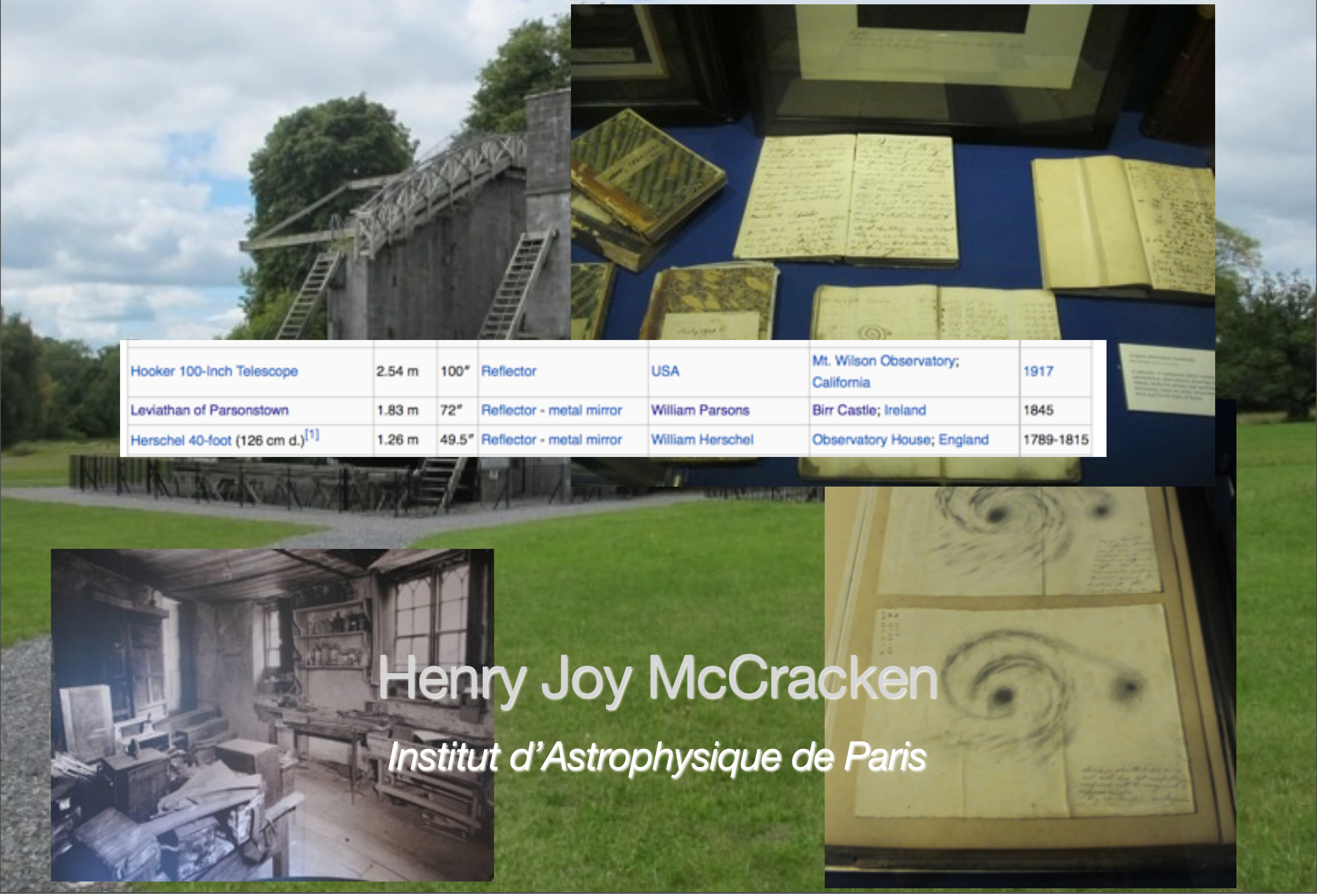
Watching haloes fill with galaxies

Henry Joy McCracken

Institut d’Astrophysique de Paris

Watching haloes fill with galaxies



Hooker 100-Inch Telescope	2.54 m	100″	Reflector	USA	Mt. Wilson Observatory; California	1917
Leviathan of Parsonstown	1.83 m	72″	Reflector - metal mirror	William Parsons	Birr Castle; Ireland	1845
Herschel 40-foot (126 cm d.)[1]	1.26 m	49.5″	Reflector - metal mirror	William Herschel	Observatory House; England	1789-1815

Henry Joy McCracken

Institut d'Astrophysique de Paris

Collaborators

Martin Kilbinger: *HOD modelling, parameter estimation*

Jean Coupon, *HOD modelling, CFHTLS*

Olivier Ilbert : *Photometric redshifts*

Yannick Mellier: *CFHTLS*

Nick Scoville: *COSMOS*

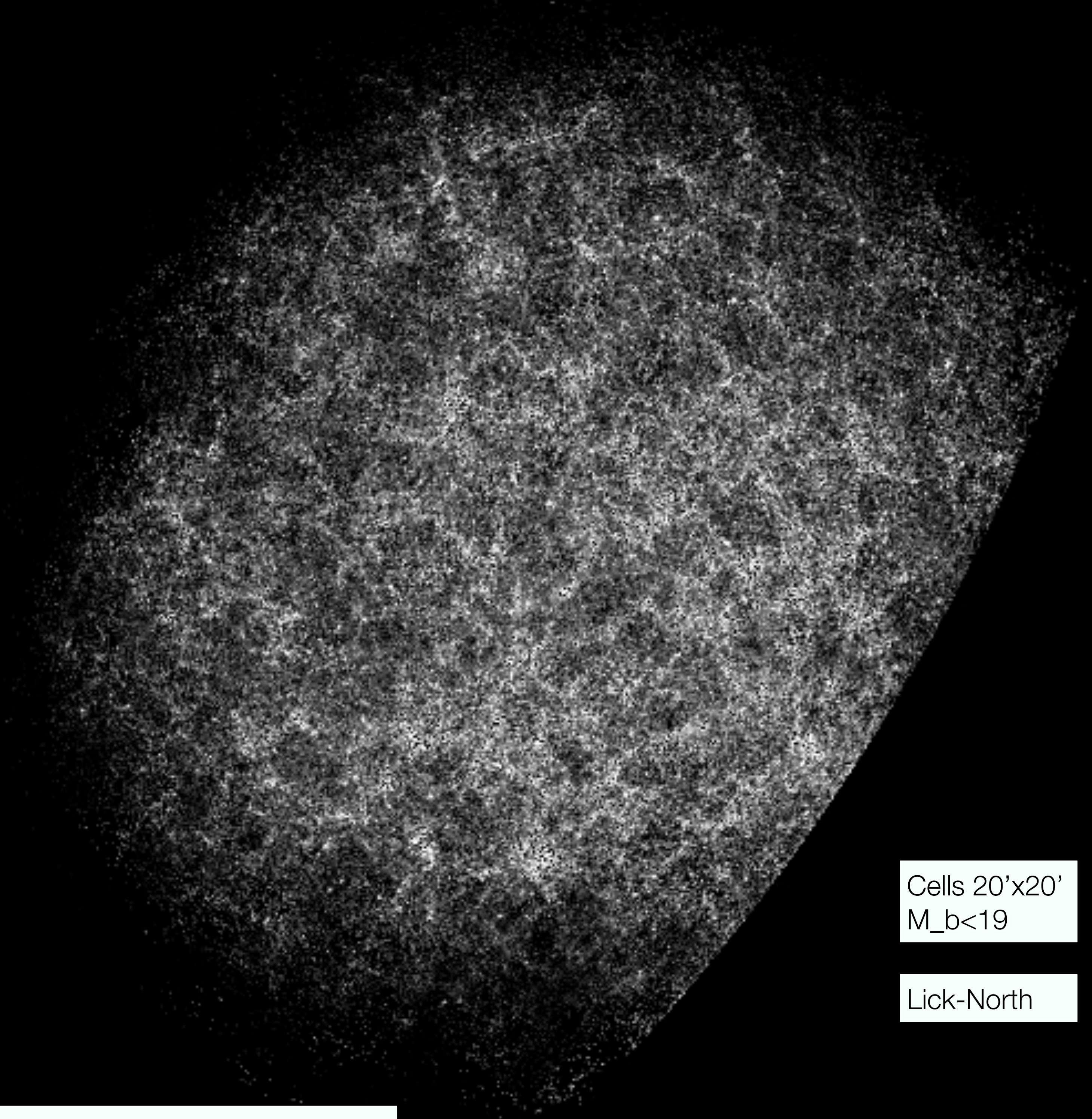
Peter Capak: *COSMOS catalogue production*

Herve Aussel: *COSMOS catalogue production*

Mara Salvato: *COSMOS*

Gigi Guzzo, Olivier Le Fevre: *Spectroscopic surveys*

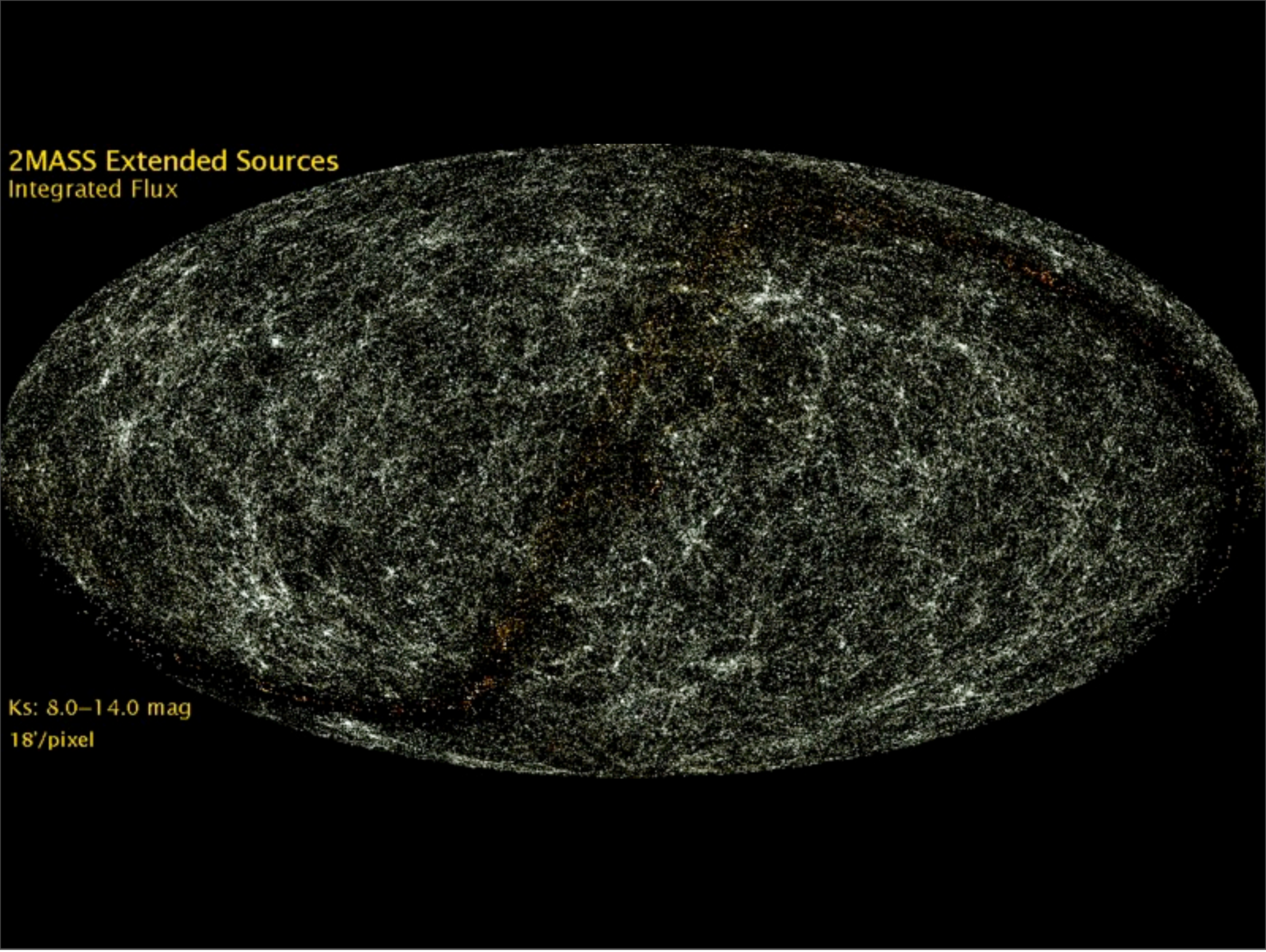
Emanuele Daddi, Patrick Hudelot: *Near IR data*

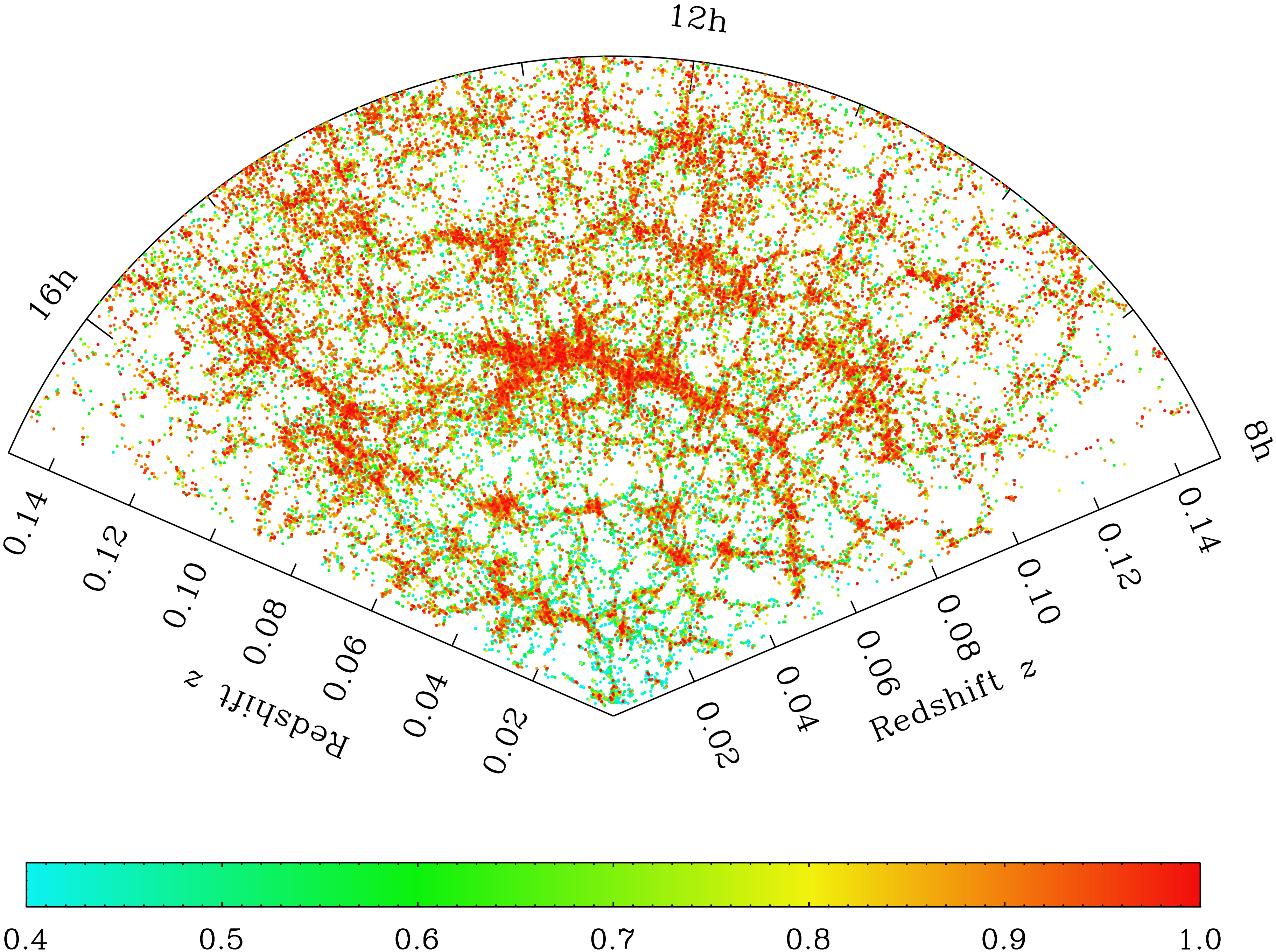


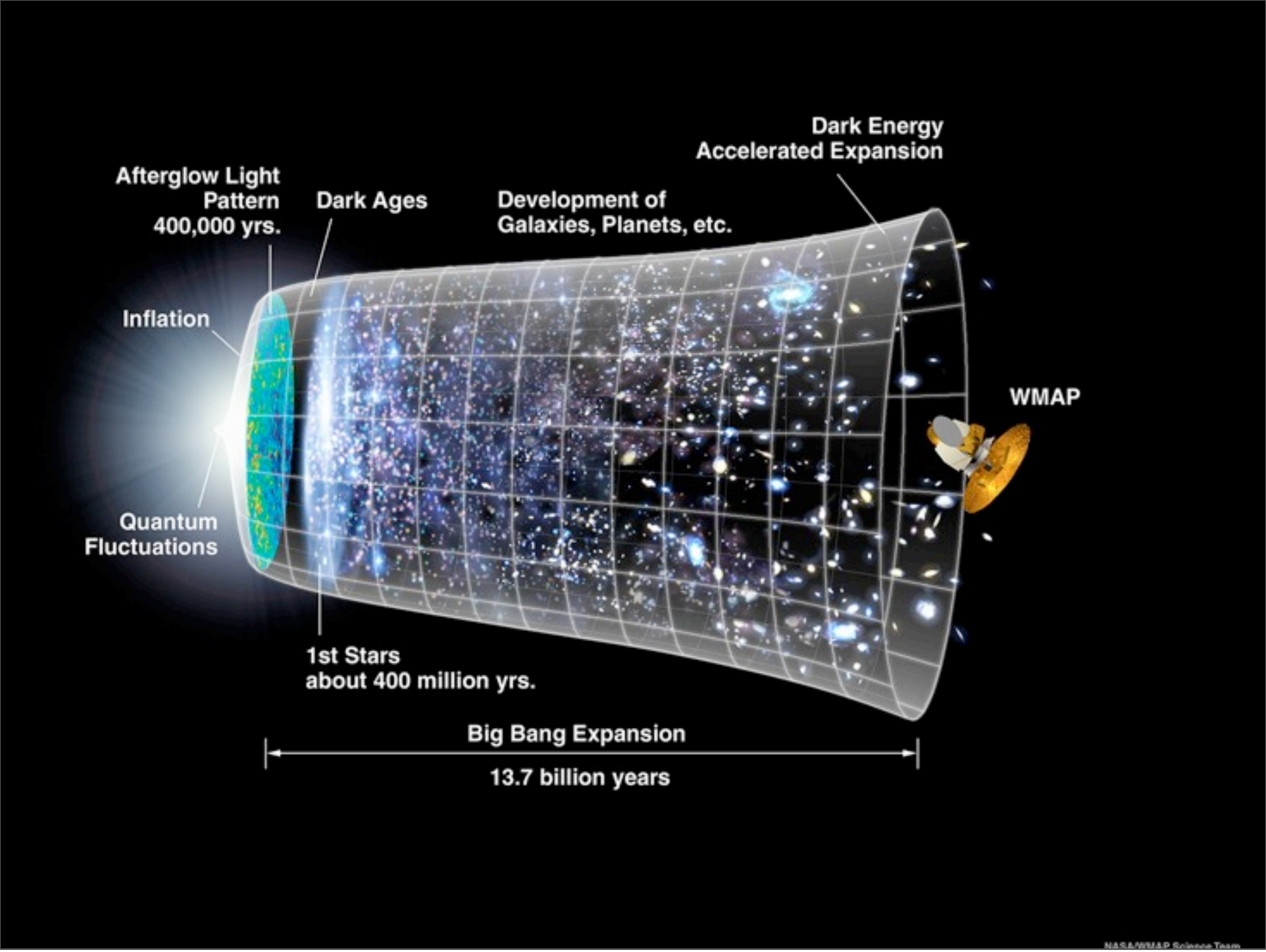
Cells 20’x20’
M_b<19

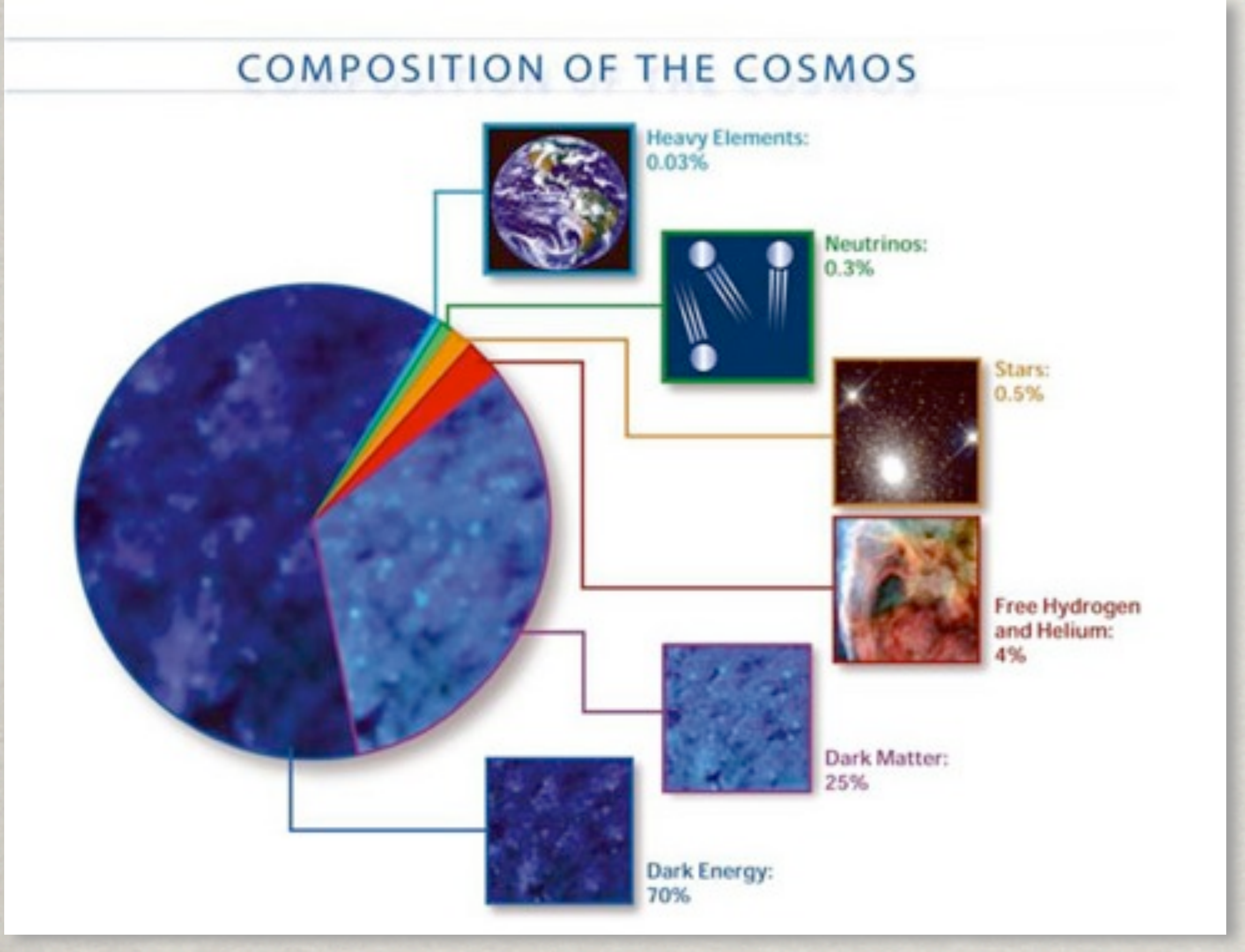
Lick-North

Seldner, Siebers, Groth, Peebles 1977


2MASS Extended Sources
Integrated Flux
Ks: 8.0–14.0 mag
18′/pixel

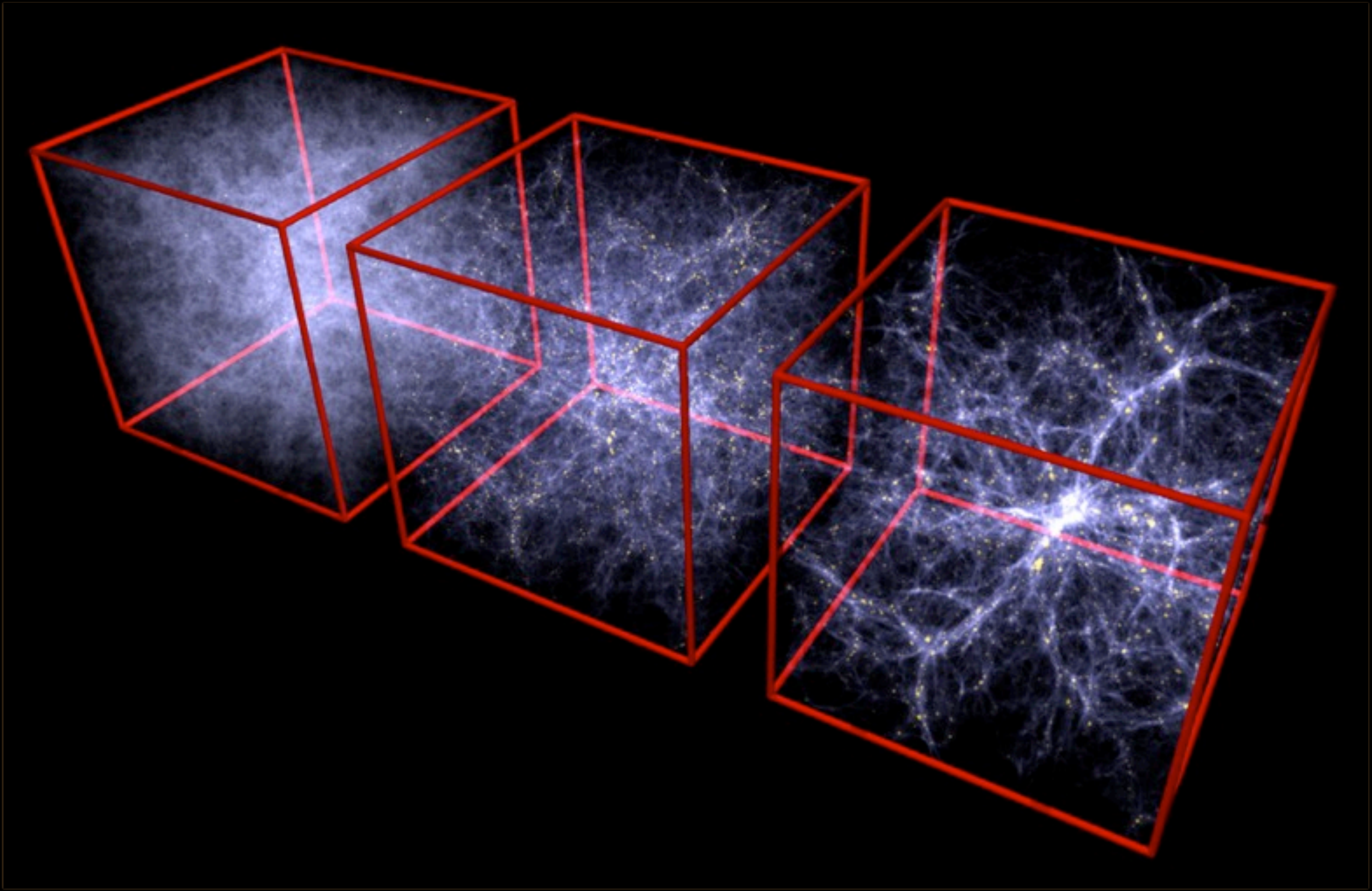

12h
16h
8h
0.14
0.12
0.10
0.08
0.06
0.04
0.02
Redshift z
0.02
0.04
0.06
0.08
0.10
0.12
0.14
Redshift z
0.4
0.5
0.6
0.7
0.8
0.9
1.0
g−r


Dark Energy
Accelerated Expansion
Afterglow Light
Pattern
400,000 yrs.
Dark Ages
Development of
Galaxies, Planets, etc.
Inflation
WMAP
Quantum
Fluctuations
1st Stars
about 400 million yrs.
Big Bang Expansion
13.7 billion years
NASA/WMAP Science Team



Luminous matter (like stars and galaxies) comprises less than 1% of the energy content of the Universe.

Galaxy formation and evolution is dominated by dark matter



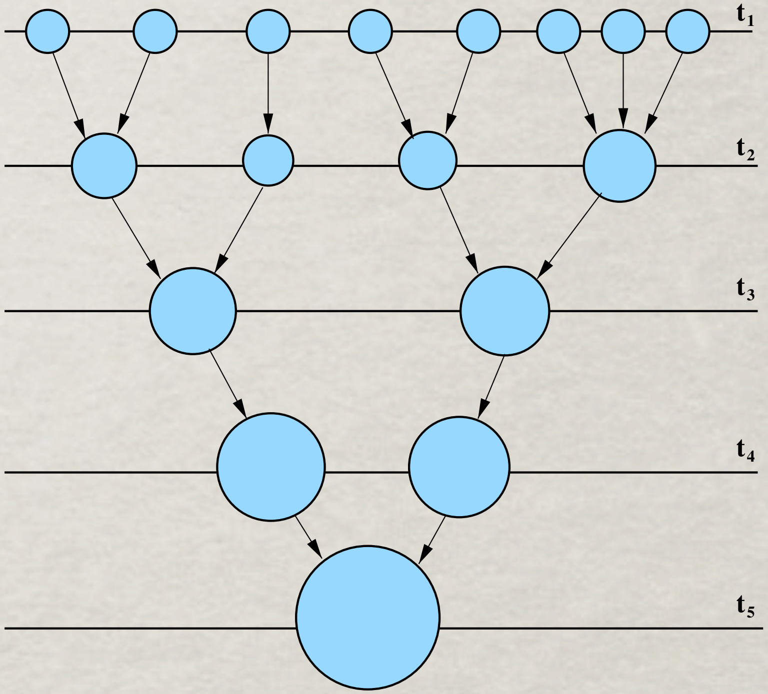
On galaxy formation

Our current galaxy formation paradigm:

Haloes of dark matter accrete and grow under the action of gravity

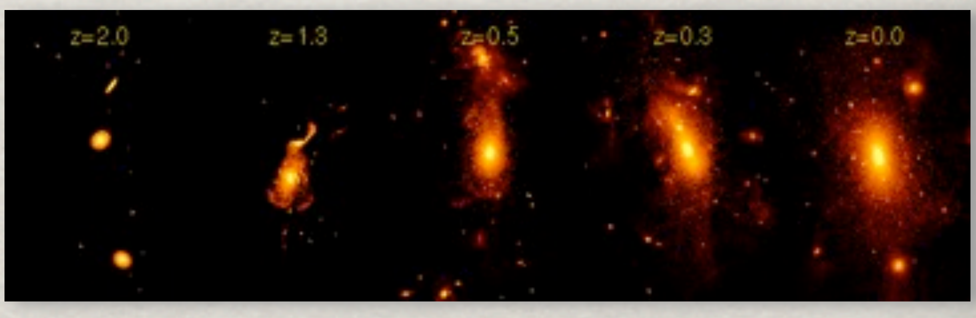
Baryonic matter gathers in these potential wells and forms stars and galaxies

The properties of dark matter haloes provide important information in addressing galaxy formation



Two key observable quantities:

- **When did most galaxies form stars?**
- **When did most mass assemble into galaxies?**
- **What is the relationship between the dark matter haloes and the galaxies?**



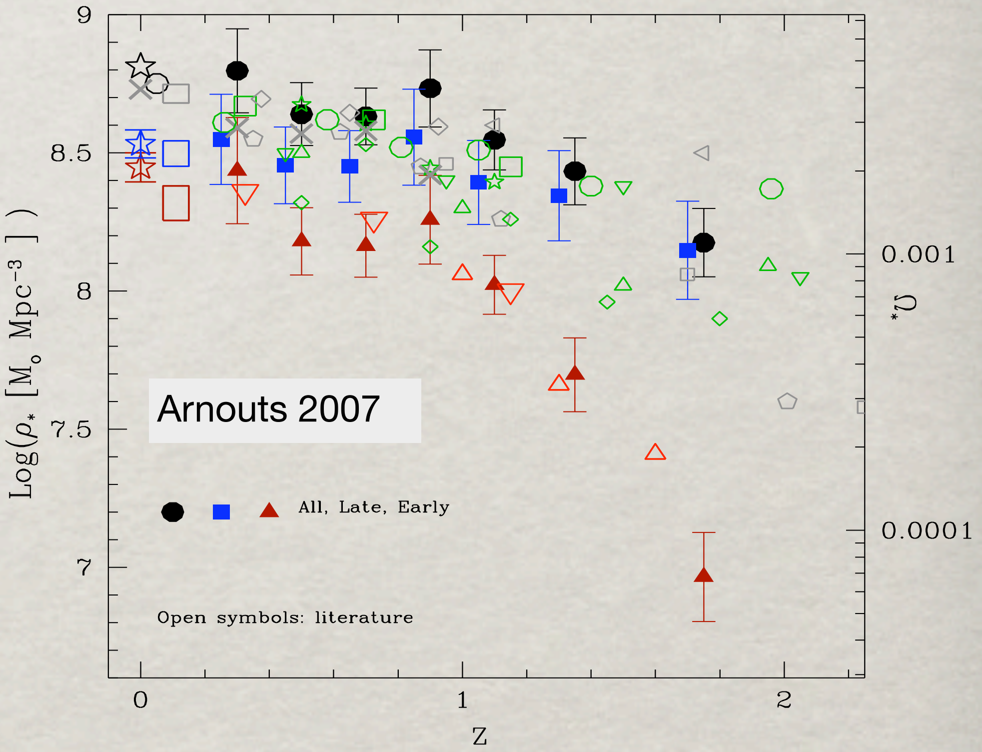
Galaxies, haloes and star-formation

What stops star-formation in massive red galaxies? Why do the sites of star-formation migrate from high-mass to low mass haloes?

How does star-formation rate relate to halo mass?

How does the halo mass where star-formation is most efficient evolve from 0<z<2? How does the peak luminosity of galaxies undergoing star-formation change with redshift?

What is the the role of environment in quenching star-formation? Are galaxies in over-dense regions more strongly quenched?



Arnouts 2007

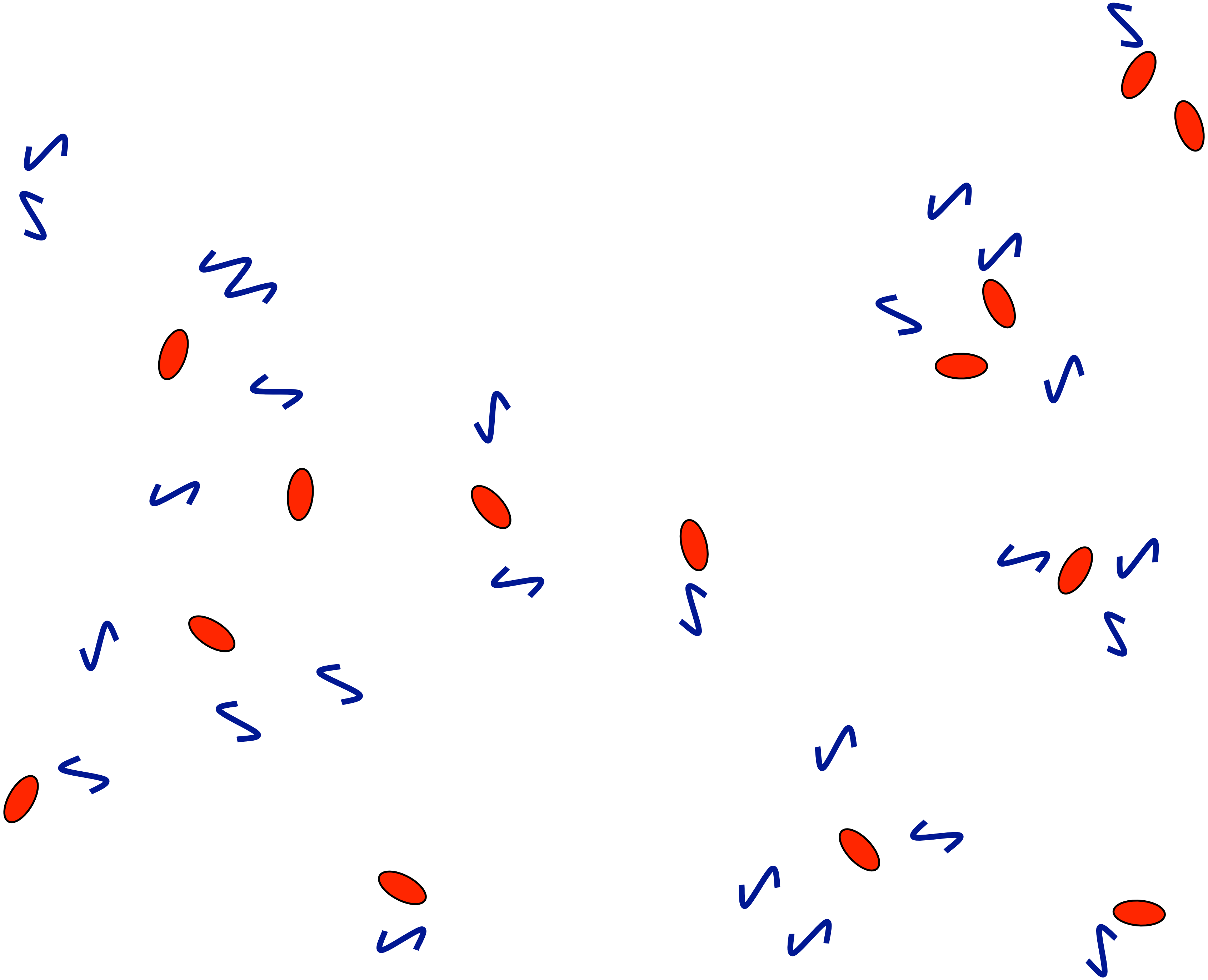
Relating galaxies and dark matter

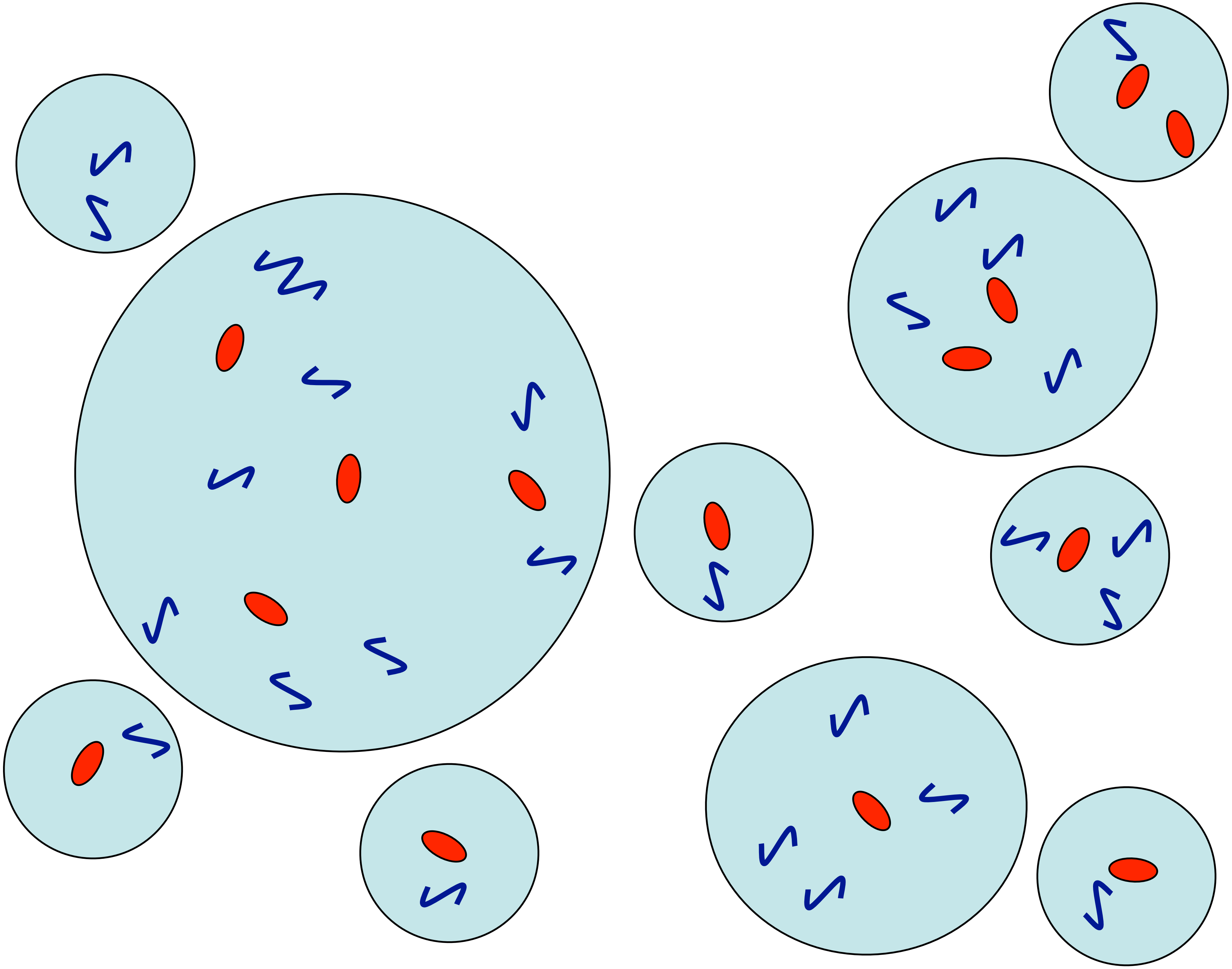
We would like to find a way to relate observed, visible galaxies with the underlying dark matter.

In ideal world: **full hydrodynamic simulations**: but these are complicated, reserved for experts, and have difficult reaching the required resolution at the current day.

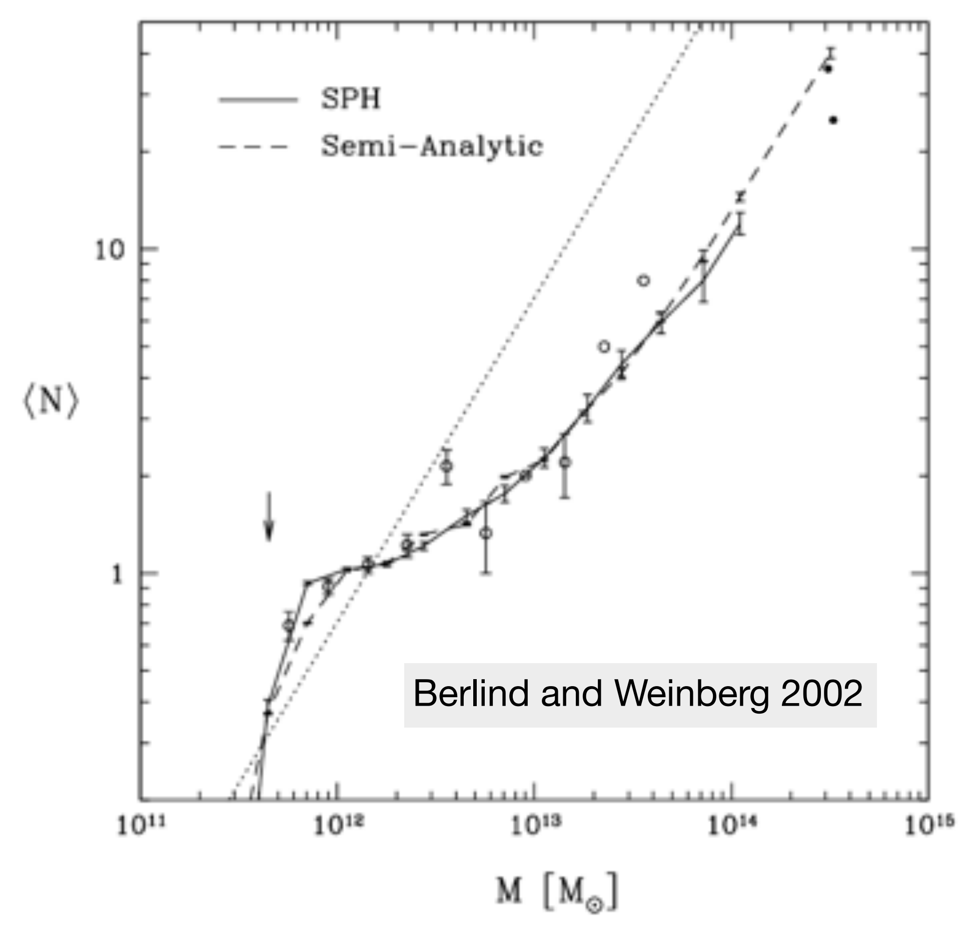
Another possibility are **semi-analytic models** in which galaxies are "painted on" to the dark matter haloes using a series of analytic recipes. Promising, but still for experts.

A simpler phenomenological model that observers could use...

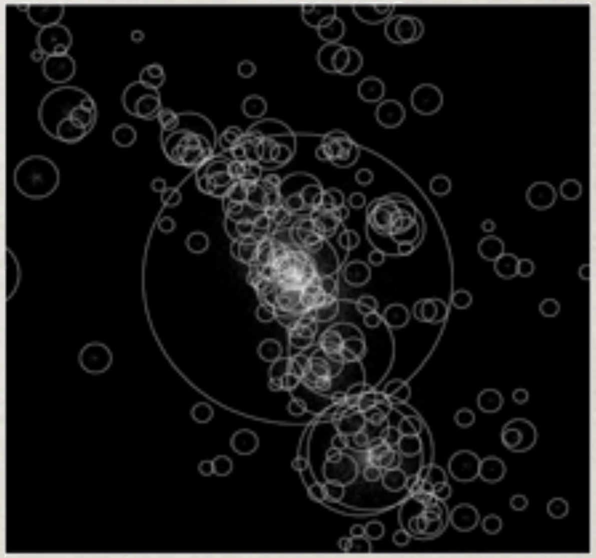




How do galaxies populate haloes?



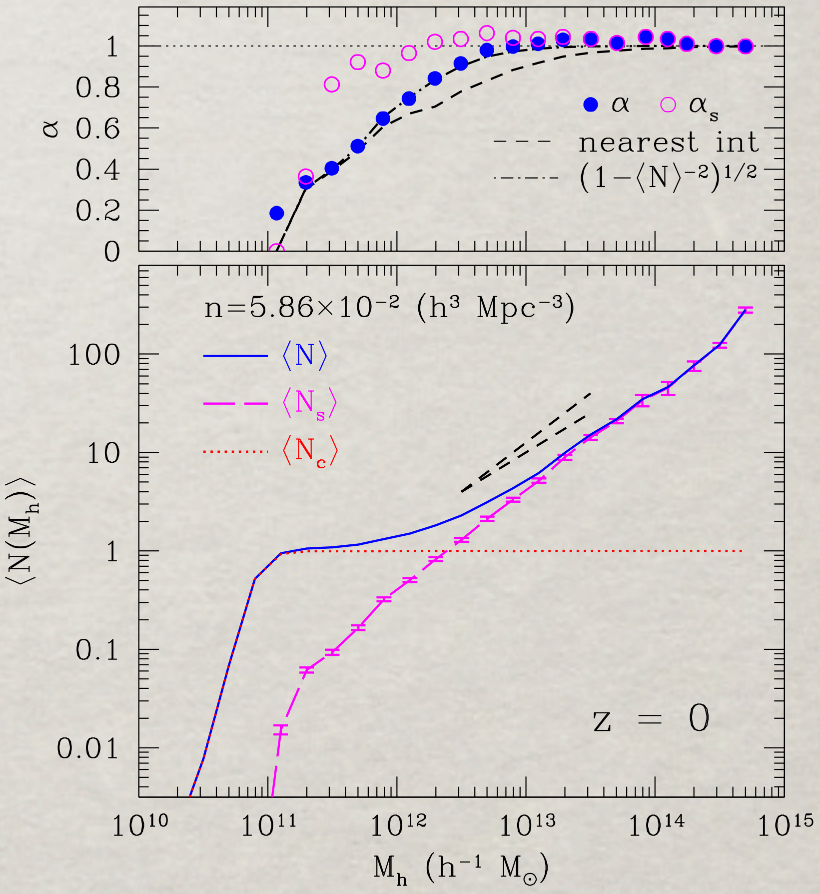
Berlind and Weinberg 2002



- “Semi-analytic” models and SPH models give a remarkably similar form for the mean number of galaxies as a function of halo mass
- Assume that “assembly bias” and environment are not important effects (which seems to be true for the time being)

Towards a parametrisation of N(M)

- Separating galaxy populations into “central” (the most massive galaxy at the center of the halo) and “satellite” populations considerably simplifies the analysis
- Satellites are poisson, central nearest integer
- In massive haloes, most galaxies are satellite galaxies
- In less massive haloes, the central galaxy is dominant
- The “satellite fraction” is an important measurable quantity

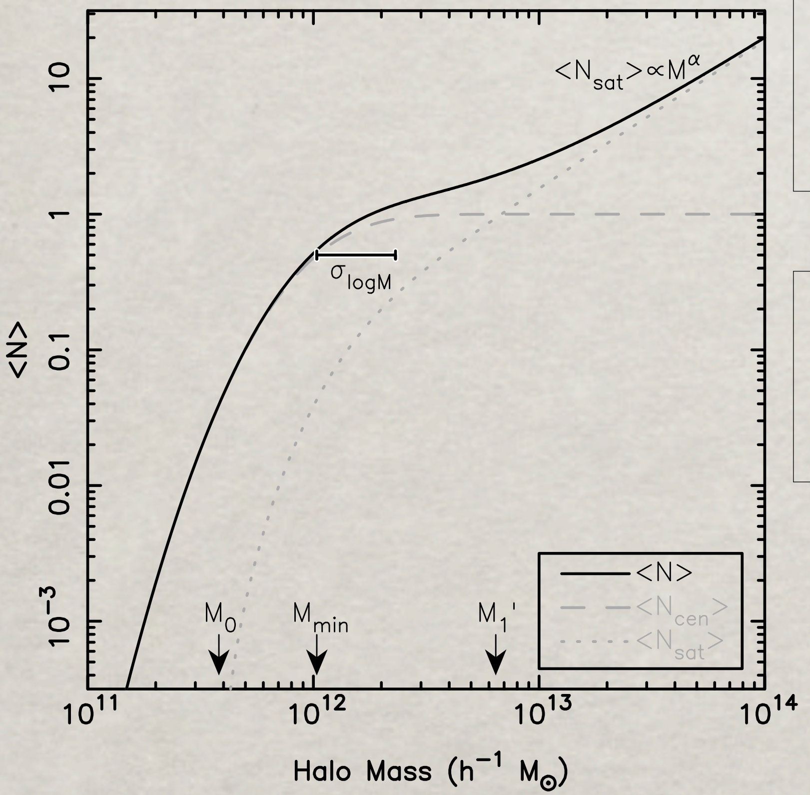


Kravstov et al 2004

The analytic halo occupation function

- If *P(N|M)* is specified by only the halo mass ... then we don't need to a do a full semi-analytic / SPH simulation to determine it!

Zheng et al 2005



Number of central galaxies per halo

$$\langle N_{\rm cen}(M) \rangle = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log M_{\rm min}}{\sigma_{\log M}}\right)\right], \qquad (8)$$

Number of satellites per halo

$$\langle N_{\rm sat}(M) \rangle = \langle N_{\rm cen}(M) \rangle \left(\frac{M - M_0}{M_1'}\right)^{\alpha} .$$

$$\langle N \mid M \rangle = \langle N_{\rm c} \mid M \rangle (1 + \langle N_{\rm s} \mid M \rangle)$$

A word on two-point correlation functions...

Two-point correlation functions give the excess probability for finding a neighbour a distance r from a given galaxy:

$$\delta P = n^2 \delta V_1 \delta V_2 [1 + \xi(r_{12})]$$

- In projected surveys, w(θ) is the simply the projection of xi(r) on the sky and depends (amongst other things) on the source redshift distribution

$$w(\theta) = \frac{DD - 2DR + RR}{RR}$$

- Measurement of w is simple -- just count the number of pairs as a function of angular scale between data catalogues *D* and random catalogues *R*

Deriving the galaxy correlation function

- Galaxy clustering statistics measure the number of pairs in excess of a random distribution
- In the “halo model” we suppose that the pair counts come from galaxies **inside the same halo** and **galaxies in separate haloes**
- The relative importance of each term depends on the angular scale and the size of the haloes
- In galaxy-galaxy lensing, we remove any additional uncertainty on the halo profiles, but only works at lower redshifts

“Halo models” are not meant as a replacement for traditional models of galaxy formation but are simply a technique to extract additional information from the observations

The one and two-halo terms

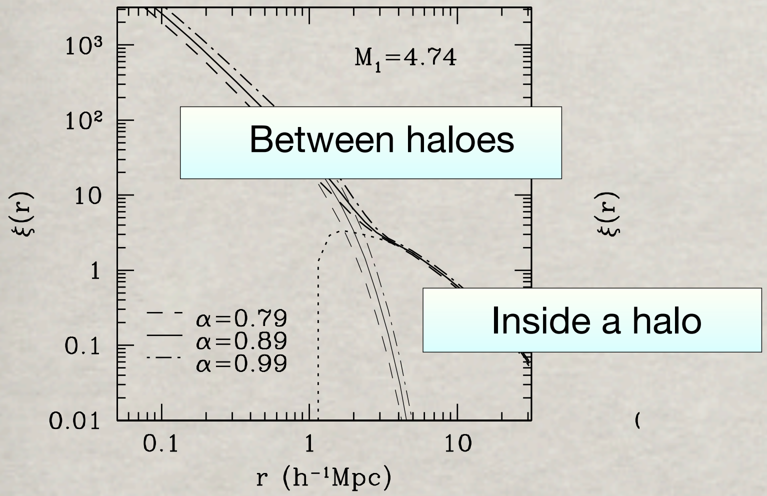
halo profile

$$1+\xi_{gg}^{1h}(r)=\frac{1}{2}\bar{n}_g^{-2}\int n(M)\langle(N(N-1)\rangle_M\lambda(r|M)dM$$

halo mass function

$$\xi_{gg}^{2h}=\xi_{dm}^{lin}(r)\bar{n}^{-2}\int n(M_1)b_h(M_1)\langle N\rangle_{M_1}dM_1$$

$$\times\int n(M_2)b_h(M_2)\langle N\rangle_{M_2}\lambda(r|M_1M_2)dM_2$$



Between haloes

Inside a halo

- At large scales this becomes the simple linear bias term
- 1h term = central-sat term + ss term
- “The transition scale” between one-halo and two-halo terms is sensitive to the effects of “halo exclusion” and it’s important to model this accurately because most of the signal in galaxy surveys is on these scales

How to compute the expected galaxy clustering using the halo model

Three ingredients are necessary to generate a prediction of w, the projected two-point correlation function:

1. **An accurate representation of the non-linear power spectrum of dark matter, the density profile of dark matter haloes and the number of dark matter haloes as a function of halo mass**
2. **A prescription (“guess”) for how the numbers of galaxies and pairs of galaxies which inhabit each dark matter halo depend on the halo mass.**
3. **Knowledge of the redshift selection function for each sample.**

By matching the **observed clustering of galaxies** with predictions of this model one can derive:

The typical mass of the underlying dark matter haloes

The fraction of galaxies which are satellites

The average bias (how much dark and luminous matter there is)

Parameter estimation with PMC

- Our *highly optimised* halo model code can compute *w* for a five-parameter halo model in less than 1s

$$\chi^2 = \sum_{ij}^{n} \left[(w^{\mathrm{obs}}(\theta_i) - w^{\mathrm{m}}(\theta_i))(C_{ij})^{-1}(w^{\mathrm{obs}}(\theta_j) - w^{\mathrm{m}}(\theta_j)) \right] + \frac{[\log n_{\mathrm{g}}^{\mathrm{obs}} - \log n_{\mathrm{g}}^{\mathrm{model}}]^2}{\sigma^2_{\log n_{\mathrm{g}}}},$$

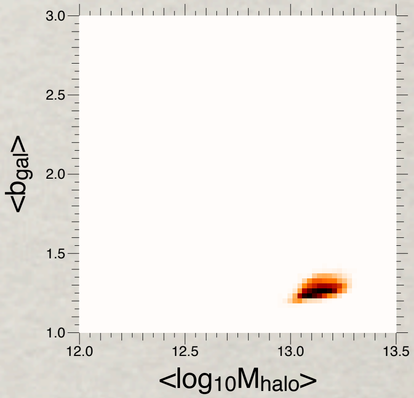
- **Models must reproduce the observed number density of galaxies; this a very strong constraint**
- Our covariance matrix C is derived using jack-knife resampling techniques
- We use the parallel **“population monte carlo**” (Wraith et al, Benabed, Kilbinger et al) technique to carry out a complete sampling of parameter space and derive realistic errors on fitted and derived parameters. On a cluster like magique2 we can sample 200,000 points in ~ 30 mins.

Derived parameters

Once we have found a best fitting halo model we can derive these additional parameters by marginalisation:

$$b_g(>M_{\rm min}) \equiv \frac{\int {\rm d}M\, n_{\rm halo}(M) N_g(M) b(M)}{\int {\rm d}M\, n_{\rm halo}(M) N_g(M)}.$$

Average halo bias



$$\langle M_{\rm host} \rangle = \frac{\int_{M_{\rm min}}^{\infty} {\rm d}M\, M\, N_g(M) n_{\rm halo}(M)}{\int_{M_{\rm min}}^{\infty} {\rm d}M\, N_g(M) n_{\rm halo}(M)},$$

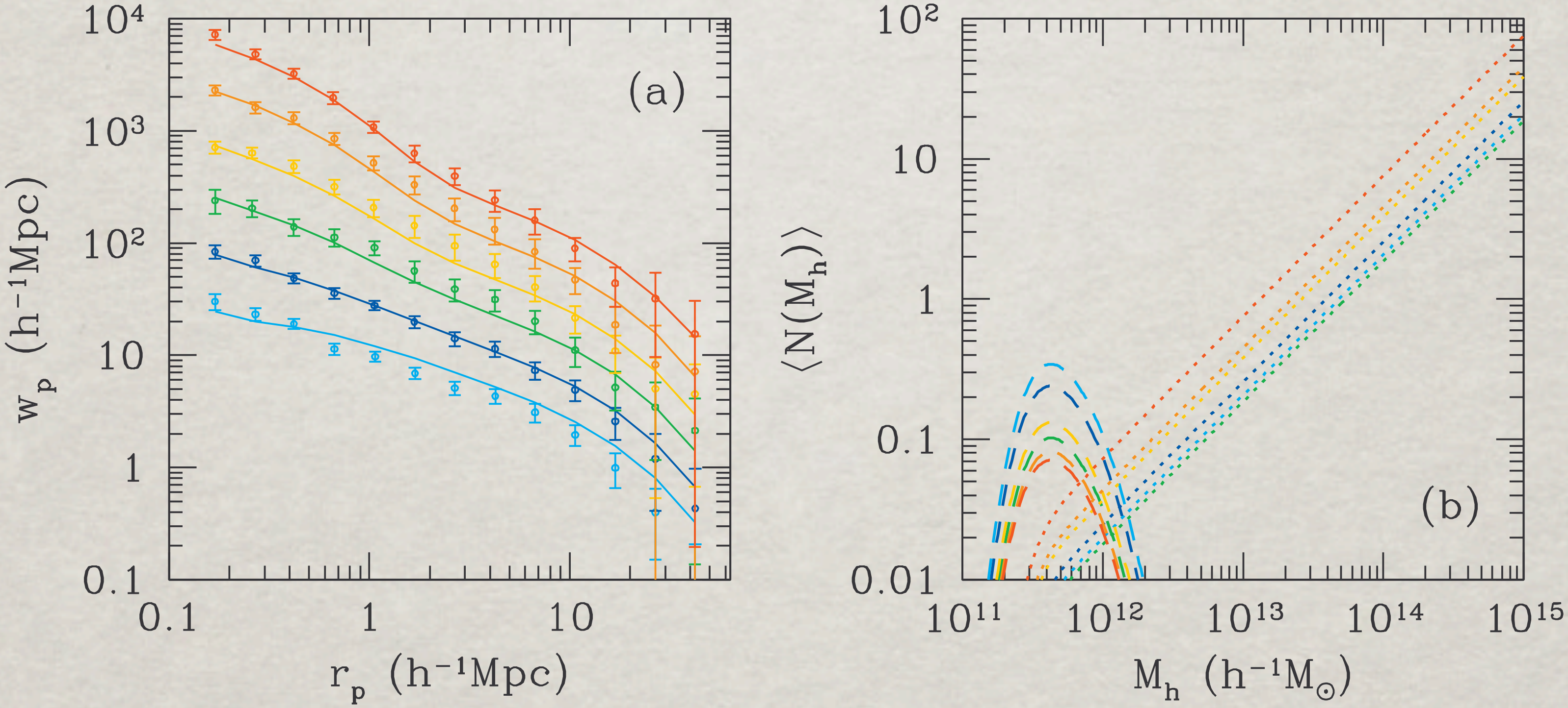
Average halo mass

$$\langle N_{\rm g} \rangle \equiv \frac{\int_{M_{\rm min}}^{\infty} N_{\rm g}(M) n_{\rm halo}(M,z)\, {\rm d}M}{\int_{M_{\rm min}}^{\infty} n_{\rm halo}(M,z)\, {\rm d}M}.$$

Average number of galaxies per halo

What the halo parameters actually mean

Zehavi et. al 2010, SDSS clustering



- Fixed alpha, M_min; only M1 allowed to vary
- More luminous objects have more pronounced one-halo term and large fraction of satellites

What kind of surveys do we need?

Low redshift surveys can now tell us a lot about the local universe.

But how do the galaxies seen in high redshift surveys evolve into the present day populations?

Red sequence / blue cloud

How does the properties of the hosting dark matter halo depend on luminosity and colour selection?

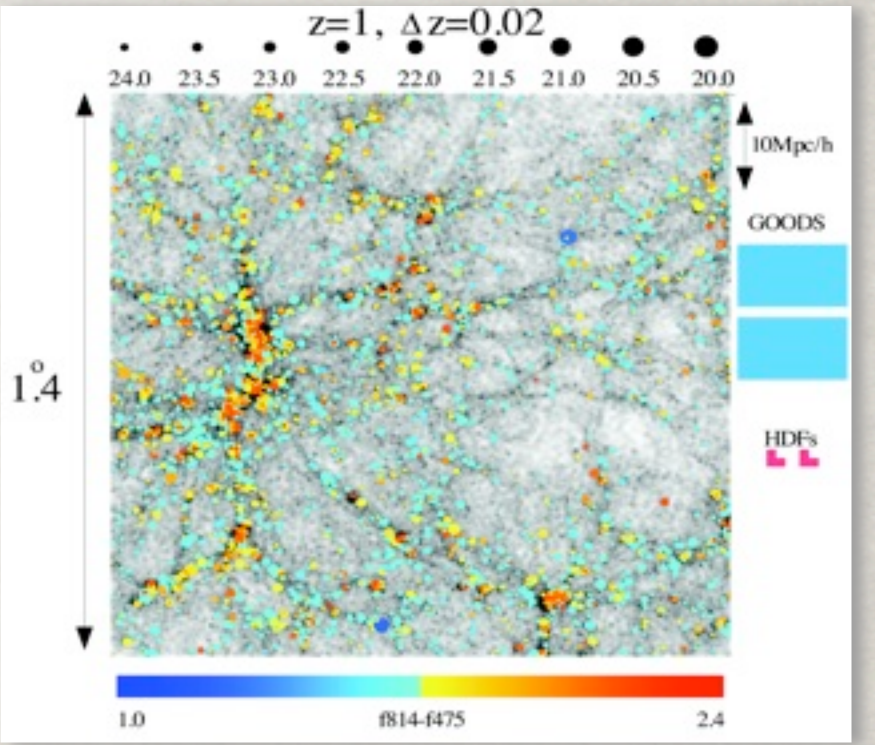
To answer this question we need surveys which can probe **a large enough range in densities at scales** at 1-10 h^{-1} Mpc at z~1 to a depth of at least $0.1L_*$

This means a field size of *at least* 2 $\deg^2$

And enough filters to compute photometric redshifts to 5% accuracy.

90 Mpc

A slice of a simulated Universe at z~1



The CFHTLS and COSMOS surveys

- The CFHTLS and COSMOS surveys are two unique and complimentary probes of the distant Universe
- The COSMOS survey covers **~2 deg^2** with very deep, multi-colour data (almost all bands and wavelengths); most precise photometric redshifts available
- CFHTLS covers **~130 deg^2** in *ugriz* : a unique probe of the Universe at z~1; over one million galaxies; CFHTLS can also access higher mass haloes. We can probe over a much larger range in redshift than SDSS
- COSMOS galaxies have stellar mass estimates and importantly can access the important $1<z<2$ redshift range.

Photometric redshifts: a cheap way to get galaxy distances (Ilbert et al. 2006,8 and Coupon et al. 08)

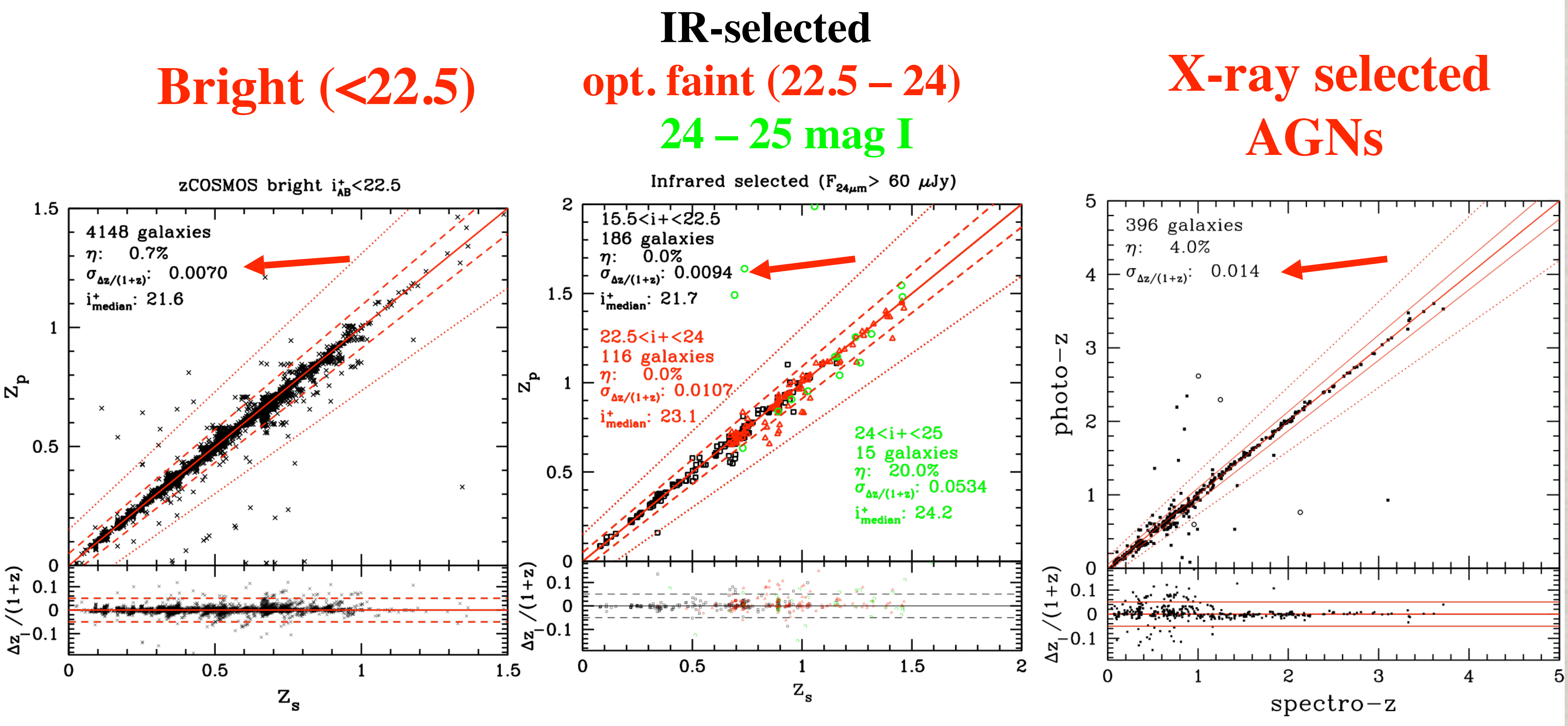
$$\chi^2(z,T,A) = \sum_{f=1}^{N_f} \left(\frac{F^f_{\mathrm{obs}} - A \times F^f_{\mathrm{pred}}(z,T)}{\sigma^f_{\mathrm{obs}}} \right)^2 ,$$

“Photometric redshifts” are computed by comparing observed spectral energy distributions with a set of template SEDs.

For many years the accuracy of photometric redshifts was difficult to assess because of the lack of large (>10k objects) spectroscopic training sets (it turned out that a lot of photo-zeds computed without training sets were **actually wrong**!)

Wide-field cameras with precise photometric calibration (like **MegaCam**) combined with wide field spectrographs producing training sets of ~10k galaxies (like **VIMOS**) makes estimating photometric redshifts for millions of galaxies with percent-level accuracy possible

COSMOS 30-band phot-zeds (ilbert et al 08)


IR-selected

Bright (<22.5)

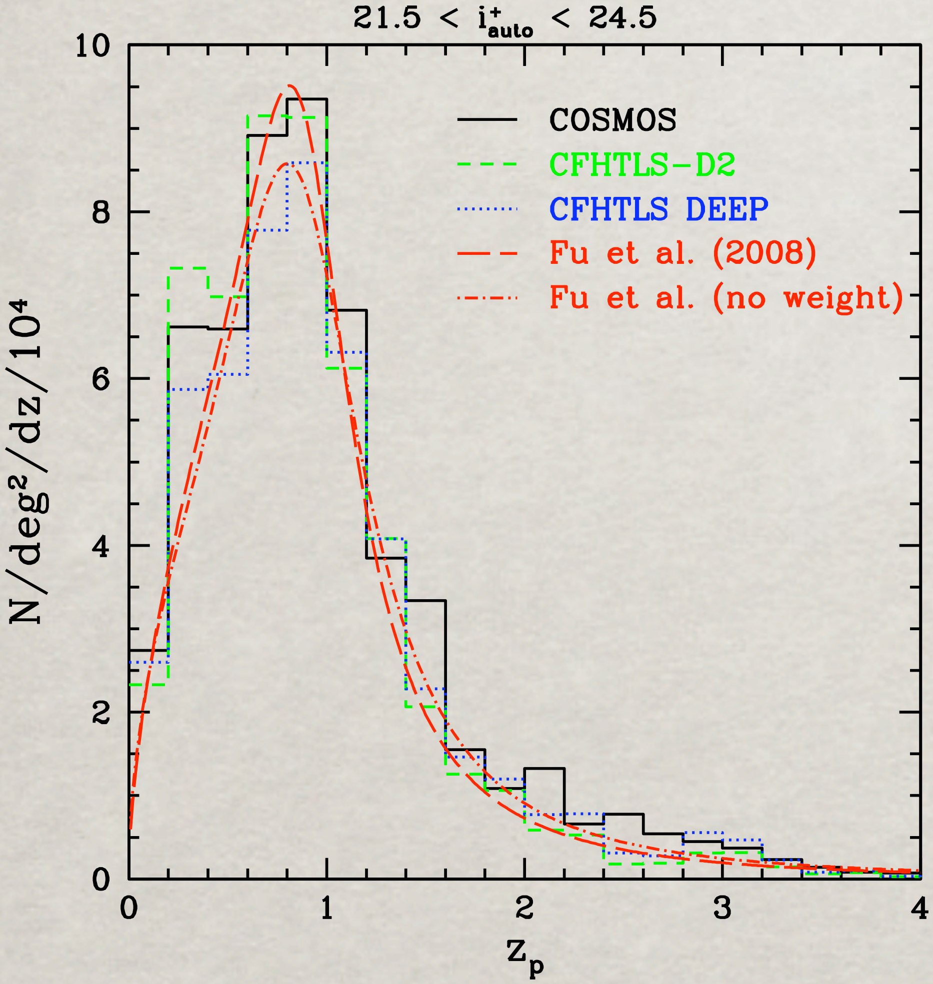
opt. faint (22.5 – 24)

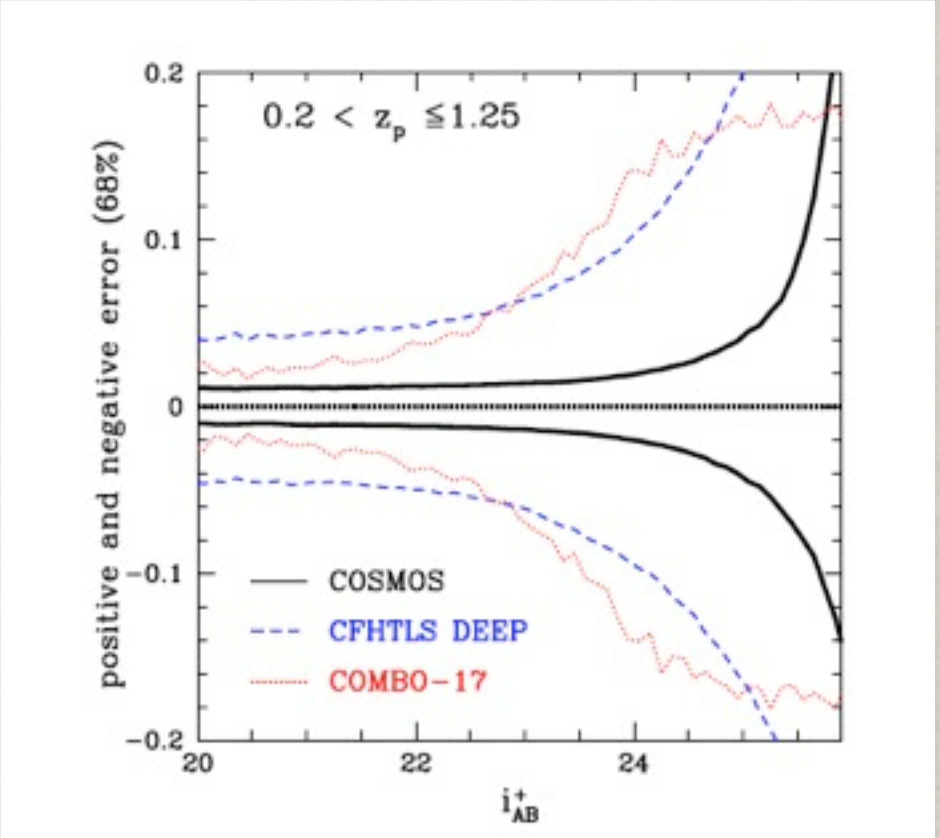
24 – 25 mag I

X-ray selected AGNs

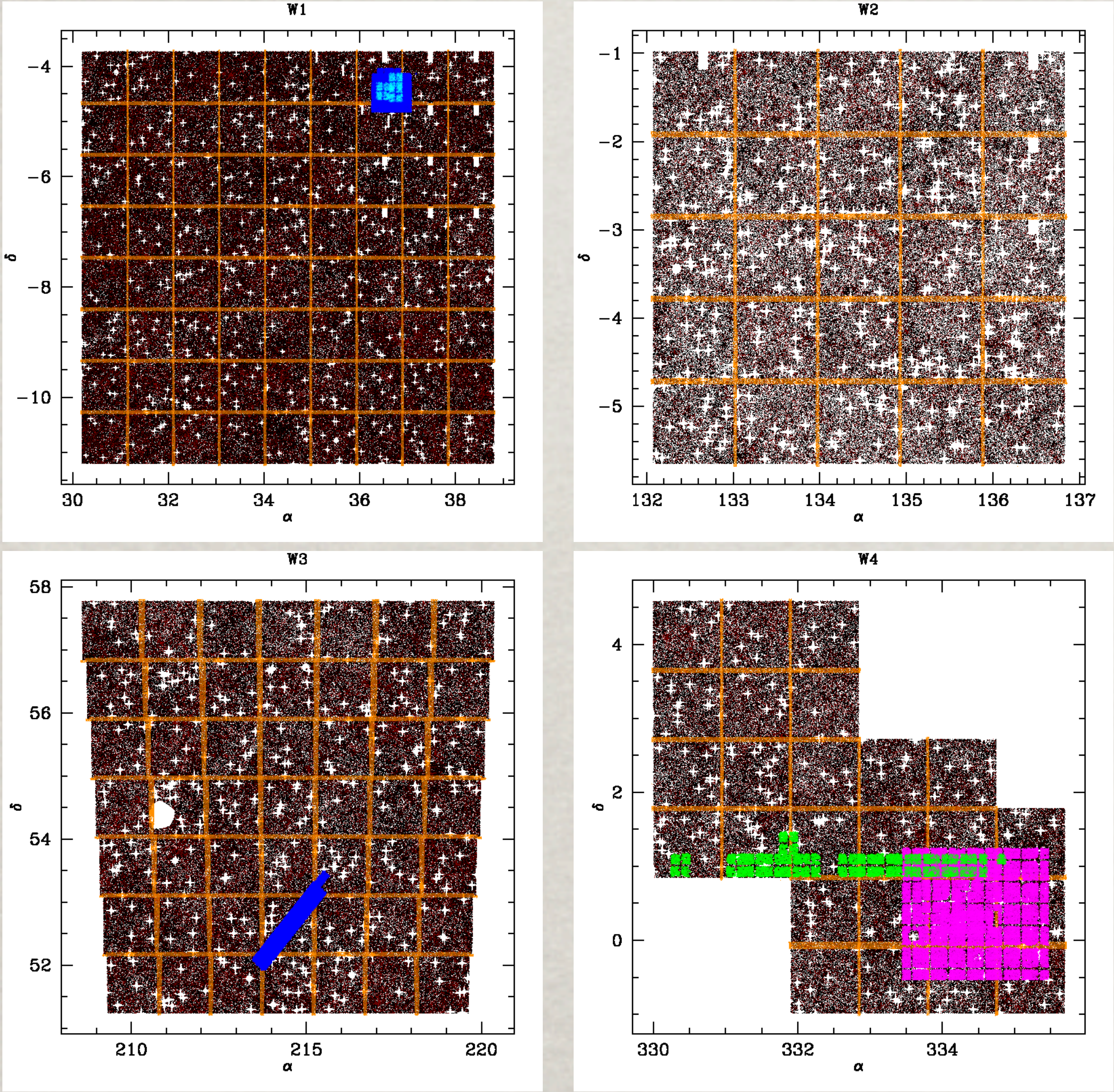
- ~0.7%-1% accuracy photometric redshifts over 0.2<z<1.2
- Large spectroscopic training sample: VLT/zCOSMOS (Lilly); also smaller samples from Magellan (Trump); and Keck (Capak)

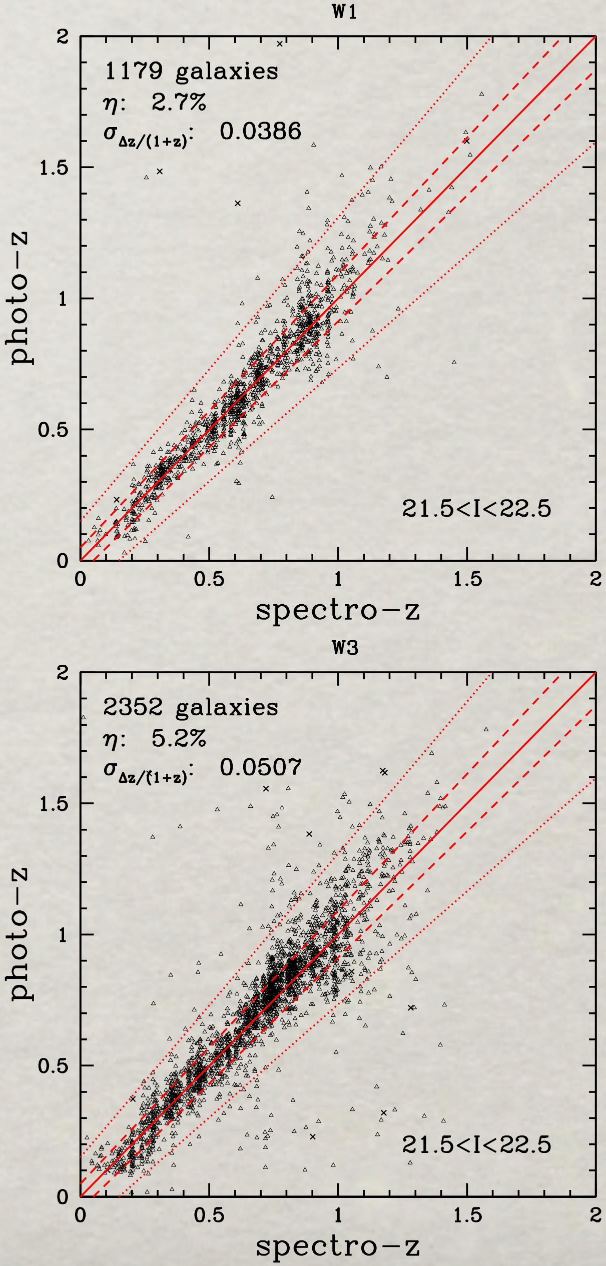
N(z) and photo-zed accuracy in COSMOS





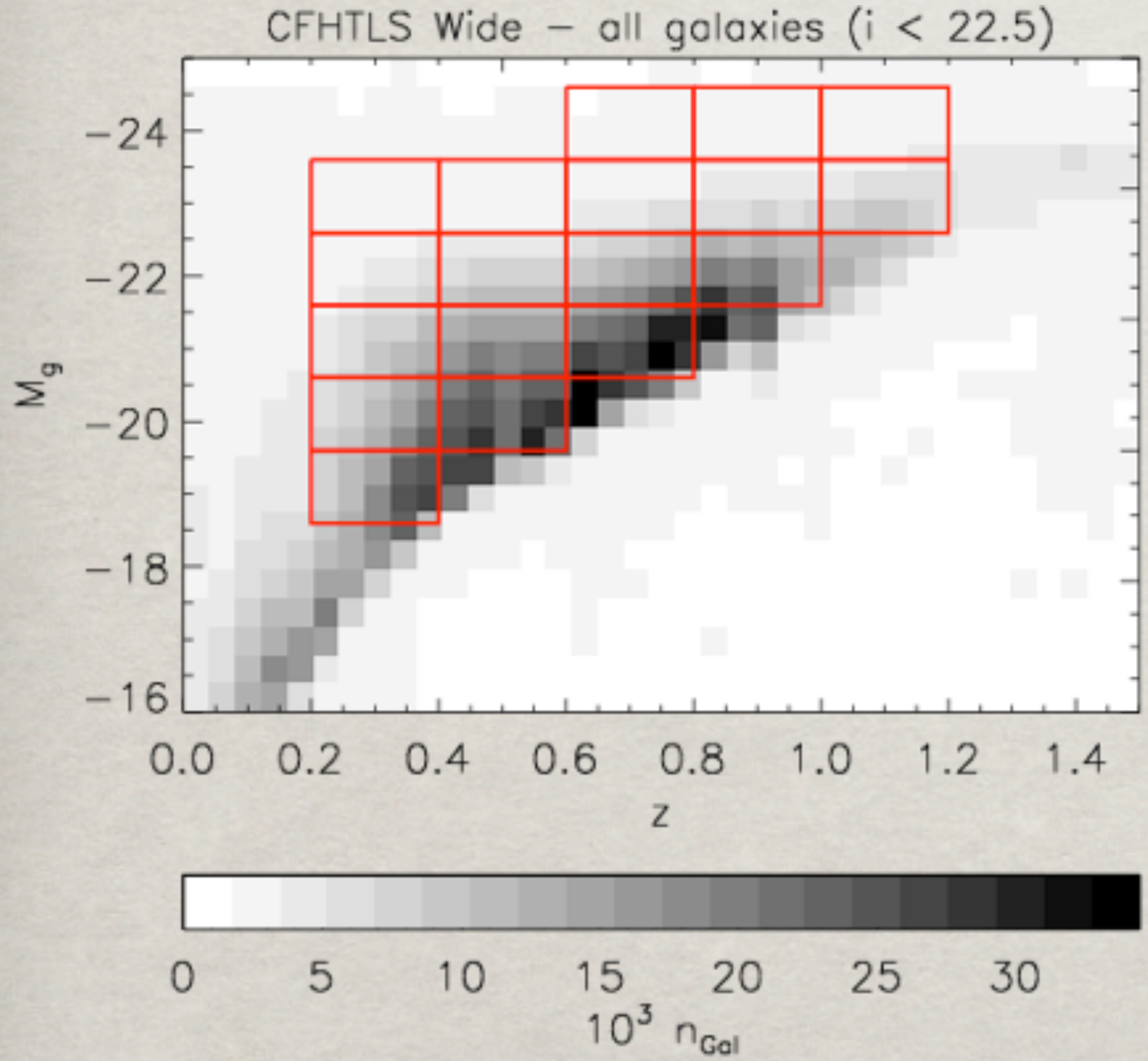
The CFHTLS-WIDE

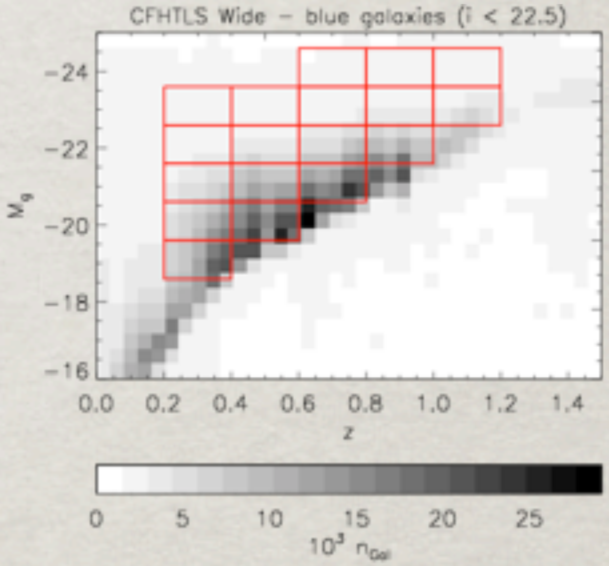


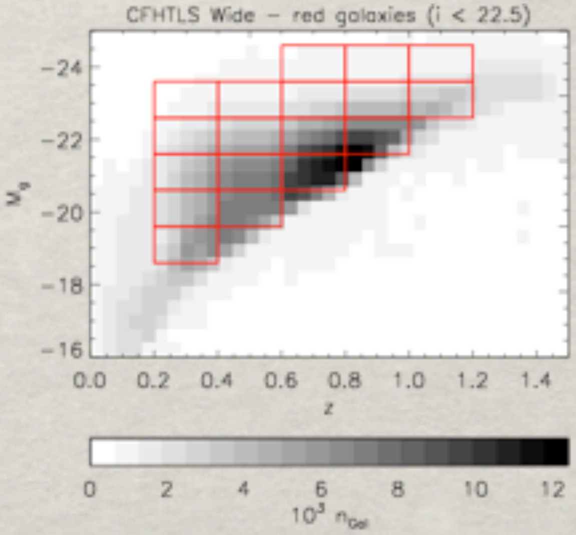


Coupon et al in prep

The CFHTLS sample selection



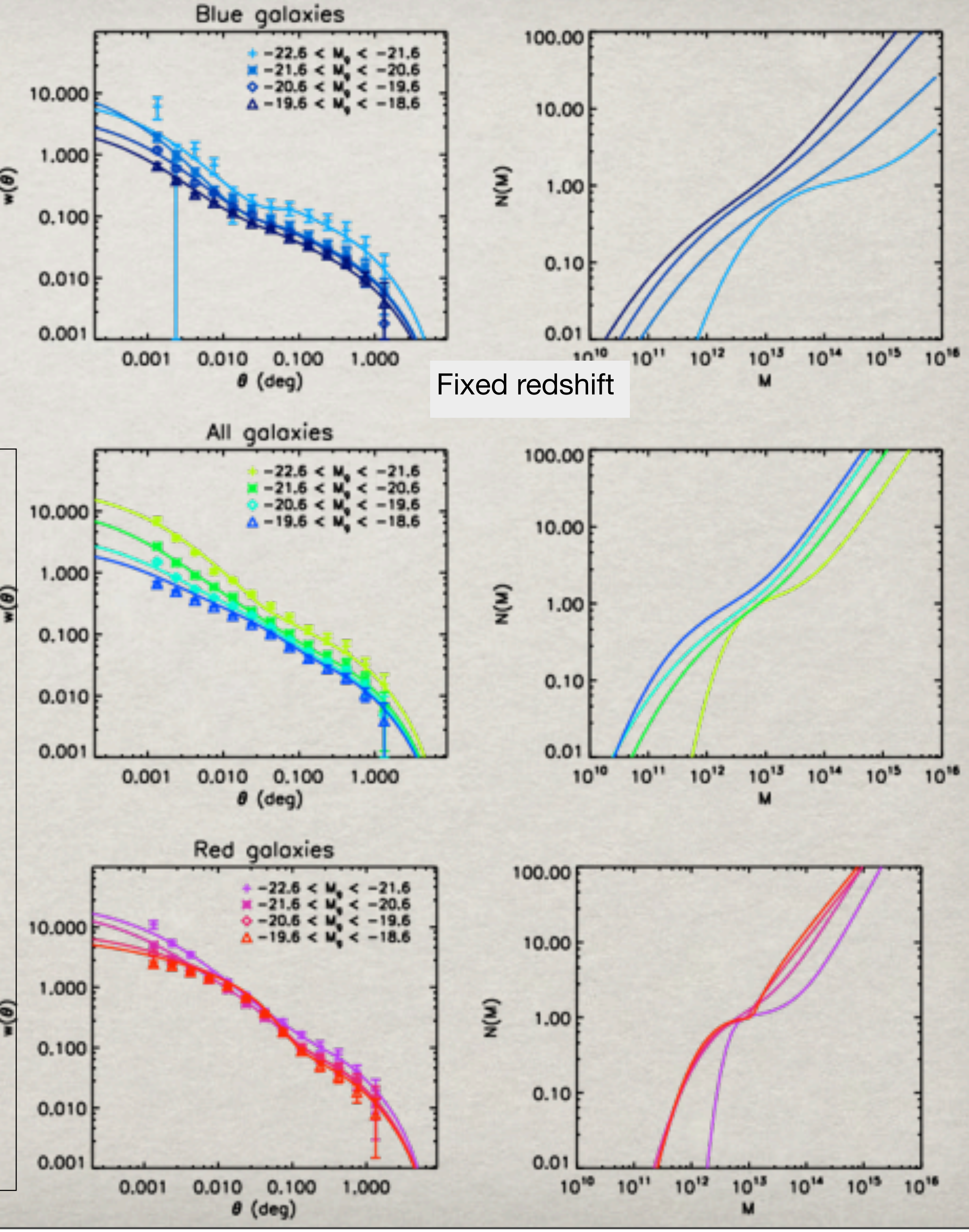




Coupon et al in prep

Colour-selected galaxies in the CFHTLS-wide

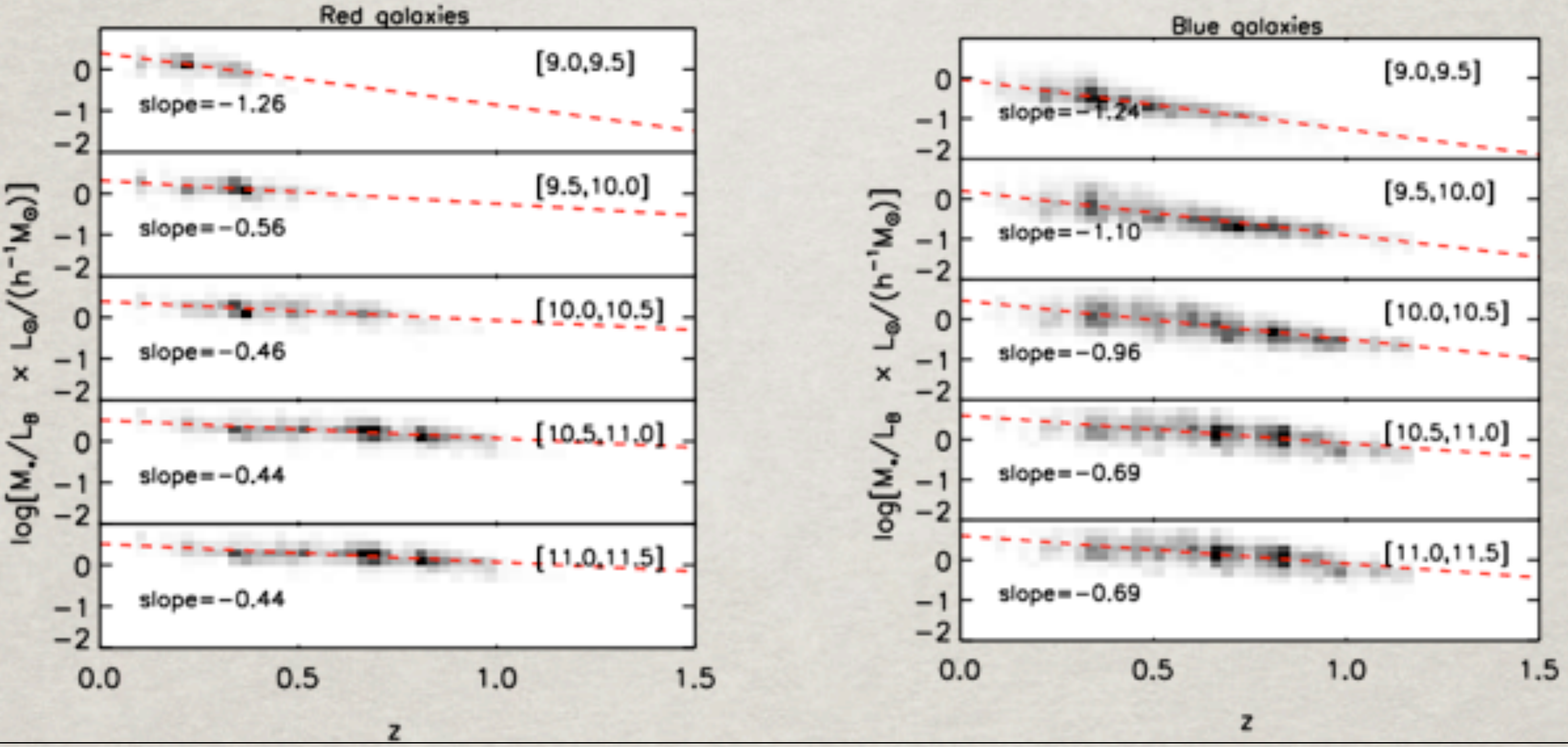
- Bluer galaxy samples are dominated by central galaxies
- Redder galaxies have are dominated by satellite populations
- More luminous galaxies are hosted in more massive haloes



Fixed redshift

Coupon et al in prep

Making mass-limited samples

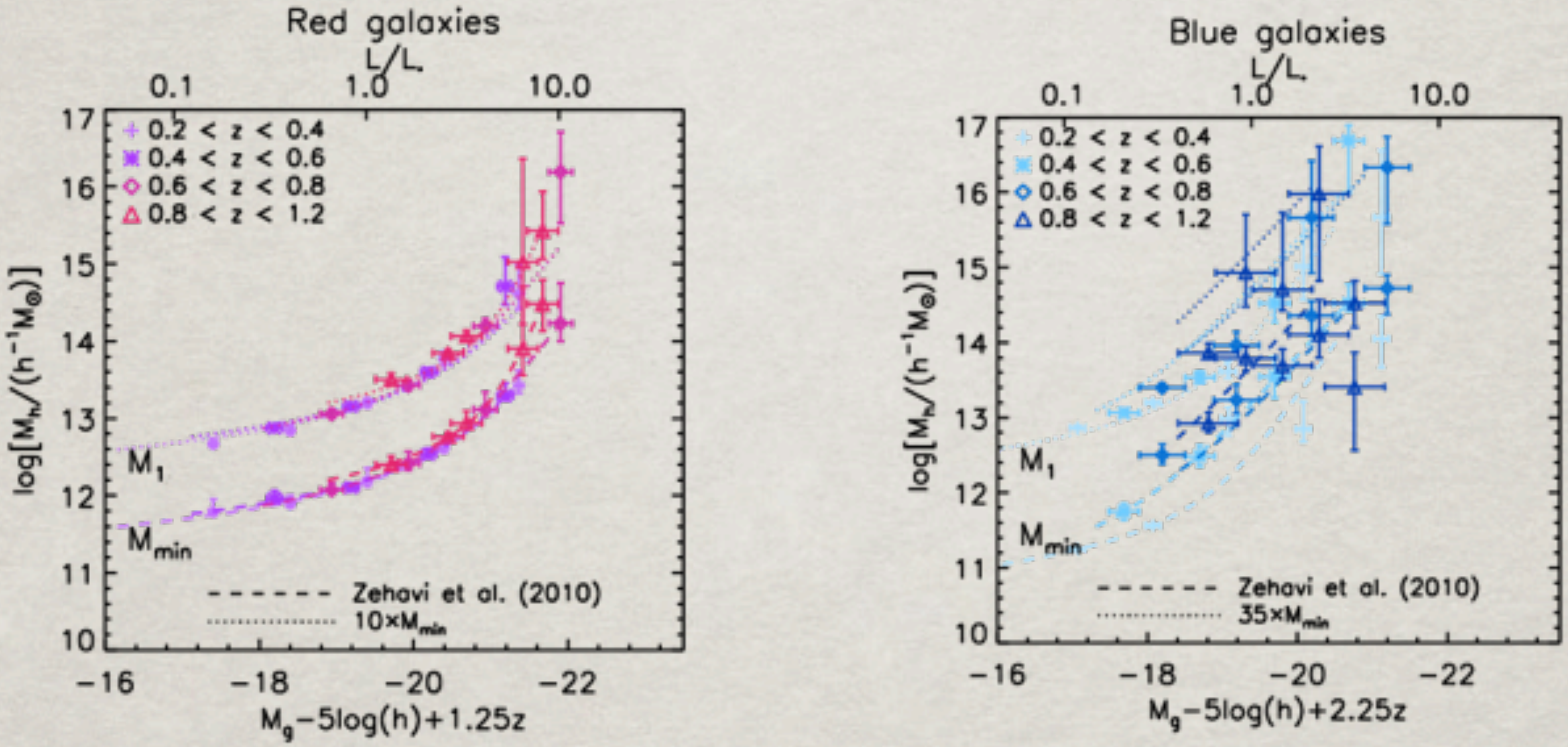


As we only have five optical bands we cannot reliably compute stellar masses, but we can try to derive an empirical relationship between mass and B-band luminosity

As expected, this works reasonably well for red populations, but for “blue” and “full” populations the slope of the redshift-M/l relationship depends strongly on mass; for these populations we use an intermediate relationship.

Coupon et al in prep

Observed halo mass scales



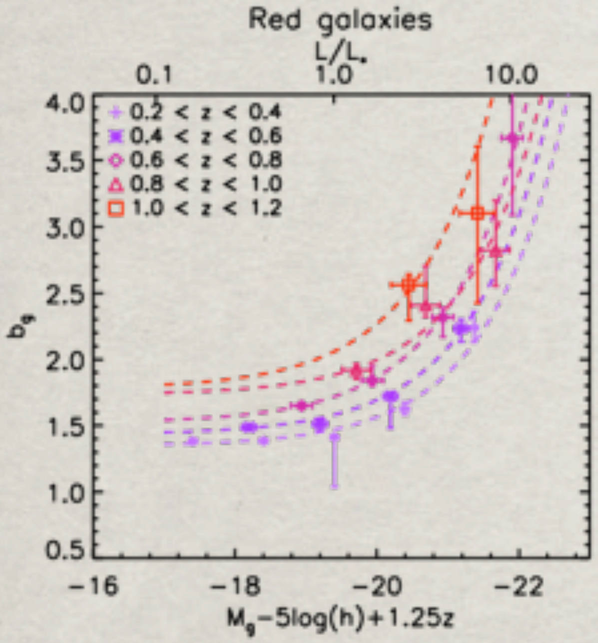
M_min and M1 corresponds to the halo masses required to host one and two central galaxies respectively

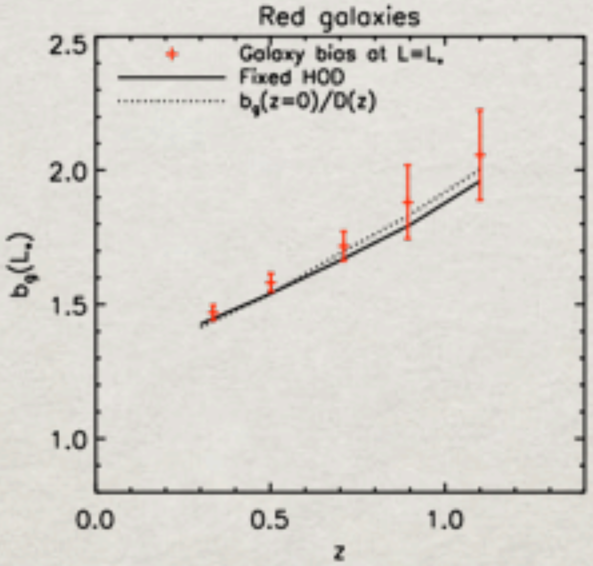
We use these empirical corrections to transform our samples into approximately mass-limited ones. We fit M_min and M_1 as a function of redshift and luminosity

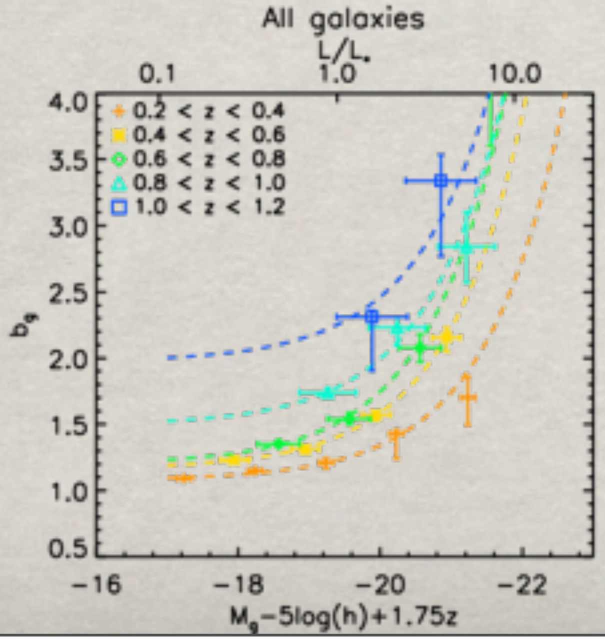
At each redshift, more luminous / massive galaxies are hosted by more massive haloes

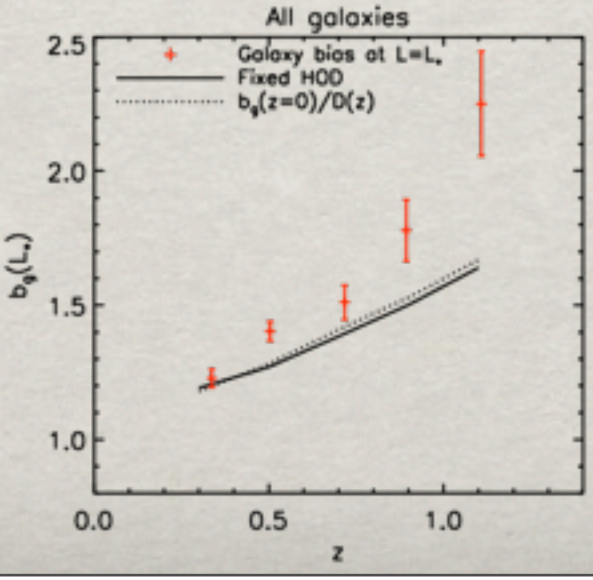
Coupon et al in prep

Evolution of halo bias with redshift









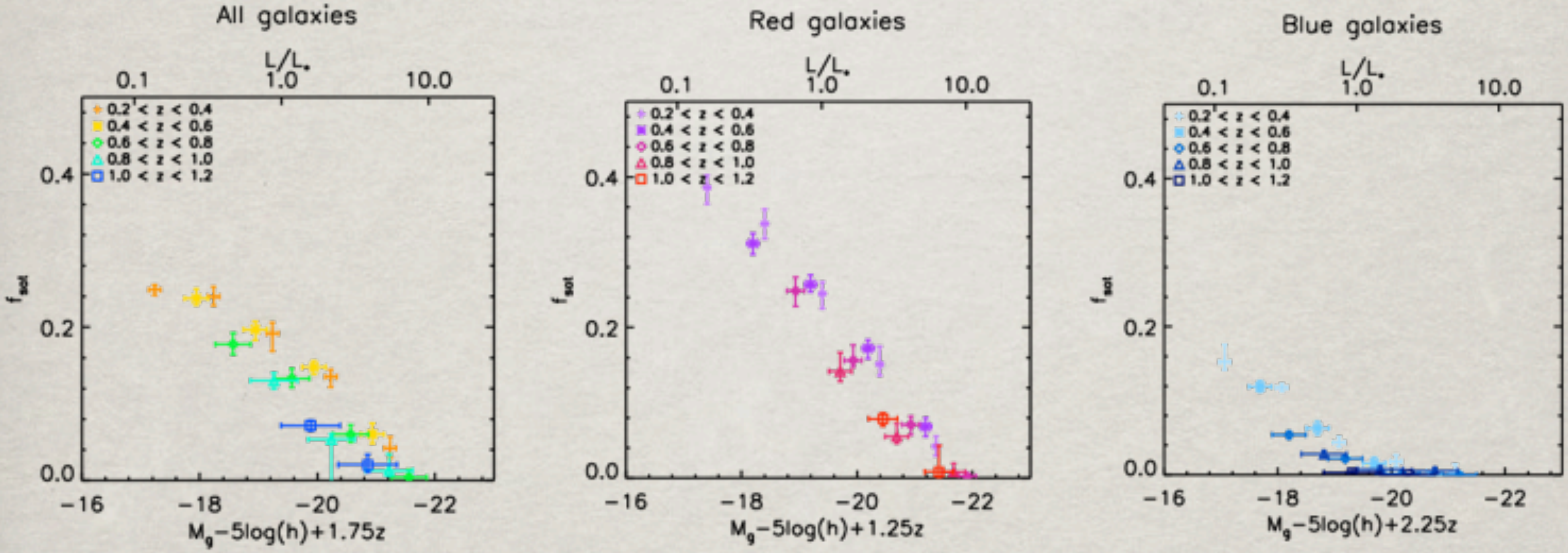
Red galaxies follow closely the “dm only” evolution.

Interpretation of these results is complicated by changing red/blue luminosity functions

More luminous objects are expected to have a different bias evolution

Coupon et al in prep

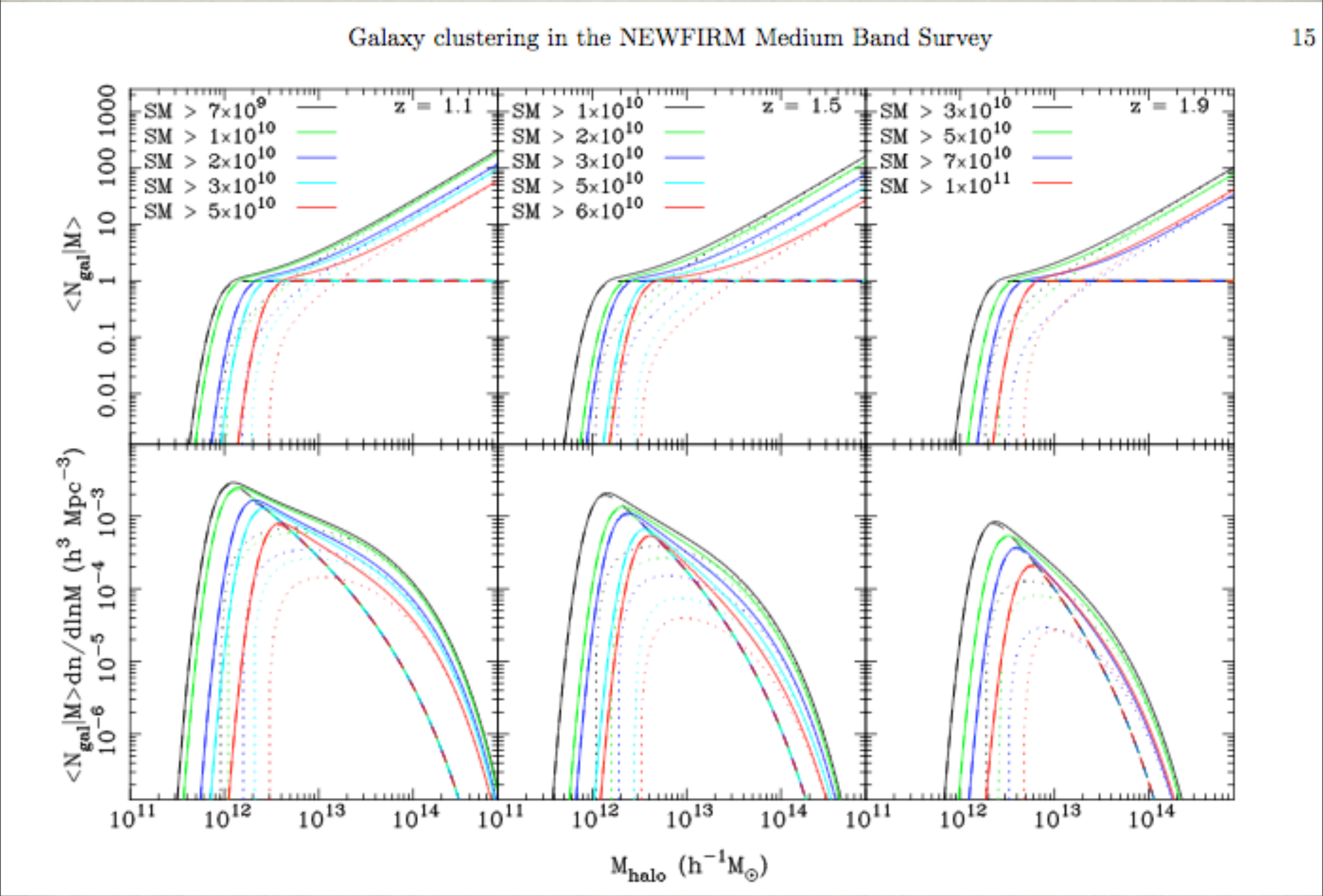
Satellite fractions



From our best-fitting halo model parameters we can compute the fraction of galaxies which are “satellites” and which are “centrals”

The satellite fraction is essentially set by a combination of the halo mass function and the halo occupation distribution N(M); halo mass function drops rapidly at high and low masses; N(M) for blue galaxies dominated by central galaxies

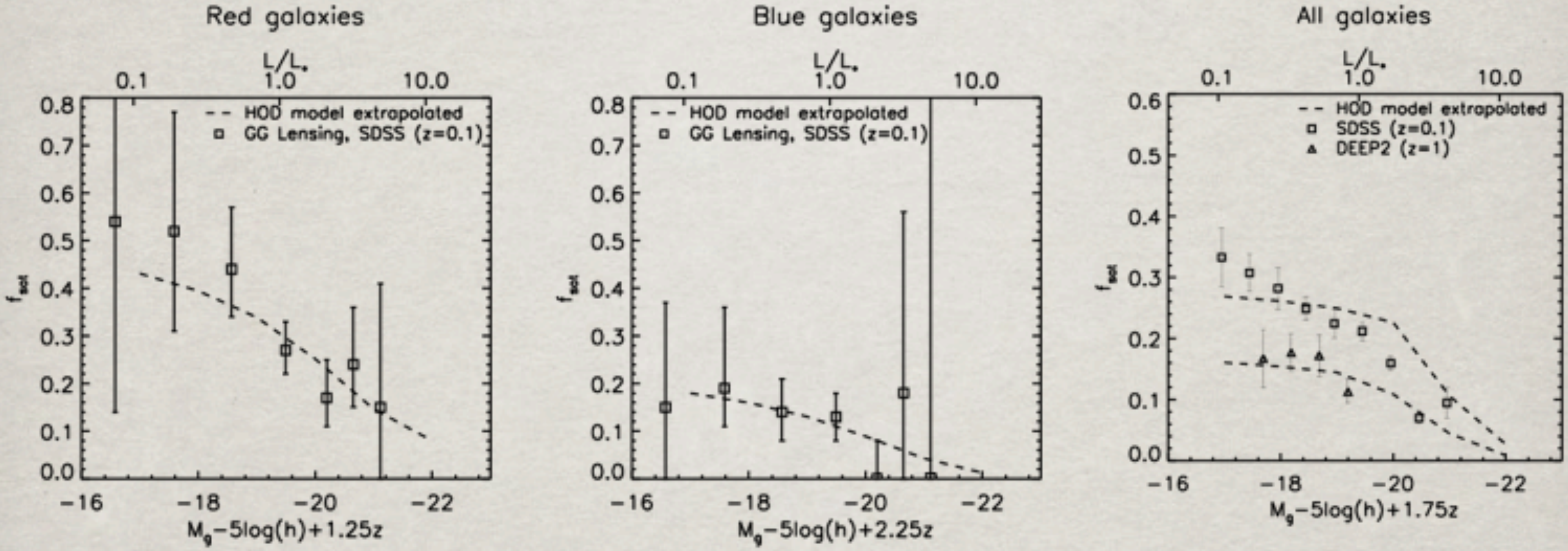
Can satellite fraction measurements provide useful information about galaxy evolution?



SM > 7×10^9
SM > 1×10^10
SM > 2×10^10
SM > 3×10^10
SM > 5×10^10
z = 1.1
SM > 1×10^10
SM > 2×10^10
SM > 3×10^10
SM > 5×10^10
SM > 6×10^10
z = 1.5
SM > 3×10^10
SM > 5×10^10
SM > 7×10^10
SM > 1×10^11
z = 1.9
<N_gal|M>
0.01
0.1
1
10
100
1000
<N_gal|M>dn/dlnM (h^3 Mpc^-3)
10^-6
10^-5
10^-4
10^-3
10^11
10^12
10^13
10^14
M_halo (h^-1 M_⊙)

Coupon et al in prep

Redshift evolution of the sat. fraction



We use our observations to fit a relationship between M1/M_min and M_g and then use our halo model to extrapolate our measurements to low redshift

Observations are consistent with this extrapolated halo model: this is consistent with the observation that mergers are not a dominant process at z<1 (otherwise we would over-predict the satellite fraction at z~0).

Agreement with the lensing results are reassuring.

Conclusions from CFHTLS

1. Our model fits observations for red galaxy samples relatively well, blue samples less well. “Full” samples contain a mix of blue and red galaxy populations which depend on redshift and luminosity

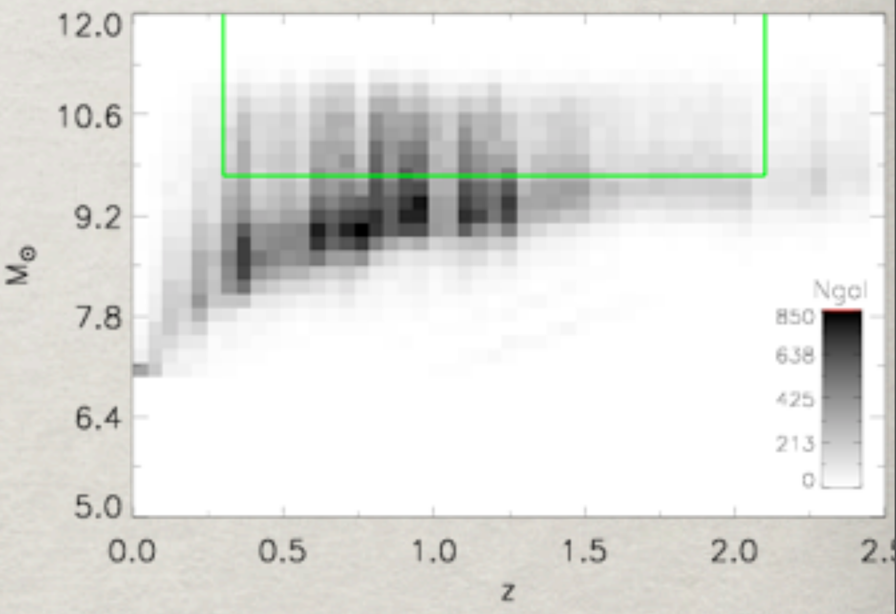
2. The observed evolution of galaxy properties with redshift at $z<1$ is (largely) consistent with dark matter evolution for L^* galaxies

3. Some hints for evolution for more massive objects

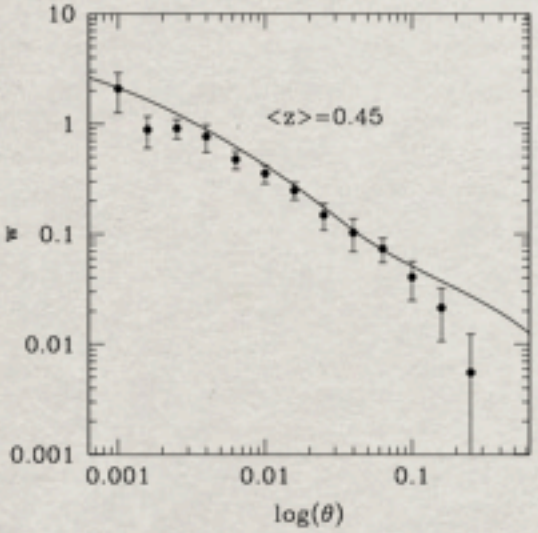
4. Interpretation of results are complicated by conversion between mass and luminosity

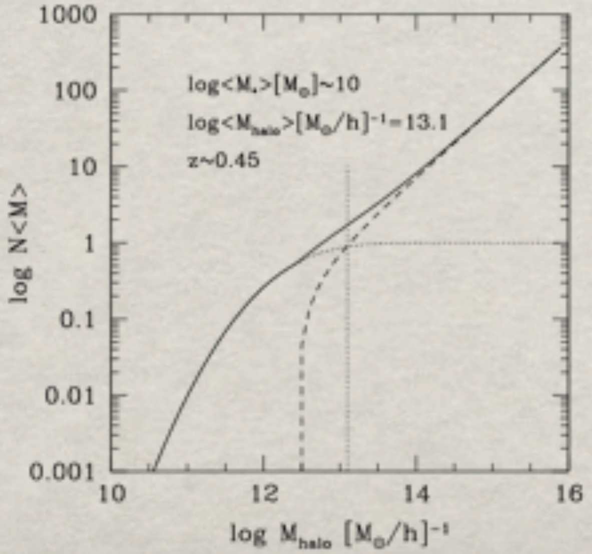
Mass-selected catalogues in COSMOS

- In the CFHTLS data, our luminosity selected samples can under considerable evolution over the redshift range of our sample: in g* it is ~1.5 magnitudes!
- In the COSMOS galaxies we can **select by stellar mass** which removes uncertainty in M/L conversion
- However, the size of the COSMOS field means number of high mass haloes are rare, which complicates interpretation

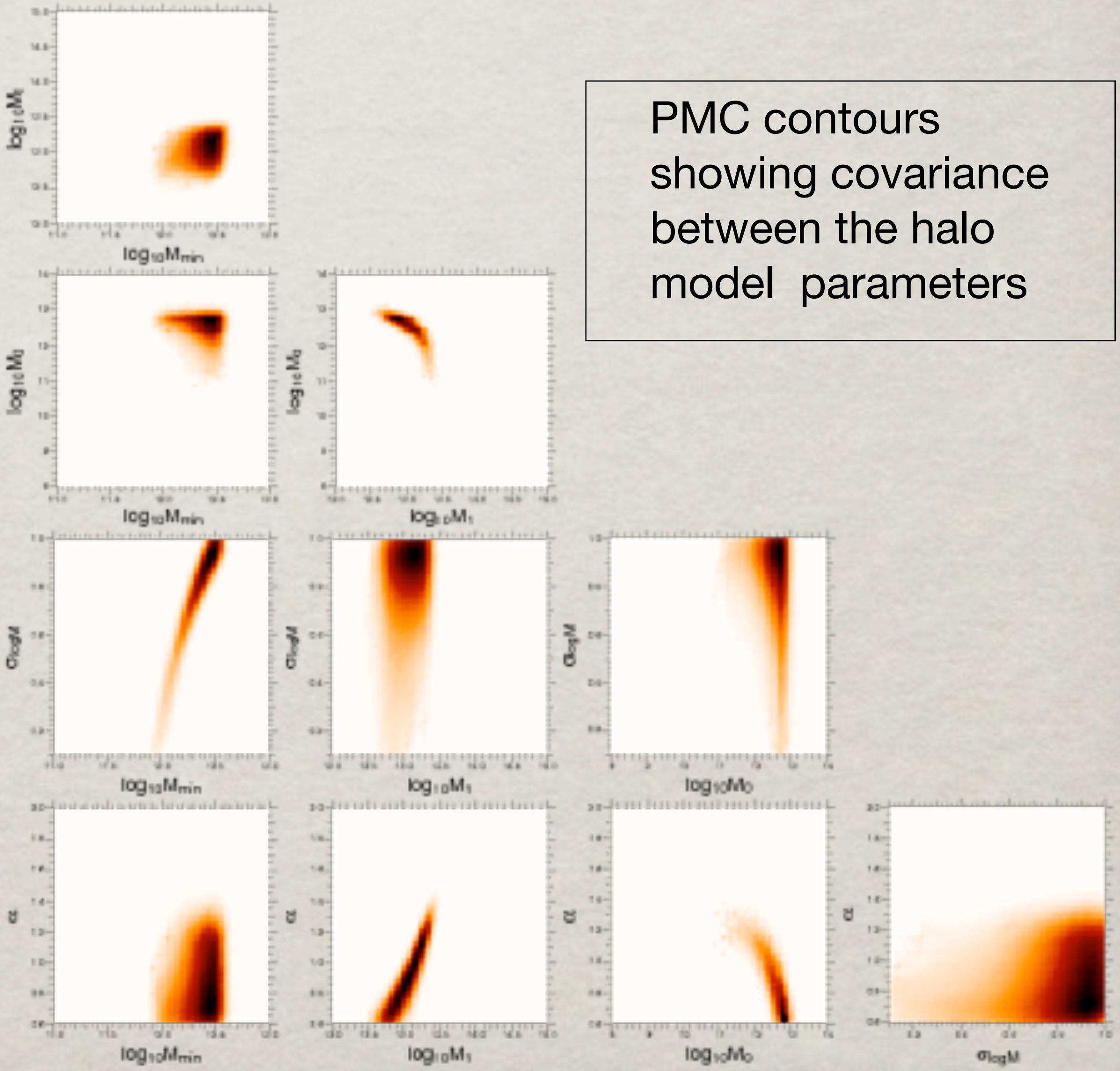


Stellar mass-selected catalogue at z~0.4





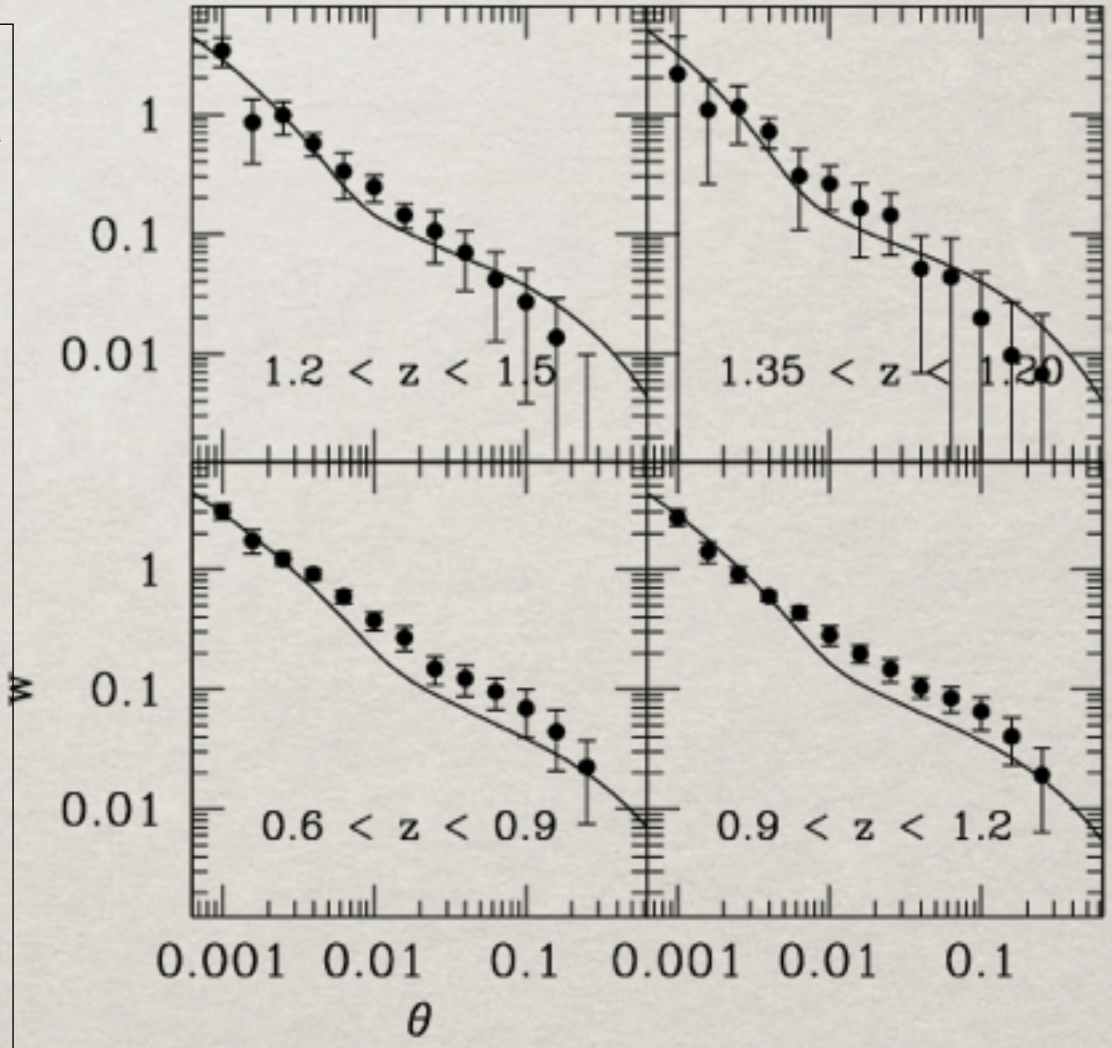
COSMOS_subaru_iab26_kab24_v7c_all_0.45



PMC contours showing covariance between the halo model parameters

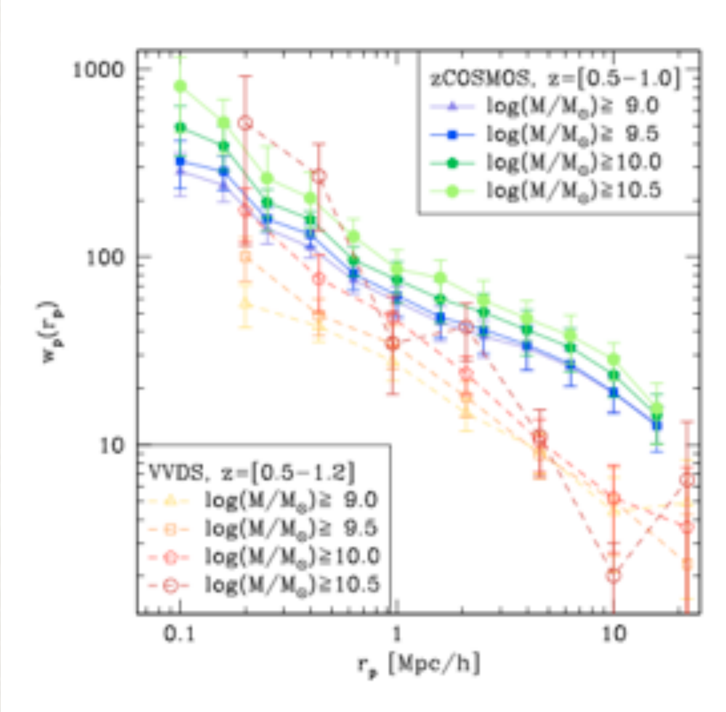
Fits for mass-selected samples

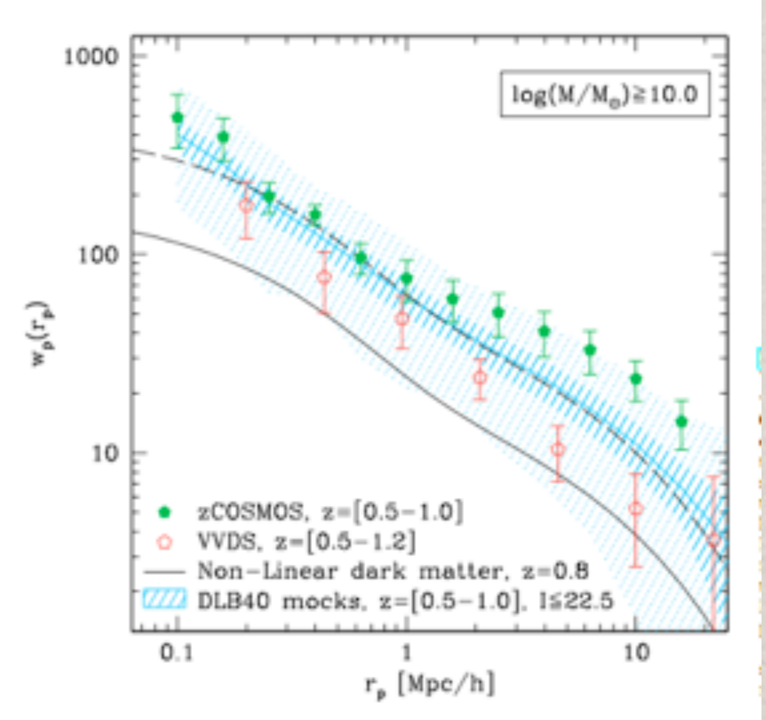
- The halo model cannot always reproduce perfectly the observed clustering signal, in particular on larger scales; in the COSMOS field this is because there is a large structure at z~0.8
- The transition between one-and two-halo terms becomes more pronounced at higher redshift



High clustering amplitudes at large scales in COSMOS

Meneux et al 2009

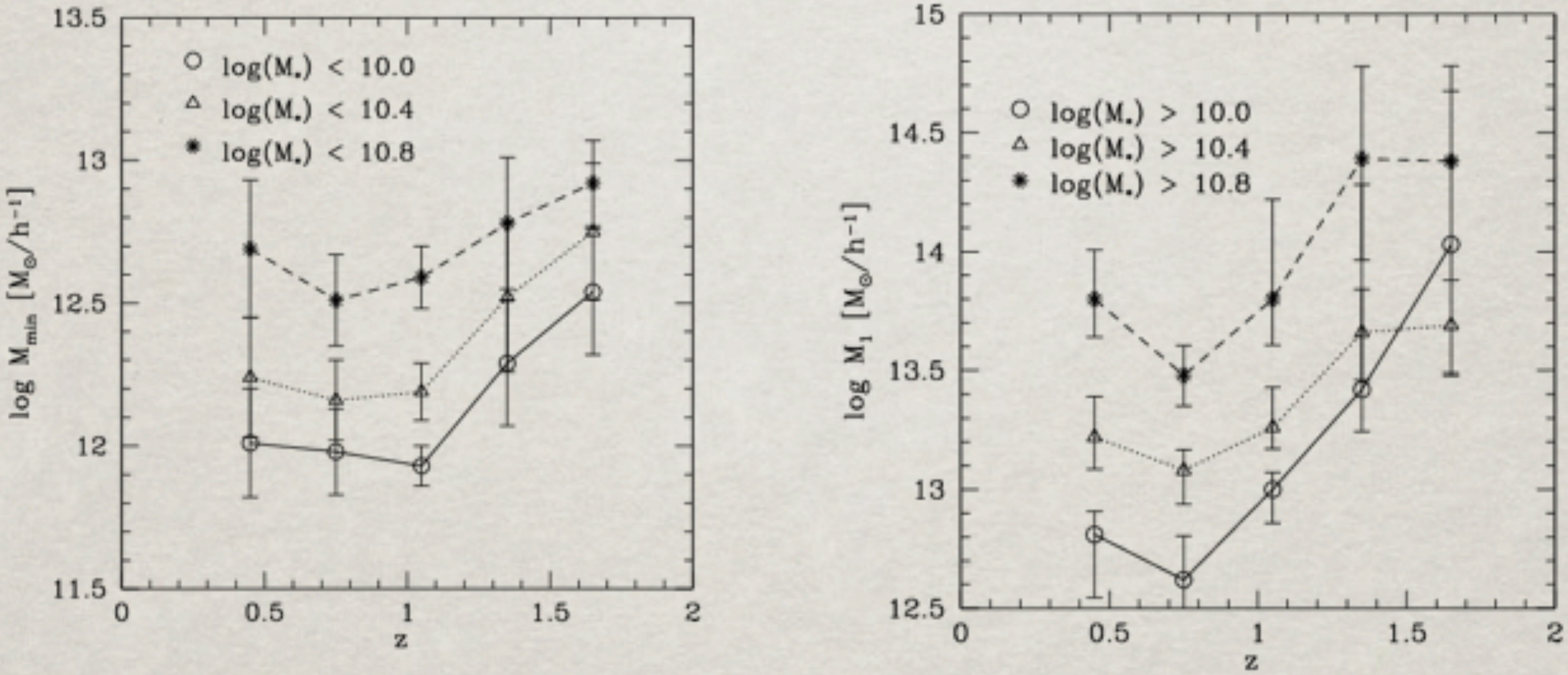




At z~1 at scales larger than 10 h{-1}Mpc, COSMOS has more power on large scales compared to other fields

This seems to be caused by the presence of rich structures in the field at z~1

Halo mass scale evolution at z~2

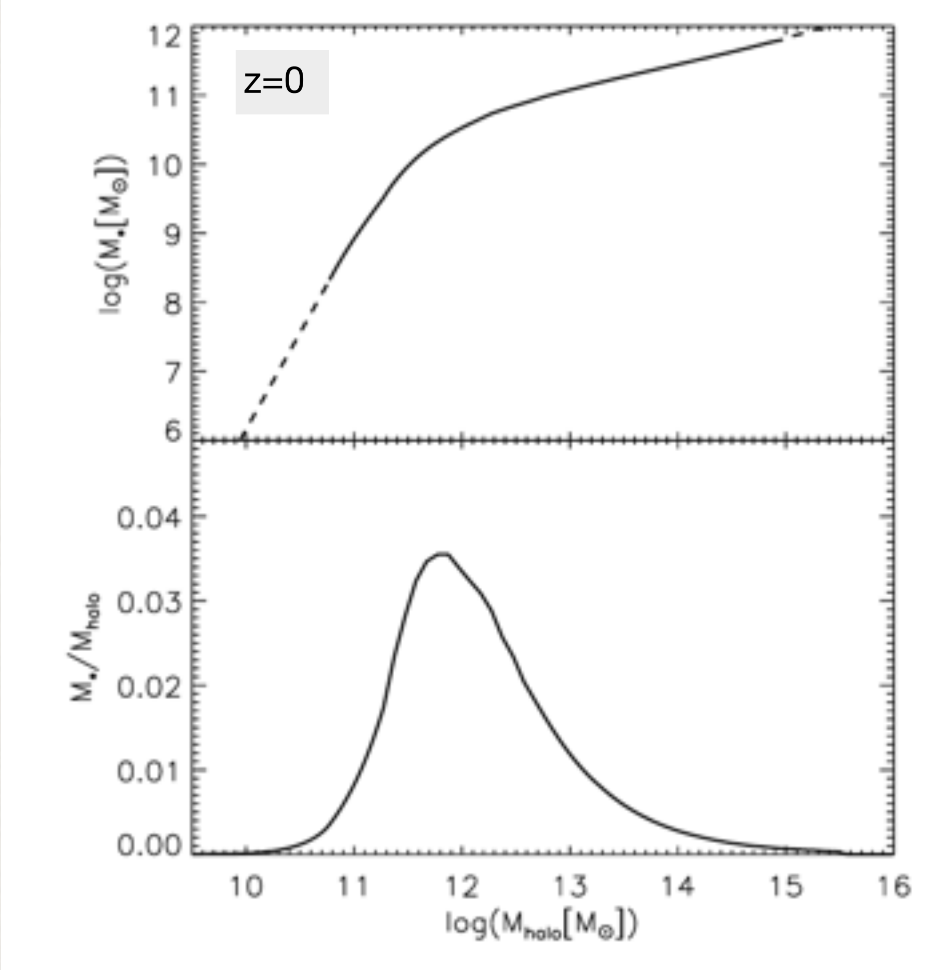


At lower redshifts, our results are consistent with CFHTLS measurements, namely halo masses at M1/M_min which remains constant 0<z<1

However, in all our samples we detect **strong evolution** of M_min and M1 over the redshift range 1<z<2; at higher redshifts, M1 and M_min rise rapidly

Can understand these results in terms of the **evolution of the stellar mass function**

Halo mass / stellar mass relationship



Guo et al 2010

Efficiency of star-formation depends on halo mass

Different physical processes act in low mass and high mass haloes to reduce star-formation efficiency

In high-mass haloes, AGN feedback suppresses star formation

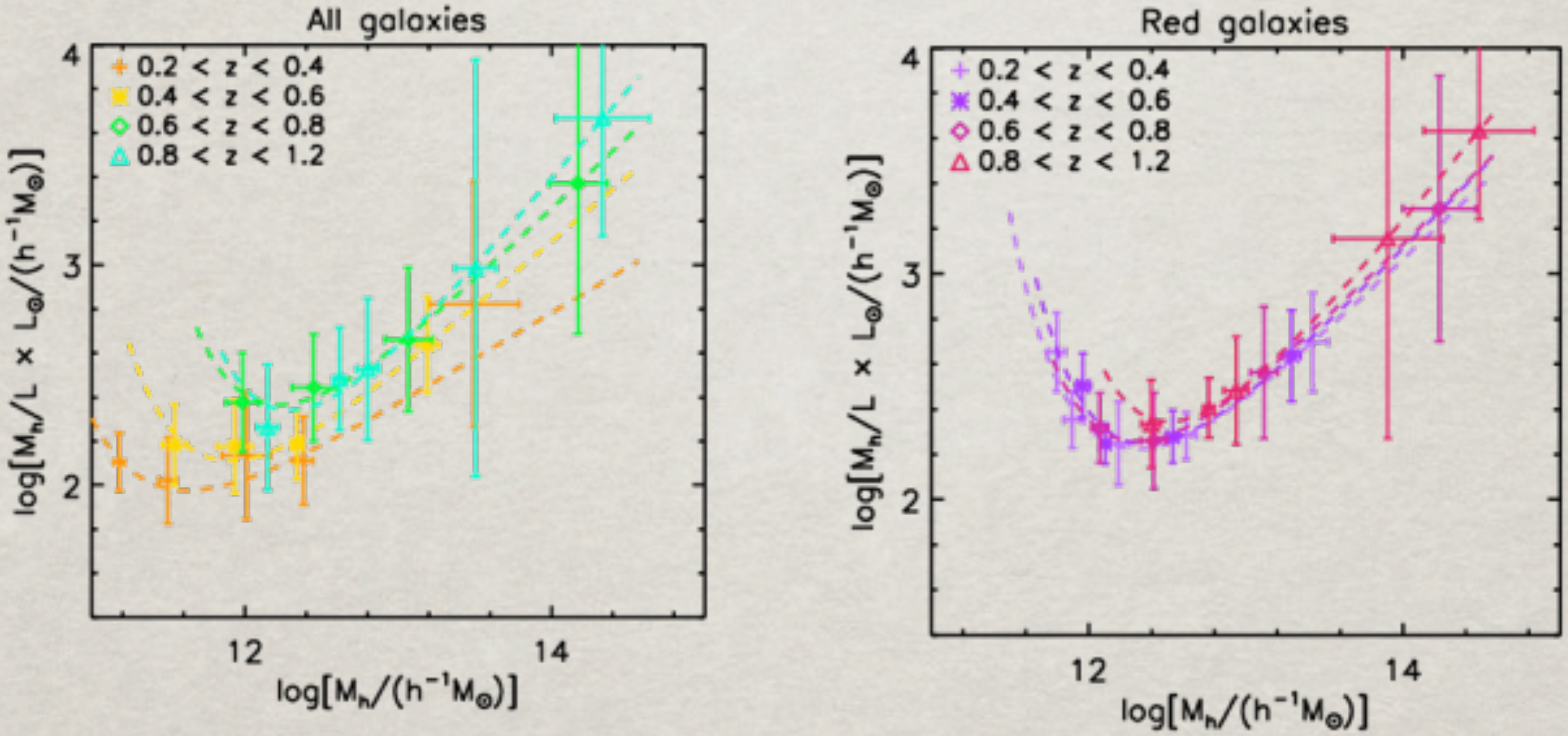
In lower-mass haloes, supernovae-driven winds can have the same effect

In addition, the halo mass at which star-formation is the most efficient can move from high halo masses to low halo masses at lower redshifts (another manifestation of the “downsizing” phenomena)

Can we make measurements of these phenomena in the CFHTLS/COSMOS?

Coupon et al in prep

How do haloes fill with galaxies?



$$\frac{M_h}{L_{\rm cen}} = \left(\frac{M_h}{L_{\rm cen}}\right)_{M_t} \left(\frac{M_h}{M_t}\right)^{1-\alpha_M} \exp\left(\frac{M_t}{M_h} - 1\right)$$

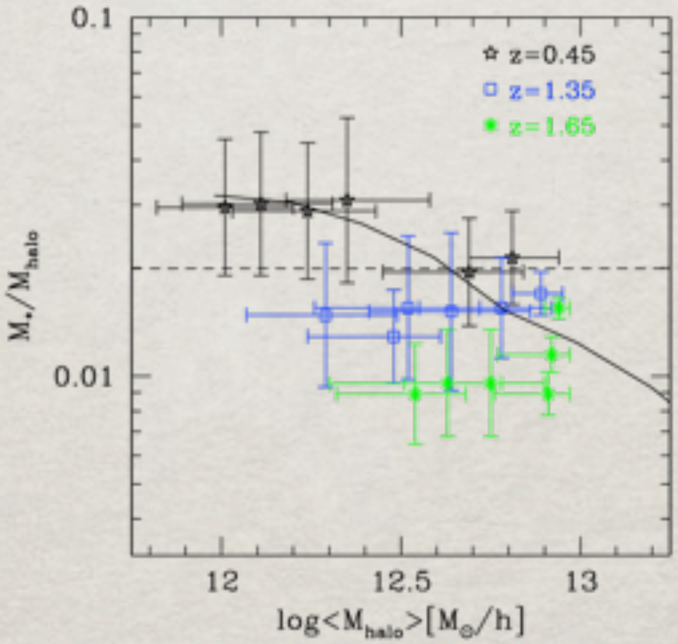
Zehavi 2010

M_t represents the “transition mass” which moves to progressively higher halo masses at higher redshifts

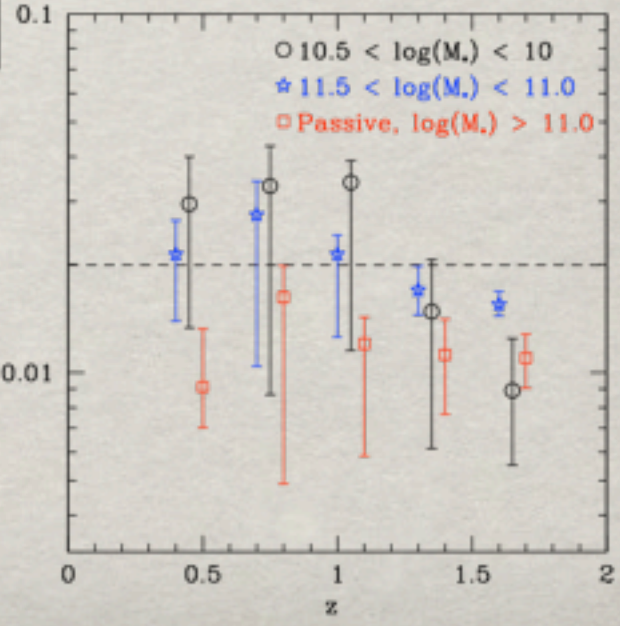
How haloes fill with galaxies (2): extension to high redshifts

In our mass-selected samples, at higher redshifts, our samples deviate from from the local-redshift relationship

In higher mass haloes, star-formation is suppressed (by AGN feedback?)

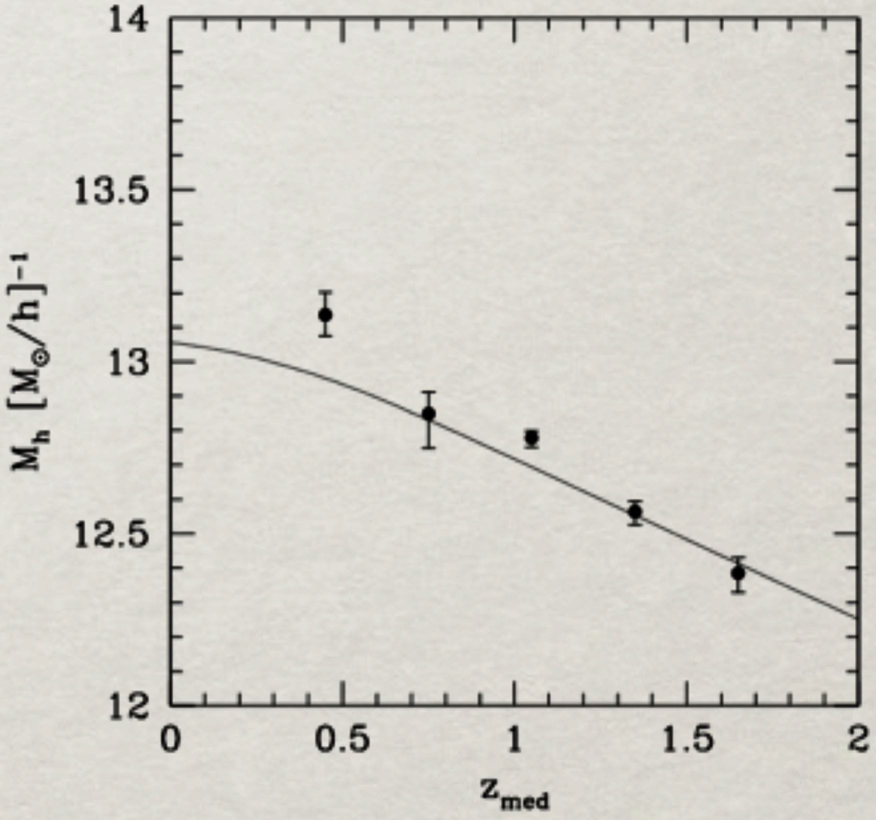


The baryonic mass fraction in massive haloes does not change



Halo mass evolution

- Samples are selected to have a constant stellar mass with redshift
- Is the change in halo mass with redshift consistent with hierarchical merging seen in large N-body simulations?
- We can compare with the fitting formula given by Zhao et al. (2009) based on merger trees in numerical simulations; average halo mass evolution agrees well with this.



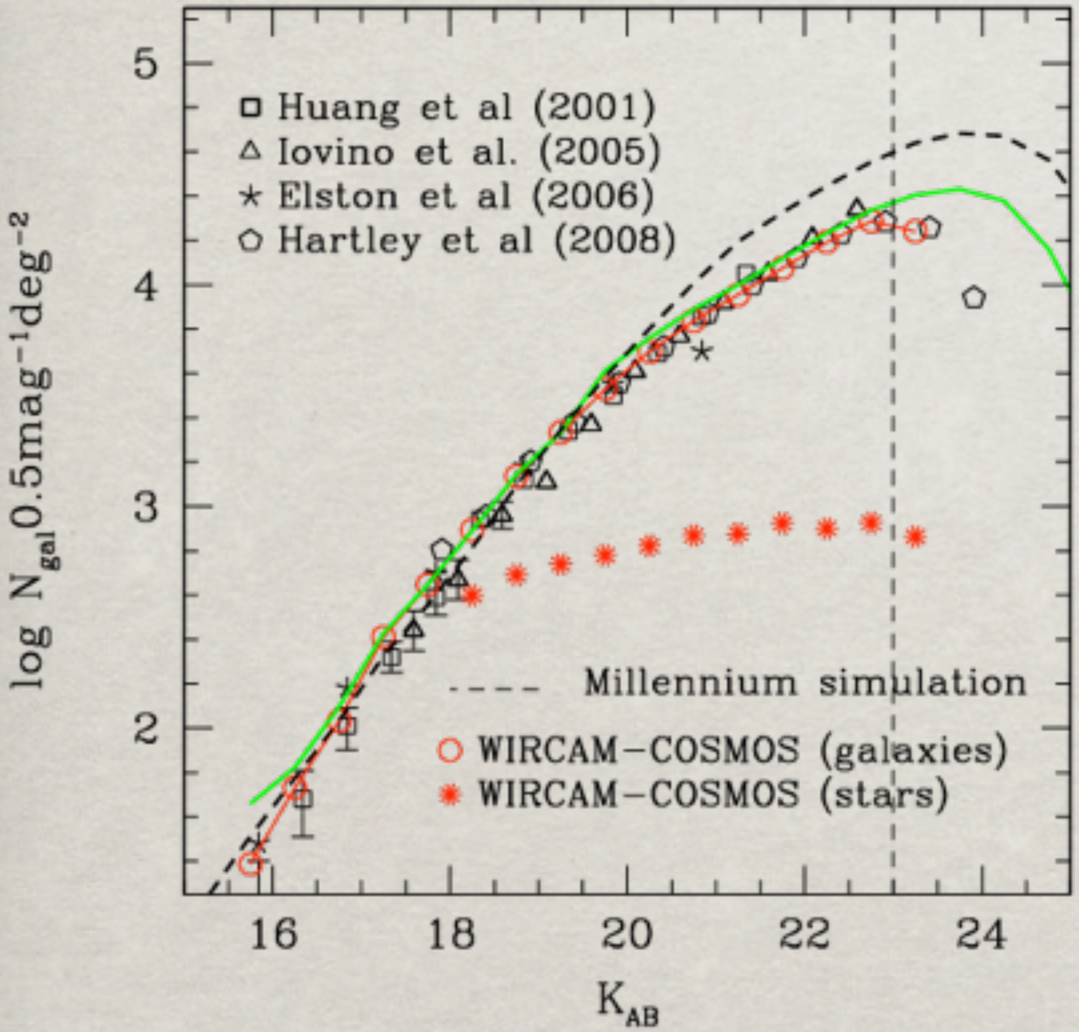
Conclusions and prospects

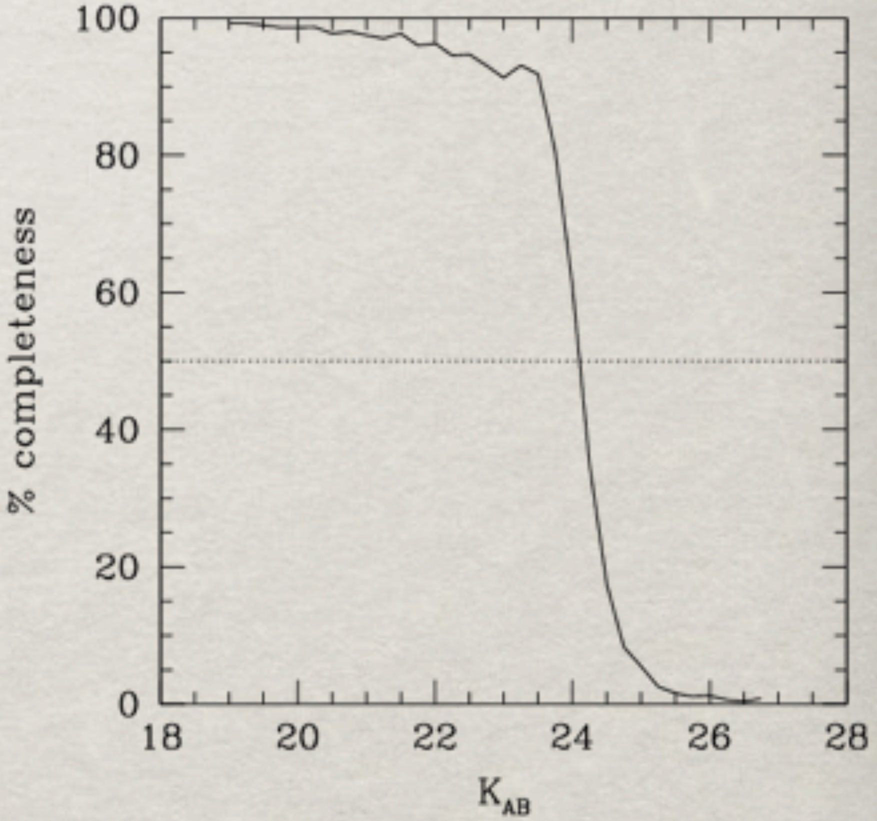
Clear trends are observed between luminosity, halo mass and satellite fraction in the CFHTLS but interpreting these results is “complicated” by the presence of significant luminosity evolution

At z<1 halo evolution follows closely the dark matter evolution

Mass-selected samples evolve at higher redshifts: mass fraction decreases at higher redshifts.

What’s next: Add near IR data to the CFHTLS survey and increase the depth of Near-IR COSMOS: UIltravista survey. First year of ultra-vista data is collected. Precise stellar mass estimates for a **wider range** of halo masses


Huang et al (2001)
Iovino et. al. (2005)
Elston et al (2006)
Hartley et al (2008)
Millennium simulation
WIRCAM−COSMOS (galaxies)
WIRCAM−COSMOS (stars)
$\log N_{gal} 0.5 mag^{-1} deg^{-2}$
2
3
4
5
16
18
20
22
24
K_{AB}


% completeness
0
20
40
60
80
100
18
20
22
24
26
28
K_{AB}