XXL Survey groups and clusters in the Hyper Suprime-Cam Survey. Scaling relations between X-ray properties and weak lensing mass

Mauro Sereno, Keiichi Umetsu, Stefano Ettori, Dominique Eckert, Fabio Gastaldello, Paul Giles, Maggie Lieu, Ben Maughan, Nobuhiro Okabe, Mark Birkinshaw, et al.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **492,** 4528–4545 (2020)
Advance Access publication 2019 December 19

doi:10.1093/mnras/stz3425

XXL Survey groups and clusters in the Hyper Suprime-Cam Survey. Scaling relations between X-ray properties and weak lensing mass

Mauro Sereno [1,2]★ Keiichi Umetsu [3] Stefano Ettori [1,2] Dominique Eckert,[4] Fabio Gastaldello [5] Paul Giles,[6] Maggie Lieu [7] Ben Maughan [8] Nobuhiro Okabe [9,10,11] Mark Birkinshaw,[7] I-Non Chiu,[3] Yutaka Fujita [12] Satoshi Miyazaki,[13,14] David Rapetti [15,16,17] Elias Koulouridis[18,19] and Marguerite Pierre[18]

Affiliations are listed at the end of the paper

Accepted 2019 November 25. Received 2019 November 25; in original form 2019 September 26

ABSTRACT

Scaling relations trace the formation and evolution of galaxy clusters. We exploited multi-wavelength surveys – the XXL survey at *XMM-Newton* in the X-ray band, and the Hyper Suprime-Cam Subaru Strategic Program for optical weak lensing – to study an X-ray selected, complete sample of clusters and groups. The scalings of gas mass, temperature, and soft-band X-ray luminosity with the weak lensing mass show imprints of radiative cooling and active galactic nucleus feedback in groups. From the multi-variate analysis, we found some evidence for steeper than self-similar slopes for gas mass ($\beta_{m_g|m} = 1.73 \pm 0.80$) and luminosity ($\beta_{l|m} = 1.91 \pm 0.94$) and a nearly self-similar slope for the temperature ($\beta_{t|m} = 0.78 \pm 0.43$). Intrinsic scatters of X-ray properties appear to be positively correlated at a fixed mass (median correlation factor $\rho_{X_1X_2|m} \sim 0.34$) due to dynamical state and merger history of the haloes. Positive correlations with the weak lensing mass (median correlation factor $\rho_{m_{wl}X|m} \sim 0.35$) can be connected to triaxiality and orientation. Comparison of weak lensing and hydrostatic masses suggests a small role played by non-thermal pressure support (9 ± 17 per cent).

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

1 INTRODUCTION

Scaling relations between integrated properties of galaxy clusters open a window on the main mechanisms shaping the formation and evolution of cosmic structure (Voit 2005). They are also a key and often puzzling ingredient in cosmological studies of abundances and growth evolution, see e.g. Planck Collaboration et al. (2016) and Pacaud et al. (2018, hereafter XXL Paper XXV).

In the self-similar scenario, virialization is driven by gravity. The expected relations in the virial region between total mass (M), gas mass ($M_{\rm gas}$), temperature ($T_{\rm X}$), and soft-band X-ray luminosity ($L_{\rm X}$) are (Kaiser 1986; Giodini et al. 2013; Ettori 2015)

$$M_{\rm gas} \propto M \tag{1}$$

$$T_{\rm X} \propto E_z^{2/3} M^{2/3}, \tag{2}$$

$$L_{\rm X} \propto E_z^2 M. \tag{3}$$

Secondary infall and continuous mass accretion from the surrounding matter can perturb virial equilibrium (Bertschinger 1985), but scaling relations preserve the power-law structure (Fujita et al. 2018). Diversity of central structure and age of clusters contributes to the scatter of the $T_{\rm X} - M$ and $L_{\rm X} - M$ relations (Fujita et al. 2018; Fujita & Aung 2019). The mass dependence of the halo concentration and the Fundamental Plane relation of galaxy clusters make the $L_{\rm X} - T_{\rm X}$ and $L_{\rm X} - M$ relations shallower than the self-similar predictions (Fujita & Aung 2019). However, less massive objects formed earlier than more massive objects in the hierarchical structure formation scenario, and the effects of secondarily infall are weaker for less massive systems and at high redshifts, where haloes can be nearer to conventional virial equilibrium.

Baryonic physics can more significantly alter scaling relations (Maughan 2014; Barnes et al. 2017; Farahi et al. 2018; Truong et al. 2018). Radiative cooling, when dense gas cools to produce stars, causes a relatively stronger effect in low-mass systems. Active galactic nucleus (AGN) feedback is an inside-out process affecting primarily regions at small radii. This activity has more impact on the lowest mass systems whereas the binding energy of

★ E-mail: mauro.sereno@inaf.it

massive systems is so large that only the inner core is affected, thus leaving the integrated properties within the virial radius essentially unaltered. AGN feedback can balance radiative cooling and prevent the overcooling and the consequent removal of gas from the hot phase. These processes depend on mass and modify the slope of the scaling relations, making the $M_{\rm gas} - M$ and the $L_{\rm X} - M$ steeper and the $T_{\rm X} - M$ relation shallower.

Non-thermal pressure and incomplete thermalization are more significant at high redshift. Clusters increasingly violate the assumption of HE and require a lower temperature at a given mass to balance gravitational collapse, which leads to a lower normalization. On the other hand, merger induced shocks can heat the gas with a temperature increase which is larger than the total mass variation.

Different processes cause intrinsic scatter around the mean relations. The luminosity is sensitive to the entire merger history of the clusters (Mantz et al. 2016a). The most significant deviations from the $L_{\rm X} - M$ relation originate from recent massive mergers (Torri et al. 2004). During minor mergers, the gas content of smaller and colder substructures is efficiently stripped and mixed due to stellar and AGN feedback (Truong et al. 2018). Radiative phenomena can also perturb the cluster luminosity at fixed mass. In fact, the scatter in luminosity and temperature due to processes in the intra cluster medium (ICM) increases by 20–40 percent and 15–20 per cent, respectively, when the core is considered (Truong et al. 2018).

Additional scatter in X-ray luminosity at fixed mass is caused by radiative phenomena such as stellar or AGN feedback which diversify the cluster luminosity. The presence or absence of compact, bright cores found at the cluster center strongly affect the luminosity and, to a smaller extent, the temperature, but have a smaller role for scatter in $M_{\rm gas}$, which is mostly sensitive to larger spatial scales.

Cluster properties form and evolve due to the same physical processes and some correlation between intrinsic scatters is expected. Numerical simulations show that X-ray quantities are positively correlated at any redshift under a large range of physical assumptions (Stanek et al. 2010; Truong et al. 2018). Correlation is due to baryonic processes and to the merger and accretion history.

It is challenging to measure scaling relations of galaxy clusters. Mass estimates through proxies require complete calibration samples but most of the cluster samples at our disposal are incomplete, heterogeneous, or small (Sereno 2015). Scaling relations can be used to forecast the properties of the not-observed faint end of the halo mass function (Sereno & Ettori 2017), which may require extrapolation. Observed samples are usually affected by selection effects, Malmquist/Eddington biases, or large measurement uncertainties, which require a careful statistical treatment (Kelly 2007; Maughan 2014; Sereno & Ettori 2015a; Mantz 2016).

In this paper, we study the scaling relations between the X-ray properties of the ICM and the mass down to small groups. We take advantage of multi-wavelength surveys, which uniformly scan large areas of sky. The XXL Survey, one of the largest *XMM-Newton* programmes to date (Pierre et al. 2016, hereafter XXL Paper I), covers an area of ∼50 square degrees with an average effective completeness limit of $F_{\rm X,comp} \sim 1.3 \times 10^{-14}$ erg s^{-1} cm^{-2} in the observer-frame [0.5–2.0] keV band within a $1'$ radius aperture for extended sources (Pacaud et al. 2016, hereafter XXL Paper II). The survey has already uncovered nearly four hundreds galaxy clusters and groups out to redshift ∼2 (Adami et al. 2018,hereafter XXL Paper XX) over a wide range of nearly two decades in mass (Lieu et al. 2016, hereafter XXL Paper IV).

Hyper Suprime-Cam is an optical wide-field imager with a field-of-view of 1.77 deg^2 mounted on the prime focus of the 8.2 m Subaru telescope (Miyazaki et al. 2018; Komiyama et al. 2018; Furusawa et al. 2018; Kawanomoto et al. 2018). The Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP, Miyazaki et al. 2018; Aihara et al. 2018a, b) has been carrying out a multi-band imaging survey in five optical bands (*grizy*) with unprecedented depth ($i \sim 26$ at the 5σ limit within a $2''$ diameter aperture), aiming at observing ∼ 1400 deg^2 on the sky in its Wide layer (Aihara et al. 2018b). The survey is optimized for weak lensing (WL) studies (Mandelbaum et al. 2018; Hikage et al. 2019; Miyatake et al. 2019; Hamana et al. 2019) and overlaps with XXL in the XXL-North field.

WL masses are regarded as the most accurate mass estimates for galaxy clusters (Applegate et al. 2014; Umetsu et al. 2014; Okabe & Smith 2016; Melchior et al. 2017; Sereno et al. 2017). They are in principle independent of the equilibrium state of the cluster but can still be affected by their own systematics (Meneghetti et al. 2010; Becker & Kravtsov 2011; Rasia et al. 2012; Svensmark, Wojtak & Hansen 2015).

This is the second paper in a series exploiting the combined HSC-SSP and XXL surveys. In our companion paper (Umetsu et al. 2019), we present a systematic WL analysis of the XXL cluster sample using HSC data. Here, we study the relations between WL mass ($M_{\rm WL}$), $M_{\rm gas}$, $T_{\rm X}$, $L_{\rm X}$, and X-ray masses based on the hydrostatic equilibrium (HE) hypothesis ($M_{\rm HE}$), for the X-ray selected clusters. Our joint multi-variate analysis uses the WL mass measurements obtained by Umetsu et al. (2019).

Bayesian hierarchical models have been efficiently developed to derive scaling relations (D'Agostini 2005; Kelly 2007; Andreon & Bergé 2012; Maughan 2014; Mantz 2016; Sereno 2016). Here, we rely on the CoMaLit (COmparing MAsses in LITerature) approach to scaling relations, wherein we have been applying Bayesian hierarchical methods with latent variables to deal with heteroscedastic and possibly correlated measurement errors and intrinsic scatter, upper and lower limits, missing data, time evolution, and selection effects. For a detailed description, we refer to Sereno & Ettori (2015b, hereafter CoMaLit-I), Sereno, Ettori & Moscardini (2015, hereafter CoMaLit-II), Sereno (2015, hereafter CoMaLit-III), Sereno & Ettori (2015a, hereafter CoMaLit-IV), and Sereno & Ettori (2017, hereafter CoMaLit-V). The method was extended to multi-dimensional analyses in Sereno et al. (2019, hereafter XXL Paper XXXVIII).

The paper is as follows. In Section 2, we detail the statistical scheme used for regression. In Section 3, we introduce the sample and the data-set. The selection function is discussed in Section 4. Results are presented in Section 5. Section 6 is devoted to some final considerations. In Appendix A, we validate our method with mock data. Appendix B details how the probability distribution of latent variables is affected by observational thresholds. In Appendix C, we discuss systematic errors due to pre-determined scaling relations to measure the overdensity radius. Priors on the scatter covariance matrix are discussed in Appendix D.

1.1 Notation and conventions

As reference cosmological model, we assume a flat ΛCDM (Λ and Cold Dark Matter) universe with density parameter $\Omega_{\rm M} = 0.28$, and Hubble constant $H_0 = 70$ km s^{-1} Mpc^{-1}, as found from the study of the nine-year cosmic microwave background (CMB) observations of the Wilkinson Microwave Anisotropy Probe satellite (WMAP9), combined with baryon acoustic oscillation measurements and constraints on H_0 from Cepheids and type Ia supernovae (Hinshaw et al. 2013).

Throughout the paper, O_Δ denotes a global property of the cluster measured within the radius r_Δ which encloses a mean over-

density of Δ times the critical density at the cluster redshift, $\rho_{\rm cr} = 3H(z)^2/(8\pi G)$, where $H(z)$ is the redshift dependent Hubble parameter and G is the gravitational constant. We also define $E_z \equiv H(z)/H_0$.

The notation 'log ' represents the logarithm to base 10 and 'ln' is the natural logarithm. Scatters in natural logarithm can be quoted as per cents. Throughout the paper, unless otherwise noted, we denote σ as the intrinsic scatter in log (decimal) quantities and use δ to represent log (decimal) measurement uncertainty.

Unless stated otherwise, central values and dispersions of the parameter distributions are computed using the bi-weighted statistics (Beers, Flynn & Gebhardt 1990) of the marginalized posterior distributions.

2 REGRESSION

In this section, we describe the statistical method employed to fit the scaling relations. The regression scheme for two measurable cluster properties was detailed in the CoMaLit series (CoMaLit-I, CoMaLit-II, CoMaLit-IV) and in Sereno (2016). This scheme allows for the consistent treatment of time-evolution, correlated intrinsic scatters, and selection effects (Malmquist/Eddington biases). The method was extended to multi-observables with dimension $D \geq 2$ in XXL Paper XXXVIII. Here, we summarize the main features.

2.1 Scheme

We assume that the cluster properties scale as power laws of the cluster mass,

$$O_\Delta = 10^\alpha M_\Delta^\beta E_z^\gamma. \tag{4}$$

Hereafter, we focus on the logarithms of the quantities, which are thus linearly related. In a nutshell, we take the mass as the basic cluster feature (denoted by Z as the reasoning would apply to choices other than the mass as well). For any property, e.g. the temperature or the WL mass, we distinguish three variables: (i) y, the result of the real measurement process; (ii) Y, the quantity that would be measured in a Gedanken experiment with infinite accuracy and precision (Feigelson & Babu 2012); (iii) Y_Z, the quantity that is exactly linked to Z through a functional relation $Y_Z(Z)$ (Maughan 2014).

The measured y is manifest whereas Y, Y_Z, and Z are latent. As defined, Y is intrinsically scattered with respect to Y_Z and does not lie on the ideal linear relation with Z. The measured y differs from Y because of the observational uncertainty. The variable Y_Z is a rescaled version of the underlying Z. An analogous treatment of multiple response variables can be found in Maughan (2014) and Mantz (2016).

This scheme can be generalized to an arbitrary number of clusters and cluster properties. Here, the index i runs through the D cluster properties; the index n runs through the N clusters in the sample. Then, y_{in} is the n-th measurement of the i-th observable, Y_{in} is the true value, and $Y_{Z,\,in}$ is the latent unscattered quantity. If the latent variables $Y_{Z,\,i}$ are linearly related to Z, they are linearly related to each other.

2.2 Distributions

As a result of the observations of the n-th cluster, the $\{y_{in}\}_{i\,=\,1,\,..,\,D}$ and the related uncertainty covariance matrix $\mathbf{V}_{\delta,n}$ are known.[1] On the other hand, $\{Y_{Z,\,in}\}$, $\{Y_{in}\}$, and the covariance matrix of the intrinsic scatter $\mathbf{V}_{\sigma,n}$ are unknown variables to be determined under the assumption of linearity.

The scaling relation of the i-th property is expressed as

$$Y_{Z,i} = \alpha_{Y_i|Z} + \beta_{Y_i|Z} Z + \gamma_{Y_i|Z} \log F_z, \tag{5}$$

where α denotes the normalization, the slope β accounts for the dependence on Z, and the slope γ accounts for the redshift evolution. F_z is the renormalized Hubble parameter, $F_z = E_z/E_z(z_{\rm ref})$.

The measured y and the latent values Y are related as

$$P(y_{1,n},\, y_{2,n},\, ...|Y_{1,n},\, Y_{2,n},\, ...) \propto \\ \times \mathcal{N}^{\rm nD}\left(\{Y_{1,n},\, Y_{2,n},\, ...\},\, \mathbf{V}_{\delta,n}\right) \times \prod_i \mathcal{H}(y_{{\rm th},in}), \tag{6}$$

where $\mathcal{N}^{\rm D}$ is the multivariate Gaussian distribution, $\mathcal{H}$ is the Heaviside function, $\mathbf{V}_{\delta,n}$ is the $D \times D$ uncertainty covariance matrix of the n-th cluster whose diagonal elements are denoted as $\delta^2_{y,in}$, and whose off-diagonal elements are denoted as $\rho_{yin,\,yjn}\delta_{y,\,in}\delta_{y,\,jn}$.

The probability distribution is truncated for $y_{in} < y_{{\rm th},\,in}$, which accounts for selection effects where only clusters above the observational thresholds (in the response variables) are included in the sample. This corrects for the Malmquist bias (CoMaLit-II).

The threshold $y_{\rm th}$ may not be exactly known, as when the quantity which the selection procedure is based on differs from the quantity used in the regression. This can be accounted for with the additional relation

$$P(y_{{\rm th},in}|y_{{\rm th,obs},in}) = \mathcal{N}\left(y_{{\rm th,obs},in},\, \delta^2_{y_{{\rm th},in}}\right), \tag{7}$$

where $y_{{\rm th,\,obs},in}$ is the estimated observational threshold and $\delta_{y_{\rm th,in}}$ is the related uncertainty.

The intrinsic scatters shape the distribution of the true quantities $\{Y_{in}\}$ around the model predictions $\{Y_{Z,\,in}\}$. We assume that the scatters are Gaussian, as well supported by numerical simulations (Stanek et al. 2010; Fabjan et al. 2011; Angulo et al. 2012) and observational studies of core-excised (Maughan 2007) or core-included luminosities (Vikhlinin et al. 2009). For the n-th cluster

$$P(Y_{1,n},\, Y_{2,n},\, ...|Y_{Z,1n},\, Y_{Z,2n},\, ...) \propto \\ \times \mathcal{N}^{\rm D}\left(\{Y_{Z,1n},\, Y_{Z,2n},\, ...\},\, \mathbf{V}_{\sigma,n}\right) \times \prod_i \mathcal{H}(Y_{{\rm th},in}), \tag{8}$$

where $\mathbf{V}_{\sigma,n}$ is the $D \times D$ scatter covariance matrix whose diagonal elements are the intrinsic variances, $\sigma^2_{Y_i|Z}$, and whose off-diagonal elements can be expressed in terms of the correlations as $\rho_{Y_iY_j|Z}\sigma_{Y_i|Z}\sigma_{Y_j|Z}$.

The scatter can be mass- or time-dependent, hence the subscript n in the scatter covariance matrix. However, the inference of the scatter evolution requires larger data-sets than currently available to us (Sereno 2016) and we neglect it, $\mathbf{V}_{\sigma,n} = \mathbf{V}_\sigma$. The adopted parameterization can be easily extended to time dependent scatter and correlations (CoMaLit-IV).

Even if the selection procedure is based only on the value of the measured y rather than the value of Y, any threshold in y affects all the conditional probability distributions, see Appendix B. In fact, we do not sample a generic distribution of clusters but we select them and we have to model the distribution of the sampled objects. Hence, the distribution of Y given Z for a generic cluster from the full population differs from the distribution specific to a selected sample, which follows equation (8) and it is truncated too, see Appendix B. The threshold for the n-th measurement of the i-th response variable is denoted as $Y_{{\rm th},\,in}$. This is related to the threshold

[1] The regression scheme can also deal with missing data (CoMaLit-V).

for the measured value as

$$P(Y_{\mathrm{th},in}|y_{\mathrm{th,in}}) = \mathcal{N}\left(y_{\mathrm{th},in}, \delta^2_{y,in}\right), \quad (9)$$

where $\delta_{y,\,in}$ is the uncertainty associated to y_{in}. In the absence of Malmquist biases ($Y_{\mathrm{th},\,in} \to -\infty$), Y_Z is the mean value of $p(Y_1|Z)$, i.e. $\langle Y_1 \rangle = Y_Z$.

The distribution of the reference Z variable is modeled as a Gaussian function or a mixture (Kelly 2007, CoMaLit-II, CoMaLit-IV). We adopt the simplest but still effective case of one component (Sereno 2016),

$$P(Z) = \mathcal{N}\left(\mu_Z, \sigma^2_Z\right). \quad (10)$$

Most of the parent populations of astronomical quantities can be well approximated with this scheme, see Appendix A. Here, we model only the shape of the distribution and we do not fit the halo abundance and the observed number count of clusters, see e.g. Murata et al. (2019) and Giles et al. (2016, hereafter XXL Paper III).

The evolution of the (mean of the) Z-distribution can be modeled as (CoMaLit-IV),

$$\mu_Z(z) = \mu_{Z,0} + \gamma_{\mu_Z,D} \log D_z, \quad (11)$$

where $\mu_{Z,\,0}$ is the mean at the reference redshift z_{ref} and D_z is the luminosity distance. We renormalize the distances such that D_z is equal to one at the reference redshift z_{ref}.

The dispersion of the Z-distribution evolves as

$$\sigma_Z(z) = \sigma_{Z,0} D_z^{\gamma_{\sigma_Z}}. \quad (12)$$

The dependence on D_z accounts for the redshift evolution. This is justified by theoretical predictions based on the self-similar model, by results of numerical simulations, and by observational fits (CoMaLit-IV). The explicit dependence on the cosmological distance for the evolution of the covariate distribution, see equation (11), accounts for the completeness of a sample selected according to either flux or signal-to-noise (CoMaLit-IV). The redshift dependence in equation (11) is general enough to address even more complicated cases. More general parameterizations for time-evolution can be found in CoMaLit-IV.

2.3 Priors

The final piece of the statistical treatment is the explicit declaration of the priors, which have to be conveniently non-informative (CoMaLit-I). If we do not know the value of Z, some slopes and normalizations in equation (5) may be redundant (XXL Paper XXXVIII). If Z is the mass, we can break the degeneracy thanks to the WL mass which is a reliable, nearly unbiased proxy of the true mass.

In our analysis, the statistical uncertainty on the WL masses accounts for shape noise, cosmic noise due to uncorrelated large-scale structures, intrinsic variations of the projected cluster lensing signal at fixed mass due to variations in halo concentration, cluster asphericity, and the presence of correlated haloes (Umetsu et al. 2019). Unaccounted sources of errors in cluster mass calibration can be due to dilution of the WL signal by residual contamination from foreground and cluster members, bias in the source photometric redshifts, and systematic uncertainty in the mass modeling, which sum up to a total systematic uncertainty of ∼5 per cent in ensemble mass calibration of the XXL sample (Umetsu et al. 2019). As WL priors, we can then consider

$$\alpha_{Y_{MWL}|Z} = \mathcal{N}\left(0, \delta^2_{\mathrm{sys},m_{\mathrm{WL}}}\right) \quad (13)$$

$$\beta_{Y_{MWL}|Z} = 1\,, \quad (14)$$

$$\gamma_{Y_{MWL}|Z} = 0\,, \quad (15)$$

where $\delta_{\mathrm{sys},m_{\mathrm{WL}}} = 0.05/\ln(10)$. Fixing the parameters of the M_{WL}-M rather than the T-M relation is just a matter of rescaling which does not affect the analysis of the intrinsic scatters. Any residual bias suffered by the WL mass (i.e. $\langle \alpha_{Y_{MWL}|Z} \rangle \neq 0$) affects the estimated overall normalization of the scaling relations. The data analysis can only constrain the relative bias between the mass proxies Y_i (CoMaLit-I).

For the other X-ray observables, the priors on the intercepts $\alpha_{Y|Z}$ and on the mean $\mu_{Z,\,0}$ are flat,

$$\alpha_{Y_i|Z},\ \mu_{Z,0} \sim \mathcal{U}(-1/\epsilon, 1/\epsilon), \quad (16)$$

where ϵ is a small number. In our calculations, $\epsilon = 10^{-4}$.

For the slopes and the time-evolutions of the relations including observables other than the WL mass, as well as for the Z covariate distribution, we consider uniformly distributed direction angles, $\arctan\beta$ and $\arctan\gamma$, i.e. we model the prior probabilities as a Student's t_1 distribution with one degree of freedom,

$$\beta_{Y|Z},\ \gamma_{Y|Z},\ \gamma_{\mu_Z,D},\ \gamma_{\sigma_Z} \sim t_1. \quad (17)$$

A non-informative prior on the variance has to show a very long tail to large values. This can be obtained with the nearly scale-invariant Gamma distribution for the precision, i.e. the inverse of the variance,

$$1/\sigma^2_{Z,0} \sim \Gamma(r, \lambda), \quad (18)$$

where the rate r and the shape parameter λ are fixed as $r = \lambda = \epsilon$.

For more than two observed properties, $D \geq 2$, we express the prior on the (inverse of the) intrinsic scatter matrix in terms of the scaled Wishart distribution (Huang & Wand 2013),

$$\mathbf{V}_\sigma^{-1} \sim \mathbf{I_W}(\mathbf{S}, f), \quad (19)$$

where f is the number of degrees of freedom and $\mathbf{S}$ is the $D \times D$ scale matrix. The scaled Wishart distribution implicitly defines a prior on the variance-covariance: the standard deviation of the element i of the multivariate normal, $\sigma_i = \Sigma_{ii}^{1/2}$, has a half-t distribution with scale $\mathbf{S}_i$ and f degrees of freedom. The density is flat for $\sigma_i \ll \mathbf{S}_i$ and has a long tail at large values. In our computations, we consider the identity matrix as $\mathbf{S}$. We take $f = 2$, so that all correlation parameters have a marginal uniform prior distribution between -1 and 1. The prior in equation (19) differs from XXL Paper XXXVIII (see their equations 26 and 27). Final results are stable with respect to the choice of the priors, as far as they are non informative, see Appendix D.

For just two observed properties, $D = 2$, correlations cannot be constrained for small data-sets and we consider $\rho_{Y_1Y2} = 0$ and a Gamma distribution for each intrinsic scatter as in equation (18).

3 SAMPLE

The XXL-365-GC catalogue from the second XXL data release (DR2) is described in XXL Paper XX. The XXL selection function depends on the size, shape, and count rate of the emitting source, and on the local *XMM-Newton* sensitivity (XXL Paper II). A validation of the candidates by human confirmation is applied too. The C1 population is designed to be free of contamination by spurious detections or blended point sources, while the C2 population is more complete but less pure, with an initial selection of ∼ 50 per cent of spurious sources (XXL Paper I). A third C3 class contains known

heterogeneous clusters not detected by the automatic pipeline. Confirmed XXL clusters are cleaned up a posteriori by optical spectroscopic observations or detailed comparison of X-ray and optical observations.

We consider the subsample of 302 confirmed clusters of class C1 and C2. The exclusion of C3 clusters improves the statistical completeness of the sample. Of the sample under consideration, 265 clusters have measured gas mass, 235 clusters have spectroscopically derived luminosity and temperature, 227 clusters have an estimated mass based on the HE assumption, and 136 clusters in the North have a measured WL mass. A subsample of 97 clusters has all four properties measured.

We consider the temperature within 300 kpc, $T_{300\mathrm{kpc}}$, and the rest frame [0.5–2] keV luminosity $L^{\mathrm{XXL}}_{500,\mathrm{MT}}$ and gas mass $M_{\mathrm{gas},\,500}$ within r_{500}. In the following, we summarize the measurement process for luminosity, temperature, and gas mass, which are detailed in XXL Paper XX and references therein, and the HE masses. The WL analysis is detailed in a companion paper (Umetsu et al. 2019).

3.1 Luminosity and temperature

Luminosity and temperature are measured with a spectral analysis of the cluster single best pointing (XXL Paper XX). Spectra are extracted for each of the *XMM-Newton* cameras from the region within 300 kpc and fitted in the [0.4 − 11.0] keV band with the absorbed APEC (Astrophysical Plasma Emission Code) model (v2.0.2) in XSPEC (Dorman, Arnaud & Gordon 2003), with a fixed metal abundance of $Z = 0.3Z_{\odot}$. The background was modelled following Eckert, Molendi & Paltani (2011).

Luminosities $L^{\mathrm{XXL}}_{500,\mathrm{MT}}$ within $r_{500,\,\mathrm{MT}}$ in the [0.5–2.0] keV band (cluster rest frame), where $r_{500,\,\mathrm{MT}}$ was calculated using the mass-temperature relation of XXL Paper IV, are extrapolated from 300 kpc assuming a β-profile with a core radius $r_{\mathrm{c}} = 0.15r_{500,\,\mathrm{MT}}$ and a slope $\beta = 2/3$.

X-ray temperatures could not be measured for all clusters. Several cluster observations were affected by flaring, had very low counts, were contaminated by point sources, or were at very low redshift with bad spatial coverage.

3.2 Gas mass

Gas masses for clusters with known redshift are computed following the method outlined in Eckert et al. (2016, hereafter XXL Paper XIII) and XXL Paper XX. Surface-brightness profiles are extracted in the [0.5–2] keV band starting from the X-ray peak from mosaic images of the XXL fields instead of individual pointings. The surface-brightness profiles are decomposed on a basis of multiscale parametric forms, deprojected, and then converted into gas density profiles using X-ray cooling functions calculated using the APEC plasma emission code.

The gas mass within an overdensity r_{500} is computed with an iterative procedure to compute r_{500} and the temperature from the surface brightness profile exploiting the $f_{\mathrm{gas}} - M$ relation from XXL Paper XIII.

3.3 Hydrostatic mass

For an ideal fluid where thermal conductivity and viscosity have no significant roles, under the assumption that it has a spherically-symmetric geometry and no internal motions, we can use the HE equation of the ICM in a gravitational potential to recover the radial profile of the total mass (Ettori et al. 2013):

$$M_{\mathrm{tot}}(<r) = -\frac{r^2}{G\rho_{\mathrm{gas}}}\frac{dP_{\mathrm{gas}}}{dr} \tag{20}$$

$$= -\frac{k_{\mathrm{B}}T_{\mathrm{gas}}(r)\,r}{\mu m_{\mathrm{u}}G}\left(\frac{\mathrm{d}\log T_{\mathrm{gas}}}{\mathrm{d}\log r} + \frac{\mathrm{d}\log n_{\mathrm{gas}}}{\mathrm{d}\log r}\right). \tag{21}$$

where gas mass density and pressure are related through the perfect gas law,

$$P_{\mathrm{gas}} = \rho_{\mathrm{gas}}\,k_{\mathrm{B}}T_{\mathrm{gas}}/(\mu m_{\mathrm{u}}) = n_{\mathrm{gas}}\,k_{\mathrm{B}}T_{\mathrm{gas}}, \tag{22}$$

where k_{B} is the Boltzmann constant, $m_{\mathrm{u}} = 1.66 \times 10^{-24}$ g is the atomic mass unit, and $\mu \sim 0.6$ is the mean molecular weight in a.m.u..

Gas density profiles are measured as described above, see Section 3.2.

More critical is the gas temperature profile, considering that we have estimated only a value from the counts collected within 300 kpc. Instead of adopting the assumption of an isothermal gas, which can significantly bias the recovered mass profile (up to 30–40 per cent; see e.g. Rasia et al. 2006), we use a 'universal' pressure profile, appropriately rescaled for the object mass M_{500}, to specify the radial profile of the temperature:

$$P(r) = \left(\frac{M_{500}}{10^{15}M_{\odot}}\right)^{2/3}\frac{a}{[(bx)^c(1+(bx)^d]^{\frac{e-c}{d}}}, \tag{23}$$

where $x = r/r_{500}$, $(a, b, c, d, e) = (5.68, 1.49, 0.43, 1.33, 4.40)$ (see table 3 in Ghirardini et al. 2019);

The 3D temperature profile is then recovered through equation (22). For each object, $T(r)$ is rescaled by the factor $T_{300\mathrm{kpc}}/\overline{T}(<300\,\mathrm{kpc})$, where $\overline{T}$ is the emission-weighted temperature. Because of the dependence of these profiles on the radius and mass estimated at the overdensity $\Delta = 500$, we iterate the procedure till convergence at a level < 5 per cent is obtained on r_{500}.

The errors on the spectroscopic measurement of $T_{300\mathrm{kpc}}$ are propagated to the temperature profile according to the signal-to-noise ratio estimated from the emission measure recovered from the gas density profile. These errors are used for a Monte Carlo sampling of the hydrostatic mass profile. The 16th and 84thpercentile of 100 Monte Carlo realizations define the uncertainty associated with the mass estimates at each radius where the gas density has been recovered.

3.4 Covariance uncertainty matrix

The knowledge of the covariance uncertainty matrix is crucial to obtain unbiased estimates of intrinsic scatters and correlations. Measurements of luminosity and temperature are based on the analysis of a single spectrum of the core region within 300 kpc. The luminosity estimate mostly depends on the normalization of the spectrum, whereas the temperature is strictly related to the shape. This makes their measurements rather independent, but being the result of a single measurement process some correlation persists.

On the other hand, the gas mass measurement process exploits the photometry and the surface brightness profile in annular regions up to larger radii (on average, nearly double the limiting radius used for the spectra) and is largely independent of the spectral output. Furthermore, the gas mass measurement exploits mosaic images, whereas the spectra are taken from individual pointings.

The HE mass measurements are based on temperature profiles normalized to $T_{300\mathrm{kpc}}$ and on the shape of the gas density profiles. This makes HE masses strongly correlated with temperatures and gas masses and we prefer to exclude them from the multi-variate analysis.

The aperture radius, i.e. the radius within which properties are measured, is estimated independently for each X-ray observable. This is done on purpose to minimize the correlation between measurements. X-ray properties measured within the same WL determined r_{500} would be strongly correlated with the mass. Our procedure is standard in large surveys where WL masses are usually available only for small subsamples and independent methods are needed to approximate the virial radius. The downside is that the intrinsic scatter of quantities determined in this way is increased.

Finally, the WL measurement is independent of the X-ray observations (apart from the cluster coordinates and redshift).

We then consider the luminosity and the temperature as the only correlated measurements. To estimate the uncertainty correlation we proceed in the following way. Luminosity and temperature within 300 kpc are estimated in a single measurement process. Their correlation is an output of the spectroscopic fit. We approximate the probability distribution of the observed luminosity and temperature as a bivariate Gaussian.

The luminosity is then extrapolated by assuming a distribution of radial β-profiles. This is approximated as a bivariate Gaussian $\mathcal{N}(r_c/r_{500}, \beta)$ with mean core radius $r_c = 0.15\, r_{500}$ and mean $\beta = 2/3$ (XXL Paper III), and scatter in slope of $\sigma_\beta \sim 0.1$, scatter in core radius of $\sigma(r_c/r_{500}) \sim 0.1$, and correlation $\rho_{\beta r_c} \sim 0.66$ as representative of the sample of 45 bright nearby galaxy clusters in Mohr, Mathiesen & Evrard (1999). The outer radius r_{500} is estimated with the $M - T_X$ from XXL Paper IV.

For each cluster, we extract 10^4 couples of correlated luminosity and temperature within 300 kpc from the approximated bivariate normal distribution. Based on the $M - T_X$ relation, we then derive the related values of r_{500}. The gas profiles used to extrapolate the luminosity are randomly sampled by extracting correlated values r_c and β from the distribution of radial profiles. The final uncertainty correlation matrix is computed from 10^4 sampled pairs of temperature, $T_{300\mathrm{kpc}}$, and luminosity, $L^{\mathrm{XXL}}_{500,\mathrm{MT}}$.

4 SELECTION FUNCTION

The full knowledge of the selection function is crucial in cosmological studies of number counts and abundance evolution, when observed properties have to be related to the underlying mass function. In XXL Paper XXV, the selection function was expressed in terms of the true cluster parameters rather than in terms of their measured counterparts affected by measurement errors and intrinsic scatters.

Here, we are interested in scaling relations and our primary need is a safe treatment of any selection bias. We can model the population of observed clusters rather than starting from the halo mass function, see Appendix A, and we just need to account for clusters to be included or excluded from the sample based on observed properties, see Appendix B.

The size and shape dependence of the XXL selection function is mostly meant to distinguish between point sources and extended emission from clusters. Here, we already deal with a nominally pure, spectroscopically confirmed sample. The main remaining dependence is on the X-ray flux. In fact, the isophotes of the XXL completeness function and the sky coverage in the source parameter space, i.e. cluster core radius versus total XMM count rate, follow at first order the curves of equal rest-frame flux (XXL Paper II).

Based on the above considerations, we approximate the selection function in terms of a complementary error function for the luminosity $L^{\mathrm{XXL}}_{500,\mathrm{MT}}$, see Appendix B. The luminosity threshold as a function of redshift is computed as the lower smoothed envelope in the $L^{\mathrm{XXL}}_{500,\mathrm{MT}} - z$ plane, see Fig. 1.

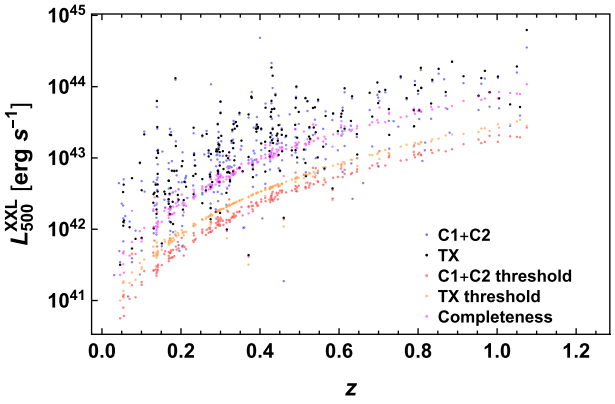

Figure 1. Luminosities of the XXL clusters as a function of redshift. Black and blue points are the luminosities derived by the spectral analysis or the scaling relations, respectively. Red and orange points mark the luminosity thresholds for the full and the spectroscopic sample with measured temperature, respectively. The magenta points follow the effective completeness limit.

Since spectroscopic luminosity and temperature are measured only for a bright subsample, we consider the luminosities derived from the count rate by adopting a convenient set of scaling relations (XXL Paper XX, see section 4.3), which are available for the full C1 + C2 sample. Derived thresholds are shown in Fig. 1. Thresholds for the bright subsample are derived likewise considering the lower envelope of the clusters with measured temperature.

The scale-length of the approximated selection function accounts for both statistical uncertainties in the flux measurements (weighted by the number of clusters used to measure the envelope) and the other aspects of the XXL selection function not accounted for by the flux.

The flux limit of the DR2 catalogue is $F_{\mathrm{X,cut}} \sim 3.2 \times 10^{-15}$ erg s^{-1} cm^{-2} in the observer-frame [0.5–2.0] keV band within a $1'$ radius aperture. The effective completeness limit averaged across the entire survey area is $F_{\mathrm{X,comp}} \sim 1.3 \times 10^{-14}$ erg s^{-1} cm^{-2}. These limits can be converted to a standard deviation assuming that they delimit the 5σ detection range.

5 RESULTS

In this section, we describe the results of our regression procedure on the HSC-XXL clusters. We work in the log space, where scaling relations are expected to be linear. The normalized WL mass, gas mass, temperature, and luminosity in logarithmic units are written as

$$m_{\mathrm{wl}} = \log(M_{\mathrm{WL},500}/10^{14}/\mathrm{M}_\odot), \tag{24}$$

$$m_{\mathrm{he}} = \log(M_{\mathrm{HE},500}/10^{14}/\mathrm{M}_\odot), \tag{25}$$

$$m_{\mathrm{g}} = \log(M_{\mathrm{gas},500}/10^{14}/\mathrm{M}_\odot), \tag{26}$$

$$t = \log(T_{300\mathrm{kpc}}/\mathrm{keV}), \tag{27}$$

$$l = \log(L^{\mathrm{XXL}}_{500,\mathrm{MT}}/10^{43}/\mathrm{erg/s}^{-1}), \tag{28}$$

respectively.

We either fit each X-ray property versus the WL mass ($D = 2$), see Table 1, or we perform a multi-scaling analysis of the subsample with complete information ($D = 4$), see Table 2. The $M - T_X$ relation

Table 1. Observed scaling relations for the HSC-XXL sample as derived from the bivariate ($D = 2$) analysis. Conventions and units are as in Section 2. The weak lensing mass is the variable $Y_1 = X$. Col. 1: variable $Y_2 = Y$ of the regression procedure. Col. 2: number of fitted clusters. Cols. 3, 4, and 5: intercept, slope, and time evolution of the scaling relation. Cols. 6, and 7: scatter of Y and X, respectively. Values in square brackets correspond to parameters kept fixed in the regression.

		Intercept	Slope	Time-evolution	Intrinsic scatters	
Y	n	$\alpha_{Y\|Z}$	$\beta_{Y\|Z}$	$\gamma_{Y\|Z}$	$\sigma_{Y\|Z,0}$	$\sigma_{X\|Z,0}$
l	105	-0.10 ± 0.19	1.06 ± 0.35	2.10 ± 2.08	0.55 ± 0.13	0.07 ± 0.08
l	105	-0.09 ± 0.19	1.07 ± 0.37	[2]	0.55 ± 0.11	0.08 ± 0.09
l	105	-0.09 ± 0.15	[1]	[2]	0.54 ± 0.09	0.05 ± 0.06
t	105	0.44 ± 0.09	0.85 ± 0.31	0.18 ± 0.66	0.06 ± 0.05	0.31 ± 0.08
t	105	0.42 ± 0.07	0.75 ± 0.27	[2/3]	0.07 ± 0.05	0.29 ± 0.09
t	105	0.41 ± 0.05	[2/3]	[2/3]	0.07 ± 0.04	0.29 ± 0.06
$m_{\rm g}$	118	-1.08 ± 0.11	1.35 ± 0.36	1.76 ± 1.22	0.11 ± 0.10	0.25 ± 0.09
$m_{\rm g}$	118	-1.01 ± 0.12	1.55 ± 0.30	[0]	0.06 ± 0.06	0.28 ± 0.07
$m_{\rm g}$	118	-1.11 ± 0.07	[1]	[0]	0.24 ± 0.06	0.19 ± 0.11
$m_{\rm HE}$	100	-0.04 ± 0.08	[1]	[0]	0.31 ± 0.05	0.37 ± 0.06

Table 2. Observed scaling relations from the multi-variate analysis ($D = 4$) of 97 HSC-XXL groups. Conventions and units are as in Section 2. Results for the weak lensing mass reflect the priors and are not showed. Col. 1: variable. Cols. 2, 3, and 4: intercept, slope, and time evolution of the scaling relation.

	Intercept	Slope	Time-evolution
Y	$\alpha_{Y\|Z}$	$\beta_{Y\|Z}$	$\gamma_{Y\|Z}$
$m_{\rm g}$	-1.10 ± 0.19	1.73 ± 0.80	0.16 ± 1.13
t	0.42 ± 0.11	0.78 ± 0.43	0.11 ± 0.57
l	0.28 ± 0.23	1.91 ± 0.94	3.12 ± 1.35

Table 3. Properties of the covariance matrix of the intrinsic scatters from the multi-scaling analysis ($D = 4$) of the HSC-XXL clusters for the case of free time evolution. Diagonal (bold face): posterior bi-weight estimators of the intrinsic scatter of each property at fixed mass. Upper triangle: posterior bi-weight estimators of property pair correlation coefficients at fixed mass. Lower triangle: statistical significance (in per cents) of the positiveness of the estimated correlation.

	$m_{\rm wl}$	$m_{\rm g}$	t	l
$m_{\rm wl}$	$\mathbf{0.32 \pm 0.08}$	0.51 ± 0.27	0.15 ± 0.38	0.34 ± 0.39
$m_{\rm g}$	93%	$\mathbf{0.27 \pm 0.12}$	0.20 ± 0.46	0.55 ± 0.42
t	63%	64%	$\mathbf{0.12 \pm 0.04}$	0.35 ± 0.44
l	77%	84%	75%	$\mathbf{0.23 \pm 0.14}$

is discussed in the companion paper (Umetsu et al. 2019). The $L_{\rm X} - T_{\rm X}$ relation is discussed in XXL Paper XX. As reference redshift, we consider $z_{\rm ref} = 0.3$, close to the median redshift of the sample.

Results from the $D = 2$ and the $D = 4$ analyses can be compared only with a caveat. The luminosity– and temperature–mass relations are measured only for the clusters with known temperature, whereas the gas mass–total mass relation also considers less bright clusters. For the multi-variate ($D = 4$) analysis, we consider the subset where all X-ray and WL properties are known.

5.1 Hydrostatic bias

To measure the hydrostatic bias, we assume that the hydrostatic mass $M_{\rm HE}$ is a scattered, biased proxy of the mass (CoMaLit-I),

$$\log M_{\rm HE} = \log[(1 - b_{\rm HE})M] \pm \sigma_{m_{\rm HE}|m}. \quad (29)$$

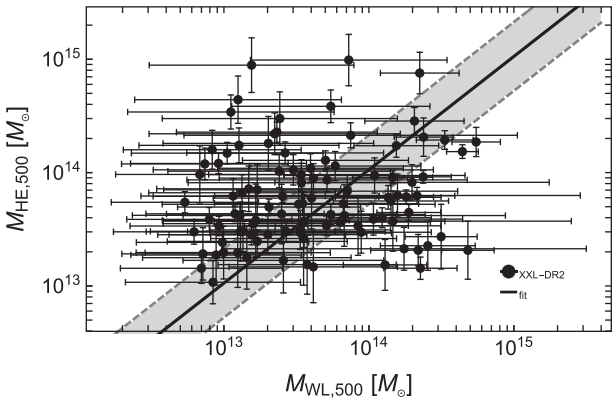

Figure 2. The hydrostatic versus weak lensing masses of the HSC-XXL clusters in the case of fixed slope, $\beta_{m_{\rm he}|m} = 1$. The dashed black lines show the bisector (full black line) plus or minus the intrinsic scatter at the median redshift $z = 0.30$. The shaded grey region encloses the 68.3 per cent confidence region around the median relation due to intrinsic scatter and uncertainties on the not-fixed scaling parameters.

We do not find any strong statistical evidence for significant non thermal pressure, see Fig. 2. The level of inferred bias strongly depends on the calibration sample and on the applied methodology. Observed values range from $\gtrsim 0$ to 40–50 per cent (CoMaLit-I). Eckert et al. (2019) found a median non-thermal pressure fraction of ~ 6 per cent at r_{500} in a sample of 12 nearby Planck selected clusters with high-quality *XMM-Newton* observations out to the virial radius.

Theoretical estimates from numerical simulations are strongly dependent on the adopted scheme. After disentangling bulk from small-scale turbulent motions in high-resolution simulations of galaxy clusters, Vazza et al. (2018) constrained the gas kinetic energy effectively providing pressure support in the cluster gravitational potential and reported a bias of the order of $\sim$10 per cent at R_{200} in low mass clusters. The typical non-thermal pressure support is $\sim$5 per cent in the centre of clusters, and it is $\sim$15 per cent in the outskirts (Angelinelli et al. 2019).

Our result is compatible with a small contribution of non thermal pressure in low mass groups, $b_{\rm HE} = 9 \pm 17$ per cent. We find that $b_{\rm HE} \lesssim 33(44)$ per cent at the 95.45 (99.73) per cent level. Strong

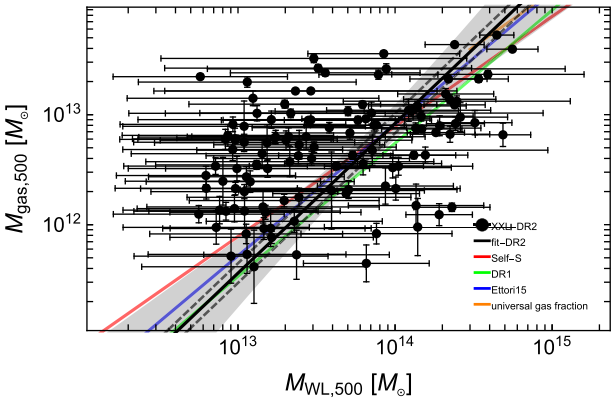

Figure 3. The gas versus total mass relation of the HSC-XXL clusters in the case of free time evolution. The dashed black lines show the median scaling relation (full black line) plus or minus the intrinsic scatter at the median redshift $z = 0.30$. The shaded grey region encloses the 68.3 per cent confidence region around the median relation due to intrinsic scatter and uncertainties on the scaling parameters. The red line is the fit for self-similar parameters. The green and blue lines plot the relations from the XXL Paper XIII, and Ettori (2015), respectively. The orange line follows the expected universal gas fraction (Eckert et al. 2019) and it is plotted only for the mass range where it holds ($M_{500} \gtrsim 2 \times 10^{14}\,{\rm M}_\odot$).

conclusions are hampered by the large statistical uncertainty. In fact, results are consistent with no bias too.

5.2 Gas versus total mass

The $M_{\rm gas} - M$ relation appears to be steeper than the self-similar prediction ($\beta_{m_{\rm g}|m} = 1$), see Figs 3 and 4, due to the role played by radiative cooling and AGN feedback in lesser systems. Conversion of hot gas into stars is more efficient in low mass systems and steepens the $M_{\rm gas} - M$ relation.

For the $D = 4$ analysis, $\beta_{m_{\rm g}|m} = 1.73 \pm 0.80$. We find a steeper than self-similar relation with a probability of ∼84 per cent or 80 per cent for the $D = 2$ and $D = 4$ fit with free time evolution, respectively. For the $D = 2$ fit, we assume that the correlation between gas mass and luminosity scatters is not extreme and we do not model the Malmquist bias for the gas mass ($m_{\rm g,\,th} \to 0$).

There is no conclusive statistical evidence for time-evolution. The probability for positive evolution ($\gamma_{m_{\rm g}|m} > 0$) is ∼95 (56) per cent for the $D = 2$ (4) fit.

Results are stable if we consider no evolution, i.e. the strong prior $\gamma_{m_{\rm g}|m} = 0$. In this case, the probability of a slope steeper than the self-similar value is ∼98 per cent for the $D = 2$ fit. In the following, we quote the more conservative results with free evolution.

The gas fraction is consistent with predictions. Eckert et al. (2019) estimated the expected gas fraction of galaxy clusters by considering the universal baryon fraction from the CMB power spectrum, the baryon depletion factor predicted by numerical simulations, and the stellar fraction from a compilation of recent results. They found $f_{\rm gas,\,500} = 0.131 \pm 0.009$ for $M_{500} \gtrsim 2 \times 10^{14}\,{\rm M}_\odot$, in agreement with our result, $f_{\rm gas,\,500} = 0.11 \pm 0.05$ for $M_{500} = 2 \times 10^{14}\,{\rm M}_\odot$ at $z = 0.3$.

Our relation is in good agreement with previous observational results. The measured slope agrees well with XXL Paper XIII, who found $\beta_{m_{\rm g}|m} = 1.21^{+0.11}_{-0.10}$ based on the first XXL data release. On the other hand, the normalization and the related baryonic fraction that we find are higher, see Fig. 3. In fact, even though the gas masses are consistent (XXL Paper XX), the WL masses here used are smaller (Umetsu et al. 2019).

Ettori (2015) considered a self-similar framework where apparent deviations are due to three physical mass-dependent quantities: the gas clumpiness, the gas mass fraction, and the slope of the thermal pressure. Normalization and mass dependence of the gas mass fraction were constrained with samples with observed gas masses, temperatures, luminosities, and Compton parameters in local clusters. At $z = 0.3$ and for $M_{500} = 3 \times 10^{13}\,{\rm M}_\odot$, we find $f_{\rm gas,\,500} = 0.057 \pm 0.019$ in agreement with Ettori (2015), who found $f_{\rm gas,\,500} \sim 0.065$. Lovisari, Reiprich & Schellenberger (2015) analyzed *XMM-Newton* observations for a complete sample of local ($z < 0.034$), flux-limited galaxy groups selected from the *ROSAT* All-Sky. Exploiting hydrostatic masses to calibrate the relation, they found $f_{\rm gas,\,500} \sim 0.076$ for $M_{500} = 5 \times 10^{13}\,{\rm M}_\odot$, in agreement with our result ($f_{\rm gas,\,500} = 0.053 \pm 0.015$).

There is some evidence that these results are stable with respect to mass and redshift range and selection methods (Chiu et al. 2016, 2018a). Chiu et al. (2018a) analyzed 91 massive galaxy clusters ($M_{500} \gtrsim 1.5 \times 10^{14}\,{\rm M}_\odot$) selected by the South Pole Telescope SPT-SZ survey. Exploiting masses estimated from the SZE (Sunyaev-Zel'dovich effect), they found a slope of ∼1.3 and no significant redshift evolution over a large redshift range ($0.2 < z < 1.25$). This agreement may be related to the small hydrostatic bias, see Section 5.1.

5.3 Luminosity versus mass

AGN activity heats the gas of the smallest systems and cooling can be counterbalanced by intense feedback. Baryonic processes reduce the amount of gas in the smallest systems, and thus their total luminosity. Removal of dense gas in small groups due to efficient radiative cooling steepens the $L_{\rm X} - T_{\rm X}$ and the $L_{\rm X} - M$ relations. In fact, slopes are steeper if the core is included. At very high redshift, the main driver for gas removal is AGN feedback (Truong et al. 2018).

We find evidence for a luminosity–mass relation steeper ($\beta_{l|m} = 1.91 \pm 0.94$) than the self-similar expectation ($\beta_{l|m} = 1$) with a probability of ∼81 per cent for the $D = 4$ fit assuming free time evolution, see Fig. 4. The results are in agreement but less significant for the $D = 2$ fit, see Fig. 5, with a probability of ∼57 (∼57) per cent for the fit with free (fixed) time evolution.

The mass dependence of the halo concentration and dark matter processes can make the theoretical $L_{\rm X} - M$ relation shallower ($\beta_{l|m} < 1$) according to the Fundamental Plane relation of galaxy clusters (Fujita & Aung 2019). Our comparison to the self-similar expectation ($\beta_{l|m} = 1$) is then conservative in highlighting the effects of baryonic processes, which go in the opposite direction.

The larger uncertainties with respect to the gas mass analysis are due to the Malmquist bias and to the smaller sample. There is inconclusive evidence for negative time-evolution, with $\gamma_{l|m} < 2$ at the ∼50 per cent probability level.

As a comparison, we considered the MCXC (Meta-Catalogue of X-ray detected Clusters of galaxies, Piffaretti et al. 2011), which comprises 1743 unique X-ray clusters with measured X-ray luminosities measured in the [0.1–2.4] keV band collected from available *ROSAT* All Sky Survey-based and serendipitous cluster catalogues. For our tests, we fixed the unquoted MCXC statistical uncertainty to 10 per cent.

Masses were retrieved from the Literature Catalogs of weak Lensing Clusters (LC^2, CoMaLit-III), whose release v3.6 comprises 601 unique clusters with reported coordinates, redshift, and WL

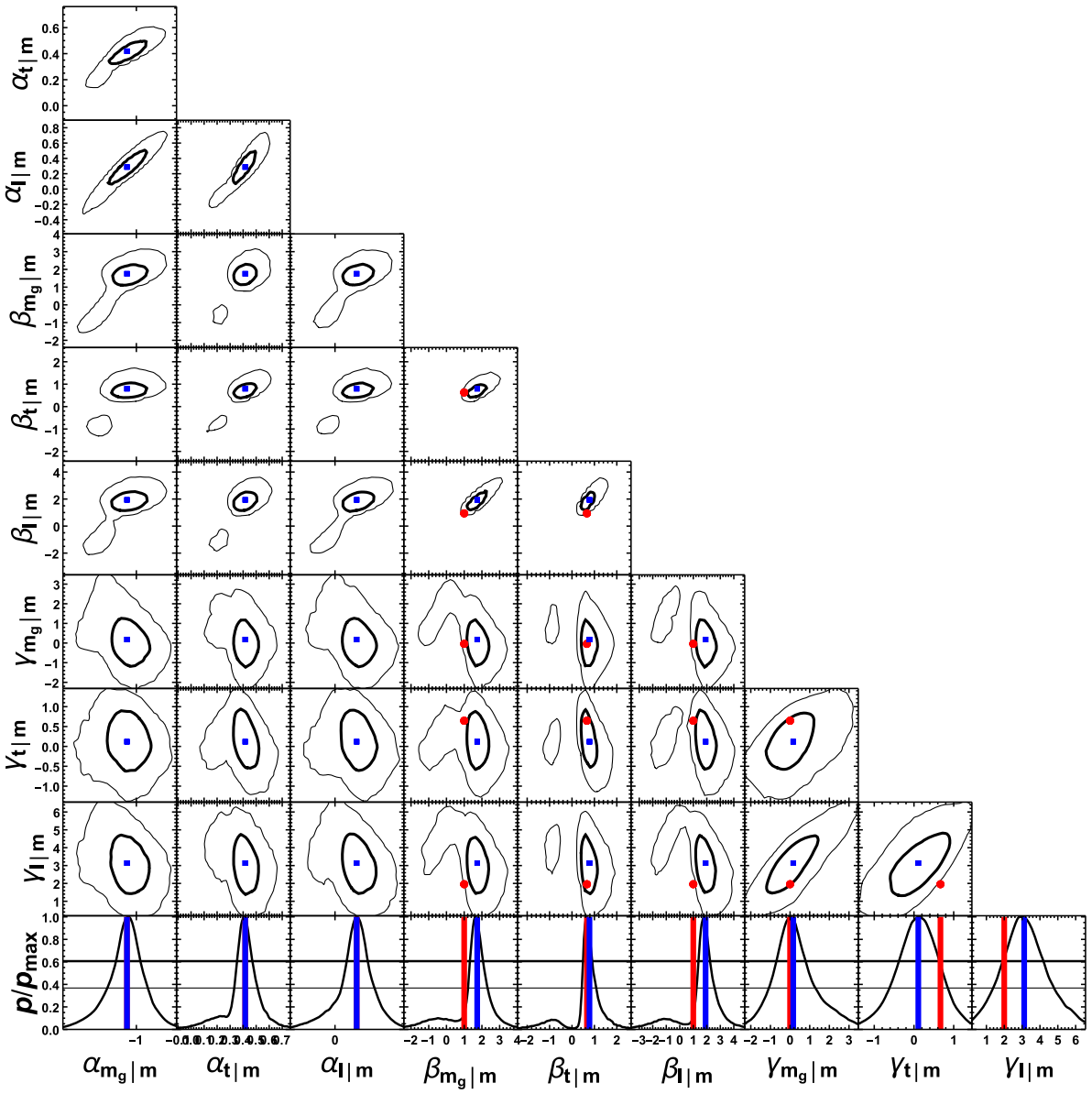

Figure 4. Probability distributions of the parameters of the scaling relations as obtained from the joint multi-variate regression ($D = 4$). The intercepts, slopes, and time-evolutions are denoted as α, β, and γ, respectively. The thick and thin black contours include the 1-σ and 2-σ confidence regions in two dimensions, here defined as the regions within which the probability is larger than $\exp(-2.3/2)$ and $\exp(-6.17/2)$ of the maximum, respectively. The bottom row shows the marginalized 1D distributions, renormalized to the maximum probability. The thick and thin black levels denote the confidence limits in one dimension, i.e. $\exp(-1/2)$ or $\exp(-4/2)$ and of the maximum. The blue symbols mark the biweight estimator. The red symbols mark the self-similar prediction.

masses.[2] MCXC and LC^2-single were cross-matched by coordinates, with a maximum allowed separation of $2'$, and redshift, with a maximum separation of $\Delta z = 0.05$, to find 216 clusters with complete info, see Fig. 6. This heterogeneous sample contains more massive clusters than XXL, with a median mass of $\sim 5 \times 10^{14}\,\mathrm{M}_\odot$, but the inferred scaling relation ($\alpha_{l|m} = 0.50 \pm 0.21$, $\beta_{l|m} = 1.55 \pm 0.28$, $\gamma_{l|m} = -0.50 \pm 0.61$) agrees with the HSC-XXL sample within the statistical uncertainties. It is noteworthy that we retrieve negative evolution for the MCXC sample, in agreement with CoMaLit-IV. However, we remark that we could not correct for any Malmquist bias in the MCXC sample.

[2]The catalogues are available at http://pico.oabo.inaf.it/~sereno/CoMaLit/LC2/

A more rigorous comparison is with Lovisari et al. (2015), who studied a comparable mass range. Our results agree, see Fig. 5.

Kettula et al. (2015) presented a WL and X-ray analysis of 12 low mass clusters from the Canada-France-Hawaii Telescope Lensing Survey (CFHTLS) and the XMM-CFHTLS combined with high mass systems from the Canadian Cluster Comparison Project and low-mass systems from Cosmic Evolution Survey. After correcting for Malmquist and Eddington bias, they found a slope of $\sim$1.27 for the core excised $L_X - M_{WL}$ relation, consistent with ours even though they did not consider the scatter in the WL mass.

For temperatures below $\sim$2keV, the contribution of line emission to the luminosity becomes significant. As a consequence, if clusters followed the self-similar predictions (i.e. there were no feedback effects), then the observed $L_X - T_X$ relation would flatten below $\sim$2 keV (Zou et al. 2016). To estimate the effect, we follow Zou et al.

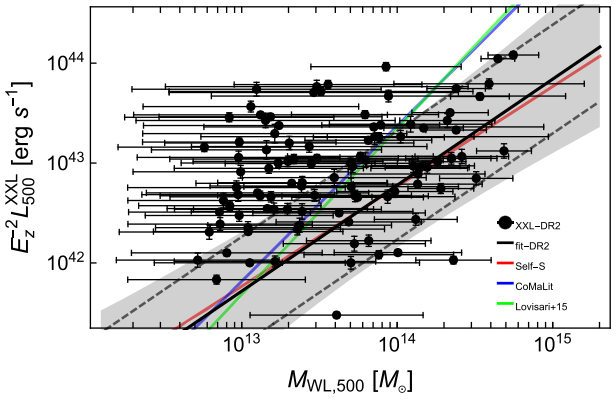

Figure 5. The luminosity–mass relation of the HSC-XXL clusters in the case of free time evolution. The dashed black lines show the median scaling relation (full black line) plus or minus the intrinsic scatter at the median redshift $z = 0.30$. The shaded grey region encloses the 68.3 per cent confidence region around the median relation due to intrinsic scatter and uncertainties on the scaling parameters. The red line is the fit for self-similar parameters. The blue and green lines plot the relations from the LC2-MCXC sample and Lovisari et al. (2015), respectively.

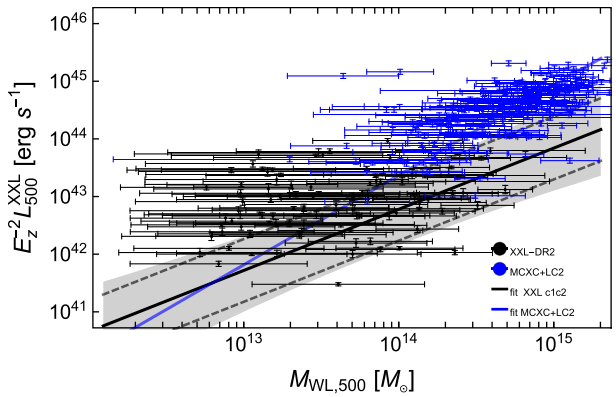

Figure 6. The luminosity–mass relation of the HSC-XXL clusters (black points) versus the LC2-MCXC sample (blue points). Black symbols and lines are as in Fig. 5.

(2016) and we measure the luminosity of APEC spectra first with a metal abundance of $Z = 0.3$, and then setting $Z = 0$ without changing any other parameters to approximate pure bremsstrahlung emission. The declining contribution of bremsstrahlung to the soft-band luminosity for $T < 2$ keV can be approximated with a power law of the form $L_{\rm brem}/L_{\rm X} \propto T^{0.21}$. Considering a self-similar evolution for the mass-temperature relation, we obtain $L_{\rm brem}/L_{\rm X} \propto M^{0.14}$. This effect can partially mask processes that are removing gas from the inner regions in low mass systems. Self-similar expectations for the slope of the $L_{\rm X} - M$ relation should then be lowered by $\sim 0.1 - 0.2$ when comparing with our result, which makes the steepening observed here more significant.

5.4 Temperature versus mass

The scaling relation between mass and temperature is presented in Umetsu et al. (2019), who also showed the agreement of the Bayesian analysis with alternative methods based on stacked signals usually employed for low signal-to-noise detections. For convenience, we report the results of the $D = 2$ analysis in Table 1 and we refer to Umetsu et al. (2019) for a detailed discussion.

Constraints on possible deviations from self-similarity are less significant than for other relations. Removal of low entropy gas from the hot phase by radiative cooling leads to higher temperatures in lower mass groups and flattens the $T_{\rm X} - M$ relation. For the $D = 4$ analysis with free time evolution, the slope of the temperature–mass relation is $\beta_{t|m} = 0.78 \pm 0.43$ and the relation is shallower than the self-similar expectation ($\beta_{t|m} = 2/3$) with a probability of $\sim$45 per cent. The probability is $\sim$22 (41) per cent for the $D = 2$ fit with free (fixed) time evolution.

5.5 Multi-scaling analysis

We find coherent evidence for effects of radiative cooling and AGN feedback from the multi-dimensional ($D = 4$) analysis, which full accounts for the effects of correlated intrinsic scatters. The probability that the $m_{\rm g} - m$, $t - m$, and $l - m$ relations are steeper, shallower, and steeper than the self-similar prediction, respectively, at the same time is $\sim$26 per cent. The probability for two (one) met criteria out of three is of $\sim$80 ($\sim$99) per cent.

These probabilities significantly exceed those associated to random fluctuations. The probability of three, or at least two, or at least one successful coin tosses out of three attempts is of 1 out of eight (12.5 per cent), one half (50 per cent), or 7 out of eight (87.5 per cent), respectively.

The total level of statistical significance is dragged down by the $T_{\rm X} - M$ relation, which is in agreement with self-similarity within the uncertainty. If we consider only the $L_{\rm X} - M$ and the $M_{\rm gas} - M$ relations, where we expect more prominent effects of AGN feedback and radiative cooling, the probability for the two slopes being more steeper than self-similarity is $\sim$78 per cent well in excess of the random probability of 25 per cent.

5.6 Intrinsic scatters and correlations

The intrinsic scatter of gas mass ($\sigma_{m_{\rm g}|m} = 0.105 \pm 0.105$, i.e. 0.24 ± 0.24 per cent from the $D = 2$ analysis) and temperature ($\sigma_{t|m} = 0.064 \pm 0.050$, i.e. 0.15 ± 0.11 per cent) at a given mass are of the order of ten–twenty per cent whereas scatter in luminosity is larger ($\sigma_{l|m} = 0.55 \pm 0.13$). These results agree with XXL Paper XXXVIII even though one major difference has to be emphasized. Here, samples for the $D = 2$ analysis are not the same for the different X-ray properties, since gas mass is measured for next to all XXL groups whereas temperature and luminosity are measured only for the brightest ones. XXL Paper XXXVIII only considered the X-ray properties of the 100 brightest clusters detected in the XXL Survey. Furthermore, WL masses were not available for XXL Paper XXXVIII and properties were measured within a fixed radius of 300 kpc.

Notwithstanding the statistical uncertainties ($\delta\rho_{Y_1Y_2} \sim 0.4$, see Table 2), we find evidence of positively correlated scatters at fixed mass, see Fig. 7. The correlation factors between gas mass and temperature, gas mass and luminosity, and temperature and luminosity are positive at the $\sim$64, 84, and 75 per cent confidence level, respectively.

The probability that at least one, at least two, or all three correlations are positive at the same time is of $\sim$56, $\sim$71, or $\sim$96 per cent, well in excess of random fluctuations. This is consistent with expectations from the dynamical state and the assembly history. The $L_{\rm X} - M$ relation is very dependent on the gas content (Truong et al. 2018). ICM processes reduce the amount of gas in the smallest systems, and at the same time their total luminosity. Simulations show that clusters move coherently along the $L_{\rm X} - T_{\rm X}$ relation

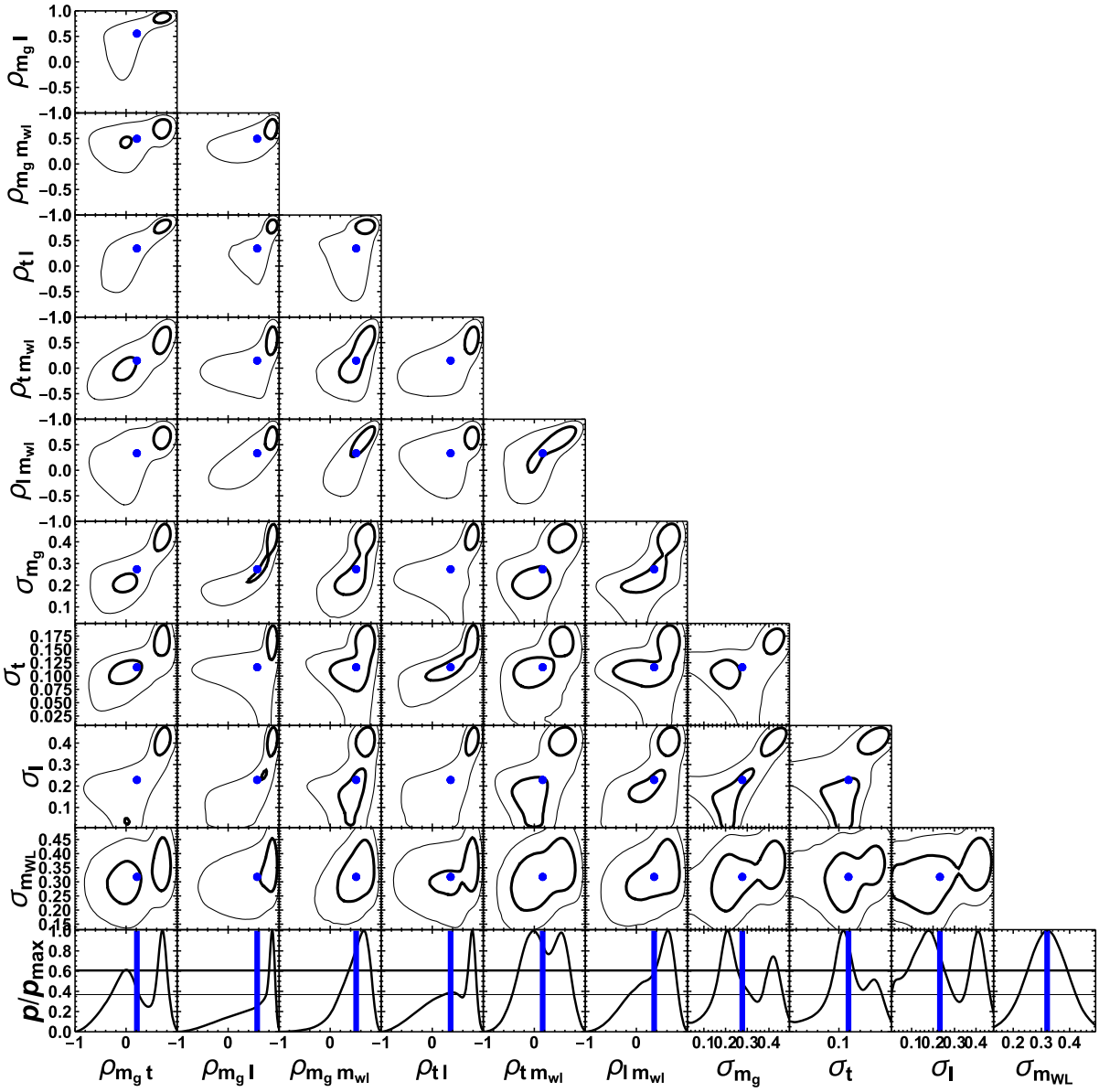

Figure 7. Probability distributions of the parameters of the intrinsic scatter covariance matrix from the joint multi-variate analysis ($D = 4$). Symbols are as in Fig. 4.

during mergers (Rowley, Thomas & Kay 2004; Hartley et al. 2008), implying correlated scatter in the $T_X - M$ and $L_X - M$ relations.

Gas mass and temperature react to merger and accretion with different time-scales. The two quantities increase at subsequent time, which lessens the expected correlation. Dark matter and gas increase in mass by the same fraction when the cluster growth happens via slow accretion, due to the constant ratio between gas and dark matter densities in the cluster outskirts, or via major mergers, due to a relatively constant gas fraction in systems of comparable mass. However, the kinetic energy of the hot gas takes more time to convert to thermal energy. Apart from the possibility of a transient shock that heats the gas with a temperature enhancement which is greater than the variation of the total mass, post-merger clusters can exhibit a lower value of temperature at fixed total mass.

The dynamical state of the cluster can cause a positive correlation between the intrinsic scatter of luminosity and temperature at fixed mass (Mantz et al. 2016b). Temperature traces mass most reliably in regular clusters that are close to virial equilibrium, whereas it can be depressed in merging clusters where energy in bulk motions has not yet virialized. Similarly, the luminosity takes some time to reach its equilibrium value, even though it is boosted by the addition of the subcluster during the merger. On the other hand, dense, bright cores occur preferentially in dynamically relaxed, hot clusters which show higher than average luminosities and approximately average temperatures. The positive correlation can be counterbalanced by strong AGN bursts, common phenomena at very high redshifts, which cause an increase of temperature along with a temporary decrease of luminosity, due to the gas displaced as ejected material (Truong et al. 2018).

According to our results, intrinsic scatters in X-ray properties appear to be correlated with the intrinsic scatter in WL mass too. In principle, WL estimates do not depend on the dynamical state or radiative processes but they are related to the mass structure of the halo and they are affected by the presence of substructures and

triaxiality (Meneghetti et al. 2010; Becker & Kravtsov 2011; Rasia et al. 2012). However, relaxed clusters are usually morphologically regular, which links scatters of WL mass and X-ray properties. The presence of substructures in the cluster surroundings may either dilute or enhance the signal (Meneghetti et al. 2010; Giocoli et al. 2014). Scatter in the WL mass can come from either massive subclumps (Meneghetti et al. 2010) or uncorrelated large-scale matter projections along the line of sight (Becker & Kravtsov 2011), and it can be inflated in morphologically complex haloes. Clumps in the gas distribution can boost the X-ray luminosity whereas dark matter substructures can dilute the weak lensing shear signal. On the other hand, the WL signal is boosted in morphologically regular and concentrated systems, which are typically dynamically relaxed and with a hotter average temperature.

Triaxiality and other deviations from spherical symmetry are major sources of WL scatter (Limousin et al. 2013; Sereno et al. 2013). Observed properties depend on the orientation of the cluster with respect to the line of sight (Gavazzi 2005; Oguri et al. 2005; Sereno 2007; Sereno & Umetsu 2011; Limousin et al. 2013; Sereno et al. 2018; Umetsu et al. 2018; Chiu et al. 2018b). Systems whose major axis points toward the observer are typically over-represented in signal-limited samples. Their lensing masses and optical or X-ray luminosities derived under the standard assumption of spherical symmetry are over-estimated. On the other hand, the majority of randomly oriented clusters are elongated in the plane of the sky and measured properties are under-estimated.

We find positive correlation between weak lensing mass and gas mass or temperature or luminosity, at the ∼93, 63, and 77 per cent confidence level, respectively. All three conditions are fulfilled at the same time with a probability of ∼56 per cent. This is most likely due to mass structure and triaxiality.

5.6.1 *Previous results*

XXL Paper XXXVIII found positive correlation between scatters in gas mass and temperature ($\rho_{m_{\rm g}t} = 0.35 \pm 0.52$), gas mass and luminosity ($\rho_{m_{\rm g}l} = 0.40 \pm 0.43$), and temperature and luminosity ($\rho_{tl} = 0.07 \pm 0.70$). Due to the lack of mass measurements, correlations were measured at a fixed latent property, related to the mass. Here, thanks to the WL masses, we can measure the correlations at a given mass.

Other literature results mostly focus on more massive clusters, i.e. $T_{\rm X} \gtrsim 4-5$ keV. Maughan (2014) applied the PICACS model to two X-ray samples, i.e. a local sample and REXCESS, with measured core-excised temperatures, gas masses, and either hydrostatic masses or luminosities within r_{500}. He found weak statistical evidence for moderate positive correlation between the scatter in T and $M_{\rm g}$ ($\rho_{m_{\rm g}t|m} = 0.31 \pm 0.30$), and between the scatter in T and the core excluded bolometric luminosity ($\rho_{tl_{\rm ce}|m} = 0.37 \pm 0.30$), and evidence for strong positive correlation in the scatter in $M_{\rm g}$ and $L_{\rm X}$ ($\rho_{l_{\rm ce}m_{\rm g}|m} = 0.85 \pm 0.14$).

Literature results on correlations between intrinsic scatters can be somewhat inconsistent, even within the same group. Mantz et al. (2015) and Mantz et al. (2016b) reported results conflicting to some degree. Mantz et al. (2015) constrained the cosmological parameters with a number count analysis of a sample of X-ray selected clusters detected in the *ROSAT* All-Sky Survey. They used follow-up measurements of soft band X-ray luminosity, temperature, and gas mass within r_{500}. WL measurements were available for a subsample of massive clusters. Under the very strong assumption of gas mass being uncorrelated with temperature ($\rho_{m_{\rm g}t|m} = 0$) and luminosity ($\rho_{m_{\rm g}l|m} = 0$), and WL mass being uncorrelated with the X-ray properties, i.e. all the off-diagonal covariance terms but $\rho_{tl|m}$ fixed to zero, they found the correlation of intrinsic scatter in L and T to be consistent with zero ($\rho_{tl|m} = 0.11 \pm 0.19$). This is consistent with Mantz et al. (2010), who analyzed a sample of 238 clusters drawn from three samples based on the *ROSAT* All-Sky Survey. However, using the same analysis method but incorporating more follow-up measurements and an updated calibration for X-ray observations than in Mantz et al. (2015), Mantz et al. (2016b) found a stronger correlation between the intrinsic scatters of luminosity and temperature at fixed mass ($\rho_{tl|m} = 0.53 \pm 0.10$).

Based on the analysis of 40 clusters identified as being dynamically relaxed and hot with measured gas mass, core excised temperature, core-excised or core-included soft-band [0.1–2.4] keV intrinsic luminosity within r_{500}, and hydrostatic masses as the true, unbiased masses, Mantz et al. (2016a) found that $\rho_{tl|m} = -0.06 \pm 0.24$ and $\rho_{tm_{\rm g}|m} = -0.18 \pm 0.28$, consistent with zero, and positive correlation between the core-included luminosity and the gas mass, $\rho_{lm_{\rm g}|m} = 0.43 \pm 0.22$. The correlation is stronger considering core-excised luminosity, $\rho_{l_{\rm ce}m_{\rm g}|m} = 0.88 \pm 0.06$. Since this sample is relaxed, results are not easily compared with Mantz et al. (2015) and Mantz et al. (2016b), where clusters were not selected based on their equilibrium status.

Okabe et al. (2010) analyzed 12 LoCuSS (Local Cluster Substructure Survey) clusters and derived a 68.3 per cent confidence lower limit of $\rho_{tm_{\rm g}|m} \gtrsim 0.185$, suggesting positive correlation between the intrinsic scatters of temperature and gas mass. The analysis was later extended to the full LoCuSS sample (Mulroy et al. 2019; Farahi et al. 2019). Under the hypothesis that the scatter of the weak lensing mass is uncorrelated from the X-ray properties, Farahi et al. (2019) found positive correlation between the intrinsic scatters of core–excised temperature and gas mass ($\rho_{tm_{\rm g}|m} = 0.13 \pm 0.03$), core–excised temperature and luminosity ($\rho_{tl|m} = 0.49^{+0.13}_{-0.16}$), and gas mass and core–excised luminosity ($\rho_{m_{\rm g}l|m} = 0.76^{+0.09}_{-0.13}$). They found also remarkable anti-correlation between the X-ray properties tracing the hot gas and the galaxy luminosity or richness tracing the cold stellar phase, which confirms that the highest-mass systems retain the cosmic fraction of baryons.

6 CONCLUSIONS

We analyzed the mass-observable scaling relations at the low mass end of the halo mass function down to groups with $M \gtrsim 10^{13}\,{\rm M}_{\odot}$. Our analysis favours gas mass– and the luminosity–total mass relations steeper than the self-similar model whereas the temperature–mass relation is consistent within statistical uncertainty. This picture is consistent with significant AGN feedback and radiative cooling in low mass systems. The measured hydrostatic bias is consistent with a small role of non-thermal pressure, even though the large statistical uncertainty does not exclude larger deviations from equilibrium.

Scatters of gas mass and temperature at a given mass are of the same order even though gas mass can be measured in a larger sample whereas temperature measurements are limited to the more luminous clusters. Luminosity is much more affected by the presence or the absence of a cool core. Each ranking of properties based on scatter size only holds for the specific operative definition of the quantities used in the analysis. A different measurement process can imply different scatters. For the XXL survey, gas masses are measured for the full sample whereas temperatures are available only for a bright subsample of clusters with sufficient photon count.

The analysis of the correlation between intrinsic scatters can unveil new features of the formation and evolution of galaxy clusters. This investigation is still in its infancy, mostly due to numerical problems in sampling the covariance matrix, which is symmetric and positive definite. Previous analyses often resorted

to simplifying assumptions (Mantz et al. 2015, 2016b; Farahi et al. 2019; Mulroy et al. 2019). Furthermore, the scatter of the weak lensing mass is often assumed as uncorrelated to simplify the treatment. Here, we found evidence for positive correlation between intrinsic scatter of ICM properties and weak lensing mass.

The positive correlation between the intrinsic scatters of X-ray quantities at fixed mass can be understood in terms of the dynamical states and the merger history of the clusters. We found also marginal evidence for a positive correlation between the X-ray quantities and the WL mass, which points to the role played by triaxiality and mass structure. Even though the sources of scatter are diverse and most processes determining X-ray properties are related to gas physics and temperature distribution that have a small impact on weak lensing estimates, triaxiality and sub-structures in the dark matter halo can correlate the scatters of WL estimates and X-ray properties at fixed mass. Asphericity can coherently affect luminosity, gas mass, and WL estimates, which are over-estimated for clusters elongated along the line of sight.

The positive correlation between X-ray properties should be taken into account to build unbiased selection functions for cosmological studies. Due to correlation, masses estimated from $M_{\rm gas}$ or $T_{\rm X}$ in an X-ray flux-limited sample would be biased high, with implications in number count analyses (Maughan 2014).

ACKNOWLEDGEMENTS

MS thanks the Hiroshima University for hospitality. We thank Lucio Chiappetti and August Evrard for comments and Cristian Vignali for arrangements. This paper is partly subsidized by Hiroshima University under the 'Program for Promoting the Enhancement of Research Universities'. MS and SE acknowledge financial contribution from contract ASI-INAF n.2017-14-H.0 and INAF 'Call per interventi aggiuntivi a sostegno della ricerca di main stream di INAF'. KU acknowledges support from the Ministry of Science and Technology of Taiwan (grant MOST 106-2628-M-001-003-MY3) and from the Academia Sinica Investigator Award (grant AS-IA-107-M01). DR was supported by a NASA (National Aeronautics and Space Administration) Postdoctoral Program Senior Fellowship at the NASA Ames Research Center, administered by the Universities Space Research Association under contract with NASA. The Saclay group acknowledges long-term support from the Centre National d'Etudes Spatiales (CNES). YF was supported by MEXT KAKENHI Nos. 18K03647.

XXL is an international project based around an XMM Very Large Programme surveying two 25 deg^2 extragalactic fields at a depth of $\sim 6 \times 10^{-15}$ erg s^{-1} cm^{-2} in the [0.5-2] keV band for point-like sources. The XXX website is http://irfu.cea.fr/xxl/. Multi-band information and spectroscopic follow-up of the X-ray sources are obtained through a number of survey programmes, summarised at http://xxlmultiwaves.pbworks.com/.

This research has made use of NASA's Astrophysics Data System (ADS) and of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

Based on observations obtained with XMM-Newton, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA.

REFERENCES

Adami C. et al., 2018, A&A, 620, A5 (XXL Paper XX)
Aihara H. et al., 2018a, PASJ, 70, S8
Aihara H. et al., 2018b, PASJ, 70, S8
Alvarez I., Niemi J., Simpson M., 2014, preprint (arXiv:1408.4050)
Andreon S., Bergé J., 2012, A&A, 547, A117
Angelinelli M., Vazza F., Giocoli C., Jones T. W., Ettori S., Brunetti G., Brüggen M., Eckert D., 2019, preprint (arXiv:1905.04896)
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Applegate D. E. et al., 2014, MNRAS, 439, 48
Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017, MNRAS, 465, 213
Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25
Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32
Bertschinger E., 1985, ApJS, 58, 39
Chiu I. et al., 2016, MNRAS, 455, 258
Chiu I. et al., 2018a, MNRAS, 478, 3072
Chiu I.-N., Umetsu K., Sereno M., Ettori S., Meneghetti M., Merten J., Sayers J., Zitrin A., 2018b, ApJ, 860, 126
D'Agostini G., 2005, preprint (arXiv:physics/0511182)
Dorman B., Arnaud K. A., Gordon C. A., 2003, in Bulletin of the American Astronomical Society, Vol. 35, AAS/High Energy Astrophysics Division #7, Americal Astronomical Society (United States), p. 641
Eckert D., Molendi S., Paltani S., 2011, A&A, 526, A79
Eckert D. et al., 2016, A&A, 592, A12 (XXL Paper XIII)
Eckert D. et al., 2019, A&A, 621, A40
Ettori S., 2015, MNRAS, 446, 2629
Ettori S., Donnarumma A., Pointecouteau E., Reiprich T. H., Giodini S., Lovisari L., Schmidt R. W., 2013, Space Sci. Rev., 177, 119
Fabjan D., Borgani S., Rasia E., Bonafede A., Dolag K., Murante G., Tornatore L., 2011, MNRAS, 416, 801
Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2018, MNRAS, 478, 2618
Farahi A. et al., 2019, Nat Commun, 10, 2504
Feigelson E. D., Babu G. J., 2012, Modern Statistical Methods for Astronomy. Cambridge Univ. Press, Cambridge
Fujita Y., Aung H., 2019, ApJ, 875, 26
Fujita Y., Umetsu K., Ettori S., Rasia E., Okabe N., Meneghetti M., 2018, ApJ, 863, 37
Furusawa H. et al., 2018, PASJ, 70, S3
Gavazzi R., 2005, A&A, 443, 793
Ghirardini V. et al., 2019, A&A, 621, A41
Giles P. A. et al., 2016, A&A, 592, A3 (XXL Paper III)
Giocoli C., Meneghetti M., Metcalf R. B., Ettori S., Moscardini L., 2014, MNRAS, 440, 1899
Giodini S., Lovisari L., Pointecouteau E., Ettori S., Reiprich T. H., Hoekstra H., 2013, Space Sci. Rev., 177, 247
Hamana T. et al., 2019, preprint (arXiv:1906.06041)
Hartley W. G., Gazzola L., Pearce F. R., Kay S. T., Thomas P. A., 2008, MNRAS, 386, 2015
Hikage C. et al., 2019, PASJ, 71, 43
Hinshaw G. et al., 2013, ApJS, 208, 19
Huang A., Wand M., 2013, Bayesian Anal, 8, 439
Kaiser N., 1986, MNRAS, 222, 323
Kawanomoto S. et al., 2018, PASJ, 70, 66
Kelly B. C., 2007, ApJ, 665, 1489
Kettula K. et al., 2015, MNRAS, 451, 1460
Komiyama Y. et al., 2018, PASJ, 70, S2
Lieu M. et al., 2016, A&A, 592, A4 (XXL Paper IV)
Lieu M., Farr W. M., Betancourt M., Smith G. P., Sereno M., McCarthy I. G., 2017, MNRAS, 468, 4872
Limousin M., Morandi A., Sereno M., Meneghetti M., Ettori S., Bartelmann M., Verdugo T., 2013, Space Sci. Rev., 177, 155
Lovisari L., Reiprich T. H., Schellenberger G., 2015, A&A, 573, A118
Mandelbaum R. et al., 2018, PASJ, 70, S25
Mantz A. B., 2016, MNRAS, 457, 1279
Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, MNRAS, 406, 1759
Mantz A. B. et al., 2015, MNRAS, 446, 2205
Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., 2016a, MNRAS, 456, 4020
Mantz A. B. et al., 2016b, MNRAS, 463, 3582

Maughan B. J., 2007, ApJ, 668, 772
Maughan B. J., 2014, MNRAS, 437, 1171
Melchior P. et al., 2017, MNRAS, 469, 4899
Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, A&A, 514, A93
Miyatake H. et al., 2019, ApJ, 875, 63
Miyazaki S. et al., 2018, PASJ, 70, S1
Mohr J. J., Mathiesen B., Evrard A. E., 1999, ApJ, 517, 627
Mulroy S. L. et al., 2019, MNRAS, 484, 60
Murata R. et al., 2019, PASJ, 71, 107
Oguri M., Takada M., Umetsu K., Broadhurst T., 2005, ApJ, 632, 841
Okabe N., Smith G. P., 2016, MNRAS, 461, 3794
Okabe N., Zhang Y.-Y., Finoguenov A., Takada M., Smith G. P., Umetsu K., Futamase T., 2010, ApJ, 721, 875
Pacaud F. et al., 2016, A&A, 592, A2 (XXL Paper II)
Pacaud F. et al., 2018, A&A, 620, A10 (XXL Paper XXV)
Pierre M. et al., 2016, A&A, 592, A1 (XXL Paper I)
Piffaretti R., Arnaud M., Pratt G. W., Pointecouteau E., Melin J.-B., 2011, A&A, 534, A109
Planck Collaboration et al., 2014, A&A, 571, A20
Planck Collaboration et al., 2016, A&A, 594, A24
Rasia E. et al., 2006, MNRAS, 369, 2013
Rasia E. et al., 2012, New J Phys, 14, 055018
Rowley D. R., Thomas P. A., Kay S. T., 2004, MNRAS, 352, 508
Sereno M., 2007, MNRAS, 380, 1207
Sereno M., 2015, MNRAS, 450, 3665 (CoMaLit-III)
Sereno M., 2016, MNRAS, 455, 2149
Sereno M., Ettori S., 2015a, MNRAS, 450, 3675 (CoMaLit-IV)
Sereno M., Ettori S., 2015b, MNRAS, 450, 3633 (CoMaLit-I)
Sereno M., Ettori S., 2017, MNRAS, 468, 3322 (CoMaLit-V)
Sereno M., Umetsu K., 2011, MNRAS, 416, 3187
Sereno M., Ettori S., Umetsu K., Baldi A., 2013, MNRAS, 428, 2241
Sereno M., Ettori S., Moscardini L., 2015, MNRAS, 450, 3649 (CoMaLit-II)
Sereno M., Covone G., Izzo L., Ettori S., Coupon J., Lieu M., 2017, MNRAS, 472, 1946
Sereno M., Umetsu K., Ettori S., Sayers J., Chiu I. N., Meneghetti M., Vega-Ferrero J., Zitrin A., 2018, ApJ, 860, L4
Sereno M., Ettori S., Eckert D., Giles P., Maughan B. J., Pacaud F., Pierre M., Valageas P., 2019, A&A, 632, A54 (XXL Paper XXXVIII)
Stanek R., Rasia E., Evrard A. E., Pearce F., Gazzola L., 2010, ApJ, 715, 1508
Svensmark J., Wojtak R., Hansen S. H., 2015, MNRAS, 448, 1644
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Torri E., Meneghetti M., Bartelmann M., Moscardini L., Rasia E., Tormen G., 2004, MNRAS, 349, 476
Truong N. et al., 2018, MNRAS, 474, 4089
Umetsu K. et al., 2014, ApJ, 795, 163
Umetsu K. et al., 2018, ApJ, 860, 104
Umetsu K. et al., 2019, preprint (arXiv:1909.10524)
Vazza F., Angelinelli M., Jones T. W., Eckert D., Brüggen M., Brunetti G., Gheller C., 2018, MNRAS, 481, L120
Vikhlinin A. et al., 2009, ApJ, 692, 1033
Voit G. M., 2005, Rev Mod Phys, 77, 207
Zou S., Maughan B. J., Giles P. A., Vikhlinin A., Pacaud F., Burenin R., Hornstrup A., 2016, MNRAS, 463, 820

APPENDIX A: MASS DISTRIBUTION

The intrinsic distribution of most astronomical quantities, e.g. the halo mass function or the luminosity function, are locally exponential (in log-space), i.e. $P_{\rm parent}(Z) \sim \exp(-aZ)$. However, in most cases, we do not have to model the full distribution. We have to model just the distribution of the clusters included in the sample (Kelly 2007). Once the parent population is filtered by the selection process, a Gaussian distribution provides a reliable approximation (CoMaLit-IV).

Table A1. Scaling relations and scatters of the simulated sample. Input and recovered parameters are listed in col. 2 and 3, respectively.

Parameter	Input	Recovered
$\alpha_{m_{\rm g}\vert m}$	− 1.000	− 1.100 ± 0.084
$\alpha_{t\vert m}$	0.320	0.290 ± 0.047
$\alpha_{l\vert m}$	0.510	0.430 ± 0.160
$\beta_{m_{\rm g}\vert m}$	1.200	1.100 ± 0.130
$\beta_{t\vert m}$	0.600	0.540 ± 0.071
$\beta_{l\vert m}$	2.100	1.900 ± 0.240
$\gamma_{m_{\rm g}\vert m}$	0.000	0.540 ± 0.330
$\gamma_{t\vert m}$	0.670	0.910 ± 0.200
$\gamma_{l\vert m}$	2.300	3.100 ± 0.650
$\rho_{m_{\rm g}t}$	0.310	0.360 ± 0.370
$\rho_{m_{\rm g}l}$	0.850	0.064 ± 0.410
$\rho_{m_{\rm g}m_{\rm wl}}$	0.000	0.210 ± 0.400
ρ_{tl}	0.530	0.160 ± 0.370
$\rho_{tm_{\rm wl}}$	0.000	− 0.004 ± 0.400
$\rho_{lm_{\rm wl}}$	0.000	0.020 ± 0.380
$\sigma_{m_{\rm g}\vert m}$	0.078	0.090 ± 0.043
$\sigma_{t\vert m}$	0.039	0.060 ± 0.019
$\sigma_{l\vert m}$	0.180	0.190 ± 0.070
$\sigma_{m_{\rm wl}\vert m}$	0.130	0.130 ± 0.068

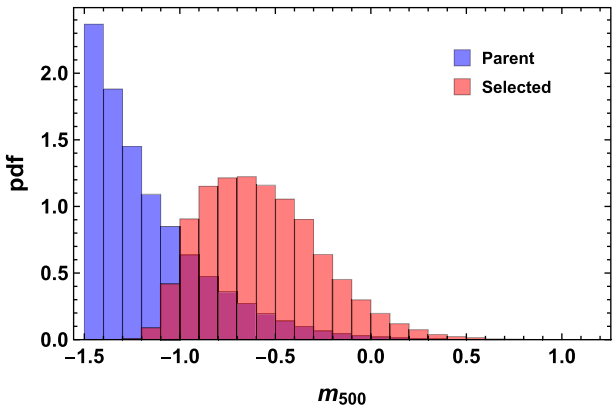

Figure A1. Renormalized mass distribution of the simulated sample. The halo mass function is in blue. The luminosity-selected subsample is in red. m_{500} is the logarithm of the mass in units of $10^{14}\,{\rm M}_{\odot}$.

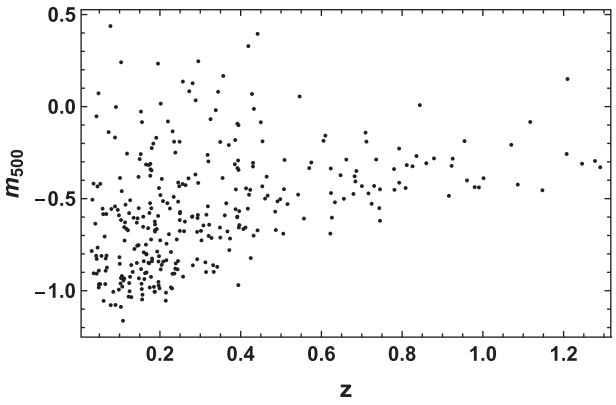

Figure A2. Masses versus redshift for the simulated sample. m_{500} is the logarithm of the mass in units of $10^{14}\,{\rm M}_{\odot}$.

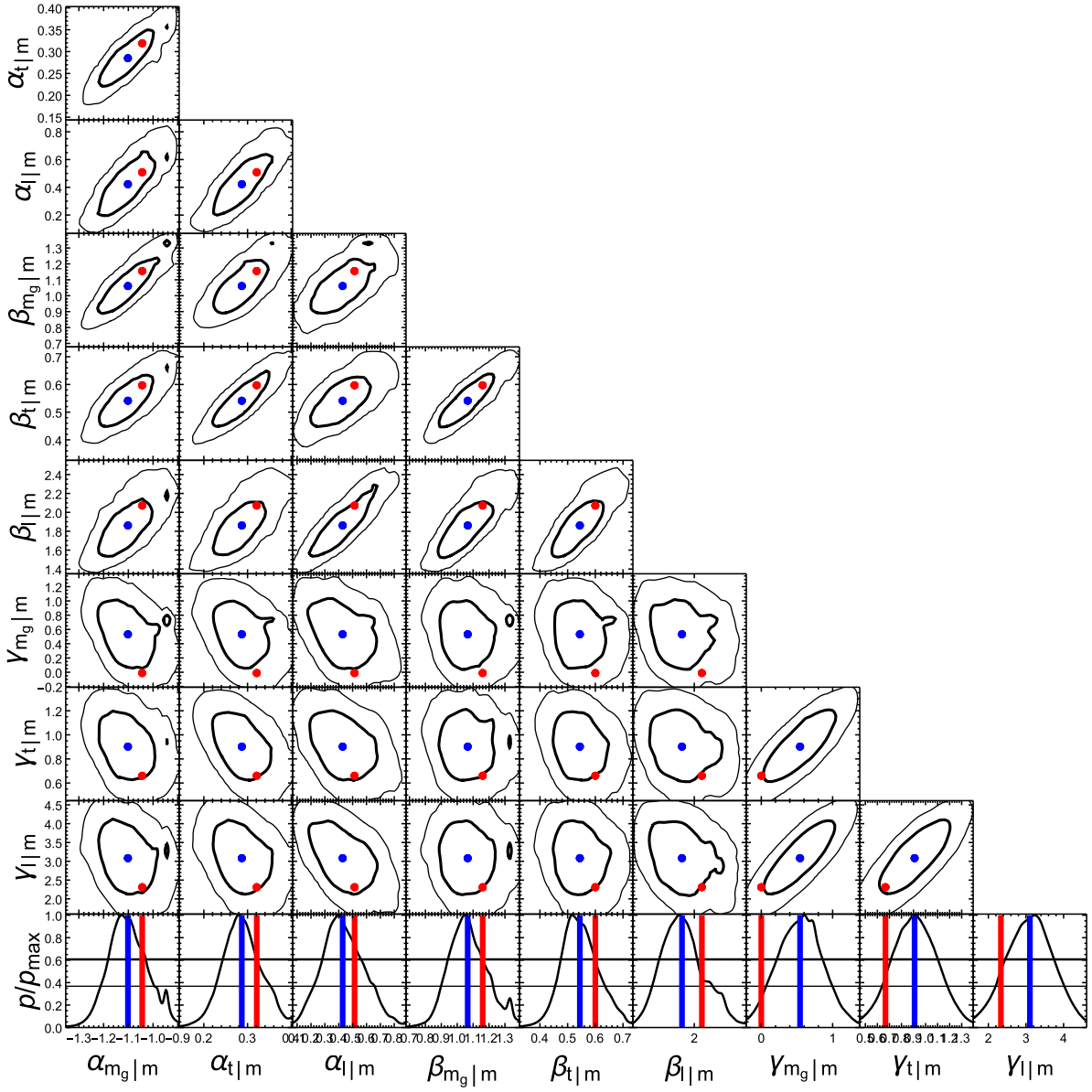

Figure A3. Probability distributions of the parameters of the scaling relations of the simulated sample from the joint multi-variate regression. The thick (thin) lines include the 1-(2-)σ confidence region in two dimensions, defined as the region within which the value of the probability is larger than a given fraction of the maximum. Red and blue symbols denote the input or the recovered bi-weight estimators of the parameters, respectively.

Paradoxically, if the parent distribution (which is more populated at very small masses) is used instead of the distribution of selected clusters (which has a long tail at small values) but the completeness is not properly accounted for, the Eddington bias is exacerbated. In Bayesian statistics, if you use misplaced priors, you get worse results.

Even though the mathematical aspects have already been shown elsewhere (Kelly 2007), it is still instructive to verify this with a numerical simulation.

We first simulate a population of clusters from the halo mass function modelled as in Tinker et al. (2008). We then generate their observable properties considering correlated intrinsic lognormal scatters and observational uncertainties and, finally, we select the sample in observed luminosity. Input values are mostly taken from Maughan (2014) and are summarized in Table A1. We retain only clusters above a cut in luminosity, $l > -1.5$, which corresponds to a rest-frame flux larger than $F_{\rm X,cut} \sim 3.2 \times 10^{-15}$ erg s^{-1} cm^{-2}.

The marginalized mass distribution of the parent and of the selected sample are compared in Fig. A1. Due to the selection process, the less massive groups at low redshift are excised, see Fig. A2. Whereas the parent population follows approximately a power-law, the selected clusters follow a peaked distribution, see Fig. A1. It is this distribution that has to be used to eliminate the Eddington bias. Even though the distribution is skewed, we approximate it by a simple Gaussian function, which proves a good approximation within the statistical uncertainties.

As measurement uncertainties, we adopted a 15 per cent accuracy on gas temperature, luminosity and gas mass, and an uncertainty on the WL mass varying linearly (in log-space) from 100 per cent at $M_{\rm WL} = 10^{13}\,{\rm M}_\odot$ to 25 per cent at $M_{\rm WL} = 3 \times 10^{14}\,{\rm M}_\odot$.

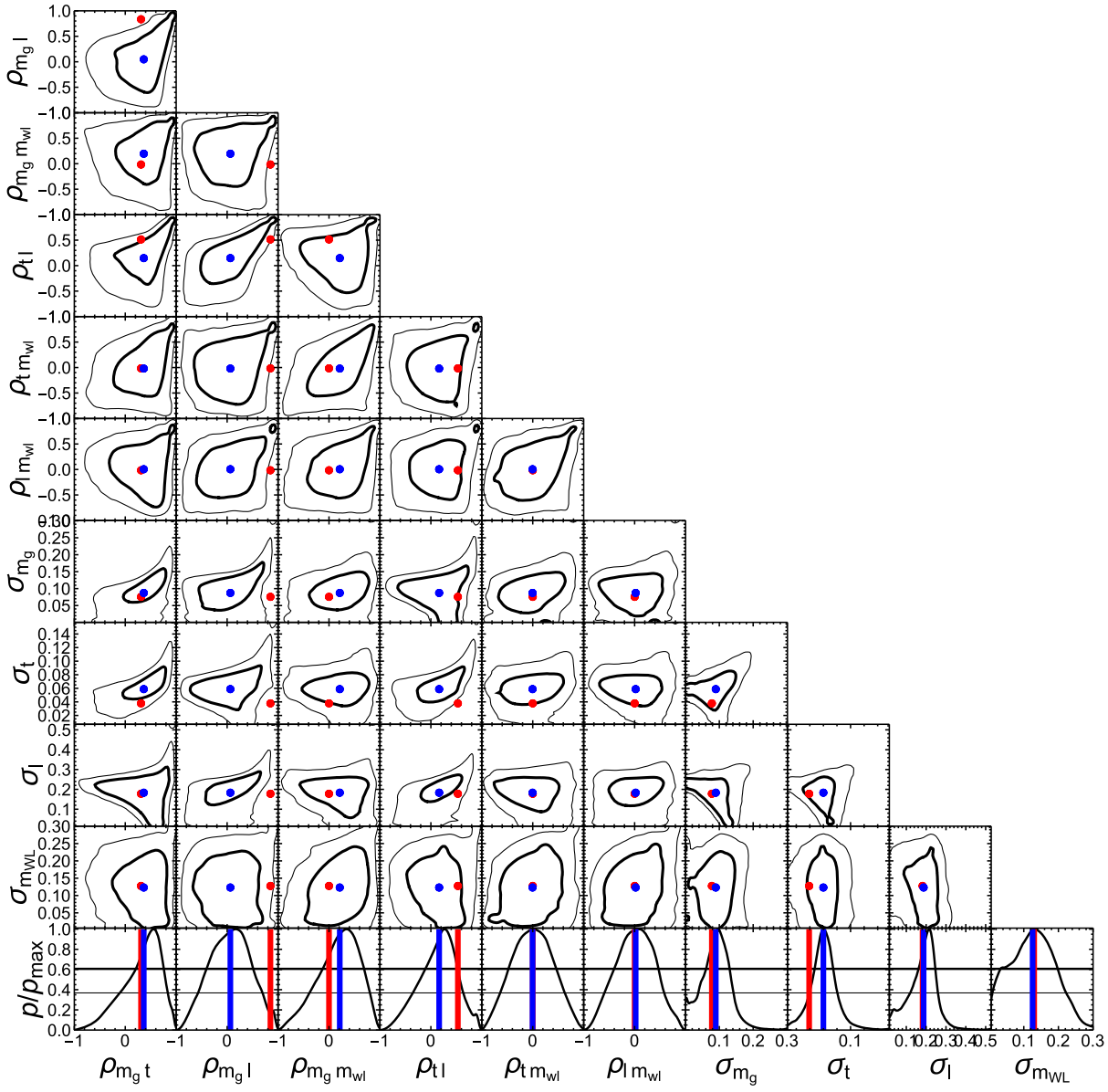

Figure A4. Same as Fig. A3 for the recovered intrinsic scatters and correlations of the simulated sample.

The simulation follows the scheme depicted in Section 2. Now, the time evolving Gaussian distribution for the mass distribution is just an approximation for the true filtered halo mass function.

As sample size, we reproduce the same number of clusters as XXL-365-GC with either known WL mass, or gas mass, temperature and luminosity. The regression procedure recovers the input parameters within the statistical uncertainties, see Table A1 and Figs A3 and A4.

APPENDIX B: THRESHOLDS FOR MANIFEST AND LATENT VARIABLES

The treatment of truncated probability distributions in presence of measurement errors can be insidious. To simplify the discussion, we consider the problem in one dimension ($D = 1$) in the following.

We first think about the marginalized distributions where we integrate out the unobserved quantity (Y). For example, we have a quantity Y that depends on some model parameters θ (e.g. Y could be normally distributed with θ comprising the mean and standard deviation). In the absence of measurement errors, if we were just truncating Y with a threshold $Y_{\rm th}$, then we can write (with $P^{'}$ indicating the truncated distribution)

$$P'(Y|\theta) = \frac{P(Y|\theta)\,\mathcal{H}(Y - Y_{\rm th})}{\int_{Y_{\rm th}}^{\infty} P(Y|\theta)dY}, \tag{B1}$$

where the integral in the denominator normalises the probability.

If we have an observation y of Y with measurement error given by $P(y|Y)$ we can marginalise out the unknown Y to write

$$P(y|\theta) = \int P(y|Y)P(Y|\theta)dY. \tag{B2}$$

Now if we apply the threshold $y_{\rm th}$ to the observed y, we have to write the truncated distribution as

$$P'(y|\theta) = \frac{P(y|\theta)\,\mathcal{H}(y - y_{\rm th})}{\int_{y_{\rm th}}^{\infty} P(y|\theta)dy} \tag{B3}$$

The integral in the denominator of equation (B3) contains the integral over Y from equation (B2). In other words, the truncation is applied to the marginalised distribution of $P(y|\theta)$ and not to the conditional distribution $P(y|Y)$. This point can be lost when the normalization terms are not explicitly written down .

The treatment can be more complicated if we want to treat the unobserved Y as a parameter of the model without marginalising it, as in the CoMaLit scheme. We cannot compute the normalization term in the denominator without integrating out Y. The correct form for the truncated distribution is

$$P'(y|Y,\theta) = \frac{P(y|Y)P(Y|\theta)\mathcal{H}(y-y_{\rm th})}{\int_{y_{\rm th}}^{\infty}\int P(y|Y)P(Y|\theta)dYdy} \tag{B4}$$

The CoMaLit solution to this problem is to propagate the threshold on y to also apply to Y (and Z). This may incorrectly seem at first like wrongly accounting for the same selection effect several times. However it is actually approximating the correct truncated distribution.

The above considerations are general, but some formulae can be worked in convenient cases. Let us consider truncated Gaussian distributions. In the CoMaLit scheme, we introduce a latent variable Z and its rescaled and shifted deterministic version Y_Z; the latent variable Y, i.e. a scattered proxy of Y_Z; and the manifest variable y, that is the measured realization of Y affected by noise. To account for Malmquist bias, we can truncate the distribution of y. Accordingly, the distribution of Y is affected too. In the following, to simplify the notation we identity Y_Z with Z.

Thanks to the chain rule, the full probability distribution can be written as

$$p_{\rm parent}(y,Y,Z) = p(y|Y,Z)p(Y|Z)p(Z). \tag{B5}$$

For lognormal scatters and uncertainties,

$$p_{\rm parent}(y,Y,Z) = \mathcal{N}(y|Y,\delta_y)\mathcal{N}(Y|Z,\sigma_{Y|Z})p_{\rm parent}(Z). \tag{B6}$$

The joint probability distribution for the selected sample, where we only retain objects if $y > y_{\rm th}$, is

$$p(y,Y,Z) = C_{\rm th}p_{\rm parent}(y,Y,Z)\mathcal{H}(y-y_{\rm th}), \tag{B7}$$

where $C_{\rm th}$ is a normalization constant.

Relations of interest are simply derived from the definition of marginal probability and the chain rule. We find

$$p(y|Y) = \frac{\mathcal{H}(y-y_{\rm th})}{\chi_{\rm erf}(Y,y_{\rm th},\delta_y)}\mathcal{N}(y|Y,\delta_y) \tag{B8}$$

where $\mathcal{N}(x|\mu,\sigma)$ is the Gaussian function with mean μ and standard deviation σ and $\chi_{\rm erf}$ is the completeness for a Gaussian variable, see e.g. Planck Collaboration et al. (2014),

$$\chi_{\rm erf}(x,\mu,\sigma) = \frac{1}{2}\left[1+{\rm erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right]. \tag{B9}$$

The conditional probability of Y given Z is a smoothly truncated Gaussian distribution

$$p(Y|Z) = \frac{\chi_{\rm erf}(Y,y_{\rm th},\delta_y)}{\chi_{\rm erf}\left(Z,y_{\rm th},\sqrt{\delta_y^2+\sigma_{Y|Z}^2}\right)}\mathcal{N}(Y|Z,\sigma_{Y|Z}). \tag{B10}$$

The final distribution of Z is the filtered parent distribution,

$$p(Z) = C_{\rm th}\chi_{\rm erf}\left(Z,y_{\rm th},\sqrt{\delta_y^2+\sigma_{Y|Z}^2}\right)p_{\rm parent}(Z). \tag{B11}$$

The normalization $C_{\rm th}$ depend on the shape of the parent population and assures that $p(Z)$ is properly normalized.

APPENDIX C: IMPACT OF PRE-DETERMINED SCALING RELATION

The measurement of the over-density radius is elusive and proxies can be used to approximate it. Some circularity can be then in place when we measure a proxy within r_{500}, $Y(<r_{500})$. The only unbiased way to measure r_{500} is to know the mass M_{500},

$$r_{500} \propto M_{500}^{1/3}. \tag{C1}$$

However, M_{500} can be unknown and we would like to approximate it with our proxy $Y(<r_{500})$. Some external scaling relations are often used. The effect of such approximations can be evaluated if the radial profile of the quantity under scrutiny is known,

$$Y \propto r^{\eta}. \tag{C2}$$

If M_{500} is approximated with an external calibration then,

$$M_{500} \propto X_{500}^{\beta}. \tag{C3}$$

We find that

$$Y_{500} \propto X_{500}^{\beta\eta/3}, \tag{C4}$$

This is the case of the luminosity estimates used in this paper, which are based on an external mass-temperature relation. The effects, which are anyway small since $\beta_{M-T}\sim 3/2$ and $\eta\sim 0.15$, have been considered for the luminosity error budget of the luminosity and for the estimated measurement correlation between luminosity and temperature, see Section 3.4.

Alternatively, we can use a previously determined relation for the very same quantity we want to measure,

$$Y_{500,\delta_\beta} \sim M_{500}^{\beta(1+\delta_\beta)}, \tag{C5}$$

where β is the underlying true slope and δ_β is the fractional error on the slope of the scaling relation we use to iteratively measure r_{500} and Y_{500} at the same time. Assuming that the error is small, $\delta_\beta \ll 1$, we found that

$$Y_{500,\delta_\beta} \sim Y_{500}^{1-\frac{\delta_\beta}{1-\eta/(3\beta)}}, \tag{C6}$$

$$\sim M_{500}^{\beta\left(1-\frac{\delta_\beta}{1-\eta/(3\beta)}\right)}. \tag{C7}$$

For our gas mass measurements, we assume the relation from XXL Paper XIII, $M_{\rm Gas}\propto M_{500}^{1.21}$. This is fully consistent with our estimate of $\beta = 1.4\pm0.4$ ($D=2$) and we expect a negligible systematic error. For the gas profile $\eta\sim 1$. Assuming a self-similar slope of $\beta = 1$, the assumption of $\beta = 1.21$, would bias low the measured slope by $\Delta\beta\sim -0.3$, which makes our detection of a steeper than self-similar relation even more significant.

APPENDIX D: PRIORS ON THE SCATTER COVARIANCE MATRIX

One of the main issue in multi-scaling analyses is the proper treatment of the scatter covariance matrix. Pros and cons on priors on scatter matrices can be found in Alvarez, Niemi & Simpson (2014), Lieu et al. (2017) and references therein. In order to simplify the problem, Mantz et al. (2015) and Mantz et al. (2016b) considered most properties as uncorrelated and reduced the problem to the treatment of a matrix of dimension 2×2. Farahi et al. (2019) and Mulroy et al. (2019) adopted priors for the correlation factors which speed up the computation but can break the requirement of positive definiteness.

Table D1. Observed scaling relations from the multi-variate analysis ($D = 4$) of 97 HSC-XXL groups. Same as Table 2 for alternative priors on the covariance matrix, see equation (D1).

Y	Intercept $\alpha_{Y\|Z}$	Slope $\beta_{Y\|Z}$	Time-evolution $\gamma_{Y\|Z}$
	Free time evolution		
m_g	-1.09 ± 0.15	1.77 ± 0.48	-0.01 ± 0.95
t	0.41 ± 0.07	0.73 ± 0.27	0.14 ± 0.53
l	0.28 ± 0.17	1.92 ± 0.59	2.96 ± 1.18

Table D2. Properties of the covariance matrix of the intrinsic scatters from the multi-scaling analysis ($D = 4$) of the HSC-XXL clusters. Same as Table 2 for alternative priors on the covariance matrix, see equation (D1).

	m_{wl}	m_g	t	l
m_{wl}	$\mathbf{0.29 \pm 0.08}$	0.44 ± 0.29	0.00 ± 0.36	0.18 ± 0.43
m_g	89%	$\mathbf{0.22 \pm 0.10}$	0.01 ± 0.42	0.31 ± 0.49
t	49%	51%	$\mathbf{0.11 \pm 0.03}$	0.16 ± 0.46
l	65%	70%	63%	$\mathbf{0.15 \pm 0.12}$

As reference prior, we consider the scaled Wishart distribution, see equation (19). As an alternative, we follow XXL Paper XXXVIII, where the prior on the (inverse of the) intrinsic scatter matrix is expressed in terms of the Wishart distribution,

$$\mathbf{V}_\sigma^{-1} \sim \mathbf{W}(\mathbf{S}, d), \tag{D1}$$

where d is the number of degrees of freedom and $\mathbf{S}$ in the $D \times D$ scale matrix. Here, $d = D + 1$, and the marginalized prior distribution of the correlation factors is uniform between -1 and 1. The diagonal elements of $\mathbf{S}$ are modelled as

$$\mathbf{S}_{aa} \sim \Gamma(\epsilon, \epsilon). \tag{D2}$$

Even though this prior is non-informative, it can favour high variance in case of high correlation. Final results are stable with respect to the two different priors, see Tables D1 and D2. Here, we favour the scaled Wishart distribution since it facilitates chain convergence and speeds up computation time.

[1]*INAF - Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Piero Gobetti 93/3, I-40129 Bologna, Italia*
[2]*INFN, Sezione di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy*
[3]*Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), No. 1, Section 4, Roosevelt Road, Taipei 10617, Taiwan*
[4]*Department of Astronomy, University of Geneva, ch. d'Ecogia 16, CH-1290 Versoix, Switzerland*
[5]*INAF - IASF Milano, via A. Corti 12, I-20133 Milano, Italy*
[6]*Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK*
[7]*European Space Astronomy Centre, ESA, Villanueva de la Cañada, E-28691 Madrid, Spain*
[8]*H. H. Wills Physics Laboratory, University of Bristol, Tyndall Ave, Bristol BS8 1TL, UK*
[9]*Department of Physical Science, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan*
[10]*Hiroshima Astrophysical Science Center, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan*
[11]*Core Research for Energetic Universe, Hiroshima University, 1-3-1, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan*
[12]*Department of Earth and Space Science, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan*
[13]*National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[14]*SOKENDAI (The Graduate University for Advanced Studies), 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[15]*Center for Astrophysics and Space Astronomy, Department of Astrophysical and Planetary Science, University of Colorado, Boulder C0 80309, USA*
[16]*NASA Ames Research Center, Moffett Field, CA 94035, USA*
[17]*Universities Space Research Association, Mountain View, CA 94043, USA*
[18]*AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France*
[19]*Institute for Astronomy & Astrophysics, Space Applications & Remote Sensing, National Observatory of Athens, GR-15236 Palaia Penteli, Greece*

This paper has been typeset from a TEX/LATEX file prepared by the author.