

# ComPRASS: a Combined *Planck*-RASS catalogue of X-ray-SZ clusters<sup>★</sup>

P. Tarrío<sup>1,2</sup>, J.-B. Melin<sup>3</sup>, and M. Arnaud<sup>1,2</sup>

<sup>1</sup> IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

<sup>2</sup> Université Paris Diderot, AIM, Sorbonne Paris Cité, CEA, CNRS, 91191 Gif-sur-Yvette, France  
e-mail: paula.tarrío-alonso@cea.fr

<sup>3</sup> IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

Received 24 December 2018 / Accepted 28 February 2019

## ABSTRACT

We present the first all-sky catalogue of galaxy clusters and cluster candidates obtained from joint X-ray-SZ detections using observations from the *Planck* satellite and the ROSAT all-sky survey (RASS). The catalogue contains 2323 objects and has been validated by careful cross-identification with previously known clusters. This validation shows that 1597 candidates correspond to already known clusters, 212 coincide with other cluster candidates still to be confirmed, and the remaining 514 are completely new detections. With respect to *Planck* catalogues, the ComPRASS catalogue is simultaneously more pure and more complete. Based on the validation results in the SPT and SDSS footprints, the expected purity of the catalogue is at least 84.5%, meaning that more than 365 clusters are expected to be found among the new or still-to-be-confirmed candidates with future validation efforts or specific follow-ups.

**Key words.** catalogs – galaxies: clusters: general – X-rays: galaxies: clusters – methods: data analysis – techniques: image processing

## 1. Introduction

Galaxy cluster catalogues with high purity and completeness are fundamental to study clusters both from an astrophysical and a cosmological perspective. Since the first galaxy cluster catalogue constructed by [Abell \(1958\)](#) by analysing photographic plates, which only contained low-redshift clusters, numerous catalogs have been compiled using observational data sets at different wavelengths, from microwaves to X-rays, trying to improve the completeness and purity at all redshift and mass ranges.

The first cluster catalogues were built from optical data sets, where clusters are identified as overdensities of galaxies. This approach was also used later with infrared data. Clusters can also be detected in X-ray observations, where they appear as bright extended sources. In these images we see the emission of the hot gas of the intracluster medium (ICM). Finally, over the last decade, this gas has also begun to be detected thanks to the characteristic spectral distortion it produces on the cosmic microwave background (CMB) due to Compton scattering of the CMB photons by the ICM electrons. This effect is known as the Sunyaev-Zeldovich (SZ) effect ([Sunyaev & Zeldovich 1970, 1972](#)).

To date, cluster catalogues have usually been constructed through single-survey data, each probing a different region of the electromagnetic spectrum. However, this approach does not leverage the possible synergies that different sources of information might offer if a joint detection approach were used. A joint approach would allow to reduce contamination and thus increase the purity of the constructed catalogues, since a contaminant

source at a given wavelength will not be generally present at a different wavelength. On the other hand, it would also improve the detection efficiency: an object below the detection threshold of two independent surveys might be detected when combining the two, since its signal will be boosted.

Although multi-wavelength, multi-survey detection of clusters was theoretically conceived some years ago ([Maturi 2007; Pace et al. 2008](#)), it is a very complex task and until recently had only been attempted in the pilot study of [Schuecker et al. \(2004\)](#) on the X-ray ROSAT All-Sky Survey (RASS; [Truemper 1993; Voges et al. 1999](#)) and the optical Sloan Digital Sky Survey (SDSS; [York et al. 2000](#)). Recognising the opportunity of more efficiently exploiting existing observations to detect new galaxy clusters, we recently developed a novel cluster-detection method based on the combination of SZ and X-ray surveys ([Tarrío et al. 2016, 2018](#)).

The method is based on matched multifrequency filters (MMF; [Herranz et al. 2002; Melin et al. 2006, 2012](#)) and performs a true joint X-ray-SZ detection by treating the X-ray image as an additional frequency to be simultaneously filtered with the SZ frequency maps. In order to do so, the expected physical relation between SZ and X-ray fluxes is leveraged. The X-ray-SZ joint detection method can be seen as an evolution of the MMF3 detection method, one of the MMF methods used to detect clusters from *Planck* observations, which incorporates X-ray observations to improve the detection performance. In [Tarrío et al. \(2018\)](#), we evaluated this method in the area of the sky covered by the SPT survey using data from the RASS and *Planck* surveys. We showed that thanks to the addition of the X-ray information the method is able to simultaneously achieve better purity, better detection efficiency, and better position accuracy than its predecessor, the *Planck* MMF3 detection method. We also showed that if the redshift of a cluster is known by any

<sup>★</sup> Table A.1, Figure 1, and the X-ray map are only available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](ftp://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/626/A7>.

other means, the joint detection provides a good estimation of its mass.

In this paper, we present ComPRASS, a Combined *Planck*-RASS catalogue of X-ray–SZ sources. This all-sky catalogue contains 2323 galaxy cluster candidates and was constructed by applying the X-ray–SZ joint detection method proposed in Tarrío et al. (2018) on all-sky maps from the *Planck* and RASS surveys. We present an external validation of the catalogue using existing X-ray, SZ, and optical cluster catalogues. From this validation, we confirm 1597 ComPRASS candidates. The 726 remaining objects are either new objects (514) or are associated to an already known but yet unconfirmed cluster candidate (212). ComPRASS catalogue is simultaneously more pure and more complete than *Planck* catalogues, with an expected purity greater than 84.5% and at least 365 real clusters, unknown to date, among the new or yet-to-be-confirmed candidates.

The structure of the paper is as follows. Section 2 describes the construction of the ComPRASS catalogue, including a brief description of the input data and the detection method. In Sect. 3 we present the external validation of the catalogue, which is based on cross-identification with previously known clusters and cluster candidates from SZ, X-ray, and optical catalogues. Section 4 evaluates some of the properties of the catalogue by comparing it to other catalogues constructed from the same observations. Finally, we conclude the paper and discuss ongoing and future research directions in Sect. 5. The full catalogue is available in machine readable format. A full description of the available information is given in Appendix A.

Throughout, we adopt a flat  $\Lambda$ CDM cosmological model with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  and  $\Omega_M = 1 - \Omega_\Lambda = 0.3$ . We define  $R_{500}$  as the radius within which the average density of the cluster is 500 times the critical density of the universe,  $\theta_{500}$  as the corresponding angular radius, and  $M_{500}$  as the mass enclosed within  $R_{500}$ .

## 2. Construction of the ComPRASS catalogue

The ComPRASS catalogue was obtained by applying the X-ray–SZ joint detection method proposed in Tarrío et al. (2018) to all-sky maps from the *Planck* and RASS surveys. We note that the method of Tarrío et al. (2018) has not been modified for this paper. In the rest of this section, we briefly describe the observations that we used, we summarise the X-ray–SZ detection method indicating the chosen values for various selectable parameters, we present the resulting catalogue, and we provide an estimation of its completeness as a function of mass and redshift.

### 2.1. Input data

The ROSAT All-Sky Survey (RASS) is the only full-sky X-ray survey performed to date (Truemper 1993; Voges et al. 1999). The RASS data release<sup>1</sup> consists of 1378 fields that provide the exposure map and the X-ray counts in three different bands: TOTAL (0.1–2.4 keV), HARD (0.5–2.0 keV), and SOFT (0.1–0.4 keV). Each field covers  $6.4 \text{ deg} \times 6.4 \text{ deg}$  of sky and has a size of  $512 \times 512$  pixels, yielding a resolution of  $0.75 \text{ arcmin pixel}^{-1}$ .

To construct the ComPRASS catalogue, we used an X-ray all-sky HEALPix map that we built from the RASS HARD band information and the RASS exposure maps. This map has a

resolution of  $0.86 \text{ arcmin pixel}^{-1}$ , which is HEALPix resolution closest to the RASS resolution. The details of its construction can be found in Appendix B of Tarrío et al. (2016). This map is available at the CDS.

*Planck* is the only all-sky SZ survey. It observed the sky in nine frequency bands with two instruments: the Low Frequency Instrument (LFI), which covered the 30, 44, and 70 GHz bands, and the High Frequency Instrument (HFI), which covered the 100, 143, 217, 353, 545, and 857 GHz bands.

To construct the ComPRASS catalogue, we use the six temperature channel maps of HFI, which are the same channels used by the Planck Collaboration to produce their cluster catalogues (Planck Collaboration VIII 2011; Planck Collaboration XXIX 2014; Planck Collaboration XXVII 2016). In particular, we used the latest version of these maps; a description can be found in Planck Collaboration VIII (2016). The resolution of the published maps is  $1.72 \text{ arcmin pixel}^{-1}$ . As in Tarrío et al. (2016) and Tarrío et al. (2018), to make them directly compatible with the all-sky X-ray map mentioned above, we up-sampled them to a resolution of  $0.86 \text{ arcmin pixel}^{-1}$  by zero-padding in the spherical harmonics domain, that is, by adding new modes with zero power.

### 2.2. Joint detection method

The X-ray–SZ joint detection method proposed in Tarrío et al. (2018) is based on considering the X-ray map as an additional frequency to be simultaneously filtered with the SZ frequency maps. In order to do so, the X-ray map needs to be converted into an equivalent SZ map at a reference frequency  $\nu_{\text{ref}}$ , leveraging the expected physical relation between SZ and X-ray fluxes, namely the  $F_X/Y_{500}$  relation. The details of this conversion are described in Appendix B of Tarrío et al. (2016). The reference frequency  $\nu_{\text{ref}}$  is just a fiducial value with no effect on the detection algorithm. In our case, we took  $\nu_{\text{ref}} = 1000 \text{ GHz}$ . For the  $F_X/Y_{500}$  relation we assumed the relation found by the Planck Collaboration I (2012), fixing the redshift to a reference value of  $z_{\text{ref}} = 0.8$ , as done in Tarrío et al. (2018). As shown in the right panel of Fig. 16 of Tarrío et al. (2016), a change in the normalization of the assumed  $F_X/Y_{500}$  relation by a factor of two only impacts the S/N measurement by up to 5%, which makes the detection robust against possible errors in the assumed relation.

After this conversion, the complete set of maps (the original  $N_\nu$  SZ maps obtained at sub-millimetre frequencies  $\nu_1, \dots, \nu_{N_\nu}$ , and the additional map at the reference frequency  $\nu_{\text{ref}}$  obtained from the X-ray map) are filtered using the following X-ray–SZ MMF<sup>2</sup>:

$$\Psi_{\theta_s}(\mathbf{k}) = \sigma_{\theta_s}^2 \mathbf{P}^{-1}(k) \mathbf{F}_{\theta_s}(\mathbf{k}), \quad (1)$$

with

$$\sigma_{\theta_s}^2 = \left[ \sum_{\mathbf{k}} \mathbf{F}_{\theta_s}^T(\mathbf{k}) \mathbf{P}^{-1}(k) \mathbf{F}_{\theta_s}(\mathbf{k}) \right]^{-1}, \quad (2)$$

and

$$\mathbf{F}_{\theta_s}(\mathbf{k}) = [j(\nu_1)T_1(\mathbf{k}), \dots, j(\nu_{N_\nu})T_{N_\nu}(\mathbf{k}), Cj(\nu_{\text{ref}})T_{\theta_s}^x(\mathbf{k})]^T. \quad (3)$$

<sup>2</sup> We use  $\mathbf{k}$  to denote the 2D spatial frequency, corresponding to the 2D position  $\mathbf{x}$  in the Fourier space. We use  $k$  to denote its modulus. All the variables expressed as a function of  $\mathbf{k}$  or  $k$  are thus to be understood as variables in the Fourier space.

<sup>1</sup> [ftp://legacy.gsfc.nasa.gov/rosat/data/pspc/processed\\_data/rass/release](ftp://legacy.gsfc.nasa.gov/rosat/data/pspc/processed_data/rass/release), or <http://www.xray.mpe.mpg.de/rosat/survey/rass-3/main/help.html#ftp>

Here,  $\Psi_{\theta_s}$  is a  $(N_v + 1) \times 1$  column vector whose  $i$ th component will filter the map at observation frequency  $\nu_i$ ;  $\mathbf{P}(k)$  is the noise power spectrum;  $j(\nu_i)$  is the SZ spectral function at frequency  $\nu_i$ ;  $T_i(\mathbf{k}) = \tilde{T}_{\theta_s}^x(\mathbf{k})B_{\nu_i}(\mathbf{k})$  and  $T_{\theta_s}^x(\mathbf{k}) = \tilde{T}_{\theta_s}^x(\mathbf{k})B_{X\text{-ray}}(\mathbf{k})$  are the convolutions of the normalized cluster 2D spatial profiles (SZ and X-ray components, respectively) with the point spread function (PSF) of the instruments at the different frequencies; and the constant  $C$  is a geometrical factor that accounts for the different shapes of the SZ and X-ray 3D profiles (Eq. (25) of Tarrío et al. 2016). A more elaborated description of the filter can be found in Tarrío et al. (2016) and Tarrío et al. (2018).

The filtering is done in the Fourier space as  $\hat{y}(\mathbf{k}) = \Psi_{\theta_s}^T(\mathbf{k}) \cdot \mathbf{M}(\mathbf{k})$ , where  $\mathbf{M}(\mathbf{k}) = [M_1(\mathbf{k}), \dots, M_{N_v}(\mathbf{k}), M_{\text{ref}}(\mathbf{k})]^T$  is the Fourier transform of the  $N_v + 1$  input maps. After transforming the result back to the real space, we obtain the filtered map  $\hat{y}(\mathbf{x})$  and the S/N map  $(\hat{y}(\mathbf{x})/\sigma_{\theta_s})$ .

We note that the properties of the noise are not the same in different sky regions; therefore,  $\mathbf{P}(k)$  has to be calculated locally at each position. This is done in practice from the X-ray and SZ images themselves, assuming that they contain mostly noise. Since this assumption may not hold in some X-ray images due to bright X-ray sources with strong signals, we apply the X-ray source mask defined in Tarrío et al. (2018) for the calculation of  $\mathbf{P}(k)$ .

This MMF approach relies on the knowledge of the normalized SZ and X-ray cluster profiles ( $\tilde{T}_{\theta_s}(\mathbf{k})$  and  $\tilde{T}_{\theta_s}^x(\mathbf{k})$ ). These profiles are not known in practice, and therefore they need to be approximated by the theoretical profiles that best represent the clusters we want to detect. As in Tarrío et al. (2016) and Tarrío et al. (2018), we assume the generalised Navarro-Frenk-White (GNFW) profile (Nagai et al. 2007) given by

$$f(x) \propto \frac{1}{(c_{500}x)^\gamma [1 + (c_{500}x)^\alpha]^{(\beta-\gamma)/\alpha}}, \quad (4)$$

with parameters given by

$$[\alpha, \beta, \gamma, c_{500}] = [1.0510, 5.4905, 0.3081, 1.177], \quad (5)$$

for the 3D pressure profile (Arnaud et al. 2010), and

$$[\alpha, \beta, \gamma, c_{500}] = [2.0, 4.608, 1.05, 1/0.303], \quad (6)$$

for the square of the gas density profile (Piffaretti et al. 2011), respectively. Here,  $x = r/R_{500}$  represents the 3D distance to the centre of the cluster in  $R_{500}$  units, and  $R_{500}$  relates to the characteristic cluster scale  $R_s$  through the concentration parameter  $c_{500}$  ( $R_s = R_{500}/c_{500}$ ). The cluster SZ and X-ray profiles as a function of the scaled angular radius ( $\theta/\theta_{500} = x$ ) are then obtained by numerically integrating these 3D GNFW profiles along the line of sight.

Finally, these normalized cluster profiles need to be convolved by the instrument beams ( $B_{\nu_i}(\mathbf{k})$  and  $B_{X\text{-ray}}(\mathbf{k})$ ). For the SZ components, we use a Gaussian PSF with FWHM depending on the frequency, as shown in Table 6 of Planck Collaboration VIII (2016). For the X-ray component, we use a PSF that was estimated numerically by stacking observations of X-ray point sources from the Bright Source Catalogue (Voges et al. 1999). These are the same instrument beams used in Tarrío et al. (2016) and Tarrío et al. (2018), since we have the same input observations.

To implement the detection procedure in practice, we proceed in two phases.

1. First phase: We project the all-sky maps into 504 small  $10^\circ \times 10^\circ$  tangential patches, as was done in MMF3

(Planck Collaboration VIII 2011; Planck Collaboration XXIX 2014; Planck Collaboration XXVII 2016). Each patch is filtered by the X-ray-SZ filter  $\Psi_{\theta_s}$  (Eq. 1) using  $N_s = 32$  different sizes, covering the expected range of radii. In our case, we vary  $\theta_{500}$  from 0.94 to 35.31 arcmin, in  $N_s = 32$  steps equally spaced in logarithmic scale. For each size, we obtain a filtered map and a S/N map. We then construct a list with the peaks in these maps that are above a specified S/N threshold  $q$ . The iterative procedure for finding the peaks in each set of  $N_s$  S/N maps is described in Tarrío et al. (2018). Subsequently we merge the 504 lists into a single preliminary all-sky list of candidates by merging peaks that are close to each other by less than 10 arcmin, as was done in MMF3 (Planck Collaboration VIII 2011; Planck Collaboration XXIX 2014; Planck Collaboration XXVII 2016).

2. Second phase: We create a set of maps centered at each candidate from the first phase, and we re-apply the MMF (see details of the procedure in Tarrío et al. 2018). This second phase allows a better estimation of the candidate properties and S/N. Only the candidates whose new S/N is above the specified threshold  $q$  are kept in the final list.

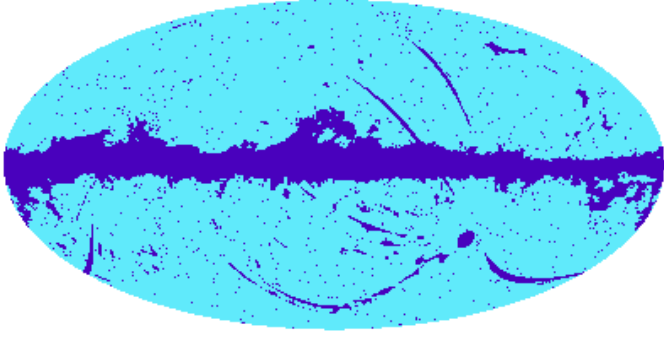
A key point of this X-ray-SZ detection algorithm is the selection of the threshold  $q$  that defines which S/N peaks are kept in the first and second phases. As explained in Tarrío et al. (2018), the X-ray-SZ method uses an adaptive threshold that depends on the noise characteristics of each region, and that is determined numerically by means of Monte Carlo simulations. These simulations establish, as a function of the mean Poisson noise level in the X-ray map, the mean Gaussian noise level in the SZ maps and the filter size ( $\theta_s$ ), the joint S/N threshold  $q_J$  corresponding to a given false-alarm probability  $P_{\text{FA}}$  (i.e. the probability that a detection is due to a random fluctuation). The value of  $P_{\text{FA}}$  serves to select the operational point of the detection method, with higher values resulting in more complete but less pure catalogues and lower values yielding more pure but less complete catalogues. To construct the ComPRASS catalogue, we used  $P_{\text{FA}} = 3.4 \times 10^{-6}$ , which corresponds to a cut at  $4.5\sigma$  in a zero-mean Gaussian distribution. For reasons of simplicity, the adaptive threshold  $q_J$  is applied after obtaining the final candidate list from the second phase. The threshold  $q$  to be applied in the first and second phases is set to  $q = 4$ , a sufficiently low value so that it does not introduce any different selection effect, that is, it does not discard any candidates above  $q_J$ .

Finally, the catalogue produced at the second phase needs to be further cleaned to discard detections in contaminated regions of the sky, in regions with poor statistics, or that correspond to non-cluster objects (mainly AGNs). To this end, we apply an SZ mask to avoid SZ contaminated regions and a X-ray exposure time mask ( $t_{\text{exp}} > 100$  s) to avoid X-ray regions with poor statistics. These two masks are defined in Sects. 3.4.1 and 3.4.2 of Tarrío et al. (2018), respectively. Figure 1 shows the combination of these two masks, which provide a final footprint of 80.96% of the sky. To discard AGN detections, we establish an additional criterion based on the SZ part of the S/N: we only keep the detections satisfying  $(S/N)_{\text{SZ}} > 3$ , as described in Sect. 3.4.3 of Tarrío et al. (2018).

### 2.3. ComPRASS catalogue

The ComPRASS catalogue was obtained by running the blind joint X-ray-SZ detection algorithm summarised in Sect. 2.2 on the RASS and *Planck* all-sky maps described in Sect. 2.1. It contains 2323 candidates, distributed in the sky as shown in Fig. 2. The sky is not covered homogeneously: there are more candidates in the regions where the RASS exposure time is





**Fig. 1.** Mask applied to the second phase detections. This results from the combination of the SZ mask and the X-ray exposure time mask. The dark-blue areas are masked, the light-blue area is the final footprint, which accounts for 80.96% of the sky. The electronic version of this mask is available at the CDS.

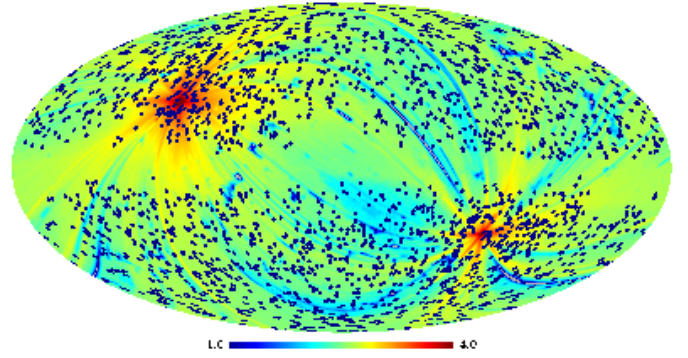
higher and where the *Planck* noise is lower. This is expected, since in those regions both surveys are deeper.

Table A.1 summarises the main properties of the ComPRASS candidates. For each candidate, the catalogue provides its position, the joint S/N:  $(S/N)_J$ , and the SZ and X-ray components of this S/N:  $(S/N)_{SZ}$  and  $(S/N)_{XR}$ .  $(S/N)_J$  is the value of the joint S/N map  $(\hat{s}(\mathbf{x})/\sigma_{\theta_s})$  at the position and size of the detection.  $(S/N)_{SZ}$  ( $(S/N)_{XR}$ ) corresponds to the S/N of the SZ (X-ray) filtered maps, that is, the SZ (X-ray) filtered maps divided by the background noise of the SZ (X-ray) maps at the position and size of the blind joint detection. It also provides a value for the significance of each detection. This value is defined in Tarrío et al. (2018) as the significance value in a Gaussian distribution corresponding to the probability that the detection is due to noise, which is calculated from the results of the Monte Carlo simulations performed to calculate the joint S/N threshold  $q_J$ .

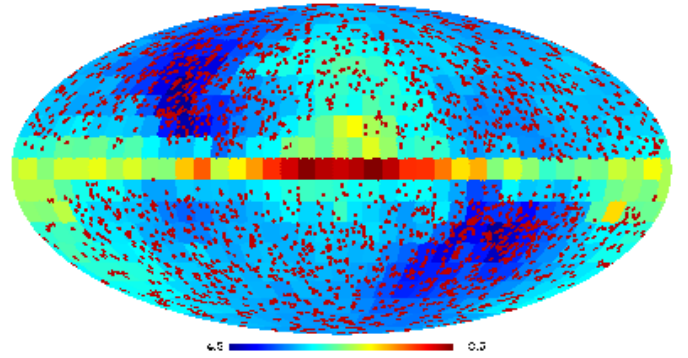
The electronic version of the catalogue (Table A.1) also contains the mass-redshift degeneracy curves for each candidate. These curves provide the mass ( $M_{500}^{XSZ}$ ) estimated by the joint X-ray–SZ extraction for various values of redshift between 0.01 and 1.2. They are obtained by (1) re-extracting the candidate signal at the position given by the blind detection using different values for the reference redshift  $z_{ref}$  (0.01, 0.1, 0.2, 0.4, 0.7 and 1.2); (2) interpolating between the six resulting  $Y_{500}(\theta_{500})$  degeneracy curves to obtain the curve corresponding to a given value of redshift between 0.01 and 1.2; and (3) breaking the size-flux degeneracy using the  $M_{500} - D_A^2 Y_{500}$  scaling relation from Planck Collaboration XX (2014):

$$E^{-2/3}(z) \left[ \frac{D_A^2(z) Y_{500}}{10^{-4} \text{Mpc}^2} \right] = 10^{-0.19} \left[ \frac{M_{500}}{6 \times 10^{14} M_\odot} \right]^{1.79}, \quad (7)$$

which relates  $\theta_{500}$  and  $Y_{500}$  when  $z$  is known, as explained in Planck Collaboration XXIX (2014) and Tarrío et al. (2018). This relation does not take into account possible bias between X-ray derived mass and true mass. The upper and lower bounds of the 68% confidence interval in the mass-redshift degeneracy curves are also provided in the electronic version of the catalogue. These bounds are derived from the bounds in the  $Y_{500}(\theta_{500})$  curves, using the procedure explained above. The errors on the  $Y_{500}(\theta_{500})$  curves are calculated from Eq. (32) of Tarrío et al. (2016). This expression corrects the error on  $Y_{500}$  obtained by the filter (Eq. (28) of Tarrío et al. 2016) by a factor that depends on the size  $\theta_{500}$  and that accounts for the additional dispersion produced by the mismatch between the real cluster profiles and



(a)



(b)

**Fig. 2.** Sky distribution of the 2323 ComPRASS candidates (blue/red dots). The sky map is colour-coded according to (a) the logarithm of the RASS exposure time and (b) the logarithm of the *Planck* noise standard deviation map.

the ones used in the filter. We note that the error on  $Y_{500}$  obtained by the filter (Eq. (28) of Tarrío et al. 2016) does not only depend on the S/N, but also on the variance of the filtered Poisson fluctuations on the signal.

The catalogue also contains the information obtained from the external validation (see details in Sect. 3). If the candidate is associated with a previously known cluster or cluster candidate, the name of the associated object is indicated. Moreover, if the redshift of the associated cluster is known, it is also included in the ComPRASS catalogue. In the case of multiple associations with different redshifts, the order of priority defined in Sect. 3.6 is applied. When the redshift of the cluster candidate is known from the external validation, the catalogue provides an estimation of its mass ( $M_J$ ), and the upper and lower bounds of its 68% confidence interval, calculated from the mass- $z$  degeneracy curves at the redshift of the cluster.

A full description of all the fields provided in the ComPRASS catalogue is given in Appendix A.

#### 2.4. Completeness estimation

In this section we provide an estimation of the completeness to illustrate the performance of the ComPRASS catalogue at different mass and redshift ranges.

The completeness has been calculated by injecting simulated clusters into the real RASS and *Planck* maps described in Sect. 2.1 and extracting them with the filter described

in Sect. 2.2. The clusters whose extracted S/Ns satisfy the thresholds imposed in the blind detection ( $S/N_J > q_J$  and  $S/N_{SZ} > 3$ ) are considered as detected. The completeness is then calculated as the fraction of detected clusters with respect to the total number of clusters injected outside the masked region of the sky (see Fig. 1). For simplicity, we fixed the positions, sizes, and redshifts of the clusters for the extraction step. Figure 3 shows the resulting completeness as a function of mass and redshift. For a mass of  $5 \times 10^{14} M_\odot$ , the catalogue should contain 80% of the clusters at  $z = 0.5$  and 60% of the clusters at  $z = 1$ .

For these simulations, we considered four mass bins: 2–4, 4–6, 6–8, and 8– $10 \times 10^{14} M_\odot$  and eight redshift bins: 0.1–0.3, 0.3–0.5, 0.5–0.6, 0.6–0.7, 0.7–0.8, 0.8–0.9, 0.9–1.0, and 1.0–1.1, resulting in 32 bins in the redshift-mass plane. For each bin, we injected 1000 clusters at random positions of the sky, with  $z$  and  $M_{500}$  uniformly distributed in the bin. The resulting 32 000 clusters were then re-binned into smaller mass-redshift bins for visualisation purposes in Fig. 3. The clusters were simulated as in Tarrío et al. (2016; Sects. 2.2.1 and 4.2.1), assuming the GNFW profile defined in Eq. (4) with the parameters given by Eqs. (5) and (6), for the SZ and X-ray components, respectively. The amplitudes of these simulated clusters are estimated according to the expected values of  $Y_{500}$  (for SZ) and  $L_{500}$  (for X-ray). The luminosities  $L_{500}$  and the corresponding flux  $F_X$  were calculated from the L–M relation in Arnaud et al. (2010), Planck Collaboration XI (2011), including the scatter  $\sigma_{\log L} = 0.183$ . The flux  $Y_{500}$  was calculated from the nominal flux  $F_X$  assuming the  $F_X/Y_{500}$  relation found by the Planck Collaboration I (2012), implicitly neglecting the scatter of the Y–M relation.

We note that this selection function should not be used for precise cosmological purposes. A precise selection function would require a full autoconsistent modelling of the scatter in the Y–M relation, the deviations from the assumed cluster profile, the uncertainty on the redshift at the detection step, and so on, as a function of the position on the sky.

### 3. External validation

The ComPRASS catalogue is validated by identifying the candidates that are associated with previously known clusters. To this end, we use existing SZ, X-ray, and optical cluster catalogues.

In particular, we took several SZ-selected catalogues, namely the three *Planck* catalogues: ESZ (Planck Collaboration VIII 2011), PSZ1 (Planck Collaboration XXIX 2014; Planck Collaboration XXXII 2015), and PSZ2 (Planck Collaboration XXVII 2016), the SPT catalogue (Bleem et al. 2015), and the ACT and ACTPol catalogues (Hasselfield et al. 2013; Hilton et al. 2018). The redshifts in the original PSZ1 and PSZ2 catalogues were updated with follow-up data from MegaCam at CFHT (van der Burg et al. 2016), NOT (Dahle et al., in prep.), WHT (Buddendiek et al. 2015), the PSZ1 follow-up program developed at Roque de los Muchachos Observatory (Barrena et al. 2018)<sup>3</sup>, the high- $z$  PSZ2 follow up program of Burenin et al. (2018), the optical follow-up of PSZ2 clusters of Boada et al. (2018), and from some private communications on specific clusters (see notes in ComPRASS catalogue). We also consider the extension of the *Planck* catalogue presented in Burenin (2017).

We also took as reference the X-ray selected MCXC catalogue (Piffaretti et al. 2011). This is a metacatalogue of

<sup>3</sup> We only updated the clusters with definitely confirmed optical counterparts (validation flag = 1) and those with no counterpart found (validation flag = ND).

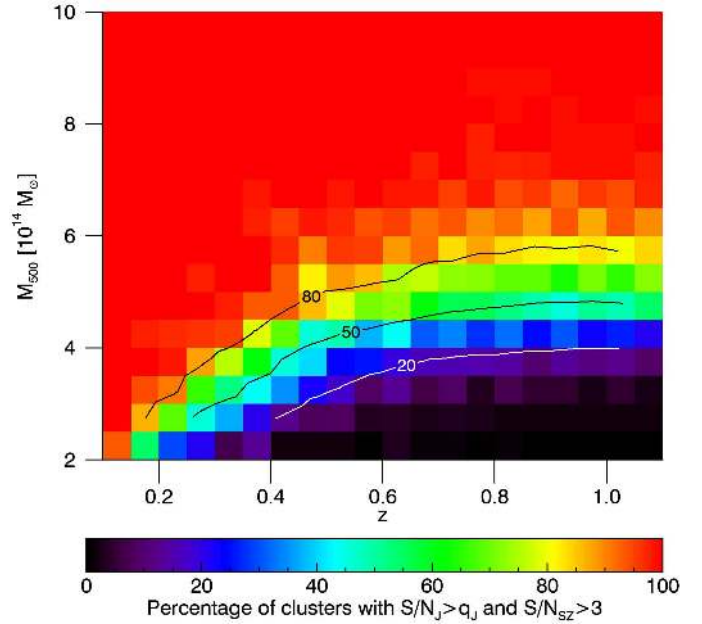


Fig. 3. Estimated completeness of the ComPRASS catalogue in different mass-redshift bins.

X-ray detected clusters that was constructed from publicly available cluster catalogues of two kinds: RASS-based catalogues, obtained from the RASS survey data, and serendipitous catalogues, based on deeper pointed X-ray observations. It contains only clusters with an available redshift at that date. We used an updated version of the catalogue (Sadibekova et al., in prep.), which includes the 1743 original clusters plus 125 new clusters from 160SD, NORAS, WARPS, SGP, REFLEXII (Chon & Böhringer 2012) and MACS (including the ones in Mann & Ebeling 2012 and Repp & Ebeling 2018). This new version takes into account the new spectroscopic redshifts published in the literature for the ROSAT clusters. We furthermore supplement the updated MCXC with 44 unpublished MACS clusters observed by *XMM-Newton* or *Chandra*.

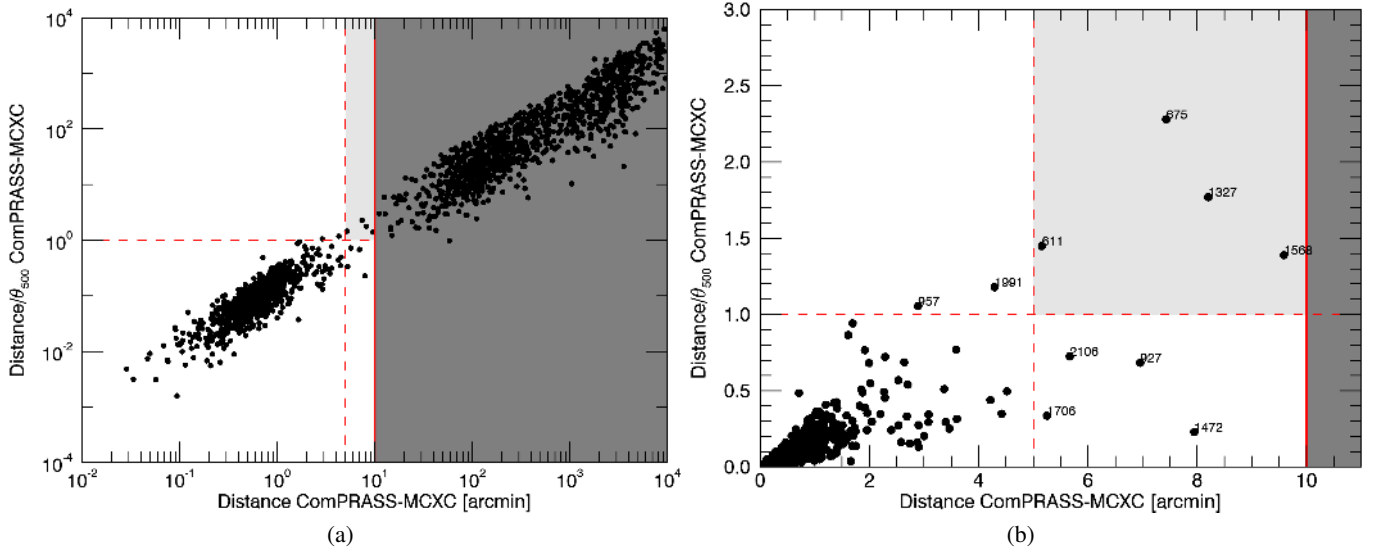
In addition, we considered several optically selected cluster catalogues: the Abell catalogue (Abell et al. 1989), the Zwicky catalogue (Zwicky et al. 1961), and five catalogues based on SDSS data, namely, the redMaPPer catalogue (Rykoff et al. 2014), the MaxBCG catalogue (Koester et al. 2007), the GMBCG catalogue (Hao et al. 2010), the AMF catalogue (Szabo et al. 2011), and the WHL catalogue (Wen et al. 2012).

Finally, we also took into account three recently published high-redshift cluster catalogues (Buddendiek et al. 2015; Gonzalez et al. 2019 and Wen & Han 2018) and performed a search in NED and SIMBAD databases for possible missing associations.

Based on the identification with clusters or candidates of these catalogues, the ComPRASS candidates are classified into three classes:

- *Confirmed* (class 1): candidates that are associated to an already confirmed cluster.
- *Identified not confirmed* (class 2): candidates that are associated to candidates in X-ray or SZ catalogues that have not been yet confirmed.
- *New* (class 3): candidates that are not associated to any cluster or cluster candidate.

In the rest of this section, we describe in detail this external validation.



**Fig. 4.** Positional criteria for matching ComPRASS candidates to X-ray-detected MCXC clusters. The distance between each ComPRASS candidate and its closest MCXC cluster is plotted against their relative distance in terms of  $\theta_{500}$ . The white areas define the regions where the associations are considered to be correct. *Panel a:* complete view in logarithmic scale, while *panel b:* zoom of the low-distance region.

### 3.1. Identification with X-ray clusters

To identify the ComPRASS candidates that correspond to already known X-ray clusters, we first determined the closest MCXC cluster for each ComPRASS candidate. We then relied on two quantities to do the association: the angular distance between the candidate and the closest cluster ( $d$ ), and their relative distance in terms of the size of the cluster ( $d/\theta_{500}$ ), calculated from the mass and redshift reported in the MCXC catalogue. Figure 4a shows a scatter plot of the angular distance versus the relative distance between all the ComPRASS candidates and their closest MCXC cluster. We observe two main clouds of points: those with a small distance in absolute and in relative terms, and those with a long distance in absolute and in relative terms. The first cloud of points represents good associations, whereas the second cloud corresponds to the candidates that are randomly distributed with respect to the considered known clusters. There are also a few points between the two main clouds (see zoom in Fig. 4b) for which the association is not clear. These intermediate cases were studied individually (see Appendix B for details) to conclude that only four of them did not correspond to a true association: candidates 611, 675, 1327, and 1568. From this analysis, we decided to use the following association rules:

- If  $d > 10'$  the candidate is not associated with the cluster (dark grey area in Fig. 4).
- If  $5' < d < 10'$  and  $d/\theta_{500} > 1$  the candidate is not associated with the cluster (light grey area in Fig. 4).
- Otherwise ( $5' < d < 10'$  and  $d/\theta_{500} < 1$ , or  $d < 5'$ ) the candidate is associated with the cluster (white area in Fig. 4).

These association rules give us 737 matches of ComPRASS candidates with MCXC clusters. 696 candidates are associated with MCXC clusters with known redshift and mass. The remaining 41 are associated with MCXC clusters for which we do not know  $\theta_{500}$  (7 clusters without redshift and 34 MACS clusters without a published  $L_{500}$ , and thus, without  $M_{500}$ ). Although we cannot calculate the relative distance for these 41 associations, all of them have  $d < 3$  arcmin, meaning that they are considered to be correct. Of these 737 associations, 730 are classified as confirmed (class 1), while the remaining 7, corresponding to

the 7 clusters without redshift, are classified as identified but not confirmed (class 2).

Following the procedure described in Sect. 2.3 and using the  $M_{500} - D_A^2 Y_{500}$  relation proposed in Planck Collaboration XX (2014; see Eq. (7)), we estimated the mass  $M_{500}$  for the 696 detections matching confirmed MCXC clusters with published values for their mass. Figure 5 shows the relation between the estimated mass and the published mass for the corresponding clusters.

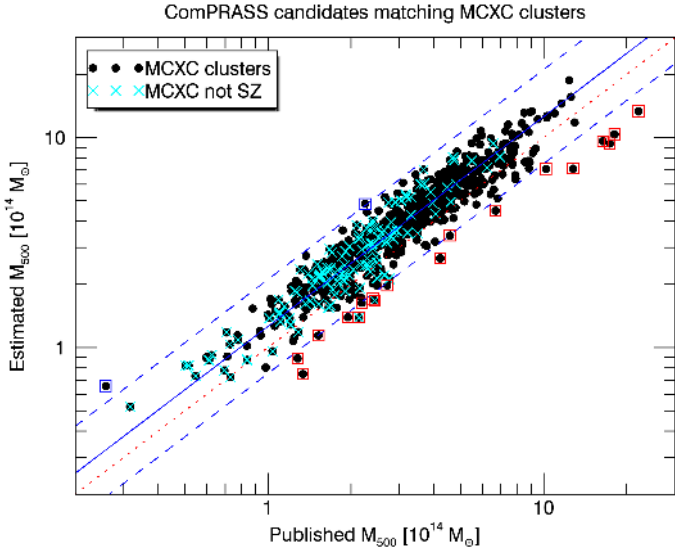
This mass comparison shows that the ratio between the estimated mass and the published mass is on average slightly greater than one, with a median value of 1.26. The same value is found for the ratio between the published SZ mass and the published MCXC mass for the same clusters. This behaviour was also observed by the Planck Collaboration XXVII (2016) when they compared the SZ mass and the X-ray luminosity of common PSZ2-MCXC objects. We identified twenty outliers that are at more than  $2.5\sigma$  from the median ratio: two with overestimated mass and eighteen with underestimated mass.

The two ComPRASS candidates with overestimated mass (1490 and 2195) match also PSZ2 clusters with known redshift and mass. The published masses for the corresponding PSZ2 clusters are higher than those of the associated MCXC clusters, and compatible with the joint mass estimates. The redshift of the MCXC and the PSZ2 clusters coincide, so we can conclude that these outliers are due to a difference between the X-ray and the SZ mass estimates, perhaps because they are X-ray-underluminous clusters.

Regarding the 18 candidates with underestimated mass, 10 of them are very extended and present an offset between the SZ and the X-ray peaks, which results in an underestimated mass. The remaining 8 match PSZ2 clusters with published SZ masses smaller than the MCXC masses; they are probably clusters with a stronger X-ray emission than the average for a given mass (2 of them are classified as cool-core clusters in Rossetti et al. 2017), yielding a higher X-ray mass estimate.

This discussion is in agreement with results from hydrodynamical simulations, which indicate that the SZ signal is more tightly correlated to the mass than the X-ray flux. We cannot exclude that some outliers are pathological cases (e.g.





**Fig. 5.**  $M_{500}$  estimated from the joint detection for the 696 ComPRASS candidates matching a confirmed MCXC cluster vs. the published  $M_{500}$  of the corresponding cluster. The dotted red line indicates the line of zero intercept and unity slope. The solid blue line indicates the median ratio. The dashed blue lines indicate the interval of  $\pm 2.5\sigma$  around the median ratio. Outliers are highlighted with a blue (high estimated mass) or red (low estimated mass) square; see text. Cyan crosses indicate MCXC clusters that are not in *Planck* or SPT.

very strong mergers) where the SZ signal is very different from what is expected from the true mass.

### 3.2. Identification with SZ clusters

To identify the ComPRASS candidates that correspond to already known SZ clusters, we followed a procedure analogous to the one described in Sect. 3.1 for the association with MCXC clusters. To this end, we selected only the objects in the considered SZ catalogues (excluding Burenin 2017) with known redshift and mass (i.e. confirmed clusters). Figure 6a shows a scatter plot of the angular distance versus the relative distance between all the ComPRASS candidates and their closest SZ confirmed cluster (with known redshift and mass). As in the association with MCXC clusters, we observe two main clouds of points: those corresponding to good associations and those corresponding to bad associations. There are also a few points between the two main clouds (see zoom in Fig. 6b) for which the association is not clear. These intermediate cases were studied individually to conclude that only four of them did not correspond to a true association: candidates 1300, 1587, 1718, and 2094 (see Appendix B for details). From this analysis, we decided to discard the associations with  $d > 5'$  and  $d/\theta_{500} > 1$ , as in the association with MCXC clusters. This association rule gives us 1060 matches of ComPRASS candidates with SZ clusters with known redshift and mass. These candidates are considered as confirmed (class 1).

Apart from confirmed SZ clusters, the considered SZ catalogues also contain cluster candidates that have not yet been confirmed to date. The redshift of these objects is not known, and therefore we cannot calculate  $\theta_{500}$  and the relative distance for the possible associations. For this reason we need to define a different association rule for these cases. To this end, we did a positional matching within a  $10'$  search radius and we found 239 candidates matching a “not confirmed” SZ candidate. Most

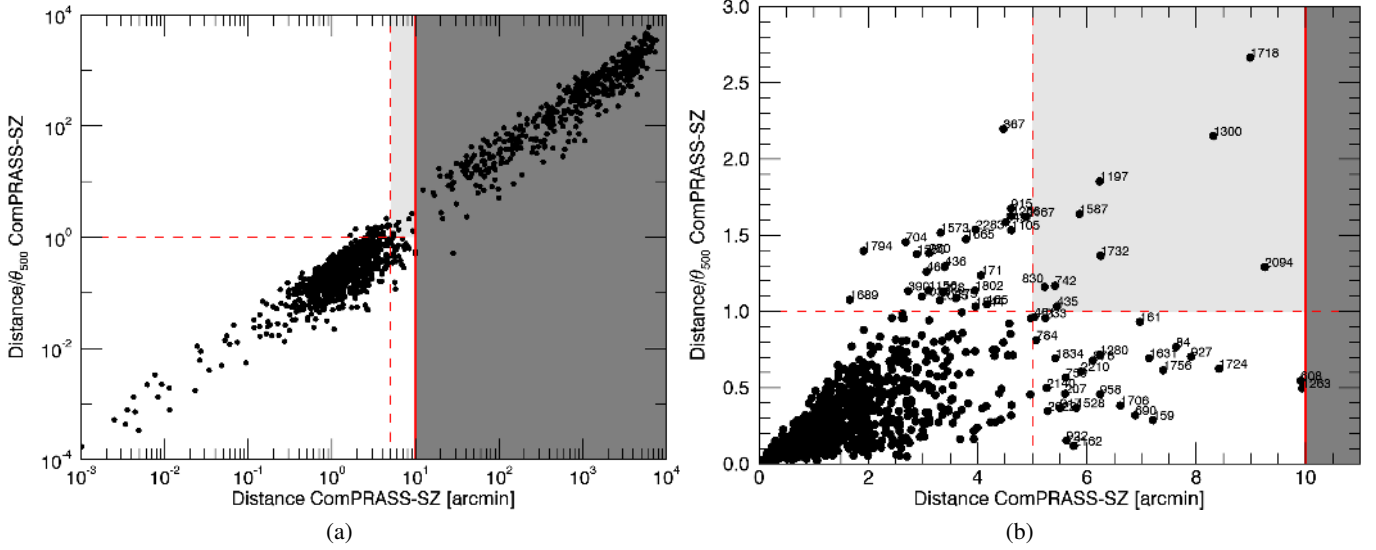
of these associations (233) have  $d < 5$  arcmin, and therefore based on the above study they are considered to be correct. The remaining six (candidates 153, 809, 1487, 1904, 1934, 2152) have  $d > 5$  arcmin and were analysed individually. Based on their X-ray and SZ S/N maps, it seemed that most of these associations were correct, and so we decided to use the following rule for the associations with SZ candidates: If  $d < 10'$  the ComPRASS candidate is associated with the SZ candidate; otherwise it is not associated. The candidates associated in this way to an SZ candidate are classified as class 2 (identified, not confirmed). If they were already confirmed by the X-ray identification, they remain in class 1 (3 out of 239).

We estimated the mass  $M_{500}$  for the 1060 detections matching confirmed SZ clusters. Figure 7 shows the relation between the estimated mass and the published mass for the corresponding clusters. This mass comparison shows that the ratio between the estimated mass and the published mass is very close to one, with a median value of 0.95. We identified 30 outliers that are at more than  $2.5\sigma$  from the median ratio: 22 with overestimated mass and 8 with underestimated mass.

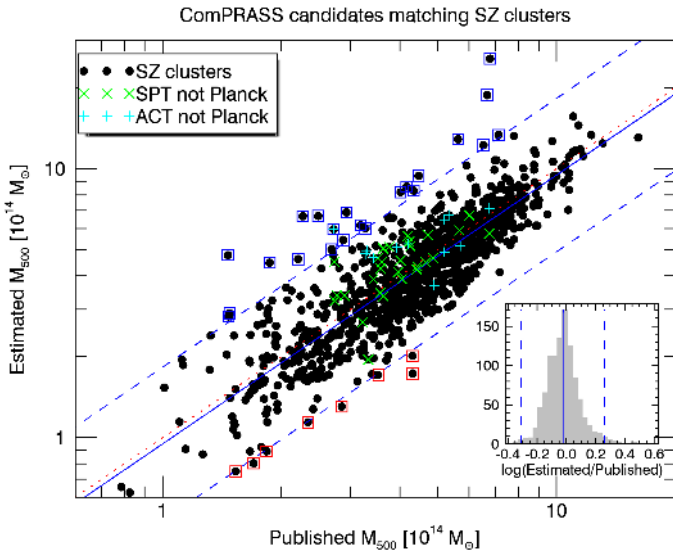
Regarding the 22 ComPRASS candidates with overestimated mass, we found that 18 of them are cool-core clusters, according to Rossetti et al. (2017), Andrade-Santos et al. (2017), Ebeling et al. (2010), Pratt et al. (2009), Hudson et al. (2010) and Morandi et al. (2015). Another two candidates are also cool-core clusters according to Bartalucci (priv. comm.) and Reiprich & Schellenberger (priv. comm.). Bartalucci analysed the XMM observation of candidate 1053 and obtained a central density at  $0.01R_{500}$  of 0.058, higher than the limit of 0.015 considered in Hudson et al. (2010), indicating that it is a cool-core. Candidate 2019 is also a cool-core cluster according to the temperature profile obtained from *Chandra* observations by Reiprich and Schellenberger (priv. comm.). Since our mass estimate relies both on the X-ray and SZ signal, and the X-ray emission for these clusters is stronger than the one expected from the SZ emission, our joint mass estimate is higher than the mass estimated from the SZ signal only, which is very close to the published SZ mass. This explains why our joint mass estimates are higher than the published SZ masses for these cool-core clusters. Regarding the remaining two ComPRASS candidates with overestimated mass, one is not a cool-core according to Andrade-Santos et al. (2017), but appears as a weak cool-core in Hudson et al. (2010). Finally, for the last one we do not know whether it is a cool-core cluster or not.

Regarding the eight candidates with underestimated mass, five of them are very extended and present an offset between the SZ and the X-ray peaks. This results in an underestimated mass because the extracted signal is not centred at the SZ peak. The three remaining candidates are probably clusters that are underluminous in X-ray, since the estimated mass using just the extracted SZ signal is perfectly compatible with the published SZ mass.

Finally, we also considered the catalogue of galaxy clusters presented in Burenin (2017) detected in the *Planck* all-sky Compton parameter maps and identified using data from the WISE and SDSS surveys. Since this catalogue contains a large number of projections (about 37%), we decided to be conservative and consider these clusters as not-yet-confirmed SZ candidates (even though a large percentage of them will be real clusters). We found 604 matches at less than 10 arcmin distance from ComPRASS candidates. These candidates are classified as identified but not confirmed (class 2), unless they are confirmed from another catalogue. The information about these associations is included in the ComPRASS catalogue (see Appendix A).



**Fig. 6.** Positional criteria for matching ComPRASS candidates to SZ clusters. The distance between each ComPRASS candidate and its closest SZ cluster is plotted against their relative distance in terms of  $\theta_{500}$ . The white areas define the regions where the associations are considered to be correct. *Panel a*: complete view in logarithmic scale, and *panel b*: zoom of the low-distance region.



**Fig. 7.**  $M_{500}$  estimated from the joint detection for the 1060 ComPRASS candidates matching a confirmed SZ cluster vs. the published  $M_{500}$  of the corresponding cluster. The dotted red line indicates the line of zero intercept and unity slope. The solid blue line indicates the median ratio. The dashed blue lines indicate the interval of  $\pm 2.5\sigma$  around the median ratio. Outliers are highlighted with a blue (high estimated mass) or red (low estimated mass) square; see text. Green crosses indicate SPT clusters that are not in *Planck*. Cyan crosses indicate ACT clusters that are not in *Planck*. The inset panel shows the histogram of the logarithm of the ratio between the estimated and the published  $M_{500}$ .

### 3.3. Identification with optical clusters

We considered the seven optically selected cluster catalogues mentioned before: Abell, Zwicky, redMaPPer, MaxBCG, GMBCG, AMF, and WHL. For the first two catalogues, there is no information readily available about the mass of the clusters, and therefore we based our associations on positional criteria exclusively. The redMaPPer catalogue provides the richness of its clusters. The corresponding mass can be estimated using the richness–mass relation of [Rozo et al. \(2015\)](#). We used

this information, together with positional criteria to define our associations. For the four remaining SDSS-based catalogues, the richness is available, but the scatter of the richness–mass relation is too big, so we select our associations based directly on richness and positional criteria. In the remainder of this section, we describe in detail how the identification with the different catalogues was done.

#### 3.3.1. redMaPPer

The redMaPPer catalogue contains clusters that were detected on SDSS data by looking for spatial over-densities of red-sequence galaxies, and provides an estimate of the richness and the photometric redshift for all the objects, and the spectroscopic redshift of the brightest central galaxy (BCG) for some of them. We used the latest version of the publicly available redMaPPer catalogue (v6.3), which contains 26111 objects.

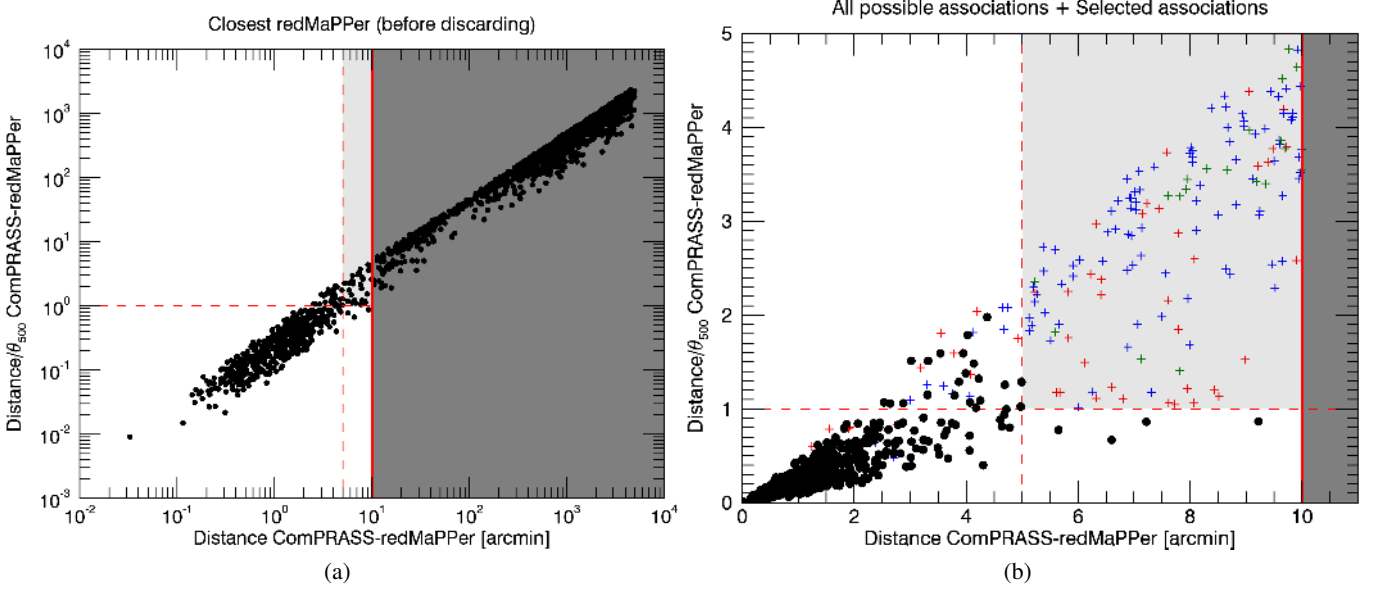
The association of ComPRASS candidates with redMaPPer clusters was done in several steps, following a similar procedure to the one used by [Planck Collaboration XXVII \(2016\)](#), but with some differences.

First, for each ComPRASS candidate, we look for all the redMaPPer clusters within a distance of 10 arcmin and we keep up to three redMaPPer clusters for each candidate. For each possible association, we estimate the mass  $M_J$  of the candidate by following the procedure described in Sect. 2.3 assuming the redshift of the corresponding redMaPPer cluster. We used the spectroscopic redshift when available, otherwise we took the photometric redshift. This yields up to three masses for each candidate. Then, as in the association with X-ray and SZ clusters, we discard all the pairs for which  $5' < d < 10'$  and  $d/\theta_{500} > 1$ . In this case,  $\theta_{500}$  is estimated from the richness and redshift reported in the redMaPPer catalogue, using the richness–SZ mass relation from [Rozo et al. \(2015\)](#):

$$\langle \ln \lambda | M_\lambda \rangle = a + \alpha \ln \left( \frac{M_\lambda}{M_p} \right), \quad (8)$$

with  $a = 4.572 \pm 0.021$ ,  $\alpha = 0.965 \pm 0.067$ ,  $M_p = 5.23 \times 10^{14} M_\odot$  and  $\sigma_{\ln \lambda | M_\lambda} = 0.266 \pm 0.017$ . This relation was calibrated





**Fig. 8.** Positional criteria for matching ComPRASS candidates to redMaPPer clusters. *Panel a*: closest redMaPPer cluster to each ComPRASS candidate. *Panel b*: all the possible redMaPPer counterparts within 10 arcmin of each ComPRASS candidate. Red, blue, and green symbols represent the closest, second-closest, and third-closest counterpart (if it exists), respectively. The black circles identify the pairs that are finally selected. The blue and red coloured crosses in the  $d < 5$  region correspond to associations that satisfy the positional criteria, but that are finally discarded by the richness cut.

with *Planck* SZ masses, and therefore we expect it to be well adapted to our validation. As done in the PSZ2 catalogue (Planck Collaboration XXVII 2016), we also discard the pairs for which the estimated mass of the candidate,  $M_J$ , and the mass of the redMaPPer counterpart,  $M_\lambda$ , calculated from the richness using Eq. (8), differ significantly. In particular, we used the following rule:  $|\ln(M_J/M_\lambda)| > 3\sigma$ , where  $\sigma = 0.399$  is the dispersion of  $\ln(M_J/M_\lambda)$ . This value was calculated from the closest association to each candidate, after discarding those for which  $5' < d < 10'$  and  $d/\theta_{500} > 1$ , and using an iterative procedure to discard the  $3\sigma$  outliers. We note that this value is higher than the one used by Planck Collaboration XXVII (2016), because we need to take into account both the dispersion of the richness-mass relation and the dispersion of our estimated mass  $M_J$  with respect to the true mass ( $\sigma^2 = \sigma_{\ln \lambda}^2 |M_\lambda/\alpha^2 + \sigma_{M_\lambda}^2$ ). After these two cuts in separation and mass, 538 candidates remain with an associated redMaPPer cluster, including 9 with two possible associations. Finally, if there is more than one possible association for a given candidate, we select the closest redMaPPer. We note that this choice is different from the one used in the PSZ2 catalogue, where the richest redMaPPer cluster was selected. This is justified because the position provided by the joint detection method is better than that provided by *Planck*. We also checked in detail these nine cases and, taking into account the position and size of the SZ and X-ray peaks together with the position of the two possible redMaPPer counterparts and the SDSS data on these regions, we concluded that the closest redMaPPer is always a consistent match.

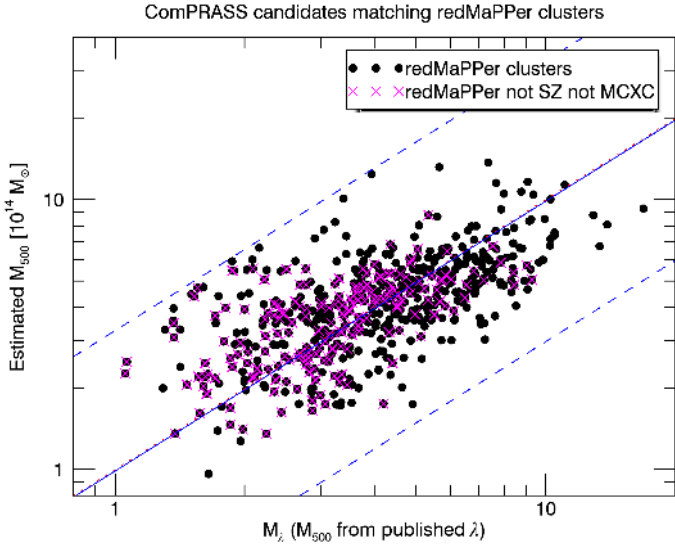
Figure 8 shows the distribution of possible redMaPPer counterparts for ComPRASS candidates. Panel a shows a scatter plot of the angular distance versus the relative distance between all the ComPRASS candidates and their closest redMaPPer counterpart. In this case, we do not observe clearly the two clouds corresponding to good and bad associations. Panel b shows all the possible redMaPPer counterparts within 10 arcmin of each ComPRASS candidate. The black circles identify the pairs that are finally selected according to the above procedure (angular separation cut and richness cut).

We see that the richness cut discards some of the associations that satisfy the angular separation criteria (blue and red colour crosses in the left part ( $d < 5$  arcmin) of Fig. 8b).

Following the procedure described in Sect. 2.3 and using the  $M_{500} - D_A^2 Y_{500}$  relation in Eq. (7), we estimated the mass  $M_{500}$  for the 538 detections matching redMaPPer clusters. Figure 9 shows the relation between the estimated mass and the mass of the corresponding clusters, calculated from the richness-mass relation of Rozo et al. (2015; Eq. (8)). This mass comparison shows that the ratio between the estimated mass and the published mass is very close to one, with a median value of 0.99.

### 3.3.2. Abell and Zwicky

Since there is no information about the mass of Abell (Abell et al. 1989) and Zwicky (Zwicky et al. 1961) clusters, and therefore about their size, we based our associations exclusively on the angular distance  $d$  between the ComPRASS candidates and its closest Abell or Zwicky cluster. If this distance is lower than 5 arcmin, we consider that the ComPRASS candidate is associated to the Abell or Zwicky cluster. Indeed, taking into account their number, the chance association within this radius is negligible (e.g.  $< 0.25\%$  for Abell clusters). Otherwise they are not associated. With this simple rule we find that 693 ComPRASS candidates are associated to an Abell cluster and 332 to a Zwicky cluster. These candidates are considered as confirmed (class 1), since Abell and Zwicky clusters are expected to be real clusters, even if the spectroscopic redshift is not available. This information is included in the ComPRASS catalogue (Table A.1). We note that 558 of the 693 candidates associated to an Abell cluster had already been confirmed with MCXC, SZ, or redMaPPer, 36 had been identified with SZ or MCXC candidates, and 99 had not yet been identified. For the 332 candidates associated to a Zwicky cluster, 283 had already been confirmed with MCXC, SZ, or redMaPPer, 11 had been identified with SZ candidates, and 38 had not yet been identified.



**Fig. 9.**  $M_{500}$  estimated from the joint detection for the 538 ComPRASS candidates matching a redMaPPer cluster vs. the  $M_{500}$  of the corresponding cluster calculated from the richness (Eq. (8)). The dotted red line indicates the line of zero intercept and unity slope. The solid blue line indicates the median ratio. The dashed blue lines indicate the interval of  $\pm 3\sigma$  around the median ratio. Magenta crosses indicate redMaPPer clusters that are not in *Planck*, SPT, ACT or MCXC.

### 3.3.3. Other SDSS-based catalogues

Apart from redMaPPer, we have used four additional cluster catalogues based on SDSS data: (1) the MaxBCG catalogue (13 823 objects, Koester et al. 2007); (2) the GMBCG catalogue (55 424 objects, Hao et al. 2010); (3) the AMF catalogue (69 173 objects, Szabo et al. 2011); and (4) the WHL catalogue (132 684 objects, Wen et al. 2012).

Unlike Abell and Zwicky, these catalogues contain many low-mass clusters, and therefore we cannot directly associate our candidates to any SDSS cluster found nearby. Namely, the mass of the optical counterpart should be compatible with that of the X-SZ candidate, as in the case of redMaPPer (Sect. 3.3.1). Each of the SDSS-based catalogues provides an estimated richness, which could be used to calculate a mass proxy using the richness–mass relation found in Planck Collaboration XXIX (2014). However, there is a large amount of scatter in this relation. To be conservative we decided to only consider the SDSS clusters with a richness above a certain threshold, as done in Planck Collaboration XXIX (2014).

To chose this richness threshold, we looked at all the potential associations between ComPRASS candidates (not yet confirmed with the previous validation) and SDSS clusters within an angular distance of 5 arcmin. Using optical images from PanStarrs, SDSS, and WISE, together with available ancillary X-ray images from *XMM-Newton* and *Swift*, and the filtered maps obtained as output of our detection method, we saw that all the potential associations with clusters with richness above 50 seem to be correct. Thus, the final criterion to associate a ComPRASS candidate with a SDSS cluster is an angular distance  $d$  between the ComPRASS candidate and the SDSS cluster lower than 5 arcmin; and a cluster homogenized richness<sup>4</sup>  $\lambda \geq 50$ . This

<sup>4</sup> We homogenized the richness estimates provided by the different SDSS-based catalogues to that of WHL, by applying the correcting factors found in Planck Collaboration XXIX (2014): 1.52, 1.75, and 0.74 for MaxBCG, GMBCG, and AMF, respectively.

gives 499 ComPRASS candidates associated with a cluster from the considered SDSS-based catalogues. They are classified as confirmed (class 1). We note that 483 of the 499 had already been confirmed with X-ray, SZ, redMaPPer, Abell or Zwicky catalogues, 9 had been identified with SZ candidates, and 7 had not yet been identified with previous catalogues.

### 3.3.4. High-redshift optical–infrared catalogues

Since the above-considered optical catalogues do not contain high-redshift clusters, we decided to consider three additional cluster catalogues which specifically targeted this type of object.

Firstly, we considered the sample of 1959 massive clusters of galaxies in the redshift range of  $0.7 < z < 1.0$  presented in Wen & Han (2018), which were found by searching around spectroscopically confirmed  $z > 0.7$  luminous red galaxies (LRGs) in SDSS. We found two matches within a 5 arcmin radius with ComPRASS candidates that were not confirmed in the previous external validation: PSZRX G180.69+46.41 corresponds to cluster J092829.4+410715 at  $z = 0.8194$  and had not been identified before; and PSZRX G271.62+61.69 corresponds to J115417.3+022124 at  $z = 0.7118$ , and had been identified (but not confirmed) with one of the clusters of Burenin (2017). The joint masses assuming these redshifts are  $M_{500} = (5.9 \pm 1.1) \times 10^{14} M_{\odot}$  and  $M_{500} = (7.0 \pm 1.3) \times 10^{14} M_{\odot}$ , respectively. The masses reported by Wen & Han (2018),  $M_{500} = 7.37 \times 10^{14} M_{\odot}$  and  $M_{500} = 3.05 \times 10^{14} M_{\odot}$ , respectively, do not have an associated error, but taking into account the possible uncertainties in the mass estimation at those high redshifts, they are compatible. By visually inspecting available ancillary images from WISE, PanSTARRS, SDSS, and *Swift*, both appear to be good associations. We therefore classified these two candidates as confirmed (class 1).

We found 14 additional matches with already confirmed ComPRASS candidates: the redshift in the catalogue of Wen & Han (2018) is compatible (difference lower than 0.06) with the redshift of the other ComPRASS counterpart for six of them, and is higher (difference bigger than 0.19) for eight of the matches. Appendix C.1 includes detailed notes on these eight associations; they correspond to either chance associations or real counterparts that contribute to our total joint signal.

Secondly, we considered the high-redshift clusters ( $0.7 < z < 1.5$ ) from the Massive and Distant Clusters of WISE Survey (MaDCoWS) presented in Gonzalez et al. (2019). We found eight matches of the ComPRASS candidates with this catalogue within a radius of 5 arcmin, one of them corresponding to one of the “identified but not confirmed” candidates: candidate PSZRX G113.26+48.41, already identified with PSZ2 G113.27+48.39, and situated at 0.7 arcmin from MOO J1359+6725. Since the redshift and richness of MOO J1359+6725 are not available, our ComPRASS candidate remains not confirmed. The other seven matches correspond to already confirmed (class 1) candidates. Two of them have a redshift that is higher than the redshift of the other candidate counterpart, and they are located at larger angular separation. They may therefore correspond to chance associations. Appendix C.2 includes detailed notes on these two associations. The remaining five matches do not have a redshift in the MaDCoWS catalogue.

Finally, we took the 44 high- $z$  ( $0.5 < z < 1.0$ ) clusters presented in Buddendiek et al. (2015), initially selected by cross-correlating the RASS faint and bright source catalogues with red galaxies from SDSS DR8, and then followed-up with optical telescopes. We found 12 matches of the ComPRASS candidates with this catalogue within a radius of 5 arcmin. Two of them

**Table 1.** Some interesting galaxy clusters found close to the new candidates (class 3) of the ComPRASS catalogue.

Id.	ComPRASS name	Name	Redshift	Separation [arcmin]
779	PSZRX G115.09+28.55	PLCKESZ G115.12+28.56	0.169	1.47
1307	PSZRX G210.74+08.03	PSZ1 G210.76+08.02	0.296	1.56
1463	PSZRX G234.29+20.47	PLCKESZ G234.2-20.5	0.27	0.36
1556	PSZRX G246.41+67.77	ZwCl 4333	0.08057	0.87
1579	PSZRX G249.37+40.82	RX J1020.5-0550	0.404	1.25
1587	PSZRX G249.99+24.23	PSZ1 G250.02+24.15	0.400	4.62
1638	PSZRX G255.85+41.58	SPT-CL J0438-4907	0.24	0.72
2018	PSZRX G302.72+25.82	Abell 3527-bis	0.20	0.99

**Notes.** The search was done in the NED and SIMBAD databases.

correspond to unidentified candidates (class 3), and ten to already confirmed candidates (class 1).

High-redshift cluster CIG-J120958.9+495352 (Buddendiek et al. 2015), at  $z = 0.902$ , is close to candidate PSZRX G139.37+65.89. The joint mass assuming this redshift is  $M_{500} = 5.64 \times 10^{14} M_{\odot}$ , which is very close to the mass reported by Buddendiek et al. (2015) ( $M_{500} = (5.3 \pm 1.5) \times 10^{14} M_{\odot}$ ). We therefore associated these objects and classified our candidate as confirmed (class 1).

Cluster CIG-J231520.6+090711, at  $z = 0.725$ , is at less than 1 arcmin from candidate PSZRX G086.90+46.91. Although Buddendiek et al. (2015) does not provide the mass of this cluster, it seems that the association is correct. We therefore classified our candidate as confirmed (class 1).

For the ten matches with an already confirmed candidate, the redshift in the catalogue of Buddendiek is compatible (difference smaller than 0.06) with the redshift of the other ComPRASS counterpart in six of the cases, and is higher (with a difference greater than 0.08) for four of the matches. We studied these four cases in detail (see notes in Appendix C.3) and found that two of them seem to be correctly associated both to the Buddendiek clusters and to the previously found counterparts (they are indeed the same object), and that the redshift provided by Buddendiek is more accurate. In the other two matches there is a superposition of two different clusters: the Buddendiek cluster (closer to the detection) and a redMaPPer cluster (more distant). In these cases, the association with the higher- $z$  Buddendiek cluster is confirmed, whereas the lower- $z$  redMaPPer cluster may contribute to total signal or may simply be a chance association.

### 3.4. Cross-identification with NED and SIMBAD databases

Finally, we used NED<sup>5</sup> and SIMBAD<sup>6</sup> databases to avoid missing a few additional associations. In particular, we looked for other known galaxy clusters around the ComPRASS candidates that have not yet been identified with the previous validation. Considering a search radius of 5 arcmin we found possible cluster counterparts for 60 ComPRASS candidates. Since most of these objects are small-mass optical clusters, it is difficult to determine if they really are the counterparts of our candidates. We focus instead on the ones that could be missed associations or particularly interesting cases.

In particular, we found four *Planck* clusters, one SPT cluster, one Zwicky cluster, one Abell cluster, and one X-ray cluster. Table 1 summarises these possible missed associations. We stud-

ied these cases individually to determine why we missed some of them and whether or not they are good associations.

PLCKESZ G115.12+28.56 was not included in the published *Planck* catalogues, but it was confirmed with optical follow-up observations at the Canary Island observatories (Planck Collaboration Int. XXXVI 2016). The distance between this cluster and ComPRASS candidate PSZRX G115.09+28.55 is 1.47 arcmin, which means that they are associated. Consequently, we classified this candidate as confirmed.

PSZ1 G210.76+08.02 was not included in the published PSZ1 catalogue because it was below the  $4.5\sigma$  limit. However, it was confirmed with optical observations from the Russian-Turkish 1.5 m telescope as part of the follow-up programme of the *Planck* collaboration (Planck Collaboration XXVI 2015). By looking at the *Swift* observation of this cluster and knowing that it is 1.56 arcmin away from ComPRASS candidate PSZRX G210.74+08.03, we can conclude that they are associated. We therefore classified this candidate as confirmed.

PLCKESZ G234.2-20.5 was not included in the published ESZ catalogue because it was detected with a S/N lower than 6, which was the limit for the published catalogue. However, it was confirmed with *XMM-Newton* observations in Planck Collaboration I (2012). By looking at the available *XMM-Newton* and *Swift* observations of this cluster and knowing that it is only 0.36 arcmin away from ComPRASS candidate PSZRX G234.29+20.47, we can conclude that they are the same cluster. Therefore, we classified this candidate as confirmed. Considering the redshift of PLCKESZ G234.2-20.5 ( $z = 0.27$ ), the mass estimated for PSZRX G234.29+20.47 is  $M_{500} = 5.07 \times 10^{14} M_{\odot}$ , which is compatible with the mass of PLCKESZ G234.2-20.5 reported in Planck Collaboration I (2012) ( $M_{500} = (4.5 \pm 0.1) \times 10^{14} M_{\odot}$ ) in the terms considered in Fig. 7.

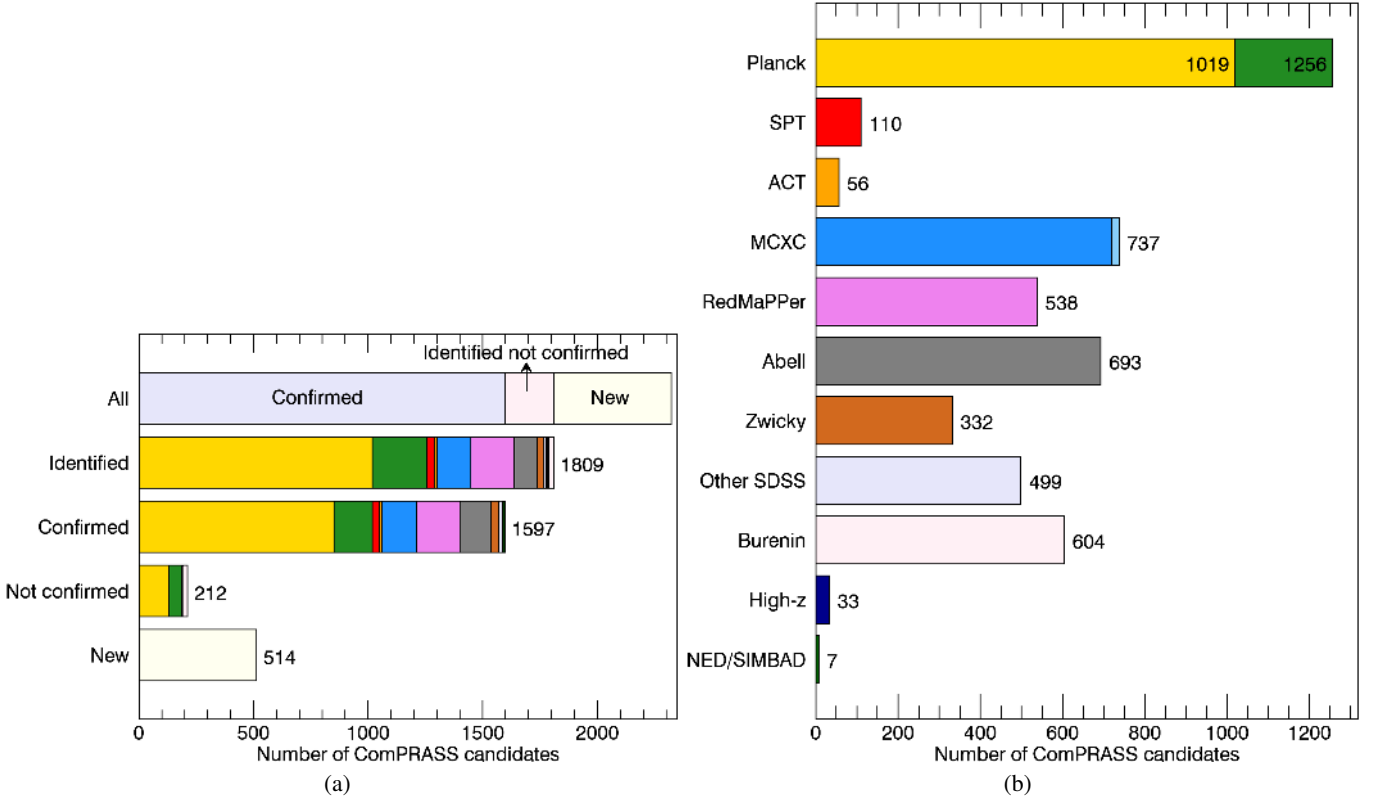
PSZ1 G250.02+24.15 was considered in Sect. 3.1, but not associated with candidate 1587 because it was one of the matches falling in the grey zone of Fig. 6 (see notes on Appendix B).

SPT-CL J0438-4907 was not included in the published SPT catalogue of Bleem et al. (2015) because its significance was lower than 4.5, which was the limit for the published catalogue. However, it was detected at lower significance and confirmed with optical observations in Saro et al. (2015). The presence of this cluster at only 0.72 arcmin from candidate PSZRX G255.85+41.58 and the *Swift* observation of this cluster indicate that this candidate is a real cluster (Tarrío et al. 2018). Moreover, the mass estimated assuming the redshift of the SPT cluster is  $M_{500} = (3.0 \pm 0.5) \times 10^{14} M_{\odot}$ , very close to the mass published by Saro et al. (2015) for the SPT cluster ( $M_{500} = (3.13 \pm 0.81) \times 10^{14} M_{\odot}$ ). We therefore classified this candidate as confirmed.

<sup>5</sup> <http://nedwww.ipac.caltech.edu/>

<sup>6</sup> <http://simbad.u-strasbg.fr/simbad>





**Fig. 10.** *Panel a:* number of candidates in the ComPRASS catalogue that were either confirmed or not and/or identified with the external validation described in Sect. 3. *Confirmed* (class 1) refers to the candidates that are associated to an already confirmed cluster; *Identified not confirmed* (class 2) indicates the candidates that are associated to candidates in X-ray or SZ catalogues that have not yet been confirmed; and *New* (class 3) represents the candidates that are not associated to any cluster or cluster candidate. *Identified* includes candidates in class 1 and class 2. The different colours in the three middle bars correspond to the catalogue in which they were confirmed and/or identified: MMF3 (yellow), *Planck* (green), SPT (red), ACT (orange), MCXC (blue), redMaPPer (violet), Abell (grey), Zwicky (brown), other SDSS-based catalogues (lavender), high-redshift catalogues (dark blue), additional clusters found in NED/SIMBAD (dark green), and Burenin (2017) (pale pink). Since one candidate may have been identified with several catalogues, the indicated colour refers to the first one in the previous list. For example, red indicates the candidates identified/confirmed by SPT and not identified/confirmed by *Planck*. *Panel b:* number of candidates in the ComPRASS catalogue that were identified with the different catalogues. This includes both class 1 and class 2 candidates. For the *Planck* catalogues, the candidates identified with MMF3 are indicated in yellow, while the others are shown in green. For the MCXC catalogue, RASS and serendipitous clusters are indicated in medium and light blue, respectively.

The cluster ZwCl 4333 was not associated to ComPRASS candidate PSZRX G246.41+67.77 because they are separated by 13.73 arcmin. However, the position of ZwCl 4333 in SIMBAD is not the same as in the Zwicky catalogue. With this updated position, ZwCl 4333 and PSZRX G246.41+67.77 are separated by 0.87 arcmin, and therefore would be associated. We therefore classified this candidate as confirmed.

Abell 3527-bis is not a cluster in the Abell catalogue. It was discovered by de Gasperin et al. (2017) and named Abell 3527-bis due to its proximity (20 arcmin) and similar redshift ( $z = 0.20$ ) to the cluster Abell 3527. Candidate PSZRX G302.72+25.82 is at less than 1 arcmin from this cluster. Furthermore, the mass estimated assuming the redshift of this cluster is  $M_{500} = 3.02 \times 10^{14} M_{\odot}$ , very close to the mass published by de Gasperin et al. (2017) for Abell 3527-bis ( $M_{500} = (3.3 \pm 0.3) \times 10^{14} M_{\odot}$ ). Therefore, we decided to classify this candidate as confirmed (class 1).

X-ray source RX J1020.5–0550 is galaxy cluster [ATZ98] C027 (Appenzeller et al. 1998), at redshift  $z = 0.404$ . Candidate PSZRX G249.37+40.82 is only at 1.25 arcmin from this cluster and the detection peak coincides with the optical cluster. We therefore classified this candidate as confirmed.

### 3.5. Validation summary

Figure 10 and Table 2 summarise the results of the identification of the ComPRASS candidates with the considered SZ, X-ray, and optical cluster catalogues. From the 2323 candidates of the catalogue, 1597 correspond to known confirmed clusters, 212 are associated to a yet unconfirmed cluster candidate and 514 correspond to new objects. These  $212 + 514 = 726$  unconfirmed candidates could either be real clusters or false detections. Considering that the value of the purity estimated for the X-ray–SZ detection method in the SPT footprint ( $>83.1\%$ ) (Tarrío et al. 2018) is approximately valid for the whole sky, we expect to have at least 1930 real clusters in the catalogue, that is, at least 333 real clusters among the 726 unconfirmed candidates.

Figure 11 shows the distribution in the  $M_{500}-z$  plane of the candidates in the ComPRASS catalogue that were validated from the external catalogues. The mass  $M_{500}^{XZ}$  is the mass estimated by the X-ray–SZ detection method when the redshift of the associated cluster is assumed. Each cluster is colour-coded according to the catalogue in which it was identified (see legend). If the cluster belongs to several catalogues, the following order is chosen: MMF3, other *Planck*, SPT, ACT, MCXC,

**Table 2.** Number of previously known clusters or cluster candidates that are associated to the ComPRASS candidates.

Cluster catalogue	Clusters in the considered region	Clusters detected	Percentage (%)
all MMF3	1202	1019	84.8
confirmed MMF3	909	850	93.5
MMF3 candidates	293	169	57.7
all PSZ2	1554	1200	77.2
all <i>Planck</i>	1810	1256	69.4
all SPT	647	110	17.0
ACT	206	56	27.2
MCXC	1751	737	42.1
RASS	1294	719	55.6
Serendipitous	457	18	3.9
redMaPPer	25 333	538	2.1
Abell	4995	693	13.9
Zwicky	8873	332	3.7
MCXC not SZ	1122	150	13.4
all SPT not <i>Planck</i>	558	31	5.6
all SZ	2522	1299	51.5
all SZ + MCXC	3643	1449	39.8
Total number of detections		2323	
Detections matching any cluster candidate		1809	
New detections		514	
Detections matching a confirmed cluster		1597	
Detections matching an unconfirmed candidate		212	

**Notes.** *Planck* refers to the combination of the three *Planck* catalogues (Planck Collaboration VIII 2011; Planck Collaboration XXIX 2014; Planck Collaboration XXVII 2016), whereas PSZ2 refers only to the last one. MMF3 is the subsample of objects in the PSZ2 catalogue that were detected using the MMF3 detection algorithm. RASS refers to the subsample of objects in the MCXC catalogue that were detected from RASS observations. SZ refers to the combination of all the SZ catalogues (*Planck*, SPT, and ACT).

redMaPPer, Abell, other SDSS-based catalogues, and high-redshift catalogues. This figure evidences the ability to recover clusters that were not included in the MMF3 catalogue (yellow circles). In particular, at lower redshifts we gain lower-mass clusters discovered with X-rays, optical (redMaPPer), or even deeper SZ surveys (SPT and ACT). At high redshift, we also recover some objects detected by deeper SZ surveys or deep X-ray observations. In particular, above  $z = 0.75$ , we recover one cluster detected in the deep X-ray North Ecliptic Pole (NEP) survey (Henry et al. 2006), and four clusters detected by SPT and/or ACT. In comparison to MMF3, these additional clusters tend to populate the low-mass regions at any redshift, pushing the detection limit of MMF3 towards a lower mass for any redshift.

### 3.6. Redshift assembly

The ComPRASS catalogue provides the redshift of all the candidates that have been confirmed by a cluster with an available redshift. If the ComPRASS candidate is associated to only one cluster, we provide the redshift of this cluster. In the case of multiple associations with different redshifts, we provide a unique redshift, which is chosen by applying the following order of priority: MCXC, SZ, redMaPPer (favouring spectroscopic over photometric redshift, when available), Abell, WHL, AMF, MAXBCG, GMBCG, and finally, the considered high-redshift catalogues. We have privileged homogeneity for the sources of redshift rather than a case-by-case assembly of the most accurate redshift. We have checked that all the redshifts available for a given candidate are in general consistent. Some notes on specific cases where a discrepancy was found are included in Appendix C.

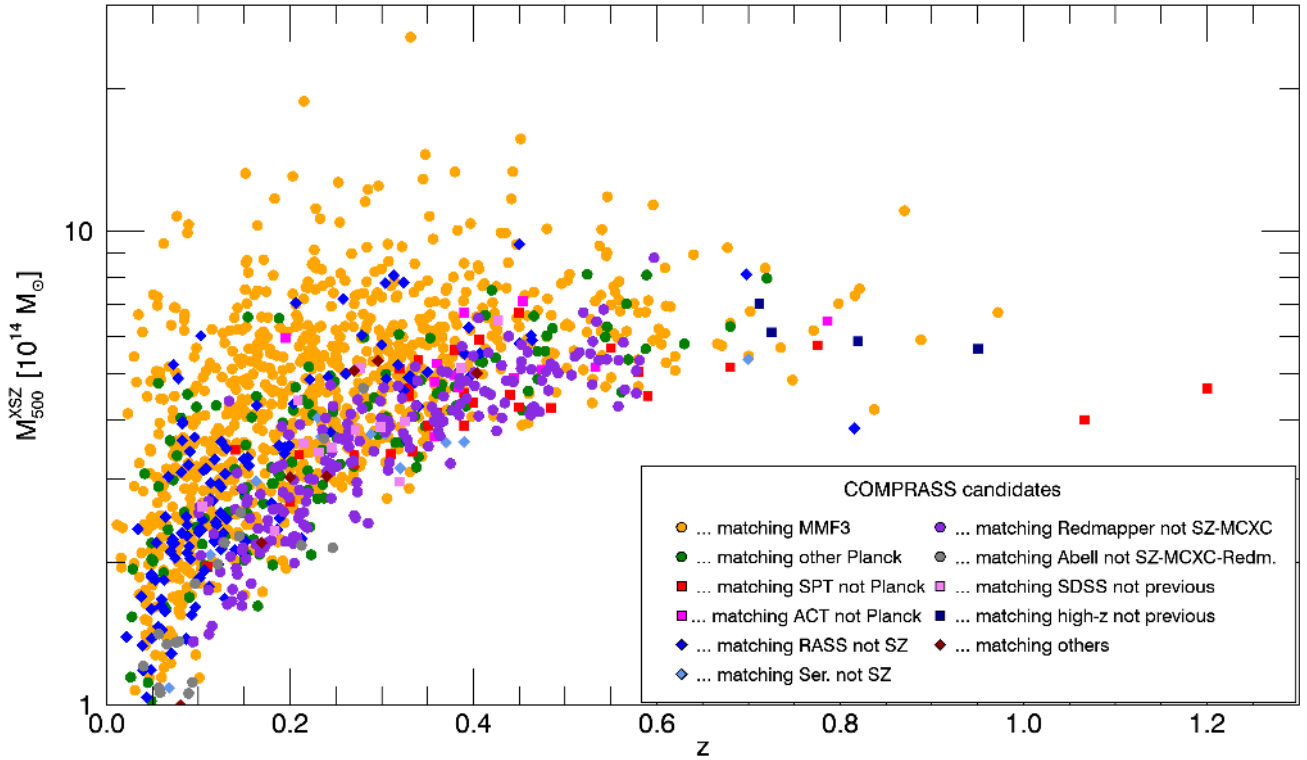
## 4. Evaluation

### 4.1. Comparison with MMF3 clusters

Since the X-ray–SZ detection method was built as an extension of the MMF3 detection method, we expect it to have a better performance than that of its predecessor.

As shown in Table 2, the proposed method is able to recover 84.8% of the MMF3 candidates (confirmed and not confirmed) included in the PSZ2 catalogue and situated in the considered region of the sky. There are 183 MMF3 candidates missing, 59 of which are confirmed clusters. Six of these 59 confirmed MMF3 clusters are detected within a radius between 5 and 10 arcmin but were not considered to be associated (grey zone in Fig. 6b). Forty-two of the 59 missing clusters were initially detected (in the second phase), but then 20 were discarded because the joint S/N was not high enough and 22 were discarded because they had  $(S/N)_{SZ} < 3$ . One was detected in the first phase, but lost in the second phase because the joint S/N is below the initial threshold  $q = 4$ . Finally, the remaining ten were not detected in the first phase of the algorithm due to masking by another nearby detection (6) or due to a joint S/N below the initial threshold  $q = 4$  (4). Figure 12 shows the distribution of the detected and missing MMF3 confirmed clusters in the mass-redshift plane.

The 124 MMF3 unconfirmed candidates that do not appear in our candidate list are missing for several reasons: 68 were initially detected in the second phase, but then 45 were discarded because the joint S/N was not high enough, 22 were discarded because they had  $(S/N)_{SZ} < 3$  and 1 was discarded because it was in the masked region (survey mask). A further 10 were detected in the first phase, but lost in the second phase because



**Fig. 11.** Distribution in the  $M_{500}-z$  plane of validated ComPRASS candidates. Each candidate is colour-coded according to its associated cluster. Yellow circles represent ComPRASS candidates matching confirmed MMF3 clusters, green circles represent ComPRASS candidates matching other confirmed *Planck* clusters (not MMF3), red squares represent ComPRASS candidates matching confirmed SPT clusters not detected by *Planck*, magenta squares represent ComPRASS candidates matching confirmed ACT clusters not detected by *Planck* or SPT, blue diamonds represent ComPRASS candidates matching confirmed MCXC clusters that do not match any of the previously mentioned catalogues (dark blue for RASS, light blue for serendipitous), purple circles correspond to candidates matching redMaPPer clusters not in SZ or MCXC catalogues, grey circles correspond to candidates matching an Abell cluster not in the previous catalogues, pink squares correspond to candidates matching other SDSS clusters not in the previous catalogues, dark blue squares correspond to candidates matching other high-redshift clusters from Wen & Han (2018) and Buddendiek et al. (2015) not in the previous catalogues, and dark red diamonds correspond to candidates matching the additional clusters found in NED and SIMBAD databases. The mass is estimated from the X-SZ signal and the cluster  $z$  as described in Sect. 2.3.

the joint S/N is below the initial threshold  $q = 4$ . Finally, the remaining 46 were not detected in the first phase of the algorithm due to masking by another nearby detection or due to a joint S/N below the initial threshold  $q = 4$ .

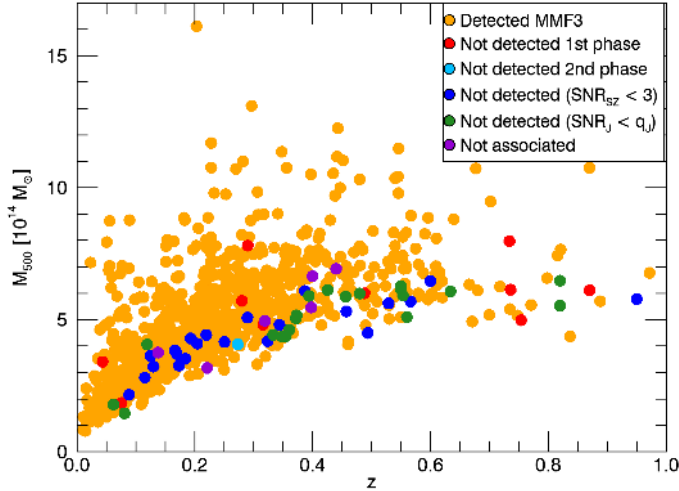
There are a total of 44 MMF3 objects (22 confirmed, 22 not confirmed) that we discard because  $(S/N)_{SZ} < 3$ . This may seem unexpected, since MMF3 objects have  $S/N > 4.5$  in the original extraction done by Planck Collaboration XXVII (2016). However, our value of  $(S/N)_{SZ}$  is measured at the position and size determined by the joint detection method, whereas the S/N value of original MMF3 objects is measured at the position and size determined by the MMF3 algorithm. Another difference with respect to MMF3 is that the resolution of the maps is different, since we are using a resolution of  $0.86 \text{ arcmin pixel}^{-1}$  instead of the original resolution of *Planck* maps ( $1.72 \text{ arcmin pixel}^{-1}$ ). We used the MMF3 method of Planck Collaboration XXVII (2016) to extract the S/N of these 44 objects from *Planck* maps in different ways. First, we centred the extraction at the position determined by the joint detection method, but we set the size as a free parameter. Second, we centred the extraction at the MMF3 position and set the size as a free parameter. Finally, we used the maps with the original resolution ( $1.72 \text{ arcmin pixel}^{-1}$ ) and extracted the signal at the MMF3 position with free size. We found that 16 of the 44 MMF3 are lost due to the different size determined by the joint method, 17 are lost due to the different position determined by the joint method, and 11 are lost

because of the different resolution (plus the different position and size).

Even though the ComPRASS catalogue misses a small fraction of the MMF3 confirmed clusters (6.5%), it includes other previously known clusters that are missed by MMF3 (see Table 2 and Fig. 11). In particular it includes 168 additional *Planck* clusters, 42 SPT clusters, and ACT clusters that were not detected by *Planck*, 150 MCXC clusters that were not detected by *Planck*, SPT or ACT, and 382 additional optical clusters not included in the considered SZ and X-ray catalogues. The overall effect is an improvement of the purity-detection efficiency performance with respect to the reference catalogue PSZ2-MMF3 (see also Fig. 13 of Tarrío et al. 2018).

Finally, it is interesting to understand the quality of the MMF3 detections that were missed or not by ComPRASS. The PSZ2 catalogue provides a quality flag,  $Q_{\text{NEURAL}}$ , that indicates the quality of each detection. A value of  $Q_{\text{NEURAL}} < 0.4$  identifies low-reliability detections with a high degree of success (Planck Collaboration XXVII 2016). For the 1202 MMF3 candidates that fall in our unmasked region of the sky, the percentage of low-reliability detections is 4.4%. If we consider only the 1019 MMF3 candidates contained in the ComPRASS catalogue, the percentage of low-reliability detections falls down to 1.6%, almost three times lower. The percentage of low-reliability candidates among the 183 MMF3 candidates that were missed by ComPRASS is 20.2%. This indicates the





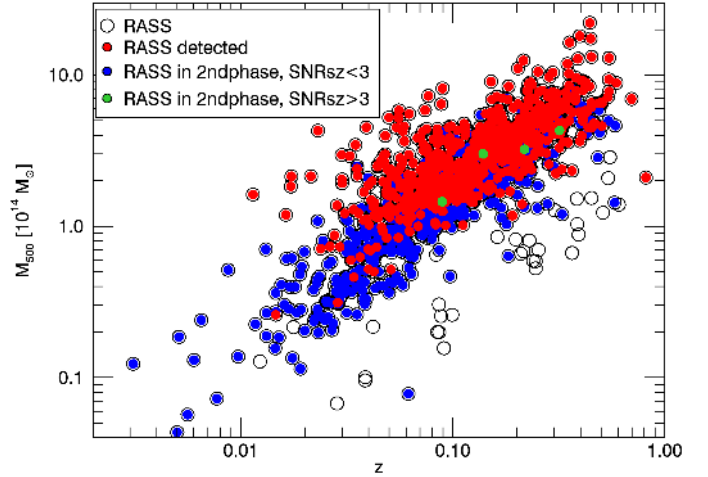
**Fig. 12.** Mass and redshift of confirmed clusters in the MMF3 catalogue. Yellow-filled circles represent MMF3 clusters that are detected by the joint detection algorithm; red-filled circles represent MMF3 clusters that were lost in the first phase of the algorithm; the light blue filled circle represents the MMF3 cluster that was lost in the second phase of the algorithm; dark blue filled circles represent MMF3 clusters detected in the second phase of the algorithm, but discarded due to a low  $(S/N)_{SZ}$ ; and green-filled circles represent MMF3 clusters detected in the second phase of the algorithm, but discarded because the joint S/N was not high enough.

ability of the ComPRASS catalogue to clean the MMF3 catalogue from probably false detections, thanks to the incorporation of the X-ray information.

#### 4.2. Comparison with RASS clusters

Since the proposed joint detection method uses RASS observations, it is interesting to check whether it is able to recover known clusters that have been detected using the same observations. Table 2 shows that we detect 719 of the 1294 RASS clusters situated in the considered region (RASS exposure time greater than 100 s, outside the PSZ2 masked region), which corresponds to 55.6%. There are 575 RASS clusters that we do not recover. Most of them (512) were in fact included in the list of detections provided by the second phase of the algorithm, but discarded for various reasons: 508 were discarded because their  $(S/N)_{SZ}$  was lower than 3, a further 2 were discarded because the RASS exposure time at the position of detection is lower than 100 s, 1 is discarded because it is outside the SZ mask, and 1 was discarded because the joint S/N does not reach the adaptive threshold.

There are 63 RASS clusters that were not originally detected by the joint algorithm. Most of them (49) were not detected because their joint S/N does not reach the threshold of  $q = 4$ . Although these clusters are visible in RASS, the joint signal is not sufficiently high to be detected because their SZ signal is very faint: when we extract them from *Planck* maps using the MMF3 method of *Planck Collaboration XXVII (2016)* they have a very low S/N, lower than 3 for 48 of them (meaning that they would have been discarded anyway after the second phase of the algorithm) and 3.3 for the last one. The remaining 14 were masked by a nearby brighter source during the iterative peak detection procedure in the first phase of the algorithm. The clusters that are lost due to masking by close-by detections could be recovered in the future with an improved version of the iterative peak-finding process. Two possible solutions would be to use a smaller mask or to simultaneously search for all the peaks above



**Fig. 13.** Mass and redshift of the clusters in the RASS catalogue. Open circles represent RASS clusters in the considered region; red filled circles represent RASS clusters that are detected by the joint detection algorithm; blue filled circles represent RASS clusters that were detected in the second phase of the algorithm, but discarded due to a low  $(S/N)_{SZ}$ ; and the green-filled circles represent the RASS cluster that was detected in the second phase of the algorithm, but discarded due to other reasons (see text).

a given threshold. In any case, a procedure to deblend sources would be needed afterwards.

In summary, the joint detection method is able to recover almost all the RASS clusters, as expected, but some of them are discarded later with the  $(S/N)_{SZ}$  threshold that we impose to remove AGN detections. Figure 13 illustrates this comparison by showing the RASS clusters and the joint detections in the mass-redshift plane. The  $(S/N)_{SZ}$  threshold mainly eliminates low-mass clusters that are detectable in RASS, especially at low redshift (blue filled circles). RASS clusters with very low mass at any redshift (empty circles) are mainly lost because their joint S/N does not reach the threshold of  $q = 4$ , as explained before.

The clusters of the Massive Cluster Survey (MACS; *Ebeling et al. 2001, 2007, 2010; Repp & Ebeling 2018*) are a particular subset of RASS clusters that were confirmed by a systematic optical follow-up of X-ray sources from the ROSAT bright source catalogues (*Voges et al. 1999; Boller et al. 2016*). Thus, they are clusters that are visible in the RASS maps, but which do not have a clear extended emission to be automatically detected as clusters. There are 109 MACS clusters in the considered region of the sky. The ComPRASS catalogue contains 75 of these clusters (68.8%), whereas the PSZ2 catalogue contains only 55 (50.4%). This shows that the addition of the X-ray information allows for the detection of more MACS clusters than those detected by *Planck*. The 34 MACS clusters that do not appear in the ComPRASS catalogue are missing due to various reasons: 1 was masked by a nearby brighter source in the first phase of the algorithm, 2 were lost in the second phase because the joint S/N did not reach the threshold of  $q = 4$ , and the remaining 31 were detected in the second phase but later discarded (1 because it was inside the SZ mask and 30 because their  $(S/N)_{SZ}$  was lower than 3).

#### 4.3. Invalidated *Planck* candidates

*Planck* published cluster catalogues contain confirmed clusters as well as not-yet-confirmed candidates that could correspond

to false detections or to low-mass haloes boosted by SZ noise peaks. Since the publication of these catalogues, some of the candidates have been invalidated by specific follow-up observations. In particular, the MegaCam follow-up of van der Burg et al. (2016) has invalidated 3 PSZ2 candidates and 8 PSZ1 candidates; and the optical follow-up observational programme developed at Roque de los Muchachos Observatory (Barrena et al. 2018) has invalidated (labelled as “Non detections”, indicating that the counterpart of the SZ detection was not found) 49 PSZ1 candidates, 2 of which were already in the list of van der Burg et al. (2016).

The ComPRASS catalogue contains only 5 of these 58 invalidated objects: PSZ2 G037.67+15.71, PSZ2 G157.07–33.63, PSZ1 G029.79–17.37 (or PSZ2 G029.80–17.40), PSZ1 G044.83+10.02 (or PSZ2 G044.83+10.02) and PSZ1 G096.44–10.40. Since the percentage of *Planck* candidates in the ComPRASS catalogue is 69.4% and the percentage of invalidated *Planck* candidates is only 8.6%, this indicates that we tend to clean the *Planck* catalogues from false candidates or low-mass systems. For this reason, we expect that the mass proxy provided by the ComPRASS catalogue will be less biased than the SZ mass proxy of the *Planck* catalogues at low significance values.

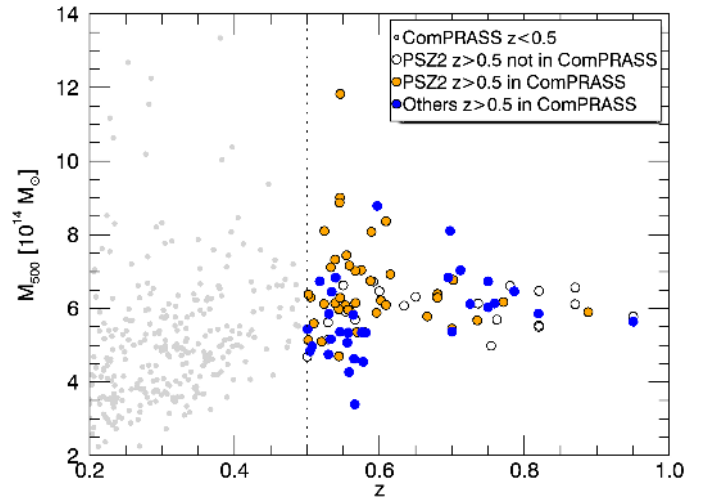
#### 4.4. Performance at high redshift

In this section we evaluate the performance of the ComPRASS catalogue in the high-redshift regime ( $z > 0.5$ ). The catalogue includes 125 confirmed clusters in this regime, with more than half (73) located in the SDSS footprint, which covers one third of the sky. The higher proportion of high-redshift clusters in the SDSS region is simply due to the fact that most of the optical catalogues considered during the validation cover only this area. We therefore focus on this region to compare the performance of the ComPRASS catalogue to that of the PSZ2 catalogue. We note however that SDSS-based catalogues are mostly limited to  $z < 0.6$ , so we certainly have an incomplete list of high-redshift ComPRASS clusters.

Figure 14 shows the ComPRASS and PSZ2 high-redshift clusters in the mass–redshift plane. ComPRASS and PSZ2 contain 41 high- $z$  clusters in common (orange circles), PSZ2 contains 20 high- $z$  clusters that are not detected in ComPRASS (empty circles), and ComPRASS contains 32 high- $z$  clusters not detected in PSZ2 (blue circles). The amount of high-redshift clusters that ComPRASS adds is slightly higher than the amount of high-redshift clusters that it loses with respect to the PSZ2 catalogue. However, due to the limited size of the sample and the lack of a systematic follow-up (some of the unconfirmed ComPRASS candidates could indeed be high-redshift clusters), it is difficult to extrapolate this behaviour to other regions of the sky.

#### 4.5. Minimum expected purity

Given that the external validation was performed with cluster catalogues that do not cover all the sky homogeneously, it is interesting to see the differences between the regions of the sky that are better and less covered by the considered catalogues. Two particularly well-covered regions are the SDSS footprint and the SPT footprint. All the clusters in the five considered SDSS-based catalogues (redMaPPer, MAXBCG, GMBCG, AMF, WHL) are contained in the former; whereas the SPT clusters are all located in the latter. This means that our validation is more exhaustive in these regions.



**Fig. 14.** Mass and redshift of known high-redshift ( $z > 0.5$ ) clusters in the SDSS footprint. Orange circles represent ComPRASS candidates that have been associated to a PSZ2 cluster at  $z > 0.5$ . Blue circles represent ComPRASS candidates that have been associated to other  $z > 0.5$  clusters. Empty circles represent PSZ2 clusters at  $z > 0.5$  that are not included in the ComPRASS catalogue. Smaller grey circles show ComPRASS candidates confirmed at  $z < 0.5$ . The mass represented in the figure corresponds to the joint mass estimated by the X-ray–SZ method for the grey, orange, and blue circles; and to the published mass in the PSZ2 catalogue for the empty circles.

Table 3 summarises the number and percentage of candidates that are classified in the three different classes (confirmed, identified not confirmed, and new) in different regions: inside and outside the SDSS footprint, inside and outside the SPT footprint, inside and outside the two footprints combined, and in all the sky. The percentage of candidates that are classified as confirmed is higher in the SDSS and SPT regions (84.5%) than in the rest of the sky (53.4%). Since the chances that a ComPRASS candidate is not a real cluster should not depend on being or not in one of these regions, we expect to have at least the same percentage of real clusters in the rest of the sky as in the SDSS and SPT footprints. Therefore, with future validation efforts, we expect to find at least  $(1175 \times 0.845) - 627 = 365$  real clusters among the  $383 + 165 = 548$  not confirmed and new candidates situated outside these two footprints.

## 5. Conclusions

We present ComPRASS, the first all-sky catalogue of X-ray–SZ sources constructed using a joint X-ray–SZ detection method (Tarrío et al. 2018) on RASS and *Planck* maps. This joint detection method can be seen as an evolution of the MMF3 detection method, one of the MMF methods used to detect clusters from *Planck* observations, which incorporates X-ray observations to improve the detection performance.

The catalogue, which contains 2323 cluster candidates, has been validated by careful cross-identification of its candidates with existing X-ray, SZ, and optical cluster catalogues. With this validation, we have classified the ComPRASS candidates into three classes: (1) Confirmed (1597), which corresponds to the candidates that are associated to an already known cluster, (2) Identified not confirmed (212), which includes the candidates associated to an already known candidate from SZ or X-ray catalogues that has not yet been confirmed as a real cluster, and (3) New (514), which contains the rest of the candidates

**Table 3.** Number and percentage of candidates that are classified in the three different classes (confirmed, identified not confirmed, and new) in different regions: inside and outside the SDSS footprint; inside and outside the SPT footprint; inside and outside the SDSS and SPT footprints combined; and in all sky.

	SDSS footprint				SPT footprint			
	Inside		Outside		Inside		Outside	
	Number	Percentage	Number	Percentage	Number	Percentage	Number	Percentage
Total	923		1400		225		2098	
Confirmed	786	85.2	811	57.9	184	81.8	1413	67.3
Not confirmed	41	4.4	171	12.2	6	2.7	206	9.8
New	96	10.4	418	29.9	35	15.6	479	22.8
	SDSS/SPT footprint				All sky			
	Inside		Outside		Number	Percentage		
	Number	Percentage	Number	Percentage	Number	Percentage		
Total	1148		1175		2323			
Confirmed	970	84.5	627	53.4	1597	68.7		
Not confirmed	47	4.1	165	14.0	212	9.1		
New	131	11.4	383	32.6	514	22.1		

that have not yet been associated to any known cluster or cluster candidate.

With respect to the reference catalogue PSZ2-MMF3, the ComPRASS catalogue is simultaneously more pure and more complete, since although it misses a small fraction of the MMF3 clusters (6.5%), it includes many other additional clusters (747) that are missed by MMF3. In particular, it includes 382 optical clusters not included in purely SZ or purely X-ray catalogues. In addition, the ComPRASS catalogue contains a much smaller percentage of low-reliability detections than the MMF3 catalogue, and a much smaller percentage of invalidated *Planck* candidates than the average recovery rate of *Planck* candidates. This indicates that ComPRASS cleans the *Planck* catalogues from false candidates (or low-mass systems). We thus expect to provide a less-biased mass proxy than the one given in the *Planck* catalogues, especially at low significance values where the SZ mass proxy is more affected by Eddington bias (van der Burg et al. 2016).

Regarding the X-ray catalogues, we recover 55.6% of the RASS clusters. The rest, which are mostly at low mass and low redshift, are lost mainly due to the cut imposed to avoid AGN detections, but they could be incorporated by relaxing this constraint. Interestingly, the ComPRASS catalogue contains 68.8% of the considered MACS clusters (more than in PSZ2). Since these clusters are not seen as extended sources in RASS, and they have been confirmed by systematic follow-up of tens of thousands of X-ray sources, the ComPRASS catalogue provides a way to find some of these clusters in a more direct way.

Considering the results in the regions of the sky where our validation is more exhaustive (SPT and SDSS footprints, where many more clusters are already known), the expected purity for the ComPRASS catalogue is greater than 84.5%. This means that we expect to have more than 365 real clusters, unknown to date, among the new or yet-to-be-confirmed candidates, especially among those situated outside these two footprints.

We are currently working on the validation of the ComPRASS candidates that have not yet been confirmed with the external validation presented here. To this end we are performing visual inspection and quantitative analysis on ancillary data from WISE, SDSS, PanSTARRS, *XMM-Newton* and *Swift*. The results of this study will be part of a future paper.

*Acknowledgements.* This research is based on observations obtained with *Planck* (<http://www.esa.int/Planck>), an ESA science mission with instruments and contributions directly funded by ESA Member States, NASA, and Canada. This research has made use of the ROSAT all-sky survey data which have been processed at MPE. The authors acknowledge the use of the HEALPix package (Górski et al. 2005). This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n° 340519. The authors would like to thank Iacopo Bartalucci for his help in the preparation of the *Swift* images that were used to check some individual candidates.

## References

- Abell, G. O. 1958, *ApJS*, 3, 211  
 Abell, G. O., Corwin, Jr., H. G., & Olowin, R. P. 1989, *ApJS*, 70, 1  
 Andrade-Santos, F., Jones, C., Forman, W. R., et al. 2017, *ApJ*, 843, 76  
 Appenzeller, I., Thiering, I., Zickgraf, F. J., et al. 1998, *ApJS*, 117, 319  
 Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010, *A&A*, 517, A92  
 Barrena, R., Streblyanska, A., Ferragamo, A., et al. 2018, *A&A*, 616, A42  
 Bleem, L. E., Stalder, B., de Haan, T., et al. 2015, *ApJS*, 216, 27  
 Boada, S., Hughes, J. P., Menanteau, F., et al. 2018, *ApJ*, accepted [arXiv:1809.06378]  
 Boller, T., Freyberg, M. J., Trümper, J., et al. 2016, *A&A*, 588, A103  
 Buddendiek, A., Schrabback, T., Greer, C. H., et al. 2015, *MNRAS*, 450, 4248  
 Burenin, R. A. 2017, *Astron. Lett.*, 43, 507  
 Burenin, R. A., Vikhlinin, A., Hornstrup, A., et al. 2007, *ApJS*, 172, 561  
 Burenin, R. A., Bikmaev, I. F., Khamitov, I. M., et al. 2018, *Astron. Lett.*, 44, 297  
 Chon, G., & Böhringer, H. 2012, *A&A*, 538, A35  
 de Gasperin, F., Intema, H. T., Ridl, J., et al. 2017, *A&A*, 597, A15  
 Ebeling, H., Edge, A. C., & Henry, J. P. 2001, *ApJ*, 553, 668  
 Ebeling, H., Barrett, E., Donovan, D., et al. 2007, *ApJ*, 661, L33  
 Ebeling, H., Edge, A. C., Mantz, A., et al. 2010, *MNRAS*, 407, 83  
 Gonzalez, A. H., Gettings, D. P., Brodwin, M., et al. 2019, *ApJS*, 240, 33  
 Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759  
 Hao, J., McKay, T. A., Koester, B. P., et al. 2010, *ApJS*, 191, 254  
 Hasselfield, M., Hilton, M., Marriage, T. A., et al. 2013, *J. Cosmol. Astropart. Phys.*, 7, 8  
 Henry, J. P., Mullis, C. R., Voges, W., et al. 2006, *ApJS*, 162, 304  
 Herranz, D., Sanz, J. L., Hobson, M. P., et al. 2002, *MNRAS*, 336, 1057  
 Hilton, M., Hasselfield, M., Sifón, C., et al. 2018, *ApJS*, 235, 20  
 Hudson, D. S., Mittal, R., Reiprich, T. H., et al. 2010, *A&A*, 513, A37  
 Koester, B. P., McKay, T. A., Annis, J., et al. 2007, *ApJ*, 660, 239  
 Lakhchaura, K., Singh, K. P., Saikia, D. J., & Hunstead, R. W. 2011, *ApJ*, 743, 78  
 Mann, A. W., & Ebeling, H. 2012, *MNRAS*, 420, 2120



- Maturi, M. 2007, *Astron. Nachr.*, 328, 690
- Melin, J.-B., Bartlett, J. G., & Delabrouille, J. 2006, *A&A*, 459, 341
- Melin, J.-B., Aghanim, N., Bartelmann, M., et al. 2012, *A&A*, 548, A51
- Miller, N. A., Owen, F. N., Hill, J. M., et al. 2004, *ApJ*, 613, 841
- Morandi, A., Sun, M., Forman, W., & Jones, C. 2015, *MNRAS*, 450, 2261
- Nagai, D., Kravtsov, A. V., & Vikhlinin, A. 2007, *ApJ*, 668, 1
- Pace, F., Maturi, M., Bartelmann, M., et al. 2008, *A&A*, 483, 389
- Piffaretti, R., Arnaud, M., Pratt, G. W., Pointecouteau, E., & Melin, J.-B. 2011, *A&A*, 534, A109
- Planck Collaboration VIII. 2011, *A&A*, 536, A8
- Planck Collaboration XI. 2011, *A&A*, 536, A11
- Planck Collaboration I. 2012, *A&A*, 543, A102
- Planck Collaboration XX. 2014, *A&A*, 571, A20
- Planck Collaboration XXIX. 2014, *A&A*, 571, A29
- Planck Collaboration XXVI. 2015, *A&A*, 582, A29
- Planck Collaboration XXXII. 2015, *A&A*, 581, A14
- Planck Collaboration VIII. 2016, *A&A*, 594, A8
- Planck Collaboration XXVII. 2016, *A&A*, 594, A27
- Planck Collaboration Int. XXXVI. 2016, *A&A*, 586, A139
- Pratt, G. W., Croston, J. H., Arnaud, M., & Böhringer, H. 2009, *A&A*, 498, 361
- Repp, A., & Ebeling, H. 2018, *MNRAS*, 479, 844
- Rossetti, M., Gastaldello, F., Eckert, D., et al. 2017, *MNRAS*, 468, 1917
- Rozo, E., Rykoff, E. S., Bartlett, J. G., & Melin, J.-B. 2015, *MNRAS*, 450, 592
- Rykoff, E. S., Rozo, E., Busha, M. T., et al. 2014, *ApJ*, 785, 104
- Saro, A., Bocquet, S., Rozo, E., et al. 2015, *MNRAS*, 454, 2305
- Schuecker, P., Böhringer, H., & Voges, W. 2004, *A&A*, 420, 61
- Sunyaev, R. A., & Zeldovich, Y. B. 1970, *Comm. Astrophys. Space Phys.*, 2, 66
- Sunyaev, R. A., & Zeldovich, Y. B. 1972, *Comm. Astrophys. Space Phys.*, 4, 173
- Szabo, T., Pierpaoli, E., Dong, F., Pipino, A., & Gunn, J. 2011, *ApJ*, 736, 21
- Tarrío, P., Melin, J.-B., Arnaud, M., & Pratt, G. W. 2016, *A&A*, 591, A39
- Tarrío, P., Melin, J. B., & Arnaud, M. 2018, *A&A*, 614, A82
- Truemper, J. 1993, *Science*, 260, 1769
- van der Burg, R. F. J., Aussel, H., Pratt, G. W., et al. 2016, *A&A*, 587, A23
- Voges, W., Aschenbach, B., Boller, T., et al. 1999, *A&A*, 349, 389
- Wen, Z. L., & Han, J. L. 2018, *MNRAS*, 481, 4158
- Wen, Z. L., Han, J. L., & Liu, F. S. 2012, *ApJS*, 199, 34
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579
- Zwicky, F., Herzog, E., Wild, P., Karpowicz, M., & Kowal, C. T. 1961, *Catalogue of Galaxies and of Clusters of Galaxies* (Pasadena: California Institute of Technology), I

## Appendix A: Description of the ComPRASS catalogue

The ComPRASS catalogue contains 2323 candidates. For each candidate, the catalogue provides the following fields (see Table A.1). The median position error estimated from the distances to SPT positions is 54 arcsec (Tarrío et al. 2018).

- ID: Identifier of the candidate.
- Name: Name of the candidate.
- raj2000: Right ascension (J2000) in degrees.
- dej2000: Declination (J2000) in degrees.
- glon: Galactic longitude in degrees.
- glat: Galactic latitude in degrees.
- SNR\_J: Joint signal-to-noise ratio  $((S/N)_J)$  obtained with the best filter size.
- SNR\_SZ: SZ component of the S/N:  $(S/N)_{SZ}$ .
- SNR\_XR: X-ray component of the S/N:  $(S/N)_{XR}$ .
- Significance: Significance of the detection, which is defined in Tarrío et al. (2018) as the significance value in a Gaussian distribution corresponding to the probability that the detection is due to noise.
- M500\_list: Vector containing different values of joint mass, in units of  $10^{14} M_{\odot}$ , corresponding to different values of redshift (given in  $z\_list$ ). Together with  $z\_list$ , this vector provides the mass-redshift degeneracy curve of the candidate.
- M500\_list\_err\_upp: Vector containing the upper bound of the 68% confidence interval on the mass-redshift degeneracy curve, in units of  $10^{14} M_{\odot}$ .
- M500\_list\_err\_low: Vector containing the lower bound of the 68% confidence interval on the mass-redshift degeneracy curve, in units of  $10^{14} M_{\odot}$ .
- $z\_list$ : Vector containing the different values of redshift that correspond to the values in  $M\_list$ ,  $M\_list\_err\_upp$ , and  $M\_list\_err\_low$ .
- Class: Result from the classification of the candidate: 1 (confirmed), 2 (identified not confirmed), 3 (new).
- $z$ : Redshift of the candidate, which is taken from the cluster associated to the candidate. In case of multiple associations, this redshift is chosen as described in 3.6. If no cluster with known redshift is associated to the candidate, the default value -1 is shown.
- $z\_ref$ : Indicates the origin of the redshift: MCXC, SZ, redMaPPerSpec, redMaPPerPhot, WHL, AMF, MAXBCG, GMBCG, WenHan2018, Buddendiek2015, MadCOWS or NED/SIMBAD.
- M500: Estimation of the cluster joint mass ( $M_J$ ), in units of  $10^{14} M_{\odot}$ . This field is set to -1 when the redshift of the cluster is unknown and thus the mass cannot be calculated.
- M500\_err\_upp: Upper bound of 68% confidence interval on the cluster joint mass, in units of  $10^{14} M_{\odot}$ . This field is set to -1 when the redshift of the cluster is unknown and thus the mass cannot be calculated.
- M500\_err\_low: Lower bound of 68% confidence interval on the cluster joint mass, in units of  $10^{14} M_{\odot}$ . This field is set to -1 when the redshift of the cluster is unknown and thus the mass cannot be calculated.
- Planck: In case of association with a cluster/candidate in one of the *Planck* catalogues, this field indicates the name of the *Planck* cluster/candidate.
- SPT: In case of association with a cluster in the SPT catalogue, this field indicates the name of the SPT cluster.
- ACT: In case of association with a cluster in the ACT catalogue, this field indicates the name of the ACT cluster.
- MCXC: In case of association with a cluster in the MCXC catalogue, this field indicates the name of the MCXC cluster.
- redMaPPer: In case of association with a cluster in the redMaPPer catalogue, this field indicates the name of the redMaPPer cluster.
- Abell: In case of association with a cluster in the Abell catalogue, this field indicates the name of the Abell cluster.
- Zwicky: In case of association with a cluster in the Zwicky catalogue, this field indicates the name of the Zwicky cluster.
- MaxBCG: In case of association with a cluster in the MAXBCG catalogue, this field indicates the name of the MAXBCG cluster.
- GMBCG: In case of association with a cluster in the GMBCG catalogue, this field indicates the name of the GMBCG cluster.
- AMF: In case of association with a cluster in the AMF catalogue, this field indicates the name of the AMF cluster.
- WHL: In case of association with a cluster in the WHL catalogue, this field indicates the name of the WHL cluster.
- Wen2018: In case of association with a cluster in the Wen & Han (2018) catalogue, this field indicates the name of the Wen & Han (2018) cluster.
- Buddendiek2015: In case of association with a cluster in the Buddendiek et al. (2015) catalogue, this field indicates the name of the Buddendiek et al. (2015) cluster.
- Gonzalez2018: In case of association with a cluster in the Gonzalez et al. (2019) catalogue, this field indicates the name of the Gonzalez et al. (2019) cluster.
- id\_burenin: In case of association with a cluster in the catalogue of Burenin (2017), this field indicates the identifier of the cluster in this catalogue. Otherwise, it is set to -1.
- Other\_clusters: In case of association with a cluster in table 1, this field indicates the name of the cluster.
- Notes: Notes on specific candidates.

**Table A.1.** Cluster candidates in the ComPRASS catalogue.

Id	Name	raj2000	dej2000	glon	glat	SNR <sub>J</sub> (S/N) <sub>J</sub>	SNR <sub>SZ</sub> (S/N) <sub>SZ</sub>	SNR <sub>XR</sub> (S/N) <sub>XR</sub>	Significance	M500_list	M500_list_err_low	M500_list_err_upp	z_list
0	PSZRX G000.06+45.19	229.157	-0.969	0.061	45.192	10.05	6.40	7.76	7.57				
1	PSZRX G000.21+28.06	298.014	-39.718	0.211	-28.056	6.02	3.13	6.21	5.03				
2	PSZRX G000.35+77.81	203.781	20.153	0.346	77.814	5.27	3.82	3.75	4.87				
3	PSZRX G000.40+41.86	316.089	-41.354	0.403	-41.864	19.08	9.61	19.11	14.20				
4	PSZRX G000.62+48.15	324.445	-41.074	0.622	-48.152	6.72	3.11	6.34	4.89				
5	PSZRX G000.76+35.70	307.971	-40.613	0.757	-35.699	13.11	6.50	13.83	10.06				
6	PSZRX G001.50+35.78	308.172	-40.026	1.499	-35.778	5.86	3.71	5.45	4.87				
7	PSZRX G001.56+32.35	303.776	-39.454	1.555	-32.350	6.59	3.01	6.87	5.59				
8	PSZRX G001.84+46.92	322.772	-40.310	1.844	-46.923	12.88	4.51	13.61	9.20				
9	PSZRX G002.08+22.13	291.345	-36.478	2.083	-22.128	5.70	4.00	4.26	4.85				
10	PSZRX G002.20+66.99	213.118	14.001	2.199	66.992	6.85	3.85	5.94	5.54				

Id	Class	z	z_ref	M500 M <sub>500</sub>	M500_err_low	M500_err_upp	Planck	SPT	ACT	MCXC	redMaPPer
0	1	0.120	MCXC	2.52	0.76	0.67	PSZ2 G000.04+45.13			RXC J1516.5-0056	
1	3			-1	-1	-1					
2	3			-1	-1	-1					
3	1	0.165	MCXC	5.25	0.55	0.54	PSZ2 G000.40-41.86			RXC J2104.3-4120	
4	3			-1	-1	-1					
5	1	0.342	MCXC	6.53	0.83	0.79	PSZ2 G000.77-35.69	SPT-CLJ2031-4037		RXC J2031.8-4037	
6	1			-1	-1	-1					
7	1			-1	-1	-1					
8	1	0.421	MCXC	6.42	1.13	1.03		SPT-CLJ2131-4019		SMACSJ2131.1-4019	
9	2			-1	-1	-1	PSZ2 G002.04-22.15				
10	1	0.140	RedmapperSpec	1.94	0.54	0.48					RMJ141231.8+140041.1

Id	Abell	Zwicky	MaxBCG	GMBCG	AMF	WHL	Wen2018	Buddendiek2015	Gonzalez2018	id_burenin	Other_clusters	Notes
0	ABELL 2051			J229.18966-00.99227		J151644.1-005809						-1
1												-1
2												-1
3	ABELL 3739											-1
4												-1
5												-1
6	ABELL 3688											-1
7	ABELL 3671											-1
8												-1
9												-1
10	ABELL 1875		MAXBCG 5433									1868

**Notes.** The different fields are described in Appendix A. This table is available in its entirety in a machine-readable form at the CDS.

## Appendix B: Notes on the association of individual candidates

### B.1. Grey zone on X-ray associations

*Candidate 611.* This candidate is considered to be not associated with its closest MCXC cluster (RXC J0012.9-0853), since it is in the light grey zone of Fig. 4. This detection is in a complex X-ray region with three very close emission peaks. The candidate is centred at one of the peaks, coinciding with PSZ2 G094.46-69.65, ACT-CL J0012.8-0855 and RMJ001257.7-085829.5. However, RXC J0012.9-0853 is centred on the peak just next to it. This indicates that the association is incorrect.

*Candidate 675.* This candidate is considered to be not associated with its closest MCXC cluster (RXC J1540.1+6611), since it is in the light-grey zone of Fig. 4. This detection is inside the core of Abell 2125, which is a complex system. However, the position of RXC J1540.1+6611, a cluster from the 400SD catalogue (Burenin et al. 2007), is not on the core, but on LSBXE, a more diffuse X-ray emission to the southwest (see Fig. 4 of Miller et al. 2004). This confirms the non-association. This detection also matches with Zwicky 7603.

*Candidate 1327.* This candidate is considered to be not associated with its closest MCXC cluster (RXC J0347.4-2149), since it is in the light-grey zone of Fig. 4. This detection coincides with Abell 3138 and PSZ2 G215.19-49.65, but REFLEX II cluster RXC J0347.4-2149 is centred on a double structure just below A3168 (Chon & Böhringer 2012). This confirms the non-association.

*Candidate 1568.* This candidate is considered to be not associated with its closest MCXC cluster (RXC J0956.4-1004), since it is in the light-grey zone of Fig. 4. The RASS image shows three X-ray extended emissions. The detection is centred on one of them, while RXC J0956.4-1004 corresponds to another one, which confirms the non-association. This detection also matches with Abell 901 and PSZ2 G247.97+33.52. The association with the PSZ2 cluster is considered correct, since the detection is in the same SZ peak as the PSZ2 cluster. In the PSZ2 catalogue, however, PSZ2 G247.97+33.52 is associated with RXC J0956.4-1004.

### B.2. Zone I on X-ray associations

We define Zone I as the area in Fig. 4 where  $d < 5$  arcmin and  $d/\theta_{500} > 1$ . According to the X-ray association criteria defined



in Sect. 3.1, candidate-cluster pairs in this region are considered to be associated.

*Candidate 957.* This candidate is considered to be associated with its closest MCXC cluster (RXC J1241.5+3250). A detailed look at the X-ray and SZ S/N maps shows that the SZ signal contributes with a double structure, which causes the detection to slightly move away from the RXC J1241.5+3250 position (situated between the two peaks), but staying inside the same X-ray emission region. Although the relative distance  $d/\theta_{500} > 1$ , the angular distance is only 3 arcmin, so we can consider that the association is correct.

*Candidate 1991.* This candidate is considered to be associated with its closest MCXC cluster (RXC J0129.4–6432). A detailed look at the X-ray and SZ S/N maps shows that this candidate and RXC J0129.4–6432 are inside the same SZ peak and the same X-ray emission region, and therefore we can consider that the association is correct.

### B.3. Zone II on X-ray associations

We define Zone II as the area in Fig. 4 where  $d > 5$  arcmin and  $d/\theta_{500} < 1$ . According to the X-ray association criteria defined in Sect. 3.1, candidate-cluster pairs in this region are considered to be associated.

*Candidate 927.* This candidate has a very extended emission, both in X-ray and in SZ, corresponding to a very low redshift cluster ( $z = 0.065$ ). Although the candidate position is further than 5 arcmin away from the closest MCXC cluster (RXC J1200.3+5613), it is well inside the same X-ray emission peak. We can therefore consider that the association is correct.

*Candidate 1472.* This candidate has a very extended emission, both in X-ray and in SZ. The candidate position is inside the same X-ray emission region as the MCXC cluster RXC J1144.6+1945, a very low redshift cluster at  $z = 0.021$ . Therefore, we can consider that the association is correct.

*Candidate 1706.* The X-ray and SZ filtered maps of this candidate show two bright and extended X-ray sources in the same extended SZ peak. The candidate is centred at one of the X-ray sources, while the closest MCXC cluster (RXC J0627.2–5428) is situated between the two. This candidate corresponds to merging cluster A3395 (Lakhchaura et al. 2011). The candidate is centred at the SW emission region, while RXC J0627.2–5428 is centred closer to the NE region. Everything is part of the same system, so we can consider that the association is correct.

*Candidate 2106.* This candidate has a very extended emission, both in X-ray and in SZ. It corresponds to a very low-redshift cluster ( $z = 0.099$ ). The candidate position is inside the same X-ray emission peak as the MCXC cluster RXC J2359.3–6042, and therefore we can consider that the association is correct.

### B.4. Grey zone on SZ associations

*Candidate 435.* There are two clusters around this candidate: Zwicky 8586 and PSZ1 G070.59–30.48. The PSZ1 position is in the centre of a complex SZ structure that contains two SZ peaks. The candidate is situated at the stronger peak and therefore it seems to be correctly associated to PSZ1 G070.59–30.48. Furthermore, the candidate is closer to Zwicky 8586, to which PSZ1 G070.59–30.48 is associated in the PSZ1 catalogue.

*Candidate 742.* The SZ S/N map around this candidate shows two SZ peaks. While the candidate is situated at one of the peaks, the closest SZ cluster (PSZ1 G266.19+19.06) is between the two, meaning that the PSZ1 detection covers both of them. We therefore consider that the association is correct.

*Candidate 830.* A detailed look at the X-ray and SZ S/N maps around this cluster show that the candidate is inside the SZ emission centred at PSZ2 G121.13+49.64, but its position is biased towards a strong X-ray emission. In addition, the candidate is associated with a Zwicky cluster (Zwicky 5680), which is associated with PSZ2 G121.13+49.64 in the PSZ2 catalogue. For these reasons, we consider that the association is correct.

*Candidate 1197.* This candidate is clearly associated with an MCXC cluster (RXC J0728.9+2935). Its closest SZ cluster (PSZ2 G189.23+20.55) was associated with the RXC J0728.9+2935 in the PSZ2 catalogue, and therefore we consider that the association is correct. A detailed look at the X-ray and SZ S/N maps shows 1 peak in the X-ray map and 2 peaks in the SZ map, one coinciding with the X-ray peak. The candidate and the MCXC cluster are situated at the X-ray peak, while the PSZ2 cluster is situated closer to the other peak. All of them seem to be part of the same (complex) structure.

*Candidate 1300.* This candidate is clearly associated with an MCXC cluster (RXC J0959.7+2223). Its closest SZ cluster (PSZ2 G210.01+50.85) was not associated with RXC J0959.7+2223 in the PSZ2 catalogue, and therefore we consider that the association is incorrect.

*Candidate 1587.* The position of this candidate is shifted with respect to the SZ cluster PSZ2 G250.04+24.14 due to the presence of a bright X-ray source next to the cluster (but probably outside the cluster). We therefore consider that the association is incorrect.

*Candidate 1718.* The SZ and X-ray S/N maps around this candidate show two strong X-ray emissions which appear as a single object in the SZ maps. The candidate is situated at the southern X-ray peak, closely matching one MCXC cluster (RXC J0330.0–5235) and one Abell cluster (Abell 3128), both at very low redshift. Its closest SZ cluster (PSZ2 G264.60–51.07, SPT-CLJ0330–5228, ACT-CL J0330–5227) is situated at the northern X-ray peak and has an estimated redshift of 0.44. We therefore conclude that there are two different clusters and that the candidate corresponds to the lower-redshift cluster, and should not be associated with the higher-redshift SZ cluster.

*Candidate 1732.* This candidate is associated with Abell 3436. Its closest SZ cluster (PSZ1 G266.19+19.06) was not associated with Abell 3436 in the PSZ1 catalogue. However, a detailed look at the X-ray and SZ S/N maps show that the three objects belong to the same structure. We therefore consider that the association is correct.

*Candidate 2094.* This candidate is at the same position of Abell 3666 and MCXC cluster RXC J2016.2–8047. Its closest SZ cluster (PSZ2 G313.00–30.01) sits at another X-ray peak where there is another MCXC cluster and another Abell cluster. The association of this candidate with PSZ2 G313.00–30.01 is therefore incorrect.

## Appendix C: Notes on multiple associations

*Candidate 610.* Double cluster. This candidate is associated to a PSZ2 cluster at  $z = 0.453$  and to an MCXC and Abell cluster at  $z = 0.225$ . The two clusters contribute to the detection.

### C.1. Multiple associations with [Wen & Han \(2018\)](#) clusters

*Candidate 497.* There are two different clusters within 5 arcmin. One is a PSZ2 cluster at  $z = 0.04$  and another one is a Wen cluster at  $z = 0.81$ . The X-ray and SZ images correspond clearly to a very extended (very low-redshift) cluster.

*Candidate 647.* There are two different clusters within 5 arcmin. One is a redMaPPer cluster at  $z = 0.18$  and another one is a Wen cluster at  $z = 0.96$ . The redMaPPer cluster is much closer and the X-ray image corresponds better to this cluster.

*Candidate 919.* There are two different clusters within 5 arcmin. One is a redMaPPer and Abell cluster at  $z = 0.21$  and another one is a Wen cluster at  $z = 0.82$ . The redMaPPer cluster is much closer. The X-ray image shows the emission from the redMaPPer clusters.

*Candidate 953.* There are three different clusters within 5 arcmin. One is a PSZ2 cluster at  $z = 0.36$ , another is an Abell and redMaPPer cluster at  $z = 0.28$ , and the third is a Wen cluster at  $z = 0.70$ . The PSZ2 and the Abell-redMaPPer appear to be the most probable counterparts.

*Candidate 1020.* There are two different clusters within 5 arcmin. One is a MCXC cluster at  $z = 0.07$  and another one is a Wen cluster at  $z = 0.77$ . The X-ray and SZ images correspond clearly to a very extended (very low-redshift) cluster.

*Candidate 1131.* There are two different clusters within 5 arcmin. One is a SZ-MCXC-redMaPPer cluster at  $z = 0.19$  and another one is a Wen cluster at  $z = 0.86$ . The first one is the clear association by looking at *Swift* and XMM images.

*Candidate 1260.* There are two different clusters within 5 arcmin. One is a redMaPPer cluster at  $z = 0.57$  and another one is a Wen cluster at  $z = 0.76$ . Both of them have a mass of 2.5. The X-ray image does not allow us to discriminate which is the best position. We may be seeing a superposition of the two clusters.

*Candidate 1308.* There are two different clusters within 5 arcmin. One is a redMaPPer cluster at  $z = 0.41$  and another one is a Wen cluster at  $z = 0.71$ . The redMaPPer cluster is more massive and located at a smaller angular separation. There is another cluster at  $z$  around 0.2 visible within the 5 arcmin.

### C.2. Multiple associations with [Gonzalez et al. \(2019\)](#) clusters

*Candidate 1143.* There are two different clusters within 5 arcmin. One is a cluster at  $z = 0.21$  that is included in MCXC, PSZ2, Abell, and redMaPPer catalogues and that coincides in position with our candidate. The other one is a MADCowS cluster at  $z = 1.12$  further from our candidate (at 4 arcmin). The correct association is therefore the lower-redshift cluster.

The MADCowS cluster does not appear to contribute greatly to the total filtered signal.

*Candidate 1165.* There are two different clusters within 5 arcmin. One is a redMaPPer cluster at  $z = 0.50$  and the other one is a MADCowS cluster at  $z = 0.89$ . The redMaPPer cluster is situated at the same position as our candidate (distance of 0.3 arcmin), while the MADCowS cluster is further away (at 3.6 arcmin). The correct association is therefore the redMaPPer cluster. The MADCowS cluster does not appear to contribute greatly to the total filtered signal.

### C.3. Multiple associations with [Buddendiek et al. \(2015\)](#) clusters

*Candidate 1121.* There is a redMaPPer cluster at  $z = 0.60$  and a [Buddendiek et al. \(2015\)](#) cluster at  $z = 0.67$  coinciding in position with the candidate and with the X-ray peak in RASS maps. Given that the redMaPPer redshift is photometric, it appears that both clusters are indeed the same object. The association is therefore correct for both. Since the redshift provided by [Buddendiek et al. \(2015\)](#) is spectroscopic (and several spectroscopic redshifts consistent with it are also available in SDSS) we can consider that this latter redshift is more precise.

*Candidate 1130.* There is a PSZ2 cluster at  $z = 0.57$  and a [Buddendiek et al. \(2015\)](#) cluster at  $z = 0.70$  coinciding in position with the candidate and with the X-ray peak in RASS maps. By looking at the SDSS images, both clusters seem to be the same object. Both associations are therefore correct. The redshift provided by Buddendiek is compatible with the photometric redshifts available in SDSS, but the redshift of the PSZ2 cluster, which comes from a redMaPPer photometric redshift, is lower. This may indicate that it was underestimated.

*Candidate 1232.* There is a superposition of two different clusters: the Buddendiek cluster (closer to the detection) and a redMaPPer cluster (more distant). Candidate 1232 has a joint estimated mass of 5.55 assuming the redshift of the redMaPPer counterpart, while the redMaPPer mass is 2.77. If we assume the redshift of the Buddendiek cluster, the joint mass would be 6.73, whereas the mass reported by [Buddendiek et al. \(2015\)](#) for this cluster is 9.8. It is possible that both clusters contribute to the detection, or that we have only detected the high- $z$  cluster. In that case, the redMaPPer clusters would not be correctly associated.

*Candidate 1411.* There is a superposition of two different clusters: the Buddendiek cluster (closer to the detection) and a redMaPPer cluster (more distant). Our candidate has a joint estimated mass of 3.97 assuming the redshift of the redMaPPer counterpart, while the redMaPPer mass is 1.81. It is possible that both clusters contribute to the detection, or that we have only detected the high- $z$  cluster. In that case, the redMaPPer clusters would not be correctly associated.