**Cosmology Final Project** - *Measauring Shear using Precision Weak Lensing*
Spencer Scott

## 1 Introduction and Motivation

When probing the mass density of the universe, galaxies are commonly used as tracer particles for the full mass distribution. However, this only accounts for luminous matter, while the majority of the matter density in the universe is contained in dark matter. For years, weak gravitational lensing has enabled us to probe the full underlying mass distribution due to its sensitivity to all mass distributions, luminous and dark. However, as discussed in class, traditional weak lensing is susceptible to systematics such as shape noise, which requires a large number of observed galaxies to minimize. Blain (2002) first proposed using the velocity fields of individual weakly lensed galaxies to measure shear. While this idea has been around for many years, the first "precision weak lensing" measurements were not made until the work of Gurri et al. (2020).

For my final project, I set out to reproduce the results in Gurri et al. (2020). The rest of this report is organized as follows:

- In Section 2, I briefly discuss the data used in the project.
- In Section 3, I explain the model used to model the effects of weak lensing as well as the MCMC parameters used to fit the data.
- In Section 4, I present the results of my Markov-Chain Monte Carlo (MCMC) analysis.
- Finally, in Section 5, I discuss possible ways to improve my analysis and potential uses for this model.

## 2 Data

The data used for this project were originally taken for Gurri et al. (2020). While more details on the data reduction can be found therein, I will briefly describe the data below.

The data were obtained during 6 observing runs in 2018-2019 using the WiFeS spectrograph (Dopita et al., 2007, 2010) at the Australian National University 2.3m telescope. WiFeS is an optical IFU with a field of view of 38"x25" and spatial resolution 0.5"x1". Galaxies were typically measured 6 times each with 20 minutes exposure, for 120 minutes of average exposure time per galaxy. To measure velocities, Gurri et al. (2020) fit Gaussian profiles to each absorption and emission line presented in Morton (1991). The published velocities are maximum likelihood fits, with uncertainties calculated using the Fisher information matrix. Gurri et al. (2020) only consider pixels for which the uncertainty in velocity is $<$50km/s, a convention I also follow in this report. The data are publicly available at `http://dx.doi.org/10.26185/5f488683e4867`.

## 3 Methods

In this section, I will discuss the ingredients that go into the model of the data that allow me to attempt to reproduce observations. First, I will discuss the equations that go into the model, and then I will briefly discuss the parameters of the MCMC fits.

### 3.1 Lensing Model

Before discussing the lensing model, we must first establish a coordinate system. The coordinate system used in this work, which follows the convention from Miralda-Escude (1991) is illustrated in Figure 1. $\xi$ and $\eta$ are coordinates within the detector; $x$ and $y$ are coodinates centered on the imaged galaxy; $u$ and $v$ are the tangential and radial lensing directions with respect to the lens; $\theta$ is the angle the galaxy makes with respect to detector coordinates; and $\delta$ is the angle between the lensing axis and the detector coordinates.

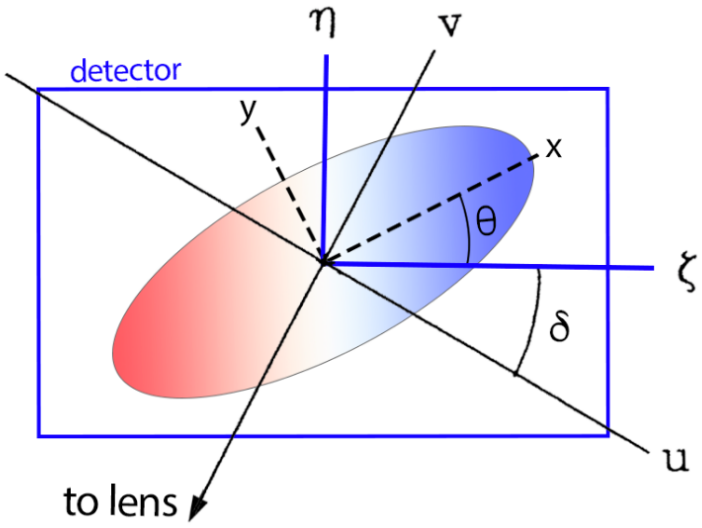


Figure 1: A visual representation of the coordinate system used in this work, adapted from Gurri et al. (2020). The definition of this coordinate system follows Miralda-Escude (1991).

In order to measure shear with precision weak lensing, we must be able to transform from the lensed image in the detector coordinates ($\vec{d}$:=$(\xi, \eta)$) to an unlensed image in a galactocentric coordinate system ($\vec{g}$:=(x,y)). Without lensing, this transformation depends on the inclination of the galaxy with respect to the observer, $i$, and $\theta$ defined above. Mathematically, this is stated as:

$$\vec{d} = \mathscr{R}\mathscr{I}\vec{g} \tag{1}$$

with $\mathscr{R}$ and $\mathscr{I}$ the rotation and inclination transformations, defined as follows:

$$\mathscr{R} := \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \mathscr{I} := \begin{bmatrix} 1 & 0 \\ 0 & \sin(i) \end{bmatrix} \tag{2}$$

However, we measure lensed galaxies in detector coordinates, so we must take into account the effects of lensing and then invert this transformation. Following Gurri et al. (2020) we can write the lensing matrix $\mathscr{A}$ in the detector coordinate system:

$$\mathscr{A} := \begin{bmatrix} 1 - \kappa - \gamma\cos(\theta) & -\gamma\sin(\theta) \\ -\gamma\sin(\theta) & 1 - \kappa + \gamma\cos(\theta) \end{bmatrix} \tag{3}$$

The full coordinate transformation is then

$$\vec{g} = \mathscr{I}^{-1}\mathscr{R}^{-1}\mathscr{A}\vec{d} \tag{4}$$

Once in the unlensed galactocentric coordinates, we can model the velocity field following Courteau (1997) and Green et al. (2014). The main assumption of this model is that galaxy rotation curves are axially symmetric and asymptotically flat. Under these assumptions, the galaxy rotation curve is characterized by a maximum rotation velocity $V_{max}$ and a scale radius $r_t$:

$$V(x,y) = \frac{2V_{max}}{\pi}\arctan\left(\frac{|R|}{r_t}\right)\sin(\omega) \tag{5}$$

where $R$ and $\omega$ are the equivalent polar coordinates to the galactocentric coordinates $\vec{g}$, defined as $R^2 = x^2 + y^2$ and $\omega = \arctan(y/x)$.

In order to take into account the systemic velocity of the galaxy with respect to the observer when fitting to data, we add an extra $v_0$ term to $V(x,y)$. The full equation for the velocity is then

$$V(x,y) = \frac{2V_{max}}{\pi}\arctan\left(\frac{|R|}{r_t}\right)\sin(\omega) + v_0 \tag{6}$$

### 3.2 MCMC

With a model for the velocity field of a galaxy and transormations from detector coordinates, we can now compare the observed velocity field $V_{\xi\eta}$ to the modeled velocity field $V(\xi,\eta)$. Using the likelihood function defined in Equation 7, we can perform an MCMC sampling of parameter space.

$$\ln(\mathscr{L}) = -\frac{1}{2}\sum_{\xi\eta}\left(\frac{V_{\xi\eta} - V(\xi,\eta)}{\sigma_{\xi\eta}}\right)^2 \tag{7}$$

For my MCMC sampling, I utilize EMCEE Python package (Foreman-Mackey et al., 2019). Each chain has 300 walkers which take 7000 steps, of which the first 5000 are discarded for a "burn-in" period. The chains probe 8 dimensions of parameter space, with each parameter described in Table 1. The initial positions of the walkers are randomly distributed in parameter space, with random distributions centered as follows:

- Spatial coordinates are centered on the middle of the detector.
- $v_0$ is centered on the mean velocity in the data.
- $V_{max}$ is centered on the maximum of the velocity in the data minus $v_0$.
- $\theta$ and $i$ are centered on $\frac{\pi}{4}$ radians.
- $\gamma$ is centered on 0
- $r_t$ is centered on 2.5.

For their fits, Gurri et al. (2020) follow the prescription in Hogg et al. (2010) to account for outliers in their data. That is beyond the scope of this project, so I neglect to do so here.

| Variable | Definition |
|---|---|
| $\xi_0$ | Center of galaxy in horizontal coordinates in the detector |
| $\eta_0$ | Center of galaxy in vertical coordinates in the detector |
| $\theta$ | Angle galaxy makes with respect to detector coordinates |
| $i$ | Galaxy inclination |
| $\gamma$ | Cosmic shear |
| $V_{max}$ | Asymptotic velocity of galaxy rotation |
| $v_0$ | Systemic velocity of galaxy with respect to observer |
| $r_t$ | Velocity scale radius |

Table 1: List of parameters used in MCMC fit and brief definitions of each.

## 4 Results

I performed an MCMC sampling of each of the 21 galaxies made available by Gurri et al. (2020). I present the medians of the posterior distributions of the samples as the "best-fit" results in Table 2, as well the values published in by Gurri et al. (2020) for comparison. The values measured for this project are denoted with a subscript $p$, while the literature values are denoted with a subscript $l$. While there are some notable successes in fitting some of the parameters for some of the galaxies, I was not able to successfully reproduce all of the published parameters for any galaxy.

Figure 2 shows the posterior distributions from an example galaxy fit, which is labeled as "ID6" in Table 2. Figure 3 shows the best fit values from this posterior as compared to the data, as well as residuals. It is interesting to note that despite a well-behaved posterior, low residuals, and passing the eye test, my best-fit values, even for this galaxy, are very different from the Gurri et al. (2020) published values. Also of note, my measured values for $i$ are often $\approx$1.57 radians, or $90^{\circ}$. There are a number of possible causes for these discrepancies, which I will discuss in Section 5.

| Object ID | Object Designation | $\xi_{0,l}$ | $\xi_{0,p}$ | $\eta_{0,l}$ | $\eta_{0,p}$ | $\theta_l$ | $\theta_p$ | $i_l$ | $i_p$ | $\gamma_l$ | $\gamma_p$ | $V_{max,l}$ | $V_{max,p}$ | $v_{0,l}$ | $v_{0,p}$ | $r_{t,l}$ | $r_{t,p}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ID1 | GGL100035+032254 | 11.01 | 20.27 | 22.18 | 15.08 | 0.08 | 1.78 | 0.19 | 1.68 | 0.02885 | -0.27 | 263.75 | 346.99 | 31168.86 | 31242.88 | 1.12 | 0.01 |
| ID2 | GGL130200-065613 | 9.73 | 10.22 | 22.61 | 22.52 | -0.07 | 0.06 | 0.30 | 1.57 | -0.0083 | -0.12 | 360.93 | 224.61 | 21214.76 | 21220.50 | 3.00 | 2.45 |
| ID3 | GL221230-241804 | 11.09 | 3.004 | 21.67 | 15.08 | 2.96 | 0.72 | 0.60 | 1.56 | -0.0044 | -0.18 | 353.64 | 623.86 | 31896.24 | 31509.012 | 1.82 | 0.01 |
| ID4 | GGL030700-071242 | 8.85 | 8.60 | 21.45 | 21.31 | 2.02 | 4.26 | -0.04 | 1.56 | -0.0301 | 0.28 | 282.89 | 123.19 | 28267.29 | 36815.55 | 2.31 | 0.33 |
| ID5 | GGL110920-000655 | 7.61 | 0.14 | 20.98 | 21.58 | -0.15 | 0.80 | 0.0017 | 1.67 | 0.02478 | -0.74 | 69.89 | 65.20 | 21485.88 | 21506.44 | 0.55 | 0.40 |
| ID6 | GGL100049+044553 | 11.13 | 11.14 | 21.92 | 22.90 | 2.76 | 3.52 | 0.32 | 1.57 | -0.0028 | -0.38 | 469.20 | 283.58 | 37259.62 | 37229.56 | 3.90 | 1.24 |
| ID7 | GGL140313+140843 | 7.86 | 7.41 | 22.46 | 22.99 | 2.47 | 3.95 | 0.24 | 1.56 | 0.07481 | 0.17 | 378.37 | 193.84 | 30305.15 | 30280.17 | 1.92 | 0.01 |
| ID8 | GGL231916-223942 | 12.32 | 24.97 | 22.69 | 24.83 | -3.38 | 4.45 | 0.05 | 1.57 | 0.04140 | -0.39 | 302.23 | 847.17 | 31657.99 | 32163.99 | 2.61 | 2.72 |
| ID9 | GGL103439+110321 | 8.05 | 7.99 | 22.35 | 22.96 | -0.94 | 1.02 | 0.13 | 1.57 | -0.0199 | 0.29 | 239.22 | 120.51 | 34154.04 | 34167.91 | 0.90 | 0.00 |
| ID10 | GGL020928-100653 | 5.64 | 5.58 | 22.93 | 22.77 | -1.76 | 1.83 | -0.24 | 1.57 | 0.02047 | -0.19 | 140.00 | 74.52 | 25777.36 | 25780.52 | 0.78 | 0.01 |
| ID11 | GGL003402-094532 | 11.50 | 11.48 | 22.52 | 0.26 | -0.03 | 0.03 | 0.21 | 1.57 | -0.0284 | -0.93 | 282.89 | 413.88 | 28267.29 | 28017.79 | 1.36 | 0.01 |
| ID12 | GGL003402-094532 | 11.50 | 11.50 | 22.52 | 0.25 | -0.04 | 0.03 | 0.21 | 1.43 | 0.04619 | -0.94 | 282.89 | 413.25 | 28267.29 | 28018.40 | 2.31 | 0.01 |
| ID13 | GGL003400-094434 | 11.83 | 12.37 | 20.85 | 20.02 | 2.95 | 3.33 | 0.31 | 1.57 | -0.0248 | -0.66 | 358.81 | 328.68 | 28485.95 | 28523.42 | 2.29 | 1.38 |
| ID14 | GGL223150+002627 | 12.68 | 12.53 | 22.77 | 22.95 | -3.32 | 3.27 | 0.31 | 1.57 | 0.01907 | 0.34 | 394.73 | 222.63 | 39551.96 | 39541.43 | 1.46 | 0.39 |
| ID15 | GGL103257+120139 | 9.91 | 10.15 | 19.90 | 20.00 | -1.42 | 1.36 | 0.26 | 0.87 | 0.02141 | 0.18 | 363.83 | 261.08 | 42950.12 | 42944.03 | 2.55 | 1.419 |
| ID16 | GGL214705-012125 | 10.95 | 16.19 | 21.83 | 23.75 | -3.10 | 3.41 | 0.46 | 1.57 | 0.02581 | -0.16 | 530.02 | 330.17 | 39510.89 | 39518.76 | 4.58 | 2.79 |
| ID17 | GGL134940+110621 | 11.33 | 17.31 | 22.34 | 0.60 | 1.83 | 0.02 | 0.27 | 1.51 | 0.02750 | 0.81 | 321.32 | 267.08 | 44616.56 | 44349.52 | 1.44 | 0.35 |
| ID18 | GGL120439+014609 | 8.95 | 9.05 | 20.52 | 20.42 | -2.67 | 2.76 | -0.10 | 0.40 | 0.06458 | -0.67 | 186.18 | 114.07 | 23339.5 | 23338.91 | 2.04 | 1.21 |
| ID19 | GGL120439+014609 | 10.31 | 10.23 | 22.17 | 22.25 | -2.77 | 2.817 | 0.17 | 2.96 | 0.01443 | -0.903 | 190.88 | 128.07 | 23412.55 | 23410.48 | 2.13 | 1.31 |
| ID20 | GGL235410+002339 | 8.50 | 8.04 | 23.2 | 23.53 | 2.51 | 3.83 | -0.03 | 1.57 | 0.05806 | 0.25 | 200.97 | 89.28 | 18446.20 | 18423.59 | 1.31 | 0.00 |
| ID21 | GGL235410+002339 | 10.13 | 13.05 | 21.35 | 26.55 | 2.51 | 3.11 | 0.11 | 1.537 | 0.06669 | -0.55 | 188.50 | 348.02 | 18445.17 | 18233.71 | 1.87 | 1.295 |

Table 2: The results of my MCMC fitting, compared to the values made publicly available in Gurri et al. (2020). I present here the median values of the posterior distributions as "best-fit" values. I denote the literature values of parameters with a subscript $l$, while I denote the parameters fit for this project with a subscript $p$. All angles are in radians, and all velocities are in km/s. For uncertainties on best-fit values, I refer the reader to Gurri et al. (2020) for the literature uncertainties, and the figures available on GitHub for uncertainties on the values measured for this work. Of note, due to tiling of data, some figures may have values of $\eta_0$ that are twice what are quoted here. For discussion on why this might be, see Section 5.

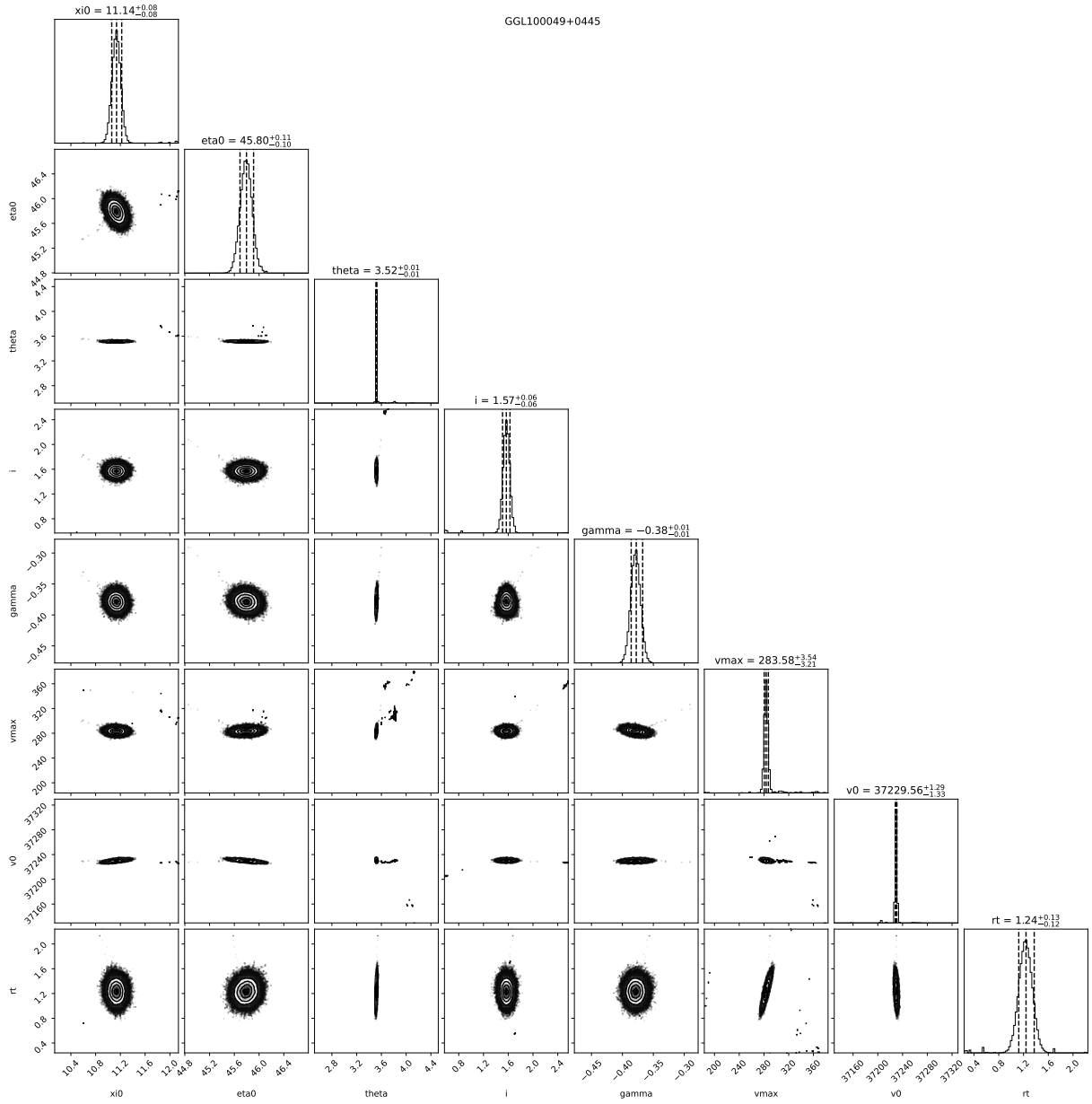


Figure 2: A sample of one of the posterior distributions from the MCMC fit. I could not figure out how to get the labels to adjust size in the `corner` module (Foreman-Mackey, 2016), so the parameters are, from left to right: $\xi_0$, $\eta_0$, $\theta$, $i$, $\gamma$, $V_{max}$, $v_0$, and $r_t$. Median values are quoted, in the titles, as well as the $16^{th}$ and $84^{th}$ percentile ranges. Best fit values, in this case, have converged well and established posteriors, but do not match values in Gurri et al. (2020). I have uploaded a gzipped tarball of the full size version of this figure, as well as corner plots and burn-in plots for every observed galaxy in a separate file, as the file was too large to upload to GitHub.

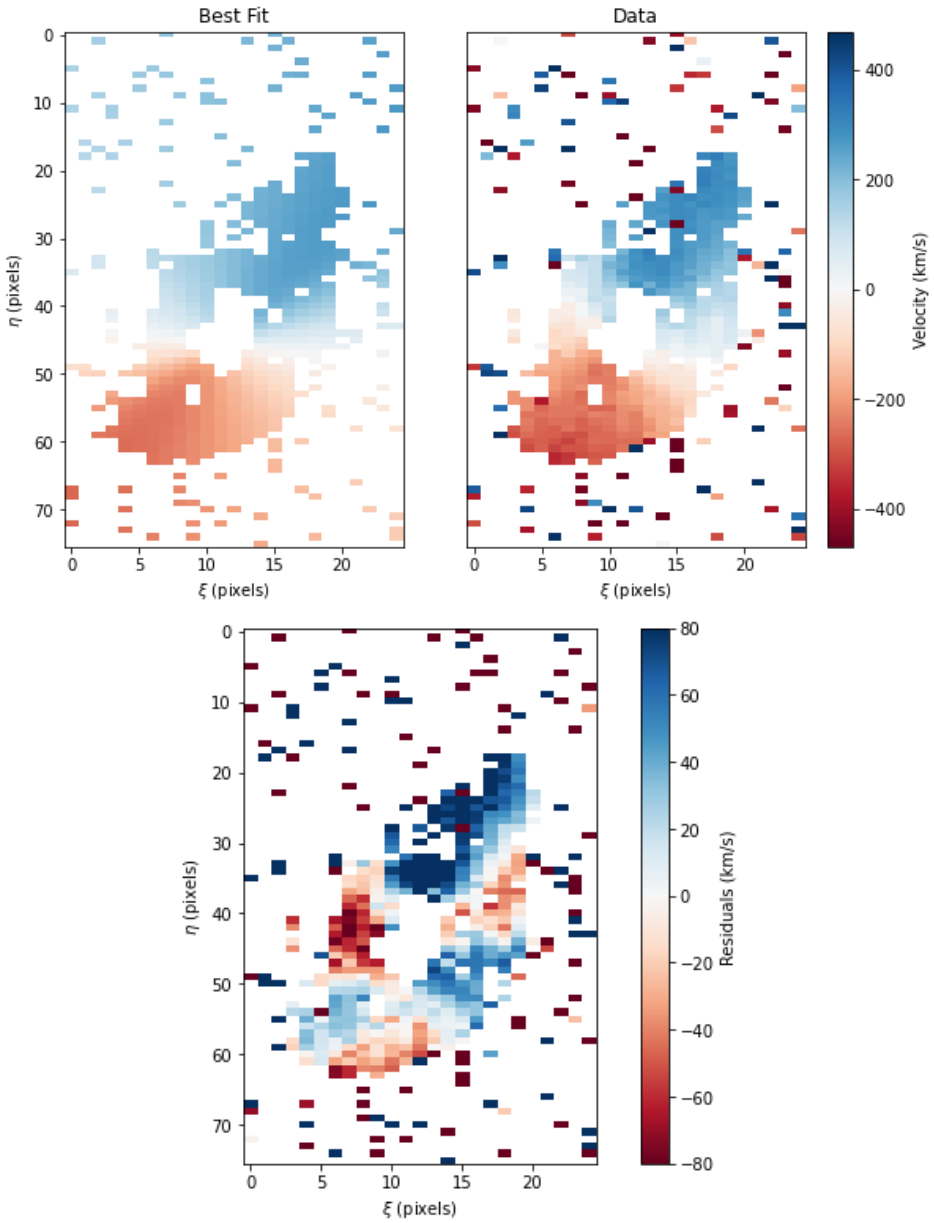


Figure 3: Best fit model (top left), data (top right), and residuals (bottom middle) for the fit shown in Figure 2. Despite the disagreement between the Gurri et al. (2020) measured values and my best fit values, the best fit model appears to match observations very well. There could be numerous reasons for this, as I discuss in Section 5. It seems likely that I have some error in my modeling code which leads to best fit values that match the data, but are not physically correct.

## 5 Discussion and Future Directions

In this report, I have outlined the basic model needed for precision weak lensing measurements and attempted to use this model to perform an MCMC sampling of parameter space to reproduce the results of Gurri et al. (2020). While my results, summarized in Table 2, do not match published values, the parameters from these fits seem to match the data reasonably well for at least a subset of observations. This is demonstrated in Figure 3.

There are a number of possible reasons for this discrepancy, including but not limited to:

- bugs in code
- improper modeling
- differing constraints on parameter space
- not modeling outliers
- inconsistencies in published values and models

In the following, I will very briefly discuss each of these potential issues, and possible solutions.

As in any code, bugs may exist. While I have extensively checked for bugs in my modeling code, it is impossible to cover every case, especially given limited time. The solution for this is obvious: run more tests and test more edge cases. Another possible problem is that I did not properly implement the model as it was stated in the paper. One indication that this may be the case is the preference for my fits to $i$ to be centered on $\pi/2$. Even given the best fit values for the model from the literature, I struggled to make models match observations. Like testing for bugs, this requires going back and testing more cases, which I did not have time for in the course of completing this project.

Also, while creating Table 2, I realized that some of the constraints I placed on the MCMC walkers differed from those used by Gurri et al. (2020). Specifically, I limited $\theta$ and $i$ to the interval [0,2$\pi$], which Gurri et al. (2020) clearly did not, as indicated by the existence of negative values. This should not be a considerable issue, however, as negative angles are easily translated to their positive counterparts, fits to coordinates and parameters like $v_0$ are reasonably close to published values, and for $\gamma$ (the measurement of which is the goal of this work) I placed no constraints.

Finally, while performing my analysis and comparing to results, I discovered some inconsistencies in published values, models, and data. For example, the .fits file for the galaxy labeled "ID4" states a best fit $\gamma = -0.02816$, while Table 3 in Gurri et al. (2020) states a best fit $\gamma = -0.0301$. Also, the paragraph before Equation 6 in Gurri et al. (2020) states that the velocity model is an arctanh function, rather than an arctan function, as it is elsewhere in literature. While it is likely that these issues are simply typos, there were other inconsistencies in the public data files that make these potential issues worth double checking. The authors of the study have been contacted about these points, but I have not heard back yet. I plan to explore independently whether these could be the cause of disagreement between our results, but that was beyond the scope of this project and the time allotted to it.

The original work included nuisance parameters meant to account for outliers in the data in their MCMC sampling. They also suggested a simpler sigma-clipping-like method for accounting for outliers. I incorporated neither into my work for this project. Because the effect of weak lensing is very small, the effect of outliers on measurement may make matching published values impossible if not properly accounted for.

Despite these issues, precision weak lensing remains a promising tool to measure the matter distribution in the universe. Precision weak lensing is not as susceptible to the shape noise that traditional weak lensing is, which enables shear measurements for individual objects rather than ensembles. Because this method of utilizing precision weak lensing methods requires data from an IFU currently, issues with photometric redshifts are also avoided.

Finally, I would like to thank the corresponding author of the original work, Pol Gurri, for correspondence during this project, without which my results may have been even more wrong.

## References


Blain, A.W. Detecting Gravitational Lensing Cosmic Shear from Samples of Several Galaxies Using Two-dimensional Spectral Imaging. ApJ, 570(2):L51–L54, May 2002. doi: 10.1086/341103. https://ui.adsabs.harvard.edu/abs/2002ApJ...570L..51B.

Courteau, S. Optical Rotation Curves and Linewidths for Tully-Fisher Applications. AJ, 114:2402, Dec. 1997. doi: 10.1086/118656. https://ui.adsabs.harvard.edu/abs/1997AJ....114.2402C.

Dopita, M., Hart, J., McGregor, P., Oates, P., Bloxham, G. et al. The Wide Field Spectrograph (WiFeS). Ap&SS, 310(3-4):255–268, Aug. 2007. doi: 10.1007/s10509-007-9510-z. https://ui.adsabs.harvard.edu/abs/2007Ap&SS.310..255D.

Dopita, M., Rhee, J., Farage, C., McGregor, P., Bloxham, G. et al. The Wide Field Spectrograph (WiFeS): performance and data reduction. Ap&SS, 327(2):245–257, June 2010. doi: 10.1007/s10509-010-0335-9. https://ui.adsabs.harvard.edu/abs/2010Ap&SS.327..245D.

Foreman-Mackey, D. corner.py: Scatterplot matrices in python. *The Journal of Open Source Software*, 1(2):24, jun 2016. doi: 10.21105/joss.00024.

Foreman-Mackey, D., Farr, W., Sinha, M., Archibald, A., Hogg, D. et al. emcee v3: A Python ensemble sampling toolkit for affine-invariant MCMC. *The Journal of Open Source Software*, 4 (43):1864, Nov. 2019. doi: 10.21105/joss.01864. https://ui.adsabs.harvard.edu/abs/2019JOSS....4.1864F.

Green, A.W., Glazebrook, K., McGregor, P.J., Damjanov, I., Wisnioski, E. et al. DYNAMO - I. A sample of H$\alpha$-luminous galaxies with resolved kinematics. MNRAS, 437(2):1070–1095, Jan. 2014. doi: 10.1093/mnras/stt1882. https://ui.adsabs.harvard.edu/abs/2014MNRAS.437.1070G.

Gurri, P., Taylor, E.N. and Fluke, C.J. The first shear measurements from precision weak lensing. MNRAS, 499(4):4591–4604, Sept. 2020. doi: 10.1093/mnras/staa2893. https://ui.adsabs.harvard.edu/abs/2020MNRAS.499.4591G.

Hogg, D.W., Bovy, J. and Lang, D. Data analysis recipes: Fitting a model to data. *arXiv e-prints*, art. arXiv:1008.4686, Aug. 2010. https://ui.adsabs.harvard.edu/abs/2010arXiv1008.4686H.

Miralda-Escude, J. Gravitational Lensing by Clusters of Galaxies: Constraining the Mass Distribution. ApJ, 370:1, Mar. 1991. doi: 10.1086/169789. https://ui.adsabs.harvard.edu/abs/1991ApJ...370....1M.

Morton, D.C. Atomic Data for Resonance Absorption Lines. I. Wavelengths Longward of the Lyman Limit. ApJS, 77:119, Sept. 1991. doi: 10.1086/191601. https://ui.adsabs.harvard.edu/abs/1991ApJS...77..119M.