# Characterizing Dark Matter Concentrations Through Magnitude Distortions due to Gravitational Lensing

Andre Menck

December 13, 2012

## 1 Problem Statement

In addition to observing astronomical objects through the electromagnetic radiation they give off, astronomers also study the universe by analyzing gravitational effects on mater. More specifically, one can work back from these observed effects to determine the mass distribution in the universe (both inside and outside galaxies). In this project, we worked on one specific gravitational effect: weak gravitational lensing–that is, the bending of light due to medium-sized clusters of galaxies. In particular, we attempted to measure the mass of unseen matter (also called dark matter) that should be present to produce the observed lensing effects. We analyzed data from the Dark Energy galaxy Survey (DES) simulations, found out where weak lensing occurs, and attempted to fit our model of mass distribution to the data in the simulations. While this problem has been attempted before (with varying degrees of success), the current techniques usually rely on the shape distortion of galaxies caused by weak lensing. In this project, we focused on the magnitude distortions, which should provide a roughly independent measure of lens mass. However, the intrinsic systematic uncertainty associated with this method is larger than in methods using shape distortion calculations, due to the higher amount of noise in the galaxy magnitude distribution.

## 2 Available Data

Given a cluster of galaxies (and an associated dark matter halo)[1], we possess list of galaxies $\{u^{(i)} : i = 1, ..., N\}$ near it[2]. Each galaxy data point consists of a list of parameters $u^{(i)} = (m^{(i)}, r^{(i)}, z^{(i)})$, defined as follows:

(1) $m^{(i)}$ is a vector of magnitudes in the G, R, I, Z, and Y bands (that is, each element of $m^{(i)}$ is the "brightness" of the galaxy, averaged over a range in wavelengths).

(2) $r^{(i)}$ is the distance (on the 2-dimensional sky plane) between the galaxy and the cluster center, given in arcminutes.

(3) $z^{(i)}$ is the redshift of the galaxy[3]

[1] Physically, what we observe is that clusters of galaxies and halos of dark matter go hand-in-hand, so that dark matter in the halos is what causes the gravitational lensing.

[2] Strictly speaking, these are the galaxies with a distance $r < R_{max}$ from the cluster center, where $r$ is the distance on the sky-plane, measured in arcminutes. Note that, in the three-dimensional picture, this includes galaxies that are "behind" the cluster–these are precisely the galaxies that will be affected by weak lensing.

[3] For the purpose of this project, we can think of redshift as the distance between the Earth and the given galaxy

## 3 Model and Relevant Physics

From the list of galaxies above, we wish to derive some prediction that can be parameterized by $\rho(\vec{r})$, the density of dark matter. To do so, we introduce the galaxy number density function $\phi$. If there were no gravitational lensing, $\phi$ would have some natural functional form in magnitude $m$–let us call this function $\phi_0$. Through some fairly involved calculations, it can be shown that, in the presence of a gravitational lens of (small) lensing parameter $\kappa$, the observed number density function is [1]:

$$\phi(r, m, z; \rho) = \frac{1}{1 + 2\kappa}\phi_0(m + 2\kappa, z)$$

Where the lensing parameter $\kappa = \kappa(r, z; \rho)$ is a complicated function of $r$ and $z$ that can be calculated from the mass distribution $\rho$ [2]. At this point, we will make a simplifying (but physical) assumption, that the mass distribution follows a Navarro, Frenk and White (NFW) mass distribution [3]:

$$\rho(r; R_S, c) = \frac{\delta_c(c)\rho_c(z)}{\frac{r}{R_S}\left(1 + \frac{r}{R_S}\right)^2}$$

This enables us to write down the dark matter distribution as a function parameterized by $R_S$, the virial radius. The total halo mass is then obtained by integrating the given mass distribution. This simplifies our problem to fitting this parameter to the available data. Note that, in order to recover any predictive power whatsoever, this assumption is necessary– after all, if we were fitting a general mass distribution, our parameter space would be infinite-dimensional.

## 4 Probabilistic Setup

The first problem we face is to find the functional form of $\phi_0$. For this, we take the galaxies at a radius $r > R_{outer}$ away from the center of the cluster, and consider them to be unlensed[4]. In the relevant range of magnitudes, we have found that a shifted power law provides a good fit to the number density of galaxies, as a function of $m$. More specifically, given a redshift range $z$, we model the probability of finding an unlensed galaxy of magnitude $m$ as: $p_z(m) = \beta(m-a)^\alpha$ for $m \in (a, m_{max})$[5]. Using this, we can write down the maximum likelihood estimate for the value of the parameters $\alpha$ and $a$:

$$\hat{\alpha} = \frac{1}{\log(m_{max} - a) - \frac{1}{N}\sum_{i=1}^{N}\log(m^{(i)})} - 1 \text{ and } \hat{a} = \min_i(m^{(i)})$$

Where $\beta$ acts as a normalization constant. Since $\phi_0$ is just the galaxy number density, we can relate these quantities as:

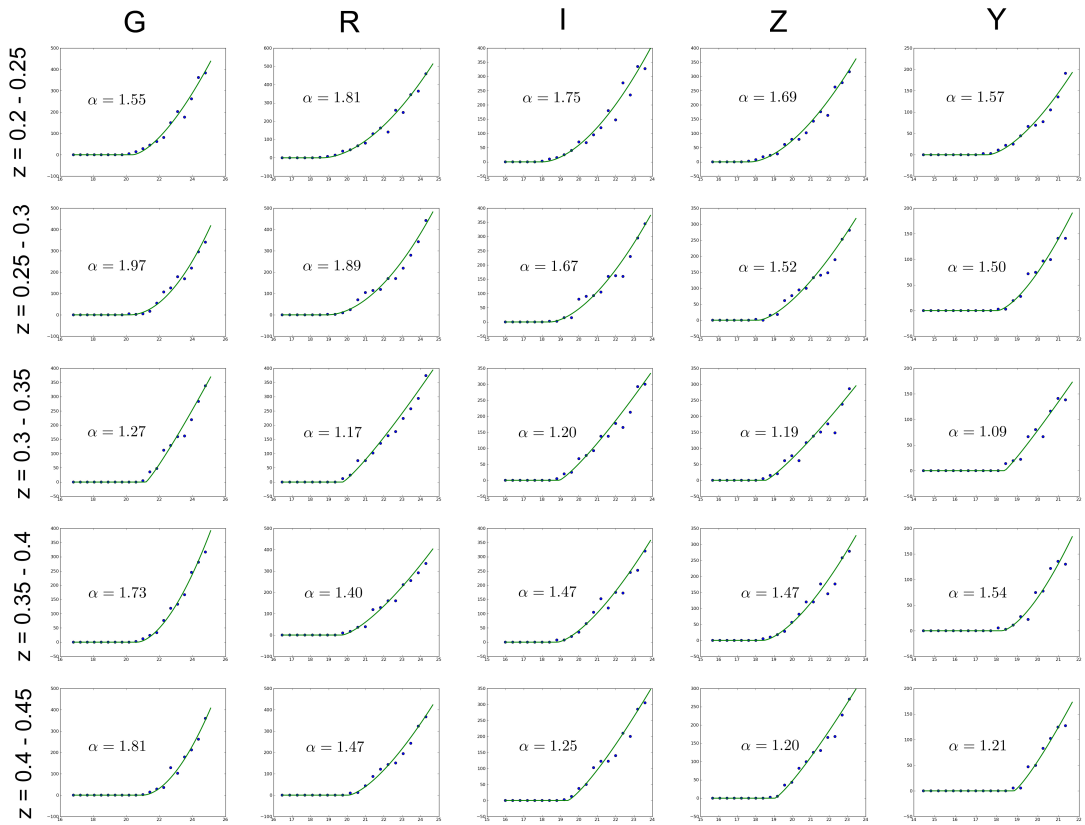
$$\phi_0(m) = N\frac{\beta}{\pi R^2}(m - a)^\alpha = \hat{\beta}(m - a)^\alpha$$

While this fit was only an intermediate step in order to reach our final goal, we will present a sample of the results. In the graphs to below, the fit $\phi_0(m)$ obtained is plotted alongside a few values calculated from the data[6], for each magnitude band and range of galaxy redshift:

[4] This can be justified by noting that the value of $\kappa$ at these radii is negligible, given halos of typical $R_s$

[5] We denote $p_z$ to make explicit that, for each redshift range, there is a different density function. To achieve this, the galaxies must be divided up into redshift bins.

[6] Note that these calculated data points are never used in the fit (that is, our fit only used individual galaxy magnitudes). In a loose sense, these points are the "histogram" of the galaxy magnitudes.




Once this fit is performed, we can move on to analyze the data of lensed galaxies ($r^{(i)} < R_{outer}$). Given a lensed galaxy at a radius $r$ and magnitude $m$, we want to find the probability $p(r, m)$ of observing this data point:

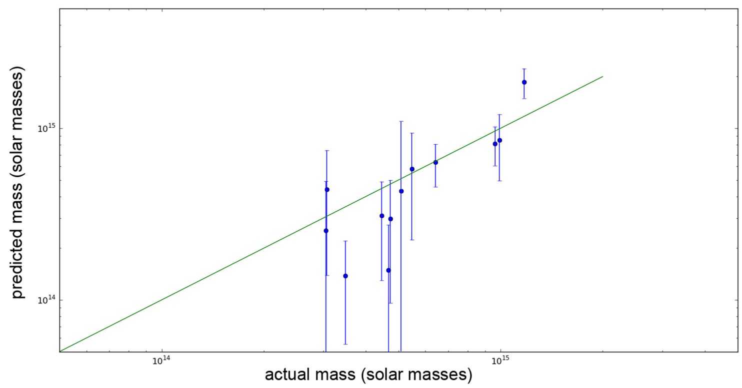
$$p(r, m) = \frac{E[\#\text{ of galaxies with magnitude} = m \text{ and radius} = r]}{E[\text{total }\#\text{ of galaxies}]} = \frac{\phi(r, m, z; R_S, c)}{\int_0^{R_{max}} \int_0^{m_{max}} \phi(r, m, z; R_S, c) dr dm}$$

The probability above enables us to (numerically) write down a log likelihood function for our observed data points $l(R_s)$. While this function can be expressed and differentiated analytically with respect to $R_s$, the resulting expression in unimaginably complex (due to dependence of $\kappa$ on $R_s$). Therefore, in this project it was decided to instead compute the likelihood function numerically, as in the expression above. Thus, to maximize $l$ with respect to $R_s$, we applied gradient ascent numerically (that is, the derivative is computed numerically), which results in a relatively efficient optimization process.

## 5 Results

After setting the fitting parameters[7], our initial attempt to implement the methods described above produced the results shown in the graph below. As can be seen, the weak lensing signal is quite weak for the less massive halos–the only conclusive result is for halos of mass larger than

[7] Interestingly, we found that the fit was relatively robust to changes in the maximum fitting radius (within reason, of course). However, the essential parameter to which the fitting is extremely sensitive to is the number of bins in galaxy redshift, set to 24.

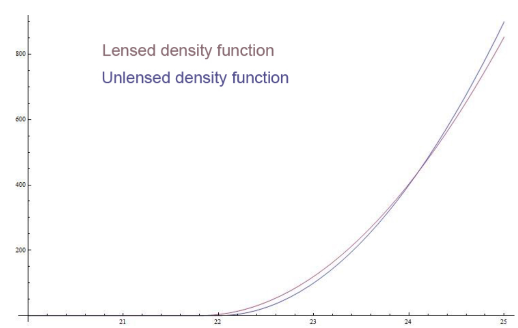



$5 \times 10^{15}$ solar masses. In studies of weak lensing, this is usually the best one can hope for [4], seeing as how the signal present in the galaxy magnitude data is often overwhelmed by noise. Specifically, the error of $\sim 30\%$ on more massive halos is competitive with those seen in galaxy shear (shape) distortion studies [4]. To readers who are unfamiliar with weak lensing studies, this might seem like a rather unsatisfying result–after all, why shouldn't we be able to recover the mass of a dark matter halo? To answer this question, we bring up two important factors that impact our analysis:

(1) The non-uniformity in the galaxy background $\phi_0$. In our model of this situation, we assumed that galaxies are roughly uniformly distributed in the sky–however, this is far from true.

(2) As depicted graphically below, the two shifts in the function (scaling and shift to the left) have an almost null combined effect, so that it is increasingly difficult to differentiate between lensed and unlensed magnitude distributions (for small values of $\kappa$).

## 6 Future Studies - The Color-Magnitude Map




In the interest of producing a better fit for the halo masses, this project will explore a possible remedy for the second of the two factors above, which might be applied to future studies. Specifically, if we were able to classify the galaxies by their color, then we could perform the analysis described above separately in each color group. The advantage that we wish to explore is that different colored galaxies, when grouped together, might produce a different galaxy background function $\phi_0$. If this is indeed the case, then one should be able to analyze the *combined* shift in each magnitude function, producing a stronger mass estimate (in other words, random noise is less likely to affect each function individually in such a way that mimicks lensing).

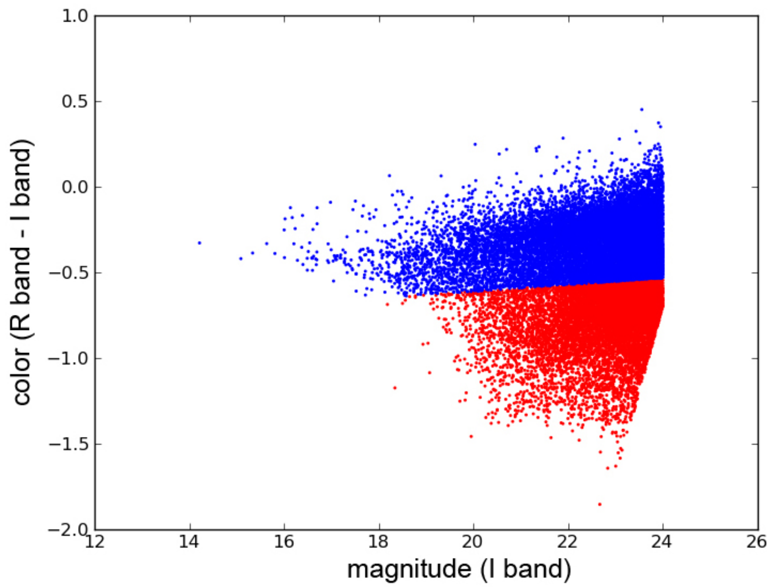
For this classification problem, we take galaxy color and magnitude as our input features, where color is defined as the difference between two observable wavelength bands. Then, we proceed to apply k-means clustering to classify the data.[8] One important factor that we ran into while applying this algorithm is the range of the data in each coordinate–while magnitude varies between 14 and 24, the color is usually between 0.5 and -1.5. Thus, in order for the algorithm to weigh each coordinate equally, we used the following kernel:

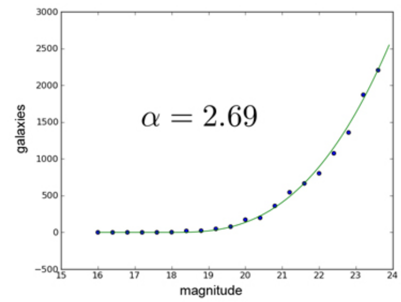
$$x^T y \mapsto \frac{x_1 y_1}{w} + x_2 y_2$$

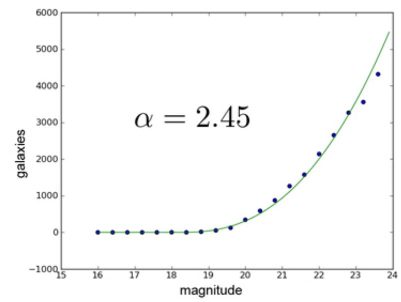
Where the weight $w$ was computed by equating the squared spread of each coordinate: $w = \frac{10^2}{2^2} = 25$. This method produced reasonable results (similar to those found in the literature when a

[8] In the literature, it is customary to apply Gaussian Discriminant Analysis to this problem–we choose to use k-means clustering because no training set was available for this given simulation.

training set is available [5]), depicted bellow. Also depicted are the fitted and numerical galaxy background functions for the I-band (averaged over all redshift)–there is an apparent difference in the two, reflected in the parameter estimates.




One would expect, due to this difference, that this classification will enable us to achieve better errors on our halo mass estimation problem. However, there is an additional problem that comes with separating the data as we have: namely, after binning the galaxies in redshift, each bin is left with very few galaxies, which then compromises the galaxy background function fit. The number of observed galaxies in our data set proved much too small to make this method work adequately, producing extremely poor fits for the galaxy background function in each bin of redshift (and thus, an extremely poor fit for the halo masses). More specifically, over 70% of redshift bins have less than 20 galaxies, making the power-law background fit unviable. This secondary result, however, is relevant to inform the direction that future studies of weak lensing might take. Given that the number of background galaxies is increased (say, by improvements in data-gathering techniques), one should expect a significant improvement in the halo mass calibration. In other words, a future experiment that used this technique in the data analysis would expect a reduction in the statistical error (that is, such a study should check if the statistical error in the mass estimation does indeed decrease faster than $O(1/\sqrt{N})$ due to this technique).

# 7 References

**[1]** Bartelmann, Matthias, and Peter Schneider. "Weak Gravitational Lensing."
**[2]** Candace, and Tereasa Brainerd. "Gravitational Lensing by NFW Halos."
**[3]** Navarro, Julio F., Carlos S. Frenk, and Simon D. M. White. "The Structure of Cold Dark Matter Halos." The Astrophysical Journal 462 (1996): 563
**[4]** Burke, David. et. all "Weighing the Giants III: Methods and Measurements of Accurate Galaxy Cluster Weak-Lensing Masses.""
**[5]** Wyder, Ted K., et. all "The UV-Optical Galaxy Color-Magnitude Diagram. I. Basic Properties." The Astrophysical Journal Supplement Series 173.2 (2007): 293-314"
A special thanks to David Burke, for his guidance and assistance in the entirety of this project.