Connecting galaxies and dark matter halos: methods, insights, and challenges

Risa Wechsler
KIPAC @ Stanford/SLAC

KICC 10th Anniversary Symposium
September 17, 2019
What are models of the galaxy-halo connection good for?

- Parameterize the galaxy-halo connection; marginalize over it to constrain cosmological parameters
- Model the galaxy-halo connection to make realistic realizations of large surveys
- Constrain the parameters (and parameterization) of a galaxy-halo model to understand galaxy formation physics
- Parameterize the galaxy-halo connection; marginalize over it to constrain dark matter physics
Approaches to modeling the galaxy-halo connection
Approaches to modeling the galaxy–halo connection

<table>
<thead>
<tr>
<th>Physical models</th>
<th>Empirical models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrodynamical simulations</td>
<td>Simulate halos and gas; star formation and feedback recipes</td>
</tr>
<tr>
<td>Semianalytic models</td>
<td>Evolution of density peaks plus recipes for gas cooling, star formation, feedback</td>
</tr>
<tr>
<td>Empirical forward modeling</td>
<td>Evolution of density peaks plus parameterized star formation rates</td>
</tr>
<tr>
<td>Subhalo abundance modeling</td>
<td>Density peaks (halos and subhalos) plus assumptions about galaxy–(sub)halo connection</td>
</tr>
<tr>
<td>Halo occupation models</td>
<td>Collapsed objects (halos) plus model for distribution of galaxy number given host halo properties</td>
</tr>
</tbody>
</table>
Subhalo abundance matching

Recent models parameterize scatter and concentration dependence of galaxy properties
Constraining the halo occupation with galaxy clustering

Can also use group & cluster statistics, galaxy-galaxy lensing, velocities in groups & clusters, counts in cells, void statistics...

Zehavi et al 2011
Combine observations with halo statistics and growth

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Choose a SFR - (v_{\text{max}}) (SFRVM) relation from parameter space.</td>
</tr>
<tr>
<td>2.</td>
<td>This gives the probability distribution of SFRs for any chosen halo (v_{\text{max}}) and redshift.</td>
</tr>
<tr>
<td>3.</td>
<td>Map these SFR distributions to halos such that earlier-forming halos get lower SFRs.</td>
</tr>
<tr>
<td>4.</td>
<td>Integrate SFRs along halo merger trees to infer galaxy growth.</td>
</tr>
<tr>
<td>5.</td>
<td>Predict observables: SMFs, correl. functions, lensing, etc. for SF, quenched galaxies.</td>
</tr>
<tr>
<td>6.</td>
<td>Apply effects to simulate observational errors and biases.</td>
</tr>
<tr>
<td>7.</td>
<td>Compare to data and calculate likelihood of the chosen SFRVM relation.</td>
</tr>
</tbody>
</table>

**Behroozi, RW, Hearin & Conroy 2019**
(see also Becker 2015; Rodriguez-Puebla et al 2016; Chon 2017; Moster et al 2018)
Based on framework first explored in Conroy & Wechsler 2009 parameterized SFR (mass accretion rate) ... constrain with SMF, SFR, clustering.
Hydrodynamical simulations

Best understanding of the physics
Can test key assumptions in parametric models
Expensive; hard to explore parameter space

Chavez-Montaro et al 2016; Eagle simulations
Interplay between physical and empirical modeling approaches AND interplay between high fidelity and inexpensive approaches is essential.

- Current hydro simulations are not accurate, flexible, or fast enough to be used for direct modeling of full galaxy surveys.
- Current empirical models are missing key correlations and some important galaxy properties; need to test them with more physical models.
- Galaxies probably do all live in dark matter halos! But those halos are not always properly resolved in dark matter simulations; can be modified or destroyed by baryons...
Interplay between approaches

- Test simple parameterizations against best hydro simulations and semi-analytic models
- Test sensitive systematics on best models possible
- Train models on high-fidelity small volume simulations and apply to large volume simulations
What have we learned about the galaxy-halo connection?
Galaxy formation is inefficient; feedback is important.

Wechsler & Tinker 2018 Annual Reviews of Astronomy & Astrophysics
Galaxy masses are tightly connected to the masses of their dark matter halos.

Scatter in stellar mass at a given halo mass is small. ~0.2 dex at high mass; probably still < 0.3 dex at small mass...

Scatter is larger in current semi-analytic models than in current hydrodynamical simulations...

RW & Tinker 2018
Most stars in the Universe form in a small range of halo mass, around the mass of the Milky Way

- Efficiency of forming stars given incoming gas is roughly constant over 10 billion years.
- Decline in star formation rate due primarily to decline in mass accretion onto halos

massive galaxies: start forming early, peaked at $z \sim 2-4$, then quenched. Halo growth continues after galaxy growth slows.

low mass galaxies: extended star formation histories, start later and continue longer at a low rate.

Behroozi, Conroy & RW 2014; Behroozi, RW, Hearin & Conroy 2019; RW & Tinker, ARAA 2018
Quenched fraction is a strong function of mass.

fraction of galaxies that are quenched increases rapidly with time at a fixed number density, but is not a strong function of time at fixed halo mass.
Quenched fraction changes with environment and central / satellite identity

But this can mostly be explained by quenching that is just a function of halo mass and how long something has been a satellite...
Halo mass is not enough (Assembly bias)

Wechsler 2001; Gao et al 2005; Wechsler et al 2006; Gao et al 2007; Croton et al 2007; Dalal et al 2008; Zentner et al 2014; Mao et al 2018; Mansfield & Kravtsov 2019 many more...Plot from Wechsler & Tinker 2018;

Fitting for assembly bias is not straightforward, and may increasingly may matter in cosmological modeling...
Still uncertainty in some basics!

At a given halo mass, are red or blue galaxies more massive?

RW & Tinker 2018
Modeling small scales can be very powerful... requires joint fitting of cosmology and halo occupation parameters.
Galaxy-halo connection at the smallest masses
Discoveries in the Milky Way

New generation of optical surveys with wider and deeper imaging:
DES, SMASH, MagLiteS, ATLAS, Pan-STARRS, HSC, Gaia
and others...

Start of SDSS, digitized datasets and algorithmic searches

“Classical” Milky Way Satellites found in photographic plates

Timeline from Bechtol & Drlica-Wagner
New search of nearly full sky using DES Y3 + PanSTARRS

DES Y3 search covers 10% more area than Y2; total sky is 32,476 deg$^2$
Covers 82% (25,327 deg$^2$) of the non-dusty sky.
New search of 3/4 of the sky using DES + PanSTARRS

- Automated search using two independent algorithms
- Majority of statistically significant “hot spots” are concentrated in problematic regions of the survey (e.g., high dust extinction, bright stars and galaxies).
- Apply a geometric mask based on reddening maps and external catalogs.
- Recover vast majority of known satellites with an automated pipeline; no new satellites detected

17/19 known sats in DES footprint
21/31 known sats in PanSTARRS footprint
9 known sats outside these footprints

DES Milky Way Working Group, in preparation
Observational Selection Function

Run two satellite search algorithms on stellar populations injected into survey data

Number of bright stars and surface brightness are the most important properties.

\[ \text{Projected detection probability in physical parameter space} \]

DES Milky Way Working Group, in preparation
1. Resimulate Milky Way-like halos from large cosmological volume.

2. Paint satellite galaxies onto subhalos using galaxy—halo model.

3. Apply observational selection effects based on imaging data.

4. Calculate likelihood of observed satellites given galaxy—halo connection parameters.

Ethan Nadler, RW & the DES MW Working group in prep
Satellite Galaxy-Subhalo Model

Here “galaxy” means anything brighter than $M_v=0$

Nadler, RW & the DES MW Working group in prep
See also Nadler et al 2019 for model details
Feed model through observational selection function

Selection function from DES + PanSTARRS

E. Nadler + DES Milky Way Working Group, in preparation
Presence of the LMC?

- Does the LMC matter for understanding the luminosity of MW satellites and the halos they inhabit?
- Here we consider two samples:
  - “LMC analogs”: two best fits to LMC.
    - $\geq 1$ subhalos with $v_{\text{peak}} > 55 \text{ km/s}$, $30 \text{ kpc} < d < 70 \text{ kpc}$
    - Accreted within 2 Gyr
  - “No LMC”: 4 hosts with same $c$ distribution, no LMC
• Very difficult to explain anisotropy between DES and PS1 without LMC-associated satellites
• Predict 8 +/- 2 (out of 39) observed DES (6) + PS1 (2) satellites to be LMC-associated.
Luminosity Function and Minimum Halo Mass

Faint end slope: consistent with global LF

Galaxy-halo scatter: consistent with massive systems

$M_{50}$, mass at which 50% of halos are occupied: 95% upper limit = $2 \times 10^8$ $M_{\odot}$

Impact of baryons: no impact ruled out, consistent with destruction tuned to FIRE simulations
Galaxy-halo connection

What is the smallest mass that can form a galaxy?

$M_{50}$, mass at which 50% of halos are occupied = few $\times 10^7$ (95% upper limit = 2$x 10^8$ $M_{\odot}$)

$f_{\text{gal}} = \frac{1}{2} \left(1 + \text{erf} \left( \frac{M_{\text{peak}} - M_{50}}{\sqrt{2} \sigma} \right) \right)$
What does this tell us about dark matter?
Dark matter-baryon interaction constrained by factor of 1000 more constraining than previous cosmological analysis, using SDSS + classical satellites only. Preliminary estimate for new analysis: constraints will improve by ~factor of two in cross-section; $m_{\text{WDM}} > 4\text{keV}$
Lots of progress to be made soon...
Our current galaxy map (z~0.3)

DESI Bright Galaxy Survey map, z~0.3
Things you should know about how galaxies and halos are connected:

• Galaxy formation is inefficient: efficiency of turning gas into stars peaks at 20–30% of the baryon fraction in halos of $10^{12} \, M_\odot$, significantly lower for higher- & lower-mass halos.

• Stellar masses of central galaxies are a strong function of DM halo mass below $10^{12} \, M_\odot$, scaling like $M_\star \sim M^{2-3}$, and a weaker function above this pivot point, $M_\star \sim M^{1/3}$.

• The scatter in galaxy stellar mass or luminosity at fixed halo mass (before halos are stripped by a larger system) has a scatter of less than 0.2 dex for objects above $10^{12} \, M_\odot$; this likely increases at lower halo mass but by no more than a factor of two.

• For most of the Universe’s history from $z \sim 8$–0, the bulk of all star formation occurs in galaxies that live in a narrow range of halo mass around $10^{12} \, M_\odot$.

• Most galaxies at any stellar mass are the central galaxies in their own dark matter halo; the fraction of satellite galaxies at a given galaxy stellar mass declines from $\sim 30\%$ at low mass to $\sim 5\%$ at high mass.

• Most trends of galaxy properties with large-scale environment are well explained by the fact that the halo mass function and average halo properties vary with environment, combined with a galaxy–halo connection that is mostly independent of environment.

• Halo mass is the dominant determinant of the state of the galaxy occupying it, but there is statistically significant evidence that some galaxy properties are influenced by other halo properties and that this manifests in galaxy clustering properties (assembly bias).