



HAL
open science

MILCANN: A tSZ map for galaxy cluster detection assessed using a neural network

G. Hurier, N. Aghanim, M. Douspis

► **To cite this version:**

G. Hurier, N. Aghanim, M. Douspis. MILCANN: A tSZ map for galaxy cluster detection assessed using a neural network. *Astronomy and Astrophysics - A&A*, 2021, 653, pp.A106. 10.1051/0004-6361/201730534 . hal-03347111

HAL Id: hal-03347111

<https://hal.science/hal-03347111>

Submitted on 16 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MILCANN: A tSZ map for galaxy cluster detection assessed using a neural network[★]

G. Hurier¹, N. Aghanim², and M. Douspis²

¹ Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Plaza de San Juan, 1, Planta 2, 44001 Teruel, Spain

² Université Paris Saclay, CNRS, Institut d'Astrophysique Spatiale, Bâtiment 121 Campus Paris-Sud, 91405 Orsay, France
e-mail: nabila.aghanim@ias.u-psud.fr

Received 31 January 2017 / Accepted 14 December 2019

ABSTRACT

We present the first combination of a thermal Sunyaev-Zel'dovich (tSZ) map with a multi-frequency quality assessment of the sky pixels based on artificial neural networks with the aim being to detect tSZ sources from submillimeter observations of the sky by *Planck*. We present the construction of the resulting filtered and cleaned tSZ map, MILCANN. We show that this combination leads to a significant reduction of noise fluctuations and foreground residuals compared to standard reconstructions of tSZ maps. From the MILCANN map, we constructed a tSZ source catalog of about 4000 sources with a purity of 90%. Finally, we compare this catalog with ancillary catalogs and show that the galaxy-cluster candidates in our catalog are essentially low-mass (down to $M_{500} = 10^{14} M_{\odot}$) high-redshift (up to $z \leq 1$) galaxy cluster candidates.

Key words. large-scale structure of Universe – galaxies: clusters: general – methods: data analysis

1. Introduction

As the largest virialized structures in the Universe, galaxy clusters are excellent tracers of matter distribution. Their abundance can be used to constrain the cosmological model in a way that is independent from and complementary to the cosmic microwave background (CMB; see e.g., [Planck Collaboration XX 2014](#); [Planck Collaboration XIII 2016](#); [Hurier & Lacasa 2017](#); [Salvati et al. 2018](#)). Galaxy clusters are composed of dark matter, of the galaxies themselves, as well as hot ionized intra-cluster medium (ICM). Consequently, they can be identified in the optical bands as concentrations of galaxies (see e.g., [Abell et al. 1989](#); [Gladders & Yee 2005](#); [Koester et al. 2007](#); [Rykoff et al. 2014](#)), and can be observed in X-rays thanks to the bremsstrahlung emission produced by the ionized ICM (see e.g. [Bohringer et al. 2000](#); [Ebeling et al. 2000, 2001](#); [Böhringer et al. 2001](#)). The same hot ICM also creates a distortion in the black-body spectrum of the CMB through the thermal Sunyaev-Zel'dovich (tSZ) effect ([Sunyaev & Zeldovich 1969, 1972](#)), an inverse-Compton scattering between the CMB photons and the ionized electrons in the ICM (see e.g. [Birkinshaw 1999](#); [Carlstrom et al. 2002](#), for reviews). The tSZ Compton parameter in a given direction, \mathbf{n} , on the sky is given by

$$y(\mathbf{n}) = \int n_e \frac{k_B T_e}{m_e c^2} \sigma_T ds, \quad (1)$$

where ds is the distance along the line of sight, \mathbf{n} , and n_e and T_e are the electron number density and temperature, respectively. In units of CMB temperature the contribution of the tSZ effect at a

frequency ν is

$$\frac{\Delta T_{\text{CMB}}}{T_{\text{CMB}}} = g(\nu) y. \quad (2)$$

Neglecting relativistic corrections, we have

$$g(\nu) = \left[x \coth\left(\frac{x}{2}\right) - 4 \right], \quad (3)$$

with $x = h\nu/(k_B T_{\text{CMB}})$. At $z = 0$, where $T_{\text{CMB}}(z = 0) = 2.726 \pm 0.001$ K, the tSZ effect is negative below 217 GHz and positive for higher frequencies.

Recent large cluster catalogs based on measurements of the tSZ effect have been produced from *Planck* ([Planck Collaboration VIII 2011](#); [Planck Collaboration Int. XXXII 2015](#)), ACT ([Marriage et al. 2011](#); [Hilton et al. 2018](#)), and SPT ([Bleem et al. 2015](#)) data. Several detection algorithms targeting tSZ sources (see e.g., [Melin et al. 2006](#); [Carvalho et al. 2009](#)) have been proposed and compared ([Melin et al. 2012](#)). These latter authors demonstrated that a multi-filter approach based on the use of optimal filters for tSZ detection is more robust than a tSZ map-based approach. The latter relies on the construction of a y map with component separation methods such as MILCA ([Hurier et al. 2013](#)) or NILC ([Remazeilles et al. 2011](#)). These methods are devised to mitigate contamination of the tSZ signal from other astrophysical emissions, such as radio, infra-red point sources, and cosmic infra-red background ([Dunkley et al. 2011](#); [Shirokoff et al. 2011](#); [Reichardt et al. 2012](#); [Sievers et al. 2013](#); [Planck Collaboration XXI 2014](#)). The combination of high- and low-resolution SZ surveys, as was performed in PACT ([Aghanim et al. 2019](#)), is another way of efficiently reducing the contamination of the y map but this relies on the availability of complementary public data. In all cases, the reconstructed y map cannot be totally immune from contamination that can produce spurious galaxy cluster detections and/or

[★] The list of candidate clusters is only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/653/A106>

a significant bias in the measured tSZ fluxes. Melin et al. (2012) showed that tSZ map-based detection methods suffer from a larger number of spurious tSZ sources than multi-filter methods, leading to a significantly lower level of purity of the produced catalogs. As a result, an a posteriori quality assessment of the tSZ signal from galaxy clusters is required to produce high-purity galaxy cluster samples of tSZ clusters.

Planck Collaboration Int. XXXII (2015) and Planck Collaboration XXVII (2016) describe a method for the quality assessment of the tSZ sources, which they refer to as a “validation”. The method includes several steps, from cross-matches with catalogs to the search for counterparts in galaxy or X-ray surveys, including visual inspection of tSZ sources. Automatic assessments of the quality of tSZ-detected sources can also be performed. A method based on artificial neural networks (ANN) was proposed by Aghanim et al. (2015). This uses the *Planck* multi-frequency data to assess the quality of the tSZ sources by decomposing the measured signal into the different astrophysical components contributing to the *Planck* frequencies. This new quality assessment method was applied to validate the *Planck* cluster catalog (Planck Collaboration XXVII 2016). Moreover, the use of the ANN method showed that the tSZ-source catalog (Planck Collaboration Int. XXXII 2015) suffers from contamination by galactic CO sources and infra-red emission. A recent follow-up of *Planck* tSZ sources in the optical (van der Burg et al. 2016) demonstrated the efficiency of this ANN-based quality assessment by confirming the sources with poor quality criteria as actual spurious tSZ sources.

In this work, we extend the use of the ANN method initially developed by Aghanim et al. (2015) to provide a quality assessment of individual tSZ sources and of each pixel in a reconstructed full-sky y map. The resulting ANN-weighted y map is then used for cluster detection, providing a new catalog of tSZ sources. The paper is organized as follows. In Sect. 2, we present the different datasets. In Sect. 3, we detail the construction of the ANN-based weights. In Sect. 6, we construct a sample of galaxy cluster candidates and present a detailed characterization of this sample. Finally in Sect. 7, we perform a multi-wavelength assessment of the detected galaxy cluster candidates.

Throughout the paper we use the following cosmological parameters derived from the results of the Planck Collaboration 2015 (Planck Collaboration XIII 2016): $\Omega_m = 0.316$, $H_0 = 67.26$, $\sigma_8 = 0.83$, $n_s = 0.9652$, and $\Omega_b h^2 = 0.02222$.

2. Data

For the present analysis we used several publicly available datasets (catalogs and surveys) either to describe astrophysical source properties or to characterize the galaxy cluster candidates detected in the ANN-weighted y maps.

We use the *Planck* intensity maps from 70 to 857 GHz (Planck Collaboration I 2016). We also use the spectral responses given in Planck Collaboration IX (2014). We assume the *Planck* beams are Gaussian with values given in Planck Collaboration VII (2014). We also use the *Planck* full-sky CMB-lensing map (Planck Collaboration XV 2016). The Planck Collaboration has published source catalogs that we use for our analysis. Namely, the catalog of tSZ sources detected in *Planck* data (PSZ2 hereafter, see Planck Collaboration XXVII 2016) and the catalogs of point-sources detected at 30 GHz and 353 GHz from the *Planck* Catalogue of Compact Sources (Planck Collaboration XXVIII 2014). All the *Planck* data can be retrieved from the *Planck* Legacy Archive¹.

¹ <https://pla.esac.esa.int/>

We also use the reprocessed IRAS maps, IRIS (Improved Reprocessing of the IRAS Survey, Miville-Deschênes & Lagache 2005), and the AllWISE Source Catalog² (Wright et al. 2010; Mainzer et al. 2011). Finally, we use catalogs of clusters detected in X-rays (MCXC, Piffaretti et al. 2011, and reference therein) and in the SDSS survey, namely WHL12 (Wen et al. 2012), WHL15 (Wen & Han 2015), WHY18 (Wen et al. 2018), and redMaPPer (Rykoff et al. 2014).

3. Artificial neural network

Machine learning, and in our specific case ANNs, enable us to learn the characteristic signature of “true” tSZ sources and spurious signals directly from the data using a reference sample of astrophysical sources. As shown by Aghanim et al. (2015), van der Burg et al. (2016), Planck Collaboration Int. XXXII (2015), and Planck Collaboration XXVII (2016), this method allows the identification of spurious tSZ sources from catalogs of cluster candidates. In the following, we adapt the approach used in Aghanim et al. (2015) in order to extend the machine-learning-based quality assessment to each pixel of the sky maps rather than to samples of individually detected tSZ candidates.

For clarity, we first summarize the key elements of the ANN method (a more detailed description is provided in Aghanim et al. 2015).

We focus on the astrophysical emissions that most affect the tSZ signal in multifrequency experiments and we model the flux at each frequency, taking into account the tSZ effect (neglecting relativistic corrections), CMB, and CO emission. We also add an effective IR component, which represents the contamination by galactic dust, cold Galactic sources, and Cosmic Infrared Background (CIB) fluctuations, and an effective radio component, which accounts for diffuse radio and synchrotron emission and radio sources. The flux in frequency is then written as

$$F_\nu = A_{SZ}F_{SZ}(\nu) + A_{CMB}F_{CMB}(\nu) + A_{IR}F_{IR}(\nu) + A_{RAD}F_{RAD}(\nu) + A_{CO}F_{CO}(\nu) + N(\nu), \quad (4)$$

where $F_{SZ}(\nu)$, $F_{CMB}(\nu)$, $F_{IR}(\nu)$, $F_{RAD}(\nu)$, and $F_{CO}(\nu)$ are the spectra of tSZ, CMB, IR, radio, and CO emissions; A_{SZ} , A_{CMB} , A_{IR} , A_{RAD} , and A_{CO} are the corresponding amplitudes; and $N(\nu)$ is the instrumental noise.

In order to improve the photometry of tSZ sources, we compute the flux for each pixel and each frequency using a matched-filter in ℓ rather than the aperture photometry used in Aghanim et al. (2015). To build the matched-filter in ℓ equivalent to a Wiener filter, we compute the power spectrum, y_ℓ , of a tSZ signal from a single cluster with $R_{500} = 5'$ (point-like with respect to the *Planck* experiment) assuming a universal pressure profile (Arnaud et al. 2010). We also consider the power spectrum of the CMB, C_ℓ^{CMB} , computed using *Planck* best-fit cosmology (Planck Collaboration XIII 2016), and the power spectrum of the noise, C_ℓ^{NN} , in the 100 GHz channel (estimated from the half-ring map difference). Considering the most relevant frequencies for the tSZ flux estimation (100 to 217 GHz), relevant angular scales ($\ell \in [1000, 2000]$), and focusing on the cleanest sky area (high galactic latitudes, $b > 20$), we neglect the thermal dust contribution from the Milky Way and therefore do not include it in the computation of the matched-filter. The resulting filter is

² <http://wise2.ipac.caltech.edu/docs/release/allwise/expsup/>

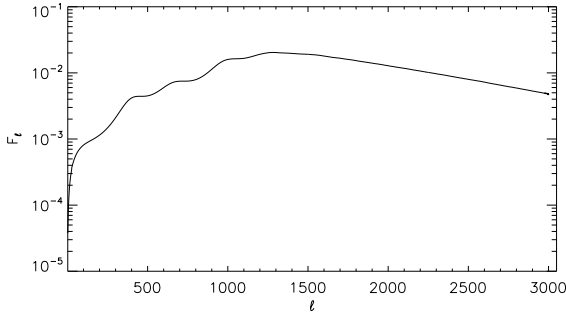


Fig. 1. Matched filter in ℓ space (F_ℓ) used for ANN-based quality assessment.

therefore defined as:

$$F_\ell = \sqrt{\frac{y_\ell}{C_\ell^{\text{CMB}} + C_\ell^{\text{NN}}}}, \quad (5)$$

and is shown in Fig. 1. The filter is applied to the *Planck* intensity maps from 70 to 857 GHz convolved with a 13' beam (lowest resolution associated with the 70 GHz map). The resulting harmonic space coefficients are thus given by $a'_{\ell,m} = F_\ell a_{\ell,m}$. The amplitudes of F_ν (Eq. (4)), namely A_{SZ} , A_{CMB} , A_{IR} , A_{RAD} , and A_{CO} , are linearly fitted for each pixel from the map constructed with the filtered harmonic coefficients. We verified that the results do not significantly depend on the chosen amplitude of C_ℓ^{NN} for the matched-filter computation. Figure 2 (right panel) shows the correlation matrix of the fitted SED parameters. The left panel displays the correlation matrix of measured fluxes (via the matched-filter) from 70 to 857 GHz. The correlation matrix of the SED amplitudes shows that the matched-filter photometry allows us to achieve measurements of the fluxes with a higher signal-to-noise ratio and therefore a better separation of the various contributions to the SED as compared to the aperture photometry used in Aghanim et al. (2015). The difference is particularly striking for the tSZ component that is now only significantly correlated with the radio component. This correlation has two origins, a physically motivated spatial correlation with radio-loud AGNs and the similarities of shape between the tSZ and radio SED at low frequency, making their distinction hard to achieve. We observe another significant correlation between the CO and the thermal dust components due to spatial correlation. We stress that these two components are among the major sources of spurious detection in the *Planck* catalog: the CO emission produces a rise of the intensity at 100 GHz, and the signal from the dust emission increases with frequency. These two trends together mimic the tSZ spectral signature. In this case, the 70 GHz channel is of great use to separate CO emission, which only affects the 100 GHz channel, from a tSZ emission, which should present consistent 70 and 100 GHz channels. Additionally, by construction, the matched-filter selects specific scales of the tSZ emission that correspond to the cluster scales, allowing us to reduce the contamination by large-scale emissions (i.e., galactic thermal dust, CMB).

Following Aghanim et al. (2015), we consider a standard three-layer back-propagation ANN to separate pixels of the sky maps into three populations of reliable quality, unreliable quality (i.e., false detection), and noisy sources (referred to as good, bad, and ugly in Aghanim et al. 2015). The inputs of the neural network consist of the five SED parameters per pixel, computed here with a more precise photometry based on the matched-filter, from which we derive three full-sky maps associated with the three quality classes.

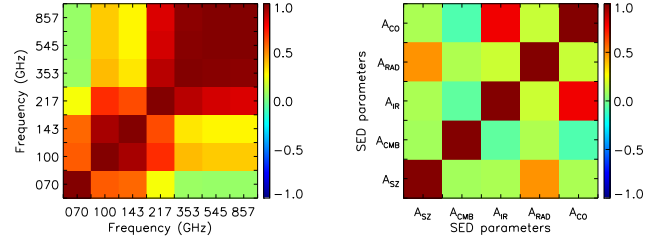


Fig. 2. *Left panel:* correlation matrix of the measured fluxes from 30 to 857 GHz estimated on 2000 random positions over the sky. *Right panel:* correlation matrix of fitted SED parameters from the same positions.

To train the neural network and to assess the tSZ signal quality for each sky-pixel, we use the same sample for Good, Bad, and Ugly classes as in Aghanim et al. (2015), namely: galaxy clusters; infra-red, radio, and cold galactic sources; and random (noise) estimates.

4. Construction of the tSZ map

4.1. MILCA *Planck* tSZ map

Independently from the ANN quality assessment, we reconstructed a tSZ map with the MILCA method (Hurier et al. 2013) using *Planck* HFI from 100 to 857 GHz, after verifying that including frequencies from 30 to 70 GHz does not significantly change the reconstructed map, especially at galaxy cluster scales.

We performed the construction of the tSZ map using eight bins in spherical harmonic space. For the first three bins, we used two constraints (tSZ and CMB), and for the last five bins we only used a constraint on the tSZ SED. The map reconstruction was performed with an effective FWHM of 7 arcmin. For all bins, two degrees of freedom were used to minimize the variance of the noise (see Hurier et al. 2013, for a detailed description of the MILCA method).

In Fig. 3, we show the MILCA full-sky map at 7 arcmin FWHM. Figures 4 and 5 show a zoom on two regions of $8.5 \times 8.5^\circ$ and two regions of $4.25 \times 4.25^\circ$ where we can observe bright galaxy clusters; a typical region with *Planck* clusters, a region with known optically selected clusters (redMaPPer catalog with $\lambda > 50$); a region showing low-S/N cluster candidates, and a region showing significant CO contamination triggering spurious detection in the PSZ2 catalog. In the full-sky map (Fig. 3), we observe a significant amount of foreground residuals near the galactic plane, where synchrotron and free-free residuals appear as negative biases in the tSZ y -map signal. We also observe contamination by bright galactic cirrus correlated with the zodiacal light. As shown in previous works (Hurier et al. 2013; Planck Collaboration XXII 2016), the main sources of contamination in tSZ maps built from *Planck* intensity maps are radio point sources, CO, and CIB emission.

For consistency with the ANN quality assessment, we convolve the reconstructed tSZ map, noted \bar{y} , by the matched-filter used for the SED fitting. The obtained filtered map, \bar{y}_f , has a transfer function consistent with the maps used to perform the ANN classification.

4.2. ANN weighting

In Fig. 3, we observed that the reconstructed tSZ map suffers from bias due to residuals from other astrophysical emission. Using the ANN-based quality assessment presented in Sect. 3, we can estimate the quality of the tSZ signal for each line of

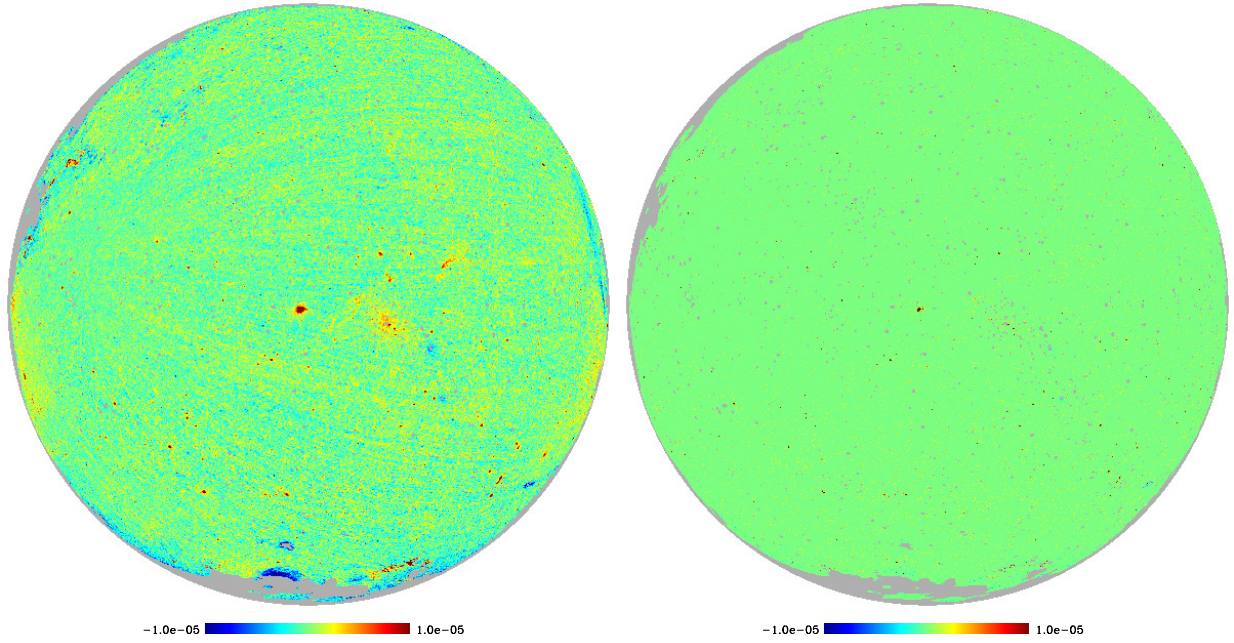


Fig. 3. *Left panel:* Planck tSZ MILCA map. *Right panel:* MILCANN map, for the northern sky in orthographic projection. Gray regions are masked due to galactic foreground or point-source contamination.

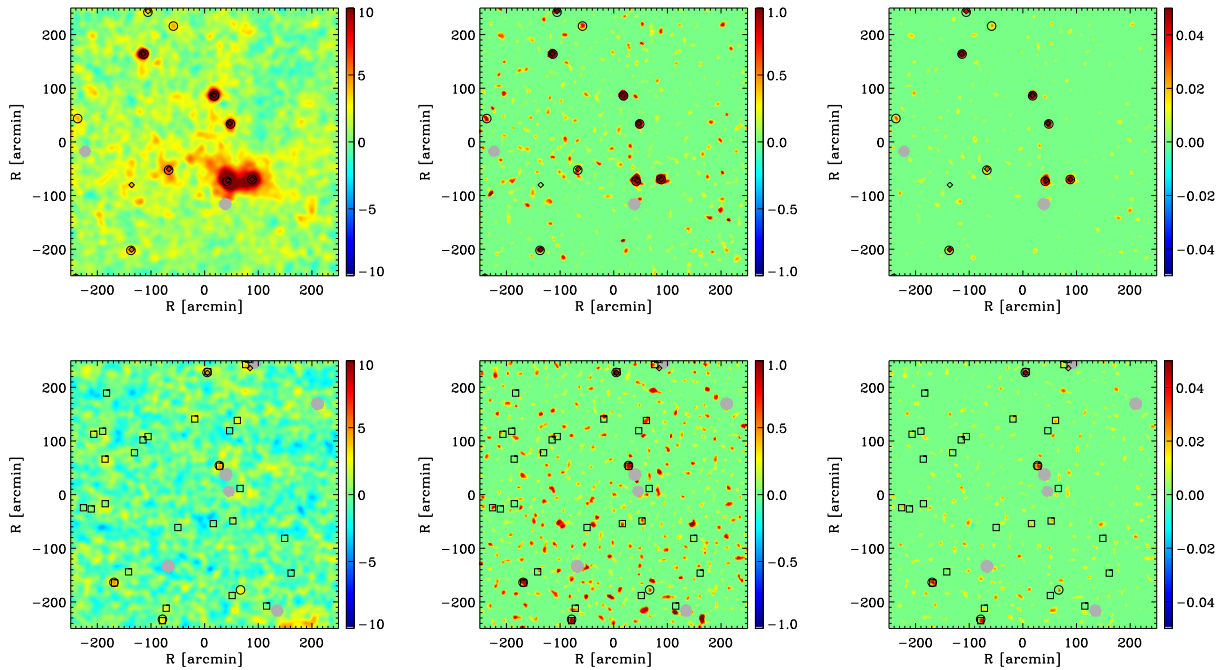


Fig. 4. *Upper panels:* a region of the sky including a nearby multiple-cluster system, *lower panels:* a region with a large number of clusters identified in the optical within SDSS data. *From left to right:* Planck tSZ MILCA map, neural network weight, match-filtered MILCANN map, for patches of $8.5 \times 8.5^\circ$ in gnomonic projection centered on galactic coordinates $(l, b) = (264^\circ, -24^\circ)$ and $(l, b) = (11.5^\circ, 70^\circ)$. Gray regions are masked due to point-source contamination. We display objects from the HAD catalog (presented in this paper) as black circles, from the PSZ2 catalog as black diamonds, and from redMaPPer (when available) as black squares.

sight in the sky. First, we define an ANN-based weight, Q_N , as

$$Q_N = Q_{\text{GOOD}}(1 - Q_{\text{BAD}}), \quad (6)$$

where Q_{GOOD} and Q_{BAD} are the ANN classification output values for the Good and Bad classes. By construction, this ANN-weight ranges from 0 to 1, with values close to 1 for pixels that present a high-quality tSZ signal. As verified by the optical follow-up of tSZ candidates in van der Burg et al. (2016), this ANN-based weight provides a good proxy for the robustness of a tSZ signal.

4.3. MILCANN map

Finally, we construct a new map (noted MILCANN) from the input tSZ map, filtered and cleaned from the contamination as:

$$M_{\text{MILCANN}} = \widehat{y}_t Q_N. \quad (7)$$

In Figs. 4 and 5, we display the ANN weight described in Sect. 3 and the MILCANN map. We observe a clear spatial correlation between the ANN weight and the input tSZ map. In Fig. 3,

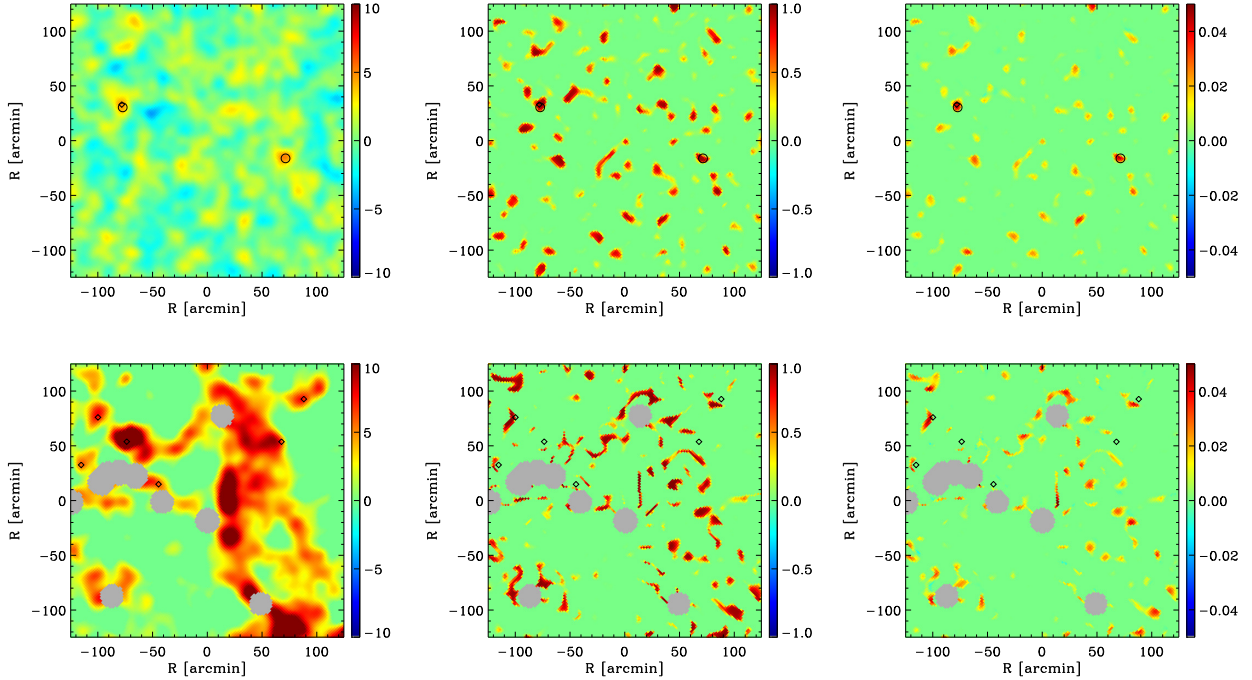


Fig. 5. Region of the sky without foreground contamination (*upper panels*) compared with an area of the sky dominated by galactic contamination (*lower panels*). From left to right: same as Fig. 4 but for patches of $4.25 \times 4.25^\circ$ centered on galactic coordinates $(l, b) = (44.5^\circ, -27.5^\circ)$ and $(l, b) = (122^\circ, 24.5^\circ)$.

the MILCANN map shows that the ANN weight has significantly removed the foreground contamination, especially near the galactic plane where synchrotron and free-free contamination have been completely suppressed. Similarly, the contamination produced by high-latitude galactic dust has been reduced by the ANN weighting procedure. On Figs. 4 and 5, we also observe that the MILCANN map presents a significantly reduced overall background noise, with the bright tSZ pixels conserved.

We observe that redMaPPer clusters previously undetected via their tSZ signal are seen in the MILCANN map (see second row of Fig. 4). We also note that the ANN weighting process allows us to avoid the contamination by spurious CO sources that were present in the PSZ2 catalog (see bottom row of Fig. 5). In general, Figs. 4 and 5 show that the ANN weighting process leading to the MILCANN map is a significant improvement for galaxy cluster detection when approaching the tSZ noise limit; it allows us to lower the threshold of detection without causing us to obtain a significant number of spurious sources.

However, it is worth noting that the ANN classification and subsequent weighting process is not a perfect procedure (see also Aghanim et al. 2015). As a matter of fact, the ANN-weight map in Fig. 4 presents values of close to 1 even for some pixels where no clear tSZ signal can be observed. These misclassified pixels are produced by chance alignment between noise structure and tSZ spectral signature. As a consequence, a tSZ-source detection performed directly on the ANN-weighted map would lead to false detections. Furthermore, the value of Q_N is not exactly 1 for all bona fide galaxy clusters. A high-flux galaxy cluster will have $Q_N \simeq 1$, whereas low-flux galaxy clusters will have $Q_N < 1$. As a result, the ANN weighting process does not conserve the shape of the tSZ sources. It modifies the intensity of faint tSZ pixels in the outskirts of galaxy clusters.

To estimate the transfer function of the ANN weighting procedure, we randomly selected 10 000 pixels with no signif-

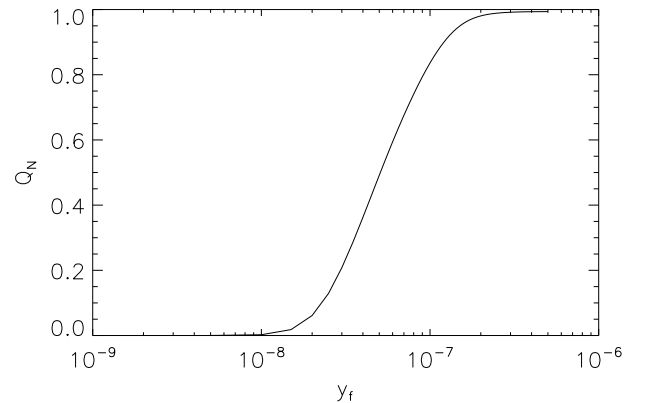


Fig. 6. Average ANN weight value as a function of the match-filtered intensity.

icant astrophysical emission within the 84% sky area used for the detection (see Planck Collaboration Int. XXXII 2015, for a description of the mask). In the filtered frequency maps, we added a given tSZ signal to the corresponding pixels. We then computed Q_N for these 10 000 pixels by applying the ANN to the modified frequency map pixels, and averaged them as a function of the injected tSZ signal. This approach allows us to properly account for real sky background and noise level in the transfer function. In Fig. 6, we show the average value of the ANN weight as a function of the injected tSZ signal intensity after filtering, y_f . We observe that the ANN response presents a steep transition: all signal below $y_f = 10^{-8}$ is completely suppressed by the ANN weighting process, whereas signal above $y_f = 2 \times 10^{-7}$ is almost not affected. We stress here that these Compton y values are obtained after filtering and are not comparable with the tSZ intensity in the input y map.

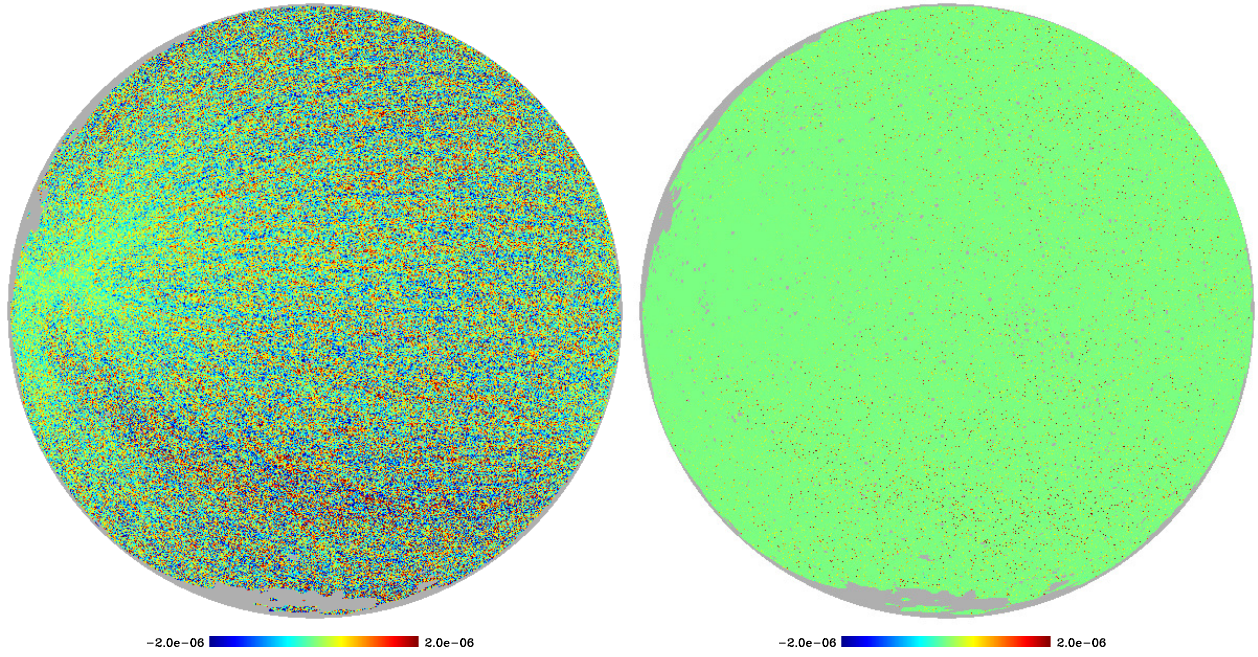


Fig. 7. *Left panel:* simulated *Planck* tSZ MILCA noise map. *Right panel:* simulated MILCANN noise map for the northern sky in orthographic projection. Gray regions are masked due to galactic foregrounds or point-source contamination.

5. Uncertainty and systematic errors in the MILCANN map

5.1. Noise and CIB-residual simulations

We have shown qualitatively that the MILCANN tSZ map presents a significantly reduced background compared to the input MILCA tSZ map. In this section, we describe our modeling of the noise and CIB residuals in the MILCA and MILCANN tSZ maps to effectively quantify the improvement obtained by the ANN weighting process.

The tSZ maps derived from component separation methods are constructed through linear combination of *Planck* frequency maps that depends on the angular scale and the pixel, p , as

$$\hat{y} = \sum_{i,\nu} w_{i,p}(\nu) T_{i,p}(\nu), \quad (8)$$

where $T_{i,p}(\nu)$ is the *Planck* map at frequency ν for the angular filter i , and $w_{i,p}(\nu)$ are the weights of the linear combination (MILCA weights in this study). The CIB contamination (or leakage) in the y -map reads,

$$y_{\text{CIB}} = \sum_{i,\nu} w_{i,p}(\nu) T_{i,p}^{\text{CIB}}(\nu), \quad (9)$$

where $T^{\text{CIB}}(\nu)$ is the CIB emission at frequency ν . Using the weights $w_{i,p}(\nu)$, and considering the CIB luminosity function, it is possible to compute the expected CIB leakage as a function of redshift by propagating the SED through the weights used to build the tSZ map. As shown by [Planck Collaboration XXIII \(2016\)](#), the CIB at low- z leaks with a small amplitude in the tSZ map, whereas high- z CIB produces a higher, dominant level of leakage.

The CIB power spectra have been constrained in previous analyses (see e.g., [Planck Collaboration XXX 2014](#)); they can be used to predict the expected CIB leakage. To do so, we performed 200 Monte-Carlo simulations of multi-frequency CIB maps that follow the CIB auto- and cross-power spectra. We

then added instrumental noise to the simulated CIB maps before applying the MILCA weights used to build the input tSZ map. We obtained 200 realizations of instrumental noise and CIB in the MILCA tSZ map, consistent with noise and CIB observed in the *Planck* frequency maps.

It is important to stress that the noise in the input MILCA tSZ map is by construction correlated with the noise in the frequency maps. Therefore, the noise on the ANN weights is also correlated with the noise in the MILCA tSZ map. Consequently, to produce a fair description of the noise, we trained other ANNs on the simulated maps to reproduce the correlation feature between the noise in the MILCA map and the noise in the ANN weights. For completeness and considering that the training of an ANN is a nonlinear process, we also added CMB, point sources, and thermal dust to the noise+CIB simulations during the training process.

Finally, we built and applied these noise-based ANN weights to the MILCA noise+CIB-residuals simulation. In [Fig. 7](#), we present a simulation of noise+CIB residuals before and after applying the noise-based ANN weights. We observe that the weighting process allows us to significantly reduce the noise level in the simulated MILCA map.

In [Fig. 8](#), we compare the intensity distributions in MILCA and MILCANN maps. For the MILCA map, we observe a significant tail of pixels with negative intensity (mainly associated with radio-source contamination). We do not observe this contamination in the MILCANN map because the ANN weighting process significantly reduces radio source contamination. We also observe that the noise in the MILCANN map is lower than in the MILCA map by a factor of five. However, we note that the intensity of the brightest pixels in the map is not affected by the ANN weighting process.

Considering the correlation between the ANN weight and the noise in MILCA map, the noise in the MILCANN map does not present symmetric distribution. Therefore, we are dealing with a nonGaussian, inhomogeneous correlated noise with an asymmetric distribution. As observed in [Fig. 8](#), the noise is

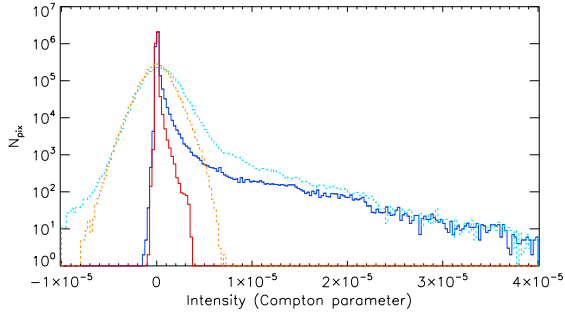


Fig. 8. Intensity distribution of pixels in the tSZ maps for the MILCA map (light blue), the MILCANN map (dark blue), the simulation of noise+CIB in the MILCA map (orange), and the simulation of noise+CIB in the MILCANN map (red).

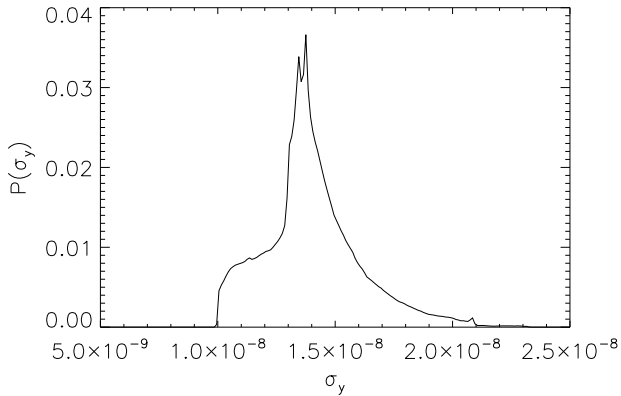


Fig. 9. Distribution of the noise standard deviation, σ_y , across the MILCANN full-sky map.

more likely to produce positive rather than negative values in the MILCANN map, implying a nonzero expectation value. Consequently, in the following we used and propagated the complete noise distribution.

5.2. Noise inhomogeneities

Due to the *Planck* scanning strategy, the noise level on the sky is inhomogeneous. The most noticeable feature is the fact that ecliptic poles present a higher redundancy of observations and thus a significantly lower noise level (Planck Collaboration I 2016). The MILCANN map obtained from the product of the filtered MILCA map and the ANN weight therefore exhibits noise inhomogeneities amplified from the input noise in MILCA map.

From the instrumental noise+CIB MILCANN simulated map, we derived the standard deviation of the noise in MILCANN map, σ_y , by computing the local standard deviation of the MILCANN simulated noise map within a four-degree Gaussian window. The distribution of noise standard deviation, σ_y , is shown in Fig. 9. This latter represents the distribution of the pixel-dependent noise levels.

We constructed a map of the signal-to-noise ratio, \widehat{y}_σ , as

$$\widehat{y}_\sigma = \frac{\widehat{y}_f}{\sigma_y}. \quad (10)$$

We note that using a unique threshold on \widehat{y}_σ is equivalent to using a pixel-dependent threshold on \widehat{y}_f .

Figure 10 presents the distributions of the MILCANN map and the MILCANN noise simulation as a function of y_σ and Q_N . We observe that the MILCANN noise simulation does not

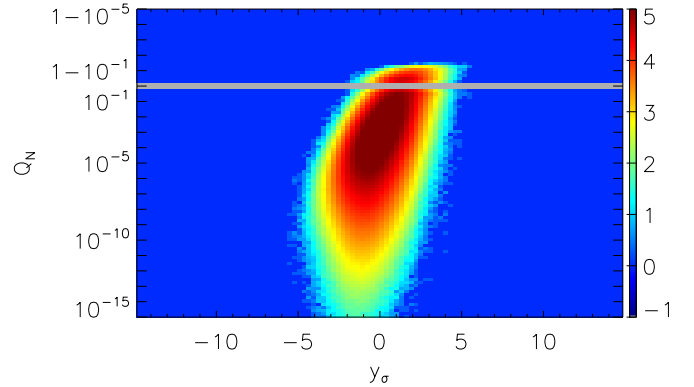
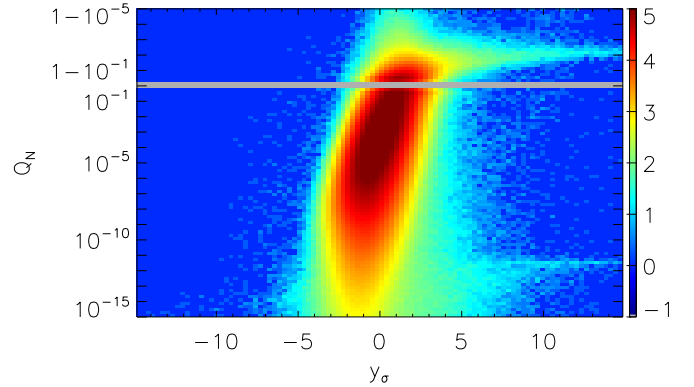


Fig. 10. Distribution of the MILCANN map (*top panel*) and MILCANN noise simulation (*bottom panel*) pixels as a function of Q_N and normalized intensity, y_σ . The colors show the number of pixels in logarithmic scale. To display the large dynamical range of the y axis, we divided the axis into two log scales with a “regular” log scale for small Q_N values and an “inverted” log scale for large Q_N values.

show high- Q_N and high- y_σ pixels ($Q_N > 0.9$, which are associated with real tSZ signal). We also observe that for $Q_N \approx 0$ the MILCANN map presents a significantly larger distribution of y_σ than the noise simulation (for $Q_N \in [10^{-11}, 10^{-15}]$). This extended distribution is produced by foreground residuals that are present in the MILCA tSZ map. We verified that these residuals are strongly correlated with galactic latitude, implying that they are related to systematic effects. We do not observe a similar behavior at larger values of Q_N . This confirms that foreground residuals are strongly reduced using the ANN weighting procedure, as already observed in Fig. 8.

6. tSZ candidate detection

As shown by Melin et al. (2012), tSZ map-based galaxy cluster detection methods suffer from a high level of contamination by spurious sources since it is difficult to disentangle real tSZ emission from biases in the tSZ map induced by residuals from astrophysical emissions. Given its significantly reduced residual signals, the MILCANN map may be better suited to detection of galaxy clusters than standard reconstructed y maps. In this section, we therefore use the improved tSZ map obtained after applying the ANN-based weight to perform a basic cluster detection. We also characterize the purity and completeness of the derived sample of cluster candidates to assess the improvement compared to a previous cluster catalog derived from *Planck* data. However, we stress that the MILCANN map cannot be used to

provide accurate estimates of the flux or the shape of tSZ sources considering that the tSZ signal is affected by the ANN weighting response.

6.1. Methodology

To detect sources in the MILCANN map, we applied a mask of the galactic plane and point sources detected by *Planck* keeping 84% of the sky as defined in [Planck Collaboration Int. XXXII \(2015\)](#). Subsequently, using the noise standard-deviation map computed in Sect. 5, we divided the MILCANN map by its local noise level map to perform the detection in signal-to-noise units. We applied a threshold of $y_\sigma > 3$ to the MILCANN map and considered any adjacent pixels found to be above the threshold as a candidate tSZ source. We discarded all sources detected on less than 5 adjacent pixels of 1.7×1.7 arcmin² to avoid detections produced by anomalous pixels. We cleaned multiple detections of the same source by merging all sources in a radius of 10 arcmin. We obtained a sample of 3969 tSZ sources that we refer to as the HAD catalog³.

6.2. Characterization of the MILCANN detection method

In this section, we present a detailed description of each step in the computation of the selection function of the HAD catalog, that is, a detailed description of the impact of the transfer function on the galaxy-cluster signal.

6.2.1. Fourier space filtering response

Before applying the ANN weight to build the MILCANN map, we filtered the MILCA map with the matched-filter presented in Fig. 1. Here, we present the estimation of the transfer function of this filtering process. To do so, we first build a mock map of a sky-projected tSZ signal from a galaxy cluster with $Y_{500} = 1$ arcmin² assuming a Generalized Navarro-Frenk-White (GNFW) pressure profile ([Arnaud et al. 2010](#)) with 1000 pixels per R_{500} . We then convolve the tSZ mock map by the instrumental beam and by the matched-filter presented in Sect. 3. We perform this procedure for values of R_{500} ranging from 0.1 to 100 arcmin. Finally, we extract the tSZ intensity at the center of the galaxy cluster on the convolved mock map.

Figure 11 presents the tSZ intensity, $y_f^{(1)}$, after applying the matched-filter for a galaxy cluster with a universal pressure profile and a flux $Y_{500} = 1$ arcmin² as a function of the apparent size on the sky, θ_{500} . The matched-filter we use selects compact objects (of typical size 5 arcmin) and thus presents a response that significantly reduces the flux of extended galaxy clusters. However, this is not an important limitation since our main goal is to detect compact tSZ sources associated with new galaxy clusters that are either low-mass or high- z . Considering the resolution of *Planck* tSZ maps (roughly 7 arcmin) such galaxy clusters are point-like. Furthermore, for these clusters or for more extended ones, we can compute their tSZ signal directly from the MILCA map or from the frequency maps.

6.2.2. Completeness

Given all the steps detailed above, we can express the complete processing we applied as

³ The catalog is available in the download section of <http://szcluster-db.ias.u-psud.fr>

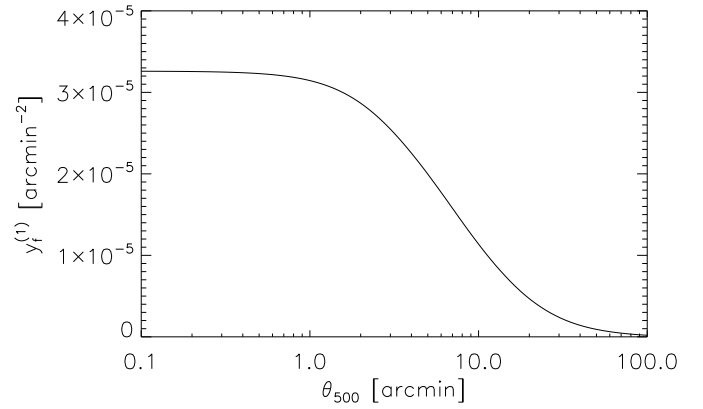


Fig. 11. Central intensity of the SZ signal in a filtered map for a galaxy cluster with $Y_{500} = 1$ arcmin² as a function of typical galaxy cluster radius θ_{500} .

$$y_\sigma = Y_{500} \frac{y_f^{(1)}(\theta_{500}) Q_N(y_f)}{\sigma_y} + N_\sigma, \quad (11)$$

where Y_{500} is the tSZ flux of the galaxy cluster, $y_f^{(1)}$ is the match-filtered central intensity for a cluster with $Y_{500} = 1$ arcmin², as shown in Fig. 11, the dependency of Q_N with y_f is presented in Fig. 6, the distribution of the noise across the map, σ_y , is shown in Fig. 9, and N_σ is the homogenized noise⁴ in the MILCANN map.

The galaxy cluster signal probability distribution is obtained by the convolution of: (i) the $M_{500} - Y_{500}$ relation intrinsic scatter, (ii) the distribution of σ_y (noise inhomogeneity), and (iii) the distribution of the noise N_σ (noise probability distribution). We assumed that the relation $M_{500} - \theta_{500}$ does not present any scatter.

The completeness, $C(y_\sigma)$, is then given by the ratio of the integral of y_σ distribution, $\mathcal{P}(y_\sigma)$, above the detection threshold normalized by the integral of the full distribution,

$$C(y_\sigma) = \frac{1}{\int_0^\infty \mathcal{P}(y_\sigma) dy_\sigma} \int_t^\infty \mathcal{P}(y_\sigma) dy_\sigma, \quad (12)$$

where t is the detection threshold applied on the y_σ map ($t = 3$ in our case).

Figure 12 shows the completeness as a function of mass, M_{500} , and redshift, z , of a given cluster. We observe that, with a very basic detection method applied on a filtered and cleaned tSZ map, we can detect clusters down to a typical mass of $M_{500} = 1 \times 10^{14} M_\odot$ with percent level completeness. We also observe that for very large mass ($> 2 \times 10^{15} M_\odot$) the completeness is slightly smaller than one. This effect is produced by the matched-filter that significantly reduces the tSZ effect produced by extended (massive) sources.

6.2.3. Purity

We estimated the purity of the catalog by performing the detection of tSZ sources on the MILCANN map and the MILCANN noise simulation from Sect. 5. This estimate does not account for all foreground residuals in the MILCANN map which therefore may slightly overestimate the purity of the HAD catalog. We found $N_{\text{det}}^{(1)}$ detections for the MILCANN map and $N_{\text{det}}^{(2)}$ for

⁴ That can be modeled through noise+CIB residuals simulations normalized by the standard deviation map σ_y .

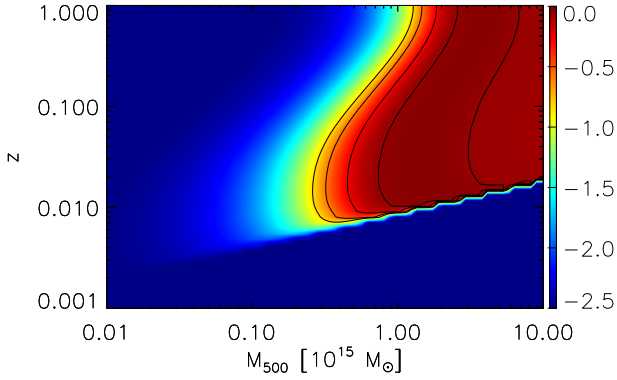


Fig. 12. Completeness of the HAD catalog as a function of M_{500} and z . The color scale is logarithmic and ranges from $10^{-2.5}$ (dark blue) to 1 (red). Black contours show the 10, 20, 50, and 90% completeness level.

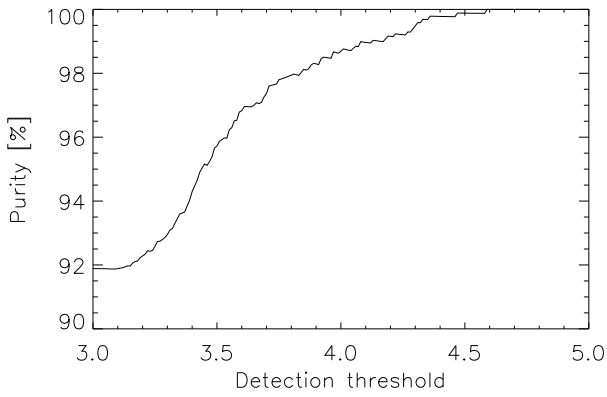


Fig. 13. Purity of the HAD catalog as a function of the detection threshold.

the simulated noise map. We performed the detection for several detection thresholds. The purity is obtained as $P = \frac{N_{\text{det}}^{(1)} - N_{\text{det}}^{(2)}}{N_{\text{det}}^{(1)}}$.

In Fig. 13, we present the purity as a function of the detection threshold. For the threshold $D_t = 3$ used in the construction of the HAD catalog, we derived an estimated purity of above 90%.

6.3. Comparison with reference galaxy cluster catalogs

In this section, we present a brief comparison of the HAD cluster candidate catalog and other reference catalogs. We used two approaches for the comparison. First, we compared the numbers of cluster candidates in the HAD catalog, with the predicted number of clusters assuming *Planck*-SZ cosmology ([Planck Collaboration XX 2014](#)) while considering the completeness and purity of the HAD catalog (see above). This predicted number is found to be 4082 ± 700 (the uncertainty is obtained by propagating the uncertainty over Ω_m and σ_8 to galaxy cluster number count). Thus, the 3969 detected candidates in the HAD catalog are consistent with this prediction. However, this number has to be considered carefully as the assumptions on the scaling relation and completeness may not encompass the full complexity of the cluster physics.

We then performed a cross-match with reference galaxy cluster catalogs. We compared our catalog of candidates with the PSZ2 catalog. The distribution of distance to nearest neighbor is shown in Fig. 14. Among the 3969 HAD sources, we find 1243 in common with the PSZ2 sources among which 997 are confirmed galaxy clusters all with an ANN quality factor $Q_{\text{Neural}} > 0.6$ as defined in [Aghanim et al. \(2015\)](#). The HAD catalog

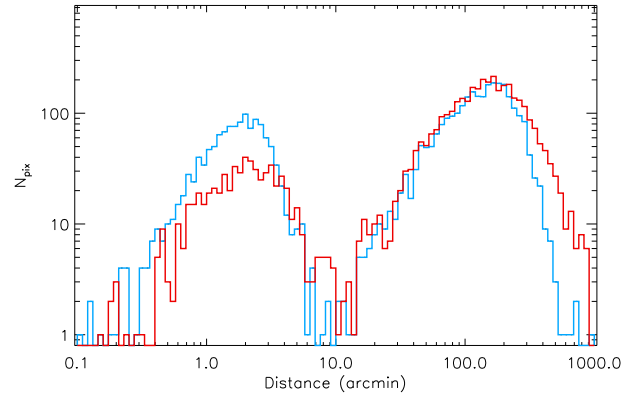


Fig. 14. Nearest-neighbor distance distribution between HAD sources and PSZ2 public catalog sources (blue histogram) and MCXC public catalog sources (red histogram).

additionally contains 496 known clusters that are missing from PSZ2 (they are contained in the ACT, SPT, redMaPPer, Wen+12, or MCXC catalogs). Figure 14 clearly shows the two populations of sources that are in common and not in common with PSZ2. For the sources common to HAD and PSZ2 the typical separation is below 4 arcmin, consistent with the *Planck* resolution. For very extended sources, the position mismatch between HAD and PSZ2 can reach up to 10 arcmin; it strongly depends on the exact methodology used to define the galaxy cluster position. Nevertheless, considering the number density in the *Planck* SZ and HAD catalogs, within a radius of 10 arcmin the number of chance association is 15 ($\approx 1\%$ of the PSZ2 sample) and within a radius of 4 arcmin this number is 4. This implies that a 10 arcmin matching distance still provides a robust association between HAD and PSZ2 sources.

Figure 14 presents the distribution of nearest-neighbor distance between HAD and MCXC objects. The matching procedure of the HAD catalog outputs the following positional associations with reference catalogs:

- 1243 HAD sources in common with PSZ2.
- 601 HAD sources matched with known X-ray clusters from MCXC (including 92 objects not in PSZ2).
- 687 HAD sources matched with over-density of galaxies ($N_{200} > 25$) from WHL12 (including 276 objects not in PSZ2); we estimated a maximum of 20 chance associations at 99% confidence level.
- 115 HAD sources matched with over-density of galaxies ($N_{500} > 8$) from new clusters of WHL15 (including 79 objects not in PSZ2); we estimated a maximum number of 18 chance associations at 99% confidence level.
- 1400 HAD sources matched with over-density of galaxies ($R_L > 10$) from WHY18 (including 649 objects not in PSZ2); we estimated a maximum number of 50 chance associations at 99% confidence level.
- 469 HAD sources matched with over-density of galaxies ($\Lambda > 50$) from redMaPPer (including 179 objects not in PSZ2).
- 35 HAD sources matched with SPT clusters.
- 43 HAD sources matched with ACT clusters.

The cross-matched numbers are summarized in Table 1. Objects present in the PSZ2 catalog that are not present in the HAD catalog (418 sources in total) are extended sources (smeared out by our matched-filter) or sources with a very-low-quality flag from the ANN. A low-quality-ANN quality assessment implies either that these sources are spurious detections or that they show contamination by at least another type of astrophysical emission.

Table 1. Cross-matched number of objects between HAD and reference catalogs.

	HAD	PSZ2	SPT	ACT	MCXC	WHL12	WHL15	WHY18	redMaPPer	TOTAL
N_{obj}	3969	1653	224	182	1743	9951	8625	47594	5540	
HAD	/	1243	35	43	601	687	115	1400	469	1803
PSZ2	1235	/	27	30	556	427	48	881	312	1134

Notes. We also performed the same cross-match for the PSZ2. The total column refers to the number objects from HAD or PSZ2. For the Wen+12 catalog we imposed $N_{200} > 25$, for Wen+15 we imposed $N_{500} > 8$, and for redMaPPer we imposed $\lambda > 50$.

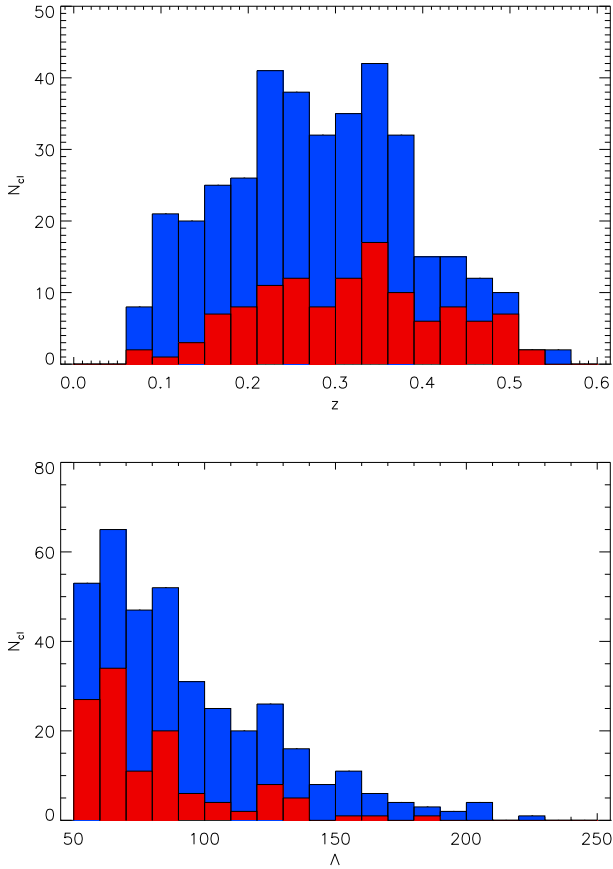


Fig. 15. *Top panel:* redshift distribution for HAD candidates matching redMaPPer overdensities of galaxies (dark blue) and for HAD candidates not contained in the PSZ2 catalog (red). *Bottom panel:* richness, Λ , distribution for HAD candidates matching redMaPPer overdensities of galaxies (dark blue) and for HAD candidates not contained in PSZ2 catalog (red).

We used the redMaPPer red-sequence-based redshifts to compute the redshift and richness, Λ , distributions of the matching population between HAD and redMaPPer catalogs in two cases: (i) considering all cluster candidates in the HAD catalog that match redMaPPer sources and (ii) candidates in the HAD catalog not contained in the PSZ2 catalog that match redMaPPer sources.

Figure 15 present the redshift distribution for sources contained both in HAD and redMaPPer catalogs. We observe that new tSZ sources in the HAD catalog, that is those not contained in the PSZ2 catalog are at higher redshift than PSZ2 clusters on average. This redshift distribution confirms that the MILCANN tSZ map enables the detection of high- z galaxy clusters that are not contained in the PSZ2 catalog. From this redshift distribution, we find that new HAD sources are at redshifts ranging from

0.2 to 0.5. Figure 15 also shows the richness distribution for all HAD sources and HAD sources not contained in the PSZ2 that match redMaPPer galaxy overdensities. We observe that the HAD sources not contained in the PSZ2 catalog are preferentially at lower richness.

7. Multi-wavelength statistical characterization of the HAD sample

In this section, we perform a stacking analysis to unveil the average properties of the HAD sources. In particular, we focus on the average submillimeter SED, on the stacked lensing signal, and on color-color diagnostic using the WISE galaxy catalog.

7.1. Stacked SED of cluster candidates

For cluster candidates in HAD not seen in PSZ2, we stack the *Planck* maps per frequency and the IRIS full-sky map at $100 \mu\text{m}$ (Miville-Deschênes & Lagache 2005). We then measure the flux through aperture photometry. Figure 16 presents the obtained SED exhibiting both tSZ and IR emissions. Consistently with Planck Collaboration XXIII (2016), we modeled the IR emission with a modified black-body SED assuming $\beta_d = 1.75$, a dust temperature $T_d = 24 \text{ K}$, and a mean redshift $\bar{z} = 0.4$. Comparing with previous analysis (Planck Collaboration XXIII 2016), the amount of IR emission toward new HAD tSZ sources is compatible with IR emission observed toward PSZ2 confirmed galaxy clusters. We observed that the IR contribution is negligible at 100 and 143 GHz. At 353 and 545 GHz, the IR emission contributes 25 and 80% of the total signal, respectively. At higher frequency, the tSZ contribution is negligible.

7.2. CMB lensing

We stacked the CMB lensing convergence measured by the Planck Collaboration (Planck Collaboration XV 2016) for all sources from the HAD catalog that are not included in the PSZ2 catalog. In Fig. 17, we present the derived stacked signal. We detect an excess in convergence at a 5σ confidence level. Assuming a typical redshift in the range 0.2–0.5 for galaxy clusters, and correcting for purity we compute an average mass of about $2 \times 10^{14} M_\odot$ per source (see e.g., Melin & Bartlett 2015, for a detailed description of the convergence to mass conversion).

7.3. WISE catalog

We also searched for counterparts of the HAD SZ candidates in the AllWISE full-sky catalog (Cutri et al. 2013). We stacked the sources in the AllWISE catalog toward HAD sources that have no PSZ2 counterpart. The stacked density of sources is presented in Fig. 18. We observe a significant AllWISE source overdensity

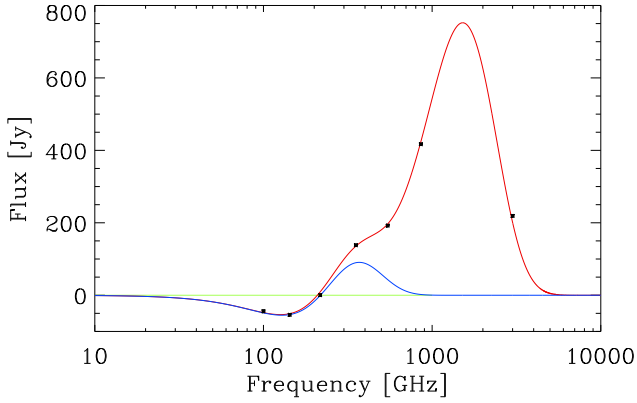


Fig. 16. Stacked SED of galaxy-cluster candidates (black sample). The tSZ effect contribution is shown as a solid blue line, and the total SED accounting for tSZ and infra-red emission is shown as a solid red line.

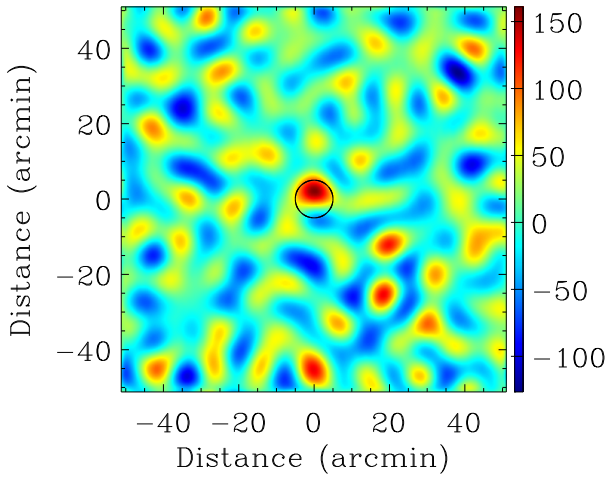


Fig. 17. Stacked CMB-lensing convergence map toward HAD SZ candidates not contained in the PSZ2 catalog.

of 23 ± 1^5 per HAD source inside an aperture of 10 arcmin; the background level of the AllWISE source-density map was estimated to be between 10 and 15 arcmin.

We also studied the distribution of the AllWISE matching sources in the AllWISE color-color plane. The surface density of HAD-source member galaxies in the AllWISE catalog is small compared to the total density of all objects. Consequently, we cannot estimate the AllWISE colors for each member galaxy individually and we estimated the color-color distribution of the cluster-member galaxies.

We first compute the color-color distribution of AllWISE sources, D_{in} , within a radius of 10 arcmin around the HAD sources. We then compute the same distribution, D_{out} , for sources located between 10 and 15 arcmin from the HAD sources. Assuming that the background and foreground objects are uniformly distributed on a 15 arcmin scale, we estimate the distribution for member galaxies by computing the sky-area weighted difference between the two distributions,

$$D_{\text{cl}} = D_{\text{in}} - \frac{A_{\text{in}}}{A_{\text{out}}} D_{\text{out}}, \quad (13)$$

⁵ The uncertainty provided is on the average over-density of sources and does not account for intrinsic scatter between sources in the stacked sample.

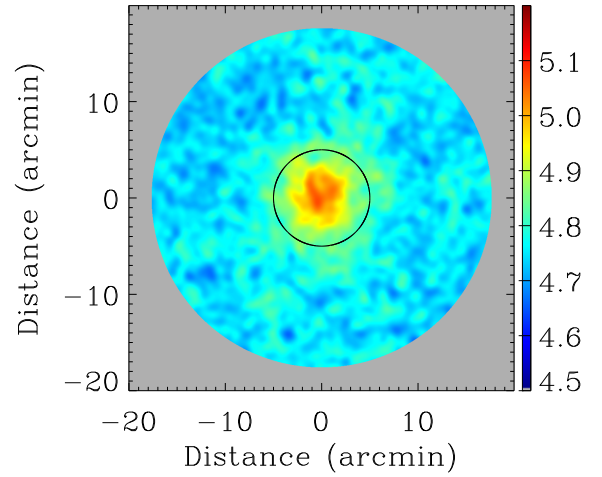


Fig. 18. Stacked number of sources in the AllWISE catalog toward SZ candidates not contained in the PSZ2 catalog.

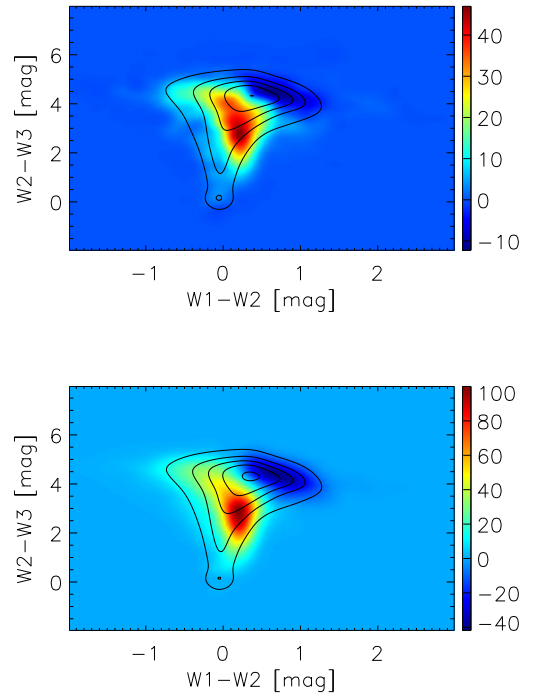


Fig. 19. WISE color-color plane for HAD SZ sources not contained (top panels) and contained (bottom panels) in the PSZ2. The color scale shows the distribution for member galaxies. For comparison, the distribution for background objects is represented by black contours.

where D_{cl} is the color-color distribution of galaxy cluster members, and A_{in} and A_{out} correspond to the areas of the regions used to compute D_{in} and D_{out} .

In Fig. 19, we present the color-color distributions of member galaxies of HAD sources included and not included in the PSZ2 catalog. We observe that both populations present similar distributions in the $(W2-W3)-(W1-W2)$ color-color plane. These distributions are significantly different from the distribution of background in the WISE catalog. They show a positive excess for $W1-W2 \approx 0.2$ and a lack of objects for $W2-W1 \approx 0.8$.

We use the SWIRE galaxy template library (Polletta et al. 2007) and compute the expected tracks in the WISE color-color plane for 25 galaxy SED templates. When comparing with tracks

for various galaxy types, we observe that the negative decrement in the WISE color–color plane is essentially populated by AGNs or high-redshift IR sources. This implies that tSZ-based catalogs present a selection bias toward clusters not hosting bright AGNs or clusters that have a strong IR source in the foreground and/or background. Indeed, for clusters that host a bright radio-loud AGN, the tSZ effect cannot be recovered in the microwave or submillimeter domain. A similar argument applies when the tSZ effect from a galaxy cluster is contaminated by bright IR sources. From these color–color distributions, we can conclude that the cluster candidates in the HAD catalog are populated by low-redshift ($z < 1$) elliptical, spiral galaxies, and Luminous Red Galaxies (LRGs) as expected for tSZ samples derived with the *Planck* experiment.

8. Conclusion

Previous studies have shown that a tSZ-map-based approach is not optimal and is less-efficient than a multi-frequency based approach (Melin et al. 2012) to detect clusters of galaxies via their tSZ signal. However, we have demonstrated that an ANN quality-assessed tSZ map, MILCANN, enables the construction of a competitive tSZ source catalog even with a simple detection method.

The matched-filtering and the ANN weighting process involved in the construction of the MILCANN tSZ map means that its use is specifically tailored to tSZ-cluster detection. In particular, the MILCANN tSZ map presents a significantly lower level of noise and foreground residuals than standard tSZ maps. However, the ANN weighting procedure produces a distortion of the tSZ signal both in shape and flux. Consequently, the MILCANN map can only be used for cluster-detection purposes and is not suited for other analyses such as tSZ scaling relations, profiles, or angular power spectra.

From the MILCANN tSZ map, we constructed the HAD source catalog containing 3969 cluster candidates with an estimated purity of 90%. This catalog is more than twice as large as the *Planck* catalog (Planck Collaboration XXVII 2016), achieving cluster detection down to $M_{500} = 10^{14} M_{\odot}$, and reaches the same purity level.

We verified that the number of sources in the HAD catalog is consistent with the expected cluster abundance. Additionally, comparing the HAD catalog with ancillary catalogs, we demonstrate that the HAD galaxy clusters catalog contains new tSZ detections at high redshift and low richness. Finally, we show that the sources detected in the MILCANN map present an excess of convergence in the *Planck* CMB lensing map, compatible with cluster masses of $2 \times 10^{14} M_{\odot}$. These sources host galaxies with the same spectral behavior as *Planck* PSZ2 galaxy cluster member galaxies (ellipticals, spiral galaxies, and LRGs at $z < 1$).

Acknowledgements. The authors thank an anonymous referee for the comments that improved the publication. They thank the International space science institute (ISSI) for hosting initial discussions during the meetings of the international team meeting “SZ effect in the *Planck* era”. GH acknowledges the support of the Agence Nationale de la Recherche through grant ANR-11-BS56-015. This publication used 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. It made use of the

SZ-Cluster Database (http://szcluster-db.ias.u-psud.fr/sitools/client-user/SZCLUSTER_DATABASE/project-index.html) operated by the Integrated Data and Operation Centre (IDOC) at the Institut d’Astrophysique Spatiale (IAS) under contract with CNES and CNRS. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme grant agreement ERC-2015-AdG 695561.

References

- Abell, G. O., Corwin, H. G., Jr., & Olowin, R. P. 1989, *ApJS*, 70, 1
- Aghanim, N., Hurier, G., Diego, J.-M., et al. 2015, *A&A*, 580, A138
- Aghanim, N., Douspis, M., Hurier, G., et al. 2019, *A&A*, 632, A47
- Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010, *A&A*, 517, A92
- Birkinshaw, M. 1999, *Phys. Rep.*, 310, 97
- Bleem, L. E., Stalder, B., de Haan, T., et al. 2015, *ApJS*, 216, 27
- Bohringer, H., Voges, W., Huchra, J. P., et al. 2000, *VizieR Online Data Catalog: J/ApJS/129/435*
- Böhringer, H., Schuecker, P., Guzzo, L., et al. 2001, *A&A*, 369, 826
- Carlstrom, J. E., Holder, G. P., & Reese, E. D. 2002, *ARA&A*, 40, 643
- Carvalho, P., Rocha, G., & Hobson, M. P. 2009, *MNRAS*, 393, 681
- Cutri, R. M., Wright, E. L., Conrow, T., et al. 2013, *Explanatory Supplement to the ALLWISE Data Release Products*, 1
- Dunkley, J., Hlozek, R., Sievers, J., et al. 2011, *ApJ*, 739, 52
- Ebeling, H., Edge, A. C., Allen, S. W., et al. 2000, *VizieR Online Data Catalog: J/MNRAS/318/333*
- Ebeling, H., Edge, A. C., & Henry, J. P. 2001, *ApJ*, 553, 668
- Gladders, M. D., & Yee, H. K. C. 2005, *ApJS*, 157, 1
- Hilton, M., Hasselfield, M., Sifón, C., et al. 2018, *ApJS*, 235, 20
- Hurier, G., & Lacasa, F. 2017, *A&A*, 604, A71
- Hurier, G., Macías-Pérez, J. F., & Hildebrandt, S. 2013, *A&A*, 558, A118
- Koester, B. P., McKay, T. A., Annis, J., et al. 2007, *ApJ*, 660, 239
- Mainzer, A., Bauer, J., Grav, T., et al. 2011, *ApJ*, 731, 53
- Marriage, T. A., Acquaviva, V., Ade, P. A. R., et al. 2011, *ApJ*, 737, 61
- Melin, J.-B., & Bartlett, J. G. 2015, *A&A*, 578, A21
- 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
- Miville-Deschênes, M.-A., & Lagache, G. 2005, *ApJS*, 157, 302
- 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 VII. 2014, *A&A*, 571, A7
- Planck Collaboration IX. 2014, *A&A*, 571, A9
- Planck Collaboration XX. 2014, *A&A*, 571, A20
- Planck Collaboration XXI. 2014, *A&A*, 571, A21
- Planck Collaboration XXVIII. 2014, *A&A*, 571, A28
- Planck Collaboration XXX. 2014, *A&A*, 571, A30
- Planck Collaboration I. 2016, *A&A*, 594, A1
- Planck Collaboration XIII. 2016, *A&A*, 594, A13
- Planck Collaboration XV. 2016, *A&A*, 594, A15
- Planck Collaboration XXII. 2016, *A&A*, 594, A22
- Planck Collaboration XXIII. 2016, *A&A*, 594, A23
- Planck Collaboration XXVII. 2016, *A&A*, 594, A27
- Planck Collaboration Int. XXXII. 2015, *A&A*, 581, A14
- Polletta, M., Tajer, M., Maraschi, L., et al. 2007, *ApJ*, 663, 81
- Reichardt, C. L., Shaw, L., Zahn, O., et al. 2012, *ApJ*, 755, 70
- Remazeilles, M., Delabrouille, J., & Cardoso, J.-F. 2011, *MNRAS*, 410, 2481
- Rykoff, E. S., Rozo, E., Busha, M. T., et al. 2014, *ApJ*, 785, 104
- Salvati, L., Douspis, M., & Aghanim, N. 2018, *A&A*, 614, A13
- Shirokoff, E., Reichardt, C. L., Shaw, L., et al. 2011, *ApJ*, 736, 61
- Sievers, J. L., Hlozek, R. A., Nolta, M. R., et al. 2013, *JCAP*, 10, 60
- Sunyaev, R. A., & Zeldovich, Y. B. 1969, *Nature*, 223, 721
- Sunyaev, R. A., & Zeldovich, Y. B. 1972, *Comments Astrophys. Space Phys.*, 4, 173
- van der Burg, R. F. J., Aussel, H., Pratt, G. W., et al. 2016, *A&A*, 587, A23
- Wen, Z. L., & Han, J. L. 2015, *ApJ*, 807, 178
- Wen, Z. L., Han, J. L., & Liu, F. S. 2012, *ApJS*, 199, 34
- Wen, Z. L., Han, J. L., & Yang, F. 2018, *MNRAS*, 475, 343
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868