Progress and Questions in Galaxy Formation

(a biased view)

Risa Wechsler

CCAPP Symposium
Oct 14, 2009

z=0
z=3
z=1
z=7

~milky way
M>3e12M/

~virgo cluster
M>3e14M/

expansion scalefactor

halo merger histories

- extract merger histories from simulations
- get merger histories analytically from "Extended Press-Schechter"

time

mass of most massive progenitor (a biased view)
Big Questions in Galaxy Formation:

- how are dark matter, gas, and stars distributed in the Universe?
- what determines the masses and luminosities of galaxies?
- what sets the star formation rate of galaxies?
  - how is star formation triggered, regulated, and quenched?
  - does the nature of star formation change with time?
- how are galaxies assembled?
  - how does stellar mass build up over time?
  - how do galaxies get their gas?
  - how is galaxy build up related to dark matter halo buildup?
- what determines how galaxies are clustered in space? what is the role of large and small scale environment?
- what determines the sizes and shapes of galaxies?
- how do energetic processes influence galaxies? what is the interplay between galaxies and supermassive black holes?
- when and how did the first galaxies arise, how did they reionize the Universe?
- how are galaxies at different epochs connected to each other?
Galaxy Formation is not separable from Cosmology and Particle Astrophysics

- **Dark Energy**
  - 5 popular probes for measuring dark energy; 3 of them depend on galaxies (also CMB, SN)
  - Galaxy Clusters: requires understanding the mass-observable relation (obs=stars, gas)
  - Weak Lensing: galaxy formation impacts power spectrum, galaxy alignments, photo-z’s
  - Baryon Acoustic Oscillations: requires understanding non-linearities at the <1% level

- **Dark Matter**
  - Galaxy power spectrum is the best constraint on neutrino mass, measurable in the next decade
  - Probes of the small-scale power spectrum: number and profiles of dwarf galaxies
  - Dark Matter in the Milky Way: e.g. dwarf galaxy profiles impact indirect detection, MW shape impacts direct detection formation and astrophysics of the MW impacts particle propagation,
  - Galaxy Formation is **determined by** the growth of structure
  - what insight can we get into galaxy formation by studying the growth of structure (e.g. growth, abundance and clustering of dm halos)?
Cosmology is not separable from galaxy formation, but it’s mostly “settled”...

for the purposes of understanding galaxy formation basics...

except at very high redshift and for the lowest masses

Rozo, RW et al 2009

these are the most important parameters for the abundance and clustering of galaxies.

Mantz et al 2009

XLF

WMAP5
simulations by A. Klypin & B. Allgood

A. Klypin & B. Allgood

simulating the required dynamic range is still very challenging, but predictions from structure formation are now extremely solid*

*at the ~5-10% level.
dark matter halos and substructure

- dm halo schematically: dark matter which has reached virial equilibrium
- dm halo functionally: typically defined by an overdensity delta (e.g. 200) with respect to the background density or the critical density of the Universe.
- “subhalo” or “satellite” halo: a self-bound object within the virial radius of a larger halo
basic statistics of the dark matter distribution are well determined for a given cosmological model

halo abundance

Tinker et al 2008

halo clustering

Tinker et al 2009
hierarchical galaxy formation

in the cold dark matter model, small things form first and merge to become larger systems.

- in collapsed dark matter halos, where the density is high, gas cools and sinks to the center to form a galaxy
- expect a galaxy at the center of each density peak massive enough to form stars

movie by S. Gottloeber

Wechsler et al. 2002

Wechsler et al. 2002
galaxy formation details are very difficult to simulate from first principles

- requires understanding the complex interplay between a variety of uncertain physical processes over a vast array of length and time scales (pc to Gpc)

most popular approaches:

- hydrodynamical simulations: simulate gravity and hydrodynamics, with “sub grid” prescriptions for star formation and feedback. *uncertain unresolved physics, computationally challenging.*

- semi-analytic models: combine knowledge of dark matter evolution with parameterized physical prescriptions. *lots of free parameters, may miss important physics.*

just a couple of issues in current models:

- baryon budget in clusters
  - too much mass in stars in clusters (okay in groups)
  - not enough massive galaxies at high z

- evolution of the stellar mass density
Semi-Empirical Galaxy Formation Modeling

**Theory:** We now have a reasonable accounting for the distribution of mass in the Universe, *and how this mass assembles over time*, from numerical simulations. Now that the parameters of LCDM are pinned down, this part is very solid!

**Observations:** We now have statistically measured galaxy populations at various epochs. Fairly well measured with spectroscopic surveys from $z\sim1$ to the present. With larger error bars, there are constraints back to $z \sim 4$ (even $z\sim7-8$ for the rarest galaxies).

Use these two pieces of information to get empirical constraints on the connection between galaxies and halos at given epochs, and then to infer how galaxies grow and form stars over time.

Despite the complexity of galaxy formation, many basic trends can be understood within a simple framework where galaxy properties are primarily determined by the masses of their dark matter halos.
the SHAM (SubHalo Abundance Matching) a one-parameter galaxy formation model.

- The basic idea:
  - Galaxies live in halos (including subhalos)
  - There is a property of halos that is tightly correlated with some property of galaxies.

- But some details matter:
  - which property, how tightly
  - including satellite galaxies properly

- bottom line: galaxy masses are tightly correlated to the maximum mass its halo ever had. (current mass($v_{\text{max}}$) for central galaxies; mass at accretion for satellites)

Despite its simplicity,

- With a very few tweaks this “galaxy formation model” can describe nearly all known galaxy statistics that we’ve tested.

- It can be extended to get real insight into galaxy evolution and galaxy formation physics, especially when combined with information about halo accretion histories to connect epochs.
assign galaxies to halos by matching \( n(>M \text{ or } v_{\text{max}}) \) to \( n(>M^* \text{ or } L) \)

key assumptions: one galaxy per dark matter clump; galaxy mass/luminosity tightly correlated with halo mass/velocity
required high-res cosmological simulations with substructure to properly test

e.g. Kravtsov, Berlind, Wechsler, et al 2004; Tasitsiomi, Kravtsov, Wechsler & Primack 2005
Distinct Halo Evolution

Subhalo Evolution

Accretion epoch

Wechsler et al 2002 hypothesis: maximum/accretion mass/velocity is tightly correlated with stellar mass/total luminosity
galaxy-galaxy correlation function

**SDSS, z=0**

Data: Zehavi et al 2004

**Subaru, z=4-5**

Data: Ouchi et al 2005

(similar agreement for Lee et al 2006)

The scale & luminosity dependence of galaxy clustering at high z are very well explained by this simple approach!
additional clustering statistics provide further tests

- model makes predictions for a variety of statistics, e.g.:
  1. galaxy-mass correlations (Tasitsiomi et al 2004)
  2. close pair statistics & evolution (Berrier et al 2006)
  3. 3pt statistics (Marin et al 2008)
  4. high redshift galaxy clustering (Lee et al 2009)
  5. properties of galaxies in clusters (RW et al in prep, Vale & Ostriker)

- a simple model with one (well-constrained) parameter can explain almost all tested galaxy clustering statistics.

- Gravity and the dynamical evolution of dark matter halos and subhalos are the primary drivers of galaxy clustering.

- Galaxies have a tight correspondence to their dark matter hosts or host subhalos
Connection between galaxy mass and halo mass

inputs: HMF(z) & GSMF(z)

The peak of integrated star formation efficiency is in ~MW halos today and somewhat more massive halos in the past.

This mass also marks the transition between dominant halo growth and dominant stellar mass growth.

Galaxy formation is never very efficient! ~ 20% in the most efficient halos.

\( M^* \sim M^{2-3} \) at low mass; \( M^* \sim M^{0.3} \) at high mass.

(take dashed lines with a barrel of salt, given uncertainties in GSMF)
Scatter in the galaxy halo connection: constraints

**HOD constraints from 2 point clustering**

*Zheng, Coil & Zehavi 07*

**Constraints from satellite dynamics**

*More et al 09*

\[ \sigma_{\text{log} L} = 0.16 \pm 0.04 \]

**“uber comparison”**

*RW in prep*

**Constraints from BCGs**

*Hansen, Sheldon, RW & Koester 09*
How well is this connection really constrained?

Behroozi, Conroy & RW (to be submitted)

includes poisson and cosmic variance errors, uncertainty due to scatter in M*-M

includes random errors in stellar masses, possible systematic errors in the stellar mass function

uncertainty from cosmology is smaller than uncertainty in stellar mass functions at z=0
Consistency with other studies

alternative ways to measure: dynamics of satellites, lensing --> mean mass ($M^*$); identify clusters, measure stellar mass; internal dynamics; clustering/HOD constraints
Connection between galaxy mass and descendent halo mass

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09

Conroy & Wechsler 09
mass builds-up hierarchically: small halos get more fractional mass earlier than larger mass halos

stellar mass build-up proceeds the other way: stars in massive galaxies in place by $z=1$; still rapidly forming in small galaxies
The star formation rate as a function of stellar mass

- Specific star formation rate declines roughly equally for galaxies of various masses.
- Star formation slows for masses greater than $M \sim$ few $10^{12}$ halos today.
- Peak of SFR($M$) declines slightly with time.

Conroy & Wechsler 09
Typical galaxy trajectories

- Massive halos host more massive galaxies (they spend more time rapidly forming stars).
- Massive halos host older galaxies (they start forming stars earlier).
- Massive halos host higher surface brightness galaxies (they formed more stars when the universe was denser).
- Low mass galaxies always form galaxies slowly.
Star formation slows for masses greater than $M \sim 1 \times 10^{12}$ halos (roughly the scale where galaxy bimodality sets in) today, but this transition isn’t sharp.

~ MW mass halos are most efficient today, and this mass is only a weak function of redshift.

What does this imply about the physics?

- AGN feedback
- shock heating
- inefficient cooling
- SN feedback
- feedback from massive stars
- photoionization
Constraining physical models

(Hutcheson, RW, Somerville in prep)

- Typical procedure in semi-analytic galaxy formation
  - describe physical processes with a set of simple scaling laws, either empirical or theoretical
  - normalize to properties of local galaxies.
  - hope everything else works out.

- Our procedure: explore parameter space!
  (Bayesian Semi-analytic modeling)
  - example: SN feedback.
  - cold gas reheated by SN via
    \[
    \dot{m}_{\text{th}} = \epsilon_0^{\text{SN}} \left( \frac{200 \text{ km s}^{-1}}{V_{\text{disc}}} \right)^{\alpha_{\text{th}}} \dot{m}_*,
    \]
  - constrain this relation using the M*/M(M) relation

first lesson: lots of degeneracies!
second lesson: what works at z=0 does not work at z=1!
What about at very low masses?

\[
\log_{10}(M_\odot / M_{h}) \leq -4
\]

Our study, \( < M_\ast / M_h \mid M_h > \)

Moster et al. 2009
Guo et al. 2009
Hansen et al. 2009
Klypin et al. in prep.
Yang et al. 2008
Zheng et al. 2007
Blanton et al. 2008
Mandelbaum et al. 2006
Extending to the lowest masses: a simple model

\[ z_{\text{reion}} = 6-12 \]

| star formation; stellar mass set by \( v_{\text{peak}} \) before reionization |
|--------------------------|--------------------------|
| star formation           | no star formation        |

threshold for star formation after reionization

\[ T_{\text{vir}} \sim 1e5K \]
\[ M_{\text{vir}} \sim \text{few } e^9 \]

threshold for star formation before reionization

\[ T_{\text{vir}} \sim 8e3K \]
\[ M_{\text{vir}} \sim \text{few } e^7 \]
At low mass: model appears to work for dwarf galaxies...
...with $M^* - M$ scaling extrapolated from higher masses

Busha et al 09

using via Lactea II subhalos

$M_{300} \sim M_p^{0.25}$
(no scatter in $M^* - M$)
Results depend a bit on the thresholds

Changes the shape of the dwarf velocity function (later reionization = more small dwarfs)

Threshold for star formation before reionization

\[ T \sim 8 \times 10^3 \text{K} \]

This largely sets the total number of dwarfs

Threshold for star formation after reionization

\[ T \sim 2 \times 10^4 - 1 \times 10^5 \text{K} \]

This changes the fraction of dwarfs that have recent star formation
Open question: how much scatter and how sharp is the SF threshold?
What halos are actually probed by current data?

Cosmology uncertainties are more important at high $z$. 
Efficiency of galaxy formation at very high redshift

The peak of galaxy formation efficiency today is about 5%.

At intermediate redshift, galaxy formation becomes less efficient in \( M = 10^{11} \, M_\odot \) halos, presumably due to feedback or inefficient star formation.

Result from Gonzalez, Labbe et al 2009 at \( z \sim 7 \): detection of 11 galaxies with masses \( 1e9-1.3e10 \). Number density corresponds to \( M \sim 1e11 \) halos.
general lessons learned

- Stellar masses and star formation rates appear to be tightly correlated with the masses of their dark matter halos/subhalos -- **mass appears to be the dominant parameter setting galaxy properties.**

- We have a simple model for galaxy clustering which works extremely well!

- Integrated galaxy formation is inefficient even in the most efficient halos (~milky way), and is most efficient in \( M = 1 \times 10^{12} \) at all redshifts.

- At a given stellar mass, **most galaxies are central galaxies** (70-80%).

- Low mass galaxies of a given stellar mass live in more massive halos at higher redshift (implying **star formation is more extended than halo growth**), \( M^* \sim M^{2-3} \), and stellar mass growth of low mass galaxies is dominated by star formation (not merging).

- High mass galaxies live in nearly the same mass halos at all redshifts, implying **halo growth is much more extended than star formation** and merging galaxies must get disrupted; \( M^* \sim M^{0.3} \).

- Scatter in (maximum) halo mass is small for galaxies above \( M^* \sim 1 \times 10^9 \) and is probably mostly explained by scatter in mass accretion histories.

- Although dark matter halos form “hierarchically” (small halos form first), what astronomers typically call “downsizing” (large galaxies form their stars first, smaller galaxies form later), as well as most observed **clustering/environmental trends are a natural outcome of halo growth + inefficient galaxy formation at high and low masses.**
open questions for the future:

It is still very difficult to produce a more physical galaxy formation model or hydrodynamical simulation that reproduces galaxy clustering and the evolution of stellar masses and star formation rates as well as this empirical model. What physics is missing?

- At typical galaxy masses, how precisely do we need to make the galaxy-halo connection to do precision cosmology?

- Which processes/properties are fundamentally set by halo mass vs. merger history, angular momentum, large scale environment, etc? What determines the scatter: merging, mass accretion history, halo structure?

- At the lowest masses, is the threshold for forming galaxies sharp or gradual? Scatter in the galaxy halo connection probably increases for very low mass galaxies and this is likely explained by some manifestation of feedback. With current data this is degenerate with the low mass scaling.

- Galaxy formation efficiency different at very high redshifts? Does increased merger rate lead to much more scatter between galaxy and halo properties? Is feedback operating differently? Can we build a self-consistent model for galaxy formation and reionization?

- Can we make a similar empirically-constrained model that self-consistently incorporates gas? (Data will increase rapidly in the next decade with ALMA, 21-cm, etc)