Modelling the properties of galaxies and clusters

Kay, Scott Thomas

Modelling the Properties of Galaxies and Clusters

Scott Thomas Kay

A thesis submitted to the University of Durham
in accordance with the regulations for
admittance to the Degree of Doctor of Philosophy.

Department of Physics
University of Durham
June 5, 2000

Abstract

Modelling the Properties of Galaxies and Clusters

by Scott Thomas Kay

This thesis examines various properties of galaxies and clusters within hierarchical models of structure formation.

A simple model based on analytical scaling relations is applied to X-ray observations of clusters at low and high redshift, in an attempt to constrain cosmological parameters from their evolutionary properties. It is found that the density parameter, Ω_0 cannot be constrained using the data alone. Two independent constraints on the slope of the linear power spectrum, infer values of $\Omega_0 < 0.7$ at 95 per cent confidence.

The remainder of this thesis concentrates on the method of cosmological simulation, a self-consistent approach to the modelling of structure in the Universe. A parameter-space study is performed for the simplest model of galaxy formation: the radiative cooling of baryons within the cores of dark matter haloes. It is found that the properties of the galaxies in the simulations are insensitive to the range of parameters studied, with the exception of those that affect the cooling rate of the gas. For modest resolution and reasonable choices of physical parameters, the amount of baryons in galaxy material is around a factor of 2 too high.

An investigation is then performed for including the effects of star formation and energy from supernovae (feedback) within cosmological simulations, to reduce the amount of gas that cools. The star formation rate is driven by the minimum density for which the stars formation occurs and, for high star formation efficiencies, is limited by the cooling rate of the gas. A successful model for feedback is found to require the prevention of reheated gas from cooling for a short period of time, as an attempt to mimick the properties of a multiphase medium.

Finally, preliminary results are presented for simulations of a galaxy cluster, including the effects of radiative cooling, star formation and feedback. The properties of the cluster are found to vary significantly between models with and without feedback, due to the feedback reducing the star formation rate by reheating gas that cools.

Declaration

The work described in this thesis was undertaken between 1996 and 1999 whilst the author was a research student under the supervision of Professor C.S. Frenk in the Department of Physics at the University of Durham. This work has not been submitted for any other degree at the University of Durham or at any other University. The work in Chapter 2 was accomplished in collaboration with Dr R.G. Bower. Chapter 4 was carried out in collaboration with Dr F.R.Pearce, Dr A. Jenkins, Prof. C.S. Frenk, Prof. S.D.M. White, Dr P.A. Thomas and Prof. H.M.P. Couchman. Chapters 5 and 6 benefitted considerably from discussions with Prof. C.S. Frenk, Dr F.R. Pearce and Dr A. Jenkins. However, the majority of the material presented here is the author's own work.

Portions of the work presented here have appeared in the following papers:

- Kay S.T., Bower R.G., 1999, MNRAS, 308, 664 (Chapter 2)
- Kay S.T., Pearce F.R., Jenkins A., Frenk C.S., White S.D.M., Thomas P.A., Couchman H.M.P., 2000, MNRAS, in press (Chapter 4)

To my parents and Hilary.

Acknowledgements

Sincere thanks go to my supervisor, Carlos Frenk, for his wealth of ideas and encouraging support over the years. I would also like to express my gratitude to Richard Bower, Frazer Pearce and Adrian Jenkins for the large part they played in shaping my understanding of all things numerical and astrophysical.

My acknowledgments would not be complete without thanking all of the people who helped made doing my PhD more worthwhile: Scott, Dave, Doug, Steve, Simon, Ale, Henry, Russell, Luis, Andrew, Mikey, Peder, Fiona, James, Cesario, Katherine, Roger, Mark - far too many reasons to mention why.

Contents

Chapter 1

Introduction

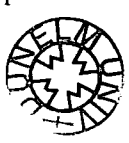
1.1 Background

Understanding the origin and nature of structure in the Universe is one of the fundamental aims of modern cosmology. It is only over the past few decades that a "standard model" for structure formation has started to emerge.

Cosmology, as a science, is a very young subject. Our understanding of the Universe was revolutionised in the early twentieth century, when two important advances were made: one theoretical and the other observational. The first was Einstein's General Relativity, a new theory for the force of gravity. General Relativity predicts that the curvature of spacetime is directly related to the amount of mass-energy within it. On cosmological scales, gravity is the strongest force, so the geometrical properties of the Universe must be determined by its content.

Around the same time as Einstein's discovery, Hubble and his colleagues confirmed that several observed faint nebulae were actually large aggregates of stars (galaxies), external to our own Milky Way. The galaxies were also observed to be receding with velocities that increased with their distance from us (Hubble's Law; Hubble & Humason 1931), regardless of position on the sky. (Of course, this is only true in an average sense, otherwise galaxy clusters would not exist!) Hubble's observations led to the idea, known as the Cosmological Principle, that the Universe was isotropic (the same in all directions) and homogeneous (the same at all positions) on large scales. (Homogeneity is a conceptual idea, based on the Copernican Principle, that we do not inhabit a special position in the Universe. However, it naturally explains Hubble's Law, and as was argued by Milne, it would be difficult to conceive of a Universe with physical properties that varied from place to place.)

Both General Relativity and the Cosmological Principle led to the formation of theories for the dynamical evolution of the Universe. The most popular of these was the Big Bang model,



based on the idea that the Universe is expanding from a hot dense initial state, much smaller than an atomic nucleus. Such a theory implies that the Universe had a violent past, resulting in several key predictions for observed properties at the present. However, many people were opposed to the idea of a Big Bang, and an alternative theory was proposed, based on the idea that the Universe should remain the same at all times (the Perfect Cosmological Principle). The so-called Steady State theory (e.g. Hoyle 1954), satisfies the Perfect Cosmological Principle by assuming that matter is created at a continuous rate that is sufficient to preserve the mean density of the Universe. The Steady State model has since fallen out of favour, largely due to the observation of a diffuse background radiation in the microwave region of the electromagnetic spectrum (the cosmic microwave background; Penzias & Wilson 1965), predicted by the Big Bang model as a relic from the hot dense past, when the density of the Universe had dropped sufficiently to allow the photons to become decoupled from matter in the Universe (the epoch of last scattering). The Steady State theory, by definition, makes no such prediction.

Another major success of the Big Bang model is its prediction for the abundances of light elements that formed before the first stars (nucleosynthesis). The predicted values are in remarkable agreement with observations (e.g. Copi, Schramm, & Turner 1995).

Despite its triumphs, the Big Bang model is not without its flaws. An example, of particular importance to theories of structure formation, is known as the horizon problem. The remarkable discovery of tiny fluctuations in the temperature of the microwave background by the *COBE* satellite (Smoot et al. 1992) was crucial evidence of evolution in the Universe. These fluctuations were imprinted by primordial density perturbations at the epoch of last-scattering, which have subsequently grown to form the galaxies and clusters observed today. However, temperature fluctuations were observed by *COBE* on angular scales much larger than the regions able to come into contact with one another (known as the horizon) at the epoch of last scattering. The Big Bang model provides no explanation for this.

A breakthrough came from the theory of *inflation*, (Guth & Pi 1982 and references therein), an extension to the Big Bang model that solves the horizon problem amongst others. Inflation predicts that a region much larger than our observable Universe underwent a brief period of exponential expansion a short time after the Big Bang. One of the consequences of inflation is that it allows the horizon scale to be significantly expanded, allowing the observable region of the Universe to come into causal contact early on. Another consequence of inflation is that it provides an explanation for the origin of structure itself. Inflation allows tiny irregularities, such as quantum fluctuations, in the energy-density to be amplified to the sizes inferred from *COBE* observations.

The nature of this structure cannot be completely baryonic. Nucleosynthesis places an upper limit on the baryon density in the Universe that is around an order of magnitude lower than the total contribution, inferred from a range of observational studies (e.g. see Dekel, Burstein, & White 1997 and references therein). Hence, the majority of material in the Universe must be non-baryonic and dark. Two candidate theories of structure formation emerged, based on the properties of the dark matter. The first, known as the Hot Dark Matter (HDM) model predicts that the dark matter consists of relativistic particles, such as neutrinos, whose thermal motions smooth out density fluctuations on scales below ~ 50 Mpc (considerably larger than the scale of galaxy clusters). Structure in the HDM model forms "top-down", i.e. large scales collapse first, subsequently fragmenting into smaller objects.

An alternative model, based on Cold Dark Matter (CDM) was proposed for which there are currently no known particle candidates, although it was heavily influenced by advances made in particle physics theories, such as super-symmetry. Examples of possible CDM particles are the axion or the gravitino. CDM is slow-moving (hence its name) with no absolute cut-off scale for the density fluctuations. Structure in this model forms hierarchically ("bottom-up"), i.e. the smallest fluctuations collapse and form objects first, which merge to form larger objects.

The CDM model is currently the most popular paradigm for structure formation, mainly due to comparative studies of the clustering properties of HDM and CDM, using N-body simulations (see Frenk 1991 and references therein). The HDM model predicts a clustering scale that is far too high to plausibly relate to the clustering properties of galaxies observed in the real Universe. (The relation between the mass distribution and the galaxy distribution is historically known as *bias*.) The CDM model was able to achieve the required clustering scale using only a simple bias prescription, based on the initial peaks of the density field. A huge amount of work has been done that is based on the CDM paradigm, mainly using N-body simulations. Of particular importance is the relationship between dark matter haloes (the basic elements of virialised structure in the CDM model) and the formation of galaxies and clusters of galaxies. The haloes provide a natural "framework" for the galaxies and clusters. As shown by White & Rees (1978), the characteristic densities and temperatures of galaxy-size haloes co-incide with gas cooling times of the order of the age of the Universe or shorter, where cluster-sized haloes have cooling times much longer than the age of the Universe. This at least hints at the widely different properties of galaxies, mainly consisting of stars, and clusters, where the majority of baryons are a hot diffuse intracluster medium.

Self-consistent approaches to the modelling of baryons within dark matter haloes are presently our "best bet" for understanding the properties of galaxies and clusters. The overall aim of this thesis is to investigate various aspects of the galaxy and cluster formation process using one of these approaches: the method of numerical simulation. Beginning with the rest of this chapter, the basic mathematical elements of cosmology are outlined, that are of relevance to the work presented in this thesis. Chapter 2 draws on these results to measure cosmological parameters using a simple model based on the X-ray evolution of clusters. The rest of this thesis concentrates on using numerical simulations to model galaxies and clusters within the CDM model. The method of *cosmological* simulation is described in Chapter 3. Chapters 4 and 5 apply this method to modelling the formation of galaxies within a small volume, and Chapter 6 presents results obtained from applying the method to modelling a cluster of galaxies. Conclusions are drawn in Chapter 7.

1.2 Basic cosmology

Applying the Cosmological Principle to General Relativity requires the construction of a metric that satisfies both homogeneity and isotropy. Such a metric (known as the Robertson-Walker metric) in its most general form is

$$\begin{aligned} ds^2 &= c^2dt^2 - a^2(t)\left[d\chi^2 + f(\chi)^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right] \\ &= c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-\kappa r^2} + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right] \end{aligned} \tag{1.1}$$

where χ, θ and ϕ are the spatial comoving co-ordinates, and t is *proper* time (measured from the Big-Bang), which must be the same everywhere. The function, $a(t)$ is the expansion factor and $r = f(\chi)$, where

$$f(\chi) = \begin{cases} \sin\chi & \kappa = +1; \\ \chi & \kappa = 0; \\ \sinh\chi & \kappa = -1. \end{cases} \tag{1.2}$$

The constant, κ, can take three values: $1, 0$ and -1, describing space with positive, zero (Euclidean) and negative curvature respectively.

The expansion factor is determined from the solution of the Friedmann equation,

$$\ddot{a} = -\frac{4}{3}\pi G\left(\rho + 3\frac{P}{c^2}\right)a + \frac{\Lambda}{3}c^2 a, \tag{1.3}$$

where P and ρ are the pressure and density source terms and Λ is the cosmological constant. The equation of state for the expanding fluid can be written as $P = w\rho c^2$, where $0 < w < 1$

is the Zel'dovich interval. For photons, $w = 1/3$ and for a pressureless fluid (dust), $w = 0$. Another implication of the Cosmological Principle is that the expansion is adiabatic, hence $\rho = \rho_0(a/a_0)^{-3(1+w)}$, where the subscript, 0 refers to the present day. For the range of epochs relevant to galaxies and clusters, the dust model provides a good approximation; the density of pressureless material evolves as $\rho \propto a^3$. The adiabatic assumption allows Equation 1.3 to be cast in the form

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 + \frac{\Lambda}{3}a^2 - Kc^2, \tag{1.4}$$

where K is the curvature, with $\kappa = K/\mathrm{mod}(K)$. A special model, known as the Einstein-de Sitter model, is the case of $K = \Lambda$=0. The density is just sufficient to close the Universe and is known as the critical density, $\rho_{\mathrm{cr}} = 3H^2/(8\pi G)$, where $H = \dot{a}/a$ is the Hubble parameter (the present value of this function is known as the Hubble constant, H_0). The density is commonly written in terms of the critical density, and is known as the density parameter, Ω. Defining $\Lambda_0 = \Lambda/(3H_0^2)$, the curvature of the Universe can be derived from the current value of the density parameter, cosmological constant and Hubble constant

$$K = \frac{H_0^2}{c^2}(\Omega_0 + \Lambda_0 - 1), \tag{1.5}$$

which also gives the current value of the expansion factor, since $a_0 = 1/\sqrt{|K|}$. Hence, a Euclidean universe satisfies $\Omega_0 + \Lambda_0 = 1$, a universe with positive spatial curvature (known as a *closed* universe) satisfies $\Omega_0 > 1 - \Lambda_0$ and a model with negative spatial curvature (an *open* universe) satisfies $\Omega_0 < 1 - \Lambda_0$.

Equation 1.4 can then be re-written as

$$\dot{a}^2 = H_0^2\left[\Omega_0(a^{-1} - 1) + \Lambda_0(a^2 - 1) + 1\right]. \tag{1.6}$$

Hence, values of Ω_0, Λ_0 and H_0 determine specific solutions to Equation 1.6.

An alternative variable to the expansion factor that is directly observable, is redshift, z. The redshift of a luminous source is measured as

$$z = \frac{\lambda_0 - \lambda_e}{\lambda_e}, \tag{1.7}$$

where λ_0 and λ_e are the wavelengths of a photon at the time when it was observed and emitted respectively. Redshift is directly related to the expansion factor, as $1 + z = a_0/a$. Hence, objects observed in our past light cone have values of $z > 0$.

Equations 1.1 and 1.6 allow physical quantities such as distances and lookback times to be calculated for arbitrary choices of cosmological parameters. For example, the angular diameter

distance to a luminous source is $d_a(z) = a_0\, f(\chi)$, where χ is calculated by considering the path taken by a photon emitted at a redshift of z

$$\chi(z) = \frac{c}{a_0}\int_0^z \frac{dz'}{H(z')}. \tag{1.8}$$

A related quantity to d_a is the luminosity distance to a source, simply defined as $d_L(z) = d_a(z)(1+z)$. Knowledge of d_a allows the comoving volume element, $dV(z)$, to be calculated, using

$$\frac{dV}{dz} = 4\pi \frac{c\, d_a(z)^2}{H(z)}, \tag{1.9}$$

assuming that the volume subtends the whole sky.

References

Copi C. J., Schramm D. N., Turner M. S., 1995, ApJ, 455, 95

Dekel A., Burstein D., White S. D. M., 1997, in Turok N., ed, Critical Dialogues in Cosmology (Princeton 250th Anniversary). World Scientific, p. 175

Frenk C. S., 1991, Physica Scripta, T36, 70

Guth A. H., Pi S. Y., 1982, Physical Review Letters, 49, 1110

Hoyle F., 1954, The Observatory, 74, 253

Hubble E., Humason M. L., 1931, ApJ, 74, 43

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Smoot G. F. et al., 1992, ApJ, 396, L1

Chapter 2

Measuring Ω_0 from the entropy evolution of clusters

Abstract. In this chapter, the entropy-driven model of cluster evolution developed by Bower is extended in order to predict the X-ray evolution of galaxy clusters in a range of cosmological scenarios. The model is then applied to recent measurements of evolution in the $L_x - T$ relation and in the X-ray luminosity function to place constraints on cosmological parameters. The data is insufficient to determine a particular cosmological model; an additional constraint is required to break the degeneracy between the effective slope of the power spectrum, n, and the density parameter, Ω_0. Two methods for achieving this are examined: a theoretical calculation of the relation between Ω_0 and n predicted by the CDM paradigm, and a direct measurement of n from galaxy clustering. The first gives $\Omega_0 < 0.55$ ($0.1 < \Omega_0 < 0.7$ for $\Omega_0 + \Lambda_0 = 1$), while the second gives $\Omega_0 < 0.6$ ($\Omega_0 < 0.65$ for $\Omega_0 + \Lambda_0 = 1$), both at the 95% confidence level. The rate of entropy evolution is insensitive to the values of Ω_0 considered, although it is sensitive to changes in the distribution of the intracluster medium.

2.1 Introduction

Clusters of galaxies are the largest virialised mass concentrations in the present-day Universe. Evolutionary studies therefore offer a unique method for directly determining the rate at which these structures grow in mass. This is influenced by the competing effects of the steepness of the density fluctuation power spectrum, characterised by the effective slope, n, and the current values of the cosmological parameters, Ω_0 (the density parameter) and Λ_0 (the cosmological constant, in units of $3H_0^2$, where $H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ is the Hubble constant). This chapter investigates the possibility that the value of Ω_0, or the combination, $\Omega_0 + \Lambda_0 = 1$ (which produces a flat geometry), can be robustly determined from considerations based solely on the X-ray evolution of galaxy clusters. Such an approach can only be successful if is able to overcome two substantial obstacles. Firstly, the correspondence between cluster mass and X-ray luminosity is not straightforward, being sensitive to the way in which the gas fills the cluster's gravitational potential well.

Secondly, the degeneracy between the cosmological parameters and the effective slope of the power spectrum on cluster scales, n, must be broken before a value of Ω_0 can be determined.

The first problem can be tackled by measuring the rate of luminosity evolution and calibrating the efficiency of X-ray emission by some other means of mass estimation, such as the luminosity–temperature relation (e.g. Mushotzky & Scharf 1997) or gravitational lensing effects (Smail et al. 1997; Bower & Smail 1997). The second problem can be approached in several ways. The most straightforward is to adopt a fluctuation spectrum on the grounds of a physical hypothesis, for example the CDM model (e.g. Bardeen et al. 1986). Alternatively, an empirical measurement of the power spectrum could be used, for example from the large-scale distribution of rich clusters obtained in redshift surveys (e.g. Tadros, Efstathiou, & Dalton 1998). Perhaps the most appealing approach would be to use X-ray observations spanning a wide range in redshift and luminosity to separate the models purely on the basis of their observed evolution.

Numerous approaches to the determination of Ω_0 have already been presented in the literature (e.g. see Dekel, Burstein, & White 1997, for a recent review), which range from direct arguments based on the peculiar motions of galaxies in the local Universe, through indirect methods using the baryon fraction in clusters of galaxies, to measurements of the acoustic peak in the microwave background spectrum. Each of these approaches has its own strengths, measuring different aspects of the overall cosmological model. A value based on the rate of cluster mass growth is appealing, however, since it measures Ω_0 on the basis of its large-scale effect over a modest factor of the Universe's expansion. Convergence of all these methods will act as confirmation that our global cosmological picture is valid and that no crucial additional physical processes have been omitted.

Discussion of the implications of cluster evolution is also timely given the growing area of sky that has now been exploited in X-ray surveys. These range in strategy from wide-area projects based on the *ROSAT* all-sky survey (Ebeling et al. 1997; De Grandi et al. 1997), EMSS (Henry et al. 1992; Gioia & Luppino 1994) to smaller solid-angle surveys based on serendipitous sources identified in deep *ROSAT* fields (e.g. Rosati et al. 1998; Scharf et al. 1997; Collins et al. 1997). Some constraints are also available from very deep survey fields (Hasinger et al. 1998; McHardy 1998; Bower 1996) although the area covered by these fields is currently very small. Reliable temperature data are also becoming available for clusters spanning a range of redshifts and luminosities, from the *ASCA* satellite (e.g. Tsuru et al. 1996; Mushotzky & Scharf 1997; Markevitch 1998). These recent advances have motivated this study, but there is also a clear need to identify the places where future observations using, for example,

the *Chandra* and *XMM* satellites are best targeted. One crucial question is whether more progress is to be made by going to lower flux levels, or more uniform surveys covering a wider area of sky.

This work is related to that of several others in the recent literature, representing a range of possible approaches to the problem of cluster evolution. This chapter uses a phenomenological model that separates the evolution of clusters into factors depending on the evolution of the mass spectrum and processes resulting in heating and cooling of the intracluster gas. The foundations of the model were discussed extensively in Bower 1997 (hereafter Paper I), providing a physical basis and interpretation. This allows additional theoretical input to be minimized, by using simple scaling relations to translate the properties of cluster samples at low redshift into their equivalents at high redshift. The approaches used by Mathiesen & Evrard (1998), Reichart et al. (1998) and Blanchard & Bartlett (1998) are related, but use an empirical model for the X–ray luminosity calibration and evolution, combined with the Press–Schechter method (Press & Schechter 1974) for the distribution of cluster masses. The approach of Kitayama & Suto (1997) is more different in the sense that they achieve a match to the luminosity function data through varying the epoch at which the clusters form. In the case of very low values of Ω_0, the distinction between the epoch at which a cluster is observed and that when it was formed becomes important. Therefore, the model from Paper I has been developed to incorporate this effect. Finally, a method of direct deconvolution has been outlined by Henry (1997, see also Eke et al. 1998) using recent *ASCA* temperature data. A wide variety of approaches to this topic is clearly desirable in order to indicate the robustness of the underlying principles. The discussion therefore includes a comparison between this work and others, and areas of uncertainty are outlined which can be considerably improved from further observations.

The layout of this chapter is as follows. In Section 2.2, the model on which this chapter is based is outlined and the manner in which it is readily extended to incorporate evolution in both open ($\Omega_0 < 1$, $\Lambda_0 = 0$) and flat ($\Omega_0 < 1$, $\Omega_0 + \Lambda_0 = 1$), sub-critical density universes is discussed. The approach used to incorporate the effects of the cluster formation epoch is also detailed. Section 2.3 summarises the constraints placed on the model parameters for a range of cosmologies using currently available X-ray cluster data. The degeneracy between Ω_0 and n is discussed and whether the methods mentioned earlier can be used to empirically distinguish between different cosmological models. The possibility of measuring evolution in the cluster core radius to place further constraints on the model is also investigated. In Section 2.4, the results of this study are summarised and the robustness of the assumptions made in the model are discussed. The results are then compared with the values that have been obtained by other authors, and differences between

the proposed models are explored, leading to the consideration of the overall accuracy of the method. With this in mind, key strategies for forthcoming X-ray surveys are identified. Finally, conclusions are reiterated in Section 2.5.

2.2 X-ray evolution in different cosmologies

2.2.1 X–ray emission and the cluster core

X-ray emission from the intracluster medium (ICM) is dominated by collisional processes. For hot clusters, the most important of these is thermal bremsstrahlung. The emissivity (power radiated per unit volume) scales as $\rho^2 T^\alpha$, where $\alpha = 0.4\,[\pm 0.1]$ is used for bolometric/wide-band detectors and $\alpha = 0\,[\pm 0.1]$ for low energy band-passes, although it is found that the results are not affected when making changes to these values within the limits quoted in square brackets. Hence, detections of galaxy clusters in the X-ray region of the electromagnetic spectrum are sensitive to the distribution of the ICM. Surface brightness profiles are traditionally fitted assuming the following (so-called β – model) density profile

$$\rho(r) = \rho_c \left[1 + \left(\frac{r}{r_c}\right)^2\right]^{-3\beta/2}, \tag{2.1}$$

(Cavaliere & Fusco-Femiano 1976), with r_c defining the effective core-size of the ICM and β the rate at which the density falls off with radius. A value of $\beta = 2/3$ is adopted for the main results in this chapter, appropriate for a non-singular isothermal distribution, in agreement with the observational average (e.g. Jones & Forman 1984). The effect of departures from this asymptotic slope is discussed in Section 2.4.1. The emission is thus characterised by a flattening at small scales (typically $r_c \sim 100\,\mathrm{kpc}$), departing from the distribution expected if the gas traces the underlying dark matter. A plausible physical interpretation was given by Evrard & Henry (1991, hereafter EH; see also Kaiser 1991), hypothesising that the intracluster gas was pre–heated prior to the cluster's formation and has retained the entropy acquired from this pre–collapse phase. Assuming that any dissipation was negligible, this provides an entropy *floor*, $\mathrm{s_{min}}$ [1], forcing the gas to build up a mass distribution that reflects this constraint. For an isothermal distribution, $\mathrm{s_{min}}$ is directly related to the core density, ρ_c. In Paper I, this idea was developed further by allowing s_{min} to evolve, using the parameterisation

$$s_{\mathrm{min}} = s_{\mathrm{min}}(z=0) + c_v\,\epsilon\,\ln(1+z), \tag{2.2}$$

[1] specific entropy is defined as $s = c_v \ln(T\rho^{1-\gamma})$, γ is the ratio of specific heats at constant volume (c_v) and constant pressure (c_p); its value is taken to be $5/3$, appropriate for a monatomic ideal gas

where ϵ determines the rate of core entropy evolution, which could be dominated by, for example, the gas being shock–heated as a result of merging (implying negative values) or conversely, radiative cooling (positive values). The value $\epsilon = 0$ corresponds to no net evolution of the core entropy.

2.2.2 *Evolution in an* $\Omega = 1$ *universe*

The X–ray luminosity of an individual cluster is calculated by integrating over the virialised region

$$L_{\rm x} \propto \int_0^{r_{\rm vir}} r^2 \rho^2 T^{\alpha} dr. \tag{2.3}$$

Equation 2.1 is used to describe the density profile, and it is assumed that the characteristic temperature of the ICM is proportional to the virial temperature of the system. In practice, the integral is dominated by the contribution from within a few core radii, and thus the scaling properties of this integral depend only weakly on the shape of density profile. As in EH, extracting the scaling properties from the integral produces

$$L_{\rm x} \propto \rho_c^2 r_c^3 T^{\alpha}. \tag{2.4}$$

Using Equation 2.2 to parameterise the evolution of the core entropy, the core density of the ICM evolves as

$$\rho_c \propto T^{3/2} (1+z)^{-3\epsilon/2}. \tag{2.5}$$

The core density can also be related to the virial radius, virial density and the core radius, using the asymptotic profile of the gas distribution, $\rho_c \sim \rho_{\rm vir}(r_{\rm vir}/r_c)^{3\beta}$. Assuming that the virial radius scales as $r_{\rm vir} \propto (M/\rho_{\rm vir})^{1/3}$, where M is the virial mass of the cluster, leads to a relation for the core radius

$$r_c \propto M^{1/3} T^{-1/(2\beta)} (1+z)^{-1+(1+\epsilon/2)/\beta}. \tag{2.6}$$

The above relations are not appropriate for scaling the physical properties of individual clusters. However, assuming the Weak Self-Similarity Principle (Paper I), these relations are accurate in determining the mean evolution of a population. This is fully justified if a scale-free power spectrum is adopted to describe initial fluctuations, $P(k) \propto D^2(z)\, k^n$ (Kaiser 1986; Paper I), where D is the linear growth factor and the range $-3 < n \leq 1$, appropriate for hierarchical models, is considered in this study. For an $\Omega = 1$ Universe, $D \propto (1+z)^{-1}$ (Peebles 1980). Since linear evolution contains no fixed physical scale, the only scale of interest at a given epoch is that

at which structure becomes non-linear (i.e. as the overdensity, $\delta \rightarrow 1$). This translates to a mass scale by considering the variance of linear mass fluctuations inside spheres of comoving radius, R

$$\sigma_R^2 \propto D^2(z) \int_0^\infty \tilde{W}_R^2(k)\, k^{n+3}\, d\ln k, \tag{2.7}$$

where $\tilde{W}_R(k)$ is the Fourier transform of the top-hat window function. A specific value of R (and thus σ_R) will correspond to the scale of objects that are now rich clusters (Evrard 1989; White, Efstathiou, & Frenk 1993). This leads to a simple scaling relation for the mass of such objects

$$M \propto (1+z)^{-6/(n+3)}, \tag{2.8}$$

given that $M = (4\pi/3)R^3\bar{\rho}$. Hence as $n \rightarrow -3$, structure formation becomes progressively more rapid.

The gas virial temperature will correspondingly evolve as

$$T \propto M^{2/3}\,(1+z). \tag{2.9}$$

This and the above results forms a relation for the X-ray luminosity

$$L_x \propto (1+z)^{\frac{3}{2}(\epsilon-1)(2-\frac{1}{\beta})}\, T^{\alpha+\frac{3}{2}(3-\frac{1}{\beta})}. \tag{2.10}$$

The two key observables that are used in this chapter are the X-ray luminosity function (XLF) and the luminosity-temperature ($L_{\rm x} - T$) relation. The first requires the number density evolution of clusters to be modelled, which is generally done using Monte-carlo realisations of merger trees within the extended Press-Schechter formalism (Lacey & Cole 1993), or directly using numerical simulations. However, since only a characteristic value is sought at any given epoch, it is assumed that the the number density scales inversely with M, i.e. $n(L) \propto (1+z)^{6/(n+3)}$, a necessary requirement to conserve mass for a self-similar mass distribution. For the $L_{\rm x} - T$ relation, only the evolution in normalisation can be constrained by the model discussed here. Attempting to constrain the slope using this approach requires the assumption of strong self-similarity, which is deliberately avoided in this work (see Paper I for a more detailed discussion).

2.2.3 Extension to $\Omega_0 < 1$ cosmologies

For a sub-critical density Universe, several major modifications are required to be made to the model. It can no longer be assumed that the dynamical evolution can be written in time-independent form. Hence, the simple relations for the evolution of the virial quantities will no

longer suffice. Even for scale-free fluctuation spectra, additional knowledge is required to relate the typical density of collapsed objects to the background density. For a low density universe, the Weak Self-similarity Principle cannot be rigorously justified since the internal mass structure of the clusters may not be homologous, depending on the epoch at which they collapse. However, these deviations are small in N-body simulations (e.g. Eke, Navarro, & Frenk 1998, for the $\Omega_0 = 0.3, \Lambda_0 = 0.7$ CDM scenario). The aim of this chapter is to *compare* the properties of clusters in flat and open universes, and thus the approach is fully adequate.

To relate the dynamical properties of cluster populations to the background cosmology, the results of the spherical top-hat collapse model (e.g. Lahav et al. 1991) are used. This predicts that overdense spherically symmetric perturbations depart from the linear regime, turn around and collapse, forming virialised structures with mean internal densities given by the formula $\rho_{\rm vir} = \Delta_{\rm vir}\, \rho_{\rm cr}$, where $\rho_{\rm cr}$ is the density required to close the Universe. The quantity $\Delta_{\rm vir}$ is a function of the background cosmology and can be approximated using the following fitting functions (Eke, Navarro, & Frenk 1998)

$$\begin{aligned} \Delta_{\rm vir} &= 178\, \Omega^{0.30} \, (\Lambda_0 = 0) \\ \Delta_{\rm vir} &= 178\, \Omega^{0.45} \, (\Omega_0 + \Lambda_0 = 1) \end{aligned} \tag{2.11}$$

which returns the value of 178 for $\Omega = 1$. Knowledge of Ω_0, Λ_0 and the redshift of collapse, z_f is sufficient to provide an evaluation.

Another difference with $\Omega_0 < 1$ models is that the linear growth factor, $D(z)$, evolves more slowly than $(1+z)^{-1}$. The appropriate relations are as follows (Peebles 1980)

$\boldsymbol{\Omega_0 < 1}$:

$$\begin{aligned} D(x) &= 1 + \frac{3}{x} + \frac{3(1+x)^{\frac{1}{2}}}{x^{\frac{3}{2}}} \ln[(1+x)^{\frac{1}{2}} - x^{\frac{1}{2}}], \\ x &= \frac{\Omega_0^{-1} - 1}{1+z}. \end{aligned}$$

$\boldsymbol{\Omega_0 + \Lambda_0 = 1}$:

$$\begin{aligned} D(y) &= \frac{(y^3+2)^{\frac{1}{2}}}{y^{\frac{3}{2}}} \int_0^y \frac{u}{u^3+2}^{\frac{3}{2}} \, du, \\ y &= \frac{2(\Omega_0^{-1} - 1)^{\frac{1}{3}}}{1+z}. \end{aligned} \tag{2.12}$$

Accounting for these differences produces the following modified scaling relation for the X-ray luminosity

$$L_{\rm x} \propto T^a \, (1+z)^b \, (1+z_f)^c \, [\Delta_{\rm vir}(z_f)/\Omega(z_f)]^d; \tag{2.13}$$

$$a = \alpha + 3(3 - 1/\beta)/2;$$

$$b = 3\epsilon(1/(2\beta) - 1);$$

$$c = 3(1/\beta - 3/2);$$

$$d = 1/\beta - 3/2.$$

Hence for $\beta = 2/3$, $c = d = 0$. The temperature scaling relation becomes

$$T \propto D(z)^{4/(n+3)} (1 + z_f) [\Delta_{\rm vir}(z_f)/\Omega(z_f)]^{1/3} \tag{2.14}$$

and the number density scales as $n(L) \propto D(z)^{-6/(n+3)}$.

2.2.4 The epoch of cluster formation

The above modifications are required to be made for $\Omega_0 < 1$ models, since the linear growth of structure 'freezes out' when $z \sim \Omega_0^{-1} - 1$ (Peebles 1980). Producing the present-day abundance of clusters requires structure to have formed at progressively earlier epochs for lower density Universes. Consequently, the epoch of cluster formation, z_f can be significantly different from the redshift range of an observed sample, leading to inaccuracies in the scaling properties (Kitayama & Suto 1996). This problem is addressed by defining the epoch of formation to be such that the cluster has acquired a fraction, f of its total mass when observed. The characteristic mass can then be scaled as follows

$$M(z_f(z)) = f\, M(z). \tag{2.15}$$

Material that has been accumulated between z_f and z is assumed to have negligible effects on the normalisation of the mass distribution, even though the total mass (and therefore temperature) of clusters can change.

The effect that varying f has on the X-ray luminosity is illustrated in Figure 2.1 for a typical median redshift of current high redshift cluster samples ($z = 0.4$). The following function is plotted

$$F(\Omega_0, \Lambda_0, z) = F_0\, (1 + z_f)^{11/4} [\Delta_{\rm vir}(z_f)/\Omega(z_f)]^{11/12}, \tag{2.16}$$

as a function of Ω_0, which is the part of Equation 2.13 (substituting Equation 2.14 for T and assuming $\beta = 2/3$) that depends on z_f. The constant, F_0 normalises the relation to the corresponding value at the present day ($z_f(z = 0)$). Each relation is scaled using the $\Omega = 1$ value to illustrate the relative effect. For the $\Omega = 1$ case, the formation epoch scales in direct proportion to the observed

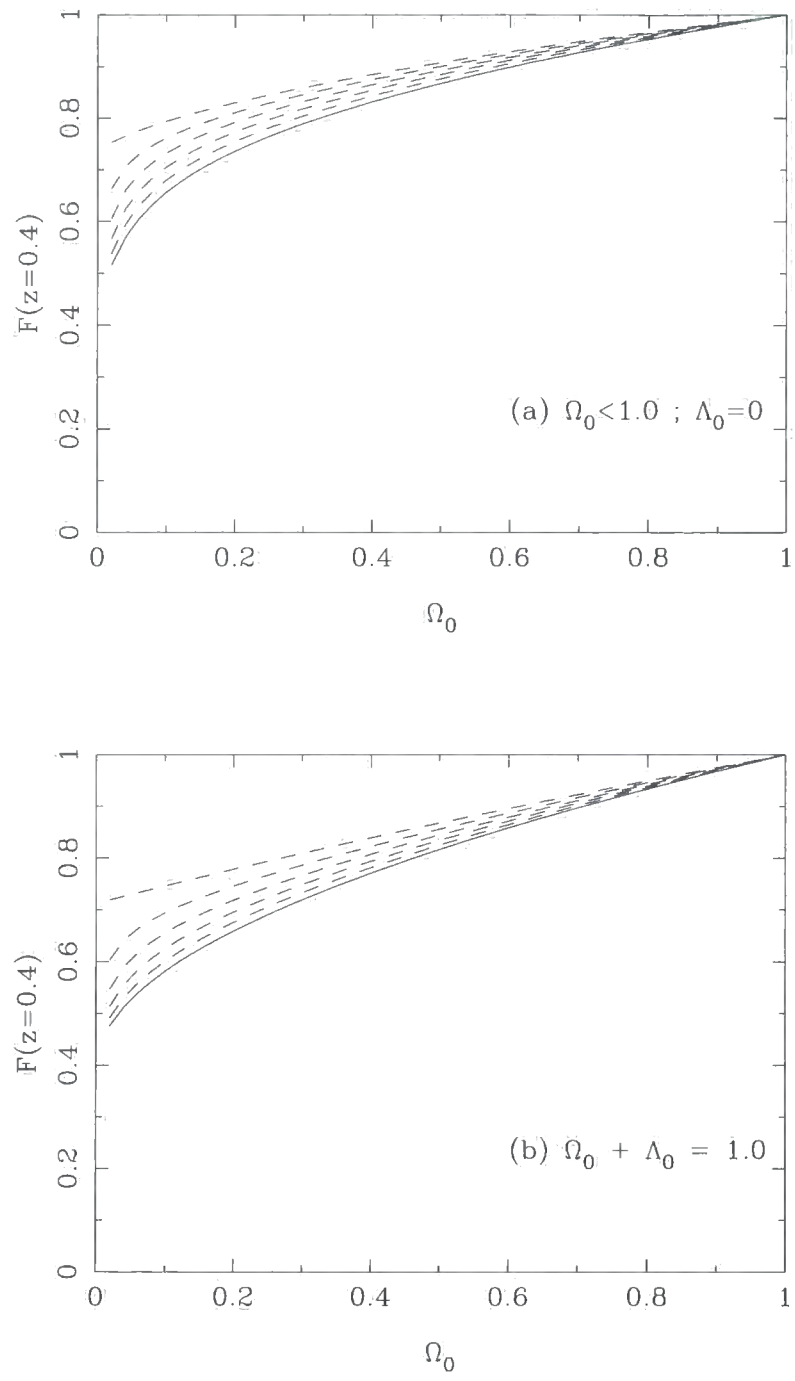


Figure 2.1: The X-ray luminosity factor that depends on the cluster formation redshift, z_f, as a function of the background cosmology, normalised to the $\Omega = 1$ value. Results are illustrated for both open ($\Omega_0 < 1$) and flat ($\Omega_0 + \Lambda_0 = 1$) models, for the case $z = 0.4$. The solid line illustrates the variation of F for a mass fraction of 1 and the dashed lines show progressively lower values, in steps of 0.1, to $f = 0.5$.

Table 2.1. Values of the parameters used to fit the (temperature) evolution of the $L_x - T$ relation, assuming a power law dependence on T with slope, λ and an evolution term between low and high redshift parameterised as $(1+z)^{\eta}$. Shown are the best-fit values and 95 per cent confidence levels of the slope and evolution parameters, for the median redshift of the high-z sample, $\langle z \rangle = 0.3$, and a value of $\Omega_0 = 0$.

Parameter	Best Fit	95 per cent (min,max)
η	-0.10	(-0.63,0.43)
λ	3.45	(3.03,3.85)

redshift. As the density drops below the critical value, there is a decrease in about a factor of 2 for $\Omega_0 = 0.1$, although the dependence on f is relatively weak (even for $\Omega = 0.1$) : reducing f from $1 \rightarrow 0.5$ produces only a 40 per cent change in F, even in this extreme case. Therefore, a fiducial value of $f = 0.5$ has been selected for the results that follow.

2.3 Using the evolution of clusters to constrain Ω_0

The essence of this approach is to scale characteristic properties of a cluster population from one epoch to another. Therefore, data based on local samples is used as the basis for the scaling transformations described in the last section. The likelihood distribution of parameters can then be determined, by fitting the scaled relations to the data available at higher redshift. At this stage, it has been assumed that the free parameters are the slope of the power spectrum, n, and the entropy evolution parameter, ϵ. The cosmological density parameter has been fixed at four values (with and without the cosmological constant) : $\Omega_0 = 0.1, 0.3, 0.5, 1.0$, in order to clearly illustrate the effects of a varying cosmological background on the physical evolution of clusters. Below, the data–sets used to constrain the set (n, ϵ) are discussed as well as the corresponding results. All data presented have been compiled with an assumed Hubble constant of $H_0 = 50\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2.3.1 The Luminosity–Temperature relation

Constraints are placed on the temperature evolution of clusters by making use of the $L_x - T$ relation, which is at present, adequately described by a power law, $L_x \propto T^{\lambda}$, and an intrinsic scatter. The evolution of the $L_x - T$ relation is determined by fitting a maximum likelihood model to the combined low and high redshift data given by David et al. (1993) and Mushotzky & Scharf (1997), compiled mainly from the *ASCA* and *Einstein* satellites. The distribution of X-ray temperatures is assumed to be Gaussian, and the 90% range in systematic errors in the temperature measurements

(quoted by the authors) are converted to an equivalent 1σ value. In order to ensure that there is good overlap in luminosity between the high and low redshift data-sets, the comparison is limited to clusters with bolometric luminosities greater than $10^{44.5}\,\mathrm{erg\,s^{-1}}$, although the results are not particularly sensitive to this choice. The model for the $L_\mathrm{x} - T$ correlation that was fitted includes an adjustable zero-point, slope and intrinsic scatter, as well as a redshift-dependent normalisation term that is parameterised as $\Delta \log T_0 = \eta \log(1+z)$ (where T_0 is a reference temperature). Confidence limits are determined by minimizing over the other parameters and calculating the distribution of C values, where

$$C_\mathrm{LT} \equiv -2 \sum_{i=1}^{N} \ln P_i(\eta), \tag{2.17}$$

where the subscript i ranges over the cluster sample, P_i is the probability of measuring the cluster with a given temperature, luminosity and value of η. It is assumed that ΔC is distributed as χ^2 with one free parameter (Cash 1979).

This treatment ignores any change in the slope of the relation, which is true if the luminosities of the clusters do not evolve and the evolution of the $L_\mathrm{x} - T$ relation is due to the temperature evolution of the clusters alone. In practice, however, a particular choice of the parameters (n, ϵ) implies correlated changes in luminosity and temperature. For simplicity and consistency with other work, the evolution in both quantities has been phrased as a change in temperature at fixed luminosity. This correction involves the slope of the relation. However, it must be emphasised that this is an artifact of the way in which the normalisation is quoted. In principle, each model requires that the likelihood be examined as a function of a combination of η and the fitted slope, λ (which changes with the choice of (n, ϵ) parameters); in practice, however, the slope of the relation is sufficiently well constrained that almost identical results are obtained if the likelihood is examined as a function of the single parameter, η and finding the slope of the best-fit value.

The preferred values of η and λ are given in Table 2.1, for the median redshift of the high–redshift sample ($\langle z \rangle = 0.3$) and $\Omega_0 = \Lambda_0 = 0$. These values are converted to the appropriate cosmologies when required, since the normalisation is affected by the assumed value of Ω_0 through the distance dependence of the luminosity. In $\Omega_0 < 1$ cosmologies, the luminosities of clusters are brighter, pushing the evolution of the temperature normalisation in the positive direction for higher values of Ω_0. The uncertainties in the evolution rate are dominated by the intrinsic scatter in the relation. It is therefore noted that while the evolution of the relation is statistically well defined, it is sensitive to systematic error and selection (based on X-ray flux) which may have tended to exclude hotter (or colder) clusters.

Table 2.2. Details of the distant cluster survey XLF's used for constraining evolution of the *ROSAT* BCS XLF. The EMSS 2 data are not binned (*) but rather taken as a point on the cumulative luminosity function.

Survey	N_{bins}	$\langle z \rangle$	N_{clus}
SHARC (Burke et al. 1997)	4	0.44	16
WARPS (Jones et al. 1998)	3	0.47	11
RDCS (Rosati et al. 1998)	3	0.60	14
EMSS 1(Henry et al. 1992)	4	0.33	23
EMSS 2(Luppino & Gioia 1995)	*	0.66	6

2.3.2 The X–ray Luminosity Function

The second X–ray observable used in this chapter is the cluster X–ray Luminosity function. For simplicity, the XLF is assumed to take the form of a Schechter function

$$\frac{dn}{dL_{\rm x}} = A \exp\left(-\frac{L_{\rm x}}{L_{\rm x}^*}\right) L_{\rm x}^{-\alpha}. \tag{2.18}$$

The local XLF used in this chapter is the one given by Ebeling et al. (1997) based on the *ROSAT* Brightest Cluster Sample (BCS), which contains clusters with $z < 0.3$, selected and flux–limited at X–ray wavelengths. The authors' best fit parameters to Equation 2.18 have been adopted, namely $A = 3.32 \times 10^{-7}\,\mathrm{Mpc}^{-3}\Delta L^{-1}; \alpha = 1.85; L_{\rm x}^* = 5.7 \times 10^{44}\,\mathrm{erg\,s}^{-1}$ for the $0.5 - 2.0$ keV band, and $A = 4.95 \times 10^{-7}\,\mathrm{Mpc}^{-3}\Delta L^{-1}; \alpha = 1.82; L_{\rm x}^* = 1.07 \times 10^{45}\,\mathrm{erg\,s}^{-1}$ for the $0.3 - 3.5$ keV band. The functions are plotted as solid lines in Figure 2.2. Since they find no significant evolution ($< 1.8\sigma$) in their sample, this function is assumed to represent the XLF at the present day.

Currently available high redshift cluster samples, used to place constraints on the evolution of the XLF, are summarised in Table 2.2. This data are also illustrated in Figure 2.2, evaluated for $\Omega = 1$. Figure 2.2(a) illustrates the binned, non–parametric XLF's that have been compiled from *ROSAT* data (SHARC , WARPS and RDCS), with luminosities evaluated in the $0.5 - 2.0$ keV band. None of these samples shows significant evolution in the cluster XLF, out to $z \sim 0.5$. Figure 2.2(b) shows the EMSS 1 XLF in the $0.3 - 3.5$ keV band. Clearly, this hints at negative evolution (i.e. a lower space density of clusters of given luminosity at higher redshift), although it must be noted that higher luminosity bins usually have greater median redshift values and hence applying one redshift to the whole sample leads to an overestimation of the evolution. Finally, Figure 2.2(c) shows the EMSS 2 cumulative point, also in the $0.3-3.5$ keV band. Since this point is significantly lower than the BCS XLF, it should provide a tight constraint on n and ϵ.

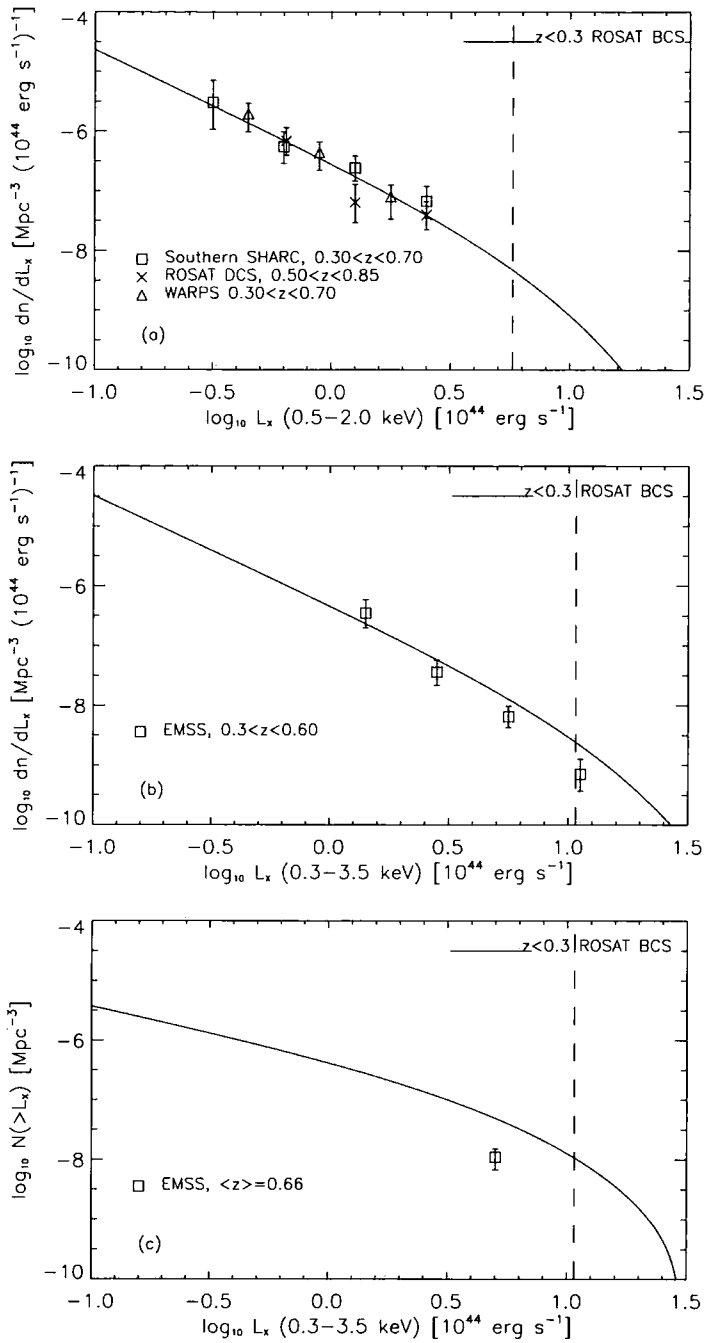


Figure 2.2: The X–ray luminosity functions used in this chapter compiled from data observed using the *ROSAT* and *Einstein* satellites. The solid line in all three plots is the best–fit Schechter function to the *ROSAT* BCS $z < 0.3$ sample as determined by Ebeling et al. (1997). Plot (a) illustrates the high–redshift non–parametric XLF data from *ROSAT* surveys, evaluated in the $0.5 - 2.0$ keV pass band. The square points are from the Southern SHARC survey (Burke et al. 1997), the crosses from the RDCS survey (Rosati et al. 1998, private communication) and the triangles from the WARPS survey (Jones et al. 1998, private communication). Plot (b) illustrates the EMSS cluster X–ray luminosity function from the work of Henry et al. (1992) and plot (c) shows the cumulative XLF point from the EMSS distant cluster sample ($N(> 5 \times 10^{44}\,\mathrm{erg\,s^{-1}}) = 1.1 \pm 0.43 \times 10^{-8}\,\mathrm{Mpc^{-3}}$; Luppino & Gioia 1995), both evaluated in the $0.3 - 3.5$ keV band. The vertical dashed line marks the position of L_{x}^{*}. Data is presented assuming $\Omega = 1$.

A likelihood method is used to compare the data–points with each model prediction, defined by (n, ϵ), for the assumed cosmology. Ideally, the model parameters would be constrained using a global maximum likelihood approach. Starting from the local XLF, individual likelihood probabilities could be assigned to each of the observed clusters (and non–detections) by combining the model X–ray luminosity function at the appropriate redshift with the selection function defined by each of the surveys. This approach would avoid all problems related to redshift binning of the available data, and allow cosmological corrections to be consistently applied. Unfortunately, the detailed survey selection functions are not generally available, so an approximate approach must be adopted. Since little evolution is observed, binning the data in redshift is not likely to result in significant bias.

The constraints placed on the model parameters (n, ϵ) from evolution of the various high–redshift XLF's (relative to the BCS XLF) for a given cosmology were generated as follows. Firstly, the high–redshift XLF data points were rescaled to the values appropriate for the assumed cosmology. Luminosity scales as d_L^2, where d_L is the luminosity distance (Section 1.2), therefore clusters are more luminous for $\Omega_0 < 1$ cosmologies than for $Omega = 1$. The inferred space density of clusters decreases as Ω_0 is lowered since the volume out to any given redshift is larger. The space density is assumed to scale as $1/V_{\rm max}$, where

$$V_{\rm max} = \int_0^{z_{\rm lim}} \left| \frac{dV}{dz} \right| dz, \tag{2.19}$$

and $z_{\rm lim}$ is the redshift at which the luminosity infers a flux equal to the flux limit of the survey. For the results presented here, the space densities are always scaled using the assumption that all clusters are at the flux limit of the survey. However, the relative change in $V_{\rm max}$ for clusters being a factor of two brighter than the flux limit is ~ 1 per cent. It was then assumed that the number of clusters in the i^{th} bin ($N_{clus,i}$) were drawn from a Poisson distribution, with the expected number being a function of the *true* values (n^T, ϵ^T). The C–statistic was then calculated for the range of model (n, ϵ) such that $-2.5 \leq n \leq 1.0$ and $-6 \leq \epsilon \leq 4$, where C is defined as

$$C_{\rm XLF} = -2 \sum_{i=1}^{N} \ln P_i \left(N_{clus,i}; E_i(n, \epsilon) \right), \tag{2.20}$$

N is the number of bins in the high–redshift XLF and P_i is the probability of observing $N_{clus,i}$ clusters given E_i, the expected number calculated from the model, at the median redshift of the data point. Confidence regions were then calculated by differencing C with respect to the minimum value (i.e. corresponding to the most probable set, (n_0, ϵ_0)), $\Delta C = C - C_{\rm min}(n_0, \epsilon_0)$. If (n, ϵ) were independent parameters, ΔC would be distributed as χ^2 with two degrees of freedom

(Cash 1979), but this is not assumed to be the case. Instead, a large number of Monte-carlo realisations of each XLF data–set were generated, for all the cosmologies studied. Each realisation involved generating a set of XLF points at the same luminosities as the real data–set, drawn from Poisson distributions with mean values set to the numbers of clusters predicted by the best–fit model. The C distribution was then taken for each realisation and the difference in C values was calculated, between this distribution and the one produced by the best–fit model, at the fixed point (n_0, ϵ_0). This method allows a likelihood distribution of ΔC values to be constructed that resembles the true distribution. Confidence levels were chosen such that they enclosed 68 per cent and 95 per cent of the total number.

2.3.3 Results

For each of the fiducial values of Ω_0 (in both open and flat Universes), the likelihood distribution was determined as a function of n and ϵ. Initially, the different data–sets were analysed independently. However the results show similar trends, albeit less constrained than the overall likelihood distribution, generated from combining all of the XLF data. Therefore the discussion is focussed on results from the combined XLF data–set which are shown in Figure 2.3. (Results from individual data–sets can be found in Appendix A.) Clearly, this illustrates how the constrained area of parameter space varies with the background cosmology. This is the result of an interplay between two contributing factors, namely the model scaling relations used to fit the high redshift data and the positions of the data–points relative to the local XLF and $L_{\rm x} - T$ relation.

Constraints from the evolution of the $L_{\rm x} - T$ normalisation show a weak degeneracy between n and ϵ for high values of Ω_0. As $\Omega_0 \rightarrow 0$, the range of ϵ becomes independent of n (centred on $\epsilon = 0$). The range of n is unconstrained for all values of Ω_0. The reason for the degeneracy is straightforward. The data imply no net evolution of the $L_{\rm x} - T$ normalisation for an empty Universe ($\Omega_0 = 0$). Increasing the value of Ω_0 pushes the high redshift sample to lower luminosities, prompting positive evolution in temperature for a fixed value of $L_{\rm x}$. In an $\Omega = 1$ Universe with no entropy evolution ($\epsilon = 0$) and a power spectrum with a negative slope, the model over–predicts the cluster luminosities, hence positive entropy evolution is required to match the data. As n increases, the over–prediction becomes less severe (hence ϵ decreases). For $\Omega_0 < 1$ cosmologies, the amount of evolution required becomes progressively weaker as Ω_0 decreases. Since the growth of structure is also slower, the data become more and more consistent with $\epsilon = 0$ for all values of n.

The results from constraining the evolution of the combined XLF data–sets show a degeneracy in the opposite sense: higher values of n demand higher values of ϵ. This relation can be clearly

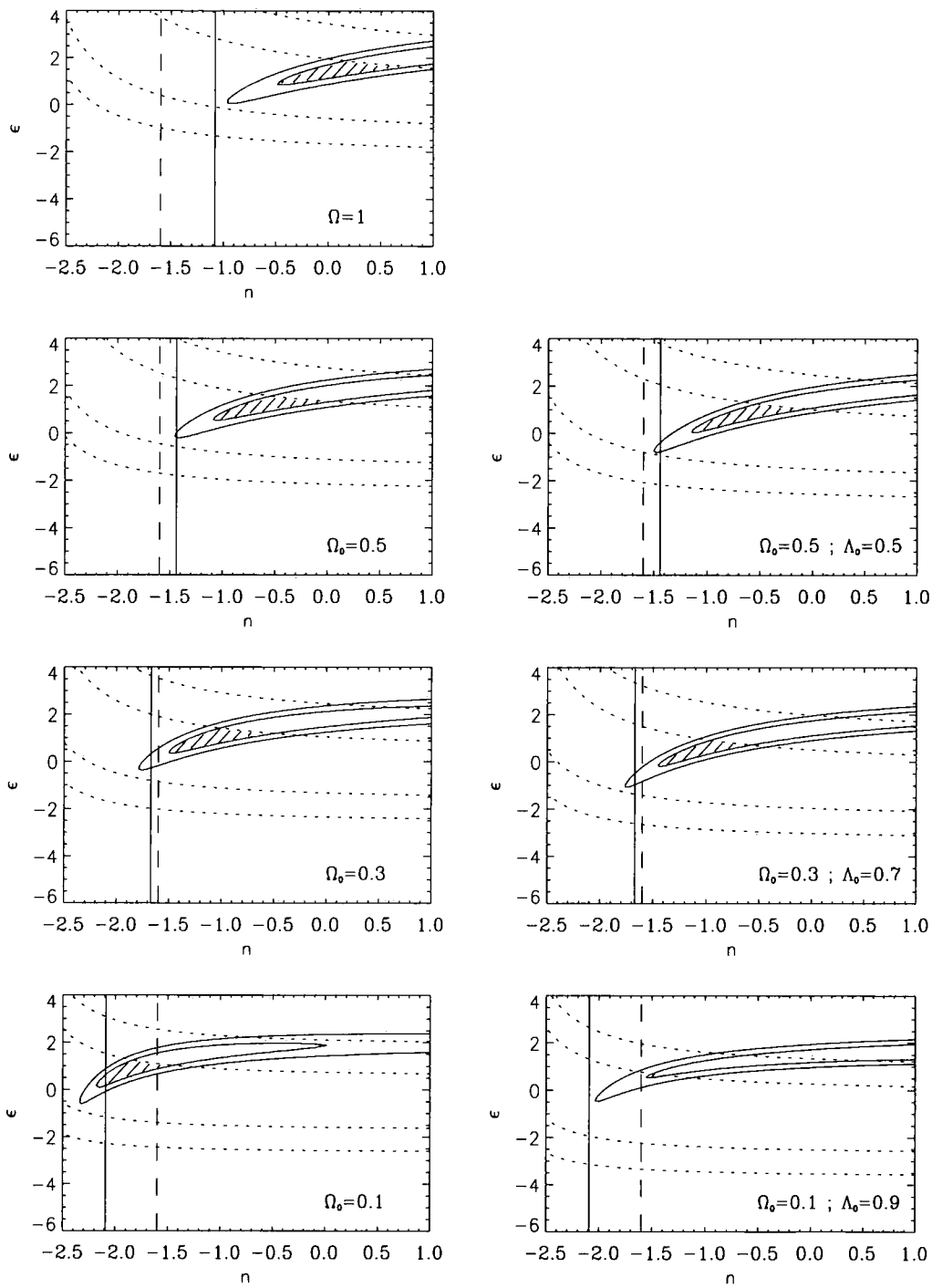


Figure 2.3: Constraints on the slope of the power spectrum (n) and the entropy evolution parameter (ϵ) for seven different cosmological models. Solid contours are the 68 per cent and 95 per cent confidence regions from the evolution of the X–ray luminosity function, using data from all of the surveys studied. Dotted contours represent the 68 per cent and 95 per cent confidence regions for evolution of the cluster luminosity–temperature relation. Shaded regions are for parameters consistent with the evolution within 68 per cent limits (of both XLF and $L_x - T$). The vertical solid line is the value of n appropriate to the CDM model calculation in Section 2.3.4 and the vertical dashed line is a measured value of n from the APM rich cluster sample.

explained from the $\Omega = 1$ plot alone. Again, the model scales the XLF in such a way that for $\epsilon = 0$, the result is an over–prediction of the cluster luminosities. However, this effect becomes less pronounced as n *decreases*, leading to lower values of ϵ. Eventually, the local luminosity function is shifted so much that the shape of the function becomes important, leading to an under–prediction of the abundance of clusters at the faint–end, where most of the data–points lie. This effect places a lower limit on the value of n. For lower values of Ω_0, the data–points are pushed to brighter luminosities, as well as lower values of dn/dL_x. Although the faint–end data are consistent with no evolution for $\Omega = 1$, the re–scaling forces the points below the local XLF for lower values of Ω_0. The brighter data (in particular the EMSS 2 point) is less sensitive to the change in cosmology, due to the steeper slope of the XLF. Coupled with the fact that the evolution of structure is weaker, an adequate fit to the data is allowed for more negative n.

In comparison to the open models, the flat ($\Omega_0 + \Lambda_0 = 1$) models show slight variations in the constrained region of parameter space. The reason for the difference is two–fold. Firstly, the data–points themselves are scaled differently between cosmologies than their open counterparts. This is because the distance to a fixed redshift is larger for flat models (hence also increasing the volume element). This has an effect akin to having an open model with a lower value of Ω_0. Secondly, the growth of structure in a flat model is more rapid than the corresponding open case, which affects the model scaling relations not unlike an open model with a higher effective Ω_0.

Combining the constraints from the $L_x - T$ and XLF measurements does not favour any particular value of Ω_0 (with or without Λ_0): there is always a region of parameter space consistent with both observations. The same conclusion applies to all of the individual XLF data–sets (see Appendix A). Specifically, the faint–end data–points from the *ROSAT* surveys are unable to set tight constraints; however they are able to set a lower limit on the slope of the power spectrum that is sensitive to the value of Ω_0. The EMSS samples, particularly the EMSS 2 cumulative point, start to provide tighter limits on n and ϵ, since these data start to probe the more sensitive luminosity range: brighter than L_x^*. Therefore, future surveys will be more efficient at extracting information on the evolution of clusters by covering larger areas of sky, probing luminosities around and beyond L_x^*.

2.3.4 Breaking the n–Ω_0 degeneracy

It is evident that the model is unable to constrain Ω_0 from the XLF and $L_x - T$ normalisation constraints alone – a plausible range of n and ϵ can always be found for any value of Ω_0. However, since the allowed range of n is sensitive to Ω_0, combining an independent determination of n with

the X-ray data will place a significant constraint on the value of Ω_0. There are two ways in which this can be achieved: using a theoretical prediction of the primordial power spectrum that gives the (cosmology–dependent) slope on cluster scales or using a direct measurement of n from redshift surveys.

For the theoretical case, the slope of the power spectrum on cluster scales is required, $n_{\rm eff}$, such that

$$n_{\rm eff} = \left.\frac{d\ln P}{d\ln k}\right|_{k_{\rm eff}}. \tag{2.21}$$

The CDM model is used, which predicts an initial power spectrum with $n \sim 1$ on large scales, turning over to $n \sim -3$ on small scales. This work uses an approximation to the power spectrum, using the transfer function given by Bardeen et al. (1986). The turnover scale is determined by the shape parameter, Γ, which is in turn, determined by the cosmological model

$$\Gamma = \Omega_0 {\rm h} \exp\left[-\frac{\Omega_b}{\Omega_0}\left(\Omega_0 + \sqrt{2{\rm h}}\right)\right] \tag{2.22}$$

(Sugiyama 1995). The value $\Omega_b = 0.013\,{\rm h}^{-2}$ is adopted from nucleosynthesis constraints (Copi, Schramm, & Turner 1995), although Ω_b only makes a small contribution to the value of Γ. The shape of the power spectrum is independent of the cosmological constant.

To estimate $k_{\rm eff}$, it is assumed that the virial mass of a typical rich cluster in the local Universe is $M \sim 5 \times 10^{14}\, h^{-1}\,{\rm M}_\odot$. If the cluster formed by the top–hat collapse of a spherically overdense region, the comoving radius required to contain this mass (assuming present background density) would be $R \sim 8\, h^{-1}\,{\rm Mpc}$ (Eke, Cole, & Frenk 1996). This is only true for a critical density Universe but can easily be modified for lower densities, since $R \propto \Omega_0^{-\frac{1}{3}}$. The value of R can then be converted to an effective scale on the linear power spectrum by using the following formula (Peacock & Dodds 1994)

$$k_R = \left[\frac{\left[\frac{1}{2}(n+1)\right]!}{2}\right]^{\frac{1}{n+3}} \frac{\sqrt{5}}{R}, \tag{2.23}$$

which should iteratively converge on the values $(n_{\rm eff}, k_{\rm eff})$, using Equation 2.21. The calculated values of n are represented as solid lines in Figure 2.3. Note that the value of n becomes more negative with decreasing values of Ω_0. Although the effective scale of galaxy clusters is larger in a low mass–density Universe, the change in shape of the CDM power spectrum itself (parametrised by Γ) dominates the shift in the value of n. As Ω_0 decreases, the position of the turnover moves to larger scales, leading to a more negative slope on cluster scales. Figure 2.3 shows the result of this calculation, plotted as a solid vertical line. Analysing the likelihood distributions for a

Table 2.3. Upper limits on the value of Ω_0 (with and without Λ_0), using the constraint that $n = -1.6$, from the APM survey (Tadros, Efstathiou & Dalton, 1997). Results from both individual and combined XLF data–sets are presented. Limits are set using the 95 per cent confidence limits consistent with both the XLF and $L_x - T$ constraints. Ranges in square brackets are for quoted uncertainties in n (± 0.3)

Survey	Ω_0(upper; $\Lambda_0 = 0$)		Ω_0(upper; $\Lambda_0 = 1 - \Omega_0$)	
SHARC	1.0	$[0.6, > 1.0]$	1.0	$[0.5, > 1.0]$
WARPS	1.0	$[0.75, > 1.0]$	1.0	$[0.7, > 1.0]$
RDCS	1.0	$[0.65, > 1.0]$	1.0	$[0.6, > 1.0]$
EMSS 1	1.0	$[0.7, > 1.0]$	1.0	$[0.6, > 1.0]$
EMSS 2	0.7	$[0.4, > 1.0]$	0.65	$[0.3, > 1.0]$
COMBINED DATA	0.4	$[0.25, 0.6]$	0.4	$[0.2, 0.65]$

continuous range of Ω_0, it is found that the CDM prediction is only consistent with the X–ray data for $\Omega_0 < 0.55$ ($\Lambda_0 = 0$) and $0.1 < \Omega_0 < 0.7$ ($\Omega_0 + \Lambda_0 = 1$), at 95 per cent confidence. A lower limit can be placed on Ω_0 in the flat model since the predicted value moves to more negative n faster than the XLF contours, as Ω_0 decreases.

Another approach is to constrain n directly by measuring the power spectrum on rich cluster scales from wide–area sky surveys. One example (Tadros, Efstathiou, & Dalton 1998) comes from the APM cluster sample, which has a median redshift $\langle z \rangle = 0.09$, giving a slope, $n_{\rm APM} = -1.6 \pm 0.3$. It is assumed that the underlying power spectrum has negligible curvature on the range of scales appropriate to this work. Ideally, cluster populations should be separated into redshift bins, and each population should be scaled relative to the preceeding one, using the slope of the spectrum at that epoch. Imminent surveys such as the Sloan Digital Sky Survey and the 2dF Galaxy Redshift Survey should provide the necessary data to constrain the slope at higher redshifts. Furthermore, the APM result assumes that the bias (i.e. the ratio of luminous to total mass clustering amplitudes; see Davis et al. 1985) is approximately constant on scales of the mean inter–cluster separation.

The range of Ω_0 was constrained by taking the 95% limits placed on n and ϵ from the evolution of the XLF and $L_x - T$ normalisation and subsequently fixing the value of n to -1.6. For the XLF data, it was found that the model fits the data more and more adequately as Ω_0 decreases (for the range of n considered): this is shown in the plot by the shift in the contours. Hence, this prevents a lower limit from being placed on Ω_0. In the flat models the geometry weakly alters the shape of the likelihood distribution but the overall trend is the same. Since both constraints

produce contours that become more aligned in lower Ω_0 cosmologies, further conclusions cannot be drawn regarding the lower limit.

What the data do infer is an upper limit for the value of Ω_0; the results are given in Table 2.3, which displays the values that are consistent with the 95 per cent confidence limits from the XLF and $L_x - T$ relations. Results for both the individual and combined XLF data–sets are presented. The ranges in brackets were determined by varying the value of n by ± 0.3. It is evident that the value of Ω_0 is unconstrained using the APM value, for all of the individual data-sets except the EMSS 2 survey. The flat models give similar results to their open counterparts, with the range in the upper limits being slightly larger. Combining the XLF surveys sets a tighter constraint: $\Omega_0 < 0.6$ (or equivalently $\Omega_0 < 0.65$ for $\Omega_0 + \Lambda_0 = 1$).

2.3.5 Evolution in cluster core sizes

Finally, another constraint that could be readily obtained from X–ray data is the evolution of cluster core–radii. Using the model developed in Section 2.2, the characteristic core-radius of a cluster population scales as

$$r_c \propto D^a (1+z)^b (1+z_f)^c (\Delta_{vir}/\Omega)^d; \tag{2.24}$$

$$a = (2/(n+3))(1-1/\beta),;$$

$$b = \epsilon/(2\beta),;$$

$$c = 1/(2\beta) - 1,;$$

$$d = (1/3)(1/(2\beta) - 1).$$

Even though the core radius depends on both n and ϵ, its evolution is much more sensitive to changes in the central entropy of the gas rather than to the extent of structural evolution. Figure 2.4 illustrates this, showing regions of parameter space consistent with various amounts of core radius evolution, out to a redshift $z = 0.5$. The background cosmology and spectral index have only a weak influence on the evolution of r_c, hence observations should quite easily place a constraint on the value of ϵ – the two extremes of the shaded regions represent scaling in the core by a factor of ~ 0.3 and 3 respectively. Core sizes that shrink with redshift demand negative values of ϵ (i.e. the core entropy of clusters decreases with redshift) whereas core sizes that grow with redshift demand values of ϵ that are weakly negative or higher. This approach is useful in the sense that the rate of entropy evolution in the core can be directly measured and can be tested for consistency with the other methods presented above.

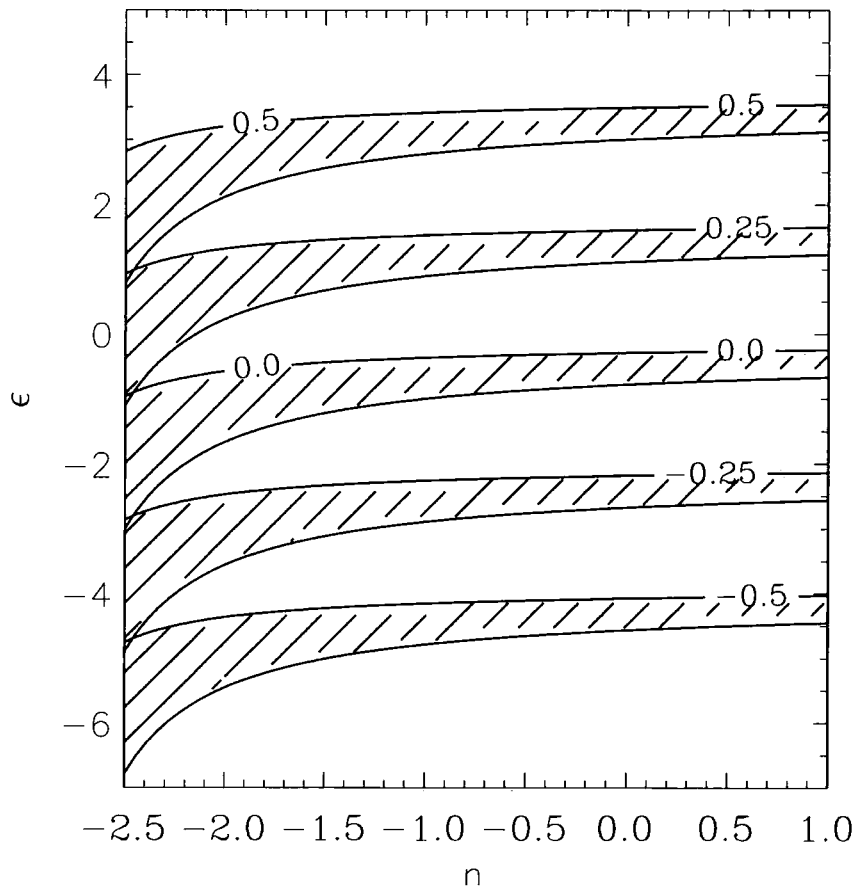


Figure 2.4: The region of parameter space covered by evolution of the cluster core radius for various models. The shaded regions depict the range of cosmologies, from $\Omega = 1$ (for which structure evolves most rapidly) to $\Omega = 0.1, \Lambda_0 = 0.0$ (for which structure evolves least rapidly, for the models considered in this chapter). Each shaded area corresponds to evolution between $z = 0$ and $z = 0.5$, with the labels indicating (positive and negative) values of $\Delta \log r_c$.

Analyses of evolution in the cluster core size have now started to appear in the literature. Results are currently suggesting that there is no significant evolution in the core size out to $z \sim 0.5$ (Vikhlinin et al. 1998). Given the uncertainties, the results of this chapter are consistent with this, but positive evolution in the cluster core radii ($\epsilon > 0$) is preferred. At present, the sample sizes are small; a more comprehensive measurement of core radius evolution will serve as an important test of the validity of this model.

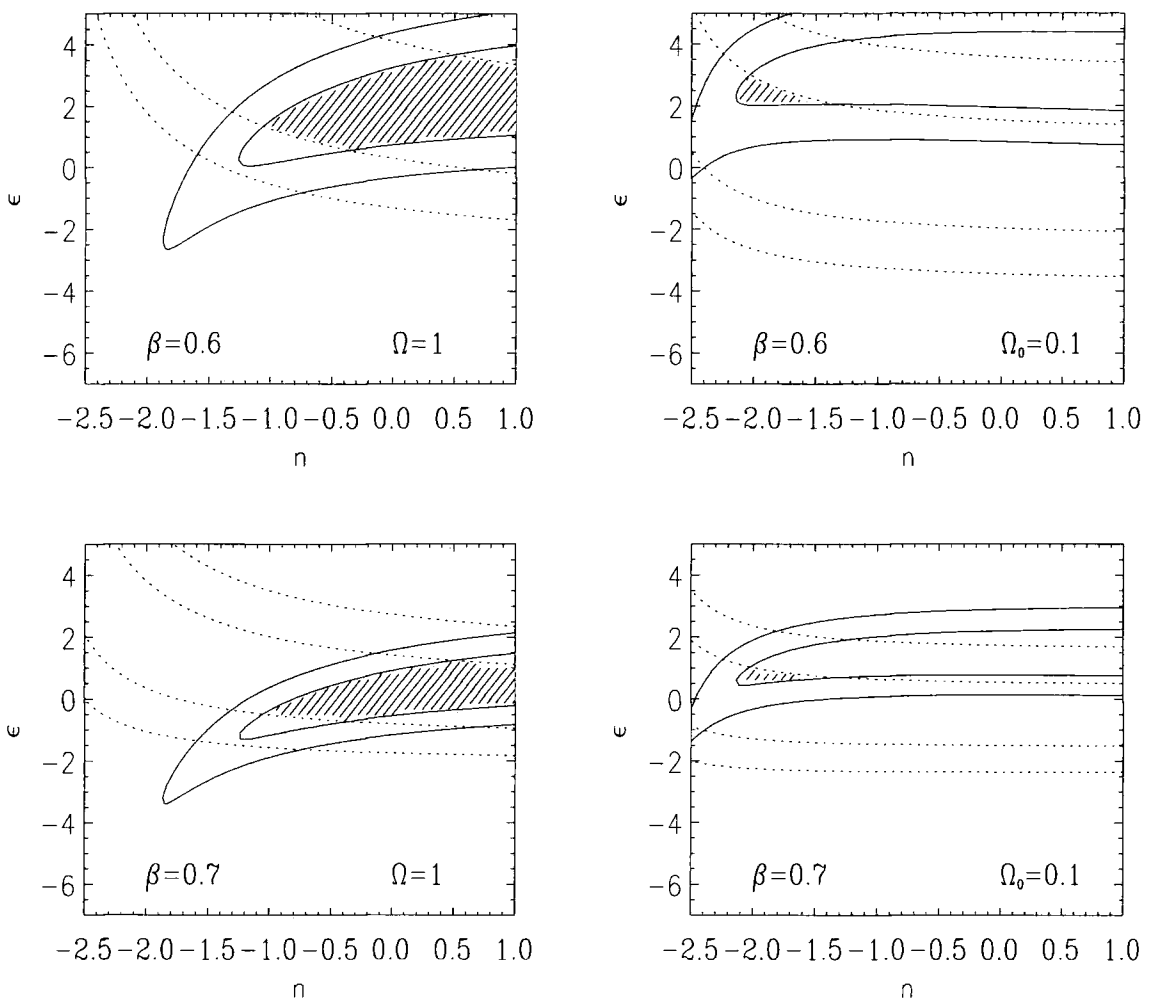


Figure 2.5: An example of the effect of varying β using the SHARC XLF data. The left column is for $\Omega = 1$ (the most mass evolution) and the right column $\Omega_0 = 0.1$ (the least mass evolution). The top and bottom rows for the considered limits for β (0.6 and 0.7 respectively). The effect on n is insignificant, although raising β lowers the range of ϵ.

2.4 Discussion

Several issues regarding this work are worthy topics for discussion. Firstly, the results presented in this chapter are based on several important assumptions made when modelling the X–ray evolution of clusters, so these choices and their robustness require to be discussed. Secondly, particularly due to the recent publication of *ROSAT* and *ASCA* data, several results constraining Ω_0 using the evolution of X–ray clusters have now appeared in the literature. Therefore it is necessary to compare the results from this work with those found by other groups. Finally, it is important to outline a strategy for future X–ray cluster surveys (particularly with *Chandra* and *XMM*), hence areas are pin–pointed that will maximize future success in this field.

2.4.1 Model assumptions

Although general and based on physically motivated scaling relations, the model is based on several key assumptions. One example is the β profile: the results presented in Section 2.3 assume that the X–ray emission profiles are well fit by the β – model, where the density of the intracluster gas tends to a constant value at small r (the core) and the isothermal result at large r, with an asymptotic slope of $2/3$. Individual measurements of β do not necessarily give this value, although it is assumed that the average value over the population at a given epoch is well represented by $\langle\beta\rangle = 2/3$. However, it is important to investigate the tolerance of the model when varying the value of β. Figure 2.5 illustrates the effect on the constrained range of n and ϵ that variations in β induces. Using the SHARC data as an example, two values have been taken, $\beta = 0.6$ and $\beta = 0.7$. Two cosmological models have been chosen : $\Omega = 1$ and $\Omega_0 = 0.1$, for which structure evolves the most and least rapidly (for fixed n) out of all the models considered in this study. Changing β can significantly change the range of ϵ values but has relatively little effect on the values of n. This, coupled with the work on the core radius discussed in the last section, reinforces the idea that both n and ϵ contribute separately to the evolution of clusters: the entropy evolution is sensitive to the way in which the gas is distributed, with β and r_c defining the shape of the gas density profile, while the mass evolution is tied in with the background density. Hence, the dependence of Ω_0 on β is weak, although certain models may be ruled out on the basis that the range of ϵ may be implausible. This is evident in the figure, where the combination $(\beta = 0.6, \Omega_0 = 0.1)$ pushes allowed values of ϵ into the limit set by the maximum cooling calculation ($\epsilon = 2.0$) given in Paper I.

With the *Chandra* satellite now launched, it will become possible to study the emission profiles of distant clusters in more detail. These studies may show that the cluster profile evolves with redshift. For example, simulations of cluster evolution with pre–heated gas by Mohr & Evrard (1997) suggest that the entropy injected reduces the overall slope of the gas density profile rather than just that of the profile of the very centre. This effect can be parameterised by a much slower role over near the core radius but the scaling of the solution remains the same, since the X–ray emission still scales as $\rho_c^2 r_c^3 T^\alpha$. Thus, as Figure 2.5 shows, the β profile and the ϵ parameter are interconnected – while observations of the emission profiles of high redshift clusters would prompt the revision of the physical interpretation of a particular ϵ value, and re–examine the balance of heating and cooling within the cluster, it would not invalidate the phenomenological description provided by the model.

The same argument may be applied to account for the non–isothermal temperature structure of clusters (as suggested by Markevitch et al. 1998). The model only requires that the temperature structure is normalised by the virial temperature of the dark matter halo. The temperature structure may affect the rate at which the cluster core evolves for a given entropy change, but this should again taken into account when the ϵ parameter is interpreted.

The model has been developed under the assumption that clusters of galaxies are in dynamical equilibrium. Clusters that are seen during major mergers may have X–ray luminosities and average temperatures that are far from their hydrostatic values. Nevertheless, the description that has been developed here applies to the cluster population as a whole, and the properties of individual clusters may differ wildly from the average. In this sense, the effects of mergers are already incorporated through the entropy evolution described by the model, and the disruption of hydrostatic equilibrium is mimicked by an increase in core entropy. The cluster temperatures are also susceptible to these departures, however, since they are weighted by the X-ray emission. Ideally, cluster temperatures should be measured from the outer regions of clusters where the gas temperature more closely reflects the virial temperature of the gravitational potential. These are not available for distant clusters at present, but the results will still be valid so long as the evolution of emission-weighted temperature reflects the evolution of the cluster potential. This is a problem common to all methods based on the X–ray temperatures of clusters.

The slope of the luminosity–temperature relation does not explicitly enter the determination for the evolution of clusters in this work. All that is required by the model is the evolution of the normalisation. It has been implicitly assumed that the slope remains fixed at its present–day value. This assertion, which corresponds to the assumption that the entropy evolution parameter (ϵ) is independent of cluster temperature, is unavoidable with the limited temperature data that is available for high redshift clusters. Consequently, it is important that the evolution of the normalisation is determined by using consistent slopes for high and low redshift data, and preferably using high redshift clusters representative of the clusters used to determine the luminosity function evolution. At present, temperature measurements are available only for the brightest clusters at high redshift. This situation should improve using the *XMM* satellite (Lumb et al. 1996); its large collecting area is ideally matched to the determination of temperatures for the lower luminosity, distant clusters. The slope of nearby ($z < 0.05$) clusters has been most recently determined by Markevitch (1998), using *ROSAT* luminosities and *ASCA* temperatures, measuring a value, $\lambda \sim 2.65$. However, this work has made use of the older sample based on *Einstein* MPC data (David et al. 1993) for two reasons. Firstly, constraints on the evolution of the $L_{\rm x} - T$ relation are provided by Mushotzky

& Scharf (1997), who use this (local) sample to contrast with their higher redshift ($\langle z \rangle \sim 0.3$) *ASCA* data. Also, the slope supplied by Markevitch has been calculated for clusters with cooling flows removed, in order to directly compare with models that exclude radiative components in the plasma evolution. Since the model presented here is based on evolution of the central entropy, the cooling flow clusters cannot be excluded, which probably contribute to most, if not all of the negative entropy evolution in cluster populations. It is argued that the dominant cause for the intrinsic scatter in the relation is due to each cluster having its own particular core entropy, coming from their individual formation histories.

The considerations discussed above relate to the distribution of gas within the dark matter potential. In a low density universe, it is necessary to explicitly account for the distinction between the background density of the Universe when the cluster is formed, and it's value when the cluster is observed. These two epochs were related by assuming that the mass of the cluster grows by a factor of two (Kitayama & Suto, 1996). However, for $\Omega_0 > 0.1$, this factor produces only a small change in the halo evolution.

2.4.2 Comparison with other methods

A number of papers have recently appeared in the literature dealing with the constraints on Ω_0 from the X–ray evolution of clusters. One of the most popular methods uses the evolution of the cluster abundances (e.g. Henry 1997; Eke et al. 1998; Viana & Liddle 1996; Mathiesen & Evrard 1998; Oukbir & Blanchard 1997; Reichart et al. 1999). To extract the value of Ω_0, an estimator for the cluster virial mass is required. This can be achieved by calculating the cluster temperature function at any given epoch and converting this to a mass function using the Press–Schechter formalism with the assumption that the cluster haloes form via spherical infall. The result from this method is still slightly unclear – at present the majority of authors find values of Ω_0 typically in the region $0.3 - 0.5$ with or without Λ, although some authors find higher values including $\Omega = 1$ (e.g. Sadat, Blanchard, & Oukbir 1998, Blanchard & Bartlett 1998). An extensive discussion on the existence of this clash in results can be found in Eke et al. (1998). In particular, they highlight the discrepancy in the abundance of low–redshift clusters, particularly those at the highest temperatures, where the measured value is uncertain. Those favouring high Ω_0 assume a higher abundance of clusters in this region than those favouring low values of Ω_0.

This work differs from these treatments in that it does not assume a detailed model for the abundance of clusters or for their temperature or X–ray luminosity distribution. Instead, the scaling relations have been emphasised, that must relate the properties of clusters at one epoch to those

at another. By introducing an additional parameter that describes the evolution of the core gas entropy, the approach allows cluster samples to be compared at different epochs with minimal further assumptions. Comparing these results with those obtained by other authors illuminates the constraints which are general and those which are model–specific. As an example, Reichart et al. (1998) also look at luminosity function evolution by adopting the Press–Schechter approach for the evolution of the mass function , but use the $L_{\rm x} - T$ relation to calibrate their mass to X–ray luminosity relation. In addition, their model differs from the one presented in this Chapter since it involves the slope of the $L_{\rm x} - T$ relation explicitly, resulting in a purely empirical calibration of the luminosity–mass conversion. By contrast, this approach uses the Weak Self–Similarity Principle (plus the entropy evolution) to scale the properties of clusters between epochs. As discussed extensively in Paper I, this is completely separate from the scaling relations between clusters of different mass at the same epoch. Trying to equate the two requires the Strong Self–Similarity Principle to be adopted; something for which there is little physical justification. Reichart et al. adopt the same parameterization as Evrard & Henry (1991), notably $L_{\rm x} \propto T^p(1+z)^s$; assuming $\beta = 2/3$, their parameters relate to n and ϵ as $p = 13/6$ and $s = (13 - 3\epsilon)/4$. However, it is emphasised, the way in which the parameters are constrained here (particularly ϵ in this context) is fundamentally different, and it should not be unexpected if identical results are not obtained.

The results presented here are compared with the work of Eke et al. (1998). These authors derive a value of Ω_0 from the evolution of the cluster temperature function, using emission–weighted temperatures obtained from *ASCA* data (Henry 1997). Since the cluster sample is X–ray flux limited, they have to correct for incompleteness in their estimation of the temperature function. This is done by assuming a non–evolving $L_{\rm x} - T$ relation (i.e. fixed slope, normalisation and scatter). The advantage of this method is that it is related to the underlying mass distribution in a more direct manner than the approach developed in this chapter, hence their work places a strong constraint on Ω_0: $\Omega_0 = 0.43 \pm 0.25$ for $\Lambda_0 = 0$ and $\Omega_0 = 0.36 \pm 0.25$ with a non–zero cosmological constant. This was achieved with a substantially smaller data–set. Furthermore, the authors were able to internally constrain the shape of the power spectrum. This is possible because they directly link the form of the temperature function to the mass function through the Press–Schechter model (assuming a CDM power spectrum). They arrive at a value of Γ of 0.08 ± 0.07 ($\Gamma = 0.09 \pm 0.08$ for $\Omega_0 + \Lambda_0 = 1$).

To compare the results of this work to those of Eke et al., their constrained range of Γ is converted to an equivalent range of n, using the same method discussed in Section 2.3.4. Figure 2.6 illustrates the results. The range of Ω_0 is then found that is both consistent with the X–ray data and

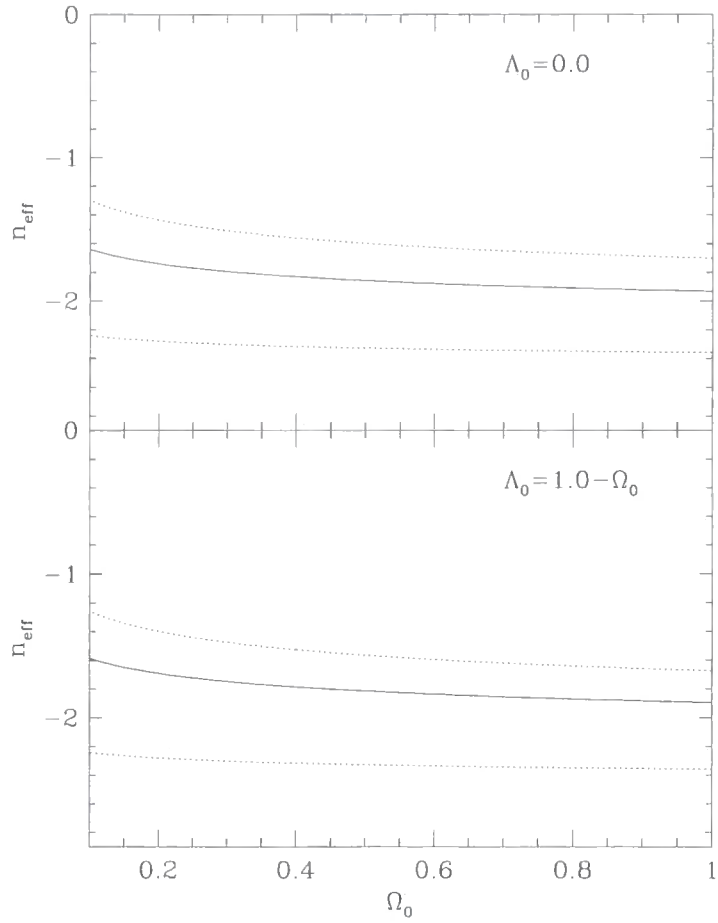


Figure 2.6: The relation between n and Ω_0 for the constrained range of Γ given by Eke et al. (1998), for the range of cosmological scenarios studied in this chapter (at $z = 0$). The solid lines are for the best-fit values and the dotted lines are the results using the quoted limits by the authors.

the calculated range of n. It is found that $\Omega_0 < 0.45 (\Omega_0 < 0.5$ for $\Omega_0 + \Lambda_0 = 1)$ at 95 per cent confidence, in agreement with their upper limit. Again, a constraint on the lower limit of Ω_0 is not possible. This is initially surprising, but seems to be due to systematic differences in the data. It is found that, for a low Ω_0 cosmology, the available data are consistent with weak negative evolution of the XLF and no evolution in the $L - T$ normalisation. By contrast, Eke et al.'s temperature function data (corrected for cosmology) shows a decline in temperature function amplitude at high redshift, suggesting that a very low value of Ω_0 is unacceptable due to the deficit of clusters at this epoch. Such dependence on the particular data set is worrying, but emphasises the importance of using independent data–sets to establish consistent conclusions and reduce systematic biases. It is satisfying to find that the method presented here finds similar results, at least in determining the upper limit.

2.4.3 Future directions

Since the results presented in this chapter were based on scaling the *ROSAT* BCS XLF to fit the higher redshift surveys, it is assumed that the two populations are well separated in redshift. This is not immediately obvious, since the *ROSAT* BCS sample has clusters out to $z \sim 0.3$, where the other samples start around this value. The saving grace is that the higher–redshift surveys have detected lower luminosity clusters ($L_x < L_x^*$), due to the small solid angles of the surveys, whereas the BCS clusters in this regime are at much lower redshifts. However, this range of luminosities places a much weaker constraint on the model than if high redshift data were available at the brighter end ($L_x > L_x^*$). Improved coverage will result in a better internal constraint, since the effects of evolving cluster number density and cluster luminosity will be more easily separated. As a pointer for the future, what is required is a more statistically sound description of the high–redshift XLF from L_x^* and above. As a hypothetical example, assume that the combined data–set used in this chapter is a good description of the XLF for luminosities around and below L_x^*. A measurement is then made at $L_x(0.5-2.0\,\mathrm{keV}) \sim 10^{45}\,\mathrm{erg\,s^{-1}}$, detecting 10 clusters at $\langle z \rangle = 0.5$ in a solid angle that resulted in a cumulative abundance in agreement with the BCS cumulative XLF ($N(>L_x) \sim 0.25 \times 10^{-8}\,\mathrm{Mpc^{-3}}$). This strongly suggests a detection of no evolution, particularly for the $\Omega = 1$ geometry. Using the APM constraint on n then implies $\Omega_0 < 0.2$ at the 99 per cent level. In practice, the *XMM* satellite is particularly suited to this task, with its large collecting area and high spatial resolution, that will allow one to easily make the distinction between clusters and AGN. What is required, however, is a wide area survey covering at least 500 square degrees, rather than a survey that is particularly deep.

2.5 Conclusions

In this chapter, the entropy–driven model for cluster evolution, detailed in Bower (1997) for an $\Omega = 1$ Universe, has been modified to account for evolution in different cosmological scenarios, specifically open models and flat, non–zero Λ models. This allows us to separate contributions made by the hierarchical growth of structure (controlled by the slope of the power spectrum, n) and changes in the core entropy of the intracluster gas (controlled by the entropy evolution parameter, ϵ).

Constraints were then placed on n and ϵ for seven reasonable cosmological models, using the current wave of X–ray measurements of the X–ray Luminosity Function and the Luminosity–Temperature relation. For the $L_x - T$ relation, the compiled low and high redshift samples

from David et al. (1993) and Mushotzky & Scharf (1997) respectively were used, determining the best–fit slope and scatter. Subsequently, 68 and 95 per cent confidence limits were placed on the evolution of the temperature normalisation, which is slope–dependent. For the XLF, recently available *ROSAT* samples were used, specifically the SHARC (Burke et al. 1997), WARPS (Jones et al. 1998) and RDCS (Rosati et al. 1998) high–redshift, non–parametric determinations, as well as the older EMSS samples (Henry et al. 1992, Luppino & Gioia 1995). The amount of evolution in these measurements was quantified by comparing them with the local *ROSAT* BCS XLF (Ebeling et al. 1997). Using the X–ray data alone, acceptable regions of parameter space are found for every cosmological model considered, to a high degree of confidence. As the density parameter decreases, more negative values of n are tolerated, due to the weakening growth rate of clusters. Constrained values of ϵ are appropriate to modify the luminosity evolution to compensate for the measured change in abundance, moving weakly from positive values to zero (no evolution) for low–density Universes.

To break the degeneracy between n and Ω_0, the slope of the power spectrum was calculated based on the CDM paradigm, by taking the typical mass of a rich cluster and calculating the corresponding fluctuation scales that give rise to such objects. We find that the CDM calculation sets a limit of $\Omega_0 < 0.55$ ($0.1 < \Omega_0 < 0.7$ for $\Omega_0 + \Lambda_0 = 1$). Using the value determined by the APM survey ($n = -1.6 \pm 0.3$), it can be concluded that $\Omega_0 < 0.6$ (< 0.65 for $\Omega_0 + \Lambda_0 = 1$). All limits were calculated to 95 per cent confidence. In general, a lower limit for the value of Ω_0 cannot be determined, since the model fits the data much more comfortably, at fixed n, for lower density Universes. The evolution of the cluster core–radius was also considered, where it was showed that it depends much more sensitively on the entropy evolution of clusters than the structural evolution (determined by n and Ω_0). Present data suggest no strong evolution out to $z \sim 0.5$ (Vikhlinin et al. 1998), roughly consistent with our results given the quoted uncertainties.

Finally, the choice and robustness of assumptions made in the model were discussed, particularly the effect of varying the mass distribution of the gas (controlled by the value of β). Again, this has a much stronger effect on the constrained range of ϵ than the rate of structural evolution, and hence only weakly affects the determination of Ω_0. The results were then compared to those determined by other groups, particularly with Eke et al. (1998). It was found that the upper limit constrained from this work is consistent with theirs ($\Omega_0 < 0.45$, or $\Omega_0 < 0.5$ for $\Omega_0 + \Lambda_0 = 1$ at 95 per cent confidence) although a lower limit cannot be placed on Ω_0.

In order to provide tighter limits on the value of Ω_0, it will be essential to obtain an accurate measurement of the bright end of the XLF out to redshifts at least comparable to present distant

cluster samples, ideally from a wide angle survey. It is hoped that this will be provided by the *XMM* satellite.

References

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Blanchard A., Bartlett J. G., 1998, A&A, 332, L49

Bower R. G. et al., 1996, MNRAS, 281, 59

Bower R. G., 1997, MNRAS, 288, 355

Bower R. G., Smail I., 1997, MNRAS, 290, 292

Burke D. J., Collins C. A., Sharples R. M., Romer A. K., Holden B. P., Nichol R. C., 1997, ApJ, 488, L83

Cash W., 1979, ApJ, 228, 939

Cavaliere A., Fusco-Femiano R., 1976, A&A, 49, 137

Collins C. A., Burke D. J., Romer A. K., Sharples R. M., Nichol R. C., 1997, ApJ, 479, L117

Copi C. J., Schramm D. N., Turner M. S., 1995, ApJ, 455, 95

David L. P., Slyz A., Jones C., Forman W., Vrtilek S. D., 1993, ApJ, 412, 479

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

De Grandi S., Molendi S., Böhringer H., Chincarini G., 1997, ApJ, 486, 738

Dekel A., Burstein D., White S. D. M., 1997, in Turok N., ed, Critical Dialogues in Cosmology (Princeton 250th Anniversary). World Scientific, p. 175

Ebeling H., Edge A. C., Fabian A. C., Allen S. W., Crawford C. S., Böringer H., 1997, ApJ, 479, L101

Eke V. R., Cole S. M., Frenk C. S., 1996, MNRAS, 282, 263

Eke V. R., Cole S. M., Frenk C. S., Henry J. P., 1998, MNRAS, 298, 1145

Eke V. R., Navarro J. F., Frenk C. S., 1998, ApJ, 503, 569

Evrard A. E., 1989, ApJ, 341, L71

Gioia I. M., Luppino G. A., 1994, ApJS, 94, 583

Hasinger G., Burg R., Giacconi R., Schmidt M., Trumper J., 1998, A&A, 329, 482

Henry J. P., 1997, ApJ, 489, L1

Henry J. P., Gioia I. M., Maccacaro T., Morris S. L., Stocke J. T., Wolter A., 1992, ApJ, 386, 408

Jones C., Forman W., 1984, ApJ, 276, 38

Jones L. R., Scharf C., Perlman E., Ebeling H., Horner D., Wegner G., Malkan M., McHardy I., 1998, AN, 319, 87

Kaiser N., 1986, MNRAS, 222, 323

Kaiser N., 1991, ApJ, 383, 104

Kitayama T., Suto Y., 1996, ApJ, 469, 480

Lacey C., Cole S., 1993, MNRAS, 262, 627

Lahav O., Rees M. J., Lilje P. B., Primack J. R., 1991, MNRAS, 251, L128

Lumb D., Eggel K., Laine R., Peacock A., 1996, SPIE, 2808, 326

Luppino G. A., Gioia I. M., 1995, ApJ, 445, L77

Markevitch M., 1998, ApJ, 504, 27

Markevitch M., Forman W. R., Sarazin C. L., Vikhlinin A., 1998, ApJ, 503, 77

Mathiesen B., Evrard A. E., 1998, MNRAS, 295, 769

McHardy I. M. et al. ., 1998, AN, 319, 51

Mohr J. J., Evrard A. E., 1997, ApJ, 491, 38

Mushotzky R. F., Scharf C. A., 1997, ApJ, 485, L65

Oukbir J., Blanchard A., 1997, A&A, 317, 10

Peacock J. A., Dodds S. J., 1994, MNRAS, 267, 1020

Peebles P. J. E., 1980, The large scale structure of the Universe. Princeton University Press

Press W. H., Schechter P., 1974, ApJ, 187, 425

Reichart D. E., Nichol R. C., Castander F. J., Burke D. J., Romer A. K., Holden B. P., Collins C. A., Ulmer M. P., 1999, ApJ, submitted

Rosati P., Ceca R. D., Norman C., Giacconi R., 1998, ApJ, 492, L21

Sadat R., Blanchard A., Oukbir J., 1998, A&A, 329, 21

Scharf C. A., Jones L. R., Ebeling H., Perlman E., Maklan M., Wegner G., 1997, ApJ, 477, 79

Smail I., Ellis R. S., Dressler A., Couch W. J., Oemler A., Butcher H., Sharples R. M., 1997, ApJ, 479, 70

Sugiyama N., 1995, ApJS, 100, 281

Tadros H., Efstathiou G., Dalton G., 1998, MNRAS, 296, 995

Tsuru T., Koyama K., Hughes J. P., Arimoto N., Kii T., Hattori M., 1996, in The 11th international colloquium on UV and X-ray spectroscopy of Astrophysical and Laboratory Plasmas

Viana V. L., Liddle A. R., 1996, MNRAS, 281, 323

Vikhlinin A., McNamara B. R., Forman W., Jones C., Quintana H., Hornstrup A., 1998, ApJ, 498, L21

White S. D. M., Efstathiou G., Frenk C. S., 1993, MNRAS, 262, 1023

Chapter 3

Cosmological simulations: methodology

Abstract. In this chapter, the method of cosmological simulation is outlined. An overview of current techniques is first presented before introducing the system of equations necessary to follow the evolution of a cosmological fluid. The procedure for generating initial conditions for cosmological simulations follows. The remainder of the chapter focuses on the algorithms implemented in HYDRA, the code used to generate results for the remainder of this thesis. In particular, the methods used to model gravity and hydrodynamics are presented, as well as the numerical integration scheme.

3.1 Overview

Cosmological simulations are an extremely powerful method for modelling the growth of structure in the universe. The strength of this approach lies in their ability to self-consistently model the non-linear evolution of density fluctuations, necessary for studying the formation of objects such as galaxies and clusters. Analytical methods are generally not possible without introducing oversimplifying assumptions, such as artificial symmetries, into the problem.

The minimum requirement for all cosmological simulations is to calculate the evolution of the gravitational potential. (For modelling the evolution of the dark matter, this is all that is required.) On cosmological scales, Newtonian gravity suffices since general relativistic effects produce corrections that are negligible. The most popular method for modelling the effects of gravity is the N-body simulation. This method involves representing the density field by a set of N point-masses, or *particles*, each carrying its own position and velocity. Gravitational forces are calculated at the positions of each particle due to the presence of the other $N - 1$ particles. The forces are then used to update the positions and velocities of the particles via a numerical integration scheme.

Early N-body simulations calculated gravitational forces directly by summing the pairwise contributions (e.g. Aarseth 1963). The problem with this method is that it rapidly becomes impractical

for systems containing N greater than a few thousand, since the calculation time scales rapidly, as $\sim O(N^2)$. One solution is to use special-purpose hardware designed to rapidly compute N^2 calculations. An example of this is the GRAPE hardware (e.g. Sugimoto et al. 1990); several groups interested in cosmological problems have made use of this device (e.g. Eke, Navarro, & Frenk 1998; Steinmetz & Müller 1995).

Another method is to develop a more efficient algorithm. The majority of groups have chosen this option. In particular, there have been two types of algorithms developed, each with its own strengths. One of the methods is the *tree* algorithm (e.g. Barnes & Hut 1986; Hernquist & Katz 1989). This algorithm hinges on the assumption that the level of detail required in calculating the potential should decrease with distance. The calculation proceeds by recursively subdividing the simulation volume into cells until each particle is housed within a separate cell. The hierarchy of cells is known as the tree, each subdivision is a branch and cells containing a single particle are leaves. While the tree is being constructed, multipole terms up to a finite order are calculated and stored for each branch. For small separations, the forces on each particle are calculated directly, whereas long range forces are calculated using the tree. Starting at a top branch, an *opening angle* is calculated, $\Delta = l/x$, where l is the length of the branch cell and x is the distance to the centre of the branch cell from the position of the current particle. If Δ is smaller than some tolerance, only the multipole moments for that branch are used and all sub-branches are discarded. Otherwise the branch is *opened* and the same procedure is applied to all sub-branches until either the opening angle becomes sufficiently small or all leaves have been reached.

The tree force calculation scales $\sim O(N \log N)$, and is therefore much more efficient for large numbers of particles than direct summation. For cosmological simulations, periodic boundary conditions are required in order to satisfy the Cosmological Principle, i.e. the large-scale mass distribution should be homogeneous and isotropic. Treecodes do not automatically satisfy this condition. However this can be rectified by using the Ewald summation mechanism (Ewald 1921).

The other popular method is known as the P^3M (Particle-Particle, Particle-Mesh) algorithm (e.g. Hockney & Eastwood 1981; Efstathiou et al. 1985; Evrard 1988). This method splits the force calculation into two sections. Large scale forces are calculated using the Particle-Mesh algorithm (PM). This proceeds by smoothing the particles onto a regular cubic mesh and solving Poisson's equation in Fourier space (Hockney & Eastwood 1981; Efstathiou & Eastwood 1981). It is a very efficient method: the use of FFT's determine a computational speed which scales as $\sim O(L^3 \log L)$, where L^3 is the total number of mesh cells. For a fixed ratio of N/L^3, this

translates to $\sim O(N \log N)$ and, like the tree force calculation, becomes rapidly more efficient with increasing N than using direct summation methods.

The disadvantage of using the PM algorithm alone is that the fixed mesh makes it very inflexible for cosmological problems. The mesh imposes a global minimum length scale (the Nyquist scale), below which no useful information can be obtained. Improving the force resolution by increasing the size of the mesh is both impractical and inefficient; the memory requirement of the mesh scales rapidly (L^3), and becomes extremely wasteful in low density environments. The P^3M algorithm rectifies this by calculating smaller scale forces via a direct sum over near neighbours (the PP method).

The P^3M method is particularly suited to cosmological simulations since using FFT's naturally provides periodic boundary conditions. Its main drawback is that it performs increasingly poorly as the level of clustering rises. The formation of structure leads to the majority of the work shifting from the PM algorithm (when the particles are well-separated), to the slower PP algorithm. An adaptive version of P^3M was developed (AP^3M; Couchman 1991) which redistributes work between the PM and PP algorithms when such problems occur. Regions that are highly clustered are identified and segregated from the normal PP calculation. A submesh (or refinement) is then placed over each region and the P^3M calculation is individually applied. The algorithm is recursive, hence refinements can themselves contain refinements.

For many cosmological problems the evolution of the baryons is required to be included explicitly in the modelling, since it is the signatures of this species that are directly observed. For most purposes it is assumed that the baryons behave like an ideal gas, therefore they are able to interact collisionally. For cosmological simulations, the most common method for calculating hydrodynamical forces is the Smoothed Particle Hydrodynamics algorithm (SPH; Gingold & Monaghan 1977; Lucy 1977). Detailed reviews of this technique are available in the literature (e.g. Monaghan 1992; Benz 1990). Like N-body methods, the SPH technique is based on using a set of particles to represent the baryon distribution. Quantities are calculated at the positions of particles by summing smoothed contributions from near neighbours; unlike gravity, gas forces are short-range. The SPH algorithm will be discussed in more detail below.

The code used to perform the simulations for this thesis is HYDRA (Couchman, Thomas, & Pearce 1995), which uses the AP^3M algorithm to calculate gravitational forces and SPH for the hydrodynamics. The rest of the chapter is structured around the implementations of the algorithms Version 3 of this code. However, it is appropriate firstly to introduce the system of equations that govern the evolution of self-gravitating fluids within the

context of an expanding background; this is the purpose of Section 3.2. The method for generating initial conditions is outlined in Section 3.3. The main algorithms in HYDRA are then discussed in more detail in Sections 3.4 (AP^3M), 3.5 (SPH) and 3.6 (numerical integration). This is not intended to be an exhaustive review since the method of HYDRA is already well documented in the literature (Couchman 1991; Thomas & Couchman 1992; Couchman, Thomas, & Pearce 1995; Pearce & Couchman 1997; Thacker et al. 1998). Finally, the chapter is summarised in Section 3.7.

3.2 The system of equations

For studies of structure formation it is useful to write the system of equations in the comoving reference frame (i.e. at rest with respect to the large scale expansion of the universe) using the coordinate transformation $\mathbf{x} = \mathbf{r}/a$, where $\mathbf{r}$ is a physical position and a is the expansion factor, normalised to the present value. The peculiar velocity field is then $a\mathbf{v} = (\dot{\mathbf{r}} - \dot{a}\mathbf{x})$, where dotted superscripts refer to derivatives in time. Forces are calculated using the following equation of motion (evaluated at position $\mathbf{x}$; see Peebles 1980)

$$a\frac{D\mathbf{v}}{Dt} + 2\dot{a}\mathbf{v} + \nabla\phi + \frac{1}{\rho}\nabla P = 0, \tag{3.1}$$

where ϕ, ρ and P are the gravitational potential, mass density and pressure of the fluid respectively. The second term in Equation 3.1 is present due to the change to the comoving reference frame; its effect is known as the Hubble drag, since it opposes the motion of objects as a result of the Hubble expansion. The third term is the gravitational acceleration, $\nabla\phi = \mathbf{f}(\mathbf{x})/a^2$; for a Newtonian force law,

$$\mathbf{f}(\mathbf{x}) = -G\int \frac{\rho(\mathbf{x}')\,(\mathbf{x}' - \mathbf{x})d\mathbf{x}'}{|\mathbf{x}' - \mathbf{x}|^3}, \tag{3.2}$$

assuming a density field continuous over all space. The pressure is calculated assuming the equation of state

$$P_i = w\rho_i u_i, \tag{3.3}$$

where u is the specific thermal energy and w is a constant between 0 and 1, known as the Zel'dovich interval. For collisionless material such as cold dark matter, $w = 0$ and for gas, $w = \gamma - 1$, where γ is the ratio of specific heats at constant pressure and volume respectively. The work in this thesis assumes an equation of state for a monatomic ideal gas (i.e. $\gamma = 5/3$).

For the gas, the evolution of the thermal energy is determined by the first law of thermodynamics, i.e.

$$\frac{Du}{Dt} + 2Hu + a\frac{P}{\rho}\nabla.\mathbf{v} - \frac{\Gamma - \Lambda}{\rho} = 0, \tag{3.4}$$

where $H = \dot{a}/a$. The second term comes from the change to comoving co-ordinates: the gas loses thermal energy to the adiabatic expansion of the universe. The third term is the contribution from additional adiabatic processes while the final term in Equation 3.4 are the contributions from heat sources (Γ) and sinks (Λ), able to change the entropy of the gas.

3.3 Initial conditions

Initial conditions are generated for cosmological simulations assuming that mass fluctuations are small relative to the background, and can therefore be adequately described by linear theory, in which analytical solutions are readily obtained (Peebles 1980). Since the gas pressure is weak at early times, all that is required to specify the initial positions and velocities is the gravitational potential. In order to calculate this, it is useful to define the overdensity field as

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}, \tag{3.5}$$

where $\bar{\rho}$ is the mean density of the background. Hence, for linear fluctuations, $\delta \ll 1$. Due to the discrete nature of simulations, the overdensity field is not required to be known everywhere, but rather it is sampled at regular intervals on a cubic mesh, with L cells per side. The field can then be fully specified using a linear combination of Fourier modes

$$\delta(\mathbf{x}) = \frac{1}{V} \sum_{\mathbf{n}=-\infty}^{\infty} \tilde{\delta}(\mathbf{k}) \exp(-i\mathbf{k}.\mathbf{x}), \tag{3.6}$$

where $V = S^3$ is the volume of the region under study, $\mathbf{k} = (2\pi/S)\mathbf{n}$ and $\mathbf{n}$ is an integer triple. In order to prevent *aliasing*, the limits of the sum are set such that $1 \leq |\mathbf{n}| \leq L/2$, i.e. physically from S to the Nyquist interval, $\Delta = 2S/L$. The Fourier co-efficients, $\tilde{\delta}(\mathbf{k})$, are complex variables and hence can be separated into (real) amplitude and phase components, i.e.

$$\tilde{\delta}(\mathbf{k}) = |\tilde{\delta}(\mathbf{k})| \exp(i\,\phi_{\mathbf{k}}), \tag{3.7}$$

with the phase, $\phi_{\mathbf{k}}$, being in the range, $0 \leq \phi_{\mathbf{k}} \leq 2\pi$. Assuming the initial fluctuations are Gaussian (e.g. Bardeen et al. 1986) simplifies matters since the physical information only enters through the amplitudes; the phase is a random variable. Amplitudes are drawn from a Rayleigh distribution with mean equal to the desired power spectrum, $\left\langle |\tilde{\delta}(\mathbf{k})|^2 \right\rangle = P(k)$. The following approximation to the CDM linear power spectrum is used (Bond & Efstathiou 1984)

$$\begin{aligned} P(k) &= P_0\, k\, T^2(k) \\ &= \frac{P_0\, k}{[1 + (6.4q + (3.0q)^{1.5} + (1.7q)^2)^{1.13}]^{\frac{2}{1.13}}}, \end{aligned} \tag{3.8}$$

where P_0 is the power spectrum amplitude and $T(k)$ is the CDM transfer function that modifies the primordial Harrison-Zel'dovich spectrum, $P(k) \propto k$ (which results from inflation). The variable $q = k/\Gamma$ where $\Gamma \sim \Omega_0 h$ is the shape parameter. The power spectrum is usually normalised such that it satisfies

$$\sigma_8^2 = \frac{V}{2\pi^2} \int_0^\infty k^2 P(k)\, \tilde{W_8}^2(k) dk. \tag{3.9}$$

The quantity σ_8^2 is the variance of linear mass fluctuations in spheres of radius $8\,h^{-1}$ Mpc, obtained by multiplying the power spectrum with the Fourier transform of a top-hat filter, $\tilde{W_8}$; its value is estimated from the local abundance of rich clusters (Evrard 1989; White, Efstathiou, & Frenk 1993; Eke, Cole, & Frenk 1996; Viana & Liddle 1996).

In Fourier space, the density field, $\delta(\tilde{\mathbf{k}})$ is then multiplied by the Green's function of the Laplacian to obtain the potential (see Efstathiou et al. 1985)

$$\tilde{\phi}(\mathbf{k}) = \tilde{\delta}(\mathbf{k})\, \tilde{G}(\mathbf{k}). \tag{3.10}$$

Performing an inverse Fourier transform gives the potential tabulated at each mesh point. This is then differenced to obtain gravitational accelerations, $\mathbf{f}(\mathbf{x})$.

Knowledge of the accelerations allows both a displacement and velocity field to be constructed. In the linear regime, the initial displacement field, Ψ is related to the initial ($\mathbf{q}$) and subsequent ($\mathbf{x}(\mathbf{q}, \mathbf{t})$) positions of the mass distribution as

$$\Psi(\mathbf{q}) = \frac{\mathbf{x}(\mathbf{q}, \mathbf{t}) - \mathbf{q}}{D(t)}, \tag{3.11}$$

(Zel'Dovich 1970). The function, $D(t)$ is the normalised linear growing mode solution (refer to Section 2.2.3). The velocity field, $\mathbf{v}(\mathbf{x}, \mathbf{t}) = \Psi(\mathbf{q})\dot{D}(t)$ and acceleration field, $\mathbf{a}(\mathbf{x}, \mathbf{t}) = \Psi(\mathbf{q})\ddot{D}(t)$ follow. Substituting both these equations into Equation 3.1 implies that

$$\Psi(\mathbf{q}) = -\frac{\mathbf{f}(\mathbf{q})}{a^2(a\ddot{D} + 2\dot{a}\dot{D})}. \tag{3.12}$$

The simplest method for assigning positions and velocities to the particles is to perturb them from the same positions on the grid where the displacement field was evaluated. An alternative method is to create a relaxed distribution of particles known as a *glass* (White 1993). This has the advantage that there is no artificial symmetry imposed on the initial distribution (as is the case with the grid). A glass is created by first placing the particles in the volume at random positions. They are then evolved by calculating gravitational forces with a reversed sign, until their velocities become vanishingly small. The displacements and velocities (evaluated on a grid) are then interpolated onto the particles using methods identical to those discussed in the next section.

For the gas, the additional variables are the density, ρ and the thermal energy u (or equivalently, the temperature, T). Initial values for the gas density field are calculated by the SPH algorithm, based on the particle positions. Values for the initial gas temperatures depend on the problem under study, but are usually chosen such that the gas is placed on an adiabat below the virial temperature of any systems resolved by the simulation.

3.4 Gravity

3.4.1 The AP^{3}M *algorithm*

As discussed earlier, the AP^{3}M algorithm consists of three tasks: the PM calculation, which calculates long range forces; the PP calculation, which calculates short range forces; and refinement placing, to speed up the calculation within highly overdense regions. The PM calculation can be broken up into a number of steps:

1. **Mass assignment:** firstly, an estimate of the density field is obtained, sampled at every mesh interval (or grid point), by smoothing the particles onto the grid

$$\rho(\mathbf{n}) = \sum_{i=1}^{N} m_i W(\mathbf{x}_i - \mathbf{n}), \tag{3.13}$$

where subscript i labels each of the N particles contributing to the estimate of ρ_i at grid point $\mathbf{n}$ and W is the smoothing kernel used. Let $x_\alpha = \mathbf{x}.\mathbf{e}_\alpha$ and $n_\alpha = \mathbf{n}.\mathbf{e}_\alpha$, where $\mathbf{e}_\alpha(\alpha = 1, 2, 3)$ are orthonormal vectors along the x, y and z directions respectively. Also assume that particle positions are in the range $0 \leq \mathbf{x_i} \leq L$. A simple example of mass assignment is the Nearest Grid Point (NGP) method, where

$$W_\alpha = \begin{cases} 1, & L|x_\alpha - n_\alpha| \leq \frac{1}{2}; \\ 0, & \text{otherwise}, \end{cases} \tag{3.14}$$

where $W = W_1 W_2 W_3$. Hence, particles only contribute to the estimate at one cell. The problem with the NGP method is that the sharpness of the kernel shape results in a discontinuous density field (Efstathiou et al. 1985). Smoother fields are obtained by successively performing a self-convolution on the kernel (the limit being a Gaussian), at the expense of more computational effort as the range of the kernel increases. HYDRA uses the Triangular Shaped Cloud (TSC) kernel

$$W_\alpha = \begin{cases} \frac{3}{4} - L^2|x_\alpha - n_\alpha|^2 & L|x_\alpha - n_\alpha| \leq \frac{1}{2}; \\ \frac{1}{2}\left(\frac{3}{2} - L|x_\alpha - n_\alpha|\right)^2 & \frac{1}{2} \leq L|x_\alpha - n_\alpha| \leq \frac{3}{2}; \\ 0 & \text{otherwise}. \end{cases} \tag{3.15}$$

This produces continuity in both the density field and its first derivative at the expense of each particle contributing to the local 27 grid points.

2. **Potential Calculation:** in order to calculate the potential, the density field must be convolved with an appropriate Green's function. It is faster to do this in Fourier space ($\sim O(L^3 logL)$), since the convolution becomes a trivial multiplication. The density field is Fourier transformed using a standard Fast-Fourier Transform (FFT) algorithm

$$\tilde{\rho}(\mathbf{k}) = \sum_{\mathbf{n}} \rho(\mathbf{n}) \exp(i\mathbf{k}.\mathbf{n}). \tag{3.16}$$

This is multiplied by the Fourier transform of the chosen Green's function to obtain the Fourier equivalent of the potential

$$\tilde{\phi}(\mathbf{k}) = \tilde{\rho}(\mathbf{k})\,\tilde{G}(\mathbf{k}). \tag{3.17}$$

before being inverse transformed, giving $\phi(\mathbf{n})$.

3. **Force Calculation:** the force at each mesh point is obtained by calculating potential differences with neighbouring cells, using a finite (10-point) differencing scheme (Couchman 1995).

4. **Reverse interpolation:** finally, the forces are interpolated back on to the positions of the particles by performing the reverse of step one. The same kernel must be used in order to conserve momentum.

The PM force is augmented by the PP force on separations below a critical length scale, r_e, defined by the maximum error of the force calculation. For separations much larger than the Nyquist wavelength of the FFT mesh, the PM force accurately matches the desired force law. However, as the separation decreases, an error is introduced, mainly from the inability of the PM force to resolve density fluctuations on small scales. A balance is therefore sought between accuracy and efficiency: a larger value of r_e leads to more accurate forces but is computationally more expensive since the amount of PP work increases. HYDRA uses a value such that the maximum deviation from the calculated force to the true force is 7.7 per cent (Couchman, Thomas, & Pearce 1995). Typical force errors are much smaller than this, the rms being ~ 0.5 per cent. This sets the value of r_e to be approximately 2.16 FFT mesh spacings.

The PP force is calculated by searching for neighbours that lie within a distance r_e from each particle. HYDRA achieves this by defining a chaining mesh (the PP mesh), with the number of cells, L_s set equal to the nearest integer below $L/2.16$, where L is the number of FFT cells/side

(e.g. $L = 64 \rightarrow L_s = 29$). Link lists are then constructed for each cell. The PP force contribution at position $\mathbf{x_i}$ is therefore

$$\mathbf{F}_{PP}(\mathbf{x_i}) = \sum_{j \neq i} \mathbf{F}(|\mathbf{x_i} - \mathbf{x_j}|) - \mathbf{F}_{PM}(|\mathbf{x_i} - \mathbf{x_j}|), \tag{3.18}$$

where the first term on the right hand side is the sum of pairwise forces and the second term is the contribution from the PM force.

Regions with a high level of clustering are isolated from the PP calculation with the refinement placing algorithm. This is done by highlighting cells with constituent numbers of particles larger than some optimum value. Refinements (which at present, can only be cubical and non-overlapping) are then placed by finding peaks in the cell counts and growing regions until all of the hot cells are covered. This is then passed through a set of optimization routines, designed to search for and optimize various pathological cases (for example, merging clusters of particles). The P^3M algorithm is then performed within these regions. Hence, much of the work that would have been done by the PP algorithm is passed onto the faster PM algorithm. This procedure is recursive, hence refinements can have refinements and so on.

3.4.2 The force law

It is not practical to model the gravitational force using the Newtonian inverse-square law, for two reasons:

1. as particles approach one another the resultant accelerations become arbitrarily large. This has disastrous implications for the size of the timestep, since as $|\Delta \mathbf{x}| \rightarrow 0, \Delta t \rightarrow 0$ in order to preserve numerical accuracy.

2. particles in close proximity become more susceptible to the effects of two-body encounters, relaxing on a timescale

$$\tau_2 \sim \frac{\sigma^3}{8\pi G^2 \rho m \ln \Lambda}, \tag{3.19}$$

where σ is the velocity dispersion, m is the particle mass and $\ln\Lambda$ is the Coulomb logarithm, with Λ being the ratio of the maximum and minimum length scale of the system under consideration. Clearly two-body relaxation should be viewed as unphysical since in cosmological simulations the point masses do not represent individual entities but are merely tracers of the mass distribution. Therefore the smallest value of τ_2 should be large compared to the age of the Universe.

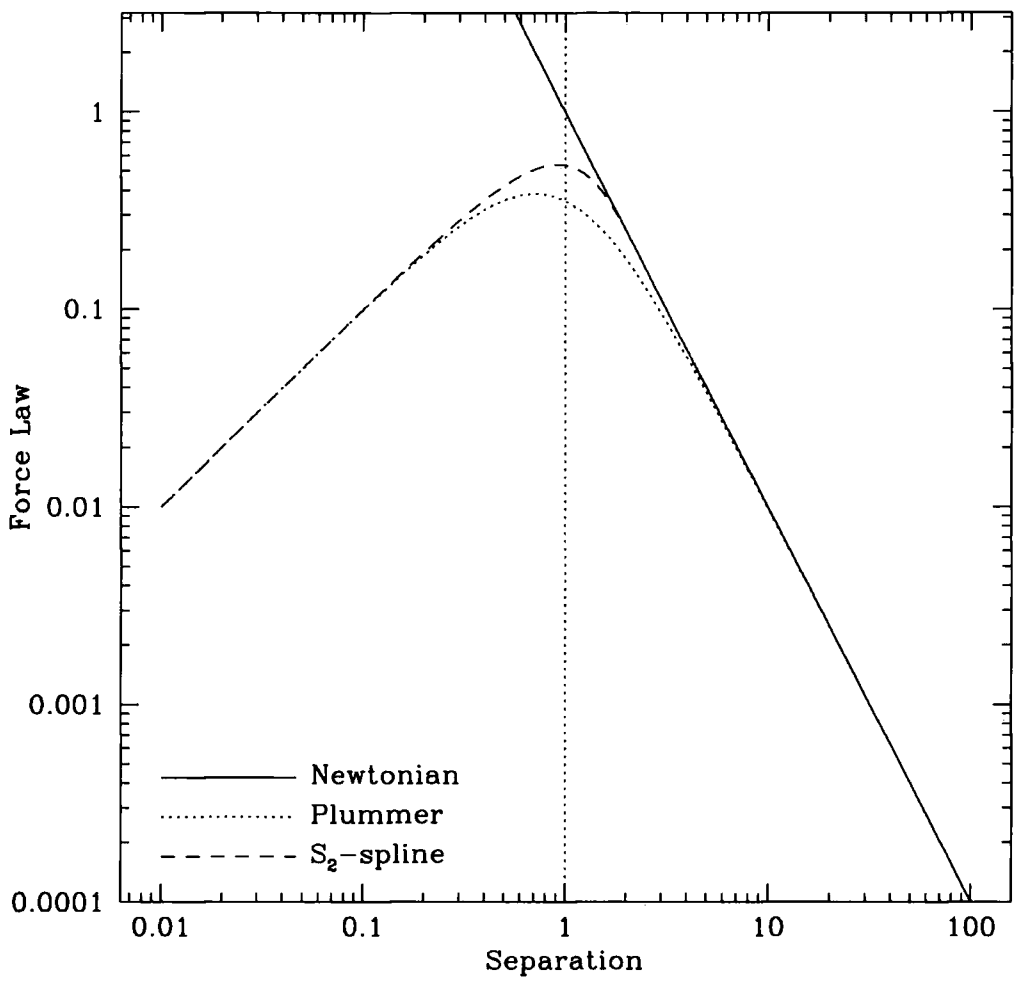


Figure 3.1: The Newtonian (solid line), Plummer (dotted line) and S_2 (dashed line) force shapes plotted against separation, in units of the Plummer softening, ϵ. The S_2 law is used by HYDRA.

Both problems are alleviated by *softening* the force law on small scales. This removes the singularity of the Newtonian law whilst increasing the value of τ_2 through the value of $\ln\Lambda$. HYDRA uses an S_2-spline softening function (Couchman 1991), parameterised by the S_2 softening length, s. Figure 3.1 illustrates the S_2 force shape compared to the Newtonian and Plummer laws (Binney & Tremaine 1987). The Plummer softening length, ϵ, has been chosen such that $s = 2.34\epsilon$, which produces asymptotically matching values on both large and small scales. Clearly the S_2 shape approaches the Newtonian law at smaller radii than the Plummer law, making it a more desirable choice.

The evolution of the softening length (in comoving co-ordinates) follows

$$s(z) = \min\left(s_{\max}, s_0(1+z)\right), \tag{3.20}$$

where z is the current redshift and s_0 and $s_{\max}$ are the redshift zero and maximum softenings. Before starting the simulation, the value of s_0 is input and the value of $s_{\max}$ is set by the code to be equal to 0.3 FFT cells. For a typical calculation, the softening stays constant at $s_{\max}$ until $z \sim 1$, when it shrinks, reaching the value s_0 at the present day.

3.5 Smoothed Particle Hydrodynamics

SPH is an interpolation method, hence the estimate of a scalar, vector or tensor field at position x is a convolution of the *true* field with a smoothing kernel, W. For example, the estimate of a scalar field, A(**x**) is

$$\langle A(\mathbf{x})\rangle = \int_{\mathbf{x}'} A(\mathbf{x}')\, W(|\mathbf{x}-\mathbf{x}'|, h)\, d\mathbf{x}', \tag{3.21}$$

where W is generally spherically symmetric and h is the kernel smoothing length. The kernel must satisfy two conditions. Firstly, it must be normalised to satisfy conservation laws. Secondly, it must approach the Dirac delta function for vanishingly small width ($W(|\mathbf{x}-\mathbf{x}'|, h) \rightarrow \delta(|\mathbf{x}-\mathbf{x}'|$ as $h \rightarrow 0$). The smoothing kernel used in HYDRA is the B_2–spline kernel (Monaghan & Lattanzio 1986)

$$W(u,h) = \frac{1}{4\pi h^3}\begin{cases} 4-6u^2+3u^3, & 0 \le u \le 1 \\ (2-u)^3, & 1 < u \le 2 \\ 0, & u > 2 \end{cases} \tag{3.22}$$

where $u = |\Delta\mathbf{x}|/h$. The kernel gradient is modified such that particles feel an enhanced repulsive force when sufficiently close to one another, to prevent artificial clustering (Thomas & Couchman 1992)

$$\frac{dW}{du}(u,h) = -\frac{1}{4\pi h^3}\begin{cases} 4, & 0 \le u \le 2/3 \\ 3u(4-3u), & 2/3 < u \le 1 \\ 3(2-u)^2, & 1 < u \le 2 \\ 0, & u > 2. \end{cases} \tag{3.23}$$

Within the context of simulations, the continuous field must be replaced by a set of discrete points and therefore a discrete approximation to Equation 3.21 is used

$$\langle A(\mathbf{x}_i)\rangle = \sum_j m_j \frac{A_j}{\rho_j} W(|\mathbf{x}_i - \mathbf{x}_j|, h), \tag{3.24}$$

where the index j labels the particles which have mass m_j, position $\mathbf{x}_j$, density ρ_j. Hence particles are used both as interpolation centres ($\mathbf{x_i}$) and interpolation contributors ($\mathbf{x_j}$). As an example, the SPH estimate of the gas density at position $\mathbf{x_i}$ is simply

$$\langle \rho_i(\mathbf{x_i}) \rangle = \sum_j m_j \, W(|\Delta \mathbf{x}|, h), \tag{3.25}$$

with $\Delta \mathbf{x} = \mathbf{x_j} - \mathbf{x_i}$. It is also possible to write associated vector derivatives of functions in this formalism to the same degree of accuracy (i.e. $\sim O(h^2)$), involving only the kernel gradient (e.g. Hernquist & Katz 1989; Benz 1990). Thus, SPH is used to approximate solutions to the thermal energy equation (Equation 3.4) and the momentum equation (Equation 3.1). The energy equation (minus the heating and cooling terms) is used in HYDRA in the following form

$$\frac{du_i}{dt} = \sum_j m_j \, \Delta \mathbf{v}.\nabla W_{ij} \left(\frac{P_i}{\rho_i} + \frac{\Pi_{ij}}{2} \right), \tag{3.26}$$

with $\Delta \mathbf{v} = \mathbf{v}_j - \mathbf{v}_i$. Similarly for the momentum equation (minus the gravitational term)

$$\frac{d\mathbf{v}_i}{dt} = -\sum_j m_j \left(\frac{P_i}{\rho_i^2} + \frac{P_j}{\rho_j^2} + \Pi_{ij} \right) \nabla W_{ij} \tag{3.27}$$

The Π_{ij} term in both equations represents an artificial viscosity, necessary to capture shocks within the flow and efficiently convert bulk kinetic energy into heat. The viscosity is modelled using the standard function (Monaghan & Gingold 1983)

$$\Pi_{ij} = \frac{-\alpha \bar{c}_{ij} \mu_{ij} + \beta \mu_{ij}^2}{\tilde{\rho}_{ij}}, \tag{3.28}$$

where

$$\mu_{ij} = \begin{cases} \bar{h}_{ij} \mathbf{v}_{ij}.\mathbf{r}_{ij} / (r_{ij}^2 + \nu^2), & \mathbf{v}_{ij}.\mathbf{r}_{ij} < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3.29}$$

The values of α and β are chosen to equal 1 and 2 respectively, which determine the strength of the first (bulk) and second (Von Neumann-Richtmeyer) contributions to the viscosity of the fluid. The term $\bar{c}_{ij} = 1/2(c_i + c_j)$ is the average adiabatic sound speed of particle pairs ($c_i = \gamma(\gamma - 1)u$), $\bar{h}_{ij}$ is their average SPH smoothing length and $\nu = 0.01 \bar{h}_{ij}^2$ is a correction term to prevent numerical divergence when particles are in close proximity of one another. The approximation for the average density, $\tilde{\rho}_{ij} = \rho_i(1 + (h_i/h_j)^3)/2$ is used, which avoids expensive re-calculations of neighbour lists (Thacker et al. 1998).

For systems with attractive forces such as gravity, it is not ideal to use a fixed value of h for every particle. As regions collapse, the number of particles within a fixed radius will increase dramatically. Therefore each particle carries its own smoothing length that is adaptive in both

space and time. The value of h is set such that there are approximately $N_{\rm SPH}$ neighbours within a radius of $2h$ of each particle (the range of the kernel). Typical values for cosmological simulations are $N_{\rm SPH} = 30 - 100$. Hence, particles in low density regions will have larger values of h than particles in high density regions.

To avoid extensive computational effort, the value of h is set using an iterative procedure which converges on the correct number of neighbours for each particle, providing there are no abrupt changes in the flow. HYDRA incorporates a weighted- average scheme (Thacker et al. 1998) which removes instabilities present in the solutions of previous algorithms (Hernquist & Katz 1989).

The introduction of individual smoothing lengths violates the conservation laws since particles may have very different values of h which leads to an imbalance in the forces. To prevent this, quantities are calculated between particle pairs using either an average of their smoothing lengths or an average of the kernel itself. The latter, known as kernel averaging or symmetrization, is used in HYDRA , specifically

$$W_{ij} = \frac{1}{2}\left[W(|\Delta \mathbf{x}_{ij}|, h_i) + W(|\Delta \mathbf{x}_{ij}|, h_j)\right] \tag{3.30}$$

(Hernquist & Katz 1989; Steinmetz & Müller 1993).

3.6 Numerical integration

Solutions to Equations 3.1 & 3.4 are approximated using a numerical integration scheme. The numerical integration algorithm used in HYDRA is a simple PEC (Predict, Evaluate, Correct) method (Couchman, Thomas, & Pearce 1995), which utilises only the current particle positions and velocities (plus densities and thermal energies for the gas). This has the advantage of allowing for arbitrary changes in the timestep, which can happen where there are abrupt changes in the forces, for example in the vicinity of shocks.

Firstly, the size of the timestep is evaluated. To obtain this, HYDRA calculates four timescales at various appropriate stages of the force calculation; the first two are relevant to all particles while the third is relevant to the gas particles.

1. $\Delta t_a = (\epsilon/a_{\rm max})^{1/2}$, where ϵ is the current (Plummer) softening length and $a_{\rm max}$ is the magnitude of the largest particle acceleration.

2. $\Delta t_v = \epsilon/v_{\rm max}$, where $v_{\rm max}$ is the current largest particle speed.

3. $\Delta t_h = \min(h/|\Delta \mathbf{v}|)$, where h is the gas particle's SPH smoothing length and $|\Delta \mathbf{v}|$ is the magnitude of the velocity difference between each gas particle and one of its neighbours.

4. $\Delta t_H = 1/H$, where $H = \dot{a}/a$ is the Hubble parameter.

The chosen timestep is then

$$\Delta t = \kappa \min(0.4\Delta t_v, 0.25\Delta t_a, 0.2\Delta t_h, 0.0625\Delta t_H). \tag{3.31}$$

The constants are chosen to achieve sufficient numerical accuracy for the integration (Couchman, Thomas, & Pearce 1995); the value of κ is typically set to unity for adiabatic simulations (Thomas & Couchman 1992). Note that a single timestep is used to advance all particles - in principle, a further gain in efficiency could be obtained by implementing individual timesteps for each particle.

The particles are then advanced using the PEC scheme. Let the current step be denoted by the subscript n. Particle positions and velocities are first predicted for step $n+1$ using the current accelerations

$$\mathbf{x}'_{n+1} = \mathbf{x}_n + \mathbf{v}_n \Delta t + \mathbf{f_n} \Delta t^2/2, \tag{3.32}$$

$$\mathbf{v}'_{n+1} = \mathbf{v}_n + \mathbf{f}_n \Delta t. \tag{3.33}$$

Accelerations and thermal energy derivatives ($\dot{u} \equiv du/dt$) are then evaluated for step $n+1$, based on the new positions and velocities

$$\mathbf{f}'_{n+1} = \mathbf{f}(\mathbf{x}'_{n+1}, \mathbf{v}'_{n+1}), \tag{3.34}$$

$$\dot{u}'_{n+1} = \dot{u}(\mathbf{x}'_{n+1}, \mathbf{v}'_{n+1}). \tag{3.35}$$

Finally, the quantities are then corrected to include the new accelerations

$$\mathbf{x}_{n+1} = \mathbf{x}'_{n+1} + \alpha\,(\mathbf{f}'_{n+1} - \mathbf{f}_n)\,\Delta t^2/2, \tag{3.36}$$

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \beta\,(\mathbf{f}'_{n+1} - \mathbf{f}_n)\,\Delta t, \tag{3.37}$$

$$u_{n+1} = u_n + \beta\,(\dot{u}'_{n+1} - \dot{u}'_n)\,\Delta t. \tag{3.38}$$

The values of α and β are set to $1/3$ and $1/2$ respectively (Couchman, Thomas, & Pearce 1995).

3.7 Summary

In this Chapter, the method of cosmological simulations has been outlined with particular reference to HYDRA (Couchman, Thomas, & Pearce 1995). The various stages of this method were then discussed, starting with the procedure of generating the initial conditions, before introducing the algorithms used to calculate the gravitational forces ($\mathrm{AP^3M}$) and hydrodynamical forces for the gas (SPH). Finally, the numerical integration scheme was outlined.

References

Aarseth S. J., 1963, MNRAS, 126, 223

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Barnes J., Hut P., 1986, Nature, 324

Benz W., 1990, in Buchler J. R., ed, Numerical Modelling of Stellar Pulsations: Problems and Prospects. Kluwer, Dordrecht, p. 269

Binney J., Tremaine S., 1987, Galactic Dynamics. Princeton Series in Astrophysics

Bond J. R., Efstathiou G., 1984, ApJ, 285, L45

Couchman H. M. P., 1991, ApJ, 368, L23

Couchman H. M. P., 1995, in Numerical Methods in Astrophysics (New York: Springer), Vol. 2

Couchman H. M. P., Thomas P. A., Pearce F. R., 1995, ApJ, 452, 797

Efstathiou G., Davis M., Frenk C. S., White S. D. M., 1985, ApJS, 57, 241

Efstathiou G., Eastwood J. W., 1981, MNRAS, 194, 503

Eke V. R., Cole S. M., Frenk C. S., 1996, MNRAS, 282, 263

Eke V. R., Navarro J. F., Frenk C. S., 1998, ApJ, 503, 569

Evrard A. E., 1988, MNRAS, 235, 911

Evrard A. E., 1989, ApJ, 341, L71

Ewald P. P., 1921, Ann. Phys., 64, 253

Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375

Hernquist L., Katz N., 1989, ApJS, 70, 419

Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles. McGraw-Hill

Lucy L., 1977, AJ, 82, 1013

Monaghan J. J., 1992, ARA&A, 30, 543

Monaghan J. J., Lattanzio J. C., 1986, A&A, 158, 207

Pearce F. R., Couchman H. M. P., 1997, New Ast., 2, 411

Peebles P. J. E., 1980, The large scale structure of the Universe. Princeton University Press

Steinmetz M., Müller E., 1993, A&A, 268, 391

Steinmetz M., Müller E., 1995, MNRAS, 276, 549

Sugimoto D., Chikada Y., Makino J., Ito T., Ebisuzaki T., Umemura M., 1990, Nature, 345, 33

Thacker R. J., Tittley E. R., Pearce F. R., Couchman H. M. P., Thomas P. A., 1998, MNRAS, submitted (astro-ph/9809221)

Thomas P. A., Couchman H. M. P., 1992, MNRAS, 257, 11

Viana V. L., Liddle A. R., 1996, MNRAS, 281, 323

White S. D. M., 1993, in Formation and Evolution of Galaxies: Les Houches Lectures

White S. D. M., Efstathiou G., Frenk C. S., 1993, MNRAS, 262, 1023

Zel'Dovich Y. B., 1970, A&A, 5, 84

Chapter 4

Parameter tests within simulations of galaxy formation

Abstract. Simulations of galaxy formation require a number of parameters. Some of these are intrinsic to the numerical algorithms (e.g. the timestep), while others describe the physical model (e.g. the gas metallicity). In this chapter, results are presented for a systematic exploration of the effects of varying a subset of these parameters. A fiducial model is compared to 24 different simulations, in which one parameter at a time is varied. For reasonable choices of values, many parameters have relatively little effect on the galaxies, with the notable exception of the parameters that control the resolution of the simulation and the efficiency with which gas cools.

4.1 Introduction

Over the past few years, the use of numerical simulations to model the formation of galaxies within a proper cosmological context has become reality. There are two reasons for this. Firstly, there now exists sufficient computing power in order to probe the dynamic range required for resolving galaxies within large volumes (at least four orders of magnitude in length scale). This is due to the progress in both computer hardware, particularly massively parallel machines, and the development of fast algorithms to calculate forces (as discussed in the previous chapter).

The second reason is that there now exists a physical framework in which the modelling of galaxy formation can take place. It was evident in early studies (Hoyle 1953) that the formation of stars and galaxies is driven by dissipative processes, notably radiative cooling. Several studies subsequently applied this to the collapse and fragmentation of inhomogeneous gas clouds (Binney 1977; Rees & Ostriker 1977; Silk 1977). The results were encouraging, since they predicted that the largest mass able to cool within a free–fall time is comparable to that of bright galaxies ($\sim 10^{12}\mathrm{M}_{\odot}$). Further work (White & Rees 1978) placed this in a more physically realistic context, in which galaxies formed via the condensation of gas within the cores of dark matter haloes. This naturally explains the discontinuity in mass between galaxies and clusters, as well

as predicting a galaxy luminosity function in rough agreement with the observations, at least for the bright galaxies. An excess of faint galaxies was predicted, first revealing the need for non-gravitational heating processes (such as *feedback* from stars) to circumvent overcooling, and hence too much star formation at high redshift.

The advances made in the understanding of the large-scale structure (dominated by collisionless dark matter) allowed galaxy formation to be modelled within a proper cosmological context. N-body simulations were used to calculate the clustering evolution of collisionless dark matter (e.g. Frenk, White, & Davis 1983; Klypin & Shandarin 1983; Centrella & Melott 1983; Davis et al. 1985), leading to the idea of *biased* galaxy formation (e.g. Frenk et al. 1988 and references therein). The statistics of Gaussian random fields (see Bardeen et al. 1986), and the formulation of conditional halo mass functions, based on Press-Schecter theory (Press & Schechter 1974; Bond et al. 1991; Bower 1991) were also developed. Thus, cosmological, *ab initio* galaxy formation was born (White & Frenk 1991; Cole 1991). These *semi–analytical* approaches have now become into an industry (e.g. Cole et al. 1994; Kauffmann, Guiderdoni, & White 1994; Baugh, Cole, & Frenk 1996; Somerville & Primack 1998; Cole et al. 1999) in which many successful predictions have been related to the wealth of new observations of distant objects, for example the Lyman–break galaxies (Governato et al. 1998) and quasars (Haehnelt & Kauffmann 1999). However, the techniques rely on several physical assumptions, many which can only be reliably tested using numerical simulations.

First attempts at modelling cosmological galaxy formation using numerical techniques have concentrated on the simplest possible model (e.g. Katz, Hernquist, & Weinberg 1992; Cen & Ostriker 1992; Evrard, Summers, & Davis 1994; Pearce et al. 1999). The evolution of baryons and CDM are followed within a suitably large volume. The dark matter only interacts gravitationally whereas the gas (with a non-zero equation of state) can also interact via pressure forces (adiabatic heating and cooling). Furthermore, it is able to change its entropy by the ability to shock–heat and cool radiatively. In general, the latter process presents a formidable numerical challenge because in hierarchical clustering theories, cooling is most efficient in small objects at high redshift where cooling times are very short. As a result, the outcome of a simulation depends on its resolution. Hence modelling the effects of reheating/feedback is also beneficial from a numerical perspective; the hope is that convergent results can be obtained.

Although resolution is the most important numerical concern in simulations of galaxy formation, several other numerical and physical parameters influence the outcome of a simulation. Examples of the former are the gravitational softening or the size of the timestep; examples of the latter are

the metallicity of the gas or the initial redshift. Often parameter values are set in the absence of rigorous criteria. In this chapter a systematic study is undertaken of a *limited* subset of the parameters required in simulations of galaxy formation. Various relevant quantities are measured, for example the amount of gas that cools, the number and masses of the galaxies that form, etc. This is not intended as an exhaustive investigation, but rather as a guide to the sensitivity of simulation results to these parameters. For simplicity, this chapter ignores the effects of photo–ionization (studied previously, for example, by Weinberg, Hernquist, & Katz 1997 and Navarro & Steinmetz 1997) and feedback.

The rest of this chapter is organized as follows. In Section 4.2 the numerical method is outlined, in particular the approach to radiative cooling is discussed. In Section 4.3 a "fiducial" simulation is defined its evolution is described, focussing on the distribution of gas in the temperature–density plane and on the properties of "galaxies." In Section 4.4 this fiducial simulation is compared to a set of simulations in which one parameter at a time is varied. In particular, the correspondence of galaxies in different simulations is considered as well as properties such as their mass distribution and the distribution of baryons within dark matter haloes. The results are summarised in Section 4.5 and conclusions are drawn in Section 4.6.

4.2 Numerical method

All simulations have been performed using Version 1 of HYDRA , which uses identical gravity algorithms to those described in the previous chapter. However, there are several differences in the SPH technique for the version of HYDRA used here, as opposed to the rest of this thesis (which uses Version 3). For more details of the SPH in Version 1 of HYDRA , refer to Thomas & Couchman (1992).

The basic minimum in modelling galaxy formation is the incorporation of radiative cooling effects, so that the gas can achieve high enough density contrasts where star formation can proceed. Radiative cooling requires including a sink in the thermal energy equation, as discussed previously (see Equation 3.4). It is assumed that the gas is locally in collisional ionization equilibrium (CIE), i.e. that no additional particle species are being created or destroyed. This allows the emissivity (power radiated per unit volume) to be factored into two independent functions of density and temperature

$$\Lambda(n,T) = n^2\,\Lambda_c(T), \tag{4.1}$$

where n is the number density of free particles and $\Lambda_c(T)$ is the cooling function. Gas temperature is defined in the usual manner, i.e. $T = (\gamma - 1)\mu m_p\, u$, where γ is the ratio of specific heats, m_p is the mass of a proton and μ is the mean atomic mass, assumed to take the value for a primordial cosmological composition, $\mu \sim 0.59$. The adopted cooling function is a simple series of power law fits to the optically-thin radiative cooling code of Raymond, Cox & Smith (1976)

$$\begin{aligned}\Lambda_c(T < 10^5\mathrm{K}) &= (\alpha_1 + \alpha_2\, Z)T, \\ \Lambda_c(T \geq 10^5\mathrm{K}) &= \alpha_3\, T^{0.5} + \frac{\alpha_4}{T} + \frac{\alpha_5\, Z}{T^{0.8}},\end{aligned} \tag{4.2}$$

where T is the temperature in Kelvin and the α_i's are (dimensional) constants,

$$\begin{aligned}\alpha_1 &= 6.1 \times 10^{-28}\mathrm{erg\,s^{-1}\,cm^3\,K^{-1}}, \\ \alpha_2 &= 7.4 \times 10^{-27}\mathrm{erg\,s^{-1}\,cm^3\,K^{-1}}, \\ \alpha_3 &= 2.3 \times 10^{-27}\mathrm{erg\,s^{-1}\,cm^3\,K^{-0.5}}, \\ \alpha_4 &= 6.1 \times 10^{-18}\mathrm{erg\,s^{-1}\,cm^3\,K}, \\ \alpha_5 &= 7.4 \times 10^{-18}\mathrm{erg\,s^{-1}\,cm^3\,K^{0.8}}.\end{aligned}$$

Contributions to Λ_c come from thermal bremsstrahlung and line cooling due to Hydrogen, Helium and a variable content due to heavier elements, parameterised by the gas metallicity, Z in solar units. Three examples of Λ_c for $Z/Z_\odot = 0.0, 0.5, 1.0$ are illustrated in Figure 4.1. Overlaid is the cooling function taken from the work of Sutherland & Dopita (1993).(Version 3 of HYDRA uses the latter cooling function.)

In cosmological simulations, the cooling timescale is often shorter than the dynamical or sound-crossing timescales. However it is not computationally feasible to shrink the timestep and follow cooling explicitly. One alternative is to restrict the timestep in such a way that each particle only loses up to half of its thermal energy in one step (Katz & Gunn 1991; Gerritsen & Icke 1997). The method used in HYDRA takes a simpler approach: hydrodynamical forces are calculated for the gas particles using the adiabatic and viscous contributions to the thermal energy equation. Energy is then removed, at constant density, due to radiative cooling at the end of the step. The amount subtracted, Δu_i, is calculated such that

$$\int_{u_i}^{u_i - \Delta u_i} \frac{du}{\Lambda(n_i, T_i)} = -\frac{\Delta t}{\rho_i} \tag{4.3}$$

with subscript i denoting the i^{th} particle and u, its specific thermal energy. This results in stable and physically reasonable behaviour, even for relatively long timesteps.

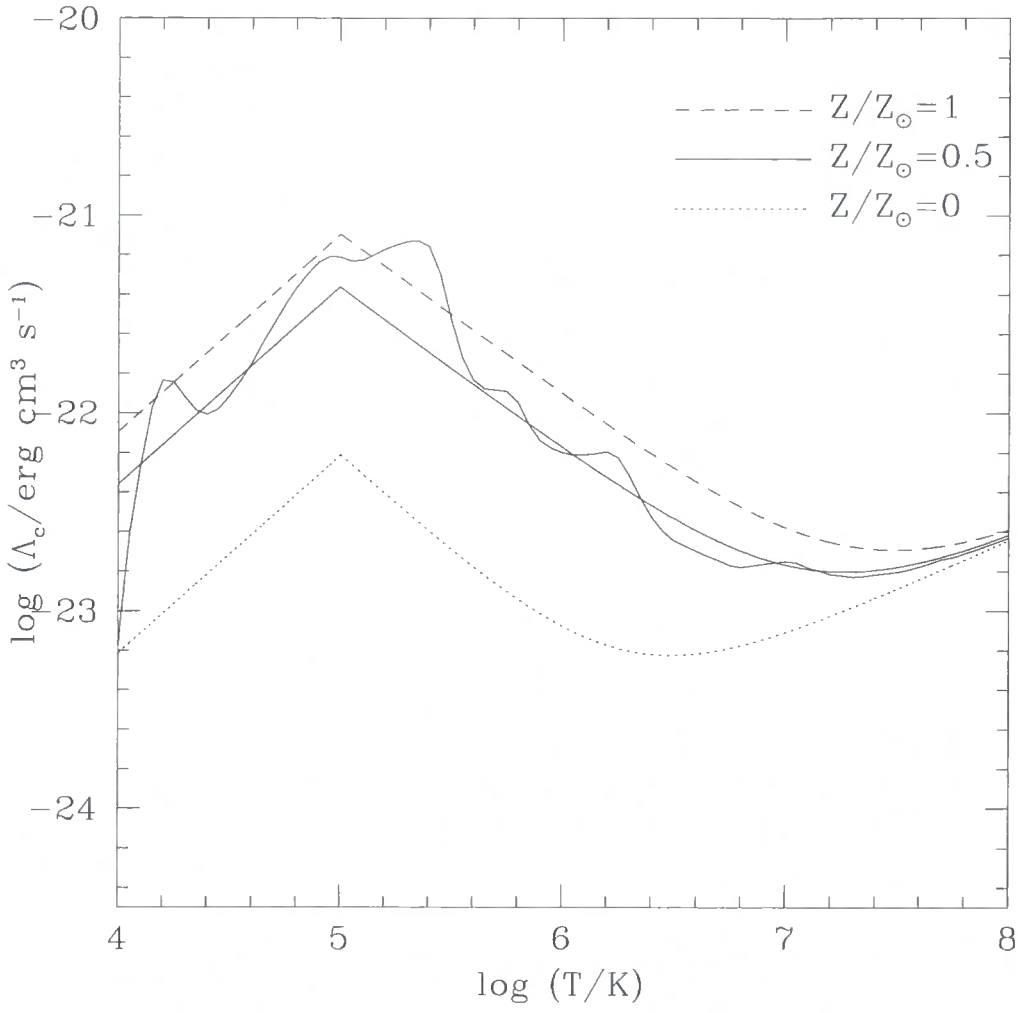


Figure 4.1: Illustrations of the cooling functions used in this chapter, which are power law fits (to functions generated by the code of Raymond, Cox, & Smith 1976). Three variations for gas metallicities, $Z/Z_{\odot} = 0.0, 0.5, 1.0$, are shown. Overlaid for the $Z = 0.5Z_{\odot}$ case is a tabulated cooling function from the work of Sutherland & Dopita (1993). The functions are only plotted for temperatures down to 10^4 K since this is where the cooling is switched off. (Physically, negligible thermal energy is lost below this value.)

4.3 The fiducial simulation

4.3.1 Initial conditions

Throughout this chapter the effects of varying various numerical and physical parameters are examined relative to a fiducial simulation. The fiducial initial conditions were chosen to correspond to the standard cold dark matter model (SCDM). This sets the value of the cosmic density parameter, $\Omega = 1$, the cosmological constant, $\Lambda = 0$ and the Hubble constant, $h = 0.5$ [1]. Although

[1] $H_0 = 100\,h\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$

Table 4.1. Values of the parameters chosen for the fiducial simulation that were subsequently varied. N is the total number of particles, $N_{\rm SPH}$ is the number of gas particles over which the SPH smoothing is carried out, L is the size of the FFT mesh, $s_{\rm max}$ is the maximum comoving size of the softening in units of FFT cells, ϵ_0 is the effective Plummer softening at $z = 0$, κ is the timestep normalisation, $\Omega_{\rm b}$ is the baryon fraction, Z is the gas metallicity, z_i is the initial redshift and T_i is the initial gas temperature. The choice of these parameters is discussed in more detail in the text.

Parameter	Value
N	2×32^3
$N_{\rm SPH}$	32
L	64
$s_{\rm max}$	0.3
ϵ_0	$10\ h^{-1}$ kpc
κ	1
$\Omega_{\rm b}$	0.06
Z	$0.5 Z_\odot$
z_i	24
T_i	100K

this model is now out of favour (for example, its power spectrum cannot simultaneously fit both the large–scale *COBE* normalisation and the amplitude of small–scale galaxy clustering), it is a well studied model with moderately large power on all scales relevant to galaxy formation. In any case, the precise form of the initial conditions is not of concern here.

The initial mass distribution was generated as described in Chapter 3, by perturbing a uniform distribution of particles using the Zel'dovich approximation. The amplitude of the power spectrum was set such that $\sigma_8 = 0.6$, consistent with measurements of the local abundance of rich clusters (White, Efstathiou, & Frenk 1993, Viana & Liddle 1996; Eke, Cole, & Frenk 1996). The size of the simulation volume (in comoving coordinates) was fixed at $S = 10\ h^{-1}$ Mpc, with an FFT mesh resolution of $L = 64$ (implying a cell–length of $156.25\ h^{-1}$ kpc). Although the volume is not big enough to follow the evolution of large scale structure accurately (density fluctuations on the scale of the box are not evolving linearly at the final time), this is not important for this study since the results are only generated for comparison with each other and not real galaxy surveys.

Table 4.1 summarises the fiducial values of all the parameters that are studied in this chapter. The mass distribution was represented by 32^3 gas and 32^3 dark matter particles. The initial softening was set to 0.3 FFT cells (equivalent to a comoving Plummer softening of $\sim 20\ h^{-1}$ kpc), but after

$z \sim 1$ it switched to a fixed physical softening of $10\ h^{-1}$ kpc. The number of neighbours over which the SPH algorithm smooths was set to $N_{\rm SPH} = 32$. The fiducial simulation was run from an initial epoch, $z_i = 24$, to the present ($z = 0$), with the data output at regular intervals.

Regarding the physical parameters, the baryon fraction was set equal to $\Omega_{\rm b} = 0.06$, consistent with the results from nucleosynthesis calculations (Copi, Schramm, & Turner 1995). The global gas metallicity was chosen to be unevolving, with a value of $Z = 0.5 Z_{\odot}$, similar to results obtained from spectral observations of intracluster gas (e.g. Mushotzky et al. 1996). The initial gas temperature was set to 100K, well below the minimum virial temperature of resolved haloes in the simulation ($T \sim 5 \times 10^5$K).

4.3.2 The baryon phase diagram at $z = 0$

The two objects that are studied in this chapter are the dark matter haloes and the cold, dense clumps of gas subsequently referred to as the "galaxies". All comparisons are carried out at $z = 0$. The selection criteria used to catalogue the haloes and the galaxies are described in detail in section 4.3.4, but as motivation for the galaxy selection procedure, the correspondence between the spatial distribution and the thermal evolution of the baryons is developed in more detail. Figure 4.2 illustrates the temperature–density distribution of the gas at $z = 0$, spanning 7 orders of magnitude in density and 9 orders of magnitude in temperature. The thermal distribution is split into three *phases*, labelled as the *uncollapsed* phase, the *shocked* phase and the *galaxy* phase. Accompanying this figure are three dot–plots, showing the corresponding spatial distribution of each phase.

4.3.2.1 The uncollapsed phase

The uncollapsed phase is defined as the regime in which gas particles have $T \leq 1$K and $\rho < 10^3 \langle\rho\rangle$. The boundaries for this phase are chosen such that they enclose the particles that lie on the power law relation, which is an adiabat. These particles are still largely in free expansion and are, on average, cooling adiabatically with $T \propto (1+z)^2$. Spatially, the particles are homogeneously distributed, as a result of occupying a small range in density. The hard lower limit on this quantity (visible for all temperatures) is artificial, imposed by the algorithm due to the maximum distance it can search for $N_{\rm SPH}$ neighbours. For the fiducial simulation, $\rho_{\rm min} \sim 1.9 \langle\rho\rangle$, clearly an undesirably large value. As a result, the low density gas is forced to unphysically high values (uncollapsed gas should have $\rho < \langle\rho\rangle$). The latest version of HYDRA , which is used in subsequent chapters) now properly accounts for low density gas and it has been checked that this makes no significant difference to the properties of the high density material.

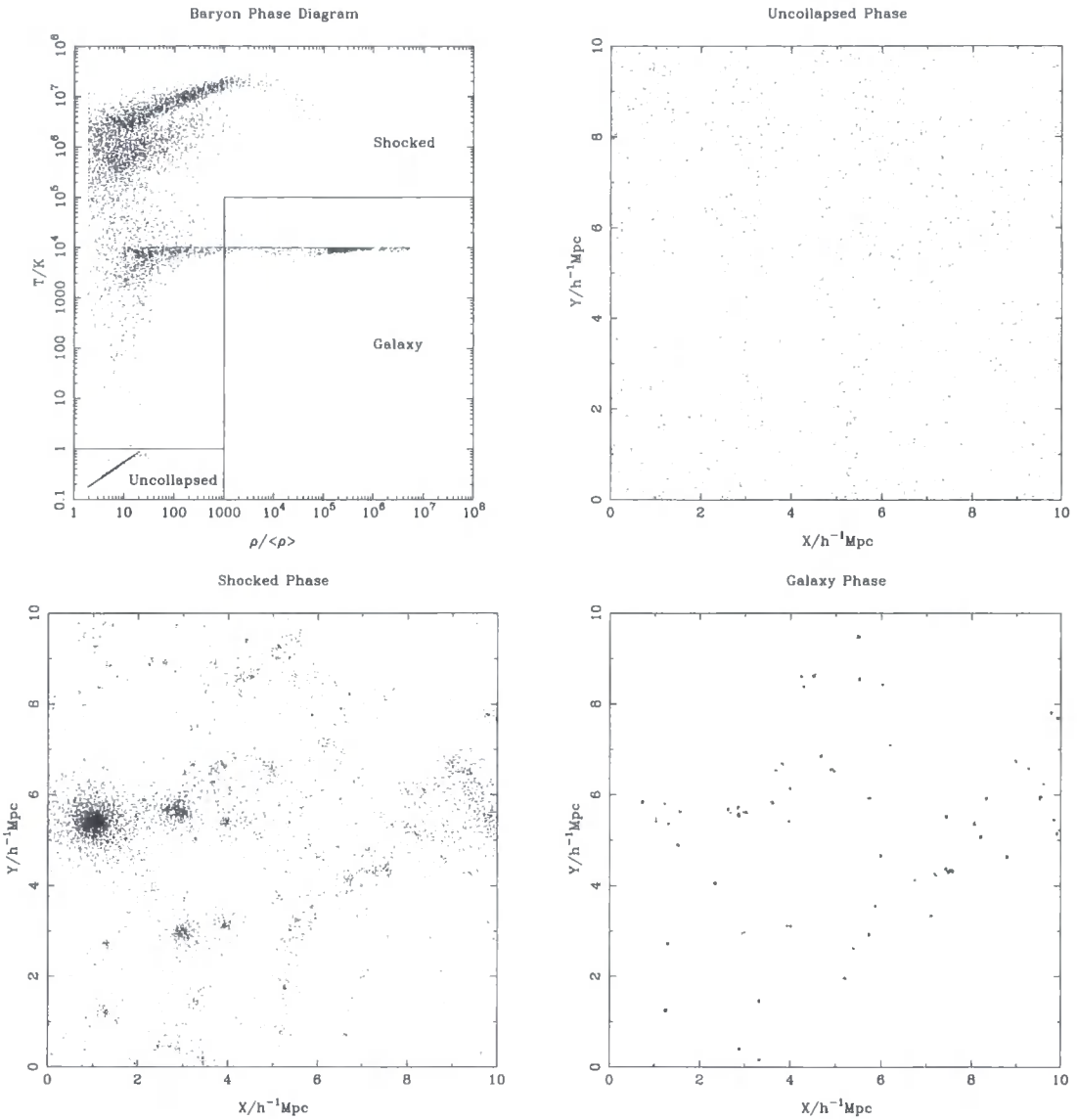


Figure 4.2: (top–left) The temperature–density phase diagram of all baryons in the fiducial simulation at the final output time ($z = 0$). The density is displayed in units of the mean gas density and the temperature in Kelvin. Positions of the gas particles obeying the various selection criteria defined in the text are shown in the other three plots, specifically the uncollapsed, shocked and galaxy phases. It is evident that each phase shows a considerable difference in structure: the uncollapsed gas exhibits a large degree of homogeneity, the shocked phase particles show large fluctuations in density and the galaxy phase is exclusively made up of tight clumps of particles.

4.3.2.2 The shocked phase

The shocked gas phase contains particles with $\rho < 10^3 \langle\rho\rangle$ and $T > 1\mathrm{K}$ as well as all particles with $T > 10^5\mathrm{K}$. Gas entered this phase due to the collapse and violent relaxation of gravitationally unstable regions, in which the bulk kinetic energy of the gas was transferred, via shocking, into

heat. As a result, the spatial distribution of this phase exhibits large fluctuations in density: halo atmospheres lie in this regime and the shocked gas is almost completely absent from the voids.

4.3.2.3 The galaxy phase

The particles in the galaxy phase lie in the regime, $\rho \geq 10^3 \langle\rho\rangle$, and $T \leq 10^5\mathrm{K}$. The presence of a strong horizontal feature at $10^4\mathrm{K}$ is due to the fact that the cooling rate is taken to be zero at this temperature. Gas Particles with higher temperatures must lose a significant fraction of their thermal energy via radiative cooling before reaching the main locus, The plot of the positions of the galaxy phase particles highlights the fact that cold, dense gas forms very tight clumps, which have spatial extents of the order of the gravitational softening. To select the "galaxies" themselves, only particles that lie in this region of the phase diagram are used.

4.3.3 Evolution of the baryon phase diagram

To illustrate the global evolution that led to the baryon phase diagram of Figure 4.2, the mass fraction in each phase is plotted in Figure 4.3, as a function of the expansion factor (normalised to the present value). All phase boundaries are fixed, with the exception of the uncollapsed phase – the temperature threshold is modified to $T < (1+z)^2\mathrm{K}$, which accounts for the adiabatic expansion of the uncollapsed gas. Initially, all the gas is in this phase, with a temperature of $100\mathrm{K}$. As the simulation progresses, some of the gas transfers to the shocked phase, as the first virialised haloes become resolved ($z \sim 10$). Gas does not start to enter the galaxy phase (forming the first objects) until later ($z \sim 3$). Both the shocked and galaxy phases continue to accrete more and more mass as the uncollapsed phase becomes progressively depleted. By the present day, ~ 18 per cent of the gas is still in the uncollapsed phase whilst the shocked gas phase contains about 64 per cent of the gas mass and the remaining 18 per cent has cooled into the galaxy phase.

Another important question we have addressed is to determine the precise trajectories in the $\rho - T$ plane, of the particles that finally end up in the galaxy phase. Figure 4.4 illustrates the two main paths that the galaxy particles take, for the whole duration of the fiducial simulation. The marked points indicate the redshifts $z = (10, 5, 3, 2, 1, 0.5, 0.1)$ for each trajectory.

The first example (solid line) shows the expected evolution of a galaxy particle – it becomes initially shock heated to over $10^6\mathrm{K}$ before cooling back down on to the $10^4\mathrm{K}$ boundary and, on average, progressively increasing in density until the final time. The second example (dotted line) takes a different path – again the particle is shock heated, but only as far as $10^4\mathrm{K}$. Both

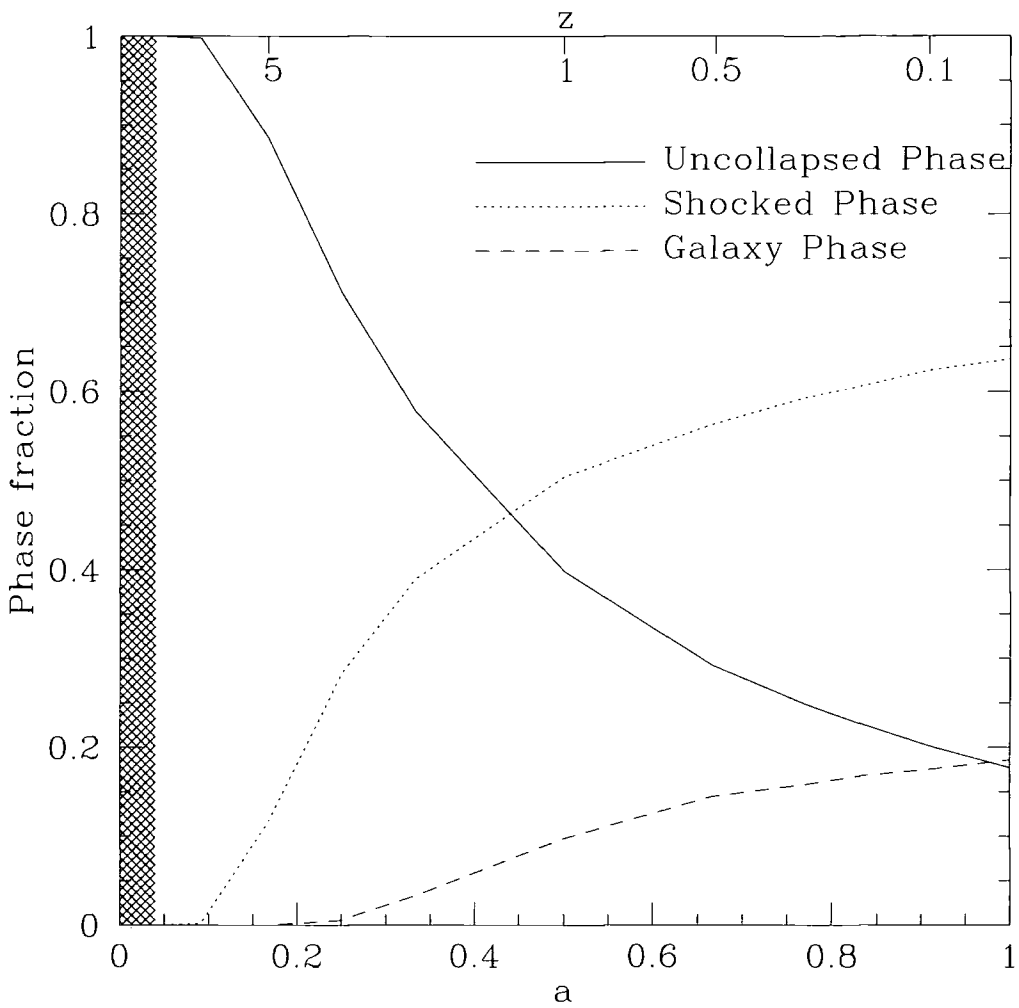


Figure 4.3: The evolution of the uncollapsed, shocked and galaxy phases, plotted as the fraction of gas particles in each phase (with respect to the total mass of gas particles), as a function of the expansion factor, $a = (1+z)^{-1}$. The shaded region marks the time before the start of the simulation ($a = 0.04$). Redshift is indicated along the top of the figure.

particles undergo a series of further heating events and, notably, the second one reaches its largest temperature at some later stage when a major galaxy merger takes place.

Quantitatively, it is found that only $\sim$ 11 per cent of the particles which end up in the galaxy phase were initially heated above 10^5K for a minimum of one timestep ($\sim$ 32 per cent initially reached temperatures above 12500K). The majority of the baryons that ends up in galaxies reach high density in the smallest haloes that are close to the resolution limit of the simulation. Consequently, their very short cooling times cause them to be heated above the 10^4K locus and subsequently cool back onto it within one timestep.

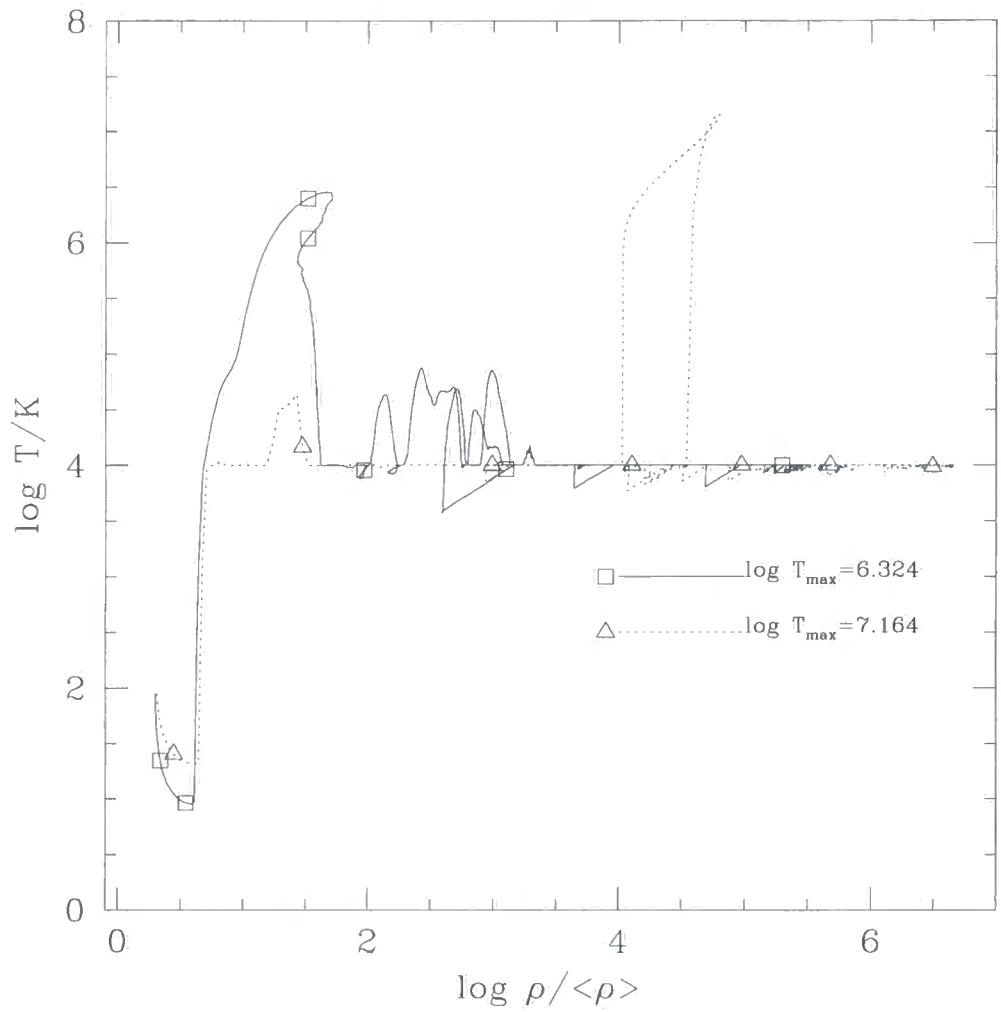


Figure 4.4: Two examples of particle trajectories in the $\rho - T$ plane, from the initial epoch ($z = 24$) to the present day. The legends depict their maximum temperatures. The solid path illustrates a particle that is heated to high temperature on initial heating (i.e. the first time it crosses the 10^4 K boundary), whereas the dotted path illustrates the evolution of a particle which is only initially heated as far as 10^4 K. Marked points indicate redshifts $z = (10, 5, 3, 2, 1, 0.5, 0.1)$.

4.3.4 Selection criteria

To select the galaxies, the subset of gas particles that lie in the defined galaxy phase (Figure 4.2) were selected. These particles were then run through a friends–of–friends group finder (Davis et al. 1985), which links particles together to form groups until no unlinked particles remain that are closer to a group than the linking length, $l = b\overline{n}^{-1/3}$, where $\overline{n}$ is the global mean number density. The value $b = 0.1$ ($l \sim 30\,h^{-1}$ kpc) was adopted, which gives a linking length of the order of the final S_2 softening length. Since the clumps are very tight and are well separated, the

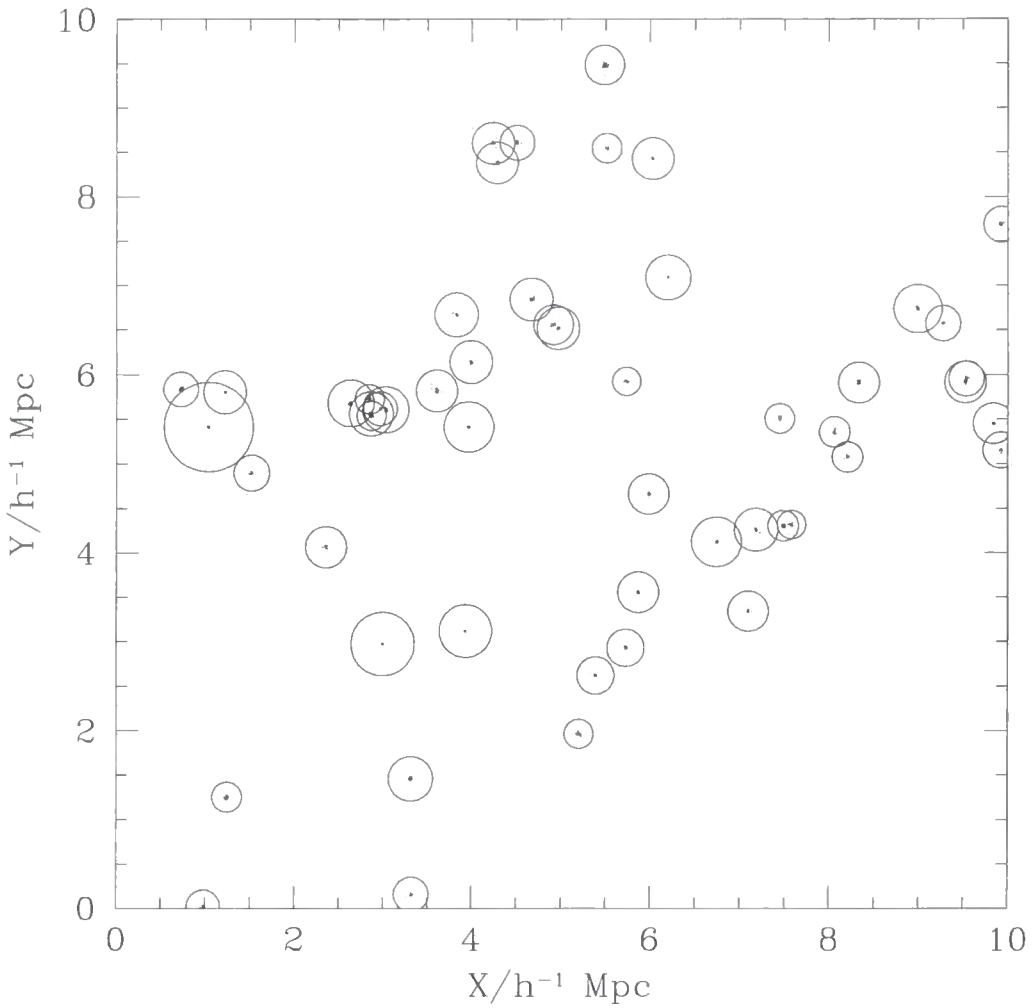


Figure 4.5: Projected positions of the fiducial set of 53 galaxies. The dots show the particles that are linked together by the group finder and the circles illustrate the mass of the galaxies, centred on the median position of each galaxy with radii proportional to the cube root of the galaxy mass.

output of the group finder is insensitive to variations in the choice of b: changes of the order of a few percent in mass were noted when varying b by a factor of two in each direction.

Once this catalogue was generated it was truncated by throwing away objects containing fewer than $N_{\rm SPH}$ particles. Objects below this mass are poorly defined because the SPH algorithm does not sample their properties well. For the fiducial simulation, 16 groups out of 69 were discarded in this manner, making up ~ 4 per cent of the mass in the galaxy phase. This constitutes most of the residual material in the galaxy phase that is not part of the final galaxy catalogue.

The effects of varying the defined limits of the galaxy phase were also studied. The total mass of gas in this phase is used for some of the comparisons, therefore it is important to ensure that

a region has been selected within which most of the particles are part of objects in the galaxy catalogue. The most important boundary is the lower limit of the gas density ($\rho/\langle\rho\rangle = 10^3$). Since the 10^4K feature is visible down to the lowest densities, there are particles that have gone through the cooling process that lie outside the galaxy phase. Lowering the density limit by an order of magnitude caused only a 3 per cent increase in the total number of particles grouped into galaxies. On the other hand, the fraction of the galaxy phase present in the final galaxy catalogue dropped from 96 per cent to 80 per cent. Reducing the density by the same amount again caused no further change in the former quantity but the completeness again decreased to 56 per cent, indicating that at these low densities there is contamination from the diffuse halo gas. The temperature boundary for the galaxy phase includes all relevant particles for $\rho > 10^3 \langle\rho\rangle$, since the hotter particles are still cooling down. As a check however, the upper limit on the temperature was lowered from 10^5K to 12000K. This produced a change of less than 1 per cent to the completeness and lowered the number of galaxies in the catalogue from 53 to 52.

The positions of the fiducial galaxies are illustrated in Figure 4.5. The dots are the projected positions of the particles linked together by the group finder. The circles illustrate the mass of the objects, centred on the median position of the linked particles in each group, with radii proportional to the cube root of the number of particles linked.

The dark matter haloes were identified by using a spherical overdensity algorithm (Lacey & Cole 1994) on the positions of the dark matter particles in the fiducial simulation endstate. The algorithm estimates the density field which is then used to find the centres of overdense regions. Successive spheres are placed around each centre with growing radii, until the mean internal density equals some value – the usual value of 178 times the mean density is assumed, as predicted by the spherical collapse model for the virial radii of haloes in an $\Omega = 1$ universe. Iterations are performed, with the centre redefined as the centre of mass of the object, until the difference becomes within some tolerance. Substructure within haloes was ignored for the purposes of this chapter. The halo finder produces a catalogue of all haloes with more than $N_h = 100$ particles in the fiducial case, corresponding to a mass of $\sim 8 \times 10^{11} h^{-1}\,\mathrm{M}_\odot$. Since the dark matter dominates the dynamical evolution, a lower limit in halo mass (rather than particle number) was applied when comparing different simulations. Hence, for simulations with poorer mass resolution, the lower limit on the number of particles was reduced in order to keep the halo mass threshold constant. This limit never went below 10 particles, approximately the lowest number that can be trusted (Efstathiou et al. 1988).

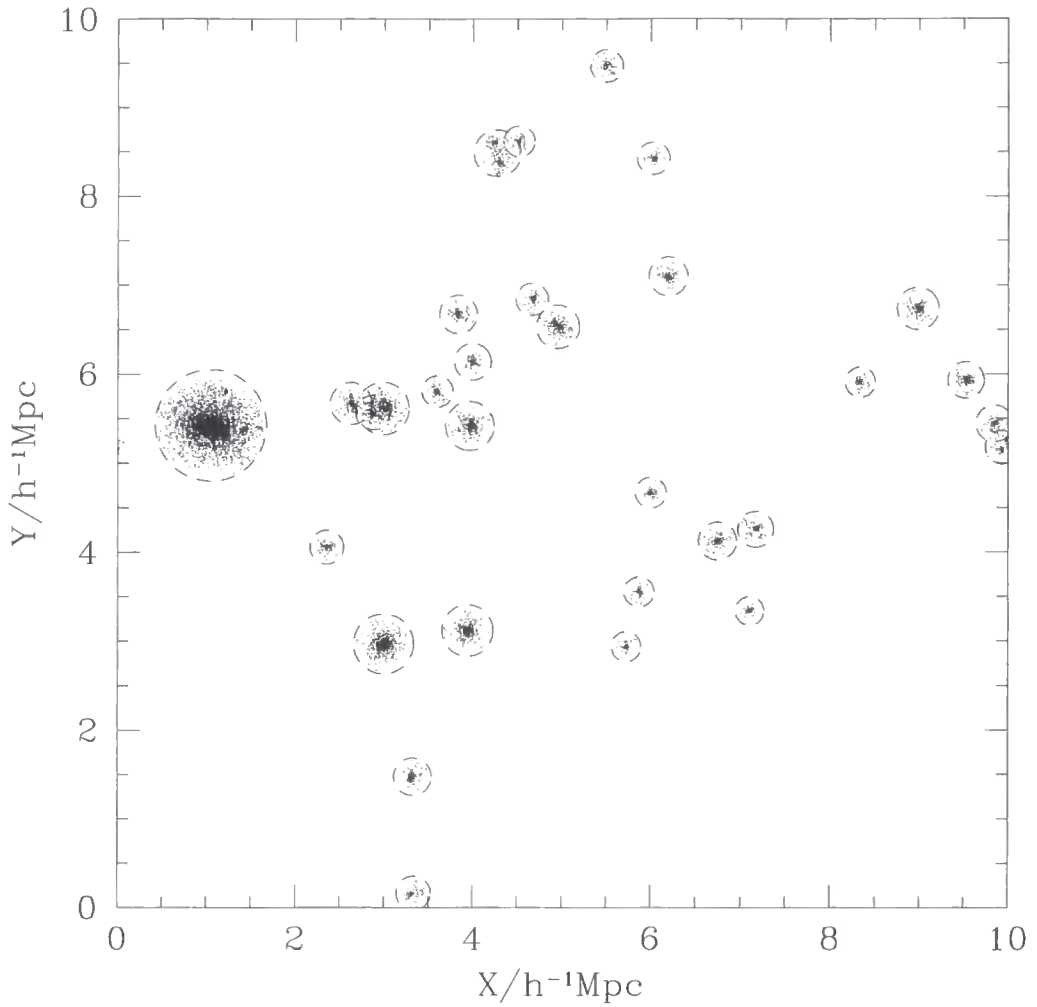


Figure 4.6: Projected positions of the dark matter particles in the fiducial simulation selected by the halo finder, Only haloes with more than 100 particles are shown. The circles are centred on each halo's centre of mass, and their radii are set equal to the distance from the centre to the most outlying particle of the halo. There are 31 haloes in total in the fiducial catalogue.

The output of the halo finder for the fiducial simulation is illustrated in figure 4.6. The circles are centred on the halo centre of mass and have radii equal to the virial radius of each halo. The dots represent the positions of all the dark matter particles within the virial radius of each halo. The largest halo has a total mass of $\sim 5 \times 10^{13} h^{-1}\ \mathrm{M}_{\odot}$, which corresponds to a virial temperature of $\sim 10^{7}\mathrm{K}$.

The $\rho - T$ distribution of the baryons located within the halo virial radii for the fiducial simulation are shown in Figure 4.7. Only gas within the shocked and galaxy phases are found in haloes as expected. The range of temperatures for the majority of the gas hotter than $10^{4}\mathrm{K}$ agrees well with

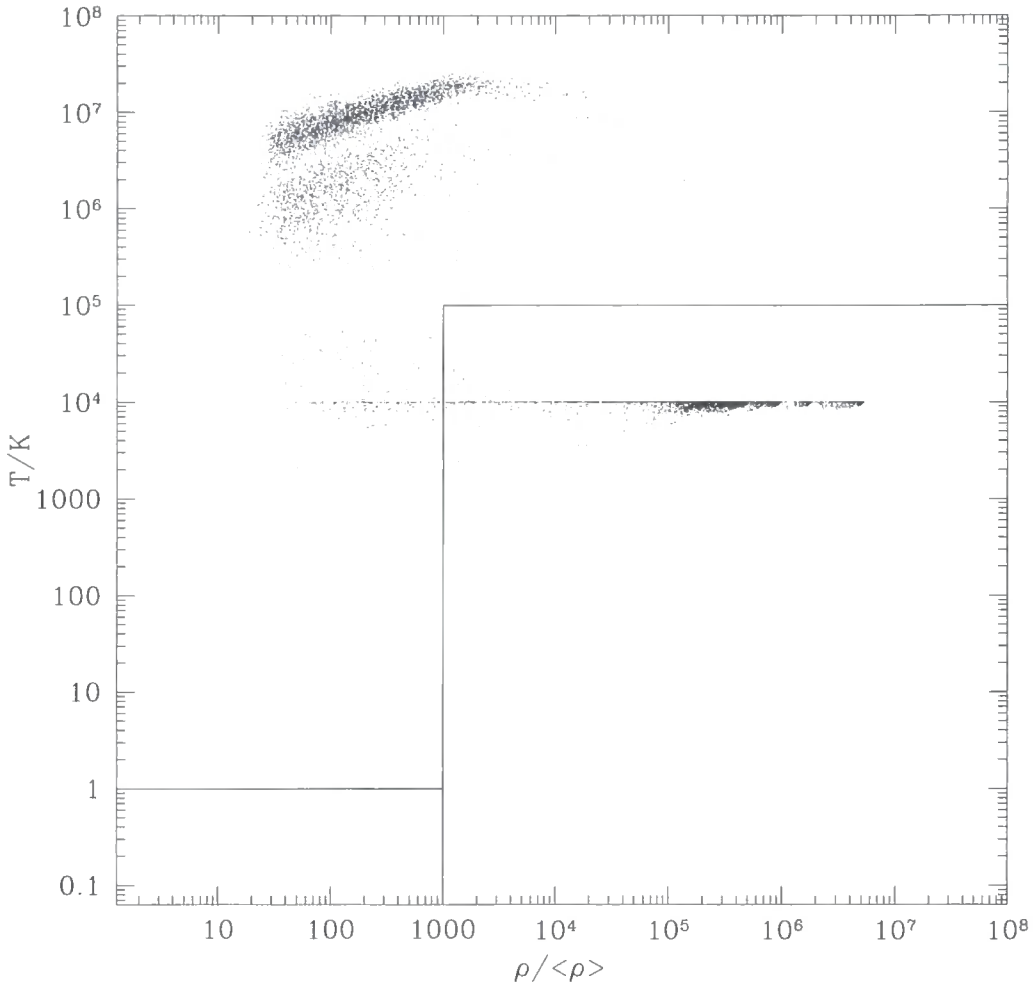


Figure 4.7: Phase diagram of all the gas particles within the selected halo virial radii in the fiducial simulation. Overlaid are the boundaries which we have defined in order to split the gas into separate phases.

the corresponding range of virial temperatures of the haloes, for the mass range of the simulation (from $\sim 5 \times 10^5$ to $\sim 10^7$K).

4.4 Simulation comparisons

4.4.1 Global features

The complete set of simulations that have been analysed are listed in Table 4.2, split into three sections. The first section includes simulations in which features other than simple parameter values were varied. The second section contains simulations in which numerical parameters (i.e. those that affect the algorithm) were varied. Finally, the third section contains simulations in which

the physical parameters were altered. For most of the parameters, there is a set of three simulations: the fiducial case and two additional simulations with the appropriate parameter varied about the fiducial value. The columns indicate the comparison number (14 in total); the alteration made to the fiducial parameter set; the fraction of particles identified in the uncollapsed, shocked and galaxy phases respectively; the number of galaxies in the final catalogue; the number of haloes in the final catalogue; and the *completeness*, defined as the fraction of particles in the galaxy phase that appear in the final object set. For discussion, the sections are labelled *general*, *numerical* and *physical* respectively.

4.4.1.1 General comparisons

The first comparison was designed to probe effects that might be present due to runs being performed on different computer architectures. The fiducial simulation was run on a Sun Sparcstation, whereas most of the comparison runs were performed on a DEC Alpha. Hence, an identical simulation to the fiducial case was also performed, using the latter platform. Differences can arise from system–specific variations of the code due to the compiler.

Comparison 2 was performed in order to investigate the numerical effects introduced by altering the initial positions of the particles. Two simulations were set up to test this. The first tested the periodicity of the algorithm, by translating all the particles by a fixed amount (in this case 1 FFT cell) before starting the simulation, then moving them back at the final time. The second was performed in order to examine how small inaccuracies in the initial positions can affect the results, by adding a random perturbation (tiny compared to the initial displacements) to each particle position at the initial time. In both this and the architecture comparison, no significant differences are observed in the phase fractions when compared to the fiducial simulation. These first two comparisons set a baseline for the size of variations that can be introduced solely by noise. Simulations that only differ by this amount should be viewed as being indistinguishable from the fiducial case.

For comparison 3, a run was added that uses a cooling function in tabulated form, as given by Sutherland & Dopita (1993), rather than a series of power–law fits. The two cooling functions used are shown in Figure 4.1 for the fiducial metallicity, $Z/Z_{\odot} = 0.5$. The power–law cooling function fits the tabulated function relatively well for the range of relevant temperatures ($10^4 - 10^8$K), although the latter shows a significant enhancement in $\Lambda(T)$ around 10^5K, (as much as a factor of 3) due to helium line cooling. However, as Table 4.2 shows, there is only a slight enhancement in the amount of gas in the galaxy phase.

Table 4.2.: Details of the set of simulations analysed in this chapter, following the various definitions and selection procedures discussed in Section 4.3. The first column is the comparison number given to the simulation, with the second column indicating the alteration made to the fiducial initial conditions. The next three columns give the fraction of baryons in each of the three phases defined in the text. Columns six and seven give the number of galaxies and haloes respectively that remain in the catalogue after the various selection procedures have been applied. The final column is the *completeness*, i.e. the fraction of baryons in the galaxy phase that are present in the final galaxy catalogues. Bold type has been used to denote values that deviate significantly from the fiducial case (two per cent for the phase fractions, five percent for the galaxies, haloes and completeness).

No.	Alteration	Uncollapsed Phase	Shocked Phase	Galaxy Phase	Galaxies	Haloes	Completeness
0	Fiducial Simulation	0.18	0.64	0.18	53	31	0.96
1	Architecture	0.16	0.65	0.19	51	32	0.96
2	Position Offset (1 cell)	0.17	0.65	0.18	50	31	0.96
2	Position Offset (random)	0.18	0.64	0.18	55	30	0.97
3	New Cooling Function	0.18	0.62	0.20	55	31	0.98
4	Cold gas decoupled	0.18	**0.67**	**0.15**	51	31	0.95
5	$s_{\max} = 0.6$	0.17	0.64	0.19	53	30	0.96
6	$\kappa = 0.5$	0.18	0.63	0.19	52	31	0.97
6	$\kappa = 2.0$	**0.06**	**0.78**	0.18	53	31	0.95
7	$\epsilon_0 = 5\,h^{-1}$ kpc	0.16	**0.69**	**0.15**	**45**	31	0.95
7	$\epsilon_0 = 20\,h^{-1}$ kpc	0.18	0.63	0.19	53	30	0.97
8	$N_{\rm SPH} = 16$	0.18	**0.56**	**0.26**	**97**	31	0.98

continued from previous page

No.	Alteration	Uncollapsed Phase	Shocked Phase	Galaxy Phase	Galaxies	Haloes	Completeness
8	$N_{\rm SPH} = 64$	0.17	**0.70**	**0.13**	**28**	31	**0.89**
9	$N = 2 \times 16^3$	0.20	**0.80**	**0**	**0**	**29**	0
9	$N = 2 \times 24^3$	0.17	**0.74**	**0.09**	**14**	30	**0.71**
9	$N = 2 \times 48^3$	0.17	**0.57**	**0.26**	**141**	**34**	0.98
10	$1 + z_{\rm i} = 10$	0.17	0.65	0.18	55	32	0.97
10	$1 + z_{\rm i} = 50$	0.19	0.63	0.18	51	**29**	0.96
11	$\Omega_{\rm b} = 0.03$	0.18	**0.69**	**0.13**	**49**	31	0.94
11	$\Omega_{\rm b} = 0.12$	0.18	**0.58**	**0.24**	53	31	0.97
12	$Z = 0.0 Z_{\odot}$	0.18	**0.78**	**0.04**	**9**	30	**0.44**
12	$Z = 1.0 Z_{\odot}$	0.18	**0.60**	**0.22**	54	31	0.97
13	$T_{\rm i} = 0{\rm K}$	0.18	0.64	0.18	51	31	0.95
13	$T_{\rm i} = 10^7{\rm K}$	**0.22**	**0.61**	0.17	50	30	0.96

The final comparison that was made in this section was between the fiducial simulation and the version of HYDRA used by Pearce et al. (1999). The latter contains a change in the SPH algorithm to reduce cooling in large haloes (see also Thacker et al. 1998). It is based on a *multiphase* approach, in which the cold, dense particles are decoupled from the hot medium: gas above 10^5K does not see gas below 12000K for the purposes of calculating the gas density. For all other gas particles, the SPH forces are calculated in the standard way. This procedure ensures that infalling cold diffuse particles still feel the halo's outer accretion shock and allows hot haloes to cool at a rate determined by only the hot gas and galaxies to dissipate energy, merge and feel drag as they move around the halo environment. The effects of this alteration on the phase distribution are evident in the table: the fraction of gas particles in the galaxy phase is reduced by 3 per cent while the fraction in the hot phase increases to compensate.

4.4.1.2 Varying numerical parameters

The first numerical parameter varied (comparison 5) was the initial (maximum) softening, $s_{\max}$. Increasing $s_{\max}$ moves the changeover from a comoving to a physical softening to higher redshift. The value of the maximum softening was doubled to $s_{\max} = 0.6$, so that for the same final softening length, the changeover now occurred at $z \sim 3$ rather than at $z \sim 1$. More importantly, the softening now has a larger comoving value (equivalent to a Plummer softening of $\sim 40\ h^{-1}$ kpc) before the changeover. The galaxies that form before this epoch are more loosely bound (increasing their effective volume), and are therefore more susceptible to the effects of ram–pressure stripping, tidal disruptions or accretion as they move through the hot halo gas. No significant changes in the final phase fractions are present however.

The timestep normalisation value, κ (comparison 6) was varied by a factor of 2 about the fiducial value of 1. This effectively doubles or halves the length of each timestep, causing the simulation to take roughly twice or half as many timesteps to run to $z = 0$. Doubling the fiducial value introduces inaccuracies in both the positions and velocities of the particles. The phase fractions show that there is significantly more shocked gas in the $\kappa = 2$ simulation than in the fiducial run, at the expense of the uncollapsed gas. The galaxy phase fraction and the number of galaxies in the final catalogue do not change. The increased efficiency at which the gas is heated can be explained by the fact that the larger timesteps in the $\kappa = 2$ run increased the error in particle positions and velocities, which leads to larger values in the velocity divergence, increasing the efficiency of gas heating from viscous interactions.

As was hoped, the value $\kappa = 0.5$ had no significant effect on the endstate properties of the galaxies and haloes, vindicating CTP95's choice of $\kappa = 1$.

Comparison 7 consisted of halving and doubling the value of the final softening length about the fiducial value, $\epsilon_0 = 10\ h^{-1}$ kpc. Changing the value of ϵ_0 only has an effect once the physical softening length has dropped below the maximum defined by $s_{\max}$. For $\epsilon_0 = 5\ h^{-1}$ kpc this occurs at $z = 3$ whilst for $\epsilon_0 = 20\ h^{-1}$ kpc the softening is always comoving. Simulations with smaller softening lengths have more tightly bound galaxies which may be less susceptible to disruption. Doubling the fiducial value of ϵ_0 has no significant effect on the final number of galaxies and the fraction of gas in each phase. However halving the final softening length systematically decreases the number of galaxies by $\sim$ 15 per cent, with the residual material still in the shocked phase. It is possible that this is due to the effects of two–body relaxation artificially heating gas particles (Steinmetz & White 1997). For a dark matter halo with 50 particles forming at $z = 0$, the two–body relaxation times are approximately 10.4, 7.5 & 5.9 Gyr for $\epsilon_0 = 20, 10\,\&\,5\,h^{-1}$ kpc respectively. For $\epsilon_0 = 5\ h^{-1}$ kpc the relaxation time is less than half the age of the universe.

The value of $N_{\rm SPH}$ was varied in comparison 8. Since this parameter sets the resolution at which the gas density field is evaluated, large overdensities are progressively smoothed out as the value of $N_{\rm SPH}$ increases. It then becomes increasingly difficult to form galaxies, since the emissivity is proportional to the square of the gas density. On the other hand, using a smaller value of $N_{\rm SPH}$ increases the noise arising from the discrete nature of the mass distribution. Thus a compromise in the value is sought. It is common to assume $N_{\rm SPH} = 32$, which is the fiducial value, and this comparison looks at the effects introduced by changing this number by a factor of two. The number of galaxies in each simulation clearly shows the expected trend, with more galaxy phase gas and more galaxies if $N_{\rm SPH} = 16$ than if $N_{\rm SPH} = 64$. This material comes from the shocked phase as the uncollapsed phase fraction is more or less constant in all cases. The significant drop in galaxy number and in completeness for the $N_{\rm SPH} = 64$ simulation is due to 11 groups being discarded, with $N_g < N_{\rm SPH}$.

Comparison 9 consisted of changing the mass resolution by varying the total number of particles. The initial power spectrum was truncated appropriately to account for the varying Nyquist frequency of the particle mesh. The four runs have $N = 2\times(16, 24, 32, 48)^3$ particles respectively, where the 2 indicates the number of species. The masses of the two particle species in each of these runs are given in Table 4.3. The three runs with the largest N were selected to quantify the result of increasing the mass resolution of the simulation. A run with $N = 2 \times 16^3$ was also added to probe the regime where artificial two–body heating of the gas particles by the dark matter particles should significantly affect the cooling rate of the gas. Figure 4.8 shows the critical dark matter particle mass (defined as the mass for which the cooling time of the gas equals the two–

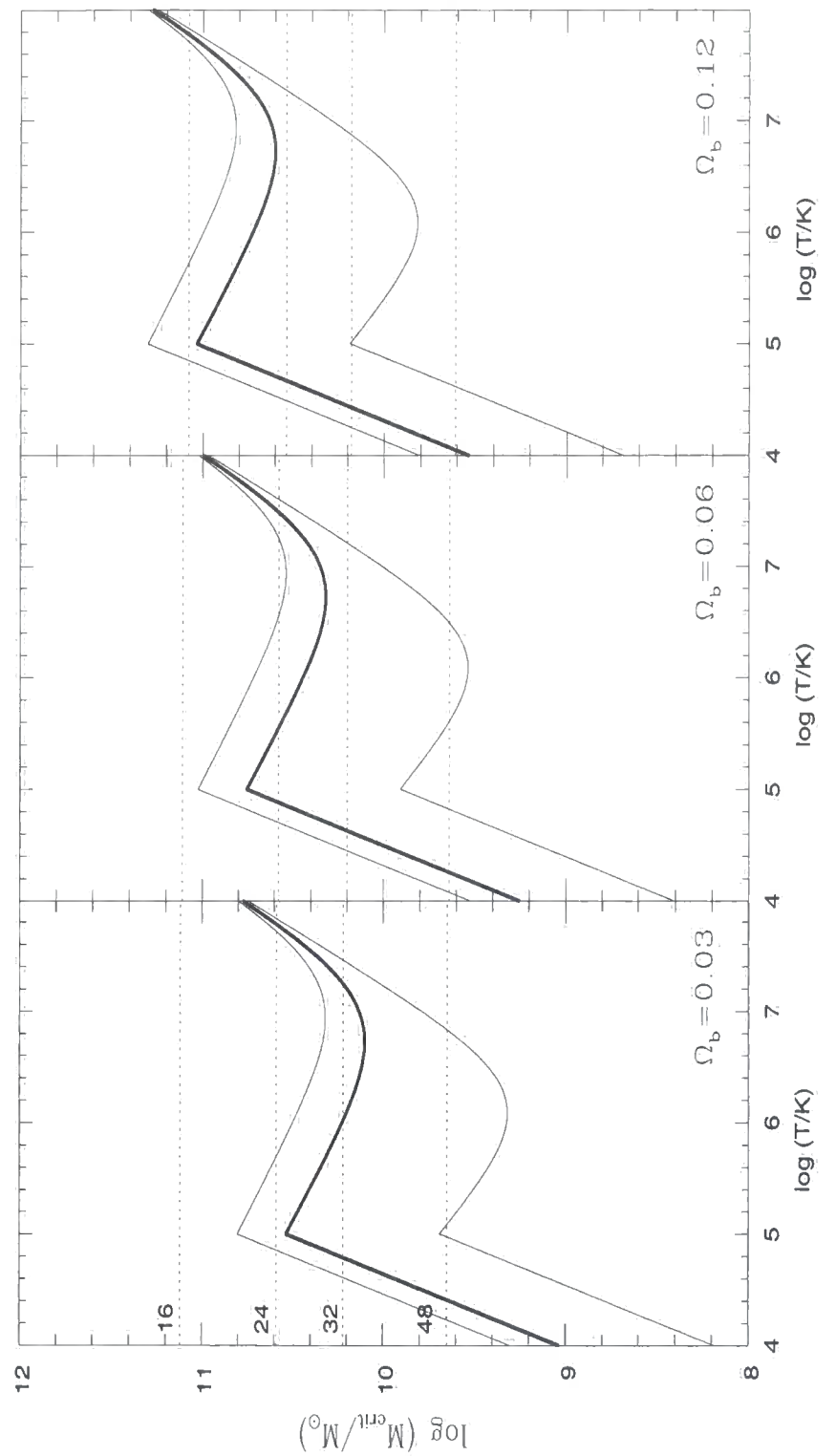


Figure 4.8: The critical dark matter particle mass for spurious two–body heating as a function of temperature (Steinmetz & White 1997). The three solid curves in each plot correspond to different gas metallicities (from top to bottom, $Z/Z_{\odot} = 1.0, 0.5, 0.0$, with the fiducial metallicity, 0.5, in bold). The different panels assume a local gas fraction, f, corresponding to the values of the global baryon fraction analysed in this chapter, $\Omega_b = 0.03, 0.06, 0.12$. The dotted horizontal lines give the masses of the dark matter particles for $N = 2 \times X^3$, where X is the label on each line. A value of 5 was assumed for the Coulomb logarithm.

Table 4.3. The masses of each particle species for the runs performed with varying values of N.

N	$M_{\rm dark}(10^{10}h^{-1}\,{\rm M}_\odot)$	$M_{\rm gas}(10^{10}h^{-1}\,{\rm M}_\odot)$
2×16^3	6.38	0.41
2×24^3	1.89	0.12
2×32^3	0.80	0.05
2×48^3	0.24	0.015

body heating time) as a function of the gas temperature, for the metallicities and baryon fractions studied in this chapter (Steinmetz & White 1997). The gas fraction is assumed to be equal to the appropriate value of $\Omega_{\rm b}$, although the values can be scaled, if desired. The Coulomb logarithm is assumed to be 5, which is appropriate for the largest mass haloes in the simulations studied in this chapter (and therefore demonstrating the worst–case scenario for this effect). Overlaid as dotted horizontal lines are the mass of the dark matter particles for the simulations in this comparison. For $\Omega_b = 0.06$ and $Z = 0.5Z_\odot$, both the fiducial simulation and the $N = 2 \times 48^3$ simulation are acceptable for the range of temperatures relevant to this work, and the $N = 2 \times 24^3$ simulation is borderline. Since the $N = 2 \times 16^3$ run lies well above the critical line, the amount of gas cooled in this simulation should be significantly affected by two–body heating. The effect is indeed severe: only 24 particles make it into the galaxy phase and hence no galaxies form at all. (Note that the cooling time at the resolution limit of this simulation is $\sim 10^8$ years – a small fraction of the age of the universe – therefore in the absence of two–body heating effects, a population of cooled objects should have formed.)

For the remaining three simulations, the amount of uncollapsed gas remains roughly constant, and the biggest change is in the increasing amount of gas that cools from the shocked phase into the galaxy phase with larger N. This is reflected both in the galaxy fractions and in the number of galaxies present in the final catalogues. This effect is expected when varying the mass resolution of the simulations. When N is increased, the sampling of the density field improves and hence smaller mass haloes are resolved. The smallest haloes in the simulation will, on average, form first, and therefore have higher physical densities than the haloes that form later. As discussed in Section 4.3.3, most of the galaxy mass is accumulated in these objects.

Finally, the effect of changing the initial redshift of the simulation has been investigated by considering runs with $1 + z_i = 10, 25$ and 50 respectively (comparison 10). The main potential pitfall when choosing z_i is that choosing too small a value will suppress the amount of small–

scale power, causing objects on small scales to form too late. Hence the initial redshift should be high enough that scales of the order of the initial softening should not have already become non–linear (i.e. $\delta \ll 1$). For the $1 + z_i = 50$ run, the fiducial boundaries of the uncollapsed phase are sufficient to segregate the uncollapsed gas from the rest, while for the $1 + zi = 10$ run, the temperature boundary has been rescaled by a factor of 6.25, i.e. $(25/10)^2$. Evidently, there are no significant changes in any of the gas phases or in the number of galaxies, when comparing all three runs.

4.4.1.3 Physical comparisons

The mass fraction of baryons was varied in comparison 11, forming the set of values $\Omega_\mathrm{b} = (0.03, 0.06, 0.12)$. Varying this parameter alters the masses of both the gas and dark matter particles, and therefore the number density of atomic particles, n. The latter affects the cooling rate of the gas since the emissivity scales as n^2. Again (see Figure 4.8), it is important that the baryon fraction should be consistent with the critical mass of the dark matter particles, given both the metallicity and the value of N. Clearly for the choice of other fiducial parameters, the value $\Omega_\mathrm{b} = 0.03$ ($Z = 0.5$, $N = 2 \times 32^3$) is borderline and this is reflected in a reduction in the number of galaxies formed. Increasing the baryon fraction to $\Omega_\mathrm{b} = 0.12$ has little effect upon the number of galaxies but they are systematically heavier.

Comparison 12 was performed in order to examine the effect of changing the metallicity of the gas. Each particle is assigned a constant amount of *metals* (i.e. of nuclei that have an atomic number greater than 2). The metallicity affects the shape of the cooling function, as illustrated in Figure 4.1. Also affected is the critical mass of the dark matter particles (which depends on the cooling function) as shown in Figure 4.8. The simulation with $Z = 0$ is clearly on the wrong side of the line and the galaxy phase fractions are consequently affected: slightly more material cools with the higher metallicity, but the run with no metals shows a catastrophic drop in the amount of material in the galaxy phase, the number of galaxies and the completeness.

Finally, in comparison 13, a set of runs were performed with different initial gas temperatures, in order to measure how the dynamical and thermal evolution of the gas depends on its initial thermal state. Two additional runs have been included, one with $T_i = 0\mathrm{K}$ and the other with $T_i = 10^7\mathrm{K}$. The fiducial boundaries of the uncollapsed phase are sufficient for the run with $T_i = 0\mathrm{K}$. However the boundaries of the uncollapsed phase are redrawn for the run with $T_i = 10^7\mathrm{K}$, such that $s < 2.5$ and $\rho / \langle\rho\rangle < 50$, where $s = \log_{10}(T/(\rho\,/\,\langle\rho\rangle)^{2/3})$ measures the specific entropy of the gas. (Using these constraints makes it easier to pick out the uncollapsed gas since it is isentropic.) For the run with $T_i = 0\mathrm{K}$, there are no significant differences in the measured quantities. However,

the run with $T_i = 10^7$K shows a small increase in the uncollapsed gas phase at the expense of the shocked gas. For this run, the gas cools as expected from adiabatic expansion until $z \sim 5$. However by $z = 3$ the uncollapsed phase is at temperatures below 10^4K, over an order of magnitude below the predicted temperature from adiabatic expansion. At these redshifts, the density and the cooling rate of the gas is high enough to allow the gas to radiate a significant fraction of its energy away. Consequently, only a small fraction of the gas is too hot to withstand gravitational collapse.

4.4.2 The spatial distribution of galaxies

The distribution of the galaxies at $z = 0$ for all of the simulations are illustrated in Figures 4.9 and 4.10. The positions of the circles represent the projected positions of the galaxies, with their radii set proportional to the cube root of their masses. The number of galaxies in each catalogue are also given.

Evident from the figures is that qualitatively similar clustering patterns are observed for all simulations with appreciable numbers of galaxies. This is not surprising since the sites of galaxy formation are determined by the structure of the underlying dark matter, whose average properties do not change from simulation to simulation. A more quantitative measurement of the clustering, for example using the two–point correlation function, is too noisy to gain anything useful (the number of objects in each volume is too small).

4.4.3 The mass distribution of galaxies

Figure 4.11 gives plots of the cumulative mass functions of galaxies (i.e. the number density of galaxies greater than a specified baryon mass, M_g), with the fiducial simulation always plotted as the solid line. The dynamic range for these simulations is about 2 orders of magnitude in abundance and about 1.5 orders of magnitude in mass. The functions all have a characteristic shape: a shallow slope at low masses, turning over to a steeper slope at higher masses and a large tail at the high mass end. The latter feature is simply due to the existence of a single massive galaxy that has a baryonic mass of $\sim 5 \times 10^{11} h^{-1}\,\mathrm{M}_\odot$. Galaxies this massive are extremely rare in the Universe; this object is discussed below.

The biggest differences in the mass functions occur in the physical comparisons, in which the gas cooling rate varies. The simulation with $Z = 0$ fails to produce as many galaxies as the fiducial simulation – the abundance is down by a factor of 5 and the galaxies that do form are much less massive. The large differences seen in the runs with varying baryon fraction are mainly due to the difference in masses of the individual gas particles, varying by a factor of two. However, when the

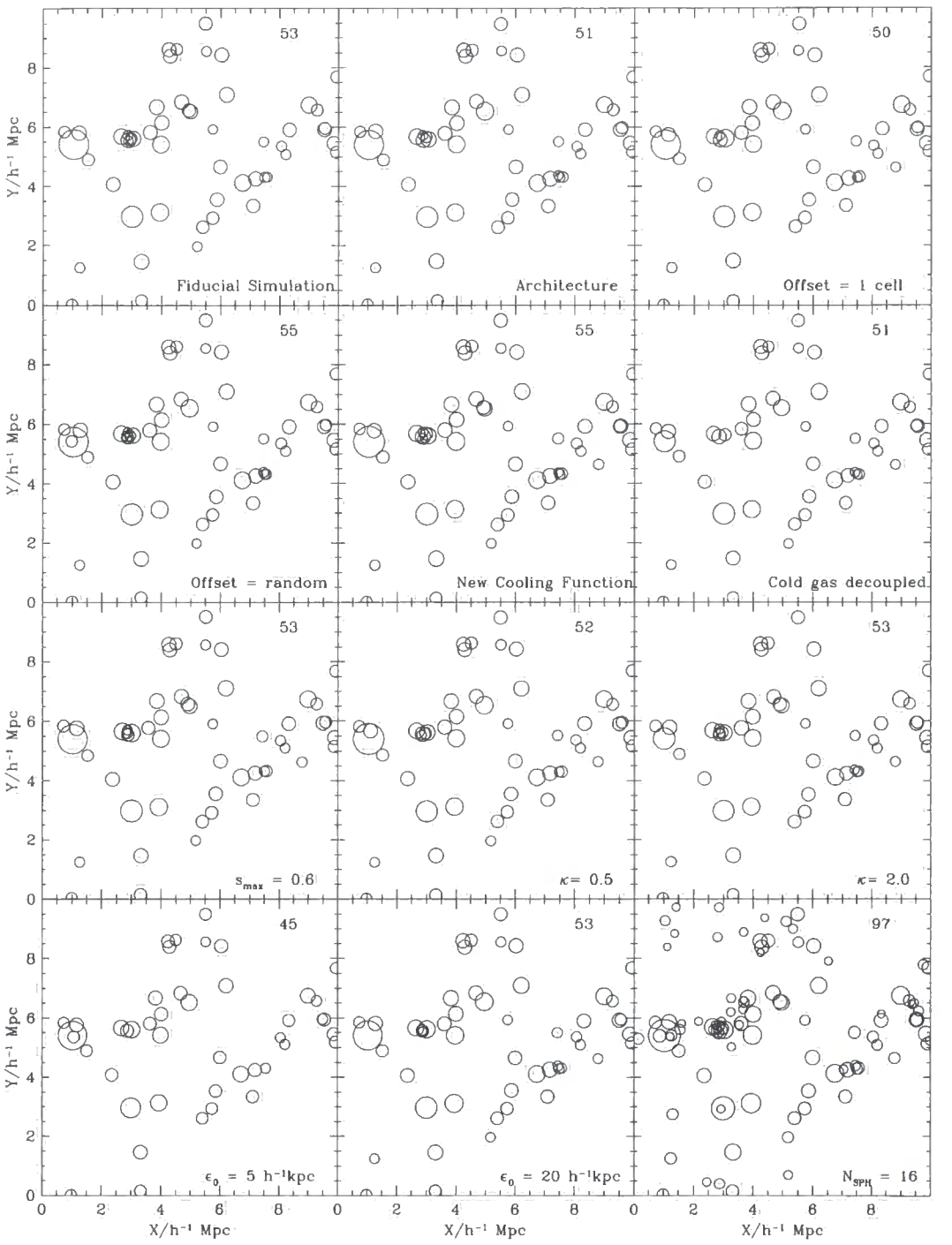


Figure 4.9: Projected positions of the galaxies in the final catalogues for all of the simulations analysed in this chapter. The radii of the circles are proportional to the cube root of the galaxy masses. The number of galaxies in the catalogues are printed at the top right of each panel.

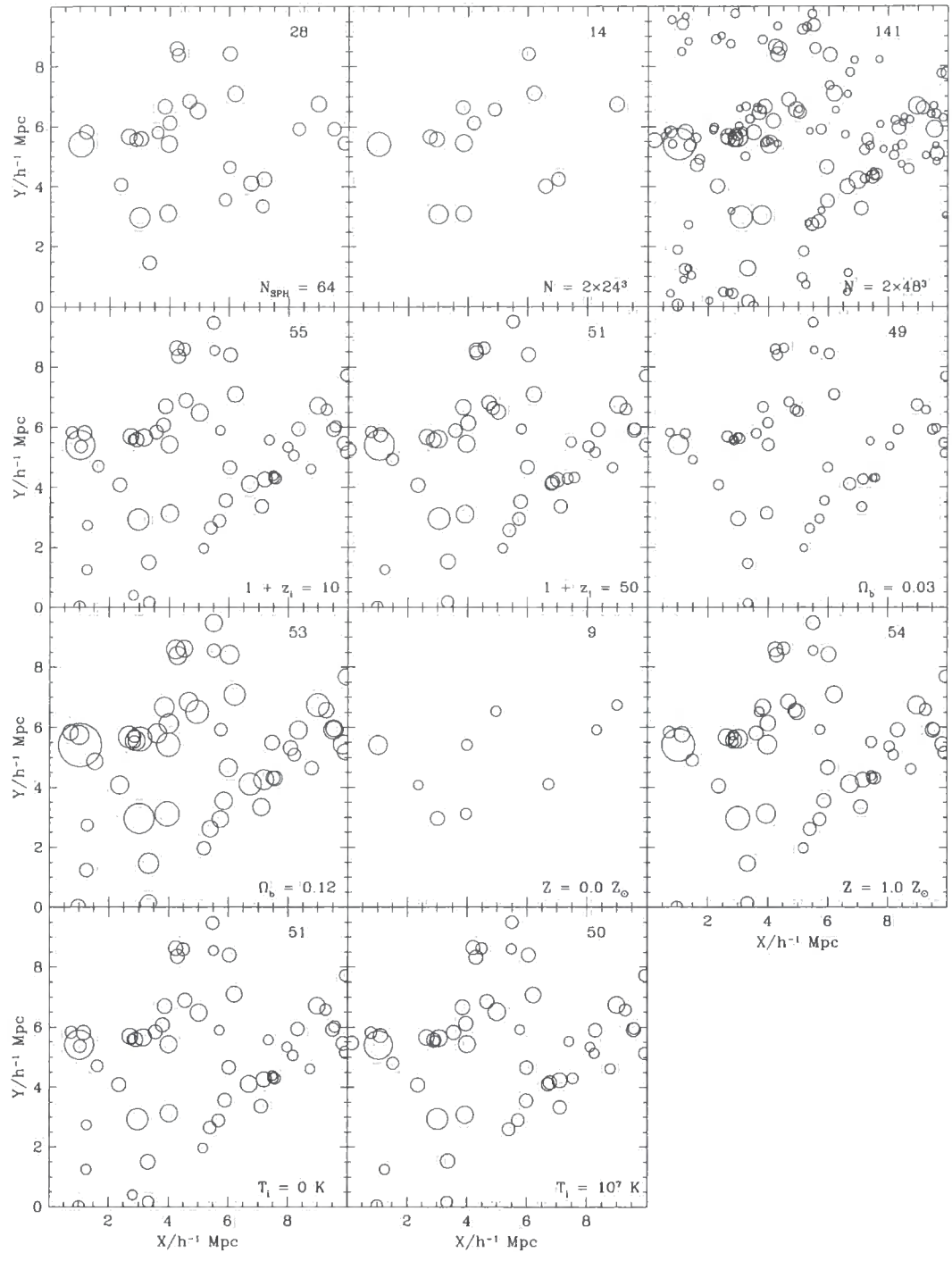


Figure 4.10: Projected positions of the galaxies (continued).

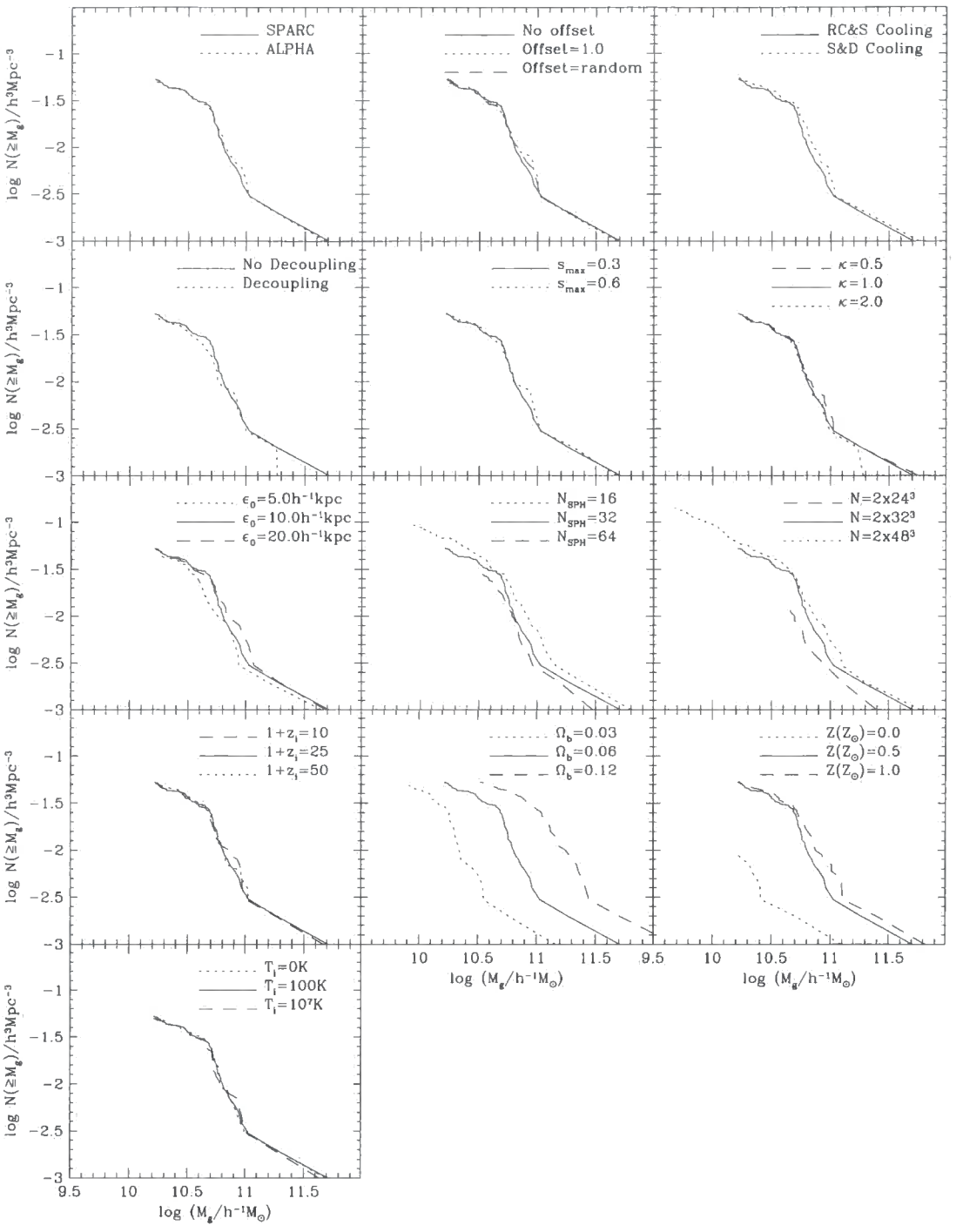


Figure 4.11: Cumulative galaxy mass functions for the full set of simulation comparisons. The plot gives the comoving number density of galaxies with baryon mass greater or equal to M_g. The fiducial simulation is always plotted with a solid line–type.

galaxy masses are scaled to take out this effect, there is still a systematic difference between each function. This is due to the effect the baryon fraction has on the cooling rate of the gas.

Varying the number of SPH neighbours shows a progression in the smallest resolved galaxy mass due to the selection criterion that the number of constituent particles $N_g > N_{\rm SPH}$. There are fewer galaxies for larger $N_{\rm SPH}$ and they are systematically less massive. This is a reflection of the fact that the algorithm has to smooth over a larger range, which consequently leads to a poorer resolution of the density field. Since emissivity scales as n^2, the lower cooling rate inhibits the galaxy formation process - an entirely numerical effect. Varying the value of N shows similar results for the same reason, i.e. a larger number of particles enables the SPH algorithm to resolve higher gas densities.

The remaining differences are relatively small. Changing the cooling function to the tabulated form of Sutherland & Dopita (1993) produces more cooling around 10^5K, boosting the cooling rate and systematically increasing the mass of the galaxies by ~ 10 per cent. However, there is no significant difference in the shape of the mass function.

Comparison 4 (in which the cold gas is decoupled from the hot gas) also shows a small difference in the mass functions. This difference is dominated by the largest object in the simulation volume, which is reduced in mass by a factor of 3 when the cold, dense gas is decoupled from the hot gas. Notably, this alteration to the algorithm has a negligible effect on the smaller objects.

The comparison set in which the final softening length is varied shows an offset for $\epsilon_0 = 5\,h^{-1}$ kpc. This simulation produces galaxies that are systematically lighter than the simulations with larger values of ϵ_0, as expected if they are being affected by two–body heating from the dark matter particles.

Table 4.4.: Statistical measures of the scatter present in all simulation comparisons of the matched galaxies and haloes. The first two columns detail the particular comparison and the third column lists the number of galaxies in each catalogue that were matched with the fiducial simulation, using the method detailed in the text. The rest of the columns list the median and semi-interquartile range (siqr) for the measured differences in galaxy separation (in h^{-1} kpc), galaxy mass, halo dark matter mass and halo baryon mass respectively.

No.	Alteration	Matches	Δr(galaxies)		$\Delta \log M$(galaxies)		$\Delta \log M$(dark matter)		$\Delta \log M$(baryons)	
			median	siqr	median	siqr	median	siqr	median	siqr
1	Architecture	50	17.	6.0	0.00	0.01	0.003	0.01	−0.002	0.01
2	Position Offset (1 cell)	49	21.	5.4	0.00	0.02	0.007	0.02	0.0	0.02
2	Position Offset (random)	52	15.	5.5	0.00	0.02	0.004	0.009	0.002	0.01
3	New Cooling Function	53	14.	5.8	0.04	0.01	0.0	0.01	0.01	0.02
4	Cold gas decoupled	49	19.	4.0	−0.03	0.03	−0.008	0.01	−0.02	0.03
5	$s_{\max} = 0.6$	51	24.	7.3	0.01	0.02	0.002	0.02	0.02	0.02
6	$\kappa = 0.5$	51	19.	7.1	0.00	0.01	−0.002	0.01	0.004	0.01
6	$\kappa = 2.0$	50	22.	7.4	0.00	0.03	0.004	0.01	−0.002	0.02
7	$\epsilon_0 = 5\,h^{-1}$ kpc	44	23.	8.7	−0.08	0.03	−0.007	0.02	−0.07	0.02
7	$\epsilon_0 = 20\,h^{-1}$ kpc	50	15.	4.9	0.01	0.02	0.0	0.01	0.01	0.01
8	$N_{\rm SPH} = 16$	53	18.	5.9	0.09	0.04	−0.003	0.02	−0.04	0.02
8	$N_{\rm SPH} = 64$	28	16.	4.3	−0.06	0.04	−0.003	0.02	−0.02	0.03
9	$N = 2 \times 24^3$	14	95.	62.	−0.13	0.06	−0.01	0.05	0.07	0.05

Continued from previous page

No.	Alteration	Matches	Δr(galaxies)		$\Delta \log M$(galaxies)		$\Delta \log M$(dark matter)		$\Delta \log M$(baryons)	
			median	siqr	median	siqr	median	siqr	median	siqr
9	$N = 2 \times 48^3$	51	93.	55.	0.02	0.06	0.005	0.03	0.01	0.05
10	$1 + z_\mathrm{i} = 10$	50	46.	16.	0.01	0.03	0.0	0.04	0.01	0.03
10	$1 + z_\mathrm{i} = 50$	48	39.	23.	0.00	0.03	−0.001	0.03	−0.004	0.03
11	$\Omega_\mathrm{b} = 0.03$	48	26.	7.5	−0.41	0.03	−0.006	0.01	0.34	0.03
11	$\Omega_\mathrm{b} = 0.12$	51	18.	4.9	0.37	0.04	−0.02	0.01	0.40	0.02
12	$Z = 0.0 Z_\odot$	9	19.	1.7	−0.56	0.04	0.004	0.01	0.10	0.04
12	$Z = 1.0 Z_\odot$	51	14.	4.7	0.05	0.03	0.0	0.02	0.03	0.01
13	$T_\mathrm{i} = 0\mathrm{K}$	51	18.	6.7	0.00	0.02	0.0	0.01	0.003	0.01
13	$T_\mathrm{i} = 10^7\mathrm{K}$	49	41.	18.	−0.02	0.02	−0.01	0.03	0.002	0.03

4.4.4 Comparison of matched galaxies and haloes

Individual properties of the galaxies and haloes have been compared on an object–to–object basis between each simulation and the fiducial run. Since the number of objects in each catalogue is different in nearly every case, there is not a straightforward one–to–one correspondence between the sets. To circumvent this, the smallest object set was always taken first and subsequently compared to the larger set. When more than one object in the larger set was matched with the same object in the smaller set, the closest pair was selected. Generally, this was not a problem since the discarded objects are usually at considerably larger distances than their matched object. However, this method will inevitably produce outliers in both separation and mass differences due to the fact that, on occasions, an object in one simulation is fragmented into smaller objects in another. Therefore the median and semi–interquartile ranges were used as statistical measures of the offset and scatter of any relative quantity measured. Positions of the galaxies were defined as the median position of the linked particles and, for the haloes, their centre of mass.

4.4.4.1 Galaxy–galaxy displacements

The first property examined was the scatter in the positions of the matched galaxies. The results are illustrated in Figures 4.12 and 4.13, which show the displacement between each matched pair (in h^{-1} kpc) plotted against the average object mass. The solid line illustrates the median separation for each comparison. The number of galaxies matched and the median plus semi–interquartile ranges of the displacements are presented in Table 4.4, columns 3, 4 & 5. It is clear from the comparison between the simulations run on different computer architectures that there is an intrinsic offset in matched positions at the level of a few softening lengths. Furthermore, a significant number of matches exceed this level by as much as a factor of four. As a consequence of this, only those comparisons that have median offsets in excess of the level seen for these two simulations are regarded as significantly different.

The first set to show a significant change is the simulations with different values of N. Varying the mass resolution causes differences in the number of objects formed and also affects the background dark matter distribution, so it is of no surprise that the displacements show a significantly larger amount of scatter in these comparisons. Varying the value of $N_{\rm SPH}$ also produces very different numbers of galaxies but this does not affect the median separation at all, demonstrating that it is the difference in the dark matter that primarily drives the increased scatter in the galaxy positions rather than mergers or interactions between the different numbers of galaxies.

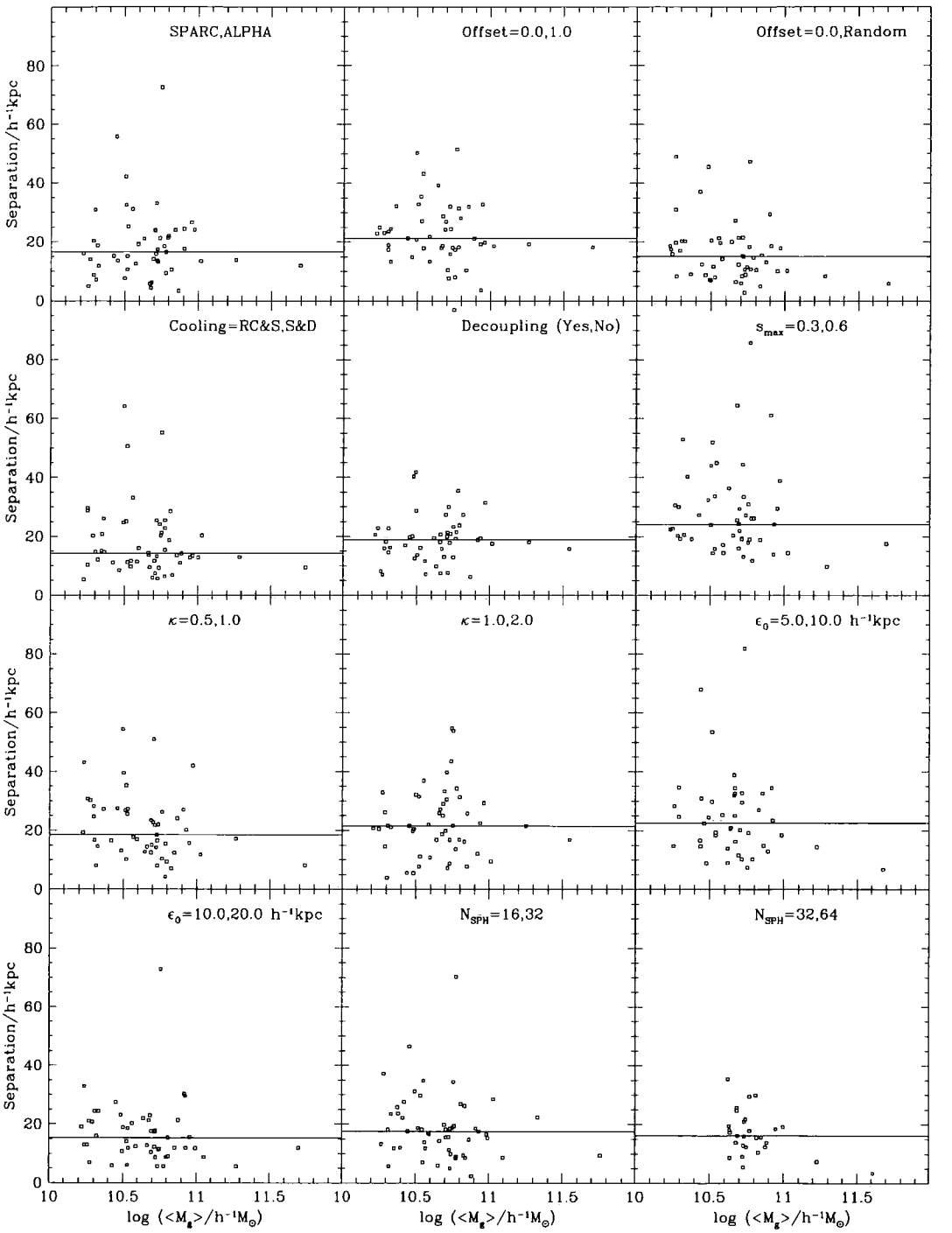


Figure 4.12: Separations between matched pairs of galaxies in the fiducial simulation and all other simulations, plotted against the average galaxy mass.The labels indicate the parameter that varies in each plot. The median scatter is shown as a solid horizontal line. Matches above $100\,h^{-1}$ kpc are not plotted.

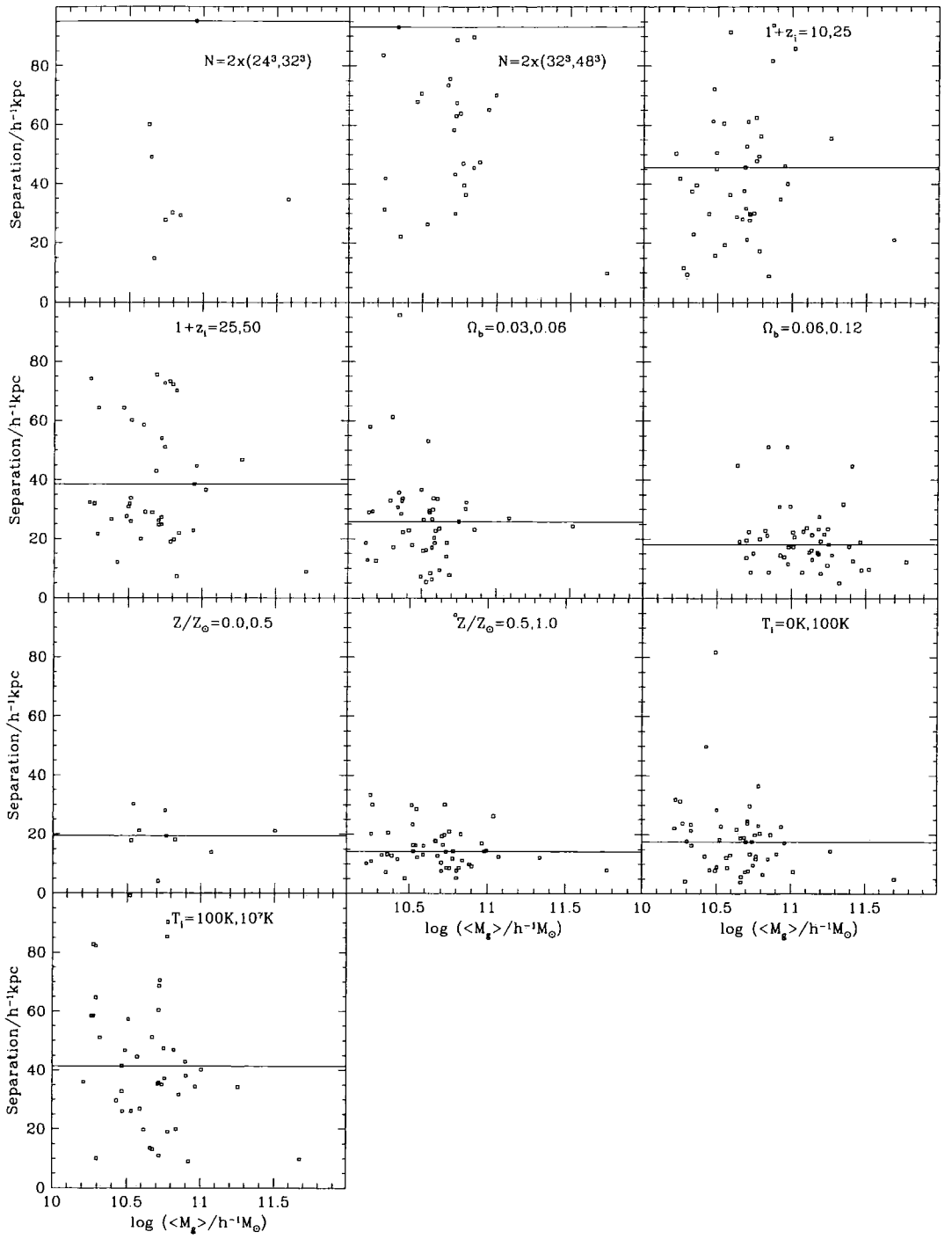


Figure 4.13: Separations between matched pairs of galaxies (continued).

A systematic difference in the matched positions is also introduced when the initial redshift of the simulation is changed: the median displacement increases by a factor of two in both comparisons. Variations in the initial positions and velocities of the particles leads to asynchrony in the subsequent spatial trajectories of the galaxies.

Regarding the physical comparisons, substantially increasing the initial temperature of the gas leads to a significant effect. Particles that have been supplied with this much thermal energy naturally have greater pressure support, which directly affects the equation of motion, and therefore the particle trajectories.

4.4.4.2 Galaxy–galaxy masses

Another measurement that has been performed is the differences in the masses of the matched galaxy pairs. Figures 4.14 and 4.15 illustrate the results by showing the logarithm of the mass ratio of each matched pair plotted against the average baryonic mass, always defined as the variant relative to the fiducial case. The solid line illustrates a ratio of unity (i.e. equal masses) and the dashed line illustrates the median value of the mass ratio. The median and semi–interquartile ranges are given in columns 6 & 7 of Table 4.4.

Adopting the tabulated cooling function causes an increase in the mass of each galaxy, because of an overall enhancement in the cooling rate around the temperatures most appropriate for these simulations ($\sim 10^5 - 10^7$K). Decoupling the hot gas from the cold gas causes the masses of the galaxies to decrease slightly with only a few exceptions, notably the largest object. Decreasing the softening length from $\epsilon_0 = 10$ to $5\,h^{-1}$ kpc also has the effect of decreasing the masses of the galaxies. Similarly, a larger value of $N_{\rm SPH}$ causes the masses of the galaxies to be systematically lower, due to its effect on the density field and hence the cooling rate.

The dispersion in the mass ratio for the comparisons with varying N is large because of the large scatter in object positions for these runs. Thus, spurious matches much more likely. For $N = 2 \times 24^3$ the objects are significantly less massive, whilst for $N = 2 \times 48^3$ the objects are only a little more massive than the fiducial simulation.

The largest differences in galaxy mass result from altering the cooling rate via the physical parameters, $\Omega_{\rm b}$ and Z. For the comparisons with varying baryon fraction, the galaxies are more than a factor of two heavier for the larger value of $\Omega_{\rm b}$. Furthermore, there appears to be a slight trend for increasing mass excess for heavier galaxies but this is difficult to measure due to the small dynamic range in mass. Varying the metallicity produces a similar effect to varying $\Omega_{\rm b}$.

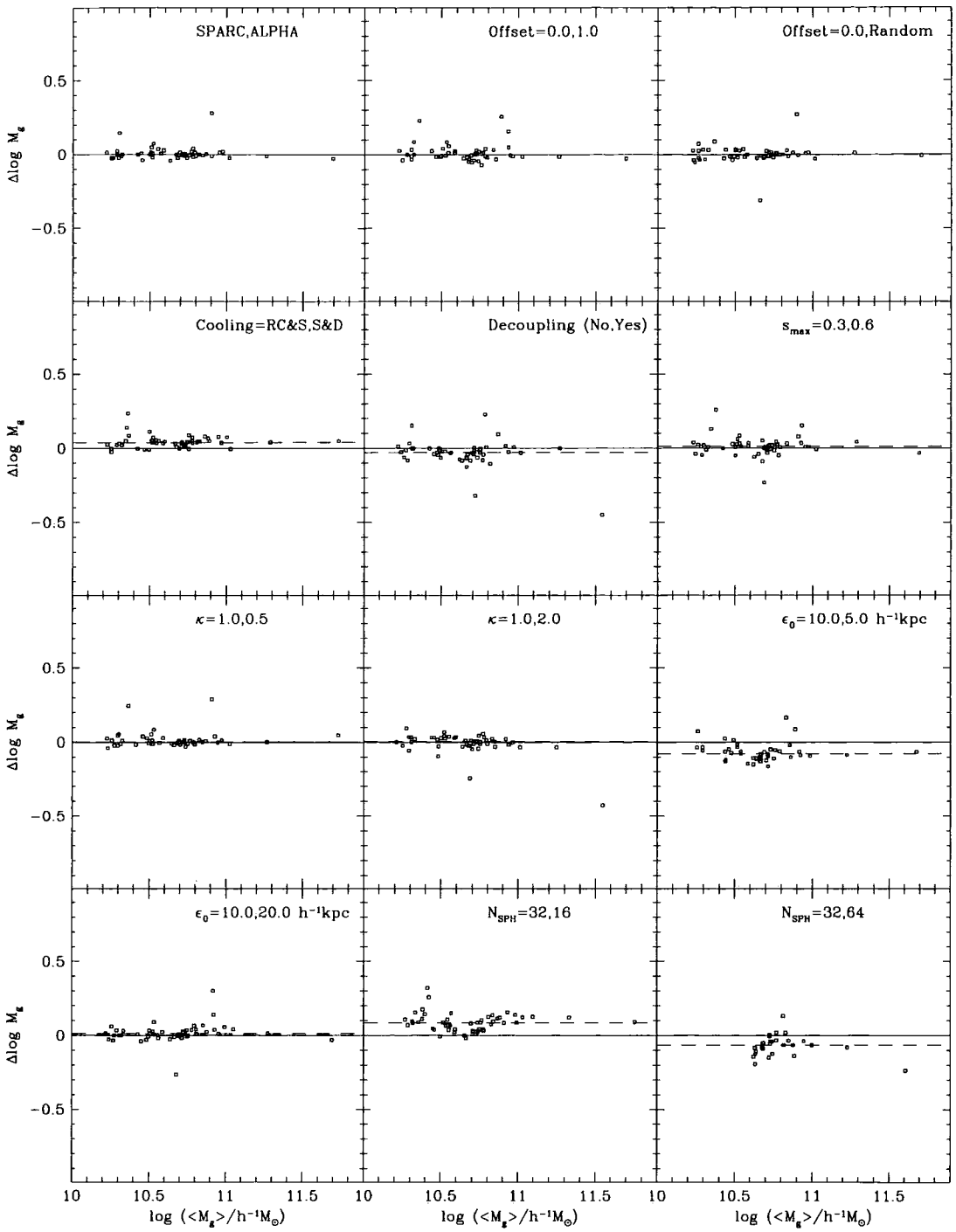


Figure 4.14: The logarithm of the ratio of the masses of matched pairs of galaxies plotted against the logarithm of the average galaxy mass. Labels indicate the parameter corresponding to each comparison (with the relative difference defined as the second minus the first value). The solid line corresponds to equal masses and the dashed line illustrates the median value for each comparison.

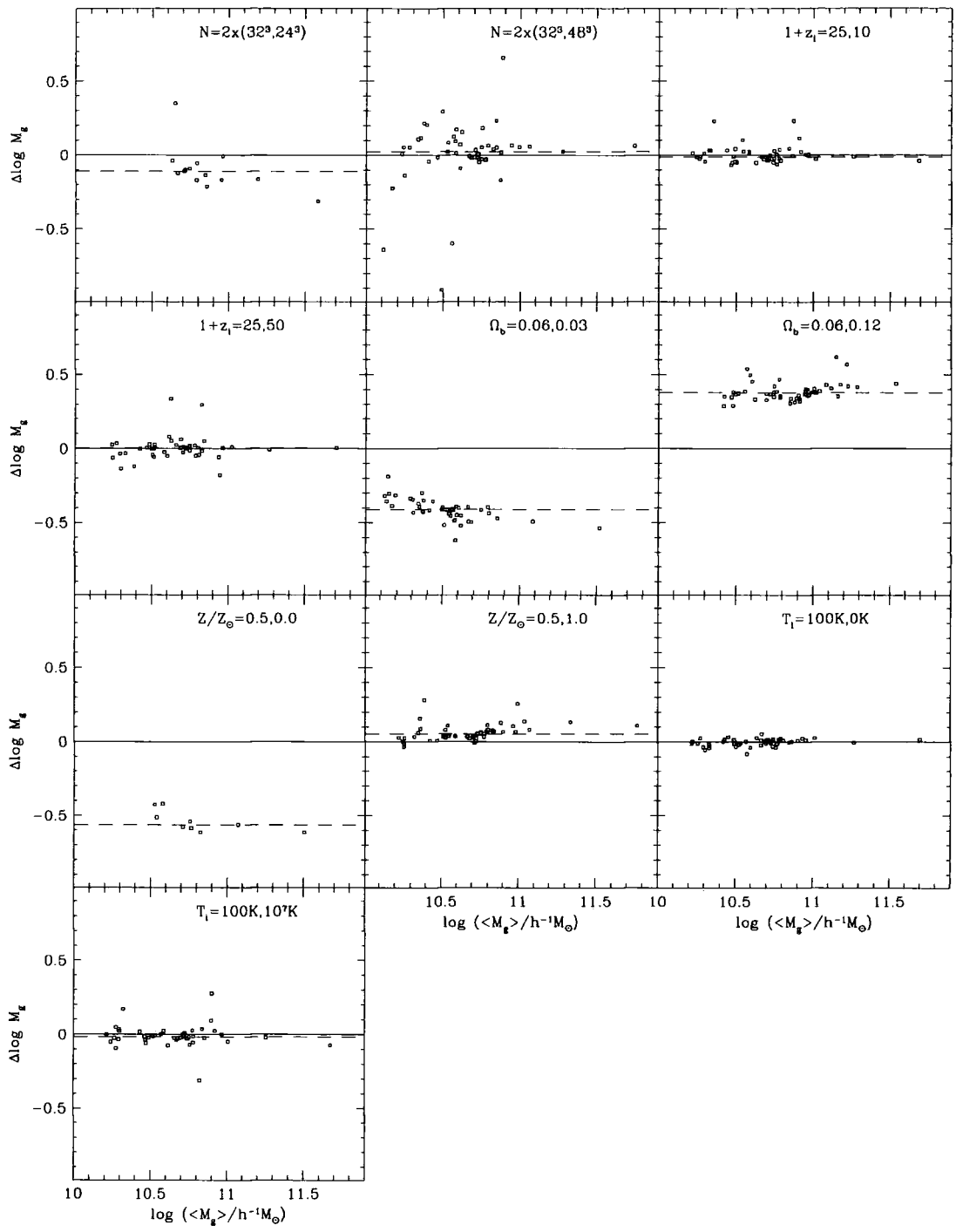


Figure 4.15: The logarithm of the ratio of the masses of matched pairs of galaxies (continued).

4.4.4.3 Dark matter halo–halo masses

The method used for comparing masses of matched galaxy pairs was also performed on the dark matter haloes. Results are illustrated in Figures 4.16 and 4.17; columns 8 & 9 in Table 4.4 quantify the differences. Overall, very tight correlations are seen for most of the simulations, as the dark matter is the dominant source of the gravitational potential and so is not significantly affected by changes in the gas dynamics. Changing the FFT resolution affects the force calculation through the softening and thus affects the dynamical evolution of the dark matter. Changing the cooling rate has a small effect on the size of the dark matter haloes for the largest baryon fraction considered. In this case, the larger amount of cooled material slightly deepens the potential, making it harder for the haloes to suffer from tidal disruption as well as causing an increase in the accretion rate.

4.4.5 The distribution of baryons in haloes

The final item investigated in this study is how parameter variations affect the final distribution of the baryons within haloes, as a function of the halo virial mass. In particular, the mass fraction of baryons present in the galaxy phase, the ratio of baryon masses for matched haloes and the halo baryon fractions have been analysed for each simulation.

4.4.5.1 Fraction of baryons in galaxies

Figure 4.18 illustrates the fraction of baryons in the galaxy phase as a function of the halo virial mass, for each comparison. Because of the small samples, the data is binned into histograms; the last bin corresponds to the single largest object. This statistic is remarkably stable; the first two comparisons show virtually no difference. The trend is for higher mass haloes to have less of their mass in galaxies, reflecting their longer (average) cooling times than the less massive haloes. (Since larger mass systems form later, on average, than lower mass systems, their characteristic densities are lower, and the cooling time, $\tau_{\rm cool} \propto \rho_{\rm gas}^{-1}$.) The mass fraction varies from ~ 0.65 in the smallest haloes down to ~ 0.2 in the highest mass halo. The comparison between the fiducial simulation and the case in which the galaxy phase is decoupled from the hot gas reflects the reduction in the total amount of galaxy phase baryons in the latter – about a factor of three for the galaxy in the largest halo ($\sim 5 \times 10^{13} h^{-1}\,{\rm M}_\odot$), but only ~ 10 per cent for haloes with masses of $\sim 10^{11} h^{-1}\,{\rm M}_\odot$. Other comparisons that give significant changes in the mass fractions show nothing unexpected.

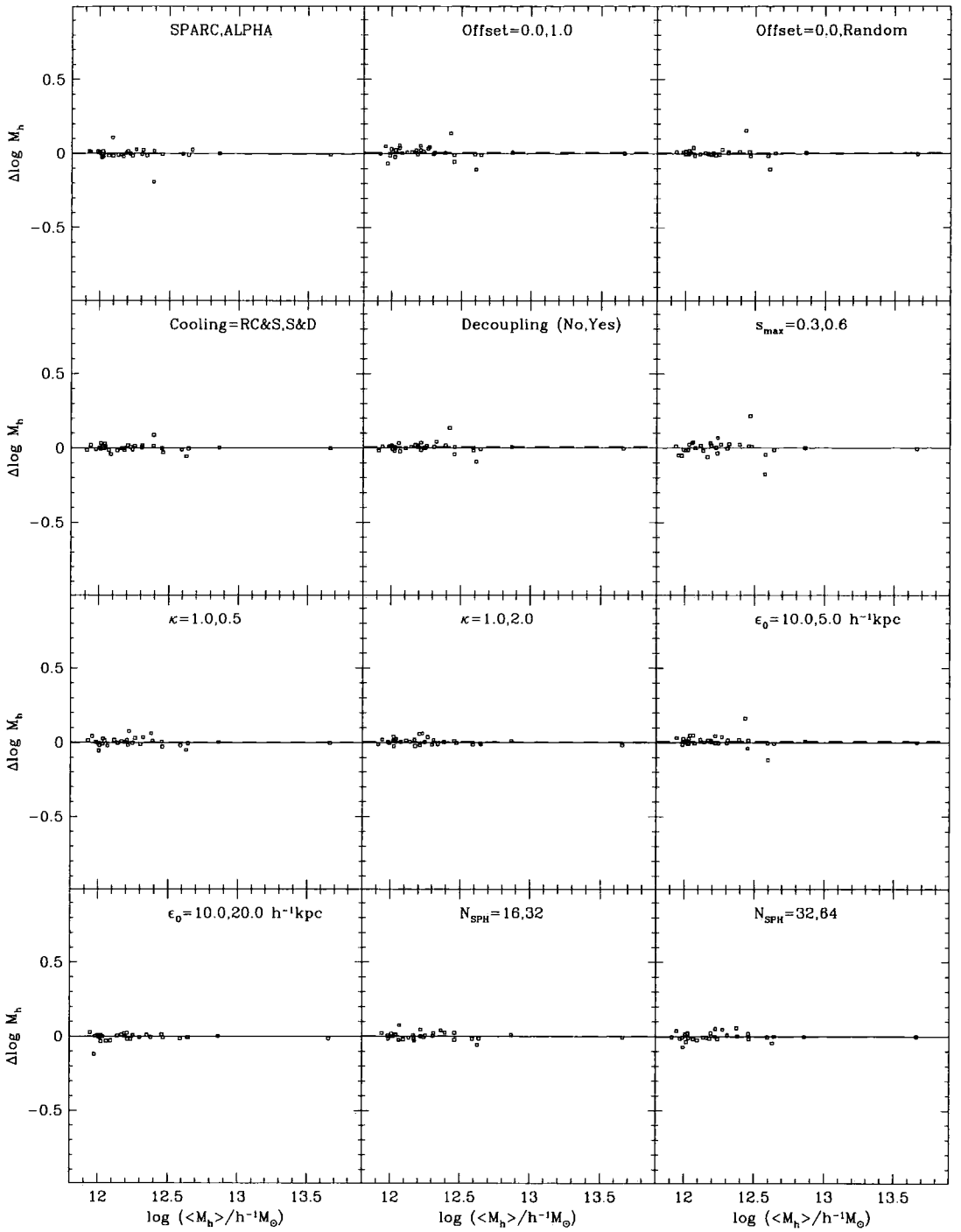


Figure 4.16: The logarithm of the ratio of the dark matter masses of matched pairs of haloes plotted against the logarithm of the average mass for each pair. Labels indicate the parameter corresponding to each comparison (with the relative difference defined as the second minus the first value). The solid line corresponds to equal masses and the dashed line illustrates the median value for each comparison.

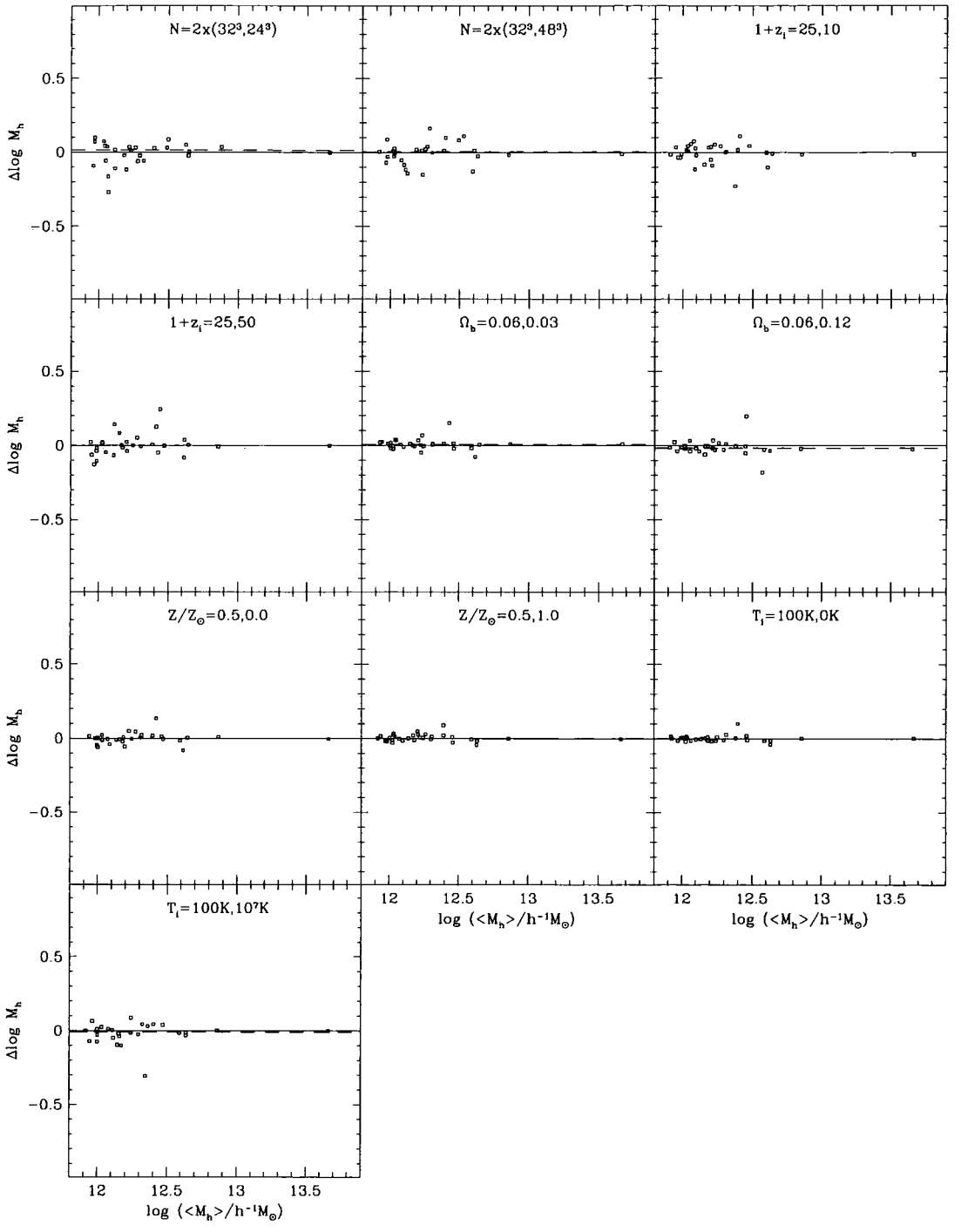


Figure 4.17: The logarithm of the ratio of the dark matter masses of matched pairs of haloes (continued).

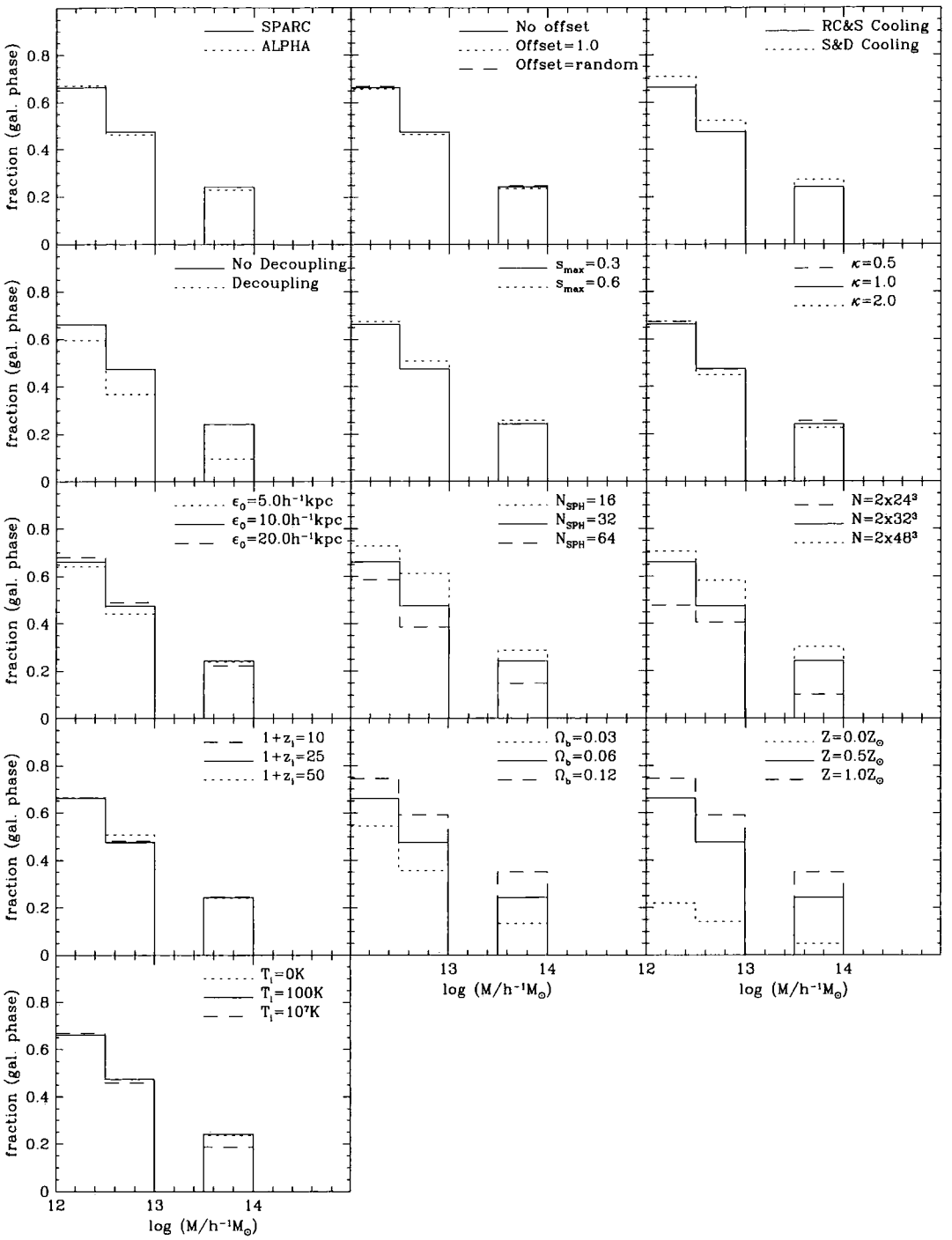


Figure 4.18: The mass fraction of baryons in the galaxy phase plotted as a function of the total halo mass. The fiducial simulation is the solid histogram.

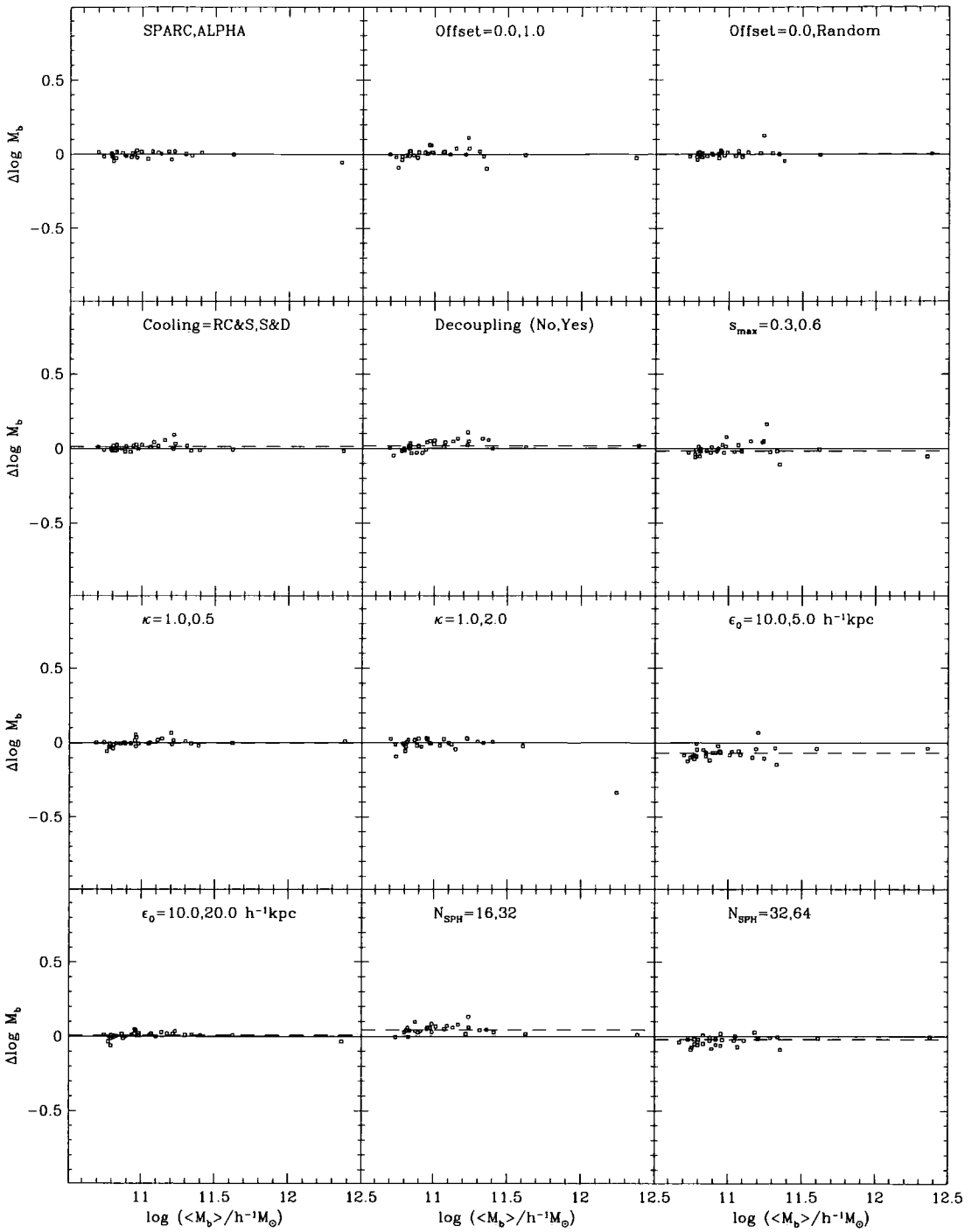


Figure 4.19: The logarithm of the ratio of the baryon masses of matched pairs of haloes plotted against the logarithm of the average baryon mass for each pair. Labels indicate the parameter corresponding to each comparison (with the relative difference defined as the second minus the first value). The solid line corresponds to equal masses and the dashed line illustrates the median value for each comparison.

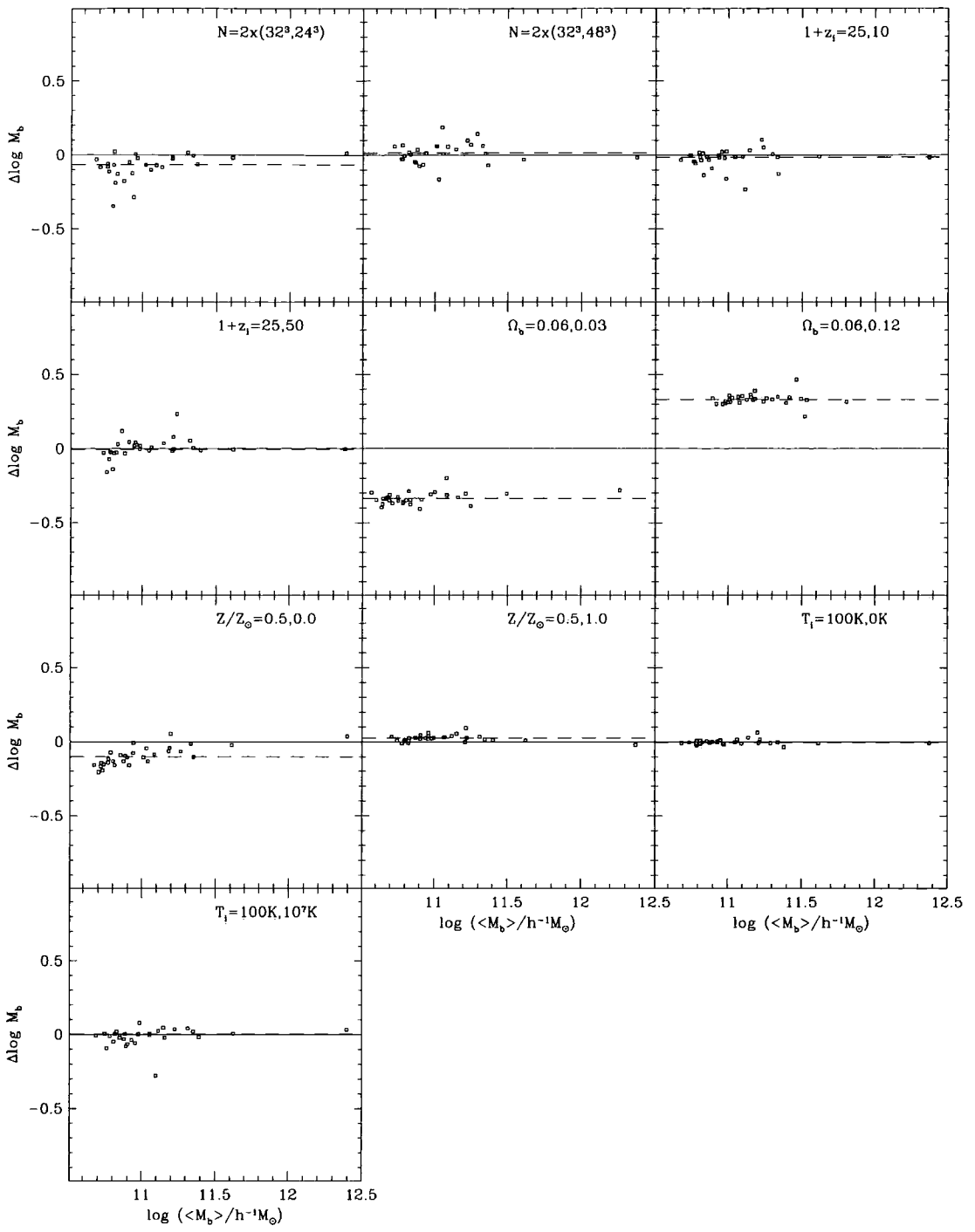


Figure 4.20: The logarithm of the ratio of the baryon masses of matched pairs of haloes (continued).

4.4.5.2 Local halo baryon fraction

The variations in the total mass of baryons within the virial radius of matched haloes are quantified in columns 10 & 11 of Table 4.4 (see also Figures 4.19 and 4.20). The median differences are slightly higher than for the dark matter, but smaller than the galaxies alone. Interestingly, the sign of the median difference is not always the same as the equivalent comparison for the galaxies, indicating that fluctuations in the total baryon content are not simply due to the galaxies alone. Significant differences are most evident in the simulations in which parameter changes have affected the cooling rate. Inducing a higher cooling rate causes the gas to lose some of its pressure support and therefore to become more compressed within its halo. The simulation with $\epsilon_0 = 5\,h^{-1}\,\mathrm{kpc}$ shows a significant change in the same direction as the change in the galaxy masses, supporting the idea that the gas is being more strongly heated from two–body encounters with the dark matter particles.

The total baryon fraction in haloes has also been analysed as a function of the halo virial mass. The results are displayed in Figure 4.21. The fraction is normalised to the global value of Ω_b for each simulation. The fiducial simulation has halo baryon fractions ranging from $\sim$ 93 per cent of the global value in the smallest haloes to $\sim$ 82 per cent in the largest. All simulations agree with the fiducial value to within $\sim$ 20 per cent, and are consistent with the results as found by Pearce et al. (1999). Notably, not all simulations show a monotonically decreasing baryon fraction with halo mass. However, the scatter is considerable because of the small number of objects in each bin (with only one in the highest mass bin). For almost all of the simulations and across the entire range of halo masses, the baryon fraction is less than the global value.

4.4.5.3 The mass to light ratio

Combining results from Figures 4.18 & 4.21 gives the variation in the fraction of the total mass that is locked up in galaxies, i.e. $M_\mathrm{gal}/M_\mathrm{tot} = f_\mathrm{gal}\Omega_\mathrm{b}$, with halo mass. Although the normalisation is sensitive to the amount of gas that cools (and hence it would be unfair to compare values with observations), all of the simulations presented here exhibit a trend of decreasing mass fraction in galaxies with halo mass, due to the behaviour of f_gal. This trend implies that the mass to (optical) light ratio (which scales inversely with $M_\mathrm{gal}/M_\mathrm{tot}$) increases with halo mass. Using the fiducial results - purely as an example - and assuming that the galaxies have $M/L_B \sim 5$ ($M/L_B \sim 8$ for E/S0 galaxies and $M/L_B \sim 3$ for spiral galaxies), the value of M/L_B is around 150 for $M \sim 10^{12}\mathrm{M}_\odot$ increasing to $\sim$ 400 for the largest object ($M \sim 10^{14}\mathrm{M}_\odot$).

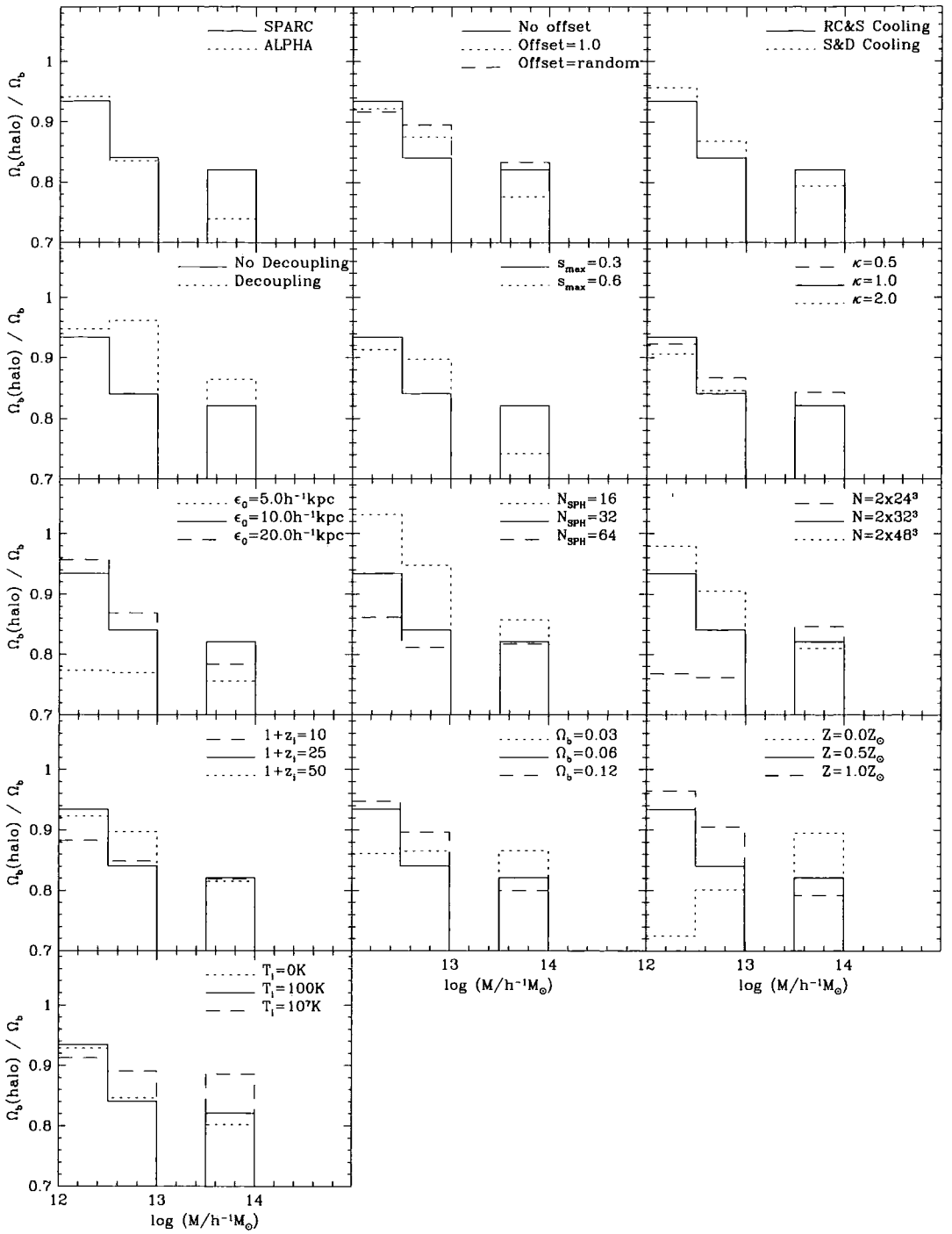


Figure 4.21: The baryon fraction in haloes as a function of the total halo mass, in units of the global baryon fraction. The fiducial simulation is the solid histogram.

Observations of the mass to light ratio from poor groups to rich clusters (i.e. for $M \sim 10^{12} - 10^{16}\mathrm{M}_{\odot}$) have yet to agree on whether the mass to light ratio increases with halo mass or is constant. For example, results from the Northern CfA Redshift Survey indicate that the median M/L_B in groups is around 250, rising by up to a factor of 4 in rich clusters (Ramella, Pisani, & Geller 1997). However, results using masses determined from X-ray observations suggests that M/L_B is approximately constant (at around 150-200) for $M = 10^{12} - 10^{15}\mathrm{M}_{\odot}$ (David, Jones, & Forman 1995) Imminent results from large redshift surveys (such as the Sloane Digital Sky Survey and the 2dF Galaxy Redshift Survey) as well as X-ray observations from the new Chandra and XMM satellites should provide a significantly more robust determination of the mass to light ratio for a wide range of halo masses.

4.5 Summary

The main effects observed in the comparison are now summarised.

In comparison 1, the effects introduced by running the simulation on different computer architectures were studied. No significant differences were found between the two simulations. However, the displacements between matched pairs of galaxies at $z = 0$ were found to be of the order of an S_2 softening length. It is therefore incorrect to trust either the positions of individual galaxies or their trajectories on scales smaller than this.

Comparison 2 was intended to look at the effects of small displacements in the positions of the particles at the start of the simulation. One simulation was run with the particles initially displaced by 1 FFT cell to test the periodic nature of the algorithm. Another simulation was performed with tiny random perturbations applied to the initial particle positions, to test the level at which differences can grow over the course of a few thousand timesteps (the typical length of each simulation). Again, no significant changes were detected amongst the three simulations.

The cooling function was then changed from the series of power–law fits, used in the fiducial simulation, to a set of tabulated values from Sutherland & Dopita ((Sutherland & Dopita 1993)). A slight enhancement in the galaxy masses, at the 10 per cent level, was produced in the latter case, although no significant changes were produced in the positions of the galaxies.

In comparison 4, the fiducial simulation was compared to one in which the galaxy phase gas is partially decoupled from hot gas in the SPH algorithm (see Pearce et al. 1999). This modification provides a better estimate of the hot gas density near galaxies. Specifically it prevents an overestimate leading to an artificial enhancement in the cooling rate. Only the largest galaxy in the simulation was significantly affected, with its mass being reduced by about a factor of three.

Comparison 5 consisted of doubling the initial (maximum) softening length. This results in a period of time in which the softening is larger than in the fiducial simulation. This led to slight differences in the positions of the galaxies, although it left the mass distribution unaffected.

The size of the timesteps was altered by a factor of two each way, in comparison 6, by changing the constant, κ. Increasing the timestep resulted in much more shocked gas at the expense of the uncollapsed phase. This illustrates the effect of the timestep on the dynamics of the particles: introducing significant errors into their positions and velocities leads to enhanced viscous heating of the gas. The galaxies themselves remained largely unaffected. No significant change was evident when halving the value of κ, thus justifying the value $\kappa = 1$ which was adopted as fiducial.

Comparison 7 consisted of changing the final value of the softening length, ϵ_0. In the simulation with $\epsilon_0 = 5\,h^{-1}$ kpc, the galaxies were slightly less massive than in the fiducial simulation ($\epsilon_0 = 10\,h^{-1}$ kpc), possibly due to discreteness effects caused by harder two–body encounters. A larger value of $\epsilon_0 = 20\,h^{-1}$ kpc produced no significant differences.

In comparison 8, the value of $N_{\rm SPH}$ was varied, making it larger and smaller by a factor of two, thus looking at the effects produced when the SPH algorithm uses a different number of neighbours when calculating the smoothing length for each gas particle. These changes had a dramatic effect on the simulations. The value of $N_{\rm SPH}$ determines the minimum mass for which the gas density can be resolved: smaller values will resolve smaller masses at the expense of larger random fluctuations. It was found that the number of galaxies increases strongly with smaller $N_{\rm SPH}$, as did the masses of the matched galaxy pairs, although without significantly affecting their positions.

Comparison 9 consisted of changing the number of particles in the simulation, thus altering the mass resolution. To demonstrate the effect of spurious heating by dark matter particles, a simulation was run with a dark matter particle mass well above the critical value calculated by Steinmetz & White (1997). In this case, gas is unable to cool and form galaxies, even though cooling times in the smallest haloes are significantly shorter than the age of the universe. For the main set of comparisons, a suite of runs with $N = 2 \times [24, 32, 48]^3$ were compared. The number and mass of galaxies increased with N. Improving the mass resolution also affected the final distribution of dark matter, due to the extra small–scale power. The two simulations with the highest resolution were in better agreement than the two with the lowest resolution, albeit that the discrepancies in the galaxy phase material (i.e. the fraction of cooled gas) were still large.

In comparison 10, the initial redshift of the simulations was changed ($1 + z_i = 10, 25, 50$). The masses of the objects were not significantly altered by this change but small changes in the spatial distribution of the galaxies were produced.

Comparison 11 involved changing the value of the baryon fraction ($\Omega_{\rm b}$=0.03,0.06,0.12). The higher value of $\Omega_{\rm b}$ causes the cooling rate to increase (since emissivity is proportional to $\Omega_{\rm b}^2$). The mass of each galaxy also increases, since m is proportional to Ω_b. The run with $\Omega_{\rm b} = 0.03$ has dark matter masses of the order of the critical mass for two–body heating, although the reduction in the cooling rate produced by this effect is confused by the fact that the cooling rate is significantly smaller anyway. Overall, large differences were noted in all analyses in which the baryon fraction was varied. In comparison 12, the value of the (unevolving) gas metallicity was varied ($Z = 0.0, 0.5, 1.0$ in solar units). The run with no metals differed dramatically from the fiducial simulation, producing very few objects by $z = 0$. This simulation had a dark matter mass well above the critical value for two–body heating. The run with $Z = 1.0$ showed an enhancement in galaxy masses relative to the fiducial simulation, although the change was far less severe.

Finally, in comparison 13 the initial gas temperature was changed from 100K to 0K and 10^7K respectively. No significant changes were measured between the former two. Although the uncollapsed gas in the latter simulation was on a higher adiabat ($T \sim 10^3$K at $z = 0$), this had very little effect on the final state of the gas, with the exception of introducing significant displacements between the matched galaxy pairs.

4.6 Conclusions

This chapter has explored how variations in simulation parameters and certain aspects of the physical assumptions affect the population of galaxies which form in cosmological simulations. For this process, a minimum of physical assumptions was applied: the ability of the gas to cool radiatively to achieve suitably high density contrasts. A simulation was run with a fiducial set of carefully chosen parameters and a number of variants were performed each differing by the value of one parameter. Galaxy and halo catalogues were selected in a consistent manner before analysing the effects that the parameter changes had on the thermodynamic state of the gas, on the mass and spatial distribution of the galaxies, and on the baryon content of haloes.

By carefully tracking the phase evolution of gas particles it has been shown that most galaxy material (around 90 per cent in the fiducial simulation) originally cools in small haloes where the virial temperature is $\sim 10^5$K. As Figure 4.18 shows, the percentage of halo baryons that are cold rises as the halo mass falls. Only a small fraction of the baryons cool at later stages of the

evolution. Mergers of these small fragments can reheat the gas briefly as massive galaxies are built up. This implies that the most dynamically interesting part of the process occurs at the resolution threshold of the simulation and is therefore sensitive to many numerical and physical parameters.

Numerical noise leads to a scatter in the final galaxy positions that is of the order of the (S_2) softening for most galaxies but can be much higher in some cases. Individual galaxy trajectories should therefore be treated with caution.

The length of timestep recommended by Couchman, Thomas & Pearce (1995) is adequate. Halving it caused no significant change in the galaxy catalogues, but choosing a larger value led to a significant increase in the amount of shocked gas at the expense of the uncollapsed gas. This can be attributed to errors in the particle positions and velocities, resulting in an artificial enhancement in the viscous heating of the gas.

It is confirmed that simulations with dark matter particle masses greater than the Steinmetz & White (1997) critical mass for two–body heating do indeed show a significant reduction in the final amount of cooled gas and in the number of galaxies, even when the shortest cooling time is a small fraction of the age of the universe. The choice of dark matter particle mass should be comfortably below this critical value. The mass in galaxies is also strongly affected by the choice of baryon fraction and metallicity.

Pre–heating of the gas to a high initial temperature (10^7K at $z = 24$) produces small changes in the final positions of the galaxies but has little effect on their final mass. Gas on a high adiabat ($T > 10^4$K) at high redshift can radiate a significant fraction of its energy away which significantly weakens its effect on the cooling rate of the gas.

Increasing the mass resolution by using a larger number of particles causes both the number of galaxies and their masses to increase. Improved resolution of the density field enhances the cooling rate within haloes, particularly at high redshift when the first objects form.

Lowering the number of neighbours used by the SPH algorithm to calculate hydrodynamical quantities has a similar effect on the amount of cooled material to increasing the number of particles in the simulation. Smaller values of $N_{\rm SPH}$ enable higher overdensities to be resolved (therefore enhancing the cooling rate in such regions), at the expense of more contamination from sampling noise.

The fraction of the baryonic material that cools within a halo and forms galaxies is a very stable quantity which varies inversely with halo mass. Reducing the total amount of gas affects all scales in a similar way. The ratio of baryonic to dark material within the virial radius of haloes is around $80 - 90$ per cent of the global value, for the range of halo masses studied.

This study involved using four different prescriptions for the gravitational softening. The initial comoving softening was varied from $20\,h^{-1}$ kpc to $40\,h^{-1}$ kpc and the final physical softenings were varied from $5\,h^{-1}$ kpc to $10\,h^{-1}$ kpc and $20\,h^{-1}$ kpc. With the exception of the $5\,h^{-1}$ kpc case, this variation had very little effect on the galaxy population. The smaller final softening systematically reduced both the mass of galaxies and the mass of baryons within haloes by ~ 20 per cent. This effect is probably due to the artificial heating of the gas from two–body encounters with the dark matter particles (Steinmetz & White 1997).

Finally, in almost every simulation an overly massive galaxy was produced at the centre of the largest halo. With a baryonic mass of $\sim 5 \times 10^{11} h^{-1}\,\mathrm{M}_{\odot}$, such an object is not expected within such a small volume of the universe. Decoupling the central dense gas from the hot halo material effectively reduces the mass of this object to a more reasonable value (see also Pearce et al. 1999).

In conclusion, since nearly all of the baryons found in galaxies cooled in haloes at the resolution threshold of the simulation (where the cooling times are generally shortest), modelling the galaxy formation process is extremely susceptible to numerical effects. Remarkably most of the parameters studied did not have a severe effect on the final mass and spatial distributions of the galaxies. The most important factor is the global amount of gas which cools to high density. Large differences are easily obtained by changing the cooling rate through the physical parameters, Ω_b and Z, as well as by altering the mass resolution of the simulation. Although the presence of a resolution threshold *itself* limits the cooling rate of primordial gas, this limitation is clearly unphysical. Future simulations will undoubtedly have much improved resolution. It is therefore of paramount importance that physically–motivated heating processes (such as thermal feedback from SNe) are included in order to inhibit overcooling at early times. Without proper inclusion of such processes it will be impossible to obtain convergent results, independent of numerical artifacts, for simulations of galaxy formation, within the now standard paradigm for structure formation.

References

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Baugh C. M., Cole S., Frenk C. S., 1996, MNRAS, 282, L27

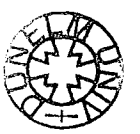
Binney J., 1977, ApJ, 215, 483

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Bower R. G., 1991, MNRAS, 248, 332

Cen R., Ostriker J. P., 1992, ApJ, 393, 22

Centrella J., Melott A. L., 1983, Nature, 305, 196



Cole S., 1991, ApJ, 367, 45

Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 1999, MNRAS, submitted

Copi C. J., Schramm D. N., Turner M. S., 1995, ApJ, 455, 95

David L. P., Jones C., Forman W., 1995, ApJ, 445, 578

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Efstathiou G., Davis M., Frenk C. S., White S. D. M., 1988, MNRAS, 235, 715

Eke V. R., Cole S. M., Frenk C. S., 1996, MNRAS, 282, 263

Evrard A. E., Summers F., Davis M., 1994, ApJ, 422, 11

Frenk C. S., White S. D. M., Davis M., 1983, ApJ, 271, 417

Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, ApJ, 327, 507

Gerritsen J. P. E., Icke V., 1997, A&A, 325, 972

Governato F., Baugh C. M., Frenk C. S., Cole S., Lacey C. G., Quinn T., Stadel J., 1998, Nature, 392, 359

Haehnelt M. G., Kauffmann G., 1999, MNRAS, submitted (astro-ph/9906493)

Hoyle F., 1953, ApJ, 118, 513

Katz N., Gunn J. E., 1991, ApJ, 377, 365

Katz N., Hernquist L., Weinberg D. H., 1992, ApJ, 399, L109

Kauffmann G., Guiderdoni B., White S. D. M., 1994, MNRAS, 267, 981

Klypin A. A., Shandarin S. F., 1983, MNRAS, 204, 891

Lacey C., Cole S., 1994, MNRAS, 271, 676

Mushotzky R., Loewenstein M., Arnaud K. A., Tamura T., Fukazawa Y., Matsushita K., Kikuchi K., Hatsukade I., 1996, ApJ, 466, 686

Navarro J. F., Steinmetz M., 1997, ApJ, 478, 13

Pearce F. R. et al., 1999, ApJ, 521, L99

Press W. H., Schechter P., 1974, ApJ, 187, 425

Ramella M., Pisani A., Geller M. J., 1997, AJ, 113, 483

Raymond J. C., Cox D. P., Smith B. W., 1976, ApJ, 204, 290

Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541

Silk J., 1977, ApJ, 211, 638

Somerville R. S., Primack J. R., 1998, MNRAS, submitted (astro-ph/9802268)

Steinmetz M., White S. D. M., 1997, MNRAS, 288, 545

Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253

Thacker R. J., Tittley E. R., Pearce F. R., Couchman H. M. P., Thomas P. A., 1998, MNRAS, submitted (astro-ph/9809221)

Viana V. L., Liddle A. R., 1996, MNRAS, 281, 323

Weinberg D. H., Hernquist L., Katz N., 1997, ApJ, 477, 8

White S. D. M., Efstathiou G., Frenk C. S., 1993, MNRAS, 262, 1023

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Chapter 5

Modelling star formation and feedback within cosmological simulations

Abstract. A crucial ingredient of any hierarchical model of galaxy formation is a non-gravitational heating source, able to prevent baryons from cooling catastrophically and forming stars at high redshift. This chapter investigates the modelling of such processes in cosmological simulations of galaxy formation, specifically the formation of stars and the effects of the energy associated with their supernovae (generically referred to as "feedback"). First, the overcooling problem is illustrated explicitly, studied using a series of simulations with increasing numerical resolution. The box-size and mass ratio of gas to dark matter particles are also considered. Second, various prescriptions for turning gas into stars are investigated, demonstrating that the global star formation rate is determined by the cooling rate of the gas, providing most of the baryons in galaxies are in stars. The latter quantity is most sensitive to the densities at which star formation is allowed to proceed. Third, both thermal and kinetic feedback prescriptions are studied in various implementations. Only the thermal method gives encouraging results, providing the reheated gas sustains a short period without cooling.

5.1 Introduction

The incorporation of feedback effects into cosmological simulations presents a formidable challenge, with no group claiming to have found a satisfactory solution. The reason the problem is so difficult is that the physical processes responsible for the redistribution of energy in the interstellar medium (ISM) occur on a scale that is several orders of magnitude below what can presently be resolved. However, as was discussed in the previous chapter, the nature of hierarchical galaxy formation demands additional, non-gravitational heat sources in order to prevent gas from cooling catastrophically and turning into stars at high redshift. Even at the resolutions currently attainable with simulations, the amount of cold dense gas that cools exceeds current estimates of the *cosmic baryon budget*. This suggests that only $\sim 10 - 20$ per cent of the baryons in the local universe are cold dense galactic material (mainly stars and HI atomic gas); the majority of the baryons are presumably in the form of a hot diffuse plasma between groups and clusters of galaxies (e.g. see

Fukugita, Hogan, & Peebles 1998). Furthermore, the inclusion of feedback will not only serve to avoid the *cooling catastrophe* but also to produce numerically convergent results.

A number of other problems with hierarchical models have been attributed to the exclusion of feedback effects. For example, the formation of disks occur naturally within galaxy-sized haloes, providing the disk does not lose a significant amount of angular momentum to the halo (Fall & Efstathiou 1980). Such a loss is witnessed in numerical studies, as a result of the merging process (e.g. Frenk et al. 1985; Navarro & White 1993). However, if feedback can successfully keep the gas hot long enough for the merger rate to settle down, the angular momentum problem can be prevented (Eke, Efstathiou, & Wright 1999). On larger scales, the X-ray properties of galaxy clusters do not obey scaling relations expected from simple gravitational instability models (Kaiser 1986). As discussed in Chapter 2, the observed cores in their ICM distribution may be an artifact of pre-heating due to feedback effects, before the cluster formed (Kaiser 1991; Evrard & Henry 1991; Ponman, Cannon, & Navarro 1999).

The main source of feedback is believed to be the supernova ejecta of stars, capable of effectively disrupting the cooling process in pre-galactic objects (e.g. Dekel & Silk 1986) and consequently reducing the faint end of the predicted luminosity function to within reasonable agreement with observations (White & Frenk 1991; Cole 1991). Within the context of numerical simulations, feedback has been studied mainly in its applications to individual galaxies, to address the angular momentum problem, as well as its ability to self-regulate quiescent star formation (e.g. Katz 1992; Navarro & White 1993; Steinmetz & Müller 1995; Gerritsen & Icke 1997; Buonomo et al. 1999). The true nature of the ISM is multiphase (McKee & Ostriker 1977) and is complicated by, at least, the presence of magnetic fields and turbulence. Hence, the treatment of both star formation and feedback processes in simulations is necessarily crude. Star formation is approximated by the conversion of gas into a collisionless fluid and feedback is modelled by injecting supernovae energy, using simple energetics arguments, into surrounding gas.

Attempts to include the effects of star formation and feedback within cosmological simulations of random volumes have been made by several groups (e.g. Katz, Weinberg, & Hernquist 1996; Tissera, Lambas, & Abadi 1997; Pearce 1998). "Galaxies" within these simulations are dense blobs of cooled gas with internal properties hitherto unresolved. The aim of this chapter is to explore the effects of incorporating a selection of star formation and feedback prescriptions (applied to various kinds of simulations) into cosmological simulations, with the sole objective of reducing the global amount of gas that cools. All simulations performed in this chapter use Version 3 of HYDRA , as detailed in Chapter 3. The evolution of both baryonic (gas) and dark

matter species are followed using a total of N particles. Both species interact via gravity, whilst the gas can adiabatically heat and cool, shock heat and radiatively cool. The cooling function used is identical to that used in Simulation 3 of the previous chapter (taken from the work of Sutherland & Dopita 1993l; see also Figure 4.1).

The remainder of this chapter is organised as follows. In Section 5.2 the efficiency of radiative cooling is examined by studying its dependence on numerical resolution, the size of the simulation volume and the mass ratio of gas to dark matter particles. The effects of star formation are then studied in Section 5.3, comparing a set of twelve simulations, each testing different methods for converting gas into stars. A similar study is performed in Section 5.4, this time comparing simulations with differing feedback implementations. In Section 5.5, conclusions are drawn.

5.2 The efficiency of radiative cooling

5.2.1 Simulation details

This section extends work accomplished in the previous chapter regarding the effects of numerical resolution on the thermal distribution of gas. Of primary interest is the change in the efficiency of radiative cooling as the resolution is varied. The effects of varying the size of the simulation volume and the ratio of dark matter to gas particle masses are also studied.

Initial particle positions and velocities were generated in an identical fashion to those used in the previous chapter. To summarise, the standard cold dark matter (SCDM) model was adopted, hence setting $\Omega = 1$ (the density parameter), $\Lambda = 0$ (the cosmological constant) and $h = 0.5$ (Hubble's constant in units of $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$). The amplitudes of the initial (linear) density fluctuations were determined by setting $\sigma_8 = 0.6$ (the rms value for spheres of radius $8\,h^{-1}$ Mpc), in agreement with observations of rich clusters in the local universe. For the simulations with $S = 10\,h^{-1}$ Mpc (as used in the previous chapter), the identical sequence of random phases were used so that the same large scale structure was produced. This relatively small box-size can easily produce fluctuations in any statistical measurement when the structure is changed.

Table 5.1 lists the simulations that were performed and analysed in this chapter. Column 1 lists the run number while Columns 2 and 3 give the number of gas particles (N_{gas}) and total number of particles ($N = N_{\mathrm{gas}} + N_{\mathrm{dark}}$) used to sample the mass distribution. Columns 4-7 list the values of parameters that were varied in this study, namely the baryon fraction, box size, final (Plummer) softening length and the number of cells/side in the FFT mesh used by the PM algorithm. All other parameters were fixed at the values listed in Table 4.1.

Table 5.1. Various properties of the simulations analysed in Section 5.2. Column 1 indexes the simulations while Columns 2 and 3 give the number of gas and the total number (gas + dark matter) particles. Columns 4-7 lists values of the parameters that were varied in this study, namely the baryon fraction (Column 4), the simulation box size (Column 5), the final (Plummer) softening (Column 6) and the number of cells/side in the FFT mesh (Column 7). The gas mass resolution of each simulation (defined in the text) is given in Column 8. Finally, columns 9-11 list the gas fractions in the uncollapsed (f_u), shocked (f_s) and galaxy (f_g) phases at redshift zero.

Run	N_{gas}	N	Ω_b	$S/h^{-1}Mpc$	$\epsilon_0/h^{-1}kpc$	L	$\Re$	f_u	f_s	f_g
01	16^3	2×16^3	0.06	10.0	20.0	32	0.12	0.01	0.99	0.00
02	24^3	2×24^3	0.06	10.0	13.0	32	0.42	0.02	0.75	0.23
03	32^3	2×32^3	0.06	10.0	10.0	64	1.0	0.02	0.63	0.35
04	48^3	2×48^3	0.06	10.0	7.0	64	3.3	0.02	0.56	0.42
05	64^3	2×64^3	0.06	10.0	5.0	128	7.9	0.02	0.55	0.43
06	64^3	2×64^3	0.06	20.0	10.0	128	1.0	0.03	0.65	0.32
07	16^3	32^3	0.125	10.0	20.0	32	0.12	0.03	0.70	0.27
08	24^3	48^3	0.125	10.0	13.0	32	0.42	0.03	0.60	0.37
09	32^3	64^3	0.125	10.0	10.0	64	1.0	0.03	0.56	0.41
10	16^3	2×16^3	0.125	10.0	20.0	32	0.12	0.01	0.98	0.01
11	24^3	2×24^3	0.125	10.0	13.0	32	0.42	0.02	0.66	0.32
12	32^3	2×32^3	0.125	10.0	10.0	64	1.0	0.02	0.58	0.40

The simulations have been subdivided into three groups, separated by solid horizontal lines in Table 5.1. The first set (runs 1 to 5) was chosen in order to analyse the effects of increasing the resolution of the simulation within a fixed volume. This is similar to comparison 9 of Chapter 4, making two changes in addition to using a later implementation of HYDRA . Firstly, the range of N has been extended to include a simulation with $N = 2 \times 64^3$; the range in gas particle masses is now 64, from $\sim 6.4 \times 10^7 - 4.1 \times 10^9 h^{-1}\,M_\odot$. Secondly, the final softening length was varied as $\epsilon_0 \propto m^{1/3}$, where m is the particle mass. This choice of scaling for the softening is commonly adopted since it allows for increased force resolution with larger N, without changing the typical timescales for two-body encounters The value of ϵ_0 was normalised in which it was set equal to $10\,h^{-1}$ kpc for run 3. The evolution of the softening is also determined by the number of FFT cells (the initial softening is set to the length of 0.3 FFT cells), which can only take values that are integer powers of two. The value of L was set equal to the nearest integer power of two greater

than or equal to $2N_{\rm dark}^{1/3}$. Hence for runs 1,3 and 5 the softening scales proportional to $N_{\rm dark}^{-1/3}$ at all times.

The second group (run 6, along with run 3) was devised to test the effect of varying the box size at fixed resolution. The difference in volume between the two runs is a factor of 8. A larger box will resolve more larger mass haloes at a frequency determined by the underlying mass function, itself determined by the choice of initial power spectrum. The number of FFT cells is set as above and the softening is fixed between the runs, since the mass resolution is fixed.

Finally, runs 7-12 were performed in order to study the effects of using unequal mass particles between the two species (the mass of a gas particle is typically Ω_b/Ω_0 times the mass of a dark matter particle in cosmological simulations, such that the mass distribution for each species is sampled using equal numbers of particles). This was achieved by labelling every eighth particle as a gas particle, the remaining being dark matter. A value of $\Omega_{\rm b} = 0.125$ was chosen in order to give equal masses. Although this is around twice the value constrained by nucleosynthesis measurements (for the SCDM model), the choice is made for comparative purposes only. A second set of runs was also performed (10-12) identical to runs 1-3 except having $\Omega_{\rm b} = 0.125$ in order to provide a fair comparison to runs 7-9. The number of FFT cells was chosen for runs 7-12 as above; the final softenings were scaled with the gas particle mass.

5.2.2 *Gas properties at $z = 0$*

As in Chapter 4, the baryons are split into three phases depending on their densities and temperatures. Figure 5.1 shows the density-temperature phase diagram of the baryons at $z = 0$ for run 3. Contrasting this plot with the top-left panel of Figure 4.2, it is evident that the hard edge at low density is no longer present, since the uncollapsed gas now resides on an adiabat at densities below the mean value. This is because the SPH neighbour search is no longer limited to the local 27 PP cells. The smoothing length defines a sphere that now always contains, on average, $N_{\rm SPH}$ particles for every gas particle.

The phase boundaries are defined similarly to those in shown in Figure 4.2. The uncollapsed phase is defined to be that with $T < 0.3$K and $\rho < \langle\rho\rangle$, hence picking out the particles on the adiabat. Galaxy phase material is defined as the region with $T < 3 \times 10^4$K and $\rho > 500\,\langle\rho\rangle$. The rest of the gas is marked as the shocked phase. The choice of boundaries are appropriate for all of the runs listed in Table 5.1.

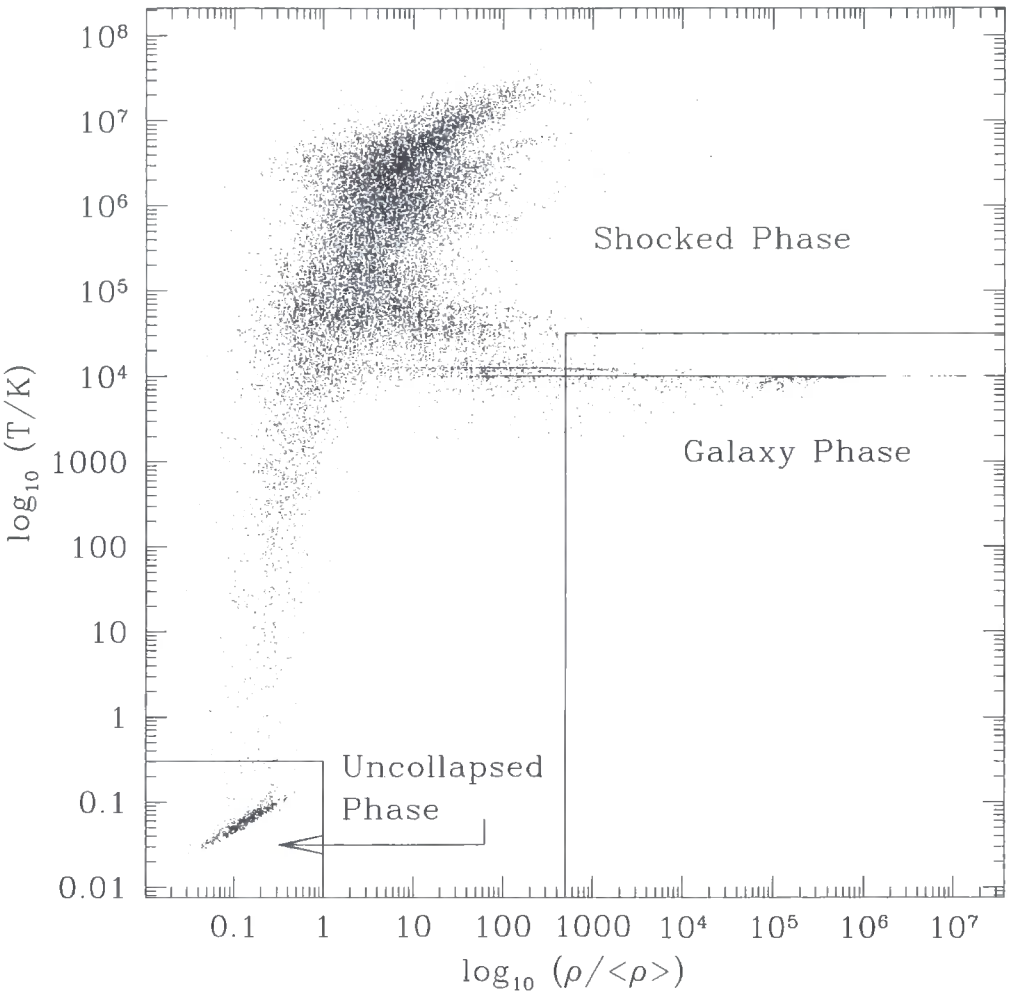


Figure 5.1: The $\rho - T$ phase diagram for baryons in run 3 at redshift zero. Density is in units of the mean value of the gas while temperature is in Kelvin. Boundaries for the three defined phases are also marked.

The gas mass resolution of each simulation is defined by the following parameter

$$\Re \equiv \frac{10^9}{m_{\rm gas}/{\rm M}_\odot}, \tag{5.1}$$

hence a gas particle with a mass of $10^9 {\rm M}_\odot$ will have a value of $\Re = 1$. Values of $\Re$ for each simulation are given in Column 8 of Table 5.1. For the first group, simulations with values of $\Re$ below unity have dark matter particle masses above the critical mass for significant two-body heating of the gas by the dark matter particles (Steinmetz & White 1997; see also Figure 4.8). Hence only runs 4 and 5 have dark matter particle masses *comfortably* below the critical value. Run 6 has the same critical mass as run 3, while for runs 7-12, the critical masses are smaller than their counterparts with the same number of gas particles, since the baryon fraction is higher.

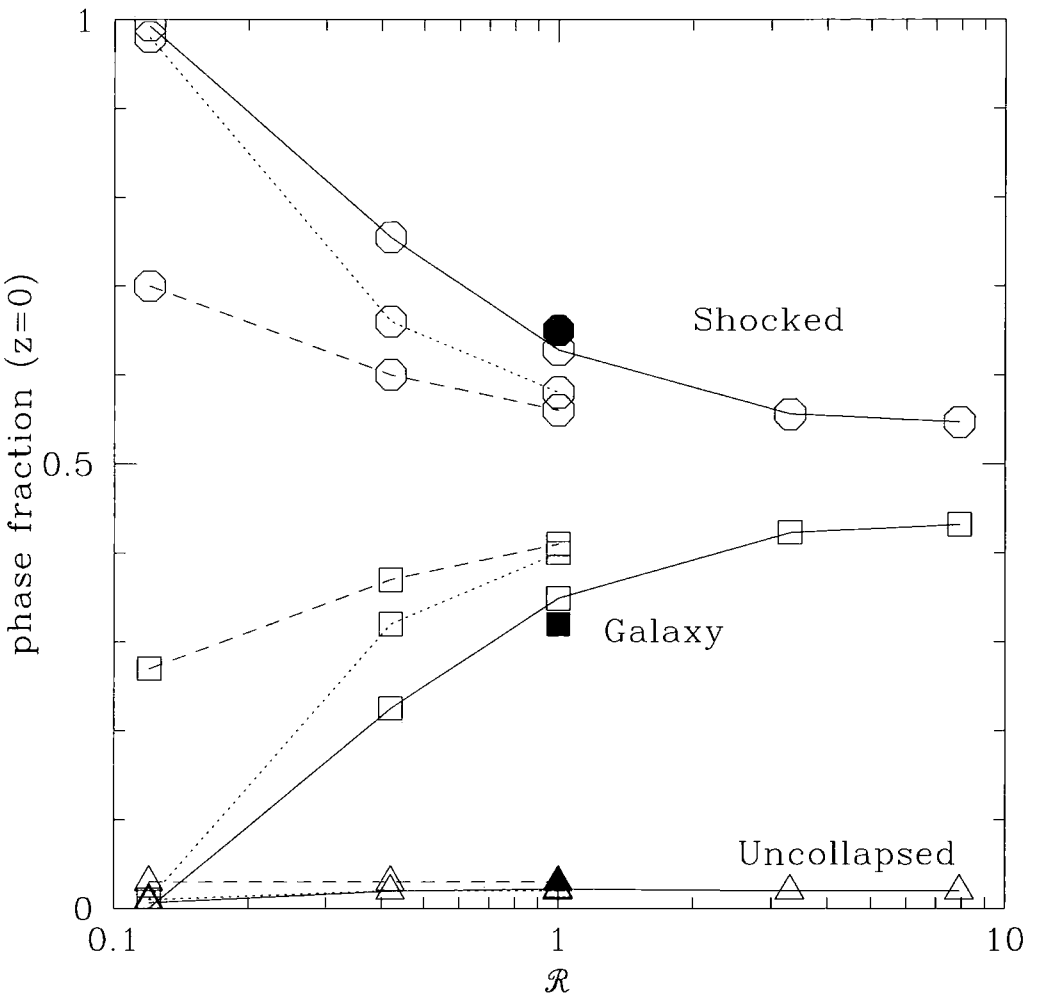


Figure 5.2: Fractions of baryons in the three phases at redshift zero as a function of the gas mass resolution. Values for the uncollapsed phase are plotted as triangles, the shocked phase as octagons and the galaxy phase as squares. The solid shapes are the results from run 6, while values from runs 1-5 are joined by solid lines, runs 7-9 are joined by dashed lines and runs 10-12 by dotted lines.

Furthermore, runs 7-9 have equal gas and dark matter particle masses, causing an additional increase in the value of the critical mass.

Columns 9-11 of Table 5.1 list the $z = 0$ gas fractions in the uncollapsed (f_u), shocked (f_s) and galaxy (f_g) phases. These are also illustrated in Figure 5.2, where values are plotted against $\Re$. Results for runs 1-5 are joined by solid lines, results for run 6 are shown as solid shapes and results for runs 7-9 and 10-12 are joined by dashed and dotted lines respectively. For uncollapsed gas, measured values are marked using triangles, shocked gas fractions are represented using octagons and galaxy phase gas fractions by squares.

The amount of uncollapsed gas is effectively constant for all runs, at 2 per cent of the total. This is far less than the amount of uncollapsed gas present in the fiducial simulation in Chapter 4 and is due to the different SPH implementations. The version of HYDRA used here has improved shock capturing through the use of a different algorithm for the artificial viscosity (see Thacker et al. 1998), leading to more efficient heating of the gas within haloes at the resolution limit of simulations.

Significant differences between the fractions in each phase are only evident for the shocked and galaxy phases, with the amounts controlled by the efficiency with which the gas can cool. Firstly, all of the runs have values of $f_{\rm g}$ that are significantly different from the amount of cold gas and stars observed in nearby galaxies. Ignoring runs 1 and 10 (with drastically low values), the minimum fraction of gas within the galaxy phase is 23 per cent. For an $\Omega = 1$ universe, the observed fraction is around a factor of two lower (Fukugita, Hogan, & Peebles 1998).

As expected (and indeed demonstrated for a smaller set of simulations in the previous chapter), the fraction of gas in the galaxy phase rises with resolution for all comparisons, at the expense of the shocked gas fraction. For the first comparison (runs 1-5) the amount of galaxy phase gas increases from zero at the lowest resolution to about 43 per cent for the highest resolution. Comparing run 6 to 3 ($\Re = 1$) where the size of the simulation volume differs by a factor of 8, the fractions of gas in each phase are within a few per cent of one another. Note that an intrinsic level of scatter will be present due to the small size of both volumes.

For runs 10-12 (i.e. the dotted lines in Figure 5.2) the fraction of galaxy phase gas is larger than for runs 1-3, with the difference increasing with resolution. The larger baryon fraction in the former simulations has enabled more gas to cool, as found in the previous chapter. Comparing runs 10-12 to runs 7-9 (where the gas and dark matter particles have equal masses), the difference between the galaxy phase fractions (comparing runs with equal $\Re$) decreases with increasing resolution. This can be explained as follows. For the smallest value of $\Re$ the amount of two-body heating of the gas by the dark matter particles is less severe in the run with equal mass particles. Thus significantly more gas has cooled in this case. As the resolution increases, two-body heating becomes less effective and the use of unequal mass particles becomes less important.

Another significant point of discussion regarding Figure 5.2 is the apparent convergence in the galaxy phase fraction (and also the shocked phase fraction) at the highest resolutions studied, in particular the two points with $\Re = 3.3$ and $\Re = 7.9$. The simplest explanation to the relatively small difference between these two values is that it reflects the additional mass in resolved haloes in the $\Re = 7.9$ simulation. Since the mean cooling time in haloes at the resolution limit of both

simulations (around 10 particles) is considerably smaller than the age of the Universe, it would be expected that the additional mass fraction reflects the additional amount of cooled gas (of course this argument neglects halo substructure). A useful gauge for calculating the mass in haloes is to use the Press-Schechter *ansatz* (Press & Schechter 1974). This predicts that only a further ~ 1 per cent of the total mass in the volume is resolved into haloes in the $\Re = 7.9$ simulation, in agreement with the additional fraction of gas in the galaxy phase. Hence, the apparent convergence is somewhat illusory, due the steep drop at lower resolutions where the level of artificial two-body heating becomes significant. A thoroughly consistent measurement of the galaxy phase fractions at both higher resolutions and larger simulation volumes will require a substantial increase in computational effort.

5.2.3 Evolutionary properties of the gas

The amount of gas in the uncollapsed, shocked and galaxy phases for runs 2-6 is presented in Figure 5.3, plotted against both time and redshift. Similarly to Section 4.3.3, the temperature threshold was modified to $T < 0.3(1+z)\mathrm{K}$, allowing for the adiabatic expansion of the uncollapsed gas. The evolution of the phases is clearly illustrated for $z < 10$ (the first output time of the simulations). All simulations plotted show qualitatively similar trends. All of the gas is initially uncollapsed, but more than half becomes shocked by $z = 5$, due to the formation of the first resolved structures. By this time, the fraction of gas that has cooled to high density (and hence enters the galaxy phase) starts to rise, continuing to do so until the present day. Both the shocked and uncollapsed phases are progressively depleted for the majority of the time. The former happens as more and more gas falls onto collapsed objects, however the consequent increase in the shocked gas is outweighed by the amount cooling into the galaxy phase.

Comparing this scenario (in particular, run 3 with $\Re = 1$) to Figure 4.3, there are apparent differences in the evolution of each phase. The enhanced shock capturing ability in the version of HYDRA used in this chapter increases the amount of gas in the shocked phase by more than 20 per cent by $z = 5$. This also caused the amount of galaxy phase gas to rise. As a consequence, the amount of uncollapsed gas is greatly reduced.

When the resolution is increased, no significant difference is made to the amount of uncollapsed gas at any time. The amount of gas within the shocked phase does not vary significantly until the gas begins to cool, causing the shocked fraction to peak and start decreasing around $z = 5$. Simulations with higher resolution cause this peak to move to lower values at earlier times, due to the presence of smaller denser haloes, within which the gas can cool more efficiently. After

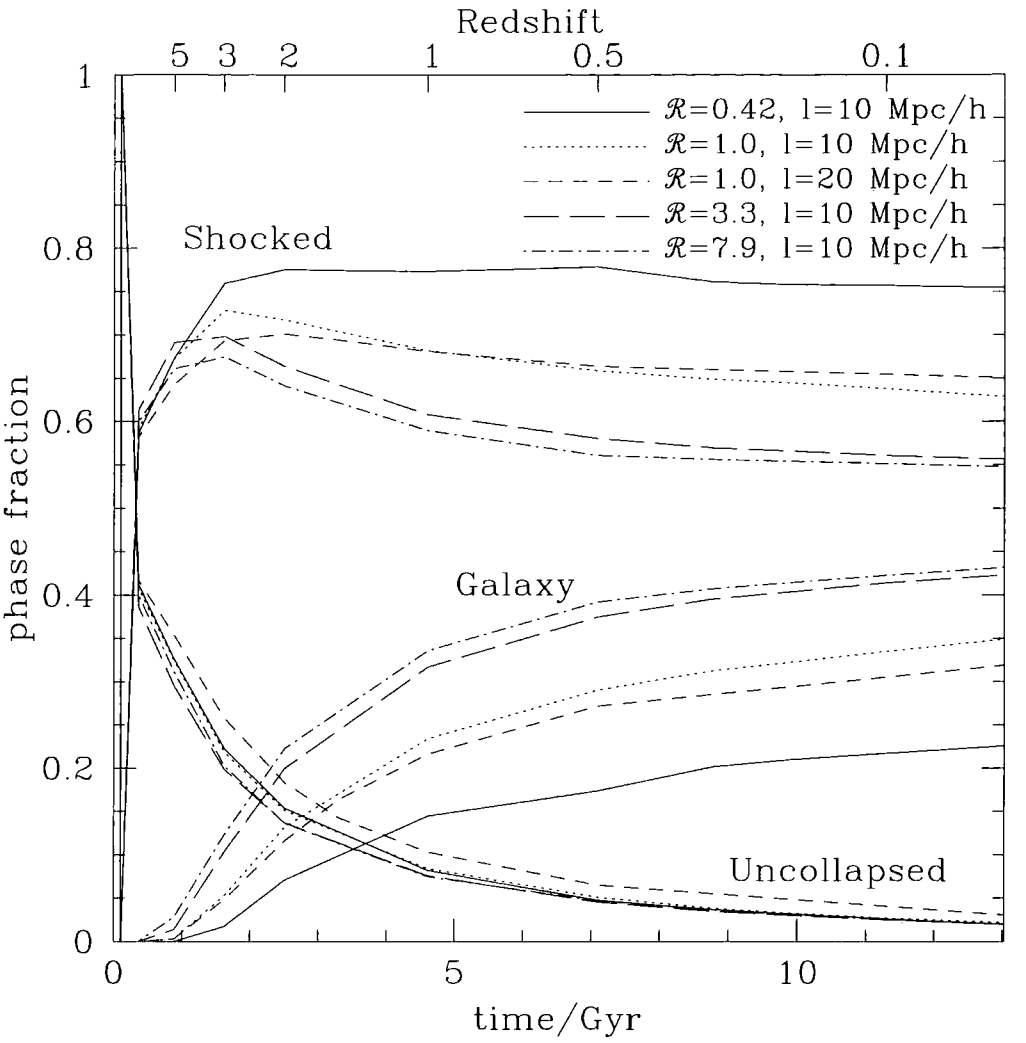


Figure 5.3: Values of the uncollapsed, shocked and galaxy phase fractions as a function of redshift for runs 1-6. Both time (in Gyr) and redshift are plotted along the horizontal axis.

$z \sim 1$, all of the simulations have galaxy phase fractions increasing more or less at the same rate. For the simulations with $\Re = 3.3$ and $\Re = 7.9$ (the two highest values), the evolution of the phases fractions are similar. The same can be said for the runs with $S = 10\,h^{-1}$ Mpc and $S = 20\,h^{-1}$ Mpc: only differences within a few percent are evident at all times.

The gas cooling rate is shown explicitly for runs 2-6 in Figure 5.4, again plotted against both time and redshift. All of the simulations plotted exhibit a strong peak in the cooling rate around $z \sim 3$, before an exponential-like decay until $z = 0$. Increasing the resolution moves the peak to higher values at earlier times, as more of the gas cools in the newly resolved objects. Doubling the box-size makes no significant difference to the cooling rate.

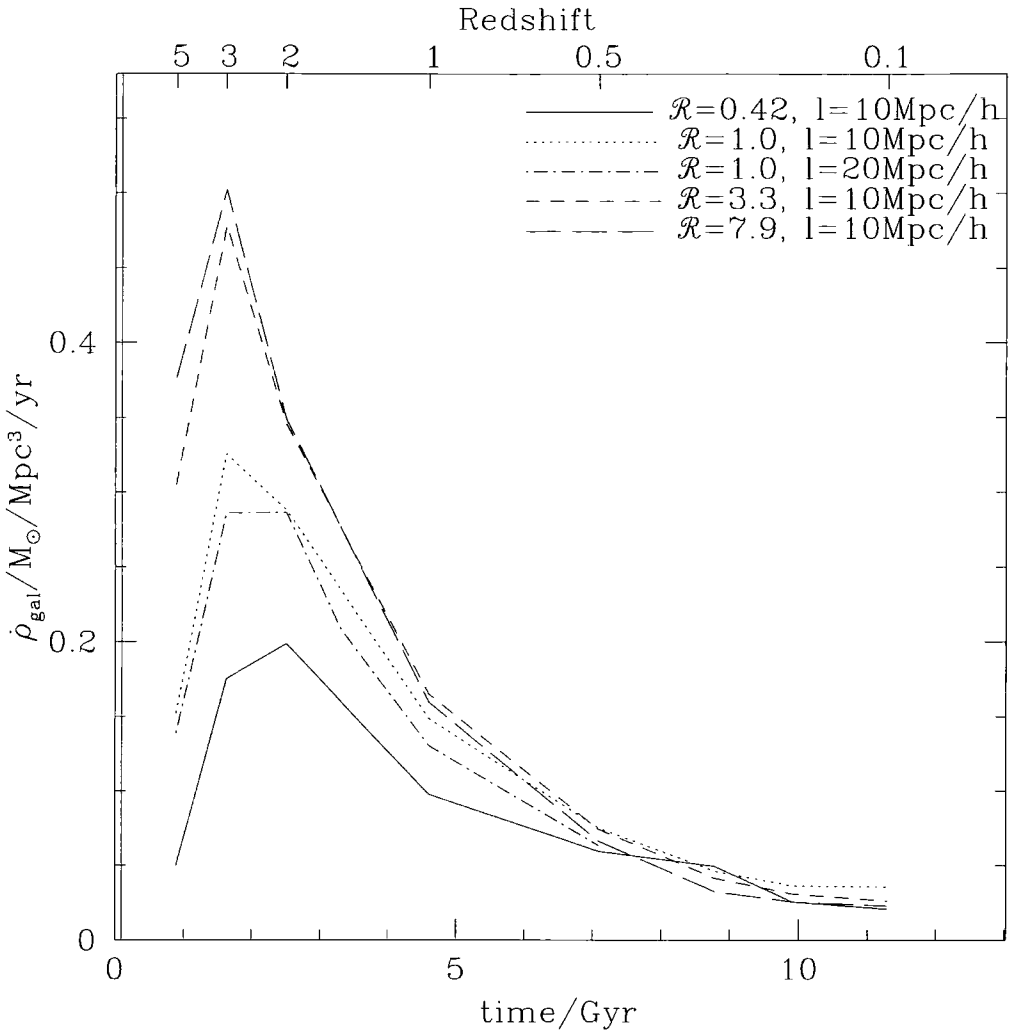


Figure 5.4: The gas cooling rate density (in $M_\odot$/ Mpc^3/yr) for runs 1-6. Time (in Gyr) and redshift are labelled on the horizontal axis.

5.2.4 Summary

In this section, several issues pertaining to the numerical method of cosmological simulations were examined, specifically to test the effects they have on the radiative cooling efficiency of the gas. Three sets of simulations were performed to analyse the effects of resolution, the size of the simulation volume and the ratio of gas to dark matter particle masses. As in Chapter 4, the gas was split into three phases: uncollapsed gas, shocked gas and galaxy phase gas (material that has cooled to high density).

Increasing the resolution of the simulation allows smaller objects to be resolved, in which more gas can cool at early times. Significantly more gas has cooled by $z = 0$ than is observed in the local universe for the majority of the simulations studied. The cooling rate of the gas is strongly

peaked around $z = 3$, moving to larger values and higher redshift for higher resolution. For the lowest resolution simulations studied, the rise in the fraction of cooled gas is steeper than it is for the highest resolution simulations. The former is within the regime where two-body heating of the gas by the dark matter particles is significant.

Increasing the simulation volume by a modest factor (doubling the length of the box) makes no significant differences to the amount of gas in each phase at any time, even though these values are affected by intrinsic scatter due to the small volumes that were used.

Finally, using a mass ratio of unity between the gas and dark matter particles showed that significantly more gas cooled in this case for the lowest resolution simulation than when equal numbers of gas and dark matter particles were used. As the resolution increases (and two-body heating effects become less severe) the results converge.

5.3 Star formation

In this section, various methods for turning gas into stars are compared to each other within the context of cosmological simulations. All of the star formation prescriptions follow the same basic procedure. At the end of each timestep, a subset of gas particles are identified that are eligible to form stars. A fraction of this gas is then converted into stars, determined by a local star formation rate. Subsequent evolution of the stars is then due to the force of gravity only.

To provide a fair comparison, identical initial conditions to run 3 of Section 5.2 were used for every simulation in this study. Although the resolution of the simulations is modest, it allowed the necessary comparative set of tests to be completed on a workstation within a reasonable timescale. The effects of resolution will be investigated in future work.

5.3.1 Overview of prescriptions

5.3.1.1 The Pearce method

The first two simulations studied, which will hereafter be referred to as PEARCE1 and PEARCE2, use the $\rho + T$ method implemented by Pearce (1998), who also used it to study the effects of star formation within cosmological simulations. Gas particles are eligible to form stars if they satisfy an overdensity criterion, $\delta > \delta_*$ and a temperature criterion, $T < T_*$. The critical temperature is $T_* = 10^5$K and the critical overdensity is $\delta_* = 5000$ for PEARCE1 and $\delta_* = 50000$ for PEARCE2. (Note that the choice of parameters isolates material that must have radiatively cooled, falling within the galaxy phase defined in the previous section.) A star formation rate is not imposed, rather every

gas particle that satisfies the above criteria is *instantaneously* converted into a collisionless star particle

The overdensity threshold chosen by Pearce (1998) was $\delta_* = 550$, who justified this value by arguing that the fraction of gas converted into stars by $z = 0$ was close to the observed value. For the parameters used in his simulations, the threshold is only a factor of a few smaller than the maximum resolved gas density. (This can be estimated by considering a gas particle surrounded by a sphere of radius $2h_{\rm min}$, which encloses a further $N_{\rm SPH}$ neighbours.) In this study, the effective maximum density is around the value chosen for PEARCE2; PEARCE1 used a threshold 10 times smaller.

5.3.1.2 The hmin method

The third simulation (HMIN) used an even simpler prescription. As in the Pearce method, gas particles are identified with $T < 10^5$K, as well as those with SPH smoothing lengths equal to the minimum value ($h_{\rm min} = s/2$, where s is the gravitational softening length), and converted into stars. The only additional free parameter (when compared to a simulation with no star formation) is the temperature threshold. The smoothing length constraint acts like a density threshold: particles with $h = h_{\rm min}$ will have densities around the maximum resolved by the SPH algorithm. Since $h_{\rm min}$ evolves proportional to the gravitational softening, at times before $z \sim 1$ it acts like an overdensity threshold (i.e. fixed in comoving co-ordinates) and afterwards a physical density threshold.

5.3.1.3 The Summers method

The fourth simulation, labelled SUMMERS1, used the method detailed by Summers (1994), who also studied the effects of star formation within cosmological simulations. In his method, gas particles are instantaneously converted into stars if they satisfy a temperature criterion, $T < 3 \times 10^4$K and a physical (number) density criterion, $n > 0.1\,{\rm cm}^{-3}$. Additionally, eligible gas particles must also be within overdense regions, $\delta > 10$, so that star formation does not occur within shallow potentials at high redshift, when the mean density of the universe is much higher ($\langle\rho\rangle \propto (1+z)^3$).

The density threshold was chosen by Summers since he argued its value is around the minimum where star formation is observed to occur within disk galaxies. For the simulations performed in this study, the threshold corresponds to a baryon overdensity, $\delta \sim 3.5 \times 10^5$ at $z = 0$ (assuming a gas of primordial composition with $\mu = 0.59$). This is around an order of magnitude higher than the maximum density reliably resolved by the SPH algorithm, for the simulations studied in this chapter. This prompted the study of an additional simulation (SUMMERS2), performed with the lower threshold, $n > 0.01\,{\rm cm}^{-3}$.

5.3.1.4 The Schmidt method

The sixth simulation carried out (labelled SCHMIDT) used a prescription based on the Schmidt law (Schmidt 1959; see also Kennicutt 1998 and references therein), a parameterisation of the density dependence of the star formation rate, inferred from the atomic gas distributions of nearby galaxies

$$\Sigma_{\rm SFR} = A\Sigma_{\rm gas}^{N}, \tag{5.2}$$

where $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ are the surface gas and star formation rate densities, A is a normalisation constant, controlling the overall efficiency of star formation and N is typically 3/2. For this study, Equation 5.2 is converted into a relation for the mass density of the gas following the work of Mihos & Hernquist (1994), who used it to study the star formation properties within disk galaxies

$$\frac{\dot{M}_{\rm gas}}{M_{\rm gas}} = A\,\rho_{\rm gas}^{1/2}, \tag{5.3}$$

where $\dot{M}_{\rm gas}/M_{\rm gas}$ is the mass fraction of gas converted into stars per unit time interval. Gas particles are eligible to form stars if they are within overdense regions ($\delta > 10$, as in the Summers method) and a convergent flow, $\nabla.\mathbf{v} < \mathbf{0}$, where v is the peculiar velocity of a gas particle. Equation 5.3 is applied to each eligible particle by calculating the fraction of its mass that would form stars over the timestep, $p = A\,(\rho/\langle\rho\rangle)^{1/2}\Delta t$. If the condition, $r/p < 1$ (where r is a random number drawn between 0 and 1) is satisfied, the gas particle is converted into a star particle. For this study, a normalisation constant of $A = 5 \times 10^{-4}\,{\rm cm}^{3/2}{\rm g}^{-1/2}{\rm s}^{-1}$ was adopted, inferring a star formation rate of around $2{\rm M}_{\odot}{\rm yr}^{-1}$ at an overdensity of $\sim 5 \times 10^{4}$.

5.3.1.5 The Navarro method

The seventh simulation, labelled NAVARRO1, the method of Navarro & White (1993) was adopted, who used it in their study of the formation and evolution of a centrifugally supported disk from a rotating isothermal sphere. This is a popular method that has been used by several other groups with minor modifications (e.g. Steinmetz & Müller 1995; Tissera, Lambas, & Abadi 1997; Carraro, Lia, & Chiosi 1998). Gas particles are eligible to form stars if they are within a convergent flow ($\nabla.\mathbf{v} < \mathbf{0}$) and have cooling times shorter than the local dynamical time

$$\tau_{\rm cool} = \frac{3\rho_{\rm gas}kT}{2\mu m_{\rm p}\Lambda(T)} < \tau_{\rm dyn} = \sqrt{\frac{3\pi}{16G\rho_{\rm gas}}}, \tag{5.4}$$

where $\Lambda(T)$ is the cooling function. Navarro & White note that Equation 5.4 is readily satisfied providing the gas density exceeds a critical value, which, for a primordial composition is $\rho > 7 \times 10^{-26}\,{\rm g\,cm}^{-3}$. (For this study, the same density constraint was applied, even though the gas was assumed to have a fixed metallicity of $Z = 0.5Z_{\odot}$. Other authors, for example

Carraro, Lia, & Chiosi 1998 prefer to use Equation 5.4, which explicitly includes the metallicity of the gas.) The density threshold is almost identical to that used in SUMMERS1 ($n = 0.1\,\mathrm{cm}^{-3}$ is approximately equivalent to $\rho = 10^{-25}\,\mathrm{g\,cm}^{-3}$).

Once a particle has been identified as eligible for star formation, its star formation rate is calculated by the formula,

$$\frac{\mathrm{d}\rho_{\mathrm{gas}}}{\mathrm{dt}} = -\frac{\mathrm{d}\rho_*}{\mathrm{dt}} = -\frac{c_*\rho_{\mathrm{gas}}}{t_*}, \tag{5.5}$$

where $t_* = \tau_{\mathrm{dyn}}$ and $c_* = 1$. Stars are formed in quantized masses of factors of two, hence the gas density halves in the time interval, $\Delta t_* = \ln 2 t_*$. At a time Δt_* after the gas particle is eligible or has last formed stars, its mass is halved into one gas and one star particle, which both inherit the phase space co-ordinates of their parents. When the mass of the gas is less than 5 per cent of its original mass, the remaining fraction is converted into stars.

The prescription was repeated for an additional simulation (NAVARRO2), with one exception. Rather than create new particle masses, gas particles were converted into stars determined by the rate equation

$$p = 1 - \exp\left(\frac{-c_*\Delta t}{t_*}\right). \tag{5.6}$$

As in SCHMIDT, this is applied to each particle as a probability by drawing a random number between 0 and 1; the particle is converted into a star particle if the ratio $r/p < 1$.

For simulation 9 (NAVARRO3), the previous method was repeated with a critical density of $\rho > 10^{-26}\,\mathrm{g\,cm}^{-3}$, equivalent to that used in SUMMERS2.

5.3.1.6 The Katz method

The tenth simulation, labelled KATZ1, replicates the star formation method used in the cosmological simulations studied by Katz, Weinberg & Hernquist (1996; see also Katz 1992). Four criteria are used to determine whether a gas particle is eligible to form stars: the divergence criterion, as used in SCHMIDT; the physical density threshold used in SUMMERS1; an overdensity threshold, $\delta > 55.7$ and a "Jeans instability" condition. The Jeans condition is satisfied if

$$\frac{h}{c_s} > \frac{1}{\sqrt{4\pi G\rho_{\mathrm{gas}}}}, \tag{5.7}$$

where h is the particle's SPH smoothing length and c_s is the local sound-speed of the gas. This condition depends on the resolution of the simulation since h is limited by the gravitational softening.

If a gas particle is eligible to form stars, the same star formation rate is applied as in NAVARRO1, except that $t_* = \max(\tau_{\mathrm{dyn}}, \tau_{\mathrm{cool}})$. For particles with $T < 30000\mathrm{K}$, $t_* = \tau_{\mathrm{dyn}}$, since the cooling

Table 5.2. Details of the star formation prescriptions tested in this chapter. Column 1 lists the simulations performed while Columns 2 and 4 give details of the star formation methods, namely the eligibility criterion and the mechanism for converting gas into stars (details in the text).

Simulation	Eligibility Criteria	Conversion Mechanism
PEARCE1	$\delta > 5000$; $T < 10^5$K	instantaneous + conversion
PEARCE2	$\delta > 50000$; $T < 10^5$K	instantaneous + conversion
HMIN	$h \leq h_{\rm min}$; $T < 10^5$K	instantaneous + conversion
SUMMERS1	$\delta > 10$; $n > 0.1\,{\rm cm}^{-3}$; $T < 3 \times 10^4$K	instantaneous + conversion
SUMMERS2	$\delta > 10$; $n > 0.01\,{\rm cm}^{-3}$; $T < 3 \times 10^4$K	instantaneous + conversion
SCHMIDT	$\nabla.\mathbf{v} < 0$; $\delta > 10$	s.f. rate + conversion
NAVARRO1	$\nabla.\mathbf{v} < 0$; $\rho > 7 \times 10^{-26}\,{\rm g\,cm}^{-3}$	s.f. rate + creation
NAVARRO2	$\nabla.\mathbf{v} < 0$; $\rho > 7 \times 10^{-26}\,{\rm g\,cm}^{-3}$	s.f. rate + conversion
NAVARRO3	$\nabla.\mathbf{v} < 0$; $\rho > 10^{-26}\,{\rm g\,cm}^{-3}$	s.f. rate + conversion
KATZ1	$\nabla.\mathbf{v} < 0$; $\delta > 10$; $n > 0.1\,{\rm cm}^{-3}$; $\tau_{\rm J} > \tau_{\rm dyn}$	s.f. rate + dual identity
KATZ2	$\nabla.\mathbf{v} < 0$; $\delta > 10$; $n > 0.1\,{\rm cm}^{-3}$; $\tau_{\rm J} > \tau_{\rm dyn}$	s.f. rate + conversion
KATZ3	$\nabla.\mathbf{v} < 0$; $\delta > 10$; $n > 0.01\,{\rm cm}^{-3}$; $\tau_{\rm J} > \tau_{\rm dyn}$	s.f. rate + conversion

time becomes longer than the age of the Universe as $T \rightarrow 10^4$K. Stars are created if $r/p < 1$, using Equation 5.6. Rather than convert whole gas particles into star particles, this method gives each particle a dual identity (the so-called *Schizophrenic* method ; see also Mihos & Hernquist 1994), by allowing it to carry both a gas mass and a stellar mass. For every gas or dual particle satisfying $r/p < 1$, a fraction, $\epsilon_* = 1/3$ is converted into collisionless material. Gravitational forces are calculated using the total mass and hydrodynamical forces using only the gas mass. Once the gas mass has dropped below 5 per cent of its original mass, the particle becomes purely collisionless.

Simulation 10 (KATZ2) used the previous method with $\epsilon_* = 1$. Hence at any time, particles were either gas or stars but not both. The final simulation run (KATZ3) was identical to KATZ2 except that the density threshold was lowered to that used in SUMMERS2, i.e. $n > 0.01\,{\rm g\,cm}^{-3}$.

5.3.1.7 Summary

The various star formation prescriptions are summarised in Table 5.2. Column 1 lists the simulations run and columns 2 & 3 give the eligibility criteria and mechanisms for incorporating stars.

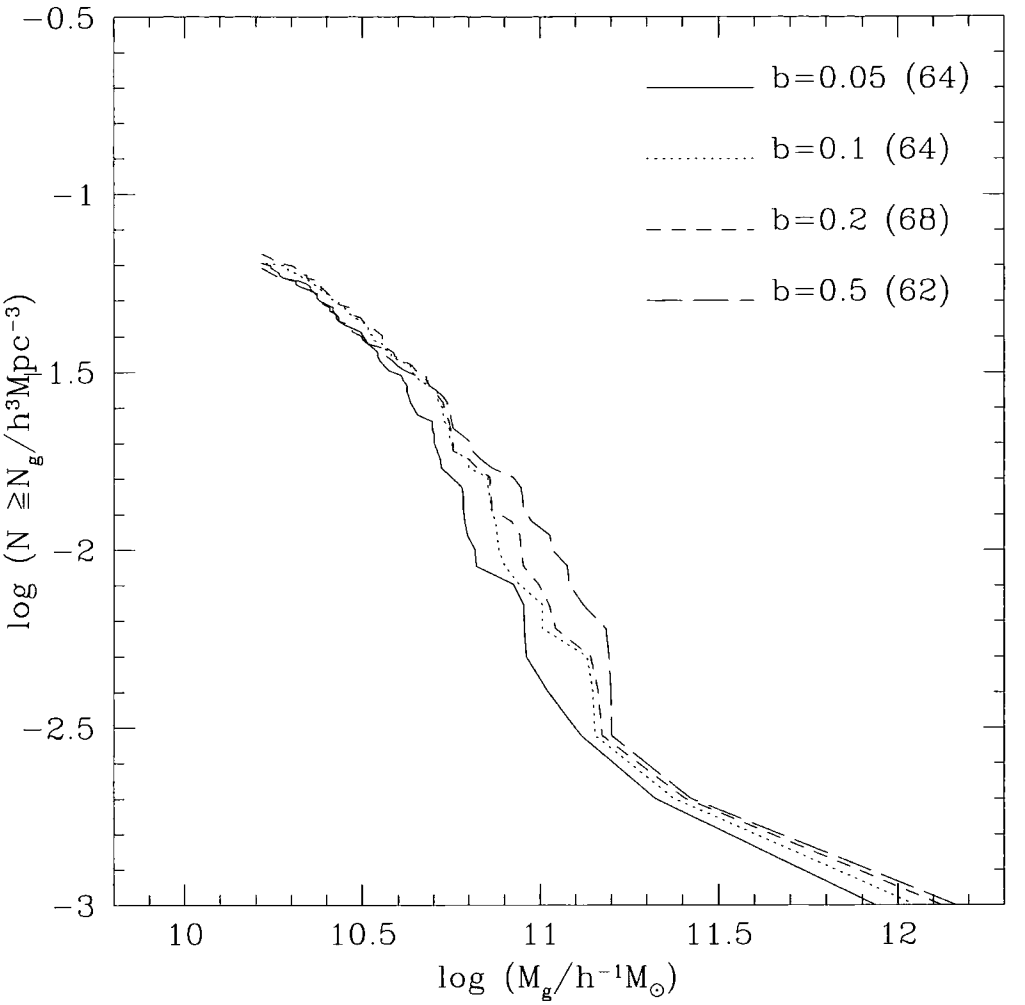


Figure 5.5: Cumulative mass functions of the galaxy population for different values of the (dimensionless) linking length, b, for PEARCE1. Numbers in brackets correspond to the total number of galaxies found by the group-finder.

5.3.2 Galaxy selection

Galaxies were selected by extracting a subset of particles consisting of all star particles plus the gas particles with $\rho > 500\,\langle\rho\rangle$ and $T < 3 \times 10^4$K (these limits were used to define the galaxy phase Section 5.2). This subset was then passed through a friends-of-friends group-finder, linking particles together within a fixed separation, parameterised by the dimensionless parameter, b (c.f. Section 4.3.4). Any linked objects with constituent particle numbers smaller than $N_{\rm SPH}$ (i.e. 32) were discarded, producing the final galaxy catalogue.

Figure 5.5 illustrates the cumulative mass function of galaxies at $z = 0$ in PEARCE1 for $b = (0.05, 0.1, 0.2, 0.5)$. The number of galaxies in each catalogue is printed in the top-left of the figure. Unlike simulations with no star formation, the mass functions are sensitive to the choice

of b. The lack of dissipation in stellar material makes them more susceptible to disruption by tidal forces and mergers, producing objects that are less bound than their purely gaseous counterparts. To some extent this problem is numerical, since the potential wells of the galaxies are artificially shallow due to the large softening length employed. Hence, higher resolution simulations should at least partly alleviate the problem.

One method for choosing b is that which maximises the number of galaxies in the catalogue. However this procedure must be treated with caution, since distinct objects within close proximity may be artificially linked together, even though the total number of galaxies increases. For this study, a value of $b = 0.1$ was chosen, in agreement with that used in Chapter 4. The number of galaxies in this catalogue is around 6 per cent less than the maximum value plotted in Figure 5.5.

5.3.3 Results

5.3.3.1 General properties at $z = 0$

Table 5.3 lists some general properties of the twelve simulations at $z = 0$, plus the equivalent simulation without star formation (run 3 of Section 5.2; hereafter referred to as GAS). Columns 2 and 3 list measurements of the efficiency with which each simulation was carried out. The former lists the number of timesteps taken to complete the simulation (from $z = 24$ to $z = 0$), while the latter gives the amount of CPU time used for each run, relative to GAS. To provide a fair comparison, the machine dependent clock rate was removed from each time by dividing the total time taken by the average time for the first ten steps.

The least efficient method is NAVARRO1, taking around a third longer to complete than GAS. The increase in the number of particles due to star formation increases the computational time spent calculating gravitational forces. The rest of the simulations all take around half of the time to complete than GAS, with the shortest being PEARCE1, with one third.

Column 5 gives the fraction of baryons (gas+stars) in the galaxy phase, $f_{\rm g}$, at $z = 0$. For all simulations, including the run with no star formation, $f_{\rm g} = 0.34 \pm 0.01$. As was found by Pearce (1998), the conversion of cold dense gas into stars does not significantly affect the cooling rate of the gas.

In Column 6, the fraction of baryons in stars, $f_{\rm stars}$, is listed for each run. The range is much wider for $f_{\rm stars}$ than it is for $f_{\rm g}$, varying from 18 per cent (SUMMERS1) to 28 per cent (PEARCE1). Increasing the overdensity threshold from 5000 (PEARCE1) to 50000 (PEARCE2) decreased the fraction in stars by 0.03, a relative change of $\sim$ 10 per cent. When gas particles initially cool to 10^4K, the majority of them have densities below 500 $\langle\rho\rangle$ (the minimum density of the galaxy

Table 5.3. General properties of the simulations at $z = 0$. Column 1 lists the simulation; Column 2 gives the number of timesteps taken to complete each simulation; Column 3 lists the total CPU time used (relative units); Column 4 gives the fraction of baryons in the galaxy phase; Column 5 gives the fraction of baryons in stars and Column 6 lists the number of galaxies in each simulation.

Simulation	$N_{\rm steps}$	CPU	$f_{\rm g}$	$f_{\rm stars}$	$N_{\rm gal}$
GAS	4366	1.00	0.35	0.00	74
PEARCE1	2427	0.32	0.33	0.28	64
PEARCE2	3581	0.51	0.34	0.25	70
HMIN	2804	0.38	0.34	0.26	68
SUMMERS1	3732	0.56	0.35	0.18	70
SUMMERS2	2626	0.37	0.34	0.26	66
SCHMIDT	3227	0.48	0.35	0.24	64
NAVARRO1	4686	1.33	0.35	0.19	69
NAVARRO2	3864	0.49	0.35	0.18	71
NAVARRO3	3318	0.47	0.34	0.25	70
KATZ1	4223	0.67	0.35	0.18	71
KATZ2	3975	0.63	0.35	0.18	75
KATZ3	3095	0.43	0.34	0.25	78

phase). However, the sudden lack of pressure support causes regions of cooled gas to collapse further under gravity, increasing in density (as illustrated in Figure 4.4). Eventually the gas will cross the density threshold for star formation and a higher threshold lengthens the time taken for a given particle to make this transition. HMIN has a similar stellar fraction to the previous two simulations. This is of no surprise since the smoothing length criterion acts like a density threshold, roughly equivalent to $\delta = 50000$ at $z = 0$. Since $h_{\rm min}$ varies as the gravitational softening, the threshold was initially ~ 8 times lower until $z = 1$, where it increased to the present value.

A considerable difference in $f_{\rm stars}$ is evident between SUMMERS1 and SUMMERS2. The latter run has approximately the same fraction of stars as in the previous simulations, while the former run has $f_{\rm stars} = 0.18$, a relative decrease of about 30 per cent. As was the case for PEARCE1 and PEARCE2, the difference can be explained by the choice of density threshold - the value of the threshold in SUMMERS1 (at $z = 0$) is around an order of magnitude higher than for PEARCE2. Furthermore, this threshold is beyond the densities that the SPH algorithm can reliably resolve, and so the evolution of particles in this regime are purely determined by numerics. The density

threshold is fixed in physical co-ordinates, so at higher redshift, the equivalent overdensity is lower, decreasing proportional to $(1+z)^3$. At $z = 3$ the threshold is around 5500, similar to PEARCE1. Similarly for SUMMERS2, the $z = 0$ overdensity threshold is approximately equivalent to PEARCE2, decreasing to ~ 550 at $z = 3$.

For SCHMIDT, the method of converting gas into stars is qualitatively different from the previous simulations since the probability of star formation at fixed density is non-zero for all gas particles, so long as they are within an overdense, converging region (as is the case when they are cooling within haloes). The number of stars at any given time is controlled by the value of the normalisation constant, A. The choice of A ($5 \times 10^{-4}\,\mathrm{cm}^{3/2}\mathrm{g}^{-1/2}\mathrm{s}^{-1}$) was made purely by trial and error, so that the fraction of stars at $z = 0$ agreed with that in PEARCE2.

NAVARRO1 and NAVARRO2 produce very similar stellar fractions even though the first creates stars as separate particles. Both simulations produce the same value of f_{stars} as SUMMERS1, while NAVARRO3 has value as SUMMERS2. Again, the choice of density threshold has determined the fraction of cooled material in stars at $z = 0$. The use of a divergence criterion (rather than a temperature threshold) as well as a star formation rate made no difference whatsoever.

The methods of KATZ1 and KATZ2 produce the same fraction of stars, even though the former uses particles representing both gas and stars. Both these runs produced the same value as in NAVARRO1 and NAVARRO2, hence the "Jeans condition" has no significant effect on the amount of star formation that has occurred by $z = 0$. For gas particles with minimum smoothing lengths, where the condition is most stringent, the density threshold is always higher than the minimum density required to satisfy the Jeans condition. The same is true for KATZ3, which produces the same fraction of stars as NAVARRO3.

The last column in Table 5.3 lists the number of galaxies in each catalogue. All runs contain ~ 70 galaxies at $z = 0$ to within 10 per cent. There are no obvious systematic trends in the number of galaxies with other general properties. For example, a higher density threshold produces a larger number of galaxies when comparing SUMMERS1 to SUMMERS2. However, the converse is true when comparing KATZ2 to KATZ3.

5.3.3.2 Star formation rates

For every star formation event in the simulations, the time and mass converted per particle was stored, allowing complete star formation histories to be constructed. These are illustrated in Figure 5.6, plotted as star formation rate densities (in $\mathrm{M}_\odot/\,\mathrm{Mpc}^3/\mathrm{yr}$) against time, in Gyr. To reduce noise, the data was binned, with each bin 0.2 Gyr wide.

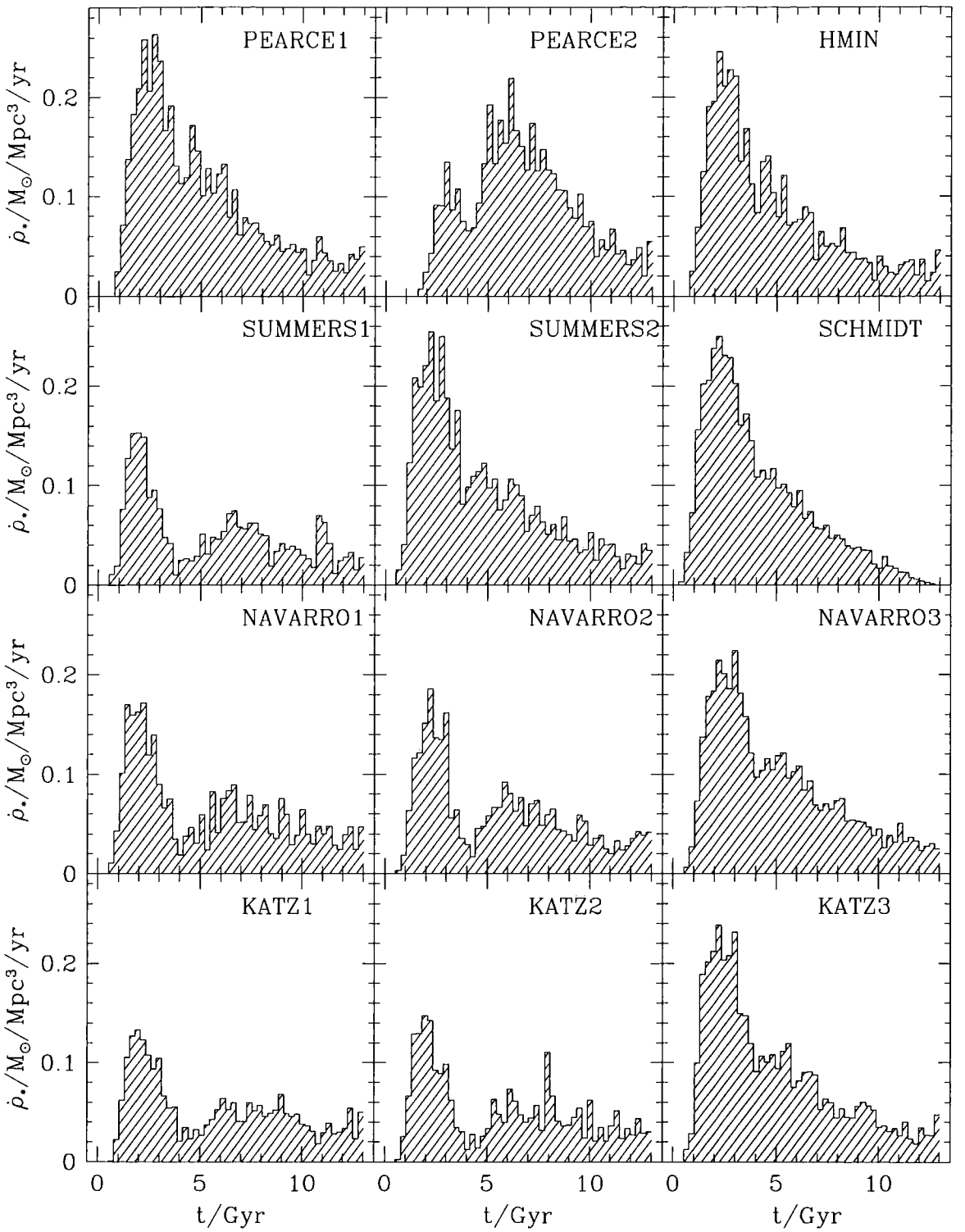


Figure 5.6: Star formation rate densities (in $M_\odot$/ Mpc^3/yr) for all 12 simulations plotted against time, in Gyr. Each bin is 0.2 Gyr wide.

All simulations begun forming stars within the first Gyr or so. The exact value depends on two factors: the resolution of the simulation, which determines the time when the first haloes collapse, and the range of densities where star formation efficiently occurs. Since the resolution was fixed for all simulations studied, differences in the initial time are due to the second factor. Note, however, that increasing the resolution of the simulations will allow smaller mass haloes to collapse at earlier times, transferring more gas from the shocked phase to the galaxy phase (as discussed in Section 5.2). Since these haloes will have, on average, higher densities, the amount of star formation will also increase.

The star formation rate of PEARCE1 rises sharply at $t \sim 1$Gyr, peaks at $t \sim 2 - 3$Gyr, before decaying for ~ 10Gyr to the present day. Star formation was still occurring at the end of the simulation. The shape of the curve is very similar to the cooling rates illustrated in Figure 5.4, hence a substantial fraction of the gas was rapidly converted into stars once it had cooled. Contrary to this, PEARCE2 shows a more *hump-like* function, with the bulk of the star forming activity occurring at later times (the main peak is around 6 Gyr). The area under the curve is also smaller than for PEARCE1 as it is proportional to the value of $f_{\rm stars}$. As discussed earlier, the higher overdensity threshold in this simulation induces a delay in the conversion of many of the cooled gas particles into stars. This also acts to reduce the amount of stars at any given time. The depression in the PEARCE2 star formation rate at $t \sim 3$Gyr is a statistical fluctuation, due to the small box-size employed. The majority of the gas that initially cooled within the simulation (forming the peak of the cooling rate at $z \sim 3$) is related to the formation of a single object, the largest halo in the volume at $z = 0$. For PEARCE2, a substantial fraction of this gas did not form stars until a time after the formation of other haloes, causing a brief drop in the star formation rate.

The star formation rate of HMIN is similar to PEARCE1 even though the fraction of baryons in stars by $z = 0$ is the same as PEARCE2. Before $z = 1$ the value of $h_{\rm min}$ is equivalent to an overdensity threshold similar to the value used in PEARCE1, hence the star formation rates are similar at these times (approximately $t < 5$Gyr). After $z = 1$ the value of $h_{\rm min}$ shrinks proportional to $(1 + z)$, increasingly restricting star formation until $z = 0$, where its value is equivalent to the threshold used in PEARCE2.

The high physical density threshold used in SUMMERS1 considerably suppresses the star formation rate. The peak is evident, followed by a depression and a smaller peak at later times. Similarly to PEARCE2, the two peaks are caused by the delay converting cooled gas into stars. However, the first peak is stronger than the second in SUMMERS1 due to the threshold being constant in physical

co-ordinates, allowing more stars to form at earlier times. In SUMMERS2, where the threshold is 10 times lower, the familiar shape of the star formation rate is recovered.

SCHMIDT used a continuous star formation rate over the range of densities relevant to overdense, converging regions. Again, much of the stars formed at early times, although the star formation rate drops to zero by the present. All gas at high densities (where the rate is highest) was used up and the maximum density of available gas determined a star formation rate that became vanishingly small.

Simulations NAVARRO1 and NAVARRO2 exhibit very similar histories, demonstrating that the spawning of new particles makes no significant difference to the average star formation rate. The high physical density threshold in both simulations produces a similar rate to SUMMERS1. The same applies to NAVARRO2 and SUMMMERS2, albeit using a threshold that was around an order of magnitude smaller.

KATZ1 and KATZ2 also have similar star formation rates, showing that the use of particles with dual masses makes no difference to the average star formation rate. Again, these two simulations and KATZ3 show similar results to those with the same choice of density threshold.

5.3.3.3 Galaxy mass functions

Mass functions for the galaxy population at $z = 0$ are given in Figure 5.7. Due to the small numbers of galaxies in the catalogues, the functions are shown in cumulative form. For comparison, the mass function of GAS is displayed in all panels. All mass functions are in extremely good agreement with each other, demonstrating that these simulations are insensitive to any changes brought on by including star formation.

The mass functions are also plotted with the cold gas subtracted from each object in Figure 5.8. Again, GAS is shown in all panels. The functions for all simulations with star formation are lower than the GAS mass function at all masses, with the deviation increasing in magnitude as the mass decreases. On average, galaxies contain higher gas to stars ratios at lower masses than they do at higher masses. The merging process, prominent in all hierarchical theories, leads to the trend of larger mass objects being older. The number of galaxies with a fixed stellar mass varies strongly with the choice of density threshold employed. Simulations with the highest thresholds (and hence lowest values of $f_{\rm stars}$) show the largest deviations from the total mass function. The only exception to this is SCHMIDT, showing significantly larger deviations even though it has a comparable value of $f_{\rm stars}$. This simulation suffered from more disruption than the others, with only 75 per cent of the stars in galaxies where the others have typically 90-95 per cent. Doubling the linking length increased the former only by $\sim$ 5 per cent. NAVARRO2 and NAVARRO3 have

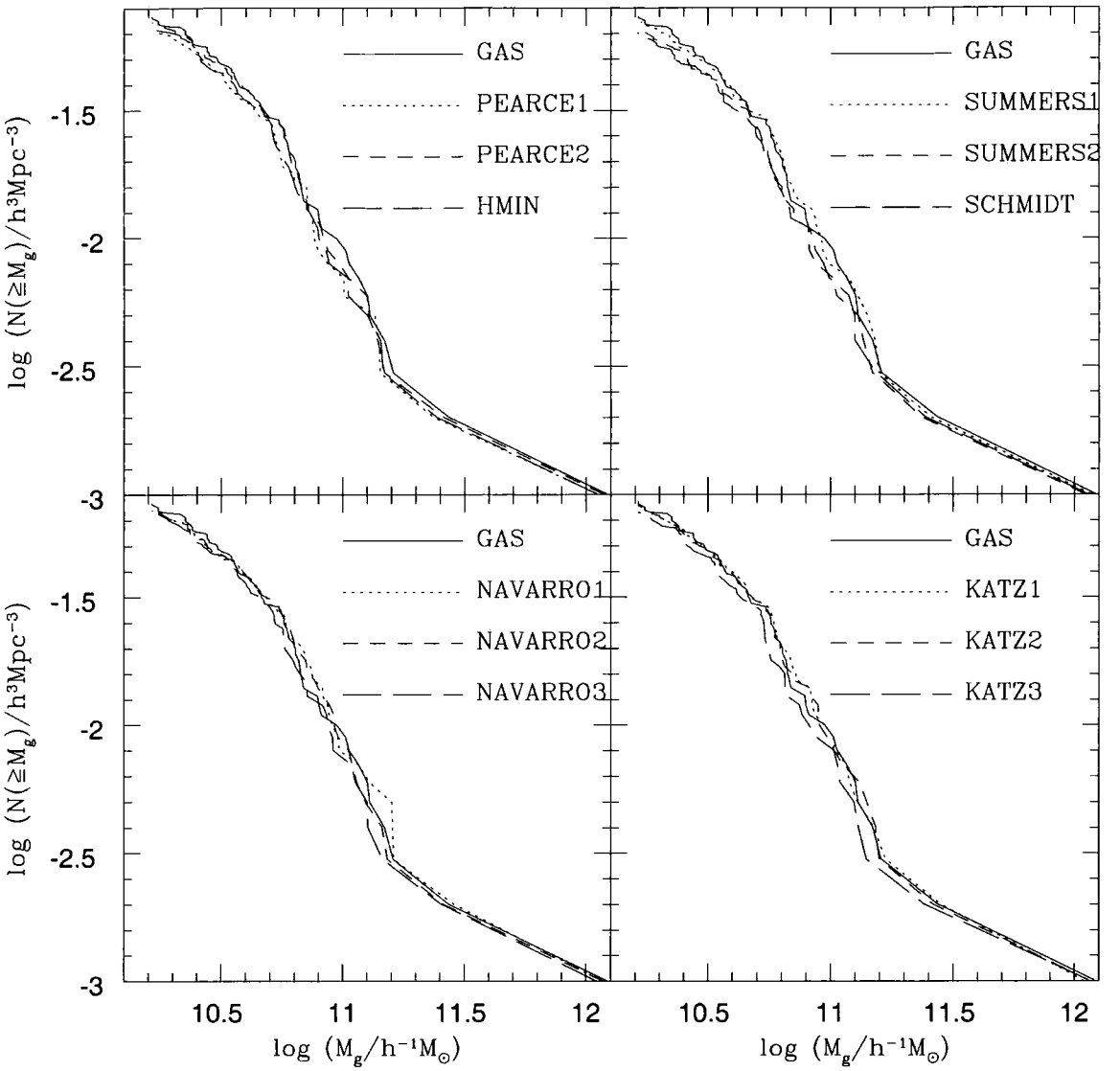


Figure 5.7: Cumulative mass functions of the galaxies (cold gas plus stars) at $z = 0$ for simulations with star formation. The mass function for GAS is also plotted in each panel.

very similar functions, even though their conversion mechanisms are different; the same is the case for KATZ2 and KATZ3.

5.3.4 Summary

In this section, various prescriptions for turning gas into stars within cosmological simulations were studied and compared to an identical simulation without star formation (run 3 of Section 5.2).

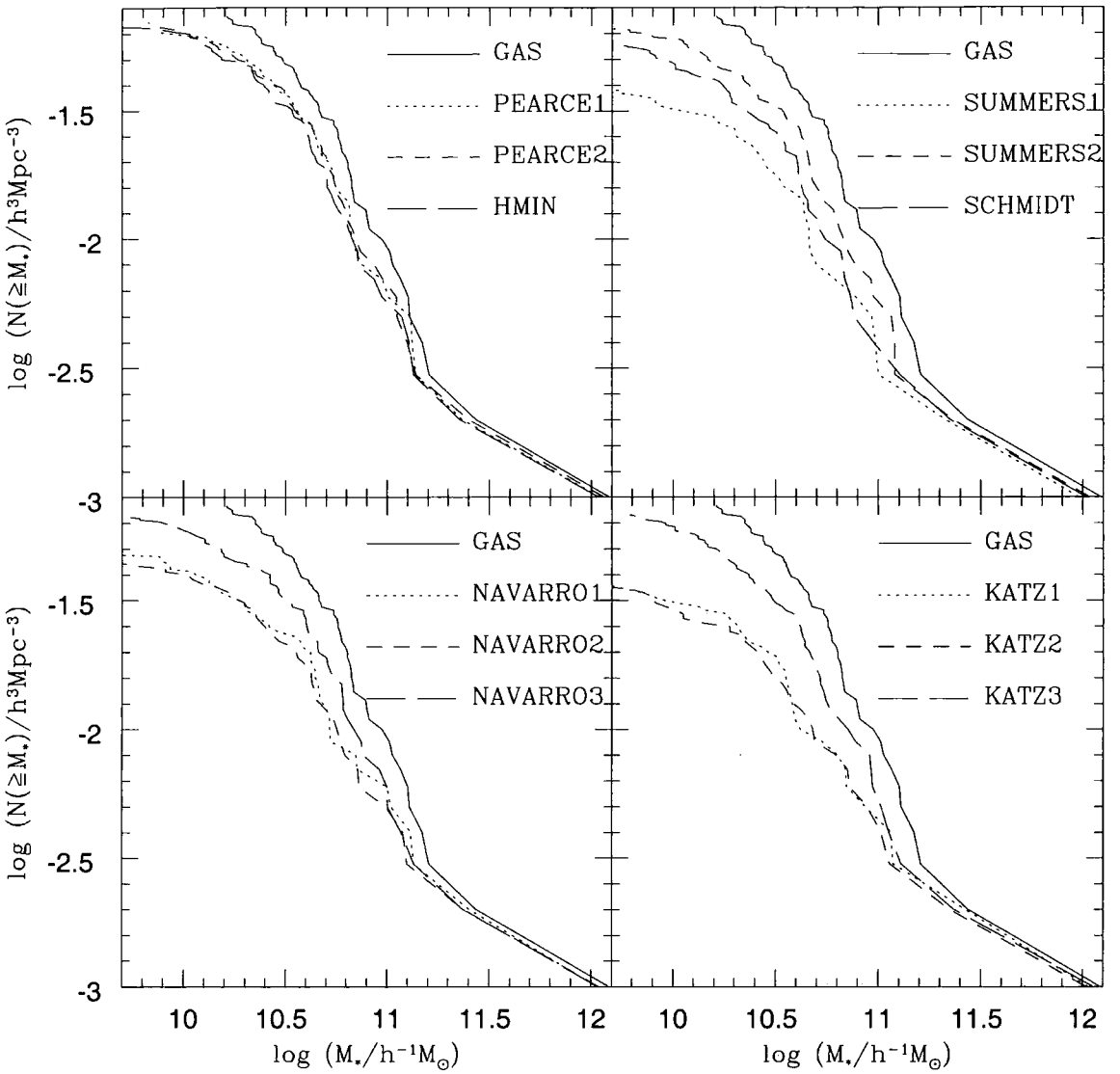


Figure 5.8: Cumulative mass functions of the galaxies (stars only) at $z = 0$ for all simulations with star formation. The mass function for GAS is also plotted in each panel.

Twelve simulations were run to study various criteria for converting gas into stars, many which have already been implemented in the literature.

It was found that the amount of gas cooled by $z = 0$ in simulations with star formation did not vary significantly (less than 5 per cent) relative to the run with no star formation. The global fraction of baryons in stars by $z = 0$, as well as the star formation history, were primarily controlled by the choice of density threshold; a higher threshold lengthens the timescale of the star formation process within cooled objects. Simulations with prescriptions which efficiently

converted gas into stars produced star formation rates that were limited by the cooling rate of the gas.

The number of galaxies formed within the simulations was constant to within $\sim$ 10 per cent. Cumulative mass functions of the galaxies (containing both stars and cold gas) were in excellent agreement for all simulations studied, with no difference between these and the mass function of the simulation with no star formation. When the cold gas was subtracted for each galaxy (leaving only the stars), the mass functions turned over at lower masses, signifying a higher cold gas to stars ratio for lower mass galaxies than for higher mass galaxies. In general, this ratio increases with the density threshold, for a given mass. The fraction of stars in galaxies was around 90 per cent or higher for all simulations except SCHMIDT, where the value was around 75 per cent.

Many additional constraints on the star formation process served to make prescriptions unnecessarily complicated. Examples of these are a star formation rate based on the dynamical time of the gas, the use of separate particles to represent stars and a "Jeans condition" to identify regions unstable to gravitational collapse.

5.4 Feedback

5.4.1 Method

In this section, a set of 9 simulations was performed to test various mechanisms for implementing feedback within cosmological simulations. The main requirement for any feedback prescription is a mechanism for transferring energy from the stars (once they have been around long enough for supernovae to occur) to the surrounding gas. These mechanisms can be split into 2 categories: those which heat the surrounding gas by supplying feedback as thermal energy, and those which transfer energy in kinetic form, by adding perturbations to the gas velocity. Both methods are studied in this chapter, although emphasis is placed on the thermal method, widely believed to be the more plausible of the two for the range of scales considered (see Katz, Weinberg, & Hernquist 1996 for a discussion on this subject).

Each simulation was run from the same initial conditions used in the previous section and gas was converted into stars following the PEARCE1 prescription. The primary reason for choosing this method is its simplicity. However as was shown, the majority of the methods studied in the previous section converted gas into stars at a rate dictated by the choice of the density threshold. The PEARCE1 mechanism instantaneously converts gas particles into star particles if the former have overdensities exceeding 5×10^3 and temperatures below 3×10^4K.

5.4.1.1 Energetics

The maximum amount of energy available from supernovae was calculated using the same arguments as Katz, Weinberg & Hernquist (1996) for all the simulations studied. Stars with masses greater than $8M_\odot$ are assumed to release 10^{51} erg of energy into the surrounding gas. The mass frequency of new stars was calculated using the initial mass function (IMF) given by Miller & Scalo (1979), with lower and upper mass limits of $0.1M_\odot$ and $100M_\odot$ respectively. This implies that the specific energy released in supernovae is $\sim 3.7 \times 10^{15}\,\mathrm{erg\,g^{-1}}$, or equivalently a temperature of $2.9 \times 10^7\mathrm{K}$ [1].

The size of the timesteps in the simulations (typically $3 \times 10^6\mathrm{yr}$) is considerably smaller than the lifetime of an $8M_\odot$ star ($\tau_8 \sim 2 \times 10^7\mathrm{yr}$). For this reason, the feedback energy was released gradually, using the following function

$$u_{\mathrm{SN}}(\Delta t) = 3.7 \times 10^{15} \exp(-\Delta t/\tau_8)\ \mathrm{erg\,g^{-1}}, \tag{5.8}$$

where Δt is the time since the star particle was formed. Each star particle releases energy until $\Delta t = 200\mathrm{Myr}$, where u_{SN} becomes negligible.

5.4.1.2 Simulations with thermal feedback

The first simulation, labelled THERMAL1, used a feedback prescription based on the standard SPH smoothing technique to distribute u_{SN} from each star particle to the surrounding gas particles. This method was originally studied by Katz (1992) and has since been used by various groups (e.g. Navarro & White 1993; Summers 1993; Katz, Weinberg, & Hernquist 1996). When a gas particle is converted into stars, both its density and neighbour list are still maintained by the SPH algorithm (even though hydrodynamic forces are no longer applied). Each neighbour, providing it is a gas particle, receives a weighted contribution of the energy available at that time. Hence, a gas particle (labelled with subscript i) receives a total contribution to its thermal energy

$$\Delta u_i = \sum_{i \neq j} u_{\mathrm{SN}}(\Delta t)_j \frac{m_j}{\rho_j} W_{ij}, \tag{5.9}$$

where the sum is over all neighbouring star particles, labelled with subscript j, with masses m_j and densities ρ_j. The smoothing kernel, W_{ij} is averaged between pairs (Equation 3.30). Note that the efficiency of energy transfer in this method is resolution dependent. A larger number of neighbours

[1] $T = (\gamma - 1)\mu m_{\mathrm{p}}/k$, where $\gamma = 5/3$ is the ratio of specific heats at constant volume and constant pressure, for a monatomic ideal gas; $\mu = 0.59$ is the mean molecular weight of a gas of primordial composition; m_{p} and k are the mass of a proton and Boltzmann's constant respectively.

($N_{\rm SPH}$), or a higher local density will increase the fraction of energy received by the surrounding gas.

The efficiency problem was circumvented by the method used in the second simulation, THERMAL2. Rather than smoothing over near neighbours, each star particle adds a fraction, $f_{\rm t}$, of its energy to the nearest gas particle. A similar method was implemented by Gerritsen & Icke (1997) in their simulations of disk galaxies. A value of $f_{\rm t} = 1$ was adopted for this study, i.e. adding all the available energy at a given time to the nearest gas particle. Such an efficiency is highly unlikely in real supernovae environments. However a value of unity was chosen to test the maximum possible effect the feedback could have.

The third and fourth simulations, labelled DECOUPLE1 and DECOUPLE2, were identical to the first two, except that the cold gas was decoupled from the hot gas. Densities for gas particles with temperatures above 10^5K are calculated by only summing over near neighbours with temperatures above 12000K. This technique has already been applied to large cosmological simulations (see Pearce et al. 1999) to reduce the gas cooling rate around the largest galaxies (c.f. comparison 4 in the previous chapter). Within the context of feedback, it was studied to observe if it had any effect on the cooling rate of the reheated gas.

The next two simulations, TWOPHASE1 and TWOPHASE2 were performed to further extend the method of DECOUPLE2. As in the latter simulation, feedback energy is supplied to the nearest gas particle which can then be decoupled from surrounding cold gas if is hot enough. Two additions have been made, one to the conversion of gas into stars and one in the heating of the gas from feedback. Firstly, when the gas is eligible to form stars, a star formation rate is applied using Equation 5.6, with $t_* = \tau_{\rm dyn}$; $\tau_{\rm dyn}$ is the gas dynamical time given by Equation 5.4 and c_* is a free parameter. This introduces a short delay in the formation of new star particles, at a frequency controlled by c_*: for TWOPHASE1, $c_* = 1$ and for TWOPHASE2, $c_* = 0.1$. Hence cold gas is prevented from turning into stars in one timestep, which causes the feedback to lose its effectiveness for reheating cold gas *before* it can turn into stars.

The second change is to inhibit the gas from cooling for a duration of $\tau_{\rm dyn}/c_*$, allowing reheated gas more time to equilibrate with any surrounding hot gas. The no-cooling approach was also used by Gerritsen (1997) who argued its validity essentially from the failure to model the multiphase nature of the ISM. As each supernova shell explodes, most of the energy is left behind in a hot, ambient medium with considerably lower density (and therefore longer cooling times) than surrounding cold gas (McKee & Ostriker 1977). It is evidently impossible to model such behaviour within cosmological simulations, where internal properties of the galaxies are

unresolved. Furthermore, the feedback scheme has to be most effective at the resolution limit of the simulation (where cooling times are shortest) with objects containing a few tens of particles. Therefore, the hope is that the thermal effects of supernovae are mimicked, when smoothed over scales of 10 kpc or so.

5.4.1.3 Simulations with kinetic feedback

The final three simulations (KINETIC1, KINETIC2 and KINETIC3) assume that the feedback energy is transferred via kinetic perturbations. As in THERMAL1, the energy is smoothed over near neighbours

$$\Delta v_i^2 = \sum_j \frac{2}{m_j} \frac{m_j W_{ij}}{\rho_j} f_{\rm v} u_{\rm SN} (\Delta t)_j, \tag{5.10}$$

where the label j runs over all neighbour particles and f_v is an efficiency parameter. For KINETIC1, $f_v = 0.1$, KINETIC2, $f_v = 0.01$ and KINETIC3, $f_v = 0.005$. Kinetic feedback was first studied by Navarro & White (1993) and subsequently by Mihos & Hernquist (1994).

5.4.2 Results

5.4.2.1 General properties at $z = 0$

Details of the simulations and some of their $z = 0$ properties are summarised in Table 5.4. The first three columns summarise the simulations studied and pertinent details of the feedback schemes.

Column 4 gives the number of timesteps taken to complete each simulation. Results for the simulations with thermal feedback scatter between $\sim 2500 - 3000$ steps, similar to the number in the simulations including star formation only. All of the runs with kinetic feedback took more steps, with the number increasing with the value of f_v. The velocity perturbations added to the gas are large enough to cause significant disruption of the galaxies. This leads to particles being ejected from the objects completely, with their high velocities (and the consequent shocking of surrounding gas) shrinking the size of the timestep.

Columns 5 and 6 give the fraction of baryons in the galaxy phase, $f_{\rm g}$ and the fraction of stars, $f_{\rm stars}$, in each simulation. For the first four runs, the values of $f_{\rm g}$ and $f_{\rm stars}$ are within 2 per cent of one another, whereas for the remaining five they are significantly smaller. Several conclusions can already be drawn from the data in Table 5.4. Solely adding thermal energy to the gas makes no significant difference to the global cooling and star formation rates. As discussed in Katz (1992), gas receiving feedback energy has high enough density to radiate it away rapidly. When all of the feedback energy is added to a star particle's nearest neighbour, thus increasing the efficiency of

Table 5.4. Details and properties of the cooled gas at the end of each simulation with star formation and feedback. Column 1 gives the label used for each simulation; column 2 lists the form of feedback used; column 3 gives the value of parameters used; column 4 lists the number of timesteps taken to evolve each simulation to $z = 0$; column 5 gives the fraction of baryons in the galaxy phase; column 6 gives the fraction of baryons in stars and column 7 lists the number of galaxies in each simulation.

Simulation	Form	Parameter	$N_{\rm steps}$	$f_{\rm g}$	$f_{\rm stars}$	$N_{\rm gal}$
THERMAL1	thermal		3188	0.34	0.28	67
THERMAL2	thermal		2585	0.33	0.28	74
DECOUPLE1	thermal		2616	0.32	0.26	67
DECOUPLE2	thermal		2541	0.32	0.26	62
TWOPHASE1	thermal	$c_* = 1.0$	2753	0.20	0.15	46
TWOPHASE2	thermal	$c_* = 0.1$	3001	0.17	0.13	48
KINETIC1	kinetic	$f_{\rm v} = 0.1$	5060	0.04	0.02	1
KINETIC2	kinetic	$f_{\rm v} = 0.01$	3618	0.11	0.08	10
KINETIC3	kinetic	$f_{\rm v} = 0.005$	3117	0.14	0.12	22

energy transfer, again no difference is made. The same is true when the hot gas is decoupled from the cold gas.

The first simulations in which the cooling and star formation activity is significantly reduced are TWOPHASE1 and TWOPHASE2. These runs have cooled gas fractions in reasonable agreement with the observations. Reducing c_* by a factor of 10 allowed less gas to cool.

The runs with kinetic feedback show drastic reductions in both $f_{\rm g}$ and $f_{\rm stars}$. The velocity perturbations added to the neighbours of star particles are sufficiently large to unbind material in each galaxy, preventing more of the gas from forming stars. Ejected particles are then shock heated and become part of the diffuse halo gas. The amount of material ejected from galaxies will be overestimated for two reasons. Firstly, the softening length (and therefore spatial resolution) of the simulations is larger than the scale lengths of observed galaxies with similar masses (a few kpc or so). Thus, the softened potential makes it easier for the perturbed gas to escape the gravitational confinement of the galaxy. Secondly, a significant degree of thermalization should occur in the reheated gas via shocks. Since these shocks are unresolved, none of the feedback energy can ever be dissipated. Hence, the value of f_v is sensitive to the resolution of the simulation.

In Column 8, the number of galaxies (identified using the method described in Section 5.3.2 with a linking length of $b = 0.1$) are listed for each run. The first 4 runs produced similar numbers

of galaxies, but the last 5 simulations produced significantly fewer. In simulations TWOPHASE1 and TWOPHASE2, the number of galaxies was reduced by $\sim$ 25 per cent, relative to the first 4. Simulation KINETIC1, with the largest value of f_v, produced only 1 object by $z = 0$. The remaining two (KINETIC2 and KINETIC3) also produced fewer number of galaxies, $\sim$ 16 per cent and $\sim$ 34 per cent of the number without feedback (PEARCE1, in the previous section, has 64 galaxies). Although KINETIC3 has roughly the right fraction of baryons in stars, the fraction of stars in galaxies is around 46 per cent.

5.4.2.2 Star formation rates

Figure 5.9 illustrates the global star formation histories in all of the simulations, plotted as the star formation rate density (in units $M_\odot$/ Mpc^3/yr) against time, in Gyr. To reduce noise, the data are shown as histograms with the width of each bin equal to 0.2 Gyr.

The first four simulations have star formation rates similar to that found when no feedback is included (run PEARCE1 in the previous section). Hence, feedback was ineffective at reducing the amount of star formation for these runs. The remainder of the runs show a decrease in the star formation rate, with the largest reduction occurring near the peak (at $\sim 2-3$ Gyr). The greater the star formation rate, the more energy becomes available to reheat the surrounding gas and prevent it from forming stars. Note that the characteristic shape is preserved in all runs with reduced star formation, i.e. the peak at early times followed by a relatively long decay.

5.4.2.3 Mass functions

Cumulative galaxy mass functions (including both cold dense gas and stars) are illustrated in Figure 5.10. The results have been split into four panels for ease of comparison; also plotted are the cumulative galaxy mass functions of PEARCE1 and GAS. The first four feedback runs (which show no significant decrease in the star formation activity) have mass functions that are not significantly different from the case when feedback is not included. Simulations TWOPHASE1 and TWOPHASE2 have galaxy mass functions of considerably smaller amplitude at all masses, with the former slightly higher than the latter (reflecting the amount of star formation). Hence, feedback is effective for the whole range of masses studied. Run KINETIC1 is shown as a point since it contains only one galaxy. Runs KINETIC2 and KINETIC3 show the same trend as the previous panel.

Cumulative mass functions of the galaxies with the cold gas removed are plotted in Figure 5.11. Again, the results have been split into four panels and the (gas) function of PEARCE1 has been overlaid. The first four runs show no significant difference in the stellar mass function when

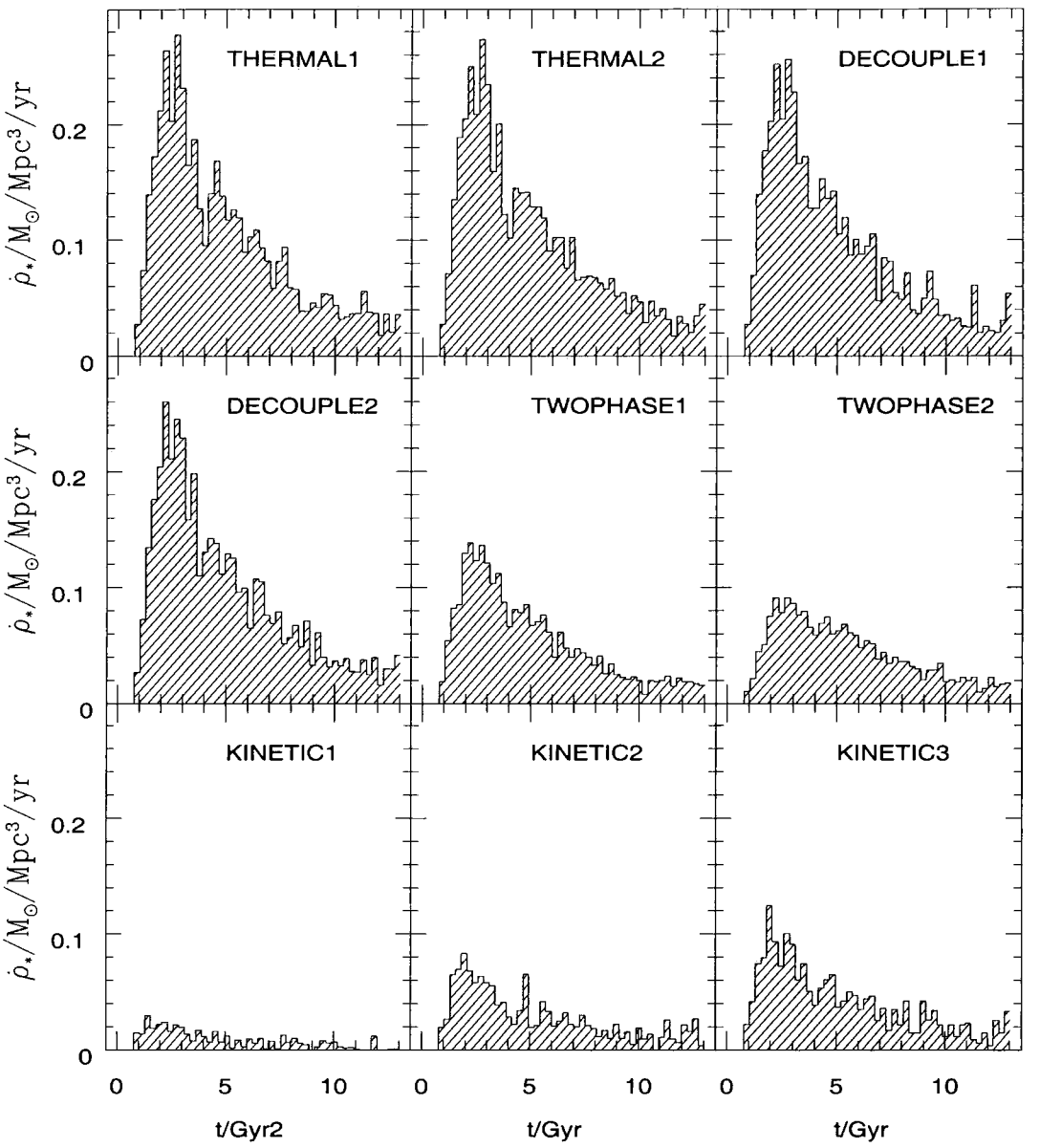


Figure 5.9: The star formation rate density (in $M_\odot$/ Mpc^3/yr) plotted against time, in Gyr for each of the 9 simulations studied with feedback. Each bin has a width of 0.2 Gyr.

compared to the run without feedback. Runs TWOPHASE1 and TWOPHASE2 have similar shape but reduced amplitude. Run KINETIC1 is shown as a point. For the remaining two (KINETIC2 and KINETIC3), the amplitude is lower, but the shape is similar at the high mass end: the low mass end flattens off suggesting that the kinetic feedback is more effective within low mass objects (for which the binding energy is generally lower).

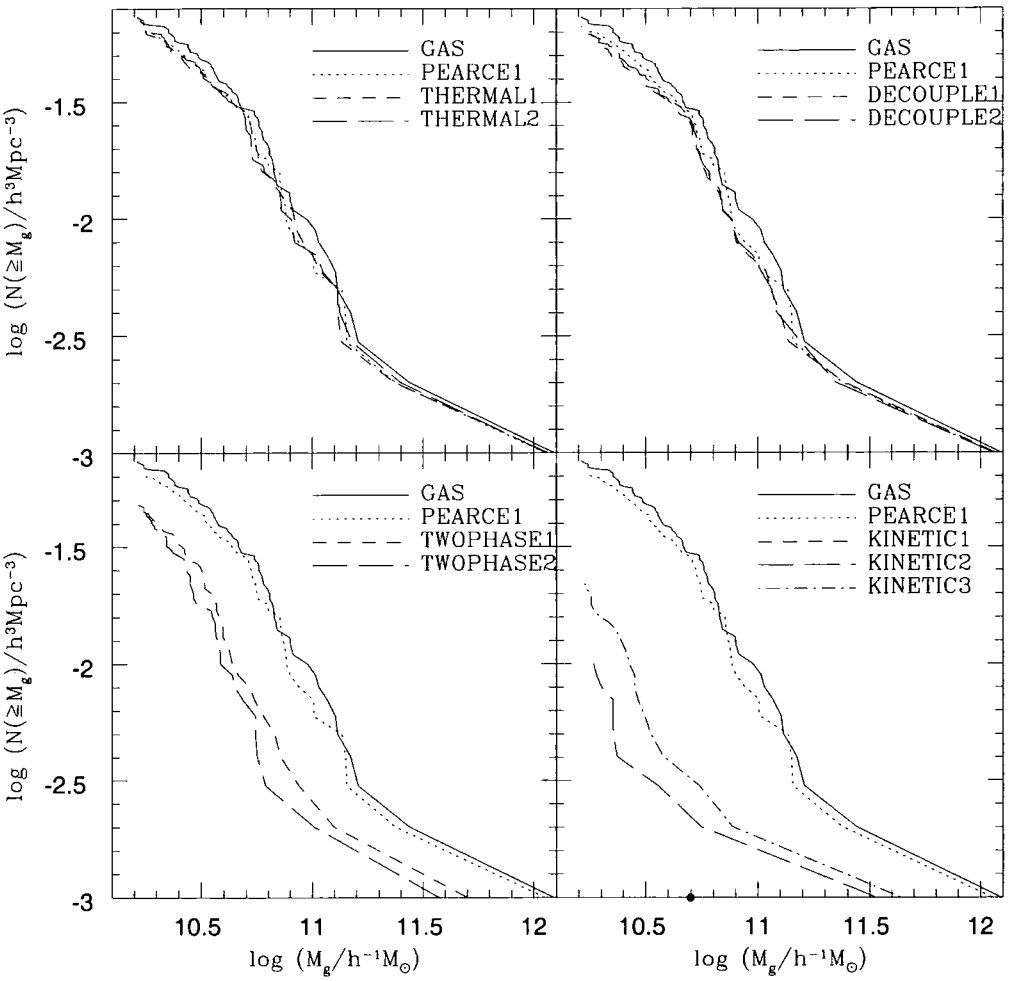


Figure 5.10: Cumulative galaxy mass functions for each of the 9 feedback runs. Also shown (in all panels) is the galaxy mass function for an identical run except with no star formation or feedback.

5.4.2.4 Luminosity functions

Figure 5.12 illustrates the cumulative K-band luminosity function of the galaxies for each of the 9 simulations with feedback. The luminosity of each galaxy was calculated by first calculating the K-band luminosity per unit mass, $l_{\rm K}$, of each star particle using stellar population models (Bruzual A. & Charlot 1993); for this, all that is required is the lookback time to when each star particle formed. The K-band magnitude of the galaxy is then

$$M_{\rm K} = -2.5\log(l_{\rm K}\Upsilon m_*) + M_{\rm K,0}, \tag{5.11}$$

where Υ is the mass-to-light ratio (i.e. $\Upsilon \geq 1$, allowing some fraction of stars to be brown dwarfs), m_* is the mass of each star particle and $M_{\rm K,0}$ is the zero-point magnitude. Also shown in the Figure

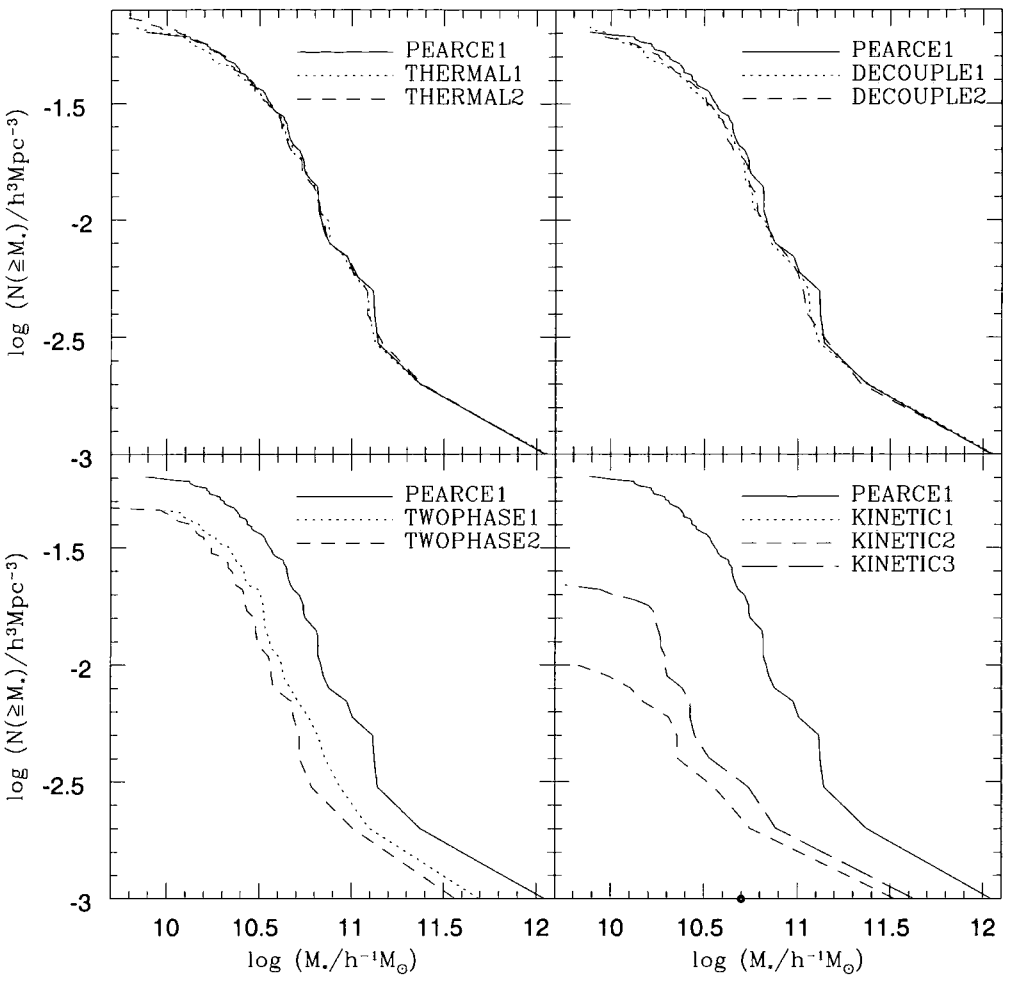


Figure 5.11: Cumulative stellar mass functions for each of the 9 feedback runs. Also shown (in all panels) is the galaxy mass function for an identical run except with no star formation or feedback.

is the best-fit Schechter parameterization of the K-band luminosity function from the Stromlo-APM Redshift Survey (Loveday 2000),

$$\phi(M_{\rm K})dM_{\rm K} = 0.921\phi^* \exp(-L/L^*)\,(L/L^*)^{1+\alpha}dM_{\rm K}, \tag{5.12}$$

where $L/L^* = \exp(-0.921(M_{\rm K} - M^*_{\rm K}))$, $M^*_{\rm K} = -23.58 + 5\log h$, $\alpha = -1.16$ and $\phi^* = 0.012\,h^3\,{\rm Mpc}^{-3}$.

A value of Υ has been chosen for each simulation in order to fit the bright-end (around $M^*_{\rm K}$) of each luminosity function. The first four simulations with thermal feedback fit the bright end but require $\Upsilon = 3$. Since the feedback is ineffective at reducing the amount of cooled gas in these models, the galaxies are too bright. Models TWOPHASE1 and TWOPHASE2 require more

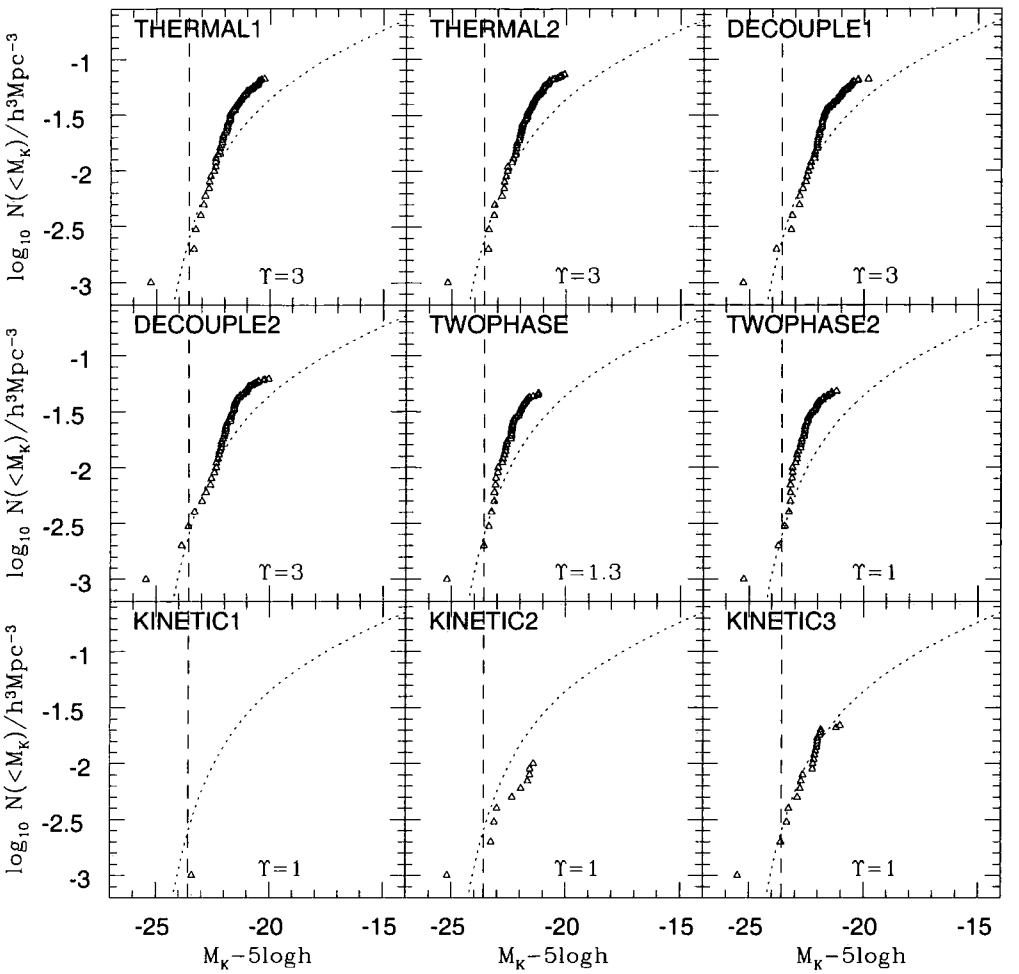


Figure 5.12: Cumulative K-band luminosity functions for each of the 9 feedback runs (points). The dotted curve is the best-fit Schechter function to the data presented in Loveday (2000) and the dashed vertical line illustrates the value of $M_{\rm K}^*$. Also given, in the bottom-right of each panel is the mass-to-light ratio adopted for each function.

modest values of Υ (the latter providing a good fit to the bright end with all the mass in visible stars). However, note that they also suffer from an excess of galaxies below $M_{\rm K}^*$. Model KINETIC3 provides the best-fit with $\Upsilon = 1$ but as discussed above, more than half of the stars have been stripped out of the galaxies in this run.

5.5 Conclusions

In this chapter the inclusion of star formation and feedback within cosmological simulations has been investigated in an attempt to reduce the amount of baryons that cool radiatively to an observationally acceptable value. Firstly, the efficiency of the radiative cooling process was studied

in detail, testing its sensitivity to the resolution of the simulation, the size of the volume and the mass ratio of gas to dark matter particles. One of these simulations with moderate resolution was then rerun with twelve different prescriptions of turning gas into stars. A second set of 9 tests was performed, exploring the effects of supernovae feedback. Conclusions from the work presented in this chapter can be summarised as follows

1. Increasing the resolution of the simulation increases the fraction of gas that cools, as expected (and demonstrated previously, in Chapter 4). The fraction varies weakly with resolution for simulations with the highest values. However, as the resolution is lowered, the fraction of cooled material decreases dramatically. Low resolution simulations with equal mass gas and dark matter particles (achieved by increasing the number of dark matter particles relative to the number of gas particles) shows a much weaker trend than for their equivalents with equal numbers of particles. The amount of shocked gas varies to compensate for changes in the amount of cooled gas; only 2 per cent of the baryons is uncollapsed and in voids, a value insensitive to the resolutions studied.

2. The cooling rate of the gas shows a strong peak at around $z = 2 - 3$ which both increases and shifts (weakly) to earlier epochs when the resolution is increased.

3. Doubling the box-size of the simulation makes no significant changes to the phase fractions or cooling rate of the gas, even though intrinsic scatter, arising from the volumes studied, is present at some level.

4. As found by Pearce (1998), a simple prescription for turning gas into stars, based on the densities and temperatures of the gas particles suffices. A suitable choice of temperature and density threshold picks out the gas that has cooled to the highest density, where star formation is most likely to occur. Additional constraints are unnecessary, as subsequent processes leading to star formation (such as Jeans instability) occur on scales well below what are presently attainable with cosmological simulations.

5. The star formation rate is primarily controlled by the efficiency with which the gas can cool, providing the mass in galaxies is dominated by stars. The ratio of gas to stars in galaxies is only sensitive to the densities at which stars can efficiently form.

6. The cooling rate of the gas is unaffected by star formation. No differences in the resulting mass functions of galaxies (cold gas + stars) are observed for simulations with and without star

formation. Mass functions excluding the cold gas have lower amplitudes, with larger deviations for smaller masses.

7. Incorporating stars by spawning child particles or giving particles dual identity (as gas and stars) makes no significant difference to any of the quantities measured, at the expense of increasing computational effort relative to simply converting whole particles into collisionless star particles. The latter method can reduce the total CPU time by a factor of 2 when compared to a simulation with no star formation.

8. Thermal feedback is ineffective at reducing the amount of cooled material since most of the energy is radiated away. Giving energy to one particle rather than smoothing over near neighbours does not solve the problem. Neither does solely decoupling the hot gas from the cold gas, as an attempt to lengthen the cooling times of reheated gas particles.

9. Kinetic feedback has catastrophic effects on the galaxy population. While successfully reducing the amount of gas that cools and forms stars, it easily disrupts the objects themselves, reducing the number of galaxies by a factor of 3 for an efficiency of 0.5 per cent; only 1 galaxy remains at $z = 0$ for an efficiency of 10 per cent.

10. The only scheme that successfully reduced the fraction of cooled gas and stars, while producing reasonable numbers of galaxies was a scheme that prevented reheated gas from cooling for a brief period of time, based on the local dynamical time. The overall shape of the star formation rate function and the mass function are similar to the case with no feedback, however this method over-predicts the number of galaxies with K-band magnitudes fainter than $M_{\rm K}^{*} = -23.58$.

Future work will require to look at the effects of each feedback model at various resolutions - the most important test for any model is whether it stabilises the amount of gas that cools. As smaller haloes are resolved with shorter dynamical times, the feedback should be able to regulate the amount of star formation.

References

Bruzual A. G., Charlot S., 1993, ApJ, 405, 538

Buonomo F., Carraro G., Chiosi C., Lia C., 1999, MNRAS, submitted (astro-ph/9909199)

Carraro G., Lia C., Chiosi C., 1998, MNRAS, 297, 1021

Cole S., 1991, ApJ, 367, 45

Dekel A., Silk J., 1986, ApJ, 303, 39

Eke V. R., Efstathiou G., Wright L., 1999, MNRAS, submitted (astro-ph/9908294)

Evrard A. E., Henry J. P., 1991, ApJ, 383, 95

Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189

Frenk C. S., White S. D. M., Efstathiou G., Davis M., 1985, Nature, 317, 595

Fukugita M., Hogan C. J., Peebles P. J. E., 1998, ApJ, 503, 518

Gerritsen J. P. E., Icke V., 1997, A&A, 325, 972

Kaiser N., 1986, MNRAS, 222, 323

Kaiser N., 1991, ApJ, 383, 104

Katz N., 1992, ApJ, 391, 502

Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19

Kennicutt J., R. C., 1998, ApJ, 498, 541

Loveday J., 2000, MNRAS, 312, 557

McKee C. F., Ostriker J. P., 1977, ApJ, 218, 148

Miller G. E., Scalo J. M., 1979, ApJS, 41, 513

Navarro J. F., White S. D. M., 1993, MNRAS, 265, 271

Pearce F. R., 1998, A&A, submitted (astro-ph/9803133)

Pearce F. R. et al., 1999, ApJ, 521, L99

Ponman T. J., Cannon D. B., Navarro J. F., 1999, Nature, 397, 135

Press W. H., Schechter P., 1974, ApJ, 187, 425

Schmidt M., 1959, ApJ, 129, 243

Steinmetz M., Müller E., 1995, MNRAS, 276, 549

Steinmetz M., White S. D. M., 1997, MNRAS, 288, 545

Summers F. J., 1993, Ph.D. thesis, AA(California Univ.)

Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253

Tissera P. B., Lambas D. G., Abadi M. G., 1997, MNRAS, 286, 384

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

Chapter 6

Simulations of a galaxy cluster

Abstract. Preliminary results are presented for three simulations of a galaxy cluster within a flat, low density universe. The cluster was chosen from a cold dark matter cosmological simulation and resimulated with gas at higher resolution. Gas in the first simulation could only change its entropy through shocks. The second simulation also included the modelling of radiative cooling processes. The third additionally used a model for star formation and feedback. Various properties of all three clusters are studied, such as their density and temperature profiles, as well as the mass distribution of the galaxies that have formed in the simulations with cooling.

6.1 Introduction

The origin and nature of galaxy clusters is a problem that has been attacked with numerical simulations for nearly thirty years; their relatively simple structure (when compared to the galaxies themselves) being a primary motivation. Early studies using collisionless N-body simulations demonstrated the feasibility of galaxy clusters as gravitationally bound systems, formed via gravitational instability in the expanding Universe (e.g. Aarseth 1963; Peebles 1970; White 1976). However, it was the development of models of the large scale structure in the universe, particularly the cold dark matter paradigm, that placed galaxy clusters within a robust cosmological context (Peebles 1980). Collisionless CDM simulations verified the existence of clusters in these models, as the most recent objects to have formed from the merging hierarchy of dark matter haloes (e.g. Davis et al. 1985 and references therein).

The success of the CDM simulations, as well as the discovery of the X-ray emitting intracluster medium (ICM) prompted the extension of numerical models to include gas. Both Eulerian and Lagrangian codes were developed and subsequently applied to modelling the structure of the ICM, confirming the general belief that most of the gas is in hydrostatic equilibrium within the cluster potential (e.g. Evrard 1990; Thomas & Couchman 1992; Cen & Ostriker 1994; Navarro, Frenk, & White 1995; Anninos & Norman 1996; Bryan & Norman 1998;

Eke, Navarro, & Frenk 1998). Such studies formed the basis for the Santa Barbara Cluster Comparison Project, demonstrating the general widespread agreement of results, regardless of the numerical code used (Frenk et al. 1999). Such studies have also led to a variety of useful applications, such as testing the validity of cluster scaling relations in various cosmologies and the shape of the X-ray temperature function. The first example was an integral part of the model developed in Chapter 2, and both have been used particularly to discriminate between different cosmological scenarios.

One of the key assumptions in the majority of these studies is that the intracluster medium (ICM) can be approximated as non-radiative, an assumption that is partly justified by the long cooling times in all but the central regions of clusters (Fabian, Nulsen, & Canizares 1991). At the very least, such a model cannot be reliably used to predict cluster properties based on their X-ray luminosity. The emission is a collisional process (thermal bremsstrahlung), scaling with the square of the density. Hence, most of the flux comes from the central regions of clusters where cooling is important. Work incorporating the effects of radiative cooling is still in its infancy, largely due to the increase in computational effort required to perform the simulations. Nevertheless, simulations have already begun to quantify the extent to which cooling affects the structure of the ICM (Thomas & Couchman 1992; (Katz & White 1993); Frenk et al. 1996; Suginohara & Ostriker 1998; Lewis et al. 1999; Pearce et al. 1999). Such work is crucial for comparisons with, for example, the X-ray luminosity-temperature relation and the X-ray luminosity function of clusters. Unlike the non-radiative models, convergence has not been established between groups over how the cooling should affect the X-ray properties of clusters, mainly due to the way in which the process is modelled. For example, Suginohara & Ostriker and Lewis et al. both conclude that cooling *increases* the luminosity of clusters when compared to estimates from non-radiative models. However Pearce et al. argue that the opposite is true, because the former models overestimate the central cooling rate of the gas.

Of equal importance are the effects of galaxy formation on the cluster. As discussed in the previous two chapters, radiative cooling already provides a first step, allowing gas to condense in the cores of galactic-sized haloes in less than a Hubble time (White & Rees 1978). Several groups have also examined the interaction of the cold dense knots of gas ("galaxies") with their environment and each other, circumventing overcooling by failing to resolve the haloes where the cooling times are drastically short (Katz & White 1993; Frenk et al. 1996) Understanding the effects of the cluster environment on the galaxies is important since it provides information about the galaxy formation process itself (e.g. Dressler 1980; Butcher & Oemler 1984; Poggianti et al. 1999).

A vital link that is still missing between simulations and observations is the inclusion of a model for star formation and feedback. Star formation allows direct studies of, for example, the star formation rates and colour distributions of cluster galaxies. Stars also affect the dynamical evolution of the galaxies, since unlike gas, they do not suffer from the effects of ram pressure and drag as they move through the ICM. The importance of feedback processes is perhaps even more important within the cluster environment. Direct observational evidence already exists in the high metallicities of the ICM, and many other effects have also been attributed to the effects of feedback, such as constant density cores in the ICM mass distribution (see Chapter 2) and the slope of the X-ray luminosity-temperature relation. It must be re-iterated that feedback may also allow numerically stable results to be achieved, preventing the gas from cooling in small haloes at high redshift.

This chapter presents preliminary work for a study of galaxy clusters, self-consistently modelling the properties of both the dark matter and gas (baryons) within a flat, low density CDM model. Results are presented for three simulations of a single object. The first simulation (referred to as A-clus) models the gas as non-radiative, which can adiabatically heat and cool as well as increase its entropy from shocks. The second simulation (G-clus) additionally models the effects of radiative cooling. Finally, a third simulation was run (S-clus) incorporating a model for both star formation and feedback that was studied in the previous chapter. It must be stressed that work on this project is still in progress and hence a full understanding of many of the results will require more analysis.

The remainder of this chapter is organised as follows. In Section 6.2, the numerical method for simulating the cluster is presented, detailing the procedure used for *resimulating* haloes in large cosmological simulations at higher resolution. Results are given in Section 6.3 and is summarised in Section 6.4.

6.2 The simulations

Several methods exist for simulating the formation of clusters, the most popular being cosmological simulations of large, random volumes. The advantage of this method is that there is usually a reasonable number of clusters of varying masses in the box, ideal for examining statistical properties of the objects. Furthermore, the simulations also contain the correct large-scale tidal forces, crucial in determining the dynamical properties of haloes. Unfortunately, the size of the volume also limits the resolution of the simulation, for two reasons. Firstly, only a finite number of particles can be stored in memory at any one time. Secondly, the cycle time for each timestep

scales with the number of particles to varying degrees (see Chapter 3). The problem is even harder if the model includes gas.

One solution to the problem, other than waiting for faster, larger computers, is to resimulate individual objects one at a time, selected from previously run cosmological simulations (known as *parent* simulations). Since the position, size and mass of haloes are primarily determined from evolution of the dark matter, parent simulations can be purely collisionless. The region that forms the object of interest can then be resampled at higher resolution (as well as the addition of baryons and additional physical processes, such as radiative cooling), and the resolution of the remaining volume can be degraded, since it is only used to provide the external forces.

The cluster that is studied in this chapter was simulated using the resimulation method. The remainder of this section details the procedure that was followed.

6.2.1 The parent simulation

The parent simulation was selected from the set of GIF collisionless cold dark matter (CDM) simulations run by the Virgo Consortium, designed to be used for combined N-body and semi-analytic studies of galaxy formation (e.g. Kauffmann et al. 1999). The chosen simulation used identical cosmological parameters to the ΛCDM model of Jenkins et al. 1998, i.e. $h = 0.7$ (the Hubble constant in units of $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$); $\Omega_0 = 0.3$ (the present value of the density parameter) and $\Lambda_0 = 0.7$ (the cosmological constant in units of $3H_0^2$). Structure was normalised using $\sigma_8 = 0.9$, in agreement with the local abundance of rich clusters (Eke, Cole, & Frenk 1996; Viana & Liddle 1996).

The parent simulation contains $N = 16777216$ dark matter particles within a box of length, $S = 141.3\,h^{-1}\,\mathrm{Mpc}$. The mass of each dark matter particle is therefore $\sim 1.4 \times 10^{10} h^{-1}\,\mathrm{M}_\odot$. The (Plummer) softening length was set to $\sim 70\,h^{-1}\,\mathrm{kpc}$ at $z = 50$, where the simulation was started. The softening evolved such that it was fixed in comoving co-ordinates until $z \sim 1.4$, where it decreased to $\epsilon_0 = 30\,h^{-1}\,\mathrm{kpc}$ at $z = 0$ (see Equation 3.20).

6.2.2 The choice of cluster

Candidate haloes were identified within the parent simulation by running a spherical overdensity groupfinder on the particles at $z = 0$ (see Section 4.3.4 for an overview of the groupfinder method). The cut-off radius (which will hereafter be referred to as the virial radius) was chosen such that the mean internal density was equal to $324\,\langle\rho\rangle$, as predicted from the spherical top-hat collapse model for the cosmological parameters adopted here. Only haloes with circular velocities in the range,

$V_c = 600 - 800\,\mathrm{km\,s^{-1}}$ were considered, which are identified with modest-sized clusters. The haloes were then subjected to the added constraint of being at a minimum distance of $5\,h^{-1}$ Mpc from other haloes with $V_c > 500\,\mathrm{km\,s^{-1}}$. Finally, the haloes all had to be regular in appearance, with no visible signs of recent major mergers.

The chosen cluster has a virial radius, $r_{\rm vir} \sim 1.1\,h^{-1}$ Mpc, containing a mass of $M_{\rm vir} \sim 1.49 \times 10^{14} h^{-1}\,\mathrm{M}_\odot$; the circular velocity, V_c is therefore $\sim 750\,\mathrm{km\,s^{-1}}$.

6.2.3 Resimulation details

The first stage that was performed for the generation of the resimulation initial conditions was to identify the region to be resampled at higher resolution. All particles were identified within a sphere, centred on the cluster at $z = 0$, of radius, $2r_{\rm vir}$. (The choice of radius was examined by Eke (1996), who used a similar technique to resimulate clusters, and concluded that twice the virial radius was adequate for preserving bulk cluster properties.) The particles were then traced back to their initial positions at $z = 50$, defining the shape of the region that collapsed to form the cluster. A cube was placed around this region that was just big enough to enclose the particles.

Using a cubic shape for the initial high resolution region is inefficient as the region defined by the cluster particles is more spherical. Hence, the cube demands using as much as a factor of 2 more particles than a perfect sphere. The shape of the high resolution region was defined by subdividing the cube into 32^3 cells and highlighting those containing the cluster particles (occupied cells). Any cells completely surrounded by occupied cells ("holes") were also marked as occupied. Additionally, empty cells adjacent to occupied cells were marked as border cells, creating a *shell* around the high-resolution region, used to protect new cluster particles from the surrounding mass. However, if an occupied cell was adjacent to the edge of the cube, the cube was grown iteratively and the procedure repeated until there was at least one empty cell between each border cell and the edges of the cube. The shape was stored as a three-dimensional mask of the high resolution region (including border cells). For the cluster studied here, the length of the high resolution cube is $19.159\,h^{-1}$ Mpc, hence the length of each mask cell is $\sim 0.6\,h^{-1}$ Mpc. The mass resolution was chosen such that each mask cell would contain 8 dark matter particles, an increase of a factor of ~ 6.3. Hence the high resolution cube would contain 64^3 particles.

The next stage was to compute the displacement fields for the resimulation, using the technique described in Section 3.3. The displacement field for the parent simulation is required for particles external to the cluster region (hereafter referred to as the low resolution region). The displacement field for the high resolution region had to be calculated in two parts. Firstly, the displacement field

for the high resolution cube was calculated for the range of power used in the parent simulation (also using the same phases). The increase in particle density required the original displacement field to be interpolated. Secondly, additional displacements were generated from the appropriate power spectrum, to extend the range of power capable of being resolved by the high resolution particles. The two were then added together.

After the displacement fields were calculated, a uniform distribution of particles had to be created, which would be perturbed with the displacements to produce initial positions and velocities. The first part was to create the low resolution distribution. For this, the full volume was subdivided into 472^3 cells, which is the number of high resolution particles required to fill the full volume. A sphere of radius $(\sqrt{3}/2)S \sim 122.4\,h^{-1}$ Mpc was then placed on the centre of the volume, just large enough to enclose it. Assuming spherical polar co-ordinates, the sphere was subdivided along the radial (r) and angular directions (θ and ϕ). The number of bins in each of the three orthogonal directions were defined singularly by the value $\Delta\theta$. The length of each bin along r was scaled as $\Delta r = r\Delta\theta$, so that the radial size increased in proportion to the angular size. The number of bins in the ϕ direction was obtained using the relation, $\Delta\phi = \min(\pi/2, \Delta\theta \sin\theta)$, preventing the azimuthal angular size of each bin from becoming vanishingly small. A uniform mass distribution was then created by adding the centroid position of each of the 472^3 cells to the corresponding spherical bin, for all cells lying outside the high resolution mask, which was placed in the centre of the volume. The resulting mass distribution is one that increases with distance from the centre of the volume and hence the cluster, hence producing a drastic reduction in the number of particles external to the region of study. The uniform distribution for the high resolution region was simply a grid of particles within the occupied and border cells.

The particles were perturbed in the low resolution region via interpolation of the low resolution displacement field onto the uniform distribution. A CIC interpolation scheme was used, as described in Section 3.3. For the high resolution region, the uniform distribution of particles are in exactly the same positions as the calculated high resolution displacement field, hence the operation was a trivial addition. Velocities were prescribed using the Zel'dovich formalism (again, discussed in Section 3.3).

Figure 6.1 illustrates the initial particle distribution of the resimulated cluster. The full box, with both high and low resolution regions is shown, as well as a zoom shot of the high resolution region. The low resolution region contains 27612 dark matter particles and the high resolution contains 127488 dark matter particles. Hence the high resolution region contains less than twice the number of particles than if the cube was used.

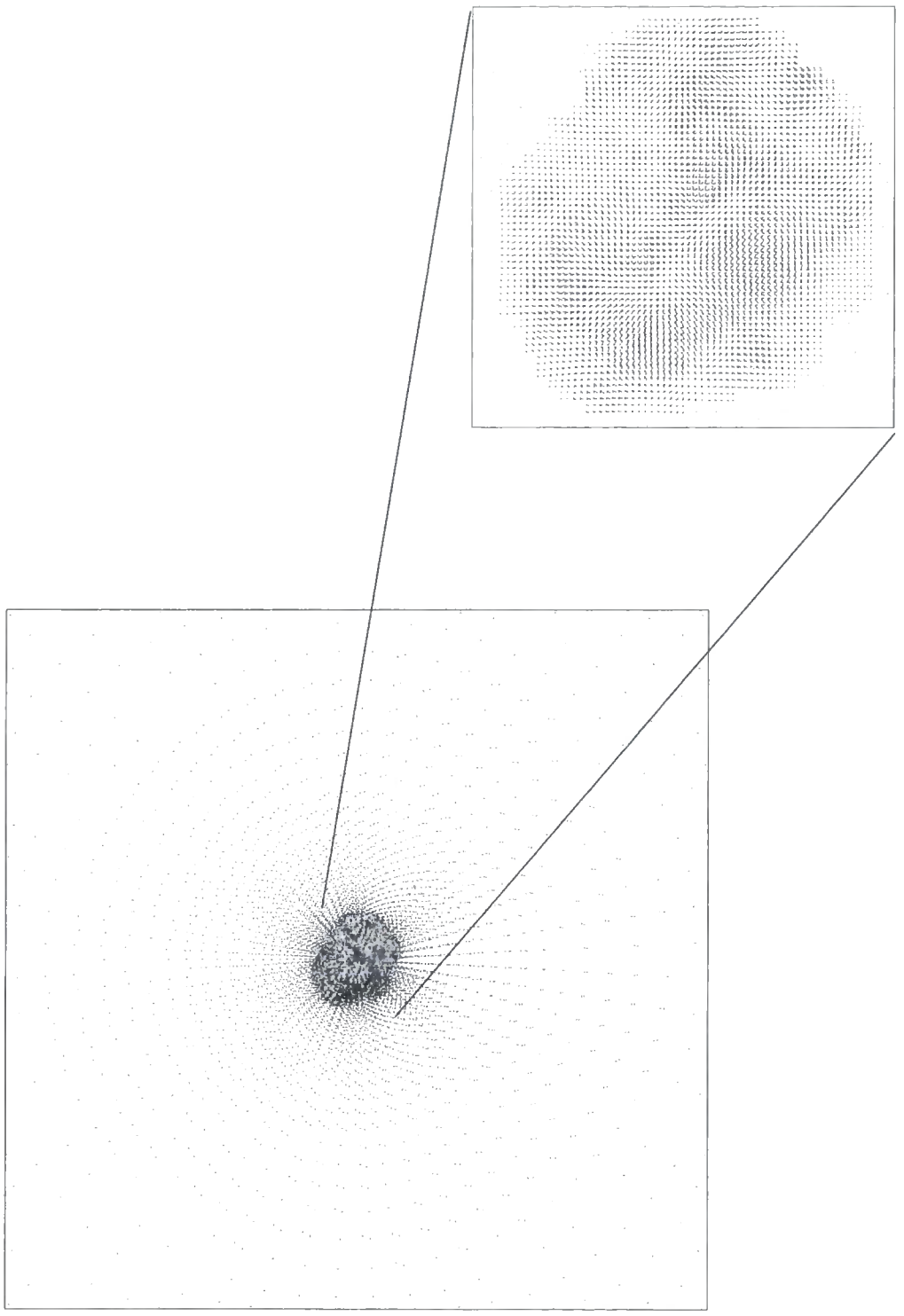


Figure 6.1: An example of the initial particle distribution for the resimulated cluster. The smaller box illustrates a zoom factor of ~ 7.

Gas particles were given the same initial positions and velocities as the dark matter, with masses fixed assuming a baryon fraction, $\Omega_b = 0.03$. This choice is consistent with nucleosynthesis calculations (Copi, Schramm, & Turner 1995) as well as the fraction of baryons measured in clusters, which makes life difficult for an $\Omega = 1$ universe (White et al. 1993; see also Gunn & Thomas 1996). The mass of each gas and dark matter particle is therefore $m_{\rm dark} = 2.23 \times 10^9 h^{-1}\,{\rm M}_\odot$ and $m_{\rm gas} = 2.23 \times 10^8 h^{-1}\,{\rm M}_\odot$ respectively. The mass of the dark matter particle is comfortably below the critical mass for two-body heating of the gas by the dark matter (Steinmetz & White 1997). The gas was given an initial temperature of 100K and a fixed metallicity of $Z = 0.3 Z_\odot$ was assumed, in agreement with X-ray observations of the ICM (e.g. Mushotzky et al. 1996).

The resimulations were run from $z = 50$ to $z = 0$ on a linux workstation using Version 3 of HYDRA , as detailed in Chapter 3. However, to complete the simulations within a reasonable timescale, the timesteps were fixed so that each simulation took approximately 10,000 steps to complete. Hence, the length of each step was ~ 1.36Myr (the age of the universe is ~ 13.6Gyr). At early times, the step was allowed to decrease to accurately account for the Hubble expansion. In practice, the timestep was short enough for all but the first 5 steps, after which it decreased by around a factor of 2.

For the simulations that explicitly followed the cooling rate of the gas (i.e. G-clus and S-clus), Equation 4.3 was applied to each gas particle at the end of the step. A cooling function was adopted using tabulated values from Sutherland & Dopita (1993), interpolated for a metallicity of $Z = 0.3 Z_\odot$. The S-clus simulation also included a model for star formation and feedback. For star formation, the method used in the PEARCE1 simulation (Section 5.3) was adopted, converting gas particles with overdensities greater than 5000 and temperatures less than 30000K into collisionless star particles. Feedback was implemented using the method of TWOPHASE2, discussed in detail in Section 5.4. To briefly summarise, energy from each star particle is added to the thermal energy of the nearest gas particle, which is then prevented from cooling for a further 10 dynamical times, estimated from the local gas density.

6.3 Results

6.3.1 Definitions

The centre of each cluster was defined using a simple algorithm based on the maximum density of the gas. Firstly, the density of each particle was estimated (using a similar technique to the

SPH method, as discussed in Section 3.5). For epochs after the main cluster had formed (i.e. a distinct central peak in the mass distribution), a sphere was centred on the gas particle with the maximum density and its radius was set to equal $100(1+z)\,h^{-1}$ kpc (i.e. constant size in physical co-ordinates). The procedure was repeated, using the centre of mass of the sphere, until the centre was displaced by an amount less than $10(1+z)\,h^{-1}$ kpc. For epochs before the main cluster formed, the initial centre was guessed by eye. (Quantitative results are only presented for $z=0$ in this chapter).

The virial radius of each cluster was fixed at the value calculated for the original cluster, i.e. $r_{\rm vir}=1.1\,h^{-1}$ Mpc. This resulted in a mean internal density that was within a few per cent of the original value ($\sim 324\,\langle\rho\rangle$). None of the results presented in this section are sensitive to using this definition.

The baryons in the cluster were split into various components : cold gas (all gas with $T<30000$K), hot gas (all gas with $T>30000$K) and for S-clus, stars. Galaxies were defined using the method described in Section 5.3.2, using a linking length $b=0.1(1+z)$ in units of The redshift factor was included so that the linking length is constant in physical co-ordinates. This was necessary since the softening length (which defines the approximate size of the galaxies) scales in this way for the range of redshifts studied.

6.3.2 The assembly of cluster material

Figure 6.2 illustrates the dark matter distribution within the virial radius of the original cluster and G-clus at $z=0$. The resimulated cluster contains more particles as well as exhibiting a larger degree of substructure. The higher resolution of the resimulated clusters results in more small scale power, leading to the formation of more smaller mass haloes that subsequently collapse with the main cluster. Furthermore, the binding energy of haloes has increased due to the smaller softening employed (and for this case, the condensation of cooled gas within the cores of the haloes). This gives smaller objects a larger chance of surviving the effects of tidal stripping as they orbit the cluster potential.

The evolution of the dark matter is shown in Figure 6.3 for G-clus at six epochs: $z=3,2,1,0.5,0.1,0$. Each box has a length of $2.2(1+z)\,h^{-1}$ Mpc, i.e. twice the $z=0$ virial radius fixed in physical co-ordinates. The filamentary origin of the cluster is evident for the $z=3$ snapshot, which collapses to form a twin core system by $z=1$. The two cores have merged by $z=0.5$ and the cluster spends the remaining time accreting mass bound in smaller objects (the substructure present at $z=0$).

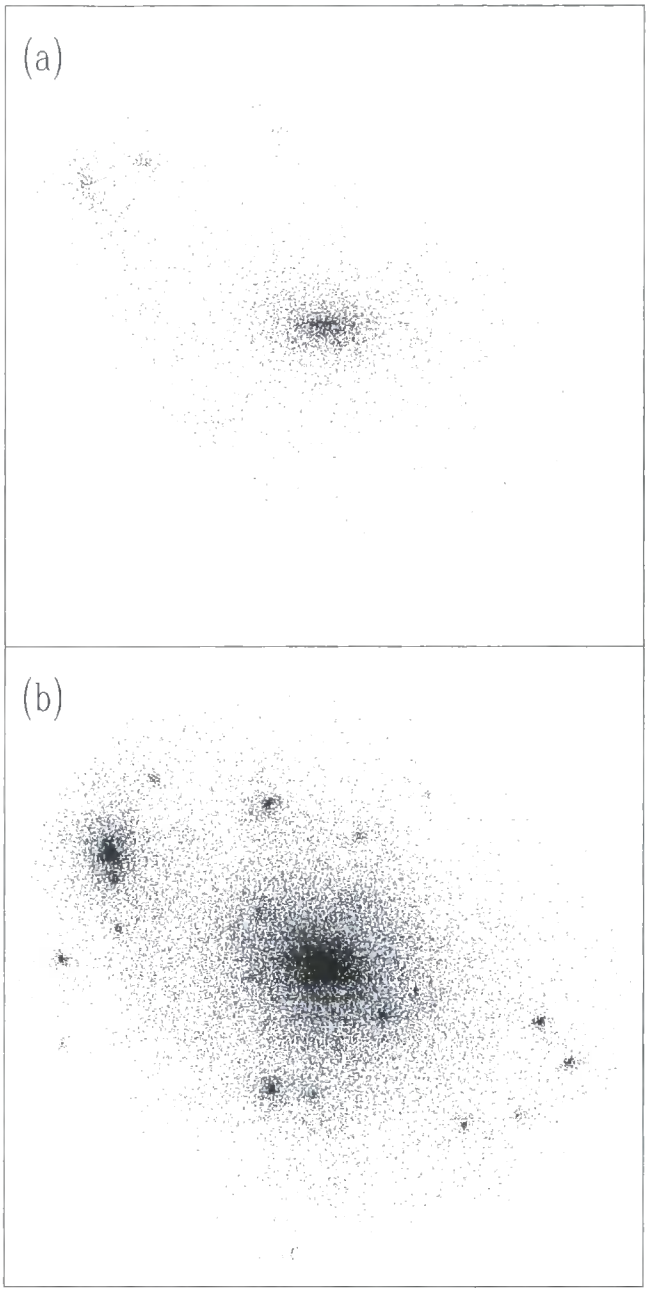


Figure 6.2: Projected positions of dark matter particles within the virial radius of (a) the original cluster and (b) the resimulated cluster (G-Clus) at $z = 0$.

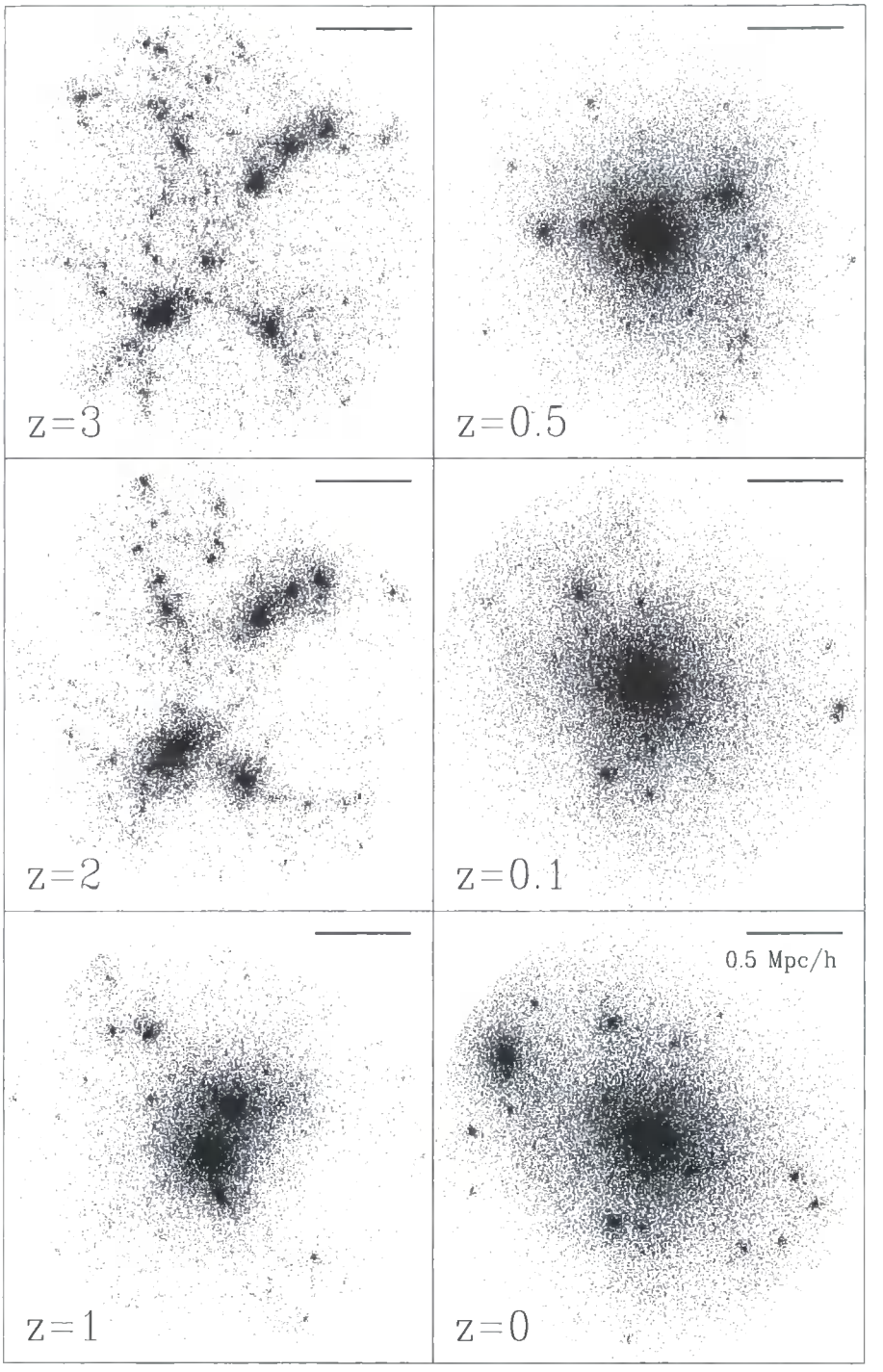


Figure 6.3: Projected positions of the dark matter particles within G-clus at various epochs.

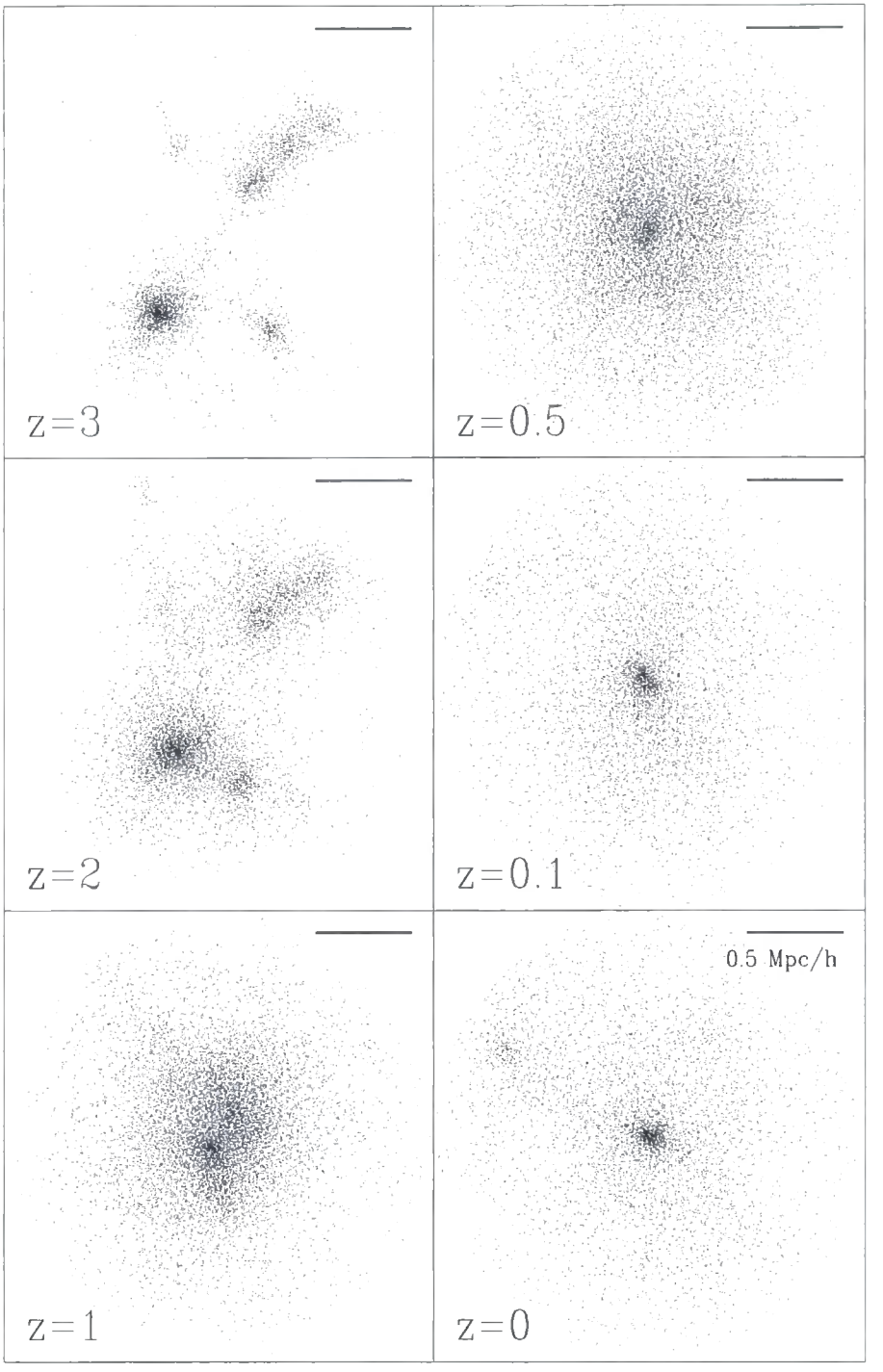


Figure 6.4: Projected positions of hot gas particles ($T > 30000$K) within G-clus at various epochs.

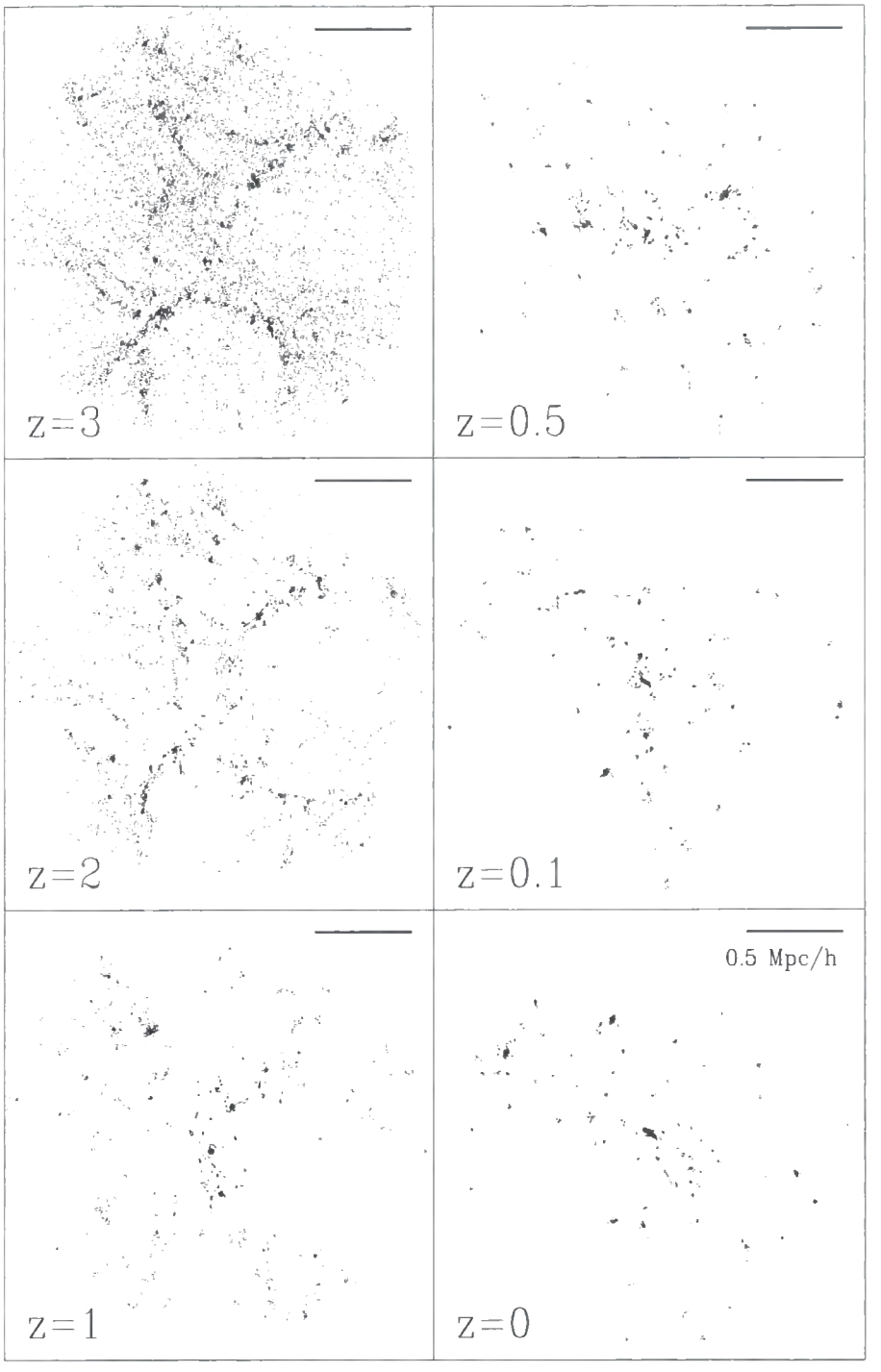


Figure 6.5: Projected positions of cold gas particles ($T < 30000$K) within G-clus at various epochs.

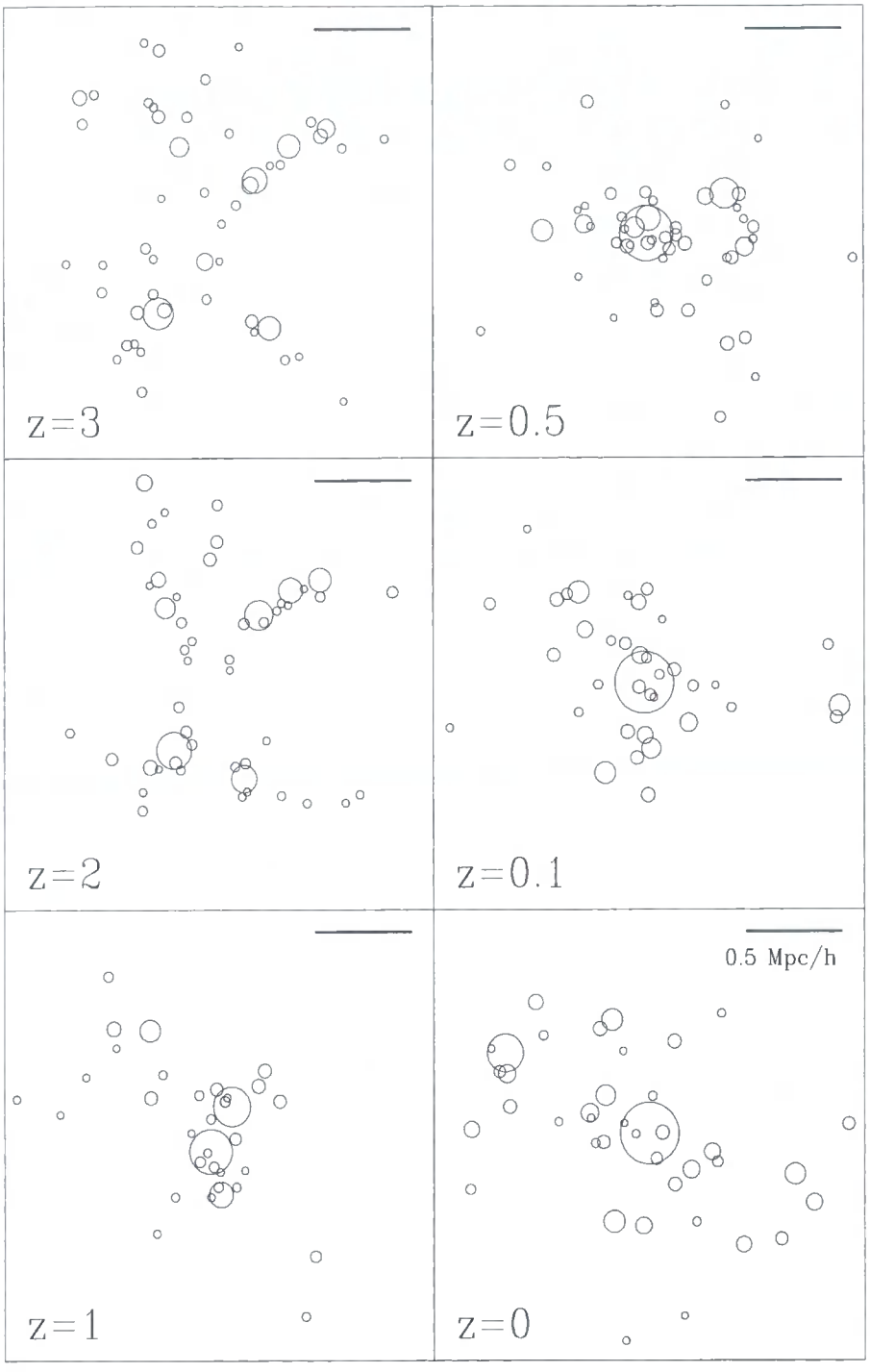


Figure 6.6: Projected positions of the galaxies within G-clus at various epochs.

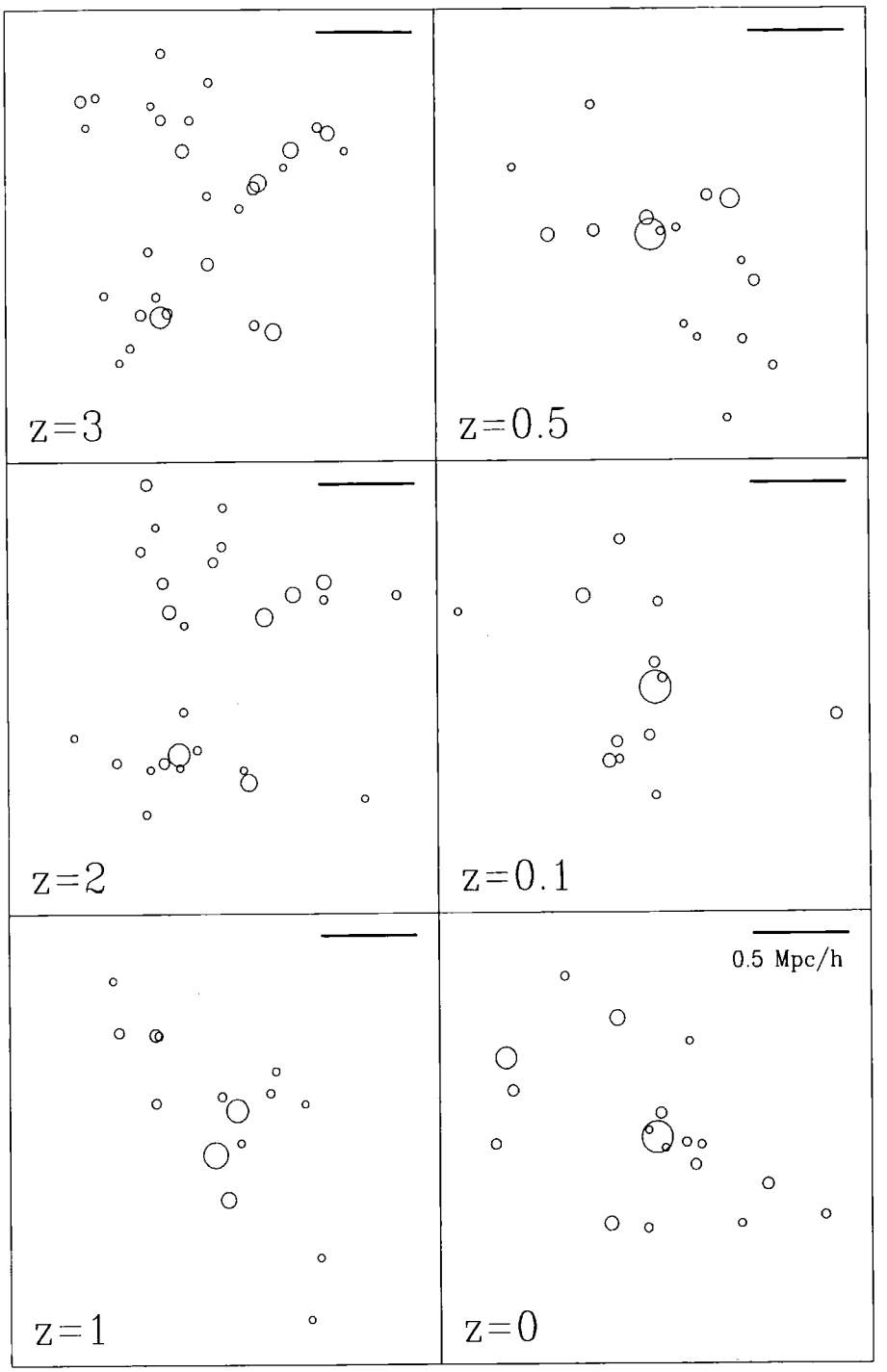


Figure 6.7: Projected positions of the galaxies within S-clus at various epochs.

The equivalent plot for the hot gas is given in Figure 6.4. Clearly the largest haloes at each redshift are traced by hot gas. This is because the gas is shock heated in these objects, which have considerably longer cooling times than the smaller haloes, where most of the gas will have cooled. The hot gas also exhibits a much smoother distribution than the dark matter, due to the isotropic nature of the pressure forces which is supporting the gas from further gravitational collapse.

Figure 6.5 illustrates the evolution of the cold gas in G-clus. At $z = 3$, some of the gas is diffuse (yet to collapse) while some of it has cooled within small haloes, forming the tight knots identified with the galaxies. By $z = 1$ almost all of the diffuse gas has been shocked above 30000K, and subsequent evolution is determined by the rate at which galaxies are falling into and merging within the cluster potential.

Projected positions of the galaxies themselves are illustrated in Figure 6.6 for G-clus and Figure 6.7 for S-clus. Shown are circles that are centred on the centre of mass of each object with radii proportional to the cube root of their mass. Note that the larger galaxies in both plots (particularly the most massive at the centre of the clusters) are broken into smaller objects at higher redshifts, following the merger history of their parent haloes in Figure 6.3.

There are two important differences between the galaxy populations in G-clus and S-clus. Firstly, at each epoch, the number of galaxies is considerably higher in G-clus than in S-clus (by $z = 0$, there are 40 galaxies in G-clus and 18 within S-clus). Secondly, many of the galaxies in G-clus can be identified with less-massive counterparts in S-clus, at all times. Both differences can be explained by the effects of feedback in S-clus. The feedback model reduces the mass of each object by reheating gas into the halo before it can form more stars. Although the feedback is more effective in smaller mass haloes (where the virial temperature is generally smaller), larger mass galaxies formed themselves from mergers of smaller objects. The number of galaxies decreases since many of the galaxies which are near the resolution limit (32 particles) in G-clus, are unresolved in S-clus.

6.3.3 Bulk properties

Table 6.1 summarises the bulk properties of the three clusters at $z = 0$. Comparing the virial mass of each cluster to the original cluster ($M_{\rm vir} = 1.49 \times 10^{14} h^{-1}\,{\rm M}_{\odot}$), all of the values agree to within a few per cent. The values of $N_{\rm dark}$ are all extremely close, increased by a factor of ~ 6.3 over the original simulation. The number of baryon particles ($N_{\rm gas} + N_{\rm stars}$) show much greater variations between each simulation, also evident in the values of $f_{\rm bary}$. For A-clus, $f_{\rm bary}$ is around 15 per cent lower than the assumed global value ($\Omega_b/\Omega_0 = 0.1$), due to the gas being

Table 6.1. Bulk properties of the resimulated clusters at $z = 0$. In each column, the following quantities are listed: (1) the simulation label; (2) the virial mass in $10^{14}h^{-1}\,\mathrm{M}_{\odot}$; (3) the number of dark matter particles; (4) the number of gas particles; (5) the number of star particles; (6) the bolometric X-ray luminosity in $10^{44}h^{-2}\,\mathrm{erg\,s^{-1}}$; (7) the emission-weighted X-ray temperature in keV; (8) the fraction of the total mass in baryons; (9) the fraction of baryons in gas with $T > 30000\mathrm{K}$; (10) the fraction of baryons in gas with $T < 30000\mathrm{K}$; (11) the fraction of baryons in stars.

Simulation	M_{vir}	N_{dark}	N_{gas}	N_{stars}	L_{x}	kT_{x}	f_{bary}	f_{hot}	f_{cold}	f_{stars}
A-clus	1.47	66870	56801	0	0.58	1.55	0.086	1.00	0.00	0.00
G-clus	1.45	66869	48178	0	0.88	1.14	0.074	0.26	0.74	0.00
S-clus	1.47	66876	41871	16259	2.19	0.88	0.088	0.71	0.01	0.28

more extended than the dark matter, resulting from a net transfer of energy in the merging process (e.g. Navarro, Frenk, & White 1995). However, in G-clus, where the gas is allowed to cool, the fraction drops from 8.6 per cent of the total mass to 7.4 per cent; the effects of cooling are counter-intuitive. Including feedback increases the f_{bary} to a value comparable to the non-radiative result.

The fraction of baryons in hot and cold gas again varies from simulation to simulation. For A-clus, all of the baryons have temperatures above 30000K. From a simple argument based on spherical symmetry and isothermality, the minimum resolved halo (assuming 32 gas and dark matter particles) should have a virial temperature, $T \sim 8.5 \times 10^4\mathrm{K}$. Hence, all of the baryons, including those in the small haloes, have been shocked well above 30000K. In comparison, the hot gas fraction in G-clus has decreased to only a quarter of the total value. From Figure 6.4 it is clear that all baryons in haloes within the cluster, with the exception of the atmosphere of the cluster halo itself, have cooled. This is an enormously large fraction, clearly demonstrating the overcooling problem in the absence of feedback. The fractions are almost reversed for S-clus, where ~ 71 per cent of the baryons is in hot gas. Out of the remaining fraction, only ~ 1 per cent has not formed stars.

Bolometric X-ray luminosities are estimated using the formalism of Navarro, Frenk & White (1995)

$$L_{\mathrm{x}} = 3.05 \times 10^{31} \sum_i \rho_i T_i^{1/2} h^{-2}\,\mathrm{erg\,s^{-1}}, \tag{6.1}$$

where the sum extends over all hot gas particles within the virial radius, and ρ_i and T_i are their densities (in units of the mean gas density) and temperatures (in Kelvin) respectively. The inclusion of radiative cooling increases L_{x} for the cluster by around 35 per cent. The increase in emission comes from a strong cooling flow at the centre of the cluster, onto the central galaxy. This effect

was also observed by several other groups (e.g. Katz & White 1993; Suginohara & Ostriker 1998; Lewis et al. 1999), however it is argued by Pearce et al. (1999) that the cooling rate of the hot gas around the central galaxy is too high, due to SPH overestimating their densities. Pearce et al. show that when the hot gas is fully decoupled from the cold galactic gas (see also Chapter 4, Comparison 4), the X-ray luminosity *decreases* when compared to the case without cooling.

The feedback model makes the cluster considerably brighter (around a factor of 3 or so). This is due to a small group of particles that have cooled onto the central cluster galaxy and have subsequently been reheated and temporarily prevented from cooling back down. Hence this gas is both hot and dense enough to dominate X-ray emission. Such strong emission is not seen within such galaxies, and it clear that the feedback model is not representing a physically reasonable scenario. Switching off the cooling is done mimick the effects of the hot diffuse medium that is created by supernovae in the multiphase ISM (McKee & Ostriker 1977). Hence the densities should not really be used in the calculation, whilst the gas is prevented from cooling.

Emission-weighted temperatures are calculated as

$$kT_{\rm x} = 8.62 \times 10^{-8} \frac{\sum_i \rho_i T_i^{3/2}}{\sum_i \rho_i T_i^{1/2}} \ {\rm keV}, \tag{6.2}$$

again summing over all hot gas particles within the virial radius of the cluster. The value of $kT_{\rm x}$ calculated for G-clus is around 25 per cent lower than the value for A-clus, due to the inclusion of cooling. The emission-weighted temperature drops in G-clus, relative to A-clus because the dominant contribution from gas in G-clus comes from particles with lower temperatures than the virial temperature of the cluster, since they are within a cooling flow. In S-clus, the emission-weighted temperature is even lower. As was discussed above, the emission comes primarily from reheated particles within the central galaxy, which are colder than the particles contributing to most of the emission in G-clus.

6.3.4 Density profiles

Figure 6.8 illustrates both dark matter and (hot) gas density profiles at $z = 0$ for all three simulations. Each profile was constructed by dividing the cluster into 50 spherical shells, such that the thickness of each shell, $\Delta \log r$ =constant. The radius is scaled to the virial radius and the density is in units of the mean density of each species. The softening scale and the virial radius are also marked with vertical lines.

For the dark matter, all profiles agree extremely well except within the inner 10 per cent of the virial radius. The profile steepens slightly for S-clus compared to A-clus and the effect is more

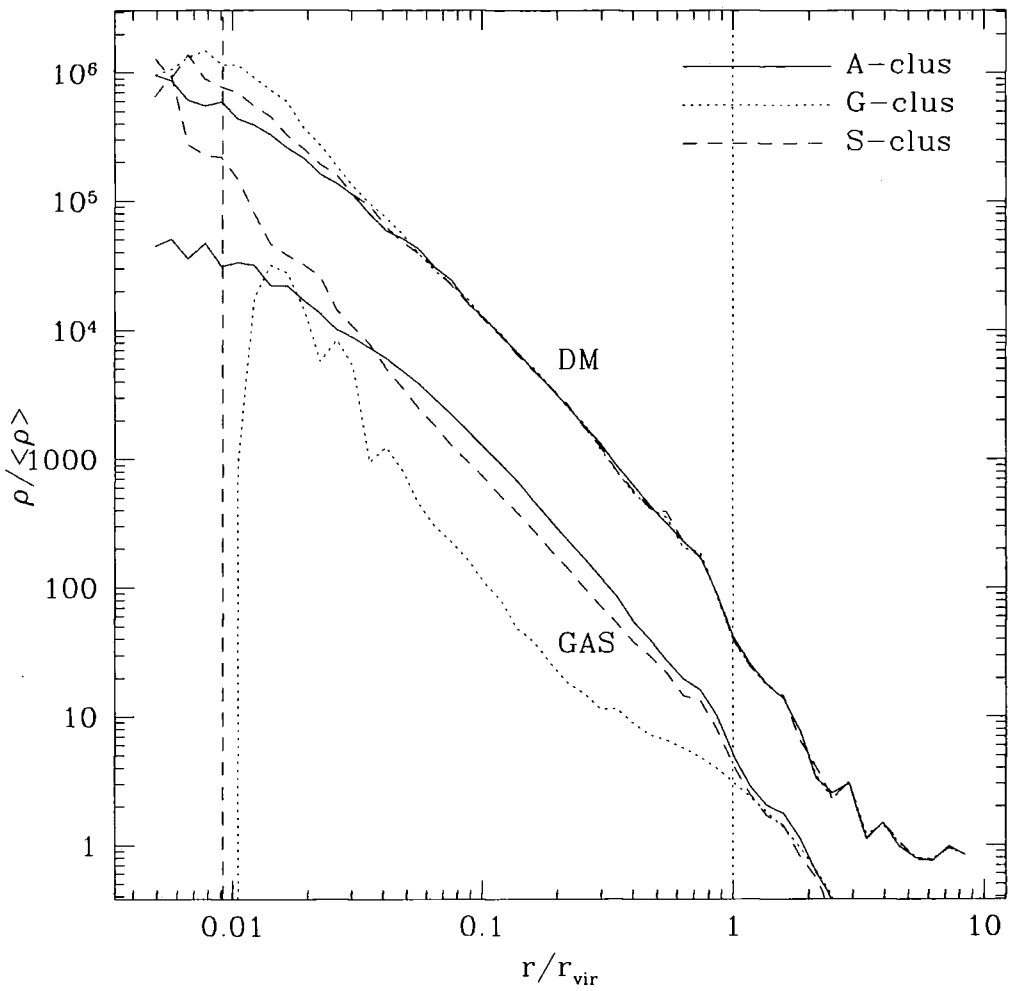


Figure 6.8: Density profiles at $z = 0$, plotted in units of the mean density of each species. The vertical dashed line is the softening scale while the vertical dotted line marks the virial radius.

pronounced for G-clus. The presence of the central galaxy has deepened the central potential of the cluster, which has drawn in more dark matter particles.

Significant variations are present in the density profiles of the gas. First of all, the density profile for A-clus is flatter than its dark matter equivalent, for the inner 10 per cent of the virial radius. As discussed above, a net transfer of energy occurred from the dark matter to the gas within mergers. This caused gas in the centre to adiabatically expand, pushing the rest of the gas to larger radii.

The gas density profile of G-clus is considerably different from that of A-clus. From around the virial radius, the G-clus profile becomes considerably flatter (and hence is lower than the A-clus profile), before turning back up at around 20 per cent of the virial radius. Further in, it becomes even steeper, reaching a similar density as A-clus at around 2 per cent of the virial radius, before

dropping to zero in the centre. The same characteristic differences are noted in Pearce et al. (1999). Radiative cooling provides a sink for hot gas in the centre of the cluster, allowing gas to flow in from outside the cluster. The gas that flows in has a higher entropy than the gas it replaces, hence reducing the density. However the effects begin to reverse when the cooling time becomes short enough for the gas to lose a considerable fraction of its entropy within the lifetime of the cluster. This transition occurs around where the slope of the profile becomes steeper. Further in, the increased cooling rate causes the density profile to rise dramatically until eventually, at the centre, all the hot gas has cooled.

The gas density profile for S-clus is much straighter (and closer to the A-clus profile), for radii greater than a few percent of the virial radius, when compared to the previous two. The feedback model has reduced the amount of high entropy gas falling into the cluster. The large increase within the inner few per cent of the virial radius is due to the reheated gas. As discussed above, gas particles with the highest density are within the central galaxy, resulting from the cooling of the hot halo gas onto the object.

6.3.5 Temperature profiles

Figure 6.9 illustrates the temperature profiles of the hot gas at $z = 0$. The shape of the A-clus profile is that expected for a non-radiative simulation (e.g. Navarro, Frenk, & White 1995; Eke, Navarro, & Frenk 1998). The temperature at the virial radius is around a factor of 2 lower than the peak value. For G-clus, the temperature profile is higher than for A-clus, before declining and eventually dropping below it then going to zero. Again, the difference is caused by the inflow of higher entropy gas. Gas flowing through the virial radius is less dense than the gas it was replacing, hence it is heated to a higher temperature to maintain pressure support. The temperature starts decreasing around where the slope of the density profile steepens, as the gas is able to radiate a significant fraction of its energy within the lifetime of the cluster. A similar, though less pronounced effect is present for the S-clus temperature profile; the feedback model has reduced the effect of the cooling flow on the hot gas.

6.3.6 The mass distribution of galaxies

Figure 6.10 shows cumulative mass functions for the galaxies in G-clus and S-clus at $z = 0$. Since the former has more than a factor of 2 more objects, the amplitude of the function is correspondingly higher. The most massive galaxy in G-clus has a baryon mass of $\sim 4.6 \times 10^{12} h^{-1}\,\mathrm{M}_{\odot}$, which is reduced in S-clus to $\sim 6.3 \times 10^{11} h^{-1}\,\mathrm{M}_{\odot}$, nearly an order of magnitude. The former is mass

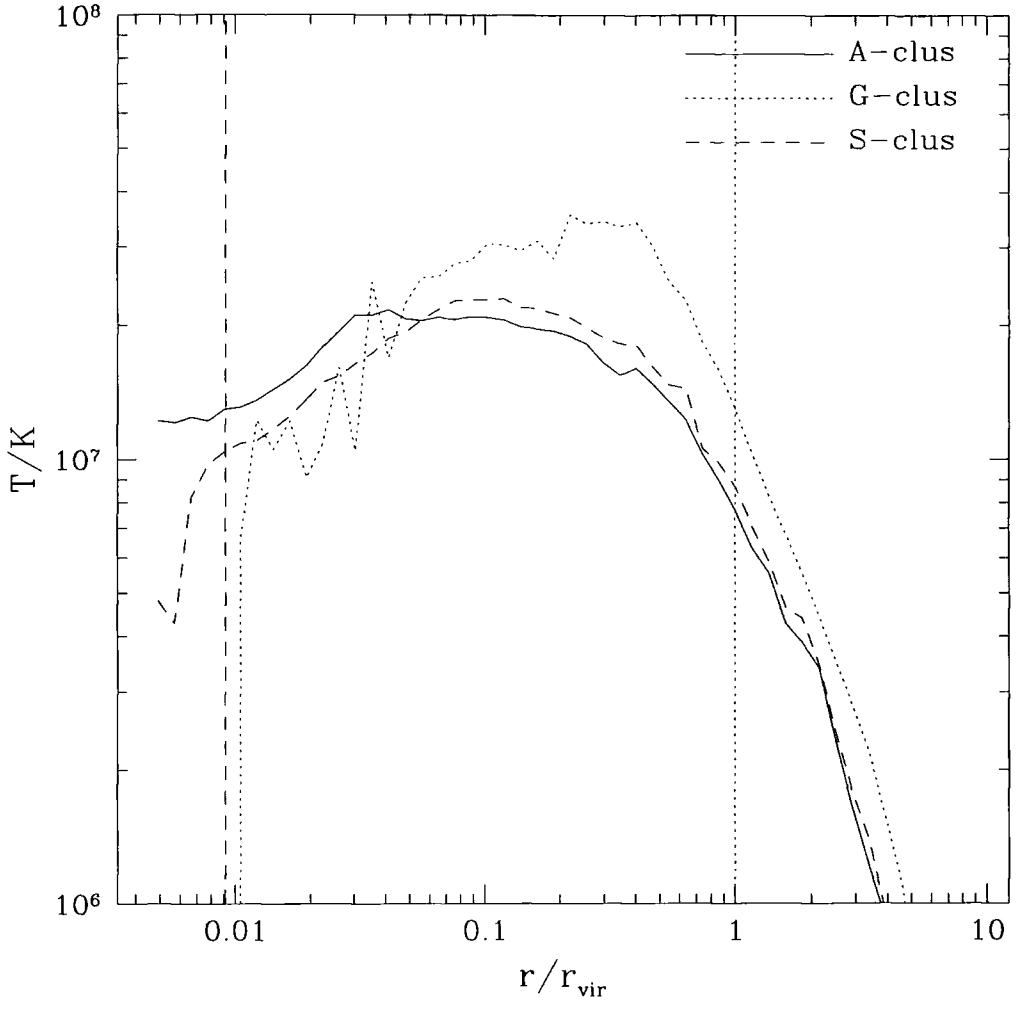


Figure 6.9: Gas temperature profiles for the clusters at $z = 0$. The vertical dashed line is the softening scale while the vertical dotted line marks the virial radius.

is similar to that observed by Katz & White (1993), who stress the effects of *overmerging* in their simulation. However, a significant fraction of the central object has come from the large cooling flow (c.f. Comparison 4 in Chapter 4, where the largest galaxy had its mass reduced by 2/3 when the hot gas was decoupled from the cold gas). Hot gas in S-clus will not cool onto the central galaxy so quickly since its density is decoupled from the cold gas (which is negligible anyway since the galaxy mainly consists of stars). However, as noted earlier, the mass of the central galaxy was significantly reduced at an earlier time, before it formed from the mergers of smaller objects.

The shape of the two mass functions are very similar, allowing for statistical fluctuations, particularly at the high mass end. Hence, the effect of feedback is to *rescale* the mass function in a mass independent manner (for the range of masses studied).

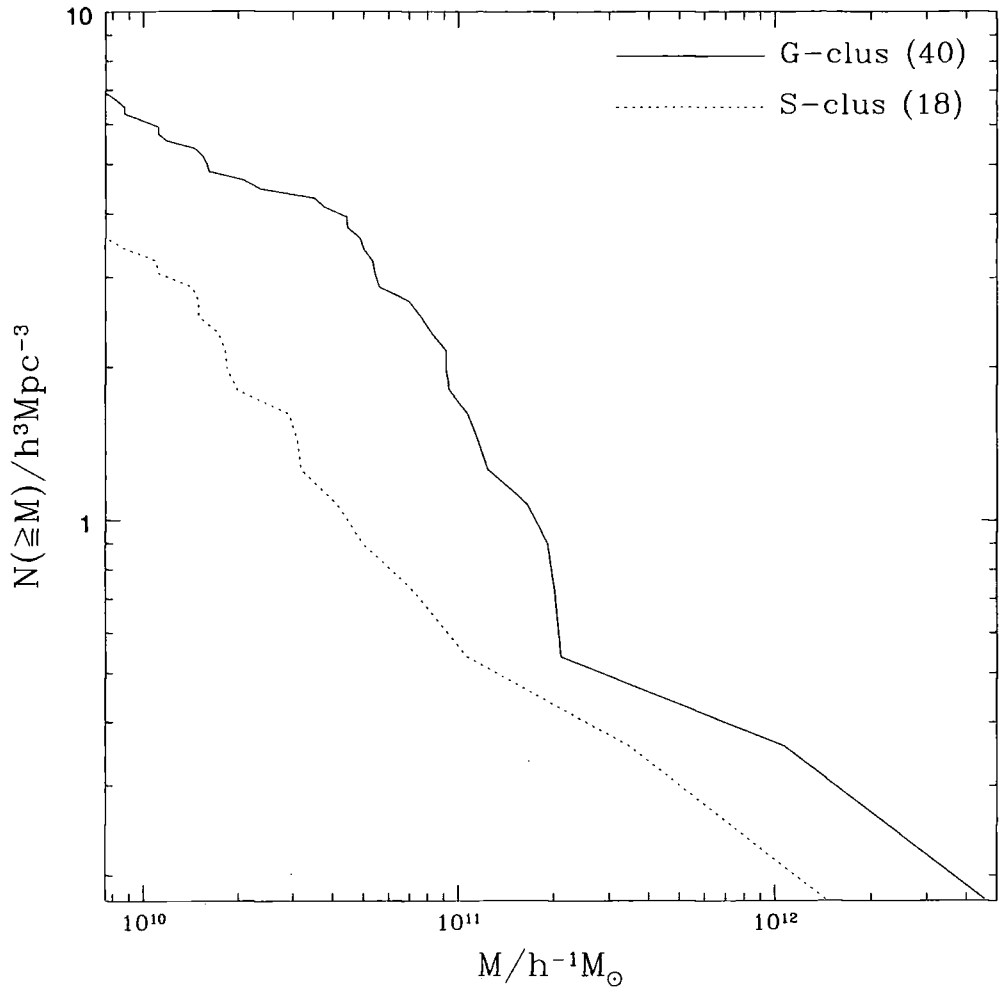


Figure 6.10: Cumulative mass functions of the galaxies in G-clus and S-clus at $z = 0$.

Another difference between the two simulations (evident when comparing Figure 6.10 to Columns 11 & 12 in Table 6.1) is that the fraction of galaxy phase material in bound objects is significantly different. For G-clus, the fraction is 0.99 where for S-clus, the fraction 0.65. The lower value of the former is due to the relative ease with which star particles are stripped from galaxies compared to the gas particles. This is due to the former not being able to dissipate energy when perturbed by their neighbours. Part of this effect may be due to the effects of feedback and part will be due to the interaction of galaxies within the harsh environment of the cluster. The latter may be enhanced by the large softening scale in the simulations and it is important that future simulations will investigate the sensitivity of the rate at which stars are stripped to the resolution of the simulation.

6.4 Summary

In this chapter, preliminary results are presented for three high resolution simulations of a single galaxy cluster. A cluster with a circular velocity of 750 $\mathrm{km\,s^{-1}}$ was selected from a previously run collisionless cold dark matter simulation of a random volume within a flat, low density universe. The object was then resimulated using techniques that allows the cluster to be simulated with higher resolution, while degrading the resolution of the remaining volume, used to provide accurate tidal forces. The first resimulation explicitly included baryons, able to adiabatically heat and cool and additionally shock-heat via collisional interactions. The second resimulation additionally allowed the gas to radiatively cool, using the cooling function of Sutherland & Dopita (1993) with a fixed gas metallicity, $Z = 0.3Z_{\odot}$. Finally, the third resimulation included a model for star formation and feedback, namely that used in the TWOPHASE2 simulation studied in Chapter 5. Main conclusions follow.

1. The virial mass of all three simulations are the same as the original cluster to within a few per cent. The discrepancies are due to fluctuations in the mass in baryons, whose fraction is around 20 per cent lower than the global value.

2. Almost 75 per cent of the gas within the virial radius of the cluster had cooled by $z = 0$, demonstrating the cooling catastrophe present in all hierarchical models without a non-gravitational heat source. Including star formation and feedback lowered this value by over a factor of 2, with nearly all the gas in stars.

3. Including radiative cooling increased the bolometric luminosity of the cluster by ~ 35 per cent and decreased the emission-weighted temperature by ~ 25 per cent. A strong cooling flow onto the central galaxy was responsible for both changes. As discussed in Pearce et al. (1999), decoupling the cold gas from the hot gas reduces the cooling flow, tending to cause the opposite changes in the luminosity and emission-weighted temperature. The feedback model caused a further increase in the X-ray luminosity by a more than a factor of two, as well as a 30 per cent decrease in the emission-weighted temperature. Both effects are due to a subset of (unphysically) hot dense particles reheated from stars within the central galaxy.

4. All dark matter density profiles agree very well except within the inner few per cent of the virial radius, where a small degree of steepening occurs due to the amount of cold dense gas and stars at the centre of the cluster. The (hot) gas profile differs significantly between radiative and non-radiative models. As found by Pearce et al., cooling causes material from outside the

cluster to be drawn in, which has a higher entropy. Consequently, the density is significantly lower in the outer parts of the cluster but its gradient steepens considerably at smaller radii, where cooling is more efficient. Introducing feedback produces a similar profile to the non-radiative cluster, except with a significantly higher density within a few per cent of the virial radius. Again, this is due to dense particles being reheated by stars in the central galaxy.

5. Significant differences are seen in the hot gas temperature profiles. When the gas is allowed to cool (with no star formation or feedback), the higher entropy material causes temperature to be higher when compared to a simulation without cooling, outside a few per cent of the virial radius. For the inner portion, the temperature drops below the non-radiative case due to the rise in the cooling rate. The cluster with feedback exhibits similar, albeit less pronounced behaviour than the radiative case.

6. An extremely massive central galaxy is produced in the simulation with cooling, containing over 40 per cent of the gas that lies within the virial radius of the cluster. The mass of this galaxy is reduced by around a factor of 10 when star formation and feedback is incorporated.

7. The $z = 0$ mass function of galaxies within the cluster is similar in shape for both runs with cooling, with the run including feedback lower in amplitude due to the total number of galaxies having decreased from 40 to 18. The feedback is effective in reducing the mass, particularly in small objects, which consequently reduce the mass of the galaxies to values below the resolution limit of the simulation.

References

Aarseth S. J., 1963, MNRAS, 126, 223

Anninos P., Norman M. L., 1996, ApJ, 459, 12

Bryan G. R., Norman M. L., 1998, ApJ, 495, 80

Butcher H., Oemler J., A., 1984, ApJ, 285, 426

Cen R., Ostriker J. P., 1994, ApJ, 429, 4

Copi C. J., Schramm D. N., Turner M. S., 1995, ApJ, 455, 95

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Dressler A., 1980, ApJ, 236, 351

Eke V. R., Cole S. M., Frenk C. S., 1996, MNRAS, 282, 263

Eke V. R., Navarro J. F., Frenk C. S., 1998, ApJ, 503, 569

Evrard A. E., 1990, ApJ, 363, 349

Fabian A. C., Nulsen P. E. J., Canizares C. R., 1991, A&AR, 2, 191

Frenk C. S., Evrard A. E., White S. D. M., Summers F. J., 1996, ApJ, 472, 460

Frenk C. S. et al., 1999, ApJ, 525, 554

Gunn K. F., Thomas P. A., 1996, MNRAS, 281, 1133

Jenkins A. et al., 1998, ApJ, 499, 20

Katz N., White S. D. M., 1993, ApJ, 412, 455

Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188

Lewis G. F., Babul A., Katz N., Quinn T., Hernquist L., Weinberg D. H., 1999, ApJ, submitted (astro-ph/9907097)

McKee C. F., Ostriker J. P., 1977, ApJ, 218, 148

Mushotzky R., Loewenstein M., Arnaud K. A., Tamura T., Fukazawa Y., Matsushita K., Kikuchi K., Hatsukade I., 1996, ApJ, 466, 686

Navarro J. F., Frenk C. S., White S. D. M., 1995, MNRAS, 275, 720

Pearce F. R., Thomas P. A., Couchman H. M. P., Edge A. C., 1999, MNRAS, submitted (astro-ph/9912013)

Peebles P. J. E., 1970, AJ, 75, 13

Peebles P. J. E., 1980, The large scale structure of the Universe. Princeton University Press

Poggianti B. M., Smail I., Dressler A., Couch W. J., Barger A. J., Butcher H., Ellis R. S., Oemler J., A., 1999, ApJ, 518, 576

Steinmetz M., White S. D. M., 1997, MNRAS, 288, 545

Suginohara T., Ostriker J. P., 1998, ApJ, 507, 16

Thomas P. A., Couchman H. M. P., 1992, MNRAS, 257, 11

Viana V. L., Liddle A. R., 1996, MNRAS, 281, 323

White S. D. M., 1976, MNRAS, 177, 717

White S. D. M., Navarro J. F., Evrard A. E., Frenk C. S., 1993, Nature, 366, 429

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Chapter 7

Conclusions

In this thesis, several issues concerning the properties of galaxies and clusters within hierarchical models of structure formation were addressed.

Chapter 2 applied a simple phenomenological model, based on analytical scaling relations, to X-ray observations of clusters at low and high redshift. The aim of this chapter was to determine whether current measurements of the X-ray evolution of clusters are sufficient to discriminate between various open and flat cosmological scenarios. The model was developed as an extension to the entropy-driven model for cluster evolution (Bower 1997), incorporating the effects of a low-density universe. It was then applied to measurements of evolution in the X-ray luminosity function of clusters, as well as the normalisation of the luminosity-temperature relation.

It was found that the data is insufficient to constrain the value of the density parameter, Ω_0 (for open and flat models), due to the degeneracy between Ω_0 and n, the effective slope of the power spectrum on cluster scales. To break the degeneracy, two independent methods were considered: a theoretical relation between n and Ω_0 from the CDM model and an observational measurement of n from the APM galaxy survey (Tadros, Efstathiou, & Dalton 1998). The first method gives $\Omega_0 < 0.55$ ($0.1 < \Omega_0 < 0.7$ for $\Omega_0 + \Lambda_0 = 1$), while the second method gives $\Omega_0 < 0.6$ ($\Omega_0 < 0.65$ for $\Omega_0 + \Lambda_0 = 1$), with both results at the 95 per cent confidence level. Tighter constraints on the cosmological parameters will require, in particular, the amount of evolution in the bright end of the luminosity function to be measured. Forthcoming X-ray surveys of clusters with the *Chandra* and *XMM* satellites should fulfill this task.

The remainder of this thesis concentrated on using the method of cosmological simulation, a self-consistent approach to modelling galaxies and clusters within a robust cosmological framework. Details of the method were outlined in Chapter 3, describing the techniques that were used for generating initial conditions, as well as the algorithms used to calculate gravitational and hydrodynamical forces.

In Chapter 4, a systematic study was undertaken of the effects of varying a subset of numerical and physical parameters, within cosmological simulations of galaxy formation. A simple fiducial model was devised, to investigate the evolution of CDM and baryon density fluctuations within a small volume. Besides gravity, the baryons were also able to adiabatically heat and cool, shock heat and radiatively cool. The latter process gives rise to clumps of cold dense material, which were identified as "galaxies". The fiducial simulation was compared to 24 variants, each differing in the value of one parameter. Various properties of the simulations were then compared, such as the mass and spatial distributions of the galaxies and the amount of gas that cools within the dark matter haloes.

The results were stable for the range of parameter space that was explored, with the exception of the parameters that affect the cooling rate of the gas. Values for some of these parameters can be justified by other means (e.g. the global baryon fraction, from nucleosynthesis constraints) however some are an inherent part of the numerical model, controlling the resolution of the simulation. As the resolution is increased, smaller mass haloes are resolved, within which the cooling times are generally shorter, leading to a *cooling catastrophe*, in which the amount of cooled gas in the simulations far exceeds observed values. This problem is common to all hierarchical models of galaxy formation (White & Rees 1978), with the most popular solution being the inclusion of additional energy input from the supernovae of stars (feedback). Future simulations with higher resolution will require a model for feedback, not only to avoid overcooling but also to enable numerically convergent results to be obtained.

In Chapter 5, a first step was taken to include the effects of star formation and feedback within cosmological simulations. The first part of the chapter explored the overcooling problem in detail, by varying the resolution of the simulation, the box-size and the ratio of dark matter to baryon particle masses (for a fixed baryon fraction). The amount of gas cooled by $z = 0$ was in excess of observed values for nearly all of the simulations studied, even when the dark matter particles were more massive than the critical threshold for artificial two-body heating (Steinmetz & White 1997). For simulations with resolutions above this threshold, varying the box-size or the mass ratio of dark matter to baryon particle masses made no significant difference to the cooling efficiency of the gas.

The chapter continued, with an examination of various methods for turning gas into stars in the simulations. A total of 12 simulations were compared at fixed resolution, testing 6 different methods. It was found that a simple prescription, based on the densities and temperatures of the gas particles was sufficient to obtain a reasonable star formation rate. Additional assumptions, such as attempting to identify regions that are Jeans unstable, are unnecessary. The ratio of cold

gas to stars within the galaxies are sensitive to the densities where star formation is able to occur. For a low ratio of gas to stars, the star formation rate is primarily controlled by the cooling rate of the gas.

One of the simplest prescriptions for star formation (similar to that used by Pearce 1998) was then used to investigate the effects of feedback. Both thermal and kinetic feedback models were considered, using a set of 9 simulations. Kinetic feedback was able to reduce the amount of cold gas and stars in the simulations, however this feedback significantly affected the binding energy of the galaxies, destroying most of the small mass objects. Thermal feedback could not reduce the cooling rate of the gas due to the high densities enabling them to rapidly radiate the energy away. This problem was circumvented by preventing reheated gas from cooling for a short period of time. The shape of the galaxy mass function and the star formation rate for this model are both preserved.

Finally, Chapter 6 presented preliminary results for simulations of a single, modest-sized galaxy cluster, within a flat, low density universe ($\Omega_0 = 0.3, \Lambda_0 = 0.7, H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$). Three simulations were performed in this study: a non-radiative model and two radiative models, with and without star formation and feedback. Various properties were analysed, such as the density and temperature profiles of the clusters and the mass distribution of galaxies in the radiative models. The feedback significantly reduced both the mass and number of galaxies formed in the cluster, without which nearly three-quarters of the gas had cooled by $z = 0$. All three simulations produced significantly different results in the central gas distribution. Hence, the X-ray properties of the cluster varied depending on the model used. However, it has to be stressed that more analysis is required before the results of this chapter are fully understood.

To conclude, work in this thesis has merely scratched the surface of a subject that will occupy cosmologists for many years to come. Although the basic tools for a viable model of structure formation are now in place, many of the intricate details that influence the evolution of structure on a wide range of scales have yet to be understood. As has been stressed in this thesis, an important example of this is feedback. A satisfactory explanation for how the effects of feedback are propagated from sub-galactic scales (the ISM) to the scales of galaxies and clusters remains elusive. However, the process is crucial for the survival of hierarchical models. It may explain many, if not all, of the model's shortcomings: the formation of disks and the angular momentum problem; the faint end of the galaxy luminosity function and the overcooling problem; the slope of the X-ray luminosity temperature relation and the core in their gas distributions etc..

Work is now in progress to extend the results found in this thesis. A more thorough investigation of the effects of feedback is under study, to determine to what extent it is able to resolve the aforementioned problems. Hence, the effects of feedback on other scales will also have to be considered. The success of a feedback model in Chapter 5 was gauged by the amount of cooled material in the volume at $z = 0$. However, it is not obvious that a model tuned to produce a specific stellar baryon fraction will solve, for example, the angular momentum problem in disk galaxies. Many of the complications in the modelling of feedback arise from the sensitivity of the methods to numerical resolution. Hence, from a numerical perspective, it is important to understand the effects of feedback models within simulations with varying resolution. Ultimately, a resolution-independent model for feedback will be crucial, so that future high resolution simulations achieve numerically convergent results, and that the physical interpretation of these models can be trusted.

References

Bower R. G., 1997, MNRAS, 288, 355

Pearce F. R., 1998, A&A, submitted (astro-ph/9803133)

Steinmetz M., White S. D. M., 1997, MNRAS, 288, 545

Tadros H., Efstathiou G., Dalton G., 1998, MNRAS, 296, 995

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Appendix A

X-ray evolution of clusters: results from individual XLF datasets

The five figures in this appendix illustrate constraints on the free parameters, n and ϵ, of the entropy evolution model discussed in Chapter 2. Again, the results are shown for seven cosmological scenarios, the $\Omega = 1$ model and both open and flat models with $\Omega_0 = (0.5, 0.3, 0.1)$. Constraints from the evolution of the $L_x - T$ normalisation are applied to all figures (dotted contours), while the XLF constraints are from individual datasets (solid contours). Figure A.1 illustrates constraints using the SHARC dataset (Burke et al. 1997), Figure A.2 the WARPS dataset (Jones et al. 1998), Figure A.3 the RDCS dataset (Rosati et al. 1998), Figure A.4 the EMSS 1 dataset (Henry et al. 1992) and Figure A.5 the EMSS 2 dataset (Luppino & Gioia 1995). Shaded regions illustrate areas of parameter space consistent with both the $L_x - T$ and XLF data at the 68 per cent level. The vertical solid line is the measured value of n from the CDM theoretical calculation, while the vertical dashed line is the measured value of n from the APM rich cluster survey (Tadros, Efstathiou & Dalton 1998). Vertical dashed lines illustrate the range of n consistent with the Γ constraint set by Eke et al. 1998, using the evolution of the temperature function (Figure 2.6).

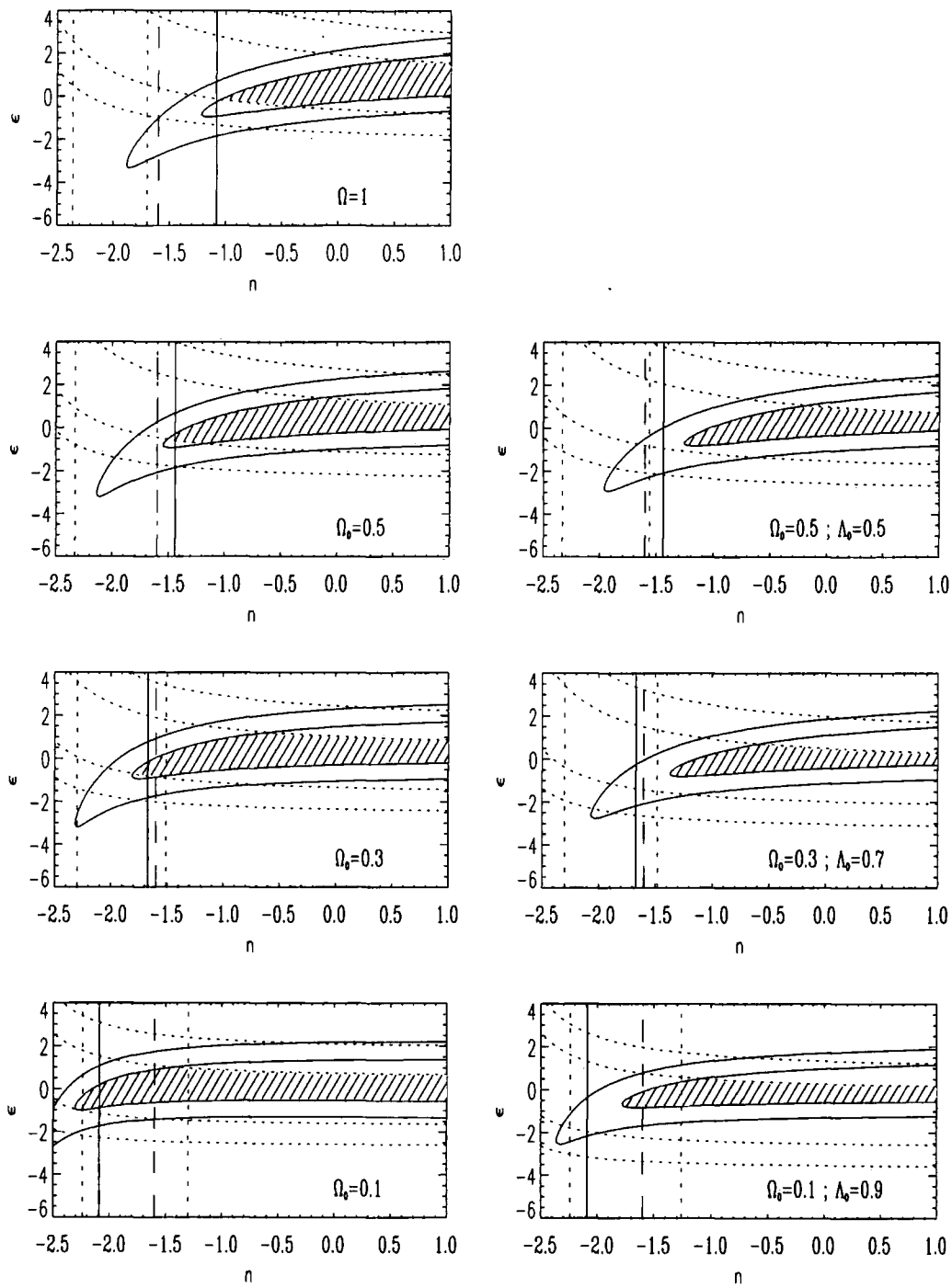


Figure A.1: Model constraints using the XLF data fromt the SHARC survey (Burke et al. 1997).

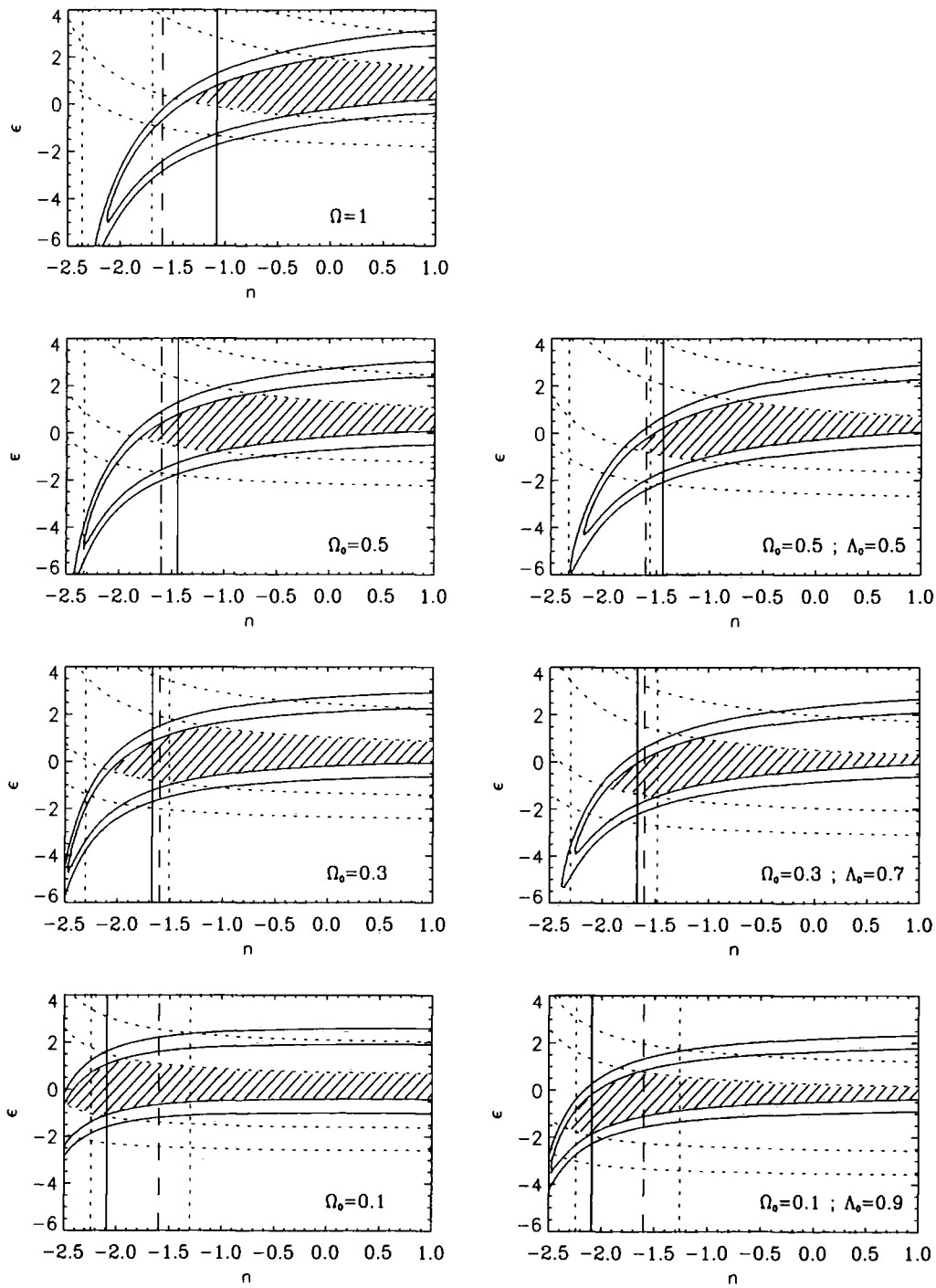


Figure A.2: Model constraints using the XLF data from the WARPS survey, given by Jones et al. (1998).

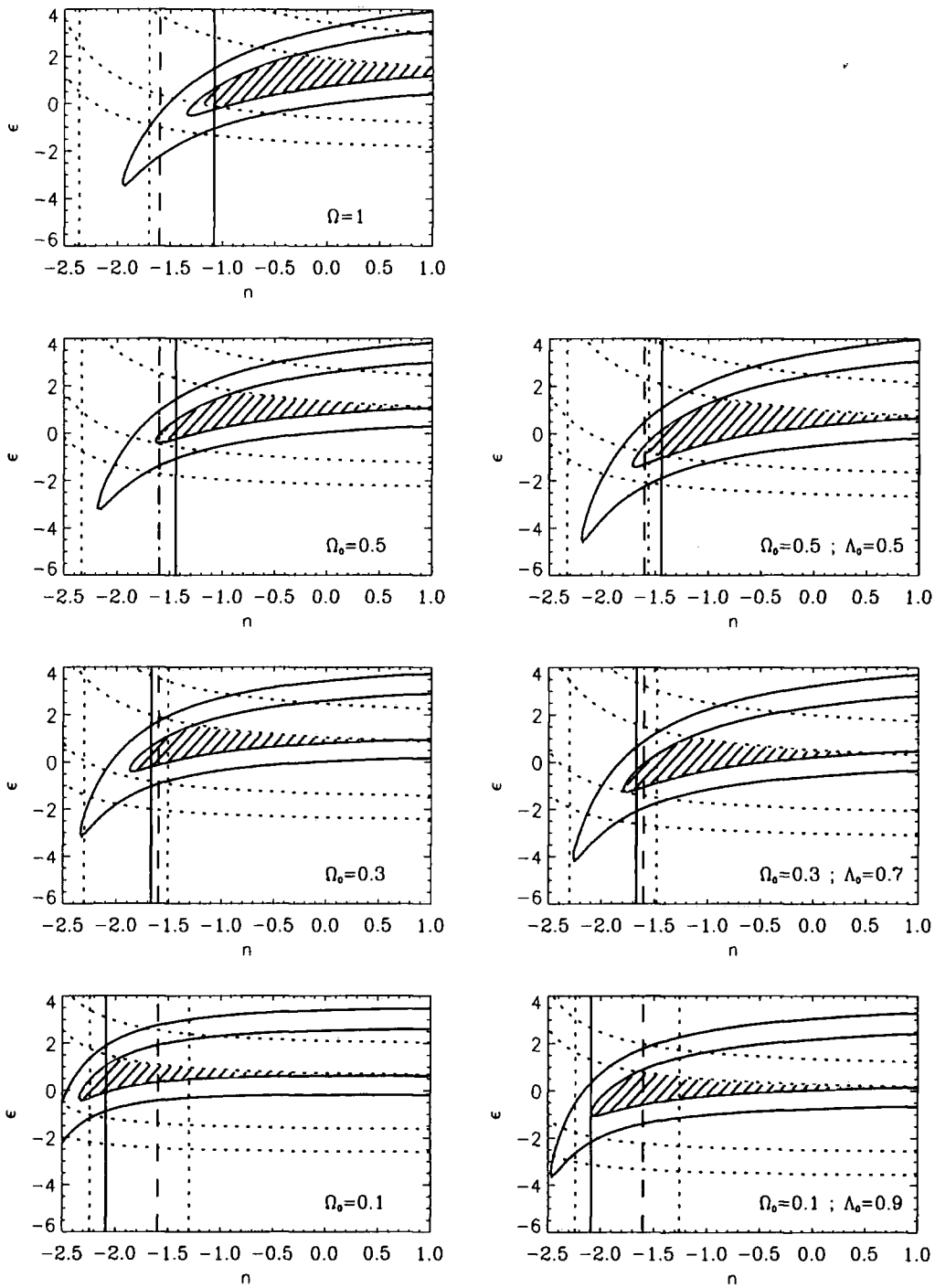


Figure A.3: Model constraints using the XLF data from the RDCS survey, given by Rosati et al. (1998).

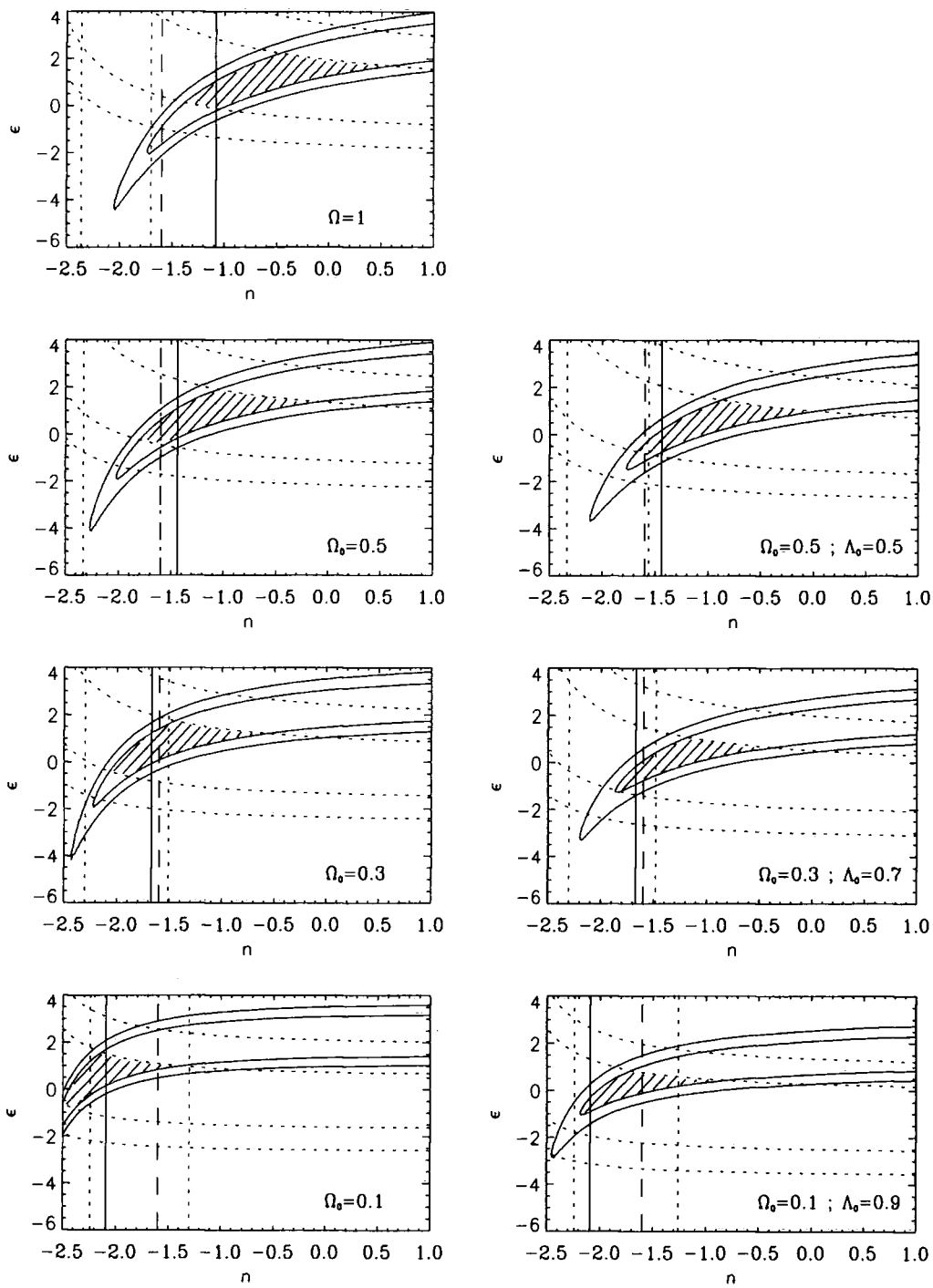


Figure A.4: Model constraints using the XLF data from the EMSS survey, given by Henry et al. (1992).

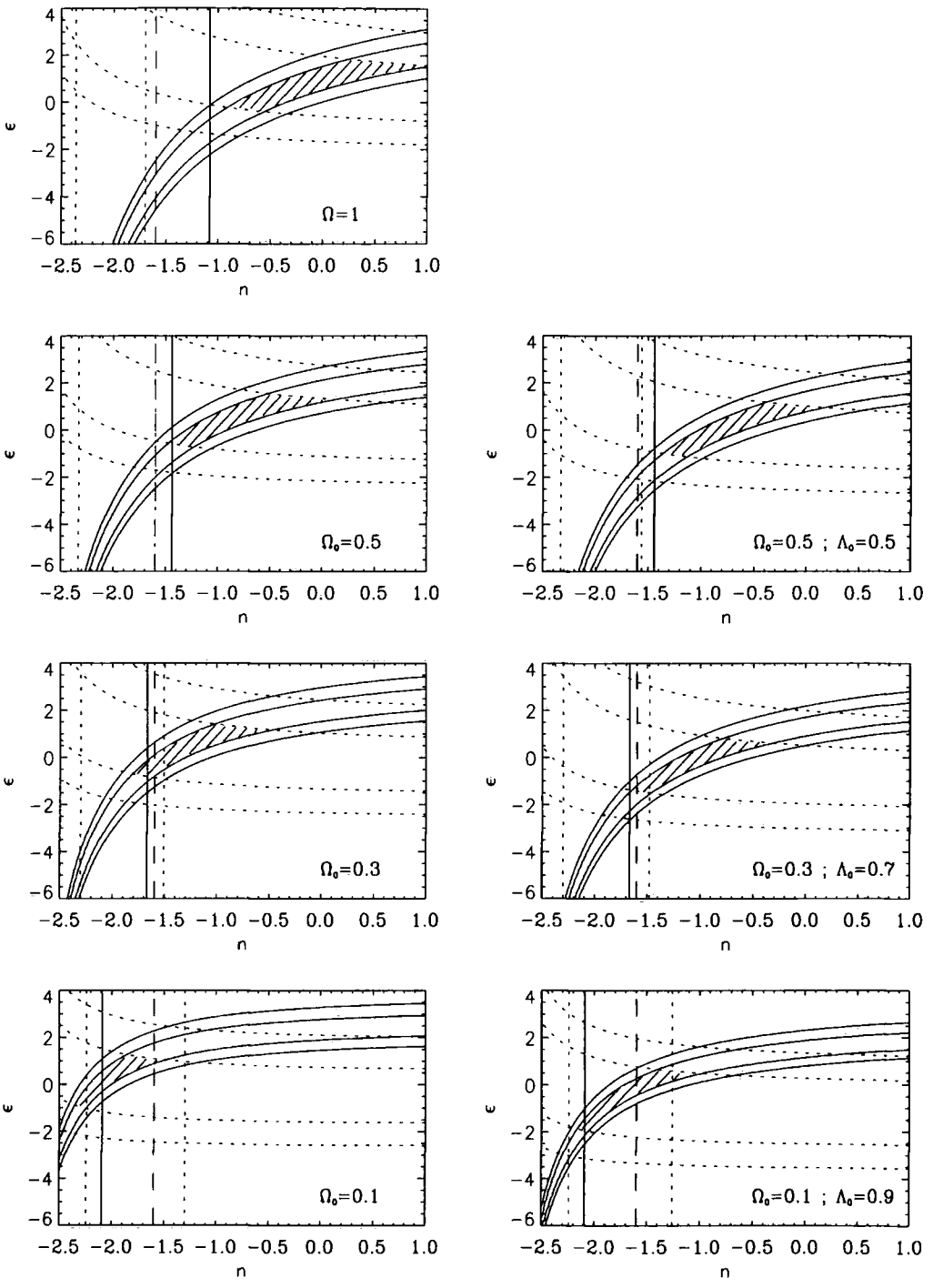


Figure A.5: Model constraints using the XLF data from the EMSS distant cluster sample, given by Luppino & Gioia (1995).