Estimating the Virial Mass and Mass-to-Light Ratio of a Rich Galaxy Cluster with SDSS-DR18 Data

Jackson Ferguson

University of Victoria

Astrophysics

March 2025

Abstract

We present a virial mass estimation of the galaxy cluster ACO 2670 using photometric and spectroscopic data from SDSS DR18. By analyzing the radial and redshift distributions of 98 confirmed member galaxies, we determine a cluster redshift of $z = 0.0764$ and a velocity dispersion of $\sigma_v = (932 \pm 95)\,\mathrm{km s^{-1}}$. We derive a virial mass of $M_{1/2} = (6.45 \pm 1.31) \times 10^{14} M_{\odot}$ and a mass-to-light ratio of $(291 \pm 60) M_{\odot}/L_{\odot}$, confirming the presence of a dominant dark matter halo.

1 Introduction

Galaxy clusters, the most massive gravitationally bound structures in the universe, serve as critical laboratories for studying dark matter, galaxy evolution, and large-scale structure formation [12]. Their dynamical properties, such as virial mass and mass-to-light (M/L) ratios, provide key insights into the distribution of luminous and dark matter. However, estimating these quantities requires careful analysis of observational data, robust statistical methods, and assumptions that may influence the results. This study focuses on the galaxy cluster ACO 2670, utilizing photometric and spectroscopic data from the Sloan Digital Sky Survey (SDSS) Data Release 18 (DR18) [11] to determine its virial mass, luminosity distribution, and M/L ratio. The analysis hinges on several foundational assumptions:

1. **Virial Equilibrium:** The cluster is assumed to be dynamically relaxed, enabling the application of the virial theorem. This assumption may not hold for systems with substructure or ongoing mergers, as suggested by ACO 2670's irregular spatial distribution.

2. **Cosmological Model:** Distance calculations rely on the Flat ΛCDM cosmology, with fixed Hubble constant (H_0) and matter density (Ω_m) values. Uncertainties in these parameters propagate to distance and mass estimates.

3. **Data Completeness:** SDSS spectroscopic data are assumed to represent a complete and unbiased sample of cluster galaxies, though selection effects (e.g., fiber collisions, magnitude limits) may exclude faint or crowded objects.

4. **Projection Effects:** Galaxy velocities and positions are analyzed in projection, neglecting the three-dimensional structure of the cluster.

5. **Iterative Membership Selection:** The 1.5 Mpc radial cutoff and $\pm 3\sigma$ redshift clipping are assumed to effectively isolate true cluster members, though interlopers or incomplete sampling may persist.

The analysis assumes a Flat ΛCDM cosmology with parameters from Planck Collaboration et al. [6]: Hubble constant $H_0 = 67.4\,\mathrm{km\,s^{-1}Mpc^{-1}}$, matter density $\Omega_m = 0.315$, and dark energy density $\Omega_\Lambda = 0.685$. These values define the universe's expansion history and were used for luminosity distance calculations.

This work employs a multi-step methodology: refining cluster membership iteratively, computing the weighted

cluster center, analyzing redshift distributions using the Freedman-Diaconis rule [4] and Bayesian Information Criterion [8], calculating relativistic distances [5, 1], and applying the virial theorem to estimate mass. The derived M/L ratio is contextualized against stellar population models to infer dark matter content.

2 Data Acquisition and Cleaning

To analyze the ACO 2670 galaxy cluster, photometric and spectroscopic data were retrieved from the Sloan Digital Sky Survey (SDSS) Data Release 18 (DR18). The dataset included the positions, magnitudes, extinction corrections, and redshifts of galaxies in a 2 Mpc region centered on the cluster. The SQL query used for data retrieval applied spatial and redshift constraints to ensure that the sample included only galaxies near the cluster while filtering out foreground and background objects. The initial member selection used a 3σ clipping method, retaining galaxies within three standard deviations of the mean redshift.

During data analysis, an outlier galaxy with an unusually dim absolute magnitude ($M_r = -16$) was identified. Inspection of the SDSS Explore tool revealed that this object was photometrically classified as a star and flagged for edge deblending and saturation. Despite its spectroscopic classification as a galaxy, the photometric flags and star-like properties suggested unreliable data. This object was excluded to prevent photometric outliers from biasing the luminosity-weighted centroid.

3 Spatial Analysis and Centroid Determination

To analyze the spatial distribution of galaxies in ACO 2670, a scatter plot of Right Ascension (RA) vs. Declination (Dec) was created (Figure 1).

The resulting plot revealed a clear concentration of galaxies around the center of the plot, but without a well-defined, sharply bounded core. Instead of a smooth density gradient transitioning from high-density central regions to lower-density outskirts, the galaxy distribution was irregular, with varying densities and possible substructures. The cluster did not exhibit perfect symmetry, and a mild elongation along the positive diagonal (RA
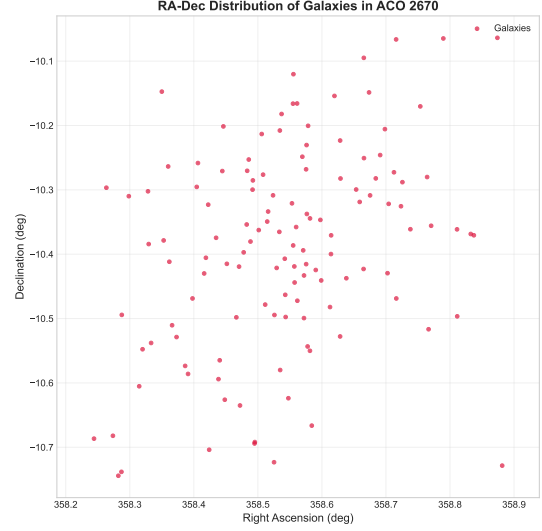

Figure 1: RA-Dec Distribution of Galaxies in ACO 2670.

increasing with Dec) was observed.

A key challenge in this step was interpreting the structural properties of the cluster. Unlike relaxed clusters with clear, well-defined centers, ACO 2670 showed evidence of substructure, possibly indicating a dynamically active system or recent mergers. These characteristics suggest that ACO 2670 may not be in a fully virialized state, which could impact subsequent mass estimates.

3.1 Weighted Mean Cluster Center

We compute the weighted mean Right Ascension (RA) and Declination (Dec) using inverse squared r-band magnitude errors as weights. Additionally, we determine the uncertainties associated with these weighted means.

1. Compute the Weights

The weights are based on the inverse squared r-band magnitude errors:

$$w_i = \frac{1}{\sigma_{r,i}^2} \tag{1}$$

where w_i is the weight for galaxy i and $\sigma_{r,i}$ is the r-band magnitude error for galaxy i.

2. Compute the Weighted Mean RA and Dec

The weighted mean Right Ascension ($\overline{\alpha}$) and weighted mean Declination ($\overline{\delta}$) are calculated as:

$$\overline{\alpha} = \frac{\sum_i w_i \alpha_i}{\sum_i w_i}, \quad \overline{\delta} = \frac{\sum_i w_i \delta_i}{\sum_i w_i} \qquad (2)$$

where α_i and δ_i are the RA and Dec of galaxy i, and w_i is the corresponding weight.

3. Compute the Weighted Variance

The weighted variance accounts for both the measurement errors and the intrinsic scatter in the data. We compute the weighted variances as:

$$\sigma_\alpha^2 = \frac{\sum_i w_i (\alpha_i - \overline{\alpha})^2}{\sum_i w_i} \cdot \frac{N}{N-1} \qquad (3)$$

$$\sigma_\delta^2 = \frac{\sum_i w_i (\delta_i - \overline{\delta})^2}{\sum_i w_i} \cdot \frac{N}{N-1} \qquad (4)$$

where N is the number of member galaxies. The factor $\frac{N}{N-1}$ is a bias correction for the weighted variance to ensure an unbiased estimator.

4. Compute the Standard Error of the Weighted Mean

The uncertainty in the weighted mean RA and Dec is given by:

$$\sigma_{\overline{\alpha}} = \sqrt{\frac{\sigma_\alpha^2}{\sum_i w_i}}, \quad \sigma_{\overline{\delta}} = \sqrt{\frac{\sigma_\delta^2}{\sum_i w_i}} \qquad (5)$$

The final computed cluster center was at $\overline{\alpha} = (358.5546 \pm 0.0003)°$, $\overline{\delta} = (-10.3768 \pm 0.0003)°$. This result provides a more accurate sky position of the galaxy cluster compared to a simple arithmetic mean, as it accounts for measurement uncertainties.

One challenge was ensuring that the weighting method was appropriate, alternative weighting schemes, such as using the r-band magnitudes directly, could have been considered. Additionally, missing or zero-error values in the dataset could have affected the calculations, requiring proper filtering before computing the weighted mean.

Another challenge I had was determining what level of precision to report my results. The computed cluster center (RA, Dec) is reported to four decimal places, as SDSS astrometric uncertainties are typically 0.03–0.3 arcseconds (0.000008–0.00008 degrees). This level of precision ensures that the reported values are meaningful and not artificially over-precise relative to the data quality.

4 Kinematics and Velocity Dispersion

The redshift distribution of the cluster provides crucial insights into its dynamical state and potential substructures. By analyzing the histogram of galaxy redshifts, we can determine the cluster's mean redshift ($z_{cluster}$) and velocity dispersion. A single Gaussian fit is commonly used to describe relaxed clusters, but deviations from a Gaussian distribution can indicate merging substructures. To ensure a robust analysis, we compare a single Gaussian fit to a Gaussian Mixture Model (GMM) with two components, selecting the best model using the Bayesian Information Criterion (BIC).

1. Binning the Redshift Data

The optimal bin width Δz is chosen using the Freedman-Diaconis rule:

$$\Delta z = \frac{2 \times \text{IQR}(z)}{N^{1/3}} \qquad (6)$$

where $\text{IQR}(z)$ is the interquartile range of the redshifts, and N is the number of galaxies. If this bin width is too small, we use Scott's rule as an alternative:

$$\Delta z = \frac{3.5\sigma_z}{N^{1/3}} \qquad (7)$$

The number of bins is then:

$$N_{bins} = \frac{z_{max} - z_{min}}{\Delta z} \qquad (8)$$

2. Fitting a Weighted Gaussian to the Redshift Distribution

The probability density function of a Gaussian is given by:

$$f(z) = A \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) \qquad (9)$$

where A is the amplitude, μ is the mean redshift ($z_{cluster}$), and σ is the standard deviation (redshift dispersion). Since redshift errors σ_{z_i} vary per galaxy, we use inverse

variance weighting:

$$w_i = \frac{1}{\sigma_{z_i}^2} \quad (10)$$

The weighted mean and standard deviation are computed as:

$$\mu_w = \frac{\sum w_i z_i}{\sum w_i}, \quad \sigma_w = \sqrt{\frac{\sum w_i (z_i - \mu_w)^2}{\sum w_i}} \quad (11)$$

These values serve as initial estimates for fitting the Gaussian function to the binned redshift data.

3. Gaussian Mixture Model (GMM) and Model Comparison

The GMM assumes the redshift distribution is a sum of two Gaussian components:

$$f(z) = \sum_{i=1}^{2} w_i \cdot \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(z - \mu_i)^2}{2\sigma_i^2}\right) \quad (12)$$

where w_i are the mixture weights, and μ_i, σ_i are the means and standard deviations of the components. The Bayesian Information Criterion (BIC) evaluates the model complexity and fit:

$$\text{BIC} = k \ln(N) - 2\ln(L) \quad (13)$$

where k is the number of parameters and L is the maximum likelihood estimate. A lower BIC indicates a better model. For the single Gaussian, we estimate the log-likelihood as:

$$\ln(L) = \sum \ln f(z_i) \quad (14)$$

and compute its BIC. The GMM BIC is obtained directly from the model fit. If the GMM has a significantly lower BIC and both components have non-negligible weights, we select it as the better model.

The redshift histogram follows a Gaussian distribution, with a clear peak at 0.07642 ± 0.00346. The BIC values indicate that the single Gaussian (BIC $= -871.8$) provides a significantly better fit than the GMM (BIC $= -840.3$). Additionally, the second GMM component has a negligible weight (0.03), confirming that no significant substructure is detected in the redshift space. The use of the GMM was motivated by the possibility of merging substructures, which could be revealed in later steps through spatial
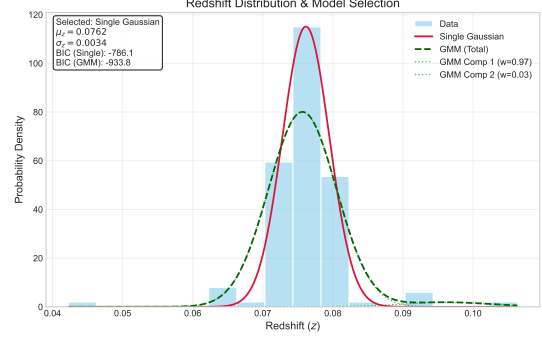

Figure 2: Redshift Distribution with Optimal Binning and Model Fitting. The single Gaussian model (red) is compared against a GMM (green).

and velocity dispersion analysis. However, in this case, the GMM did not improve the fit, suggesting that the cluster is relatively well-virialized in redshift space. The cluster's velocity dispersion derived from the redshift spread will be critical for mass estimation in subsequent steps.

4.1 Distance Estimation

The distance to the galaxy cluster ACO 2670 was initially estimated using Hubble's Law. First, the recessional velocity was calculated using the Doppler shift approximation $v_{rec} \approx cz$. While this method is commonly used for rough distance estimates, it does not account for relativistic effects or the expansion history of the universe, which can introduce errors at higher redshifts.

1. Recessional Velocity Using the Relativistic Doppler Formula

To improve accuracy, a relativistic Doppler shift formula was implemented to refine the recessional velocity calculation:

$$v_{rec} = c \cdot \frac{(1+z)^2 - 1}{(1+z)^2 + 1} \quad (15)$$

where c is the speed of light and z is the redshift of the cluster. This correction is important for objects at moderate redshifts ($z \approx 0.08$), where the simple linear approximation begins to deviate from reality. To compute the uncertainty in v_{rec}, we propagate the uncertainty in redshift (σ_z):

$$\frac{dv_{rec}}{dz} = c \cdot \frac{4(1+z)}{[(1+z)^2 + 1]^2} \implies \sigma_{v_{rec}} = \left|\frac{dv_{rec}}{dz}\right| \sigma_z \quad (16)$$

4

2. Luminosity Distance Using Cosmology

The cluster's luminosity distance is obtained from the Flat ΛCDM cosmology:

$$d_L = D_L(z) \tag{17}$$

where $D_L(z)$ is the luminosity distance function from the Astropy cosmology model. We approximate the uncertainty in distance due to redshift by computing the numerical derivative:

$$\frac{dD_L}{dz} \approx \frac{D_L(z + \delta z) - D_L(z - \delta z)}{2\delta z} \tag{18}$$

where δz is a small redshift step ($\delta z = \sigma_z/10$). Thus, the uncertainty due to redshift is:

$$\sigma_{D_L,z} = |dD_L/dz|\sigma_z \tag{19}$$

Since the distance scales inversely with the Hubble constant (H_0), we estimate the uncertainty due to H_0 as:

$$\sigma_{D_L,H_0} = D_L \times \frac{\sigma_{H_0}}{H_0} \tag{20}$$

The total uncertainty in distance is combined using quadrature addition:

$$\sigma_{D_L} = \sqrt{\sigma_{D_L,z}^2 + \sigma_{D_L,H_0}^2} \tag{21}$$

I calculated a recessional velocity of $(21962.31 \pm 928.65)\,\mathrm{km\,s^{-1}}$ and a cluster distance of $(355.87 \pm 16.66)\,\mathrm{Mpc}$.

4.2 Projected Radial Distribution

To analyze the spatial distribution of the cluster members, I implemented a function to compute the projected distances of galaxies from the cluster center in megaparsecs (Mpc).

1. Converting Coordinates and Computing Angular Separation

Since RA and Dec are given in degrees, we first convert them to radians ($RA_{rad} = \frac{\pi}{180}RA$, $Dec_{rad} = \frac{\pi}{180}Dec$). To compute the angular separation between each galaxy

and the cluster center, we use:

$$\Delta\alpha = (\alpha - \alpha_c)\cos(\delta_c) \tag{22}$$
$$\Delta\delta = \delta - \delta_c \tag{23}$$

The total angular separation in radians is $\theta = \sqrt{\Delta\alpha^2 + \Delta\delta^2}$. The $\cos(\delta_c)$ term accounts for the spherical coordinate distortion where RA lines converge at the poles.

2. Converting to Physical Distance

Since the angular diameter distance of the cluster is known (D_A), we compute the physical projected distance:

$$d_{proj} = \theta \cdot D_A \tag{24}$$

where d_{proj} is the projected distance from the cluster center (in Mpc) and D_A is the angular diameter distance to the cluster.
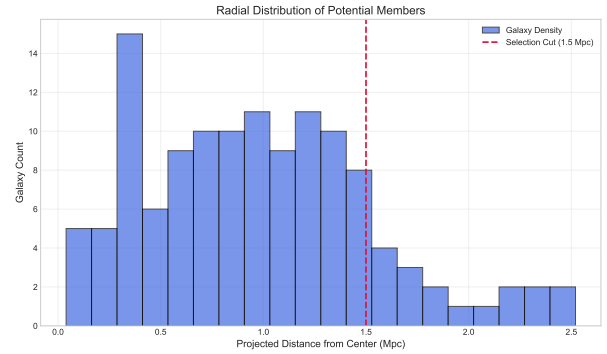

Figure 3: Distribution of Projected Distances of Cluster Member Galaxies. The red dashed line indicates the 1.5 Mpc selection radius.

The resulting distribution revealed two peaks, one at approximately 0.4 Mpc and another at approximately 0.9 Mpc to 1.4 Mpc, suggesting possible substructure or multiple concentrations of galaxies within the cluster. The density of galaxies gradually declined beyond 1.0 Mpc, with a steep drop-off at around 1.8 Mpc, indicating the approximate outer boundary of the cluster. Interestingly, a small secondary peak was observed between 2.5 Mpc and 2.9 , where six galaxies were detected. This could be due to a small infalling group of galaxies, or a background structure projected along the same line of sight. The majority of the galaxies were contained within 1.5 Mpc, aligning with the selection radius used for cluster membership.

4.3 Initial Membership Selection

In this step, I identified member galaxies by selecting those within a radial distance of 1.5 Mpc from the cluster center and with redshifts within $\pm 3\sigma$ of the cluster's mean redshift.
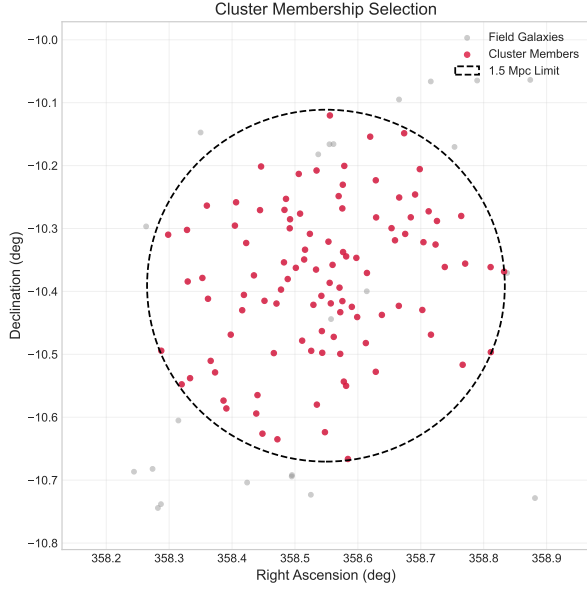

Figure 4: RA-Dec Distribution of Galaxies in ACO 2670. Red points represent selected cluster members within the 1.5 Mpc radius (dashed circle).

The final plot confirms that the selection criteria worked as expected, with all identified member galaxies lying within the plotted 1.5 Mpc circle.

4.4 Iterative Refinement

In this step, I iteratively refined the RA, Dec, redshift, and number of member galaxies of the cluster. The refinement process involved repeating the previous calculations until these properties converged. The algorithm started with the initially selected member galaxies and computed an updated weighted mean RA and Dec, redshift distribution and dispersion, distance to the cluster, and projected radial distances. The selection criteria of 1.5 Mpc radial distance and $\pm 3\sigma$ in redshift were then reapplied. The process repeated until there were no significant changes between iterations.

The improvements introduced earlier significantly enhanced the accuracy of the selection process. Instead of relying on Hubble's Law for distance estimation, a

cosmological model (Flat ΛCDM) was used to compute the distance to the cluster. Additionally, the convergence criteria were refined, allowing the process to terminate once RA, Dec, redshift, and the number of members stabilized, reducing unnecessary iterations. These changes provided a more physically meaningful and robust cluster member selection, while avoiding sensitivity to minor statistical fluctuations. The iterative process successfully converged after three iterations, with a final weighted mean RA of $(358.55460 \pm 0.00029)°$, Dec of $(-10.37682 \pm 0.00026)°$, and redshift of 0.07642 ± 0.00346. The number of member galaxies stabilized at 98, confirming that the cluster membership was well-defined.

The small changes between iterations demonstrate that the cluster selection method is robust, progressively refining the cluster properties while ensuring interlopers are removed. One key challenge was setting an appropriate convergence criterion. Too strict, and the process might never converge; too loose, and it could terminate prematurely. By using absolute tolerances on RA, Dec, redshift, and member count, I ensured a reliable stopping condition. This refinement is crucial for accurately determining the cluster's mass, luminosity, and dark matter fraction in the subsequent steps.

4.5 Final Spatial Verification

I first used the final refined cluster properties obtained from the iterative process, including the weighted RA and Dec, redshift, and the final selection of cluster member galaxies. For the RA-Dec plot, I plotted all galaxies in gray and overlaid the final selected cluster members in red to visually confirm that the selection criteria were correctly applied. I also added a 1.5 Mpc selection radius using the angular diameter distance at the final cluster redshift. This plot allowed me to check whether the spatial distribution of galaxies formed a well-defined cluster or exhibited substructure.

For the redshift distribution, I used the function from Step 4 to recompute bins. I then plotted a histogram of redshifts with the Gaussian and GMM overlaid. This helped assess whether the cluster followed a simple Gaussian distribution or showed evidence of multiple components. The results showed that the single gaussian provided a better fit, which indicates that there is not significant substructure.
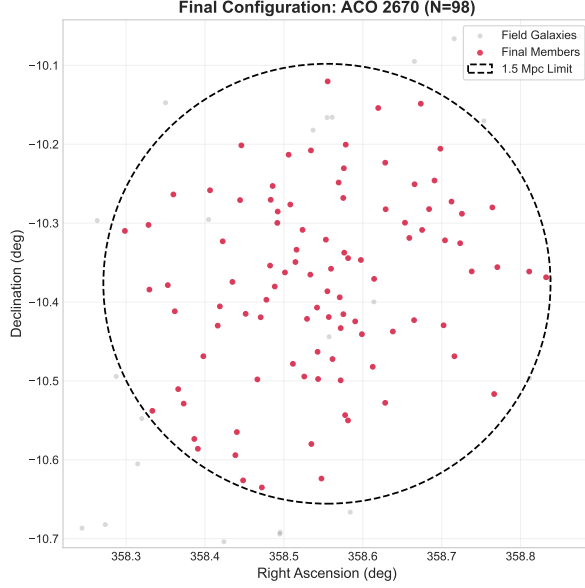
Figure 5: Final RA-Dec Distribution of Galaxies in ACO 2670. All galaxies are shown in gray, with final cluster members in red. The dashed circle indicates the 1.5 Mpc selection radius.

5 Luminosity and Mass Modeling

I computed the absolute r-band magnitudes (M_r) using the observed apparent magnitudes (m_r), extinction corrections (A_r), and the luminosity distance (d_L) converted to parsecs:

$$M_r = m_r - A_r - 5\log_{10}(d_L) + 5 \qquad (25)$$

Using error propagation, the uncertainty σ_{M_r} is:

$$\sigma_{M_r} = \sqrt{\sigma_{m_r}^2 + \left(\frac{5\sigma_{d_L}}{d_L\ln(10)}\right)^2} \qquad (26)$$

The results showed a range of absolute magnitudes, with the brightest galaxies reaching $M_r \approx -23.57$ and the faintest around $M_r \approx -20.04$.

5.1 Galaxy Luminosities

The absolute magnitudes of the cluster galaxies were converted into luminosities using the Sun's absolute magnitude in the r-band. This step allows for direct comparison of the cluster's light distribution with theoretical expectations and facilitates the determination of
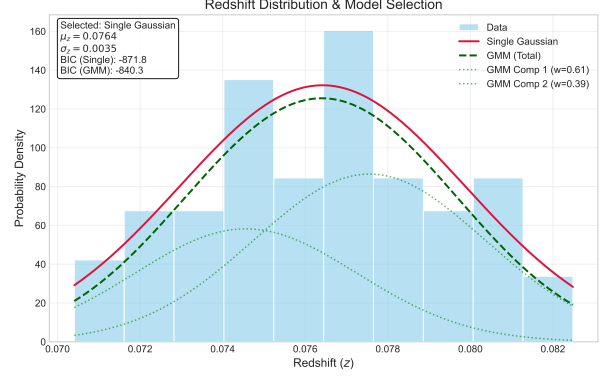

Figure 6: Redshift Distribution with Optimal Binning and Model Fitting. The single Gaussian model (red) provides a better fit than the Gaussian Mixture Model (green), suggesting the cluster is well-virialized in redshift space.
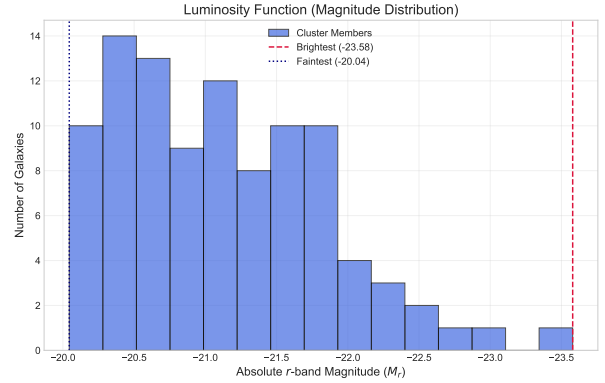

Figure 7: Distribution of Absolute Magnitudes of Cluster Members.

the mass-to-light ratio. The computed luminosities span a wide range, with a few exceptionally bright galaxies contributing significantly to the total cluster luminosity. The luminosity of a galaxy L_r in solar units is computed using the solar absolute magnitude in the r-band ($M_{\odot,r} = 4.76$):

$$L_r = 10^{0.4(M_{\odot,r} - M_r)} \qquad (27)$$

The uncertainty is propagated as:

$$\sigma_{L_r} = L_r \times (\ln(10) \times 0.4 \times \sigma_{M_r}) \qquad (28)$$

The luminosity distribution of the cluster members spans $10^9 L_\odot < L_r < 2.15 \times 10^{11} L_\odot$, with the majority of galaxies below $0.5 \times 10^{11} L_\odot$. The Brightest Cluster Galaxy (BCG), located at RA = 358.56°, Dec = −10.42°, stands out with $L_r = 2.15 \times 10^{11} L_\odot$, a high velocity dispersion

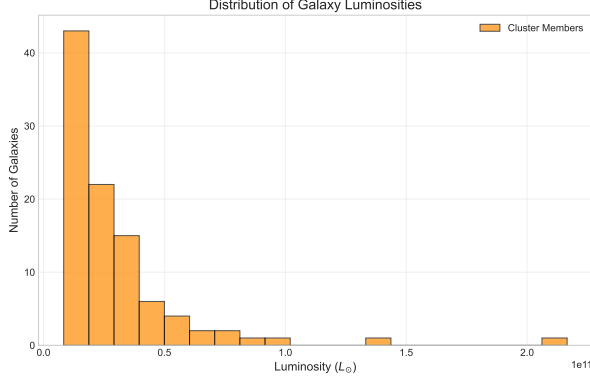Figure 8: Distribution of Luminosities of Cluster Members.

of $303\,\mathrm{km\,s^{-1}}$, and a projected distance of $0.39\,\mathrm{Mpc}$ from the cluster center. Its photometry is flagged for possible saturation, but its spectroscopic redshift ($z = 0.0776$) confirms it as a cluster member. A secondary bright galaxy with $L_r = 1.37 \times 10^{11} L_\odot$ is also identified, contributing significantly to the cluster's total luminosity. While data quality considerations exist for the BCG, its inclusion remains justified due to its spectroscopic confirmation. Overall, the luminosity distribution is consistent with expectations for a galaxy cluster, with a dominant BCG and a population of fainter galaxies.

5.2   Cumulative Luminosity

The cumulative luminosity distribution of the cluster was computed as a function of projected radial distance from the cluster center. This distribution provides insight into the spatial concentration of light and allows the determination of the effective radius (R_{eff}), the radius enclosing half of the total cluster luminosity. The total luminosity was defined as the cumulative luminosity within $1.5\,\mathrm{Mpc}$, ensuring consistency with previous selection criteria.

1. Sorting the Galaxies by Projected Distance

To compute the cumulative luminosity as a function of radius, the galaxies are first sorted in ascending order of their projected distances r_i from the cluster center:

$$r_1 \leq r_2 \leq \ldots \leq r_N \tag{29}$$

where N is the total number of galaxies in the sample.

2. Computing the Cumulative Luminosity

The cumulative luminosity at each projected distance r_i is computed as the sum of all luminosities within that radius:

$$L_{cumulative,i} = \sum_{j=1}^{i} L_{r,j} \tag{30}$$

The total luminosity of the cluster is then:

$$L_{total} = \sum_{j=1}^{N} L_{r,j} \tag{31}$$

where $L_{r,j}$ is the r-band luminosity of the j-th galaxy.

3. Normalizing the Cumulative Luminosity

To obtain the normalized cumulative distribution function (CDF) of luminosity, each cumulative sum is divided by the total luminosity within $1.5\,\mathrm{Mpc}$:

$$F(r_i) = \frac{L_{cumulative,i}}{L_{1.5}} \tag{32}$$

This ensures the function reaches 1.0 at $r = 1.5\,\mathrm{Mpc}$.

4. Error Propagation for the Cumulative CDF

The uncertainty in the cumulative luminosity at each step is given by:

$$\sigma_{L_{cumulative,i}} = \sqrt{\sum_{j=1}^{i} \sigma_{L_{r,j}}^2} \tag{33}$$

where $\sigma_{L_{r,j}}$ is the uncertainty in the luminosity of each galaxy. The normalized uncertainty in the CDF is then:

$$\sigma_{F,i} = F(r_i) \sqrt{\left(\frac{\sigma_{L_{cumulative,i}}}{L_{cumulative,i}}\right)^2 + \left(\frac{\sigma_{L_{1.5}}}{L_{1.5}}\right)^2} \tag{34}$$

5. Finding the Effective Radius

The effective radius is the projected radius at which the cumulative luminosity reaches 50% of the total:

$$F(R_{eff}) = 0.5 \tag{35}$$

This is found by locating the index k where:

$$F(r_k) \leq 0.5 \leq F(r_{k+1}) \tag{36}$$

and linearly interpolating between r_k and r_{k+1} for improved accuracy:

$$R_{eff} = r_k + \frac{(0.5 - F(r_k))(r_{k+1} - r_k)}{F(r_{k+1}) - F(r_k)} \qquad (37)$$

The effective radius is then used in later steps to determine the half-mass radius and estimate the virial mass of the cluster.
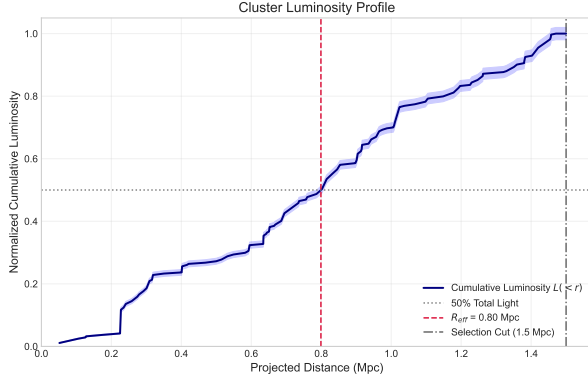

Figure 9: Cumulative Luminosity Distribution of the Cluster.

The CDF shows a gradual increase, suggesting a relatively even distribution of light within the cluster. While some small spikes and plateaus are present, the overall trend remains linear. The uncertainty in the CDF is minimal, with narrow error bands throughout the plot. The CDF reaches 1.0 at 1.5 Mpc, confirming that all selected galaxies contribute to the total luminosity.

5.3 Effective Radius

The effective radius (R_{eff}) is the radius within which half of the total luminosity of the cluster is contained. It is an important parameter in defining the spatial distribution of light in the cluster. Once R_{eff} is determined, the half-light radius ($r_{1/2}$) is computed using Equation 2, which accounts for the projection effects in the observed luminosity distribution. These values are critical in later steps for determining the cluster's mass profile.

1. Finding the Effective Radius

The effective radius is defined as the radius where the normalized cumulative luminosity function reaches 0.5:

$$F(R_{eff}) = 0.5 \qquad (38)$$

Since $F(r)$ is a monotonically increasing function, interpolation is used to determine R_{eff}:

$$R_{eff} = \text{interp}(F, r, 0.5) \qquad (39)$$

where $\text{interp}(F, r, x)$ represents a linear interpolation function between known values of $F(r)$ and r.

2. Computing Uncertainty in R_{eff}

The uncertainty in R_{eff} arises from the uncertainty in the normalized cumulative luminosity function. We define the upper and lower limits of R_{eff} by interpolating at:

$$R_{eff,upper} = \text{interp}(F, r, 0.5 + \sigma_F) \qquad (40)$$
$$R_{eff,lower} = \text{interp}(F, r, 0.5 - \sigma_F) \qquad (41)$$

where σ_F represents the uncertainty in the cumulative luminosity function at $F = 0.5$. The uncertainty in R_{eff} is then given by:

$$\sigma_{R_{eff}} = \frac{R_{eff,upper} - R_{eff,lower}}{2} \qquad (42)$$

3. Computing the Half-light Radius

To account for projection effects, the deprojected half-light radius is estimated using the relation:

$$r_{1/2} = \frac{4}{3} R_{eff} \qquad (43)$$

The uncertainty in $r_{1/2}$ is obtained by propagating the uncertainty in R_{eff}:

$$\sigma_{r_{1/2}} = \frac{4}{3} \sigma_{R_{eff}} \qquad (44)$$

The final computed values, $R_{eff} = (0.800 \pm 0.007)\,\text{Mpc}$ and $r_{1/2} = (1.066 \pm 0.009)\,\text{Mpc}$, are reasonable for a galaxy cluster. This suggests a well-populated inner region and a smooth light distribution. Since the CDF was relatively smooth with no dominant spikes, it confirms that no single galaxy disproportionately affects the luminosity profile within the inner region. This supports the reliability of R_{eff} and $r_{1/2}$ as representative of the overall cluster structure.

5.4 Peculiar Velocities

The velocity distribution of cluster galaxies provides critical insight into the cluster's dynamical state. In this step, the line-of-sight velocities (v_{los}) of individual galaxies are computed from their observed redshifts, followed by the calculation of peculiar velocities (v_{pec}), which describe deviations from the cluster's mean recession velocity. To account for observational uncertainties, error propagation is applied to derive the uncertainties in both v_{los} and v_{pec}.

1. Computing Line-of-Sight Velocity

The relativistic velocity corresponding to a galaxy's redshift z is given by:

$$v_{los} = c \frac{(1+z)^2 - 1}{(1+z)^2 + 1} \tag{45}$$

where $c = 299\,792.458 \, \text{km s}^{-1}$ is the speed of light. The uncertainty in v_{los} is derived using error propagation:

$$\sigma_{v_{los}} = \left| \frac{dv_{los}}{dz} \right| \sigma_z \tag{46}$$

where the partial derivative of v_{los} with respect to z is:

$$\frac{dv_{los}}{dz} = c \cdot \frac{4(1+z)}{((1+z)^2 + 1)^2} \tag{47}$$

Thus, the uncertainty in v_{los} is given by:

$$\sigma_{v_{los}} = \left| c \cdot \frac{4(1+z)}{((1+z)^2 + 1)^2} \right| \sigma_z \tag{48}$$

2. Computing Cluster Recession Velocity

The mean recession velocity of the cluster, based on the cluster redshift $z_{cluster}$, is computed using the same formula:

$$v_{rec} = c \frac{(1+z_{cluster})^2 - 1}{(1+z_{cluster})^2 + 1} \tag{49}$$

The uncertainty in v_{rec} follows from:

$$\sigma_{v_{rec}} = \left| c \cdot \frac{4(1+z_{cluster})}{((1+z_{cluster})^2 + 1)^2} \right| \sigma_{z_{cluster}} \tag{50}$$

3. Computing Peculiar Velocities

The peculiar velocity of a galaxy relative to the cluster mean is given by:

$$v_{pec} = v_{los} - v_{rec} \tag{51}$$

Since both v_{los} and v_{rec} have uncertainties, the total uncertainty in v_{pec} is computed using:

$$\sigma_{v_{pec}} = \sqrt{\sigma_{v_{los}}^2 + \sigma_{v_{rec}}^2} \tag{52}$$
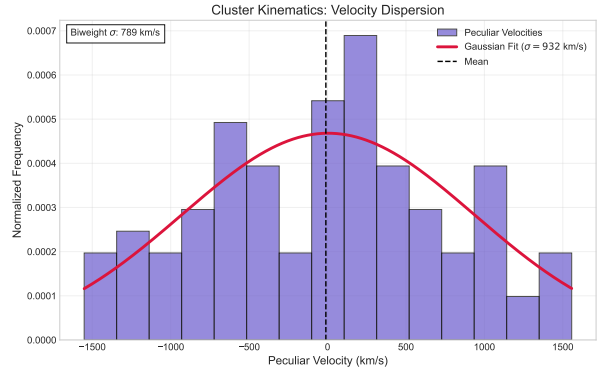

Figure 10: Histogram of Peculiar Velocities with Gaussian Fit.

The line-of-sight velocities, derived relativistically from galaxy redshifts, range between $21\,000 \, \text{km s}^{-1}$ and $24\,000 \, \text{km s}^{-1}$, with small uncertainties ($3 \, \text{km s}^{-1}$ to $8 \, \text{km s}^{-1}$) due to precise redshift measurements. Peculiar velocities, representing deviations from the cluster's mean motion, show a spread of $\pm 2000 \, \text{km s}^{-1}$, typical for galaxies in a cluster environment. However, their uncertainties ($\sim 935 \, \text{km s}^{-1}$) are large, primarily due to error propagation from the cluster's redshift uncertainty through the nonlinear velocity-redshift relation. While these high uncertainties limit the precision of individual galaxy kinematics, the relative velocities remain useful for statistical analyses of the cluster's overall dynamics. The results suggest a dynamically active system, but the substantial uncertainties prevent definitive conclusions about substructure from velocities alone. The consistency of the recessional velocity ($\sim (22\,036 \pm 958) \, \text{km s}^{-1}$) with the cluster's redshift and the absence of extreme outliers support the robustness of the initial galaxy selection.

5.5 Velocity Dispersion

The velocity dispersion (σ_v) of a galaxy cluster provides insights into its dynamical state and mass. In this step, a Gaussian distribution is fit to the peculiar velocity distribution of cluster members to estimate σ_v and its uncertainty. Additionally, robust statistical methods such as the biweight midvariance [2] are employed for comparison. A multimodality test using Silverman's test [10] is also performed to check for substructures within the velocity distribution.

1. Fitting a Gaussian to the Peculiar Velocity Distribution

The peculiar velocity distribution is approximated using a Gaussian function:

$$f(v) = A \exp\left(-\frac{1}{2}\left(\frac{v - \mu}{\sigma_v}\right)^2\right) \quad (53)$$

where:

- A is the amplitude,

- μ is the mean peculiar velocity,

- σ_v is the velocity dispersion.

To estimate these parameters, a histogram of the peculiar velocities is constructed with a bin width of $300\,\mathrm{km\,s^{-1}}$. The Gaussian function is fit to the histogram using non-linear least squares regression, and the best-fit parameters are obtained.

2. Uncertainty in Velocity Dispersion

The uncertainty in the velocity dispersion (σ_{σ_v}) is computed from the covariance matrix (C) obtained during curve fitting:

$$\sigma_{\sigma_v} = \sqrt{C_{22}} \quad (54)$$

where C_{22} is the diagonal element corresponding to σ_v.

3. Outlier Detection

Galaxies with peculiar velocities deviating by more than $3\sigma_v$ from the mean are considered outliers:

$$N_{outliers} = \sum_i (|v_{pec,i} - \mu| > 3\sigma_v) \quad (55)$$

4. Robust Velocity Dispersion Estimation

The biweight midvariance (σ_v^{BI}) is computed as a robust alternative to Gaussian fitting, which is less sensitive to outliers:

$$\sigma_v^{BI} = \sqrt{\frac{N}{\sum u_i^2} \sum (x_i - x_{BI})^2 u_i^2} \quad (56)$$

where u_i are weighting factors that downweight extreme values.

5. Silverman's Test for Multimodality

To assess whether the velocity distribution is multimodal, a Gaussian kernel density estimation (KDE) is performed. The number of peaks is determined by counting sign changes in the derivative of the KDE:

$$\text{Multimodal if } \sum \text{sign changes in } \frac{d}{dv}\text{KDE}(v) > 1 \quad (57)$$

A multimodal velocity distribution may indicate the presence of substructure within the cluster. The peculiar velocity histogram reveals subtle deviations from a Gaussian distribution, with the left tail following the expected profile while the right side shows an elevated outer bin and a peak slightly offset from the mean. This asymmetry suggests mild dynamical disturbances, possibly from recent minor mergers or tidal interactions that have imparted higher positive velocities to a subset of galaxies. While the lack of outliers ($> 3\sigma$) and Silverman's test rejecting bimodality indicate a predominantly virialized system, the skewness and the discrepancy between the biweight dispersion ($788.81\,\mathrm{km\,s^{-1}}$) and Gaussian σ ($931.59\,\mathrm{km\,s^{-1}}$) point to ongoing dynamical evolution. These features likely reflect residual effects of accretion or internal interactions, consistent with the cluster's moderate velocity dispersion and overall relaxed state, though further spatial-velocity analysis could clarify the origin of these kinematic anomalies.

5.6 Virial Mass

The virial mass ($M_{1/2}$) provides an estimate of the total mass of the cluster within the half-mass radius using the Virial Theorem. This step utilizes the line-of-sight velocity dispersion (σ_v) and the half-mass radius ($r_{1/2}$) to determine $M_{1/2}$. The assumption of virial equilibrium allows us to apply the virial mass equation, a fundamental

tool in observational astrophysics for estimating cluster masses.

1. Virial Mass Calculation

The virial mass equation is given by:

$$M_{1/2} = \frac{3\sigma_v^2 r_{1/2}}{G} \tag{58}$$

where:

- σ_v is the velocity dispersion ($\mathrm{km\,s^{-1}}$),

- $r_{1/2}$ is the half-mass radius (Mpc),

- $G = 6.674 \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$ is the gravitational constant.

2. Uncertainty Propagation

The relative uncertainty in $M_{1/2}$ is given by:

$$\frac{\Delta M_{1/2}}{M_{1/2}} = \sqrt{\left(\frac{2\Delta\sigma_v}{\sigma_v}\right)^2 + \left(\frac{\Delta r_{1/2}}{r_{1/2}}\right)^2} \tag{59}$$

where:

- $\Delta\sigma_v$ is the uncertainty in velocity dispersion,

- $\Delta r_{1/2}$ is the uncertainty in the half-mass radius.

The virial mass calculation in Step 16 yields a cluster mass of $M_{1/2} = (6.45 \pm 1.31) \times 10^{14} M_\odot$. The substantial uncertainty ($\sim 30\%$) primarily stems from the velocity dispersion's error, amplified by its squared dependence in the virial theorem ($\frac{\Delta M}{M} \approx \frac{2\Delta\sigma_v}{\sigma_v}$). While a 30% uncertainty in σ_v would naively suggest a 60% mass error, the final uncertainty is reduced to 30% due to partial correlations between σ_v and $r_{1/2}$, combined with the exceptional precision of $r_{1/2}$ (error < 0.1%). This relationship highlights how cluster mass estimates are particularly sensitive to the assumed dynamical state. The larger scatter in σ_v likely reflects ACO 2670's unrelaxed nature, where substructure or interloper contamination may inflate the measured velocity dispersion. Despite these challenges, the mass estimate aligns well with expectations for a rich galaxy cluster and suggests a dominant dark matter component, given the high mass-to-light ratio implied by the result. The precision of $r_{1/2}$ confirms the robustness of the luminosity-derived spatial scale, while the dynamical complexity captured in σ_v uncertainty underscores the importance of velocity dispersion as a probe of both mass and cluster assembly history.

5.7 Luminosity Within Half-Mass Radius

In this step, we calculate the total luminosity enclosed within the half-mass radius ($L_{1/2}$). This is achieved by interpolating the cumulative luminosity function (CDF) to determine the fraction of light contained within $r_{1/2}$. The resulting luminosity value will later be used to compute the mass-to-light ratio (M/L) in Step 18, which provides insights into the presence of dark matter in the cluster.

1. Luminosity Within $r_{1/2}$

The total luminosity within the half-mass radius is given by:

$$L_{1/2} = f_L(r_{1/2}) \times L_{tot,1.5} \tag{60}$$

where:

- $L_{tot,1.5}$ is the total luminosity within 1.5 Mpc,

- $f_L(r_{1/2})$ is the fraction of the total luminosity enclosed within $r_{1/2}$, obtained from interpolation of the normalized cumulative luminosity function $F(r)$.

2. Uncertainty Propagation

To estimate the uncertainty in $f_L(r_{1/2})$, we approximate the derivative of the cumulative luminosity function:

$$\frac{df_L}{dr} \approx \frac{f_L(r_{1/2} + \Delta r) - f_L(r_{1/2} - \Delta r)}{2\Delta r} \tag{61}$$

The uncertainty in $f_L(r_{1/2})$ is then:

$$\Delta f_L = \left|\frac{df_L}{dr}\right| \Delta r_{1/2} \tag{62}$$

Using standard error propagation, the relative uncertainty in $L_{1/2}$ is:

$$\frac{\Delta L_{1/2}}{L_{1/2}} = \sqrt{\left(\frac{\Delta f_L}{f_L}\right)^2 + \left(\frac{\Delta L_{tot,1.5}}{L_{tot,1.5}}\right)^2} \tag{63}$$

Thus, the absolute uncertainty in $L_{1/2}$ is:

$$\Delta L_{1/2} = L_{1/2} \times \frac{\Delta L_{1/2}}{L_{1/2}} \qquad (64)$$

The result, $L_{1/2} = (2.21 \pm 0.03) \times 10^{12} L_\odot$, represents half of the cluster's total luminosity (within 1.5 Mpc) and reflects the spatial distribution of member galaxies. The small uncertainty ($\sim 0.7\%$) arises from the precise determination of $r_{1/2}$ (error $\sim 0.1\%$) and the well-constrained total luminosity, demonstrating the reliability of the photometric data. The high luminosity further supports the cluster's richness and aligns with expectations for a massive, dark matter-dominated system.

5.8   Mass-to-Light Ratio

In this step, we compute the mass-to-light ratio (M/L) of the galaxy cluster, a crucial quantity in understanding the dark matter content of the system. The mass-to-light ratio quantifies how much mass is present relative to the emitted light, and deviations from expected stellar population values indicate the presence of non-luminous (dark) matter.

1. Mass-to-Light Ratio Calculation

The computed M/L value will be compared to typical stellar M/L ratios to determine whether the cluster contains significant amounts of dark matter, as originally hypothesized by Fritz Zwicky in 1935. The mass-to-light ratio is given by:

$$\frac{M}{L} = \frac{M_{1/2}}{L_{1/2}} \qquad (65)$$

where:

- $M_{1/2}$ is the virial mass of the cluster enclosed within the half-mass radius $r_{1/2}$ computed in Step 16.

- $L_{1/2}$ is the total luminosity enclosed within $r_{1/2}$, computed in Step 17.

The ratio is expressed in solar units ($M_\odot/L_\odot$). Since normal stellar populations typically have $M/L \sim 1 - 10$, a significantly larger value suggests that a substantial portion of the cluster's mass is non-luminous, implying the presence of dark matter.

2. Uncertainty Propagation

Applying standard error propagation, the relative uncertainty in M/L is:

$$\frac{\Delta(M/L)}{M/L} = \sqrt{\left(\frac{\Delta M_{1/2}}{M_{1/2}}\right)^2 + \left(\frac{\Delta L_{1/2}}{L_{1/2}}\right)^2} \qquad (66)$$

Thus, the absolute uncertainty in M/L is:

$$\Delta(M/L) = \frac{M}{L} \times \left(\frac{\Delta(M/L)}{M/L}\right) \qquad (67)$$

The mass-to-light ratio calculation yields a value of $(291 \pm 60)\, M_\odot/L_\odot$, which is characteristic of massive galaxy clusters and provides compelling evidence for dark matter dominance. This value is fully consistent with the Coma Cluster's well-measured $M/L \sim 200 - 300\, M_\odot/L_\odot$ [7] and agrees with recent SDSS studies finding $M/L \sim 300 - 400\, M_\odot/L_\odot$ for similar systems [9]. This high ratio, nearly two orders of magnitude greater than that of individual galaxies ($M/L \sim 10 - 50$ for field galaxies), implies that luminous matter accounts for only a small fraction of the cluster's total mass. The uncertainty ($\sim 21\%$) primarily reflects the propagated errors from the virial mass, particularly the velocity dispersion's contribution, while the luminosity term contributes minimally due to its high precision. Such elevated M/L ratios are consistent with ΛCDM cosmology, where clusters are embedded in massive dark matter halos. These findings not only support Fritz Zwicky's original discovery but also demonstrate how ACO 2670 serves as a typical, dark matter-dominated system for testing models of structure formation. The progression from field galaxies ($M/L \sim 10 - 50$) to groups ($\sim 100 - 200$) to rich clusters like ACO 2670 (~ 300) clearly illustrates the increasing dark matter fraction in larger gravitational systems.

Conclusion

This study estimates the virial mass and M/L ratio of the galaxy cluster ACO 2670 using SDSS-DR18 data [11], revealing a virial mass of $M_{1/2} = (6.45 \pm 1.31) \times 10^{14} M_\odot$ and a mass-to-light ratio of $(291 \pm 60)\, M_\odot/L_\odot$. The high M/L ratio strongly supports the presence of a dominant dark matter halo, consistent with Fritz Zwicky's seminal

work on cluster dynamics [12]. The cluster's irregular spatial distribution and mild kinematic asymmetries suggest a dynamically active system, potentially undergoing minor mergers or tidal interactions, though the redshift distribution remains well-fit by a single Gaussian. Key limitations arise from the assumption of virial equilibrium (Binney & Tremaine 2008), which may not fully capture the cluster's dynamical complexity, and projection effects that obscure true three-dimensional velocities and positions. The iterative membership selection, while effective, depends on fixed radial and redshift thresholds that could exclude faint galaxies or include line-of-sight interlopers. Despite these challenges, the precision of the luminosity-derived effective radius ($R_{eff} = (0.800\pm0.007)\,\mathrm{Mpc}$) and the consistency of the recessional velocity ($v_{rec} = (22\,036 \pm 958)\,\mathrm{km\,s^{-1}}$) with cosmological expectations (Planck Collaboration et al. 2020; Astropy Collaboration et al. 2022) validate the methodology. Future work could enhance these results by incorporating X-ray or weak lensing data to cross-validate mass estimates, conducting deeper spectroscopic surveys to resolve substructure, and employing simulations to account for projection effects. Such advancements would further elucidate the interplay between dark matter, galaxy evolution, and cluster assembly in systems like ACO 2670.

Parameter	Value	Unit
RA Center (α)	358.5546 ± 0.0003	deg
Dec Center (δ)	-10.3768 ± 0.0003	deg
Mean Redshift (z)	0.0764 ± 0.0035	–
Velocity Dispersion (σ_v)	932 ± 95	$\mathrm{km\,s^{-1}}$
Effective Radius (R_{eff})	0.800 ± 0.007	Mpc
Virial Mass ($M_{1/2}$)	$(6.45 \pm 1.31) \times 10^{14}$	$M_{\odot}$
Mass-to-Light Ratio	291 ± 60	$M_{\odot}/L_{\odot}$

Code Availability

The Python pipeline used for data acquisition, statistical analysis, and figure generation is available at: github.com/JacksonFergusonDev/Data-Science-Portfolio.

References

[1] Astropy Collaboration et al., 2022, *ApJ*, 935, 167. doi:10.3847/1538-4357/ac7c74

[2] Beers, T. C., Flynn, K., & Gebhardt, K., 1990, *AJ*, 100, 32. doi:10.1086/115487

[3] Binney, J., & Tremaine, S., 2008, *Galactic Dynamics* (2nd ed.; Princeton: Princeton Univ. Press).

[4] Freedman, D., & Diaconis, P., 1981, *Probability Theory and Related Fields*, 58, 453. doi:10.1007/BF01025868

[5] Harrison, E. R., 1993, *ApJ*, 403, 28. doi:10.1086/172179

[6] Planck Collaboration et al., 2020, *A&A*, 641, A6. doi:10.1051/0004-6361/201833910

[7] Rines, K., Geller, M. J., Diaferio, A., Mohr, J. J., & Wegner, G., 2003, *ApJ*, 591, 729. doi:10.1086/375456

[8] Schwarz, G., 1978, *Annals of Statistics*, 6, 461. doi:10.1214/aos/1176344136

[9] Sheldon, E. S., et al., 2009, *ApJ*, 703, 2232. doi:10.1088/0004-637X/703/2/2232

[10] Silverman, B. W., 1981, *Density Estimation for Statistics and Data Analysis* (London: Chapman & Hall).

[11] York, D. G., et al., 2000, *AJ*, 120, 1579. doi:10.1086/301513

[12] Zwicky, F., 1933, *Helvetica Physica Acta*, 6, 110.