Planck 2013 results. XX. Cosmology from Sunyaev-Zeldovich cluster counts


Total number of authors: 254

Published in: Astronomy and Astrophysics

Link to article, DOI: 10.1051/0004-6361/201321521

Publication date: 2013

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Planck 2013 results. XX. Cosmology from Sunyaev–Zeldovich cluster counts


(Affiliations can be found after the references)

Received 20 March 2013 / Accepted 4 February 2014

ABSTRACT

We present constraints on cosmological parameters using number counts as a function of redshift for a sub-sample of 189 galaxy clusters from the Planck SZ (PSZ) catalogue. The PSZ is selected through the signature of the Sunyaev-Zeldovich (SZ) effect, and the sub-sample used here has a signal-to-noise threshold of seven, with each object confirmed as a cluster and all but one with a redshift estimate. We discuss the completeness of the sample and our construction of a likelihood analysis. Using a relation between mass M and SZ signal Y calibrated to X-ray measurements, we derive constraints on the power spectrum amplitude σ8 and matter density parameter Ωm in a flat ΛCDM model. We test the robustness of our estimates and find that possible biases in the Y–M relation and the mass function are larger than the statistical uncertainties from the cluster sample. Assuming the X-ray determined mass to be biased low relative to the true mass by between zero and 30%, motivated by comparison of the observed mass scaling relations to those from a set of numerical simulations, we find that σ8 = 0.75 ± 0.03, Ωm = 0.29 ± 0.02, and σ8(Ωm/0.27)1/3 = 0.764 ± 0.025. The value of σ8 is degenerate with the mass bias; if the latter is fixed to a value of 20% (the central value from numerical simulations) we find σ8(Ωm/0.27)1/3 = 0.78 ± 0.01 and a tighter one-dimensional range σ8 = 0.77 ± 0.02. We find that the larger values of σ8 and Ωm preferred by Planck’s measurements of the primary CMB anisotropies can be accommodated by a mass bias of about 4%. Alternatively, consistency with the primary CMB constraints can be achieved by inclusion of processes that suppress power on small scales relative to the ΛCDM model, such as a component of massive neutrinos. We place our results in the context of other determinations of cosmological parameters, and discuss issues that need to be resolved in order to make further progress in this field.

Key words. cosmological parameters – large-scale structure of Universe – galaxies: clusters: general

* Corresponding author: M. Douspis e-mail: marian.douspis@ias. upasd.fr
1. Introduction

This paper, one of a set associated with the 2013 release of data from the Planck\(^1\) mission (Planck Collaboration I 2014), describes the constraints on cosmological parameters using number counts as a function of redshift for a sample of 189 galaxy clusters.

Within the standard picture of structure formation, galaxies aggregate into clusters of galaxies at late times, forming bound structures at locations where the initial fluctuations create the deepest potential wells. The study of these galaxy clusters has played a significant role in the development of cosmology over many years (see, for example, Perrenod 1980; Oukbir & Blanchard 1992; White et al. 1993; Carlberg et al. 1996; Voit 2005; Henry et al. 2009; Vikhlinin et al. 2009b; Allen et al. 2011a). More recently, as samples of clusters have increased in size and variety, number counts inferred from tightly-selected surveys have been used to obtain detailed constraints on the cosmological parameters.

The early galaxy cluster catalogues were constructed by eye from photographic plates with a “richness” (or number of galaxies) attributed to each cluster (Abell 1958; Abell et al. 1989). As time has passed, new approaches for selecting clusters have been developed, most notably using X-ray emission due to thermal Bremsstrahlung radiation from the hot gas that makes up most of the baryonic matter in the cluster. X-ray cluster surveys include both the NORAS (Böhringer et al. 2000) and REFLEX (Böhringer et al. 2004) surveys, based on ROSAT satellite observations, which have been used as source catalogues for higher-precision observations by the Chandra and XMM-Newton satellites, as well as surveys with XMM-Newton, including the XMM Cluster Survey (XCS, Mbrightness et al. 2012) and the XMM Large Scale Structure survey (XMM-LSS, Willis et al. 2013).

To exploit clusters for cosmology, a key issue is how the properties used to select and characterize the cluster are related to the total mass of the cluster, since this is the quantity most readily predicted using theoretical models. Galaxies account for a small fraction of the cluster mass, and the scatter between richness and mass appears to be large. However, there are a number of other possibilities. In particular, there are strong correlations between the total mass and both the integrated X-ray surface brightness and X-ray temperature, making them excellent mass proxies.

The Sunyaev-Zeldovich (SZ) effect (Sunyaev & Zeldovich 1970; Zeldovich & Sunyaev 1969) is the inverse Compton scattering of cosmic microwave background (CMB) photons by the hot gas along the line of sight, and this is most significant when the line of sight passes through a galaxy cluster. It leads to a decrease in the overall brightness temperature in the Rayleigh-Jeans portion of the spectrum and an increase in the Wien tail, with a null around 217 GHz (see Birkinshaw 1999 for a review). The amplitude of the SZ effect is given by the integrated pressure of the gas within the cluster along the line of sight. Evidence both from observation (Marrone et al. 2012; Planck Collaboration Int. III 2013) and from numerical simulations (Springel et al. 2001; da Silva et al. 2004; Moll et al. 2005; Nagai 2006; Kay et al. 2012) suggests that the SZ effect is an excellent mass proxy. A number of articles have discussed the possibility of using SZ-selected cluster samples to constrain cosmological parameters (Barbosa et al. 1996; Aghanim et al. 1997; Haiman et al. 2001; Holder et al. 2001; Weller et al. 2002; Diego et al. 2002; Battey & Weller 2003).

This paper describes the constraints on cosmological parameters imposed by a high signal-to-noise (S/N) sub-sample of the Planck SZ Catalogue (PSZ, see Planck Collaboration XXIX 2014, henceforth Paper I, for details of the entire catalogue) containing nearly 200 clusters (shown in Fig. 1). This sub-sample has been selected to be pure, in the sense that all the objects within it have been confirmed as clusters via additional observations, either from the literature or undertaken by the Planck collaboration. In addition all objects but one have a measured redshift, either photometric or spectroscopic. This is the largest SZ-selected sample of clusters used to date for this purpose. We will show that it is the systematic uncertainties from our imperfect knowledge of cluster properties that dominate the overall uncertainty on cosmological constraints.

The Planck cluster sample is complementary to those from observations using the South Pole Telescope (SPT, Carlstrom et al. 2011) and the Atacama Cosmology Telescope (ACT, Swetz et al. 2011), whose teams recently published the first large samples of SZ-selected clusters (Reichardt et al. 2013; Hasselfield et al. 2013). The resolution of Planck at the relevant frequencies is between 5 and 10 arcmin, whereas that for ACT and SPT is about 1 arcmin, but the Planck sky coverage is much greater. This means that Planck typically finds larger, more massive, and lower-redshift clusters than those found by SPT and ACT.

Our strategy is to focus on number counts of clusters, as a function of redshift, above a high S/N threshold of seven and to explore the robustness of the results. We do not use the observed SZ brightness of the clusters, due to the significant uncertainty caused by the size-flux degeneracy as discussed in Paper I. Accordingly, our theoretical modelling of the cluster population is directed only at determining the expected number of clusters in each redshift bin exceeding the S/N threshold. The predicted and observed numbers of clusters are then compared in order to obtain the likelihood. In the future, we will make use of the SZ-estimated mass and a larger cluster sample to extend the analysis to broader cosmological scenarios.

This paper is laid out as follows. We describe the theoretical modelling of the redshift number counts in Sect. 2, while Sect. 3 presents the PSZ cosmological sample and selection function used in this work. The likelihood we adopt for putting constraints on cosmological parameters is given in Sect. 4. Section 5 presents our results on cosmological parameter estimation and assesses their robustness. We discuss how they fit in with other cluster and cosmological constraints in Sect. 6, before providing a final summary. A detailed discussion of our calibration of the SZ flux versus mass relation and its uncertainties is given in Appendix A.

2. Modelling cluster number counts

2.1. Model definitions

We parameterize the standard cosmological model as follows. The densities of various components are specified relative to the present-day critical density, with \(\Omega_X = \rho_X / \rho_{crit}\) denoting that for component \(X\). These components always include matter, \(\Omega_m\), and a cosmological constant \(\Omega_{\Lambda}\). For this work we assume that the Universe is flat, that is, \(\Omega_m + \Omega_{\Lambda} = 1\), and the optical depth to reionization is fixed at \(\tau = 0.085\) except in the CMB and
The present-day expansion rate of the Universe is quantified by the Hubble constant $H_0 = 100 \text{ h km s}^{-1} \text{ Mpc}^{-1}$.

The cluster number counts are very sensitive to the amplitude of the matter power spectrum. When studying cluster counts it is usual to parametrize this in terms of the density variance in spheres of radius $8 \, h^{-1} \text{ Mpc}$, denoted $\sigma_8$, rather than overall power spectrum amplitude, $A_s$. In cases where we include primary CMB data we use $A_t$ and compute $\sigma_8$ as a derived parameter. In addition to the parameters above, we allow the other standard cosmological parameters to vary: $n_s$, representing the spectral index of density fluctuations; and $\Omega_b h^2$ quantifying the baryon density.

The number of clusters predicted to be observed by a survey in a given redshift interval $[z_i, z_{i+1}]$ can be written

$$n_i = \int_{z_i}^{z_{i+1}} \frac{dN}{dz}, \quad (1)$$

with

$$\frac{dN}{dz} = \int d\Omega \int dM_{500} \hat{\chi}(z, M_{500}, l, b) \frac{dN}{dz} dM_{500} d\Omega, \quad (2)$$

where $d\Omega$ is the solid angle element and $M_{500}$ is the mass within the radius where the mean enclosed density is 500 times the critical density. The quantity $\hat{\chi}(z, M_{500}, l, b)$ is the survey completeness at a given location $(l, b)$ on the sky, given by

$$\hat{\chi} = \int dY_{500} \int d\theta_{500} P(z, M_{500}|Y_{500}, \theta_{500}) \chi(Y_{500}, \theta_{500}, l, b). \quad (3)$$

Here $P(z, M_{500}|Y_{500}, \theta_{500})$ is the distribution of $(z, M_{500})$ for a given $(Y_{500}, \theta_{500})$, where $Y_{500}$ and $\theta_{500}$ are the SZ flux and size of a cluster of redshift and mass $(z, M_{500})$.

This distribution is obtained from the scaling relations between $Y_{500}$, $\theta_{500}$, and $M_{500}$, discussed later in this section. Note that $\chi(z, M_{500}, l, b)$ depends on cosmological parameters through $P(z, M_{500}|Y_{500}, \theta_{500})$, while the completeness in terms of the observables, $\chi(Y_{500}, \theta_{500}, l, b)$, does not depend on the cosmology as it refers directly to the observed quantities.

For the present work, we restrict our analysis to the quantity $dN/dz$ that measures the total counts in redshift bins. In particular, we do not use the blind SZ flux estimated by the cluster candidate extraction methods that, as detailed in Planck Collaboration VIII (2011), is found to be significantly higher than the flux predicted from X-ray measurements. In contrast to the blind SZ flux, the blind S/N is in good agreement with the S/N measured using X-ray priors. Figure 2 shows the blind S/N ($S/N_{\text{blind}}$) versus the S/N re-extracted at the X-ray position using the X-ray size, for the X-ray detections of Planck clusters that are associated with known X-ray clusters in the reference cosmological sample. In contrast to the blind SZ flux, the blind S/N is in good agreement with the S/N measured using X-ray priors.

$^2$ The two signal-to-noises are actually estimated at two different positions on the sky (blind SZ and X-ray position), leading to different values of both the signal and the noise. It thus happens that the recomputed S/N is higher than the blind SZ.
To carry out a prediction of the counts expected in a survey, given cosmological assumptions, we therefore need the following inputs:

- a mass function that tells us the number distribution of clusters with mass and redshift;
- scaling relations that can predict observable quantities from the mass and redshift;
- the completeness of the survey in terms of those observables, which tells us the probability that a model cluster would make it into the survey catalogue.

These are described in the remainder of this section and in the next.

### 2.2. Mass function

Our main results use the mass function from Tinker et al. (2008), giving the number of haloes per unit volume:

$$\frac{dN}{dM_{500}}(M_{500}, z) = f(\sigma) \rho_m(z = 0) \ln(\sigma^{-1})M_{500}^{-1}dM_{500}, \quad (4)$$

where

$$f(\sigma) = A \left[ 1 + \left( \frac{\sigma}{\sigma_500} \right)^{\alpha_500} \right] \exp \left(-\frac{\sigma}{\sigma_500} \right), \quad (5)$$

and $$\rho_m(z = 0)$$ is the mean matter density at $$z = 0$$. The coefficients $$\alpha$$, $$\beta$$, and $$\sigma$$ are tabulated in Tinker et al. (2008) for different overdensities, $$\Delta_{\text{mean}}$$, with respect to the mean cosmic density, and depend on $$z$$. Here we use $$\Delta_{\text{critical}} = 500$$ relative to the critical density, so we compute the relevant mass function coefficients by interpolating the Tinker et al. (2008) tables for haloes with $$\Delta_{\text{mean}} = \Delta_{\text{critical}}/\Omega_m(z) = 500/\Omega_m(z)$$, where $$\Omega_m(z)$$ is the matter density parameter at redshift $$z$$.

The quantity $$\sigma$$ is the standard deviation, computed in linear perturbation theory, of the density perturbations in a sphere of radius $$R$$, which is related to the mass by $$M = 4\pi\sigma_m^3R^3/3$$. It is given by

$$\sigma^2 = \frac{1}{2\pi^2} \int dk k^2 P(k, z) W(kR)^2, \quad (6)$$

where $$P(k, z)$$ is the matter power spectrum at redshift $$z$$, which we compute for any given set of cosmological parameters using CAMB (Lewis et al. 2000), and $$W(z) = 3(\sin x - x \cos x)/x^2$$ is the filter function of a spherical top hat of radius $$R$$.

The quantity $$dN/dM_{500}(z)/dM_{500}$$ by the volume element $$dV/(dz\,dM_{500})$$ in Eq. (2) is computed by multiplying the mass function $$dN(M_{500}, z)/dM_{500}$$ by the volume element $$dV/(dz\,dM_{500})$$.

As a baseline we use, except where stated otherwise, the Tinker et al. (2008) mass function, but we consider an alternative mass function as a cross-check. In a recent publication by Watson et al. (2013), a new mass function is extracted from the combination of large cosmological simulations (typical particle numbers of 3000$^3$ to 6000$^3$) with a very large dynamic range (size from 11 $h^{-1}$ to 6000$ h^{-1}$ Mpc), which extends the maximum volume probed by Tinker et al. by two or orders of magnitude. The two mass functions agree fairly well, except in the case of the most massive objects, where Tinker et al.’s mass function predicts more clusters than Watson et al.’s. The Tinker et al. mass function might be derived from volumes that are not large enough to properly sample the rarer clusters. These rare clusters are more relevant for Planck than for ground-based SZ experiments, which probe smaller areas of the sky. The Watson et al. mass function is used only in Sect. 5.3, which deals with mass function uncertainties.

### Table 1. Summary of scaling-law parameters and error budget.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\log Y$</td>
<td>$-0.19 \pm 0.02$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$1.79 \pm 0.08$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$0.66 \pm 0.50$</td>
</tr>
<tr>
<td>$\sigma_{\log Y}$</td>
<td>$0.075 \pm 0.01$</td>
</tr>
</tbody>
</table>

**Notes.** $\beta$ is kept fixed at its central value except in Sect. 5.3.

### 2.3. Scaling relations

A key issue is to relate the observed SZ flux, $$Y_{500}$$, to the mass $$M_{500}$$ of the cluster. As we show in Sect. 5, cosmological constraints are sensitive to the normalization, slope and scatter of the assumed $$Y_{500} - M_{500}$$ relation. We thus paid considerable attention to deriving the most accurate scaling relations possible, with careful handling of statistical and systematic uncertainties, and to testing their impact on the derived cosmological parameters.

The baseline relation is obtained from an observational calibration of the $$Y_{500} - M_{500}$$ relation on one-third of the cosmological sample. The calibration uses $$M_{500}^\text{YX}$$, the mass derived from the X-ray $$Y_X - M_{500}$$ relation, as a mass proxy. Here $$Y_X$$ is the X-ray analogue of the SZ signal introduced by Kravtsov et al. (2006), as defined in Appendix A. $$Y_{500}$$ is then measured interior to $$R_{500}^X$$, the radius corresponding to $$M_{500}^\text{YX}$$. The mean bias between $$M_{500}^\text{YX}$$ and the true mass, $$(1 - b)$$, is assumed to account for all possible observational biases (departure from hydrostatic equilibrium (HE), absolute instrument calibration, temperature inhomogeneities, residual selection bias, etc.) as discussed in full in Appendix A. In practice, the plausible range for this mean bias $$(1 - b)$$ was estimated by comparing the observed relation with predictions from several sets of numerical simulations, as detailed in Appendix A.

The large uncertainties on $$(1 - b)$$ are due to the dispersion in predictions from the various simulation sets. This is a major factor limiting the accuracy of our analysis. A value $$(1 - b) = 0.8$$ could be considered as a best guess given available simulations, with no clear dependence on mass or redshift. From one cluster to the next the ratio of $$M_{500}^\text{YX}$$ to the true mass is expected to be stochastic, contributing to the scatter in the $$Y_{500} - M_{500}$$ relation given below. A conspiracy of all possible sources of bias (departure from HE, absolute instrument calibration, temperature inhomogeneities, residual selection bias) would seem necessary to lead to a significantly lower value of $$(1 - b)$$. This apparently implausible possibility needs to be excluded through tests using other probes such as baryon and gas fractions, gas pressure, etc. As a baseline we take $$(1 - b) = 0.8$$ with a flat prior. We also consider, when analysing systematic uncertainties on the derived cosmological parameters, a case where the bias is fixed to the value $$(1 - b) = 0.8$$. As detailed in Appendix A, we derive a baseline relation for the mean SZ signal $$\bar{Y}_{500}$$ from a cluster of given mass and redshift in the form

$$E^{-\beta}(z) \left[ \frac{D_A^2(z)}{10^{14} \text{Mpc}^2} \right] = Y_\star \left[ \frac{h}{0.7} \right]^{2 + \alpha} \left[ \frac{(1 - b) M_{500}}{3 \times 10^{14} M_\odot} \right]^\alpha, \quad (7)$$

where $$D_A(z)$$ is the angular-diameter distance to redshift $$z$$ and $$E^{-\beta}(z) = \Omega_m(1 + z)^2 + \Omega_\Lambda$$. The coefficients $$Y_\star$$, $$\alpha$$, and $$\beta$$ are given in Table 1.
We have implemented three independent versions of the computation of counts and constraints. The differences in predicted counts are of the order of a few percent, which translates to less than a tenth of $1\sigma$ on the cosmological parameters of interest.

### 3. The Planck cosmological samples

#### 3.1. Sample definition

The reference cosmological sample is constructed from the PSZ Catalogue published in Planck Collaboration XXIX (2014) and made public with the first release of Planck cosmological products. It is based on the SZ detections performed with the matched multi-filter (MMF) method MMF3 (Melin et al. 2006), which relies on the use of a filter of adjustable width $\theta_{500}$ chosen to maximize the S/N of the detection. In order to ensure a high purity and to maximize the number of redshifts, the cosmological sample was constructed by selecting the SZ detections above a S/N threshold of 7 outside Galactic and point source masks covering 35% of the sky, as discussed in Paper I. From the original PSZ, only the information on S/N (for the selection) and redshift are used.

This sample contains 189 candidates. All but one are confirmed bona fide clusters with measured redshifts, including 184 spectroscopic redshifts. Among these confirmed clusters 12 were confirmed with follow-up programmes conducted by the Planck collaboration (see Paper I for details). The remaining non-confirmed cluster candidate is a high-reliability CLASS candidate, meaning that its characterization as a cluster is supported by data in other wavebands (see Paper I for details). It is thus considered as a bona fide cluster. The distribution on the sky of this baseline cosmological sample is shown in Fig. 1.

In addition to our reference sample, we consider two other samples drawn from the PSZ for consistency checks. One is based on the detections from the second implementation of the MMF algorithm, MMF1, described in Paper I. It contains 188 clusters with S/N > 7 and no missing redshifts, with almost complete overlap with the baseline sample (187 clusters in common). The third sample considered in the present study is also based on MMF3 detections but with a higher S/N cut of S/N > 8. It allows us to test selection effects and to probe the consistency of the results as a function of the S/N cut. It contains 136 clusters, all with measured redshifts.

The selection function for each of these samples is constructed as described in the next section.

#### 3.2. Completeness

The completeness of the reference cosmological sample is computed with two distinct and complementary approaches: a semi-analytic approach based on the assumption of Gaussian uncertainties, and a computational approach based on Monte Carlo cluster injection into real sky maps.

The completeness $\chi$ can be evaluated analytically by setting the probability of the measured SZ flux, $Y_{500}$, to be Gaussian distributed with a standard deviation equal to the noise, $\sigma_{Y_{500}}(\theta_{500}, l, b)$, computed for each size $\theta_{500}$ of the MMF filter and at each position $(l, b)$ on the sky:

$$
\chi_{\text{err}}(Y_{500}, \theta_{500}, l, b) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{Y_{500} - X_{\text{err}}(\theta_{500}, l, b)}{\sqrt{2} \sigma_{Y_{500}}(\theta_{500}, l, b)} \right) \right],
$$

(10)
Fig. 4. Noise map $\sigma_{Y_{500}}(\theta_{500})$ for $\theta_{500} = 6$ arcmin. The PSZ is limited by instrumental noise at high ($|\theta| > 20^\circ$) Galactic latitude (deeper at ecliptic poles) and foreground noise at low Galactic latitude. The scale of the map ranges from 0.5 to 2 times the mean noise of the map, which is $\langle \sigma_{Y_{500}}(6 \text{ arcmin}) \rangle = 2.2 \times 10^{-4} \text{ arcmin}^2$.

where $X = 7$ is the S/N threshold and the error function is defined as usual by

$$\text{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u \exp(-t^2) \mathrm{d}t. \quad (11)$$

$X_{\text{erf}}(Y_{500}, \theta_{500}, l, b)$ thus lies in the range $[0, 1]$ and gives the probability for a cluster of flux $Y_{500}$ and size $\theta_{500}$ at position $(l, b)$ to be detected at $S/N \geq X$.

The noise estimate $\sigma_{Y_{500}}(\theta_{500}, l, b)$ is a by-product of the detection algorithm and can be written in the form (see e.g., Melin et al. 2006)

$$\sigma_{Y_{500}}(\theta_{500}, l, b) = \left[ \int \hat{d}^3k \, F_{\theta_{500}}(k) \cdot P^{-1}(k, l, b) \cdot F_{\theta_{500}}(k) \right]^{-1/2}, \quad (12)$$

with $F_{\theta_{500}}(k)$ being a vector of dimension $N_{\text{freq}}$ (the six highest Planck frequencies here) containing the beam-convolved cluster template scaled to the known SZ frequency dependence. The cluster template assumed is the non-standard universal pressure profile from Arnaud et al. (2010). $P(k, l, b)$ is the noise power spectrum (dimension $N_{\text{freq}} \times N_{\text{freq}}$) directly estimated from the data at position $(l, b)$. Figure 4 shows $\sigma_{Y_{500}}(\theta_{500}, l, b)$ for $\theta_{500} = 6$ arcmin in a Mollweide projection with the Galactic mask used in the analysis applied. As expected, the noise at high Galactic latitude is lower than in the areas contaminated by diffuse Galactic emission. The ecliptic pole regions have the lowest noise level, reflecting the longer Planck integration time in these high-redundancy areas.

The Monte Carlo (MC) completeness is calculated by injecting simulated clusters into real sky maps following the method presented in Paper I, with the modifications that the 65% Galactic dust mask and a $S/N > 7$ threshold are applied to match the cosmological sample definition. The Monte Carlo completeness encodes effects not probed by the erf approximation, including the variation of cluster pressure profiles around the fiducial pressure profile used in the MMF, spatially-varying and asymmetric effective beams, and the effects of correlated non-Gaussian uncertainties in the estimation of $(Y_{500}, \theta_{500})$. As shown in Fig. 5, the erf-based formula for the completeness is a good approximation to the Monte Carlo completeness. The agreement is best for the typical sizes probed by Planck (5 to 10 arcmin), though the two determinations of the completeness start to deviate for small and large sizes, due to beam and profile effects, respectively. For simplicity, we chose the erf formulation as the baseline. The effect of using the Monte Carlo completeness instead is discussed in Sect. 5.2.

4. Likelihood and Markov chain Monte Carlo

4.1. The likelihood

We now have all the information needed to predict the counts in redshift bins for our theoretical models. To obtain cosmological constraints with the PSZ sample presented in Sect. 3, we construct a likelihood function based on Poisson statistics (Cash 1979):

$$\ln L = \ln P(N_i|n_i) = \sum_{i=1}^{N_i} \left[ \ln(n_i) - n_i - \ln(N_i!) \right], \quad (13)$$

where $P(N_i|n_i)$ is the probability of finding $N_i$ clusters in each of $N_i$ bins given an expected number of $n_i$ in each bin in redshift. The likelihood includes bins that contain no observed clusters. As a baseline, we assume bins in redshift of $\Delta z = 0.1$ and we checked that our results are robust when changing the bin size between 0.05 and 0.2. The modelled expected number $n_i$ depends on the bin range in redshift and on the cosmological parameters, as described in Sect. 2. It also depends on the scaling relations and the selection function of the observed sample. The parameters of the scaling relations between flux (or size) and mass and redshift are taken to be Gaussian distributed with central values and uncertainties stated in Table 1, and with the scatter in $Y_{500}$ incorporated into the method via the log-normal distribution with width $\sigma_{Y_{500}}$. In the PSZ, the redshifts have been collected from different observations and from the literature. Individual uncertainties in redshift are thus spread between 0.001 and 0.1. Most
of the clusters in the cosmological sample have spectroscopic redshifts (184 out of 189) and we checked that the uncertainties in redshift are not at all dominant in our error budget; they are thus neglected. The cluster without known redshift is incorporated by scaling the counts by a factor 189/802 = 0.23; the cluster without known redshift is incor-

4.2. Markov chain Monte Carlo

In order to impose constraints on cosmological parameters from our sample(s) given our modelled expected number counts, we modified CosmoMC (Lewis & Bridle 2002) to include the likelihood described above. We mainly study constraints on the spatially-flat ΛCDM model, varying Ω_m, σ_8, Ω_b, H_0, and n_s, but also adding in the total neutrino mass, Σ_m_ν, in Sect. 6. In each of the runs, the nuisance parameters (Y_α, α, σ_logy) follow Gaussian priors, with the characteristics detailed in Table 1, and are marginalized over. The bias (1 − b) follows a flat prior in the range [0.7, 1]. The redshift evolution of the scaling, β, is fixed to its reference value unless stated otherwise.

4.3. External datasets

When probing the six parameters of the ΛCDM model, we combine the Planck clusters with the Big Bang nucleosynthesis (BBN) constraints from Steigman (2008), Ω_m h^2 = 0.022 ± 0.002. We also use either the H_0 determination from HST by Riess et al. (2011), H_0 = (73.8 ± 2.4) km s^{-1} Mpc^{-1}, or baryon acoustic oscillation (BAO) data. In the latter case we adopt the combined likelihood of Hinshaw et al. (2013) and Planck Collaboration XVI (2014), which uses the radial BAO scales observed by 6dFGRS (Beutler et al. 2011), SDSS-DR7-rec and SDSS-DR9-rec (Padmanabhan et al. 2012; Anderson et al. 2012), and WiggleZ (Blake et al. 2012).

5. Constraints from Planck clusters: ΛCDM

5.1. Results for Ω_m and σ_8

Cluster counts in redshift for our Planck cosmological sample are not sensitive to all parameters of the ΛCDM model. We focus first on (Ω_m, σ_8), assuming that n_s follows a Gaussian prior centred on the best-fit Planck CMB value4 (n_s = 0.9603 ± 0.0073). We combine our SZ counts likelihood with the BAO and BBN likelihoods discussed earlier. We also incorporate the uncertainties on scaling parameters in Table 1. Allowing the bias to range uniformly over the interval [0.7, 1.0], we find the expected degeneracy between the two parameters, σ_8(Ω_m/0.27)^{0.3} = 0.764±0.025,5 with central values and uncertainties of Ω_m = 0.29±0.02 and σ_8 = 0.75±0.03 (Table 2 and Fig. 6, red contours). The cluster counts as a function of redshift for the best-fit model are plotted in Fig. 7. When fixing the bias to (1 − b) = 0.8, the constraint on Ω_m remains unchanged while the constraint on σ_8 becomes stronger: σ_8(Ω_m/0.27)^{0.3} = 0.78 ± 0.01 and σ_8 = 0.77 ± 0.02 (Fig. 8).

To investigate how robust our results are when changing our priors, we repeat the analysis substituting the HST constraints on H_0 for the BAO results. Figure 6 (black contours) shows that the main effect is to change the best-fit value of H_0, leaving the (Ω_m, σ_8) degeneracy almost unchanged. In the following robustness tests, we assume a fixed mass bias, (1 − b) = 0.8, to better identify the effect of each of our assumptions.

5.2. Robustness to the observational sample

To test the robustness of our results, we performed the same analysis with different sub-samples drawn from our cosmological sample or from the PSZ, as described in Sect. 3, following that section’s discussion of completeness. Figure 9 shows the likelihood contours of the three samples (blue, MMF3 S/N > 8; red, MMF3 S/N > 7; black, MMF1 S/N > 7) in the (Ω_m, σ_8) plane. There is good agreement between the three samples. Obviously the three samples are not independent, as many clusters are common, but the noise estimates for MMF3 and MMF1 are different leading to different selection functions. Table 2 summarizes the best-fit values.

We perform the same analysis as on the baseline cosmological sample (SZ and BAO and BBN), but employing a different computation of the completeness function using the Monte Carlo method described in Sect. 3. Figure 9 shows the change in the 2D likelihoods when this alternative approach is adopted. The Monte Carlo estimation (in purple), being close to the analytic one, gives constraints that are similar, but enlarge the contour along the (Ω_m, σ_8) degeneracy.

5.3. Robustness to cluster modelling

A key ingredient in the modelling of the number counts is the mass function. Our main results adopt the Tinker et al. mass function as the reference model. We compare these to results from the Watson et al. mass function to evaluate the impact of uncertainty in predictions for the abundance of the most massive and extreme clusters. Figure 8 shows the 95% contours

---

4 Table 2 of Planck Collaboration XVI (2014).

5 We express it this way to ease comparison with other work.
Fig. 6. *Planck* SZ constraints (+BAO+BBN) on ΛCDM cosmological parameters in red. The black lines show the constraints upon substituting HST constraints on $H_0$ for the BAO constraints. Contours are 68 and 95% confidence levels.

Fig. 7. Distribution in redshift for the Planck cosmological cluster sample. The observed number counts (red), are compared to our best-fit model prediction (blue). The dashed and dot-dashed histograms are the best-fit models from the Planck SZ power spectrum and Planck CMB power spectrum fits, respectively. The cyan long dashed histogram is the best fit CMB and SZ when the bias is free (see Sect. 6.3). The uncertainties on the observed counts, shown for illustration only, are the standard deviation based on the observed counts, except for empty bins where we show the inferred 84% upper limit on the predicted counts assuming a Poissonian distribution. See Sect. 6 for more discussion.

Fig. 8. Comparison of the constraints using the mass functions of Watson et al. (black) and Tinker et al. (red), with $b$ fixed to 0.8. When relaxing the constraints on the evolution of the scaling law with redshift (blue), the contours move along the degeneracy line. Allowing the bias to vary uniformly in the range [0.7, 1.0] enlarges the constraints perpendicular to the $\sigma_8$–$\Omega_m$ degeneracy line due to the degeneracy of the number of clusters with the mass bias (purple). Contours are 95% confidence levels here.

when adopting the different mass functions. The main effect is to change the orientation of the degeneracy between $\Omega_m$ and $\sigma_8$, moving the best-fit values by less than 1σ.

We also relax the assumption of standard evolution of the scalings with redshift by allowing $\beta$ to vary with a Gaussian prior taken from Planck Collaboration X (2011), $\beta = 0.66 \pm 0.5$. Once again, the contours move along the $\sigma_8$–$\Omega_m$ degeneracy direction (shown in blue in Fig. 8).
Table 3. Constraints from clusters on $\sigma_8(\Omega_m/0.27)^{0.3}$.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>CPPP$^a$</th>
<th>MaxBCG$^b$</th>
<th>ACT$^c$</th>
<th>SPT</th>
<th>Planck SZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
<td>Vikhlinin et al. (2009b)</td>
<td>Rozo et al. (2010)</td>
<td>Hasselfield et al. (2013)</td>
<td>Reichardt et al. (2013)</td>
<td>This work</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>49+37</td>
<td>~13 000</td>
<td>15</td>
<td>100</td>
<td>189</td>
</tr>
<tr>
<td>Redshift range</td>
<td>[0.025, 0.25] and [0.35, 0.9]</td>
<td>[0.1, 0.3]</td>
<td>[0.2, 1.5]</td>
<td>[0.3, 1.35]</td>
<td>[0.0, 0.99]</td>
</tr>
<tr>
<td>Median mass (10^{14} h^{-1} M_{\odot})</td>
<td>2.5</td>
<td>1.5</td>
<td>3.2</td>
<td>3.3</td>
<td>6.0</td>
</tr>
<tr>
<td>S/N cut</td>
<td>5</td>
<td>(N_{\text{MC}} &gt; 11)</td>
<td>5</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Scaling</td>
<td>$Y_X - T_X$, $M_{\text{gas}}$</td>
<td>$N_{\text{MC}} - M_{\text{cool}}$</td>
<td>several</td>
<td>$L_X - M$, $Y_X$</td>
<td>$Y_{\text{GC}} - Y_X$</td>
</tr>
<tr>
<td>$\sigma_8(\Omega_m/0.27)^{0.3}$</td>
<td>0.784 ± 0.027</td>
<td>0.806 ± 0.033</td>
<td>0.768 ± 0.025</td>
<td>0.767 ± 0.037</td>
<td>0.764 ± 0.025</td>
</tr>
</tbody>
</table>

Notes. $^{(a)}$ The degeneracy is $\sigma_8(\Omega_m/0.27)^{0.37}$. $^{(b)}$ The degeneracy is $\sigma_8(\Omega_m/0.27)^{0.41}$. $^{(c)}$ For ACT we choose the results assuming the scaling law derived from the universal pressure profile in this table (constraints using other scaling relations are shown in Fig. 10).

Fig. 9. 95% contours for different robustness tests: MMF3 with S/N cut > 7 in red; MMF3 with S/N cut > 8 in blue; and MMF1 with S/N cut > 7 in black; and MMF3 with S/N cut at 7 but adopting the MC completeness in purple.

As shown in Appendix A, the estimation of the mass bias is not trivial and there is a large scatter amongst simulations. The red and purple contours compare the different constraints when fixing the mass bias to 0.8 and when allowing it to vary uniformly in the range [0.7, 1.0] respectively. Modelling of the cluster observable-mass relation is clearly the limiting factor in our analysis.

6. Discussion

Our main result is the constraint in the $(\Omega_m, \sigma_8)$ plane for the standard ΛCDM model imposed by the SZ counts, which we have shown is robust to the details of our modelling. We now compare this result first to constraints from other cluster samples, and then to the constraints from the Planck analysis of the sky-map of the Compton y-parameter (Planck Collaboration XXI 2014) and of the primary CMB temperature anisotropies (Planck Collaboration XVI 2014).

6.1. Comparison with other cluster constraints

We restrict our comparison to some recent analyses exploiting a range of observational techniques to obtain cluster samples and mass calibrations.

Benson et al. (2013) used 18 galaxy clusters in the first 178 deg^2 of the SPT survey to find $\sigma_8(\Omega_m/0.25)^{0.3} = 0.785 \pm 0.037$ for a spatially-flat model. They break the degeneracy between $\sigma_8$ and $\Omega_m$ by incorporating primary CMB constraints, deducing that $\sigma_8 = 0.795 \pm 0.016$ and $\Omega_m = 0.255 \pm 0.016$. In addition, they find that the dark energy equation of state is constrained to $w = -1.09 \pm 0.36$, using just their cluster sample along with the same HST and BBN constraints used here. Subsequently, Reichardt et al. (2013) reported a much larger cluster sample and used this to improve on the statistical uncertainties on the cosmological parameters (see Table 3). Hasselfield et al. (2013) use a sample of 15 high S/N clusters from ACT, in combination with primary CMB data, to find $\sigma_8 = 0.786 \pm 0.013$ and $\Omega_m = 0.250 \pm 0.012$ when assuming a scaling law derived from the universal pressure profile.

Strong constraints on cosmological parameters have been inferred from X-ray and optical richness selected samples. Vikhlinin et al. (2009b) used a sample of 86 well-studied X-ray clusters, split into low- and high-redshift bins, to conclude that $\Omega_m > 0$ with a significance about 5σ and that $w = -1.14 \pm 0.21$. Rozo et al. (2010) used the approximately 10^3 clusters in the Sloan Digital Sky Survey (SDSS) MaxBCG cluster sample, which are detected using a colour-magnitude technique and characterized by optical richness. They found that $\sigma_8(\Omega_m/0.25)^{0.41} = 0.832 \pm 0.033$. The fact that this uncertainty is similar to those quoted above for much smaller cluster samples signifies, once again, that cluster cosmology constraints are now limited by modelling, rather than statistical, uncertainties.

Table 3 and Fig. 10 show some current constraints on the combination $\sigma_8(\Omega_m/0.27)^{0.3}$, which is the main degeneracy line in cluster constraints. This comparison is only meant to be indicative, as a more quantitative comparison would require full consideration of modelling details which is beyond the scope of this work. Cosmic shear (Kilbinger et al. 2013), X-rays (Vikhlinin et al. 2009b), and MaxBCG (Rozo et al. 2010) each have a different slope in $\Omega_m$, being 0.6, 0.47, and 0.41, respectively (instead of 0.3), as they are probing different redshift ranges. We have rescaled when necessary the best value and errors to quote numbers with a pivot $\Omega_m = 0.27$. Hasselfield et al. (2013) have derived “cluster-only” constraints from ACT by adopting several different scaling laws, shown in blue and dashed blue in Fig. 10. The constraint assuming the universal pressure profile is highlighted as the solid symbol and error bar. For SPT we show the “cluster-only” constraints from Reichardt et al. (2013). For our own analysis we show our baseline result.
for SZ and BAO and BBN with a prior on \( 1 - b \) distributed uniformly in [0.7, 1]. The figure thus demonstrates good agreement amongst all cluster observations, whether in optical, X-rays, or SZ. Table 3 compares the different data and assumptions of the different cluster-related publications.

6.2. Consistency with the Planck y-map

In a companion paper (Planck Collaboration XXI 2014), we performed an analysis of the SZ angular power spectrum derived from the Planck y-map obtained with a dedicated component-separation technique. For the first time, the power spectrum has been measured at intermediate scales (50 \( \leq \ell \leq 1000 \)). The same modelling as in Sect. 2 and Taburet et al. (2009, 2010) has been used to derive best-fit values of \( \Omega_m \) and \( \sigma_8 \), assuming the universal pressure profile (Arnaud et al. 2010), a bias \( 1 - b = 0.8 \), and the best-fit values for other cosmological parameters from Planck Collaboration XVI (2014)\(^6\). The best model obtained, shown in Fig. 7 as the dashed line, demonstrates the consistency between the PSZ number counts and the signal observed in the y-map.

6.3. Comparison with Planck primary CMB constraints

We now compare the PSZ cluster constraints to those from the analysis of the primary CMB temperature anisotropies given in Planck Collaboration XVI (2014) (see Footnote 6). In that analysis \( \sigma_8 \) is derived from the standard six \( \Lambda \)CDM parameters.

The Planck primary CMB constraints, in the \( (\Omega_m, \sigma_8) \) plane, differ significantly from our own, in particular through favouring a higher value of \( \sigma_8 \), (see Fig. 11). For \( 1 - b = 0.8 \), this leads to a factor of 2 larger number of predicted clusters than is actually observed (see Fig. 7). There is therefore some tension between the results from the Planck CMB analysis and the current cluster analysis. Figure 10 illustrates this with a comparison of three analyses of primary CMB data alone (Planck Collaboration XVI 2014; Story et al. 2013; Hinshaw et al. 2013) and cluster constraints in terms of \( \sigma_8(\Omega_m(0.27)^{0.3}) \).

It is possible that the tension results from a combination of some residual systematics with a substantial statistical fluctuation. Enough tests and comparisons have been made on the Planck data sets that it is plausible that at least one discrepancy at the two to three sigma level will arise by chance. Nevertheless, it is worth considering the implications if the discrepancy is real.

As we have noted, the modelling of the cluster gas physics is the most important uncertainty in our analysis, in particular through its impact on the mass bias \( 1 - b \). While we have argued for a preferred value of \( 1 - b \approx 0.8 \) based on comparison of our \( Y_{\text{SZ}}-M_{\text{500}} \) relation to those derived from a number of different numerical simulations, and we suggest a plausible range of \( 1 - b \) from 0.7 to 1, a significantly lower value would substantially alleviate the tension between CMB and SZ constraints. We have undertaken a joint analysis using the CMB likelihood presented in Planck Collaboration XV (2014) and the cluster likelihood presented in the present paper, sampling \( 1 - b \) in the range [0.1, 1.5]. This results in a “measurement” of \( 1 - b = 0.59 \pm 0.05 \). We show in Fig 7 the SZ cluster counts predicted by the Planck’s best-fit primary CMB model for \( 1 - b = 0.59 \). Clearly, this substantial reduction in \( 1 - b \) is enough to reconcile our observed SZ cluster counts with Planck’s best-fit primary CMB parameters.

Such a large bias is difficult to reconcile with numerical simulations, and cluster masses estimated from X-rays and from weak lensing do not typically show such large offsets (see Appendix A). Systematic discrepancies in the relevant scaling relations have, however, been identified and studied in stacking analyses of X-ray, SZ, and lensing data for the very large MaxBCG cluster sample, e.g., Planck Collaboration XII (2011), Biesiadzinski et al. (2012), Draper et al. (2012), Rozo et al. (2012), and Sehgal et al. (2013), suggesting that the issue is not yet fully settled from an observational point of view. The uncertainty reflects the inherent biases of the different mass estimates. Systematic effects arising from instrument calibration constitute

\[ \text{Fig. 10. Comparison of constraints (68% confidence interval) on } \sigma_8(\Omega_m(0.27)^{0.3}) \text{ from different experiments of large-scale structure (LSS), clusters, and CMB. The solid line ACT point assumes the same universal pressure profile as this work. Probes marked with an original power of } \Omega_m \text{ different from 0.3. See text and Table 3 for more details.} \]

\[ \text{Fig. 11. 2D } \Omega_m-\sigma_8 \text{ likelihood contours for the analysis with Planck CMB only (red); Planck SZ and BAO and BBN (blue) with (1 - b) in [0.7, 1].} \]
a further source of uncertainty – in X-ray mass determinations, temperature estimates represent the main source of systematic uncertainty in mass, as the mass scales roughly with $T^{3/2}$. Other biases in the determination of mass-observable scaling relations come from the object selection process itself (e.g., Mantz et al. 2010; Allen et al. 2011b; Angulo et al. 2012). This may be less of a concern for SZ selected samples because of the expected small scatter between the measured quantity and the mass. An improbable conspiracy of all sources of bias seems required to lead to a sufficiently-low effective value of $(1-b)$ to reconcile the SZ and CMB constraints. This possibility needs to be carefully examined with probes based on a variety of physical quantities and derived from a wide range of types of observation, including masses, baryon and gas fractions, etc.

A different mass function may also help reduce the tension. Mass functions are calibrated against numerical simulations that may still suffer from volume effects for the largest haloes, as shown in the difference between the Tinker et al. (2008) and Watson et al. (2013) mass functions. This does not seem sufficient, however, given the results presented in Fig. 8.

One might instead ask whether the Planck data analysis could somehow have missed a non-negligible fraction of the total number of clusters currently predicted to have $S/N > 7$, resulting in a lower observed number count distribution. This is linked to a possible underestimate of the true dispersion about the $Y_{500}$–$Y_X$ relation at a given $M_{500}$. It would be necessary for Planck to have missed ~40 percent of the clusters with predicted SZ $S/N > 7$ in order for the SZ and CMB number count curves in Fig. 7 to be in agreement. Increasing the dispersion about the $Y_{500}$–$M_{500}$ relation and allowing it to correlate strongly with the scatter in X-ray properties (in particular, $Y_X$) would raise the possibility that our calibration procedures (which are based on X-ray and SZ selected clusters assuming the scatter in $Y$ and $Y_X$ at fixed $M_{500}$ to be small and to be uncorrelated with cluster dynamical state) might produce a relation which is biased high. A sufficiently-large effect seems, however, to require a level of scatter and a degree of correlation with cluster structure which are inconsistent with the predictions of current hydrodynamical simulations (see the discussion in Appendix A).

Alternatively, the discrepancy may reflect a need to extend the minimal ΛCDM model in which the $\sigma_8$ constraints are derived from the primary CMB analysis. Any extension would need to modify the power spectrum on the scales probed by clusters, while leaving the scales probed by primary CMB observations unaffected. The inclusion of neutrino masses, quantified by their sum over all families, $\sum m_\nu$, can achieve this (see Marulli et al. 2011 and Burenin 2013 for reviews of how cosmological observations can be affected by the inclusion of neutrino masses). The SPT collaboration (Hou et al. 2014) recently considered such a possibility to mitigate their tension with WMAP-7 primary CMB data. There is an upper limit of $\sum m_\nu < 0.93\, eV$ from the Planck primary CMB data alone\textsuperscript{7}. If we combine the Planck CMB (Planck+WP) likelihood and the cluster count data using a fixed value $(1-b) = 0.8$, then we find a $2.8\sigma$ preference for the inclusion of neutrino masses with $\sum m_\nu = (0.53 \pm 0.19)\, eV$, as shown in Fig. 12. If, on the other hand, we adopt a more conservative point of view and allow $(1-b)$ to vary between 0.7 and 1.0, this preference drops to $1.9\sigma$ with $\sum m_\nu = (0.40 \pm 0.21)\, eV$. Adding BAO data to the compilation lowers the value of the required mass but increases the significance, yielding $\sum m_\nu = (0.20 \pm 0.09)\, eV$, due to a breaking of the degeneracy between $H_0$ and $\sum m_\nu$.

\textsuperscript{7} Planck Collaboration XVI (2014), Table 10, Col. 3.

As these results depend on the value and allowed range of $(1-b)$, better understanding of the scaling relation is the key to further investigation. This provides strong motivation for further study of the relationship between $Y$ and $M$. Over the past few years we have moved into an era where systematic uncertainties dominate to an increasing extent over statistical uncertainties. Observed mass estimates using different methods are of improving quality; for instance, X-ray mass proxies can be measured to better than 10% precision (e.g., Vikhlinin et al. 2006; Arnaud et al. 2007). In this context, systematic calibration uncertainties are playing an increasingly important role when using the cluster population to constrain cosmology.

7. Summary

We have used a sample of nearly 200 clusters from the PSZ, along with the corresponding selection function, to place strong constraints in the $(\Omega_m, \sigma_8)$ plane. We have carried out a series of tests to verify the robustness of our constraints, varying the observed sample choice, the estimation method for the selection function, and the theoretical methodology, and have found that our results are not altered significantly by those changes.

The relation between the mass and the integrated SZ signal plays a major role in the computation of the expected number counts. Uncertainties in cosmological constraints from clusters are no longer dominated by small number statistics, but by the gas physics and sample selection biases. Uncertainties in the $Y–M$ relation include contributions from X-ray instrument calibration, X-ray temperature measurements, inhomogeneities in cluster density and temperature profiles, and selection effects. Considering several ingredients of the gas physics of clusters, numerical simulations predict scaling relations with 30% scatter in amplitude (at a fiducial mass of $6 \times 10^{14} \, M_\odot$), and suggest a mass bias between the true mass and the estimated mass...
of \((1 - b) = 0.8^{+0.2}_{-0.1}\). Adopting the central value we found constraints on \(\Omega_m\) and \(\sigma_8\) that are in good agreement with previous measurements using clusters of galaxies.

Comparing our results with Planck primary CMB constraints within the ΛCDM cosmology reveals some tension. This can be alleviated by permitting a large mass bias \((1 - b = 0.60)\), which is however significantly larger than expected. Alternatively, the tension may indicate a need for an extension of the base ΛCDM model that modifies its power spectrum shape. For example the inclusion of non-zero neutrino masses helps in reconciling the primary CMB and cluster constraints, a fit to Planck CMB and SZ and BAO yielding \(\sum m_\nu = (0.20 \pm 0.09)\) eV.

Cosmological parameter determination using clusters is currently limited by the knowledge of the observable–mass relations. In the future our goal is to increase the number of dedicated follow-up programmes to obtain better estimates of the mass proxy and redshift for most of the \(S/N > 5\) Planck clusters. This will allow for improved determination of the scaling laws and the mass bias, increase the number of clusters that can be used, and allow us to investigate an extended cosmological parameter space.

Acknowledgements. The development of Planck has been supported by: ESA; CNES and CNRS/INSU/IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSCIC: MICEAN, JA and RES (Spain); Tekes, AoF and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); and PRACE (EU). A description of the Planck Collaboration and a list of its members, including the technical or scientific activities in which they have been involved, can be found at http://www.sciops.esa.int/index.php?project=planck&page=Planck_Collaboration

Appendix A: Calibration of the \(Y_{500}-M_{500}\) relation

A cluster catalogue is a list of positions and measurements of observable physical quantities. Its scientific utility depends largely on our ability to link the observed quantities to the underlying mass, in other words, to define an observable proxy for the mass. Planck detects clusters through the SZ effect. This effect is currently the subject of much study in the cluster community, chiefly because numerical simulations indicate that the spherically-integrated SZ measurement is correlated particularly tightly with the underlying mass. In other words, this measurement potentially represents a particularly valuable mass proxy.

To establish a mass proxy, one obviously needs an accurate measurement both of the total mass and of the observable quantity in question. However, even with highly accurate measurements, the correlation between the observable quantity and the mass is susceptible to bias and dispersion, and both of these effects need to be taken into account when using cluster catalogues for scientific applications.

The aim of this Appendix is to define a baseline relation between the measured SZ flux, \(Y_{500}\), and the total mass \(M_{500}\). The latter quantity is not directly measurable. On an individual cluster basis, it can be inferred from dynamical analysis of galaxies, from X-ray analysis assuming hydrostatic equilibrium (HE), or from gravitational lensing observations. However, it is important to note that all observed mass estimates include inherent biases. For instance, numerical simulations suggest that HE mass measurements are likely to underestimate the true mass by 10–15 percent due to neglect of bulk motions and turbulence in the intra-cluster medium (ICM, e.g., Nagai et al. 2007; Piffaretti & Valdarnini 2008; Meneghetti et al. 2010), an effect that is commonly referred to in the literature as the “hydrostatic mass bias”. Similarly, simulations indicate that weak lensing mass measurements may underestimate the mass by 5 to 10 percent, owing to projection effects or the use of inappropriate mass models (e.g., Becker & Kravtsov 2011). Instrument calibration systematic effects constitute a further source of error. For X-ray mass determinations, temperature estimates represent the main source of systematic uncertainty, as the mass at a given density contrast scales roughly with \(T^{3/2}\). Other biases in the determination of mass-observable scaling relations come from the object selection process itself (e.g., Allen et al. 2011b; Angulo et al. 2012). A classic example is the Malmquist bias, where bright objects near the flux limit are preferentially detected. This effect is amplified by Eddington bias, the mass function dictating that many more low-mass objects are detected compared to high-mass objects. Both of these biases depend critically on the distribution of objects in mass and redshift, and on the dispersion in the relation between the mass and the observable used for sample selection. This is less of a concern for SZ selected samples than for X-ray selected samples, the SZ signal having much less scatter at a given mass than the X-ray luminosity. However for precise studies it should still be taken into account.

On the theoretical side, numerous \(Y_{500}\)–\(M_{500}\) relations have been derived from simulated data, as discussed below. The obvious advantage of using simulated data is that the relation between the SZ signal and the true mass can be obtained, because the “real” value of all physical quantities can be measured. The disadvantage is that the “real” values of measurable physical quantities depend strongly on the phenomenological models used to describe the different non-gravitational processes at work in the ICM.

Nevertheless, the magnitude of the bias between observed and true quantities can only be assessed by comparing multi-wavelength observations of a well-controlled cluster sample to numerical simulations. Thus, ideally, we would have full follow-up of a complete Planck cluster sample. For large samples, however, full follow-up is costly and time consuming. This has led to the widespread use of mass estimates obtained from mass-proxy relations. These relations are generally calibrated from individual deep observations of a subset of the sample in question (e.g., Vikhlinin et al. 2009a), or from deep observations of objects from an external dataset (e.g., the use of the REXCESS relations in Planck Collaboration XI 2011).

For the present paper, we will rely on mass estimates from a mass-proxy relation. In this context, the \(M_{500}\)–\(Y_{X}\) relation is clearly the best to use. \(Y_{X}\), proposed by Kravtsov et al. (2006), is defined as the product of \(M_{500}\), the gas mass within \(R_{500}\), and \(T_X\), the spectroscopic temperature measured in the [0.15–0.75] \(R_{500}\) aperture. In the simulations performed by Kravtsov et al. (2006), \(Y_X\) was extremely tightly correlated with the true cluster mass, with a logarithmic dispersion of only 8 percent. Observations using masses derived from X-ray hydrostatic analysis indicate that \(Y_X\) does indeed appear to have a low dispersion (Arnaud et al. 2007; Vikhlinin et al. 2009a). Furthermore, the local \(M_{500}\)–\(Y_{X}\) relation for X-ray selected relaxed clusters has been calibrated to high statistical precision (Arnaud et al. 2010; Vikhlinin et al. 2009a), with excellent agreement achieved between various observations (see e.g., Arnaud et al. 2007). Since simulations suggest that the \(Y_{500}\)–\(M_{500}\) relation is independent of dynamical state, calibrating the \(Y_{500}\)–\(M_{500}\) relation via a low-scatter mass proxy, itself calibrated on clusters for which the HE bias is expected to be minimal, is a better approach than using HE mass estimates for the full sample, since the latter can be highly biased for very unrelaxed objects.
We approach the determination of the $Y_{500} - M_{500}$ relation in two steps. We first calibrate the $Y_{500}$–proxy relation. This is combined with the X-ray calibrated relation, between the proxy and $M_{500}$, to define an observation-based $Y_{500} - M_{500}$ relation. In the second step, we assess possible biases in the relation by directly comparing the observation-based relation with that from simulations. This approach, rather than directly assessing the HE mass bias, allows us to avoid complications linked to the strong dependence of the HE bias on cluster dynamical state, and thus on the cluster sample (real or simulated). The final output from this procedure is a relation between $Y_{500}$ and $M_{500}$, with a full accounting of the different statistical and systematic uncertainties that go into its derivation, including bias.

In the following, all relations are fit with a power law in log-space using the orthogonal BCES method (Akritas & Bershady 1996), which takes into account the uncertainties in both variables and the intrinsic scatter. All dispersions are given in $\log_{10}$.

### A.1. Baseline mass-proxy relation

As a baseline, we use the relation between $Y_{X}$ and the X-ray hydrostatic mass $M_{500}^{\text{HE}}$, established for 20 local relaxed clusters by Arnaud et al. (2010):

$$E^{-2/3}(z) \left[ \frac{Y_{X}}{2 \times 10^{14} M_{\odot} \text{keV}} \right] = 10^{0.376 \pm 0.018} \left[ \frac{M_{500}^{\text{HE}}}{6 \times 10^{14} M_{\odot}} \right]^{1.78 \pm 0.06},$$

where $Y_{X}$ is the observed temperature and gas mass profile, as described in Planck Collaboration XI (2011). The re-extraction procedure uses matched multi-filters (MMF) and assumes that the ICM pressure follows the universal profile shape derived by Arnaud et al. (2010) from the combination of the REXCESS sample with simulations. The extraction is undertaken on the 15.5 month Planck survey data set, and so statistical precision on the $Y_{X}$–$M_{500}$ slope and normalization of the relation, as well as intrinsic uncertainties to the X-ray temperature and the gas mass profile.

The SZ signal is estimated within a sphere of radius $R_{\text{500}}$, centred on the position of the X-ray peak, as detailed in e.g., Planck Collaboration XI (2011). The $Y_{500}$–proxy relation is thus

$$E^{-2/3}(z) Y_{SX} = 10^{4.4 \pm 0.1} (1 - b) M_{500},$$

where $b$ is the bias between the true mass and the observed mass used to calibrate the relation. The bias is a poorly-known stochastic variable with substantial variation expected between clusters. In our case, $b$ represents the mean bias between the observed mass and the true mass.

The mass proxy $M_{500}^{Y_{SX}}$ is defined from the best-fit $Y_{X} - M_{500}^{\text{HE}}$ relation

$$E^{-2/3}(z) Y_{SX} = 10^{4} \left[ M_{500}^{Y_{SX}} \right]^{a}.$$  

For any cluster, $M_{500}^{Y_{SX}}$, together with the corresponding $Y_{X}$ and $R_{500}^{Y_{SX}}$, can be estimated iteratively from this relation from the observed temperature and gas mass profile, as described in Kravtsov et al. (2006). The calibration of the $Y_{X} - M_{500}$ relation is thus equivalent to a calibration of the $M_{500}^{Y_{SX}} - M_{500}$ relation, which relates the mass proxy $M_{500}^{Y_{SX}}$ to the mass via

$$M_{500}^{Y_{SX}} = 10^{0.175 \pm 0.011} (1 - b) M_{500} \left[ \frac{1}{6 \times 10^{14} M_{\odot}} \right]^{1.77 \pm 0.06}.$$  

In addition to the bias factor, there are statistical uncertainties on the slope and normalization of the relation, as well as intrinsic scatter around the relation, linked to the corresponding statistical uncertainties and scatter of the $Y_{X} - M_{500}^{\text{HE}}$ relation.

### A.2. Relation between $Y_{500}$ and $M_{500}^{Y_{X}}$

#### A.2.1. Best-fit relation

We first investigate the relationship between $Y_{500}$ and $M_{500}^{Y_{X}}$, the mass estimated iteratively from Eq. A.4, with parameters given by the best-fit Arnaud et al. (2010) relation (Eq. (A.1)). Full X-ray follow-up of the Planck SZ cosmological cluster sample is not yet available. Our baseline sample is thus a subset of 71 detections from the Planck cosmological cluster sample, detected at $S/N > 7$, for which good quality $\text{XMM-Newton}$ observations are available. The sample consists of data from our previous archival study of the Planck Early SZ (ESZ) clusters (Planck Collaboration XI 2011), of Planck-detected LoCuSS clusters presented by Planck Collaboration Int. III (2013), and from the $\text{XMM-Newton}$ validation programme (Planck Collaboration IX 2011; Planck Collaboration Int. I 2012; Planck Collaboration Int. IV 2013). The corresponding sub-samples include 58, 4, and 9 clusters, respectively. The X-ray data were re-analysed in order to have a homogeneous data set; measurement differences are negligible with respect to previously-published values. Uncertainties on $Y_{X}$, $R_{500}^{Y_{SX}}$, and $M_{500}^{Y_{SX}}$ include those due to statistical errors on the X-ray temperature and the gas mass profile.

The relation resulting for these 71 clusters from the cosmological sample is

$$E^{-2/3}(z) \left[ \frac{D_{A} Y_{SX}}{10^{4} \text{Mpc}^{2}} \right] = 10^{0.175 \pm 0.011} \left[ \frac{M_{500}^{Y_{SX}}}{6 \times 10^{14} M_{\odot}} \right]^{1.77 \pm 0.06}.$$  

This agrees within 1$\sigma$ with the results from the sample of 62 clusters from the ESZ sample with archival $\text{XMM-Newton}$ data published in Planck Collaboration XI (2011). The slope and normalization are determined at slightly higher precision, due to the better quality SZ data. The derived intrinsic scatter (Table A.1) is significantly smaller. This is a consequence of: a more robust treatment of statistical uncertainties; propagation of gas mass profile uncertainties in the $Y_{X}$ error budget; and, to a lesser extent, the propagation of $R_{500}^{Y_{SX}}$ uncertainties to $Y_{500}$ estimates.

#### A.2.2. Effects of Malmquist bias

The fitted parameters are potentially subject to selection effects such as Malmquist bias, owing to part of the sample lying close to the selection cut. For the present sample, we use an approach...
Table A.1. Parameters for the $Y_{500}$–$M_{500}$ relation, expressed as $E^{-2/3}(z) \left[ D_x^2 Y_{500} / 10^{-4} \text{Mpc}^2 \right] = 10^\alpha \left[ M_{500} / 6 \times 10^{14} \, M_\odot \right]^{\beta}$.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$N_c$</th>
<th>MB</th>
<th>Mass</th>
<th>$A$</th>
<th>$\alpha$</th>
<th>$[\sigma_{500\text{YM}}] \text{ int}$</th>
<th>$[\sigma_{500\text{YM}}] \text{ raw}$</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>XMM-ESZ PEPXI</td>
<td>62</td>
<td>N</td>
<td>$M_{500}^{\text{X}}$</td>
<td>$-0.19 \pm 0.01$</td>
<td>$1.74 \pm 0.08$</td>
<td>$0.10 \pm 0.01$</td>
<td>...</td>
<td>A.2.1</td>
</tr>
<tr>
<td>Cosmo sample</td>
<td>71</td>
<td>N</td>
<td>$M_{500}^{\text{X}}$</td>
<td>$-0.175 \pm 0.011$</td>
<td>$1.77 \pm 0.06$</td>
<td>$0.065 \pm 0.010$</td>
<td>$0.080 \pm 0.009$</td>
<td>A.2.1</td>
</tr>
<tr>
<td>Cosmo sample</td>
<td>71</td>
<td>Y</td>
<td>$M_{500}^{\text{X}}$</td>
<td>$-0.186 \pm 0.011$</td>
<td>$1.79 \pm 0.06$</td>
<td>$0.063 \pm 0.011$</td>
<td>$0.079 \pm 0.009$</td>
<td>A.2.2</td>
</tr>
<tr>
<td>XMM-ESZ</td>
<td>62</td>
<td>Y</td>
<td>$M_{500}^{\text{X}}$</td>
<td>$-0.19 \pm 0.01$</td>
<td>$1.75 \pm 0.07$</td>
<td>$0.065 \pm 0.011$</td>
<td>$0.079 \pm 0.009$</td>
<td>A.2.3</td>
</tr>
<tr>
<td>S/N &gt; 7</td>
<td>78</td>
<td>Y</td>
<td>$M_{500}^{\text{X}}$</td>
<td>$-0.18 \pm 0.01$</td>
<td>$1.72 \pm 0.06$</td>
<td>$0.063 \pm 0.010$</td>
<td>$0.078 \pm 0.008$</td>
<td>A.2.3</td>
</tr>
<tr>
<td>Cosmo sub-sample A</td>
<td>10</td>
<td>Y</td>
<td>$M_{500}^{\text{HE}}$</td>
<td>$-0.15 \pm 0.04$</td>
<td>$1.6 \pm 0.3$</td>
<td>...</td>
<td>$0.08 \pm 0.02$</td>
<td>A.3.2</td>
</tr>
<tr>
<td>Cosmo sub-sample B</td>
<td>58</td>
<td>Y</td>
<td>$M_{500}^{\text{HE}}$</td>
<td>$-0.19 \pm 0.03$</td>
<td>$1.7 \pm 0.2$</td>
<td>$0.25 \pm 0.06$</td>
<td>$0.27 \pm 0.06$</td>
<td>A.3.2</td>
</tr>
</tbody>
</table>

Notes. Column 1, considered sample; Col. 2, number of clusters in the sample; Col. 3, Malmquist bias correction; if this column contains Y, a mean correction for Malmquist bias has been applied to each point before fitting; Col. 4, mass definition; Cols. 5 and 6, slope and normalization of the relation; Cols. 7 and 8, intrinsic and raw orthogonal scatter around the best-fit relation at a given mass; Col. 9, Section in which sub-sample is discussed. The Cosmo sample highlighted in bold represents the baseline relation (see text for details).

The correction decreases the effective $Y_{500}$ values at a given mass, an effect larger for clusters closer to the S/N threshold. The net effect is small, a roughly 1σ decrease of the normalization and a slight steepening of the power-law slope (Table A.1).

A.2.3. Stability of slope and normalization

The slope and normalization of this relation are robust to the choice of sample (Table A.1). We compared our results to those obtained from:

- An extended sample of 78 clusters with S/N > 7 (71 in common with the baseline sample). This is built from all objects falling in the 84% sky mask used to define the SZ catalogue Planck Collaboration XXIX (2014), and for which XMM-Newton data have been published by the Planck Collaboration (Planck Collaboration IX 2011; Planck Collaboration XI 2011; Planck Collaboration Int. I 2012; Planck Collaboration Int. III 2013; Planck Collaboration Int. IV 2013).
- The original 62 clusters from the ESZ sample published in Planck Collaboration XI (2011), with updated SZ signal measurements obtained from 15.5 month Planck data (62 in common with the baseline sample). These objects are all known from X-ray surveys and all lie at $z < 0.5$. We use them to test fit robustness to the inclusion of non-X-ray selected, higher-redshift systems.

As indicated in Table A.1, there is agreement within 1σ between the various samples. The results are also in agreement with the relation obtained from a simple combination of the $Y_{500}$–$Y_{X}$ relation (discussed in Paper I) and the $Y_{X}$–$M_{500}^{\text{HE}}$ relation (Eq. (A.1) above).

A.3. The observation-based $Y_{500}$–$M_{500}$ relation

A.3.1. Combination of the $Y_{500}$–$M_{500}^{\text{X}}$ and the $M_{500}^{\text{HE}}$–$M_{500}$ relations

We now combine Eq. (A.7) with the $M_{500}^{\text{X}}$–$M_{500}$ relation. This will not change the best-fit parameters, but will increase their uncertainties. As the determinations of the two relations are independent, we added quadratically the uncertainties in the...
best-fit parameters of the \(Y_{500} - M_{500}^{\text{HE}}\) (Eq. (A.6)) and \(M_{500}^{\text{abs}} - M_{500}^{\text{HE}}\) (Eq. (A.5), with values from Eq. (A.1)) relations. Our best-fit \(Y_{500} - M_{500}^{\text{HE}}\) relation is then

\[
E^{-2/3}(c) \left[ \frac{D_{\Lambda}^2 Y_{500}}{10^{15} \text{ Mpc}^2} \right] = 10^{-0.19 \pm 0.02} \left( \frac{(1 - b) M_{500}}{6 \times 10^{14} M_\odot} \right)^{1.79 \pm 0.08}. \tag{A.8}
\]

Thus inclusion of the statistical uncertainty in the \(M_{500}^{\text{HE}} - M_{500}^{\text{abs}}\) relation doubles the uncertainty on the normalization and increases the uncertainty on the slope by 40%. Note that we have implicitly assumed here that the scatter around the two relations is uncorrelated.

### A.3.2. Effect of use of an external dataset

The above results assume a mass estimated from the baseline \(Y_{500} - M_{500}\) relation, derived by Arnaud et al. (2010) from an external dataset of 20 relaxed clusters (Eq. (A.1)). How does this relation compare to the individual hydrostatic X-ray masses of the Planck cosmological cluster sample? Of the 71 clusters in the baseline sample:

- 58 objects have temperature profile information extending to various fractions of \(R_{500}\), of which
- 10 cool-core objects have temperature profile measurements at least out to \(R_{500}\).

Thus, while spatially-resolved temperature profiles are available for 58 of the 71 clusters with XMM-Newton observations, we must be careful in interpretation of these data. The Arnaud et al. relation was derived from a carefully chosen data set consisting of relaxed, cool-core objects having well-constrained temperature profiles out to around \(R_{500}\), i.e., the type of object for which it makes sense to undertake a hydrostatic mass analysis. Many clusters of the Planck sample are merging systems for which such an analysis would give results that are difficult to interpret. In addition, few of the Planck sample have spatially-resolved temperature profiles out to \(R_{500}\). However, as given in Table A.1, the best-fit \(Y_{500} - M_{500}^{\text{HE}}\) relation for the 10 cool-core clusters that are detected to \(R_{500}\) agrees with Eq. (A.8) within 1σ. Moreover, the relation for the 58 Planck clusters with HE mass estimates, derived regardless of dynamical state and radial detection extent, also agrees within 1σ (albeit with greatly increased scatter). We are thus confident that the masses estimated from an externally-calibrated \(Y_{500} - M_{500}^{\text{HE}}\) relation are applicable to the present data set.

### A.3.3. Dispersion about the observed relations

A key issue is the dispersion around the mean relation. We first estimate the intrinsic scatter of the \(Y_{500} - M_{500}^{\text{HE}}\) relation by combining the intrinsic scatter of the \(Y_{500} - M_{500}^{\text{abs}}\) relation and that of the \(M_{500}^{\text{HE}} - M_{500}^{\text{HE}}\) relation. This estimate is applicable to relaxed objects only, since the \(Y_{500} - M_{500}^{\text{HE}}\) relation has been measured using a sample of such systems. If the scatter about the input relations is independent, this gives

\[
\sigma^2 = \sigma_{Y_{500}M_{500}^{\text{HE}}}^2 + 2\cos^2(\tan^{-1} \alpha) \sigma_{M_{500}^{\text{HE}}Y_{500}}^2, \tag{A.9}
\]

where \(\alpha\) is the slope of the \(Y_{500} - M_{500}^{\text{HE}}\) relation. As the HE mass estimate introduces extra scatter as compared to the true mass (Kay et al. 2012), the dispersion about the \(Y_{500} - M_{500}^{\text{HE}}\) relation is expected to be smaller than that of the \(Y_{500} - M_{500}^{\text{HE}}\) relation (although this will depend on correlations between the scatter in the \(M_{500}^{\text{HE}} - M_{500}^{\text{HE}}\) and \(M_{500}^{\text{HE}} - Y_{500}\) relations). The above expression thus also provides an estimate of the scatter of the \(Y_{500} - M_{500}^{\text{HE}}\) relation, again for relaxed objects. While merging events are expected to induce shocks in the ICM, leading to higher temperatures and thus an increase in \(Y_{500}\), current simulations suggest that this is a weak effect. This may be due to the relatively short duration of the shocking phase during a merger (e.g., Ricker & Sarazin 2001; Ritchie & Thomas 2002; Poole et al. 2007). Thus, further assuming that the intrinsic scatter of the \(Y_{500} - M_{500}^{\text{HE}}\) relation is the same for the total relaxed and unrelaxed population, as indicated by numerical simulations (Kravtsov et al. 2006; Kay et al. 2012), Eq. (A.9) gives a conservative estimate of the intrinsic scatter of the \(Y_{500} - M_{500}^{\text{HE}}\) relation.

The intrinsic dispersion about our baseline \(Y_{500} - M_{500}^{\text{HE}}\) relation (Eq. (A.1)), taken from Arnaud et al. (2010), is not measurable; neither is it measurable for the best-fit Chandra \(Y_{500} - M_{500}^{\text{HE}}\) relation published in Vikhlinin et al. (2009a). Using a smaller sample of 10 systems, Arnaud et al. (2007) measured an intrinsic scatter of \(\sigma_{\log M_{500}^{\text{HE}}Y_{500}} = 0.039\) (9 percent), in excellent agreement with the results of the simulations of Nagai et al. (2007) for the scatter of the \(M_{500}^{\text{HE}} - Y_{500}\) relation for relaxed clusters (8.7 percent, their Table 4). It is somewhat larger than the intrinsic scatter of the relation between the true mass and \(Y_{500}\) derived by Kravtsov et al. (\(\sigma_{\log M_{500}^{\text{HE}}Y_{500}} = 5\pm7\) percent) but close to the results of Fabjan et al. (2011), who find \(\sigma_{\log M_{500}^{\text{HE}}Y_{500}} = 0.036 \pm 0.046\). We thus take as a conservative estimate \(\sigma_{\log M_{500}^{\text{HE}}Y_{500}} = 0.05\). The intrinsic dispersion about the \(Y_{500} - M_{500}^{\text{HE}}\) relation for our data is \(\sigma_{\log Y_{500}M_{500}^{\text{HE}}} = 0.065 \pm 0.01\). This value is three times larger than the results of Kay et al. (2012). Partly this is due to the presence of outliers in our dataset (as discussed in Paper I), and it may also be due to projection effects in observed data sets (Kay et al. 2012).

Our final observational estimate of the intrinsic scatter is then \(\sigma_{\log Y_{500}M_{500}^{\text{HE}}} < 0.074\) or 18 percent, similar to the predictions from Kay et al. (2012) and Sehgal et al. (2010). These predictions depend both on the numerical scheme and specific physics assumptions, with values varying by a factor of two in the typical range 0.04 to 0.08 (references in Sect. A.4.1 below).

### A.4. Assessing the bias from comparison with numerical simulations

The final piece of the jigsaw consists of assessing the bias \(b\) in Eq. (A.2). Since the relation has been calibrated using the HE mass for a sample of relaxed clusters, \(b\) represents the bias between \(M_{500}^{\text{HE}}\) and the true mass for this category of clusters. In principle, this can be assessed through comparison with numerical simulations. However, this approach is hampered by two difficulties. The first is the exact definition of “relaxed”, since it is almost impossible to select such clusters from observations and simulations according to the same criteria. The second is the specific implementation of the HE equation, which can differ substantially between observations (e.g., the use of forward fitting using parametric models, etc.) and simulations (e.g., the use of mock observations, etc.). Thus the amplitude of the bias that is found will depend not only on physical departures from HE, but also on technical details in the approach to data analysis.
Here we use a different approach that avoids these pitfalls, assessing the bias $b$ by comparing directly the estimated $Y_{500}\sim M_{500}$ relations with those found from numerical simulations. We then discuss the consistency of the resulting bias estimate with the HE bias expected from simulations and from absolute calibration uncertainties.

A.4.1. Comparison of simulated $Y_{500}\sim M_{500}$ relations and data

We first compared the $Y_{500}\sim M_{500}$ relations from 14 different analyses done by seven groups (Nagai 2006; Yang et al. 2010; Sehgal et al. 2010; Krause et al. 2012; Battaglia et al. 2012; Kay et al. 2012; Sembolini et al. 2013). We translated these simulations results into a common cosmology and, where necessary, converted cylindrical relations into spherical measurements assuming a ratio of $Y_{500, cy} / Y_{500, sph} = 0.74/0.61 \approx 1.2$, as given by the Arnaud et al. (2010) pressure profile.

The left-hand panel of Fig. A.2 shows the different $Y_{500}\sim M_{500}$ relations rescaled to our chosen cosmology. The simulations use various different types of input physics, and the resulting $Y_{500}\sim M_{500}$ relations depend strongly on this factor. The only obvious trend is a mild tendency for adiabatic simulations to find nearly self-similar slopes (1.66). Runs with non-gravitational processes tend to find slightly steeper slopes, but this is not always the case (e.g., the Krause et al. 2012 simulations). The right-hand panel of Fig. A.2 shows the ratio of each simulation $Y_{500}\sim M_{500}$ relation to the Planck $Y_{500}\sim M_{500}^{\text{HE}}$ relation given in Eq. (A.8). All results have been rescaled to account for the differences in baryon fraction between simulations. At our reference pivot point of $M_{500} = 6 \times 10^{14} M_{\odot}$, all simulations are offset from the measured relation. There is also a clear dependence on mass arising from the difference in slope between the majority of the simulated relations and that of the Planck relation. The Planck slope is steeper, possibly indicating the stronger effect of non-gravitational processes in the real data.

A.4.2. Quantification of the mass bias

We define the mass bias $b$ between the “true” and observed $M_{500}$ values, following Eq. (A.2) and explicitly allowing for possible mass dependence of the bias, i.e., $b = b(M_{500})$. Both masses are defined at a fixed density contrast of 500, so that the relations between observed and true mass and radius read

$$M_{500}^{\text{obs}} = [1 - b(M_{500}^{\text{true}})] M_{500}^{\text{true}}$$

(A.10)

$$R_{500}^{\text{obs}} = [1 - b(M_{500}^{\text{true}})]^{1/3} R_{500}^{\text{true}}$$

(A.11)

where “true” denotes simulated quantities, and “obs” denotes quantities estimated at the apertures derived from observations. The corresponding $Y_{500}\sim M_{500}$ relations are

$$Y^{\text{true}} = A_{\text{true}} M_{500}^{\text{true} \beta}$$

(A.12)

$$Y^{\text{obs}} = A_{\text{obs}} M_{500}^{\text{obs} \alpha}$$

(A.13)

In our case, $Y_{500}$ is measured interior to $R_{500}^{\text{obs}}$ as opposed to $R_{500}^{\text{true}}$. The ratio $Y^{\text{true}} / Y^{\text{obs}}$ depends on the radial variation of $Y_{500}$ for scaled radii, $r / R_{500}^{\text{true}} = R_{500}^{\text{true}} / R_{500}^{\text{obs}} = (1 - b)^{-1/3}$, which is close to 1. For a GNFW universal profile (Arnaud et al. 2010), we find that it can be well fit by a power law of the form $(1 - b)^{1/4}$. Combining Eq. (A.10), Eq. (A.12) and Eq. (A.13) we then arrive at

$$[1 - b(M_{500}^{\text{true}})] = \left[ A_{\text{true}} M_{500}^{\text{true} \beta} \right]^{-1/4+\alpha} \left[ A_{\text{obs}} M_{500}^{\text{obs} \alpha} \right]^{1/4+\alpha}$$

(A.14)

The relation makes it clear that a mass dependence of the bias naturally translates into a different slope of the observed and true $Y_{500}\sim M_{500}$ relations.

The bias $b$ can then be estimated from a comparison of observed and simulated relations, with the caveat that differences can also arise from imperfect modelling of cluster physics.
within the simulations. For the ensemble of simulations shown in Fig. A.2, the right panel shows the ratio of the observed and simulated relations as a function of mass. Figure A.3 shows the corresponding variation of \((1 - b)\) as a function of mass \(M\) from Eq. (A.14) for the observed slope \(a = 1.79\). This is mass dependent due to the difference in slopes between the simulated and observed relations. At a pivot point of \(M_{500} = 6 \times 10^{14} M_\odot\), the median value of \(A_{\text{true}}/A_{\text{obs}}\) is 0.74, implying \((1 - b) = 0.81\). However, there is a large amount of scatter in the predictions from simulations. As a consequence, \((1 - b)\) can vary from 0.74 to 0.97 at \(M_{500} = 6 \times 10^{14} M_\odot\). Note that the above results depend significantly on the baryon fraction \(f_b\). For example, assuming the WMAP-7 value \(f_b = 0.167\), the median value of \((1 - b)\) is 0.86 at the pivot point of \(M_{500} = 6 \times 10^{14} M_\odot\).

### A.4.3. Consistency with HE bias predictions and absolute calibration uncertainties

Taken at face value, the bias we derive above of \((1 - b) \approx 0.8\) implies that the HE mass used to calibrate the \(Y_{500} - M_{500}^{\text{HE}}\) relation is offset from the true mass by around 20 percent. Is this reasonable?

We can first compare HE X-ray and weak lensing (WL) masses. Although as mentioned above both measurements are expected to be biased, such comparisons are useful because the mass measurements involved are moderately large sample sizes of systems and are now starting to appear in the literature. However, at present there is little consensus, with some studies finding good agreement between HE and WL masses (e.g., Vikhlinin et al. 2009a; Zhang et al. 2010), some finding that HE masses are lower than WL masses, (e.g., Mahdavi et al. 2008), and some even finding that HE masses are higher than WL masses (Planck Collaboration Int. III 2013). The key point in such analyses is rigorous data quality on both the X-ray and optical sides. Most recent work points to relatively good agreement between X-ray and WL masses, with \(M^{\text{HE}}/M^{\text{WL}} \approx 0.9\) on average, and \(M^{\text{HE}}/M^{\text{WL}} \approx 1\) for relaxed systems (Mahdavi et al. 2013; von der Linden et al. 2014).

According to cosmological numerical simulations, the measurement bias induced by X-ray measurements relative to the “true” values can be caused by two main effects. The first is the classic “hydrostatic mass bias” due to non-thermal pressure support from turbulence/random motions, etc. However, the exact details are very model-dependent. The HE bias expected from simulations varies substantially, depending on the details of the numerical scheme, the input physics, and the approach used to calculate the HE masses (e.g., Rasia et al. 2012). In addition, the amount of bias is different depending on the dynamical state of the object, relaxed systems having less bias than unrelaxed systems. The majority of numerical simulations predict HE biases of 10 to 20 percent (Nagai et al. 2007; Piffaretti & Valdarnini 2008; Lau et al. 2009; Kay et al. 2012; Rasia et al. 2012).

Temperature inhomogeneities constitute the second contributor to X-ray measurement bias. In the presence of large amounts of cool gas, a single-temperature fit to a multi-temperature plasma will yield a result that is biased towards lower temperatures (e.g., Mazzotta et al. 2004). The presence of temperature inhomogeneities will depend on the dynamical state. While this effect can be investigated with simulations, estimates of its impact vary widely, owing to differences in numerical schemes and the different implementations of the input physics. For instance, simulations with heat conduction consistently predict smoother temperature distributions, thus X-ray spectroscopic biases are minimal in this case. On the other hand, “adiabatic” simulations predict long-lasting high-density cool-core type phenomena, which will lead to significant biases in single-temperature fits. Estimates of biasing due to temperature inhomogeneities can range up to 10 or 15 percent (e.g., Rasia et al. 2012).

Finally, for HE mass estimates obtained from X-ray data, instrument calibration uncertainties also play a significant role in introducing uncertainties in mass estimates. For instance, the difference in calibration between XMM-Newton and Chandra can induce differences in \(Y_X\). This is typically 5 percent, from a comparison of XMM-Newton based values published by Planck Collaboration XI (2011) to Chandra values for 28 ESZ clusters by Rozo et al. (2012). This can lead to differences of up to 10 percent in the mass \(M_{500}^{\text{HE}}\) derived from \(Y_X\), owing to the dependence of the mass on \(Y_X\).

Thus our adopted baseline value of \((1 - b) = 0.8\), ranging from 0.7–1, appears to encompass our current ignorance of the exact bias.

### A.5. Conclusions

In summary the baseline is

\[
E^{-2/3}(z) \left[ \frac{D_s^2}{L^2} Y_{500}^{\text{obs}} \right] = 10^{-0.19 \pm 0.02} \left( \frac{(1-b)M_{500}}{6 \times 10^{14} M_\odot} \right)^{1.79 \pm 0.08},
\]

with an intrinsic scatter of \(\sigma_{\log Y} = 0.075\) and a mean bias \((1 - b) = 0.80^{+0.21}_{-0.20}\). The statistical uncertainty on the normalization is about 5% and the error budget is dominated by the systematic uncertainties.

### References
