Galaxy formation and evolution through simulations and SAMs

Toni Sagristà Sellés
Galaxy formation and evolution
Departament d'Astronomia i Meteorologia
Universitat de Barcelona

17 May 2011

Abstract

The current document provides an outline on the main techniques commonly used to study the sensible topic of galaxy formation and evolution, that is, SPH simulations and semi-analytic models (SAMs). We then focus on the strengths and drawbacks of each approach and the various implementation techniques commonly used, to finalize with a careful description of the current state of the work in each field.

Contents

Chapter 1

Introduction

The cosmological constant dominated dark matter model (ΛCDM) provides a successful theoretical framework for understanding the process of galaxy formation. In this framework the large-scale structure of dark matter we infer today is developed through a chain of processes of collision and accretion starting from small-scale density fluctuations in a bottom-up scenario driven only by means of gravitational forces. The baryonic component then slides down the gravitational wells created by dark matter concentrations, shock-heats, cools and condenses into structure such as galaxies [WR78]. The successive grouping of dark matter halos can be tracked with *merger trees*, which are a good theoretical data structure to account for the bottom-up hierarchical formation model, that starts off with multiple entities that merge into larger ones as one advances deep down in the tree. These trees are usually based on analytic approaches, such as Monte-Carlo methods using the extended Press-Schechter formalism [HNS+02]. Also, numerical simulations can be used to effectively follow the dynamical evolution of collisionless systems of particles, provided enough resolution is available to resolve relevant scales. The construction of merger trees to identify halos and sub-halos in these simulations is considerably more demanding in terms of computer resources, but it is still feasible. Simulations of the formation of galaxies inside these DM halos are even more costly, for processes and phenomena other than gravity are to be taken into account, such as gas dynamics, radiative cooling (collapse through internal energy decrease), star formation, gas restitution (return of chemically enriched material from evolved stars to ISM), supernova feedback (energy released by supernovae may decrease star formation efficiency, etc.), AGN outflows and many others.

Two main approaches exist to study the evolution of galaxies, *semi-analytic models* (SAMs) and **N-body/hydrodynamical simulations** (SIMs). SAMs start off from physical assumptions and build up on them, being better suited to understand the underlying processes driving galaxy formation. Usually SAMs are computationally less demanding than simulations, making them very good candidates to test various sets of parameters to find which better adjusts to observational data. However, SAMs do not follow directly the interaction between baryonic matter and dark matter, and very often the models are very simplified and use a large number of parameters to account for different observations simultaneously.

On the other hand, direct cosmological simulations of galaxy formation are based on the direct numerical integration of the equations of hydrodynamics for a set of discrete particles and can record the evolution of dark matter and gas explicitly. They treat the dynamics of fluids (gas) and collisionless systems more realistically than SAMs, but their current resolution is not accurate enough to simulate intermediate and low mass galaxies in large cosmological volumes. Additionally, some processes key to successfuly model the evolution of a galaxy such as star formation or chemical enrichment must be implemented using meshes and interpolations in a sub-resolution fashion, that ultimately use models that need, again, parameters that have to be "guessed".

Both approaches make definite predictions for the evolution of galaxy properties at various masses. Due to their greater computational efficiency, SAMs usually include more models for physical processes, and due to their greater flexibility it's been possible to tune them accordingly to reach a very good agreement with current observational data.

In the next chapters we'll dive into these two techniques to inquire further in their philosophy and implications. Also, we'll give a snapshot of the current state of work in the field.

Chapter 2

Semi-analytic models

An important part of our understanding of the processes involved in galaxy formation comes from semi-analytic models. SAMs are made of recipes for galaxy formation and evolution. They rely on the idea of a bottom-up hierarchical clustering to build a merger tree into which the model will run. Their output is very good to statistically predict galaxy properties and counts. These merger trees are pre-calculated using either the *Extended Press-Schechter* formalism or direct cosmological simulations.

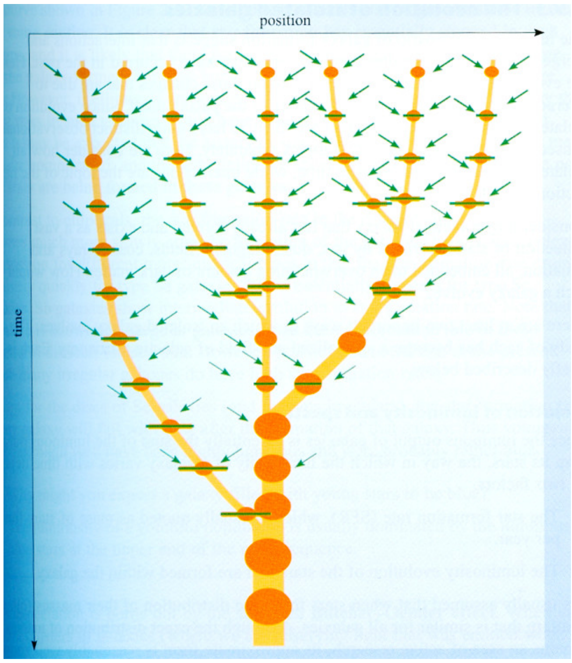

Figure 2.1: Merger tree that shows the schematic formation history of a single giant elliptical galaxy by the merger of many smaller galaxies.

Modern SAMs are quite successful at reproducing observerd statistical properties of galaxies (such as luminosity and stellar mass functions, star formation rates, number counts and so on) in large cosmological volumes over a large range of galaxy masses and redshifts. Particularly, they have proved to produce very accurate numbers for high redshift ($z \leq 6$) and massive galaxies ($M_{star} \geq 10^{10} M_{\odot}$). Therefore, we should expect SAMs to do a better job at reproducing the observable universe than simulations, but we should worry about them doing it so for the wrong reasons due to the great uncertainty existent in many of the processes involved. Also, most of the physical recipes contain free parameters. For example, if a physical process is modeled inadequately it is possible to compensate by modifying a competing process, such as radiative cooling and energy feedback. The latest tendency to run models inside merger trees generated by n-body simulations in order to constrain the evolution of the dark matter component (n-body simulations are good at reproducing the dynamics and evolution of collisionless systems only driven by gravity, such as dark matter), one can isolate various physical processes and attempt to improve the accuracy of the recipes.

A complete SAM must include recipes for all of the following processes involved in galaxy formation and evolution:

Cosmology (h, Ω_Λ, Ω_m, Ω_r, etc.). Nowadays, with 'recent' WMAP results these parameters have changed to the commonly established $\Omega_\Lambda = 0.7, \Omega_m = 0.3$ with $h = 0.7$. This ΛCDM model is widely accepted and it actually saved a lot of time, for now SAMs do not need to be run using different cosmologies to see which better sticks to observations.

Dark halos This is a basic process in galaxy evolution, which is classically described using the Press-Schechter formalism. Nowadays, more and more SAMs are using N-body simulations to account for dark halo structure.

Gas cooling Gas is shock heated as it falls to the potential well of the dark halo producing a hot gas halo. Then there's a net decrease in the internal energy due to radiative non conservative processes of gas cooling. Conservation of angular momentum ensures the cold gas forms a rotationally supported disk.

Star formation Model the first and subsequent generation of stars (Pop-III and Pop-I/II). We need an IMF to determine the masses and some initial conditions. This process is largely speculative for the stellar processes involved in Pop-III stars are not clear.

Feedback Cold gas is heated and removed from the galactic disk due to supernovae winds, AGN and other energy injections.

Historically, the semi-analytic approach to galaxy formation was initially formulated in White & Frenk (1991) [WF91], work that was an elaboration of the ideas exposed in [WR78]. They constructed an analytical model consistent with their understanding of N-body works to the date on dissipationless clustering and of numerical and analytic work on gas evolution and cooling. They also employed standard models for the evolution of stellar populations and constructed new models to represent the chemical enrichment of surrounding gas by evolved stars. They used a cosmology with $\Omega = 1$ and $H_0 = 50 Kms^{-1}Mpc^{-1}, h = 0.5$.

This pulled the trigger for the race of SAMs, for multiple models have appeared since then, the most important being the Munich group (Kauffmann, White, & Guider- doni 1993; Kauffmann, Guiderdoni, & White 1994; Kauffmann 1995; Kauffmann 1996a; Kauffmann 1996b; Kauffmann, Nusser, & Steinmetz 1997; Kauffmann & Charlot 1996) and the Durham group (Cole et al. 1994; Heyl et al. 1995; Baugh, Cole, & Frenk 1996a; Baugh, Cole, & Frenk 1996b; Baugh et al. 1997).

Differences between these models are mainly in the parameters with normalisation and details of star formation, gas cooling and feedback.

2.1 Cosmological SAMs

In this section we'll briefly discuss some cosmological semi-analytic models.

2.1.1 AMIGA: Analytic Model for IGM and GAlaxy evolution

AMIGA is a new SAM whose main purpose is to investigate the cosmological evolution of the IGM at high redshifts. It has some fundamental differences that single it out among the others, like it being fully analytical (makes no use of simulations whatsoever) and it following the metallicity evolution of all baryonic matter from the beginning.

2.1.2 Santa Cruz

This model was first presented in [SP99], which presents a detailed description, very academic, of the basic ingredients they used to build their model.

- Monte-Carlo approach for the dark matter merger trees, using the EPS to extract the probability that a halo of a given mass M_0 at redshift z_0 has a progenitor of mass M_1 at some larger redshift z_1.
- Adapted current cosmology standards to fit COBE data. They run models with $\Omega_0 = 0.3$, $\Omega_0 = 0.5$, and $\Omega_0 = 1$.
- Provide recipes for gas cooling, star formation, chemical evolution, galaxy merging, etc.

2.1.3 GECO: Galaxy Evolution COde

This model, by Riccardelly and Franceschinni [RF10], is also Monte-Carlo based and uses the EPS formalism and include classical recipes for all the processes described avobe. Additionally, they follow the parallel growth of black holes with time and model their feedback on the hosting galaxies. They set the model free parameters by matching data on local stellar mass functions and the relation between galaxy bulge and black-hole mass at $z = 0$.

2.1.4 GALFORM

GALFORM is a semi-analytic model developed by the Durham group for calculating the formation and evolution of galaxies in hierarchical clustering cosmologies [CLBF00]. It improves upon, and extends, the earlier scheme developed by Cole et al.

The model employs a new Monte-Carlo algorithm to follow the merging evolution of dark matter haloes with arbitrary mass resolution. It incorporates realistic descriptions of the density profiles of dark matter haloes and the gas they contain; it follows the chemical evolution of gas and stars, and the associated production of dust; and it includes a detailed calculation of the sizes of discs and spheroids.

2.1.5 Galacticus

And finally, Galacticus is a free, **open source** model of galaxy formation by Andrew J. Benson from Caltech [Ben10]. It can be downloaded from `https://launchpad.net/galacticus`. The Galacticus model was designed to be highly modular to facilitate expansion and the exploration of alternative descriptions of key physical ingredients. For example, it provides various implementations for some modules, such as *Press-Schechter*, *Sheth-Tormen* and *Tikner* for the halo mass function.

It is written in Fortran 90 and contains APIs for C++, Fortran 90, Phyton and Perl. The package also contains a graphical viewer and explorer to start getting results easily.

Chapter 3

N-Body/SPH simulations

3.1 N-Body Simulations

N-body simulations are simulations of dynamical systems of particles that are under the influence of physical forces such as gravity. Tipically and ideally, the force exerted on a particle arises from its interaction with all the other particles in the system. However, we'll see this direct approach is practically unfeasible due to computing restrictions.

There are two parts common and key to any n-body code:

The force routine This is were forces are computed, and our choice here will greatly impact the performance of the simulation.

The integrator This is where equations of motion are solved using the forces calculated by the force routine.

At each step in the simulation, the force routine updates the forces acting on the particles and then the integrator triggers to update the particle's physical properties (position, velocity, acceleration, etc.). A very good article that provides a careful analysis on the most popular techniques to implement these simulations can be found at [Bag05].

The figure 3.1 shows a schematic flow chart of the main modules an N-body code is composed of.

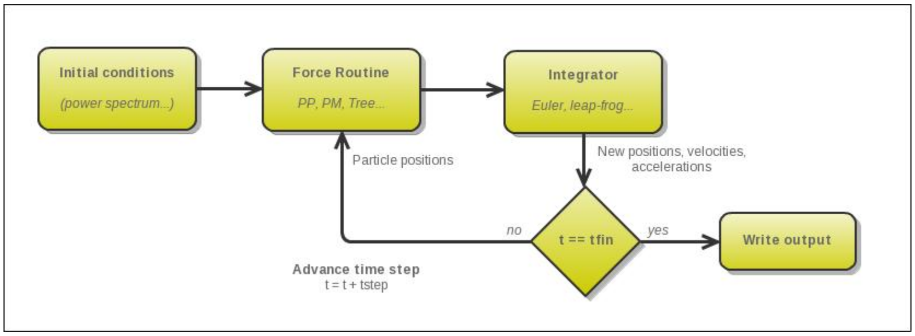

Figure 3.1: Flow chart for an N-body code.

3.1.1 Force routines

There are various options to implement the force routine.

Particle-Particle In this approach each particle must compute its total exerted force using all the other particles in the system. Obviously, this is the most realistic and straightforward method but also the most inefficient ($O(n^2)$), not suitable to simulate systems bigger than an average star cluster.

Particle-Mesh In PM codes, the simulation space is divided using a mesh of density values which sets our limit resolution. In this approach particles no longer interact with other particles, but with the mesh's vertexes, so forces are computed based on the values of the density of the vertexes of the cell they're in. It is computationally better than the PP method, $O(n + n_g log(n_g))$ where n_g is the number of divisions.

Tree In tree codes the space is divided using a multipole principle, and data structures like Quadtrees (2D) and Octal-trees are used. In this approach each node of the tree contains the mass and position of the CM of the region they describe. Computationally efficient, $O(nlog(n))$.

Particle-Particle Particle-Mesh P^3M [EDWF85] is a hybrid method that uses PM for the low resolution space and PP for the precise interactions.

3.1.2 Integrators

Also, multiple integrator methods are available to integrate differential equations from forces to find out $x_i, \dot{x}_i, \ddot{x}_i$.

Euler 1st order integrator, all values are calculated at each time step.

Leapfrog 2nd order integrator, velocities and positions are calculated interleaved in time.

Runge-Kutta

Symplectic Method based on canonical transformations.

3.2 SPH

SPH (Smoothed-Particle Hydrodynamics) is a computational method used to simulate fluid flows. In this technique, the fluid is divided into a set of discrete particles or *tracers*, from which continuous properties are derived. These tracers move with the fluid and have a smoothing length h, over which physical properties are smoothed by a kernel function. The contribution of a particle to the calculation of a physical property is wheighted accordingly to the distance and density. Usually, the kernel function is a gaussian function or a cubic spline,

$$F(r_i) = \sum_{j=1}^{n} F_j \frac{m_j}{\rho_j} W(|r_{ij}|, h_i)$$

where W is the kernel function, n the number of particles and h_i the distance of the particle i to the tracer.

3.3 Cosmological simulations

Below we'll take a glance at some of the existent cosmological simulations.

3.3.1 Cosmological simulation code GADGET-2

GADGET-2 is a free, **open source** code for cosmological N-body/SPH simulations on massively parallel computers with distributed memory. GADGET-2 computes gravitational forces with a hierarchical tree algorithm (optionally in combination with a particle-mesh scheme for long-range gravitational forces) and represents fluids by means of smoothed particle hydrodynamics (SPH). The code can be used for studies of isolated systems, or for simulations that include the cosmological expansion of space, both with or without periodic boundary conditions. GADGET-2 is the starting point for most cosmological simulations.

3.3.2 The Millenium Simulation

This N-body simulation, member of the Virgo consortium, used more than 10 billion particles to trace the evolution of the matter distribution in a cubic region of the Universe over 2 billion light-years on a side. The first results, published in the June 2 2005 issue of Nature, show how comparing such simulated data to large observational surveys can improve our understanding of the physical processes underlying the buildup of real galaxies and black holes. The Millenium Simulation [SWJ+05] is the largest N-body simulation ever carried out. The Millenium Simulation code is based on the GADGET-2.

Chapter 4

Comparison studies

According to [BPF$^+$01], which compares the semi-analytic model of Cole et al. and the SPH simulation of Pearce et al., the two models produce an ensemble of galaxies with very similar properties, altough there are some differences in the gas cooling rates and in the number of galaxies that populate halos of different mass (50% less in SAMs).

SAMs have effectively no resolution limit, one of the biggest problems of N-body/SPH simulations, for very important processes such as star formation or radiative cooling fall beneath the resolution threshold and must be introduced using recipes that are computed at defined triggers in the N-body code, restricting the accessible dynamic range. On the other hand, N-body simulations provide an accurate description of the evolution of structure into the highly non-linear regime where dark matter halos form. To solve this, SAMs use the analytic Press-Schechter theory, which predicts within 50% the distributions of halo masses found in N-body simulations for a specified cosmology, whilst theoretically motivated fitting functions do even better. Extensions of this theory predict the hierarchical build-up of haloes through the mergers of smaller progeintors.

The behaviour of the baryonic matter is less well understood, since the dynamics of the gas are not determined by gravity alone but also hydrodynamical forces and radiative processes, and since gas must cool into dense lumps before it can turn into stars, these processes are crucial for galaxy formation. In this fashion, both techniques require a number of simplifying assumptions in order to model the evolution of cooling gas. SAMs, for example, assume that dark matter haloes and their associated gas component are spherically symmetric, and that gas is efficiently shock-heated when haloes collapse. SPH, as we explained in section 3.2, assumes that gas is well represented by a set of discrete tracer particles. The two methods have different strengths and limitations. SAMs can follow a large dynamic range of scales and is flexible enough to explore the effects of varying assumptions. SPH, on the other hand, does not impose any re-

strictions on geometry and solves directly the approximate evolution equations for gravitationally coupled dark matter and dissipative gas.
Summing up, we can present the following scheme.

SPH Limitations The key assumption is that the evolution of gas can be approximated by the evolution of a set of particles. To find the properties in the continuum, one has to smooth and interpolate using the tracer particles.

SAM Limitations Semi-analytic models make several assumptions on the treatment of gas in order to obtian simple, analytic solutions to complex hydrodynamical processes. Assumptions like the spherical symmetry and the shock-heating of the gas to the virial temperature of its associated halo. The hot gas is then assumed to settle into a distribution with a universal form.

Chapter 5

Conclusions

In this work I've inquired the study of galaxy formation and evolution using two different approaches. We've see either has advantages and drawbacks and they should be complementary to each other. Actually, as we have seen, some SAMs make use of simulations to build their merger trees, and some simulations make use of some recipes developed by SAMs to include into their modellings of sub-resolution processes to enhance their effectiveness and get results they couldn't have achieved otherwise.

Bibliography

[Bag05] J.S. Bagla, *Cosmological n-body simulation: Techniques, scope and status*, Current Science **88** (2005), 1088–1100.

[Ben10] A. J. Benson, *Galacticus: A semi-analytic model of galaxy formation*, New Astronomy (2010), 35.

[BPF+01] A.J. Benson, F.R. Pearce, C.S. Frenk, C.M. Baugh, and A. Jenkins, *A comparison of semi-analytic and smoothed particle hydrodynamics galaxy formation*, MNRAS **320** (2001), 261–280.

[CLBF00] S. Cole, C.G. Lacey, C.M. Baugh, and C.S. Frenk, *Hierarchical galaxy formation*, MNRAS **319** (2000), 168–204.

[EDWF85] G. Efstathiou, M. Davis, S.D.M. White, and C.S. Frenk, *Numerical techniques for large cosmological N-body simulations*, Astrophysical Journal Supplement Series **57** (1985), 241–260.

[HNS+02] M. Hirschmann, T. Naab, R. Somerville, A. Burkert, and L. Oser, *Galaxy formation in semi-analytic models and cosmological hydrodynamic zoom simulations*, MNRAS **000** (2002), 1–23.

[RF10] E. Ricciardelli and A. Franceschini, *Geco: Galaxy evolution code - a new semi-analytical model of galaxy formation*, Astronomy and Astrophysics **518** (2010), A14+.

[SP99] R.S. Somerville and J.R. Primack, *Semi-analytic modelling of galaxy formation: the local universe*, MNRAS **310** (1999), 1087–1110.

[SWJ+05] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, D. Thacker, R. Croton, J. Helly, J. A. Peacock, Thomas P. Cole, S., H. Couchman, A. Evrard, J. Colberg, and F. Pearce, *Simulating the joint evolution of quasars, galaxies and their large-scale distribution*, Nature **435** (2005), no. astro-ph/0504097, 629–636. 42 p.

[WF91] S. D. M. White and C. S. Frenk, *Galaxy formation through hierarchical clustering*, Astrophysical Journal **379** (1991), 52–79.

[WR78] S.D.M. White and M.J. Rees, *Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering*, MNRAS **183** (1978), 341–358.