

The evolution of massive galaxies in semi-analytical models of galaxy formation

Carlton M. Baugh

Institute for Computational Cosmology, Department of Physics,
Durham University, Science Laboratories, South Road, Durham, DH1 3LE, UK.
email: c.m.baugh@durham.ac.uk

Abstract. Massive galaxies with old stellar populations have been put forwards as a challenge to models in which cosmic structures grow hierarchically through gravitational instability. I will explain how the growth of massive galaxies is helped by features of hierarchical models. I give a brief outline of how the galaxy formation process is modelled in hierarchical cosmologies using semi-analytical models, and illustrate how these models can be refined as our understanding of processes such as star formation improves. I then present a brief survey of the current state of play in the modelling of massive galaxies and list some outstanding challenges.

Keywords. galaxies:formation, galaxies:evolution, galaxies: high-redshift, mehtds:numerical.

1. Introduction

The observational evidence in support of the emergence of massive galaxies at early epochs in the history of the Universe is now overwhelming. Different techniques have successfully isolated galaxies at substantial lookback times, which have stellar masses similar to those of the largest galaxies seen today. Often the methods used to identify such galaxies use colour selection, and target galaxies which either have old stellar populations with little recent star formation, or which may have substantial star formation activity, but which appear red because they are enshrouded in dust.

At the same time, the past twenty years have seen a hierarchical cosmology – the cold dark matter model with a cosmological constant (Λ CDM) – cemented in place as the “standard model” of cosmology. In Λ CDM, very small fluctuations in the density of the Universe, as revealed in maps of the cosmic microwave background radiation, are amplified due to gravitational instability. The assembly of self-gravitating structures called dark matter haloes, is hierarchical, proceeding via steady accretion and mergers of more substantial structures. Analytic calculations and N-body simulations have been applied to study the hierarchical clustering of the dark matter to high accuracy. The hierarchical growth of the abundance of dark matter haloes is well established in such calculations.

The apparent dilemma then faced by theoreticians is how do we reconcile the observations of seemingly old, massive galaxies in place at early epochs in the history of the Universe with the steady growth of dark matter haloes, whereby the largest structures are forming now?

In this review I will describe how massive galaxies fit into hierarchical models, explaining how their mass is built up. I will then outline the semi-analytical approach to modelling galaxy formation, setting out the pros and cons of the technique. I will highlight results from different groups running semi-analytical models; due to space restrictions, this will be an unavoidably incomplete overview of the current literature, but

will hopefully serve to whet the appetite of the reader to investigate the predictions of this class of model further.

2. How do galaxies acquire their mass?

Two features of hierarchical clustering cosmologies help to build big galaxies at early epochs. These effects are not peculiar to the approach used to model galaxy formation, but instead arise from the process of gravitational instability operating in a model with fluctuations over a broad range of mass scales.

These effects are:

Formation bias. Regions of the Universe which are overdense evolve faster than those that are underdense. The higher density means a faster turnaround from the background expansion and an earlier gravitational collapse, compared with the evolution of a similar sized object in an underdense part of the Universe. Hence within a large overdensity, dark matter haloes which are able to retain gas that has been heated by supernovae will tend to form sooner and so galaxy formation will be advanced compared to an average patch of the Universe. Therefore, if massive galaxies form in massive haloes, then there is a natural tendency for these objects to form earlier, purely due to gravity.

Formation time versus assembly time. In hierarchical models, it is not necessary to form the mass of a galaxy in situ, in one object. The task of accumulating stellar mass can be delegated to a number of progenitor galaxies, which may individually form stars at modest rates, but collectively build stellar mass more prodigiously. Massive galaxies display a distinct shift between the formation time of their stars, e.g. the time when half of the present day mass was in place in all of the progenitors, and the assembly time, when the progenitors merge into one object (Baugh, Cole & Frenk 1996; Kauffmann 1996; De Lucia *et al.* 2006).

3. Semi-analytical modelling of galaxy formation

By itself, the cold dark matter model says nothing directly about galaxy formation. Inferences can be drawn about the sequence of galaxy formation, based on how structures grow in the dark matter. However, without an attempt at a physically motivated calculation of the fate of baryons in a cold dark matter Universe, there is little hope of learning much about galaxy formation or of understanding the implications of observations of high redshift galaxies for the cold dark matter cosmology (see recent reviews, Baugh 2006 and Benson 2010).

White & Rees (1978) argued that galaxy formation is a two-phase process, with the bulk of the mass undergoing a dissipationless collapse which is responsible for building the gravitational potential wells, or halos, in which galaxies form. The baryonic component of the Universe is able to dissipate energy, and therefore to collapse down to smaller scales, forming denser units. This model was able to explain the appearance of clusters of galaxies. However, without an additional process to reduce the efficiency of galaxy formation in shallow gravitational potential wells, the predicted luminosity function is much steeper than is observed at the faint end.

This pioneering work, along with a clutch of papers published around the same time looking at the radiative cooling of gas within gravitational potential wells, laid the groundwork for modern galaxy formation theories. The time taken for all of the gas within a halo to cool radiatively increases with halo mass. This is because cooling is a two-body process (collisionally excited radiative transitions or bremsstrahlung) which depends on the square of the gas density. In hierarchical models, more massive haloes

tend to form later when the density of the Universe is lower. It is possible for the cooling time of the gas to exceed the Hubble time, thus limiting the supply of cold gas to form a galaxy (see the review of Fred Hoyle's contributions to galaxy formation theory by Efstathiou 2003).

The first papers to incorporate these ideas fully into the cold dark matter cosmology, introducing the semi-analytical methodology, were published in 1991 (White & Frenk 1991; Cole 1991, Lacey *et al.* 1991). This approach tries to follow a wide range of the processes which are thought to be important in determining the fate of the baryons. This is a daunting task. At the time, theories of star formation were rudimentary at best. There has been much progress in this area since 1991, but we are still a long way from having a reliable description of the process which underpins galaxy formation. The regulation of star formation efficiency comes from the stars themselves. Stars above $\approx 5-8$ times the mass of the Sun end their life in a Type II supernova, which injects substantial amounts of energy and momentum into the interstellar medium. This alters the state of the gas in the interstellar medium (ISM), perhaps leading to the ejection of gas from the galactic disk or even the dark matter halo. This process is known as supernova feedback and is critical to the success of any model of galaxy formation.

The absence of a precise description of a key process, such as star formation or supernova feedback, may lead one to consider giving up any hope of ever understanding galaxy formation. Instead, in semi-analytical modelling an attempt is made to write down the differential equation which gives the current best bet model of how the system behaves. As our understanding develops, or when new observations clarify how a process works, then the model can be improved. The differential equation may contain a free parameter. Often there is little guidance as to the appropriate range of values to take for the parameter. In such instances, the only approach is to be pragmatic and see what the model predicts for different parameter values. By comparing the model predictions to observations, the value of the parameter is set as the one which gives the most faithful reproduction of the data. This procedure is exactly what physicists undergo when attempting to describe complex phenomena: start off with a simple model, which can be adjusted or refined to improve the description of observations. I will give an example of this principle in action below.

The semi-analytical framework allows us to model a range of processes together, within the cosmological setting of the formation of structure in the dark matter. The ability to follow the interplay between processes is essential in studying galaxy formation. The models solve the set of differential equations which govern the flow of mass and metals between different reservoirs of baryons: hot gas, cold gas and stars. The output of the models is the full star formation and chemical enrichment histories for a wide range of galaxies, including mergers between galaxies.

Semi-analytical modelling has some features which might be perceived as limitations or drawbacks. One example is the generality of the assumptions which are needed to be able to calculate the fate of the baryonic component (e.g. the assumption of a spherically symmetric hot halo gas profile to compute the cooling rate). Another is the "deterministic" way in which processes such as supernova feedback are modelled. In the semi-analytical model, the mass loading of the supernova driven wind is specified by choosing model parameters, and precisely this amount of gas is ejected from the ISM. In a gas dynamics simulation in which the wind is fully coupled to the hydrodynamics equations, the same number of supernovae could result in a very different mass of gas being ejected. The mass loading could be intricately linked to the resolution of the simulation.

Nevertheless, despite the progress made over the past twenty years, there is still widespread mistrust of semi-analytical modelling. Recently this has led to something

akin to grassroots insurgency in the subject, effectively a “tea party” movement for galaxy formation, whereby simplified models have been devised with the aim of elucidating how galaxies form. Examples include the “bathtub” and “reservoir” models. These calculations are inspired by models of supply and demand from economics, and track the inflow (sources) of gas into halos and the “sinks” of cold gas in star formation and supernova feedback. In their simplest form, the models follow one galaxy per halo, and invoke ad-hoc efficiency factors to specify the inflow of gas as a function of halo mass, without any attempt to calculate the rate at which gas can cool or to explain the form of the efficiency factor. Galaxy mergers are ignored. This class of calculation effectively takes one of the equations which has been considered within semi-analytical models for more than a decade and solves it in isolation.

The desire for a better grasp of how galaxy properties are shaped by different processes is understandable, but it is not clear that it can be usefully gained from stripped-down approaches. The perceived “complexity” of semi-analytical modelling is actually the great strength of the technique. The ability to model the interplay between processes is the key to building a realistic model of galaxy formation. By taking a more complete view of galaxy formation rather than a selective one, the consequences of the calculation – the predictions of the model – are more far reaching and therefore more tightly constrained by observations. If the model seems complex, then this is simply a reflection of the nature of the underlying processes, such as star formation and heating by supernovae.

4. An illustration of semi-analytical modelling: the star formation rate in galaxies

An illustration of how semi-analytical models work can be obtained by considering how star formation is modelled within a galaxy.

The bulk of semi-analytical models attempt to predict the global star formation rate within a galaxy. The early modelling of the star formation rate was essentially based on dynamical arguments, with loose motivation coming from a comparison to the Kennicutt-Schmidt law (Bell *et al.* 2003). The star formation rate, ψ , is often parametrised as

$$\psi = \epsilon \frac{M_{\text{cold}}}{\tau},$$

where M_{cold} is the total mass of cold gas, and ϵ is an efficiency factor, which controls the fraction of cold gas which is turned into stars in the timescale τ . The timescale for star formation is generally assumed to scale with the dynamical time within the galaxy:

$$\tau = t_{\text{dyn}} f(v_{\text{disk}}).$$

In some models, $f(v_{\text{disk}}) = 1$; in the Cole *et al.* (2000) model, an explicit scaling of the star formation timescale with the circular velocity of the disk was implemented, to allow the model to produce a better match to the observed gas fraction luminosity relation for spiral galaxies: $f(v_{\text{disk}}) = (v_{\text{disk}}/200\text{kms}^{-1})^{\alpha_*}$. Hence in the most general case, two parameters are required to set the star formation rate: ϵ and α_* . These parameters are set by choosing values, running the model and then comparing the model predictions to observables. The key observables for constraining the values of these star formation parameters are the gas fraction - luminosity relation, the galaxy luminosity function and the colour magnitude relation.

High resolution imaging of galaxies at different wavelengths has revealed that star formation activity correlates better with the molecular hydrogen content of galaxies than

with the overall cold gas mass. Lagos *et al.* (2011a) investigated more general star formation models in the GALFORM semi-analytical model, implementing different empirical and theoretically motivated star formation laws (see also Cook *et al.* 2010; Fu *et al.* 2010). The most successful of these was the empirical star formation law proposed by Blitz & Rosolowsky (2006), who suggested that the observational data could be explained if the ratio of molecular to atomic hydrogen is set by the pressure in the mid-plane of galactic disks; gas disks with higher pressure have a higher fraction of H₂.

This work illustrates the modularity of semi-analytical modelling and how it provides a framework in which new and improved descriptions of various processes can be readily implemented. The Blitz & Rosolowsky star formation law involves two observationally determined “parameters”. Whereas in the original parameterization of the star formation rate there was little guidance about the range of parameter values which should be considered, there is now a much smaller volume of parameter space to search (at least once the Blitz & Rosolowsky law has been adopted). Furthermore, as the modelling of the star formation becomes more sophisticated, the predictions that can be made by the model expand. Rather than simply outputting the cold gas mass of galaxies, the atomic and molecular hydrogen contents are now predicted, meaning that the model should also be able to reproduce the mass functions of HI and H₂, their evolution and their relation to other galaxy properties (Lagos *et al.* 2011b). By combining GALFORM with the photon dominated region model of Bell *et al.* (2006), it is also possible to predict the different carbon monoxide transitions, and to make contact with observations from ALMA (Lagos *et al.* 2012).

Hence by adopting the improved star formation model, the parameter space open to the model has shrunk in volume and the constraints on the model have increased through the capability to make new predictions which must match the available observations.

5. Massive galaxies in hierarchical models

I now review some of the predictions made by semi-analytical models for the abundance of massive galaxies at different epochs.

5.1. *The massive end of the stellar mass function*

Traditionally semi-analytical models have used observational estimates of the local galaxy luminosity function to set the values of many of the model parameters. The comparison with the luminosity function requires modeling which is already part of the majority of semi-analytical models. This involves taking the star formation and chemical enrichment histories predicted for each galaxy, including bursts of star formation, and combining this with a simple stellar population model, and a model for the attenuation of starlight by dust. The stellar population model depends on the choice of stellar initial mass function. This choice in turn influences the galaxy formation model, by setting the yield of metals and the mass of gas recycled into the ISM during an episode of star formation.

With the advent of large galaxy redshift surveys with multi-band photometry, the present day stellar mass function has now been determined to high accuracy in so far as sampling errors are concerned (Cole *et al.* 2001; Bell *et al.* 2003; Li & White 2009). The Li & White determination of the stellar mass function is based on stellar masses inferred from SDSS photometry and redshifts, assuming a simple model for the star formation history, a universal stellar initial mass function (IMF) and a model for dust extinction.

Hence, the comparison between the model predictions with the observational estimate requires an additional layer of modelling when considering the stellar mass function compared with the luminosity function. A similar set of assumptions to those already made

independently in the galaxy formation model (the form of the IMF, the star formation history, the chemical enrichment history, the dust attenuation) is required to process the direct observables in order to infer stellar masses. Complications arise when there is a mis-match between the assumptions made in the observational estimate and the predictions and choices made in the semi-analytical model.

Guo *et al.* (2011) present a prediction for the stellar mass function which is in very good agreement with the observationally inferred mass function of Li & White over around two and a half decades in mass around the break in the mass function. The model overpredicts the mass function for galaxies with mass $10^{11.5} M_{\odot}$ by up to a factor of two. Guo *et al.* argue that this discrepancy could be reduced if stripping of stellar mass and cold gas from satellite galaxies is included in massive haloes.

At high redshift, the Guo *et al.* model agrees with the observationally estimated stellar mass function around $10^{11} M_{\odot}$. The model overpredicts the abundance of lower mass galaxies and underpredicts the number density of more massive galaxies. Two possible solutions have been put forward to boost the number of massive galaxies at high redshift, which do not require changes to the galaxy formation model. The first is to revise the stellar population synthesis model used. Currently there is some debate over the handling of asymptotic giant branch stars in the SSP models (see the contributions of Claudia Maraston and Gustavo Bruzual to this symposium). The model of Maraston (2005; see also Maraston *et al.* 2006) predicts more light in the near infrared for stellar populations with ages around ~ 1 Gyr, than, for example, the models of Bruzual & Charlot (2003); this would boost the bright end of the luminosity function predicted by the model at high redshift. The second solution discussed by Guo *et al.* is to take into account the errors in the determination of the stellar mass. Due to the steepness of the stellar mass function at the massive end, the errors preferentially scatter galaxies from lower to higher mass bins, increasing the abundance of high mass galaxies. To show this, Guo *et al.* convolve the raw stellar mass function output by their model with a Gaussian which represents the estimated errors. In practice, the error could be systematic as well as random (Marchesini *et al.* 2009).

Ideally, to carry out a fair comparison between models and observations, one should attempt to estimate the stellar masses of the model galaxies by feeding their broad band photometry into the same estimation software applied to the observations (Pforr, Maraston & Tonini 2012; Mitchell *et al.* in preparation). In their estimation of stellar masses from the photometry output by GALFORM, Mitchell *et al.* show that, surprisingly, the complexity in the star formation histories predicted by the model (as compared to the simple exponentially declining star formation rates traditionally assumed in performing the fitting) is not responsible for the largest contribution to the error in the estimated stellar mass. Instead, substantial systematic errors arise from the way in which the variation in the metallicity of the stellar populations is dealt with and in the modelling of dust extinction. (Note that GALFORM assumes that the stars and dust are mixed together which gives different attenuation of the starlight compared to foreground slab models or the commonly used Calzetti law – see González-Pérez *et al.* 2013 for a discussion.)

5.2. The abundance of massive red galaxies

Various colour selections have been applied to isolate galaxies at intermediate and high redshifts. The galaxies selected are typically massive, due to the combination of apparent magnitude limit and the target redshift range, but can display a range of star formation activity. In this subsection, we focus on three examples: 1) Luminous Red Galaxies (LRGs) which lie at $z = 0.2 - 0.7$; 2) BzK galaxies in the redshift interval $1.4 < z < 2.5$; and 3) Extremely Red Objects (EROs), which have median redshifts of $z > 1$.

1) Luminous Red Galaxies (LRGs)

LRGs are identified in the Sloan Digital Sky Survey on the basis of their colour, with a selection designed to isolate galaxies with passively evolving stellar populations (Eisenstein *et al.* 2001). Almeida *et al.* (2008) applied the same colour selection to model galaxies in two variants of the GALFORM model, Baugh *et al.* (2005) and Bower *et al.* (2006). Without any adjustment to the model parameters, both models come within 15% of matching the observed space density of SDSS LRGs at $z = 0.24$. The models fare less well at higher redshift, either under- (Baugh *et al.* or over-(Bower *et al.*) predicting the abundance of 2SLAQ LRGs (a deeper sample using SDSS photometry and 2dFGRS spectroscopy) by a factor of ≈ 2 . The Bower *et al.* model gives a remarkably good match to the observed correlation function of LRGs measured by Masjedi *et al.* (2006), which is a power-law over three and a half decades in pair separation. Almeida *et al.* demonstrate how the stellar mass of LRGs is built up in the model, with 50% of the stars in place in by $z = 2.2$ in all progenitors of the $z = 0.24$ LRG, with half the final mass in one progenitor a billion years later by $z = 1.6$,

2) BzK galaxies.

The BzK colour selection targets the spectral features close to 4000\AA : the 3646 \AA Balmer break which is seen in A stars (and O and B stars) and the accumulation of absorption lines due to ionised metals at 4000\AA (mainly from F, G and K stars). With the B, z and K, filters these break features can be used to isolate galaxies in the redshift interval $1.4 < z < 2.5$ (Daddi *et al.* 2004). Merson *et al.* (2013) applied the BzK colour selection to the model of Bower *et al.* (2006). The BzK technique isolates $\sim 75\%$ of the model galaxies in the target redshift interval, with some interlopers from other redshifts. BzK galaxies are further divided into star forming and passive subsamples, depending on their location in the BzK plane. The Bower *et al.* model gives an excellent match to the number of star forming BzK galaxies. Passive BzK galaxies are observed to be less common than star forming BzK galaxies; the model does less well at matching the slope of the counts of passive BzK galaxies.

3) Extremely Red Objects (EROs).

EROs are selected to have very red $R - K$ or $i - K$ colours, and for $K \sim 20$ have redshifts $z \sim 1$. Early generations of semi-analytical models underpredicted the number of EROs by an order of magnitude (Smith *et al.* 2002). González-Pérez *et al.* (2009) examined the predictions of published models for the abundance and nature of EROs. The Baugh *et al.* model was found to fail to match the ERO counts by a factor of 10. This model adopts a long duration for starbursts, to help match the number counts of galaxies at $850\mu\text{m}$. This leads to blue colours for bursting galaxies. The Bower *et al.* model gives a very good match to the abundance of EROs; in this model the burst duration is shorter than in the Baugh *et al.* model and the overall amount of star formation in bursts is lower, both of which will lead to redder bursts in the Bower *et al.* model. Other semi-analytical models have also matched the counts of EROs, but were either not able to model the whole of the galaxy population to the present day, or overpredicted the number of massive galaxies at the present day (Granato *et al.* 2004; Kang *et al.* 2006; Menci *et al.* 2006; Fontanot *et al.* 2007)

6. Conclusions and remaining problems

Semi-analytical models are a powerful tool which allow galaxy formation to be modelled in the context of cosmologies in which structure grows hierarchically structure. The models are modular, so as our understanding of various phenomena improves, the

treatment of these processes can be overhauled within the semi-analytical machinery. The models make direct contact with observations, predicting broad band magnitudes which can be subjected to the same colour selection criteria as the data.

I have presented a very incomplete survey of the predictions of current models for the number and nature of massive galaxies, both locally and at high redshift. The models are in reasonably good shape, matching the number of massive galaxies today and the observed abundance of various populations of red galaxies.

However, some outstanding challenges to the models remain, including:

1. The growth of brightest cluster galaxies. Stott *et al.* (2010) argue that the stellar mass of brightest cluster galaxies has changed little since $z \sim 1.5$ (see also Collins *et al.* 2009). Semi-analytical models predict that the stellar mass of such galaxies should change over this interval, even if this is solely due to mergers between progenitors, with little in situ star formation (De Lucia & Blaizot 2007). The resolution of this discrepancy may need the treatment of cluster galaxies to be extended in the models to include the modelling of the tidal disruption of the stars and cold gas in satellite galaxies and the build-up of intra-cluster light (Tecce *et al.* 2010).

2. The abundance of star forming galaxies at high redshift. Semi-analytical models successfully reproduce the luminosity function in the rest-frame ultra-violet, estimated from samples of galaxies constructed using the Lyman break drop-out technique (Lacey *et al.* 2011). The predicted dust extinction is substantial at such short wavelengths and difficult to estimate empirically; nevertheless, the colours of model Lyman-break galaxies match those observed (González-Pérez *et al.* 2013). Matching the abundance *and* redshift distribution of galaxies whose star formation is enshrouded by dust is much more challenging. To date, this has only been possible (at the same time as reproducing the properties of present-day galaxies) by invoking a top-heavy IMF in bursts of star formation (Baugh *et al.* 2005; Lacey *et al.* 2008). This assumption continues to generate controversy, despite building circumstantial evidence in favour of IMF variation.

3. The chemical composition of massive ellipticals. There is a tight correlation between the α -element content of elliptical galaxies and their velocity dispersion (e.g the Mg- σ relation measured by Bender, Burstein & Faber 1992). This relation has been interpreted as a constraint on the duration of star formation in ellipticals (e.g. Thomas, Greggio & Bender 1999) as α -element production is a tracer of Type-II supernovae and the iron content is a probe of Type-Ia supernovae, which occur a much longer time after the star formation episode. The α to Fe ratio is also sensitive to the form of the IMF (Nagashima *et al.* 2005). A recent study of the mass-metallicity relation in the current generation of semi-analytical models has highlighted a problem in reproducing the steepness of the observed relation, though this is sensitive to the form of IMF adopted and the modelling of supernova feedback (De Lucia & Borgani 2012).

Continued progress in the subject will come from a combination of approaches, including semi-analytical modelling, gas dynamics simulations, both of cosmological volumes and of the formation and evolution of individual galaxies, and ever improving observational data which can guide the development of the models.

References

- Almeida, C., *et al.* 2008, *MNRAS*, 386, 2145
 Bender, R., Burstein, D., & Faber, S. M. 1992, *ApJ*, 399, 462
 Baugh, C. M., Cole, S., & Frenk, C. S. 1996, *MNRAS*, 283, 1361
 Baugh, C. M., *et al.* 2005, *MNRAS*, 356, 1191
 Baugh, C. M. 2006, *Rep Prog Phys*, 69, 3101

- Bell, E. F., *et al.* 2003, *MNRAS*, 343, 367
Bell, E. F., *et al.* 2003, *ApJS*, 149, 289
Bell, T. A., *et al.* 2006, *MNRAS*, 371, 1865
Benson, A. J. 2010, *Phys Rep*, 495, 33
Blitz, L. & Rosolowsky, E. 2006, *ApJ*, 650, 933
Bower, R. G., *et al.* 2006, *MNRAS*, 370, 645
Bruzual, G. & Charlot, S. 2003, *MNRAS*, 344, 1000
Cole, S., 1991, *ApJ*, 367, 45
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, *MNRAS*, 319, 168
Cole, S., *et al.* 2001, *MNRAS*, 326, 255
Collins, C. A., *et al.* 2009, *Nature*, 458, 603
Cook, M., *et al.* 2010, *MNRAS*, 402, 941
Daddi, E., *et al.* 2004, *ApJ*, 617, 746
De Lucia, G., *et al.* 2006, *MNRAS*, 366, 499
De Lucia, G. & Blaizot J. 2007, *MNRAS*, 375, 2
De Lucia, G. & Borgani, S. 2012, *MNRAS*, 426, L61
Efstathiou, G. 2003, Talk in memory of Fred Hoyle, (arXiv:astro-ph/0303623)
Eisenstein, D. J., *et al.* 2001, *AJ*, 122, 2267
Fontanot, F., *et al.* 2007, *MNRAS*, 382, 903
Fu, J., Kauffmann, G., & Krumholz, M. R. 2010, *MNRAS*, 409, 515
González-Pérez V., *et al.* 2009, *MNRAS*, 398, 497
González-Pérez V., *et al.* 2013, *MNRAS*, 429, 1609
Granato, G. L., *et al.* 2004, *ApJ*, 600, 580
Guo, Q., *et al.* 2011, *MNRAS*, 413, 101
Kang, X. & Jing, Y. P., Silk J. 2006, *ApJ*, 648, 820
Kauffmann, G. 1996, *MNRAS*, 281, 487
Lacey, C. G. & Silk, J. 1991, *ApJ*, 381, 14
Lacey, C. G., *et al.* 2008, *MNRAS*, 385, 1155
Lacey, C. G., *et al.* 2011, *MNRAS*, 412, 1828
Lagos, C., Del, P., *et al.* 2011a, *MNRAS*, 416, 1566
Lagos, C., Del, P., *et al.* 2011b, *MNRAS*, 418, 1649
Lagos, C., Del, P., *et al.* 2012, *MNRAS*, 426, 2142
Li, C. & White, S. D. M. 2009, *MNRAS*, 398, 2177
Maraston, C. 2005, *MNRAS*, 362, 799
Maraston, C., *et al.* 2006, *ApJ*, 652, 85
Marchesini, D., *et al.* 2009, *ApJ*, 701, 1765
Masjedi, M., *et al.* 2006, *ApJ*, 644, 54
Menci, N., *et al.* 2006, *ApJ*, 647, 753
Merson, A. I., *et al.* 2013, *MNRAS*, 429, 556
Nagashima, M., *et al.* 2005, *MNRAS*, 363, L31
Pforr, J., Maraston, C., & Tonini, C. 2012, *MNRAS*, 422, 3285
Smith, G. P., *et al.* 2002, *MNRAS*, 333, L16
Stott, J. P., *et al.* 2010, *ApJ*, 718, 23
Tecce, T. E., *et al.* 2010, *MNRAS*, 408, 2008
Thomas, D., Greggio, L., & Bender, R. 1999, *MNRAS*, 302, 537
White, S. D. M. & Rees, M. J. 1978, *MNRAS*, 183, 341
White, S. D. M. & Frenk, C. S. 1991, *ApJ*, 379, 52