Comparing Masses in Literature (CoMaLit)-I. Bias and scatter in weak lensing and X-ray mass estimates of clusters

Mauro Sereno1*, Stefano Ettori2,3
1Dipartimento di Fisica e Astronomia, Università di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italia
2INAF, Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italia
3INFN, Sezione di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italia

31 July 2014

ABSTRACT
The first building block for using galaxy clusters in astrophysics and cosmology is an accurate determination of their mass, which can be estimated with weak lensing (WL) determinations or X-ray analyses assuming hydrostatic equilibrium (HE). By comparing the two mass proxies in well observed samples of rich clusters, we determined the intrinsic scatters, \( \sigma_{\text{WL}} \sim 15 \) per cent for WL masses and \( \sigma_{\text{HE}} \sim 25 \) per cent for HE masses. The certain assessment of the bias is hampered by differences as large as \( \sim 40 \) per cent in either WL or HE mass estimates reported by different groups. If the scatter in the mass proxy is not considered, the slope of any scaling relation ‘mass–observable’ is biased towards shallower values, whereas the intrinsic scatter of the scaling is over-estimated.

Key words: galaxies: clusters: general – gravitational lensing: weak – galaxies: clusters: intracluster medium – methods: statistical

1 INTRODUCTION
Usage of clusters of galaxies in cosmology and astrophysics relies on precise determination of their masses (Voit 2005; Limousin et al. 2013). In the context of ongoing and future large surveys (Laureijs et al. 2011), cluster properties which can easily measured, such as optical richness, X-ray luminosity, Sunyaev-Zel’dovich (SZ) flux, ..., are used as mass proxies. This requires an accurate calibration of the observable through comparison with direct mass estimates (Andreon & Bergé 2012; Ettori 2013).

The assessment of scaling relations is the foundation for investigating the physics of the baryons and of the dark matter at the cluster scale (Pratt et al. 2009; Arnaud et al. 2010; Giodini et al. 2013). Cosmological parameters can be constrained with cluster abundances and the observed growth of massive galaxy clusters (Mantz et al. 2010; Planck Collaboration et al. 2013b) or with gas fractions (Ettori et al. 2009).

Two of the most well regarded mass estimates are the weak lensing (WL) mass and the X-ray mass exploiting hydrostatic equilibrium (HE). Weak lensing observations of the shear distortion of background galaxies trace the gravitational field of the matter distribution of the lens (Hoekstra et al. 2012; von der Linden et al. 2014; Umetsu et al. 2014). The physics behind gravitational lensing is very well understood and WL provides unbiased estimates of the total mass along the line of sight. The problem is to single out the contribution of the lens and to de-project the information to get the intrinsic mass, which can then be confronted with theoretical predictions.

Under the assumption that hydrostatic equilibrium holds between the intracluster medium (ICM) and the gravitational potential, the cluster mass can be recovered from observations of the spatially resolved spectroscopic data and the X-ray surface brightness (LaRoque et al. 2006; Donahue et al. 2014). However, deviations from equilibrium or non-thermal contributions to the pressure are difficult to quantify and can bias the mass estimate.

Methods based on spectroscopic measurements of galaxies velocities, such as the caustic technique (Rines & Diaferio 2006) or approaches exploiting the Jeans equation (Lemze et al. 2009; Biviano et al. 2013), can be effective too but they are hindered by the very expensive observational requirements and are mostly limited to low redshift halos.

In principle either WL or HE can provide accurate and unbiased mass measurements, but the approximations that have to be used (spherical symmetry, smooth density distributions, thermal equilibrium,...) may bias and scatter the results. These effects must be accurately quantified to calibrate other mass proxies.

Numerical studies argued that lensing masses obtained from the fit of the cluster tangential shear profiles with Navarro-Frenk-White (Navarro, Frenk & White 1996, NFW) functionals are biased low by \( \sim 5–10 \) per cent with a scatter of \( \sim 10–25 \) per cent (Meneghetti et al. 2010; Becker & Kravtsov 2011; Rasia et al. 2012). The main sources of uncertainty in deprojected WL mass measurements are due to the presence of substructures and triaxiality. Lensing properties depend on the orientation of the cluster...
with respect to the line of sight (Oguri et al. 2005; Sereno 2007; Sereno & Umetsu 2011; Limousin et al. 2013). For systems whose major axis points toward the observer, 3D masses derived under the standard assumption of spherical symmetry are typically over-estimated. The opposite occurs for clusters elongated in the plane of the sky, which are in the majority if the selected sample is randomly oriented.

The presence of substructures in the cluster surroundings may dilute the tangential shear signal (Meneghetti et al. 2010; Giocoli et al. 2012, 2014). Severe mass under-estimations may come from either massive sub-clumps (Meneghetti et al. 2010) or uncorrelated large-scale matter projections along the line of sight (Becker & Kravtsov 2011).

The scatter is less significant in optimally selected clusters either having regular morphology or living in substructure-poor environments (Rasia et al. 2012).

The origins of bias and scatter of X-ray masses are well understood too, even though they are difficult to quantify (Rasia et al. 2012). They are strictly connected to non-thermal sources of pressure in the gas, to temperature inhomogeneity and, to a lesser degree and mainly in the external regions, to the presence of clumps. Even if the cluster is in hydrostatic equilibrium, the assumption that all the pressure is thermal biases the HE mass low. Large-scale, unvitalized bulk motions and subsonic turbulence contribute kinetic pressure (Battaglia et al. 2012).

Furthermore, structures in the temperature distribution bias low the temperature estimate. In fact, the X-ray detectors of Chandra and XMM-Newton have a higher efficiency in the soft band and, thus, weight more colder gas (Mazzotta et al. 2004).

Numerical simulations showed that X-ray masses based on hydrostatic equilibrium are biased low by a large amount of ∼25–35 per cent (Piffaretti & Valdarnini 2008; Rasia et al. 2012, 2014). The bias grows from the inner to the outer regions of the clusters, where the presence of non-thermal sources of pressure in the ICM and temperature inhomogeneity play a larger role (Rasia et al. 2012).

Since the intrinsic scatters in either WL or HE masses have different origins, they are mostly uncorrelated. Scatter in WL masses is mainly due to triaxiality and sub-structures in the dark matter halo. However, the gas distribution approximately follows the gravitational potential and is rounder than the dark matter one. Dark matter substructures are not necessarily associated to gas clumps. On the other hand, the sources which cause scatter in the HE masses are more related to gas physics and temperature distributions than to the total matter distribution and have a small impact on WL estimates.

On the observational side, the certain assessment of cluster masses is further complicated by instrumental and methodological sources of errors which may cause systematic uncertainties in data analysis (Rozo et al. 2014b).

The main sources of systematics in WL masses are due to selection and redshift estimate of background galaxies, which can be obtained through accurate photometric redshifts and color-color selection methods (Medezinski et al. 2010), and to the calibration of the shear signal. A small calibration correction of the order of just a few percents translates into a typical error of ∼10 per cent in the estimate of the virial mass (Umetsu et al. 2014).

Instrumental uncertainty has long been recognized as one of the main source of systematics plaguing HE masses. XMM cluster temperatures are systematically smaller by 10-20 per cent than Chandra estimates (Nevalainen, David & Guainazzi 2010; Donahue et al. 2014). On the other hand, Chandra and XMM measurements of the gas distribution are highly consistent with one another (Rozo et al. 2014b; Donahue et al. 2014).

The picture on inconsistencies between Chandra and XMM results is still debated. Donahue et al. (2014) found that Chandra and XMM temperatures of the very massive CLASH clusters agree in the core, where photon fluxes are considerable, whereas the regions where the temperature differences are larger are typically ∼1 arcmin from the much brighter cluster core. Temperature differences persist even in outer regions with large signal-to-background ratio. These temperature discrepancies caused analogue offset-s in the HE mass.

Martino et al. (2014) compared the mass profiles of 21 LoCuSS (Local Cluster Substructure Survey) clusters that were observed with both satellites, extracting surface brightness and temperature profiles from identical regions of the respective datasets and including analytic models that predict the spatial variation of the Chandra and XMM-Newton backgrounds to ≤2 and ≤5 per cent precision, respectively. Notwithstanding global XMM spectroscopic temperatures lower by ∼10 per cent, they obtained consistent results for the gas and total hydrostatic cluster masses. Martino et al. (2014) explained this counterintuitive result noticing that temperature discrepancies were significant only above a value of 6 keV. In the outer regions, most of the estimated temperatures were lower than this threshold and Chandra and XMM temperatures were in good agreement. Furthermore, they argued that larger errors bars are associated to highest temperature, due to the larger difficulty to distinguish the hottest spectra having a flatter shape from the background. The relative statistical weight in a fitting procedure is then lower.

This is the first in a series of papers focused on CoMParing MAsses in LITerature (CoMaLit). Here, we look for systematic differences in WL and HE masses obtained from independent analyses and we assess the overall level of bias and intrinsic scatter. According to numerical simulations, the scatter in X-ray masses is supposedly smaller than in weak-lensing masses but a definite assessment of the values of bias and scatter of HE masses is still lacking, due to uncertainties in the treatment of the gas physics and to variability caused by the hydrodynamical scheme adopted in numerical simulations (Rasia et al. 2014). In this paper, we provide the first measurements of intrinsic scatters for WL and HE masses.

In the second paper of the series (Sereno, Ettori & Moscardini 2014), we discuss the scaling relation between the SZ flux and the cluster mass in a Planck selected clusters of galaxies (Planck Collaboration et al. 2013).

The present paper is structured as follows. In Sec. 2 we discuss how the scatter in mass proxies can be estimated and how it impacts the calibration of scaling relations. Samples of clusters used in the analysis are introduced in Sec. 3. Comparison among either WL or HE masses from different groups is investigated in Sec. 4. Section 5 is devoted to the measurements of scatter and biases affecting the mass proxies. Discussion of results is contained in Sec. 6. Final considerations can be found in Sec. 7.

Throughout the paper, we assume a fiducial flat ΛCDM cosmology with density parameter Ω_M = 0.3, and Hubble constant H_0 = 70 km s^{-1} Mpc^{-1}; M_{500} denotes the mass within the radius r_{500}, which encloses a mean overdensity of 500 times the critical density at the cluster redshift, ρ_{cr} = 3H(z)^2/(8πG); H(z) is the redshift dependent Hubble parameter. When H_0 is not specified, h is the Hubble constant in units of 100 km s^{-1} Mpc^{-1}.

The presence of the superscript ‘WL’, ‘HE’, and ‘Tr’, means that M_{500} and r_{500} were determined using the mass estimate from the WL analysis, the X-ray measurements or the knowledge of the
true mass (which is available only for simulated clusters), respectively. \( \log \) is the logarithm to base 10 and \( \ln \) is the natural logarithm.

2 BIASES AND SCATTER INDUCED BIASES

The biases and the scatters of two mass proxies can be estimated by comparing the proxies in a cluster sample. The lensing and the hydrostatic mass approximate the true mass as

\[
\ln M^{WL} \pm \delta_{WL} = \sigma_{WL} \ln M^{TV} \pm \sigma_{WL}, \quad (1)
\]

\[
\ln M^{HE} \pm \delta_{HE} = \sigma_{HE} \ln M^{TV} \pm \sigma_{HE}, \quad (2)
\]

where the \( \alpha \)'s quantify the bias and the \( \beta \)'s embody any deviation from linearity. The relations in Eqs. (1) and (2) can be modelled with normal distributions. The intrinsic scatter \( \sigma_{WL} \) and \( \sigma_{HE} \) are due to different physical processes and are assumed to be uncorrelated. The actual WL (HE) mass is known save for a measurement error \( \delta_{WL} \) (\( \delta_{HE} \)).

Bias and scatter in logarithmic variables slightly differ from analogue quantities in linear variables. We adopt (natural) logarithmic variables for coherence with the standard derivation of scaling relations.

2.1 Eddington-like bias

Intrinsic scatter in the mass proxy induces systematic effects alike to the Eddington bias [Eddington 1913].\cite{Eddington1913}\cite{Jeffreys1938}\cite{Eddington1940}. Due to scatter, the average value of an observed quantity differs from the true intrinsic mean for objects of the same class, see Fig. [1]. When a subsample is selected according to the measured values of the proxy, \( X_{Proxy} \), the distribution of the differences between the proxy and the true values, \( X_{True} \), may be biased.

For quantities drawn from a limited range, border and selection effects have to be considered. Near a threshold, the asymmetry between objects that are scattered into a range of observed values from above and objects that are scattered into from below is broken. This can be accounted for by assuming that the true masses are drawn from a normal rather than a uniform distribution.\cite{Kelly2007}

Let us assume that the proxy \( X_{True} \) is the WL mass. Due to selection effects, the observed sample may be poor in clusters below a given threshold. At the tail at low values, more objects with larger \( X_{True} \) are scattered into the subsample from the right side, than from the left side where the \( X_{True} \)'s are smaller, see Fig. [1].\cite{Kelly2007}

In a sample with steep bounds in true mass, clusters with very low values of \( M^{WL} \) are then of two main kinds. They are either intrinsically less massive clusters, i.e., with low values of \( M^{TV} \) and nearly unbiased values of \( M^{WL} \), or clusters with higher values of \( M^{TV} \) that are scattered to lower values of observed \( M^{WL} \). The mean \( M^{TV} \) is then biased low with respect to the measured \( M^{WL} \).

On the other hand, a second proxy such as the HE mass is not biased by selections based on the first proxy. If the clusters are selected according to their \( M^{WL} \), \( M^{HE} \) is still an unbiased scattered proxy of the true mass, since the scatters in measurements of WL and X-ray masses are uncorrelated. As a consequence, the mass ratio \( M^{HE}/M^{WL} \) is biased high for clusters with small WL masses.

The opposite happens at large masses, where the more massive the clusters the rarer. The mass ratio \( M^{HE}/M^{WL} \) is then biased low for clusters with large WL masses.

2.2 Biased slope

The intrinsic scatter in the mass estimate can make the slope of any scaling relation calibrated with either WL or HE masses shallower if the true masses in the selected sample are not uniformly distributed. This is a ripple effect of the Eddington-like bias. The Eddington bias was first discussed in relation to observational uncertainties [Eddington 1913].\cite{Jeffreys1938}\cite{Eddington1940}. Due to measurement errors, the observed variance of the mass proxies is larger than the variance of the true masses. Slope estimators have to correct for this by de-biasing the sample variance [Akritas & Bershady 1996]. Here, we are emphasizing the similar effect of the intrinsic scatter. Analog treatments, which are often focused on observational errors rather than intrinsic scatter, have already been discussed [Andreon & Bergé 2012].

The distribution of the observed mass proxy is smoothed and it has a larger dispersion than the true masses. Due to the finite range, very large (small) measured WL or HE masses likely correspond to smaller (larger) true masses (in arbitrary units), whose observed WL or HE mass were scattered to the tails. If this is not accounted for, the derived slope of the scaling relation is biased toward flatter values.

Let us consider an unbiased (but scattered) proxy of the true mass,

\[
M^{Pr} \pm \delta_{Pr} = M^{TV} \pm \sigma_{Pr}, \quad (3)
\]

and a second observable quantity \( Y \) we want to calibrate,

\[
Y \pm \delta_Y = \alpha_Y + \beta_Y M^{TV} \pm \sigma_Y. \quad (4)
\]

What we usually do is to compare the observable \( Y \) to the mass proxy,

\[
Y \pm \delta_Y = \alpha_Y^{Pr} + \beta_Y^{Pr} M^{Pr} \pm \sigma_Y^{Pr}. \quad (5)
\]

Due to the intrinsic scatter in the mass proxy, \( \alpha_Y^{Pr} \) and \( \beta_Y^{Pr} \) are biased estimates of \( \alpha_Y \) and \( \beta_Y \). The effect can be studied through a simple simulation. Let us consider a sample of 100 true (logarithmic) masses drawn from a Gaussian distribution with mean \( \mu_X = 1.0 \) and \( \sigma_X = 0.35 \). The Gaussian distribution provides a good approximation for signal-selected samples. In fact, at low masses the number of clusters is limited by the selection threshold. At high masses, there are a few clusters because of the steepness of the mass function. The resulting distribution is then approximately normal for realistic cases [Andreon & Bergé 2012].

The observed mass proxies differ from the true values due to the intrinsic scatter \( \sigma_{Pr} = 0.15 \) and an observational uncertainty \( \delta_{Pr} = 0.05 \). The proxy \( Y \) is linearly related to \( M^{TV} \) with \( \alpha_Y = -0.2 \) and \( \beta_Y = 1.5 \) and a scatter \( \sigma_Y = 0.2 \). The observational uncertainty is \( \delta_Y = 0.05 \). We model scatters and errors with normal distributions.

We performed the linear regression with the Bayesian package JAGS.\cite{Plummer2003} Previous applications of this Bayesian technique to astronomical contexts can be found in [Andreon & Hurn 2012]. The intrinsic distribution of the independent variable was approximated with a Gaussian function of mean \( \mu \) and standard deviation \( \sigma \). We adopted uniform priors for the intercept and the mean \( \mu \). For the variances, i.e., the squared scatters, we considered inverse Gamma distribution [Andreon & Hurn 2010]. For the slope \( \beta \), we assumed a Student’s \( t \) distribution, which is equivalent to a uniform prior on the direction angle \( \arctan \beta \).

1 http://mcmc-jags.sourceforge.net/.
As a first step, we verified that the regression retrieves unbiased parameters when we compare observable and true mass, see Eq. (4). We found $\alpha_Y = -0.22 \pm 0.07$, $\beta_Y = 1.49 \pm 0.07$ and $\sigma_Y = 0.17 \pm 0.02$. Fit results are statistically consistent with the input parameters.

As far as the evolution of $Y$ with the mass proxy is concerned, see Eq. (5), we found $\alpha_Y = 0.00 \pm 0.08$, $\beta_Y = 1.25 \pm 0.07$ and $\sigma_Y = 0.23 \pm 0.02$. The relation is flatter than the intrinsic one and the estimated scatter is larger. On turn, the flatter relation causes a higher intercept $\alpha$.

To avoid biases, we have to consider that $M_{\text{Pr}}$ is a scattered proxy of the true mass. Equations (3) and (4) have to be fitted simultaneously. We then performed the regression adding a further parameter $\sigma_Y$. We found $\alpha_Y = -0.18 \pm 0.16$, $\beta_Y = 1.43 \pm 0.15$, $\sigma_Y = 0.15 \pm 0.07$ and $\sigma_Y = 0.11 \pm 0.05$. Intrinsic parameters are well recovered even though statistical uncertainties are larger.

More details on this statistical model are provided in Appendix B which also provides some ready-to-use approximate corrections. Correcting $\alpha_Y$ and $\beta_Y$ as suggested in Eqs. (B10), we found $\alpha_Y \sim -0.22$ and $\beta_Y \sim 1.48$, in agreement with the input parameters.

### 3 CLUSTER SAMPLES

We looked in literature for public catalogs compiled in the last few years with either WL or HE masses. The main properties of the samples, which we are going to discuss in the following, are summarized in Table 1.

When quoted mass values were provided with asymmetric errors, we estimated the mean value and the standard deviation as suggested in D’Agostini (2004). All the considered masses refer to the fiducial cosmological model. Conversions were performed as described in App. A.

#### 3.1 Simulated sample

Rasia et al. (2012) compared the weak-lensing and X-ray properties of a sample of numerically simulated massive clusters at redshift $z = 0.25$. The haloes were the most massive ($M_{\text{vir}} > 5 \times 10^{14} M_{\odot} h^{-1}$) from a set of radiative simulations in a cosmological simulation of volume 1 (Gpc$/h)^{3}$, evolved in the framework of a WMAP-7 normalized cosmology (Fabjan et al. 2011).

Each cluster was later re-simulated at higher resolution and with more complex gas physics. The simulations included: metal-dependent radiative cooling and cooling/heating from a spatially uniform and evolving UV background; a star-formation model where a hot ionized phase coexists in pressure equilibrium with a cold phase, which is the reservoir for star formation; a description of metal enrichment from different stellar populations; the effect of supernovae feedback through galactic winds.

The clusters were finally processed to generate optical and X-ray mock observations along three orthogonal projections. The final sample consists of 60 cluster realizations. WL and HE masses are estimated within $r_{500}$, the over-density radius corresponding to the true mass.

#### 3.2 Canadian Cluster Comparison Project

The Canadian Cluster Comparison Project (CCCP, Mahdavi et al. 2013) assembled a sample of 50 rich clusters of galaxies in the redshift range $0.15 < z < 0.55$. All of the clusters were observable from the Canada-France-Hawaii Telescope (CFHT), which restricts the sample to systems at $15 \text{ deg} < \text{declination} < 65 \text{ deg}$. Most of them were selected to have an ASCA (Advanced Satellite for Cosmology and Astrophysics) temperature $k_B T_X > 3 \text{ keV}$. X-ray properties were measured either with Chandra or XMM-Newton. Weak lensing studies for 5 additional clusters without X-ray analyses can be found in Hoekstra et al. (2012).

Lensing masses were determined with aperture statistics (Hoekstra et al. 2012). This approach relies on shear measurements at large radii. Contamination by cluster members is suppressed. The 3D masses were computed from the model-independent 2D aperture masses with a de-projection method based on a NFW density profile.

Mahdavi et al. (2013) performed an X-ray analysis of the sample using both Chandra and XMM observations. They found that due to temperature discrepancies, the XMM cluster masses were systematically $\sim 15$ per cent smaller than Chandra masses. In order to combine the data, Mahdavi et al. (2013) down-weighted the
Table 1. Characteristics of the X-ray and WL samples used in the analysis. $N_{\text{cl}}$ is the number of clusters in the sample.

| Acronym   | $N_{\text{cl}}$ | WL instrument | WL reference | X-ray instrument | X-ray reference | notes |
|-----------|----------------|---------------|--------------|-----------------|----------------|-------|
| RA12      | 60             | —             | —            | —               | —              | —     |
| CCCP-WL   | 55             | CFHT          | Hoekstra et al. (2012) | —               | —              | —     |
| CCCP-HE   | 50             | —             | Chandra, XMM | Mahdavi et al. (2013) | —              | —     |
| WTG       | 51             | Subaru, CFHT  | Applegate et al. (2014) | —               | —              | —     |
| CLASH-WL  | 20             | Subaru        | Umetsu et al. (2014) | —               | —              | —     |
| CLASH-CXO | 25             | —             | Chandra       | Donahue et al. (2014) | —              | —     |
| CLASH-XMM | 18             | —             | XMM           | Donahue et al. (2014) | —              | —     |
| E10       | 44             | —             | XMM           | Ettori et al. (2010) | —              | —     |
| L13       | 35             | —             | Chandra       | Landry et al. (2013) | —              | —     |
| B12       | 25             | —             | Chandra       | Bonamente et al. (2012) | Additional SZ data from SZA | —     |

High-energy effective area of Chandra. X-ray quantities were estimated either within $r_{500}^{\text{WL}}$, the radius evaluated from the weak-lensing mass measurement, or $r_{500}^{\text{HE}}$, as evaluated from the mass estimate assuming hydrostatic equilibrium. Values of $M_{500}^{\text{WL}}$ and $M_{500}^{\text{HE}}$ were provided by the authors.

Mahdavi et al. (2013) identified a subsample of 20 cool core systems with core entropy at 20 kpc smaller than 70 keV cm$^{-2}$ and 8 systems with low offsets between the brightest cluster galaxy (BCG) and the X-ray surface brightness peak, $D_{\text{BCG}} < 10$ kpc.

3.3 Cluster Lensing And Supernova survey with Hubble

The Cluster Lensing And Supernova survey with Hubble (CLASH, Postman et al. 2012) has been mapping the matter distribution of 25 rich clusters drawn largely from the Abell and Massive Cluster Survey (MACS, Ebeling et al. 2010) cluster catalogs. Umetsu et al. (2014) performed a joint shear-and-magnification weak-lensing analysis of a sub-sample of 16 X-ray regular and 4 high-magnification galaxy clusters in the redshift range $0.19 \lesssim z \lesssim 0.60$. A complementary analysis exploiting strong lensing data was presented in Merten et al. (2014).

To make the comparison with the other data samples easier, we will use the mass estimates in Umetsu et al. (2014), whose methodology exploits only the weak-lensing regime whereas results in Merten et al. (2014) strongly relies on information from the inner regions. In fact, mass estimates in Umetsu et al. (2014) were based on joint weak lensing shear plus magnification measurements based on ground-based wide-field Subaru data. On the other hand, the analysis in Merten et al. (2014) combined the Subaru shear profile with weak-lensing constraints from the Hubble Space Telescope (HST) in the intermediate regime and strong lensing constraints from HST.

All of the CLASH clusters have been observed with the Chandra satellite (Postman et al. 2012). A subsample of 18 clusters was targeted by XMM too. The X-ray analysis was presented in Donahue et al. (2014), which computed HE masses and gas fractions. Based on Chandra data, Donahue et al. (2014) identified 10 clusters (9 of them with WL mass) with a strong cool core, i.e., with an excess core entropy smaller than 30 keV cm$^{-2}$.

3.4 Weighing the Giants

The Weighing the Giants (WTG, von der Linden et al. 2014) program targeted 51 X-ray luminous clusters from the MACS and the Brightest Cluster Survey (BCS, Ebeling et al. 2000). The clusters span a large range in redshift $(0.15 \lesssim z \lesssim 0.7)$ and dynamical state. 7 clusters are classified as relaxed (von der Linden et al. 2014).

We derived $M_{500}^{\text{WL}}$ and $r_{500}^{\text{WL}}$ using the NFW density profile adopted in the WTG analysis. The values of the scale radius and the concentration are provided in Applegate et al. (2014, table 4).

3.5 X-ray samples

Ettori et al. (2010) E10) studied a sample of 44 X-ray luminous galaxy clusters observed with XMM-Newton in the redshift range $0.1 \lesssim z \lesssim 0.3$. They applied two different techniques (the ‘method 1’, which we take as the reference method, and the ‘method 2’) to recover the gas and the dark mass properties, described with a NFW profile. Clusters were classified according to their core properties. E10 identified a subsample of 16 low-entropy-core systems, which represent the prototype of a relaxed cluster with a well defined cool core at low entropy.

Landry et al. (2013) L13) presented Chandra X-ray measurements of the hydrostatic mass and of the gas mass fraction out to $r_{500}$ for the complete sample of the 35 most luminous clusters from the BCS and its extension at redshift $0.15 \lesssim z \lesssim 0.30$. The clusters span a large range of dynamical states. The data were analysed using two independent pipelines and two different models for the gas density and temperature, the ‘Polytropic’ (which we take as our reference case) or the ‘Vikhlinin’ model.

Bonamente et al. (2012) B12) derived the hydrostatic masses and the pressure profile of a sample of 25 massive relaxed galaxy clusters with a simultaneous analysis of SZ data from the Sunyaev-Zel’dovich array (SZA) and archival Chandra observations.

4 MASS COMPARISON

Even though in principle WL and X-ray masses could be unambiguously determined from a given set of observations, calibration issues and hidden systematics make these measurements very difficult.

In this section we compare either WL or HE masses from different catalogs. It is nowadays customary to quote masses within a given overdensities and to derive scaling relations in terms of them. These masses can be related to the virial mass and most cluster properties are expected to be self-similar if rescaled by their value at $r_{\Delta}$.

To limit extrapolation of published results, we then considered the masses within $r_{500}$, rather than extrapolating the results up to a fixed length.
Table 2. Comparison of WL masses from independent analyses. Entries are in the format: \((N_{cl}), M_{\Delta}(\pm \delta M) \pm \sigma(\pm \delta \sigma)\), where \(N_{cl}\) is the number of clusters in common between the samples, \(M\) is the central estimate of the difference in natural logarithm \(\Delta \ln \left(\frac{M_{\text{row}}}{M_{\text{col}}}\right)\), with associated uncertainty \(\pm \delta M\), \(\sigma\) is the dispersion with associated uncertainty \(\pm \delta \sigma\). \(M_{\text{row}}\) \((M_{\text{col}})\) refers to the sample indicated in the corresponding row (column). Quoted values are the bi-weight estimators of the mass ratios.

| CLASH-WL   | WTG   |
|------------|-------|
| (6)        | (17)  |
| −0.45(±0.12)| −0.31(±0.05) |
| ±0.25(±0.10) | ±0.21(±0.09) |

On the other hand, the relationship between \(M_\Delta\) and \(r_\Delta\) exacerbates problems connected to aperture differences, which complicate the comparison between different samples. Since the total mass within a fixed radius scales nearly linearly with the radius, differences in mass within a given overdensity are inflated by \(\sim 100/3\) per cent with respect to differences within a fixed physical radius.

Differences among properties measured within a fixed length are not inflated but they refer to physically different regions in differently sized clusters. A promising alternative is to express the results in terms of the circular velocity \(v_c^2 = GM(r)/r\). In fact, the circular velocity is almost independent of cosmology and results in terms of the circular velocity are not inflated but they refer to physically different regions in different samples. Since the total mass within a fixed radius scales nearly linearly with the radius, differences in mass within a given overdensity are inflated by \(\sim 100/3\) per cent with respect to differences within a fixed physical radius.

Differences among properties measured within a fixed length are not inflated but they refer to physically different regions in differently sized clusters. A promising alternative is to express the results in terms of the circular velocity \(v_c^2 = GM(r)/r\). In fact, the circular velocity is almost independent of cosmology and results in terms of the circular velocity are not inflated but they refer to physically different regions in different samples. Since the total mass within a fixed radius scales nearly linearly with the radius, differences in mass within a given overdensity are inflated by \(\sim 100/3\) per cent with respect to differences within a fixed physical radius.

To compare different samples we considered the (natural) logarithm of mass ratios (Rozo et al. 2014b). The central estimate and the scatter were computed as bi-weight estimators of the distribution. Uncertainties were estimated with bootstrap resampling with replacement. The main advantage in using logarithms is that their difference is (anti-)symmetric. This solves the problem affecting those estimators of ratios which are not symmetric with respect to an exchange of the numerator and denominator.

Quoted errors in compiled catalogs may account for different sources of statistical and systematic uncertainties. Furthermore, it can be argued that the published uncertainties are unable to account for the actual variance seen in sample pairs (Rozo et al. 2014b). We then conservatively performed unweighted analyses.

4.1 WL masses

In principle, the WL mass calibration could be determined to an accuracy of \(\lesssim 8\) per cent (von der Linden et al. 2014; Umetsu et al. 2014), but differences between masses reported by different groups are off by \(\approx 20\)–\(50\) per cent (Applegate et al. 2014; Umetsu et al. 2014). On the other hand, comparisons show that mass measurements correlate quite tightly (Applegate et al. 2014).

The CCCP and the WTG samples share 17 clusters, see Table 2. \(M_{500}\) from CCCP are smaller by \(\approx 30\) per cent with a scatter of \(\approx 20\) per cent. This difference is way larger than the claimed mass calibration uncertainty and highlights the difficulties connected to unbiased calibrations in WL measurements. We found no trend with redshift, see Fig. 2.

The CCCP masses are nearly unaffected by aperture problems (Donahue et al. 2014). In fact, the circular velocity is almost independent of cosmology and results in terms of the circular velocity are not inflated but they refer to physically different regions in different samples. Since the total mass within a fixed radius scales nearly linearly with the radius, differences in mass within a given overdensity are inflated by \(\sim 100/3\) per cent with respect to differences within a fixed physical radius.

Differences among properties measured within a fixed length are not inflated but they refer to physically different regions in differently sized clusters. A promising alternative is to express the results in terms of the circular velocity \(v_c^2 = GM(r)/r\). In fact, the circular velocity is almost independent of cosmology and it is nearly unaffected by aperture problems (Donahue et al. 2014). Within a given over-density radius \(v_\Delta^2 \propto M_\Delta^{2/3}\). Quoted results for central estimate and scatter of \(\Delta \ln M_{500}\), as well as fractional changes, can be translated in analogue results for \(\Delta \ln v_{500}\) by simply multiplying by the factor 2/3.

To compare different samples we considered the (natural) logarithm of mass ratios (Rozo et al. 2014b). The central estimate and the scatter were computed as bi-weight estimators of the distribution. Uncertainties were estimated with bootstrap resampling with replacement. The main advantage in using logarithms is that their difference is (anti-)symmetric. This solves the problem affecting those estimators of ratios which are not symmetric with respect to an exchange of the numerator and denominator.

Quoted errors in compiled catalogs may account for different sources of statistical and systematic uncertainties. Furthermore, it can be argued that the published uncertainties are unable to account for the actual variance seen in sample pairs (Rozo et al. 2014b). We then conservatively performed unweighted analyses.

4.2 X-ray masses

X-ray properties of galaxy clusters reported by competing groups may reach discrepancies of 50 per cent (Rozo et al. 2014b). Here, we consider the off-set and the scatter in the estimate of HE masses. Discrepancies may stem from either differences in the considered data sets (mainly if taken with different instruments), or from non-consistent data reduction pipelines or, finally, from different techniques to recover the mass.

This last issue can be quantified by comparing mass estimates obtained from the same data-sets but with different methodologies. This is the case of the analyses in either E10 or L13, for which we could compare the scatter in the mass estimate due to the different modelling, see Table 3.

The typical statistical error in a HE mass estimate is of the order of \(\approx 15\) per cent. The observed scatter in the mass ratios is then consistent with the propagation of this error. This comparison suggests that mass estimates are not biased due to different tech-
Table 4. Comparison of HE masses from independent analyses. For the CCCP-HE sample, we considered masses within $r_{500}^{\text{HE}}$. Entries are as in Table 2.

| Method     | CCCP-HE      | CLASH-XMM     | CLASH-CXO     | L13     | B12     |
|------------|--------------|---------------|---------------|---------|---------|
|            | (11)         | (3)           | (3)           | (11)    | (6)     |
| E10        | 0.22(±0.08)  | ~ 0.17        | ~ −0.15       | 0.35(±0.14) | 0.25(±0.10) |
|            | ±0.28(±0.11) | ±(~)0.06      | ±(~)0.28      | ±(~)0.30 | ±(~)0.19 |
| CCCP-HE    | —            | 0.03(±0.21)   | −0.38(±0.14)  | 0.12(±0.07)  | 0.24(±0.22) |
|            | (5)          | ±0.29(±0.16)  | ±0.34(±0.21)  | ±0.33(±0.14) | ±0.35(±0.17) |
| CLASH-XMM  | —            | —             | −0.38(±0.09)  | ~ −0.05   | 0.14(±0.15) |
|            | (18)         | ±0.35(±0.10)  | ±(~)0.18      | ±(~)0.18  | ±(~)0.30 |
| CLASH-CXO  | —            | —             | —             | 0.31      | 0.45(±0.14) |
|            | (4)          | ±(~)0.01      | ±(~)0.01      | ±0.37(±0.13) |
| L13        | —            | —             | —             | —         | 0.03     |
|            | (4)          |                |                | ±(~)0.06  |          |

Figure 3. Masses in the RA12 sample. Top panel: WL mass vs. HE mass. Clusters are grouped in four bins in true mass (black points). Lower panel: bias of the proxy as a function of the true mass. Black (blue) points correspond to the bias of the WL (HE) mass. The solid error-bars denote the 1-σ uncertainties for the central estimate. The dashed error-bars denote the dispersion. All masses are computed within $r_{500}^{\text{WL}}$.

Niques, whose associated variance is negligible with respect to the statistical uncertainty.

Larger variations are mainly related to different data-sets, see Table 2. Discrepancies of order of $\gtrsim 30$ per cent may be in place. This may be the case for results based on Chandra (CLASH-CXO, B12, L13) versus XMM analyses (E10, CLASH-XMM), whose temperature estimates may disagree at large radii (Donahue et al. 2013).

Each method/analysis may systematically either over- or underestimate the cluster mass. X-ray masses in the CLASH sample based on Chandra (XMM) data are systematically larger (smaller) than other estimates. On the other hand, masses from B12 and L13 are lower than other samples.

A significant role can be played by additional data-sets exploited in the analysis. The inclusion of SZ data, which are more sensitive to the outer regions, might lower the mass values in B12.

The large differences in estimated masses and the large scatter suggest that quoted formal statistical uncertainties in HE masses, usually of the order of $\sim 10$–15 per cent, might be underestimated.

5 REGRESSION RESULTS

We measured biases and intrinsic scatters of WL and HE masses through the statistical model detailed in Sec. 2. To simplify the analysis, we assumed that the lensing and the hydrostatic masses scale linearly with the true mass, $\beta_{\text{WL}} = 1$ and $\beta_{\text{HE}} = 1$.

The true masses are known only in simulations. For observed samples, we could estimate only the relative bias between WL and HE masses and we fixed $\alpha_{\text{WL}} = 0$. The effective bias $M_{500}^{\text{HE},\text{WL}}/M_{500}^{\text{WL}}$ can be defined as $\exp(\alpha_{\text{HE},\text{WL}})$. The relative bias $M_{500}^{\text{HE}}/M_{500}^{\text{WL}}$ can be defined as $\exp(\alpha_{\text{HE}} - \alpha_{\text{WL}})$. Bias and scatter are largely uncorrelated. We tested that results do not change if we consider $\alpha_{\text{HE}} = 0$ rather than $\alpha_{\text{WL}} = 0$.

The intrinsic distribution of the independent variable, $\ln M_{500}^{\text{WL}}$, was approximated with a Gaussian function of mean $\mu$ and standard deviation $\tau$, as suitable for flux selected samples of rich clusters (Andreon & Bergé 2012; Sereno, Ettori & Moscardini 2014). We tested that results based on more complex distributions, such
as mixture of Gaussian functions, were indistinguishable from the simplest case.

We chose priors as less informative as possible. We adopted uniform priors for the intercept, $\sigma_{\text{HE}}$, and the mean $\mu$. For $\sigma_{\text{WL}}^2$, $\sigma_{\text{HE}}^2$ and $\tau$, we considered an inverse Gamma distribution (Andernach & Huterer 2010). The regression was implemented with JAGS.

### 5.1 Simulated sample

As a first step, we analyzed the simulated sample from RA12. In the realm of simulations, we know the true masses of the clusters and we can exploit this information to compute the bias and the intrinsic scatter of each mass proxy. WL and HE masses can be compared to true masses autonomously. Regression results are in agreement with the original analysis in Rasia et al. (2012) and are summarized in Table 5. Note that differently from Rasia et al. (2012), we estimated the intrinsic rather than the total scatter and we focused on logarithmic variables.

The level of bias for each proxy is approximately constant with respect to the true mass, see Fig. 3.

The intrinsic scatter plays an important role when we analyze the bias as a function of the mass proxy, see Fig. 4. The decreasing logarithmic variables estimated the intrinsic rather than the total scatter and we focused on the intrinsic scatter.

Due to the combined action of selection effect and intrinsic scatters, at small (large) values of WL masses, $M_{500}^{WL}$ is biased low (high) with respect to the true mass whereas $M_{500}^{HE}$ is unbiased. As a consequence, the ratio $M_{500}^{HE}/M_{500}^{WL}$ decreases with $M_{500}^{WL}$. For intermediate values, clusters can be scattered into a given range in $M_{500}^{WL}$ from either above or below, and the ratio $M_{500}^{HE}/M_{500}^{WL}$ is not biased. The larger the scatter, the steeper the trend in the mass ratio. The scatter also flattens the $M_{500}^{WL} - M_{500}^{HE}$ relation towards larger values of $M_{500}^{WL}$.

For similar reasons and being $M_{500}^{HE}$ at the denominator, the ratio $M_{500}^{WL}/M_{500}^{HE}$ increases as a function of $M_{500}^{WL}$.

The simple statistical model summarized in Eqs. (1, 2) accounts for these effects and can be validated by comparing its predictions to the RA12 sample. We generated a sample of simulated clusters whose true masses are drawn from a Gaussian distribution. The corresponding measured WL and HE masses were simulated assuming intrinsic scatters and observational uncertainties as measured in the RA12 sample. This model successfully reproduces the trends in the observed mass ratio, see Fig. 4.

As a second step, we tested the regression algorithm with the mock observations of the simulated RA12 sample, see Table 6. Different from the first step, we used the estimated values of the mass proxies but we did not exploit the information on the true mass. We could then not calibrate the bias in either $M_{500}^{WL}$ or $M_{500}^{HE}$ in an absolute way, but we had to normalize one bias relatively to the other one. We assumed $\sigma_{\text{WL}} = 0$, i.e., we measured $M_{500}^{HE}/M_{500}^{WL}$ in units of $M_{500}^{WL}/M_{500}^{HE}$. The level of bias and the intrinsic scatters are recovered well within the statistical uncertainties.

### 5.2 Observed samples

We considered varigiate samples of clusters with observed WL and HE masses: i) the CCCP sample; ii) the CLASH sample with X-ray estimates based on either Chandra or XMM data; iii) the WTG clusters with either HE masses from B12 (‘WTG-B12’) or from L13 (‘WTG-L13’). For the CCCP sample, we could consider either masses within the same radius, i.e., $r_{500}^{WL}$, or alternatively WL masses within $r_{500}^{WL}$ and HE masses within $r_{500}^{HE}$.

Results for the real clusters are summarized in Table 6. We relied on the CCCP sample or by $\lesssim 25$–35 per cent if we consider the WTG estimates. The results for the CLASH sample depend on the X-ray analysis. The bias is $\sim 10$ per cent for Chandra data and $\sim 30$ per cent for XMM data.

The difference in the level of the bias among the various samples reflects the different absolute mass calibrations in the WL and the X-ray samples, see Sec. 4. The bias ascertained with either the WTG or the CLASH-XMM sample is in agreement with results from numerical simulations whereas results based on the CCCP and CLASH-CXO slightly under-estimate it.

Apart from the overall normalization, results from different data-sets are qualitatively and quantitatively consistent. The intrinsic scatter on WL masses is of order of $\sim 10$–15 per cent, in very good agreement with numerical simulations. On the other hand, the estimated scatter on HE masses is $\sim 25$ per cent, a factor of two larger than theoretical predictions. The large value of $\sigma_{\text{HE}}$ is evident in the plots of $M_{500}^{WL}/M_{500}^{HE}$ versus the HE mass. The observed ratio increases much more steeply than the predictions from simulations.

For the CCCP and the CLASH samples, we could restrict the analysis to either cool-core (CC) clusters or systems with low offsets between the BCG and the X-ray surface brightness peak. The HE mass of CC clusters in the CCCP sample is less biased. Due to the limited number of clusters, we could not confirm this result for the CLASH sample. There is no evident trend for the low-offsets clusters.

We verified a posteriori how well the statistical model reproduces the observed trends of the mass ratio, see Figs. 5, 6. We generated a number of ‘true’ masses from the normal distribution derived in the regression analyses and the corresponding ‘true’ simulated WL and HE masses, scattered according to the measured $\sigma_{\text{WL}}$ and $\sigma_{\text{HE}}$. The simulated WL and HE masses were finally generated considering the measured statistical uncertainties. These simulated masses were then binned as the real clusters. The observed trends in bias and scatter are well recovered.

### 6 DISCUSSION

The main sources of bias and scatter in WL mass measurements are due to the presence of substructures and triaxiality (Rasia et al. 2012; Giocoli et al. 2014). These effects are dominated by the dark matter component and are more easily reproduced in numerical simulations than the more complex processes involving gas physics. Reassuringly, the level of scatter we ascertained from observations is in very good agreement with the theoretical prediction of $\sigma_{\text{WL}} \lesssim 10$ per cent.

The relative bias we measured between WL and X-ray masses
Table 6. Biases and intrinsic scatters of the WL and the HE mass. Col. 1: sample; col. 2: number of clusters in the sample, $N_{cl}$; cols. 3, 4: radius within which the WL lensing and the HE mass were computed; col. 5: effective ratio between the true mass and the WL mass; the WL mass is assumed to be an unbiased proxy; col. 6: intrinsic scatter on $\ln M_{500}^{WL}$/$M_{500}^{HE}$; cols. 7, 8: effective ratio $M_{500}^{HE}$/$M_{500}^{WL}$ and intrinsic scatter (as in col. 6). Quoted values are bi-weight estimators of the posterior probability distribution.

| Sample         | $N_{cl}$ | $r_{500}^{WL}$ | $r_{500}^{HE}$ | $M_{500}^{WL}$/$M_{500}^{HE}$ | $\sigma_{WL}$ | $M_{500}^{HE}$/$M_{500}^{WL}$ | $\sigma_{HE}$ |
|----------------|----------|----------------|----------------|--------------------------------|---------------|--------------------------------|---------------|
| RA12           | 60       | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.14 \pm 0.04$ | $0.75 \pm 0.03$                | $0.13 \pm 0.04$ |
| CCCP           | 50       | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.14 \pm 0.06$ | $0.85 \pm 0.05$                | $0.24 \pm 0.07$ |
| CCCP-Cool Core | 16       | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.18 \pm 0.10$ | $0.93 \pm 0.11$                | $0.24 \pm 0.12$ |
| CCCP-Low Offset| 20       | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.18 \pm 0.10$ | $0.82 \pm 0.09$                | $0.30 \pm 0.11$ |
| CCCP           | 50       | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.20 \pm 0.09$ | $0.81 \pm 0.07$                | $0.45 \pm 0.07$ |
| CLASH-CXO      | 20       | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.17 \pm 0.09$ | $0.78 \pm 0.09$                | $0.34 \pm 0.12$ |
| CLASH-CXO-Cool Core | 9    | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.22 \pm 0.14$ | $0.77 \pm 0.14$                | $0.31 \pm 0.17$ |
| CLASH-CXO-Low Offset | 8   | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.31 \pm 0.17$ | $0.70 \pm 0.15$                | $0.34 \pm 0.17$ |
| CLASH-XMM      | 16       | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.17 \pm 0.10$ | $0.56 \pm 0.08$                | $0.45 \pm 0.14$ |
| WTG-L13        | 14       | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.32 \pm 0.14$ | $0.64 \pm 0.09$                | $0.16 \pm 0.08$ |
| WTG-B12        | 14       | $r_{500}^{WL}$ | $r_{500}^{HE}$ | 1                              | $0.19 \pm 0.12$ | $0.47 \pm 0.07$                | $0.34 \pm 0.15$ |
is consistent within the statistical errors with predictions from simulations (Rasia et al. 2012). Based on a suite of different numerical simulations (Battaglia et al. 2012; Kay et al. 2012; Planck Collaboration et al. 2013b) estimated $b = 1 - M_{\text{HE}}/M_{500} = 0.2^{+0.2}_{-0.2}$. However, the inherent uncertainty in the HE/WL calibrations prevents firmer conclusions.

The measured intrinsic scatter in X-ray masses is notably larger than the theoretical prediction. The disagreement might hinge on several plausible causes. The formal statistical uncertainty in X-ray mass estimates is usually of the order of $\sim 10–15$ per cent. However, the observed discrepancies among mass estimates from independent analyses are as large as 45 per cent ($\sim 30$ per cent within the same physical radius). The under-estimation of the formal error on the HE masses could determine an over-estimation of the intrinsic scatter.

Secondly, scatter in simulations might be under-estimated due to their current limits (Rasia et al. 2014). Estimates of bias and scatter from numerical simulations are still uncertain, showing dependences on the physical treatment of the gas, and, possibly, on the hydrodynamical scheme adopted. Each simulation suite has well defined prescriptions for gas physics. Different treatments of radiative cooling and heating from the UV background play an important role. Thermal conduction in hot clusters may be effective in removing cold blobs and in making the thermal structure of the ICM more homogeneous. This leads to an increase of the spectroscopic temperature and therefore of the hydrostatic mass. Feedback from active galactic nuclei and supernovae can significantly reduce the temperature inhomogeneity.

The impact of each ingredient is significant and each process may be more or less effective in different clusters. Theoretical predictions based on specific descriptions may then significantly under-estimate the intrinsic scatter in the HE mass.

Some disagreement among theoretical predictions is also caused by the adopted simulation scheme. Smoothed-particle-hydrodynamics (SPH) simulations produce larger temperature variations connected to the persistence of both substructures and their stripped cold gas than adaptive-mesh-refinement (AMR) codes (Stracki, Springel & Haehnelt 2011; Rasia et al. 2014), which lead to a more efficient mixing of gas entropy. Low-entropy gas residing in high-density clumps is more efficiently mixed to the high-entropy ICM than in SPH simulations. The simulated temperature distribution is then more homogenous and the relative bias introduced in the estimate of X-ray temperature is smaller (Vazza et al. 2011). Around $r_{500}$, the temperature inhomogeneities of the SPH simulations can generate twice the typical hydrostatic-equilibrium mass bias of the AMR sample (Rasia et al. 2014).

These variations between simulation schemes makes predictions less certain. A better understanding of the physical processes responsible for the complex thermal structure in ICM requires improved resolution and high sensitivity observations, first of all for higher temperature systems and larger cluster-centric radii (Rasia et al. 2014).
A further source of disagreement might be ascribed to any dependence of the bias on cluster mass. Neglecting such dependence might inflate the estimate of the scatter. The massive objects are expected to be the most disturbed ones, and they should have a complex temperature structure (Rasia et al. 2012). This would imply a bias larger for the more massive clusters. We tested this hypothesis by repeating the analysis of Sec. 5 without fixing the slope $\beta_{\text{HE}}$ to unity.

Due to the addiction of a new free parameter to be determined with the regression, we could obtain well constrained results only for the two richer samples. For the CCCP sample (all masses within $r_{500}^{\text{WL}}$), we obtained $\sigma_{\text{WL}} = 0.17 \pm 0.07$, $\beta_{\text{HE}} = 1.19 \pm 0.24$ and $\sigma_{\text{HE}} = 0.20 \pm 0.09$. For the CLASH-CXO sample, we obtained $\sigma_{\text{WL}} = 0.21 \pm 0.10$, $\beta_{\text{HE}} = 1.29 \pm 0.63$ and $\sigma_{\text{HE}} = 0.30 \pm 0.16$.

For both samples, with respect to the results obtained fixing $\beta_{\text{HE}} = 1$, the measured $\sigma_{\text{WL}}$ is slightly larger whereas $\sigma_{\text{HE}}$ is smaller. However, $\sigma_{\text{HE}}$ is still larger than both $\sigma_{\text{WL}}$ and the scatter predicted by numerical simulations.

The estimated $\beta_{\text{HE}}$ is slightly larger than unity, but still consistent within the statistical uncertainty. This scenario would then imply a still larger than expected scatter in HE masses at the expense of a not so plausible bias decreasing with mass ($\beta_{\text{HE}} > 1$). This alternative scenario is then more complex but does not solve the main incongruences it was supposed to solve. Since the estimated $\beta_{\text{HE}}$ is consistent with unity within the errors, we then disfavor this scenario.

7 CONCLUSIONS

In this paper, the first in a series which aims to revise critically the status quo in measuring cluster masses and calibrating scaling relations, we studied the biases and the intrinsic scatters of weak lensing or hydrostatic masses. Either WL or HE masses determined from different groups may differ by $\sim 40$ per cent, which hinders the absolute calibration of any scaling relation and the assessment of the relative bias between WL and HE masses.

We found that the intrinsic scatter of WL masses is of the order of $\sim 10–15$ per cent, in line with theoretical predictions. The intrinsic scatter of HE turned out to be larger, $\sim 20–30$ per cent, at odds with results from numerical simulations. The discrepancies might hinge on under-estimated statistical uncertainties in HE masses. A better understanding of the physical processes responsible for the complex thermal structure in the ICM and improved simulation schemes are also required to improve the theoretical predictions.

Most of the sources of scatter in the estimates of WL and HE masses are of well known origin. The assumption of spherical symmetry causes an over or under estimate of the WL mass whether the cluster is elongated in the plane of the sky or towards the observer, respectively. Departures from hydrostatic equilibrium or the difficult assessment of non-thermal contribution to the pressure limit the accuracy of HE masses.

Over-simplified modelling inflates the intrinsic scatter. The joint analysis of multi-wavelength observations, from the X-ray to the optical band to the SZE in the radio, can provide unbiased es-

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Same as Fig. 5 but for the clusters in the CLASH sample with Chandra-based X-ray analyses. WL (HE) masses are measured within $r_{500}^{\text{WL}}$ ($r_{500}^{\text{HE}}$).}
\end{figure}
Figure 7. Same as Fig. 5 but for the clusters in the CLASH sample with XMM-based X-ray analyses. WL (HE) masses are measured within $r_{500}$ ($r_{500}^\text{HE}$).

Acknowledgements

The authors thank Lauro Moscardini for useful discussions. MS acknowledge financial contributions from contracts ASI/INAF I/023/12/0, by the PRIN MIUR 2010-2011 ‘The dark Universe and the cosmic evolution of baryons: from current surveys to Euclid’ and by the PRIN INAF 2012 ‘The Universe in the box: multiscale simulations of cosmic structure’.

References

Akritas M. G., Bershady M. A., 1996, ApJ, 470, 706
Andreon S., Bergé J., 2012, A&A, 547, A117
Andreon S., Hurn M. A., 2010, MNRAS, 404, 1922
Andreon S., Hurn M. A., 2012, ArXiv: 1210.6232
Applegate D. E. et al., 2014, MNRAS, 439, 48
Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92
Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2012, ApJ, 758, 74
Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25
Biviano A. et al., 2013, A&A, 558, A1
Bonamente M. et al., 2012, New Journal of Physics, 14, 025010
D’Agostini G., 2004, ArXiv physics/0403086
De Filippis E., Sereno M., Bautz M. W., Longo G., 2005, ApJ, 625, 108
Donahue M. et al., 2014, ArXiv: 1405.7876
Ebeling H., Edge A. C., Allen S. W., Crawford C. S., Fabian A. C., Huchra J. P., 2000, MNRAS, 318, 333
Ebeling H., Edge A. C., Mantz A., Barrett E., Henry J. P., Ma C. J., van Speybroeck L., 2010, MNRAS, 407, 83
Eddington A. S., 1913, MNRAS, 73, 359
Eddington, Sir A. S., 1940, MNRAS, 100, 354
Ettori S., 2013, MNRAS, 435, 1265
Ettori S., Gastaldello F., Leccardi A., Molendi S., Rossetti M., Buote D., Meneghetti M., 2010, A&A, 524, A68
Ettori S., Morandi A., Tozzi P., Balestra I., Borgani S., Rosati P., Lusso I., Terenzi F., 2009, A&A, 501, 61
Fabjan D., Borgani S., Rasia E., Bonafede A., Dolag K., Murante G., Tornatore L., 2011, MNRAS, 416, 801
Giocoli C., Meneghetti M., Ettori S., Moscardini L., 2012, MNRAS, 426, 1558
Giocoli C., Meneghetti M., Metcalf R. B., Ettori S., Moscardini L., 2014, MNRAS, 440, 1899
Giodini S., Lovisari L., Pointecouteau E., Ettori S., Reiprich T. H., Hoekstra H., 2013, Space Science Reviews, 177, 247
Hoekstra H., Mahdavi A., Bulbul A., Bildfell C., 2012, MNRAS, 427, 1298
Jeffreys H., 1938, MNRAS, 98, 190
Kay S. T., Peel M. W., Short C. J., Thomas P. A., Young O. E., Jeffreys H., 1938, MNRAS, 98, 190
LaRoque S. J., Bonamente M., Carlstrom J. E., Joy M. K., Nagai D., Reese E. D., Dawson K. S., 2006, ApJ, 665, 1489
Lemze D., Broadhurst T., Rephaeli Y., Barkana R., Umetsu K., 2009, ApJ, 701, 1336
Limousin M., Morandi A., Sereno M., Meneghetti M., Ettori S., Bartelmann M., Verdegu T., 2013, Space Science Reviews, 177, 155
Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, ApJ, 767, 116
Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, MNRAS, 406, 1759
Martino R., Mazzotta P., Bourdin H., Smith G. P., Bartalucci L., Marrone D. P., Finoguenov A., Okabe N., 2014, ArXiv: 1406.6831
Mazzotta P., Rasia E., Moscardini L., Tormen G., 2004, MNRAS, 354, 10
Medezinski E., Broadhurst T., Umetsu K., Oguri M., Rephaeli Y., Benitez N., 2010, MNRAS, 405, 257
Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, A&A, 514, A93
Merten J. et al., 2014, ArXiv: 1404.1376
Morandi A. et al., 2012, MNRAS, 425, 2069
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Nevalainen J., David L., Guainazzi M., 2010, A&A, 523, A22
Oguri M., Takada M., Umetsu K., Broadhurst T., 2005, ApJ, 632, 841
Piffaretti R., Valdarnini R., 2008, A&A, 491, 71
Planck Collaboration et al., ArXiv: 1303.5089
Planck Collaboration et al., ArXiv: 1303.5080
Postman M. et al., 2012, ApJS, 199, 25
Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, A&A, 498, 361
Rasia E. et al., 2014, New Journal of Physics, 16, 055018
Rines K., Diaferio A., 2006, AJ, 132, 1275
Rozo E., Bartlett J. G., Evrard A. E., Rykoff E. S., 2014a, MNRAS, 438, 78
Rozo E., Rykoff E. S., Bartlett J. G., Evrard A., 2014b, MNRAS, 438, 49
Sereno M., 2007, MNRAS, 380, 1207
Sereno M., De Filippis E., Longo G., Bautz M. W., 2006, ApJ, 645, 170
Sereno M., Ettori S., Baldi A., 2012, MNRAS, 419, 2646
Sereno M., Ettori S., Moscardini L., 2014, MNRAS, submitted
Sereno M., Ettori S., Umetsu K., Baldi A., 2013, MNRAS, 428, 2241
Sereno M., Umetsu K., 2011, MNRAS, 416, 3187
Sijacki D., Springel V., Hachnelt M. G., 2011, MNRAS, 414, 3656
Umetsu K. et al., 2014, ArXiv: 1404.1375
Vazza F., Dolag K., Ryu D., Brunetti G., Gheller C., Kang H., Pfrommer C., 2011, MNRAS, 418, 960
Voit G. M., 2005, Reviews of Modern Physics, 77, 207
von der Linden A. et al., 2014, MNRAS, 439, 2

APPENDIX A: MASSES AND COSMOLOGY

Masse estimates depend on the cosmological model. A conversion from other cosmological parameters may be required to convert to a reference model. The mass within a given cosmological over-density $\Delta$ is defined as

$$M_\Delta = \frac{4\pi}{3} \Delta \rho_c (D_\Delta \theta_\Delta)^3,$$

where $\theta_\Delta$ is the angular radius enclosing the overdensity and $D_\Delta$ is the cluster angular diameter distance.

Lensing 3D masses within a radius $r = D_\Delta \theta$, where $\theta$ is the aperture radius, scale as

$$M^{WL} \propto \Sigma_{cr}(D_\Delta \theta_\Delta)D_\Delta \theta f(\theta) \tag{A2}$$

where $\Sigma_{cr} \equiv (c^2 D_\Delta)/(4\pi G D_\Delta D_{ds})$ is the critical surface density for lensing, $\theta_\Delta$ is the angular Einstein radius and $D_\Delta$ and $D_{ds}$ are the source and the lens-source angular diameter distances, respectively. The function $f(\theta) \sim \theta^{\gamma}$ quantifies the deviation of the mass profile from the isothermal case. At $r_{500}$, mass profile are nearly isothermal, i.e., $\delta \gamma \sim 0$.

Solving for Eqs. (A1) and (A2), we obtain

$$M^{WL} \propto D_\Delta^{3+\delta \gamma} \left(\frac{D_{ds}}{D_\Delta}\right)^{1+\delta \gamma} H(z)^{-1+\delta \gamma/2} \tag{A3}$$

$$\left(\frac{D_{ds}}{D_\Delta}\right)^{-3/2} H(z)^{-1} \text{ for } \delta \gamma = 0. \tag{A4}$$

Hydrostatic masses within $\theta$ scales as

$$M^{HE} \propto D_\Delta^{3+\delta \gamma}, \tag{A5}$$

the HE mass within a given cosmological over-density is then

$$M^{HE} \propto D_\Delta^{3+\delta \gamma} H(z)^{-1+\delta \gamma/2} \tag{A6}$$

$$H(z)^{-1} \text{ for } \delta \gamma = 0. \tag{A7}$$

For $\delta \gamma = 0$, $M^{HE}/r_\Delta$ is independent of the adopted cosmology. When it was required, we used the above relations with $\delta \gamma = 0$ to make the proper conversion from different cosmological parameters.
If the intrinsic scatter is log-normal, as it is usually the case, in the above discussion \( X_{\text{Proxy}} \) can be read, for example, as \( \ln M^{\text{NL}} \) or \( \ln M^{\text{HE}} \), whereas \( X \) stands for \( \ln M^{\text{TV}} \).

**APPENDIX C: BCES**

The Bivariate Correlated Errors and Intrinsic Scatter (BCES) method is a well known regression technique with good performance for data-sets with heteroscedastic and even correlated errors on both axis as well as an intrinsic scatter (Akrivas & Bershady [1996]). The slope of the linear relation in Eq. (B5) can be estimated as

\[
\beta_Y(Y|X) = \frac{\sum_i (\tilde{Y}_i - \langle \tilde{Y} \rangle)(\tilde{X}_i - \langle \tilde{X} \rangle)}{\sum_i (\tilde{X}_i - \langle \tilde{X} \rangle)^2 - \sum_i \delta^2_{X,i}},
\]

where \( \tilde{X}_i \) (\( \tilde{Y}_i \)) is the observed values of \( X_i \) (\( Y_i \)), with associated observational uncertainty \( \delta_{X,i} \) (\( \delta_{Y,i} \)), and \( \delta_{X,Y,i} \) is the covariance between errors; \( \langle \tilde{X} \rangle \) (\( \langle \tilde{Y} \rangle \)) are the mean values.

If the covariate is scattered too, see Eq. (B1), the slope in Eq. (C1) is biased. For \( \alpha_{Pr} = 0 \) and \( \beta_{Pr} = 1 \), the unbiased slope is

\[
\beta_Y(Y|X) = \frac{\sum_i (\tilde{Y}_i - \langle \tilde{Y} \rangle)(\tilde{X}_{Pr,i} - \langle \tilde{X}_{Pr} \rangle) - \sum_i \delta_{X_{Pr,Y},i}}{\sum_i (\tilde{X}_{Pr,i} - \langle \tilde{X}_{Pr} \rangle)^2 - \sum_i (\delta^2_{X_{Pr,i}} + \sigma^2_{\beta_{Pr}})}.
\]

Since the intrinsic scatter \( \sigma_{Pr} \) is unknown, Eq. (C2) cannot be applied straightway to the data. Nevertheless, the modification of the BCES estimator can be used to quantify the bias in the slope determination when this estimator is employed for linear regression.

\[
\begin{align*}
\alpha_{Y} & = \alpha_Y + \beta_Y \mu_X, \\
\beta_{Y} & = \beta_Y + \beta_Y \mu_X, \\
\end{align*}
\]

but we have only measurements of the proxy \( X_{\text{Proxy}} \) instead of \( X \), we can not just study the relation

\[
Y = \alpha^0_Y + \beta^0_Y X_{\text{Proxy}},
\]

and take \( \beta^0_Y \) as an unbiased estimator of \( \beta_Y \). In the Gaussian case,

\[
\begin{align*}
\alpha^0_Y & = \alpha_Y + \mu_X b_g \beta_Y, \\
\beta^0_Y & = b_g \beta_Y.
\end{align*}
\]