Università degli Studi di Milano
Facoltà di Scienze Matematiche, Fisiche e Naturali
Dottorato di Ricerca in
Fisica, Astrofisica e Fisica Applicata

THERMODYNAMIC AND CHEMICAL PROPERTIES OF THE INTRA-CLUSTER MEDIUM IN THE OUTER REGIONS OF GALAXY CLUSTERS

Settore scientifico disciplinare: FIS/05

Coordinatore: Prof. Gianpaolo Bellini

Tutore: Dott. Silvano Molendi

Tutore interno: Prof. Giuseppe Bertin

Tesi di dottorato di:

Alberto Leccardi

Ciclo XXI

Anno Accademico 2008–2009

Contents

1 A concise overview of galaxy clusters **3**
- 1.1 General properties of clusters . . . 3
  - 1.1.1 Optical band . . . 3
  - 1.1.2 X-ray band . . . 5
- 1.2 The physics of the ICM . . . 6
  - 1.2.1 Emission mechanisms . . . 6
  - 1.2.2 Heating and cooling . . . 9
  - 1.2.3 Transport processes . . . 10
- 1.3 The spatial distribution of X-ray emission . . . 11
- 1.4 Spatially resolved spectroscopy . . . 14
  - 1.4.1 The temperature structure . . . 14
  - 1.4.2 The metallicity structure . . . 18
  - 1.4.3 The cluster classification: cool core vs. non-cool core . . . 21

2 The *XMM-Newton* satellite **25**
- 2.1 The X-ray telescopes . . . 25
  - 2.1.1 The point spread function (PSF) . . . 27
  - 2.1.2 The effective area . . . 28
- 2.2 European Photon Imaging Camera (EPIC) . . . 29

3 The EPIC background and its characterization **33**
- 3.1 The EPIC background . . . 33
  - 3.1.1 The internal background . . . 33
  - 3.1.2 The soft protons . . . 34
  - 3.1.3 The astrophysical background . . . 36
- 3.2 The background characterization . . . 36
  - 3.2.1 The analysis of "closed" observations . . . 36
  - 3.2.2 The analysis of "blank field" observations . . . 39

4 Data analysis **43**
- 4.1 Spectra preparation . . . 43
  - 4.1.1 Preliminary data preparation . . . 43
  - 4.1.2 Quiescent soft proton contamination . . . 44
  - 4.1.3 Spectra accumulation . . . 44
- 4.2 Spectral analysis . . . 44
  - 4.2.1 Estimate of background parameters . . . 45
  - 4.2.2 Spectral fit in concentric rings . . . 46

5 The bias on the ML temperature estimator **49**
- 5.1 The source-only case . . . 49
- 5.2 The source-plus-background case . . . 52
- 5.3 Attempts to correct the bias . . . 56

5.3.1 Using different estimators . . . 57
5.3.2 Fitting with a log-normal function . . . 58
5.3.3 A semi-empirical method: summing three distributions . . . 58

6 Radial profiles **61**
6.1 The sample . . . 61
6.2 Temperature and metallicity profiles . . . 65

7 Discussion of systematic effects **71**
7.1 Evaluation of systematics on temperature measures . . . 71
7.1.1 "A priori" tests . . . 71
7.1.2 "A posteriori" tests . . . 75
7.1.3 A budget for systematics . . . 78
7.2 Evaluation of systematics on metallicity measures . . . 80

8 Results **85**
8.1 The mean temperature profile . . . 85
8.1.1 Characterizing the profile . . . 85
8.1.2 Redshift evolution of temperature profiles . . . 87
8.1.3 Cool core and non-cool core clusters . . . 89
8.1.4 REFL04 and LP07 subsamples . . . 89
8.1.5 Comparison with hydrodynamic simulations . . . 90
8.1.6 Comparison with previous observations . . . 92
8.2 Metallicity profiles . . . 94
8.2.1 Cool core and non-cool core clusters . . . 94
8.2.2 Redshift evolution of metallicity profiles . . . 94
8.2.3 Comparison with hydrodynamic simulations . . . 96
8.2.4 Comparison with previous works . . . 97

9 Conclusions and future perspectives **99**
9.1 Main results . . . 99
9.2 Future perspectives . . . 100
9.2.1 The shock front in Abell 754 . . . 100
9.2.2 Deriving other physical properties . . . 101

10 Preliminary results on additional work in progress **103**
10.1 The sample and the analysis procedure . . . 104
10.2 Defining interesting quantities . . . 106
10.3 Chemical properties . . . 110
10.3.1 Chemical vs. thermodynamic quantities . . . 110
10.3.2 Metallicity profiles . . . 112
10.4 Conclusions . . . 113

Introduction

Clusters of galaxies are the most massive gravitationally bound systems in the universe. They are permeated by a hot, X-ray emitting, intra-cluster medium (ICM), which represents the dominant baryonic component. Galaxies contribute to the gravitating mass of the cluster only for few percent and the ICM for roughly 10%; the remaining is dark matter. The key ICM observable quantities are its density, temperature, and chemical composition (in particular, the so-called metallicity). Measuring the gas density is relatively simple, whereas measuring the temperature and the metallicity of the ICM requires accumulating a spectrum and fitting it with a plasma model. The aim of my thesis is the characterization of the radial properties (especially temperature and metallicity) of the ICM, for a large sample of clusters of galaxies observed with *XMM-Newton*. Particular attention was paid to investigate and characterize the systematic effects that may affect measurements.

The outline of the thesis is the following. In Chapter 1, I will give a non-exhaustive overview of the properties of galaxy clusters and introduce all concepts used to obtain my results. Although clusters emit broadband spectrum electromagnetic radiation from radio to gamma rays, in this thesis I will focus almost exclusively on the X-ray emission from the intra-cluster medium. In Chapter 2 I will give a brief overview of the main characteristics of the *XMM-Newton* satellite and its on-board instrumentation (in particular EPIC). In Chapter 3 I will characterize the EPIC background properties, obtained from my detailed analysis of several "blank field" and "filter closed" observations. In Chapter 4 my data analysis procedure, from the spectra preparation to the spectral analysis, will be outlined. In Chapter 5 I will describe simulations employed to examine how best to analyze spectra with poor statistical quality and find an unbiased estimator of the temperature; this chapter is based on Leccardi & Molendi (2007). In Chapter 6 I will report the sample properties and present the radial profiles for temperature and metallicity. In Chapter 7 I will describe the detailed analysis of a number of systematic effects. In Chapter 8 I will report the characterization of the temperature profile decline and the dependence of temperature and metallicity profile shape from physical properties (e.g. the redshift); I will also compare profiles with hydrodynamic simulations and previous observational works. In Chapter 9 I will summarize the results of this work and outline possible future perspectives. Finally, in Chapter 10 I will report the preliminary results obtained by investigating thermodynamic and chemical properties of cluster cores for an enlarged sample from $z = 0.02$ to $z = 0.25$; these results will be discussed in a forthcoming paper (Leccardi et al., in prep.). The bulk of my thesis (i.e. Chapters 4, 6, 7, 8, and part of Chapter 3) is based on two published papers, namely Leccardi & Molendi (2008a) and Leccardi & Molendi (2008b).

Chapter 1

A concise overview of galaxy clusters

1.1 General properties of clusters

1.1.1 Optical band

Galaxy clusters were first discovered in the optical as exceptional concentrations of *nebulae* by Charles Messier in 1784 and F. Wilhelm Herschel in 1785; however, only at the beginning of the XX century did it become clear that nebulae are extragalactic objects and the concentrations of nebulae were called *clusters of galaxies*. In the optical band, clusters of galaxies appear as high overdensity systems of galaxies.

By traditional definition (Abell 1958), clusters contain at least 50 galaxies brighter than $m_3 + 2$ (where m_3 is the magnitude of the third brightest cluster member), within a radius of $R \approx 1.5$ h^{-1} Mpc from the cluster center. This galaxy count is usually defined as the richness of the cluster. A number of various cluster properties (e.g. concentration of galaxies, symmetry, galactic content) have been used to classify them from the morphological point of view. Most classifications are highly correlated, so that clusters can be roughly represented as a one-dimensional sequence, running from regular to irregular systems (Abell 1965). Regular clusters tend to be compact, rich, spiral poor, and to have a central dominant galaxy. Such a morphological sequence suggests that regular clusters may be dynamically more relaxed.

The redshift of a cluster is determined from the mean radial velocity (along the line of sight) of its galaxies. The radial velocities, v_r, of the individual galaxies are distributed around this mean and, to first approximation, one can describe the velocity distribution through the dispersion,

$$\sigma_r = \langle (v_r - \langle v_r \rangle)^2 \rangle^{1/2}. \tag{1.1}$$

The velocity dispersion, thus calculated along the line of sight, is usually assumed as representative for other directions, through the "pressure isotropy" hypothesis.

The most regular clusters show a smooth spatial distribution of galaxies with a more concentrated core (see Fig. 1.1). A number of models have been proposed to fit this distribution. The simplest is the isothermal sphere, which assumes a gaussian and isotropic radial velocity distribution, independent of the position. The isothermal sphere describes fairly well cluster central regions, but fails at large radii because the total number of galaxies and the total mass would diverge linearly with the radius. King (1962) showed that the three-dimensional galaxy distribution is

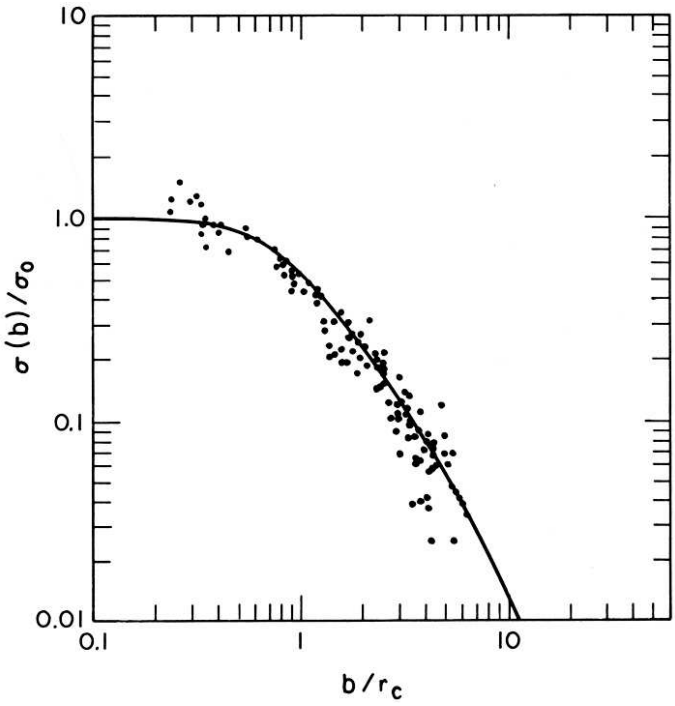

Figure 1.1: Projected galaxy number density (dots) observed in 12 regular clusters Bahcall (1975). The observed number densities are normalized to the central surface number density and given as a function of the projected radius, b, divided by the core radius, r_c.

better approximated by the analytic function,

$$n(r) = n_0 \left[1 + (r/r_c)^2\right]^{-3/2}, \tag{1.2}$$

where n_0 is the central density of galaxies and the core radius, r_c, is the size of the central core. Also for the King model, the cluster mass and the galaxy number diverge, although more slowly, at large radii; therefore, such density distribution must be truncated at some finite radius. Astronomers usually measure cluster properties in a limited region of radius, r_Δ, characterized by an over-density, Δ, with respect to the critical density of the universe; Δ usually spans the range between 100 and 2500. The greater the over-density, the smaller the region. The virial radius in the currently adopted ΛCDM cosmology is characterized by $\Delta \approx 100$ (Eke et al. 1998).

Regular clusters should have had the time to relax dynamically. Since this process involves the spatial motion of the galaxies, a lower limit to the relaxation time is the galaxy crossing time,

$$t_{cr} = \frac{R}{\sigma_r} \simeq 1 \left(\frac{R}{1\ \mathrm{Mpc}}\right) \left(\frac{\sigma_r}{10^3\ \mathrm{km\,s^{-1}}}\right)^{-1} \mathrm{Gyr}. \tag{1.3}$$

In a Hubble time, clusters that did not suffer major mergers should have had enough time to relax dynamically in their inner regions ($r \lesssim 1\ \mathrm{h^{-1}\ Mpc}$), while the cluster outskirts are expected to be less regular, as is indeed observed.

The total mass of a cluster can be estimated by measuring the galaxy velocity dispersion, σ_r (1.1), and assuming that clusters are bound, self-gravitating systems.

Under these conditions one can apply the virial theorem,

$$2K + W = 0, \tag{1.4}$$

where K and W are the sum over all galaxies of the kinetic energy of the individual galaxies and of the gravitational potential produced by other cluster galaxies:

$$K = \frac{1}{2}\sum_i m_i v_i^2, \tag{1.5}$$

$$W = -\frac{1}{2}\sum_{i\neq j}\frac{G m_i m_j}{r_{ij}}. \tag{1.6}$$

From (1.4) one can derive the cluster total mass,

$$M_{tot} = \frac{R_G\langle v^2\rangle}{G}, \tag{1.7}$$

where the mass-weighted velocity dispersion is

$$\langle v^2\rangle = \frac{\sum_i m_i v_i^2}{M_{TOT}} \tag{1.8}$$

and the gravitational radius is

$$R_G = 2\,M_{TOT}^2\left(\sum_{i\neq j}\frac{m_i m_j}{r_{ij}}\right)^{-1}. \tag{1.9}$$

The quantities $\langle v^2\rangle$ and R_G can be measured from the radial velocity distribution and the projected spatial distribution of a fair sample of galaxies. From (1.7) one can write:

$$M_{tot} = 7\times 10^{14} M_\odot \left(\frac{\sigma_r}{10^3\ \mathrm{km\,s^{-1}}}\right)^2\left(\frac{R_G}{1\ \mathrm{Mpc}}\right). \tag{1.10}$$

Since $\sigma_r \approx$ 700-800 km s^{-1} and $R_G \approx$ 1-3 Mpc, the total mass is on the order of 10^{14}-$10^{15}\,M_\odot$, which is 10-100 times greater than the mass of all the galaxies of a cluster. With a similar calculation using the galaxies in the cluster of Coma, Zwicky (1933) pointed out the *missing mass problem* and suggested the existence of *dark matter*, which provides enough mass to hold the cluster together. Such a problem was only partially overcome with the discovery of the X-ray emitting intra-cluster medium (see Sect. 1.3), which contains most of the baryons in a cluster.

1.1.2 X-ray band

In 1966, X-ray emission was detected from the region around the galaxy M87, at the center of the Virgo cluster (Byram et al. 1966), followed five years later by the detection in the direction of Coma (Fritz et al. 1971; Gursky et al. 1971a) and Perseus clusters (Meekins et al. 1971; Gursky et al. 1971b). Since these are the richest clusters known, Cavaliere et al. (1971) suggested that clusters of galaxies could be generally associated with X-ray emission; however, observational evidence was poor, because these early detections were all made with balloon- or rocket-borne detectors.

A great advance in the study of X-ray clusters (and in X-ray astronomy in general) was made with the launch of the first X-ray satellite, *Uhuru*, which gave us the first complete survey of the sky in X-rays (Giacconi et al. 1972). Observations

by *Uhuru* suggested that many clusters of galaxies are bright X-ray sources with luminosity on the order of 10^{43}-10^{45} erg s^{-1}.

The emission from these sources was extended, rather than point-like, and its size was comparable to that of the distribution of galaxies (Kellogg et al. 1972; Forman et al. 1972); moreover, there was no evidence for temporal variations and low-energy photo-absorption. The above mentioned characteristics indicated that the emission was truly diffuse, and not the result of the superposition of many compact sources.

Observations of the X-ray spectra of galaxy clusters have played a critical role in establishing the emission mechanisms. The two most popular mechanisms were the thermal bremsstrahlung from a hot diffuse intra-cluster gas (Felten et al. 1966) and the inverse Compton scattering from a population of relativistic electrons (Brecher & Burbidge 1972; Bridle & Feldman 1972; Harris & Romanishin 1974; Rephaeli 1977). To distinguish between the two expected spectral shapes (i.e. exponential cut-off and power law) high quality data were required. Actually, what established the thermal nature of the primary emission mechanism was the detection of line emission from highly ionized gas in the X-ray spectra of Perseus (Mitchell et al. 1976), Coma and Virgo (Serlemitsos et al. 1977).

It is now firmly established that the largest amount of luminous matter which we observe in clusters is the hot diffuse plasma which fills the deep potential well of clusters (Sarazin 1988).

1.2 The physics of the ICM

The intra-cluster medium (ICM) has temperature, T, and electron number density, n_e, on the order of 10^7-10^8 K and 10^{-2}-10^{-4} cm^{-3}, respectively. It contains not only primordial elements from the Big-Bang nucleosynthesis, but also a fraction of heavy elements, which must be produced by stars. At temperatures on the order of several million K, the ICM is almost fully ionized: Hydrogen and Helium are completely ionized, while heavier elements retain only a few of their electrons in their inner shells.

By using the typical densities of the ICM and the cross section for Thompson scattering ($\sigma_T = 6.65 \times 10^{-25}$ cm^{-2}), which is the main source of opacity under these conditions, one can calculate the optical depth, $\tau \simeq n_e \, \sigma_T \, R$, which is much smaller than unity for typical cluster sizes, $R \approx 1$ Mpc. Therefore, clusters are transparent to X-rays (i.e. optically thin) and even photons from the core region can be collected; however, as a consequence, the observed spectrum is the superposition of spectra generated at different depths (i.e. with different temperature and density). To derive the three-dimensional cluster properties, deprojection techniques have been developed; however, they usually require strong geometrical assumptions (e.g. spherical symmetry) which are not always fulfilled in galaxy clusters.

1.2.1 Emission mechanisms

Ionized plasma produces large amounts of X-ray photons, which allow us to observe and measure several physical properties of the emitting intra-cluster medium. The most important emission processes are bremsstrahlung and discrete line emission related to K- and L-shell transitions.

Thermal bremsstrahlung Bremsstrahlung radiation is due to the acceleration of free electrons in the Coulomb field of an ion. If the electrons have a thermal distribution with temperature T, and ions have charge Z the bremsstrahlung emissivity

at frequency ν is:

$$\epsilon_b(\nu) = \frac{2^5\pi e^6}{3m_e c^3}\left(\frac{2\pi}{3m_e k}\right)^{1/2} Z^2 n_e n_i\, g(Z,T,\nu)\, T^{-1/2} \exp\left(-\frac{h\nu}{kT}\right), \tag{1.11}$$

where e and m_e are the electron charge and mass, k and h are the Boltzmann and Planck constants, n_e and n_i are the electron and ion density (Rybicki & Lightman 1979). The Gaunt factor $g(Z,T,\nu)$ corrects for quantum mechanical effects and is a slowly varying function of temperature and frequency (Rybicki & Lightman 1979). The spectrum is rather flat at low frequency and has an exponential cutoff at $\nu_{cut} \simeq kT/h$, whose position provides an estimate of the electron temperature, T. Typical values for clusters are between 2 and 15 keV[1]. By integrating the emissivity (1.11) over all frequencies, over the solid angle of the source, and along the line of sight, one obtains the observed flux:

$$F = \int d\nu \int d\Omega \int dl\ \epsilon_b(\nu) \propto \frac{EI}{D_A^2}\, T^{1/2}, \tag{1.12}$$

where D_A is the angular distance to the source and EI is the emission integral

$$EI = \int dV n_e n_i, \tag{1.13}$$

which can be measured directly as the normalization of the observed spectrum. From the redshift of the source and a cosmological model, it is possible to determine D_A and, with some assumptions on symmetry and geometry, the volume V of the source. By making use of (1.13), one can estimate the average electron density, n_e, whose typical values are 10^{-2}-10^{-4} cm^{-3}, as mentioned before.

Discrete line emission Discrete emission results from atomic processes such as collisional excitation, radiative recombination, dielectric recombination and resonant excitation. In the intra-cluster medium, ionization and emission result primarily from collisions of ions with electrons (ion-ion collisions can be ignored). Moreover, the time scales for ionization and recombination are usually much smaller than the age of the cluster ($\approx$ the Hubble time) and any other relevant hydrodynamic time scale, thus the plasma is assumed to be in ionization equilibrium. All processes mentioned above have been incorporated in public available codes[2] (e.g. MEKAL in XSPEC), which are used to study collisionally ionized spectra.

For hotter clusters (k$T \gtrsim 3$ keV), the most prominent feature in the spectrum is a blend of lines around 7 keV (see Fig. 1.2, bottom panel), which is made up of the K_α fluorescent lines of the hydrogen-like Iron (Fe XXVI) at 6.9 keV and of the helium-like Iron (Fe XXV) at 6.7 keV. In cooler clusters, the Fe-K_α blend is less evident and the most prominent feature is the Fe-L complex, located around 1 keV (see Fig. 1.2, top panel), which is related to the L-shell transition of Iron in different ionization states. Weaker lines associated with other elements (e.g. S, Si, Mg, Ne, O, N, C) have also been detected. However, with the current observations, it is very difficult to measure elements other than Iron outside the core, where the temperature is relatively high and data are characterized by poor statistical quality.

It is remarkable that all the abundant elements, which were synthesized in stars after the primordial nucleosynthesis, have the energy of their K- and L-shell transitions in the spectral band accessible to modern X-ray telescopes. Most of the observed emission lines in the ICM arise from the well understood hydrogen- and

[1] Hereafter I will express the temperature in energy units (keV), setting to unity the Boltzmann constant.

[2] http://heasarc.nasa.gov/docs/xanadu/xspec/xspec11/index.html

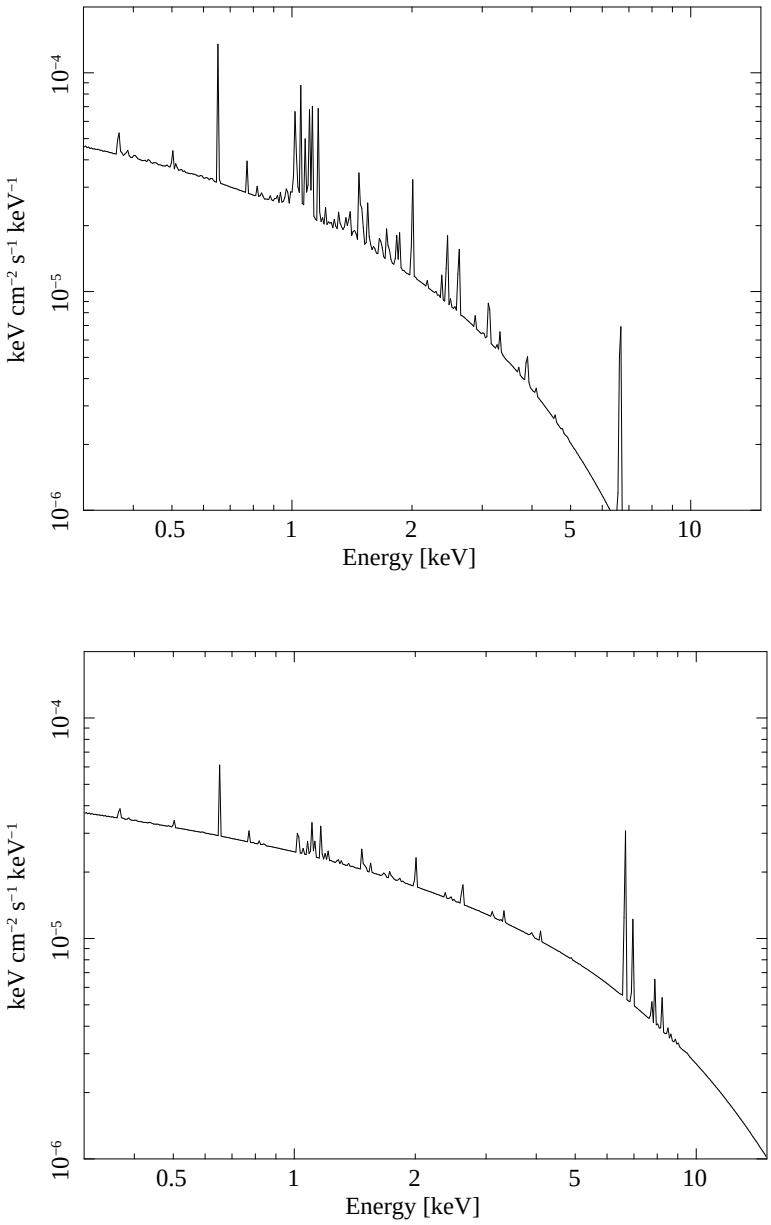

Figure 1.2: Simulated spectra produced with a MEKAL model in XSPEC. For the top panel the plasma temperature is 2 keV, for the bottom panel 6 keV. The metallicity is 0.3 solar (typical value for the ICM) and the normalization 10^{-4} in XSPEC units, for both spectra. Note the Fe-L complex around 1 keV in the spectrum of the cooler gas and the Fe-K_α complex at 6.7-6.9 keV in the spectrum of the hotter gas.

helium-like ions, and their equivalent widths can, under the reasonable assumption of collisional equilibrium, be directly converted into the elemental abundance of the corresponding element. The equivalent width, EW, is measured from the spectrum:

$$EW = \int_{\Delta\nu} \frac{I_{line(\nu)}}{I_c(\nu)}, \tag{1.14}$$

where $I_{line}(\nu)$ and $I_c(\nu)$ are the intensity of the line and of the continuum emission at a frequency ν. Since the bremsstrahlung emission (1.11) is mainly due to the contribution of hydrogen ions, $I_c(\nu) \propto n_e n_H$, while $I_{line}(\nu) \propto n_e n_{X^i}$, thus the EW is proportional to the abundance of the X^i element (n_{X^i}/n_H). The typical mean abundance of the Iron derived in galaxy clusters is $n_{Fe}/n_H \approx 2 \cdot 10^{-5}$, about one-third of the solar value. I recall that the presence of heavy elements in the ICM proves that the intergalactic gas is not of primordial origin but must have been processed in stars. This requires mechanisms, which are not clear yet, to transport the metals from galaxies to the ICM.

1.2.2 Heating and cooling

As discussed in Sect. 1.2.1, the primary cooling process for the intra-cluster plasma is the emission of radiation by free-free (i.e. thermal bremsstrahlung) and free-bound transitions. The cooling time scale can be calculated as the ratio of the internal energy density to the X-ray emissivity (Eq. 1.11 for all the ions i):

$$t_{cool} \simeq \frac{n_e T}{\epsilon} = 9 \times 10^{10} \left(\frac{n_e}{10^{-3}\ \mathrm{cm}^{-3}}\right)^{-1} \left(\frac{T}{10\ \mathrm{keV}}\right)^{1/2} \mathrm{yr}. \tag{1.15}$$

This time scale is usually longer than the Hubble time, suggesting that radiative cooling cannot play a great role in the global physics of the ICM. However, in the center of some clusters, where the density is high, t_{cool} may be shorter than the Hubble time and cooling should be taken into account (Sect. 1.4.3).

Although the ICM is very hot, no major on-going heating is necessary because the cooling process is quite slow. The thermal energy was accumulated during the process of cluster formation and accretion. Let us consider a sub-cluster approaching and falling into the main cluster potential well, its potential energy is converted to kinetic energy, which is in turn converted to thermal energy, when the infalling gas is shocked by the collisions with the hot ICM. Infall and compression can produce temperatures such that

$$\frac{3}{2}\frac{kT}{\mu m_p} \approx -\Phi, \tag{1.16}$$

where Φ is the gravitational potential of the cluster. The infall velocity of the colliding subclusters is expected to be comparable with the escape velocity (Sarazin 2002), so that the motions are expected to be moderately supersonic and to drive shock waves into the ICM. Let v_s be the velocity of the shock wave relative to the pre-shock, undisturbed intra-cluster gas. The sound speed is

$$c_s = \left(\frac{5}{3}\frac{P}{\rho}\right)^{1/2} \sim 1000\ \mathrm{km\,s}^{-1}, \tag{1.17}$$

where P is the pressure of the gas and ρ its mass density. The Mach number of the shock,

$$M \equiv \frac{v_s}{c_s}, \tag{1.18}$$

is expected to be only slightly greater than unity. Numerical simulations of cluster mergers (e.g. Gabici & Blasi 2004) have confirmed that almost all shocks are expected to be found in galaxy clusters with $M < 3$.

1.2.3 Transport processes

Mean free paths The mean free paths of electrons and ions in a plasma without magnetic field are determined by Coulomb collisions and can be quantified as (Spitzer 1956):

$$\lambda_e = \frac{3^{3/2}(kT)^2}{4\pi^{1/2} n_e e^4 \ln\Lambda} \simeq 30 \left(\frac{T}{10\ \mathrm{keV}}\right)^2 \left(\frac{n_e}{10^{-3}\ \mathrm{cm}^{-3}}\right)^{-1} \mathrm{kpc}, \tag{1.19}$$

where the Coulomb logarithm $\ln\Lambda \approx 38$. The typical mean free path of electrons (protons have nearly the same value) is some tens of kpc, much smaller than the cluster size, indicating that the ICM is collisional and can be treated as a fluid satisfying hydrodynamic equations. It is worth noting, however, that the mean free path is comparable to the size of a galaxy. Therefore, in the interactions between the ICM and the individual galaxies, the plasma may be nearly collisionless.

Effects of the magnetic field It is now believed that the ICM is a weakly magnetized plasma. There are strong hints from X-rays and radio (measurements of synchrotron emission and Faraday rotation) that a large scale magnetic field exists, and the upper-limits on its strength are on the order of 10 μG. Very little is known about the origin and the structure of such magnetic fields; however, simple predictions can be made on the possible effects of a magnetic field in the ICM. Charged particles rotate around magnetic field lines following spiral trajectories with a radius (the gyroradius) of

$$r_g = \frac{m v_\perp}{Ze\,B}, \tag{1.20}$$

where m is the particle mass, $v_\perp$ is the component of its velocity perpendicular to the magnetic field, Ze is the particle charge, and B is the magnetic field strength. If $v_\perp = (2kT/m)^{1/2}$, which is the RMS value in a thermal plasma, then

$$r_g = \frac{3\times10^8\ \mathrm{cm}}{Z} \left(\frac{T}{10^8\ K}\right)^{1/2} \left(\frac{m}{m_e}\right)^{1/2} \left(\frac{B}{1\,\mu G}\right)^{-1} \tag{1.21}$$

which is much smaller than any length scale of interest in clusters, and is also much smaller than λ_e (1.19). Then, the effective mean free path for diffusion perpendicular to the magnetic field is only on the order of r_g^2/λ_e (Spitzer 1956). Due to their larger gyroradii, the ions are more effective in transport processes perpendicular to the magnetic field; however, in practice, the gyroradii are always so small that diffusion perpendicular to the magnetic field in the ICM can be ignored.

Equilibration time scales The electron-electron collision time associated with λ_e (1.19) is:

$$t_{e-e} = 3\times10^5 \left(\frac{T}{10\ \mathrm{keV}}\right)^{3/2} \left(\frac{n_e}{10^{-3}\ \mathrm{cm}^{-3}}\right)^{-1} \mathrm{yr}. \tag{1.22}$$

After few t_{e-e} an electron population is expected to relax to a Maxwell-Boltzmann distribution. The corresponding time scale for protons is

$$t_{p-p} \simeq (m_p/m_e)^{1/2}\, t_{e-e} \simeq 43\, t_{e-e} \tag{1.23}$$

and that necessary to reach equipartition (equal temperature) among both populations is

$$t_{e-p} \simeq (m_p/m_e)\, t_{e-e} \simeq 1870\, t_{e-e}. \tag{1.24}$$

Since all these time scales are shorter than the typical cluster age ($\approx$ Gyr), the energy distribution for ions and electrons should, to first order, be well described by a Maxwellian with a single temperature T.

Conduction In a plasma with a gradient in the electron temperature, heat is conducted down the temperature gradient. If the scale length of the temperature gradient is much longer than the mean free path of electrons, λ_e (1.19), as usual in galaxy clusters, the heat flux is given by

$$Q = -\kappa \nabla T, \tag{1.25}$$

where the thermal conductivity (Spitzer 1956) is:

$$\kappa = 1.31\, n_e\, \lambda_e\, k \left(\frac{kT}{m_e}\right)^{1/2} \approx 3 \times 10^{13} \left(\frac{T}{10\ \mathrm{keV}}\right)^{5/2} \mathrm{erg\,s^{-1}\,cm^{-1}\,K^{-1}}. \tag{1.26}$$

Thermal conduction is mainly due to electrons, because of the inverse dependence on the particle mass, and depends very strongly on the temperature. Within the intra-cluster medium, thermal conduction transfers heat from hot to cold regions in order to make the temperature constant, in absence of any other competing effect. There are examples (e.g. cold fronts, see Markevitch & Vikhlinin 2007, for a review) for which the conduction is strongly inhibited; one of the possible reasons is the presence of magnetic fields, which could play a role as discussed in Sect. 1.2.3.

Convection One possible source of mixing motions in the gas is convection. If the intra-cluster gas were hydrostatic but had a steep temperature gradient,

$$-\frac{d\ln T}{dr} > -\frac{2}{3}\frac{d\ln n_p}{dr}, \tag{1.27}$$

it would be unstable to convective mixing; (1.27) can also be expressed in terms of the entropy gradient,

$$\frac{dS}{dr} < 0. \tag{1.28}$$

If the entropy profile of the gas were declining, convective motions would start and mixing would occur within several sound crossing times in the cluster. This is a rather short time (1.3) and actually observations tell us that ICM entropy profiles have a positive gradient (e.g. Voit 2005; Pratt et al. 2006).

1.3 The spatial distribution of X-ray emission

In Sect. 1.2.3, it has been shown that the elastic collision times for ions and electrons (1.22) in the ICM are much shorter than any dynamical process, therefore the gas can be treated as a fluid. The time required for a sound wave in the ICM to cross a cluster is given by

$$t_s = 6 \times 10^8 \left(\frac{D}{1\ \mathrm{Mpc}}\right) \left(\frac{T}{10\ \mathrm{keV}}\right)^{-1/2} \mathrm{yr}, \tag{1.29}$$

where T is the temperature of the ICM and D the cluster size. The sound wave crossing time is on the same order of the galaxy crossing time (1.3) because the galaxies and the hot gas have roughly the same specific kinetic energy. Since this time is short with respect to the age of clusters and the cooling is not effective (see Eq. 1.15), the gas is assumed to be hydrostatic. The pressure distribution of the gas, P, is thus determined by the gravitational potential, Φ, through the equation of hydrostatic equilibrium:

$$\nabla P = -\rho_g \nabla \Phi, \tag{1.30}$$

where $P = \rho T/\mu m_p$, ρ is the gas mass density, and μm_p the mean density per particle. By assuming, in addition, that the cluster is spherically symmetric, (1.30) reduces to:

$$\frac{1}{\rho}\frac{dP}{dr} = -\frac{d\Phi}{dr} = -\frac{GM(r)}{r^2}, \tag{1.31}$$

where r is the radial coordinate (i.e. the distance from the cluster center) and $M(r)$ is the mass enclosed within a sphere of radius r.

The hydrostatic equilibrium is a reasonable assumption as long as the cluster is stationary (the gravitational potential does not change on a sound crossing time, i.e no major mergers in act), forces other than gas pressure and gravity (e.g. magnetic fields) are not important, and any motion of the gas is subsonic. Under these circumstances, by making use of (1.31) one can derive the total mass of the cluster, $M(r)$, from temperature and density profiles of the ICM (Sarazin 1988):

$$M(r) = -\frac{T(r)\,r}{G\mu m_p}\left(\frac{d\ln\rho}{d\ln r} + \frac{d\ln T}{d\ln r}\right). \tag{1.32}$$

It is not surprising that the temperature is a good tracer of the total mass, because it is strictly related to the depth of the gravitational potential well (1.16).

The cluster masses obtained through measurements based on X-ray and optical (e.g. using gravitational lensing) observations show a good agreement (e.g. Zhang et al. 2008). Despite the discovery of the ICM, which is responsible for most of the cluster baryons, the *missing mass problem* is not yet overcome, because the total mass is still larger than the mass of the ICM and all cluster galaxies. Observations of nearby and intermediate redshift, rich clusters show that the average mass fraction in stars (in galaxies and intra-cluster light combined) is roughly one-sixth of the X-ray gas mass fraction measured within the virial radius (Lin & Mohr 2004). In the central regions the contribution of galaxies to the visible mass is more significant, but the contribution to the total mass is at most of 3% at every radius. The X-ray gas mass fraction depends on the radius and spans the range between 5% and 15% (see e.g. Allen et al. 2008, and references therein). Summarizing, galaxies and ICM are responsible for $\approx 15\%$ of the total mass for rich clusters; the remaining $\approx 85\%$ appears to be *dark matter* (DM). The nature of DM is still under discussion.

The most recent and strongest evidence in favor of the existence of DM comes from the analysis of the "bullet" cluster, i.e. 1E 0657-56 (Markevitch 2006; Clowe et al. 2006). This is a spectacular example of an ongoing major merger of two massive galaxy clusters. The system is observed shortly after the first core-passage of the infalling sub-cluster, which moves approximately in the plane of the sky and is preceded by a prominent bow shock with Mach number ≈ 3 (1.18). By combining measurements from optical and X-ray band the authors found a displacement of the collisional ICM with respect to the non-collisional galaxies and the cluster potential well obtained with gravitational lensing (see Fig. 1.3). In theories without dark matter, such as Modified Newtonian Dynamics (MOND), the lensing would be expected to follow the baryonic matter, i.e. the X-ray plasma. However, the lensing is strongest in two regions separated from the visible matter, providing support for the idea that most of the mass in the cluster is in the form of collisionless DM. It is remarkable that both the first and the strongest indications in favor of the existence of DM are related to galaxy clusters.

Isothermal gas distributions: the β-model The simplest distribution of gas temperature would be an isothermal distribution. Let us consider both gas and galaxies in the gravitational potential[3] of the cluster, Φ, and assume spherical sym-

[3]The source of the potential (e.g. dark matter) is not important.

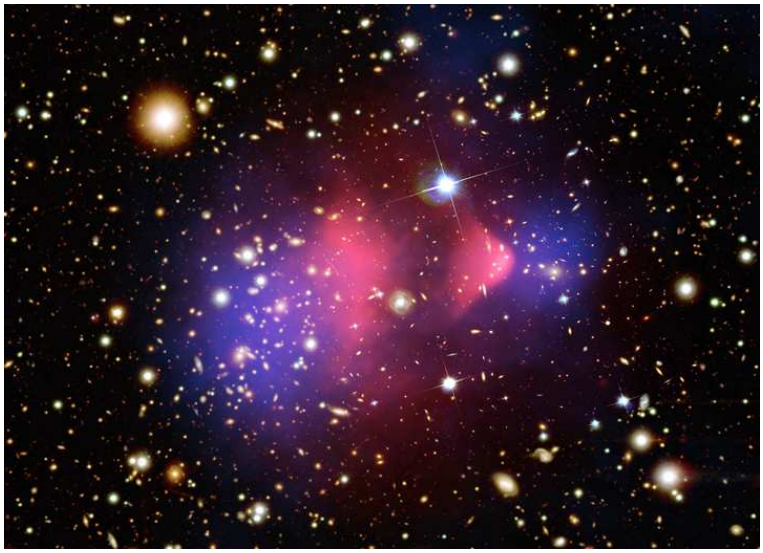

Figure 1.3: The "bullet" cluster 1E 0657-56. Superposed on the HST image, the density of the ICM from *Chandra* measurements is reported in red and the strength of the potential well from lensing measurements in blue.

metry. One can write the Jeans equations for both gas and galaxies

$$\frac{1}{\rho_{gas}}\frac{dP_{gas}}{dr} = -\frac{d\Phi}{dr}, \tag{1.33}$$

$$\frac{1}{\rho_{gal}}\frac{dP_{gal}}{dr} = -\frac{d\Phi}{dr}. \tag{1.34}$$

Since the right sides of (1.33) and (1.34) are equal, they cancel out together with the potential dependence. By giving an equation of state for gas and galaxies, $P_{gas} = \rho_{gas}\,\sigma^2_{gas}$ and $P_{gal} = \rho_{gal}\,\sigma^2_{gal}$, where $\sigma_{gas} = (kT/\mu m_p)^{1/2}$ and $\sigma_{gal} = \sigma_r$ both independent of r, and by defining the parameter $\beta = \sigma^2_{gal}/\sigma^2_{gas}$, one obtains the so-called β-model (Cavaliere & Fusco-Femiano 1978):

$$\rho_{gas} \propto \rho^{\beta}_{gal} = \rho_0 \left[1 + \left(\frac{r}{r_c}\right)^2\right]^{-3\beta/2}, \tag{1.35}$$

where the galaxy density in (1.2) is used.

This self-consistent isothermal model assumes that the gas and the galaxy distributions are both static and isothermal. Moreover, it assumes that the galaxy mass density ρ_{gal} is proportional to the galaxy density $n(r)$ in (1.2). Even if none of these assumptions is fully justified and the gas is not isothermal, the β-model has the advantage that the resulting gas distribution is analytic and that nearly all the integrals needed to compare the model to the observations are also analytic. For this reason, the β-model has been largely used to fit gas density profiles.

Adiabatic and polytropic gas distributions The ICM will be isothermal if thermal conduction is sufficiently rapid (see Sect. 1.2.3). On the other hand, if thermal conduction is slow, but the ICM is well-mixed, then the entropy per atom in the gas will be constant (see Sect. 1.2.3). In an adiabatic gas, pressure and density are simply related,

$$P \propto \rho^{\gamma}, \tag{1.36}$$

where γ is the usual ratio of specific heats and equals to 5/3 for a monatomic ideal gas. Although the value of 5/3 would be expected to apply if the ICM were strictly adiabatic, (1.36) is often used to parameterize the thermal distribution of plasma, with γ taken to be a fitting parameter. For example, $\gamma = 1$ implies that the gas distribution is isothermal. ICM models with an arbitrary value of γ are often referred to as "polytropic" models, and γ is called the polytropic index. Models with the polytropic index $\gamma > 5/3$ are convectively unstable (1.27); thus, hydrostatic polytropic models must have $1 \leq \gamma \leq 5/3$.

Hydrodynamical simulations Outside the central regions, where a not-well defined heating source should balance the radiative losses, the physics of the X-ray emitting ICM is expected to be mainly driven by adiabatic compression and shocks, taking place during the collapse of the cosmic baryons into accreting dark matter halos (see Sect. 1.2.2). Such processes can be properly treated in hydrodynamical simulations, using Tree+SPH codes like e.g. GADGET-2 (Springel 2005), and allow us to investigate cluster outskirts.

Roncarelli et al. (2006) used a set of hydrodynamical simulations adopting four different physical prescriptions for nine galaxy clusters, and found that the volume density profile for the gas steepens in the outskirts changing the slope of the power law from ≈ 2.5 to ≈ 3.4 at roughly 1.2 R_{200}. Moreover, they found that the behavior of the profiles in the external regions of clusters does not depend significantly on the presence or absence of cooling and supernova feedback, confirming the expectation that the mean behavior of the ICM in the cluster outskirts is mainly due to the gravitational force.

1.4 Spatially resolved spectroscopy

Up to 1990, the available instruments (i.e. *HEAO 1*, *Einstein*, *EXOSAT* and *Ginga*) could provide only global measurements of cluster temperature and metallicity; for simplicity galaxy clusters were assumed to be roughly isothermal. With the advent of *ROSAT*, *ASCA*, and *BeppoSAX*, and more recently of *Chandra* and *XMM-Newton*, it became possible to perform spatially resolved spectroscopy and obtain radial profiles and two-dimensional maps of both temperature and metallicity. This allowed us to discover and observe the effects of important phenomena such as cooling flows, merger shocks, cold fronts, etc. (see Peterson & Fabian 2006; Markevitch & Vikhlinin 2007, for extensive reviews), with a great progress in the study of the physical processes in the ICM. Moreover, by assuming hydrostatic equilibrium, deprojected radial temperature profiles allowed an X-ray measurement of the cluster total mass (see Sect. 1.3). The measure of the total mass allows us not only to have an independent confirmation of the missing mass (i.e. the *dark matter*), but also to use clusters as cosmological tools, because the mass is the main parameter on which all theoretical models of structure formation are based.

1.4.1 The temperature structure

Temperature profiles are usually obtained by extracting spectra in concentric annuli centered on the X-ray emission peak (see Chapter 4); then, they are rescaled by the mean cluster temperature and a scale length (the virial radius, R_{180}, R_{200}, etc.; see Sect. 6.2). In Fig. 1.4 the mean temperature profiles for two samples of clusters observed with *BeppoSAX* and *XMM-Newton* are reported. The first sample is further divided into two groups: those showing a temperature decline towards the center, i.e. the cool core clusters, and those showing a roughly isothermal core, i.e. the non-cool core clusters (see Sect. 1.4.3 for further details). A temperature drop

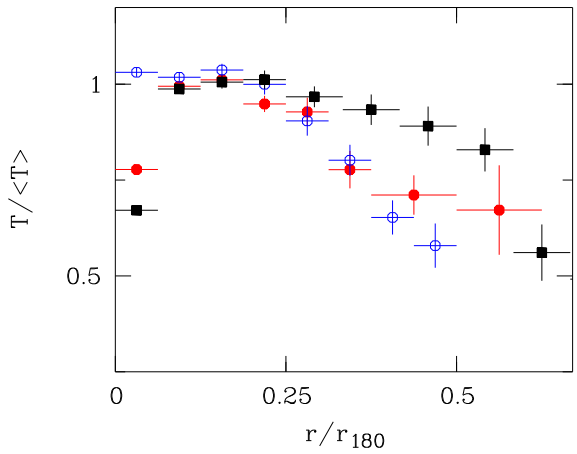

Figure 1.4: Mean temperature profiles from *XMM-Newton* (black square) and *BeppoSAX* (circles) data. The *BeppoSAX* sample is divided in cool core (red filled circles) and non-cool core (blue empty circles) clusters.

in the core is expected when the cluster is not disturbed by major mergers and has the time to relax.

Beyond $\approx 0.2\ R_{180}$ all profiles show a decline. The cluster outer regions are rich in information and interesting to study, because clusters are still forming there by accretion (e.g. Tozzi et al. 2000; Borgani et al. 2004). Moreover, far from the core it is easier to compare simulations with observations, because non-gravitational effects (often called "feedback") are less important (e.g. Borgani et al. 2004; McNamara et al. 2005; Roncarelli et al. 2006), so that the cluster radial profiles can be used to test the actual theories of the structure formation. Unfortunately, while cluster surface-brightness rapidly declines with radius (1.35), background (of instrumental, solar, local, and cosmic origin) is roughly constant over the detector (see Chapter 3). For this reason, spectra accumulated in the outer regions are characterized by poor statistics and high background, especially at high energies, where the instrumental background dominates other components (see Fig. 4.1). Such conditions make temperature measurements at large distances from the center a technically challenging task, requiring an adequate treatment of both statistical and systematic errors (see Chapter 5).

Given the technical difficulties, early measurements of cluster temperature profiles have been controversial. At the end of the *ASCA* and *BeppoSAX* era, the shape of the profiles at large radii was still the subject of debate (Markevitch et al. 1998; Irwin et al. 1999; White 2000; Irwin & Bregman 2000; Finoguenov et al. 2001; De Grandi & Molendi 2002).

Markevitch et al. (1998) analyzed azimuthally averaged radial temperature profiles for 30 clusters observed with *ASCA*, finding that nearly all clusters show a significant temperature decline at large radii (see Fig. 1.5, left panel). By fitting the composite temperature profile for symmetric clusters with a polytropic relation up to roughly half of the virial radius, the authors found that the temperature decline corresponded to a polytropic index (see Sect. 1.3) of 1.24 on average.

However, doubts on the universality and the steepness of this profile have been raised by subsequent studies. Irwin et al. (1999) compared results by Markevitch et al. (1998) with those by other authors (see references in Irwin et al. 1999), and found isothermal temperature profiles, even for those clusters where Markevitch et al. found a decline. After the analysis of a sample of 106 clusters observed

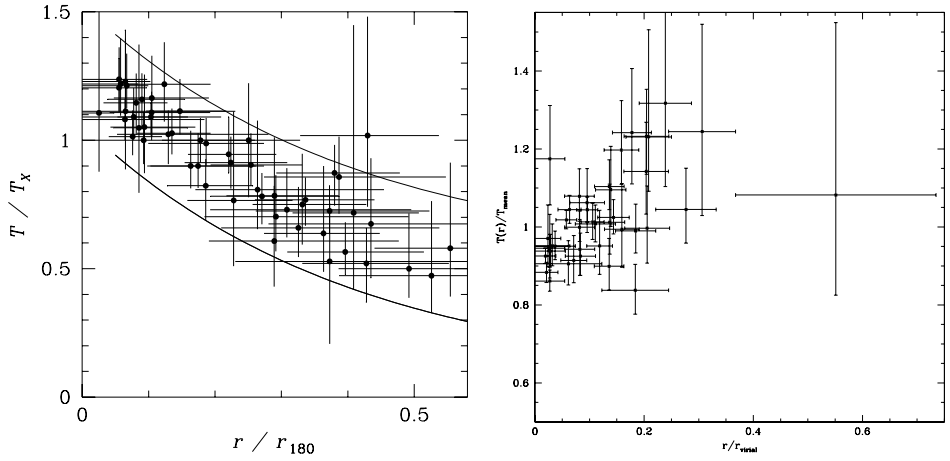

Figure 1.5: Left panel: radial temperature profiles for 30 clusters observed with *ASCA* (Markevitch et al. 1998). Right panel: radial temperature profiles for 11 nearby clusters observed with *BeppoSAX* (Irwin & Bregman 2000).

with *ASCA*, White (2000) concluded that 90% of the temperature profiles were consistent with being flat; however, some of the differences between White (2000) and other works can be attributed to White's use of a PSF model that overestimated scattering at low energies. Moreover, White typically did not extend measurements to large radii because of the large uncertainty inherent in his image deconvolution method; in the overlapping radial range most of his temperature profiles are in fact consistent with Markevitch et al. (1998). For a sample of cool clusters (i.e. k$T \gtrsim$ 4 keV) observed with *ASCA* however, Finoguenov et al. (2001) derived temperature profiles similar to those obtained by Markevitch et al. (1998), although the radial range explored was smaller.

Although *ASCA* has been the first X-ray instrument able to perform spatially resolved spectroscopy in hot clusters, given its adequate energy range (1–10 keV), its large and strongly energy-dependent point-spread function (PSF) required complicated correction procedures for the spectral analysis of extended sources. Different works on temperature measurements with *ASCA* have applied different methods to correct for the PSF effects. In this light *BeppoSAX* was a more suitable instrument to investigate temperature structures in galaxy clusters and supplied an independent dataset with respect to *ASCA*. The Medium-Energy Concentrator Spectrometer (MECS) on board *BeppoSAX* works in a similar energy range of *ASCA*, but has a sharper PSF (HPR $\approx 1'$), which is radially symmetric and almost energy independent. Irwin & Bregman (2000), who analyzed a sample of 11 clusters observed with *BeppoSAX*, placed further arguments supporting an isothermal gas. They claimed that the temperature profiles were generally flat or even increase slightly out to $\approx$ 30% of the virial radius (see Fig. 1.5, right panel); however, De Grandi & Molendi (2002) pointed out a technical error in their data analysis. Instead, De Grandi & Molendi (2002) analyzed 21 nearby clusters and found declining temperature profiles outside the core (see Fig. 1.6), in good agreement with results by Markevitch et al. (1998).

Recent observations with current telescopes (i.e. *XMM-Newton* and *Chandra*) have clearly shown that cluster temperature profiles decline beyond 15-20% of the virial radius (Piffaretti et al. 2005; Vikhlinin et al. 2005; Pratt et al. 2007). I report, as an example, the temperature profiles obtained by Vikhlinin et al. (2005) for 13

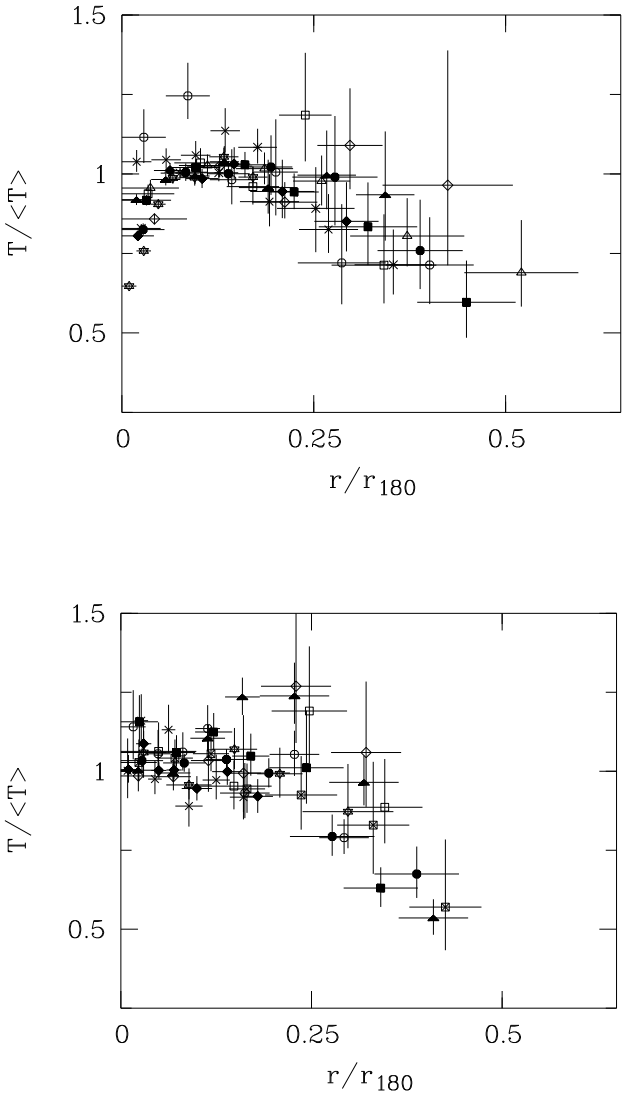

Figure 1.6: Radial temperature profiles for 21 clusters observed with *BeppoSAX* (De Grandi & Molendi 2002). Top panel: profiles for cool core clusters; bottom panel: profiles for non-cool core clusters.

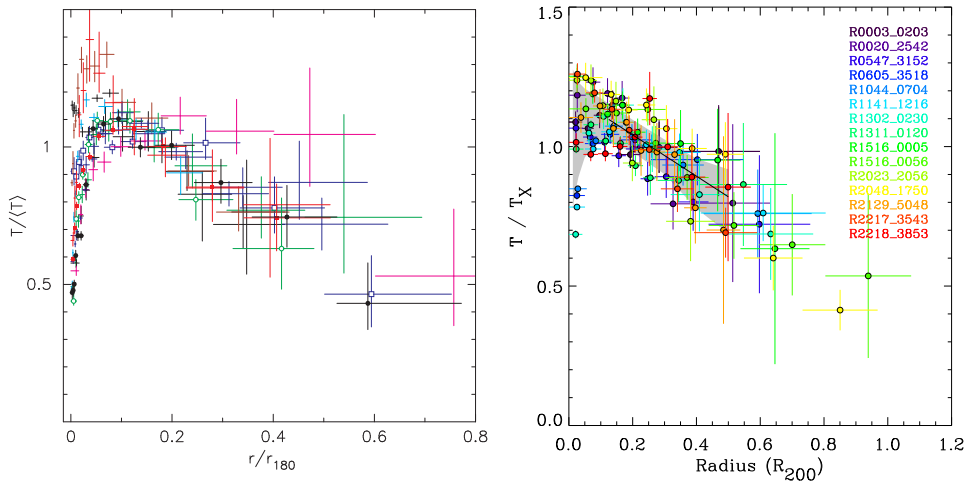

Figure 1.7: Left panel: radial temperature profiles for 13 nearby clusters observed with *Chandra* (Vikhlinin et al. 2005). Right panel: radial temperature profiles for 15 nearby clusters observed with *XMM-Newton* (Pratt et al. 2007).

nearby clusters observed with *Chandra*, and those obtained by Pratt et al. (2007) for a sample of 15 nearby clusters observed with *XMM-Newton* (see Fig. 1.7). However, most of these measurements might still be unreliable at very large radii (larger than one-half of the virial radius), because they are probably affected by a number of systematics related to the analysis technique and the background treatment.

The main purpose of my thesis is not only to measure the mean temperature profile for a large sample of clusters, exploiting at best the instrumentation on-board *XMM-Newton*, but also to investigate with particular attention where the systematic errors come from and how large they are, by making use of extensive montecarlo simulations.

With the advent of *Chandra* and *XMM-Newton*, also detailed two-dimensional temperature maps of galaxy clusters have been produced (e.g. Finoguenov et al. 2005; Fabian et al. 2006; Rossetti et al. 2007), showing complex structures (e.g. spiral features) in the ICM. Great efforts have been devoted to compare observed 2D temperature maps with those produced from hydrodynamic simulations (e.g. Tittley & Henriksen 2005; Ascasibar & Markevitch 2006).

1.4.2 The metallicity structure

The discovery of the Fe-K line emission in the spectrum of the Perseus cluster by the *Ariel V* satellite (Mitchell et al. 1976) and in Coma and Virgo by *OSO-8* (Serlemitsos et al. 1977) confirmed that the X-ray emission of galaxy clusters is predominantly thermal radiation from hot intra-cluster plasma rather than inverse Compton radiation (see Sect. 1.1.2). These observations also showed that the ICM is made up by a significant fraction of processed gas, which was ejected from stars in the cluster galaxies. This was not a priori obvious because, in principle, the ICM could be made of Hydrogen and Helium only, since clusters are the last structures in the universe to have clearly decoupled from the Hubble flow. The presence of heavy elements establishes an important connection between galaxies and ICM; however, while the origin of metals is clearly related to stellar processes and supernova explosions, the mechanisms through which heavy elements are transferred from stars

in galaxies to the ICM is not clear yet. Such mechanisms are extremely important because they give us a link between the hot diffuse baryons in the ICM and the cold gas locked in the stellar phase.

Most of the metals from O up to the Fe-group are produced by supernovae. The supernovae can be broadly divided into two groups: Type Ia supernovae (SN Ia) and core collapse supernovae (SNcc). SN Ia are most likely thermonuclear explosions of accreting white dwarfs. When the white dwarf reaches the Chandrasekhar limit, carbon ignition in the central region leads to a thermonuclear runaway. SN Ia produce a large amount of Fe, Ni, and Si-group elements (Si, S, Ar, and Ca), but only very small amounts of O, Ne, and Mg. SNcc occur when the iron core of massive stars collapses under its own gravity and produce a large amount of O, Ne, and Mg, unlike SN Ia. The time scales for the formation of the two types of supernovae are also different. Already after $\approx 10^7$ years the most massive stars exploded as SNcc, rapidly causing massive enrichment through galactic winds. Instead white dwarfs formed and in some cases produced SN Ia after $\approx 10^8$ years. By measuring the relative abundance of various elements, one can investigate which type of supernova enrichment is dominant (e.g. de Plaa et al. 2007).

First spectroscopic analyses of cluster samples revealed that the ICM has an Iron abundance of about one-third of the solar value (Mushotzky 1984). By combining spectra obtained by *Einstein* and *Ginga*, White et al. (1994) found the first indication of a centrally enhanced metallicity in four cooling flow clusters. Until the launch of *ASCA* in 1993, Iron was the only element for which the abundance was accurately measured in a large number of clusters. *ASCA* allowed us to detect the emission features from O, Ne, Mg, Si, S, Ar, Ca, Fe, and Ni in the spectra of a number of clusters. Moreover, *ASCA* data clearly showed that in cool core clusters the metallicity increases towards the center (Fukazawa et al. 1994), and that SN Ia are responsible for the strongest contribution to the cluster core enrichment. *ASCA* observations also allowed us to accurately determine the mean iron abundance out to $z \approx 0.5$, revealing a lack of evolution out to $z \approx 0.4$ (Mushotzky & Loewenstein 1997). However, its large and energy dependent PSF did not allow us to investigate the spatial abundance distribution in detail.

De Grandi & Molendi (2001) took advantage of the better spatial resolution of *BeppoSAX* and measured the radial iron abundance profiles for a sample of 17 rich, nearby clusters of galaxies. They found that the eight non-cool core clusters in their sample have rather flat or slightly decreasing profiles, while the iron abundance presents an enhancement in the central regions of cool core clusters (see Fig. 1.8). De Grandi et al. (2004) showed that the iron mass associated with the central abundance excess can be entirely produced by the brightest cluster galaxy (BCG) observed in almost all cool core clusters.

Observations with *XMM-Newton* and *Chandra* confirmed the central peak and the flat distribution of metals for cool core and non-cool core clusters respectively (Vikhlinin et al. 2005; Pratt et al. 2007; Tamura et al. 2004; Baldi et al. 2007), previously found with *BeppoSAX*; however, how the central metallicity gradients are distributed for a representative sample of galaxy clusters is still an open issue. *XMM-Newton* and *Chandra* also confirmed the strong contribution of SN Ia to the enrichment of cluster cores found by *ASCA*. By using *XMM-Newton* data, Böhringer et al. (2004a) found that long enrichment times (> 5 Gyr) are needed to produce the observed central abundance peaks. As mentioned in Sect. 1.4.1 for the temperature measurements, the outer regions (i.e. beyond 0.2 R_{180}), although information rich, are much less studied than the cluster core, as far as the metallicity is concerned. In such regions, comparing observed profiles with those obtained from hydrodynamic simulations is straightforward, because the complicated phenomena related to the core are almost negligible.

One of the main results of my thesis work is the first measure, after the one

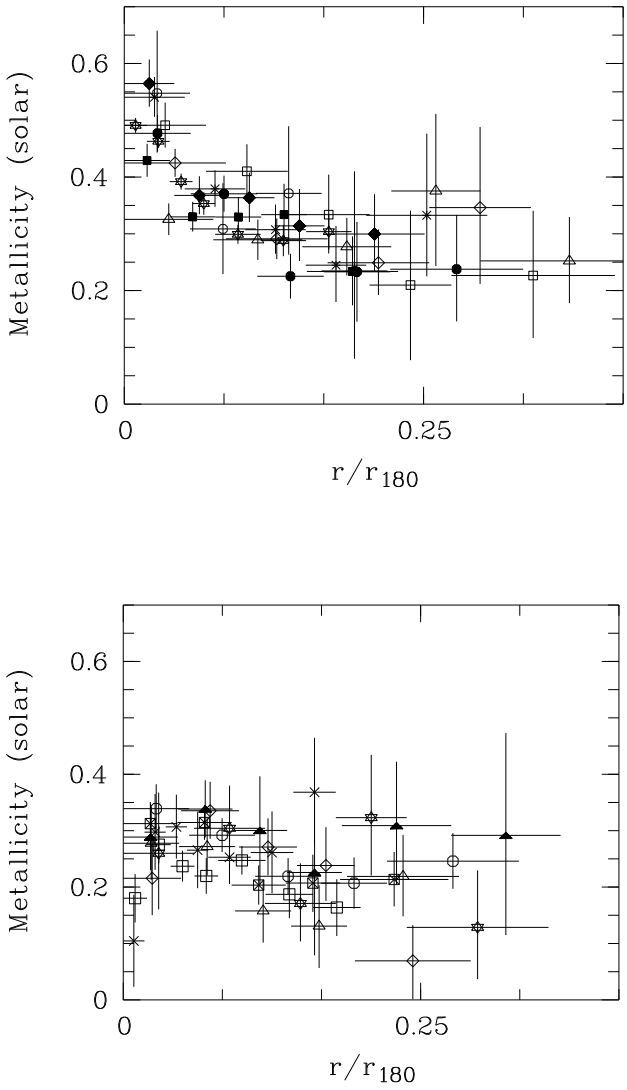

Figure 1.8: Radial metallicity profiles for 17 clusters observed with *BeppoSAX* (De Grandi & Molendi 2001). Top panel: profiles for cool core clusters; bottom panel: profiles for non-cool core clusters.

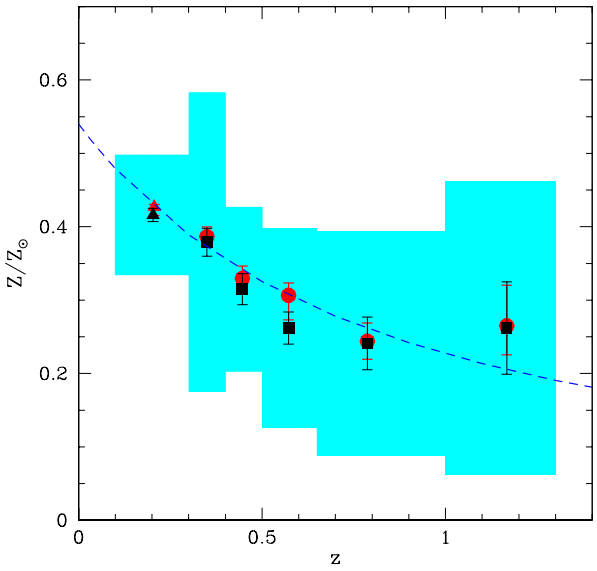

Figure 1.9: Evolution of the mean iron abundance (red circles) with redshift from Balestra et al. (2007). Error bars refer to the one-sigma confidence level. Shaded areas show the RMS dispersion. The dashed line indicates the best fit with a power law.

obtained with *BeppoSAX*, of the mean metallicity profile for a large, representative sample of galaxy clusters out to 0.4 R_{180}.

Recent *XMM-Newton* and *Chandra* measurements are further expanding the redshift range of precise iron abundance measurements in clusters. Balestra et al. (2007) analyzed *Chandra* data of 56 clusters with $0.3 \leq z \leq 1.3$ and found that the iron abundance within $\approx 0.3\,R_{vir}$ significantly declines with redshift out to $z \approx 0.5$, remaining then constant out to $z \approx 1.3$ (see Fig. 1.9). Maughan et al. (2008) analyzed *Chandra* data of 115 clusters with $0.1 \leq z \leq 1.3$ and found that the abundance evolution is still present (although less significant) even if the core regions are excluded, indicating that the observed evolution is not only due to a possible variation of the number of cool core clusters with redshift. However, the abundance evolution is still under discussion, because the analysis of low metallicity, statistically poor spectra could be affected by a number of systematics (see e.g. Sect. 7.2), which have not yet been well investigated.

Deep observations of bright, nearby clusters of galaxies with *XMM-Newton* and *Chandra* also allowed to map the 2D distribution of metals in the ICM (e.g. Sanders et al. 2004; Sauvageot et al. 2005; Sanders & Fabian 2006). The distribution of metals in clusters is usually not azimuthally symmetric, and in several cases shows complex patterns (see Fig. 1.10).

1.4.3 The cluster classification: cool core vs. non-cool core

Observations show that the X-ray emission from many clusters of galaxies is sharply peaked around the central brightest galaxy (see Fig. 1.11). The inferred radiative cooling time of the gas in that peak (1.15) is much shorter than the age of the cluster, suggesting the existence of a cooling flow there (Fabian et al. 1994). The

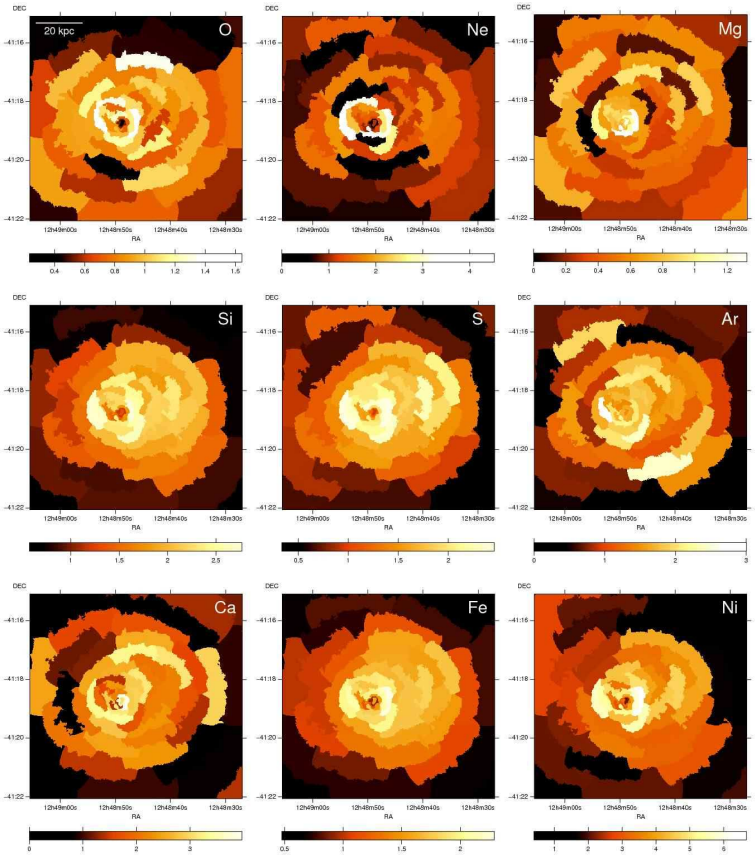

Figure 1.10: Abundance map for several elements in the Perseus cluster obtained by Sanders & Fabian (2006) with *Chandra* data.

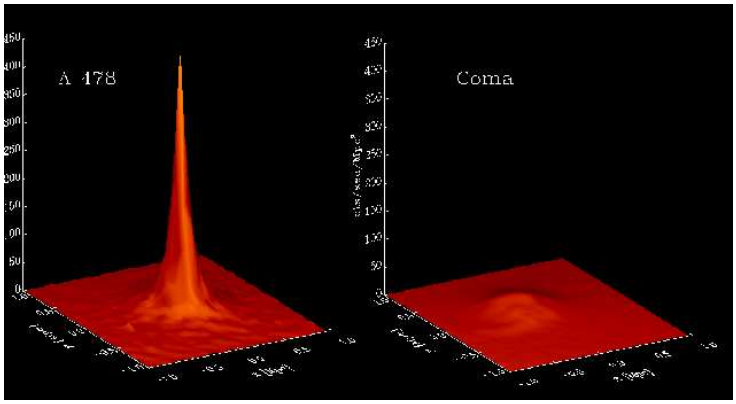

Figure 1.11: Comparison of the X-ray surface-brightness between a cool core cluster (A478, left panel) and a cluster without a cool core (the Coma cluster, right panel). Although the luminosity of the two clusters is roughly the same (within a factor of two) the X-ray emission from the cool core cluster is much more peaked.

cooling gas should slowly move towards the center due to the weight of the overlying ICM layers. The temperature drop in the very central regions was expected to be a factor of 100; however, over the last years, X-ray spectroscopy showed that the temperature drop is limited to about a factor of three (Peterson et al. 2001). Just when the gas should be cooling most rapidly, it appears not to be cooling at all; therefore, a heating source must exist able to balance the radiative losses. Several mechanisms have been proposed and investigated (e.g. thermal conduction, AGN heating, ICM turbulence), but none are fully satisfying and astronomers are still debating the issue.

The question of how to classify astronomical objects is widely present in the literature. In optical, clusters are usually represented as a one-dimensional sequence, from regular to irregular systems (see Sect. 1.1.1); conversely, in X-rays, they are commonly classified according to the presence of merger evidence or on the basis of the core properties. As far as the core is concerned, clusters are divided into two groups, cool core (CC) and non-cool core (NCC) clusters, according to the presence or the absence of a surface-brightness peak and a temperature drop in the center. It is not yet clear if cluster distribution with respect to temperature and surface-brightness presents a bimodality; indeed, it was found a non-negligible number of objects with intermediate characteristics. The structure in surface-brightness and temperature is usually regular and shows azimuthal symmetry for CC; conversely, NCC clusters appear as irregular and have a rather flat distribution of surface-brightness (see Fig. 1.11, right panel) and temperature.

The correlation between thermodynamic and chemical properties has not yet been widely studied; as mentioned in Sect. 1.4.2, previous works found that CC are always characterized by a central metallicity enhancement, while NCC clusters usually show mild (if present) metallicity gradients (see Fig. 1.8). The investigation of such correlation could provide us a number of clues about the evolution of clusters, from the thermodynamic and the chemical point of view.

Chapter 2

The *XMM-Newton* satellite

In this chapter, I will give a brief overview of the main characteristics of the *XMM-Newton* satellite. *XMM-Newton* (X-ray Multi-mirror Mission, Fig. 2.1) is an ESA satellite, launched on December 10th, 1999 (Jansen et al. 2002). It carries three X-ray telescopes, with different instruments in their focal plane, and a 30 cm optical/UV telescope. There are three types of science instruments, which operate simultaneously and independently:

- **European Photon Imaging Camera (EPIC),** three CCD cameras (two MOS and one pn) for X-ray imaging and spectroscopy with moderate spectral resolution;
- **Reflection Grating Spectrometer (RGS),** two spectrometers for high resolution X-ray spectroscopy;
- **Optical Monitor (OM),** for optical/UV imaging.

A sketch of the *XMM-Newton* payload is shown in the bottom panel of Fig. 2.1. Its most important characteristics for the study of extended sources such as galaxy clusters are the following.

- **High sensitivity:** *XMM-Newton* carries the X-ray telescopes with the largest effective area ($A_{\rm eff}$) of a focusing telescope ever, i.e. 4650 cm^2 at 1 keV total.
- **Good angular resolution:** the point spread function (PSF) has a full width at half maximum (FWHM) of about $6''$, and a half energy width (HEW, i.e. the radius at which a half of the total energy is encircled) of about $15''$.
- **Moderate and high spectral resolution:** the resolving power is moderate (i.e. $E/\Delta E \simeq 20 - 50$) for EPIC cameras, and high (i.e. $E/\Delta E \simeq 200 - 800$) for RGS spectrometers.
- **Long continuous target visibility:** a highly elliptical orbit offers continuous target visibility up to 40 hours, but unfortunately also a higher background level if compared to lower circular orbits.

In Table 2.1 I compare the basic characteristics of *XMM-Newton* and other recent X-ray missions.

2.1 The X-ray telescopes

The three X-ray telescopes on board *XMM-Newton* are co-aligned with a relative astrometry between the three EPIC cameras calibrated to better than $1''$-$2''$ across

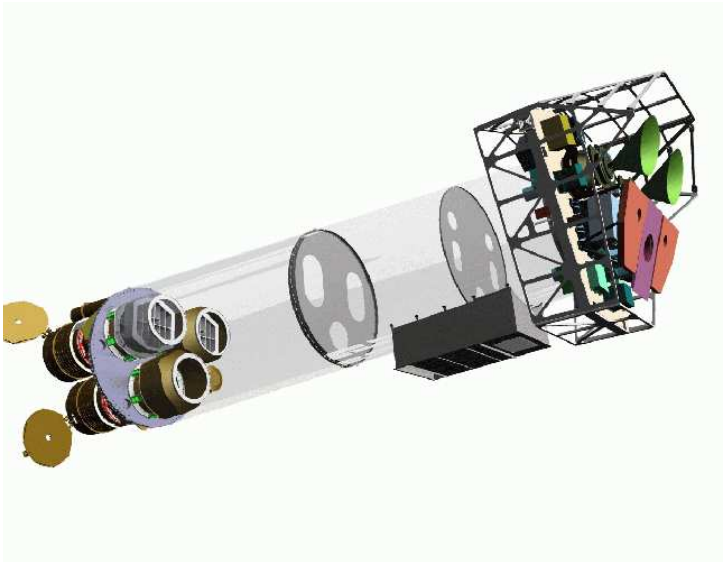

Figure 2.1: Top panel: a picture of *XMM-Newton*. Bottom panel: a sketch of the *XMM-Newton* payload. The mirror modules, two of which are equipped with Reflection Grating Arrays, are visible at the lower left. At the right end of the assembly, the focal X-ray instruments are shown: the EPIC MOS cameras with their radiators (black/green "horns"), the radiator of the EPIC pn camera (violet) and those of the (light blue) RGS detectors (in pink). The OM telescope is obscured by the lower mirror module.

Satellite	FWHM[a]	HEW[a]	Energy range[b]	$A^c_{\rm eff}$
XMM-Newton	6	15	0.15–15	4650
Chandra	0.2	0.5	0.1–10	800
ROSAT	3.5	7	0.1–2.4	400
ASCA	73	174	0.5–10	350
BeppoSAX	60	75	1.3–10	240

Notes: [a] in arcmin; [b] in keV; [c] in $\rm cm^2$ at 1 keV.

Table 2.1: Comparison between the main characteristics of *XMM-Newton* and other X-ray missions.

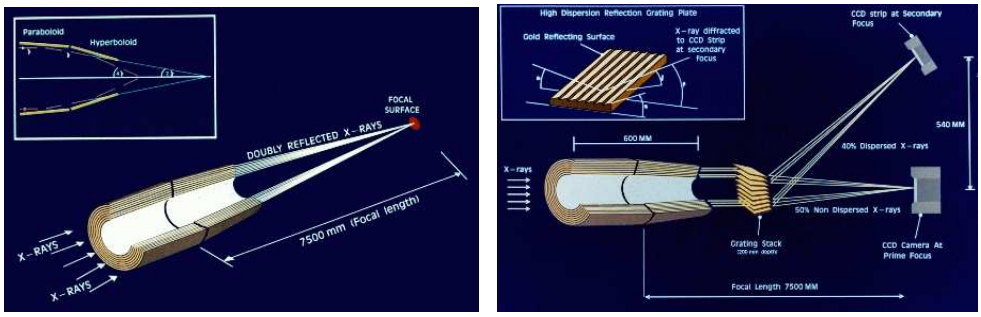

Figure 2.2: Left panel: the light path of the telescope with the pn camera in the focus. Right panel: the light path in the two *XMM-Newton* telescopes with grating assemblies. The fraction of non intercepted radiation which passes to the primary MOS focus is 44%, while 40% of the incident light is intercepted by the grating plates.

the FOV. The telescope hosting the pn camera in its focus has a light path as shown on the left of Fig. 2.2, while the others have grating assemblies, which diffract about half of the incoming radiation onto their secondary focus (Fig. 2.2, right panel).

2.1.1 The point spread function (PSF)

The first critical parameter for determining the quality of an X-ray mirror is its ability to focus photons. This is one of the most important characteristics of *XMM-Newton*: the core of its PSF is narrow and varies little over a large energy range (0.1–4.0 keV). At higher energies, the PSF is slightly more energy dependent. A wide set of in orbit calibration data sets have been considered to characterize the properties of the PSF (on-axis and off-axis) and of the encircled energy fraction (EEF), i.e. the fraction of energy of a point source collected within a given radius (Ghizzardi 2001, 2002).

Each of the three telescopes has its own PSF. The shape of the PSF is quite complex but the azimuthally averaged radial profile can be represented by analytical functions. In Fig. 2.3 I show the EEF for an on-axis source as a function of the collecting radius for several different energies. The off-axis PSF depends mainly on the off-axis angle (i.e. the distance from the center of the field of view) and slightly on the azimuthal angle. In general, the energy dependence of the PSF increases with the off-axis angle.

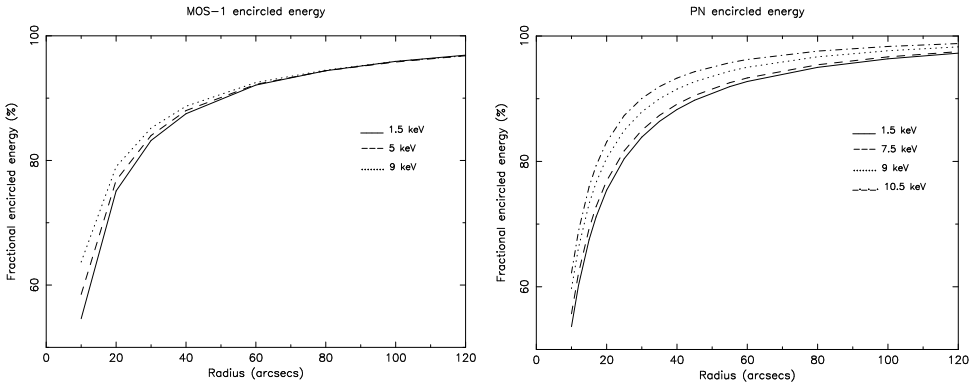

Figure 2.3: The MOS1 (left panel) and pn (right panel) EEF as a function of angular radius (on-axis) at different energies.

2.1.2 The effective area

Another important characteristic of the telescope performance is its effective area (A_{eff}), which represents the ability of the mirrors to collect photons with different energies (Fig. 2.4). The combination of the EPIC instrument (mirrors and detectors) is effective in the energy range between 0.1 and 10 keV; the effective area reaches its maximum at 1.5 keV and has a pronounced edge near 2 keV due to the gold on the telescope surface. The effective areas of the MOS cameras are lower than that of the pn, because only a fraction of the incoming photons reaches the detectors, which are partially obscured by the RGS gratings (Fig. 2.2, right panel). The effective

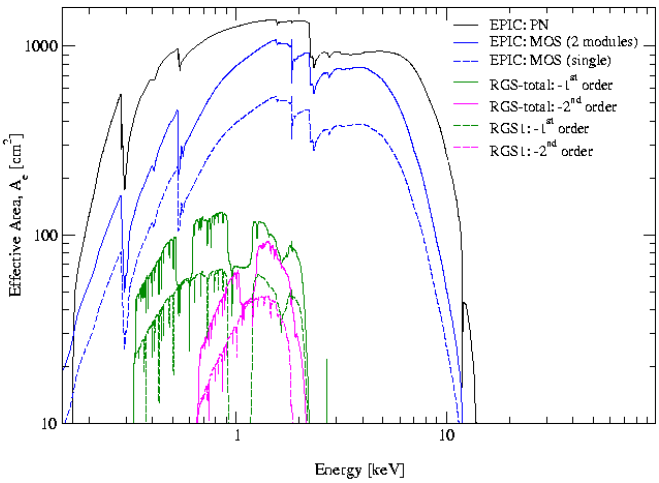

Figure 2.4: The net effective area of all *XMM-Newton* X-ray telescopes, combined with the response characteristics of the focal X-ray instruments, EPIC and RGS.

area depends on the off-axis angle within the mirror field of view (FOV). This effect is known as *vignetting*: when increasing the off-axis angle, a smaller fraction of the photons entering the telescopes actually reaches the focal plane. In Fig. 2.5 I report the vignetting of the telescopes, described by the decline of the effective area as a function of the off-axis angle.

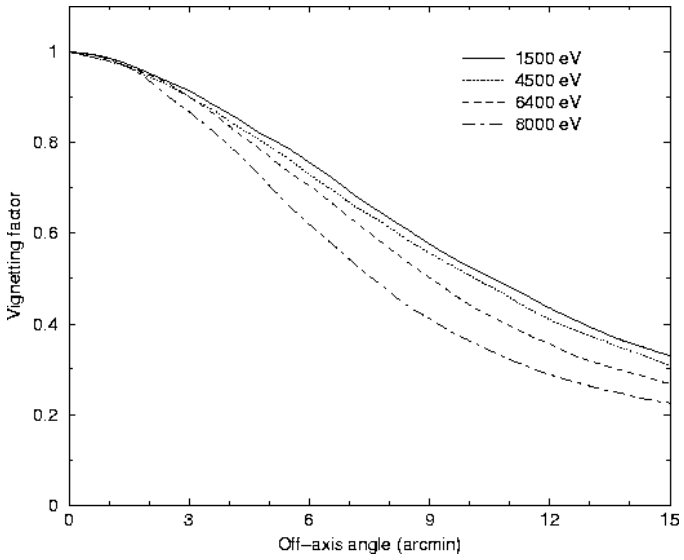

Figure 2.5: Vignetting as a function of the off-axis angle, at few selected energies, of the X-ray telescope in front of the pn camera (Ehle et al. 2003).

	EPIC MOS	EPIC pn
Bandpass	0.15–10 keV	0.15–10 keV
Sensitivity[a]	$\sim 10^{-14}$	$\sim 10^{-14}$
Field of view	$30'$	$30'$
PSF[b]	$5''/14''$	$6''/15''$
Pixel size	40 μm ($1.1''$)	150 μm ($4.1''$)
Timing resolution	1.5 ms	0.03 ms
Spectral resolution at 1 keV	~ 70 eV	~ 80 eV
Spectral resolution at 6.4 keV[c]	~ 150 eV	~ 150 eV

Notes: [a] after 10 ks in the range 0.15–15 keV, in units of erg cm^{-2} s^{-1}; [b] FWHM and HEW; [c] the energy of the Fe K$_\alpha$.

Table 2.2: Main characteristics of the European Photon Imaging Cameras.

2.2 European Photon Imaging Camera (EPIC)

Two of the X-ray telescopes on board *XMM-Newton* are equipped with EPIC MOS (*Metal Oxide Semi-conductor*) CCD arrays (Turner et al. 2001), the third carries a different CCD camera called EPIC pn (Strüder et al. 2001). The detector layout is shown in Fig. 2.6: the shaded area corresponds to the FOV ($r \approx 15'$). During revolution 961 (on March 9th, 2005), an event was registered in the focal plane of the EPIC MOS1 instrument, probably due to a micrometeoroid impact scattering debris, damaging the CCD6 which has been switched off few days later.

The two types of EPIC cameras differ in many properties, besides the geometry of the chips shown in Fig. 2.6. The readout of the pn camera is much faster than that of the MOS, since each column has its own readout node. Moreover, the pn quantum efficiency is greater than that of MOS, because the pn CCDs are back-illuminated, while the MOS chips are front illuminated. In Table 2.2 I compare the main characteristics of the detectors.

The EPIC cameras feature several modes of data acquisition, to allow observations of a large class of sources. Some of these modes have a short readout time

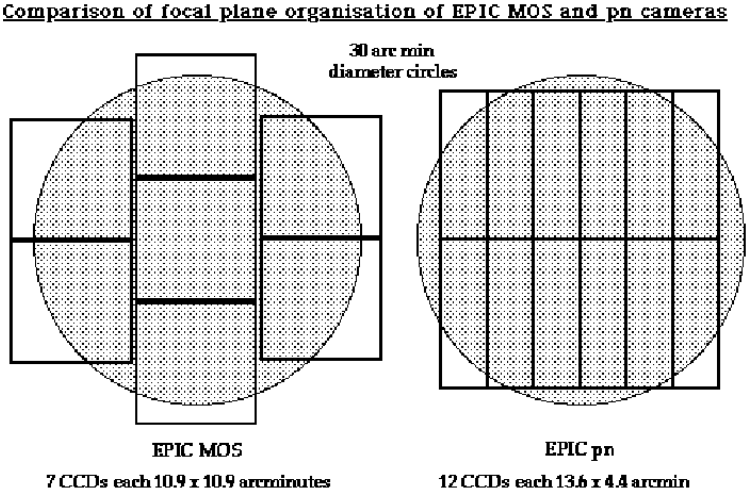

Figure 2.6: A rough sketch of the field of view of the two types of EPIC cameras; MOS (left) and pn (right). The pn chip array is slightly offset with respect to the optical axis of its telescope, so that the nominal on-axis observing position does not fall on the gaps between the central CCDs. The two MOS cameras are rotated by 90° with respect to each other.

to overcome the pile-up problem plaguing very bright sources[1] and perform timing studies. Since galaxy clusters are stationary and low surface-brightness sources, the full frame and the extended full frame (available only for the EPIC pn) modes are always used for observing clusters with EPIC. When these operative modes are set, all pixels of all CCDs are read out and the full FOV is covered. For the extended full frame mode the collection time is longer than in the full frame mode (time resolution of 200 ms instead of 73.4 ms): this is useful when correcting for pn out-of-time events (i.e. events registered during the readout process).

The angular resolution of EPIC is determined by the PSF of the mirrors, since the pixel size is smaller than the FWHM of the PSF (namely, MOS and pn pixel sizes are respectively 40 μm and 150 μm which correspond to angles of 1.1″ and 4.1″ on the sky).

The resolving power of the EPIC cameras is determined by the intrinsic energy resolution of individual pixels. In Fig. 2.7 I show the spectral resolution as a function of the energy for a point source. In the same figure, I also plot the measured in-flight FWHM of the on board calibration lines of the MOS1 (Al- and Mn K_α), after the correction for the charge transfer inefficiency (CTI), i.e. the loss of charge during the transport to the amplifiers. The resolving power for the MOS camera is reduced of about 13% since launch, due to an increase in the CTI of the CCDs. Conversely, the resolving power for the pn camera is substantially unchanged.

The EPIC effective area depends on the effective area of the telescopes, on the quantum efficiency of the CCD, and on the filter transmission power. The quantum efficiency (QE) describes the response of the CCD chips to the incident radiation; in Fig. 2.8 I show the different QE for MOS and pn. The pn is much more sensitive at low energies because is back-illuminated. Since the CCDs are not only sensitive to X-ray photons, but also to IR, visible and UV light, the EPIC cameras are equipped with a set of three filters (thick, medium, and thin) for minimizing the optical light

[1]The pile-up effect is the arrival of more than one photon in a pixel, or in adjacent pixels, before the signal is read out.

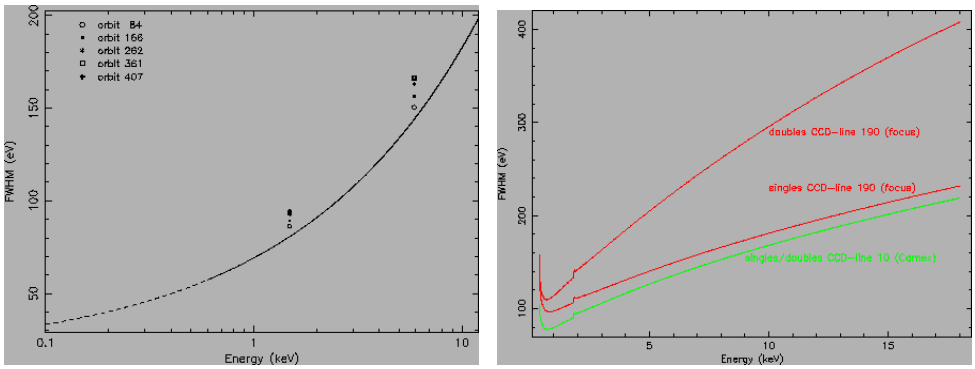

Figure 2.7: Left panel: the MOS resolving power (FWHM) as a function of the energy. The solid curve is a best-fit $E^{1/2}$ function to ground calibration data. The points are the measured in-flight FWHM of the Al-K_α (at 1.487 keV) and Mn-K_α (at 5.893 keV) lines, after CTI correction, for five different epochs. Right panel: the pn resolving power (FWHM) as a function of the energy. Curves are given for single and double events at the focus position as well as at a position 10 pixels away from the readout node (Ehle et al. 2003).

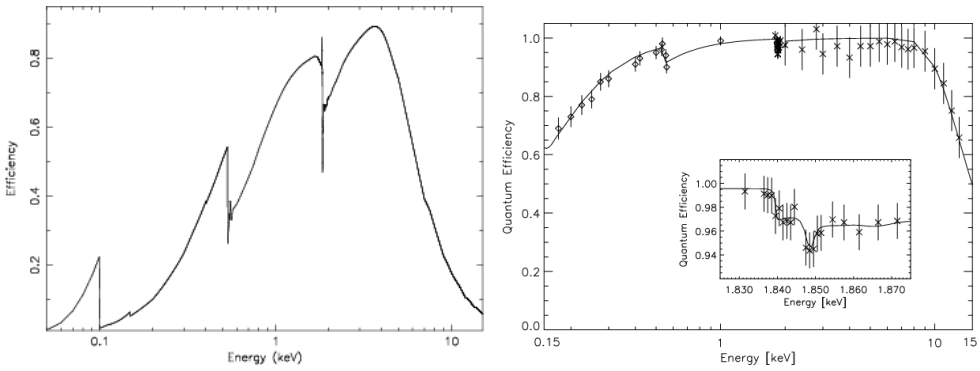

Figure 2.8: Quantum efficiency for MOS (left) and pn (right) as a function of the photon energy (Ehle et al. 2003).

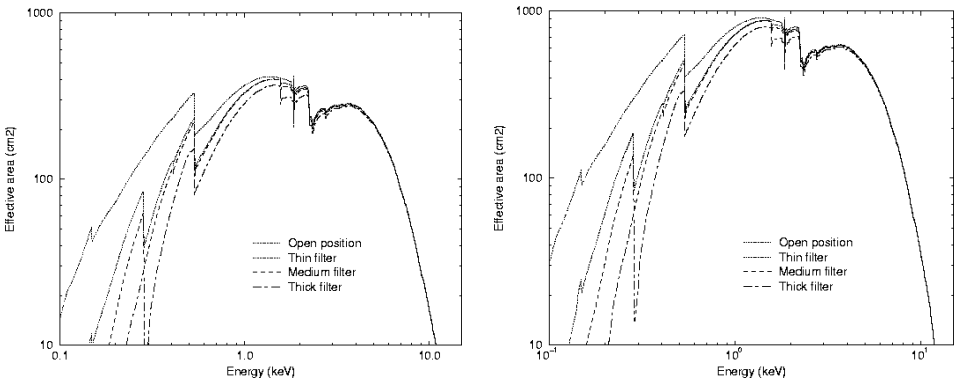

Figure 2.9: The effective area for MOS (left) and pn (right) for each of the optical blocking filters and without a filter (Ehle et al. 2003).

contamination; however, the use of a filter also reduces the softest X-ray response (see Fig. 2.9).

Chapter 3

The EPIC background and its characterization

When analyzing data from low surface-brightness sources as galaxy clusters, it is crucial to characterize background properties as well as possible. I have been carefully characterized the EPIC background by using data sets taken with the filter wheel in the closed position (hereafter referred to as "closed") and from unexposed regions (hereafter OUT-FOV), and using observations of "blank fields" at high galactic latitude. In Sect. 3.1 I describe the components which combine to constitute the EPIC background, and in Sect. 3.2 I report the detailed analysis of "closed" and "blank field" observations.

3.1 The EPIC background

The EPIC background is due to several components, which can be separated in three groups.

- **Particles:** the events due to the soft protons channeled by the telescopes towards the detectors and the internal (cosmic-ray induced) background, created directly by particles penetrating the CCDs and indirectly by the fluorescence of satellite material to which the detectors are exposed.
- **Photons:** the astrophysical background, dominated by thermal emission from within our own Galaxy at lower energies ($E \lesssim 1$ keV) and by a power law emission (primarily from unresolved cosmological sources) at higher energies.
- **Electronic noise:** bright pixels, hot columns, and readout noise ($E \lesssim 0.2$ keV).

3.1.1 The internal background

The internal cosmic-ray induced background is due to high energy ($E \gtrsim 1$ MeV) particles interacting with the detector and producing events similar to those produced by genuine X-ray photons. Cosmic rays also interact with the structure surrounding the detectors; they excite atoms, which then emit X-ray radiation when returning to their ground state. The monochromatic energy of the emitted radiation depends on the particular transition and element.

The internal background can be studied by analyzing "closed" observations, whose events are solely induced by cosmic rays (see Sect. 3.2.1). Its spectrum is characterized by a continuum plus several fluorescence emission lines: in Fig. 3.1 I report the spectra obtained by stacking data from ≈ 50 (≈ 30) closed observations

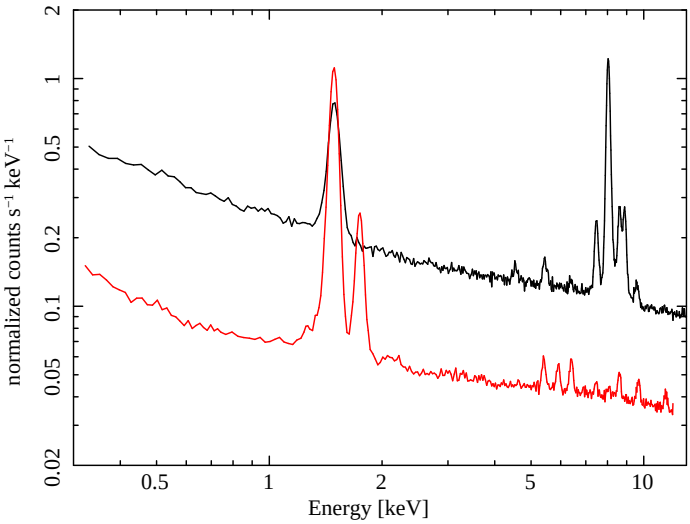

Figure 3.1: MOS1 (red) and pn (black) spectra obtained by stacking data from closed observations.

with the EPIC MOS (pn) camera. The MOS spectrum features two most intense lines around 1.5 keV and 1.8 keV, due to Al-K$_\alpha$ and Si-K$_\alpha$ respectively; the pn spectrum features the Al-K$_\alpha$ line and the Ni-Cu-Zn multiplet around 8 keV. The difference between the spectra from MOS and pn depends on the detector structure (EPIC pn is monolithic and complanar, thus Si-K$_\alpha$ radiation is self-absorbed) and position (the pn detector is very close to an electronic board responsible for the Ni-Cu-Zn emission, see also Fig. 3.2).

Above 2 keV the spectra are well fitted by a single power law (index 0.2-0.3), instead below 2 keV they become softer. As showed in Sect. 3.2.1, the internal background is roughly flat over the MOS detectors; conversely, the pn camera suffers from the proximity of the electronic board mentioned above, which blocks a fraction of the particles, and the spatial distribution of the events mirrors rather precisely the shape of the board (see Fig. 3.2). The intensity of the continuum shows temporal variations on the order of 10-15% of the mean for MOS and pn (Katayama et al. 2002; Leccardi & Molendi 2008a), while the intensity of emission lines is rather stable. More details will be given in Sect. 3.2.

3.1.2 The soft protons

The soft protons (hereafter SP) are particles with energy on the order of few hundred keV, channeled by the telescopes towards the detectors. They probably have solar origin and are accelerated by magnetospheric reconnection events. This background component is characterized by a strong variability, appearing in the light curve of the observations as sudden flares, when the count rate increases by a factor of 100 or more. In Fig. 3.3 I report an example of light curve for an extremely polluted observation. The spectra of SP flares are variable and no clear correlation has been found between intensity and spectral shape. The current understanding is that soft protons are organized in clouds in the Earth's magnetosphere. The number of such clouds encountered by *XMM-Newton* is variable and depends upon many factors (e.g. the altitude of the satellite, its position with respect to the magnetosphere and the amount of solar activity). Unfortunately, the presence of this troublesome background component was not properly anticipated before the launch. The light curves for MOS1 and MOS2 are very similar, while for pn some flares appear to

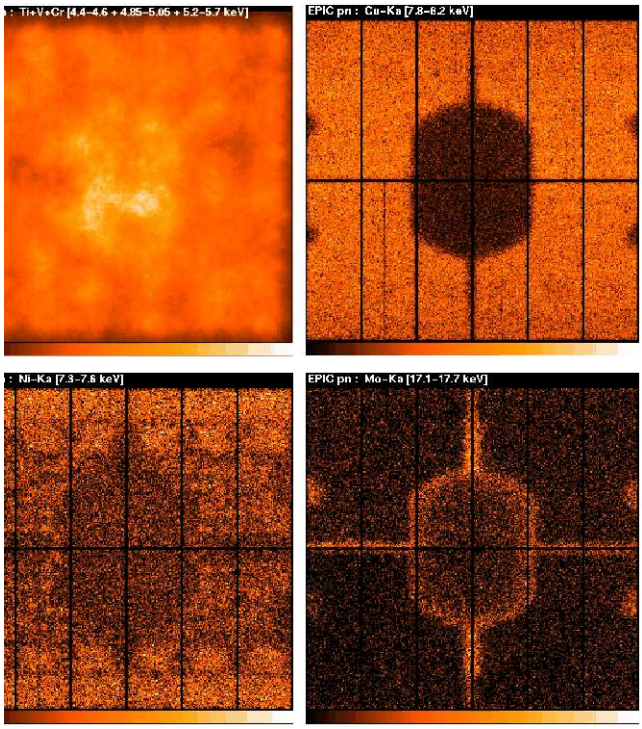

Figure 3.2: Background images for the pn camera with spatially inhomogeneous fluorescent lines: smoothed image in the Ti+V+Cr-K_α lines (top left), full resolution image in Copper (7.8–8.2 keV) (top right), Nickel (7.3–7.6 keV) (bottom left) and Molybdenum (17.1–17.7 keV) (bottom right) (Ehle et al. 2003).

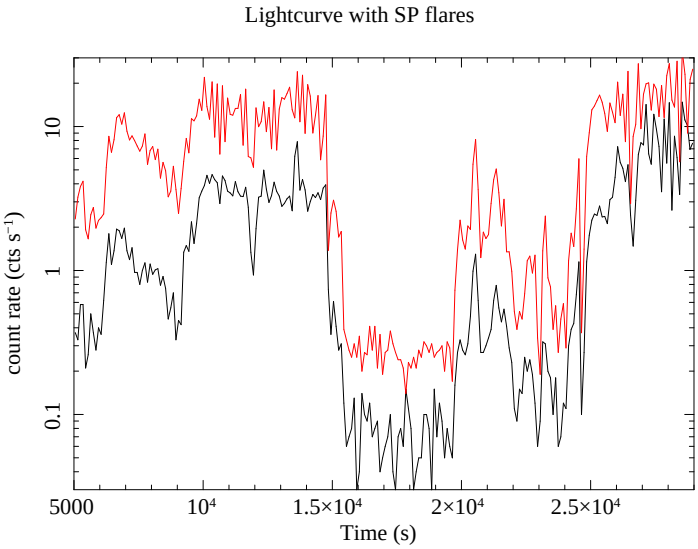

Figure 3.3: MOS1 (black) and pn (red) light curves for an extremely polluted observation. Note the logarithmic scale on the y-axis. Only 5 ks remain after the exclusion of soft proton flares.

be much more intense and to have more extended tails. Data collected during the flares are unusable and must be rejected with *good time interval* (GTI) filtering (see Chapter 4). After the filtering process a fraction of soft protons (called quiescent soft protons, QSP) can survive. To quantify the amount of this component, I made use of the "IN over OUT" diagnostic described in Sect. 4.1.2.

3.1.3 The astrophysical background

The astrophysical background beyond ≈ 3 keV has extragalactic origin: the so-called cosmic X-ray background (CXB) is isotropic on large angular scales and is due to the superposition of a large number of unresolved AGN. Its spectrum is well fitted by a power law with a photon index of 1.4 (De Luca & Molendi 2004). The total CXB flux is the sum of the contribution of the unresolved sources within a given area of the sky; for this reason, it may vary on the sky because of Poissonian variations of the number of sources, intrinsic variability of the source fluxes, and the large scale structure of the universe (see e.g. Fabian & Barcons 1992).

At lower energies, the photon background is a mixture of Galactic, heliospheric, and geocoronal diffuse emission, plus the extragalactic emission from point-like sources and warm-hot intergalactic medium (WHIM). This component is strongly spatially variable and also changes spectrally (Snowden et al. 1997; Lumb et al. 2002); it is therefore very difficult to characterize and model.

3.2 The background characterization

In Sect. 3.2.1 I present the results of my analysis of ≈ 50 MOS observations with the filter wheel in the closed position. I also performed a preliminary analysis of ≈ 30 closed observations with the pn camera; however, when dealing with cluster observations I considered only EPIC MOS data because of the following reasons. The region outside the pn field of view (FOV) is much smaller than the MOS one, therefore the uncertainties on the intensity of the NXB are large. The presence of a non negligible fraction of out-of-time events causes events from different parts of the instrument to be mixed together; this is troublesome, because the spectral components inside and outside the FOV have different shapes and a standard statistical correction does not work properly. The pn NXB is much less stable than the MOS one: especially below 2 keV, it shows variations not only in intensity but also in shape. Finally the pn instrument has further drawbacks due to the electronic board near the detector: the NXB spatial distribution is not flat and the emission due to Ni-Cu-Zn lines is more intense in the outer regions (see also Fig. 3.2).

In Sect. 3.2.2 I show the results obtained from the analysis of ≈ 30 "blank field" observations with the MOS camera.

3.2.1 The analysis of "closed" observations

I analyzed ≈ 50 MOS closed observations, exposure times of individual observations span the 5–100 ks range, for a total exposure time of ≈ 650 ks. For each observation, I selected six concentric rings ($0'$–$2.75'$, $2.75'$–$4.5'$, $4.5'$–$6'$, $6'$–$8'$, $8'$–$10'$, and $10'$–$12'$) centered on the detector center. For each instrument (i.e. MOS1 and MOS2) and each ring, I produced the total spectrum by summing, channel by channel, spectral counts accumulated during all observations. I associated the appropriate RMF with each total spectrum and performed a minimal grouping to avoid channels with no counts. In Fig. 3.4 I report the total spectra accumulated in the $10'$–$12'$ ring, for MOS1 and MOS2, in the 0.2–11.3 keV band.

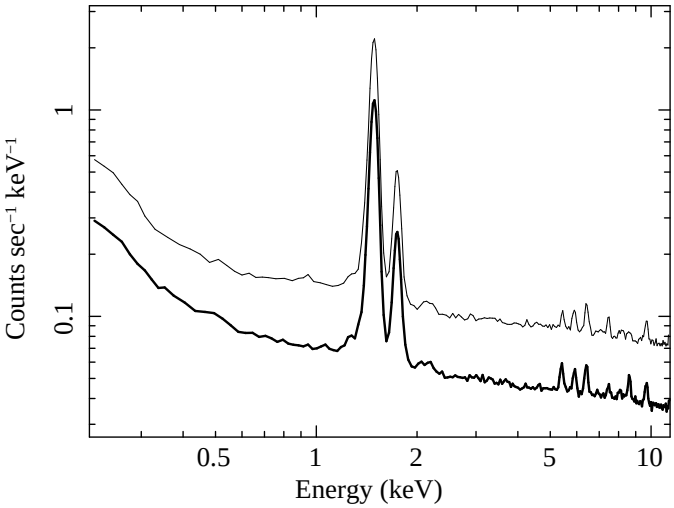

Figure 3.4: MOS1 (thick) and MOS2 (thin) spectra from closed observations in the whole energy band, i.e. 0.2–11.3 keV. MOS2 spectrum is scaled by a factor of 2 for clarity. Spectra were accumulated in the $10'$–$12'$ ring. The total exposure time is $\approx$ 650 ks.

	Γ_1	$E_{\rm B}$ [keV]	Γ_2
MOS1	0.22	7.0	0.05
MOS2	0.32	3.0	0.22

Table 3.1: Best fit parameters for the broken power law fit. Γ_1 and Γ_2 are the slopes below and above the break energy, $E_{\rm B}$.

As mentioned above, closed observation events are solely due to the internal background, which is characterized by a cosmic-ray induced continuum plus several fluorescence emission lines. Beyond 2 keV, the continuum was well fitted by a single power law (index 0.24 and 0.23 for MOS1 and MOS2 respectively); instead, for the 0.7–10.0 keV range, a broken power-law (see Table 3.1) was more appropriate. Emission lines were modeled by Gaussians. Note that the models for particle background components were not multiplied by the effective area.

In Table 3.2 I list the emission lines in my background model with their rest frame energies. Normalization values are always reported in XSPEC units. Lines are determined by 3 parameters: peak energy, intrinsic width, and normalization. The energy of Al-K$_\alpha$, E$_{\rm Al}$, was left free to allow for a small shift in the energy scale; the energies of Al-K$_\beta$, Si-K$_\alpha$, Si-K$_\beta$, and Au-M lines were linked to E$_{\rm Al}$ in such a way that a common shift, ΔE, can be applied to all lines. Similarly, the energy of Cr-K$_\alpha$, E$_{\rm Cr}$, was free and the energies of all other lines above 5 keV were linked to E$_{\rm Cr}$. The intrinsic width was always fixed to zero, except for Al and Si lines; for these lines it was fixed to 2.2 eV to allow for minor mismatches in energy calibrations for different observations. Normalizations for all K$_\alpha$ lines, Al-K$_\beta$, and Si-K$_\beta$ were free, normalizations of other K$_\beta$ lines were forced to be one-seventh of the correspondent K$_\alpha$ line (Keith & Loomis 1978). The correlation between broken power-law and Gaussian parameters was very weak.

As noticed by Kuntz (2006), there are observations in which the count rate of some CCDs is very different, especially at low energies, indicating that the spectral shape of the NXB is not constant over the detector. In particular, the count rate of MOS1 CCD-4 and CCD-5 and MOS2 CCD-2 and CCD-5 may be very high. Since my procedure requires background parameters to be rescaled from the outer to the

Line	E [keV]	Line	E [keV]
Al-K_α	1.487	Mn-K_β	6.490
Al-K_β	1.557	Fe-K_β	7.058
Si-K_α	1.740	Ni-K_α	7.472
Si-K_β	1.836	Cu-K_α	8.041
Au-M_α	2.110	Ni-K_β	8.265
Au-M_β	2.200	Zn-K_α	8.631
Cr-K_α	5.412	Cu-K_β	8.905
Mn-K_α	5.895	Zn-K_β	9.572
Cr-K_β	5.947	Au-L_α	9.685
Fe-K_α	6.400		

Table 3.2: Instrumental emission lines in the 0.7–10.0 keV energy band.

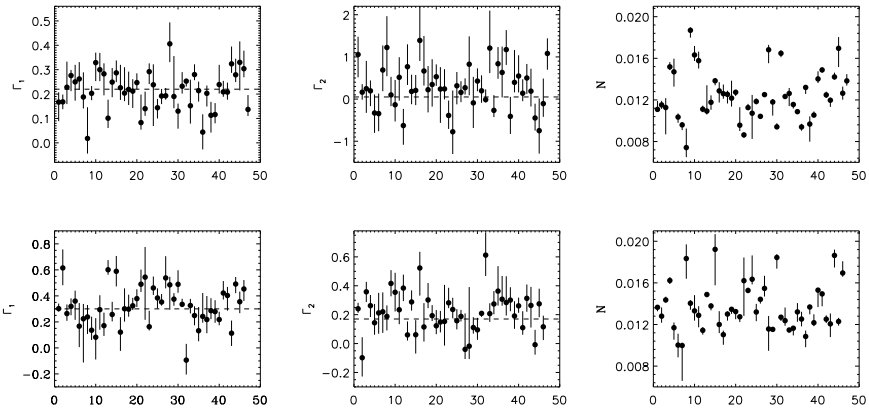

Figure 3.5: Γ_1, Γ_2, and N values for MOS1 (top) and MOS2 (bottom) for all closed observations analyzed. The dotted lines are the best fit values reported in Table 3.1. For Γ_1 and Γ_2 the scatter is comparable with the uncertainties, while for N there is an intrinsic scatter of $\approx 20\%$. Values for N are reported in XSPEC units.

inner rings (see Sect. 4.2), I always excluded the above mentioned "bright" CCDs from data analysis when using the 0.7–10.0 keV band (see Sect. 4.1.1). When using the band above 2 keV all CCDs are used, because the effect is negligible for almost all observations.

After the exclusion of the bright CCDs, I fit spectra accumulated in the $10'$–$12'$ ring for different closed observations, to check for temporal variations of the NXB. In Fig. 3.5 I report the values of broken power-law free parameters (namely the slopes, Γ_1 and Γ_2, and the normalization, N) for MOS1 and MOS2 in the 0.7–10.0 keV band. The scatter of Γ_1 and Γ_2 values is the same order of magnitude as the statistical uncertainties, while the scatter of the N values ($\approx 20\%$) is not purely statistic; i.e., NXB normalization varies for different observations.

I also checked for spatial variations of the internal background. As explained above, I accumulated the total spectrum for each of the 6 rings and for each instrument. I defined the surface-brightness, *SB*, as the ratio between N and the area of the ring. In Fig. 3.6 I report MOS1 and MOS2 best fit values of *SB* as a function of the distance from the center, by fixing Γ_1 and Γ_2. The spatial variations are greater than statistical errors but less than 5%. To a first approximation, the NXB is flat over the detector. When fitting spectra above 2 keV, I find similar results, both in terms of temporal and spatial variations.

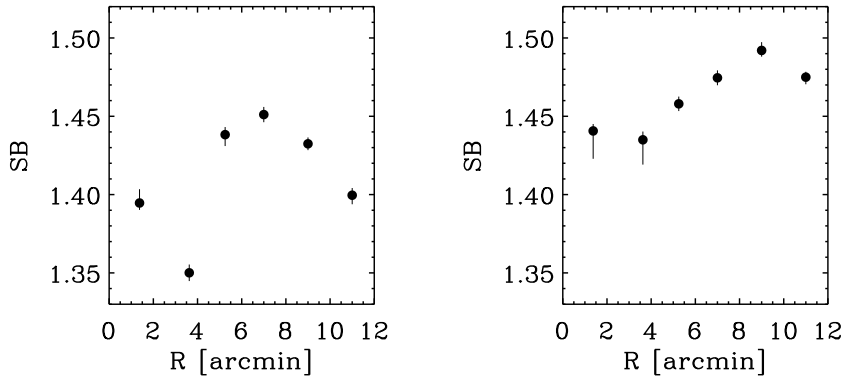

Figure 3.6: Surface-brightness best fit values for MOS1 (left) and MOS2 (right) as a function of the distance from the detector center.

Emission lines showed rather weak temporal variations and most of them (namely, all except for Al, Si, and Au) have a uniform distribution over the detector. Al lines are more intense in the external CCDs, while Si lines are more intense in the central CCD. Conversely, Au lines are very localized in the outer regions of the field of view, thus I model them only when analyzing rings beyond 3.5′.

3.2.2 The analysis of "blank field" observations

A large number (≈ 30) of "blank field" observations were analyzed to characterize the spectrum of other background components. Exposure times of individual observations cover between 30 and 90 ks for a total exposure time of ≈ 600 ks. Almost all observations have different pointing in order to maximize the observed sky region and minimize the cosmic variance of the X-ray background.

Data were prepared and cleaned as described in Sects. 4.1.1 and 4.1.2. For each instrument (i.e. MOS1 and MOS2) and each filter (i.e. THIN1 and MEDIUM), I produced total spectra by summing, channel by channel, spectral counts accumulated during all observations after the selection of the same rings used for closed observations (see Sect. 3.2.1). I associated the appropriate RMF and ARF with each spectrum and performed a minimal grouping to avoid channels with no counts. I also estimated the QSP contamination by calculating the average $R_{\rm SB}$ (see Sect. 4.1.2), which is 1.09 ± 0.01 for both filters and both detectors.

Inside the field of view, the spectral components beyond 0.7 keV are the following (see Fig. 3.7):

- the galactic component likely related to the emission from the Halo (HALO),
- the cosmic X-ray background (CXB),
- the quiescent soft protons (QSP),
- the cosmic ray induced continuum (NXB),
- the fluorescence emission lines.

The photon components only (i.e. HALO and CXB) were multiplied by the effective area and absorbed by our Galaxy. The equivalent hydrogen column density along the line of sight, $N_{\rm H}$, was fixed to the 21 cm measurement (Dickey & Lockman 1990), averaged over all fields. I selected blank field observations pointed at high

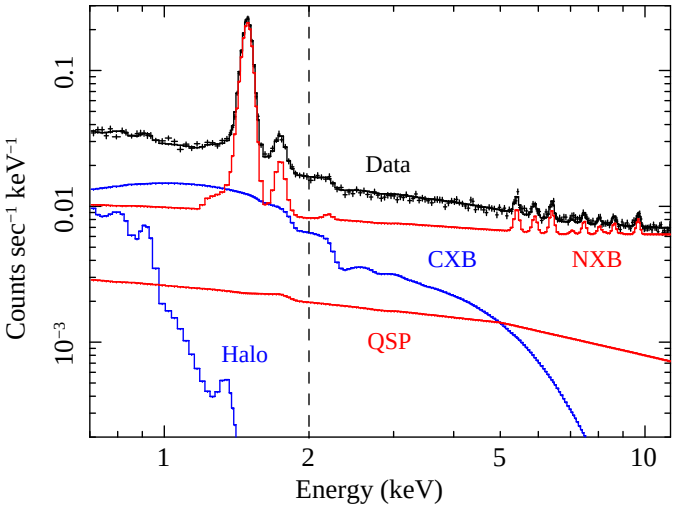

Figure 3.7: MOS1 spectrum from blank field observations in the 10′–12′ ring. Above 2 keV the spectrum is simpler. Data are shown in black, particle background in red and astrophysical background in blue.

galactic latitude, so that $N_{\rm H}$ is $< 10^{21}$ cm^{-2} and the absorption effect is negligible above 1 keV.

In the 0.7–10.0 keV band, the total model is composed of a thermal component (HALO), a power law (CXB), two broken power laws (QSP and NXB), and several Gaussians (fluorescence emission lines). The thermal model (APEC in XSPEC) parameters are: kT = 0.197 keV, Z = 1.0 $Z_{\odot}$, and z = 0.0 (Kuntz & Snowden 2000). The slope of the CXB power law was fixed to 1.4 (De Luca & Molendi 2004), and the normalization was calculated at 3 keV to minimize the correlation with the slope. The QSP broken power law has a break energy at 5.0 keV, and the slopes were fixed to 0.4 and 0.8 respectively below and above 5 keV. The model parameters for the internal background are the same as reported in Sect. 3.2.1. In the 2.0–10.0 keV band the model is simpler (namely, three power laws and several Gaussians) and more stable. The HALO component is negligible above 2 keV, the CXB model is the same as in the 0.7–10.0 keV band, the slope of the QSP power law was fixed to 1.0, and the model parameters for the internal background are those reported in Sect. 3.2.1.

Most components have rather similar spectral shapes (see Fig. 3.7), therefore a high degree of parameter degeneracy is present. In such cases, it is useful to constrain as many parameters as possible. Events outside the field of view are exclusively due to the internal background, therefore the spectrum accumulated in this region provides a good estimate of the NXB normalization, $N_{\rm NXB}$. By analyzing closed (CL) observations I find that the ratio between $N_{\rm NXB}$ calculated in two regions of the detector is independent of the particular observation:

$$\frac{N_{\rm NXB}(R_1;O_1)}{N_{\rm NXB}(R_2;O_1)} = \frac{N_{\rm NXB}(R_1;O_2)}{N_{\rm NXB}(R_2;O_2)}\,, \tag{3.1}$$

where $R_{1,2}$ are any two regions of the detector, and $O_{1,2}$ are any two observations. By using the region outside the field of view (OUT), for each ring (R) of blank field (BF) observations, I estimated and fixed $N_{\rm NXB}$ from Eq. 3.1:

$$N_{\rm NXB}(R;BF) = N_{\rm NXB}(R;CL) \times \frac{N_{\rm NXB}(OUT;BF)}{N_{\rm NXB}(OUT;CL)}\,. \tag{3.2}$$

Instr.	Filter	$N_{\rm HALO}$ $[10^{-4}]$	$N_{\rm QSP}$ $[10^{-3}]$	$N^{\rm a}_{\rm CXB}$ $[10^{-2}]$
MOS1	THIN1	1.7±0.1	2.4±0.1	5.1±0.1
MOS2	THIN1	1.6±0.1	2.5±0.1	5.0±0.1
MOS1	MEDIUM	1.4±0.1	2.6±0.1	6.0±0.1
MOS2	MEDIUM	1.6±0.1	2.4±0.1	5.8±0.1

Note: [a] calculated at 3 keV.

Table 3.3: Best fit results for the analysis of blank field observations in the 10′–12′ ring.

Ring	HALO		CXB	
	MOS1	MOS2	MOS1	MOS2
0′–2.75′	0.62	0.68	0.80	0.91
2.75′–4.5′	0.74	0.70	0.70	0.78
4.5′–6′	0.63	0.65	0.89	0.95
6′–8′	0.74	0.71	0.89	0.92

Table 3.4: Correction factors, $K(r)$, for each ring and instrument.

In Table 3.3 I report the best fit values for the normalization of the HALO, $N_{\rm HALO}$, of the QSP, $N_{\rm QSP}$, and of the CXB, $N_{\rm CXB}$, in the 10′–12′ ring, for MOS1 and MOS2 instruments and for THIN1 and MEDIUM filters. I fit spectra in the 0.7–10.0 keV energy band. I stress the remarkably good agreement between MOS1 and MOS2 for all parameters. Moreover, I point out that, when comparing observations with different filters, values for $N_{\rm HALO}$ and $N_{\rm QSP}$ agree, while values for $N_{\rm CXB}$ are significantly different ($\approx 20\%$) because of the cosmic variance ($\approx 15\%$ expected for the considered solid angles).

By construction (see Eq. 4.1) there is a relation between $R_{\rm SB}$ and $N_{\rm QSP}$, so that the higher $R_{\rm SB}$, the higher $N_{\rm QSP}$. For observations that are not contaminated by QSP, $R_{\rm SB} \approx 1.0$ and $N_{\rm QSP} \approx 0.0$ are expected. Since $R_{\rm SB}$ values span a relatively narrow range (roughly between 1.0 and 1.5), I approximated the relation between $R_{\rm SB}$ and $N_{\rm QSP}$ with a linear function: $N_{\rm QSP} = A \times (R_{\rm SB} - 1)$. The scaling factor, $A \approx 0.03$, was determined from the analysis of blank fields observations, for which I measured $R_{\rm SB} = 1.09 \pm 0.01$ and $N_{\rm QSP} = (2.5 \pm 0.1) \times 10^{-3}$. Thus, for each observation I model the bulk of the QSP component by deriving $N_{\rm QSP}$ from $R_{\rm SB}$ (see Sects. 4.2.1 and 4.2.2). In Sects. 7.1.1 and 7.1.2 I will discuss possible systematics related to QSP, and show that the linear approximation used above is satisfactory.

When analyzing cluster observations, I estimated the normalizations of the background components in the 10′–12′ ring, and rescaled them in the inner rings (see Sect. 4.2.1). When considering the 0.7–10.0 keV energy band, a simple rescaling by the area ratio was too rough and caused systematic errors, especially in the outer regions where cluster emission and background fluctuations are comparable. To overcome this problem, I have proceeded in the following manner. I fit blank field spectra, by fixing $N_{\rm NXB}$ and $N_{\rm QSP}$, and determined $N_{\rm CXB}$ and $N_{\rm HALO}$ best fit values. For each ring and instrument, I defined a correction factor, $K(r)$:

$$K(r) = \frac{N_{\rm obs}}{N_{\rm exp}}, \tag{3.3}$$

where $N_{\rm obs}$ is the best fit value that I just obtained, and $N_{\rm exp}$ was derived by rescaling the value measured in the 10′–12′ ring by the area ratio. In Table 3.4 I report the values for $K(r)$ for all cases. $K(r)$ is a second order correction, because the contribution of CXB and HALO components to the total flux is relatively small:

when considering the 0.7–2.0 keV band only (i.e. the energy range in which these components are more intense), the HALO-to-total and the CXB-to-total flux ratios are $\approx 5\%$ and $\approx 20\%$, respectively. Thus, the effective correction is of a few percent only, for both cases.

Different observations have different centers in detector coordinates, and the intensity of the various components depends on the particular observation. These facts could cause some discrepancies; however, since I analyzed a large number of blank-field and cluster observations, I expect only a few percent systematic effect on the mean profile. When considering the band above 2 keV, the statistical quality of the data is poorer; therefore, rescaling by the area ratio (i.e. no correction factor) can be considered a good approximation for both CXB and NXB. The QSP value was rescaled by the soft proton vignetting profile (Kuntz 2006) and did not require any correction factor.

Chapter 4

Data analysis

In this chapter I will describe the data analysis procedure in detail. The preparation of spectra comprises the following major steps:

- preliminary data processing,
- good time interval (GTI) filtering to exclude periods of high soft proton flux,
- filtering according to pattern and flag criteria,
- excision of brightest point-like sources,
- calculation of the "IN over OUT" ratio,
- extraction of spectra in concentric rings.

The spectral analysis is structured as follows:

- estimate of background parameters from a peripheral ring of the field of view;
- spectral fitting using the Cash statistic and modeling the background, rather than subtracting it, as commonly done;
- production of surface-brightness, temperature, and metallicity profiles.

All these points are described in detail in the following sections.

I recall that I used only EPIC MOS data because a robust characterization of EPIC pn background was not possible (see Chapter 3 for further details).

4.1 Spectra preparation

4.1.1 Preliminary data preparation

Observation data files (ODF) were retrieved from the *XMM-Newton* archive and processed in a standard way with the Science Analysis System (SAS) v6.1.

The soft proton cleaning was performed using a double filtering process. I extracted a light curve in 100 second bins in the 10–12 keV energy band by excluding the central CCD, applied a threshold of 0.20 cts s^{-1}, produced a GTI file and generated the filtered event file accordingly. This first step allows most flares to be eliminated, however softer flares may exist such that their contribution above 10 keV is negligible. I then extracted a light curve in the 2–5 keV band and fit the histogram obtained from this curve with a Gaussian distribution. Since most flares had been rejected in the previous step, the fit was usually very good. I calculated

the mean count rate, μ, and the standard deviation, σ, applied a threshold of $\mu+3\sigma$ to the distribution, and generated the filtered event file.

After soft proton cleaning, I filtered the event file according to `PATTERN` and `FLAG` criteria (namely `PATTERN`$\leq$12 and `FLAG`==0). I excluded observations that are highly affected by soft proton flares, when their good (i.e. after flare cleaning) exposure time is not long enough (less than 16 ks when summing MOS1 and MOS2) to measure reliable temperature and metallicity profiles out to external regions.

When fitting spectra in the 0.7–10.0 keV band (see Sect. 4.2), I also excluded the "bright" CCDs, i.e. CCD-4 and CCD-5 for MOS1 and CCD-2 and CCD-5 for MOS2 (see Sect. 3.2.1 for the discussion).

Brightest point-like sources were detected, using a procedure based on the SAS task `edetect_chain`, and excluded from the event file. I estimated a flux limit for excluded sources on the order of 10^{-13} erg cm^{-2} s^{-1}; after the source excision, the cosmic variance of the residual X-ray background on the entire field of view (FOV) is expected to be $\approx 20\%$.

4.1.2 Quiescent soft proton contamination

A quiescent soft proton (QSP) component can survive the double filtering process (see Sect. 4.1.1). To quantify the amount of this component, I made use of the "IN over OUT" diagnostic[1] (De Luca & Molendi 2004). I measured the surface-brightness, $\mathrm{SB_{IN}}$, in an outer region of the FOV, where the cluster emission is negligible, and compared it to the surface-brightness, $\mathrm{SB_{OUT}}$, calculated outside the FOV in the same energy range (i.e. 6–12 keV). Since soft protons are channeled by the telescope mirrors inside the FOV and the cosmic ray induced background covers the whole detector, the ratio

$$R_{\mathrm{SB}} = \frac{SB_{\mathrm{IN}}}{SB_{\mathrm{OUT}}} \tag{4.1}$$

is a good indicator of the intensity of residual soft protons and was used for background modeling (see Sects. 4.2.2 and 3.2.2). The values for R_{SB} roughly span the range from 1.0 (negligible contamination) to 1.5 (high contamination). The typical uncertainty in measuring R_{SB} is a few percent.

4.1.3 Spectra accumulation

The cluster emission was divided in 10 concentric rings (namely 0′–0.5′, 0.5′–1′, 1′–1.5′, 1.5′–2′, 2′–2.75′, 2.75′–3.5′, 3.5′–4.5′, 4.5′–6′, 6′–8′, and 10′–12′). The center of the rings was determined by surface-brightness isocontours at large radii and is not necessarily coincident with the X-ray emission peak. I prefer that azimuthal symmetry be preserved at large radii, where I am interested in characterizing profiles, at the expense of central regions.

For each instrument (i.e MOS1 and MOS2) and each ring, I accumulated a spectrum and generated an effective area (ARF). For each observation I generated one redistribution function (RMF) for MOS1 and one for MOS2. I performed a minimal grouping to avoid channels with no counts, as required by the Cash statistic.

4.2 Spectral analysis

Spectral fitting was performed within the XSPEC v11.3 package. The choice of the energy band for the spectral fitting was not trivial. I fit spectra in the 0.7–10.0 keV

[1] A public script is available at http://xmm.esac.esa.int/external/xmm_sw_cal/background/epic_scripts.shtml

and in the 2.0–10.0 keV energy bands, by using the Cash statistic, with an absorbed thermal plus background model. The high-energy band has the advantage of requiring a simplified background model (see Sects. 3.2.1 and 3.2.2); however, a significant fraction of source counts was excluded and the statistical quality of the measurement was substantially reduced. Due to the paucity of source counts, there is a strong degeneracy between source temperature and normalization, and the temperature is systematically underestimated; therefore, when using the 2.0–10.0 keV band, an "a posteriori" correction was required (Leccardi & Molendi 2007). In contrast, in the 0.7–10.0 keV band, the statistical quality of the data is good, but the background model is more complicated and background components are less stable and affected by strong degeneracy (see Sect. 3.2.1 and 3.2.2). I excluded the band below 0.7 keV because the shape of the internal background is very complicated and variable with time and because the source counts reach their maximum at ≈ 1 keV. The 0.7–10.0 keV energy band seems to be the best trade-off for this kind of measurements. Hereafter, all considerations are valid for both energy bands, unless otherwise stated.

In conditions of poor statistics (i.e. few counts/bin) and high background, the Cash statistic (Cash 1979) is more suitable than the χ^2 with reasonable channel grouping (Leccardi & Molendi 2007). The Cash statistic requires the number of counts in each channel to be greater than zero (Cash 1979); thus, the background cannot be subtracted. In the case at hand, the total background model is the sum of many components, each one characterized by peculiar temporal, spectral, and spatial variations (see Sect. 3.2.2). When subtracting the background, the information on single components was lost. Conversely, background modeling allows one to preserve the information and to manage all components appropriately. Moreover, I recall that the background modeling does not require strong channel grouping, error propagation, or renormalization factors.

4.2.1 Estimate of background parameters

To model the background, a careful characterization of all its components is mandatory. Ideally, one would like to estimate background parameters in the same region and at the same time as the source. Since this was not possible, I estimated background parameters in the external $10'$–$12'$ ring and rescaled them in the inner rings, by making reasonable assumptions on their spatial distribution tested by analyzing blank-field observations (see Sect. 3.2.2). The $10'$–$12'$ ring often contains a weak cluster emission that, if neglected, may cause a systematic underestimate of temperature and normalization in the inner rings (see Sect. 7.1.1). In this ring the spectral components in the 0.7–10.0 keV band are:

- the thermal emission from the cluster (GCL),
- the emission from the Galaxy Halo (HALO),
- the cosmic X-ray background (CXB),
- the quiescent soft protons (QSP),
- the cosmic ray induced continuum (NXB),
- the fluorescence emission lines.

The HALO component is negligible when considering the 2.0–10.0 keV range. The model is the same one as used when analyzing blank-field observations (see Sect. 3.2.2 for further details) plus a thermal component for the GCL.

I fixed most parameters (namely all except for the normalization of HALO, CXB, NXB, and fluorescence lines) to reduce the degeneracy due to the presence of

different components with similar spectral shapes. All cluster parameters were fixed: the temperature, kT, and the normalization, $N_{\rm S}$, were extrapolated from the final profiles through an iterative procedure. The metallicity, Z, was fixed to 0.2 solar[2] and the redshift, z, was fixed to the optical value. The QSP normalization, $N_{\rm QSP}$, was calculated from $R_{\rm SB}$ (see Sect. 3.2.2) and fixed. Minor discrepancies in shape or normalization with respect to the real QSP spectrum are possible, the model accounts for them by slightly changing the normalization of other components, i.e. $N_{\rm HALO}$, $N_{\rm CXB}$, and $N_{\rm NXB}$ (see Sects. 7.1.1 and 7.1.2 for the discussion of the systematic effects related to QSP).

Summarizing, in the 10′–12′ ring, I have determined the range of variability, [$N_{\rm min}$,$N_{\rm max}$], (i.e. the best fit value $\pm1\sigma$ uncertainty) for the normalization of the main background components, i.e. $N_{\rm HALO}$, $N_{\rm CXB}$, and $N_{\rm NXB}$. Once properly rescaled, this information allowed us to constrain background parameters in the inner rings.

4.2.2 Spectral fit in concentric rings

I fit spectra in internal rings with the same model as adopted in the 10′–12′ ring case (see Sect. 4.2.1). In Fig. 4.1 I compare spectra and best fit models for two different regions of the same cluster. In the inner ring (1′–1.5′) source counts dominate, while in the outer ring (4.5′–6′) background counts dominate.

The equivalent hydrogen column density along the line of sight, $N_{\rm H}$, was fixed to the 21 cm measurement (Dickey & Lockman 1990). Since clusters in our sample are at high galactic latitude ($|b| > 20^{\circ}$), the $N_{\rm H}$ is $< 10^{21}$ cm^{-2} and the absorption effect is negligible above 1 keV. I have always left the temperature, kT, and the normalization, $N_{\rm S}$, free to vary. The metallicity was constrained between ±5 $Z_{\odot}$ below $\approx 0.4R_{180}$ and fixed to 0.2 solar beyond, because there the source-to-background count rate ratio is too small and the measurements are unreliable. The redshift was allowed to vary between $\pm7\%$ of the optical measurement in the two innermost rings and, in the other rings, was fixed to the average value of the first two rings, by considering independently MOS1 and MOS2 spectra. The main reason for this choice is to allow for EPIC calibration uncertainties, and for possible discrepancies between X-ray and optical derived redshift values. Typical shift values are on the order of 2%.

The normalization of HALO, CXB, and NXB for the inner rings were obtained by rescaling the best-fit values in the 10′–12′ ring (see Sect. 4.2.1) by the area ratio and the correction factor, $K(r)$, obtained from blank field observations (see Table 3.4 in Sect. 3.2.2):

$$N^{\rm int} = N^{\rm ext} \times \frac{Area^{\rm int}}{Area^{\rm ext}} \times K(r)\ , \tag{4.2}$$

for NXB $K = 1$ for all rings. Then, $N^{\rm int}_{\rm HALO}$, $N^{\rm int}_{\rm CXB}$, and $N^{\rm int}_{\rm NXB}$ were free to vary within a certain range: the lower (upper) limit of this range was derived by rescaling the best-fit value minus (plus) the 1σ-error calculated in the 10′-12′ ring. The local background should have a variation length scale of some degrees (Snowden et al. 1997); conversely, $N_{\rm CXB}$ may have large (i.e. 20-100%) variations between different rings due to the cosmic variance. However, extensive simulations show that these statistical fluctuations do not introduce systematics in the temperature measurement, when averaging on a large sample (see Sect. 7.1.1). I obtained $N^{\rm int}_{\rm QSP}$ by rescaling the value adopted in the 10′-12′ ring by the area ratio and by the QSP vignetting profile (Kuntz 2006), and fixed it for all rings. Normalizations of instrumental fluorescence emission lines were free to vary within a limited range

[2]The solar abundances were taken from Anders & Grevesse (1989).

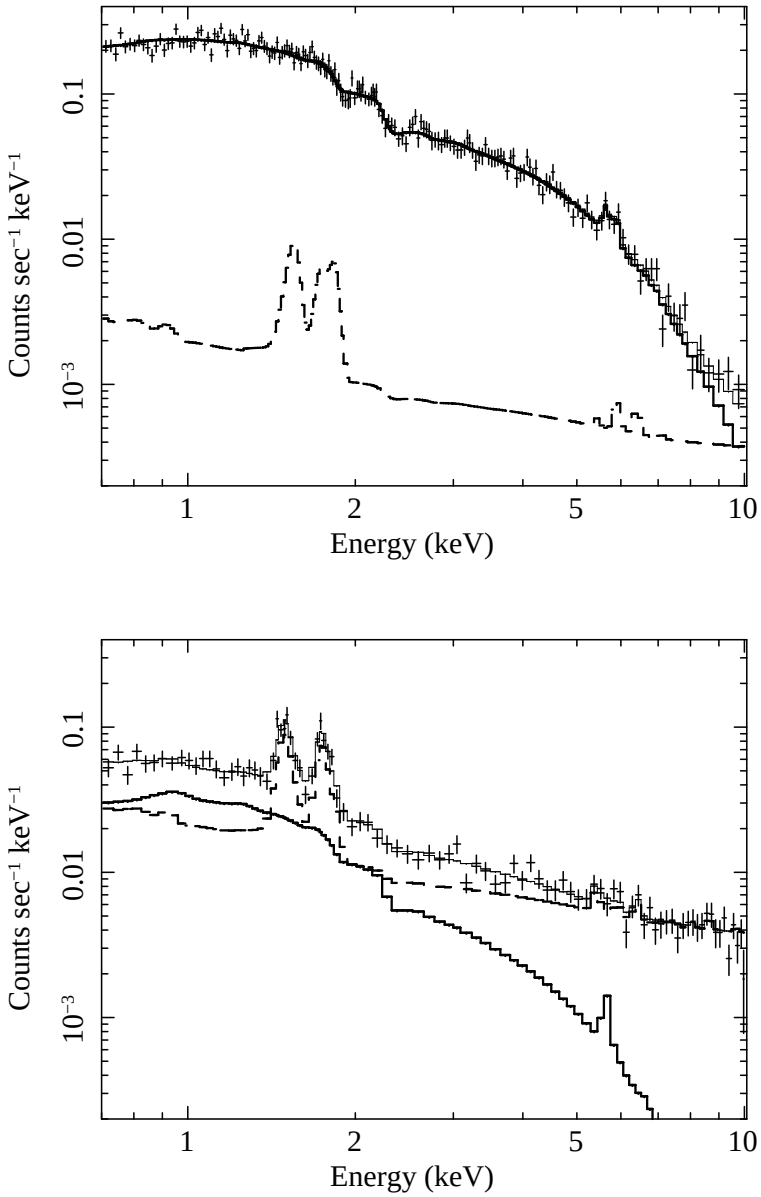

Figure 4.1: Spectra and best fit models for the $1'$–$1.5'$ (top) and the $4.5'$–$6'$ (bottom) rings of Abell 1689. The solid thick and the dashed thick lines respectively represent the thermal and the total background model. The solid thin line represents the total (i.e. thermal + background) model. In the inner ring, source counts dominate background ones, in the outer the opposite is true.

determined from the analysis of closed observations and have an almost negligible impact on final measurements.

For each ring, when using the 0.7–10.0 keV energy band, I determined kT, Z, and $N_{\rm S}$ best fit values and one-sigma uncertainties for each MOS and calculated the weighted average. Conversely, when using the 2.0–10.0 keV band, I combined temperature measurements from different instruments as described in Leccardi & Molendi (2007), to correct for the bias that affects the temperature estimator (see Chapter 5). In the 0.7–10.0 keV band, there are many more source counts, the temperature estimator is much less biased, and the weighted average returns a slightly ($\approx 3\%$ in an outer ring) biased value (see the $F = 1.0$ case in Sect. 7.1.1).

Finally, surface-brightness (i.e. normalization over area), temperature, and metallicity profiles were produced for each cluster.

Chapter 5

The bias on the ML temperature estimator

When exploring the physical properties of the intra-cluster medium (ICM) in the outer regions of galaxy clusters, both statistical and systematic issues need to be addressed. Typically, the spectra have poor statistics (i.e. few counts/bin) and a high background, especially at high energies, where the instrumental background dominates other components, also due to the sharp decrease of the effective area of the experiments (see Chapter 2). In this chapter I employ simulations to examine how best to analyze this kind of spectra, focusing on the treatment of statistical errors only. More specifically the question I wish to address is the following: *"What are the effects of pure statistical uncertainties in determining interesting parameters of highly non linear models (e.g. the temperature of the ICM), when we analyze spectra accumulated from low surface-brightness regions using current X-ray experiments?"* To deal with this, I performed a set of simulations. First, I chose the input values for model parameters and produced the expected spectrum; then, I generated a large number of perturbed spectra representing a large set of measurements; at last, I analyzed them with different techniques based on the method of maximum-likelihood (hereafter ML) and compared results. The choices of simulation parameters (e.g. spectral model, fixed parameters, etc.) are justified by the practical issue of determining the temperature in the outer regions of massive galaxy clusters. The analysis is mainly focused on *XMM-Newton*; however, most results are valid in all cases when analyzing low-count Poisson-distributed data.

From a more general perspective, this may be viewed as an attempt to quantify the significance of the bias of ML estimators commonly adopted by X-ray astronomers to determine spectral parameters. I will show that the most common ML estimators, indeed all those available within XSPEC, are characterized by a substantial bias when applied to our specific case (see Sect. 5.2). A long-term solution to the problem requires an unbiased, or perhaps a less biased, estimator to be found and implemented within standard fitting packages (see Sect. 5.3). Another, faster, solution involves correcting the bias a posteriori by making use of extensive Montecarlo simulations (see Sect. 5.3.3).

5.1 The source-only case

At first, I addressed the idealized source-only case. I represented the source with an absorbed thermal model (`WABS*MEKAL` in XSPEC). The parameter values were the following: the equivalent hydrogen column density along the line of sight, $N_{\rm H}$, is 2.5×10^{20} cm^{-2}; the metallicity, Z, and the redshift, z, were respectively 0.25

solar and 0.2; the temperature, kT, was 7 keV and the normalization, $N_{\rm S}$, was 3.5×10^{-3} in XSPEC units. Redistribution matrix (RMF) and effective area (ARF) have been produced from the observation number 0093030101 of the galaxy cluster Abell 1689 with the EPIC MOS1 instrument; the angular size of the region from which the simulated spectra were accumulated was ≈ 4 arcmin^2, corresponding to the ring between $1.0'$ and $1.5'$ centered on the cluster emission peak. The exposure times considered are 5 ks, 10 ks, 100 ks, and 1000 ks and the total counts in the 2.0–10.0 keV band are respectively about 600, 1200, 12000, and 120000. For each channel I perturbed the number of counts with a Poisson distribution centered on the expected value. I repeated this step $N_{\rm meas}$ times (with $N_{\rm meas}$ very large) to obtain $N_{\rm meas}$ spectra, simulating $N_{\rm meas}$ independent measurements of the source.

I fit simulated spectra using the χ^2 and the Cash statistics, where the latter is more suitable for analyzing spectra with a few counts per channel (Cash 1979; Nousek & Shue 1989; Mighell 1999; Arzner et al. 2006). I recall that each measurement can be represented by the number of counts, $O_{\rm i}$, observed in each channel i ($i = 1, ..., N$ where N is the number of channels). The probability, Q, of obtaining this particular measurement (i.e. this particular spectrum) is the product of Poisson distributions and can be expressed as a function of the expected counts, $E_{\rm i}$, which depend[1] on the particular set of model parameters, α (e.g. in this case $\alpha = (N_{\rm H}, {\rm k}T, Z, z, N_{\rm S})$):

$$Q(\alpha) = \prod_{i=1}^{N} \frac{{E_{\rm i}}^{O_{\rm i}} \, {\rm e}^{-E_{\rm i}}}{O_{\rm i}!} \ . \tag{5.1}$$

The associated log-ML function C (Cash 1979) is defined as

$$C(\alpha) = -2 \ \ln Q(\alpha) = -2 \sum_{i=1}^{N} \ (O_{\rm i} \ \ln E_{\rm i} - E_{\rm i} - \ln O_{\rm i}!) \ . \tag{5.2}$$

The best set of parameters is determined by maximizing Q (i.e. minimizing C) with respect to α. Conversely, the χ^2 statistic is based on the hypothesis that each spectral bin contains a sufficient number of counts for the deviations of the $O_{\rm i}$ from the $E_{\rm i}$ to have a Gaussian distribution. This hypothesis is satisfied for large $O_{\rm i}$, when Q can be approximated by a product of Gaussian distributions, and the associated log-ML function χ^2 is defined as

$$\chi^2(\alpha) = \sum_{i=1}^{N} \ \frac{(O_{\rm i} - E_{\rm i})^2}{{\sigma_{\rm i}}^2} \ , \tag{5.3}$$

where $\sigma_{\rm i}$ is usually the uncertainty in the i-th bin ($\sigma_{\rm i} = {O_{\rm i}}^{1/2}$). The larger $O_{\rm i}$, the better the approximation of the Gaussian regime. Channel grouping is a widely used strategy that allows the bias introduced by this approximation to be reduced. I grouped channels in order to have at least 25 counts per bin, which is a commonly adopted compromise. Conversely, when using the Cash statistic I performed a minimal grouping to avoid channels with no counts; i.e. the spectrum was substantially unbinned and no spectral information was lost. Each spectrum was fitted between 2.0 and 10.0 keV with the absorbed thermal model mentioned above. The $N_{\rm H}$ was fixed to the input value, z was allowed to vary between 0.186 and 0.214 ($\pm 7\%$ of the input value), while kT, Z, and $N_{\rm S}$ were free. I determined best-fit values and one-sigma uncertainties for all parameters.

In Table 5.1 I compared the weighted average of the $N_{\rm meas}$ measured temperatures to the input value, ${\rm k}T_0 = 7$ keV, for different exposure times and statistics. I chose $N_{\rm meas}$ in order to have similar uncertainties on the average ($N_{\rm meas} = 1200$ for

[1] In the following equations the dependency of $E_{\rm i}$ from α is omitted for clarity.

		χ^2		Cash	
Exp.[a]	kT_0^b	kT^c	$\Delta T/T_0^d$	kT^c	$\Delta T/T_0^d$
1000	7.00	6.89±0.01	−1.6%	7.00±0.01	+0.0%
100	7.00	6.83±0.01	−2.4%	7.03±0.01	+0.4%
10	7.00	6.76±0.03	−3.4%	6.91±0.02	−1.3%
5	7.00	6.59±0.04	−5.9%	6.81±0.03	−2.7%

Notes: [a] exposure time in ks; [b] input temperature in keV; [c] measured temperature in keV; [d] relative difference.

Table 5.1: Weighted averages of temperature best fit values compared to the input value and relative differences $\Delta T/T_0$, using different exposure times and statistics.

N_{bin}^a	kT_0^b	kT^c	$\Delta T/T_0^d$
400	7.00	6.99±0.01	−0.1%
100	7.00	6.95±0.01	−0.7%
25	7.00	6.89±0.01	−1.6%

Notes: [a] counts per bin; [b] input temperature in keV; [c] measured temperature in keV; [d] relative difference.

Table 5.2: Weighted averages of temperature best-fit values compared to the input value and relative differences $\Delta T/T_0$, using different channel groupings.

5 and 10 ks, $N_{meas} = 300$ for 100 and 1000 ks). In almost all cases, the true temperature is underestimated by a few percent and the effect becomes more evident for shorter exposure times. I recall that both χ^2 and Cash statistics are based on the ML method. Although X-ray astronomers make extensive use of ML estimators, it is well known from the literature (e.g. Cowan 1998) that: 1) ML estimators may be biased, i.e. the expectation value may be different from the true value of the quantity to estimate; 2) ML estimators are usually Gaussian and unbiased only in the asymptotic limit. In the case at hand, the asymptotic limit is approached when the total number of counts becomes large. The results reported in Table 5.1 show that: 1) both ML estimators are biased; 2) both estimators are asymptotically unbiased; 3) the Cash estimator tends to the true value more quickly than the χ^2 one.

As I have just pointed out, the χ^2 is significantly more biased than the Cash estimator (i.e. the difference between the expected and the true value is greater). This is because the approximation of the Gaussian regime fails for few counts per bin. The obvious implication is that, to improve the precision of the χ^2 estimates, one needs to increase the number of counts in each bin, N_{bin}. In Table 5.2 I compare the results obtained using the χ^2 with different channel groupings (note that this is not necessary when using the Cash statistic). The input temperature was 7 keV, the exposure time 1000 ks, and the number of measurements 300. As expected, I find that the greater the number of counts in each bin, N_{bin}, the smaller the bias. However, in practice, grouping of a large number of channels is not desirable, because it causes loss of spectral information; 25 counts per bin is a commonly adopted compromise. I have to mention the existence of an alternative way of reducing the bias that affects the χ^2 estimator for few counts per bin. Some authors (e.g. Churazov et al. 1996; Gehrels 1986; Kearns et al. 1995) have chosen different statistic weights (σ_i in Eq. 5.3) instead of the standard $O_i^{1/2}$. By analyzing again all spectra using all the alternative weights implemented in XSPEC, I obtained results somewhat similar to those already discussed for the Cash statistic.

My results are summarized as follows. The standard χ^2 statistic works well only in the Gaussian regime, which is reached when performing a strong channel

grouping (see Table 5.2). When using a realistic grouping (e.g. 25 counts per bin) the measured temperature, kT, is lower than the true temperature, kT_0 (see Table 5.1). The Cash statistic (Cash 1979) works better, because it is based on the ML function for Poisson processes; however, when the total number of spectral counts is small, kT is lower than kT_0 by a few percent (see Table 5.1). This means that the Cash ML estimator is only asymptotically unbiased (for a review about parameter estimation and ML concepts see Cowan 1998). Many efforts (e.g. Cash 1979; Wachter et al. 1979; Baker & Cousins 1984; Gehrels 1986; Nousek & Shue 1989; Kearns et al. 1995; Churazov et al. 1996; Jading & Riisager 1996; Mighell 1999; Hauschild & Jentschel 2001; Bergmann & Riisager 2002; Arzner et al. 2006) have been devoted to extending the standard theories about curve fitting to the case of low-count spectra (best fit parameters and confidence intervals estimation, goodness-of-fit test, etc.); however, no definitive solution has been found yet.

5.2 The source-plus-background case

I also considered a more realistic situation by introducing a simplified instrumental background. The model is a power law (`PEGPWRLW/b` in XSPEC) convolved with the RMF but not multiplied by the ARF. The power-law slope, $\Gamma_{\rm B}$, is fixed to 0.25; the normalization, $N_{\rm B}$, is calculated at the center of the energy band to minimize the correlation with $\Gamma_{\rm B}$.

There are two ways of analyzing spectra with background: subtract it using a spectrum from blank field observations or model it. The background subtraction using the χ^2 statistic is a widely used technique; however, in the previous section (see Sect. 5.1) I showed that, for low-count spectra, the Cash statistic is more suitable than the χ^2 with reasonable grouping. Since the Cash statistic requires the number of counts in each channel to be greater than zero, the background has to be modeled (Cash 1979). I analyzed simulated spectra in both ways and compared the results. Hereafter I call "sub-χ^2" the standard analysis technique and "mod-C" the analysis using the Cash statistic and the background modeling.

I proceeded as for the source-only case, considering the Abell 1689 observation mentioned in Sect. 5.1 as a guideline. I produced a simulated spectrum choosing realistic input values for an absorbed thermal (see Sect. 5.1) plus background (see above) model (`WABS*MEKAL+PEGPWRLW/b` in XSPEC) and produced $N_{\rm meas}$ different measurements with a Poissonian perturbation of the number of counts in each channel. In the mod-C case, each spectrum was associated with the RMF and the ARF and was fitted between 2.0 and 10.0 keV with the `WABS*MEKAL+PEGPWRLW/b` model. The $N_{\rm H}$ and $\Gamma_{\rm B}$ were fixed to the input values, z was allowed to vary between 0.186 and 0.214 ($\pm$7% of the input value), kT, Z, $N_{\rm S}$, and $N_{\rm B}$ were free. I determined best-fit values and one-sigma uncertainties for all parameters. At last, I computed the weighted average of all $N_{\rm meas}$ values for each parameter using one-sigma uncertainties, and compared it with the input value. In the sub-χ^2 case, I simulated a background-only spectrum with a long exposure time. I considered a `PEGPWRLW/b` model (slope and normalization were equal to those of the power law in the source observation mentioned above) and I perturbed the expected spectrum as explained above. The adopted background spectrum was the same for all $N_{\rm meas}$ measurements, and its exposure time was 1000 ks. For each of the $N_{\rm meas}$ source spectra, I grouped the spectral channels to have at least 25 counts per bin and associated the background spectrum, the RMF, and the ARF with the binned spectrum. I fit the net spectrum with a thermal model only (`WABS*MEKAL` in XSPEC) in the 2.0–10.0 keV band to determine the best-fit values, then computed the weighted averages and compared them with the input values.

I considered two spatial regions: the ring between 1.0′ and 1.5′ centered on the

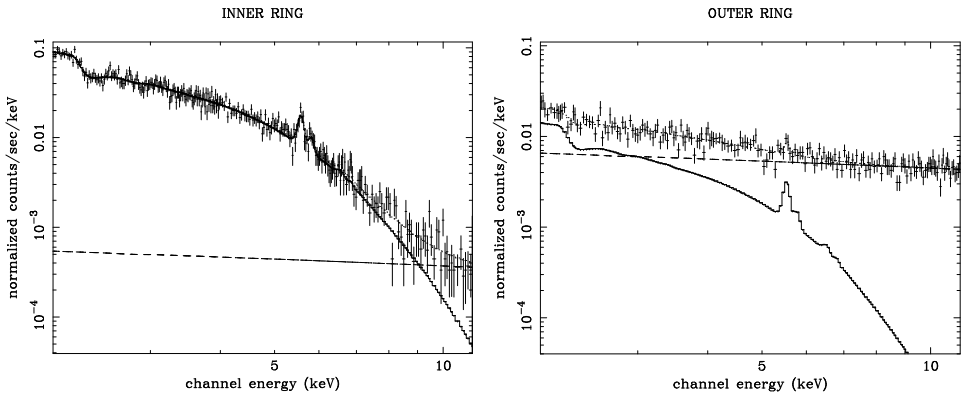

Figure 5.1: Simulated spectra accumulated in an inner ring (left panel) and in an outer ring (right panel). The solid line is the source contribution, the dashed line the background, and the dotted the sum. In the outer ring, beyond 3 keV, background counts dominate source counts. See text for further details and for model parameters.

cluster emission peak, where the source dominates the background (see Fig. 5.1, left panel) and the ring between $4.5'$ and $6.0'$, where the background dominates (see Fig. 5.1, right panel). Input values for the normalizations of both components were the best-fit values measured in the two rings of the Abell 1689 observation mentioned in Sect. 5.1. More specifically, in the inner ring $N_{\rm S} = 3.5 \times 10^{-3}$ and $N_{\rm B} = 1.5$, and in the outer ring $N_{\rm S} = 7.0 \times 10^{-4}$ and $N_{\rm B} = 17.5$ (XSPEC units). For each ring, I considered three input temperatures (namely 5, 7, and 9 keV) and two exposure times (10 and 100 ks).

In Table 5.3 I show the comparison between the two data analysis techniques described above (i.e. sub-χ^2 and mod-C). At first, I considered the inner ring, where the source dominates the background. The results are very similar to the case without background (see Sect. 5.1, Table 5.1). For the 100 ks case, mod-C returns the correct temperature and sub-χ^2 slightly underestimates it. For shorter exposure times, both techniques return a slightly biased value (bias $\approx 3\%$). No significant trend with the input temperature, kT_0, is found. Conversely, when considering the outer ring, where the background dominates, I find the true temperature to be strongly underestimated in all cases. There is a clear trend with the input temperature: the higher kT_0, the stronger the bias. For long exposure times, mod-C (bias $\approx 10\%$) works better than sub-χ^2 (bias $\approx 30\%$). For short exposure times, both techniques underestimate the true temperature by a factor of 2. These results are qualitatively similar to those found for the source-only case (see Table 5.1), but the bias is much stronger.

I have repeated the same analysis as described above for a particular set of simulated spectra (namely in the outer ring, with exposure time of 10 ks and $kT_0 =$ 7 keV) modeling the source with a bremsstrahlung rather than with a `MEKAL`. The bremsstrahlung model is simpler and can be expressed as an analytic function of its two free parameters (i.e. the temperature and the normalization). Conversely, the `MEKAL` model has two further parameters (metallicity and redshift), and its complicated dependency on the parameters is not expressed in an analytic form: the expected values are tabulated on a finite grid as a function of all parameters. For this particular set of spectra, the bias for `MEKAL` and bremsstrahlung models is the same, suggesting that the bias is not related to the approximation of a finite grid of values.

			sub-χ^2		mod-C	
Ring	Exp.[a]	kT_0[b]	kT[c]	$\Delta T/T_0$[d]	kT[c]	$\Delta T/T_0$[d]
1.0′–1.5′	100	5.00	4.84±0.01	−3.2 %	4.96±0.01	−0.8%
1.0′–1.5′	100	7.00	6.78±0.02	−3.1 %	6.97±0.02	−0.4%
1.0′–1.5′	100	9.00	8.69±0.02	−3.4 %	8.97±0.03	−0.3%
1.0′–1.5′	10	5.00	4.81±0.03	−3.8 %	4.82±0.03	−3.6%
1.0′–1.5′	10	7.00	6.78±0.05	−3.1 %	6.79±0.05	−3.0%
1.0′–1.5′	10	9.00	8.68±0.11	−3.6 %	8.62±0.08	−4.2%
4.5′–6.0′	100	5.00	3.95±0.01	−21.0 %	4.71±0.02	−5.8%
4.5′–6.0′	100	7.00	5.24±0.02	−25.1 %	6.45±0.03	−7.9%
4.5′–6.0′	100	9.00	6.43±0.02	−28.6 %	8.09±0.04	−10.1%
4.5′–6.0′	10	5.00	3.02±0.03	−39.6 %	3.20±0.03	−36.0%
4.5′–6.0′	10	7.00	3.68±0.04	−47.4 %	3.94±0.04	−43.7%
4.5′–6.0′	10	9.00	4.11±0.05	−54.3 %	4.52±0.06	−49.8%

Notes: [a] exposure time in ks; [b] input temperature in keV; [c] measured temperature in keV; [d] relative difference.

Table 5.3: Comparison between the results obtained using the sub-χ^2 and the mod-C data analysis techniques.

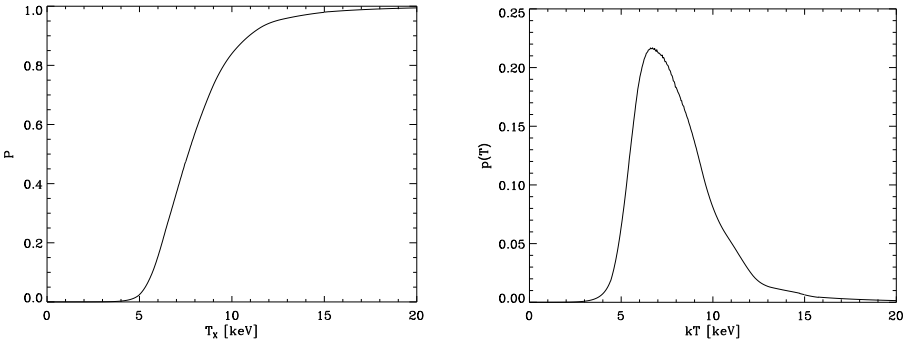

Figure 5.2: A cumulative distribution function (left panel) and the associated probability density function (right panel).

Some insight into the origin of the bias can be gained by inspecting the probability density function (p.d.f.) of the parameter of interest (in this case the temperature). Here I deal with the Cash statistic, and similar considerations apply to the χ^2. For each measurement, I defined as $C_{\rm min}$ the absolute minimum value of the function $C(\alpha)$. As in the previous section (see Sect. 5.1), I minimized $C(\alpha)$ (see Eq. 5.2) to determine the best estimate, $\alpha_{\rm best}$, of the parameter set ($C_{\rm min} \equiv C(\alpha_{\rm best})$). Cash (1979) has shown that the function ΔC (i.e. $C - C_{\rm min}$) follows a χ^2 distribution, therefore the confidence intervals can be generated in a standard way (e.g. using the XSPEC command ERROR). For each free parameter (here I considered only the temperature) I produced the function C with the XSPEC command STEPPAR. I calculated $\Delta C(T)$ and, since it is χ^2 distributed, I could associate with each temperature, $T_{\rm X}$, the probability that the true value is less or equal to $T_{\rm X}$, i.e. the cumulative distribution function (c.d.f.) of the temperature, $P(T_{\rm X})$. In Fig. 5.3 I compare representative $p(T)$ for single measurements in different conditions. For each case, the $p(T)$ is chosen randomly from the $N_{\rm meas}$ different measurements; therefore, the attention should be focused on the shapes of the p.d.f., rather than on temperature values. Clearly the curves become less symmetric and less similar

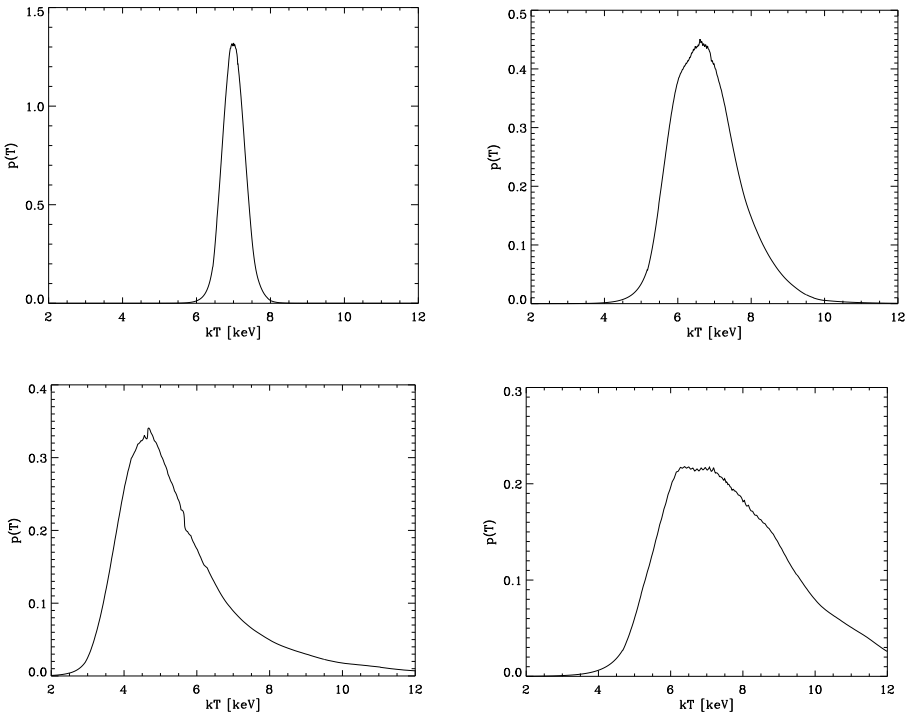

Figure 5.3: A comparison between representative $p(T)$ for single measurements extracted randomly in different conditions. The attention should be focused on the shapes of the p.d.f., rather than on temperature values. Top panels: the inner ring, where the source dominates over the background. Bottom panels: the outer ring, where the background is dominant. For left panels the exposure time is 100 ks, for right panels 10 ks. The input temperature is always 7 keV. Note that the scales in ordinate are different. Clearly the curves become less symmetric and less similar to Gaussians, as the exposure time decreases and the background contribution increases.

to Gaussians, as the exposure time decreases and the background contribution increases. The input temperature also plays a role: the higher the temperature, the less symmetric the curve. Summarizing, the poorer the statistical quality of the data, the more asymmetric the $p(T)$ and the stronger the bias.

The way measurements are combined does not change the result. I have experimented with two different methods: the weighted average of individual measurements and the product of individual p.d.f.. The weighted average roughly approximates the $p(T)$ to a Gaussian function and implies the contribution of high temperature tails be neglected. A more appropriate way to join information from different and independent measurements is to multiply single p.d.f.. The best value for the parameter corresponds to the maximum of the joined p.d.f.. I multiplied all $N_{\rm meas}$ p.d.f., computed as explained above, and still find a discrepancy between the best fit and true values. The bias is only slightly weaker than when computing a weighted average. I have also tried computing the $p(T)$ in a different way, i.e. using the parametric bootstrap technique (Press et al. 1992), which consists in creating and analyzing a large number of fake datasets starting from model best-fit values. I essentially obtain the same results.

In Table 5.3 I have showed that the strength of the bias mainly depends on the total number of counts and on the background contribution. One possibility for increasing the total counts is to extend the band to lower energies. I have explored it by analyzing one set of simulated spectra (namely in the outer ring, with an exposure time of 10 ks and $kT_0 = 7$ keV) between 0.5 and 10.0 keV. In this energy band the correlation between kT and $N_{\rm S}$ is weaker and the uncertainty on both parameters for a single measurement is smaller. When using the mod-C technique, the bias in the broad band is smaller ($\approx 10\%$ vs. $\approx 40\%$) than in the narrow band, suggesting that also the parameter degeneracy could play an important role when fitting in the 2.0–10.0 keV band. However, in practice, enlarging the band to lower energies has also drawbacks, such as the imperfect calibration of EPIC instruments and the presence of the galactic X-ray background (negligible beyond 2 keV), which introduce systematic effects that are hard to take into account. Moreover, broadband spectra are substantially more contaminated by emission from low temperature components located on the same line of sight as the dominant component (Mazzotta et al. 2004; Vikhlinin et al. 2005). For the analysis of real clusters I used the 0.7–10.0 keV energy band (see Chapter 4).

When considering the realistic case of a thermal source with a background, a stronger bias is expected (Eadie et al. 1971; Bergmann & Riisager 2002). As in the source-only case (see Sect. 5.1), I actually find that 1) the χ^2 and the Cash estimators are strongly biased, and that 2) the Cash estimator is less biased than the χ^2 one, especially for long exposure times. The strength of the bias depends mainly on two factors: the total number of counts and the background contribution.

5.3 Attempts to correct the bias

Having established that neither sub-χ^2 nor mod-C return acceptable results, one is faced with two alternative ways to proceed. A long-term solution to the problem requires that an unbiased, or perhaps a less biased, ML estimator be found and implemented within standard fitting packages (i.e. XSPEC). Another, faster, solution involves correcting the bias a posteriori, making use of extensive Monte Carlo simulations. In the following sections I show the main results obtained by exploring both approaches.

	100 ks		20 ks		10 ks	
Est.[a]	kT[b]	$\Delta T/T_0$[c]	kT[b]	$\Delta T/T_0$[c]	kT[b]	$\Delta T/T_0$[c]
T	6.44±0.03	−8.0%	4.96±0.05	−29.1%	4.04±0.07	−42.3%
T^{-1}	6.96±0.03	−0.6%	7.46±0.08	+6.6%	8.83±0.14	+26.1%
$T^{-1/2}$	6.88±0.03	−1.7%	6.59±0.06	−5.9%	6.36±0.09	−9.1%
$T^{-1/4}$	6.80±0.03	−2.9%	6.24±0.06	−10.9%	5.72±0.08	−18.3%

Notes: [a] temperature estimator; [b] measured temperature in keV; [c] relative difference.

Table 5.4: Comparison between the results obtained using different estimators of the temperature. The input temperature is 7 keV.

5.3.1 Using different estimators

From the literature (e.g. Cowan 1998), it is known that even if $\hat{X}$ is an unbiased estimator of X, $f(\hat{X})$ is not necessarily an unbiased estimator of $f(X)$. By reversing the argument one can argue that if $\hat{T}$ is a biased estimator of T, a transformation, f, may exist such that $f(\hat{T})$ is an unbiased (or at least less biased) estimator of $f(T)$. To test this idea I defined an analytic model, which I dubbed `BREM2`, similar to bremsstrahlung (1.11), by using the MDEFINE command within XSPEC.

$$S_{\rm T}(E) = N_{\rm S}\, E^{-4/3}\, T^{-1/2} \exp\left(-\frac{E}{T}\right) , \tag{5.4}$$

where the energy, E, is expressed in keV. The normalization, $N_{\rm S}$, is chosen to reproduce the same flux as a `MEKAL` with no metals. The Gaunt factor is well approximated by $E^{-4/3}$.

I simulated 3 sets of 3000 thermal plus background (`BREM2+PEGPWRLW/b`) spectra with the following input parameters: $N_{\rm S} = 7.2 \times 10^{-4}$, $N_{\rm B} = 17.5$, and $T = 7$ keV. These parameters correspond to those adopted in the case of the outer region (see Sect. 5.2). Each set had a different exposure time: 10, 20, and 100 ks. I defined 3 different estimators of the temperature,
$A = T^{-1}$,
$B = T^{-1/2}$,
$C = T^{-1/4}$,
and their respective models:
$S_{\rm A}(E) = N_{\rm S}\, E^{-4/3}\, A^{1/2} \exp\left(-A\, E\right)$,
$S_{\rm B}(E) = N_{\rm S}\, E^{-4/3}\, B \exp\left(-B^2\, E\right)$,
$S_{\rm C}(E) = N_{\rm S}\, E^{-4/3}\, C^2 \exp\left(-C^4\, E\right)$.
For simplicity I have considered only power laws as different $f(T)$. I fit each set of spectra with the listed models and measured the best estimate of $f(T)$. I computed the weighted average of the 3000 $f(T)$ and calculated T by using the inverse function, f^{-1}. In Table 5.4 I report the results of this analysis for different exposure times and estimators. The choice of the estimator strongly affects the bias. When using T as estimator, I obtained very similar results to those obtained with a `MEKAL`. This was expected, because the model `BREM2` is very similar to a bremsstrahlung (see Eq. 5.4); note also that in Sect. 5.2 I showed that the bias is roughly the same when using a `MEKAL` or a bremsstrahlung, as a model. When considering the bias as a function of the power-law index, I find a minimum corresponding to $T^{-1/2}$, which is the best estimator among those considered. For short exposure times (i.e. 10 ks) the use of $T^{-1/2}$ instead of T reduces the bias by a factor of 4. I suggest that this could be related to the degree of complexity of the derivative of $S(E)$ with respect to the estimator. Note also that, when slightly increasing the statistic (e.g. when considering 20 ks of exposure time), the bias associated with the $T^{-1/2}$ estimator is almost negligible if compared to typical statistical uncertainties. Such

Ring	Exp.[a]	kT_0[b]	kT[c]	$\Delta T/T_0$[d]
1.0′–1.5′	100	5.00	4.96±0.01	−0.8 %
1.0′–1.5′	100	7.00	6.97±0.01	−0.4 %
1.0′–1.5′	100	9.00	8.97±0.01	−0.3 %
1.0′–1.5′	10	5.00	4.88±0.02	−2.4 %
1.0′–1.5′	10	7.00	6.90±0.05	−1.4 %
1.0′–1.5′	10	9.00	8.81±0.13	−2.1 %
4.5′–6.0′	100	5.00	4.90±0.02	−2.0 %
4.5′–6.0′	100	7.00	6.77±0.04	−3.3 %
4.5′–6.0′	100	9.00	8.51±0.09	−5.4 %
4.5′–6.0′	10	5.00	4.68±0.13	−6.4 %
4.5′–6.0′	10	7.00	5.90±0.24	−15.7 %
4.5′–6.0′	10	9.00	7.67±0.51	−14.8 %

Notes: [a] exposure time in ks; [b] input temperature in keV; [c] measured temperature in keV; [d] relative difference.

Table 5.5: Results obtained fitting $p(T)$ with a log-normal distribution.

results encourage exploration of this approach (i.e. to consider different estimators) in order to find a rigorously derived unbiased estimator of the temperature.

5.3.2 Fitting with a log-normal function

The shape of the $p(T)$ resembles the log-normal function, which is the p.d.f. of any random variable whose logarithm is normally distributed. If X is a random variable with a normal distribution, then $x \equiv \exp(X)$ has a log-normal distribution. The log-normal distribution has p.d.f.

$$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}}\, \mathrm{e}^{-(\ln x-\mu)^2/2\sigma^2} \tag{5.5}$$

for $x > 0$, where μ and σ are respectively the mean and the standard deviation of the variable's logarithm. The expected value is

$$\mathrm{E}(X) = \mathrm{e}^{\mu+\sigma^2/2}\,, \tag{5.6}$$

and the variance is

$$\mathrm{var}(X) = (\mathrm{e}^{\sigma^2} - 1)\, \mathrm{e}^{2\mu+\sigma^2}\,. \tag{5.7}$$

I fit each $p(T)$ with a log-normal function, $f(x;\mu,\sigma)$ (see Eq. 5.5) and calculated the best values of μ_i and σ_i. I computed a weighted average of μ_i using σ_i^{-2} as weights and calculated the expected value (see Eq. 5.6) and the uncertainty, i.e. the variance (see Eq. 5.7) divided by the square root of the number of measurements. In Table 5.5 I report the results of these test calculations. In all cases, this method provides better results than a simple weighted average (see Table 5.3 for a comparison). There is still a bias of a few percent, except for the case of the outer ring with 10 ks: in this case the bias is greater than 10%. Thus, when the background contribution is small, the log-normal distribution provides a good estimate, while the result is still biased when the background is dominant, especially for few total counts, but much less than when using the standard techniques.

5.3.3 A semi-empirical method: summing three distributions

The three EPIC instruments (MOS1, MOS2, and pn) on board *XMM-Newton* provide three simultaneous and independent measurements of the same target; therefore,

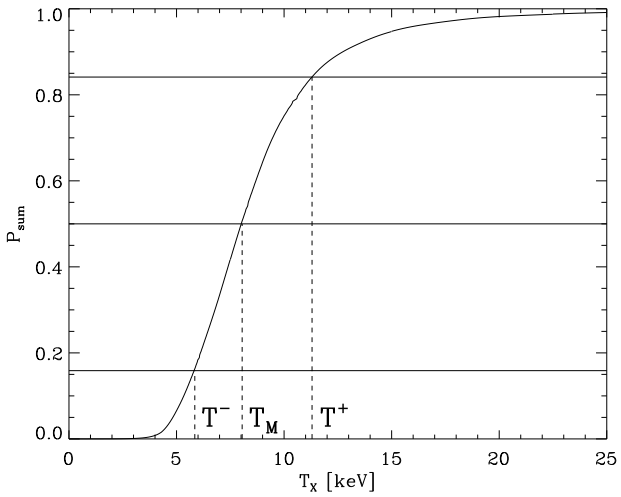

Figure 5.4: Visual representation of the definition of T^-, $T_{\rm M}$, and T^+ from a joined cumulative distribution function.

when dealing with EPIC data, one needs to correctly combine these three measurements. A weighted average is the simplest procedure, however in Sect. 5.2 I have shown that it leads to biased results. In Sect. 5.2 I have also shown that the strength of the bias is related to the shape of the p.d.f. and in Sect. 5.3.2 that a fit with a log-normal function does not return sufficiently accurate results. In this section, I chose a different procedure, emphasizing the contribution of p.d.f. tails at high temperature. I derived $N_{\rm meas}$ measurements of the temperature with their corresponding p.d.f., as in the mod-C case described in Sect. 5.2. I divided the $N_{\rm meas}$ measurements into groups of three; then, for each group, I considered the three p.d.f., $p_{\rm i}(T)$ (i=1,2,3), and combined them in a non-standard way by calculating the sum, rather than the product, of the single p.d.f.. In practice, it is equivalent, but more useful, to sum the c.d.f., $P_{\rm i}(T_{\rm X})$ directly. The sum is renormalized by dividing it by 3. I defined $P_{\rm sum}(T_{\rm X})$ as

$$P_{\rm sum}(T_{\rm X}) = \frac{1}{3}\sum_{i=1}^{3} P_{\rm i}(T_{\rm X}) \ , \tag{5.8}$$

which is a sort of joined c.d.f. of three measurements, and the associated p.d.f. is usually more symmetric than the single $p_{\rm i}(T)$. I defined as T^-, $T_{\rm M}$, and T^+ the temperatures that correspond to a probability, $P_{\rm sum}(T_{\rm X})$, of 0.1587, 0.5000, and 0.8413, respectively (see Fig. 5.4). I considered $T_{\rm M}$ as the best estimate for the three joined measurements, $\mathrm{d}T^- \equiv (T_{\rm M} - T^-)/\sqrt{3}$ as the lower uncertainty and $\mathrm{d}T^+ \equiv (T^+ - T_{\rm M})/\sqrt{3}$ as the upper uncertainty. Thus I have $N_{\rm meas}/3$ "triplet" measurements: $T_{\rm M}{}^{+\mathrm{d}T^+}_{-\mathrm{d}T^-}$. I computed the weighted average of the $N_{\rm meas}/3$ "triplets", and in Table 5.6 compare the results with the input values. In almost all cases this semi-empirical method (hereafter "triplet" method) provides excellent results: the discrepancy is lower than 2% and often comparable with the statistical uncertainty.

I have tried joining different numbers of measurements together. Simulations show that, when considering two measurements at a time, the temperature results slightly underestimated; instead, when considering five measurements, I obtain substantially correct results, as when using the "triplets". This suggests that the effectiveness of this a posteriori correction depends on the number of measurements combined.

Ring	Exp.[a]	kT_0[b]	kT[c]	$\Delta T/T_0$[d]
1.0′–1.5′	100	5.00	4.97±0.01	−0.6 %
1.0′–1.5′	100	7.00	6.97±0.02	−0.4 %
1.0′–1.5′	100	9.00	8.98±0.03	−0.2 %
1.0′–1.5′	10	5.00	4.90±0.04	−2.0 %
1.0′–1.5′	10	7.00	6.94±0.07	−0.9 %
1.0′–1.5′	10	9.00	8.98±0.11	−0.2 %
4.5′–6.0′	100	5.00	5.00±0.02	−0.0 %
4.5′–6.0′	100	7.00	6.90±0.04	−1.4 %
4.5′–6.0′	100	9.00	8.91±0.05	−1.0 %
4.5′–6.0′	10	5.00	5.04±0.06	+0.8 %
4.5′–6.0′	10	7.00	6.97±0.09	−0.4 %
4.5′–6.0′	10	9.00	8.96±0.13	−0.4 %

Notes: [a] exposure time in ks; [b] input temperature in keV; [c] measured temperature in keV; [d] relative difference.

Table 5.6: Results obtained with the semi-empirical "triplet" method.

I want to stress that this technique is not rigorously derived from principles of statistics, but it is able to correct the bias and returns the expected temperature under very different situations (e.g. different background contributions and exposure times) when using the 2.0–10.0 keV band. This could be related to the fact that joined p.d.f. are usually more symmetric than the single ones.

Chapter 6

Radial profiles

6.1 The sample

I selected from the *XMM-Newton* archive a sample of hot ($kT > 3.3$ keV), intermediate redshift ($0.1 \lesssim z \lesssim 0.3$), and high galactic latitude ($|b| > 20°$) clusters of galaxies. Upper and lower limits to the redshift range are determined, respectively, by the cosmological dimming effect and the size of the EPIC field of view ($\approx 15'$ radius). Indeed, our data analysis technique requires that the intensity of background components be estimated in a peripheral region, where the cluster emission is almost negligible (see Sect. 4.2.1). I retrieved from the public archive all observations of clusters satisfying the above selection criteria, performed before March 2005 (when the CCD6 of EPIC MOS1 was switched off[1]) and available at the end of May 2007. Unfortunately, 23 of these 86 observations are highly affected by soft proton flares (see Table 6.1) and I excluded them from the sample, as mentioned in Sect. 4.1. Furthermore, I excluded 14 observations of clusters that show evidence of recent and strong interactions (see Table 6.2). For such clusters, a radial analysis is not appropriate, because the gas distribution is far from being azimuthally symmetric. Finally, I find that the target of observation 0201901901, which is classified as a cluster, is probably a point-like source; therefore, I excluded this observation too from the sample.

In Table 6.3 I list the 48 observations that survived the above mentioned selection criteria and report cluster physical properties (i.e. redshift, temperature, and scale radius) and observation characteristics (i.e. exposure time, intensity of residual soft protons, filter). The redshift value (from optical measurements) is taken from the NASA Extragalactic Database[2]; $kT_{\rm M}$ and R_{180} are derived from the analysis (see Sect. 6.2). In Fig. 6.1 I report the cluster distribution in the redshift-temperature space. The only selection effect I detect is the paucity of cool ($kT_{\rm M} < 5$ keV) clusters at high ($z > 0.2$) redshift. In Fig. 6.2 I report the histograms of the frequency distribution for observation exposure times and for $R_{\rm SB}$ values (see Sect. 4.1.2). Since the observation 0084230401 of Abell 267 is extremely polluted by QSP ($R_{\rm SB} = 1.8$), I excluded it from the sample. Observations were performed with THIN1 and MEDIUM filters, as reported in Table 6.3.

All results are given assuming a ΛCDM cosmology with $\Omega_{\rm m} = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70$ km s^{-1} Mpc^{-1}.

[1] http://xmm.esac.esa.int/external/xmm_news/items/MOS1-CCD6/index.shtml
[2] http://nedwww.ipac.caltech.edu

Name	Obs ID
RXCJ0303.8–7752	0042340401
RXCJ0516.7–5430	0042340701
RXCJ0528.9–3927	0042340801
RXCJ2011.3–5725	0042341101
Abell 2537	0042341201
RXCJ0437.1+0043	0042341601
Abell 1302	0083150401
Abell 2261	0093030301
Abell 2261	0093030801
Abell 2261	0093030901
Abell 2261	0093031001
Abell 2261	0093031101
Abell 2261	0093031401
Abell 2261	0093031501
Abell 2261	0093031601
Abell 2261	0093031801
Abell 2219	0112231801
Abell 2219	0112231901
RXCJ0006.0–3443	0201900201
RXCJ0145.0–5300	0201900501
RXCJ0616.8–4748	0201901101
RXCJ0437.1+0043	0205330201
Abell 2537	0205330501

Table 6.1: Observations excluded from the sample due to high soft proton contamination.

Name	Obs ID
Abell 2744	0042340101
Abell 665	0109890401
Abell 665	0109890501
Abell 1914	0112230201
Abell 2163	0112230601
Abell 2163	0112231501
RXCJ0658.5–5556	0112980201
Abell 1758	0142860201
Abell 1882	0145480101
Abell 901	0148170101
Abell 520	0201510101
Abell 2384	0201902701
Abell 115	0203220101
ZwCl2341.1+0000	0211280101

Table 6.2: Observations of clusters that show evidence of recent and strong interactions.

Name	Obs ID	z[a]	k$T_{\rm M}$[b]	R_{180}[c]	Exp.[d]		$R_{\rm SB}$[e]	F[f]
RXCJ0043.4–2037	0042340201	0.2924	6.8	1.78	11.9	11.3	1.25	T
RXCJ0232.2–4420	0042340301	0.2836	7.2	1.85	12.1	11.7	1.08	T
RXCJ0307.0–2840	0042340501	0.2534	6.8	1.82	11.4	12.6	1.08	T
RXCJ1131.9–1955	0042341001	0.3072	8.1	1.93	12.4	12.3	1.08	T
RXCJ2337.6+0016	0042341301	0.2730	7.2	1.86	13.4	13.1	1.19	T
RXCJ0532.9–3701	0042341801	0.2747	7.5	1.90	10.9	10.5	1.09	T
Abell 68	0084230201	0.2550	7.2	1.88	26.3	25.9	1.37	M
Abell 209	0084230301	0.2060	6.6	1.85	17.9	17.8	1.19	M
Abell 267	0084230401*	0.2310	4.5	1.49	17.0	16.5	1.79	M
Abell 383	0084230501	0.1871	4.4	1.52	29.3	29.8	1.33	M
Abell 773	0084230601	0.2170	7.5	1.96	13.6	15.5	1.16	M
Abell 963	0084230701	0.2060	6.5	1.83	24.4	26.0	1.19	M
Abell 1763	0084230901	0.2230	7.2	1.92	13.0	13.2	1.08	M
Abell 1689	0093030101	0.1832	9.2	2.21	36.8	36.8	1.14	T
RX J2129.6+0005	0093030201	0.2350	5.5	1.66	36.0	37.5	1.21	M
ZW 3146	0108670101	0.2910	7.0	1.81	52.9	52.9	1.07	T
E1455+2232	0108670201	0.2578	5.0	1.56	35.3	35.8	1.11	M
Abell 2390	0111270101	0.2280	11.2	2.37	9.9	10.3	1.11	T
Abell 2204	0112230301	0.1522	8.5	2.16	18.2	19.5	1.06	M
Abell 1413	0112230501	0.1427	6.7	1.92	25.4	25.4	1.10	T
Abell 2218	0112980101	0.1756	6.5	1.86	18.2	18.2	1.17	T
Abell 2218	0112980401	0.1756	7.0	1.93	13.7	14.0	1.42	T
Abell 2218	0112980501	0.1756	6.1	1.80	11.3	11.0	1.07	T
Abell 1835	0147330201	0.2532	8.6	2.05	30.1	29.2	1.16	T
Abell 1068	0147630101	0.1375	4.5	1.58	20.5	20.8	1.09	M
Abell 2667	0148990101	0.2300	7.7	1.96	21.9	21.6	1.48	M
Abell 3827	0149670101	0.0984	7.1	2.02	22.3	22.4	1.16	M
Abell 3911	0149670301	0.0965	5.4	1.77	25.8	26.1	1.43	T
Abell 2034	0149880101	0.1130	7.0	1.99	10.2	10.5	1.16	T
RXCJ0003.8+0203	0201900101	0.0924	3.7	1.47	26.3	26.6	1.10	T
RXCJ0020.7–2542	0201900301	0.1424	5.7	1.78	14.8	15.4	1.02	T
RXCJ0049.4–2931	0201900401	0.1080	3.3	1.37	19.2	18.8	1.28	T
RXCJ0547.6–3152	0201900901	0.1483	6.7	1.92	23.3	24.0	1.12	T
RXCJ0605.8–3518	0201901001	0.1410	4.9	1.65	18.0	24.1	1.07	T
RXCJ0645.4–5413	0201901201	0.1670	7.1	1.95	10.9	10.9	1.11	T
RXCJ1044.5–0704	0201901501	0.1323	3.9	1.47	25.7	25.9	1.03	T
RXCJ1141.4–1216	0201901601	0.1195	3.8	1.46	28.4	28.6	1.03	T
RXCJ1516.3+0005	0201902001	0.1183	5.3	1.73	26.7	26.6	1.13	T
RXCJ1516.5–0056	0201902101	0.1150	3.8	1.46	30.0	30.0	1.08	T
RXCJ2014.8–2430	0201902201	0.1612	7.1	1.96	23.0	23.4	1.05	T
RXCJ2048.1–1750	0201902401	0.1470	5.6	1.75	24.6	25.3	1.07	T
RXCJ2149.1–3041	0201902601	0.1179	3.3	1.37	25.1	25.5	1.11	T
RXCJ2218.6–3853	0201903001	0.1379	6.4	1.88	20.2	21.4	1.11	T
RXCJ2234.5–3744	0201903101	0.1529	8.6	2.17	18.9	19.3	1.31	T
RXCJ0645.4–5413	0201903401	0.1670	8.5	2.13	11.5	12.1	1.51	T
RXCJ0958.3–1103	0201903501	0.1527	6.1	1.83	8.3	9.4	1.16	T
RXCJ0303.8–7752	0205330101	0.2742	7.5	1.89	11.7	11.5	1.18	T
RXCJ0516.7–5430	0205330301	0.2952	7.5	1.87	11.4	11.7	1.19	T

Notes: [a] redshift taken from the NASA Extragalactic Database; [b] mean temperature in keV derived from the analysis; [c] scale radius in Mpc derived from the analysis; [d] MOS1 and MOS2 good exposure time in ks; [e] intensity of residual soft protons (see Eq. 4.1); [f] filter wheel: THIN1 (T) or medium (M); * excluded due to high residual soft proton contamination.

Table 6.3: Physical properties and observation details for the 48 clusters analyzed.

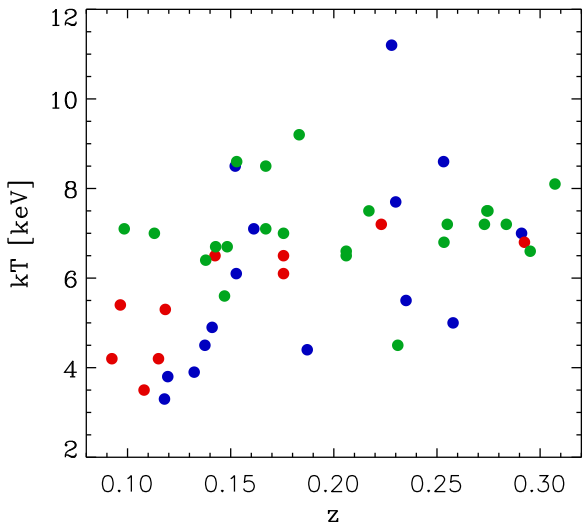

Figure 6.1: Distribution of selected clusters in the redshift-temperature space. I distinguish cool core (blue), non-cool core (red) and uncertain (green) clusters, as defined in Sect. 6.2. There is no evidence of selection effects, except for a weak positive correlation between redshift and temperature.

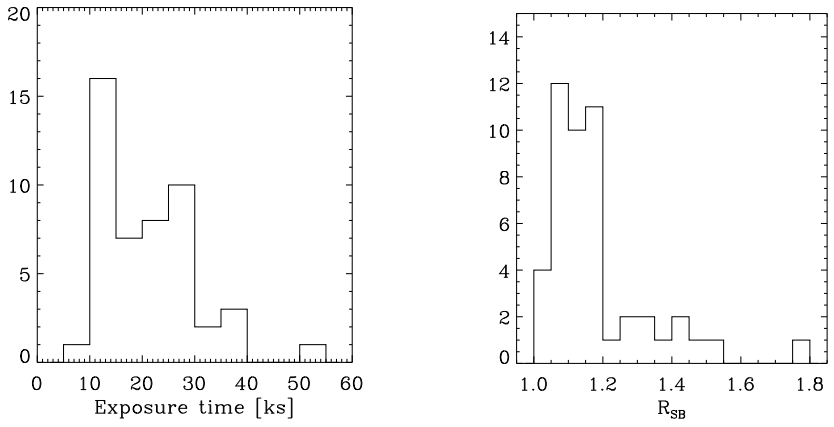

Figure 6.2: Histograms of the frequency distribution for averaged MOS exposure time (left panel) and $R_{\rm SB}$ (right panel) values.

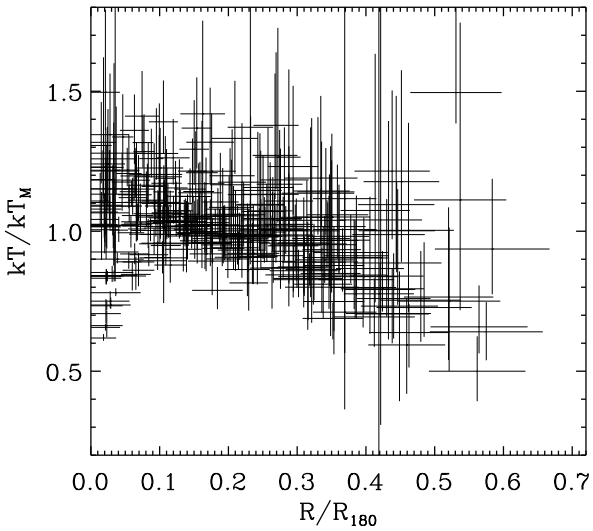

Figure 6.3: Radial temperature profiles for all clusters in my sample rescaled by R_{180} and $kT_{\rm M}$.

6.2 Temperature and metallicity profiles

Clusters in my sample have different temperatures and redshifts, therefore it is not trivial to identify one (or more) parameters that indicate the last ring where the temperature measurement is reliable. I define an indicator, I, as the source-to-background count rate ratio in the energy band used for the spectral fitting. For each observation and each ring, I calculated I: the higher I, the more the source contribution, the more reliable the measurement in that particular ring. The indicator I is affected by an intrinsic bias; i.e., upward statistical fluctuations of the temperature are associated with higher I (because of the difference in spectral shape between source and background models); therefore, near a threshold, the mean temperature is slightly overestimated. This systematic effect is almost negligible when considering the whole sample, but may appear when analyzing a small number of objects; however, I note that, although present, it does not affect results obtained when dividing the whole sample into subsamples (e.g. Sects. 7.1.2 and 8.1.2).

In Fig. 6.3 I show the radial temperature profiles for all clusters in my sample by setting a lower limit $I_0 = 0.6$; spectra are fitted in the 0.7–10.0 keV band. Each profile is rescaled by the cluster mean temperature, $kT_{\rm M}$, computed by fitting the profile with a constant after the exclusion of the core region (i.e. for $R > 0.1\ R_{180}$). The radius is rescaled by R_{180}, i.e. the radius encompassing a spherical density contrast of 180 with respect to the critical density. I computed R_{180} from the mean temperature and the redshift (Arnaud et al. 2005):

$$R_{180} = 1780 \left(\frac{kT_{\rm M}}{5\,{\rm keV}}\right)^{1/2} h(z)^{-1}\ {\rm kpc}, \tag{6.1}$$

where $h(z) = (\Omega_{\rm m}(1+z)^3+\Omega_\Lambda)^{1/2}$. The scale radius, R_{180}, is a good approximation to the virial radius in an Einstein-De Sitter universe and has been largely used to rescale cluster radial properties (e.g. Markevitch et al. 1998; De Grandi & Molendi 2002; Vikhlinin et al. 2005). I then used 180 as an over-density for comparing my results with previous works (see Sect 8.1.6), even if the virial radius encloses a

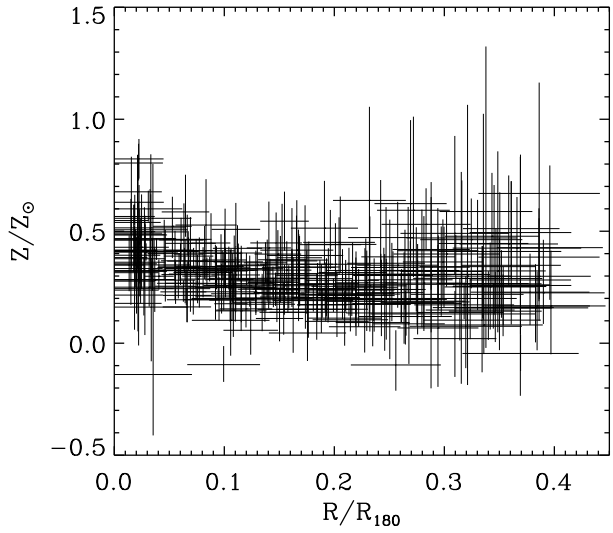

Figure 6.4: Radial metallicity profiles for all clusters in my sample. Abundances are expressed in Anders & Grevesse (1989) solar values and radii in units of R_{180}.

spherical density contrast of about 100 in the current adopted cosmology (Eke et al. 1998).

In Fig. 6.4 I show radial metallicity profiles for all clusters in my sample. All metallicity measurements are relative to the solar values published by Anders & Grevesse (1989). Even if these have been superseded by more recent values (Grevesse & Sauval 1998; Asplund et al. 2005), they allow straightforward comparison with most of the literature (e.g. De Grandi & Molendi 2001; Balestra et al. 2007; Baldi et al. 2007). Grevesse & Sauval (1998) and Asplund et al. (2005) introduced a 0.676 and 0.60 times lower iron solar abundance respectively, while other elements are substantially unchanged. A simple scaling by 0.676 and 0.60 converts measures from the Anders & Grevesse (1989) iron abundance to the Grevesse & Sauval (1998) and Asplund et al. (2005) abundances. The radius is rescaled by R_{180}, as for the temperature profiles. The profiles are limited to $\approx 0.4\ R_{180}$ because beyond that radius the source-to-background count rate ratio is too small and the measurements are unreliable.

It is worth noting that, especially for temperature profiles, the error bars are usually strongly asymmetric; i.e., the upper bar is larger than the lower. Moreover, the higher the temperature, the larger the error bars. The reason is that most of the information on the temperature is located around the energy of the exponential cut-off. Due to the spectral shapes of source and background components and to the sharp decrease of the effective areas at high energies, the source-to-background count rate ratio strongly depends on the energy band (see for example Fig. 4.1); i.e., the higher the cut-off energy, the lower the source-to-background ratio and the larger the uncertainties.

In Fig. 6.5 I report the mean profiles binned in units of R_{180}; these values have been computed by performing, for each new bin, a weighted average of temperature (or metallicity) values in the original bins which have a non-zero intersection with the new bin. The weight is the product of two components: one is the inverse squared errors, the other depends on the intersection between the original bin and the new one. If the original bin is totally included into the new one, the weight is equal to one. If the original bin has only a partial intersection with the new one,

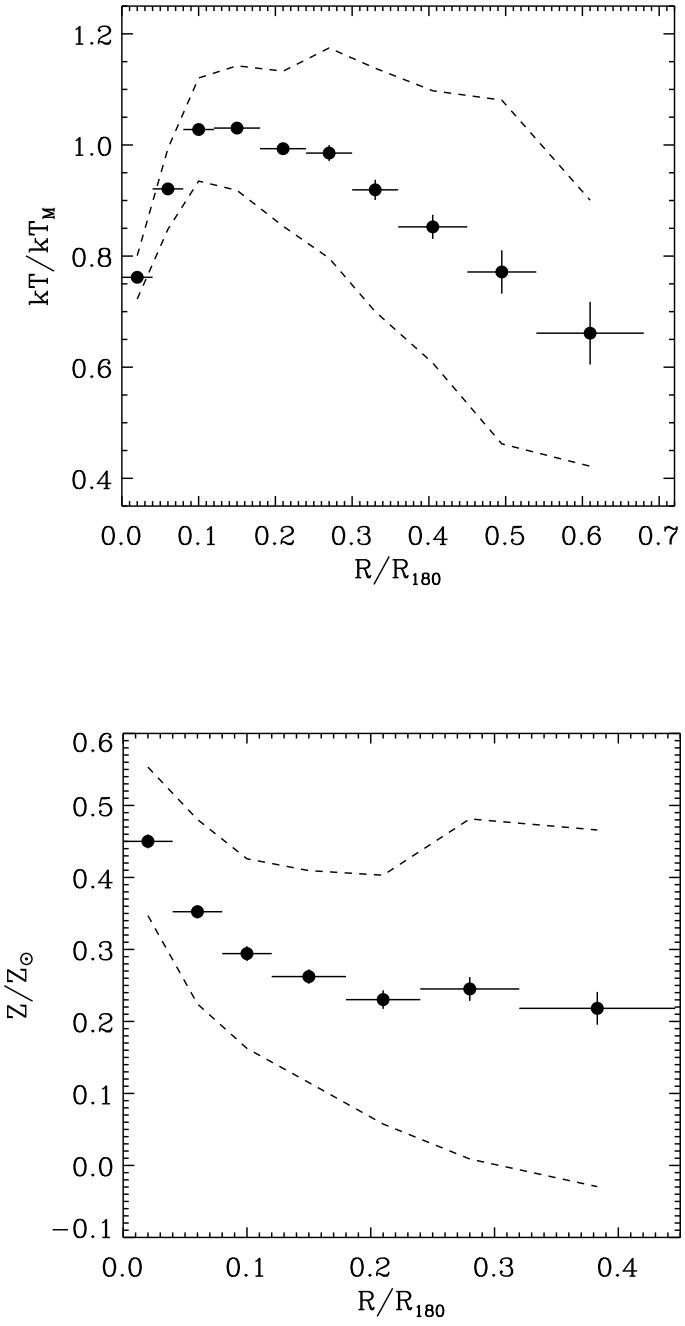

Figure 6.5: Mean radial temperature (top panel) and metallicity (bottom panel) profiles averaged over all clusters. Temperatures are rescaled by k$T_{\rm M}$, abundances are expressed in Anders & Grevesse (1989) solar values, and radii are in units of R_{180}. The dotted lines show the one-sigma scatter of the values around the average, for both profiles.

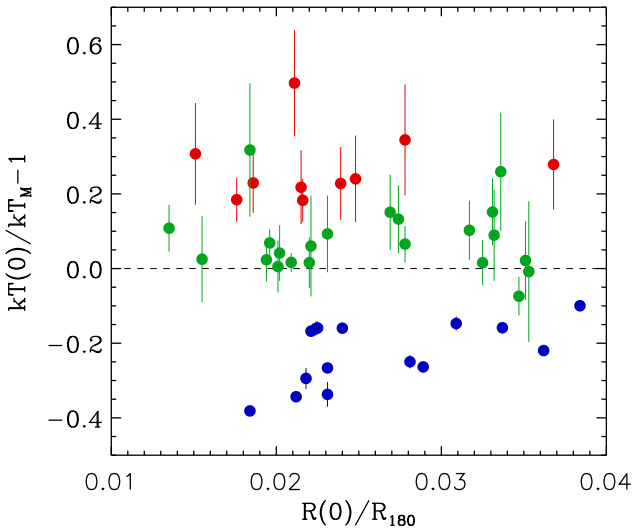

Figure 6.6: Temperature vs. radius for the innermost ring respectively scaled by kT_M and R_{180}. Clusters for which the temperature is significantly (at least 3σ) lower than kT_M are defined as cool cores (blue circles); those for which the temperature profile does not significantly (at least 2σ) decrease are defined as non-cool cores (red circles); other clusters, whose membership is not clearly determined, are classified as uncertain (green circles). When considering distant clusters, which fill the right side of the panel, smaller gradients are expected due to the lower spatial resolution.

the weight is the fraction of the original bin that belongs to the new one. Possible blurring effects associated either with the original binning in angular units or with the recasting in units of R_{180} should be minimal. As far as the original binning is concerned, I note that the size of central bins (i.e. $30''$) is comparable to the *XMM-Newton* PSF. As far as the recasting is concerned, what I show in Fig. 6.5 is the result of various trials specifically aimed at reaching the best compromise between resolution and statistical quality; moreover, the size of central bins (i.e. 0.04 R_{180}) corresponds to $\approx 30''$ for clusters at $z \approx 0.2$ with a temperature of 7 keV.

The temperature profiles (see Fig. 6.3) show a clear decline beyond $\approx 0.2\ R_{180}$ and toward the center, because of the presence of cool core clusters (see Sect. 1.4.3). The measurements extend out to $\approx 0.6\ R_{180}$. The large scatter of values is mostly of statistical origin; however, a maximum likelihood test shows that, when excluding the region below $0.2\ R_{180}$, profiles are characterized by a 6% intrinsic dispersion, which is comparable to the systematics (see Sect. 7.1.3), so that the existence of a universal cluster temperature profile is still an open issue. The scatter in the inner region is mostly due to the presence of both cool core and non-cool core clusters, but also to the choice of preserving the azimuthal symmetry at large radii (see Sect. 4.1.3). In Fig. 6.6 I report temperature and radius of the innermost ring scaled by kT_M and R_{180} for all clusters. I define (i) cool core (hereafter CC) clusters as those for which the temperature is significantly (at least 3σ) lower than kT_M, (ii) non-cool core (hereafter NCC) clusters as those for which the temperature profile does not significantly (at least 2σ) decrease, and (iii) uncertain (hereafter UNC) clusters as those for which the membership is not clearly determined. In Table 6.4 I report the membership for each cluster.

The mean metallicity (see Fig. 6.5, right) is 0.45 $Z_\odot$ in the center and decreases out to $\approx 0.2\ R_{180}$; beyond this radius the profile is consistent with being flat, a fit

CC[a]	NCC[b]	UNC[c]
Abell 383	RXCJ0043.4–2037	RXCJ0232.2–4420
RX J2129.6+0005	Abell 1763	RXCJ0307.0–2840
ZW 3146	Abell 2218	RXCJ1131.9–1955
E1455+2232	Abell 3911	RXCJ2337.6+0016
Abell 2390	RXCJ0003.8+0203	RXCJ0532.9–3701
Abell 2204	RXCJ0020.7–2542	Abell 68
Abell 1835	RXCJ0049.4–2931	Abell 209
Abell 1068	RXCJ1516.3+0005	Abell 773
Abell 2667	RXCJ1516.5–0056	Abell 963
RXCJ0605.8–3518		Abell 1689
RXCJ1044.5–0704		Abell 1413
RXCJ1141.4–1216		Abell 3827
RXCJ2014.8–2430		Abell 2034
RXCJ2149.1–3041		RXCJ0547.6–3152
RXCJ0958.3–1103		RXCJ0645.4–5413
		RXCJ2048.1–1750
		RXCJ2218.6–3853
		RXCJ2234.5–3744
		RXCJ0303.8–7752
		RXCJ0516.7–5430

Notes: [a] cool core clusters; [b] non-cool core clusters; [c] uncertain clusters.

Table 6.4: Cool core, non-cool core, and uncertain clusters.

with a constant for $R > 0.2\ R_{180}$ gives $Z = 0.23 \pm 0.01\ Z_{\odot}$. The profiles show a large scatter, which is mostly of statistical origin. In the central regions ($R < 0.2\ R_{180}$) I find an intrinsic scatter of $22 \pm 2\%$ related to the presence of cool core clusters; in the outer regions ($R > 0.2\ R_{180}$) the intrinsic scatter is only $14 \pm 8\%$ (i.e. $\approx 0.03\ Z_{\odot}$), the same order of magnitude as the systematics (see Sect. 7.1). Past works (e.g. De Grandi & Molendi 2001) have shown that the abundance profiles of cool core and non-cool core clusters differ in the central regions. I found qualitatively similar results, discussed in Sect. 8.2.1 and in Chapter 10.

Chapter 7

Discussion of systematic effects

7.1 Evaluation of systematics on temperature measures

I carefully checked my results, searching for possible systematic effects. Prior to the analysis, I made use of extensive simulations to quantify the impact of different spectral components on a simulated temperature profile ("a priori" tests). After the analysis, I investigated how the measured temperature profile changed, when choosing different key parameters ("a posteriori" tests).

7.1.1 "A priori" tests

I performed simulations that reproduced our analysis procedure as closely as possible (see Chapter 4). I considered two rings: the external $10'$–$12'$, $R_{\rm ext}$, where the background parameters are estimated, and the $4.5'$–$6'$, $R_{\rm int}$, where the temperature is measured. The exposure time for each spectrum is always 20 ks, i.e. a representative value for my sample (see Fig. 6.2). As in Chapter 5, I used the Abell 1689 EPIC MOS1 observation as a guideline, for producing RMF and ARF, and for choosing typical input parameters. The simulation procedure is structured as follows:

- choice of reasonable input parameters,
- generation of 300 spectra in $R_{\rm ext}$,
- generation of 500 spectra in $R_{\rm int}$,
- estimate of background parameters in $R_{\rm ext}$,
- rescaling background parameters and fitting spectra in $R_{\rm int}$.

Simulation details are described in each subsection. I tested the effect of the cosmic variance, of an inaccurate estimate of the cluster emission in $R_{\rm ext}$, and of the QSP component. All results were obtained by fitting spectra in the 0.7–10.0 keV band. I also conducted a similar analysis for the 2.0–10.0 keV band, finding that the systematics for the two bands are the same order of magnitude. I recall, however, that the hard band is characterized by the worst statistics; therefore, in this case systematic errors are masked by statistical ones and have less impact on the final measurement.

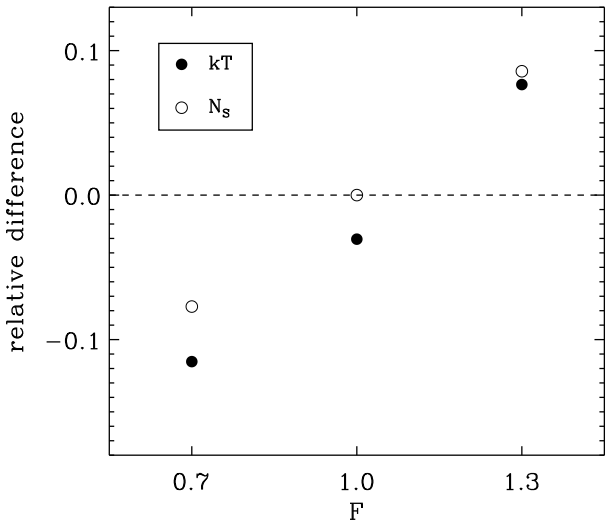

Figure 7.1: Relative differences between measured and input values for the source temperature, kT, and normalization, $N_{\rm S}$, as a function of the factor F, which simulates the fluctuation due to the cosmic variance (see text for details). Uncertainties are smaller than the circle size. 30% fluctuations cause $\approx 10\%$ variations in kT and $N_{\rm S}$. For a positive/negative fluctuation, the measured kT and $N_{\rm S}$ are higher/lower than the input values.

The cosmic variance

I employed a simulation to quantify the effect of the cosmic variance on temperature and normalization measurements. In this simulation I neglected the soft proton contribution; the background components are the HALO, the CXB, and the NXB, and are modeled as for MOS1 in Sect. 3.2.2. In $R_{\rm ext}$ there were only background components, while in $R_{\rm int}$ there was also the thermal source. Normalization[1] input values in $R_{\rm ext}$ were: $N^{\rm ext}_{\rm HALO} = 1.6 \times 10^{-4}$, $N^{\rm ext}_{\rm CXB} = 5.0 \times 10^{-2}$, and $N^{\rm ext}_{\rm NXB} = 1.0 \times 10^{-2}$; input values in $R_{\rm int}$ were obtained by rescaling the values in $R_{\rm ext}$ by the area ratio (i.e. as in Eq. 4.2 with $K(r) = 1.0$). Then, $N_{\rm CXB}$ was also multiplied by a factor, F, which simulates the fluctuation due to the cosmic variance between $R_{\rm int}$ and $R_{\rm ext}$. After the excision of brightest point-like sources (see Sect. 4.1.1), one-sigma fluctuations were expected to be $\approx 30\%$. I then considered 3 cases: a null ($F = 1.0$), a positive ($F = 1.3$), and a negative ($F = 0.7$) fluctuation. In the first case the input value for CXB in $R_{\rm int}$ was equal to that rescaled by the area ratio, in the second it was 30% higher, and in the third 30% lower. Input parameters for the thermal model in $R_{\rm int}$ were k$T = 6$ keV, $Z = 0.2\ Z_{\odot}$, $z = 0.2$, and $N_{\rm S} = 7.0 \times 10^{-4}$. In $R_{\rm ext}$, Z and z were fixed to the input values, while kT and $N_{\rm S}$ were free. For this particular choice of the parameters, the source-to-background count rate ratio, I, is 1.13 (see Sect. 6.2). As explained in Sects. 4.2.1 and 4.2.2, I determined the ranges of variability for $N_{\rm HALO}$, $N_{\rm CXB}$, and $N_{\rm NXB}$ and rescaled them in $R_{\rm int}$. I then fit spectra in the 0.7–10.0 keV band and calculated the weighted averages of kT and $N_{\rm S}$ over the 500 simulations.

In Fig. 7.1 I show the relative differences between measured and input values for the temperature, kT (filled circles), and the normalization, $N_{\rm S}$ (empty circles). A positive fluctuation of CXB normalization (i.e. $F = 1.3$) returns higher temperature

[1]Normalization values are always reported in XSPEC units

and normalization, because the excess of counts due to the CXB is modeled by the thermal component, which is steeper than the CXB power law. For the $F = 1.0$ case, while N_S returns exactly the input value, kT returns a slightly ($\approx 3\%$) underestimated value, probably due to the bias on the temperature estimator (see Chapter 5). The effect of the cosmic variance is roughly symmetric on both kT and N_S, making it almost negligible when averaging over a large sample. I also performed simulations for my worst case, i.e. $I = 0.6$ (see Sect. 6.2), finding qualitatively the same results. For the $F = 1.0$ case, the bias on the temperature is $\approx 8\%$ rather than $\approx 3\%$ and the bias on the normalization is negligible.

The cluster emission in the 10′–12′ ring

The source contribution in the 10′–12′ ring, which mainly depends on cluster redshift and emission measure, is difficult to estimate with accuracy. I employed a simulation to determine how an inaccurate estimate could affect the measurement of cluster temperature, kT, and normalization, N_S. Soft protons were neglected in this case, too; background components and their input values were the same as for the $F = 1.0$ case of the cosmic variance tests. Input parameters for the thermal model in R_{int} were the same as in that case, instead in R_{ext} were k$T^{ext} = 4$ keV, $Z^{ext} = 0.2\ Z_\odot$, $z^{ext} = 0.2$, and $N_S^{ext} = 2.5 \times 10^{-4}$. For this particular choice of the parameters, the source-to-background count rate ratio, I, is 1.13 (see Sect. 6.2). When fitting spectra in R_{ext}, all thermal parameters were fixed: namely, the temperature, the metallicity, and the redshift were fixed to the input values, while for N_S^{ext} I considered 4 cases. In the first case, I neglected the source contribution ($N_S^{ext} = 0$); in the other cases, the normalization was fixed to a value lower ($N_S^{ext} = 1.0 \times 10^{-4}$), equal ($N_S^{ext} = 2.5 \times 10^{-4}$), and higher ($N_S^{ext} = 4.0 \times 10^{-4}$) than the input value. Normalizations of all background components (namely N_{HALO}, N_{CXB}, and N_{NXB}) were free parameters. For each case, I computed the weighted average of N_{HALO}, N_{CXB}, and N_{NXB} over the 300 spectra in R_{ext} and compared them to the input values (see Fig. 7.2). Both N_{NXB} and N_S^{ext} are weakly correlated; instead, N_{HALO} and, in particular, N_{CXB} show a strong negative correlation with the input value for N_S^{ext}, which depends on their spectral shapes. Note that, if I correctly estimate N_S^{ext}, then N_{HALO}, N_{CXB}, and N_{NXB} converge to their input values.

For each input value of N_S^{ext} in R_{ext}, I fit spectra in R_{int} in the 0.7–10.0 keV band after the usual rescaling of background parameters (see Sect. 4.2.2), calculated the weighted averages of the source temperature, kT, and normalization, N_S, over the 500 simulations, and compared them to the input values (see Fig. 7.3). Values of kT and N_S measured in R_{int} show a positive correlation with the value of N_S^{ext} fixed in R_{ext}. This is indeed expected because of the broad similarity in the spectral shapes of thermal and CXB models. In R_{ext} an overestimate of N_S^{ext} implies an underestimate of N_{CXB} (see Fig. 7.2); N_{CXB} is then rescaled by the area ratio, hence underestimated in R_{int} too. This results in an overestimate of kT and N_S in R_{int}, as for the $F = 1.3$ case of the cosmic variance simulations. Typical uncertainties ($\approx 50\%$) on N_S^{ext} cause systematic 5% and 7% errors on kT and N_S (see Fig. 7.3). Note that, after the correction for the $\approx 3\%$ bias mentioned when analyzing the cosmic variance simulations, the effect on N_S and kT is symmetric; thus, when averaging on a large sample, the effect on the mean profile should be almost negligible. Note also that if I were to neglect the cluster emission in the 10′–12′ ring ($N_S^{ext} = 0$), I would cause a systematic underestimate of kT and N_S on the order of 7-10% (see Fig. 7.3).

In a real case I deal with a combination of fluctuations and cannot treat each one separately, so I employed a simulation to investigate how fluctuations with different origins combine with each other. I combined effects due to the cosmic variance and to an inaccurate estimate of the cluster emission in the 10′–12′ ring,

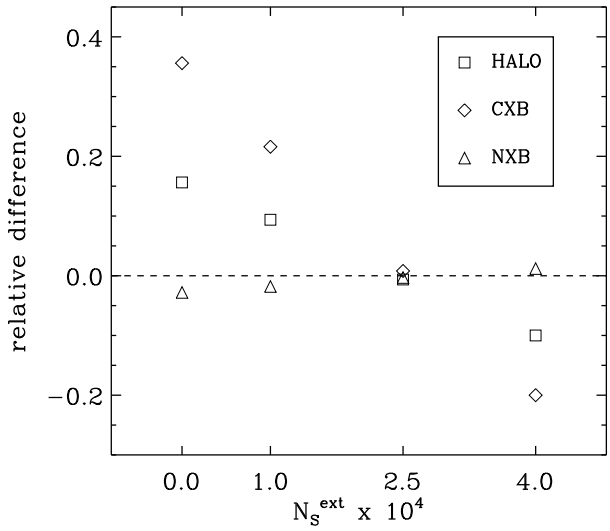

Figure 7.2: Relative differences between measured and input values for the normalization of background components (namely $N_{\rm HALO}$, $N_{\rm CXB}$, and $N_{\rm NXB}$) as a function of the input value for cluster normalization in $R_{\rm ext}$, $N_{\rm S}^{\rm ext}$. Uncertainties are smaller than the symbol size, and $N_{\rm CXB}$ shows the strongest (negative) correlation with $N_{\rm S}^{\rm ext}$.

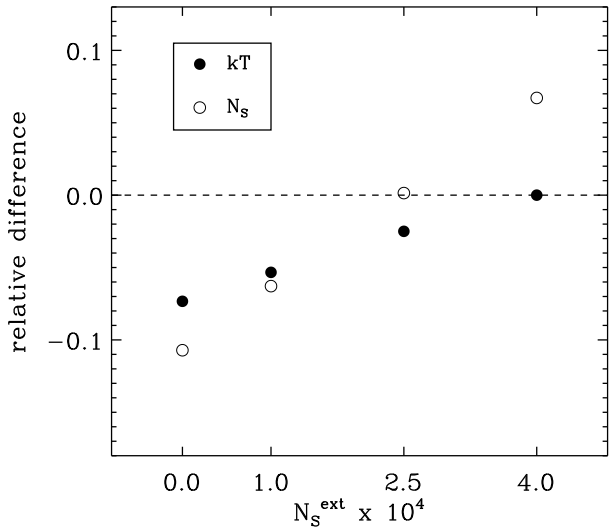

Figure 7.3: Relative differences between measured and input values for the source temperature, kT, and normalization, $N_{\rm S}$, as a function of the input value for cluster normalization in $R_{\rm ext}$, $N_{\rm S}^{\rm ext}$. Uncertainties are smaller than the symbol size. An underestimate/overestimate of $N_{\rm S}^{\rm ext}$ causes kT and $N_{\rm S}$ to be underestimated/overestimated.

by considering the $F = 0.7$, $F = 1.0$, and $F = 1.3$ cases mentioned when analyzing the cosmic variance simulations and $N_{\rm S}^{\rm ext} = 1.0 \times 10^{-4}$, $N_{\rm S}^{\rm ext} = 2.5 \times 10^{-4}$, and $N_{\rm S}^{\rm ext} = 4.0 \times 10^{-4}$ mentioned in this section. The simulation procedure was the same as described before. For the cluster normalization, I find that fluctuations combine in a linear way and that effects are highly symmetric with respect to the zero case ($F = 1.0$ for the cosmic variance and $N_{\rm S}^{\rm ext} = 2.5 \times 10^{-4}$ for the cluster emission in the $10'$–$12'$ ring). For the cluster temperature, I find again the $\approx 3\%$ bias related to the estimator; once accounted for this 3% offset, results are roughly similar to those found for the normalization case. To be more quantitative, when averaging on a large sample, the expected systematic on the temperature measurement is $\approx 3\%$ due to the biased estimator and $\lesssim 2\%$ due to deviations from the linear regime.

The QSP component

A careful characterization of the QSP component is crucial for our data analysis procedure. I employed a simulation to quantify how an incorrect estimate of the QSP contribution from the "IN over OUT" diagnostic; i.e., the $R_{\rm SB} = 1.10$ (see Sect. 4.1.2) could affect the measurements. The spectral components and their input values were the same as for the $F = 1.0$ case of the cosmic variance simulations, plus the QSP component in both rings. The model for QSP was the same as described in Sect. 3.2.2. I chose two input values for $N_{\rm QSP}$ corresponding to a standard ($R_{\rm SB} = 1.10$) and a high ($R_{\rm SB} = 1.40$) level of QSP contamination. For these particular choices of the parameters, the source-to-background count rate ratio, I, is 1.06 for $R_{\rm SB} = 1.10$ and 0.77 for $R_{\rm SB} = 1.40$ (see Sect. 6.2). For each input value I considered 2 cases: an underestimate ($R_{\rm SB} = 1.05 - 1.35$) and an overestimate ($R_{\rm SB} = 1.15 - 1.45$) of the correct value. By fitting spectra in $R_{\rm ext}$ in the 0.7–10.0 keV band, I determined the range of variability of $N_{\rm HALO}$, $N_{\rm CXB}$, and $N_{\rm NXB}$, and rescaled it in $R_{\rm int}$ (see Sect. 4.2.2). I then fit spectra in $R_{\rm int}$ and compared the weighted averages of cluster temperature, kT, and normalization, $N_{\rm S}$, to their input values (see Fig. 7.4).

When considering $N_{\rm S}$, the relative difference between measured and input values is less than 5% for all cases and the effect is symmetric, so the impact on the mean profile obtained from a large sample should be very weak. Instead, kT strongly depends on the estimate of the QSP component: the relative difference is $\approx 5\%$ for $R_{\rm SB} = 1.10$ and $\approx 20\%$ for $R_{\rm SB} = 1.40$. When overestimating $R_{\rm SB}$, kT is underestimated, because of the broad similarity in the spectral shapes of the two components. In the $R_{\rm SB} = 1.40$ case, the values corresponding to an overestimate and an underestimate, although symmetric with respect to zero, are characterized by different uncertainties (errors in the first case are twice those in the second); thus, a weighted average returns a 10% underestimated value.

7.1.2 "A posteriori" tests

In this subsection I investigate how the mean profile is affected by a particular choice of key parameters: the last ring for which I measure a temperature, the energy band used for the spectral fitting, and the QSP contamination.

The truncation radius

In Sect. 6.2 I introduced the indicator I to determine the last ring where the temperature measurement is reliable. I produced the mean temperature profiles by averaging over all measurements or which $I > I_0$, for different values of the threshold I_0. In Fig. 7.5 I report the profiles obtained in the 0.7–10.0 keV band for different choices of I_0 (namely 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0). As expected, the smaller the

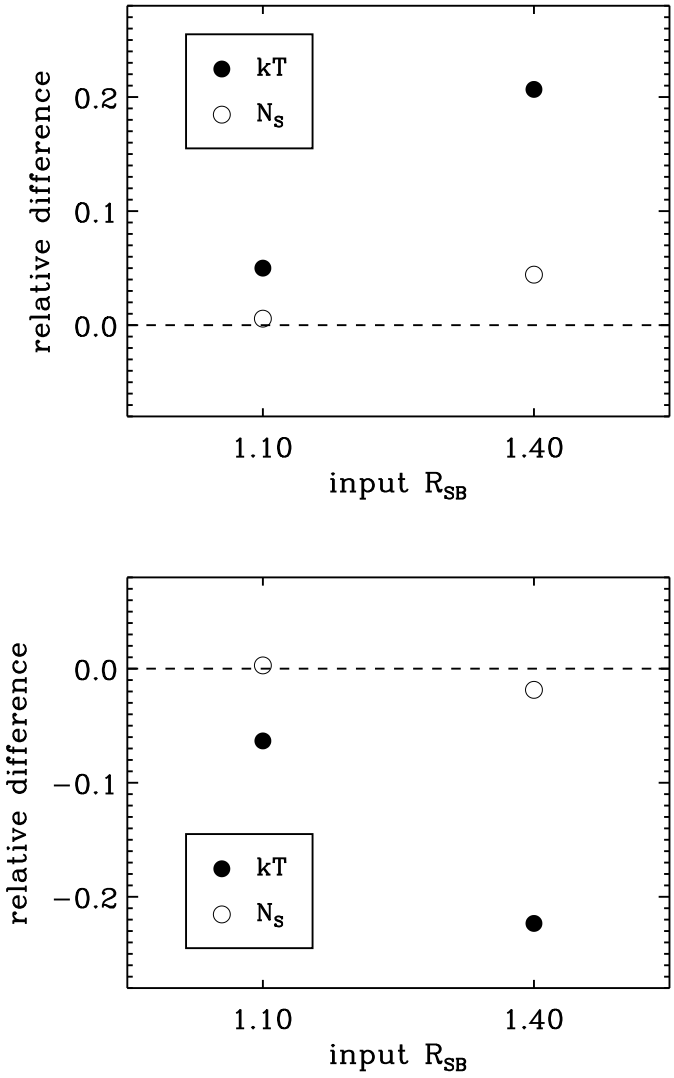

Figure 7.4: Relative differences between measured and input values for the source temperature, kT, and normalization, $N_{\rm S}$, as a function of the input value for the QSP contribution, $R_{\rm SB}$. Uncertainties are smaller than the circle size. Top panel: $R_{\rm SB}$ is underestimated to 1.05 and 1.35 with respect to 1.10 and 1.40. Bottom panel: $R_{\rm SB}$ is overestimated to 1.15 and 1.45 with respect to 1.10 and 1.40. See text for the discussion.

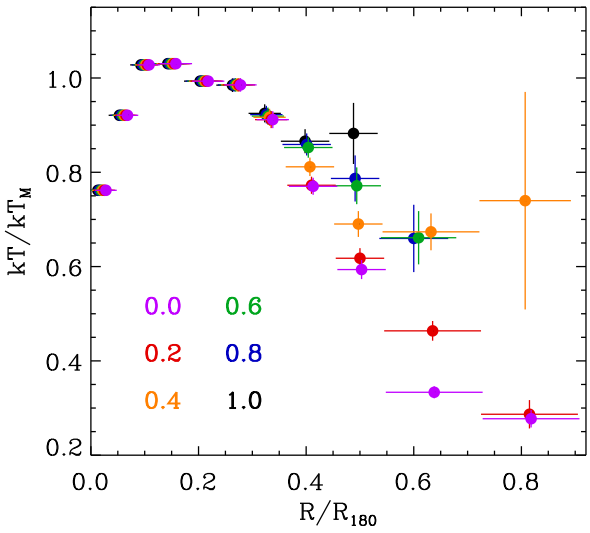

Figure 7.5: Mean temperature profiles computed by choosing different values for the threshold I_0 (defined in Sect. 6.2) plotted with different colors. There is a clear systematic effect: the smaller the threshold, the steeper the profile. The radii have been slightly offset in the plot for clarity.

threshold, the farther the mean profile extends. Let us focus on the points between 0.3 and 0.6 of R_{180}, a clear systematic effect is noticed: the smaller the threshold, the lower the temperature; meaning that, on average, the temperature is lower in those rings where the background is more important. Such a systematic effect becomes evident where cluster emission and background fluctuations are comparable and is probably related to small imperfections in my background modeling and to the bias on the temperature estimator (see Sect. 7.1.1). The imperfections of my background model becomes the dominant effect for low values of I (namely $I \lesssim 0.4$). Thus, under a certain threshold, I_0, the measurements are no longer reliable. In Fig. 7.5 I show that $I_0 = 0.6$ represents a good compromise. Indeed, when considering the region between 0.4 and 0.5 of R_{180} and comparing the average value for kT obtained for a threshold $I_0 = 0.6$ and for $I_0 = 1.0$, I find a small ($4\% \pm 3\%$) relative difference.

Fitting in different bands

I fit spectra in two different energy bands (i.e. 0.7–10.0 keV and 2.0–10.0 keV), each one characterized by different advantages and drawbacks (see Sect. 4.2). The indicator, I, defined in Sect. 6.2 depends on the band in which the count rate is calculated: more precisely, I(0.7–10.0) is roughly 1.5 times greater than I(2.0–10.0) for low values (i.e. $I \lesssim 2.0$). The threshold $I_0 = 0.6$ in the 0.7–10.0 keV band corresponds to $I_0 = 0.4$ in the 2.0–10.0 keV band. In Fig. 7.6 I compare the mean temperature profiles obtained in the 0.7–10.0 keV band ($I_0 = 0.6$) and in the 2.0–10.0 keV band ($I_0 = 0.4$). The profiles are very similar, except for the innermost point. The uncertainties in the 0.7–10.0 case are much smaller at all radii, even if the total number of points (i.e. the number of rings for all cluster) is the same, because the higher statistical quality at low energies allows a substantial reduction of the errors on single measurements.

In the most internal point, a high discrepancy between the two measurements

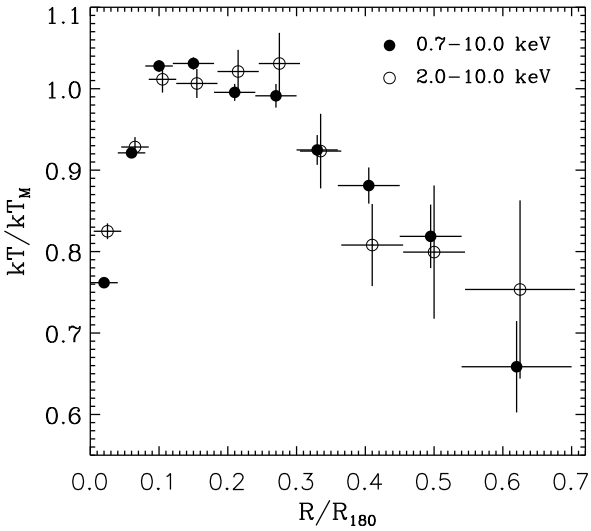

Figure 7.6: Mean temperature profiles obtained by fitting spectra in the 0.7–10.0 keV (filled circles) and in the 2.0–10.0 keV band (empty circles). The profiles are very similar, except for the innermost point. The radii have been slightly offset in the plot for clarity.

is present, although in that region the background is negligible. This is due to the superposition, along the line of sight, of photons emitted by optically thin ICM with different densities and temperatures (see Sect. 1.2). When looking at the center of cool core clusters, the line of sight intercepts regions characterized by strong temperature gradients, so the accumulated spectrum is the sum of many components at different temperatures. In this case, the best fit value for the temperature strongly depends on the energy band (i.e. the harder the band, the higher the temperature), because the exclusion of the soft band implies the exclusion of most of the emission from cooler components (Mazzotta et al. 2004).

Contamination from QSP

I divided clusters in my sample into four groups, according to the QSP contamination estimated from $R_{\rm SB}$ (see Sect. 4.1.2). In Fig. 7.7 I report the mean temperature profiles for the four groups, by fitting spectra in the 0.7–10.0 keV band and fixing $I_0 = 0.6$. When dividing clusters into subsamples, I chose larger bin sizes to reduce the error bars. When $R_{\rm SB}$ was high, my selection criterion based on the source-to-background count rate ratio (see Sects. 6.2) excluded the outer rings, indeed the red profile only extends out to 0.5 R_{180}. No correlation is found between the shape of the profiles and $R_{\rm SB}$; i.e., the four profiles are fully consistent. The discrepancy in the innermost ring is due to the presence of a different number of cool core clusters in each group. I therefore conclude that the systematic error associated with the QSP contamination is smaller than statistical errors ($\approx 7\%$ beyond 0.4 R_{180}).

7.1.3 A budget for systematics

In this subsection I summarize the main results for what concern systematic errors associated with the mean profile. I compare expected systematics computed from "a priori" tests with measured systematics from "a posteriori" tests.

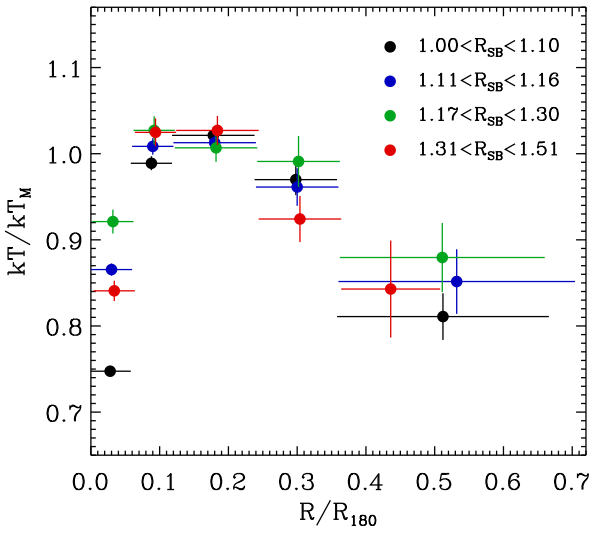

Figure 7.7: Mean temperature profiles as a function of the QSP contamination, $R_{\rm SB}$. The four profiles are fully consistent, and no correlation is found between the shape of the profiles and $R_{\rm SB}$. The radii have been slightly offset in the plot for clarity.

The $F = 1.0$ case of the cosmic variance simulations and the $N_{\rm S}^{\rm ext} = 2.5 \times 10^{-4}$ case of the cluster outer ring simulations show that our analysis procedure is affected by a 3% to 8% systematic underestimate of the temperature when analyzing the outermost rings. The bias is probably related to the temperature estimator as described in Chapter 5. In contrast, the normalization estimator is unbiased. In Sect. 7.1.1, I also found that i) the effects of the cosmic variance and of an inaccurate estimate of the cluster emission in the external ring are symmetric for both the temperature, kT, and the normalization, $N_{\rm S}$; ii) the effects due to fluctuations with different origins combine in a linear way and, when averaging over a large sample, the systematic associated with the mean profile is almost negligible for $N_{\rm S}$ and $\lesssim 2\%$ for kT. Thus, the expected systematic for kT is $\lesssim 5\%$.

Moreover, in Sect. 7.1.1 I found that, for a standard level of contamination ($R_{\rm SB} = 1.10$), a typical 5% error in the estimate of $R_{\rm SB}$ causes negligible effects on both measurements of cluster temperature and normalization. The same error causes negligible effects on $N_{\rm S}$ measurements for a high level of contamination ($R_{\rm SB} = 1.40$). In contrast, effects on kT for $R_{\rm SB} = 1.40$ are important: the same 5% error causes a 10% underestimate of kT, also when averaging over a large sample. However, at the end of Sect. 7.1.2, in particular from Fig. 7.7, I have concluded that, when considering the whole sample, the systematic error associated with the QSP contamination is smaller than statistical errors ($\approx 7\%$ beyond 0.4 R_{180}). The difference between expected and measured systematic errors is only apparent. Indeed, when analyzing my sample, I averaged measurements that span a wide range of values for $R_{\rm SB}$ and I; conversely, the 10% systematic error is expected for an unfavorable case, i.e. $R_{\rm SB} = 1.40$ and $I = 0.77$.

In Sect. 7.1.2 I compared the mean temperature value obtained for a threshold $I_0 = 0.6$ and for $I_0 = 1.0$ in an outer region (i.e. between 0.4 and 0.5 of R_{180}). In this ring the mean value for the indicator I is 1.14, thus the expected bias related to the temperature estimator is $\approx 3\%$ (see Sect. 7.1.1). Actually, I measured a $4\% \pm 3\%$ temperature discrepancy, which is consistent with the expected bias. As pointed

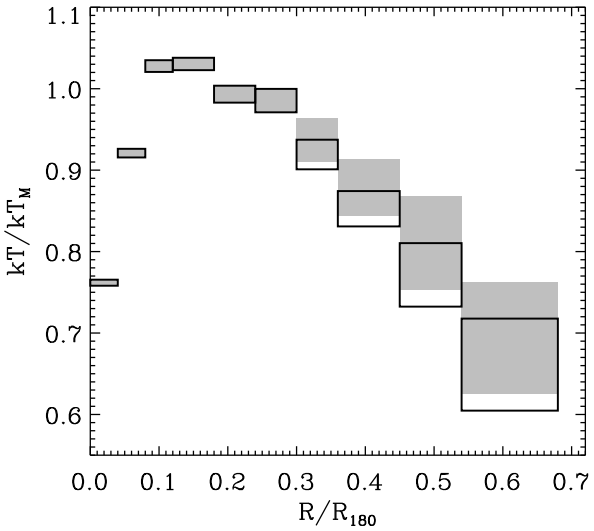

Figure 7.8: Mean temperature profile rescaled by R_{180} and k$T_{\rm M}$. For each ring, empty boxes and shaded regions indicate one-sigma uncertainties respectively before and after the bias correction.

out in Sect. 7.1.2, the discrepancy could also be due to small imperfections in my background model. I am not able to quantify the amount of this contribution, but it is expected to be small when considering $I > 0.6$.

To summarize, in external regions the measurements of the cluster temperature are affected by systematic effects, which depend on the radius through the factor I, i.e. the source-to-background count rate ratio. For each ring, I calculated the mean value for I, estimated the expected bias from simulations, and applied a correction to the mean profile. The expected bias is negligible for internal rings out to 0.30 R_{180} (for which $I > 3$), is 2-3% for 0.30–0.36 and 0.36–0.45 bins, and is $\approx 5\%$ for the last two bins (i.e. 0.45–0.54 and 0.54–0.70). I associated with the correction an uncertainty close to the correction itself, accounting for the limited knowledge from my "a posteriori" tests of the precise value of the bias. In Fig. 7.8 I show the mean temperature profile before and after the correction for the bias. The uncertainty is the quadrature sum of the statistical error and of the error associated with the correction. Hereafter, I will consider the mean profile corrected for the bias, unless otherwise stated. Note that the bias is always comparable to the statistical uncertainties. For this reason, mine can be considered as a definitive work, for what concerns the measurement of radial temperature profiles of galaxy clusters with *XMM-Newton*. I have reached the limits imposed by the instrument and by the analysis technique, so that increasing of the number of objects will not improve the quality of the measurement.

7.2 Evaluation of systematics on metallicity measures

I made use of montecarlo simulations to test the reliability of the metallicity measurements. The simulation procedure is similar to that described for the source-only case in Sect. 5.1. I considered a thermal (MEKAL in XSPEC) spectrum only, without a background. Input parameters were 6 keV temperature, 0.25 $Z_{\odot}$ metallicity,

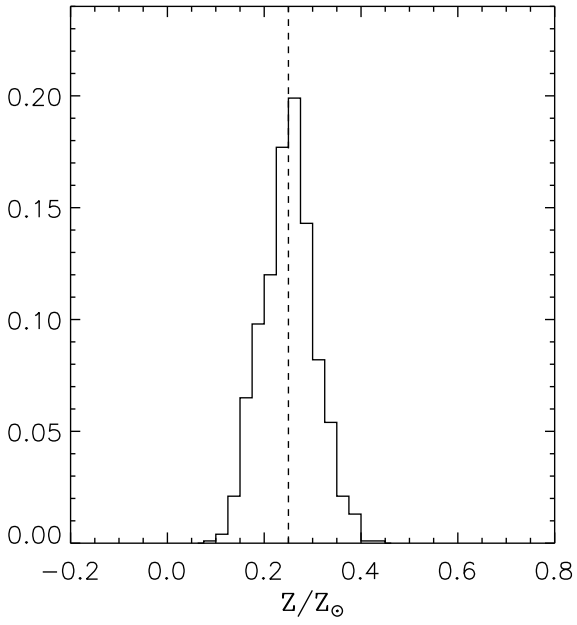

Figure 7.9: Frequency distribution for metallicity (expressed in solar units) best fit values for a 100 ks exposure time. The dashed line indicates the input value, i.e. 0.25 $Z_{\odot}$.

0.2 redshift, and 9×10^{-4} normalization. As usual, I used the Abell 1689 EPIC MOS1 observation as a guideline, for producing RMF and ARF, and for choosing typical input model parameters. I considered the $3.5'$–$4.5'$ ring only; i.e., the outer ring where I am usually able to measure a reliable metallicity. For each channel, I perturbed the number of counts with a Poisson distribution centered on the expected value, and repeated this procedure 1000 times to obtain 1000 spectra, which simulate 1000 independent measurements of the source. I fit simulated spectra with a MEKAL model in the 0.7–10.0 keV energy band using the Cash statistic. Temperature, metallicity, redshift, and normalization were allowed to vary within their XSPEC standard ranges. Here I focused on the metallicity, for which the standard allowed range is between 0 and 10^3 in solar units. For each measurement, I determined the best fit value and the one-sigma uncertainties.

In Fig. 7.9 I report the frequency distribution for metallicity best fit values for a 100 ks exposure time (i.e. for spectra with high statistical quality); as expected, the curve is very similar to a narrow Gaussian peaked around the input value (i.e. 0.25 $Z_{\odot}$). The mean, the median, and the weighted (over one-sigma uncertainties) average are all close to 0.25 $Z_{\odot}$, namely the weighted average is 0.246 ± 0.002 $Z_{\odot}$. In the left panel of Fig. 7.10 I report the same histogram for a 20 ks exposure time (i.e. for spectra with standard statistical quality). The curve is peaked on the input value, but the variance, which mainly depends on the number of counts around the energy of the emission line, is higher. Since the metallicity is forced to be positive, the curve is truncated and a dozen points pile up near the boundary (i.e. zero). In this case, the mean and the median are close to 0.25 $Z_{\odot}$, instead the weighted average is 0.193 ± 0.003 $Z_{\odot}$, i.e. $\approx 25\%$ lower than the input value. Part of this discrepancy is due to a boundary effect. For measurements characterized by a negative fluctuation of the metallicity, the minimum of the χ^2 curve lies outside

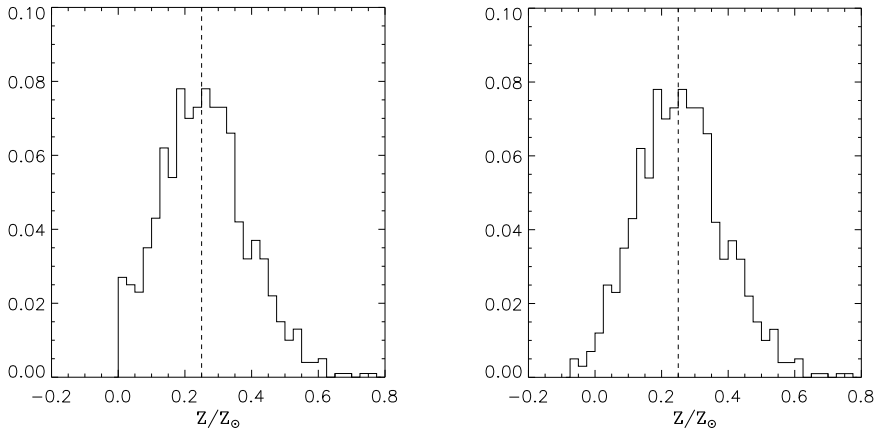

Figure 7.10: Frequency distribution for metallicity (expressed in solar units) best fit values for a 20 ks exposure time. The dashed lines indicate the input value, i.e. 0.25 $Z_{\odot}$. For the left panel, the metallicity is forced to be positive; for the right panel, it can assume negative values too. The two histograms only differ around zero.

the allowed range. The best fit values are then forced to be above zero and the uncertainties have unreliable small values; therefore, when computing a weighted average, these measurements have the highest statistical weight, and the net effect is an underestimate of the real metallicity. When fitting exactly the same spectra, allowing the metallicity to be negative, I obtain the histogram reported in the right panel of Fig. 7.10. The tails of the curve are more symmetric, all measurements have roughly the same uncertainty, as expected, and the weighted average is $0.229 \pm 0.004\ Z_{\odot}$, i.e. $\approx 8\%$ lower than the input value. This simple solution allowed us to correct for most of the underestimate; however, a small (i.e. $\approx 0.02\ Z_{\odot}$) systematic still affects the measurements, especially in the outer regions. I also performed more realistic simulations, by introducing a background, and obtain substantially similar results.

Although allowing observables to assume unphysical values is against common sense, there are measurement procedures that can yield unphysical values. Deciding to accept only physical values and reject others will clearly result in a bias. This is a general issue, which does not pertain to astrophysics alone. An interesting example I found concerns the analysis of data from the Collider Detector at Fermilab[2]. A statistical committee specifically appointed to provide guidelines for the analysis of the Collider Detector data recommends a treatment similar to the one I propose here, i.e. unphysical values can be used in statistical procedures.

The boundary effect I have just pointed out should play an important role when comparing subsamples characterized by a different statistical quality of the data. For example, in Sect. 8.2.2 I compare the metallicity obtained for near and distant clusters. In my sample, near cluster spectra usually have a better statistical quality for various reasons (e.g. longer observations, cosmological dimming effect). If the metallicity is allowed to vary between 0 and 5 $Z_{\odot}$, the mean metallicity beyond 0.20 R_{180} is $0.249 \pm 0.011\ Z_{\odot}$ and $0.188 \pm 0.014\ Z_{\odot}$, for near and distant clusters respectively. The measured discrepancy of $0.061 \pm 0.018\ Z_{\odot}$ has a significance of more than three-sigma. Conversely, if the metallicity is allowed to vary between $\pm 5\ Z_{\odot}$ (see Sect. 8.2.2), the discrepancy of $0.029 \pm 0.019\ Z_{\odot}$ is consistent with a

[2]http://www-cdf.fnal.gov/physics/statistics/statistics_faq.html#ssel1

purely statistical fluctuation.

I then warn X-ray astronomers about the existence of this kind of systematics, which could affect the measurement of the metallicity, especially for the case of low metallicity, statistically poor spectra.

Balestra et al. (2007) and Maughan et al. (2008) have adopted a procedure that is alternative to ours to estimate mean metal abundances for a sample of clusters. These authors performed a simultaneous spectral fit, leaving temperature and normalization free to vary for each object and using a unique metallicity value for all clusters in each redshift bin. It is worth noting that when modeling the background, as we do, a joint fit is infeasible, because of the large number of model parameters. Moreover, we are not aware of any detailed work that investigates the impact of systematic errors, possibly affecting individual measurements, on the final result of a joint fit.

Chapter 8

Results

In this chapter I will characterize the mean temperature and metallicity profiles, investigate their dependency from physical properties (e.g. the redshift), and compare them with hydrodynamic simulations and previous observational works.

I summarize here my main results: in Table 8.1 and Fig. 8.1 I report the mean temperature profile, and in Table 8.2 and Fig. 8.2 I report the mean metallicity profile, obtained by averaging individual cluster profiles as explained in Sect. 6.2.

8.1 The mean temperature profile

8.1.1 Characterizing the profile

I fit all individual temperature profiles reported in Fig. 6.3 beyond 0.2 R_{180} with a linear model and a power law to characterize the profile decline. By using a linear model

$$\frac{\mathrm{k}T}{\mathrm{k}T_{\mathrm{M}}} = \mathrm{A} - \mathrm{B}\left(\frac{R}{R_{180}} - 0.2\right), \tag{8.1}$$

I find $\mathrm{A} = 1.02 \pm 0.01$ and $\mathrm{B} = 0.77 \pm 0.11$. By using a power law

$$\frac{\mathrm{k}T}{\mathrm{k}T_{\mathrm{M}}} = \mathrm{N}\left(\frac{R}{0.2\ R_{180}}\right)^{-\mu}, \tag{8.2}$$

Ring[a]	Temperature[b]
0.00–0.04	0.762±0.004
0.04–0.08	0.921±0.005
0.08–0.12	1.028±0.007
0.12–0.18	1.030±0.008
0.18–0.24	0.993±0.010
0.24–0.30	0.985±0.014
0.30–0.36	0.938±0.026
0.36–0.45	0.878±0.035
0.45–0.54	0.810±0.058
0.54–0.70	0.694±0.069

Notes: [a] in units of R_{180}; [b] in units of $\mathrm{k}T_{\mathrm{M}}$.

Table 8.1: Mean temperature values rescaled by $\mathrm{k}T_{\mathrm{M}}$ and corrected for the biases discussed in Chapter 5, for each interval in units of R_{180}.

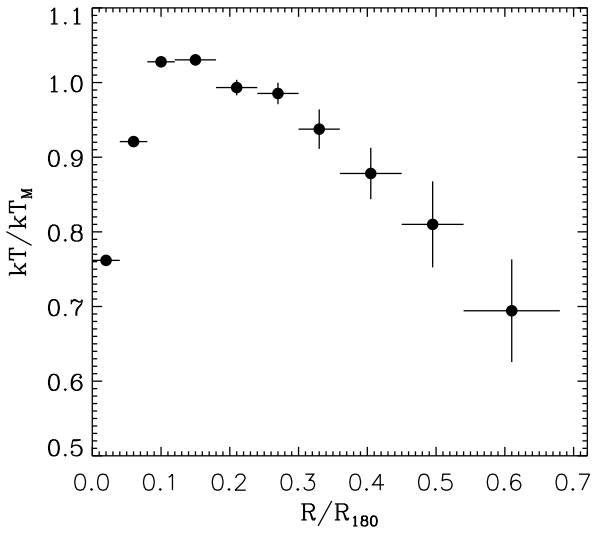

Figure 8.1: Mean temperature profile rescaled by R_{180} and k$T_{\rm M}$.

Ring[a]	Metallicity[b]
0.00–0.04	0.450±0.010
0.04–0.08	0.352±0.009
0.08–0.12	0.294±0.010
0.12–0.18	0.262±0.010
0.18–0.24	0.230±0.013
0.24–0.32	0.245±0.017
0.32–0.45	0.218±0.023

Notes: [a] in units of R_{180}; [b] in solar units (Anders & Grevesse 1989).

Table 8.2: Mean metallicity values in solar units (Anders & Grevesse 1989) for each interval in units of R_{180}.

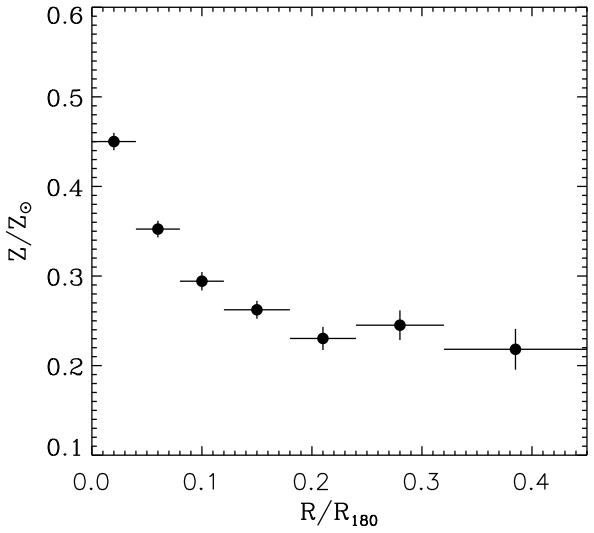

Figure 8.2: Mean metallicity profile averaged over all clusters. Abundances are expressed in Anders & Grevesse (1989) solar values and radii in units of R_{180}.

I find N $= 1.03 \pm 0.01$ and $\mu = 0.24 \pm 0.04$. If the gas can be approximated by a polytrope (see Sect. 1.3), one can derive its index, γ, from the slope of projected temperature profiles, μ (De Grandi & Molendi 2002):

$$\gamma = 1 + \mu/2, \tag{8.3}$$

under the assumption that, at large radii, three-dimensional gas temperature and density profiles be well described, respectively, by a power law and a β-model with $\beta = 2/3$. For $R > 0.2\ R_{180}$, I measured $\gamma = 1.12 \pm 0.02$, which is an intermediate value between those associated with isothermal ($\gamma = 1.0$) and adiabatic ($\gamma = 1.67$) gas. However, note that the power-law best-fit parameters depend on the chosen region (see Fig. 8.3), as well as the derived γ, thus the above values should be taken with some caution.

8.1.2 Redshift evolution of temperature profiles

I divided clusters in my sample into four groups according to the redshift, to investigate a possible evolution of temperature profiles with cosmic time. In Fig. 8.4 I report the mean temperature profiles for the four groups. Spectra were fitted in the 0.7–10.0 keV band and $I_0 = 0.6$ (see Sect. 6.2). As in the following Sects. 8.1.3 and 8.1.4, when dividing clusters into subsamples, the profiles were not corrected for biases (see Sect. 7.1.3), because when comparing subsamples one is not interested in determining the absolute value of the temperature, but in searching for relative differences. Moreover, in Figs. 8.4 and 8.6 I chose larger bin sizes to reduce the error bars (as in Fig. 7.4). The four profiles are very similar: the discrepancy in the outer regions is comparable to statistical and systematic errors, and the difference in the central region is due to a different fraction of cool core clusters. I fit each group of profiles with a power law beyond $0.2\ R_{180}$ and report the results in Fig. 8.5. Since there is no clear correlation between the two parameters and the redshift, the conclusion from the analysis of my sample is that there is no indication of profile evolution up to $z = 0.3$.

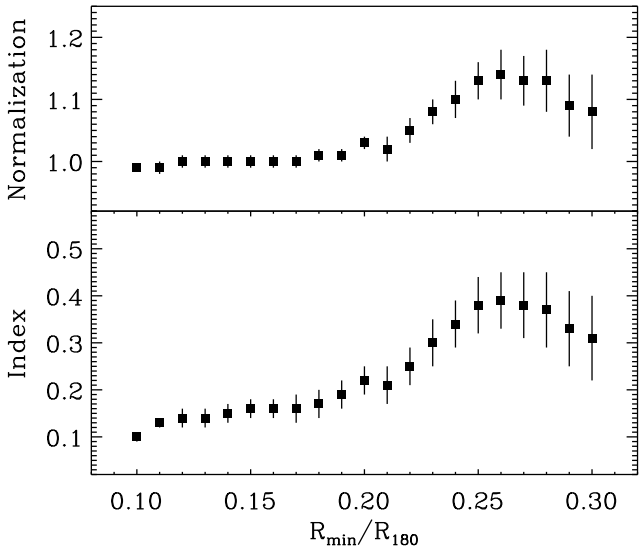

Figure 8.3: Power-law best-fit parameters obtained by fitting profiles beyond a variable radius, $R_{\rm min}$, in units of R_{180}. The normalization is calculated at 0.2. The index best-fit value is not constant with $R_{\rm min}$, thus the ICM cannot be considered as a polytrope.

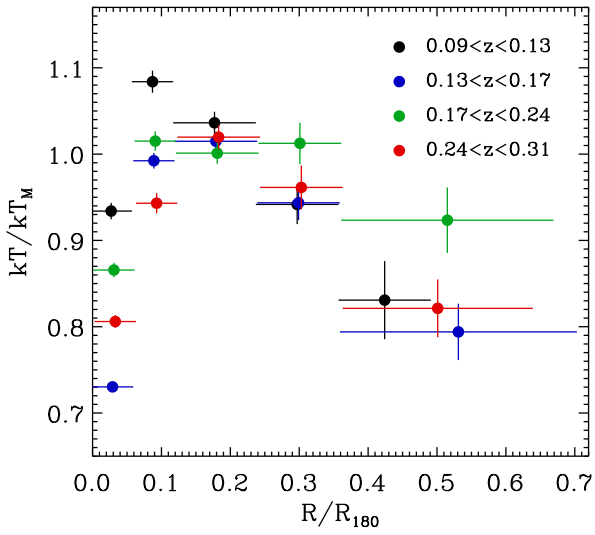

Figure 8.4: Mean temperature profiles for the four z-binned groups of clusters. There is no indication of profile evolution. The radii have been slightly offset in the plot for clarity.

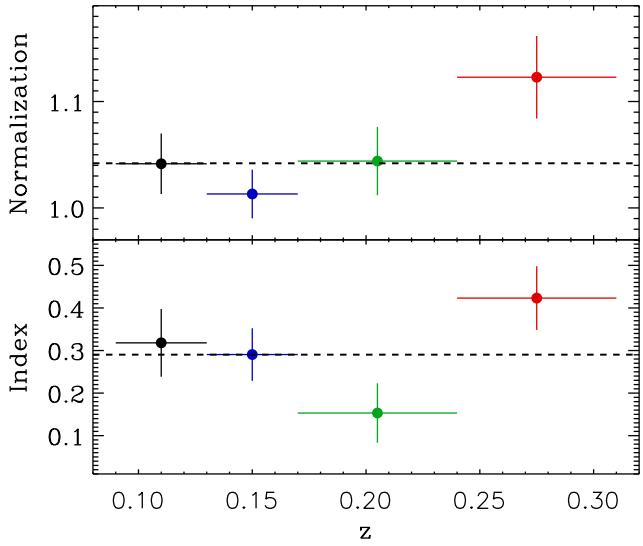

Figure 8.5: Best fit parameters obtained by fitting each group of profiles with a power law beyond 0.2 R_{180}. The normalization is calculated at 0.2 R_{180}. The dashed lines indicate the best fit values for the whole sample. No clear correlation is found between power-law parameters and the redshift.

8.1.3 Cool core and non-cool core clusters

In Sect. 6.2 I defined three groups: clusters that clearly host a cool core, clusters with no evidence of a cool core, and uncertain clusters. In Fig. 8.6 I show the mean temperature profiles for the three groups. Spectra were fitted in the 0.7–10.0 keV band and $I_0 = 0.6$. Profiles differ by definition in the core region and are consistent beyond $\approx 0.1\ R_{180}$.

8.1.4 REFL04 and LP07 subsamples

My sample is not complete with respect to any property. However, most of its clusters ($\approx 2/3$) belong to the REFLEX Cluster Survey catalog (Böhringer et al. 2004b), a statistically complete X-ray flux-limited sample of 447 galaxy clusters, and a dozen objects belong to the *XMM-Newton* Legacy Project sample (Pratt et al. 2007), which is representative of an X-ray flux-limited sample with $z < 0.2$ and k$T > 2$ keV. I then selected two subsamples from our sample: clusters that belong to the REFLEX catalog (REFL04 subsample) and to the Legacy Project sample (LP07 subsample). The smaller (i.e. the LP07) was derived from Pratt's parent sample, by applying our selection criteria based on cluster temperature and redshift (see Sect. 6.1). I also excluded cluster observations that are heavily affected by soft proton contamination; however, the latter selection should be equivalent to a random choice and introduce no bias. Thus, I expect the LP07 subsample to be representative of an X-ray flux-limited sample of galaxy clusters with $0.1 < z < 0.2$ and k$T > 3.3$ keV. The larger (i.e. the REFL04) subsample includes the LP07 one. Clusters that belong to the REFL04, but not to the LP07, were observed with *XMM-Newton* for different reasons. They are not part of a large program and almost all observations have different PIs. Thus, there are no obvious reasons to believe that the REFL04 subsample is significantly biased with respect to any fundamental cluster property. A similar reasoning leads to the same conclusion for

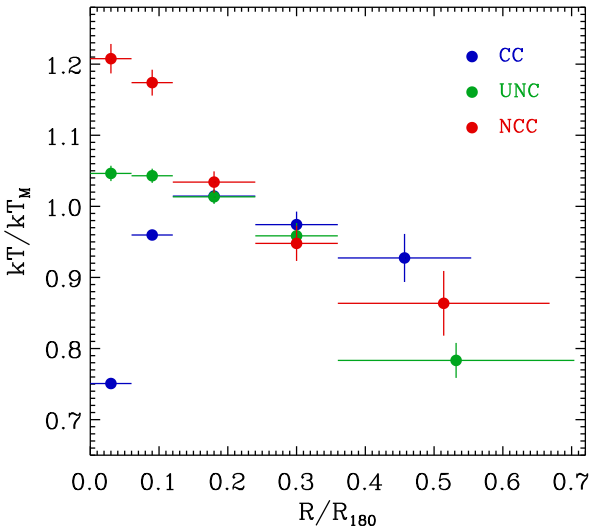

Figure 8.6: Mean temperature profiles for cool core (blue), non cool core (red), and uncertain (green) clusters. Profiles differ by definition in the core region and are consistent in the outer regions.

my whole sample.

In Fig. 8.7 I compare mean temperature profiles obtained from the two subsamples and the whole sample. The three profiles are fully consistent beyond $\approx 0.1\ R_{180}$, the difference in the central region is due to a different fraction of CC clusters. The conclusion is that my whole sample is representative of hot, intermediate redshift clusters with respect to temperature profiles, i.e. the quantity I am interested in.

8.1.5 Comparison with hydrodynamic simulations

In this subsection I compare the mean temperature profile obtained in this work with the one (hereafter B04) derived from cluster hydrodynamic simulations by Borgani et al. (2004). The authors used the TREE+SPH code GADGET (Springel et al. 2001) to simulate a concordance cold dark-matter cosmological model ($\Omega_{\rm m} = 0.3$, $\Omega_\Lambda = 0.7$, $\sigma_8 = 0.8$, and $h = 0.7$) within a box of 192 h^{-1} Mpc on a side, 480^3 dark-matter particles, and as many gas particles. The simulation included radiative cooling, star formation, and supernova feedback. Simulated cluster profiles were scaled by the emission-weighted global temperature and R_{180} calculated from its definition (i.e. the radius encompassing a spherical density contrast of 180 with respect to the critical density). In Fig. 8.8 I compare my observed profile to the projected mean profile obtained by averaging over simulated clusters with $kT > 3$ keV. The evident mismatch between the two profiles is most likely due to a different definition for the scaling temperature: actually, it is known that the emission weighted temperature is higher than the mean temperature obtained from observational data (Mazzotta et al. 2004). By rescaling the B04 profile by 10%, I find good agreement between simulation and our data beyond $\approx 0.25\ R_{180}$, confirming the expectation that the mean behavior of the ICM in the cluster outskirts is mainly due to the gravitational force. Conversely in the core region, where many not yet well understood physical processes also play a role, simulations are not able to reproduce the observed profile shape.

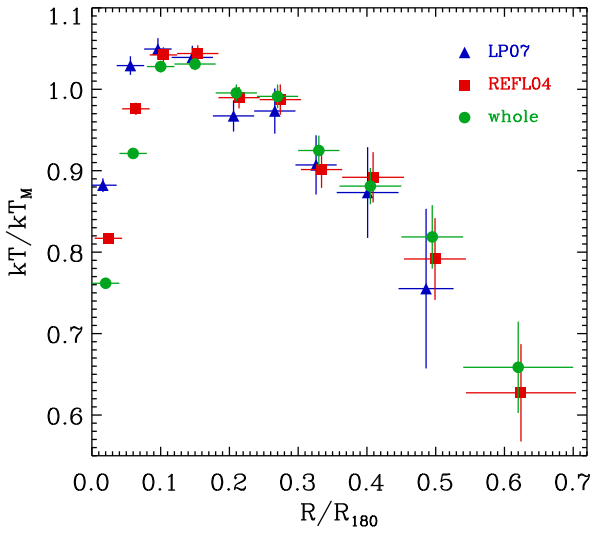

Figure 8.7: Mean temperature profiles obtained from the LP07 subsample (blue triangles), the REFL04 subsample (red squares), and the whole sample (green circles). The three profiles are fully consistent in the outer regions. The radii have been slightly offset in the plot for clarity.

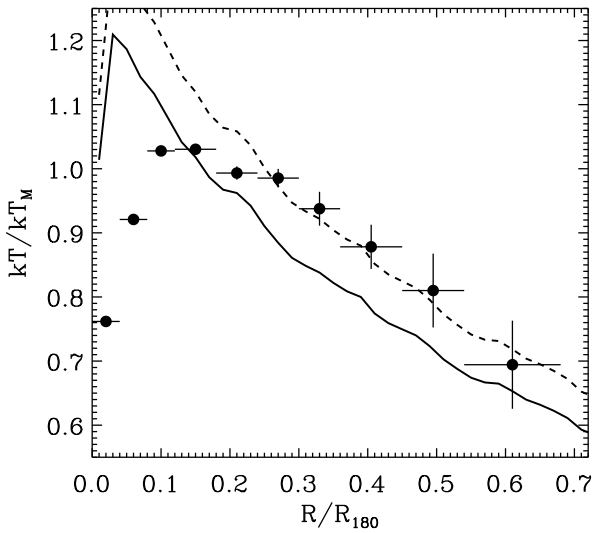

Figure 8.8: Comparison between the observed mean profile obtained in this work (circles) and the one derived from hydrodynamic simulations (Borgani et al. 2004) by averaging over clusters with k$T > 3$ keV (solid line). The dashed line is obtained by rescaling the solid one by 10%.

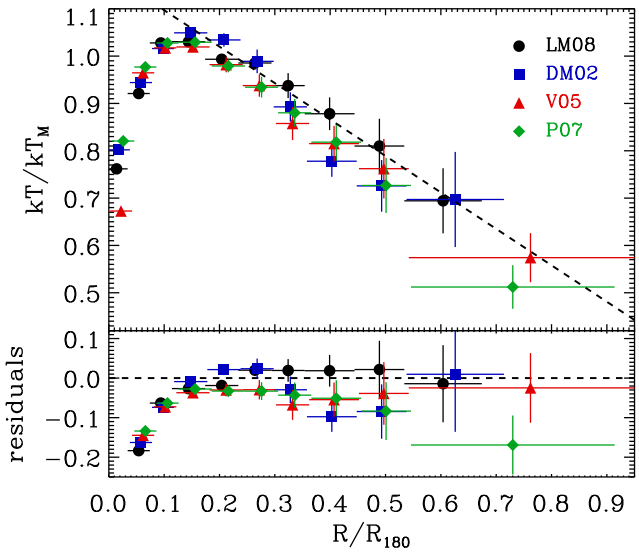

Figure 8.9: Top panel: mean temperature profiles obtained from this work (black circles, LM08), by De Grandi & Molendi (blue squares, DM02), by Vikhlinin et al. (red upward triangles, V05), and by Pratt et al. (green diamonds, P07). All profiles are rescaled by k$T_{\rm M}$ and R_{180} as defined in Sect. 6.2. The dashed line shows the best fit with a linear model beyond 0.2 R_{180} (see Sect. 8.1.1) and is drawn to guide the eye. Bottom panel: residuals with respect to the linear model. The LM08 profile is the flattest one.

8.1.6 Comparison with previous observations

In this section I compare my mean temperature profile (LM08) with those obtained by other authors, namely De Grandi & Molendi (2002), Vikhlinin et al. (2005), and Pratt et al. (2007). De Grandi & Molendi (DM02) have analyzed a sample of 21 hot (k$T > 3.3$ keV), nearby ($z \lesssim 0.1$) galaxy clusters observed with *BeppoSAX*. Their sample included both CC and NCC clusters. Vikhlinin et al. (V05) have analyzed a sample of 13 nearby ($z \lesssim 0.2$), relaxed galaxy clusters and groups observed with *Chandra*. I selected from their sample only the hottest (k$T > 3.3$ keV) 8 clusters, for a more appropriate comparison with my sample. Pratt et al. (P07) analyzed a sample of 15 hot (k$T > 2.8$ keV), nearby ($z \lesssim 0.2$) clusters observed with *XMM-Newton*. Clusters of their sample present a variety of X-ray morphology.

Comparing different works is not trivial. Cluster physical properties, instrumental characteristics, and data analysis procedures may differ. Moreover, each author uses his own recipe to calculate a mean temperature and to derive a scale radius. I rescaled temperature profiles obtained by other authors by using the standard cosmology (see Sect. 6.1) and calculating the mean temperature, k$T_{\rm M}$, and the scale radius, R_{180}, as explained in Sect. 6.2; the aim is to reduce all inhomogeneities as much as possible.

In Fig. 8.9 I compare the four mean temperature profiles, rescaled by k$T_{\rm M}$ and R_{180}. Due to the correction for the biases described in Sect. 7.1.3, our mean profile is somewhat flatter than others beyond 0.2 R_{180}. Discrepancies in the core region are due to a different fraction of CC clusters. The outermost point of the P07 profile is $\approx 20\%$ lower; however, it is only constrained by two measurements beyond ≈ 0.6 R_{180}. The indicator I (see Sect. 6.2) warns about the reliability of these two measurements, for which $I \approx 0.3$, i.e. a half of the threshold I chose, $I_0 = 0.6$. In

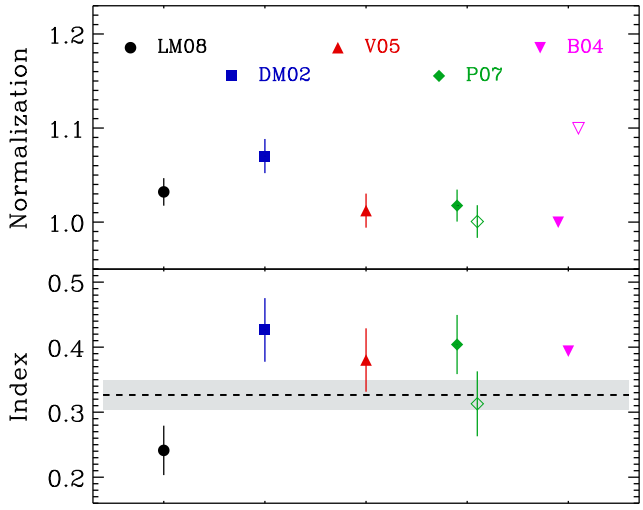

Figure 8.10: Best fit parameters, obtained by fitting observed and simulated cluster profiles with a power law, beyond 0.2 R_{180}. In the upper panel I report the normalization, in the lower the index. I use the same symbols as in Fig. 8.9 for observed clusters and a violet downward triangle for Borgani's work (B04). The normalization is calculated at 0.2 R_{180}. For P07 I report two values, empty diamonds indicate index and normalization obtained when excluding the two outermost measurements (see text for details). The empty downward triangle indicates the normalization of the B04 rescaled profile (see Sect. 8.1.5). In the lower panel, the dashed line and the shaded region represent the weighted average and its one-sigma confidence interval derived from the observed profiles only (for P07 we use the lower value, i.e. the empty diamond). As previously noted from Fig. 8.9, the LM08 profile is the flattest one, but all indices of observed profiles are consistent within two sigma. Conversely, the B04 profile seems to be significantly steeper, but in this case I am not able to provide an estimate of parameter uncertainty.

Fig. 7.5 I showed that, when using our analysis technique, lower values of I are associated with a bias on the temperature measurement. I assume that a somewhat similar systematic may affect the P07 analysis technique, too. When excluding these two measurements, the P07 mean profile only extends out to 0.6 R_{180} and is consistent with LM08 (see also Fig. 8.10). It is possible that measurements obtained with other experiments be also affected by a similar kind of systematics, which make the profiles steeper.

I fit observed and simulated cluster profiles with a power law beyond 0.2 R_{180} and in Fig. 8.10 report best fit parameters. The LM08 profile is the flattest one; however, all observed profile indices are consistent within 2-3 sigma. In Sect. 7.1.3 I have quantified the systematic underestimate on the temperature measurement associated with our procedure. Since it depends on the indicator I, which itself depends on the radius, I also expect a net effect on the profile index, μ; namely, I expect μ to be overestimated. For this reason, it is possible that the discrepancy between indices obtained from different works (reported in Fig. 8.10) may not have a purely statistical origin. I calculated an average profile index, $\mu = 0.31 \pm 0.02$, which is significantly lower than obtained from the B04 profile, $\mu = 0.39$; however, for the simulation case I am not able to provide an estimate of parameter uncertainty.

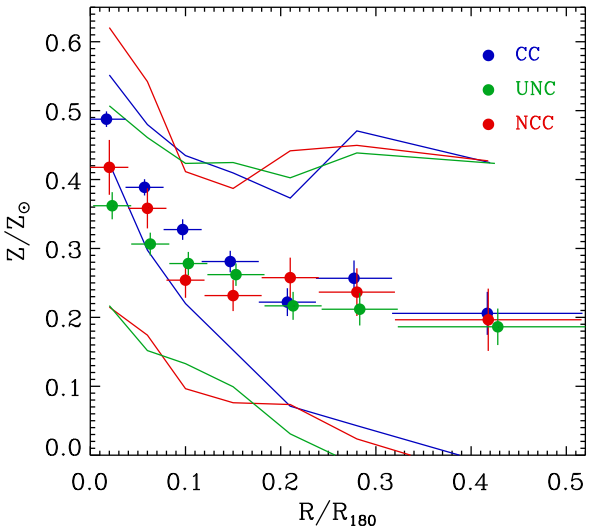

Figure 8.11: Mean metallicity profiles for cool core (blue), non-cool core (red), and uncertain (green) clusters. Lines indicate the scatter for each class. All profiles show a central peak, which is more pronounced for cool core clusters.

8.2 Metallicity profiles

The mean metallicity profile shows a central peak (see Fig. 8.2), due to the presence of cool core clusters, where the mean metallicity is 0.45 $Z_{\odot}$. The profile decreases out to $\approx$ 0.2 R_{180}; beyond this radius the profile is consistent with being flat ($Z \approx 0.2$).

8.2.1 Cool core and non-cool core clusters

In Sect. 6.2 I defined three groups: clusters that clearly host a cool core, clusters with no evidence of a cool core, and uncertain clusters. In Fig. 8.11 I show the mean temperature profiles for the three groups. Profiles are consistent beyond $\approx 0.2\ R_{180}$. In the center, as expected, cool core clusters show a strong central peak, but unexpectedly also profiles for the other two classes show a mild enhancement. This original result will be discussed in more detail in Chapter 10.

8.2.2 Redshift evolution of metallicity profiles

I divided clusters in our sample into two groups to investigate a possible evolution of the metal abundance with redshift: near (distant) clusters are characterized by a redshift lower (greater) than 0.2. In Fig. 8.12 I show the mean metallicity profiles for the two groups. The single points are consistent within one- or two-sigma, except for the core region where the spatial resolution plays an important role. In my procedure I fixed the size of the central ring to 30″, which corresponds to $\approx 0.03\ R_{180}$ for nearest ($z \approx 0.1$) and to $\approx 0.07\ R_{180}$ for most distant ($z \approx 0.3$) objects; thus, the metallicity peak for distant clusters is blurred over a larger region. Most of the three-sigma discrepancy in the region within 0.2 R_{180} (see shaded regions in Fig. 8.12) is most likely due to a different fraction of cool core and non-cool core clusters within the two subsamples (see Sect. 1.4.3).

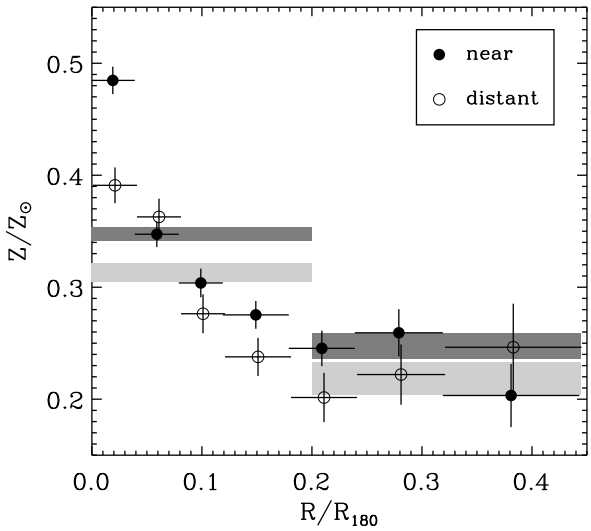

Figure 8.12: Mean metallicity profiles for near ($z < 0.2$, filled circles) and distant ($z > 0.2$, empty circles) clusters. The dark and the light shaded regions indicate the average metallicity within one-sigma uncertainties below and beyond 0.20 R_{180} for near and distant clusters, respectively. The radii have been slightly offset in the plot for clarity.

In the outer regions, if the metallicity is allowed to vary between 0 and 5 $Z_\odot$ as commonly done, I measure a discrepancy between two profiles of 0.061 ± 0.018 $Z_\odot$ with a more than three-sigma significance (see Sect. 7.2). Conversely, if the metallicity is allowed to vary between ± 5 $Z_\odot$ as in our data analysis procedure (see Sect. 4.2), near and distant clusters have a mean metallicity of 0.248 ± 0.011 $Z_\odot$ and 0.219 ± 0.015 $Z_\odot$, respectively. For this case, the discrepancy (i.e. ≈ 1.5 sigma, see shaded regions in Fig. 8.12) is consistent with a purely statistical fluctuation, and is the same order of magnitude as systematics I have estimated. I stress that the systematic effect described in Sect. 7.2, which affects in particular low metallicity, statistically poor spectra, if unaccounted for, could cause a false detection of the metallicity evolution.

Two recent works (Balestra et al. 2007; Maughan et al. 2008) have investigated the evolution in the iron content of the ICM, by analyzing data from the *Chandra* archive. As mentioned in Sect. 7.2, these authors performed a simultaneous spectral fit, leaving temperature and normalization free to vary for each object and using a unique metallicity value for all clusters in each redshift bin. Unfortunately, I cannot compare my results with those obtained by Balestra et al. (2007), because they considered different regions from cluster to cluster. Instead, Maughan et al. (2008) analyzed the region within R_{500} (≈ 0.6 R_{180}) with and without the core (i.e. 0.15 R_{500}) and obtained respectively $Z \approx 0.4$ $Z_\odot$ and $Z \approx 0.35$ $Z_\odot$ between $z = 0.1$ and $z = 0.3$. When analyzing roughly the same regions for my clusters, I obtain $Z = 0.32$ $Z_\odot$ and $Z = 0.26$ $Z_\odot$ respectively, with negligible uncertainties. The relative difference is, thus, consistent with the one found by Maughan et al. (2008), even if the absolute values for the metal abundance are significantly lower ($\approx 20\%$). A possible explanation for the discrepancy between absolute values may be related to different weights on the averaging procedure: I averaged weighting over the inverse squared errors, while Maughan et al. (2008) measured metal abundances from individual spectra extracted from the entire region of interest. Summarizing, when

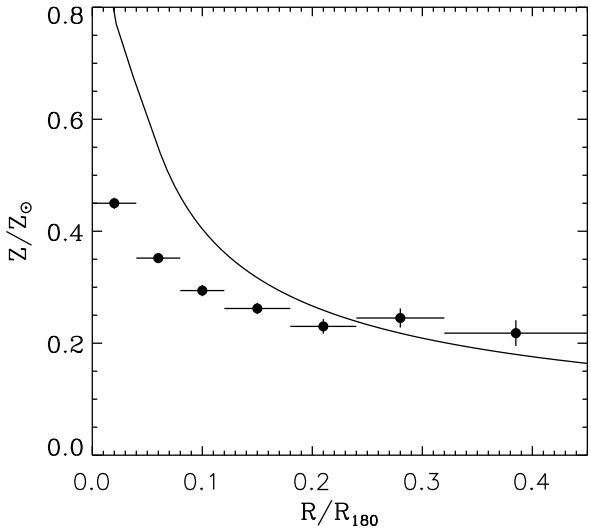

Figure 8.13: Comparison between our observed mean profile (circles) and the one derived from hydrodynamic simulations (solid line) by Fabjan et al. (2008). Abundances are expressed in Anders & Grevesse (1989) solar values and radii in units of R_{180}.

considering the same regions (i.e. beyond 0.15 R_{500}), I measure a difference in metal abundance, between near and distant sample, consistent with the result obtained by Maughan et al. (2008). When considering regions where the metallicity profile is roughly flat (i.e. beyond $\approx 0.2\ R_{180}$), because of the poor statistical quality of the measurements, I find a result that is consistent with that found beyond 0.15 R_{500}, but also with the absence of evolution.

8.2.3 Comparison with hydrodynamic simulations

I compare our (hereafter LM08) mean metallicity profile with the one (hereafter F08) derived from hydrodynamic simulations of four relaxed clusters by Fabjan et al. (2008).

The simulations were performed using the hydrodynamical TREE-SPH code GADGET-2 (Springel 2005) with the implementation of chemical enrichment by Tornatore et al. (2007). The authors used the emission-weighted definition of metallicity, with emissivity of each gas particle computed in the 0.5–10.0 keV energy band. In principle, for a comparison with observational data, one should extract synthetic spectra from the simulated clusters and then measure the metallicity by fitting these spectra with a single-temperature and single-metallicity plasma model (e.g. the MEKAL model in XSPEC). A recent work presented by Rasia et al. (2008) showed that, at least for Iron, the emission-weighted estimator of the metallicity gives results quite close (within about 10%) to those obtained from the spectral-fitting analysis.

In Fig. 8.13 I compare LM08 and F08 profiles. Note the differences at both small and large radii. In the central region, the F08 profile is much more peaked; however, this difference likely results from two factors: namely that Fabjan et al. (2008) analyzed only relaxed clusters and that LM08 results are limited by the *XMM-Newton* PSF. At large radii, while the F08 profile shows a constant decrease, LM08 is consistent with being flat beyond 0.2 R_{180}. This discrepancy could be due to issues

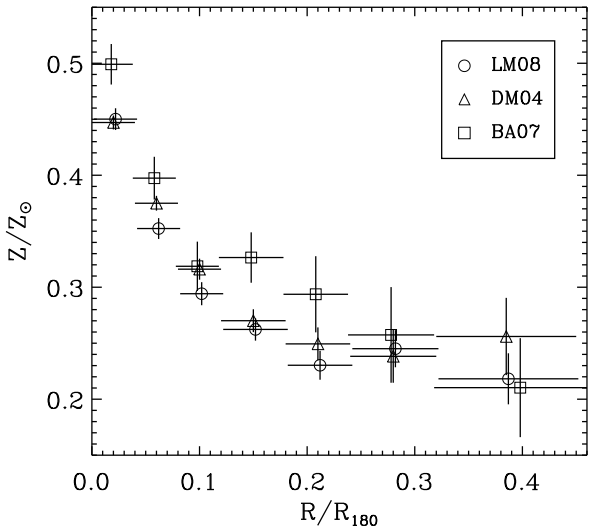

Figure 8.14: Mean metallicity profiles obtained from this work (LM08, circles), by De Grandi et al. (2004) (DM04, triangle), and Baldi et al. (2007) (BA07, squares). Abundances are expressed in Anders & Grevesse (1989) solar values and radii in units of R_{180}. The radii have been slightly offset in the plot for clarity.

related to observations, simulations, or both. More precisely, observations could be affected by unknown systematic effects, and simulations could underestimate the metallicity at large radii by underestimating possible motions in the ICM, which might be responsible of mixing metals.

8.2.4 Comparison with previous works

I compare my results with those obtained by De Grandi et al. (2004) and Baldi et al. (2007). De Grandi et al. (hereafter DM04) have analyzed a sample of 21 hot (k$T \gtrsim 3.5$ keV), nearby ($z \lesssim 0.1$) galaxy clusters observed with *BeppoSAX*. Baldi et al. (hereafter BA07) have analyzed 12 very hot (k$T \gtrsim 6$ keV), intermediate redshift ($0.1 \lesssim z \lesssim 0.3$) clusters observed with *Chandra*.

As mentioned in Sect. 8.1.6, comparing results obtained from different works is not trivial; I have thus rescaled DM04 and BA07 profiles by using the standard cosmology (see Sect. 6.1) and deriving the scale radius, R_{180}, as explained in Sect. 6.2, in order to reduce all inhomogeneities as much as possible. I also converted DM04 abundances from Grevesse & Sauval (1998) to Anders & Grevesse (1989) solar values.

In Fig. 8.14 I compare LM08, DM04, and BA07 mean metallicity profiles. Even if the three samples cover a different redshift range and the instruments (i.e. *XMM-Newton*, *BeppoSAX*, and *Chandra*) present different characteristics, the mean profiles are remarkably similar over the entire radial range.

Chapter 9

Conclusions and future perspectives

9.1 Main results

I have analyzed a sample of $\approx$ 50 hot, intermediate redshift galaxy clusters observed with *XMM-Newton* to measure their radial properties. The sample should be representative of hot, intermediate redshift clusters, at least with respect to the temperature and metal abundance profiles.

The results concerning temperature measurements are summarized as follows.

- The mean temperature profile declines with radius in the 0.2–0.6 R_{180} range.
- When excluding the core region, the profiles are characterized by an intrinsic dispersion (6%) comparable to the estimated systematics.
- There is no evidence of profile evolution with redshift from $z = 0.1$ to $z = 0.3$.
- The profile slope in the outer regions is independent of the presence of a cool core.
- The slope of the mean profile is broadly similar to those obtained from hydrodynamic simulations. I find a discrepancy of $\approx 10\%$ in normalization probably due to a different definition for the scaling temperature, while the slopes are roughly consistent.
- When compared to previous works, my profile is somewhat flatter, probably due to a different level of characterization of systematic effects, which become very important in the outer regions and tend to steepen the profile.

The results concerning metallicity measurements are summarized as follows.

- The mean metallicity is 0.45 $Z_\odot$ in the center and decreases out to $\approx 0.2\ R_{180}$; beyond 0.2 R_{180} the metallicity is consistent with being flat at $0.23 \pm 0.01\ Z_\odot$.
- The profiles show a large scatter, which is mostly of statistical origin. In the central regions the scatter (i.e. $22 \pm 2\%$) is also related to the presence of cool core clusters, in the outer regions it (i.e. $14 \pm 8\%$) is comparable to systematics.
- There is no clear evidence of profile evolution from $z = 0.1$ to $z = 0.3$ in the outer regions (i.e. beyond $\approx 0.2\ R_{180}$);

- When considering the same regions analyzed by Maughan et al. (2008) using *Chandra*, I obtain consistent results as far as the relative difference between near and distant clusters is concerned, but $\approx$ 20% lower absolute values, possibly due to the different method used to calculate the mean metallicity.

- I point out the existence of a systematic effect, affecting in particular distant clusters, which, if unaccounted for, could cause a false detection of evolution.

- When comparing my mean profile to the one derived from hydrodynamic simulations by Fabjan et al. (2008), I find differences at small and large radii. In particular, while the profile obtained by Fabjan et al. (2008) shows a constant decrease, mine is consistent with being flat beyond 0.2 R_{180}.

- When comparing my mean profile to those obtained by recent works with *BeppoSAX* (De Grandi et al. 2004) and *Chandra* (Baldi et al. 2007), I find remarkable agreement over the entire radial range.

The results reported above were obtained using a novel data analysis technique, which includes two major improvements. First, the use of the background modeling, rather than the background subtraction, and the Cash statistic rather than the χ^2. This method requires a careful characterization of all background components, but allows us to properly deal with each individual component. Second, the detailed assessment of systematic effects with Montecarlo simulations allows us to give an estimate of the systematics affecting the mean temperature profile, and propose a correction to the profile itself. I performed two groups of tests to quantify the impact of different components on simulated spectra, and investigate how the measured temperature profile changes when choosing different key parameters.

My work not only provides a confirmation of previous results. For the first time, we believe we know where the systematics come from and how large they are. Indeed, my work allows us to not only constrain cluster temperature profiles with confidence in the outer regions, but also, from a more general point of view, to explore the limits of *XMM-Newton* instrumentation and find new ways to exploit at best the large amount of data in the *XMM-Newton* archive.

9.2 Future perspectives

9.2.1 The shock front in Abell 754

My work will allow us to look forward to ambitious new measurements: an example is the attempt to measure the putative shock in Abell 754, for which we have obtained a $\approx$ 200 ks observation with *XMM-Newton* in AO7. Detection and characterization of shock fronts in clusters are limited to a tiny number of objects. One of the major difficulties is that shocks occur in low surface-brightness regions, where available instrumentation has considerable difficulty in making reliable measurements. The novel observational and data analysis strategy described in my thesis should allow us to make a reliable measurement of the shock front in A754. The sensitivity afforded by our observation will also allow us to characterize parts of the outer regions of A754 in unprecedented detail. From a more general perspective ours may be viewed as an attempt of demonstrating the capability of the EPIC experiment to provide substantially improved measurements of cluster outer regions, which are likely to remain unsurpassed for years to come. The observation was performed on the last May and we are currently analyzing data.

9.2.2 Deriving other physical properties

The radial temperature profiles obtained in the framework of this thesis will be used to derive other physical properties of great interest (e.g. mass and entropy). Before, however, the profiles must be deprojected to derive the 3D cluster properties. A number of deprojection techniques have been developed (e.g. Pizzolato et al. 2003), even if they usually require strong geometrical assumptions (e.g. spherical symmetry) which are not always fulfilled in galaxy clusters. The deprojection starts from the most external ring for which a measurement is available and continues inward recursively subtracting the contribution of external rings to the inner (the process is called "onion peeling").

Once obtained 3D density (from the surface-brightness) and temperature profiles of the ICM, one can derive the total mass profile from (1.32), by assuming the hydrostatic equilibrium (e.g. Voigt & Fabian 2006). Theoretical models for the formation of structures predict cluster properties as a function of mass and redshift; therefore, a reliable measure of the mass for a large sample of clusters gives the possibility of using clusters as cosmological probes. Moreover, by measuring the total mass, it is possible to investigate the scaling relations (e.g. Ettori et al. 2002) between physical quantities that are directly measurable (e.g. X-ray luminosity) and those requiring complicated procedures of measurement (e.g. temperature and mass). Scaling relations can be used to better exploit the large amount of data from present and future X-ray surveys by converting the cluster luminosity function into the mass function and studying its evolution with redshift.

From deprojected temperature and density profiles, it is common to derive also the entropy, $S = T \times n^{-2/3}$, in analogy with the thermodynamic entropy. The characterization of the entropy distribution represents one of the most straightforward way to investigate non-gravitational processes. Previous works have found that, during the cluster formation and accretion, the ICM arranges itself in order to have a minimum of entropy in the center. However, cooler (and less massive) clusters show an entropy excess both in the core, where non-gravitational processes are more important, and sometimes in the outer regions, as a signature of possible pre-heating processes (e.g. Pratt et al. 2006).

Chapter 10

Preliminary results on additional work in progress

The issue of how to classify astronomical objects is widely discussed in the literature. The goal is to find few observable properties, that provide valuable information on the physics of the system at hand. As mentioned in Sect. 1.1.1, clusters of galaxies in optical can be roughly represented as a one-dimensional sequence, from regular to irregular systems, where regular clusters should be dynamically more relaxed. In X-rays, clusters are usually classified according to the presence of merger evidence or on the basis of the core properties. Clues of on-going or recent mergers are provided by the X-ray emission morphology (i.e. the presence of sub-clusters, irregular features in surface-brightness and temperature distribution, etc.) and by the presence of strong non-thermal emission (mainly detectable in the radio band). As far as the core properties are concerned, clusters are roughly divided according to the presence or the absence of a surface-brightness peak and a temperature drop in the center, usually referred as cool core (see Sect. 1.4.3). The two classifications are highly correlated in the sense that clusters showing evidence of recent interactions usually do not present a cool core, and vice-versa.

I focused on the central regions of clusters and, there, investigated thermodynamic and chemical properties of the intra-cluster medium. In Sect. 6.1 I showed that, when dividing clusters of my sample on the basis of the central temperature drop, about one-third does not clearly belong to cool core (those for which the temperature is significantly lower than the mean temperature) nor to non-cool core clusters (those for which the temperature profile does not significantly decrease), suggesting the existence of a population with intermediate characteristics. In Sect. 8.2.1 I showed that cool core clusters present strong metallicity peaks in their center, but also that the mean metallicity profile for non-cool core clusters shows a mild positive gradient towards the center. A key thermodynamic observable in describing clusters is the entropy (e.g. Ponman et al. 2003; Voit 2005; Pratt et al. 2006), which is commonly defined as $T_{3D} \times n_{3D}^{-2/3}$, where T_{3D} and n_{3D} are the deprojected temperature and density. In the literature (e.g. Rossetti et al. 2007), it is usual to define a pseudo-entropy from projected quantities as:

$$s = T \times EM^{-1/3}; \tag{10.1}$$

entropy should be the thermodynamic observable which better correlates with the metallicity. Given all these points, in this chapter I will show the preliminary results obtained when subdividing clusters of an enlarged sample, on the basis of their central pseudo-entropy gradient and correlating thermodynamic and chemical properties.

Name	z[a]	k$T_{\rm OUT}$[b]	Exp. time[c]	$R_{\rm SB}$[d]	Filter
Abell 4038	0.0300	3.0	78.1	1.45	MEDIUM
Abell 2199	0.0301	4.1	38.4	1.14	THIN1
2A 0335+096	0.0349	3.6	230.8	1.02	THIN1
Abell 2052	0.0355	2.9	85.1	0.99	THIN1
Abell 576	0.0390	3.8	42.7	1.54	MEDIUM
Abell 3571	0.0391	6.3	43.5	1.54	MEDIUM
Abell 119	0.0442	6.0	54.1	1.62	THIN1
MKW03	0.0450	3.3	99.2	1.08	THIN1
Abell 3376	0.0456	3.9	58.9	1.03	MEDIUM
Abell 1644	0.0470	4.2	42.0	1.40	THIN1
Abell 4059	0.0475	4.0	64.9	1.11	THIN1
Abell 3558	0.0480	5.2	126.4	1.00	THICK
Abell 3562	0.0480	4.3	116.6	1.85	THIN1
Triangulum Austr.	0.0510	9.2	27.3	1.15	MEDIUM
Hydra A	0.0538	3.4	52.2	1.59	THIN1
Abell 754	0.0542	8.7	40.6	1.12	MEDIUM
Abell 85	0.0551	5.5	34.6	1.04	MEDIUM
Abell 2319	0.0557	9.2	44.6	1.80	MEDIUM
Abell 3158	0.0597	4.9	54.0	1.58	THIN1
Abell 1795	0.0625	5.4	97.3	1.29	THIN1
Abell 399	0.0720	6.0	27.9	1.84	THIN1
Abell 401	0.0740	7.3	34.7	1.76	MEDIUM
Abell 3112	0.0750	4.3	64.6	1.23	MEDIUM
Abell 2029	0.0773	6.2	30.8	1.17	THIN1
Abell 2255	0.0806	6.2	25.1	1.37	THIN1
Abell 1650	0.0838	5.4	75.0	1.29	MEDIUM
Abell 2597	0.0852	3.5	144.3	1.07	THIN1

Notes: [a] redshift taken from the NASA Extragalactic Database; [b] reference temperature in keV derived from our analysis; [c] total good exposure time in ks; [d] intensity of residual soft protons.

Table 10.1: Physical properties and observation details for local clusters.

10.1 The sample and the analysis procedure

I enlarged my sample to about 60 objects from $z = 0.02$ to $z = 0.25$, by adding a number of local clusters (see Rossetti et al., in prep.). All clusters were observed with *XMM-Newton* , have a high temperature ($kT > 3$ keV) and a high galactic latitude ($|b| > 20°$). In Tables 10.1 and 10.2 I list the observations of local and distant clusters respectively and report cluster physical properties (e.g. redshift and temperature) and observational technical characteristics (e.g. total exposure time and filter). The redshift value (from optical measurements) is taken from the NASA Extragalactic Database[1]; k$T_{\rm OUT}$ is derived from our analysis. Each observation is performed using THIN1 or MEDIUM filters. I excluded from the sample observations that are highly affected by soft proton flares, so that the total (i.e. MOS1+MOS2+pn) exposure time for all observations is at least 20 ks. I also excluded observations of extremely disturbed clusters for which it was impossible to define a center; for what concerns double clusters, I analyzed the most luminous only.

Observation data files (ODF) were retrieved from the *XMM-Newton* archive and processed in a standard way with the Science Analysis System (SAS) v7.0. The event files were prepared as described in Sect. 4.1.1.

[1] http://nedwww.ipac.caltech.edu

Name	z[a]	k$T_{\rm OUT}$[b]	Exp. time[c]	$R_{\rm SB}$[d]	Filter
RXCJ0049.4–2931	0.1080	3.3	44.8	1.28	THIN1
Abell 2034	0.1130	7.0	27.9	1.16	THIN1
RXCJ1516.5-0056	0.1150	3.8	83.6	1.08	THIN1
RXCJ2149.1–3041	0.1179	3.3	71.2	1.11	THIN1
RXCJ1516.3+0005	0.1183	5.3	77.3	1.13	THIN1
RXCJ1141.4–1216	0.1195	3.8	82.0	1.03	THIN1
RXCJ1044.5–0704	0.1323	3.9	72.4	1.03	THIN1
Abell 1068	0.1375	4.5	56.3	1.09	MEDIUM
RXCJ2218.6–3853	0.1379	6.4	54.6	1.11	THIN1
RXCJ0605.8–3518	0.1410	4.9	58.3	1.07	THIN1
RXCJ0020.7–2542	0.1424	5.7	41.9	1.02	THIN1
Abell 1413	0.1427	6.7	71.7	1.10	THIN1
RXCJ2048.1–1750	0.1470	5.6	71.1	1.07	THIN1
RXCJ0547.6–3152	0.1483	6.7	67.1	1.12	THIN1
Abell 2204	0.1522	8.5	51.2	1.06	MEDIUM
RXCJ0958.3–1103	0.1527	6.1	22.9	1.16	THIN1
RXCJ2234.5–3744	0.1529	8.6	42.8	1.31	THIN1
RXCJ2014.8–2430	0.1612	7.1	64.8	1.05	THIN1
RXCJ0645.4–5413	0.1670	7.1	59.2	1.11	THIN1
Abell 1914	0.1712	8.7	62.9	1.17	THIN1
Abell 2218	0.1756	6.5	117.0	1.17	THIN1
Abell 1689	0.1832	9.2	106.7	1.14	THIN1
Abell 383	0.1871	4.4	82.3	1.33	MEDIUM
Abell 115	0.1971	5.1	103.2	1.20	MEDIUM
Abell 2163	0.2030	15.5	29.2	1.07	THIN1
Abell 963	0.2060	6.5	69.4	1.19	MEDIUM
Abell 209	0.2060	6.6	49.3	1.19	MEDIUM
Abell 773	0.2170	7.5	45.6	1.16	MEDIUM
Abell 1763	0.2230	7.2	36.3	1.08	MEDIUM
Abell 2390	0.2280	11.2	29.4	1.11	THIN1
Abell 2667	0.2300	7.7	59.9	1.48	MEDIUM
RX J2129.6+0005	0.2350	5.5	102.0	1.21	MEDIUM

Notes: [a] redshift taken from the NASA Extragalactic Database; [b] reference temperature in keV derived from our analysis; [c] total good exposure time in ks; [d] intensity of residual soft protons.

Table 10.2: Physical properties and observation details for distant clusters.

To investigate cluster properties in the central regions, I accumulated spectra in a circle of radius 0.05 R_{180} centered on the X-ray emission peak, and there measured the normalization, $N_{\rm IN}$, the temperature, $T_{\rm IN}$, and the metallicity, $Z_{\rm IN}$, of the core. Since I was interested in characterizing a variation of such quantities, I also measured the normalization, $N_{\rm OUT}$, and the temperature, $T_{\rm OUT}$, in the 0.05–0.20 R_{180} ring (the outer radius is limited by the apparent size of nearest clusters). For each EPIC instrument and each region, I accumulated a spectrum and generated an effective area (ARF), and for each observation generated redistribution functions (RMF) for MOS1, MOS2, and pn.

Spectra accumulated in the central regions of clusters have almost always high statistical quality; therefore, the complicated procedures developed for dealing with the background in the outer regions (see Chapter 3) are not strictly necessary and also EPIC pn data have been used. For all three detectors (namely MOS1, MOS2, and pn), channels were assembled in order to have at least 25 counts for each group, as commonly done when using the χ^2 statistic. I merged nine blank-field observations to accumulate background spectra, as commonly done. For each cluster observation, I calculated the count rate ratio, Q, between source and background observations beyond 9 keV in an external ring ($10'$–$12'$) of the field of view. I scaled background spectra by Q and, for each region, subtracted it from the cluster observation. This rough rescaling accounted for temporal instrumental background variations.

The spectral fitting was performed in the 0.5–10.0 keV energy band, by using the χ^2 statistic, with an absorbed thermal model (WABS*MEKAL in XSPEC). I fit spectra leaving the temperature and the normalization free to vary. The metallicity was constrained between $\pm 5\ Z_\odot$ (see Sect. 7.2). The redshift was constrained between $\pm 5\%$ of the optical measurement. The equivalent hydrogen column density along the line of sight, $N_{\rm H}$, was fixed to the 21 cm measurement (Dickey & Lockman 1990). Finally, for each quantity I computed the average over the three (MOS1, MOS2, pn) values and derived the projected emission measure, EM, as the ratio between the normalization and the area of the region in square arcmin.

10.2 Defining interesting quantities

To characterize the gradient of temperature and emission measure, I should compare the central to a global value for such quantities; however, for local clusters it was only possible to perform reliable measurements out to a small ($\approx 20\%$) fraction of R_{180}. As a temperature reference, I used $T_{\rm OUT}$ (see Sect. 10.1), which can be gathered to be a good proxy of the global temperature from Fig. 6.3. For what concerns the emission measure, I was faced with two alternatives, namely using $EM_{\rm OUT}$ (see Sect. 10.1) or the self-similar scaling (Arnaud et al. 2005).

A reference value for the emission measure should be measured at $\approx 0.4\ R_{180}$, where profiles show a remarkable degree of similarity (see Fig. 10.1). Here I considered the subsample of distant ($z > 0.1$) clusters, for which it is possible to measure EM out to 0.4 R_{180}. I defined as EM_0 the emission measure calculated in the 0.2–0.4 R_{180} ring, and as

$$EM_{\rm SSS} = \Delta_{\rm z}^{3/2}\,(1+z)^{9/2}\,\left(\frac{{\rm k}T_{\rm OUT}}{10\,{\rm keV}}\right)^{1/2}, \qquad (10.2)$$

the self-similar scaled emission measure. For each distant cluster, I calculated our "ideal" ratio $EM_{\rm IN}/EM_0$, determined directly from the data, the self-similar scaled ratio $EM_{\rm IN}/EM_{\rm SSS}$, and the standard ratio $EM_{\rm IN}/EM_{\rm OUT}$. In Fig. 10.2 I compare the self-similar scaled (left panel) and the standard (right panel) ratios to the ideal ratio. For both cases I find a good correlation, but the scatter is smaller (16% vs.

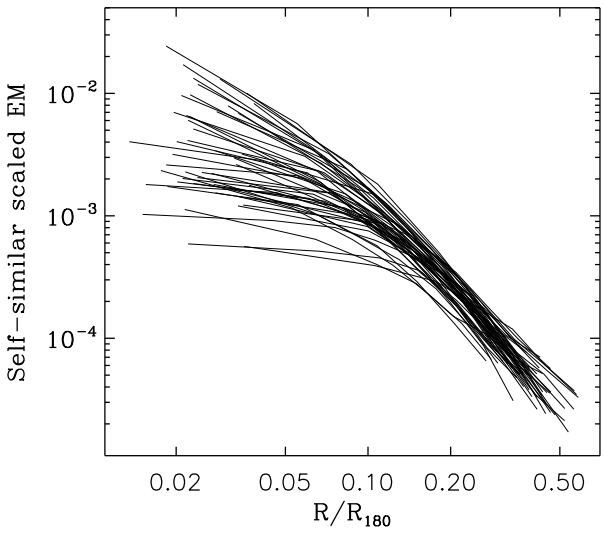

Figure 10.1: Self-similar scaled, projected emission measure profiles for the intermediate-redshift (i.e. $z > 0.1$) clusters. The radius is in units of R_{180} and the emission measure is scaled by $EM_{\rm SSS}$ (see Eq. 10.2).

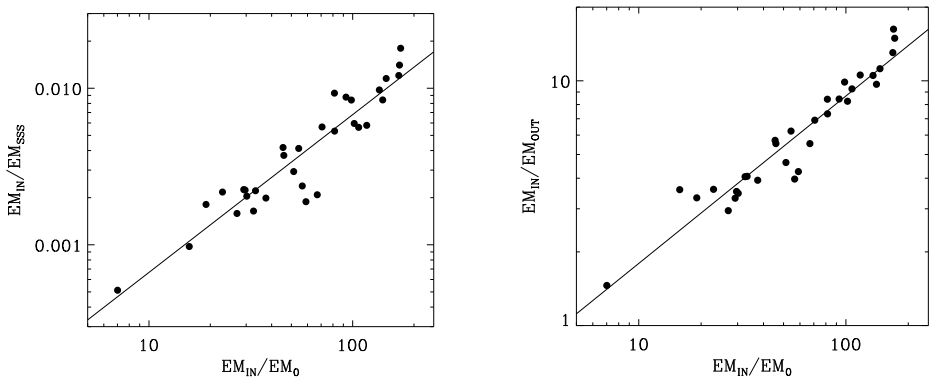

Figure 10.2: Correlations between various definitions for the *EM* ratios (see text for details). Left panel: self-similar scaled vs. ideal ratios. Right panel: standard vs. ideal ratios. The solid curve is the best fit with a power law. The uncertainties are smaller than the point size. The scatter is 29% and 16% for left and right panels respectively.

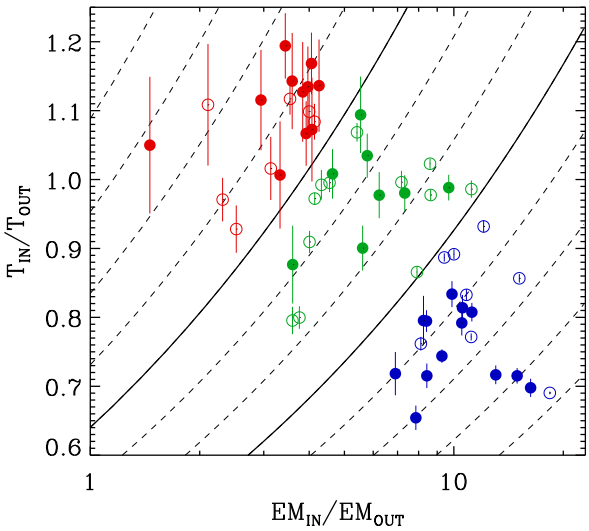

Figure 10.3: Comparison of temperature and emission measure ratios for all clusters. The dashed and the solid curves indicate the regions of the plot where the pseudo-entropy ratio, σ, is constant; clusters with the strongest pseudo-entropy variations populate the bottom-right corner. The solid curves indicate the thresholds used to divide clusters in high- (red circles), medium- (green circles), and low- (blue circles) entropy core systems. For each class, open and filled circles indicate local (i.e. z < 0.1) and distant (i.e. $z > 0.1$) clusters respectively.

29%) when using the standard ratio. For clarity, in the following plots, I will make use of $EM_{\rm IN}/EM_{\rm OUT}$, but I point out that all the results are largely independent of this particular choice.

In Fig. 10.3 I compare the temperature ratio, $T_{\rm IN}/T_{\rm OUT}$, with the emission measure ratio, $EM_{\rm IN}/EM_{\rm OUT}$. As expected, there is a clear, but quite scattered, correlation: more precisely, the stronger the emission measure peak, the stronger the temperature drop. From Eq. 10.1 I use emission measure and temperature ratios to define a pseudo-entropy ratio,

$$\sigma = \frac{T_{\rm IN}}{T_{\rm OUT}} \times \left(\frac{EM_{\rm IN}}{EM_{\rm OUT}} \right)^{-1/3} . \tag{10.3}$$

The pseudo-entropy ratio has been found to be well correlated with the entropy ratio (Rossetti et al. 2009, in preparation). In Fig. 10.3 the solid and the dashed curves indicate the regions where σ is constant; clusters with the strongest variations (i.e. lower ratios) of pseudo-entropy fill the bottom-right corner and are usually known as cool core clusters. From Fig. 10.3 there is not a clear bimodality in entropy, clusters form a continuous population of objects extending from low-entropy cores (LEC) through intermediate objects (MEC) to high-entropy cores (HEC). I arbitrarily chose two reasonable thresholds to divide clusters into three classes, so that each one contains roughly the same number of objects. In Table 10.3 I report the membership for each cluster. In Fig. 10.4 I show the histogram of the frequency distribution for the entropy ratio, σ. The distribution shows a weak indication of bimodality, but is also consistent with being flat.

The appearance of the plot in Fig. 10.3 and the classification based on the

LEC[a]	MEC[b]	HEC[c]
2A 0335+096	Abell 1644	Abell 119
Abell 1795	Abell 1650	Abell 2255
Abell 2052	Abell 2029	Abell 3158
Abell 2199	Abell 2319	Abell 399
Abell 2597	Abell 3558	Abell 401
Abell 3112	Abell 3562	Triangulum Austr.
Abell 85	Abell 3571	RXCJ0547.6–3152
Hydra A	Abell 4038	RXCJ0020.7–2542
Abell 383	Abell 4059	RXCJ2048.1–1750
RX J2129.6+0005	Abell 576	RXCJ2234.5–3744
Abell 2390	RXCJ0049.4–2931	RXCJ1516.3+0005
Abell 2204	RXCJ0645.4–5413	RXCJ1516.5–0056
Abell 1068	MKW03	Abell 773
Abell 2667	Abell 209	Abell 1763
RXCJ0605.8–3518	Abell 1413	Abell 2163
RXCJ1044.5–0704	Abell 963	Abell 2218
RXCJ1141.4–1216	Abell 1689	Abell 2034
RXCJ2014.8–2430	Abell 1914	Abell 3376
RXCJ2149.1–3041	Abell 754	
RXCJ0958.3–1103		
Abell 115		

Notes: [a] low entropy core clusters; [b] medium entropy core clusters; [c] high entropy core clusters.

Table 10.3: Low, medium, and high entropy core clusters.

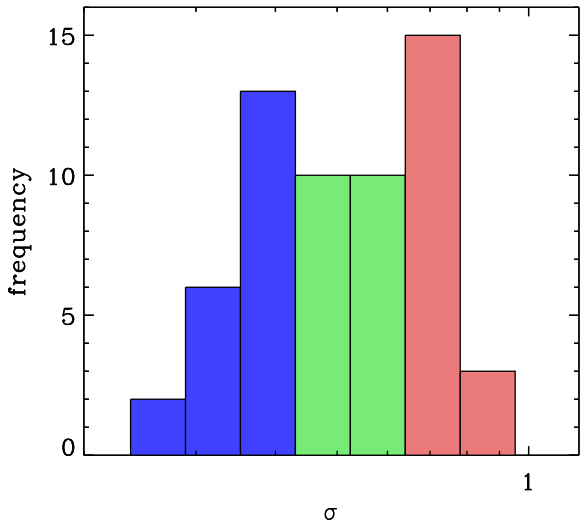

Figure 10.4: Histogram of the frequency distribution for the entropy ratio, σ. The color code is the same as used in Fig. 10.3.

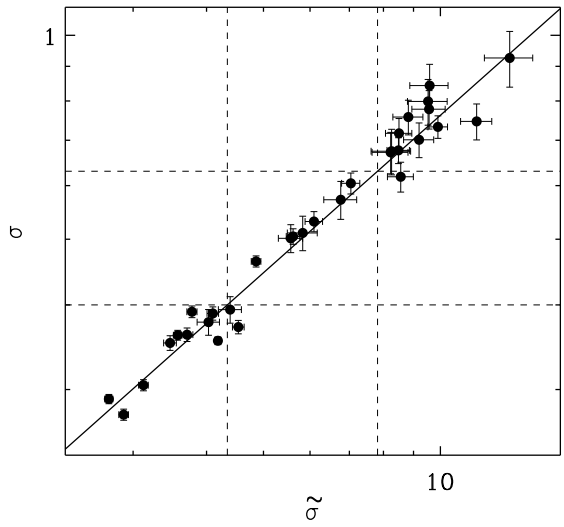

Figure 10.5: Correlation between σ and $\tilde{\sigma}$ (see text for definitions) for clusters with $z > 0.1$. The solid line shows the best-fit power law; the scatter is $\approx 6\%$. The dashed lines indicate the thresholds used to divide clusters into LEC, MEC, and HEC systems.

entropy ratio is very similar, also when using

$$\tilde{\sigma} = \frac{T_{\rm IN}}{T_{\rm OUT}} \times \left(\frac{EM_{\rm IN}}{EM_{\rm SSS}} \right)^{-1/3}, \tag{10.4}$$

which differs from σ for the use of the self-similar scaled EM. In Fig. 10.5 I show the correlation between σ and $\tilde{\sigma}$ for the subsample of distant clusters; I find a $\approx 6\%$ scatter around the best-fit power law (represented with a solid line). The dashed lines indicate the thresholds used to define the three classes (i.e. LEC, MEC, and HEC clusters). Only few border-line objects change their class from MEC to LEC or HEC and vice-versa, when switching from σ to $\tilde{\sigma}$; this fact is a confirmation of the robustness of our classification.

The central cooling time is another quantity largely used in the literature (e.g. Peres et al. 1998) to estimate the relaxation degree of clusters. Similarly as done for the entropy, I defined a pseudo-cooling-time, $t_{\rm cool} = T^{1/2} \times EM^{-1/2}$, and a pseudo-cooling-time ratio, $\tau = (T_{\rm IN}/T_{\rm OUT})^{1/2} \times (EM_{\rm IN}/EM_{\rm OUT})^{-1/2}$. When separating clusters according to τ, I find essentially the same results as when using σ.

10.3 Chemical properties

10.3.1 Chemical vs. thermodynamic quantities

In Fig. 10.6 I plot the central metallicity, $Z_{\rm IN}$, vs. the pseudo-entropy ratio, σ, for all clusters. I find σ and $Z_{\rm IN}$ to have a negative correlation, namely: the stronger the pseudo-entropy gradient, the stronger the metallicity peak.

I have further investigated the relationship between metallicity and pseudo-entropy shown in Fig. 10.6, by computing the main properties for the metal abundance distributions within 0.05 R_{180} for the three entropy classes. The computation has been performed both with and without few outliers, which have been identified

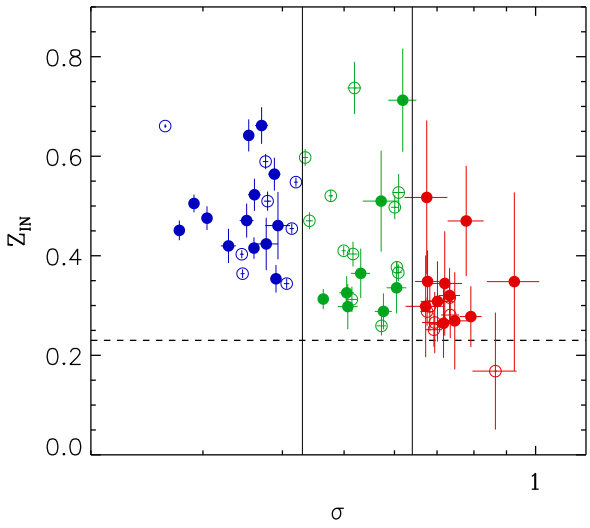

Figure 10.6: Central metallicity peak, $Z_{\rm IN}$, vs. the pseudo-entropy ratio, σ. Symbol and color codes are the same as in Fig. 10.3. Abundances are expressed in Anders & Grevesse (1989) solar values. The dashed line at 0.23 $Z_{\odot}$ indicates the metallicity in the outer regions of clusters (see Table 8.2).

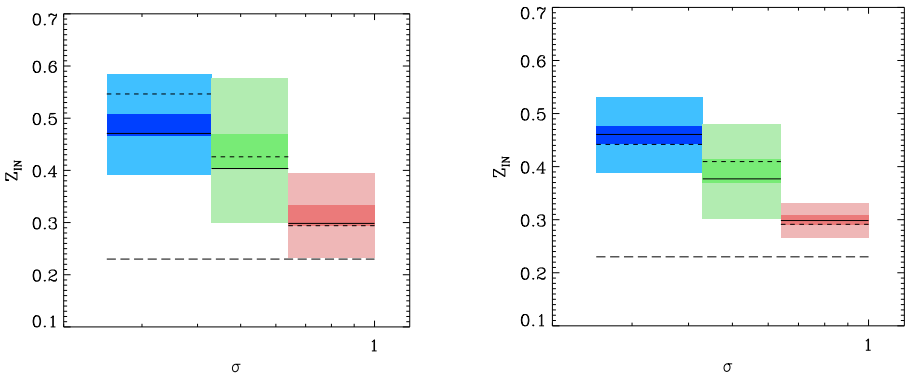

Figure 10.7: Summary of the main properties of the metallicity distributions for the three entropy classes. The color code is the same as used in Fig. 10.3. Abundances are expressed in Anders & Grevesse (1989) solar values. For each group, the dashed and the solid lines indicate the weighted average and the median respectively. The clearer area encloses the mean value plus or minus the standard deviation (i.e. indicates the one-sigma scatter of the values around the mean), and the darker area encloses the mean value plus or minus its uncertainty (calculated as the standard deviation over the square root of the number of elements). The long-dashed line at 0.23 $Z_{\odot}$ indicates the metallicity in the outer regions. In the right panel, few outliers are excluded (see text for details).

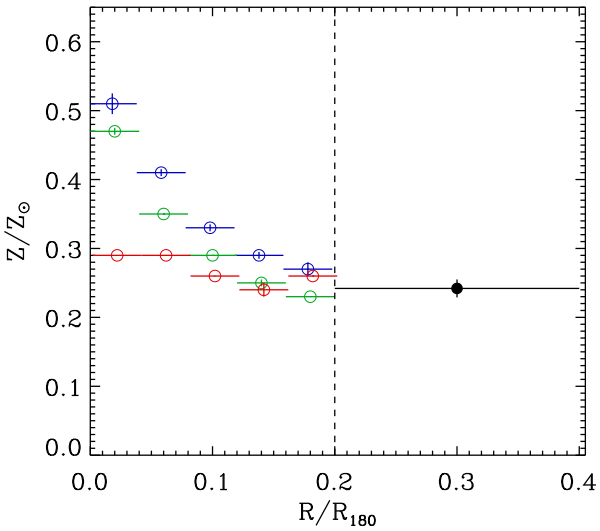

Figure 10.8: Mean metallicity profiles for LEC (blue circles), MEC (green), and HEC (red) clusters. The dashed line indicates 0.2 R_{180}, within that radius profiles are obtained from all clusters, the black point beyond that radius is the average value from distant (i.e. $z > 0.1$) clusters.

through a recursive sigma clipping procedure with a threshold of 1.5 σ. Results are reported in Fig. 10.7: in the left panel I considered all clusters, in the right one only those that survived the sigma clipping. When excluding the outliers, the correlation between metallicity and pseudo-entropy becomes more evident, the scatter is strongly reduced, and the shape of the distribution is much more symmetric (i.e. median, mean, and weighted average become consistent). Last but not least, an abundance excess with respect to the value measured in the outer regions (i.e. 0.23 $Z_{\odot}$, see Table 8.2) is found for all classes, whether the outliers were included or not. For HEC systems the excess is small, but nonetheless statistically significant.

10.3.2 Metallicity profiles

To provide a more comprehensive characterization of chemical properties I have also produced the mean radial metallicity profiles for each entropy class. In Fig. 10.8 I show the mean profiles for LEC, MEC, and HEC clusters (defined in Sect. 10.2). The binning used is in units of R_{180} and was computed as explained in Sect. 6.2; for each bin the average is calculated by allowing the existence of an intrinsic dispersion (see Fig. 10.7). I have performed measurements of the metal abundance for all clusters out to 0.2 R_{180} only. Beyond that radius I only averaged over distant clusters and obtained a metallicity of 0.242 ± 0.013 $Z_{\odot}$, that is remarkably consistent with the value obtained by De Grandi & Molendi (2001) for a local sample of relaxed clusters observed with *BeppoSAX*.

Within 0.1 R_{180}, a region typically associated with the core, all profiles show an abundance excess. The excess is strongest for LEC, somewhat weaker for MEC, and wekaest for HEC clusters. Between 0.1 and 0.2 R_{180} the profiles for the three classes are roughly consistent with one another and, at least for LEC systems, show a significant excess with respect to the mean value measured in the outskirts. Between 0.2 and 0.4 R_{180}, where I have data for the intermediate-redshift sample only, the three profiles are consistent with being flat and equal with each other.

10.4 Conclusions

In this section I will briefly describe the two main results of this work in progress, their discussion will be presented in a forthcoming paper (Leccardi et al., in prep.).

I find that most clusters, although showing no evidence of a temperature decline toward the center, exhibit a central metallicity peak. This may suggest that most clusters have spent a significant amount of time as a cool core (i.e. almost undisturbed) having had the chance of yielding a large quantity of metals in the central regions.

I also investigated the correlation between the central metallicity and the pseudo-entropy ratio, namely: the stronger the pseudo-entropy ratio, the stronger the metallicity peak. Despite the correlation, there are a few outliers with rather weak pseudo-entropy variations, but mild and strong metallicity peaks. This quite interesting result leads to a problem. Simulations (Poole et al. 2006) show that when a cool core is formed, it is almost impossible to destroy its entropy structure. Moreover, it is known that the main way for a cluster to produce a central metallicity peak is through a cool core phase (De Grandi et al. 2004). A possible explanation for such outliers is that the strongest interactions are able to destroy the entropy structure, but not all cool core evidences; the central metallicity peak is resistant and remains as a relic of the past cool core phase.

Bibliography

Abell, G. O. 1958, ApJS, 3, 211

Abell, G. O. 1965, ARA&A, 3, 1

Allen, S. W., Rapetti, D. A., Schmidt, R. W., et al. 2008, MNRAS, 383, 879

Anders, E. & Grevesse, N. 1989, Geochim. Cosmochim. Acta, 53, 197

Arnaud, M., Pointecouteau, E., & Pratt, G. W. 2005, A&A, 441, 893

Arzner, K., Güdel, M., Briggs, K., et al. 2006, astro-ph/0609193

Ascasibar, Y. & Markevitch, M. 2006, ApJ, 650, 102

Asplund, M., Grevesse, N., & Sauval, A. J. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 336, Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis, ed. T. G. Barnes, III & F. N. Bash, 25–+

Bahcall, N. A. 1975, ApJ, 198, 249

Baker, S. & Cousins, R. D. 1984, Nucl. instr. and meth. A, 221, 437

Baldi, A., Ettori, S., Mazzotta, P., Tozzi, P., & Borgani, S. 2007, ApJ, 666, 835

Balestra, I., Tozzi, P., Ettori, S., et al. 2007, A&A, 462, 429

Bergmann, U. C. & Riisager, K. 2002, Nucl. instr. and meth. A, 489, 444

Böhringer, H., Matsushita, K., Churazov, E., Finoguenov, A., & Ikebe, Y. 2004a, A&A, 416, L21

Böhringer, H., Schuecker, P., Guzzo, L., et al. 2004b, A&A, 425, 367

Borgani, S., Murante, G., Springel, V., et al. 2004, MNRAS, 348, 1078

Brecher, K. & Burbidge, G. R. 1972, ApJ, 174, 253

Bridle, A. H. & Feldman, P. 1972, Nature Physical Science, 235, 168

Byram, E. T., Chupp, T. A., & Friedman, H. 1966, Science, 152, 66

Cash, W. 1979, ApJ, 228, 939

Cavaliere, A. & Fusco-Femiano, R. 1978, A&A, 70, 677

Cavaliere, A., Gursky, H., & Tucker, W. H. 1971, Nature, 231, 437

Churazov, E., Gilfanov, M., Forman, W., & Jones, C. 1996, ApJ, 471, 673

Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, ApJ, 648, L109

Cowan, G. 1998, Statistical data analysis (Oxford Science Publications)

De Grandi, S., Ettori, S., Longhetti, M., & Molendi, S. 2004, A&A, 419, 7

De Grandi, S. & Molendi, S. 2001, ApJ, 551, 153

De Grandi, S. & Molendi, S. 2002, ApJ, 567, 163

De Luca, A. & Molendi, S. 2004, A&A, 419, 837

de Plaa, J., Werner, N., Bleeker, J. A. M., et al. 2007, A&A, 465, 345

Dickey, J. M. & Lockman, F. J. 1990, ARA&A, 28, 215

Eadie, W. T., Drijard, D., James, F. E., Roos, M., & Sadoulet, B. 1971, Statistical Methods in Experimental Physics (Amsterdam: North-Holland Publishers)

Ehle, M., Breitfellner, M., Gonzales Riestra, R., et al. 2003, XMM-Newton Users' Handbook, http:/xmm.vilspa.esa.es/external/xmm_user_support/documentation/uhb/index.html

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

Ettori, S., De Grandi, S., & Molendi, S. 2002, A&A, 391, 841

Fabian, A. C., Arnaud, K. A., Bautz, M. W., & Tawara, Y. 1994, ApJ, 436, L63

Fabian, A. C. & Barcons, X. 1992, ARA&A, 30, 429

Fabian, A. C., Sanders, J. S., Taylor, G. B., et al. 2006, MNRAS, 366, 417

Fabjan, D., Tornatore, L., Borgani, S., Saro, A., & Dolag, K. 2008, MNRAS, 386, 1265

Felten, J. E., Gould, R. J., Stein, W. A., & Woolf, N. J. 1966, ApJ, 146, 955

Finoguenov, A., Arnaud, M., & David, L. P. 2001, ApJ, 555, 191

Finoguenov, A., Böhringer, H., & Zhang, Y. 2005, A&A, 442, 827

Forman, W., Kellog, E., Gursky, H., Tananbaum, H., & Giacconi, R. 1972, ApJ, 178, 309

Fritz, G., Davidsen, A., Meekins, J. F., & Friedman, H. 1971, ApJ, 164, L81

Fukazawa, Y., Ohashi, T., Fabian, A. C., et al. 1994, PASJ, 46, L55

Gabici, S. & Blasi, P. 2004, Astroparticle Physics, 20, 579

Gehrels, N. 1986, ApJ, 303, 336

Ghizzardi, S. 2001, XMM-SOC-CAL-TN-29, http://xmm.vilspa.esa.es/external/xmm_sw_cal/calib/documentation.shtml

Ghizzardi, S. 2002, XMM-SOC-PS-TN-22, http://xmm.vilspa.esa.es/external/xmm_sw_cal/calib/documentation.shtml

Giacconi, R., Murray, S., Gursky, H., et al. 1972, ApJ, 178, 281

Grevesse, N. & Sauval, A. J. 1998, Space Science Reviews, 85, 161

Gursky, H., Kellog, E., Leong, C., Tananbaum, H., & Giacconi, R. 1971a, ApJ, 165, L43

Gursky, H., Kellogg, E., Murray, S., et al. 1971b, ApJ, 167, L81+

Harris, D. E. & Romanishin, W. 1974, ApJ, 188, 209

Hauschild, T. & Jentschel, M. 2001, Nucl. instr. and meth. A, 457, 384

Irwin, J. A. & Bregman, J. N. 2000, ApJ, 538, 543

Irwin, J. A., Bregman, J. N., & Evrard, A. E. 1999, ApJ, 519, 518

Jading, Y. & Riisager, K. 1996, Nucl. instr. and meth. A, 372, 289

Jansen, F., Lumb, D., Altieri, B., et al. 2002, A&A, 365, L1

Katayama, H., Takahashi, I., Ikebe, Y., Matsushita, K., & Freyberg, M. J. 2002, astro-ph/0202480

Kearns, K., Primini, F., & Alexander, D. 1995, in ASP Conf. Ser. 77: Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes, 331–+

Keith, H. D. & Loomis, T. C. 1978, X-ray spectrometry, 7, 217

Kellogg, E., Gursky, H., Tananbaum, H., Giacconi, R., & Pounds, K. 1972, ApJ, 174, L65

King, I. 1962, AJ, 67, 471

Kuntz, K. 2006, ftp://epic3.xra.le.ac.uk/pub/cal-pv/meetings/mpe-2006-05/kk.pdf

Kuntz, K. D. & Snowden, S. L. 2000, ApJ, 543, 195

Leccardi, A. & Molendi, S. 2007, A&A, 472, 21

Leccardi, A. & Molendi, S. 2008a, A&A, 487, 461

Leccardi, A. & Molendi, S. 2008b, A&A, 486, 359

Lin, Y.-T. & Mohr, J. J. 2004, ApJ, 617, 879

Lumb, D., Warwick, R. S., Page, M., & de Luca, A. 2002, A&A, 389, 93

Markevitch, M. 2006, in ESA Special Publication, Vol. 604, The X-ray Universe 2005, ed. A. Wilson, 723–+

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

Markevitch, M. & Vikhlinin, A. 2007, Phys. Rep., 443, 1

Maughan, B. J., Jones, C., Forman, W., & Van Speybroeck, L. 2008, ApJS, 174, 117

Mazzotta, P., Rasia, E., Moscardini, L., & Tormen, G. 2004, MNRAS, 354, 10

McNamara, B. R., Nulsen, P. E. J., Wise, M. W., et al. 2005, Nature, 433, 45

Meekins, J. F., Gilbert, F., Chubb, T. A., H., F., & Henry, R. C. 1971, Nature, 231, 107

Mighell, K. J. 1999, ApJ, 518, 380

Mitchell, R. J., Culhane, J. L., Davison, P. J. N., & Ives, J. C. 1976, MNRAS, 175, 29

Mushotzky, R. F. 1984, Physica Scripta Volume T, 7, 157

Mushotzky, R. F. & Loewenstein, M. 1997, ApJ, 481, L63+

Nousek, J. A. & Shue, D. R. 1989, ApJ, 342, 1207

Peres, C. B., Fabian, A. C., Edge, A. C., et al. 1998, MNRAS, 298, 416

Peterson, J. R. & Fabian, A. C. 2006, Phys. Rep., 427, 1

Peterson, J. R., Paerels, F. B. S., Kaastra, J. S., et al. 2001, A&A, 365, L 104

Piffaretti, R., Jetzer, P., Kaastra, J. S., & Tamura, T. 2005, A&A, 433, 101

Pizzolato, F., Molendi, S., Ghizzardi, S., & De Grandi, S. 2003, ApJ, 592, 62

Ponman, T. J., Sanderson, A. J. R., & Finoguenov, A. 2003, MNRAS, 343, 331

Poole, G. B., Fardal, M. A., Babul, A., et al. 2006, MNRAS, 373, 881

Pratt, G. W., Arnaud, M., & Pointecouteau, E. 2006, A&A, 446, 429

Pratt, G. W., Böhringer, H., Croston, J. H., et al. 2007, A&A, 461, 71

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1992, Numerical recipes (Cambridge University Press)

Rasia, E., Mazzotta, P., Bourdin, H., et al. 2008, ApJ, 674, 728

Rephaeli, Y. 1977, ApJ, 212, 608

Roncarelli, M., Ettori, S., Dolag, K., et al. 2006, MNRAS, 373, 1339

Rossetti, M., Ghizzardi, S., Molendi, S., & Finoguenov, A. 2007, A&A, 463, 839

Rybicki, G. B. & Lightman, A. P. 1979, Radiative processes in astrophysics (New York, Wiley-Interscience, 1979. 393 p.)

Sanders, J. S. & Fabian, A. C. 2006, MNRAS, 371, 1483

Sanders, J. S., Fabian, A. C., Allen, S. W., & Schmidt, R. W. 2004, MNRAS, 349, 952

Sarazin, C. L. 1988, X-ray emission from clusters of galaxies (Cambridge Astrophysics Series, Cambridge: Cambridge University Press, 1988)

Sarazin, C. L. 2002, in ASSL Vol. 272: Merging Processes in Galaxy Clusters, 1–38

Sauvageot, J. L., Belsole, E., & Pratt, G. W. 2005, A&A, 444, 673

Serlemitsos, P. J., Smith, B. W., Boldt, E. A., Holt, S. S., & Swank, J. H. 1977, ApJ, 211, L63

Snowden, S. L., Egger, R., Freyberg, M. J., et al. 1997, ApJ, 485, 125

Spitzer, L. 1956, Physics of Fully Ionized Gases (Physics of Fully Ionized Gases, New York: Interscience Publishers, 1956)

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., Yoshida, N., & White, S. D. M. 2001, New Astronomy, 6, 79

Strüder, L., Briel, U., Dennerl, K., & et al. 2001, A&A, 365, L18

Tamura, T., Kaastra, J. S., den Herder, J. W. A., Bleeker, J. A. M., & Peterson, J. R. 2004, A&A, 420, 135

Tittley, E. R. & Henriksen, M. 2005, ApJ, 618, 227

Tornatore, L., Borgani, S., Dolag, K., & Matteucci, F. 2007, MNRAS, 382, 1050

Tozzi, P., Scharf, C., & Norman, C. 2000, ApJ, 542, 106

Turner, M. J. L., Abbey, A., Arnaud, M., & et al. 2001, A&A, 365, L27

Vikhlinin, A., Markevitch, M., Murray, S. S., et al. 2005, ApJ, 628, 655

Voigt, L. M. & Fabian, A. C. 2006, MNRAS, 368, 518

Voit, G. M. 2005, Advances in Space Research, 36, 701

Wachter, K., Leach, R., & Kellogg, E. 1979, ApJ, 230, 274

White, D. A. 2000, MNRAS, 312, 663

White, III, R. E., Day, C. S. R., Hatsukade, I., & Hughes, J. P. 1994, ArXiv Astrophysics e-prints

Zhang, Y.-Y., Finoguenov, A., Böhringer, H., et al. 2008, A&A, 482, 451

Zwicky, F. 1933, Helv. Phys. Acta, 6, 110