A&A 441, 893–903 (2005)
DOI: 10.1051/0004-6361:20052856

Astronomy & Astrophysics

The structural and scaling properties of nearby galaxy clusters

II. The *M–T* relation

M. Arnaud[1], E. Pointecouteau[1], and G. W. Pratt[2]

[1] CEA/DSM/DAPNIA Service d'Astrophysique, CE Saclay, L'Orme des Merisiers, Bât. 709, 91191 Gif-sur-Yvette, France
e-mail: marnaud@discovery.saclay.cea.fr
[2] MPE, Giessenbachstraße, 85748 Garching, Germany

Received 10 February 2005 / Accepted 19 June 2005

ABSTRACT

Using a sample of ten nearby ($z \lesssim 0.15$), relaxed galaxy clusters in the temperature range [2–9] keV, we have investigated the scaling relation between the mass at various density contrasts ($\delta = 2500, 1000, 500, 200$) and the cluster temperature. The masses are derived from NFW-type model fits to mass profiles, obtained under the hydrostatic assumption using precise measurements, with *XMM-Newton*, at least down to $\delta = 1000$. The logarithmic slope of the *M–T* relation is well constrained and is the same at all δ, reflecting the self-similarity of the mass profiles. At $\delta = 500$, the slope of the relation for the sub-sample of hot clusters ($kT > 3.5$ keV) is consistent with the standard self-similar expectation: $\alpha = 1.49 \pm 0.15$. The relation steepens when the whole sample is considered: $\alpha = 1.71 \pm 0.09$. The normalisation of the relation is discrepant (by ∼30 per cent), at all density contrasts, with the prediction from purely gravitation based models. Models that take into account radiative cooling and galaxy feedback are generally in better agreement with our data. We argue that remaining discrepancies, in particular at low δ, are more likely due to problems with models of the ICM thermal structure rather than to an incorrect estimate of the mass from X-ray data.

Key words. cosmology: observations – cosmology: dark matter – galaxies: clusters: general – galaxies: intergalactic medium – X-rays: galaxies: clusters

1. Introduction

From a theoretical point of view, galaxy clusters are characterised by their mass. Models of structure formation predict the space density, spatial distribution and physical properties of clusters (internal structure, radius, temperature, luminosity, etc) as a function of mass and redshift (see Bertschinger 1998, for a review). However, observationally, the mass is not easily measured, and the observed scaling relations are in fact expressed in terms of the temperature T, rather than the mass M. These scaling relations are important sources of information on the physics of cluster formation (e.g. Voit & Ponman 2003). For the information to be complete, we must determine the *M–T* relation itself, which provides the missing link between the gas properties and the mass. Furthermore, measures of the cosmological parameters, such as σ_8, Ω_m and w, from cluster abundance or spatial distribution, rely heavily on this relation to link the mass to the X-ray observables available from X-ray cluster surveys. The present error on the value of σ_8, as determined from X-ray observations, is dominated by uncertainty on the *M–T* relation (Pierpaoli et al. 2003; Viana et al. 2003; Henry 2004), and a precise calibration of this relation is mandatory if we want to do "precision" cosmology with clusters (Borgani 2003).

The average temperature is expected to be closely related to the mass, via the virial theorem. We can first define M_δ as the mass within the radius R_δ, inside which the mean mass density is δ times the critical density, $\rho_c(z) = 3h(z)^2 H_0/8\pi G$, at the cluster redshift. We then expect $h(z)M_\delta = A(\delta)T^{3/2}$, *if* clusters are in hydrostatic equilibrium *and* they obey self-similarity. Here $h(z)$ is the Hubble constant normalised to its local value and $A(\delta)$ depends on the internal structure. The above relation is remarkably well verified by adiabatic numerical simulations, down to $\delta \sim 200$, which roughly corresponds to the virialised part of clusters (e.g. Evrard & Gioia 2002).

For relaxed clusters, the mass can be derived from X-ray observations of the gas density and temperature profile and the hydrostatic equilibrium equation. In recent years, a sustained observational effort to measure the local *M–T* relation has been undertaken using *ROSAT*, *ASCA* and *BeppoSAX*, but no definitive picture has yet emerged. It is unclear whether the mass scales as $T^{3/2}$ as expected (Horner et al. 1999; Ettori et al. 2002; Castillo-Morales & Schindler 2003); or if this is true only in the high mass regime ($kT \gtrsim 4$ keV), with a steepening at lower mass (Nevalainen et al. 2000; Xu et al. 2001; Finoguenov et al. 2001); or even if the slope is higher over the entire mass range (Sanderson et al. 2003). The derived normalisations of the *M–T* relation derived from *ASCA* data are

generally lower than predicted by adiabatic numerical simulations (e.g. Nevalainen et al. 2000; Finoguenov et al. 2001), typically 40 per cent below the prediction of Evrard et al. (1996). On the other hand, using *BeppoSAX* data, Ettori et al. (2002) found a normalisation consistent with the predictions (although the errors were large).

These studies had to rely largely on extrapolation to derive the virial mass, and were limited by the low resolution and statistical quality of the temperature profiles. With *XMM-Newton* and *Chandra* we can now measure the mass profile of clusters with unprecedented accuracy. Using *Chandra* observations, Allen et al. (2001) derived an M–T relation slope of 1.51 ± 0.27, consistent with the self-similar model, and confirmed the offset in normalisation. However, their sample comprised only 5 hot (i.e massive) clusters ($kT > 5.5$ keV), and, due to the relatively small *Chandra* field of view, their M–T relation was established at R_{2500} (i.e. about $\sim 0.3 R_{200}$).

In a recent paper (Pointecouteau et al. 2005, hereafter, Paper I), we measured the integrated mass profiles of ten relaxed, nearby clusters observed with *XMM-Newton*. The sample has an excellent temperature coverage, from 2 to 9 keV. The mass profiles cover a wide range of radii (from $0.01 R_{200}$ to $0.7 R_{200}$), and are particularly well constrained between $0.1 R_{200}$ and $0.5 R_{200}$. In Paper I, we studied the structural properties of the mass profiles, in order to test current scenarios for the Dark Matter clustering. In this paper, these data are used to establish a precise M–T relation up to the virial radius. In Sect. 2, we describe how we derive the temperature and mass data. In Sect. 3 we present and compare the M–T relations at various density contrasts. We discuss the reliability of the X-ray mass estimates in Sect. 4. The derived M–T relations are discussed with respect to pre-*Chandra*/*XMM-Newton* results in Sect. 5, and with expectations from models in Sect. 6. Our conclusions are presented in Sect. 7.

Throughout the paper, results are given for the currently-favoured ΛCDM cosmology, with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.3$ and $\Omega_\Lambda = 0.7$.

2. The data

2.1. The mass at various density contrasts

We used the mass profiles determined in Paper I to estimate the mass at four density contrasts: $\delta = 2500, 1000, 500$ and 200[1]. We recall that the mass profiles were derived from the observed density and temperature profiles (corrected for PSF and projection effects) using the hydrostatic equilibrium equation. The mass and errors at each radius of the temperature profile were calculated using a Monte Carlo method.

In Paper I, we found that the mass profiles are well described by an NFW-type model (Navarro et al. 1997):

$$M_{\rm NFW}(r) = M_{200} \frac{\ln(1 + c_{200}x) - c_{200}x/(1 + c_{200}x)}{\ln(1 + c_{200}) - c_{200}/(1 + c_{200})} \quad (1)$$

where $x = r/R_{200}$. The concentration c_{200}, and the mass M_{200}, are free parameters determined by the fitting procedure. R_{200} is related to M_{200} via $M_{200}/R_{200}^3 = (4\pi/3)200\rho_{\rm c}(z)$. The best fitting M_{200} and R_{200} values are recalled in Table 1. The concentration parameters are given in Paper I. The corresponding NFW mass, M_δ, at any density contrast δ, can be derived from the best fitting c_{200} and M_{200} values using Eq. (1) ($M_\delta = M_{\rm NFW}(r_\delta)$ with $M_\delta/r_\delta^3 = (4\pi/3)\delta\rho_{\rm c}(z)$). However, the computation of the error is not straightforward since the uncertainties in c_{200} and M_{200} are correlated. To avoid this problem, for each δ under consideration, we refitted the observed mass profile data with an NFW model using M_δ and the corresponding concentration as free parameters in the fit. The errors on M_δ can then be derived from standard $\Delta\chi^2$ criteria for one interesting parameter, the concentration being optimised (minimum χ^2) for any given value of M_δ. The resulting M_{2500}, M_{1000} and M_{500} values are also shown in Table 1. We also list the outermost radius (in units of R_{200}), and the corresponding density contrast $\delta_{\rm obs}$, reached by the temperature profiles (and thus the outermost extent of the measured mass profiles).

We will use these mass estimates to study the M_δ–T relation at various density contrasts. Our study is thus based on a parametric model of the observed mass profiles, rather than directly on the measured mass data. Let us discuss this point in more detail. All of the clusters are observed down to at least $\delta = 1000$ ($R_{1000} = (0.47 \pm 0.02)\, R_{200}$, averaged over the whole sample), the only exceptions being A1983 and MKW9 ($\delta_{\rm obs} \sim 1400$). At $\delta = 1000$ and $\delta = 2500$, using the best fitting model rather than the data is simply equivalent to 'smoothing' the data (without data extrapolation). We checked that this does not introduce a bias in the following way: for each cluster, we estimated the mass at $\delta = 2500$ by interpolating the observed profile expressed as a function of density contrast[2] in the log-log plane. We then compared the interpolated value to that derived from the NFW fit to the mass profile. In all cases the values are consistent within their 1σ errors. The ratio of the two values has a median value of 0.99 across the sample, and there is no significant correlation with mass. This reflects the fact that the NFW model is a good fit to these data, particularly in the $0.1 R_{200}$–$0.5 R_{200}$ range (see Paper I).

As explained above, the estimates of M_{2500} and M_{1000} are made (almost) without data extrapolation. However, the mass estimates at $\delta = 500$ and $\delta = 200$ do involve extrapolation of the data. The M_{500} and M_{200} estimates rely on the assumption that the best fitting NFW model remains a good representation of the cluster mass profile beyond $\delta_{\rm obs}$. We further discuss the reliability of this assumption in Sect. 4.

2.2. Overall temperature

To investigate the M–T relation, we need to define a global temperature. For this quantity, we used the overall spectroscopic temperature of the $0.1 R_{200} \leq r \leq 0.5 R_{200}$ region. The lower boundary of $0.1\ R_{200}$ was chosen so as to avoid most of the cooling core, where a large dispersion is observed in the

[1] We chose these δs for the following reasons: $\delta = 2500$ allows a direct comparison with previous *Chandra* results; $\delta = 1000$ is the density contrast limit of our observations; $\delta = 500$ corresponds to the edge of the virialized part of clusters in a conservative approach; $\delta = 200$ is the classical "virial" radius in an $\Omega = 1$ universe.

[2] $M(\delta)$ is readily derived from $M(r)$, using $\delta = 3M(r)/4\pi\rho_{\rm c}(z)r^3$ at each discrete radial value of the mass profile.

Table 1. Physical cluster parameters. Masses are in units of $10^{14}\,M_\odot$, and are given for a ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $H_0 = 70$ km s^{-1} Mpc^{-1}. Errors are 1σ errors.

Cluster	z	kT (keV)	R_{200} (kpc)	M_{2500}	M_{1000}	M_{500}	M_{200}	R_{obs}/R_{200}	δ_{obs}
A 1983	0.0442	2.18 ± 0.09	1103 ± 136	0.43 ± 0.09	0.77 ± 0.22	1.09 ± 0.37	1.59 ± 0.61	0.38	1455
MKW 9	0.0382	2.43 ± 0.24	1006 ± 84	0.41 ± 0.07	0.66 ± 0.14	0.88 ± 0.20	1.20 ± 0.30	0.41	1401
A 2717	0.0498	2.56 ± 0.06	1096 ± 44	0.45 ± 0.04	0.79 ± 0.08	1.10 ± 0.12	1.57 ± 0.19	0.54	727
A 1991	0.0586	2.71 ± 0.07	1106 ± 41	0.58 ± 0.05	0.91 ± 0.09	1.20 ± 0.12	1.63 ± 0.18	0.60	655
A 2597	0.0852	3.67 ± 0.09	1344 ± 49	1.08 ± 0.07	1.69 ± 0.14	2.22 ± 0.22	3.00 ± 0.33	0.57	713
A 1068	0.1375	4.67 ± 0.11	1635 ± 47	1.47 ± 0.07	2.69 ± 0.16	3.87 ± 0.28	5.68 ± 0.49	0.58	622
A 1413	0.1430	6.62 ± 0.14	1707 ± 57	2.33 ± 0.13	3.66 ± 0.27	4.82 ± 0.42	6.50 ± 0.65	0.79	339
A 478	0.0881	7.05 ± 0.12	2060 ± 112	3.12 ± 0.31	5.43 ± 0.70	7.57 ± 1.11	10.8 ± 1.8	0.58	650
PKS 0745-191	0.1028	7.97 ± 0.28	1999 ± 77	3.32 ± 0.23	5.41 ± 0.49	7.27 ± 0.75	10.0 ± 1.2	0.57	694
A 2204	0.1523	8.26 ± 0.22	2075 ± 77	3.62 ± 0.22	6.11 ± 0.51	8.39 ± 0.81	11.8 ± 1.3	0.61	580

Columns: (1) Cluster name; (2) Redshift; (3) Spectroscopic temperature of the $0.1 < r < 0.5\ R_{200}$ region; (4) R_{200}, roughly the virial radius in numerical simulations; (5, 6, 7, 8) Total mass at density contrast $\delta = 2500, 1000, 500, 200$, derived from an NFW fit to the observed mass profile; (9) Fraction of R_{200} spectroscopically observed (outer radius of the final temperature bin); (10) Corresponding density contrast.

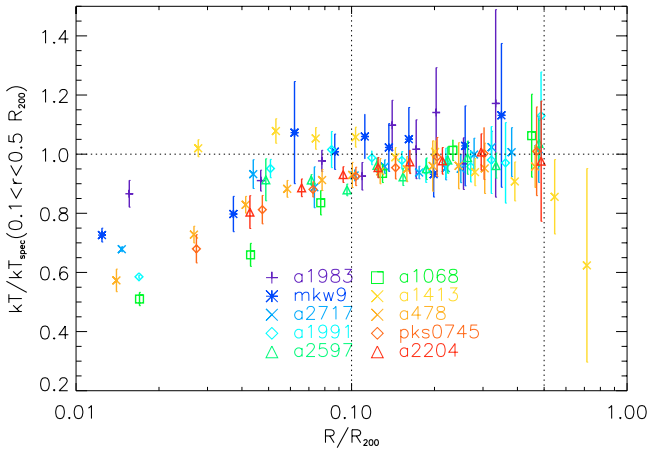

Fig. 1. Temperature profiles. The temperatures have been normalised to the spectroscopic temperature measured in $0.1 < r < 0.5\ R_{200}$ region; the radius has been scaled to R_{200}. The profiles have been corrected for PSF and projection effects (see Paper I for details).

temperature profiles (Fig. 1). The upper boundary is limited by the quality of the spectroscopic data. An upper boundary of $0.5\ R_{200}$ appeared a good compromise. Only the data from A1983 and MKW9 do not quite reach this radius; they are however detected up to $\sim 0.4\ R_{200}$ (see Table 1). Note that $0.5\ R_{200}$ corresponds roughly to $\delta = 1000$.

For each cluster, we performed an isothermal fit of the spectrum extracted within the $[0.1–0.5]R_{200}$ range, R_{200} being derived from the best fitting NFW model (see above). In the fit, the abundance was let free and the N_H was fixed to the 21 cm value (except for A478, see Pointecouteau et al. 2004). We corrected the derived value for PSF blurring and projection effects using the ratio of the mean emission-measure weighted value of the temperature profile in the $[0.1–0.5]\ R_{200}$ region after PSF/projection correction to the mean value before correction (see Paper I for details on the correction procedure). The correction factor is generally negligible and is always less than 5 per cent. The resulting temperature values are given in Table 1.

We could have estimated a "mass-weighted" temperature in the $0.1R_{200} < r < 0.5R_{200}$ region from the temperature profile. However, this temperature would still be a "spectroscopic" temperature since it would be derived from averaging over measured X-ray temperatures. It would not, strictly speaking, be equivalent to the "mass-weighted" temperature derived from numerical simulations. We thus preferred to use the overall spectroscopic temperature, which is a directly measured quantity and can also easily be estimated in numerical simulations. Since the region is defined in scaled radius, it can be derived from the simulated temperature profiles, using for instance the approach proposed recently by Mazzotta et al. (2004). In addition, we note that only a global spectroscopic temperature can be usually estimated for high z clusters. In our approach, the extraction region can be similarly defined and our definition thus allows a consistent study of the evolution of the M–T relation.

We note that Allen et al. (2001) use a mass-weighted temperature, T_{2500}, estimated from the temperature profile in the $r < R_{2500} \sim 0.3R_{200}$ region. In practice, their definition is equivalent to ours because i) the temperature profiles are fairly flat beyond the cooling core region ($r > 0.1R_{200}$) in both studies, and ii) the cooling core does not contribute much in mass to the average. This can be checked from Fig. 1 of Allen et al. (2001), where T_{2500} cannot be distinguished from the spectroscopic temperature of the region beyond the cooling core.

3. The *M–T* relation

3.1. The M_{2500}–T relation

In order to check the consistency of our *XMM-Newton* results with the *Chandra* study of Allen et al. (2001), we first investigated the M_{2500}–T relation. For our sample, $\delta = 2500$ corresponds to an average radius of $[0.29 \pm 0.02]\ R_{200}$, where the mass is particularly well constrained for all clusters.

The sample studied by Allen et al. (2001) comprises hot lensing clusters (5.5 to 15 keV). We thus considered only the sub-sample of clusters with moderate to high temperatures (i.e.

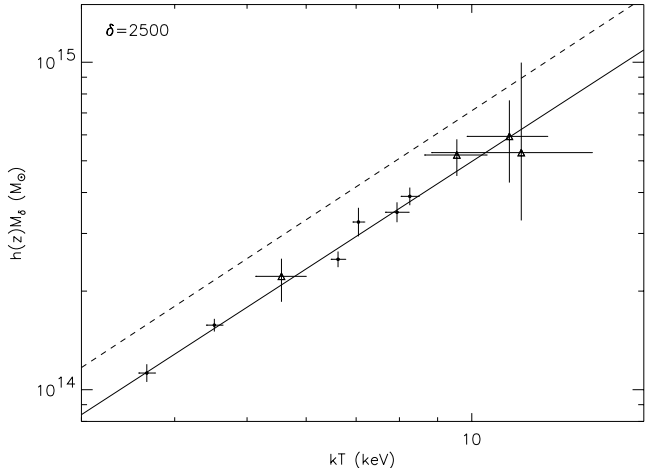

Fig. 2. The M–T relation at $\delta = 2500$ as seen by *XMM-Newton* from the observation of 6 hot ($kT > 3.5$ keV), relaxed clusters. Filled squares show the *XMM-Newton* data points; the full line shows the best fitting power law. The data on 4 published Chandra clusters (triangles) have been added to the fit but due to their larger uncertainties, they do not change the parameters of the fit to the *XMM-Newton* data only (see text). The dashed black line is the prediction from adiabatic numerical simulations (Evrard et al. 1996).

$T > 3.5$ keV), and fitted the M_{2500}–T relation using a power law model of the form:

$$h(z)M_\delta = A_\delta \left[\frac{kT}{5\ \mathrm{keV}}\right]^\alpha . \quad (2)$$

Here and in the following, the fit is performed using linear regression in the log–log plane, and the goodness of fit is calculated using a χ^2 estimator taking into account the errors on both mass and temperature. We used the routine FITEXY from numerical recipes (Press et al. 1992). Note that, as in the study of Allen et al. (2001), the masses are scaled by $h(z)$, which corrects for the evolution expected in the standard self-similar model. This scaling factor is small in our z range ($h(z) = 1.07$ at $z = 0.15$) but varies between ∼1.05 and ∼1.28 for the *Chandra* clusters located at higher redshifts ($0.1 < z < 0.46$).

The data are well fitted by a power law (χ^2/d.o.f = 2.5/4). The slope, $\alpha = 1.51 \pm 0.11$, is perfectly consistent with the expectation from the standard self-similar model, and with the results from *Chandra* observations ($\alpha = 1.51 \pm 0.27$). The derived normalisation, $A = (1.79 \pm 0.06) \times 10^{14}\ M_\odot$, is also consistent with the *Chandra* normalisation (see Table 3). As noted by Allen et al. (2001), such a normalisation is discrepant with the value derived from numerical simulations including gravitational heating only: our measured value is about ∼30 per cent below the prediction of Evrard et al. (1996). When the *Chandra* data for 4 of the 5 clusters studied by Allen et al. (2001) are added to the present data set[3], the best fitting values are almost unchanged ($\alpha = 1.52 \pm 0.1$ with the same intercept). This is due to the larger uncertainties in the *Chandra* temperature and mass determinations compared to those measured here (see Fig. 2). Figure 2 shows the best fit for the combined *XMM-Newton* and *Chandra* data compared to the expectations from the adiabatic numerical simulations of Evrard et al. (1996).

Still working at $\delta = 2500$, we performed a fit over the whole *XMM-Newton* sample, i.e. now including the four low mass systems. We obtain $\alpha = 1.70 \pm 0.07$, and a normalisation $A = (1.79 \pm 0.06) \times 10^{14}\ M_\odot$. The fit is acceptable, although formally less good (χ^2/d.o.f = 9.33/8). The slope now differs significantly from the expected value of $\alpha = 1.5$, and is just barely consistent with it at a 3σ level. This is further discussed in Sect. 6.2.

3.2. The M_δ–T relations up to the virial radius

Figure 3 shows the M_δ–T relations at various δ, together with the best fitting power law (Eq. (2)) in each case, and the prediction from the numerical simulations of Evrard et al. (1996). The best fitting slopes and normalisations are listed in Table 2, together with the standard errors. The best fits are listed and plotted both for the whole sample, and for the sub-sample of hot clusters. The corresponding R_δ–T relations are also given in the Table.

The normalisation and slope are nearly independent parameters for the whole sample. The covariance in $\log(A_\delta)$ and α, normalised to the product of their standard errors, is small: 0.045, 0.071, 0.092 and 0.11 for $\delta = 2500, 1000, 500$ and 200, respectively. This is due to our choice of the pivot of the M_δ–T relation, $kT = 5$ keV (Eq. (2)), close to the mean temperature of the whole sample (4.8 keV) or the median value of its temperature range (5.2 keV). The normalisation for other pivots can be derived using our best fitting values from Eq. (2). For instance, for a pivot at 6 keV, used in several works on cluster scaling relations, the normalisation A_δ is ∼36 per cent higher. The relative error, $\sigma_{A_\delta}/A_\delta$, is increased by ∼10 per cent.

We have also computed the raw and intrinsic scatter about the best fitting relations in the log-log plane. They are given in Table 2. To estimate the raw scatter, we used the orthogonal distances to the regression line, weighted by the error[4]. The intrinsic scatter is computed from the quadratic difference between the raw scatter and the scatter expected from the statistical errors.

The behaviour of the M_{2500}–T relation is reproduced at all other density contrasts. The slope is stable on all spatial scales: the variation is at most 10 per cent of the statistical error. It is always consistent with the expected $\alpha = 1.5$ value for the subsample of $T > 3.5$ keV clusters, whereas it steepens to $\alpha = 1.7$ when the cool clusters are included. Similarly, the normalisation remains ∼ 30 per cent below the value from adiabatic numerical simulations of Evrard et al. (1996) at all δ.

This stable behaviour is a direct consequence of the self-similarity of the mass profiles (Paper I). For an NFW type profile, the ratio of the masses at different density contrasts

[3] The fifth cluster is PKS 0745-191, which is common to both samples. We use only the *XMM-Newton* measurement here.

[4] For a linear relation of the form $Y = aX + b$, and a sample of N data points (Y_i, X_i) with errors σ_{Y_i} and σ_{X_i}, an estimate of the square of the raw scatter is: $\sigma^2_{\mathrm{raw}} = \frac{1}{N-2}\sum_{i=1}^{N} w_i (Y_i - aX_i - b)^2$, where $w_i = \frac{1/\sigma_i^2}{(1/N)\sum_{i=1}^{N} 1/\sigma_i^2}$ with $\sigma_i^2 = \sigma^2_{Y_i} + a^2\sigma^2_{X_i}$. Here $Y = \log(M_\delta), X = \log(T)$.

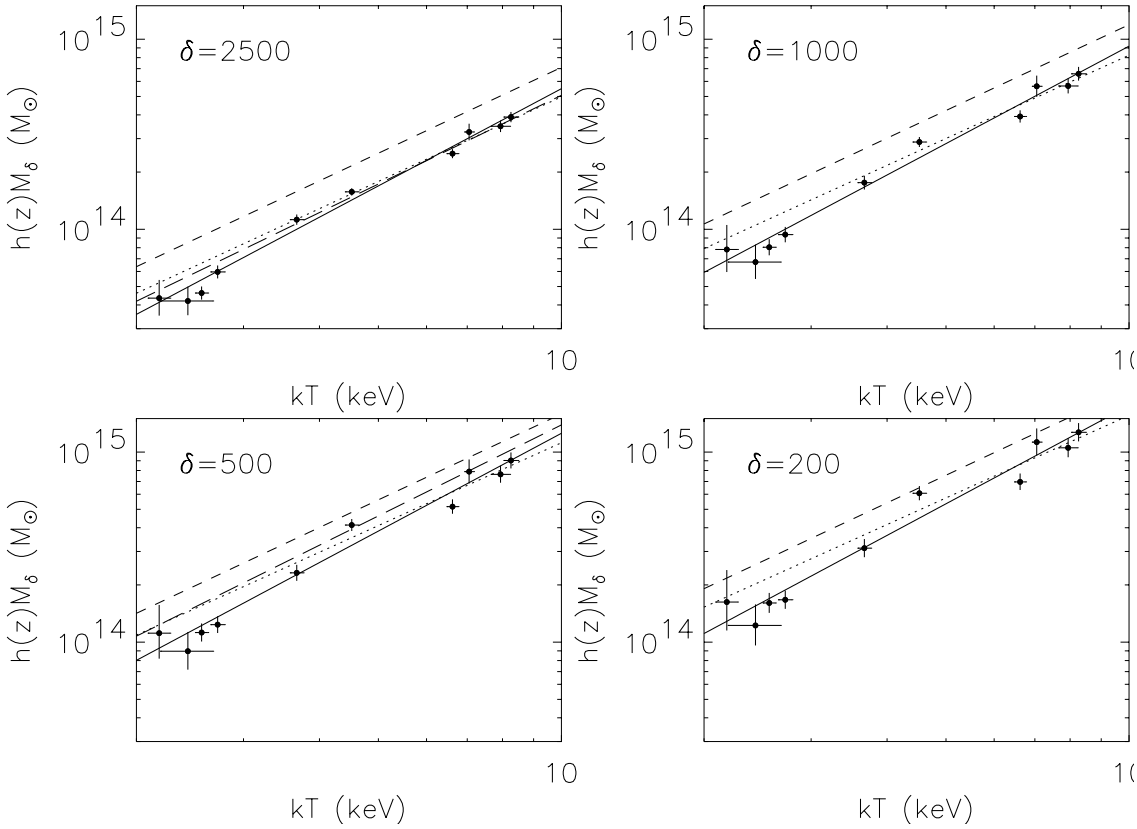

Fig. 3. The M–T relation as seen by *XMM-Newton* from a sample of 10 clusters covering a temperature range from 2 to 9 keV. From top to bottom and left to right, the M–T relation is given at the density contrasts δ of 200, 500, 1000 and 2500 with respect to the critical density of the Universe. Measurements are plotted with error bars. In each panel, the best fit for the whole sample is overplotted as a solid line, and the best fit for the hot cluster subsample is plotted as a dotted line. The predicted relation from adiabatic numerical simulations (Evrard et al. 1996) is overplotted as a dashed line. The long-dashed line (panels $\delta = 500$ and $\delta = 2500$) is the relation derived from a numerical simulation including radiative cooling, star formation and SN feedback (Borgani et al. 2004).

only depends on the concentration parameter, and $M_\delta = M_{2500}F(c,\delta)$. If all clusters had exactly the same concentration parameter (i.e., perfect self-similarity) the M_δ–T relations at various δ should differ only in their normalisation: $M_\delta = A(2500)F(c,\delta)T^\alpha$, for $M_{2500} = A(2500)T^\alpha$. The clusters in our sample are not perfectly self-similar but there is no significant variation of c with mass, and thus temperature (see Fig. 3 of Paper I). This explains the observed invariance of the slope. Furthermore, the observed concentration parameter is consistent with theoretical expectations (Paper I). As a consequence, the variation with δ of the normalisation, $A(\delta) = A(2500)F(c,\delta)$, follows expectations, and the offset with respect to simulations observed at $\delta = 2500$ remains the same at all δ (see also Fig. 4, and Fig. 13 of Pratt & Arnaud 2002).

However, the quality of the power law fit decreases with decreasing δ (see Table 2 and also Fig. 3). The reduced χ^2 increases and the corresponding null hypothesis probability for the whole sample varies from 0.32 at $\delta = 2500$ to 0.07 at $\delta = 200$. This behaviour corresponds to an increase of the intrinsic scatter in the observed M_δ–T relation (see Table 2). For the whole sample, the worst fit and largest intrinsic scatter is observed at $\delta = 200$. The regression method we used (Sect. 2.1) is strictly valid only if the intrinsic scatter is negligible as compared to the statistical scatter. This is not always the case (see Table 2), and we first checked if our results could be affected, using the M_{200}–T relation for the whole sample (i.e., the worst case). We refitted the data using the orthogonal BCES method (Akritas & Bershady 1996). While this is the least-biased regression method when both measurement errors and intrinsic scatter are present, it is less accurate that the χ^2 method when the intrinsic scatter is negligible. The best fitting values remain unchanged (within 0.5 per cent) and the standard error estimates are only slightly larger (by ∼15 per cent).

The regression method we used is thus justified. Nevertheless, the derived intrinsic scatter should not be over-interpreted. The cluster sample is small and is certainly not representative of the entire cluster population. In particular, it is heavily biased towards the more relaxed clusters. Moreover, the increased intrinsic scatter at low δ may be an artifact of the method we used to derive the various M_δ. These were derived from an NFW fit to the observed mass profile, using M_δ as a free parameter (see Sect. 2.1). In the NFW fit, M_{2500} is extremely well constrained by the data around $\delta = 2500$, quasi-independently of the shape of the observed mass profile. On the other hand, M_δ at low δ, beyond the maximum radius of

Table 2. Results of power law fits to the M_δ–T and R_δ–T relation at various density contrasts δ. The data are fitted with a power law of the form $h(z)M_\delta = A_\delta \times (kT/5\ \mathrm{keV})^\alpha$ and $h(z)R_\delta = B_\delta \times (kT/5\ \mathrm{keV})^\beta$, where kT is the overall spectroscopic temperature of the $[0.1R_{200}$–$0.5R_{200}]$ region. A ΛCDM cosmology is assumed: $\Omega_\mathrm{m} = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70$ km s^{-1} Mpc^{-1}.

	M_δ–T relation						R_δ–T relation	
δ	A_δ ($10^{14}\ M_\odot$)	α	$\sigma_{\log,\mathrm{raw}}$	$\sigma_{\log,\mathrm{int}}$	χ^2(d.o.f.)	nhp	B_δ (kpc)	β
Whole sample								
200	5.34 ± 0.22	1.72 ± 0.10	0.077	0.051	14.49(8)	0.07	1674 ± 23	0.57 ± 0.02
500	3.84 ± 0.14	1.71 ± 0.09	0.064	0.039	12.65(8)	0.12	1104 ± 13	0.57 ± 0.02
1000	2.82 ± 0.09	1.71 ± 0.08	0.053	0.027	10.74(8)	0.22	791 ± 8	0.57 ± 0.02
2500	1.69 ± 0.05	1.70 ± 0.07	0.041	0.016	9.33(8)	0.32	491 ± 4	0.56 ± 0.02
$T > 3.5$ keV								
200	5.74 ± 0.30	1.49 ± 0.17	0.081	0.064	10.45(4)	0.03	1714 ± 30	0.50 ± 0.05
500	4.10 ± 0.19	1.49 ± 0.15	0.064	0.046	8.34(4)	0.08	1129 ± 17	0.50 ± 0.05
1000	3.00 ± 0.12	1.49 ± 0.14	0.048	0.027	5.91(4)	0.21	807 ± 10	0.50 ± 0.04
2500	1.79 ± 0.06	1.51 ± 0.11	0.025	–	2.50(4)	0.65	500 ± 5	0.50 ± 0.03

Columns: (1) density contrast δ; (2), (3) intercept and slope for the M_δ–T relation: $h(z)M_\delta = A_\delta \times (kT/5\mathrm{keV})^\alpha$ with standard errors; (4), (5) raw and intrinsic scatter about the best fitting relations in the log–log plane (see Sect. 3.2); (6) Chi-squared and degree of freedom; (7) null hypothesis probability associated with the best fit; (8), (9) Intercept and slope for the $R_\delta - T$ relation: $h(z)R_\delta = B_\delta \times (kT/5\mathrm{keV})^\beta$.

observation, can be viewed as an "extrapolation" of the NFW model best fitting the observed mass profile. It thus depends both on the normalisation of the observed profile (basically M_{2500}) and on its shape. The shape parameter (the concentration c) is very sensitive to the data at small radii, in particular in the cooling core region, where the mass profile is least well constrained, and where there could be systematic errors due to the PSF/projection correction. As a result we expect, as observed, increasing statistical errors on M_δ as δ decreases, and a corresponding increase of the raw scatter in the M_δ–T relation (reflecting the scatter in c, see Fig. 3 of Paper I). The increase might be larger than that expected purely from the increase of statistical errors due to intrinsic scatter in c and/or systematic errors on c.

4. Reliability of X-ray mass estimates

As discussed in Paper I, there is an excellent *quantitative* agreement in *shape* between the X-ray mass profiles used in this work and the profile predicted by numerical simulations. The observed scaled profiles are well-described by the quasi-universal cusped profile (NFW-type) now found in all CDM simulations, and have concentration parameters as expected for their mass. As concluded in Paper I, this suggests that the Dark Matter collapse is well understood, at least down to the cluster scale. In turn, this gives us confidence in the *XMM-Newton* mass estimates, not only in the observed radial range, but also where we have extrapolated beyond it (i.e. at $\delta < \delta_\mathrm{obs} \sim 1000$). By using the best fitting NFW model to estimate M_{500} and M_{200}, we have implicitly assumed that this model remains valid beyond δ_obs (Sect. 2.1). It would be surprising if this were not the case since i) it is consistent with the theoretical predictions above δ_obs, and ii) in one case (A1413), we were even able to check the validity of the NFW profile down to $\delta < 500$.

Strictly speaking, the good agreement between the observed and predicted shape of the mass profiles does not mean that the *absolute* value of the X-ray mass is correct. It could be subject to systematic errors. However, for the correct universal shape of the mass profile to be recovered, this systematic error would have to be the same, within the statistical errors at all observed δ, whatever the cluster temperature.

One possible source of such systematic error is a departure from hydrostatic equilibrium (HE). The recent simulations of Kay et al. (2004a) suggest that the mass determined from the HE equation underestimates the true mass, due to residual gas motion. The effect is about the same at all radii up to $\delta = 500$. It is of the order of 15 per cent for adiabatic models and of 10 per cent for models including cooling and feedback, with typical variations of ±5 per cent. Such variations would not significantly change the shape of the X-ray mass profiles, taking into acccount our statistical errors. Thus the measured M_{2500} and M_{1000} values, and thus the corresponding normalisation of theM_{2500}–T and M_{1000}–T relations, could well be ∼10–15 per cent too low. The offset would be the same for the M_{500}–T and M_{200}–T relations, since they are derived from "extrapolation" of the NFW model. Note that this is probably an upper limit, since we focus on particularly relaxed clusters.

Another possible source of systematic error is that associated with errors in estimated temperatures from uncertainties in the instrument calibration. Possible errors are of the order of 10 per cent. This value is consistent with the systematic difference observed between temperature derived with *XMM-Newton* and *Chandra* with the former being on average 0.92 ± 0.08 times the latter (Kotov & Vikhlinin 2005). Since the mass derived from the HE equation scales as $M \propto T$, error in T would translate into a systematic error on the estimate of the 'real' mass of clusters. However that would not change the normalisation of the "observed" M–T relation, since that depends only on the *shape* of the temperature profile.

5. Comparison with previous determinations

In the present study, as discussed above, an NFW profile has been used to describe the integrated mass profile, derived, assuming HE, from the observed density and temperature

profiles. A similar approach[5] was used in the *Chandra* study of Allen et al. (2001). Previous *ROSAT*/*ASCA* studies also estimated the mass from the HE equation, but assumed a β-model for the gas density profile and a polytropic (or even isothermal) temperature profile (Horner et al. 1999; Nevalainen et al. 2000; Finoguenov et al. 2001; Xu et al. 2001; Castillo-Morales & Schindler 2003; Sanderson et al. 2003). Data were of poorer spatial resolution and statistical quality, and had less radial extent, thus requiring more extrapolation, particularly for low mass clusters. This could introduce systematic errors and biases, particularly at low δ, since a mass profile derived from an isothermal or polytropic β-model is not consistent with an NFW profile at large radii (see Fig. 11 of Neumann & Arnaud 1999). On the other hand, the latest *ROSAT*/*ASCA* studies of the M–T relation (Finoguenov et al. 2001; Sanderson et al. 2003) are superior in terms of the size of the cluster samples, and their wide and homogeneous coverage in temperature. Their results are compared to ours in Table 3.

Finoguenov et al. (2001) established the M_{500}–T relation for 39 clusters with *ASCA* temperature profiles. Interestingly, their results are consistent with ours within the uncertainties (Table 3). This suggests that systematic errors are not dominant over statistical errors at $\delta = 500$ in this *ROSAT*/*ASCA* study. However, our normalisation is on the upper side of their allowed values. The slope they find for their hot cluster subsample ($T > 3$ keV) is, as we have found, consistent with the expected $\alpha = 1.5$ value. When Finoguenov et al. (2001) included all clusters (down to $T \sim 0.9$ keV), they found a steepening of the M_{500}–T relation: $\alpha = 1.78 \pm 0.1$. The effect is larger than in our case: we find $\alpha = 1.71 \pm 0.09$ (although the difference is not significant). However, our sample does not reach quite such low temperatures and the difference could also reflect a progressive steepening of the M–T relation toward low masses.

The same remark holds for the results of Sanderson et al. (2003), who derived a slope of $\alpha = 1.84 \pm 0.06$ for the M_{200}–T relation, which is barely consistent with our value. Their large sample includes 66 clusters in the [0.5–15] keV temperature range. Furthermore, their normalisation at $\delta = 200$ is significantly lower than ours (by 15 per cent, Table 3). This may reflect the introduction of systematic errors when extrapolating polytropic models down to δ as low as 200. However, such systematic errors could not explain the discrepancy between our results and theirs at $\delta = 2500$. At that density contrast, their slope for a hot cluster ($T > 5.5$ keV) sub-sample is $\alpha = 1.84 \pm 0.14$, a result which is inconsistent with ours at the ∼95 per cent confidence level. Their normalisation at $\delta = 2500$, $(1.2 \pm 0.4) \times 10^{14}\ M_\odot$, is lower than our value of $(1.69 \pm 0.05) \times 10^{14}\ M_\odot$ and is only barely consistent with the *Chandra* results.

Sanderson et al. (2003) have suggested that the discrepancy (with respect to the *Chandra* results) might be related to the dynamical state of the clusters in the different samples. Both the *Chandra* study and the present *XMM-Newton* study focus on particularly relaxed clusters, which is not the case in Sanderson et al.'s study. However, we note that Finoguenov et al.'s sample does not discriminate in terms of dynamical state, and their results are in good agreement with ours. That said, the M–T relation could well depend on the exact dynamical states of the clusters in the sample in question, an effect which is not trivial to predict. The numerical simulations of Rowley et al. (2004) show that clusters with substructure tend to lie below the mean

[5] The method differs lightly from ours. Allen et al. (2001) predict the temperature profile corresponding to a given NFW mass profile and the observed surface brigthness profile and fit it to the observed temperature profile.

Table 3. Comparison of the present results with $M - T$ relations from the literature. α is the logarithmic slope of the relation and A is the normalisation at $kT = 5$ keV, in units of $10^{14}\ M_\odot$ for $H_0 = 70$ km s^{-1} Mpc^{-1}.

Reference[a]	A	α	Method[b]
$\delta = 2500$			
Observation			
Present work	1.69 ± 0.05	1.70 ± 0.07	$M_{\rm NFW}$, $T > 2.0$ keV
Present work	1.79 ± 0.06	1.51 ± 0.11	$M_{\rm NFW}$, $T > 3.5$ keV
ASF01	1.88 ± 0.34^c	1.51 ± 0.27	$M_{\rm NFW}$, $T > 5.5$ keV
SPF03	1.2 ± 0.4^c	1.84 ± 0.14	$M_{\beta\gamma}$, $T > 5.5$ keV
Theory			
EMN96	2.5	1.5	$T_{\rm em}$, adiabatic simul.
BMS04	1.73 ± 0.35	1.55 ± 0.05	$T_{\rm m}$,$T > 2.0$ keV
KSA04	1.97 ± 0.03	1.54 ± 0.05	$T_{\rm m}$
$\delta = 500$			
Observation			
Present work	3.84 ± 0.14	1.71 ± 0.09	$M_{\rm NFW}$, $T > 2.0$ keV
Present work	4.10 ± 0.19	1.49 ± 0.15	$M_{\rm NFW}$, $T > 3.5$ keV
FRB01	3.26 ± 0.60^c	1.48 ± 0.11	$M_{\beta\gamma}$, $T > 3$ keV
FRB01	3.31 ± 0.45^c	1.78 ± 0.10	$M_{\beta\gamma}$, $T > 0.9$ keV
Theory			
EMN96	5.6	1.5	$T_{\rm ew}$, adiabatic simul.
VBB02	3.6	∼1.7	$T_{\rm em}$
BMS04	4.6 ± 0.2	1.59 ± 0.05	$T_{\rm em}$,$T > 0.7$ keV
RMB05	7.2 ± 0.5	1.66 ± 0.09	$T_{\rm sl}$, $T > 1$ keV
RMB05	4.2 ± 0.2	1.53 ± 0.05	$M_{\beta\gamma}$, $T_{\rm sl}$; $T > 1$ keV
$\delta = 200$			
Observation			
Present work	5.34 ± 0.22	1.72 ± 0.10	$M_{\rm NFW}$, $T > 2$ keV
Present work	5.74 ± 0.30	1.49 ± 0.17	$M_{\rm NFW}$, $T >$ 3.5keV
SPF03	4.5 ± 0.3^c	1.84 ± 0.06	$M_{\beta\gamma}$, $T > 0.6$ keV
Theory			
EMN96	7.4	1.5	$T_{\rm em}$, adiabatic simul.
MTK02	7.6	1.61	$T_{\rm em}$,$T > 2.0$ keV

Notes: [a] References: (ASF01) Allen et al. (2001); (BMS04) Borgani et al. (2004); (EMN96) Evrard et al. (1996); (FRB01) Finoguenov et al. (2001); (KSA04) Kay et al. (2004b); (MTK02) Muanwong et al. (2002); (PW) present work; (RMB05) Rasia et al. (2005); (SPF03) Sanderson et al. (2003); (VBB02) Voit et al. (2002); [b] Method: ($M_{\rm NFW}$): mass estimated using an NFW model to describe the mass profile; ($M_{\beta\gamma}$): mass estimated using a polytropic β-model for the gas distribution; ($T_{\rm m}$): mass-weighted temperature; ($T_{\rm em}$): emission-weighted temperature; ($T_{\rm sl}$): spectroscopic-like temperature as defined in Mazzotta et al. (2004); [c] the normalisation at 5 keV is derived from the published normalisation at 1 keV and the best fitting slope. The fractional error has been assumed to be the same, which is conservative.

T–M relation, probably due to incomplete thermalisation (their Fig. 15). However, for the same reason, we would expect the X-ray mass to underestimate the true mass (Kay et al. 2004a), moving unrelaxed clusters back closer to the mean relation. A *XMM-Newton* study of the M–T relation for an unbiased sample of clusters is needed to assess the effect of cluster dynamical state on the measured M–T relation. The relation should ideally be compared to numerical simulations and lensing mass data. In any case, we do not confirm Sanderson et al.'s results, at least for the relaxed clusters considered here.

Finally, it is of interest to compare the present results with those of Ettori et al. (2002), who also use an NFW model to estimate masses[6]. Their M–T relation for a relaxed subsample of 12 clusters with $T > 3$ keV can be directly compared with our relation for hot clusters. At $\delta = 2500$, they found $\alpha = 1.88 \pm 0.27$, which is marginally consistent with our value, and $\alpha = 2.3 \pm 0.4$ at $\delta = 500$, a value clearly rejected by our data. However, there is a large scatter in their M–T relation at $\delta = 2500$, which becomes dramatic at $\delta = 500$ (their Fig. 5). We do not observe such a scatter. It may reflect systematic errors connected to the extrapolation of the NFW model. As we discussed above in Sect. 3.2, the precision on extrapolated mass depends on the precision on the concentration parameter (i.e., the shape of the mass profile, especially in the center), which is more difficult to constrain with *BeppoSAX* than with *XMM-Newton*, in particular due to the larger *BeppoSAX* PSF.

6. Comparison with theoretical predictions

The temperature structure of the ICM in a cluster is the result of the complex interplay between gravitational processes (i.e., the evolution of the gas in the Dark Matter potential), and of any other process that can affect the gas entropy (e.g., radiative cooling and heating from galaxy feedback). The theoretical M–T relation – which should be viewed rather as a T versus M relation when predicted from theoretical studies – depends on the exact modelling of all these processes. Moreover, as we discuss below, comparison of observations and theory also depends on the exact definition of the "average" temperature, since the gas is never perfectly isothermal.

6.1. The normalisation of the $M_\delta - T$ relation at 5 keV

The normalisation of the M–T relation is particularly well constrained by our study, the statistical error now being less than ∼5 per cent at 5 keV. The value of the normalisation depends on the (sub)sample considered because it is correlated with the slope. However by choosing a reference temperature of 5 keV, close to the median temperature, we minimize the effect, and the difference is of the order of the statistical error. We can thus first compare our results with the predicted values, quasi-independently of the slope issue.

[6] In their approach, the predicted temperature profile was fitted to the observed *BeppoSAX* temperature profile.

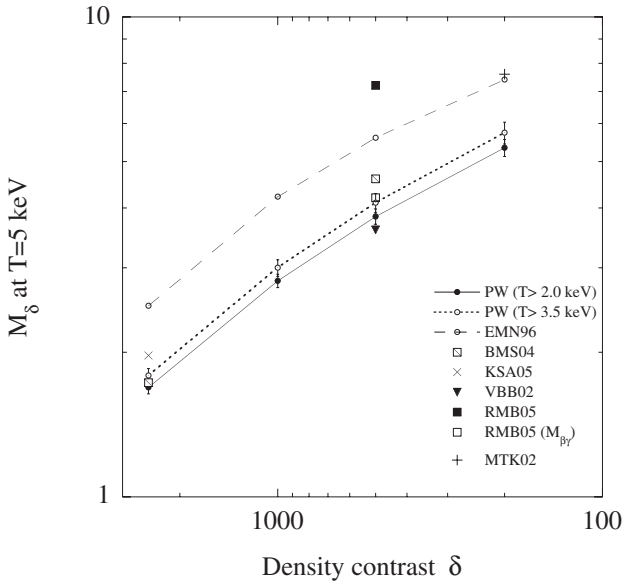

Fig. 4. The normalisation of the M_δ–T relation at $T = 5$ keV for various δ. The results of theoretical works (see Table 3 for references) are compared to the values derived in the present work (PW) .

6.1.1. Adiabatic models

The difference with the prediction from adiabatic models is not dramatic, particularly when the dispersion among various adiabatic simulations is taken into account. The normalisation is 30 per cent below the prediction of Evrard et al. (1996), compared to typically more than 50 per cent difference in the normalisation derived from different adiabatic simulations (Table 3; Henry 2004). However, as discussed by Muanwong et al. (2002), higher resolution simulations tend to predict higher normalisation, thus exacerbating the discrepancy with our results. Adiabatic simulations probably fundamentally fail to predict the correct normalisation of the M–T relation. The observed discrepancy could in principle be due to incorrect modelling of the dark matter component itself, since it is this which drives the potential. However, we have already argued that this is not the case from the excellent agreement between the observed and simulated mass profiles (Paper I and Sect. 4). A more likely explanation is that non-gravitational processes affect the M–T relation, as they affect all other scaling relations.

6.1.2. Non-adiabatic models

From theoretical arguments and from observations of the gas entropy, it is now clear that both cooling and galaxy feedback have to be taken into account when discussing relations involving the ICM (Voit & Ponman 2003; Pratt & Arnaud 2005). We thus focus on published theoretical studies that include both phenomena. Their predictions are compared with our results in Table 3 and in Fig. 4. All of the quoted simulations are made in the "concordance" ΛCDM cosmology, and use the same definition of M_δ as in the present work. However, they do not always use the same definition of the temperature (see below).

At a given mass, cooling and galaxy feedback increase the gas entropy as compared to the value attained by pure gravitational heating. In many scenarios, these processes increase the temperature and, as expected, a lower normalisation is found when these processes are included in a given numerical simulation (Muanwong et al. 2002; Thomas et al. 2002) or analytical model (Voit et al. 2002). Recent non-adiabatic simulations seem to be quite successful at reproducing the observed M_{2500}–T relation. The normalisation derived by Borgani et al. (2004), for a sample of simulated clusters with temperatures $T > 2$ keV, is in perfect agreement with our value (see also Fig. 3), while the normalisation derived by Kay et al. (2004b, see also Thomas et al. 2002) is only ∼10 per cent too high. These works used the mass weighted temperature, $T_{\rm m,2500}$, estimated within R_{2500}, which is directly linked to the thermal energy. As discussed in Sect. 2.2 our spectroscopic temperature, $T_{\rm s}$, should be close to the mass weighted temperature in that region. Thus, the good agreement between the observed M_{2500}–$T_{\rm s}$ and predicted M_{2500}–$T_{\rm m}$ relations is encouraging, and suggests that the thermal energy content in this central region is roughly correctly modelled.

Up until recently, the standard temperature definition used to mimic X-ray observations in numerical simulation studies was the emission-weighted temperature $T_{\rm em}$. Using this temperature definition at $\delta = 500$, the normalisation of Borgani et al. (2004) is now too high by about 12–20 per cent, the normalisation found by Voit et al. (2002) is marginally too low and the normalisation found by Muanwong et al. (2002) at $\delta = 200$ is 30 per cent too high (Fig. 4). It is likely that the variability of these results is linked to differences in the various physical models used.

Independent of the physics, a crucial point seems to be the exact definition of the temperature. Recently, Mazzotta et al. (2004) introduced the spectroscopic-like temperature ($T_{\rm sl}$) in order to better reproduce the temperature obtained from spectral fits when the ICM is multi-temperature. Mazzotta et al. show that $T_{\rm sl}$ is biased towards the lower values of the dominant thermal component, and that in general $T_{\rm em}$ overestimates $T_{\rm sl}$. Unfortunately, this exacerbates the disagreement between observed and simulated M_{500}–T normalisations. Using $T_{\rm sl}$, Rasia et al. (2005) over-predict a normalisation, relative to our values, by a factor as large as ∼1.8. Note that Rasia et al. use the same physical model as Borgani et al. (2004) and yet their normalisation is ∼50 per cent higher (see Table 3 and also their Fig. 2). We note that the earlier work of Mathiesen & Evrard (2001), based on spectroscopic temperatures of adiabatic numerical simulations, showed a smaller effect. Strictly speaking these temperatures were estimated with R_{500}, whereas our spectroscopic temperature measurement $T_{\rm s}$ is interior to $0.5R_{200}$ ($\sim R_{1000}$). A1413 is the only cluster for which we have data up to $\delta = 500$. The spectroscopic temperature within $\delta = 500$, $T_{\rm s,500}$ is only slightly smaller (by 3 per cent) than $T_{\rm s}$. This would increase the normalisation of the M_{500}-T relation by less than 5 per cent if we used $T_{\rm s,500}$ (assuming the same correction factor for all clusters).

It thus appears that there is a genuine disagreement between observed and predicted normalisation of the M_{500}–T relation. One interpretation, as proposed by Rasia et al. (2005), is that the X-ray mass underestimates the "true" mass (see also Borgani et al. 2004; Muanwong et al. 2002). Using $T_{\rm sl}$ temperatures, they estimated the value of M_{500} that an X-ray observer would derive from their simulation using the HE equation and a polytropic β-model. The resulting M_{500}–$T_{\rm sl}$ is indeed now in good agreement with our observation (Table 3 and Fig. 4). However, the normalisation is $4.2 \pm 0.2 \times 10^{14}\ M_\odot$, as compared to $7.2 \pm 0.5 \times 10^{14}\ M_\odot$ when using the "true" theoretical mass. This corresponds to a very serious underestimate of the mass by X-ray observations: the "true" M_{500} mass of clusters would be a factor $\sim 7.2/4.2 = 1.7$ higher than the X-ray mass. We think this is very unlikely, at least for the masses estimated as in the present work. Firstly, our approach – fitting an NFW model and extrapolating the mass profiles – is more sophisticated than the simple polytropic β-model approach. Secondly, we note again the excellent quantitative agreement of our mass profiles with theoretical predictions. If we have underestimated the "true" M_{500} by a factor of 1.7, we should also have underestimated the "true" M_{2500} by the same factor. This is unlikely: from combined lensing/X-ray studies, Allen et al. (2001) conclude that systematic uncertainties are less than 20 per cent[7]. Conversely, for the X-ray mass profiles to have the correct universal shape, as we have observed, the predicted difference between the X-ray mass estimates and the true mass should be roughly constant with radius. This is not what is expected if the difference is important at $\delta = 500$: it is linked to differences between the temperature profile derived from projected $T_{\rm sl}(r)$ values and the true profile, which depends on the ICM structure along the line of sight (Mazzotta et al. 2004; Rasia et al. 2005), and thus a priori on radius. It would be interesting to check this point with numerical simulations.

A more likely explanation is that numerical simulations do not correctly describe the gas thermal structure at large scale, at least for relaxed clusters considered here. We note that numerical simulations predict temperature profiles decreasing with radius, by nearly a factor 2 at $\delta = 500$ (Borgani et al. 2004, Fig. 6), while the observed profile is flatter (Fig. 1). This would bias low the $T_{\rm sl}$ as compared to the X–ray temperature of real clusters, and thus increase the normalisation of the M_{500}–$T_{\rm sl}$ relation[8]. It would be interesting to compare the theoretical M_{2500}–$T_{\rm sl}$ and M_{2500}–$T_{\rm m}$ relations, and investigate if there is continued good agreement with observations. We expect this to be the case since the predicted temperature variations are not dramatic – less than 20 per cent variations within $\delta = 2500$ (see Fig. 6 of Borgani et al. 2004 and Fig. 9 of Kay et al. 2004a), so that $T_{\rm sl}$ should be close to $T_{\rm m}$.

6.2. The slope of the M_δ–T relation

The observed M–T relation slope is consistent with the self-similar expectation for the sub-sample of hot clusters ($T > 3.5$ keV): $\alpha = 1.51 \pm 0.11$ at $\delta = 2500$, where it is best

[7] Note that our M_{2500}–T relation is the same, both in slope and normalisation, as that found by Allen et al. (2001).

[8] Since $T_{\rm em}$ is less affected, the predicted M_{500}–$T_{\rm em}$ relations would be in better agreement with observations.

constrained. The slope is significantly higher when the whole sample ($T > 2$ keV) is considered: $\alpha = 1.71 \pm 0.07$.

A value of $\alpha = 1.5$ is expected from the virial theorem if clusters obey self-similarity. All adiabatic simulations confirm this value (Evrard et al. 1996; Pen 1998; Eke et al. 1998; Bryan & Norman 1998; Yoshikawa et al. 2000; Thomas et al. 2001), including when a wide bandpass spectral temperature, as measured with *Chandra* or *XMM-Newton*, is used to establish the M–T relation (Mathiesen & Evrard 2001). Numerical simulations including cooling and feedback do predict a slightly higher slope. However, the effect is smaller than we observe ($\Delta(\alpha) = 0.05$–0.1) and is generally not significant (Table 3). The only exception is the M_{500}–$T_{\rm sl}$ relation derived by Rasia et al. (2005): $\alpha = 1.66 \pm 0.09$; however the normalisation is then much too high (see above). It is also worth noting that the phenomenological analytical model of Voit et al. (2002) yields a steeper slope. We obtained $\alpha \sim 1.7$ by fitting their relation (their Fig. 22) in our temperature range. This is in good agreement with the observed value; however, in this case the larger slope is mostly due to the variation of the concentration of the Dark Matter with mass in their model (Voit et al. 2002), which is larger than we observe (Paper I).

As a final remark, we want to emphasise that the observed discrepancy with the standard self-similar value is actually small. The slope increase, observed when including cool clusters, is significant at most at the ∼85 per cent confidence level. Furthermore, at 2 keV, the limiting temperature of our sample, this corresponds to only −20 per cent difference in mass as compared to the extrapolation of the best fitting M–T relation for hotter clusters (see also Fig. 3). There is scatter in the M–T relation, and our sample comprises only 4 cool clusters. We thus cannot exclude that the steepening is an artefact of our particular choice of clusters. We also note that the quality of the power law fit decreases when including low mass systems. This may indicate that the M–T relation is actually convex in the log-log plane, either across the entire temperature range, or below a "break" temperature. We lack clusters in the intermediate temperature range to assess this issue. Clearly, a possible discrepancy between predicted and observed slopes needs to be confirmed and better specified by considering a larger cluster sample.

7. Conclusion

Using a sample of ten relaxed galaxy clusters observed with *XMM-Newton*, we have calibrated the local M_δ–T relation, in the temperature range [2–9] keV, at four density contrasts, $\delta = 2500, 1000, 500, 200$. We used the spectroscopic temperature estimated within $0.5R_{200}$ ($\delta \sim 1000$), excluding the cooling core region, and derived the masses at various δ from NFW profile fits to precise mass profiles measured up to at least $\delta = 1000$. We argue that our measured masses are particularly reliable. The logarithmic slope of the M_δ–T relation is the same at all δ, reflecting the self-similarity of the mass profiles. The slope is well constrained and is consistent with the standard self-similar expectation, $\alpha = 1.5$, for the sub-sample of hot clusters ($T > 3.5$ keV). The relation steepens to $\alpha \sim 1.7$ when the whole sample ($T > 2$ keV) is considered. The normalisation of the M–T relation is measured with a precision better than ±5 per cent and is 30 per cent below the value predicted by the adiabatic numerical simulations of Evrard et al. (1996).

Models that take into account radiative cooling and galaxy feedback are now in good agreement with the observed M_{2500}–T relation. We argue that remaining discrepancies at $\delta = 500$ and lower are more likely to be due deficiencies in models of the ICM thermal structure, to which the spectroscopic-like temperature seems to be very sensitive, rather than to an incorrect estimate of the mass from X-ray data.

More detailed comparisons are needed to understand the origin of the discrepancies between the predicted and observed M–T relations. Our directly measured M_{1000}–T relation now provides the most direct constraint at large scale for numerical simulations. Simulations of mass profiles, as would be determined by an X-ray observer using modern *Chandra* and *XMM-Newton* techniques, are also needed. This would be particularly interesting for relaxed cluster sub-samples, and using better representations of observed temperatures (e.g. as proposed by Mazzotta et al. 2004). Such data could be directly compared to observed mass profiles. This would provide information on i) possible overall systematic errors in X-ray mass estimates, and ii) further test the reliability of simulations to correctly reproduce the ICM structure.

On the observational side, study of a much larger, unbiased, sample is needed to i) determine the exact shape of the local M–T relation; ii) study its intrinsic scatter, and iii) assess the effect of cluster dynamical state on the M–T relation.

Acknowledgements. We thank A. Evrard for useful comments on the manuscript. We thank M. Bershady for providing the BCES software and for useful discussions on the statistical analysis. We are grateful to J. Ballet and H. Bourdin for further useful discussions on statistical analysis. The present work is based on observations obtained with *XMM-Newton* an ESA science mission with instruments and contributions directly funded by ESA Member States and the USA (NASA). E.P. acknowledges the financial support of CNES (the French space agency). G.W.P. acknowledges funding from a Marie Curie Intra-European Fellowship under the FP6 programme (Contract No. MEIF-CT-2003-500915).

References

Akritas M. G., & Bershady M. A 1996, ApJ, 470, 706
Allen, S. W., Schmidt, R. W., & Fabian, A. C. 2001, MNRAS, 328, L37
Bertschinger, E. 1998, Ann. Rev. Astron. Ap., 36, 599
Borgani, S., 2003, in The Riddle of Cooling Flows in Galaxies and Clusters of galaxies, ed. T. Reiprich, J. Kempner, & N. Soker, http://www.astro.virginia.edu/coolflow/
Borgani, S., Murante, G., Springel, V., et al. 2004, MNRAS, 348, 1078
Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80
Castillo-Morales, A., & Schindler, S. 2003, A&A, 403, 433
Ettori, S., De Grandi, S., & Molendi, S. 2002, A&A, 391, 841
Evrard, A. E., Metzler, C. A., & Navarro, J. F. 1996, ApJ, 469, 494
Evrard, A. E., & Gioia, I. 2002, in Merging Processes in Galaxy Clusters, ed. L. Feretti, I. M. Gioia, G. Giovannini, Astrophysics and Space Science Library, 272, 253
Finoguenov, A., Reiprich, T. H., & Böhringer, H. 2001, A&A, 368, 749

Eke, V. R., Navarro, J. F., &Frenk, C. S. 1998, ApJ, 503, 569
Henry, J. P. 2004, ApJ, 609, 603
Horner, D. J., Mushotzky, R. F., & Scharf, C. A. 1999, ApJ, 520, 78
Kay, S. T., Thomas, P. A., Jenkins, A., & Pearce, F. R. 2004a, MNRAS, 355, 1091
Kay, S. T., Thomas, P. A., Jenkins, A., & Pearce, F. R. 2004b, [arXiv:astro-ph/0411650]
Kotov, O., & Vikhlinin, A. 2005, ApJ, submitted, [arXiv:astro-ph/0504233]
Mathiesen, B. F., & Evrard, A. E. 2001, ApJ, 546, 100
Mazzotta, P., Rasia, E., Moscardini, L., & Tormen, G. 2004, MNRAS, 354, 10
Muanwong, O., Thomas, P. A., Kay, S. T., & Pearce, F. R. 2002, MNRAS, 336, 527
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Neumann, D. M., & Arnaud, M. 1999, A&A, 348, 711
Nevalainen, J., Markevitch, M., & Forman, W. 2000, ApJ, 532, 694
Pen, U. L. 1998, ApJ, 498, 60
Pierpaoli, E., Borgani, S., Scott, D., & White, M. 2003, MNRAS, 342, 163
Pointecouteau, E., Arnaud, M., Kaastra, J., & de Plaa, J. 2004, A&A, 423, 33
Pointecouteau, E., Arnaud, M., & Pratt, G. W. 2005, A&A, 435, 1,
Pratt, G. W., & Arnaud, M. 2002, A&A, 394, 375
Pratt, G. W., & Arnaud, M. 2005, A&A, 429, 791
Press, W. H., Teukolsky, S. A., Vetterling, S. A., & Flannery B.P. 1992, Numerical Recipes in Fortran 77, Second edition, 660
Rasia, E., Tormen, G., & Moscardini, L. 2004, MNRAS, 351, 237
Rasia, E., Mazzotta, P., Borgani, S., et al. 2005, ApJ, 618, L1
Rowley, D. R., Thomas, P. A., & Kay, S. T. 2004, MNRAS, 352, 508
Sanderson, A. J. R., Ponman, T. J., Finoguenov, A., Lloyd-Davies, E. J., & Markevitch, M. 2003, MNRAS, 340, 989
Thomas, P. A., Muanwong, O., Kay, S. T., et al. 2001, MNRAS, 324, 450
Thomas, P. A., Muanwong, O., Kay, S. T., & Liddle, A. R. 2002, MNRAS, 330, L48
Viana, P., Kay, S. T., Liddle, A. R., Muanwong, O., & Thomas, P. A. 2003, MNRAS, 346, 319
Voit, G. M., Bryan, G. L., Balogh, M. L., & Bower, R.G. 2002, ApJ, 576, 601
Voit, G. M., & Ponman, T. J. 2003, ApJ, 594, L75
Xu, H., Jin, G., & Wu, X. P. 2001, ApJ, 553, 78
Yoshikawa, K., Jing, Y. P., & Suto, Y. 2000, ApJ, 535, 593