THESIS

# COSMOLOGICAL IMPLICATIONS OF THE ABUNDANCES OF CLUSTERS OF GALAXIES

Tetsu Kitayama

Dissertation submitted to
Department of Physics, School of Science,
The University of Tokyo

December, 1997

# Abstract

This thesis studies a number of implications of the abundances of clusters of galaxies on cosmology. We assess theoretically the cluster abundances based mainly on the Press-Schechter formalism, and attempt to constrain viable cosmological models from the latest and future observations. We discuss in detail theoretical uncertainties that may change our predictions. For this purpose, extensions of the original PS formalism are also incorporated so as to take into account explicitly the formation and subsequent evolution of clusters.

Particular attention is paid to the number counts, so called the log$N$–log$S$ relation, of clusters of galaxies, in addition to the conventional X-ray temperature and luminosity functions. Since the log$N$–log$S$ data are determined more directly from observations with a minimum of assumptions, they attain smaller statistical and systematic uncertainties. In fact, the log$N$–log$S$ relation recently obtained by the *ROSAT* satellite surveys (Rosati et al. 1995, 1998; Ebeling et al. 1997, 1998) comprises the largest and most complete samples of clusters of galaxies to date. We find that there exists a set of cosmological models that reproduces remarkably well the above log$N$–log$S$ data. Such models agree simultaneously with the observed present-day X-ray temperature and luminosity functions. By comparing our theoretical predictions with the observations, we derive constraints on the density parameter $\Omega_0$ and the fluctuation amplitude $\sigma_8$. If our results are combined with the observations of the cosmic microwave background radiation by the *COBE* satellite, low-density cold dark matter models with $\Omega_0 \sim 0.3 - 0.5$ are strongly favoured. The limits derived from the log$N$–log$S$ relation and the X-ray temperature function are consistent with each other, while the former is more stringent due to the better statistical significance of the observed data.

Given rather tight constraints on $\Omega_0$ and $\sigma_8$ from the currently observed soft X-ray log$N$–log$S$ relation, we discuss further implications of the cluster abundances, including the possibilities to break the degeneracy of these parameters by future observations. Firstly, we combine the cluster number counts at different wavelengths in several X-ray and submm bands. This is of great significance in relation to the future observations by *ASCA*, *AXAF*, and *Astro-E* in the hard X-ray bands, and *PLANCK* and *LMSA* in radio bands. We demonstrate explicitly that one can break the degeneracy between $\Omega_0$ and $\sigma_8$ by combining the X-ray and submm number counts. Secondly, we consider the redshift evolution of the cluster number counts and compare it with that of the X-ray brightest Abell-type clusters (Ebeling et al. 1996a,b). The results, although still preliminary, point to low-density universes with $\Omega_0 \sim 0.3$. Finally, we confirm that the contributions of clusters to the X-ray and submm background radiations are insignificant in any model compatible with the *ROSAT* soft X-ray log$N$–log$S$ relation.

# Contents


**1 Introduction** 4
- 1.1 Purpose . . . 4
- 1.2 Clusters of galaxies as a probe of cosmology . . . 4
- 1.3 Significance of the cluster abundance . . . 5
- 1.4 Methodology . . . 7

**2 Standard Picture of Structure Formation in the Universe** 9
- 2.1 Overview . . . 9
- 2.2 Homogeneous and isotropic universe . . . 11
- 2.3 Linear perturbation theory . . . 14
- 2.4 Characteristic scales for the structure formation . . . 16
- 2.5 Power spectrum of density fluctuations . . . 19
- 2.6 Spherical model of the non-linear collapse . . . 22

**3 Mass Function of Virialized Structures** 26
- 3.1 Press-Schechter theory . . . 26
  - 3.1.1 Derivation of the mass function . . . 26
  - 3.1.2 Validity of the Press-Schechter theory . . . 29
- 3.2 Excursion set approach . . . 30
  - 3.2.1 Solving the cloud-in-cloud problem . . . 31
  - 3.2.2 Merger probabilities . . . 32
  - 3.2.3 Estimate of halo formation epochs . . . 33
- 3.3 Formation and destruction rates . . . 34
  - 3.3.1 Motivation . . . 34
  - 3.3.2 Formal procedure . . . 36
  - 3.3.3 A realistic prescription . . . 37
  - 3.3.4 Estimate of halo formation epochs . . . 41

**4 X-ray Temperature and Luminosity Functions** 44
- 4.1 Observational backgrounds . . . 44
- 4.2 Theoretical formulation . . . 46
- 4.3 Physical descriptions of X-ray clusters . . . 47
  - 4.3.1 Characteristic time-scales . . . 47
  - 4.3.2 Temperature . . . 48
  - 4.3.3 X-ray emission . . . 50

4.3.4 Summary of parameters . . . . . 56
4.4 Basic features of the cluster mass function . . . . . 56
4.5 Predictions of the X-ray temperature and luminosity functions . . . . . 58
4.6 Sensitivity to intracluster gas parameters . . . . . 60
4.7 Effects of formation epochs . . . . . 64
4.8 Evolution of cluster number densities . . . . . 67

**5 Cluster Number Counts in X-ray and Radio Wavebands 72**
5.1 Theoretical formulation . . . . . 72
5.1.1 X-ray flux . . . . . 73
5.1.2 Radio flux — The Sunyaev-Zel'dovich effect — . . . . . 73
5.2 Predictions of the X-ray cluster log$N$–log$S$ relation . . . . . 76
5.3 Constraints on the CDM models . . . . . 78
5.4 Cluster number counts in different wavebands . . . . . 84
5.5 Evolution of cluster number counts . . . . . 90
5.6 Contribution to the X-ray and submm background radiations . . . . . 93

**6 Conclusions 97**

**Acknowledgments 100**

**A Kinematic properties of Expanding Universe 101**
A.1 Cosmological redshift . . . . . 101
A.2 Luminosity distance . . . . . 102
A.3 Angular diameter distance . . . . . 104

**B Fitting formulae for the Lacey & Cole distribution function 105**

**References 107**

# Chapter 1

# Introduction

## 1.1 Purpose

Our universe today contains a variety of structures; stars are accumulated in galaxies, and galaxies spread inhomogeneously to make up groups and clusters of galaxies. One of the main goals of modern cosmology is to understand the process of formation and evolution of these structures. Any theories of structure formation must then be examined with care by being confronted with a number of observational aspects. For this purpose, the abundances of clusters of galaxies serve as one of the most powerful probes of structure formation models. In the present thesis, we perform detailed comparisons between theoretical predictions and observations by means of the cluster abundances. Our principal aims are to clarify the implications that clusters of galaxies have on cosmology and to derive possible constraints that they assert on viable structure formation models.

## 1.2 Clusters of galaxies as a probe of cosmology

Clusters of galaxies are the largest virialized systems known in the universe. Their typical observed properties (Bahcall 1996 and references therein) are summarized in Table 1.1. They cover a wide range of mass scale from poor groups made up of only a few galaxies to rich clusters which contain over a thousand of galaxies. To our current knowledge of the universe, there is no fundamental physics that determines a-priori the mass scale of clusters, unlike that of galaxies. According to a standard picture of structure formation, the observed cosmic structures including clusters have grown via gravitational instability from tiny density fluctuations in the early universe. In this context, it is likely that the structures are built up successively from small objects, and this process, called the hierarchical clustering, happens to have now reached the scale of galaxy clusters. Clusters of galaxies are expected to be relatively young and growing with the age of the universe.

It has long been realized that clusters of galaxies are powerful tools for the study of galaxy formation, dark matter, large-scale structures, and cosmology. Their utility stems from several reasons. Firstly, owing to their large size and dynamical youth, clusters of galaxies well reflect the properties of the global universe. A typical cluster mass of order $10^{15}\,\mathrm{M}_{\odot}$ corresponds to the mass enclosed within a comoving length of order 10 Mpc in the nearly homogeneous universe. This means that clusters are closely related to the initial

Table 1.1: Typical observed properties of clusters of galaxies

| Property | Typical value |
|---|---|
| Total mass | $10^{14} \sim 10^{16}\,M_\odot^*$ |
| | dark matter $60 \sim 90\%$ |
| | intracluster gas $10 \sim 30\%$ |
| | galaxies and stars $\lesssim 10\%$ |
| Radius | $\sim$Mpc$^\dagger$ |
| Radial velocity dispersion | $400 \sim 1400$ km s$^{-1}$ |
| Gas temperature | $10^7 \sim 10^8$ K |
| X-ray luminosity | $10^{43} \sim 10^{46}$ erg s$^{-1}$ |
| Electron number density | $10^{-5} \sim 10^{-2}$ cm$^{-3}$ |
| Iron abundance | $\sim 0.3$ solar$^\ddagger$ |

* $1\,M_\odot = 1.989 \times 10^{33}$ g
† 1 pc $= 3.086 \times 10^{18}$ cm
‡ $(\mathrm{Fe/H})_\odot = 4.68 \times 10^{-5}$ in number

density fluctuation field, which is considered as the "seeds" of cosmic structures, at this length scale.

Secondly, physics that governs clusters of galaxies is relatively simple and well-understood. As will be discussed in §4.3, the intracluster gas is approximately in thermal and hydrostatic equilibria, and the radiative processes such as cooling and star formation are not very important. Recent intensive numerical simulations (e.g. Evrard 1990; Navarro, Frenk & White 1997; Bryan & Norman 1998; Yoshikawa, Itoh & Suto 1998) have succeeded in reproducing a number of observed features of clusters. Although there are several problems which still remain to be resolved, such as origins of metals and the formation of core regions, the overall properties of clusters are fairly well accounted for. Clusters of galaxies are therefore suited for calibrating the structure formation theories and related parameters.

Thirdly, there are several means to observe clusters of galaxies. Optical, radio, and X-ray observations have been developed to probe the intracluster medium and individual galaxies in clusters. More recently, the gravitational lensing effect has revealed itself as a useful tool of detecting the mass distribution within clusters. These observations are rapidly increasing today both in quality and quantity to give us better knowledge of clusters of galaxies.

## 1.3 Significance of the cluster abundance

By the term "cluster abundance" we mean the number of clusters of galaxies present in a certain volume in the universe. Most commonly used statistics include the X-ray temperature function (XTF) and the X-ray luminosity function (XLF), which describe the comoving

number density of clusters as a function of X-ray temperature and X-ray luminosity, respectively. We will also consider the number counts, the number of clusters in a unit solid angle, of clusters. As the number counts are usually measured as a function of flux $S$ and plotted in log scales, they are often called the log$N$–log$S$ relation.

Thanks to the rapid progress of observational techniques, size and quality of available cluster samples are dramatically improving in recent years. For instance, the local XLF data are now constructed from about 200 clusters (Ebeling et al. 1997), which is an order of magnitude larger in size than the previous ones (e.g. Henry & Arnaud 1991). Most of the statistics used to describe the cluster abundance have so far been obtained in X-ray wavebands. Since X-ray identifications of clusters of galaxies are largely free from the projection effect which notoriously plagues optically selected cluster catalogues, X-ray observations are suitable for constructing homogeneous cluster samples. Furthermore, it is very likely that homogeneous cluster samples will also become available in other wavebands in the near future. For example, the intracluster gas will scatter off the cosmic microwave background (CMB) photons, and this effect, called the Sunyaev & Zel'dovich (1972, hereafter SZ) effect, can be observed in radio bands. This effect has now been detected for several clusters from ground-based telescopes (e.g. Grainge et al. 1993; Birkinshaw & Hughes 1994) and will be measured for even larger number of clusters in the near future by, for example, the *PLANK* satellite mission. Since the SZ effect has an interesting characteristic that the surface brightness does not suffer the cosmological dimming effect (§5.1.2), it has a great potential of probing distant clusters.

What can we learn from the cluster abundances? The present-day abundance of clusters reflects the amplitude and shape of the power spectrum of density fluctuations on a scale of order 10 Mpc, and can be used to measure these quantities with a minimum of assumptions. This measure is of crucial importance in cosmology because the density fluctuation spectrum is the fundamental quantity that characterizes a structure formation scenario. It tells us not only about the structure formation processes themselves, but also about the origin of density fluctuations, which is closely related to the nature of the dark matter and the physics of the early universe.

The anisotropies in the CMB detected by the *COBE* satellite offer another independent way of measuring the amplitude and shape of the density fluctuation spectrum (e.g. Bunn & White 1997). It should be noted that *COBE* probes the scale greater than Gpc, which is over two orders of magnitude larger than the cluster scale. At scales of interest for the structure formation ( $\lesssim 10$ Mpc), the *COBE* normalized fluctuation spectrum depends largely on the adopted model parameters, such as the primordial spectral index. By contrast, the spectrum determined from the cluster abundance is fairly robust because clusters probe these scales more directly. These two methods are therefore complementary to each other in fixing the fluctuation spectrum.

In addition to their present-day abundance, one can also make use of the time-evolution of the abundances of galaxy clusters to further constrain cosmological models. The rate at which the number density of clusters change with time is sensitive to the rate of growth of density fluctuations. This in turn depends primarily on the mean density of the universe which governs the total gravitational force that affects the fluctuations. Consequently, the evolution of the cluster abundance gives us a clue to measuring the density parameter $\Omega_0$, which is one of the fundamental cosmological parameters.

## 1.4 Methodology

In this thesis, we assess theoretically the cluster abundances and attempt to deduce constraints on cosmology from the latest and future observations. The theoretical predictions require first to describe how clusters are formed from tiny density fluctuations and to relate the distribution of real clusters to the fluctuation spectrum. For this purpose, we mainly use an analytical formalism developed originally by Press & Schechter (1974, hereafter PS). Based on some simplified assumptions on the non-linear evolution of fluctuations, PS derived a simple formula for the mass function, the comoving number density as a function of mass, of bound virialized objects. For the identical purpose, one could alternatively perform intensive numerical simulations and trace the 3-dimensional development of clusters in a more straightforward manner. Numerical simulations, however, still suffer from limitations of computation time, small-scale resolutions, and a dynamic range. The PS formalism has a great advantage of making it possible to explore much wider range of the parameter space and to clarify the underlying key physics. In addition, the PS mass function has been extensively tested against numerical simulations and found to show good agreements over the mass range of our interest (e.g. Efstathiou et al. 1988; Carlberg & Couchman 1989; White, Efstathiou & Frenk 1993; Lacey & Cole 1994; Eke, Cole & Frenk 1996).

There are still some cautions regarding the PS theory. One of the most fundamental problems is that the dynamical history of each object, including its formation epoch, is not delineated by it. It is in fact of essential importance to trace such a history in order to take into account properly the evolutionary effects of individual objects. As will be shown in §4.3, for example, one needs to specify the formation epoch of a cluster to relate its mass to other properties, such as the temperature. This point has been neglected in previous analyses on the cluster abundances (e.g. Evrard & Henry 1991; Oukbir & Blanchard 1992, 1997; Eke, Cole & Frenk 1996; Barbosa et al. 1996; Colafrancesco et al. 1997). Therefore, in addition to the original PS theory, we incorporate some extended formalisms to describe explicitly the formation and subsequent evolution of virialized clusters (Kitayama & Suto 1996 a,b; see also Lacey & Cole 1993 and Sasaki 1994).

The above formalisms are combined with models of the intracluster gas to yield the abundance of clusters as a function of some observable, such as temperature or X-ray luminosity. In comparing these predictions with observation, we pay particular attention to the number counts, the $\log N$–$\log S$ relation, of clusters. In contrast to previous studies based mainly on the XTF or XLF (e.g. Evrard 1989; Frenk et al. 1990; Henry & Arnaud 1991; Bahcall & Cen 1993; Hanami 1993; White et al. 1993; Viana & Liddle 1996; Pen 1998), our approach has the following advantages. 1) Since the observed number counts are constructed from a larger number of clusters, they attain smaller statistical uncertainties. The $\log N$–$\log S$ relation recently obtained by the *ROSAT* satellite surveys (Rosati et al. 1995, 1998; Ebeling et al. 1997, 1998) comprises nearly 300 clusters in total, whereas the most commonly used XTF data (Henry & Arnaud 1991) are estimated from only about 25 clusters. 2) The cluster number counts are largely free from a bias that may contaminate the XTF. The observed XTF is likely to include preferentially luminous species, especially near its low temperature end, because the sample constructing it is selected on the basis of the X-ray flux, not the temperature. Such a bias is likely to become even stronger for distant cluster samples. 3) One can discuss the evolution of the cluster abundance in a more straightforward manner

using the number counts. In fact, the "observed" XLF and XTF in a distant universe are model-dependent statistics, because the translation of the observed X-ray flux to the absolute luminosity, and of the observed number per solid angle to the physical number density can only be done by assuming specific values for the fundamental cosmological parameters such as $\Omega_0$.

By means of our theoretical predictions, we first clarify the impacts that the latest observations have on cosmology. Constraints on cosmological parameters are derived from the above mentioned *ROSAT* $\log N$–$\log S$ relation, which is constructed from the largest and most complete samples of clusters to date. We then discuss implications that future observations may provide. The cluster abundances in different wavebands or at different epochs are combined to deduce further constraints on cosmology. Specifically, we extend our analysis to a radio band and predict the number of clusters expected to be observed by the SZ effect. The emissions from the intracluster gas in X-ray and radio bands are originated from completely different physical mechanisms; the former is mainly due to thermal bremsstrahlung, and the latter is due to the inverse-Compton scattering of the CMB photons. The cluster abundances in these bands will therefore behave in a very different manner. We also consider the time-evolution of cluster numbers in these bands and make suggestive predictions for the future observations.

This thesis is organized as follows. Chapter 2 briefly reviews a standard picture of structure formation. Chapter 3 describes the analytical formalisms for the mass function including the PS theory and its extensions. Chapter 4 presents the predictions for the XTF and XLF. Chapter 5 performs detailed comparisons between theory and observation for the X-ray $\log N$–$\log S$ relation to derive constraints on cosmology. Chapter 5 also discusses further implications of cluster abundances, such as the predictions in the radio band, the evolution of the cluster number counts, and the contribution of clusters to the X-ray background (XRB) and submm background radiation (SBR). Finally, Chapter 6 summarizes our conclusions.

# Chapter 2

# Standard Picture of Structure Formation in the Universe

At the bottom line of our analysis is a standard picture of structure formation in the universe; the observed cosmological structures including galaxies and their clusters grew from small density perturbations in the early universe. This chapter reviews briefly such a scenario in order to provide background knowledge for later discussions as well as to summarize relevant cosmological equations. More complete presentations of the structure formation theories can be found in many textbooks (e.g., Peebles 1980; Kolb & Turner 1990; Padmanabhan 1993; Peebles 1993).

## 2.1 Overview

Our universe today appears to be highly inhomogeneous with rich structures: stars are accumulated in galaxies, galaxies make up groups of galaxies, clusters of galaxies, and even superclusters. On the other hand, the observations of the cosmic microwave background radiation (CMB) support that the universe was homogeneous and isotropic to a good precision in the past (Smoot et al. 1992). These facts provide two stringent conditions that any plausible theories of structure formation should satisfy simultaneously. It is thus generally considered that the inhomogeneous cosmological structures have emerged out of primordial density perturbations (small enough not to violate the CMB observations) present in the early universe. The origin of such perturbations is often ascribed to quantum fluctuations produced during an inflation era or to topological defects generated during cosmological phase transitions (e.g. Kolb & Turner 1993).

Once the primordial fluctuations are produced, they start to evolve mainly due to the gravitational instability. As long as the fluctuations are small, their gravitational growth is well described by linear perturbation theory (§2.3). Details of the evolution in reality depend on the number and type of particles present in the universe, and several characteristic scales appear accordingly (§2.4) to determine the shape of the power spectrum of density fluctuations (§2.5). There is a good deal of observational evidence that the matter content of the universe is dominated by the invisible dark matter (e.g. Peebles 1993). Although it has not yet been possible to determine what the dark matter is specifically, one is able to investigate its nature from the structure formation scenarios. If, for instance, the dark matter is purely

baryonic, then it confronts a serious difficulty in accounting for both the observed CMB anisotropies and the large-scale structures, because the growth of the baryonic perturbations is suppressed until the photon-baryon decoupling epoch (§2.4). The baryon dominated universe is further in conflict with the predictions of the theories of Big Bang nucleosynthesis (e.g. Walker et al. 1991). On the contrary, if hot dark matter (HDM) with large velocity dispersions (e.g. light neutrinos) is dominant, it smears out small-scale perturbations and fails to account for the formation of sub-galactic scale objects. Alternatively, if the universe is made up of cold dark matter (CDM) with negligibly small velocity dispersions (e.g. axions or super-symmetric particles), the structure formation occurs from small scales and it seems to agree with a number of observational aspects. This is in fact why CDM models have been so widely discussed in the past decade. In the present thesis, we focus our attention on the CDM dominated universe.

As the perturbations continue to grow, they enter a non-linear regime, decouple from the cosmic expansion, contract, and collapse to form bound virialized halos (§2.6). In the CDM dominated universe, the amplitude of density perutubations increase with decreasing scales (§2.5). Thus virialized halos start to form from small scales. As time goes on, the perturbations go non-linear at larger and larger scales, producing successively bigger halos through tidal interaction and mergers of smaller halos. This process is called the hierarchical clustering. The development of structures in this context has been studied extensively by means of both numerical simulations (e.g. Davis et al. 1985; Navarro et al. 1996) and analytical formalisms (e.g. Monaco 1997 for review; see also Chapter 3).

The hierarchical clustering of non-linear structures mentioned above is simply a consequence of the gravitational interaction among collisionless dark matter particles which dominate the universe. However, in order to form any luminous observable structures, physics that governs baryons plays an essential role. In contrast to dark matter, baryonic gas feels pressure that opposes condensation on very small scales. Thus there emerges a *minimum* mass scale for the aggregation by gravity of baryonic gas along with the dark matter, which is essentially the Jeans mass (eq. [2.46]). Once baryons contract into bound systems above this scale, they are shock-heated to high temperature, converting the kinetic energy into the thermal energy. The heated baryonic gas then loses energy via radiative cooling. In order to form stars or other compact objects, it is necessary for the gas to cool efficiently to low temperature. Alternatively, if the cooling is not efficient, the hot gas simply remains in equilibrium. This cooling condition gives rise to a *maximum* mass scale for the gas condensation and the subsequent star formation (Silk 1977; Rees & Ostriker 1977). The two scales mentioned above roughly account for the observed lower and upper limits in the mass of galaxies. As the hierarchical clustering proceeds, galaxies once formed will be incorporated into even larger halos to build up groups or clusters of galaxies. Since clusters are in general larger than the maximum mass scale for the gas condensation, radiative cooling and star formation are relatively unimportant in the intracluster gas (see also §4.3).

## 2.2 Homogeneous and isotropic universe

In order to study the dynamics of the global universe, we first consider a homogeneous and isotropic universe. It is described by a metric of the Robertson-Walker form:

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -c^2\mathrm{d}t^2 + a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right], \tag{2.1}$$

where $g_{\mu\nu}$ is the metric tensor, $c$ is the speed of light, $a(t)$ is the cosmic scale factor, and $k$ is the spatial curvature constant. Note that equation (2.1) is invariant under the scaling $a(t) \to \xi a(t), r \to r/\xi, k \to \xi^2 k$, where $\xi$ is an arbitrary constant. This allows us to normalize $k$ or $a(t)$ arbitrarily. Thus, unless otherwise stated explicitly, we set

$$a_0 \equiv a(t_0) = 1, \tag{2.2}$$

where a subscript 0 denotes the present value.

The energy-momentum tensor of the homogeneous and isotropic universe should take the same form as for a perfect fluid (e.g. Weinberg 1972):

$$T^\mu{}_\nu = \mathrm{diag}(-\rho c^2, p, p, p), \tag{2.3}$$

where $\rho c^2$ is the energy density and $p$ is the pressure.

Now the Einstein field equations read

$$R^\mu{}_\nu - \frac{1}{2}Rg^\mu{}_\nu + \Lambda g^\mu{}_\nu = \frac{8\pi G}{c^4}T^\mu{}_\nu, \tag{2.4}$$

where $R^\mu{}_\nu$ is the Ricci tensor, $R \equiv R^\mu{}_\mu$ is the scalar curvature, and $\Lambda$ is the cosmological constant. Substituting equations (2.1) and (2.3) into equation (2.4) with some rearrangement yields

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda c^2}{3}, \tag{2.5}$$

$$\dot{\rho}c^2 = -\frac{\dot{a}}{a}(\rho c^2 + p), \tag{2.6}$$

where a dot stands for the derivative with respect to time $t$. Let us here define some parameters as follows :

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)} \quad : \text{Hubble parameter}, \tag{2.7}$$

$$\Omega(t) \equiv \frac{\rho(t)}{\rho_{\mathrm{c}}(t)} \equiv \frac{8\pi G\rho(t)}{3H^2(t)} \quad : \text{density parameter}, \tag{2.8}$$

$$\lambda(t) \equiv \frac{\Lambda c^2}{3H^2(t)} \quad : \text{dimensionless cosmological constant}, \tag{2.9}$$

$$K(t) \equiv \frac{kc^2}{a^2(t)H^2(t)} \quad : \text{curvature parameter}, \tag{2.10}$$

where $\rho_{\mathrm{c}}(t)$ is called the critical density of the Universe. Using these parameters, equation (2.5) can be rewritten as

$$\Omega(t) + \lambda(t) - K(t) = 1. \tag{2.11}$$

Evaluating this relation at the present epoch $t_0$ gives a relation:

$$\Omega_0 + \lambda_0 - K_0 = 1. \tag{2.12}$$

Using this relation, equation (2.5) becomes

$$\left(\frac{H}{H_0}\right)^2 = \Omega_0 \frac{\rho}{\rho_0} + \frac{1 - \Omega_0 - \lambda_0}{a^2} + \lambda_0. \tag{2.13}$$

Equation (2.13) shows explicitly that the dynamics of the universe is determined by $\Omega_0$, $\lambda_0$, and $H_0$, in addition to the energy density $\rho(t)$ governed by equation (2.6).

For an equation of state of the form $p = w\rho c^2$, equation (2.6) yields $\rho \propto a^{-3(1+w)}$. Since the content of the universe is roughly divided into non-relativistic matter ($w = 0$) and radiation ($w = 1/3$), we have

$$\begin{aligned} \rho(t) &= \rho_{\rm m}(t) + \rho_{\rm r}(t) \\ \rho_{\rm m}(t) &= \rho_{\rm m0}\ a^{-3}(t) = \rho_{\rm m0}\ (1+z)^3 \ : \text{non-relativistic matter}, \\ \rho_{\rm r}(t) &= \rho_{\rm r0}\ a^{-4}(t) = \rho_{\rm r0}\ (1+z)^4 \ : \text{radiation}, \end{aligned} \tag{2.14}$$

where $z \equiv a_0/a - 1$ is the cosmological redshift (see also Appendix A.1). The present universe is dominated by non-relativistic matter (baryons and dark matter) with a small fraction of radiation (photons and massless neutrinos) as

$$\begin{aligned} \rho_0 &= \rho_{\rm m0} + \rho_{\rm r0} \simeq \rho_{\rm m0} \\ \rho_{\rm m0} &= 1.88 \times 10^{-29} \Omega_{\rm m0} h^2 \ \ {\rm g\ cm^{-3}}, \\ \rho_{\rm r0} &= 7.86 \times 10^{-34} T_{2.73}^4 \ \ {\rm g\ cm^{-3}}, \end{aligned} \tag{2.15}$$

or equivalently,

$$\begin{aligned} \Omega_0 &= \Omega_{\rm m0} + \Omega_{\rm r0} \simeq \Omega_{\rm m0} \\ \Omega_{\rm m0} &= \frac{\rho_{\rm m0}}{\rho_{\rm c0}} \\ \Omega_{\rm r0} &= \frac{\rho_{\rm r0}}{\rho_{\rm c0}} \simeq 4.18 \times 10^{-5} T_{2.73}^4 h^{-2}, \end{aligned} \tag{2.16}$$

where $h \equiv H_0/(100\ {\rm km\ sec^{-1}\ Mpc^{-1}})$ is the dimensionless Hubble constant, $\rho_{\rm c0} = 3H_0^2/8\pi G = 1.88 \times 10^{-29} h^2$ g cm$^{-3}$ is the present-day critical density, and $T_{2.73} \equiv T_{\gamma 0}/(2.73\ {\rm K})$ is the normalized present-day CMB temperature. Current observations suggest $\Omega_0 \sim 0.1 - 1$ and $h \sim 0.5 - 1$ (e.g. Peebles 1993), while theories of the Big Bang nucleosynthesis (e.g. Walker et al. 1991) predict the present-day baryon density of

$$\Omega_{\rm B} \equiv \frac{\rho_{\rm B0}}{\rho_{\rm c0}} \simeq (0.01 - 0.03) h^{-2}. \tag{2.17}$$

This points to that the total density of the universe is dominated by dark matter. Equations (2.14)–(2.16) indicate that the radiation and matter should have had the equal energy density at some epoch in the past:

$$1 + z_{\rm eq} \sim \frac{\Omega_{\rm m0}}{\Omega_{\rm r0}} \sim 2.4 \times 10^4 \Omega_0 h^2 T_{2.73}^{-4} \tag{2.18}$$

Inserting equations (2.14)–(2.16) into (2.13), one obtains the dynamical equation of the universe:

$$\left(\frac{H}{H_0}\right)^2 = \frac{\Omega_{\rm r0}}{a^4} + \frac{\Omega_{\rm m0}}{a^3} + \frac{1-\Omega_0-\lambda_0}{a^2} + \lambda_0. \tag{2.19}$$

Equation (2.19) can be solved for $a(t)$ in some specific cases. In the radiation dominated era ($z > z_{\rm eq}$), the scale factor $a \simeq z^{-1}$ becomes very small, and the r.h.s. of equation (2.19) is governed by the first term, yielding

$$a \simeq (4\Omega_{\rm r0})^{1/4}\,\tau^{1/2}, \tag{2.20}$$

where $\tau \equiv H_0 t$ is the dimensionless cosmic time. In the matter dominated era ($z < z_{\rm eq}$), it is convenient to consider the following two regimes separately. First, as long as $a(t)$ is sufficiently small ($0 \ll z < z_{\rm eq}$), the r.h.s. of equation (2.19) is dominated by the second term and we have

$$a \simeq \left(\frac{9}{4}\Omega_0\right)^{1/3}\tau^{2/3}. \tag{2.21}$$

Second, at more recent epochs ($z \sim 0$), the last two terms start to contribute and the solution of equation (2.19) varies with $\Omega_0$ and $\lambda_0$. We show below the exact solutions in some representative cases that are of cosmological interest:

1. The Einstein-de Sitter model: $\Omega_0 = 1$, $\lambda_0 = 0$

$$a = \left(\frac{3}{2}\tau\right)^{2/3}. \tag{2.22}$$

2. Open models without cosmological constant: $\Omega_0 < 1$, $\lambda_0 = 0$

$$a = \frac{\Omega_0}{2(1-\Omega_0)}(\cosh\eta - 1), \qquad \tau = \frac{\Omega_0}{2(1-\Omega_0)^{3/2}}(\sinh\eta - \eta). \tag{2.23}$$

3. Flat models with positive cosmological constant: $\Omega_0 < 1$, $\lambda_0 = 1 - \Omega_0$

$$a = \left(\frac{\Omega_0}{1-\Omega_0}\right)^{1/3}\sinh^{2/3}\left(\frac{3\sqrt{1-\Omega_0}}{2}\tau\right). \tag{2.24}$$

Models with zero curvature, i.e. $\Omega_0 + \lambda_0 = 1$ (models 1 and 3), are mainly favored from a theoretical point of view (e.g. inflationary scenario, Sato 1981; Guth 1981), whereas low density universes, i.e. $\Omega_0 < 1$ (models 2 and 3), are supported by observations (e.g. Peebles 1993). In the rest of this thesis, we restrict our attention to the above three cosmological models.

An important scale in cosmology is the Hubble time, or equivalently the Hubble horizon radius, defined by

$$\begin{aligned} t_{\rm H}(t) &\equiv \left(\frac{a}{\dot{a}}\right) = H(t)^{-1}, \\ d_{\rm H}(t) &\equiv c t_{\rm H}(t) = cH(t)^{-1}. \end{aligned} \tag{2.25}$$

the former is the time-scale over which the size of the universe changes, and the latter sets the length-scale within which physical processes operate coherently. The latter is also the scale at which general relativistic effects become important (e.g. Peebles 1980); at the length much smaller than $d_H(t)$, the Newtonian approximation works adequately well. From equation (2.19), the Hubble time and the Hubble horizon radius (in comoving coordinates defined by eq. [2.29]) at some characteristic epochs are

- Matter-radiation equality: $z_{\rm eq} \sim 2.4 \times 10^4 \Omega_0 h^2 T_{2.73}^{-4}$

$$\begin{aligned} t_{\rm H}(t_{\rm eq}) &\sim 3 \times 10^3 (\Omega_0 h^2)^{-2} T_{2.73}^6 \ {\rm yr}, \\ \frac{d_{\rm H}(t_{\rm eq})}{a(t_{\rm eq})} &\sim 20 (\Omega_0 h^2)^{-1} T_{2.73}^2 \ {\rm Mpc}, \end{aligned} \tag{2.26}$$

- Photon-baryon decoupling: $z_{\rm dec} \sim T_{\rm dec}/T_{\gamma 0} \sim 1100 (T_{\rm dec}/3000{\rm K}) T_{2.73}^{-1}$

$$\begin{aligned} t_{\rm H}(t_{\rm dec}) &\sim 3 \times 10^5 (\Omega_0 h^2)^{-1/2} \left( \frac{T_{\rm dec}}{3000{\rm K}} \right)^{1/2} T_{2.73}^{-1/2} \ {\rm yr}, \\ \frac{d_{\rm H}(t_{\rm dec})}{a(t_{\rm dec})} &\sim 90 (\Omega_0 h^2)^{-1/2} \left( \frac{T_{\rm dec}}{3000{\rm K}} \right)^{-1/2} T_{2.73}^{1/2} \ {\rm Mpc}, \end{aligned} \tag{2.27}$$

- Present: $z = 0$

$$\begin{aligned} t_{\rm H}(t_0) &\sim 10^{10} h^{-1} \ {\rm yr}, \\ d_{\rm H}(t_0) &\sim 3000 h^{-1} \ {\rm Mpc}, \end{aligned} \tag{2.28}$$

where the photon-baryon decoupling epoch $t_{\rm dec}$ is the time at which photons and baryons lose thermal contact at the temperature $T_{\rm dec} \sim 3000$K, and we have used $z \sim a^{-1} \propto T_\gamma$ (see Appendix A). The Hubble times listed above give rough estimates of the age of the universe at these epochs apart from a numerical factor.

## 2.3 Linear perturbation theory

Having described the dynamics of the global universe, we next study the gravitational evolution of density fluctuations. As we are mainly interested in the sub-horizon scale fluctuations, the Newtonian gravity is applied. For simplicity, we approximate the cosmological matter as a perfect fluid without dissipation. The dissipational effects will be discussed separately in §2.4.

In the present analysis, we work in a coordinate $\boldsymbol{x}$ comoving with the expansion (or contraction) of the global universe, which is related to the physical coordinate $\boldsymbol{X}$ via

$$\boldsymbol{X} = a(t)\, \boldsymbol{x}. \tag{2.29}$$

This yields the following relation for the velocity:

$$\dot{\boldsymbol{X}} = \dot{a}\boldsymbol{x} + a\dot{\boldsymbol{x}} = H(t)\, \boldsymbol{X} + \boldsymbol{v}(\boldsymbol{x}, t), \tag{2.30}$$

where the first term on the r.h.s. stands for the Hubble flow induced by the cosmic expansion, and the second term is the peculiar velocity. In this coordinate, the density contrast and its Fourier decomposition are defined by

$$\begin{aligned}\delta(\boldsymbol{x},t) &\equiv \frac{\rho(\boldsymbol{x},t)-\bar{\rho}(t)}{\bar{\rho}(t)},\\ &\equiv \frac{1}{(2\pi)^3}\int \delta_k(t)\exp\left(-i\boldsymbol{k}\cdot\boldsymbol{x}\right)\mathrm{d}^3k \end{aligned} \tag{2.31}$$

where $\rho(\boldsymbol{x},t)$ is the density at position $\boldsymbol{x}$ at time $t$, $\bar{\rho}(t)$ is the mean density of background universe, and $\boldsymbol{k}$ is the comoving wave vector which is related to the comoving length scale of $2\pi/|\boldsymbol{k}|$. Note that, since $\delta(\boldsymbol{x})$ is real, $\delta_k$ is complex but satisfies $\delta_k^* = \delta_{-k}$. In the Einstein-de Sitter universe, for instance, $\bar{\rho}(t)$ in the matter dominated era is evaluated from equations (2.8) and (2.22) as

$$\bar{\rho}^{\mathrm{EdS}}(t) = \frac{1}{6\pi G t^2}. \tag{2.32}$$

For non-relativistic ideal (dissipationless) fluid, the equations for the perturbations in the comoving coordinate read

$$\frac{\partial\delta}{\partial t} + \frac{1}{a}\boldsymbol{\nabla}\cdot[(1+\delta)\boldsymbol{v}] = 0 \qquad : \text{continuity equation}, \tag{2.33}$$

$$\frac{\partial\boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} + \frac{1}{a}(\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{v} + \frac{1}{\rho a}\boldsymbol{\nabla}p + \frac{1}{a}\boldsymbol{\nabla}\phi = 0 : \text{Euler's equation}, \tag{2.34}$$

$$\boldsymbol{\nabla}^2\phi = 4\pi G\bar{\rho}\,a^2\delta \qquad : \text{gravitational field equation}, \tag{2.35}$$

where $\boldsymbol{\nabla}$ is the derivative with respect to $\boldsymbol{x}$, and $p$ and $\phi$ are respectively the perturbed pressure and gravitational potential. Taking the Fourier decomposition of physical quantities as in equation (2.31), the above formulae give, to the first order in perturbations,

$$\boldsymbol{v}_k = \boldsymbol{v}_k^{\parallel} + \boldsymbol{v}_k^{\perp} = -\frac{i\boldsymbol{k}}{k^2}a\dot{\delta}_k + \frac{\text{const.}}{a}, \tag{2.36}$$

$$\ddot{\delta}_k + 2\frac{\dot{a}}{a}\dot{\delta}_k + \left(\frac{c_{\mathrm{s}}^2k^2}{a^2} - 4\pi G\bar{\rho}\right)\delta_k = 0, \tag{2.37}$$

where $c_{\mathrm{s}}^2 \equiv \partial p/\partial\rho$ is the sound velocity, and $\boldsymbol{v}_k^{\parallel}$ and $\boldsymbol{v}_k^{\perp}$ are respectively the rotation-free ($\boldsymbol{k}\times\boldsymbol{v}_k^{\parallel} = 0$) and divergence-free ($\boldsymbol{k}\cdot\boldsymbol{v}_k^{\perp} = 0$) part of the peculiar velocity field (the latter decays as $\propto a^{-1}$). Given the zeroth order solutions for the background universe, i.e. $a(t)$ and $\bar{\rho}(t)$, equation (2.37) governs the behaviour of linear density perturbations. It is obvious from this equation that each $k$-mode $\delta_k$ evolves independently of other modes in the linear regime.

The solution to equation (2.37) is classified by the Jeans wave number defined by

$$k_{\mathrm{J}}^2 \equiv \frac{4\pi G\bar{\rho}a^2}{c_{\mathrm{s}}^2}. \tag{2.38}$$

For $k \gg k_{\mathrm{J}}$, the pressure dominates the gravity and equation (2.37) has stable oscillating solutions. For $k \ll k_{\mathrm{J}}$, the pressure becomes negligible and the perturbations become gravitationally unstable, yielding the solution of the form:

$$\delta_k(t) = A_+(\boldsymbol{k})D_+(t) + A_-(\boldsymbol{k})D_-(t), \tag{2.39}$$

where the subscripts + and − represent the growing and decaying modes, respectively. We consider only the growing mode which dominates the solution. In the matter dominated era ($\bar{\rho} \propto a^{-3}$), the expressions for the linear growth factor $D(t)$ for the three cosmological models mentioned in §2.2 are:

1. Einstein-de Sitter model: $\Omega_0 = 1$, $\lambda_0 = 0$

$$D(t) \propto a(t) \propto t^{2/3}. \tag{2.40}$$

2. Open model with no cosmological constant: $\Omega_0 < 1$, $\lambda_0 = 0$

$$D(t) \propto 1 + \frac{3}{x} + 3\sqrt{\frac{1+x}{x^3}} \ln(\sqrt{1+x} - \sqrt{x}), \tag{2.41}$$

$$x \equiv (\Omega_0^{-1} - 1)a(t).$$

3. Flat model with positive cosmological constant: $\Omega_0 < 1$, $\lambda_0 = 1 - \Omega_0$

$$D(t) \propto \sqrt{1 + \frac{2}{y^3}} \int_0^y \mathrm{d}u \left(\frac{u}{2+u^3}\right)^{3/2}, \tag{2.42}$$

$$y \equiv 2^{1/3}(\Omega_0^{-1} - 1)^{1/3} a(t).$$

Equations (2.41) and (2.42) approach equation (2.40) at $1 + z = a^{-1} \gg \Omega_0^{-1} - 1$ and $1 + z \gg (\Omega_0^{-1} - 1)^{1/3}$, respectively. As $D(t)$ can be normalized arbitrarily, we hereafter set $D(t_0) = 1$.

We plot in Figure 2.1 the linear growth factor of density fluctuations given by equations (2.40)–(2.42) for some representative choices of $\Omega_0$ and $\lambda_0$. This figure indicates that the density fluctuations continue to grow at a rapid rate at recent times for $\Omega_0 = 1$, while their growth slows down for lower values of $\Omega_0$. For a fixed value of $\Omega_0$, the growth slows down more significantly in the open models with $\lambda_0 = 0$ than the spatially flat models with $\lambda_0 = 1 - \Omega_0$.

## 2.4 Characteristic scales for the structure formation

The results of the last section are correct as long as the cosmological matter can be treated as an ideal fluid. In reality, however, there are several physical processes that operate to affect the growth of perturbations, such as the free streaming damping and Silk damping of small scale fluctuations. These processes, together with the Jeans instability discussed in the last section, give rise to characteristic scales in the distribution of fluctuations. In this section, we summarize such scales that are of importance for the structure formation in the universe.

For the baryonic component, there are mainly two important scales, the Jeans scale and the collisional damping (Silk damping) scale. The former scale in the physical coordinate is defined using equation (2.38) as

$$l_{\mathrm{J}} \equiv \frac{2\pi a}{k_{\mathrm{J}}} = c_{\mathrm{s}} \sqrt{\frac{\pi}{G\rho}}. \tag{2.43}$$

For scales larger than this, the gravitational force overcomes the pressure support and the density perturbations can grow via gravitational instability. At $t < t_{\mathrm{eq}}$, the pressure and

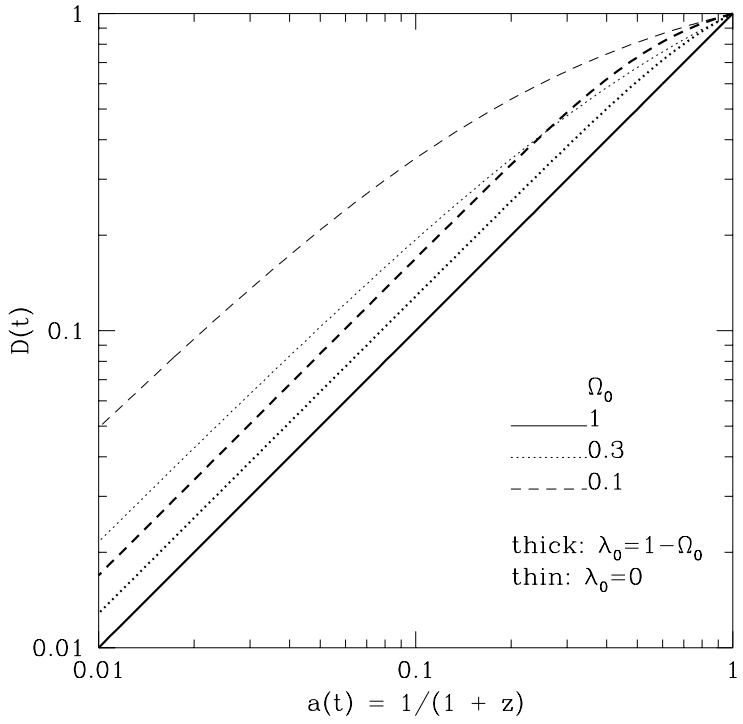


Figure 2.1: Linear growth factor of density fluctuations in the cases of $\Omega_0 = 1$ (solid line), $\Omega_0 = 0.3$ (dotted), and $\Omega_0 = 0.1$ (dashed). Thick lines are the spatially flat models with $\lambda_0 = 1 - \Omega_0$, while thin lines are the open models with $\lambda_0 = 0$.

gravity are both supplied by photons, and the Jeans scale is roughly equal to the Hubble horizon radius:

$$l_{\rm J}(t) \simeq \frac{2\sqrt{2}\pi}{3} d_{\rm H}(t) \propto t \propto a^2 \quad \text{at} \quad t < t_{\rm eq}, \tag{2.44}$$

where $c_{\rm s}^2 = c^2/3$ and $H^2 \simeq 8\pi G\rho/3$ are used. This means that the growth of the fluctuations is suppressed within the horizon scale during the radiation dominated era. At $t_{\rm eq} < t < t_{\rm dec}$, baryons are still tightly coupled with photons and the Jeans scale depends on the amount of baryons present in the universe. If the baryon density is large and dominates the density of the photon-baryon plasma, the sound velocity starts to decline as $c_{\rm s}^2 \propto \rho_\gamma/\bar{\rho} \propto a^{-1}$, and the Jeans scale becomes $l_{\rm J} \propto a$, i.e. constant in the comoving coordinate. Alternatively, if baryon density is small, the sound velocity is kept at $c_{\rm s}^2 = c^2/3$, and the Jeans length continues to grow as $l_{\rm J}(t) \propto d_{\rm H}(t) \propto t \propto a^{3/2}$. In either cases, the comoving Jeans scale lies between $d_{\rm H}(t_{\rm eq})/a(t_{\rm eq})$ and $d_{\rm H}(t_{\rm dec})/a(t_{\rm dec})$, and the fluctuations below this scale are suppressed to grow. At $t > t_{\rm dec}$, however, the photon pressure is no longer effective and the Jeans length is

significantly reduced as

$$\frac{l_{\rm J}(t)}{a(t)} \sim 10^{-2}(\Omega_0 h^2)^{-1/2}\left(\frac{T_{\rm B}}{T_\gamma}\right)^{1/2} \ {\rm Mpc} \quad {\rm at} \quad t > t_{\rm dec}, \tag{2.45}$$

where $T_{\rm B}$ and $T_\gamma$ are respectively the baryon and photon temperatures, and $c_s^2 \simeq 5k_{\rm B}T_{\rm B}/3m_{\rm p}$ is used. The baryon mass enclosed inside the above length defines the baryonic Jeans mass:

$$\begin{aligned} M_{\rm J}^{\rm B}(t) &\equiv \frac{4\pi\rho_0}{3}\left(\frac{\Omega_{\rm B}}{\Omega_0}\right)\left[\frac{l_{\rm J}(t)}{a(t)}\right]^3 \\ &\sim 10^6\left(\frac{\Omega_{\rm B}}{\Omega_0}\right)(\Omega_0 h^2)^{-1/2}\left(\frac{T_{\rm B}}{T_\gamma}\right)^{3/2} \ {\rm M}_\odot \quad {\rm at} \quad t > t_{\rm dec}. \end{aligned} \tag{2.46}$$

The second scale for the baryonic component, the collisional damping (Silk damping) scale, is due to the photon diffusion around the time of photon-baryon decoupling. As the photon mean free path becomes larger, the photons diffuse out of overdense regions into underdense regions, thereby damping small scale inhomogeneities. Considering the random walk of photons, this damping scale is estimated as (e.g. Kolb & Turner 1990)

$$\begin{aligned} \frac{l_{\rm Silk}}{a_{\rm dec}} &\sim 4\left(\frac{\Omega_0}{\Omega_{\rm B}}\right)^{1/2}(\Omega_0 h^2)^{-3/4} \ {\rm Mpc}, \\ M_{\rm Silk} &\equiv \frac{4\pi\rho_0}{3}\left(\frac{l_{\rm Silk}}{a_{\rm dec}}\right)^3 \sim 6\times 10^{12}\left(\frac{\Omega_0}{\Omega_{\rm B}}\right)^{3/2}(\Omega_0 h^2)^{-5/4}\,{\rm M}_\odot. \end{aligned} \tag{2.47}$$

Thus, if the matter in the universe is dominated by baryons, the first cosmological object to be formed is larger than the galactic scales.

For collisionless dark matter components, important scales are the free streaming (collisionless damping) scale and the stagspansion scale. The former is the scale under which the density perturbations are smoothed out by the phase mixing of collisionless particles. Since the relativistic particles move freely within the Hubble horizon, the free streaming scale is roughly the horizon radius at which dark matter particles cease to be relativistic:

$$\begin{aligned} \frac{l_{\rm FS}}{a(t_{\rm nr})} &\sim \frac{d_{\rm H}(t_{\rm nr})}{a(t_{\rm nr})} \sim 15\left(\frac{m_{\rm DM}}{30{\rm eV}}\right)^{-1}\frac{T_{\rm DM}}{T_\gamma} \ {\rm Mpc}, \\ M_{\rm FS} &\equiv \frac{4\pi\rho_0}{3}\left(\frac{l_{\rm FS}}{a_{\rm nr}}\right)^3 \sim 4\times 10^{15}(\Omega_0 h^2)\left(\frac{m_{\rm DM}}{30{\rm eV}}\right)^{-3}\left(\frac{T_{\rm DM}}{T_\gamma}\right)^3 {\rm M}_\odot, \end{aligned} \tag{2.48}$$

where $m_{\rm DM}$ and $T_{\rm DM}$ are respectively the mass and temperature of the dark matter particles, and $t_{\rm nr}$ is the time at which the particles become non-relativistic. We define $t_{\rm nr}$ as as the epoch at which $T_{\rm DM} \sim m_{\rm DM}c^2/3$, assuming that the particles were initially in thermal equilibrium. For the HDM, such as light neutrinos ($T_{\rm DM}/T_\gamma \simeq 0.71$) with mass $m_{\rm DM} \sim 30$eV, perturbations at scales under clusters of galaxies will be wiped out. For the CDM, on the other hand, the free streaming scale becomes smaller than the galaxy scale either because the particle mass is larger than ∼keV (e.g. neutralinos) or because the velocity dispersion of the particles is initially small (e.g. axions).

The stagspansion occurs when the dark matter perturbations are dragged by the gravity of more dominant components. During the radiation dominated era, as the sub-horizon scale fluctuations of the photon-baryon plasma are unable to grow, the dark matter perturbations are also suppressed within the horizon scale. Since this suppression continues until the matter-radiation equality, the stagspansion scale is equal to the Hubble horizon radius at this epoch, $d_{\rm H}(t_{\rm eq})/a(t_{\rm eq})$ (eq. [2.26]). For the CDM, this is the only scale that is of astrophysical interest.

## 2.5 Power spectrum of density fluctuations

So far we have considered the physics that may affect the density fluctuations of a given scale (i.e. a given $k$-mode). We next discuss statistical descriptions of density fluctuations.

It is generally assumed that the primordial density fluctuation field is random Gaussian, whose statistical properties are specified completely by a power spectrum defined as

$$P(k,t) \equiv \langle |\delta_k(t)|^2 \rangle. \tag{2.49}$$

For a rigorous definition of the random Gaussian field, see for instance Adler (1981). The primordial power spectrum produced in the early universe is often assumed to have a scale-free power-law form with initial index $n_{\rm i}$:

$$P(k,t_{\rm i}) \propto k^{n_{\rm i}}. \tag{2.50}$$

The case of $n_{\rm i} = 1$ is called the Harrison-Zel'dovich spectrum and has a unique property that the fluctuation enters the horizon with the constant amplitude. Conventional inflationary models predict $n_{\rm i}$ close to 1 (e.g. Liddle & Lyth 1993). As seen in the previous section, the linear growth of density fluctuations is affected by several physical processes and the linear power spectrum is altered accordingly from its primordial shape. After the relevant physical processes cease to affect the spectral shape, the fluctuations simply grow according to the linear growth factor $D(t)$ (eqs [2.40]–[2.24]) at all scales until they become non-linear. The processed power spectrum is then written as

$$P(k,t) = P(k,t_{\rm i})T(k,t)\frac{D^2(t)}{D^2(t_{\rm i})}, \tag{2.51}$$

where $T(k,t)$ is a transfer function representing the changes in the spectral shape.

The specific form of the transfer function is computed by solving numerically the Boltzmann equations for photon, baryon and dark matter distribution functions. For the universe dominated by CDM, it is well fitted by (Bardeen et al. 1986)

$$T_{\rm CDM}(k) = \left[\frac{\ln(1+2.34q)}{2.34q}\right]^2 \left[1+3.89q+(16.1q)^2+(5.46q)^3+(6.71q)^4\right]^{-1/2} \quad \text{at } t > t_{\rm dec}, \tag{2.52}$$

where $q \equiv k/(\Gamma h\ {\rm Mpc}^{-1})$, and $\Gamma$ is the shape parameter given by (Peacock & Dodds 1994; Sugiyama 1995)

$$\Gamma \equiv \Omega_0 h (T_{\gamma 0}/2.7\ {\rm K})^{-2} \exp[-\Omega_{\rm B}(1+\sqrt{2h}\Omega_0^{-1})]. \tag{2.53}$$

Figure 2.2(a) shows the linear power spectrum of density fluctuations in the universe dominated by CDM and HDM. For the HDM model, as discussed in §2.4, fluctuations below

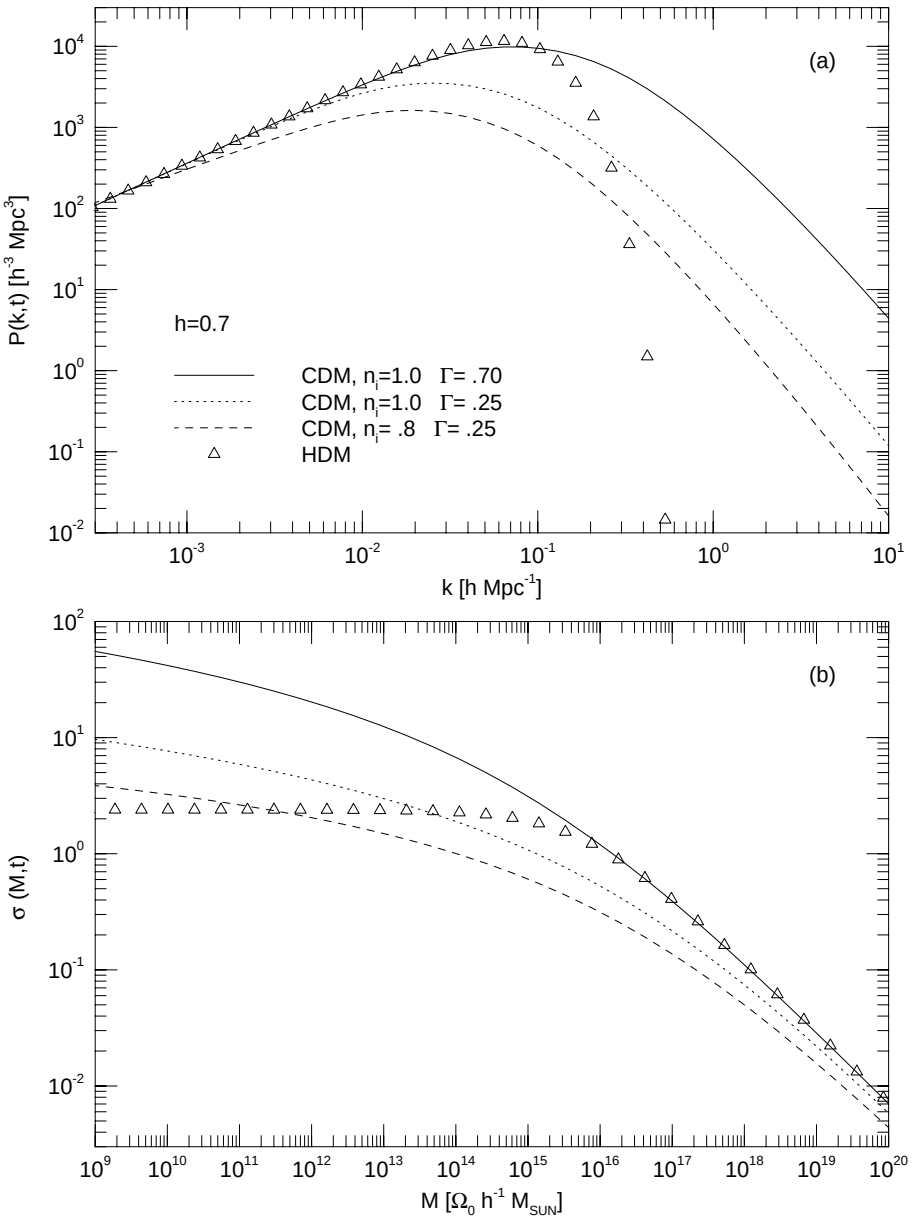


Figure 2.2: (a) Linear power spectrum of density fluctuations in CDM models with $(n_{\rm i}, \Gamma) = (1.0, 0.7)$ (solid line), $(1.0, 0.25)$ (dotted), and $(0.8, 0.25)$ (dashed) computed from equation (2.52). Also plotted is that of the HDM model (triangles) from Bond & Szalay (1983) with $m_{\rm DM} = 30$eV .The Hubble constant is fixed at $h = 0.7$. (b) Mass variance of density fluctuations. Parameters are the same as panel (a).

the free streaming scale (eq. [2.48]) are smoothed out and disappear. In the CDM dominated universe, the characteristic scale appears at $k_{\rm eq} \equiv 2\pi a(t_{\rm eq})/d_{\rm H}(t_{\rm eq})$ (see also eq. [2.26]), and fluctuations below this scale (i.e. $k > k_{\rm eq}$) are suppressed. For larger $\Gamma(\simeq \Omega_0 h)$, $k_{\rm eq}$ becomes larger, since the matter-radiation equality gets earlier as indicated by equation (2.18) to make $d_{\rm H}(t_{\rm eq})$ smaller. Asymptotic forms of the CDM spectrum are given by

$$P_{\rm CDM}(k,t) \propto \begin{cases} k^{n_{\rm i}} & k \ll k_{\rm eq}, \\ k^{n_{\rm i}-4} & k \gg k_{\rm eq}. \end{cases} \tag{2.54}$$

The difference of $k^4$ between these limits is understood as follows. For a mode with $k \gg k_{\rm eq}$, the perturbation enters in the horizon scale at $t_{\rm ent} < t_{\rm eq}$ and stays almost constant until $t_{\rm eq}$. Then its amplitude $\delta_k$ is suppressed by a factor $[a(t_{\rm ent})/a(t_{\rm eq})]^{-2}$. On the other hand, the comoving wave number corresponding to the horizon size varies as $k \propto a/d_{\rm H} \propto a^{-1}$. As a result, the power spectrum $P(k) \propto \delta_k^2$ is suppressed by $[k(t_{\rm ent})/k(t_{\rm eq})]^4$ .

It now remains to fix the amplitude of the power spectrum. There are conventionally two ways of doing this. One is to use the observed temperature fluctuations of the CMB at the photon-baryon decoupling $t_{\rm dec}$. Currently available CMB normalization comes from the *COBE* satellite which measures the fluctuations at $\gtrsim 7^\circ$, corresponding to the comoving scale of $\gtrsim 1$ Gpc. As this is larger than the horizon radius at the matter-radiation equality, *COBE* gives a direct clue to the primordial spectrum. Having fixed the spectrum at $t_{\rm dec}$ in this way, one can simply evolve it to an arbitrary epoch by $P(k,t) = P(k,t_{\rm dec})D^2(t)/D^2(t_{\rm dec})$. One should note, however, that scales of interest for structure formation, $\lesssim 10h^{-1}$ Mpc, are over two orders of magnitude smaller than the scale probed by *COBE*. Thus, if the *COBE* normalization is adopted, the power spectrum at $\lesssim 10h^{-1}$ Mpc is highly sensitive to small changes in the primordial spectral index $n_{\rm i}$ and the shape of the transfer function $T(k)$, as can be seen in Figure 2.2(a).

The other way to fix the amplitude of the power spectrum is to use the present-day distribution of density fluctuations. Consider a density fluctuation that is averaged over a volume $V$ containing mass $M = \bar{\rho}V$:

$$\begin{aligned} \delta(M,\boldsymbol{x},t) &= \int \delta(\boldsymbol{y},t) W_M(|\boldsymbol{x}-\boldsymbol{y}|) {\rm d}^3 y, \\ &= \frac{1}{(2\pi)^3} \int \delta_k(t) \hat{W}_M(k) \exp(-i\boldsymbol{k}\cdot\boldsymbol{x}) {\rm d}^3 k, \end{aligned} \tag{2.55}$$

where the spatial window function $W_M(r)$ and its Fourier transform $\hat{W}_M(k)$ are related to each other as in equation (2.31). A top-hat form is often adopted for the spatial window function:

$$\begin{aligned} W_M(r) &= \begin{cases} 3/(4\pi R^3) & r < R, \\ 0 & r > R, \end{cases} \\ \hat{W}_M(k) &= \frac{3}{(kR)^3} \left[\sin(kR) - kR\cos(kR)\right], \end{aligned} \tag{2.56}$$

where $R = (3M/4\pi\bar{\rho})^{1/3}$ is the filtering radius. Then the mass variance, the rms fluctuation smoothed over the mass scale $M$, is given by

$$\sigma^2(M,t) \equiv \langle |\delta(M,\boldsymbol{x},t)|^2 \rangle = \frac{1}{(2\pi)^3} \int P(k,t) \hat{W}_M^2(k) {\rm d}^3 k. \tag{2.57}$$

It is rather conventional to use $\sigma_8$, or equivalently the biasing factor $b$, to represent the normalization of the present-day power spectrum:

$$\sigma_8 \equiv \sigma(R = 8h^{-1}\mathrm{Mpc}, t_0) \equiv \frac{1}{b}. \tag{2.58}$$

Then the spectrum at an arbitrary epoch is computed by $P(k,t) = P(k,t_0)D^2(t)/D^2(t_0)$ and $\sigma(M,t) = \sigma(M,t_0)D(t)/D(t_0)$.

The mass variance at a given scale is essentially determined by the amount of power $k^3P(k,t)$ *above* that scale. If $k^3P(k,t)$ increases monotonically with decreasing scale (i.e. increasing $k$), as in the case of the CDM spectrum with $n_\mathrm{i} \geq 1$, then the largest contribution to the mass variance comes from the fluctuations right at that scale. In this case, it is often convenient to approximate the power spectrum locally by a power-law form as $P(k,t) \propto k^n$ using an effective index $n$. Then the mass variance is expressed as

$$\sigma^2(M,t) \propto k^{n+3} \propto R^{-(n+3)} \propto M^{-(n+3)/3} \tag{2.59}$$

From equation (2.54), the CDM model predicts the following asymptotic forms for the mass variance:

$$\sigma^2_\mathrm{CDM}(M,t) \propto \begin{cases} M^{-(n_\mathrm{i}+3)/3} & M \gg M_\mathrm{eq}, \\ M^{-(n_\mathrm{i}-1)/3} & M \ll M_\mathrm{eq}, \end{cases} \tag{2.60}$$

where $M_\mathrm{eq}$ is the mass enclosed within a sphere of comoving radius $d_\mathrm{H}(t_\mathrm{eq})/a(t_\mathrm{eq}) = 2\pi/k_\mathrm{eq}$. On the contrary, if $k^3P(k,t)$ has a maxima at some scale as in the case of the HDM spectrum, then the mass variance below that scale will become almost constant.

The above mentioned features are clearly shown in Figure 2.2(b). For the HDM spectrum, the mass variance increases with decreasing scale down to the free-streaming scale and stays constant thereafter. For the CDM spectrum, the mass variance increases monotonically with decreasing scale over the plotted mass range. It is important to note for our later analyses that the CDM mass variance has steeper slope at the cluster scale if $\Gamma \simeq \Omega_0 h$ is larger.

In the rest of this thesis, unless otherwise stated explicitly, we adopt the CDM fluctuation spectrum given by equations (2.50)–(2.53) with the Harrison-Zel'dovich primordial spectrum ($n_\mathrm{i} = 1$). In this case, we have found that the mass variance computed via equations (2.56) and (2.57) is well fitted by the following formula:

$$\sigma_\mathrm{CDM} \propto \left[1 + 2.208m^p - 0.7668m^{2p} + 0.7949m^{3p}\right]^{-\frac{2}{9p}}, \tag{2.61}$$

where $p = 0.0873$, and $m \equiv M(\Gamma h)^3/(\Omega_0 h^2)/10^{12}\,\mathrm{M}_\odot$. The above approximation and its derivative with respect to $m$ simultaneously give accurate fits to $\sigma$ and $\mathrm{d}\sigma/\mathrm{d}M$ within a few percent in the range $10^{-7} \lesssim m \lesssim 10^5$.

## 2.6 Spherical model of the non-linear collapse

Once perturbations enter the non-linear regime, it is very difficult to fully trace their evolution analytically. However, one can obtain a good insight into this subject by studying the fate of spherically symmetric perturbations, for which the gravitational collapse can be solved exactly (Gunn & Gott 1972; Peebles 1980).

Suppose that a spherical mass shell with a physical radius $r = r(t)$ is moving with the expansion of background universe. For simplicity, we assume that the background universe is described by the Einstein-de Sitter model. In the absence of shell crossings, the time evolution of such a shell is simply described by the equation of motion :

$$\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -\frac{GM}{r^2}, \tag{2.62}$$

where $M = M(< r)$ is the mass enclosed within the radius $r$. The collapsing solution to the above equation is given in a parametric form as

$$r = \frac{GM}{C}(1 - \cos\theta), \qquad t = \frac{GM}{C^{3/2}}(\theta - \sin\theta), \tag{2.63}$$

where $C$ is an arbitrary constant. At an early epoch ($\theta \ll 1$), the mean density within the mass shell is given by expanding equation (2.63):

$$\bar{\rho}(< r, t) = \frac{3M}{4\pi r^3} = \frac{1}{6\pi G t^2}\left[1 + \frac{3C}{20}\left(\frac{6t}{GM}\right)^{2/3} + \cdots\right]. \tag{2.64}$$

Equation (2.64) expresses the density as a perturbation from the Einstein-de Sitter universe (eq. [2.32]). Then from equations (2.32) and (2.64), the mean linear density contrast of this region is

$$\delta(< r, t) = \frac{\bar{\rho}(< r, t) - \bar{\rho}(t)}{\rho_{\mathrm{b}}(t)} = \frac{3C}{20}\left(\frac{6t}{GM}\right)^{2/3} \propto D(t), \tag{2.65}$$

where $D(t)$ is the linear growth factor given in equation (2.40).

Equation (2.63) implies that the mass shell initially expanding with the background universe will gradually slow down, reach a maximum "turn-around" radius $r_{\mathrm{ta}}$ at $\theta = \pi$, and then start to contract. Equation (2.63) further suggests that all the mass in the shell will collapse to a point with infinite density at $\theta = 2\pi$. In reality, however, the mass shell is thought to undergo a process called violent relaxation and reach virial equilibrium instead of contracting to zero radius. Violent relaxation is essentially due to the time-varying gravitational potential during collapse (Lyndel-Bell 1967). Since the potential fluctuates with time, the energy of individual particles is not conserved and thereby the random velocities of particles are enhanced. As a result, shell crossings take place and the collisionless component of the density will relax to attain a radius given by the virial theorem:

$$r_{\mathrm{vir}} = \frac{r_{\mathrm{ta}}}{2} = \frac{r(\theta = \pi)}{2}, \tag{2.66}$$

The time at which the system reaches virial equilibrium, $t_{\mathrm{vir}}$, is approximately the time corresponding to $\theta = 2\pi$. From equations (2.32) and (2.63), the ratio of the density of the spherical region to that of the background universe at this epoch is

$$\Delta_{\mathrm{c}} \equiv \frac{\bar{\rho}(< r_{\mathrm{vir}}, t_{\mathrm{vir}})}{\bar{\rho}(t_{\mathrm{vir}})} = 18\pi^2. \tag{2.67}$$

It is often convenient to consider the linear density contrast "extrapolated" to this epoch using equation (2.65):

$$\delta_{\mathrm{c}} \equiv \delta(< r_{\mathrm{vir}}, t_{\mathrm{vir}}) = \frac{3}{20}(12\pi)^{2/3} \tag{2.68}$$

This is the density contrast that a collapsed region should have at $t_{\rm vir}$ if it had always evolved according to linear theory. In other words, if the density contrast predicted by linear theory reaches this value, the region is thought to have collapsed at that time.

The above discussion is similarly extended to other cosmological models (Peebles 1980; Lahav et al. 1991; Lacy & Cole 1993; Nakamura & Suto 1997). In the following, we summarize the results of the spherical collapse model in the three cosmological models mentioned in §2.2:

1. Einstein-de Sitter model: $\Omega_0 = 1$, $\lambda_0 = 0$

$$\begin{aligned} \Delta_{\rm c} &= 18\pi^2 \simeq 178, \\ \delta_{\rm c} &= \frac{3(12\pi)^{2/3}}{20} \simeq 1.69, \end{aligned} \tag{2.69}$$

2. Open models with no cosmological constant: $\Omega_0 < 1$, $\lambda_0 = 0$

$$\begin{aligned} \Delta_{\rm c} &= 4\pi^2 \frac{(\cosh\eta_{\rm vir} - 1)^3}{(\sinh\eta_{\rm vir} - \eta_{\rm vir})^2}, \\ \delta_{\rm c} &= \frac{3}{2}\left[\frac{3\sinh\eta_{\rm vir}(\sinh\eta_{\rm vir} - \eta_{\rm vir})}{(\cosh\eta_{\rm vir} - 1)^2} - 2\right]\left[1 + \left(\frac{2\pi}{\sinh\eta_{\rm vir} - \eta_{\rm vir}}\right)^{2/3}\right], \end{aligned} \tag{2.70}$$

3. Flat models with positive cosmological constant: $\Omega_0 < 1$, $\lambda_0 = 1 - \Omega_0$

$$\begin{aligned} \Delta_{\rm c} &= \left(\frac{r_{\rm ta}}{r_{\rm vir}}\right)^3 \frac{2w_{\rm vir}}{\chi}, \\ &\simeq 18\pi^2(1 + 0.4093 w_{\rm vir}^{0.9052}), \\ \delta_{\rm c} &= \frac{3}{5} F\left(\frac{1}{3}, 1, \frac{11}{6}; -w_{\rm vir}\right)\left(\frac{2w_{\rm vir}}{\chi}\right)^{1/3}\left(1 + \frac{\chi}{2}\right), \\ &\simeq \frac{3(12\pi)^{2/3}}{20}(1 + 0.0123\log_{10}\Omega_{\rm vir}). \end{aligned} \tag{2.71}$$

Equation (2.70) is from Lacey & Cole (1993), and equation (2.71) is from Nakamura & Suto (1997). In these equations, $\eta_{\rm vir} \equiv \arccosh(2/\Omega_{\rm vir} - 1)$, $w_{\rm vir} \equiv 1/\Omega_{\rm vir} - 1$, $\chi \equiv \lambda_0 H_0^2 r_{\rm ta}^3/(GM)$, $F$ is the hyper-geometric function of type (2,1), and $\Omega_{\rm vir}$ is the density parameter at $t_{\rm vir}$. The above results are plotted as a function of $\Omega_{\rm vir}$ in Figure 2.3.

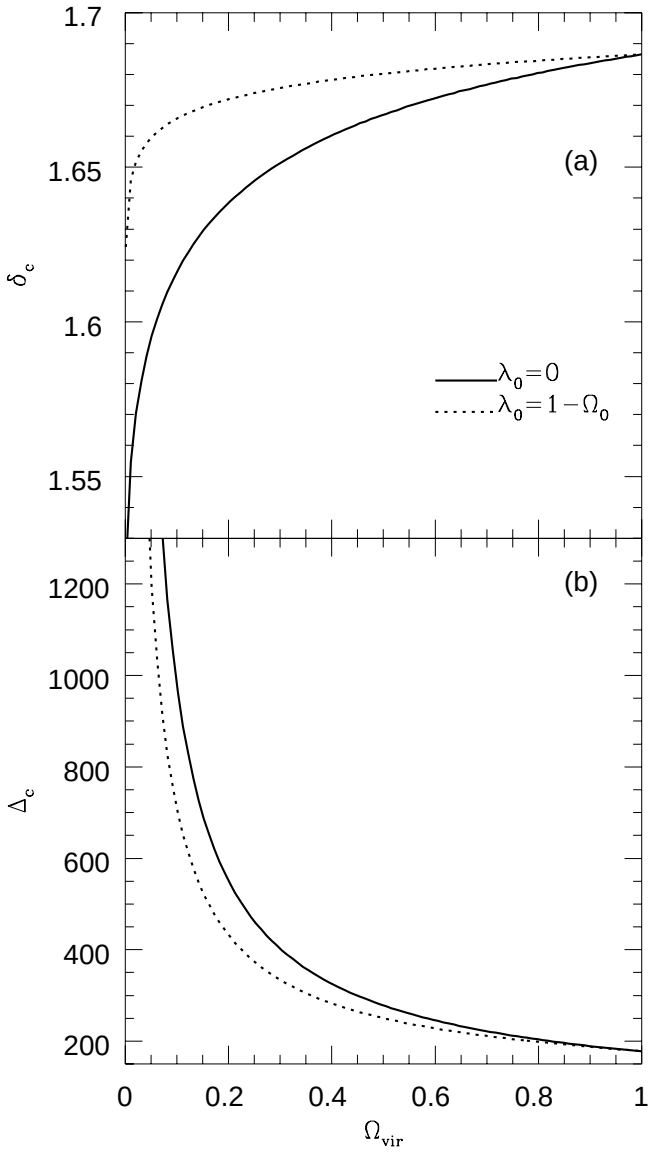


Figure 2.3: Critical overdensity $\delta_c$ and $\Delta_c$ as a function of density parameter at virialization $\Omega_{vir}$. Solid lines indicate $\lambda_0 = 0$, while dotted $\lambda_0 = 1 - \Omega_0$.

# Chapter 3

# Mass Function of Virialized Structures

In order to make definite predictions for the cluster abundances, it is necessary to quantify the hierarchical development of non-linear structures. This can be done most directly by means of intensive numerical simulations (e.g. Hockney & Eastwood 1981; Davis et al. 1985; Cen & Ostriker 1992, 1993; Suto 1993; Kang et al. 1994; Bryan et al. 1994; Jing et al. 1994; Suginohara 1994; Navarro et al. 1995, 1996, 1997). Alternatively, there have been a number of attempts for analytical prescriptions (see Monaco 1998 for review). In particular, the theory by Press & Schechter (1972, PS) and its extensions (e.g. Bond et al. 1991; Lacey & Cole 1993) yield simple ways of describing the building-up of dark matter halos that also agree well with the numerical simulations (e.g. Efstathiou et al. 1988; Carlberg & Couchman 1989; White, Efstathiou & Frenk 1993; Lacey & Cole 1994; Eke, Cole & Frenk 1996). This chapter describes these formalisms including that proposed by the present author (Kitayama & Suto 1996a).

## 3.1 Press-Schechter theory

### 3.1.1 Derivation of the mass function

Based on the spherical model of nonlinear collapse discussed in §2.6, one can derive a mass function, which describes how many virialized systems of a given mass are present in the universe at a given epoch. From equations (2.65) and (2.68), the condition for a spherical region to collapse by the time $t$ is that the region has had at some early epoch $t_{\rm i}$ a linear overdensity greater than

$$\delta_{\rm c}(t,t_{\rm i}) \equiv \delta_{\rm c}\frac{D(t_{\rm i})}{D(t)}, \tag{3.1}$$

where $t_{\rm i}$ is arbitrary as long as it is after the shape of the density fluctuation spectrum is fixed and the perturbations start to grow at all scales in proportion to $D(t)$.

Consider then an initial density field $\delta(M,\boldsymbol{x},t_{\rm i})$ smoothed over a spherical region that contains mass $M$ around a comoving point $\boldsymbol{x}$ at time $t_{\rm i}$. For a random Gaussian field, the probability distribution at any given point is given by

$$P_{\rm G}[\delta(M,t_{\rm i})]{\rm d}\delta = \frac{1}{\sqrt{2\pi\sigma^2(M,t_{\rm i})}}\exp\left[-\frac{\delta^2(M,t_{\rm i})}{2\sigma^2(M,t_{\rm i})}\right]{\rm d}\delta, \tag{3.2}$$

where $\sigma^2(M,t_{\rm i})$ is the mass variance defined by equation (2.57) at $t_{\rm i}$. The fraction of the total mass in the universe which is bound in collapsed objects of mass greater than mass $M$ at time $t$ is calculated from equations (3.1) and (3.2) as

$$\begin{aligned} F(M,t) &= \int_{\delta_{\rm c}(t,t_{\rm i})}^{\infty} P_{\rm G}[\delta(M,t_{\rm i})]{\rm d}\delta \\ &= \frac{1}{2}{\rm erfc}\left(\frac{\delta_{\rm c}(t,t_{\rm i})}{\sqrt{2}\sigma(M,t_{\rm i})}\right), \end{aligned} \tag{3.3}$$

where ${\rm erfc}(u)$ is the complimentary error function defined by

$${\rm erfc}(u) \equiv \frac{2}{\sqrt{\pi}}\int_u^{\infty} \exp(-t^2){\rm d}t. \tag{3.4}$$

Equation (3.3) is independent of the initial epoch $t_{\rm i}$ since $\delta_{\rm c}(t,t_{\rm i})$ and $\sigma(M,t_{\rm i})$ are both proportional to $D(t_{\rm i})$. This allows one to replace $\delta_{\rm c}(t,t_{\rm i})$ and $\sigma(M,t_{\rm i})$ in equation (3.3) with the values extrapolated according to the linear theory at an arbitrarily fixed reference epoch. It is often convenient to choose the present time $t_0$ as such a reference epoch and express $F(M,t)$ as

$$F(M,t) = \frac{1}{2}{\rm erfc}\left(\frac{\delta_{\rm c}(t)}{\sqrt{2}\sigma(M)}\right), \tag{3.5}$$

where

$$\delta_{\rm c}(t) \equiv \delta_{\rm c}(t,t_0) = \frac{\delta_{\rm c}}{D(t)}, \tag{3.6}$$

and

$$\sigma^2(M) \equiv \sigma^2(M,t_0). \tag{3.7}$$

Note that we have normalized the linear growth factor $D(t)$ to give $D(t_0)=1$. The above treatment is simply equivalent to extrapolating the whole initial density field to the reference epoch $t_0$, and counting the fraction of regions with $\delta > \delta_{\rm c}(t)$ at a mass scale $M$.

From equation (3.5), the comoving number density of collapsed systems of mass $M$ at time $t$, per interval ${\rm d}M$, is expressed as (Press & Schechter 1974, PS)

$$\begin{aligned} \frac{{\rm d}n(M,t)}{{\rm d}M} &\equiv 2\frac{\rho_0}{M}\left|\frac{\partial F}{\partial M}\right| \\ &= \sqrt{\frac{2}{\pi}}\frac{\rho_0}{M}\frac{\delta_{\rm c}(t)}{\sigma^2(M)}\left|\frac{{\rm d}\sigma(M)}{{\rm d}M}\right|\exp\left[-\frac{\delta_{\rm c}^2(t)}{2\sigma^2(M)}\right], \end{aligned} \tag{3.8}$$

where $\rho_0$ is the mean mass density of the universe at present. The fudge factor 2 on the right-hand side of the first line was heuristically introduced by PS to correct for the normalization, so that

$$\int_0^{\infty} M\frac{{\rm d}n(M,t)}{{\rm d}M}{\rm d}M = \rho_0. \tag{3.9}$$

Physical meanings of this factor will be discussed in more detail in §3.1.2 and 3.2.1. Equation (3.8) is often called the PS mass function. Given cosmological parameters, the present day mass variance $\sigma^2(M)$, and the linear growth rate of density fluctuations $D(t)$, one can explicitly compute this quantity.

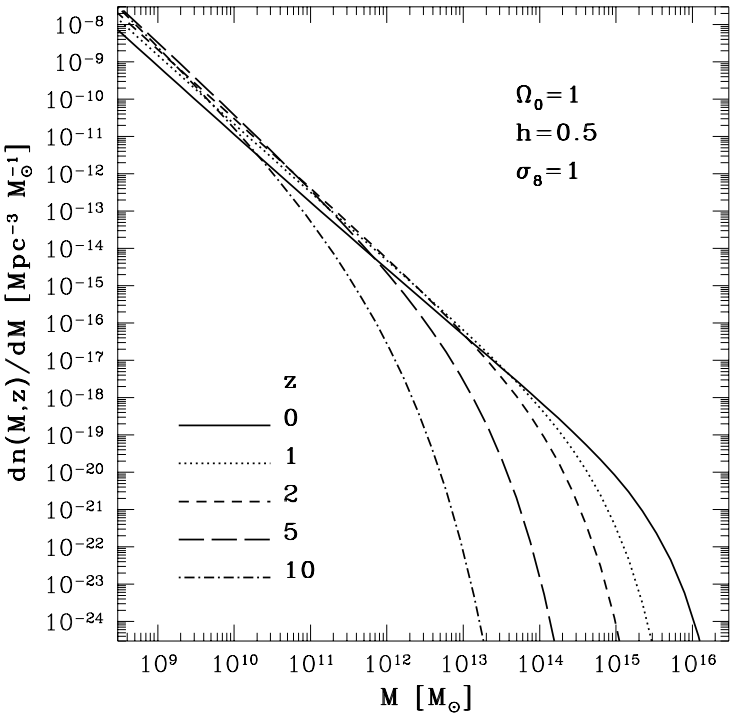


Figure 3.1: The Press-Schechter mass function in the CDM model with $\Omega_0 = 1$, $h = 0.5$, and $\sigma_8 = 1$, evaluated at $z = 0$ (solid line), $z = 1$ (dotted), $z = 2$ (short dashed), $z = 5$ (long dashed), and $z = 10$ (dot-dashed).

In Figure 3.1 we plot the PS mass function in a standard CDM model with $\Omega_0 = 1$, $h = 0.5$, and $\sigma_8 = 1$. The number density of bound objects decreases with time at small masses, while it increases at large masses. Thus the PS mass function evaluated for the CDM model correctly reflects the "bottom up picture" of hierarchical clustering, in which small-scale structures form first and large-scale structures appear later.

### 3.1.2 Validity of the Press-Schechter theory

The PS mass function has been extensively used in a variety of fields in cosmology (e.g. Efstathiou & Rees 1988; White & Frenk 1991; Evrard & Henry 1991; Fukugita & Kawasaki 1994; Mo & Fukugita 1996). This is mainly because of two reasons. One is that a number of authors have compared the PS mass function with the results of N-body simulations and found very good agreements (e.g., Efstathiou at al. 1988; Carlberg & Couchman 1989; White, Efstathiou & Frenk 1993; Lacey & Cole 1994; Eke et al. 1996). Such comparisons have proved the validity of the PS mass function for various cosmological models, power spectra, and halo identification schemes. The other is that the PS mass function has such a simple analytic form. It is thus very easily combined with models of specific astrophysical objects, such as clusters of galaxies, to make quantitative comparisons with observations. The simplicity of the PS formula also allows one to perform extensive parameter surveys and to grasp the underlying key physics.

Nonetheless, there are several cautions regarding the PS mass function which should be kept in mind. The first one is that the collapse epoch estimated by the spherical non-linear model may not be realistic in general cases, and that the pancake or cylindrical collapse may be more adequate. In order to incorporate such geometrical effects, ellipsoidal collapse has been studied using the Lagrangian dynamic models, such as the Zel'dovich approximation (Bartelmann, Ehlers & Schneider 1993; Monaco 1995; Lee & Shandarin 1997). It has been reported that, as the time taken for the first axis collapse in the ellipsoidal collapse is shorter than the spherical case, the mass function is modified in a way that is roughly fitted by the PS formula with a lower value of $\delta_c (\simeq 1.5)$. It is, however, not clear whether real virialized clumps correspond to the first axis collapse rather than second or third. Audit, Teyssier& Alimi (1997) finds that the mass function is very similar to the PS one if the second axis collapse is adopted, while the halo formation will even be delayed compared with the spherical case for the third axis collapse. The significance of the geometrical effects are thus still in question, although they do not seem to alter the mass function in a dramatic way.

The second point is closely related to the fudge normalization factor of 2 introduced by hand in equation (3.8). In the formalism of the last section, if a region has $\delta > \delta_c(t)$ when filtered over a mass scale $M$, it is attributed to an bound object of mass greater than $M$ present at time $t$. However, even if a region has $\delta < \delta_c(t)$ at a mass scale $M$, it may have $\delta > \delta_c(t)$ when filtered over some larger mass scale. Such a region should also be counted as an object with $> M$ present at $t$. Equation (3.5) misses the contributions from these regions and therefore underestimates the fraction of collapsed objects. The normalization factor of 2 was rather heuristically adopted in order to compensate such underestimation. This difficulty, so-called the "cloud-in-cloud problem", was solved by Peacock & Heavens (1990) and Bond et al. (1991). They reformulated the original PS theory and justified "the factor of 2" by taking properly account of underdense regions. We will study such an approach in §3.2.1.

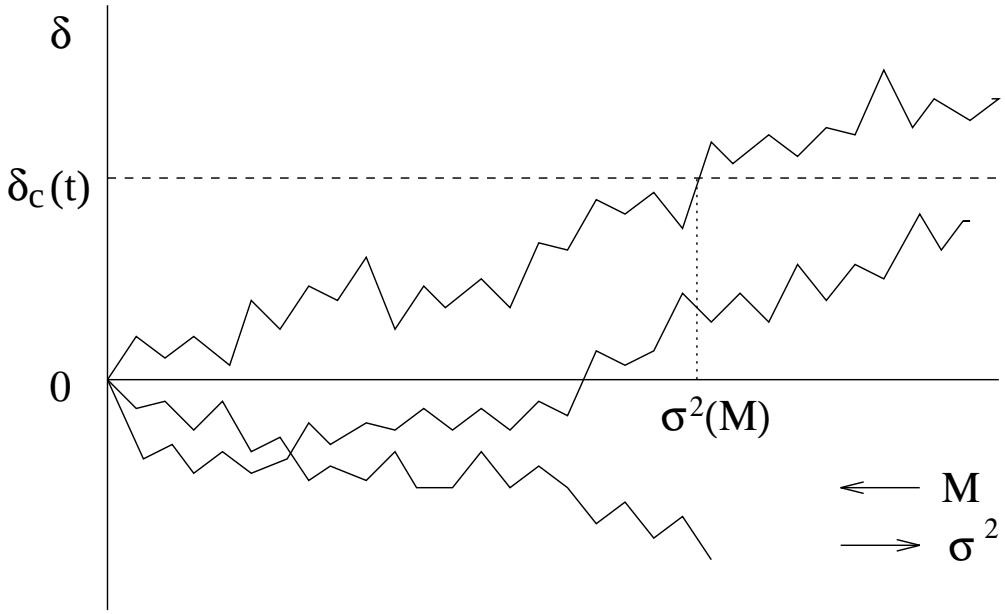


Figure 3.2: Trajectories of $\delta$ versus $\sigma^2$. The origin corresponds to $M = \infty$ or $\sigma^2 = 0$. Some of the trajectories make their first upcrossing of a barrier of height $\delta_c(t)$ at $\sigma^2(M)$.

The third point is more fundamental. As can be seen from its derivation, the PS mass function simply describes the statistics at a given epoch and does not trace the dynamical history of each virialized object. In a number of circumstances, however, such history (e.g. halo formation times) plays a crucial role in describing the subsequent formation and evolution of luminous objects such as quasars, galaxies, and clusters of galaxies (e.g. Kauffmann & White 1993). For this reason, several authors have investigated extensions of the PS theory. One of the most successful is the "excursion set approach" by Bond et al. (1991) which we will describe in §3.2.1. Interestingly, this approach solves simultaneously the cloud-in-cloud problem mentioned above and a part of the dynamical problem brought up here. We will also attempt to formulate the formation times of virialized halos based on this approach in the subsequent sections.

## 3.2 Excursion set approach

Bond et al. (1991) developed a method to extend the original PS theory solving at the same time the cloud-in-cloud problem mentioned in the last section. Their method is based upon the assumptions that the initial density field is random Gaussian and that the variance $\sigma^2(M)$ of density fluctuations is a monotonically decreasing function of $M$. As indicated in §2.5, both of these assumptions are satisfied in the present context.

### 3.2.1 Solving the cloud-in-cloud problem

Consider an initial density field $\delta(\boldsymbol{x})$ that is extrapolated to the present epoch $t_0$, and choose an arbitrary fixed location in space. One starts with smoothing around this point with a sufficiently large mass scale. In the limit $M \to \infty$, the mean overdensity $\delta$ within this scale and its variance $\sigma^2(M) = \langle|\delta|^2\rangle$ both approach zero. As one successively reduces the smoothing scale $M$, the density contrast $\delta$ begins to wonder away from zero and perform trajectories as illustrated in Figure 3.2. Since the variance $\sigma^2(M)$ monotonically increases with decreasing mass scale $M$, the trajectories of $\delta$ versus $M$ can also be looked upon as those versus $\sigma^2(M)$. We only consider trajectories at a fixed location $\boldsymbol{x}$ and do not attempt to relate different locations. Then the trajectories of $\delta$ become something like random walks with respect to $\sigma^2(M)$ (their details will be discussed shortly).

Now suppose that at some $\sigma^2(M)$ a trajectory of overdensity $\delta$ exceeds a critical value $\delta_{\rm c}(t)$ defined by equation (3.6) for the first time (Fig. 3.2). One then attributes the point $\boldsymbol{x}$ to an object of mass $M$ at time $t$. Note that the point $\boldsymbol{x}$ does not belong to any higher mass, since $\delta$ has not exceeded $\delta_{\rm c}(t)$ at any larger filtering scale. Thus, $M$ is indeed the largest mass that the point $\boldsymbol{x}$ is attributed to at time $t$. In this way, one can take care of the cloud-in-cloud problem mentioned in the previous section.

Next step is to calculate the fraction of the trajectories of $\delta$ that make their first upcrossings through the threshold $\delta_{\rm c}(t)$ at $\sigma^2(M)$. In order to do so, one needs to know more details about the trajectories of $\delta$. The properties of the trajectories depend on the correlation between each step, namely, the form of the window function chosen. Although the analyses of these trajectories for general window functions are quite complicated, one can obtain a simple analytic description in the case of sharp $k$-space filtering, i.e. spherical top-hat filtering in $k$-space:

$$\hat{W}_M(k) \equiv \begin{cases} 1 & \text{for } k < k_{\rm s}(M) \\ 0 & \text{for } k > k_{\rm s}(M) \end{cases}, \tag{3.10}$$

where $k_{\rm s}$ is the cut off wave number which corresponds to a mass scale $M$. In this case, the increment to $\delta$ when one decreases $M$ (or increases $\sigma^2$) is totally uncorrelated with any of the previous steps, and the trajectories of $\delta$ reduces to the true Brownian motion described by the diffusion equation:

$$\frac{\partial Q(\sigma^2,\delta)}{\partial\sigma^2} = \frac{1}{2}\frac{\partial^2 Q(\sigma^2,\delta)}{\partial\delta^2}, \tag{3.11}$$

where $Q(\sigma^2,\delta){\rm d}\delta$ stands for the probability that a trajectory reaches $(\sigma^2, \delta \sim \delta + d\delta)$.

The problem is now identical to solving equation (3.11) with the boundary condition that there exists an absorbing barrier at $\delta = \delta_{\rm c}(t)$. The solution to this problem is given by (Chandrasekhar 1943)

$$Q[\sigma^2,\delta,\delta_{\rm c}(t)] = \frac{1}{\sqrt{2\pi\sigma^2(M)}}\left[\exp\left\{-\frac{\delta^2}{2\sigma^2(M)}\right\} - \exp\left\{-\frac{[\delta-2\delta_{\rm c}(t)]^2}{2\sigma^2(M)}\right\}\right]. \tag{3.12}$$

This is the probability distribution of trajectories that reach $(\sigma^2, \delta \sim \delta + {\rm d}\delta)$ without crossing the barrier of height $\delta_{\rm c}(t)$. Then the probability that a trajectory will be absorbed by the barrier at $\sigma^2$ per interval ${\rm d}\sigma^2$ is calculated as

$$\frac{{\rm d}P[\sigma^2,\delta_{\rm c}(t)]}{{\rm d}\sigma^2} = \frac{\partial}{\partial\sigma^2}\int_{\delta_{\rm c}(t)}^{\infty} Q[\sigma^2,\delta,\delta_{\rm c}(t)]{\rm d}\delta$$

$$= \frac{1}{\sqrt{2\pi}} \frac{\delta_{\rm c}(t)}{\sigma^3(M)} \exp\left[-\frac{\delta_{\rm c}^2(t)}{2\sigma^2(M)}\right]. \tag{3.13}$$

Equation (3.13) represents the probability that a given point in space resides in an object of mass corresponding to $\sigma^2(M)$ at time $t$. Therefore, the comoving number density of virialized systems of mass $M$ at time $t$, per interval $\mathrm{d}M$, is expressed as

$$\begin{aligned} \frac{\mathrm{d}n(M,t)}{\mathrm{d}M} &\equiv \frac{\rho_0}{M} \frac{\mathrm{d}P[\sigma^2, \delta_{\rm c}(t)]}{\mathrm{d}\sigma^2} \left|\frac{\mathrm{d}\sigma^2(M)}{\mathrm{d}M}\right| \\ &= \sqrt{\frac{2}{\pi}} \frac{\rho_0}{M} \frac{\delta_{\rm c}(t)}{\sigma^2(M)} \left|\frac{\mathrm{d}\sigma(M)}{\mathrm{d}M}\right| \exp\left[-\frac{\delta_{\rm c}^2(t)}{2\sigma^2(M)}\right]. \end{aligned} \tag{3.14}$$

This exactly reproduces equation (3.8) without introducing any fudge numerical factor by hand.

For general non-sharp $k$-space filters (e.g. Gaussian or top hat), the process is non-Markovian and the trajectories of $\delta$ no longer obey equation (3.11). Bond et al. (1991) showed, by numerical calculations, that the different filtering choices in the above analysis will give mass functions with differing low mass slopes, and may change the abundance at higher masses by a factor of $\sim 2$.

### 3.2.2 Merger probabilities

The above approach is readily extended to describing transitions of an arbitrary mass to another. Consider trajectories of $\delta$ against $\sigma^2(M)$ as in Section 3.2.1. Suppose now that one of the trajectories makes its first upcrossing through a barrier of height $\delta_{\rm c}(t_2)$ at $\sigma^2(M_2)$, and then continues its way until it eventually crosses a second barrier of height $\delta_{\rm c}(t_1)\,[> \delta_{\rm c}(t_2)]$ at $\sigma^2(M_1)\,[> \sigma^2(M_2)]$ as shown in Figure 3.3. This trajectory represents the process that a point in space residing in an object of mass $M_1$ at time $t_1$ is merged into an object of mass $M_2 (> M_1)$ at time $t_2 (> t_1)$. (Note that $\delta_{\rm c}(t)$ and $\sigma^2(M)$ are decreasing functions of $t$ and $M$ respectively.) One can obtain the conditional probability that a trajectory, starting from $[\sigma^2(M_2), \delta_{\rm c}(t_2)]$, will make its first upcrossing through $\delta_{\rm c}(t_1)$ at $\sigma^2(M_1)$, by simply replacing $\sigma^2(M)$ with $\sigma^2(M_1) - \sigma^2(M_2)$, and $\delta_{\rm c}(t)$ with $\delta_{\rm c}(t_1) - \delta_{\rm c}(t_2)$ in equation (3.13):

$$\frac{\mathrm{d}P_1(M_1, t_1 | M_2, t_2)}{\mathrm{d}M_1} = \frac{1}{\sqrt{2\pi}} \frac{\delta_{\rm c1} - \delta_{\rm c2}}{(\sigma_1^2 - \sigma_2^2)^{3/2}} \left|\frac{\mathrm{d}\sigma_1^2}{\mathrm{d}M_1}\right| \exp\left[-\frac{(\delta_{\rm c1} - \delta_{\rm c2})^2}{2(\sigma_1^2 - \sigma_2^2)}\right], \tag{3.15}$$

where simplified notations $\sigma_i \equiv \sigma(M_i)$ and $\delta_{{\rm c}i} \equiv \delta_{\rm c}(t_i)$ are used. Equation (3.15) corresponds to the conditional probability per interval $\mathrm{d}M_1$ that a mass element is part of an object of mass $M_1$ at time $t_1$, provided it becomes part of a larger object of mass $M_2$ at later time $t_2$. The same expression was obtained in a more heuristical approach by Bower (1991). Alternatively, the reverse conditional probability that a trajectory with a first upcrossing of $\delta_{\rm c}(t_1)$ at $\sigma^2(M_1)$ makes its first upcrossing of $\delta_{\rm c}(t_2)$ at $\sigma^2(M_2)$ is, from equations (3.13) and (3.15),

$$\frac{\mathrm{d}P_2(M_2, t_2 | M_1, t_1)}{\mathrm{d}M_2} = \frac{\mathrm{d}P_1(M_1, t_1 | M_2, t_2)/\mathrm{d}M_1 \times \mathrm{d}P(M_2, t_2)/\mathrm{d}M_2}{\mathrm{d}P(M_1, t_1)/\mathrm{d}M_1}$$

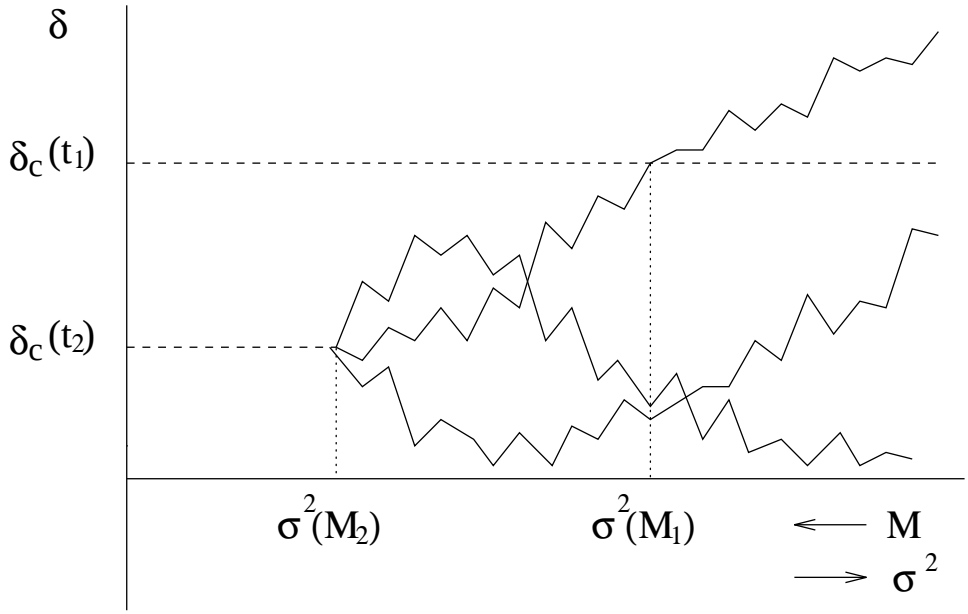


Figure 3.3: Trajectories of $\delta$ versus $\sigma^2$ which make their first upcrossing of a barrier of height $\delta_{\rm c}(t_2)$ at $\sigma^2(M_2)$ and then continue. Some of the trajectories cross through the second barrier of height $\delta_{\rm c}(t_1)(>\delta_{\rm c}(t_2))$ at $\sigma^2(M_1)(>\sigma^2(M_2))$.

$$= \frac{1}{\sqrt{2\pi}} \frac{\delta_{\rm c2}(\delta_{\rm c1}-\delta_{\rm c2})}{\delta_{\rm c1}} \left[\frac{\sigma_1^2}{\sigma_2^2(\sigma_1^2-\sigma_2^2)}\right]^{3/2} \left|\frac{{\rm d}\sigma_2^2}{{\rm d}M_2}\right| \exp\left[-\frac{(\sigma_2^2\delta_{\rm c1}-\sigma_1^2\delta_{\rm c2})^2}{2\sigma_1^2\sigma_2^2(\sigma_1^2-\sigma_2^2)}\right], \qquad (3.16)$$

where ${\rm d}P(M,t)/{\rm d}M = {\rm d}P[\sigma^2,\delta_{\rm c}(t)]/{\rm d}\sigma^2 \times |{\rm d}\sigma^2/{\rm d}M|$ is given by equation (3.13). Equation (3.16) is the conditional probability per interval ${\rm d}M_2$ that an element is part of an object of mass $M_2$ at time $t_2$, provided it has been part of a smaller object of mass $M_1$ at earlier time $t_1$.

Lacey & Cole (1994) performed extensive N-body simulations with $128^3$ particles to check the validity of the above formulae. They found that the analytical merger probabilities agree well with the numerical simulations over almost three orders of magnitude in mass, which is close to the full dynamical range of their simulations.

### 3.2.3 Estimate of halo formation epochs

Based upon the merger probabilities in the last subsection, Lacey & Cole (1993, hereafter LC) proposed a way of estimating halo formation epochs. They defined a "formation epoch" of an object as the time when at least half of its mass is first assembled, and computed the probability distribution function of such epochs as follows.

Consider objects of mass $M$ at time $t$, and their parents of mass $M_1(<M)$ at time $t_1(<t)$. If $M_1$ is greater than $M/2$, then each of the parents should be the largest parent which evolves into a distinct object of mass $M$. The number density of such parents per ${\rm d}M_1$ and ${\rm d}M$ is

written as

$$\frac{\mathrm{d}^2 n(M_1,t_1,M,t)}{\mathrm{d}M_1 \mathrm{d}M} = \frac{\mathrm{d}P_2(M,t|M_1,t_1)}{\mathrm{d}M}\frac{\mathrm{d}n(M_1,t_1)}{\mathrm{d}M_1} \quad \text{for } M_1 > M/2. \tag{3.17}$$

Therefore the probability that an object of mass $M$ at time $t$ has the largest parent of mass greater than $M/2$ at time $t_1$ is given by

$$\begin{aligned} p(>M/2,t_1|M,t) &= \int_{M/2}^{M} \left[\frac{\mathrm{d}^2 n(M_1,t_1,M,t)}{\mathrm{d}M_1 dM} \Big/ \frac{\mathrm{d}n(M,t)}{\mathrm{d}M}\right] \mathrm{d}M_1, \\ &= \int_{M/2}^{M} \left(\frac{M}{M_1}\right) \frac{\mathrm{d}P_1(M_1,t_1|M,t)}{\mathrm{d}M_1} \mathrm{d}M_1, \end{aligned} \tag{3.18}$$

where we have used the first line of equation (3.16). This is equal to the probability that the formation epoch $t_\mathrm{f}$ according to the above definition is before time $t_1$. The distribution of $t_\mathrm{f}$ is thus written as

$$\begin{aligned} \frac{\mathrm{d}p}{\mathrm{d}t_\mathrm{f}}(M,t_\mathrm{f},t) &\equiv -\frac{\partial}{\partial t_1} p(>M/2,t_1|M,t) \\ &= \frac{1}{\sqrt{2\pi}} \int_0^1 \mathrm{d}\tilde{S} \frac{M}{M_1(\tilde{S})} \frac{1}{\tilde{S}^{3/2}} \left(1-\frac{\tilde{\omega}_\mathrm{f}^2}{\tilde{S}}\right) \exp\left(-\frac{\tilde{\omega}_\mathrm{f}^2}{2\tilde{S}}\right) \left[-\frac{\mathrm{d}\tilde{\omega}_\mathrm{f}}{\mathrm{d}t_\mathrm{f}}\right], \end{aligned} \tag{3.19}$$

where

$$\tilde{\omega}_\mathrm{f}(M,t_\mathrm{f},t) \equiv \frac{\delta_\mathrm{c}(t_\mathrm{f}) - \delta_\mathrm{c}(t)}{\sqrt{\sigma^2(M/2)-\sigma^2(M)}}, \tag{3.20}$$

$$\tilde{S}(M_1,M) \equiv \frac{\sigma^2(M_1)-\sigma^2(M)}{\sigma^2(M/2)-\sigma^2(M)}, \tag{3.21}$$

and $M_1(\tilde{S})$ is computed by solving equation (3.21) for $M_1$. Figure 3.4 shows $\mathrm{d}p/\mathrm{d}z_\mathrm{f} \equiv \mathrm{d}p/\mathrm{d}t_\mathrm{f} \times |\mathrm{d}t_\mathrm{f}/\mathrm{d}z_\mathrm{f}|$ in CDM models with $\lambda_0 = 0$, $h = 0.7$ and $\sigma_8 = 1$. If cosmological parameters are fixed, the distribution of formation epochs has its mean at higher redshift with larger dispersion, as mass becomes smaller. Alternatively, if the mass of an object is fixed, the mean formation epoch becomes earlier for lower $\Omega_0$.

As noted by LC, there is a problem with the formation epoch distribution derived here. In the case of power-law matter fluctuation spectra $P(k) \propto k^n$ with index $n > 0$, equation (3.19) gives a negative probability at $t_\mathrm{f} \simeq t$. This implies that the above definition of the halo formation distribution function is not completely self-consistent. Since the effective power index of the observed fluctuation spectrum is negative below the cluster scale (e.g. Peacock & Dodds 1994), this would not be a serious problem in practice for most astrophysically interesting objects. It is nevertheless important to explore other prescriptions for computing the distribution of halo formation epochs. We will propose one of such alternatives in the next section.

## 3.3 Formation and destruction rates

### 3.3.1 Motivation

One of the key quantities in describing the dynamical history of dark matter halos is the rate at which halos of a given mass are *created* at a given epoch. While the PS mass function

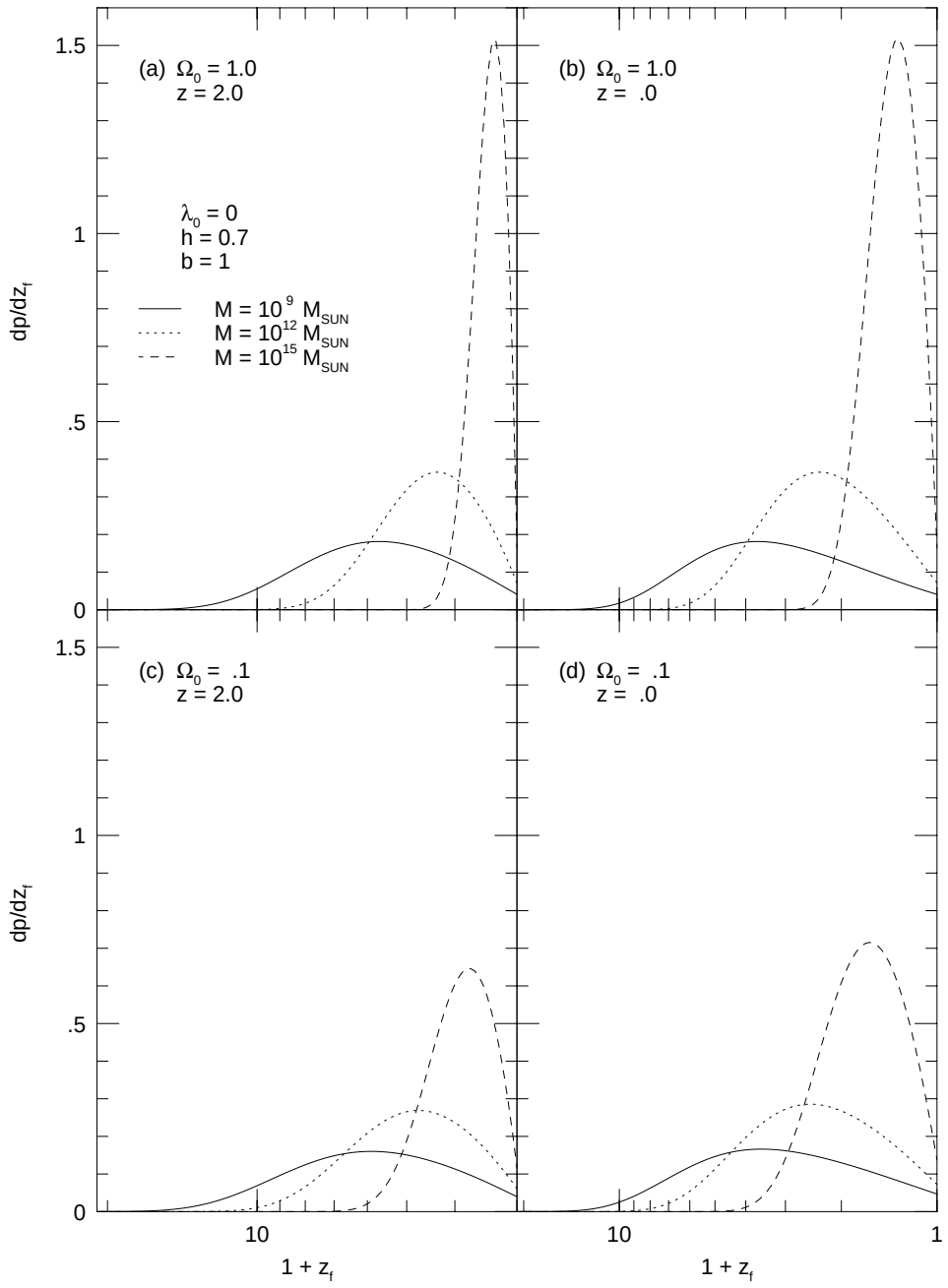


Figure 3.4: The formation epoch distribution function $\partial p/\partial z_{\rm f}$ by LC in CDM models; (a) $\Omega_0 = 1$, $z = 2$, (b) $\Omega_0 = 1$, $z = 0$, (c) $\Omega_0 = 0.1$, $z = 2$, (d) $\Omega_0 = 0.1$, $z = 0$. In all panels, $\lambda_0 = 0$, $h = 0.7$, and $b = 1/\sigma_8 = 1$. Lines indicate $M = 10^9\,{\rm M}_\odot$ (solid), $10^{12}\,{\rm M}_\odot$ (dotted), and $10^{15}\,{\rm M}_\odot$ (dashed).

provides a reasonable approximation to the number of halos *present* at a given epoch, it is not directly related to this quantity. One may argue naively that the time derivative of the PS mass function would give the number of halos formed in unit time. In fact, several authors have used such a prescription to predict the rate of formation of astrophysical objects (e.g., Efstathiou & Rees 1988; Haehnelt & Rees 1993; Fukugita & Kawasaki 1994). However, one needs to be very careful to do so. Let us differentiate the PS mass function (3.8) with respect to time:

$$\begin{aligned} R_{\rm PS}(M,t) &\equiv \frac{\rm d}{{\rm d}t}\left[\frac{{\rm d}n(M,t)}{{\rm d}M}\right], \\ &= \left[\frac{\delta_{\rm c}(t)}{\sigma^2(M)} - \frac{1}{\delta_{\rm c}(t)}\right]\left[-\frac{{\rm d}\delta_{\rm c}(t)}{{\rm d}t}\right]\frac{{\rm d}n(M,t)}{{\rm d}M}. \end{aligned} \tag{3.22}$$

The above expression is positive only for $\sigma(M) < \delta_{\rm c}(t)$ since ${\rm d}\delta_{\rm c}(t)/{\rm d}t < 0$. It becomes negative if $\sigma(M)$ exceeds $\delta_{\rm c}(t)$ at low masses as is the case of almost all physically plausible spectra including that of the CDM model. This clearly indicates that $R_{\rm PS}(M,t)$ should not be simply interpreted as the formation rate of bound objects, but rather corresponds to the net rate of change in the number density. In other words, one can write down $R_{\rm PS}(M,t)$ in terms of the formation rate $R_{\rm form}(M,t)$ at which objects of a given mass $M$ are formed by mergers of smaller mass objects, and the destruction rate $R_{\rm dest}(M,t)$ at which they are destroyed by coalescence to produce more massive systems (Cavaliere, Colafrancesco & Scaramella 1991; Blain & Longair 1993):

$$R_{\rm PS}(M,t) \equiv R_{\rm form}(M,t) - R_{\rm dest}(M,t). \tag{3.23}$$

The negative $R_{\rm PS}(M,t)$ at low masses is then ascribed to an excess of the destruction rate over the formation rate. On the other hand, the formation rate dominates at higher masses resulting in the positive $R_{\rm PS}(M,t)$.

Blain & Longair (1993a,b) and Sasaki (1994) correctly realized the above point, and attempted to obtain the rates of formation and destruction separately, assuming some empirical forms for the rates. In the following, we derive more general formulae for these rates on the basis of merger probabilities discussed in §3.2.2.

### 3.3.2 Formal procedure

From equations (3.15) and (3.16), one can compute instantaneous transition rates from a certain mass to another. Setting $t_1 = t - \Delta t$, $t_2 = t$ and $M_2 = M$ in equation (3.15), and taking the limit $\Delta t \to 0$, we obtain

$$\begin{aligned} \frac{{\rm d}^2 P_1(M_1|M;t)}{{\rm d}M_1 {\rm d}t} &\equiv \lim_{\Delta t\to 0}\frac{{\rm d}P_1(M_1, t-\Delta t|M,t)/{\rm d}M_1}{\Delta t} \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{(\sigma_1^2-\sigma^2)^{3/2}}\left[-\frac{{\rm d}\delta_{\rm c}(t)}{{\rm d}t}\right]\left|\frac{{\rm d}\sigma_1^2}{{\rm d}M_1}\right|. \end{aligned} \tag{3.24}$$

This quantity is interpreted as the rate per ${\rm d}M_1$ and ${\rm d}t$ at which an object of mass $M(>M_1)$ is formed from an object of mass $M_1$. Similarly from equation (3.16), one finds

$$\frac{{\rm d}^2 P_2(M_2|M;t)}{{\rm d}M_2 {\rm d}t} \equiv \lim_{\Delta t\to 0}\frac{{\rm d}P_2(M_2, t+\Delta t|M,t)/{\rm d}M_2}{\Delta t}$$

$$= \frac{1}{\sqrt{2\pi}} \left[\frac{\sigma^2}{\sigma_2^2(\sigma^2-\sigma_2^2)}\right]^{\frac{3}{2}} \left[-\frac{\mathrm{d}\delta_{\rm c}(t)}{\mathrm{d}t}\right] \left|\frac{\mathrm{d}\sigma_2^2}{\mathrm{d}M_2}\right| \exp\left[-\frac{(\sigma^2-\sigma_2^2)\delta_{\rm c}^2(t)}{2\sigma^2\sigma_2^2}\right]. \tag{3.25}$$

Equation (3.25) corresponds to the rate per $\mathrm{d}M_2$ and $\mathrm{d}t$ at which an object of mass $M(<M_2)$ is incorporated into a larger object of mass $M_2$.

Now we can write down the rates of formation and destruction of gravitationally bound systems using the transition rates derived above. From equation (3.24), we formally define the rate of formation of an object of mass $M$ at time $t$ as

$$\begin{aligned} R_{\rm form}(M,t) &\equiv \int_0^M \mathrm{d}M_1 \frac{\mathrm{d}^2P_1(M_1|M;t)}{\mathrm{d}M_1\mathrm{d}t} \frac{\mathrm{d}P(M,t)}{\mathrm{d}M} \frac{\rho_0}{M} \\ &= \left[\frac{1}{\sqrt{\pi}} \frac{\delta_{\rm c}(t)}{\sigma^2(M)} \int_0^\infty \frac{\mathrm{d}x}{x^2}\right] \left[-\frac{\mathrm{d}\delta_{\rm c}(t)}{\mathrm{d}t}\right] \frac{\mathrm{d}n(M,t)}{\mathrm{d}M}. \end{aligned} \tag{3.26}$$

Equation (3.25) similarly defines the formal destruction rate:

$$\begin{aligned} R_{\rm dest}(M,t) &\equiv \int_M^\infty \mathrm{d}M_2 \frac{\mathrm{d}^2P_2(M_2|M;t)}{\mathrm{d}M_2\mathrm{d}t} \frac{\mathrm{d}P(M,t)}{\mathrm{d}M} \frac{\rho_0}{M} \\ &= \left[\frac{1}{\delta_{\rm c}(t)} + \frac{1}{\sqrt{\pi}} \frac{\delta_{\rm c}(t)}{\sigma^2(M)} \int_0^\infty \frac{\mathrm{e}^{-y^2}}{y^2}\mathrm{d}y\right] \left[-\frac{\mathrm{d}\delta_{\rm c}(t)}{\mathrm{d}t}\right] \frac{\mathrm{d}n(M,t)}{\mathrm{d}M}. \end{aligned} \tag{3.27}$$

Clearly, these formal expressions diverge, reflecting the divergence of equations (3.24) and (3.25) in the limits $M_1 \to M$ and $M_2 \to M$, respectively. Such a divergence would simply be related to the fact that the above formalism identifies the formation and destruction of objects even if the associated change of mass is infinitesimal. In reality, however, such a small change in mass hardly corresponds to the formation or destruction in the usual sense; it should rather be regarded as an accretion or evolution of the same system. In order to incorporate a more realistic and physically reasonable description of these events, we should replace the above definitions of formation and destruction with more appropriate ones. Then the divergence associated with equations (3.26) and (3.27) would disappear naturally.

We argue that the divergence is ascribed to counting the transition of a certain object to "itself", by showing explicitly that equations (3.26) and (3.27) diverge in a similar manner. In fact, one can prove that the divergences in the two rates exactly cancel out and their difference reduces to the time derivative of the PS mass function:

$$\begin{aligned} R_{\rm form}(M,t) - R_{\rm dest}(M,t) &= \left[\frac{\delta_{\rm c}(t)}{\sigma^2(M)} - \frac{1}{\delta_{\rm c}(t)}\right] \left[-\frac{\mathrm{d}\delta_{\rm c}(t)}{\mathrm{d}t}\right] \frac{\mathrm{d}n(M,t)}{\mathrm{d}M} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\mathrm{d}n(M,t)}{\mathrm{d}M}\right] \end{aligned} \tag{3.28}$$

The above result also suggests that the divergence in equations (3.26) and (3.27) can be removed by redefining the formation and destruction events properly. We will propose such a prescription below.

### 3.3.3 A realistic prescription

Since the divergence mentioned above originates from the ambiguity in the definition of formation and destruction, the simplest remedy is to introduce threshold masses for these

processes. More specifically, we replace equations (3.26) and (3.27) with

$$R_{\rm form}(M,t;M_{\rm f}) \equiv \int_0^{M_{\rm f}} {\rm d}M_1 \frac{{\rm d}^2 P_1(M_1|M;t)}{{\rm d}M_1 {\rm d}t} \frac{{\rm d}P(M,t)}{{\rm d}M} \frac{\rho_0}{M}, \tag{3.29}$$

and

$$R_{\rm dest}(M,t;M_{\rm d}) \equiv \int_{M_{\rm d}}^{\infty} {\rm d}M_2 \frac{{\rm d}^2 P_2(M_2|M;t)}{{\rm d}M_2 {\rm d}t} \frac{{\rm d}P(M,t)}{{\rm d}M} \frac{\rho_0}{M}. \tag{3.30}$$

The above procedure defines the formation and destruction of an object of mass $M$ if the mass of the progenitor $M_1$ is smaller than $M_{\rm f}$ and the mass of the merged object $M_2$ is greater than $M_{\rm d}$, respectively. Equations (3.29) and (3.30) count the number of objects corresponding to the fraction of mass density that has come from progenitors with $M_1 < M_{\rm f}$ or is going into objects with $M_2 > M_{\rm d}$.

Adopting such a prescription, the above formulae are integrated to give

$$R_{\rm form}(M,t;M_{\rm f}) = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\sigma^2(M_{\rm f}) - \sigma^2(M)}} \left[ -\frac{{\rm d}\delta_{\rm c}(t)}{{\rm d}t} \right] \frac{{\rm d}n(M,t)}{{\rm d}M}, \tag{3.31}$$

and

$$\begin{aligned} R_{\rm dest}(M,t;M_{\rm d}) = & \left[ \left( \frac{1}{\delta_{\rm c}(t)} - \frac{\delta_{\rm c}(t)}{\sigma^2(M)} \right) {\rm erfc}(Z) + \frac{1}{\sqrt{\pi}} \frac{\delta_{\rm c}(t)}{\sigma^2(M)} \frac{{\rm e}^{-Z^2}}{Z} \right] \\ & \times \left[ -\frac{{\rm d}\delta_{\rm c}(t)}{{\rm d}t} \right] \frac{{\rm d}n(M,t)}{{\rm d}M}, \end{aligned} \tag{3.32}$$

where ${\rm erfc}(u)$ is the complimentary error function (eq. [3.4]), and $Z$ is defined as

$$Z(M,t;M_{\rm d}) \equiv \delta_{\rm c}(t) \sqrt{\frac{\sigma^2(M) - \sigma^2(M_{\rm d})}{2\sigma^2(M)\sigma^2(M_{\rm d})}}. \tag{3.33}$$

In the above prescription, it is important to choose physically reasonable values for $M_{\rm f}$ and $M_{\rm d}$. Admittedly it is inevitably artificial. We here propose to set $M_{\rm f} = M/2$ and $M_{\rm d} = 2M$, mainly for the following two reasons. One is similar to that mentioned in §3.2.3; the identity of an object of mass $M$ is clearly retained between $M/2$ and $2M$. If a new object of mass $M$ were formed from a parent object of mass greater than $M/2$, the parent object would make up the dominant part of the new object. Therefore such a process should correspond to evolution of the parent object rather than formation of a new object. The same is true for the destruction of an object of mass $M$ which is incorporated into an object of mass less than $2M$.

The other reason is more practical: equations (3.31) and (3.32) are rather insensitive to the specific choice of $M_{\rm f}$ and $M_{\rm d}$ at $M_{\rm f} \simeq M/2$ and $M_{\rm d} \simeq 2M$, and their difference agrees fairly well with $R_{\rm PS}(M,t)$. To see such features more clearly, we plot $R_{\rm form}(M,z;M_{\rm f})$, $R_{\rm dest}(M,z;M_{\rm d})$ and their difference against $M/M_{\rm f}$ or $M_{\rm d}/M$ in Figure 3.5 (each quantity is normalized by $|R_{\rm PS}(M,z)|$). The figure shows that $R_{\rm form}/|R_{\rm PS}|$ and $R_{\rm dest}/|R_{\rm PS}|$ are fairly insensitive to $M/M_{\rm f}$ or $M_{\rm d}/M$ for $M/M_{\rm f}, M_{\rm d}/M \gtrsim 1.5$, and that $R_{\rm form} - R_{\rm dest} \simeq R_{\rm PS}$ within 30 per cent for $M/M_{\rm f}, M_{\rm d}/M \lesssim 3$. Thus, for practical purposes as well, the choices of $M_{\rm f} = M/2$ and $M_{\rm d} = 2M$ seem reasonable.

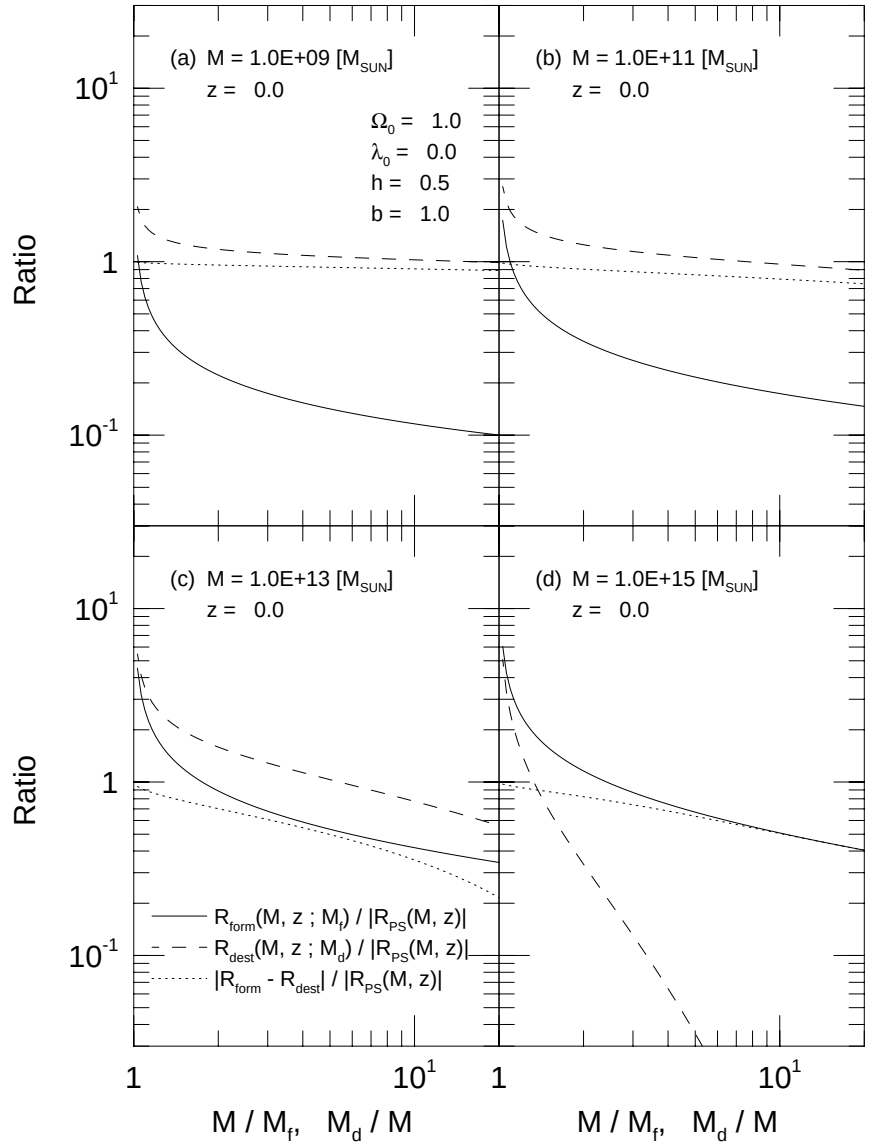


Figure 3.5: Dependence of $R_{\rm form}(M,z;M_{\rm f})$ (solid curve), $R_{\rm dest}(M,z;M_{\rm d})$ (dashed), and their difference (dotted) on $M/M_{\rm f}$ or $M_{\rm d}/M$ at $z=0$ in the standard CDM model with $\Omega_0=1$, $h=0.5$, and $b=1/\sigma_8=1$ (each quantity is normalized by $|R_{\rm PS}(M,z)|$): (a) $M=10^9\,{\rm M}_\odot$, (b) $M=10^{11}\,{\rm M}_\odot$, (c) $M=10^{13}\,{\rm M}_\odot$, (d) $M=10^{15}\,{\rm M}_\odot$.

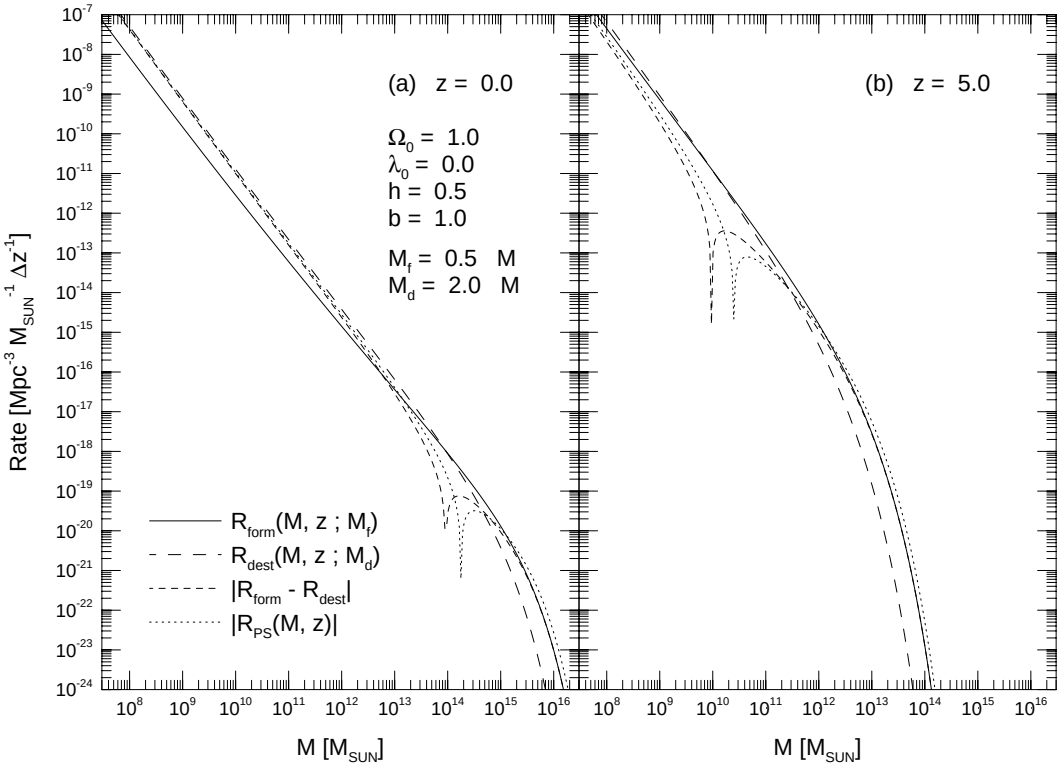


Figure 3.6: Rates of formation $R_{\rm form}(M,z;M_{\rm f})$ (solid curve) and destruction $R_{\rm dest}(M,z;M_{\rm d})$ (long-dashed) in the case of $M_{\rm f}=M/2$ and $M_{\rm d}=2M$, as a function of mass in the standard CDM model with $\Omega_0=1$, $h=0.5$, and $b=1/\sigma_8=1$. Also plotted are the absolute values of their difference (short-dashed) and $R_{\rm PS}(M,z)$ (dotted): (a) $z=0$, (b) $z=5$.

Figure 3.6 illustrates $R_{\rm form}(M,z;M/2)$, $R_{\rm dest}(M,z;2M)$, their difference, and $|R_{\rm PS}(M,z)|$ versus $M$ in the standard CDM model. Both panels demonstrate that $R_{\rm form} < R_{\rm dest}$ for small masses and $R_{\rm form} > R_{\rm dest}$ for large masses. The critical mass $M_{\rm c}$ at which $R_{\rm form} = R_{\rm dest}$ increases with time: $M_{\rm c} \simeq 10^{10}\,{\rm M}_\odot$ at $z = 5$, and $M_{\rm c} \simeq 10^{14}\,{\rm M}_\odot$ at $z = 0$ in this example. The difference between two rates is very close to $R_{\rm PS}$ except at $M \simeq M_{\rm c}$. The deviation arises from the fact that we set $M_{\rm f}, M_{\rm d} \neq M$, but is not important in practice because its amplitude is several orders of magnitude smaller than that of $R_{\rm form}$ or $R_{\rm dest}$. With these results in mind, we hereafter adopt $M_{\rm f} = M/2$ and $M_{\rm d} = 2M$, and simply use $R_{\rm form}(M,t)$ and $R_{\rm dest}(M,t)$ to denote $R_{\rm form}(M,t;M/2)$ and $R_{\rm dest}(M,t;2M)$ respectively.

### 3.3.4 Estimate of halo formation epochs

Using the rate of formation derived in the previous subsection, we now attempt to compute the distribution function for the halo formation epochs. In order to do so, we need a probability that an object once formed remains, without destructed, until some later epoch. For this purpose, we adopt "the survival probability" evaluated by LC from equation (3.16) as

$$P_{\rm surv}(M,t_1,t_2) \equiv \int_M^{2M} \frac{{\rm d}P_2(M_2,t_2|M,t_1)}{{\rm d}M_2}{\rm d}M_2$$
$$= 1 - \frac{1}{2}\left[\frac{\delta_{\rm c1}-2\delta_{\rm c2}}{\delta_{\rm c1}}\right]\exp\left[\frac{2\delta_{\rm c2}(\delta_{\rm c1}-\delta_{\rm c2})}{\sigma^2(M)}\right]{\rm erfc}[X(M,t_1,t_2)] - \frac{1}{2}{\rm erfc}[Y(M,t_1,t_2)], \quad (3.34)$$

where $X(M,t_1,t_2)$ and $Y(M,t_1,t_2)$ are defined respectively as

$$X(M,t_1,t_2) \equiv \frac{\sigma^2(2M)[\delta_{\rm c1}-2\delta_{\rm c2}]+\sigma^2(M)\delta_{\rm c2}}{\sqrt{2\sigma^2(M)\sigma^2(2M)[\sigma^2(M)-\sigma^2(2M)]}}, \quad (3.35)$$

and

$$Y(M,t_1,t_2) \equiv \frac{\sigma^2(M)\delta_{\rm c2}-\sigma^2(2M)\delta_{\rm c1}}{\sqrt{2\sigma^2(M)\sigma^2(2M)[\sigma^2(M)-\sigma^2(2M)]}}. \quad (3.36)$$

Equation (3.34) gives the probability that an object of mass $M$ at time $t_1$ remains to have mass less than $2M$ at later time $t_2 (> t_1)$. Strictly speaking, the above definition of the survival probability is based on a slightly different definition for the destruction of an object from our current approach. According to the definition in §3.3.3, an object is thought to have destructed if its mass doubles *instantaneously*, rather than *continuously* as is assumed in equation (3.34). Since the latter definition is more strict than the former, equation (3.34) would underestimate the probability based on the definition in §3.3.3. Nevertheless, without any better substitute, we adopt equation (3.34) as a rough estimation.

From the above formulae and equation (3.31), we formulate the number density per ${\rm d}M$ and ${\rm d}t_{\rm f}$ of bound systems that form with mass $M$ at time $t_{\rm f}$ and survive without destructed until a later time $t$ as (Kitayama & Suto 1996a)

$$\frac{{\rm d}^2F_{\rm KS}}{{\rm d}M{\rm d}t_{\rm f}}(M,t_{\rm f},t) \equiv R_{\rm form}(M,t_{\rm f}) \times P_{\rm surv}(M,t_{\rm f},t)$$
$$= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\sigma^2(M/2)-\sigma^2(M)}}\left[2-\left(\frac{\delta_{\rm c}(t_{\rm f})-2\delta_{\rm c}(t)}{\delta_{\rm c}(t)}\right)\exp\left(\frac{2\delta_{\rm c}(t)[\delta_{\rm c}(t_{\rm f})-\delta_{\rm c}(t)]}{\sigma^2(M)}\right)\right.$$
$$\left.\times{\rm erfc}[X(M,t_{\rm f},t)] - {\rm erfc}[Y(M,t_{\rm f},t)]\right]\left[-\frac{{\rm d}\delta_{\rm c}}{{\rm d}t}(t_{\rm f})\right]\frac{{\rm d}n(M,t_{\rm f})}{{\rm d}M}. \quad (3.37)$$

Given mass $M$ and time $t$, equation (3.37) is regarded as the distribution function of the formation epoch $t_{\rm f}$. This quantity is positive-definite by construction unlike equation (3.38) below.

By contrast, the quantity corresponding to equation (3.37) but based on the definition by LC (§3.2.3) can be formulated as

$$\frac{{\rm d}^2 F_{\rm LC}}{{\rm d}M\,{\rm d}t_{\rm f}}(M, t_{\rm f}, t) \equiv 2\frac{{\rm d}p}{{\rm d}t_{\rm f}}(2M, t_{\rm f}, t)\frac{{\rm d}n}{{\rm d}M}(2M, t). \tag{3.38}$$

Note that, in the definition of the halo formation epochs by LC, the mass of a halo at $t_{\rm f}$ and $t$ is different at most by a factor of 2. Equation (3.38) assumes that this factor is close to 2, based on the fact that the mergers are in general dominated by steady accretion of small objects.

As discussed above and also in §3.2.3, neither equation (3.37) or equation (3.38) is a completely self-consistent expression for the distribution of halo formation epochs. In practice, however, these are in good agreement except at $z_{\rm f} \sim z$ as shown in Figure 3.7. The small difference is ascribed to different criteria of formation in the two approaches, because the formation epochs are not uniquely defined for halos which increase their mass quiescently via accretion. We will study the effects that different definitions between equations (3.37) and (3.38) have on the predictions for the cluster abundances in §4.7 and 4.8.

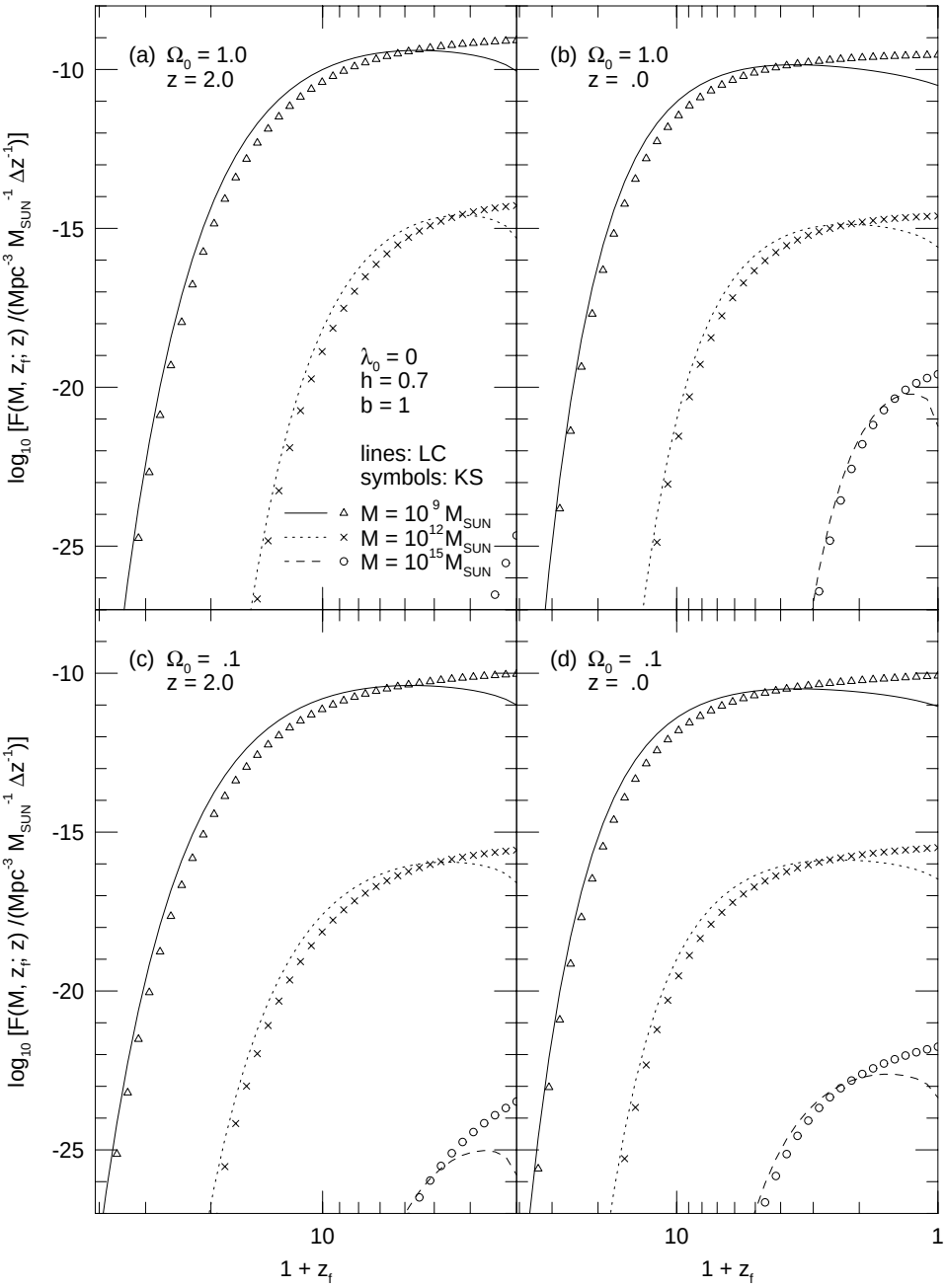


Figure 3.7: The distribution function of halo formation epochs in CDM models; (a) $\Omega_0 = 1$, $z = 2$, (b) $\Omega_0 = 1$, $z = 0$, (c) $\Omega_0 = 0.1$, $z = 2$, (d) $\Omega_0 = 0.1$, $z = 0$. In all panels, $\lambda_0 = 0$, $h = 0.7$, and $b = 1/\sigma_8 = 1$ are assumed. Lines and symbols indicate equations (3.38) (symbols) (3.37) respectively for $M = 10^9\,\mathrm{M}_\odot$ (solid line, triangles), $10^{12}\,\mathrm{M}_\odot$ (dotted line, crosses), and $10^{15}\,\mathrm{M}_\odot$ (dashed line, circles).

# Chapter 4

# X-ray Temperature and Luminosity Functions

On the basis of materials presented in the previous chapters, we are now in a position to make quantitative predictions for the abundances of clusters of galaxies. In this chapter, we start out with studying the X-ray temperature and luminosity functions (XTF and XLF). We particularly focus on demonstrating characteristic features of the cluster abundances by means of these statistics in order to clarify the physics that governs them. This will also provide a useful grounding for more detailed comparisons between theory and observation using cluster number counts discussed in Chapter 5.

## 4.1 Observational backgrounds

A standard way to construct the observed XLF is so-called the "$1/V_{\rm max}$ technique" (Schmidt 1968). Given a complete flux-limited sample of galaxy clusters at redshift $z$, the space density of the $i$-th cluster with X-ray luminosity $L_i$ is written as

$$\mathrm{d}n_i(L_i,z) = \frac{1}{V_{{\rm max},i}}, \tag{4.1}$$

where $V_{{\rm max},i}$ is the maximum volume in which the $i$-th cluster could be detected given the flux limit and geometric boundaries of the survey. For each luminosity bin of width $\Delta L$, the differential number density of clusters with X-ray luminosity $L$ is computed by

$$\frac{\mathrm{d}n(L,z)}{\mathrm{d}L} = \frac{1}{\Delta L}\sum_i \mathrm{d}n_i(L_i,z), \tag{4.2}$$

where the sum is over the clusters within the luminosity bin. The cumulative number density of clusters with X-ray luminosity greater than $L$ is

$$n(>L,z) = \sum_{L<L_i} \mathrm{d}n_i(L_i,z). \tag{4.3}$$

The determination of the XTF is in principle not so straightforward as that of the XLF, since the observed sample is most often flux limited rather than temperature limited. However, a good correlation found between X-ray luminosity and temperature (Mushotzky 1984;

Edge & Stewart 1990; David et al 1993) ensures that there should be no serious selection effect at least for the local cluster samples. For this reason, the XTF is usually computed from the flux limited sample simply by replacing $L$ with $T$ in the above equations (4.1)–(4.3). Some caution is still required as the scatter in the $L-T$ correlation could result in missing systematically clusters near the limiting temperature. This effect can be avoided by considering only clusters with sufficiently high temperatures.

The cluster XLF has been compiled by a number of authors, both for local and high redshift samples (e.g. Edge et al. 1990; Gioia et al. 1990; Henry & Arnaud 1991; Henry et al. 1992; Nichol et al. 1997; Collins et al. 1997; Burke et al. 1997; Ebeling et al. 1997; Rosati et al. 1998). The local XLF that contains the largest number of clusters comes from the *ROSAT* Brightest Cluster Sample (hereafter BCS) of Ebeling et al. (1997). The BCS consists of 199 clusters with redshifts $z \leq 0.3$ comprising the largest X-ray flux limited (90% complete down to $S = 4.5 \times 10^{-12}\mathrm{erg\,cm^{-2}\,s^{-1}}$in the 0.1–2.4keV band) and X-ray selected sample to date. The highest redshift XLF so far is from the *ROSAT* Deep Cluster Survey (hereafter RDCS) of Rosati et al. (1998). The RDCS sample is 97% complete down to $S = 4 \times 10^{-14}\mathrm{erg\,cm^{-2}\,s^{-1}}$in the 0.5–2keV band and contains 70 clusters over the redshift range $z = 0.05 - 0.8$.

There has been a claim mainly based on the *Einstein* Extended Medium Sensitivity Survey (EMSS, Gioia et al. 1990; Henry et al. 1992) that the cluster XLF shows strong negative evolution (decreasing number towards past) near its bright end. However, more recent observations by *ROSAT* (including the BCS and RDCS sample mentioned above) find no evidence for such evolution and suggest that the local XLF remains fairly constant up to $z \sim 0.8$ (Collins et al. 1997; Burke et al. 1997; Ebeling et al. 1997; Rosati et al. 1998). The discrepancy might be ascribed to the limited number of high redshift clusters in the EMSS sample and the errors in measuring the extended X-ray emissions by a small size ($2.4' \times 2.4'$) detection box of the *Einstein* Imaging Proportional Counter. In fact, Nichol et al. (1997) reanalysed the same EMSS clusters using the new X-ray/optical data and found that the sample is consistent with non-evolving XLF.

Compared to the XLF, the local XTF data (Edge et al. 1990; Henry & Arnaud 1991) contain a smaller number of clusters, since it is more difficult to measure the temperature of a cluster than the luminosity. Edge et al.'s XTF sample consists of 46 clusters with temperatures measured mainly by either *Einstein* or *EXOSAT*, while Henry & Arnaud (1991) compiled 25 clusters and redetermined their temperatures combining *Einstein*, *EXOSAT*, *HEAO 1* and *OSO* results. Both samples are more than 90% complete down to $S(2\text{–}10\ \mathrm{keV}) \sim 3\times10^{-11}\mathrm{erg\,cm^{-2}\,s^{-1}}$, and their XTF's are in reasonable agreement with each other. Eke et al. (1996) reanalysed the Henry & Arnaud's sample and pointed out that there are errors in their analysis but these errors happen to nearly compensate each other to give an approximately correct XTF.

The XTF at redshift other than zero has become available only recently. Henry (1997) determined the XTF at $z \sim 0.3$ from 10 clusters in the EMSS sample with temperatures remeasured by *ASCA*. The data indicate fairly weak or no evolution of the XTF between $z \sim 0.3$ and $z = 0$. This is consistent with both the little evolution of the XLF mentioned above and that of the $L-T$ correlation (e.g. Mushotzky & Scarf 1997) discussed in §4.3.3.

While the high redshift XLF and XTF do serve as an important probe of the evolution of galaxy clusters, one caution should be noted. As is clear from equations (4.1)–(4.3), one

has to assume specific values of $\Omega_0$ and $\lambda_0$ to determine $V_{\mathrm{max},i}$ and hence the XTF and XLF ($\Omega_0 = 1$ and $\lambda_0 = 0$ are conventionally assumed). For the XLF, in addition, the translation of the observed flux into the absolute luminosity requires to specify $\Omega_0$ and $\lambda_0$. The differences due to specific choices of these parameters are not negligible at high redshifts; at $z \sim 0.5$, for instance, choosing $\Omega_0 = 1$ or $\Omega_0 = 0.3$ would make a difference of greater than a factor of 2 in the XLF constructed from an identical sample. Thus, even if the XLF seems not to evolve for $\Omega_0 = 1$, it could evolve mildly negative in low density universes. This point becomes especially important if one attempts to constrain cosmological parameters from the high redshift XLF or XTF.

## 4.2 Theoretical formulation

As will be shown in §4.3, based upon some physically plausible assumptions and empirical relations, the temperature $T$ and the X-ray luminosity $L$ of a cluster observed at redshift $z$ are related to its mass $M$ and the formation redshift $z_{\mathrm{f}}$. One can therefore translate the mass function derived in Chapter 3 into the XTF and XLF by means of these relations. As discussed in §3.1.2, however, the traditional PS theory does not allow to incorporate explicitly the formation redshift $z_{\mathrm{f}}$. A conventional way to compute the XTF and XLF is then simply to assume $z_{\mathrm{f}} = z$ and use the PS mass function given by equation (3.8) as (e.g. Henry & Arnaud 1991; Eke et al. 1996)

$$\begin{aligned} \frac{\mathrm{d}n(T,z)}{\mathrm{d}T} &= \frac{\mathrm{d}n(M,z)}{\mathrm{d}M}\,\frac{\mathrm{d}M}{\mathrm{d}T}\bigg|_{M=M(T,z)}, \\ \frac{\mathrm{d}n(L,z)}{\mathrm{d}L} &= \frac{\mathrm{d}n(M,z)}{\mathrm{d}M}\,\frac{\mathrm{d}M}{\mathrm{d}L}\bigg|_{M=M(L,z)}, \end{aligned} \tag{4.4}$$

where $\mathrm{d}n(T,z)$ is the comoving number density of clusters with $T \sim T + \mathrm{d}T$ at $z$, and likewise for $\mathrm{d}n(L,z)$.

To be strict, the redshift $z$ at which one observes a cluster should be conceptually distinguished from its formation redshift $z_{\mathrm{f}}$. For this purpose, one can use the alternative formalisms to the PS theory derived in Chapter 3 which take account of this difference explicitly (Kitayama & Suto 1996a,b):

$$\begin{aligned} \frac{\mathrm{d}n(T,z)}{\mathrm{d}T} &= \int_z^\infty \mathrm{d}z_{\mathrm{f}} \frac{\mathrm{d}F}{\mathrm{d}M\mathrm{d}z_{\mathrm{f}}}(M,z_{\mathrm{f}};z)\frac{\mathrm{d}M}{\mathrm{d}T}\bigg|_{M=M(T,z_{\mathrm{f}},z)}, \\ \frac{\mathrm{d}n(L,z)}{\mathrm{d}L} &= \int_z^\infty \mathrm{d}z_{\mathrm{f}} \frac{\mathrm{d}F}{\mathrm{d}M\mathrm{d}z_{\mathrm{f}}}(M,z_{\mathrm{f}};z)\frac{\mathrm{d}M}{\mathrm{d}L}\bigg|_{M=M(L,z_{\mathrm{f}},z)}, \end{aligned} \tag{4.5}$$

where $\mathrm{d}F/\mathrm{d}M\mathrm{d}z_{\mathrm{f}}$ denotes the number density of clusters that form with mass $M$ at redshift $z_{\mathrm{f}}$ and survive until redshift $z$. We compute this quantity from either equation (3.37) or (3.38).

In the following, we will consider both of the above approaches in order to compare different assumptions regarding the cluster formation epochs.

## 4.3 Physical descriptions of X-ray clusters

In order to make predictions for the XTF and XLF, one needs appropriate models to relate the cluster mass to its temperature and X-ray luminosity. We describe below our modelling of the intracluster gas properties.

### 4.3.1 Characteristic time-scales

Dynamics of galaxy clusters is characterized by a time-scale defined by

$$t_{\rm dyn} \equiv \frac{1}{\sqrt{G\rho}} \sim 10^9 \left(\frac{\rho}{200\rho_{\rm c0}}\right)^{-1/2} \quad {\rm yr}, \tag{4.6}$$

where $\rho_{\rm c0} = 1.88\times10^{-29}h^2$ g cm$^{-3}$ is the present-day critical density of the universe. Equation (4.6) is an order of magnitude smaller than the Hubble time at present given by equation (2.28). This means that the galaxy clusters are gravitationally bound and decoupled from the cosmic expansion. The observed radial velocity dispersion of galaxies is also consistent with the value expected from the total mass and size of clusters (Table 1.1), which supports that clusters are indeed virialized systems.

Once a cluster is formed and the intracluster gas is heated up, the gas particles will approach Maxwellian velocity distribution via Coulomb collisions. For electron-electron collisions, the time-scale to achieve the equilibrium is written as

$$t_{\rm eq}(e,e) \equiv \frac{l_{\rm e}}{\langle v_{\rm e}\rangle} \sim 3\times10^5 \left(\frac{T_{\rm e}}{10^8{\rm K}}\right)^{3/2} \left(\frac{n_{\rm e}}{10^{-3}\ {\rm cm}^{-3}}\right)^{-1} \quad {\rm yr}, \tag{4.7}$$

where $l_{\rm e} \simeq 23{\rm kpc}(T_{\rm e}/10^8{\rm K})^2(n_{\rm e}/10^{-3}{\rm cm}^{-3})^{-1}$ is the mean free path of electrons given by Spitzer (1956), $\langle v_{\rm e}\rangle$ is the rms electron velocity, $T_{\rm e}$ is the electron temperature, and $n_{\rm e}$ is the electron number density. The time-scale for protons to equibrate among themselves is $t_{\rm eq}(p,p) \sim (m_{\rm p}/m_{\rm e})^{1/2}t_{\rm eq}(e,e) \sim 43\,t_{\rm eq}(e,e)$. The time-scale for electrons and protons to reach the equal temperature is $t_{\rm eq}(p,e) \sim (m_{\rm p}/m_{\rm e})t_{\rm eq}(e,e) \sim 1870\,t_{\rm eq}(e,e)$. Since these time-scales are in general shorter than the Hubble time and the cooling time-scale discussed below, the intracluster gas can be thought of being in a thermal equilibrium characterized by a single kinetic temperature $T$.

As the density of the intracluster gas is relatively low ($n_{\rm e} \lesssim 10^{-2}$ cm$^{-3}$), the gas is optically thin and the transport of the radiation field can be ignored. At temperatures of $T \sim 10^7 - 10^8$K, the gas is almost fully ionized and produces X-rays mainly due to thermal bremsstrahlung (free-free) emission. The cooling time-scale of the intracluster gas is then estimated as

$$t_{\rm cool} \equiv \frac{n_{\rm e}k_{\rm B}T}{\epsilon^{\rm ff}} \sim 4\times10^{10} \left(\frac{n_{\rm e}}{10^{-3}\ {\rm cm}^{-3}}\right)^{-1} \left(\frac{T}{10^8\ {\rm K}}\right)^{1/2} {\rm yr}, \tag{4.8}$$

where $\epsilon^{\rm ff} \simeq 2\times10^{-27}(T/{\rm K})^{1/2}(n_{\rm e}/{\rm cm}^{-3})^2$ erg cm$^{-3}$ s$^{-1}$ is the thermal bremsstrahlung emissivity, and $k_{\rm B}$ is the Boltzmann constant. In most clusters, the cooling time-scale is longer than either the dynamical time-scale or the Hubble time, and thus radiative cooling is not very important. In a high density central region of some clusters, the cooling time-scale becomes shorter than the Hubble time and the cooled gas component has been actually detected

(e.g., Ikebe 1995). While such cooled gas component may lead to the excess X-ray emission near the cluster center, it contributes only a small fraction of the total gas mass. We neglect the effects of radiative cooling in the rest of this thesis.

As long as the time-scales of collisions among ions and electrons (eq. [4.7]) are short compared with the time-scales for dynamical or radiative processes (eqs [4.6], [4.8]), the intracluster gas can be treated as a fluid. The time required for a sound wave to cross a cluster is given by

$$t_{\rm s} \equiv \frac{R}{c_{\rm s}} \sim 7 \times 10^8 \left(\frac{T}{10^8\ {\rm K}}\right)^{-1/2} \left(\frac{R}{\rm Mpc}\right)\ {\rm yr}, \tag{4.9}$$

where $R$ is the physical size of a cluster, and $c_{\rm s}$ is the sound velocity. As the above sound crossing time-scale is usually shorter than the Hubble time and the cooling time-scale, the gas is often assumed to be hydrostatic.

The above arguments show that the physics of clusters of galaxies are fairly simple compared with those of other astrophysical objects such as stars or galaxies, in which non-equilibrium thermal and dynamical processes play essential role. This fact in turn makes clusters of galaxies a good probe of cosmology.

### 4.3.2 Temperature

We model that the cluster is spherically symmetric and the intracluster gas is in an approximate hydrostatic equilibrium with the binding cluster potential. Under these assumptions, the gas distribution obeys

$$\frac{1}{\rho_{\rm gas}} \frac{{\rm d}p_{\rm gas}}{{\rm d}r} = -\frac{GM_r}{r^2}, \tag{4.10}$$

or equivalently,

$$\frac{k_{\rm B}T}{\mu m_{\rm p}} \left(\frac{{\rm d}\ln \rho_{\rm gas}}{{\rm d}\ln r} + \frac{{\rm d}\ln T}{{\rm d}\ln r}\right) = -\frac{GM_r}{r}, \tag{4.11}$$

where $\rho_{\rm gas}$ is the gas density, $p_{\rm gas} = \rho_{\rm gas} k_{\rm B} T/\mu m_{\rm p}$ is the gas pressure, $r$ is the radius from the cluster center, $M_r$ is the total cluster mass within $r$, $\mu$ is the mean molecular weight, and $m_{\rm p}$ is the proton mass. For a fully ionized plasma dominated by hydrogen (H) and helium (He), the mean molecular weight is given by

$$\mu = \frac{m_{\rm H} n_{\rm H} + m_{\rm He} n_{\rm He}}{m_{\rm H}(2n_{\rm H} + 3n_{\rm He})} = \frac{4}{8 - 5Y}, \tag{4.12}$$

where $m$ and $n$ stand for the mass and number density of individual species, and $Y$ is the helium fraction by weight. Throughout this thesis, we adopt $Y = 0.24$, which gives $\mu = 0.59$.

The galaxy distribution is governed by the same gravitational potential as in equation (4.11) and obeys the Jeans equation:

$$\sigma_{\rm v}^2 \left(\frac{{\rm d}\ln \rho_{\rm gal}}{{\rm d}\ln r} + \frac{{\rm d}\ln \sigma_{\rm v}^2}{{\rm d}\ln r}\right) = -\frac{GM_r}{r}, \tag{4.13}$$

where $\sigma_{\rm v}$ is the radial velocity dispersion, and $\rho_{\rm gal}$ is the galaxy density. If the cluster is further approximated to be isothermal, i.e. ${\rm d}\ln T/{\rm d}\ln r \simeq 0$ and ${\rm d}\ln \sigma_{\rm v}^2/{\rm d}\ln r \simeq 0$, then

equations (4.11) and (4.13) indicate that

$$\beta_{\rm spec} \equiv \frac{\sigma_{\rm v}^2}{k_{\rm B}T/\mu m_{\rm p}}, \tag{4.14}$$

and

$$\beta_{\rm fit} \equiv \frac{{\rm d}\ln\rho_{\rm gas}}{{\rm d}\ln\rho_{\rm gal}}, \tag{4.15}$$

would be equal. The former quantity $\beta_{\rm spec}$ represents the ratio of specific kinetic energy of the galaxies to the specific thermal energy of the gas. Observations of a large sample of clusters yield a mean best-fit value of $\langle\beta_{\rm spec}\rangle \simeq 1.0 \pm 0.1$ (Mushotzky 1984; Lubin & Bahcall 1993), indicating that on average the gas and galaxies follow each other with comparable energies. On the other hand, the observed average profile ${\rm d}\ln\rho_{\rm gas}/{\rm d}\ln r \simeq -2.0 \pm 0.3$ and ${\rm d}\ln\rho_{\rm gal}/{\rm d}\ln r \simeq -2.4 \pm 0.2$ yield $\langle\beta_{\rm fit}\rangle \simeq 0.84 \pm 0.1$ (Bahcall & Lubin 1994). Thus, equations (4.14) and (4.15) are roughly equal to each other. This consistency supports the assumption that the gas is in an approximate hydrostatic equilibrium with the cluster potential.

From equations (4.13) and (4.14), the isothermal temperature of the intracluster gas is expressed as

$$k_{\rm B}T = \frac{\mu m_{\rm p}\sigma_{\rm v}^2}{\beta_{\rm spec}} = -\frac{\mu m_{\rm p} G M_r}{\beta_{\rm spec} r}\left(\frac{{\rm d}\ln\rho_{\rm gal}}{{\rm d}\ln r}\right)^{-1}. \tag{4.16}$$

Based on the spherical model of non-linear collapse described in §2.6, we evaluate this relation at the virial radius within which the mass $M$ is enclosed:

$$\begin{aligned} r_{\rm vir}(M, z_{\rm f}) &= \left(\frac{3M}{4\pi\bar{\rho}\Delta_{\rm c}}\right)^{1/3}, \\ &= 1.69(1+z_{\rm f})^{-1}\left(\frac{\Delta_{\rm c}}{18\pi^2}\right)^{-1/3}\left(\frac{M}{10^{15}h^{-1}\,{\rm M}_\odot}\right)^{1/3}\Omega_0^{-1/3}h^{-1}\ {\rm Mpc}, \end{aligned} \tag{4.17}$$

where $\bar{\rho}$ is the mean density of the universe, and $\Delta_{\rm c}(\Omega_0, \lambda_0, z_{\rm f})$ is the mean density of a cluster virialized at $z_{\rm f}$ in units of $\bar{\rho}$ (eqs [2.69]–[2.71]). Then equation (4.16) is rewritten in terms of $M$ and $z_{\rm f}$ as

$$\begin{aligned} k_{\rm B}T &= \gamma\frac{\mu m_{\rm p} G M}{3 r_{\rm vir}} \\ &= 5.2\ \gamma\left(\frac{\Delta_{\rm c}}{18\pi^2}\right)^{1/3}\left(\frac{M}{10^{15}h^{-1}\,{\rm M}_\odot}\right)^{2/3}(1+z_{\rm f})\,\Omega_0^{1/3}\ {\rm keV}, \end{aligned} \tag{4.18}$$

where $\gamma$ is a fudge factor of order unity introduced to take account of uncertainties in both $\beta_{\rm spec}$ and ${\rm d}\ln\rho_{\rm gal}/{\rm d}\ln r$. The above mentioned observed values, $\langle\beta_{\rm spec}\rangle \simeq 1.0 \pm 0.1$ and ${\rm d}\ln\rho_{\rm gal}/{\rm d}\ln r \simeq -2.4 \pm 0.2$, yield $\gamma \simeq 1.2 \pm 0.2$. We will in the following adopt $\gamma = 1.2$ as our canonical choice and also study the cases of $\gamma = 1$ and $\gamma = 1.5$ to take into account possible systematic effects that may change the amplitude of equation (eq. [4.18]).

The $T - M$ relation derived above (eq. [4.18]) is also supported by the results of recent numerical simulations at least for massive clusters. For example, Bryan & Norman (1998) have evaluated the correlation between the emission weighted temperature and the virial mass

of simulated clusters at redshifts $z = 0$, 0.5, and 1, assuming that the formation redshift $z_{\rm f}$ of each cluster is equal to the redshift $z$ at which it is viewed. Their simulations reproduce well the scaling relation (4.18) at $T \gtrsim 3$ keV in both the standard ($\Omega_0 = 1$) and open ($\Omega_0 = 0.4$) CDM models with almost the fixed best-fit value of $\gamma \simeq 1.2$, while the scatter from the best-fit relation seem to increase at lower temperatures. This agreement further supports that equation (4.18) gives a reasonable approximation to the emission weighted temperature of clusters, which is the quantity usually obtained in X-ray observations.

Strictly speaking, the $T - M$ relation of equation (4.18) should be evaluated at $z_{\rm f}$, that is conceptually different from $z$. The relation is, however, often applied *assuming* that $z$ is equal to $z_{\rm f}$ as done above in the simulations and in equation (4.4). Although the aforementioned agreements seem to support this assumption at least for massive clusters, its validity including smaller clusters should be examined carefully. In fact, the scatters found at low temperatures in the simulated $T - M$ relation by Bryan & Norman (1998) are consistent with the fact that smaller clusters have earlier formation epochs with larger dispersion (Fig. 3.4). Thus, in order to compare different assumptions regarding the cluster formation epochs, we will consider the following two cases in predicting the cluster abundances. One is to assume conventionally $z_{\rm f} = z$, evaluate equation (4.18) at an arbitrary $z$, and compute the XTF and XLF by equation (4.4). The other is to distinguish explicitly $z_{\rm f}$ from $z$, evaluate equation (4.18) at $z_{\rm f}$ for each cluster, and use equation (4.5).

In the latter approach, one can not only separate $z_{\rm f}$ from $z$ conceptually, but also incorporate quantitatively the effects of evolution of each cluster between $z_{\rm f}$ and $z$, induced, for instance, by quiescent accretion of matter. As such evolution is not yet fully resolved theoretically, we simply use the following power-law form to model it:

$$T(M, z_{\rm f}, z) = \left(\frac{1+z_{\rm f}}{1+z}\right)^s T(M, z_{\rm f}), \tag{4.19}$$

where $T(M, z_{\rm f})$ is the temperature evaluated at $z_{\rm f}$ by equation (4.18), and $s$ is the parameter introduced to control the strength of the evolution. Hydrodynamical simulations indicate that the temperature of clusters, once formed, does not change drastically after virialization and are roughly consistent with $0 \lesssim s \lesssim 1$ (e.g. Evrard 1990; Navarro et al. 1995; Eke, Navarro & Frenk 1998).

### 4.3.3 X-ray emission

The intracluster gas produces copious X-ray emission, mainly due to thermal bremsstrahlung and line emissions. These emission processes give emissivities that are proportional to the ion and electron densities, and otherwise depend only on the temperature. The total emissivity (emitted energy per unit time, unit volume and unit frequency) then becomes

$$\begin{aligned} \epsilon_\nu &= \sum_i \Lambda_\nu(i,T) n_i n_{\rm e}, \\ &= \sum_i \Lambda_\nu(i,T) f(i,T)\ n_{\rm e}^2 \end{aligned} \tag{4.20}$$

where $\Lambda_\nu(i,T)$ is the emitted energy at $T$ per the $i$-th ion, unit time, unit electron density, and unit frequency. The second line of equation (4.20) assumes that the intracluster gas is

in an ionization equilibrium and the ratios $f(i,T) \equiv n_i/n_{\rm e}$ depend only on $i$ and $T$. For the thermal bremsstrahlung (free-free) component, the emissivity can be expressed by a simple analytic form (e.g. Rybicki & Lightman 1979)

$$\epsilon_\nu^{\rm ff} = \frac{2^5\pi e^6}{3m_{\rm e}c^3}\left(\frac{2\pi}{3m_{\rm e}k_{\rm B}}\right)^{1/2} T^{-1/2}\exp\left(-\frac{h_{\rm P}\nu}{k_{\rm B}T}\right)\bar{g}(T,\nu)n_{\rm e}\sum_i Z_i^2 n_i,$$
$$= 6.84\times10^{-38}\left(\frac{2}{2-Y}\right)T^{-1/2}\exp\left(-\frac{h_{\rm P}\nu}{k_{\rm B}T}\right)\bar{g}(T,\nu)\ n_{\rm e}^2 \quad {\rm erg\,s^{-1}\,cm^{-3}\,Hz^{-1}}, \quad (4.21)$$

where $h_{\rm P}$ is the Planck constant, $Z_i$ is the charge of ion $i$, and $\bar{g}(T,\nu)$ is the velocity-averaged Gaunt factor which is roughly approximated by $\bar{g}(T,\nu) = 0.9(h_{\rm P}\nu/k_{\rm B}T)^{-0.3}$ (Evrard & Henry 1991). The second line in equation (4.21) is evaluated in CGS units, assuming that the gas is fully ionized and dominated by hydrogen and helium.

Figure 4.1 illustrates the emissivities computed by the Masai model (Masai 1984) as a function of temperature. Relative composition of heavy metals is assumed to be the same as the sun, and the total metal mass fraction is varied between $Z = 0$ and $0.8\,{\rm Z}_\odot$ ($1\,{\rm Z}_\odot = 0.017$). The contribution from metal lines to the total emissivity is shown to become important at $T \lesssim 2$ keV especially at soft energies.

The X-ray luminosity is obtained by integrating the emissivity over the total volume of a cluster:

$$L_\nu = \int \epsilon_\nu {\rm d}V = \sum_i [\Lambda_\nu(i,T)f(i,T)] \times \int n_{\rm e}^2 {\rm d}V, \quad (4.22)$$

where the second equality assumes that the intracluster gas is nearly isothermal. Since the emissivity is proportional to the square of the electron density, the X-ray luminosity is very sensitive to the spatial distribution of the intracluster gas.

### Self-similar model

In order to model the X-ray luminosity by equation (4.22), one needs to specify the density profile of the intracluster gas. It is conventional to adopt a form that fits well the observed surface brightness profile (e.g. Sarazin 1988):

$$\rho_{\rm gas}(r) = \rho_{\rm gas}^0\left[1+\left(\frac{r}{r_{\rm c}}\right)^2\right]^{-1}, \quad (4.23)$$

where $\rho_{\rm gas}^0$ is the central gas density, and $r_{\rm c}$ is the core radius. Since it is still very difficult to predict $r_{\rm c}$ theoretically for an arbitrary cluster, we adopt a simplest self-similar assumption, in which $r_{\rm c}$ varies in proportion to $r_{\rm vir}(M, z_{\rm f})$:

$$r_{\rm c} = 0.15h^{-1}{\rm Mpc}\ \times \frac{r_{\rm vir}(M,z_{\rm f})}{r_{\rm vir}(10^{15}\,{\rm M}_\odot, z_{\rm f}=0)}, \quad (4.24)$$

where the normalization is chosen to match the local observations (Abramopoulos & Ku 1983; Jones & Forman 1984; Hatsukade 1989). Now the central gas density is fixed by the relation:

$$\int_0^{r_{\rm vir}} \rho_{\rm gas}(r)\,4\pi r^2 {\rm d}r = M\left(\frac{\Omega_{\rm B}}{\Omega_0}\right). \quad (4.25)$$

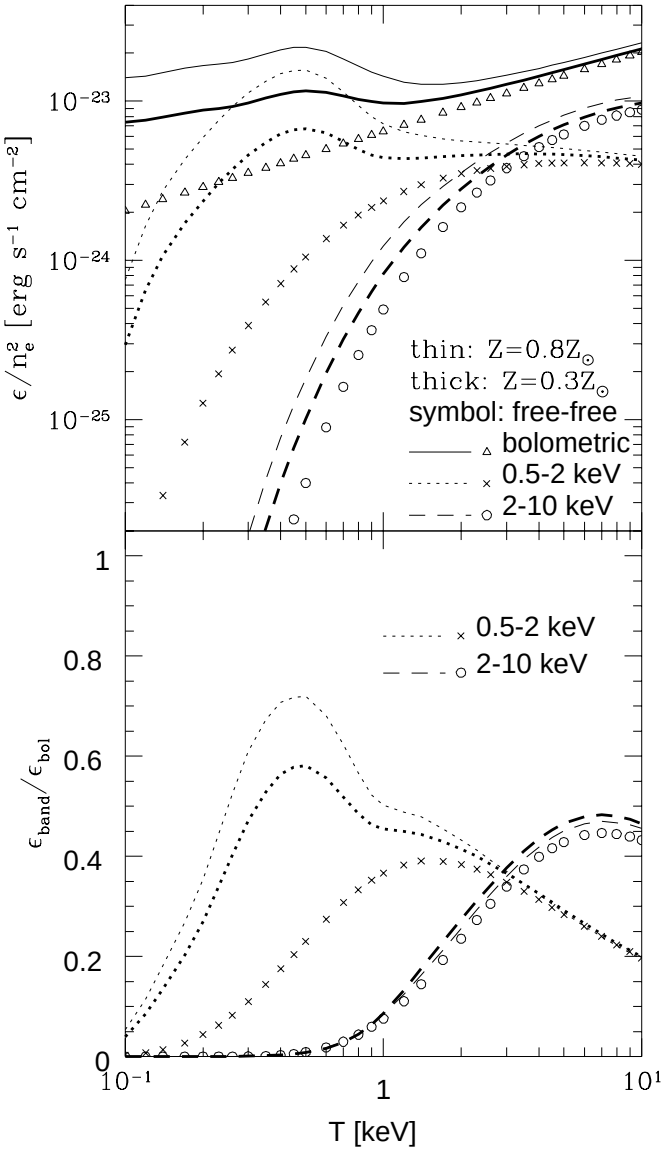


Figure 4.1: X-ray emissivities computed by the Masai model (Masai 1984): (a) bolometric (solid lines, triangles), 0.5–2 keV (dotted, crosses), and 2–10 keV (dashed, circles) emissivities; (b) ratio of the emissivity in a given band to the bolometric one. Thick lines indicate the heavy metal fraction of $Z = 0.3\,Z_\odot$, thin lines $Z = 0.8\,Z_\odot$, and symbols the free-free component with $Z = 0$. The helium mass fraction is fixed at $Y = 0.24$, and the hydrogen fraction is $X = 1 - Y - Z$.

We adopt $\Omega_{\rm B} = 0.0125h^{-2}$ to be consistent with the primordial nucleosynthesis (Walker et al. 1991).

Using equations (4.21)– (4.23), one obtains the thermal bremsstrahlung component of the bolometric luminosity as

$$L_{\rm bol}^{\rm ff} = \int_0^\infty {\rm d}\nu \int_0^{r_{\rm vir}} {\rm d}r 4\pi r^2 \epsilon_\nu^{\rm ff},$$
$$= 1.2\times 10^{45} \frac{A}{\gamma^{3/2}} \left(\frac{\Omega_{\rm B}/\Omega_0}{0.1}\right)^2 \left(\frac{\Delta_{\rm c}}{18\pi^2}\right)^{1/2} \left(\frac{T}{6{\rm keV}}\right)^2 (1+z_{\rm f})^{3/2}\ \Omega_0^{1/2} h\ \ {\rm erg\ sec^{-1}}, \quad (4.26)$$

where $A$ is a fudge factor of order unity which depends on the ratio $r_{\rm vir}/r_{\rm c}$. For the normalization of $r_{\rm c}$ adopted in equation (4.24), $A$ is equal to 0.86 in the case of $(\Omega_0, \lambda_0, h) = (1, 0, 0.5)$, and 1.1 in the case of $(\Omega_0, \lambda_0, h) = (0.1, 0, 0.7)$.

Equation (4.26) explicitly shows that the bolometric luminosity scales as $L_{\rm bol}^{\rm ff} \propto T^2$. This is, however, in conflict with the observed correlation $L_{\rm bol} \propto T^3$. Taking account of line emissions will make matters worse, because the emissivity will increase at low temperatures and the resulting $L-T$ relation will become even shallower (see also Fig. 4.2). Furthermore, at a given temperature, the bolometric luminosity predicted by equation (4.26) would evolve as $(1+z_{\rm f})^{3/2}$, apart from a weak $z_{\rm f}$ dependence of $\Delta_{\rm c}$. This suggests much stronger evolution in the amplitude of the $L_{\rm bol}-T$ relation than what is actually found in the recent observations (Henry et al. 1994; Tsuru et al. 1996; Mushotzky & Scarf 1997). Thus the above rather conventional modelling of the X-ray luminosity is not very adequate.

Since the X-ray emissivity is proportional to the square of the gas density, most of the cluster luminosity comes from the central core region. Therefore the self-similar assumption of $r_{\rm c}$ (eq. [4.24]) is thought to be largely responsible for the above discrepancy. In fact, this is currently a generic problem to which we do not yet have a clear answer. A number of numerical simulations (e.g. Navarro et al. 1995) face the same problem of predicting too shallow slope of the $L-T$ relation. It has been suggested that the discrepancy may be reconciled if the gas attains large entropy at high redshift prior to the cluster collapse by, for instance, the energy ejection from supernovae (Kaiser 1991; Evrard 1990; Evrard & Henry 1991). Such intuitive models, however, still need further investigation. Instead of relying on any uncertain theoretical ansatz, we next discuss incorporating an empirical $L-T$ relation which is determined directly from observations.

### Empirical $L-T$ relation

It has been recognized that there exists a good correlation between the observed luminosity and temperature of local clusters (e.g. Edge & Stewart 1991a,b; David et al. 1993). Recent *ASCA* experiments further suggest that this correlation is likely to remain almost unchanged at least up to $z \sim 0.5$ (Henry et al. 1994; Tsuru et al. 1996; Mushotzky & Scarf 1997). Unfortunately, a simple self-similar model for the X-ray emission discussed in the previous section fails to account for both the slope and (no) evolution found in the observations. Thus we now attempt to model the X-ray emission based on the empirical $L-T$ relation.

We model the correlation between $L_{\rm bol}$ and $T$ as follows:

$$L_{\rm bol} = L_{44} \left(\frac{T}{6\,{\rm keV}}\right)^\alpha (1+z)^\zeta\ \ 10^{44} h^{-2}\ {\rm erg\ sec^{-1}}, \quad (4.27)$$

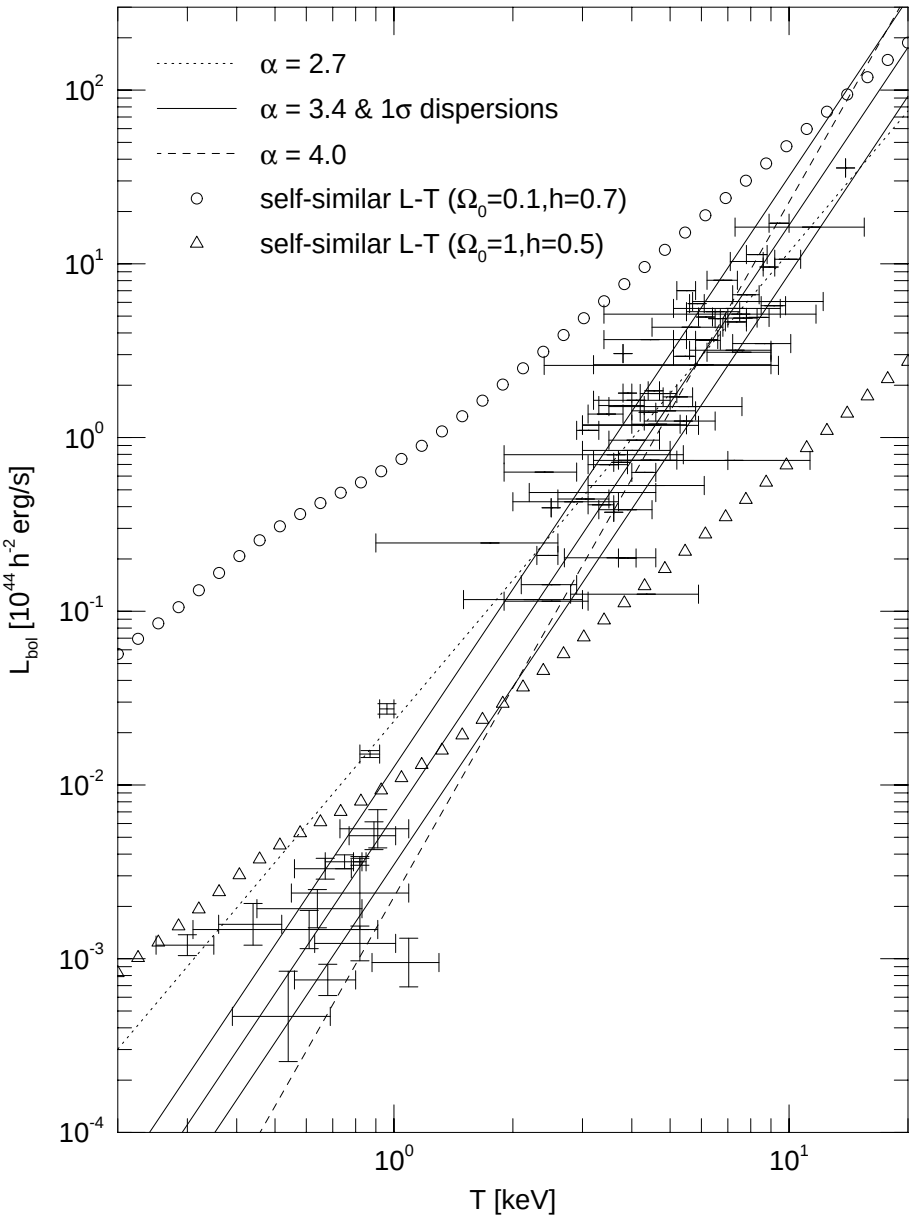


Figure 4.2: The $L-T$ relation of X-ray clusters at $z=0$. The data points at $T \gtrsim 1.5$ keV are from X-ray brightest Abell-type clusters (XBACs, Ebeling et al. 1996) while those at $T \lesssim 1$ keV are from Hickson's compact groups (HCGs, Ponman et al. 1996). Solid lines are our canonical $L-T$ relation (4.27) with $L_{44}=2.9$ and $\alpha=3.4$ (David et al. 1993) and its $1\sigma$ scatter computed from the data points. Dotted line is that with $L_{44}=2.9$ and $\alpha=2.7$, and the dashed line with $L_{44}=2.9$ and $\alpha=4$. Also plotted are the theoretical $L-T$ relations based on the self-similar assumption for $(\Omega_0,\lambda_0,h)=(1,0,0.5)$ (open triangles) and $(0.1,0,0.7)$ (open circles).

where $L_{44}$, $\alpha$ and $\zeta$ are parameters which we determine from observations. For the amplitude $L_{44}$ and the slope $\alpha$, we adopt as our canonical choice $L_{44} = 2.9$ and $\alpha = 3.4$ from the observed present-day $L - T$ correlation of David et al. (1993). We separately consider the cases of $L_{44} = 1.5$ and 5.5, and $\alpha = 2.7$ and 4 in order to take account of the scatter in the observed data. As regards $\zeta$, which specifies the redshift evolution of the $L - T$ relation, we take $\zeta = 0$ (no evolution) as canonical and also examine the cases of mild evolution $\zeta = -1$ and 1 to bracket the possible evolutionary effect.

Figure 4.2 compares our model $L - T$ relation at $z = 0$ with recent observations; the data at higher temperatures ($T \gtrsim 1.5$ keV) are of the X-ray brightest Abell-type clusters (XBACs; Ebeling et al. 1996) and the ones at lower temperatures ($T \lesssim 1$ keV) are of Hickson's compact groups (HCGs; Ponman et al. 1996). For both samples, we only plot the clusters with measured X-ray temperatures; 73 clusters from XBACs and 16 from HCGs. For XBACs, we adopt the X-ray temperatures from the compilation of David et al. (1993) and convert the 0.1-2.4 keV band fluxes of Ebeling et al. (1996) to the bolometric luminosities using the Masai model (1984).

Figure 4.2 shows that our model $L - T$ relation (eq. [4.27]) with $L_{44} = 2.9$ and $\alpha = 3.4$ is consistent with the observations of rich clusters as well as small groups over almost two orders of magnitude in temperature. The best fit value for the slope $\alpha$ of the $L - T$ relation remains almost unchanged; $\alpha = 3.5$ from the XBACs sample only, and $\alpha = 3.3$ from combined samples of XBACs and HCGs. The range over which we vary $L_{44}$, $1.5 \sim 5.6$, corresponds to the $\pm 1\sigma$ scatter of the observed data when $\alpha$ is fixed at 3.4. Figure 4.2 also shows that varying $\alpha$ from 2.7 to 4.0 while fixing $L_{44} = 2.9$ roughly covers the scatter in the observed $L - T$ for the low temperature systems.

Since we focus on the bolometric luminosity function in the present Chapter, there is no need for the band correction of the X-ray luminosity. Nevertheless, for later convenience, we also describe here our modelling of the band correction. Given the relation between $L_{\rm bol}$ and $T$, we compute the luminosity in a given energy band $[E_1,E_2]$ as

$$L_{\rm band}[T, E_1, E_2] = L_{\rm bol}(T) \times f[T, E_1, E_2], \tag{4.28}$$

where $f[T, E_1, E_2]$ is the band correction factor defined by

$$f[T, E_1, E_2] \equiv \frac{\int_{\nu(E_1)}^{\nu(E_2)} \epsilon_\nu(T) {\rm d}\nu}{\int_0^\infty \epsilon_\nu(T) {\rm d}\nu}. \tag{4.29}$$

Wherever needed, the emissivity $\epsilon_\nu(T)$ is computed by the Masai model (Masai 1984). The helium mass fraction is fixed at $Y = 0.24$, the heavy metal fraction is fiducially taken as $Z = 0.3\,{\rm Z}_\odot$, and the hydrogen fraction is $X = 1 - Y - Z$. Relative composition of heavy metals is assumed to be the same as the sun. In order to take account of the uncertainties in the observed metal abundances, we also study the cases of $Z = 0.8\,{\rm Z}_\odot$ and of only the free-free component with $Z = 0$.

It should be noted that the amplitude of the $L - T$ in the self-similar model depends sensitively on the value of $\Omega_{\rm B}/\Omega_0$, i.e., the gas mass fraction of the cluster, as clearly exhibited in Figure 4.2. The approach based on the observed $L-T$ relation (eq. [4.27]), on the contrary, is entirely independent of it.

Table 4.1: Summary of parameters for the intracluster gas properties

| Parameter | Meaning | Fiducial value | Varied range |
|---|---|---|---|
| $\gamma$ | amplitude of the $T - M$ relation | 1.2 | $1.0 \sim 1.5$ |
| $\alpha$ | index of the $L - T$ relation | 3.4 | $2.7 \sim 4.0$ |
| $L_{44}$ | amplitude of the $L - T$ relation | 2.9 | $1.5 \sim 5.6$ |
| $\zeta$ | evolution of the $L - T$ relation | 0 | $-1.0 \sim 1.0$ |
| $Z$ | metal abundance | $0.3 Z_\odot$ | $0^* \sim 0.8\,\mathrm{Z}_\odot^\dagger$ |
| $s$ | evolution of temperature | – | $0 \sim 1$ |

* free-free only
† $1\,\mathrm{Z}_\odot = 0.017$

### 4.3.4 Summary of parameters

The parameters in our models for the intracluster gas properties are summarized in Table 4.1.

One remark is that the effects of evolution between $z_\mathrm{f}$ and $z$, whenever considered explicitly, are modeled only in the temperature via the parameter $s$. Once the temperature of a cluster is given, the X-ray luminosity based on the empirical $L - T$ relation does not depend on $z_\mathrm{f}$. It is thus not necessary to model the evolution of the X-ray luminosity separately in the current framework.

## 4.4 Basic features of the cluster mass function

Before predicting the XTF and XLF, let us first consider the cluster mass function (MF). Admittedly, the direct observation of the cluster MF is not yet available since it is still very difficult to determine accurately the mass of galaxy clusters. Nevertheless, we consider it here prior to any other statistics of clusters because of the following two reasons. 1) The MF is the most fundamental quantity derived directly from the theories of structure formation, and predictions of other statistics such as the XTF and XLF are based upon it. So it is instructive to study how the predicted cluster mass function depend on cosmology. This would also help us in our later discussion to segregate the effects of cosmology and adopted cluster models on the cluster abundances. 2) It is likely that the cluster MF becomes available from future X-ray and gravitational lensing observations. It would thus be meaningful to make predictions even at this stage.

Figure 4.3 shows our predictions of the cluster MF computed from equation (3.8) at $z = 0$ for various cosmological parameters. Figure 4.3(a) is plotted varying $\Omega_0$ and $\lambda_0$ for $h = 0.7$ and $\sigma_8 = 1$ in the CDM model. Obviously, the predicted MF depends strongly on $\Omega_0$. This is primarily because the mean mass density of the universe is proportional to $\Omega_0$. The higher

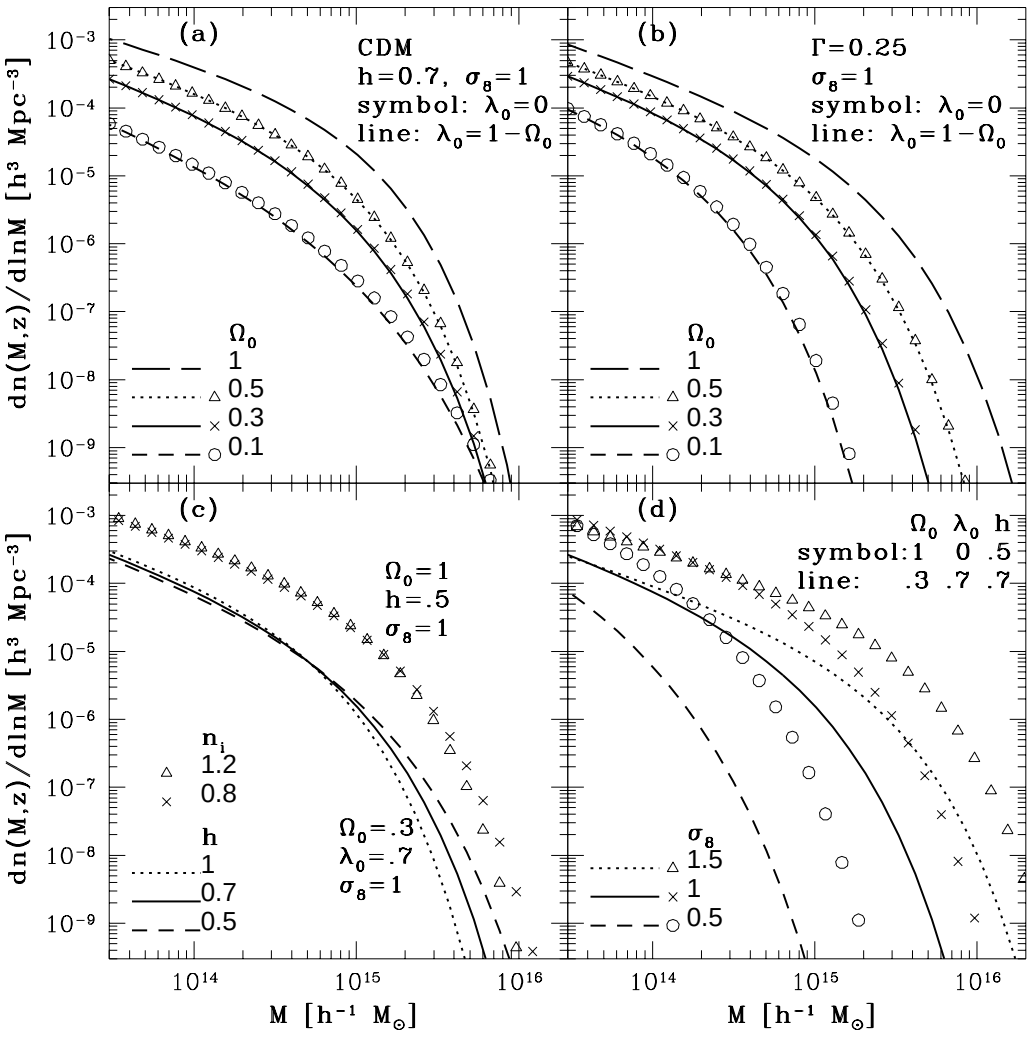


Figure 4.3: The cluster mass function in CDM models. (a) Predictions for various $\Omega_0$ and $\lambda_0$. Lines indicate flat models with $\lambda_0 = 1 - \Omega_0$, while symbols indicate open models with $\lambda_0 = 0$. $\sigma_8 = 1$ and $h = 0.7$ are fixed. (b) Same as panel (a) except for the CDM-like models with the fixed shape parameter $\Gamma = 0.25$. (c) Predictions for various $h$ (lines) and $n_{\rm i}$ (symbols), in the cases of $(\Omega_0, \lambda_0, \sigma_8) = (0.3, 0.7, 1)$ and $(\Omega_0, \lambda_0, h, \sigma_8) = (1, 0, 0.5, 1)$, respectively. (d) Predictions for various $\sigma_8$ in the cases of $(\Omega_0, \lambda_0, h) = (1, 0, 0.5)$ and $(\Omega_0, \lambda_0, h) = (0.3, 0.7, 0.7)$.

the mean mass density, the greater the number of virialized objects. By contrast, the MF at $z = 0$is insensitive to $\lambda_0$. Open models with $\lambda_0 = 0$ predict slightly larger number of clusters than the flat models with $\lambda_0 = 1 - \Omega_0$ simply because the critical density threshold $\delta_c$ is smaller (eqs [2.70], [2.71]).

The change in $\Omega_0$ and $h$ further shift the CDM spectrum of density fluctuations via the shape parameter $\Gamma \simeq \Omega_0 h$ (eq. [2.53]) and change the slope of the MF. For larger $\Omega_0 h$, the spectrum is shifted towards the smaller scale because the matter-radiation equality time becomes earlier (eq. [2.18]). This makes the mass variance steeper at the cluster scale as shown in Figure 2.2 and leads to a steeper MF. In order to segregate the effect of the change in the spectral shape, we consider in Figure 4.3(b) the "CDM-like" models, in which the shape parameter $\Gamma$ is fixed at a certain value regardless of the assumed values of $\Omega_0$ and $h$. The choice $\Gamma = 0.25$ is inferred from the observed galaxy correlation function (e.g. Peacock 1996). As expected, fixing $\Gamma = 0.25$ makes the high $\Omega_0$ predictions shallower and low $\Omega_0$ ones steeper. For intermediate values of $\Omega \sim 0.3 - 0.5$ with $h = 0.7$, the predictions remain almost unchanged.

As illustrated in Figure 4.3(c), the MF is rather insensitive to $h$ in the range $0.5 \leq h \leq 1.0$, if the MF is plotted in units of $h^3 \mathrm{Mpc}^{-3}$ against $h^{-1}\,\mathrm{M}_\odot$. The MF also depends only weakly on the primordial spectral index $n_i$ for $0.8 \leq n_i \leq 1.2$. On the contrary, Figure 4.3(d) shows that the MF changes very drastically with $\sigma_8$. Since larger $\sigma_8$ leads to earlier formation of virialized objects of given mass, the MF rises greatly above its exponential cut off but drop below the cut off. Changing $\sigma_8$ by a factor of $\sim 2$ results in the change in the number of clusters by several orders of magnitude.

## 4.5 Predictions of the X-ray temperature and luminosity functions

Having studied the basic features of the MF, we now make predictions in Figure 4.4 for the cluster XTF and XLF in three representative models, namely, the standard CDM (SCDM) model with $(\Omega_0, \lambda_0, h) = (1, 0, 0.5)$, the lambda CDM (LCDM) model with $(\Omega_0, \lambda_0, h) = (0.3, 0.7, 0.7)$, and the open CDM (OCDM) model with $(\Omega_0, \lambda_0, h) = (0.3, 0, 0.7)$. We have used the PS mass function via equation (4.4) and chosen our fiducial set of cluster gas parameters ($\alpha = 3.4$, $L_{44} = 2.9$, $\gamma = 1.2$). The effects of varying these parameters and taking account of the cluster formation epochs will be discussed later in §4.6 and §4.7, respectively. Also plotted in the figures are the boxes indicating the ranges covered by the observed local XTF of Henry & Arnaud (1991) and the BCS XLF of Ebeling et al. (1997). These boxes are estimated from the errors assigned to each data point in Henry & Arnaud (1991) and Ebeling et al. (1997) and should be regarded as rough sketches of the observed ranges. Owing to its larger sample size as mentioned in §4.1, the BCS XLF has smaller statistical errors than the XTF.

Figure 4.4 indicates that, like the MF, the XTF and XLF are very sensitive to $\sigma_8$. The SCDM model seems to reproduce roughly the observed XTF if $\sigma_8 \sim 0.5 - 0.7$, although it predicts much steeper XLF than what is actually observed for the same $\sigma_8$. We will investigate possibilities to make the slope of the SCDM XLF shallower to fit the observation in the next section. The *COBE* normalized SCDM model with $\sigma_8 \sim 1.2$ overpredicts the

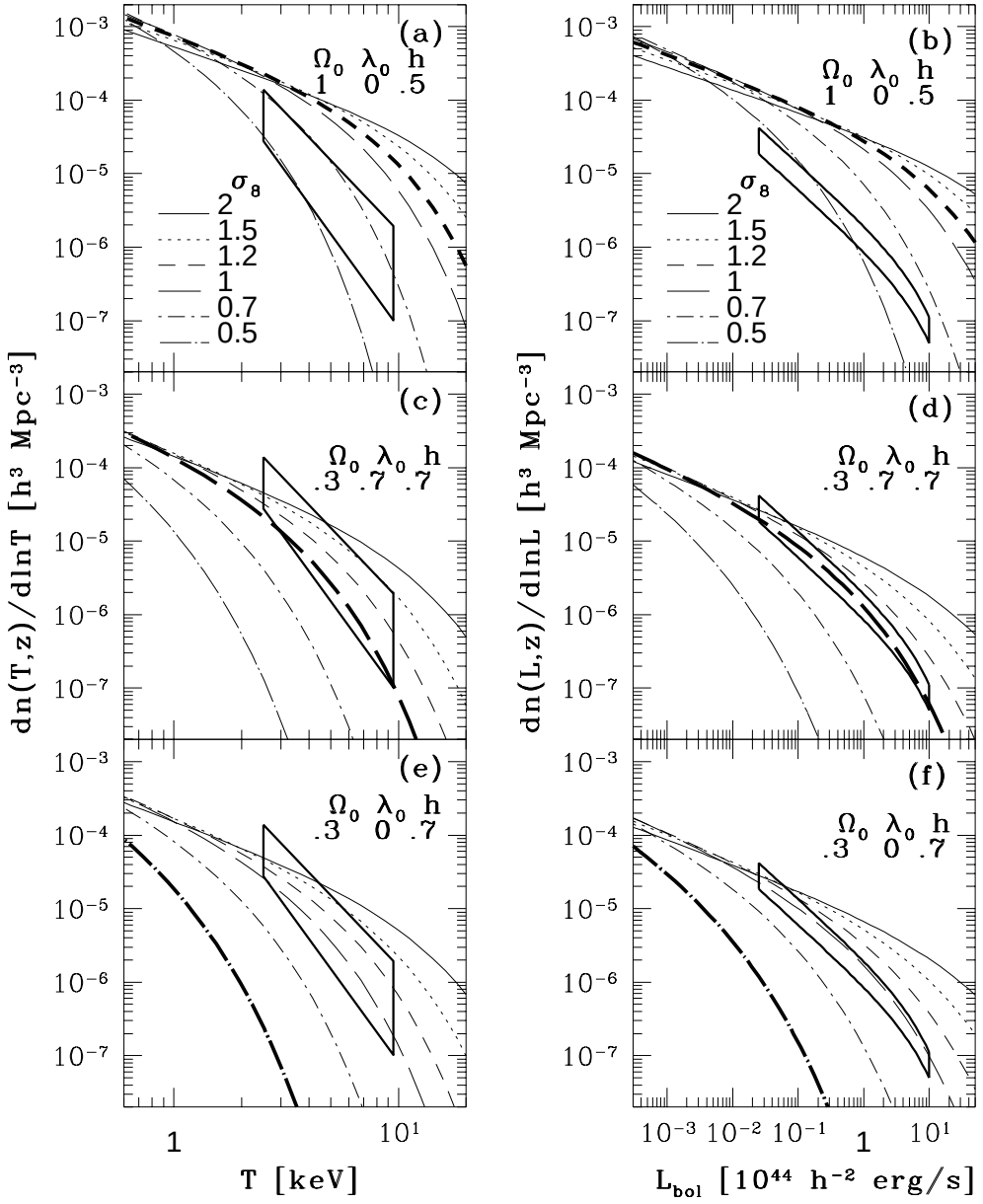


Figure 4.4: Predictions for the cluster XTF (left panels) and XLF (right panels) at $z = 0$ in (a)(b) the SCDM model with $(\Omega_0, \lambda_0, h) = (1, 0, 0.5)$, (c)(d) the LCDM model with $(0.3, 0.7, 0.7)$, and (e)(f) the OCDM model with $(0.3, 0, 0.7)$. Lines are plotted for $\sigma_8 = 2, 1.5, 1.2, 1.0, 0.7, 0.5$ as indicated in the panel. Thick lines represent *COBE* normalized values of $\sigma_8$ in each model from (Bunn & White 1997).

observed number of clusters by a wide margin. Such a discrepancy gets even worse for if $h$ is larger, because the *COBE* normalized value of $\sigma_8$ increases rapidly with $h$ ($\sigma_8 \sim 1.8$ for $h = 0.7$). On the contrary, both the XTF and XLF data are reproduced simultaneously by the *COBE* normalized LCDM model with $\sigma_8 \sim 1$. For the OCDM model considered here, $\sigma_8 \sim 1$ also fits the observed XTF and XLF well, while the *COBE* normalized $\sigma_8(\sim 0.5)$ significantly underpredicts the observations. As will be discussed in §5.3, it is possible for an open CDM model with $h \sim 0.5 - 0.8$ to reproduce simultaneously the cluster abundance and the *COBE* data if $\Omega_0 \sim 0.4 - 0.6$.

Figure 4.5 summarizes how the predicted XTF and XLF vary with $\Omega_0$, $\lambda_0$, and the spectral shape. Figures 4.5(a) and (b) tell us that the predictions depend more strongly on $\Omega_0$ than on $\lambda_0$. Compared with the MF in Figure 4.3, however, the difference between open and flat models for a given $\Omega_0$ is slightly larger in the XTF and XLF. This is due to the fact that $\Delta_c$ in equation (4.18) depends on $\Omega_0$ and $\lambda_0$ (eqs. [2.69]–[2.71]). As the cosmological constant adds a term that works as a repulsive force to the equation governing the evolution of spherical perturbation, the flat models predict slightly lower density of final collapsed objects than the open ones (Fig. 2.3). For a fixed shape parameter $\Gamma = 0.25$ (Figures 4.5 c,d), low density models with $(\Omega_0, h, \sigma_8) = (0.3, 0.7, 1)$ with/without $\lambda_0$ reproduce the XTF and XLF data, indicating that these models are capable of accounting simultaneously for both the cluster abundance and the galaxy correlation function. Finally, Figures 4.5(e) and (f) show that the predicted XTF and XLF are rather insensitive to $h$ if plotted in the units shown in the figures. The XTF and XLF also depend only weakly on the primordial spectral index $n_{\rm i}$ in the range $0.8 \leq n_{\rm i} \leq 1.2$.

## 4.6 Sensitivity to intracluster gas parameters

In addition to the cosmological parameters, our analysis incorporates several parameters that specifies the cluster gas properties (Table 4.1). The analysis of the last section was all done assuming our fiducial set for these parameters; $\alpha = 3.4$, $L_{44} = 2.9$, and $\gamma = 1.2$. Let us now study the effects of possible changes in them (effects of varying $s$ and $\zeta$ will be discussed separately in §4.7 and §4.8, respectively).

In our XTF predictions based on equation (4.4), only relevant parameter is the amplitude $\gamma$ of the $T - M$ relation. If this parameter changes, it simply shifts the XTF prediction horizontally by the same factor. As this is a trivial effect, we do not demonstrate it specifically here.

On the other hand, the XLF predictions at $z = 0$ are affected by $\alpha$, $L_{44}$, and $\gamma$. Figure 4.6 illustrates how the XLF in the LCDM model with $(\Omega_0, \lambda_0, h, \sigma_8) = (0.3, 0.7, 0.7, 1)$ behaves when each of these parameters is varied from the fiducial choice. First, changes in the index $\alpha$ of the $L - T$ relation affect the slope of the XLF; steeper $L - T$ predicts larger (smaller) $L$ for high (low) $T$ clusters and thus makes the XLF shallower. As the self-similar model yields a $L - T$ index of $\sim 2$, it predicts a very steep XLF. Second, changes in the amplitude $L_{44}$ of the $L - T$ relation shift the XLF prediction horizontally as shown in Figure 4.6 (b). Changes in $\gamma$ also move the XLF horizontally. Incidentally, varying $L_{44}$ in the range $L_{44} = 1.5 \sim 5.5$ shifts the XLF by roughly the same amount as varying $\gamma$ in the range $\gamma = 1 \sim 1.5$, owing to the fact that $(1.5/1)^{3.4} \simeq 4$ is close to $5.5/1.5 \simeq 3.7$.

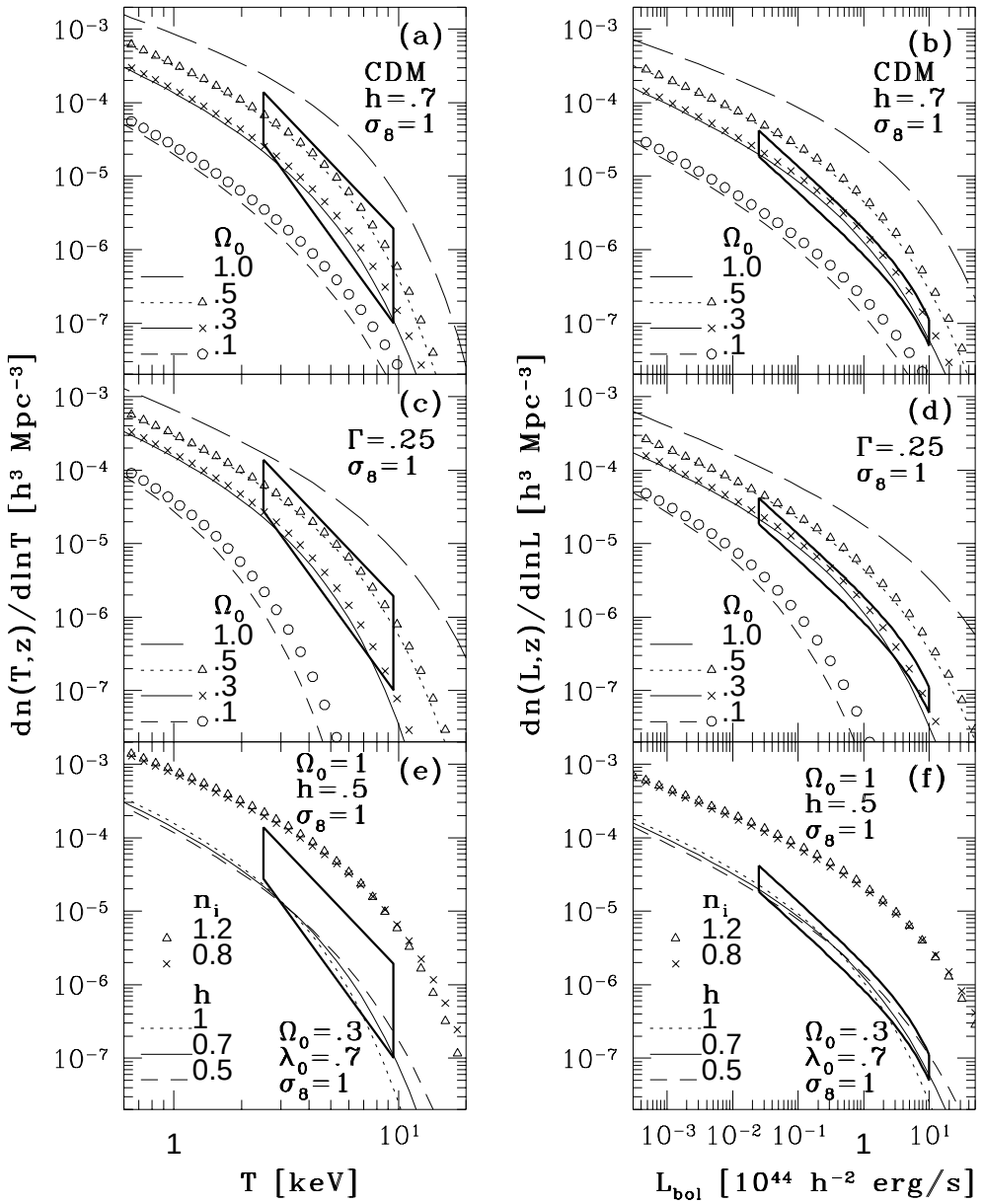


Figure 4.5: Dependence of the cluster XTF (left panels) and XLF (right panels) on $\Omega_0$, $\lambda_0$, and the spectral shape. (a)(b) Predictions for various $\Omega_0$ and $\lambda_0$ in the CDM model. $h = 0.7$ and $\sigma_8 = 1$ are fixed. Lines indicate $\lambda_0 = 1 - \Omega_0$, while symbols $\lambda_0 = 0$. (c)(d) Same as (a)(b) except for the CDM-like models with the fixed shape parameter $\Gamma = 0.25$. (e)(f) Predictions for various $h$ and $n_{\rm i}$. Lines indicate $h = 1$ (dotted), $h = 0.7$ (solid), and 0.5 (dashed) in the case of $(\Omega_0, \lambda_0, \sigma_8) = (0.3, 0.7, 1)$, while symbols $n_{\rm i} = 1.2$ (triangles) and 0.8 (crosses) for $(\Omega_0, \lambda_0, h, \sigma_8) = (1, 0, 0.5, 1)$.

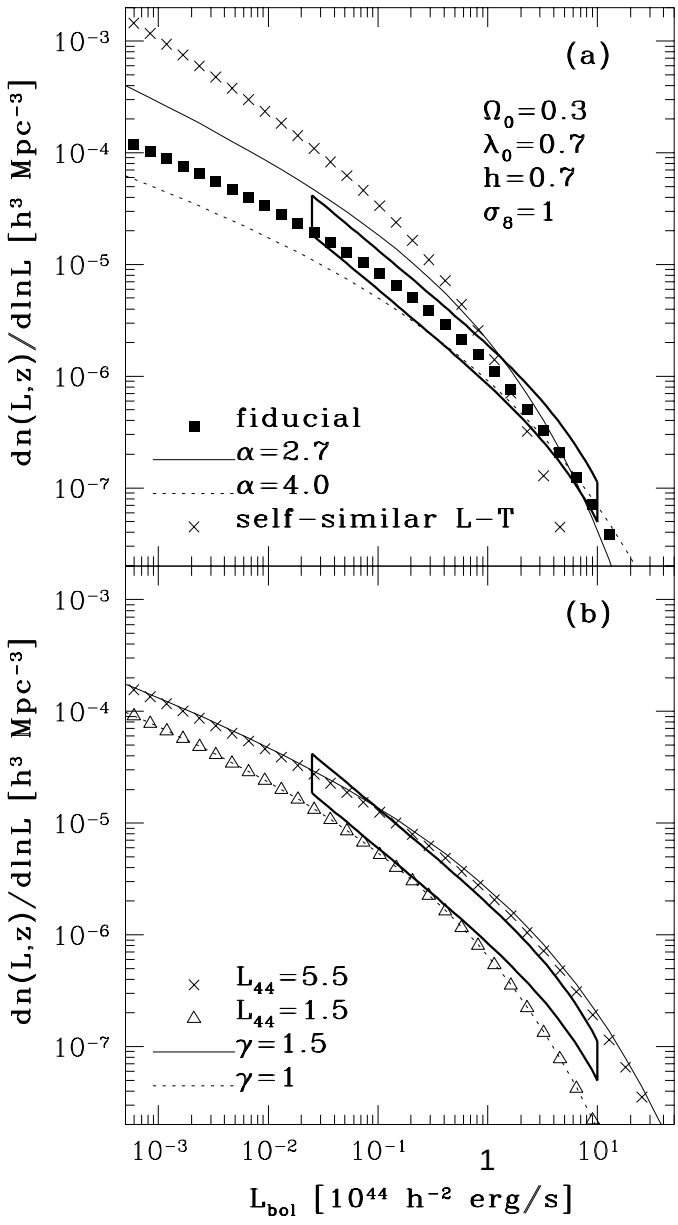


Figure 4.6: Predictions of the cluster XLF when the parameters for the intracluster gas properties are varied from the fiducial set ($\alpha = 3.4$, $L_{44} = 2.9$, $\gamma = 1.2$). Cosmological parameters are fixed at $(\Omega_0, \lambda_0, h, \sigma_8) = (0.3, 0.7, 0.7, 1)$.

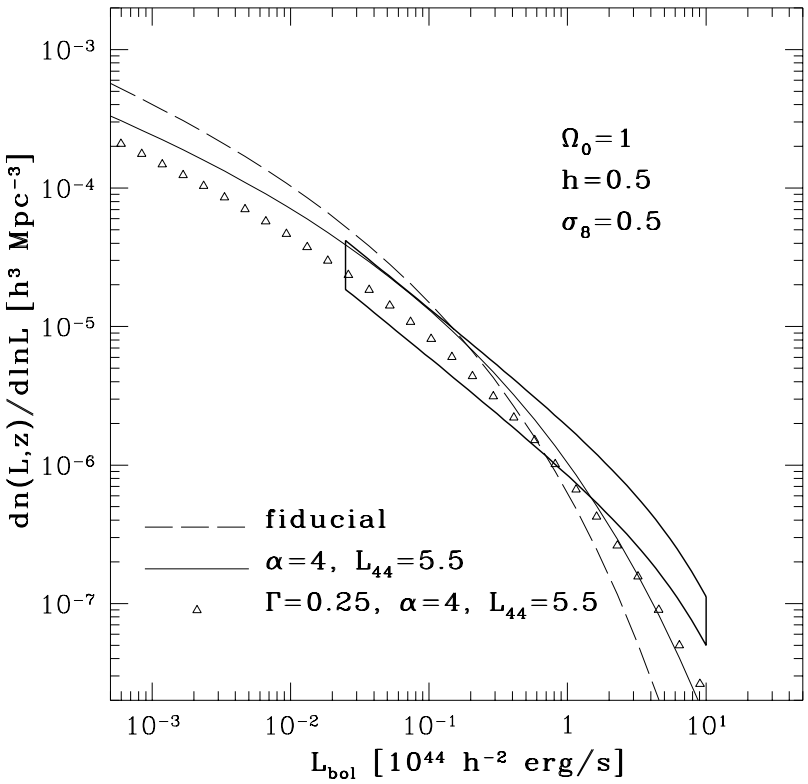


Figure 4.7: Predictions of the cluster XLF in the SCDM models with $(\Omega_0, h, \sigma_8) = (1, 0.5, 0.5)$.

The above results suggest that different choices of the cluster gas parameters can change the shape of predicted XLF to some extent. Let us now investigate whether the steep slope of the XLF predicted in the SCDM model (Figure4.4b) can be made shallower to fit the observation by changing the gas parameters. Figure 4.7 shows the predictions for the models with $(\Omega, h, \sigma_8) = (1, 0.5, 0.5)$. In order to make the XLF prediction as shallow as possible, both $\alpha = 4$ and $L_{44} = 5.5$ are adopted. In addition, as an even more extreme case, we also fix the shape parameter to $\Gamma = 0.25$. The predicted XLF's are nevertheless still steeper than the observed one. Note that the XLF becomes even steeper if $h$ is larger. Therefore, the $\Omega_0 = 1$ CDM model with $h \gtrsim 0.5$ is in conflict with the present-day XLF.

## 4.7 Effects of formation epochs

Thus far we have only used the conventional PS theory via equation (4.4) to predict the abundances of X-ray clusters. We now take into account explicitly the cluster formation epochs and the subsequent temperature evolution via equation (4.5).

Figures 4.8 and 4.9 show the number densities of clusters computed on the basis of equations (3.37) and (3.38), respectively. Compared to the conventional PS approach that assumes $z_{\rm f} = z$, these predictions have been affected by the following three factors. First is the difference between the redshift of formation $z_{\rm f}$ and the redshift of observation $z$. This effect by itself increases the temperature for a given mass of a cluster by a factor $(1+z_{\rm f})/(1+z)$ in the case of $\Omega_0 = 1$ (eq. [4.18]). Since $z_{\rm f}$ is in general larger for smaller clusters, the temperature increase becomes more significant at lower $T$ and $L$. Second is the difference in the mass of a cluster at $z_{\rm f}$ and $z$. Equations (3.37) and (3.38) assume by definition that the mass of a cluster at $z_{\rm f}$ is smaller than the mass at $z$ by up to a factor of 2, while no mass change is considered in the PS approach. This effect tends to decrease the cluster temperature evaluated at $z_{\rm f}$ by up to a factor of $(1/2)^{2/3} \simeq 0.63$. Third is the temperature evolution between $z_{\rm f}$ and $z$ modelled by equation (4.19); the temperature of a cluster evolves more rapidly for larger $s$ as $(1+z_{\rm f})^s/(1+z)^s$. In general, the $s = 0$ predictions are rather close to the PS ones because the first and second effects almost cancel each other on the cluster scales. On the other hand, the $s = 1$ predictions are larger than the PS ones by a factor of $2 \sim 10$ in the range where the observed data are available.

The quantitative details of the above three effects of course depend on the definition of the formation epochs as well as adopted cosmological parameters. Figures 4.8 and 4.9, however, imply that they are rather insensitive to the choice of the different definitions for the formation epochs between equations (3.37) and (3.38). The small difference between the two models shown in Figure 3.7 has been almost canceled out in the integration over $z_{\rm f}$ (eq. [4.5]). The main difference is that the evolutionary effect represented by the parameter $s$ becomes slightly enhanced in Figure 4.9 than in Figure 4.8. This is primarily because equation (3.38) predict higher mean $z_{\rm f}$ than equation (3.37) as indicated in Figure 3.7.

As to the dependences on cosmological parameters, the mean formation redshift of a cluster of a given mass becomes higher if $\sigma_8$ is larger or $\Omega_0$ is smaller. If $\sigma_8$ and $\Omega_0$ are both fixed, the formation redshifts are larger for the open models with $\lambda_0 = 0$ than the flat models with $\lambda_0 = 1 - \Omega_0$. Figures 4.8 and 4.9 exhibit clearly that the predicted curves are shifted rightwards by a greater amount in the models with higher formation redshifts.

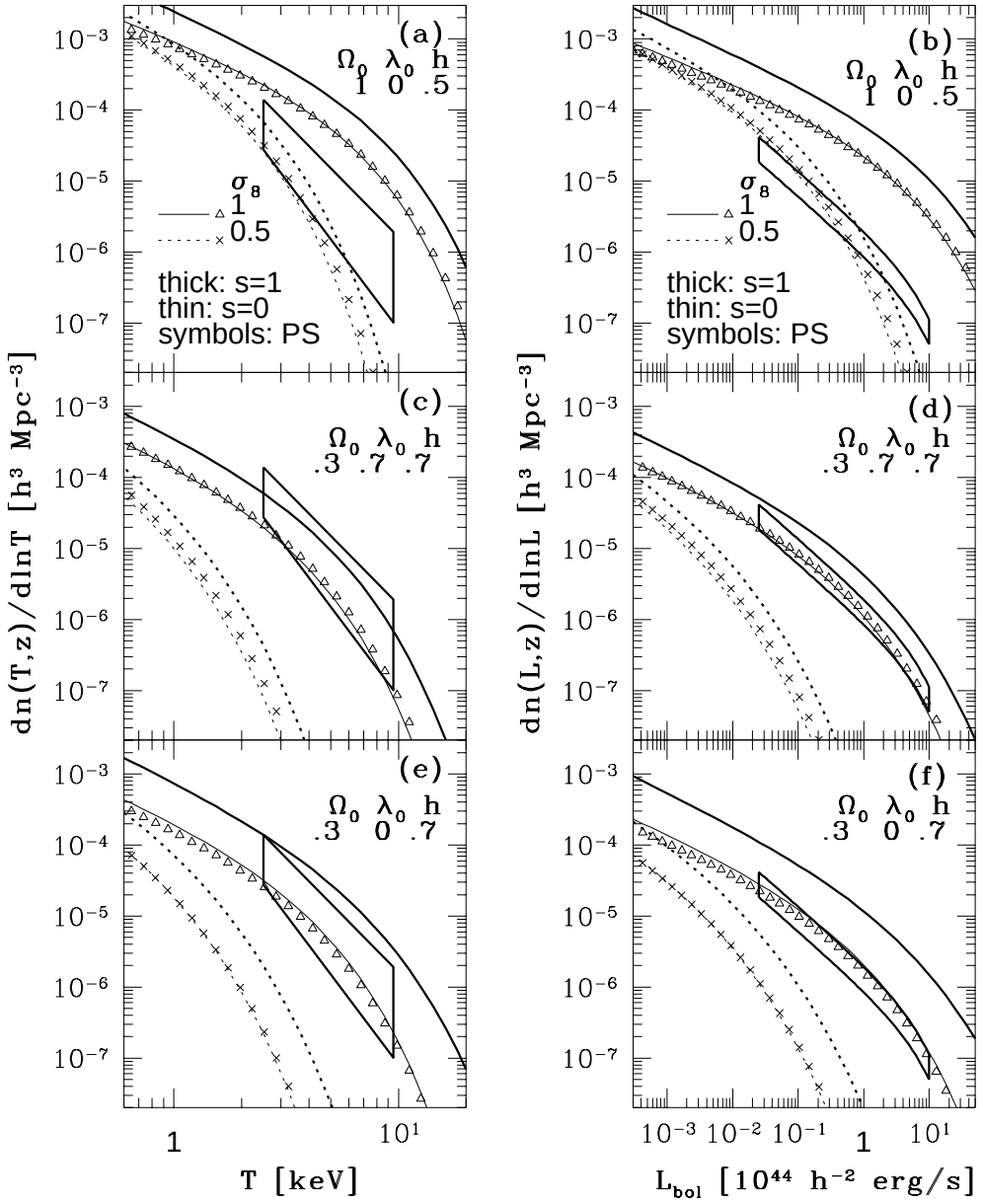


Figure 4.8: Effects of halo formation epochs on the cluster XTF (left panels) and XLF (right panels). Lines indicate the predictions based on equation (3.37) with $s = 1$ (thick) and $s = 0$ (thin), while symbols indicate those of the conventional PS model. The cases of $\sigma_8 = 1$ (solid lines, triangles) and 0.5 (dotted lines, crosses) are plotted for the same cosmological models as in Figure 4.4.

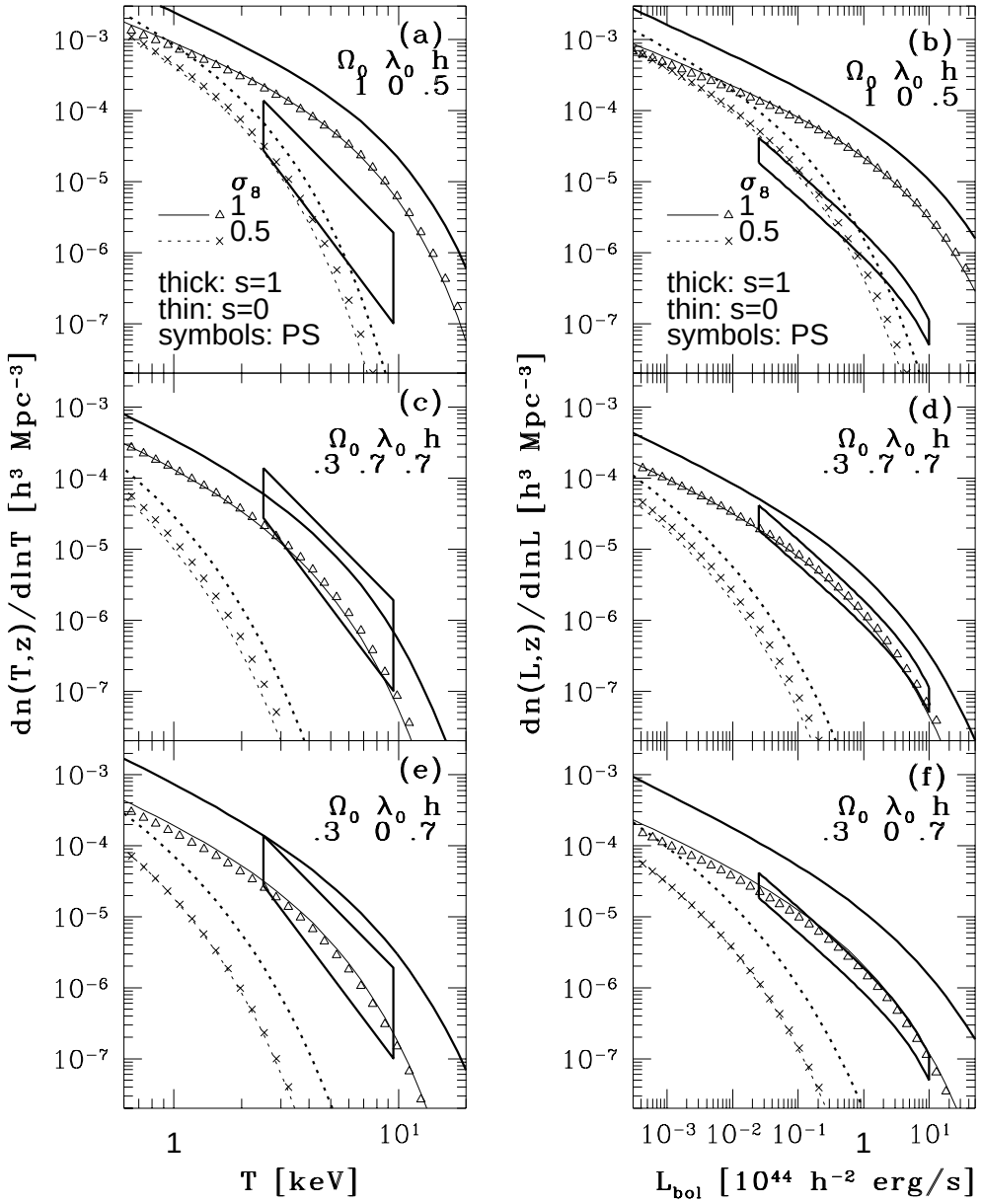


Figure 4.9: Same as Figure 4.8 except for using equation (3.38).

## 4.8 Evolution of cluster number densities

Evolution of the cluster abundances provides one of the most direct measures of the growth rate of density fluctuations. As the size of high redshift cluster samples are rapidly increasing, it is of crucial importance to discuss its implications on cosmology. We therefore make predictions for the evolution of the XTF and XLF in this section. As notified in §4.1, the XTF and XLF at high redshifts are model-dependent statistics and are not very suitable for detailed comparison between theory and observation. Here we simply aim to illustrate basic features of the evolution of the cluster abundances by means of these statistics. More quantitative analysis based on the cluster number counts will be presented in the next chapter.

Figure 4.10 show how the predicted XTF and XLF evolve between $z = 1$ and $z = 0$. We have used the PS mass function and chosen our fiducial set of intracluster gas parameters ($\alpha = 3.4$, $L_{44} = 2.9$, $\gamma = 1.2$, $\zeta = 0$). As was shown in Figure 2.1, the growth of density fluctuations continues at a rapid rate at recent times for $\Omega_0 = 1$, while it slows down significantly for smaller $\Omega_0$. As a result, the XTF and XLF evolves rapidly in the SCDM model but more slowly in the low $\Omega_0$ models. The evolution is also sensitive to the value of $\sigma_8$ and becomes faster for smaller $\sigma_8$. Among the models that roughly reproduce the present-day XTF and XLF, the SCDM model with $\sigma_8 = 0.5$ predicts that the cluster number densities drop by over an order of magnitude between $z = 0$ and 0.5 at $T \gtrsim 5$ keV and $L_{\rm bol} \gtrsim 10^{44} h^{-2}$ erg s$^{-1}$. This is already in conflict with the recent findings of the almost non-evolving XLF (constructed assuming $\Omega_0 = 1$) by *ROSAT* up to $z \sim 0.8$ (Rosati et al. 1998). On the contrary, the LCDM and OCDM models with $\sigma_8 = 1$ yield much milder evolution of the cluster number densities, which is consistent with the current observations within error bars. Note that the "non-evolving" XLF constructed assuming $\Omega_0 = 1$ would show weakly negative evolution (decreasing number towards the past) if instead a lower value of $\Omega_0$ is assumed.

The above mentioned features remain essentially identical if one takes into account the cluster formation epochs as shown in Figures 4.11 and 4.12. Only change from the PS approach is that the evolution becomes slightly more negative. This is mainly because the ratio $(1 + z_{\rm f})/(1 + z)$ approaches unity as one goes to higher redshifts and the evolution between $z_{\rm f}$ and $z$ becomes less important. Thus the discrepancy between the theory and observation for the SCDM model becomes even worse. Again, different definitions for the formation epochs between equations (3.37) and (3.38) do not affect our predictions in a significant manner.

It should be noted that the above arguments assume the invariance of the $L - T$ relation ($\zeta = 0$) at all redshifts. In order to study the effects of possible changes in this relation, we plot in Figure 4.13 the XLF at high redshifts in the cases of $\zeta = -1$ and $\zeta = 1$. If $\zeta$ is negative (positive), it makes the predicted evolution of the XLF more negative (positive) compared to the $\zeta = 0$ case. Even in a rather extreme choice of $\zeta = 1$, however, the SCDM model with $\sigma_8 = 0.5$ still predicts rather strong negative evolution of the XLF. Thus, taken together with the results shown in Figure 4.7, the SCDM model disagrees with both the shape and the evolution of the observed XLF within the range of parameters considered here.

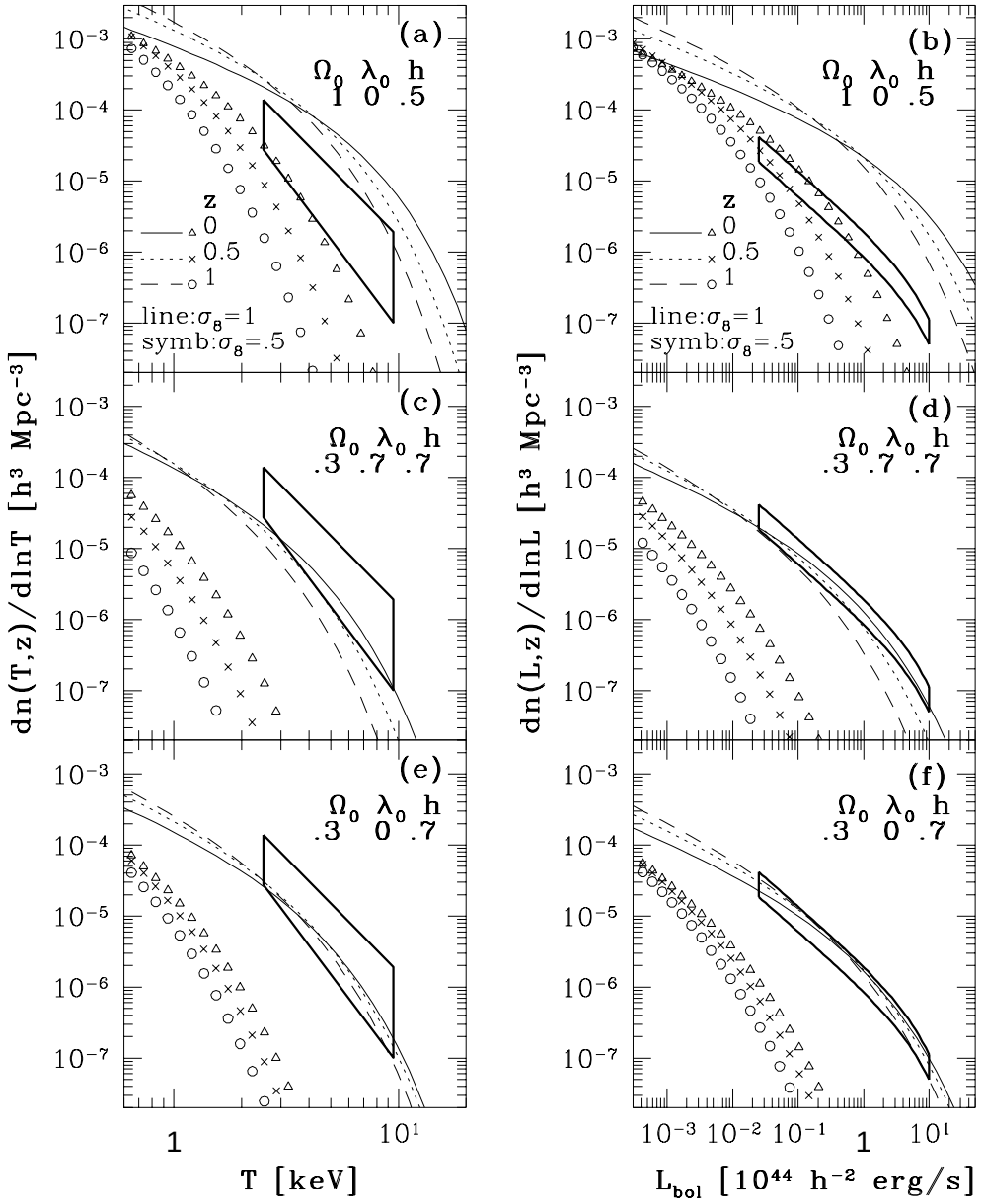


Figure 4.10: Evolution of the XTF (left panels) and XLF (right panels). The PS mass function and the fiducial set of intracluster gas parameters ($\alpha = 3.4$, $L_{44} = 2.9$, $\gamma = 1.2$, $\zeta = 0$) are used. Lines indicate $\sigma_8 = 1$ while symbols $\sigma_8 = 0.5$. The cases of $z = 0$ (solid lines, triangles) and $z = 1$ (dashed lines, circles) are plotted for the same cosmological models as in Figure 4.4.

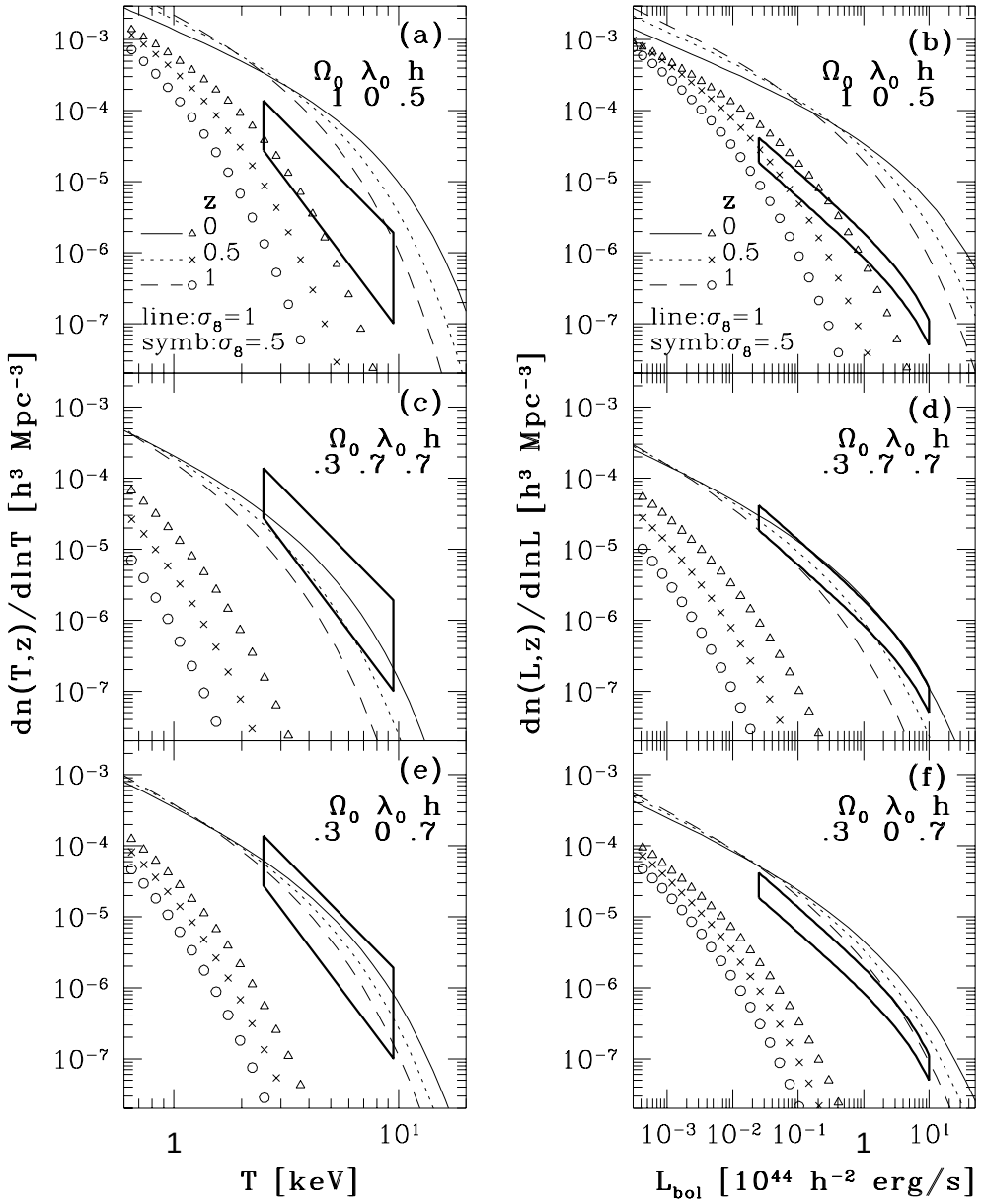


Figure 4.11: Same as Figure 4.10 except for using equation (3.37) with $s = 0.5$.

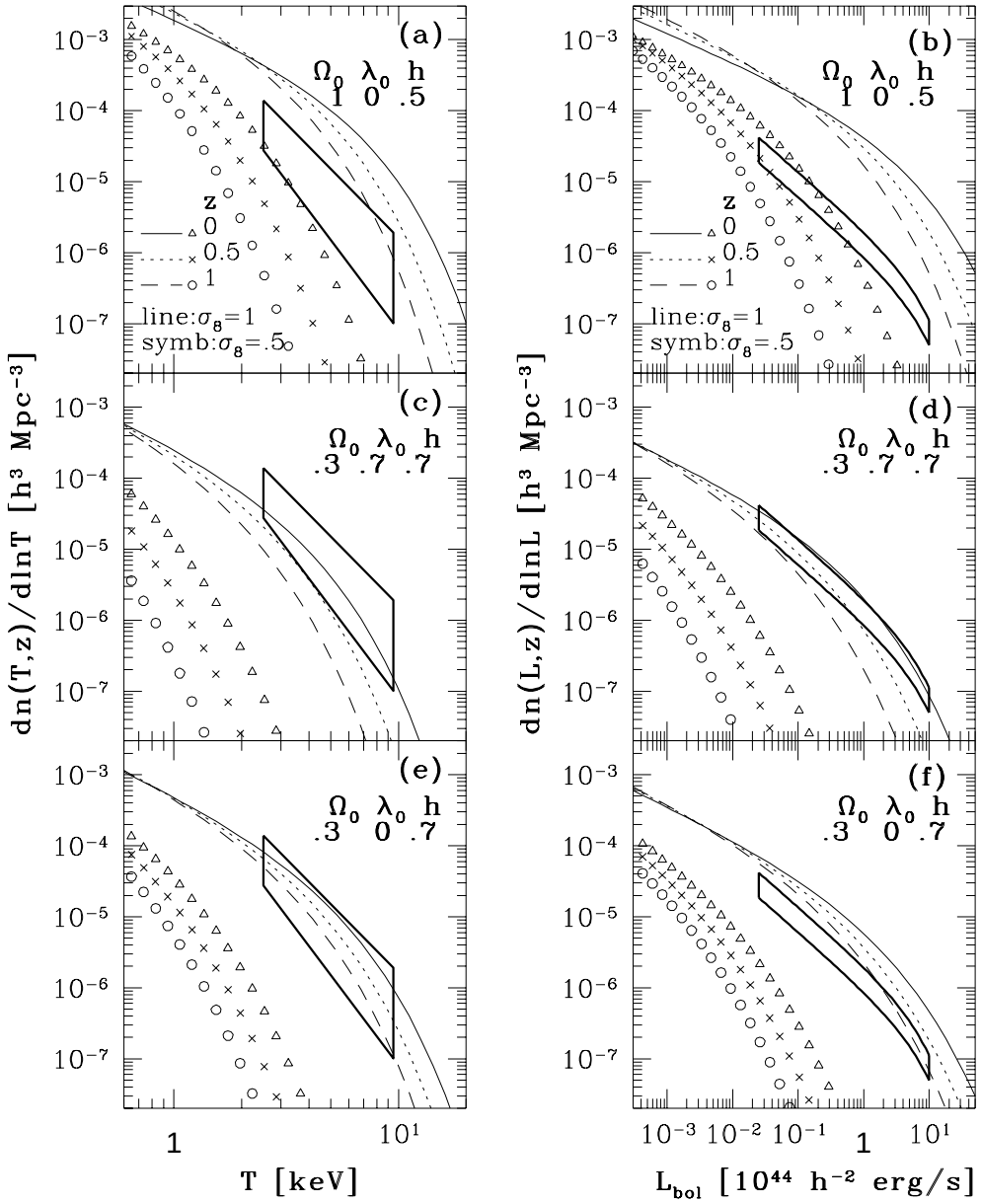


Figure 4.12: Same as Figure 4.10 except for using equation (3.38) with $s = 0.5$.

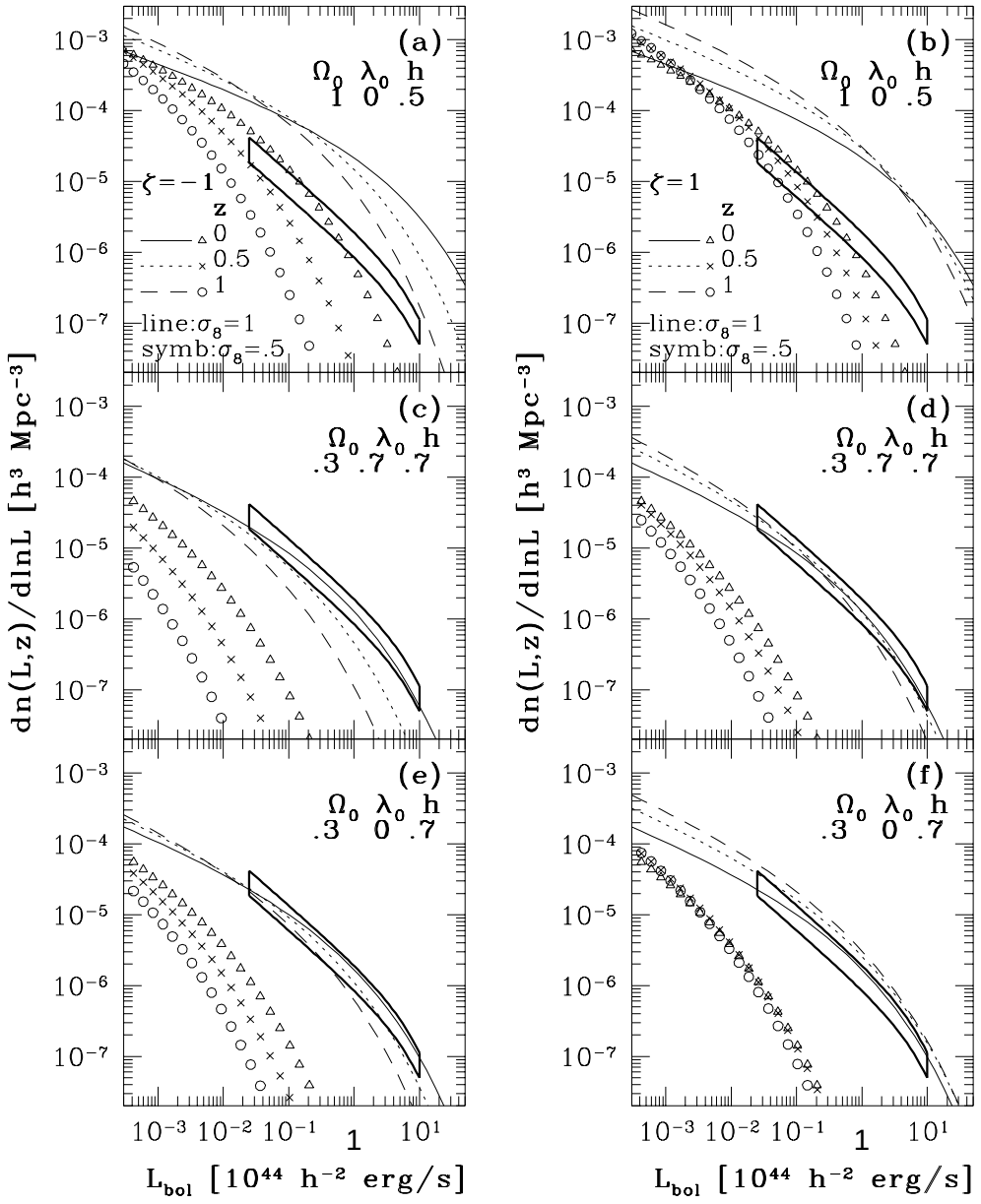


Figure 4.13: Evolution of the XLF predicted by the conventional PS theory for $\zeta = -1$ (left panels) and $\zeta = 1$ (right panels). Lines indicate $\sigma_8 = 1$ while symbols $\sigma_8 = 0.5$. The cases of $z = 0$ (solid lines, triangles) and $z = 1$ (dashed lines, circles) are plotted for the same cosmological models as in Figure 4.4.

# Chapter 5

# Cluster Number Counts in X-ray and Radio Wavebands

In this chapter, we study extensively the cosmological implications of cluster number counts. As mentioned in §1.4, cluster number counts are bias-free and model independent statistics even at high redshifts, in contrast to the X-ray temperature and luminosity functions discussed in the last chapter. Moreover, the X-ray $\log N$–$\log S$ relation recently determined by the *ROSAT* surveys (Ebeling et al. 1997; Rosati et al. 1998) comprises the largest and most complete samples of galaxy clusters to date. Cluster number counts are thus very suited for detailed comparisons between theoretical predictions and observations.

Our analysis in this chapter proceeds as follows. Firstly, we make predictions for the soft X-ray $\log N$–$\log S$ relation and derive constraints on viable cosmological models from the *ROSAT* data mentioned above. Secondly, we combine the cluster number counts at different wavelengths, in X-ray and radio bands, to deduce further constraints on cosmology from future observations. Thirdly, we predict the evolution of cluster number counts and compare them with the X-ray brightest Abell-type clusters (Ebeling et al. 1996a,b), which is the largest cluster sample available with measured redshifts. Finally, we discuss the contribution of clusters of galaxies to the X-ray background (XRB) and the submm background radiation (SBR).

## 5.1 Theoretical formulation

Given the relation between the flux $S$ and the mass $M$ of a cluster as shown in §5.1.1 and §5.1.2, we compute the number of clusters per unit solid angle with flux greater than $S$ as

$$N(>S) = \int_0^\infty \mathrm{d}z \frac{\mathrm{d}V}{\mathrm{d}z} \int_S^\infty \mathrm{d}S' \frac{\mathrm{d}n(M,z)}{\mathrm{d}M} \frac{\mathrm{d}M}{\mathrm{d}S'}\bigg|_{M=M(S',z)} \tag{5.1}$$

where $\mathrm{d}V/\mathrm{d}z$ is the comoving volume element per unit solid angle described in Appendix A.3. As in equation (4.4), the above formula uses the PS mass function given by (3.8) and assumes that clusters virialize at the epoch they are observed. By contrast, if one considers explicitly the formation epoch $z_\mathrm{f}$ of clusters, the number counts are expressed as

$$N(>S) = \int_0^\infty \mathrm{d}z \frac{\mathrm{d}V}{\mathrm{d}z} \int_S^\infty \mathrm{d}S' \int_z^\infty \mathrm{d}z_\mathrm{f} \frac{\mathrm{d}F}{\mathrm{d}M\,\mathrm{d}z_\mathrm{f}}(M, z_\mathrm{f}; z) \frac{\mathrm{d}M}{\mathrm{d}S'}\bigg|_{M=M(S',z_\mathrm{f},z)}, \tag{5.2}$$

where $\mathrm{d}F/\mathrm{d}M\mathrm{d}z_{\rm f}$ denotes the number density of clusters that form with mass $M$ at redshift $z_{\rm f}$ and survive until redshift $z$. We have shown in §4.7 and §4.8 that while equations (3.37) and (3.38) lead to very similar predictions of the cluster abundances, the differences from the conventional PS approach seem to be slightly enhanced in the latter definition. In the present Chapter, we use for definiteness equation (3.38) in order to estimate the maximal effects of formation epochs on our predictions.

### 5.1.1 X-ray flux

The X-ray flux observed in an arbitrary energy band $[E_a,E_b]$ is related to the source luminosity $L_{\rm band}$ of a cluster at $z$ in the corresponding band $[E_a(1+z),E_b(1+z)]$ as

$$S[E_a,E_b]=\frac{L_{\rm band}[E_a(1+z),E_b(1+z)]}{4\pi d_{\rm L}^2(z)}, \tag{5.3}$$

where $d_{\rm L}$ is the luminosity distance defined in Appendix A.2. The source luminosity $L_{\rm band}$ is modelled in the same way as described in §4.3. One remark is that $L_{\rm band}$ here is given in bands that vary with redshift, and this makes the band correction factor of equation (4.29) a complicated function of $T$ and $z$. In equations (5.1) and (5.2), the X-ray flux $S$ is first converted into $L_{\rm band}$ by equation (5.3), then equations (4.27)–(4.29) are solved iteratively to yield $T$, and finally equations (4.18) and (4.19) are used to recover $M$.

### 5.1.2 Radio flux — The Sunyaev-Zel'dovich effect —

The inverse-Compton scattering of the CMB photons by high temperature intracluster electrons leads to distortion of the CMB spectrum (Sunyaev & Zel'dovich 1972, SZ). This process, called the SZ effect, would provide another independent way of probing clusters of galaxies in radio wavebands.

Suppose that an isotropic radiation field is scattered by electrons that are non-relativistic and have isotropic Maxwellian distribution. Then the time evolution of the photon occupation number $\tilde{n}$ is described by the Kompaneets equation (e.g. Rephaeli 1995):

$$\begin{aligned}\frac{\partial\tilde{n}}{\partial t}&=\frac{k_{\rm B}T_\gamma}{m_{\rm e}c^2}\frac{\sigma_{\rm T}n_{\rm e}c}{x^2}\frac{\partial}{\partial x}\left[x^4\left(\frac{T}{T_\gamma}\frac{\partial\tilde{n}}{\partial x}+\tilde{n}+\tilde{n}^2\right)\right],\\&\simeq\frac{k_{\rm B}T}{m_{\rm e}c^2}\frac{\sigma_{\rm T}n_{\rm e}c}{x^2}\frac{\partial}{\partial x}\left(x^4\frac{\partial\tilde{n}}{\partial x}\right)\end{aligned} \tag{5.4}$$

where $x\equiv h_{\rm P}\nu/k_{\rm B}T_\gamma$, and the cross section $\sigma_{\rm T}$ is that of the Thomson scattering because $h_{\rm P}\nu\ll m_{\rm e}c^2$. The second line in the above equation follows since the gas temperature $T$ is much larger than the CMB photon temperature $T_\gamma$. Substituting the incident blackbody spectrum $\tilde{n}_{\rm BB}=1/(e^x-1)$ into equation (5.4) and integrating along a line-of-sight through the cluster, one obtains

$$\begin{aligned}\tilde{n}&=\tilde{n}_{\rm BB}+\Delta\tilde{n},\\\Delta\tilde{n}&=\frac{xe^x}{(e^x-1)^2}\left[x\cot\left(\frac{x}{2}\right)-4\right]y,\end{aligned} \tag{5.5}$$

where $y$ is the Compton $y$-parameter defined by

$$y \equiv \int \frac{k_{\rm B}T}{m_{\rm e}c^2}\sigma_{\rm T}n_{\rm e}{\rm c}dt. \tag{5.6}$$

Then the specific intensity at the rest frame of the cluster becomes

$$\begin{aligned} I_\nu^{\rm NR} &= \frac{2h_{\rm P}\nu^3}{c^2}\tilde{n} = i_0(1+z)^3x^3n_{\rm BB} + \Delta I_\nu^{\rm NR}, \\ \Delta I_\nu^{\rm NR} &= i_0(1+z)^3g(x)y, \end{aligned} \tag{5.7}$$

where

$$i_0 = \frac{2(k_{\rm B}T_{\gamma 0})^3}{(h_{\rm P}c)^2} = 2.29\times10^4T_{2.73}^3 \ \ {\rm mJy\ arcmin^{-2}}, \tag{5.8}$$

$$g(x) = \frac{x^4e^x}{(e^x-1)^2}\left(x\coth\frac{x}{2}-4\right), \tag{5.9}$$

where $T_{2.73} \equiv T_{\gamma 0}/(2.73\ {\rm K})$. There are several points to be noted here. First, the spectral distortion that we would observe due to the SZ effect is independent of cluster redshift since $x$ is invariant to the redshift. Second, $g(x)$ becomes negative at $x < 3.83$, i.e. at observed frequencies $> 217$ GHz or wavelength $< 0.14$ cm, while it becomes positive at $x > 3.83$, i.e. at $< 217$ GHz or $> 0.14$ cm, as illustrated in Figure 5.1. Clusters of galaxies therefore serve as negative (absorbing) sources at mm wave bands, while they become positive (emitting) sources at submm bands. Third, the redshift dependence inferred from equation (5.7), $\nu\Delta I_\nu \propto (1+z)^4$, implies that the SZ intensity is free from the cosmological dimming effect which would weaken the intensity by a factor $(1+z)^{-4}$ due to the expansion of the universe. This makes the SZ effect a particularly important and powerful probe of distant clusters. Fourth, the SZ intensity depends less strongly on the electron density as $\int n_{\rm e}{\rm d}l$ than the X-ray intensity which is proportional to $\int n_{\rm e}^2{\rm d}l$. This means that the SZ intensity is less sensitive to the gas density profile and thus easier to model. Furthermore, as a larger relative contribution to the total intensity comes from the outer regions of a cluster, one may be able to probe such low density regions by means of this effect.

From equation (5.7), the observed flux of the cluster is given by

$$\begin{aligned} S_\nu^{\rm NR} &= \frac{1}{d_{\rm L}^2(z)}\int \Delta I_\nu^{\rm NR}{\rm d}A, \\ &= 4.9g(x)\frac{2-Y}{2}\left[\frac{d_{\rm A}(z)}{cH_0^{-1}}\right]^{-2}\left(\frac{M}{10^{15}h^{-1}\,{\rm M}_\odot}\right)\left(\frac{T}{\rm keV}\right)\frac{\Omega_{\rm B}h}{\Omega_0}\ {\rm mJy}, \end{aligned} \tag{5.10}$$

where the integral $\int {\rm d}A$ is taken over the projected area of the cluster.

At $T \gtrsim 10$ keV, relativistic effects of electrons become important and the above expressions need to be modified. We thus apply the relativistic correction to the intensity change derived by Rephaeli (1995; see also Rephaeli & Yankovitch 1997):

$$\Delta I_\nu^{\rm R} = i_0\,x^3n_{\rm BB}(x)\,\eta\,y\,[\Phi(x,\eta)-1], \tag{5.11}$$

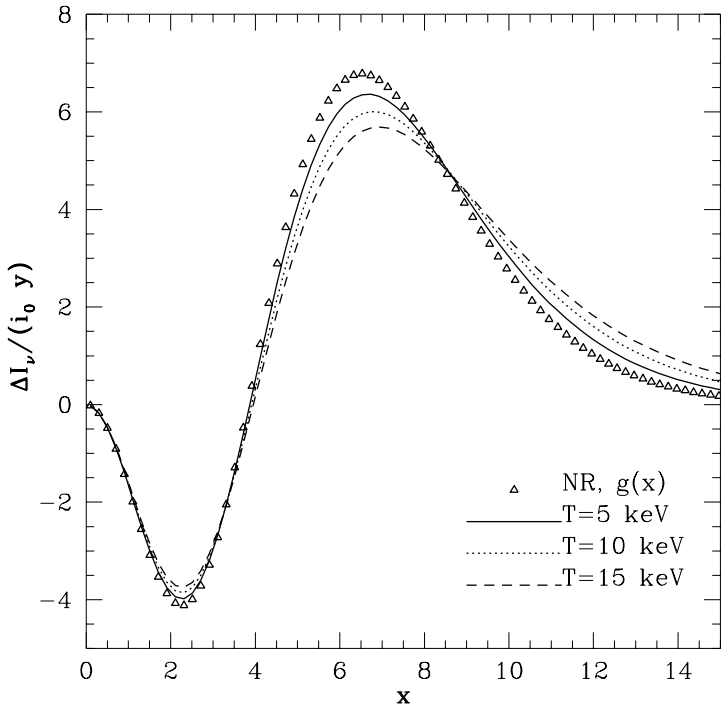


Figure 5.1: The spectral dependence of the intensity change $\Delta I_\nu/(i_0 y)$ induced by the SZ effect. The non-relativistic expression $g(x)$ is shown as triangles, whereas the relativistic results are plotted as lines for $T = 5$ keV (solid), 10 keV (dotted), and 15 keV (dashed).

with

$$\Phi(x,\eta) = \int_0^1 \left[\frac{1}{e^{xu}-1} + \frac{1}{u^3(e^{x/u}-1)}\right](e^x-1)\,\xi(u,\eta)\,\mathrm{d}u \Big/ \left[4\int_0^1 \beta_\mathrm{e}^2\gamma_\mathrm{e}^5 \mathrm{e}^{-\eta\gamma_\mathrm{e}+\eta}\,\mathrm{d}\beta_\mathrm{e}\right],$$
$$\xi(u,\eta) = \int_{(1-u)/(1+u)}^1 [(1+\beta_\mathrm{e})u - 1 + \beta_\mathrm{e}]\gamma_\mathrm{e}^3 \mathrm{e}^{-\eta\gamma_\mathrm{e}+\eta}\mathrm{d}\beta_\mathrm{e}, \tag{5.12}$$

where $\eta \equiv m_\mathrm{e}c^2/k_\mathrm{B}T$, $\beta_\mathrm{e}$ is the electron velocity in units of $c$, $\gamma_\mathrm{e} \equiv 1/\sqrt{1-\beta_\mathrm{e}^2}$, and $u \equiv \nu/\nu'$ is the fractional frequency shift of photons scattered from $\nu$ to $\nu'$. The scattering is assumed to be isotropic in the electron rest frame. Figure 5.1 plots the spectral dependence of the intensity change given by equations (5.7) and (5.11). In what follows we consider mainly the submm band at the observed wavelength of 0.85 mm [$\nu_0 = 350$ GHz, $x = 6.2$, and $g(x) = 6.7$], where the emission is fairly strong. Ground observations are also feasible at this wavelength, since it coincides with the atmospheric window of the earth.

From equations (5.7) and (5.11), the relativistic correction factor is

$$q(x,\eta) = \frac{\Delta I_\nu^\mathrm{R}}{\Delta I_\nu^\mathrm{NR}} = \frac{\eta}{Q(x)}\left[\Phi(x,\eta) - 1\right], \tag{5.13}$$

where

$$Q(x) = \frac{xe^x}{e^x-1}\left(x\coth\frac{x}{2} - 4\right). \tag{5.14}$$

At the observed wavelength of 0.85 mm, the above result leads to 4%, 11% and 16% reduction in flux at $T = 3$ keV, 8 keV and 12 keV, respectively. We find that equation (5.13) at this wavelength is well fitted by

$$q(x,\eta) = \left[1 - 0.013\left(\frac{T}{1\ \mathrm{keV}}\right)\right] \qquad \text{at 0.85 mm, i.e. } x = 6.2. \tag{5.15}$$

While clusters with $T \gtrsim 10$ keV are rare and the above correction does not make significant difference, we take it into account for completeness.

## 5.2 Predictions of the X-ray cluster log$N$–log$S$ relation

In Figure 5.2, we plot our predictions for the cluster soft X-ray log$N$–log$S$ relation in various CDM models. We use the PS mass function via equation (5.1) and adopt our canonical set of parameters ($\alpha = 3.4$, $L_{44} = 2.9$, $\zeta = 0$, $\gamma = 1.2$, $Z = 0.3\,\mathrm{Z}_\odot$) to evaluate the X-ray flux. The observed data at fainter fluxes ($S < 10^{-12}\mathrm{erg\,cm^{-2}\,s^{-1}}$) are taken from the *ROSAT* Deep Cluster Survey (RDCS, Rosati et al. 1995, 1998; Rosati & Della Ceca 1998) and those at brighter fluxes ($S > 10^{-12}\mathrm{erg\,cm^{-2}\,s^{-1}}$) are from the *ROSAT* Brightest Cluster Sample (BCS, Ebeling et al. 1997, 1998). In the analysis below, we use the RDCS data of Rosati et al. (1998) including the systematic errors according to Rosati (private communication) in addition to the statistical errors. The systematics come from the incompleteness in the optical identification of clusters at faint flux levels and from an uncertainty in the flux determination on the basis of the wavelet analysis adopted in the RDCS. The former has typically increased the upper error bar of $N(>S)$ by +15% at a given flux in the range of $2 \lesssim S \lesssim 3 \times 10^{-14}\mathrm{erg\,cm^{-2}\,s^{-1}}$,

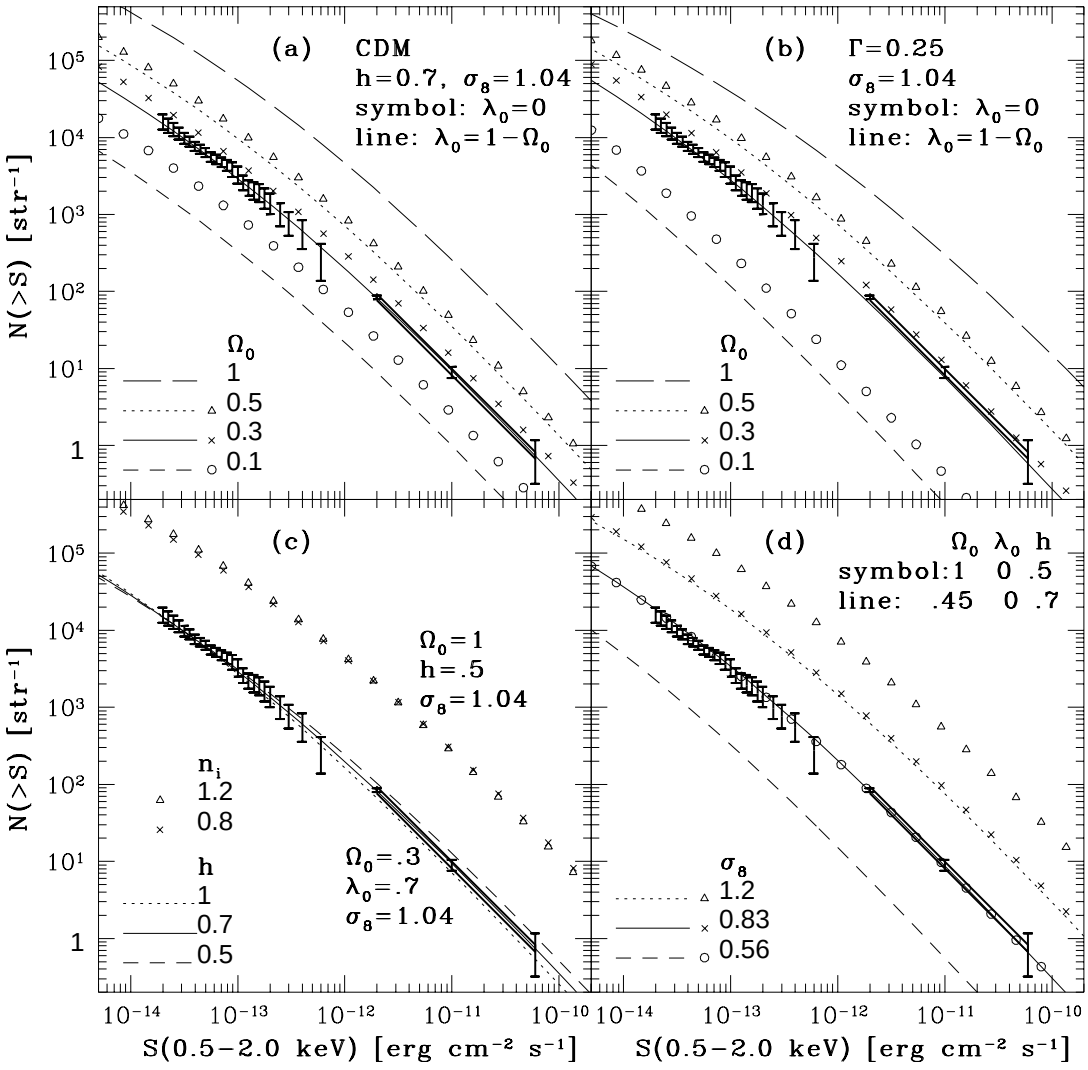


Figure 5.2: The cluster soft X-ray log$N$–log$S$ relation in CDM models. (a) Predictions for various $\Omega_0$ and $\lambda_0$. Lines indicate flat models with $\lambda_0 = 1 - \Omega_0$, while symbols open models with $\lambda_0 = 0$. $\sigma_8 = 1.04$ and $h = 0.7$ are fixed. (b) Same as panel (a) except for the CDM-like models with the fixed shape parameter $\Gamma = 0.25$. (c) Predictions for various $h$ (lines) and $n_{\rm i}$ (symbols), in the cases of $(\Omega_0, \lambda_0, \sigma_8) = (0.3, 0.7, 1)$ and $(\Omega_0, \lambda_0, h, \sigma_8) = (1, 0, 0.5, 1)$, respectively. (d) Predictions for various $\sigma_8$ in the cases of $(\Omega_0, \lambda_0, h) = (1, 0, 0.5)$ and $(\Omega_0, \lambda_0, h) = (0.45, 0, 0.7)$. Data points at $S \lesssim 10^{-12}$ erg cm$^{-2}$ s$^{-1}$are from the RDCS, including both statistical and systematic errors as mentioned in the main text. The error box at $S \gtrsim 2 \times 10^{-12}$ represents a power-law fitted region from the BCS, for which Poisson errors are also plotted at $S = 2 \times 10^{-12}$, $1 \times 10^{-11}$ and $6 \times 10^{-11}$erg cm$^{-2}$ s$^{-1}$.

while the latter has changed $S$ typically by $+8\%$ which is also included in the error of $N(>S)$. The error box at $S > 10^{-12}$ erg cm$^{-2}$ s$^{-1}$is drawn from the best-fit power-law representation of the BCS data (Ebeling, private communication). Since this error box simply represents the fitting errors, we have assigned to the data $\pm 1\sigma$ Poisson errors estimated from the real number of clusters within the survey area of the BCS (4.136 steradian). These errors will be used in the statistical analysis in §5.3.

The predicted log$N$–log$S$ relations are affected by a number of factors determined by cosmology, such as the shape of the local XLF, evolution of the XLF at high redshift, and the volume per unit solid angle. If $\Omega_0$ is larger, for instance, the local XLF becomes steeper and more abundant (§4.5), while the number of clusters decreases more rapidly at high redshift (§4.8) and the volume per unit solid angle is smaller(§A.3). Alternatively, if the models with $\lambda_0 = 1 - \Omega_0$ and $\lambda_0 = 0$ are compared for a fixed $\Omega_0$, the former predicts more rapid decrease of the cluster number at high redshift but larger volume per unit solid angle. These effects all combine to produce the features presented in Figure 5.2. As a result, the predicted log$N$–log$S$ relations are shown to depend most strongly on $\Omega_0$ and $\sigma_8$, mildly on $\lambda_0$, and only weakly on $h$ and $n_{\rm i}$.

Figures 5.2(a) and (d) indicate that the *COBE* normalized CDM models with $(\Omega_0, \lambda_0, h, \sigma_8)$ $= (0.3, 0.7, 0.7, 1.04)$ and $(0.45, 0, 0.7, 0.83)$ reproduce remarkably well the observed log$N$–log$S$ over almost four orders of magnitude in flux, provided that we adopt the canonical set of intracluster gas parameters. In the case of the standard CDM model with $(\Omega_0, \lambda_0, h) = (1, 0, 0.5)$, however, the *COBE* normalization ($\sigma_8 = 1.2$) significantly overproduces the number of clusters. This discrepancy becomes even worse for a larger value of $h$ as the *COBE* normalized $\sigma_8$ becomes greater. Therefore the standard CDM model is compatible with the cluster number counts only if $\sigma_8 = 0.56$, a factor of 2 smaller than the COBE normalization.

Figure 5.3 exhibits how the predicted log$N$–log$S$ depends on different choices of parameters for the gas properties. We also plot the results taking account of the formation epochs via equation (5.2), instead of the standard PS theory. Figures 5.3 (a) and 5.3 (b) indicate that varying $\alpha$ or $\zeta$ mainly changes the slope of the predicted log$N$–log$S$, while different values of $L_{44}$, $\gamma$ or $s$ affect the amplitude as well and may shift the log$N$–log$S$ predictions by a factor up to 5. Our predictions are shown to be insensitive to the assumed metallicity. This is because the metallicity only affect our predictions through the band correction factor (eq. [4.29]), and this factor depends only weakly on the metallicity at $T \gtrsim 1$ keV as illustrated in Figure 4.1(b).

## 5.3 Constraints on the CDM models

Having studied the features of the predicted log$N$–log$S$ relation, we now derive the constraints on the cosmological parameters. As shown in the previous section, the predictions are the most sensitive to $\Omega_0$ and $\sigma_8$, while they depend to a much lesser extent on $h$, $\lambda_0$, or any other parameters for the intracluster gas properties. In other words, one can constrain $\Omega_0$ and $\sigma_8$ rather independently of other parameters by comparing the predictions and observations.

Figure 5.4 summarizes the constraints on $\sigma_8$ and $\Omega_0$ from the cluster log$N$–log$S$, XTF and *COBE* 4 year results (Bunn & White 1997) in CDM universes with $h = 0.7$ and our

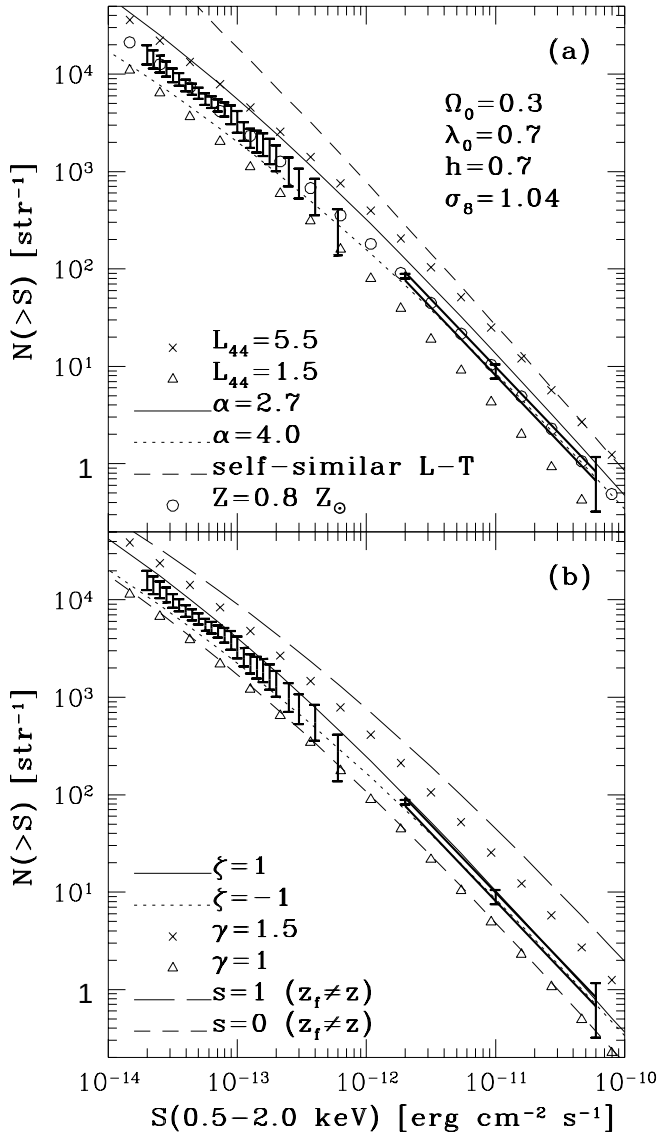


Figure 5.3: Predictions for the log$N$–log$S$ of X-ray clusters when the parameters of the gas properties are varied from the fiducial set ($\alpha = 3.4$, $L_{44} = 2.9$, $\zeta = 0$, $\gamma = 1.2$, $Z = 0.3\,Z_\odot$). We also plot the results taking account of the cluster formation epochs via equation (5.2) with $s = 0$ and 1, instead of the standard PS theory.

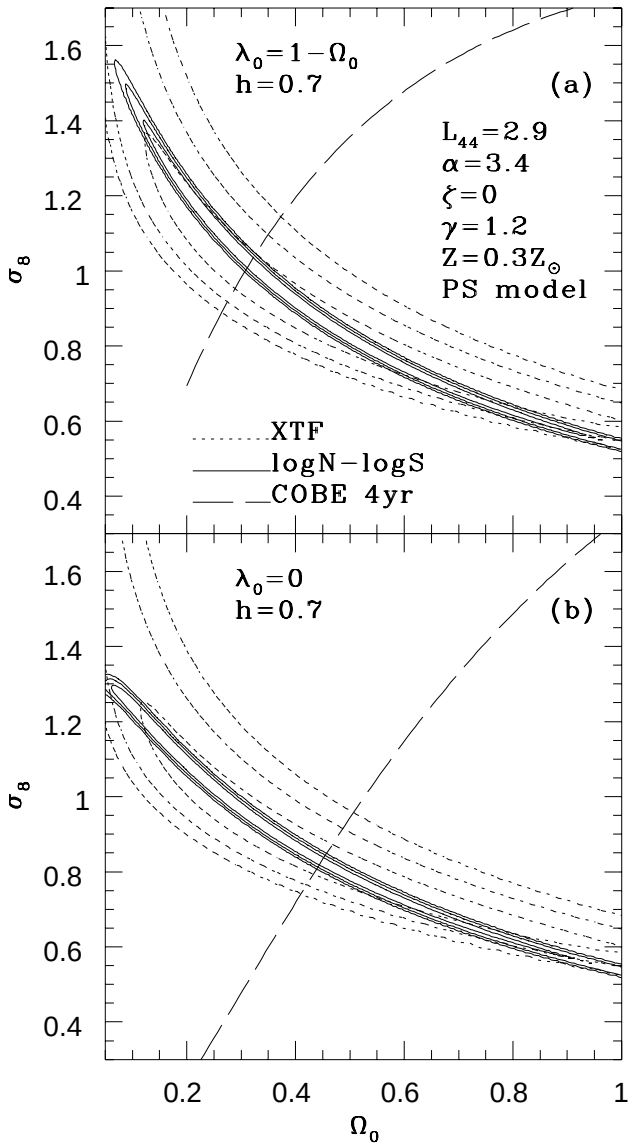


Figure 5.4: Limits on $\Omega_0$ and $\sigma_8$ in CDM models ($h = 0.7$, $n_{\rm i} = 1$) with (a) $\lambda_0 = 1 - \Omega_0$, and (b) $\lambda_0 = 0$. Constraints from cluster log$N$–log$S$ (solid) and XTF (dotted) are plotted as contours at $1\sigma(68\%)$, $2\sigma(95\%)$ and $3\sigma(99.7\%)$ confidence levels. The fiducial set of gas parameters ($\alpha = 3.4$, $L_{44} = 2.9$, $\zeta = 0$, $\gamma = 1.2$, $Z = 0.3\,{\rm Z}_\odot$) and the PS mass function are adopted. Dashed lines indicate the *COBE* 4 year results computed from the fitting formulae at $0.2 < \Omega_0 \leq 1$ by Bunn & White (1997).

standard cluster model ($\alpha = 3.4$, $L_{44} = 2.9$, $\zeta = 0$, $\gamma = 1.2$, $Z = 0.3\,Z_\odot$). We perform a $\chi^2$ test of the log$N$–log$S$ using the six data points; at $S$[0.5–2 keV] $= 4 \times 10^{-14}$, $1.2 \times 10^{-13}$ and $3 \times 10^{-13}$erg cm$^{-2}$ s$^{-1}$from the RDCS (Rosati et al. 1998), $2 \times 10^{-12}$, $1 \times 10^{-11}$ and $6 \times 10^{-11}$erg cm$^{-2}$ s$^{-1}$from the BCS (Ebeling et al. 1997, 1998) with appropriate statistical and systematic errors as discussed in the previous section. Strictly speaking, each data point of the cumulative number counts discussed here is not independent, but we treat all the above data points as independent. Since we have selected the six data points where the cluster numbers are different by a factor of $3 \sim 10$ from their neighboring points, we expect that this is not a bad approximation. In fact, we found that the constraints on $\Omega_0$ and $\sigma_8$ are essentially determined by the two data points; at $2 \times 10^{-12}$erg cm$^{-2}$ s$^{-1}$from the BCS survey and at $4 \times 10^{-14}$erg cm$^{-2}$ s$^{-1}$from the RDCS. We have in addition repeated the same analysis using twelve data points at different fluxes and made sure that the resulting constraints are insensitive to the choice. For comparison, we perform a $\chi^2$ test using the XTF data points and associated errors at $T = 3$, 4.2, and 6.2 keV from Eke et al. (1996) who reanalysed the original data of Henry & Arnaud (1991).

Figure 5.4 indicates that constraints from the cluster log$N$–log$S$ relation are consistent with, but stronger than, those from the XTF, because the observed log$N$–log$S$ has smaller error bars than the XTF and covers wider dynamic range. Our $1\sigma$ (68%) confidence limits from the cluster log$N$–log$S$ relation are well fitted by

$$\sigma_8 = (0.54 \pm 0.02) \times \begin{cases} \Omega_0^{-0.35-0.82\Omega_0+0.55\Omega_0^2} & (\lambda_0 = 1 - \Omega_0), \\ \Omega_0^{-0.28-0.91\Omega_0+0.68\Omega_0^2} & (\lambda_0 = 0), \end{cases} \tag{5.16}$$

where the quoted errors include only those due to the observed log$N$–log$S$ relation. As will be discussed below, the systematic uncertainty of the above fit due to the theoretical modelling of cluster luminosity is estimated to be $\sim 15\%$. The *COBE* normalized $\Omega_0 = 1$ model is inconsistent with the cluster number counts at more than $3\sigma$ (99.7%) confidence. Observed cluster abundances and *COBE* normalization are simultaneously accounted for by the CDM model with $(\Omega_0, \lambda_0, \sigma_8) \simeq (0.3, 0.7, 1)$ and $(0.45, 0, 0.8)$ in the case of $h = 0.7$.

Figures 5.5 and 5.6 exhibit the robustness of the above results against our model assumptions in the $\lambda_0 = 1 - \Omega_0$ and $\lambda = 0$ cases, respectively. Panel (a) of these figures indicates that varying $L_{44}$ from our canonical value 2.9 to 1.5 (5.5) will systematically increase (decrease) the best-fit $\sigma_8$ value for a given $\Omega_0$ by about 15%. The range of $L_{44}$ considered here roughly corresponds to the $\pm 1\sigma$ scatter in the observed $L - T$ relation (see also Fig.4.2). Although this scatter may be partly due to the observational uncertainties in determining the temperature, we conservatively interpret it as an intrinsic scatter in the $L - T$ relation which results in the systematic error for the best-fit $\sigma_8$ value by $\sim 15\%$.

Panels (b)–(f) of Figures 5.5 and 5.6 show the systematic effects due to the other parameters for the intracluster gas properties. The best-fit $\Omega_0 - \sigma_8$ relation (eq. [5.16]) is shown to be rather robust against $\alpha$ and $\zeta$ over the ranges considered here; the changes in these parameters merely move the contours along the best-fit relation. This is because $\alpha$ and $\zeta$ mainly affect the slope of the predicted log$N$–log$S$ (Fig.5.3) and such changes are compensated by altering the CDM fluctuation spectrum along with $\Omega_0$ and $\sigma_8$. In addition, our results are shown to be very insensitive to the assumed metallicity $Z$. On the other hand, allowing for the changes of $1 < \gamma < 1.5$ and $0 < s < 1$, the best-fit $\sigma_8$ value shifts in a comparable

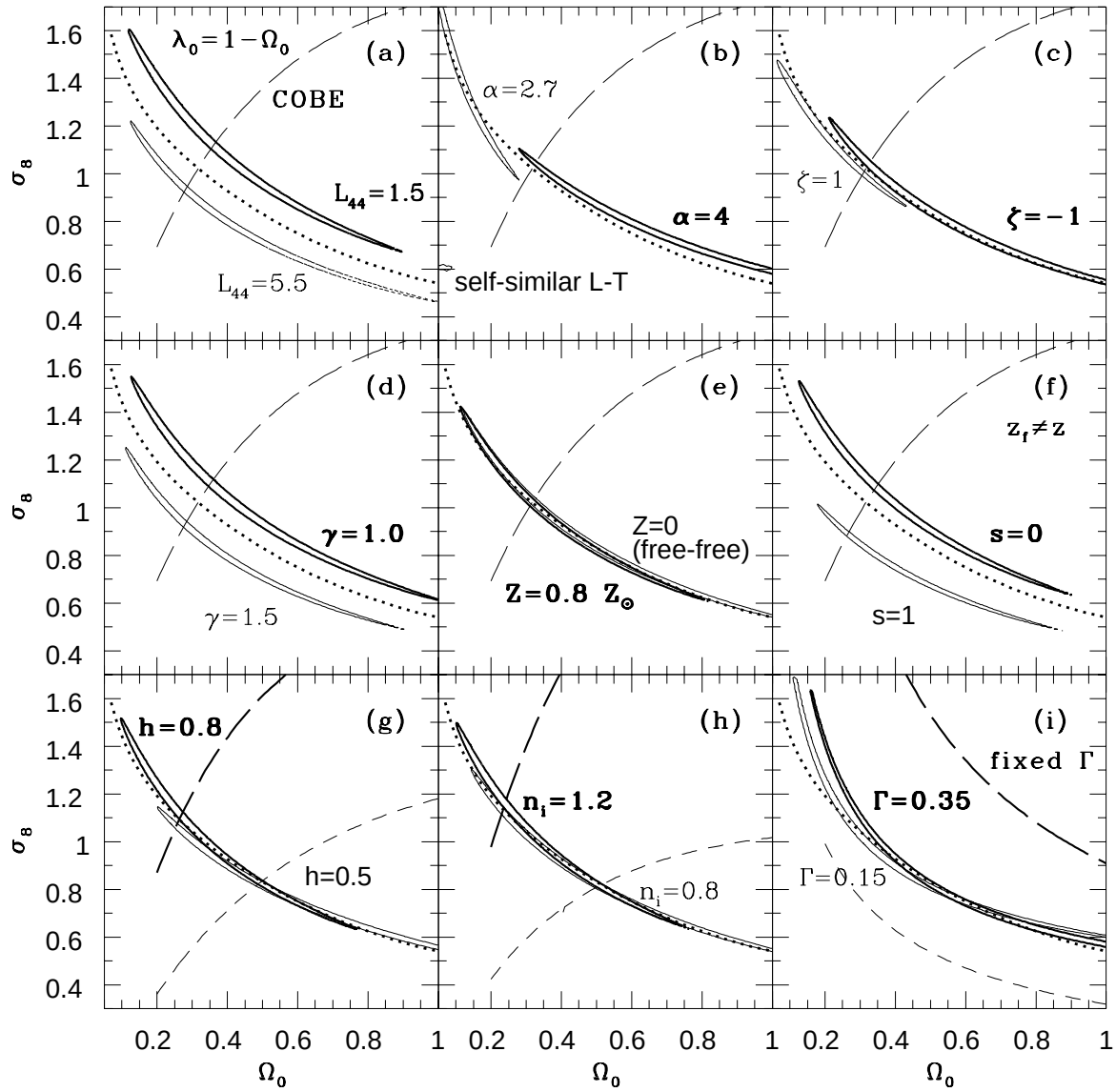


Figure 5.5: Systematic effects on the $\Omega_0 - \sigma_8$ constraints in the case of $\lambda_0 = 1 - \Omega_0$. The $1\sigma(68\%)$ confidence contours from the cluster $\log N$–$\log S$ are plotted for different (a) $L_{44}$, (b) $\alpha$, (c) $\zeta$, (d) $\gamma$, (e) $Z$, (f) $s$ ($z_{\rm f} \neq z$), (g) $h$, (h) $n_{\rm i}$, and (i) $\Gamma$ (fixed $\Gamma$). Except for the parameter varied in each panel, we adopt the PS mass function, the CDM spectrum with $n_{\rm i} = 1$, $h = 0.7$, $\Gamma \simeq \Omega_0 h$ (eq. [2.53]), and canonical cluster parameters ($L_{44} = 2.9$, $\alpha = 3.4$, $\zeta = 0$, $\gamma = 1.2$, $Z = 0.3\,{\rm Z}_\odot$). Dotted line represents the best-fit relation for the set of parameters mentioned above (eq. [5.16]). Dashed lines indicate the *COBE* 4 year results from Bunn & White (1997).

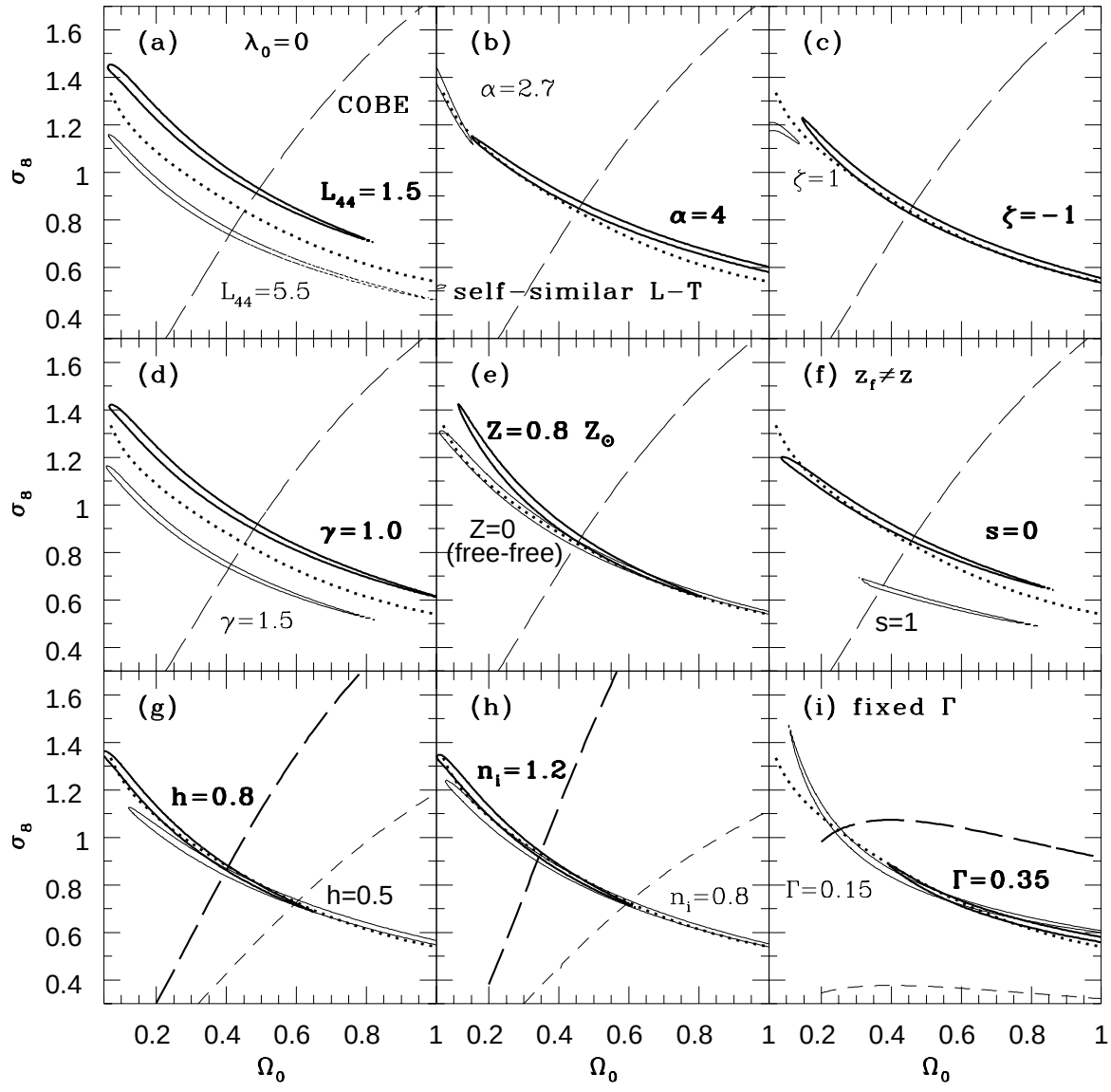


Figure 5.6: Same as Figure 5.5, except for $\lambda_0 = 0$.

Table 5.1: CDM model parameters from the *ROSAT* X-ray log$N$–log$S$.

| Model | $\Omega_0$ | $\lambda_0$ | $h$ | $\sigma_8$ | $\alpha$ | $\gamma$ |
|---|---|---|---|---|---|---|
| L03 | 0.3 | 0.7 | 0.7 | 1.04 | 3.4 | 1.2 |
| O045 | 0.45 | 0 | 0.7 | 0.83 | 3.4 | 1.2 |
| E1 | 1.0 | 0 | 0.5 | 0.56 | 3.4 | 1.2 |
| L03$\gamma$ | 0.3 | 0.7 | 0.7 | 0.90 | 3.4 | 1.5 |
| L01$\alpha$ | 0.1 | 0.9 | 0.7 | 1.47 | 2.7 | 1.2 |

amount to that due to the changes of $1.5 < L_{44} < 5.5$. Note that a change in $\gamma$ or $s$ affect the XTF and log$N$–log$S$ contours in a similar manner, and the resulting constraints from the log$N$–log$S$ and XTF remain consistent with each other.

The ranges of parameters $\alpha$, $\zeta$, $\gamma$, $Z$, and $s$ considered in Figures 5.5 and 5.6 are, unlike that of $L_{44}$, not directly related to definite statistical consideration. Furthermore, it is difficult to judge quantitatively how their intrinsic uncertainties correlate with one another. So we simply illustrate their individual effects in the figures, and quote only the representative systematic error due to $L_{44}$ for definiteness and simplicity. As is clear from Figures 5.5 and 5.6, with the ranges of the parameters considered here, the error due to $L_{44}$ represents a reasonable estimate for the total systematic uncertainty in the best-fit $\sigma_8$ value.

Panels (g)–(i) of Figures 5.5 and 5.6 indicate that the constraints on $\Omega_0$ and $\sigma_8$ from the cluster number counts depend only weakly on the parameters $h$, $n_{\rm i}$, or $\Gamma$, which specify the shape of the fluctuation spectrum. This is in contrast to the fact that the *COBE* normalized $\sigma_8$ is very sensitive to these parameters.

Another interesting application of X-ray cluster number counts can be found in probing the underlying $L-T$ relation. Figures 5.4, 5.5 and 5.6 indicate that the log$N$–log$S$ and XTF contours overlap with each other at the $\pm1\sigma$ level for our canonical $L-T$ relation with $2.7 < \alpha < 4$ or $-1 < \zeta < 1$. On the other hand, they are in good agreement with each other only if $L_{44} = 2.9$, and merely marginally consistent at the $2\sigma$ level if $L_{44} = 1.5$ or 5.5. Incidentally, if we use the theoretical $L-T$ relation based on the self-similar assumptions, the log$N$–log$S$ and XTF contours do not agree with each other even at the $3\sigma$ level. These reflect the fact that the predicted log$N$–log$S$ is sensitive to the adopted $L-T$ relation (Fig. 5.3). Thus, with more accurate determination of the log$N$–log$S$ and XTF by the future observations, one will be able to constrain the $L-T$ relation more tightly.

## 5.4 Cluster number counts in different wavebands

In §5.3, we have shown that CDM models with a certain range of parameters reproduce the *ROSAT* log$N$–log$S$ data as well as the XTF (Henry & Arnaud 1991) and the *COBE* 4 year data (Bunn & White 1997). To be definite, we hereafter consider five models with different sets of parameters as summarized in Table 5.1, and see if one can break the degeneracy of

the models by comparing the log$N$–log$S$ relations in different wavebands.

Figure 5.7 shows the number counts of clusters of galaxies in the soft X-ray (0.5–2 keV) band, the hard X-ray (2–10 keV) band, and the submm (0.85 mm) band. In the soft X-ray band, all of the adopted models by definition reproduce well the observed log$N$–log$S$ from the *ROSAT* Deep Cluster Survey (RDCS, Rosati et al. 1995, 1998) and the *ROSAT* Brightest Cluster Sample (BCS, Ebeling et al. 1997, 1998). In addition, we plot contributions from clusters in certain mass ranges and redshift ranges to the log$N$–log$S$ relations in Figures 5.8 and 5.9, respectively. It is shown that the cluster log$N$–log$S$ relation in these bands consist mainly of clusters of mass $M \gtrsim 10^{14}\,\mathrm{M}_\odot$ and at redshift $z \lesssim 0.5$. In general, contributions from massive clusters as well as from distant clusters become larger in low $\Omega_0$ models.

In the hard X-ray band, massive clusters make a relatively larger contribution to the predicted log$N$–log$S$ than in the soft band (Fig. 5.8). Thus, models with lower value of $\Omega_0$, i.e. flatter fluctuation spectra (more power at large scales), predict larger number of clusters in Figure 5.7(b). All the models considered here are shown to lie under the upper limit inferred from the number of *all* X-ray sources detected by *ASCA* (Cagnoni et al. 1998; Ueda et al. 1998). In addition, all the models except E1 are consistent with the number of clusters observed by *HEAO 1* (Piccinotti et al. 1982) within the $1\sigma$ errors. The predicted count in model E1 is smaller than the *HEAO 1* data at the $3\sigma$ level. Possible incompleteness in the *HEAO 1* result would simply raise the observed data point and does not reconcile the discrepancy.

In the submm band, the flux produced by the SZ effect is related to the mass, temperature, and redshift of clusters in a completely different manner (eq. [5.10]) from the X-ray flux. As already mentioned, the flux diminishes less slowly at high redshifts than in X-ray. In addition, at given redshift, the SZ flux depends less strongly on mass and temperature; $S \propto MT$ in submm, whereas $S \propto T^\alpha \sim T^{3.4}$ in X-ray. For the models which produce almost identical soft X-ray number counts, these two effects combine to yield greater submm counts in low $\Omega_0$ models (Fig. 5.7c). Compared with the hard X-ray band, the degeneracy of the models are broken in much greater extent in the submm band. Therefore, future determination of the submm cluster counts will enable us to constrain the models more tightly.

To see this clearly, we plot in Figure 5.10 contour maps of the cluster log$N$–log$S$ in different bands on the $\Omega_0 - \sigma_8$ plane. For the 0.5–2 keV band, the $1\sigma$ significance level derived in §5.3 is plotted, while for the 2–10 keV and submm bands, the contours of the number of clusters per steradian ($10^2$, $10^3$, $10^4$ from bottom to top) at $S$(2–10 keV) $= 10^{-13}$ erg cm$^{-2}$ s$^{-1}$and $S_\nu(0.85\ \mathrm{mm}) = 50$ mJy are plotted, respectively. The $\sigma_8$ values derived from the *COBE* 4 year data (Bunn & White 1997) are also given for reference. Panels (a) and (b) show that the contours for the 2–10 keV band counts run almost parallel to the $\chi^2$ contour of the 0.5–2 keV band counts. In this sense, future log$N$–log$S$ data in the hard X-ray band will provide an independent consistency check of the *ROSAT* soft X-ray data. The shape of the submm log$N$–log$S$ contours, on the other hand, is quite different, especially at high $\sigma_8$, and thus should place complementary constraints on $\Omega_0$ and $\sigma_8$. If the fixed shape parameter ($\Gamma = 0.25$) is adopted, the corresponding log$N$–log$S$ contours are similar to the conventional CDM ($\Gamma \simeq \Omega_0 h$) case, except at $\Omega_0 \lesssim 0.2$ and $\Omega_0 \gtrsim 0.8$. By contrast, the *COBE* normalized $\sigma_8$ is very sensitive to any changes in the spectral shape.

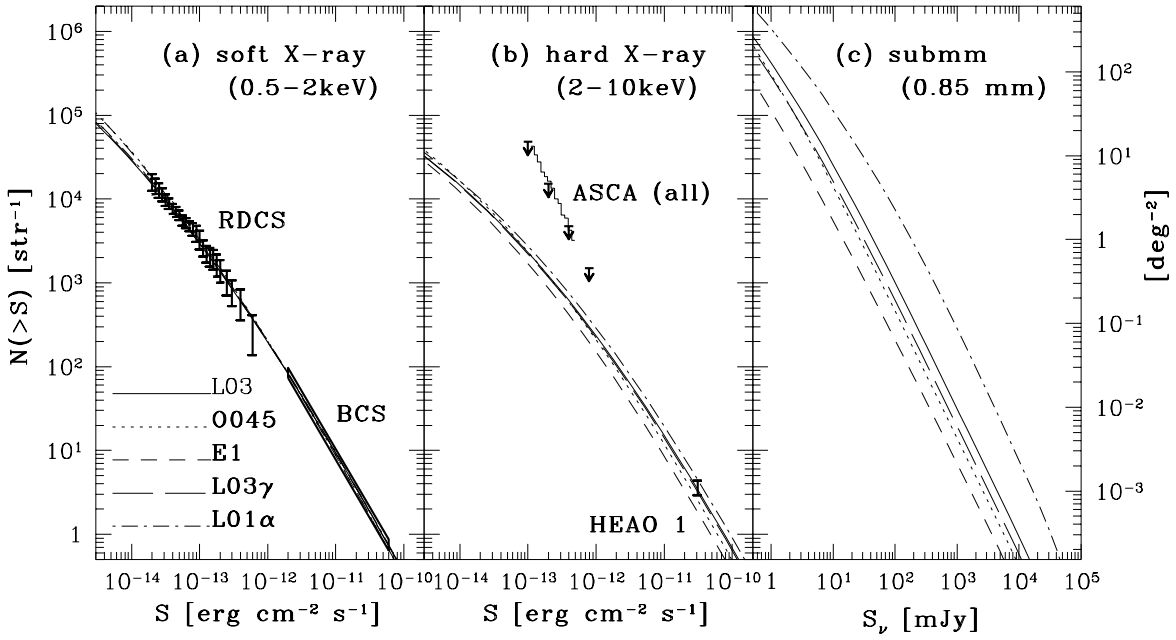


Figure 5.7: Log$N$–log$S$ relations of galaxy clusters for CDM models in (a) the soft X-ray (0.5–2 keV) band, (b) the hard X-ray (2–10 keV) band, and (c) the submm (0.85 mm) band. Lines represent the models listed in Table 5.1; L03 (solid), O045 (dotted), E1 (short dashed), L03$\gamma$ (long dashed), and L01$\alpha$ (dot-dashed). Also shown in panel (a) are the 1$\sigma$ error bars from the RDCS, and the error box from the BCS. In panel (b), the number of clusters observed by *HEAO 1* (Piccinotti et al. 1982) is indicated by the 1$\sigma$ error bar, and the counts of *all* X-ray sources detected by *ASCA* are plotted for reference as a histogram (Ueda et al. 1998) and downward arrows (Cagnoni et al. 1998).

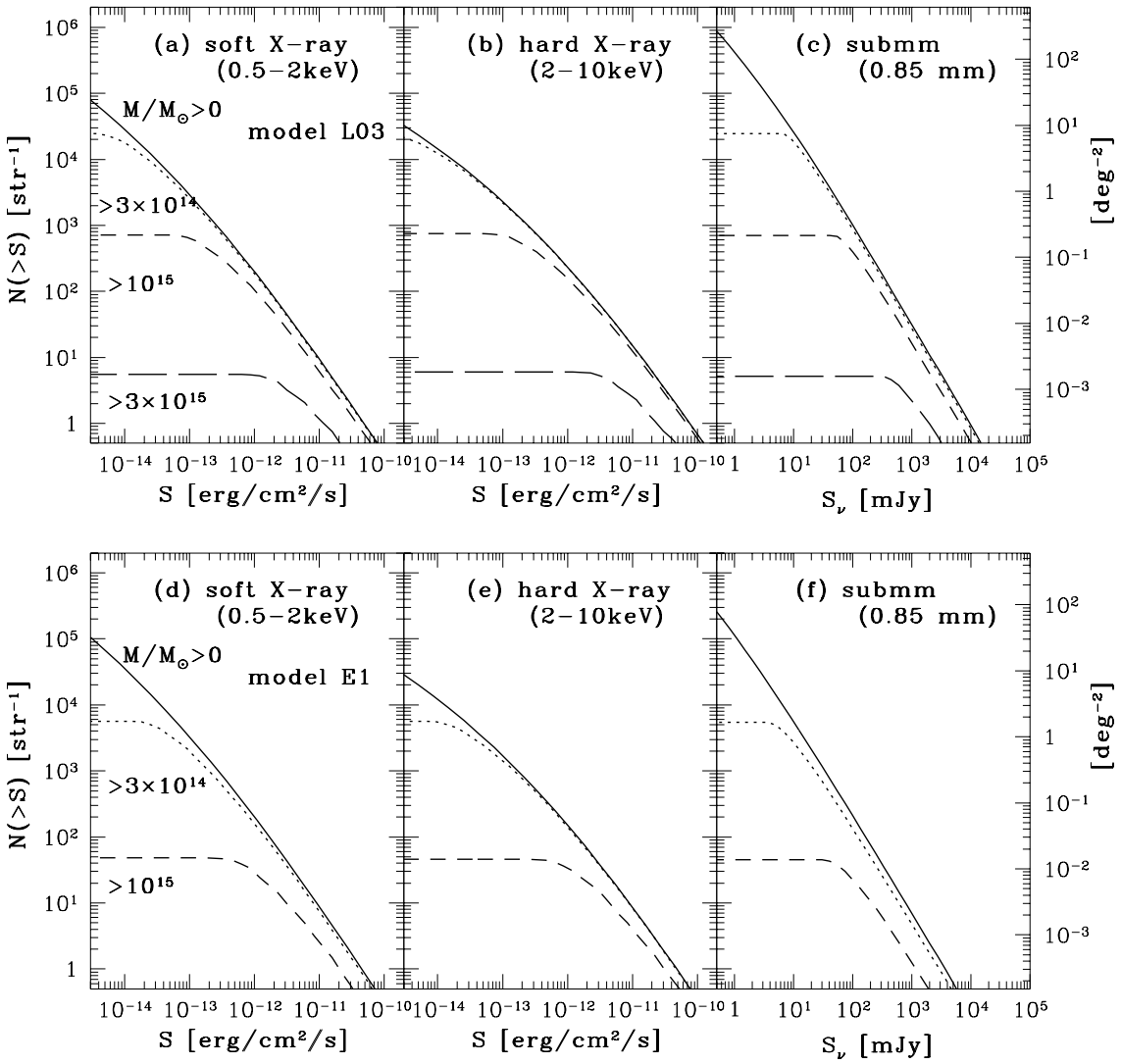


Figure 5.8: Contributions from clusters in certain mass ranges to the log$N$–log$S$ relations in (a)(d) the soft X-ray (0.5–2 keV) band, (b)(e) the hard X-ray (2–10 keV) band, and (c)(f) the submm (0.85 mm) band. Upper panels are for model L03, while lower panels are for model E1. Lines indicate the number of clusters with $M > 0$ (solid), $M > 3 \times 10^{14}\,\mathrm{M}_\odot$ (dotted), $M > 10^{15}\,\mathrm{M}_\odot$ (short dashed), and $M > 3 \times 10^{15}\,\mathrm{M}_\odot$ (long dashed).

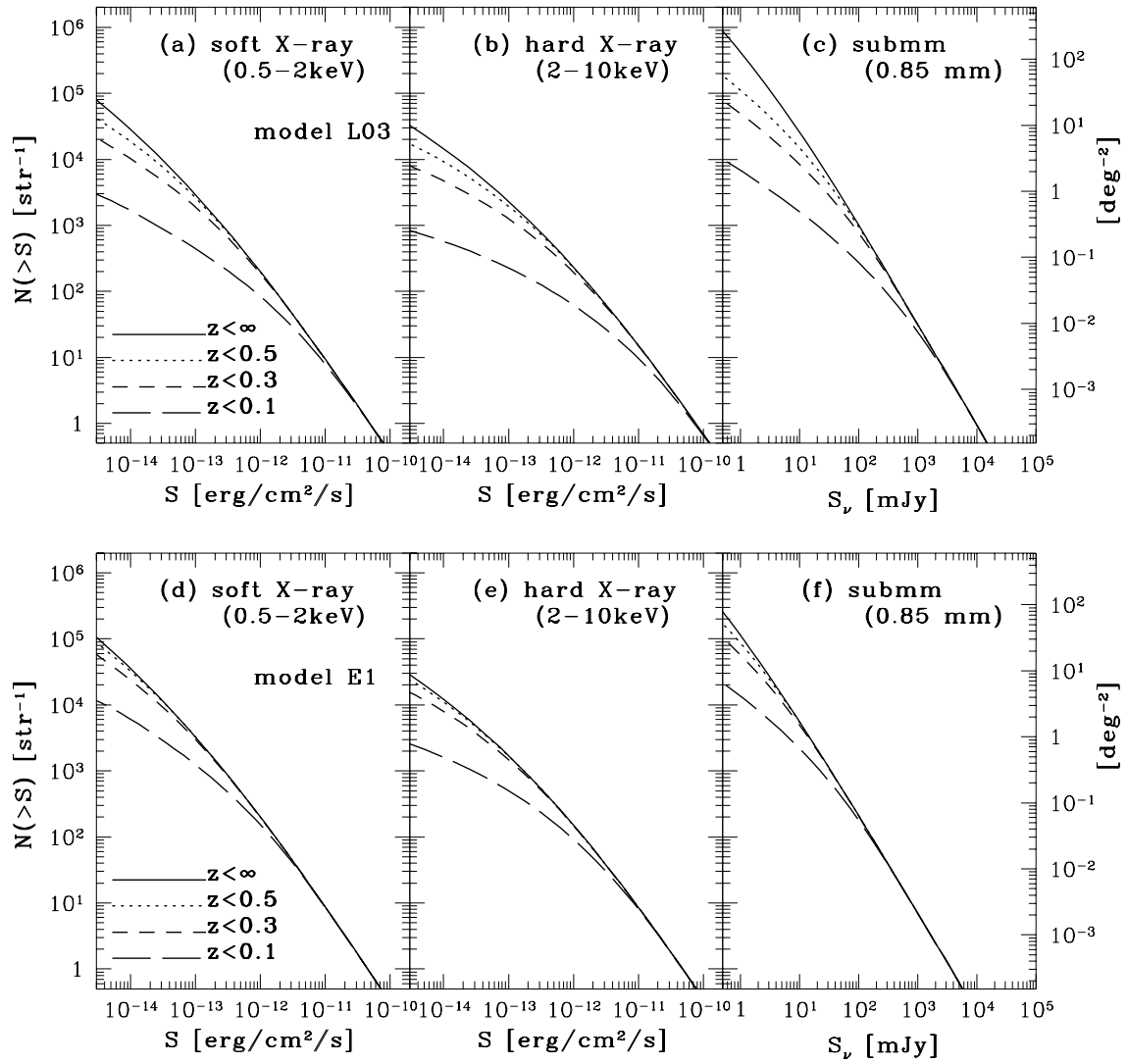


Figure 5.9: Contributions from clusters in certain redshift ranges to the log$N$–log$S$ relations in (a)(d) the soft X-ray (0.5–2 keV) band, (b)(e) the hard X-ray (2–10 keV) band, and (c)(f) the submm (0.85 mm) band. Upper panels are for model L03, while lower panels are for model E1. Lines indicate the number of clusters with $z < \infty$ (solid), $z < 0.5$ (dotted), $z < 0.3$ (short dashed), and $z < 0.1$ (long dashed).

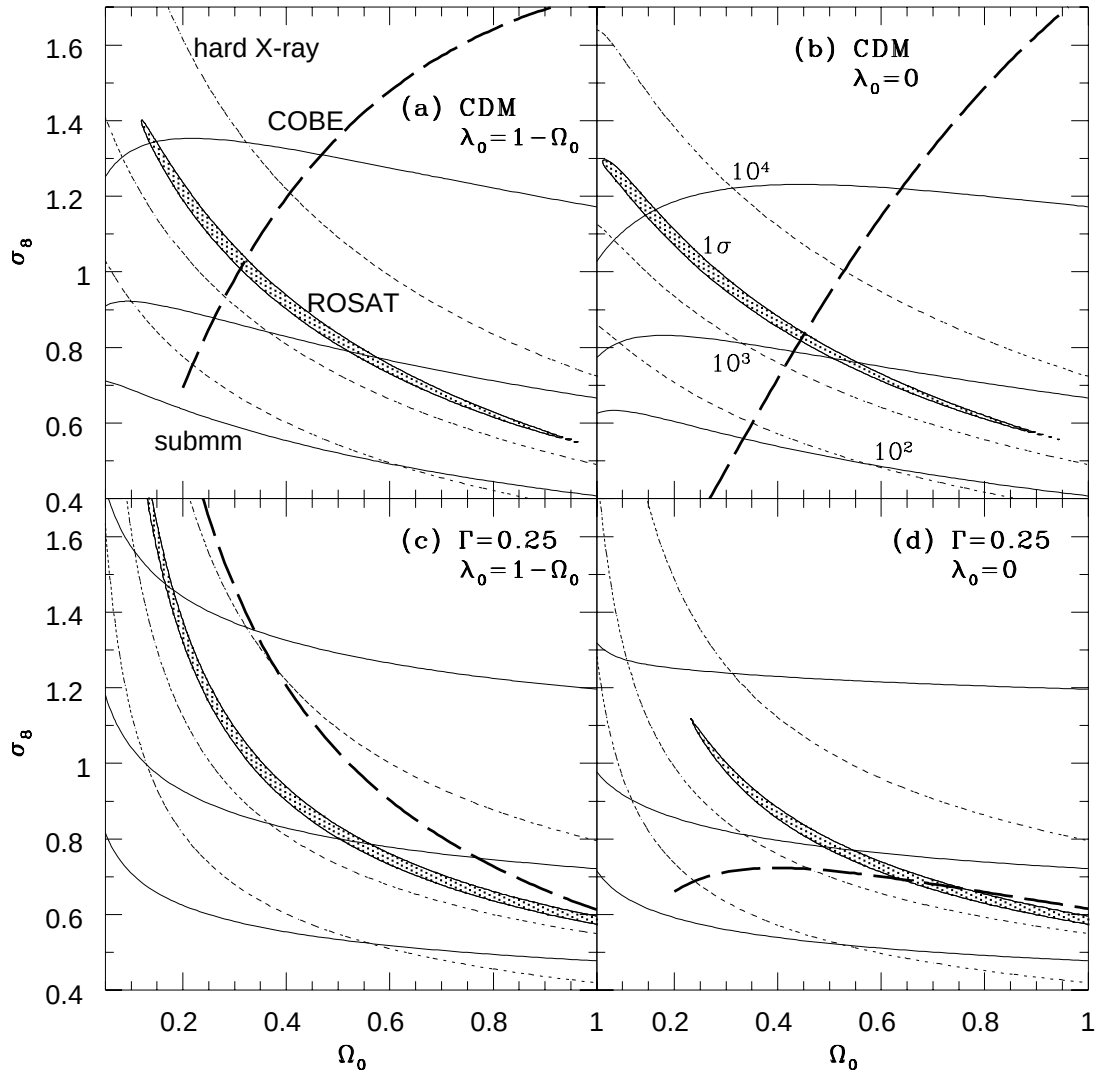


Figure 5.10: Contour maps on the $\Omega_0$-$\sigma_8$ plane in (a) spatially flat ($\lambda_0 = 1 - \Omega_0$) CDM models, (b) open ($\lambda_0 = 0$) CDM models, (c) spatially flat CDM-like models with the fixed shape parameter ($\Gamma = 0.25$), and (d) open CDM-like models with $\Gamma = 0.25$. In all cases, $h = 0.7$ and the fiducial set of intracluster gas parameters is assumed. Shaded regions represent the $1\sigma$ significance contours derived in §5.3from the soft X-ray (0.5–2 keV) $\log N$–$\log S$. Dotted and solid lines indicate the contours of the number of clusters greater than $S$ per steradian ($10^2$, $10^3$, $10^4$ from bottom to top) with $S = 10^{-13}$ erg cm$^{-2}$ s$^{-1}$in the hard X-ray (2–10 keV) band and with $S_\nu = 50$ mJy in the submm (0.85 mm) band, respectively. Thick dashed lines represent the *COBE* 4 year results by Bunn & White (1997).

## 5.5 Evolution of cluster number counts

Evolutionary behavior of the cluster abundance would provide a useful probe of the cosmological parameters, as discussed by a number of authors (e.g., Jing & Fang 1994; Viana & Liddle 1996; Eke, Cole, Frenk 1996; Kitayama & Suto 1996b; Oukbir & Blanchard 1997; Carlberg et al. 1997; Henry 1997; Sadat, Blanchard & Oukbir 1998; Mathiesen & Evrard 1998). It is of particular importance because the present and near future observations will provide us with a rapidly increasing amount of information concerning high redshift clusters. There are already a few (candidate) clusters found at $z \lesssim 1$ (e.g. Hattori et al. 1997; Yamada et al. 1997). We thus consider the evolution of the cluster distribution and discuss whether one can break the degeneracy among the models inferred from the soft X-ray log$N$–log$S$ relation (Table 5.1).

Figure 5.11 exhibits the redshift evolution of the number of clusters in different bands. As expected, the evolutionary behavior strongly depends on the values of $\Omega_0$ and $\sigma_8$; the fraction of low redshift clusters becomes larger for greater $\Omega_0$ and smaller $\sigma_8$. It is indicated that one may be able to distinguish among these models merely by determining the redshifts of clusters up to $z \sim 0.2$ (see also discussion below).

In order to characterize the physical properties of galaxy clusters, we have so far focused on their flux, because it is the simplest quantity which can be determined primarily from observations. Another major quantity of clusters is their temperature. Although the sample of clusters with measured temperature is still limited, an increasing amount of temperature information is going to become available from the *ASCA* and future observations. Furthermore, the temperature has an advantage in that it is insensitive to the detailed distribution of intracluster gas and much easier to model, compared to the X-ray flux or luminosity. This is, in fact, the main reason why many of the previous analysis of cluster abundance and its evolution are based upon the temperature rather than the X-ray luminosity (White et al. 1993; Viana & Liddle 1996; Eke et al. 1996). However, the observed samples of clusters are usually flux limited, which needs to be kept in mind when one compares the theoretical predictions with the observed statistics. Within our present framework, it is possible to explicitly incorporate the effects of limiting fluxes in the analysis based upon the cluster temperature.

Figure 5.12 shows the cumulative distribution of clusters with a temperature greater than $T$, a redshift less than $z$, and with/without an X-ray flux limit. It is apparent that the presence of a flux limit affects significantly the temperature distribution and its evolution, especially at low $T$ and high $z$. If the flux limit is $S = 10^{-12}$ erg cm$^{-2}$ s$^{-1}$in the 0.5–2 keV band, for instance, the number of clusters observed with $T > 2$ keV and $z < 0.3$ per unit solid angle is an order of magnitude less than what would have been expected without any flux limit.

The above discussion has clarified that the redshift evolution even at low $z$ and the temperature distribution of a given cluster sample with a specific flux limit can provide a useful probe of cosmological models. Unfortunately, we do not yet have a X-ray cluster sample with complete redshift and temperature information. However, one may still demonstrate the power of this approach using the currently available "best" sample. For this purpose, we tentatively adopt the X-ray brightest Abell-type clusters (XBACs, Ebeling et al. 1996a,b), which is about 80% complete and consists of 242 clusters with $z < 0.2$ and $S > 5 \times 10^{-12}$

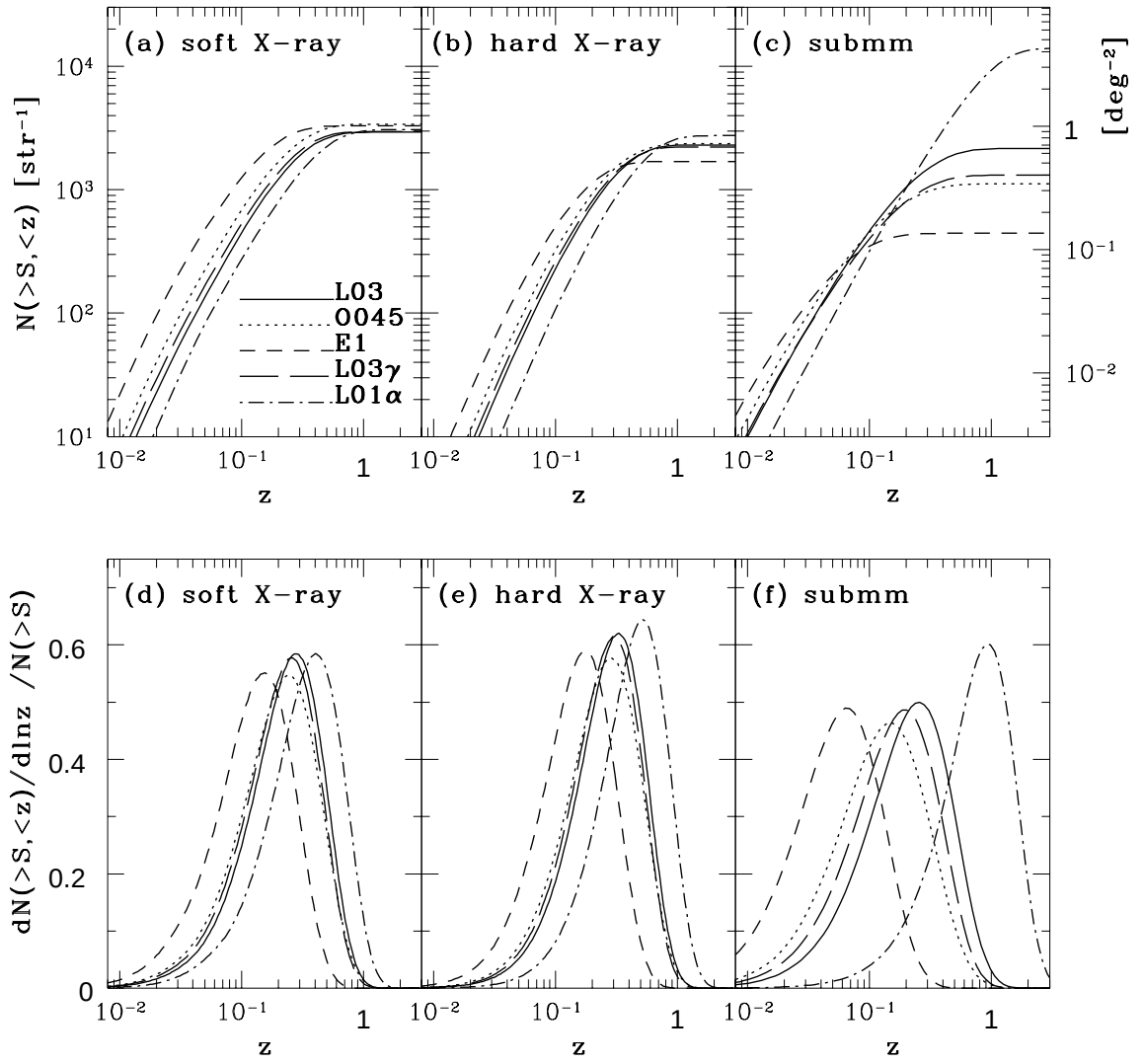


Figure 5.11: Redshift evolution of the number of galaxy clusters. Upper panels show the cumulative number $N(>S,<z)$ against $z$ in (a) the soft X-ray (0.5–2 keV) band with $S = 10^{-13}$ erg cm$^{-2}$ s$^{-1}$, (b) the hard X-ray (2–10 keV) band with $S = 10^{-13}$ erg cm$^{-2}$ s$^{-1}$, and (c) the submm (0.85 mm) band with $S_\nu = 50$ mJy. Lower panels (d)–(f) are similar to (a)–(c) except for plotting the differential distribution $\mathrm{d}N(>S,<z)/\mathrm{d}\ln z$ normalized by $N(>S)$.

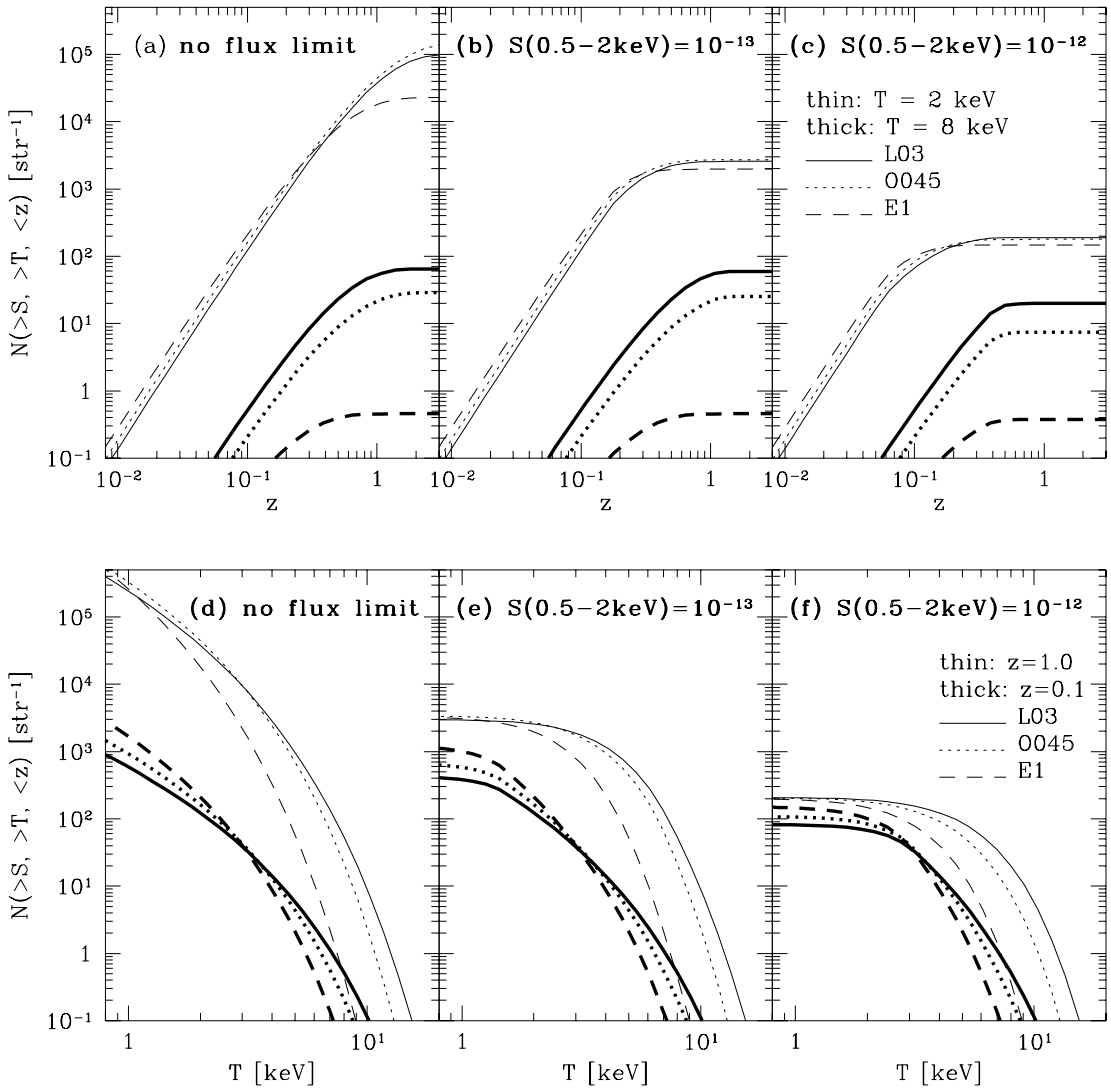


Figure 5.12: Effects of limiting fluxes on the redshift and temperature distributions of galaxy clusters. Upper panels exhibit the number $N(>S,>T,<z)$ against $z$ for $T = 2$ keV (thin lines) and $T = 8$ keV (thick) in the cases of (a) $S = 0$ (no flux limit), (b) $S(0.5\text{–}2\ \mathrm{keV}) = 10^{-13}$ erg cm$^{-2}$ s$^{-1}$, and (c) $S(0.5\text{–}2\ \mathrm{keV}) = 10^{-12}$ erg cm$^{-2}$ s$^{-1}$. Lower panels (d)–(f) are similar to (a)–(c) except for plotting $N(>S,>T,<z)$ versus $T$ in the cases of $z = 1$ (thin) and $z = 0.1$ (thick).

erg cm$^{-2}$ s$^{-1}$ in the 0.1–2.4 keV band. We use the temperature and redshift data compiled in table 3 of Ebeling et al. (1996a,b). About 92% and 30% of all the clusters listed in this table have a measured redshift and temperature, respectively. For the rest of the sample, the table also provides the redshifts estimated from the magnitude of the 10th-ranked cluster galaxy, and the temperatures estimated from the empirical $L-T$ relation. Keeping in mind the incompleteness of the sample and uncertainties especially in the estimated temperature data, we simply intend to perform a crude comparison with our predictions.

The results are plotted in Figure 5.13. The sky coverage of the XBACs is hard to quantify, mainly due to the uncertain volume incompleteness of the underlying optical catalogue, as noted by Ebeling et al. (1996a). In Figure 5.13, therefore, we simply plot the real numbers of the XBACs and normalize all of our model predictions to match the total number of the XBACs at its flux limit $S(0.1\text{–}2.4\ \mathrm{keV}) = 5\times10^{-12}$ erg cm$^{-2}$ s$^{-1}$and redshift limit $z = 0.2$. In general, the model predictions are shown to be capable of reproducing well the shape and amplitude of the observed distributions $N(>S, >T, <z)$, even when $T$ and $z$ are varied. Taking into account the incompleteness of the observed data and large statistical fluctuations at low numbers, the agreements with models L03 and L03$\gamma$ (both has $\Omega_0 = 0.3$) are rather remarkable. Since the shapes and amplitudes of the predicted curves in these figures are primarily determined by the value of $\Omega_0$, this result provides a further indication for a low $\Omega_0$ universe.

## 5.6 Contribution to the X-ray and submm background radiations

The X-ray background radiation (XRB, Giacconi et al. 1962) and the submm background radiation (SBR, Puget et al. 1996) provide useful probes of the cosmic structures at high redshifts. While their origins are yet to be resolved, the observed number counts of clusters of galaxies place stringent constraints on their contribution to the XRB and SBR. In this section, we derive such constraints based on the results presented in the previous sections.

From the observed log$N$–log$S$ relation in Figure 5.7(a), the cluster contribution to the XRB is estimated roughly as

$$\begin{aligned} I^{\rm cl}(0.5\text{–}2\ \mathrm{keV}) &= \int_{S_{\rm min}}^{\infty} S\left|\frac{dN}{dS}\right| dS \\ &\sim 2\times10^{-10}\left(\ln\frac{S_{12}}{S_{\rm min}}+4.3\right)\ \mathrm{erg\,cm^{-2}\,s^{-1}str^{-1}}, \end{aligned} \tag{5.17}$$

where $S_{12}\equiv10^{-12}$ erg cm$^{-2}$ s$^{-1}$, and we have fitted the 0.5–2 keV log$N$–log$S$ by a broken power-law: $N(>S)\propto S^{-1}$ at $S<S_{12}$ and $N(>S)\propto S^{-1.3}$ at $S>S_{12}$. The above estimate depends only weakly on the faint-end flux $S_{\rm min}$, and taking for instance $S_{\rm min}=10^{-15}$ erg cm$^{-2}$ s$^{-1}$yields $I_{\rm cl}(0.5\text{–}2\ \mathrm{keV})\sim2\times10^{-9}$erg cm$^{-2}$ s$^{-1}$str$^{-1}$, which is less than 10% of the observed XRB (Gendreau et al. 1995; Suto et al. 1996).

To be more specific, we compute numerically the contribution of clusters to the XRB. The XRB intensity from clusters $I_\nu^{\rm cl}(E_0)$ at energy $E_0$, as seen by an observer at present, is

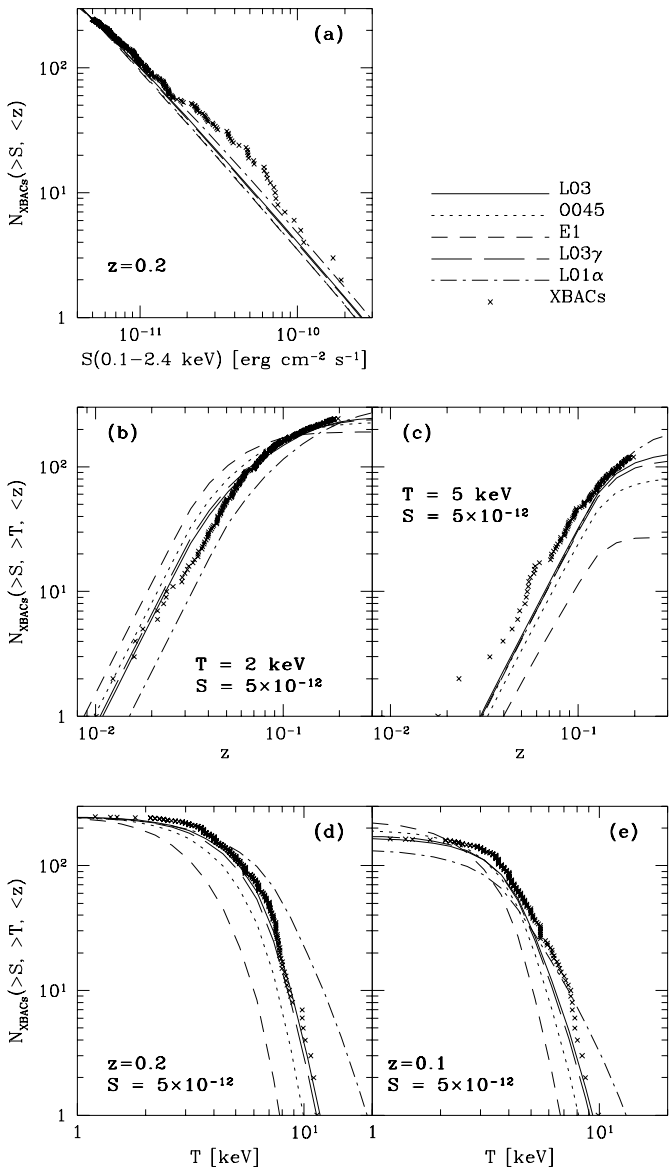


Figure 5.13: Tentative comparison with the XBACs (Ebeling et al. 1996a,b). Upper panel (a) shows the log$N$–log$S$ in the 0.1–2.4 keV band with the redshift limit of $z = 0.2$ (models L03 and O045 almost overlap with L03$\gamma$ and E1 respectively). Middle panels exhibit $N(>S,>T,<z)$ versus $z$ for $S(0.1\text{–}2.4\ \mathrm{keV}) = 5\times10^{-12}$ erg cm$^{-2}$ s$^{-1}$in the cases of (b) $T = 2$ keV and (c) $T = 5$ keV. Lower panels plot $N(>S,>T,<z)$ versus $T$ for $S(0.1\text{–}2.4\ \mathrm{keV}) = 5\times10^{-12}$ erg cm$^{-2}$ s$^{-1}$in the cases of (d) $z = 0.2$ and (e) $z = 0.1$. All the model predictions are normalized to reproduce the total number of the XBACs at its flux limit $S(0.1\text{–}2.4\ \mathrm{keV}) = 5\times10^{-12}$ erg cm$^{-2}$ s$^{-1}$and redshift limit $z = 0.2$.

given by

$$I_\nu^{\rm cl}(E_0) = \int_0^\infty \left(\frac{{\rm d}V}{{\rm d}z}\right) \frac{(1+z)\, J_\nu[E_0(1+z),z]}{4\pi d_{\rm L}^2(z)}\, {\rm d}z$$
$$= \frac{c}{4\pi H_0} \int_0^\infty \frac{J_\nu[E_0(1+z),z]}{(1+z)\sqrt{\Omega_0(1+z)^3+(1-\Omega_0-\lambda_0)(1+z)^2+\lambda_0}}\, {\rm d}z, \quad (5.18)$$

where ${\rm d}V/{\rm d}z$ is the comoving volume element per unit solid angle given in equation (A.15), and $J_\nu(E,z)$ is the comoving space-averaged volume emissivity at redshift $z$. We estimate the latter by

$$J_\nu(E,z) = \int_0^\infty L_\nu(E,M,z)\, \frac{{\rm d}n(M,z)}{{\rm d}M}\, {\rm d}M. \quad (5.19)$$

It should be noted that the observed XRB spectrum is not derived from all sky surveys (in particular, at lower energy band), but from small regions in the sky where there are very few known bright X-ray sources. In order to compare with such observations, therefore, we also need to omit very bright sources in our theoretical predictions by equation (5.18). For this purpose, we only consider contributions from clusters with flux below the critical value $S_{\rm crit}(0.5\text{–}2\ {\rm keV}) = 10^{-13}\ {\rm erg\,cm^{-2}\,s^{-1}}$, which roughly corresponds to the flux of the brightest X-ray source observed in the Lockman Hole by *ROSAT* (Hasinger et al. 1993).

Figure 5.14(a) exhibits the XRB intensity from clusters of galaxies in the cases of models L03 and E1, both of which reproduce the observed $\log N$–$\log S$ relation in the soft X-ray (0.5–2 keV) band. In these models, clusters contribute to the XRB only less than 20% at $E_0 \lesssim 2$ keV, which is consistent with the rough estimate described above. Thus, clusters of galaxies cannot be the major sources for the XRB, even in soft X-ray bands (Evrard & Henry 1991; Oukbir, Bartlett & Blanchard 1997). As is clear from the above analysis, this conclusion is based on the observed $\log N$–$\log S$ relation, and is almost independent of the assumed cosmological parameters. In fact, if one adopts a theoretical $L_{\rm bol} - T$ relation inferred from the self-similar assumption (Kaiser 1986), which yields $\alpha = 2$ in equation (4.27), there exist some models in which clusters account for the bulk of the soft XRB (Silk & Tarter 1973; Schaeffer & Silk 1988; Blanchard et al. 1992; Burg, Cavaliere & Menci 1993; Kitayama & Suto 1996a). Such models, however, are simply in conflict with both the observed $L_{\rm bol} - T$ relation ($\alpha = 3 - 3.5$) and the *ROSAT* $\log N$–$\log S$ relation.

A similar analysis can be performed for the SBR. Since we do not yet have the observed $\log N$–$\log S$ in the submm band, we fit the predicted $\log N$–$\log S$ of model L03 and obtain

$$\nu I_\nu^{\rm cl}(0.85\ {\rm mm}) \sim 10^{-12} \left(\frac{S_{\nu,{\rm min}}}{100\ {\rm mJy}}\right)^{-0.5} {\rm W\ m^{-2}\ str^{-1}}, \quad (5.20)$$

where we have assumed a single power law with $N(>S_\nu) \propto S_\nu^{-1.5}$. Taking $S_{\nu,{\rm min}} = 1$ mJy gives $\nu I_\nu^{cl}(0.85\ {\rm mm}) \sim 10^{-11}{\rm W\ m^{-2}\ str^{-1}}$, which is about 3% of the detected SBR by Puget et al. (1996). The results from numerical integration shown in Figure 5.14(b) confirm that the clusters of galaxies contribute only less than $\sim 5\%$ of the SBR.

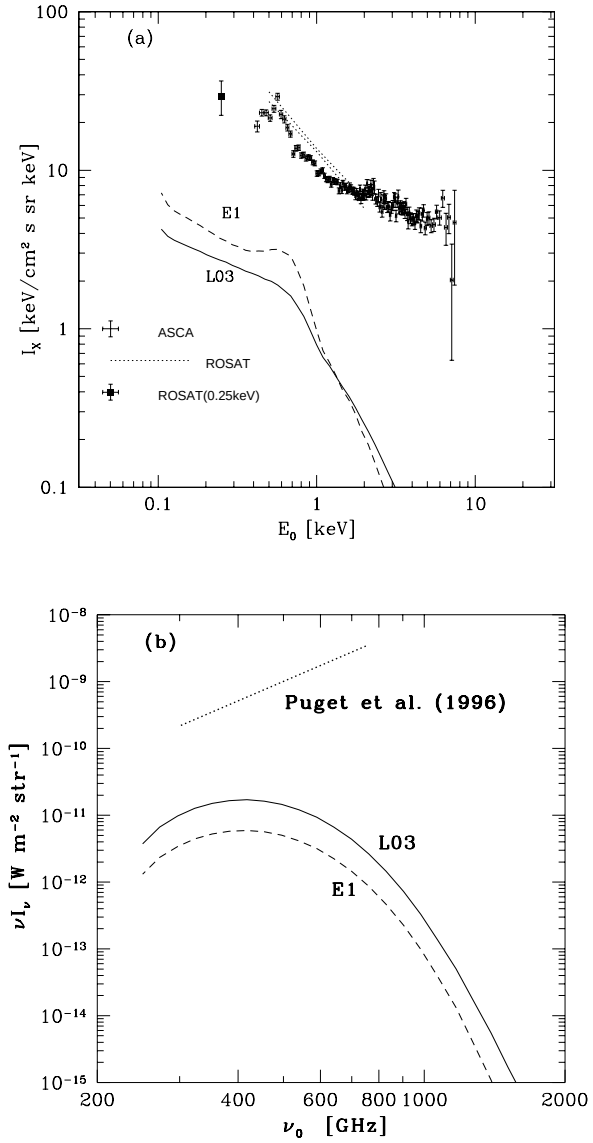


Figure 5.14: Contribution of clusters of galaxies to (a) the XRB, and (b) the SBR, in models L03 (solid) and E1 (dashed). Also shown are the observed XRB data of *ASCA* (Gendreau et al. 1995), *ROSAT* (Hasinger 1992), *ROSAT* at 0.25 keV (Barber, Roberts & Warwick 1996), and the tentative detection of the cosmic far-infrared background by *COBE* (Puget et al. 1996).

# Chapter 6

# Conclusions

In this thesis, we have studied the implications that the abundances of clusters of galaxies have on cosmology. We have made theoretical predictions for the cluster abundances based on the hierarchical clustering scenario, a standard picture of structure formation, and physical models of the intracluster gas. These predictions are then compared with the latest observations to derive constraints on the power spectrum of density fluctuations and cosmological parameters. We have also investigated possibilities to further constrain the structure formation models from future observations. Our conclusions are summarized as follows.

First of all, we have shown that there exist a set of cosmological models that reproduce remarkably well the observed abundances of clusters. Such models agree simultaneously with the observed XTF and XLF discussed in Chapter 4 and the X-ray $\log N$–$\log S$ relation in Chapter 5 within their error bars. We stress that the agreement with the X-ray $\log N$–$\log S$ is especially notable, considering its high degree of accordance over almost four orders of magnitude in flux. We have discussed in detail theoretical uncertainties that may change our predictions. For this purpose, we further incorporate alternative formalisms to the PS mass function, developed in Chapter 3, to take into account explicitly cluster formation epochs. Consequently, the agreement between the theory and observations is found to be quite robust.

As discussed in Chapters 4 and 5, the predicted present-day XTF and XLF as well as the X-ray $\log N$–$\log S$ are the most sensitive to the density parameter $\Omega_0$ and the fluctuation amplitude $\sigma_8$, while they depend to a much lesser extent on the cosmological constant $\lambda_0$, the Hubble constant $h$, the primordial spectral index $n_{\rm i}$, or any other parameters for the intracluster gas properties. Thus, by comparing the predictions and observations, one can constrain $\Omega_0$ and $\sigma_8$ rather independently of other parameters. The resulting constraints ($1\sigma$) derived from the X-ray $\log N$–$\log S$ relation in Chapter 5 read $\sigma_8 = (0.54 \pm 0.02)\Omega_0^{-0.35-0.82\Omega_0+0.55\Omega_0^2}$ for $\lambda_0 = 1-\Omega_0$, and $\sigma_8 = (0.54 \pm 0.02)\Omega_0^{-0.28-0.91\Omega_0+0.68\Omega_0^2}$ for $\lambda_0 = 0$. The quoted errors indicate only the statistical ones from the observed $\log N$–$\log S$, and the systematic uncertainty from our theoretical modelling of X-ray flux in the best-fit value of $\sigma_8$ is estimated to be about 15%. The above values depend only weakly on the shape of the power spectrum. The constraints from the $\log N$–$\log S$ relation and the XTF are consistent with each other, while the former is more stringent due to the better statistical significance of the observed data.

The observations of the CMB by *COBE* provide a completely independent way of measuring the density fluctuation spectrum. If the *COBE* 4 year data (Bunn & White 1997) are combined with the above derived constraints, one finds that low density CDM models with

$(\Omega_0, \lambda_0, h, \sigma_8) \simeq (0.3, 0.7, 0.7, 1)$ and $(0.45, 0, 0.7, 0.8)$ in particular simultaneously account for the cluster log$N$–log$S$, XTF, and the *COBE* results. It should be noted that since *COBE* probes a scale of $\gtrsim 1$ Gpc, which is over two orders of magnitude larger than $8h^{-1}$Mpc, its constraints on the $\Omega_0 - \sigma_8$ plane are indeed sensitive to other model parameters such as $h$ and $\lambda_0$. It is nevertheless noteworthy that the CDM model with $(\Omega_0, \lambda_0, h, \sigma_8) \simeq (0.3, 0.7, 0.7, 1)$ is a specific example of the currently most successful cosmological model that explains as well the cosmic age, two-point correlation functions of galaxies and clusters, and small-scale velocity dispersions (e.g. Efstathiou, Sutherland & Maddox 1990; Suginohara & Suto 1991; Suto 1993; Watanabe, Matsubara & Suto 1994; Peacock 1997). On the other hand, our results confirm that the *COBE* normalized CDM model with $\Omega_0 = 1$ and $h \gtrsim 0.5$ significantly overpredicts the observed cluster abundances. This discrepancy cannot be reconciled by any of the systematic effects considered in this thesis and is indeed one of the strongest evidences against this model.

Given rather tight constraints on $\Omega_0$ and $\lambda_0$ from the recently observed cluster abundances, we investigate in Chapter 5 further possibilities to break the degeneracy of these parameters by future observations. One way of doing this is to combine the cluster abundances in different wavebands. We demonstrate this approach specifically in the submm band (0.85 mm) together with the hard (2–10 keV) and soft (0.5–2 keV) X-ray bands. In the submm band, models which reproduce the observed soft X-ray log$N$–log$S$ by *ROSAT* predict $N(> S_\nu) \sim (10^2 - 10^4)(S_\nu/100\mathrm{mJy})^{-1.5}$ str$^{-1}$ at $S_\nu \gtrsim 100$mJy. The amplitude of the above relation depends sensitively on $\Omega_0$ and $\sigma_8$, and in a substantially different manner from the soft X-ray log$N$–log$S$. Therefore, by combining the cluster number counts in the X-ray and submm bands, one will be able to break the degeneracy in the acceptable parameter region on the $\Omega_0 - \sigma_8$ plane. This indicates that the future observations by, for instance, the European *PLANCK* mission and the Japanese *LMSA* project would provide powerful probes of these parameters. In the hard X-ray band, the same models predict $N(> S) \sim 200(S/10^{-12}\mathrm{erg\,cm^{-2}\,s^{-1}})^{-1.3}$ str$^{-1}$ at $S \gtrsim 10^{-12}\mathrm{erg\,cm^{-2}\,s^{-1}}$, which is consistent with the upper limit inferred from the number of all X-ray sources observed by *ASCA* (Cagnoni et al. 1998; Ueda et al. 1998). If we compare our predictions with the earlier determination of the hard X-ray cluster counts from *HEAO 1* (Piccinotti et al. 1982), the model with $\Omega_0 = 1$ disagrees with the data at the $3\sigma$ level, while low $\Omega_0$ models reproduce the data rather well. We expect that future determination of the hard X-ray cluster log$N$–log$S$ by *ASCA*, *AXAF*, or *Astro-E* missions, would provide an important cross-check for our interpretation of the *ROSAT* log$N$–log$S$ data.

The number counts of clusters of galaxies also provide tight constraints on their contribution to the background radiation in the corresponding energy band. Based on the X-ray log$N$–log$S$ relation observed by *ROSAT* in the 0.5–2 keV band, we conclude that clusters of galaxies contribute at most $\sim 20\%$ of the total XRB and less than $\sim 5\%$ of the SBR.

The evolution of the cluster abundances offers another way of constraining cosmological models. We emphasize that the cluster number counts are much more suited for this purpose than the conventional XLF or XTF, because number counts are cosmological model-independent and bias-free observable at high redshifts. We have performed a tentative comparison between our theoretical predictions and the recent compilation of the XBACs by Ebeling et al. (1996a,b), which is the largest sample of galaxy clusters available to date with measured redshifts. While the incompleteness of the sample still makes it difficult to draw any definite conclusions from this comparison, it is noteworthy that our predictions reproduce

well the evolutionary features of the XBACs and that the results, although preliminary, seem to favor low density universes with $\Omega_0 \sim 0.3$. Interestingly, this result and the above mentioned constraints from the X-ray $\log N$–$\log S$ plus *COBE* observations independently point to $\Omega_0$ lower than unity. As a number of cluster surveys are now ongoing to identify high redshift clusters (e.g. Scharf et al. 1997; Collins et al. 1997; Burke et al. 1997; Jones et al. 1998; Rosati & Della Ceca 1998), it is very promising that we will be able to perform more quantitative analyses by means of the cluster evolution in the near future.

To conclude, the abundances of clusters of galaxies are proven to be an excellent probe of cosmology. We have shown the power of this approach which will become even a stronger tool as the observations attain greater precision.

# Acknowledgments

I express my sincere gratitude to Yasushi Suto, who is my supervisor and a collaborator of the present work, for his genuine support and encouragement. Special thanks are as well due to Katsuhiko Sato for his continuous help, and Shin Sasaki, a collaborator of a part of the present work, for stimulating discussions. I am deeply obliged to Harald Ebeling, Patrick Henry, and Piero Rosati for generously providing their X-ray data and useful correspondences. I am also grateful to Tadayuki Takahashi for his careful reading of the manuscript and helpful suggestions. My appreciation goes further out to Nobuo Arimoto, Alain Blanchard, Yasuhiko Furihata, Masahiro Kawasaki, Kazuo Makishima, Takahiko Matsubara, Takahiro Nakamura, Sadanori Okamura, John Peacock, Kazuhiro Shimasaku, Tetsuya Shiromizu, Tatsushi Suginohara, Naoshi Sugiyama, Hajime Susa, Takahiro Tanaka, Tomonori Totani, and Masahide Yamaguchi for discussions and comments. I am indebted to Keita Ikumi and Shinya Tsubaki for their technical assistance with the computer network. Finally, I thank my family for their support and encouragement.

This research is supported in part by the Grants-in-Aid for Scientific Research from Ministry of Education, Science and Culture of Japan (09-7408). I acknowledge the fellowships from the Nippon Telegraph and Telephone Corporation and the Japan Society for the Promotion of Science for Young Scientists.

# Appendix A

# Kinematic properties of Expanding Universe

In this chapter, we briefly summarize kinematic effects of the expansion on the propagation of light in an isotropic universe.

## A.1 Cosmological redshift

Suppose that a wavecrest of the light is emitted from a source located at comoving coordinate $(r,\theta,\phi) = (r_1,0,0)$ at time $t_1$, and arrives at a detector at the origin $(r,\theta,\phi) = (0,0,0)$ at time $t_0$. As the massless wave travels along a geodesic ($\mathrm{d}s^2 = 0$), the coordinate distance and time are related by

$$\int_{t_1}^{t_0} \frac{\mathrm{d}t}{a(t)} = \int_0^{r_1} \frac{\mathrm{d}r}{(1-kr)^{1/2}}, \tag{A.1}$$

where the metric (2.1) is used. The next wavecrest emitted at time $t_1 + \delta t_1$ will arrive at the detector at time $t_0 + \delta t_0$. Since the source and detector positions are both fixed in the comoving coordinate system, the right-hand-side of equation (A.1) is constant and we have

$$\int_{t_1}^{t_0} \frac{\mathrm{d}t}{a(t)} = \int_{t_1+\delta t_1}^{t_0+\delta t_0} \frac{\mathrm{d}t}{a(t)}. \tag{A.2}$$

In most of astrophysically interesting cases, $\delta t_1$ and $\delta t_0$ are much shorter than the Hubble time over which $a(t)$ is changed. Then equation (A.2) yields

$$\frac{\delta t_1}{a(t_1)} = \frac{\delta t_0}{a(t_0)} \tag{A.3}$$

Thus the emitted and observed frequencies are related by

$$\frac{\nu_1}{\nu_0} = \frac{\delta t_0}{\delta t_1} = \frac{a(t_0)}{a(t_1)}. \tag{A.4}$$

The cosmological redshift is now expressed as

$$z \equiv \frac{\nu_1}{\nu_0} - 1 = \frac{a(t_0)}{a(t_1)} - 1. \tag{A.5}$$

If the universe is expanding, then $a(t_0) > a(t_1)$ and the light is red-shifted to give positive $z$. On the other hand, if the universe is contracting, then $a(t_0) < a(t_1)$ and the light is blue-shifted to give negative $z$.

It is evident from the above results and equation (2.14) that the radiation field which has the Plank spectrum $I_\nu \propto \nu^3/[\exp(h_{\rm P}\nu/k_{\rm B}T) - 1]$ will retain the spectral shape during the expansion, with the "effective" temperature $T$ varying as $T \propto a^{-1}$.

## A.2 Luminosity distance

If the source at coordinate $r_1$ emits $\delta N$ photons with frequency $\nu_1 \sim \nu_1 + \delta\nu_1$ at time $t_1 + \delta t_1$, the intrinsic luminosity in this frequency range is written as

$$\delta L = h_{\rm p}\nu_1 \frac{\delta N}{\delta t_1}. \tag{A.6}$$

Since the photon number is conserved, the detector at the origin $r = 0$ will receive at time $t_0$ the flux given by

$$\delta f = \frac{h_{\rm p}\nu_0}{4\pi a^2(t_0) r_1^2} \frac{\delta N}{\delta t_0} = \frac{\delta L}{4\pi a^2(t_0) r_1^2 (1+z)^2}, \tag{A.7}$$

where equations (A.4), (A.5) and (A.6) are used. Note that equation (A.7) is independent of the emitted frequency. The luminosity distance $d_{\rm L}$ is now defined as

$$d_{\rm L} \equiv \left(\frac{\delta L}{4\pi\delta f}\right)^{1/2} = a(t_0) r_1 (1+z) \tag{A.8}$$

Solving equation (A.1) in the case of matter dominated era, one obtains

$$d_{\rm L}(z) = \frac{c}{H_0}(1+z) \begin{cases} \sin(\sqrt{K_0}\chi)/\sqrt{K_0} & (K_0 > 0) \\ \chi & (K_0 = 0) \\ \sinh(\sqrt{-K_0}\chi)/\sqrt{-K_0} & (K_0 < 0) \end{cases}, \tag{A.9}$$

where $K_0 = \Omega_0 + \lambda_0 - 1$, and $\chi(z)$ is

$$\chi(z) = \int_0^z \frac{{\rm d}z}{\sqrt{\Omega_0(1+z)^3 + (1-\Omega_0-\lambda_0)(1+z)^2 + \lambda_0}}. \tag{A.10}$$

In the case of $\lambda_0 = 0$, equation (A.9) reduces, regardless of $K_0$, to

$$d_{\rm L}(z) = \frac{2c}{H_0\Omega_0^2}[\Omega_0 z + (\Omega_0 - 2)(\sqrt{\Omega_0 z + 1} - 1)] \tag{A.11}$$

The luminosity distances in some representative cosmological models are plotted in Figure A.1 (a). For a given redshift, the luminosity distance is shown to become larger in low density models.

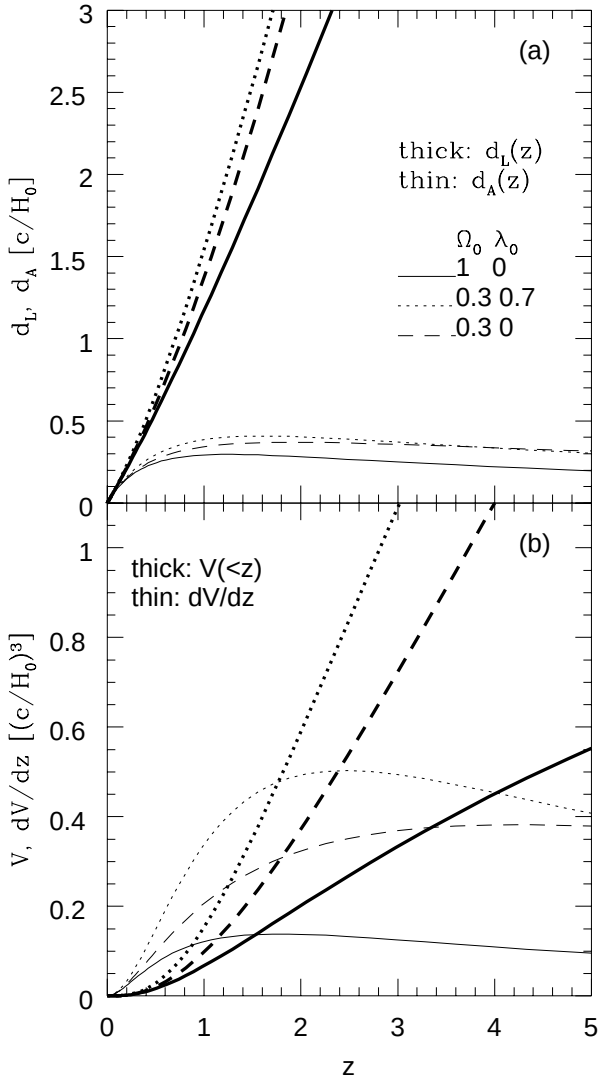


Figure A.1: (a) The luminosity distance (thick lines) and the angular diameter distance (thin lines) in units of $c/H_0$ as a function of redshift. Cosmological models are $(\Omega_0, \lambda_0) = (1, 0)$ (solid), $(0.3, 0.7)$ (dotted), and $(0.3, 0)$ (dashed). (b) The comoving volume $V(< z)$ (thick lines) and $dV/dz$ (thin lines) per unit solid angle in units of $(c/H_0)^3$ as a function of redshift. Cosmological models are the same as panel (a).

## A.3 Angular diameter distance

Consider an object of proper diameter $D$ located at coordinate $r_1$ that emits light at time $t = t_1$. From the metric (2.1), it follows that the angular diameter of the source $\theta$ observed at the origin $r = 0$ is related to $D$ by

$$\theta = \frac{D}{a(t_1)r_1} \tag{A.12}$$

The angular diameter distance $d_{\rm A}$ is defined as

$$d_{\rm A} \equiv \frac{D}{\theta} = a(t_1)r_1. \tag{A.13}$$

Thus the angular diameter distance (eq. [A.13]) and the luminosity distance (eq. [A.8]) are related via equation (A.5) as

$$d_{\rm A} = \frac{d_{\rm L}}{(1+z)^2} \tag{A.14}$$

The angular diameter distance in some representative cosmological models are plotted in Figure A.1 (a).

Using the angular diameter distance defined above, one can express the comoving volume element $\mathrm{d}V/\mathrm{d}z$ per unit solid angle on the sky at redshift $z$ as

$$\begin{aligned} \frac{\mathrm{d}V}{\mathrm{d}z} &= (1+z)^3 d_{\rm A}^2(z) c \left| \frac{\mathrm{d}t}{\mathrm{d}z} \right| \\ &= \frac{c}{H_0} \frac{(1+z)^2 d_{\rm A}^2(z)}{\sqrt{\Omega_0(1+z)^3 + (1-\Omega_0-\lambda_0)(1+z)^2 + \lambda_0}}, \end{aligned} \tag{A.15}$$

where $\mathrm{d}t/\mathrm{d}z = -H^{-1}(1+z)^{-1}$ and equation (2.19) are used. Figure A.1 (b) illustrates the above volume element $\mathrm{d}V/\mathrm{d}z$ together with the integrated volume up to $z$. The volume is shown to become larger for lower $\Omega_0$.

# Appendix B

# Fitting formulae for the Lacey & Cole distribution function

Since the distribution function of the halo formation epochs derived by LC (eq. [3.19]) requires a time-consuming numerical integration, the following fitting formulae are used in the present thesis.

Consider first the case of power-law fluctuation spectra $P(k) \propto k^n$, i.e. $\sigma^2(M) \propto M^{-\alpha}$ where $\alpha \equiv (n+3)/3$. Then equation (3.19) reduces to

$$\frac{\partial p}{\partial \tilde{\omega}_{\rm f}}(\alpha, \tilde{\omega}_{\rm f}) = \sqrt{\frac{8}{\pi}} {\rm e}^{-\tilde{\omega}_{\rm f}^2/2} - \frac{2(2^\alpha - 1)}{\sqrt{\pi}\alpha} \tilde{\omega}_{\rm f} \int_{\frac{\tilde{\omega}_{\rm f}}{\sqrt{2}}}^{\infty} {\rm d}y \frac{{\rm e}^{-y^2}}{y^2} \left(1 + \frac{2^\alpha - 1}{2} \frac{\tilde{\omega}_{\rm f}^2}{y^2}\right)^{(1-\alpha)/\alpha}. \quad \text{(B.1)}$$

For $\tilde{\omega}_{\rm f} \ll 1$, equation (B.1) becomes

$$\frac{\partial p}{\partial \tilde{\omega}_{\rm f}}(\alpha, \tilde{\omega}_{\rm f}) \simeq \sqrt{\frac{8}{\pi}} \left[1 - \frac{\sqrt{2^\alpha - 1}}{\alpha} \int_0^{\sqrt{2^\alpha - 1}} (1 + x^2)^{(1-\alpha)/\alpha} {\rm d}x\right] + \frac{2(2^\alpha - 1)}{\alpha} \tilde{\omega}_{\rm f} + O(\tilde{\omega}_{\rm f}^2), \quad \text{(B.2)}$$

For $\tilde{\omega}_{\rm f} \gg 1$, on the other hand, the numerical results for equation (B.1) are insensitive to $\alpha$ and well approximated by the solution in the $\alpha = 1$ case:

$$\frac{\partial p}{\partial \tilde{\omega}_{\rm f}}(\alpha = 1, \tilde{\omega}_{\rm f}) = 2\tilde{\omega}_{\rm f} \, {\rm erfc}\left(\frac{\tilde{\omega}_{\rm f}}{\sqrt{2}}\right). \quad \text{(B.3)}$$

On the basis of equations (B.2) and (B.3), we constructed the following interpolation formula which gives a very good approximation to equation (B.1):

$$\frac{\partial p}{\partial \tilde{\omega}_{\rm f}}(\alpha, \tilde{\omega}_{\rm f}) = \frac{A(\alpha)}{1 + B(\alpha)\tilde{\omega}_{\rm f}} {\rm e}^{-5\tilde{\omega}_{\rm f}^2} + 2C(\alpha)\tilde{\omega}_{\rm f} \, {\rm erfc}\left(\frac{\tilde{\omega}_{\rm f}}{\sqrt{2}}\right), \quad \text{(B.4)}$$

where

$$A(\alpha) \equiv \sqrt{\frac{8}{\pi}}(1 - \alpha)(0.0107 + 0.0163\alpha), \quad \text{(B.5)}$$

$$B(\alpha) \equiv \frac{2}{A(\alpha)} \left[C(\alpha) - \frac{2^\alpha - 1}{\alpha}\right], \quad \text{(B.6)}$$

$$C(\alpha) \equiv 1 - \frac{1 - \alpha}{25}. \quad \text{(B.7)}$$

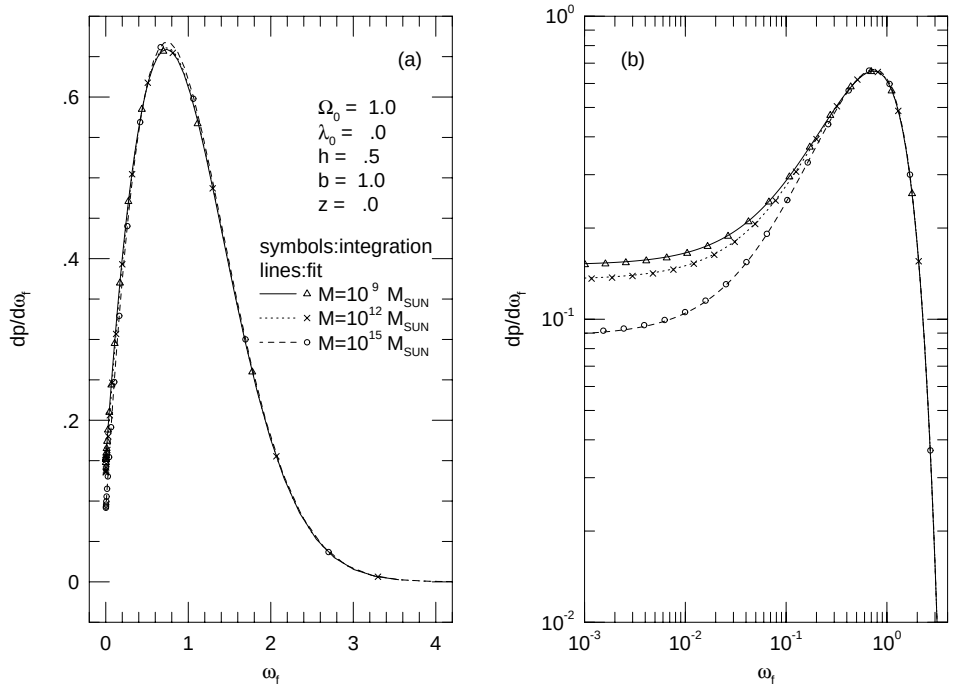


Figure B.1: The scaled distribution function of halo formation epochs, $\partial p/\partial\tilde{\omega}_{\rm f}$, in the standard CDM model with $\Omega_0 = 1$, $\lambda_0 = 0$, $h = 0.5$, and $b = 1/\sigma_8 = 1$; (a) linear, and (b) log-log plots. Symbols and lines indicate results of a numerical integration of equation (3.19) and a fit by equations (B.4)–(B.8), respectively; $M = 10^9\,{\rm M}_\odot$ (triangles, solid line), $10^{12}\,{\rm M}_\odot$ (crosses, dotted line), and $10^{15}\,{\rm M}_\odot$ (circles, dashed line).

Since equations (B.1)–(B.3) are valid only when the underlying spectra are of power-law form, the above fitting formulae may not be applicable to more general spectra. In fact, in the case of the CDM spectrum, a straightforward substitution of the effective spectral index $\alpha_{\rm eff} \equiv -{\rm d}\ln\sigma^2_{\rm CDM}/{\rm d}\ln M$ into equation (B.4) does not work very well, because $\alpha_{\rm eff}$ varies with the mass scale and so cannot be approximated by a constant in the integral of equation (3.19). Nevertheless, we found that equation (B.4) still gives an excellent empirical fit to equation (3.19) for the CDM spectrum, if $\tilde{\alpha}_{\rm eff}$ below is adopted rather than $\alpha_{\rm eff}$:

$$\tilde{\alpha}_{\rm eff} \equiv \alpha_{\rm eff}(0.6268 + 0.3058\alpha_{\rm eff}). \qquad \text{(B.8)}$$

The accuracy of this approximation is within a few percent in the range $10^6\,{\rm M}_\odot \lesssim M(\Omega_0 h^2)^2 \lesssim 10^{16}\,{\rm M}_\odot$ (see Fig. B.1).

# References


Abramopoulos, F., Ku, W. 1983, ApJ, 271, 446

Aghanim, A., De Luca, A., Bouchet, F. R., Gispert, R., Puget, J. L. 1997, A&A, 325, 9

Adler, R. J. 1981, The Geometry of Random Fields. Wiley, New York

Audit, E., Teyssier,R., Alimi,J.-M. 1997, A&A, 325, 439

Barber, C. R., Roberts, T. P., Warwick, R. S. 1996, MNRAS 282, 157

Barbosa, D., Bartlett, J. G., Blanchard, A., Oukbir, J. 1996, A&A 314, 13

Bahcall, N. A. 1996, in Formation of Structure in the Universe, 13th Jerusalem Winter School (astro-ph/9611148)

Bahcall, N. A., Lubin, L. M. 1994, ApJ, 426, 513

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

Bartelmann, M., Ehlers, J., Schneider, P. 1993, A&A, 280, 351

Birkinshaw, M., Hughes, J. P. 1994, ApJ, 420, 33

Blain, A. W., Longair, M. S. 1993, MNRAS, 265, L21

Blanchard, A., Silk, J. 1991, in Proc. 26th Rencontres de Moriond in Astrophysics, ed. B. Rocca-Volmerange, J. M. Deharveng, J. Tran Than Van, Gif-sur-Yvette; Editions Frontières, 93

Blanchard, A., Wachter, K., Evrard, A.E., Silk, J. 1992, ApJ, 391, 1

Blumenthal, G. R., Faber, S. M., Primack, J. R., Rees, M. J., 1984, Nature, 311, 517

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

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

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

Bryan, G. L., Cen, R., Norman, M. L., Ostriker, J. P., Stone, J. M. 1994, ApJ, 428, 405

Bryan, G. L., Norman, M. L. 1998, ApJ, 495, 80

Bunn, E. F., White, M. 1997, ApJ, 480, 6

Burg, R., Cavaliere, A., Menci, N. 1993, ApJ, 404, L55

Burke, D. J., Collins, C. A., Sharples, R. M., Romer, A. K., Holden, B. P., Nichol, R. C. 1997, ApJ, 488, 83

Cagnoni, I., Della Ceca, R., Maccacaro, T. 1998, ApJ, 493, 54

Carlberg, R. G., Couchman, H. M. P. 1989, ApJ, 340, 47

Carlberg, R. G., Morris, S.L., Yee, H.K.C., Ellingson, E. 1997, ApJ, 479, 19

Cavaliere, A., Colafrancesco, S., Scaramella, R. 1991, ApJ, 380, 15

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

Cen, R., Ostriker, J. P. 1993, ApJ, 417, 404

Colafrancesco, S., Mazzotta, P., Rephaeli, Y., Vittorio, N. 1997, ApJ, 479, 1

Collins, C. A., Burke, D. J., Romer, A. K., Sharples, R. M., Nichol, R. C. 1997, ApJ, 479, L117
Chandrasekhar, S. 1943, Rev. Mod. Phys., 15, 2
David, L. P., Slyz, A., Jones, C., Forman, W., Vrtilek, S. D. 1993, ApJ 412, 479
Davis, M., Efstathiou, G., Frenk, C., White, S. 1985, ApJ, 292, 371
Ebeling, H., Voges, W., Böhringer, H., Edge, A. C., Huchra, J. P., Briel, U. G. 1996a, MNRAS, 281, 799
Ebeling, H., Voges, W., Böhringer, H., Edge, A. C., Huchra, J. P., Briel, U. G. 1996b, MNRAS, 283, 1103
Ebeling, H., Edge, A. C., Fabian, A.C., Allen, S. W., Crawford, C. S. 1997, ApJ, 479, L101
Ebeling, H., Edge, A. C., Böhringer, H., Allen, S. W., Crawford, C. S., Fabian, A. C., Voges, W., Huchra, J. P. 1998, MNRAS, submitted
Edge, A. C., Stewart, G. C. 1991a, MNRAS, 252, 414
Edge, A. C., Stewart, G. C. 1991b, MNRAS, 252, 428
Efstathiou, G., Frenk, C. S., White, S. D. M., Davis, M. 1988, MNRAS, 235, 715
Efstathiou, G., Rees, M. J. 1988, MNRAS, 230, 5p
Efstathiou, G., Sutherland, W. J., Maddox, S. J. 1990, Nature, 348, 705
Eke, V. R., Cole, S., Frenk, C. S. 1996, MNRAS 282, 263
Eke, V. R., Navarro, J. F., Frenk, C. S. 1998, ApJ, in press
Evrard, A. E. 1989, ApJ, 341, L71
Evrard, A. E. 1990, ApJ, 363, 349
Evrard, A. E., Henry, J. P. 1991, ApJ 383, 95
Freedman, W. L., et al. 1994, Nature, 371, 757
Fukugita, M., Kawasaki, M. 1994, MNRAS, 269, 563
Gendreau, K. C., Mushotzky, R., Fabian, A. C., Holt, S. S., Kii, T., Serlemitsos, P. J., Ogasaka, Y., Tanaka, Y., et al. 1995, PASJ 47, L5
Giacconi, R., Gursky, H., Paolini, F. R., Rossi, B. B. 1962, Phys. Rev. Lett., 9, 439
Gioia, I. M., Henry, J. P., Maccacaro, T., Morris, S. L., Stocke, J. T., Wolter, A. 1990, ApJ, 356, L35
Grainge, K., Jones, M., Pooley, G., Saunders, R., Edge, A. 1993, MNRAS, 265, L57
Gunn, J. E., Gott, J. R. 1972, ApJ, 176, 1
Guth, A. H. 1981, Phys. Rev., D23, 347
Haehnelt, M. G., Rees, M. J. 1993, MNRAS, 263, 168
Hanami, H. 1993, ApJ, 415, 42
Hasinger, G. 1992, in The X-ray Background, eds. Barcons, X., Fabian, A. C., Cambridge Univ. Press, Cambridge, 229
Hasinger, G., Burg, R., Giacconi, R., Hartner, G., Schmidt, M., Trümper, J., Zamorani, G. 1993, A&A, 275, 1
Hatsukade, I. 1989, Osaka University PhD thesis (unpublished)
Hattori, M., et al. 1997, Nature, 388, 146
Henry, J. P. 1997, ApJ, 489, L1
Henry, J. P., Arnaud, K. A. 1991, ApJ, 372, 410
Henry, J. P., Jiao, L., Gioia, I. M. 1994, ApJ, 432, 49
Hockney, R. W., Eastwood, J. W. 1981, Computer Simulation Using Particles., McGraw-Hill, New York

Ikebe, Y. 1995, University of Tokyo PhD thesis (unpublished)
Jing, Y. P., Mo, H. J., Börner, G., Fang, L. Z. 1994, A&A, 284, 703
Jing, Y. P., Fang, L. Z. 1994, ApJ, 432, 438
Jones, C., Forman, W. 1984, ApJ, 276, 38
Jones,L.R., Scharf,C.A., Ebeling,H., Perlman,E., Wegner,G., Malkan,M., Horner, D. 1998, ApJ, 495, 100
Kaiser, N. 1986, MNRAS, 222, 323
Kaiser, N. 1991, ApJ, 383, 104
Kang, H., Cen, R., Ostriker, J. P., Ryu, D. 1994, ApJ, 428, 1
Kauffmann, G., White, S. D. M. 1993, MNRAS, 261, 921
Kitayama, T., Suto, Y. 1996a, MNRAS 280, 638
Kitayama, T., Suto, Y. 1996b, ApJ 469, 480
Kitayama, T., Suto, Y. 1997, ApJ 490, 557
Kitayama, T., Sasaki, S., Suto, Y. 1998, PASJ, 50, 1
Kolb, E. W., Turner, M. S., 1990, The Early Universe. Addison-Wesley, Redwood City, Calif.
Lacey, C. G., Cole, S. 1993, MNRAS, 262, 627 (LC)
Lacey, C. G., Cole, S. 1994, MNRAS, 271, 676
Lahav, O., Lilje, P. B., Primack, J. R., Rees, M. J. 1991, MNRAS, 251, 128
Lee, J., Shandarin, S. F. 1997, Bull. American Astron. Soc., 191, 87.07
Liddle, A. R., Lyth, D. H. 1993, Phys. Rep., 231, 1
Lilje, P. B. 1992, ApJ, 386, L33
Lubin, L., Bahcall, N. A. 1993, ApJ, 415, L17
Lyndel-Bell, D. 1967, MNRAS, 136, 101
Markevitch, M., Mushotzky, R., Inoue, H., Yamashita, K., Furuzawa, A., Tawara, Y. 1996, ApJ, 456, 437
Masai, K. 1984, Ap&SS 98, 367
Mathiesen, B., Evrard, A. E. 1998, MNRAS 295, 769
Mo, H. J., Fukugita, M. 1996, ApJ, 467, L9
Monaco, P. 1995, ApJ, 447, 23
Monaco, P. 1998, Fund. Cosm. Phys., 19, 157
Mushotzky, R. F. 1984, Phys. Scr., T7, 157
Mushotzky, R. F., Scharf, C. A. 1997, ApJ, 482, L13
Nakamura, T. T., Suto, Y. 1997, Prog. Theor. Phys., 97, 49
Navarro, J. F., Frenk, C. S., White, S. D. M. 1995, MNRAS, 275, 720
Navarro, J. F., Frenk, C. S., White, S. D. M. 1996, ApJ, 462, 563
Navarro, J. F., Frenk, C. S., White, S. D. M. 1997, ApJ, 490, 463
Nichol, R. C., Holden, B. P., Romer, A. K., Ulmer, M. P., Burke, D. J., Collins, C. A. 1997, ApJ, 481, 644
Oukbir, J., Blanchard, A. 1997, A&A, 317, 10
Oukbir, J., Bartlett, J. G., Blanchard, A. 1997, A&A 320, 365
Padmanabham, T. 1993, Structure Formation in the Universe. Cambridge Univ. Press, Cambridge
Peacock, J. A. 1996, MNRAS, 284, 885
Peacock, J. A., Dodds, S. J. 1994, MNRAS, 267, 1020
Peacock, J. A., Heavens, A. F. 1990, MNRAS, 243, 133

Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press Princeton
Peebles, P. J. E. 1993, Principles of Physical Cosmology. Princeton Univ. Press, Princeton
Pen, U. 1998, ApJ, 498, 60
Piccinotti, G., Mushotzky, R. F., Boldt, E. A., Holt, S. S., Marshall, F. E., Serlemitsos, P. J., Shafer, R. A. 1982, ApJ, 253, 485
Ponman, T. J., Bourner, P. D. J., Ebeling, H., Böhringer, H. 1996, MNRAS, 283, 690
Press, W. H., Schechter, P. 1974, ApJ 187, 425 (PS)
Puget, J.-L., Abergel, A., Bernard, J.-P., Boulanger, F., Burton, W. B., Désert, F.-X., Hartmann, D. 1996, A&A, 308, L5
Rees, M. J., Ostriker, J. P., 1977, MNRAS, 179, 541
Rephaeli, Y. 1995, ARA&A 33, 541
Rephaeli, Y., Yankovitch, D. 1997, ApJ, 481, L55
Rosati, P., Della Ceca, R., Burg R., Norman, C., Giacconi, R. 1995, ApJ, 445, L11
Rosati, P., 1996, Bull. American Astron. Soc., 189, 122.13
Rosati, P., Della Ceca, R., Norman, C., Giacconi, R. 1998, ApJ, 492, L21
Rosati, P., Della Ceca, R. 1998, in preparation
Rybicki, G. B., Lightman, A. P. 1979, Radiative Processes in Astrophysics., Wiley, New York
Sadat,R., Blanchard,A., Oukbir,J. 1998, A&A, 329, 21
Sarazin, C. L. 1988, X-ray Emission from Clusters of Galaxies. Cambridge, Cambridge Univ. Press
Sasaki, S. 1994, PASJ, 46, 427
Sasaki, S., Takahara, F. 1994, Prog. Theor. Phys., 91, 699
Sasaki, S., Takahara, F., Suto, Y. 1993, Prog. Theor. Phys., 90, 85
Sato, K. 1981, MNRAS, 195, 467
Schaeffer, R., Silk, J. 1988, ApJ, 332, 1
Scharf,C.A., Jones,L.R., Ebeling,H., Perlman,E., Malkan,M., Wegner,G. 1997, ApJ, 477, 79
Shimasaku, K. 1997, ApJ, 489, 501
Silk, J. 1977, ApJ, 211, 638
Silk, J., Tarter, J. 1973, ApJ, 183, 387
Smoot, G. F., et al. 1992, ApJ, 396, L1
Spitzer, L., Jr. 1956, Physics of Fully Ionized Gases. Interscience, New York
Squires, G., Kaiser, N., Babul, A., Fahlman, G., Woods, D., Neumann, D.M., Böhringer, H. 1996, ApJ, 461, 572
Suginohara, T., Suto, Y. 1991, PASJ, 43, L17
Suginohara, T. 1994, PASJ, 46, 441
Sugiyama, N. 1995, ApJS, 100, 281
Sunyaev R. A., Zel'dovich Ya. B. 1972, Commts. Astrophys. Space Phys., 4, 173 (SZ)
Suto, Y. 1993, Prog. Theor. Phys., 90, 1173
Suto, Y., Makishima, K., Ishisaki, Y., Ogasaka, Y. 1996, ApJL, 461, L33
Tanvir, N. R., Shanks, T., Ferguson, H. C., Robinson, D. R. T. 1995, Nature, 377, 27
Tsuru, T., Koyama, K., Hughes, J., Arimoto, N., Kii, T., Hattori, M. 1996, in UV and X-ray Spectroscopy of Astrophysical and Laboratory Plasmas, eds. Yamashita, K., Watanabe, T., Universal Academy Press, Tokyo, Japan, 375

Ueda, Y., Takahashi, T., Inoue, H., Tsuru, T., Sakano, M., Ishisaki, Y., Ogasaka, Y., Makishima, K., Yamada, T., Ohta, K., Akiyama, M. 1998, Nature, 391, 866
Viana, P. T. P., Liddle, A. R. 1996, MNRAS 281, 323
Walker T. P., Steigman G., Schramm D. N., Olive K. A., Kang H. -S. 1991, ApJ, 376, 51
Watanabe, T., Matsubara, T., Suto, Y. 1994, ApJ, 432, 17
Weinberg, S. 1972, Gravitation and Cosmology. Wiley, New York
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., Navarro, J. F., Evrard, A. E., Frenk, C. S. 1993, Nature, 366, 429
Yamada, T., Tanaka, I., Aragon-Salamanca, A., Kodama, T., Ohta, K., Arimoto, N. 1997, ApJ, 487, L25
Yoshikawa, K., Itoh, M., Suto, Y. 1998, PASJ, 50, 203