Astronomy & Astrophysics manuscript no. output
April 29, 2026

Galaxy cluster temperature maps from joint X-ray and tSZ maps with The Three Hundred hydrodynamical simulations

R. Wicker[1], M. De Petris[1,2], A. Ferragamo[3,1], I. Bartalucci[4], G.Yepes[5,6], E. Rasia[7,8,9], R. Adam[10], W. Cui[5,6,11], F. Mayet[12], L. Perotto[12], and M. Muñoz-Echeverría[13]

[1] Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, I-00185 Roma, Italy
e-mail: raphael.wicker@uniroma1.it
[2] INAF-Osservatorio Astronomico di Roma, Via Frascati 33, I-00078 Monteporzio Catone, Italy
[3] Physics Department "Ettore Pancini", Università degli studi di Napoli "Federico II", Via Cintia 21, I-80126 Napoli, Italy
[4] INAF/IASF-Milano, Via A. Corti 12, 20133 Milano, Italy
[5] Departamento de Física Teórica, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain
[6] Centro de Investigación Avanzada en Física Fundamental (CIAFF), Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain
[7] INAF-Osservatorio Astronomico di Trieste, via Tiepolo 11, I-34131, Trieste, Italy
[8] IFPU-Institute for Fundamental Physics of the Universe, via Beirut 2, 34151, Trieste, Italy
[9] Department of Physics, University of Michigan, 450 Church St, Ann Arbor, MI 48109, USA
[10] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, France
[11] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[12] Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 53, avenue des Martyrs, 38000 Grenoble, France
[13] Institut de Radioastronomie Millimétrique (IRAM), Avenida Divina Pastora 7, Local 20, E-18012 Granada, Spain

Received XXX; accepted XXX

ABSTRACT

Galaxy clusters can be used as powerful cosmological probes, provided one can obtain accurate mass estimates, which requires a precise knowledge of the underlying astrophysics of galaxy clusters. For these purposes, spatially resolved measurements of the thermodynamic properties of the intracluster medium (ICM), such as density and temperature, are necessary. In particular, temperature estimates are traditionally obtained through spatially resolved X-ray spectroscopy. Such measurements suffer from their sensitivity to the chosen energy calibration, may exhibit inherent biases, and are especially hard to perform at high redshift as they require deep observations. In recent years, however, millimeter wavelength data with high spatial resolution, comparable to that of current X-ray telescopes, have begun to be available. This has enabled the development of new methods to infer and map the cluster temperature in individual clusters, using the combination of density maps from X-ray data and pressure maps from millimeter data. In this paper, we present the first systematic validation of this approach on a large sample of synthetic clusters generated in The Three Hundred hydrodynamical simulations. We show that we are able to recover theoretical estimates of the temperature, namely the mass-weighted and spectroscopic-like temperatures, within biases of the order of $\lesssim 1\%$ in the best cases, up to $\sim 10\%$ in average, with scatters on the order of 10%. To prepare the application of this approach to observed data, we discuss the modeling of the effective length $l_{\rm eff}$, a key quantity necessary for the combination of X-ray and thermal Sunyaev-Zel'dovich projected data. In particular we provide templates calibrated on simulations for this quantity, and investigate their impact on the recovery of the temperature map, compared to other standard models.

Key words. Galaxies: clusters: intracluster medium – X-rays: galaxies: clusters – SZ effect: galaxies: clusters – Methods: numerical – Methods: statistical

1. Introduction

Sitting at the nodes of the cosmic web as massive, virialized structures, galaxy clusters can be used as robust cosmological probes (White et al. 1993; Allen et al. 2011), notably due to the dependence of the halo mass function (HMF) on the geometry of the universe. However, the quality of cluster-based cosmological constraints is extremely dependent on the mass calibration (Pratt et al. 2019), and as such accurate measurements of the galaxy cluster masses are crucial.

The most common estimates of the cluster mass using observations of the gas content, i.e. using the X-ray emission of the intracluster medium (ICM, Sarazin 1988) or relying on the thermal Sunyaev-Zel'dovich effect (tSZ, Sunyaev & Zeldovich 1972), are carried out under the assumption of the hydrostatic equilibrium (HE). However, a variety of astrophysical effects and nonthermal processes in the gas, such as turbulence, bulk motions, shocks, magnetic fields, or cosmic rays (Lau et al. 2009; Vazza et al. 2009; Battaglia et al. 2012; Nelson et al. 2014; Shi et al. 2015; Biffi et al. 2016; Gianfagna et al. 2023), may be the source of departures from the ideal conditions of HE. The resulting mass estimates will be underestimated in consequence (see e.g. Rasia et al. 2006, 2012; Gianfagna et al. 2021; Wicker et al. 2023; Muñoz-Echeverría et al. 2024), with a strong impact on the subsequent cosmological constraints (see the reviews by Voit 2005; Pratt et al. 2019). As the estimation of the hydrostatic mass relies on using radial profiles of the gas density, gas pressure, and gas temperature, a deep understanding and characterization of the ICM and its thermodynamical properties are

Article published by EDP Sciences, to be cited as https://doi.org/10.1051/0004-6361/202558025

thus critical. In that respect, temperature measurements of the ICM are of particular interest and importance. Indeed, 1D temperature profiles are necessary when computing the hydrostatic mass of a cluster from X-ray data. If in addition one wants to study the astrophysical effects in the gas including turbulence, gas sloshing, or shocks, a 2D mapping of the temperature allows access to the spatial distribution of the gas temperature, providing invaluable insight into ICM physics (Markevitch & Vikhlinin 2007).

The most standard estimate of the cluster gas temperature in observations is obtained using spatially resolved X-ray spectroscopy. However, although widely used, these measurements are plagued by numerous limitations. First, the X-ray flux is heavily dependent on redshift, decreasing as $(1 + z)^{-4}$, making high z observations extremely costly in terms of integration time, with objects appearing extremely faint. As a result, a 2D mapping of the temperature in high redshift clusters proves to be extremely challenging to carry out in a systematic manner. In addition, X-ray estimates of the temperature will be affected by systematics of various sources. On the one hand, the X-ray signal is sensitive to the squared gas density, which naturally leads X-ray observations to preferentially probe only the denser, colder regions of the core. As a result, the temperature measured by spectroscopy will actually be a weighted mean of the temperature along the line of sight, where the weights are a nonlinear combination of gas density and temperature (Mazzotta et al. 2004; Vikhlinin et al. 2006; Biffi et al. 2014). Another source of systematics stems from the fact that the X-ray spectroscopy and the subsequent temperature estimates are extremely dependent on the energy calibration of the instruments used in the observations. This leads the temperature estimates of different X-ray observatories to show strong discrepancies between one another (Schellenberger et al. 2015; Migkas et al. 2024). These disagreements may go up to 15% above 10keV between *Chandra* and XMM-*Newton* (Schellenberger et al. 2015), and even above 30% at $\sim$ 10keV between SRG/eROSITA and the two aforementioned observatories (Migkas et al. 2024). Finally, the uncertainties in temperature measurements increase with the temperature. As a result, the X-ray estimates of the ICM temperatures, notably with XMM-*Newton* and *Chandra* are especially hard to obtain at high temperatures.

As a result of these shortcomings, alternative methods to infer galaxy cluster temperatures need to be proposed and investigated. In order to infer 1D temperature profiles, or even directly mass profiles, works such as Ameglio et al. (2007, 2009); Mroczkowski et al. (2009); Adam et al. (2015, 2016); Ruppin et al. (2017); Ghirardini et al. (2018) have started using a combination of X-ray and tSZ observations, using the density information from the former and the pressure information from the latter.

Adam et al. (2017) (which we will note A17 in the rest of the paper) went a step further and used a combination of XMM-*Newton* and high angular resolution NIKA tSZ data to infer a temperature map in the cluster MACS J0717.5+3745. Similarly, Artis et al. (in prep) combine XMM-*Newton* and NIKA2 observations and manage to highlight an overheated region in the cluster PSZ2G091.83+26.11. This approach not only requires spatially resolved X-ray and SZ data but also relies heavily on the so-called "effective length". The X-ray signal is sensitive to the projected squared density, while the mapping of the temperature requires a map of the projected density. As a result, a conversion from the former to the latter is necessary, through the effective length. The effective length is sensitive to the distribution of the gas along the line of sight and is thus sensitive to clumping and to the presence of substructures along the direction of the projection. As such, the effective length is not a trivial quantity to compute and needs to be modeled. A17 and Artis et al. (in prep.) both modeled the effective length in various ways, based on direct observations, without resorting to hydrodynamical simulations.

The purpose of this work, built as a series of two papers, is to validate this joint X-ray/tSZ mapping approach on a large simulated sample of clusters and then apply it in a systematic manner to an observed sample of galaxy clusters, for which both X-ray imagery and high-resolution tSZ imagery are available, namely the NIKA2 tSZ Large Program (Mayet et al. 2020; Perotto et al. 2022). The current paper focuses on the theoretical work in a simulated setup, mainly on the modeling of the effective length, while the second paper will focus on the application of the method to realistic mock maps and to observed data. In Section 2 we describe our simulated dataset and the maps that we use for this work. The full theoretical modeling necessary to the approach is detailed in Section 3, while we detail our results in Section 4, with a focus on the comparison between our reconstruction of the temperature and other theoretical estimates of the temperature. We discuss these results and the possible systematics in Section 5, before presenting our main conclusions.

2. Simulation dataset

2.1. The NIKA2 LPSZ Twin Samples

The cluster sample used in this work was selected from The Three Hundred (The300 hereafter) suite of hydrodynamical simulations (Cui et al. 2018). The300 simulations consist of zoom-in re-simulations of the 324 host regions of the most massive cluster-size halos from the dark-matter-only MultiDark Simulation run carried out with Planck cosmology (Planck Collaboration et al. 2016b). Each cluster sits at the center of boxes of $15h^{-1}$ comoving Mpc of radius, and was simulated using runs of several different codes, at different resolutions. We used the clusters from the runs performed using the Gadget-X code, at low resolution (3840^3 dark matter particles, thus labeled Gadget-X 3k).

The selection of the cluster sample used in this work, dubbed the NIKA2 LPSZ Twin Samples (NIKA2-TS, or NTS, hereafter), was chosen to match the statistical properties of the observed NIKA2 SZ Large Program (NIKA2 LPSZ) (Mayet et al. 2020; Perotto et al. 2022). The NIKA2 LPSZ consists of a representative sample of 38 intermediate-redshift clusters ($z \in [0.5, 0.9]$) with masses spanning an order of magnitude from 10^{14} to $10^{15} M_\odot$, of which we give the distribution in the mass-redshift plane in Fig. A.1 in appendix. These clusters were originally detected in the *Planck* and Atacama Cosmology Telescope (ACT) SZ cluster catalogs (Planck Collaboration et al. 2016a; Hasselfield et al. 2013). These clusters have been observed with the NIKA2 camera, in two bands of 150 GHz and 260 GHz. The angular resolution achieved by NIKA2 reaches $\sim$ 17.6" full width half maximum (FWHM) at 150 GHz, and $\sim$ 11.1" FWHM at 260 GHz. In this work, we only used synthetic maps that mimic the 150 GHz observations, since the tSZ signal from the clusters is stronger at this frequency.

Selected within four snapshots between $z = 0.49$ and $z = 0.817$, to match the original NIKA2 LPSZ redshift coverage, the Twin Samples are three samples containing 38 clusters each. The three Twin Samples, denoted here NTS-M_{tot}, NTS-M_{HE}, and NTS-Y_{500} have been selected to match the redshifts of the NIKA2 LPSZ clusters and their total mass $M_{tot,500}$, hydrostatic

mass $M_{HE,500}$, and the integrated Compton parameter Y_{500}, respectively. Some of the objects are, in addition, common between two or more NTS, leading to a final total of 96 simulated clusters. The complete description of the selection and production of the Twin Samples can be found in Paliwal et al. (2022).

2.2. Simulated maps

This work required the use of different types of map, which have been produced using The300 simulated data from the GADGET-X 3k runs. To produce the maps, we selected only the gas particles that count as ICM, which could be responsible for X-ray and tSZ signals. We thus selected gas particles with a density below $10^{-1}\mathrm{cm}^{-3}$, below the star-formation density threshold, and with temperatures higher than 0.5 keV. We focused on cubic regions centered on the main cluster of each simulation box, of $4R_{200} \times 4R_{200} \times 4R_{200}$ [1] in size. We then produced theoretical maps of the tSZ signal (we label these maps tSZ or y_{tSZ} interchangeably within the paper), the X-ray signal (in truth the emission measure, EM_{n_e}), the temperature, and of the effective length l_{eff}.

Theoretical tSZ maps, i.e. simulated maps of the Compton parameter y_{tSZ} (defined in Eq. 5), were produced using the `pymsz` (Chluba et al. 2012, 2013; Cui et al. 2018) Python package. The code retrieves the thermodynamic quantities, including the density and temperature, to compute the pressure and in turn the theoretical Compton parameter y_{tSZ} for all the selected particles. Then a smoothed particle hydrodynamics (SPH) smoothing kernel is applied to the Compton-y_{tSZ} data.

Theoretical maps of the emission measure were produced from the particles using the code `Smac` (Dolag et al. 2005; Ansarifard et al. 2020). The code smooths the field by applying an SPH kernel, then computes and projects the distribution of the squared gas density ρ^2_{gas} along the line of sight, in units of $[\mathrm{gr}^2\mathrm{cm}^{-6}\mathrm{kpc}^2\mathrm{cm}]$. The output needs then to be converted to standard units of the projected emission measure $[\mathrm{cm}^{-6}\mathrm{Mpc}]$, using the mean molecular weight $\mu = 0.618$, the proton mass, and the electron-to-proton abundance ratio $n_e/n_p \sim 1.16$ (with slight variations in the field). We refer the reader to Veronesi et al. (2025) for more details on the mapping process. In addition, Sect 2.2.2. of this paper discusses how in the considered temperature range, the temperature and metallicity dependence of the cooling function is very weak, so that the emission measure map is very close to the actual X-ray signal map.

Finally, we produce maps of the effective length that we define in Eq. 4. To produce these maps, we start directly from the simulation particles' densities. We run the SPH code SPLASH (Price 2007) on the simulation files to obtain smoothed cubes, which we project to generate maps of density and of squared density simultaneously.

As it appears then, we use three different kinds of map ($y_{tSZ}, \mathrm{EM}_{n_e}, l_{\mathrm{eff}}$) to obtain the temperature, produced using three different codes. These maps have different properties, with different pixel numbers, pixel sizes, and different angular resolutions. Indeed, while the pixel number and pixel size for the y_{tSZ} maps are dictated by the properties of the true NIKA2 maps, which span $16.65' \times 16.65'$ with a pixel size of 3", the default mapping setting for the emission measure maps is to have 200 pixels within R_{500}. In order to avoid inconsistencies between the maps due to these differences, we transform the EM_{n_e} and l_{eff} maps to match the properties of the y_{tSZ} map. Namely, we first resample the maps so that the angular pixel size of all maps matches that of the tSZ map, i.e. with a pixel size of 3". We then match the spatial extension of the different maps by cropping them to correspond to the extension of the tSZ map. Finally, we smooth all our input maps with a Gaussian kernel to match the angular resolution of the synthetic maps to the resolution of the NIKA2 camera at 150GHz of $\sim$ 17.6".

[1] R_Δ correspond to the radius within which the mean density corresponds to Δ times the critical density of the universe, $\rho_c = 3H^2/8\pi G$. Similarly, M_Δ is the mass enclosed within R_Δ

3. Theoretical modeling

3.1. Reconstruction of the temperature

The temperature measurement in a given sky direction resulting from the combination of tSZ and X-ray data, denoted T_{SZX} hereafter, can be written as the ratio of a projected "effective pressure" $\overline{P_e}$, obtained from tSZ, and "effective density" $\overline{n_e}$ obtained from X-ray

$$k_B T_{SZX} = \frac{\overline{P_e}}{\overline{n_e}}, \tag{1}$$

where k_B is the Boltzmann constant. These effective quantities link to the projected electronic pressure and electronic density as follows

$$\overline{P_e} = \frac{1}{l_{\mathrm{eff}}} \int_{l.o.s.} P_e \mathrm{d}l \tag{2}$$

and

$$\overline{n_e} = \frac{1}{l_{\mathrm{eff}}} \int_{l.o.s.} n_e \mathrm{d}l, \tag{3}$$

through the effective length l_{eff}, discussed in the introduction. Following the work of A17, the effective length is defined as

$$l_{\mathrm{eff}} = \frac{\left(\int_{l.o.s.} n_e \mathrm{d}l\right)^2}{\int_{l.o.s.} n_e^2 \mathrm{d}l}, \tag{4}$$

effectively allowing for the conversion of the projected squared density obtained from X-ray data into the projected density. In a real observational setup, the maps of the projected pressure and squared density are obtained respectively from the tSZ signal y_{tSZ} and X-ray surface brightness S_X maps, with

$$y_{\mathrm{tSZ}} = \frac{\sigma_T}{m_e c^2} \int_{l.o.s.} P_e \mathrm{d}l, \tag{5}$$

where σ_T is the Thomson cross-section, m_e the electron mass, c the celerity of light, and

$$S_X = \frac{1}{4\pi(1+z)^4} \int_{l.o.s.} \Lambda(T_e, Z) n_e^2 \mathrm{d}l, \tag{6}$$

with $\Lambda(T_e, Z)$ the cooling function depending on gas temperature T_e and metallicity Z, taking into account the interstellar absorption and the instrument response. As a result, the complete expression for the reconstruction of T_{SZX} from Eq. 1, can be expanded following Eqs. 2 to 6 as :

$$k_B T_{SZX} = \frac{\overline{P_e}}{\overline{n_e}} = \frac{m_e c^2}{\sigma_T} y_{\mathrm{tSZ}} \sqrt{\frac{\Lambda(T_e, Z)}{4\pi(1+z)^4 S_X \cdot l_{\mathrm{eff}}}}. \tag{7}$$

However, since this work relies on a purely theoretical setup, without observational contributions, the complete formula for

$k_B T_{SZX}$ differs slightly from Eq. 7. Indeed, we obtain the theoretical effective density $\overline{n_e}$ from the emission measure maps, computed inside Smac as :

$$\mathrm{EM}_\rho = \sum_{i \in l.o.s.} \rho_{g,i}^2 V_i = \sum_{i \in l.o.s.} \rho_{g,i} m_{g,i}, \tag{8}$$

where ρ_g is the gas mass density and m_g the gas mass. We need to convert this map to a map of the projected squared electronic density,

$$\mathrm{EM}_{n_e} = \sum_{i \in l.o.s.} n_{e,i}^2 l_i. \tag{9}$$

We use to this effect the relation $\rho_g = \mu m_p (n_e + n_p)$, so that the final maps we use are in units of [cm^{-6}Mpc].

The resulting effective density thus writes :

$$\overline{n_e} = \sqrt{\frac{\mathrm{EM}_{n_e} \cdot 10^3}{l_{\mathrm{eff}}}}, \tag{10}$$

where the factor 10^3 is simply a conversion factor that accounts for the emission measure in [cm^{-6}Mpc], with l_{eff} being in kpc. As a result, the expression for $k_B T_{SZX}$ in the current setup is written as:

$$k_B T_{SZX} = \frac{\overline{P_e}}{\overline{n_e}} = \frac{m_e c^2}{\sigma_T} \frac{1}{\sqrt{l_{\mathrm{eff}}}} \frac{y_{tSZ}}{\sqrt{\mathrm{EM}_{n_e} \cdot 10^3}} \tag{11}$$

To assess the quality of the temperature map inference using the joint X-ray and tSZ maps, we need to compare it to other standard estimates of the temperature. In this paper, we focus on the mass-weighted temperature T_{mw}, of which T_{SZX} should be the closest approximation (A17), and on the spectroscopic-like temperature T_{sl}. Indeed, these two theoretical computations of the temperature should be the closest estimates of the total thermal energy in the ICM, and of an observed X-ray temperature, respectively (Mazzotta et al. 2004; Biffi et al. 2014; Henden et al. 2018; Pellissier et al. 2023; ZuHone et al. 2023). However, it is worth noting that depending on the simulation, the spectroscopic-like temperature may be biased with respect to the X-ray temperature obtained by fitting synthetic spectra, whether it is only for low temperatures or the full temperature range (Biffi et al. 2014; Hahn et al. 2017; Henden et al. 2019; Pellissier et al. 2023; ZuHone et al. 2023).

The mass-weighted temperature is defined as follows from Biffi et al. (2014)

$$T_{mw} = \frac{\sum_{i \in l.o.s.} T_i m_i \mathrm{d}l}{\sum_{i \in l.o.s.} m_i \mathrm{d}l}, \tag{12}$$

where m is the mass of gas particles, while the spectroscopic-like temperature, from Mazzotta et al. (2004), can be defined as :

$$T_{sl} = \frac{\sum_{i \in l.o.s.} T_i w_i \mathrm{d}l}{\sum_{i \in l.o.s.} w_i \mathrm{d}l}, \tag{13}$$

where the weight w_i is defined as

$$w_i = \frac{n_{e,i}^2}{T_i^{3/4}}. \tag{14}$$

We produce spectroscopic-like and mass-weighted temperature maps using Splash, in a similar manner to the l_{eff} maps. In order to carry out a meaningful comparison, we in addition process the maps in the same way we process the input $(y_{tSZ}, \mathrm{EM}_{n_e}, l_{\mathrm{eff}})$ maps, i.e. resizing, cropping, and convolving them by a Gaussian kernel to match the properties of the y_{tSZ} map.

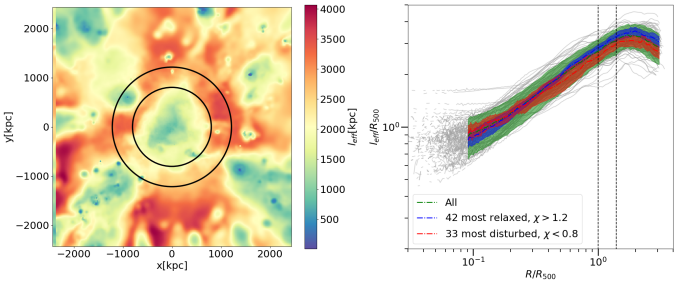

Fig. 1. *Left*: Effective length map for the main cluster of region 211 of the snapshot 101, at $z = 0.817$. The two rings correspond to R_{500} and R_{200}. *Right :* 1D profiles of the effective length for all clusters in our sample. We give in addition the median l_{eff} profiles computed for the full sample (green), the most relaxed (blue) and the most disturbed (red) clusters of the sample. The vertical dashed lines correspond to the $1.0R_{500}$ and $1.4R_{500} \sim R_{200}$, while the shaded areas correspond to the 1σ dispersion regions.

3.2. Modeling of the effective length

In the framework of hydrodynamical simulations, we have access to the true map of effective length for each cluster, from which we can reconstruct T_{SZX}. However, this quantity is not directly observable and, as such, preparing an application of the reconstruction to real data first requires an effort to model l_{eff} at the required level of accuracy.

The most basic modeling that can be adopted for the effective length is simply assuming a constant value l_0 for the quantity, as was done as a test, for example, in A17 and Artis et al. (in prep.). The results of these two works suggest that this assumption may impact the normalization of the temperature map (Artis et al. in prep) as well as the recovery of the substructures in the image (A17), although the magnitude of this effect is not fully understood. As a result, when trying to infer temperature maps with good recovery of their spatial structure *and* accurate absolute measurements, robust modeling of l_{eff} is crucial.

A17 and Artis et al. (in prep.) proposed a variety of different models in the observational data, relying on the use of β density profiles (Cavaliere & Fusco-Femiano 1976) fitted in the X-ray images. In this work, which is focused on the use of hydrodynamical simulations, we have access to "true" effective length maps that we can use to infer and test a more accurate modeling of l_{eff}. However, a downside of the use of simulations is the integration length along the line of sight, which is not infinite. This will limit the validity of our modeling within a certain range of radii. In this precise case, with an integration length of $4 \times R_{200}$, our modeling of the effective length is valid up to slightly outside R_{200}. For this reason, in this paper we only show the result maps up to R_{200}.

For all clusters in the sample, the effective length exhibits a similar behavior, with low values of l_{eff} in the center and gradually increasing towards the outskirts, as we show in Fig.1. All clusters also exhibit a flattening of the effective length, and in some cases a decrease for radii $\gtrsim R_{200}$, at the limit of our range of validity. This global shape, common to all clusters, was described in A17 with a shape factor Q, related to the mean density and to clumping along the line of sight. The quantity l_{eff}/R_{500}, plotted in Fig. 1, corresponds here to $1/Q^2$, directly linking the shape of the l_{eff} profiles to the properties of the gas distribution in the clusters. The low values of l_{eff} in the center closely correlate with R_{500}, where the mean density is high, the gas is clumped, and other lower-density structures along the line of sight will only weakly contribute to the value of l_{eff}. As the mean density

Table 1. Parameter values for the different GHP templates calibrated using the three combined Twin Samples

Parameter	Median template	Relaxed template	Disturbed template
$\mathbf{A_0}$	$3.973^{+1.122}_{-0.493}$	$4.012^{+1.004}_{-0.346}$	$3.437^{+0.438}_{-0.154}$
$\mathbf{B_0}$	0.463 ± 0.156	$0.534^{+0.088}_{-0.168}$	$0.500^{+0.058}_{-0.112}$
α	$0.501^{+0.038}_{-0.031}$	$0.540^{+0.013}_{-0.011}$	$0.447^{+0.010}_{-0.009}$
β	$5.531^{+5.435}_{-2.542}$	$5.037^{+3.599}_{-1.692}$	$8.508^{+4.170}_{-3.305}$
γ	$1.418^{+1.920}_{-0.659}$	$1.083^{+1.130}_{-0.293}$	$1.037^{+0.981}_{-0.257}$
Number of clusters	96	42	33

Notes. The parameter values are given with the 16th and 84th percentiles from our posterior distributions.

decreases towards the outskirts, the impact of outside structures (background and foreground groups, smaller clusters, filaments) will increase and introduce more variability in the shapes of the profiles. This will result in slightly different slopes, maximal values of the shape factor, or radii at which the effective length flattens, and starts to decrease.

In order to be able to fit this particular shape, as well as to account for the variety of possible $l_{\rm eff}$ profiles, we propose a functional form written as:

$$l_{\rm eff}(R) = A_0 \mathrm{R}_{500} \frac{\left(B_0 \frac{R}{\mathrm{R}_{500}}\right)^{\alpha}}{\left(1 + \left(B_0 \frac{R}{\mathrm{R}_{500}}\right)^{\beta}\right)^{\gamma/\beta}}. \quad (15)$$

The parameters A_0 and R_{500}/B_0 represent respectively the amplitude and cutoff radius of the $l_{\rm eff}$ profile, α the main power-law slope of the profile, and β and γ dictate the shape of the flattening and eventual decrease of the profile in the outskirts.

This functional form was derived as a "generalization" of the functional form for the transfer function of a first order high pass filter, which exhibits the same features but restricted to the case $(\alpha = 1, \beta = 2, \gamma = 1)$, hence the name of "generalized high pass" (GHP) model that we use in the rest of the paper for simplicity. This functional form can also be found in the temperature profile proposed by Vikhlinin et al. (2006).

An advantage of this approach is that once calibrated in simulations, this template profile, which can be propagated in a 2D template map, only requires the estimation of the R_{500} of a cluster to be applied directly in observations. In addition, this parametrization, albeit complex, accounts for a wide variety of shapes, allowing the calibration of the templates to be fine-tuned.

We calibrate the model by performing an MCMC, in order to fit the model to the true effective length profiles. We compute in particular three different templates from this GHP model, calibrated on the full sample, on the most relaxed clusters, and the most disturbed clusters of the sample. These are our Median, Relaxed, and Disturbed templates, respectively. We base our estimation of the dynamical state of the cluster on the relaxation parameter χ from De Luca et al. (2021), itself derived from the more general definition of Haggar et al. (2020). The definition of χ used in this paper is thus :

$$\chi = \sqrt{\frac{2}{\left(\frac{\Delta_r}{0.1}\right)^2 + \left(\frac{f_s}{0.1}\right)^2}}, \quad (16)$$

where Δ_r is the center-of-mass offset and f_s is the sub-halo mass-fraction. To segregate the relaxed and disturbed clusters used to calibrate the templates, we use these values computed at R_{500}, and select as relaxed the clusters for which $\chi > 1.2$, and disturbed clusters for which $\chi < 0.8$. Using this selection, the Relaxed template has been calibrated using 42 clusters and the Disturbed template using 33 clusters, out of the 96 of the full sample (see Fig.1).

In Table 1 we give the value of the fitted parameters for the three calibrated templates. With values of $A_0 = 3.973^{+1.122}_{-0.493}$ and $A_0 = 4.012^{+1.004}_{-0.346}$ respectively, the median and relaxed templates exhibit slightly larger amplitudes than the disturbed ones, with $A_0 = 3.437^{+0.438}_{-0.154}$. However, all templates seem to exhibit compatible cutoff radii, with $(B_{0,median}, B_{0,relaxed}, B_{0,disturbed}) = (0.463 \pm 0.156, 0.534^{+0.088}_{-0.168}, 0.500^{+0.058}_{-0.112})$, meaning that the flattening in the effective-length profile generally occurs around $2 \times \mathrm{R}_{500}$. The main difference between the different templates lies in the value of the main slope α, with the disturbed effective length profile appearing flatter, with $\alpha = 0.447^{+0.010}_{-0.009}$, than the relaxed one where $\alpha = 0.540^{+0.013}_{-0.011}$. Logically, the slope of the median template is in between the two extremes, with $0.501^{+0.038}_{-0.031}$. The steeper slope of the relaxed template can be easily understood by the gas distribution of relaxed clusters, which is more concentrated than that of disturbed clusters. However, due to degeneracies, the parameters β and γ that encapsulate the shape of the knee in the $l_{\rm eff}$ profile are poorly constrained and do little to help segregate the different templates.

3.3. Comparative performance between GHP and n_e-model inferred $l_{\rm eff}$

We want to assess here the advantages of using a GHP template map calibrated on simulations with respect to the more standard approach based on the use of fitted density profiles. To that end, we infer $\mathrm{T_{SZX}}$ temperature maps using both our GHP model and using 3D density profiles fitted using two different models, a β model (Cavaliere & Fusco-Femiano 1976, given in Eq. 17) and a Vikhlinin et al. (2006) model (given in Eq. 18)

$$n_{e,\beta}(r) = n_0 \cdot \left(1 + \left(\frac{r}{r_c}\right)^2\right)^{-3\beta/2}, \quad (17)$$

$$n_{e,Vikhlinin}(r) = \left[n_{0,1}^2 \left(\frac{r}{r_{c,1}}\right)^{-\alpha} \cdot \left(1 + \left(\frac{r}{r_{c,1}}\right)^2\right)^{\frac{\alpha}{2} - 3\beta_1} \cdot \left(1 + \left(\frac{r}{r_a}\right)^{\gamma}\right)^{-\frac{\epsilon}{2\gamma}} + n_{0,2}^2 \left(1 + \left(\frac{r}{r_{c,2}}\right)^2\right)^{-3\beta_2} \right]^{1/2}. \quad (18)$$

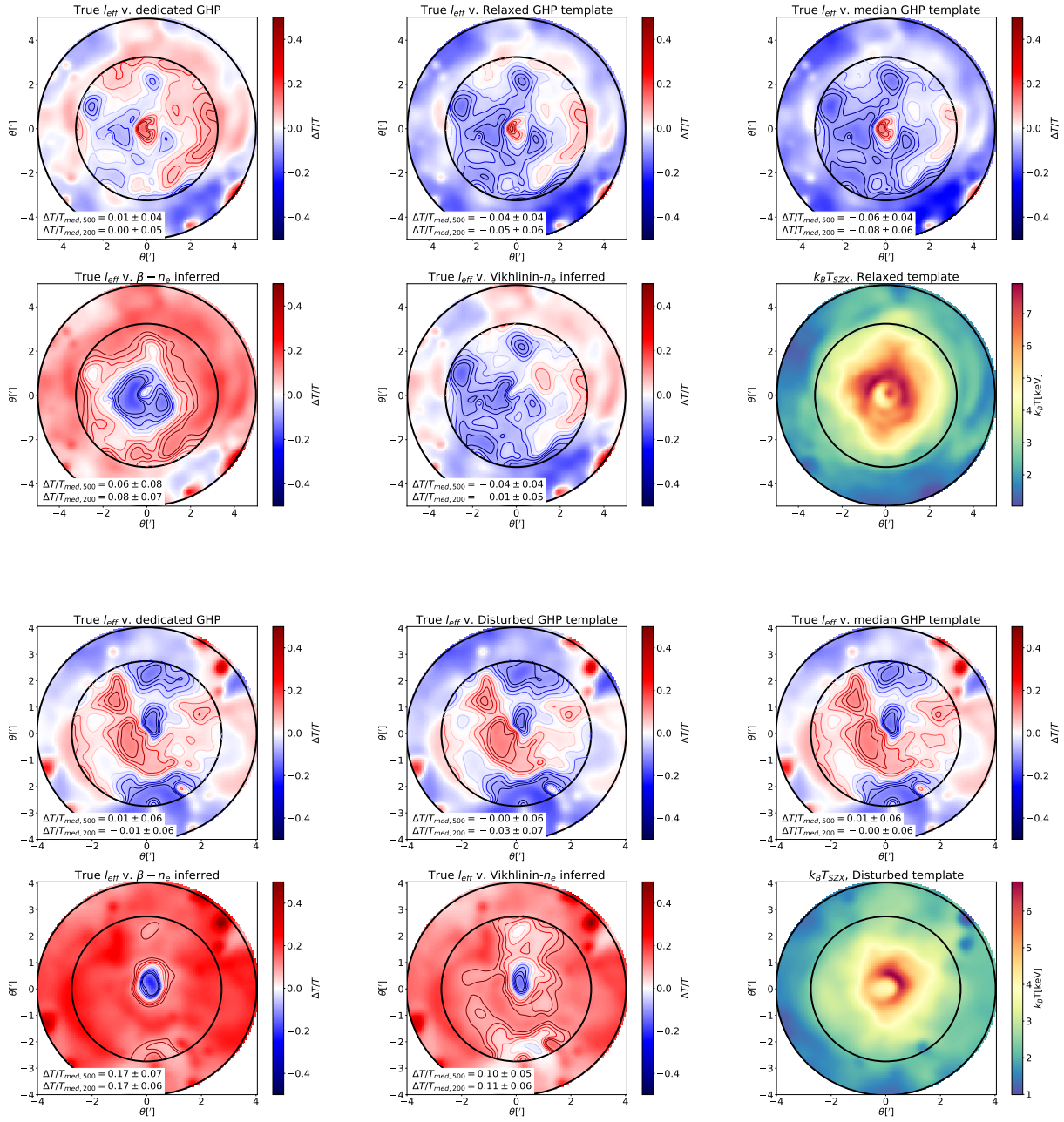

Fig. 2. Two examples of comparison between T_{SZX} derived using the true l_{eff} and T_{SZX} obtained using a modeled effective length. The two rings correspond to R_{500} and R_{200}. The color scale of the comparative maps ranges from -50% to +50%. The levels of the contours inside R_{500} range from -10% to +10%. For each of the two clusters, the final panel shows the T_{SZX} map obtained using the optimal (Relaxed, Disturbed or Median) GHP template.
Top two rows: Main cluster of the region 75 of the snapshot 110, at z =0.490. *Bottom two rows:* Main cluster of the region 135 of the snapshot 107, at z =0.592.

We then compare the T_{SZX} maps reconstructed using parametric maps (i.e. using a GHP template or using density profiles) to the T_{SZX} map obtained using the true l_{eff} map. We estimate the bias and scatter around the "true" T_{SZX} resulting from the use of our models, with the bias $\Delta T/T$ defined as:

$$\Delta T/T = \frac{T_{true} - T_{template}}{T_{true}}, \tag{19}$$

where T_{true} is the temperature map obtained using the true effective length map, while $T_{template}$ is the temperature map obtained with different l_{eff} models.

In the end, the effective length models that we consider here are labeled : "Dedicated GHP", "Optimal GHP", "Vikhlinin n_e", and "$\beta - n_e$". We give the definition of these different models in Table 2, and also compare them with the Median GHP template. We show in Fig.2 the resulting maps for such a comparison in two particular cases (one relaxed cluster and a disturbed one), and in Fig.3 the biases and scatters for all models of l_{eff}, for all clusters.

We show that the effective lengths obtained using GHP templates are statistically better suited to recover the true T_{SZX} that one would infer using the true effective length. Indeed, even if the GHP reconstruction may be biased or very scattered around

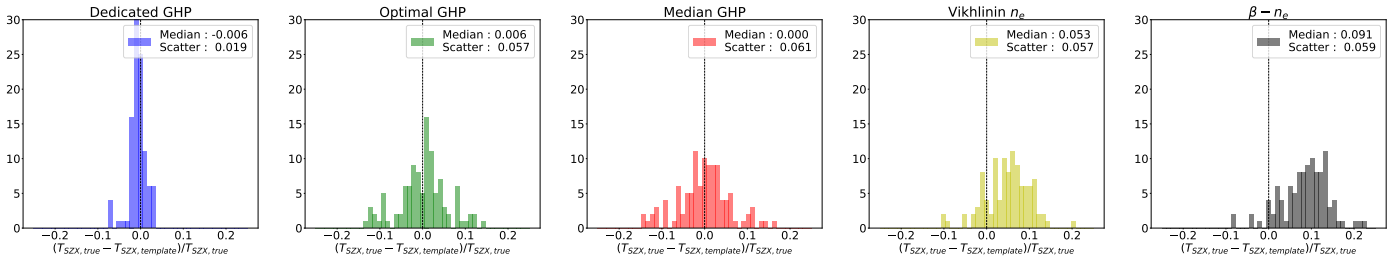

Fig. 3. Recovered biases between $T_{SZX,true}$ and $T_{SZX,template}$ within R_{500}, for all the clusters in the sample, for all different modelings. A more complete version of these statistics, accounting for their evolution with redshift, mass, and temperature, is available in Appendix C.

Table 2. Definition of the different modelings used for the effective length.

Template label	Definition
Dedicated GHP	GHP model fitted directly on the $l_{\rm eff}$ profile of each cluster.
Optimal GHP	Best GHP template for the considered cluster : Relaxed, Disturbed, or Median template.
Vikhlinin n_e	Effective length profiles obtained from 3D Vikhlinin profiles fitted to each cluster.
$\beta - n_e$	Effective length profiles obtained from 3D β profiles fitted to each cluster.

the true T_{SZX} for some particular clusters, and if the density profile $l_{\rm eff}$ can recover T_{SZX} with very low bias and scatter for some clusters (top two rows of Fig. 2), the GHP templates generally outperform their n_e profile counterparts, achieving high accuracy, as reflected in the lower median bias and the comparable scatter, as shown in Fig. 3. Indeed, whereas the median bias observed in the sample is below the percent level for the GHP templates, with a scatter of the order of 6%, the temperatures inferred using a fitted density profile are biased by up to 5% for the Vikhlinin model, and even up to 10% for the β model, with a comparable scatter. It is worth noting that even if the $l_{\rm eff}$ model based on a Vikhlinin fit of the density profile manages a better recovery of T_{SZX} than when using a β profile, it still underperforms with respect to the GHP templates. In particular, the T_{SZX} values obtained using an effective length computed from fitted density profiles seem to be systematically underestimated. This would be a hint of an overestimation of $l_{\rm eff}$ within the chosen apertures of R_{500} and R_{200}. When analyzing the synthetic effective length maps for different models (see Fig. B.1 in Appendix B), we show that the synthetic $l_{\rm eff}$ maps based on density profiles seem to fail to drop to values as low in the center as the GHP templates. In addition the maps obtained using parametric density profiles fail to reproduce the flattening and then decrease of the effective length, but rather continue to increase beyond R_{200}.

This effect, particularly in the central regions, may be a signature of gas clumping that is not well accounted for in the density profiles, but that we are able to capture when starting directly from the $l_{\rm eff}$ maps or profiles, to calibrate the GHP templates. This explains to some extent the apparent overestimation of $l_{\rm eff}$, leading to an underestimated T_{SZX} when using the $l_{\rm eff}$ estimates based on density profiles. In addition, it is worth noting that this behavior does not come with a dependence on the redshift, the mass, or the mean temperature of the cluster, as shown in Fig.3.

We also want to stress here that the density profiles used to implement this approach are 3D density profiles computed in thin radial bins (we used 150 logarithmically spaced radial bins from $0.005\,R_{200}$ to $2R_{200}$). In a more realistic, observational-like setup, where the density profiles are generally obtained using 2D deprojected data computed in fewer radial bins, this effect could be even stronger, highlighting the potentiality of using templates calibrated in simulations.

4. Results

We give in Fig.4 the temperature maps for all the clusters in the combined Twin Samples, ordered in terms of their dynamical state. It appears that we are able to recover a comprehensive view of temperature maps for a great variety of cluster morphologies. In addition, we give in Appendix D the comparative maps between T_{mw} and the T_{SZX} maps obtained with the optimal GHP template, and with the true effective length map. We notably show the presence of several obvious cool cores, that can be seen as paler cavities in the center of the concerned clusters, with cool cores being more represented towards the more relaxed clusters, as could be expected. We can also highlight overheated regions in several of those clusters. Although a deeper analysis would be necessary to confirm their nature with absolute certainty, such regions may be related to shocks. The potential of using the T_{SZX} approach to investigate the presence of shocks in galaxy clusters has incidentally been investigated by Artis et al. (in prep.).

We also want to discuss the particular case of the CL0110 cluster of snapshot 104 (7th row, 6th column of Fig.4). This particular cluster seems to exhibit a very peculiar temperature distribution, strongly mis-centered with respect to the alleged coordinates of the cluster center. After careful investigation of this particular cluster, it appears that this effect arises purely from the T_{SZX} derivation. Indeed, both the effective quantities $(\overline{P_e}, \overline{n_e})$ and the theoretical estimates of the temperature (T_{mw}, T_{sl}) appear to have well centered distributions inside R_{500}. This cluster is a low mass system ($M_{500} = 5.93 \times 10^{13} M_\odot$) that shows a quite peculiar and asymmetric distribution of its gas density, connected to a neighboring halo of similar size. As a result of this complex architecture, the density distribution shows a nonnegligible decrement in the lower left corner of the map, artificially increasing the reconstructed T_{SZX}. This example highlights that in such peculiar systems, the reconstruction of T_{SZX} must be carried out with extreme caution. It may also be a hint that in very low mass clusters, with lower gas densities, the T_{SZX} approach might not be so well suited to easily recover robust and reliable temperature estimates, as we develop later on.

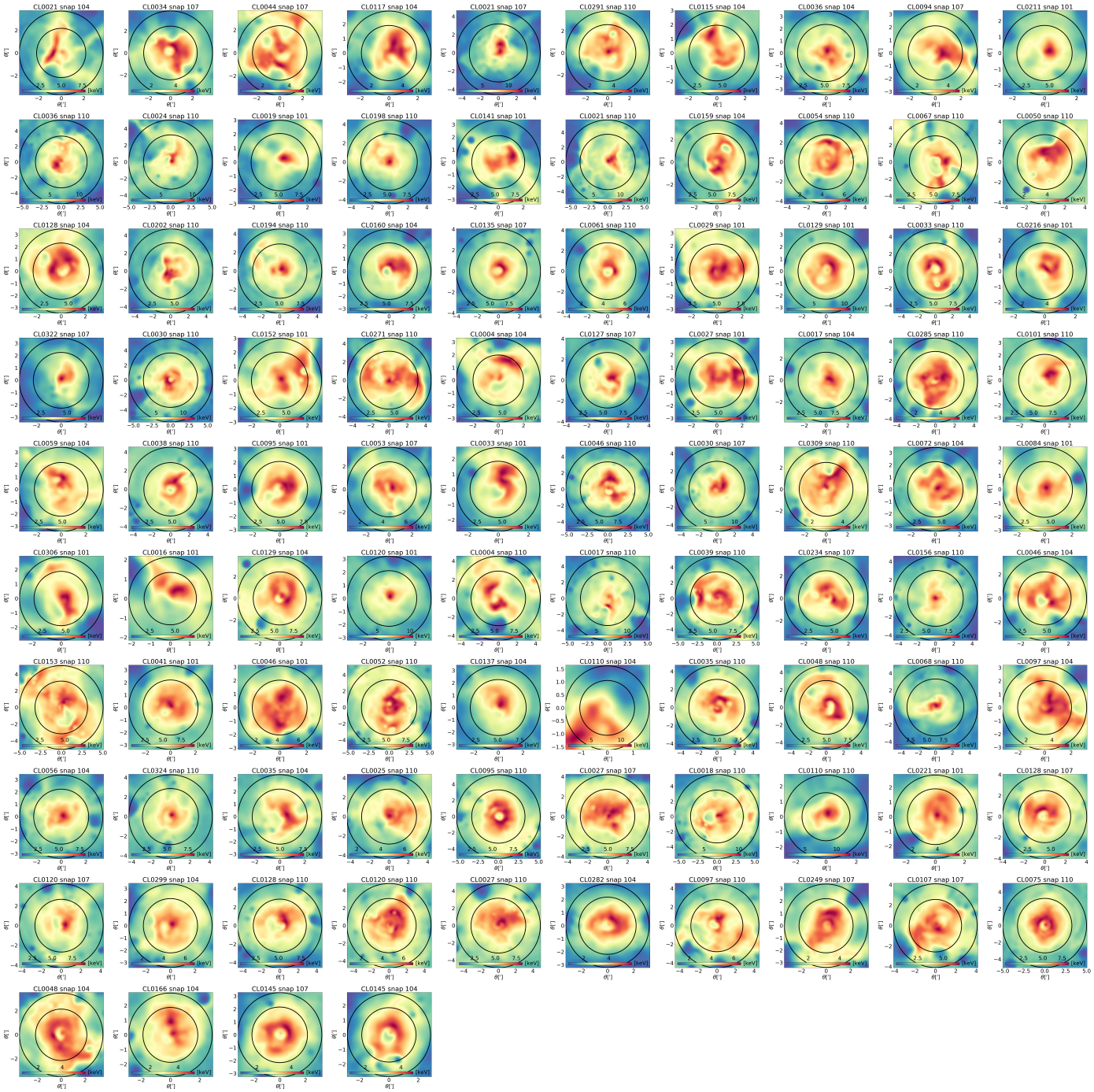

Fig. 4. T_{SZX} maps obtained for all the clusters in our sample, using the optimal GHP template (Relaxed, Disturbed or Median) for each cluster. The maps are ordered by relaxation parameter at R_{500}, χ_{500}. The top left clusters have the lowest chi_{500}, i.e. are the most disturbed clusters, up to the bottom right clusters which are the most relaxed systems.

In Figures 5 and 6 we compare T_{SZX} with the mass-weighted temperature T_{mw} and the spectroscopic-like temperature T_{sl}, respectively. We show that there exists a slight offset between the theoretical estimates (T_{mw}, T_{sl}) and our reconstruction of T_{SZX}. Indeed, T_{SZX} seems to be slightly underestimated with respect to the theoretical estimates of the temperature, of the order on 5-7 % when using GHP models, with a scatter on the order of 10 %. Once again the approach relying on GHP templates outperforms the use of density profiles, for which the recovered temperatures are biased by 13 to 15 %, i.e. twice as much as the GHP models, for an equivalent level of scatter. The complete comparative plots given in Appendix C show that when considering the results cluster by cluster, the dispersion of temperatures around the median bias is significantly larger between T_{sl} and T_{SZX} than between T_{mw} and T_{SZX}. This can be appreciated in comparative maps, of which we give an example in Fig. 7. This is the sign that T_{SZX} is, indeed, a closer approximate of the mass-weighted temperature than of the spectroscopic temperature measure by X-ray observatories, preferentially sensitive to denser regions of the gas distribution. Despite all these effects, nevertheless we show that our reconstruction of T_{SZX} using GHP templates always remains compatible with T_{mw} and T_{sl} within 1σ. We also

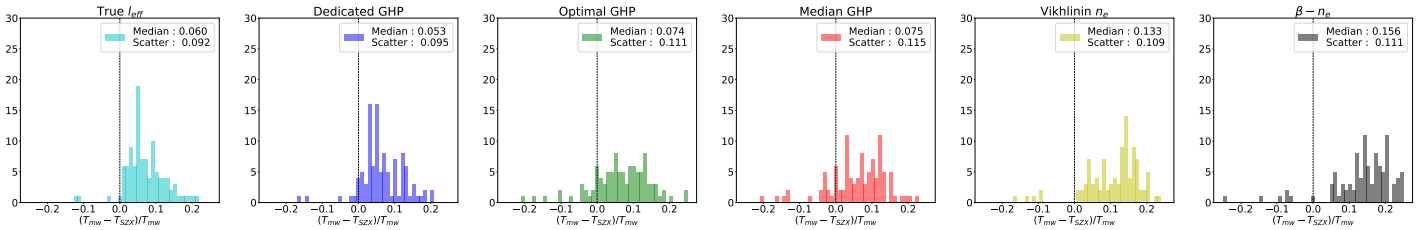

Fig. 5. Recovered biases between T_{SZX} and T_{mw} within R_{500}, for all the clusters in the sample, for all different modelings. A more complete version of these statistics, accounting for their evolution with redshift, mass, and temperature, is available in Appendix C.

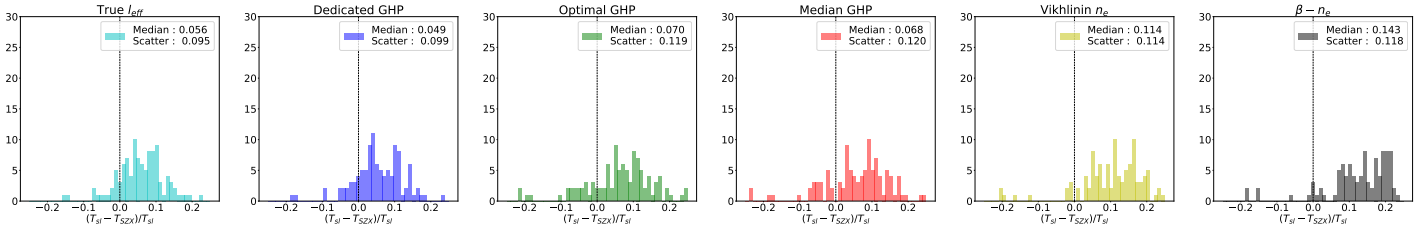

Fig. 6. Recovered biases between T_{SZX} and T_{sl} within R_{500}, for all the clusters in the sample, for all different modelings. A more complete version of these statistics, accounting for their evolution with redshift, mass, and temperature, is available in Appendix C.

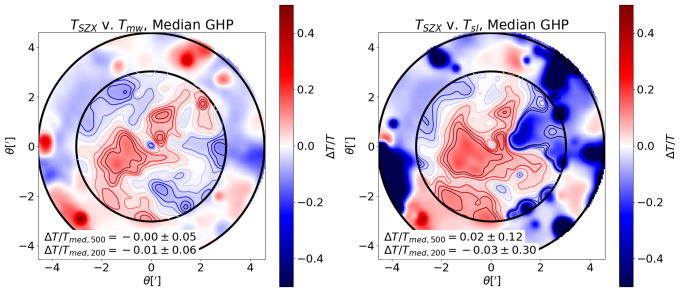

Fig. 7. Comparison of T_{SZX} with the theoretical estimates of the temperature T_{mw} (left) and T_{sl} (right), for the main cluster of region 156 of the snap 110, at z =0.49. The color bar spans the ±50% range, the contours within R_{500} span the ±10% range.

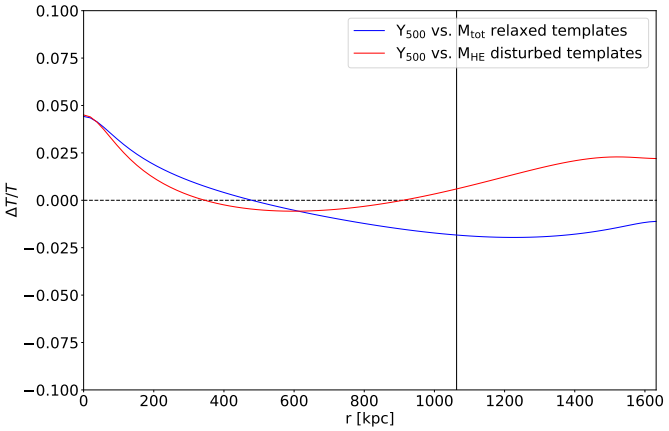

Fig. 8. Radial comparison of the temperature maps obtained using GHP templates calibrated on different Twin Samples, for the main cluster of the region 25 of the snapshot 110, at z = 0.49. We stress that the comparison shown here remains valid for all other clusters at all redshifts in our sample. The blue line shows the comparison of the temperature maps obtained using the relaxed GHP templates calibrated on the Y_{500} and M_{tot} Twin Samples. The red line shows the comparison of the T_{SZX} maps obtained using the disturbed GHP templates calibrated on the Y_{500} and M_{HE} Twin Samples. The vertical black line delimits R_{500}, the radii range up to R_{200}.

find no suggestion of dependencies on mass and temperature (see appendix C), with only a hint of redshift dependence that stems from an artifact of the pixel size. We discuss this effect in Sect. 5.3.

However, we show that for a few isolated clusters, the recovered temperature biases are very large, on the order of 15-20%, beyond the level of scatter, for all choices of l_{eff} modeling. The more complete figures in appendix C show that these clusters are exclusively in the low-mass limit of our sample, including the very low-mass system CL0110 of snapshot 104, discussed earlier in the paper. This is another indication that although in our current mass range we see no evolution of the bias with mass, in the case of lower mass clusters, with lower densities and stronger sensitivity of the gas distribution to astrophysical effects, the reconstruction of T_{SZX} should be carried out with particular caution. In particular, it seems that the capabilities of the T_{SZX} approach seem to start to break down for cluster masses $\lesssim \times 10^{14} M_{\odot}$. This may be entirely driven by our choice of simulations, as the low resolution of the GADGET-X 3k runs does not allow us to fully resolve physical processes in the smallest systems with masses $\lesssim \times 10^{13.5} M_{\odot}$. Although outside the scope of this paper, investigating the potentialities of the T_{SZX} approach in lower mass systems, in different simulations, may be needed to assess whether a "range of validity" in mass exists for the T_{SZX} approach.

5. Discussion

5.1. Calibration of the GHP templates

In this work, we calibrate the GHP templates for l_{eff} in the three combined NIKA2 LPSZ Twin Samples. However, we have *a priori* no reason to be sure that this calibration is valid for different samples. As our end goal is to apply the T_{SZX} method in real observations, using the GHP templates calibrated in this paper, we need to test the adaptability of the templates to different samples. To check for the sample dependency of our template calibration, we perform it separately on the three individual NTS. The parameters for each fit are given in Table 3.

We show that the median profiles (i.e. calibrated on all clusters of the considered sample) of each NTS are all in perfect agreement with each other and in agreement with the median

Table 3. Parameter values for the GHP template calibrations in the different Twin Samples.

	A_0	B_0	α	β	γ
Combined NTS					
Relaxed Template	$4.012^{+1.004}_{-0.346}$	$0.534^{+0.088}_{-0.168}$	$0.540^{+0.013}_{-0.011}$	$5.037^{+3.599}_{-1.692}$	$1.083^{+1.130}_{-0.293}$
Median Template	$3.973^{+1.122}_{-0.493}$	0.463 ± 0.156	$0.501^{+0.038}_{-0.031}$	$5.531^{+5.435}_{-2.542}$	$1.418^{+1.920}_{-0.659}$
Disturbed Template	$3.437^{+0.438}_{-0.154}$	$0.500^{+0.058}_{-0.112}$	$0.447^{+0.010}_{-0.009}$	$8.508^{+4.170}_{-3.305}$	$1.037^{+0.981}_{-0.257}$
M_{tot} NTS					
Relaxed Template	$5.168^{+1.028}_{-1.075}$	$0.326^{+0.176}_{-0.092}$	0.529 ± 0.009	$3.280^{+1.263}_{-0.584}$	$2.062^{+1.881}_{-1.134}$
Median Template	$4.010^{+1.189}_{-0.528}$	0.452 ± 0.157	$0.507^{+0.034}_{-0.028}$	$5.274^{+5.471}_{-2.346}$	$1.400^{+1.964}_{-0.661}$
Disturbed Template	$3.838^{+0.478}_{-0.339}$	$0.392^{+0.088}_{-0.083}$	$0.460^{+0.010}_{-0.009}$	$6.821^{+4.091}_{-2.039}$	$1.642^{+1.901}_{-0.784}$
M_{HE} NTS					
Relaxed Template	$3.660^{+0.177}_{-0.110}$	$0.578^{+0.039}_{-0.049}$	0.544 ± 0.008	$10.433^{+3.055}_{-3.288}$	$0.825^{+0.190}_{-0.111}$
Median Template	$3.882^{+1.060}_{-0.484}$	$0.475^{+0.167}_{-0.159}$	$0.506^{+0.044}_{-0.038}$	$5.715^{+5.351}_{-2.628}$	$1.370^{+1.934}_{-0.631}$
Disturbed Template	$3.139^{+0.068}_{-0.065}$	$0.659^{+0.034}_{-0.033}$	0.433 ± 0.010	$12.768^{+1.605}_{-2.535}$	$0.791^{+0.093}_{-0.081}$
Y_{500} NTS					
Relaxed Template	$4.156^{+0.656}_{-0.320}$	$0.551^{+0.073}_{-0.111}$	$0.573^{+0.013}_{-0.011}$	$5.028^{+2.047}_{-1.301}$	$1.383^{+0.758}_{-0.310}$
Median Template	$4.060^{+1.280}_{-0.582}$	$0.466^{+0.179}_{-0.167}$	$0.506^{+0.034}_{-0.027}$	$4.847^{+5.388}_{-2.160}$	$1.498^{+1.984}_{-0.721}$
Disturbed Template	$6.674^{+1.736}_{-1.399}$	$0.180^{+0.089}_{-0.047}$	$0.495^{+0.026}_{-0.023}$	$1.698^{+0.403}_{-0.281}$	$3.211^{+1.170}_{-1.229}$

Notes. The parameter values are given with the 16th and 84th percentiles from our posterior distributions.

template of the combined samples. However, this agreement is not so large between the relaxed and disturbed templates for the three different Twin Samples. In particular, the relaxed templates of the M_{tot} and Y_{500} NTS and the disturbed templates of the Y_{500} and M_{HE} NTS show the largest discrepancies, with some parameters disagreeing at more than 2σ. These differences in the relaxed and disturbed templates, while the median templates are in perfect agreement, may stem from the selection of the Twin Samples, which has been performed solely on mass (or Compton parameter) and redshift, with no focus on the dynamical state.

This suggests that generalization of the relaxed and disturbed templates to other samples should not be so trivial. This result needs to be taken into account when moving to the application on observed data. We thus try to quantify the effect of these discrepancies on the final recovered temperature maps. To that end, we compare the temperature maps obtained using the two extreme relaxed templates (Y_{500} and M_{tot}) and disturbed templates (Y_{500} and M_{HE}). We show in Fig.8 this comparison for one of the clusters of our sample. As expected, the relative difference between the temperature maps using the two extreme templates depends on the radius. It appears that the maximum difference between the two maps is in the very center of the cluster, with a sharp increase in the innermost regions. However, the relative difference remains quite negligible, always on the order of $\sim 1\%$, increasing to slightly more than 3% in the innermost regions. This effect is due to the degeneracies between some of the parameters of the GHP model, which allows the recovery of similar shapes of the l_{eff} profile for different sets of parameters. This behavior, which we show for one cluster, has been observed in all the clusters of the sample. Compared to the median bias and scatter between T_{SZX} and T_{mw} in our sample, which is on the order of 10%, this contribution from the choice of calibration for the GHP templates thus appears to be negligible.

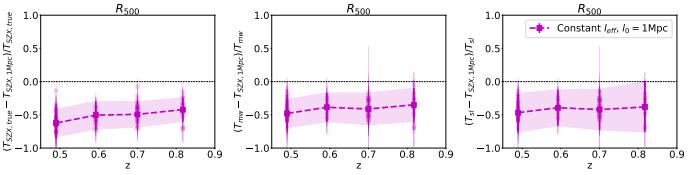

Fig. 9. Comparison between T_{SZX} obtained using an effective length l_0 = 1Mpc and T_{SZX} obtained using the true l_{eff} map (left), T_{mw} (center), and T_{sl} (right). The biases and scatter are computed within R_{500}.

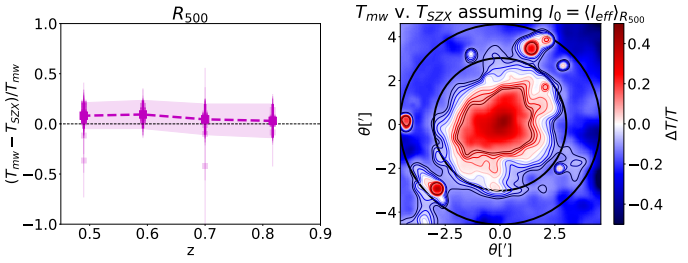

Fig. 10. Comparison between T_{SZX} obtained using an effective length set to the mean of the true effective length at R_{500} and T_{mw} for the full sample (left), and in the particular case of one cluster, the main cluster of region 156 of snapshot 110 at $z = 0.49$ (right). The biases and scatter are computed within R_{500}.

5.2. A word about the modeling of l_{eff}

Throughout this work, we assumed a variety of different models for the effective length, either relying on fitting a density profile, propagating it to a map of l_{eff}, or using template maps calibrated on simulations, obtained with the GHP profile of l_{eff}. Past works (A17, Artis et al. in prep.), however, have sometimes used a simpler modeling for l_{eff}, assuming that the quantity is a constant

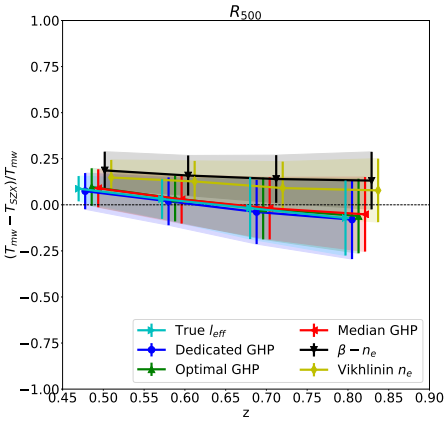

Fig. 11. Comparison between T_{SZX} and T_{mw} for the clusters of snapshot 110, within R_{500}, displaced at redshifts $z = 0.49$ (original redshift), $z = 0.592$ (matching snapshot 107), $z = 0.7$ (matching snapshot 104) and $z = 0.817$ (matching snapshot 101). The points have been slightly shifted on the x axis for clarity.

$l_0 \sim 1$Mpc ($l_0 = 1.4$Mpc in A17, based on estimates of Sayers et al. (2013) and Mroczkowski et al. (2012)). Here we test the effect of this assumption on the recovery of temperature maps.

We show in Fig.9 that assuming a constant $l_0 = 1$Mpc leads to a bias on the order of $\sim 50\%$, with a high scatter on the order of 25%. Of course, one may argue that setting an effective length at 1Mpc when such a low value is barely reached in most of our $l_{\rm eff}$ maps is much too coarse an approximation. In order to use a more realistic modeling for a constant $l_{\rm eff}$, we test setting the value of $l_{\rm eff}$ at the mean value of the effective length within a given aperture (e.g. R_{500}), instead of 1 Mpc. Such a value is not a direct observable; however, it could be derived, for instance, from a density profile. We show in Fig.10 that in this case the bias between different reconstructions of temperatures is very low. However, the scatter is still quite high, around $\sim 15\%$. The most concerning issue with this modeling is that we show that the bias depends on the radius, with little of the global dynamics of the map being recovered properly (right panel of Fig.10).

This highlights that the simple modeling choice of assuming a constant effective length of ~ 1Mpc should *not* be considered when trying to infer temperature maps, whether to study the structure or the normalization of the maps. However, it may be enough to highlight overheated regions in maps, similar to the approach of Artis et al. (in prep.). Even when considering the mean value of the effective length, the goal of a study using this modeling should be clearly defined. If the purpose is to simply infer the normalization of a map, for instance, to estimate integrated temperature values inside a given aperture, then assuming a constant effective length set at its mean value can be sufficient. However, when trying to study in addition the features and spatial distribution of the gas temperature, a more complex modeling for the effective length *must* be considered, as assuming a constant $l_{\rm eff}$ will inevitably induce radially dependent biases in the final reconstruction.

5.3. Impact of the resolution

We mentioned in Sect. 4 the possible existence of a mild trend with redshift of the bias between T_{SZX} and the theoretical estimates of temperature T_{mw} and T_{sl}. We explore here the origin of this observed trend and suggest that this behavior is not a physical dependence on redshift, but rather purely due to the resolution and the sampling of the maps. In our current setup, all maps have an angular pixel size of 3", and are smoothed with a Gaussian kernel of 18" FWHM. For a cluster with $R_{500} = 1000$kpc at $z = 0.49$ (the lowest redshift of our simulated sample), ~ 9020 pixels will cover the area delimited by R_{500}, and the area delimited by the FWHM of the PSF will cover 0.3% of the said area. For the same cluster taken at $z = 0.817$ (the highest redshift available in our case), the area πR^2_{500} will be covered by ~ 5770 pixels and the area delimited by the FWHM of the PSF will cover 0.5% of it. Such a change is nonnegligible and may impact the final comparison between the reconstructed T_{SZX} and the theoretical (T_{mw}, T_{sl}).

In order to investigate the impact of this effect, we chose to focus on the clusters at the lowest redshift of our sample, in snapshot 110, and pass them through our pipeline while displacing them at higher redshifts. We stress here that we are always considering the same clusters and simply gradually reducing the sampling and resolution by artificially modifying their redshifts (thus modifying their angular diameter distance) in the pipeline, rather than actually looking at the same halos but taken at different steps of their evolution. To remain comparable with the rest of the study, we chose to set the clusters at the same four redshifts as in the main analysis, i.e. $z = 0.49$ (snapshot 110, original redshift of the clusters), $z = 0.592$ (snapshot 107), $z = 0.7$ (snapshot 104), and $z = 0.817$ (snapshot 101, highest redshift available in the sample).

We show that by taking the same clusters and gradually lowering their sampling and resolution by mimicking the placement at different redshifts, we are indeed able to replicate a negative bias trend between T_{SZX} and T_{mw} with redshift. The T_{SZX} maps obtained using density profiles seem to be slightly less sensitive to this effect than the T_{SZX} maps obtained using GHP templates or using the true $l_{\rm eff}$ map. As the former $l_{\rm eff}$ maps are "simpler" than the latter ones, with fewer distinctive features, it is possible that the gradual degradation of the maps induces a slightly lower loss of information. We thus claim the slight redshift trend with z of the bias between T_{SZX} and (T_{mw}, T_{sl}) that we highlighted in Sect. 4 is not a physical effect, but rather is mainly triggered by different sampling and resolution at different redshifts.

6. Conclusions

Mapping of the galaxy cluster gas temperature, although critical for the characterization of the ICM and cluster mass estimates, remains a challenge. Few previous works have highlighted the potentiality of using a combination of X-ray and tSZ data to derive temperature profiles, and fewer still have tried to apply this method to map the temperature in 2D. However, this approach focused on 2D maps had yet to be validated and tested on a large sample. In this work, we prepare the systematic application of this method to real data, by validating the theoretical aspects of the T_{SZX} reconstruction inside The300 suite of hydrodynamical simulations. We start from maps of the theoretical y_{tSZ} signal and of the gas emission measure, serving as proxies for the pressure and squared density maps, respectively.

In addition, we have to take into account the effective length of the cluster, allowing us to convert the map of the integrated squared density obtained from the emission measure to a map of the integrated density. As this quantity is not a direct observable, we proposed a model, dubbed generalized high pass (GHP). We calibrated our model on simulated maps of the true effective length, deriving three different template parameterizations calibrated on relaxed clusters, on disturbed ones, and on all clusters to be applied as a median template. We compared the T_{SZX} obtained using our modeling with other T_{SZX} maps obtained using

more standard modelings, i.e. assuming either a constant value or parametric density profiles to infer the effective length. We show that using the GHP model of $l_{\rm eff}$ for the reconstruction of T_{SZX} systematically outperforms other more standard models of the effective length, with a lower bias and scatter around the mean temperature value. In particular, we highlight how modeling the effective length with a single constant value should only be done with extreme caution with respect to the goals of the analysis, if at all. Indeed, we show that carelessly choosing a constant value for $l_{\rm eff}$ may lead to extremely high biases and scatter in the recovered temperature, and that even with a carefully chosen constant value, a nonnegligible radial bias remained in the T_{SZX} reconstruction.

Comparing our T_{SZX} reconstructions with theoretical estimates of the temperature, we find a good agreement with the mass-weighted and spectroscopic-like temperatures T_{mw} and T_{sl}. We find a median bias of at most on the order of 10% when using GHP templates, with a scatter on the same order of magnitude, with no hint of evolution with the mass or the mean gas temperature. We also highlight a slight hint of a redshift dependence of the bias between T_{SZX} and the theoretical temperatures. However, this is fully driven by the sampling of the maps at different redshifts, with higher redshift clusters having a lower number of pixels within R_{500} and the PSF covering a larger fraction of the radius.

This present work is the first of two papers in which the global aim is to validate the T_{SZX} approach and apply it to high resolution tSZ maps, the NIKA2 LPSZ sample. The next paper will focus on the application to the said data, using our own modeling of the effective length, and will discuss the impact of the instrumental effects on the final temperature maps. The effects to be discussed in particular will be that of the noise and the tSZ transfer function. Indeed, due to the scanning strategy of the tSZ data, large scales will be filtered out. This needs to be properly modeled and understood to avoid introducing biases, notably radially dependent ones, in the final temperature maps. On the basis of the recovered temperature maps, we will investigate the properties of the gas distribution, including highlighting potential shocks and cool cores. The comparison of our approach on one LPSZ cluster, PSZ2G091.83+26.11, already analysed in the work of Artis et al., should be interesting. They derived the temperature map for one of the LPSZ clusters, using a simpler modeling for the effective length, and used it to highlight an overheated region of the cluster, trying to link it to a shock.

Acknowledgements. The authors thank the anonymous referee for their useful comments and remarks. The authors acknowledge membership to The Three Hundred collaboration, thanks to which this work was made possible. The simulations used in this paper have been performed in the MareNostrum Supercomputer at the Barcelona Supercomputing Center, thanks to CPU time granted by the Red Española de Supercomputación. As part of The Three Hundred project, this work has received financial support from the European Union's Horizon 2020 Research and Innovation programme under the Marie Skłodowska-Curie grant agreement number 734374, the LACEGAL project. R.W., M.D.P. and A.F. acknowledge financial support from PRIN-MUR grant 20228B938N *"Mass and selection biases of galaxy clusters: a multi-probe approach"* funded by the European Union Next generation EU, Mission 4 Component 2 CUP B53D23004790006. R. Adam acknowledges support from the Programme National Cosmologie et Galaxies (PNCG) of CNRS/INSU with INP and IN2P3, co-funded by CEA and CNES. R. Adam was supported by the French government through the France 2030 investment plan managed by the National Research Agency (ANR), as part of the Initiative of Excellence of Université Côte d'Azur under reference number ANR-15-IDEX-01. GY and WC would like to thank Ministerio de Ciencia e Innovación (MICINN) for financial support under the project grant PID2021-122603NB-C21. WC is further supported by Atracción de Talento fellowship no. 2020-T1/TIC19882 granted by the Comunidad de Madrid and by the Consolidación Investigadora no. CNS2024-154838 granted by the Agencia Estatal de Investigación (AEI) in Spain, ERC: HORIZON-TMA-MSCA-SE for supporting the LACEGAL-III (Latin American Chinese European Galaxy Formation Network) project with grant number 101086388. M.M.E. acknowledges the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-22-CE31-0010.

References

Adam, R., Arnaud, M., Bartalucci, I., et al. 2017, A&A, 606, A64
Adam, R., Comis, B., Bartalucci, I., et al. 2016, A& A, 586, A122
Adam, R., Comis, B., Macías-Pérez, J. F., et al. 2015, A&A, 576, A12
Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, ARA&A, 49, 409
Ameglio, S., Borgani, S., Pierpaoli, E., & Dolag, K. 2007, MNRAS, 382, 397
Ameglio, S., Borgani, S., Pierpaoli, E., et al. 2009, MNRAS, 394, 479
Ansarifard, S., Rasia, E., Biffi, V., et al. 2020, A&A, 634, A113
Battaglia, N., Bond, J. R., Pfrommer, C., & Sievers, J. L. 2012, ApJ, 758, 74
Biffi, V., Borgani, S., Murante, G., et al. 2016, ApJ, 827, 112
Biffi, V., Sembolini, F., De Petris, M., et al. 2014, MNRAS, 439, 588
Cavaliere, A. & Fusco-Femiano, R. 1976, A&A, 49, 137
Chluba, J., Nagai, D., Sazonov, S., & Nelson, K. 2012, MNRAS, 426, 510
Chluba, J., Switzer, E., Nelson, K., & Nagai, D. 2013, MNRAS, 430, 3054
Cui, W., Knebe, A., Yepes, G., et al. 2018, MNRAS, 480, 2898
De Luca, F., De Petris, M., Yepes, G., et al. 2021, MNRAS, 504, 5383
Dolag, K., Hansen, F. K., Roncarelli, M., & Moscardini, L. 2005, MNRAS, 363, 29
Ghirardini, V., Ettori, S., Eckert, D., et al. 2018, A&A, 614, A7
Gianfagna, G., De Petris, M., Yepes, G., et al. 2021, MNRAS, 502, 5115
Gianfagna, G., Rasia, E., Cui, W., et al. 2023, MNRAS, 518, 4238
Haggar, R., Gray, M. E., Pearce, F. R., et al. 2020, MNRAS, 492, 6074
Hahn, O., Martizzi, D., Wu, H.-Y., et al. 2017, MNRAS, 470, 166
Hasselfield, M., Hilton, M., Marriage, T. A., et al. 2013, JCAP, 2013, 008
Henden, N. A., Puchwein, E., Shen, S., & Sijacki, D. 2018, MNRAS, 479, 5385
Henden, N. A., Puchwein, E., & Sijacki, D. 2019, MNRAS, 489, 2439
Lau, E. T., Kravtsov, A. V., & Nagai, D. 2009, ApJ, 705, 1129
Markevitch, M. & Vikhlinin, A. 2007, PhR, 443, 1
Mayet, F. et al. 2020, EPJ Web Conf., 228, 00017
Mazzotta, P., Rasia, E., Moscardini, L., & Tormen, G. 2004, MNRAS, 354, 10
Migkas, K., Kox, D., Schellenberger, G., et al. 2024, A&A, 688, A107
Mroczkowski, T., Bonamente, M., Carlstrom, J. E., et al. 2009, ApJ, 694, 1034
Mroczkowski, T., Dicker, S., Sayers, J., et al. 2012, ApJ, 761, 47
Muñoz-Echeverría, M., Macías-Pérez, J. F., Pratt, G. W., et al. 2024, A&A, 682, A147
Nelson, K., Lau, E. T., & Nagai, D. 2014, ApJ, 792, 25
Paliwal, A., Artis, E., Cui, W., et al. 2022, in EPJ Web of Conferences, Vol. 257, mm Universe @ NIKA2 - Observing the mm Universe with the NIKA2 Camera, 00036
Pellissier, A., Hahn, O., & Ferrari, C. 2023, MNRAS, 522, 721
Perotto, L., Adam, R., Ade, P., et al. 2022, in EPJ Web of Conferences, Vol. 257, mm Universe @ NIKA2 - Observing the mm Universe with the NIKA2 Camera (EDP), 00038
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016a, A&A, 594, A27
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, A&A, 594, A13
Pratt, G. W., Arnaud, M., Biviano, A., et al. 2019, Space Sci. Rev., 215, 25
Price, D. J. 2007, PASA, 24, 159
Rasia, E., Ettori, S., Moscardini, L., et al. 2006, MNRAS, 369, 2013
Rasia, E., Meneghetti, M., Martino, R., et al. 2012, New J. Phys., 14, 055018
Ruppin, F., Adam, R., Comis, B., et al. 2017, A&A, 597, A110
Sarazin, C. L. 1988, X-ray emission from clusters of galaxies
Sayers, J., Mroczkowski, T., Zemcov, M., et al. 2013, ApJ, 778, 52
Schellenberger, G., Reiprich, T. H., Lovisari, L., Nevalainen, J., & David, L. 2015, A&A, 575, A30
Shi, X., Komatsu, E., Nelson, K., & Nagai, D. 2015, MNRAS, 448, 1020
Sunyaev, R. A. & Zeldovich, Y. B. 1972, Comments Astrophys. Space Phys., 4, 173
Vazza, F., Brunetti, G., Kritsuk, A., et al. 2009, A&A, 504, 33
Veronesi, I., Bartalucci, I., Rasia, E., et al. 2025, A&A, 694, A107
Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, ApJ, 640, 691
Voit, G. M. 2005, Reviews of Modern Physics, 77, 207
White, S. D. M., Navarro, J. F., Evrard, A. E., & Frenk, C. S. 1993, Nature, 366, 429
Wicker, R., Douspis, M., Salvati, L., & Aghanim, N. 2023, A&A, 674, A48
ZuHone, J., Bahar, Y. E., Biffi, V., et al. 2023, A&A, 675, A150

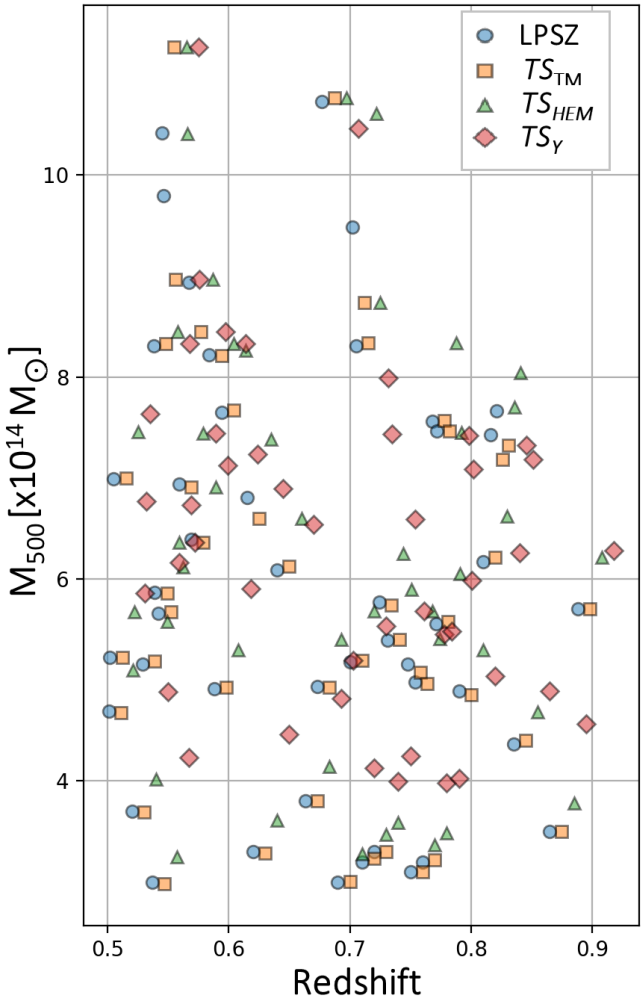

Fig. A.1. Distribution of the NIKA2 LPSZ Twin Sample in the mass-redshift plane, figure taken from (Paliwal et al. 2022). The blue circles represent the true LPSZ clusters, while the other markers correspond to their simulated twins, within the three different Twin Samples. The orange squares correspond to the total mass twin sample, the green triangles to the hydrostatic mass twin sample, and the red diamonds to the Y_{500} twin sample. The redshifts of the simulated twins have been artificially set to the redshifts of their real-life counterparts.

Appendix A: The NIKA2 LPSZ Twin Samples

We give in Fig. A.1 the mass-redshift distribution of the NIKA 2 LPSZ Twin Sample simulated clusters, compared to the distribution of their real-life LPSZ counterparts.

Appendix B: GHP effective length maps compared to n_e-inferred ones

We show here the shape comparison between different effective length templates, obtained using the GHP model (defined in Eq. 15) and using the Vikhlinin model (defined in Eq. 18). As the Vikhlinin model gives the most precise estimation of $l_{\rm eff}$ out of the two models based on the fitting of the density profile (Sect. 3.3), we do not consider the β model here.

For each cluster the GHP effective length map is obtained following one of the three calibrations of the GHP model (Relaxed, Disturbed or Median), while the Vikhlinin effective length is obtained by directly fitting the density profile for each cluster.

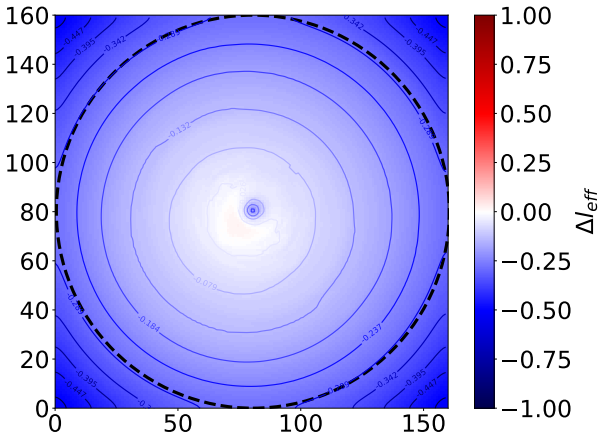

Fig. B.1. Stacked comparative map of the GHP-inferred and Vikhlinin-inferred effective length maps. The dashed line represents the spatial extension of the map, covering R_{200} for each cluster. The contours span the range ±50% while the colorbar spans the range ±100%.

We show below the stacking S of all the bias maps between the GHP and the Vikhlinin-inferred effective length maps, i.e.

$$S = \mathrm{med}\left(\frac{l_{eff,GHP} - l_{eff,Vik}}{l_{eff,GHP}}\right). \quad \text{(B.1)}$$

We show that on average, the effective lengths recovered using the Vikhlinin model are higher than those obtained using the GHP model. This effect is particularly strong in the outskirts of the clusters at radii $\gtrsim R_{200}$, as the Vikhlinin model cannot replicate the flattening of $l_{\rm eff}$, and in the innermost regions, with a strong peak towards low $l_{\rm eff}$ values in the GHP model that cannot be reproduced with the Vikhlinin model. The Vikhlinin and GHP effective lengths show agreement within 10% only in the inner regions of the cluster below $\sim 0.5R_{200}$, excluding the central peak. As discussed in the bulk of the paper, this effect is a signature of the gas clumping, poorly recovered when fitting 1D density profiles. This also explains the underestimated T_{SZX} values that we obtained when using $l_{\rm eff}$ maps derived from density profiles (see the discussion in Sect. 3.3).

Appendix C: Redshift, mass, and temperature dependence of the reconstruction biases

We give here in Figures C.1, C.2, and C.3 the more complete comparisons between different T_{SZX} models, as well as their comparison with T_{mw} and T_{sl}, showing their trends with redshift, mass and temperature. The error bars displayed in these figures are computed as the standard deviation of the bias values around the mean bias, within R_{500}. For the sake of clarity, the data points in the redshift panels have been shifted along the x axis.

Appendix D: Comparison between T_{SZX} and T_{mw}, for all clusters in the NIKA2 LPSZ Twin Samples

We show here the comparison between the mass-weighted temperature T_{mw} and T_{SZX}, assuming the optimal GHP template (Fig. D.1) and using the true effective length map (Fig. D.2). We show that using a GHP template, though on average less accurate than using directly the true effective length map, manages to reproduce competitive results.

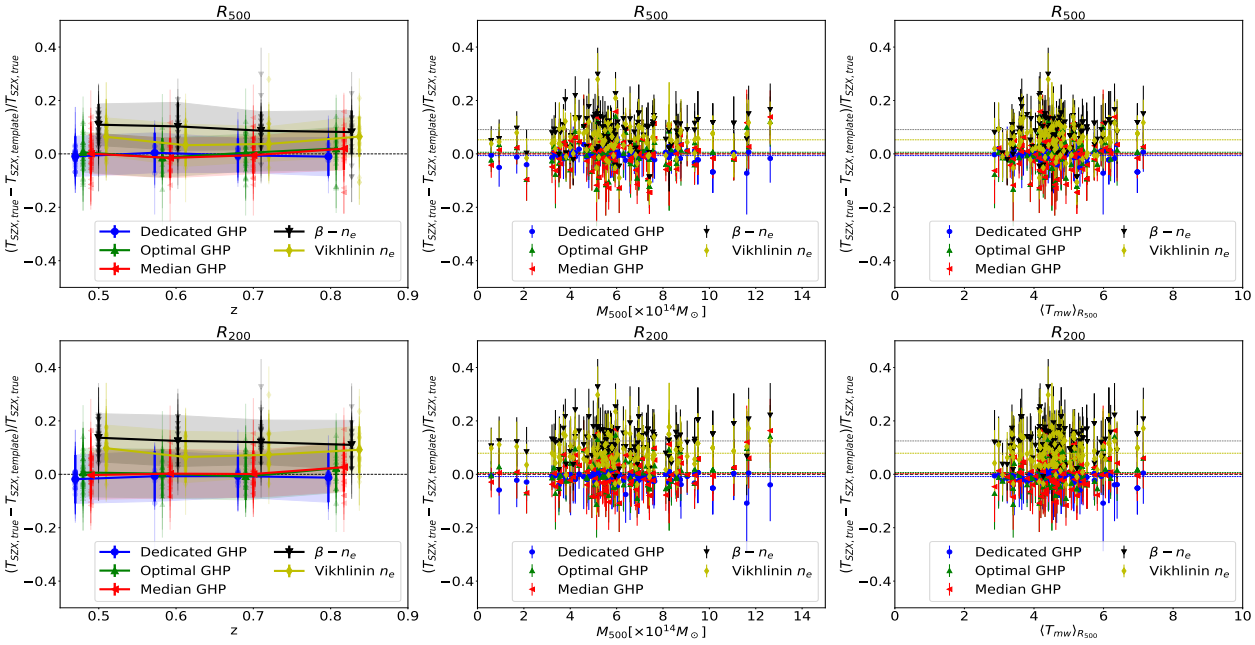

Fig. C.1. Median bias and scatter between $T_{SZX,true}$ and $T_{SZX,template}$, for all the clusters in the sample, depending on z (left column), M_{500} (middle column) and $\langle T_{mw} \rangle_{R_{500}}$ (right column). The coloured dashed lines in the mass and temperature panels correspond to the median biases for each model. *Top row:* Biases and scatters taken at R_{500}. *Bottom row:* Biases and scatters taken at R_{200}.

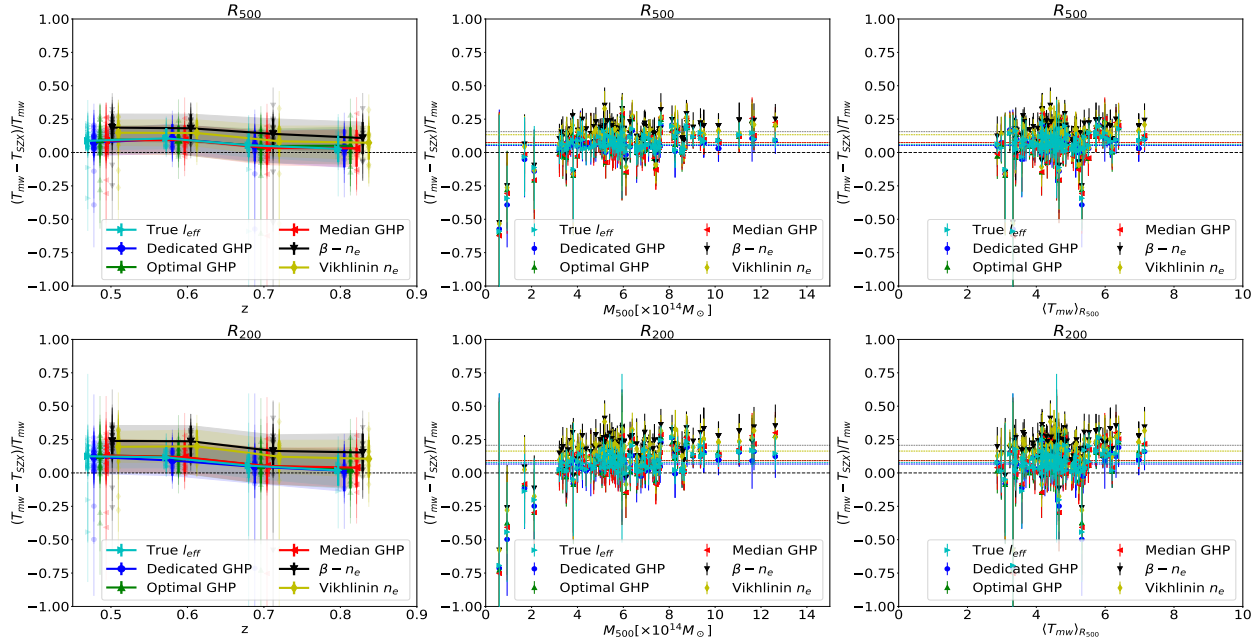

Fig. C.2. Comparison between T_{SZX}, obtained using different models, and T_{mw} for all the clusters in our sample, depending on their redshift, mass, and mean temperature at R_{500}. The coloured dashed lines in the mass and temperature panels correspond to the median biases for each model. *Top row :* Biases and scatters measured at R_{500}. *Bottom row :* Biases and scatters measured at R_{200}.

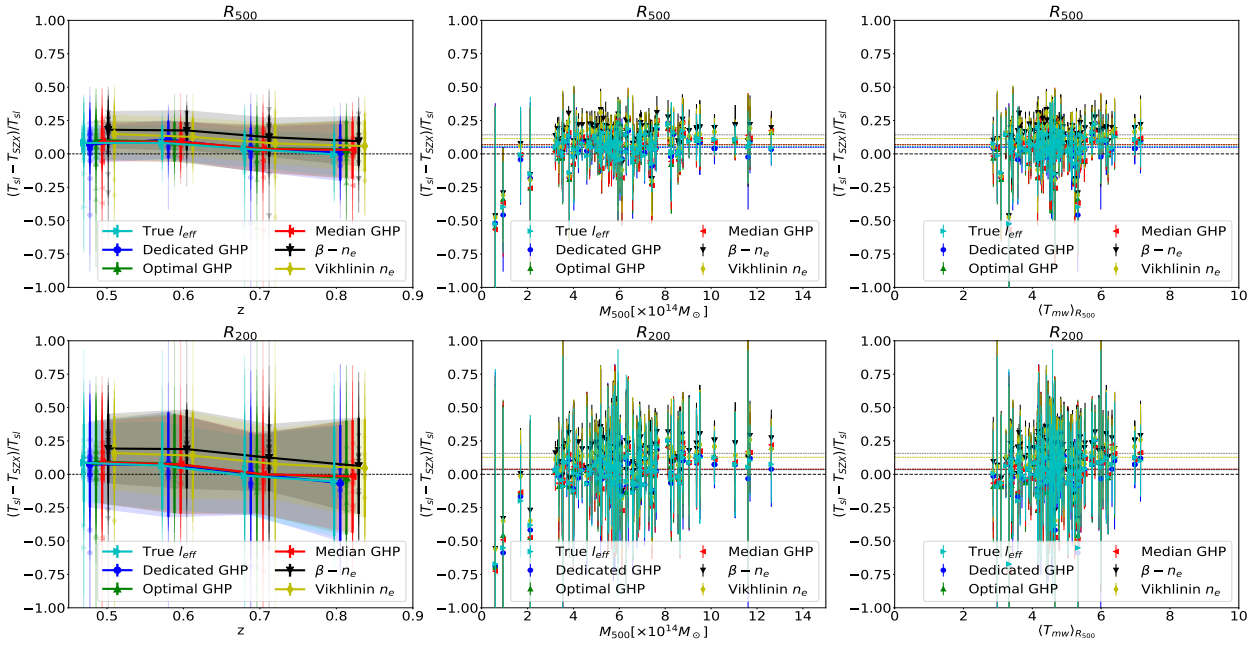

Fig. C.3. Comparison between T_{SZX}, obtained using different models, and T_{sl} for all the clusters in our sample, depending on their redshift, mass, and mean temperature at R_{500}. The coloured dashed lines in the mass and temperature panels correspond to the median biases for each model. *Top row :* Biases and scatters measured at R_{500}. *Bottom row :* Biases and scatters measured at R_{200}.

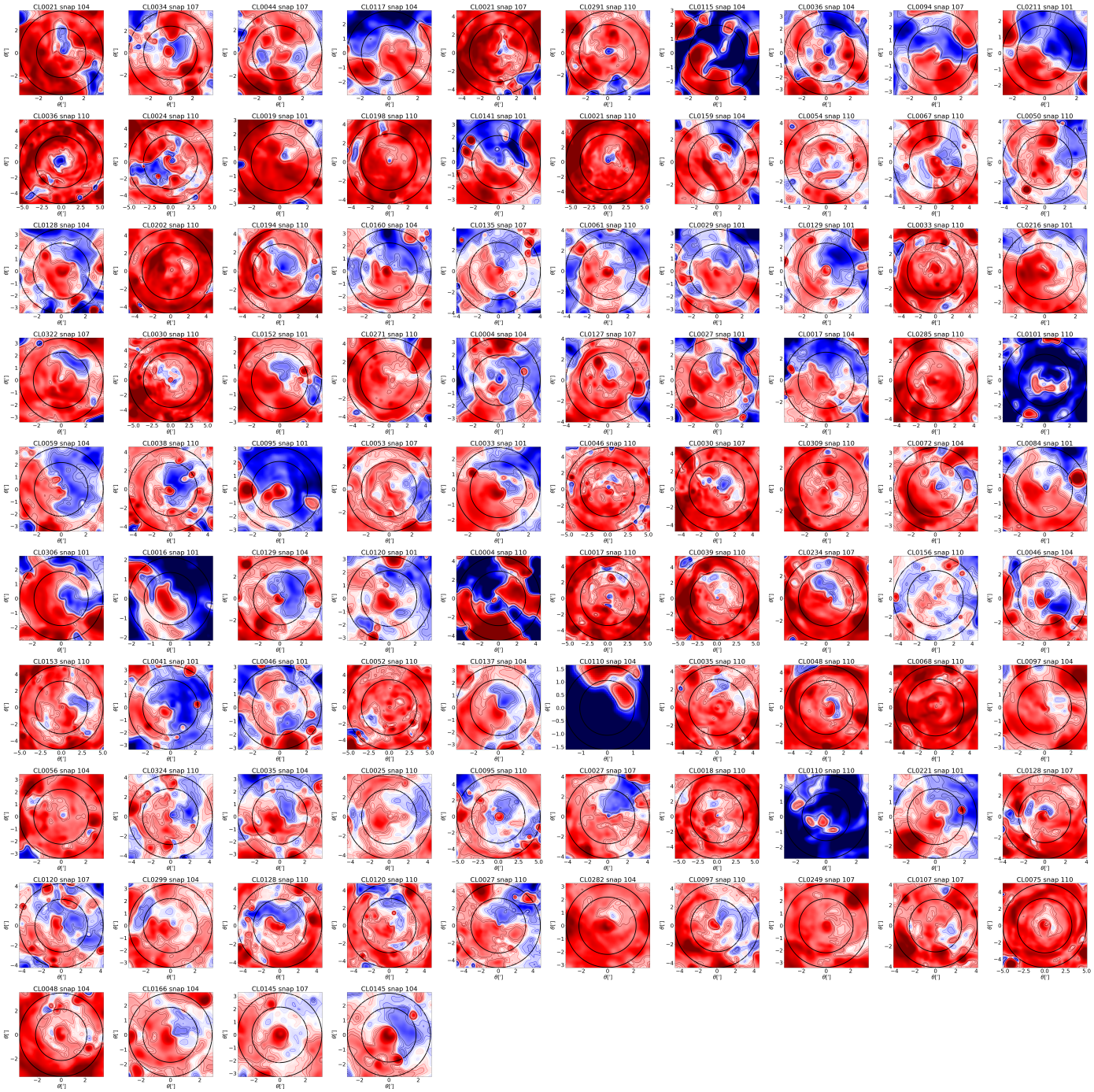

Fig. D.1. Comparison between T_{SZX} and T_{mw} maps obtained for all the clusters in our sample, using the optimal GHP template (Relaxed, Disturbed or Median) for each cluster. The maps are ordered by relaxation parameter at R_{500}, χ_{500}. The top left clusters have the lowest chi_{500}, i.e. are the most disturbed clusters, up to the bottom right clusters which are the most relaxed systems. The colormaps span the range ±50%, while the contours span the range ±10%.

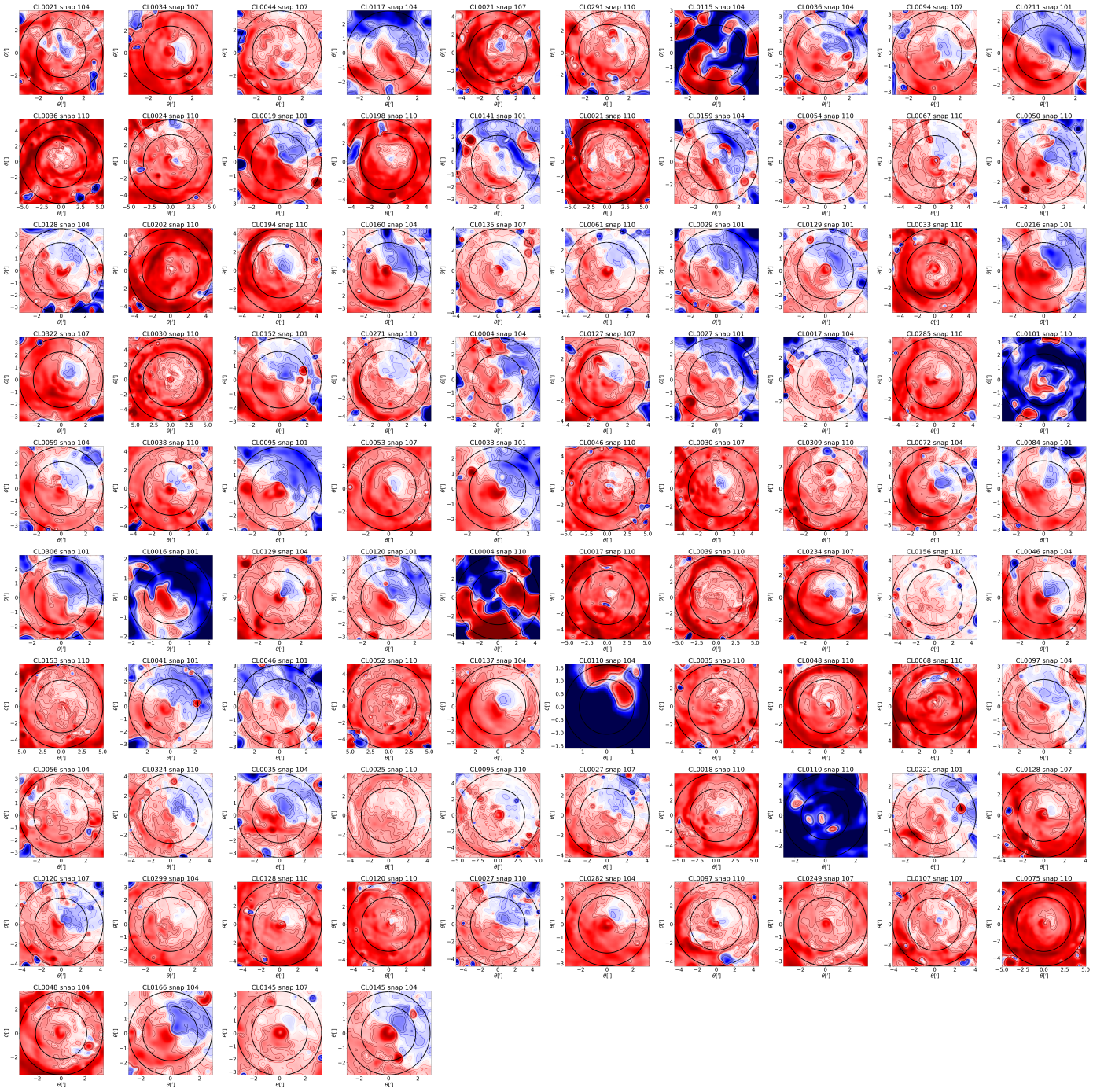

Fig. D.2. Comparison between T_{SZX} and T_{mw} maps obtained for all the clusters in our sample, using the true effective length map for each cluster. The maps are ordered by relaxation parameter at R_{500}, χ_{500}. The top left clusters have the lowest chi_{500}, i.e. are the most disturbed clusters, up to the bottom right clusters which are the most relaxed systems. The colormaps span the range ±50%, while the contours span the range ±10%.