

# Exploring the contamination of the DES-Y1 cluster sample with SPT-SZ selected clusters

S. Grandis,<sup>1,2\*</sup> J. J. Mohr,<sup>1,2,3</sup> M. Costanzi,<sup>4,5</sup> A. Saro,<sup>4,5,6,7</sup> S. Bocquet,<sup>1,2</sup> M. Klein,<sup>1,3</sup> M. Aguena,<sup>8,9</sup> S. Allam,<sup>10</sup> J. Annis,<sup>10</sup> B. Ansarinejad,<sup>11</sup> D. Bacon,<sup>12</sup> E. Bertin,<sup>13,14</sup> L. Bleem,<sup>15,16</sup> D. Brooks,<sup>17</sup> D. L. Burke,<sup>18,19</sup> A. Carnero Rosel,<sup>20,21</sup> M. Carrasco Kind,<sup>22,23</sup> J. Carretero,<sup>24</sup> F. J. Castander,<sup>25,26</sup> A. Choi,<sup>27</sup> L. N. da Costa,<sup>9,28</sup> J. De Vicente,<sup>29</sup> S. Desai,<sup>30</sup> H. T. Diehl,<sup>10</sup> J. P. Dietrich,<sup>1</sup> P. Doel,<sup>17</sup> T. F. Eifler,<sup>31,32</sup> S. Everett,<sup>33</sup> I. Ferrero,<sup>34</sup> B. Floyd,<sup>35</sup> P. Fosalba,<sup>25,26</sup> J. Frieman,<sup>10,16</sup> J. García-Bellido,<sup>36</sup> E. Gaztanaga,<sup>25,26</sup> D. Gruen,<sup>18,19,37</sup> R. A. Gruendl,<sup>22,23</sup> J. Gschwend,<sup>9,28</sup> N. Gupta,<sup>38</sup> G. Gutierrez,<sup>10</sup> S. R. Hinton,<sup>39</sup> D. L. Hollowood,<sup>33</sup> K. Honscheid,<sup>27,40</sup> D. J. James,<sup>41</sup> T. Jeltema,<sup>33</sup> K. Kuehn,<sup>42,43</sup> O. Lahav,<sup>17</sup> C. Lidman,<sup>44,45</sup> M. Lima,<sup>8,9</sup> M. A. G. Maia,<sup>9,28</sup> M. March,<sup>46</sup> J. L. Marshall,<sup>47</sup> P. Melchior,<sup>48</sup> F. Menanteau,<sup>22,23</sup> R. Miquel,<sup>24,49</sup> R. Morgan,<sup>50</sup> J. Myles,<sup>18,19,37</sup> R. Ogando,<sup>9,28</sup> A. Palmese,<sup>10,16</sup> F. Paz-Chinchón,<sup>23,51</sup> A. A. Plazas,<sup>48</sup> C. L. Reichardt,<sup>38</sup> A. K. Romer,<sup>52</sup> E. Sanchez,<sup>29</sup> V. Scarpine,<sup>10</sup> S. Serrano,<sup>25,26</sup> I. Sevilla-Noarbe,<sup>29</sup> P. Singh,<sup>4,5</sup> M. Smith,<sup>53</sup> E. Suchyta,<sup>54</sup> M. E. C. Swanson,<sup>23</sup> G. Tarle,<sup>55</sup> D. Thomas,<sup>56</sup> C. To,<sup>18,19,37</sup> J. Weller,<sup>3,57</sup> R. D. Wilkinson<sup>52</sup> and H. Wu<sup>58</sup>

*Affiliations are listed at the end of the paper*

Accepted 2021 March 19. Received 2021 March 17; in original form 2021 January 15

## ABSTRACT

We perform a cross validation of the cluster catalogue selected by the red-sequence Matched-filter Probabilistic Percolation algorithm (redMaPPer) in Dark Energy Survey year 1 (DES-Y1) data by matching it with the Sunyaev–Zel’dovich effect (SZE) selected cluster catalogue from the South Pole Telescope SPT-SZ survey. Of the 1005 redMaPPer selected clusters with measured richness  $\hat{\lambda} > 40$  in the joint footprint, 207 are confirmed by SPT-SZ. Using the mass information from the SZE signal, we calibrate the richness–mass relation using a Bayesian cluster population model. We find a mass trend  $\lambda \propto M^B$  consistent with a linear relation ( $B \sim 1$ ), no significant redshift evolution and an intrinsic scatter in richness of  $\sigma_\lambda = 0.22 \pm 0.06$ . By considering two error models, we explore the impact of projection effects on the richness–mass modelling, confirming that such effects are not detectable at the current level of systematic uncertainties. At low richness SPT-SZ confirms fewer redMaPPer clusters than expected. We interpret this richness dependent deficit in confirmed systems as due to the increased presence at low richness of low-mass objects not correctly accounted for by our richness-mass scatter model, which we call contaminants. At a richness  $\hat{\lambda} = 40$ , this population makes up  $> 12$  per cent (97.5 percentile) of the total population. Extrapolating this to a measured richness  $\hat{\lambda} = 20$  yields  $> 22$  per cent (97.5 percentile). With these contamination fractions, the predicted redMaPPer number counts in different plausible cosmologies are compatible with the measured abundance. The presence of such a population is also a plausible explanation for the different mass trends ( $B \sim 0.75$ ) obtained from mass calibration using purely optically selected clusters. The mean mass from stacked weak lensing (WL) measurements suggests that these low-mass contaminants are galaxy groups with masses  $\sim 3\text{--}5 \times 10^{13} M_\odot$  which are beyond the sensitivity of current SZE and X-ray surveys but a natural target for SPT-3G and *eROSITA*.

**Key words:** methods: statistical – galaxies: clusters: general – large-scale structure of Universe.

## 1 INTRODUCTION

Extraction of cosmological information from the number counts of galaxy clusters is critically sensitive to the contamination of the selected samples and to their completeness as a function of mass (Aguena & Lima 2018). Over the last decade, the method of direct

mass calibration has been established as an empirical approach to the modelling of completeness of cluster samples. By constraining the mean relation between selection observable and mass, and the scatter around this relation, the thresholds applied in selection observable can be transformed into completeness as a function of mass (see for instance Melin, Bartlett & Delabrouille 2005; Grandis et al. 2020). While systematic uncertainties on this mapping are still large, they can be faithfully traced and propagated on to the cosmological constraints via self-consistent and simultaneous analysis of the

\* E-mail: [s.grandis@lmu.de](mailto:s.grandis@lmu.de)

number counts and the mass calibration (e.g. Mantz et al. 2015; Bocquet et al. 2019; Abbott et al. 2020).

Concerning the purity of cluster samples, traditionally, the focus was on adjusting cluster selection in such a way as to limit contamination. The Sunyaev–Zel’dovich effect (SZE; Sunyaev & Zeldovich 1972), for instance, introduces a distinct spectral feature in the cosmic microwave background (CMB). Multifrequency matched filtering of CMB maps (Melin, Bartlett & Delabrouille 2006) therefore can provide pure cluster samples (Bleem et al. 2015, 2020; Planck Collaboration XXVII 2016; Hilton et al. 2018, 2021; Huang et al. 2020). In X-rays only extended sources are typically considered (Vikhlinin et al. 1998; Böhringer et al. 2001, 2004; Romer et al. 2001; Pacaud et al. 2006; Clerc et al. 2014), limiting the contamination by point-like non-cluster sources. In both of these methods, which rely on the intracluster medium (ICM) for detection, contamination is further controlled by optical confirmation, which is also required to determine the clusters redshifts (for the latest applications, see Klein et al. 2018, 2019; Bleem et al. 2020).

In the case of optical cluster selection via overdensities of red galaxies, no additional multiwavelength data are required to determine the selection observable and the redshift (see for instance Rykoff et al. 2014). At least in principle, every overdensity of red galaxies is expected to be associated with a halo of some mass (Cohn et al. 2007; Farahi et al. 2016). This is due to the fact that every galaxy lives in a halo. The presence of red galaxies thus guarantees the presence of at least one halo, the host halo of the brightest red galaxy. Assigning the most massive of the – possibly more than one – haloes to the optical structure theoretically ensures a one-to-one mapping between optically selected clusters and haloes. Several effects need to be accounted for, when modelling the mapping between halo mass and observed richness that results from this mapping. First, the colour filter for the red galaxies, which follows the red sequence calibrated on spectroscopic data, sweeps a large range of projected distances along the line of sight. This leads to significant projection effects from structures surrounding the cluster (Cohn et al. 2007; Song et al. 2012; Costanzi et al. 2019a). Furthermore, the optically selected cluster centre can be significantly displaced from the actual halo centre, leading to a lower measured richnesses. Other important effects are masking of cluster galaxies, and percolation effects (the merging or splitting of optically selected clusters, García & Rozo 2019).

Costanzi et al. (2019a) and Abbott et al. (2020) have recently calibrated the impact of these effects on richness estimates for clusters selected by the red-sequence Matched-filter Probabilistic Percolation algorithm (hereafter redMaPPer; Rykoff et al. 2014) in the Dark Energy Survey<sup>1</sup> (DES; Dark Energy Survey Collaboration 2016) year 1 data by combining properties extracted from real data with simulations. In this work, we further test this scatter model. We quantify the fraction of optically selected clusters for which the scatter model fails, calling them *contaminants*. Physically speaking, these objects are low-mass haloes which suffered projection, percolation, miscentring, or masking effects that are larger than expected from the simulations used to calibrate the scatter model. This type of contamination is in stark contrast with the more traditional use of the term ‘contaminant’ in SZE and X-ray cluster searches, where contaminants are random noise fluctuations or misclassified point sources. In the context of optical cluster finding, contaminants are low-mass haloes which are not described by the adopted richness–mass modelling.

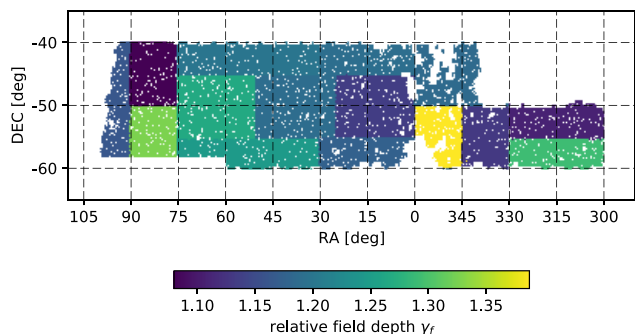
<sup>1</sup><https://www.darkenergysurvey.org>

Empirical constraints on the optical contamination by low-mass systems can be obtained by cross matching ICM selected clusters with optically selected clusters. The matched systems are likely higher mass clusters given their multiwavelength signature. In contrast, low-mass contaminants would be associated with shallower potential wells filled with less and cooler gas, resulting in weaker SZE signals and X-ray emission. In light of this, contamination of optically selected cluster samples can be studied by investigating the X-ray and SZE properties of these objects (Roza & Rykoff 2014; Roza et al. 2015; Saro et al. 2015, 2017; Farahi et al. 2019b). In this work, we shall focus on the validation of optically selected clusters using SZE information. This is motivated by the large overlap of the survey footprints of DES and the South Pole Telescope (SPT; Carlstrom et al. 2011) SZ survey (Story et al. 2013; Bleem et al. 2015). This study is further facilitated by the extensive cosmological and astrophysical work establishing that the empirically calibrated SZE masses derived from the cosmological studies (Bocquet et al. 2015, 2019; de Haan et al. 2016) are consistent with the weak lensing signal, the projected phase-space density of galaxy members, and the hot gas and stellar content of the SPT selected clusters (see for instance Saro et al. 2015; Hennig et al. 2017; Chiu et al. 2018; Bulbul et al. 2019; Capasso et al. 2019a, and references therein). Specifically, the posteriors on the SZ-signal–mass relation from the weak lensing calibrated cluster number counts by Bocquet et al. (2019), which we use in this work, provide reliable mass information with the appropriate systematic uncertainties for all SPT selected clusters.

In this work, we build on previous studies by Saro et al. (2015) and Bleem et al. (2020), which match SPT-SZ (and SPT-ECS) selected clusters (Bleem et al. 2015) to clusters selected by redMaPPer. We first present the employed cluster samples (Section 2) and the modelling framework (Section 3). To this end, we set up the likelihood for each SPT selected cluster to have a given measured richness conditional on the parameters of the richness–mass relation. We then use the SZE mass information to infer the most likely richness–mass relation parameters. Then, for each redMaPPer selected object, one can compute the probability that the object is confirmed by SPT-SZ. Contamination levels can be estimated by comparing these probabilities with the actual occurrence of matches (Section 4.1). Given considerations about the incompleteness introduced by the optical cleaning of the SPT-SZ candidates compared to the redMaPPer selection, these studies are limited to richnesses larger than 40. We then extrapolate the richness–mass relation we derive to lower richnesses, and investigate if its prediction of the redMaPPer number counts is consistent with previous measurements (Section 4.2). We also utilize the mean mass from stacked weak lensing measurements to estimate the mean mass of the contaminants. Finally, we discuss our results in comparison with the literature and outline several future prospects of the analysis method we present here (Section 5). Throughout this work, we adopt a flat Lambda cold dark matter ( $\Lambda$ CDM) cosmology with  $H_0 = 70.6 \text{ kms}^{-1} \text{ Mpc}^{-1}$  (Rigault et al. 2020),  $\Omega_M = 0.276$ , and  $\sigma_8 = 0.781$  (Bocquet et al. 2019), except where otherwise stated. Masses are computed within a radius at which the mass density is 500 times the critical density of the Universe at the redshifts of the clusters.

## 2 CLUSTER SAMPLES

In this work, we investigate the scatter model, contamination fraction, and mean scaling relation of the optically selected cluster sample based on the DES-Y1 data (Drlica-Wagner et al. 2018). These measurements are performed by cross matching and cross calibrating the optical sample with clusters selected in the SPT-SZ survey (Bleem



**Figure 1.** Joint SPT-SZ-DES-Y1 footprint with colours indicating the relative field depth of SPT-SZ observations as reported in de Haan et al. (2016). The relative field depth allows us to scale the ‘unbiased significance’–mass relation to fields of varying depth relative to the definition introduced by Reichardt et al. (2013).

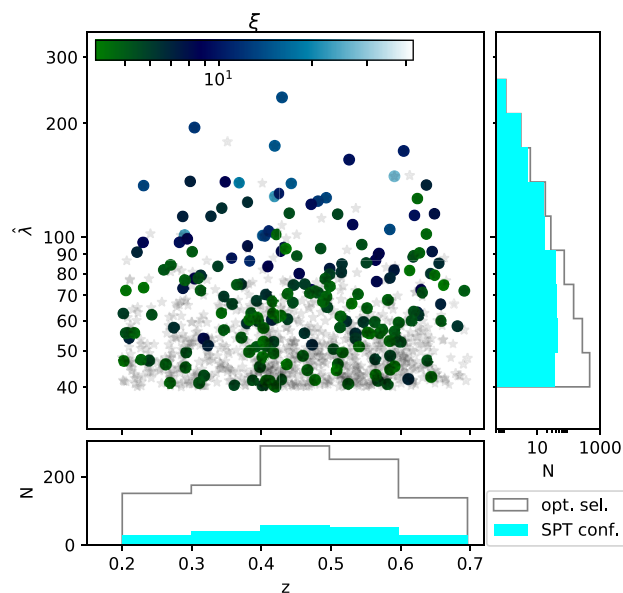
et al. 2015). This limits the analysis to the joint footprint of SPT-SZ and DES-Y1, which is shown in Fig. 1 with the relative SPT field depth colour coded. It comprises an area of 1463 deg<sup>2</sup>. In the following, we will touch on the main aspects of the two samples relevant to our analysis.

## 2.1 Optically selected samples

We employ the optically selected cluster sample extracted from DES-Y1 data (Drlica-Wagner et al. 2018) with the redMaPPer algorithm (Rykoff et al. 2014, 2016; Abbott et al. 2020), which was used for cosmological analyses (Abbott et al. 2020). This sample provides a measured cluster richness  $\hat{\lambda}^i$  with associated measurement uncertainty  $\delta\lambda^i$ , and photometric redshift  $z^i$  for each cluster  $i$ . The photometric redshifts display per cent scatter around spectroscopic redshifts (McClintock et al. 2019). Over the entire DES-Y1 footprint, when selected by  $\hat{\lambda} > 20$ , the redMaPPer sample comprises 7066 objects and spans the redshift range  $z \in (0.2, 0.7)$ . In our cross matching studies, we restrict ourselves to the joint DES-Y1 x SPT-SZ footprint and to  $\hat{\lambda} > 40$ . This sub-sample consists of 1005 objects, shown as grey points in Fig. 2. For objects in the latter sample, we extract the relative SPT field depth  $\gamma_f^i$  (de Haan et al. 2016).

## 2.2 SPT matched sample

We match the  $\hat{\lambda} > 40$  redMaPPer sample with the SPT-SZ sample selected above SZE signal to noise  $\xi > 4$  (Bleem et al. 2015). To reduce the contamination by noise fluctuations, we employ the SPT-SZ catalogue that was cleaned by the automated cluster confirmation and redshift measurement tool (Multicomponent Matched Filter, hereafter MCMF) by Klein et al. (in preparation) (see also Klein et al. 2018, 2019, for recent applications), using DES-Y3 photometric data. MCMF computes the richness of ICM selected cluster candidates using the ICM signature as a prior for the position and the aperture. The photometric data are filtered using spectroscopically calibrated red-sequence models. Comparison of the candidates richness to the richness distribution along random lines of sight allows one to select clusters based on the chance of random superposition  $f_{\text{cont}} < 0.1$ , which enables us to produce an uncontaminated SZE selected sample down to signal to noise  $\xi = 4$ . However, in the low SZE signal to noise regime the cut in the chance of random superposition can exclude clusters that would otherwise have passed the ICM selection, introducing optical incompleteness. We show in Appendix A that for objects with  $\hat{\lambda} > 40$  this incompleteness is always  $< 2.5$  per cent.



**Figure 2.** Central panel: Distribution in measured richness and redshift of the optically selected redMaPPer clusters (grey points) with measured richness  $\hat{\lambda} > 40$  in the joint SPT-SZ–DES-Y1 footprint together with their counterpart SPT-SZ selected clusters, colour coded by their SZE detection significance. Lower panel: redshift histogram of the optically selected redMaPPer clusters with measured richness  $\hat{\lambda} > 40$  in the joint SPT-SZ–DES-Y1 footprint (opt. sel., in grey) and the redshift histogram of the SPT-SZ confirmed sample (cyan). The fraction of SPT-SZ confirmed objects is constant with redshift. Right-hand panel: richness histogram of the optically selected redMaPPer clusters, and richness histogram of the SPT-SZ confirmed redMaPPer clusters (cyan). The fraction of SPT-SZ confirmed clusters decreases strongly towards lower richness. Quantitatively studying this fraction is the aim of this work.

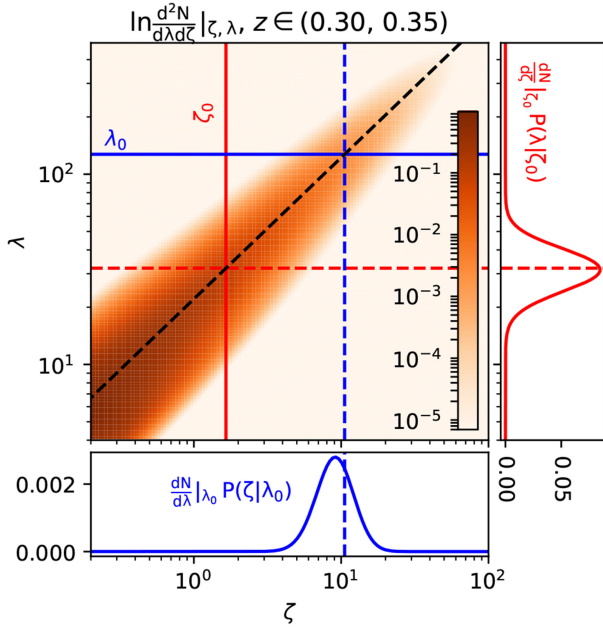
That means that any redMaPPer-object with a measured richness  $\hat{\lambda} > 40$  is (almost) certain to make it past the optical cleaning step in the SPT-SZ selection. Thus, for the redMaPPer- $(\hat{\lambda} > 40)$  sample, the probability of having an SPT-SZ counterpart is essentially given only by the SZE signal to noise.

We define counterparts in the two samples as those optical and SZE selected clusters lying within a projected radius of 1.0 Mpc and having consistent redshifts, i.e.  $|z_{\text{SPT}} - z_{\text{RM}}| < 0.05(z_{\text{SPT}} + z_{\text{RM}})/2$ . We match 207 objects, shown as coloured circles in Fig. 2, where the colour represents their SZE signal to noise. Two of the SPT clusters (SPT-CLJ0202-5401 and SPT-CLJ0143-4452) are matched to multiple (2) redMaPPer- $(\hat{\lambda} > 40)$  systems. We confirm that in both cases, both redMaPPer objects correspond to a significant detection (i.e. not a random superposition) by MCMF of an optical structure along the line of sight to the SPT cluster. The redshift and the MCMF-richness of these objects match the redMaPPer redshifts and richnesses. The SZE signal from these objects might have contributions from both clusters along the line of sight. Consequently, the SZE signature is likely biased high. However, they do not appear as outliers in the richness – SZ signal scatter plot. Given the rarity of these objects, we simply select the redMaPPer objects corresponding to the lowest  $f_{\text{cont}}$  MCMF peak (lowest probability of being a random superposition).

## 3 METHODS

In this section, we outline our modelling framework, presenting the hierarchical Bayesian cluster population model used in this work to





**Figure 3.** Graphic representation of the differential number density in space of intrinsic observables SZE signal  $\zeta$  and richness  $\lambda$  (central panel). In red is the result of conditioning that distribution on a given value of SZE observable  $\zeta_0$  (solid red line), with the resulting conditional probability distribution in richness (right inset). The shape of this distribution represents the richnesses consistent with  $\zeta_0$ , while the amplitude is the differential number of objects at  $\zeta_0$ . The dashed line marks the richness obtained by inverting the richness–SZE signal relation (intersection of solid red line and dashed black line). Note that this richness is larger than the mean expected richness. This effect is caused by Eddington bias, and results from there being more low-mass than high-mass systems (and thus more systems scattering up than down). The same argument holds when conditioning on a richness  $\lambda_0$  in blue and in the lower inset, and explains the offset between the blue curve and blue dashed line in the lower inset when conditioning on richness  $\lambda_0$ .

constrain the richness–mass modelling and the failure fraction of that model from data. We also discuss how the number counts and mean WL masses are used to compare our results to observations.

### 3.1 General cluster population model

The cluster population model adopted in this work follows closely the model presented in Grandis et al. (2020), which builds on work by Bocquet et al. (2015). Given the abundance of clusters  $\frac{dN}{dM}|_{M,z}$  as a function of mass  $M$  and redshift  $z$ , we express the abundance of clusters in the joint space of intrinsic SZE signal to noise  $\zeta$  and richness  $\lambda$  as

$$\frac{d^2N}{d\zeta d\lambda}|_{z,\gamma_f} = \int dM P(\zeta, \lambda|M, z, \gamma_f) \frac{dN}{dM}|_{M,z}, \quad (1)$$

where  $P(\zeta, \lambda|M, z, \gamma_f)$  describes the distribution of SZE signal and mean richness for a given mass and redshift. Because the SZE signature is a detection significance, the expression includes  $\gamma_f$ , which is the normalized field depth at the cluster position. Furthermore, here we do not yet include measurement noise on the observables. Thus,  $(\zeta, \lambda)$  are *intrinsic* observables, as opposed their *measured* counterparts  $(\xi, \hat{\lambda})$ , discussed in Section 3.2.

A graphic representation of the abundance of objects as a function of the two intrinsic observables (with  $\gamma_f = 1$ ) is shown in the central panel of Fig. 3. For these mean relations, we adopt the same power-

law behaviour as outlined in Benson et al. (2013), reading

$$\langle \ln \zeta \rangle = \ln \gamma_f A_{\text{SZE}} + B_{\text{SZE}} \ln \left( \frac{M}{M_0} \right) + C_{\text{SZE}} \ln \left( \frac{E(z)}{E(z_0)} \right), \quad (2)$$

for the intrinsic SZE signal, with  $M_0 = 3.0 \times 10^{14} M_\odot h^{-1}$  and  $z_0 = 0.6$ . We define an analogous relation for the mean intrinsic richness, following Saro et al. (2015)

$$\langle \ln \lambda \rangle = \ln A_\lambda + B_\lambda \ln \left( \frac{M}{M_0} \right) + C_\lambda \ln \left( \frac{E(z)}{E(z_0)} \right). \quad (3)$$

The scatter in SZE signal is modelled as a lognormal distribution with dispersion  $\sigma_{\text{SZE}}$ , while the scatter in richness has both a lognormal component together with a Poisson contribution  $\sigma_{\lambda,\text{tot}}^2 = \sigma_\lambda^2 + (\lambda - 1)/\lambda^2$ . We thus have four free parameters for each relation: an amplitude  $A_{\text{SZE}/\lambda}$ , a mass slope  $B_{\text{SZE}/\lambda}$ , a redshift evolution  $C_{\text{SZE}/\lambda}$ , and an intrinsic scatter  $\sigma_{\text{SZE}/\lambda}$ . We also introduce the correlation coefficient  $\rho$  between the intrinsic scatter in SZE signal and the intrinsic scatter in richness as a free parameter of our analysis.

### 3.2 Observational errors

We also account for the observational uncertainties affecting the measured SZE signal and the richness. For the SZE signal, the measured signal  $\xi$  follows the distribution established by Vanderlinde et al. (2010), which reads

$$P(\xi|\zeta, \gamma_f) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} (\xi - \sqrt{\zeta^2 + 3})^2 \right\}. \quad (4)$$

For the richness, we follow two prescriptions. The first follows the method used in Saro et al. (2015). Together with the measured richness  $\hat{\lambda}^i$ , the redMaPPer cluster catalogue provides an estimate on the error of the richness  $\delta\lambda^i$  for each entry  $i$ , which is interpreted as a Gaussian standard deviation, yielding

$$P_{\text{bkg}}(\hat{\lambda}^i|\lambda, \delta\lambda^i) = \frac{1}{\sqrt{2\pi}(\delta\lambda^i)^2} \exp \left\{ -\frac{1}{2} \left( \frac{\hat{\lambda}^i - \lambda}{\delta\lambda^i} \right)^2 \right\}. \quad (5)$$

For applications where the average measurement uncertainty as a function of arbitrary measured richness  $\hat{\lambda}$  is required, we adopt the extrapolation scheme presented in Grandis et al. (2020), Appendix A, to estimate  $\delta\lambda(\hat{\lambda}, z)$  directly from the catalogue. This model accounts only for the photometric uncertainties in the background subtraction. We call this model ‘background’ (‘bkg’).

A detailed study of projection effects on simulations by Costanzi et al. (2019a), expanded the prescription above to provide an accurate description of the impact of correlated structures, masking, and percolation on the mapping between intrinsic and measured richness. This effect is summarized by the fitted probability density function  $P_{\text{proj}}(\hat{\lambda}|\lambda, z)$ . For the exact definition of this function see equation (15) in Costanzi et al. (2019a). This model is called ‘projection’ (‘proj’).

All analysis steps that follow are performed for both models in an attempt to grasp the impact projection effects might have on our inference.

### 3.3 SPT cross calibration and priors

All objects of the matched redMaPPer–SPT cross-matched sample have a redshift  $z^i$ , an observed SZE signal  $\xi^i$ , and a measured richness  $\hat{\lambda}^i$ . For each set of scaling relation parameters, we then use the SZE signal  $\xi^i$  to predict the expected distribution of intrinsic richnesses  $\lambda$ , by convolving the joint distribution of intrinsic observables with

the measurement uncertainty of the SZE signal, i.e.

$$P(\lambda|\xi^i, \gamma_f^i, z^i) \propto \int d\zeta P(\xi^i|\zeta) \frac{d^2N}{d\zeta d\lambda} \Big|_{z^i, \gamma_f^i}. \quad (6)$$

This equation depends on the scaling relation parameters through the last factor from equation (1). As can be seen in Fig. 3, for each intrinsic SZE  $\xi^i$  this expression defines a range of intrinsic richness at a particular redshift. Given the typical measured mass-observable scaling relation parameters, higher SZE  $\xi^i$  corresponds to higher  $\lambda$ , and the scatter about both underlying mass-observable relations (equations 2 and 3) together with the covariance in this scatter leads to the width in the richness distribution for a given SZE  $\xi^i$ .

To account for observational uncertainties on richness, equation (6) can be convolved with the optical error model

$$P(\hat{\lambda}|\xi^i, \gamma_f^i, z^i) \propto \int d\lambda P(\hat{\lambda}|\lambda, z) P(\lambda|\xi^i, \gamma_f^i, z^i). \quad (7)$$

The proportionality constant is determined by ensuring that the equation above is properly normalized for all possible measured richnesses:  $\int_{\lambda_0}^{\infty} d\hat{\lambda} P(\hat{\lambda}|\xi^i, \gamma_f^i, z^i) = 1$ . Note that this normalization cancels any possible dependence on the absolute number of objects, strongly reducing the cosmological dependence in this analysis.

Evaluation of the properly normalized equation (7) at the measured richness  $\hat{\lambda}^i$  gives the likelihood of the measured observables for each cluster in the sample

$$\ln \mathcal{L}^i = \ln P(\hat{\lambda}^i|\xi^i, \gamma_f^i, z^i) \quad (8)$$

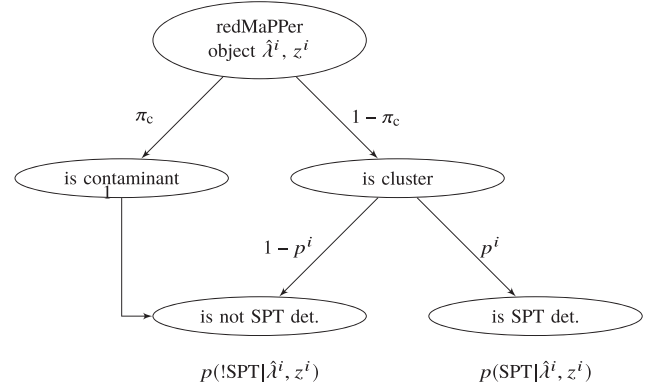
The total log-likelihood then results from summing the log-likelihoods of the individual clusters.

We proceed with this model by first constraining the richness-mass relation parameters by adopting priors on the SZE-mass scaling relations from a recent, weak lensing informed analysis (Bocquet et al. 2019). In that work, constraints on the SZE signal-mass scaling relation parameters are derived by jointly fitting the number counts of cluster sample together with mass calibration information derived from pointed weak lensing follow up measurements. Effectively, we adopt the recent SPT-SZ analysis results (with cosmological constraints in good agreement with other probes) and ask: given the adopted form of the richness-mass relation (equation 3), what parameters are required for consistent SZE signals and richnesses of the cross-matched clusters?

Specifically, the priors on the SZE scaling relation parameters in the baseline  $\nu$ - $\Lambda$ CDM model read  $A_{\text{SZE}} = 5.24 \pm 0.85$ ,  $B_{\text{SZE}} = 1.534 \pm 0.100$ ,  $C_{\text{SZE}} = 0.465 \pm 0.407$ , and  $\sigma_{\text{SZE}} = 0.161 \pm 0.080$  [symmetrized versions of the constraints reported by Bocquet et al. (2019), Table 3, 2nd column ‘ $\nu$ - $\Lambda$ CDM, SPTcl’]. These priors encode the systematic mass uncertainty on the SZE derived masses. We like to stress here that these masses do not assume hydrostatic equilibrium, but are instead empirically calibrated using number counts and weak lensing information. When predicting the redMaPPer number counts we also use  $\Omega_M = 0.276 \pm 0.047$  and  $S_8 = \sigma_8(\Omega_M/0.3)^{0.2} = 0.766 \pm 0.025$  (Bocquet et al. 2019) to properly account for uncertainties in these cosmological parameters. All these priors are modelled as Gaussian distributions in the likelihood inference.

### 3.4 Constraining contamination

The analysis outlined in the previous section can only be performed on a matched SPT-redMaPPer sample, as for any cluster both a measurement of the richness and the SZE signal are required. Considering instead the entire redMaPPer sample above  $\hat{\lambda} > 40$ , we



**Figure 4.** Probability tree describing the different possibilities for a redMaPPer object with measured richness  $\hat{\lambda}^i$  and redshift  $z^i$  to be detected by SPT  $p(\text{SPT}|\hat{\lambda}^i, z^i)$  and not being detected  $p(!\text{SPT}|\hat{\lambda}^i, z^i)$ . These probabilities depend not only on the raw detection probabilities  $p^i$  as obtained from the observable-mass relations and the selection function, but also on the contamination fraction  $\pi_c$ , which models the fraction of objects for which our richness-mass scatter model fails.

can view being matched or not being matched by SPT as a boolean measurement. We can also seek to predict the outcome of this event for each single redMaPPer cluster  $i$  based on the observed richness  $\hat{\lambda}^i$ , the redshift  $z^i$ , and the values of the scaling relation parameters. Indeed, given a  $\hat{\lambda}^i$ , we can predict the probability for a redMaPPer cluster to have a given intrinsic SZE signal by computing

$$P(\zeta|\hat{\lambda}^i, z^i, \gamma_f^i) \propto \int d\lambda P(\hat{\lambda}^i|\lambda, z^i) \frac{d^2N}{d\zeta d\lambda} \Big|_{z^i, \gamma_f^i}, \quad (9)$$

where in this case the proportionality constant is set by  $\int d\zeta P(\zeta|\hat{\lambda}^i, z^i, \gamma_f^i) = 1$ , as no selection of SZE properties was performed. For a graphical representation of this equation, see Fig. 3, where we highlight how conditioning on a given richness ( $\lambda_0$  denoted as blue horizontal line) selects a range of intrinsic SZE signal to noises (bottom panel), based on the joint abundance. Note that in this case, too, the normalization cancels the dependence on the absolute number of objects, strongly reducing the cosmological dependence in this analysis.

We can compute the probability  $p^i$  that the cluster will have a measured SZE signal  $\xi > 4$  as

$$p^i = \int d\zeta P(\xi > 4|\zeta) P(\zeta|\hat{\lambda}^i, z^i, \gamma_f^i). \quad (10)$$

This is referred to hereafter as the confirmation probability.

In this simple example, the likelihood of matching a system is given by the probability of being confirmed  $p^i$ , as we remind the reader that in a Bayesian context the likelihood is defined as the probability of the data given the model – albeit that in this case the datum is a boolean (being or not being matched). Similarly, the likelihood of a not matched system is given by the probability of not being confirmed, that is  $1 - p^i$ .

As in Grandis et al. (2020), we extend this formalism to investigate different properties of the selection function. Following the probability tree shown in Fig. 4 from top to bottom, we first entertain the possibility that the redMaPPer object is not described by our richness-mass scatter model, and thus is a contaminant. This is modelled by the contamination fraction  $\pi_c$ . Every redMaPPer object has only a chance  $1 - \pi_c$  to be well described by of richness-mass scatter model and, conversely, a chance  $\pi_c$  of being a contaminant. If the object is a cluster, the detection probability  $p^i$  affects whether it is detected by SPT or not.

Following the probability tree and adding up the weight of all the branches leading to a redMaPPer cluster ending as a non-detection (represented by the notation ‘!SPT’), the likelihood is

$$p(\text{!SPT}|\hat{\lambda}^i, z^i) = \pi_c + (1 - \pi_c)(1 - p^i). \quad (11)$$

For matched objects (denoted as ‘SPT’), the likelihood is

$$p(\text{SPT}|\hat{\lambda}^i, z^i) = (1 - \pi_c)p^i. \quad (12)$$

The total likelihood for the redMaPPer sample can then be obtained by summing the log-likelihood of the individual clusters based on whether they have been detected by SPT or not, reading

$$\ln \mathcal{L} = \sum_{i \in \text{matched}} \ln p(\text{SPT}|\hat{\lambda}^i, z^i) + \sum_{i \in \text{!matched}} \ln p(\text{!SPT}|\hat{\lambda}^i, z^i). \quad (13)$$

This likelihood depends on the scaling relation parameters as well as on the contamination fraction.

We employ this model to investigate the case of a richness dependent contamination (abbreviated ‘cont’ in the following), where the contamination probability is modelled as

$$\pi_c(\hat{\lambda}) = \frac{A(\hat{\lambda})}{1 + A(\hat{\lambda})} \text{ with } A(\hat{\lambda}) = A_c \left( \frac{\hat{\lambda}}{45} \right)^{B_c}. \quad (14)$$

In this parametrization  $A_c > 0$  and arbitrary  $B_c$  lead to values of  $0 < \pi_c(\hat{\lambda}) < 1$  for any value of richness  $\hat{\lambda}$ .  $A_c$  and  $B_c$  are free parameters of our analysis in this case.

### 3.5 Model predictions of observables

After having determined the richness–mass scaling relation parameters, we employ the different posteriors on these parameters to predict several quantities which we compare with the data: (1) the fraction of SPT detected redMaPPer clusters as a function of redshift and measured richness, (2) the mean mass of clusters in redshift and richness bins, and (3) the number of redMaPPer clusters in redshift and measured richness bins. These are each described in more detail below.

#### 3.5.1 SPT confirmation fraction in richness/redshift bin

The probability of a cluster of intrinsic richness  $\lambda$  and redshift  $z$  falling into an observed richness and redshift bin  $j$ , defined by  $\hat{\lambda}_-^j < \hat{\lambda} < \hat{\lambda}_+^j$  and  $z_-^j < z < z_+^j$ , is

$$\begin{aligned} P(j|\lambda, z) &= \chi_{z_-^j}^{z_+^j}(z) \int_{\hat{\lambda}_-^j}^{\hat{\lambda}_+^j} d\hat{\lambda} P(\hat{\lambda}|\lambda, z) \\ &= \chi_{z_-^j}^{z_+^j}(z) [P(\hat{\lambda} > \hat{\lambda}_-^j|\lambda, z) - P(\hat{\lambda} > \hat{\lambda}_+^j|\lambda, z)]. \end{aligned} \quad (15)$$

Here,  $\chi_{z_-^j}^{z_+^j}(z) = 1$  for  $z_-^j < z < z_+^j$ , and 0 elsewhere. Note that in the limit of infinitely small bins around a measured richness  $\hat{\lambda}^i$  and redshift  $z^i$ , that is  $\hat{\lambda}_{+/-}^j \rightarrow \hat{\lambda}^i$  and  $z_{+/-}^j \rightarrow z^i$ , this equation tends towards the error model for a single cluster (equation 5).

The expected distribution of intrinsic SZE signal  $\zeta$  in the bin then follows closely the expression for a single cluster given in equation (9)

$$P(\zeta|j, \gamma_f) \propto \int dz \int d\lambda P(j|\lambda, z) \frac{d^2 N}{d\zeta d\lambda} \Big|_{z, \gamma_f}, \quad (16)$$

where the normalization is given by the condition  $\int d\zeta P(\zeta|j, \gamma_f) = 1$ . As above, this normalization makes this expression lose much of its dependence on cosmology.

The fraction of the redMaPPer clusters in bin  $j$  that are then also in the SPT sample,  $f(\text{SPT}|j)$ , is obtained by convolving the predicted distribution of intrinsic SZE signal with the SPT selection function given by the condition  $\xi > 4$ ,

$$f(\text{SPT}|j) = \frac{1}{\sum_f \Omega_f} \sum_f \Omega_f \int d\zeta P(\xi > 4|\zeta) P(\zeta|j, \gamma_f), \quad (17)$$

where  $\Omega_f$  is the solid angle of the field  $f$  in the joint footprint. The weighted sum over the fields properly accounts for the spatially varying SPT-SZ survey depth.

Following from equations (12) and (11), the effects of contamination are included by substituting

$$f(\text{SPT}|j) \mapsto (1 - \pi_c(\hat{\lambda}^j))f(\text{SPT}|j), \quad (18)$$

where we choose  $\hat{\lambda}^j$  as the geometrical mean of the  $\hat{\lambda}_{+/-}^j$ .

#### 3.5.2 Mean mass in richness/redshift bin

The prediction for the mean mass of clusters in an observed richness and redshift bin  $j$  defined by  $\hat{\lambda}_-^j < \hat{\lambda} < \hat{\lambda}_+^j$  and  $z_-^j < z < z_+^j$  can be computed from the predicted distribution of masses

$$P(M|j) \propto \int dz \int d\lambda P(j|\lambda, z) P(\lambda|M, z) \frac{dN}{dM} \Big|_{M, z}, \quad (19)$$

where the normalization is given by  $\int dM P(M|j) = 1$ . The mean mass  $\bar{M}(j)$  is then simply

$$\bar{M}(j) = \int dM M P(M|j). \quad (20)$$

Assuming that a fraction  $\pi_c(\hat{\lambda}^j)$  of the objects in the bin are contaminants, we can investigate the mean mass of the contaminants  $\bar{M}_c^j$ . The mean measured weak lensing mass  $\hat{M}_{\text{WL}}^j$  in this bin, as inferred in stacked WL analyses (McClintock et al. 2019; Abbott et al. 2020), can then be expressed as

$$\hat{M}_{\text{WL}}^j = (1 - \pi_c(\hat{\lambda}^j))\bar{M}(j) + \pi_c(\hat{\lambda}^j)\bar{M}_c^j \pm \delta \hat{M}_{\text{WL}}^j, \quad (21)$$

where  $\delta \hat{M}_{\text{WL}}^j$  is the reported measurement error on the mean mass. We convert the measurements of the mean mass in redMaPPer richness and redshift bins within the DES collaboration (McClintock et al. 2019; Abbott et al. 2020) to the mass definition of  $M_{500c}$  by assuming an NFW mass profile and the mass concentration relation from Bhattacharya et al. (2013). Together with the predictions on  $\bar{M}(j)$  and  $\pi_c(\hat{\lambda}^j)$  from this work, we can predict the mean mass of the contaminants

$$\bar{M}_c^j = \bar{M}(j) + \pi_c^{-1}(\hat{\lambda}^j)(\hat{M}_{\text{WL}}^j - \bar{M}(j)) \pm \pi_c^{-1}(\hat{\lambda}^j)\delta \hat{M}_{\text{WL}}^j. \quad (22)$$

The difference between the contaminants mean mass  $\bar{M}_c^j$  and the expected mean mass  $\bar{M}(j)$  is thus sourced by the difference between the measured WL mass  $\hat{M}_{\text{WL}}^j$  and the expected mean mass  $\bar{M}(j)$ , modulated by the inverse of the contamination. Few contaminants with masses very different from the expected mass have the same impact on the measured mean mass as many contaminants with masses closer to the mean expected mass.

Physically, the hypothesis is that contaminants in optical cluster selection are low-mass groups or even individual massive red galaxies residing in a line of sight crowded by other red galaxies at approximately the same photometric redshift.



### 3.5.3 Number counts in richness/redshift bin

The number of redMaPPer clusters  $N(j)$  in a richness and redshift bin  $j$  defined by  $\hat{\lambda}_-^j < \hat{\lambda} < \hat{\lambda}_+^j$  and  $z_-^j < z < z_+^j$  is given by

$$N(j) = \int dz \Omega(z) \int d\lambda P(j|\lambda, z) \quad (23)$$

$$\int dM P(\lambda|M, z) \frac{dN}{dM} \Big|_{M,z}, \quad (24)$$

where  $\Omega(z)$  is the solid angle in units of degrees in which an object with redshift  $z$  can be detected (Abbott et al. 2020). This is expressed as a function of redshift to encode that redMaPPer clusters are only presented within particular redshift ranges. We compare our prediction for the number counts to other code available within DES (Abbott et al. 2020) and find that the numerical differences are clearly smaller than the systematic uncertainties on the quantity derived from marginalizing over our richness–mass relation and the cosmological parameters.

In the case of contamination, we can assume that the total number of objects  $N_{\text{tot}}(j)$  in the bin  $j$  is the sum of the number of clusters  $N(j)$  and contaminants  $N_c(j)$ . Using  $N_c(j) = \pi_c(\hat{\lambda}^j)N_{\text{tot}}(j)$ , we find that

$$N_{\text{tot}}(j) = \frac{N(j)}{1 - \pi_c(\hat{\lambda}^j)} = (1 + A(\hat{\lambda}^j))N(j), \quad (25)$$

where the last transformation is made using the parametrization of the richness dependence of the contamination given by equation (14). This provides a physical interpretation for that parametrization:  $A(\hat{\lambda})$  is the ratio between the number of contaminants and the number of clusters at the richness  $\hat{\lambda}$ , as  $N_c(j) = A(\hat{\lambda}^j)N(j)$ .

### 3.6 Code and model validation

We validate our code in two stages. First, we test for potential biases induced by the model fitting. To do this, we create synthetic data sets drawn from our models assuming specific input parameters. We then run our likelihood analysis on the synthetic data and compare the posteriors on the parameters to the known input values. Here, we draw a synthetic cluster sample in a footprint that is 10 times larger than the actual survey footprint. This reduces statistical noise by a factor of  $\sqrt{10}$ . With this approach, we demonstrate that the code recovers the input parameters at better than  $1\sigma$ . Biases in our analysis of the real data are thus smaller than  $\sim 0.3\sigma$ .

The second stage of the validation tests whether the adopted model is an adequate description of the observations. This question in an inherently scientific question that can only be answered by analysing the actual data. In the course of this work, we then pursue three lines of argument to assess the adequacy of our models. First, we investigate if our model provides a good fit to the data. Given that our likelihoods are not Gaussian linear models, we cannot use a  $\chi^2$ -test. Generalizations of the  $\chi^2$ -test to arbitrary Bayesian likelihood analysis are beyond the scope of this work, so here we employ visual comparison between the data and the model prediction. In order to generalize  $\chi^2$ -test to arbitrary Bayesian likelihood analysis, one needs to draw a large amount of mock data sets, and run full likelihood analyses on them (Nicola, Amara & Refregier 2019; Doux et al. 2021). Our prior experience with simplified versions of such methods in Bocquet et al. (2015, 2019) revealed that in all cases where advanced Bayesian method detected ‘bad’ fits to the data, this was also abundantly clear by visual inspection. Furthermore, these methods are of questionable efficacy from a statistical and philosophical perspective (Kerscher & Weller 2019). In the cases of interest to this work, specifically the question whether our prediction

for the SPT confirmation fraction matches the measured confirmation fraction, visual inspection provides sufficient discriminatory power. Another element to assess the adequacy of the model is comparison to independent external results. Lastly, if a model is physically plausible, this provides further support for it, or, conversely, if the model makes implausible predictions, that would provide reason to reject it. Either way, exploring which predictions your model makes, and how these could be tested, is of importance for its validation.

## 4 RESULTS

In this section, we present our main results, starting with the calibration of the richness–mass relation based on the SPT cross matching. We then proceed to constrain the redMaPPer contamination fraction. We finally seek to extend the measurement of the redMaPPer contamination to lower richnesses by comparing measured and predicted number counts of optically selected clusters and their stacked WL signal.

### 4.1 Cross calibration of $\lambda$ –mass relation

We present our calibration of the richness–mass scaling relation parameters and then examine what our sample can tell us about the sample contamination.

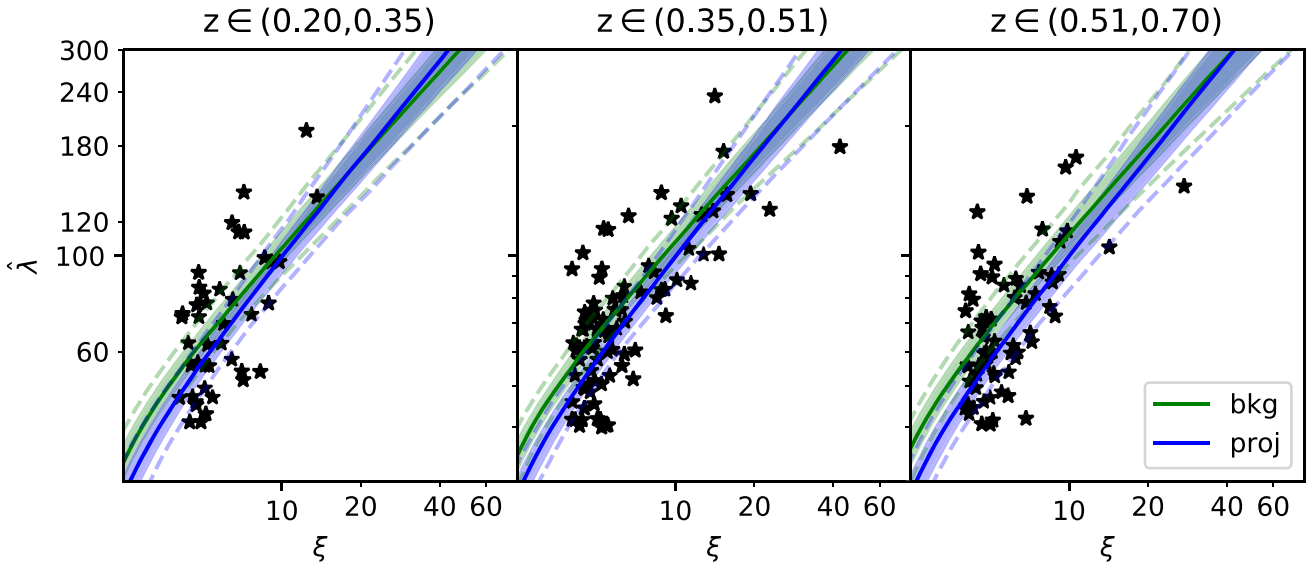
#### 4.1.1 Scaling relation parameters

As outlined in Section 3.3, the measured richnesses and SZE signals for clusters in the cross matched redMaPPer–SPT sample enable us to transfer the calibration of the SZE signal–mass relation given by published priors (Bocquet et al. 2019) to the richness–mass relation. The scatter plot of the matched sample in redshift bins is shown in Fig. 5 with black points.

The resulting posteriors on the parameters of the richness–mass scaling relation are summarized in Table 1 and in Fig. 6. In blue, we show the constraints from adopting the projection optical error model and sampling the cross-calibration likelihood, while in green we show the posteriors when using the background optical error model. The constraints are in very good agreement, highlighting that at the level of statistical constraining power of the cross matched sample, the two error models are not distinguishable. The minor changes induced by the projection effects can be seen by comparing the two mean relations shown in Fig. 5. For comparison in Fig. 6, we also plot in grey bands previous results by Bleem et al. (2020) from analysing the redMaPPer richness–mass relation of the SPT–SZ and the SPTpol Extended Cluster Survey. This analysis is not independent on our work as the data partially overlap: Bleem et al. (2020) used the entire SPT–SZ sample together with the SPTpol Extended Cluster Survey and DES Y3 data, while we use DES Y1 data and restrict ourselves to the part of the SPT–SZ sample in the DES–Y1 footprint. The good agreement with our results is thus more a consistency check for the accuracy of our analysis method than an assessment on the adequacy of the assumed model.

#### 4.1.2 Correlated scatter

The correlation coefficient  $\rho$  between the scatters in observables carries important astrophysical information on the physical processes inside galaxy clusters (as shown in simulations by Stanek et al. 2010; Angulo et al. 2012; Wu et al. 2015; Farahi et al. 2018; Truong et al.



**Figure 5.** The richness as a function of SZE signal of the matched sample (black dots) in three different redshift bins with mean redshift increasing from left-hand panel to right-hand panel. Overplotted is the mean richness–SZE signal relation and associated 68 (filled area), and 95 per cent (dashed line) uncertainties resulting from the analyses assuming the two error models (blue: projection model, green: background only model). The small deviation of solid-blue lines from solid-green lines reflects the impact of projection effects of the mean relations.

**Table 1.** Mean and standard deviation estimated from the single parameter posteriors for the richness–mass scaling relation when cross-calibrated using priors on the SZE–mass relation under the assumption of two different  $\lambda$  measurement noise models: (1) background (equation 5) and (2) projection (Costanzi et al. 2019a). ‘+ cont’: includes confirmation fraction. ‘–’ appears for the optical–SZE scatter correlation coefficient  $\rho$ , because the parameter is not constrained.

	$A_\lambda$	$B_\lambda$	$C_\lambda$	$\sigma_\lambda$	$\rho$	$A_c$	$B_c$
Background							
SPT calibr	$83.3 \pm 11.2$	$1.03 \pm 0.10$	$0.69 \pm 0.42$	$0.22 \pm 0.06$	–		
+ cont	$82.7 \pm 11.6$	$1.06 \pm 0.10$	$0.38 \pm 0.31$	$0.22 \pm 0.07$	–	$0.94 \pm 0.58$	$-1.69 \pm 0.89$
Projection							
SPT calibr	$72.4 \pm 10.6$	$1.15 \pm 0.12$	$0.41 \pm 0.50$	$0.23 \pm 0.06$	–		
+ cont	$71.9 \pm 10.7$	$1.23 \pm 0.12$	$0.03 \pm 0.39$	$0.23 \pm 0.08$	–	$1.15 \pm 0.66$	$-2.26 \pm 1.01$

2018). Ignoring this correlation in cluster population studies can lead to biases in the inferred scaling relation (Allen, Evrard & Mantz 2011; Angulo et al. 2012). In this work, the correlation coefficient is free to vary between  $-1$  and  $1$ . Unfortunately, our data do not provide any significant constraint on  $\rho$ , because our data only determine the total scatter among two observables. While the observational contribution to that total scatter can be accounted for, the total intrinsic scatter in richness at a given SZE signal is

$$\sigma_{\text{tot}}^2 \approx \sigma_\lambda^2 + \left( \frac{B_\lambda}{B_{\text{SZE}}} \sigma_{\text{SZE}} \right)^2 - 2\rho\sigma_\lambda \left( \frac{B_\lambda}{B_{\text{SZE}}} \sigma_{\text{SZE}} \right). \quad (26)$$

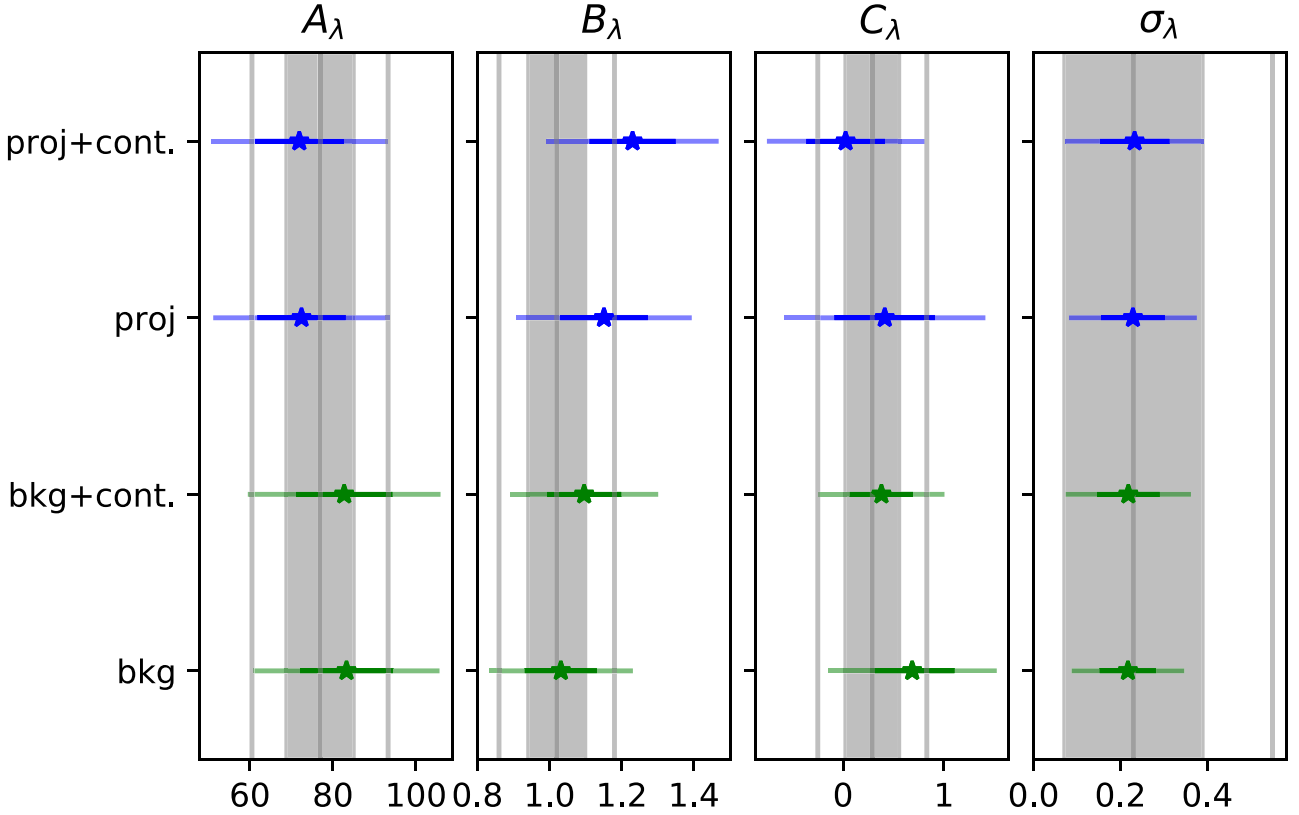
Imposing a prior on the intrinsic scatter  $\sigma_{\text{SZE}}$  and the mass trend  $B_{\text{SZE}}$  of the SZE observable, one can constrain the slope  $B_\lambda$  of the richness–mass relation from the average trend in the data (cf. Fig 5). This results in a degeneracy between the intrinsic scatter in richness  $\sigma_\lambda$  and the correlation coefficient  $\rho$ , which we show in Fig. 7. Larger intrinsic scatter in richness can be accommodated by positive correlation coefficients, while smaller scatter requires negative correlation coefficients. Due to this fundamental degeneracy, we do not expect to measure the correlation coefficient. Recent reported detections of correlations in the scatter of two observables (Mulroy et al. 2019;

Farahi et al. 2019a) are likely sourced by the assumption that the intrinsic scatter of the weak lensing inferred mass is perfectly known. The lack of a detection of the correlation coefficient in this work thus reflects a more refined handling of systematic uncertainties.

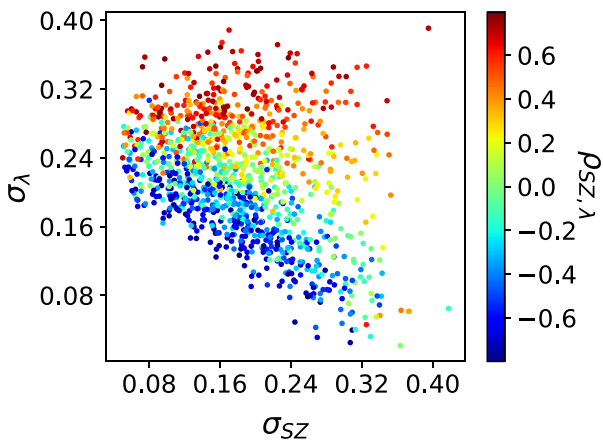
#### 4.1.3 SPT confirmation probabilities

As discussed in Section 3.4, the probability of confirming a redMaPPer cluster in SPT is very sensitive to the respective scaling relation parameters of the two selection observables. Consequently, one needs to marginalize over a reasonable range of scaling relation parameters when inferring the contamination level of the redMaPPer sample with the likelihood given in equation (13) (see also equation 10). To do this, we sample the likelihood of confirmation probabilities simultaneously with the cross-calibration likelihood (equation 8). This ensures proper accounting for the systematic uncertainties on the scaling relations when inferring the contamination fractions. The resulting posteriors, depending on the assumed optical error model, are shown in green and blue in Fig. 6 for the background (‘bkg’; equation 5) and projection (‘proj’; Costanzi et al. 2019a) model, respectively, and summarized in Table 1.





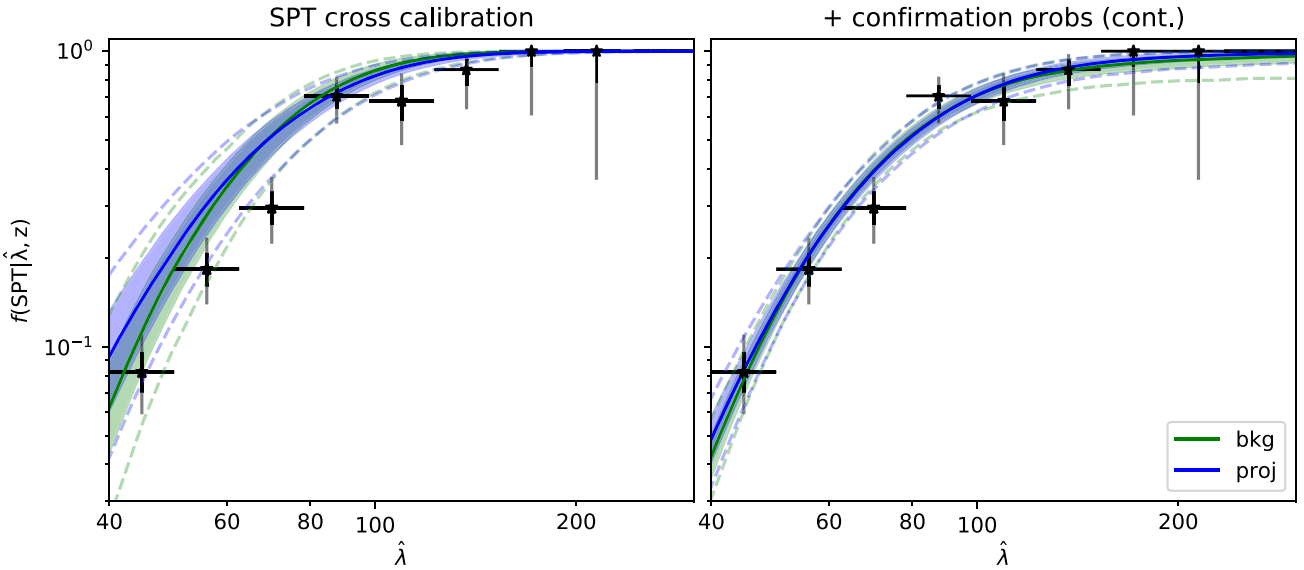
**Figure 6.** Means and standard deviations estimated from the single parameter posteriors for the richness–mass scaling relation when cross-calibrated using priors on the SZE–mass relation under the assumption of two different  $\hat{\lambda}$  measurement noise models: (1) ‘background’ in green (equation 5) and (2) ‘projection’ in blue (Costanzi et al. 2019a). ‘+ cont’: includes confirmation fraction (richness-dependent failure fraction of the richness–mass scatter model). The full errorbars denote one standard deviation, while the shaded ones denote two standard deviations grey bands denote the mean and standard deviation of the richness–mass relation parameters reported by Bleem et al. (2020), with the vertical lines denoting two standard deviations. The results of the different analyses are mutually consistent.



**Figure 7.** The data primarily constrain the total richness–SZE signal scatter, so there are significant degeneracies between the three scatter terms in the model, the intrinsic scatters in richness and SZE signal as well as the correlation coefficient between these two scatters. Assuming a prior on the SZE scatter, constraints on the correlation coefficient are not to be expected. We illustrate this degeneracy by plotting a random subset of the posterior chain of the background only cross-calibration analysis.

The posteriors on the scaling relation parameters are generally in good agreement with the results without the confirmation probability likelihood. The detection probability likelihood slightly alters the posteriors on the scaling relation parameters, having the largest impact on the mass trend and the redshift trend. This is not surprising when considering that not detecting a redMaPPer object in SPT is equivalent to the measurement  $\xi < 4$ , which given priors on the SZE signal–mass relation carries some mass information, at least in the form of an upper limit. This information is, however, quite weak, as can be seen by examining the change in measurement uncertainties. Furthermore, it is consistent with the information recovered from the matched sample alone, because the shifts are mean values also do not exceed  $1\sigma$ .

Using equation (17), we predict the SPT confirmation fraction of the redMaPPer ( $\hat{\lambda} > 40$ ) sample as a function of measured richness, shown in Fig. 8. The shaded region (dashed line) reflects the  $1$  ( $2$ )  $\sigma$  systematic uncertainty as propagated from the posterior samples for the background error model (green), and the projection model (blue). We also plot the measured confirmation rate as black point with errorbars. Note also here that the difference between the predictions for the two error models is small. In the left-hand panel, we show the prediction for the baseline SPT calibration with no contamination, while in the right-hand panel we show the case of richness dependent contamination. The case without contamination tends to overpredict

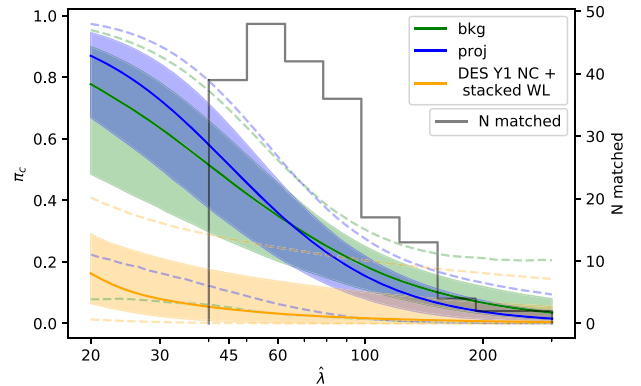


**Figure 8.** SPT confirmation fraction as a function of measured richness and associated 68 (filled area), and 95 per cent (dashed line) uncertainties, predicted from the posteriors on the scaling relation parameters for the background error model (‘bkg’, green), overlaid with the measurement as black points (black error bars are  $1\sigma$ , grey error bars  $2\sigma$  uncertainty). Left-hand panel shows the predictions for the SPT calibration only, and the right-hand panel shows the addition of the confirmation fraction likelihood in the contamination case. Clearly, the confirmation fraction resulting from our richness–mass model alone is insufficient to predict the actual confirmation fraction. We interpret this as the presence of a population of redMaPPer selected objects for which our scatter model fails. We call this population contaminants. The contamination fraction is richness dependent and provides a better fit to the SPT confirmation fraction of redMaPPer objects, while the choice of ‘bkg’ or ‘proj’ makes little difference.

the confirmation fraction. In contrast, the richness dependent contamination provides a better description of the data. Thus, the case with richness dependent contamination provides a better fit to the SPT confirmation of redMaPPer objects with measured richness  $\hat{\lambda} > 40$ , than the case without contamination.

In Fig. 9, we show our posterior on the contamination fraction as a function of richness for the two optical error models considered (blue: accounting for projection effects, green: background only). The two posteriors are in good agreement with each other when considering the uncertainties we derive. We also show in grey the number of SPT-SZ confirmed redMaPPer. This indicates that the bulk of the constraining power of the confirmation fraction comes from richnesses  $\hat{\lambda} \sim 60$ . We also present the indirect constraint on the contamination fraction that has been derived independently (Abbott et al. 2020) using the number counts of redMaPPer objects, mean weak lensing mass measurements of redMaPPer objects, and cosmological priors from the shear and red galaxy auto- and cross-correlations (Abbott et al. 2018). That analysis predicts a significantly lower contamination than our results at about the  $2\sigma$  level.

Importantly, the external analysis (Abbott et al. 2020) assumes that the mean contaminants have zero mass, and the contamination fraction they derive provides a poor fit to the mean WL masses in the lowest richness bins of the redMaPPer sample. Furthermore, assuming the lower contamination fraction as priors, the cosmological constraints from refitting the number counts, and mean masses were not in agreement with the cosmological constraints from the shear and red galaxy auto- and cross-correlations (Abbott et al. 2018). In summary, there are apparent qualitative and quantitative differences between our constraints on the redMaPPer contamination fraction, and those from the external analysis (Abbott et al. 2020). The contamination levels we find at lower richness are surprisingly large. We will demonstrate in the following that these high contamination levels are physically plausible.



**Figure 9.** Constraints on the fraction of failures in our richness–mass scatter model, which we call contaminants, as a function of measured richness from our analyses of the SPT confirmation fraction of redMaPPer objects with two different optical error models (blue: accounting for projection effects, green: background only), with associated 68 (filled area), and 95 per cent (dashed line) uncertainties. In yellow, the result from Abbott et al. (2020), when fitting mean weak lensing masses and number counts of redMaPPer objects with cosmological priors from the shear and red galaxy auto- and cross-correlations (Abbott et al. 2018). These measures are consistent with our results at less than  $2\sigma$ . In grey, the richness distribution of redMaPPer systems confirmed by SPT, indicating that the bulk of our constraining power comes richnesses  $\hat{\lambda} \sim 60$ .

Finally, we stress that ‘contaminants’ in this work simply reflect objects that at a given richness do not conform with our SZE derived mass distribution. The Bayesian population model allows us to perform such a derivation while taking account of several biases [mainly the *Eddington bias* (e.g. Mortonson, Hu & Huterer 2011)

and the *Malmquist bias* (e.g. Allen, Evrard & Mantz 2011)]. Both of these biases depend on scatter between observables and mass that occurs also below the selection thresholds. Extrapolating an inadequate scatter model can thus bias the mean relation and alter our definition of ‘contaminant’. An instance of such an effect is described in the discussion section below (cf. Section 5.5).

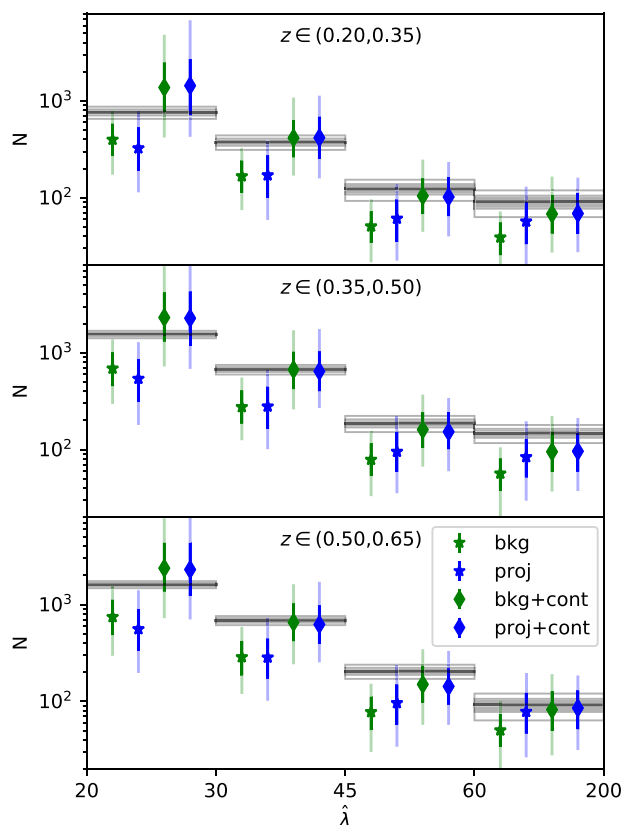
## 4.2 Comparison to redMaPPer observables

In the previous section, we determined the systematic uncertainty on the richness–mass relation in the regime of measured richness  $>40$ , which is at the intermediate to high-mass end. We also established that a sizeable amount of contamination provides a better fit to the observed SPT confirmation fraction than alternative models. In this section, we address whether such large contamination fractions are consistent with the measured number of redMaPPer objects as a function of richness and redshift. We then employ the mean mass from stacked weak lensing measurements in richness–redshift bins to estimate the mean mass of the contaminants. We consider richness and redshift bins with edges at  $\lambda \in (20, 30, 45, 60, 200)$  and redshift bins with edges  $z \in (0.20, 0.35, 0.50, 0.60)$  (the same used in Abbott et al. 2020), from which we also take the measured masses and uncertainties.

### 4.2.1 Number counts

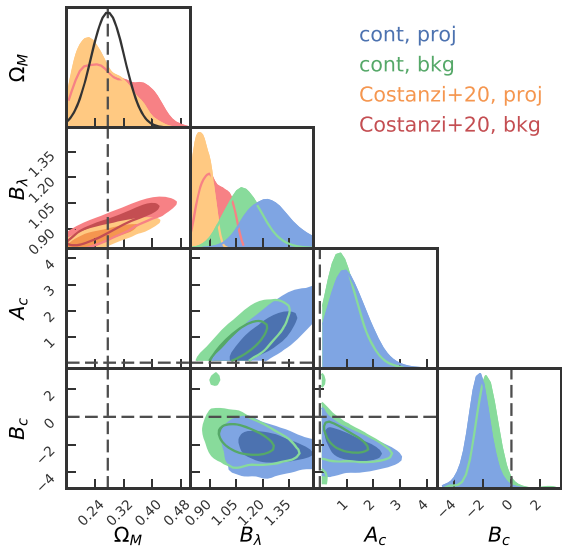
We predict the number of redMaPPer clusters in richness and redshift bins using the posteriors on the richness–mass relation derived with and without the contamination fraction model (equations 23 and 25, respectively). As stated above, we not only propagate the uncertainty on the richness–mass scaling relation parameters, but also those on the cosmological parameters by sampling the cosmological parameters within the priors reported in Section 3.3. This results in a prediction of the number of objects with large systematic uncertainties, as shown in Fig. 10 for the projection model (blue), and the background model (green). Stars denote the prediction assuming no contamination, while diamonds denote the prediction accounting for our constraints on the contamination. In the same figure, we also plot as grey bands the number of redMaPPer objects with their statistical uncertainties (Abbott et al. 2020), which are considerably smaller than the systematic uncertainties.

The predictions with richness dependent contamination provide a better description than the predictions without contamination. This shows that the contamination fraction that we fit is not in contradiction with the measured number of redMaPPer clusters. While uncertainties on the contamination fraction are still large, especially the lower limit we place on the contamination fraction is clearly compatible with measured number counts. Conversely, the maximum likelihood value of our contamination posterior overpredicts the number counts in the lowest richness bin. This statement also holds true if instead of the SPT cosmology (Bocquet et al. 2019), we use the cosmological constraints from the study of the cosmic shear and galaxy auto- and cross-correlations in DES-Y1 data (Abbott et al. 2018), as shown in Appendix B. We would like to stress here, that we do not fit for the number counts or redMaPPer objects, but that we extrapolate our scaling relation (that is calibrated for  $\hat{\lambda} > 40$ ) and SPT cosmological results (derived for  $M \gtrsim 3 \times 10^{14} M_{\odot}$ ) to lower masses and richnesses. Indeed, the better agreement of external data with our prediction in the presence of contamination provides further support for the presence of a considerable richness dependent contamination in the redMaPPer sample.



**Figure 10.** The number of redMaPPer objects in bins of richness and redshift are shown as predictions (coloured points) derived from our richness–mass relation constraints and SPT cluster number counts cosmology for the different optical error models (blue projection, green background). Full error bars show the 16th/84th-percentile, while shaded error bars show the 2.5th/97.5th percentile. Stars denote predictions without contamination, while diamonds denote predictions with contamination. Overplotted in grey is the number of redMaPPer objects with the associated statistical uncertainty (empty boxes showing the  $2\sigma$  uncertainty). When considering the systematic error bars, the prediction with richness dependent contamination (see Fig. 9) provides a better match to the data than the prediction without contamination.

An interesting comparison can be drawn with the work by Costanzi et al. (2020). In that work, the redMaPPer number counts were fitted jointly to the SPT WL follow-up data (Bocquet et al. 2019) that was used to derive the priors on the SZ scaling relation parameters for our analysis, while the SPT confirmation fraction was not fit. A strong degeneracy between the matter density  $\Omega_M$  and the mass trend of the richness,  $B_{\lambda}$ , was found, as shown in Fig. 11. Note that in that work, the authors did not entertain the possibility that the richness–mass model would fail, leading to contamination. When comparing to their constraints on the richness–mass trends, we find them to be lower than but consistent with our results. Assuming the same matter density prior as we did (Fig. 11, in black) would improve the agreement. More interestingly, we find in our analysis of the SPT confirmation fraction of redMaPPer objects a strong degeneracy between the amplitude of contamination  $A_c$  and the mass trend of the richness  $B_{\lambda}$ , suggesting that shallower mass trends would lead to less contamination. Furthermore, a degeneracy between the contamination fraction and the cosmological constraints is to be expected when fitting the number counts. Future joint analysis of the SPT confirmation fraction and the redMaPPer number counts are



**Figure 11.** Marginal posteriors on the matter density  $\Omega_M$ , the mass trend of the richness mass relation  $B_\lambda$ , and the parameters amplitude  $A_c$  and richness trend  $B_c$  of the contamination fraction from our work (green: ‘bkg’, blue: ‘proj’), in comparison with the joint analysis of the SPT multiwavelength data and the redMaPPer number count by Costanzi et al. (2020) (red and yellow), which did not fit the SPT confirmation fraction. Black is the  $\Omega_M$  posterior from Bocquet et al. (2019). Clearly, the amplitude of the contamination fraction is degenerate with the mass trend of the richness. In the context of redMaPPer number counts fits, the matter density, and the richness slope are degenerate.

expected to correctly explore these degeneracies in the likelihoods while plausibly putting tighter constraints on the contamination fraction.

#### 4.2.2 Estimating the mean mass of contaminants

As discussed earlier, our richness dependent contamination model provides a good fit to the SPT confirmation fraction of redMaPPer objects, as well as an improved prediction for the number of redMaPPer objects in richness and redshift bins. Knowing the mean mass of the clusters from our SPT calibration and the contamination fraction from our study of SPT confirmations we can employ the measured mean mass in richness–redshift bins to estimate the mean mass of the contaminants (equation 22). Mean cluster masses, contamination fractions, and resulting mean contaminant masses are reported in Table 2 and in Fig. 12. In the lowest richness bin  $20 < \hat{\lambda} < 30$ , we find that  $66 \pm 21$  per cent of the objects are contaminants with a mean mass of  $\sim 3 \times 10^{13} M_\odot$ , compared to a cluster population with a mean mass of  $\sim 1 \times 10^{14} M_\odot$ . These constraints vary little with redshift. In the next lowest bin ( $30 < \hat{\lambda} < 45$ ), the contaminants make up  $51 \pm 18$  per cent of the population and have a mean mass of  $\sim 5 \times 10^{13} M_\odot$ , compared to the cluster populations mean mass of  $\sim 1.6 \times 10^{14} M_\odot$ . These mean masses of clusters and contaminants are roughly located at the extremes of the mass constraints derived from stacked weak lensing (Abbott et al. 2020, see Fig. 9), supporting their overall physical plausibility.

For the higher richness bins, the central value of the mean contaminants mass approaches the expected mean mass. In light of equation (22), this is the natural consequence of the smaller difference between the expected mean mass and the measured mean WL mass. Furthermore, the errorbars on the mean contaminants

mass become very large at higher richness. As can be seen again in equation (22), the error scales like the inverse contamination fraction  $\pi_c(j)^{-1}$ . At larger richness, the contamination fraction tends to zero, increasing the error on the mean contaminants mass.

From a physical perspective, it is both convenient and reasonable to refer to these objects as ‘contaminants’, because their masses indicate that these are galaxy groups rather than galaxy clusters. Our measurement is consistent with the notion that every optically selected object is associated with a halo of some mass. We find that for the redMaPPer sample, these contaminants constitute a significant fraction of the objects and have masses lying in the group mass range ( $\sim 3\text{--}5 \times 10^{13} M_\odot$ ).

## 5 DISCUSSION

In the following, we discuss the major results of our work. We first compare our results on the scatter and the mass trend of the richness–mass relation to other works. Subsequently, the impact of SZE contamination on our results is discussed. We then present other evidence for and against the contamination fraction and mean contaminant masses we have measured. Finally, alternative scenarios to the richness dependent contamination case are outlined, together with future prospects of discriminating among them.

### 5.1 Comparison to literature

Our comparison to the literature focuses on the two scaling relation parameters which are most closely linked to the richness dependent contamination: the intrinsic scatter and the mass trend or power-law slope. In the following, we consider also results based on a redMaPPer selected sample obtained from *Sloan Digital Sky Survey* data (SDSS; see Rykoff et al. 2014, for the discussion of the redMaPPer application), because those richnesses are consistent with the richnesses extracted from DES (see McClintock et al. 2019, equations 66–67). We will not include constraints based on the application of other optical cluster finders or richness measurement methods, as it is unclear how they are impacted by projection effects compared to redMaPPer.

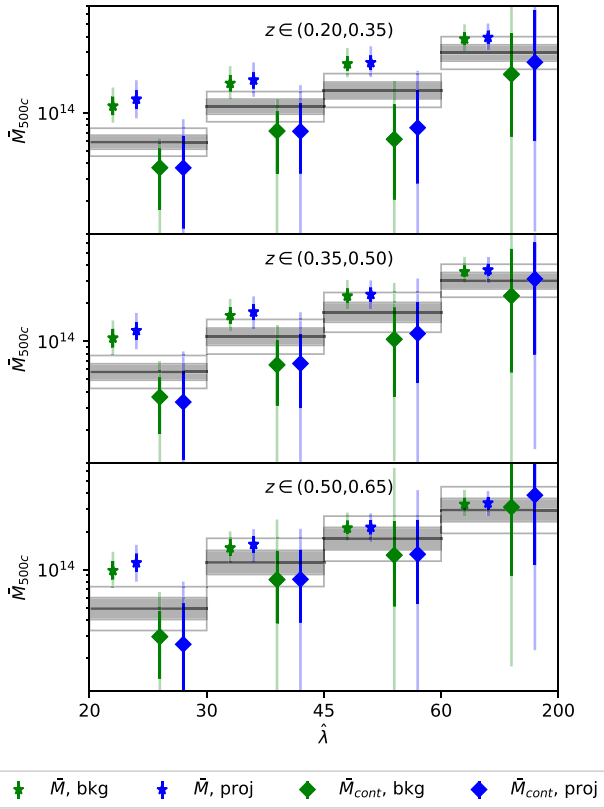
#### 5.1.1 Constraints on scatter

The intrinsic scatter in the observable–mass relation directly affects the mass incompleteness of cluster samples. Measurements of the scatter are made even more important in many studies of optically selected clusters, because weak lensing mass constraints are extracted from stacked observations of many clusters in bins of richness and redshift (e.g. Simet et al. 2017; Murata et al. 2018, 2019; McClintock et al. 2019). Due to the stacking, such studies lose leverage on the scatter, while Bayesian population modelling retains more of the constraining power in the data (Grandis et al. 2019). Noticeably, Sereno et al. (2020) recently suggested a Bayesian modelling approach to stacking that retains some constraining power on the intrinsic scatter. Cross calibration with ICM based mass proxies – the technique applied in this work – is a useful, and more widely used technique for constraining the scatter. Rozo et al. (2015) cross calibrated the SDSS–redMaPPer sample (Rykoff et al. 2014) with the cluster catalogue selected via SZE in the first all sky Planck survey (Planck Collaboration XXXII 2015), investigating the scaling between the richness and the SZE-inferred masses of 191 clusters.



**Table 2.** Mean cluster masses  $\bar{M}$ , contamination fractions  $\pi_c(\hat{\lambda}^j)$ , and resulting mean contaminants masses  $\bar{M}_c$  in richness–redshift bins. These values reflect the systematic uncertainty on the richness–mass relation and the richness dependent contamination from the posteriors of the SPT calibration and the SPT confirmation probability fits for the background only model.

Richness bins	(20, 30)	(30, 45)	(45, 60)	(60, 200)
Redshift bins				
$\log_{10} \bar{M}$				
(0.2, 0.35)	14.06 ± 0.07	14.23 ± 0.06	14.39 ± 0.06	14.58 ± 0.05
(0.35, 0.50)	14.02 ± 0.07	14.20 ± 0.06	14.36 ± 0.06	14.55 ± 0.05
(0.50, 0.60)	14.00 ± 0.07	14.17 ± 0.06	14.33 ± 0.06	14.52 ± 0.05
$\pi_c(\hat{\lambda}^j)$	0.66 ± 0.21	0.51 ± 0.18	0.38 ± 0.14	0.16 ± 0.06
$\log_{10} \bar{M}_c$				
(0.2, 0.35)	13.5 ± 0.3	13.7 ± 0.3	13.7 ± 0.5	14.1 ± 0.6
(0.35, 0.50)	13.5 ± 0.3	13.7 ± 0.4	13.9 ± 0.4	14.3 ± 0.5
(0.50, 0.60)	13.5 ± 0.3	13.8 ± 0.3	14.0 ± 0.4	14.4 ± 0.6



**Figure 12.** Mean mass of the clusters (stars) and contaminants (diamonds) in redshift and richness bins predicted from our posteriors for the background error model (green), and the projection error model (blue). Full error bars show the 16th/84th percentile, while shaded error bars show the 2.5th/97.5th percentile. Overplotted as grey bands the mean masses reported by Abbott et al. (2020) from stacked weak lensing (empty boxes showing the  $2\sigma$  uncertainty).

The scatter around that relation was  $\sigma_{\ln \lambda | M_{\text{SZE}}} = 0.266 \pm 0.017$ . Investigating the relation between the DES-redMaPPer richness and the temperature in 58 archival *Chandra* observations and 110 *XMM* observations, Farahi et al. (2019b) find  $\sigma_{\ln \lambda | M} = 0.20^{+0.10}_{-0.08}$ . Both of these measurements are in good agreement with our results, for instance  $\sigma_{\lambda} = 0.22 \pm 0.06$ .

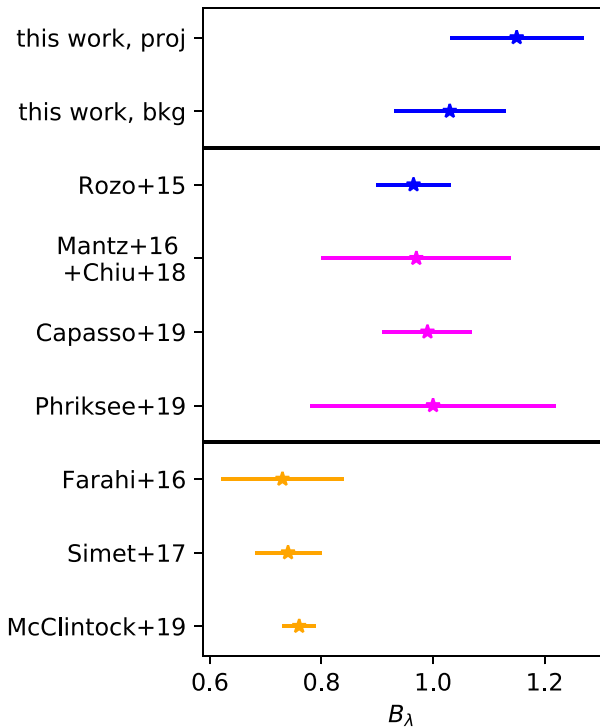
### 5.1.2 Mass trend of richness

Our estimates of the mean contaminant mass are sourced mainly by the mismatch between our prediction for the mean mass in richness–redshift bins and its measurement through stacked weak lensing by McClintock et al. (2019). Because the contamination fraction is larger at lower richnesses (see Fig. 9), this contamination – if ignored in a scaling relation analysis – can manifest itself as an apparent difference in the mass trend parameter  $B_{\lambda}$  (see also Fig. 11). Moreover, if different methods of cluster selection led to different levels of contamination, one might expect differences in the derived mass trend.

As an example, McClintock et al. (2019) find a mass trend of  $B_{\lambda, \text{McC19}} = 0.73 \pm 0.03$ . The tension with our results is representative of the tension with various other results reported in recent years (see McClintock et al. 2019; Capasso et al. 2019b; Bleem et al. 2020, for summaries). Examination of these results indicates a tendency for the mass trends in optically selected samples to be shallower than those derived using samples that have ICM based selection.

In the following and in Fig. 13 we present a selection of those results. In the class of works based on ICM selection, we point the reader to (1) Rozo et al. (2015) cross calibrating of *SDSS*–redMaPPer with clusters selected via SZE in the first all sky Planck survey that found  $B_{\lambda} = 0.965 \pm 0.067$ ; (2) Mantz et al. (2016) determining  $\lambda \sim M_{\text{gas}}^{0.75 \pm 0.12}$  on a sample *ROSAT* selected cluster, where combining this with the scaling  $M_{\text{gas}} \sim M^{1.29 \pm 0.09}$  for the gas mass content of SPT selected clusters (Chiu et al. 2018) results in  $B_{\lambda} = 0.97 \pm 0.17$ . Note here that this inference is based upon the relation between halo mass and gas mass. It is generally accepted from an empirical and theoretical perspective that the gas mass fraction of low-mass systems is lower than that for high mass systems; see, e.g. Mohr, Mathiesen & Evrard (1999) or the discussion in Bulbul et al. (2019, Section 5.2.2.); (3) a dynamical analysis of *ROSAT* selected, *SDSS*–redMaPPer confirmed cluster by Capasso et al. (2019b), reports  $B_{\lambda} = 0.99 \pm 0.08$ ; and (4) a stacked weak lensing analysis of the same sample of *ROSAT* selected, *SDSS*–redMaPPer confirmed clusters by Phriksee et al. (2020), reports  $B_{\lambda} = 1.00 \pm 0.22$ .

On the other hand, constraints obtained on purely optically selected samples include: (1) stacked spectroscopic analysis of *SDSS*–redMaPPer cluster by Farahi et al. (2016), which finds  $B_{\lambda} = 0.76 \pm 0.11$ ; (2) stacked weak lensing of *SDSS*–redMaPPer clusters by Simet et al. (2017), with  $B_{\lambda} = 0.74 \pm 0.06$ ; and (3) stacked weak lensing of DES–redMaPPer clusters by McClintock et al. (2019) with  $B_{\lambda} = 0.73 \pm 0.03$ .



**Figure 13.** Comparison of the constraints on the redMaPPer-richness–mass relation mass trend or power-law slope from different works to our own results. The colour coding represents the selection method used: blue for SZE, magenta for X-ray, and yellow for solely optical selection. The discrepancy between the results based on ICM selection (SZE and X-rays) and optical selection can clearly be seen. This difference may be due to low richness sample contamination in analyses based on optically selected samples.

Note that the stacked analysis of the CMB lensing signal around DES-redMaPPer clusters by Baxter et al. (2018) and Raghunathan et al. (2019a, b) currently does not constrain the mass trend of richness, and we do not include works that use number counts together with a fixed cosmology to constrain the mass slope (e.g. Murata et al. 2018), because the choice of cosmology can affect the mass slope. We also exclude from this discussion the study of the DES redMaPPer number counts and the large-scale auto- and cross-correlations between galaxies, clusters, and cosmic shear by To et al. (2021), as it is unclear if the mass information presented therein comes from the halo bias–mass relation, or from the combination of the number counts and the cosmological constraints from the auto- and cross-correlations.

Because the full cosmological analyses of the DES-redMaPPer sample (Abbott et al. 2020) and the *SDSS*-redMaPPer sample (Costanzi et al. 2019b) use the mass calibration derived by McClintock et al. (2019) and Simet et al. (2017), respectively, we do not discuss them separately in this section.

Noticeably, constraints on ICM selected samples are in good agreement with our own results and generally indicate  $B_\lambda \sim 1$ , whereas, results based in stacking methods of optically selected samples suggest  $B_\lambda \sim 0.75$ , as can be clearly seen in Fig. 13. The redMaPPer sample contamination presented here could be the underlying driver of the apparent tension in these mass trend measurements. Specifically, a contamination fraction which increases towards low richnesses would increase the  $B_\lambda$  value inferred from the (stacked) WL data. The mass trend is of special importance to

optical cluster cosmology; Abbott et al. (2020) identified this trend as the single most important source of the large tensions between redMaPPer based cluster cosmology constraints and those from other cosmological probes (including cosmological constraints from ICM selected samples). Specifically, a slope around 0.9 would make the number counts consistent with cosmological constraints from the shear and red galaxy auto- and cross-correlations (Abbott et al. 2018).

## 5.2 Impact of excess incompleteness in the SZE signal

In this work, we implicitly did not explicitly include the impact of radio and dust emission on the SZE signal. We shall argue in the following why this assumption is justified and is not expected to change our inferred low-mass contamination fraction. Two possible regimes of radio and dust emission need to be distinguished: (1) emission that is weak compared to the overall SZE signal of a halo, leading to an alteration of the SZE signal–mass relation, and (2) strong radio or dust emission that wipes out the cluster SZE signature and leads to excess incompleteness in the SZE.

The first form of SZE contamination is implicitly included in this work. Our priors on the SZE scaling relation parameters have been derived empirically for the total SZE signal (Bocquet et al. 2019). Consequently, our priors include the impact of radio and dust emission, as long as these emissions do not suppress a majority of the haloes SZE signal. The frequency of the latter has been studied by Gupta et al. (2017) in the case of radio emission. In that work, the authors found that at the mass range we are interested in, the resulting SZE excess incompleteness is  $< 5$  per cent. That means that at most 5 per cent of the SPT-SZ cluster that should have been detected are lost due to radio emission. The implications on this work are that the predicted SPT confirmation fraction of redMaPPer objects  $f(\text{SPT}|\hat{\lambda}, z)$  could at worst be 0.95 times the value that we computed assuming no excess SZE incompleteness from radio emission. Comparison to Fig. 8 shows that this is insufficient to describe the difference between the predicted and the measured confirmation fraction. Given that we interpret this difference as the presence of low-mass contaminants, our estimation of the low-mass contamination fraction is practically unaltered even by the worst case radio emission scenarios.

## 5.3 Comments on the contamination

Previously, we describe how our constraints on redMaPPer sample contamination provide a natural explanation for the tendency of analyses of optically selected samples to lead to shallower mass trends. Here, we investigate the expected physical properties of the contaminants. Direct measurement of contamination in the richness range  $20 < \hat{\lambda} < 40$  lies below the sensitivity of current wide area SZE based cluster surveys (Bleem et al. 2015), although this is a regime opening up to improved RASS+DES samples such as MARD-Y3 (Klein et al. 2019) and smaller but deeper SPT surveys (Huang et al. 2020).

However, the physical properties of the contaminants can be investigated using structure formation simulations. Barnes et al. (2017) studied the observable cluster properties in zoom-in hydrodynamical simulations of 30 galaxy cluster, using the *EAGLE* galaxy formation formalism. They found that galaxy groups of a mass  $\sim 3 \times 10^{13} M_\odot$  should have a soft X-ray ([0.5–2] keV band) luminosity of between  $\sim 2 \times 10^{42}$  and  $\sim 1 \times 10^{43} \text{ erg s}^{-1}$  and ICM temperatures of  $\sim 1 \text{ keV}$ . Given sufficient X-ray sensitivity, we thus expect to see diffuse X-ray emission from the ICM of the contaminant systems. For instance, a

follow up of *SDSS*-redMaPPer objects in the redshift range  $0.08 < z < 0.12$  with  $20 < \hat{\lambda} < 30$  by (von der Linden et al., in preparation) with the X-ray telescope *Swift* will shed light on the luminosity distribution at low richnesses. Analysis of that data could potentially test our findings. Future prospects for detection with *eROSITA* are discussed below (cf. Section 5.5).

According to the analysis of the number of satellite galaxies in haloes performed by Anbajagane et al. (2020) on three suites of independent cosmological simulations of galaxy formation in a cosmological context (BAHAMAS + MACSIS, TNG300 of the IllustrisTNG suite, and Magneticum Pathfinder, all implementing different sub-grid physics prescriptions), our contaminants would have from 0 to 12 satellite galaxies above a stellar mass of  $10^{10} M_{\odot}$ . It is unclear which fraction of those galaxies would appear as red sequence galaxy and thus be picked up by redMaPPer. Nevertheless, one would need to invoke projection effects to justify that a fraction of them might have richnesses of  $>20$ . This is, however, at odds with the projection effect model we employed in this work (Costanzi et al. 2019a) which has been calibrated for clusters with intrinsic richnesses as low as 5, and was set-up to account exactly for this effect. Given that we find that the projection effect model has very similar results compared to the background only model, this would imply that either the simulations or the projection model is inaccurate. That is, either contaminants must have many more galaxies than expected for the mass we estimate or the projection effects must be much stronger.

The projection model is derived assuming two main components: a redshift kernel that reflects the width of the red-sequence and a realization of the distribution of red galaxy in an  $N$ -body simulation. The first component can be reliably derived from spectroscopic studies of cluster galaxies and the DES photometric properties. The latter was simulated by assigning richnesses to haloes based on a richness–mass relation previously derived from stacked WL studies. A possible issue is that this relation might be biased by the presence of projection and selection effects. Another possible caveat is that the distribution of red galaxies as a function of local density in the simulation used to determine the projection model was not accurate outside of high-mass haloes. Such a mis-estimation of the density of red galaxies might bias the strength of projection effects. Empirical validation of the red galaxies distribution in dense environments predicted by the simulation, e.g. by comparing to measurements of the small-scale bias of red galaxies, might improve our confidence in the accuracy of the projection model. Alternatively, direct empirical confirmation of the projection model is currently undertaken by direct spectroscopic studies of *SDSS*-redMaPPer objects in the redshift range  $0.08 < z < 0.12$  (Myles et al. 2020).

An effect that is degenerate with contamination as discussed in this work is the possibility that the weak lensing signal and the richness of clusters at the same mass are biased in the same (or opposite) direction by the same physical process (see for instance Angulo et al. 2012, for simulation based work). This would lead to correlated intrinsic scatter of the WL signal and of the richness (not unlike the correlated scatter among richness and SZE signal we discuss in Section 4.1.2). In the case of a correlation of the scatter between WL signal and richness physical, intuition suggests that miscentring, halo triaxiality (Dietrich et al. 2014), and project effects (Sunayama et al. 2020) might lead to the correlation coefficient to be positive, as red galaxies trace dark matter tightly. This signal was already detected in simulations as ‘weak lensing selection bias’ by Abbott et al. (2020), but proved insufficient to account for the discrepancy in mass trends with respect to ICM based studies, as the correction did not show any strong richness trend. It is furthermore included in the mass estimates

we used in this work, suggesting that it cannot account for the signal that we interpret as contamination.

#### 5.4 Implications for redMaPPer cluster cosmology

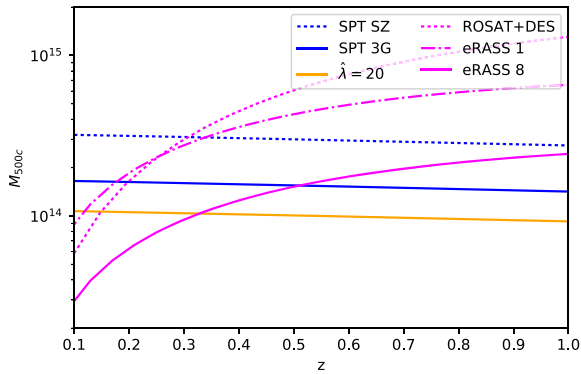
The implications of our result on the redMaPPer cluster cosmology which jointly fits for the redMaPPer number counts and stacked WL signals (such as the analysis carried out by Abbott et al. 2020), deserve some further discussion. First, our work implies that the richness–mass model used in that work, and derived by Costanzi et al. (2019a), is an incomplete description of the actual scatter between richness and mass. We investigate the possibility that the unresolved low-richness systematic advocated in Abbott et al. (2020) is related to such flawed modelling of the scaling relation. Specifically, we show that stronger projection effects are to be plausibly expected at low richness, as highlighted by our findings of a sizeable low mass population in the low richness range. We describe the possible reasons of this mis-calibration, and outline how the richness–mass modelling can be calibrated empirically (cf. Section 5.3).

Furthermore, we anticipate that using our constraints on the contamination fraction and mean contaminants mass as priors for a combined analysis of cluster abundance and stacked WL is likely to result in a strong deterioration of the cosmological constraints. This is due to the fact that our inferred posteriors on the contamination fraction and mean contaminants masses are very broad. Especially allowing for the mean contaminants mass in each redshift–richness bin to vary within our posterior range is likely to strongly dilute the mass information gained from the stacked WL analysis, resulting in a weaker cosmological inference from the number counts. None the less, we demonstrated in Section 4.2.1 that such weaker cosmological constraints will be consistent with the measurements for SPT-SZ number counts (Bocquet et al. 2019) and the auto- and cross-correlations of galaxies and cosmic shear in DES-Y1 data (Abbott et al. 2018).

#### 5.5 Alternative explanations and future prospects

A possible question that we left unanswered is whether the low-richness bins are actually populated by two populations with distinct mass distributions or rather by one single population with a significantly larger intrinsic scatter in mass. The latter option cannot be excluded from our analysis of the SPT-SZ confirmation fraction or the cross-matched sample, because the SPT-SZ selection only selects the mid- and high-mass portion of the mass distribution at any richness. This is precisely what the SPT-SZ confirmation fraction measures. Inspecting Fig. 8 one sees that less than a tenth of the objects with richness  $\hat{\lambda} = 40$  are detected. If the intrinsic scatter in richness were to increase strongly with decreasing richness or mass, this would not be visible in any of our observables, because we only observe the massive systems in the distribution of masses at richness  $\hat{\lambda} = 40$ . The location of the mean of the distribution is thus degenerate with the width of the distribution (the manifestation in cluster studies of the Malmquist bias). As such, we cannot exclude that at low richnesses the population we labelled ‘clusters’ and the population we labelled ‘contaminants’ actually merge into a single population with large intrinsic scatter. Current constraints on the scatter (cf. Section 5.1.1) are only applicable to the higher richness case.

This situation is, however, going to change in the next years as deeper SZE and X-ray surveys begin operations. In the SZE regime, this can already be seen by the lower mass limit in the SPTpol survey (Huang et al. 2020). Further improvements are expected by the third generation SPT camera currently performing the SPT-3G survey over



**Figure 14.** Current (dotted lines) and future (solid and dash–dotted) mass limits in magenta for X-ray surveys and in blue for SZE surveys. Yellow shows the mass corresponding to  $\hat{\lambda} > 20$ . Deeper photometric surveys do not lower the mass limit of optical selection but enable detections to higher redshift. As can be seen in the projected mass limits, *eROSITA*, and SPT-3G will extend the ICM leverage on low-mass systems significantly.

1500 deg<sup>2</sup> (Benson et al. 2014). This will lower the limiting mass of SZE detection by a factor of 2 compared to SPT-SZ. In X-rays, the recent ‘first light’<sup>2</sup> of the *eROSITA*<sup>3</sup> X-ray telescope (Predehl et al. 2010; Merloni et al. 2012) on board the Russian ‘Spectrum-Roentgen-Gamma’ satellite will dramatically improve the sensitivity of X-ray cluster surveys. We present in Fig. 14, the mass limits for 40 counts in the first *eROSITA* All-Sky-Survey (eRASS 1, 0.5 yr of observation) and eighth *eROSITA* All-Sky-Survey (eRASS 8, 4 yr of observation) following the prediction by Grandis et al. (2019). This would allow us to follow up optically selected objects up to  $z \sim 0.35$  individually and to much higher redshifts in stacks of lower richness objects. The presence or absence of X-ray emission in *eROSITA* from optically selected objects will be a powerful tool to extend the kind of analysis presented in this work to lower richnesses. It will most likely help to discriminate the different scenarios we identified earlier.

## 6 CONCLUSIONS

In this work, we empirically validate the redMaPPer selected DES-Y1 survey by cross calibration with SPT-SZ selected clusters. We first limit ourselves to the high-richness regime ( $\hat{\lambda} > 40$ ) to avoid optical incompleteness in the SPT confirmation. We produce a matched sample by positional matching between the redMaPPer ( $\hat{\lambda} > 40$ ) objects and the SPT-SZ selected clusters (Bleem et al. 2015). Of the 1005 redMaPPer selected cluster with measured richness  $\hat{\lambda} > 40$  in the joint footprint, 207 are confirmed by SPT-SZ. On the matched sample, we model the distribution in SZE-signal, richness, and redshift with a Bayesian cluster population model. The free parameters of this model, the parameters controlling the scaling between richness and mass, and the scatter around this relation, are constrained from our analysis. We adopt priors from previous SPT studies on the parameters on the SZE signal–mass relation, effectively transferring the vetted SPT mass calibration to the redMaPPer richness–mass relation.

In an attempt to explore the impact of projection effects on the richness–mass relation, we employ two different error models: the first uses the error bars reported in the redMaPPer catalogue and

accounts only for the photometric uncertainty in the background subtraction, while the second includes projection effects calibrated on simulations (Costanzi et al. 2019a). Our cross calibration of the richness–mass relation and the scatter around it is not significantly affected by the optical error model. Furthermore, our derived parameters are consistent with those reported by previous cross-calibration studies (e.g. Bleem et al. 2020).

We then turn to exploiting the information contained in the fact that some redMaPPer ( $\hat{\lambda} > 40$ ) in the joint SPT-DES Y1 footprint have been confirmed by SPT-SZ and some have not. Taking into account the relative SPT field depth at the redMaPPer positions, we employ the mass information contained in the richness and the SPT selection function to predict the detection probability by SPT for each redMaPPer cluster. We explore a model of the probability of SPT detection that accounts for possible redMaPPer contamination with respect to the scatter around the mean relation. Comparing these detection probabilities with the actual occurrence of matches constrains the contamination fraction.

Our investigation of the contamination fraction indicates that a model with a large contamination fraction of up to 50 per cent for richness 40 and a strong trend increasing to lower richness provides a better fit to the SPT confirmation fraction of redMaPPer clusters. Furthermore, the prediction of the redMaPPer number counts down to richnesses  $\hat{\lambda} > 20$  with a richness dependent contamination fraction are a better description of the measured number of clusters when compared to the case without contamination. This provides both internal and external evidence for considerable contamination, increasing to lower richnesses. Adopting our posterior on the contamination, our prediction for the clusters mean mass and the weak lensing constraints on mean mass in richness–redshift bins, we predict that the contaminants have mean masses of  $\sim 3 \times 10^{13} M_{\odot}$  ( $\sim 5 \times 10^{13} M_{\odot}$ ) for the range  $20 < \hat{\lambda} < 30$  ( $30 < \hat{\lambda} < 45$ ). The presence of group scale contaminants might be an explanation for the fact that in the cosmological study of redMaPPer objects (Abbott et al. 2020) the WL mass measurements of richness-selected samples are biased low at low richness.

We discuss possible explanations for why analyses of the richnesses of ICM selected cluster samples tend to produce different (steeper) mass trends than analyses that rely on optically selected cluster samples. While this effect might indeed be due to the larger contamination of purely optically selected samples, we also discuss that to date, the mass sensitivity of ICM selection does not extend to the mass range spanned by the richnesses  $20 < \hat{\lambda} < 40$ . As such, we cannot exclude alternative explanations for the tension between our relation and the stacked weak lensing masses. We then highlight how upcoming X-ray surveys like *eROSITA*, and SZE surveys like SPT-3G will improve the ICM based selection sensitivities to probe the mass regime associated with these lower richness systems, enabling improved tests of the results presented here.

## ACKNOWLEDGEMENTS

We acknowledge financial support from the MPG Faculty Fellowship program, the DFG Cluster of Excellence ‘Origin and Structure of the Universe’, the new DFG cluster ‘Origins’ and the Ludwig-Maximilians-Universität Munich. AS is supported by the ERC-StG ‘ClustersXCosmo’ grant agreement 716762, by the FARE-MIUR grant ‘ClustersXEuclid’ R165SBKTM, and by the INFN INDARK grant Numerical computations in this work relied on the PYTHON packages NUMPY (Travis 2006) and SCIPY (Jones et al. 2001). The plots were produced using the package matplotlib (Hunter 2007).

<sup>2</sup><http://www.mpe.mpg.de/7362095/news20191022>

<sup>3</sup><http://www.mpe.mpg.de/eROSITA>



Posterior samples have been drawn from the likelihood functions and the priors using `emcee` (Foreman-Mackey et al. 2013).

This work was performed in the context of the South-Pole Telescope scientific program. SPT is supported by the National Science Foundation through grant PLR-1248097. Partial support is also provided by the NSF Physics Frontier Center grant PHY-0114422 to the Kavli Institute of Cosmological Physics at the University of Chicago, the Kavli Foundation and the Gordon and Betty Moore Foundation grant GBMF 947 to the University of Chicago. This work is also supported by the U.S. Department of Energy. AAS acknowledges support by US NSF AST-1814719.

This paper has gone through internal review by the DES collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2015-71825, ESP2015-66861, FPA2015-68048, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) e-Universe (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

## DATA AVAILABILITY

The data underlying this article are proprietary to the Dark Energy Survey Collaboration and the South Pole Telescope collaboration. Where applicable, please refer to the reference and links provided in this work to obtain the data.

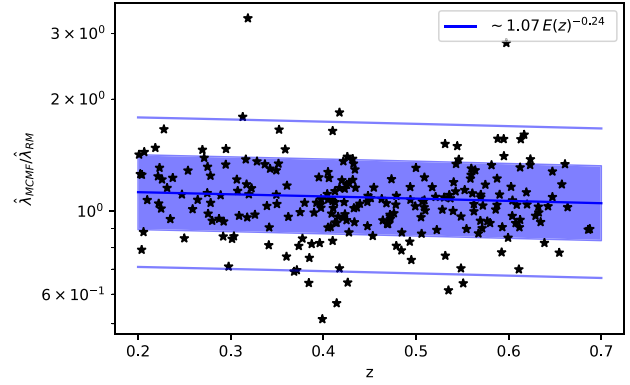
## REFERENCES

- Abbott T. M. C. et al., 2018, *Phys. Rev. D*, 98, 043526  
 Abbott T. M. C. et al., 2020, *Phys. Rev. D*, 102, 023509  
 Agüena M., Lima M., 2018, *Phys. Rev. D*, 98, 123529  
 Allen S. W., Evrard A. E., Mantz A. B., 2011, *ARA&A*, 49, 409  
 Anbajagane D., Evrard A. E., Farahi A., Barnes D. J., Dolag K., McCarthy I. G., Nelson D., Pillepich A., 2020, *MNRAS*, 495, 686  
 Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, *MNRAS*, 426, 2046  
 Barnes D. J. et al., 2017, *MNRAS*, 471, 1088  
 Baxter E. J. et al., 2018, *MNRAS*, 476, 2674  
 Benson B. A. et al., 2013, *ApJ*, 763, 147  
 Benson B., SPT T., SPTpol, SPT-3G Collaborations, 2014, American Astronomical Society Meeting Abstracts #223. p. 135.02  
 Bhattacharya S., Habib S., Heitmann K., Vikhlinin A., 2013, *ApJ*, 766, 32  
 Bleem L. E. et al., 2015, *ApJS*, 216, 27  
 Bleem L. E. et al., 2020, *ApJS*, 247, 25  
 Bocquet S. et al., 2015, *ApJ*, 799, 214  
 Bocquet S. et al., 2019, *ApJ*, 878, 55  
 Böhringer H. et al., 2001, *A&A*, 369, 826  
 Böhringer H. et al., 2004, *A&A*, 425, 367  
 Bulbul E. et al., 2019, *ApJ*, 871, 50  
 Capasso R. et al., 2019a, *MNRAS*, 482, 1043  
 Capasso R. et al., 2019b, *MNRAS*, 486, 1594  
 Carlstrom J. et al., 2011, *PASP*, 123, 568  
 Chiu I. et al., 2018, *MNRAS*, 478, 3072  
 Clerc N. et al., 2014, *MNRAS*, 444, 2723  
 Cohn J. D., Evrard A. E., White M., Croton D., Ellingson E., 2007, *MNRAS*, 382, 1738  
 Costanzi M. et al., 2019a, *MNRAS*, 482, 490  
 Costanzi M. et al., 2019b, *MNRAS*, 488, 4779  
 Costanzi M. et al., 2020, *Phys. Rev. D*, 103, 043522  
 Dark Energy Survey Collaboration, 2016, *MNRAS*, 460, 1270  
 de Haan T. et al., 2016, *ApJ*, 832, 95  
 Dietrich J. P. et al., 2014, *MNRAS*, 443, 1713  
 Doux C. et al., 2021, *MNRAS*, 503, 2688  
 Drlica-Wagner A. et al., 2018, *ApJS*, 235, 33  
 Farahi A., Evrard A. E., Rozo E., Rykoff E. S., Wechsler R. H., 2016, *MNRAS*, 460, 3900  
 Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2018, *MNRAS*, 478, 2618  
 Farahi A. et al., 2019a, *Nat. Commun.*, 10, 2504  
 Farahi A. et al., 2019b, *MNRAS*, 490, 3341  
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306  
 García R., Rozo E., 2019, *MNRAS*, 489, 4170  
 Grandis S., Mohr J. J., Dietrich J. P., Bocquet S., Saro A., Klein M., Paulus M., Capasso R., 2019, *MNRAS*, 488, 1738  
 Grandis S. et al., 2020, *MNRAS*, 498, 771  
 Gupta N. et al., 2017, *MNRAS*, 467, 3737  
 Hennig C. et al., 2017, *MNRAS*, 467, 10515  
 Hilton M. et al., 2018, *ApJS*, 235, 20  
 Hilton M. et al., 2021, *ApJS*, 253, 3  
 Huang N. et al., 2020, *AJ*, 159, 110

- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Jones E. et al., 2011, SciPy: Open source scientific tools for Python. <http://www.scipy.org/>
- Kerscher M., Weller J., 2019, *SciPost Physics Lecture Notes*, 9. Prof. J.-S. Caux, Amsterdam
- Klein M. et al., 2018, *MNRAS*, 474, 3324
- Klein M. et al., 2019, *MNRAS*, 488, 739
- Mantz A. B. et al., 2015, *MNRAS*, 446, 2205
- Mantz A. B. et al., 2016, *MNRAS*, 463, 3582
- McClintock T. et al., 2019, *MNRAS*, 482, 1352
- Melin J. B., Bartlett J. G., Delabrouille J., 2005, *A&A*, 429, 417
- Melin J. B., Bartlett J. G., Delabrouille J., 2006, *A&A*, 459, 341
- Merloni A. et al., 2012, preprint ([arXiv:1209.3114](https://arxiv.org/abs/1209.3114))
- Mohr J. J., Mathiesen B., Evrard A. E., 1999, *ApJ*, 517, 627
- Mortonson M. J., Hu W., Huterer D., 2011, *Phys. Rev. D*, 83, 023015
- Mulroy S. L. et al., 2019, *MNRAS*, 484, 60
- Murata R., Nishimichi T., Takada M., Miyatake H., Shirasaki M., More S., Takahashi R., Osato K., 2018, *ApJ*, 854, 120
- Murata R. et al., 2019, *PASJ*, 71, 107
- Myles J. T. et al., 2020, preprint ([arXiv:2011.07070](https://arxiv.org/abs/2011.07070))
- Nicola A., Amara A., Refregier A., 2019, *J. Cosmol. Astropart. Phys.*, 2019, 011
- Pacaud F. et al., 2006, *MNRAS*, 372, 578
- Phriksee A., Jullo E., Limousin M., Shan H., Finoguenov A., Komonjinda S., Wannawichian S., Sawangwit U., 2020, *MNRAS*, 491, 1643
- Planck Collaboration XXXII, 2015, *A&A*, 581, A14
- Planck Collaboration XXVII, 2016, *A&A*, 594, A27
- Predehl P. et al., 2010, in Arnaud M., Murray S. S., Takahashi T., eds, Proc. SPIE Conf. Ser. Vol. 7732, Space Telescopes and Instrumentation 2010: Ultraviolet to Gamma Ray. SPIE, Bellingham, p. 77320U
- Raghunathan S. et al., 2019a, *Phys. Rev. Lett.*, 123, 181301
- Raghunathan S. et al., 2019b, *ApJ*, 872, 170
- Reichardt C. L. et al., 2013, *ApJ*, 763, 127
- Rigault M. et al., 2020, *A&A*, 644, A176
- Romer A. K., Viana P. T. P., Liddle A. R., Mann R. G., 2001, *ApJ*, 547, 594
- Rozo E., Rykoff E. S., 2014, *ApJ*, 783, 80
- Rozo E., Rykoff E. S., Bartlett J. G., Melin J.-B., 2015, *MNRAS*, 450, 592
- Rykoff E. S. et al., 2014, *ApJ*, 785, 104
- Rykoff E. S. et al., 2016, *ApJS*, 224, 1
- Saro A. et al., 2015, *MNRAS*, 454, 2305
- Saro A. et al., 2017, *MNRAS*, 468, 3347
- Sereno M. et al., 2020, *MNRAS*, 497, 894
- Simet M., McClintock T., Mandelbaum R., Rozo E., Rykoff E., Sheldon E., Wechsler R. H., 2017, *MNRAS*, 466, 3103
- Song J., Mohr J. J., Barkhouse W. A., Warren M. S., Dolag K., Rude C., 2012, *ApJ*, 747, 58
- Stanek R., Rasia E., Evrard A. E., Pearce F., Gazzola L., 2010, *ApJ*, 715, 1508
- Story K. T. et al., 2013, *ApJ*, 779, 86
- Sunayama T. et al., 2020, *MNRAS*, 496, 4468
- Sunyaev R. A., Zeldovich Y. B., 1972, *Comments Astrophys. Space Phys.*, 4, 173
- To C. et al., 2021, *Phys. Rev. Lett.*, 126, 141301
- Travis E., 2006, *A Guide to NumPy*. Trelgol Publishing, USA
- Truong N. et al., 2018, *MNRAS*, 474, 4089
- Vanderlinde K. et al., 2010, *ApJ*, 722, 1180
- Vikhlinin A., McNamara B. R., Forman W., Jones C., Quintana H., Hornstrup A., 1998, *ApJ*, 502, 558
- Wu H.-Y., Evrard A. E., Hahn O., Martizzi D., Teyssier R., Wechsler R. H., 2015, *MNRAS*, 452, 1982

## APPENDIX A: OPTICAL COMPLETENESS OF THE SPT SAMPLE

As discussed above, we employ an SPT sample that has been confirmed with MCMF by imposing a cut  $f_{\text{cont}} < 0.1$ , in the probability of random superpositions between the SZE candidate and



**Figure A1.** Ratio between the MCMF richness  $\hat{\lambda}_{\text{MCMF}}$  employed in the optical confirmation of SPT-SZ cluster to the redMaPPer richness  $\hat{\lambda}_{\text{RM}}$  as a function redshift for the matched sample. In blue, the mean relation between the two with the intrinsic scatter show by the shaded ( $1\sigma$ ) and transparent lines ( $2\sigma$ ). Different centering and apertures lead to intrinsic scatter, a deviation from unity and redshift trend.

physically unassociated optical structures. As discussed previously (Klein et al. 2019, in preparation; Grandis et al. 2020), this cut is equivalent to a redshift dependent cut in the MCMF richness  $\hat{\lambda}_{\text{MCMF}} > \lambda_{\text{min}}(z)$ . Note that the richness  $\hat{\lambda}_{\text{MCMF}}$  extracted by MCMF is not identical to the redMaPPer richness  $\hat{\lambda} = \hat{\lambda}_{\text{RM}}$ , because MCMF employs a prior from the SZE candidate on the position and the aperture. The ratio between the two richnesses on the matched sample is shown in Fig. A1. When fitting this relation, we find

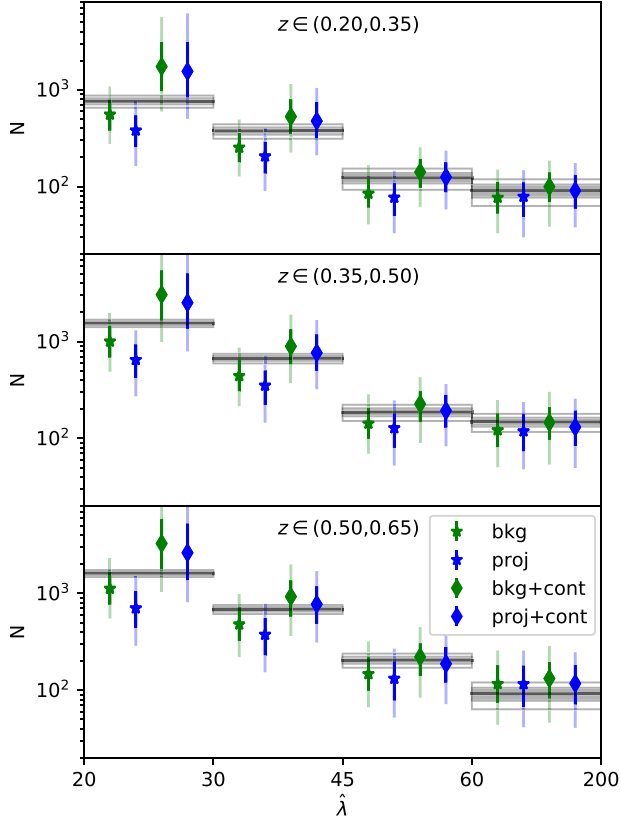
$$\frac{\hat{\lambda}_{\text{MCMF}}}{\hat{\lambda}_{\text{RM}}} = (1.07 \pm 0.02) \left( \frac{E(z)}{E(0.6)} \right)^{-0.24 \pm 0.02}, \quad (\text{A1})$$

with intrinsic lognormal scatter  $0.23 \pm 0.01$ .

Using this relation, we estimate that objects with  $\hat{\lambda}_{\text{RM}} > 40$  are always more than  $2\sigma$  above the minimal MCMF richness for all redshifts we consider. This implies at least 97.5 per cent completeness. Every redMaPPer- ( $\hat{\lambda} > 40$ ) object therefore (almost) certainly makes it past the SPT optical confirmation. In the case of the matched sample, where all objects have  $\hat{\lambda} > 40$ , we therefore can safely ignore optical incompleteness in SPT. The same holds for the study of the SPT detection of redMaPPer- ( $\hat{\lambda} > 40$ ), which we model as solely dependent on the SZE signal. We intentionally omit studying the redMaPPer detection probability of SPT objects. The probability of not finding an optically confirmed SPT cluster is given by the probability of  $\hat{\lambda}_{\text{MCMF}} > \lambda_{\text{min}}(z)$  and  $\hat{\lambda} < 40$  at the clusters SZE signal and redshift, and thus depends on the optical incompleteness of the SPT sample.

## APPENDIX B: COSMOLOGICAL DEPENDENCE OF THE PREDICTED REDMAPPER NUMBER COUNTS

To test the cosmological sensitivity of agreement of our contamination fraction with the redMaPPer number counts, we predict the latter also assuming the cosmology derived from the auto- and cross-correlations of cosmic shear and galaxies in the DES-Y1 (Abbott et al. 2018). Consequently, we used the priors  $\Omega_{\text{M}} = 0.267 \pm 0.024$  and  $S_8 = \sigma_8(\Omega_{\text{M}}/0.3)^{0.2} = 0.773 \pm 0.023$ . The resulting predicted redMaPPer number counts are shown in Fig. B1. Our qualitative assessment that the lower limit of the contamination fraction is consistent with the redMaPPer number counts holds also with these different cosmological priors.



**Figure B1.** The number of redMaPPer objects in bins of richness and redshift are shown as predictions (coloured points) derived from our richness–mass relation constraints and cosmology from the auto- and cross-correlations of cosmic shear and galaxies in the DES-Y1 data (Abbott et al. 2018) for the different optical error models (blue projection, green background). Stars denote predictions without contamination, while diamonds denote predictions with contamination. Overplotted in grey is the number of redMaPPer objects with the associated statistical uncertainty. Also in this cosmology, the model with contamination fraction is not excluded by the number counts data.

## APPENDIX C: LIST OF VARIABLES

In Table C1, we summarize the variables used in this work with units and explanation.

**Table C1.** List of most relevant variable names used in this work with definition and with units where applicable.

Symbol	Units	Explanation
$M$	$M_{\odot} h^{-1}$	Halo mass for spherical over-densities of 500 times the critical density of the universe
$z$		Redshift (no distinction between true and measured redshift is made)
$\lambda$		Intrinsic richness
$\hat{\lambda}$		Measured richness
$\zeta$		Intrinsic SPT-SZ signal to noise
$\xi$		Measured SPT-SZ signal to noise
$\gamma_f$		Effective SPT-SZ field depth
$\Omega_f$	$\text{deg}^2$	Area of the SPT-SZ field $f$
$A_{\text{SZE}, \lambda}$		Amplitude of the SZE signal/richness – mass relation
$B_{\text{SZE}, \lambda}$		Power-law index of the mass trend of the SZE signal/richness–mass relation
$C_{\text{SZE}, \lambda}$		Power-law index of the redshift trend of the SZE signal/richness–mass relation
$\sigma_{\text{SZE}, \lambda}$		Intrinsic scatter around the SZE signal/richness–mass relation
$\rho$		Correlation coefficient between the intrinsic scatters around the SZE signal and richness–mass relations
$i$		Index for the $i$ -th cluster (e.g. $\hat{\lambda}^i$ is the measured richness of the $i$ th cluster.)
$N$		Number of haloes/clusters
$\Omega_M$		Present day fractional matter density of the Universe with respect to the critical density
$\sigma_8$		Root mean square of present day matter fluctuation amplitudes at a scale of 8 Mpc/h
$\pi_c$		Contamination fraction
$j$		Index of the richness, redshift bin defined by the edges $(\hat{\lambda}_{-}^j, \hat{\lambda}_{+}^j)$ and $(z_{-}^j, z_{+}^j)$
$\bar{M}(j)$	$M_{\odot} h^{-1}$	Mean mass of the bin $j$
$\hat{M}_{\text{WL}}^j$	$M_{\odot} h^{-1}$	Measured mean WL mass in the bin $j$
$\delta \hat{M}_{\text{WL}}^j$	$M_{\odot} h^{-1}$	Measurement uncertainty on mean WL mass in the bin $j$
$\bar{M}_{\text{c}}^j$	$M_{\odot} h^{-1}$	Mean contaminants mass of the bin $j$

- <sup>1</sup>Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr 1, D-81679 Munich, Germany
- <sup>2</sup>Excellence Cluster ORIGINS, Boltzmannstr 2, D-85748 Garching, Germany
- <sup>3</sup>Max Planck Institute for Extraterrestrial Physics, Giessenbachstr, D-85748 Garching, Germany
- <sup>4</sup>INAF – Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy
- <sup>5</sup>IFPU – Institute for Fundamental Physics of the Universe, Via Beirut 2, I-34014 Trieste, Italy
- <sup>6</sup>Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy
- <sup>7</sup>INFN – National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy
- <sup>8</sup>Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP 05314-970, Brazil
- <sup>9</sup>Laboratório Interinstitucional de e-Astronomia – LInEA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ 20921-400, Brazil
- <sup>10</sup>Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA
- <sup>11</sup>CEA, Physics Department, Durham University, South Road, Durham DH1 3LE, UK
- <sup>12</sup>Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
- <sup>13</sup>CNRS, UMR 7095, Institut d’Astrophysique de Paris, F-75014 Paris, France
- <sup>14</sup>Institut d’Astrophysique de Paris, Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, F-75014 Paris, France
- <sup>15</sup>High Energy Physics Division, Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA
- <sup>16</sup>Kavli Institute for Cosmological Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
- <sup>17</sup>Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- <sup>18</sup>Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA
- <sup>19</sup>SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
- <sup>20</sup>Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain
- <sup>21</sup>Dpto. Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Spain
- <sup>22</sup>Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA
- <sup>23</sup>National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA
- <sup>24</sup>Institut de Física d’Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona) Spain
- <sup>25</sup>Institut d’Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain
- <sup>26</sup>Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain
- <sup>27</sup>Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA
- <sup>28</sup>Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ 20921-400, Brazil
- <sup>29</sup>Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Av. Complutense, 40, 28040, Madrid, Spain
- <sup>30</sup>Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India
- <sup>31</sup>Department of Astronomy/Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA
- <sup>32</sup>Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA
- <sup>33</sup>Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA
- <sup>34</sup>Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, NO-0315 Oslo, Norway
- <sup>35</sup>Department of Physics and Astronomy, University of Missouri, 5110 Rockhill Road, Kansas City, MO 64110, USA
- <sup>36</sup>Instituto de Física Teórica UAM/CSIC, Universidad Autonoma de Madrid, E-28049 Madrid, Spain
- <sup>37</sup>Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA
- <sup>38</sup>School of Physics, University of Melbourne, Parkville, VIC 3010, Australia
- <sup>39</sup>School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia
- <sup>40</sup>Department of Physics, The Ohio State University, Columbus, OH 43210, USA
- <sup>41</sup>Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
- <sup>42</sup>Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia
- <sup>43</sup>Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA
- <sup>44</sup>Centre for Gravitational Astrophysics, College of Science, The Australian National University, ACT 2601, Australia
- <sup>45</sup>The Research School of Astronomy and Astrophysics, Australian National University, ACT 2601, Australia
- <sup>46</sup>Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
- <sup>47</sup>George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA
- <sup>48</sup>Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA
- <sup>49</sup>Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain
- <sup>50</sup>Physics Department, 2320 Chamberlin Hall, University of Wisconsin-Madison, 1150 University Avenue Madison, WI 53706-1390, USA
- <sup>51</sup>Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
- <sup>52</sup>Department of Physics and Astronomy, University of Sussex, Pevensey Building, Brighton BN1 9QH, UK
- <sup>53</sup>School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK
- <sup>54</sup>Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
- <sup>55</sup>Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
- <sup>56</sup>Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
- <sup>57</sup>Fakultät für Physik, Universitäts-Sternwarte, Ludwig-Maximilians-Universität München, Scheinerstr 1, D-81679 München, Germany
- <sup>58</sup>Department of Physics, The Ohio State University, Columbus, OH 43210, USA

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.