Halo Occupation Distribution
Modeling of AGN Clustering:
An Overview

Takamitsu Miyaji

Instituto de Astronomía
Universidad Nacional Autónoma de México
Campus Ensenada
Ensenada, Baja California, Mexico

(Also Visiting Scholar UC San Diego)
Introduction

Why AGN Clustering is important for understanding the accretion history/mechanism?

- A small fraction of galaxies show AGN activity (It's an event!)
- When and where, with what mechanism the SMBH accretion occurs? Statistical properties of AGNs give observational clues.
  - Merger driven?
  - Secular Evolution/Internal to galaxy?

Observational Clues

- Luminosity functions and its cosmological evolution (AGN downsizing)
- Host galaxy properties (green valley, Merger remnant features).
- Black Hole Demography.
- Environment and underlying large scale structure (clustering).
Latest determination of X-ray Luminosity Luminosity Function/Evolution

Two redshift break structure revealed.

AGN Downsizing is still strong!

TM+ to be submitted soon
Comparison with Semi-analytical models

TM et al. 2014 in prep.

Semi-analytical models of Marulli+2005; Fanidakis+2012 overpredicts number densities of high luminosity AGNs.
Where in the Cosmic Web do AGNs occupy (accretion occurs)?

- **Observers** see the universe as galaxies, AGNs, clusters etc..
- **Theorists** see the universe as a bunch of Dark Matter Halos (DMH)--(Maybe an outdated comment!)
- How can we relate these halos with observed objects?
What are the DMHs?

\[ \delta = \frac{\Delta \rho}{\langle \rho \rangle} \]

Initial Density Fluctuation

\[ \delta = \delta_c \approx 1.69 \]

Detach from the expansion of the universe,

Collapse and Virialize

\[ \delta = \delta_c \approx 1.69 \]

Cosmic Time

comoving coordinate

Dark Matter Halos: The collapsed & Virialized structures
Two-point Correlation Function

Joint probability $\mathcal{P}$ of finding an object in both of the volume elements separated by $r$ is represented by:

$$3D: \mathcal{P} = n^2 [1 + \xi(r)] \delta V_1 \delta V_2$$

$$\xi(r) = 0 \text{ if objects are randomly distributed}$$

In the linear biasing scheme, the two point 3-D auto correlation function (ACF) is related to bias parameter by:

$$\xi_{\text{obj}}(r) = b_{\text{obj}}^2 \xi_{\text{mass}}(r),$$

and two point 3-D cross-correlation function (CCF) between catalog 1 and 2 is related to the bias parameters of 1 & 2 by:

$$\xi_{12}(r) = b_2 b_1 \xi_{\text{mass}}(r)$$
The large-scale bias of dark matter halos depends on its mass.

Here the “Halo mass” means the largest Virialized structure the object in question belongs to, and NOT represents the sub-halo mass.

Measurements of bias of a sample of AGNs is an indicator of the “typical” mass of the DMHs that the sample is associated.

This simple relation is only valid in the linear regime ($r > \sim 1-2$ h$^{-1}$ Mpc)

$$b_h(\nu) = 1 + \frac{1}{\sqrt{a\delta_c}} \left[ \sqrt{a(\nu^2)} + \sqrt{ab(\nu^2)^{1-c}} - \frac{(av^2)^c}{(av^2)^c + b(1-c)(1-c/2)} \right]$$
Modeling of the linear and non-linear regimes with Halo Occupation Distribution (HOD)

\[ \xi(r) = [1 + \xi_{1h}(r)] + \xi_{2h}(r) \]

- Model the correlation function as the sum of the contributions from pairs:
  - within the same DMHs
  - from different DMHs.
The HOD modeling is very popular in interpreting galaxy clustering

- Tinker+2005; 2010; van den Bosch+13 for recent theory

**Application to AGN 2P Correlation Functions**

- Padmanabahn+2009
  - SDSS LRG vs optically selected QSOs CCF; satellite fraction >25%
- TM, Krumpe, Coil, Aceves 2011
  - SDSS LRG vs X-ray selected AGNs CCF from ROSAT All-sky survey
- Starikova+2011
  - Chandra Boötes field. Consider both $r_p$ and $\pi$ directions. Strict upper limit on satellite fraction (<~0.1).
- Kayo & Oguri+2012 (previous talk)
- Richardson+2012, 2013
  - SDSS QSOs and Allevato+11 XMM-COSMOS ACFs.
  - Full galaxy HOD-type parameterization+MCMC parameter search.
Construction of HOD models

- In the power spectrum space $P(k)$.

Generated with “camb” (http://camb.info/)
Large Scales (approx 2-h term)

- **Matter (linear) power spectrum:** \( P_{\text{matter,lin}}(k,z) \rightarrow \xi_{\text{matter,lin}}(r,z) \)
  
  \[ P_{\text{lin,matter}}(k,z) = D(z)P_{\text{lin,matter}}(k,z=0); \quad D(z), \text{linear growth factor} \]

- **Linear biasing (i.e. Scale independent) at large scales**
  
  \( P_{\text{lin,sample}}(k,z) = b_{\text{sample}}^2 P_{\text{lin, matter}}(k,z) \)

  \( \xi_{\text{lin, sample}}(k,z) = b_{\text{sample}}^2 \xi_{\text{lin, matter}}(k,z) \)

- **DMH bias** \( b(M_h,z) \) (e.g. Sheth, Mo, Tormen '01; Tinker+'05,'10)

- **DMH mass function** \( \phi(M_h) \) (e.g. Sheth & Tormen '99; Jenkins et al. 2001; Tinker+'05)

- **\( <N(M_h)> \): Halo Occupation Distribution (HOD)**
  
  - Mean number of sample objects per DMH as a function of \( M_h \).

  - The sample bias \( b_{\text{sample}} \) is the weighted mean \( b(M_h,z) \) over DMHs

  \[ b_{\text{sample}} = \int b(M_h) <N(M_h)> \phi(M_h) dM_h / \int <N(M_h)> \phi(M_h) dM_h \]
Small Scales (1-halo term)

- \( <N(M_h)> = <N_c(M_h)> + <N_s(M_h)> \)
  - \( <N_c>(M_h) \) for the objects occupying at the **center** of the host DMH.
  - \( <N_s>(M_h) \) for “satellites”, occupying non-center location of the host DMH.

- **Assume that** the mean radial distribution of “satellite” objects follows the mass profile of the DMH (e.g. Navarro, Frenk & White [NFW] profile).

- Contribution of the same DMH pairs to \([1+\xi_{1h}(r)]\).
  - Central-satellite pairs follow the DMH mass profile
  - Satellite-satellite pairs follows the DMH mass profile convolved by itself.
  - Central-central pairs: No such pairs.
HOD Analysis of Galaxies

Example of Luminous Red Galaxies (Zheng+2009)

\( <N_c(M_h)> \) center: smoothed step function saturated to 1.
\( <N_s(M_h)> \) satellite: power-law*\( <N_c(M_h)> \) or spline
Application to SDSS Luminous Red Galaxies (LRGs) vs RASS AGNs

TM, Krumpe, Coil, Aceves (2011)

- **Galaxy Sample**
  - SDSS LRG Volume Limited Sample
  - Defined by Eisenstein et al. (2001), redrawn by us for DR4+;
    \( M_B < -21.2, 0.16 < z < 0.36 \)
  - 45899 LRGs Galaxies

- **X-ray AGN sample:**
  - ROSAT All-Sky Survey (RASS) sources matched with the SDSS broad-line AGNs (Anderson et al. 2003; 2007).
    - 1552 AGNs in 0.16 < z < 0.36
  - Excluded Narrow-line AGNs.
  - Flux limited sample.

These two samples are completely separate. No common object.
HOD of LRGs as our Tracer Set

Zheng+ 09, adjusted.

DMH
Applying HOD modeling to the AGN-LRG CCF

When modeling our CCF, we consider four HODs

• \(<N_{LRG,c}> (M_h) \) & \(<N_{LRG,s}> (M_h) \) for the central and satellite LRGs respectively.

• \(<N_{A,c}> (M_h) \) & \(<N_{A,s}> (M_h) \) and for the AGNs.

• First, we derive \(<N_{LRG,c}> (M_h) \) and \(<N_{LRG,s}> (M_h) \) using the ACF of the LRGs.
  - They can be determined with a much better statistics.

• Then, using the resulting (fixed) LRG HODS, we constrain \(<N_{A,c}> (M_h) \) & \(<N_{A,s}> (M_h) \) by fitting to the AGN-LRG CCF.
Model A: Simple model

Assumption: All AGNs that reside in halos containing LRGs (or contributing to the 1-h term) are satellites.

The 1-halo term is from AGN-LRG pairs in the same DMH.
- LRGs are in $M_h > \sim 10^{13.5} M_{\odot}$ halos.
- The 1-halo term measures AGNs in $M_h > \sim 10^{13.5} M_{\odot}$ halos.

The 2-halo term $\propto b_A b_{\text{LRG}}$.
- Determines AGN bias $b_A$
- Indicates the mean DMH mass with AGNs.
Constraints on HODs for AGNs

Simple HOD model for AGNs

Constraints roughly along $<M_h>\sim$const.

- Constraint from the 2-halo term ($b_X$)
- $\alpha<0.4$ ($\Delta\chi^2<2.3$ limit)

- Constraint from the 1-halo term

- Confidence contours (black, $\Delta\chi^2=1;2.3;4.6$)
- Mean DMH mass (green contours).
Left: number per halo.

Right: Number density

Three possible HODs within errors.

TM+2011
Model with separate central+satellite AGNs

Model B:
A model with galaxy-like central+satellite components
cf. SDSS Galaxies (e.g. Zehavi et al. 2005)
$M_1/M_{\text{min}} \approx 23$, $\alpha \approx 1.2$
Implication of the HOD Analysis

The limit on $\alpha_s<1$ means that the number of (satellite) AGNs/Halo grows slower than $M_h$.

- The HOD of satellite galaxies show $\alpha \sim 1$, i.e., number/halo $\mu M_h$ (e.g. Zehavi et al. 2010).
- AGN fraction (non-center) decreases with $M_h$.
- Long-suggested anti-correlation of emission-line AGN fraction and cluster richness (e.g. Gisler 1978; Dressler et al. 1985).
- Consistent with: AGN fraction anti-correlates with the velocity dispersion of clusters/groups (Popesso & Biviano 2006).
- X-ray AGN fraction is smaller in clusters ($M_h > 10^{14} M_{\text{sol}}$) than the field at low z. Higher at high z ($z \sim 1.5$), this trend reverses (Martini+13).
Trend Verified in direct counts/Weak lensing-based HOD studies

Allevato+12, direct count
Satellite HOD slope $\alpha_s < 0.63$

Leauthaud+2014, A. Coil's Talk
Implications -cont'd

- Possible mechanisms:
  - Merging efficiency low in high velocity encounters (Makino & Hut 1997).
  - Would AGN triggering by major merger/minor merger of sub-halos inside larger host halos explain the HOD behavior (Altamirano's talk)?
  - Ram pressure stripping/thermalevaporation of cold gas in galaxies in Intracluster/intragroup medium (Gunn & Gott 1972; Cowie & Songaila 1977).
Extended sample

RASS-AGNs extended

Schneider et al. 2010, Optically-selected Broad-line AGN sample

Paper III: Krumpe, TM, Coil Aceves 2012
HOD approach may be simply used for more accurate determination of linear bias parameters

- More accurate determination of $b_{\text{lin}}$ than power-law fits.
- Fitting 2-halo term only to $r_p > 1.5h^{-1}\text{Mpc}$ (Allevato et al. 2011, 2012)
- Fitting 1 and 2-halo terms with a simple parameterized HOD model to obtain constraints on $b$ and log $<Mh>$ (Krumpe, TM et al. 2012, 2014).

Figs: Krumpe, TM et al. (2012)
Highlight differences between $L_x$, $M_{BH}$, & $L/L_{edd}$ divided samples

High vs low $L_x$  
High vs low $M_{BH}$  
High vs low $L/L_{edd}$

Comparing biases only use data at 2-h terms  
This approach takes advantage of data at all scales.

M. Krumpe's talk yesterday
Two-halo term improvements

- Instead of simple linear PS, use non-linear PS, scale-dependent bias, and exclusion of pairs that should be counted in the 2-halo term (Zheng+'04; Tinker+'05; van den Bosch+'13)
Limitations

- Good sampling at 2-halo term ($r_p > \sim 1$ Mpc)
  - Good constraint on only one parameter: linear bias
- Poor sampling at 1-halo term ($r_p < \sim 1$ Mpc)
  - Poor constraint on the distribution of $N(M_h)$
  - CCF approach helps
- Degeneracy in the interpretation of the 1-halo term.
  - Central vs satellite pair or satellite-satellite pair?
- Do satellite AGNs follow DM profile?
  - The same problem with galaxy HOD studies, especially comparing blue vs red galaxy HODs.
Direct counts within resolved groups/clusters?

- QSO counts within rich clusters of galaxies (Martini et al. 2009; 2013)
- DMHs with $M_h \geq 13 \, h^{-1} M_\odot$ can be cataloged as groups/clusters (e.g. X-ray selected).
  - Direct counts of AGNs in these groups/clusters are possible.
  - Combine with the CFs involving AGNs that do not belong to these groups/clusters give constraints on the minimum halo mass occupied by these HODs -> Allevato+12, *(Talk by A. Finoguenov)*

- SDSS QSOs in clusters ($M_h \geq 14 \, h^{-1} M_\odot$)
  *(Talk by M. Nguyen)*
Conclusions

- The HOD analysis is a strong tool to interpret correlation functions of galaxies/AGNs to scale over linear to non-linear scales.
- From HOD analysis, we can obtain not only a single “typical” host DMH mass but also constraints on how AGNs distribute among DMHs as a function of mass.
- Applying the HOD analysis to z~0.3 SDSS LRG vs RASS AGNs, we find that solutions where AGN fraction among satellite galaxies decrease with Halo mass.
- The interpretation of the HOD analysis is limited by poor sampling at small scales (especially of AGNs) and model degeneracies.
- If we have good catalog of resolved clusters/groups, direct count of AGNs in these clusters/groups can give robust HOD measures.