Systematic effects on the upcoming NIKA2 LPSZ scaling relation

A. Moyer-Anin[1,*], R. Adam[2], P. Ade[3], H. Ajeddig[4], P. André[4], E. Artis[1,5], H. Aussel[4],
I. Bartalucci[6], A. Beelen[7], A. Benoît[8], S. Berta[9], L. Bing[7], O. Bourrion[1], M. Calvo[8],
A. Catalano[1], M. De Petris[10], F.-X. Désert[11], S. Doyle[3], E. F. C. Driessen[9], G. Ejlali[12],
A. Gomez[13], J. Goupy[8], C. Hanser[1], S. Katsioli[14,15], F. Kéruzoré[16], C. Kramer[9],
B. Ladjelate[17], G. Lagache[7], S. Leclercq[9], J.-F. Lestrade[18], J. F. Macías-Pérez[1],
S. C. Madden[4], A. Maury[4], P. Mauskopf[3,19], F. Mayet[1], A. Monfardini[8], M. Muñoz-
Echeverría[1], A. Paliwal[10], L. Perotto[1], G. Pisano[10], E. Pointecouteau[20], N. Ponthieu[11],
G. W. Pratt[4], V. Revéret[4], A. J. Rigby[21], A. Ritacco[22,23], C. Romero[24], H. Roussel[25],
F. Ruppin[26], K. Schuster[9], A. Sievers[17], and C. Tucker[3]

[1]Université Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 38000 Grenoble, France
[2]Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, France
[3]School of Physics and Astronomy, Cardiff University, CF24 3AA, UK
[4]Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191 Gif-sur-Yvette, France
[5]Max Planck Institute for Extraterrestrial Physics, 85748 Garching, Germany
[6]INAF, IASF-Milano, Via A. Corti 12, 20133 Milano, Italy.
[7]Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
[8]Université Grenoble Alpes, CNRS, Institut Néel, France
[9]Institut de RadioAstronomie Millimétrique (IRAM), Grenoble, France
[10]Dipartimento di Fisica, Sapienza Università di Roma, I-00185 Roma, Italy
[11]Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France
[12]Institute for Research in Fundamental Sciences (IPM), Larak Garden, 19395-5531 Tehran, Iran
[13]Centro de Astrobiología (CSIC-INTA), Torrejón de Ardoz, 28850 Madrid, Spain
[14]National Observatory of Athens, IAASARS, GR-15236, Athens, Greece
[15]Faculty of Physics, University of Athens, GR-15784 Zografos, Athens, Greece
[16]High Energy Physics Division, Argonne National Laboratory, Lemont, IL 60439, USA
[17]Instituto de Radioastronomía Milimétrica (IRAM), Granada, Spain
[18]LERMA, Observatoire de Paris, PSL Research Univ., CNRS, Sorbonne Univ., UPMC, 75014 Paris, France
[19]School of Earth & Space and Department of Physics, Arizona State University, AZ 85287, USA
[20]Univ. de Toulouse, UPS-OMP, CNRS, IRAP, 31028 Toulouse, France.
[21]School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK
[22]INAF-Osservatorio Astronomico di Cagliari, 09047 Selargius, Italy
[23]LPENS, ENS, PSL Research Univ., CNRS, Sorbonne Univ., Université de Paris, 75005 Paris, France
[24]Department of Physics and Astronomy, University of Pennsylvania, PA 19104, USA
[25]Institut d'Astrophysique de Paris, CNRS (UMR7095), 75014 Paris, France
[26]University of Lyon, UCB Lyon 1, CNRS/IN2P3, IP2I, 69622 Villeurbanne, France

Abstract. In cluster cosmology, cluster masses are the main parameter of inter-
est. They are needed to constrain cosmological parameters through the cluster
number count. As the mass is not an observable, a scaling relation is needed to

*e-mail: alice.moyer@lpsc.in2p3.fr

link cluster masses to the integrated Compton parameters Y, *i.e.* the Sunyaev-Zeldovich observable (SZ). Planck cosmological results obtained with cluster number counts are based on a scaling relation measured with clusters at low redshift (z<0.5) observed in SZ and X-ray. In the SZ Large Program (LPSZ) of the NIKA2 collaboration, the scaling relation will be obtained with a sample of 38 clusters at intermediate to high redshift ($0.5 < z < 0.9$) and observed at high angular resolution in both SZ and X-ray. Thanks to analytical simulation of LPSZ-like samples, we take into account the LPSZ selection function and correct for its effects. Besides, we show that white and correlated noises in the SZ maps do not affect the scaling relation estimation.

1 Introduction

Clusters of galaxies are used to constrain cosmological parameters through the cluster number count: $\dfrac{\mathrm{d}N}{\mathrm{d}M\mathrm{d}z}$ [1], where N is the number of clusters, M the cluster mass and z the redshift. To have access to cluster masses, several methods exist using for example the weak and strong lensing effects [2] or multiwavelength observations combining X-ray and millimiter data. In the millimiter domain clusters are observed with the Sunyaev-Zeldovich effect (SZ) [3], which is the inverse Compton scattering of CMB photons on cluster ionized gas. It results in a shift of the CMB spectrum to high frequencies. The intensity is linked to the Compton parameter y, defined as : $y \propto \int P_e(r)\mathrm{d}l$ where $P_e(r)$ is the electronic pressure and l is the line of sight. This effect is redshift independent and has a characteristic spectral feature that makes possible cluster detections and observations up to high redshift. Past and future large scale surveys, such as Planck [4], ACT [5], SPT [6], SO [7], CMB-S4 [8], observe in the millimeter domain to get cluster catalogs. They contain information on the observed clusters, such as the redshift, the hydrostatic equilibrium mass M_{500} and the SZ obervable Y_{500}. M_{500} is defined as the mass enclosed in a sphere with density 500 times the critical density of the universe at the cluster's redshift. Y_{500} is defined as the integral of the Compton parameter up to R_{500}. For unresolved surveys, a *Y-M* scaling relation (SR) that links the mass of the cluster to the SZ observable is needed.

2 Y-M scaling relation

The *Y-M* scaling relation is derived assuming that all clusters are spherical, in hydrostatic equilibrium (HSE) and that the intra-cluster medium (ICM) is an ideal and isothermal gas. These assumptions lead to a power law:

$$E(z)^{-\frac{2}{3}}\left(\frac{D_A^2 Y_{500}}{10^{-4}\mathrm{Mpc}^2}\right) = 10^{\alpha}\left(\frac{M_{500}}{6 \times 10^{14}\mathrm{M}_{\odot}}\right)^{\beta} \tag{1}$$

with $E(z) = H(z)/H_0$ the dimensionless Hubble parameter.

However, the above-mentionned assumptions do not represent the reality of cluster population. This is the reason why there is an intrinsic scatter σ with respect to the power law SR. In the end the *Y-M* SR is defined as follows :

$$P(\log(Y_{500})|\log(M_{500})) = \mathcal{N}(\alpha + \beta\log(M_{500}), \sigma^2) \tag{2}$$

Thus the SR is defined with three parameters: α the intercept, β the slope and σ the intrinsic scatter. Among these parameters, σ is difficult to estimate as it is hidden behind the experimental dispersion as shown in figure 1. Indeed, the observed dispersion is a combination of
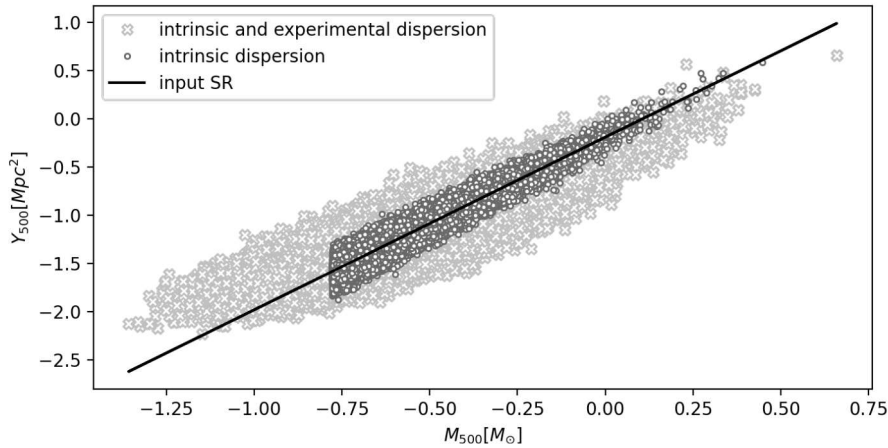
Figure 1. Simulated clusters represented in the Y-M plane. They all follow the same SR represented by the black line. Black circle points represent clusters affected by the intrinsic scatter only and grey cross points are the same clusters with an additional experimental dispersion.

the intrinsic and experimental ones. This is why, we use the LIRA software [9] to retrieve SR parameters and separate the two contributions to the total dispersion. LIRA is based on a MCMC Gibbs sampling algorithm and takes the variables, their errors and their correlations as inputs.

The NIKA2 collaboration [10] is working with its SZ Large Program (LPSZ) [11] on a new estimation of the Y-M scaling relation and of the mean pressure profile [12]. They will be obtained from a sample of 38 clusters selected from Planck [4] and ACT [5] catalogs with redshifts between $0.5 < z < 0.9$. This sample has been observed with the NIKA2 camera at the IRAM 30-m telescope in the past years. The Planck estimation [13] of the scaling relation was obtained with a sample of low redshift clusters ($z < 0.45$), whereas the LPSZ one will be obtained with intermediate to high redshift clusters. Moreover, the NIKA2 angular resolution (17.6 arcsec resolution at 2 mm) [14] allows us to have access to cluster substructures and to study cluster dynamics. This will allow us to cover effects related to more disturbed clusters, including mergers and elliptical ones.

Before studying the effects corresponding to clusters physics, we must understand and take into account systematic effects from the selection function and from the data analysis (from maps to integrated quantities).

The LPSZ selection function has two distinct contributions. On the one hand, the selection functions of Planck and ACT, since LPSZ clusters have been selected from these catalogs. And on the other hand, the effect of the box selection, which is the main contributor to the LPSZ selection function. The latter was used to force the sample to be homogeneous and not influenced by the underlying cluster's mass distribution. Boxes are 2D bins in redshift and Y_{500}. However this box selection can not be trivially processed as there are 5 thresholds in the Y_{500} axes and they cannot be treated individually. As shown by F. Kéruzoré [15], it will induce 4 Malmquist-like biases, which results in a α-bias in our sample.

3 Systematic effect study

Systematic effects are studied with simulations. The goal of these simulations is to be as realistic as possible, but at the same time, to be simple enough so that we are able to control
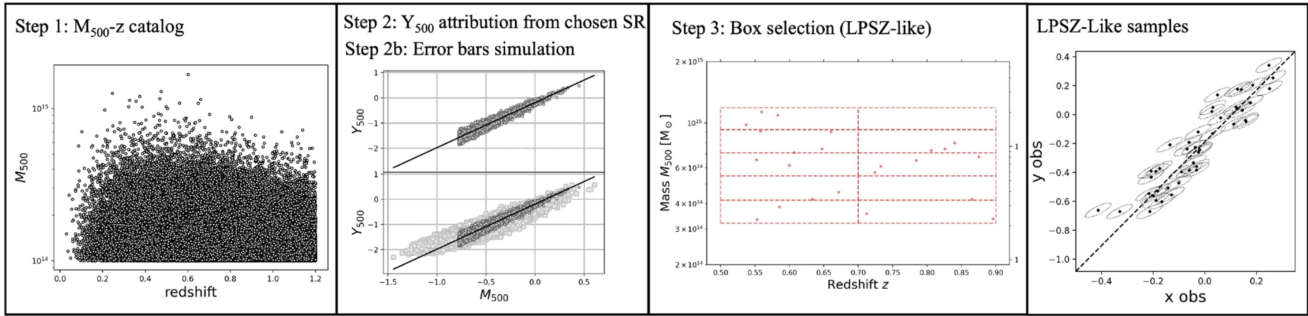
Figure 2. Steps followed to simulate a LPSZ-Like sample starting from a mass-redshift catalog. a Y_{500} is given for each simulated clusters. Then the LPSZ selection function can be applied to this catalog.

all parameters. The steps followed to simulate clusters are presented in figure 2. In step 1, we simulate a cluster population with their associated masses and redshift following a mass function [16]. In step 2, each cluster is given a Y_{500} following a chosen SR, *i.e.* with known α, β and σ. Then, to simulate observations, each Y_{500} and M_{500} value is perturbed, *i.e.* a new value is drawn inside their error bars (Step 2b). This new value follows a 2D Gaussian distribution with mean the Y_{500} and M_{500} inputs. The expected errors and correlations of the panco2 ouputs are the covariance matrix of the distribution. In step 3, the LPSZ box selection is applied to this mock catalog which gives a realistic LPSZ-like sample with a known scaling relation. After this last step, the sample is regressed by LIRA [9] thus obtaining the probability density distributions of the SR parameters.

Our goal is to study the impact of systematic effect on the estimation of SR parameters, so these steps are repeated 5000 times to avoid statistical effects. Each time, only the median value of the distributions estimated by LIRA is stored. Then, the 5000 values are compared to the input, in order to conclude on a possible bias. We repeat this study for different input SR. The Planck SR [13] was taken as a reference, which means only one parameter of Planck SR was replaced each time by the values written in table 1. The result of this study shows that, at this stage, the estimation of SR parameters are correlated with each other.

In figure 3, we present the case for which σ varies. For reasonable values of β (close to the

Table 1. Fiducial values chosen for scaling relation parameters. Planck values [13] are in bold.

α	-0.53	-0.33	**-0.19**	-0.1	-0.05	
β	1	1.66	**1.79**	2.25	2.75	3
σ	0	0.025	**0.075**	0.1	0.12	0.15

β_{Planck} estimation at the 2 σ level) and all values of α, we always retrieve input values with LIRA. As σ increases, bias on α and β increases linearly. This effect could be explained by the same mechanism as the Malmquist bias. On the right panel of figure 3 right, LIRA always retrieves the input values of the scatter σ without bias and with reasonable error bars, except for the two first points. For this sample size (around 5 clusters per box), below $\sigma \sim 0.025$ LIRA can not distinguish between the intrinsic and the experimental dispersions. Nevertheless, this should not be an issue for the LPSZ, since the value of σ is expected to be much larger ($\sigma_{Planck} = 0.075$). Knowing that one parameter will always be retrieved, we can linearly parametrize biases as a function of σ for α and β. To be more accurate, a power law parametrization has been done for α as a function of β as there is a small bias for large values of β. In figure 4, we test the bias parametrization for a given SR. On the left, we see that we
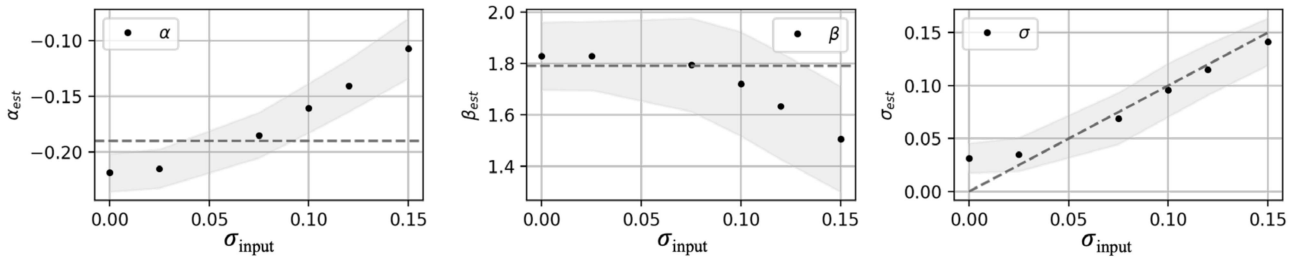
Figure 3. LIRA estimation of α, β and σ as a function of different σ input values. Black dashed lines represent input values. The points indicate the median of LIRA estimations and the shaded areas corresponds to the 68 % CL dispersion.
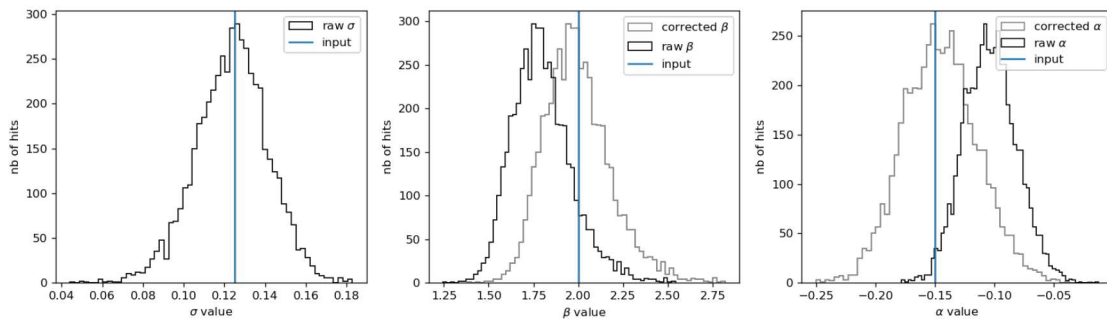
Figure 4. Distributions of the 5000 LIRA estimations of the SR parameters on samples described in section 3. Black histograms represent the raw LIRA estimation, blue ones represent the estimation corrected with the parametrization. The blue lines represent the SR input values.

do retrieve the input value of σ. For α and β the raw estimations of LIRA are biased, but after applying the above mentioned correction, the final estimations are compatible with the input SR values. Hence, the box selection function has been accounted for.

4 Effect of the pipeline analysis

Now that the selection function has been accounted for, we want to know what are the effects of the data analysis. For this purpose, we use the official LPSZ pipeline on simulated maps. We follow the steps described in section 3 (without error bar simulation). Then, maps are simulated assuming that clusters are spherical and relaxed. We assume a gNFW pressure profile [17] and typical white and correlated noises that were obtained from NIKA2-LPSZ data. We take into account other instrumental effects such as the NIKA2 beam and a typical transfer function [14] . All these simulated maps has been given to panco2 [18] which is the official software used by NIKA2-LPSZ to obtain the pressure profile and the integrated quantities from maps. As a result, panco2 gives the probability density distributions of Y_{500} and M_{500} and their correlation. These are used as input values for LIRA that gives an estimation of SR parameters to be compared with the input values.

Two map samples were simulated, one with white noise and another one with correlated noise. Regarding the estimation of the integrated quantities with panco2, we always retrieve the input values within error bars in both cases. Note that for maps with correlated noise, there is a larger dispersion around input values: 2.5% (5%) for Y_{500} (M_{500}). Whereas for

white noise we have 1.1% (2.2%) for Y_{500} (M_{500}). With these two samples, we obtain two estimations of the SR parameters. We do retrieve the input SR within two sigma confidence level from both samples but with smaller error bars for the correlated noise case[1].

5 Conclusion

Several systematic effects have been studied so far. Effects induced by the selection function can now be accounted for. Effects coming from data analysis have also been studied and no significant impacts have been observed. Understanding the effects due to the map analysis and the LPSZ selection function will help us to identify the impacts clusters physics have on the determination of the SR. The next step is to study the impact of clusters deviating from the hydrostatic equilibrium, thus taking advantage of the NIKA2 high-angular resolution. All these investigations will be useful for the upcoming LPSZ SR.

Acknowledgements

References

[1] A. Vikhlinin *et al.*, The Astrophysical Journal **692**, 1060 (2009)

[2] M. Bartelmann, IOP **27,** 233001 (2010)

[3] R. A. Sunyaev & Y. B. Zeldovich, Astro. and Space Science **7**, 20 (1970)

[4] Planck Collaboration *et al.*, Astron. Astrophys. **594**, A27 (2016)

[5] M. Hilton *et al.*, American Astronomical Society **253**, 3 (2021)

[6] L.E. Bleem *et al.*, ApJS **247**, 25 (2020)

[7] P. Ade *et al.*, Journal of Cosmology and Astroparticle Physics **2019**, 56 (2019)

[8] CMB-S4 *et al.*, American Astronomical Society **926**, 54 (2022)

[9] M. Sereno, MNRAS **455**, 2149 (2016)

[10] R. Adam *et al.*, Astron. Astrophys. **609**, A115 (2018)

[11] F. Mayet *et al.*, EPJ Web of Conferences **228**, 00017 (2020)

[12] C. Hanser *et al.*, EPJ Web of Conferences, this proceeding (2023)

[13] Planck Collaboration *et al.*, Astron. Astrophys. **571**, A20 (2014)

[14] L. Perotto, N. Ponthieu, J.-F. Macías-Pérez, *et al.*, Astron. Astrophys. **637**, A71 (2020)

[15] F. Kéruzoré, PhD thesis, Université Grenoble Alpes (2021)

[16] J. Tinker *et al.*, The Astrophysical Journal **688**, 710 (2008)

[17] D. Nagai *et al.*, The Astrophysical Journal **668**, 1 (2007)

[18] F. Kéruzoré *et al.*, Open J. Astrophys. **6** (2023), 9

---

[1]This can be explained by the fact that we add a noise covariance matrix when we analyse maps with correlated noise.