Version September 10, 2025
Preprint typeset using LaTeX style openjournal v. 09/06/15

MASS MODELS OF GALAXY CLUSTERS FROM A NON-PARAMETRIC WEAK-LENSING RECONSTRUCTION

Tobias Mistele[1], Federico Lelli[2], Stacy McGaugh[1], James Schombert[3], Benoit Famaey[4]
[1]Department of Astronomy, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106, USA
[2]INAF – Arcetri Astrophysical Observatory, Largo Enrico Fermi 5, 50125 Firenze, Italy
[3]Department of Physics, University of Oregon, 1371 E 13th Ave, Eugene, Oregon 97403, USA
[4]Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, F-67000 Strasbourg, France
Version September 10, 2025

ABSTRACT

We study the CLASH sample of galaxy clusters using a new deprojection method for weak gravitational lensing observations. This method is non-parametric, allowing us to infer mass profiles, or equivalently circular velocities, without having to assume a specific halo profile. While this method assumes spherical symmetry, we show that, on average, triaxiality is unlikely to significantly affect our results. We use this method to study the total mass profiles of the CLASH clusters, as well as the relation between their total and baryonic components: (1) We find that the implied circular velocities are consistent with being approximately flat at large radii, akin to the rotation curves of galaxies. (2) We infer radially resolved baryonic mass fractions, finding that these vary significantly from cluster to cluster and depend strongly on the details of the X-ray gas mass profiles. Since the gas mass profiles are poorly constrained at large radii, it is unclear whether the CLASH clusters reach the cosmic baryon fraction expected in ΛCDM. (3) The non-parametric masses are consistent with the stellar mass–halo mass relation expected in ΛCDM. (4) Galaxy clusters systematically deviate from the Baryonic Tully-Fisher Relation (BTFR) and the Radial Acceleration Relation (RAR) defined by galaxies, but the magnitude of the offset depends strongly on the gas mass extrapolation at large radii. Contrary to some previous results based on hydrostatic equilibrium, we find that galaxy clusters may fall on the same BTFR and RAR as galaxies if one adds a suitable positive baryonic mass component.

1. INTRODUCTION

Galaxy clusters are important astrophysical and cosmological probes. For example, in cosmological models such as ΛCDM, their abundance and masses constrain key parameters such as Ω_m and σ_8. They can also constrain models of dark matter and modified gravity, for example through features in their density profiles such as the splashback radius (e.g. Diemer & Kravtsov 2014; More et al. 2015, 2016; Adhikari et al. 2018) or through scaling relations that connect their baryonic and dynamical mass distributions (e.g. Sanders 2003; Tian et al. 2020; Eckert et al. 2022; Li et al. 2023, 2024; Famaey et al. 2025; Kelleher & Lelli 2024).

Scaling relations also play an important role in galaxies. Indeed, galaxies follow tight scaling relations such as the Baryonic Tully-Fisher Relation (BTFR, McGaugh et al. 2000; Mistele et al. 2024a) and the Radial Acceleration Relation (RAR, Lelli et al. 2017; Brouwer et al. 2021; Mistele et al. 2024b). These were predicted a priori by Modified Newtonian Dynamics (MOND, Milgrom 1983a,b,c, see Famaey & Durakovic 2025 for a recent review), but are not easily explained in ΛCDM because they must emerge from the complex and stochastic process of galaxy formation. However, MOND-inspired theories have historically struggled to explain why galaxy clusters do not seem to follow the same scaling relations as galaxies (Sanders 1999, 2003). Relativistic extension of MOND such as Aether-Scalar Tensor Theory (AeST, Skordis & Złosnik 2021) or Relativistic Khronon Theory (Blanchet & Skordis 2024) may ameliorate these issues, since they predict deviations from scaling relations such as the RAR at large masses (Mistele et al. 2023a; Durakovic & Skordis 2024), but this has not yet been demonstrated to work in quantitative detail.

All these different constraints require dynamical masses of galaxy clusters. A powerful tool to measure these is weak gravitational lensing. Indeed, unlike X-ray and galaxy kinematics measurements, weak lensing does not require hydrostatic or dynamical equilibrium. However, most existing weak-lensing measurements are based on fitting weak-lensing observations to parametric profiles such as the Navarro-Frenk-White (NFW, Navarro et al. 1996) profile, assuming a specific shape of the mass profile (but see Johnston et al. 2007; Umetsu et al. 2011, 2025). An alternative is the non-parametric method from Mistele & Durakovic (2024) which does not presume a specific profile, allowing for less biased and more model-independent measurements.

In the following we will use this method to infer non-parametric mass profiles for galaxy clusters from the Cluster Lensing and Supernova Survey with Hubble (CLASH, Postman et al. 2012) and study their properties, including the relation between the dynamical and baryonic mass components. In Sec. 2, we discuss the data we use with the methods described in Sec. 3. We present our results in Sec. 4 and after a brief discussion

in Sec. 5 we conclude in Sec. 6.

2. DATA

We consider a subset of 20 galaxy clusters from the CLASH project for which weak-lensing observations, primarily from the Subaru Suprime-Cam, are available (Umetsu et al. 2014). Of these, 16 were originally X-ray selected to be massive ($kT > 5\,\mathrm{keV}$, listed first in Table 1) and 4 were selected for their high lensing strength.

2.1. *Weak lensing data*

We adopt the azimuthally-averaged reduced shear profiles $\langle g_+ \rangle$, cluster redshifts z_l, and the corresponding critical surface densities from Umetsu et al. (2014). The shear profiles are given as functions of angular distances, which we convert to projected radii R assuming a flat ΛCDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_m = 0.27$, following Umetsu et al. (2014). We adopt this cosmology throughout this work.

The critical surface density of a given lens-source pair is defined as,

$$\Sigma^{-1}_{\mathrm{crit},ls} = \frac{4\pi G_\mathrm{N}}{c^2}\frac{D(z_l)D(z_l,z_s)}{D(z_s)} \equiv \frac{4\pi G_N}{c^2} D(z_l)\cdot\beta\,, \quad (1)$$

where G_N is the Newtonian gravitational constant and $D(z_l)$, $D(z_s)$, and $D(z_l, z_s)$ are the angular diameter distances to the lens, to the source, and between the source and the lens, respectively. For our weak-lensing analysis below, we need $\langle\Sigma^{-1}_{\mathrm{crit},ls}\rangle$ and $\langle\Sigma^{-2}_{\mathrm{crit},ls}\rangle$ where $\langle\ldots\rangle$ denotes averaging over source galaxies (see Sec. 3.1). Umetsu et al. (2014) provide estimates for $\langle\beta\rangle$ as well as $\langle\beta^2\rangle/\langle\beta\rangle^2$ from which we infer the desired quantities using the cluster redshift z_l and the assumed cosmology.

Our weak-lensing analysis makes use of the shear profiles in the form $G_+ = \langle g_+\rangle/\langle\Sigma^{-1}_{\mathrm{crit},ls}\rangle$, which we show in Fig. 1. The shear $\langle g_+\rangle$ varies as a function of projected radius while, following Umetsu et al. (2014), $\langle\Sigma^{-1}_{\mathrm{crit},ls}\rangle$ is assumed to be constant within a given cluster. The uncertainties shown in Fig. 1 are discussed in Sec. 3.3.

2.2. *Baryonic mass estimate*

We adopt the baryonic mass estimates from Famaey et al. (2025). These are available for 16 out of 20 clusters in our sample (see Table 1) to which we restrict ourselves whenever baryonic masses are required.

Famaey et al. (2025) estimate the baryonic mass profile $M_b(r)$ as a sum of the dominant contribution from the intracluster medium, $M_\mathrm{gas}(r)$, and further contributions from the brightest cluster galaxy (BCG) including companions within 50 kpc, M_{*,BCG^+}, (Burke et al. 2015) as well as contributions from other galaxies, $M_\mathrm{gal}(r)$.

Since we are only interested in radii larger than a few hundred kpc, we treat M_{*,BCG^+} as a point mass. The galaxy contribution is given as a fraction f_gal of M_gas, i.e. $M_\mathrm{gal}(r) = f_\mathrm{gal}(r)\cdot M_\mathrm{gas}(r)$. Following Famaey et al. (2025), this fraction f_gal is assumed to be the same across all galaxy clusters when normalized to the radius r_{200c}[1]

[1] In the following, M_{200c} refers to the mass within the radius r_{200c} where the galaxy cluster's average mass density drops to 200 times the critical density at the cluster's redshift.

TABLE 1
PROPERTIES OF THE CLASH CLUSTERS WE USE AND QUANTITIES INFERRED FROM OUR WEAK-LENSING ANALYSIS.

Name	z_l	R^X_max Mpc	$\log_{10} M^{2h\,\mathrm{sub}}_{200c}$ $M_\odot$	$\log_{10} M_{200c}$ $M_\odot$
Abell 383	0.187	–	–	–
Abell 209	0.206	0.651	15.26 ± 0.11	15.29 ± 0.12
Abell 2261	0.224	0.825	15.25 ± 0.11	15.28 ± 0.12
RXJ2129	0.234	0.591*	14.93 ± 0.18	14.96 ± 0.18
Abell 611	0.288	0.884	15.11 ± 0.15	15.15 ± 0.38
MS2137	0.313	0.227	15.05 ± 0.08	15.07 ± 0.08
RXJ2248	0.348	0.904	15.14 ± 0.15	15.17 ± 0.16
MACSJ1115	0.352	0.375	15.27 ± 0.18	15.32 ± 0.18
MACSJ1931	0.352	–	15.27 ± 0.46	15.33 ± 0.31
RXJ1532	0.363	–	14.58 ± 0.24	14.60 ± 0.25
MACSJ1720	0.391	0.532	15.04 ± 0.15	15.07 ± 0.11
MACSJ0429	0.399	0.292	15.02 ± 0.23	15.06 ± 0.24
MACSJ1206	0.440	0.930	15.17 ± 0.17	15.22 ± 0.17
MACSJ0329	0.450	0.377	14.79 ± 0.20	14.83 ± 0.23
RXJ1347	0.451	0.887	15.41 ± 0.11	15.44 ± 0.12
MACSJ0744	0.686	0.884	15.17 ± 0.13	15.19 ± 0.14
MACSJ0416	0.396	0.449	14.92 ± 0.09	14.94 ± 0.09
MACSJ1149	0.544	0.557	15.38 ± 0.10	15.42 ± 0.10
MACSJ0717	0.548	–	15.32 ± 0.08	15.35 ± 0.08
MACSJ0647	0.584	0.381	14.94 ± 0.13	14.97 ± 0.13

NOTE. — R^X_max denotes the radial extent of the X-ray data used in Donahue et al. (2014). *The exception is RX J2129 where Famaey et al. (2025) redid the fit with a smaller radial range than the R^X_max listed in Donahue et al. (2014). Missing R^X_max values indicate that a baryonic mass estimate from Famaey et al. (2025) is not available. The masses M_{200c} are missing for Abell 383 due to the uptick in the non-parametric mass profile (see Fig. 6).

(that we measure from weak gravitational lensing, see below).[2] Values of f_gal typically reach $O(1)$ at $r \lesssim 0.1\,\mathrm{Mpc}$ and drop to about 8% or 12% beyond r_{200c}, depending on how we extrapolate the gas density profiles (see below). This simple procedure doesn't allow precise quantitative statements about M_b, but suffices for our purposes.

The dominant contribution $M_\mathrm{gas}(r)$ is obtained by fitting double-beta profiles to X-ray observations,

$$\rho_\mathrm{gas}(r) = n_0\left(\frac{r}{r_0}\right)^{-\alpha}\left(1+\left(\frac{r}{r_{e,0}}\right)^2\right)^{-\frac{3\beta_0}{2}} + n_1\left(1+\left(\frac{r}{r_{e,1}}\right)^2\right)^{-\frac{3\beta_1}{2}}, \quad (2)$$

with free parameters n_0, r_0, α, $r_{e,0}$, β_0, n_1, $r_{e,1}$, and β_1. We adopt the fit results from Famaey et al. (2025). Most of these are originally from Laudato et al. (2022) based on data from Donahue et al. (2014), but Famaey et al. (2025) redid them for a few clusters for which the implied gas masses were unreasonably small.

The underlying Chandra X-ray observations to which these beta profiles were fit become noisy at large radii. Therefore, Donahue et al. (2014) considered only data out to a maximum radius R^X_max, which is determined by requiring at least 1500 counts of photon signal in each radial bin. In the following, we assume that the beta

[2] Specifically, Famaey et al. (2025) measure f_gal for MACS J1206 and adopt $f_\mathrm{gal}(r) = f_\mathrm{gal,J1206}\left(r\cdot r_{200c,\mathrm{J1206}}/r_{200c}\right)$ for the other clusters. Beyond the last measured data point, $f_\mathrm{gal,J1206}$ is assumed to remain constant at the last measured value.

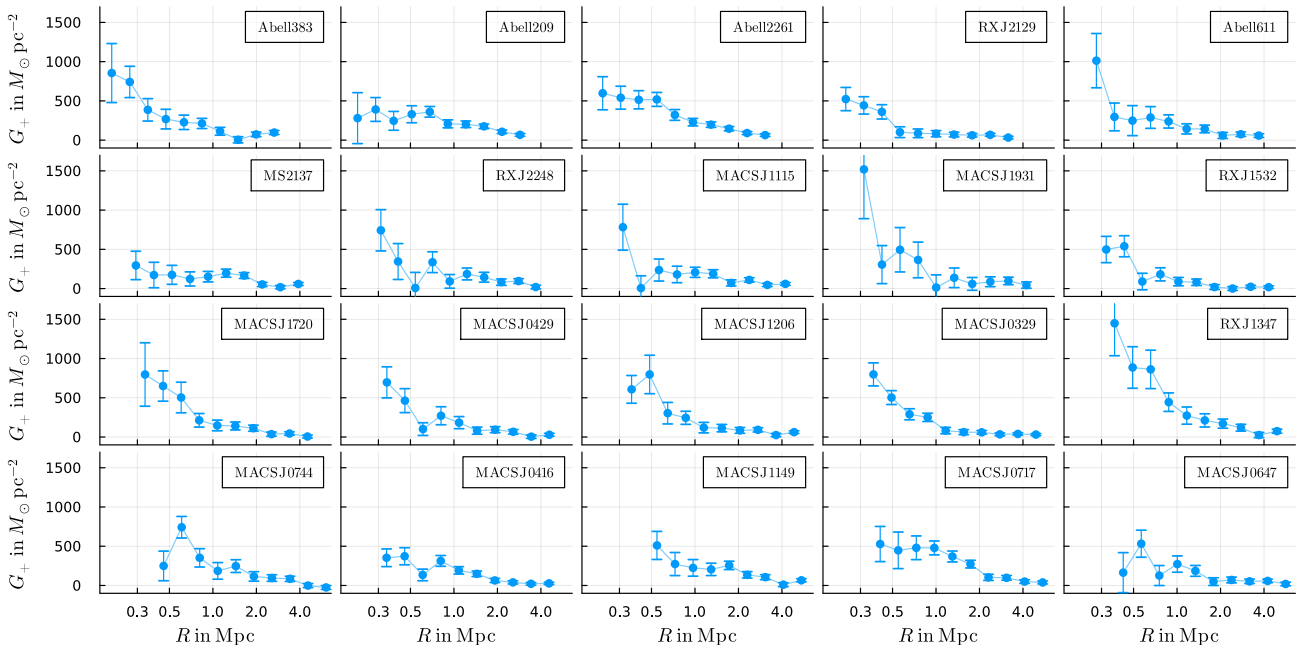

FIG. 1.— The reduced shear in terms of $G_+ = \langle g_+ \rangle / \langle \Sigma_{\rm crit}^{-1} \rangle$ for the 20 CLASH clusters from Umetsu et al. (2014). The error bars include contributions from the reduced shear $\langle g_+ \rangle$, the inverse critical surface density $\langle \Sigma_{\rm crit}^{-1} \rangle$, and from the LSS, see Sec. 3.3.

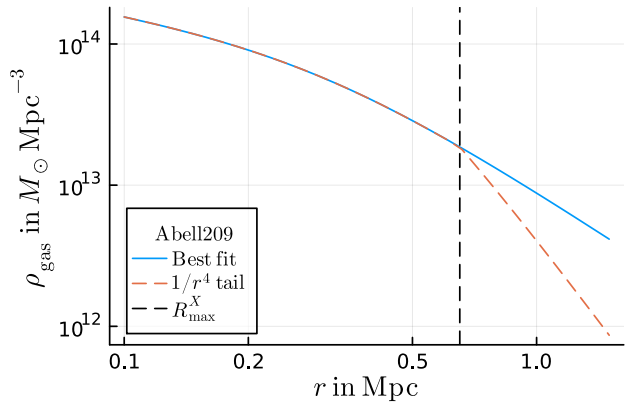

FIG. 2.— The gas density of Abell 209 implied by the double beta fit from Famaey et al. (2025). The vertical dashed line indicates $R^X_{\rm max}$, i.e. how far out this fit is reliable. We show two different ways of extrapolating beyond $R^X_{\rm max}$: Assuming the best-fit parameters to be valid even beyond $R^X_{\rm max}$ (solid blue line) and assuming a $1/r^4$ tail (dashed red line).

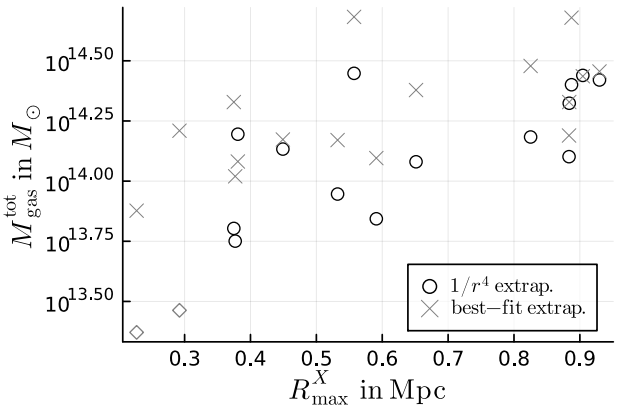

FIG. 3.— The radius $R^X_{\rm max}$ out to which the double beta profile fits of the gas densities are well constrained by observations versus the total gas mass. Circles correspond to assuming a $1/r^4$ density tail after $R^X_{\rm max}$. The two clear outliers, MS 2137 and MACS J0429, are shown as gray diamonds. Crosses correspond to assuming that the best-fit beta profiles remain valid even beyond $R^X_{\rm max}$.

profile fits are reliable up to $R^X_{\rm max}$[3]. At larger radii, beyond $R^X_{\rm max}$, we consider two options: 1) We assume that the best fit parameters remain valid even at large radii where they were not well constrained by observations and 2) we assume that the beta profiles are matched to a $1/r^4$ tail at $r = R^X_{\rm max}$.

The $1/r^4$ tail has two motivations. If we require a finite total gas mass, we need a gas density that asymptotically decays faster than $1/r^3$ so $1/r^4$ is perhaps a natural choice. Also, $1/r^4$ is the asymptotic behavior of isothermal spheres in MOND (Milgrom 1984).

The $1/r^4$ matching procedure is illustrated in Fig. 2. Beyond $R^X_{\rm max}$, we set $\beta_0 = (4-\alpha)/3$ and $\beta_1 = 4/3$ to ensure an asymptotic $1/r^4$ decay and we adjust n_0 and n_1 such that the n_0 and n_1 components of $\rho_{\rm gas}$ are both continuous. In practice, $R^X_{\rm max}$ is quite small for some clusters (see Table 1). Thus, even if $1/r^4$ is the correct asymptotic behavior, one may worry that $R^X_{\rm max}$ is not yet in that asymptotic regime, so that matching to a $1/r^4$ tail at $R^X_{\rm max}$ may underestimate the true gas mass. In that case one would expect our $1/r^4$ extrapolation to artificially induce a positive correlation between gas mass and $R^X_{\rm max}$.

[3] In the fits they redid, Famaey et al. (2025) did not impose the 1500 photon count requirement and so used X-ray observations out to larger radii for some clusters. However, since this additional data is quite noisy, it still makes sense to adopt $R^X_{\rm max}$ from Donahue et al. (2014) as an estimate of how far out to trust the beta profile fits.

To assess this effect, the white circles in Fig. 3 show the total gas mass implied by our $1/r^4$ matching procedure versus the radial range of the X-ray observations $R^X_{\rm max}$. For comparison, the gray crosses in Fig. 3 correspond to assuming that the best-fit beta profiles remain valid even beyond $R^X_{\rm max}$ with the gas mass evaluated at $r = r_{200c}$ (see Sec. 4.5). We see that assuming a $1/r^4$ tail induces only a weak additional correlation between the total gas mass and $R^X_{\rm max}$, though there are two clear outliers MS 2137 and MACS J0429, which have the smallest gas masses in our sample, and two borderline cases MACS J0329 and MACS J1115, with the 3rd and 4th lowest gas masses. Quantitatively, with our $1/r^4$ extrapolation, Pearson's r is 0.82 for the full sample and drops to 0.43 with the two outliers and the two borderline cases removed. With the best-fit extrapolation, Pearson's r is 0.66. This suggests that, if $1/r^4$ is indeed the correct asymptotic behavior, it is overall not unreasonable to match to a $1/r^4$ tail at $R^X_{\rm max}$. The gas masses for the two outliers are, however, likely to be significantly underestimated and we consider them separately below. The gas masses of clusters with $R^X_{\rm max}$ similar to the two borderline cases may also be somewhat underestimated but we do not treat them separately.

Recently, the XMM Cluster Outskirts Project (X-COP, Eckert et al. 2017; Ghirardini et al. 2019) has constrained the gas densities of 12 galaxy clusters out to about r_{200c} (for earlier measurements at large radii using the Suzaku satellite, see Walker et al. 2013, 2012; Urban et al. 2014). They find density profiles that steepen with radius. In fact, their density profiles become steeper than the asymptotic behavior of many of the best-fit beta profiles from Famaey et al. (2025), though not as steep as $1/r^4$. This suggests that the true asymptotic behavior of the gas profiles is in between the two extrapolations we use. In any case, the point of adopting multiple different extrapolations is to illustrate a range of possibilities, not to give a precise quantitative result.

3. METHOD

3.1. *Weak-lensing mass profile measurements*

Weak-lensing observations of a lens l, in our case a galaxy cluster, are based on the shapes of many background source galaxies s. From these shapes and the position angle between the sources and the lens, one can infer the azimuthally-averaged tangential reduced shear $\langle g_t \rangle$ (Bartelmann & Schneider 2001), which encodes the cluster's projected mass distribution. Indeed, assuming spherical symmetry, we have to a good approximation (Umetsu 2020; Mistele & Durakovic 2024)

$$G_+(R) = \frac{\Delta\Sigma(R)}{1 - f_c(R)\Sigma(R)}\,, \tag{3}$$

where R is the projected radius, G_+ is the azimuthally-averaged reduced tangential shear divided by the azimuthally-averaged inverse critical surface density,

$$G_+ \equiv \frac{\langle g_+ \rangle}{\langle \Sigma^{-1}_{{\rm crit},ls} \rangle}\,, \tag{4}$$

and f_c refers to the following ratio of azimuthal averages of powers of the critical surface density,

$$f_c = \frac{\langle \Sigma^{-2}_{{\rm crit},ls} \rangle}{\langle \Sigma^{-1}_{{\rm crit},ls} \rangle}\,. \tag{5}$$

Further, Σ and $\Delta\Sigma$ are, respectively, the surface mass density and the excess surface mass density of the cluster. The excess surface density is defined as

$$\Delta\Sigma(R) = \frac{M_{2D}(R)}{\pi R^2} - \Sigma(R)\,, \tag{6}$$

where $M_{2D}(R)$ is the mass enclosed by a cylinder with radius R that is oriented along the line of sight.

We will be interested in the deprojected, 3D mass profile $M(r)$. To convert observations of the shear G_+ and the critical surface density f_c into the 3D mass M, we assume spherical symmetry and use the deprojection technique of Mistele & Durakovic (2024) (see also Mistele et al. 2024b,a). This technique is based on Eq. (3) and consists of two steps. First, we convert G_+ and f_c into an excess surface density $\Delta\Sigma$,

$$\Delta\Sigma(R) = \frac{G_+(R)}{1 - f_c G_+(R)} \times \exp\left(- \int_R^\infty dR' \frac{2}{R'} \frac{f_c G_+(R')}{1 - f_c G_+(R')} \right)\,. \tag{7}$$

This step is only important at relatively small radii where the $f_c\Sigma$ term in the denominator of Eq. (3) is not negligible. This roughly corresponds to $\Sigma/\Sigma_{\rm crit}$, also known as the convergence κ, not being negligible. In contrast, at large radii where $\Sigma/\Sigma_{\rm crit}$ is small, Eq. (7) reduces to $G_+ = \Delta\Sigma$. The second step is to convert $\Delta\Sigma(R)$ from Eq. (7) into the 3D mass profile $M(r)$,

$$M(r) = 4r^2 \int_0^{\pi/2} d\theta\, \Delta\Sigma\left(\frac{r}{\sin\theta}\right)\,. \tag{8}$$

Equation (7) is mathematically valid as long as $G_+ f_c < 1$. This roughly corresponds to the condition that we are in the weak-lensing regime (Mistele & Durakovic 2024). Equation (7) further assumes that the critical surface density, i.e. f_c, is constant as a function of projected radius R. Mistele & Durakovic (2024) also provide formulas for the case with a radially varying f_c. A constant f_c is, however, often a reasonable assumption in practice and, following Umetsu et al. (2014), we adopt that assumption here.

For the integrals in the above deprojection formulas, we need to know the function $G_+(R)$ at all radii out to infinity, whereas, in practice, we measure G_+ only in a discrete set of radial bins out to some outermost bin with bin center $R_{\rm max}$. Thus, we must interpolate between the discrete radial bins and extrapolate beyond $R_{\rm max}$. In practice, the effect of the interpolation on the inferred $M(r)$ is quite minor and the extrapolation is unimportant except when r is close to the last measured data point at $R_{\rm max}$. As a result, the choices we make about how to extrapolate and interpolate are unimportant over much of the radial range we consider (Mistele & Durakovic 2024; Mistele et al. 2024b,a).

To be specific, we extrapolate assuming that G_+ follows a $1/R$ power law beyond $R_{\rm max}$ and we linearly in-

terpolate the discrete G_+ measurements in logarithmic space. We take the uncertainty in these choices into account as systematic uncertainties, see Sec. 3.3. In Appendix E, we show that extrapolating G_+ assuming an NFW profile instead of a $1/R$ power law does not significantly change our results, further confirming that our systematic uncertainty estimate is reasonable.

The above procedure is straightforward to implement numerically and runs fast, taking only a few milliseconds per galaxy cluster. We use the code provided by Mistele & Durakovic (2024)[4] to implement Eq. (7) and Eq. (8) as well as the propagation of uncertainties and covariances discussed in Sec. 3.3 below.

3.2. *Two-halo term*

The deprojection technique discussed in Sec. 3.1 assumes that all of the lensing signal is due to the cluster itself. This is a good approximation at small and moderate radii. Beyond a few Mpc, however, the signal from the cluster's local environment, the so-called two-halo term, can become important (Umetsu 2020; Oguri & Hamana 2011; Oguri & Takada 2011).

Within ΛCDM, we can estimate this contribution and subtract it. The resulting subtracted mass profiles should be a better estimate of the true mass profiles, but are more model-dependent. Below we present results with and without this two-halo subtraction, but the difference turns out not to be important for our purposes. We explain the details of our two-halo subtraction procedure in Appendix A.

3.3. *Uncertainties and covariances*

We consider two sources of systematic uncertainties, corresponding to our choices of how to extrapolate G_+ beyond the last measured data point and of how to interpolate between the discrete radial bins. For the statistical uncertainties, we take into account uncertainties in the shear measurements $\langle g_+ \rangle$, as well as uncertainties and covariances from the inverse critical surface densities $\langle \Sigma_{\rm crit}^{-1} \rangle$, and covariances induced by the large-scale structure (LSS, Hoekstra 2003). We use linear error propagation to propagate the statistical uncertainties and covariances into the reconstructed mass profiles. As a result, any shortcomings in our uncertainty and covariance estimates affect only the error bars of the reconstructed mass profiles; they do not enter the central values. The details are discussed in Appendix B.

3.4. *Non-parametric density reconstruction*

One way to obtain a 3D density profile $\rho(r)$ in a non-parametric way is to first reconstruct the 3D mass profile $M(r)$ following Sec. 3.1 and then take a numerical derivative, $\rho(r) = M'(r)/4\pi r^2$. However, numerical derivatives can be tricky. Thus, we here introduce a density reconstruction method that goes directly from shear to density without any numerical derivatives.

Specifically, as we show in Appendix C, assuming spherical symmetry we have[5]

$$\rho(r) = \frac{I_1(r) + I_2(r)}{\pi r}\,, \tag{9}$$

with

$$I_1(r) = \int_0^{\pi/2} d\theta\, \frac{\Delta\Sigma\left(\frac{r}{\sin\theta}\right) - \Delta\Sigma(r)}{\cos^2\theta}\,, \tag{10}$$

$$I_2(r) = 2\int_0^{\pi/2} d\theta\, \Delta\Sigma\left(\frac{r}{\sin\theta}\right)\,. \tag{11}$$

This replaces the second step Eq. (8) in the reconstruction method from Sec. 3.1, i.e. the step that converts $\Delta\Sigma$ to M. The first step, Eq. (7), which converts the shear G_+ to $\Delta\Sigma$, remains the same. The integrand of Eq. (10) is finite everywhere despite the $1/\cos^2\theta$ factor. This can be seen by expanding around $\theta = \pi/2$.

While this non-parametric density reconstruction does not involve numerical derivatives, it still requires relatively low-noise data for good results due to the finite difference $\Delta\Sigma(r/\sin\theta) - \Delta\Sigma(r)$ in the numerator in Eq. (10). Indeed, the $1/\cos^2\theta$ up-weights precisely the region $\theta \approx \pi/2$ of the integral where $\Delta\Sigma(r/\sin\theta)$ and $\Delta\Sigma(r)$ are close to each other so that the difference becomes small. Despite this, initial tests with mock data show that, at least in some situations, the non-parametric density reconstruction method Eq. (9) is better behaved than taking a numerical derivative of the non-parametric mass profile from Sec. 3.1.

It should be possible to adjust the two-halo subtraction procedure from Sec. 3.2 to work with this non-parametric density reconstruction method. This is left for dedicated follow-up work.

3.5. *On the assumption of spherical symmetry*

Our non-parametric deprojection method assumes spherical symmetry. As we will now discuss, this assumption can be relaxed in some important ways, making the method more widely applicable and more robust.

First, consider clusters that are approximately spherically symmetric at large radii but whose inner regions are more complex. Examples may be late-stage mergers or clusters with a complex baryonic mass distribution. As expected, our deprojection method does not work well in the inner regions of such clusters due to the lack of symmetry. However, perhaps surprisingly, the method does infer the correct enclosed mass at larger radii, outside the non-symmetric core (Mistele & Durakovic 2024). Thus, unlike methods based on fitting NFW or similar profiles, our method will not be thrown off by complexity or non-symmetry at small radii.

A related useful property of our deprojection method is that mass measurements at large radii (beyond a few times the miscentering offset) are not affected much by miscentering. This is because the effect of miscentering is essentially to induce non-symmetry at small radii, while keeping approximate spherical symmetry at large radii.

[4] https://github.com/tmistele/SphericalClusterMass.jl

[5] The factor of $1/\pi$ may be reminiscent of the Abel transform, $\rho(r) = -\pi^{-1}\int_r^\infty dR\,\Sigma'(R)/\sqrt{R^2 - r^2}$. Indeed, the Abel transform can be written in an alternative way that looks quite similar to Eq. (9) and contains no derivatives of Σ: $\rho(r) = -(\pi r)^{-1}\int_0^{\pi/2} d\theta\, \frac{1}{\cos^2\theta}\left(\Sigma\left(\frac{r}{\sin\theta}\right) - \Sigma(r)\right)$.

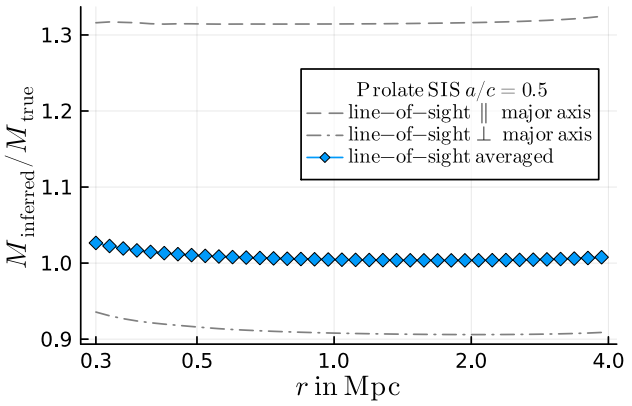

FIG. 4.— The mass profile of a prolate SIS, inferred from its reduced shear using our non-parametric deprojection method (Sec. 3.1), relative to the true mass profile (see Eq. (12)) for different orientations of the line of sight. When averaged over all line of sight orientations, the inferred mass matches the true mass at large radii where $\Sigma/\Sigma_{\rm crit}$ is negligible. At small radii, the non-linearity due to $\Sigma/\Sigma_{\rm crit}$ (see Eq. (3)) induces a small deviation from unity. Our choice of $\Sigma_{\rm crit}$ maximizes this non-linear effect. The tiny but perceptible uptick at the largest radii shown is an edge effect (see the main text).

In addition, there is a way to efficiently correct for residual miscentering effects. Below, we do not consider any miscentering effects, so for brevity we refer the reader to Mistele & Durakovic (2024) for details.

Finally, as we will now argue, our method works well, on average, even for triaxial mass distributions such as dark matter halos in ΛCDM (e.g., Bonamigo et al. 2015; Jing & Suto 2002). The important caveat here is "on average". Indeed, the mass inferred by our deprojection method for an individual triaxial halo can be off by a few 10%. This is illustrated by the gray lines in Fig. 4 for a prolate singular isothermal sphere (SIS).

Concretely, Fig. 4 assumes a 3D density $\rho(x,y,z) = \rho_{\rm SIS}(\sqrt{(x'^2+y'^2)/a^2+z'^2/c^2})$ where the coordinates x,y,z and x',y',z' are related by a rotation that determines the orientation of the line-of-sight and where $\rho_{\rm SIS}(r) \propto 1/r^2$. We use the formulas from Tessore & Metcalf (2015) to calculate the reduced shear assuming, for simplicity, a single source plane with a constant $\Sigma_{\rm crit} = 1750\,M_\odot/{\rm pc}^2$. We adopt $c=\sqrt{2}$, $a=1/\sqrt{2}$, and choose the prefactor of $\rho_{\rm SIS}$ such that, with $c=a=1$, the mass within 1 Mpc is $10^{15}\,M_\odot$. The inferred mass $M_{\rm inferred}$ is calculated by applying the deprojection formulas Eq. (7) and Eq. (8) to the reduced shear. The true mass $M_{\rm true}$ is calculated from Eq. (12) (see below). Our choice of $\Sigma_{\rm crit}$ is as small as possible, given that we are assuming *weak* lensing (we enforce $G_+ f_c < 1$ in the radial range of Fig. 4, see Sec. 3.1). This maximizes the non-linear effect at small radii discussed below.

The inferred mass $M_{\rm inferred}$ depends on how the halo is oriented relative to the line of sight (gray lines in Fig. 4). However, after averaging over all line of sight orientations, $M_{\rm infered}$ becomes very close to the true mass $M_{\rm true}$ (see the blue symbols in Fig. 4), where $M_{\rm true}$ is defined as the mass enclosed in a sphere with radius r,

$$M_{\rm true}(r) \equiv \int_{|\vec{x}'|<r} d^3\vec{x}'\,\rho(\vec{x}')\,. \tag{12}$$

In fact, at sufficiently large radii, the line of sight average of $M_{\rm inferred}$ is *exactly* $M_{\rm true}$. Sufficiently large radii here means radii where $\Sigma/\Sigma_{\rm crit} \ll 1$, which is usually a good approximation beyond a few hundred kpc. We prove this result in Appendix D and show that it holds for *any* mass distribution, not just triaxial ones (in the radial range where $\Sigma/\Sigma_{\rm crit} \ll 1$). The proof only requires one mild and reasonable additional assumption on the redshift distribution of the source galaxy population. At small radii, line-of-sight averaging does not exactly recover the true mass profile due to the non-linearity induced by $\Sigma/\Sigma_{\rm crit}$ (see Eq. (3)), but that effect seems to be quite mild in practice (Fig. 4).

Thus, on average, we expect our inferred mass profiles to be very close to the true mass profiles even for triaxial halos. This is important for statistical analyses of large samples of galaxy clusters, for example for cluster cosmology. This is particularly true for analyses based on quantities like M_{200c}, since these correspond to relatively large radii where $\Sigma/\Sigma_{\rm crit}$ is small. This result may in principle change if the line-of-sight average is incomplete, for example due to intrinsic alignments or due to selection effects. We expect such effects to be relatively minor, but leave a detailed study for future work.

Figure 4 shows a tiny but perceptible deviation from unity at the largest radii. This is an edge effect due to our choice of extrapolating G_+ beyond the last data point assuming a $1/R$ decay (see Sec. 3.1) which is not exactly true here. Such effects are taken into account in our systematic error estimate (see Sec. 3.3).

4. RESULTS

4.1. *Circular velocities and total mass profiles*

Figure 5 and Fig. 6 show the non-parametric mass profiles inferred using our non-parametric deprojection method in terms of the implied circular velocities $V_c(r) = \sqrt{GM(r)/r}$ and in terms of $M(r)$, respectively. We show results with the ΛCDM two-halo contribution subtracted (blue symbols, see Sec. 3.2) and without this subtraction (gray symbols). The light blue band indicates systematic uncertainties due to having to interpolate and extrapolate the observed shear profiles (Sec. 3.3). They become important only close to the last data point. Neighboring data points in Fig. 5 and Fig. 6 are correlated (see Sec. 3.3, Appendix B, and Fig. 16).

The circular velocities are remarkably flat, with no clear indication of a decline at large radii. This is reminiscent of galaxy rotation curves and perhaps indicates a universal pattern. The approximate flatness was previously noted in Donahue et al. (2014). In contrast to our analysis, Donahue et al. (2014) assumed a parametric NFW profile and did not subtract the two-halo contribution. Asymptotically, the circular velocities implied by NFW halos decay like $\sqrt{\ln(r)/r}$ and so we should not expect to see asymptotically flat circular velocities in ΛCDM. However, due to the relatively small concentrations of the dark matter halos of massive galaxy clusters and the limited radial range probed, our results are not very sensitive to this expected asymptotic decline.

In terms of the total mass M_{200c}, our non-parametric method gives results that are compatible with the NFW fits from Umetsu et al. (2014). There is a small and expected shift towards smaller masses (Fig. 7, left, see

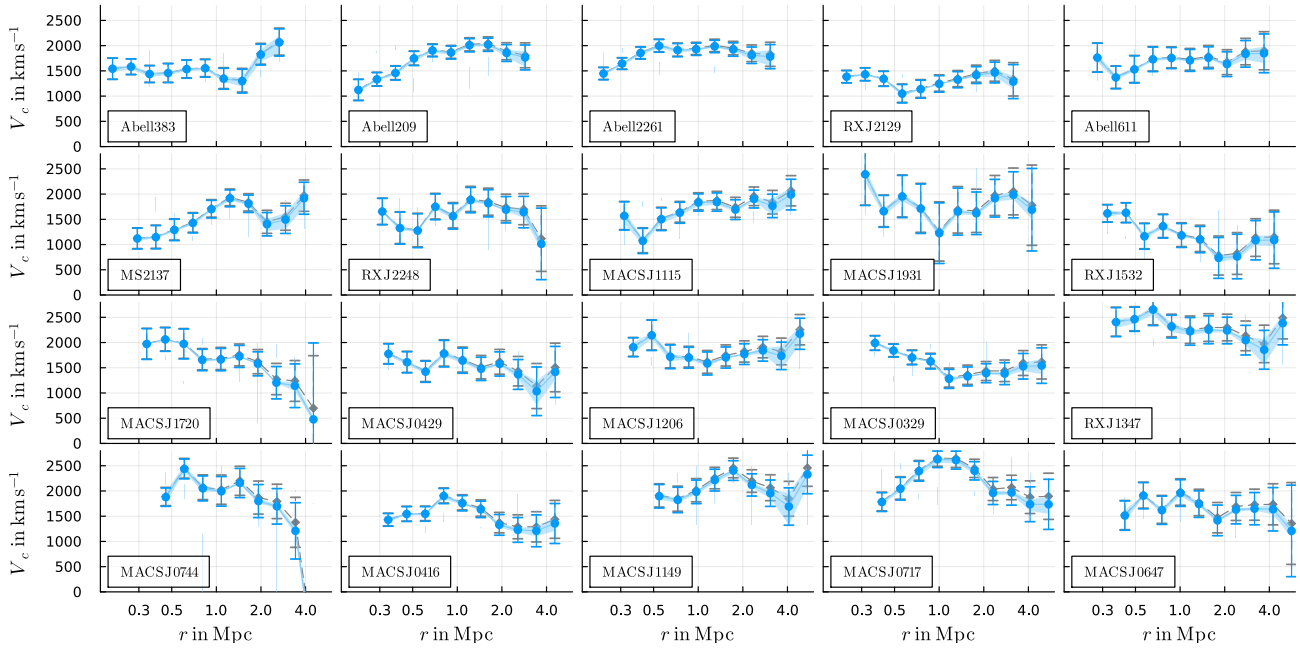

FIG. 5.— The circular velocities $V_c = \sqrt{GM(r)/r}$ inferred from the shear profile shown in Fig. 1 using our non-parametric deprojection method (Sec. 3.1). Results after subtracting the two-halo term are in blue, unsubtracted results are in gray. The two-halo subtraction is based on M_{200c} (see Sec. 3.2) for all clusters except Abell 383, where we use M_{500c} instead, because no value for M_{200c} can be determined due to the uptick at large radii. The light blue band indicates the systematic uncertainties due to choices in how to interpolate and extrapolate the shear profiles (Sec. 3.3). Error bars indicate the statistical uncertainties.

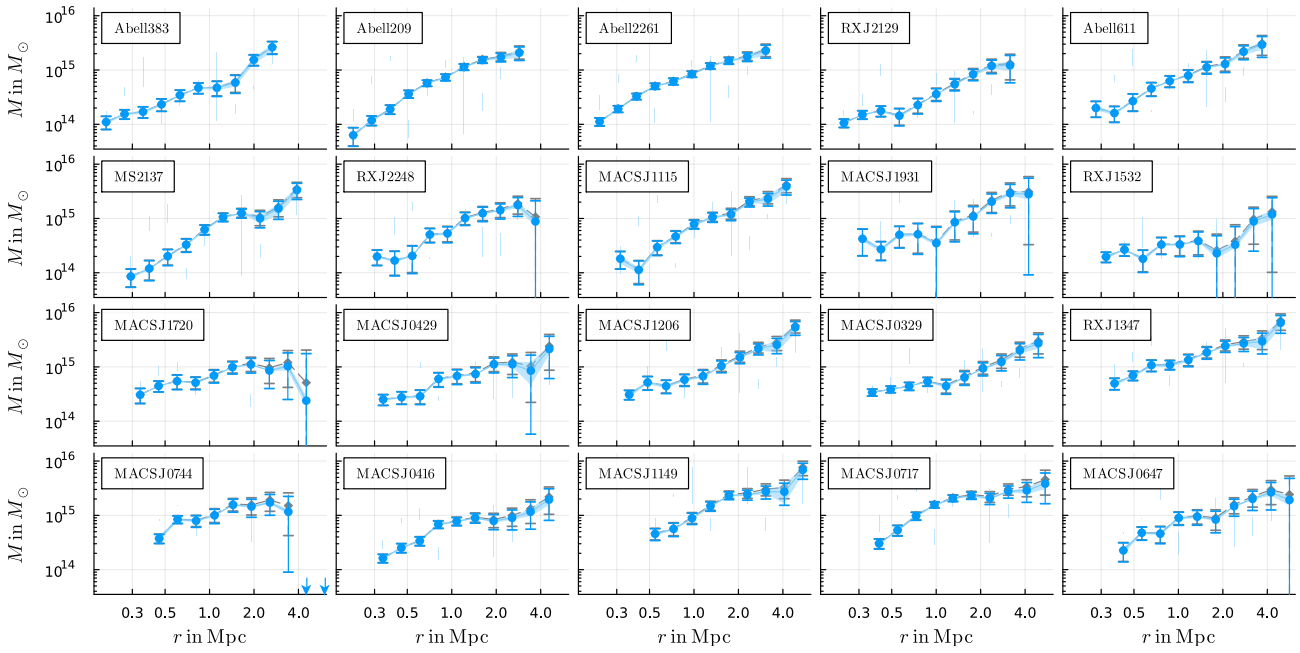

FIG. 6.— Same as Fig. 5 but in terms of masses $M(r)$. Arrows at the horizontal axis indicate negative inferred masses.

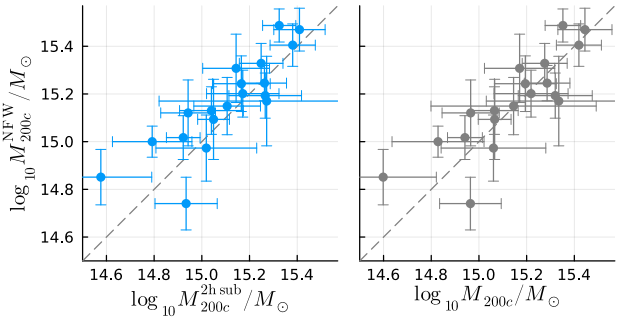

FIG. 7.— Our non-parametric M_{200c} measurements compared to the M_{200c} implied by the NFW fits from Umetsu et al. (2014) with (left) and without (right) the two-halo term subtracted (Sec. 3.2). The NFW fits from Umetsu et al. (2014) do not take into account the two-halo term. Abell 383 is excluded because no value of M_{200c} could be determined (see Fig. 6). Error bars indicate statistical uncertainties.

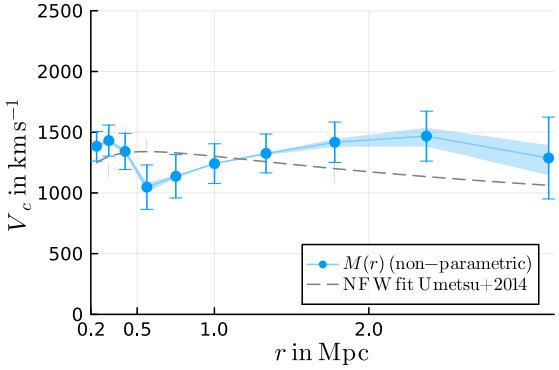

FIG. 8.— The circular velocity of RX J2129 as in Fig. 5, with the two-halo term subtracted, (blue) and the circular velocity implied by the NFW fit from Umetsu et al. (2014) (gray). A similar comparison for the other clusters can be found in Fig. 17 in Appendix E.

also Table 1) because Umetsu et al. (2014) did not take into account the two-halo term. Indeed, if we turn off our two-halo subtraction procedure, there is no longer a significant shift (Fig. 7, right). Figure 7 also shows that our uncertainties on M_{200c} are only moderately larger than those from Umetsu et al. (2014), despite making significantly fewer assumptions. On average, our statistical uncertainties are larger by 36% in the case without two-halo subtraction. This average and Fig. 7 do not include Abell 383 because we could not determine a value of M_{200c} due to the uptick at large radii (see Fig. 6).

Nevertheless, an NFW profile is not a great fit for all clusters. For example, the NFW fit from Umetsu et al. (2014) for RX J2129 seems to indicate a clearly declining circular velocity beyond ~ 0.5 Mpc (Donahue et al. 2014, see also Fig. 8). In contrast, our non-parametric mass profile implies a monotonic rise in the range $(0.5-2)$ Mpc. Inspection of Fig. 8 and the corresponding shear profile in Fig. 1 suggests that the NFW fit is thrown off by the data points at small radii, where the shear profile (and our non-parametric mass profile) show a clear change of behavior. The reason for this qualitative change in behavior may simply be statistical fluctuations or it may be real complexity in the lens' mass distribution that is not captured by the simple NFW model. RX J2129 is classified as relaxed (Donahue et al. 2016), but this does not necessarily preclude any significant non-symmetric structures in its center. For example, the temperature profile is not well described by an isothermal profile (Jiménez-Teja et al. 2024) and, using additional strong- and weak-lensing data from Hubble, Merten et al. (2015) noted some interesting morphology in the core of RX J2129. The assumption of spherical symmetry may also be violated in other clusters which are known ongoing mergers with complex X-ray emission, for example MACS J0717 (Ma et al. 2009) and MACS J0416 (Mann & Ebeling 2012).

In any case, this highlights an important advantage of our non-parametric method: Its mass estimates at large radii are not thrown off by complexity in the inner regions of a galaxy cluster. This is discussed in detail in Sec. 3.5 and in Mistele & Durakovic (2024). To be clear, if the mass distribution in the inner regions of, e.g., RX J2129 is highly non-symmetric, then our non-parametric method cannot be trusted either at these small radii. However, unlike an NFW fit, our inferred mass at larger radii will still be correct in such cases.

In contrast to NFW halos, asymptotically flat circular velocities are a prediction of alternative proposals such as MOND and, to varying degrees (Mistele et al. 2023a,b; Durakovic & Skordis 2024), its relativistic completions (e.g. Skordis & Złosnik 2021; Blanchet & Skordis 2024; Berezhiani & Khoury 2015). Qualitatively, our circular velocities are consistent with that prediction. However, the precise predictions of these models depend strongly on the baryonic mass distribution. Previous studies find that MOND-like theories underpredict the observed circular velocities of galaxy clusters, given their baryonic mass (for some recent works see, e.g., Li et al. 2023, 2024; Famaey et al. 2025; Tian et al. 2020; Kelleher & Lelli 2024; Ettori et al. 2019; Eckert et al. 2022). That said, the fact that the observed circular velocities of clusters seem to be approximately flat enables a potential solution to this discrepancy in terms of a missing baryonic mass component located at relatively small radii. We discuss this more in Sec. 4.7.

We note that our two-halo subtraction procedure is specific to ΛCDM and may not apply in other theories. Unfortunately, due to their inherent non-linearity (but see Milgrom 2025), reliably estimating the two-halo contribution in MOND-inspired theories is challenging. To the best of our knowledge, no such estimates are currently available. One specific non-linear effect that would have to be taken into account is the so-called external field effect (Bekenstein & Milgrom 1984; Haghi et al. 2016; Chae et al. 2020, 2021) which may play a role at large radii in galaxy clusters (Kelleher & Lelli 2024). Since the effect of the ΛCDM two-halo term on our results is quite modest, we might expect the same to be true for MOND-like theories. Properly addressing this question will, however, require simulations of structure formation in relativistic models such as AeST.

4.2. *Density profiles*

Figure 9 shows the 3D density profiles inferred using the non-parametric method from Sec. 3.4. This method is mathematically equivalent to taking the derivative of the non-parametric mass profiles $M(r)$ from Sec. 3.1 and dividing by $4\pi r^2$. The integral form Eq. (9) may, however, be preferable in practice because it avoids numerical derivatives. Nevertheless, reconstructing good den-

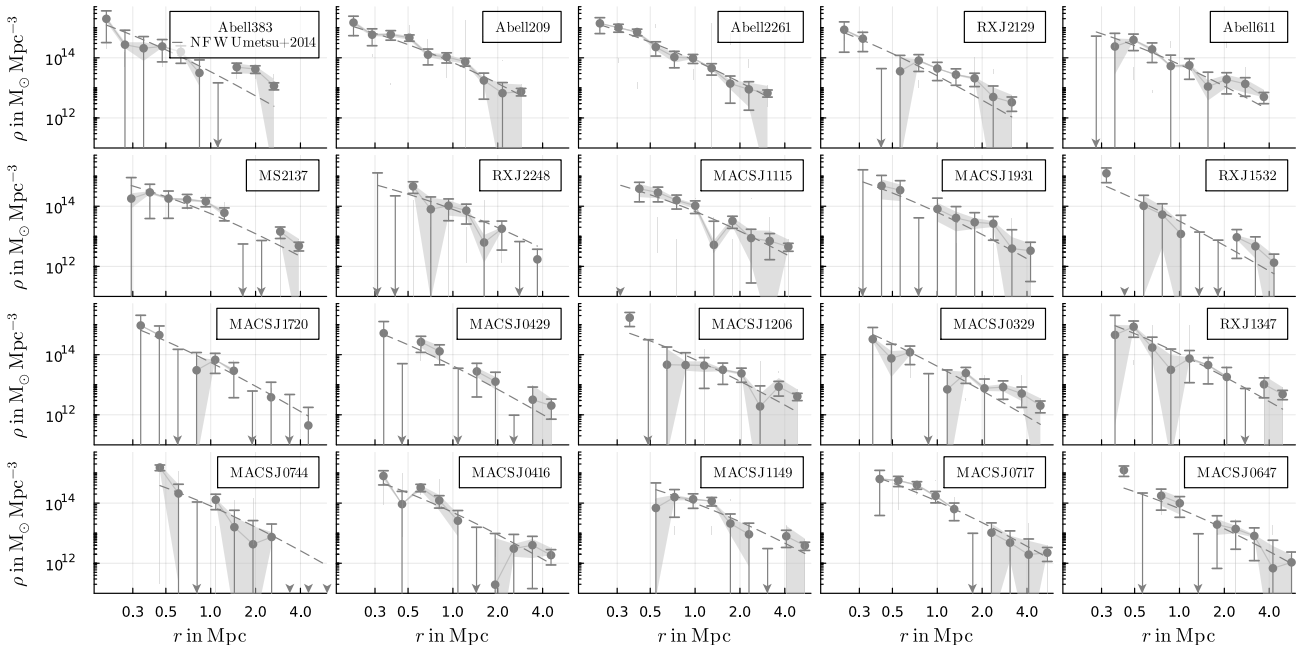

FIG. 9.— The 3D density profiles inferred using the non-parametric deprojection method from Sec. 3.4. For simplicity, no two-halo term is subtracted. Negative inferred densities are indicated by arrowheads at the horizontal axis. As in Fig. 5, light bands indicate systematic uncertainties from extrapolating and interpolating the shear profiles. Error bars indicate statistical uncertainties. Dashed gray lines show the densities implied by the NFW fits from Umetsu et al. (2014).

sity profiles requires much smaller statistical uncertainties than reconstructing good mass profiles. Indeed, for some clusters, such as Abell 209 or Abell 2261, the reconstructed densities look reasonable, but Fig. 9 also shows large fluctuations and even negative inferred densities.

That said, while the negative densities in Fig. 9 are likely just fluctuations, our deprojection method *can* properly handle negative densities. This is important for theoretical models where negative densities are real physical effects. Examples are the ghost condensate in AeST (Mistele et al. 2023a) and the "phantom dark matter" in some modified gravity models (Milgrom 1986).

Figure 9 also shows the NFW density profiles from Umetsu et al. (2014). These were obtained by fitting the weak-lensing convergence profile plus the average convergence in the cluster center and match our non-parametric density profiles reasonably well. As a further cross-check, we have fit NFW profiles to both our non-parametric mass and density profiles, finding results consistent with expectations, see Appendix F.

4.3. *Cosmic baryon fraction*

The baryon fraction $\Omega_b/\Omega_m \approx 0.16$ (Aghanim et al. 2020) plays an important role in cosmology. One may expect that this cosmic baryon fraction is consistent with the ratio $f_b \equiv M_b/(M_b + M_{\rm DM})$ in a gravitationally collapsed structure with dark matter mass $M_{\rm DM}$ (e.g., Planelles et al. 2013; Angelinelli et al. 2023; Rasia et al. 2025). This expectation is not realized in galaxies, where a much lower baryon fraction is detected. In contrast, measurements of galaxy clusters at large radii seem to detect most of the expected baryons (e.g., McGaugh et al. 2010; Wicker et al. 2023; Mantz et al. 2022).

Figures 10 and 11 show the baryon fraction implied by our analysis as a function of radius. The two-halo term is subtracted (Sec. 3.2). Our estimate of f_b is most reliable where X-ray and weak-lensing observations overlap (blue lines and symbols). In this radial range, most clusters in our sample have an f_b well below the cosmic 0.16.

At larger radii (gray lines and symbols), our results depend strongly on how we extrapolate the gas densities beyond the radius $R^X_{\rm max}$ where the beta-profile fits from Famaey et al. (2025) were well constrained by X-ray observations. If we assume a $1/r^4$ density tail (see Sec. 2.2), most clusters remain well below the cosmic baryon fraction even at large radii. This remains true even if we ignore the two clusters where $R^X_{\rm max}$ is likely too small to make a reliable estimate of the gas mass at large radii (diamond symbols). This is consistent with the expectations for a MOND isothermal sphere, which is one motivation behind considering a $1/r^4$ tail.

On the other hand, if we extrapolate the gas densities by taking the beta profile fits at face value even at large radii, where they were not well constrained by observations, f_b tends to increase with radius, with several clusters getting close to the cosmic baryon fraction around r_{200c}. That matches expectations from ΛCDM, but we caution that the beta profiles from Famaey et al. (2025) may underestimate the steepness of the gas density profiles at such large radii (see Sec. 2.2).

Beyond r_{200c} (dashed gray lines), the f_b of some clusters grow significantly beyond the cosmic baryon fraction when taking the beta profile fits at face value. However, the relative uncertainties become quite large at these extreme radii, so this effect is likely not significant.

4.4. *Stellar mass–halo mass relation*

Figure 12 shows the stellar–mass halo mass (SMHM) relation implied by our non-parametric mass profiles. We exclude Abell 383 because, as discussed above, we cannot

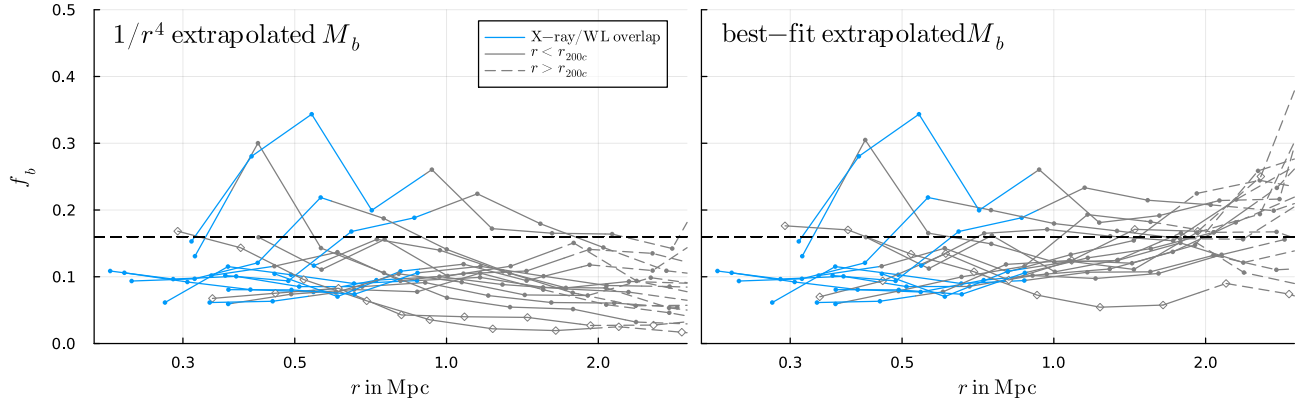

FIG. 10.— The baryon fraction $M_b(r)/M(r)$ implied by our non-parametric mass profiles as a function of radius. Blue lines indicate the radial range where weak-lensing and X-ray observations overlap, solid gray lines indicate radii beyond that but below r_{200c}, and dashed gray lines indicate radii beyond r_{200c}. We subtract the ΛCDM two-halo term (Sec. 3.2). We extrapolate the gas density profiles beyond $R^X_{\rm max}$ assuming a $1/r^4$ tail (left) and assuming the beta profile fits remain valid even at large radii where they were not well constrained by observations (right). Gray diamonds indicate the two clusters with the smallest $R^X_{\rm max}$ whose baryonic masses are likely to be underestimated with the $1/r^4$ extrapolation.

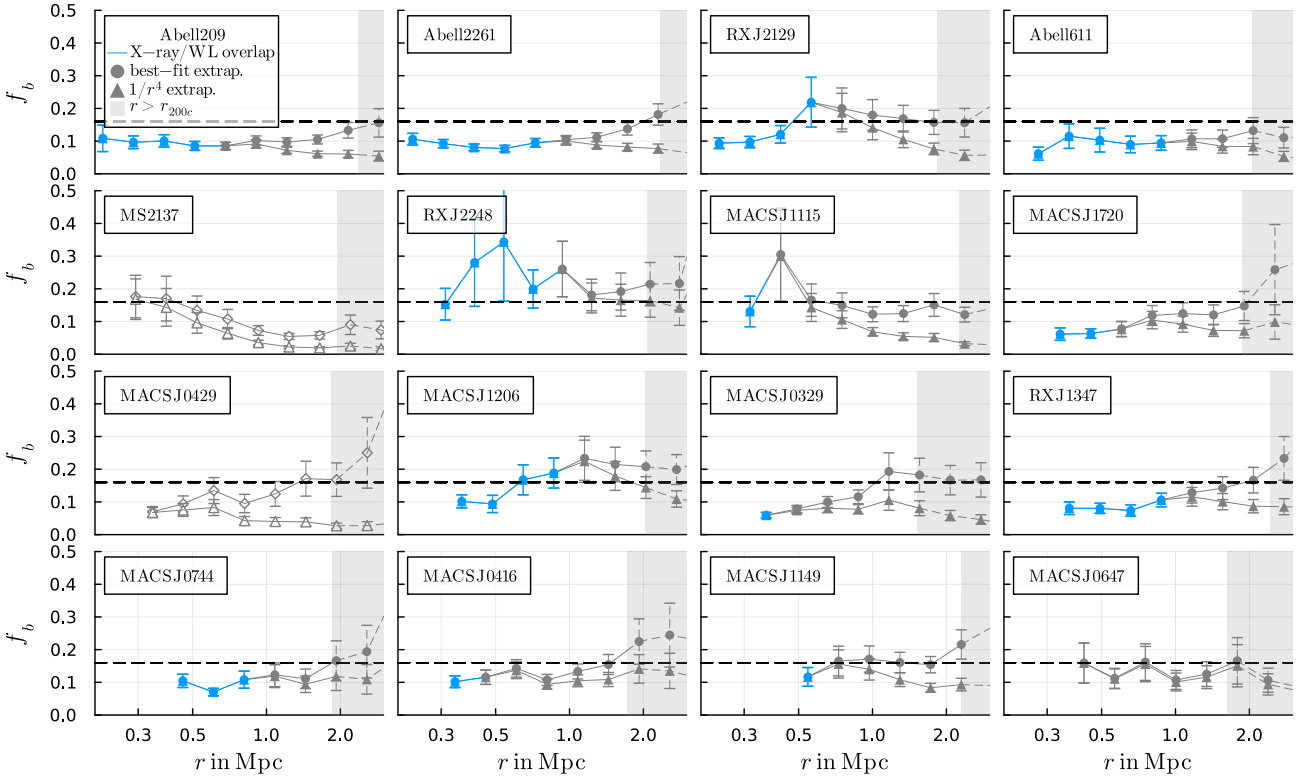

FIG. 11.— Same as Fig. 10 but separately for each cluster. Error bars indicate the statistical uncertainties. For each cluster, we show results assuming a $1/r^4$ gas density tail beyond $R^X_{\rm max}$ (smaller values of f_b, as in the left panel of Fig. 10) and assuming that the best-fit beta profiles remain valid even beyond $R^X_{\rm max}$ (larger values of f_b, as in the right panel of Fig. 10).

determine a value of M_{200c}. We include RX J1532, whose BCG's stellar mass $M_{*,\rm BCG}$ is not provided by Famaey et al. (2025), adopting its $M_{*,\rm BCG}$ directly from Burke et al. (2015).

We find very little correlation between $M_{*,\rm BCG}$ and M_{200c} (Burke et al. 2015). This may seem unexpected because, in ΛCDM, a strong correlation between these two quantities should exist. However, our cluster sample covers a relatively narrow range in total mass, so that what we see in Fig. 12 may simply be the scatter in stellar mass at an essentially fixed M_{200c}. This scatter is expected to be around 0.2 dex. Since the CLASH clusters cover a range of redshifts, the redshift evolution of the SMHM relation may add to the expected scatter. With these considerations, the Moster et al. (2013) relation seems roughly consistent with our results (dashed gray lines in Fig. 12). As a further check, we have generated simple mock data from the Moster et al. (2013)

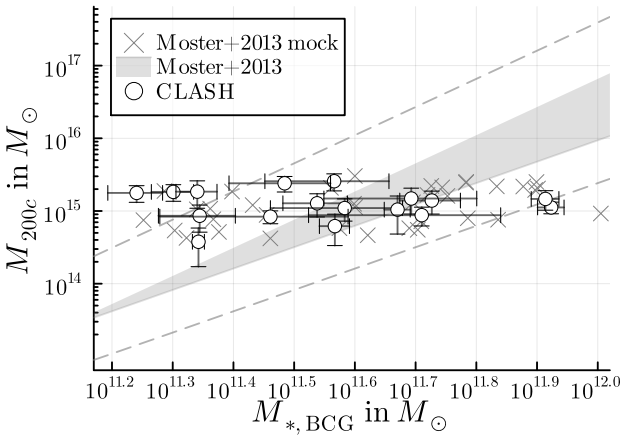

FIG. 12.— The SMHM relation implied by our non-parametric mass profiles (white symbols). The shaded gray region indicates the Moster et al. (2013) relation in the redshift range of our cluster sample. Dashed gray lines indicate 0.2 dex scatter in the direction of $M_{*,M_{\mathrm{BCG}}}$ around that region. Gray crosses show simple mock data generated from the Moster et al. (2013) relation.

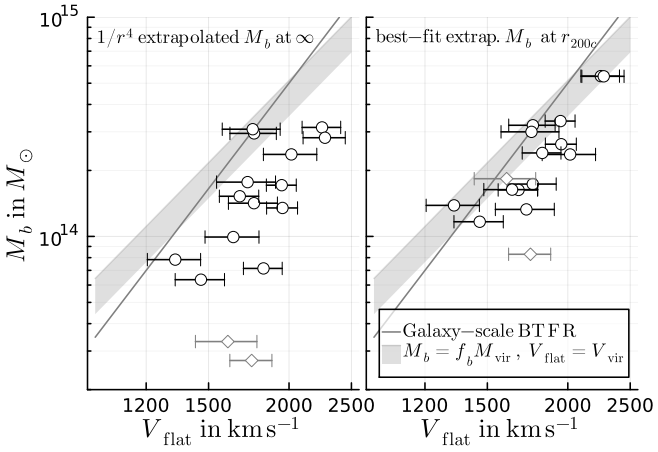

FIG. 13.— The BTFR implied by our non-parametric weak-lensing mass profiles and the baryonic mass estimates from Famaey et al. (2025). Baryonic masses assume a $1/r^4$ gas density tail (left) or extrapolate the best-fit beta profiles out to r_{200c} (right), see Sec. 2.2. Gray diamonds indicate the two clusters with the smallest radial range of X-ray observations, whose M_b are likely significantly underestimated with the $1/r^4$ extrapolation. The solid gray line is the galaxy-scale BTFR from Lelli et al. (2019). The shaded band indicates a simple ΛCDM estimate with $V_{\mathrm{flat}} = \sqrt{GM_{\mathrm{vir}}/r_{\mathrm{vir}}}$, $M_{\mathrm{vir}} = f_b M_b$, and $f_b \approx 0.16$. The size of the band corresponds to the range of redshifts of the clusters shown.

relation at $z = 0.4$, close to the mean z_l of our CLASH sample. In particular, we generated 250 equally-spaced $\log_{10} M_{200c}/M_\odot$ values between 10 and 16.5, calculated $M_{*,\mathrm{BCG}}$ according to the Moster et al. (2013) relation at $z = 0.4$, then added, respectively, 0.1 dex and 0.2 dex noise to M_{200c} and $M_{*,\mathrm{BCG}}$, and finally applied a cut $14.6 < \log_{10} M_{200c}/M_\odot < 15.5$. These mock data also seem consistent with our results (Fig. 12).

This is in contrast to massive spiral galaxies where the Moster et al. (2013) relation is in conflict with observations (e.g., Di Cintio & Lelli 2016). We also considered the SMHM relation from Kravtsov et al. (2018), finding that it is significantly offset from our measurements towards higher $M_{*,\mathrm{BCG}}$. However, the BCG stellar masses in Kravtsov et al. (2018) are defined to include stellar mass within many hundred kpc, including contributions from the intracluster light, and are in fact extrapolated to infinity. This contrasts with Burke et al. (2015) who measured $M_{*,\mathrm{BCG}}$ within 50 kpc, which may explain the offset compared to Kravtsov et al. (2018).

4.5. *BTFR*

The BTFR relates the asymptotic flat circular velocity V_{flat} to the total baryonic mass M_b. In Sec. 4.1, we saw that the circular velocities of the CLASH clusters are approximately flat. Thus, for simplicity we define V_{flat} as the weighted average of the circular velocities V_c for

$$750\,\mathrm{kpc} < r < 3\,\mathrm{Mpc}\,, \tag{13}$$

with weights given by the inverse squares of the statistical uncertainties. The circular velocities from Sec. 4.1 are not *perfectly* flat, so our V_{flat} will change somewhat if we choose a different radial range. We have verified that this effect is relatively minor and that other reasonable choices do not change our conclusions.

Regarding the second ingredient of the BTFR, the total baryonic mass M_b, we face the issue that the gas masses $M_{\mathrm{gas}}(r)$ implied by the double beta fits from Famaey et al. (2025) are divergent. This is not an issue for the $1/r^4$ extrapolation which integrates to a finite total mass (see Sec. 2.2), but we must deal with it when we extrapolate by taking the beta profile fits at face value. Thus, in the latter case, we simply evaluate M_b at the radius r_{200c} implied by our non-parametric mass profiles (instead of at $r = \infty$). This choice is somewhat ad-hoc but may have physical significance within ΛCDM, where the baryon fraction may be expected to be close to the cosmological one at the virial radius (see Sec. 4.3).

The resulting BTFRs are shown in Fig. 13. We do not subtract the two-halo term. We have verified that it is not important. How we estimate M_b is important, however. Extrapolating with a $1/r^4$ tail leads to a significant offset from the Lelli et al. (2019) galaxy-scale relation (solid line in Fig. 13), while evaluating the double beta fits at r_{200c} leads to a comparably small offset.

In either case, our results are consistent with existing literature finding that galaxy clusters may follow a parallel relation compared to galaxies (Sanders 2003; McGaugh 2015). With the $1/r^4$ extrapolation, the two clusters with the lowest M_b clearly do not follow a parallel relation, but this is expected since their M_{gas} is likely significantly underestimated (Sec. 2.2).

Figure 13 also shows a simple ΛCDM estimate based on identifying (V_{flat}, M_b) with $(\sqrt{GM_{\mathrm{vir}}/r_{\mathrm{vir}}}, f_b M_{\mathrm{vir}})$ where $f_b \approx 0.16$ is the cosmic baryon fraction. This corresponds to a slope of 3 (McGaugh 2012) which is notably different from the galaxy-scale BTFR which has a slope of about 4. A slope of about 4 seems to be a somewhat better representation of the trend in our results.

4.6. *RAR*

The RAR relates the Newtonian acceleration g_{bar} due to the baryonic mass at a given radius, $G_N M_b(r)/r^2$, to the total acceleration g_{obs}, $G_N M(r)/r^2$, at that radius. Unlike the BTFR, this relation can be evaluated locally at each radius (assuming spherical symmetry). Thus, in the radial range where X-ray and weak-lensing observa-

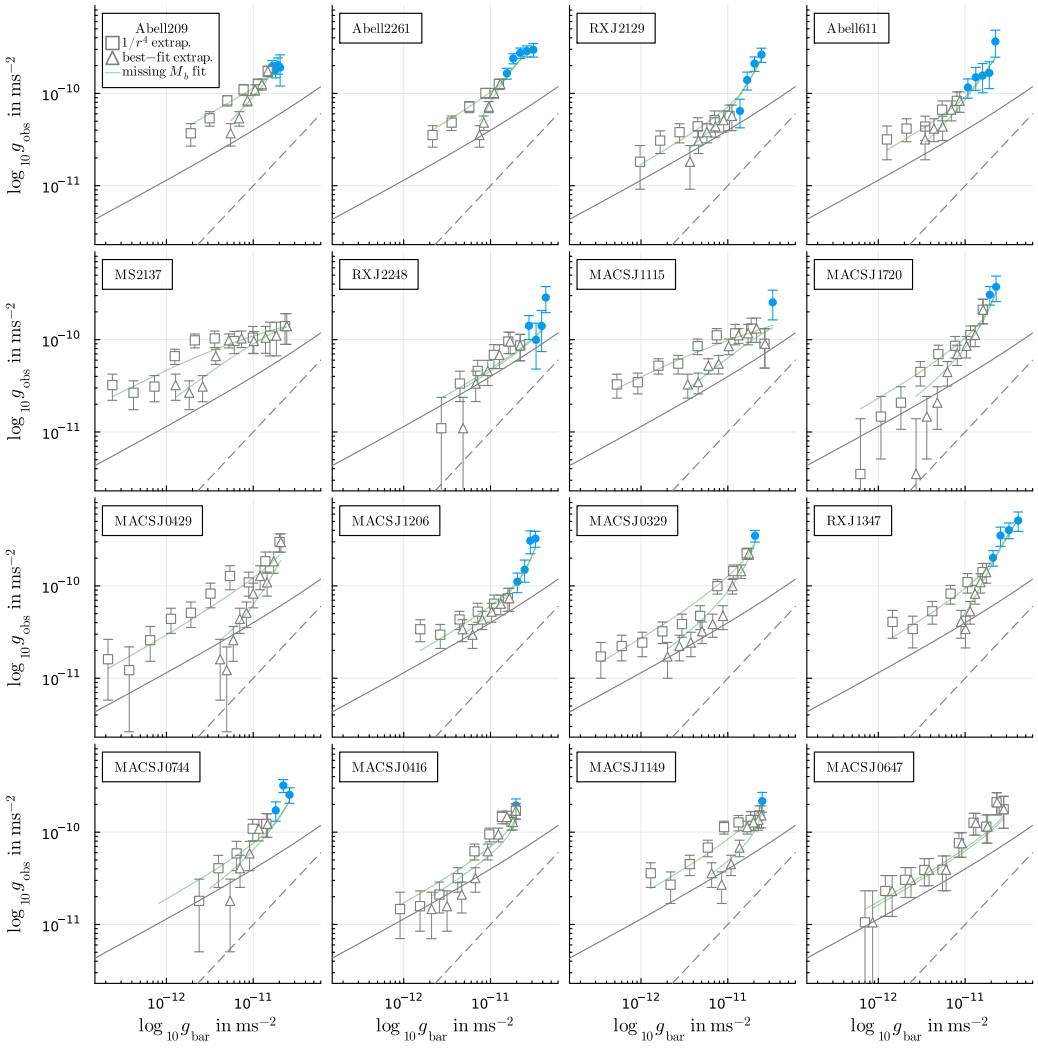

FIG. 14.— The RAR implied by our non-parametric weak-lensing mass profiles and the baryonic mass estimates from Famaey et al. (2025). Blue circles indicate the radial range where X-ray and weak-lensing observations overlap. Beyond that range, we extrapolate the gas densities assuming a $1/r^4$ tail (gray squares) or assuming the best-fit beta profiles remain valid even at large radii where they were not well constrained by observations (gray triangles). We do not subtract the two-halo term (Sec. 3.2). Two data points with negative $g_{\rm obs}$ for MACS J0744 (see Fig. 6) are omitted. The dashed gray line indicates equality of $g_{\rm obs}$ and $g_{\rm bar}$, the solid gray line indicates the galaxy-scale RAR from Lelli et al. (2017). Assuming the galaxy-scale RAR holds for clusters, one can fit a missing baryonic component (not included in $g_{\rm bar}$) to $g_{\rm obs}$ (solid green lines, see Sec. 4.7).

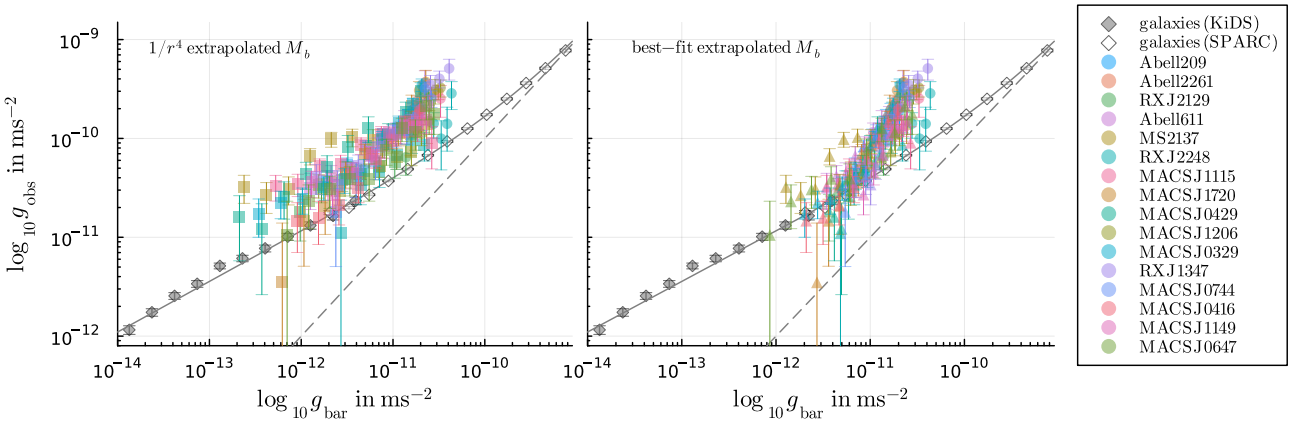

FIG. 15.— Same as Fig. 14, but all clusters combined. Left: Extrapolating the gas densities with a $1/r^4$ tail. Right: Extrapolating the gas densities assuming the best-fit beta profiles. Symbols are as in Fig. 14. Gray and white diamonds show the RAR of galaxies from SPARC (Lelli et al. 2017) and KiDS (Mistele et al. 2024b; Lelli et al. 2024), respectively.

tions overlap, we can measure the RAR without having to worry about how to extrapolate the gas profiles.

The RAR implied by our baryonic and total mass estimates of each galaxy cluster are shown in Fig. 14. Blue symbols indicate the range where X-ray and weak-lensing observations overlap. Outside this range, we separately show extrapolations of M_b assuming a $1/r^4$ gas density tail and assuming that the best-fit double beta profiles continue to be valid (Sec. 2.2). We do not subtract the two-halo term. Doing so does not significantly change the results. Figure 15 shows the data of all galaxy clusters combined.

The radial range where X-ray and weak-lensing observations overlap is quite narrow for many clusters and non-existent for 3 clusters. Thus, we cannot draw strong conclusions about the galaxy cluster RAR at large radii or, equivalently, small accelerations. We can, however, see that the galaxy-scale RAR seems to underpredict the $g_{\rm obs}$ of clusters already at relatively small radii, i.e. at relatively large $g_{\rm bar}$. For a few clusters, this may be because the weak-lensing measurements of $g_{\rm obs}$ are not trustworthy at small radii (e.g. RX J2129, see Sec. 4.1). Another hypothesis (e.g., Milgrom 2008; Kelleher & Lelli 2024; Li et al. 2023; Famaey et al. 2025) is that there is an additional baryonic mass component not captured by our estimate of $g_{\rm bar}$. That the galaxy-scale RAR underpredicts $g_{\rm obs}$ already at small radii means that, if it exists, this missing mass is not (only) to be found at large radii. We discuss this more in Sec. 4.7.

4.7. *Missing mass to recover galaxy-scale RAR*

In Sec. 4.6, we saw that the galaxy-scale RAR underpredicts the $g_{\rm obs}$ observed in clusters. Assuming that the RAR is a universal relation, this may be a sign that there is a missing baryonic mass component (Eckert et al. 2022; Kelleher & Lelli 2024; Famaey et al. 2025) such as undetected, compact clouds of cold gas (Milgrom 2008). To test this hypothesis, Kelleher & Lelli (2024) have fit the observed $g_{\rm obs}$ in clusters by adding a missing M_b component and assuming that the galaxy-scale RAR holds. They assumed a missing mass density proportional to $(M_{\rm mm,tot}/r_s^3)(1+r/r_s)^{-4}$ and found that such a profile could reasonably explain the observed $g_{\rm obs}$ within about 1 Mpc, where their data is likely not affected by hydrostatic bias. The same missing mass profile was also considered by Famaey et al. (2025) among a variety of other profiles. We here repeat the fitting procedure of Kelleher & Lelli (2024) using our measurements.

The fit has three free parameters, $M_{\rm mm,tot}$, r_s, as well as a parameter Υ_b that scales the observed baryonic mass, akin to a mass-to-light ratio. We assume flat priors for $12 < \log_{10} M_{\rm mm,tot}/M_\odot < 16$ and $1 < \log_{10} r_s/{\rm kpc} < 4$ as well as a log-normal prior around $\Upsilon_b = 1$ with 0.1 dex scatter. We do not subtract the two-halo term. Subtracting the two-halo term does not significantly change the results. For the RAR acceleration scale, we adopt $a_0 = 1.2 \cdot 10^{-10}\,{\rm m/s^2}$ (Lelli et al. 2017). The resulting best fits are shown as solid green lines in Fig. 14 and Table 2.

Overall, we find that a missing baryonic component can reasonably fit our observations. We find best-fit parameters roughly consistent with those from Kelleher & Lelli (2024) and Famaey et al. (2025). Finding a reasonable missing M_b profile fails only when the observed galaxy-cluster RAR in Fig. 14 dips below the solid gray line, i.e. below the galaxy-scale RAR. This is because, assuming the galaxy-scale RAR holds, such dips require negative missing mass densities. This mostly only happens at the very largest radii, i.e. smallest $g_{\rm bar}$, and when extrapolating the gas densities by taking the beta profile fits from Famaey et al. (2025) at face value. No significant negative mass densities are required when extrapolating the gas density with a $1/r^4$ tail, corresponding to a MOND isothermal sphere (Milgrom 1984). We also note that neighboring data points are positively correlated (see Fig. 16) so that the visual impression of neighboring data points fluctuating down together may be misleading. In any case, previous analyses have also seen this phenomenon. At such large radii, however, analyses based on gas thermodynamics may suffer from hydrostatic bias (Kelleher & Lelli 2024), $g_{\rm bar}$ may not be well constrained (Famaey et al. 2025), and, depending on the theoretical framework, an external field effect may be important (Kelleher & Lelli 2024).

As an alternative approach, assuming that the galaxy-scale RAR holds universally, we can directly calculate and study the (non-parametric) missing mass profiles $M_b^{\rm miss}(r)$ implied by our measurements of $g_{\rm obs}$ and $g_{\rm bar}$. We do so in Appendix G, finding similar results as above.

5. DISCUSSION

5.1. *Concentration and sparsity*

Our non-parametric results suggest a shift of perspective: Away from first performing a parametric fit and then calculating quantities of interest in terms of the fit results, towards directly inferring quantities of interest from the non-parametric profiles $M(r)$ or $\rho(r)$, similar to how we measured the total mass M_{200c} in Sec. 4.1 or the baryon fraction $M_b(r)/M(r)$ in Sec. 4.3. Another quantity of interest is the concentration c, which captures information about the shape of the mass profile. Traditionally, definitions of c are tied to specific parametric profiles such as the NFW profile. A concentration can, however, also be defined in a non-parametric way.

One example is the ratio $c_{0.1} \equiv r_{200c}/r_{0.1}$ where $r_{0.1}$ is the radius that encloses 10% of the total mass (Yasin et al. 2023). Indeed, in principle, we can directly measure $c_{0.1}$ using our non-parametric mass profiles. Unfortunately, just like parametrically defined concentrations (Appendix F), $c_{0.1}$ is challenging to measure in practice without extending the radial range covered by weak-lensing observations (for example using strong lensing, Merten et al. 2015; Umetsu et al. 2016, 2025). In our current sample, only 4 clusters reach sufficiently small radii to determine $c_{0.1}$.

Another example is the sparsity, defined as $M_{200c}/M_{\Delta c}$ with, for example, $\Delta = 500$ or $\Delta = 1000$ (Balmès et al. 2014). Initial results indicate that our non-parametric mass profiles tend to imply smaller sparsities than the NFW fits from Umetsu et al. (2014), consistent with the results from Balmès et al. (2014). A proper statistical analysis as well as a comparison between different parametric and non-parametric concentration measures are left for future work.

5.2. *Density profiles*

TABLE 2
FIT PARAMETERS OF THE MISSING M_b PROFILES FOLLOWING KELLEHER & LELLI (2024).

Name	$1/r^4$ extrapolated M_b				best-fit extrapolated M_b			
	$\log_{10} M_{\rm mm,tot}$ $M_\odot$	$\log_{10}\rho_s$ $M_\odot\,{\rm Mpc}^{-3}$	$\log_{10} r_s$ kpc	$\log_{10}\Upsilon_b$	$\log_{10} M_{\rm mm,tot}$ $M_\odot$	$\log_{10}\rho_s$ $M_\odot\,{\rm Mpc}^{-3}$	$\log_{10} r_s$ kpc	$\log_{10}\Upsilon_b$
Abell 209	$14.99^{+0.15}_{-0.16}$	$15.48^{+0.25}_{-0.25}$	$2.63^{+0.12}_{-0.13}$	$+0.01^{+0.10}_{-0.10}$	$14.91^{+0.16}_{-0.18}$	$15.56^{+0.27}_{-0.26}$	$2.58^{+0.13}_{-0.14}$	$-0.01^{+0.10}_{-0.10}$
Abell 2261	$14.83^{+0.13}_{-0.14}$	$16.02^{+0.24}_{-0.22}$	$2.40^{+0.11}_{-0.12}$	$+0.00^{+0.10}_{-0.10}$	$14.79^{+0.13}_{-0.15}$	$16.09^{+0.25}_{-0.22}$	$2.36^{+0.11}_{-0.12}$	$-0.03^{+0.09}_{-0.09}$
RXJ2129	$13.92^{+0.17}_{-0.16}$	$17.75^{+0.94}_{-1.08}$	$1.50^{+0.41}_{-0.34}$	$+0.00^{+0.10}_{-0.10}$	$13.88^{+0.15}_{-0.15}$	$17.89^{+0.87}_{-1.03}$	$1.45^{+0.38}_{-0.31}$	$-0.02^{+0.09}_{-0.09}$
Abell 611	$14.49^{+0.36}_{-0.33}$	$16.23^{+1.58}_{-0.91}$	$2.21^{+0.41}_{-0.64}$	$+0.02^{+0.10}_{-0.10}$	$14.39^{+0.37}_{-0.29}$	$16.48^{+1.67}_{-1.03}$	$2.09^{+0.45}_{-0.65}$	$+0.02^{+0.10}_{-0.10}$
MS2137	$14.77^{+0.29}_{-0.30}$	$15.13^{+0.57}_{-0.55}$	$2.68^{+0.26}_{-0.28}$	$+0.00^{+0.10}_{-0.10}$	$14.57^{+0.36}_{-0.48}$	$15.22^{+0.92}_{-0.74}$	$2.58^{+0.33}_{-0.45}$	$+0.02^{+0.11}_{-0.11}$
RXJ2248	$13.92^{+0.57}_{-1.08}$	$15.58^{+2.33}_{-3.14}$	$2.23^{+1.04}_{-0.85}$	$+0.05^{+0.10}_{-0.10}$	$13.84^{+0.57}_{-1.08}$	$15.58^{+2.26}_{-3.33}$	$2.19^{+1.13}_{-0.81}$	$+0.03^{+0.09}_{-0.10}$
MACSJ1115	$15.03^{+0.27}_{-0.28}$	$14.89^{+0.50}_{-0.51}$	$2.84^{+0.24}_{-0.24}$	$+0.01^{+0.10}_{-0.10}$	$14.52^{+0.49}_{-0.84}$	$15.07^{+1.35}_{-1.14}$	$2.63^{+0.45}_{-0.69}$	$+0.05^{+0.11}_{-0.11}$
MACSJ1720	$14.40^{+0.14}_{-0.15}$	$18.19^{+1.04}_{-1.25}$	$1.52^{+0.45}_{-0.36}$	$+0.00^{+0.10}_{-0.10}$	$14.38^{+0.14}_{-0.16}$	$18.23^{+0.99}_{-1.21}$	$1.49^{+0.43}_{-0.34}$	$-0.03^{+0.10}_{-0.10}$
MACSJ0429	$14.20^{+0.22}_{-0.22}$	$17.39^{+1.37}_{-1.46}$	$1.71^{+0.55}_{-0.48}$	$+0.01^{+0.10}_{-0.10}$	$14.07^{+0.21}_{-0.32}$	$17.56^{+1.23}_{-1.67}$	$1.60^{+0.60}_{-0.42}$	$-0.04^{+0.10}_{-0.09}$
MACSJ1206	$14.34^{+0.22}_{-0.19}$	$17.51^{+1.31}_{-1.22}$	$1.72^{+0.47}_{-0.47}$	$+0.01^{+0.10}_{-0.10}$	$14.29^{+0.19}_{-0.19}$	$17.79^{+1.20}_{-1.28}$	$1.61^{+0.49}_{-0.42}$	$-0.01^{+0.09}_{-0.09}$
MACSJ0329	$14.37^{+0.08}_{-0.09}$	$18.69^{+0.71}_{-0.98}$	$1.35^{+0.34}_{-0.25}$	$-0.01^{+0.10}_{-0.10}$	$14.34^{+0.08}_{-0.09}$	$18.76^{+0.65}_{-0.88}$	$1.31^{+0.31}_{-0.22}$	$-0.04^{+0.09}_{-0.09}$
RXJ1347	$14.83^{+0.19}_{-0.13}$	$17.71^{+1.45}_{-1.01}$	$1.83^{+0.39}_{-0.53}$	$+0.01^{+0.10}_{-0.10}$	$14.77^{+0.16}_{-0.12}$	$18.06^{+1.32}_{-1.09}$	$1.69^{+0.42}_{-0.47}$	$-0.02^{+0.09}_{-0.10}$
MACSJ0744	$14.55^{+0.28}_{-0.24}$	$16.78^{+1.66}_{-0.94}$	$2.05^{+0.40}_{-0.63}$	$+0.01^{+0.10}_{-0.10}$	$14.51^{+0.28}_{-0.24}$	$16.86^{+1.60}_{-1.01}$	$2.00^{+0.42}_{-0.61}$	$-0.02^{+0.10}_{-0.10}$
MACSJ0416	$14.19^{+0.26}_{-0.22}$	$16.72^{+1.45}_{-0.88}$	$1.94^{+0.37}_{-0.55}$	$+0.02^{+0.10}_{-0.10}$	$14.12^{+0.25}_{-0.20}$	$17.07^{+1.37}_{-1.05}$	$1.80^{+0.43}_{-0.50}$	$-0.01^{+0.09}_{-0.09}$
MACSJ1149	$14.82^{+0.31}_{-0.29}$	$16.19^{+1.88}_{-0.92}$	$2.34^{+0.40}_{-0.73}$	$+0.02^{+0.10}_{-0.10}$	$14.56^{+0.28}_{-0.25}$	$17.10^{+1.75}_{-1.45}$	$1.92^{+0.58}_{-0.62}$	$-0.00^{+0.09}_{-0.09}$
MACSJ0647	$14.32^{+0.47}_{-0.66}$	$15.90^{+1.97}_{-1.87}$	$2.25^{+0.71}_{-0.78}$	$+0.03^{+0.11}_{-0.10}$	$14.26^{+0.49}_{-0.75}$	$15.83^{+2.02}_{-2.18}$	$2.26^{+0.77}_{-0.79}$	$+0.04^{+0.10}_{-0.10}$

NOTE. — The listed values are the 16th, 50th, and 84th percentiles. Values of $\rho_s = 3M_{\rm mm,tot}/(4\pi r_s^3)$ are provided for easier comparison to Table 1 of Famaey et al. (2025).

In Sec. 4.2, we saw that reconstructing non-parametric density profiles is viable but noisy with current data. One way to ameliorate this may be to stack many clusters. This is especially relevant given the large cluster samples with weak-lensing data that are expected to become available from instruments such as Euclid (Euclid Collaboration et al. 2025), Roman (Spergel et al. 2015), or Rubin (LSST Science Collaboration et al. 2009). This may enable model-independent constraints on quantities such as the splashback radius (e.g. Diemer & Kravtsov 2014; More et al. 2015, 2016), allowing to distinguish between different models of dark matter and modified gravity (Adhikari et al. 2018).

For example, in the context of MOND, one might expect a much weaker splashback signal than in ΛCDM. Indeed, the splashback feature is a steep drop in density corresponding to accreting collisionless matter reaching its apocenter after first infall. In ΛCDM, a splashback feature exists in both the dark matter and galaxy density profiles since both are collisionless. Since dark matter is the dominant contribution to the gravitational potential, a splashback feature is then visible also in the weak-lensing signal. In contrast, in MOND-like theories, the dominant contribution to the gravitational potential is due to the collisional intracluster medium, with only a sub-dominant contribution from the collisionless galaxies. As a result, one may expect to see a weaker splashback feature in the weak-lensing signal.

When testing individual theoretical models, it may be easiest to follow ΛCDM-based analyses (e.g., More et al. 2016) and fit a parametric density profile that is chosen according to the specific model under consideration. Nonetheless, non-parametric constraints are desirable since they are easier to interpret and to apply to many models at once. The feasibility of such non-parametric constraints will be studied in future work.

5.3. *Baryonic mass measurements*

Compared to our total mass measurements from weak lensing, the baryonic masses are relatively poorly constrained at large radii. This limits our results regarding baryon fractions and regarding scaling relations such as the BTFR and RAR. Better knowledge of the hot gas density profiles would enable stronger constraints on cosmology, dark matter, and modified gravity. Indeed, as discussed in Sec. 2.2, results from the X-COP project suggest that, while our simple $1/r^4$ extrapolation is perhaps too steep, the beta profile fits from Famaey et al. (2025) may not be steep enough at large radii. Currently, it is unknown where our cluster sample falls between these two extremes.

It would therefore be very useful to combine our non-parametric weak-lensing measurements with improved gas density profiles from follow-up X-ray observations. Another possibility would be to stack the gas densities of a large sample of galaxy clusters, for example from eROSITA (Bulbul et al. 2024; Kluge et al. 2024), to increase the signal-to-noise ratio at large radii. These could then be combined with stacked weak-lensing observations of the same sample from, for example, Euclid.

5.4. *Cluster cosmology*

Cluster cosmology aims to constrain cosmological parameters such as Ω_m and σ_8 using the high-mass end of the halo mass function. This requires unbiased mass measurements. We expect that the non-parametric methods discussed here will be useful in minimizing the impact of a number of important biases that affect such measurements. In particular, many cluster cosmology analyses are based on measurements of M_{200c}. Non-parametric measurements of M_{200c} using our method will (see Sec. 3.5 and Mistele & Durakovic 2024): 1) Not be thrown off by baryonic effects that may change the shape of the mass profile, 2) not be thrown off by complex, non-

symmetric mass distributions at the centers of clusters (e.g., late-stage mergers), 3) not be thrown off, on average, by triaxiality, 4) be less affected by miscentering (and there is an efficient way to correct for residual miscentering effects), and 5) apply directly to models beyond ΛCDM such as models based on ultra-light dark matter that may change the shape of the halo profile. In addition, our mass reconstruction runs fast, taking only a few milliseconds per cluster to run.

The caveat is that, by making fewer assumptions, the statistical uncertainties become larger. However, as shown in Sec. 4.1, that is only a moderate effect in practice. Also, with large cluster samples from upcoming and future surveys statistical uncertainties will no longer be the main limiting factor. In fact, a trade-off in favor of significantly reduced systematic uncertainties, i.e. biases, at the cost of somewhat larger statistical error bars may be required in order to fully exploit the potential of instruments such as Euclid, Roman, or Rubin.

6. CONCLUSION

We have applied a new weak-lensing deprojection method to the CLASH sample of galaxy clusters. We have inferred non-parametric mass and density profiles, which we have studied by themselves as well as in relation to the baryonic mass components. We find: (1) The implied circular velocities are approximately flat. (2) The radially resolved baryonic mass fractions vary significantly from cluster to cluster and depend strongly on how we extrapolate the X-ray gas profiles at large radii, so it is unclear whether the CLASH clusters reach the cosmic baryon fraction expected in ΛCDM. (3) The non-parametric masses are consistent with the ΛCDM SMHM relation. (4) The CLASH clusters deviate from the BTFR and the RAR defined by galaxies, but the offset depends strongly on how we extrapolate the gas masses. Contrary to some previous results based on hydrostatic equilibrium, we find that galaxy clusters may fall on the same BTFR and RAR as galaxies if one adds a suitable positive baryonic mass component.

Several of these results are limited by the baryonic masses being relatively poorly constrained at large radii. Improving on this will unlock stronger constraints on cosmology, dark matter, and modified gravity.

ACKNOWLEDGEMENTS

We thank Amel Durakovic, Paolo Tozzi, Pengfei Li, and Konstantin Haubner for helpful discussions. This work was supported by the DFG (German Research Foundation) – 514562826.

Our non-parametric mass and density profiles, including correlation matrices, as well as the flat circular velocities implied by the mass profiles are available on Zenodo, https://dx.doi.org/10.5281/zenodo.15476959.

REFERENCES

Adhikari S., Sakstein J., Jain B., Dalal N., Li B., 2018, J. Cosmology Astropart. Phys., 2018, 033
Aghanim N., et al., 2020, Astronomy & Astrophysics, 641, A6
Angelinelli M., Ettori S., Dolag K., Vazza F., Ragagnin A., 2023, A&A, 675, A188
Balmès I., Rasera Y., Corasaniti P. S., Alimi J. M., 2014, MNRAS, 437, 2328
Bartelmann M., 1995, A&A, 303, 643
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Bekenstein J., Milgrom M., 1984, Astrophys. J., 286, 7
Berezhiani L., Khoury J., 2015, Phys. Rev., D92, 103510
Blanchet L., Skordis C., 2024, J. Cosmology Astropart. Phys., 2024, 040
Bonamigo M., Despali G., Limousin M., Angulo R., Giocoli C., Soucail G. v., 2015, MNRAS, 449, 3171
Brouwer M. M., et al., 2021, Astronomy & Astrophysics, 650, A113
Bulbul E., et al., 2024, A&A, 685, A106
Burke C., Hilton M., Collins C., 2015, MNRAS, 449, 2353
Chae K.-H., Lelli F., Desmond H., McGaugh S. S., Li P., Schombert J. M., 2020, Astrophys. J., 904, 51
Chae K.-H., Desmond H., Lelli F., McGaugh S. S., Schombert J. M., 2021, Astrophys. J., 921, 104
Covone G., Sereno M., Kilbinger M., Cardone V. F., 2014, ApJ, 784, L25
Di Cintio A., Lelli F., 2016, MNRAS, 456, L127
Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1
Donahue M., et al., 2014, ApJ, 794, 136
Donahue M., et al., 2016, ApJ, 819, 36
Durakovic A., Skordis C., 2024, J. Cosmology Astropart. Phys., 2024, 040
Dutton A. A., Macciò A. V., 2014, Mon. Not. Roy. Astron. Soc., 441, 3359
Eckert D., Ettori S., Pointecouteau E., Molendi S., Paltani S., Tchernin C., 2017, Astronomische Nachrichten, 338, 293
Eckert D., Ettori S., Pointecouteau E., van der Burg R. F. J., Loubser S. I., 2022, A&A, 662, A123
Ettori S., et al., 2019, A&A, 621, A39
Euclid Collaboration et al., 2025, A&A, 697, A1
Famaey B., Durakovic A., 2025, arXiv e-prints, p. arXiv:2501.17006
Famaey B., Pizzuti L., Saltas I. D., 2025, Phys. Rev. D, 111, 123042
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Ge H., Xu K., Ghahramani Z., 2018, in International Conference on Artificial Intelligence and Statistics, AISTATS 2018, 9-11 April 2018, Playa Blanca, Lanzarote, Canary Islands, Spain. pp 1682–1690, http://proceedings.mlr.press/v84/ge18b.html;https://dblp.org/rec/bib/conf/aistats/GeXG18
Ghirardini V., et al., 2019, A&A, 621, A41
Guzik J., Seljak U., 2001, MNRAS, 321, 439
Haghi H., Bazkiaei A. E., Zonoozi A. H., Kroupa P., 2016, MNRAS, 458, 4172
Hoekstra H., 2003, MNRAS, 339, 1155
Jiménez-Teja Y., et al., 2024, ApJ, 974, 309
Jing Y. P., Suto Y., 2002, ApJ, 574, 538
Johnston D. E., Sheldon E. S., Tasitsiomi A., Frieman J. A., Wechsler R. H., McKay T. A., 2007, ApJ, 656, 27
Kaiser N., 1995, ApJ, 439, L1
Kelleher R., Lelli F., 2024, A&A, 688, A78
Kluge M., et al., 2024, A&A, 688, A210
Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astronomy Letters, 44, 8
LSST Science Collaboration et al., 2009, arXiv e-prints, p. arXiv:0912.0201
Laudato E., Salzano V., Umetsu K., 2022, MNRAS, 511, 1878
Lelli F., McGaugh S. S., Schombert J. M., Pawlowski M. S., 2017, Astrophys. J., 836, 152
Lelli F., McGaugh S. S., Schombert J. M., Desmond H., Katz H., 2019, MNRAS, 484, 3267
Lelli F., Mistele T., McGaugh S. S., Schombert J. M., Li P., 2024, arXiv e-prints, p. arXiv:2411.02499
Lewis A., Bridle S., 2002, Phys. Rev. D, 66, 103511
Li P., et al., 2023, A&A, 677, A24
Li P., et al., 2024, A&A, 692, A253
Ma C.-J., Ebeling H., Barrett E., 2009, ApJ, 693, L56
Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, MNRAS, 391, 1940
Mann A. W., Ebeling H., 2012, MNRAS, 420, 2120
Mantz A. B., et al., 2022, MNRAS, 510, 131

McGaugh S. S., 2012, AJ, 143, 40
McGaugh S. S., 2015, Canadian Journal of Physics, 93, 250
McGaugh S. S., Schombert J. M., Bothun G. D., de Blok W. J. G., 2000, Astrophys. J. Lett., 533, L99
McGaugh S. S., Schombert J. M., de Blok W. J. G., Zagursky M. J., 2010, ApJ, 708, L14
Merten J., et al., 2015, ApJ, 806, 4
Milgrom M., 1983a, Astrophys. J., 270, 365
Milgrom M., 1983b, Astrophys. J., 270, 371
Milgrom M., 1983c, Astrophys. J., 270, 384
Milgrom M., 1984, ApJ, 287, 571
Milgrom M., 1986, ApJ, 306, 9
Milgrom M., 2008, New A Rev., 51, 906
Milgrom M., 2025, Phys. Rev. D, 111, 104033
Mistele T., Durakovic A., 2024, The Open Journal of Astrophysics, 7, 120
Mistele T., McGaugh S., Hossenfelder S., 2023a, A&A, 676, A100
Mistele T., McGaugh S., Hossenfelder S., 2023b, J. Cosmology Astropart. Phys., 2023, 004
Mistele T., McGaugh S., Lelli F., Schombert J., Li P., 2024a, ApJ, 969, L3
Mistele T., McGaugh S., Lelli F., Schombert J., Li P., 2024b, J. Cosmology Astropart. Phys., 2024, 020
Miyatake H., et al., 2019, ApJ, 875, 63
More S., Diemer B., Kravtsov A. V., 2015, ApJ, 810, 36
More S., et al., 2016, ApJ, 825, 39
Moster B. P., Naab T., White S. D. M., 2013, Mon. Not. Roy. Astron. Soc., 428, 3121
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Oguri M., Hamana T., 2011, MNRAS, 414, 1851
Oguri M., Takada M., 2011, Phys. Rev. D, 83, 023008
Planelles S., Borgani S., Dolag K., Ettori S., Fabjan D., Murante G., Tornatore L., 2013, MNRAS, 431, 1487
Postman M., et al., 2012, ApJS, 199, 25
Rasia E., et al., 2025, arXiv e-prints, p. arXiv:2505.21624
Revels J., Lubin M., Papamarkou T., 2016, arXiv:1607.07892 [cs.MS]
Sanders R. H., 1999, ApJ, 512, L23
Sanders R. H., 2003, MNRAS, 342, 901
Skordis C., Złosnik T., 2021, Phys. Rev. Lett., 127, 161302
Spergel D., et al., 2015, arXiv e-prints, p. arXiv:1503.03757
Tessore N., Metcalf R. B., 2015, A&A, 580, A79
Tian Y., Umetsu K., Ko C.-M., Donahue M., Chiu I. N., 2020, ApJ, 896, 70
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Umetsu K., 2020, A&A Rev., 28, 7
Umetsu K., Broadhurst T., Zitrin A., Medezinski E., Hsu L.-Y., 2011, ApJ, 729, 127
Umetsu K., et al., 2014, ApJ, 795, 163
Umetsu K., Zitrin A., Gruen D., Merten J., Donahue M., Postman M., 2016, ApJ, 821, 116
Umetsu K., Pizzardo M., Diaferio A., Geller M. J., 2025, ApJ, 990, 70
Urban O., et al., 2014, MNRAS, 437, 3939
Walker S. A., Fabian A. C., Sanders J. S., George M. R., 2012, MNRAS, 424, 1826
Walker S. A., Fabian A. C., Sanders J. S., Simionescu A., Tawara Y., 2013, MNRAS, 432, 554
Wicker R., Douspis M., Salvati L., Aghanim N., 2023, A&A, 674, A48
Yasin T., Desmond H., Devriendt J., Slyz A., 2023, MNRAS, 526, 5861

APPENDIX

A. TWO-HALO TERM IN ΛCDM

A.1. *Subtraction procedure*

Here, we describe the details of the two-halo subtraction procedure from Sec. 3.2. The two-halo term becomes important only at projected radii R that are sufficiently large for $\Sigma/\Sigma_{\rm crit}$ to be small. As is clear from Eq. (3), the distinction between $\Delta\Sigma$ and G_+ is then unimportant and we have to a good approximation $\Delta\Sigma = G_+$. For our two-halo subtraction procedure, it then suffices to consider only the second step Eq. (8) of our deprojection procedure, i.e. the step that converts $\Delta\Sigma$ to the mass M. The first step that converts G_+ to $\Delta\Sigma$ is unimportant.

The excess surface density we observe, $\Delta\Sigma_{\rm obs}$ contains contributions both from the galaxy cluster itself, $\Delta\Sigma_{\rm 1h}$, and from the two-halo term, $\Delta\Sigma_{\rm 2h}$,

$$\Delta\Sigma_{\rm obs}(R) = \Delta\Sigma_{\rm 1h}(R) + \Delta\Sigma_{\rm 2h}(R)\,. \tag{A1}$$

When we apply the deprojection formula Eq. (8) from Sec. 3.1 to the observed $\Delta\Sigma_{\rm obs}$, we infer a mass profile $M_{\rm obs}$ that likewise contains the desired contribution from the galaxy cluster itself, $M_{\rm 1h}$, and a two-halo contribution $M_{\rm 2h}$,

$$M_{\rm obs}(r) = M_{\rm 1h}(r) + M_{\rm 2h}(r)\,, \tag{A2}$$

where

$$M_{\rm 2h}(r) \equiv 4r^2 \int_0^{\pi/2} d\theta\, \Delta\Sigma_{\rm 2h}\left(\frac{r}{\sin\theta}\right)\,. \tag{A3}$$

We note that $M_{\rm 2h}$ is not the mass profile of any actual object. It just quantifies by how much our deprojection technique overestimates the true galaxy cluster mass profile when applied to lensing data that contains contributions from a cluster's local environment. In practice, there is a technical complication, not captured by Eq. (A3), due to the fact that we extrapolate G_+ beyond the last measured data point at $R_{\rm max}$. This requires a modification of Eq. (A3) that we discuss separately in Appendix A.2. Thus, in practice, we use Eq. (A12) instead of Eq. (A3).

A simple estimate of $\Delta\Sigma_{\rm 2h}$ within ΛCDM is (e.g. Guzik & Seljak 2001; Oguri & Hamana 2011; Covone et al. 2014)

$$\Delta\Sigma_{\rm 2h}(R) = b\frac{\bar\rho_{m,0}}{2\pi D(z_l)^2}\int_0^\infty d\ell\,\ell J_2\left(\frac{\ell R}{D(z_l)}\right) P_m(k_\ell; z_l)\,, \tag{A4}$$

where J_2 denotes the second Bessel function of the first kind, b is the bias, $\bar\rho_{m,0}$ is the mean matter density at redshift $z = 0$, and $P_m(k_\ell; z_l)$ is the linear matter power spectrum at $k_\ell = \ell/[(1+z)D(z_l)]$. We calculate P_m using CAMB and we adopt the bias from Tinker et al. (2010) which gives b as a function of the mass $M_{\Delta m}$ and redshift z_l,

$$b = b(z_l, M_{\Delta m})\,. \tag{A5}$$

Here, $M_{\Delta m}$ is the mass within the radius $r_{\Delta m}$ where the cluster's average mass density drops below Δ times the cosmological matter density $\bar{\rho}_m$ at the cluster redshift z_l. Below, we use M_{200c} instead of $M_{\Delta m}$. To avoid a mismatch, we use a value of Δ that makes the two halo mass definitions equivalent, namely $\Delta = 200\rho_{\rm crit}(z_l)/\bar{\rho}_m(z_l)$.

Our goal is to recover the galaxy cluster's mass profile $M_{\rm 1h}$ by subtracting the two-halo contribution $M_{\rm 2h}$ from $M_{\rm obs}$. To this end, we use the independent estimate of the two-halo contribution Eq. (A4). This estimate of $\Delta\Sigma_{\rm 2h}$ depends on the redshift and the total mass of the galaxy cluster. It can be converted to $M_{\rm 2h}$ using Eq. (A3) (or, rather, Eq. (A12)). We denote this estimate of $M_{\rm 2h}$ by

$$M_{\rm 2h}(r|M_{200c})\,, \tag{A6}$$

which makes the dependence on the total mass explicit. Given the observed $M_{\rm obs}$, we can determine M_{200c} by numerically solving the following equation for M_{200c},

$$M_{\rm obs}(r_{200c}) = M_{200c} + M_{\rm 2h}(r_{200c}|M_{200c})\,, \tag{A7}$$

where $M_{200c} = (4\pi/3)\,200\cdot\rho_{\rm crit}\,r_{200}^3$. This equation follows from Eq. (A2) by using our estimate $M_{\rm 2h}(r|M_{200c})$ for the two-halo contribution and by using that, by definition, $M_{\rm 1h}(r_{200c}) = M_{200c}$. Having determined M_{200c} in this way, we can extract the desired mass profile $M_{\rm 1h}(r)$ of the galaxy cluster by subtracting the two-halo contribution,

$$M_{\rm 1h}(r) = M_{\rm obs}(r) - M_{\rm 2h}(r|M_{200c})\,. \tag{A8}$$

The above procedure can easily be adapted to work with M_{500c} instead of M_{200c}. We will use M_{500c} instead of M_{200c} for one galaxy cluster, Abell 383, where a value for M_{200c} cannot be determined.

A.2. *Adjustments due to extrapolation*

The observed $\Delta\Sigma$ contains contributions from both the 1-halo and 2-halo terms, see Eq. (A1). Thus, applying our deprojection formula Eq. (8) gives

$$M_{\rm obs}(r) = 4r^2\int_0^{\pi/2} d\theta\,\left(\Delta\Sigma_{\rm 1h}\left(\frac{r}{\sin\theta}\right) + \Delta\Sigma_{\rm 2h}\left(\frac{r}{\sin\theta}\right)\right)\,, \quad \text{(not quite correct in practice)}\,. \tag{A9}$$

This is what Eq. (A3) is based on, but it is not quite what we actually infer in practice. The reason is that, after the last measured data point at $R_{\rm max}$, we no longer use $\Delta\Sigma = \Delta\Sigma_{\rm 1h} + \Delta\Sigma_{\rm 2h}$ in the integrand. Instead, we use a power law extrapolation, $G_+ \propto 1/R^n$ (see Sec. 3.1). At the large radii relevant for $R \geq R_{\rm max}$, this extrapolation also implies $\Delta\Sigma \propto 1/R^n$ to a good approximation. Thus, what we actually infer is

$$M_{\rm obs}(r) = 4r^2\left[\int_{\theta_{\rm min}}^{\pi/2} d\theta\,\left(\Delta\Sigma_{\rm 1h}\left(\frac{r}{\sin\theta}\right) + \Delta\Sigma_{\rm 2h}\left(\frac{r}{\sin\theta}\right)\right) + (\Delta\Sigma_{\rm 1h}(R_{\rm max}) + \Delta\Sigma_{\rm 2h}(R_{\rm max}))\frac{R_{\rm max}^n}{r^n}\int_0^{\theta_{\rm min}} d\theta\sin^n\theta\right]\,, \tag{A10}$$

where $\theta_{\rm min} = \arcsin(r/R_{\rm max})$. If our power law extrapolation correctly captures the behavior of $\Delta\Sigma_{\rm 1h}$, this is

$$M^{\rm obs}(r) = M_{\rm 1h}(r) + 4r^2\left[\int_{\theta_{\rm min}}^{\pi/2} d\theta\,\Delta\Sigma_{\rm 2h}\left(\frac{r}{\sin\theta}\right) + \Delta\Sigma_{\rm 2h}(R_{\rm max})\frac{R_{\rm max}^n}{r^n}\int_0^{\theta_{\rm min}} d\theta\sin^n\theta\right]\,. \tag{A11}$$

Indeed, for the extrapolation, one should choose a power law decay $1/R^n$ that plausibly matches the behavior of the shear due to the galaxy cluster itself, i.e. due to the one-halo term (*not* the total shear including the two-halo term). We can now read off the correct expression to use for $M_{\rm 2h}(r|M_{200c})$ in our two-halo subtraction procedure,

$$M_{\rm 2h}(r|M_{200c}) = 4r^2\left[\int_{\theta_{\rm min}}^{\pi/2} d\theta\,\Delta\Sigma_{\rm 2h}\left(\frac{r}{\sin\theta}\right) + \Delta\Sigma_{\rm 2h}(R_{\rm max})\frac{R_{\rm max}^n}{r^n}\int_0^{\theta_{\rm min}} d\theta\sin^n\theta\right]\,. \tag{A12}$$

This replaces Eq. (A3) and takes into account that we extrapolate beyond the last measured data point $R_{\rm max}$. See Appendix E for how to adopt this procedure when extrapolating assuming an NFW profile instead of a power law.

B. UNCERTAINTIES AND COVARIANCES

As systematic uncertainties, we consider our choices of how to extrapolate and interpolate the shear profile G_+. To estimate the effect of these choices, we first calculate mass profiles with opposite and extreme choices of how to extrapolate G_+ beyond the last data point. In particular, we consider extrapolating G_+ assuming $1/R^2$ and $1/\sqrt{R}$ power law decays. These correspond, respectively, to a decay as fast as for a point particle and to a decay significantly slower than for an SIS. These cases likely bracket the true behavior of the cluster's shear. At each radius r, we calculate

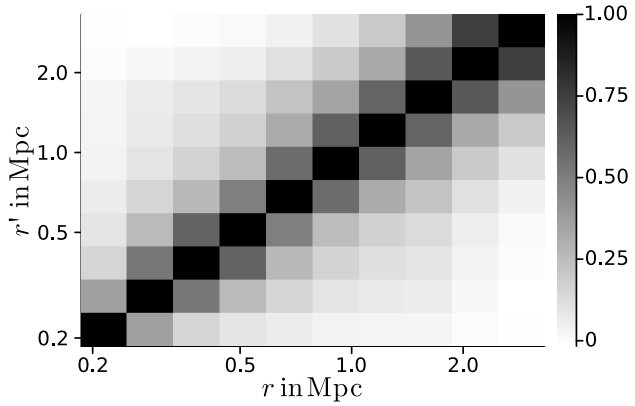

FIG. 16.— The correlation matrix of the two-halo-subtracted masses inferred for Abell 209. The correlation matrix is defined in terms of the covariance matrix as $\mathrm{Cov}(M(r), M(r'))/\sigma_{M(r)}\sigma_{M(r')}$.

the minimum and maximum mass achievable in this way. Schematically,

$$M(r)^{\max} = \max_{n\in\{\frac{1}{2},1,2\}} M(r)|^{\text{extrapolate } 1/R^n} , \tag{B1}$$

$$M(r)^{\min} = \min_{n\in\{\frac{1}{2},1,2\}} M(r)|^{\text{extrapolate } 1/R^n} . \tag{B2}$$

To take into account systematic uncertainties from interpolation, we add (subtract) the difference between the mass profiles obtained using linear and quadratic interpolation to $M^{\max}$ (from $M^{\min}$). Schematically,

$$M^{\max}_{\min} \to M^{\max}_{\min} \pm \left| M(r)|^{\text{quadratic}} - M(r)|^{\text{linear}} \right| . \tag{B3}$$

For the statistical uncertainties and covariances, we use linear error propagation to convert uncertainties and covariances on G_+ into uncertainties and covariances on the inferred mass M. Following Mistele & Durakovic (2024), we implement this error propagation by writing differentiable Julia code and then using 'ForwardDiff.jl' (Revels et al. 2016) to calculate the required Jacobians. This reduces linear error propagation to a simple matrix multiplication.

For the covariance matrix of $G_+ = \langle g_+\rangle/\langle\Sigma^{-1}_{\text{crit}}\rangle$, we take into account uncertainties in $\langle g_+\rangle$ as well as uncertainties and covariances due to $\langle\Sigma^{-1}_{\text{crit},ls}\rangle$ and due to the correlated LSS (Hoekstra 2003),

$$C^G = C^G_g + C^G_{\text{crit}} + C^G_{\text{LSS}} . \tag{B4}$$

We assume that the measurement uncertainties on $\langle g_+\rangle$ in different radial bins are uncorrelated,

$$\left(C^G_g\right)_{ij} = \delta_{ij}\frac{\sigma^2_{g_+,i}}{\langle\Sigma^{-1}_{\text{crit}}\rangle^2} , \tag{B5}$$

where i and j run over the radial bins and $\sigma_{g_+,i}$ denotes the uncertainty on $\langle g_+\rangle$ in the i-th radial bin. As discussed in Sec. 3.1, we assume $\langle\Sigma^{-1}_{\text{crit},ls}\rangle$ to be the same in all radial bins. This induces correlations between the radial bins,

$$\left(C^G_{\text{crit}}\right)_{ij} = \langle g_{+,i}\rangle\langle g_{+,j}\rangle\frac{\sigma^2_{\text{crit}}}{\langle\Sigma^{-1}_{\text{crit}}\rangle^4} , \tag{B6}$$

where σ_{crit} is the uncertainty on $\langle\Sigma^{-1}_{\text{crit},ls}\rangle$. These formulas follow from the definition Eq. (4) of G_+ in terms of g_+ and Σ_{crit}. The LSS contribution is important only at relatively large radii and following Umetsu (2020) we calculate it from the non-linear matter power spectrum produced by CAMB (Lewis & Bridle 2002). For simplicity, we assume all source galaxies to be located in a single source plane with effective redshift $z_{s,\text{eff}}$ when calculating C^G_{LSS} (see also Miyatake et al. 2019). We take $z_{s,\text{eff}}$ to be the source redshift whose critical surface density Σ_{crit} is $\langle\Sigma^{-1}_{\text{crit}}\rangle$.

In addition to G_+, our deprojection method from Sec. 3.1 also has a second input, namely $f_c = \langle\Sigma^{-2}_{\text{crit},ls}\rangle/\langle\Sigma^{-1}_{\text{crit},ls}\rangle$. For simplicity and because Umetsu et al. (2014) do not readily provide these, we do not take into account uncertainties in f_c. We expect that doing so would only have a minor effect since f_c only enters as the prefactor of $\Sigma/\Sigma_{\text{crit}}$ in the relation $G_+ = \Delta\Sigma/(1 - f_c\cdot\Sigma/\Sigma_{\text{crit}})$. Indeed, this term is unimportant at large radii and gives only moderate corrections at small radii. We expect the same to hold for corrections to the uncertainties and covariances induced by f_c. In addition, we expect the uncertainties on f_c to be less than 10%. This is smaller than the typical uncertainties on the shear G_+ at small radii, which would further reduce the importance of the f_c uncertainties.

As mentioned above, we obtain the covariance matrix of the inferred masses $M(r)$ by linearly propagating the covariance matrix of $G_+(R)$. As a representative example, Fig. 16 shows the correlation matrix of the inferred, two-halo subtracted masses $M(r)$ for Abell 209. We see that neighboring data points are correlated. In part, these correlations are already present in the inputs to our method and are just propagated into the result. But more importantly, the integrals Eq. (7) and Eq. (8) mix different radii, producing additional correlations with the characteristic pattern previously found in Mistele et al. (2024a).

C. DERIVATION OF DENSITY RECONSTRUCTION FORMULA

We start with Eq. (8) which gives $M(r)$ in terms of $\Delta\Sigma(R)$,

$$M(r) = 4r^2 \int_0^{\pi/2} d\theta\, \Delta\Sigma\left(\frac{r}{\sin\theta}\right) \equiv 4r^2 I(r)\,, \tag{C1}$$

where I denotes the θ integral. Using $\rho = M'(r)/4\pi r^2$, this gives

$$\rho(r) = \frac{I'(r)}{\pi} + \frac{2I(r)}{\pi r}\,. \tag{C2}$$

Our goal is to get rid of the derivative in the $I'(r)$ term. The idea is to pull the derivative into the integrand of the θ integral, so that it acts on $\Delta\Sigma$, and then use integration by parts. It will be useful to temporarily relax the upper integration boundary to $\pi/2 - \epsilon$ and take the limit $\epsilon \to 0$ at the end of the calculation,

$$I(r) = \lim_{\epsilon\to 0} I_\epsilon(r) \equiv \lim_{\epsilon\to 0} \int_0^{\frac{\pi}{2}-\epsilon} d\theta \Delta\Sigma\left(\frac{r}{\sin\theta}\right)\,. \tag{C3}$$

We have

$$I'_\epsilon(r) = \int_0^{\frac{\pi}{2}-\epsilon} d\theta\, \Delta\Sigma'\left(\frac{r}{\sin\theta}\right)\frac{1}{\sin\theta} = -\frac{1}{r}\int_0^{\frac{\pi}{2}-\epsilon} d\theta\, \frac{\sin\theta\tan\theta}{\sin\theta}\partial_\theta\left[\Delta\Sigma\left(\frac{r}{\sin\theta}\right)\right]\,. \tag{C4}$$

After integrating by parts, this becomes

$$I'_\epsilon(r) = -\frac{1}{r}\left[\tan\theta\,\Delta\Sigma\left(\frac{r}{\sin\theta}\right)\right]_{\theta=0}^{\theta=\frac{\pi}{2}-\epsilon} + \frac{1}{r}\int_0^{\frac{\pi}{2}-\epsilon} d\theta\,(\partial_\theta\tan\theta)\,\Delta\Sigma\left(\frac{r}{\sin\theta}\right) \tag{C5}$$

$$= -\frac{\Delta\Sigma(r)}{\epsilon\, r} + \frac{1}{r}\int_0^{\frac{\pi}{2}-\epsilon} d\theta\,\Delta\Sigma\left(\frac{r}{\sin\theta}\right)\frac{1}{\cos^2\theta} + O(\epsilon)\,. \tag{C6}$$

The $1/\epsilon$ divergence is why it was useful to relax the upper integration boundary. Importantly, a $1/\epsilon$ divergence exists not only in the boundary term but also in the remaining integral. In fact, the divergence from the remaining integral precisely cancels the $1/\epsilon$ from the boundary term. This must be the case because the expression we started with was finite in the limit $\epsilon \to 0$. The factor of $\Delta\Sigma\left(\frac{r}{\sin\theta}\right)$ in the remaining integral is well-behaved at $\theta \to \pi/2$, i.e. at the upper integration boundary in the limit $\epsilon \to 0$. The $1/\epsilon$ divergence is due only to the $1/\cos^2\theta$ factor. In fact,

$$\int_0^{\frac{\pi}{2}-\epsilon} \frac{d\theta}{\cos^2\theta} = \frac{1}{\epsilon} + O(\epsilon)\,. \tag{C7}$$

We can therefore rewrite the $1/\epsilon$ from the boundary term as this $1/\cos^2\theta$ integral and combine all integrals into one,

$$I'_\epsilon(r) = \frac{1}{r}\int_0^{\frac{\pi}{2}-\epsilon} d\theta\, \frac{\Delta\Sigma\left(\frac{r}{\sin\theta}\right) - \Delta\Sigma(r)}{\cos^2\theta} + O(\epsilon)\,. \tag{C8}$$

The integrand of this integral is finite everywhere in the interval $(0, \pi/2)$. This can be verified by expanding the integrand around $\pi/2$. The limit $\epsilon \to 0$ can now be taken,

$$I'(r) = \lim_{\epsilon\to 0} I'_\epsilon(r) = \frac{1}{r}\int_0^{\frac{\pi}{2}} d\theta\, \frac{\Delta\Sigma\left(\frac{r}{\sin\theta}\right) - \Delta\Sigma(r)}{\cos^2\theta}\,. \tag{C9}$$

Plugging this result into Eq. (C2) gives the desired result Eq. (9).

D. LINE-OF-SIGHT AVERAGE OF INFERRED MASS

As discussed in Sec. 3.5, our deprojection formulas from Sec. 3.1 were derived assuming spherical symmetry. Here we show that, nevertheless, if we average over all line-of-sight directions, our deprojection formulas Eq. (7) and Eq. (8) produce the true mass $M_{\rm true}$ (see Eq. (12)) even for non-symmetric mass distributions. This result holds given that two conditions are satisfied: (i) We restrict to the radial range where $\Sigma/\Sigma_{\rm crit}$ is negligible, and (ii) the source galaxy redshifts follow probability distributions that do not depend on the azimuth (but may depend on projected radius R).

The latter condition disallows a lopsidedness in the source galaxy population, but does allow a radial variation, for example due to obscuration towards the cluster center.

To show that this result holds, we first note that the true mass $M_{\rm true}(r) = \int_{|\vec{x}'|<r} d^3\vec{x}'\rho(\vec{x}')$ (Eq. (12)) can be written as a spherical integral over a spherical density $\bar{\rho}(r)$ obtained by averaging the 3D density $\rho(\vec{x})$ over the solid angle,

$$M_{\rm true}(r) = 4\pi\int_0^r dr' r'^2 \bar{\rho}(r') \quad \text{with} \quad \bar{\rho}(|\vec{x}|) \equiv \frac{1}{4\pi}\int d\Omega\, \rho(R(\Omega)\cdot\vec{x})\,, \tag{D1}$$

where $R(\Omega)$ is a rotation matrix that implements rotation by Ω. The proof below then proceeds roughly as follows: Due to the assumption that $\Sigma/\Sigma_{\rm crit} \ll 1$, the deprojection procedure from Sec. 3.1 becomes *linear* in the density ρ (see the denominator of the right-hand side in Eq. (3)). As a result, averaging the inferred mass $M_{\rm inferred}$ over the line of sight becomes equivalent to running the deprojection procedure for a (fictitious) cluster that has a mass density $\bar{\rho}(r)$. Since $\bar{\rho}$ is spherically symmetric, the deprojection procedure correctly infers the mass associated with $\bar{\rho}(r)$ which, according to Eq. (D1), is just $M_{\rm true}$. In the following, we work out these steps in detail.

The deprojection formulas from Sec. 3.1 are based on the observable $G_+ = \langle g_+\rangle/\langle\Sigma_{\rm crit}^{-1}\rangle$. As a first step, we will show that, in the radial range where $\Sigma/\Sigma_{\rm crit} \ll 1$, we have

$$G_+(R) = \Delta\Sigma_g(R) \quad \text{with} \quad \Delta\Sigma_g(R) \equiv \frac{M_{2D}(R)}{\pi R^2} - \langle\Sigma\rangle(R)\,, \tag{D2}$$

where $\langle\Sigma\rangle$ is the azimuthally-averaged surface density and the subscript g in $\Delta\Sigma_g$ indicates that this definition of $\Delta\Sigma$ applies more generally than the definition given in Eq. (6) which applies only in spherical symmetry. Assuming spherical symmetry and $\Sigma/\Sigma_{\rm crit} \ll 1$, we can obtain the very similar result $G_+ = \Delta\Sigma$ from Eq. (3) (with $\Delta\Sigma$ defined by Eq. (6)). The important difference is that Eq. (D2) holds without any symmetry assumptions on the mass distribution. To see this, we first consider the average $\langle g_+\rangle(R)$ of the reduced tangential shear $g_+ = \gamma_+/(1-\Sigma/\Sigma_{\rm crit})$ (Bartelmann & Schneider 2001). The average $\langle\dots\rangle$ can be understood as, first, separately at each position $(R\cos\varphi, R\sin\varphi)$ averaging over the source redshifts z_s, and then averaging azimuthally over φ,

$$\langle g_+\rangle(R) = \int\frac{d\varphi}{2\pi}\int dz_s\, p(z_s|R,\varphi)\, g_+(R,\varphi) = \int\frac{d\varphi}{2\pi}\int dz_s\, p(z_s|R,\varphi)\,\gamma_+(R,\varphi)\,, \tag{D3}$$

where we used the assumption $\Sigma/\Sigma_{\rm crit} \ll 1$, which implies that g_+ becomes γ_+. Further, our assumption that the source redshifts are drawn from probability distributions that do not depend on the azimuth means that $p(z_s|R,\varphi)$ satisfies $p(z_s|R,\varphi) = p(z_s|R)$. Thus, we can pull the z_s integral including the factor of $p(z_s|R)$ outside the φ integral,

$$\langle g_+\rangle(R) = \int dz_s\, p(z_s|R)\Sigma_{{\rm crit},ls}^{-1}\cdot\int d\varphi\,\gamma_+(R,\varphi)\Sigma_{{\rm crit},ls}\,, \tag{D4}$$

where we also judiciously introduced factors of $\Sigma_{{\rm crit},ls}$. Importantly, the azimuthal average of $\gamma_+\Sigma_{\rm crit}$ gives $\Delta\Sigma_g$ without having to assume any symmetry of the mass distribution (Kaiser 1995; Bartelmann 1995). Since $\Delta\Sigma_g$ is a property of only the lens, it is independent of the source redshifts z_s, and we can pull it outside the z_s integral,

$$\langle g_+\rangle(R) = \Delta\Sigma_g(R)\cdot\int dz_s\, p(z_s|R)\Sigma_{{\rm crit},ls}^{-1} = \Delta\Sigma_g(R)\cdot\langle\Sigma_{{\rm crit},ls}^{-1}\rangle\,. \tag{D5}$$

This is the desired result Eq. (D2) after using the definition $G_+ = \langle g_+\rangle/\langle\Sigma_{\rm crit}^{-1}\rangle$.

We note that, as discussed in Mistele & Durakovic (2024), if individual redshift estimates for the source galaxies are available, the definition of G_+ can be changed to $G_+ = \langle\Sigma_{{\rm crit},ls}\, g_+\rangle$. In this case, the desired Eq. (D2) follows even without having to assume that $p(z_s|R,\varphi)$ is independent of φ.

In any case, the input to our deprojection formulas Eq. (7) and Eq. (8) from Sec. 3.1 are now $G_+(R) = \Delta\Sigma_g(R)$ and f_c. The first of these two deprojection formulas, Eq. (7), reduces to "$\Delta\Sigma$" $= \Delta\Sigma_g$ in the radial range where $\Sigma/\Sigma_{\rm crit}$ is negligible (here, "$\Delta\Sigma$" is understood as the result of evaluating the right-hand side of Eq. (7), which is to be inserted into Eq. (8). Its formal definition Eq. (6) does not apply outside spherical symmetry). This follows from Eq. (3) by using the fact that $\Sigma/\Sigma_{\rm crit} \ll 1$ implies $f_c\Sigma \ll 1$ for any reasonably-behaved source redshift distribution. Thus, this first deprojection step is trivial in the radial range where $\Sigma/\Sigma_{\rm crit} \ll 1$. The remaining second step of the deprojection procedure is Eq. (8) with $\Delta\Sigma_g$ as the integrand. Specifically, the mass $M_{\rm inferred}$ we infer is

$$M_{\rm inferred}(r) \equiv 4r^2\int_0^{\pi/2} d\theta\,\Delta\Sigma_g\left(\frac{r}{\sin\theta}\right) = 4r^2\int_0^{\pi/2} d\theta\,\left(\frac{M_{2D}(R)}{\pi R^2} - \langle\Sigma\rangle(R)\right)\Bigg|_{R=\frac{r}{\sin\theta}}\,. \tag{D6}$$

Here, $M_{2D}(R)$ is the mass enclosed in the line-of-sight cylinder with radius R, which can be written as an integral over the azimuthally-averaged surface density $\langle\Sigma\rangle$,

$$M_{\rm inferred}(r) = 4r^2\int_0^{\pi/2} d\theta\,\left(\frac{2}{R^2}\int_0^R dR' R'\langle\Sigma\rangle(R') - \langle\Sigma\rangle(R)\right)\Bigg|_{R=\frac{r}{\sin\theta}}\,. \tag{D7}$$

The right-hand side now depends only on the azimuthally-averaged surface density $\langle\Sigma\rangle$.

Consider averaging this $M_{\rm inferred}$ over all line-of-sight directions. This amounts to calculating

$$\frac{1}{4\pi}\int d\Omega\, M^{\Omega}_{\rm inferred}(r)\,, \tag{D8}$$

where the integral over the solid angle Ω corresponds to the line-of-sight average and $M^{\Omega}_{\rm inferred}$ denotes the mass we infer when the underlying density is rotated by Ω with respect to the original mass distribution. We refer to the density of the rotated mass distribution as ρ^{Ω},

$$\rho^{\Omega}(\vec{x}) \equiv \rho(R(\Omega)\cdot\vec{x})\,. \tag{D9}$$

Similarly, we denote the azimuthally-averaged surface density corresponding to ρ^{Ω} by $\langle\Sigma\rangle^{\Omega}$. Concretely, $M^{\Omega}_{\rm inferred}$ is given by Eq. (D7) with $\langle\Sigma\rangle$ replaced by $\langle\Sigma\rangle^{\Omega}$. Let's explicitly write out this expression for $M^{\Omega}_{\rm inferred}$ in terms of ρ^{Ω},

$$4r^2\int_0^{\pi/2}d\theta\int\frac{d\varphi}{2\pi}\int dz\left(\frac{2}{(r/\sin\theta)^2}\int_0^{\frac{r}{\sin\theta}}dR'\,R'\rho^{\Omega}(R'\cos\varphi,R'\sin\varphi,z)-\rho^{\Omega}\left(\frac{r}{\sin\theta}\cos\varphi,\frac{r}{\sin\theta}\sin\varphi,z\right)\right)\,. \tag{D10}$$

This is linear in ρ^{Ω} and depends on Ω only through ρ^{Ω}. Thus, when averaging $M^{\Omega}_{\rm inferred}$ over the line-of-sight directions as in Eq. (D8), we can move the Ω integral past the other integrals and find

$$\frac{1}{4\pi}\int d\Omega\, M^{\Omega}_{\rm inferred} = 4r^2\int_0^{\pi/2}d\theta\int dz\left(\frac{2}{R^2}\int_0^{R}dR'\,R'\bar{\rho}\left(\sqrt{R'^2+z^2}\right)-\bar{\rho}\left(\sqrt{R^2+z^2}\right)\right)\Bigg|_{R=\frac{r}{\sin\theta}} \tag{D11}$$

$$= 4r^2\int_0^{\pi/2}d\theta\left(\frac{2}{R^2}\int_0^{R}dR'\,R'\bar{\Sigma}(R')-\bar{\Sigma}(R)\right)\Bigg|_{R=\frac{r}{\sin\theta}}\,, \tag{D12}$$

where $\bar{\rho}(r)=\frac{1}{4\pi}\int d\Omega\,\rho(R(\Omega)\cdot\vec{x})$ as in Eq. (D1). The integrand of the θ integral on the right-hand side is the excess surface density $\Delta\bar{\Sigma}(R)$ of a (fictitious) lens with density $\bar{\rho}$ evaluated at $R=r/\sin\theta$,

$$\frac{1}{4\pi}\int d\Omega\, M^{\Omega}_{\rm inferred} = 4r^2\int_0^{\pi/2}d\theta\,\Delta\bar{\Sigma}\left(\frac{r}{\sin\theta}\right)\,. \tag{D13}$$

Since that fictitious lens is spherically symmetric, the θ deprojection integral will infer the corresponding mass $\bar{M}$,

$$\frac{1}{4\pi}\int d\Omega\, M^{\Omega}_{\rm inferred} = \bar{M}(r)\equiv 4\pi\int_0^{r}dr'r'^2\bar{\rho}(r')\,. \tag{D14}$$

According to Eq. (D1), this is the same as $M_{\rm true}$, which is what was to be shown.

E. NFW EXTRAPOLATION

As an alternative to extrapolating G_+ beyond the last measured data point by assuming a power law, we here consider extrapolation assuming an NFW profile. In particular, we assume that, beyond the last measured data point at $R=R_{\rm max}$, the shear profile $G_+=\langle g_+\rangle/\langle\Sigma_{\rm crit}^{-1}\rangle$ is given by (see Eq. (3))

$$G_+(R>R_{\rm max}) = \frac{\Delta\Sigma_{\rm NFW}(R|M^{\rm match}_{200c})}{1-f_c\,\Sigma_{\rm NFW}(R|M^{\rm match}_{200c})}\,, \tag{E1}$$

where $\Delta\Sigma_{\rm NFW}(R|M_{200c})$ and $\Sigma_{\rm NFW}(R|M_{200c})$ denote, respectively, the excess surface density and surface density of an NFW halo with a given M_{200c}. Explicit formulas for $\Delta\Sigma_{\rm NFW}$ and $\Sigma_{\rm NFW}$ are given, for example, in Umetsu (2020). For simplicity, we here fix the concentration c_{200c} by assuming the WMAP5 mass-concentration relation from Macciò et al. (2008), so that the NFW profiles are fully specified by M_{200c} alone. Choosing a different mass-concentration relation does not significantly change our results. When extrapolating G_+ using Eq. (E1), we determine M_{200c} by matching to the observed shear G_+ at the last data point at $R_{\rm max}$ and denote the result by $M^{\rm match}_{200c}$,

$$G_+(R=R_{\rm max}) = \frac{\Delta\Sigma_{\rm NFW}(R_{\rm max}|M^{\rm match}_{200c})}{1-f_c\,\Sigma_{\rm NFW}(R_{\rm max}|M^{\rm match}_{200c})}\,. \tag{E2}$$

We note that $M^{\rm match}_{200c}$ does not necessarily coincide with the cluster's actual M_{200c} as inferred from the full, deprojected mass profile $M(r)$, hence the different notation.

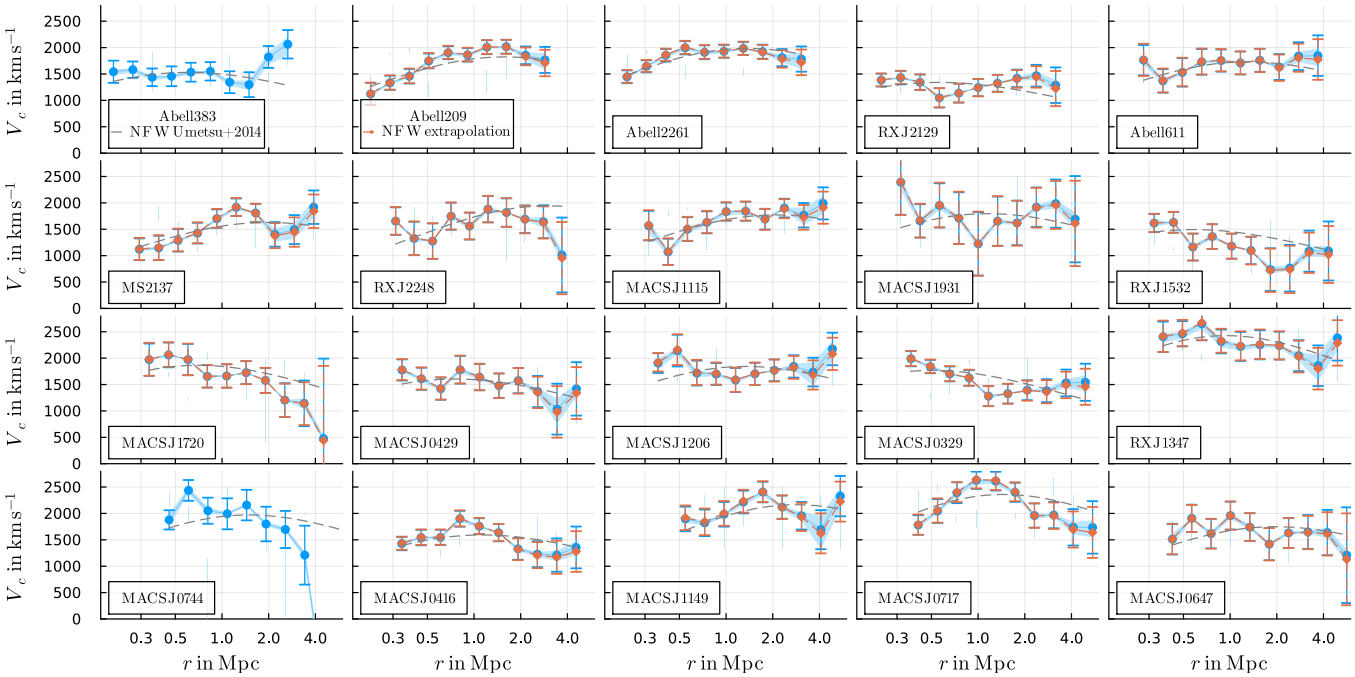

FIG. 17.— Same as Fig. 5, but additionally showing results when extrapolating the shear profile G_+ assuming an NFW profile (red symbols) instead of assuming a $1/R$ power law (blue symbols). We do not show results without the two-halo subtraction for visual clarity. Dashed gray lines are the circular velocities implied by the NFW fits from Umetsu et al. (2014), which do not take into account the two-halo term. There is no result with NFW extrapolation for Abell 383 since no value for M_{200c} can be determined from its non-parametric mass profile (see Sec. 4.1) and we cannot switch to M_{500c} because our mass-concentration relation works only for M_{200c}. Similarly, there is no result with NFW extrapolation for MACS J0744 because the last G_+ data point is negative, so our simple matching procedure does not find an NFW profile to extrapolate with.

The two-halo subtraction procedure described in Sec. 3.2 and Appendix A is mostly unchanged when extrapolating with an NFW profile. We only need to replace Eq. (A12) for the two-halo contribution to the inferred mass $M_{2\rm h}$ with

$$M_{2\rm h}(r|M_{200c}) = 4r^2 \left[\int_{\theta_{\rm min}}^{\pi/2} d\theta\, \Delta\Sigma_{2\rm h}\left(\frac{r}{\sin\theta}\right) + \int_0^{\theta_{\rm min}} d\theta \left(\Delta\Sigma_{\rm NFW}\left(\frac{r}{\sin\theta}\middle| M_{200c}^{\rm match,obs}\right) - \Delta\Sigma_{\rm NFW}\left(\frac{r}{\sin\theta}\middle| M_{200c}^{\rm match,sub}\right)\right)\right] . \quad \text{(E3)}$$

Here, $M_{200c}^{\rm match,obs}$ is determined by matching an NFW profile to the observed $G_+(R_{\rm max})$ as in Eq. (E2), and $M_{200c}^{\rm match,sub}$ is obtained by matching an NFW profile to $G_+(R_{\rm max})$ minus the two-halo contribution, i.e. by matching an NFW profile to $G_+ - G_{+,2\rm h}$ at $R = R_{\rm max}$. We note that the two-halo contribution $G_{+,2\rm h}$ depends on the actual M_{200c} (as determined by the full, deprojected mass profile $M(r)$) through the bias factor $b(z_l, M_{200c})$ (see Appendix A) and, therefore, so does $M_{200c}^{\rm match,sub}$.

In practice, since it makes the code simpler and faster, we actually determine $M_{200c}^{\rm match,sub}$ from the equation

$$\Delta\Sigma_{\rm NFW}(R_{\rm max}|M_{200c}^{\rm match,obs}) - \Delta\Sigma_{2\rm h}(R_{\rm max}) = \Delta\Sigma_{\rm NFW}(R_{\rm max}|M_{200c}^{\rm match,sub}) . \quad \text{(E4)}$$

If the two-halo term is negligibly small at $R_{\rm max}$, this gives $M_{200c}^{\rm match,obs} = M_{200c}^{\rm match,sub}$ which is the correct outcome in this case, i.e. it is what we would have obtained from matching an NFW profile to $G_+ - G_{+,2\rm h}$ at $R_{\rm max}$. When the two-halo term becomes non-negligible, we are very likely at sufficiently large radii for the difference between G_+ and $\Delta\Sigma$ to be unimportant (see Appendix A). Thus, in this case, the outcome will again match the outcome of matching to $G_+ - G_{+,2\rm h}$. This justifies our simplified procedure for determining $M_{200c}^{\rm match,sub}$.

Figure 17 shows the cirular velocities inferred using the NFW extrapolation described above, including the adjusted two-halo subtraction procedure. At large radii there is a small difference compared to our fiducial $1/R$ extrapolation. This small difference is adequately captured by our systematic uncertainty band which is spanned by extrapolation with $1/\sqrt{R}$ and $1/R^2$ power laws. For reference, Fig. 17 also shows the circular velocities implied by the NFW fits from Umetsu et al. (2014). These NFW fits do not take into account the two-halo term.

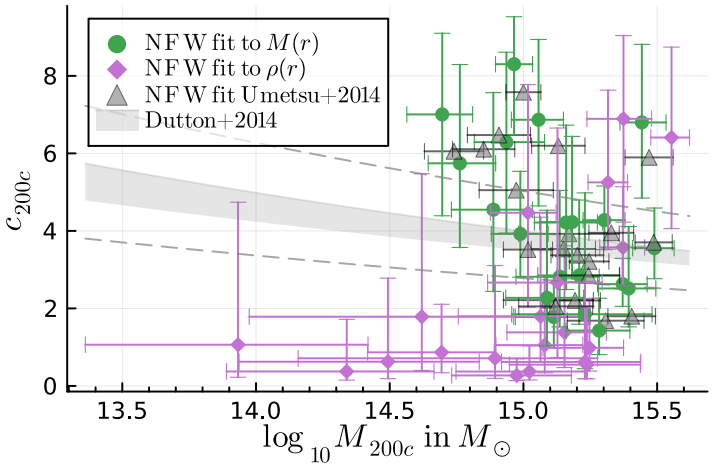

FIG. 18.— The best-fit parameters obtained by fitting NFW profiles to our non-parametric mass profiles $M(r)$ from Fig. 6 (green circles) and to our non-parametric density profiles $\rho(r)$ from Fig. 9 (purple diamonds). Unlike the mass profiles $M(r)$, the density profiles $\rho(r)$ do not contain information about the mass distribution within the smallest radial bin, leading to different fit results. No two-halo subtraction is performed, allowing for a direct comparison to the NFW fits from Umetsu et al. (2014) (gray triangles). No uncertainties on the concentrations are given in Umetsu et al. (2014) so these are not shown. The shaded gray region indicates the mass-concentration relation from Dutton & Macciò (2014) for the range of redshifts of our cluster sample. Gray dashed lines indicate 0.1 dex scatter around that region.

TABLE 3
NFW FIT RESULTS

Name	NFW fit to $M(r)$		NFW fit to $\rho(r)$	
	$\log_{10} M_{200c}$ $M_\odot$	c_{200c}	$\log_{10} M_{200c}$ $M_\odot$	c_{200c}
Abell 383	$14.94^{+0.14}_{-0.12}$	$6.29^{+2.32}_{-2.34}$	$15.02^{+0.21}_{-0.27}$	$0.37^{+0.66}_{-0.22}$
Abell 209	$15.37^{+0.09}_{-0.09}$	$2.63^{+0.67}_{-0.57}$	$15.37^{+0.08}_{-0.09}$	$3.58^{+1.56}_{-1.07}$
Abell 2261	$15.30^{+0.08}_{-0.08}$	$4.28^{+0.88}_{-0.76}$	$15.32^{+0.07}_{-0.07}$	$5.25^{+2.38}_{-1.72}$
RXJ2129	$14.76^{+0.13}_{-0.12}$	$5.74^{+2.55}_{-2.17}$	$14.69^{+0.20}_{-0.28}$	$0.87^{+1.24}_{-0.53}$
Abell 611	$15.21^{+0.15}_{-0.15}$	$2.85^{+1.94}_{-1.26}$	$15.15^{+0.16}_{-0.22}$	$1.38^{+2.23}_{-0.91}$
MS2137	$15.11^{+0.13}_{-0.14}$	$1.78^{+1.07}_{-0.80}$	$15.08^{+0.14}_{-0.18}$	$1.05^{+1.75}_{-0.71}$
RXJ2248	$15.14^{+0.15}_{-0.16}$	$2.82^{+2.23}_{-1.29}$	$15.06^{+0.19}_{-0.31}$	$1.79^{+2.56}_{-1.17}$
MACSJ1115	$15.28^{+0.12}_{-0.12}$	$1.43^{+0.84}_{-0.62}$	$15.25^{+0.13}_{-0.17}$	$0.98^{+1.27}_{-0.60}$
MACSJ1931	$15.23^{+0.25}_{-0.27}$	$1.85^{+3.13}_{-1.40}$	$14.89^{+0.38}_{-0.74}$	$0.71^{+2.39}_{-0.52}$
RXJ1532	$14.70^{+0.11}_{-0.13}$	$7.01^{+2.09}_{-2.62}$	$13.93^{+0.49}_{-0.57}$	$1.06^{+3.68}_{-0.84}$
MACSJ1720	$15.06^{+0.09}_{-0.10}$	$6.87^{+2.07}_{-2.22}$	$14.62^{+0.34}_{-0.65}$	$1.79^{+3.69}_{-1.39}$
MACSJ0429	$14.89^{+0.15}_{-0.16}$	$4.55^{+3.02}_{-2.11}$	$14.49^{+0.34}_{-0.56}$	$0.62^{+2.16}_{-0.44}$
MACSJ1206	$15.16^{+0.12}_{-0.12}$	$4.22^{+2.51}_{-1.73}$	$14.97^{+0.20}_{-0.24}$	$0.27^{+0.43}_{-0.13}$
MACSJ0329	$14.96^{+0.07}_{-0.07}$	$8.30^{+1.23}_{-1.82}$	$14.34^{+0.33}_{-0.45}$	$0.37^{+1.35}_{-0.22}$
RXJ1347	$15.44^{+0.09}_{-0.08}$	$6.80^{+2.01}_{-1.95}$	$15.23^{+0.21}_{-0.31}$	$0.62^{+2.08}_{-0.43}$
MACSJ0744	$15.18^{+0.11}_{-0.12}$	$4.23^{+2.21}_{-1.46}$	$15.37^{+0.11}_{-0.14}$	$6.89^{+2.15}_{-2.67}$
MACSJ0416	$14.99^{+0.10}_{-0.11}$	$3.92^{+1.61}_{-1.12}$	$15.02^{+0.11}_{-0.14}$	$4.47^{+3.31}_{-2.43}$
MACSJ1149	$15.39^{+0.10}_{-0.11}$	$2.52^{+1.60}_{-0.99}$	$15.24^{+0.20}_{-0.29}$	$0.55^{+1.50}_{-0.36}$
MACSJ0717	$15.49^{+0.07}_{-0.08}$	$3.55^{+1.04}_{-0.79}$	$15.55^{+0.07}_{-0.08}$	$6.41^{+2.33}_{-2.35}$
MACSJ0647	$15.09^{+0.14}_{-0.15}$	$2.27^{+2.26}_{-1.24}$	$15.13^{+0.17}_{-0.26}$	$2.67^{+3.99}_{-1.94}$

NOTE. — The listed values are the 16th, 50th, and 84th percentiles.

F. NFW FITS TO NON-PARAMETRIC MASS AND DENSITY PROFILES

We have fit NFW profiles to both our non-parametric mass profiles $M(r)$ and our non-parametric density profiles $\rho(r)$. We loosely follow Umetsu et al. (2014) in using a Bayesian fitting procedure with flat priors for $13 < \log_{10} M_{200c}/M_\odot < 17$ and $-1 < \log_{10} c_{200c} < 1$. We use the julia package 'Turing.jl' (Ge et al. 2018) with the 'Emcee()' sampler (Foreman-Mackey et al. 2013). For $\rho(r)$, but not $M(r)$, it can happen that the covariance matrix has a zero eigenvalue. For example, a direct calculation shows that, in a simple setup with just two radial bins R_1 and R_2 and linear interpolation, it can happen that the inferred density in the first bin, $\rho(r = R_1)$, is completely independent of the shear measurement at R_1, due to a cancellation of different terms in Eq. (9). The reconstructed ρ values in the two radial bins are

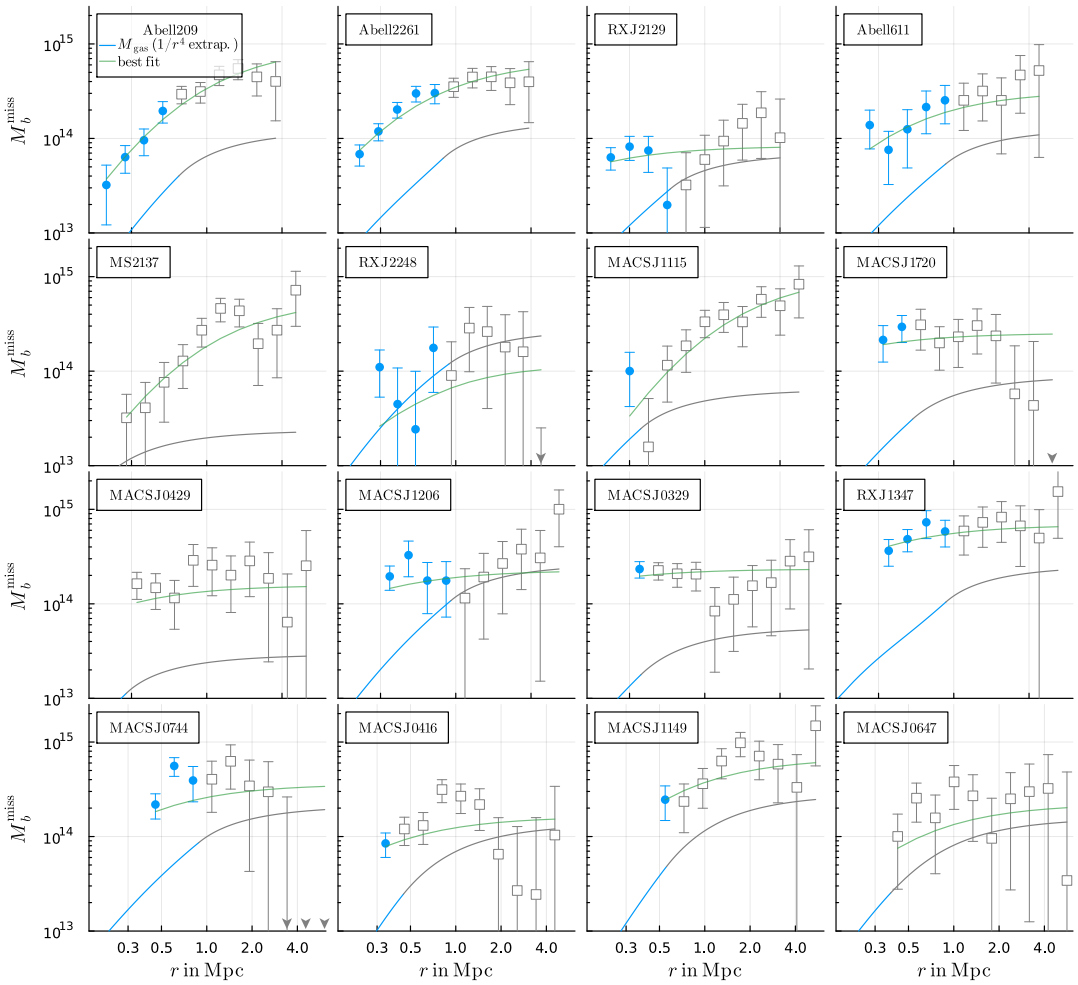

FIG. 19.— The missing baryonic mass implied by assuming that the galaxy-scale RAR holds also for galaxy clusters. Blue circles indicate the radial range where X-ray and weak-lensing observations overlap. Beyond that range, we extrapolate the gas densities assuming a $1/r^4$ tail (gray squares). Arrowheads at the horizontal axis indicate negative missing mass. We do not subtract the two-halo term. Solid blue and gray lines show the gas mass profiles. The solid green line indicates the best fit to $M_b^{\rm miss}$ following the fitting procedure of Kelleher & Lelli (2024).

then 100% correlated; both are completely determined by the shear at R_2. This leads to a zero eigenvalue in the covariance matrix. To avoid numerical issues with the inverse covariance matrix, we detect such behavior and remove the corresponding eigenvector from the fit, leaving only the orthogonal subspace of densities to be fit. We apply this removal procedure when the smallest eigenvalue of the correlation matrix is at least 100 times smaller than the second-smallest eigenvalue.

The best-fit parameters are shown in Fig. 18 and Table 3. We did not apply our two-halo subtraction procedure to allow a more direct comparison to Umetsu et al. (2014). When fitting to $M(r)$, we find best-fit parameters consistent with those of Umetsu et al. (2014).

However, fitting our non-parametric *density* profiles does not recover the same fit parameters. In particular, concentrations are systematically smaller when fitting $\rho(r)$. In addition, the statistical uncertainties on the NFW parameters are larger, with concentrations being particularly poorly constrained. This is because in going from $M(r)$ to $\rho(r)$ one loses information, unless $\rho(r)$ is measured all the way down to $r = 0$: Weak-lensing observations do not extend all the way to the centers of clusters, so there is a minimum radius $r_{\rm min}$ down to which we infer $\rho(r)$. As a result, we cannot reconstruct $M(r)$ from $\rho(r)$. We can only reconstruct $M(r) - M(r_{\rm min})$, because the density $\rho(r)$ at $r > r_{\rm min}$ does not know anything about the mass distribution within $r_{\rm min}$. In contrast, the mass profile $M(r)$ at $r > r_{\rm min}$ does. It knows the total amount of mass within $r_{\rm min}$.

There is some amount of degeneracy between concentration and mass even when fitting to $M(r)$. This may be why the Dutton & Macciò (2014) mass-concentration relation shown in Fig. 18 seems not to be followed very closely even by our fits to $M(r)$. Breaking this degeneracy may require extending the radial range covered by observations, for example using strong lensing (e.g., Merten et al. 2015; Umetsu et al. 2016, 2025).

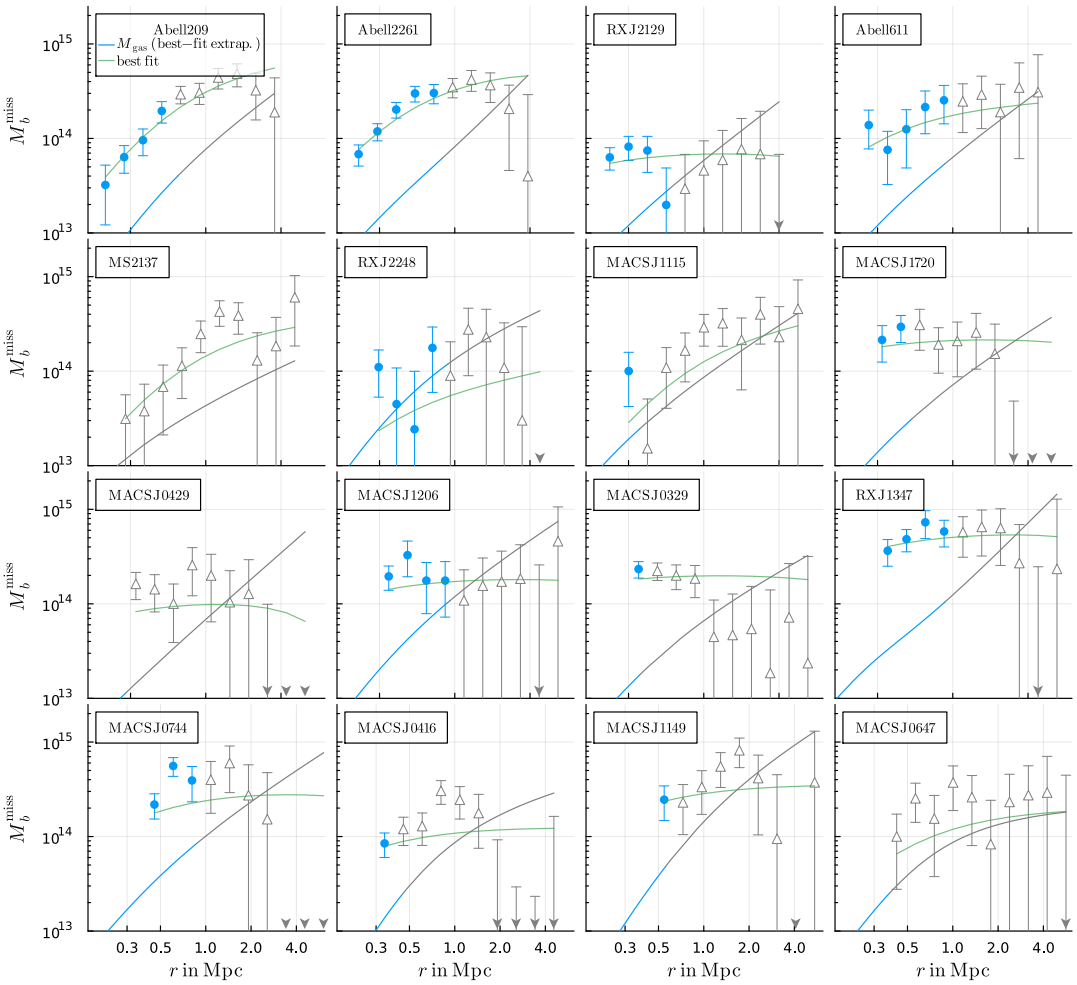

FIG. 20.— Same as Fig. 19 but extrapolating the gas densities assuming the best-fit beta profiles.

G. MISSING MASS ASSUMING THE GALAXY-SCALE RAR

The galaxy-scale RAR can be parametrized as a relation between $g_{\rm obs}$ and $g_{\rm bar}$ with $\mu(|g_{\rm obs}|/a_0)\, g_{\rm obs} = g_{\rm bar}$, where μ is the so-called interpolation function. We adopt the so-called "simple" interpolation function $\mu(x) = x/(1+x)$. If we assume the galaxy-scale RAR to hold universally, the missing baryonic mass $M_b^{\rm miss}$ is then given by

$$\frac{G_N M_b^{\rm miss}(r)}{r^2} = \mu\left(\frac{|g_{\rm obs}(r)|}{a_0}\right) g_{\rm obs}(r) - g_{\rm bar}(r)\,. \tag{G1}$$

This $M_b^{\rm miss}$ must be a monotonic function of radius if the mismatch between our measurements of $\mu(|g_{\rm obs}|)g_{\rm obs}$ and $g_{\rm bar}$ is indeed due to missing baryons.

Figure 19 shows the missing mass $M_b^{\rm miss}$ implied by our weak-lensing analysis and assuming a $1/r^4$ extrapolation for the gas densities. We do not subtract the two-halo term. Subtracting the two-halo term does not significantly change the results. In the region where X-ray and weak-lensing observations overlap, we find that $M_b^{\rm miss}$ is a monotonic function within the uncertainties. An exception are the innermost three data points of RX J2129, which, however, may not be reliable (Sec. 4.1). At larger radii, some clusters have non-monotonic and even negative $M_b^{\rm miss}$, but the error bars are large and often consistent with the monotonic mass profiles $M_{\rm mm}(r)$ (solid green lines, see below).

Figure 20 shows the $M_b^{\rm miss}$ implied by assuming the beta profile fits from Famaey et al. (2025) are valid even beyond $R_{\rm max}^X$. This increases $g_{\rm bar}$ at large radii, implying more non-monotonicities in $M_b^{\rm miss}$. Famaey et al. (2025) find that, within 1 Mpc, $M_b^{\rm miss}$ tracks the gas mass. We can confirm that result for a few clusters, for example for Abell 209 and Abell 2261, but it does not seem to hold as univerally as for the NFW fits from Famaey et al. (2025).

The solid green lines in Fig. 19 and Fig. 20 show the best fit missing mass component $M_{\rm mm}(r)$ following Kelleher & Lelli (2024) (Sec. 4.7). This additional mass component $M_{\rm mm}(r)$ is always a monotonic function of r. Despite this, the solid green lines in Fig. 20 can be non-monotonic because they correspond to $M_{\rm mm}(r) + (\Upsilon_b - 1)M_b(r)$ which is the quantity that, for good fits, should match the non-parametric missing mass profiles shown there (which have $\Upsilon_b = 1$).

This paper was built using the Open Journal of Astrophysics LaTeX template. The OJA is a journal which provides

fast and easy peer review for new papers in the `astro-ph` section of the arXiv, making the reviewing process simpler for authors and referees alike. Learn more at http://astro.theoj.org.