The hydrostatic-to-lensing mass bias from resolved X-ray and optical-IR data

M Muñoz-Echeverría, J.F Macías-Pérez, G.W Pratt, E Pointecouteau, I Bartalucci, M de Petris, A Ferragamo, C Hanser, F Kéruzoré, F Mayet, et al.

A&A, 682, A147 (2024)
https://doi.org/10.1051/0004-6361/202347584

Astronomy & Astrophysics

The hydrostatic-to-lensing mass bias from resolved X-ray and optical-IR data

M. Muñoz-Echeverría[1], J. F. Macías-Pérez[1], G. W. Pratt[2], E. Pointecouteau[3], I. Bartalucci[7], M. De Petris[6], A. Ferragamo[5,6], C. Hanser[1], F. Kéruzoré[4], F. Mayet[1], A. Moyer-Anin[1], A. Paliwal[6], L. Perotto[1], and G. Yepes[8,9]

[1] Univ. Grenoble Alpes, CNRS, LPSC-IN2P3, 53 avenue des Martyrs, 38000 Grenoble, France
e-mail: miren.munoz@lpsc.in2p3.fr
[2] Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191 Gif-sur-Yvette, France
[3] Univ. de Toulouse, UPS-OMP, CNRS, IRAP, 31028 Toulouse, France
[4] High Energy Physics Division, Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA
[5] Instituto de Astrofísica de Canarias (IAC), C/Vía Láctea s/n, 38205 La Laguna, Tenerife, Spain
[6] Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Roma, Italy
[7] INAF, IASF-Milano, Via A. Corti 12, 20133 Milano, Italy
[8] Departamento de Física Teórica and CIAFF, Facultad de Ciencias, Modulo 8, Universidad Autónoma de Madrid, 28049 Madrid, Spain
[9] Centro de Investigación Avanzada en Física Fundamental (CIAFF), Universidad Autónoma de Madrid, Cantoblanco 28049, Madrid, Spain

Received 21 July 2023 / Accepted 13 November 2023

ABSTRACT

An accurate reconstruction of galaxy cluster masses is key to use this population of objects as a cosmological probe. In this work we present a study on the hydrostatic-to-lensing mass scaling relation for a sample of 53 clusters whose masses were reconstructed homogeneously in a redshift range between $z = 0.05$ and 1.07. The M_{500} mass for each cluster was indeed inferred from the mass profiles extracted from the X-ray and lensing data, without using a priori observable-mass scaling relations. We assessed the systematic dispersion of the masses estimated with our reference analyses with respect to other published mass estimates. Accounting for this systematic scatter does not change our main results, but enables the propagation of the uncertainties related to the mass reconstruction method or used dataset. Our analysis gives a hydrostatic-to-lensing mass bias of $(1 - b) = 0.739^{+0.075}_{-0.070}$ and no evidence of evolution with redshift. These results are robust against possible subsample differences.

Key words. cosmology: observations – large-scale structure of Universe

1. Introduction

The distribution of galaxy clusters in mass and redshift is sensitive to the expansion history and matter content of the Universe, as well as to the initial conditions in the primordial Universe (Huterer et al. 2015). Thus, cluster masses are valuable tools for the use of these objects in cosmology (Vikhlinin et al. 2009; Allen et al. 2011; Costanzi et al. 2019). Recent results have shown that the cosmological analyses based on cluster number counts seem to favour more lower matter density (Ω_m) and matter power spectrum normalisation (σ_8) values than the studies based on the cosmic microwave background (CMB) power spectrum (Planck Collaboration XX 2014; Salvati et al. 2018; Costanzi et al. 2019). Given that cluster masses are not directly observable quantities and have to be estimated under several hypotheses from observations, the uncertainties and systematic errors of those estimates could be the source of tension with the CMB (Pratt et al. 2019; Salvati et al. 2020).

Some cluster number count studies (Garrel et al. 2022; Planck Collaboration XXIV 2015) have relied on cluster masses obtained from scaling relations (SRs) between the Sunyaev–Zel'dovich (SZ) effect (Sunyaev & Zeldovich 1972) or the X-ray emission of the cluster and the hydrostatic mass reconstructed from X-ray data (Arnaud et al. 2010). It has been widely investigated and proved that masses reconstructed under the hydrostatic equilibrium (HSE) hypothesis are biased low (Lau et al. 2013; Planck Collaboration XX 2014; Biffi et al. 2016; Gianfagna et al. 2021). For the cluster number count analyses based on HSE masses, the so-called HSE mass bias could be one of the possibilities to solve the mentioned Ω_m–σ_8 cosmological tension (Planck Collaboration XX 2014; Salvati et al. 2018). We define the HSE mass bias as the relative difference of HSE mass estimates to the true cluster masses, $(1 - b) = M^{\rm HSE}/M^{\rm True}$. A large value of the bias, that is, smaller $(1 - b)$, implies larger values for Ω_m and σ_8 in cluster number count analyses (Planck Collaboration XXIV 2015).

In the literature different approaches have been developed to estimate this bias. On the one hand, combined CMB power spectrum and cluster number count analyses fit the bias value that is required to get consistent results between the two probes. According to Planck Collaboration XX (2014), $(1 - b) = 0.59 \pm 0.05$ would be needed to reconcile the results from the *Planck* CMB analysis in Planck Collaboration XV (2014) with the cluster count cosmology from Planck Collaboration XX (2014). The posterior analysis of *Planck* data in Planck Collaboration XXIV (2015) obtained $(1 - b)$ values in the range between 0.54

and 0.705 considering different priors for the bias (based both on X-ray and lensing data). The updated analysis in Planck Collaboration VI (2018) provides $(1-b) = 0.62 \pm 0.03$, compatible with the $(1-b) = 0.62 \pm 0.07$ from Salvati et al. (2018). Accounting for the power spectrum of the thermal SZ (tSZ) effect together with the cluster number counts, Salvati et al. (2018) conclude that the bias needed to reconcile the CMB should be $(1-b) = 0.63 \pm 0.04$. Considering also the trispectrum in the covariance matrix of the tSZ power spectrum analysis, Bolliet et al. (2018) estimate $(1-b) = 0.58 \pm 0.06$ (68% C.L.) to be compatible with CMB data.

On the other hand, studies based on simulations have compared the HSE masses of clusters to their true masses. A large variety of simulations have been used in different works (Planck Collaboration XX 2014). Some of them computed the HSE masses by combining, under the HSE hypothesis, the true thermodynamical quantities (density, temperature, and/or pressure) from the intracluster medium in the simulations (Lau et al. 2013; Biffi et al. 2016; Gianfagna et al. 2021, 2023). Others used simulations to mimic mock observations and reconstruct the HSE masses (Rasia et al. 2012). However, they all tend to obtain a bias value of $(1-b) > 0.7$ (see Fig. 1 in Gianfagna et al. 2021, for a summary).

In an attempt to have a reliable estimate of the bias of observational HSE masses, several works have compared the HSE masses to lensing mass estimates, that is, to masses reconstructed from the lensing effect of the cluster on background sources. Under the assumption that lensing masses are unbiased estimates of the true mass of clusters, such HSE-to-lensing mass biases are good estimators of the HSE bias. Most of the studies in the literature are based on lensing masses obtained from the weak lensing signal on background galaxies. Figure 10 in Salvati et al. (2018) shows a compilation of HSE-to-lensing mass biases from different works. Despite the heterogeneity of the data and methods used in the various studies, the presented results prefer values of $M^{\rm HSE}/M^{\rm lens}$ above 0.7. Lensing mass reconstructions from a combination of weak and strong lensing data have also been used to measure the HSE-to-lensing mass bias on small samples (Penna-Lima et al. 2017; Ferragamo et al. 2022; Muñoz-Echeverría et al. 2023), obtaining $M^{\rm HSE}/M^{\rm lens}$ values that span from ∼0.6 to ∼1. The lensing of the CMB anisotropies due to the presence of clusters can also be used to estimate their mass (Melin & Bartlett 2015). A comparison of HSE and CMB lensing masses based on *Planck* data gave $1/(1-b) = 0.99 \pm 0.19$, approximately $(1-b) = 1.01^{+0.24}_{-0.16}$ (Planck Collaboration XXIV 2015). The posterior analysis in Zubeldia & Challinor (2019) jointly fitted the cosmological parameters and the HSE mass bias in the scaling relation between the SZ signal from *Planck* and cluster masses, using CMB lensing. They determined that the bias is of $(1-b) = 0.71 \pm 0.10$. According to the South Pole Telescope (SPT) data analysis in Baxter et al. (2015), the masses inferred from CMB lensing are consistent with those estimated from the SZ.

Other than lensing, some works in the literature have used the dynamical mass estimate of clusters, based on the velocity dispersion of member galaxies, to compute the bias corresponding to HSE masses (see Ferragamo et al. 2021, and references therein). According to the analysis with Sloan Digital Sky Survey (SDSS) archival data in Ferragamo et al. (2021), for the 207 galaxy clusters studied, the HSE-to-dynamical bias of *Planck* masses is $(1-b) = 0.83 \pm 0.07(\text{stat.}) \pm 0.02(\text{sys.})$. Also from optical observations, authors in Aguado-Barahona et al. (2022) measured the HSE-to-dynamical mass bias for a different sample of 297 *Planck* clusters and obtained $(1-b) = 0.80 \pm 0.04(\text{stat.}) \pm 0.05(\text{sys.})$.

In Wicker et al. (2023) authors investigated the evolution of the (total) HSE bias with mass and redshift by studying the gas mass fraction in galaxy clusters with *XMM-Newton* mass reconstructions from Lovisari et al. (2020b). The main result in Wicker et al. (2023) is that the value of the HSE bias and its dependence on mass and redshift varies significantly with the analysed cluster sample, in agreement with the conclusions in Salvati et al. (2019). However, according to Wicker et al. (2023) a value of $(1-b) \sim 0.8$ is preferred. Similarly, from the comparison of the gas fraction measured on 12 nearby clusters to the universal gas fraction value, authors in Eckert et al. (2019) concluded that the mass bias for SZ-derived estimates is $(1-b) = 0.85 \pm 0.05$, therefore, inconsistent with the bias needed to reconcile the CMB power spectrum. A different approach was taken in Hurier & Lacasa (2017), where the authors used the *Planck* galaxy cluster number counts, tSZ power spectrum, and bispectrum to constrain $(1-b) = 0.71 \pm 0.07$. This was obtained by fitting the normalisation of the SZ-mass SR, interpreting that the bias must appear in the calibration of the scaling relation. They assumed a generalised Navarro–Frenk–White (gNFW, Nagai et al. 2007; Zhao 1996) pressure profile for the gas in clusters, using the best-fitting parameter values from Arnaud et al. (2010), with the normalisation parameter computed to agree with the scaling relation in Planck Collaboration XX (2014). The choice of this particular pressure profile could affect the resulting bias value.

There are, therefore, different issues to be considered. Firstly, as stated in Planck Collaboration XXIV (2015), the main limitation of cosmological analyses with cluster number counts from SZ data is the large uncertainty on the HSE mass bias. But in spite of this large incertitude, the compilation of many studies shows that the bias values estimated with and without considering the need to reconcile CMB results have different tendencies. Such inconsistency is in line with a more general tension between results from early- and late-Universe probes (see Abdalla et al. 2022, for a review). Hence it is essential to have a deeper understanding of the HSE mass bias and its possible evolution with mass and/or redshift.

In this work, we aim to estimate the HSE-to-lensing mass bias combining single-cluster HSE and lensing M_{500} mass estimates that have been obtained by evaluating mass profiles at their corresponding radius (R_{500}). Given the large number of methods and models that can be employed to reconstruct HSE and lensing masses and the potentially different biases that they could be subjected to, we focus on a sample of clusters for which X-ray based HSE and lensing masses have been homogeneously reconstructed. The HSE masses of the homogeneous sample have been reconstructed mainly with *XMM-Newton* data and following the method described in Sect. 2.1.2. The lensing masses belong to the CoMaLit compilation of masses from the literature (Sect. 2.1.1, Sereno 2015).

As indicated by Sereno & Ettori (2015), and shown also for the well-observed CL J1226.9+3332 galaxy cluster in Fig. 1 in Muñoz-Echeverría et al. (2023), cluster mass estimates can vary up to ∼40% from one work to another. Being aware of these important differences that exist between the masses reconstructed by different studies, we gather together results from several works that have also produced estimates based on mass profiles. We use those estimates to measure the systematic dispersion with respect to our *XMM-Newton* and CoMaLit masses. In this work, we analyse a sample of clusters that spans a large redshift range, select homogeneous mass estimates, and propagate the systematic dispersion, which is one step beyond previous studies (Lovisari et al. 2020a; Sereno & Ettori 2015; Sereno et al. 2019).

This paper is structured as follows. In Sect. 2 we present the data, describing the homogeneous and comparison cluster samples. The method used to match clusters from different catalogues and the measurement of the systematic dispersion of the reference masses with respect to other estimates is described in Sect. 3. The reference sample is built in the same section. In Sect. 4 we present the HSE-to-lensing mass bias and its evolution with redshift. The scaling relation between HSE and lensing masses is obtained in Sect. 5, with all the related systematic tests in the same section. Finally, results are compared to similar works in the literature in Sect. 6 and conclusions are presented in Sect. 7. Throughout the paper 'log' corresponds to the logarithm to base 10 and 'ln' is the natural logarithm. When needed, we assume a flat ΛCDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_\mathrm{m} = 0.3$.

2. Cluster sample construction

2.1. Homogeneous sample

This study is built aiming for a clusters sample with resolved HSE and lensing masses that are comparable amongst all the objects (homogeneous reconstruction procedure) and cover the largest possible redshift range. We present in this section the mass reconstruction and regularisation procedures of the *XMM-Newton* and CoMaLit clusters, which constitute our homogeneous sample.

2.1.1. CoMaLit sample

The CoMaLit sample contains the clusters with lensing masses that we used to build the homogeneous sample. They correspond to the clusters from the Literature Catalogs of weak Lensing Clusters (LC^2) compilation presented in Sereno (2015). The LC^2 contains 806 clusters (in the 3.9 version of the LC^2-single catalogue[1]) with weak lensing masses obtained from different works in the literature, including the widely used Canadian Cluster Comparison Project (CCCP, Hoekstra et al. 2012, 2015) and Weighing the Giants (WtG, Applegate et al. 2014) clusters.

Although the masses were not derived homogeneously amongst the original works, an effort was made in Sereno (2015) to select the most comparable mass estimates. Only masses reconstructed assuming spherical symmetry were considered, clusters without optical, X-ray or SZ counterpart were excluded and when the same authors or collaborations had published several estimates for the same cluster along a refinement process, only the latest result was considered. In addition, all the masses were standardised to the same cosmology (a flat ΛCDM cosmology with $\Omega_\mathrm{m} = 0.3$ and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$) and were given at the overdensities of 2500, 500, 200 as well as at the virial radius. We define R_Δ as the radius at which the mean mass density of the cluster is Δ times the critical density of the Universe at its redshift, $\rho_\mathrm{crit} = 3H(z)^2/(8\pi G)$, with $H(z)$ the Hubble function. We consider only the masses at an overdensity of $\Delta = 500$. For some cases, the masses given in the original papers had to be extrapolated following the density profile adopted in the original paper or with a Navarro–Frenk–White (NFW, Navarro et al. 1996) model.

2.1.2. *XMM-Newton* sample with the reference X-ray pipeline

Regarding the HSE masses, we built a sample of clusters with masses reconstructed from *XMM-Newton* data and following the same procedure, hereafter *XMM-Newton* or reference X-ray pipeline. Thus, an homogeneous method was applied consistently to the full sample. This pipeline has already been used in previous works (Pratt et al. 2010; Bartalucci et al. 2017b; Ruppin et al. 2018; Kéruzoré et al. 2020; Muñoz-Echeverría et al. 2023), proving the reliability of the method.

As described in Bartalucci et al. (2017a), the X-ray raw data were processed using the standard procedures with the *XMM-Newton* Science Analysis System (SAS) pipeline. The electron density and temperature profiles were reconstructed following the correction and deprojection methods detailed in Pratt et al. (2010) and Bartalucci et al. (2018). To obtain the HSE mass profiles, the electron density and temperature were combined in the Monte Carlo procedure described in Démoclès et al. (2010). The binned HSE mass profiles were interpolated to define the M_{500} masses used in this work. Based on the same *XMM-Newton* data, two differently estimated M_{500} are available per cluster: masses derived from a X-ray calibrated scaling relation (Arnaud et al. 2010) and masses estimated from a forward NFW profile fit to the density and temperature profiles. We do not use these two masses in our main analysis, but they are employed to investigate the consistency of all three estimates in Appendix D. Amongst the clusters with *XMM-Newton* data, we distinguish three different subsamples along the redshift.

Low-z clusters: ESZ+LoCuSS. Many of the low redshift ($z < 0.5$) clusters detected by *Planck* were also observed by *XMM-Newton*. It is the case of the 62 *Planck* Early Sunyaev–Zel'dovich (ESZ) clusters (Planck Collaboration VIII 2011), whose HSE masses were reconstructed with X-ray data in Planck Collaboration XI (2011). Similarly, based on the Local Cluster Substructure Survey (LoCuSS[2]) sample, Planck Collaboration Int. III (2013) reconstructed the HSE mass of 19 clusters.

Intermediate-z clusters: LPSZ. The LPSZ stands for the NIKA2 SZ Large Programme (Mayet et al. 2020; Perotto et al. 2022). It is a high angular resolution follow-up of 45 clusters of galaxies detected with the Atacama Cosmology Telescope (ACT, Hilton et al. 2018, 2021) or the *Planck* satellite (Planck Collaboration XVIII 2015). The LPSZ follow-up combines high-resolution SZ data from the NIKA2 instrument (Adam et al. 2018; Perotto et al. 2020) with X-ray *XMM-Newton* observations and covers a redshift range between 0.5 and 0.9. Studies on individual clusters from the LPSZ sample have already been published (Ruppin et al. 2018; Kéruzoré et al. 2020; Muñoz-Echeverría et al. 2023), illustrating the joint SZ and X-ray analysis. Even though this sample was designed to be followed-up in SZ using NIKA2, we emphasise that in this work we do not make use of any SZ data for the mass estimation procedure. Instead, we consider the HSE masses obtained from *XMM-Newton* data only.

High-z clusters: Bartalucci+2018. Bartalucci et al. (2017b, 2018) were able to go beyond $z = 0.9$ and measure the HSE mass of five individual clusters from resolved mass profiles. Given the difficulties related to the high redshift of the clusters, *XMM-Newton* data were combined with *Chandra* observations. Although supplementary *Chandra* data was added, we consider these masses as homogeneous with respect to the ESZ+LoCuSS and LPSZ samples since the same reconstruction pipeline was employed. However, special care is taken in our analyses when studying the impact of these clusters. Authors

[1] http://pico.oabo.inaf.it/~sereno/CoMaLit/

[2] http://www.sr.bham.ac.uk/locuss/home.php

in Bartalucci et al. (2018) also indicate that the mass estimate for the SPT-CLJ2106−5844 cluster is not reliable, therefore, we exclude it from our analyses.

2.2. Comparison sample

The mass estimate of a cluster often varies from analysis to analysis, because of differences related to raw data or to the mass reconstruction method. In order to try to account for possible systematic biases in the CoMaLit and the reference X-ray pipeline masses, we gathered as many as possible HSE and lensing mass estimates from the literature for the clusters in our homogeneous sample. Again, we made sure that the masses in the chosen studies were measured on resolved profiles, excluding masses derived from scaling relations. We only considered HSE masses obtained from X-ray data. Comparing to HSE masses that use SZ data or scaling relations is also of great interest, but it would be an independent analysis in itself and beyond the scope of this paper (see, for example, Hoekstra et al. 2015; Sereno et al. 2017; Sereno & Ettori 2017; Schellenberger & Reiprich 2017). For lensing, in addition to the weak lensing masses, we also compared to masses reconstructed from the combination of strong and weak lensing signal.

We present in the following a brief description of this comparison sample, highlighting the distinctive characteristics of each analysis. We refer the reader to the original works for more details.

2.2.1. Ettori+2010

In Ettori et al. (2010; and the Corrigendum, Ettori et al. 2011), the authors reconstructed the HSE mass of 44 clusters with redshifts $0.092 < z < 0.307$ using *XMM-Newton* observations. They employed two different methods (M1 and M2) and gave the results in units of R_{500}. We converted the R_{500} values into M_{500} masses. The main caveat of these results is that profiles were extrapolated to reach R_{500} assuming an NFW profile. As coordinates of the assumed centres of the clusters are not given in Ettori et al. (2010), we took them from Yuan et al. (2022)[3] and when missing, from the 4XMM-DR9 source list[4].

2.2.2. Landry+2013

In Landry et al. (2013) the HSE masses of 35 clusters with redshifts between $0.152 < z < 0.3017$ were obtained using *Chandra* data. Two different mass estimates are given in the paper: either using the Vikhlinin model or the polytropic equation of state. According to the authors, the profiles of seven clusters required 'slight' extrapolation to reach R_{500}. Again, the coordinates of the assumed centres of the clusters are not given in Landry et al. (2013), so most of coordinates were taken from Ebeling et al. (1998). When missing, position coordinates of clusters were found by querying in the Simbad-CDS portal[5] with the cluster name given in Table 1 in Landry et al. (2013).

2.2.3. LoCuSS

The aforementioned LoCuSS sample contains in all 50 clusters, with $0.152 < z < 0.3$ (Smith et al. 2015). For our mass comparisons, we used the LoCuSS HSE masses published in Martino et al. (2014) and the lensing masses from Okabe & Smith (2016). The HSE masses were reconstructed with *Chandra* data for 43 clusters and with *XMM-Newton* observations for 39. For some clusters both estimates are available. Central coordinates of clusters were also taken from Martino et al. (2014). The analysis in Zhang et al. (2010) studied 12 out the 50 clusters with *XMM-Newton* and Subaru data. The lensing masses published in Zhang et al. (2010) are equivalent to those in Okabe & Smith (2016), but the HSE mass profiles were evaluated at the R_{500} corresponding to the lensing analyses. We, therefore, gave preference to the results in Okabe & Smith (2016) and Martino et al. (2014) and restricted the LoCuSS masses to the estimates in the latter two studies.

2.2.4. Mahdavi+2008

Uniformly estimated masses of 18 clusters were published in Mahdavi et al. (2008). Lensing masses were obtained as in Hoekstra (2007), but with the photometric redshift distributions from Ilbert et al. (2006). The lensing mass reconstruction was done with a method based on aperture mass estimation, that is, obtaining first projected masses, and subsequently deprojecting by assuming an NFW density model and the concentration-mass scaling relation from Bullock et al. (2001). For the HSE masses, *Chandra* observations were used. As indicated in Table 2 in Mahdavi et al. (2008), for 14 out of the 18 clusters the HSE masses at R_{500} were obtained from extrapolation and all of them were measured at the lensing R_{500}.

2.2.5. Mahdavi+2013

In Mahdavi et al. (2013) authors studied a sample of 50 clusters with redshift $0.152 < z < 0.55$. The clusters correspond to the CCCP sample. The HSE masses were reconstructed from a combined analysis of *XMM-Newton* and *Chandra* data. For the same sample, lensing estimates were obtained in Hoekstra et al. (2012), using CFH12k and Megacam data from the Canada-France-Hawaii Telescope. HSE masses were measured at the R_{500} obtained from lensing masses.

2.2.6. Israel+2014

The analysis in Israel et al. (2014) contains eight clusters with redshift $0.35 < z < 0.80$. The lensing masses were obtained from an NFW fit to the tangential shear profiles of clusters, assuming a mass-concentration relation. To reconstruct the HSE mass, the authors used the electron density profiles of individual clusters, which were estimated from *Chandra* surface brightness maps. The temperature profile of individual clusters being more challenging to obtain, the authors combined the *Chandra* data of all clusters in the sample to reconstruct a single global temperature profile for the whole sample. The HSE masses in Israel et al. (2014) were also evaluated at the R_{500} measured from lensing mass profiles.

2.2.7. LPSZ+CLASH

Within the LPSZ programme, Muñoz-Echeverría et al. (2022, 2023) and Ferragamo et al. (2022) estimated the lensing mass for three clusters in the sample in common with the Cluster Lensing And Supernova survey with *Hubble* (CLASH, Postman et al. 2012; Zitrin et al. 2015, 2013a,b, 2009). Masses were reconstructed by fitting a projected NFW mass density profile to

[3] http://zmtt.bao.ac.cn/galaxy_clusters/dyXimages/newton.html
[4] http://xmmssc.irap.omp.eu/Catalogue/4XMM-DR9/4xmmdr9_obslist.html
[5] http://simbad.u-strasbg.fr/simbad/

Table 1. Summary of the amount of clusters in the each of the comparison samples and their overlap with the homogeneous *XMM-Newton* and CoMaLit clusters.

Sample	Redshift	Type of mass	# of clusters	# of clusters in common with the *XMM-Newton* sample	# of clusters in common with the CoMaLit sample
Ettori+2010	$0.092 < z < 0.307$	HSE	44	24	
Landry+2013	$0.152 < z < 0.3017$	HSE	35	19	
LoCuSS	$0.152 < z < 0.3$	HSE and lensing	50	22	45
Mahdavi+2013	$0.152 < z < 0.55$	HSE and lensing	50	18	44
Mahdavi+2008	$0.170 < z < 0.547$	HSE and lensing	18	11	17
Israel+2014	$0.35 < z < 0.80$	HSE and lensing	8	0	8
LPSZ+CLASH	$0.55 < z < 0.89$	Lensing	3		3
Bartalucci+2018	$0.933 < z < 1.066$	Lensing	4		3
All				94	120
All without repetition				36	82

Notes. We also report the total amount of matches, that is, the data points in Fig. 1 and the number of different objects.

the publicly available CLASH convergence maps (Zitrin et al. 2015). Given that two differently modelled convergence maps were provided, for some clusters two lensing mass estimates are available, named LTM and PIEMD+eNFW following the name of the method used to reconstruct the convergence maps. We also considered the lensing masses published in Umetsu et al. (2014) and Merten et al. (2015) for the same clusters.

2.2.8. Bartalucci+2018

In Bartalucci et al. (2018) authors studied the HSE-to-lensing mass bias of five SPT clusters. The weak lensing masses were obtained in Schrabback et al. (2017) using *Hubble* Space Telescope (HST) observations. The profiles were centred in the X-ray peak or the SZ peak (indicated in Table 1 in Schrabback et al. 2017), giving two different lensing mass estimates per cluster.

3. Selection and characterisation of the sample

In this section, we present the comparison of the *XMM-Newton* and CoMaLit mass estimates to the results from other works presented in Sect. 2.2. We briefly describe the procedure used to match and select clusters from different catalogues, and then quantify the scatter based on the comparisons of several mass measurements for each cluster across our sample. Finally, we build the reference sample with the *XMM-Newton* and CoMaLit masses that we use for the rest of the analysis.

3.1. Matching clusters

We matched clusters from different catalogues on the basis of their coordinates. We considered that two entries in two distinct catalogues correspond to the same cluster for angular separations smaller than 400″. We further verified every match by checking the redshifts given in the different catalogues. We identified suspicious mismatching between A1606 ($z = 0.0963$) and A2029 ($z = 0.078$) and excluded it.

At the same time, we discarded clusters that appear as one object in some catalogue and as a combination of multiple substructures in another. For example, the cluster A1758 in Landry et al. (2013) has four entries in the LC^2-single catalogue: A1758S, A1758NW, A1758N, A1758NE. Similarly, we excluded A222, A223N, and A223S. In addition, we identified and discarded A750 (present in CoMaLit, LoCuSS, Mahdavi+2013, and Mahdavi+2008 catalogues), whose mass estimate can not be reliable since it is superimposed along the line of sight with MS0906+11 (Geller et al. 2014).

We summarise in Table 1 the overlap between the homogeneous clusters in *XMM-Newton* and CoMaLit samples and those from other works presented in Sect. 2.2. For 36 of the *XMM-Newton* and 82 of the CoMaLit clusters we identified other HSE and lensing mass estimates[6].

3.2. Estimation of systematic dispersion

We present in the left panel in Fig. 1 the relation between X-ray HSE masses obtained with the reference X-ray pipeline (homogeneous masses) with respect to other X-ray HSE masses from the literature (comparison sample). In the right panel, we show the relation between lensing masses from different works with respect to the estimates summarised in CoMaLit. Each colour represents one of the samples described in Sect. 2.2 and different estimates of the same work are differentiated with markers. The black dashed line shows the one-to-one relation.

Overall, the agreement between the samples is reasonable, with a significant dispersion around the 1:1 relation. We identify some clusters for which the mass estimates differ significantly. These are Abell521, Abell2390, Abell2163 in X-rays and RXJ1347.5−1145, CL1641, CL1701 in the lensing masses comparison. The cluster shown with the green marker on top of the left panel in Fig. 1 is Abell2390 and, despite its departure from the 1:1 relation, we do not have strong arguments for excluding it. For lensing masses (in the right panel in Fig. 1) there seem also to be a hint of some bias that we do not propagate hereafter. We verified that the bias does not correlate with a comparison sample in particular, but rather with high mass clusters. Further investigation would be needed to understand this trend.

To quantify the systematic dispersion with respect to the 1:1 relation, we followed Eqs. (3) and (4) in Pratt et al. (2009). For the $N_{\rm lens} = 120$ matched entries between the CoMaLit catalogue and the other lensing samples (Table 1), we defined the raw

[6] Since the LC^2 catalogue is a compilation of masses from many works in the literature, it is not surprising that some CoMaLit masses are directly the estimates published in other works. It is the case for some LoCuSS clusters.

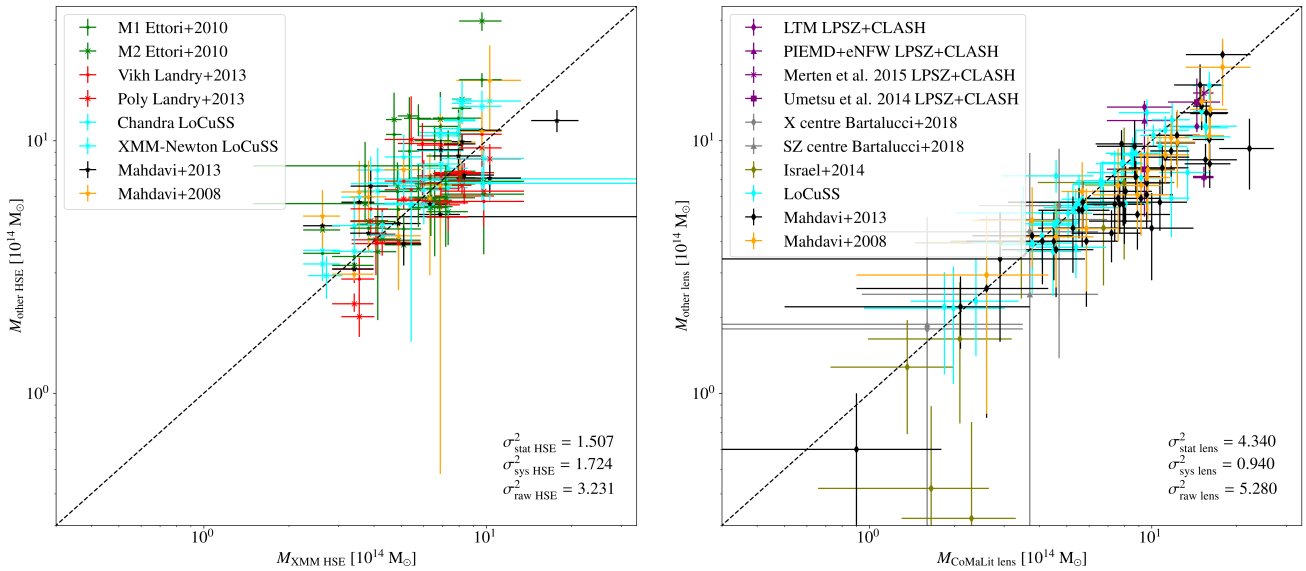

Fig. 1. Relation between HSE (left) and lensing (right) masses from the homogeneous samples in this work (*XMM-Newton* and CoMaLit) with respect to other estimates from the literature (comparison sample). Each colour indicates a different analysis and several results from the same work are differentiated by using different markers. The black dashed lines show the one-to-one relation. We give the statistical, systematic, and raw variances as defined in the text. All the variances are in units of $(10^{14}\ M_\odot)^2$.

variance as

$$\sigma^2_{\rm raw\ lens} = \frac{1}{N_{\rm lens} - 2} \sum_{i=1}^{N_{\rm lens}} w_i (M_{\rm other\ lens} - M_{\rm CoMaLit\ lens})^2, \quad (1)$$

where w_i is the weight of each cluster and $M_{\rm CoMaLit\ lens}$ and $M_{\rm other\ lens}$ are the lensing mass in the CoMaLit catalogue and in a different analysis, respectively. The weight given to each cluster is

$$w_i = \frac{1/\sigma_i^2}{1/N \sum_{j=1}^{N} 1/\sigma_j^2}, \quad (2)$$

using $\sigma_i^2 = \delta^2_{M_{\rm other\ lens}} + \delta^2_{M_{\rm CoMaLit\ lens}}$, the sum of the uncertainties related to each cluster. The $\sigma^2_{\rm raw\ HSE}$ was measured in an equivalent way using the HSE masses and uncertainties of each cluster, $M_{\rm XMM\ HSE}$ and $M_{\rm other\ HSE}$ and $\delta^2_{M_{\rm XMM\ HSE}}$ and $\delta^2_{M_{\rm other\ HSE}}$.

The statistical error associated to the masses was obtained for both lensing and X-ray masses with the above-mentioned weight, w_i, and σ_i^2:

$$\sigma^2_{\rm stat} = \frac{1}{N-2} \sum_{i=1}^{N} w_i \sigma_i^2 = \frac{1}{N-2} \sum_{i=1}^{N} \frac{N}{\sum_{j=1}^{N} 1/\sigma_j^2}. \quad (3)$$

This allowed us to define the systematic scatter, that is, the excess of scatter in the raw variance not explained by the statistical uncertainties, as

$$\sigma^2_{\rm sys} = \sigma^2_{\rm raw} - \sigma^2_{\rm stat}. \quad (4)$$

We report in Fig. 1 (and in Table A.1) the statistical, systematic, and raw scatter for the HSE and lensing masses. The raw dispersion of lensing masses ($\sigma^2_{\rm raw\ lens} = 5.280 \times (10^{14}\ M_\odot)^2$) is larger than HSE ones ($\sigma^2_{\rm raw\ HSE} = 3.231 \times (10^{14}\ M_\odot)^2$) and the uncertainties of individual lensing masses being larger, the statistical dispersion is also larger ($\sigma^2_{\rm stat\ lens} = 4.340 \times (10^{14}\ M_\odot)^2$ and $\sigma^2_{\rm stat\ HSE} = 1.507 \times (10^{14}\ M_\odot)^2$). Nevertheless, the error bars of HSE masses are not large enough to cover the excess of scatter around the 1:1 relation, making the systematic scatter for HSE masses ($\sigma^2_{\rm sys\ HSE} = 1.724 \times (10^{14}\ M_\odot)^2$) larger than for lensing ($\sigma^2_{\rm sys\ lens} = 0.940 \times (10^{14}\ M_\odot)^2$).

As mentioned in the description of each sample in Sect. 2.2, HSE masses in some works were evaluated at the R_{500} obtained from lensing. We checked the impact of excluding such estimates from the analysis. In the left panel in Fig. A.1, we present the relation between the *XMM-Newton* reference pipeline masses and X-ray masses from the comparison sample without accounting for $M^{\rm HSE}(< R^{\rm lens}_{500})$ estimates (that is, without Mahdavi et al. 2008, 2013). The statistical, raw, and systematic variances change by 0.4, 5, and 10%, respectively. Hence, taking a conservative approach, in the following sections we consider the largest systematic scatter values obtained.

3.3. Reference sample

Following the procedure described in Sect. 3.1, we matched the clusters in the CoMaLit catalogue (Sect. 2.1.1) with the clusters with HSE masses from the *XMM-Newton* reference pipeline (Sect. 2.1.2) and obtained an homogeneous sample composed of 65 clusters. Amongst the 65 clusters, 54 correspond to the ESZ+LoCuSS samples, eight clusters are from the LPSZ, and three from Bartalucci+2018. For these clusters, we performed additional checks and discarded: three clusters with senseless error bars (see Appendix B.1 for details), and ten clusters (one of them already excluded) for which X-ray and lensing mass reconstruction analyses had assumed very distant centres (see Appendix B.2).

As a result, our reference sample contains 53 clusters with homogeneous HSE and lensing masses that can be used for comparisons (see Table B.2). We present in Fig. 2 a summary of the characteristics of the sample. The histograms in the left show the number of clusters with respect to redshift, HSE mass, and lensing mass. The right panel in Fig. 2 presents the clusters in the mass–redshift plane. While very few works in the literature go above $z = 0.5$, 20% of the clusters in our sample have

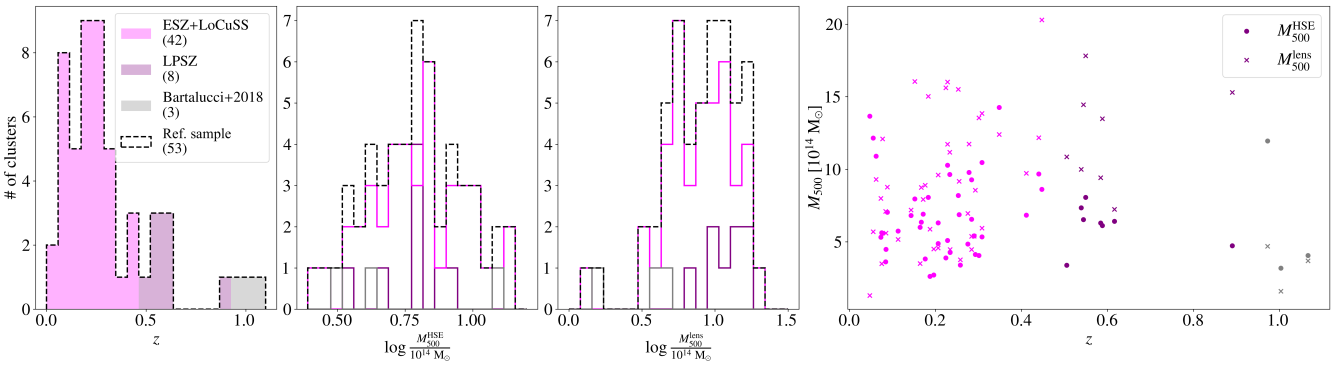

Fig. 2. Main characteristics of the 53 clusters in the reference sample. Histograms in the left panels show the redshift, HSE mass, and lensing mass distributions. We show in purple, magenta, and grey the distributions for ESZ+LoCuSS, LPSZ, and Bartalucci+2018 clusters, respectively. The black dashed lines represent the distributions of the whole sample. In the right panel we show the HSE and lensing masses as a function of redshift for all the clusters.

redshifts higher than 0.5. However, the distribution in redshift of the sample is dominated by low-z clusters.

After excluding the last 12 clusters (in Appendices B.1 and B.2) from the *XMM-Newton* and CoMaLit samples, we recalculated the scatter with respect to other HSE and lensing masses. Compared to Fig. 1, the raw, statistical, and systematic dispersions remain of the same order, but the impact of individual clusters is again noticeable in the resulting values (less than 10% of change, see Fig. A.1 and Table A.1). Therefore, we took the most conservative approach and considered that the systematic scatters to be accounted for in the *XMM-Newton* and CoMaLit masses are the largest values we have found: $\sigma^2_{\rm sys\ lens} = 1.202 \times (10^{14}\ M_\odot)^2$ and $\sigma^2_{\rm sys\ HSE} = 2.017 \times (10^{14}\ M_\odot)^2$. We note that the clusters used for these calculations are not necessarily the 53 in our reference sample, but the ones in common between *XMM-Newton* and other X-ray samples and between CoMaLit and other lensing works (summarised in Table 1). We compare in Fig. B.3 the systematic standard deviation values to the individual statistical uncertainties of the masses from the *XMM-Newton* reference pipeline and the CoMaLit catalogue.

In the following sections, we investigate how the HSE-to-lensing mass bias and scaling relation change when accounting for these systematic scatters. In order to propagate the scatters to the final results, we consider that the uncertainties in the mass of each cluster are the quadratic sum of the measurement statistical uncertainties and the systematic scatters derived in this section. Thus, we have

$$\delta^2_{\rm lens} = \delta^2_{M_{\rm CoMaLit\ lens}} + \sigma^2_{\rm sys\ lens} \tag{5}$$

for the lensing masses, and

$$\delta^2_{\rm HSE} = \delta^2_{M_{\rm XMM\ HSE}} + \sigma^2_{\rm sys\ HSE} \tag{6}$$

for the hydrostatic ones.

This is a very conservative approach that assumes that the mass estimates from the X-ray reference pipeline and the CoMaLit catalogue may have an additional error (due to, for example, the used dataset or the mass reconstruction method) that can be quantified from the distance to other estimates. Such supplementary error is usually not considered in the literature. For this reason, we also perform the study without accounting for the systematic uncertainties. An alternative approach was considered in Sereno & Ettori (2015) by separating the analysis in subsamples. A cross-validation of our results by subsamples is presented in Sect. 5.4.

4. Direct HSE-to-lensing mass bias measurement

The bias of HSE masses with respect to lensing estimates is defined from the ratio of the masses,

$$(1 - b_{\rm HSE/lens}) = M^{\rm HSE}_{500}/M^{\rm lens}_{500}. \tag{7}$$

For simplicity, in the rest of this paper we name the HSE-to-lensing mass bias without subscripts $b = b_{\rm HSE/lens}$.

As a first approach, and for comparison with other works in the literature, we directly compare the HSE-to-lensing mass ratio among the clusters of the reference sample. Following the parametrisation in Salvati et al. (2019) and Wicker et al. (2023), we describe the redshift evolution of the HSE-to-lensing mass bias as

$$M^{\rm HSE}_{500}/M^{\rm lens}_{500}(z) = (1 - b)(z) = (1 - \mathcal{B})\left(\frac{1+z}{1+z_*}\right)^{\beta_z}, \tag{8}$$

where $(1 - \mathcal{B})$ is the bias normalised at the pivot redshift, z_*, and β_z describes the evolution with redshift. As in Salvati et al. (2019), we take z_* the median redshift value of the clusters in the analysed sample. In Wicker et al. (2023) the pivot redshift is the mean of the sample.

With the homogeneous HSE and lensing masses of the 53 clusters in the reference sample, we perform a Markov chain Monte Carlo (MCMC) analysis to fit the model (Eq. (8)) to data, using the `emcee` Python package (Foreman-Mackey et al. 2019; Goodman & Weare 2010). We consider uniform priors for the parameters, $(1 - \mathcal{B}) \sim \mathcal{U}(0, 2)$ and $\beta_z \sim \mathcal{U}(-8, 8)$, and assume a Gaussian likelihood, uncorrelated between points.

We show in Fig. 3 the HSE-to-lensing mass ratio as a function of redshift for the 53 clusters in the reference sample. Here error bars include systematic scatter following Eqs. (5) and (6). The grey shaded area in the top panel indicates the 16th to 84th percentile region of the bias evolution model obtained from the posterior distributions of the fitted parameters. For comparison, the horizontal lines show the mean (dash-dotted line), median (dotted line), and error weighted mean (solid line) HSE-to-lensing mass ratio obtained with the 53 cluster masses. Posterior distributions of the fitted parameters are shown with grey contours in Fig. 4. The best-fit values and uncertainties are given in the first row in Table 2. We give $(1 - \mathcal{B})/(1 + z_*)^{\beta_z}$, which is the value of the bias at $z = 0$. We also report the results without accounting for the systematic scatter of the lensing and HSE masses. As expected, when accounting for $\sigma^2_{\rm sys}$ the uncertainties of the posterior distributions are enlarged.

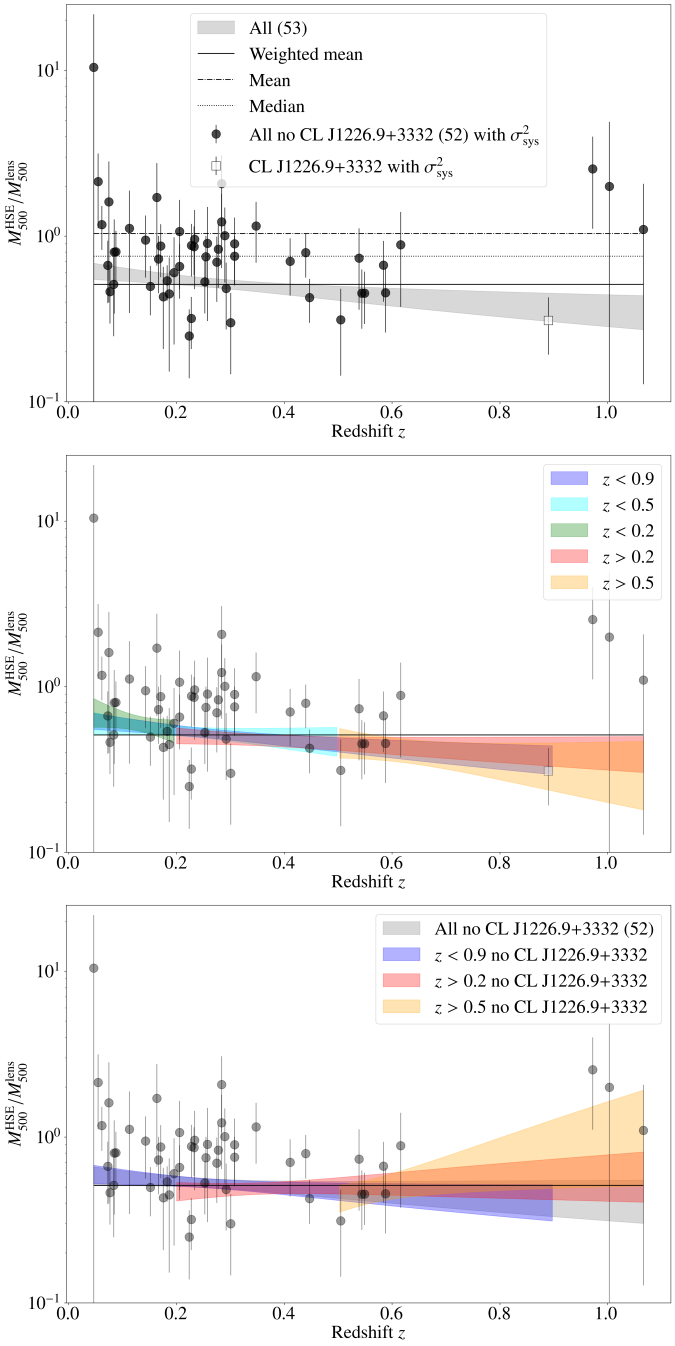

Fig. 3. HSE-to-lensing mass ratio with respect to the redshift. Markers with error bars show the ratio of each cluster in the reference sample with error bars accounting for the systematic uncertainty. Horizontal solid, dotted, and dash-dotted black lines give respectively the error weighted mean, median, and mean mass ratio for the data points. Shaded areas represent the 16th to 84th percentiles of the bias evolution model obtained by fitting different redshift ranges. Top: the bias evolution model obtained with the 53 clusters in the reference sample. Centre: different colours indicate the models fitted to clusters in different redshift ranges. Bottom: grey, blue, red, and orange shaded areas show respectively the bias evolution model fitted to clusters along all the redshift range, at $z < 0.9$, at $z > 0.2$, and at $z > 0.5$, excluding in all the cases the CL J1226.9+3332 galaxy cluster.

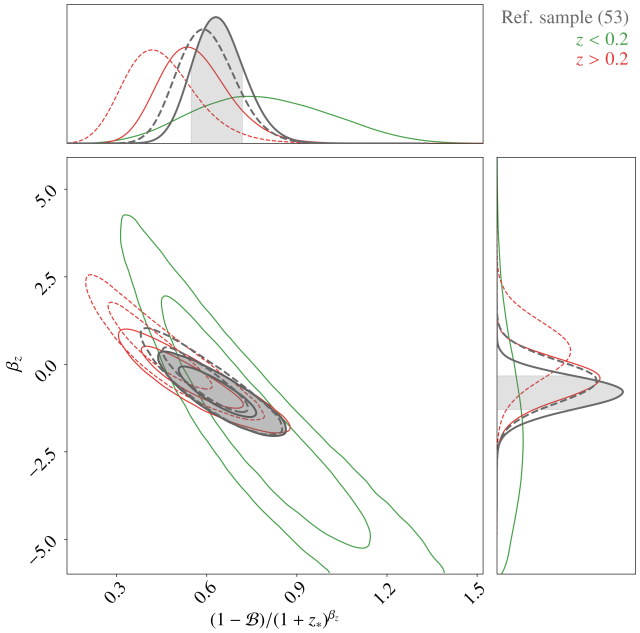

Fig. 4. One-dimensional and two-dimensional posterior distributions of the parameters in the redshift dependent mass bias model, accounting for the σ^2_{sys} in the error bars. Different colours describe the results for the various samples presented in Table 2. For a good visualisation, we only show in grey, green, and red the results for the whole sample, the $z < 0.2$, and the $z > 0.2$ ranges, respectively. Dashed distributions have been obtained excluding CL J1226.9+3332 galaxy cluster.

Due to the significant differences in the mass uncertainties and the non-uniform distribution of the clusters in redshift, certain subsamples might be driving the fit of the model. To check for these effects and investigate any dependence with redshift, we repeat the fit by considering clusters in different redshift ranges.

Considering only the clusters with $z < 0.9$ (that is, those in ESZ+LoCuSS and LPSZ samples) and only those with $z < 0.5$ (only ESZ+LoCuSS), the results are very close to the ones obtained with the reference sample. This means that the grey result in Figs. 3 and 4 is most probably dominated by ESZ+LoCuSS clusters. Best-fit values and uncertainties for these two cases are given in Table 2. The corresponding bias evolution models are shown in blue ($z < 0.9$) and cyan ($z < 0.5$) in the central panel in Fig. 3.

We find more significant differences when considering only low redshift clusters ($z < 0.2$, in green), or, when discarding them ($z > 0.2$, in red). For low redshift clusters, the HSE masses at $z = 0$ are less biased with respect to lensing masses ($(1 - \mathcal{B})/(1 + z_*)^{\beta_z}$ closer to 1) than for the reference sample, but the dependence on redshift is stronger. Exactly the opposite happens when fitting only $z > 0.2$ masses: the HSE-to-lensing mass bias is larger at $z = 0$ (smaller $(1 - \mathcal{B})/(1 + z_*)^{\beta_z}$), but the redshift evolution is weaker (the absolute value of β_z smaller). These conclusions agree with the results in Wicker et al. (2023), where the same cut in redshift is adopted. In Smith et al. (2015), the authors also reported a different tendency for *Planck* cluster masses depending on the redshift, with a larger HSE-to-lensing bias value (smaller $1 - b$) for *Planck* masses at $z > 0.3$, than the bias at $z < 0.3$. However, these masses were inferred from the SZ-mass scaling relation and not measured from profiles. Nonetheless, in our analysis β_z is compatible with no redshift evolution both for $z < 0.2$ and $z > 0.2$ subsamples (see posterior probability density contours in Fig. 4).

As shown in Fig. 2, the clusters at high redshift are rare in our sample, with a large gap between $z = 0.62$ and $z = 0.89$. Only CL J1226.9+3332, SPT-CLJ0615−5746,

Table 2. Best-fit values and uncertainties for the normalisation and redshift evolution parameters of the mass bias model in Eq. (8) obtained for different subsamples of the reference sample.

Sample	# of clusters	z_*	No $\sigma^2_{\rm sys}$		With $\sigma^2_{\rm sys}$	
			$(1-\mathcal{B})/(1+z_*)^{\beta_z}$	β_z	$(1-\mathcal{B})/(1+z_*)^{\beta_z}$	β_z
Reference sample	**53**	**0.253**	$0.585^{+0.059}_{-0.050}$	$-0.797^{+0.309}_{-0.373}$	$\mathbf{0.632^{+0.093}_{-0.074}}$	$\mathbf{-0.787^{+0.418}_{-0.529}}$
$z < 0.9$	50	0.234	$0.591^{+0.060}_{-0.050}$	$-0.846^{+0.306}_{-0.381}$	$0.642^{+0.094}_{-0.079}$	$-0.860^{+0.427}_{-0.528}$
$z < 0.5$	42	0.215	$0.578^{+0.095}_{-0.082}$	$-0.744^{+0.724}_{-0.773}$	$0.618^{+0.130}_{-0.112}$	$-0.661^{+0.887}_{-0.948}$
$z < 0.2$	19	0.113	$0.716^{+0.159}_{-0.144}$	$-1.577^{+1.589}_{-1.560}$	$0.802^{+0.216}_{-0.235}$	$-2.226^{+2.515}_{-1.993}$
$z > 0.2$	34	0.305	$0.471^{+0.071}_{-0.057}$	$-0.271^{+0.344}_{-0.424}$	$0.543^{+0.131}_{-0.092}$	$-0.392^{+0.496}_{-0.668}$
$z > 0.5$	11	0.588	$0.665^{+0.620}_{-0.279}$	$-1.043^{+1.020}_{-1.355}$	$0.692^{+1.267}_{-0.356}$	$-0.987^{+1.380}_{-2.249}$
Ref. no CL J1226.9+3332	52	0.244	$0.548^{+0.069}_{-0.060}$	$-0.467^{+0.453}_{-0.505}$	$0.594^{+0.100}_{-0.086}$	$-0.483^{+0.562}_{-0.639}$
$z < 0.9$ no CL J1226.9+3332	49	0.233	$0.560^{+0.071}_{-0.059}$	$-0.578^{+0.443}_{-0.517}$	$0.610^{+0.104}_{-0.085}$	$-0.610^{+0.548}_{-0.648}$
$z > 0.2$ no CL J1226.9+3332	33	0.301	$0.353^{+0.076}_{-0.059}$	$0.855^{+0.552}_{-0.650}$	$0.432^{+0.135}_{-0.096}$	$0.452^{+0.732}_{-0.886}$
$z > 0.5$ no CL J1226.9+3332	10	0.586	$0.060^{+0.154}_{-0.015}$	$4.593^{+0.584}_{-2.867}$	$0.068^{+0.350}_{-0.018}$	$4.398^{+0.592}_{-4.097}$

Notes. Columns 1–3 present the considered sample, the number of clusters, and the median redshift. Columns 4–7 give the best-fit values with 16th and 84th percentiles of the posterior distributions for parameters describing bias evolution, without (Cols. 4 and 5) and with (Cols. 6 and 7) the systematic scatters. In bold the values corresponding to the reference sample accounting for the systematic scatters.

SPT-CLJ0546−5345, and SPT-CLJ2341−5119 are above $z = 0.62$. For CL J1226.9+3332 the uncertainties on the bias are more than one order of magnitude smaller than the uncertainties of the three SPT clusters. We suspect that this single cluster may be forcing the bias towards lower values at high redshift. To test the impact that CL J1226.9+3332 has on the fits, we repeat the analyses excluding it. The results without CL J1226.9+3332 are shown, following the same colour scheme as before, in the bottom panel in Fig. 3 and with dashed lines in Fig. 4. We observe that β_z varies significantly when excluding CL J1226.9+3332 and it tends to be more compatible with no redshift evolution. At the same time, the bias at $z = 0$ is slightly shifted towards lower values. All the results are summarised in Table 2.

The described direct HSE-to-lensing mass bias estimation method neglects the intrinsic scatters of the HSE and lensing mass estimates. As explained in Sereno & Ettori (2015), this could influence the resulting bias that relates HSE and lensing masses. For this reason, in the next section we take a different approach to estimate the HSE-to-lensing mass bias.

5. HSE-to-lensing mass scaling relation

Estimating the scaling relation between HSE and lensing masses is an alternative way for measuring the HSE-to-lensing mass bias (Eq. (7)), together with the intrinsic scatter associated to HSE and lensing masses. We follow the methodology presented in Sereno & Ettori (2015) and consider that both the HSE and the lensing masses are scattered and biased estimates of the true mass of clusters, such that

$$\ln M^{\rm lens} \pm \delta_{\rm lens} = \alpha^{\rm lens} + \beta^{\rm lens} \ln M^{\rm True} \pm \sigma^{\rm lens}, \quad (9)$$

$$\ln M^{\rm HSE} \pm \delta_{\rm HSE} = \alpha^{\rm HSE} + \beta^{\rm HSE} \ln M^{\rm True} \pm \sigma^{\rm HSE}. \quad (10)$$

Here $\delta_{\rm lens}$ and $\delta_{\rm HSE}$ are the measurement uncertainties associated with the logarithm of the lensing and HSE mass estimates for each cluster. The natural logarithm of the bias and the deviation from linearity are α and β, respectively. The intrinsic scatter of the lensing and HSE masses with respect to the true mass are given by $\sigma^{\rm lens}$ and $\sigma^{\rm HSE}$. All the masses in the arguments of logarithms are in $10^{14}\,M_\odot$ units. Authors in Sereno & Ettori (2015) verified that the scatter and bias results do not vary if $\alpha^{\rm lens} = 0$ or $\alpha^{\rm HSE} = 0$ is considered, so following their work, we fix $\alpha^{\rm lens} = 0$.

We use the LInear Regression in Astronomy (LIRA[7], Sereno 2016) R package and the `pylira`[8] Python wrapper to perform the fit of the SR. LIRA performs the Gibbs sampling of a posterior distribution constructed from a MCMC fit based on a Bayesian hierarchical modelling. It can account for heteroscedastic measurement errors, intrinsic scatter, and time evolution of the SR.

5.1. Reference scaling relation

The SR of reference in this paper is built using the aforementioned 53 clusters in the reference sample, assuming that both the lensing and the HSE masses scale linearly with the true mass, $\beta^{\rm lens} = 1$ and $\beta^{\rm HSE} = 1$, and that there is no evolution of the SR with redshift. The MCMC sampling is performed using 200 chains and 6×10^6 steps, with a burn-in of the first half of the steps. Convergence is checked following the $\hat{R}$ test of Gelman & Rubin (1992). We take uniform priors for the free parameters: $\alpha^{\rm HSE} \sim \mathcal{U}(-4, 4)$, $\sigma^{\rm HSE} \sim \mathcal{U}(0, 10)$, $\sigma^{\rm lens} \sim \mathcal{U}(0, 10)$.

We present in the left panel in Fig. 5 the HSE-to-lensing mass scaling relation obtained with the 53 clusters of the reference sample. Data points correspond to each one of the clusters in the sample, with the ellipses in the figure indicating the error bars in both axes when considering the systematic scatter (see Eqs. (5) and (6)). We assume no correlation between both mass estimates. The grey and pink lines show respectively the scaling relation accounting and not accounting for the systematic scatter in the error bars of each cluster (Eqs. (5) and (6)). Shaded areas indicate the 1σ region. The black dashed line shows the one-to-one relation between HSE and lensing masses. In the right panel in Fig. 5, we show the posterior distributions of the fitted scaling relation parameters. The intrinsic scatter related to HSE masses is remarkably shifted towards zero when accounting for the systematic scatter in the error bars of cluster masses. This is

[7] https://cran.r-project.org/web/packages/lira/
[8] https://github.com/fkeruzore/pylira

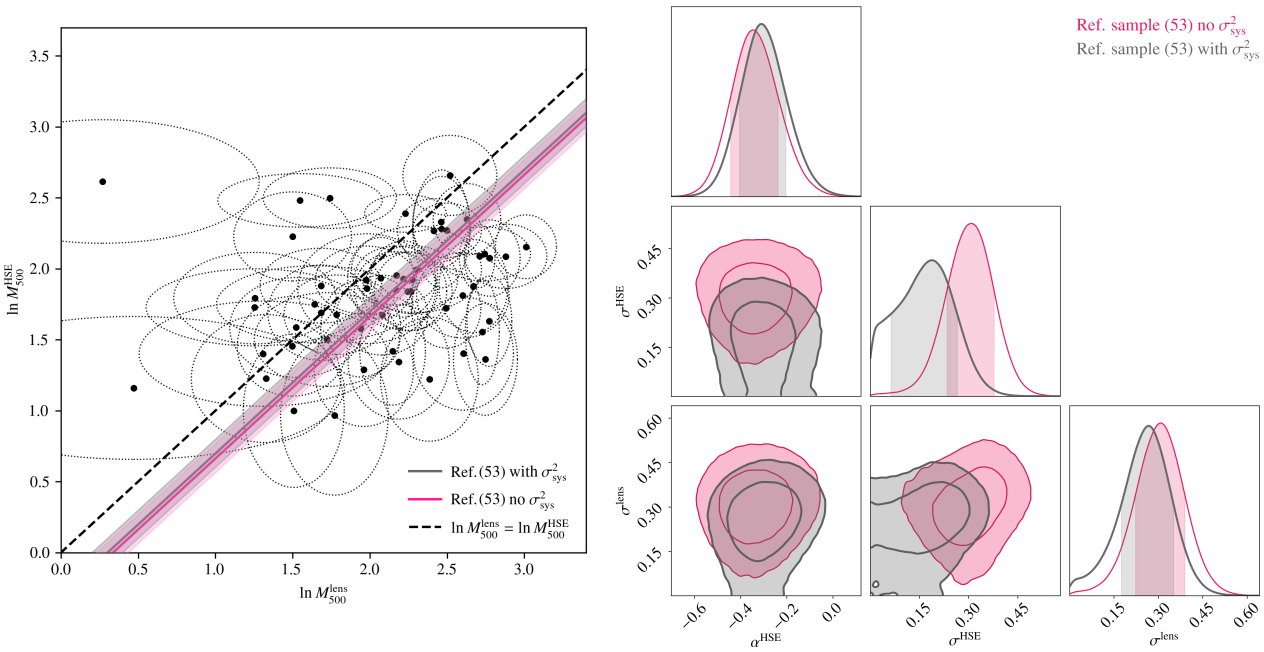

Fig. 5. Reference scaling relation ($\beta^{HSE} = 1$) between HSE and lensing masses in the reference sample. Data points with ellipses represent each cluster masses and uncertainties in both axes accounting for the systematic scatter. The pink line corresponds to the SR for the median value of parameters obtained without σ_{sys} and the solid grey line with σ_{sys}. The shaded regions show the 16th and 84th percentiles and the black dashed line gives the one-to-one relation. The corner plots in the right panel are the posterior 1D and 2D distributions of the parameters in the SR, including (grey) or not (pink) systematic scatters.

expected, since increasing the error bars of clusters reduces the need to have a dispersion around the SR. The median values with the 16th and 84th percentiles of the posterior distributions of α^{HSE}, σ^{HSE}, and σ^{lens} are given in the first row of Table 3. From α^{HSE} we compute the HSE-to-lensing mass bias at R_{500} (Eq. (7)), which gives $(1-b) = 0.739^{+0.075}_{-0.070}$ considering the systematic scatters.

Impact of particular subsamples in redshift. As for the bias model in Sect. 4, we also want to check how the SR parameters may vary depending on the chosen redshift range. Therefore, we repeat the analysis for the different redshift subsamples considered in Sect. 4. We present in Figs. 6 and C.1 and in Table 3 the different results, with and without σ^2_{sys}. Again, we observe that the bias changes for $z < 0.2$ and $z > 0.2$ clusters, in line with a $(1-b)$ value that decreases with redshift. The scaling relations with $z < 0.9$ and $z < 0.5$ samples remain almost unchanged with respect to the SR of reference. Not accounting for CL J1226.9+3332 reduces the lensing scatter for the $z > 0.5$ subsample. Overall, we find that the SRs are compatible for the different subsamples.

The posterior distribution of the SR parameters obtained for the $z < 0.5$ clusters without σ^2_{sys} (see Fig. C.2) can be directly compared to Fig. 5 in Sereno & Ettori (2015). In that work, the 50 CCCP clusters from Mahdavi et al. (2013) were used to measure the HSE-to-lensing mass scaling relation (even though the HSE masses were evaluated at the R_{500} obtained from lensing). The intrinsic scatters seem to be differently correlated in Sereno & Ettori (2015) and in this paper. However, in both cases we observe no strong correlation between α^{HSE} and the intrinsic HSE or lensing scatters. In our case, for the $z < 0.5$ clusters without (with) σ^2_{sys} we measure $(1-b) = 0.720^{+0.080}_{-0.073}$ ($(1-b) = 0.756^{+0.086}_{-0.079}$). These results (Table 3) are in line with the values reported in Table 6 in Sereno & Ettori (2015) and Table 2 in Lovisari et al. (2020a).

5.2. Investigations of possible model extensions

Beyond the reference scaling relation, for which we have assumed no redshift evolution and a linear scaling between the masses, in this section we test if relaxing some of these assumptions improves the description of the data by the scaling relation model.

Deviation from linearity. The HSE and/or lensing masses could also scale non-linearly with the true mass, meaning that the HSE-to-lensing bias would depend on the mass of the clusters. In Hoekstra et al. (2015) and von der Linden et al. (2014) authors investigated such dependence on the mass comparing *Planck* results to CCCP and WtG lensing masses, respectively. Both works found modest evidence for a mass-dependence: $M_{Planck} \propto M^{0.64\pm0.17}_{CCCP}$ with $\alpha^{HSE} \sim 0.55$ in Hoekstra et al. (2015), and $M_{Planck} \propto M^{0.68^{+0.15}_{-0.11}}_{WtG}$ with $\alpha^{HSE} \sim 0.38$ and $M_{Planck} \propto M^{0.76^{+0.39}_{-0.20}}_{WtG}$ with $\alpha^{HSE} \sim 0.19$ in von der Linden et al. (2014) for different cluster samples. Physically, this mass dependence could correspond, for example, to an impact of the baryonic physics that would depend on the strength of the clusters potential wells. In this case, low mass clusters having shallower potential wells, we can imagine that baryonic effects are stronger in them (McCarthy et al. 2011). On the contrary, simulations in Rasia et al. (2012) also indicate that massive objects are the most disturbed ones and have, probably, more complex temperature structures.

We also test this hypothesis by fitting the SR in Eqs. (9) and (10) leaving β^{HSE} as a free parameter. We take a uniform prior for $\beta^{HSE} \sim \mathcal{U}(0, 2)$ and consider the same priors for α^{HSE}, σ^{HSE}, and σ^{lens}. The resulting scaling relations are presented in Fig. 7 and the median values are given in Table 4. As shown in the corner plot in Fig. 7, α^{HSE} and β^{HSE} are completely degenerated. Nevertheless, our results are in agreement with Hoekstra et al. (2015) and

Table 3. Summary of the median values and uncertainties at 16th and 84th percentiles of the parameters for the HSE-to-lensing SR assuming linearity ($\beta^{\rm HSE} = 1$).

Cluster sample	# of clusters	No $\sigma^2_{\rm sys}$			
		$\alpha^{\rm HSE}$	$e^{\alpha^{\rm HSE}} = (1-b)$	$\sigma^{\rm HSE}$	$\sigma^{\rm lens}$
Reference sample	53	$-0.338^{+0.105}_{-0.097}$	$0.713^{+0.075}_{-0.069}$	$0.304^{+0.069}_{-0.072}$	$0.305^{+0.080}_{-0.083}$
$z < 0.9$	50	$-0.309^{+0.124}_{-0.110}$	$0.734^{+0.091}_{-0.081}$	$0.275^{+0.071}_{-0.071}$	$0.267^{+0.083}_{-0.086}$
$z < 0.5$	42	$-0.328^{+0.111}_{-0.102}$	$0.720^{+0.080}_{-0.073}$	$0.282^{+0.080}_{-0.086}$	$0.308^{+0.090}_{-0.091}$
$z < 0.2$	19	$-0.215^{+0.223}_{-0.166}$	$0.806^{+0.180}_{-0.133}$	$0.332^{+0.114}_{-0.128}$	$0.368^{+0.155}_{-0.152}$
$z > 0.2$	34	$-0.421^{+0.139}_{-0.129}$	$0.656^{+0.091}_{-0.085}$	$0.298^{+0.076}_{-0.090}$	$0.334^{+0.090}_{-0.086}$
$z > 0.5$	11	$-0.668^{+0.316}_{-0.320}$	$0.513^{+0.162}_{-0.164}$	$0.403^{+0.155}_{-0.116}$	$0.307^{+0.200}_{-0.157}$
Ref. no CL J1226.9+3332	52	$-0.350^{+0.098}_{-0.092}$	$0.705^{+0.070}_{-0.065}$	$0.294^{+0.072}_{-0.075}$	$0.295^{+0.084}_{-0.087}$
$z < 0.9$ no CL J1226.9+3332	49	$-0.338^{+0.114}_{-0.099}$	$0.713^{+0.081}_{-0.070}$	$0.273^{+0.073}_{-0.074}$	$0.274^{+0.083}_{-0.086}$
$z > 0.2$ no CL J1226.9+3332	33	$-0.430^{+0.130}_{-0.126}$	$0.651^{+0.085}_{-0.082}$	$0.289^{+0.082}_{-0.099}$	$0.320^{+0.093}_{-0.093}$
$z > 0.5$ no CL J1226.9+3332	10	$-0.629^{+0.332}_{-0.410}$	$0.533^{+0.177}_{-0.219}$	$0.446^{+0.184}_{-0.133}$	$0.189^{+0.209}_{-0.131}$
		With $\sigma^2_{\rm sys}$			
		$\alpha^{\rm HSE}$	$e^{\alpha^{\rm HSE}} = (1-b)$	$\sigma^{\rm HSE}$	$\sigma^{\rm lens}$
Reference sample	**53**	$\mathbf{-0.303^{+0.101}_{-0.095}}$	$\mathbf{0.739^{+0.075}_{-0.070}}$	$\mathbf{0.166^{+0.086}_{-0.101}}$	$\mathbf{0.257^{+0.080}_{-0.092}}$
$z < 0.9$	50	$-0.260^{+0.124}_{-0.110}$	$0.771^{+0.095}_{-0.084}$	$0.122^{+0.090}_{-0.082}$	$0.220^{+0.084}_{-0.099}$
$z < 0.5$	42	$-0.280^{+0.114}_{-0.104}$	$0.756^{+0.086}_{-0.079}$	$0.136^{+0.098}_{-0.090}$	$0.262^{+0.089}_{-0.100}$
$z < 0.2$	19	$-0.191^{+0.181}_{-0.157}$	$0.827^{+0.150}_{-0.129}$	$0.179^{+0.123}_{-0.111}$	$0.318^{+0.135}_{-0.130}$
$z > 0.2$	34	$-0.388^{+0.136}_{-0.127}$	$0.679^{+0.092}_{-0.086}$	$0.162^{+0.099}_{-0.103}$	$0.271^{+0.095}_{-0.103}$
$z > 0.5$	11	$-0.621^{+0.305}_{-0.308}$	$0.538^{+0.164}_{-0.166}$	$0.293^{+0.175}_{-0.150}$	$0.252^{+0.206}_{-0.156}$
Ref. no CL J1226.9+3332	52	$-0.312^{+0.096}_{-0.091}$	$0.732^{+0.070}_{-0.067}$	$0.154^{+0.089}_{-0.097}$	$0.249^{+0.084}_{-0.096}$
$z < 0.9$ no CL J1226.9+3332	49	$-0.287^{+0.116}_{-0.102}$	$0.750^{+0.087}_{-0.077}$	$0.120^{+0.089}_{-0.080}$	$0.226^{+0.085}_{-0.099}$
$z > 0.2$ no CL J1226.9+3332	33	$-0.397^{+0.131}_{-0.131}$	$0.672^{+0.088}_{-0.088}$	$0.151^{+0.103}_{-0.097}$	$0.264^{+0.100}_{-0.112}$
$z > 0.5$ no CL J1226.9+3332	10	$-0.587^{+0.332}_{-0.427}$	$0.556^{+0.185}_{-0.238}$	$0.339^{+0.207}_{-0.170}$	$0.190^{+0.216}_{-0.132}$

Notes. We present the results for different data subsamples, with and without accounting for the systematic uncertainties in the error bars of the masses. We show in bold the parameters for the scaling relation of reference presented in Sect. 5.

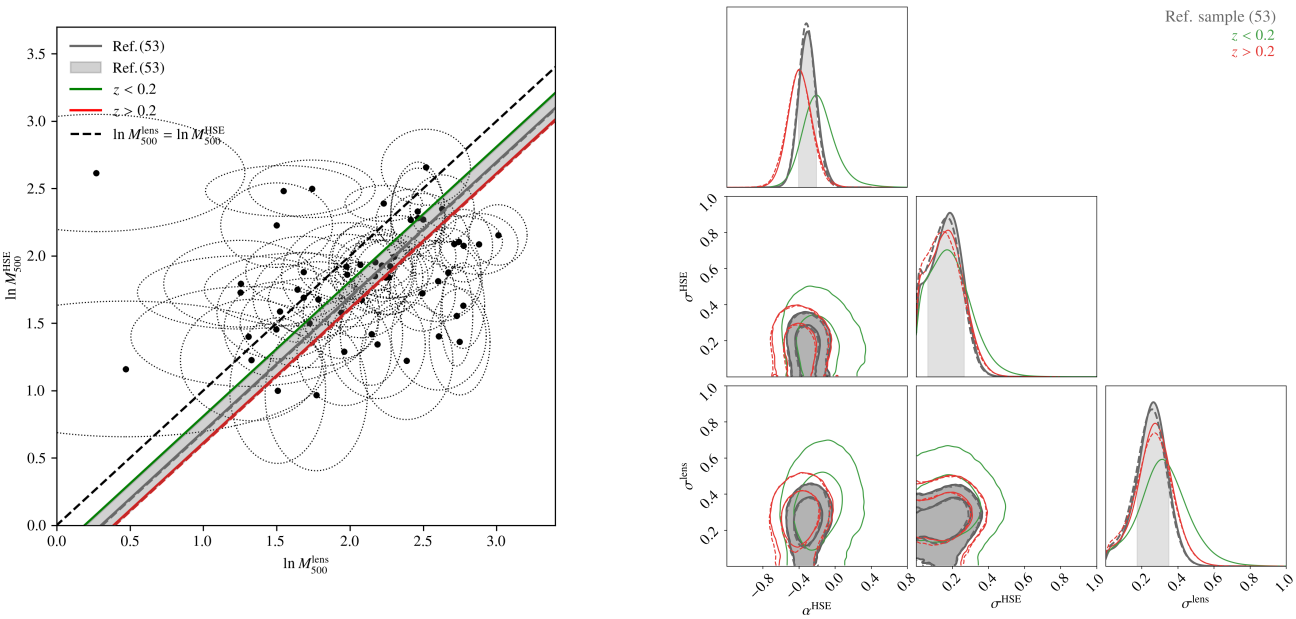

Fig. 6. Scaling relation between HSE and lensing masses for the reference sample in grey and for different subsamples in colours, all accounting for $\sigma_{\rm sys}$. Here $\beta^{\rm HSE}$ is fixed to 1. As in Fig. 4, we only show the cases for $z > 0.2$ and $z < 0.2$. Data points with ellipses represent each cluster masses and uncertainties in both axes accounting for the systematic scatters. The black dashed line shows the equality. The corner plots in the right panel are the posterior 1D and 2D distributions of the parameters in the SR.

Table 4. Summary of the median values and uncertainties at the 16th and 84th percentiles of the parameters in the HSE-to-lensing SR when considering a deviation from linearity, an offset between HSE and lensing masses or an evolution with redshift.

Cluster sample	# of clusters	No $\sigma^2_{\rm sys}$				
		$\alpha^{\rm HSE}$	$\beta^{\rm HSE}$	$\sigma^{\rm HSE}$	$\sigma^{\rm lens}$	$\gamma^{\rm HSE}$
Reference sample	53	$0.705^{+0.666}_{-1.249}$	$0.519^{+0.576}_{-0.309}$	$0.335^{+0.057}_{-0.066}$	$0.275^{+0.103}_{-0.163}$	[0]
Reference sample (BCES)	53	0.826 ± 0.886	0.481 ± 0.415	$0.326^{(*)}$	–	[0]
Reference sample	53	$-0.219^{+0.137}_{-0.134}$	[1]	$0.303^{+0.069}_{-0.072}$	$0.295^{+0.082}_{-0.085}$	$-1.742^{+1.082}_{-1.083}$
Ref. no CL J1226.9+3332	52	$-0.263^{+0.140}_{-0.136}$	[1]	$0.297^{+0.072}_{-0.073}$	$0.292^{+0.084}_{-0.087}$	$-1.168^{+1.156}_{-1.163}$
		$A^{\rm HSE}$ [$10^{14}\,M_\odot$]	$B^{\rm HSE}$	$\sigma^{\rm HSE}$ [$10^{14}\,M_\odot$]	$\sigma^{\rm lens}$ [$10^{14}\,M_\odot$]	$\gamma^{\rm HSE}$ [$10^{14}\,M_\odot$]
Reference sample	53	$0.818^{+2.313}_{-3.037}$	$0.614^{+0.351}_{-0.270}$	$1.673^{+0.413}_{-0.640}$	$3.159^{+0.701}_{-0.983}$	[0]
Reference sample (BCES)	53	4.467 ± 1.85	0.246 ± 0.19	$2.109^{(*)}$	–	[0]
Reference sample	53	$0.545^{+2.465}_{-2.925}$	$0.675^{+0.334}_{-0.291}$	$1.644^{+0.439}_{-0.700}$	$3.188^{+0.674}_{-0.854}$	$-2.685^{+4.654}_{-4.260}$
		With $\sigma^2_{\rm sys}$				
		$\alpha^{\rm HSE}$	$\beta^{\rm HSE}$	$\sigma^{\rm HSE}$	$\sigma^{\rm lens}$	$\gamma^{\rm HSE}$
Reference sample	53	$0.824^{+0.719}_{-1.087}$	$0.498^{+0.481}_{-0.326}$	$0.204^{+0.068}_{-0.082}$	$0.242^{+0.091}_{-0.135}$	[0]
Reference sample (BCES)	53	1.000 ± 0.692	0.397 ± 0.324	$0.191^{(*)}$	–	[0]
Reference sample	53	$-0.193^{+0.135}_{-0.134}$	[1]	$0.168^{+0.086}_{-0.100}$	$0.246^{+0.082}_{-0.095}$	$-1.530^{+1.071}_{-1.085}$
Ref. no CL J1226.9+3332	52	$-0.242^{+0.139}_{-0.137}$	[1]	$0.153^{+0.090}_{-0.096}$	$0.248^{+0.084}_{-0.093}$	$-0.896^{+1.154}_{-1.155}$
		$A^{\rm HSE}$ [$10^{14}\,M_\odot$]	$B^{\rm HSE}$	$\sigma^{\rm HSE}$ [$10^{14}\,M_\odot$]	$\sigma^{\rm lens}$ [$10^{14}\,M_\odot$]	$\gamma^{\rm HSE}$ [$10^{14}\,M_\odot$]
Reference sample	53	$1.603^{+1.665}_{-2.994}$	$0.522^{+0.349}_{-0.194}$	$0.949^{+0.557}_{-0.602}$	$2.867^{+0.806}_{-1.251}$	[0]
Reference sample (BCES)	53	4.644 ± 1.727	0.226 ± 0.176	$1.340^{(*)}$	–	[0]
Reference sample	53	$1.434^{+1.780}_{-3.005}$	$0.570^{+0.346}_{-0.210}$	$0.950^{+0.570}_{-0.612}$	$2.927^{+0.771}_{-1.131}$	$-2.493^{+4.535}_{-4.282}$

Notes. We present the results for the reference sample, accounting or not for the systematic scatter in the error bars of the masses. For the BCES fit we report the best-fit values and 1σ uncertainties. $^{(*)}$We also calculate the scatter with respect to the best BCES scaling relations following Eq. (4).

von der Linden et al. (2014). However, the HSE masses in those works were *Planck* masses from the SZ-mass scaling relation.

For comparison to the results obtained with LIRA, we also perform the fit of the SR using the orthogonal Bivariate Correlated Errors and intrinsic Scatter method (BCES, Akritas & Bershady 1996). BCES favours a larger deviation from linearity, that is, smaller $\beta^{\rm HSE}$. We also report the results in Table 4. Given the large uncertainties on $\alpha^{\rm HSE}$ and $\beta^{\rm HSE}$, the scaling relations obtained with LIRA and BCES are compatible.

In Fig. 8 we present the HSE-to-lensing mass ratio as a function of the lensing mass for the fitted $\alpha^{\rm HSE}$ and $\beta^{\rm HSE}$, with the green shaded area showing the 16th to 84th percentiles. The horizontal grey hatched area represents the HSE-to-lensing mass ratio measured in the previous section assuming that HSE and lensing masses scale linearly with the true mass. Given that we obtain $\beta^{\rm HSE} < 1$, on average the difference between HSE and lensing masses is larger for more massive objects. This is in agreement with the mild decreasing tendency for the HSE-to-lensing mass ratio obtained in Hoekstra et al. (2015), von der Linden et al. (2014), and Eckert et al. (2019), but different from the trend observed in Salvati et al. (2019). Nevertheless, our results are consistent with no mass dependence of the ratio. The difficulty of disentangling $\alpha^{\rm HSE}$ and $\beta^{\rm HSE}$ does not motivate further investigations of the SR model with additional free parameters. Leaving free $\alpha^{\rm lens}$ would add a free parameter to the model strongly correlated to $\alpha^{\rm HSE}$ and $\beta^{\rm HSE}$.

Considering an offset. In addition to the HSE-to-lensing mass bias defined in Eq. (7), there could be also an offset between the HSE and lensing mass estimates. Thus, the scaling relation could be defined as,

$$M^{\rm lens} \pm \delta_{\rm lens} = M^{\rm True} \pm \sigma^{\rm lens}, \tag{11}$$

$$M^{\rm HSE} \pm \delta_{\rm HSE} = A^{\rm HSE} + B^{\rm HSE} M^{\rm True} \pm \sigma^{\rm HSE}, \tag{12}$$

where $A^{\rm HSE}$ and $B^{\rm HSE}$ are the offset and the multiplicative factor, respectively. Here $\sigma^{\rm HSE}$ and $\sigma^{\rm lens}$ are again the scatter of HSE and lensing masses with respect to the SR, but in this case in units of $10^{14}\,M_\odot$.

We perform again the fit of the SR using both the LIRA and BCES methods. We present in Fig. 9 and Table 4 the results. As for the non-linear SR fit, $A^{\rm HSE}$ and $B^{\rm HSE}$ are completely degenerated. The results obtained with LIRA indicate an offset in mass completely compatible with zero. It is reassuring to verify that the data motivates a scaling relation model for which the HSE mass goes to zero in the limit $M^{\rm True} \to 0$. We show in Fig. 8 the bias evolution in blue, indicating again that there is no significant trend of the HSE-to-lensing mass ratio with cluster mass.

Evolution with redshift. LIRA enables fitting a scaling relation that evolves with redshift. Looking for such evolution can be particularly interesting with our reference sample, given the large redshift range that it covers ($0.05 < z < 1.07$). Assuming again that HSE and lensing masses scale linearly with the true mass ($\beta^{\rm lens} = \beta^{\rm HSE} = 1$), we write

$$\ln M^{\rm lens} \pm \delta_{\rm lens} = \ln M^{\rm True} \pm \sigma^{\rm lens}, \tag{13}$$

and,

$$\ln M^{\rm HSE} \pm \delta_{\rm HSE} = \alpha^{\rm HSE} + \ln M^{\rm True} \pm \sigma^{\rm HSE} + \gamma^{\rm HSE} T. \tag{14}$$

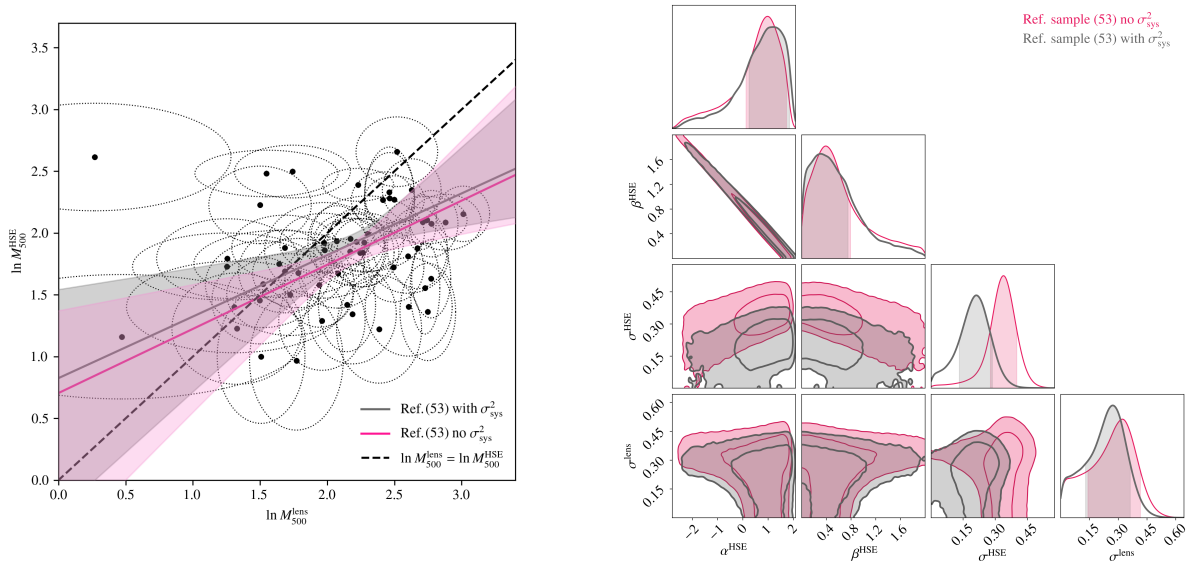

Fig. 7. Scaling relation between HSE and lensing masses in the reference sample considering a deviation from linearity. Data points with ellipses represent each cluster masses and uncertainties in both axes accounting for the systematic scatters. The pink line corresponds to the SR for the median value of parameters obtained without $\sigma_{\rm sys}$ and the solid grey line with $\sigma_{\rm sys}$. The black dashed line shows the equality and shaded regions the 16th and 84th percentiles. The corner plots in the right panel are the posterior 1D and 2D distributions of the parameters in the SR, including (grey) or not (pink) the systematic scatters.

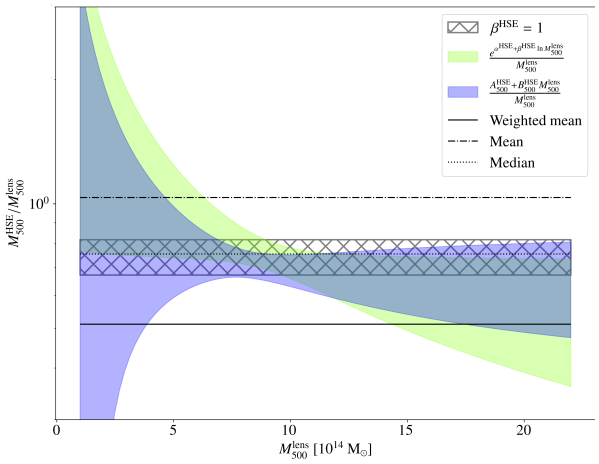

Fig. 8. HSE-to-lensing mass ratio with respect to lensing mass. The grey hatched area indicates the 16th to 84th percentiles of the bias without mass dependence, accounting for systematic scatters in the uncertainties of HSE and lensing masses. The green area shows the bias evolution when assuming a deviation from linearity of the HSE and lensing masses. Blue area indicates the bias evolution when considering an offset between HSE and lensing masses. Horizontal solid, dotted, and dash-dotted black lines give respectively the weighted mean, median, and mean mass ratio for the 53 clusters, same as in Fig. 3.

We note that T is the time evolution factor, $T = \log\left(\frac{1+z}{1+z_{\rm ref}}\right)$, with $z_{\rm ref} = 0.01$ the normalisation redshift set by default in LIRA. We take flat priors for the parameter describing the evolution with redshift: $\gamma^{\rm HSE} \sim \mathcal{U}(-10, 10)$. Similarly, we consider the evolution with redshift for the SR defined in Eqs. (11) and (12). Given the strong impact of the CL J1226.9+3332 galaxy cluster on the fits at high redshift (see Sect. 4), we repeat the analysis excluding it. All the results are summarised in Table 4.

In Fig. 10 we present the redshift evolution of the HSE-to-lensing mass ratio for the analyses performed with the reference sample and accounting for systematic uncertainties in the HSE and lensing masses. We show in grey the results obtained in Sect. 4, neglecting the intrinsic scatter of HSE and lensing masses with respect to the true masses. In blue we present the bias evolution model resulting from the scaling relation fit in this section. Darker regions show the evolution with redshift obtained when excluding CL J1226.9+3332 from the analyses.

There seem to be a tendency for a decreasing HSE-to-lensing mass ratio with redshift ($\gamma^{\rm HSE} = -1.530^{+1.071}_{-1.085}$), but it is not statistically significant when removing CL J1226.9+3332 from the sample ($\gamma^{\rm HSE} = -0.896^{+1.154}_{-1.155}$). From the comparison of the grey and blue results we observe directly the impact that accounting for the intrinsic scatters of the SRs has on the bias. Considering the intrinsic scatter reduces the difference between HSE and lensing masses and, therefore, the bias.

We present in Figs. C.3 and C.4 a comparison of the scaling relations and posterior distributions of parameters when accounting for redshift evolution (dashed lines) and not accounting for it (solid lines). The contribution of the redshift evolution factor introduces a change of the order of a few percent (or less) in the intrinsic scatters. Given the correlation of the other parameters with $\gamma^{\rm HSE}$, the change is of ∼30% for $\alpha^{\rm HSE}$ and of the order of 10% for $A^{\rm HSE}$ and $B^{\rm HSE}$. However, the results are compatible with the ones obtained without considering redshift evolution, so there is no strong evidence of redshift evolution in the data.

5.3. Comparison of SR models

In this section, we compare the tested SR models to assess which is the one preferred by the data. We define the goodness of fit of the scaling relations $\hat{\chi}^2$ following Eq. (3) in Lovisari et al. (2020b):

$$\hat{\chi}^2 = \sum_{i=1}^{N_{\rm clusters}} \frac{\left[\ln M_i^{\rm HSE} - \ln M^{\rm HSE}\left(\ln M_i^{\rm lens}, z_i, \vartheta\right)\right]^2}{\delta^2_{{\rm HSE},i} + (\sigma^{\rm HSE})^2 + (\beta^{\rm HSE})^2\left[\delta^2_{{\rm lens},i} + (\sigma^{\rm lens})^2\right]}, \quad (15)$$

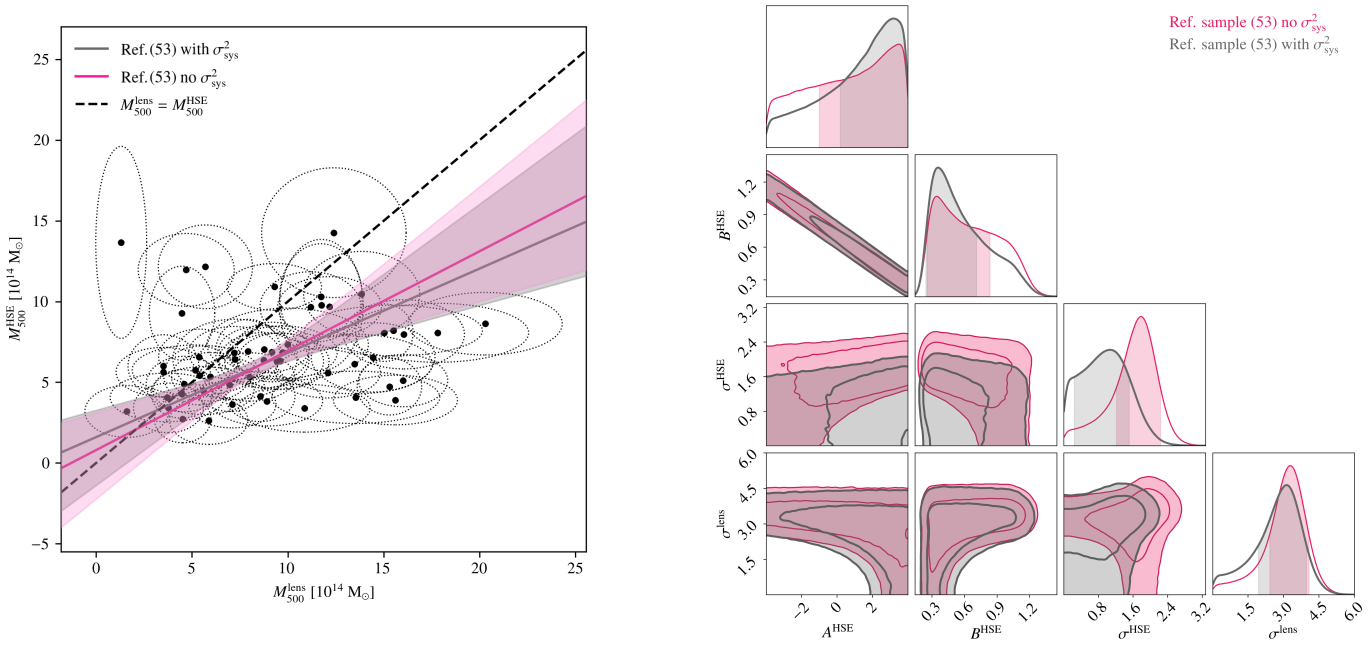

Fig. 9. Scaling relation between HSE and lensing masses in the reference sample considering an offset between both mass estimates. Data points with ellipses represent each cluster masses and the uncertainties in both axes accounting for the systematic scatter. The pink line corresponds to the SR obtained without σ_{sys} and the solid grey line with σ_{sys}. The black dashed line shows the equality. The corner plots in the right panel are the posterior 1D and 2D distributions of the parameters in the SR, including (grey) or not (pink) the systematic scatters.

where the sum is done over the $N_{clusters} = 53$ clusters in the reference sample. In Eq. (15) $\ln M^{HSE}\left(\ln M_i^{lens}, z_i, \vartheta\right)$ is the function described by Eq. (10) or (14) depending on the SR model, with the parameters $\vartheta = [\alpha^{HSE}, \beta^{HSE}, \gamma^{HSE}]$ defined accordingly. The factors $\ln M_i^{HSE}$, $\ln M_i^{lens}$, $\delta_{HSE,i}$, and $\delta_{lens,i}$ are the HSE and lensing mass of each cluster i and their associated uncertainties, and z_i is the redshift of each cluster. We compare the results obtained considering always the systematic uncertainties in the HSE and lensing mass uncertainties. We take the posterior distributions of the parameters for α^{HSE}, β^{HSE}, γ^{HSE}, σ^{HSE}, and σ^{lens}. For the scaling relations considering an offset in mass (Eqs. (11) and (12)), we replace the logarithmic masses and uncertainties by the linear values in the $\hat{\chi}^2$ definition in Eq. (15). Similarly, we take A^{HSE} and B^{HSE} instead of α^{HSE} and β^{HSE}. The $\hat{\chi}^2$ distribution for each SR model fit is shown in Fig. C.5.

According to the Akaike information criterion (AIC, Akaike 1974) and the Bayesian information criterion (BIC, Schwarz 1978), the scaling relation of reference and the one considering a deviation from linearity are almost equally probable (see Appendix C.2 for more details). Furthermore, there is statistically no gain in adding a parameter that describes an evolution with redshift. In other words, redshift evolution does not seem to be favoured by the data.

Anyhow, the intrinsic scatters being free parameters in our LIRA fits, we expect all the models to adjust the data points at the expense of increasing the scatters. From the comparison of all the σ^{HSE} and σ^{lens} (see Tables 3 and 4), there is not a statistically significant increase, nor decrease in the intrinsic scatters when changing the number of free parameters in the SR model.

In conclusion, our best scaling law between X-ray HSE and lensing masses is given by the scaling relation of reference:

$$\ln M^{lens} = \ln M^{True} \pm 0.257^{+0.080}_{-0.092}, \quad (16)$$

$$\ln M^{HSE} = -0.303^{+0.101}_{-0.095} + \ln M^{True} \pm 0.166^{+0.086}_{-0.101}, \quad (17)$$

which corresponds to a HSE-to-lensing mass bias of

$$M^{HSE}_{500}/M^{lens}_{500} = (1-b) = 0.739^{+0.075}_{-0.070}\,(\text{stat.}) \pm 0.226\,(\text{intrin. scatter}), \quad (18)$$

assuming Gaussian intrinsic scatters for lensing and HSE masses.

5.4. Caveats

The two main caveats of the analysis presented in this work are the representativity of the used sample and the inhomogeneity in the estimates of the lensing masses. The former is hardly quantifiable, given that the selection criteria of the reference sample (Sect. 2.1) are mainly a combination of the selection criteria used for the ESZ, LoCuSS, LPSZ, and Bartalucci et al. (2018) clusters. An equivalent study using a clearly defined selection criterium, as for the LPSZ (Mayet et al. 2020), would be of great interest.

Regarding the inhomogeneity of the lensing masses, we have exploited the compilation of mass estimates from different works standardised in the CoMaLit catalogue. We have treated all the CoMaLit masses equally, no matter the work from which the lensing mass has been extracted. But the different quality of the data and/or the methods used in each of the original works make the uncertainties of lensing masses not homogeneous within the CoMaLit sample. By propagating $\sigma_{sys\ lens}$ we account, to first order, for the overall error of CoMaLit masses with respect to other estimates. A possible improvement would be to measure an independent systematic scatter $\sigma_{sys\ lens}$ for each of the works used in the CoMaLit sample, but at the expenses of much poorer statistics. Instead, we quantify a posteriori the goodness of our best scaling relation (estimated with all the 53 clusters in the reference sample, Eqs. (16) and (17)) for the cluster masses extracted from each of the different works within

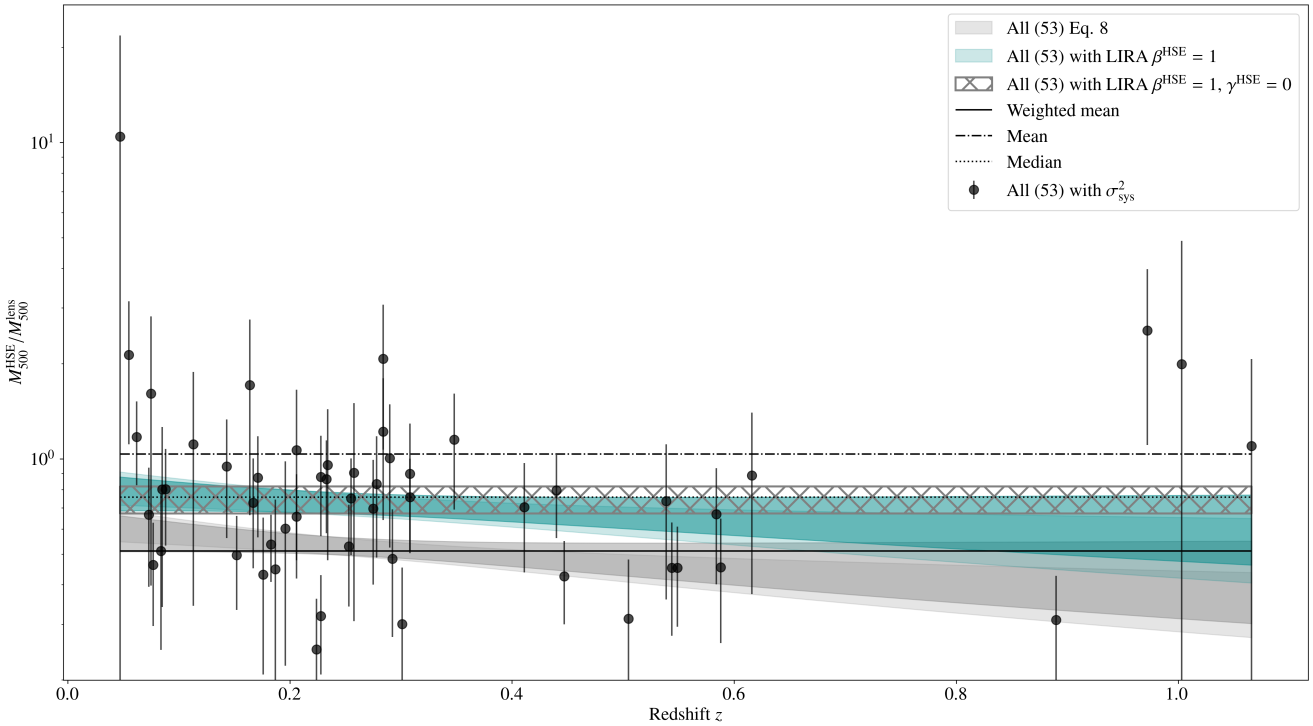

Fig. 10. HSE-to-lensing mass ratio with respect to redshift. The grey shaded area shows the evolution from Fig. 3 for all the clusters in the sample and in darker excluding CL J1226.9+3332. The blue area gives the evolution with redshift obtained from the fit of the scaling relation with the reference sample and the grey hatched area without considering the redshift evolution. The blue dark area is the evolution obtained for the reference sample excluding CL J1226.9+3332. As in Fig. 3, markers with error bars show the ratio per cluster in the reference sample with error bars accounting for the systematic uncertainty. Horizontal solid, dotted, and dash-dotted black lines give respectively the weighted mean, median, and mean mass ratio for the data points.

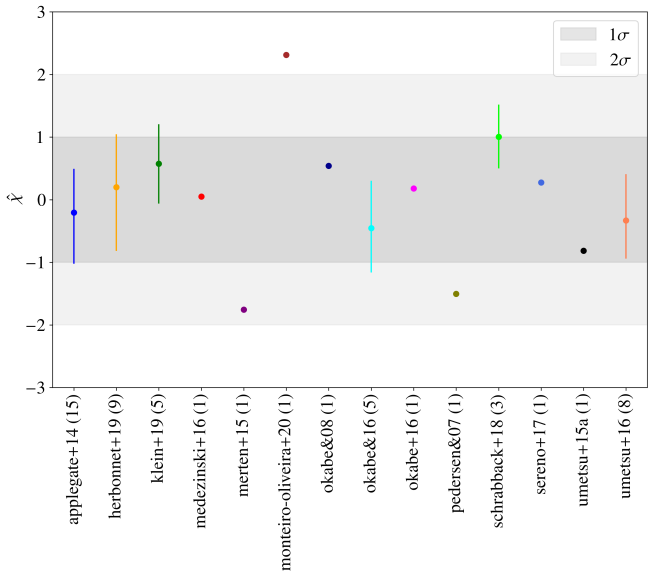

Fig. 11. Values of $\hat{\chi}$ for the clusters in the reference sample with respect to the reference scaling relation. We distinguish the results from the various lensing works used to build the CoMaLit sample with different colour markers. The number of clusters used in the reference sample from each of the works is indicated in brackets. For works with multiple clusters, we give the mean $\hat{\chi}$ and the 16th to 84th percentiles over all those clusters. Grey shaded areas indicate 1σ and 2σ regions.

the CoMaLit catalogue. In Fig. 11 we show, for the clusters obtained from each of the lensing works, the corresponding $\hat{\chi}$ defined from Eq. (15) as

$$\hat{\chi} = \frac{\ln M_i^{\rm HSE} - \ln M^{\rm HSE}\left(\ln M_i^{\rm lens}, z_i, \vartheta\right)}{\sqrt{\delta^2_{{\rm HSE},i} + \left(\sigma^{\rm HSE}\right)^2 + \left(\beta^{\rm HSE}\right)^2\left[\delta^2_{{\rm lens},i} + \left(\sigma^{\rm lens}\right)^2\right]}}. \tag{19}$$

For those works with several clusters in our reference sample, we give the mean value and the 16th to 84th percentiles over all the used clusters. We observe that only 'merten+15' (Merten et al. 2015), 'monteiro-oliveira+20' (Monteiro-Oliveira et al. 2020), and 'pedersen&07' (Pedersen & Dahle 2007) cluster masses are at more than 1σ. The cluster from 'monteiro-oliveira+20' at more than 2σ from the scaling relation is Abell1644 (on the top left of all our SRs), which is known for being a cluster in a merger scenario. Thus, we conclude that the scaling relation of reference fits well the large majority of the clusters in the reference sample. There is no hint of a too bad or too good fit to any of the subsamples in the CoMaLit catalogue.

6. Comparison to previous results

Similar studies to the one presented in this paper were previously done in the literature. However, the methods used to estimate the masses and to compute the HSE-to-lensing bias differ significantly from work to work. Thus, comparisons are delicate. In Fig. 12 we present our best bias estimate together with the HSE-to-lensing M_{500} ratios obtained in the works detailed below. We use Roman numerals to refer to each result from the literature. The different results are also summarised in Table 5.

The HSE-to-lensing mass bias was measured in Smith et al. (2015) with the 50 clusters from the LoCuSS sample

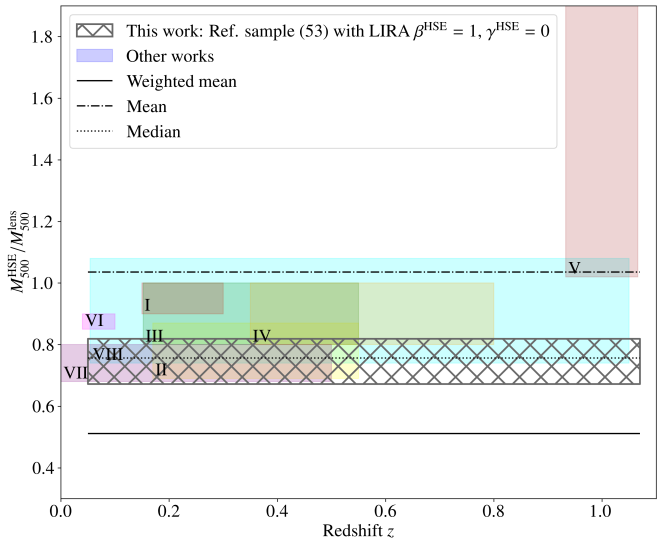

Fig. 12. HSE-to-lensing mass ratio with respect to redshift. Shaded areas indicate different results from different works in the literature. See text and Table 5 to identify Roman numerals with the works. The horizontal grey hatched area represents the HSE-to-lensing mass ratio measured in this work assuming that HSE and lensing masses scale linearly with the true mass, accounting for the systematic scatter, and considering no evolution with redshift.

($0.15 < z < 0.3$). By using resolved HSE mass estimates, they computed the weighted mean HSE-to-lensing bias: $1 - b = 0.95 \pm 0.05$ (I in Fig. 12). Uncertainties were calculated from the standard deviation of 1000 bootstrap samples geometric means. Following the equations (Eqs. (1) and (2) in Smith et al. 2015) used to calculate the weighted mean in Smith et al. (2015) we obtain for our reference sample a mean bias of: $1 - b = 0.763$ and 0.818 not including and including, respectively, the systematic error in the uncertainty of each mass estimate. Considering, as in Smith et al. (2015), only the clusters in the redshift range $0.15 < z < 0.3$, we obtain $1 - b = 0.769$ and 0.720 with and without the systematic scatter. The difference between the bias estimated in Smith et al. (2015) and the results obtained in our work could originate from the larger HSE mass estimates in Smith et al. (2015). Bright blue markers in the left panel in Fig. 1 show that HSE masses used in Smith et al. (2015) tend to be larger than the homogeneous ones.

In Mahdavi et al. (2008) authors compared the HSE and lensing masses evaluated at the same radius, in particular at the R_{500} measured from the lensing mass profile of each cluster. With a sample of 18 clusters, Mahdavi et al. (2008) concluded that at R_{500}^{lens} the ratio of masses is $M^{HSE}/M^{lens} = 0.78 \pm 0.09$ (II). Extending the analysis, the HSE-to-lensing mass bias obtained in Mahdavi et al. (2013) is consistent with no bias for cool-core clusters, while $(1 - b) \sim 0.8$ for non-cool core clusters (III). In the same line, authors in Israel et al. (2014) concluded, from the study of 8 clusters with redshifts $0.35 < z < 0.80$, that HSE and lensing masses differ by 0–20% (IV).

By using very high redshift clusters ($0.933 < z < 1.066$), Bartalucci et al. (2018) obtained that HSE masses from X-rays are a factor of $1.39^{+0.51}_{-0.37}$ (V) larger than weak lensing estimates, in contradiction with the rest of the results. The clusters in Bartalucci et al. (2018) are the highest redshift clusters in our reference sample (Sect. 2.1). Using the same HSE masses as in Bartalucci et al. (2018), but with the CoMaLit lensing estimates we obtain an error-weighted mean ratio of $M_{500}^{HSE}/M_{500}^{lens} = 1.56(1.58)$ not including (including) the systematic error in the uncertainty of each mass estimate. Instead, the error-weighted mean ratio for our full reference sample is $M_{500}^{HSE}/M_{500}^{lens} = 0.47(0.51)$. For the clusters in the X-COP sample, Eckert et al. (2022) found that HSE masses estimated using *XMM-Newton* data are 10–15% lower than the lensing estimates in Herbonnet et al. (2020) (VI). With a different approach and assuming that the gas fraction in clusters is constant, Eckert et al. (2019) obtained that HSE masses are biased (with respect to the true total mass) by 7% at R_{500}. These results differ significantly from the bias values obtained in this paper. HSE masses in Eckert et al. (2019) were reconstructed making use of excellently well resolved mass profiles, therefore, unless there are unidentified systematic effects, HSE masses in Eckert et al. (2019) should be very reliable. The small bias values obtained for the low redshift ($0.047 < z < 0.09$) clusters in Eckert et al. (2019) could then indicate that there is indeed a redshift dependence in the HSE bias and that low redshift cluster HSE masses are less biased.

Regarding also the evolution of the bias with redshift, which we have largely discussed in Sects. 4 and 5, the weak tendency for a larger bias at higher redshift seems to be in line with the results from Wicker et al. (2023) and Smith et al. (2015).

Particularly interesting are the comparisons to Sereno et al. (2019), Lovisari et al. (2020a), and Sereno & Ettori (2015) works, where the methods are equivalent to the ones employed in this paper, making use of the LIRA code and accounting for the intrinsic dispersion of HSE and lensing masses to the SR. The analysis in Lovisari et al. (2020a) compares the HSE masses obtained with *XMM-Newton* data (from Lovisari et al. 2020b) to lensing estimates in the CoMaLit LC^2 catalogue, for 62 clusters from the *Planck*-ESZ sample with $z < 0.5$. With this sample, authors obtain $1-b = 0.74 \pm 0.06$ (VII) and no redshift evolution. This is in excellent agreement with our result. In Lovisari et al. (2020a) the results found with CoMaLit lensing masses are also compared to those obtained with other lensing masses from other works in the literature: the HSE-to-lensing mass ratio spans from ~0.6 to ~1 depending on the used dataset.

Conclusions are along the same line in Sereno & Ettori (2015), where different samples with HSE and lensing mass estimates are used to measure the scaling relation and, consequently, the HSE-to-lensing bias. The effect that intrinsic scatters have on the determination of scaling relations is also studied in Sereno & Ettori (2015). They conclude that not taking into account explicitly the scatter of masses makes scaling relations flatter, as we see when using BCES instead of LIRA (also in agreement with Lovisari et al. 2020b). While the intrinsic scatter for lensing masses obtained in Sereno & Ettori (2015) is of the order of the expected values from simulations (~10−15%), the intrinsic dispersion for HSE masses is larger than expected (~20−30%). An underestimation of the statistical uncertainties in HSE masses could be the reason, according to Sereno & Ettori (2015), for this large scatter. Accounting for the systematic scatter in the uncertainty of each cluster mass, as described in this paper, could help to have more realistic uncertainties for the HSE mass estimates. The HSE-to-lensing mass ratio in Sereno & Ettori (2015) depends again on the used sample and data and spans from ~0.5 to ~1. We have reproduced the same result by separating the sample in redshift ranges.

Also Sereno et al. (2019) used the Bayesian hierarchical modelling from Sereno (2016) to fit a scaling relation between HSE masses from *XMM-Newton* data and weak lensing masses of clusters in the Hyper Suprime-Cam Survey (Pacaud et al. 2016). The median redshift of the 100 clusters in the sample is

Table 5. HSE-to-lensing mass bias values from resolved mass profiles.

Reference	Sample	# of clusters	Redshift	$(1-b) = M_{500}^{\rm HSE}/M_{500}^{\rm lens}$	Notes
	This work – Reference SR	53	$0.05 < z < 1.07$	$0.739^{+0.075}_{-0.070}$	Propagating the systematic uncertainties, accounting for intrinsic dispersion
I	Smith et al. (2015)	50	$0.15 < z < 0.3$	0.95 ± 0.05	Weighted mean
II	Mahdavi et al. (2008)	18	$0.170 < z < 0.547$	0.78 ± 0.09	$M^{\rm HSE}(R_{500}^{\rm lens})$
III	Mahdavi et al. (2013)	50	$0.152 < z < 0.55$	0.8–1	$M^{\rm HSE}(R_{500}^{\rm lens})$
IV	Israel et al. (2014)	8	$0.35 < z < 0.80$	$0.8 - 1$	Global temperature profile for the whole sample, $M^{\rm HSE}(R_{500}^{\rm lens})$
V	Bartalucci et al. (2018)	4	$0.933 < z < 1.066$	$1.39^{+0.51}_{-0.37}$	Weighted mean
VI	Eckert et al. (2022)	12	$0.047 < z < 0.09$	0.85–0.9	
VII	Lovisari et al. (2020a)	62	$z < 0.5$	0.74 ± 0.06	Accounting for intrinsic dispersion
VIII	Sereno et al. (2019)	100	$0.054 < z < 1.050$	0.91 ± 0.17	Accounting for intrinsic dispersion, temperature measured within 300 kpc

Notes. We report our reference result and different values from the literature. The last column indicates the singularity of each analysis.

$z = 0.30$, spanning from $z = 0.054$ to $z = 1.050$. Thus, the analysis in Sereno et al. (2019) is probably the closest study to our work. Nevertheless, to get temperature profiles that reach R_{500} with X-ray data (to compute then the HSE mass), in Sereno et al. (2019) a model was iteratively fitted to the integrated temperature measured per cluster within 300 kpc, well below R_{500}. Assuming $\beta^{\rm HSE} = 1$, $\beta^{\rm lens} = 1$, and $\alpha^{\rm lens} = 0$ they obtained: $\alpha^{\rm HSE} = -0.04 \pm 0.08$, $\sigma^{\rm HSE} = 0.31 \pm 0.05$, and $\sigma^{\rm lens} = 0.37 \pm 0.06$. According to Sereno et al. (2019), the difference between HSE and lensing masses is of $b = 0.09 \pm 0.17$ (VIII). The $\alpha^{\rm HSE}$ from Sereno et al. (2019) is at 3σ from our result with the full reference sample ($\alpha^{\rm HSE} = -0.338^{+0.105}_{-0.097}$ without accounting for the systematic uncertainties). Their values for $\sigma^{\rm HSE}$ and $\sigma^{\rm lens}$ agree with the intrinsic scatter values that we obtain when we do not account for the systematic uncertainties ($\sigma^{\rm HSE} = 0.304^{+0.069}_{-0.072}$ and $\sigma^{\rm lens} = 0.305^{+0.080}_{-0.083}$).

In addition, the behaviour of the HSE-to-lensing mass bias could vary with the overdensity at which masses are measured. By estimating weak lensing masses and HSE masses from X-rays at R_{200}, Jee et al. (2011) concluded that for a sample of 14 very massive and distant clusters ($0.83 < z < 1.46$), the HSE and lensing masses are compatible. However, the HSE masses were obtained from the extrapolation of a singular isothermal sphere profile to reach R_{200}, which likely limits the validity of their HSE mass estimates. Similarly, in Amodeo et al. (2016) authors compared M_{200} masses reconstructed from *Chandra* data (although the radial reach of *Chandra* is way below R_{200}) to their lensing estimates, and concluded that both mass estimates are in agreement. No evolution with redshift was detected in Amodeo et al. (2016). We prefer to avoid extrapolating the mass profiles to reach R_{200}.

7. Summary and conclusions

In this work we have investigated the HSE-to-lensing mass bias with masses inferred at R_{500} from resolved profiles. We carefully selected the clusters and obtained a reference sample with 53 clusters with redshifts spanning from $z = 0.05$ to 1.07. This is the largest redshift range analysed homogeneously with this type of data, having access to X-ray HSE masses obtained from resolved profiles. HSE masses were estimated with the *XMM-Newton* mass reconstruction reference pipeline and lensing masses were extracted from the LC^2 CoMaLit catalogue.

In order to account for possible systematic effects in the reference analysis, we compared the *XMM-Newton* and CoMaLit masses to other estimates from the literature. The obtained systematic scatters were propagated to our analyses, but all the main conclusions remain unchanged when considering or not these additional systematic dispersions on the HSE and lensing mass uncertainties.

We performed different tests in the measurement of the HSE-to-lensing mass scaling relation and bias, varying the redshift range and the scaling relation model. Our main conclusions are the following:

1. Assuming that HSE and lensing masses scale linearly with the true mass and considering $\sigma^2_{\rm sys\ HSE}$ and $\sigma^2_{\rm sys\ lens}$, we measure for the 53 clusters in the reference sample a HSE-to-lensing mass ratio of $M_{500}^{\rm HSE}/M_{500}^{\rm lens} = (1-b) = 0.739^{+0.075}_{-0.070}$ (stat.) $\pm$ 0.226 (intrin. scatter).
2. We find that the best scaling relation between HSE and lensing masses is our scaling relation of reference, where we assume that there is no evolution with mass and redshift and that HSE and lensing masses scale linearly. We obtain: $\alpha^{\rm HSE} = -0.303^{+0.101}_{-0.095}$, $\sigma^{\rm HSE} = 0.166^{+0.086}_{-0.101}$, and $\sigma^{\rm lens} = 0.257^{+0.080}_{-0.092}$.
3. When we let the SR evolve with redshift, we observe a trend towards a larger discrepancy between HSE and lensing masses at high redshift, but it is not statistically significant. In conclusion, there is no evidence of evolution with redshift. The dependence of the HSE-to-lensing mass bias on the mass of the clusters is not confirmed either.
4. Given the size of the sample, single clusters can be driving the fits and special care needs to be taken for clusters with very small uncertainties. We have investigated the case of CL J1226.9+3332 galaxy cluster, whose impact is crucial when determining the bias at high redshift.
5. Ignoring the intrinsic scatter of HSE and lensing masses with respect to the true mass of clusters introduces a bias in the measurement of the HSE-to-lensing mass bias.

Additional considerations are needed to compare the HSE-to-lensing mass bias obtained in this work to the bias needed to reconcile cluster number counts and CMB power spectrum results for several reasons: (1) the HSE masses used in cluster number count analyses are not direct HSE mass measurements, but masses obtained from a SZ (or X-ray) measurement through a SZ-mass (or X-ray–mass) scaling relation, (2) lensing masses

can also be biased with respect to the true mass of clusters (Becker & Kravtsov 2011), and (3) this sample is not representative of the cluster population in any given survey. Instead, this study provide a step forwards in our understanding of the deviation from hydrostatic equilibrium of galaxy clusters and of the impact of systematic and intrinsic errors.

Acknowledgements. We thank M. Sereno and S. Ettori for very useful comments. This project was carried out using the Python libraries `matplotlib` (Hunter 2007), `numpy` (Harris et al. 2020), `astropy` (Astropy Collaboration 2013, 2018), `scipy` (Virtanen et al. 2020), and `ChainConsumer` (Hinton 2016). This work benefited the financial support from CNRS/INSU and CNES. A.F. also thanks financial support by Universidad de La Laguna (ULL), NextGenerationEU/PRTR and Ministerio de Universidades (MIU) (UNI/511/2021) through grant "Margarita Salas".

References

Abdalla, E., Abellán, G. F., Aboubrahim, A., et al. 2022, J. High Energy Astrophys., 34, 49
Adam, R., Adane, A., Ade, P. A. R., et al. 2018, A&A, 609, A115
Aguado-Barahona, A., Rubiño-Martín, J. A., Ferragamo, A., et al. 2022, A&A, 659, A126
Akaike, H. 1974, IEEE Trans. Automatic Control, 19, 716
Akritas, M. G., & Bershady, M. A. 1996, ApJ, 470, 706
Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, ARA&A, 49, 409
Amodeo, S., Ettori, S., Capasso, R., & Sereno, M. 2016, A&A, 590, A126
Applegate, D. E., von der Linden, A., Kelly, P. L., et al. 2014, MNRAS, 439, 48
Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010, A&A, 517, A92
Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123
Baltz, E. A., Marshall, P., & Oguri, M. 2009, J. Cosmol. Astropart. Phys., 2009, 015
Bartalucci, I., Arnaud, M., Pratt, G. W., et al. 2017a, A&A, 608, A88
Bartalucci, I., Arnaud, M., Pratt, G. W., et al. 2017b, A&A, 598, A61
Bartalucci, I., Arnaud, M., Pratt, G. W., & Le Brun, A. M. C. 2018, A&A, 617, A64
Baxter, E. J., Keisler, R., Dodelson, S., et al. 2015, ApJ, 806, 247
Becker, M. R., & Kravtsov, A. V. 2011, ApJ, 740, 25
Biffi, V., Borgani, S., Murante, G., et al. 2016, ApJ, 827, 112
Bolliet, B., Comis, B., Komatsu, E., & Macías-Pérez, J. F. 2018, MNRAS, 477, 4957
Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, MNRAS, 321, 559
Costanzi, M., Rozo, E., Simet, M., et al. 2019, MNRAS, 488, 4779
Démoclès, J., Pratt, G. W., Pierini, D., et al. 2010, A&A, 517, A52
Ebeling, H., Edge, A. C., Bohringer, H., et al. 1998, MNRAS, 301, 881
Eckert, D., Ghirardini, V., Ettori, S., et al. 2019, A&A, 621, A40
Eckert, D., Ettori, S., Pointecouteau, E., van der Burg, R. F. J., & Loubser, S. I. 2022, A&A, 662, A123
Ettori, S., Gastaldello, F., Leccardi, A., et al. 2010, A&A, 524, A68
Ettori, S., Gastaldello, F., Leccardi, A., et al. 2011, A&A, 526, C1
Ferragamo, A., Barrena, R., Rubiño-Martín, J. A., et al. 2021, A&A, 655, A115
Ferragamo, A., Macías-Pérez, J. F., Pelgrims, V., et al. 2022, A&A, 661, A65
Foreman-Mackey, D., Farr, W., Sinha, M., et al. 2019, J. Open Source Softw., 4, 1864
Garrel, C., Pierre, M., Valageas, P., et al. 2022, A&A, 663, A3
Geller, M. J., Hwang, H. S., Diaferio, A., et al. 2014, ApJ, 783, 52
Gelman, A., & Rubin, D. B. 1992, Stat. Sci., 7, 457
Gianfagna, G., De Petris, M., Yepes, G., et al. 2021, MNRAS, 502, 5115
Gianfagna, G., Rasia, E., Cui, W., et al. 2023, MNRAS, 518, 4238
Goodman, J., & Weare, J. 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Herbonnet, R., Sifón, C., Hoekstra, H., et al. 2020, MNRAS, 497, 4684
Hernquist, L. 1990, ApJ, 356, 359
Hilton, M., Hasselfield, M., Sifón, C., et al. 2018, ApJS, 235, 20
Hilton, M., Sifón, C., Naess, S., et al. 2021, ApJS, 253, 3
Hinton, S. R. 2016, J. Open Source Softw., 1, 00045
Hoekstra, H. 2007, MNRAS, 379, 317
Hoekstra, H., Mahdavi, A., Babul, A., & Bildfell, C. 2012, MNRAS, 427, 1298
Hoekstra, H., Herbonnet, R., Muzzin, A., et al. 2015, MNRAS, 449, 685
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Hurier, G., & Lacasa, F. 2017, A&A, 604, A71
Huterer, D., Kirkby, D., Bean, R., et al. 2015, Astropart. Phys., 63, 23
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Israel, H., Reiprich, T., Erben, T., et al. 2014, A&A, 564, A129
Jee, M. J., Dawson, K. S., Hoekstra, H., et al. 2011, ApJ, 737, 59
Kéruzoré, F., Mayet, F., Pratt, G. W., et al. 2020, A&A, 644, A93
Kravtsov, A. V., Vikhlinin, A., & Nagai, D. 2006, ApJ, 650, 128
Landry, D., Bonamente, M., Giles, P., et al. 2013, MNRAS, 433, 2790
Lau, E. T., Nagai, D., & Nelson, K. 2013, ApJ, 777, 151
Lovisari, L., Ettori, S., Sereno, M., et al. 2020a, A&A, 644, A78
Lovisari, L., Schellenberger, G., Sereno, M., et al. 2020b, ApJ, 892, 102
Madhavacheril, M. S., Sifón, C., Battaglia, N., et al. 2020, ApJ, 903, L13
Mahdavi, A., Hoekstra, H., Babul, A., & Henry, J. P. 2008, MNRAS, 384, 1567
Mahdavi, A., Hoekstra, H., Babul, A., et al. 2013, ApJ, 767, 116
Martino, R., Mazzotta, P., Bourdin, H., et al. 2014, MNRAS, 443, 2342
Mayet, F., Adam, R., Ade, P., et al. 2020, EPJ Web Conf., 228, 00017
McCarthy, I. G., Schaye, J., Bower, R. G., et al. 2011, MNRAS, 412, 1965
Melin, J.-B., & Bartlett, J. G. 2015, A&A, 578, A21
Merten, J., Meneghetti, M., Postman, M., et al. 2015, ApJ, 806, 4
Monteiro-Oliveira, R., Doubrawa, L., Machado, R. E. G., et al. 2020, MNRAS, 495, 2007
Muñoz-Echeverría, M., Adam, R., Ade, P., et al. 2022, EPJ Web Conf., 257, 00033
Muñoz-Echeverría, M., Macias-Perez, J. F., Pratt, G. W., et al. 2023, A&A, 671, A28
Nagai, D., Kravtsov, A. V., & Vikhlinin, A. 2007, ApJ, 668, 1
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Okabe, N., & Smith, G. P. 2016, MNRAS, 461, 3794
Pacaud, F., Clerc, N., Giles, P. A., et al. 2016, A&A, 592, A2
Pedersen, K., & Dahle, H. 2007, ApJ, 667, 26
Penna-Lima, M., Bartlett, J. G., Rozo, E., et al. 2017, A&A, 604, A89
Perotto, L., Ponthieu, N., Macías-Pérez, J. F., et al. 2020, A&A, 637, A71
Perotto, L., Adam, R., Ade, P., et al. 2022, EPJ Web Conf., 257, 00038
Planck Collaboration VIII. 2011, A&A, 536, A8
Planck Collaboration XI. 2011, A&A, 536, A11
Planck Collaboration XV. 2014, A&A, 571, A15
Planck Collaboration XX. 2014, A&A, 571, A20
Planck Collaboration XXIV. 2015, A&A, 594, A24
Planck Collaboration XVIII. 2015, A&A, 594, A27
Planck Collaboration VI. 2020, A&A, 641, A6
Planck Collaboration Int. III. 2013, A&A, 550, A129
Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
Pratt, G. W., Croston, J. H., Arnaud, M., & Böhringer, H. 2009, A&A, 498, 361
Pratt, G. W., Arnaud, M., Piffaretti, R., et al. 2010, A&A, 511, A85
Pratt, G. W., Arnaud, M., Biviano, A., et al. 2019, Space. Sci. Rev., 215, 25
Rasia, E., Meneghetti, M., Martino, R., et al. 2012, New J. Phys., 14, 055018P
Ruppin, F., Mayet, F., Pratt, G. W., et al. 2018, A&A, 615, A112
Salvati, L., Douspis, M., & Aghanim, N. 2018, A&A, 614, A13
Salvati, L., Douspis, M., Ritz, A., Aghanim, N., & Babul, A. 2019, A&A, 626, A27
Salvati, L., Douspis, M., & Aghanim, N. 2020, A&A, 643, A20
Schellenberger, G., & Reiprich, T. H. 2017, MNRAS, 469, 3738
Schrabback, T., Applegate, D., Dietrich, J. P., et al. 2017, MNRAS, 474, 2635
Schwarz, G. 1978, Ann. Stat., 6, 461
Sereno, M. 2015, MNRAS, 450, 3665
Sereno, M. 2016, MNRAS, 455, 2149
Sereno, M., & Ettori, S. 2015, MNRAS, 450, 3633
Sereno, M., & Ettori, S. 2017, MNRAS, 468, 3322
Sereno, M., Covone, G., Izzo, L., et al. 2017, MNRAS, 472, 1946
Sereno, M., Umetsu, K., Ettori, S., et al. 2019, MNRAS, 492, 4528
Smith, G. P., Mazzotta, P., Okabe, N., et al. 2015, MNRAS, 456, L74
Sunyaev, R. A., & Zeldovich, Y. B. 1972, Comments Astro. Space Phys., 4, 173
Umetsu, K., Medezinski, E., Nonino, M., et al. 2014, ApJ, 795, 163
Vikhlinin, A., Kravtsov, A. V., Burenin, R. A., et al. 2009, ApJ, 692, 1060
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nat. Methods, 17, 261
von der Linden, A., Mantz, A., Allen, S. W., et al. 2014, MNRAS, 443, 1973
Wicker, R., Douspis, M., Salvati, L., & Aghanim, N. 2023, A&A, 674, A48
Yuan, Z. S., Han, J. L., & Wen, Z. L. 2022, MNRAS, 513, 3013
Zhang, Y.-Y., Okabe, N., Finoguenov, A., et al. 2010, ApJ, 711, 1033
Zhao, H. 1996, MNRAS, 278, 488
Zitrin, A., Broadhurst, T., Umetsu, K., et al. 2009, MNRAS, 396, 1985
Zitrin, A., Menanteau, F., Hughes, J. P., et al. 2013a, ApJ, 770, L15
Zitrin, A., Meneghetti, M., Umetsu, K., et al. 2013b, ApJ, 762, L30
Zitrin, A., Fabris, A., Merten, J., et al. 2015, ApJ, 801, 44
Zubeldia, I., & Challinor, A. 2019, MNRAS, 489, 401

Appendix A: Measuring the systematic scatter

We present in Fig. A.1 different measurements of the systematic scatter for HSE masses and lensing masses obtained from the reference analyses. Top panels do not account for the HSE masses obtained by evaluating the HSE mass profiles at the R_{500} from lensing. In the bottom panels in Fig. A.1 we present the systematic scatters once the cluster selection criteria from Sect. 3.3 have been applied. We summarise all the scatter values in Table A.1.

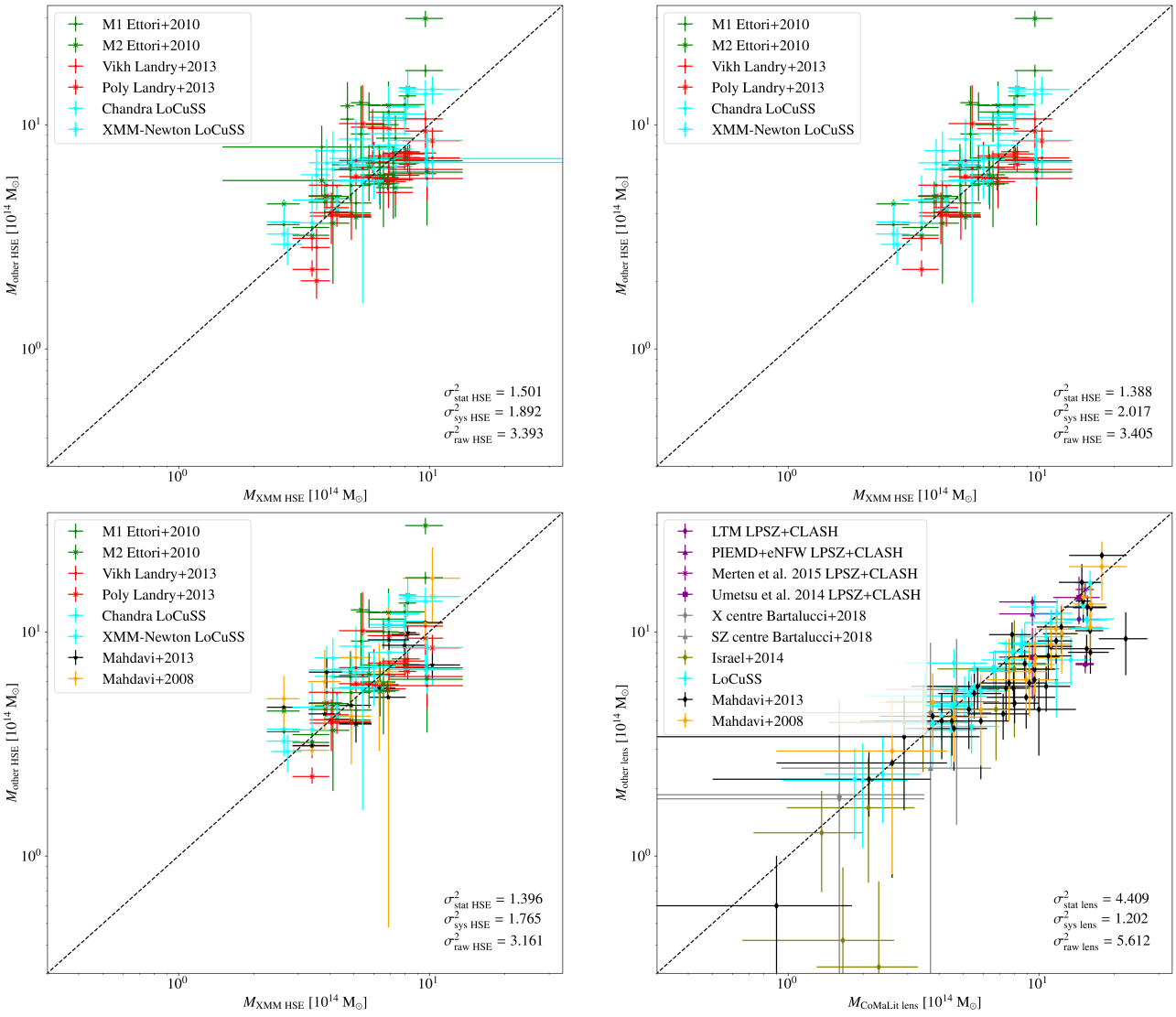

Fig. A.1. Comparison of HSE and lensing mass estimates from the homogeneous samples in this work (*XMM-Newton* and CoMaLit) with respect to other estimates from the literature (comparison sample). Top: Relation between X-ray masses from literature and from the *XMM-Newton* reference pipeline without accounting for $M^{\rm HSE}(< R^{\rm lens}_{500})$. In the left (right) the clusters with very large uncertainties and with very different *XMM-Newton* and CoMaLit centres are considered (not considered). Bottom: Same figure as Fig. 1, but not accounting for clusters with very large uncertainties and with very different *XMM-Newton* and CoMaLit centres. The dashed lines show the 1:1 relation. We give the statistical, systematic, and raw variances in units of 10^{28} M$^2_\odot$ corresponding to the data points in each figure.

Table A.1. Raw, statistical, and systematic variances of the HSE and lensing mass estimates from the homogeneous samples in this work (*XMM-Newton* and CoMaLit) with respect to other estimates from the literature (comparison samples).

Sample	$\sigma^2_{\text{raw HSE}}$ $[10^{28}\ M_\odot^2]$	$\sigma^2_{\text{stat HSE}}$ $[10^{28}\ M_\odot^2]$	$\sigma^2_{\text{sys HSE}}$ $[10^{28}\ M_\odot^2]$	Figure
All	3.231	1.507	1.724	Fig. 1
Without $M^{\text{HSE}}(< R^{\text{lens}}_{500})$	3.393	1.501	1.892	Fig. A.1
Without clusters discarded in Sect. 3.3	3.161	1.396	1.765	Fig. 1
Without $M^{\text{HSE}}(< R^{\text{lens}}_{500})$ and clusters discarded in Sect. 3.3	3.405	1.388	**2.017**	Fig. A.1
	$\sigma^2_{\text{raw lens}}$ $[10^{28}\ M_\odot^2]$	$\sigma^2_{\text{stat lens}}$ $[10^{28}\ M_\odot^2]$	$\sigma^2_{\text{sys lens}}$ $[10^{28}\ M_\odot^2]$	Figure
All	5.280	4.340	0.940	Fig. 1
Without clusters discarded in Sect. 3.3	5.612	4.409	**1.202**	Fig. 1

Notes. We report the different values depending on the sample selection criteria, showing in bold the systematic scatters considered for the rest of the analysis.

Appendix B: Additional checks for the selection of the reference sample

In this section we detail the additional checks we have performed in Sect. 3.3 to verify which clusters can be used for the HSE-lensing mass comparison and define the final sample.

B.1. Uncertainties of mass estimates

Regarding the uncertainties of the mass estimates in the homogeneous sample (the 65 clusters in Sect. 3.3), we observe that masses and their uncertainties appear correlated, error bars being larger for more massive objects. As already mentioned, lensing mass uncertainties are larger than HSE ones and this is also true for relative uncertainties when calculated with respect to the value of the measured mass. For some clusters, the uncertainties on HSE masses are suspiciously large (larger than the value of the mass) or abnormal (negative error bars) and we decide to exclude for these reasons the clusters Abell119, Abell521 (the LoCuSS and Mahdavi+2013 cluster out of the plot in the left panel in Fig. 1), and SPT-CLJ0516-54.

B.2. Difference of centres for HSE and lensing masses

For some of the clusters, the HSE and lensing masses have been reconstructed assuming different cluster centre positions. We investigate in the following if this miscentring is correlated to the mass ratio and redshift and how it may affect the mass estimates.

B.2.1. Correlation between miscentring and bias or redshift

In the left panel in Fig. B.1 we present the ratios between the HSE and lensing masses of the clusters with respect to the separations between the centres considered in the X-ray and lensing analyses. Each marker corresponds to one of the 65 clusters, showing in magenta crosses, purple squares, and grey circles the clusters from ESZ+LoCuSS, LPSZ, and Bartalucci+2018 samples, respectively. Error bars have been obtained from the propagation of the individual *XMM-Newton* and CoMaLit uncertainties. The separation between X-ray and lensing centres goes from 1.5 kpc to 700 kpc. However, there is no significative correlation between the miscentring and the HSE-to-lensing mass ratio.

The miscentring could also be related to the redshift of the cluster. In the right panel in Fig. B.1 we show the redshift with respect to the separation of the centres. There is neither indication of correlation. We do not find any significative correlation between the redshift and the uncertainties of mass ratios either.

The dynamical state of clusters is an important point to better understand the evolution of the HSE bias. In particular, relaxed clusters tend to have smaller HSE bias as compared to disturbed ones (Gianfagna et al. 2021). The offset between the centres used in lensing and X-ray analyses could be an indicator of the departure from sphericity and equilibrium of clusters, resulting in a difficulty to define the centre. However, the absence of correlation between the separation of centres and the bias or the redshift does not enable any clear dynamical classification of the clusters in our sample.

B.2.2. Simulations to quantify the impact of miscentring

It is evident that two mass estimates computed assuming very different centres are hardly comparable. Nevertheless, setting a quantitative limit of the acceptable separation between the two centres can be difficult. In this section we make use of simulated mock mass density profiles to check the impact that miscentring has on the mass profiles and, consequently, on M_{500}.

We simulate mass density profiles using the `profiley`[9] Python package, which contains already tested (Madhavacheril et al. 2020) functions that describe density profiles, as well as the corresponding miscentred density profiles. We test a variety of density profile shapes following NFW, gNFW, truncated Navarro-Frenk-White (tNFW, Baltz et al. 2009), and Hernquist (Hernquist 1990) models. To initialise the density profiles we consider different M_{500} (in the range ~ 1 to $9 \times 10^{14}\ M_\odot$) and model parameters. The parameters considered for each model are given in Table B.1.

For each density profile we build the corresponding miscentred profiles by displacing the centre by 0.0 to 0.7 Mpc from the original one. We integrate each miscentred density profile to get a mass profile and obtain the miscentred M_{500} estimate at the radius where the overdensity reaches $\Delta = 500$, M^{miscent}_{500}. In Fig. B.2 we show the relative error of the miscentred masses with respect to the separation to the true centre, for all type of profiles and for different redshifts, $z = [0.1, 0.45, 0.8, 1.15, 1.5]$. Each marker indicates a different mass density model and the colours show the true M_{500}. For visualisation purposes gNFW, tNFW, and Hernquist markers have been shifted by 0.01, 0.02,

[9] https://profiley.readthedocs.io/en/latest/index.html.

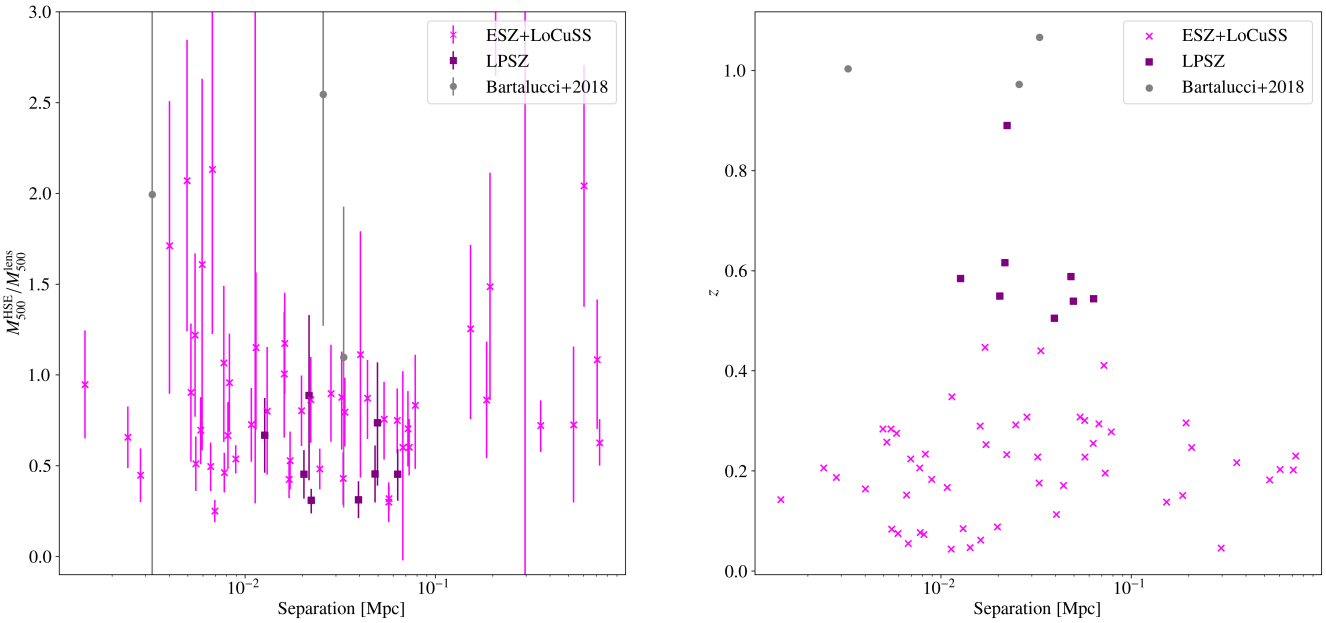

Fig. B.1. Correlation between the HSE-to-lensing mass ratio (left) and redshift (right) of the 65 clusters in the *XMM-Newton*-CoMaLit homogeneous sample with respect to the separation between the centres assumed in the X-ray and lensing analyses. Error bars in the left panel do not account for the systematic scatters.

and 0.03 Mpc. As expected, the figure shows that computing the density profile from a centre that is separated from the true centre gives biased mass reconstructions. This bias increases with the separation, less massive clusters being more sensitive to this effect. The black dashed line in Fig. B.2 indicates $(M_{500}^{\rm miscent} - M_{500})/M_{500} = 0.5$.

Considering the tested density profiles and mass range, the error due to miscentring is smaller than 50% only if the distance to the real centre is smaller than 100 kpc. Therefore, we decide to exclude clusters for which X-ray and lensing centres are separated by more than 100 kpc. These clusters are: Abell3856, Abell3888, Abell773, Abell665, Abell267, 1E0657-56, Abell521, Abell3376E, Abell520, and Abell2163.

B.3. Final sample

We present in Table B.2 the 53 clusters in the reference sample. We give their names (as named in the CoMaLit LC^2 catalogue), redshifts, masses, and mass uncertainties from the analyses of reference, that is, *XMM-Newton* mass reconstruction pipeline masses and CoMaLit estimates.

The histograms in Fig. B.3 show the distributions of the measurement uncertainties for the *XMM-Newton* and CoMaLit masses of all the clusters in the reference sample. The vertical red lines indicate the measured systematic scatter for *XMM-Newton* and CoMaLit masses.

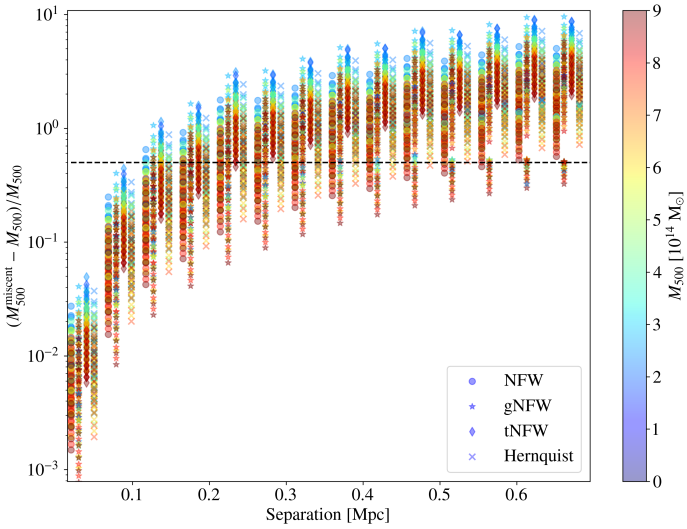

Fig. B.2. Relative error of M_{500} with respect to the separation between the true and the considered centre for the mass reconstruction. Colours indicate the true M_{500} and markers the mass density model. We consider the same separation for the four models, but gNFW, tNFW, and Hernquist markers have been shifted for visualisation purposes. The black dashed line indicates a 50% error in M_{500}.

Table B.1. Different mass density models and their corresponding parameters used to study the miscentring effect in Fig. B.2.

Model	c_{500}	α	β	γ	τ	η
NFW	1.0, 2.25, 3.5, 4.75, 6.0					
Hernquist	1.0, 2.25, 3.5, 4.75, 6.0					
gNFW	6.0	0.3, 0.85, 1.4, 1.95, 2.5	2.0, 3.0, 4.0, 5.0, 6.0	0.3, 0.475, 0.65, 0.825, 1.0		
tNFW	6.0				0.2, 1.15, 2.1, 3.05, 4.0	0.5, 1.125, 1.75, 2.375, 3.0

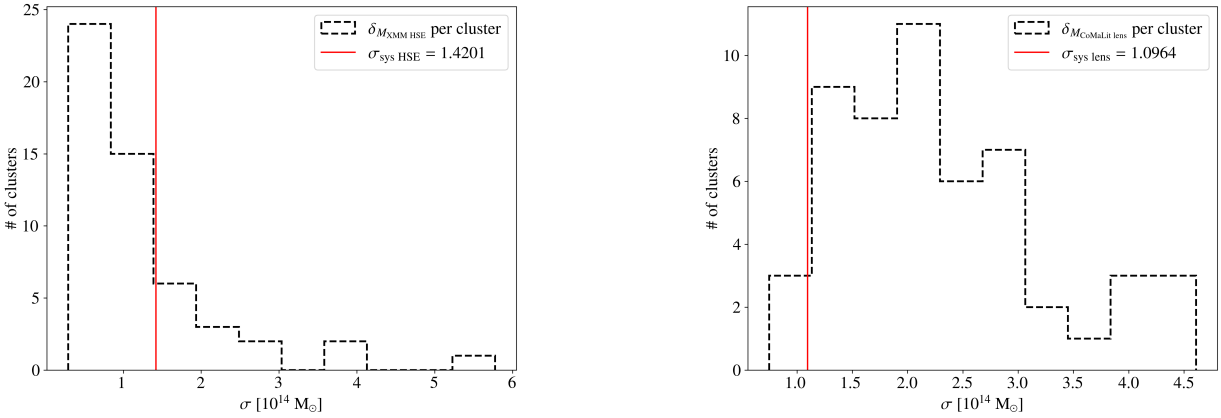

Fig. B.3. Comparison between statistical uncertainties and systematic scatter. Dashed line histograms show the statistical uncertainties of all HSE mass estimates from *XMM-Newton* reference analysis (left) and CoMaLit lensing mass estimates (right). Vertical red lines give the systematic scatter estimated with respect to other published results in Sect. 3.3.

Table B.2. Characteristics of the reference sample.

CoMaLit name	z	Centre coordinates [deg]				Masses [10^{14} $M_\odot$]			
		CoMaLit α	CoMaLit δ	XMM α	XMM δ	$M_{500}^{\text{CoMaLit lens}}$	$\delta_{M_{\text{CoMaLit lens}}}$	$M_{500}^{\text{XMM HSE}}$	$\delta_{M_{\text{XMM HSE}}}$
ABELL2744	0.3080	3.57875	-30.38944	3.57752	-30.38633	13.849	2.824	10.472	2.218
CL0016+16	0.5490	4.63933	16.43694	4.63986	16.43622	17.821	4.609	8.062	1.147
ABELL68	0.2550	9.27479	9.16000	9.27792	9.15685	9.171	1.587	6.872	1.181
ABELL85	0.0550	10.46017	-9.30306	10.46132	-9.30438	5.700	2.200	12.150	2.169
ABELL2813	0.2924	10.85167	-20.62139	10.85190	-20.62294	8.557	1.450	4.131	0.665
Abell209	0.2060	22.96892	-13.61122	22.96912	-13.61121	9.614	1.965	6.309	1.010
ABELL291	0.1960	30.43417	-2.20083	30.42942	-2.19678	4.514	0.986	2.718	0.380
RXCJ0232.2-4420	0.2840	38.07750	-44.34667	38.07721	-44.34638	5.380	1.815	6.558	0.986
Abell383	0.1870	42.01417	-3.52914	42.01418	-3.52889	5.871	1.727	2.629	0.409
ABELL478	0.0881	63.35667	10.46694	63.35581	10.46371	8.772	2.078	7.043	0.325
MS0451.6-0305	0.5389	73.54767	-3.01411	73.54705	-3.01620	9.994	4.060	7.352	1.733
RXCJ0528.9-3927	0.2840	82.22083	-39.47161	82.22109	-39.47136	4.480	1.310	9.276	2.541
RXCJ0532.9-3701	0.2750	83.23208	-37.02667	83.23226	-37.02703	6.960	1.430	4.847	0.852
SPT-CLJ0546-5345	1.0660	86.65321	-53.76039	86.65500	-53.76000	3.700	2.760	4.060	0.500
SPT-CLJ0615-5746	0.9720	93.96521	-57.77881	93.96600	-57.77960	4.700	2.252	11.960	1.750
ABELL3404	0.1670	101.37292	-54.22697	101.37122	-54.22732	8.750	2.085	6.360	0.968
MACSJ0647.7+7015	0.5840	101.95946	70.24861	101.95900	70.24810	9.427	2.493	6.296	1.014
MACSJ0911.2+1746	0.5050	137.79529	17.77539	137.79704	17.77603	10.862	3.259	3.392	0.424
ABELL963	0.2060	154.26483	39.04764	154.26530	39.04816	4.583	1.637	4.884	0.968
ABELL1300	0.3080	172.97583	-19.92772	172.97749	-19.92848	5.950	1.695	5.341	0.476
MACSJ1149.5+2223	0.5440	177.39871	22.39850	177.39763	22.40108	14.447	3.034	6.536	1.608
ABELL1413	0.1430	178.82500	23.40503	178.82495	23.40487	7.200	2.100	6.816	0.815
MACSJ1206.2-0847	0.4400	181.55062	-8.80094	181.55208	-8.80017	12.176	2.477	9.681	1.203
ZwCl1215.1+0400	0.0750	184.42137	3.65589	184.42236	3.65650	3.500	2.200	5.629	0.566
CLJ1226.9+3332	0.8900	186.74271	33.54683	186.74203	33.54627	15.298	1.275	4.732	1.042
ABELL1576	0.3010	189.24583	63.19056	189.24408	63.18711	13.543	4.243	4.064	0.782
ABELL1644S	0.0470	194.30000	-17.41306	194.29601	-17.41110	1.309	0.748	13.664	5.775
ABELL1650	0.0840	194.67287	-1.76139	194.67336	-1.76223	7.100	2.000	3.629	0.315
ABELL1651	0.0850	194.84371	-4.19603	194.84380	-4.19831	5.600	2.400	4.486	0.429
ABELL1689	0.1830	197.87300	-1.34100	197.87263	-1.34172	15.033	1.025	8.071	1.078
ABELL1763	0.2279	203.82583	40.99694	203.82979	41.00010	16.014	2.050	5.102	0.779
ABELL1795	0.0620	207.21871	26.59300	207.22115	26.58994	9.300	2.200	10.910	0.291
ABELL1835	0.2530	210.25804	2.87775	210.25916	2.87824	15.510	4.503	8.200	0.660
PSZ2G099.86+58.45	0.6160	213.69662	54.78433	213.69522	54.78396	7.242	3.043	6.421	2.038
ABELL1914	0.1712	216.50667	37.82722	216.51053	37.82434	7.929	1.293	6.909	1.390
ZwCl1454.8+2233	0.2578	224.31292	22.34278	224.31293	22.34242	3.771	1.457	3.407	0.584
Zwicky7215	0.2900	225.34483	42.34750	225.34451	42.34650	5.390	1.504	5.418	1.142
ABELL2034	0.1130	227.54875	33.51472	227.55283	33.51044	5.169	3.100	5.750	0.643
ABELL2029	0.0770	227.73371	5.74481	227.73502	5.74410	12.100	2.500	5.592	0.661
ABELL2065	0.0730	230.62150	27.70769	230.62257	27.70901	8.000	2.100	5.326	0.474
ABELL2204	0.1520	248.19650	5.57583	248.19592	5.57544	16.051	2.963	7.966	1.605
ABELL2218	0.1760	248.95329	66.21417	248.95972	66.21254	8.900	2.700	3.829	0.597
ABELL2219	0.2280	250.08475	46.70833	250.08366	46.71067	11.729	1.852	10.287	2.969
RXJ1720.1+2638	0.1640	260.04167	26.62464	260.04166	26.62503	3.510	1.485	6.006	1.314
Abell2261	0.2240	260.61325	32.13258	260.61267	32.13237	15.613	3.043	3.900	0.590
MACSJ2129.4-0741	0.5880	322.35717	-7.69189	322.35913	-7.69133	13.486	3.890	6.127	1.179
RXJ2129.7+0005	0.2340	322.41650	0.08922	322.41660	0.08861	4.470	1.158	4.277	0.491
ABELL2390	0.2330	328.40446	17.69594	328.40308	17.69493	11.183	2.396	9.652	1.668
RXJ2228.6+2037	0.4110	337.13658	20.62072	337.14047	20.62040	9.728	2.626	6.844	0.819
MACSJ2243.3-0935	0.4470	340.83933	-9.59522	340.83868	-9.59470	20.294	3.865	8.625	1.317
RXJ2248.7-4431	0.3480	342.18317	-44.53092	342.18243	-44.53054	12.400	3.605	14.262	3.762
ABELL2631	0.2780	354.40971	0.27069	354.40634	0.26678	11.748	1.888	9.785	3.807
SPT-CLJ2341-5119	1.0030	355.30092	-51.32850	355.30100	-51.32860	1.600	1.890	3.190	0.750

Notes. Column 1: cluster names from the CoMaLit catalogue (entries Comalit_Name and Comalit_Num). Column 2: redshift. Columns 3 to 6: right ascension α and declination δ of the cluster centres according to CoMaLit or X-rays. Columns 7 to 10: cluster masses and uncertainties from the CoMaLit catalogue and from the *XMM-Newton* analysis.

Appendix C: Different scaling relation models

C.1. Additional figures

In this section we present additional figures of the SR model extension fits described in Sect. 5.2. We show in Fig. C.1 and C.2 the scaling relations and posterior parameter distributions for different subsamples defined according to the redshift range, in these cases not including the systematic scatters in the error bars of each cluster. In Fig. C.1 we present the scaling relations for the samples with $z < 0.2$ and $z > 0.2$, while the panels in Fig. C.2 show the results for a $z < 0.5$ sample.

Figures C.3 and C.4 compare the SR and parameters posterior distributions when considering and not redshift evolution.

C.2. Comparison of SR models with AIC and BIC

In Fig. C.5 we present the $\hat{\chi}^2$ (defined in Eq. 15) distributions normalised by the degrees of freedom (d.o.f.) in each SR model fit as solid line histograms. In the left panel we present the results for the scaling relations fitted in $\ln M^{\rm HSE} - \ln M^{\rm lens}$ and in the right for $M^{\rm HSE} - M^{\rm lens}$. The blue histogram in the left panel shows the reduced $\hat{\chi}^2$ for the scaling relation of reference in this work and the vertical solid line is the median value of the distribution. For comparison, the dashed lines correspond to the χ^2-distributions,

$$f(\chi^2) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} e^{-\chi^2/2} (\chi^2)^{(\nu/2)-1}, \tag{C.1}$$

where ν is the number of the degrees of freedom (d.o.f.). For the scaling relation of reference we have d.o.f. = 53 − 3 = 50. The red and green results show the reduced $\hat{\chi}^2$ for the fits of the scaling relations when considering a deviation from linearity and an evolution with redshift, respectively. The histograms follow fairly well the χ^2-distributions with ν degrees of freedom.

We use the Akaike information criterion (AIC, Akaike 1974) and the Bayesian information criterion (BIC, Schwarz 1978) to compare the improvement in the $\hat{\chi}^2$ when adding parameters to the model. We calculate:

$$\mathrm{AIC} = \hat{\chi}^2_{\rm min} + 2K, \tag{C.2}$$

and

$$\mathrm{BIC} = \hat{\chi}^2_{\rm min} + K \ln N, \tag{C.3}$$

where $\hat{\chi}^2_{\rm min}$ is the minimum of the $\hat{\chi}^2$ values for each model. Here K and N are the number of free parameters in the model and the total number of data points, that is, $N = N_{\rm clusters} = 53$, respectively.

We report in Table C.1 the results for the different scaling relation models and the ΔAIC and ΔBIC differences with respect to the simplest scaling law amongst the nested models.

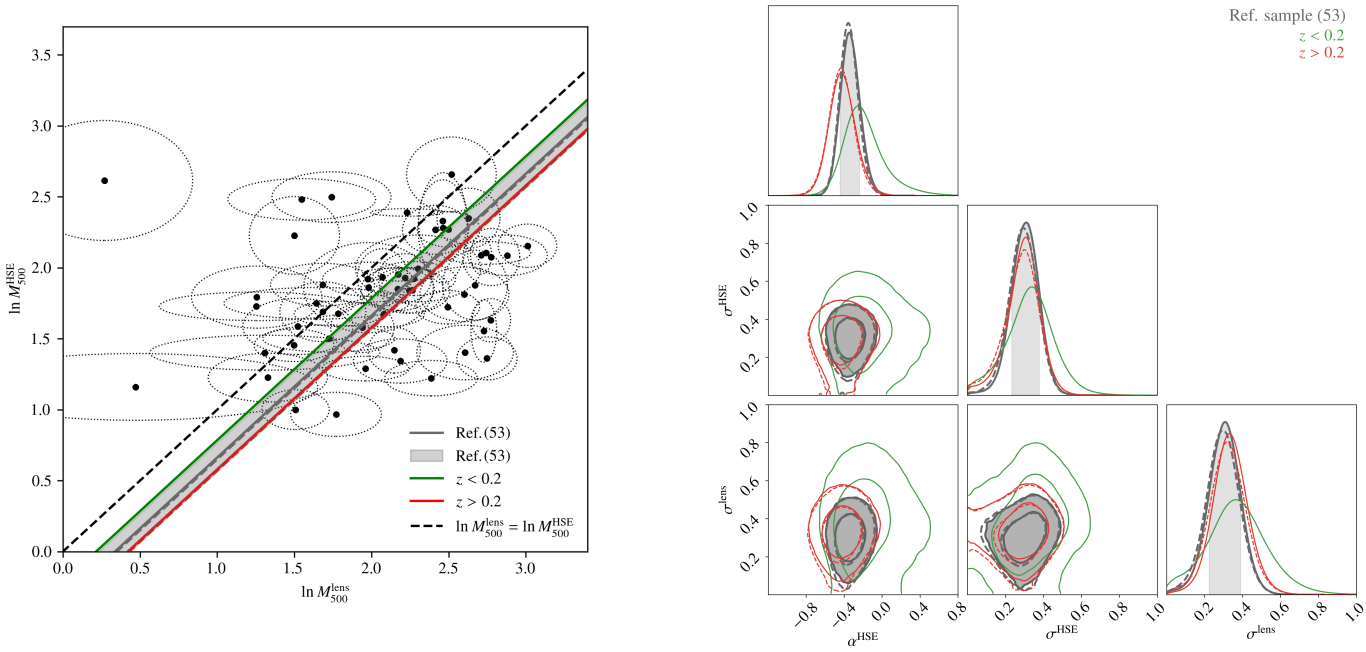

Fig. C.1. Same as Fig. 6, but without considering the systematic scatter in the fit.

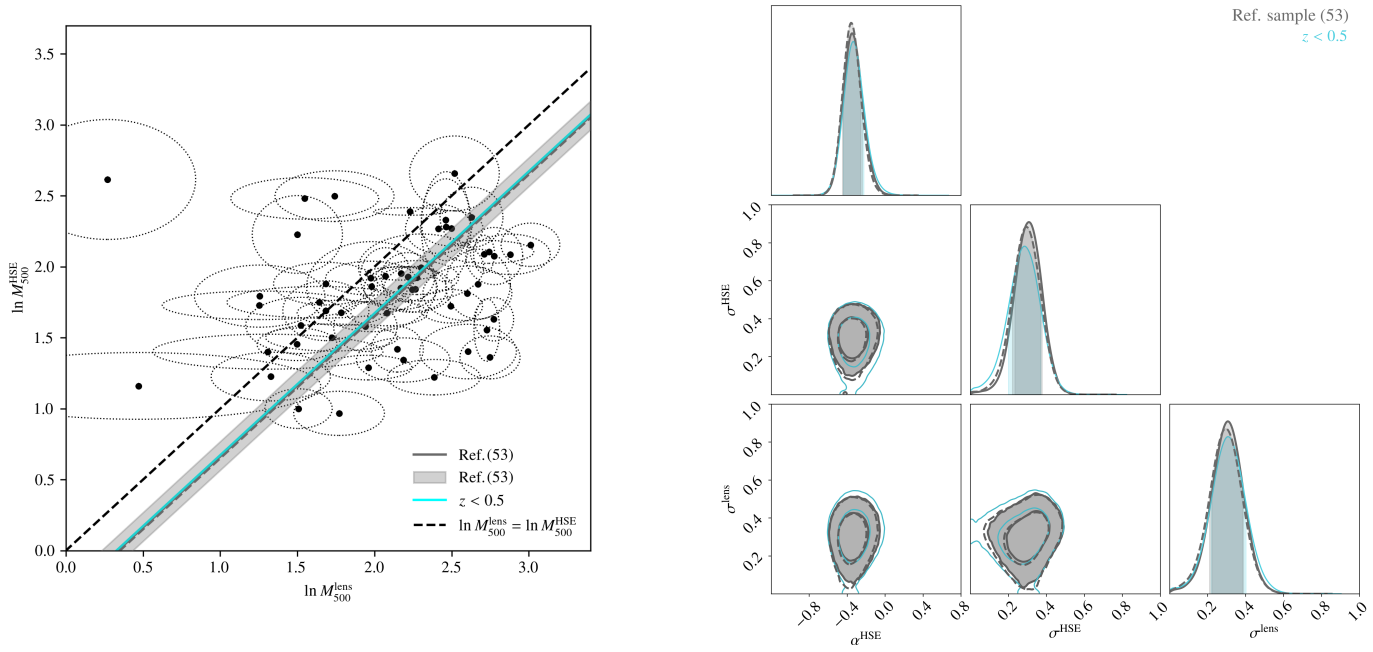

Fig. C.2. Scaling relation between HSE and lensing masses for the reference sample in grey and a subsample containing only $z < 0.5$ clusters in cyan, without considering the systematic scatter in the fit. The black dashed line shows the equality. The corner plots in the right panel are the posterior 1D and 2D distributions of the parameters in the SR. Here $\beta^{\rm HSE}$ is fixed to 1.

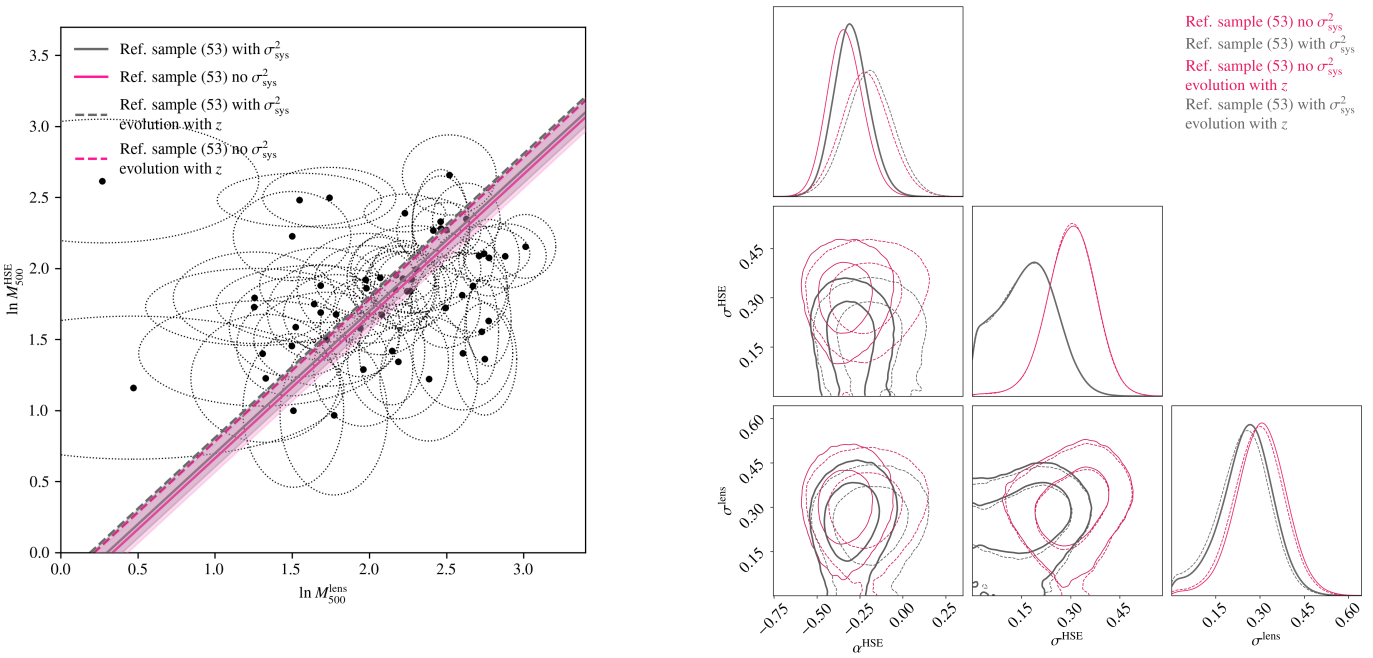

Fig. C.3. Same as Fig. 5 but with dashed lines showing the results if an evolution with redshift is considered in the scaling relation and solid lines without evolution.

Table C.1. Statistical results for the scaling relation models presented in this work.

Model	K	N	$\hat{\chi}^2_{\rm min}$	AIC	ΔAIC	BIC	ΔBIC
Reference SR	3	53	18.126	24.126	0.000	21.422	0.000
Deviation from linearity	4	53	16.020	24.020	-0.106	21.565	+0.143
Evolution with redshift	4	53	22.784	30.784	+6.658	28.329	+6.907
Considering an offset	4	53	13.929	21.993	0.000	19.475	0.000
Evolution with redshift and offset	5	53	17.992	27.992	+5.999	26.039	+6.564

Notes. We report the values for the fits accounting for the systematic scatter of HSE and lensing masses.

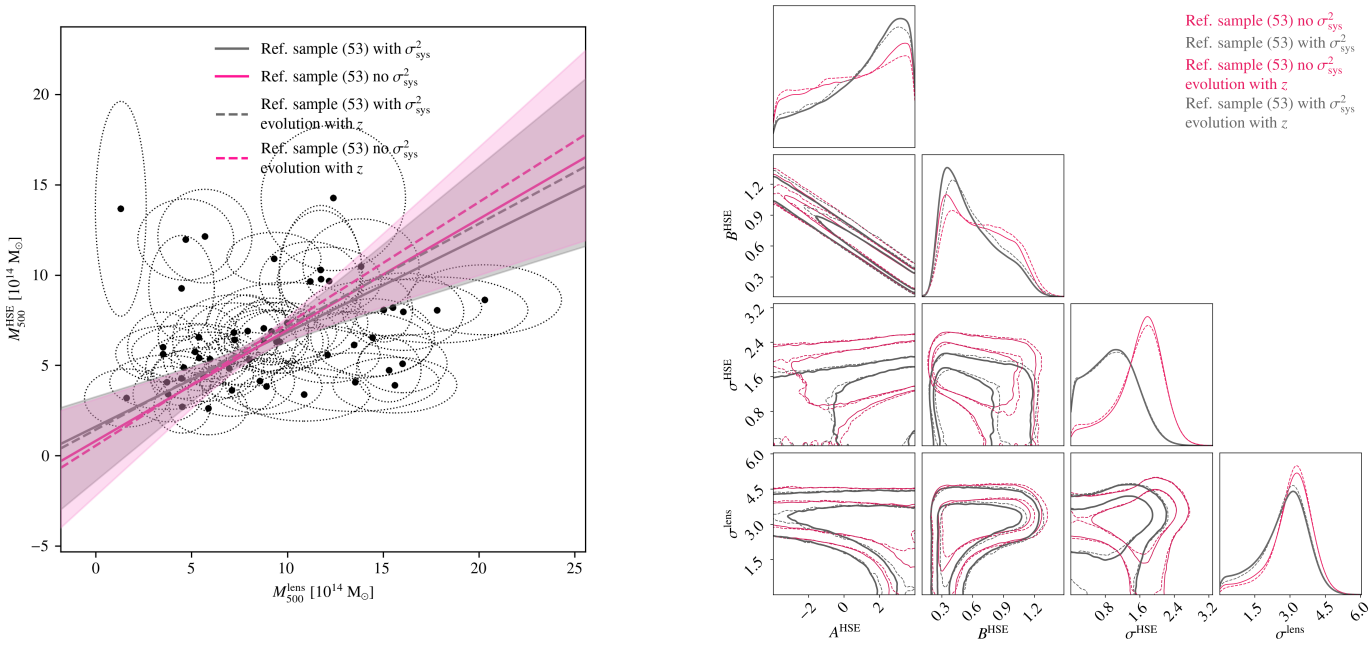

Fig. C.4. Same as Fig. 9 but with dashed lines showing the results if an evolution with redshift is considered in the scaling relation and solid lines without evolution.

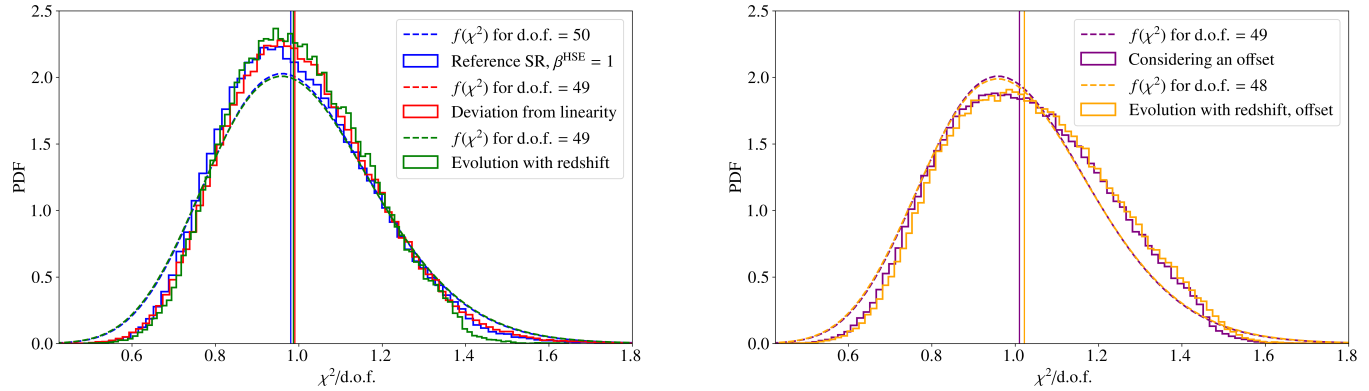

Fig. C.5. Comparison of the reduced $\hat{\chi}^2$ distributions for different scaling relation fits (solid colour histrograms) and the χ^2-distribution (Eq. C.1) for the degrees of freedom in each fit (dashed colour lines). Vertical lines show the median value of each histogram.

Appendix D: Other HSE mass estimates from *XMM-Newton* data

As described in Sect. 2.1.2, all the *XMM-Newton* HSE masses used in this work were obtained from mass profiles, by interpolating the binned profile to get M_{500} and R_{500}. For all the clusters in our reference sample we also have access to X-ray masses obtained with the $Y_X - M^{HSE}_{500}$ scaling relation from Arnaud et al. (2010). The quantity Y_X is defined as the product of the gas mass ($M_{g,500}$) and the spectroscopic temperature (T_X) and it is the X-ray analogue of the integrated Compton parameter (Kravtsov et al. 2006). We use these masses as M^{HSE} estimates and fit again the scaling relations in Eq. 9, 10, 11, and 12.

Median values and uncertainties of the fitted SR parameters are given in Table D.1. The intrinsic scatters of HSE and lensing masses with respect to the SR are reduced when using $Y_X - M^{HSE}_{500}$ estimates and the measured HSE bias is also smaller (larger $1 - b$).

For ESZ+LoCuSS and LPSZ sample clusters we have a third estimate of the HSE mass, obtained by fitting an NFW mass model to the X-ray profiles. Such masses tend to be in agreement with the interpolated ones (the reference *XMM-Newton* masses used in this work, Sect. 2.1.2), with a mean and median ratio for the 50 cluster estimates of $M_{500}/M^{NFW}_{500} \sim 1.11$ and 0.95, indicating that the HSE masses used in the analyses of reference are robust against modelling effects. Masses obtained by fitting an NFW model have in median 30% smaller uncertainties than the interpolated estimates. The scaling relations obtained with NFW fit masses (Table D.1) and interpolated ones (second row in Table 3 for the 50 clusters with $z < 0.9$) are almost identical, with α^{HSE} centred in the exact same value, but with 10 – 20% smaller intrinsic scatters when using NFW masses.

Table D.1. Summary of the median values and uncertainties at the 16th and 84th percentiles of the fitted parameters for the HSE-to-lensing mass SR for HSE masses obtained from the $Y_X - M^{HSE}_{500}$ scaling relation and from the fit of an NFW model to the X-ray data.

Sample	# of clusters	No σ^2_{sys}				
$Y_X - M^{HSE}_{500}$		α^{HSE}	$e^{\alpha^{HSE}} = (1-b)$	β^{HSE}	σ^{HSE}	σ^{lens}
Reference sample	53	$-0.237^{+0.095}_{-0.087}$	$0.789^{+0.075}_{-0.069}$	[1]	$0.105^{+0.063}_{-0.067}$	$0.247^{+0.058}_{-0.056}$
Reference sample	53	$-0.131^{+0.542}_{-0.631}$		$0.953^{+0.296}_{-0.263}$	$0.133^{+0.075}_{-0.088}$	$0.231^{+0.068}_{-0.098}$
Reference sample (BCES)	53	0.504 ± 0.262		0.662 ± 0.122	0.250^{*}	-
		A^{HSE} [10^{14} M$_\odot$]	B^{HSE}		σ^{HSE} [10^{14} M$_\odot$]	σ^{lens} [10^{14} M$_\odot$]
Reference sample	53	$0.990^{+1.748}_{-2.050}$	$0.672^{+0.235}_{-0.205}$		$0.930^{+0.508}_{-0.633}$	$2.333^{+0.594}_{-1.032}$
Reference sample (BCES)	53	2.401 ± 0.781	0.487 ± 0.095		1.480^{*}	-

NFW fit		α^{HSE}	$e^{\alpha^{HSE}} = (1-b)$	β^{HSE}	σ^{HSE}	σ^{lens}
$z < 0.9$	50	$-0.309^{+0.119}_{-0.107}$	$0.734^{+0.087}_{-0.079}$	[1]	$0.249^{+0.059}_{-0.063}$	$0.212^{+0.077}_{-0.086}$
$z < 0.9$	50	$-0.361^{+0.692}_{-0.955}$		$1.030^{+0.480}_{-0.341}$	$0.246^{+0.073}_{-0.121}$	$0.220^{+0.081}_{-0.119}$
$z < 0.9$ (BCES)	50	0.032 ± 0.485		0.836 ± 0.222	0.329^{*}	-
		A^{HSE} [10^{14} M$_\odot$]	B^{HSE}		σ^{HSE} [10^{14} M$_\odot$]	σ^{lens} [10^{14} M$_\odot$]
$z < 0.9$	50	$1.205^{+1.490}_{-2.463}$	$0.579^{+0.287}_{-0.173}$		$1.576^{+0.408}_{-0.589}$	$1.989^{+0.843}_{-1.164}$
$z < 0.9$ (BCES)	50	0.530 ± 1.343	0.646 ± 0.159		1.944^{*}	-

Notes. We present the results assuming linearity, a deviation from linearity, and an offset between HSE and lensing masses. For the BCES fit, we report the best-fit values and 1σ uncertainties. $^{(*)}$ We also calculate the scatter with respect to the best BCES scaling relations following Eq. 4.