SPT-SZ MCMF: an extension of the SPT-SZ catalogue over the DES region

M. Klein [1]★ J. J. Mohr,[1,2] S. Bocquet,[1] M. Aguena [3] S. W. Allen,[4,5,6] O. Alves,[7] B. Ansarinejad [8]
M. L. N. Ashby,[9] D. Bacon,[10] M. Bayliss,[11,12] B. A. Benson,[13,14,15] L. E. Bleem,[14,16] M. Brodwin,[17]
D. Brooks,[18] E. Bulbul,[19] D. L. Burke,[20,21] R. E. A. Canning,[22] J. E. Carlstrom,[13,14,16,23,24]
A. Carnero Rosell,[3,25,26] J. Carretero,[27] C. L. Chang,[13,14,16] C. Conselice,[28,29] M. Costanzi,[30,31,32]
A. T. Crites,[13,14,33] L. N. da Costa,[3] M. E. S. Pereira,[34] T. M. Davis,[35] J. De Vicente,[36] S. Desai,[37]
T. de Haan,[38] M. A. Dobbs,[39,40] P. Doel,[18] I. Ferrero,[41] A. M. Flores,[4,5] J. Frieman,[14,42] E. M. George,[43]
G. Giannini,[27] M. D. Gladders,[13,14] A. H. Gonzalez,[44] S. Grandis,[45] D. Gruen,[1] R. A. Gruendl,[46,47]
G. Gutierrez,[42] N. W. Halverson,[48] S. R. Hinton,[35] G. P. Holder,[39] D. L. Hollowood,[49] W. L. Holzapfel,[50]
K. Honscheid,[51,52] J. D. Hrubes,[53] N. Huang,[50] D. J. James,[54] G. Khullar,[13,14] K. Kim,[12] L. Knox,[55]
R. Kraft,[9] F. Kéruzoré,[56] A. T. Lee,[50,57] D. Luong-Van,[53] G. Mahler,[58,59] A. Mantz,[4,5] D. P. Marrone,[60]
J. L. Marshall,[61] M. McDonald,[11] J. J. McMahon,[7] J. Mena-Fernández,[62] F. Menanteau,[46,47]
S. S. Meyer,[13,14,23,24] R. Miquel,[27,63] J. Myles,[5,20,21] S. Padin,[13,14,33] A. Pieres,[3,64] A. A. Plazas Malagón,[20,21]
C. Pryke,[65] C. L. Reichardt,[8] K. Reil,[21] J. Roberson,[12] A. K. Romer,[66] C. Romero,[67] J. E. Ruhl,[68]
B. R. Saliwanchik,[69,70] L. Salvati,[31,32,71] E. Sanchez,[36] A. Saro,[30,31,32,72,73] K. K. Schaffer,[14,24,74]
T. Schrabback,[45] M. Schubnell,[7] I. Sevilla-Noarbe,[36] K. Sharon,[75] E. Shirokoff,[13,14] M. Smith,[76]
T. Somboonpanyakul,[77,78] B. Stalder,[9] S. A. Stanford,[55,79] A. A. Stark,[9] V. Strazzullo,[31,32,80] E. Suchyta,[81]
M. E. C. Swanson,[82] G. Tarle,[7] C. To,[51] K. Vanderlinde,[83,84] J. D. Vieira,[85,86] A. von der Linden,[87]
N. Weaverdyck,[7,88] R. Williamson,[13,14,33] P. Wiseman[76] and M. Young[84]

Affiliations are listed at the end of the paper

ABSTRACT
We present an extension to a Sunyaev–Zel'dovich Effect (SZE) selected cluster catalogue based on observations from the South Pole Telescope (SPT); this catalogue extends to lower signal to noise than the previous SPT–SZ catalogue and therefore includes lower mass clusters. Optically derived redshifts, centres, richnesses, and morphological parameters together with catalogue contamination and completeness statistics are extracted using the multicomponent matched filter (MCMF) algorithm applied to the S/N > 4 SPT–SZ candidate list and the Dark Energy Survey (DES) photometric galaxy catalogue. The main catalogue contains 811 sources above S/N = 4, has 91 per cent purity, and is 95 per cent complete with respect to the original SZE selection. It contains in total 50 per cent more clusters and twice as many clusters above $z = 0.8$ in comparison to the original SPT-SZ sample. The MCMF algorithm allows us to define subsamples of the desired purity with traceable impact on catalogue completeness. As an example, we provide two subsamples with S/N > 4.25 and S/N > 4.5 for which the sample contamination and cleaning-induced incompleteness are both as low as the expected Poisson noise for samples of their size. The subsample with S/N > 4.5 has 98 per cent purity and 96 per cent completeness and is part of our new combined SPT cluster and DES weak-lensing cosmological analysis. We measure the number of false detections in the SPT-SZ candidate list as function of S/N, finding that it follows that expected from assuming Gaussian noise, but with a lower amplitude compared to previous estimates from simulations.

Key words: galaxies: clusters: general – galaxies: clusters: intracluster medium – galaxies: distances and redshifts.

1 INTRODUCTION

Within the past 10–20 yr, cluster catalogues based on the detection of the intracluster medium (ICM) either via its X-ray emission or via the Sunyaev–Zel'dovich Effect (SZE) signature have grown from just tens of systems to thousands of systems (Lahav et al. 1989; Böhringer et al. 2000; Vanderlinde et al. 2010; Bleem et al. 2015, 2020; Finoguenov et al. 2020; Hilton et al. 2021; Klein et al. 2023) and will soon reach tens or even hundreds of thousands (Merloni et al. 2012; Raghunathan et al. 2022). These ICM-selected cluster catalogues

require optical follow-up to assign cluster redshifts and typically to confirm that the ICM-selected cluster candidate is physically associated with a cluster of galaxies.

With the availability of well-calibrated, large solid-angle optical surveys (e.g. SDSS, KIDS, DES, HSC-SSP, Legacy Surveys; de Jong et al. 2013; Flaugher et al. 2015; Blanton et al. 2017; Aihara et al. 2018; Dey et al. 2019) and mid-infrared surveys like that from the *Wide-Field Infrared Survey Explorer* (*WISE*; Wright et al. 2010), the confirmation and redshift assignment can be done systematically over large portions of the sky. In the past, the final cluster catalogues – especially those employed for cosmology – were often defined such that the confirmation and redshift assignment would have a negligible impact on the completeness of the original ICM candidate list (e.g. Mantz et al. 2010; Benson et al. 2013; Reichardt et al. 2013; Bocquet et al. 2015; de Haan et al. 2016). This approach is now coming to its limit, because the larger cluster samples needed for improved cosmological constraints require improved control over systematics – including even the impact of follow-up confirmation and redshift assignment on sample completeness. Moreover, using only information from the ICM-selected catalogue to produce a clean sample will lead to significantly smaller samples than would be possible using additional information from the optical follow-up.

Examples of combining X-ray-selected cluster candidate catalogues and systematic optical follow-up include the confirmation of *ROSAT*-selected clusters via DES (Klein et al. 2019), SDSS (Finoguenov et al. 2020), and the Legacy Survey DR10 (Klein et al. 2023). These efforts yielded thousands of new galaxy clusters extending to higher redshifts and with an angular density many times higher than previously selected *ROSAT* cluster samples that relied on individual cluster imaging and spectroscopy. The eFEDS X-ray survey (Brunner et al. 2022) carried out by *eROSITA* on the satellite Spektrum-Röntgen-Gamma (Predehl et al. 2021) has been analysed with multicomponent matched filter (MCMF) using HSC-SSP and Legacy Survey DR9 data yielding a 94 per cent pure sample with 477 confirmed clusters over 140 deg^2. A subset with 450 clusters was recently used for the first *eROSITA*-based cluster cosmology (Chiu et al. 2023). The usage of MCMF-based cleaning in this study allowed us to increase the sample useful for the cosmological study by more than a factor two compared to solely relying on X-ray data.

Similar systematic optical follow-up of SZE-selected cluster candidates has been carried out. The analysis of a set of SPTpol-ECS candidates was pursued with the redMaPPer algorithm (Rykoff et al. 2014) in targeted mode using DES data, supplemented with *WISE* and Panstarrs survey and pointed *Spitzer* IR and PISCO observations (Bleem et al. 2020). The ACT cluster candidate list has also been systematically followed up using DES and other data (Hilton et al. 2021). Recently, a new low S/N SZE-selected candidate list from the *Planck* mission data set has been followed up using the MCMF algorithm with DES data, resulting in the discovery of the highest redshift *Planck*-selected systems to date as well as a tripling of the number of confirmed *Planck* clusters in the DES survey footprint (Hernández-Lang et al. 2023).

In the analysis presented here, we carry out a similar study of confirming an SPT-SZ candidate list that extends to lower S/N than has been previously attempted in Bleem et al. (2015). We apply the MCMF algorithm to the SPT-SZ candidate list down to S/N = 4 using the DES and *WISE* data sets, cross-checking previously confirmed SPT-SZ clusters but also identifying many lower mass, previously undiscovered galaxy clusters. The S/N > 4.5 subsample from this paper, together with the SPT-ECS (Bleem et al. 2020) and SPTpol 500d (Bleem et al. 2024) sample build the basis of the most recent SPT-based cosmological analysis presented in Bocquet et al. (2023, 2024).

This paper is structured as follows. In Section 2, we describe the data set used in this work, and in Section 3, we outline the cluster confirmation method. The SPT-SZ MCMF cluster catalogue is presented in Section 4 and validated in Section 5. The conclusions are summarized in Section 6. Throughout this paper we adopt a flat ΛCDM cosmology with $\Omega_M = 0.3$ and $H_0 = 70 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$.

2 DATA

In this paper, we make use of the photometric catalogue from DES observations obtained within the first 3 yr of the survey (Y3) and the SPT–SZ cluster candidate list down to S/N = 4. For the high-z confirmation of cluster candidates, we further make use of mid-IR data from the *WISE* satellite (Wright et al. 2010; Mainzer et al. 2011) in the form of a matched catalogue between DES and the UnWISE catalogue (Schlafly, Meisner & Green 2019). The following subsections provide an overview of the data sets used.

2.1 The DES Y3A2 GOLD catalogue

For the optical confirmation out to $z \approx 1.3$, we make use of the DES Y3A2 GOLD catalogue, which is based on g, r, i, and z-band DECam (Flaugher et al. 2015) imaging data from DES between 2013 August and 2016 February. Details on the data reduction and data quality are given elsewhere (Abbott et al. 2018; Morganson et al. 2018).

The DES Y3A2 GOLD catalogue is a value-added version of the photometric catalogue released in the public data release 1 (DR1; Abbott et al. 2018). The catalogue covers approximately 5000 deg^2 in area with typically 3–5 exposures per band and reaches 95 per cent completeness limits of 23.72, 23.34, 22.78, and 22.25 mag in the g, r, i, and z bands, respectively. The catalogue includes additional calibration steps, flags, and types of photometry. In our work, we make use of the multi-epoch, multi-band, multi-object fitting photometry 'MOF', which is based on the NGMIX code (Sheldon 2014) and fits a galaxy model to each single epoch exposure and band at the same time, considering the different PSF shapes and sizes. Furthermore, it simultaneously fits neighbouring sources for improved deblending. In addition to MOF, we make also use of single-object fitting (SOF) photometry, which is derived in a similar way but masking neighbouring sources rather than fitting them. As SOF turned out to be more robust, while MOF provides better photometry in crowded regions, we make use of SOF photometry in cases where MOF has failed.

Out to $i = 22.2$ mag, we use the star-galaxy separator available in GOLD, which is an expanded version of that available in DES Y1A1 (Drlica-Wagner et al. 2018) and includes MOF/SOF-based extent information. For fainter magnitudes we do not apply a star-galaxy separation to maximize sensitivity to small, high-redshift cluster galaxies. The resulting impact on cluster richness from residual contamination by stars in the galaxy sample is minimized by using a local background measurement, which works well in the limit that the residual stellar density near the cluster position is nearly constant.

In addition, we make use of mask flags to exclude regions around bright stars and the 'top of the galaxy' calibration including SED-based de-reddening of sources due to interstellar dust provided in the DES Y3A2 GOLD catalogue.

2.2 The SPT-SZ cluster candidates with S/N>4

The SPT-SZ survey is based on observations with the SPT-SZ camera on the 10m SPT (Carlstrom et al. 2011), which has a 1 deg diameter field of view and a resolution of about ∼1 arcmin. The survey was

conducted from 2007 to 2011, covering 2500 deg^2 between 20 h < RA < 7 h and $-65°$ <Dec < $-40°$ and in three frequencies 95, 150, and 220 GHz. The source detection via the thermal SZE is performed on the 95 and 150 GHz maps using a matched-filter approach (Bleem et al. 2015). The SPT-SZ cluster candidate list contains 1518 sources with S/N > 4, of which 1395 (92 per cent) fall within unmasked areas of DES that are suitable for optical follow-up.

2.3 *WISE*

The *WISE* satellite is a mid-infrared telescope with a main mirror of 40 cm observing in four bands at 3.4 μm, 4.6 μm, 12 μm, and 22 μm ($w1$, $w2$, $w3$, $w4$). The *WISE* data set can be divided in three phases. The main phase performed 1.5 scans of the full sky and had sufficient cryogenic coolant to observe in the $w3$ and $w4$ bands. A second phase (NEOWISE) was performed immediately after the main campaign and with the goal of completing the second full-sky observations. Due to the lack of cryogenic cooling only observations in the $w1$ and $w2$ were possible. A third phase of *WISE* observations (NEOWISE-R) started in September 2013 when *WISE* was recommissioned after more than 2 yr of hibernation. Since then *WISE* completes one full-sky survey every ~6 months in the $w1$ and $w2$ bands.

In this work, we use the unWISE catalogue (Meisner et al. 2019) that makes use of all *WISE* data until the end of the first year of the NEOWISE-R phase. It is based on the unblurred coadds of *WISE* imaging data (unWISE Lang 2014) and includes improved source detection and deblending modelling for crowded regions. The catalogue yields a gain of 0.7 mag in depth and contains twice the number of galaxies with respect to the AllWISE catalogue (Cutri et al. 2013) that is based on solely the main and the NEOWISE phase of *WISE* observations.

3 CLUSTER CONFIRMATION METHOD

For cluster confirmation and redshift determination of the majority (>90 per cent) of SPT-SZ cluster candidates we use the multicomponent matched filter cluster confirmation tool (MCMF; see details in Klein et al. 2018, 2019) with DES photometric data. In Section 3.1, we summarize the method and describe some recent modifications.

From the previous SPT-SZ sample (Bleem et al. 2015), we expect a significant fraction (~8 per cent) to be at $z > 1$, where the DES imaging data need to be complemented with NIR or IR imaging. For that reason, we develop a high-z cluster confirmation tool, following the MCMF concept but using a combination of DES and *WISE* (mid-IR) photometry data. This is described in Section 3.2. Finally, we review the optical morphological measures that we extract for the SPT-SZ MCMF clusters in Section 3.3.

3.1 MCMF

The MCMF algorithm has been designed for the confirmation and characterization of ICM-selected cluster candidates identified in large X-ray or SZE surveys. MCMF has been successfully applied to ROSAT X-ray sources over the DES footprint (MARD-Y3; Klein et al. 2019) and more recently in combination with the Legacy Survey DR10 data set (Dey et al. 2019), it has been used to create the all-sky optically confirmed X-ray cluster catalogue (RASS-MCMF; Klein et al. 2023). In addition, it was used for the optical follow-up of the first *eROSITA*-based galaxy cluster catalogue over the early mission test field eFEDS (Klein et al. 2022). Beyond this, it has also been applied to new S/N > 3 Planck SZE-selected catalogues

over the DES region (MADPSZ; Hernández-Lang et al. 2023). In working with these different data sets, improvements and extensions to the original method have been made. In these applications, the new MCMF based catalogues significantly enhanced the number of clusters that had been previously extracted from the same X-ray or SZE data sets and followed up with cluster-by-cluster imaging and spectroscopy. In addition to enlarging the samples, the MCMF method allows one to limit the contamination of the new samples.

The MCMF algorithm includes a red-sequence technique (Gladders & Yee 2000; Rykoff et al. 2014) with redshift and magnitude-dependent colour filters in the $g − r$, $r − i$, and $i − z$ colours, a radial weighting (projected NFW profile centred at ICM-selected candidate location) and a characteristic magnitude range to estimate redshifts and richnesses for candidates. From the cluster candidate list, it makes use of the source position and an ICM-based mass proxy. The mass proxy is used to estimate the radius R_{500} within which galaxies are counted to estimate cluster richness. In this work, we make use of the SPT-SZ candidate S/N together with a calibration of the S/N-to-mass relation (Bocquet et al. 2019) to extract a cluster mass estimate for a range of hypothetical redshifts.

For each cluster candidate, the colour and radially weighted, background-subtracted richness $\lambda(z)$ within R_{500} is calculated as a function of (a priori unknown) candidate redshift. The peaks in $\lambda(z)$ are then identified and modelled with so-called peak profiles (see below). If present, multiple richness peaks (≤ 3) along the line of sight towards each candidate are recorded. Examples of peak profiles and their best fit to $\lambda(z)$ profiles of clusters are presented in previous MCMF analyses (e.g. Figs 4 and A2; Klein et al. 2019). These peaks with associated richnesses and redshifts are then collected and processed further as described in the following subsections.

Note that the peak profile models are built using renormalized stacks of individual $\lambda(z)$ profiles from clusters with spectroscopic redshift measurements (spec-z's). The clusters with spectroscopic redshift do not need to be part of the sample that is being studied. Important here is that the redshift dependency of the SZE observable-derived estimate of R_{500} be the same for the spec-z clusters and the candidates to be analysed. To ensure this, we assign a value ξ of the SZE observable S/N to the spec-z clusters that is consistent with their masses and redshifts that then can be used as input to the MCMF pipeline.

To confirm clusters, we characterize the likelihood that a given optical counterpart is a chance superposition rather than a physical counterpart to the ICM-selected cluster candidate. Doing so requires knowing the typical richness distribution of contaminants as a function of redshift within the survey region. Thus, the same exercise employed on the candidates is then repeated using random positions within the SPT-SZ footprint, using the same distribution of ξ and excluding regions containing SPT–SZ detections. These random positions provide the richness distribution of non-SZE-detected structures (noise, projections, undetected clusters). The richness distributions from the random lines of sight and true clusters are redshift dependent, because they are impacted by the selection function, the evolution of the halo–mass function and the noise in the richness estimate.

To be able to control the contamination of the final cluster sample, we calculate a quantity f_{cont}. High values of f_{cont} indicate a higher probability that the candidate in question is a chance superposition rather than a real cluster. f_{cont} is calculated using the mean richness distributions along the random lines of sight $f_{rand}(\lambda, z)$ and the richness distributions $f_{obs}(\lambda, z)$ towards the candidates. That is, for each candidate i, we calculate the number of random lines of sight

within a redshift bin with richness $\lambda \geq \lambda_i$ and divide by the number of SZE candidates within the same redshift bin with $\lambda \geq \lambda_i$. This ratio is then re-scaled according to the total number of SPT candidates and random lines of sight:

$$f_{\mathrm{cont}}(\lambda_i, z_i) = \frac{\int_{\lambda_i}^{\infty} f_{\mathrm{rand}}(\lambda, z_i)\mathrm{d}\lambda}{\int_{\lambda_i}^{\infty} f_{\mathrm{obs}}(\lambda, z_i)\mathrm{d}\lambda}, \qquad (1)$$

This f_{cont} parameter is calculated for each richness peak associated with a candidate. The peak showing the lowest value of f_{cont} is assigned as the best optical counterpart for the SPT-SZ candidate, because it is the most likely to be a real cluster.

The cluster sample itself can then be defined as those candidates showing an f_{cont} below a certain threshold value $f_{\mathrm{cont}}^{\mathrm{max}}$. The threshold value corresponds to the fraction of the contamination in the initial candidate list that makes it into the final cluster catalogue. The contamination of the resulting final cluster catalogue would then be $f_{\mathrm{SZE-cont}} \times f_{\mathrm{cont}}^{\mathrm{max}}$, where the confirmed catalogue contains all candidates with $f_{\mathrm{cont}}(\lambda_i, z_i) \leq f_{\mathrm{cont}}^{\mathrm{max}}$ and the initial contamination of the candidate catalogue is $f_{\mathrm{SZE-cont}}$. As an example, if the input catalogue is known to be 50 per cent pure and an f_{cont} threshold value $f_{\mathrm{cont}}^{\mathrm{max}} = 0.2$ is employed, then the contamination fraction of the confirmed cluster catalogue would be $0.5 \times 0.2 = 0.1$ or 10 per cent.

The version of MCMF applied here is – aside from the different mass proxy – largely the same as the version applied previously to two previous X-ray samples (Klein et al. 2019, 2022). Some minor improvement we made on the estimate of the redshift uncertainty. Based on the analysis on mock data, that include effects such as scatter in photometric calibration, intrinsic and measurement scatter of cluster member galaxy colours and structures along the line of sight, we find that cluster photo-z scatter can be reasonably well described as $\sigma_z = f(z)/\sqrt{(\lambda)}$. Here, $f(z)$ is a scale factor as function of cluster redshift that is calibrated empirically with spectroscopic redshifts. The photometric redshift uncertainties that we list are therefore redshift and richness dependent. A second improvement specific to this work is a second iteration on the estimate of the richness distributions along random lines of sight. The richness distributions along random lines of sight f_{rand} is supposed to resemble the expected richness distribution of contaminants as function of redshift. Given the correlation between ξ, cluster mass and likelihood of a source being a real cluster, the initial choice of using the same ξ distribution as the full candidate list causes the estimate of f_{cont} to be mildly biased high. To avoid this bias, we use the ξ distribution of rejected systems ($f_{\mathrm{cont}} > 0.3$) as proxy of the distribution of contaminants and to select a subsample of randoms that follows this distribution and remeasure f_{cont} for all candidates.

3.2 High-redshift extension using *WISE*

Besides the fact that passive galaxies become fainter with increasing redshift, the rest-frame wavelength range covered by the DES bands no longer brackets the 4000 Å break at redshifts $z \gtrsim 1$. Therefore, photometric redshifts become increasingly uncertain at these redshifts. For high-redshift cluster confirmation and photometric redshifts, it is therefore advantageous to move to redder bands such as the mid-IR regime covered by the *Spitzer* or *WISE* satellites. Data from both satellites were previously used for high-redshift cluster searches (Muzzin et al. 2009; Gonzalez et al. 2019) as well as for cluster confirmation (Bleem et al. 2015, 2020).

In our current analysis, we use the unWISE catalogue (Schlafly et al. 2019) that additionally includes more recent $w1$ and $w2$ band
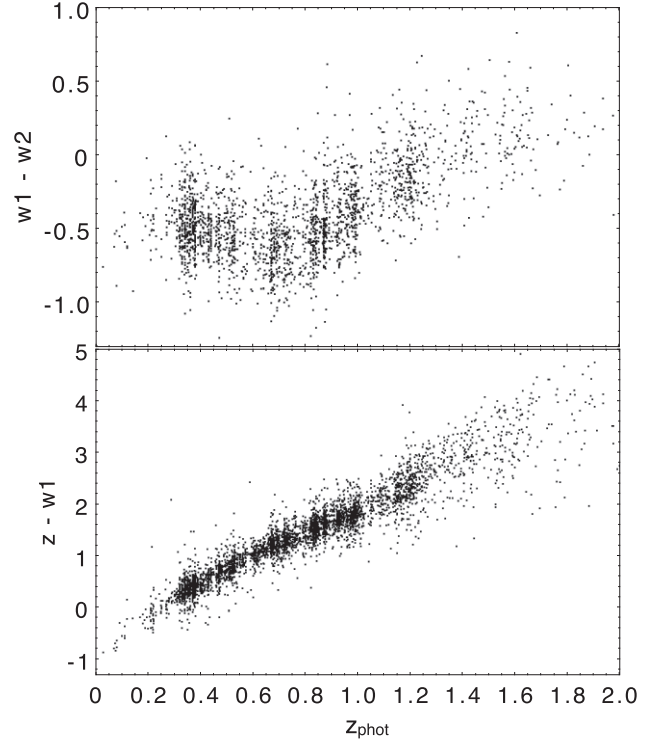
Figure 1. Passive galaxy colours versus photometric redshift (z_{phot}) in the COSMOS field, including the $w1-w2$ colour from the unWISE catalogue (top), and the DES z-band minus *WISE* $w1$ band colour (bottom). The observed galaxy colours suggest that cluster redshift constraints can be obtained out to $z \approx 1.5$ when adding *WISE* data to the DES data set.

WISE imaging data from the NEOWISE-R phase to create deep catalogues without PSF scale smoothing of the data and includes an improved modelling of crowded regions (Meisner et al. 2019). *WISE* data exist over the full sky, and therefore we match the unWISE and DES catalogues to allow for optical +IR photometry for all *WISE* sources. The optical to *WISE* galaxy colours (e.g. z-$w1$) have strong redshift dependence and are therefore well suited for getting high-quality cluster redshifts. The z-$w1$ and $w1$-$w2$ colour variation of passive galaxies with redshift is illustrated in Fig. 1 using DES measurements in the COSMOS field.

One downside of *WISE* $w1$ is the large PSF (~ 6 arcsec), which becomes a problem in dense regions such as the cores of galaxy clusters. In such dense regions, the separation of individual sources and the deblending of fluxes is a challenge. Here, the improved modeling of crowded regions in unWISE compared to previous ALLWISE catalogue becomes relevant.

Finally, we account for masks or missing data in the different surveys by deriving separate richnesses for cluster regions with different coverage (DES $+ w1$ only, *WISE* $w1$ only, *WISE* $w1$, $w2$, and DES) and sum them for the final cluster richness estimate. With this approach, we must track only the masked area in $w1$ imaging. The total richness in the high-redshift code is therefore given as

$$\lambda_{\mathrm{HZ}}(z) = \lambda_{\mathrm{DES}+w1}(z) + \lambda_{\mathrm{DES}+w1+w2}(z)$$
$$+ \lambda_{w1+w2}(z) + \lambda_{w1}(z), \qquad (2)$$

where the individual richnesses are defined in the same manner as the standard MCMF richness (see Klein et al. 2018, 2019), with the
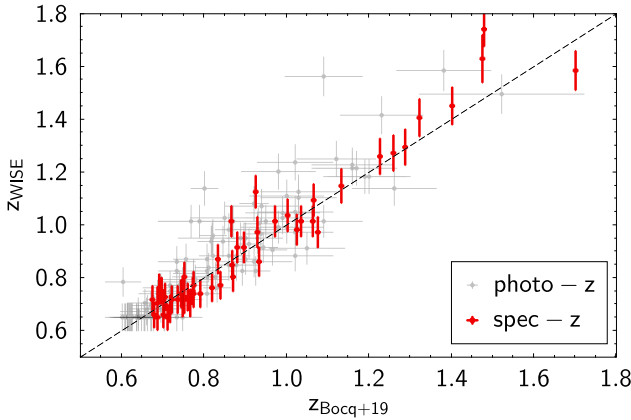
Figure 2. Comparison of redshift estimates from previous SPT-SZ catalogue (Bocquet et al. 2019) and those derived with the *WISE*-based high-z code. Spectroscopic redshifts are shown in red. Wise-based redshifts show generally good agreement with spectroscopic redshifts over the full redshift range, although they are only used for clusters at $z > 1$ in this work.

colour-weights depending on the availability of the bands (i, z, $w1$, $w2$ for $\lambda_{DES + w1 + w2}$, i, z, $w1$ for $\lambda_{DES + w1}$, $w1$, $w2$ for $\lambda_{w1 + w2}$, and no colour weight for λ_{w1}).

The high-z cluster confirmation code has been applied to all candidates and over a redshift range from $0.63 < z < 2.0$. Similar to the optically-based MCMF version, we perform runs along random lines-of-sight and calculate f_{cont} for potential counterparts to the SZE candidate. We make use of clusters with spectroscopic redshifts available for a subset of the sample and calibrate the *WISE*-based measurements. We further make use of the overlap in redshift between the optically-based MCMF and the high-z *WISE*-based run of $0.63 < z < 1.3$ to compare richnesses.

In Fig. 2, we show a comparison between the redshifts obtained with the high-z code and the redshifts provided for the previous catalogue (Bocquet et al. 2019) for clusters with $f_{cont} < 0.2$, showing reasonable agreement between *WISE* informed redshifts and the redshifts coming from dedicated optical, IR, and spectroscopic follow-up. To avoid complicated modelling of the richness-observable and richness–mass relation it is further favourable that both richnesses share an approximately similar scatter. For that reason, we investigate the ratio of richness to the SZE-based mass estimate. Using the assumption that the richness-mass slope is approximately one and that there is no redshift evolution of the scaling relation, this ratio is a measure of the scatter of the Lambda–mass relation. In Fig. 3, we show this ratio for the DES-only measurements of the $\xi > 4.25$ subsample (see Table 1) and also for the high-z code measurements, here with the additional requirement that the cluster redshift must lie at $z > 0.63$. As visible in Fig. 3, the width and the mean of the distributions for the high-redshift code and the DES-only code appear very similar. Both exhibit some deviation from a normal distribution. Fitting a normal distribution for the close-to-normal part of the distribution above $\log(\lambda/M_{500}) = 1.2$ yields consistent mean ratios ($\mu_{DES} = 1.43 \pm 0.02$, $\mu_{WISE} = 1.42 \pm 0.01$) and standard deviations ($\sigma_{DES} = 0.14 \pm 0.02$, $\sigma_{WISE} = 0.13 \pm 0.01$), providing evidence that the two richness measurements exhibit similar relations to the SZE-based mass estimates. The cross-over in the ability to confirm cluster candidates is in the $1 < z < 1.3$ regime where the mid-IR selection of *WISE* starts to see more of the cluster population than is visible in the DES data.
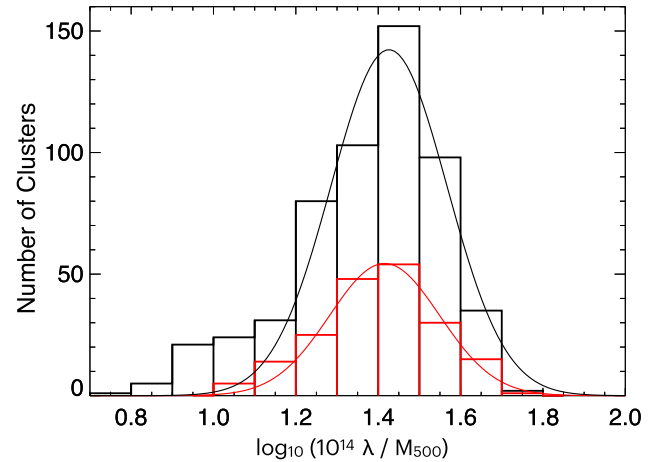
Figure 3. Distribution of richness λ over SZE-based mass estimate M_{500} for richness measurements from the DES-only MCMF code (black) and the *WISE*-based high-z code (red). Continuous curves show the best-fitting normal distributions to values above $\log(\lambda/M_{500}) = 1.2$ with best-fitting standard deviations of $\sigma = 0.14$ (black) and $\sigma = 0.13$ (red).

Table 1. Properties of the SPT-SZ MCMF cluster catalogue along with three subsamples. The table shows sample name, selection criteria f_{cont}^{max} and ξ_{min}, the expected final sample purity, the completeness with respect to the SZE selection, the total number of confirmed clusters and those above redshift $z = 0.25$.

Sample	f_{cont}^{max}	ξ_{min}	Purity (per cent)	Comp. (per cent)	N_{cl}	N_{cl} ($z > 0.25$)
SPT-SZ MCMF	0.20	4.00	91.0	95.0	811	733
$\xi > 4.25$	0.125	4.25	96.0	96.5	640	581
$\xi > 4.5$	0.13	4.50	98.0	96.5	527	480

3.3 Optical estimators of cluster dynamical state

Following our previous work on X-ray-selected clusters from *ROSAT* and *eROSITA* (Klein et al. 2019, 2022), we provide for SPT-SZ MCMF clusters estimators related to cluster morphology or dynamical state. Here, we briefly describe the different estimators and refer the interested reader to our previous work for details (see also Wen & Han 2013).

We provide six dedicated measurements related to the morphological appearance of the galaxy cluster in the optical data. Additionally, the offsets between SZE and default optical centre as well as SZE and a centre derived from fitting a 2D model to the galaxy distribution are presented and can be used as measures of the cluster dynamical state. The 2D model centre is extracted while measuring the cluster morphology estimator δ (described in Wen & Han 2013; Klein et al. 2019, 2022). The estimator δ measures the normalized deviation from a smooth two-dimensional elliptical King model (King 1962) fitted on the smoothed galaxy density map of red-sequence cluster galaxies. Besides the normalized deviation from the model, we also provide the ellipticity ϵ of the fitted model as a measure of cluster morphology. The ridge flatness β, compares the concentration of fitted one-dimensional King profiles along different angular wedges and is the ratio of the lowest concentration value to the average concentration. Low-mass clusters falling into a massive cluster will cause the radial galaxy density profile to flatten towards the merger direction causing a low value of β. A third estimator is the asymmetry factor α (Wen & Han 2013), which measures the normalized average difference between pixel values in the galaxy density maps for pixels

lying across from each other with respect to the cluster centre. All four estimators are correlated with one another; they are based on the same galaxy density maps and associated noise, and they are sensitive to similar features – primarily the asymmetry (Klein et al. 2022).

The last set of two estimators are derived by running SEXTRACTOR (Bertin & Arnouts 1996) on the passive galaxy density map. We use the resulting source list to identify nearby structures close to the main cluster and list the distance in terms of R_{500} as well as the ratio of the flux_auto measurement of the main and the second structure. The flux ratio can be thought of as a richness ratio, and it therefore serves as a proxy of the mass ratio between the two structures in question. The combination of both estimators makes it possible to select merger candidates or cluster pairs that exhibit a certain mass ratio.

4 CLUSTER CATALOGUE

We present the new SPT-SZ MCMF cluster sample in Section 4.1 and then discuss the sample contamination (Section 4.2) and completeness (Section 4.3). In Section 4.4, we discuss the impact of DES masking on the survey solid angle. Finally, we present the results of the cluster morphological or dynamical state estimators in Section 4.5.

4.1 Defining the SPT-SZ MCMF galaxy cluster sample

As detailed in Section 3.1, the MCMF f_{cont} measurements for each candidate provide a means of defining cluster samples with the desired contamination level. For the catalogue we present here, we adopt an f_{cont} threshold $f_{cont}^{max} = 0.2$, which allows for 20 per cent of the original contamination present in the SPT-SZ candidate list to slip through into the confirmed cluster sample, which we call SPT-SZ MCMF. As we will show in detail in Section 4.2 we do have measurements of the amount of contamination of the original SPT-SZ candidate catalogue as a function of SPT-SZ detection signal-to-noise ξ. For $\xi > 4.0$, we measure an original contamination $f_{SZE-cont} = 45$ per cent. The final sample is therefore expected to have $0.2 \times 0.45 = 9$ per cent contamination and it contains 811 clusters. The f_{cont} selection threshold introduces incompleteness in the catalogue at the level of 5 per cent, because while the f_{cont} selection filters out contaminants it also removes some real, low-richness clusters. Details of this sample and other subsamples described below are shown in Table 1, which lists the f_{cont} threshold, SZE S/N threshold, purity, completeness, number of confirmed clusters and number of clusters at $z > 0.25$. Here, the listing of clusters above $z = 0.25$ is of special relevance for cosmological analyses, which have typically excluded lower redshift systems due to the angular filtering in the SPT cluster selection (see e.g. de Haan et al. 2016; Bocquet et al. 2019).

In selecting the best-suited cluster sample for a given science investigation, different sample criteria can be more or less important. To guide the reader, we provide here two additional subsamples of the SPT-SZ MCMF cluster catalogue by varying f_{cont}^{max} and ξ selections.

The first of the two subsamples ($\xi > 4.25$ in Table 1) has SZE selection thresholds $\xi > 4.25$ and $f_{cont}^{max} = 0.125$. Our current understanding is that cluster subsamples with $\xi > 4.25$ are better suited for studies relying on well-behaved SZE-based cluster masses (e.g. mass-observable scaling relations or cluster counts cosmology analyses). Furthermore, when modeling cluster counts to derive cosmological constraints, it simplifies the analysis if the subsample contamination is low enough that it does not require detailed modelling. A guideline here is that the contamination fraction is at or below the level of the Poisson noise associated with the full
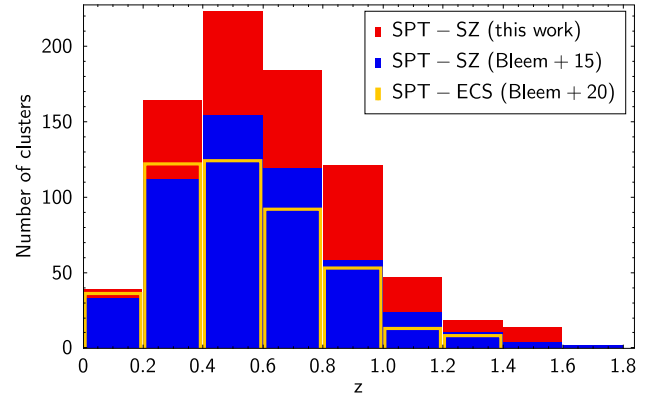
Figure 4. Redshift distribution of the new SPT-SZ MCMF cluster sample containing 811 clusters with 9 per cent contamination (red background) in comparison to the previous SPT-SZ catalogue (blue) and the SPTpol Extended Cluster Survey (SPT-ECS) catalogue in yellow.

subsample. Given the sample sizes here, the target upper limit for the contamination ranges between 3 per cent and 4 per cent, and this is met for both of the subsamples presented. A similar target can be set for the completeness of the sample, relative to its original SZE selection. As we will show in Section 4.3, the particular choice of f_{cont}^{max} used for this subsample meets requirements for both, purity and completeness. We also note that the incompleteness due to optical cleaning can be accounted for (see e.g. Grandis et al. 2020; Chiu et al. 2023).

The $\xi > 4.5$ subsample represents a more conservative selection in ξ and a looser cut in $f_{cont}^{max} = 0.13$, which remains in the well-tested ξ-regime. The contamination is low in the SZE candidate list, the predicted false detections from simulations and observations are in good agreement. The optical selection plays an insignificant role, introducing 3.5 per cent incompleteness through f_{cont} selection while providing a high (98 per cent) purity sample. The $\xi > 4.5$ subsample will further be part of the upcoming SPT cosmological analysis, which includes modelling of the impact optical cleaning on the sample selection.

The redshift distribution of the SPT-SZ MCMF cluster sample is shown in Fig. 4, where it is compared to the previously released SPT-SZ catalogue (Bleem et al. 2015) with updated redshifts from Bocquet et al. (2019) and the SPTpol Extended Cluster Survey catalogue (SPTpol-ECS Bleem et al. 2020). Of the 811 confirmed clusters in the SPT-SZ MCMF catalogue, 91 are at $z > 1$. This is a substantial increase compared to the 516 clusters in the previous catalogue, and it more than doubles the number of high-redshift clusters. The DES data cover only 92 per cent of the SPT-SZ sources, and consequently we miss 34 confirmed clusters from Bleem et al. (2015), and would expect ~69 clusters to lie outside the footprint. Adding the other published SPT-based SZE cluster catalogues, SPTpol-ECS (Bleem et al. 2020) and SPTpol 100d (Huang et al. 2020), and excluding duplicates, we obtain a combined catalogue containing 1720 clusters, exceeding the number of confirmed SZE clusters from the second Planck catalogue of 1334 Sunyaev–Zeldovich sources (Planck Collaboration XXVII 2016; Bahk & Hwang 2024), but lying below the number of S/N > 4 candidates presented by the ACT collaboration of 4195 (Hilton et al. 2021).

4.2 Initial contamination of SPT-SZ candidate lists

As discussed in Section 3.1, the expected contamination fraction of a sample selected using a particular f_{cont} threshold (i.e. clusters with
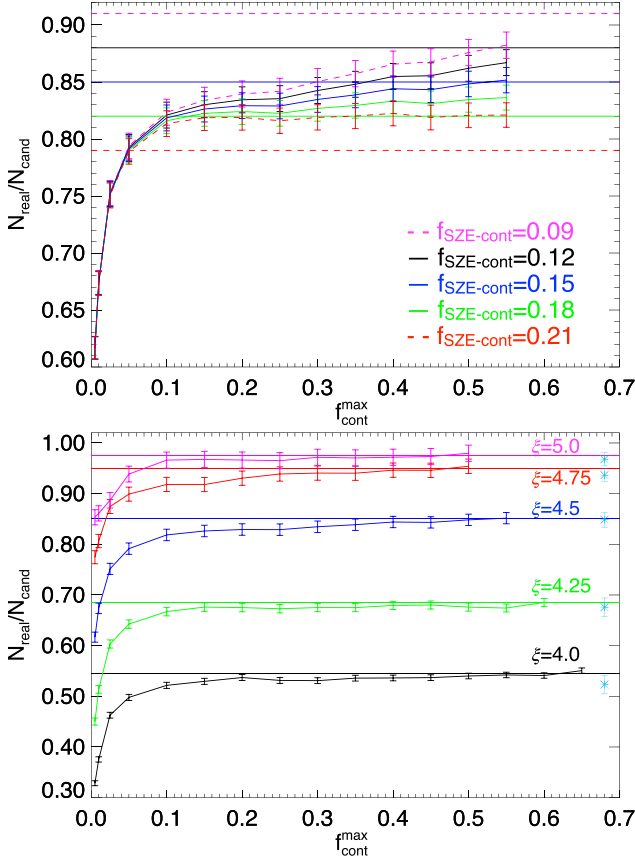
Figure 5. Empirical estimation of the initial contamination $f_{\text{SZE-cont}}$ in different SPT-SZ subsamples. Top: Example of the $f_{\text{cont}}^{\text{max}}$-based method (see discussion in Section 4.2) applied to the $\xi > 4.5$ SZE-selected sample with different assumptions for the initial contamination. The best-fitting initial contamination of 15 per cent is shown in blue. Bottom: Best-fitting results for five different SPT-SZ selection thresholds $\xi = 5.0, 4.75, 4.5, 4.25$, and 4.0 arranged from top to bottom that indicate an initial purity of 97.5 per cent, 95 per cent, 85 per cent, 69 per cent, and 55 per cent, respectively. For each case, the cyan point at $f_{\text{cont}} = 0.8$ shows an independent purity estimate from the mixture model method using the distribution of candidates in $\log_{10}(10^{14}\lambda/M_{500})$. Both methods are in good agreement with each other for all thresholds in ξ.

$f_{\text{cont}} < f_{\text{cont}}^{\text{max}}$) is $f_{\text{cont}}^{\text{max}} \times f_{\text{SZE-cont}}$. Therefore, it is crucial to know the contamination fraction in the initial candidate list. We use two methods to estimate $f_{\text{SZE-cont}}$ for the different SPT-SZ candidate lists (i.e. with different SZE selection thresholds in ξ). The first follows our previous work in (Hernández-Lang et al. 2023) and uses the fact that in $f_{\text{cont}} < f_{\text{cont}}^{\text{max}}$ selected samples, the completeness should reach 100 per cent for high values of $f_{\text{cont}}^{\text{max}} \sim 1$. The number of expected real clusters is

$$N_{\text{real}}(f_{\text{cont}}^{\text{max}}) = N_{\text{MCMF}}(f_{\text{cont}} < f_{\text{cont}}^{\text{max}}) \left[1 - f_{\text{cont}}^{\text{max}} f_{\text{SZE-cont}}\right], \qquad (3)$$

where $N_{\text{MCMF}}(f_{\text{cont}} < f_{\text{cont}}^{\text{max}})$ is the number of systems in the MCMF confirmed catalogue with f_{cont} values below $f_{\text{cont}}^{\text{max}}$. The ratio $N_{\text{real}}/N_{\text{cand}}$, where N_{cand} is the number of SPT-SZ candidates, should reach but not exceed the expected purity of the candidate list ($1 - f_{\text{SZE-cont}}$). Incorrectly estimating $f_{\text{SZE-cont}}$ would lead to inconsistencies, such as finding (1) more real clusters than allowed or (2) falling numbers of real clusters at high $f_{\text{cont}}^{\text{max}}$.

One illustrative example for an SPT-SZ subsample with $\xi > 4.5$ is shown in the top panel of Fig. 5. Here, we show the behaviour of

$N_{\text{real}}/N_{\text{cand}}$ for five different values of initial contamination fractions from 9 to 21 per cent in steps of 3 per cent. The horizontal lines show the expected purity ($1 - f_{\text{SZE-cont}}$) for the curves with the same colour. As can be seen, setting the initial contamination $f_{\text{SZE-cont}}$ too high (lowest two curves in red and green) causes an over prediction of real clusters (lines with data points) relative to that expected number given the assumed contamination level (flat line of same colour). This clear inconsistency excludes these high contamination levels.

For the lower assumed contamination cases with $f_{\text{SZE-cont}} \leq 0.12$, the curves with data points (black and magenta) continue rising over the full range of f_{cont}. This is a very unlikely scenario, given the expected richnesses of SPT clusters ($\lambda > 20$) and the richness levels probed at $f_{\text{cont}} > 0.6$ ($\lambda \approx 2$). For the initial contamination level of $f_{\text{SZE-cont}} = 0.15$ (black curve with data points), we find a stable solution where the fraction of real clusters converges to the expected contamination fraction and then remains roughly constant above $f_{\text{cont}} > 0.4$. This is a clear indication that this $\xi > 4.5$ SPT-SZ candidate list has $\approx$15 per cent contamination.

In the lower panel of Fig. 5, we show the results for five SPT-SZ candidate lists with different thresholds in signal-to-noise ξ of 5, 4.75, 4.5, 4.25, and 4. In these cases, we remeasure f_{cont} for each subsample using the appropriate signal-to-noise thresholds in the candidate and random sample. The upper limit of $f_{\text{cont}}^{\text{max}}$ considered for each signal-to-noise threshold is motivated by where the selection effectively reaches values of $\lambda \sim 5$ in richness. One can read off the purity of these samples to be 97.5 per cent (magenta), 95 per cent (red), 85 per cent (blue), 70 per cent (green), and 55 per cent (black), respectively. As expected, going to lower SPT-SZ signal-to-noise decreases the purity of the initial candidate lists. But as we have previously emphasized, the MCMF algorithm enables the removal of a large fraction of the contamination and the delivery of an overall larger cluster sample. This enlarged sample extends to lower masses at all redshifts and therefore typically also extends to higher redshift.

The second and main method to derive the level of initial contamination for different SPT-SZ candidate lists makes use of the richness distributions of the candidates and along random lines-of-sight and follows our previous work on X-ray selected clusters (Klein et al. 2022, 2023). Contrary to the first method it does not rely on f_{cont} selection or on the correct derivation of f_{cont}. To estimate the initial contamination, we model the distribution of candidates in $\log_{10}(10^{14}\lambda/M_{500})$ space as a mixture of a contamination and a cluster model. The contamination model is directly derived from the measurements along random lines of sight as a histogram in $\log_{10}(10^{14}\lambda/M_{500})$ that can be re-scaled to adopt for different amounts of contamination. The MCMF richness–mass relation typically follows a power law with a slope of approximately one and with lognormal intrinsic scatter of 15–30 per cent (Chiu et al. 2023; Bocquet et al. 2024). As the ratio $\log_{10}(10^{14}\lambda/M_{500})$ corresponds to a projection along the scaling relation, the expected distribution of clusters in $\log_{10}(10^{14}\lambda/M_{500})$ is expected to follow a normal distribution. The total model therefore consists of just four parameters, three for the normal distribution and just one, the normalization, for the contaminant distribution. The number of contaminants is then simply the integral over the best-fitting contaminant distribution. As example, the observed and the fitted model for the $\xi > 4$ candidate list is shown in Fig. 6. For the estimate of the contamination we are solely interested in the best-fitting contamination model, which is pre-dominantly determined at $\log_{10}(10^{14}\lambda/M_{500}) < 1$, where the cluster component plays no significant role.

The resulting purity estimates of this sample together with results for the other candidate lists with different ξ thresholds are shown as cyan points in the lower panel of Fig. 5. We obtain contamination
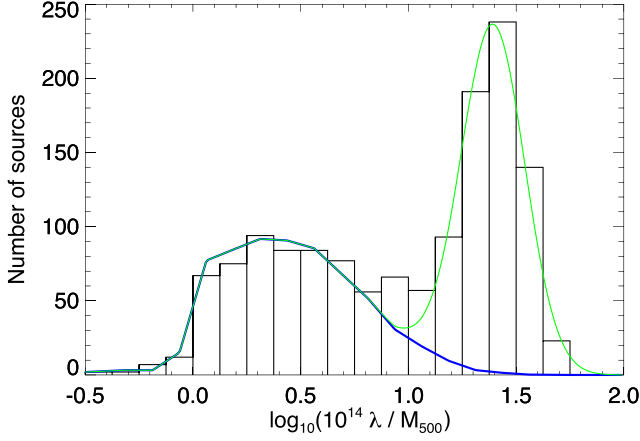
Figure 6. Example of the empirical estimation of the initial contamination $f_{SZE\text{-}cont}$ based on the distribution of candidates in $\log_{10}(10^{14}\lambda/M_{500})$ using a threshold of $\xi = 4$. The model (green) of the richness distribution of SPT-SZ candidates as a mixture of contaminants (in blue) and clusters. The composite model contains clusters modeled as a Gaussian distribution. The contaminant population is defined using measurements along random lines-of-sight.

values of 47.7 ± 1.7 per cent, 32.4 ± 1.8 per cent, 15.2 ± 1.5 per cent, 6.3 ± 1.2 per cent, and 3.2 ± 1.3 per cent for the ξ thresholds 4, 4.25, 4.5, 4.75, and 5. The results of both methods are consistent. These two methods can be used to derive the level of initial contamination in a candidate list even in the case where contaminants and clusters are not clearly separated in, e.g. a space of λ versus redshift.

4.3 SPT-SZ MCMF incompleteness due to optical cleaning

As already mentioned, the MCMF algorithm for excluding contamination can also exclude real clusters. The impact of the f_{cont}-based selection, which is essentially a redshift-dependent threshold in richness, can be modeled using the richness–mass relation (e.g. see Klein et al. 2022). In addition, a rough estimate of the overall completeness can be derived using the previously estimated initial contamination. The initial contamination defines the number of real clusters in the candidate list as well as the number of real clusters given the f_{cont} selection threshold. The differences between these two reflects the impact on the completeness of the real cluster sample. The impact of f_{cont}-based cleaning is already clearly visible in Fig. 5 (see bottom panel) as the difference between the horizontal lines that show the fraction of the candidates that are real clusters $(1 - f_{SZE\text{-}cont})$ and the curves with data points that show the recovered fraction of real clusters as a function of the f_{cont} threshold employed.

In Fig. 7, we show the expected completeness of the f_{cont} selected sample with respect to the number of real clusters in the SZE-selected sample versus the purity; this is shown for the same five SPT-SZ ξ thresholds examined previously in Fig. 5. Each curve is built by calculating the purity and completeness for a range of f_{cont} selection thresholds increasing from right to left. The purity is derived as $1 - f_{SZE-cont} \times f_{cont}^{max}$ and the completeness is the fraction of real clusters that survive the f_{cont} selection. As one can see, the impact of the optical cleaning on the completeness remains mild (≤ 5 per cent) for all SPT-SZ subsamples until one reaches a purity of 95 per cent or above, after which the completeness drops precipitously. Moreover, the highest purity samples tend to be smaller. One must consider these impacts when selecting a sample for scientific analysis.

As discussed in Section 4.1 the expected amount of sample Poisson noise (i.e. important for cluster count statistics) is simply $1/\sqrt{N_{clust}}$,
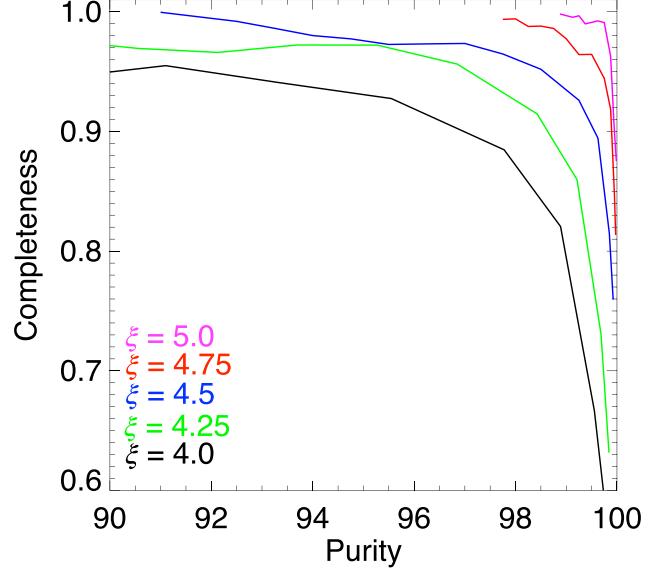

Figure 7. The purity as a function of completeness is shown for five different thresholds in the SPT-SZ selection threshold ξ. These curves are built for each sample by varying the MCMF defined optical selection f_{cont} threshold. Through tuning the f_{cont} threshold to lower values, one can create a final catalogue with an increased purity at the cost of introducing additional incompleteness.

and corresponds to 3.5–5 per cent for ξ thresholds of 4–5 in the SPT-SZ sample. This sets an upper limit for the target contamination such that contamination need not be explicitly modeled. While this translates into a completeness of 91 per cent for the $\xi > 4$ sample, it results in >96 per cent completeness for higher ξ threshold SPT-SZ subsamples, bringing the incompleteness below the level of Poisson noise of these subsamples. This means for subsamples with ξ thresholds of 4.25 or higher we are able to construct cluster samples where contamination and incompleteness are both below the expected Poisson noise, which implies that these effects will have a subdominant impact. Even in the contrary case, the sample incompleteness can be straightforwardly accounted for by modeling the selection in ξ jointly with the requirement that the richness λ exceeds the threshold corresponding to f_{cont}. A sample selected according to *both* variables ξ and λ is then complete with respect to a model that accounts for the joint selection. This modelling approach has already been successfully applied in the cosmological analysis of a real cluster sample (Chiu et al. 2023). However, currently, explicit modelling of contaminants in ICM-selected samples is still lacking. This explains the choice of a very clean (98 per cent pure) selection for the $\xi > 4.5$ subsample used in the new SPT-based cosmological study (Bocquet et al. 2024). This study explicitly includes modelling of the incompleteness due to MCMF-based cleaning but relies on high purity to avoid modelling contamination.

4.4 Impact of DES masking on SPT-SZ MCMF survey solid angle

The previous section covers the impact of optical cluster confirmation on the completeness and purity of the final cluster catalogue within the general DES footprint. One additional problem that arises with optical confirmation is that within the DES footprint there are areas with missing optical data due to, e.g. bright stars or a lack of data
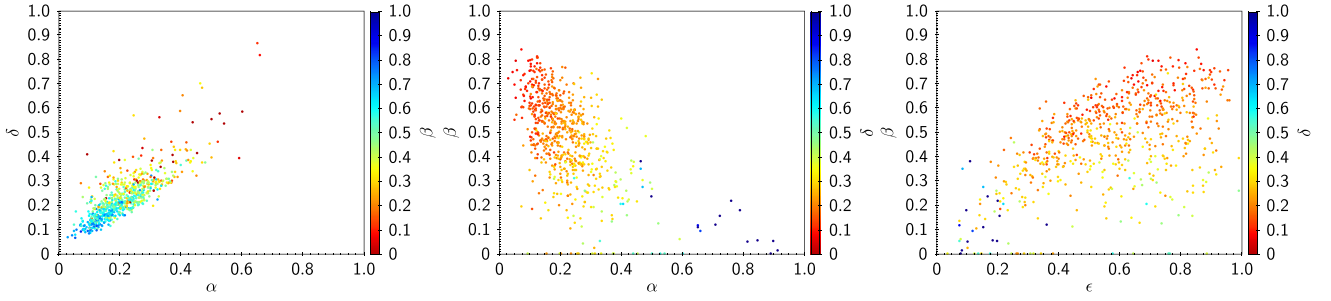
Figure 8. Comparison of different optical morphology estimators described in Section 4.5 for SPT-SZ MCMF clusters ($f_{\rm cont} < 0.2$) with $\lambda > 25$. Estimators probe different merging properties, but are well correlated.

due to poorly performing CCDs or even chip gaps. We follow these regions by building a sky mask for the optical data. The impact of missing data on cluster confirmation depends on the location and size of the masked region with respect of the SPT-SZ candidate position and effective size $\theta_{500}(z)$. Prior to confirmation the redshift and therefore the corresponding cluster size $\theta_{500}(z)$ is not well known; therefore, we use the sky masking fraction within a fixed angular distance around the candidate locations to characterize the importance of masking.

To estimate the impact of masking on cluster confirmation, we use the confirmation fraction as a function of the masking fraction. We explore three different apertures sizes with radii of one, two and 3 arcmin to test the sensitivity to masking. As a baseline we use clusters not showing masking and derive confirmation fractions $N(f_{\rm cont}<0.2)/N_{\rm cand}$ of 0.588 ± 0.014, 0.593 ± 0.014, and 0.595 ± 0.015 for the three apertures in discussion. Looking into the confirmation fractions we see that we would not have confirmed any candidate with mask fractions greater 0.56 in the two or three arcminute apertures and in total only one out of eight candidates with mask fractions greater 0.5. We therefore decided to re-define the minimum definition of a source to be considered within the DES footprint to have at least one source within 1 arcmin and a mask fraction within 2 arcmin below 0.5, effectively reducing the footprint by 0.6 per cent. There is no statistically significant impact visible on the confirmation fraction between mask fractions zero and 0.5. Taking all sources in that masking range, we find confirmation fractions of 0.52 ± 0.05, 0.52 ± 0.04, and 0.54 ± 0.03, consistent within 2σ from the baseline confirmation fractions. This residual effect can generally be accounted for by an overall re-scaling of the footprint area by 1.5 per cent. But we note that this correction is on the level of 1σ, given the uncertainty on the confirmation fraction of the unmasked clusters, this correction is likely not necessary for most studies using this sample.

4.5 Optical morphology and dynamical state

The morphology of a cluster can be an indicator of dynamical state, and so in principle the cluster morphology can be used to identify samples of clusters for the study of the dynamical evolution of clusters and cluster components. With the SPT-SZ MCMF catalogue, we include optical morphological estimators of dynamical state for all confirmed clusters; however, the quality of the measurements depends on richness and redshift. Increasing the richness selection threshold will further improve the robustness of the estimates. We therefore recommend restricting morphology analyses to the redshift range of $0.1 < z < 0.9$ and a richness $\lambda > 40$. In Fig. 8, we present comparisons among the four morphology estimators α, δ, β, and ϵ
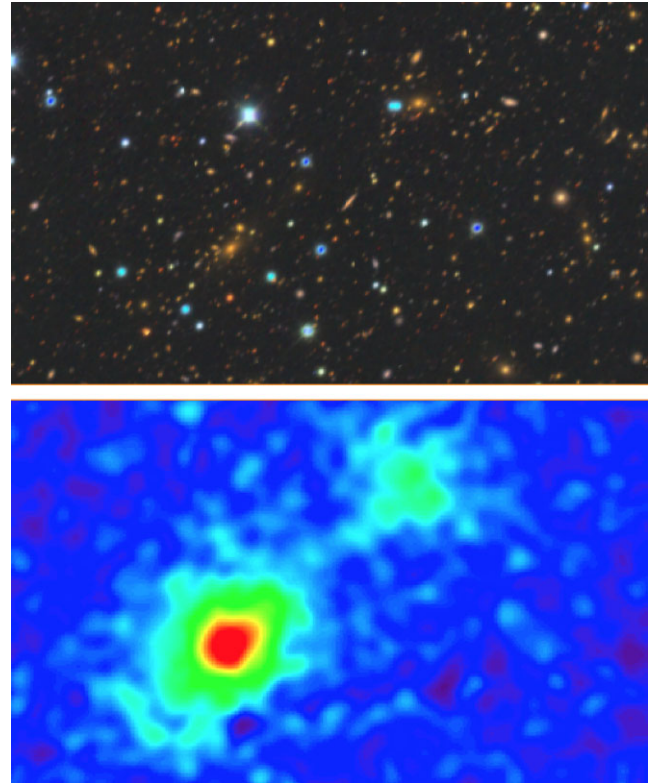
Figure 9. Top: DES RGB-image of 9 arcsec × 5.5 arcsec region around SPT-CL J0522–4818. Bottom: Smooth *Chandra* X-ray image of the same region. The cluster is classified as one of the most unrelaxed systems in optical while having a low X-ray-based disturbance estimate. The two clusters are likely in a pre-merger or early merger state, where the X-ray surface brightness distribution of the main system probed by the X-ray estimators is not yet affected by the merger process.

that are described in Section 3.3. As expected, the estimators are strongly correlated.

A preliminary comparison to X-ray morphological merger estimators that trigger on the skewness and ellipticity of the ICM distribution (e.g. Mohr, Fabricant & Geller 1993; Nurgaliev et al. 2013, 2017) for a subset of clusters that have *Chandra* or *XMM–Newton* observations shows little correlation, underscoring that optical and X-ray merger indicators are sensitive to different stages of cluster mergers and are also affected differently by projection effects. A simple example for such a case of mismatching classifications is SPT-CL J0522–4818, shown in Fig. 9. Therefore, we expect that these optical morphological estimators could be useful in combination with the

established X-ray techniques for the purpose of creating a sequence of clusters covering a broad range of dynamical state.

5 CATALOGUE VALIDATION

We validate SPT-SZ MCMF through comparison to several other catalogues in Section 5.1, carry out an examination of the contaminant distribution of the SPT-SZ candidate list in Section 5.2 and then carry out a modelling validation in Section 5.3 that employs parameter constraints from a cosmological analysis of the previous SPT-SZ sample.

5.1 Comparison of SPT-SZ MCMF to other cluster catalogues

We compare the new catalogue to three previously published SZE-selected cluster catalogues.

5.1.1 Previous SPT-SZ catalogue

To check for consistency, we compare our results to the previous release of the SPT-SZ catalogue (Bleem et al. 2015) and considering the updated redshifts provided in Bocquet et al. (2019). The expected contamination of the SPT-SZ candidate list at $\xi > 4.5$ adopted in the previous study is 15 per cent, and therefore the f_{cont} threshold 0.2 would correspond to an expected contamination in the final catalogue of 3 per cent. We find 481 clusters that have redshifts in both catalogues, and all but 4 have $f_{cont} < 0.2$. In all cases, the previously published redshift estimate is consistent with the redshift presented here.

The number of (4) unconfirmed systems, corresponding to ∼1 per cent of the previously confirmed $\xi > 4.5$ sample, is consistent with the expected incompleteness due to MCMF-based f_{cont} selection of 2 per cent given in Fig. 7. Furthermore, some of the previously confirmed clusters could indeed be chance superpositions. We fail to confirm SPT-CL J0334–4645, the highest redshift SPT cluster at $z = 1.7$ and one of the lowest redshift clusters SPT-CL J2313–4243 at $z = 0.056$. The latter is well identified in MCMF, but its richness is too low to meet the $f_{cont} < 0.2$ selection. In addition, we fail to confirm SPT-CL J0002–5557, which is listed to be at $z = 1.15$, whereas our high-z analysis places this cluster at $z = 1.37$ with an f_{cont} estimate of 0.45. The lack of red galaxies visible in the DES image indicates that this cluster needs to be beyond the MCMF DES redshift reach of $z \sim 1.3$. In *WISE*, the cluster is visible as one compact red blob, which may be the reason for the relatively high f_{cont} because counting cluster members for this compact cluster might have failed. The last cluster, SPT-CL J2005–5635, at $z = 0.2$ shows a low richness resulting in $f_{cont} = 0.31$. While the other three clusters do have matches in SZE or X-ray surveys, this cluster does not.

In Fig. 10, we show the redshifts derived from combining the MCMF outputs of the DES and the high-z runs, z_{comb}, with those published in Bocquet et al. (2019). As can be seen, there is good overall agreement for the majority of the 481 systems, but there are some outliers. The scatter between spectroscopic redshifts and MCMF redshifts is consistent with that found in our previous work using *ROSAT*-selected clusters (Klein et al. 2019). There are four prominent ($\Delta z > 5\sigma$) outlier clusters. In these four cases, we find two counterparts along the line of sight, where the second ranked one is consistent with the previously published redshift. In all four cases, the primary counterpart redshifts are coming from the DES-based run but are consistent with the counterpart from *WISE*-based MCMF run, making it unlikely that we are observing a new failure mode in
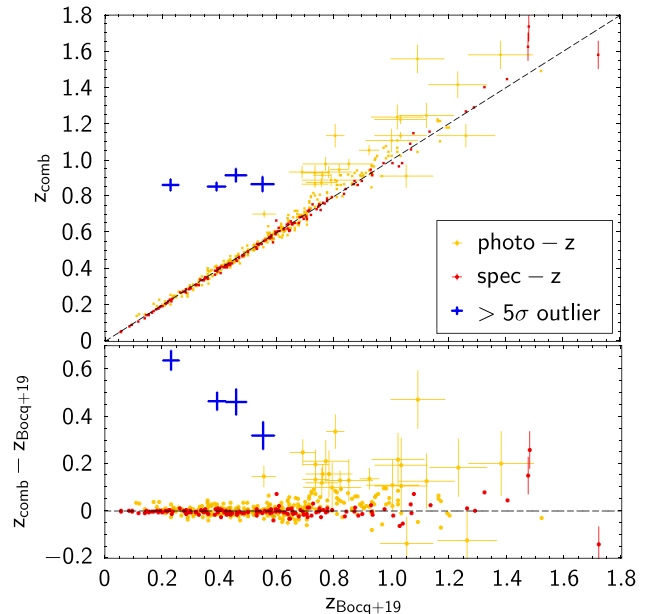
Figure 10. Comparison between redshifts given in this work and those presented Bocquet et al. (2019), split in photometric (yellow) and spectroscopic redshifts (red). The most significant outliers ($\Delta z > 5\sigma$) are shown in blue. All of those have a second ranked optical counterpart with redshifts in agreement with Bocquet et al. (2019). There is further an indication for a redshift bias in previous photo-z above $z = 0.7$. For sake of readability, we only show uncertainties for sources with $|\Delta z| > 0.1$.

one of the MCMF runs. One possible explanation for these outliers could be that the original SPT-SZ cluster by cluster follow-up may be composed of shallow observations that are sufficient to reliably detect the lower redshift counterpart but miss the higher redshift, more significant counterpart. There is further indication that there might be a mild under estimation of the redshifts given in Bocquet et al. (2019) for $z > 0.7$.

We conclude from the comparison to the previous version of the SPT-SZ cluster that there is consistency for ∼99 per cent of the overlapping sample. The number of previous systems not making our selection threshold is consistent with our estimate of incompleteness introduced by the optical cleaning, and the most prominent outliers in terms of redshift can be explained as multiple optical systems along the line of sight, where the current analysis finds a more significant richness peak than that selected in the original SPT-SZ follow-up.

5.1.2 SPTpol 100d catalogue

The comparison to the SPTpol 100 d catalogue (Huang et al. 2020) is especially interesting, because the deeper SPTpol data enable one to identify a larger number of purely SZE-selected clusters, which can then be compared to the MCMF defined catalogue from the fully overlapping but shallower SPT-SZ survey data. This 100d candidate list consists of 89 candidate clusters with a detection S/N $\xi > 4.6$. The analysis of image simulations suggests that 81 ± 2 of the candidates are real clusters, which is consistent with the number of optical-IR confirmed systems.

Using a matching radius of 150 arcsec, we find 37 matches between the 100 d and the SPT-SZ candidate catalogues, with the largest separation being 81 arcsec. Given the fact that contamination of the catalogues is mostly noise driven and the density of contaminants is estimated to be ≈0.08 deg^{-2} for SPTpol and 0.3–0.4 for SPT-SZ, it
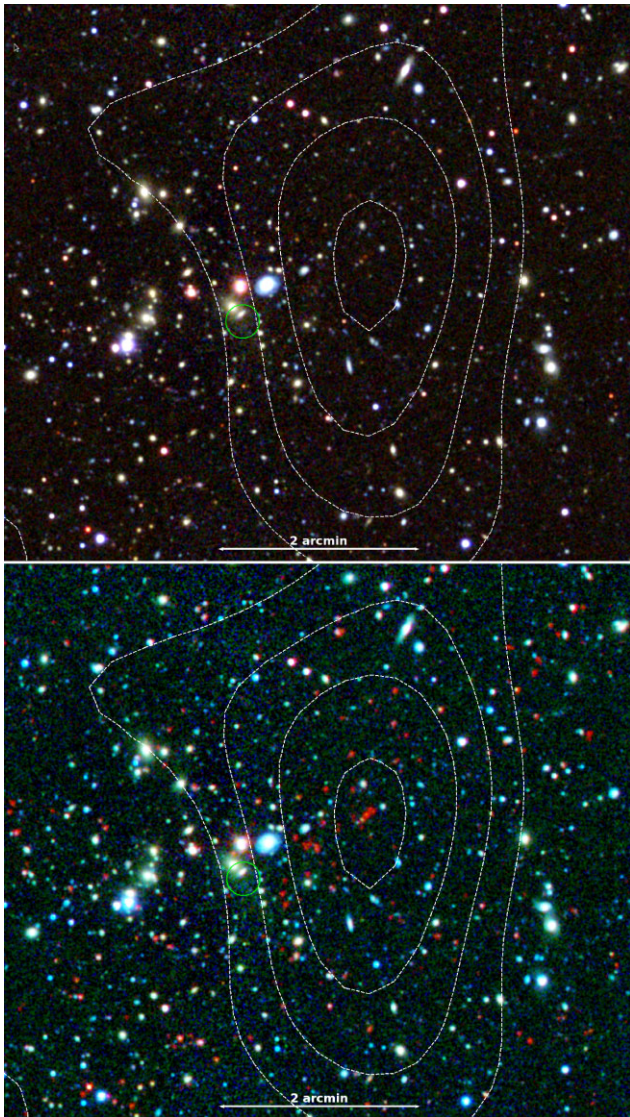
Figure 11. SPT-CL J2331–5736, the cluster with the highest redshift in SPTpol 100 d: The top image shows DES *g*, *r*, *z* colour composite image, and below is the DES *g*, *r*, and *Spitzer ch*1 colour composite image. The *Spitzer* image is taken from the SSDF (SPT *Spitzer* Deep Field). White contours show SPT-SZ S/N contours starting at 1 and increasing in steps of one. The green circle shows the location of a bright radio source detected in SUMSS. MCMF finds two counterparts, the high-z source visible only in the bottom image ($f_{cont} = 0.16$) and the low-z cluster close to the radio source ($f_{cont} = 0.005$).

is highly unlikely that we would find a chance match within this 150 arcsec search radius. Therefore, it is safe to assume that all matches correspond to real clusters.

Out of the 37 matches, we find 36 with $f_{cont} < 0.2$ that are members of the SPT-SZ MCMF cluster catalogue. The only cluster above that threshold is SPT-CLJ0002–5557, which was discussed in the previous section. Moreover, missing one cluster out of 37 matches is consistent with the expectation of 2 per cent incompleteness induced by the optical cleaning undertaken in building the SPT-SZ MCMF catalogue. Additionally, we find three clusters with $f_{cont} < 0.2$ that were not previously confirmed (Huang et al. 2020) and one cluster with a disagreement in redshift. We discuss those four systems below.

SPT-CL J2331–5736 (Fig. 11) has an S/N $\xi = 4.25$ in SPT-SZ and 8.4 in SPTpol 100 d and is the cluster with the highest redshift
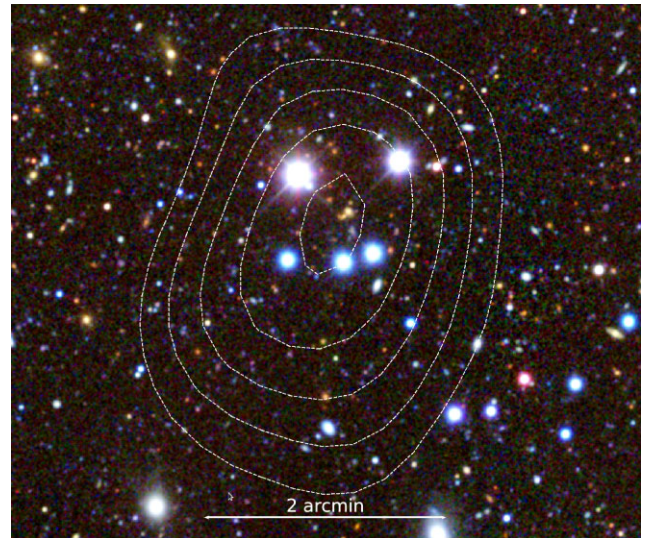
Figure 12. DES *g*, *r*, *z* colour composite image of SPT-CL J2321-5419. White contours show SPT-SZ S/N levels starting at 1 and increasing in steps of one. A bright star makes, it difficult to identify the $z = 0.79$ cluster members around the star north of the SZE peak.

in the SPTpol sample with $z = 1.38 \pm 0.1$. In Huang et al. (2020), it is noted that there is also a foreground cluster at $z = 0.29$. MCMF finds the low-z cluster to be at $z = 0.2975$ with a $f_{cont} = 0.005$ and the high-z structure at $z = 1.41$ and $f_{cont} = 0.16$. Given the f_{cont} values, both richness peaks are considered reasonable counterparts in the SPT-SZ MCMF cluster sample. The low value of f_{cont} makes it highly unlikely that the low-z structure is a chance superposition near a high-z cluster. The richness of $\lambda = 62$ is consistent with the expectation from the scaling relation. On the other hand, the tentative BCG of the high-z cluster is very close to the peak of the SZE signal. A closer investigation reveals a bright radio source with a SUMSS flux of 147.6 mJy (peak, 179.6 total) at the cluster centre of the low-z cluster, which could cause the SZE signal of this cluster to be partially diluted and its centre to be shifted.

SPT-CL J2321–5419 (Fig. 12) has an S/N $\xi = 5.26$ in SPT-SZ and 4.68 in SPTpol 100d. This cluster was not confirmed in the previous SPTpol and SPT-SZ catalogues, because of a bright star close to the SZE position. The MCMF analysis for this system indicates a redshift of 0.79 and a $f_{cont} = 0.07$. The high-z code finds a consistent redshift but does not confirm this system due to masking caused by the bright star.

SPT-CL J2357-5953 (Fig. 13) with S/N $\xi = 4.13$ in SPT-SZ and 4.66 in SPTpol is unconfirmed in SPTpol 100d, but the MCMF analysis identifies a cluster with redshift $z = 0.517$ and $f_{cont} = 0.02$. Additionally, the MCMF analysis identifies a second structure at $z = 1.11$ with $f_{cont} = 0.27$. The peak of the SZE signal is approximately in the middle of the two optical structures, which are separated from each other by 100 arcsec. The relatively large separation between the SZE and optical structure positions may have contributed to this system not being confirmed until now. The low probability of having two noise fluctuations in the two SZE surveys agree to within 29 arcsec makes it quite clear that the SZE detection itself is real. The large offset between optical and the SZE centre could be either a result of the low S/N of the detection, or it could be caused by the combination of the SZE signal from both clusters.

The last cluster is SPT-CL J0002–5214 with S/N $\xi = 4.48$ in SPT-SZ and 5.88 in SPTpol. This cluster is listed as a non-detection in
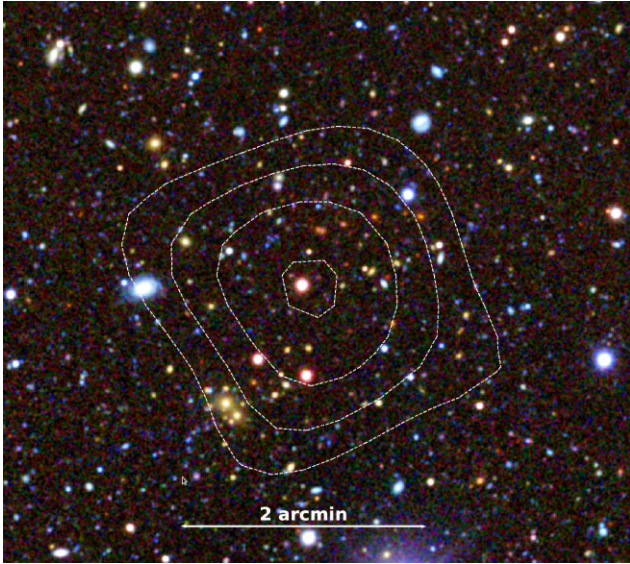
Figure 13. DES *g*, *r*, *z* colour composite image of SPT-SZ-CL J2357-5953, an SPT-SZ to SPTpol 100 d match that was not confirmed in SPTpol 100 d. White contours show SPT-SZ S/N levels starting at 1 and increasing in steps of one. There are two structures, one at $z = 0.517$ and another at $z = 1.11$ with corresponding f_{cont} values of 0.02 and 0.18, that are visible to the south-east and north-west of the SZE peak.

SPTpol, but there is a note that there is a potential group at $z = 0.44$. Noteworthy here is that according to simulations there should not be any noise fluctuations this large in the SPTpol 100d sample. The analysis with MCMF identifies two structures: one at $z = 0.41$ with $f_{cont} = 0.183$ and a second one at $z = 1.09$ with $f_{cont} = 0.198$. The high-redshift structure is also independently confirmed by the high-*z* code with a redshift of $z = 1.1$ and $f_{cont} = 0.009$. Visual inspection of Fig. 14 shows the rather compact group at intermediate redshift ($z \sim 0.4$), but the high-redshift structure is hard to identify by eye. In the DES *g*, *r*, *z* colour composite image, there is no clear cluster core, but there are a large number of high-redshift passive galaxies scattered over a region of 1.6 Mpc diameter. This becomes even clearer when using a combination of DES and *Spitzer* imaging data. We therefore conclude that this system is likely a high-redshift cluster with a low optical concentration.

In addition to checking for matched sources as above, we also check for SPT-SZ sources with low f_{cont} that do not appear in the SPTpol 100 d catalogue. Because SPTpol 100 d is substantially deeper, we do not expect many SPT-SZ confirmed clusters to be missed, but scatter in both S/N estimates and applied selection thresholds do allow for some number of missed systems. In fact, we find just one cluster in the overlapping footprints below $f_{cont} = 0.2$ that is not matched to a SPTpol 100d source. This source, SPT-SZ-CL J2342–5715 has an S/N $\xi = 4.33$ with $f_{cont} = 0.07$ and a redshift $z = 0.83$ (see Fig. 15). The DES optical image reveals a BCG that is only 33 arcsec away from the SZE peak, but the richness of the optical system $\lambda = 20.9$ is relatively low. Within a distance of 1.9 arcmin, we identify a low-*z* foreground structure harbouring a SUMMS source with a flux of ~ 60 mJy. Given the f_{cont} value, we can expect to have one contaminating source in the overlapping footprint. At the same time given the scatter in S/N in both surveys, the adopted thresholds in S/N and the low S/N of the particular system one could well find some clusters at $\xi > 4$ in SPT-SZ that are not detected in SPTpol 100 d. To summarize, we find only one SPT-SZ confirmed system that does not appear in the SPTpol 100 d catalogue, and given
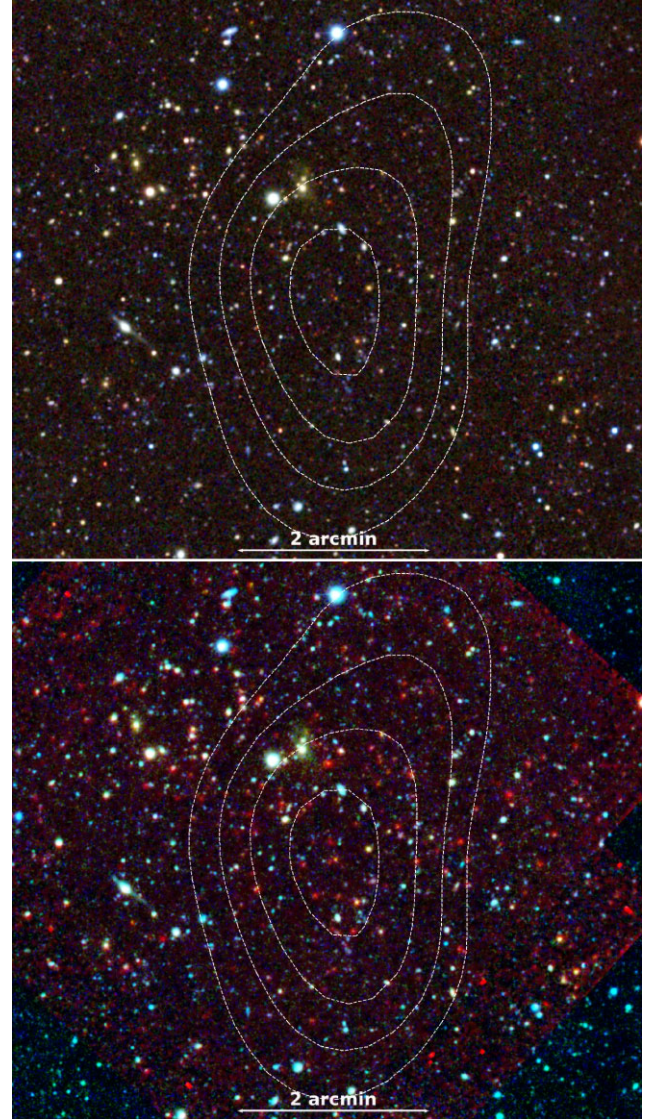

Figure 14. SPT-CL J0002-5214, an SPT-SZ match to SPTpol 100 d not confirmed in SPTpol 100 d: Top image shows DES *g*, *r*, *z* colour composite image and the bottom image shows the DES *g*, *r* and Spitzer *ch*1 colour composite image. White contours show SPT-SZ S/N contours starting at 1 and increasing in steps of one. There are two counterparts. One is at $z = 0.41$ with $f_{cont} = 0.183$ and another is at $z = 1.1$ with $f_{cont} = 0.009$.

the f_{cont} value this system could indeed be a chance superposition of an SPT-SZ noise fluctuation and an unassociated optical system.

5.1.3 ACT-DR5 cluster catalogue

The ACT-DR5 cluster catalogue (Hilton et al. 2021) is an SZE-selected cluster catalogue built using ACT survey data. ACT has similar properties to SPT. The catalogue contains 1843 clusters over the full DES footprint with ACT S/N ≥ 4. Allowing for offsets of up to 150 arcsec, we find 415 matches with our SPT-SZ MCMF catalogue, where the largest separation is 98 arcsec. Of those matches, 62 clusters have SPT-SZ S/N $\xi < 4.5$, and all of them show $f_{cont} < 0.1$, which indicates that these are very likely real clusters. Out of the full overlapping sample of 415, we find two clusters with $f_{cont} > 0.3$ and one with $0.2 < f_{cont} < 0.3$, all of them are known SPT-SZ clusters with $\xi > 5$ and would have been considered as confirmed,
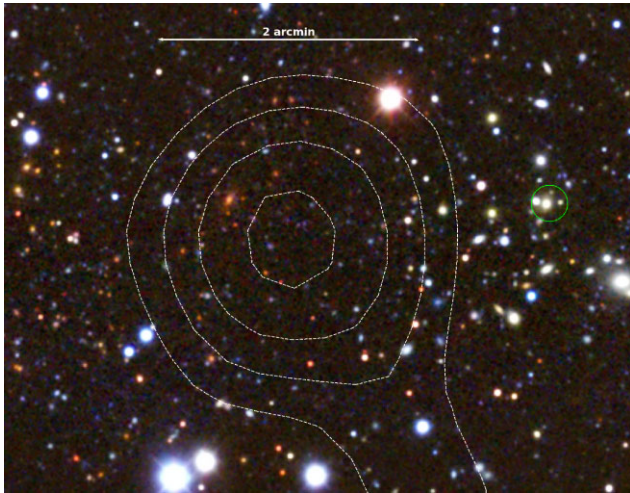
Figure 15. DES *g*, *r*, *z* colour composite image of SPT-SZ-CL J2342-5715. With $f_{cont} = 0.07$ and redshift $z = 0.83$, it is the only $f_{cont} < 0.2$ source that does not have a match in SPTpol 100 d within the overlapping footprint. White contours show SPT-SZ S/N levels starting at 1 and increasing in steps of one. The green circle shows the location of a bright radio source detected in SUMSS.

given the f_{cont} settings tuned for the $\xi > 5$ sample. We also find three systems with different redshift estimates. Two of them indicate two similarly good optical counterparts in the MCMF-based analysis where it is the MCMF second ranked system that agrees with the ACT-DR5 redshift. The remaining cluster SPT-CL J0619–5802, has only one clear MCMF counterpart at $z = 0.523$, in agreement with previous SPT-SZ work. The corresponding ACT cluster ACT-CL J0619.7-5802 is listed with a DES redMaPPer-based redshift of $z = 0.391$. Visual inspection supports the MCMF analysis with redshift $z = 0.523$.

5.2 Distribution of contaminants in SPT-SZ candidate list

We can use the MCMF algorithm to estimate the number of contaminants as a function of ξ in the initial SPT-SZ candidate list. Because the SZE is a distinct, negative signal in the 90 and 150 GHz SPT-SZ bandpasses, SZE-selected candidate catalogues contain contamination due to noise fluctuations. Because the noise is close to Gaussian, the number of false detections can be expected to follow a Gaussian noise field. The number of contaminants for the SPT–SZ catalogue were estimated previously by running the SZE-based cluster finder on source-free simulations (Bleem et al. 2015). The cumulative number of contaminants as a function of S/N ξ is shown in Fig. 16 together with the best-fitting model for Gaussian noise (red line and blue dashed line, respectively). The Gaussian model describes the number of contaminants for $\xi > 4$ with two free parameters: (1) the standard deviation of the noise and (2) a normalization parameter that is related to the ratio of the total survey solid angle to the effective solid angle of the filter functions used to detect clusters in the maps. This Gaussian model provides an excellent fit to the simulation results.

In the same figure, we show the measured number of contaminants extracted using the MCMF-based contamination analysis described in Section 4.2. Interestingly, the shape follows closely that expected from the Gaussian noise model and the image simulations, but the normalization is lower. In comparison to the Gaussian model fit to the image simulation results, the MCMF-based estimate can be
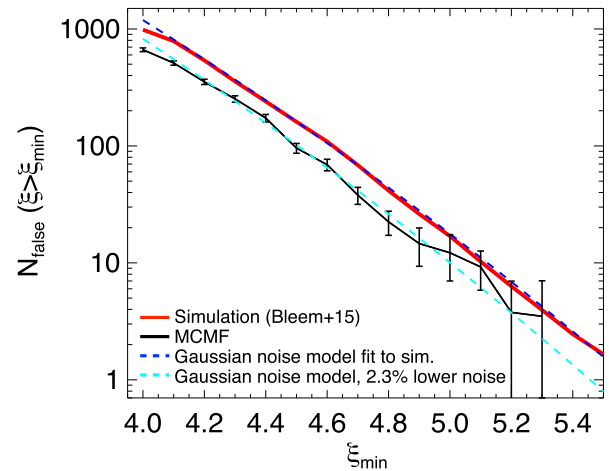
Figure 16. Cumulative number of contaminants in the SPT-SZ candidate catalogue as a function of the SZE signal to noise threshold ξ_{min} extracted from image simulations (red) and measured from the SPT-SZ catalogue using MCMF-based mixture model (black line with uncertainties). The simulation-based as well as the MCMF-based estimates can be well described by a Gaussian noise models (dashed blue and cyan lines). The higher number of contaminants in the simulations can be explained by a 2.3 per cent overestimate of the Gaussian noise used in the image simulations (see discussion in Section 4.2).

better matched if the standard deviation of the noise is reduced by 2.3 per cent. Thus, a mild overestimation of the noise in the SPT-SZ data could therefore lead to the overestimation of the contaminants apparent in Fig. 16. We note here that this difference becomes insignificant at $\xi > 5$, the threshold of the sample used in previous SPT-SZ cosmological studies, but is large compared to the Poisson uncertainties at $\xi \leq 4.7$.

In summary, the distribution of contaminants is consistent with Gaussian noise, as expected, and therefore extremely sensitive to the amplitude of that noise. There is an offset in the number of contaminants predicted by the image simulations and inferred through the MCMF-based analysis that can be explained by a 2.3 per cent change in the standard deviation of the noise. In the next section, we model the cluster counts and find evidence that points to an overestimate of the contamination in the SPT-SZ candidate list from the image simulations.

5.3 SPT-SZ MCMF validation using cluster counts

Given the new SPT-SZ MCMF cluster sample (Table 1) together with constraints on the residual contamination (Section 4.2) and incompleteness due to optical cleaning (Section 4.3), we can obtain the cluster number counts as a function of SPT-SZ S/N ξ threshold (ξ_{min}), and compare them with the prediction using the results from the cosmological analysis of the previous SPT-SZ sample with $\xi_{min} = 5$ (Bocquet et al. 2019). Here, of course, we are mainly interested in the behaviour of the new SPT-SZ MCMF clusters with S/N $\xi_{min} < 5$.

The expected number of clusters from MCMF-based mixture mode method, as well as from subtracting the simulation-based number of expected false detections from the full list of candidates is shown in Fig. 17 (left-hand panel) alongside the predicted number of clusters. Note that the uncertainties shown for the predicted cluster counts represent the Poisson noise only and do not include the error budget due to uncertainties on cosmology and scaling relation
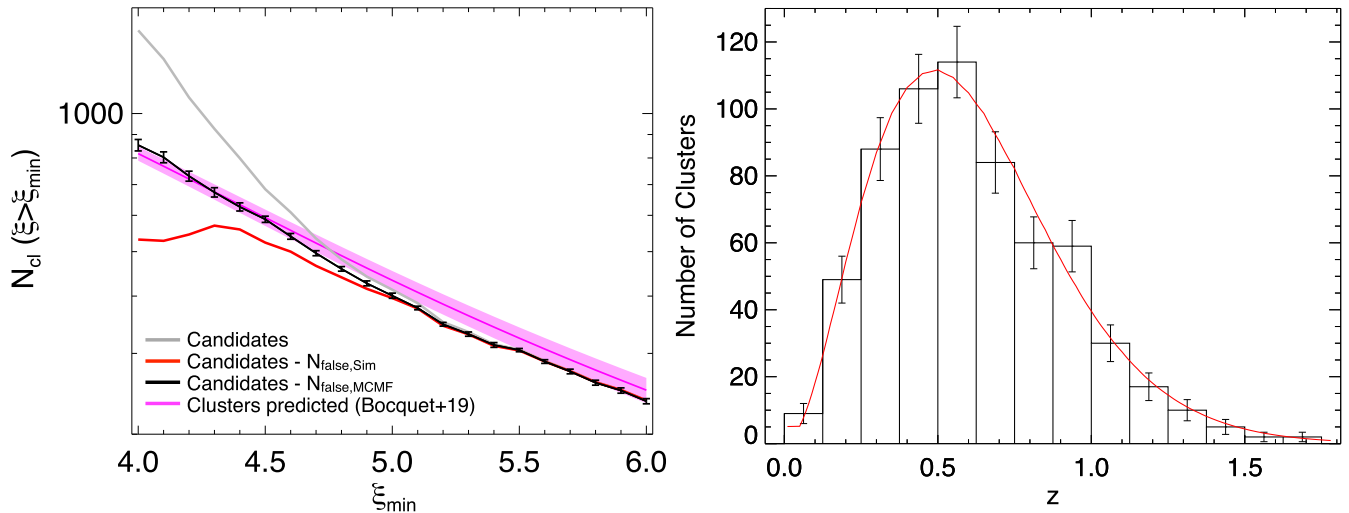
Figure 17. Left: Observed and predicted cluster counts above a given SZE selection threshold ξ_{min}. All candidates are shown in grey, candidates minus predicted contamination from image simulations in red, clusters expected from the mixture model method (see Figs 6 and 5) in black. The predicted number of clusters according to Bocquet et al. (2019) is shown in magenta (with 68 per cent confidence region only includes Poisson noise). Right: Redshift distribution of the $\xi >$ 4.25 subsample from Table 1 (black) and predicted redshift distribution according to Bocquet et al. (2019) (red). The predicted counts in ξ space are consistent with the observations, indicating that the sample is an extension of the previous $\xi > 5$ sample. The agreement of the shape of the redshift distribution with the prediction suggests that the incompleteness introduced by optical cleaning is not particularly pronounced at any redshift.

parameters. The uncertainties for the optical method predominantly depend on the number of contaminants in the sample and therefore becomes small at high ξ_{min}. As can be seen in Fig. 17, the predicted number of clusters shown in magenta agrees with the observed number using the optical method at the 1σ level and the behaviour at S/N $\xi_{min} < 5$ appears to be a meaningful extension to the high ξ_{min} regime. By contrast, the number of clusters expected from using the simulation-based appears to decrease at lower ξ_{min}, supporting the picture of a mild overestimation of the noise level in the simulations. The difference between simulation-based and optical-based estimates becomes insignificant at $\xi_{min} = 5$, which was used in previous SPT-SZ-based cosmological studies.

Using the results from Bocquet et al. (2019), we can further compare the observed and the expected redshift distributions, which we present in the right panel of Fig. 17. Here, we use the $\xi >$ 4.25 subsample, which – according to Fig. 7 – is 96 per cent pure and 96.5 per cent complete with respect to the initial SZE candidate selection. By design the f_{cont}-based selection aims to maintain a constant level of contamination as a function of redshift. Contamination therefore should not alter the shape of the redshift distribution of the sample. The red line in the right-hand panel of Fig. 17 shows the predicted redshift distribution using the results from Bocquet et al. (2019) normalized to same total number of clusters. The predicted and observed shapes of the redshift distributions agree remarkably well. Under the assumption that the contamination fraction is indeed constant over redshift this suggests that the incompleteness introduced by the $f_{cont} < 0.2$ selection is not significantly impacting the redshift distribution either.

6 CONCLUSIONS

In this paper, we present the SPT-SZ MCMF cluster catalogue with candidates selected to have SPT-SZ S/N $\xi > 4$ that are then confirmed using the MCMF algorithm. This sample represents a ≈ 50 per cent increase in size compared to the previous SPT-SZ catalogue and contains 811 clusters with 9 per cent contamination. Subsamples of

this new catalogue can be selected to have different characteristics (see Table 1). Considering an SPT-SZ S/N threshold $\xi > 4.25$ with stricter f_{cont} constraints in order to remove chance superpositions ($f_{cont} < 0.125$), we obtain 640 clusters with 96 per cent purity. This subsample has a modest 3.5 per cent incompleteness due to optical cleaning with the MCMF algorithm. This sample should meet the requirements for a cosmological analysis and corresponds to a factor of 2 increase compared to the previous SPT-SZ cluster catalogue used for cosmological analysis (Bocquet et al. 2019).

We use information derived from our MCMF-based analysis to infer the level of the initial contamination in the SZE-selected sample above several S/N thresholds as well as the purity and completeness after optical confirmation. This information can be used to select the combination of purity, sample size and completeness best suited for a given science study. Studies less impacted by contamination or that suffer from small number statistics may choose larger but more contaminated subsamples, while studies sensitive on contamination may use cleaner but smaller subsamples. The measured initial contamination, expressed in number of false detections above an S/N threshold, follows the shape expected for Gaussian noise. We find a systematic difference between our measurements and those predicted by simulations that could be explained if the noise assumed in the simulation was overestimated by a small amount (2.3 per cent). Comparing number of false detections with number of candidates, we find further evidence that the simulation-based estimates over predict the number of false detections as the number of real systems ($N_{cand} - N_{false}$) above an S/N threshold appears to decrease when lowering the threshold.

A validation test consisting of the comparison of S/N ξ and redshift z distributions of the new SPT-SZ MCMF sample to the predictions extrapolated from the previous cosmological analysis of the $\xi > 5$ subsample (Bocquet et al. 2019) shows good agreement. This gives us confidence that the new sample is well suited for an updated cosmological analysis that is carried out in combination with the DES weak lensing data set to constrain cluster masses (Bocquet et al. 2024). The subsample anticipated for that study is the more

Table 2. Column description of the SPT-SZ MCMF catalogue.

Column name	Description
SPT_ID	SPT-CL identifier
RA	Right ascension
DEC	Declination
XI	Highest detection significance in the filtered map (ξ)
THETA	Best core radius corresponding to the detection
FIELD	Name of sub-field
Z_BEST	MCMF photo-z of best counter part
Z_SPEC	Spectroscopic redshift
EZ_BEST	Photo-z uncertainty
F_CONT_BEST	f_{cont} of best counter part
LAMBDA_BEST	Richness of best counter part
ELAMBDA_BEST	Uncertainty on richness of best counter part
RA_MCMF	RA of MCMF centre
DEC_MCMF	Dec of MCMF centre
M_500_BEST	M_{500} marginalised over sys. Uncertainties as in Bocquet et al. (2019) [10^{14} M$_\odot$]
EM_500_BEST_UP	Upper uncertainty on M_500_BEST
EM_500_BEST_LO	Lower uncertainty on M_500_BEST
M500_BEST_NOSYS	M_{500} using fixed cosmology & scaling relations as in Bocquet et al. (2019) [10^{14} M$_\odot$]
EM500_BEST_NOSYS_UP	Upper uncertainty on M_500_BEST_nosys
EM500_BEST_NOSYS_LO	Lower uncertainty on M_500_BEST_nosys
RA_BCG	RA of MCMF based BCG candidate
DEC_BCG	Dec of MCMF based BCG candidate
DES_OR_HZ	0 if z_BEST is from DES MCMF else high-z (WISE) MCMF
MASKFRAC_60	Mask fraction within 1 arcmin
MASKFRAC_120	Mask fraction within 2 arcmin
MASKFRAC_180	Mask fraction within 3 arcmin
ALPHA	Dynamical state estimator: α
DELTA	Dynamical state estimator: δ
BETA	Dynamical state estimator: β
ELLIP	Dynamical state estimator: ϵ
MASSRAT	Dynamical state estimator: Sextractor based richness ratio between subclumps
GALOFF	Dynamical state estimator: Distance in r_500 between the two subclups identified by sextractor
CENTOFF_OSZ_MPC	Offset between MCMF and SZ cente in Mpc
Z_DES_1	Redshift first DES MCMF counterpart
Z_DES_2	Redshift second DES MCMF counterpart
LAMBDA_DES_1	Richness first DES MCMF counterpart
LAMBDA_DES_2	Richness second DES MCMF counterpart
F_CONT_DES_1	f_{cont} first DES MCMF counterpart
F_CONT_DES_2	f_{cont} second DES MCMF counterpart
F_CONT_BEST_XI_425	f_{cont} of best counterpart for $\xi > 4.25$ sample selection
F_CONT_BEST_XI_45	f_{cont} of best counterpart for $\xi > 4.5$ sample selection

conservative subsample that contains 480 clusters with SPT-SZ S/N $\xi > 4.5$ and $z > 0.25$ (see Table 1). Combining the SPT-SZ sample with SPT-ECS (Bleem et al. 2020), SPTpol 100d (Huang et al. 2020) and SPTpol 500d (Bleem et al. 2024), the total number of confirmed SPT-selected clusters now raises to 1720.

DATA AVAILABILITY

The catalogue will be made available as online supplement as well as on CDS. A brief description of the columns is provided Table 2.

REFERENCES

Abbott T. M. C. et al., 2018, ApJS, 239, 18
Aihara H. et al., 2018, PASJ, 70, S4
Bahk H., Hwang H. S., 2024, ApJS, 272, 7
Benson B. A. et al., 2013, ApJ, 763, 147
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Blanton M. R. et al., 2017, AJ, 154, 28
Bleem L. E. et al., 2015, ApJS, 216, 27
Bleem L. E. et al., 2020, ApJS, 247, 25
Bleem L. E. et al., 2024, Open J. Astrophys., 7, 13
Bocquet S. et al., 2015, ApJ, 799, 214
Bocquet S. et al., 2019, ApJ, 878, 55
Bocquet S. et al., 2023, preprint (arXiv:2310.12213)
Bocquet S. et al., 2024, preprint (arXiv:2401.02075)
Böhringer H. et al., 2000, ApJS, 129, 435
Brunner H. et al., 2022, A&A, 661, A1
Carlstrom J. E. et al., 2011, PASP, 123, 568
Chiu I. N., Klein M., Mohr J., Bocquet S., 2023, MNRAS, 522, 1601
Cutri R. M. et al., 2013, Explanatory Supplement to the AllWISE Data Release Products
de Haan T. et al., 2016, ApJ, 832, 95
de Jong J. T. A., Verdoes Kleijn G. A., Kuijken K. H., Valentijn E. A., 2013, Exp. Astron., 35, 25
Dey A. et al., 2019, AJ, 157, 168
Drlica-Wagner A. et al., 2018, ApJS, 235, 33
Finoguenov A. et al., 2020, A&A, 638, A114
Flaugher B. et al., 2015, AJ, 150, 150
Gladders M. D., Yee H. K. C., 2000, AJ, 120, 2148
Gonzalez A. H. et al., 2019, ApJS, 240, 33
Grandis S. et al., 2020, MNRAS, 498, 771
Hernández-Lang D. et al., 2023, MNRAS, 525, 24
Hilton M. et al., 2021, ApJS, 253, 3
Huang N. et al., 2020, AJ, 159, 110
King I., 1962, AJ, 67, 471
Klein M. et al., 2018, MNRAS, 474, 3324
Klein M. et al., 2019, MNRAS, 488, 739
Klein M. et al., 2022, A&A, 661, A4
Klein M., Hernández-Lang D., Mohr J. J., Bocquet S., Singh A., 2023, MNRAS, 526, 3757
Lahav O., Edge A. C., Fabian A. C., Putney A., 1989, MNRAS, 238, 881
Lang D., 2014, AJ, 147, 108
Mainzer A. et al., 2011, ApJ, 731, 53
Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, MNRAS, 406, 1759
Meisner A. M., Lang D., Schlafly E. F., Schlegel D. J., 2019, PASP, 131, 124504
Merloni A. et al., 2012, preprint (arXiv:1209.3114)
Mohr J. J., Fabricant D. G., Geller M. J., 1993, ApJ, 413, 492
Morganson E. et al., 2018, PASP, 130, 074501
Muzzin A. et al., 2009, ApJ, 698, 1934
Nurgaliev D., McDonald M., Benson B. A., Miller E. D., Stubbs C. W., Vikhlinin A., 2013, ApJ, 779, 112
Nurgaliev D. et al., 2017, ApJ, 841, 5
Planck Collaboration XXVII, 2016, A&A, 594, A27
Predehl P. et al., 2021, A&A, 647, A1
Raghunathan S. et al., 2022, ApJ, 926, 172
Reichardt C. L. et al., 2013, ApJ, 763, 127
Rykoff E. S. et al., 2014, ApJ, 785, 104
Schlafly E. F., Meisner A. M., Green G. M., 2019, ApJS, 240, 30
Sheldon E. S., 2014, MNRAS, 444, L25
Vanderlinde K. et al., 2010, ApJ, 722, 1180
Wen Z. L., Han J. L., 2013, MNRAS, 436, 275
Wright E. L. et al., 2010, AJ, 140, 1868

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

SPTSZ_MCMF_fPublic.fits

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

[1] *University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany*
[2] *Max Planck Institute for Extraterrestrial Physics, Giessenbachstr., D-85748 Garching, Germany*
[3] *Laboratório Interinstitucional de e-Astronomia–LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ-20921-400, Brazil*
[4] *Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA*
[5] *Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[6] *SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA*
[7] *Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[8] *School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*

[9]*Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA 02138, USA*

[10]*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK*

[11]*Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*

[12]*Department of Physics, University of Cincinnati, Cincinnati, OH 45221, USA*

[13]*Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA*

[14]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*

[15]*Fermi National Accelerator Laboratory, Batavia, IL 60510, USA*

[16]*HEP Division, Argonne National Laboratory, Argonne, IL 60439, USA*

[17]*Department of Physics and Astronomy, University of Missouri, 5110 Rockhill Road, Kansas City, MO 64110, USA*

[18]*Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK*

[19]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse 1, D-85748 Garching, Germany*

[20]*Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA*

[21]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*

[22]*Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth PO1 3FX, UK*

[23]*Department of Physics, University of Chicago, Chicago, IL 60637, USA*

[24]*Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA*

[25]*Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain*

[26]*Universidad de La Laguna, Dpto. Astrofisica, E-38206 La Laguna, Tenerife, Spain*

[27]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain*

[28]*Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK*

[29]*University of Nottingham, School of Physics and Astronomy, Nottingham NG7 2RD, UK*

[30]*Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy*

[31]*INAF–Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy*

[32]*IFPU–Institute for Fundamental Physics of the Universe, Via Beirut 2, I-34014 Trieste, Italy*

[33]*California Institute of Technology, Pasadena, CA 91125, USA*

[34]*Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, D-21029 Hamburg, Germany*

[35]*School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*

[36]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*

[37]*Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India*

[38]*High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, Japan*

[39]*Department of Physics, McGill University, Montreal, Quebec H3A 2T8, Canada*

[40]*Canadian Institute for Advanced Research, CIFAR Program in Cosmology and Gravity, Toronto, ON M5G 1Z8, Canada*

[41]*Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, NO-0315 Oslo, Norway*

[42]*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*

[43]*European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching, Germany*

[44]*Department of Astronomy, University of Florida, Gainesville, FL 32611, USA*

[45]*Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, A-6020 Innsbruck, Austria*

[46]*Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*

[47]*Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA*

[48]*Department of Astrophysical and Planetary Sciences and Department of Physics, University of Colorado, Boulder, CO 80309, USA*

[49]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*

[50]*Department of Physics, University of California, Berkeley, CA 94720, USA*

[51]*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*

[52]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*

[53]*University of Chicago, Chicago, IL 60637, USA*

[54]*Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*

[55]*Department of Physics, University of California, Davis, CA 95616, USA*

[56]*High Energy Physics Division, Argonne National Laboratory, Lemont, IL 60439, USA*

[57]*Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

[58]*Centre for Extragalactic Astronomy, Durham University, South Road, Durham DH1 3LE, UK*

[59]*Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*

[60]*Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson AZ 85721, USA*

[61]*George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA*

[62]*Centro de Investigaciones Energétches, Medioambientales y Tecnológicas (CIEMAT), E-28040 Madrid, Spain*

[63]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*

[64]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ-20921-400, Brazil*

[65]*Department of Physics, University of Minnesota, Minneapolis, MN 55455, USA*

[66]*Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton BN1 9QH, UK*

[67]*Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*

[68]*Physics Department, Center for Education and Research in Cosmology and Astrophysics, Case Western Reserve University, Cleveland, OH 44106, USA*

[69]*Department of Physics, Case Western Reserve University, Cleveland, OH 44106, USA*

[70]*Department of Physics, Brookhaven National Laboratory, Upton, NY 11973, USA*

[71]*Université Paris-Saclay, CNRS, Institut d'Astrophysique Spatiale, F-91405, Orsay, France*

[72]*INFN–National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy*

[73]*ICSC–Italian Research Center on High Performance Computing, Big Data and Quantum Computing, Italy*

[74]*Liberal Arts Department, School of the Art Institute of Chicago, Chicago, IL 60603, USA*

[75]*Department of Astronomy, University of Michigan, 1085 S. University Ave, Ann Arbor, MI 48109, USA*

[76]*School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*

[77]*Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*

[78]*Kavli Institute for Particle Astrophysics & Cosmology, P.O. Box 2450, Stanford University, Stanford, CA 94305, USA*

[79]*Institute of Geophysics and Planetary Physics, Lawrence Livermore National Laboratory, Livermore, CA 94551, USA*

[80]*INAF–Osservatorio Astronomico di Brera, Via Brera 28, I-20121, Milano, Italy & Via Bianchi 46, I-23807, Merate, Italy*

[81]*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*

[82]*Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*

[83]*Dunlap Institute for Astronomy & Astrophysics, University of Toronto, 50 St George St, Toronto, ON M5S 3H4, Canada*

[84]*Department of Astronomy & Astrophysics, University of Toronto, 50 St George Street, Toronto, ON M5S 3H4, Canada*

[85]*Astronomy Department, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA*

[86]*Department of Physics, University of Illinois Urbana-Champaign, 1110 W. Green Street, Urbana, IL 61801, USA*

[87]*Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA*

[88]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

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