Bruno Henriques

Peter Thomas

Hybrid Galaxy Evolution Modelling

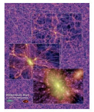
Monte Carlo Markov Chain Parameter Estimation in Semi-Analytic Models of Galaxy Formation

University of Sussex

The Millennium Simulation

- The Millennium Simulation (Springel et al. 2005) its one of the largest and highest resolution direct simulations of the evolution of dark matter structure carried to date. It traces the evolution of dark matter haloes in a cubic box of $500h^{-1}$Mpc on a side, a volume large enough to sample the richest clusters of galaxies. The simulation follows 2160^3 dark matter particles of mass $8.6\text{x}10^8 h^{-1}M_\odot$, corresponding to an halo mass resolution of $1.7\text{x}10^{10}h^{-1}\,M_\odot$ and baryonic mass of about $3.1\text{x}10^9h^{-1}\,M_\odot$.

- From the starting redshift ($z = 127$) to redshift zero, 64 snapshots are recorded, with a finer spacing towards the local Universe. For the dark matter structure a friend-of-friends algorithm identifies the halos by joining together particles into groups with a mean overdensity of about 200, approximately what is expected for a virialised group. Then substructures with more then 20 gravitationally bound particles are identified as sub-halos.



- After finding the dark matter structures and their properties at each snapshot their structural evolution is derived, by identifying each halo's descendant (unique in a hierarchically growing ΛCDM Universe), which defines the entire merger tree of each individual object. Since the halos contained in each tree are gravitationally isolated from other structures, the properties of galaxies within a tree are fully determined by the dark matter halos inside it.

A Semi-Analytic Model of Galaxy Formation - I

Gas Infall

Using the millennium simulation, the semi-analytic model (De Lucia & Blaizot 2007) assumes that each dark matter halo collapses with a fixed amount of baryons. This corresponds to a mass fraction of 17%, initially in the form of diffuse gas with primordial composition, but which will lately be distributed in different phases and fuel star formation.

Gas Cooling

In most halos, the cooling radius lies within the virial radius and the infalling gas forms a quasi-static hot atmosphere, after being shock heated to the virial temperature. This gas can cool at later times and its accretion into central regions where star formation will occur is modelled through a cooling flow.

$$t_{\rm cool}(r) = \frac{3}{2}\frac{\tilde{\mu} m_p k T_{\rm gas}}{\rho_{\rm gas}(r)\Lambda(T_{\rm gas}, Z_{\rm gas})}$$

Star Formation

The formation of stars from the cold gas occurs in the disk from 2 different channels, either quiescently or in bursts during mergers.

- Quiescently, when the surface density of the cold gas is above a certain threshold, converted into a critical mass of gas ($m_{\rm crit}$),

$$\dot{m}_\star = \alpha_{\rm SF}\frac{(m_{\rm cold} - m_{\rm crit})}{t_{\rm dyn,disk}}$$

- Or through bursts, when galaxy mergers occur,

$$\dot{m}_\star^{\rm burst} = 0.56\left(\frac{m_{\rm sat}}{m_{\rm central}}\right)^{0.7} m_{\rm gas}$$

A Semi-Analytic Model of Galaxy Formation - II

The fact that the halo mass function is much higher at both ends than the galaxy stellar mass function implies the existence of processes that suppress cooling (and hence quench star formation) in both dwarf and massive galaxies. In the model, SN feedback decreases the star formation rate in small galaxies and AGN in large systems which grow a supermassive central black hole.

Supernovae Feedback

As massive stars complete their life cycle, SN events start injecting energy into the surrounding medium, reheating the cold disk gas. If the available energy exceeds the necessary to reheat the gas into the hot phase, a fraction is ejected into an external reservoir, from which it will be reincorporated into the hot phase at later times.

$$\Delta m_{\rm reheated} = \epsilon_{\rm disk}\Delta m_\star$$

$$\Delta m_{\rm ejected} = \left(\epsilon_{\rm halo}\frac{V_{\rm SN}^2}{V_{\rm vir}^2} - \epsilon_{\rm disk}\right)\Delta m_\star$$

$$\dot{m}_{\rm ejected} = -\gamma_{\rm ej}\frac{m_{\rm ejected}}{t_{\rm dyn}}$$

AGN Feedback

Supermassive black holes grow in large galaxies mainly by the accretion of cold gas during mergers (the quasar mode).

$$\Delta m_{\rm BH,Q} = \frac{f_{\rm BH}(m_{\rm sat}/m_{\rm central})\, m_{\rm cold}}{1 + (280\,{\rm km\,s^{-1}}/V_{\rm vir})^2}$$

The radio mode is less significant for the growth of the black hole, but this quiescent accretion of hot gas, generates mechanical heating that is responsible for reducing the cooling rate.

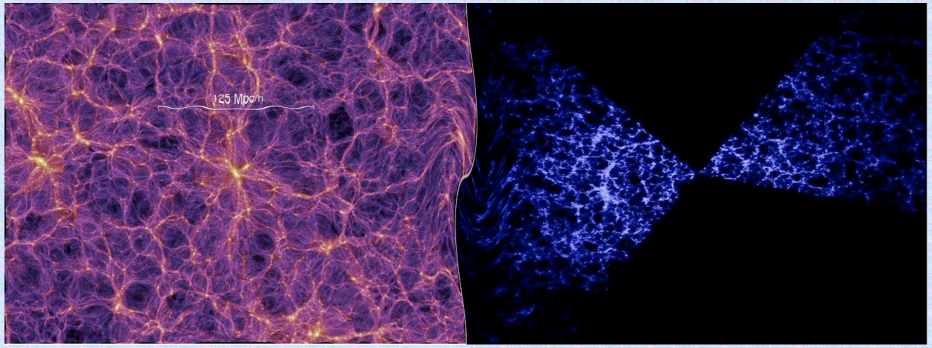
$$\dot{m}_{\rm BH,R} = k_{\rm AGN}\left(\frac{m_{\rm BH}}{10^8\,{\rm M_\odot}}\right)\left(\frac{f_{\rm hot}}{0.1}\right)\left(\frac{V_{\rm vir}}{200\,{\rm km\,s^{-1}}}\right)^3 \qquad L_{\rm BH} = \eta\,\dot{m}_{\rm BH,R}\,c^2$$

Since the heating the radio mode heating depends on the black hole mass, the quasar mode will also determine the amount of cooling suppressed.

The cold dark matter model has become the leading theoretical paradigm for the formation of structure in the Universe. Testing this model requires that precise measurements derived from galaxy surveys (such as SDSS) can be compared with equally robust theoretical calculations.

Present numerical capabilities allow reliable simulations of dark matter and diffuse gas. However, once the gas cools into halo cores, its properties are determined by small-scale processes that cannot be resolved. The current approach is to treat the dark matter evolution and galaxy physics separately. The dark matter halos merger trees are generated with N-body simulations and used as an input for semi-analytic models, which follow the evolution of the baryonic component.

Present day semi-analytics, qualitatively reproduce a vast range of observations. However, discrepancies between theory and observations are inevitable. Whenever this happens, it reveals a fundamental problem: the growing complexity of physics included makes it difficult to understand if there is a fundamental problem with the underlying galaxy formation processes or just a failure in adjusting the parameters into a maximum agreement configuration.



Monte Carlo Markov Chain methods allow us to maximize the constraining power of different observational data sets. The correlations between the parameters that emerge when we require the semi-analytic model to reproduce a given galaxy property give us insight into the meaningful physical quantities governing galaxy evolution that are represented by the parametrizations.

In addition, the MCMC technique gives us a measure of the goodness of fit of the model to the data and maps out the allowable range of parameter values. It can be used to discriminate between competing semi-analytic models.

The original semi-analytic model of De Lucia & Blaizot (2007) requires extremely efficient supernova feedback in order to correctly predict the number density of dwarf galaxies. To avoid this, we introduce satellite disruption by tidal forces in the model and obtain a new best fit using MCMC techniques. Our satellite disruption model has a likelihood four times higher than the original, and in addition predicts the correct metallicities for galaxies and a sensible amount of intracluster light.

Monte Carlo Markov Chain Sampling in Semi-Analytic Models - I

We apply MCMC techniques to sample the parameter space in a semi-analytic model of galaxy formation. Using the De Lucia & Blaizot 2007 model and a range of observational constraints, we quantify the relative impact of the different physics in the model on the predicted properties. This is achieved by sampling the parameter space and its allowed likelihood region in order to correctly predict the observational K-band, g-r colours and black hole-bulge mass relation, first separately and then combined. The correlations between the parameters give us insight on the meaningful physical processes defining each predicted property. In here we show the example of the K-band luminosity function.

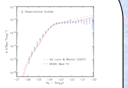


Fig. 1 - The original K-band LF and the predicted from our best fit are compared with observations.

- To reduce the number of dwarf galaxies MCMC requires the supernova feedback to be more effective than in the original model, by increasing the amount of ejected gas. This is represented by the relative strength between ejection and reincorporation (bottom right, Fig. 2).

- This alone would reduce the number density of all galaxies bellow L_*. To only suppress dwarfs, we require the SN feedback to only be effective in very small halos, by decreasing the $V_{\rm vir}$ cutoff for this process, represented by the fraction between reheating and ejection (bottom left corner).

$$\Delta m_{\rm ejected} = \left(\epsilon_{\rm halo}\frac{V_{\rm SN}^2}{V_{\rm vir}^2} - \epsilon_{\rm disk}\right)\Delta m_\star \qquad V_{\rm vir,0} = \left(\frac{\epsilon_{\rm halo}}{\epsilon_{\rm disk}}\right)^{\frac{1}{2}} V_{\rm SN}$$

- There is a degeneracy between the AGN parameters, since their product is the determinant quantity for the amount of cooling suppressed (upper left).

$$L_{\rm BH} \propto f_{\rm BH}\, k_{\rm AGN}\, m_{\rm cold}\, f_{\rm hot}$$

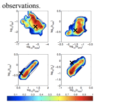


Fig. 2 - The contours follow the MCMC sampling in parameter space, while the colours represent the maximum likelihood projected along the hidden dimensions. The crosses mark the parameter values in the original model.

Monte Carlo Markov Chain Sampling in Semi-Analytic Models - II

After performing the sampling with individual constrains we rerun the MCMC requiring the model to match the three observations in order to obtain a best fit. The use of a comprehensive data set to constrain the model reduces the allowed likelihood range in parameter space into a small region (Fig. 3). In comparison to the values in the original model, the star formation efficiency and AGN feedback parameters are almost unchanged. However, new regions are found for the parameters controlling the supernova feedback, representing an increase in the efficiency of this process.

In Fig.4 and Fig. 5 we show the predictions for the best fit model of the additional properties used in the sampling, respectively the fraction of red galaxies and the black hole-bulge mass relation. Despite the better qualitative agreement achieved in comparison to the original model, our best fit still has a low likelihood value of 0.037

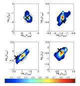


Fig. 3 - MCMC sampling with the three observational constraints. The best fit likelihood value is 0.037.

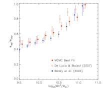


Fig. 4 - The original fraction of red galaxies and that predicted from our best fit are compared with observations.

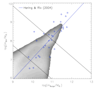


Fig. 5 – The predicted black-hole/bulge mass ratio from our best fit is compared with observations.

In order to decrease the excess of dwarf galaxies in the original model, the MCMC requires an extremely efficent supernova feedback. This seems to be unfeasible from observations of galactic outflows and indicates that the treatmeant of dwarf galaxies might need to be improved.

As a possible solution we introduce satellite disruption, following the tidal forces exerted by their central companions during mergers. Using the MCMC techniques we study the impact of this process on the physics of the model and test if it can help improving the likelihood of our best fit.

MCMC Sampling in a Model with Satellite Disruption

To quantify the amount of material disrupted during mergers, we compute the forces acting on satellite galaxies. The tidal radius, outside which material will be disrupted and transferred into the intra-cluster medium, is given by the point inside the satellite where the sum of the tidal force from the central companion and the centrifugal force equal the gravitational binding force of the satellite.

Satellite disruption affects both small and massive objects, respectively satellites that have their material disrupted and central galaxies which receive less mass from mergers. This has a clear effect on the results of the MCMC sampling (Fig. 6). The new best fit requires a much more efficient star formation efficiency in order to increase the number density of the most massive objects, and compensate for the material that they previously received that is now transferred into the intra-cluster light. On the other hand, supernova feedback is now required to be less efficient since the previous excess of dwarf galaxies is now eased by the loss of material by small satellites.

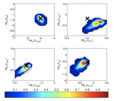


Fig. 6 - MCMC sampling for the model with satellite disruption. The likelihood of the new best fit likelihood is 0.15. The crosses represent the previous best fit.

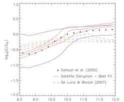


Fig. 7 – The predicted metallicity of galaxies in models with and without satellite disruption.

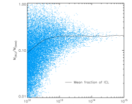


Fig. 8 - Predictions for the fraction of mass in the intra-cluster medium for the best fit model with satellite disruption. The solid line represents the mean value in each virial mass bin.

The new model produces a similar qualitative fit to the observations used to constrain the sampling. However the likelihood of the new best fit model is now 0.15, four times higher then before. This seems to indicate that disruption is indeed required by observations. Moreover, the new model has predictions for intra-cluster light, with a value of 30% in the most massive clusters matching observations (Fig. 8).

References:

-Croton D. J., Springel V., White S.D.M., et.al., 2006, MNRAS, 365, 11-28
-De Lucia G. & Blaizot J., 2007, MNRAS, 375, 2-14
-Henriques B., Thomas P., Oliver S., Roseboom I., 2009, MNRAS, in press
-Springel V., White S. D. M., Jenkins A. et. al., 2005, Nature, 435, 629-636