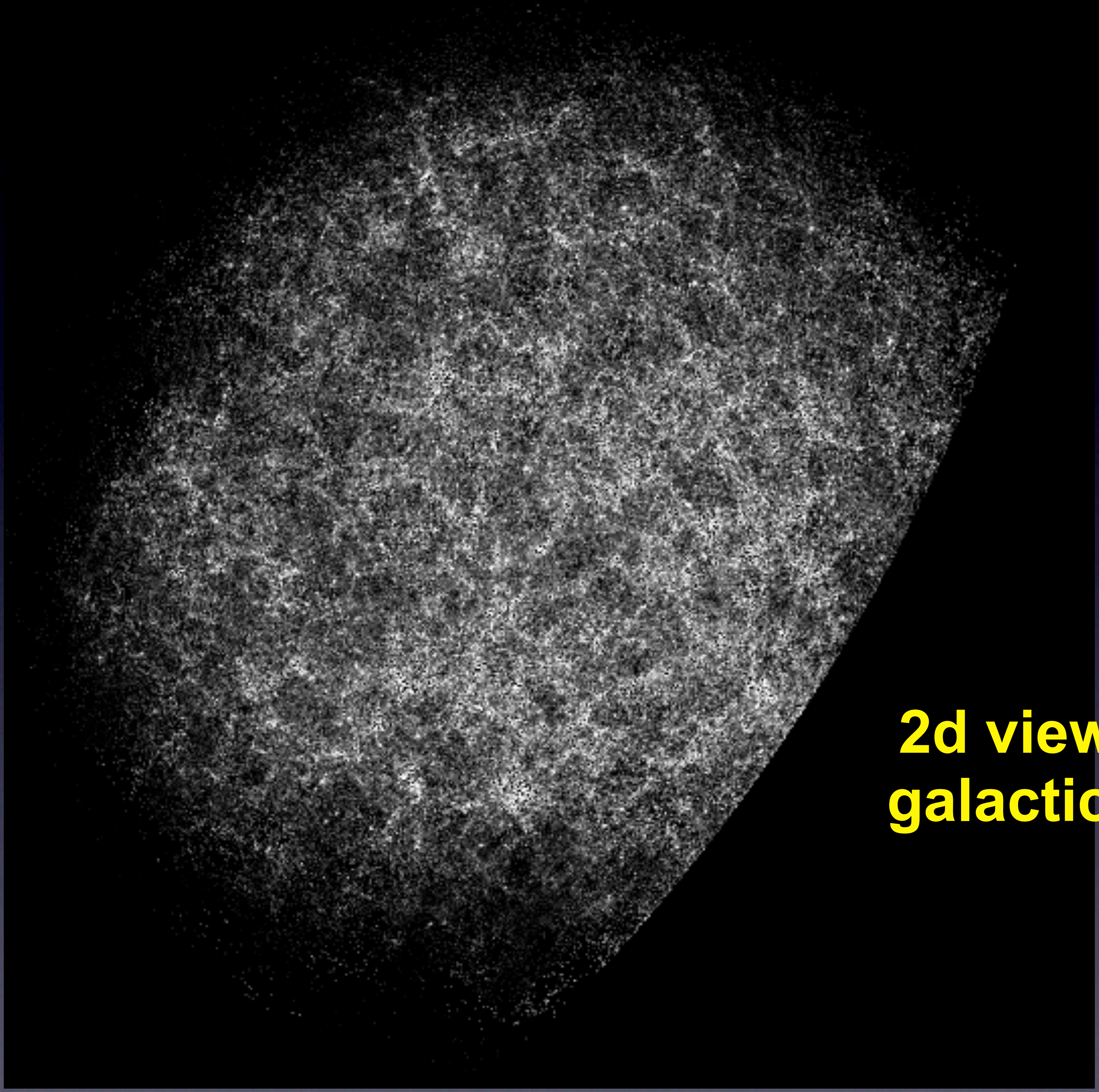
# TAUP 2013 summer school

# Observational Cosmology:
# Large-scale Structure of the Universe

Alison Coil
UCSD

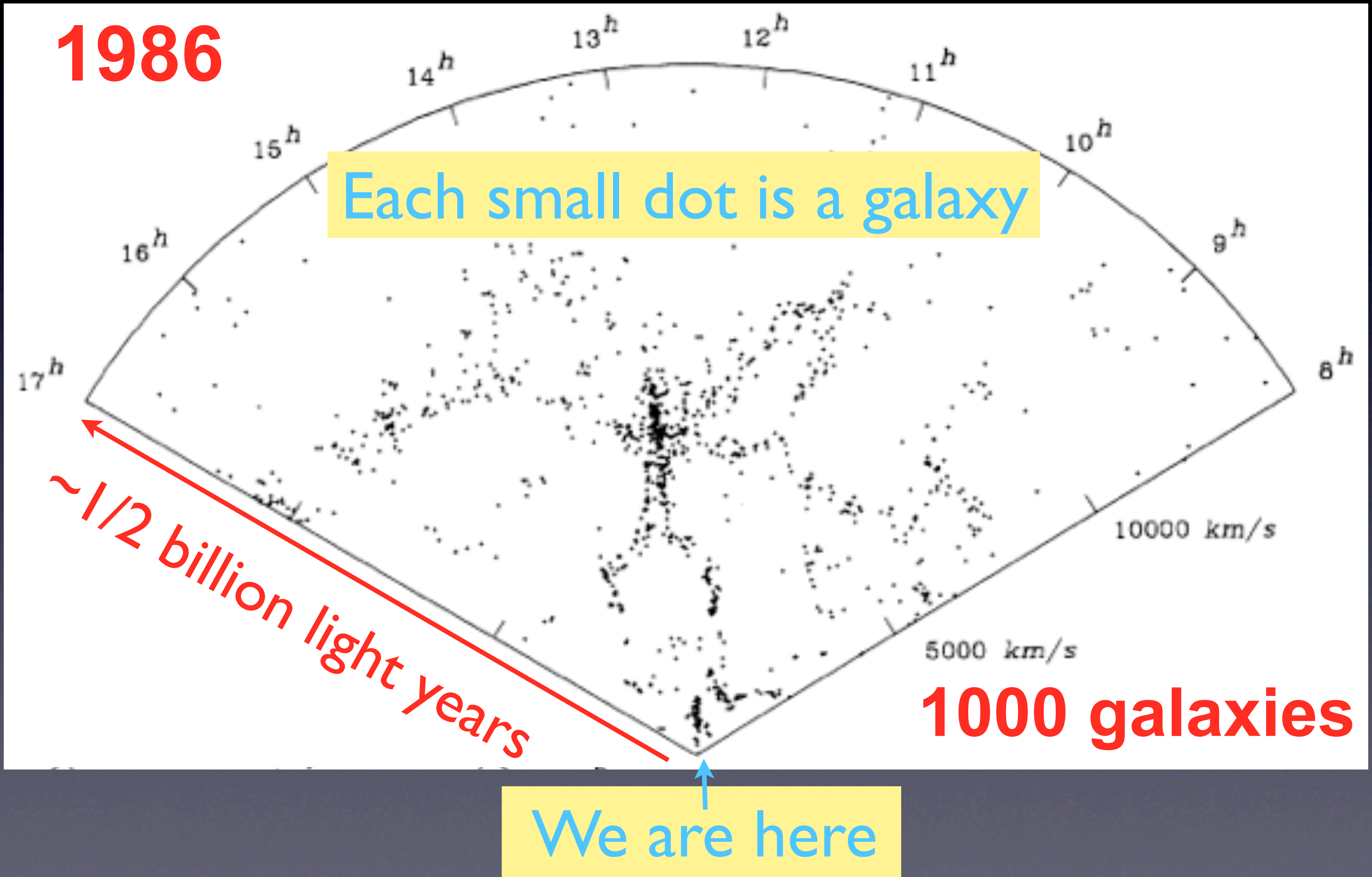
# Galaxy Distribution in Sky



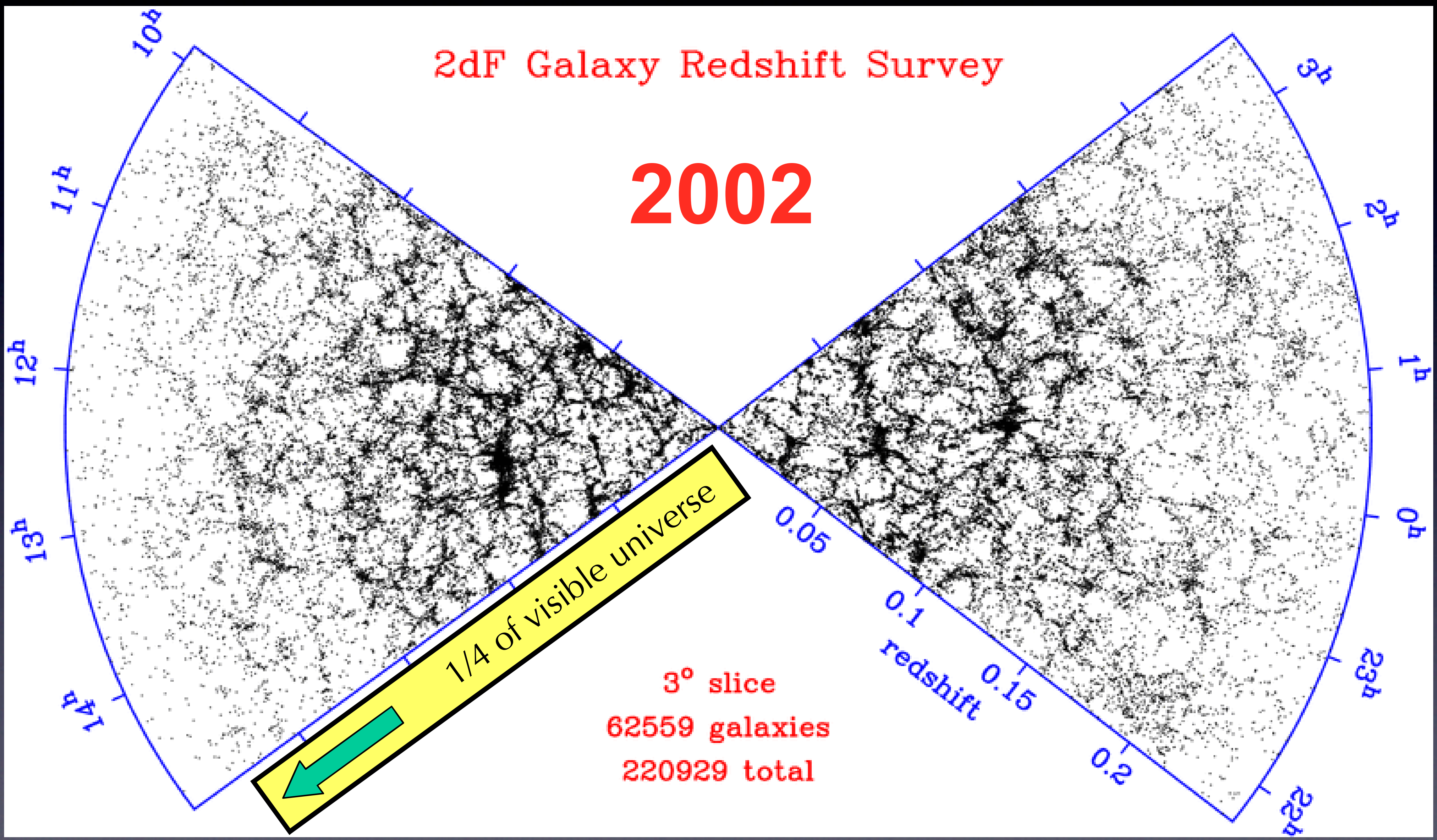
2d view of northern galactic hemisphere

Angular distribution of optical galaxies (M<19) in the north galactic pole from Lick survey - late 1970s

# Local Galaxy Distribution

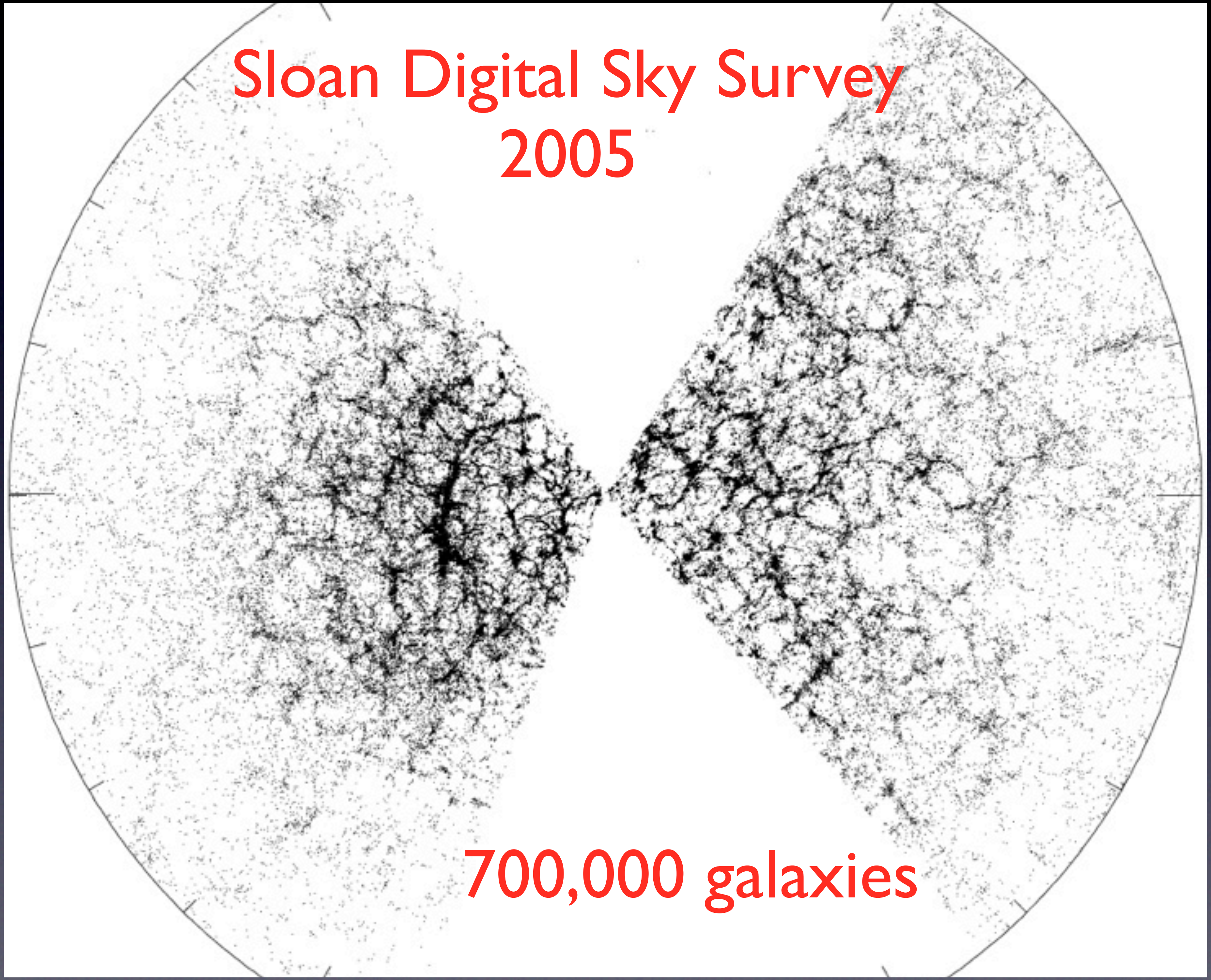


# Local Galaxy Distribution




~3 billion light years

# Local Galaxy Distribution




~3 billion light years

# Large-Scale Structure Terminology

**Groups**: have less than ~50 members, typical size ~2 Mpc. Velocity dispersions ~150 km/s, total mass ~$10^{13}$ $M_{\odot}$.

**Clusters**: have from ~50 members (a poor cluster) to ~1000 members (a rich cluster), typical size ~7 Mpc. Velocity dispersions ~1000-2000 km/s, total mass ~$10^{15}$ $M_{\odot}$.

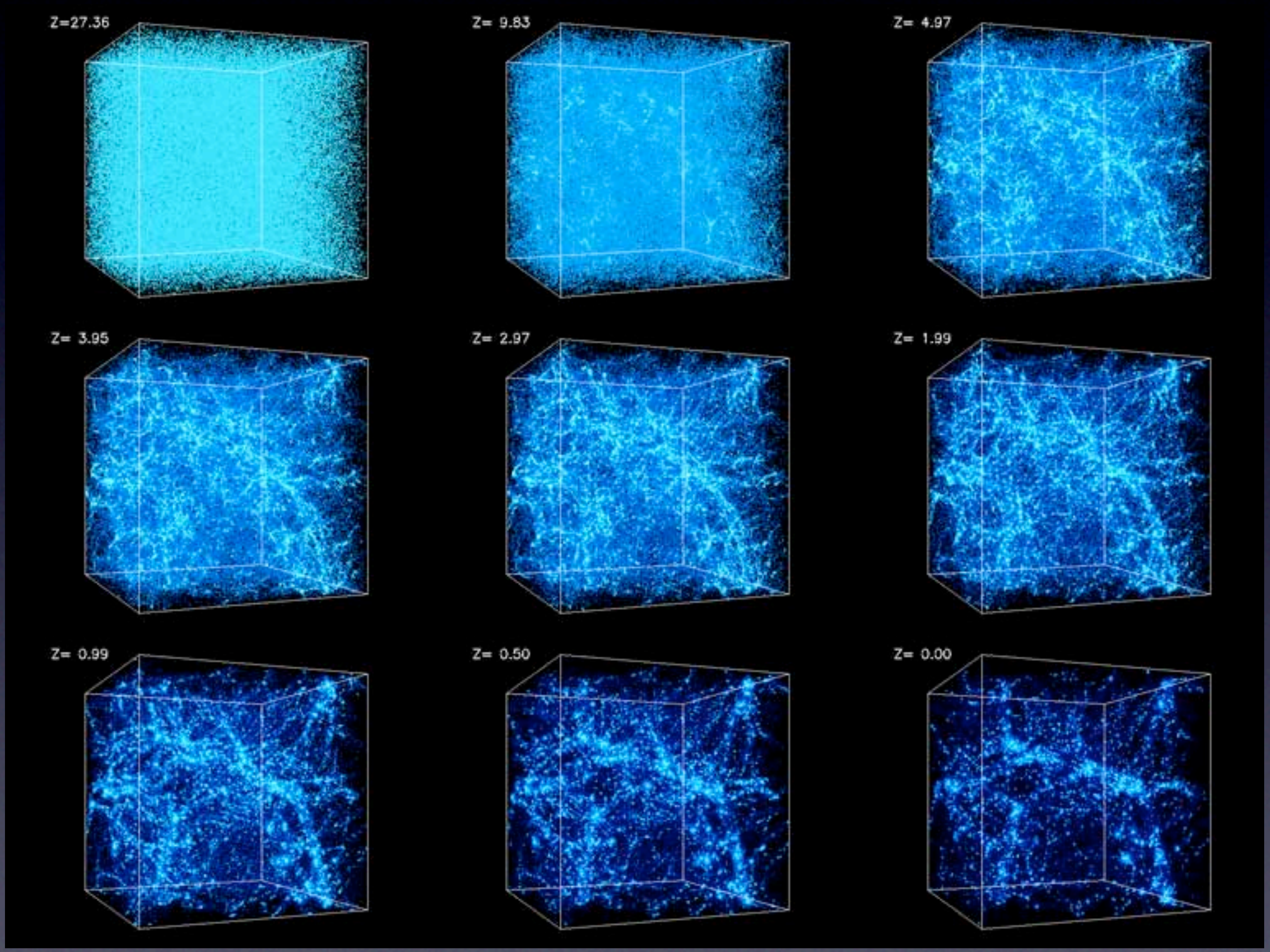
**Superclusters** are filaments of clusters.

# Structure Formation

Small initial perturbtions are amplified by gravity

~380,000 years after the Big Bang these fluctuations were 1 part in 10,000 - tiny!!

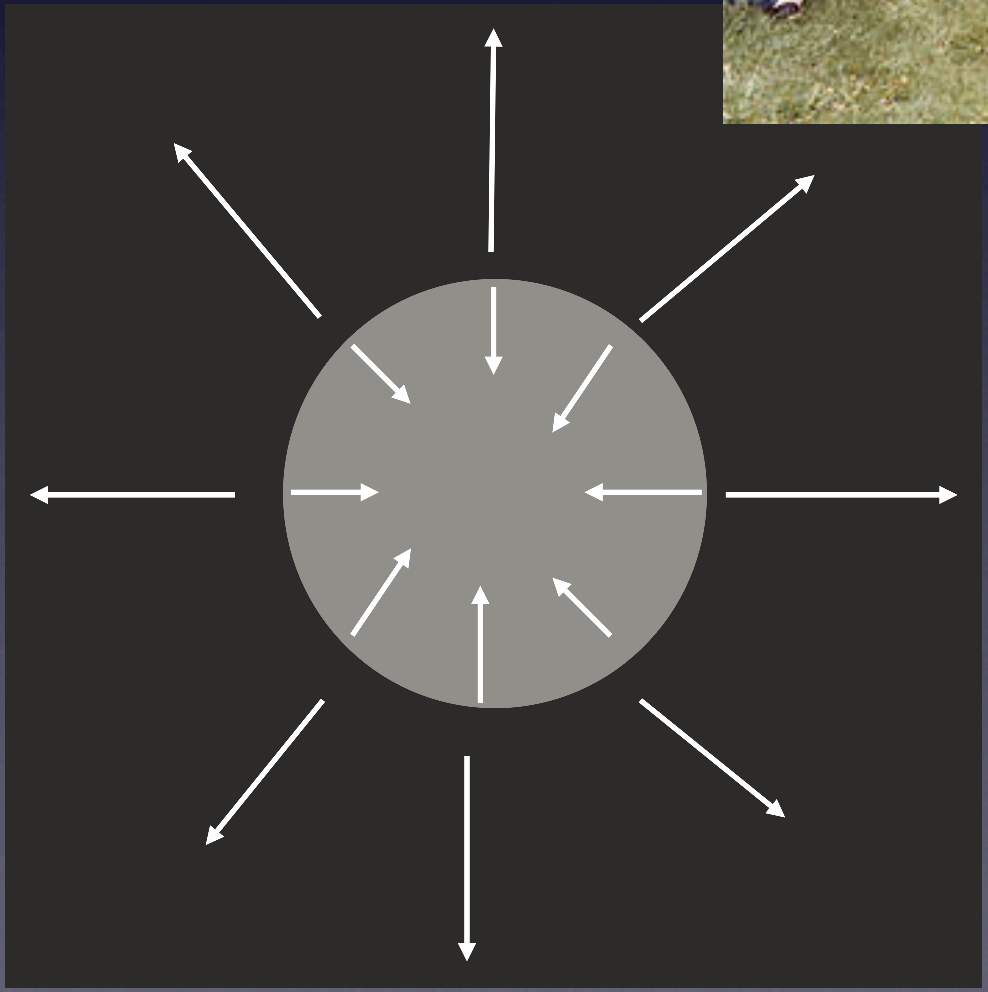
Become the very overdense galaxy groups and clusters of today



# Structure Formation

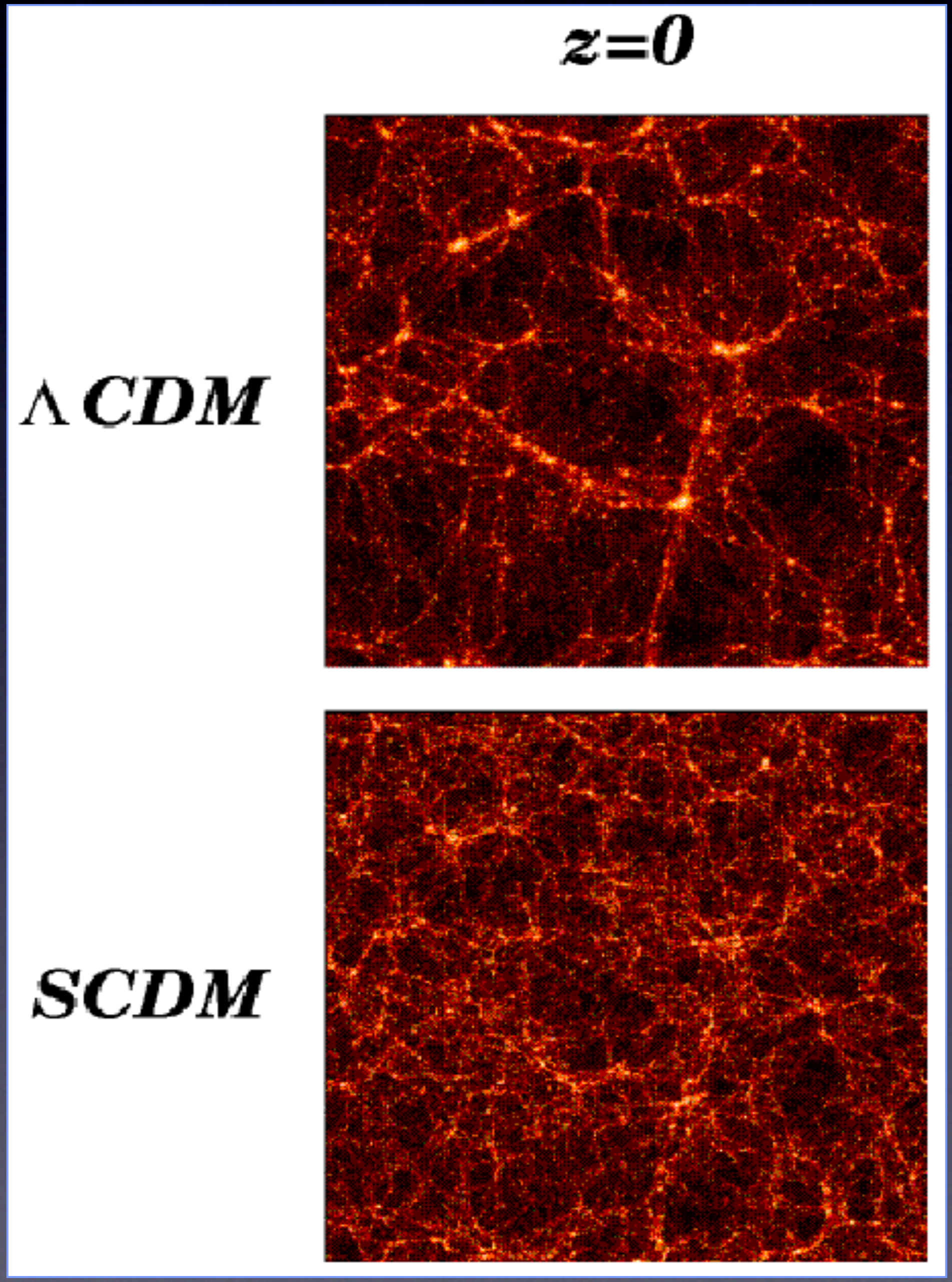
Cosmological expansion / dark energy

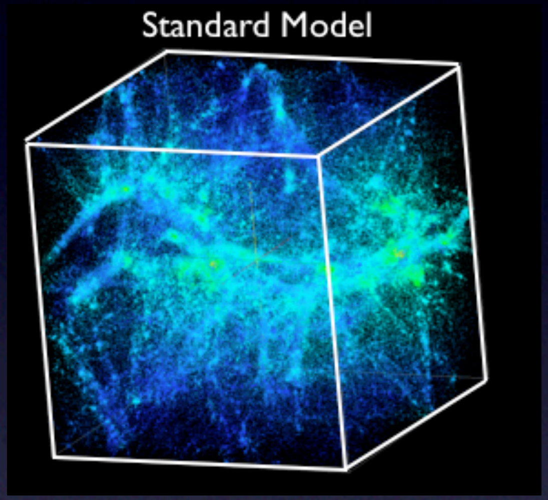
Gravity / dark matter

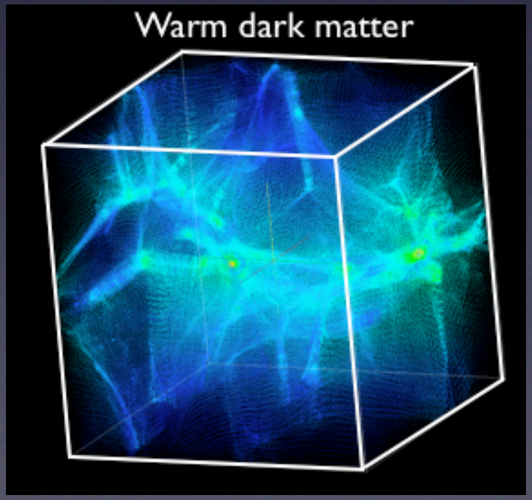


Overdensities grow into gravitationally bound structures - dark matter halos.

# Large-scale Structure Depends on Cosmology



# Clustering Primer

**Galaxy clustering reflects:**

- initial fluctuations in the mass density field + gravity
- cosmological parameters
- how galaxies populate dark matter halos

**Trace different physics on different scales:**

- small scales ($r < 100$ kpc/h):
  mergers + galaxy-galaxy interactions

- intermediate scales (100 kpc/h $< r <$ 2 Mpc/h):
  radial profiles of galaxies w/in halos / groups / clusters

- large scales ($r >$ 2 Mpc/h):
  large-scale density field / cosmology /
  host dark matter halo mass

# Clustering Primer

## Quantify clustering:

$$dP = n[1 + \xi(r)]dV$$

**two-point correlation function**

**r is separation between 2 galaxies, n is number density of galaxies**

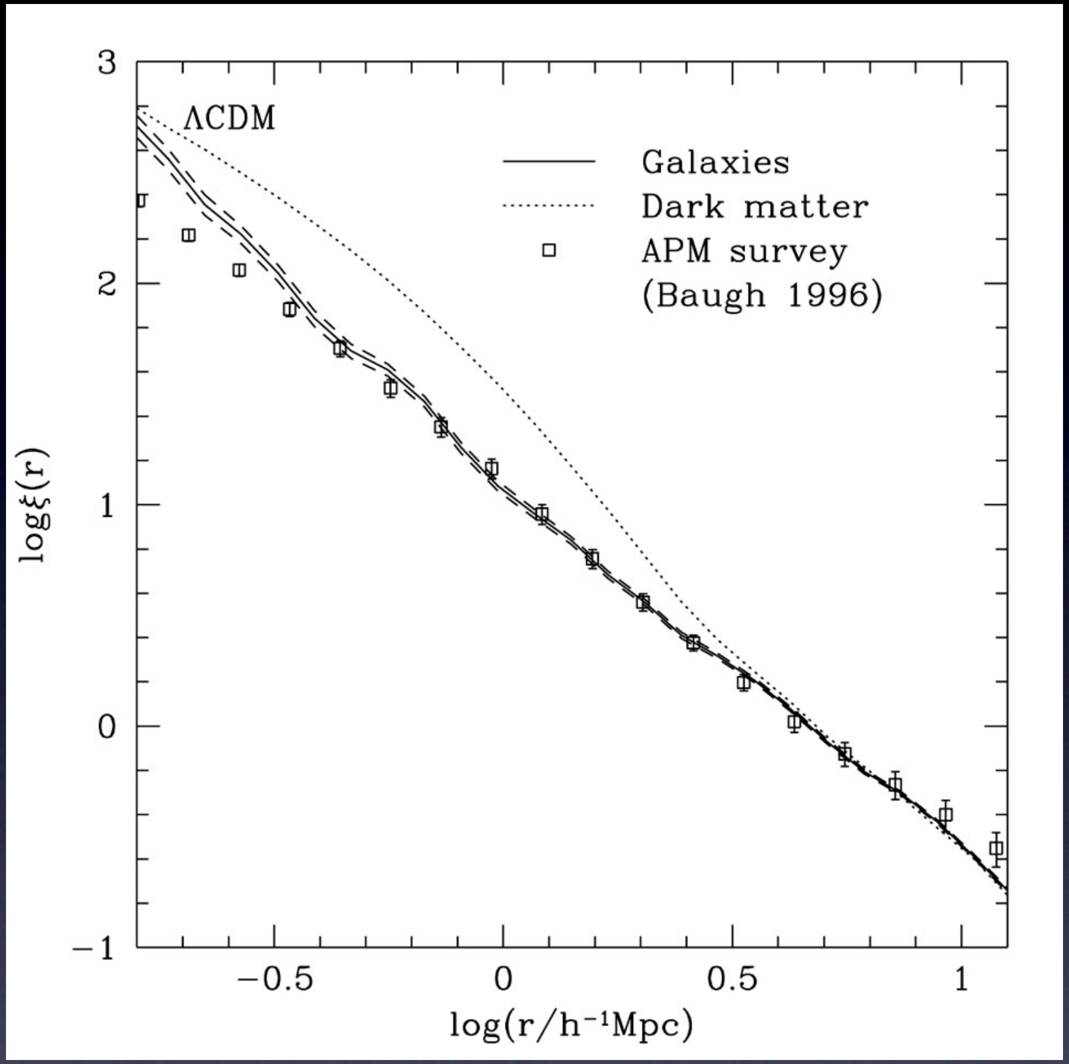
**relative to a random unclustered distribution**

**count pairs of galaxies as *f*(*r*) for data and random:**

$$\xi = \frac{n_R}{n_D}\frac{DD}{DR} - 1,$$

$$\xi = \frac{1}{RR}\left[DD\left(\frac{n_R}{n_D}\right)^2 - 2DR\left(\frac{n_R}{n_D}\right) + RR\right].$$

# Clustering Primer

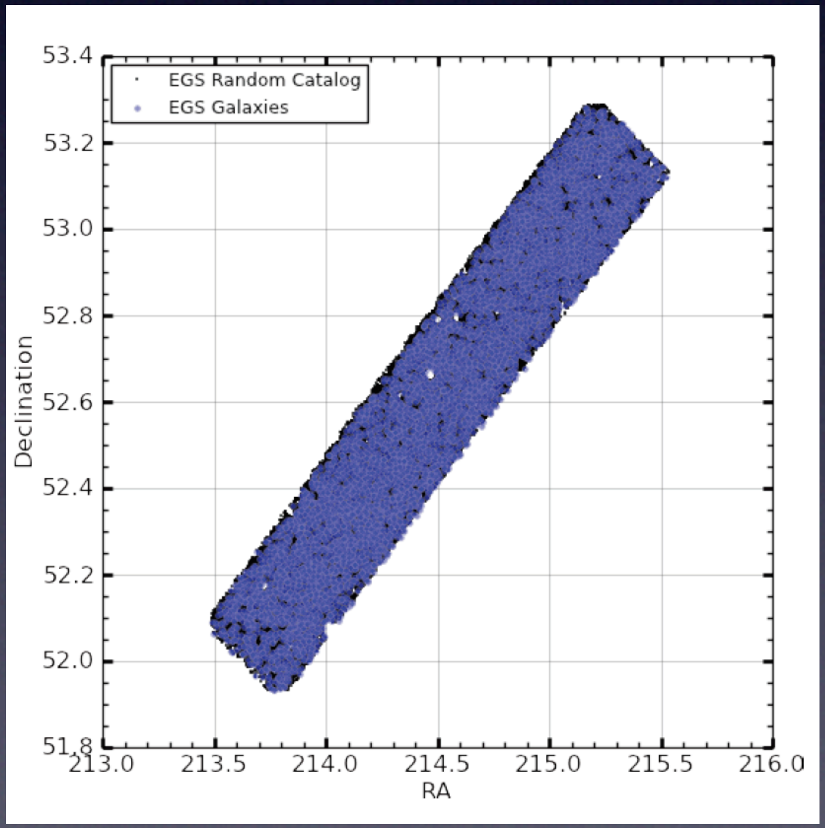


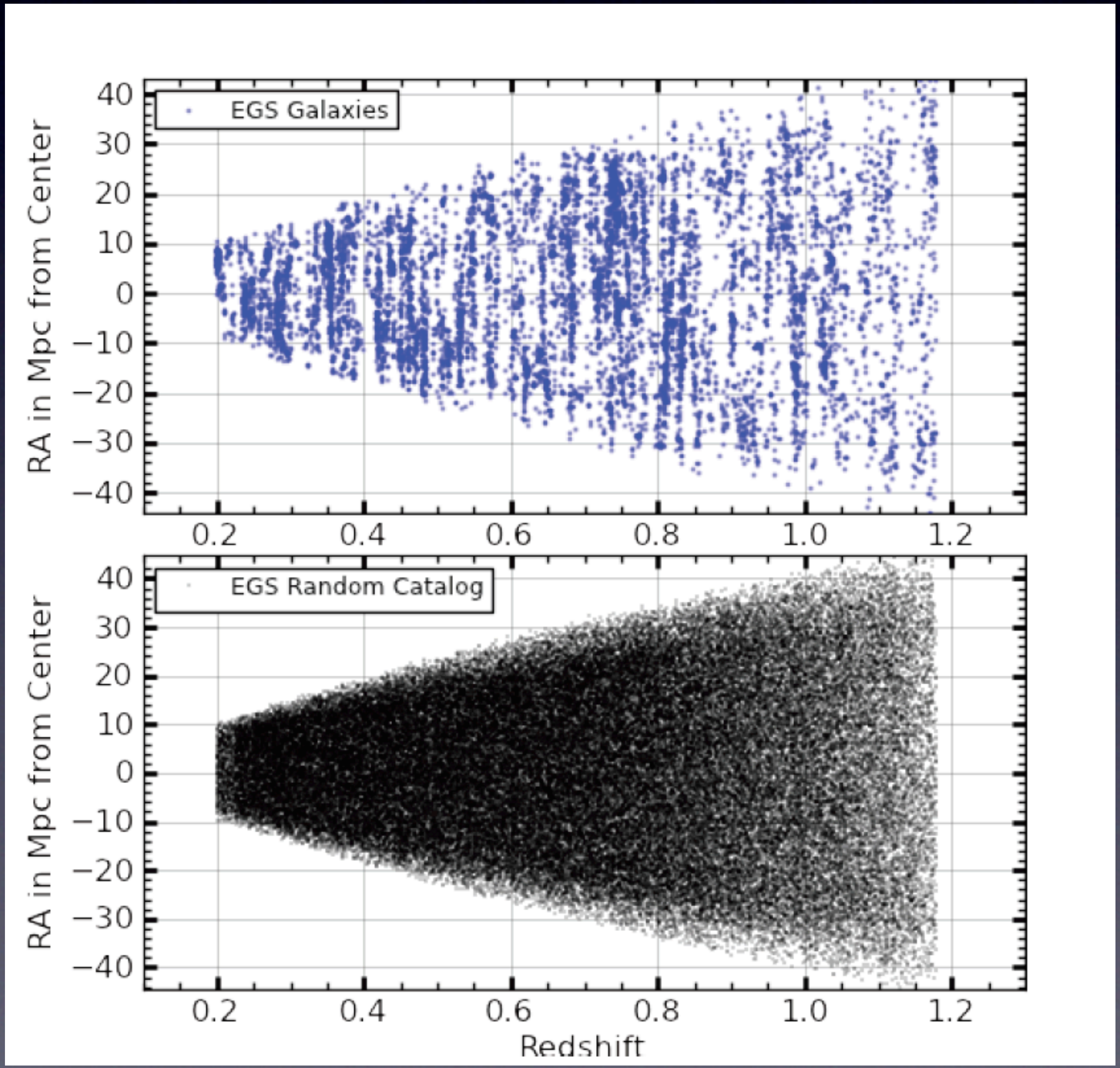

**large on small scales, small on large scales**

**roughly a power law:** $\boldsymbol{\xi(r) = (r_0/r)^\gamma}$ $\boldsymbol{r_0}$ - scale length

# Random Catalog

**Random catalog has same sky and redshift coverage as data. It must include masking of bright stars, CCD defects, vignetting in spectrograph, etc.! Must know your selection function very well, both spatially and in z.**



# Angular Clustering

**2d projected angular correlation function:**

$$dP = N[1 + \omega(\theta)]d\Omega$$

**can infer the 3d correlation function *if* you know the redshift distribution of sources**

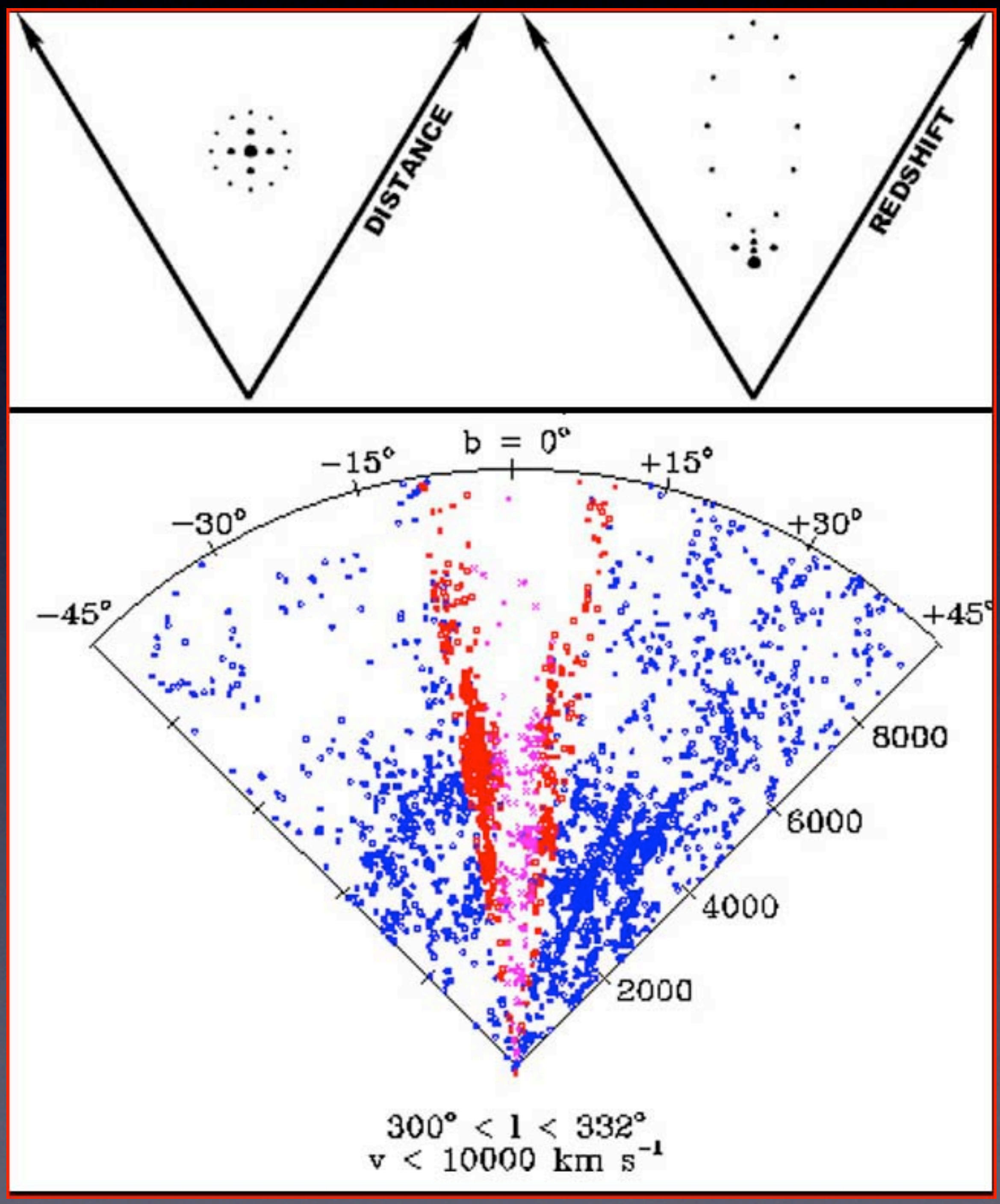
$$\omega(\theta) = A_\omega \theta^\delta$$

$$A = \frac{\int_0^\infty r_0^\gamma(z) g(z) \left(\frac{dN}{dz}\right)^2 dz}{\left[\int_0^\infty \left(\frac{dN}{dz}\right) dz\right]^2}$$

**degeneracy b/w inherent clustering amplitude and dN/dz width**
**dominant error is lack of knowledge of dN/dz**

# Redshift Space Distortions

real space    redshift space



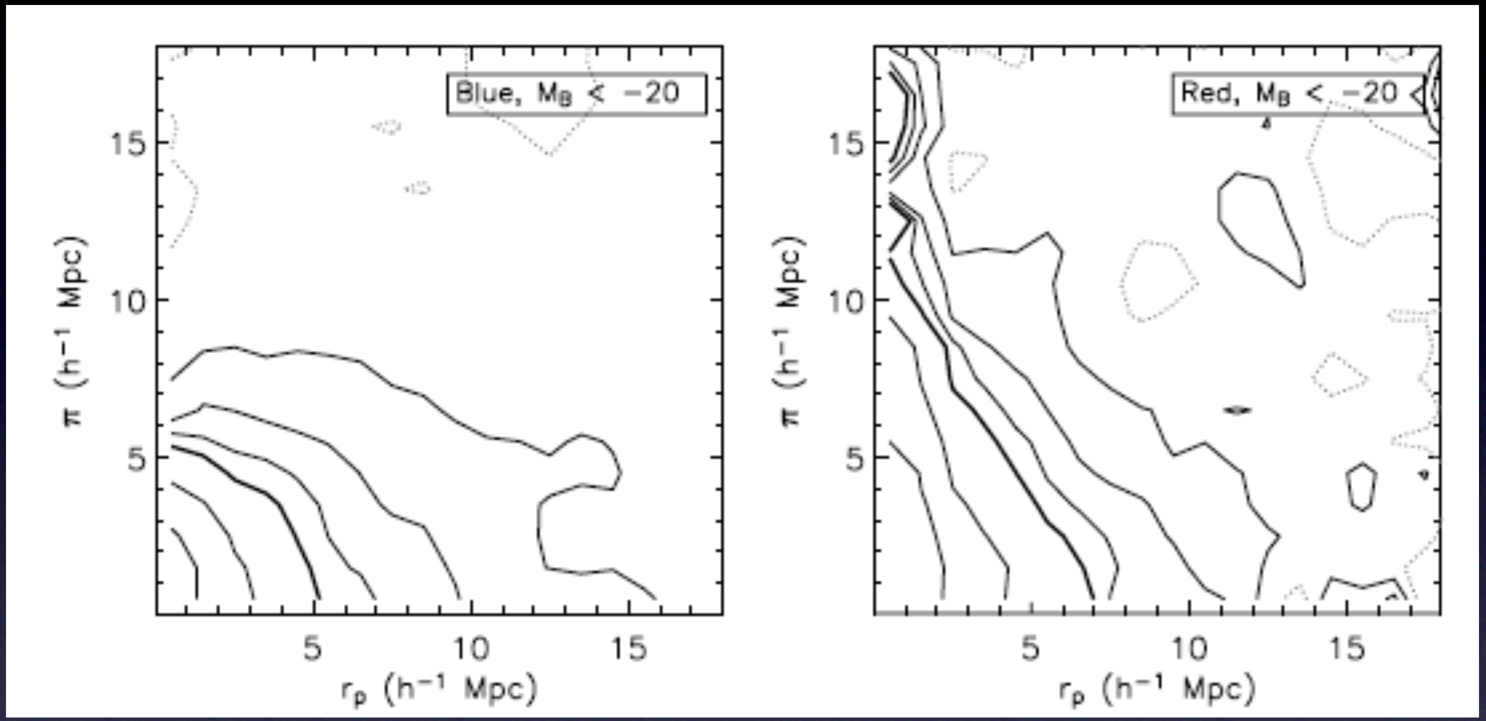

"Fingers of God"

- generally see on scales < few Mpc/h - virialized motions of galaxies w/in halos

- on larger scales see coherent infall of galaxies into forming structures (Kaiser effect)

# Redshift Space Distortions

$\xi(r_p, \pi)$



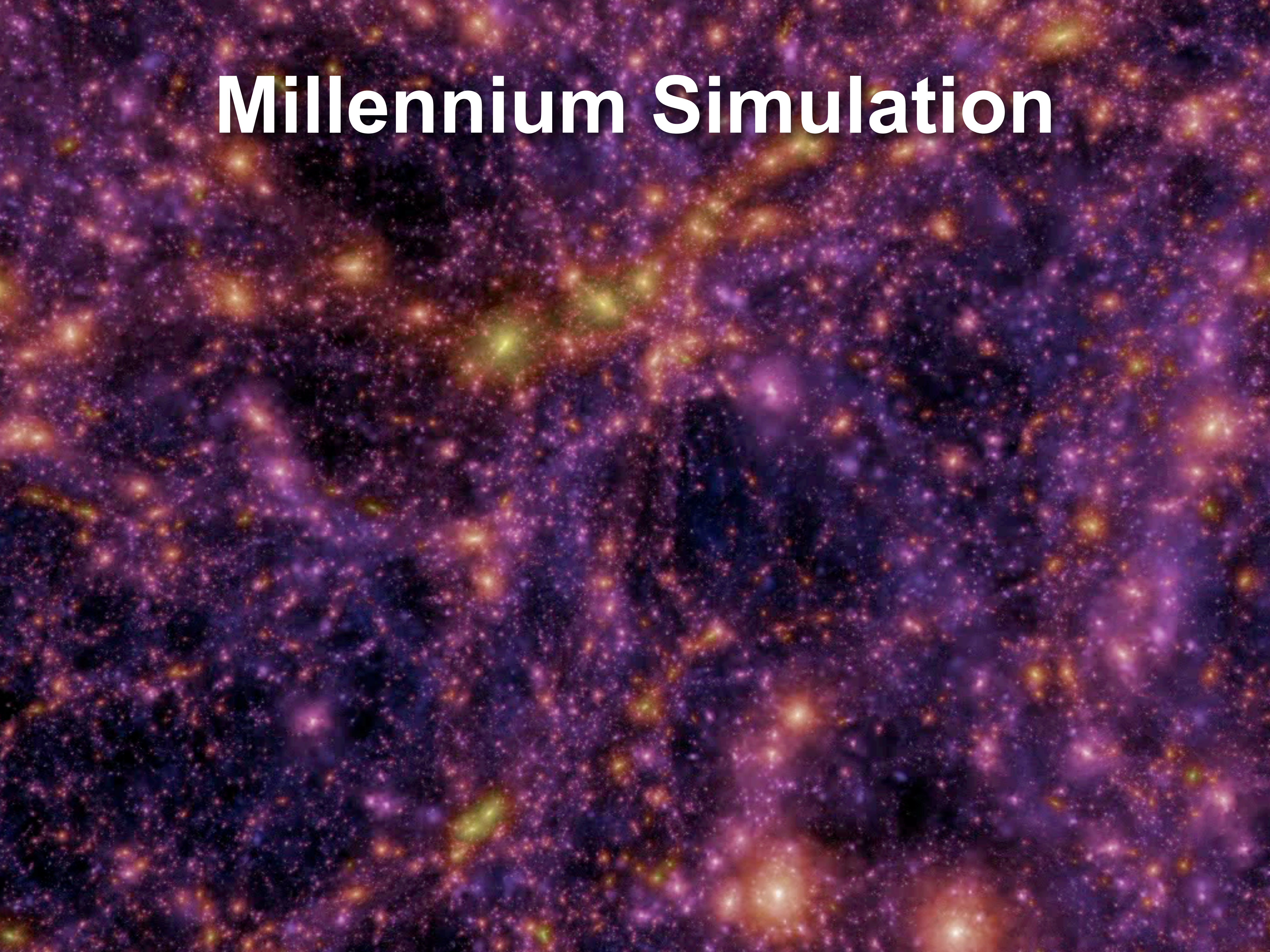

- integrate along line of sight to get projected correlation function:

$$w_p(r_p) = 2 \int_0^\infty d\pi \; \xi(r_p, \pi)$$

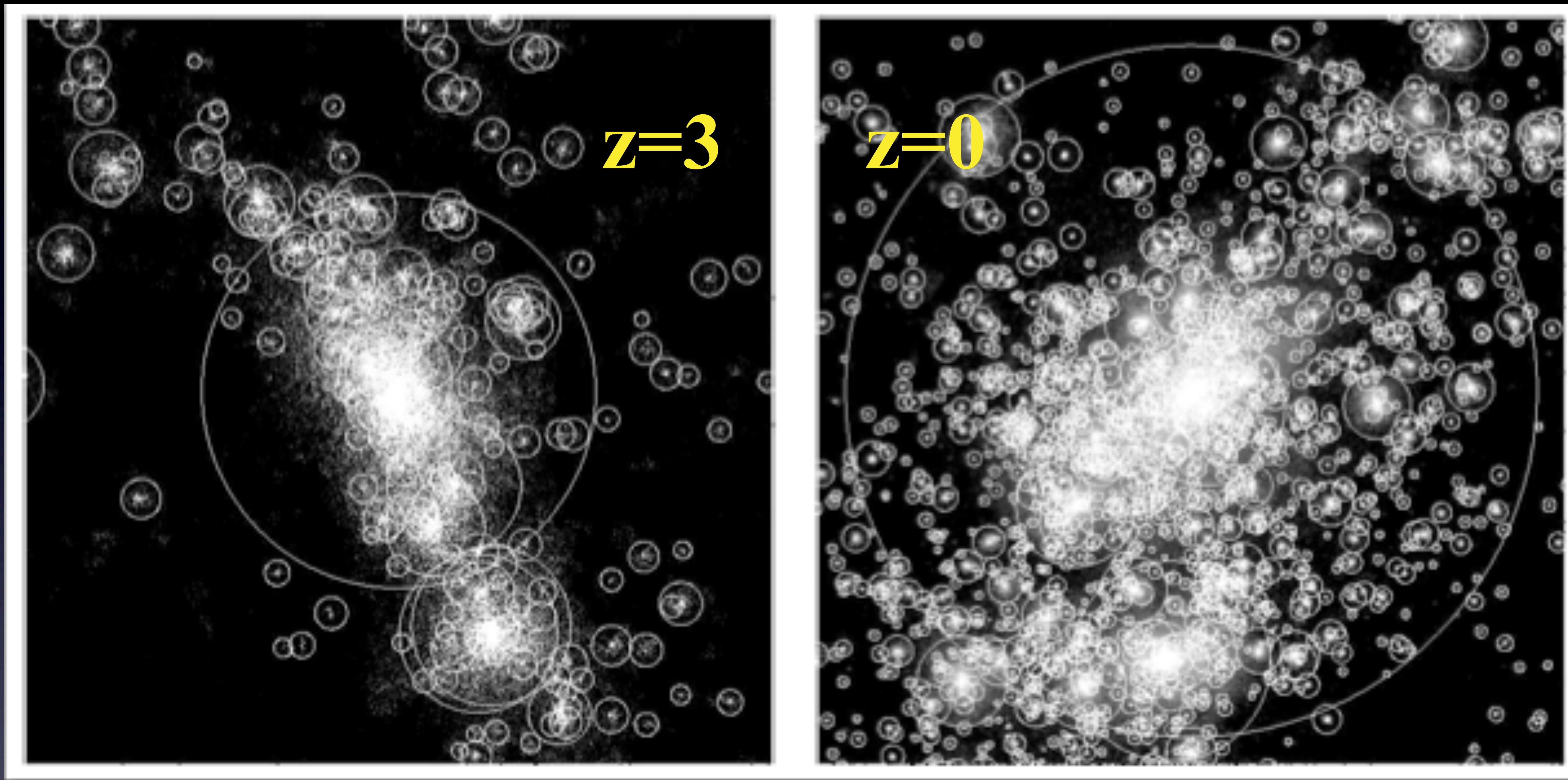
- can fit a power law to that:

$$w_p(r_p) = r_p \left(\frac{r_0}{r_p}\right)^\gamma \frac{\Gamma(\frac{1}{2})\Gamma(\frac{\gamma-1}{2})}{\Gamma(\frac{\gamma}{2})},$$

# Millennium Simulation



# Dark Matter Distribution




Kravtsov et al. 2004

Dark matter halos are collapsed overdensities in the matter distribution.
Galaxy clustering = halo clustering (cosmology) + galaxies in halos (galaxy formation)

# Galaxy Bias

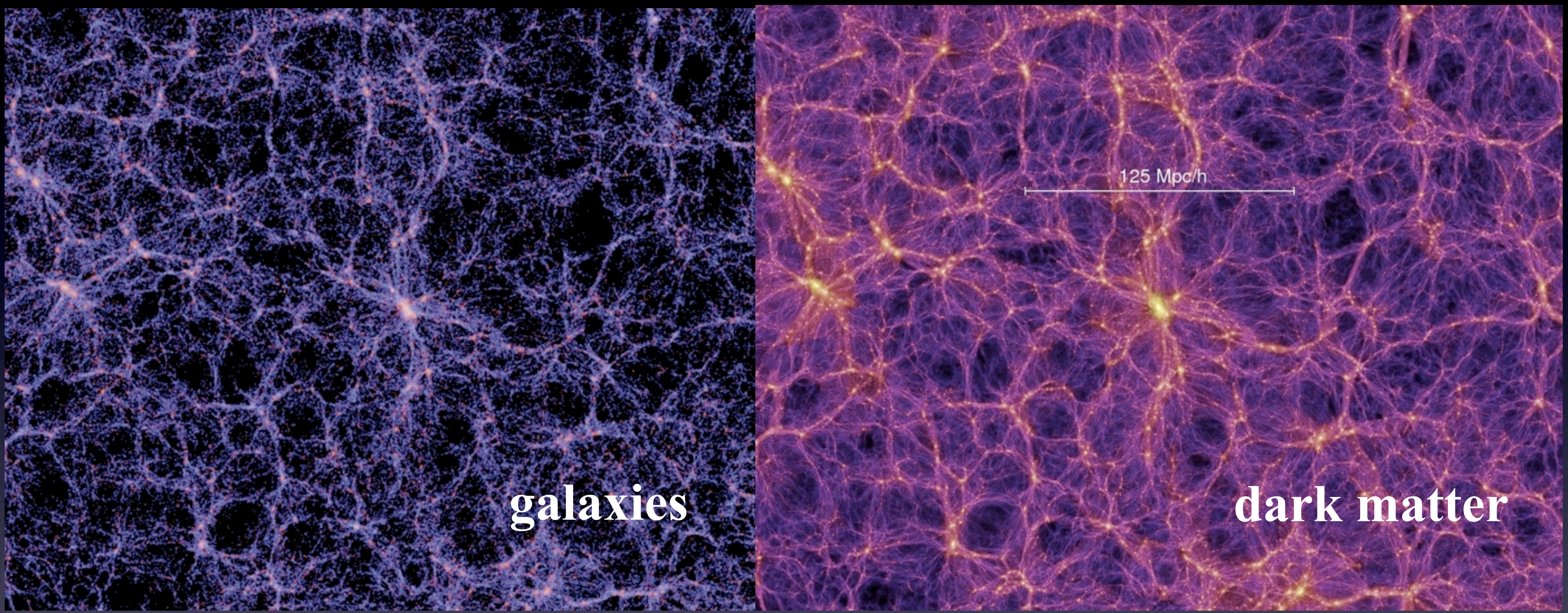
**The galaxy density field is a discrete and possibly stochastic function of the underlying dark matter density field.**

**linear galaxy bias: mean overdensity of galaxies / mean overdensity of mass**

$$b = \delta_g/\delta,$$

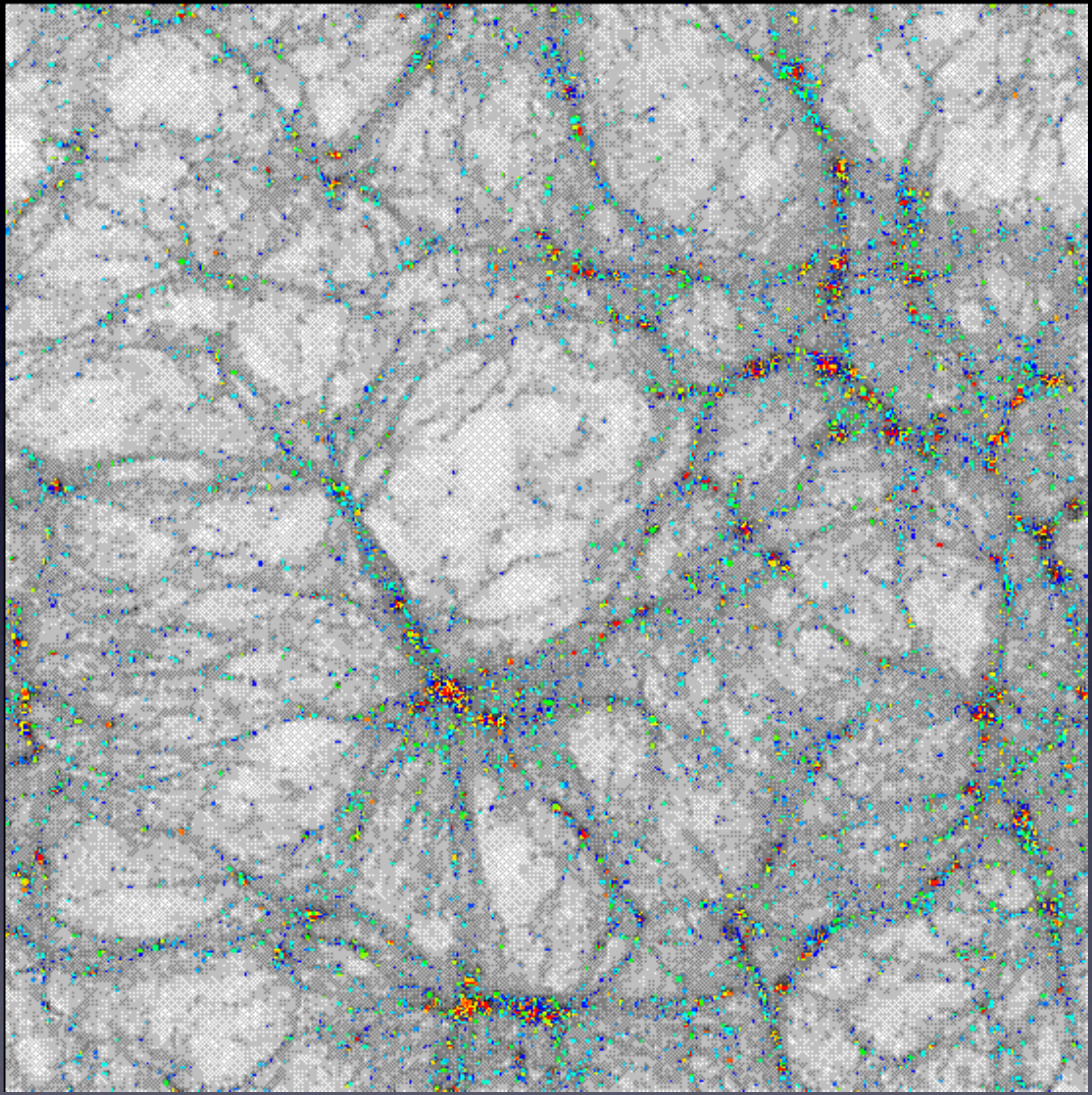
$$b = (\xi_{\rm gal}/\xi_{\rm dark\ matter})^{1/2}$$

# Bias: galaxy/dark matter clustering




z=0  Millenium Run simulation
Applied a semi-analytic model that predicts that galaxies are not very biased tracers of the mass at z=0, on large scales.

# Bias: galaxy clustering/dark matter clustering



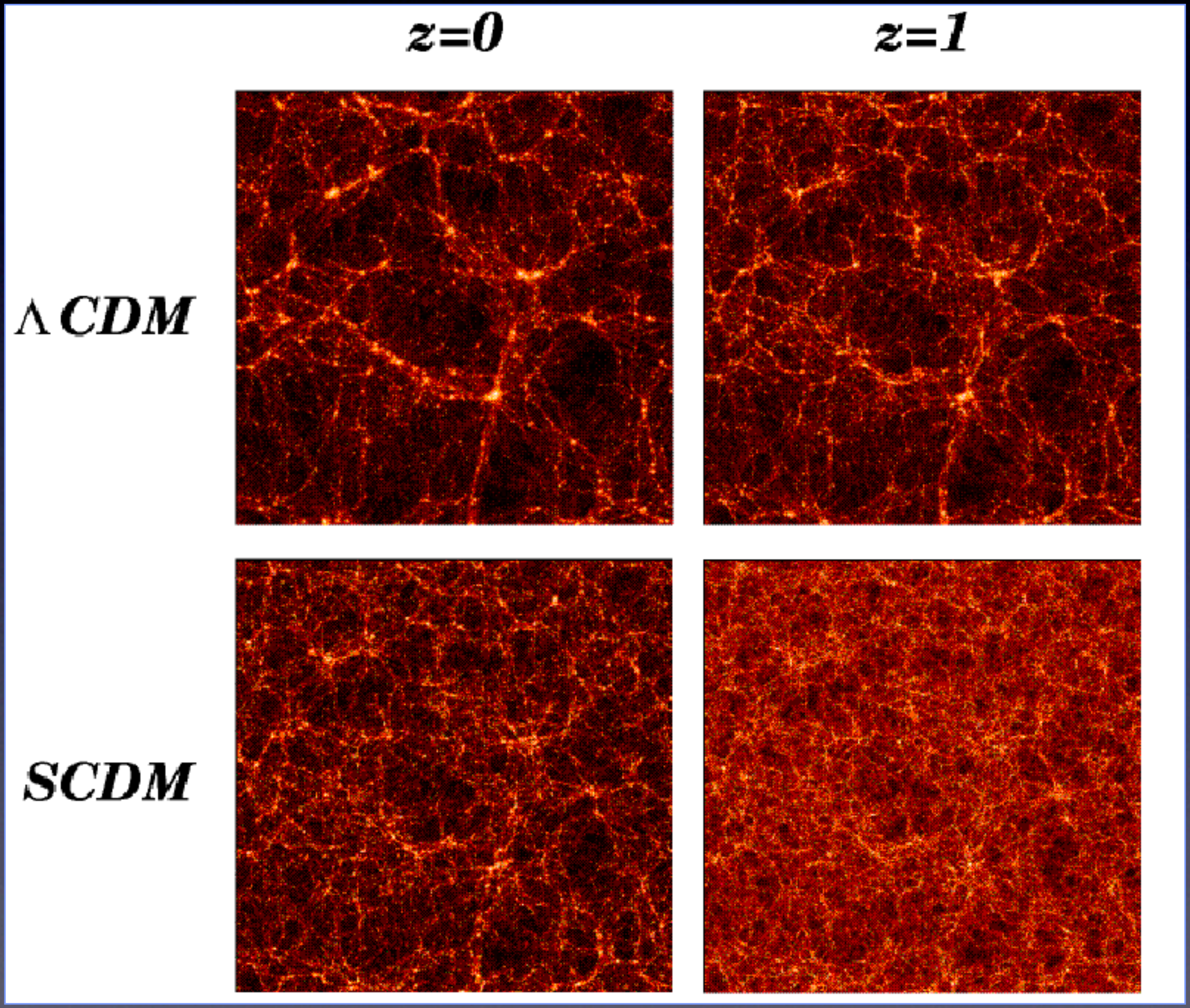
Bias is expected to evolve with redshift, as first galaxies at high-z formed in densest regions.

***Observations:***
***z=3: b~4   z=0: b~1***

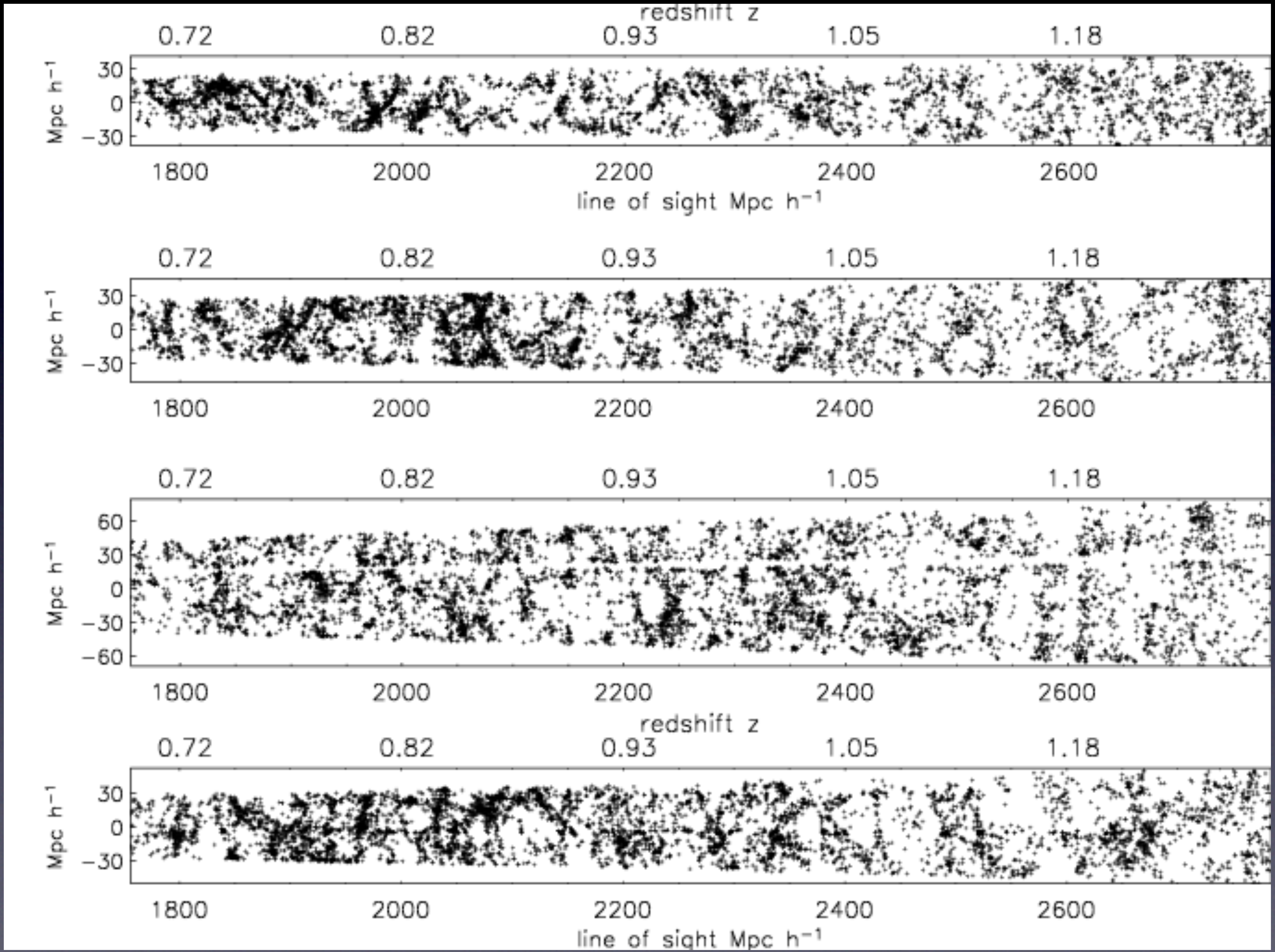
**Evolution of bias and dependence on scale and galaxy properties places strong constraints on galaxy formation theories.**

Galaxy formation sim. at z=3 by Kauffmann et al. grey=dark matter particles colors=galaxies

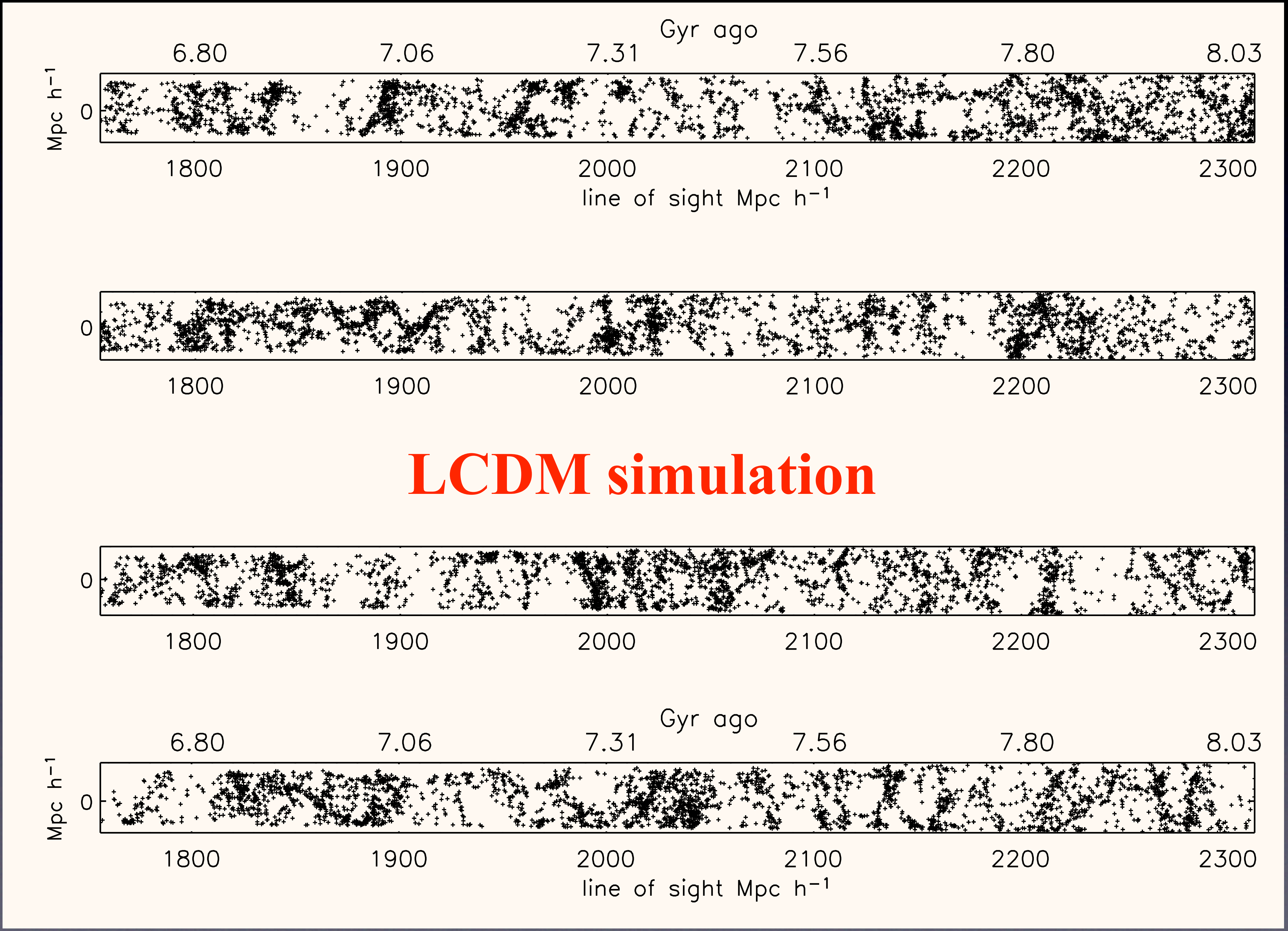
# High redshift galaxy surveys can break degeneracy w/ bias at z=0



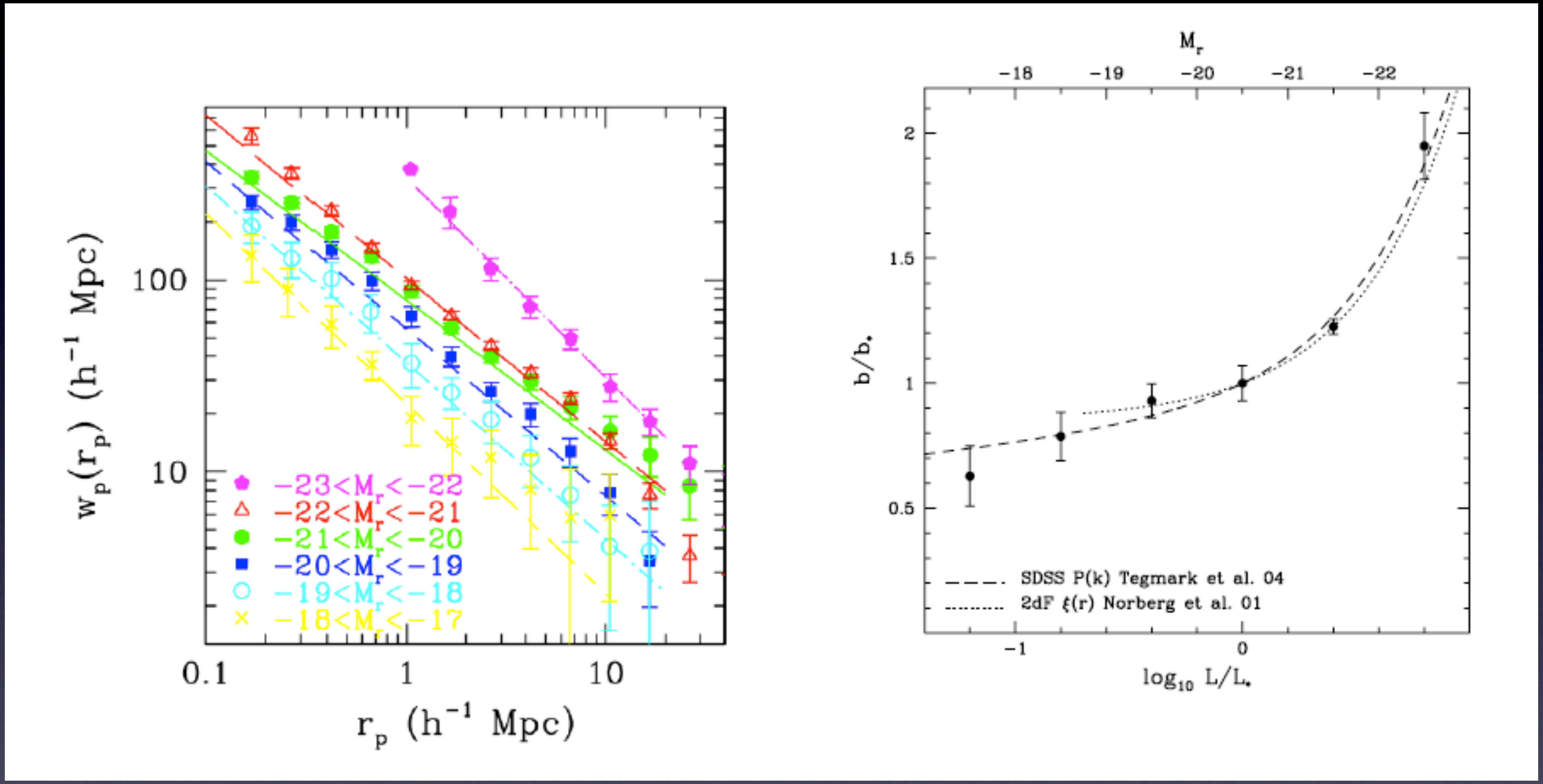
# Redshift Maps in 4 Fields at z=1



# Redshift Maps in 4 Fields at z=1



# Dependence of Clustering on Galaxy Properties




Zehavi et al. 2011

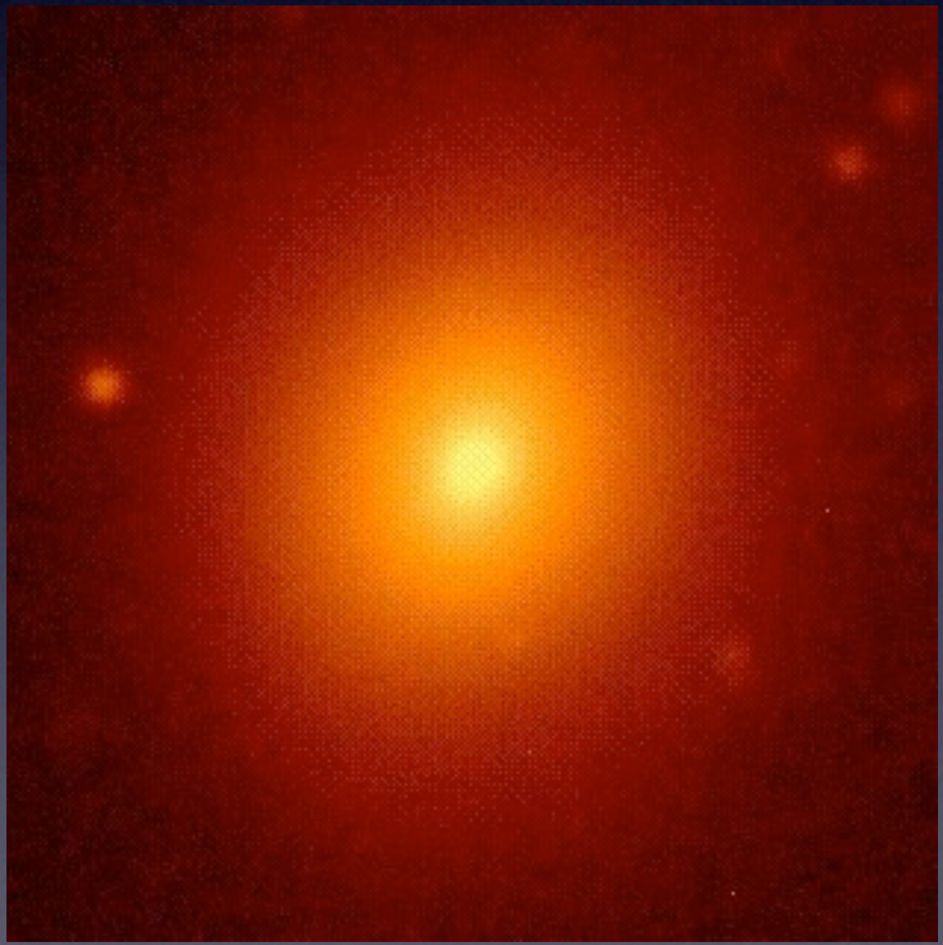
**luminosity dependence:**

**more luminous galaxies reside in more massive dark matter halos**

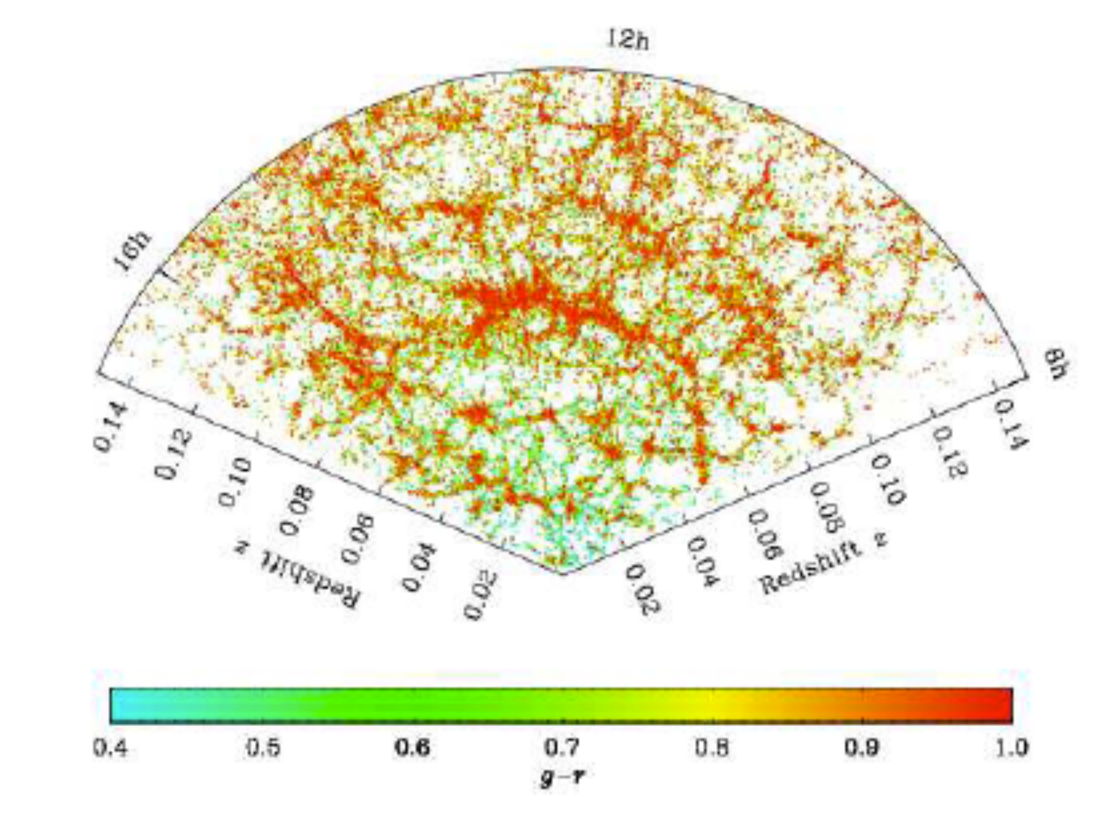
# Star-forming vs quiescent galaxies

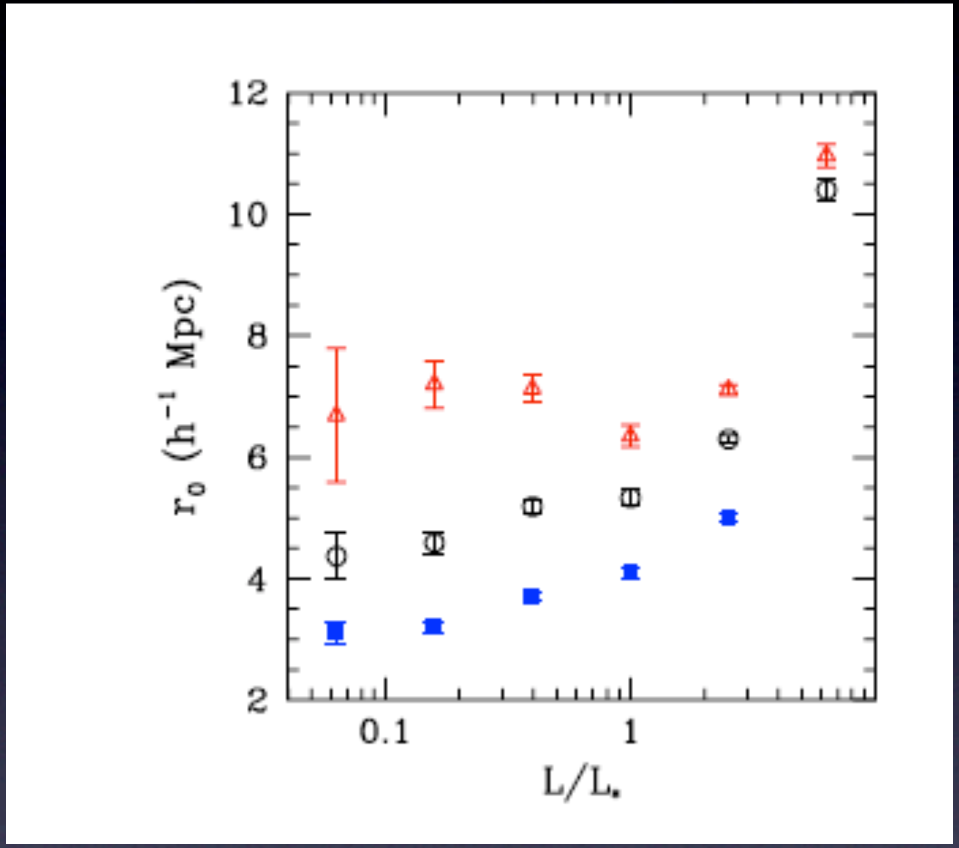
spiral, forming stars,
lots of dust and gas,
blue color

elliptical, not forming stars,
little dust and gas,
red color



# Dependence of Clustering on Galaxy Properties




Zehavi et al. 2011

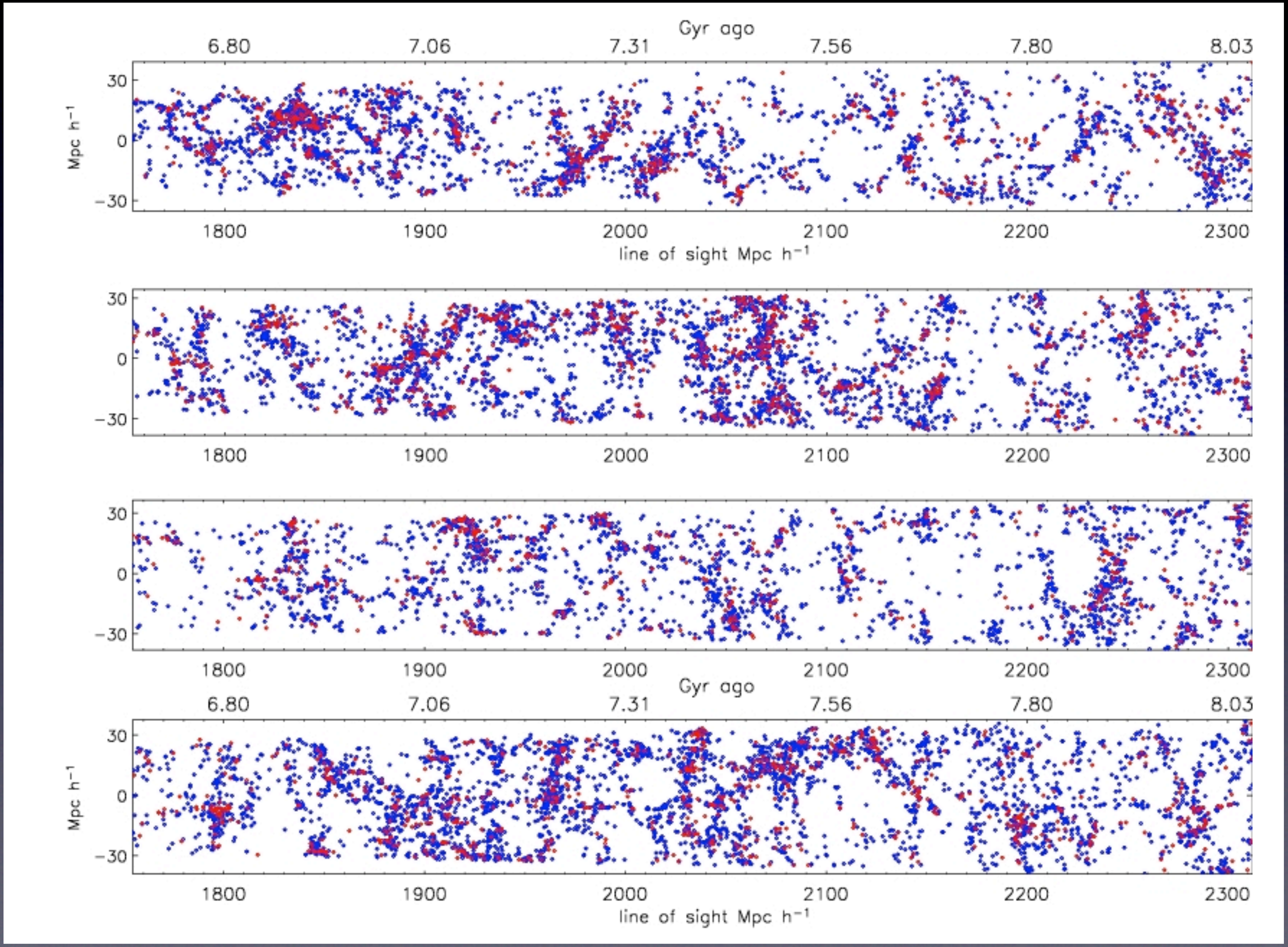
**color dependence:**

**stronger than luminosity dependence**

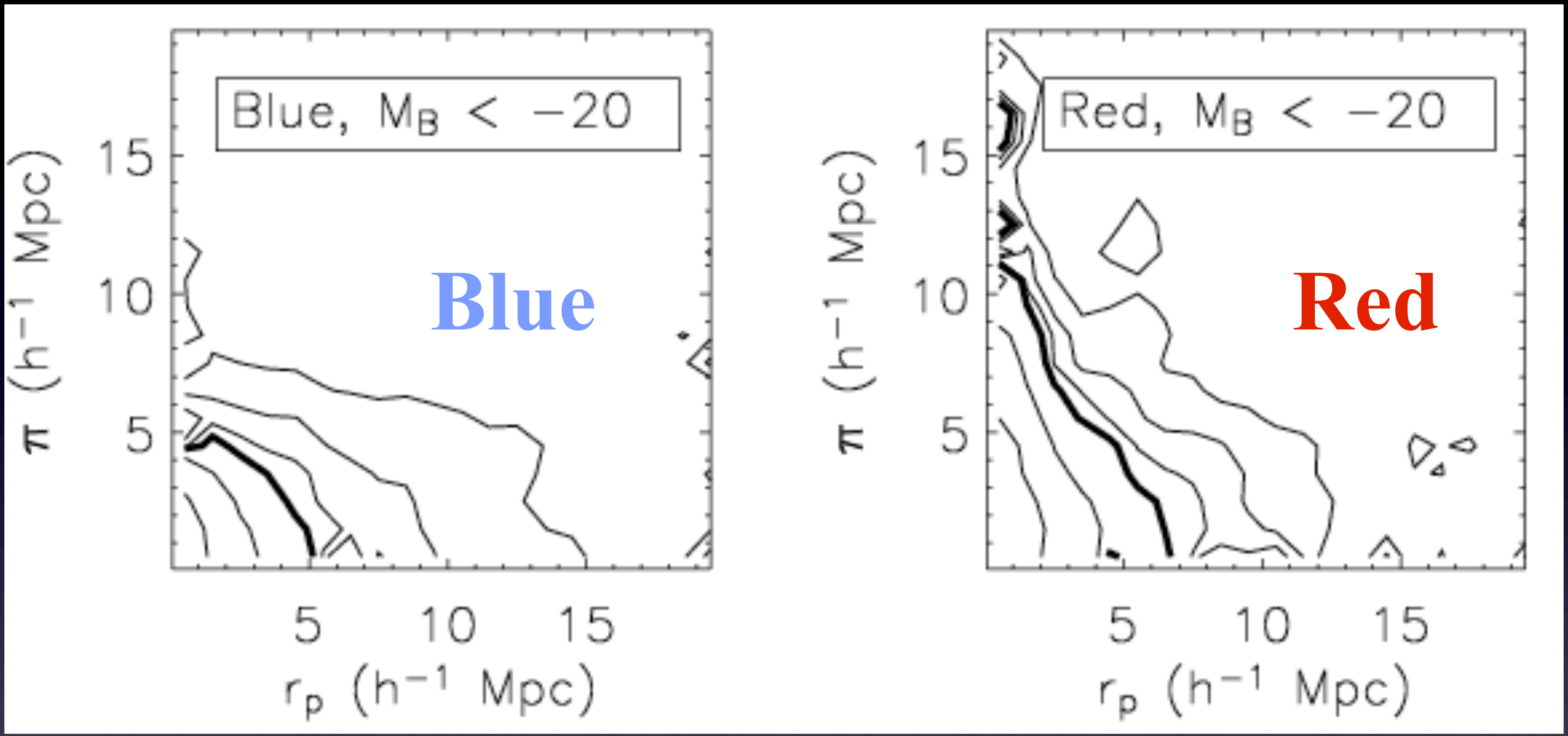
**redder galaxies more clustered than bluer galaxies - reside in more massive halos**

# Clustering at z=1

real data:



# Galaxy Clustering as a Function of Color




Coil et al. 2008

**Red galaxies have a larger $r_0$ and larger velocity dispersion: reside in more massive halos / virialized overdensities. Detect coherent infall on large scales for blue galaxies.**

# Clustering at z=1

Quantify minimum dark matter halos mass as a function of galaxy color (for $M_B$<-20):

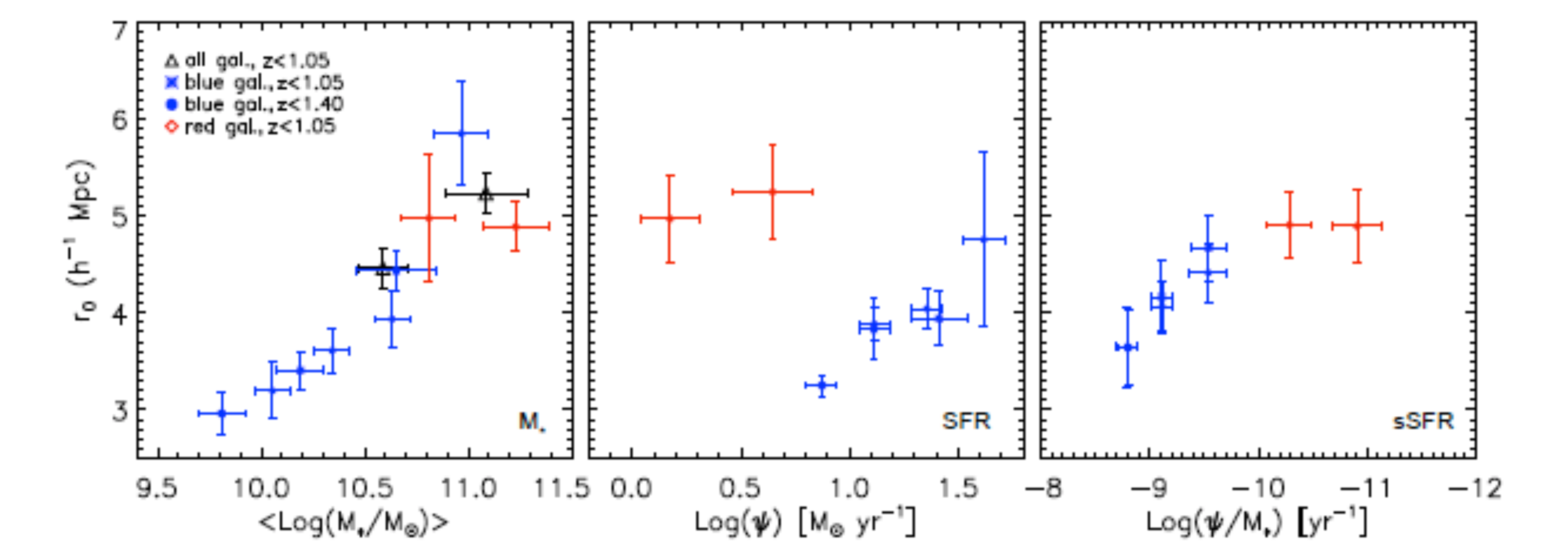
red: b=1.6, $M_{halo}$>2 $10^{12}$ $M_o$/h

blue: b=1.3, $M_{halo}$>4 $10^{11}$ $M_o$/h

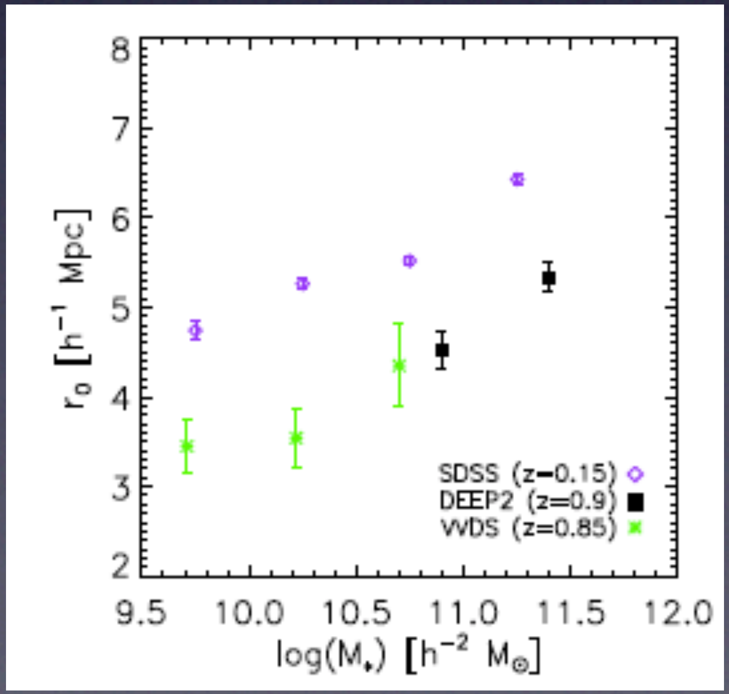
- important for color bimodality theories and simulations of gas accretion and star formation

Color-density relation is not caused by clusters. Only a few % of z=1 galaxies are in clusters. Either caused by physics in groups or intrinsic galaxy or halo property such as age / stellar mass / halo mass.

# Stellar Mass and SFR Dependence

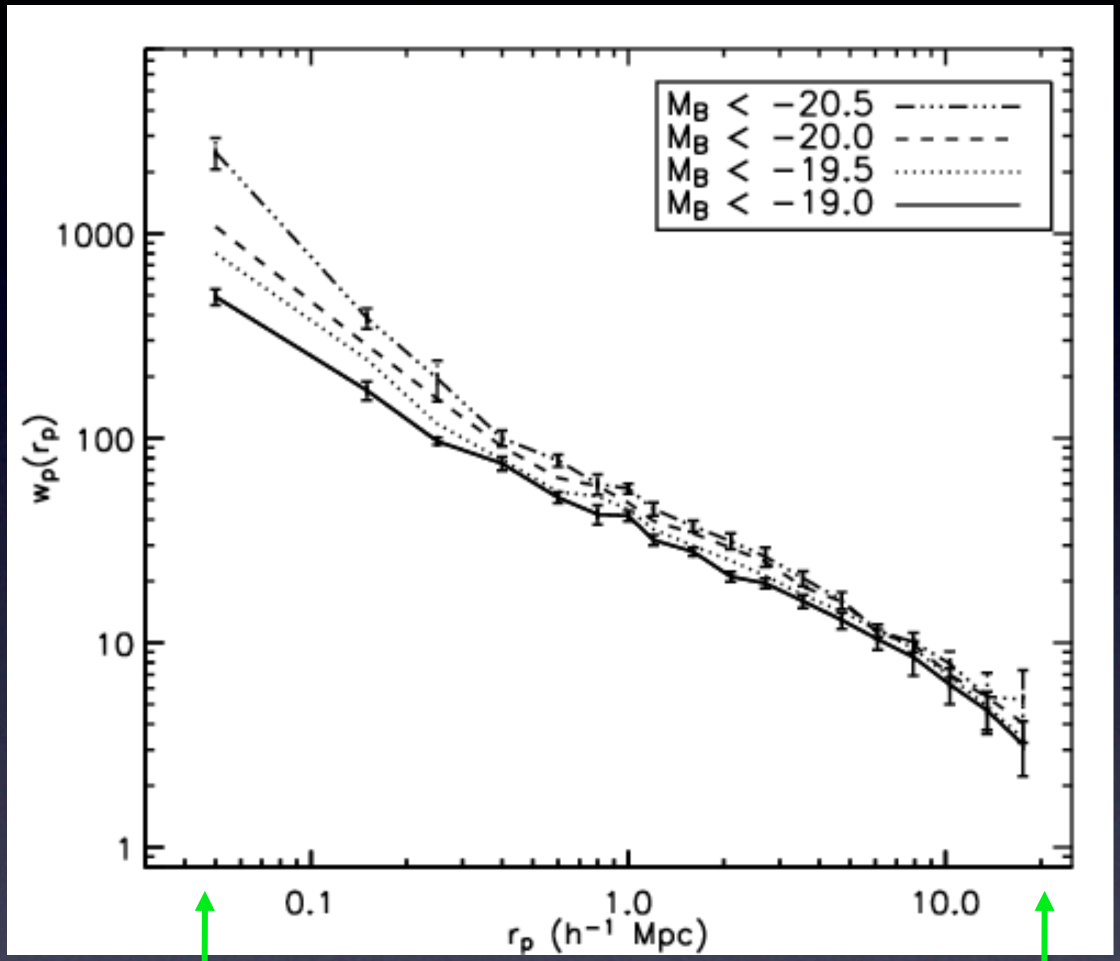



Mostek et al. 2013




- at a given stellar mass, clustering scale length grows with time

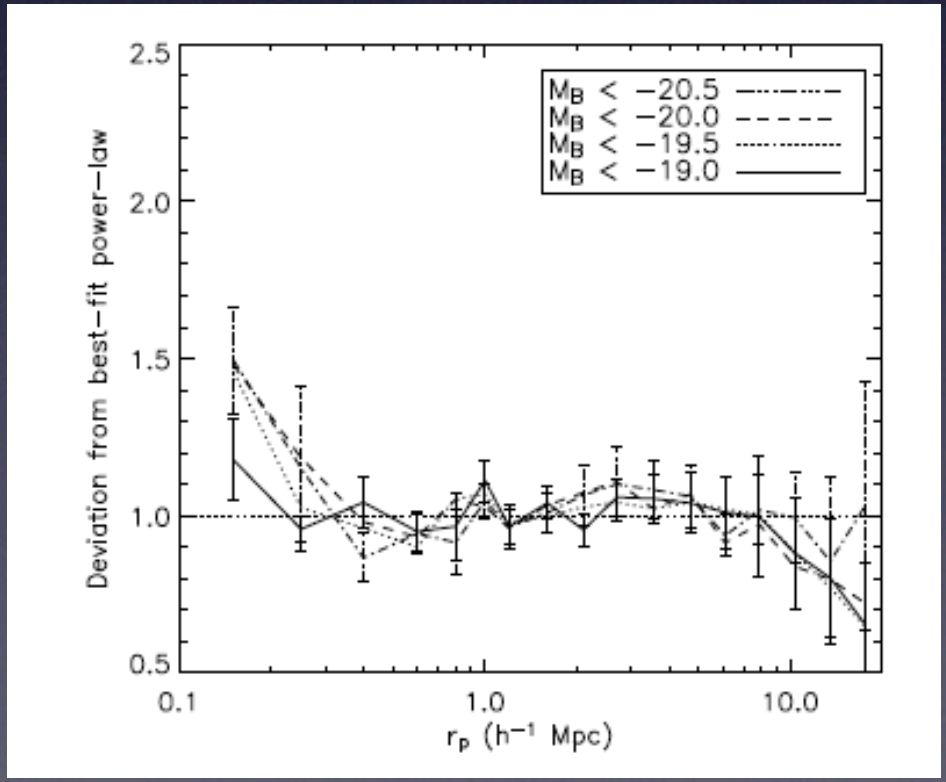
# Luminosity-dependence of clustering




**At z=1 brighter galaxies are more clustered and have steeper slopes on small scales -- preferentially found in groups -- sub-structure.**

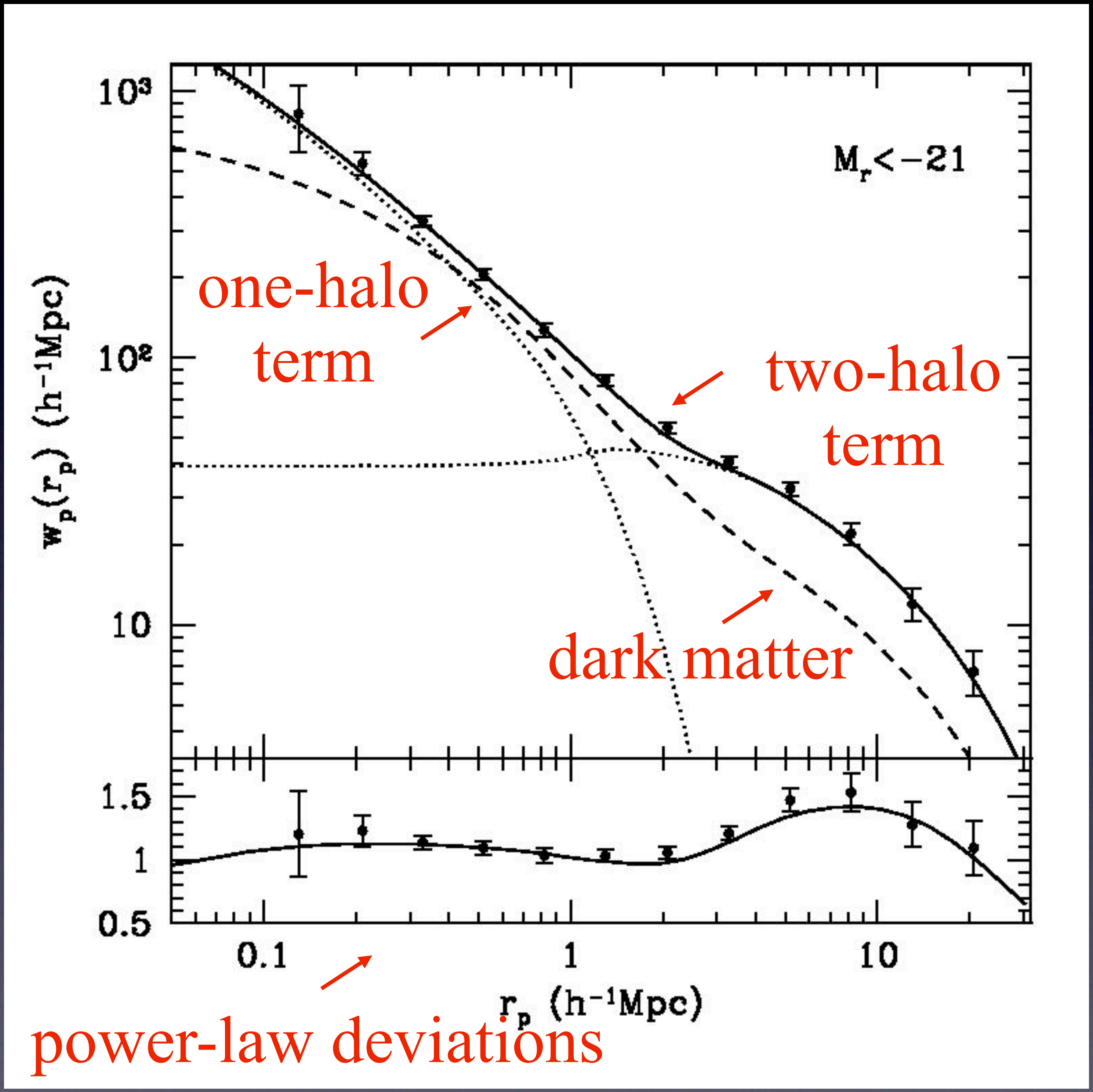
**50 kpc/h**

**20 Mpc/h**

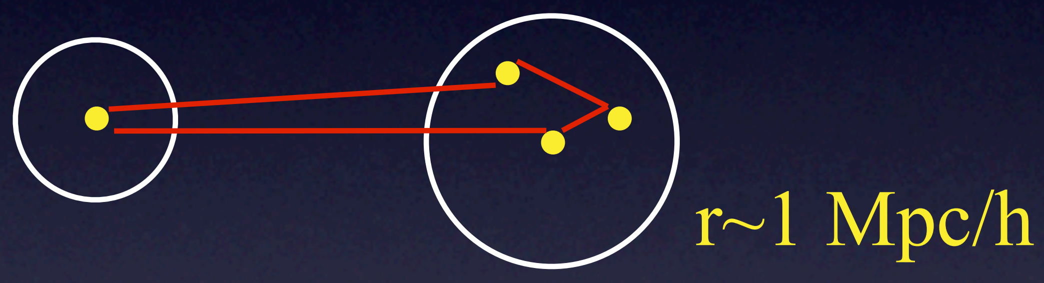



Coil et al. 2006

# Deviations from a power-law




Can fit $\xi(r)$ with an analytic halo occupation distribution (HOD) model - $N_{gals}(M_{halo})$

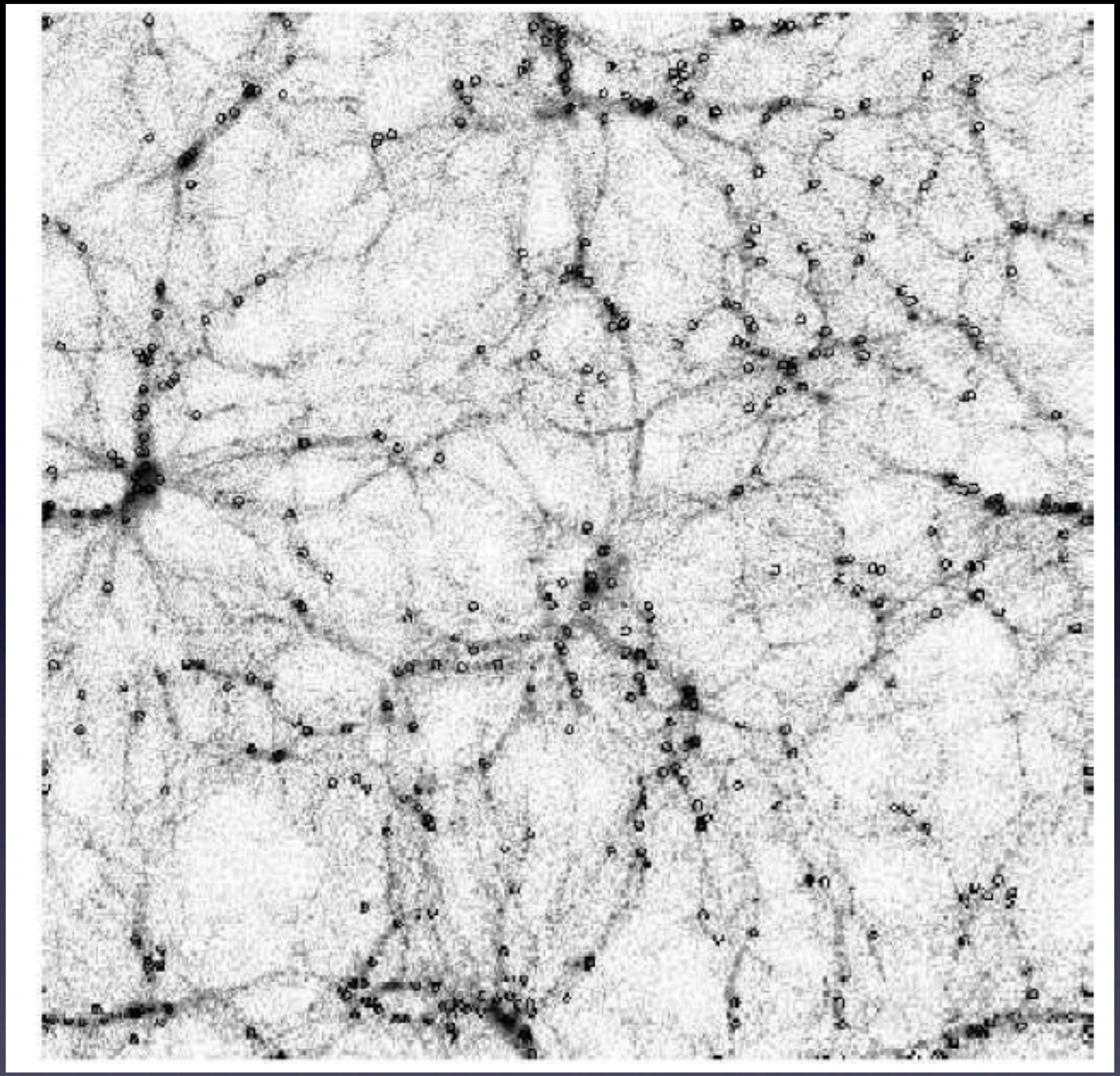



**Naturally explains ~power-law and fits the small deviations**

**Differentiates between 'central' galaxies and 'satellite' galaxies**

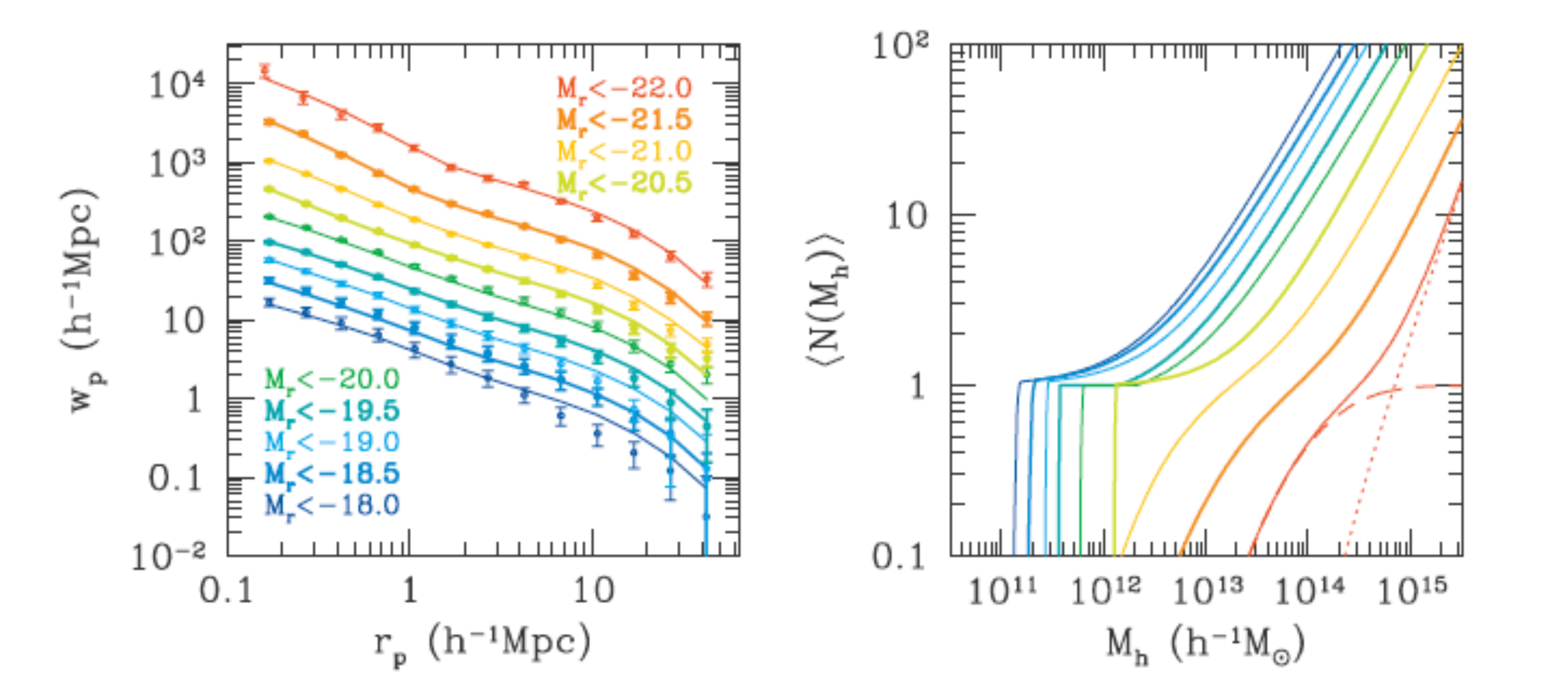
Zehavi et al. 2004

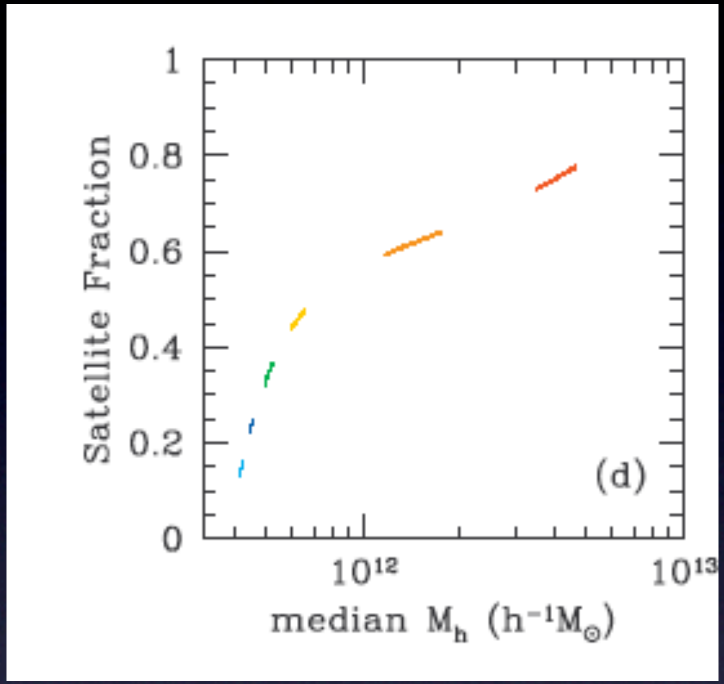
# Populating Halos with Galaxies



One ‘central’ galaxy in the middle of each halo, above a given mass threshold. Additional ‘satellite’ galaxies in higher mass halos at locations of random dark matter particles (or subhalos).

# HOD Modeling of $\xi(r)$




**Perform direct HOD fits to data in order to:**

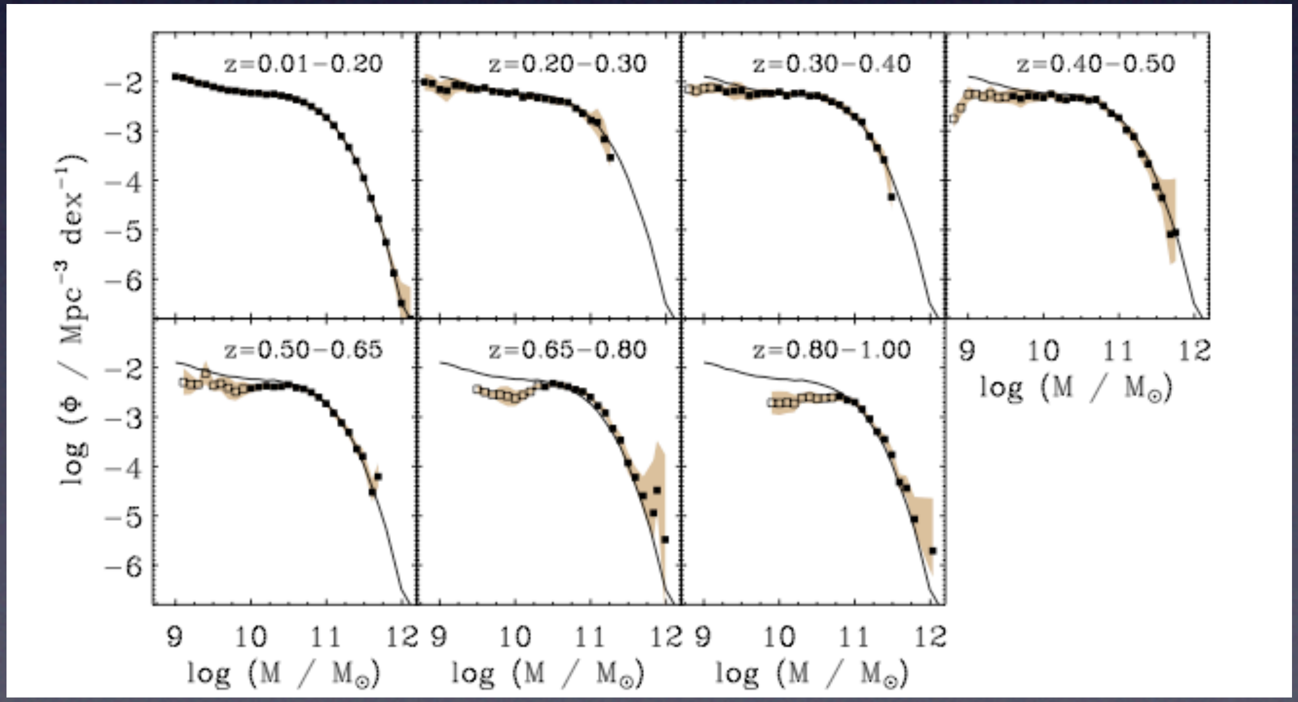
**- measure the detailed relation between luminosity and halo mass**

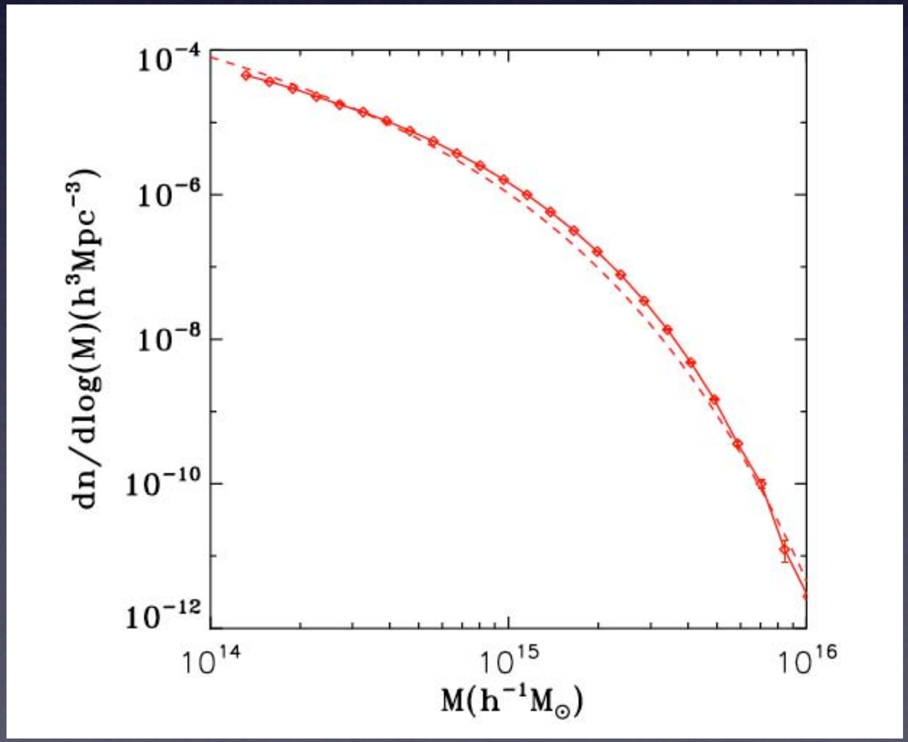
**- measure satellite fraction and how it depends on luminosity / color (ex: at z=0 25% of blue galaxies and 60% of red galaxies are satellites)**

**- determine the halo mass distributions for central vs. satellite galaxies**

# Abundance Matching

Rank galaxies by luminosity or stellar mass, assign them to halos ranked by total mass. Best to take into account scatter between luminosity or stellar mass and halo mass.




Moustakas et al. 2013

# Connecting Observed Galaxy Samples Across Time

The number densities and clustering properties of different galaxy samples observed at different z can be used to constrain progenitor / descendent populations:

- Take a galaxy sample at high-z, find halos in simulations that have same clustering and number density. Follow their merger histories to lower z to predict their clustering and number density at lower z - connect with an observed galaxy population.

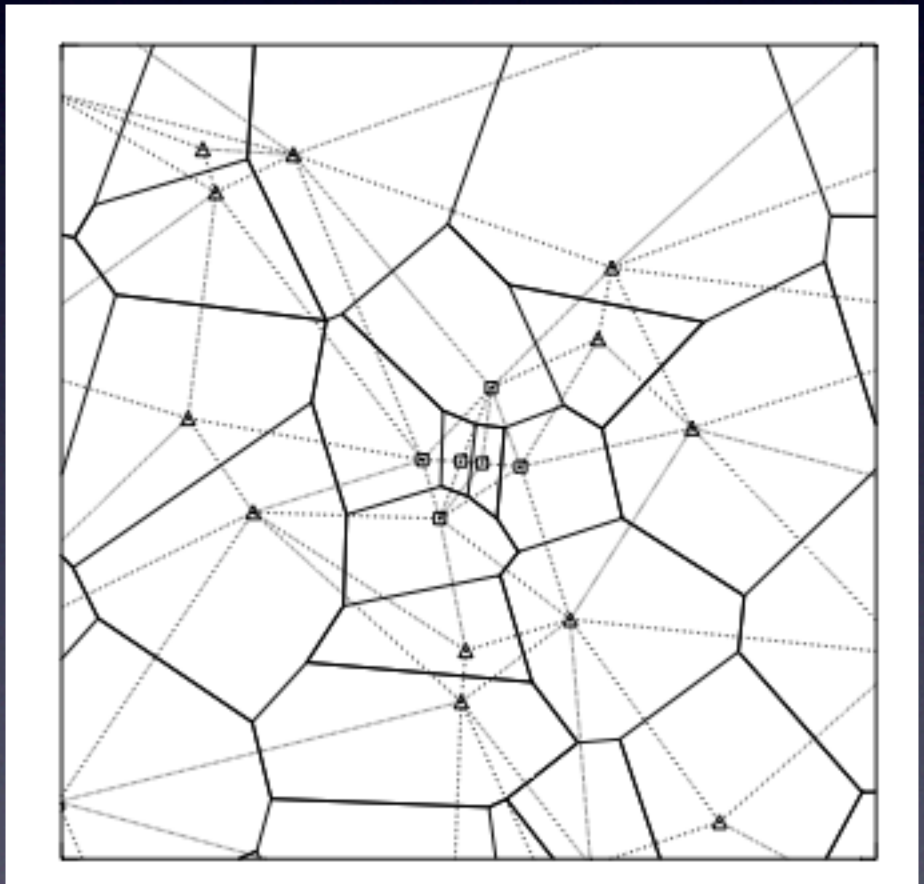
# Higher-order Clustering Measurements

Can measure 3-pt, 4-pt, etc. correlation functions.
For the 3-pt function, find triangles of galaxies with different configuration shapes. Higher-order measurements need large samples over enormous volumes - only been done at z=0 so far.

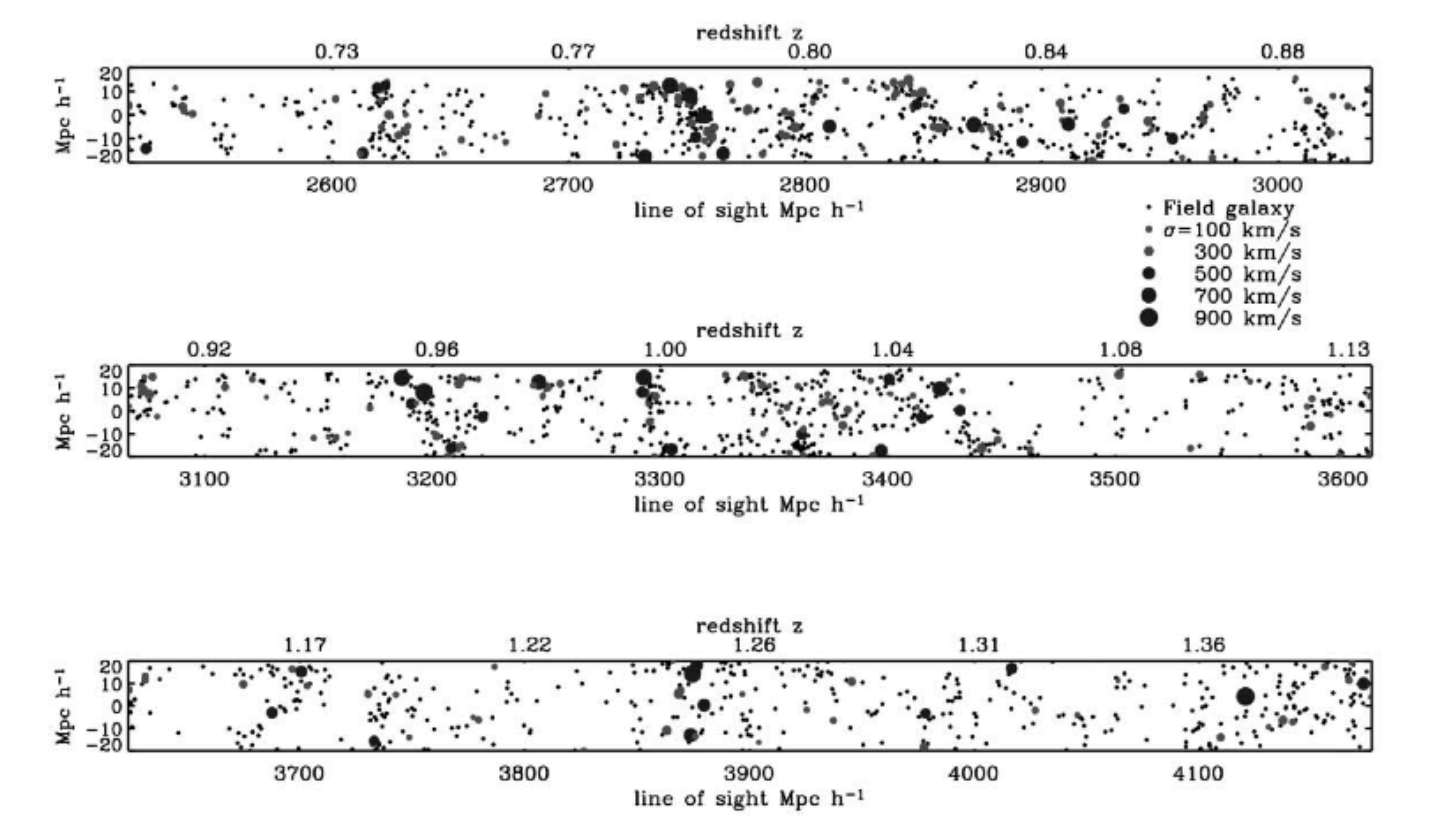
Can compare the 3-pt and 2-pt functions to measure the galaxy bias directly. Can also measure any non-linear bias.

# Identifying Galaxy Groups

Can find groups using the locations of galaxies in redshift space - no selection based on color, magnitude, etc. - just overdensity in the galaxy distribution.

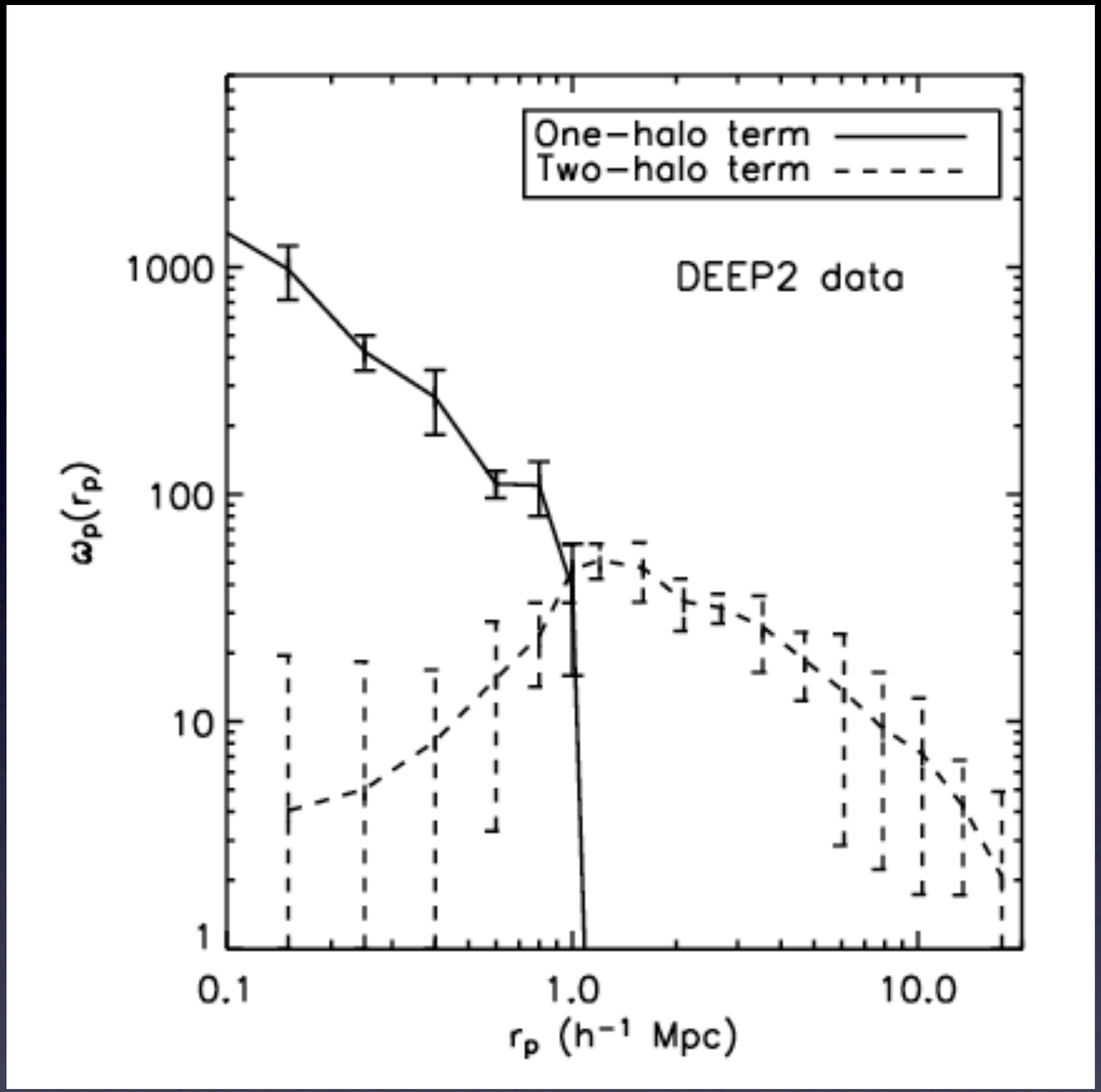


Voronoi tesselation




Groups at z=1

# Measure one-halo and two-halo terms

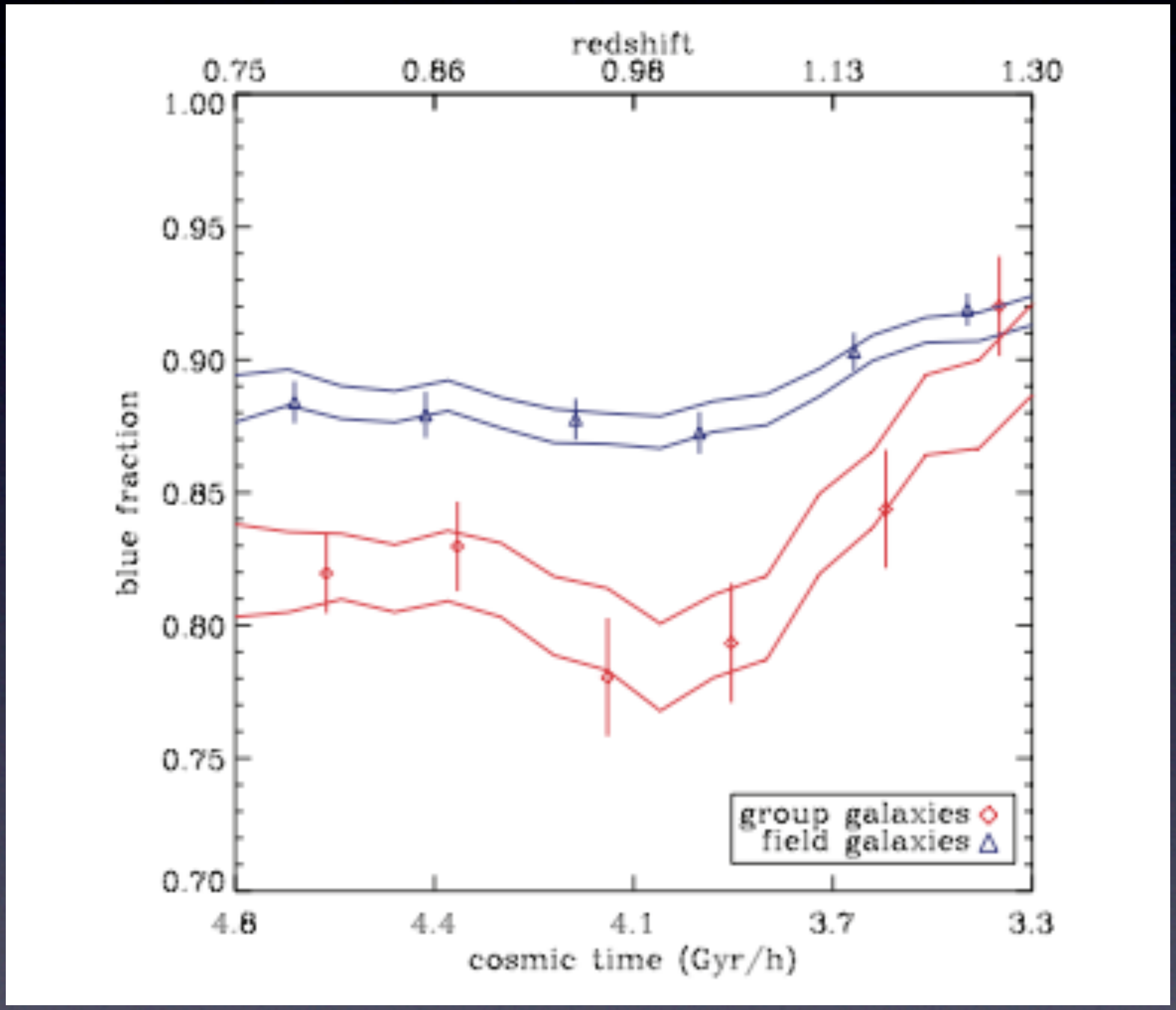



Coil et al. 2006

**Can measure the one-halo and two-halo terms directly with a group catalog!**

**One halo term constrains the radial profile of galaxies w/in dark matter halos.**

# Fraction of Blue Galaxies in Groups




Gerke et al. 2007

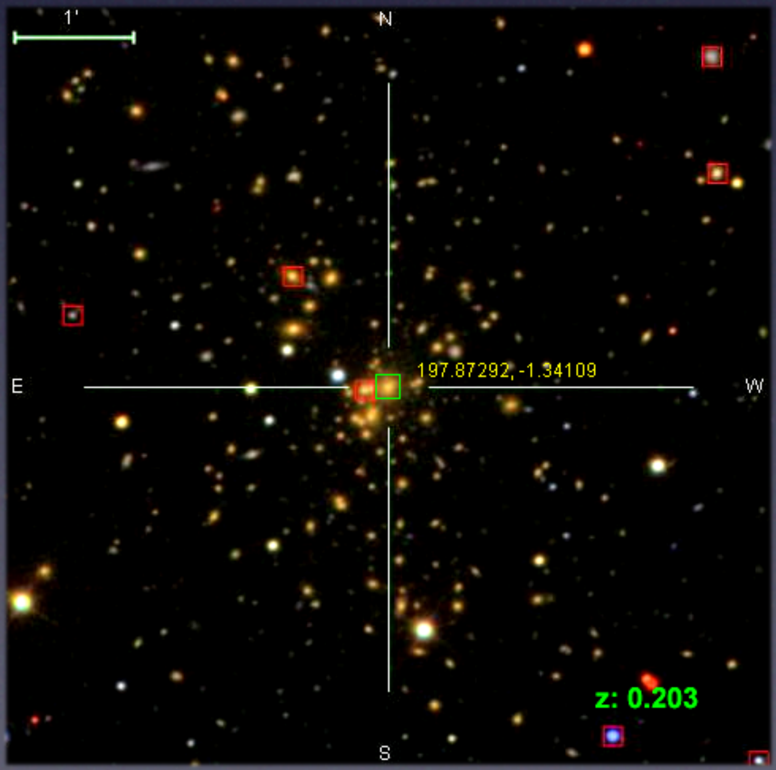
# Identifying Clusters in SDSS

'maxBCG' - optically-selected, identifies bright red galaxies and finds spatial overdensities, uses fact that most low-z clusters have a 'brightest cluster galaxy' at the center

'GMBCG' - optically-selected, finds differences between cluster galaxies and field galaxies in color space - differential detection

Table 1: Summary of optical cluster finding algorithms for photometric data

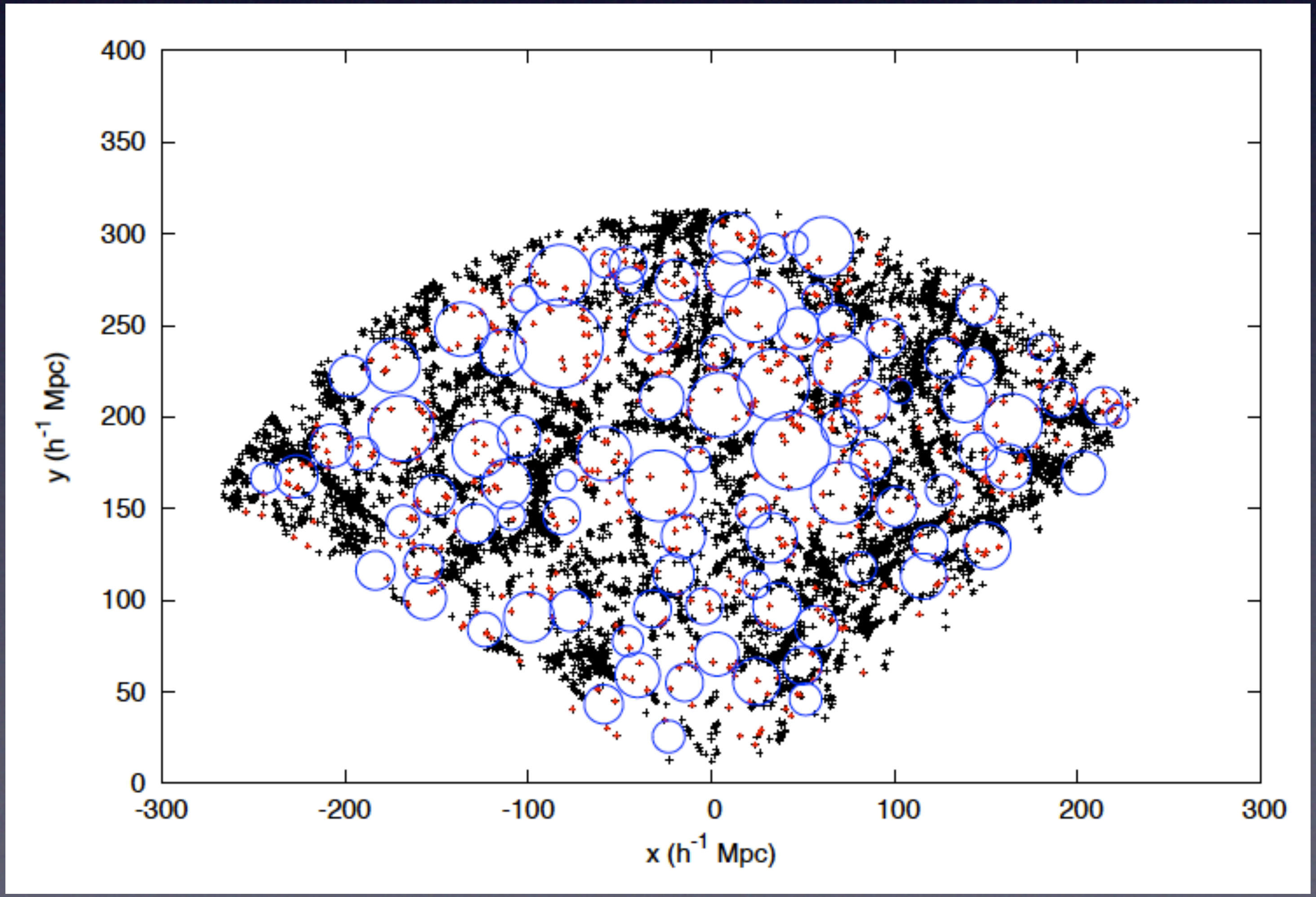
| Algorithm | Type of data applied | De-projection method |
|---|---|---|
| Percolation[a] | Single band/Simulation | Magnitude/photo-$z$ |
| Smoothing Kernels[b] | Single band | Magnitude |
| Adaptive Kernel[c] | Single band | Magnitude |
| Matched Filter [d] | Single band | Magnitude |
| Hybrid and Adaptive Matched Filter[e] | Single band | Magnitude/photo-$z$ |
| Voronoi Tessellation[f] | Single band | Magnitude |
| Cut-and-Enhance [g] | Single band | Magnitude |
| Modified Friends of Friends[h] | Multi-band | Photo-$z$ |
| C4[i] | Multi-band | All Colors |
| Percolation with Spectroscopic redshift[j] | Multi-Band | Spectroscopic Redshift |
| Cluster Red Sequence[k] | Multi-band | Red sequence |
| MaxBCG [l] | Multi-band | Red sequence |
| WHL[m] | Multi-band | Photo-$z$ |
| GMBCG[n] | Multi-band | Red sequence |



# Voids

Void statistics are closely tied to cosmological parameters.
Can test for non-Gaussianity of primordial perturbations.

Voids are not entirely devoid of galaxies - some 'void galaxies' live in very underdense regions. Can test extremes of galaxy formation using them.




Pan et al. 2011

# Voids

LCDM simulations have clear predictions for voids:

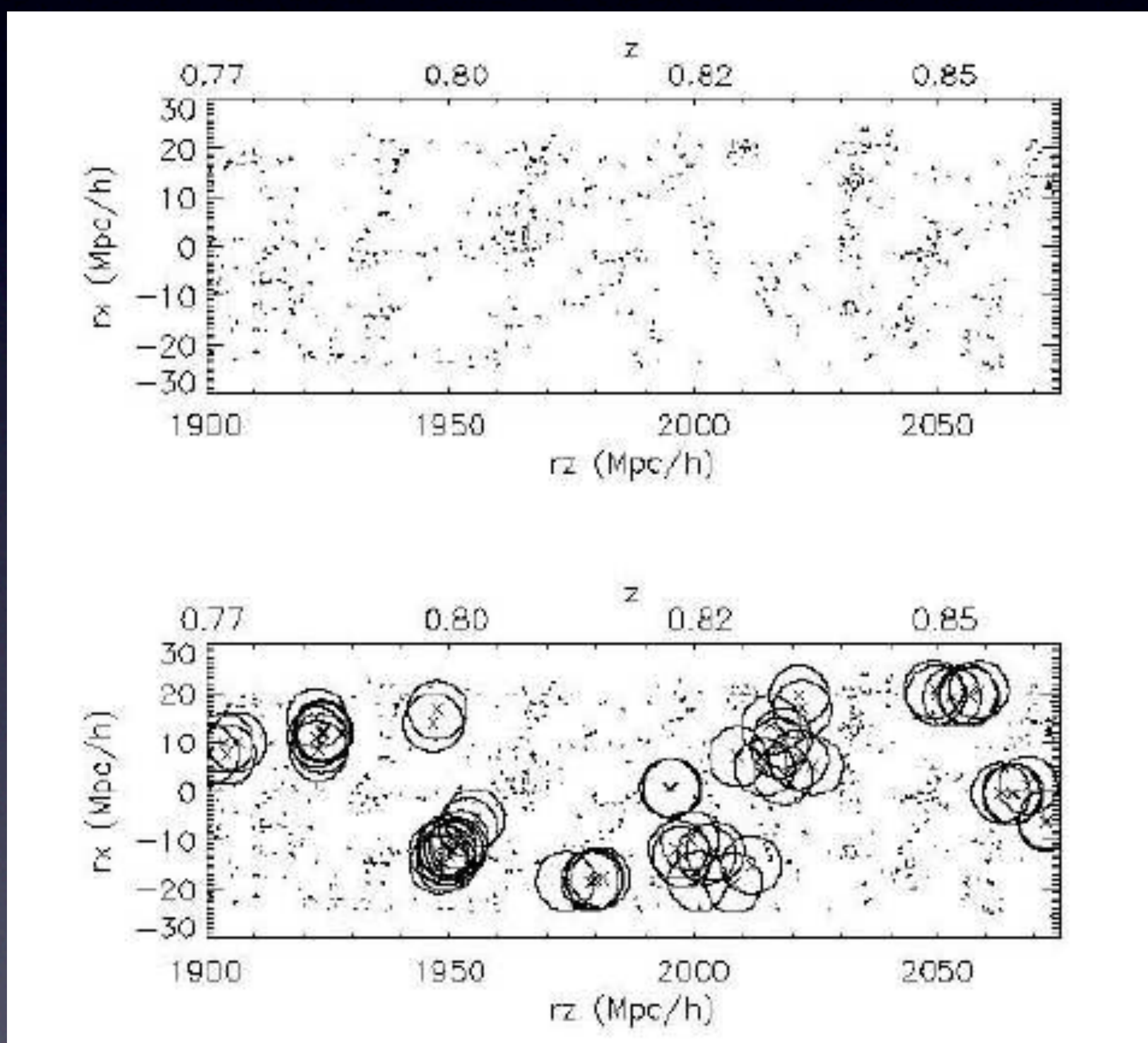
- should contain both dark matter and galaxies, though very underdense
- have well-defined, sharp edges in the dark matter density
- volume of space filled by voids evolves strongly: ~60% at z=0, 28% at z=1, 9% at z=2
- galaxies in voids should be anti-biased relative to dark matter

# Voids

Observations find:

- void sizes are typically ~15 Mpc/h
- very underdense but do have galaxies
- ~60% of volume at z=0.1 is filled by voids
- have sharp density profiles as traced by galaxies

# Void Probability Function



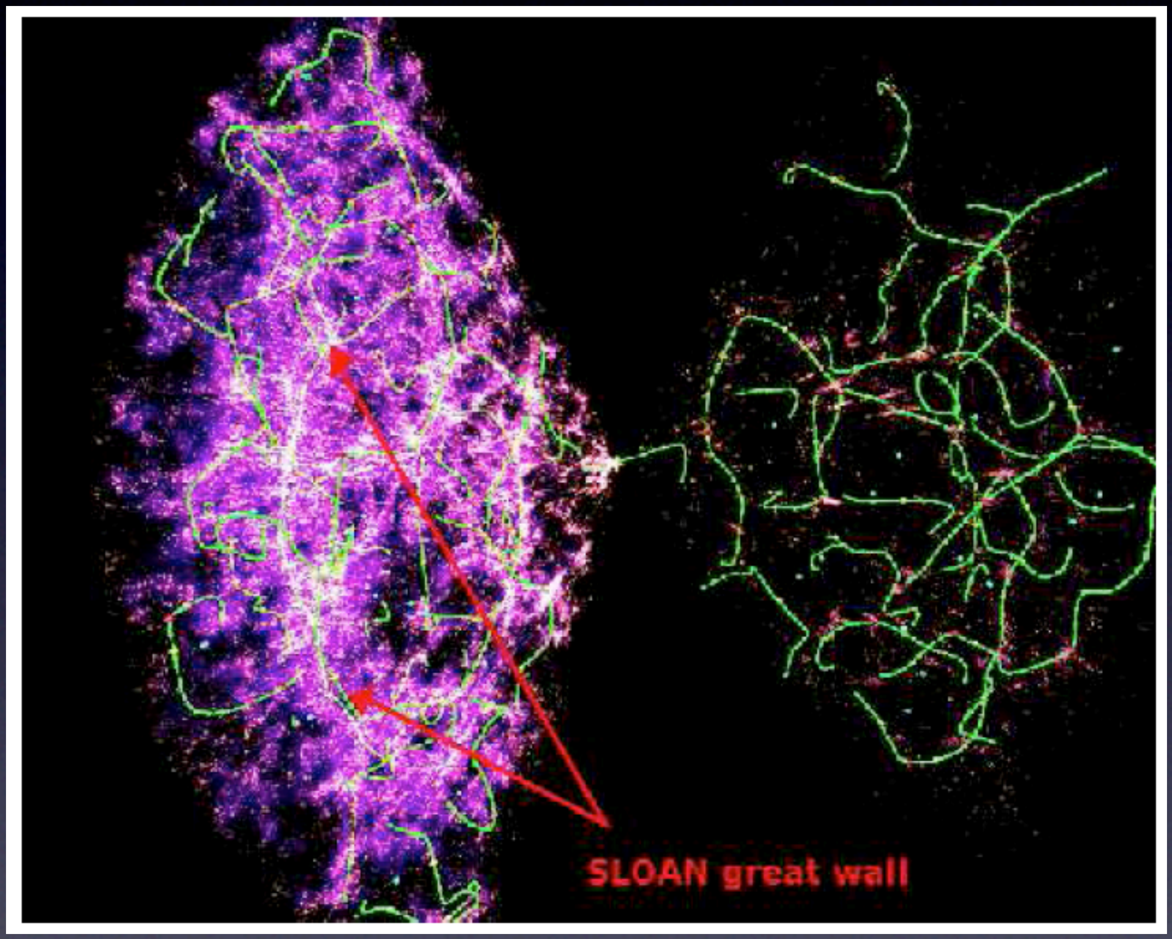

Conroy et al. 2005

The probability that a randomly placed sphere of radius R will not contain any galaxies. Defined such that it depends on the space density of points.

VPF should be higher for galaxies than for dark matter, due to bias and discretization. Observationally, it is higher for red galaxies than blue galaxies, and evolves with redshift (voids grow with time).

HOD model results agree with observed VPF - no need suppress galaxy formation in voids. Galaxy formation physics is not different in low density regions.

# Filaments




Sousbie et al. 2008

Comparisons of filament statistics - length distribution, typical width - also agree well with LCDM simulations.

Observations show that filaments don't change much from z=1 to z=0, which is expected, unlike voids.

# Limits of Current Surveys

Local Universe well sampled due to SDSS.
To z~1 largest deep galaxy surveys have ~100k redshifts, cover ~10 sq. deg.
At z~2-3 only a few thousand redshifts in small areas.

Need spectrographs that can observe ~1000 galaxies at once, on the largest telescopes. These all use fibers.

New surveys using prism / grism from ground (PRIMUS) and space (3D-HST). If use a slitmask then spectra don't overlap.