

# “Ab initio” models of galaxy formation: successes and open problems

Gabriella De Lucia<sup>1</sup>

*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748 Garching, Germany*

**Abstract.** In the past decades, different approaches have been developed in order to link the physical properties of galaxies to the dark matter haloes in which they reside. In this review, I give a brief overview of methods, aims, and limits of these techniques, with particular emphasis on semi-analytic models of galaxy formation. For these models, I also provide a brief summary of recent successes and open problems.

**Keywords:** Galaxy formation – Galaxy evolution

**PACS:** 98.62.Ai; 98.62.Ve

## INTRODUCTION

During the last decades, a number of observational tests have converged to establish the  $\Lambda$ CDM model as the *de facto* standard cosmological model for structure formation. Over the past years, it has been shown that this model can account simultaneously for the present acceleration of cosmic expansion as inferred from supernovae explosions [1], the structure seen in the  $z = 3$  Ly $\alpha$  forest [2], the power spectrum of low redshift galaxies [3], the most recent measurements of the microwave background fluctuations [4], and a number of other important observational constraints. Although the candidate particle for the non-baryonic dark matter has yet to be detected in the laboratory, and the nature of dark energy remains unknown, the fundamental cosmological parameters are now known with uncertainties of only a few per cent, removing a large part of the parameter space in galaxy formation studies.

While the basic theoretical paradigm for structure formation appears to be well established, our understanding of the physical processes that lead to the variety of observed galaxy properties is still far from complete. Although I have kept the word *ab initio* in the title of this review, as suggested by the organizers of this meeting, I would like to stress that *ab initio* treatments of the galaxy formation process are very difficult - if not unfeasible - simply because we do not have a complete understanding of the many different and complex physical processes which are at play.

Today’s models of galaxy formation find their seeds in the pioneering work by White and Rees [5] who proposed that galaxies form when gas condenses at the centre of dark matter haloes, following the radiative cooling of baryons. In the following years, three different approaches have been developed in order to link the observed properties of luminous galaxies to the dark matter haloes in which they reside.

---

<sup>1</sup> Present address: INAF - Astronomical Observatory of Trieste, via Tiepolo 11, I-34143 Trieste, Italy

In **semi-analytic models** of galaxy formation, which I will discuss in more detail in the following, the evolution of the baryonic components of galaxies is modelled using *simple yet physically and/or observationally motivated prescriptions*. Modern semi-analytic models take advantage of high-resolution N-body simulations to specify the location and evolution of dark matter haloes, which are assumed to be the birth-places of luminous galaxies. Since pure N-body simulations can handle very large number of particles, this approach can access very large dynamic ranges in mass and spatial resolution. In addition, the computational costs are limited so that the method allows a fast exploration of the parameter space and an efficient investigation of different specific physical assumption.

**Direct hydrodynamical simulations** provide an explicit description of gas dynamics. As a tool for studying galaxy formation, it is worth reminding that these methods are still limited by relatively low mass and spatial resolution, and by computational costs that are still prohibitive for simulations of galaxies throughout large volumes. In addition, and perhaps more importantly, complex physical processes such as star formation, feedback, etc. still need to be modelled as *sub-grid physics*, either because the resolution of the simulation becomes inadequate or because (and this is almost always true) we do not have a ‘complete theory’ for the particular physical process under consideration.

A third approach - usually referred to as the **Halo Occupation Distribution (HOD) models** - has become popular in more recent years. This method essentially bypasses any explicit modelling of the physical processes driving galaxy formation and evolution, and specifies the link between dark matter haloes and galaxies in a purely statistical fashion. The method is conceptually very simple and easy to implement, and it can be constrained using the increasing amount of available information on clustering properties of galaxies at different cosmic epochs. It remains difficult, however, to move from a purely statistical characterization of the link between dark matter haloes and galaxies to a more physical understanding of the galaxy formation process itself.

Clearly, each of these methods has its own advantages and weaknesses, and they should be viewed as *complementary rather than competitive*. In the following, I will focus on semi-analytic models of galaxy formation. In particular, I will provide a brief overview of the methods, aims, and limits of these techniques, and give a brief summary of their recent successes and open problems.

## METHODS, AIMS, AND LIMITS

The backbone of any semi-analytic model is provided by what in the jargon is called a dark matter ‘merger tree’, which essentially provides a representation of the assembly history of a dark matter halo. Early renditions of semi-analytic models - but this is still the case for a large number of applications today - took advantage of the extended Press-Schechter (EPS) formalism [6, 7] and Monte Carlo methods to construct representative histories of merger trees leading to the formation of haloes of a given mass. It is important to note that some recent work has demonstrated that this formalism might not provide an adequate description of the merger trees extracted directly from numerical simulations [8, 9, 10]. Although some of these studies have provided ‘corrections’ to analytic merger trees, many applications are still carried out using the classical EPS

formalism, and little work has been done so far to understand to which measure this can affect predictions of galaxy formation models.

As mentioned earlier, modern semi-analytic models (sometimes referred to as *hybrid* models) take advantage of high-resolution N-body simulations to follow the evolution of dark matter haloes in its full geometrical complexity [11, 12]. On a next level of complexity, some more recent implementations have explicitly taken into account dark matter substructures, i.e. the haloes within which galaxies form are still followed when they are accreted onto a more massive system [13, 14]. There is one important caveat to bear in mind regarding these methods: dark matter substructures are fragile systems that are rapidly and efficiently destroyed below the resolution limit of the simulation [15, 16]. Since the baryons are more concentrated than dark matter, it is to be expected that the baryonic component will be more resistant to the tidal stripping that reduces its parent halo mass. This creates a complex and strong position-dependent relation between dark matter substructures and galaxies, contrary to what was assumed in early HOD models. In addition, this treatment introduces a complication due to the presence of ‘orphan galaxies’, i.e. galaxies whose parent substructure mass has been reduced below the resolution limit of the simulation. In most of the available semi-analytic models, these galaxies are assumed to merge onto the corresponding central galaxies after a residual merging time which is given by some variation of the classical dynamical friction formula. Only a few models account for the stripping of stars due to tidal interactions with the parent halo.

Once the backbone of the model is constructed, using either N-body simulations or analytic methods, galaxy formation and evolution is ‘coupled’ to the merger trees using a set of analytic laws that are based on theoretical and/or observational arguments, to describe complex physical processes like star formation, supernovae and AGN feedback processes, etc. Adopting this formalism, it is possible to express the full galaxy formation process through a set of differential equations that describe the variation in mass of the different galactic components (e.g. gas, stars, metals). Given our limited understanding of the physical processes at play, these equations contain *free parameters* whose value is typically chosen in order to provide a reasonably good agreement with observational data in the local Universe.

One common criticism to semi-analytic models is that there are *too many* free parameters. It should be noted, however, that the number of these parameters is not larger than the number of published comparisons with different and independent sets of observational data, for any of the semi-analytic models discussed in the recent literature. In addition, these are not ‘statistical’ parameters but, as explained above, they are due to our lack of understanding of the physical processes considered. A change in any of these parameters has consequences on a number of different predictable properties, so that often there is little parameter degeneracy for a given set of prescriptions. Finally, observations and theoretical arguments often provide important constraints on the range of values that different parameters can assume.

One great advantage of hybrid methods with respect to classical techniques based on the EPS formalism, is that they provide full dynamical information about model galaxies. Using this approach, it becomes possible to construct realistic *mock catalogues* that contain not only physical information about model galaxies such as masses, star formation rates, luminosities, etc. but also dynamically consistent redshift and spatial

information, like in real redshift surveys. Using these mock catalogues, it is then possible to carry-out detailed comparisons with observational data at different cosmic epochs. These comparisons provide useful information on the relative importance of different physical processes in establishing a certain observational trend, and on the physics which is eventually missing in these models.

## RECENT SUCCESSES AND OPEN PROBLEMS

In discussing recent successes of semi-analytic models, I will start from the most fundamental description of the galaxy population: the galaxy luminosity function. Since early implementations of semi-analytic techniques, it was clear that a relatively strong supernovae feedback was needed in order to suppress the large excess of faint galaxies, due to the steep increase of low-mass dark matter haloes [17, 18]. It is interesting to note that matching the faint end of the luminosity function comes at the expenses of exacerbating the excess of luminous bright objects, due to the fact that the material reheated and/or ejected by low-mass galaxies ends up in the hot gas associated to central galaxies of relatively massive haloes. At later times, this material cools efficiently onto the corresponding central galaxies increasing their luminosities and star formation rates, at odds with observational data.

Matching the bright end of the luminosity function has proved difficult for a long time, and a good match has been achieved only recently using a relatively strong form of ‘radio-mode’ AGN feedback [19, 20]. Different prescriptions of AGN feedback have been proposed, and still much work remains to be done in order to understand if and how the energy injected by intermittent radio activity at the cluster centre is able to efficiently suppress the cooling flows. Recent observational measurements indicate that the ensemble-averaged power from radio galaxies seems sufficient to offset the mean level of cooling [21]. It is, however, important to note that not every cluster shows central radio activity, and that the steep dependence of the radiative cooling function on density makes it difficult to stabilize cooling flows at all radii.

The main reason for the success of the ‘radio-mode’ AGN feedback is that it is not connected to star formation, so that its implementation permits at the same time to suppress the luminosity of massive galaxies and to keep their stellar populations old [22]. Therefore these models seem to reproduce, at least qualitatively, the observed trend for more massive ellipticals to have shorter star formation time-scales. A good quantitative agreement has not been shown yet and is complicated by large uncertainties associated to star formation histories extracted from observed galaxy spectra.

The suppression of late cooling (and therefore star formation) does not affect, however, the assembly history of massive galaxies for which models predict an increase in stellar mass by a factor 2 to 4 since  $z \sim 1$ , depending on stellar mass [22, 23]. This creates a certain tension with the observation that the massive end of the galaxy mass function does not appear to evolve significant over the same redshift interval. Part of this tension is removed when taking into account observational errors and uncertainties on galaxy mass estimates [24, 25, see also Monaco these proceedings]. For the mass assembly of the brightest cluster galaxies (BCGs), the situation is worse: while observations seem consistent with no mass growth since  $z \sim 1$ , models predict an increase in

mass by a factor about 4 [23, 26]. One major caveat in this comparison is given by the fact that observational studies usually adopt fixed metric aperture magnitudes (which account for 25-50 per cent of the total light contained in the BCG and intra-cluster light), while models compute total magnitudes. Semi-analytic models do not provide information regarding the spatial distribution of the BCG light, so aperture magnitudes cannot be calculated. In addition, most of the available models do not take into account the stripping of stars from other cluster galaxies due to tidal and harassment effects [27, 28].

Most of the models currently available exhibit a remarkable degree of agreement with a large number of observations for the galaxy population in the local Universe (e.g. the observed relations between stellar mass, gas mass, and metallicity; the observed luminosity, colour, and morphology distribution; the observed two-point correlation functions). When analysed in detail, however, some of these comparisons show important and systematic (i.e. common to most of the semi-analytic models discussed in the literature) disagreements.

Although models are not usually tuned to match observations of galaxy clustering, they generally reproduce the observed dependence of clustering on magnitude or colour. The agreement appears particularly good for the dependence on luminosity, while the amplitude difference on colour appears greater in the models than observed [29, 30]. This problem might be (at least in part) related to the excess of small red satellite galaxies which plagues all models discussed in the recent literature (e.g. see Fig. 11 in [19] and Fig. 4 in [31]; see also Monaco these proceedings).

A generic excess of intermediate to low-mass galaxies has been discussed by Fontana et al. [32]. At low redshift, this excess is largely due to satellite galaxies that were formed and accreted early on, and that are dominated by old stellar populations. Semi-analytic models assume that when a galaxy is accreted onto a larger structure, the gas supply can no longer be replenished by cooling that is suppressed by an instantaneous and complete stripping of the hot gas reservoir. Since this process (commonly referred to as ‘strangulation’) is usually combined with a relatively efficient supernovae feedback, galaxies that are accreted onto a larger system consume their gas very rapidly, moving onto the red-sequence on quite short time-scales [33, 34, 35]. This contributes to produce an excess of faint and red satellites and a transition region (sometimes referred to as ‘green valley’) which does not appear to be as well populated as observed.

Much effort has been recently devoted to this problem. McCarthy et al. [36] have used high resolution hydrodynamic simulations to show that galaxies are able to retain a significant fraction of their hot haloes following virial crossing. Font et al. [37] incorporated a simple model based on these simulations within the Durham semi-analytic model. With this modification, a larger fraction of satellites has bluer colours, resulting in a colour distribution that is in better (but not perfect) agreement with observational data.

The completion of new high-redshift surveys has recently pushed comparisons between model results and observational data to higher redshift. I do not have time here to discuss in detail all agreements and disagreements discussed in the recent literature. I would like to stress, however, that this still rather unexplored regime for modern models is very interesting because it is at high redshift that predictions from different models differ more dramatically.

To close this section, I would like to remind that a long standing problem for hierarchical models has been to match the zero-point of the Tully-Fisher relation (the observed correlation between the rotation speed and the luminosity of spiral disks [38]) while reproducing at the same time the observed luminosity function. As discussed in Baugh [39], no model with a realistic calculation of galaxy size has been able to match the zero-point of the Tully-Fisher relation using the circular velocity of the disk measured at the half mass radius. It remains unclear if this difficulty is related to some approximation in the size calculation, or if it is related to more fundamental shortcomings of the cold dark matter model.

## CONCLUSIONS

Given our poor knowledge of most of the physical processes at play, *ab initio* treatments of the galaxy formation process are extremely difficult, if not unfeasible. In the past decades, we have developed a number of *techniques* to study galaxy formation within the currently standard cosmological model. Semi-analytic models represent the most developed of these techniques to make detailed predictions of galaxy properties at different cosmic epochs.

These models are not meant to be definitive. Rather, they need to be *falsified* against observational data, in order to gain insight on the relative importance of different physical processes, and on the physics which is eventually still missing in the models. When comparing model results with observational data, it is important to take into account observational errors and biases which are eventually introduced by a particular observational selection and/or strategy. To this aim, realistic mock catalogues can be constructed by coupling semi-analytic techniques with large cosmological N-body simulations. Recently, a number of model results have been made publicly available in the context of the modern concept of ‘Theoretical Virtual Observatory’<sup>2</sup>. Considerable interest has been shown by the astronomical community, and a large number of papers using the public database have already been published, resulting in a rapid refinement and verification of theoretical modelling.

The largest success of these techniques is to have shown that *we can study galaxy formation within the currently established hierarchical paradigm*. The largest failure is, unsurprisingly, that we have not yet solved the galaxy formation problem. Undoubtedly, however, we have learnt a great deal about how galaxies form and evolve, and how their physical properties are related to the dark matter haloes in which they reside.

## ACKNOWLEDGMENTS

I wish to thank the organizers of the conference for the invitation, and for the pleasant and stimulating atmosphere.

---

<sup>2</sup> A link to available galaxy catalogues and a documentation can be found at <http://www.mpa-garching.mpg.de/millennium/>

## REFERENCES

1. S. Perlmutter, and The Supernova Cosmology Project, *ApJ* **517**, 565–586 (1999).
2. R. Mandelbaum, P. McDonald, U. Seljak, and R. Cen, *MNRAS* **344**, 776–788 (2003).
3. W. J. Percival, R. C. Nichol, D. J. Eisenstein, J. A. Frieman, M. Fukugita, J. Loveday, A. C. Pope, D. P. Schneider, A. S. Szalay, M. Tegmark, M. S. Vogeley, D. H. Weinberg, I. Zehavi, N. A. Bahcall, J. Brinkmann, A. J. Connolly, and A. Meiksin, *ApJ* **657**, 645–663 (2007).
4. E. Komatsu, J. Dunkley, M. R.olta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, S. S. Meyer, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright (2008), *ArXiv*: 0803.0547.
5. S. D. M. White, and M. J. Rees, *MNRAS* **183**, 341–358 (1978).
6. J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser, *ApJ* **379**, 440–460 (1991).
7. R. G. Bower, *MNRAS* **248**, 332–352 (1991).
8. A. J. Benson, M. Kamionkowski, and S. H. Hassani, *MNRAS* **357**, 847–858 (2005).
9. Y. Li, H. J. Mo, F. C. van den Bosch, and W. P. Lin, *MNRAS* **379**, 689–701 (2007).
10. S. Cole, J. Helly, C. S. Frenk, and H. Parkinson, *MNRAS* **383**, 546–556 (2008).
11. G. Kauffmann, J. M. Colberg, A. Diaferio, and S. D. M. White, *MNRAS* **303**, 188–206 (1999).
12. A. J. Benson, S. Cole, C. S. Frenk, C. M. Baugh, and C. G. Lacey, *MNRAS* **311**, 793–808 (2000).
13. V. Springel, S. D. M. White, G. Tormen, and G. Kauffmann, *MNRAS* **328**, 726–750 (2001).
14. G. De Lucia, G. Kauffmann, and S. D. M. White, *MNRAS* **349**, 1101–1116 (2004).
15. G. De Lucia, G. Kauffmann, V. Springel, S. D. M. White, B. Lanzoni, F. Stoehr, G. Tormen, and N. Yoshida, *MNRAS* **348**, 333–344 (2004).
16. L. Gao, S. D. M. White, A. Jenkins, F. Stoehr, and V. Springel, *MNRAS* **355**, 819–834 (2004).
17. S. D. M. White, and C. S. Frenk, *ApJ* **379**, 52–79 (1991).
18. A. J. Benson, R. G. Bower, C. S. Frenk, C. G. Lacey, C. M. Baugh, and S. Cole, *ApJ* **599**, 38–49 (2003).
19. D. J. Croton, V. Springel, S. D. M. White, G. De Lucia, C. S. Frenk, L. Gao, A. Jenkins, G. Kauffmann, J. F. Navarro, and N. Yoshida, *MNRAS* **365**, 11–28 (2006).
20. R. G. Bower, A. J. Benson, R. Malbon, J. C. Helly, C. S. Frenk, C. M. Baugh, S. Cole, and C. G. Lacey, *MNRAS* **370**, 645–655 (2006).
21. P. N. Best, A. von der Linden, G. Kauffmann, T. M. Heckman, and C. R. Kaiser, *MNRAS* **379**, 894–908 (2007).
22. G. De Lucia, V. Springel, S. D. M. White, D. Croton, and G. Kauffmann, *MNRAS* **366**, 499–509 (2006).
23. G. De Lucia, and J. Blaizot, *MNRAS* **375**, 2–14 (2007).
24. M. G. Kitzbichler, and S. D. M. White, *MNRAS* **376**, 2–12 (2007).
25. M. J. Stringer, A. J. Benson, K. Bundy, R. S. Ellis, and E. L. Quetin (2008), *ArXiv*: 0806.2232.
26. I. M. Whiley, A. Aragón-Salamanca, G. De Lucia, A. von der Linden, S. P. Bamford, P. Best, M. N. Bremer, P. Jablonka, O. Johnson, B. Milvang-Jensen, S. Noll, B. M. Poggianti, G. Rudnick, R. Saglia, S. White, and D. Zaritsky, *MNRAS* **387**, 1253–1263 (2008).
27. P. Monaco, G. Murante, S. Borgani, and F. Fontanot, *ApJL* **652**, L89–L92 (2006).
28. C. Conroy, R. H. Wechsler, and A. V. Kravtsov, *ApJ* **668**, 826–838 (2007).
29. V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce, *Nature* **435**, 629–636 (2005).
30. A. L. Coil, J. A. Newman, D. Croton, M. C. Cooper, M. Davis, S. M. Faber, B. F. Gerke, D. C. Koo, N. Padmanabhan, R. H. Wechsler, and B. J. Weiner, *ApJ* **672**, 153–176 (2008).
31. R. S. Somerville, P. F. Hopkins, T. J. Cox, B. E. Robertson, and L. Hernquist, *MNRAS* **391**, 481–506 (2008).
32. A. Fontana, S. Salimbeni, A. Grazian, E. Giallongo, L. Pentericci, M. Nonino, F. Fontanot, N. Menci, P. Monaco, S. Cristiani, E. Vanzella, C. de Santis, and S. Gallozzi, *A&A* **459**, 745–757 (2006).
33. S. M. Weinmann, F. C. van den Bosch, X. Yang, H. J. Mo, D. J. Croton, and B. Moore, *MNRAS* **372**, 1161–1174 (2006).
34. G. De Lucia, “How ‘Hereditry’ and ‘Environment’ Shape Galaxy Properties,” in *Cosmic Frontiers*, edited by N. Metcalfe, and T. Shanks, 2007, vol. 379 of *Astronomical Society of the Pacific Conference Series*, pp. 257–+.

35. L. Wang, C. Li, G. Kauffmann, and G. De Lucia, *MNRAS* **377**, 1419–1430 (2007).
36. I. G. McCarthy, C. S. Frenk, A. S. Font, C. G. Lacey, R. G. Bower, N. L. Mitchell, M. L. Balogh, and T. Theuns, *MNRAS* **383**, 593–605 (2008).
37. A. S. Font, R. G. Bower, I. G. McCarthy, A. J. Benson, C. S. Frenk, J. C. Helly, C. G. Lacey, C. M. Baugh, and S. Cole, *MNRAS* **389**, 1619–1629 (2008).
38. R. B. Tully, and J. R. Fisher, *A&A* **54**, 661–673 (1977).
39. C. M. Baugh, *Reports of Progress in Physics* **69**, 3101–3156 (2006).