The Atacama Cosmology Telescope: Weighing Distant Clusters with the Most Ancient Light

Mathew S. Madhavacheril[1] , Cristóbal Sifón[2] , Nicholas Battaglia[3], Simone Aiola[4], Stefania Amodeo[3], Jason E. Austermann[5], James A. Beall[5], Daniel T. Becker[5], J. Richard Bond[6], Erminia Calabrese[7], Steve K. Choi[3,8] , Edward V. Denison[5], Mark J. Devlin[9], Simon R. Dicker[9] , Shannon M. Duff[5], Adriaan J. Duivenvoorden[10], Jo Dunkley[10,11], Rolando Dünner[12], Simone Ferraro[13] , Patricio A. Gallardo[8], Yilun Guan[14], Dongwon Han[15], J. Colin Hill[4,16], Gene C. Hilton[5], Matt Hilton[17,18] , Johannes Hubmayr[5], Kevin M. Huffenberger[19] , John P. Hughes[20] , Brian J. Koopman[21], Arthur Kosowsky[14] , Jeff Van Lanen[5], Eunseong Lee[22], Thibaut Louis[23], Amanda MacInnis[15], Jeffrey McMahon[24,25,26,27], Kavilan Moodley[17,18], Sigurd Naess[4], Toshiya Namikawa[28] , Federico Nati[29], Laura Newburgh[21], Michael D. Niemack[3,8], Lyman A. Page[10], Bruce Partridge[30], Frank J. Qu[28], Naomi C. Robertson[31,32], Maria Salatino[33,34], Emmanuel Schaan[13], Alessandro Schillaci[35], Benjamin L. Schmitt[36], Neelima Sehgal[15] , Blake D. Sherwin[28,32], Sara M. Simon[37], David N. Spergel[4,11] , Suzanne Staggs[10], Emilie R. Storer[10], Joel N. Ullom[5], Leila R. Vale[5], Alexander van Engelen[38], Eve M. Vavagiakis[8], Edward J. Wollack[39] , and Zhilei Xu[9]

[1] Centre for the Universe, Perimeter Institute, Waterloo, ON N2L 2Y5, Canada; mmadhavacheril@perimeterinstitute.ca
[2] Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4059, Valparaíso, Chile; cristobal.sifon@pucv.cl
[3] Department of Astronomy, Cornell University, Ithaca, NY 14853, USA
[4] Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY, 10010, USA
[5] NIST Quantum Devices Group, 325 Broadway Mailcode 817.03, Boulder, CO 80305, USA
[6] Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON, M5S 3H8, Canada
[7] School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, CF24 3AA, UK
[8] Department of Physics, Cornell University, Ithaca, NY 14853, USA
[9] Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA
[10] Joseph Henry Laboratories of Physics, Jadwin Hall, Princeton University, Princeton, NJ 08544, USA
[11] Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA
[12] Instituto de Astrofísica and Centro de Astro-Ingeniería, Facultad de Física, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile
[13] Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[14] Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA
[15] Physics and Astronomy Department, Stony Brook University, Stony Brook, NY 11794, USA
[16] Department of Physics, Columbia University, 550 West 120th Street, New York, NY 10027, USA
[17] Astrophysics Research Centre, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa
[18] School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa
[19] Department of Physics, Florida State University, Tallahassee, FL 32306, USA
[20] Department of Physics and Astronomy, Rutgers University, 136 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA
[21] Department of Physics, Yale University, New Haven, CT 06520, USA
[22] Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, University of Manchester, Manchester, UK
[23] Université Paris-Saclay, CNRS/IN2P3, IJCLab, F-91405 Orsay, France
[24] Kavli Institute for Cosmological Physics, University of Chicago, 5640 S. Ellis Avenue, Chicago, IL 60637, USA
[25] Department of Astronomy and Astrophysics, University of Chicago, 5640 S. Ellis Avenue, Chicago, IL 60637, USA
[26] Department of Physics, University of Chicago, Chicago, IL 60637, USA
[27] Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA
[28] Center for Theoretical Cosmology, DAMTP, University of Cambridge, CB3 0WA, UK
[29] Department of Physics, University of Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy
[30] Department of Physics and Astronomy, Haverford College,Haverford, PA 19041, USA
[31] Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
[32] Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 OHA, UK
[33] Physics Department, Stanford University, 382 via Pueblo, Stanford, CA 94305, USA
[34] Kavli Institute for Particle Astrophysics and Cosmology, 452 Lomita Mall, Stanford, CA 94305-4085, USA
[35] Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA
[36] Harvard-Smithsonian Center for Astrophysics, Harvard University, 60 Garden Street, Cambridge, MA 02138, USA
[37] Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
[38] School of Earth and Space Exploration and Department of Physics, Arizona State University, Tempe, AZ 85287, USA
[39] NASA/Goddard Space Flight Center, Greenbelt, MD 20771, USA

Abstract

We use gravitational lensing of the cosmic microwave background (CMB) to measure the mass of the most distant blindly selected sample of galaxy clusters on which a lensing measurement has been performed to date. In CMB data from the the Atacama Cosmology Telescope and the Planck satellite, we detect the stacked lensing effect from 677 near-infrared-selected galaxy clusters from the Massive and Distant Clusters of WISE Survey (MaDCoWS), which have a mean redshift of $\langle z \rangle = 1.08$. There are currently no representative optical weak lensing measurements of clusters that match the distance and average mass of this sample. We detect the lensing signal with a significance of 4.2σ. We model the signal with a halo model framework to find the mean mass of the population from which these clusters are drawn. Assuming that the clusters follow Navarro–Frenk–White (NFW) density profiles, we infer a mean mass of $\langle M_{500c} \rangle = (1.7 \pm 0.4) \times 10^{14} \, M_\odot$. We consider systematic uncertainties from cluster redshift errors, centering

errors, and the shape of the NFW profile. These are all smaller than 30% of our reported uncertainty. This work highlights the potential of CMB lensing to enable cosmological constraints from the abundance of distant clusters populating ever larger volumes of the observable universe, beyond the capabilities of optical weak lensing measurements.

Unified Astronomy Thesaurus concepts: Cosmology (343); High-redshift galaxy clusters (2007); Cosmic microwave background radiation (322); Gravitational lensing (670)

1. Introduction

Most of the mass in the universe is thought to consist of dark matter that does not interact with electromagnetic radiation other than through the gravitational force. Distortions in background light sources due to the gravitational influence of massive clusters of galaxies can be used to produce maps of the total matter distribution. This technique has served not only as evidence for the existence of dark matter (Trimble 1987; Massey et al. 2010), but also as a method for inferring the total mass of galaxy clusters themselves (e.g., Hoekstra et al. 2013). Such mass measurements are critical for the program of using the abundance of clusters to infer cosmological parameters like the dark energy equation of state or the mass scale of neutrinos (e.g., Allen et al. 2011; Madhavacheril et al. 2017). One key aspect of this program is our ability to constrain the abundance of clusters to high redshifts, directly probing the growth of structure through cosmic time (Voit 2005).

Sensitive large-area optical/near-infrared surveys are now allowing measurements of the mean mass of clusters up to $z \sim 1$ (Murata et al. 2019; Chiu et al. 2020) through the lensing effects induced on background galaxies. However, as the distance of the clusters increases, the number of background galaxies that are useful for weak lensing measurements decreases rapidly. Measurements of this *galaxy weak lensing* effect at large distances are therefore only possible at present through deep targeted observations with the Hubble Space Telescope (Jee et al. 2011; Schrabback et al. 2018), with which observations currently exist for only some dozens of rich, massive clusters at redshifts of $z > 0.8$. A valuable complementary probe is emerging as cosmic microwave background (CMB) measurements are becoming sensitive enough to allow measurements of weak lensing by galaxy clusters in maps of the temperature and polarization of the CMB (e.g., Baxter et al. 2015; Madhavacheril et al. 2015; Planck Collaboration & Ade 2016; Geach & Peacock 2017; Raghunathan et al. 2019; Zubeldia & Challinor 2019). The high source redshift of the CMB allows weak lensing measurements to higher redshifts than galaxy lensing. Consequently, measurements of CMB lensing by clusters are anticipated to provide more stringent constraints on the masses of high-redshift clusters than enabled by future optical surveys (e.g., Madhavacheril et al. 2017).

We provide a mass estimate using gravitational lensing of the CMB for a blindly selected sample of galaxy clusters whose average mass has not been previously determined using galaxy lensing, which is additionally the highest redshift, $\langle z \rangle = 1.08$, where a detection of gravitational lensing by galaxy clusters has been reported for a blindly selected sample to date. As opposed to targeted measurements of the most massive clusters, our work allows for inference of the average mass of a representative cluster sample. We make the code used in this analysis available at https://github.com/ACTCollaboration/madcows_lensing.

2. Data

We use a combination of CMB data from the ground-based Atacama Cosmology Telescope (ACT) and the Planck satellite at the location of galaxy clusters selected from the Massive and Distant Clusters of WISE Survey (MaDCoWS, Gonzalez et al. 2019). The MaDCoWS clusters were identified as galaxy overdensities in near-infrared imaging (at 3.4 and 4.6 μm) from the WISE all-sky survey (Wright et al. 2010), and a large number of them were followed up with the Spitzer Space Telescope (at 3.6 and 4.5 μm). At declinations $> -30°$, the addition of optical data (*grizy* bands) from the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS, Chambers et al. 2016) allowed reliable photometric redshift estimation and subsequent cluster richness measurements. As a first attempt at a mass proxy, Gonzalez et al. (2019) define the richness, λ, to be the overdensity of red-sequence galaxies (identified using both PanSTARRS and Spitzer data) brighter than 15 μJy in the Spitzer 4.5 μm band and within 1 Mpc of the brightest cluster galaxy. Photometric redshifts were estimated from the Spitzer 3.6 and 4.5 μm bands, aided with the PanSTARRS *i*-band to remove low-redshift galaxies, with an estimated scatter $\sigma_z/(1 + z) = 0.04$ and no significant bias (but with an outlier fraction potentially of order 5%).[40] We discuss the impact of photometric redshift uncertainties below. We use this subset of the MaDCoWS "WISE-PanSTARRS" sample with available photometric redshifts, imposing an additional $\lambda > 20$ cut to reduce contamination by false detections in the MaDCoWS catalog, for a total of 677 clusters after masking point sources and other artifacts in the ACT maps (which is also limited to declinations $\leqslant 20°$). We show the richness and redshift distributions of these clusters in Figure 1. The mean redshift of the sample is $\langle z \rangle = 1.08$ with a 16th and 84th percentile range of 0.93–1.26.

To reconstruct the lensing signal, we use coadded maps of ACT and Planck CMB temperature data prepared separately at 98 and 150 GHz and described in Naess et al. (2020). The coadded maps include night-time data collected during the years 2008–2018 using the MBAC (Swetz et al. 2011), ACTPol (Thornton et al. 2016), and AdvACT (Henderson et al. 2016) receivers. The Planck maps used in the coadded maps are the PR2 (2015) CMB temperature maps at 100 and 143 GHz. We also use the Planck 2018 SMICA tSZ-deprojected maps (Planck Collaboration et al. 2020) as an additional input to the lensing reconstruction (see Appendix A). This is done in order to remove the bias from the thermal Sunyaev–Zeldovich (tSZ) effect due to inverse-Compton scattering of CMB photons off ionized electrons in hot gas in massive clusters (Madhavacheril & Hill 2018). Deprojection of the tSZ is possible because the frequency dependence of the spectral distortion due to the tSZ effect is well understood.

---

[40] These statistics are based on a comparison of photometric and spectroscopic redshifts for 38 clusters for which the latter is available.
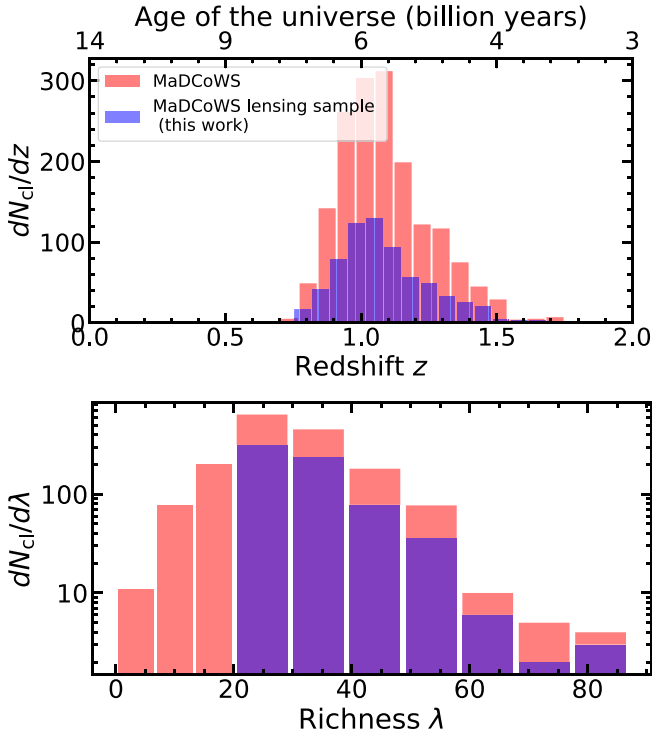
Figure 1. Distributions for cluster redshift (top), and for a measure of the number of galaxies in a cluster, richness (bottom) for the MaDCoWS WISE-PanSTARRS galaxy cluster sample used in this work. The blue histograms correspond to the 677 clusters that remain after applying the richness cut ($\lambda > 20$) and ACT mask, while the red histograms correspond to the full sample of 1676 clusters.

3. Lensing Reconstruction

We reconstruct the CMB lensing convergence κ in regions centered on the locations of MaDCoWS clusters. The lensing convergence is related to the line-of-sight integral ϕ of the underlying gravitational potential sourced by a cluster, and to the lensing deflection angle $\boldsymbol{\alpha}$ through $\nabla^2\phi = \boldsymbol{\nabla} \cdot \boldsymbol{\alpha} = -2\kappa$. The convergence is related to the surface mass density Σ through $\kappa = \Sigma/\Sigma_{\rm cr}$, where $\Sigma_{\rm cr}$ is a characteristic mass density for the formation of multiple images that depends on distances to the source and lens (see Appendix B). Since the convergence map is directly proportional to the surface mass density, it can be thought of as a mass map. Gravitational lensing of the CMB by clusters leads to a remapping of the temperature anisotropies $T(\boldsymbol{x}) = T_{\rm unlensed}(\boldsymbol{x} + \boldsymbol{\alpha})$. In the 2D Fourier space of the temperature anisotropy image, this remapping corresponds to coupling between previously independent Fourier modes (say with wavenumbers ℓ and ℓ') that is proportional to the lensing convergence: $\langle T(\ell)T(\ell')\rangle \propto \kappa(\ell + \ell')$. This allows us to reconstruct the underlying convergence mode by mode by using a "quadratic estimator," i.e., a weighted sum over products of pairs of image modes (Hu et al. 2007), which in practice can be written as the divergence of the product of the large-scale CMB gradient and the small-scale CMB fluctuations (see Appendix A). The quadratic estimator reconstruction provides an unbiased estimate of the Fourier modes of the cluster mass map within a range of scales set by the band limit of the CMB maps. Thus, the output reconstruction image $\hat{\kappa}(\boldsymbol{x})$ is effectively filtered, and we forward model this filtering when fitting to theoretical expectations. The signal-to-noise ratio for any given cluster is well below unity, and so our final
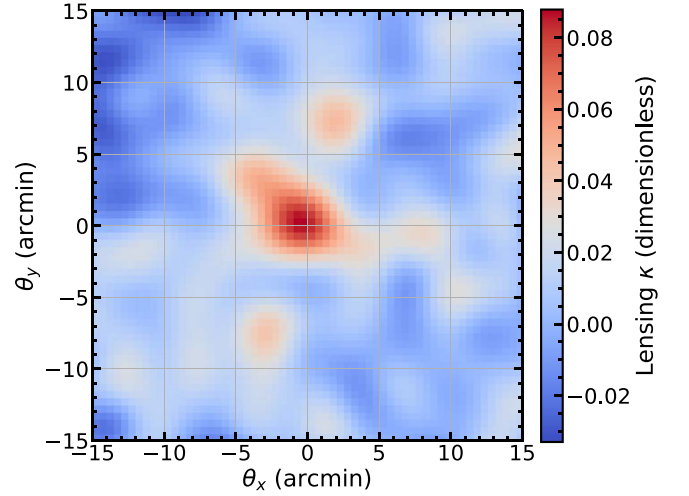

Figure 2. The stacked CMB lensing convergence mass map from the sample of 677 MaDCoWS clusters with $\langle z \rangle = 1.08$ used in this work. This reconstruction includes a low-pass filter up to a scale of 2 arcminutes and has been additionally smoothed with a Gaussian filter with FWHM of 3.5 arcminutes.

measurements are a weighted average of the mass maps and of the radial profile (azimuthal average) of the lensing convergence, where the weights are inversely proportional to the noise variance in the lensing convergence reconstruction. We estimate a weighted covariance matrix for the radial profile from the scatter among the profiles. With each cluster profile weighted by w_i, the weighted covariance of the weighted mean of the N clusters in our sample is

$$C = \frac{1}{N} \frac{\sum_{i=1}^{N} w_i (\boldsymbol{p}_i - \boldsymbol{d})^T (\boldsymbol{p}_i - \boldsymbol{d})}{V_1 - (V_2/V_1)}. \qquad (1)$$

where $\boldsymbol{p}_i$ are the individual radial profiles of each cluster mass map, $\boldsymbol{d} = \frac{1}{V_1}\sum_{i=1}^{N} w_i \boldsymbol{p}_i$ is the weighted mean profile, $V_1 = \sum_{i=1}^{N} w_i$ and $V_2 = \sum_{i=1}^{N} w_i^2$.

Details of the reconstruction process and mitigation of astrophysical foregrounds are provided in Appendix A. We show the resulting stacked CMB lensing convergence reconstruction in Figure 2.

4. Results

After averaging lensing convergence maps across all the clusters in our sample, we detect an excess relative to null within $8'$ at 4.9σ confidence. We estimate the mean mass of the sample by fitting the binned radial profile of the average mass map assuming the Navarro–Frenk–White (NFW; Navarro et al. 1996) density profile model for the distribution of matter within galaxy clusters with a fixed relation between the profile shape (or *concentration* parameter) and the mass. We account for the distribution of clusters of different masses across redshift and their clustering (i.e., the two-halo term), and apply an approximate correction for selection effects in order to construct an ensemble average of the lensing signal. This *halo model* approach, detailed in Appendix B, allows us to estimate the mean mass of the underlying sample from which the MaDCoWS sample is drawn.

We show the model predictions for the filtered lensing signal in Figure 3, color-coded by the excess χ^2 with respect to the minimum χ^2 obtained from our fit, corresponding to $\chi^2_{\rm min} = 1.4$.
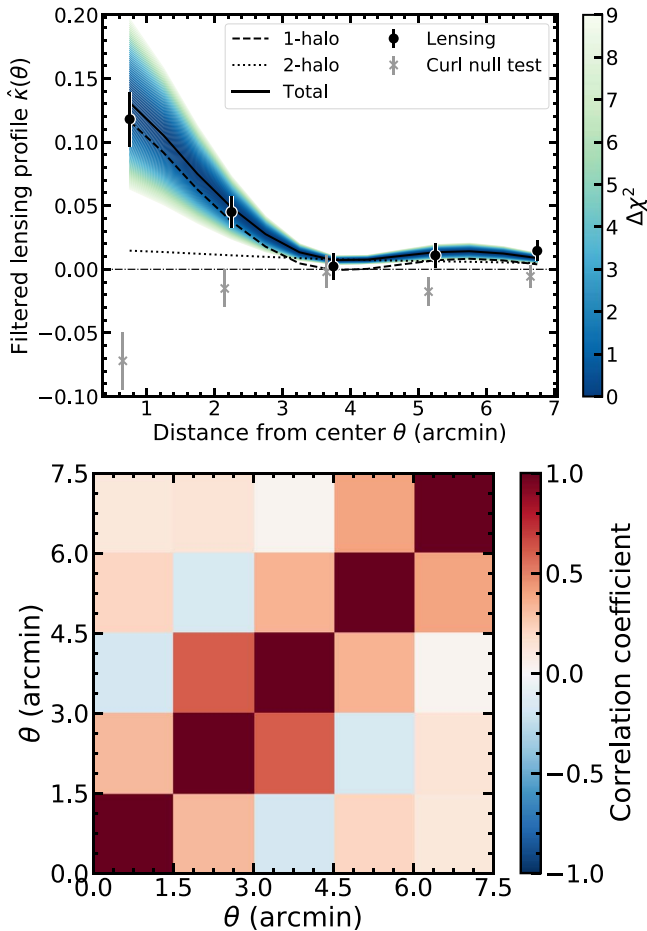
Figure 3. Top: halo model fit to the profile of the stacked lensing convergence, color-coded by $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$. With only one free parameter in the model, the 1σ, 2σ, and 3σ credible regions are limited by $\Delta\chi^2 = 1$, 4, and 9, respectively. The solid line shows the best-fit model, while the dashed and dotted–dashed lines show the contributions from the one-halo and two-halo terms to it, respectively. Gray crosses show measurements of the curl (slightly offset horizontally for clarity), which is expected to be zero (see Appendix A). Error-bars correspond to the square root of the diagonal elements of the covariance matrix. The probability-to-exceed (PTE) the χ^2 of our best-fit model is 0.93. The PTE of the curl compared to a null signal is 0.05. Bottom: the correlation coefficient matrix for the radial bins of the lensing measurement. The corresponding matrix for the curl measurement is similar. The large bin–bin correlations show that all the curl points fluctuating below zero is not unlikely.

Since our model has only one free parameter, the 1σ, 2σ, and 3σ credible ranges are given by $\Delta\chi^2 = 1$, 4, and 9, respectively. We calculate the probability-to-exceed $\chi^2_{\min}$ (PTE) by drawing one million random samples with mean zero and using the covariance of our measurements; we find a PTE of 0.93, showing that the model is an adequate description of the data. The best-fit mean mass is $\langle M_{500c} \rangle = (1.7 \pm 0.4) \times 10^{14}\, M_\odot$, with a preference for our best fit over zero mass at the 4.2σ level.

5. Discussion

The clusters in our sample have a mean redshift of $\langle z \rangle = 1.08$, the highest-redshift, blindly selected sample for which a gravitational lensing measurement has been obtained to date. Our results highlight the potential of both CMB lensing in general and ACT in particular to constrain the scaling relation between mass and other observables for high-redshift galaxy clusters.

In Appendix B, we explore systematic uncertainties from cluster redshift errors, centering errors, and the concentration parameter of the NFW profile, which are all smaller than 30% of our reported uncertainty. Since our measurement probes the overall lensing amplitude, the reported mass is also robust to assumptions about both the selection function and the scaling relation between mass and richness, which conversely we are not able to constrain.

Because we stack the lensing maps around all MaDCoWS clusters within the ACT footprint with a richness larger than 20, our measurement is representative of the full MaDCoWS sample above this cut. Targeted observations of the tSZ effect (which scales steeply with cluster mass as $\sim M^{\frac{5}{3}}$) of MaDCoWS subsamples by Gonzalez et al. (2019), Di Mascolo et al. (2020), and Dicker et al. (2020) have naturally picked preferentially the most luminous objects, and thus the mean richness of our sample, $\langle\lambda\rangle = 32$, is lower than that probed in those works. The mean mass we estimate is consistent with the scaling relations reported in those works, which however have large uncertainties and are susceptible to additional systematic effects associated with an uncertain selection function.

While looking for galaxy overdensities is an efficient means of finding galaxy clusters, characterizing the selection effects that these techniques are subject to is a challenging endeavor. No estimate of the selection function is available for MaDCoWS, which precludes the use of this sample for cosmological parameter inference. Future CMB lensing measurements of well-defined cluster samples at high redshift will yield tight constraints on the parameters that govern the growth of structure.

Appendix A
Method Details and Systematics Tests

The quadratic estimator for CMB lensing reconstruction can be recast as the divergence of the product of the small-scale CMB and the large-scale CMB gradient:

$$\hat{\kappa}(\boldsymbol{\theta}) = -\mathcal{F}^{-1}\{A^{TT}(\boldsymbol{L})\mathcal{F}\{\mathrm{Re}\left[\nabla \cdot \left[\boldsymbol{\nabla} T_g(\boldsymbol{\theta})T_h(\boldsymbol{\theta})\right]\right]\}\}, \quad \text{(A1)}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote 2D Fourier and inverse-Fourier transforms, respectively. Our analysis applies this estimator locally to cutouts of CMB data centered on each cluster location that are $128'$ wide with pixels of width 0.5 arcminutes. The large width relative to the typical arcminute size of clusters allows us to estimate the large-scale gradient.

Following Hu et al. (2007), we use a low-pass (top-hat) filtered temperature anisotropy gradient $\boldsymbol{\nabla} T_g$ with a maximum CMB multipole of $\ell_G = 2000$ to mitigate bias in massive clusters as well as to reduce contamination from foregrounds. Furthermore, following Madhavacheril & Hill (2018), for the gradient map $\boldsymbol{\nabla} T_g$, we use a CMB map from which tSZ has been explicitly deprojected using multifrequency information (the Planck PR3 SMICA tSZ-deprojected), so as to null a large tSZ-induced bias. For the high-resolution map T_h, we use an internal linear combination (ILC) of the postage stamp cutouts from the 98 and 150 GHz 2018 coadds of Planck and ACT from Naess et al. (2020), where the high-resolution ACT data dominates the information content. While the tSZ cleaning for $\boldsymbol{\nabla} T_g$ nulls the tSZ bias, a large amount of variance can be induced by the presence of large tSZ decrements in T_h. To reduce this, following Patil et al. (2020), prior to the ILC we subtract best estimates of the tSZ decrements from each of the ~4000 SZ clusters detected in the coadd (Hilton et al. 2020) since some of these clusters either appear in our sample or may appear in the postage stamps that we perform our reconstruction on. This has the effect of reducing the uncertainty in the first three bins of our measurement by ≈10%. The weights for the ILC are designed to minimize the power spectrum of the combination and are determined after fitting the total 1D power spectrum of each single-frequency stamp to a simple two parameter model that captures both an atmospheric component as well as white noise. We impose a maximum multipole cut of $\ell = 6000$ on the high-resolution map. Both maps are inverse-variance filtered as in Hu et al. (2007). For both the gradient and high-resolution map, we do not use scales below $\ell = 200$ due

to the size of our cutouts. The reconstruction is normalized with an analytic expression $A^{TT}(\boldsymbol{L})$ (Hu et al. 2007) that depends on the filters we apply. The final reconstruction is filtered to ensure it only contains modes $200 < L < 5000$, and this filter is subsequently propagated in our theoretical model when fitting the profile.

Prior to reconstruction, the gradient and high-resolution maps are multiplied by a tapering cosine window (of approximate width $18'$) to enforce periodicity required by Fourier transforms. The presence of such a window (as well as other sources of anisotropy such as inhomogeneity of instrument noise) induces a spurious lensing signal referred to as the *mean-field*. We estimate the mean-field by repeating the above stacking procedure on a large number of random locations (~300 times the number of clusters in our sample) and subtract this from our main cluster stack. The mean-field profile is roughly constant as a function of distance from the center of the stack and is 20% of the measurement in the first bin, 50% in the second bin, and larger than the measurement in subsequent bins.

We perform a number of systematics tests to ensure the robustness of our detection. This includes a curl null test, where the divergence in Equation (A1) is replaced by the curl (and normalized appropriately), shown in Figure 3. With this test, we obtain a measurement that is consistent with null (PTE of 0.05, corresponding to consistency with null at the 1.9σ level). As shown in Madhavacheril & Hill (2018), one needs correlated contaminants in the gradient and high-resolution maps of the quadratic estimator in order to be biased by these contaminants. Therefore, to test the possibility that dust emission might be biasing our measurement, we stack the Planck PR3 SMICA tSZ-deprojected map (used for the gradient) at the location of all 1676 MaDCoWS clusters with photometric redshifts (not just the 677 that fall in the ACT footprint used in this analysis). We detect no residual in this stack and thus conclude that dust contamination in the lensing reconstruction itself is unlikely. The presence of contaminants in the high-resolution map will contribute noise, which is captured in our empirically determined covariance matrix.

Appendix B
Mass Modeling

In order to infer the mean mass of the sample, we model the CMB lensing signal by assuming that all mass is contained within spherical halos, and that these halos cluster together (i.e., using the halo model; Peacock & Smith 2000; Seljak 2000). We adopt the best-fit flat ΛCDM cosmology from Planck Collaboration et al. (2016), with present-day matter density parameter $\Omega_\mathrm{m} = 0.307$ and a Hubble constant $H_0 = 67.7 \,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, also adopted by Gonzalez et al. (2019).

Within the halo model formalism, the lensing signal can be decomposed as

$$\kappa(\theta) = \kappa_{1\mathrm{h}}(\theta) + \kappa_{2\mathrm{h}}(\theta), \quad \text{(B1)}$$

where the subscripts 1h and 2h represent the one-halo (due to each halo) and two-halo (due to halo clustering) contributions. We model each component as follows.

For the one-halo term, we use an NFW density profile (Navarro et al. 1996). We use a mass definition, M_{500c}, corresponding to the mass within an overdensity of 500 times the critical density, $\rho_\mathrm{c}(z) = 3H^2(z)/8\pi G$. We model the

concentration, $c_{500c} \equiv r_{500c}/r_s$ (where r_s is the scale radius and r_{500c} is the radius containing M_{500c}), using the relation between concentration and mass from Diemer & Joyce (2019).[41] For a lens of mass m and redshift z, the convergence is related to the mass surface density, $\Sigma(m, z)$, through

$$\kappa(m, z) = \frac{\Sigma(m, z)}{\Sigma_{cr}(z)} \equiv \left(\frac{c^2}{4\pi G} \frac{D_s}{D_l(z)D_{ls}(z)} \right)^{-1} \Sigma(m, z), \quad (B2)$$

where D_s, D_l, and D_{ls} are the angular diameter distances to the last scattering surface, to the cluster, and between the cluster and the last scattering surface, respectively. The expression for the NFW surface density profile is given in Wright & Brainerd (2000).

The two-halo term arises due to the large-scale galaxy-matter power spectrum,

$$P_{gm}^{2h}(k|m, z) = b(m, z)P_m(k|z), \quad (B3)$$

where $P_m(k|z)$ is the linear matter power spectrum and $b(m, z)$ is the halo bias as calculated by Tinker et al. (2010). Note that all calculations of the two-halo term are performed in comoving coordinates. The power spectrum is the Fourier transform of the correlation function,

$$\xi_{gm}(r|m, z) = \frac{1}{2\pi^2} \int_0^\infty \frac{\sin kr}{kr} k^2 \, dk \, P_{gm}^{2h} \quad (B4)$$

from which the surface density can be calculated as

$$\Sigma_{2h}(R|m, z) = 2\bar{\rho}_m R \int_0^1 \frac{dx}{x^2\sqrt{1-x^2}} \xi_{gm}(R/x|m, z), \quad (B5)$$

where $\bar{\rho}_m$ is the (time-independent) comoving mean matter density. The two-halo term in Equation (B1) is then

$$\kappa_{2h}(\theta|m, z) = (1 + z)^2 \frac{\Sigma_{2h}(\theta|m, z)}{\Sigma_c}, \quad (B6)$$

where the additional $(1 + z)^2$ compared to Equation (B2) is due to our use of comoving coordinates in the two-halo term calculation (see, e.g., Dvornik et al. 2018). We then filter both the 1h and 2h components with the Fourier-space filter described in Appendix A to produce a prediction for $\hat{\kappa}(\theta|m, z)$.

In reality, our measurement is the mean over clusters covering a range in masses and redshifts. The halo mass function provides a prediction for the distribution of clusters in mass and redshift, which can be linked to the observable used to construct the sample (namely, richness, λ) through a mass-observable relation. However, like for any other cluster sample, due to the selection algorithm (in this case, red-sequence overdensities), MaDCoWS is a biased subsample of the underlying cluster population. We therefore follow previous galaxy–galaxy- and cluster-lensing studies (e.g., van den Bosch et al. 2013; Viola et al. 2015; Dvornik et al. 2018; Miyatake et al. 2019) and model the measured lensing signal as

$$\hat{\kappa}(\theta) = \frac{1}{\bar{N}} \int dz \, \frac{dV_{eff}}{dz} \int dm \, \frac{dn_h}{dm}$$
$$\times \int d\lambda \mathcal{P}(\ln\lambda|\ln m, z)\mathcal{S}(\lambda, z)\hat{\kappa}(\theta|m, z). \quad (B7)$$

Here $\frac{dV_{eff}}{dz} = A\frac{dN_{cl}}{dz}$ is the effective volume probed per unit redshift, which accounts for the unknown redshift component in the selection function (and A is a constant that absorbs the unit conversion, but cancels between the numerator and denominator); $\frac{dn_h}{dm}$ is the halo mass function from Tinker et al. (2010);[42] $\mathcal{P}(\ln\lambda|\ln m, z)$ is the conditional probability of richness given cluster mass (defined below), and $\bar{N}$ is the expected number of clusters, which differs from the halo mass function due to the mass-observable relation (see below) and selection function, $\mathcal{S}(\lambda, z)$, which is given by

$$\bar{N} = \int dz \, \frac{dV_{eff}}{dz} \int dm \, \frac{dn_h}{dm} \times \int d\lambda \mathcal{P}(\ln\lambda|\ln m, z)\mathcal{S}(\lambda, z). \quad (B8)$$

Consequently, in this framework mean values refer to the mean of the underlying cluster population and are calculated in analogy to Equation (B7) as

$$\langle X \rangle = \frac{1}{\bar{N}} \int dz \, \frac{dV_{eff}}{dz} \int dm \, \frac{dn_h}{dm}$$
$$\times \int d\lambda \mathcal{P}(\ln\lambda|\ln m, z)\mathcal{S}(\lambda, z) \, X. \quad (B9)$$

For simplicity, we assume that the MaDCoWS sample includes all clusters with richness larger than 20 (and we have discarded all clusters with lower richness), such that

$$\mathcal{S}(\lambda, z) = \mathcal{S}(\lambda) = \Theta(\lambda - 20), \quad (B10)$$

where Θ is the Heaviside step function. In addition, we limit the redshift integral to the range $z \in [0.7, 1.8]$. (Note that by using the observed redshift distribution, $\frac{dN_{cl}}{dz}$, instead of the comoving volume element dV/dz in Equation (B7), we are effectively introducing an additional redshift component to the selection function; this choice has no impact on our results.)

Additionally, we assume a linear relation between cluster richness, λ, and cluster mass, $M_{500c} \equiv m$:

$$\ln\lambda = \beta + \ln(m/10^{14} \, M_\odot) \equiv \beta + \ln\mu \quad (B11)$$

and assume a log-normal conditional probability of λ given cluster mass,

$$\mathcal{P}(\ln\lambda|\ln m, z) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{(\ln\lambda - \ln\mu - \beta)^2}{2\sigma^2} \right]. \quad (B12)$$

We assume a scatter on the mass-observable relation $\sigma \equiv \sigma_{\ln\lambda|\ln m} = 0.5$, typical of the scatter of different richness definitions (Rozo et al. 2014; Murata et al. 2019). The normalization β is the only free parameter in this model. While its posterior value depends on our assumptions about the selection function and the scatter above, the mean mass of the sample (calculated following Equation (B9)) is robust to these changes as it directly quantifies the amplitude of the lensing signal.

We estimate the best-fit parameters by varying the normalization β and minimizing

$$\chi^2 = (d - m)^T \cdot C^{-1} \cdot (d - m), \quad (B13)$$

---

[41] Calculated using COLOSSUS (Diemer 2018), https://bdiemer.bitbucket.io/colossus/.

[42] We use the Core Cosmology Library (Chisari et al. 2019, https://ccl.readthedocs.io/en/latest/) to calculate the halo mass function, the halo bias, and the matter power spectrum.

where d and m are the vectors of measured radial profile bins and the model expectations for them, respectively, and C is the covariance matrix of the measurement (see Equation (1)). We use five bins in our fit that encompass a region within $8'$ of the center of the stamp.

In order to demonstrate the contribution from each term in the model described above, we first fit our measurements with a single NFW profile at $z = 1.1$, including only the one-halo contribution, with the mass as the sole free parameter. This results in a best-fit mass $\langle M_{500c} \rangle = (2.1 \pm 0.5) \times 10^{14}\, M_\odot$. We then add the two-halo term, assuming all clusters have the same mass and are located at the same mean redshift, and find $\langle M_{500c} \rangle = (1.7 \pm 0.4) \times 10^{14}\, M_\odot$. Finally, we implement the full model as described above. We find $\beta = 2.4 \pm 0.3$, which translates to a population-weighted mean cluster mass $\langle M_{500c} \rangle = (1.7 \pm 0.4) \times 10^{14}\, M_\odot$, where uncertainties correspond to $\Delta\chi^2 = 1$ (i.e., 68.3% credible range). This is the main result of this paper.

For reference, we calculate the mass M_{200c} within an overdensity of 200 times the critical density for each point in the (M_{500c}, z) grid and perform the integral given by Equation (B9) over the appropriate halo mass function, and find $\langle M_{200c} \rangle = (2.5 \pm 0.6) \times 10^{14}\, M_\odot$.

As discussed above, the best-fit β depends strongly on our modeling assumptions and should not be over-interpreted, but the mean mass changes at most by 5% regardless of our assumptions about the selection function and scaling relation. Specifically, we have varied the scaling relation normalization, β, and intrinsic scatter, σ, by factors of two from the adopted values, as well as selection functions given both by step functions and error functions with midpoints in the range $\lambda = [10, 30]$, and in all cases find mean masses within 5% of the reported value. The largest shift in mass is introduced by modifying the mass-concentration relation (for reference, the population-weighted mean concentration is $c_{500c} = 2.5$): increasing (decreasing) the concentration by 20%—which brackets most mass-concentration relations in the literature—increases (decreases) the mean mass by 10%, well within the uncertainties. Redshift uncertainties also do not change our results significantly: applying a systematic shift $\Delta z = \pm\, 0.1$ (slightly larger than the 1σ scatter of photometric redshifts compared to spectroscopic redshifts, Gonzalez et al. 2019) only changes the mean mass by less than 3%. Assuming that 5% of the clusters have photometric redshifts that are wrong by $\Delta z = +0.2$ (3σ outliers, all in the same direction) has a similar impact.

Our final assessment of modeling systematics regards miscentering. The MaDCoWS algorithm defines the cluster center as the location of the peak amplitude in a smoothed galaxy density map, which can potentially be significantly offset from the center of mass (George et al. 2012; Viola et al. 2015) and results in a reduced central lensing amplitude. However, the physical resolution of our measurements is roughly 1 Mpc, which greatly reduces the impact from miscentering, and we choose not to account for it in our analysis. To test the impact of miscentering, we implement the distribution found by Johnston et al. (2007) for optically selected clusters, exploring typical offsets between the assumed and true centers of up to 0.4 Mpc (roughly $2r_s$); the change in mass is at most 5%.

ORCID iDs

Mathew S. Madhavacheril ⓘ https://orcid.org/0000-0001-6740-5350

Cristóbal Sifón ⓘ https://orcid.org/0000-0002-8149-1352
Steve K. Choi ⓘ https://orcid.org/0000-0002-9113-7058
Simon R. Dicker ⓘ https://orcid.org/0000-0002-1940-4289
Simone Ferraro ⓘ https://orcid.org/0000-0003-4992-7854
Matt Hilton ⓘ https://orcid.org/0000-0002-8490-8117
Kevin M. Huffenberger ⓘ https://orcid.org/0000-0001-7109-0099
John P. Hughes ⓘ https://orcid.org/0000-0002-8816-6800
Arthur Kosowsky ⓘ https://orcid.org/0000-0002-3734-331X
Toshiya Namikawa ⓘ https://orcid.org/0000-0003-3070-9240
Neelima Sehgal ⓘ https://orcid.org/0000-0002-9674-4527
David N. Spergel ⓘ https://orcid.org/0000-0002-5151-0006
Edward J. Wollack ⓘ https://orcid.org/0000-0002-7567-4451

References

Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, ARA&A, 49, 409
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Baxter, E. J., Keisler, R., Dodelson, S., et al. 2015, ApJ, 806, 247
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
Chisari, N. E., Alonso, D., Krause, E., et al. 2019, ApJS, 242, 2
Chiu, I. N., Umetsu, K., Murata, R., Medezinski, E., & Oguri, M. 2020, MNRAS, 495, 428
Dicker, S. R., Romero, C. E., Di Mascolo, L., et al. 2020, arXiv:2006.06703
Diemer, B. 2018, ApJS, 239, 35
Diemer, B., & Joyce, M. 2019, ApJ, 871, 168
Di Mascolo, L., Mroczkowski, T., Churazov, E., et al. 2020, A&A, 638, A70
Dvornik, A., Hoekstra, H., Kuijken, K., et al. 2018, MNRAS, 479, 1240
Geach, J. E., & Peacock, J. A. 2017, NatAs, 1, 795
George, M. R., Leauthaud, A., Bundy, K., et al. 2012, ApJ, 757, 2
Gonzalez, A. H., Gettings, D. P., Brodwin, M., et al. 2019, ApJS, 240, 33
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Henderson, S. W., Allison, R., Austermann, J., et al. 2016, JLTP, 184, 772
Hilton, M., Sifón, C., Naess, S., et al. 2020, arXiv:2009.11043
Hoekstra, H., Bartelmann, M., Dahle, H., et al. 2013, SSRv, 177, 75
Hu, W., DeDeo, S., & Vale, C. 2007, NJPh, 9, 441
Jee, M. J., Dawson, K. S., Hoekstra, H., et al. 2011, ApJ, 737, 59
Johnston, D. E., Sheldon, E. S., Wechsler, R. H., et al. 2007, arXiv:0709.1159
Madhavacheril, M., Sehgal, N., Allison, R., et al. 2015, PhRvL, 114, 151302
Madhavacheril, M., Sehgal, N., & Battaglia, N. 2017, PhRvD, 96, 103525
Madhavacheril, M. S., & Hill, J. C. 2018, PhRvD, 98, 023534
Massey, R., Kitching, T., & Richard, J. 2010, RPPh, 73, 086901
Miyatake, H., Battaglia, N., Hilton, M., et al. 2019, ApJ, 875, 63
Murata, R., Oguri, M., Nishimichi, T., et al. 2019, PASJ, 71, 107
Naess, S., Aiola, S., Austermann, J. E., et al. 2020, arXiv:2007.07290
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Patil, S., Raghunathan, S., & Reichardt, C. L. 2020, ApJ, 888, 9
Peacock, J. A., & Smith, R. E. 2000, MNRAS, 318, 1144
Planck Collaboration, Ade, P. A. R. 2016, A&A, 594, A24
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Planck Collaboration, Akrami, Y., Ashdown, M., et al. 2020, A&A, 641, A4
Raghunathan, S., Patil, S., Baxter, E., et al. 2019, ApJ, 872, 170
Rozo, E., Bartlett, J. G., Evrard, A. E., & Rykoff, E. S. 2014, MNRAS, 438, 78
Schrabback, T., Applegate, D., Dietrich, J. P., et al. 2018, MNRAS, 474, 2635
Seljak, U. 2000, MNRAS, 318, 203
Swetz, D. S., Ade, P. A. R., Amiri, M., et al. 2011, ApJS, 194, 41
Thornton, R. J., Ade, P. A. R., Aiola, S., et al. 2016, ApJS, 227, 21
Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878
Trimble, V. 1987, ARA&A, 25, 425
van den Bosch, F. C., More, S., Cacciato, M., Mo, H., & Yang, X. 2013, MNRAS, 430, 725
Viola, M., Cacciato, M., Brouwer, M., et al. 2015, MNRAS, 452, 3529
Voit, G. M. 2005, RvMP, 77, 207
Wright, C. O., & Brainerd, T. G. 2000, ApJ, 534, 34
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Zonca, A., Singer, L., Lenz, D., et al. 2019, JOSS, 4, 1298
Zubeldia, I., & Challinor, A. 2019, MNRAS, 489, 401