

Observing Galaxy Clusters with eROSITA: Simulations

Diplomarbeit

Johannes Hölzl

Dr. Karl Remeis Sternwarte Bamberg
Astronomisches Institut der Friedrich-Alexander Universität Erlangen-Nürnberg

5. Juli 2011

Betreuer: Prof. Dr. Jörn Wilms

**Friedrich-Alexander-Universität
Erlangen-Nürnberg**



ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS

The cover page shows a Composite of X-ray and optical image of the galaxy cluster Abell 1689 in a distance of 2.3 billion light years. The optical Hubble Space Telescope image is colored yellow, the intracluster medium observed by Chandra's Advanced CCD Imaging Spectrometer (ACIS) is purple.

[X-ray: NASA/CXC/MIT/E.-H Peng et al; Optical: NASA/STScI]

Abstract

The X-ray instrument eROSITA, which is developed by a German collaboration under the direction of the Max-Planck Institut für Extraterrestrische Physik, is one of the two main instruments on board of the Russian spacecraft *Spectrum-Roentgen-Gamma (SRG)*. It will be launched in 2013 to an L₂ orbit. eROSITA will perform an all-sky survey for four years, followed by a three-year period of pointed observations. eROSITA will improve the sensitivity of the *ROSAT* All-Sky Survey (RASS) by a factor of about 30. The main objectives are the observation of galaxy clusters to test cosmological models and the probing of dark energy and dark matter.

The eROSITA survey will be simulated before launch by the *Simulation Software for X-ray Telescopes (SIXT)*. In this thesis, a Monte Carlo code programmed in *Python* is presented, which generates a source catalogue of galaxy clusters for the *SIXT* simulation. The clusters are distributed according to the mass function by Tinker et al. (2008), which is based on *N*-body simulations and describes the distribution of galaxy clusters up to redshifts of $z \approx 2.5$. The Monte Carlo code generates the celestial coordinates, the mass, and the redshift of the galaxy clusters. From this, the X-ray flux is calculated with the mass-luminosity relation by Vikhlinin et al. (2009).

The final output of the simulation is a FITS file. This file is created with a *C* program. Every source entry contains a link to a X-ray image of a galaxy cluster taken with the X-ray observatory *XMM-Newton*. The image is scaled in size according to the redshift of the object.

With this catalogue, the cosmological studies to be made by eROSITA will be simulated.

Zusammenfassung

Das unter Leitung des Max-Planck Instituts für extraterrestrische Physik entwickelte Röntgeninstrument eROSITA ist eines von zwei Hauptinstrumenten an Bord des russischen Satelliten *Spectrum-Roentgen-Gamma (SRG)*. Der Start von *SRG* in einen Orbit um den Lagrange-Punkt L_2 ist für das Jahr 2013 geplant. Die ersten vier Jahre wird eROSITA eine Durchmusterung des kompletten Himmels für Energien bis ~ 10 keV durchführen. Dabei wird die Sensitivität des *ROSAT* all-sky surveys um den Faktor 30 übertroffen. Nach der Durchmusterung folgt eine dreijährige Phase, in der einzelne interessante Objekte beobachtet werden.

Eines der wichtigsten wissenschaftlichen Ziele von eROSITA ist die Bestimmung der kosmologischen Parameter durch die Beobachtung von Galaxienhaufen. Die großräumige Verteilung jener hängt von der Geometrie des Universums, die hauptsächlich von der Dunklen Energie bestimmt wird, ab. Außerdem lassen sich Rückschlüsse auf die primordialen Dichtefluktuationen im Universum ziehen. Entstehung und Entwicklung der Galaxienhaufen werden entscheidend durch Dunkle Materie beeinflusst. Außerdem soll eROSITA Aktive Galaxienkerne und galaktischen Röntgenquellen wie Röntgendoppelsterne oder Supernovaüberreste beobachten.

Vor dem Start von eROSITA wird eine Simulation des Beobachtungsprogramms durchgeführt. Dazu wird die Simulationssoftware für Röntgenteleskope *SIXT* verwendet. In dieser Arbeit wird ein in *Python* programmierter Monte-Carlo Code vorgestellt, der einen Katalog realistischer verteilter Galaxienhaufen erzeugt. Dieser Katalog dient anschließend als Input für die Simulation des Beobachtungsprogramms. Als Massenfunktion für Galaxienhaufen wurde die von Tinker et al. (2008) vorgeschlagene Massenfunktion verwendet. Diese basiert auf Mehrkörpersimulationen und beschreibt die Entwicklung der Massenfunktion bis zu Rotverschiebungen von $z \approx 2.5$. Die Monte-Carlo-Simulation erzeugt Himmelskoordinaten, Masse und Rotverschiebung der Galaxienhaufen. Daraus werden im nächsten Schritt mit der Masse-Leuchtkraft-Beziehung von Vikhlinin et al. (2009) die Leuchtkraft der Galaxienhaufen und der Röntgenfluss berechnet. Die so erzeugte Objektliste wird schließlich von einem C-Programm in eine FITS-Datei geschrieben. Dabei enthält jeder Objekteintrag einen Link zu einem vom Röntgenobservatorium *XMM-Newton* aufgenommenen Bild eines Galaxienhaufens. Dieses Bild wird entsprechend der Rotverschiebung des Galaxienhaufens skaliert.

Mit der so erzeugten FITS-Datei als Input kann anschließend, zusammen mit anderen Quellkatalogen, das Beobachtungsprogramm von eROSITA simuliert werden.



Contents

Contents	5
1. Introduction	7
2. Cosmology	9
2.1. Cosmological Principle	9
2.2. Robertson-Walker Metric	9
2.2.1. Cosmological Time	9
2.2.2. Metric	10
2.3. Friedmann Equations	11
2.4. Cosmological Parameters	13
2.5. Cosmological Distances	13
2.5.1. Redshift	14
2.5.2. Comoving Distance	14
2.5.3. Angular Diameter Distance	15
2.5.4. Luminosity Distance	15
2.5.5. Comoving Volume Element	16
3. Galaxy Clusters	17
3.1. Galaxy clusters in X-rays	17
3.2. X-ray Spectrum of Clusters	19
3.3. Dark Matter in Galaxy Clusters	19
3.4. Cosmology with Clusters	19
3.5. Density Fields	20
3.6. Correlation Function	20
3.7. Halo Identification	21
3.8. Halo Formation	21
3.8.1. Gravitational Collapse	21
3.8.2. Press-Schechter Approach	22
3.9. Mass-Luminosity-Relation	23
3.10. Cluster Mass Function	24
4. eROSITA	27
4.1. Overview	27
4.2. Spectrum-Roentgen-Gamma	27
4.3. Scientific Objectives	27
4.4. The eROSITA-instrument	28
4.4.1. Wolter Telescopes	28
4.4.2. pnCCD Detectors	29
4.5. Orbit	32
4.6. Observing Program	32

5. SIXT Simulation Software	33
6. Generating the Cluster Catalogue	35
6.1. Monte Carlo Methods	35
6.1.1. Inversion Method	36
6.1.2. Rejection Sampling	36
6.2. Python	36
6.3. General Approach	37
6.4. Pipeline	39
6.5. Parameterfile	40
6.6. Script <i>calc_massfct_grid</i>	41
6.6.1. Function <i>massfct</i>	42
6.6.2. Function <i>calc_sigma</i>	42
6.7. Script <i>integrate_shells</i>	43
6.8. Script <i>sample_catalogue</i>	44
6.9. Script <i>convert_observables</i>	45
6.10. C-program <i>simput_converter</i>	46
6.11. Results	46
6.12. Discussion	47
7. Summary and Conclusion	51
8. Outlook	51
9. Danksagungen	53
References	55
A. List of Acronyms	58
B. Typographic Conventions	59

1. Introduction

Almost a century has passed since Albert Einstein¹ presented his General Theory of Relativity to the Prussian Academy of Science, and thereby laid the theoretical foundations for modern cosmology. The bulk of scientific community embraced Einstein's theory, which was regarded as aesthetically beautiful by some, enthusiastically, but it should still take decades until mankind could engage the existential questions of origin, evolution, and future of the universe itself, which were for millenia mainly questions in the realm of philosophy and theology. While in the 1920s theorists were still sitting in their chambers to find solutions for Einstein's equations, some of them as complicated as the equations are simple, the observers were arguing if the diffuse nebulae on the sky were parts of our own milky way or remote systems of billions of stars, like our own.

But the progress gained momentum, and soon the world models allowed by the field equations were known. The problem was to obtain the parameters determining which model applies to our universe. Although there had been stunning progress in this field of research, especially in the second half of the 20th century, and up to now the parameters are known relatively precise, the refinement of the measurements is still an important task of science. Also it is known today that we live in an expanding universe which had its origin in a singularity which was once derisively named Big Bang by Fred Hoyle², who was a fervent advocat of a Steady State universe, there are still enough unanswered questions. For about ten years it is known that the expansion rate of the universe is accelerating (Perlmutter & Schmidt, 2003), driven by a mysterious dark energy whose equation of state is still unknown. Also details of the formation of structures and their origin are still ununderstood, as well as the nature of dark matter which contributes the main part to the gravitating matter in our universe.

One important approach to these questions are measurements of the Cosmic Microwave Background (CMB), on which basic properties of the universe are imprinted. A complementary method is the observation of the distribution of galaxy clusters, which is also influenced by cosmology. While the CMB observations are a domain of radio astronomers, galaxy clusters can be well observed in X-rays.

To derive precise conclusions about cosmology, it is required to observe statistically complete samples of galaxy clusters up to high redshifts. A key mission in this area will be eROSITA, an X-ray telescope whose launch is scheduled for 2013. It will perform a four-year all-sky survey (Cappelluti et al., 2011), with the main objective of obtaining a statistically complete sample of galaxy clusters up to high redshifts to perform cosmology. The eROSITA observation program will be simulated in advance. This thesis presents a Monte-Carlo code for the generation of a catalogue of galaxy clusters as an input for this simulation, based on the knowledge about their spatial distribution obtained by former missions and surveys.

¹Albert Einstein, 1879-1955

²Fred Hoyle, 1915-2001

2. Cosmology

An important goal of observing clusters is to constrain the cosmological parameters. Vice versa, the creation of a realistic mock catalogue of clusters requires some knowledge of the underlying cosmology. In the following paragraphs, the cosmological basics for this thesis are introduced. A large chapter is dedicated to distance measurement in cosmology, which is not an easy task because we live in an expanding universe, which can also be curved (although recent measurements speak against the latter). The accent lies on definition and usage of distances, the complicated task of obtaining distances by observations is not a topic of this thesis. A more complete treatise about cosmology can be found in e.g. Peacock (1999). The following considerations are also mainly oriented on this textbook.

In the next chapter, galaxy clusters are discussed. The most important intrinsic feature of a galaxy cluster is its mass, which determines its luminosity as well as temperature and density profile, i.e. the information obtained by observations. Thus, a relation between mass and luminosity calibrated by observations and a theoretical approach to a cluster mass function presented by Press & Schechter (1974) is discussed. For the simulation described in this thesis, a mass function from Tinker et al. (2008) based on the Press-Schechter formalism but calibrated by observations was used.

2.1. Cosmological Principle

Two of the basic assumptions in cosmology are subsumed as the Cosmological Principle:

1. space is homogenous, that means it looks the same everywhere (no privileged observer – Kopernican principle)
2. space is isotropic, meaning it looks the same in every direction

These principles are only valid on large scales greater ~ 50 Mpc (Press & Schechter, 1974). In our cosmic neighbourhood, for example, the matter is distributed highly inhomogeneously.

Homogeneity does not imply isotropy, but isotropy from every place in the universe implies homogeneity. (Peacock, 1999)

2.2. Robertson-Walker Metric

2.2.1. Cosmological Time

A fundamental observer in an expanding universe is defined as an observer resting in relation to the matter in his vicinity. The peculiar motion of the objects is neglected, such that the only motion results from the expansion of the universe. Such an observer can synchronise a clock with another fundamental observer by agreeing on setting the clock to a certain time when, e.g., the universe is reaching a certain mean density. This time is called Cosmological Time, further denoted as t . (Peacock, 1999)

2.2.2. Metric

The metric of the universe is obtained as a solution of the Einsteinian field equations (Einstein (1916); for a comprehensive reading: (Peacock, 1999, p. 19 ff)). Space is not necessarily Euclidean, but can also be curved. The specific shape depends on the density of the universe compared to the Critical Density (see Sect. 2.3).

In general, a line element can be written as (Peebles, 1993):

$$ds^2 = c^2 dt^2 + g_{\alpha\beta} dx^\alpha dx^\beta = dt^2 - c^2 dl^2 \quad (2.1)$$

where $g_{\alpha\beta}$ is the metric tensor, c the speed of light in vacuum, and dl the proper spatial separation between two events at the Cosmological Time t . The Greek indices α and β denote the spatial coordinates.

For a homogeneous and isotropic universe, the most general metric is the Friedman³-Lemaître⁴-Robertson⁵-Walker⁶ metric (Robertson, 1935), often simply referred to as Robertson-Walker metric. It is an exact solution of the Einsteinian field equations under the symmetry constraints mentioned above.

In our expanding universe, it is useful to describe distances by a comoving coordinate r in a coordinate system which is fixed to the expanding space and hence time-independent. The expansion is characterized by the scale factor $R(t)$.

Alternatively, a dimensionless scale-factor can be defined:

$$a(t) \equiv \frac{R(t)}{R_0} \quad (2.2)$$

where R_0 is the present scale factor. Because $a(t)$ describes the size of the universe at the time t compared to its size today, it is closely related to the cosmological redshift (see Sect. 2.5.1):

$$a = \frac{1}{1+z} \quad (2.3)$$

The line element of the Robertson-Walker metric can be written in the following form:

$$c^2 d\tau^2 = c^2 dt^2 - R^2(t) \left[f^2(r) dr^2 + g^2(r) d\psi^2 \right] \quad (2.4)$$

according to Peacock (1999), where $d\psi$ denotes the transverse part in the spherical polar coordinates:

$$d\psi^2 = d\theta^2 + \sin^2\theta d\phi^2 \quad (2.5)$$

Because of the spherical symmetry resulting from isotropy, it is sufficient to decompose the spherical polar coordinates into a radial and a transverse part. f and g are arbitrary functions of the radial coordinate.

To define the curvature of space k and scale factor $R(t)$, there are two ways frequently used in literature:

³Alexander Alexandrovich Friedman, 1888-1925

⁴Monsignor Georges Henri Joseph Edouard Lemaître, 1894-1966

⁵Howard Percy Robertson, 1903-1961

⁶Arthur Geoffrey Walker, 1909-2001

- k denotes the Gaußian curvature (for a better distinction written as upper case letter K from now on) at the epoch when $R(t) = 1$. In this case, $[K] = \text{length}^{-2}$ and $[r] = \text{length}$. $R(t)$ is dimensionless.
- $k \in \{-1, 0, 1\}$ for an open, flat respectively closed universe. In this case, r is dimensionless and $[R] = \text{length}$. $R(t)$ denotes the radius of curvature at time t .

In general, the scale factor can be chosen arbitrarily. After the renormalization from K to k , R is not arbitrary anymore. The relationship between the Gaussian curvature K and the curvature radius $R(t)$ is given as (Misner et al., 1973, chapter 27):

$$K = \frac{k}{R(t)^2} \quad (2.6)$$

where $k = \frac{K}{|K|}$. From now on, the second convention is used.

It is often convenient to define a function $S_k(r)$:

$$S_k(r) \equiv \begin{cases} \sin(r) & (k = 1) \\ r & (k = 0) \\ \sinh(r) & (k = -1) \end{cases} \quad (2.7)$$

$k = 0$ for a flat universe, which implies an Euclidian geometry, $k = -1$ for an open universe (the angles of a triangle add to less than 180°), and $k = 1$ for a closed universe (the angles of a triangle add to more than 180°). For a derivation, see e.g. Peacock (1999).

A widely used notation of the Robertson-Walker line element is (Peacock, 1999):

$$c^2 d\tau^2 = c^2 dt^2 - R^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\psi^2 \right] \quad (2.8)$$

With Eq. 2.7 and defining r in such a way that the function $f(r) = 1$ in Eq. 2.4, the line element of the Robertson-Walker metric can be written in the following form (hyperspherical coordinates):

$$c^2 d\tau^2 = c^2 dt^2 - R^2(t) \left[dr^2 + S_k^2(r) d\psi^2 \right] \quad (2.9)$$

In Eq. 2.9 and Eq. 2.8, the whole expansion and with this the time-dependence lies in the scale-factor $R(t)$, while the part in the square brackets denotes the comoving coordinates.

2.3. Friedmann Equations

The Friedmann equations, which describe the expansion dynamics of the universe, can be deduced from Einstein's field equations for the Robertson-Walker metric (Misner et al., 1973).

$$\dot{R}^2 = -kc^2 + \frac{8\pi G}{3} \rho R^2 \quad (2.10)$$

$$\ddot{R} = -\frac{4\pi G}{3c^2} R (\rho c^2 + 3p) \quad (2.11)$$

where p is the pressure caused by the content of the universe (Peacock, 1999). Eq. 2.10 shows a direct relationship between curvature k and density of the universe ρ .

By multiplying Eq. 2.10 with R^{-2} , it can alternatively be written as:

$$H^2 - \frac{8\pi G}{3}\rho = -\frac{kc^2}{R^2} \quad (2.12)$$

where H is the Hubble constant describing the expansion rate of the universe:

$$H(t) \equiv \frac{\dot{R}(t)}{R(t)} \quad (2.13)$$

Instead of H , often the dimensionless Hubble parameter h is used:

$$h \equiv \frac{H}{100 \text{ km s}^{-1} \text{ Mpc}^{-1}} \quad (2.14)$$

By setting the right-hand side of Eq. 2.12 to zero, we obtain the Critical Density ρ_c for which the universe is flat:

$$\rho_c = \frac{3H^2}{8\pi G} \quad (2.15)$$

It is convenient to write the total density as a fraction of the Critical Density:

$$\Omega = \frac{\rho}{\rho_c} \quad (2.16)$$

There are several components contributing to Ω : matter, radiation and the vacuum energy.

The cosmological constant Λ , which was originally introduced by Einstein as a constant to obtain a static universe (Einstein, 1917) and discarded after the insight of the instability of the static solution and the discovery of the expansion by Hubble⁷, is interpreted as the vacuum energy today. The vacuum energy density in terms of the cosmological constant is given as (Peacock, 1999):

$$\rho_\Lambda = \frac{\Lambda c^2}{8\pi G} \quad (2.17)$$

So the density parameters are:

$$\Omega_m \equiv \frac{\rho_m}{\rho_c} \quad \text{matter} \quad (2.18)$$

$$\Omega_r \equiv \frac{\rho_r}{\rho_c} \quad \text{radiation} \quad (2.19)$$

$$\Omega_\Lambda \equiv \frac{\rho_\Lambda}{\rho_c} \quad \text{vacuum energy} \quad (2.20)$$

The total density is the sum of the components: $\Omega = \Omega_m + \Omega_r + \Omega_\Lambda$

The present scale factor, which corresponds to the curvature radius, can be obtained by solving Eq. 2.12 for R and using Eq. 2.15 (Peacock, 1999). This is the so-called curvature length:

$$R_0 = \frac{c}{H_0} \sqrt{\frac{k}{\Omega - 1}} \quad (2.21)$$

Comparing the equation above with Eq. 2.6 one can see that $|\Omega - 1|$ equals the Gaussian curvature in units of inverse squared Hubble length' D_H^{-2} , where $D_H = \frac{c}{H_0}$.

⁷Edwin Powell Hubble, 1889-1953

In an expanding universe, the matter density decreases while the matter is dispensed into a larger volume, while, in addition, the radiation is redshifted. It is assumed that the vacuum energy density stays constant, even though there were suggestions involving a non-constant vacuum energy (Solà, 2011). Therefore, the evolution of the densities is given as:

$$\rho_m(a) \propto a^{-3} \quad (2.22)$$

$$\rho_r(a) \propto a^{-4} \quad (2.23)$$

$$\rho_\Lambda = \text{const.} \quad (2.24)$$

Hence, with Eq. 2.15 one obtains the development of the density with redshift (Peacock, 1999):

$$\frac{8\pi G\rho}{3} = H_0^2 (\Omega_\Lambda + \Omega_m a^{-3} + \Omega_r a^{-4}) \quad (2.25)$$

The evolution of the density parameters is obtained by using the definition $\Omega = \frac{\rho}{\rho_c}$ and Eq. 2.15:

$$\Omega_m(z) = \Omega_{m,0} \frac{(1+z)^3}{E^2(z)} \quad (2.26)$$

$$\Omega_r(z) = \Omega_{r,0} \frac{(1+z)^4}{E^2(z)} \quad (2.27)$$

$$\Omega_\Lambda(z) = \Omega_\Lambda \frac{1}{E^2(z)} \quad (2.28)$$

where $E(z) = \frac{H(z)}{H_0}$.

To get the dynamic of the Hubble constant with redshift, one uses the Friedmann equation 2.12 together with Eq. 2.21 and the above equation (Peacock, 1999):

$$H^2(z) = H_0^2 \left[\Omega_\Lambda + \Omega_m (1+z)^3 + \Omega_r (1+z)^4 - (\Omega - 1)(1+z)^2 \right] \quad (2.29)$$

2.4. Cosmological Parameters

Determining the cosmological parameters is the main task of cosmology. One of the best techniques for this task is the measurement of anisotropies of the CMB. Such precision measurements were first performed by *COsmic Background Explorer (COBE)* (Mather et al., 1994), followed by the *Wilkinson⁸ Microwave Anisotropy Probe (WMAP)* (Bennett et al., 2003). The most recent parameters from *WMAP* are the parameters after seven years of observations (Komatsu et al., 2011), often referred to as *WMAP7-cosmology* (see table 2.1). From this results the mean matter density today, were the critical density is given in Eq. 2.15:

$$\rho_0 = \Omega_m \rho_c = 2.7752 \cdot 10^{11} \Omega_m h_{100}^2 \text{ M}_\odot \text{ Mpc}^{-3} = 3.719 \cdot 10^{10} \text{ M}_\odot \text{ Mpc}^{-3}$$

2.5. Cosmological Distances

The following considerations are mainly oriented on Peacock (1999).

⁸David Todd Wilkinson, 1935-2002

Table 2.1: The cosmological parameters as obtained from *WMAP* after seven years of observations (Komatsu et al., 2011). n_s is the initial spectral index and σ_8 the amplitude of the initial density fluctuations.

Ω_m	0.2707
Ω_b	0.0451
h_{100}	0.703
σ_8	0.809
n_s	0.966

2.5.1. Redshift

Up to now, the scale factor $R(t)$ respectively $a(t)$ was used to describe the universe at a specific time. The most important observable when measuring cosmological distances is the redshift. The redshift is defined as

$$z \equiv \frac{\lambda_{\text{observed}} - \lambda_{\text{emitted}}}{\lambda_{\text{emitted}}} = \frac{\Delta\lambda}{\lambda} \quad (2.30)$$

The cosmological redshift does not result from the Doppler effect, but is caused by the expansion of the universe. When light travels through the expanding space, the wavelength λ expands together with space. Hence, the redshift is simple given by

$$z = \frac{R_{\text{observed}}}{R_{\text{emitted}}} - 1 = \frac{1}{a} - 1 \quad (2.31)$$

where a is the normalized scale factor (Peacock, 1999). In the following, the relation between distances and redshift is discussed.

2.5.2. Comoving Distance

The comoving distance (which is not the same as the comoving coordinate) is measured along the line of sight to an object. The infinitesimal way element of the comoving distance as a function of redshift can be calculated by the equation of motion for a photon with $R = \frac{R_0}{1+z}$ (Peacock, 1999):

$$\begin{aligned} RdR &= cdt = \frac{cdR}{\dot{R}} = \frac{cdR}{RH(z)} \\ &= \frac{cR_0 dz}{RH(z)(1+z)^2} \\ &= \frac{c}{H(z)(1+z)} dz \end{aligned} \quad (2.32)$$

So the infinitesimal radial line element is

$$R_0 dr = \frac{c}{H(z)} dz \quad (2.33)$$

and the comoving distance D_C is the integral of the above equation:

$$D_C(z) = \int_0^z \frac{c}{H(z')} dz' \quad (2.34)$$

The development of the Hubble constant $H(z)$ depends on the density parameters and is given by Eq. 2.29:

$$R_0 dr = \frac{c}{H_0} \left[\Omega_\Lambda + \Omega_m (1+z)^3 + \Omega_r (1+z)^4 - (\Omega - 1)(1+z)^2 \right]^{-\frac{1}{2}} dz \quad (2.35)$$

For practical purposes, $\Omega_r \approx 0$ for $z \leq 1000$ (Peacock, 1999, p. 84). For a matter-dominated universe with $\Omega_\Lambda = 0$ and therefore $\Omega = \Omega_m$, there exists an analytical solution to the integral, which is called Mattig's formula (Mattig, 1958):

$$R_0 S_k(r) = \frac{2c}{H_0} \frac{\Omega z + (\Omega - 2) \left[\sqrt{1 + \Omega z} - 1 \right]}{\Omega^2 (1+z)} \quad (2.36)$$

The comoving distance as a function of redshift is shown in Fig. 2.1.

2.5.3. Angular Diameter Distance

The angular diameter distance D_A is related to the apparent angular size of an object (Peacock, 1999).

$$D_A(z) = \frac{R_0 S_k(r)}{1+z} \quad (2.37)$$

And therefore for a flat universe:

$$D_A(z) = \frac{D_C}{1+z} \quad (2.38)$$

Using the angular diameter distance, an object of the size d is seen with an angular size:

$$\theta = \frac{d}{D_A} \quad (2.39)$$

For the *WMAP7*-cosmology, the angular distance reaches a maximum at $z \approx 2$, for higher redshifts it is decreasing. Therefore, high-redshifted objects appear larger objects closer to us. The angular diameter distance as a function of redshift is shown in Fig. 2.1.

2.5.4. Luminosity Distance

For the calculation of the flux received from an distant object analogous to the common $\frac{1}{r^2}$ -law, the luminosity distance D_L is defined as (Peacock, 1999)

$$D_L(z) = R_0 S_k(r) (1+z) = D_A (1+z)^2 \quad (2.40)$$

Which is for a flat universe (Peacock, 1999):

$$D_L(z) = D_C (1+z) \quad (2.41)$$

Therefore for the flux from an object with luminosity L at distance d the following equation holds:

$$F = \frac{L}{4\pi D_L^2} \quad (2.42)$$

There is one caveat concerning Eq. 2.42: It can only be applied to the bolometric flux, which is very difficult to obtain in reality. To calculate the flux in a specific band, corrections have to be applied. There is also a difference between photon flux and energy, which has to be taken into account if dealing with limited energy bands (Peacock, 1999).

The D_L increases monotonously with redshift z . The luminosity distance as a function of redshift is shown in Fig. 2.1.

2.5.5. Comoving Volume Element

The infinitesimal comoving volume element describes the volume per steradian as a function of the redshift (Peacock, 1999).

$$dV_C = [R_0 S_k(r)]^2 \cdot R_0 dr d\Omega \quad (2.43)$$

for a flat universe using Eq. 2.33:

$$dV_C = D_C^2 \cdot R_0 dr = D_C^2 \frac{c}{H(z)} dz d\Omega \quad (2.44)$$

In an isotropic universe, the integration over the solid angle simply contributes a factor 4π , so the equation above becomes (Peacock, 1999):

$$dV_C = 4\pi D_C^2 \frac{c}{H(z)} dz \quad (2.45)$$

In this volume element, a number density stays constant in an expanding universe (Peacock, 1999). Since Eq. 2.43 contains the $S_k(r)$ -term, it is possible to determine the geometry of space by observing a population of objects with known number density at different redshifts.

Because the comoving volume element is related to the angular diameter distance, it reaches a peak at $z \approx 2$ (for a *WMAP7* cosmology) and decreases for higher redshifts. The comoving volume element as a function of redshift is shown in Fig. 2.1.

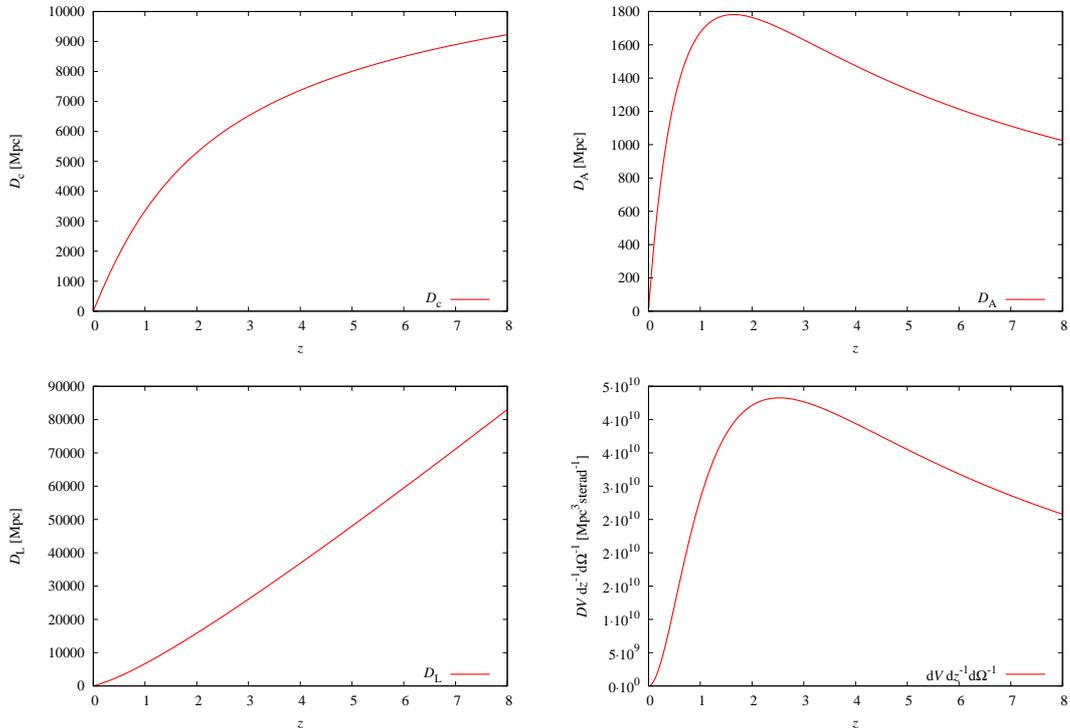


Figure 2.1: Comoving Distance (top left), Angular Diameter Distance (top right), Luminosity Distance (bottom left) and Comoving Volume Element (bottom right) for *WMAP7*-cosmology)

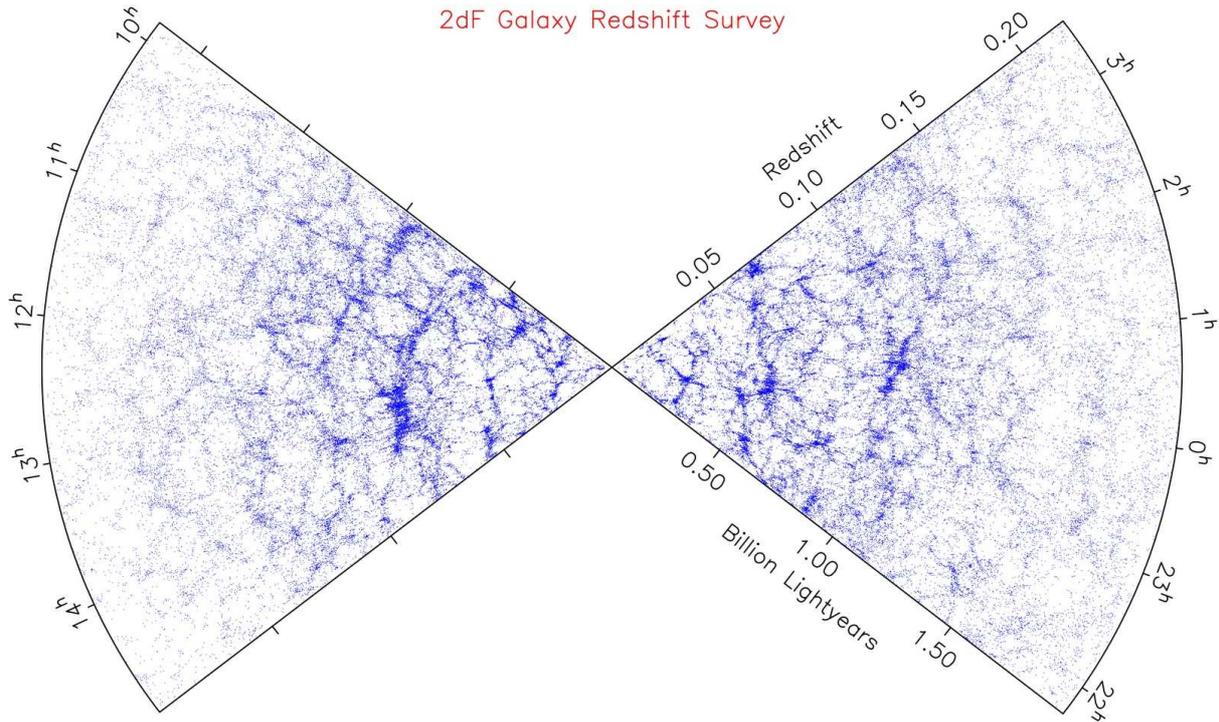


Figure 3.1: Spatial distribution of galaxies from the 2dF survey (Colless, 1999). On sufficiently large scales, the matter is distributed homogeneously, but on smaller scales structures are conspicuous. Between supermassive clusters, which are connected by long chains of galaxies, the so-called filaments, are almost empty bubbles, the voids.

3. Galaxy Clusters

After having dealt with the universe as whole in Sect. 2, the focus is now shifted to the largest coherent structures in the universe, the galaxy clusters. From the point of homogeneity, one does not expect to see any distinct features at all. But since homogeneity applies only on scales larger ~ 50 Mpc (Press & Schechter, 1974), there are lots of structures, from stars up to galaxy clusters. The galaxy distribution on large scales can be revealed by deep surveys like the 2dF survey (Colless, 1999), from which the galaxy distribution shown in Fig. 3.1 was obtained. The largest gravitational bound, virialized structures in the universe found by such surveys are the galaxy clusters. Typically, they contain some hundred galaxies (Sarazin, 1986). With a luminosity in the range 10^{43} – 10^{46} erg s $^{-1}$ (Trümper & Hasinger, 2008), they are the brightest X-ray sources next to quasars (Sarazin, 1986). After an introduction to X-ray observations and especially mass determination of galaxy clusters in Sect. 3.1 follows a discussion of the importance of clusters for cosmology in Sect. 3.4. Then, in Sect. 3.5 the description of inhomogeneous matter distributions as density fields is introduced. This is needed for the understanding of cluster identification. Afterwards the formation and mass function of clusters are discussed in Sect. 3.8-3.10.

3.1. Galaxy clusters in X-rays

The following description is mainly based on Trümper & Hasinger (2008), chap. 23. While in



Figure 3.2: Composite of X-ray and optical image of the galaxy cluster Abell 1689 at a distance of 2.3 billion light years. The optical Hubble Space Telescope image is colored yellow, the intracluster medium observed by Chandra's Advanced CCD Imaging Spectrometer (ACIS) is purple. [X-ray: NASA/CXC/MIT/E.-H Peng et al; Optical: NASA/STScI]

the optical the galaxies forming the cluster are seen, X-ray observations reveal a diffuse emission over the whole cluster (Fig. 3.2). This diffuse emission extends on scales of about 1 Mpc (Mo et al., 2010). It is caused by hot intracluster gas, which is also called intra cluster medium (ICM), with temperatures of several ten million degrees, which corresponds to X-ray energies of $kT \approx 2 - 15$ keV (Trümper & Hasinger, 2008). The force forming this intracluster gas is mainly gravitation; therefore measurement of the gas distribution allows inferences about the gravitational potential of the cluster. If one approximates that the ICM is in hydrostatic equilibrium, the ICM in the cluster potential can be described as (Trümper & Hasinger, 2008)

$$\frac{1}{\rho} \nabla P = \frac{GM(r)}{r^2} \quad (3.1)$$

where ρ is the density of the ICM, P is the pressure, G is the gravitational constant, and M is the mass enclosed in a sphere with radius r . By adding an assumption about the cluster geometry, which can in most cases be presumed as spherical, this equation can be reformulated (Trümper & Hasinger, 2008):

$$M(r) = -\frac{kT(r)}{G\mu m_p} r \left(\frac{d \log \rho}{d \log r} + \frac{d \log T_X}{d \log r} \right) \quad (3.2)$$

This equation gives the important insight that the mass enclosed by a sphere of radius r depends on the gas density and the temperature at this radius. Hence a measurement of the temperature

and density profile is needed to calculate the mass profile and from this the total mass of a cluster.

3.2. X-ray Spectrum of Clusters

The X-ray emission of the intracluster medium is mainly due to thermal bremsstrahlung, with contributions from line emission and recombination radiation (Trümper & Hasinger, 2008). According to Sarazin (1986), Eq. 5.11, the emissivity ϵ of thermal bremsstrahlung is given as:

$$\epsilon^{\text{ff}} = \frac{dL}{dVdv} = \frac{2^5 \pi e^6}{3m_e c^3} \left[\frac{2\pi}{3m_e k} \right]^{\frac{1}{2}} Z^2 n_e n_i g_{\text{ff}}(Z, T_g, \nu) T_g^{-\frac{1}{2}} \exp\left(-\frac{h\nu}{kT_g}\right) \quad (3.3)$$

where n_e is the electron density, n_i the ion density and g_{ff} a Gaunt factor correcting for quantum mechanical effects. If the Gaunt factor is assumed as constant, the spectrum is an exponential function of the energy. Because the gas temperature T_g is a parameter in the exponent, the spectral shape is mainly determined by the temperature. The chemical composition influences the spectral shape, too. Under the assumption of an ion density proportional to the electron density, the normalization of the spectrum depends on the squared gas density. There is also line emission observed in cluster spectra, especially from iron (Sarazin, 1986). This leads to the conclusion that the ICM, or parts of it, have already been processed in stars.

Because there are shells with different density and temperature along the line of sight, the observed spectrum is obtained as a convolution of three-dimensional density and temperature profile. The real observable is a so-called ‘emission measure weighted temperature’. Therefore, the mass profile can be obtained by measuring the temperature- and mass profile and deprojecting it along the line of sight, which is not a trivial task (Trümper & Hasinger, 2008).

3.3. Dark Matter in Galaxy Clusters

If the density and temperature profiles are calculated as described above, the total gravitating mass of the galaxy cluster can be obtained. This includes the non-radiating dark matter, therefore indirect dark matter observations are possible by X-ray observations of galaxy clusters. According to Trümper & Hasinger (2008), mass estimates obtained by cluster observations suggest a composition of about 87 % dark matter, while 11 % of the total mass are contributed by the ICM and only 2 % are found in the galaxies. Because the clusters consist mainly of a halo of dark matter, often the term dark matter halo or simply halo is used if referring to clusters.

3.4. Cosmology with Clusters

While the observation of single clusters provides us with information about dark matter in the cluster itself, the cosmological parameters can be constrained by means of accurate measurement of the the large-scale structures in the universe. Since galaxy clusters form from overdense regions, they trace the overall matter distribution, which constitutes the large scale structures (Trümper & Hasinger, 2008). Therefore a statistically complete sample of galaxy clusters provides us with complementary information about the cosmological parameters (Predehl et al., 2006). The mass function of clusters, which describes the number density as a function of

mass and redshift, depends on the density parameter Ω_m and the amplitude of the primordial power spectrum σ_8 . The evolution of the mass function as well as of amplitude and shape of the power spectrum $P(k)$, are strongly influenced by dark matter and dark energy. The baryonic acoustic oscillations, which allow the measurement of the curvature of space at different epochs (Predehl et al., 2006), are imprinted on the large-scale structure.

Galaxy clusters can be used as standard candles (Predehl et al., 2006), thus a high-precision measurement of their spatial distribution is possible. The constraints on the cosmological parameters obtained by cluster observations are complementary to other methods like measurements of the Cosmic Microwave Background, and degeneracies between parameters can be broken by combining observations (Mo et al., 2010).

3.5. Density Fields

Generally, the non-uniformity of the matter distribution can be described at each position \mathbf{x} as over- oder underdensity δ with respect to the mean density of the universe $\bar{\rho}$ (Peacock, 1999):

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}} \quad (3.4)$$

After recombination, these density perturbations increase linearly with time: $\delta(\mathbf{x}, t) \propto D(t)$, where $D(t)$ is the linear growth factor. Closely related to the density field is the power spectrum of the density fluctuations $P(k)$.

If the initial perturbation power spectrum is known, its development in time can be calculated. Before recombination, the shape and amplitude of the power spectrum change. This evolution is described by the linear transfer function $T(k)$. After recombination, the power spectrum at time t can be written as (Mo et al., 2010):

$$P(k, t) = P_i(k) T^2(k) D^2(t) \quad (3.5)$$

where $P_i(k)$ is the initial power spectrum. In inflationary models, the initial density perturbations arise from quantum fluctuations of the inflation scalar field, thus the power spectrum $P(k)$ of the perturbations is Gaussian (Mo et al., 2010).

3.6. Correlation Function

The characteristic scales of the clustering can be obtained by the autocorrelation function ξ of the density field, which is defined as:

$$\xi(\mathbf{r}) \equiv \langle \delta(\mathbf{x}) \delta(\mathbf{x} + \mathbf{r}) \rangle \quad (3.6)$$

where the angle brackets stand for the averaging over the normalization volume V (Peacock, 1999). Since r is independent of its direction due to isotropy, the correlation function depends only on the distance between objects. It is also shown by Peacock (1999) that the correlation function is the Fourier transform of the power spectrum $P(k)$. Davis & Peebles (1983) give an empirical correlation function for galaxies described by a power law:

$$\xi(r) = \left(\frac{r_0}{r} \right)^\gamma \quad (3.7)$$

with $\gamma = 1.77$ and $r_0 = 5.4 \pm 0.3 h^{-1} \text{ Mpc}$.

3.7. Halo Identification

Halos are identified as overdense regions relative to the cosmic density field. These overdensities can be identified by smoothing the density field on the appropriate scale with a filter function W of a characteristic radius R (Mo et al., 2010):

$$\delta(\mathbf{x}, R) \equiv \int \delta(\mathbf{x}') W(\mathbf{x} + \mathbf{x}', R) d^2\mathbf{x}' \quad (3.8)$$

For a top-hat filter function, an overdensity δ containing the mass M in a sphere with radius R is defined as

$$\delta = \frac{M(r < R)}{\frac{4}{3}\pi R^3 \bar{\rho}} \quad (3.9)$$

Halos can be identified by setting a overdensity threshold Δ and expanding the radius R around peaks in the overdensity field until the threshold is reached (Tinker et al., 2008):

$$\delta = \Delta = \frac{M_\Delta}{\frac{4}{3}\pi R_\Delta^3 \bar{\rho}} \quad (3.10)$$

The halo identified this way has the mass M_Δ enclosed by the radius R_Δ .

3.8. Halo Formation

In this section, the formation of the dark matter halos, which contribute a significant part to the total cluster mass (Trümper & Hasinger, 2008) is discussed.

3.8.1. Gravitational Collapse

The established theory explains the halo formation by collapse due to gravitational instability. Initial density perturbations grow linearly until a Critical Density is reached. At this point, gravitation is strong enough to cause a collapse decoupled from the expansion and followed by the virialization of the overdense region. The overdense region increases its matter content by accreting material from the underdense regions around (Mo et al., 2010). Different halos can merge to larger halos, such that successively larger virialized structures are formed.

In a flat universe with cosmological constant $\Lambda > 0$, the critical overdensity δ_c for a collapse is given by Mo et al. (2010)

$$\delta_c(t_{\text{col}}) = \frac{3}{5} \left(\frac{3\pi}{2} \right)^{\frac{2}{3}} - [\Omega(t_{\text{col}})]^{0.0055} \approx 1.686 [\Omega(t_{\text{col}})]^{0.0055} \quad (3.11)$$

where $\Omega(t_{\text{col}})$ is the density parameter at the time of the collapse. For a detailed discussion of the evolution of the density perturbations see (e.g. Mo et al., 2010, chapter 4).

3.8.2. Press-Schechter Approach

The formation of structures from a gas of self-gravitating particles is a problem which requires N -body simulations. Such calculations were done, e.g. by Springel et al. (2005) with $2,160^3$ particles. Because of the limitations in available calculation time, such simulations can only be done for a limited number of particles in a limited volume. To develop a global theory not restricted by these limitations, other approaches are demanded. One of the first theoretical approaches to structure formation and a cluster mass function was done by Press & Schechter (1974). This formalism is also known as Press⁹-Schechter¹⁰ formalism. Since the derivation of the Press-Schechter mass function is not mathematically rigorous, it has to be tested with numerical N -body-simulations and observations (Press & Schechter, 1974; Mo et al., 2010).

The Press-Schechter formalism (PS) presents a method to partition a continuous linear density field in disjoint regions which form the collapsed objects (Mo et al., 2010). Most generally, it describes a collisionless gas of self-gravitating particles in an expanding universe.

Press & Schechter (1974) suggest a successive formation of large structures due to nonlinear interaction of smaller particles. If the particle lumps are sufficiently bound, they are identified as single particles. The randomness in position of these particles itself acts as initial perturbation for the condensation on larger scales. It is not necessary to have additional initial perturbations. Press & Schechter (1974) show that it can be assumed that the spectrum and statistical distribution of the initial perturbations have only a very weak influence on the spectrum at late times. One of the important insights of the PS approach is that a self-similar state is reached where the functional form of halo mass distributions is reproduced at larger scales.

The Press-Schechter formalism can be applied to an expanding universe, where the characteristic particle density n_* is large enough that the mean distance between the particles $l \approx n_*^{-1/3}$ is much smaller than the light horizon L_h , so that the dynamics can be treated Newtonian (as long as the particles do not collapse to a relativistic object like a black hole, Press & Schechter (1974)). There are two processes acting against each other: The expansion of the universe dragging the particles away from each other and gravitation which attracts the particles.

The parameters characterising the behaviour of the particles are their peculiar velocity v relative to the Hubble flow (analogous to the gas temperature), the characteristic density introduced above, the characteristic particle mass m_* , and the Hubble parameter h describing the expansion of the universe (Eq. 2.14). These parameters can be combined to two dimensionless quantities:

$$q = \frac{4}{3}\pi n_* m_* \frac{G}{h^2} \quad (3.12)$$

$$N_J = n_* \left(\frac{v}{h}\right)^3 \quad (3.13)$$

q is a deceleration parameter which describes the ability of the expansion to impede condensation, while N_J is related to the particle number inside a Jeans mass and therefore quantifies the tendency to local condensation. For systems with similar q and N_J a similar behaviour is expected regardless the scale of the parameters. Thus, for q and N_J constant in time a self-similar condensation can be expected.

⁹William H. Press, *1947

¹⁰Paul Schechter, *1948

The only parameter in the Press-Schechter formalism is the characteristic mass of the initial condensations. Press & Schechter (1974) derive the functional form of the halo mass function:

$$n_R(M) \propto M^{-1-\alpha} \exp \left[-\text{const} \cdot \left(\frac{M^{1-\alpha}}{R} \right)^2 \right] \quad (3.14)$$

where $\frac{1}{3} \leq \alpha \leq \frac{1}{2}$ and R is the scale factor. Mo et al. (2010) give the Press-Schechter mass function as

$$n(M, t) dM = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M^2} \frac{\delta_c}{\sigma} \exp \left(-\frac{\delta_c^2}{2\sigma^2} \right) \left| \frac{d \ln \sigma}{d \ln M} \right| dM \quad (3.15)$$

Time enters only in $\delta_c(t)$, while the mass enters in $\sigma(M)$ and its derivative as well as in the M^{-2} -factor.

Up to a characteristic mass, the distribution varies as a power law, for higher masses it decreases exponentially. A significant number of clusters exists up to $\sigma(M) \gtrsim \delta_c(t)$, or a corresponding mass $M \lesssim M^*$, where M^* is a time-dependent characteristic mass:

$$\sigma(M^*) = \delta_c(t) = \frac{\delta_c}{D(t)} \quad (3.16)$$

where $D(t)$ is the linear growth factor discussed in Sect. 3.5. Press & Schechter (1974) show that the self-similarity is obtained for perturbations with maximal variance, i.e. an Gaussian distribution, as well as for the case with minimal variance, i.e. each particle belonging to a regular lattice site in the beginning, and all the cases between. The only influence of this initial distribution lies in the dependency of a ‘typical mass’ as a function of the expansion scale, that means the mass value around which the cluster masses are concentrated (Press & Schechter, 1974).

To apply the Press-Schechter formalism to galaxy cluster formation, it is not compulsory to assume that all structures were formed by this process, starting with the smallest possible particles as seeds. It is also possible that large objects such as galaxies were formed by other processes. Due to the self-similarity of the function, the result is the same after sufficient time (Press & Schechter, 1974).

3.9. Mass-Luminosity-Relation

The mass of an galaxy cluster can be obtained by measuring its density and temperature profile, as described above. The relation between mass and total X-ray luminosity can be calibrated by fitting a sample which contains a sufficient number of clusters to a model.

Such a mass-luminosity-relation was published by Vikhlinin et al. (2009). Its is based on a mass-limited sample of clusters at $z = 0.05$ from the *ROSAT* PSPC survey and a subsample of 36 clusters at $z = 0.35 - 0.9$ from the *ROSAT* 400 d survey with a mean redshift of $z = 0.5$. The subsample was chosen so that it is quasi-mass-limited (Vikhlinin et al., 2009). The mass- and temperature profiles of the clusters were obtained by *Chandra* observations.

Vikhlinin et al. (2009) define the masses in peaks with a overdensity Δ relative to the Critical Density at redshift z :

$$M_\Delta = M(r \leq R_\Delta) = \Delta \rho_c \frac{4}{3} R_\Delta^3 \pi \quad (3.17)$$

The mass-luminosity-relation is obtained by fitting the observations to a model which is described in Vikhlinin et al. (2009). For the overdensity a value of $\Delta = 500$ was chosen. The resulting relation is:

$$\ln L_X = (47.392 \pm 0.085) + (1.61 \pm 0.14) \ln M_{500} + (1.850 \pm 0.42) \ln E(z) - 0.39 \ln (h/0.72) \pm (0.396 \pm 0.039) \quad (3.18)$$

The last term on the right-hand site describes the scatter in the observations for a fixed M .

3.10. Cluster Mass Function

As mentioned above, the Press-Schechter formalism has to be tested on numerical simulations and observations. A function based on the PS mass function describing halo masses up to redshifts of $z \lesssim 2.5$ was presented by Tinker et al. (2008). The function was calibrated by N -body simulations calculated for volumes up to $1280 h^{-1}$ Mpc edge length. The simulations were performed with the three different codes GADGET2 (Springel et al., 2005), the hashed oct-tree (HOT) code (Warren & Salmon, 1993), and the Adaptive Refinement Technique (ART) (Kravtsov et al., 1997). The mass function is valid for halo masses ranging from $10^{11} h^{-1} M_\odot$ up to $10^{15} h^{-1} M_\odot$ (Tinker et al., 2008).

$$\frac{dn}{dM} = f(\sigma) \frac{\bar{\rho}_m}{M} \frac{d \ln \sigma^{-1}}{dM} \quad (3.19)$$

where $n = \frac{dN}{dV}$ is the number density of halos, $\bar{\rho}_m$ is the mean matter density of the universe and σ is the root mean square of the linear matter power spectrum at redshift z .

The function $f(\sigma)$ is parameterized as

$$f(\sigma) = A \left[\left(\frac{\sigma}{b} \right)^{-a} + 1 \right] e^{-\frac{c}{\sigma^2}} \quad (3.20)$$

In Tinker et al. (2008) values for A , a , b and c as well as their redshift evolution are given, which were fitted by simulations. The root mean square of the linear matter power spectrum is, based on the equation for $z = 0$ given by Reiprich & Böhringer (2002), where the linear growth factor in the numerator was inserted to account for the redshift evolution of the power spectrum

$$\sigma^2(M, z) = \sigma_8^2 \left[\frac{\int_0^\infty dk k^{2+n_s} D^2(z) T(k)^2 \hat{W}(kR(M))^2}{\int_0^\infty dk k^{2+n_s} T(k)^2 \hat{W}(k8h_{100}^{-1} \text{Mpc})^2} \right] \quad (3.21)$$

where $\hat{W}(kR(M))$ is the Fourier transform of the spherical top-hat function, which smoothes the power spectrum on a scale R (Peacock, 1999):

$$\hat{W}(kR(M)) = \frac{3}{(kR(z, M))^3} [\sin(kR(z, M)) - kR(z, M) \cos(kR(z, M))] \quad (3.22)$$

R is the radius of the sphere enclosing the overdensity:

$$R = \left(\frac{3M}{4\pi\rho_0} \right)^{\frac{1}{3}} \quad (3.23)$$

The shape of the power spectrum today is obtained by multiplying the primordial power spectrum with the transfer function $T(k)$. Here, the fitting function by Bardeen et al. (1986) for a cold dark matter model was used:

$$T(k) = \frac{\ln(1 + 2.34q(k))}{2.34q(k)} \left(1 + 3.89q(k) + (16.1q(k))^2 + (5.46q(k))^3 + (6.71q(k))^4\right)^{-0.25} \quad (3.24)$$

with $q(k) = \frac{k}{\Gamma \cdot h_{100}}$, where Γ is the shape parameter:

$$\Gamma = \Omega_m h_{100} \left(\frac{2.7 \text{ K}}{T_0}\right)^2 \exp\left(-\Omega_b - \sqrt{\frac{h_{100} \Omega_b}{0.5 \Omega_m}}\right) \quad (3.25)$$

The form of the mass function from Tinker et al. (2008) is very similar to the Press-Schechter mass function (Eq. 3.15). The exponential from the PS function is included in the fit function $f(\sigma)$, and taking the absolute of derivative in the PS-function is avoided by the -1 exponent of the σ , which guarantees a positive derivative because σ is decreasing with increasing mass.

An approximation formular for the linear growth factor, which can be used for all world models (Mo et al., 2010), is given by Carroll et al. (1992):

$$D(z) = \frac{g(z)}{1+z} \quad (3.26)$$

where $g(z)$ is given as:

$$g(z) \approx \frac{5}{2} \Omega_m(z) \left\{ \Omega_m \frac{4}{7}(z) - \Omega_\Lambda(z) + \left[1 + \frac{\Omega_m(z)}{2}\right] \left[1 + \frac{\Omega_\Lambda(z)}{70}\right] \right\} \quad (3.27)$$

According to Tinker et al. (2008), the mass function can be used for redshifts up to $z \leq 2$ and for a halo mass in the range $10^{11} h^{-1} \leq 10^{15} h^{-1} M_\odot$.

The mass function from Tinker et al. (2008) is given in terms of overdensities with respect to the mean density of the universe, while the mass-luminosity-relation from Vikhlinin et al. (2009) refers to overdensities relative with respect to the Critical Density. Hence, the overdensity to the critical mass Δ_c has to be converted to the corresponding mean-density overdensity Δ :

$$\begin{aligned} \Delta(z) &= \frac{\Delta_c \rho_c(z)}{\bar{\rho}_m(z)} \\ &= \Delta_c \frac{\rho_c(z)}{\rho_c(z) \Omega_m(z)} \\ &\stackrel{\text{Eq. 2.26}}{=} \Delta_c \frac{E(z)^2}{\Omega_{m,0} (1+z)^3} \\ &= \Delta_c \frac{\Omega_\Lambda + \Omega_{m,0} (1+z)^3}{\Omega_{m,0} (1+z)^3} \\ &= \Delta_c \left[\frac{\Omega_\Lambda}{\Omega_{m,0} (1+z)^3} + 1 \right] \\ &= \Delta_c \left[\frac{1 - \Omega_m}{\Omega_m (1+z)^3} + 1 \right] \end{aligned} \quad (3.28)$$

where G is the gravitational constant. For $E(z)$, Eq. 2.29 was used, as well as $\Omega_r \approx 0$ for $z \leq 1000$ (Peacock (1999) p. 84) and $\Omega = 1 \Rightarrow \Omega_\Lambda = 1 - \Omega_m$. Because larger overdensities result in less clusters to be found, this correction decreases the number of clusters, especially for small redshifts.

Fig. 3.3 shows the mass function for the redshifts $z = 0$, $z = 1$ and $z = 2.5$. Because the range of the mass function spans over several orders of magnitude, the ordinate is plotted with a factor of $M^2/\bar{\rho}_m$. As expected from the Press-Schechter approach (see Sect. 3.8.2) the number of clusters decreases with increasing redshift, which is not surprising since galaxy clusters evolve with time from gravitational collapse of overdense regions.

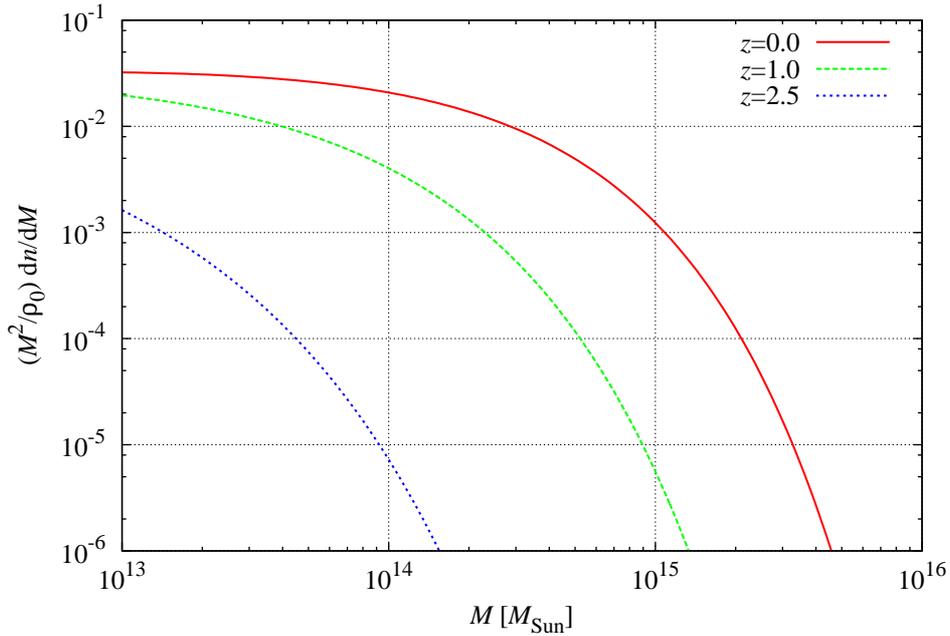


Figure 3.3: Halo mass function for $\Delta = 500$ (overdensity with respect to the mean density) and three different redshifts. The redshift evolution of the mass function is evident. For $z = 2.5$, the number density of masses is about one order of magnitude smaller than for $z = 0$.

4. eROSITA

4.1. Overview

The objective of the simulation presented in this thesis is to generate a list of galaxy clusters for the simulation of the all-sky survey which will be conducted with the instrument extended ROentgen Survey with an Imaging Telescope Array (eROSITA). eROSITA is one of the two main instruments on board of the *SRG* spacecraft. The following sections give a short introduction to *SRG* with a focus on eROSITA .

The first all-sky survey in X-rays was performed with *ROSAT* in the 1990s (Voges, 1993). Because *ROSAT* was only sensitive in soft X-rays, the missions *ABRIXAS* (Predehl, 1999) and *ROSITA* (extended ROentgen Survey with an Imaging Telescope Array, Predehl et al. (2003)) were planned to extend the *ROSAT* all-sky survey to higher energies, with the main goal of observing Active Galactic Nucleus (AGN) which are mainly obscured by gas and dust if observed in soft X-rays (Predehl et al., 2006). *ABRIXAS* failed shortly after the launch due to a malfunction in the power system (Predehl et al., 2006), while ROentgen Survey with an Imaging Telescope Array (*ROSITA*), which was designed to be attached to the International Space Station (ISS), was never realized because the scheduled launch date was 2011, one year after the planned end of the Space Shuttle program (Predehl et al., 2006). Furthermore the ISS turned out to be unsuitable for X-ray optics because of its dirty environment (Friedrich et al., 2005). After these two failed missions, it will be eROSITA which continues the work of *ROSAT*. eROSITA is based on the design of *ABRIXAS*, but with a significantly larger effective area to allow dark energy studies, which is now the central objective of the mission (Predehl et al., 2010). eROSITA is funded by the German Space Agency Deutsches Zentrum für Luft- und Raumfahrt (DLR) and the Max-Planck-Society.

4.2. Spectrum-Roentgen-Gamma

Spectrum-Roentgen-Gamma (Fig. 4.1) is a German-Russian project. The basis structure is the ‘Navigator’-platform developed by Lavochkin Association (Pavlinsky et al., 2009). The main instruments are two X-ray telescope arrays: eROSITA, which is contributed by Germany under direction of the Max-Planck Institut für extraterrestrische Physik (MPE), and the Russian hard X-ray instrument ART-XC, developed by IKI.

ART-XC consists, like eROSITA, of seven X-ray telescopes aligned parallel. The telescopes are conical approximations of the Wolter-I design and are equipped with CdTe-detectors (Predehl et al., 2010). The instrument is sensitive for higher energies than eROSITA and extends the energy band up to ~ 11 keV for the survey and ~ 30 keV for pointed observations (Pavlinsky et al., 2009).

4.3. Scientific Objectives

The scientific objectives of eROSITA are described in, e.g., Predehl et al. (2010). While *ABRIXAS* and *ROSITA* were mainly designed for the observation of Active Galactic Nuclei, the design driving science for eROSITA is the testing of cosmological models through large-scale structure observations (Predehl et al., 2010). In X-rays, galaxy clusters are good tracers

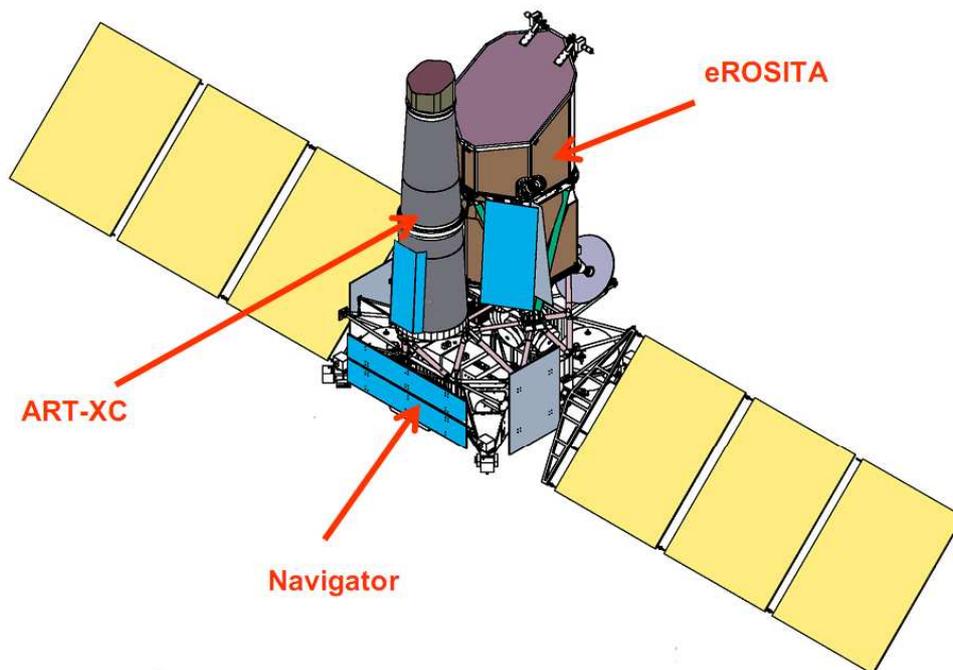


Figure 4.1: *SRG* in orbit configuration (Pavlinisky et al., 2009). The main instruments eROSITA and ART-XC are mounted on the ‘Navigator’-platform.

for these large scale structures (see Sect. 3.4). eROSITA can detect clusters up to redshifts of $z \approx 2.5$ with a precision of $\Delta z \approx 0.2$ (Pavlinisky et al., 2009). Furthermore eROSITA will help to understand dark matter and accretion physics (Pavlinisky et al., 2009).

4.4. The eROSITA-instrument

4.4.1. Wolter Telescopes

A detailed description of eROSITA can be found in, e.g., Cappelluti et al. (2011) and Predehl et al. (2010). The X-ray optics of the instrument consist of seven identical, co-aligned Wolter-I X-ray telescopes. To achieve the objectives described above, effective area and angular resolution had to be increased with respect to *ABRIXAS*. Thus the number of mirror shells was doubled. Every telescope consists of 54 gold coated, nested mirrors, where the 27 inner shells are identical to those used by *ABRIXAS* (Predehl et al., 2010). By increasing the number of mirrors, the effective area could be enhanced by a factor five for energies up to ~ 5 keV. For higher energies the outer shells do not contribute to the effective area because of the relatively large grazing angles (Predehl et al., 2006). The angular resolution improves that of *ABRIXAS* by a factor of two (Predehl et al., 2010). These improved capacity is needed for dark energy studies. Additional, the telescopes are co-aligned in contrast to those of *ABRIXAS*. A comprehensive display of the properties of eROSITA can be found in table 4.1. Each telescope has a focal length of 1600 mm (Predehl et al., 2010) and an effective area of ~ 1500 cm² at 1.5 keV (Cappelluti et al., 2011). This is about a factor of two better than *XMM-Newton* in this energy band (for a comparison of the effective areas of *XMM-Newton*, *ROSAT* PSPC, and eROSITA see Fig. 4.3). Photons from out of view reaching the detector after a single reflection on the paraboloid or hyperboloid surface are suppressed by an X-ray baffle in front of the telescopes.

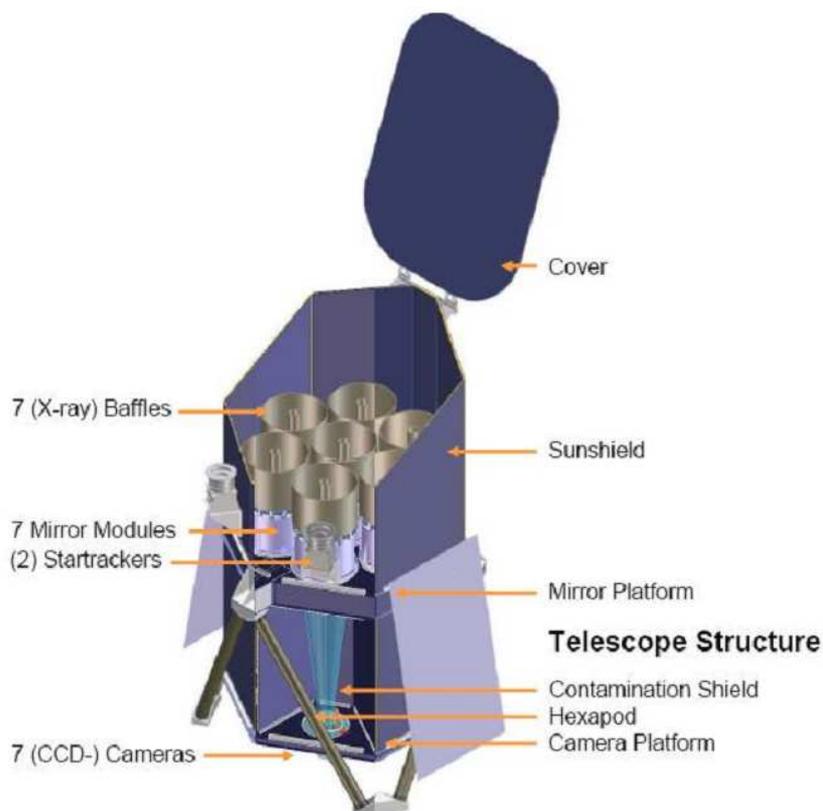


Figure 4.2: The eROSITA telescopes assembled on the carrier structure (Fürmetz et al., 2008).

Table 4.1: Properties of the eROSITA X-ray telescopes (Cappelluti et al., 2011)

Telescope design	Wolter-I
Number of telescopes	seven
Shells per telescope	54
Effective area at 1.5 keV	$\sim 1500 \text{ cm}^2$
On-axis PSF HEW	15''
Effective Angular Resolution	25 – 30''

The drawback of this design is additional vignetting, so a good compromise has to be found (Predehl et al., 2010). The mirrors have to be stabilized at $20 \pm 2 \text{ }^\circ\text{C}$ to avoid degradation of the imaging quality due to thermal deformations. This stabilization is reached by a system of heatpipes in combination with a heating system (Predehl et al., 2010).

4.4.2. pnCCD Detectors

Each of the seven Wolter telescopes is equipped with an identical pnCCD camera developed by the MPI Halbleiterlabor. The pnCCDs, which are backside-illuminated Charge Coupled Devices (CCDs), are advanced versions of the pnCCDs flying on *XMM-Newton* (Strüder et al., 2001) and have 384×384 pixels. In the energy range from 0.3 keV to 10 keV, an energy resolution close to the theoretical limit determined by Fano noise is achieved (Meidinger et al., 2009). The quantum efficiency is about 90 % (Meidinger et al., 2009). A comprehensive summary of

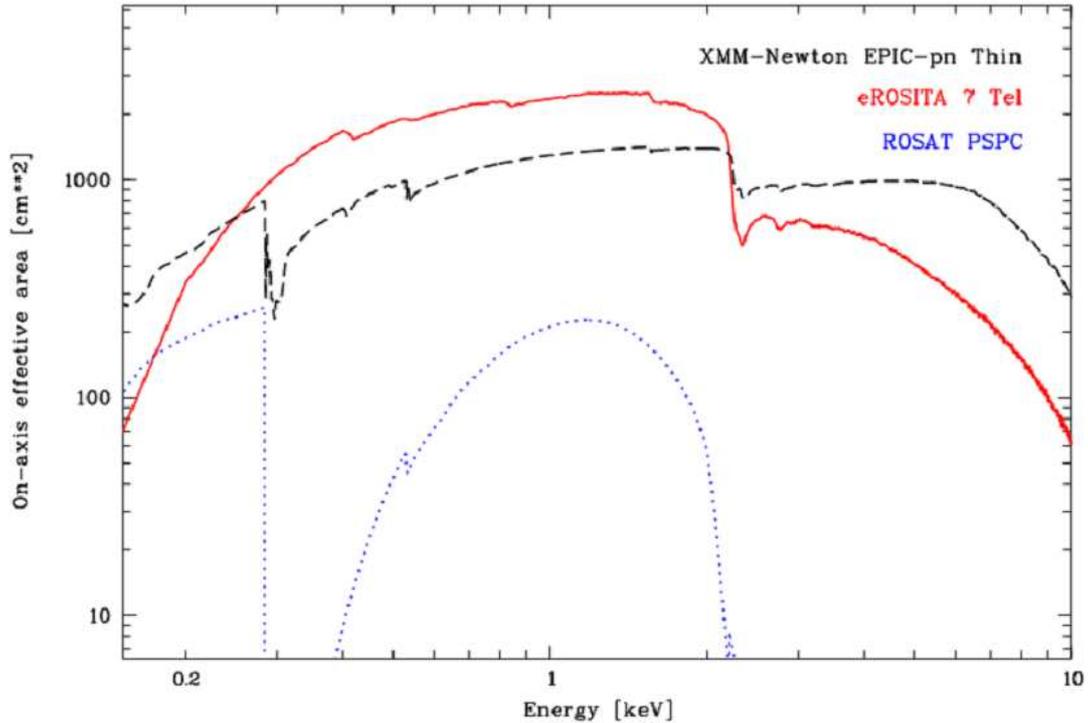


Figure 4.3: On-axis effective area for eROSITA, XMM-Newton and ROSAT-PSPC (Predehl et al., 2006)

the properties of the pnCCDs is given in table 4.2. The cameras have an imaging area and a framestore area where the image can be shifted to in less than $100 \mu\text{s}$ (Predehl et al., 2010) to minimize the probability of out-of-time events, i.e. photons recorded during readout. All 384 channels are read out simultaneously (Meidinger et al., 2009). For calibration purposes, every

Table 4.2: Properties of the eROSITA pnCCD-cameras (Meidinger et al., 2009).

Pixels	384×384
Chip size	$28.8 \text{ mm} \times 28.8 \text{ mm}$
Pixel size	$75 \mu\text{m} \times 75 \mu\text{m}$
Readout	all 384 channels parallel
Time for shifting integrated image to framestore	$\lesssim 100 \mu\text{s}$
Readout time	5 ms
Working temperature	$-80 \text{ }^\circ\text{C}$

pnCCD is equipped with a radioactive Fe^{55} source with an aluminium target. The source can be moved in and out of the field of view. The pnCCDs have to be kept at an operation temperature of $-80 \pm 0.5 \text{ }^\circ\text{C}$. This is achieved by passive elements, that is variable conductance heatpipes and radiators (Predehl et al., 2010).

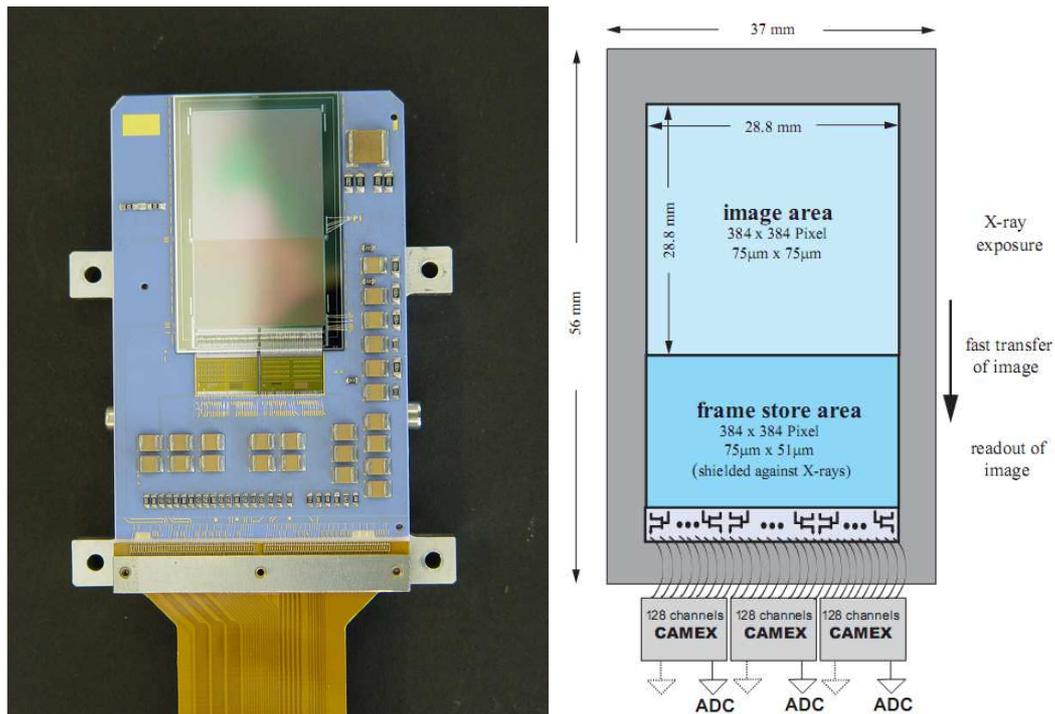


Figure 4.4: *Left image:* One of the eROSITA pnCCDs. On the bottom side the frame store area can be seen, which is smaller than the imaging area of the chip (Predehl et al., 2006). *Right image:* Schematic view of an eROSITA pnCCD (Meidinger et al., 2009)

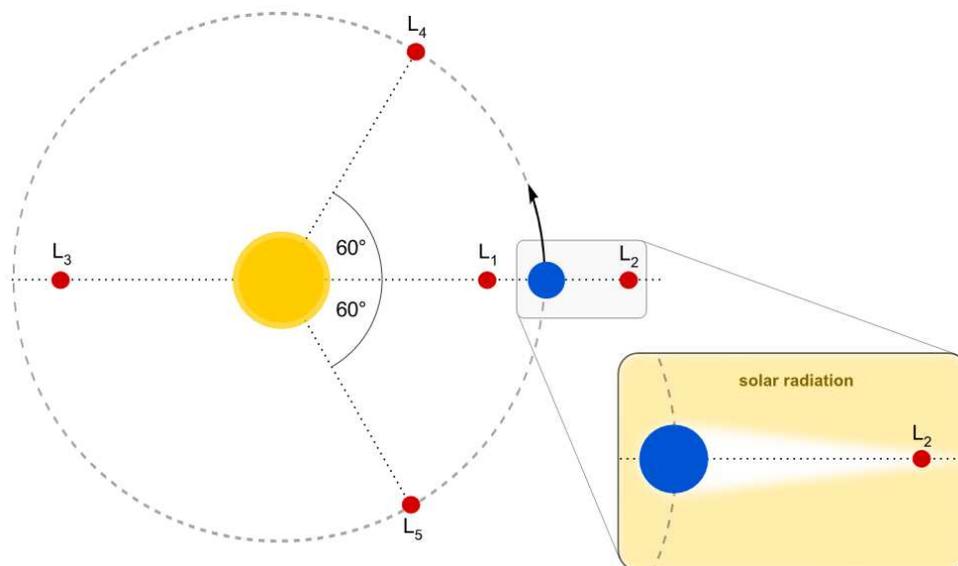


Figure 4.5: The earth-sun system with its five equilibrium (Lagrange) points L_1 - L_5 (Fig. from Wille, private communication). Earth is blue, the sun is yellow and the Lagrange points are red. eROSITA will orbit around L_2 in an elliptical orbit with a semi-major axis of $3 \cdot 10^5$ km, a semi-minor axis of $2.5 \cdot 10^5$ km, and an inclination of 35° with respect to the ecliptic.

4.5. Orbit

Launch is scheduled for 2013 with a Soyuz-Fregat from Baikonur, Kazakhstan (Pavlinisky et al., 2009). After a 110 day flight, *SRG* will reach its orbit around Lagrange point L_2 of the earth-sun-system (Fig. 4.5), 1.5 million kilometres from earth (Fürmetz et al., 2010). In L_2 , the joint gravitational force of earth and sun equals the centrifugal force on a much smaller object orbiting the sun with the same velocity as earth. Because L_2 is a saddle point of the effective gravitational potential and therefore dynamically unstable, the spacecraft has to perform course corrections (Fürmetz et al., 2010). *SRG* will be placed in an elliptical orbit around L_2 with a semi-major axis of $3 \cdot 10^5$ km, a semi-minor axis of $2.5 \cdot 10^5$ km and an inclination of 35° with respect to the ecliptic.

4.6. Observing Program

The first four years of the mission, eROSITA will conduct an all-sky survey. During this time it will rotate constantly, where the rotation axis always points to earth, so readjustments of the antenna position are not necessary. This constraint together with the orbit around L_2 results in a smearing of the scan poles to an area of a few hundred square degrees (Fürmetz et al., 2010). An exposure map is shown in Fig. 4.6.

The *ROSAT*-survey lasted only half a year, so together with the increased effective area, the eROSITA all-sky survey will improve on *ROSAT* all Sky Survey by a factor of about 30 in sensitivity (Cappelluti et al., 2011). The mean exposure time will be ~ 3 ks, while at the two scanning poles exposure times of 20 – 40 ks are reached. In the 0.5 – 2 keV band, a flux limit for clusters of $3 \cdot 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$ – $4 \cdot 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$ is expected (Predehl et al., 2010). According to Predehl et al. (2010), the eROSITA survey will reveal about 50,000 – 100,000 galaxy clusters and $3 - 10 \cdot 10^6$ AGN will be observed, including all clusters with masses above $3.5 \cdot 10^{14} h^{-1} M_\odot$ up to redshifts of $z = 2$. After the four-year survey phase, a three-year phase for pointed observations of interesting objects is planned.

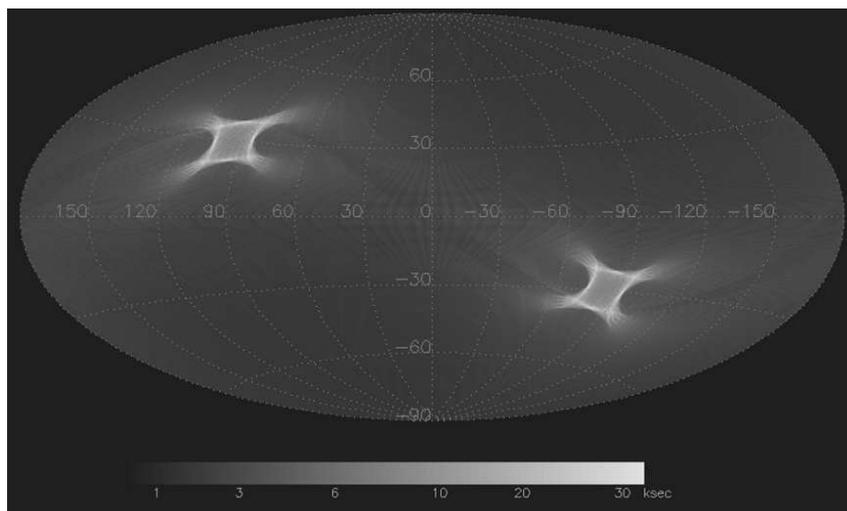


Figure 4.6: Exposure map for the eROSITA survey (Fürmetz et al., 2010). The brighter areas are the scanning poles.

5. SIXT Simulation Software

In this chapter, the *Simulation of X-ray Telescopes (SIXT)* (Schmid et al., 2010) software and the Simulation Input (SIMPUT) format (Schmid et al.), which is also the final format of the cluster catalogue, are introduced. The cluster catalogue discussed in this thesis was generated as an input catalogue for the simulation of the eROSITA survey. With increasingly complex and

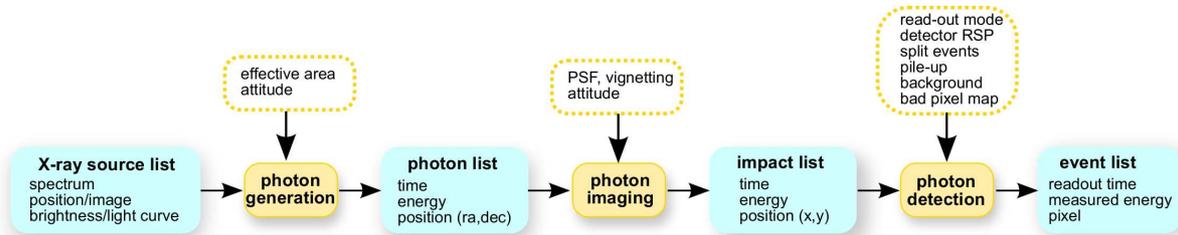


Figure 5.1: Flow chart of the *SIXT* pipeline (Schmid, 2011)

expensive missions, simulations become more and more important for planning and designing missions. The *SIXT* program for X-ray instruments was developed by Schmid et al. (2010). It is based on a Monte-Carlo algorithm and generates events for X-ray sources stored in an input catalogue. For every photon, the propagation through telescope and the response of the detector is simulated using existing calibration data like the Point Spread Function (PSF). The software is designed modularly, so it can easily be adapted to different instruments (Schmid et al., 2010). The output of the simulation is an event list as obtained by real observations. This list can be processed further to analyze properties of the instrument, before the latter one is actually constructed.

Fig. 5.1 shows the different steps of the *SIXT* pipeline. The input is a source catalogue with the corresponding spectra and optional image and light curve. With regard to pointing and effective area of the instrument, photons are generated by a Monte Carlo method. This photon list is processed together with the intrinsic properties of the instrument, like PSF and vignetting as a function of attitude. From this, an impact list of the photons hitting the detector is generated. In the last step it is accounted for the detector properties, and the event list as obtained from real observations is returned (Schmid et al., 2010).

An universal format for source catalogues to be used as input for simulations was defined by Schmid et al. with the SIMPUT format. It is a Flexible Image Transport System (FITS) file (Hanisch et al., 2001; Pence et al., 2010) with a source catalogue extension containing one or more sources. Further extension can contain images, spectra and light curves. Alternatively, these can be stored in other FITS files, in this case the source extension contains links to those files.

6. Generating the Cluster Catalogue

In the last chapters, the theoretical background about galaxy clusters, their importance for cosmology and eROSITA was introduced. In this chapter, the core of this thesis will be discussed: the sampling of a cluster catalogue via a Monte Carlo simulation. Before describing the code in detail in Sect. 6.4 - 6.10, some basics of the Monte Carlo technique are introduced in Sect. 6.1 without a focus on mathematical rigourousity. The code itself consists of several *Python*-scripts. After a short introduction to the *Python* programming language in Sect. 6.2, the general approach to the problem is presented in Sect. 6.3. An overview over the simulation pipeline is given in Sect. 6.4, before the single scripts are described in detail in Sect. 6.6 - 6.10, followed by a discussion of the results in Sect. 6.11. The simulation scripts would be relatively useless without having tested the plausibility of the results. This was done in Sect. 6.12 by comparing the calculated mass function with the results of Tinker et al. (2008) and a grid calculated independently. Finally, improvements which could be implemented in the future are discussed.

6.1. Monte Carlo Methods

Monte Carlo methods describe a class of computational algorithms working with repeated sampling of random variables. Applications are, for example, the integration of multidimensional functions with complicated boundary conditions, the simulation of complex systems with many coupled degrees of freedom like fluids or economic systems, or the sampling of values distributed according to a given probability distribution function. The latter case shall be discussed now. For conciseness, the case with only one random variable is treated without loss of generality. Let $f(x)$ be a Probability Density Function (PDF) of the variable $x \mid x \in [a, b]$ and $F(x)$ the Cumulative Distribution Function (CDF) of x .

The PDF describes the probability of finding the random variable in the infinitesimal interval $[x_0, x_0 + dx]$:

$$P(x_0, x_0 + dx) = f(x_0) dx \quad (6.1)$$

The CDF describes the probability of the random variable having a value smaller than a given y :

$$F(y) = P(x \leq y) \quad (6.2)$$

Since a PDF is normalized and $f(x) > 0 \forall x \in \mathbb{D}$, the corresponding CDF is monotonically increasing and $\lim_{y \rightarrow b} F(y) = 1$. The CDF is related to the PDF as follows:

$$F(y) = \int_{-\infty}^y f(x) dx \quad (6.3)$$

In the following, the most important techniques for generating a sample of random variables distributed according to a given PDF are discussed. For a more complete and rigorous treatise, see e.g. Deák (1990). Most random generators generate variables distributed uniformly in the unit interval. Such a variable, be it called u , can be projected any other interval $[a, b]$ simply by a linear function $\Phi(u) = a + (b - a)u$, where the uniform distribution is preserved. Therefore, from now on all considerations start with a sample of variables distributed uniformly in the unit interval.

6.1.1. Inversion Method

If the CDF is a bijective function, there exists an inverse function F^{-1} . It can be proved that for a variable u distributed uniformly in the unit interval, the variable x with $x \leftarrow F^{-1}(u)$ is distributed according to F (Deák, 1990). This is called inversion method (Deák, 1990).

For practical purpose this technique is especially applicable if F^{-1} can be calculated analytically, so the variables sampled according to $f(x)$ can easily be generated. This technique is effective because the variables can be calculated directly, but it can only be applied to a limited set of functions.

6.1.2. Rejection Sampling

The inversion method described above can only be used for invertible functions. A more generally applicable method is the rejection sampling described in Deák (1990). Let $f(x)$ be a PDF which is bounded above by c , so $f(x) \leq c \forall x \in [a, b]$. Now, the random variables can be sampled (Deák, 1990):

1. create u_1 distributed uniformly in $[a, b]$ and a random variable u_2 distributed uniformly in $[0, c]$.
2. if $u_2 \leq f(u_1)$ accept u_1 , else go back to step 1.

Because a random variable is either accepted or rejected, rejection sampling is also referred to as acceptance-rejection method. If the values of f span over a wide range, this method can be very expensive because a large part of the generated values is rejected.

Rejection sampling, as described above, can be improved by generalizing the method for an arbitrary function as upper bound. Be $g(x)$ a function with $f(x) \leq g(x) = ch(x)$, where $c \geq 1$ is a constant (Deák, 1990). In practical one selects a function $h(x)$ which is numerically easier to handle than the $f(x)$, which is the case if the bounding function $g(x)$ can be inverted analytically. Now, the random variable is sampled (Deák, 1990):

1. create u_2 distributed uniformly in the unit interval. Generate u_1 distributed according to h in $[a, b]$ via the inversion method.
2. if $u_2 \leq \frac{f(u_1)}{g(u_1)} = \frac{f(u_1)}{ch(u_1)}$, accept u_1 , else go back to step 1.

By selecting a convenient function h , the numerical costs can be reduced significantly, especially if the values of f span over a wide range. The efficiency η of the rejection sampling technique can simply be expressed by the fraction of the accepted values:

$$\eta = \frac{n_{\text{accepted}}}{n_{\text{accepted}} + n_{\text{rejected}}}$$

6.2. Python

*Python*¹¹ is an interpreted high-level programming language developed in 1991 by Guido van Rossum (Ernesti & Kaiser, 2008). It is platform independent and supports imperative, object-oriented, and functional programming paradigms (Ernesti & Kaiser, 2008). The language attaches great importance to code readability. A characteristic of *Python* is the use of indentations

¹¹The name *Python* is a homage to the British comedy group Monty Python (Ernesti & Kaiser, 2008).

as block delimiters. Besides a vast standard library, there exist some libraries for scientific purposes, the most important of them are *Scientific Python (SciPy)*, *Numerical Python (NumPy)* (Jones et al., 2001), and the plotting library *matplotlib* (Hunter, 2007). The *SciPy* and *NumPy* libraries offer a powerful array object, which does not exist in basic *Python*, and a variety of further functions.

In the descriptions below, where technical details of the scripts are discussed, one has to keep in mind that in programming languages indices mostly start with 0, which is adopted here, if algorithms are described. It should be discernible from the context which convention is used.

6.3. General Approach

The purpose of the simulation is to generate a list of galaxy clusters with the following observables:

- celestial coordinates (right ascension α , declination δ)
- redshift z
- flux (calculated as a function of the cluster total mass M and the redshift)
- angular diameter

The latter two observables are not sampled by the simulation, but are derived from the mass and redshift. Hence the Monte Carlo code generates the parameters

- celestial coordinates (right ascension α , declination δ)
- redshift z
- mass M

The direct calculation of the mass function is very expensive, mainly due to the numerical integration needed to obtain σ (root mean square of the linear matter power spectrum, see Eq. 3.21 on p. 24). Therefore, the script generating the clusters uses a grid with the mass function values, which has to be calculated before. Since between $z = 0$ and $z = 2.5$ about 23 million clusters are found for a minimal mass of $10^{13} M_{\odot}$ and, due to the usage of the rejection sampling technique, still more values have to be calculated it is much more economic to calculate a grid with some hundred thousand nodes.

Next, the number of objects in a given redshift interval is calculated. An option is simply to integrate over the whole interval, but there are some drawbacks of this solution: if the catalogue is calculated in one rush, the output file is relatively large and therefore impractically to handle compared to smaller files. Further, if the program is extended in the future to take the spatial correlation of clusters into account, it saves significant calculation time if only clusters in a small redshift interval have to be included for the calculation of the correlation (under the assumption that there is no correlation for larger distances, which seems reasonable since the correlation function can be described as a power law of the cluster distance, as was discussed in Sect. 3.6). This can be easily achieved by only considering the clusters in the neighbouring redshift shells. This possible extension of the code is discussed in Sect. 6.12. Thus, instead of integrating over the whole domain, the redshift interval is subdivided in a user-specified number of intervals (see Sect. 6.7).

When the file containing the number of objects per shell is loaded, a Poisson distribution with the read-in number as expectation value is applied to randomize the cluster number. Now the sampling of the objects starts.

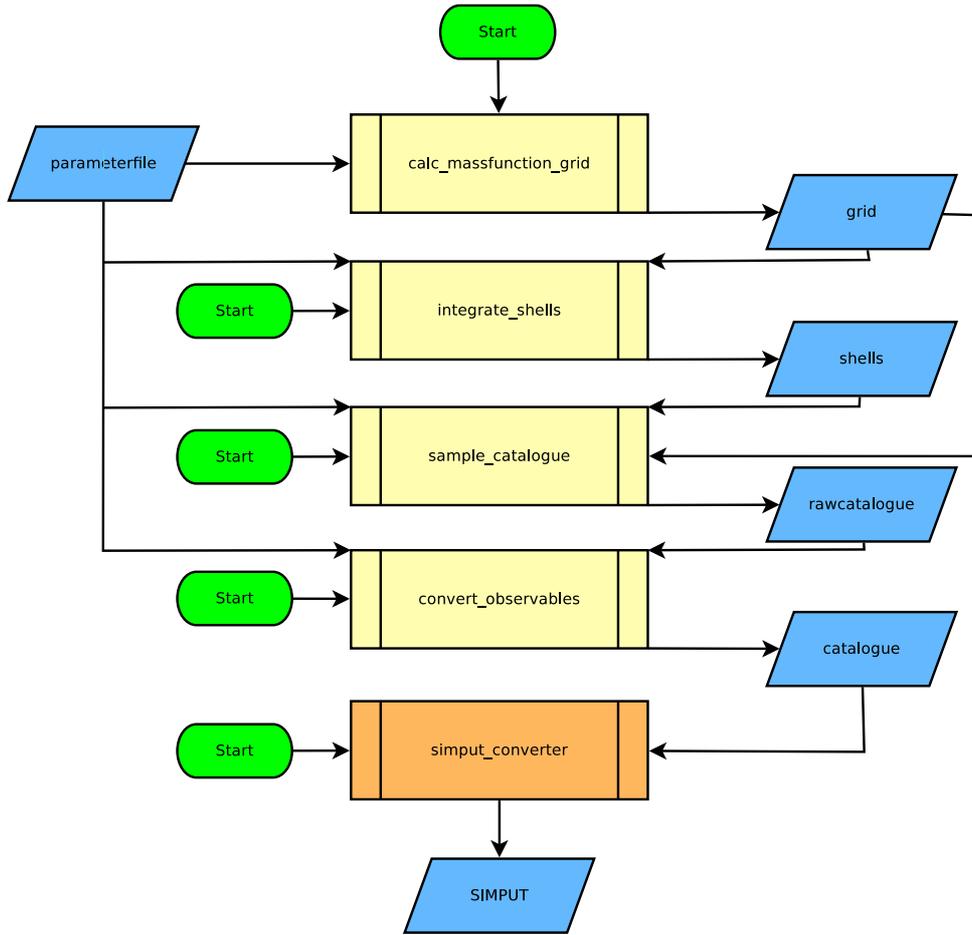


Figure 6.1: Flow chart of the pipeline for the generation of a cluster catalogue. The *Python*-scripts sampling the catalogue are filled yellow, the *C*-program to create the *SIMP*UT-files is orange. In and output files are blue, the points where the script has to be launched manually are green.

For the sampling of the celestial coordinates, the spatial distribution which is described by the correlation function is neglected (see Sect. 3.6), such that the right ascension and declination can be sampled independently. The right ascension is distributed uniformly and can be generated by using the random number generator from the *Python* standard library. The declination is obtained with rejection sampling, where the value δ for the declination is accepted if an auxiliary variable distributed uniformly in $[0, 1]$ is less or equal $\cos(\delta)$. For the mass and redshift the situation is more complicated, because these parameters are i) correlated and ii) the mass function, which is the PDF, if normalized, cannot be inverted analytically. Hence the variables have to be generated together via rejection sampling (see Sect. 6.1.2). Because the function values range over several orders of magnitude, rejection sampling with a constant as upper bound would be too ineffective, because only a insignificantly small part of the generated values would be accepted. Thus a applicable function has to found as upper bound. Up to the exponential cutoff, the mass function (Eq. 3.19) can be well approximated by a power law of the form

$$h(M) = \alpha \cdot M^\beta \quad (6.4)$$

Since the mass function decreases with increasing redshifts for all masses, a function h which depends on the mass as a power law and is independent of redshift is a upper bound to the

massfunction, as long as $h(M) > \frac{dn}{dM} \forall M \in [M_{\min}, M_{\max}]$ at $z = 0$. The exponential cutoff is located well over $M = 10^{14} M_{\odot}$, therefore the values of the minimal mass M_{\min} , which was selected for the simulation, and $M = 10^{14} M_{\odot}$ where used to define the power law. The exponent β can be calculated from

$$\beta = \left[\log \left(\frac{dn}{dM} (10^{14} M_{\odot}) \right) - \log \left(\frac{dn}{dM} (M_{\min}) \right) \right] \times \left[\log (10^{14} M_{\odot}) - \log (M_{\min}) \right]^{-1} \quad (6.5)$$

and factor α is obtained by

$$\alpha = M_{\min}^{-\beta} \frac{dn}{dM} (M_{\min}) \quad (6.6)$$

The power law from Eq. 6.4 is the PDF multiplied with a constant factor, which will later be eliminated by normalization. The CDF needed for the acceptance-rejection method with a function as upper bound, as described in Sect. 6.1.2, is the definite integral of the PDF multiplied with a constant, i.e. it is > 1 for the right side of the definition interval. To discern this function from the PDF, it is denoted as H' here

$$H'(M) = \int_{M_{\min}}^M \alpha \cdot M'^{\beta} dM' = \frac{\alpha}{1 + \beta} (M^{\beta+1} - M_{\min}^{\beta+1}) \quad (6.7)$$

Now H' is normalized for the mass interval $[M_{\min}, M_{\max}]$ in which the cluster masses will be generated:

$$H(M) = \frac{H'(M)}{H'(M_{\max})} \quad (6.8)$$

as one can see from the two equations above, $H(M)$ equals zero for M_{\min} and one for M_{\max} . That function can easily be inverted:

$$H^{-1}(y) = \left[\frac{y(\beta + 1)}{\alpha} + M_{\min}^{\beta+1} \right]^{1/(\beta+1)} \quad (6.9)$$

This is the equation used as upper bound for the rejection sampling. As well as the exponent of the power law, the factor c is determined anew for the minimal redshift of the interval of redshift in which the clusters are sampled. As a constant factor c (see Sect. 6.1.2) a value of $2 \cdot (1 + z_{\min})$ was used, so it is ensured that the mass function is still smaller than the bounding function for small deviations from the power law. The z -dependence results from the experience that for larger redshifts a higher factor is needed to hold the mentioned condition. Nevertheless, the simulation code checks if $f < ch$ for every value it samples. Now the primary four parameters described above are sampled.

Finally, the observables are calculated. The flux is obtained with the $L_X - M$ relation from Vikhlinin et al. (2009) and Eq. 2.42 on p. 15. The angular diameter distance is calculated with Eq. 2.38 on p. 15. At last the catalogue is converted into the SIMPUT-format (Schmid et al.).

6.4. Pipeline

While above the course of action was depicted without a focus on the technical details, now the function of the single scripts is explained. The complete simulation consist of four individual *Python* scripts, and a *C*-program writing the final *SIMP*UT-file. This modular approach was

Table 6.1: Description of the parameters which have to be given in the parameter file

variable name	data type	description
cat_path	string	path to the catalogue file
cat_name	string	name of the catalogue file
grid_path	string	path to the grid file
grid_name	string	name of the grid file without the .npz file extension
shells_path	string	path to the file containing the number of objects per shell file
shells_name	string	name of the file containing the number of objects per shell file
nodes_x	int	nodes of the grid on the x -axis (redshift)
nodes_y	int	nodes of the grid on the y -axis (redshift)
rho_0	float	mean matter density of the universe
cvak	float	speed of light in vacuum
Omega	float	total density parameter Ω
Omega_r	float	radiation density parameter Ω_r
Omega_m	float	matter density parameter Ω_m
Omega_b	float	baryon density parameter Ω_b
h100	float	Hubble parameter h_{100}
T0	float	CMB temperature
sigma_8	float	amplitude of the primordial power spectrum (Peacock, 1999)
n_s	float	primordial spectral index (Peacock, 1999)
mass_min	float	minimal cluster mass M_{\min}
mass_max	float	maximal cluster mass M_{\max}
z_min	float	minimal redshift
z_max	float	maximal redshift
overdensity	float	overdensity for clusters with respect to the Critical Density at z

chosen to achieve maximal flexibility and to preserve the possibility to inspect or manipulate the in- and output of every step. The output is stored in a file and read in in the next step. The variables specifying the cosmological parameters, filenames etc. are stored in a parameter file which is read in by the scripts. The pipeline is shown in Fig. 6.1.

The grid is saved in the .npz binary format provided by *NumPy* (Jones et al., 2001). The ‘shell’-file (described below) and the catalogues are saved as ASCII-files and can be inspected with every editor or used as an input for plot programs. The final output is a FITS file in the SIMPUT-format (Schmid et al.) described in Sect. 5. In the following sections, the parameter-file and the *Python*-scripts are described.

6.5. Parameterfile

All parameters relevant for the simulation are stored in a parameterfile, which is read in by the scripts. The file has to contain the variables listed in table 6.1. Blanks, empty lines and lines beginning with ‘#’ are ignored. The order of the parameters can be arbitrary. The only condition is that the variable name and value have to be separated by a ‘=’, with the variable name on the left hand side and the value on the right hand side. An example parameter file is shown in Fig. 6.2.

```

# I/O:
cat_path = ./catalogues/
cat_name = cat_wmap7
grid_path = ./grids
grid_name = grid_wmap7
shells_path = ./shelltables
shells_name = shells_wmap7

# simulation:
mass_min = 1e13
mass_max = 1e17
z_min = 0
z_max = 2.5
overdensity = 500

# grid:
nodes_x = 100
nodes_y = 1500

# cosmology:
rho_0 = 3.719e10
cvak = 2.99792e8
Omega = 1.
Omega_r = 0.0
Omega_m = 0.2707
Omega_b = 0.0451
h100 = 0.703
T0 = 2.726
sigma_8 = 0.809
n_s = 0.966

```

Figure 6.2: Example for the ASCII-file containing the parameters for the simulation

6.6. Script *calc_massfct_grid*

This function calculates the mass function (Eq. 3.19) on a rectangular, semilogarithmic grid. Because the mass goes over several orders of magnitude and the mass function can be roughly described by a power law, the nodes along the mass axis are spaced logarithmically.

The function is called with the limits of redshift and mass as well as the number of nodes M and N in redshift respectively mass. It returns three arrays with dimension 2:

- X_{mn} : contains the value for the redshift on the node (m, n) . Because the grid is rectangular, $X_{mp} = X_{mq} \forall p, q \leq N; p, q \in \mathbb{N}$
- Y_{pn} : contains the value for the mass on the node (m, n) . Because the grid is rectangular, $Y_{pn} = Y_{qn} \forall p, q \leq M; p, q \in \mathbb{N}$
- Z_{mn} : contains the function value on the node (m, n) : $Z_{mn} = f(X_{mn}, Y_{mn})$

On every node of the grid, the mass function is calculated by the function *massfct*.

6.6.1. Function *massfct*

This function calculates the mass function from Tinker et al. (2008) directly. At first the function *calc_sigma* (see section 6.6.2) is called to obtain the root mean square σ of the linear matter power spectrum.

Now the transformation from a overdensity relative to the Critical Density to one relative to the mean matter density is accomplished via Eq. 3.28.

In the next step, the fit function $f(\sigma)$ is calculated for the σ and Δ obtained above. The redshift-dependent parameters A , a , b and c are obtained via the functions given in Tinker et al. (2008), Eq. 5-8:

$$A = A_0 (1 + z)^{-0.14} \quad (6.10)$$

$$a = a_0 (1 + z)^{-0.06} \quad (6.11)$$

$$b = b_0 (1 + z)^{-\alpha} \quad (6.12)$$

$$c = c_0 \quad (6.13)$$

$$\log \alpha = - \left(\frac{0.75}{\log\left(\frac{\Delta}{75}\right)} \right)^{1.2}$$

The parameters for $z = 0$, A_0 , a_0 , b_0 and c_0 are functions of the overdensity Δ and can be calculated from the interpolation formula in Tinker et al. (2008), appendix B, Eq. B 1-B 3:

$$A_0 = \begin{cases} 0.1 \log \Delta - 0.05 & \Delta \leq 1600 \\ 0.26 & \Delta > 1600 \end{cases} \quad (6.14)$$

$$a_0 = 1.43 + (\log \Delta - 2.3)^{1.5} \quad (6.15)$$

$$b_0 = 1 + (\log \Delta - 1.6)^{-1.5} \quad (6.16)$$

$$c_0 = 1.2 + (\log \Delta - 2.35)^{1.6} \quad (6.17)$$

The authors warn about an error up to $\approx 10\%$ of this interpolation formula, which is acceptable for the purpose of this simulation.

Next, the derivative in Eq. 3.19 is calculated:

$$\frac{d \ln \sigma^{-1}(M)}{dM} = \frac{\ln \sigma^{-1}(M + \delta M) - \ln \sigma^{-1}(M - \delta M)}{2\delta M} \quad (6.18)$$

For δM , a dynamical stepsize of $10^{-4}M$ was selected. Here, some fine-tuning or the implementation of a more sophisticated algorithm could be considered. Like above, σ is calculated with the subroutine *calc_sigma*. Now the value of the mass function (Eq. 3.19) is obtained by multiplying the factors calculated above.

To ensure that no negative values of the mass function were obtained due to numerical problems or false parameters, it is tested if the result of the function is positive. If not, a warning is raised.

6.6.2. Function *calc_sigma*

The function *calc_sigma* returns the root mean square of the linear matter power spectrum, smoothed by a top-hat function with a radius dependent on the enclosed total mass (Eq. 3.22 on p. 24).

At first, the shape parameter Γ for the transfer function of the power spectrum is calculated with Eq. 3.25. The transfer function $T(k)$ and the smoothing radius corresponding to the given halo mass and redshift, given through Eq. 3.24 respectively Eq. 3.23, are defined as subfunctions.

Eq. 3.19 is splitted into integrals for the numerator and the denominator. In the denominator, the top-hat function is calculated for a constant $R = 8h_{100}^{-1}$ Mpc.

For the integration, the integrands in numerator and denominator of the mass function are defined as functions to be integrated numerically with the integrator `quad` from the `SciPy`-package, which is based on the `Fortran` library `QUADPACK` (Piessens et al., 1983). A critical point here is the upper limit of subdivisions for the integrator: especially at high overdensities, which is here corresponding to low redshifts due to Eq. 3.28, an upper limit significantly higher than the default value has to be selected to achieve convergence of the integral.

To avoid potential problems with the integral at the first and second root of the integrand in the counter, the integration is splitted in three intervalls: $[0; k_0]$, $[k_1; k_2]$, and $[k_1; \infty]$ where k_n is the n th root of the integrand. The roots are calculated via the secant method, which is implemented in the `newton`-function of the `SciPy.optimize`-package. After the integration, σ is obtained from the terms which were calculated separately above. By far the most expensive part of this function is the integration. Hence, here some fine tuning could be done.

6.7. Script `integrate_shells`

This function divides the rectangular, two-dimensional grid along the z -axis (z denotes the redshift and not a Cartesian coordinate here) in D rectangular domains and integrates over those stripes.

$$\int_{z_0}^{z_1} \int_{M_{\text{halo}}} dM'_{\text{halo}} dz f(z, M'_{\text{halo}}) V(z) \quad (6.19)$$

where z_0 and z_1 are the minimal respectively maximal redshift of the domain. The tangential part of the integration is already contained in the volume element, because it yields simply a factor of 4π in an isotropic universe. There are some constraints: the rectangular domains have to be delimited by nodes of the grid. This has two consequences:

- the number of domains D cannot be greater than the number M of nodes along the z -axis minus one (M denotes still the number of nodes along the redshift-axis and not a mass in this section! The mass, where needed, will be denoted M_{halo}). If a number D greater than the allowed maximum is selected, D is automatically adapted to $M - 1$.
- the size of the domains is not necessarily constant. If the spacing of the nodes in z is not constant, the size of the domains will vary, too. Also, if the number of nodes in z minus one is not a multiple of the number of domains D , the function will automatically adapt the boundaries of the domains to the nodes, such that it is not necessary to use interpolated values at the boundaries of the domains.

`integrate_shells` reads the cosmological parameters, input- and output filenames, etc., from the parameterfile. After this, a vector l containing the limiting indices of all the domains is created. At first, an auxiliary vector l' is defined:

$$l'_j = j \cdot \frac{M - 1}{D}, \quad 0 \leq j \leq D; \quad j \in \mathbb{N}$$

The adaption to the nodes is achieved by simply rounding down to a integer.

$$l_j = \lfloor l'_j \rfloor, \quad 0 \leq j \leq D; \quad j \in \mathbb{N}$$

The array l contains the indices for the nodes on the z -axis defining the redshift-limits of the integration domains. In the next step, it is iterated over all the domains $0 \leq i < \dim(l) - 1 = D$. For every node in the domain, the M_{Halo} -axis is integrated using the composite trapezoidal rule provided by the integrator *trapz* from *SciPy.integrate* (Jones et al., 2001). The integrator is called with a vector $\zeta_n = Z_{l_i,n}$ containing the function values and $v = Y_{l_i,n}$ for the spacing of the values in ζ . For every domain, the values of the integrals along the M_{Halo} -axis are stored in a vector κ with the dimension $l_{i+1} - l_i + 1$. In the second step the integration along the z -axis is done with the *trapz* integrator. For the redshifts corresponding to the integrals in κ , a vector ξ is generated for every domain.

$$\xi_j = X_{(l_i+j),n}, \quad 0 \leq j \leq (l_{i+1} - l_i)$$

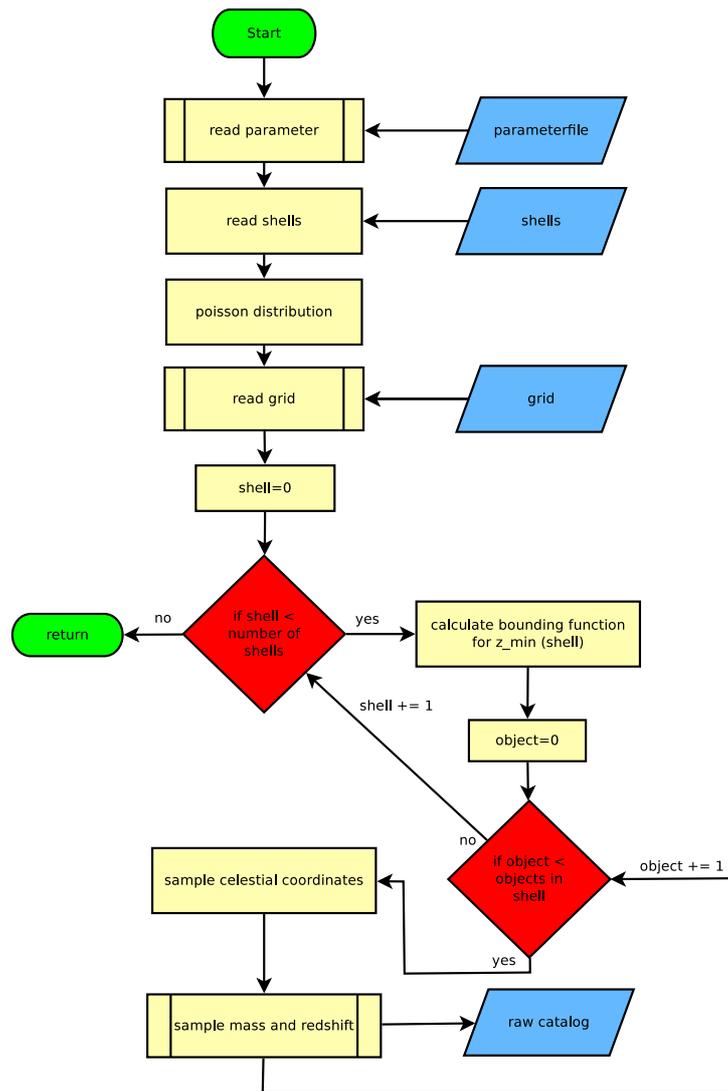
To obtain the integrand, κ is multiplied elementwise with the differential comoving volume element $\frac{dV}{dz}(\xi_j)$ (Eq. 2.45). Now the integrator is called with the vector κ containing the function values and ξ for the stepsize in the z -axis, along which the integration is done. Because the number of clusters has to be an integer, the integral is rounded. The results are written to the file, specified in the parameterfile as `shells_path/shells_name`. For an example see Fig. 6.3.

z_min	z_max	number
0.0	0.0727272727273	13406.0
0.0727272727273	0.169696969697	138541.0
0.169696969697	0.266666666667	378502.0
0.266666666667	0.363636363636	674839.0
⋮	⋮	⋮
2.03548387097	2.15161290323	411325.0
2.15161290323	2.26774193548	315030.0
2.26774193548	2.38387096774	237849.0
2.38387096774	2.5	177128.0

Figure 6.3: Example for an ASCII-file containing the number of objects per redshift interval

6.8. Script *sample_catalogue*

The program flow of this script is shown in Fig. 6.4. At first, the script reads the file with the redshifts intervals and number of clusters contained in it, which was generated by the script *integrate_shells*. To randomize the number of clusters per redshift interval, the read-in values are seen as expectation values of a Poisson-distribution. With the function *poisson* from the *numpy.random* library, a number according to this distribution is calculated for every redshift interval. Now the script iterates over the redshift shells and samples the right ascension using the random number generator from the *Python* standard library and the declination with a simple rejection sampling technique described in Sect. 6.3. Mass and redshift are generated with the

Figure 6.4: Flow chart of the script *sample_catalogue*

rejection sampling technique with a function as upper bound as described for this specific task in Sect. 6.3, and from a general point of view in Sect. 6.1.2. The bounding function is calculated anew for every new redshift interval, so it is always well adapted to the mass function. The objects are written to the file *raw_catalogue*. An example for a raw catalogue is shown in Fig. 6.5.

6.9. Script *convert_observables*

This script reads a raw catalogue and calculates the bolometric flux (Eq. 2.42) from the mass and the redshift, and the angular diameter distance as a function of redshift. The results are written to a new file with the same filename as the raw catalogue, where ‘_final’ is appended to the filename 6.3.

#	RA	DE	z	M
91.0367	-31.0818	0.4747	1.515e+13	
41.8705	-24.2324	0.4922	3.528e+14	
65.9683	-34.0029	0.5226	1.539e+13	
174.6393	-53.8424	0.5394	1.387e+13	
173.1282	-5.0177	0.4937	2.127e+14	
204.7508	57.4937	0.5517	1.955e+13	
67.0025	44.8450	0.5285	1.249e+13	
16.8363	13.5617	0.4644	1.018e+13	
281.6464	76.9146	0.4896	1.007e+13	
53.1876	-47.3649	0.4957	1.645e+13	
⋮	⋮	⋮	⋮	

Figure 6.5: Example for an ASCII-file containing the raw catalogue. The first column contains the right ascension, the second the declination, the third the redshift and the fourth the mass.

6.10. C-program *simput_converter*

The SIMPUT-Converter generates a FITS-file (Hanisch et al., 2001) in the SIMPUT-format (Schmid et al.) from the ASCII catalogue file. As was described in Sect. 5, in the SIMPUT file every object entry contains a link to an image. The *simput_converter* links the object to a *XMM-Newton* image of a cluster and to a cluster spectrum. Because the morphology of the clusters differs from object to object, therefore the image linked to every object is randomly selected from a list of clusters. The size of the cluster is defined via the IMGSCAL entry in the FITS-table. This entry denotes a linear scaling factor of the image. Since the apparent diameter of an object with diameter d is given by Eq. 2.39 as $\theta = \frac{d}{D_A}$, under the assumption that the clusters on the images have all the same diameter d_0 and appear at an angle $\theta_0 = d_0/D_A(z_{\text{img}})$, the IMGSCAL factor is given as

$$\text{IMGSCAL} = \frac{\theta_{\text{object}}}{\theta_0} = \frac{D_A(z_0)}{D_A(z_{\text{object}})} \quad (6.20)$$

where the index 0 denotes the values of the *XMM-Newton* images and ‘object’ those of the object from the simulated catalogue. The bolometric flux from the catalogue is given as an entry in the SIMPUT-objectlist.

6.11. Results

For the catalogues presented in this section, the *WMAP* seven-year parameters were used (see Sect. 2.4). Additionally, $\Omega_r = 0$, $\Omega = 1$ (Peacock, 1999) and $T_0 = 2.726$ (Mather et al., 1994) were used. A minimal cluster mass of $10^{13} M_\odot$ was assumed. If the cluster is too small, the flux is too weak for eROSITA to detect the cluster.

As grid size, 98 nodes for the redshift and 1500 logarithmically spaced nodes for the mass were selected. This is a sufficient size, as is discussed in 6.12.

To compare the object list with the mass function, a normalized histogram of the object masses was plotted against the normalized mass function in Fig. 6.8. The histogram was generated for $0.36 \leq z \leq 0.46$, while the mass function was calculated for $z = 0.41$.

#	RA	DE	z	F	DA
91.0367	-31.0818	0.4747	1.15234088498e-11	1230.9224	
41.8705	-24.2324	0.4922	1.70808941847e-09	1255.8621	
65.9683	-34.0029	0.5226	9.8414356384e-12	1296.937	
174.6393	-53.8424	0.5394	7.84400427855e-12	1318.4666	
173.1282	-5.0177	0.4937	7.51921988729e-10	1257.9549	
204.7508	57.4937	0.5517	1.30697583118e-11	1333.7232	
67.0025	44.845	0.5285	6.88479071884e-12	1304.5904	
16.8363	13.5617	0.4644	6.33792738882e-12	1215.7827	
281.6464	76.9146	0.4896	5.6267973834e-12	1252.218	
53.1876	-47.3649	0.4957	1.21100307566e-11	1260.7344	
⋮	⋮	⋮	⋮	⋮	

Figure 6.6: Example for an ASCII-file containing the final catalogue. The columns contain from the left to the right: Right ascension, declination, redshift, flux in $\text{erg s}^{-1} \text{cm}^{-2}$, Angular diameter distance in Mpc.

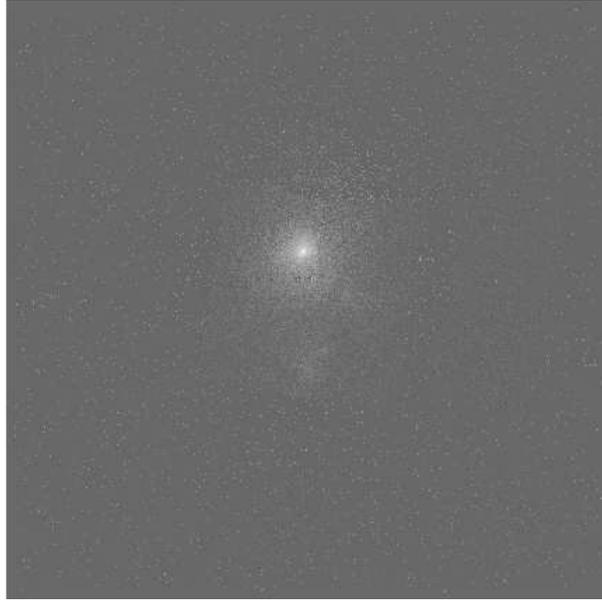


Figure 6.7: XMM-Newton image of the cluster ACO 85, as it is used in the object catalogue. SIMBAD gives a redshift of $z = 0.0521$ for this cluster.

6.12. Discussion

Before using the Monte Carlo code for the simulation of the eROSITA survey it is required to validate the code. This was done by comparing it with the results from Tinker et al. (2008) for $z = 0$ and by comparing it with a grid of the mass function provided by Thomas Reiprich (private communication). For a comparison with Tinker et al. (2008), the calculation was performed with an overdensity relative to the mean density of the universe. Apart from this single case, the overdensity is denoted with respect to the Critical Density. Fig. 6.9 shows the plot from Tinker et al. (2008) for $\Delta = 200, 800, \text{ and } 3200$ at $z = 0$. Tinker et al. (2008) define the overdensity Δ with respect to the mean matter density and use the *WMAP1* cosmological parameters. These parameters were adopted for the calculations performed for the comparison.

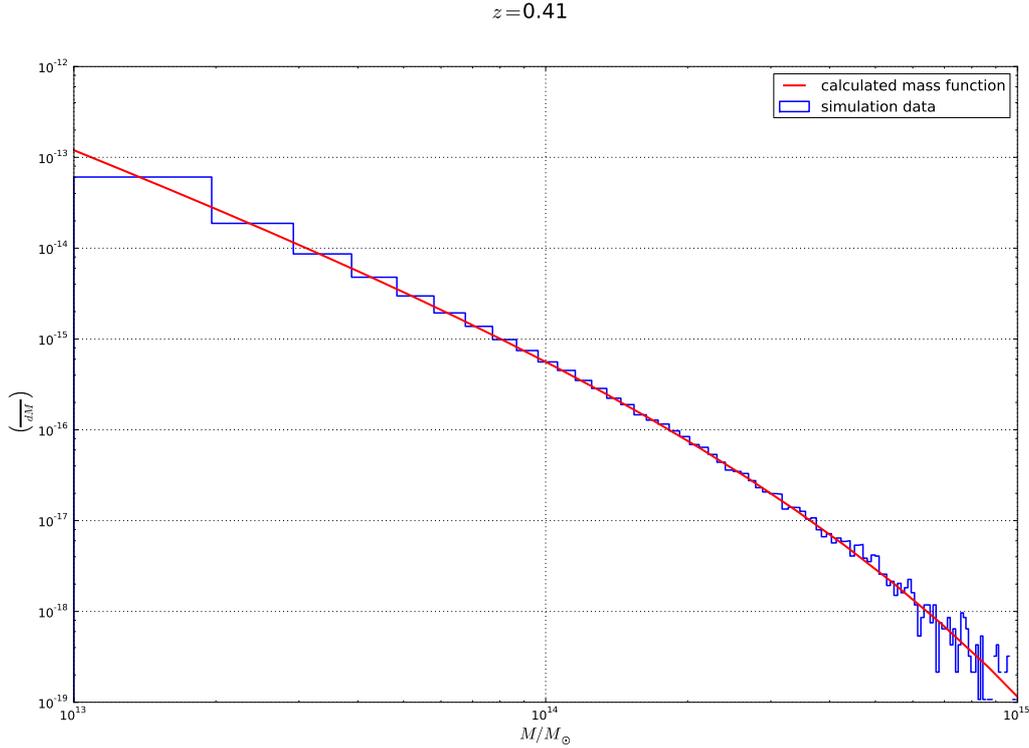


Figure 6.8: Normalized mass histogram for all objects with $0.36 \leq z \leq 0.46$, plotted against the normalized cluster mass function

The line in blue represents the calculations performed with the code presented in this thesis. The agreement is excellent.

Because of the redshift evolution of the mass function it is not sufficient to test the code for $z = 0$. To test the function up to $z = 2.5$, a comparison to a grid of the mass function provided by Thomas Reiprich was performed (Fig. 6.10). There are deviations up to $\sim 30\%$, which can probably be explained by the use of another transfer function (own calculations: Bardeen et al. (1986), Reiprich: Eisenstein & Hu (1998)), and by use of the approximation formulae for the fit parameter in the mass function as shown in Sect. 6.3 instead of the splines given by Tinker et al. (2008) as an alternative way to calculate the redshift evolution. The systematics in the deviations speak for this explanation, too. However, for the first tests the deviations are acceptable, even if improvements are desirable. Another test was performed to check the quality of the grid. For random values for the redshift in the interval $0 \leq z \leq 2.5$ and for the mass in the interval $10^{13} M_{\odot} \leq M \leq 10^{17} M_{\odot}$, the mass function was as well calculated directly as interpolated from a grid with 98×1500 nodes. For high masses, the deviation between the values obtained from interpolating the grid and the actual value increases, because the distance of the nodes increases (as mentioned above, they are spaced logarithmically), while simultaneously the dynamic of the mass function increases because the function is well over the exponential cutoff. So all values with $\left(\frac{dN}{dM}\right) < 10^{-8} \max\left(\frac{dN}{dM}\right)$ in the interval $10^{13} M_{\odot} \leq M \leq 10^{17} M_{\odot}$, $0 \leq z \leq 2.5$ were discarded for the evaluation of the grid quality. This is reasonable, because for values of the mass function this small compared to the maximal value virtually no clusters exist. The result of the comparison is shown in Fig. 6.11. For almost all values obtained by interpolation, the relative deviation to the values calculated directly is smaller than 1%. For $M \gtrsim 2 \cdot 10^{15} M_{\odot}$ all

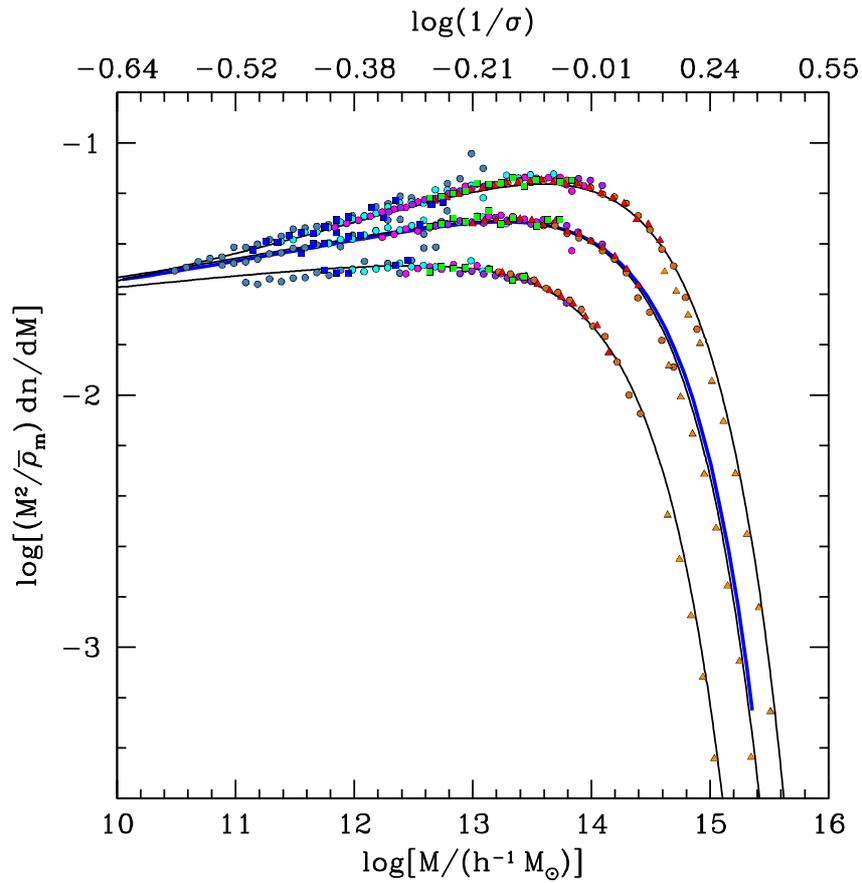


Figure 6.9: Mass function from Tinker et al. (2008) for the overdensities $\Delta = 200$, $\Delta = 800$, and $\Delta = 3200$ (top to bottom). The calculations performed for $\Delta = 800$ with the code presented in this thesis are shown in blue. The overdensity was defined with respect to the mean matter density, and the *WMAP1* cosmological parameters were used as by Tinker et al. (2008). The calculations agree excellently.

values are discarded by the mechanism described above.

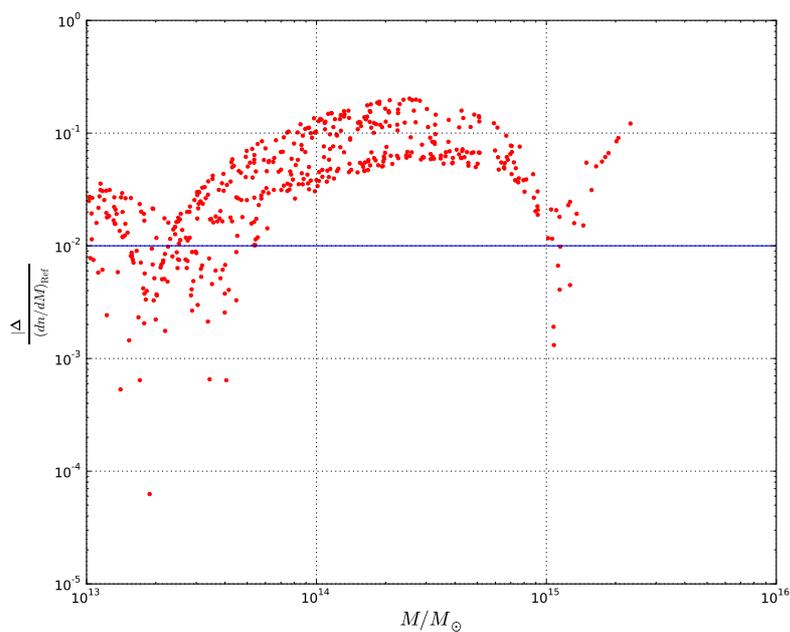


Figure 6.10: Comparison between mass function values obtained by linear interpolation of the reference grid obtained by Thomas Reiprich (private communication) and the grid calculated with the own code. The deviations go up to $\sim 30\%$, which is probably due to the use of another transfer function and the interpolation formulae for the fit parameters (Tinker et al., 2008) instead of the splines (Tinker et al., 2008).

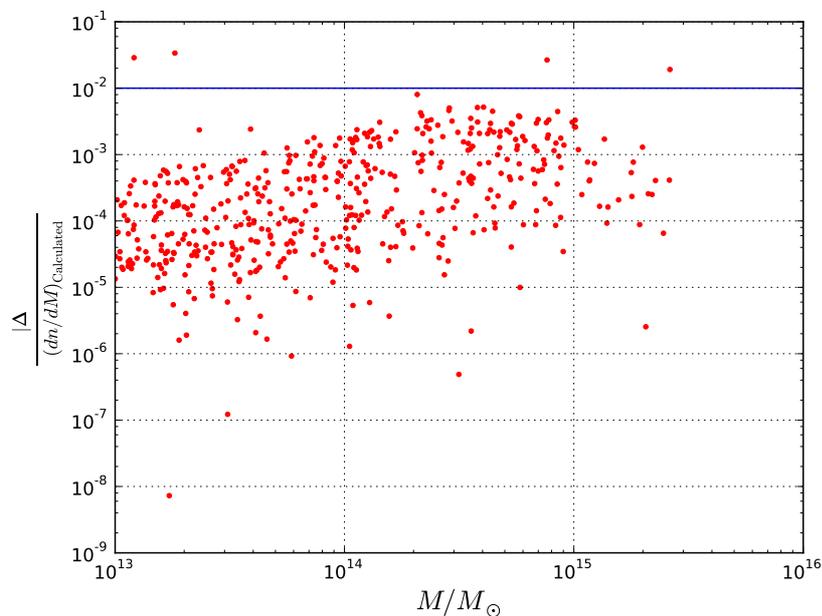


Figure 6.11: Comparison between mass function values obtained by linear interpolation of the grid and the values calculated directly. For almost all of the randomly generated points the deviation is smaller than 1% (marked by the blue line).

7. Summary and Conclusion

In this thesis, a Monte Carlo algorithm for the generation of a mock catalogue of galaxy clusters was presented. For this purpose, the halo mass function from Tinker et al. (2008) was used. The function reproduces the data obtained by N -body simulations up to redshifts of $z \leq 2$ and for halo masses in the range $10^{11}h^{-1}M_{\odot} \leq M \leq 10^{15}h^{-1}M_{\odot}$ (Tinker et al., 2008). The simulation presented here started at a lower limit of $10^{13}M_{\odot}$. The mass function can be well described by a power law up to a cutoff mass at roughly $10^{15}M_{\odot}$. The exact cutoff position depends on the redshift. Because for higher masses no clusters exist it was regarded as permissible to extend the mass range in our simulation over the $10^{15}h^{-1}M_{\odot}$ for testing purposes. Up to a redshift of 2.5, the integration of the mass function gives about 23 million halos with masses $\leq 10^{13}M_{\odot}$. With a Monte Carlo algorithm, the celestial coordinates, redshift and mass were sampled. From these variables the observables required to create a SIMPUT object list were derived: the cluster luminosity with the mass-luminosity relation from Vikhlinin et al. (2009) and a linear scaling factor for the *XMM-Newton* images. Each object was written to a SIMPUT file together with the link to an *XMM-Newton* image of a galaxy cluster scaled in size according to the redshift. This SIMPUT catalogue is used as an input catalogue for the *SIXT* simulation software presented in Sect. 5. An examination of the simulation results showed that the reproduction of the mass function was successful.

8. Outlook

Up to now, the SIMPUT catalogue was not yet used to run a simulation. This will be done soon, and the results will be compared with the estimates of 50,000 – 100,000 observed clusters by, e.g., Predehl et al. (2010). Further, it is important to implement the correlation function, because it contains important cosmological information, as was shown in Sect. 3.6. Because obtaining this information is a main objective of eROSITA, it is important to include this aspect in the simulation. Up to now the correlation was neglected because it requires a lot of calculation time, because the distance to each object which was already sampled has to be calculated and inserted into the correlation function. As mentioned above, a way out of this would be only to account for the objects in neighbouring redshift intervals, where the size of the interval has to be selected in such a way that there is virtually no correlation for the objects not taken into account. The numerical effort could be further decreased by partitioning the spherical shells, so that not all objects in one shell have to be used for correlation.

Another point which should be reviewed is the scaling of the *XMM-Newton* images: Because the images show clusters with different masses, the intrinsic diameter of the clusters is varying and Eq. 6.20 does not hold strictly. It is easily possible to implement a mass-dependent correction factor, because the radius scales as $M^{1/3}$ according to Eq. 3.7.

Since sources weaker than the flux limit of eROSITA are not observed, the total number of objects can be reduced by selecting only the objects over a flux limit for the catalogue. For this, a redshift-dependent mass limit has to be deduced and implemented in the code.

Small improvements in precision of the mass function could be done by using the splines from Tinker et al. (2008) for the calculation of the parameters in the mass function, and by using, as was suggested by Thomas Reiprich, the mass function from Eisenstein & Hu (1998) instead of Bardeen et al. (1986). Summing up, a suitable basis exists, but there is still enough room for improvements.

9. Danksagungen

Am Ende möchte ich all den zahlreichen Personen danken, die mich während der Anfertigung meiner Diplomarbeit und während meines Studiums unterstützt haben. An erster Stelle danke ich meinem Betreuer Jörn Wilms, der sich, trotz seiner unglaublichen Arbeitsbelastung, bei Fragen und Problemen stets Zeit für mich genommen hat. Auch Thomas Reiprich aus Bonn, der meine E-Mails meist noch am selben Tag beantwortete und mir damit half, manches Problem schnell zu lösen, gilt mein herzlicher Dank.

Ebenfalls nicht unerwähnt bleiben dürfen Ingo Kreykenbohm Christian Schmid, Christoph Großberger und Michael Wille die mir bei Fragen vielfältiger Art immer schnell weitergeholfen haben.

Für eine sehr angenehme Arbeitsatmosphäre danke ich meinen Bürokollegen Veronika Schaffenroth, Andreas Irrgang, Markus Firnstein und dessen Nachfolger, Raoul Gerber.

Meine Zeit an der Remeis Sternwarte wäre nur halb so schön gewesen ohne manchen gemütlichen Abend im wunderschönen Bamberg bei einem Krug Bier und angeregten Diskussionen. Dafür danke ich insbesondere Sebastian Müller, Stephan Geier, Matthias Kühnel, Lew Classen, Patrick Brünner, Moritz Böck, Andreas Konrad und, auch wenn oben bereits erwähnt, Christian Schmid und Andreas Irrgang. Jeden meiner Kollegen namentlich zu erwähnen würde den Rahmen sprengen, nichtsdestotrotz danke ich jedem Einzelnen für die großartige Zeit an der Remeis-Sternwarte. Dank Euch ist die Sternwarte nicht nur ein Ort an dem man gerne arbeitet, sondern auch Teile seiner Freizeit verbringt.

Dafür, dass diese Arbeit hoffentlich nur noch wenige Fehler enthält danke ich Andreas Dosche und Patrick Dürr, die Teile dieser Arbeit Korrektur gelesen haben sowie insbesondere Jens Rieger, der in äußerst knapper Zeit die Endkorrektur übernommen hat.

Nicht zuletzt gilt mein ganz besonderer Dank meiner Familie und insbesondere meinen Eltern, die mich immer optimal unterstützt und mich während der stressigen Phasen des Studiums ertragen haben. Wenn mein Vater mich nicht schon in früher Kindheit für die Schönheit des Sternenhimmels begeistert hätte, würde es diese Arbeit wahrscheinlich nicht geben.

References

- Bardeen J.M., Bond J.R., Kaiser N., Szalay A.S., 1986, *Astrophys. J.* 304, 15
- Bennett C.L., Bay M., Halpern M., et al., 2003, *Astrophys. J.* 583, 1
- Cappelluti N., Predehl P., Böhringer H., et al., 2011, *Memorie della Societa Astronomica Italiana Supplementi* 17, 159
- Carroll S.M., Press W.H., Turner E.L., 1992, *ARA&A* 30, 499
- Colless M., 1999, *Royal Society of London Philosophical Transactions Series A* 357, 105
- Davis M., Peebles P.J.E., 1983, *Astrophys. J.* 267, 465
- Deák I., 1990, *Random Number Generators and Simulation, Mathematical Methods of Operations Research* 4, Akadémiai Kiadó, Budapest
- Einstein A., 1916, *Annalen der Physik* 354, 769
- Einstein A., 1917, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)* 142–152
- Eisenstein D.J., Hu W., 1998, *Astrophys. J.* 496, 605
- Ernesti J., Kaiser P., 2008, *Python*, Galileo Press
- Friedrich P., Predehl P., Meidinger N., et al., 2005, In: O. Citterio & S. L. O'Dell (ed.) *Optics for EUV, X-Ray, and Gamma-Ray Astronomy II. Proc. SPIE 5900*, p.172
- Fürmetz M., Pfeiffermann E., Predehl P., et al., 2008, In: *Space Telescopes and Instrumentation 2008: Ultraviolet to Gamma Ray. Proc. SPIE 7011*
- Fürmetz M., Predehl P., Eder J., Tiedemann L., 2010, In: *Space Telescopes and Instrumentation 2010: Ultraviolet to Gamma Ray. Proc. SPIE 7732*
- Hanisch R.J., Farris A., Greisen E.W., et al., 2001, *Astron. Astrophys.* 376, 359
- Hunter J.D., 2007, *Matplotlib: A 2D Graphics Environment*
- Jones E., Oliphant T., Peterson P., et al., 2001, *SciPy: Open source scientific tools for Python*
- Komatsu E., Smith K.M., Dunkley J., et al., 2011, *Astrophys. J., Suppl. Ser.* 192, 18
- Kravtsov A.V., Klypin A.A., Khokhlov A.M., 1997, *Astrophys. J., Suppl. Ser.* 111, 73
- Mather J.C., Cheng E.S., Cottingham D.A., et al., 1994, *Astrophys. J.* 420, 439
- Mattig W., 1958, *Astron. Nachr.* 284, 109
- Meidinger N., Andritschke R., Ebermayer S., et al., 2009, In: *UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XVI. Proc. SPIE 7435*
- Misner C.W., Thorne K.S., Wheeler J.A., 1973, *Gravitation (Physics Series)*, W. H. Freeman

- Mo H., van den Bosch F., White S., 2010, *Galaxy Formation and Evolution*, Cambridge University Press, Cambridge
- Pavlinisky M., Sunyaev R., Churazov E., et al., 2009, In: *Optics for EUV, X-Ray, and Gamma-Ray Astronomy IV*. Proc. SPIE 7437
- Peacock J.A., 1999, *Cosmological Physics*, Cambridge University Press, Cambridge
- Peebles P.J.E., 1993, *Principles of Physical Cosmology*, Princeton University Press, Princeton
- Pence W.D., Chiappetti L., Page C.G., et al., 2010, *Astron. Astrophys.* 524, A42+
- Perlmutter S., Schmidt B.P., 2003, In: *Supernovae and Gamma-Ray Bursters. Lecture Notes in Physics*, Springer
- Piessens R., de Doncker-Kapenga E., überhuber C., Kahaner D., 1983, *QUADPACK: A Subroutine Package for Automatic Integration*, Springer Series in Computational Mathematics, Springer, Berlin
- Predehl P., 1999, In: O. H. Siegmund & K. A. Flanagan (ed.) *EUV, X-Ray, and Gamma-Ray Instrumentation for Astronomy X*. Proc. SPIE 3765, p.172
- Predehl P., Andritschke R., Böhringer H., et al., 2010, In: *Space Telescopes and Instrumentation 2010: Ultraviolet to Gamma Ray*. Proc. SPIE 7732
- Predehl P., Friedrich P., Hasinger G., 2003, In: J. E. Truemper & H. D. Tananbaum (ed.) *X-Ray and Gamma-Ray Telescopes and Instruments for Astronomy*. Proc. SPIE 4851, p.314
- Predehl P., Hasinger G., Böhringer H., et al., 2006, In: *Space Telescopes and Instrumentation II: Ultraviolet to Gamma Ray*. Proc. SPIE 6266
- Press W.H., Schechter P., 1974, *Astrophys. J.* 187, 425
- Reiprich T.H., Böhringer H., 2002, *Astrophys. J.* 567, 716
- Robertson H.P., 1935, *Astrophys. J.* 82, 284
- Sarazin C.L., 1986, *Rev. Mod. Phys.* 58, 1
- Schmid C., Martin M., Wilms J., et al., 2010, In: *X-ray Astronomy 2009*, Vol. 1248., p.591
- Schmid C., Smith R., Wilms J., , *SIMPUT – A File Format for Simulation Input*, Technical report, ECAP
- Solà J., 2011, *Journal of Physics Conference Series* 283, 012033
- Springel V., White S.D.M., Jenkins A., et al., 2005, *Nature* 435, 629
- Strüder L., Briel U., Dennerl K., et al., 2001, *Astron. Astrophys.* 365, L18
- Tinker J., Kravtsov A.V., Klypin A., et al., 2008, *Astrophys. J.* 688, 709
- Trümper J.E., Hasinger G., (eds.) 2008, *The Universe in X-Rays (Astronomy and Astrophysics Library)*, Springer, Berlin

Vikhlinin A., Burenin R.A., Ebeling H., et al., 2009, *Astrophys. J.* 692, 1033

Voges W., 1993, *Adv. Space Res.* 13, 391

Warren M.S., Salmon J.K., 1993, *Supercomputing '93* 12

A. List of Acronyms

- AGN** Active Galactic Nucleus
ART-XC Astronomical Roentgen Telescope – X-ray Concentrator
CCD Charge Coupled Device
CDF Cumulative Distribution Function
CMB Cosmic Microwave Background
COBE COsmic Background Explorer
DLR Deutsches Zentrum für Luft- und Raumfahrt
eROSITA extended ROentgen Survey with an Imaging Telescope Array
FITS Flexible Image Transport System
HEW Half Energy Width
IKI Space Research Institute of the Russian Academy of Sciences (originally: IKI RAN)
ICM Intra Cluster Medium
ISS International Space Station
MPE Max-Planck Institut für extraterrestrische Physik
NumPy Numerical Python
PDF Probabilty Density Function
PSF Point Spread Function
ROSITA ROentgen Survey with an Imaging Telescope Array
SciPy Scientific Python
SIMPOT Simulation Input
SIXT Simulation of X-ray Telescopes
SRG Spectrum-Roentgen-Gamma
WMAP Wilkinson Microwave Anisotropy Probe
XMM X-ray Multi-Mirror Mission

B. Typographic Conventions

The following typographic conventions were used in this thesis:

- Names of satellites and satellite missions are typeset in *slanted font* (e.g., *Spectrum-Roentgen-Gamma*)
- Names of programs and software libraries are typeset in *slanted font* (e.g., *SIXT*)
- Function names are typeset in *italic font* (e.g., *sample_catalogue*)
- Parameters and corresponding values are typeset in *typewriter font* (e.g., `Omega.m`)

ERKLÄRUNG

Hiermit erkläre ich, dass ich diese Diplomarbeit selbständig angefertigt und keine anderen als die angegebenen Hilfsmittel verwendet habe.

Bamberg, den 5. Juli 2011

(Johannes Hölzl)