The Galaxy-Dark Matter Connection

Constraining Cosmology & Galaxy Formation

Frank van den Bosch (MPIA)

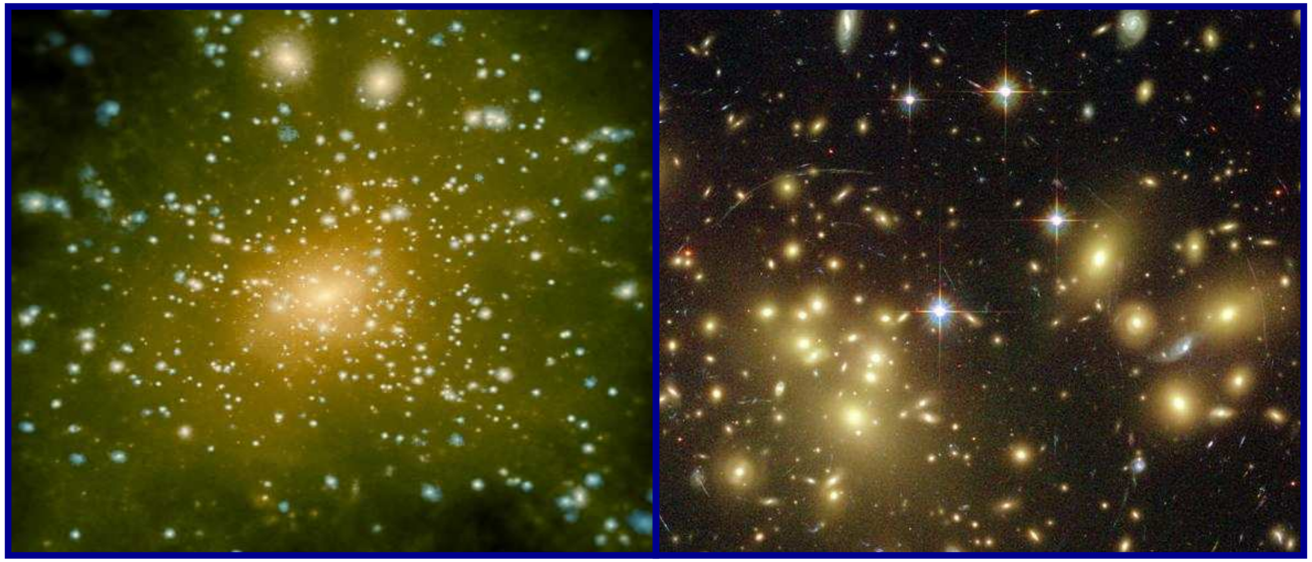
in collaboration with

Houjun Mo (UMass), Xiaohu Yang (SHAO), Simone Weinmann (Zurich)
Surhud More (MPIA), Marcello Cacciato (MPIA), Anna Pasquali (MPIA)

Motivation and Techniques

Why study the Galaxy-Dark Matter Connection?

- To constrain the physics of Galaxy Formation
- To constrain Galaxy Bias and Cosmological Parameters
- To interpret Galaxy-Galaxy Lensing and Satellite Kinematics



How to Constrain the Galaxy-Dark Matter Connection?

- Luminosity Dependent Clustering
- Galaxy Group Catalogues
- Satellite Kinematics

The Issue of Galaxy Bias

An important goal in cosmology is to probe the matter field.

Define the density perturbation field: $\delta(\vec{x},z) = \frac{\rho(\vec{x},z)-\bar{\rho}(z)}{\bar{\rho}(z)}$

Since most matter is dark, one uses galaxies as a tracer population

However, galaxies are a biased tracer of the mass distribution.

$$\xi_{\rm gg}(r) = b_g^2\, \xi_{\rm dm}(r) \quad \text{with} \quad b_{\rm g} = \langle \delta_{\rm g} / \delta_{\rm dm} \rangle$$

The Issue of Galaxy Bias

An important goal in cosmology is to probe the matter field.

Define the density perturbation field: $\delta(\vec{x}, z) = \frac{\rho(\vec{x},z) - \bar{\rho}(z)}{\bar{\rho}(z)}$

Since most matter is dark, one uses galaxies as a tracer population

However, galaxies are a biased tracer of the mass distribution.

$$\xi_{\rm gg}(r) = b_g^2 \, \xi_{\rm dm}(r) \qquad \text{with} \qquad b_{\rm g} = \langle \delta_{\rm g} / \delta_{\rm dm} \rangle$$

Bias is an imprint of galaxy formation, which is poorly understood.

Observations show that $\xi_{\rm gg}(r)$ depends on galaxy properties

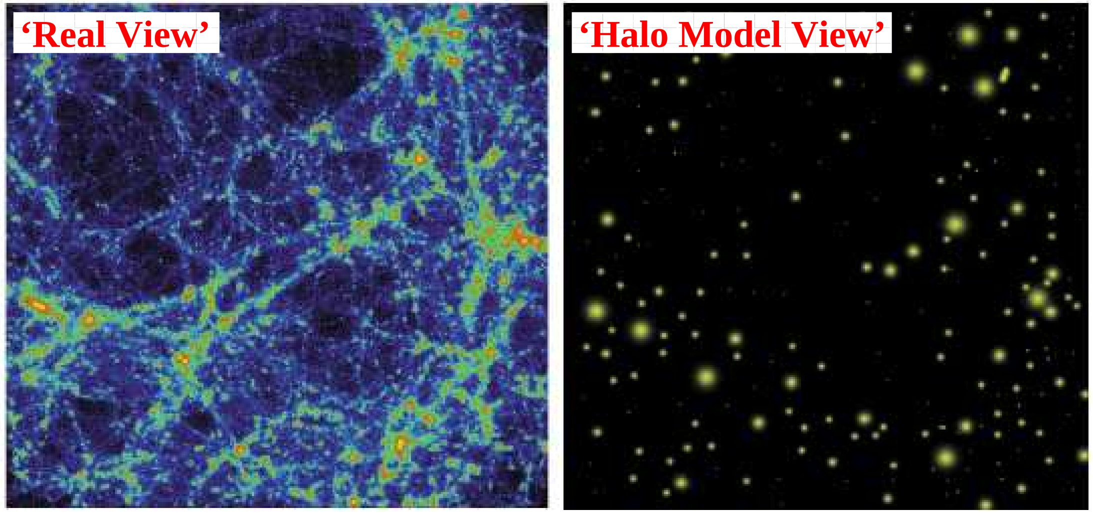
Consequently, galaxy bias $b_{\rm g}$ also depends on galaxy properties.

Consequently, little progress has been made constraining cosmology with Large-Scale Structure, despite several large redshift surveys.

How to constrain and quantify galaxy bias in a convenient way?

How to Handle Bias?

Halo Model: Describe CDM distribution in terms of halo building blocks, assuming that every CDM particle resides in virialized halo



Cooray & Sheth (2002)

- On small scales: $\delta(\vec{x}) =$ density distribution of halos
- On large scales: $\delta(\vec{x}) =$ spatial distribution of halos

Halo Bias: Dark Matter haloes are biased tracer of mass distribution.
More massive haloes are more strongly biased.

Halo Occupation Statistics: A statistical description of how galaxies are distributed over dark matter halos

Galaxy Bias = Halo Bias + Halo Occupation Statistics

The Conditional Luminosity Function

To specify Halo Occupation Statistics we introduce Conditional Luminosity Function, $\Phi(L|M)$, which is the direct link between halo mass function $n(M)$ and the galaxy luminosity function $\Phi(L)$:

$$\boxed{\Phi(L) = \int_0^\infty \Phi(L|M)\, n(M)\, \mathrm{d}M}$$

The CLF contains a lot of important information, such as:

- **The average relation between light and mass:**

$$\boxed{\langle L\rangle(M) = \int_0^\infty \Phi(L|M)\, L\, \mathrm{d}L}$$

- **The bias of galaxies as function of luminosity:**

$$\boxed{b_g(L) = \frac{1}{\Phi(L)} \int_0^\infty \Phi(L|M)\, b_h(M)\, n(M)\, \mathrm{d}M}$$

CLF is ideal statistical tool to specify Galaxy-Dark Matter Connection

Luminosity & Correlation Functions

Introduction

Galaxy Bias

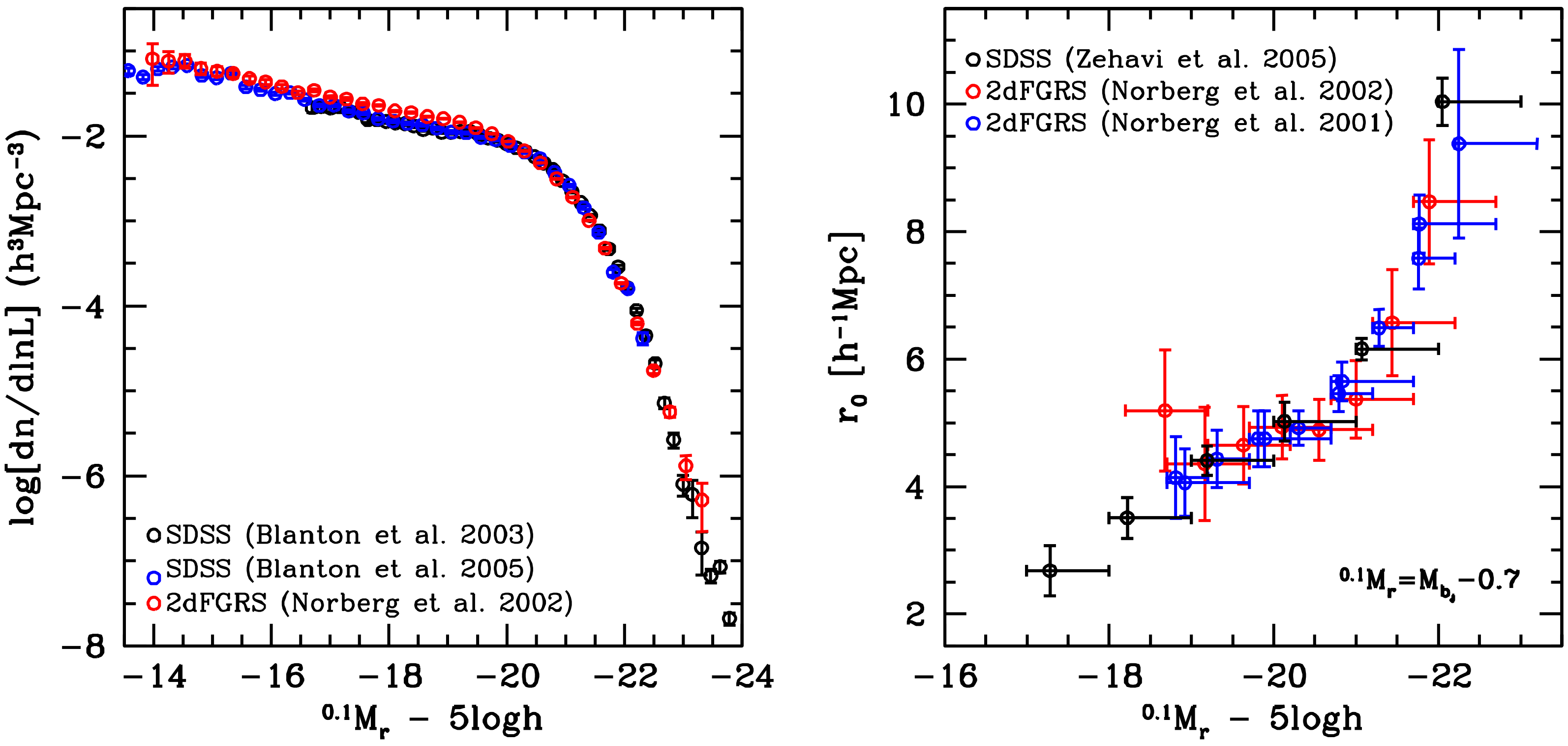
Conditional Luminosity Function
- The Conditional Luminosity Function
- Luminosity & Correlation Functions
- The CLF Model
- Best-Fit Models
- Constructing Galaxy Groups Catalogues
- The Galaxy-Dark Matter Connection

Ecology

Stochasticity

Conclusions

Extra Material



- **DATA:** More luminous galaxies are more strongly clustered.
- **ΛCDM:** More massive haloes are more strongly clustered.

More luminous galaxies reside in more massive haloes

REMINDER: Correlation length r_0 defined by $\xi(r_0) = 1$

The CLF Model

For 2dFGRS we assume that CLF has Schechter form:

$$\boxed{\Phi(L|M)\mathrm{d}L = \frac{\Phi^*}{L^*}\left(\frac{L}{L^*}\right)^{\alpha}\exp[-(L/L^*)]\,\mathrm{d}L}$$

Here Φ^*, L^* and α all depend on M.

(e.g., Yang et al. 2003; vdB et al. 2003, 2005)

For SDSS we split CLF in central and satellite components:

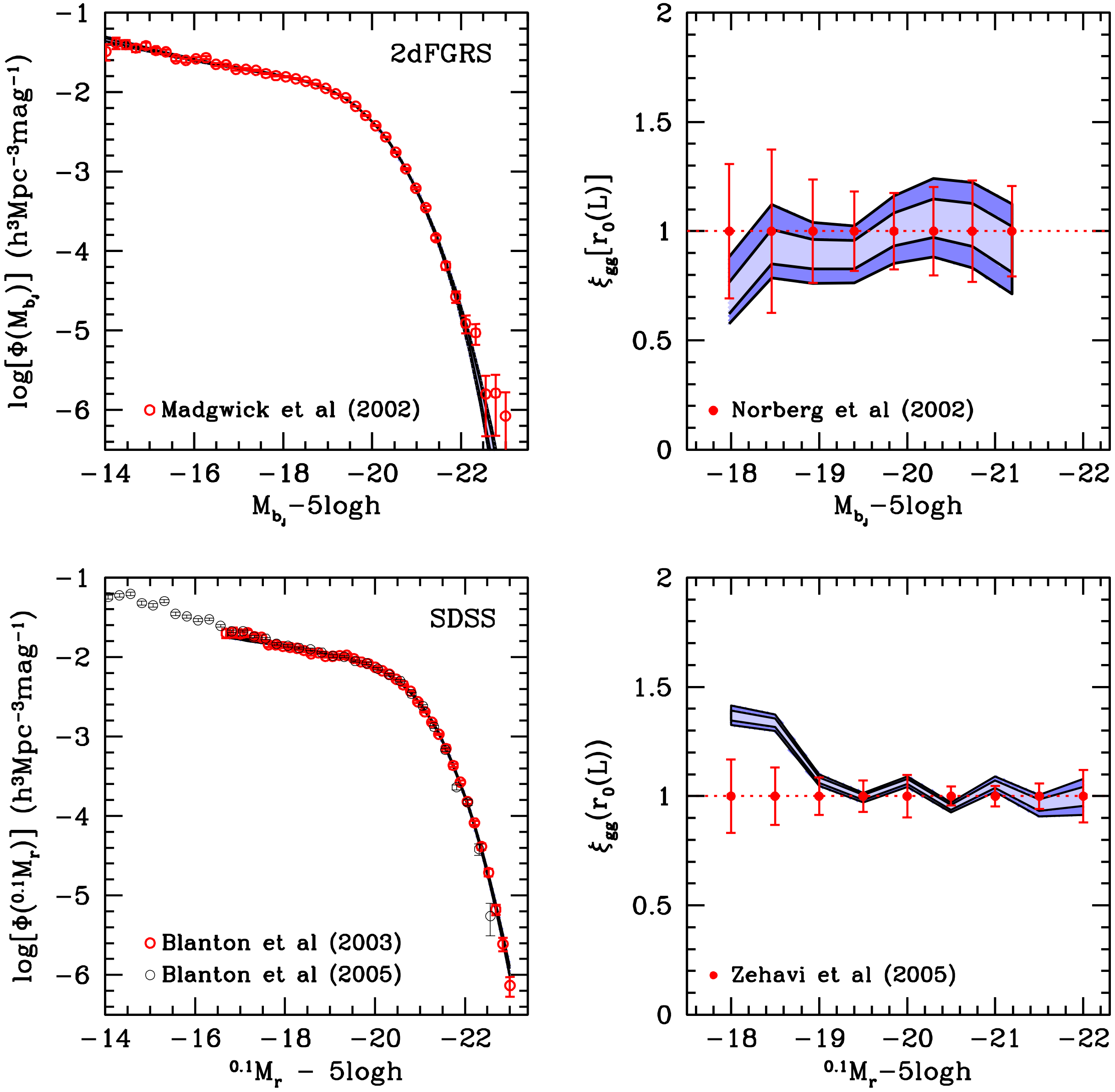
$$\begin{aligned}
\Phi(L|M)\mathrm{d}L &= \Phi_c(L|M)\mathrm{d}L + \Phi_s(L|M)\mathrm{d}L \\
\Phi_c(L|M)\mathrm{d}L &= \frac{1}{\sqrt{2\pi}\,\ln(10)\,\sigma_c}\exp\left[-\left(\frac{\log(L/L_c)}{\sqrt{2}\sigma_c}\right)^2\right]\frac{\mathrm{d}L}{L} \\
\Phi_s(L|M)\mathrm{d}L &= \frac{\Phi_s}{L_s}\left(\frac{L}{L_s}\right)^{\alpha_s}\exp[-(L/L_s)^2]\,\mathrm{d}L
\end{aligned}$$

Here L_c, L_s, σ_c, ϕ_s and α_s all depend on M

(e.g., Cooray & Milosavljevic 2005; Cooray 2005, 2006; vdB et al. 2007)

Use Monte-Carlo Markov Chain to constrain free parameters by fitting to $\Phi(L)$ and $r_0(L)$.

Best-Fit Models



2dFGRS: vdB et al. 2006 (astro-ph/0610686)

SDSS: vdB et al. 2007 (in preparation)

Constructing Galaxy Groups Catalogues

Introduction

Galaxy Bias

Conditional Luminosity Function
- The Conditional Luminosity Function
- Luminosity & Correlation Functions
- The CLF Model
- Best-Fit Models
- Constructing Galaxy Groups Catalogues
- The Galaxy-Dark Matter Connection

Ecology

Stochasticity

Conclusions

Extra Material

Galaxy-Dark Matter connection can be studied more directly by measuring the occupation statistics of galaxy groups.

Potential Problems: interlopers, (in)completeness, mass estimates

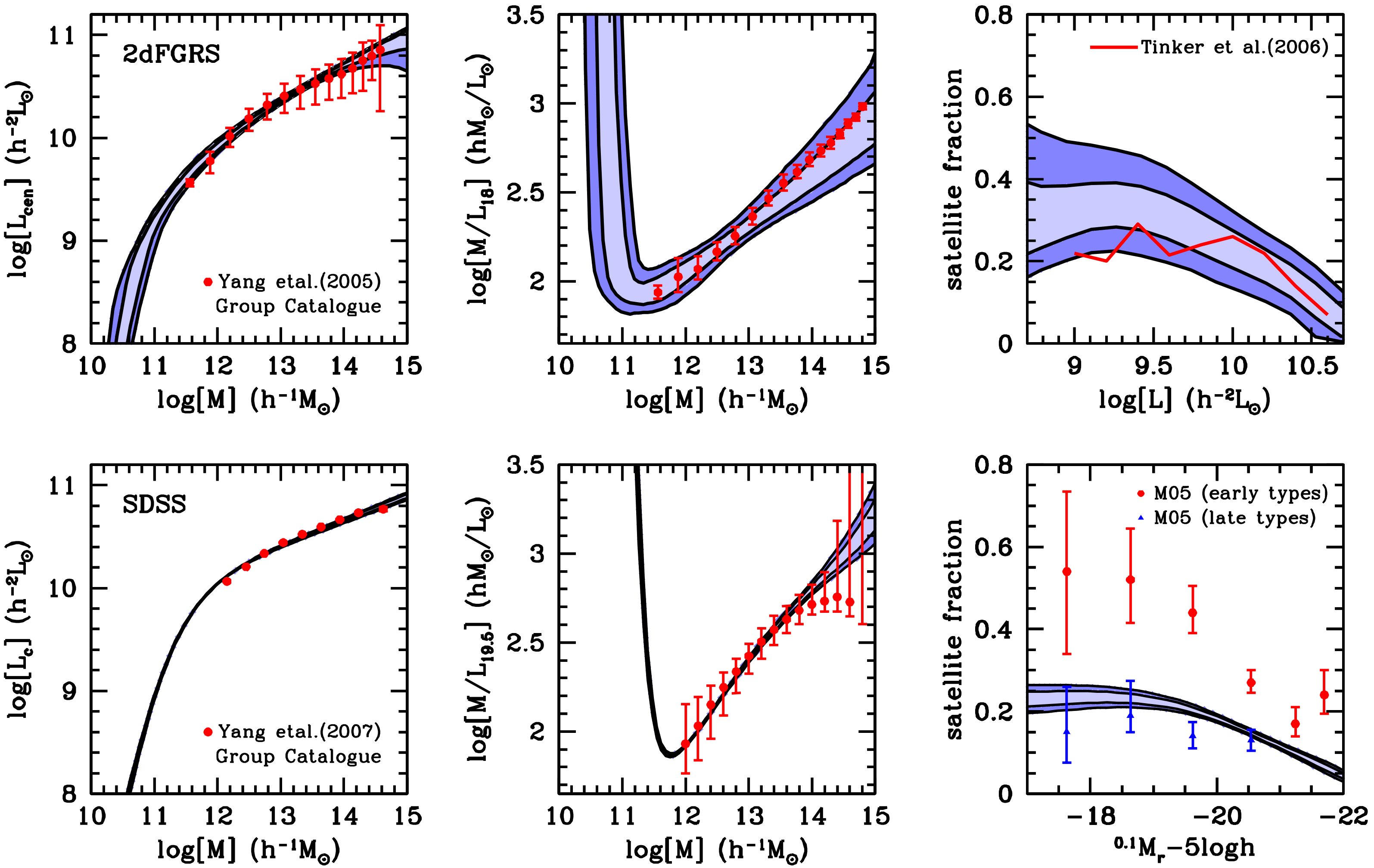
We have developed a new, iterative group finder which uses an adaptive filter modeled after halo virial properties

Yang, Mo, vdB, Jing 2005, MNRAS, 356, 1293

- **Calibrated & Optimized with Mock Galaxy Redshift Surveys**
- **Low interloper fraction ($\lesssim 20\%$).**
- **High completeness of members ($\gtrsim 90\%$).**
- **Masses estimated from group luminosities/stellar masses. More accurate than using velocity dispersion of members.**
- **Can also detect “groups” with single member**
  - ▷ **Large dynamic range ($11.5 \lesssim \log[M/\mathrm{M}_\odot] \lesssim 15$).**

Group finder has been applied to both the 2dFGRS (completed survey) and to the SDSS (NYU-VAGC DR2 + DR4; Blanton et al. 2005)

The Galaxy-Dark Matter Connection



Excellent agreement between CLF and Group results

2dFGRS: vdB et al. 2006 (astro-ph/0610686)

SDSS: vdB et al. 2007 (in preparation)

Galaxy Ecology

Many studies have investigated relation between various galaxy properties (morphology / SFR / colour) and environment

(e.g., Dressler 1980; Balogh et al. 2004; Goto et al. 2003; Hogg et al. 2004)

Environment estimated using galaxy overdensity (projected) to $n^{\rm th}$ nearest neighbour, Σ_n or using fixed, metric aperture, Σ_R.

Previous studies have found that:

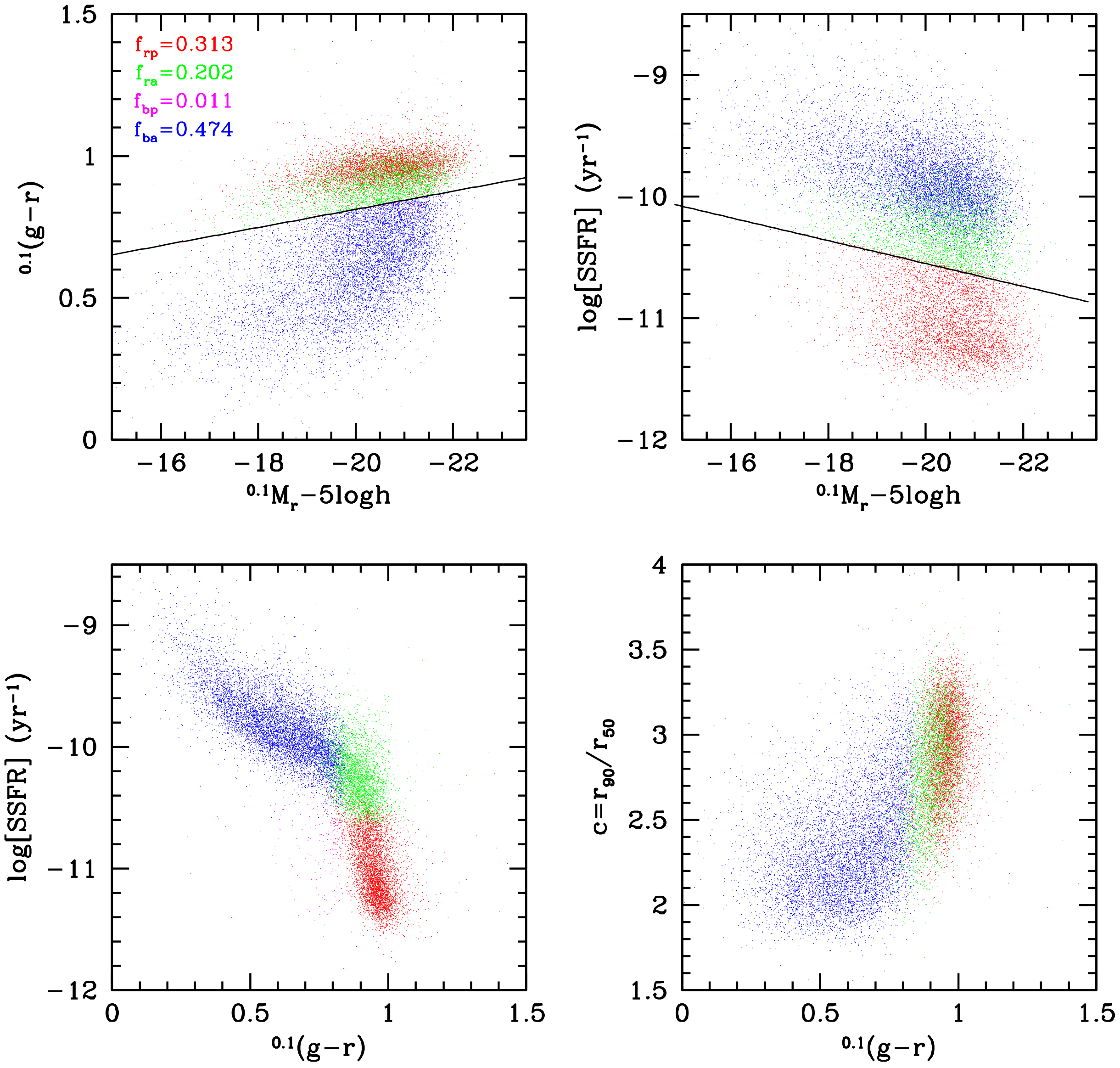
- **Fraction of early types increases with density**
- **There is a characteristic density ($\sim$ group-scale) below which the environment dependence vanishes**

Danger: **Physical meaning of Σ_n and Σ_R depends on environment**

Physically more meaningful to investigate halo mass dependence of galaxy properties. This requires galaxy group catalogues.

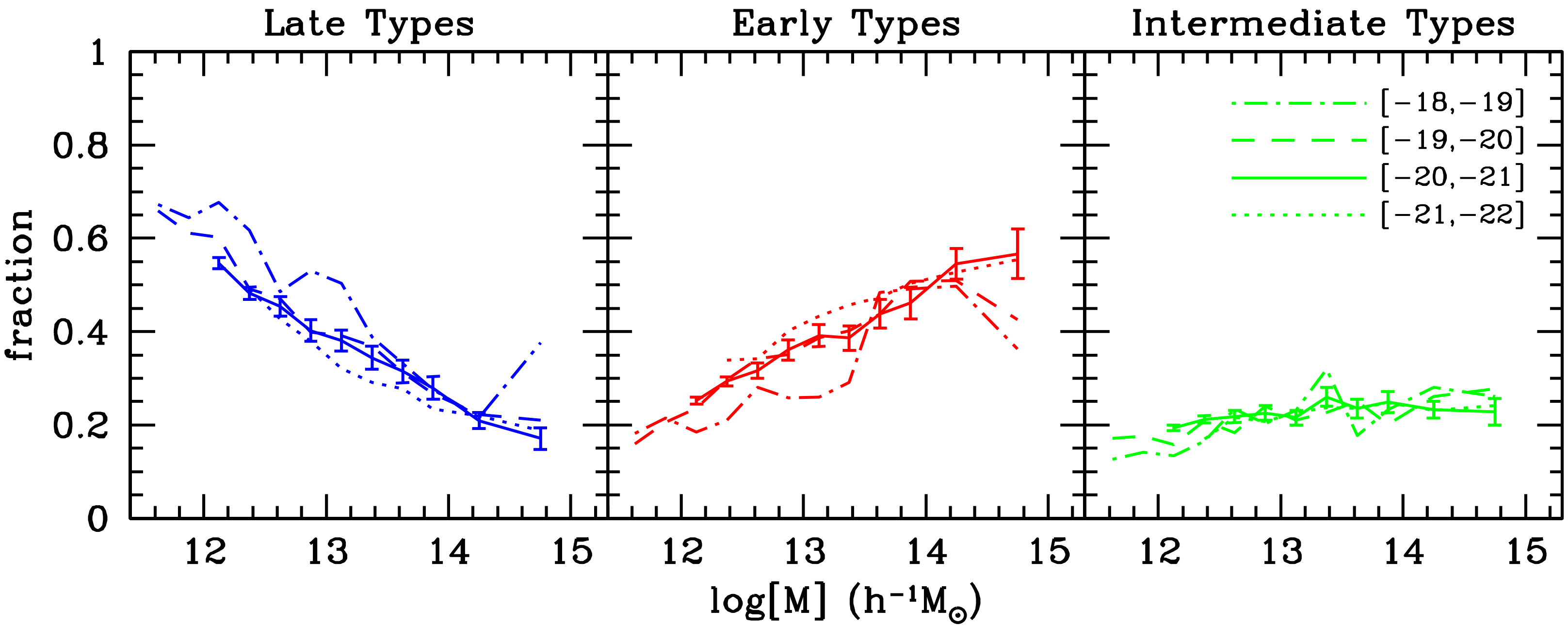
Important: **Separate L-dependence from M-dependence**

Defining Galaxy Types



SDSS-DR2 data from NYU-VAGC (Blanton et al. 2005)
SSFRs from Kauffmann et al. (2003) and Brinchmann et al. (2004)

Halo Mass Dependence



The fractions of early and late types depend strongly on halo mass.

At fixed halo mass, there is virtually no luminosity dependence.

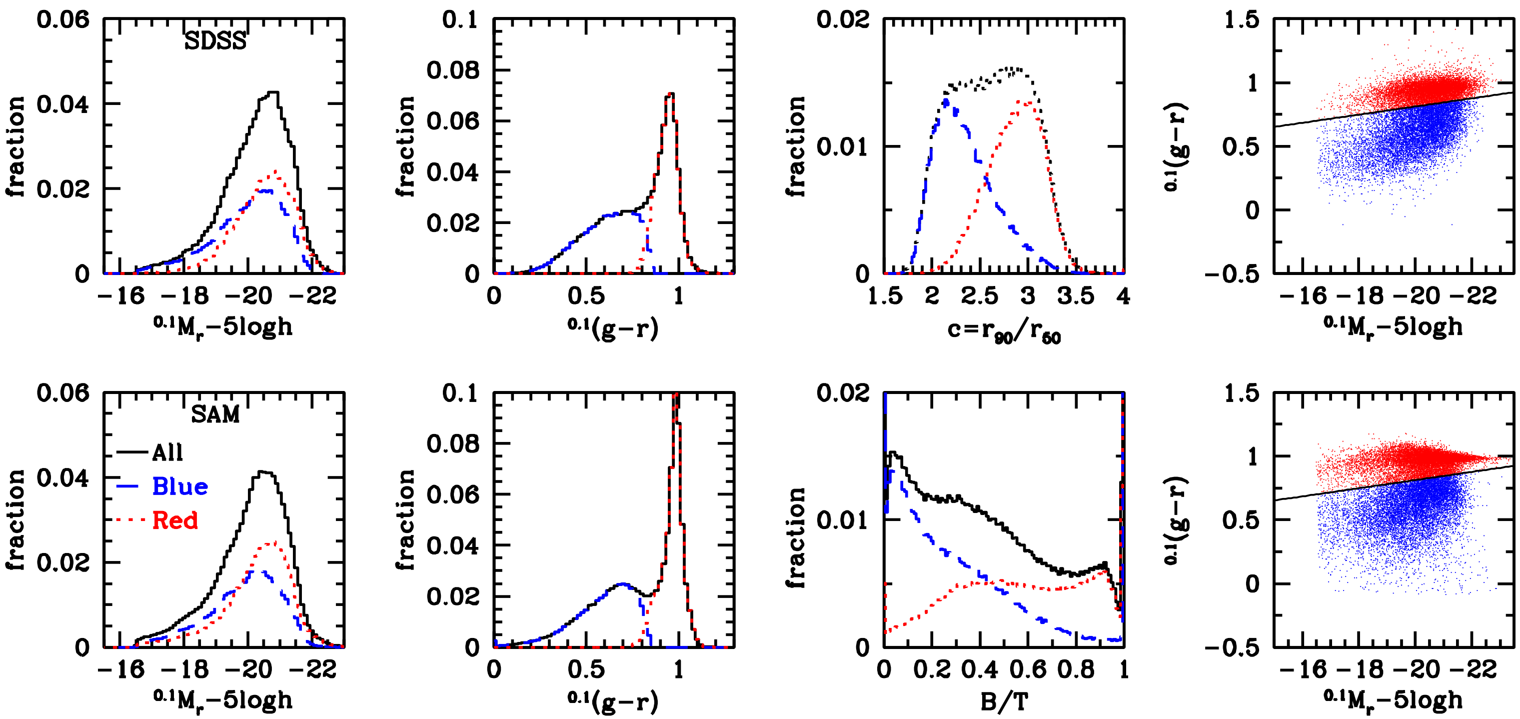
The mass dependence is smooth: there is no characteristic mass scale

The intermediate type fraction is independent of luminosity and mass.

(Weinmann, vdB, Yang & Mo, 2006)

Comparison with Semi-Analytical Model

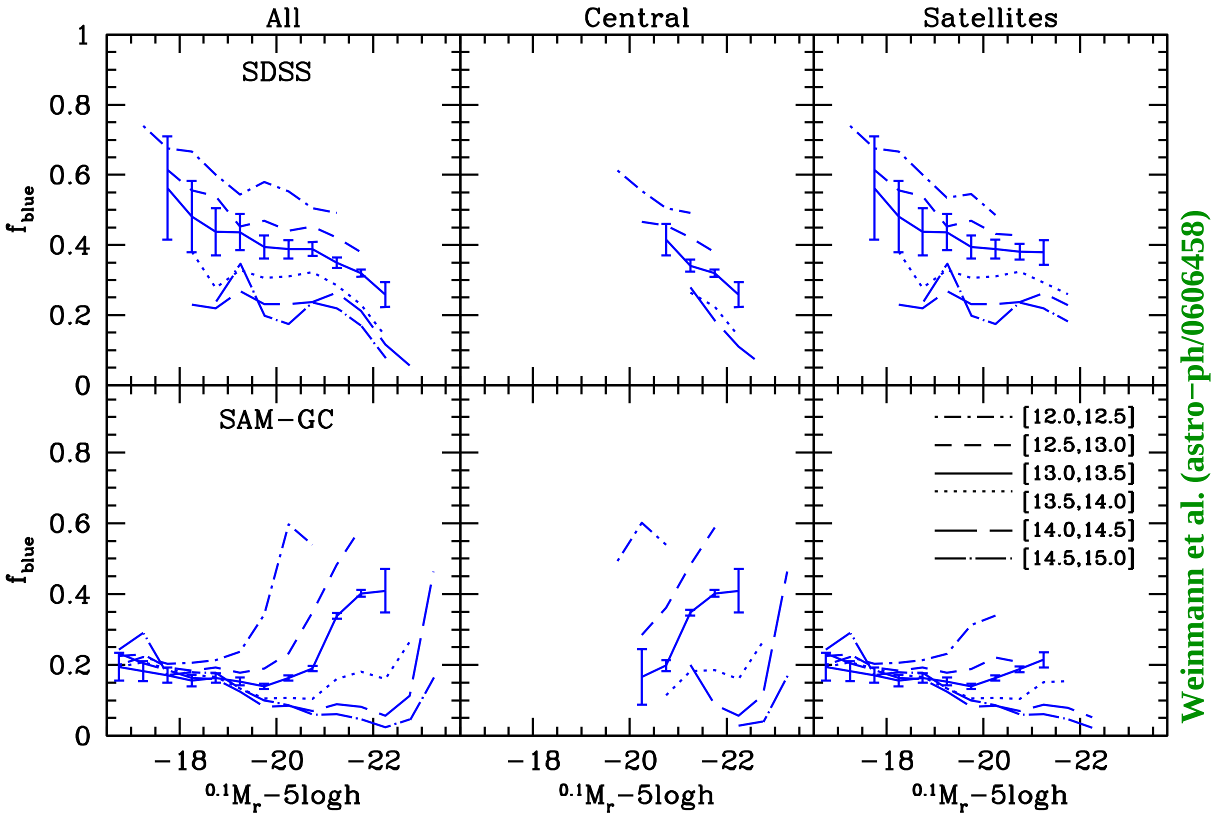
Comparison of Group Occupation Statistics with Semi-Analytical Model of Croton et al. 2006. Includes 'radio-mode' AGN feedback.



- SAM matches **global statistics** of SDSS
- LF, bimodal color distribution, and overall blue fraction
- But what about statistics as function of halo mass?

Constraining Star Formation Truncation

To allow for fair comparison, we run our Group Finder over SAM.



Weinmann et al. (astro-ph/0606458)

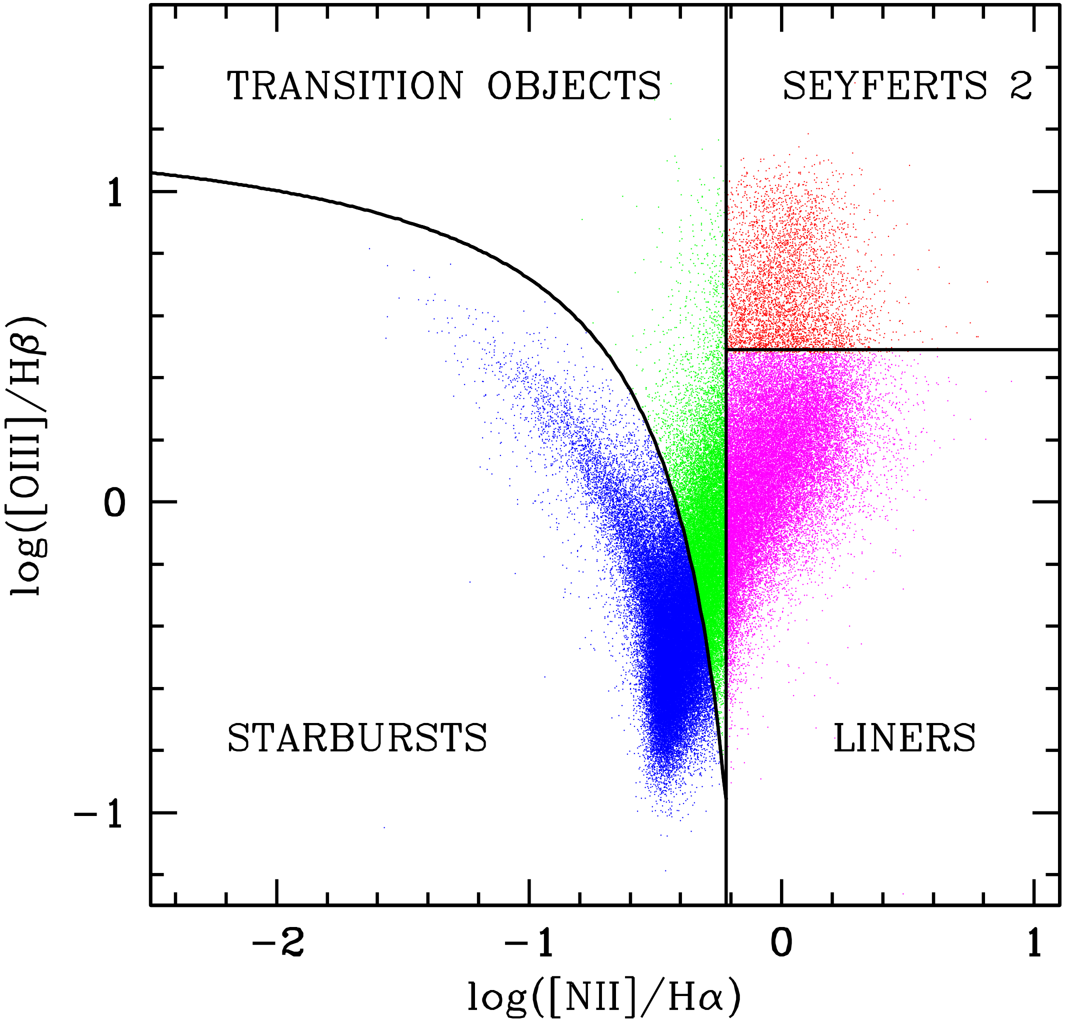
Satellites: red fraction too large: ▷ strangulation too efficient

Centrals: $f_{\rm blue}(L|M)$ wrong: ▷ Problem with AGN feedback or dust

$f_{\rm blue}(L,M)$ useful to constrain SF truncation mechanism

Defining Activity Classes

Galaxies can be classified in Seyferts, Liners and Starbursts using emission line ratios. We also use Radio detections from FIRST.

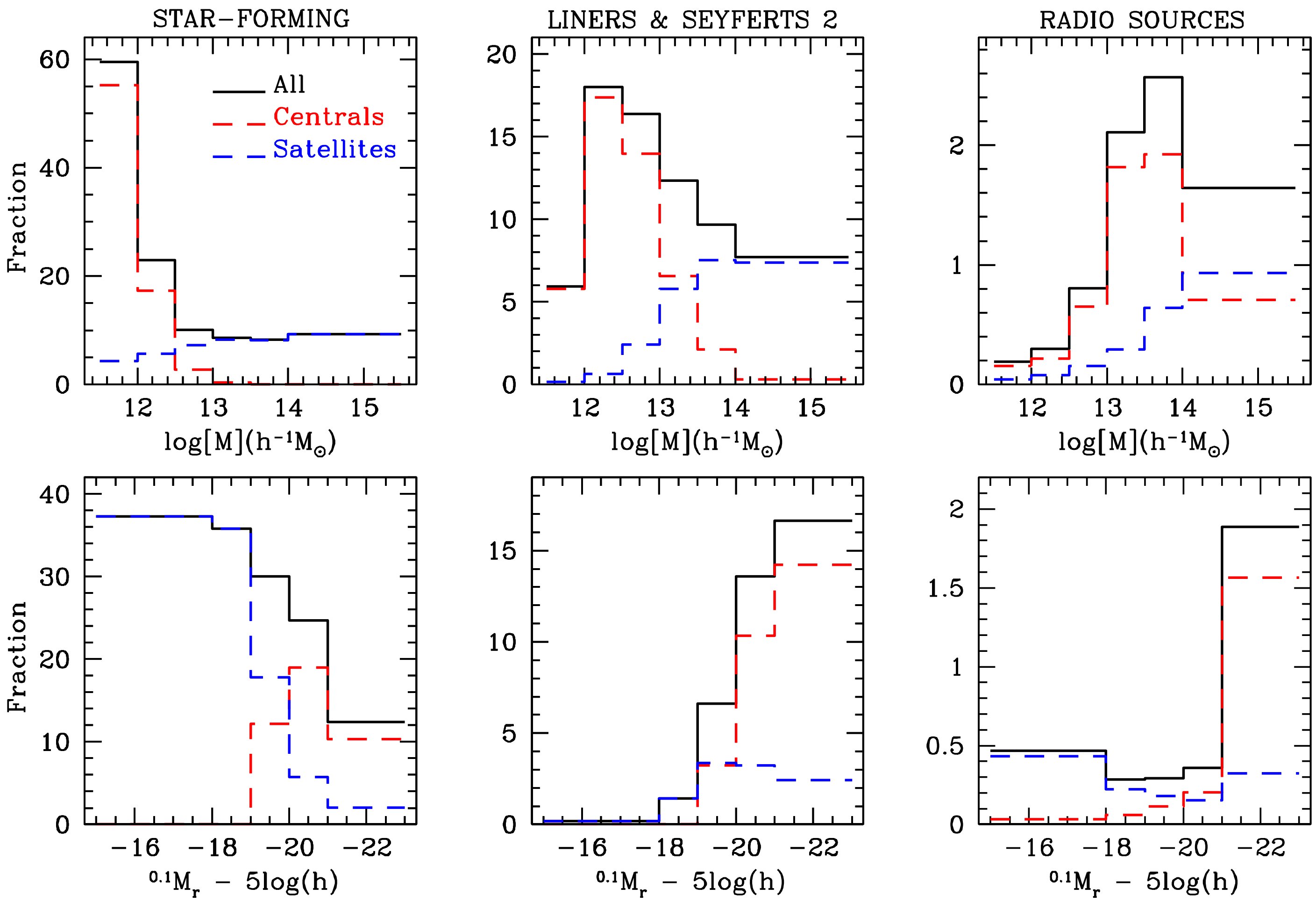


Pasquali, vdB, et al. 2007, in prep.

Ecology of AGN and Starbursts

Introduction
Galaxy Bias
Conditional Luminosity Function
Ecology
- Galaxy Ecology
- Defining Galaxy Types
- Halo Mass Dependence
- Comparison with Semi-Analytical Model
- Constraining Star Formation Truncation
- Defining Activity Classes
- Ecology of AGN and Starbursts

Stochasticity
Conclusions
Extra Material



- **Central** **SB** **activity truncated at** $M \sim 10^{12} h^{-1}\,\mathrm{M}_\odot$
- **Central** **AGN** **activity peaks at** $M \sim 3 \times 10^{12} h^{-1}\,\mathrm{M}_\odot$
- **Radio-mode** **AGN activity peaks at** $M \sim 3 \times 10^{13} h^{-1}\,\mathrm{M}_\odot$

Pasquali, vdB, et al. 2007, in prep.

Stochasticity and Stacking

To measure satelite kinematics or the weak lensing shear around galaxies, one needs to stack the signal of many galaxies.

Typically one stacks (central) galaxies in a narrow luminosity bin.

Unless $P(M|L_{\rm cen})$ is very narrow, this means stacking haloes of different masses, and signal does not reflect $\langle M\rangle(L_{\rm cen})$.

Proper interpretation of satelite kinematics and galaxy-galaxy lensing requires knowledge of $\sigma_{\rm logM}$.

How can we constrain the scatter in $P(M|L_{\rm cen})$?

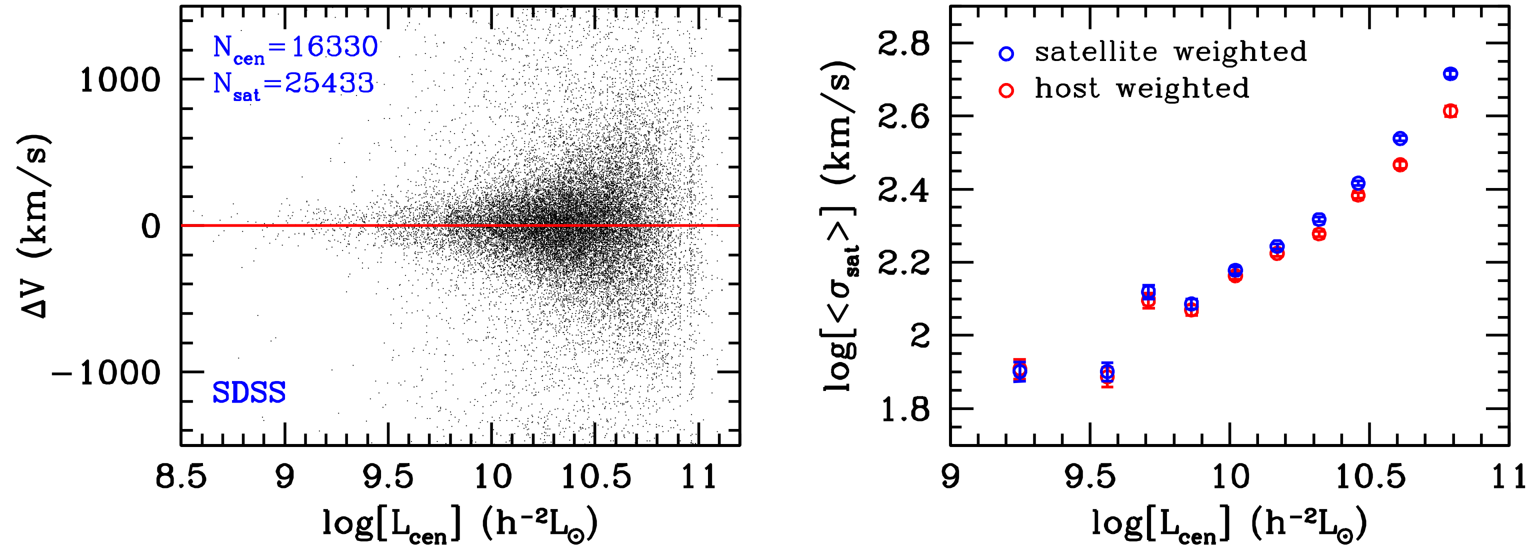
- **Use 'predictions' from semi-analytical models for galaxy formation**
- **Compute from CLF:** $P(M|L_{\rm cen}) = \frac{\Phi_c(L|M)\,n(M)}{\Phi_c(L)}$ *(Bayes Theorem)*
- **Use satellite kinematics; host-weighting *vs.* satellite weighting**

Satellite Kinematics

Introduction
Galaxy Bias
Conditional Luminosity Function
Ecology
Stochasticity
- Stochasticity and Stacking
- Satellite Kinematics
- The Scatter in $P(M|L_{\rm cen})$

Conclusions
Extra Material

Select centrals and satellites and determine $\sigma_{\rm sat}(L_{\rm cen})$, describing the width of $P(\Delta V)$ with $\Delta V = V_{\rm sat} - V_{\rm cen}$ (More, vdB, et al. 2007, in prep.)



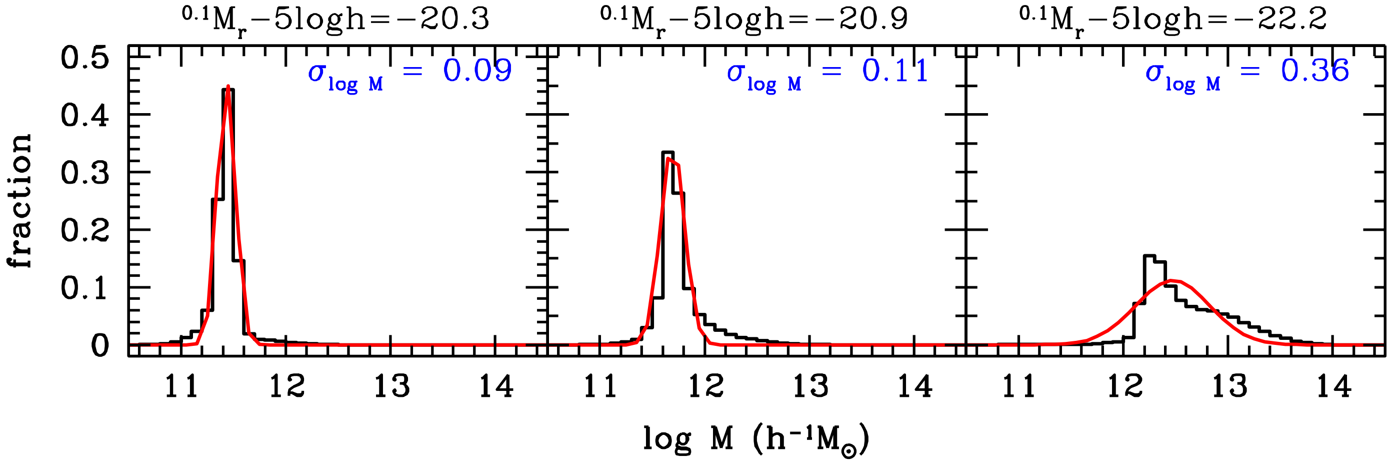
$$\langle\sigma_{\rm sat}\rangle(L_{\rm cen}) = \frac{\int P(M|L_{\rm cen})\,\langle N_{\rm sat}\rangle_M^p\,\langle\sigma_{\rm sat}\rangle_M\,{\rm d}M}{\int P(M|L_{\rm cen})\,\langle N_{\rm sat}\rangle_M^p\,{\rm d}M}$$

- $p = 1$: **satellite-weighted mean** $\langle\sigma_{\rm sat}\rangle_{\rm sw}$
- $p = 0$: **host-weighted mean** $\langle\sigma_{\rm sat}\rangle_{\rm hw}$

Unless $P(M|L_{\rm cen}) = \delta(M - \langle M\rangle)$ one has that $\langle\sigma_{\rm sat}\rangle_{\rm sw} > \langle\sigma_{\rm sat}\rangle_{\rm hw}$

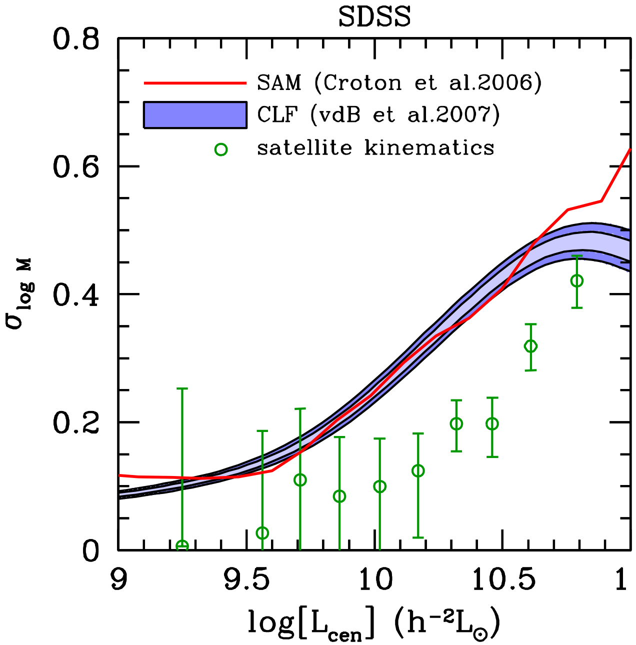
Both $\langle\sigma_{\rm sat}\rangle_{\rm sw}$ and $\langle\sigma_{\rm sat}\rangle_{\rm hw}$ can be determined from data.

The Scatter in $P(M|L_{\rm cen})$



Assuming that $P(M|L_{\rm cen})$ is a log-normal, the combination of $\langle\sigma_{\rm sat}\rangle_{\rm sw}$ and $\langle\sigma_{\rm sat}\rangle_{\rm hw}$ allows one to compute both $\langle M\rangle$ and $\sigma_{\log M}$ as function of $L_{\rm cen}$.

(More, vdB et al. 2007, in prep.)

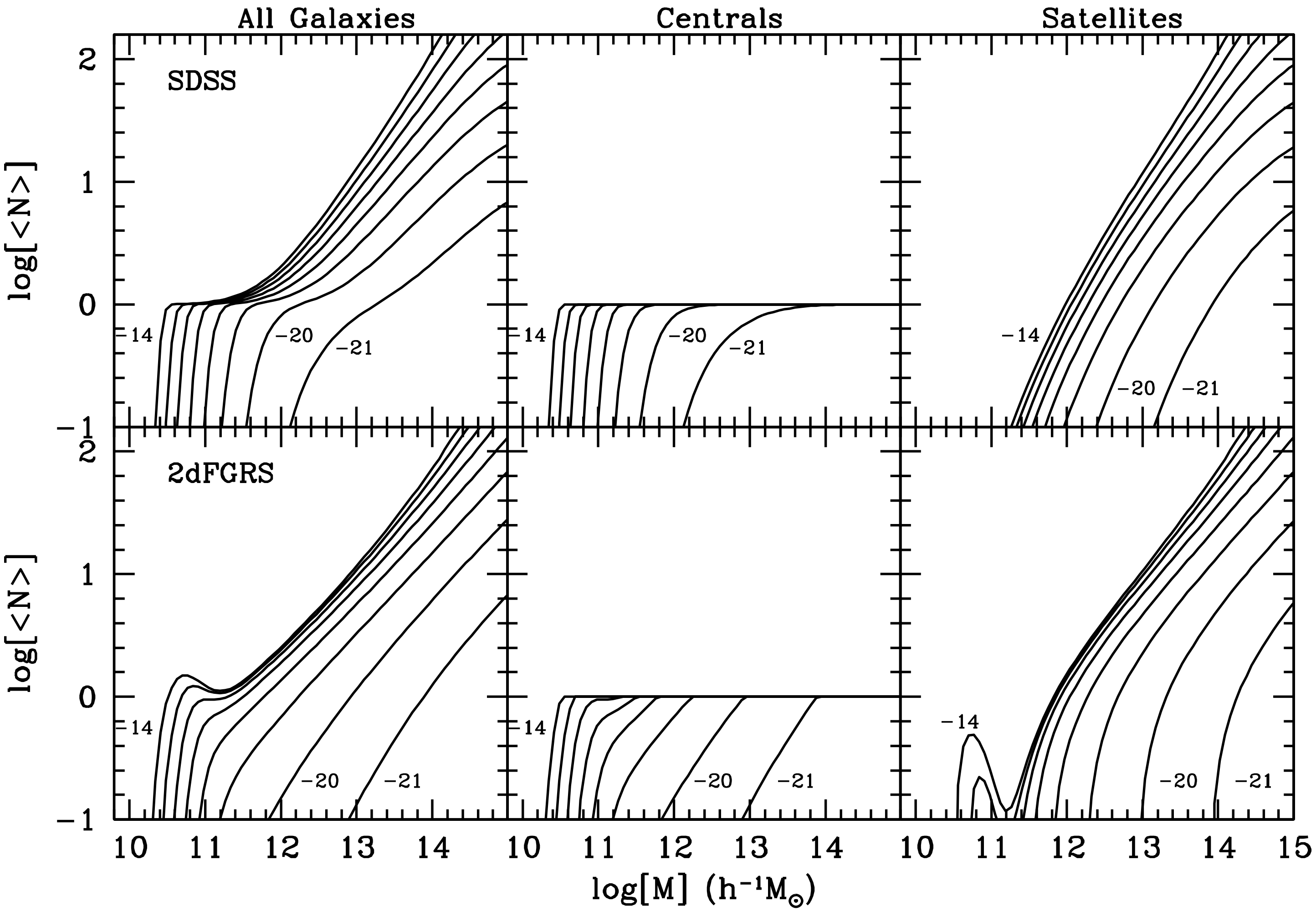


All methods agree that scatter in $P(M|L_{\rm cen})$ increases with $L_{\rm cen}$

Conclusions

- The **CLF** allows a powerful and consice treatment of **galaxy bias**.
- The **CLF** also quantifies universal relation between light and mass.
- Galaxy-Dark Matter connection inferred from **luminosity dependent clustering** in excellent agreement with results obtained from **galaxy group catalogues**.
- The **ecology** of galaxies as function of halo mass yields useful constraints on physics of galaxy formation.
- Satellite kinematics can be used to probe and quantify the **stochasticity** in galaxy formation.
- Scatter in $P(M|L_{\rm cen})$ increases strongly with increasing $L_{\rm cen}$

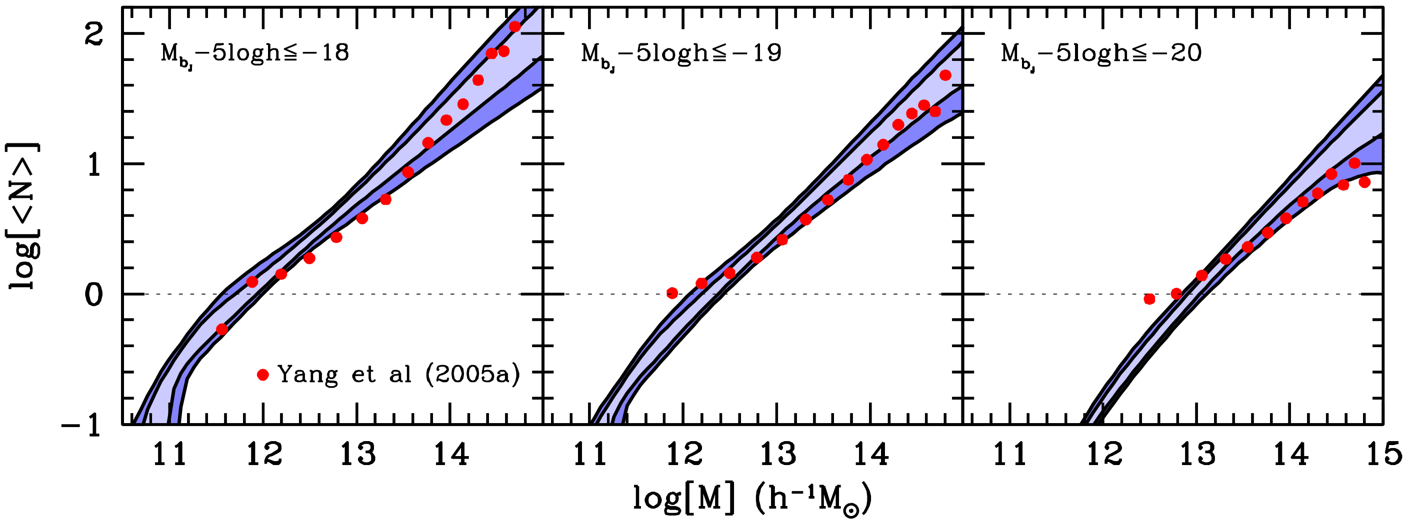
Halo Occupation Numbers



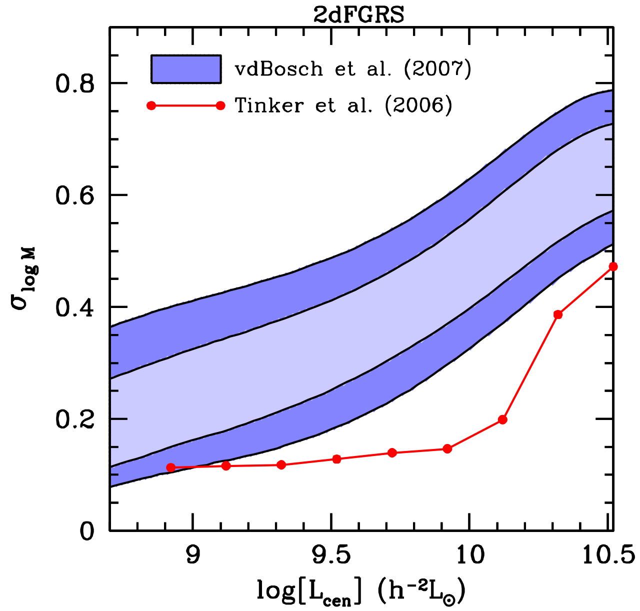
- Unlike **2dFGRS**, the **SDSS** reveals clear shoulders at $\langle N\rangle_M = 1$
- Most likely this is an ‘artefact’ of the functional form of the **CLF**

HOD results for 2dFGRS

Comparison of 2dFGRS HOD from CLF and from Group Catalogue:

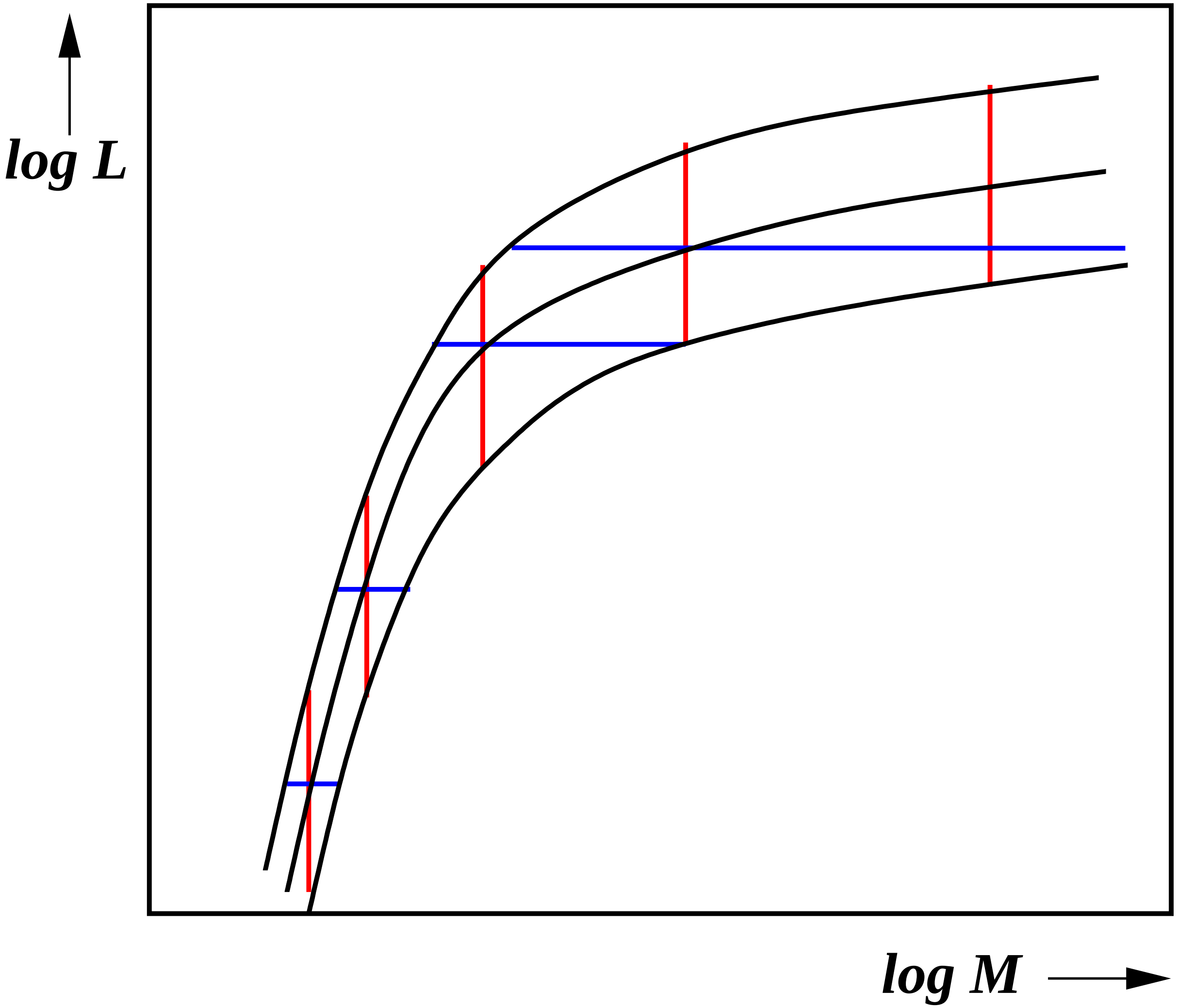


The transition $\langle N\rangle_M = 0 \rightarrow 1$ is 'smooth', indicating that $P(M|L_{\rm cen})$ is broad.



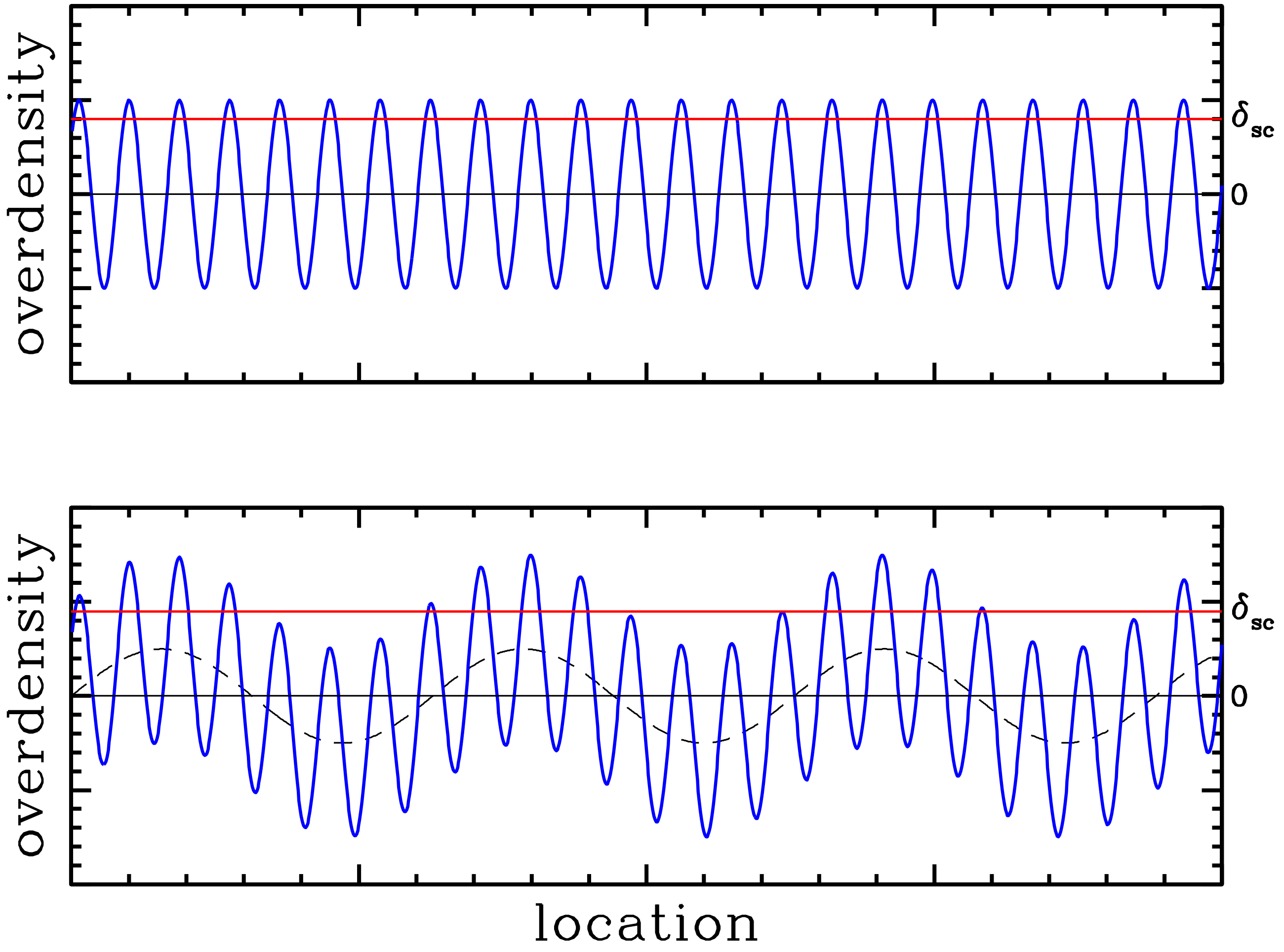
Very different from most HOD models, which often model transition as step-function, which corresponds to $P(M|L_{\rm cen}) = \delta[M - \langle M\rangle(L_{\rm cen})]$

The origin of $\sigma_{\log M}(L)$



- The scatter in $P(L_{\rm cen}|M)$ is roughly independent of M
- The scatter in $P(M|L_{\rm cen})$ increases strongly with $L_{\rm cen}$

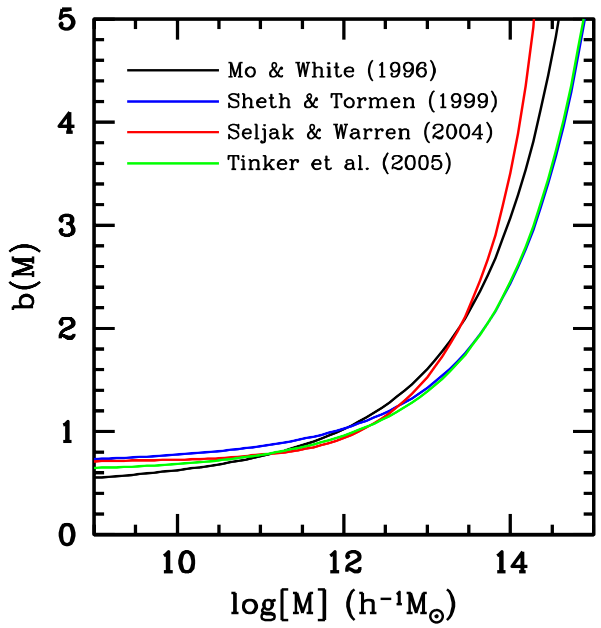
The Origin of Halo Bias



Modulation causes statistical bias of peaks (haloes)

Modulation growth causes dynamical enhancement of bias

Analytical Description of Halo Bias



Define halo bias as $\boxed{b(m) = \langle \delta_h(m)/\delta \rangle}$

Then the halo-halo correlation function for haloes of mass m can be written as

$$\boxed{\xi_{\rm hh}(r) \equiv \langle \delta_{h_1}\delta_{h_2} \rangle = b^2(m)\xi(r)}$$

More massive dark matter haloes are more strongly clustered

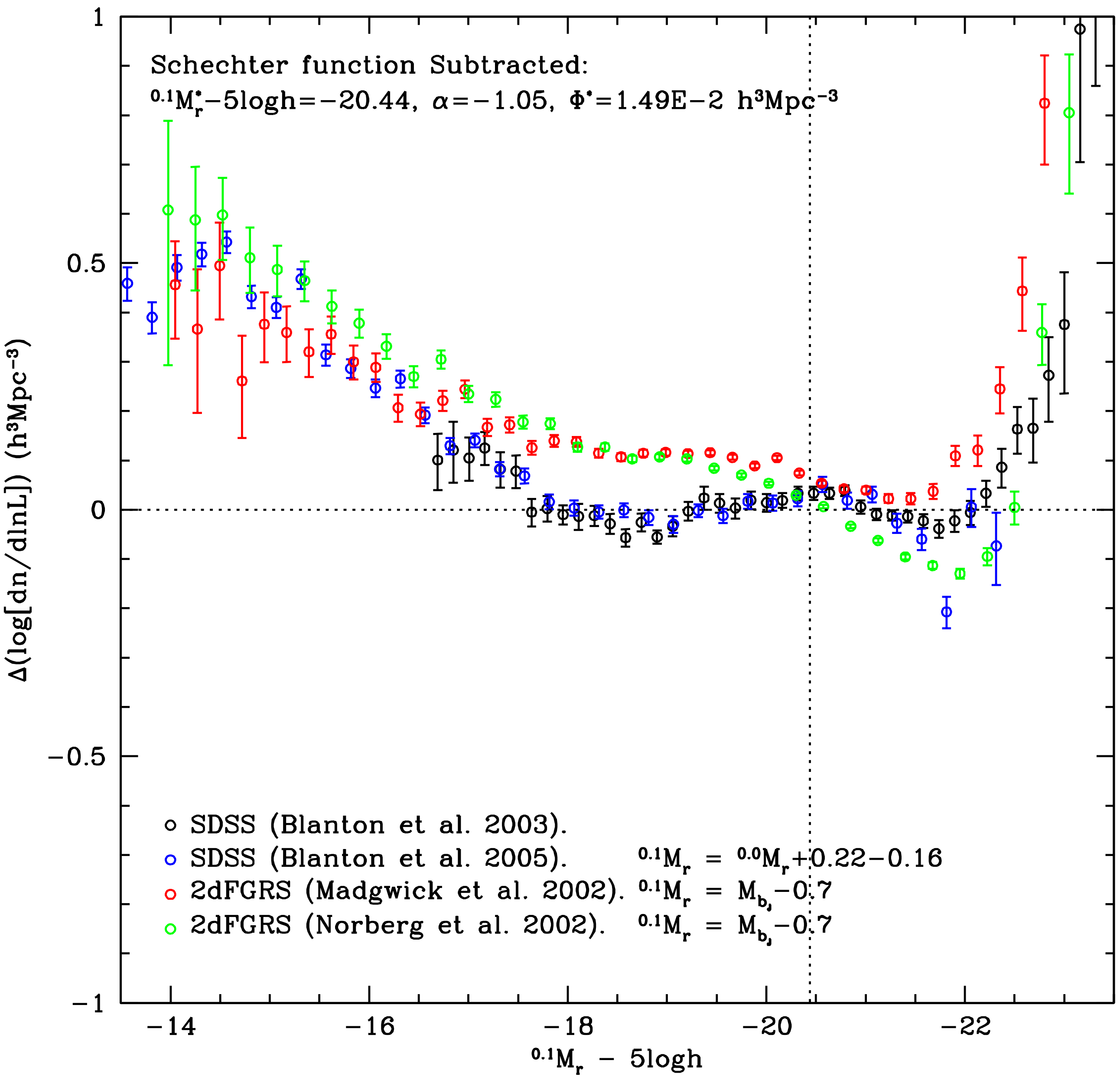
Clustering strength of galaxies is a measure of the mass of the haloes in which they reside

Halo Occupation Statistics completely specifies Halo Bias

Halo Occupation Statistics also constrain Galaxy Formation

Comparison of Luminosity Functions

Introduction
Galaxy Bias
Conditional Luminosity Function
Ecology
Stochasticity
Conclusions
Extra Material
- Halo Occupation Numbers
- HOD results for 2dFGRS
- The origin of $\sigma_{\log M}(L)$
- The Origin of Halo Bias
- Analytical Description of Halo Bias
- Comparison of Luminosity Functions
- The Relation between Light & Mass
- Cosmological Constraints
- Cosmological Parameters
- Large Scale Structure: Theory
- Large Scale Structure: The 2dFGRS
- Constructing Mock Surveys
- Mock versus 2dFGRS
- Mock versus 2dFGRS
- Galactic Conformity



The Relation between Light & Mass

Introduction

Galaxy Bias

Conditional Luminosity Function

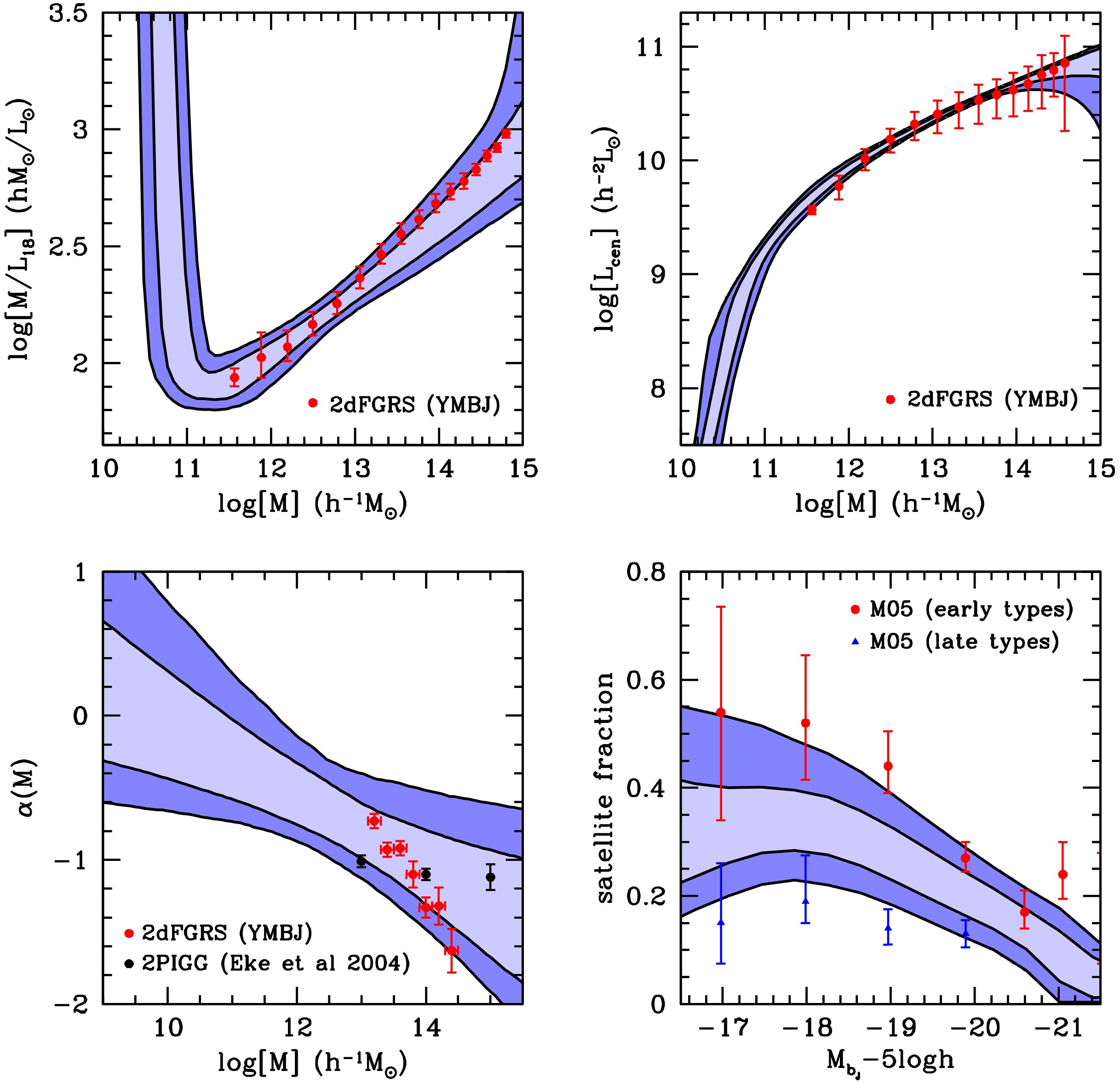
Ecology

Stochasticity

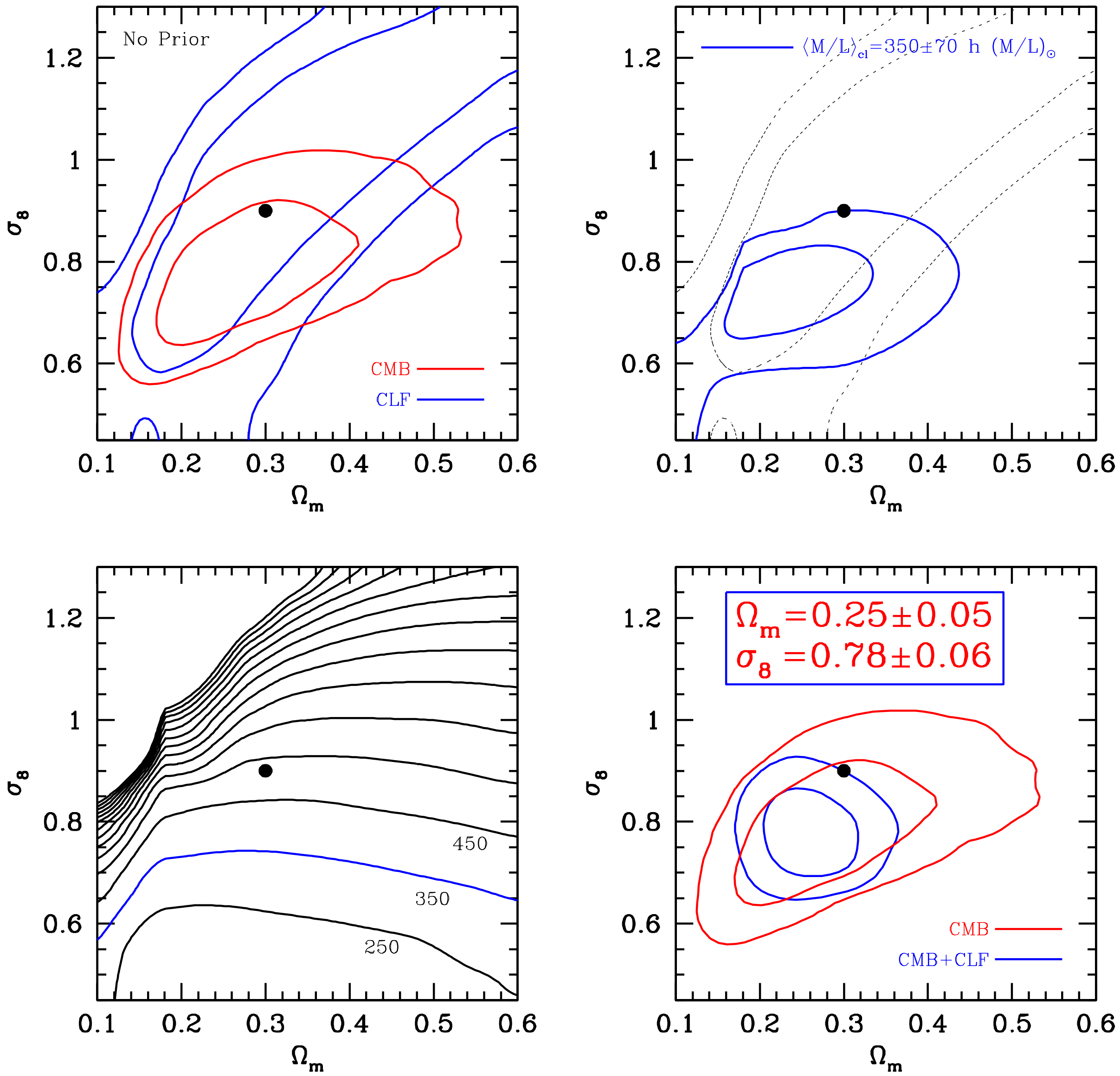
Conclusions

Extra Material

- Halo Occupation Numbers
- HOD results for 2dFGRS
- The origin of $\sigma_{\log M}(L)$
- The Origin of Halo Bias
- Analytical Description of Halo Bias
- Comparison of Luminosity Functions
- The Relation between Light & Mass
- Cosmological Constraints
- Cosmological Parameters
- Large Scale Structure: Theory
- Large Scale Structure: The 2dFGRS
- Constructing Mock Surveys
- Mock versus 2dFGRS
- Mock versus 2dFGRS
- Galactic Conformity



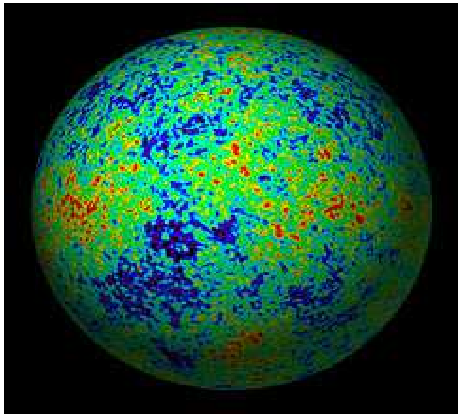
Cosmological Constraints



vdB, Mo & Yang, 2003, MNRAS, 345, 923

Cosmological Parameters

The Cosmic Microwave Background and Supernova Ia have given us precise measurements of most cosmological parameters



$\Omega_m = 0.27$

$\Omega_\Lambda = 0.73$

$\Omega_b = 0.04$

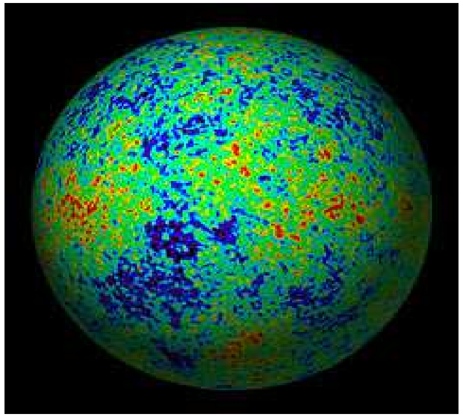
$H_0 = 72$ km/s/Mpc

$n_s = 0.95$

$\sigma_8 = 0.77$

Cosmological Parameters

The Cosmic Microwave Background and Supernova Ia have given us precise measurements of most cosmological parameters



$\Omega_m = 0.27$
$\Omega_\Lambda = 0.73$
$\Omega_b = 0.04$
$H_0 = 72$ km/s/Mpc
$n_s = 0.95$
$\sigma_8 = 0.77$

Open Questions:

- **What is the nature of dark matter; i.e., CDM *vs.* WDM?**
- **What is the nature of dark energy i.e., what is $w = P/\rho$?**
- **What is the mass of neutrinos; i.e., what is Ω_ν?**
- **What are the properties of the inflaton; i.e., what is $V(\phi)$?**

All these fundamental questions can be addressed by probing the matter perturbation field as function of redshift.

Large Scale Structure: Theory

Galaxy redshift surveys yield $\xi(r_p,\pi)$ with r_p and π the pair separations perpendicular and parallel to the line-of-sight.

redshift space CF: $\xi(s)$ **with** $s = \sqrt{r_p^2 + \pi^2}$

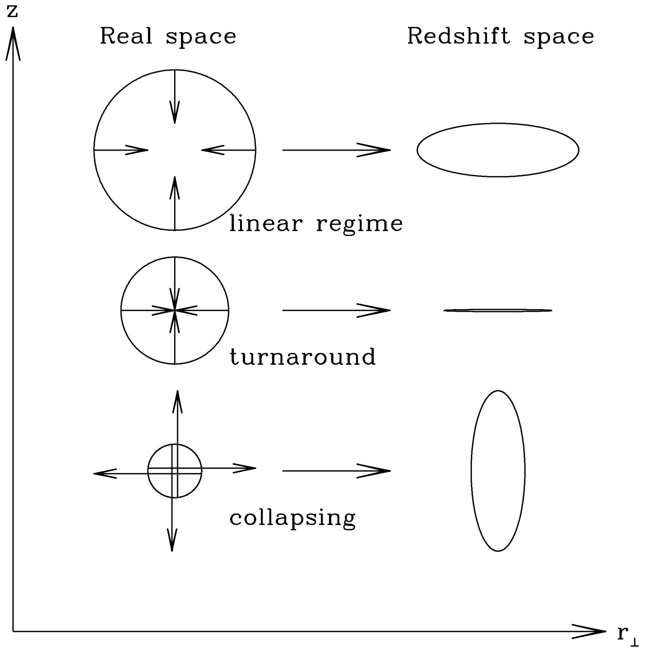
projected CF: $w_p(r_p) = \int\limits_{-\infty}^{\infty} \xi(r_p,\pi)\mathrm{d}\pi = 2\int\limits_{r_p}^{\infty} \xi(r)\,\frac{r\,\mathrm{d}r}{\sqrt{r^2-r_p^2}}$

Peculiar velocities cause $\xi(r_p,\pi)$ to be anisotropic.

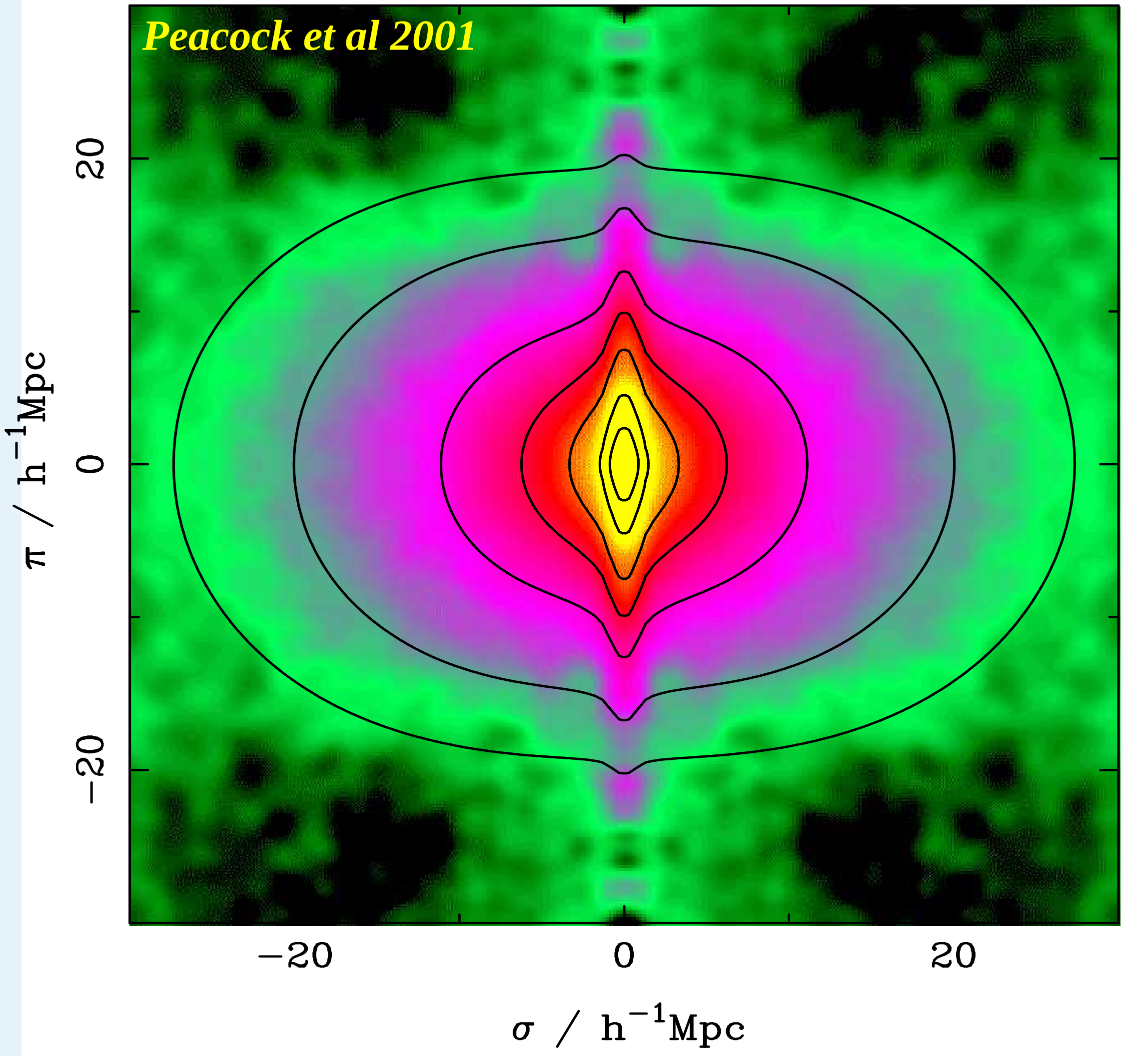
Consequently, $\xi(s) \neq \xi(r)$.

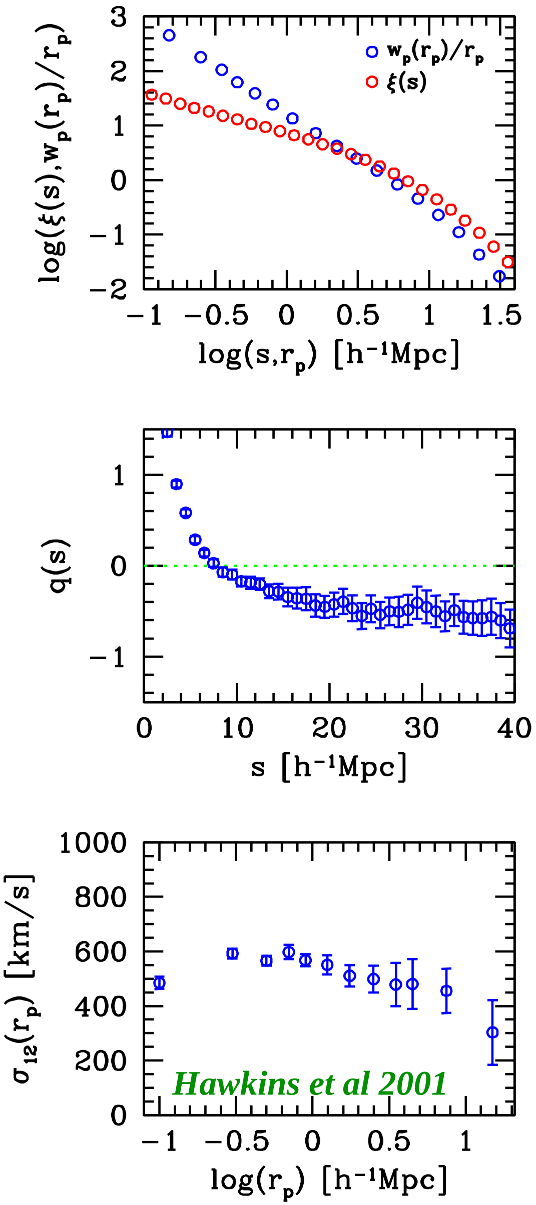
In particular, there are two effects:

- **Large Scales:** **Infall (“Kaiser Effect”)**
- **Small Scales:** **“Finger-of-God-effect”**



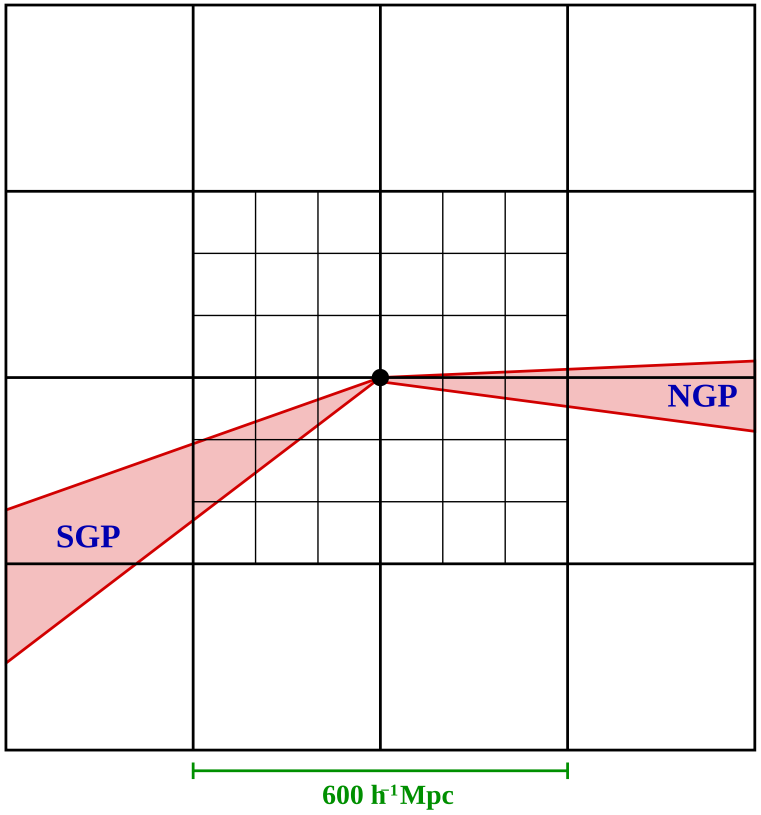
Large Scale Structure: The 2dFGRS





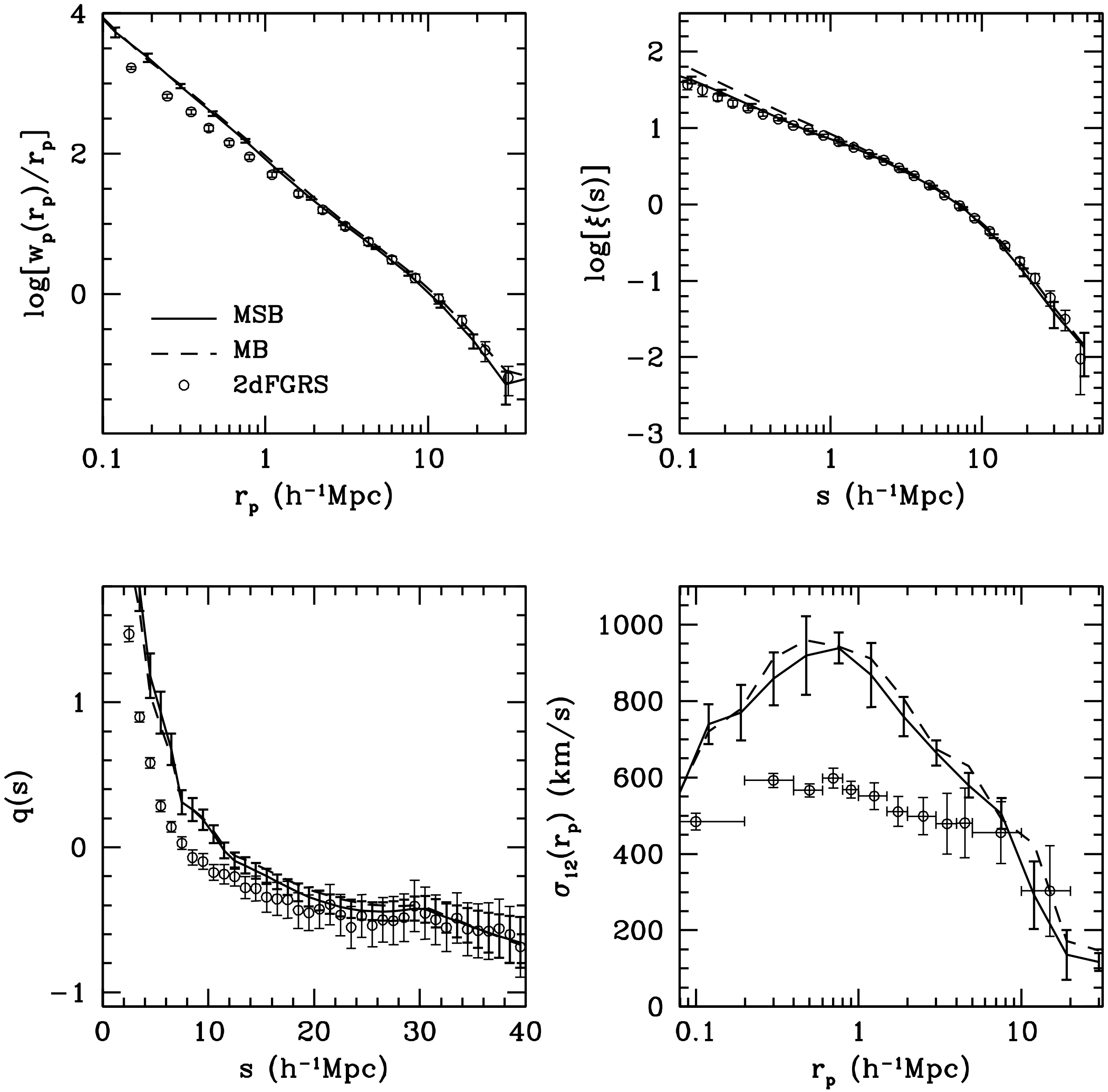
Constructing Mock Surveys

- **Run numerical simulations: ΛCDM concordance cosmology $L_{\rm box} = 100 h^{-1}$ Mpc and $300 h^{-1}$ Mpc with 512^3 CDM particles each.**
- **Identify dark matter haloes with FOF algorithm.**
- **Populate haloes with galaxies using CLF.**
- **Stack boxes to create virtual universe and mimick observations (magnitude limit, completeness, geometry, fiber collisions)**



Mock versus 2dFGRS

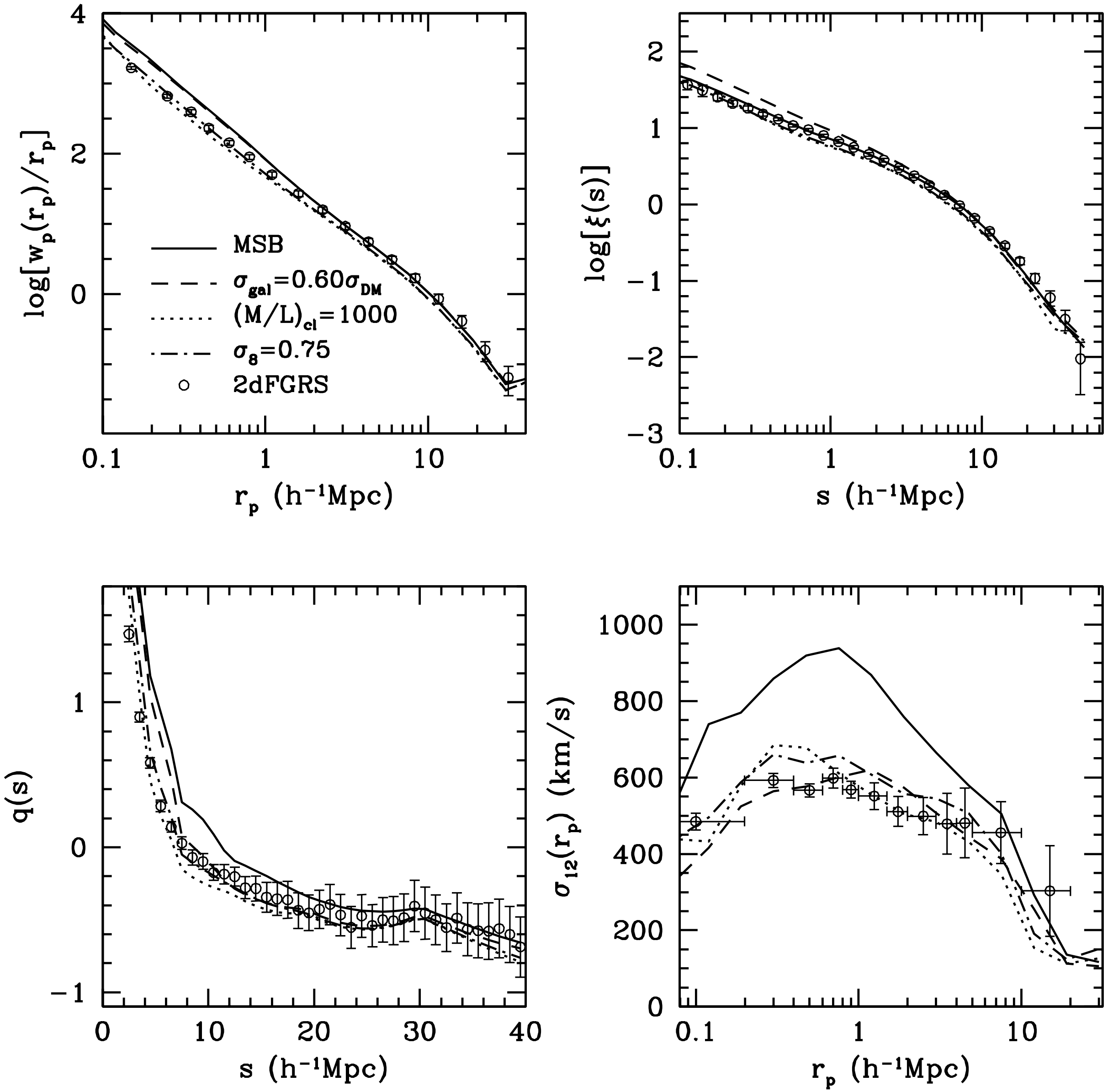
Introduction
Galaxy Bias
Conditional Luminosity Function
Ecology
Stochasticity
Conclusions
Extra Material
- Halo Occupation Numbers
- HOD results for 2dFGRS
- The origin of $\sigma_{\log M}(L)$
- The Origin of Halo Bias
- Analytical Description of Halo Bias
- Comparison of Luminosity Functions
- The Relation between Light & Mass
- Cosmological Constraints
- Cosmological Parameters
- Large Scale Structure: Theory
- Large Scale Structure: The 2dFGRS
- Constructing Mock Surveys
- Mock versus 2dFGRS
- Mock versus 2dFGRS
- Galactic Conformity



Yang, Mo, Jing, vdB & Chu, 2004, MNRAS, 350, 1153

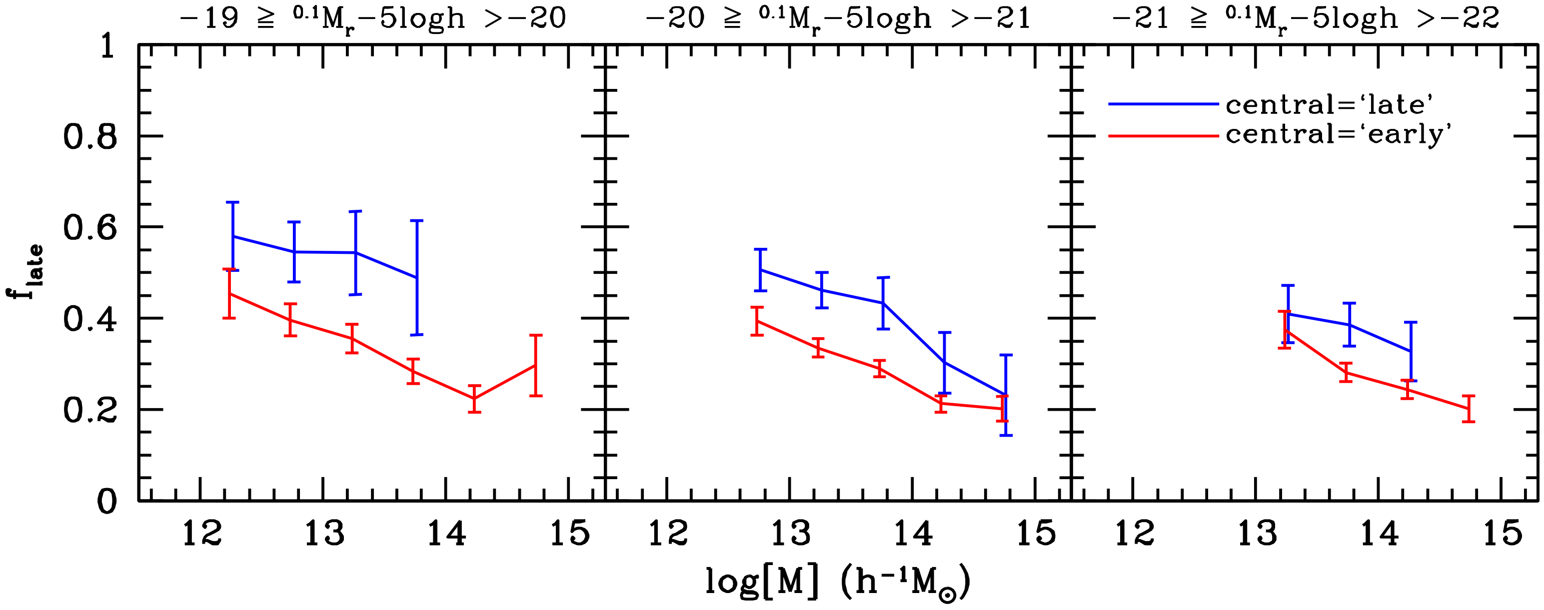
Mock versus 2dFGRS

Introduction
Galaxy Bias
Conditional Luminosity Function
Ecology
Stochasticity
Conclusions
Extra Material
● Halo Occupation Numbers
● HOD results for 2dFGRS
● The origin of $\sigma_{\log M}(L)$
● The Origin of Halo Bias
● Analytical Description of Halo Bias
● Comparison of Luminosity Functions
● The Relation between Light & Mass
● Cosmological Constraints
● Cosmological Parameters
● Large Scale Structure: Theory
● Large Scale Structure: The 2dFGRS
● Constructing Mock Surveys
● Mock versus 2dFGRS
● Mock versus 2dFGRS
● Galactic Conformity



Yang, Mo, Jing, vdB & Chu, 2004, MNRAS, 350, 1153

Galactic Conformity



Late type 'centrals' have preferentially **late type** satellites, and vice versa.

Satellite galaxies 'adjust' themselves to properties of their central galaxy

Galactic Conformity present over large ranges in luminosity and halo mass.

(Weinmann, vdB, Yang & Mo, 2006)

Large-Scale Environment Dependence

Introduction

Galaxy Bias

Conditional Luminosity Function

Ecology

Stochasticity

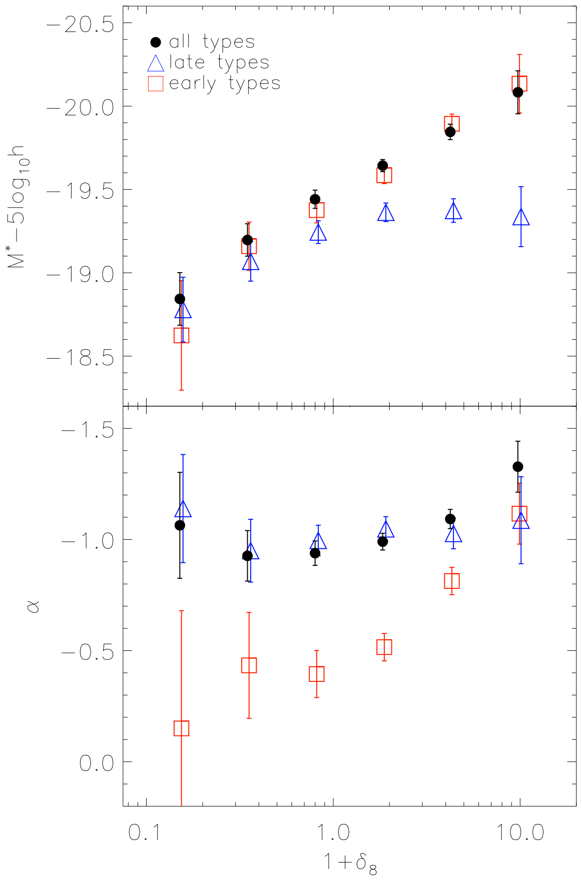
Conclusions

Extra Material

- Halo Occupation Numbers
- HOD results for 2dFGRS
- The origin of $\sigma_{\log M}(L)$
- The Origin of Halo Bias
- Analytical Description of Halo Bias
- Comparison of Luminosity Functions
- The Relation between Light & Mass
- Cosmological Constraints
- Cosmological Parameters
- Large Scale Structure: Theory
- Large Scale Structure: The 2dFGRS
- Constructing Mock Surveys
- Mock versus 2dFGRS
- Mock versus 2dFGRS
- Galactic Conformity

Inherent to CLF formalism is assumption that L depends only on M.

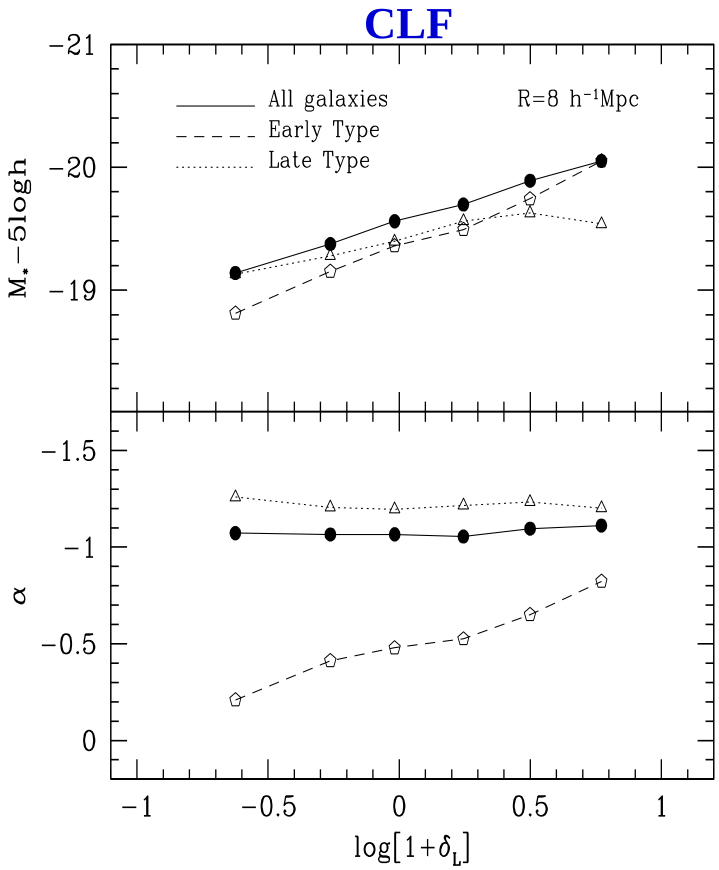
But $\Phi(L)$ has been shown to depend on large scale environment.

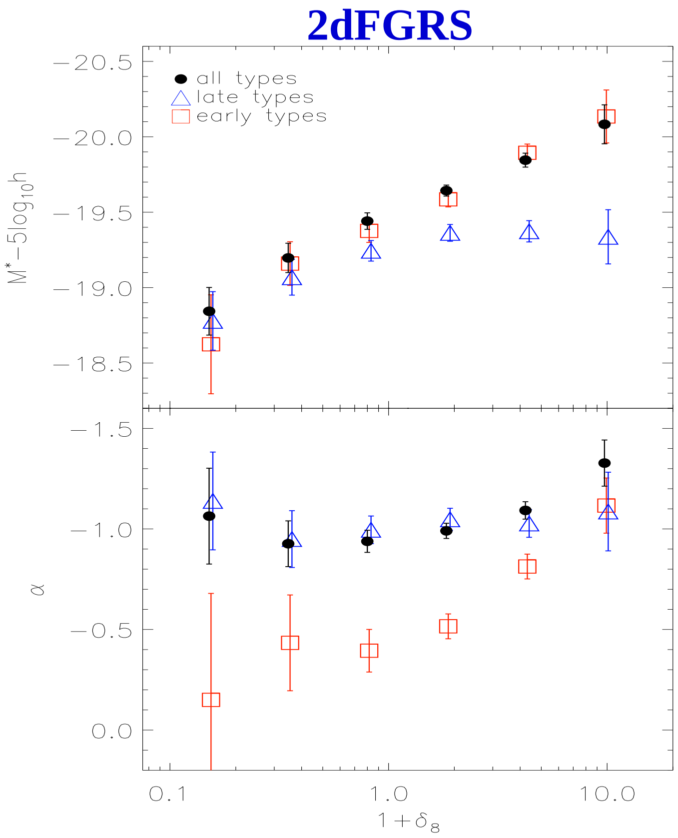


Croton et al. 2005

Does this violate implicit assumptions of CLF formalism?

Large-Scale Environment Dependence





Populate haloes in N-body simulations with galaxies using $\Phi(L|M)$

Compute $\Phi(L)$ as function of environment and type.

Since $n(M) = n(M|\delta)$, we reproduce observed trend

There is no environment dependence, only halo-mass dependence