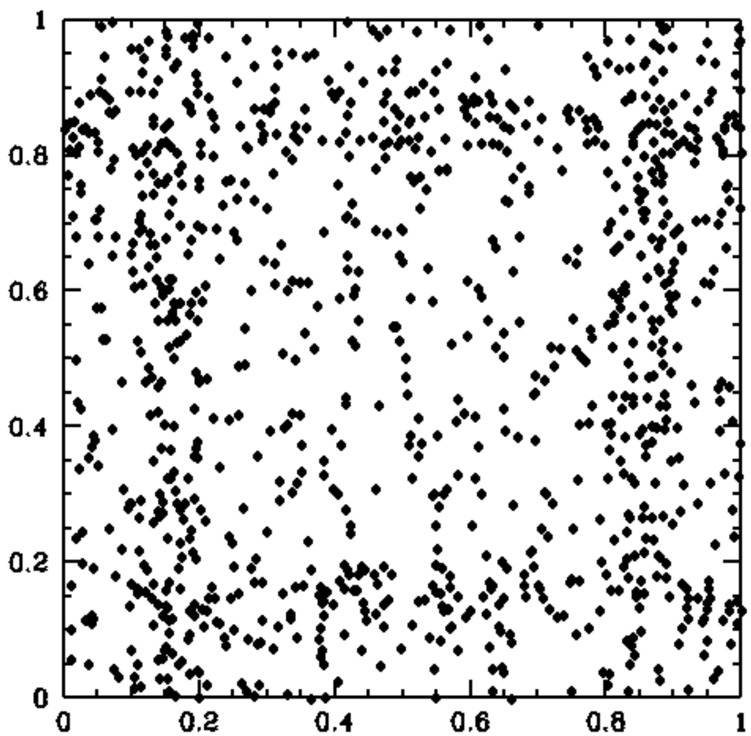
Shot noise: an example

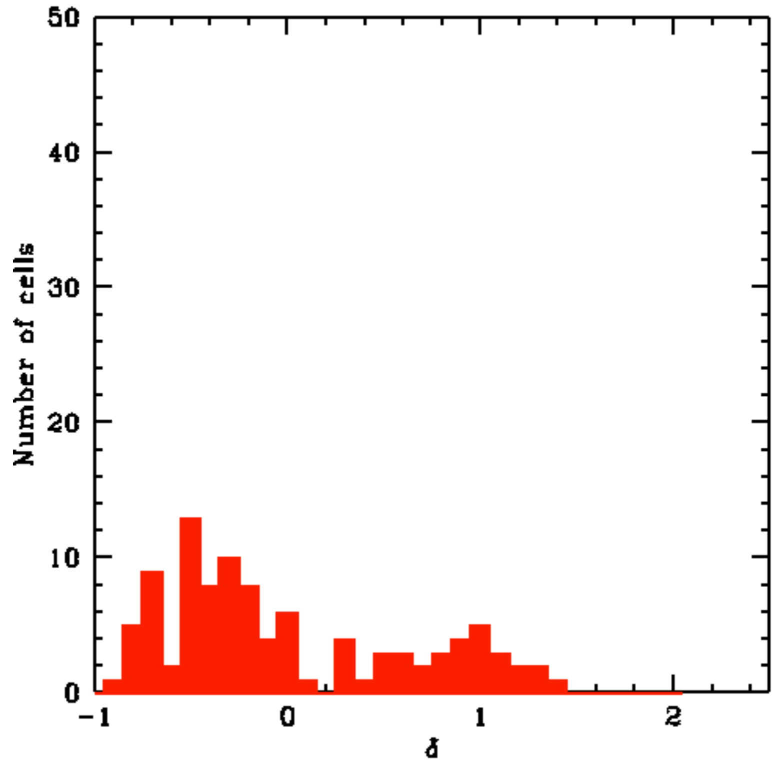
100 tracers



Shot noise: an example

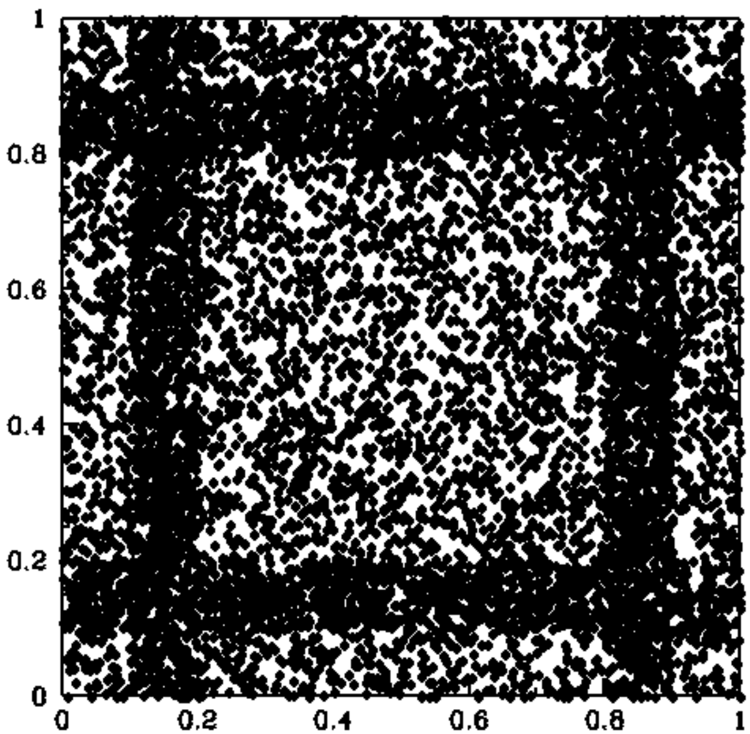
1000 tracers

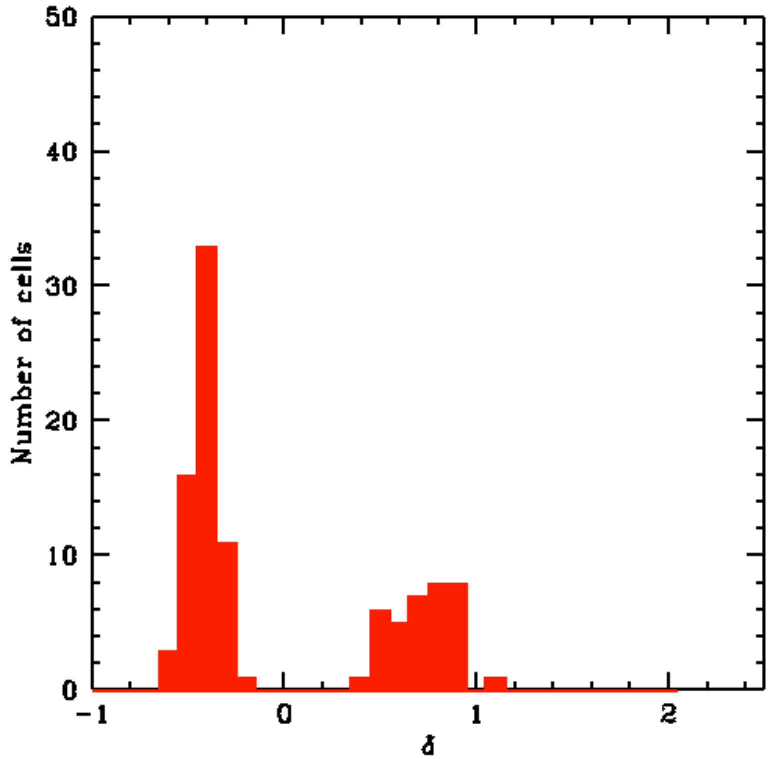




Shot noise: an example

10000 tracers





Shot noise and power spectra

- Poissonian shot noise affects power spectra in two ways

- First it adds a (white) systematic component

$$P_{obs}(k) = P(k) + \frac{1}{\bar{n}_{gal}}$$

- Second, it increases statistical uncertainties

$$\frac{\sigma_{P(k)}}{P(k)} = \left(\frac{2}{N}\right)^{1/2}\left(1 + \frac{1}{\bar{n}_{gal}P(k)}\right)$$

N = number of modes in a k-bin (scales as the volume of the survey)

Sample variance (for a Gaussian field)

Shot noise

Question

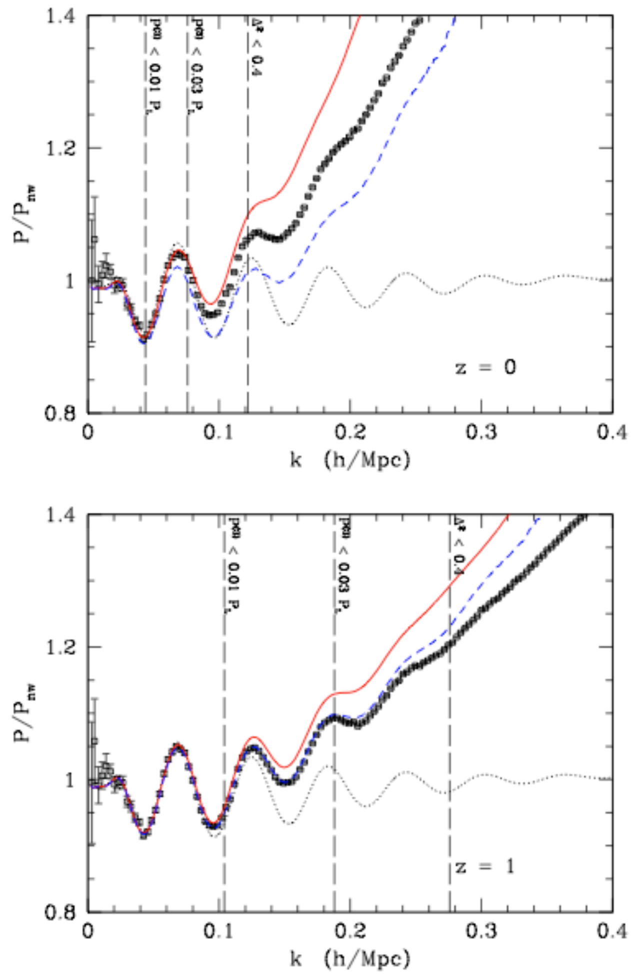
- What is the effect of shot noise on the 2-point correlation function?

Complications…

IV- Non-linear evolution

Non-linear evolution of the mass power spectrum

Carlson et al. 2009



- The non-linear growth of density perturbations changes the shape of their power spectrum from the linear one

- Current models are not very precise in recovering this behaviour for k>>0.1 h/Mpc

Outstanding question

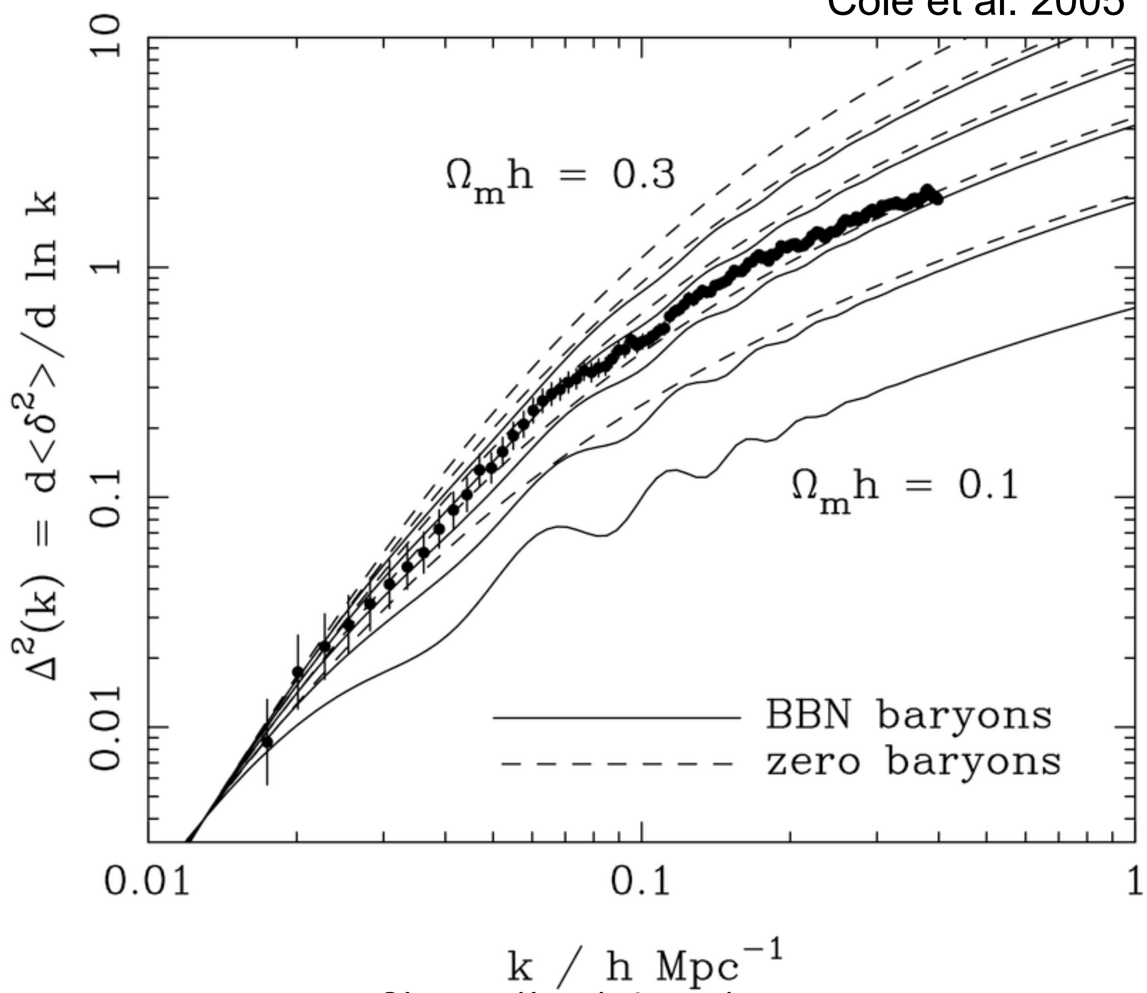
- Do uncertainties in modelling non-linearity, redshift distortions and galaxy bias compromise constraints on cosmological parameters coming from measurements of the galaxy power spectrum?

- Answer: they do not as long as we just use data on very large scales where linear models (for bias and for the evolution of perturbations) are accurate enough.

- This, however, makes errorbars of cosmological parameters big (with respect to the potential of the data) and a lot of efforts are currently made to improve the modelling of the non-linear effects

Cosmology from galaxy clustering

Cole et al. 2005



What did we learn?

- On separations larger than a few Mpc, models show that the ratio between the matter power spectrum and the galaxy power spectrum is nearly constant
- This implies that we can use the shape of the galaxy power spectrum to determine the cosmology
- Galaxy clustering gives $\Omega_m h \approx 0.2$, which for an Hubble constant h=0.7 gives $\Omega_m \approx 0.25\text{-}0.3$
- Combining this with the results of the CMB ($\Omega_{tot} \approx 1$), it suggests that 75% of the energy in the universe is in an unknown form, the so-called dark energy