We present Weak Gravitational Lensing measurements of a sample of 157 clusters within the Kilo Degree Survey (KiDS), detected with a thermal Sunyaev-Zel’dovich (SZ) signal by the Atacama Cosmology Telescope (ACT). Using a halo-model approach, we calibrated using X-ray observations, is biased with \( \frac{M_{\text{WL}}}{M_{\text{SZ}}} = (1 - b_{\text{SZ}}) \approx 0.65 \pm 0.05 \). Separating the sample into six mass bins, we find no evidence of a strong mass dependency for the mass bias, \((1 - b_{\text{SZ}})\). Adopting this ACT-KiDS SZ mass calibration would bring the Planck SZ cluster count into agreement with the counts expected from the Planck cosmic microwave background \(\Lambda\)CDM cosmological model, although it should be noted that the cluster sample considered in this work has a lower average mass \( M_{\text{SZ,uncorr}} = 3.64 \times 10^{14} M_\odot \) compared to the Planck cluster sample which has an average mass in the range \( M_{\text{SZ,uncorr}} = (5.5-8.5) \times 10^{14} M_\odot \), depending on the sub-sample used.

**Key words.** gravitational lensing: weak – large-scale structure of Universe – cosmology: observations
1. Introduction

The locations of galaxy clusters trace the position of peaks in the large-scale matter distribution. As a result, their properties – including number density, masses, baryon content and evolution – contain information about the growth of structure in our Universe and can constrain quantities such as the amplitude of matter fluctuations, $\sigma_8$, the matter density, $\Omega_m$, and the dark energy equation of state. The key statistic is their number density as a function of mass and redshift, $N(M,z)$. In practice, however, we cannot measure the mass directly and instead infer it from some observable quantity that correlates with mass (see for example, Allen et al. 2011, for a review).

Clusters can be observed across a range of wavelengths, resulting in a variety of mass proxies. For example, they emit X-rays from the hot intra-cluster gas, with the gas temperature scaling with mass (for example, Ebeling et al. 1998; Böhringer et al. 2004; Vikhlinin et al. 2009). In the optical and infrared, the cluster richness and stellar luminosity of the galaxies within the cluster also correlate with mass (for example, Gladders & Yee 2005; Rykoff et al. 2016). At millimetre wavelengths clusters are detected through the thermal Sunyaev-Zel’dovich (SZ) effect (Zel’dovich & Sunyaev 1969; Sunyaev & Zel’dovich 1972; Birkinshaw et al. 1984), which occurs when cosmic microwave background (CMB) photons are inverse Compton scattered off electrons in the hot cluster gas. In this case, the mass observable is the volume integrated intra-cluster medium (ICM) pressure, measured with the Compton–y parameter. Because the SZ effect surface brightness is independent of redshift, catalogues produced with arcminute-resolution SZ effect observations have a mass threshold that is almost constant with redshift, which makes them excellent tracers of the underlying matter distribution.

For all cluster detection methods, using clusters for cosmology is currently limited by the uncertainty in the mass-observable scaling relations which need to be calibrated. In the case of the SZ effect, a mass $M_{SZ}$ is usually derived from the Compton–y parameter by assuming the gas and density profiles estimated by Arnaud et al. (2010), from X-ray observations of local clusters, are in hydrostatic equilibrium (e.g. Hilton et al. 2021; Hasselfield et al. 2013; Planck Collaboration XXIV 2016).

Weak Gravitational Lensing provides a tool to calibrate biases in the SZ mass estimate, as it is sensitive to the total mass of the cluster, both dark and baryonic, and does not depend on its dynamical state (Hoekstra et al. 2013). Previous estimates of weak lensing masses have been made for SZ-selected clusters from the South Pole Telescope (SPT; McNamara et al. 2009; High et al. 2012; Schrabback et al. 2018; Dietrich et al. 2019; Stern et al. 2019), the Atacama Cosmology Telescope (ACT; Miyatake et al. 2013, 2019; Ji et al. 2014; Battaglia et al. 2016), and the Planck satellite (von der Linden et al. 2014; Hoekstra et al. 2015; Penna-Lima et al. 2017; Sereno et al. 2017; Medezinski et al. 2018; Herbonnet et al. 2020). Many of these studies have therefore focussed on measuring the mass bias, which quantifies the cluster mass fraction left unaccounted for in the $y$–$M$ scaling relation. The mass bias is defined as

$$1 - b_{SZ} = \frac{M_{SZ}}{M_{WL}},$$

where $M_{SZ}$ is the mass derived from the SZ Compton–y observable, assuming the X-ray derived $y$–$M$ relation from Arnaud et al. (2010), and $M_{WL}$ is the mass measured using weak gravitational lensing which we take to be the “true” mass. This definition of $M_{SZ}$ is not an “absolute” definition of the SZ mass, but is the one chosen for this analysis. There have been investigations into whether $(1 - b_{SZ})$ may depend on mass or redshift (Henson et al. 2017; Remazéilles et al. 2019; Herbonnet et al. 2020); current data are not yet conclusive given the uncertainties. Recent estimates of this bias from SZ clusters detected using Planck and ACT are reported in Medezinski et al. (2018) and Miyatake et al. (2019), with alternative parameterisations of the mass scaling reported from SPT in Bocquet et al. (2019).

In the Planck SZ cluster counts analysis (Planck Collaboration XXIV 2016) and the Planck primary CMB analysis (Planck Collaboration VI 2020) inconsistent results were obtained if hydrostatic equilibrium was assumed. This difference in results could be resolved by relaxing this assumption, yielding a mass bias of $0.62 \pm 0.03$, obtained by combining the primary CMB spectra and cluster counts likelihoods.

In the analysis presented here, we estimate the mass bias for the cluster sample from ACT using weak lensing data from the Kilo Degree Survey (KiDS; Kuijken et al. 2019; Giblin et al. 2021). The ACT experiment has been used to construct a sample of more than 4000 confirmed clusters over a 13,000 deg$^2$ region, measured during the 2008–18 observing seasons (Hilton et al. 2021). KiDS provides 774 deg$^2$ of weak lensing data in the area covered by ACT; we find 157 SZ-confirmed clusters in this region which are detected with an SZ signal-to-noise greater than 5. We estimate the mean mass of these clusters, using stacked lensing data, and the bias parameter. This allows us to probe the lower end of the SZ cluster mass range (1.72–10.4) $\times 10^{14} M_\odot$ with a larger sample size than was previously available. This study also provides a means to cross-check consistency of data from different weak lensing surveys which have survey areas in common with ACT.

The outline of this paper is as follows. We summarise the data and observable quantities in Sect. 2. Our results are presented in Sects. 3 and 4 with our conclusions following in Sect. 5. We assume a flat $\Lambda$CDM cosmology with $\Omega_m = 0.3$, $\Omega_{\Lambda} = 0.7$ and $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$ throughout, unless otherwise stated. We investigate the dependence of our result on fixing cosmology when inferring the mass, which is discussed in Sect. 3.

2. Data and observables

2.1. The Atacama Cosmology Telescope SZ-selected cluster sample

We use the ACT cluster catalogue described in Hilton et al. (2021), based on the ACT Data Release 5 (DR5) 2008–2018 coadded maps described in Naess et al. (2020). The cluster catalogue is constructed from coadded maps at 90 and 150 GHz (Naess et al. 2020). Clusters are identified using a suite of multi-frequency matched filters, and confirmed using imaging from the public large imaging and spectroscopic surveys and spectroscopic measurements reported in the literature (see Hilton et al. 2021, for details). From this optical data, the cluster redshifts are also determined and a cluster may therefore have either a spectroscopic or photometric redshift. The SZ signal-to-noise of the cluster detection is determined from a single fixed filter scale which simplifies the survey selection function. The fraction of false positives is 0.03 for clusters that are detected with an SZ signal-to-noise greater than 5 (shown in Fig. 6 of
Fig. 1. KiDS North and South survey areas are shown in light grey in the upper and lower panels, respectively. The positions of ACT clusters used in this analysis are indicated by each circle. The colour of the circles corresponds to cluster redshift, with yellow indicating high redshift and black indicating low redshift. The size of each circle scales with the SZ signal-to-noise, which is correlated with the cluster mass when the noise is the same everywhere, such that larger circles represent clusters with a greater SZ signal-to-noise than smaller circles.

Hilton et al. (2021) and rises to 0.34 for an SZ signal-to-noise cut of 4. The cluster locations are shown in Fig. 1, and the distribution of their SZ detection signal-to-noise, redshifts and SZ estimated masses, are shown in Fig. 2 in comparison to the full ACT sample. Figure 2 shows that the clusters found within the KiDS survey region are a good representation of the full ACT cluster sample (see Fig. 18 of Hilton et al. 2021). An additional selection, requiring the clusters to have a redshift between 0.1 and 0.9, is applied to the cluster sample, this is to ensure there are source galaxies behind the clusters to make the weak lensing measurements with (see Sect. 2.2).

Hilton et al. (2021) estimate the Compton–y parameter and associated SZ mass, $M_{SZ}$, for each cluster; we summarise their method here. The signal we observe, at a projected angle $\theta$ from the cluster centre, due to the SZ effect is a change in radiation intensity which can be expressed in terms of CMB temperature by

$$\frac{\Delta T(\theta)}{T_{CMB}} = f_{SZ} y(\theta),$$

where $y(\theta)$ is the line-of-sight integrated pressure profile. In the non-relativistic limit, $f_{SZ}$ only depends on the observed radiation frequency. In Hilton et al. (2021) a correction factor $f_{rel}$ is applied to include relativistic effects for gas temperature, as determined in Itoh et al. (1998).

The Compton–y signal is modelled using the universal pressure profile (Arnaud et al. 2010) which describes the intra-cluster gas pressure, $p$, as a function of scaled radius, $x = r/r_{500c}$ where $r_{500c}$ is the radius within which the density is 500 times the critical density $\rho_c$ at $z = 0$. This relation was modelled with a generalised Navarro-Frenk-White (NFW) profile from Nagai (2006) as

$$p(x) = \frac{p_0}{(c_{500} x)^\gamma [1 + (c_{500} x)^\alpha]^\beta/\alpha},$$

where $c_{500}$ is the concentration parameter and $\{\gamma, \alpha, \beta\}$ describe the slope at different radii. This includes any mass dependence in the profile shape and has been calibrated to X-ray observations of nearby clusters.
used to define the matched filter, which includes the effects of the beam (see Hasselfield et al. 2013, for full details). The location of the cluster is then defined as the peak of the SZ signal. An alternative definition of the cluster centre is the brightest cluster galaxy (BCG) which is expected to be a better tracer of the true centre compared to the SZ signal which is tracing the gas distribution. This is investigated further in Appendix A however in our main analysis we take the BCG as the cluster centre.

To obtain the cluster mass, the intrinsic scatter \( \sigma_M \) in the \( y - M \) scaling relation must be accounted for. Hilton et al. (2021) assume a log-normal distribution with dispersion \( \sigma_M = 0.2 \). This level of scatter is consistent with both numerical simulations (Bode et al. 2012) and dynamical mass measurements of clusters (Sifón et al. 2013). The distribution of SZ derived cluster masses \( (M_{SZ}) \) for our sample is shown in the lower panel of Fig. 2. We note that the values of \( M_{SZ} \) used in this analysis have no prior \((1 - b_{SZ})\) correction applied.

2.2. Weak Gravitational Lensing from the Kilo Degree Survey

KiDS is an optical survey utilising the OmegaCAM CCD mosaic camera on the VLT Survey Telescope in Chile (de Jong et al. 2015). For this analysis we are using the most recent KiDS dataset, KiDS-1000, first presented in Kuijken et al. (2019)\(^2\). This catalogue includes 30 million galaxies across two areas of sky based around the GAMA spectroscopic survey. KiDS-1000 was observed in 4 bands, \( \{u, g, r, i\} \) with an additional five bands \( \{Z, Y, J, H, K_s\} \) observed by the VISTA Kilo-Degree Infrared Galaxy survey (VIKING; Edge et al. 2013), making KiDS a deep and wide nine-band imaging dataset (Wright et al. 2019). KiDS-1000 includes 1006 deg\(^2\) of imaging with galaxy lensing measurements and accurately calibrated photometric redshifts to \( z_B < 1.2 \) (Wright et al. 2020). The redshift distributions we use in this analysis are calibrated following the methods adopted in the KiDS-1000 cosmology analyses (Asgari et al. 2021; Heymans et al. 2021) which use a self-organising map (SOM). This approach removes source galaxies from our sample because their redshift could not be accurately calibrated with spectroscopic datasets, resulting from incompleteness in the spectroscopic sample (Wright et al. 2020; Hildebrandt et al. 2021). We select source galaxies by their photometric redshift estimate, \( z_B \), which is derived from the nine-band imaging using the Bayesian Photometric Redshifts code (BPZ; Benítez 2000). For each cluster lensing measurement, source galaxies are selected to have \( z_B > z_l + 0.2 \), where \( z_l \) is the redshift of the lens. This selection aims to eliminate unlensed galaxies from our source galaxy sample, although correction factors are still required (discussed in Sect. 2.3). A redshift distribution calibrated by the SOM, based on this \( z_B \) redshift selection, is produced for each cluster. The more selective nature of the SOM method has reduced the galaxy number density and therefore increased the statistical error over the same area, however, the systematic error associated with the photometric redshift distribution has decreased. Source galaxy shape measurements are computed with the lensfit pipeline (Miller et al. 2013), which is calibrated with simulations described in Kannawadi et al. (2019) and presented in Giblin et al. (2021). KiDS-1000 has an effective number density of 6.17 arcmin\(^{-2}\), in the photometric redshift range \((0.1 < z_B < 1.2)\), and the variance of the ellipticities is on average 0.265 (see Table 1, Giblin et al. 2021). For more

---

\( \text{Fig. 2. Upper: normalised SZ signal-to-noise distribution of the ACT clusters found in the KiDS-1000 region. Middle: normalised redshift distribution of these clusters. Lower: normalised SZ-derived mass distribution of these clusters, estimated from the SZ observable Compton-\( y \) values. The blue shaded region corresponds to clusters selected to have an SZ signal-to-noise greater than 5 and in the redshift range 0.1-0.9 across the full ACT footprint. The orange line corresponds to the selection of 157 clusters used in this analysis that lie within the KiDS footprint.} \)

---

\( \text{Hilton et al. (2021) combine Eqs. (2) and (3) to produce a frequency-dependent matched filter across a range of masses and redshifts (each of which change the amplitude and width of the filter). Following Arnaud et al. (2010), the matched filter amplitude, } \bar{y}_0, \text{ can be related to cluster mass via} \)

\[ \bar{y}_0 = 10^{b_0} E^2(z) \left( \frac{M_{SZ, 500}}{M_{500}} \right)^{1+b_0} Q(M_{SZ, 500}, z) f_{rel}(M_{SZ, 500}, z). \]  \( \text{(4)} \)

\( \text{where } 10^{b_0} = 4.95 \times 10^{-5} \text{ is the normalisation, } E(z) \text{ is the ratio of the Hubble constant at a redshift } z \text{ to the present value, } B_0 = 0.08 \text{ is a scaling parameter and } M_{500} = 3 \times 10^{14} M_\odot \text{ is a mass normalisation parameter. The cluster-filter scale mismatch function } Q \text{ accounts for the discrepancy between the size of a cluster with a different mass and redshift to the reference model} \)
technical details on the survey see Kuijken et al. (2019). The footprint of the observed KiDS region is shown in Fig. 1.

2.3. The excess surface density

We summarise the formalism for estimating the weak gravitational lensing signal around a stack of clusters, following Bartelmann & Schneider (2001). Considering an isolated cluster, weak lensing introduces a systematic tangential alignment of the images of the source galaxies relative to the lens. We use the average tangential distortion, $\gamma_t$, to quantify the lensing signal.

This is calculated from the two components of shear, $[\gamma_1, \gamma_2]$, where, in a flat sky approximation, $\phi$ is the angle measured from a line of constant declination to a given source galaxy relative to the centre of the lens and $\gamma_x$ is the shear component along the $45^\circ$ rotated direction. In practice, $\gamma_1$ and $\gamma_2$ are both averaged over all angles so that a gravitational lensing signal can be detected from the average $\gamma_t$, and the average $\gamma_x$ is a useful control statistic that should be consistent with zero for a signal solely caused by gravitational lensing.

Since the tangential distortion we observe is due to intervening matter along the line of sight, the average $\gamma_t$ of source galaxies, at a projected comoving separation $R$ from the cluster lens can be expressed in terms of the excess surface density, $\Delta \Sigma$,

$$\dot{\gamma}_t(R) = \frac{\Delta \Sigma(R)}{\Sigma_c(\chi_1, \chi_l)},$$

where $\Sigma_c$ is the comoving current surface density, written in terms of comoving distances, given by

$$\Sigma_c = \frac{c^2}{4\pi G} \frac{X_s}{(X_s - \chi_1)\chi_1 (1 + z_l)}.$$

for spatially flat models only and $\chi$ is the comoving distance to the lens $\chi_1$ and the source $\chi_s$ and $z_l$ is the lens redshift. The excess surface density (ESD) is defined as

$$\Delta \Sigma(R) \equiv \Sigma(< R) - \Sigma(R),$$

with the average mass density within a radius $R$ given by

$$\Sigma(< R) = \frac{2}{R^2} \int_0^R R' \Sigma(R') dR'.$$

Equation (6) can be generalised for broad lens and source redshift distributions, $n_1$ and $n_2$ to give

$$\langle \gamma_t(R) \rangle = \int_0^\infty dz z n_1(z) \Delta \Sigma(R, z) \int_z^\infty dz' z' n_2(z') \frac{1}{\Sigma_c(z, z')}$$

For a narrow lens and source redshift distribution, $n_1(z) = \delta_\delta(z - z_l)$ and $n_2(z') = \delta_\delta(z' - z_s)$, so that the above equation reduces to Eq. (6). Equation (10) cannot be solved for in closed form, but for a narrow lens redshift distribution, averaging over a background source redshift distribution, $n(z_s)$, we find

$$\Delta \Sigma(R, z_l) \approx \langle \gamma_t(R) \rangle \left[ \int_0^\infty dz z n(z_s) \frac{1}{\Sigma_c(z_l, z_s)} \right]^{-1}.$$

This equation can be re-written as

$$\Delta \Sigma(R, z_l) = \frac{\langle \gamma_t(R) \rangle}{\Sigma_c^{-1}},$$

where, for a given dataset, the average inverse critical surface density, $\Sigma_c^{-1}$, is defined as

$$\Sigma_c^{-1} \equiv \int_0^\infty dz z n(z_s) \Sigma_c^{-1}(z_l, z_s) = \frac{4\pi G}{c^2} (1 + z_l) \chi_1 \int_0^\infty dz z n(z_s) \left( 1 - \frac{\chi_1}{\chi_s} \right).$$

From the equations derived above we can write an estimator for the ESD as

$$\Delta \Sigma = \frac{\sum_i \sum_{\ell} w_{b,\ell} e_{b,\ell} \left( \sum_{c,j} \right)^{-1} \Delta_{b,j}}{\sum_i \sum_{\ell} w_{b,\ell} \Delta_{b,j}},$$

where $\Delta_{b,j}$ is a bin selector function which is defined to be unity if the radial separation between a foreground lens $l$ and background source $s$, lies within a bin centred on $R_l$. The tangential ellipticity of source galaxy $s$ projected on to an axis perpendicular to the line between the lens and source is given by $e_{ls}$ and the weight, $w_{b,\ell}$ assigned to a given source galaxy, $s$, behind a lens, $l$, is

$$w_{b,\ell} = w_{s} \Sigma_c^{-1}.$$

Since our shape measurement pipeline uses lensfit, each source galaxy has a corresponding weight which is approximately an inverse variance weighting that accounts for total shape noise, $w_{s}^{-1} \sim \sigma_{\text{r}}^{-2} + \sigma_{\text{m}}^{-2}$, where $\sigma_{\text{r}}$ is the intrinsic shape dispersion, and $\sigma_{\text{m}}$ is the shape measurement error.

Due to the uncertainty associated with the photometric redshifts of the source galaxies, there will be some source galaxies that are a cluster member or a foreground galaxy. This leads to a dilution of the lensing signal and therefore a mass estimate that is biased low. This effect can be corrected for by making the same ESD measurements around 100 random positions, denoted with the subscript $r$, that are unclustered but still have the same selection function as the lensing clusters. This approach was first shown for the tangential shear estimator (Mandelbaum et al. 2006) but can analogously be applied to the ESD (Singh et al. 2017) as

$$\Delta \Sigma = \frac{\sum_i \sum_{\ell} w_{b,\ell} e_{b,\ell} \left( \sum_{c,j} \right)^{-1} \Delta_{b,j}}{\sum_i \sum_{\ell} w_{b,\ell} \Delta_{b,j}} (N_i - 1) - \frac{\sum_i \sum_{\ell} w_{b,\ell} e_{b,\ell} \left( \sum_{c,j} \right)^{-1} \Delta_{b,j}}{\sum_i \sum_{\ell} w_{b,\ell} \Delta_{b,j}} \Delta_{r,j}.$$

Here $N_i$ and $N_r$ are the number of lenses and random positions respectively within radial bin $\Delta_{b,j}$. Normalising with random lenses produces an unbiased estimate of the ESD unaffected by any source-lens clustering, compared to Eq. (14). This is often referred to as the “boost” factor (Sheldon et al. 2004) which is given by

$$B_l = \frac{N_i \sum_i \sum_{\ell} w_{b,\ell} \Delta_{b,j}}{N_r \sum_i \sum_{\ell} w_{b,\ell} \Delta_{b,j}},$$

for the case of lenses and random positions having no additional weighting applied. In our analysis we ensure our random sample has the same area and redshift properties as our cluster sample.

We apply an additional correction factor to account for multiplicative bias shear calibration which accounts for uncertainty in the shape measurement calibration (Kannawadi et al. 2019).
and multiplies the ESD estimator defined in Eq. (16) as
\[ \bar{m}_i = \sum_w \sum_j m_s \Delta \sigma_i \]
for
\[ C = C^{\text{stat}} + C^{\text{loss}} + C^{\text{int}}, \]
where \( C^{\text{stat}} \) accounts for the statistical uncertainty due to galaxy shapes, \( C^{\text{loss}} \) is the covariance due to projection effects from uncorrelated large-scale structure (Hoekstra 2001) and \( C^{\text{int}} \) corresponds to the intrinsic variations due to halo triaxiality and correlated halos. The contribution to the covariance from shape noise is diagonal by definition and is given by
\[ \sigma^2_{\text{stat},i} = \sigma^0_\gamma \frac{\sum_k \left( w_k \Sigma c_k \right)^2 \Delta \sigma_i}{\left( \sum_k w_k \Sigma c_k \right)^2} \]
for the estimator given in Eq. (16). The large-scale structure and intrinsic parts of the covariance are defined in Eqs. (A.1) and (A.7) in Miyatake et al. (2019).

3. Results
In our fiducial analysis we measure the ESD, \( \Delta \Sigma \), “stacking” all 157 ACT-detected clusters, as given by Eq. (14). The lensing signal is detected with a signal-to-noise of 25 in the range 0.1–10 Mpc from the cluster centres, and the ESD profile is shown in Fig. 3. The detection significance is computed as \( \sqrt{\sum_i \Delta \Sigma_i^2 / \sigma^2_i} \).

3.1. Modelling the stacked ESD
We estimate the weak lensing mass, \( M_{\text{WL}} \), of the stacked clusters using two models: a single NFW halo (Navarro et al. 1997) and a full halo model (Peacock & Smith 2000; Seljak 2000). The single NFW fit assumes that the measured lensing signal is only due to the cluster. The halo model also assumes that the central cluster halo is described by an NFW but additionally accounts for line-of-sight lensing. Furthermore, the halo model method is superior for modelling the ESD signal, as the ACT selection function can be included and we also use a modified NFW profile to account for mis-centring at small scales. We include the NFW fit to compare our results with previous analyses which have utilised a similar modelling approach.

The signal we observe is actually the reduced shear \( \kappa = \gamma / (1 - \kappa) \), where \( \kappa \) is the convergence which accounts for the change in size of an object due to gravitational lensing. At large scales \( \kappa \) is very small and so the true shear, \( \gamma \approx \kappa \). At small scales, however, \( \kappa \) is larger and the equivalence of \( \gamma \) and \( \kappa \) no longer holds. Given that we use data at small scales, we apply a correction to our model for the ESD signal to account for reduced shear following Schrabback et al. (2018).

This factor simply multiplies the model for the ESD as
\[ f_{\text{reduce}} \Delta \Sigma_{\text{model}} \]
and is given for a stack of clusters as
\[ f_{\text{reduce}} = \left[ 1 + \left( \frac{\langle \beta_i^2 \rangle}{\langle \beta_i \rangle^2} - 1 \right) \langle \beta_i \rangle k_{\text{model}} \right] \frac{1}{\left( 1 - \langle \beta_i \rangle k_{\text{model}} \right)}, \]
where \( k_{\text{model}} \) is the model convergence computed at infinite redshift and the shape weights are accounted for as
\[ \langle \beta_i \rangle = \sum \frac{\beta_i w_i}{w_i}, \quad \langle \beta_i^2 \rangle = \sum \frac{\beta_i^2 w_i}{w_i}, \]
where \( \beta_i \) is computed per lens for the estimated source redshift distribution as
\[ \beta_i = \frac{\chi_s}{\chi_\infty} \int_{z_s} dz_s n(z_s) \frac{\chi_s - \chi(z_s)}{\chi_s}. \]

The \( \langle \beta_i^2 \rangle / \langle \beta_i \rangle^2 \) in the correction term accounts for the width in the redshift distribution (Appleby et al. 2014; Hoekstra et al. 2000; Seitz & Schneider 1997). For our sample this correction factor is \( \sim 1.3 \) at the smallest scales, which correspond to 1σ shift in the ESD. For \( R > 0.5 \) Mpc the correction factor decreases to less than 1.05.

3.1.1. Single NFW
We estimate the average cluster mass, \( M_{\text{WL}} \), by assuming an NFW density profile (Navarro et al. 1997) and calculating the expected ESD using Wright & Brainerd (2000), summarised here. The NFW profile is given by
\[ \rho(r) = \frac{\delta_c \rho_c}{(r/r_s)(1 + r/r_s)^2}, \]
where the characteristic overdensity, \( \delta_c \), is related to the NFW concentration parameter, \( c_{500} \), via
\[ \delta_c = \frac{500}{3} \left[ \frac{c_{500}^2}{\ln(1 + c_{500}) - c_{500}/(1 + c_{500})} \right]. \]
In Eq. (24), \( r_s = r_{500}/c_{500} \) is the scale radius. We use the \( r_{500} \) radius to be consistent with the SZ measurements which also use this definition. The mass within a radius of \( r_{500} \), is given by

\[
M_{500} \equiv M(r_{500}) = 500 \rho_{c} \frac{4\pi}{3} r_{500}^{3}.
\]  
(26)

Within this model we fit the mass, \( M_{500} \), and the concentration, \( c_{500} \), using wide flat priors and take the average cluster redshift of \( z = 0.36 \), which is calculated by weighting the clusters using the weights defined in Eq. (15). Fitting for the concentration, rather than assuming a mass-concentration relation which has been calibrated on dark matter-only simulations, has been shown to capture the change in the ESD profile due to baryon feedback which effectively redistributes material from the centre to the outer regions of the cluster. Debackere et al. (2021) showed that when assuming a mass-concentration relation, the mass inferred from weak lensing can be overestimated by 10% at the smallest scales and underestimated by around 5% at the largest scales considered in our analysis. Allowing the concentration and mass to vary removes the bias at large scales and reduces the bias at small scales to less than 5% for the largest halos, which are most affected, and to ~1% for the halos with a mass comparable to those considered in this work. Thus this approach for mitigating the impact of baryons is sufficient in decreasing the bias in the mass estimate. Debackere et al. (2021) demonstrated however that when performing a cosmological analysis the bias in cosmological parameters increases due to a mis-match in the mass definition when the mass function is obtained from dark matter-only simulations. As a result the inferred bias cannot be directly applied as a calibration in a cosmological analysis. Similar arguments have been put forward by Cromer et al. (2022), who implemented a generalised NFW – which allows for a more flexible radial dependence of the density profile – to account for the effects of baryons, and found biases 5–10% when adopting an NFW profile. We therefore caution against over interpreting the best-fitting NFW model shown in Fig. 3, which has a \( p \)-value of 0.81. We only use the range \((0.5 < R < 5)\) Mpc to avoid any mis-centring biases at small scales (see Appendix A), and to avoid including the potentially significant signal from the large-scale structure at scales greater than 5 Mpc (Applegate et al. 2014). This upper limit was chosen apriori and the lower limit was informed by the analysis presented in Appendix A although this is consistent with the ranges used in previous cluster weak lensing analyses. We compute the average weak lensing mass to be \( M_{500}^{\text{NFW}} = (4.9 \pm 0.4) \times 10^{14} M_{\odot} \).

3.1.2. Halo model method

Cluster lensing can be understood within the framework of the halo model following Seljak (2000), Peacock & Smith (2000), Cooray & Sheth (2002). The projected surface density is completely specified by the cluster–matter cross correlation, \( \xi_{c,m} \), which can be computed from a given halo occupation model,

\[
\Sigma(R) = 2\hat{\rho} \int_{R}^{\infty} \xi_{c,m}(r) \frac{r \, dr}{\sqrt{r^{2} - R^{2}}}
\]  
(27)

where \( r \) is the 3D comoving distance \( r^{2} = R^{2} + (\chi - \chi_{c})^{2} \), for a flat Universe. The cluster–matter cross correlation power spectrum, at a wavenumber \( k \), is given in terms of \( \xi_{c,m} \) as

\[
P_{c,m}(k) = 4\pi \int_{0}^{\infty} \xi_{c,m}(r) \frac{\sin(kr)}{kr} \, r^{2} \, dr.
\]  
(28)

The power spectrum of the cluster–dark matter cross correlation can be split into two parts:

\[
P_{c,m}(k) = P_{c,m}^{1h}(k) + P_{c,m}^{2h}(k)
\]  
(29)

with the 1-halo term, which includes the non-linear regime at small scales, and the 2-halo term which describes the clustering between halos, which is dominant on large scales.

The 1-halo term, \( P_{c,m}^{1h}(k) \), describes the matter distribution inside halos hosting clusters. For a single cluster, this is

\[
P_{c,m}^{1h}(k) = \frac{1}{\bar{\rho}_{c}} \int_{0}^{\infty} P(M) \rho_{c}(k|M) \, dM,
\]  
(30)

where \( \rho_{c}(k|M) \) is the Fourier transform of the cluster density profile \( \rho(r|M) \), which we assume to be described by an NFW profile.

Since we are measuring a stack of clusters, the 1-halo term becomes

\[
P_{c,m}^{1h}(k) = \frac{1}{\bar{\rho}_{c}} \int_{0}^{\infty} P(M) \rho_{c}(k|M) \, dM,
\]  
(31)

where \( P(M) \) is the probability that a cluster within an SZ sample resides in a halo of mass \( M \), which is given

\[
P(M) \, dM = \langle N_{c}|M \rangle \frac{dn(M)}{dM} \, dM,
\]  
(32)

where \( \langle N_{c}|M \rangle \) is the average number of clusters in the SZ selection with a halo of mass \( M \), \( dn(M)/dM \) is the halo mass function and \( \bar{\rho}_{c} = \langle N_{c}|M \rangle \frac{dn(M)}{dM} dM \) is the comoving number density of clusters in the SZ selection. Equation (31) can therefore be re-written as:

\[
P_{c,m}^{1h}(k) = \frac{1}{\bar{\rho}_{c}} \int_{0}^{\infty} \int_{0}^{\infty} \frac{dn(M)}{dM} \rho_{c}(k|M_{c}) \langle N_{c}|M \rangle \, dM,
\]  
(33)

This 1-halo term is constructed so that mis-centring of clusters can be accounted for in the cluster density profile, via the \( \rho_{c}(k|M) \) term defined as

\[
\rho_{c}(k|M) = \frac{d_{\text{in}}(k|M)}{d_{\text{in}}(k|M)} \left( 1 - \rho_{\text{offset}} + \rho_{\text{offset}} e^{-0.5 \left( (\chi - \chi_{c})^{2} \right)} \right),
\]  
(34)

where, \( \rho_{\text{offset}} \) and \( f_{\text{offset}} \) is the probability of mis-centring and the size of the offset, as fraction of \( r_{s} \), respectively.

The 2-halo term, denoted by \( P_{c,m}^{2h}(k) \), describes the correlation between clusters and dark matter particles belonging to separate haloes, and is given by

\[
P_{c,m}^{2h}(k) = \frac{P_{\text{dm}}(k)}{\bar{\rho}_{c}} \int_{0}^{\infty} P(M) b_{h}(M) \, dM \int_{0}^{\infty} \frac{dn(M')}{{dM'}} \rho_{\text{of}}(k|M') \, dM',
\]  
(35)

Here \( P_{\text{dm}}(k) \) is the matter power spectrum, which we assume to be linear, and \( b_{h}(M) \) is the halo bias function. The halo bias function describes how halos of mass \( M \) are biased with respect to the overall dark matter distribution. Using Eq. (32), we can re-write Eq. (35) as

\[
P_{c,m}^{2h}(k) = \frac{P_{\text{dm}}(k)}{\bar{\rho}_{c}} \int_{0}^{\infty} \langle N_{c}|M \rangle b_{h}(M) \, dM \int_{0}^{\infty} \rho_{\text{of}}(k|M') b_{h}(M') \, dM' \, dM'.
\]  
(36)

We adopt the halo bias model from Tinker et al. (2010) and include an amplitude \( A_{h} \) that multiplies the halo bias as
shown on this figure at the lower mass range, the analyses of ACT and Planck clusters using HSC (Miyatake et al. 2019; Medezinski et al. 2018), ACT/CS82 (Battaglia et al. 2016), CLASH (Penna-Lima et al. 2017), PSZ2LenS (Sereno et al. 2017), WtG (von der Linden et al. 2014), CCCP (Hoekstra et al. 2015) which has recently been updated to CCCP and MENeaCS (Herbonnet et al. 2020) and Planck clusters with Planck CMB lensing (Zubeldia & Challinor 2019) analyses. The blue band indicates the prediction from Planck primary CMB measurements if the Eddington bias selection effect for the SZ masses derived for the ACT clusters (Hilton et al. 2021), based on the level of scatter seen in numerical simulations and dynamical mass measurements. We check that the chosen value for the scatter does not significantly impact our estimate of \((1 - b_{SZ})\) by repeating our analysis with \(\sigma_c = 0.001\) and \(\sigma_c = 0.5\). Our inferred value for the mass changes by less than 0.2\(\sigma\) which has a 2% effect on \((1 - b_{SZ})\).

In this model we fit for the following set of parameters: \(c, \ln A, A_S, p_{off}, f_{off}\). The average halo mass, which is the main parameter we are interested in, is a derived parameter, given by

\[
M_{\text{halo}} = \frac{1}{\hat{V}_c} \int_0^{\infty} \frac{S(M, z)}{M_{\text{SZ}}} e^{-\frac{\ln(M) - \ln(M_{\text{min}})}{2\sigma_c^2}} dM_{\text{SZ}},
\]

where we have assumed a log-normal scatter, \(\sigma_c\), and \(S(M, z)\) is the SZ-experiment selection function, which can be computed for a given range in SZ signal-to-noise. We fix the SZ mass scatter to \(\sigma_c = 0.2\), firstly because lensing does not constrain this well and secondly because this is the same scatter taken to correct the Eddington bias selection effect for the SZ masses derived for the ACT clusters (Hilton et al. 2021), based on the level of scatter seen in numerical simulations and dynamical mass measurements. We check that the chosen value for the scatter does not significantly impact our estimate of \((1 - b_{SZ})\) by repeating our analysis with \(\sigma_c = 0.001\) and \(\sigma_c = 0.5\). Our inferred value for the mass changes by less than 0.2\(\sigma\) which has a 2% effect on \((1 - b_{SZ})\).

In this model we fit for the following set of parameters: \(c, \ln A, A_S, p_{off}, f_{off}\). The average halo mass, which is the main parameter we are interested in, is a derived parameter, given by

\[
M_{\text{halo}} = \frac{1}{\hat{V}_c} \int_0^{\infty} \frac{S(M, z)}{M_{\text{SZ}}} e^{-\frac{\ln(M) - \ln(M_{\text{min}})}{2\sigma_c^2}} dM_{\text{SZ}},
\]

where we have assumed a log-normal scatter, \(\sigma_c\), and \(S(M, z)\) is the SZ-experiment selection function, which can be computed for a given range in SZ signal-to-noise. We fix the SZ mass scatter to \(\sigma_c = 0.2\), firstly because lensing does not constrain this well and secondly because this is the same scatter taken to correct the Eddington bias selection effect for the SZ masses derived for the ACT clusters (Hilton et al. 2021), based on the level of scatter seen in numerical simulations and dynamical mass measurements. We check that the chosen value for the scatter does not significantly impact our estimate of \((1 - b_{SZ})\) by repeating our analysis with \(\sigma_c = 0.001\) and \(\sigma_c = 0.5\). Our inferred value for the mass changes by less than 0.2\(\sigma\) which has a 2% effect on \((1 - b_{SZ})\).

In this model we fit for the following set of parameters: \(c, \ln A, A_S, p_{off}, f_{off}\). The average halo mass, which is the main parameter we are interested in, is a derived parameter, given by

\[
M_{\text{halo}} = \frac{1}{\hat{V}_c} \int_0^{\infty} \frac{S(M, z)}{M_{\text{SZ}}} e^{-\frac{\ln(M) - \ln(M_{\text{min}})}{2\sigma_c^2}} dM_{\text{SZ}},
\]

where we have assumed a log-normal scatter, \(\sigma_c\), and \(S(M, z)\) is the SZ-experiment selection function, which can be computed for a given range in SZ signal-to-noise. We fix the SZ mass scatter to \(\sigma_c = 0.2\), firstly because lensing does not constrain this well and secondly because this is the same scatter taken to correct the Eddington bias selection effect for the SZ masses derived for the ACT clusters (Hilton et al. 2021), based on the level of scatter seen in numerical simulations and dynamical mass measurements. We check that the chosen value for the scatter does not significantly impact our estimate of \((1 - b_{SZ})\) by repeating our analysis with \(\sigma_c = 0.001\) and \(\sigma_c = 0.5\). Our inferred value for the mass changes by less than 0.2\(\sigma\) which has a 2% effect on \((1 - b_{SZ})\).

In this model we fit for the following set of parameters: \(c, \ln A, A_S, p_{off}, f_{off}\). The average halo mass, which is the main parameter we are interested in, is a derived parameter, given by

\[
M_{\text{halo}} = \frac{1}{\hat{V}_c} \int_0^{\infty} \frac{S(M, z)}{M_{\text{SZ}}} e^{-\frac{\ln(M) - \ln(M_{\text{min}})}{2\sigma_c^2}} dM_{\text{SZ}},
\]
We use Bayesian inference to recover the posterior probability distribution of these five parameters using emcee, the MCMC Python package (Foreman-Mackey et al. 2013). We use wide flat priors for all parameters ($c \in [0, 10]$, $A_{b} \in [-5, 5]$, $A_{0} \in [0, 10]$, $p_{\sigma} \in [0, 1]$, $f_{\text{off}} \in [0, 1]$) and assume a Gaussian likelihood given by

$$\ln L(\Delta \Sigma_{\text{obs}}(R)|\Delta \Sigma(R, q)) = -\frac{1}{2} \left( \Delta \Sigma_{\text{obs}}(R) - \Delta \Sigma(R, q) \right)^{T} C_{R,R}^{-1} \left( \Delta \Sigma_{\text{obs}}(R) - \Delta \Sigma(R, q) \right),$$

(40)

for the set of parameters $q$. We note that since the effective concentration and halo bias parameters are nuisance parameters, which do not have a significant impact on the posterior distribution, sampling in linear or log space has no impact on our results. We fix the cosmological parameters to Planck Collaboration VI (2020), which are in excellent agreement with the recent CMB lensing results from ACT (Madhavacheril et al. 2023), for computing $P_{\text{lens}}$. Assuming a value for $\sigma_{8}$ which has been determined at a similar redshift to our lensing measurements is more robust as this is not dependent on assuming structure growth. We find our results are insensitive to this choice however, by repeating our analysis assuming cosmological parameters determined from KiDS cosmic shear measurements (Heymans et al. 2021), which are in $\sim 2\sigma$ tension with Planck, where we recover the same estimate of the cluster mass. Furthermore, if we marginalise over $\sigma_{8}$, $\Omega_{m}$ and $h$ in our fit we recover the same value for the inferred mass but our error on the mass increases by 50%. As we have a mis-centring term and a 2-halo term in this model we use the full range of scales from $(0.1-10) \text{ Mpc}$. We check that our mass estimate is robust to the choice of scales included in the analysis by re-performing our fits with the limited range of scales used in the single NFW halo fit, the mass estimate remains the same but the error increases by 20%. The halo model is evaluated at a median redshift of $z = 0.36$, calculated using the same weighting defined in Eq. (15). From our MCMC sampling analysis, we infer the average weak lensing mass estimate to be $M_{500c}^{\text{DM}} = (4.9 \pm 0.4) \times 10^{14} M_{\odot}$. Our best-fitting halo model is shown in Fig. 3 and has a $p$-value of 0.28.

### 3.2. Splitting the cluster sample

Given our large cluster sample and high signal-to-noise in the stacked cluster lensing measurement we split our sample to investigate the dependence of calibration bias on the cluster mass. As the inferred SZ cluster mass has a large associated uncertainty we split based on the SZ signal-to-noise, since the SZ signal is correlated with the cluster mass. The binning is defined in Table 1 with the corresponding number of clusters in each bin, average SZ inferred mass and average cluster redshift.

We perform both NFW and halo model fits to each SZ signal-to-noise bin following the same procedure as described for the total stack of clusters above. The best fit values are reported in Table 1 for both models. The ESD measurements for each bin with corresponding best fit models are shown in Fig. 5.

### 4. Calibration bias: $(1 - b_{\text{SZ}})$

To compute the mass bias $(1 - b_{\text{SZ}})$ we combine estimates of the SZ inferred cluster masses by taking the weighted average. Here clusters are weighted in the same way as in the stacked ESD measurement described in Eq. (15). We estimate the average SZ inferred mass to be $M_{500c}^{\text{SZ}} = (3.18 \pm 0.07) \times 10^{14} M_{\odot}$.

From our halo model estimate of the weak lensing mass, we estimate $(1 - b_{\text{SZ}}) = 0.65 \pm 0.05$. This is consistent with our estimate using a single NFW profile of $(1 - b_{\text{SZ}}) = 0.64 \pm 0.05$. The large deviation of $(1 - b_{\text{SZ}})$ from unity suggests that clusters are not in hydrostatic equilibrium and that baryon feedback plays an important role; this is consistent with recent results from kinetic...
SZ and thermal SZ analyses (for example, Koukoufilippas et al. 2020; Itiboye et al. 2022). This bias, however, cannot be solely attributed to the assumption of hydrostatic equilibrium and is dependent on other assumptions made when estimating $M_{\text{SZ}}$.

This analysis does not consider the impact of lensing mass bias as investigated in, for example, Becker & Kravtsov (2011) who found a mass bias of 5–10%. For our analysis this level of additional bias in the mass would translate into a shift of up to 1$\sigma$ in our constraint on $(1 - b_{\text{SZ}})$. Furthermore, Euclid Collaboration (2023) find a similar level of bias of less than 10%, which will be important for future analyses which are more sensitive than this analysis.

In Fig. 4 we show our new estimate together with a compilation of measurements from other data combinations. ACT and Planck cluster mass estimates can be directly compared since they are derived from the same SZ-mass scaling relations and pressure profiles. Our results are consistent with, though generally lower than, analyses of ACT and Planck clusters in a similar mass range using weak galaxy lensing data from Hyper Suprime-Cam (HSC; see Miyatake et al. 2019; Medezinski et al. 2018), CFHTLenS/RCSSZ (Sereno et al. 2017). We also find consistency with the analysis of Planck clusters with CLASH galaxy weak lensing (Penna-Lima et al. 2017) and Planck CMB lensing (Zubeldia & Challinor 2019). Compared to several of the earlier analyses, shown in Fig. 4, we have a slight improvement in constraining power which is due to the larger cluster sample.

Comparing to higher mass Planck cluster analyses, Wig (von der Linden et al. 2014), CCCP (Hoekstra et al. 2015) and its more recent re-analysis CCCP/MENaCS (Herbonnet et al. 2020), our inferred value for the mass bias is up to 3.8$\sigma$ lower. Battaglia et al. (2016) demonstrated, however, that adding a correction for Eddington bias (see Sect. 3.1) applied to the SZ mass estimates, would lower the value of $(1 - b_{\text{SZ}})$ in both the Wig and CCCP analysis (see Fig. 7 in Battaglia et al. 2016). We do apply that correction in this work. In Fig. 4 we show how our result would change if the correction was not applied, with the value we estimate for the mass bias increasing to $(1 - b_{\text{SZ}}) = 0.74 \pm 0.06$, which is a shift of 2$\sigma$. It appears that the tension between the value for $(1 - b_{\text{SZ}})$ found in this work and the recent result from Herbonnet et al. (2020), who found $(1 - b_{\text{SZ}}) = 0.84 \pm 0.04$, can be explained by uncorrected Eddington bias in the latter.

We also find consistency with the mass bias required for Planck primary CMB cosmology to be consistent with the Planck cluster count analysis, which is indicated by the blue band in Fig. 4. To be able to combine all the measurements shown in Fig. 4, we would need to account for covariances between different datasets since some clusters appear in more than one sample and the difference in the ACT and Planck selection functions.

By splitting our sample based on SZ detection signal-to-noise, a quantity correlated with the cluster mass, we investigate whether the bias is mass dependent. The results are reported in Table 1 and shown in Fig. 6. We fit a straight line to $M_{\text{SZ}}$ against $(1 - b_{\text{SZ}})$, considering two cases: the first just fitting for the y-intercept (no mass dependence) and the second fitting for both the gradient and the y-intercept. We find no evidence for any dependence on mass; with a best fit gradient of $0 \pm 0.1$. We find that repeating the fit with a gradient of zero still produces an excellent fit, a $p$-value of 0.9, and only improves the $p$-value at the sub-percent level when including a mass dependence. Similar results are obtained using the best fit mass estimates from the NFW model. To confirm that this result is not driven by the highest

![Fig. 6. Estimated $(1 - b_{\text{SZ}})$ bias parameter from the total cluster stack shown in black, and in red for six individual SZ signal-to-noise bins. These values are reported in Table 1. The blue band indicates the prediction from Planck primary CMB measurements if the $\Lambda$CDM model is correct (Planck Collaboration VI 2020), given the SZ clusters observed by Planck.](image)

5. Conclusions

We have estimated the masses of a new sample of galaxy clusters detected using the Atacama Cosmology Telescope. Stacking the signal from 157 high SZ signal-to-noise clusters, we detect the weak lensing signal at 25$\sigma$ significance. We estimate the stacked mass to be $M_{\text{WL}} = (4.9 \pm 0.4) \times 10^{14} M_{\odot}$, using an analytical halo model formalism which accounts for the ACT SZ cluster selection. By comparing this gravitational lensing estimate of mass to the mass inferred from the SZ signal itself, we provide a new estimate of the calibration of the SZ Compton–y parameter mass scaling relation.

When modelling the cluster lensing signal we adopt two methods, a simpler NFW and a full halo model. The NFW method only models the central halo term and we therefore limit the scales used for this fit to avoid mis-centring at small scales and large-scale structure at scales greater than 5 Mpc. The halo model method assumes an NFW profile for the central halo, however it also accounts for line-of-sight lensing, the ACT selection function and mis-centring. Both mass estimates are consistent within 1$\sigma$ of each other and have comparable constraining power, which may be initially surprising. This shows that even though we are able to include more data at large and small scales in the halo model fit, the uncertainty on centroiding and the large-scale structure means these extra scales do not help to constrain the mass. The comparable constraining power of these two methods reinforces that our scale cuts for the NFW model fit are correctly defined.

We find our estimate of the mass calibration bias to be consistent, given the statistical uncertainty, with estimates made using other contemporary weak lensing surveys including HSC. This measurement does not indicate a departure from the predicted value of the mass bias from Planck primary CMB measurements if the $\Lambda$CDM model is correct
Cromer, D., Battaglia, N., Miyatake, H., & Simet, M. 2022, JCAP, 2022, 034
Edge, A., Sutherland, W., Kuijken, K., et al., 2013, Messenger, 154, 32

Hoekstra, H., Bartelmann, M., Dahle, H., et al. 2013, Science, Rev., 177, 75
Naess, S., Aoila, S., Austermann, J. E., et al., 2020, JCAP, 2020, 046
Appendix A: Mis-centring of clusters

The mis-centring of clusters has been reported as one of the largest biases when estimating the cluster mass from weak gravitational lensing (e.g. Sommer et al. 2022; Ford et al. 2015). If the true cluster centre has not been identified then the amplitude of the ESD will be reduced at small radii, corresponding to a lower mass estimate for the cluster.

We consider two different cluster centre definitions: the peak of the SZ emission and the Brightest Cluster Galaxy (BCG), which has been shown to be the best tracer of the cluster centre (George et al. 2012). The SZ centre is not expected to trace the centre precisely since it is limited by the size of the beam which is 1.4 arcmin for ACT; this corresponds to $R = 0.09$ Mpc for $z_l = 0.1$ and $R = 0.62$ Mpc for $z_l = 0.9$. We also anticipate that most clusters are not relaxed and therefore the gas is not expected to trace the centre of the dark matter halo.

We find the BCG for each cluster by selecting KiDS galaxies found inside a cylinder, centred on the SZ peak, with a length $|z_B - z_l| = 2 \times 0.04(1 + z_l)$, where $z_B$ is the photometric redshift of a KiDS galaxy, and a diameter $2 \times 1.4$ arcmin. The length of the cylinder is defined to be twice the length of the uncertainty on the photometric redshift of the KiDS galaxies, and the diameter is defined based on the size of the ACT beam. From these galaxies the brightest one is chosen. We then confirm the BCG locations by visual inspection, using public imaging from KiDS, the Dark Energy Camera Legacy Survey and Hyper Suprime-Cam. We note that for some clusters (for example one of the most massive clusters in our sample, Abell 2744/ACT-CL J0014.3–3022) that are undergoing merging and therefore far from being in a virialised state, it is difficult assign an appropriate centre.

Measuring the ESD around both these centre candidates, we find at small radii the ESD measured around the BCG has a higher amplitude than measured around the SZ centre, which agrees with our expectation – this is shown in Fig. A.1. The ratio between the BCG and SZ centred signal when averaged $R > 0.5$ Mpc is unity, compared to the same ratio having a value of ~ 5 when averaged $R < 0.5$ Mpc.

Fig. A.1. Tests of mis-centring: comparison of the ESD measured around the peak in the SZ emission and the BCG. The ratio between the BCG and SZ centred signal averaged $R > 0.5$ is unity.