Alternative Gravity Models Using Galactic Rotation Curves : A Machine Learning Analysis

Submitted by: **Anurag Sanjay Kadam**

Roll no: **200121007**

under the supervision of

Dr. Sovan Chakraborty
and
Dr. Sayan Chakrabarti

Department of Physics
Indian Institute of Technology, Guwahati

Abstract

In this work, we try to use modern machine learning techniques to fit parameters of the NFW dark matter profile to account for the discrepancies in the galaxy rotation curves. We mostly build on top of the work done by Álefe de Almeida et al. who used Markov chain Monte Carlo (MCMC) method to fit these parameters. MCMC gives us quite a good fit for the profile, however these values need to be fine tuned. We achieve this by taking inspiration from [1], and implementing the Stochastic Gradient Descent optimization algorithm provided by the machine learning framework PyTorch. We use the backpropagation technique to find the gradients which is widely used in modern machine learning applications. We obtain more precise values of the DM profile parameters, which correspond to a closer fit to the data.

Acknowledgements

I would like to express my deepest gratitude to Dr. Sovan Chakraborty and Dr. Sayan Chakraborti for their invaluable guidance, unwavering support, and mentorship throughout the course of this B.Tech. project. Their expertise in the field of astrophysics and commitment to scientific inquiry have been instrumental in shaping the direction and depth of this research.

I am grateful for the opportunities provided by them to delve into the intriguing and complex realm of gravity rotation curves in galaxies. Their encouragement to explore novel concepts and methodologies has been both inspiring and intellectually enriching.

I extend my thanks to the Indian Institute of Technology Guwahati, Department of Physics for providing an environment conducive to research and academic pursuits.

This project would not have been possible without the support and encouragement of Esha Bhatia, the PhD student I am working with. Her insights and collaborative spirit have added immense value to the overall experience.

Thank you all for being an integral part of this research endeavor.

Contents

List of Figures

Chapter 1

Introduction

The **Galaxy Rotation Problem** is an astrophysical observation that challenges the prediction of Newtonian Gravity and Classical Mechanics with respect to the **rotation curves of galaxies**. These curves are plots of the orbital speeds of stars or gas as a function of their distance from the galactic centre.

Classical Mechanics and Newtonian Gravity predict that orbital speeds of the objects in the galaxy reaches a peak at a certain radius and then decreases with increasing distance. This is because as one moves farther from the centre of the galaxy, the gravitational pull from the mass within the radius decreases, forcing the stars and gas to orbit slower to stay gravitationally bound to the galaxy.

However, observations of spiral galaxies have shown that the rotation curves do not follow this expected pattern. Instead, the curves remain relatively flat, indicating that the orbital speeds of stars and gas at large distances from the galactic center are unexpectedly high. This discrepancy would suggest that there is more mass in the outer regions of galaxies than can be accounted for by the visible/baryonic matter (stars and gas) alone.

To resolve this issue, physicists have proposed the existence of a hypothetical **Dark Matter (DM)**, which is a form of matter that does not emit, absorb or reflect electromagnetic radiation. This characteristic of Dark Matter is what makes it invisible and undetectable by traditional telescopes. This added mass distribution would provide the additional gravitational pull needed to explain the observed rotation curves of galaxies.

While Dark Matter has not been directly detected, its influence on the rotation curves of galaxies, as well as other cosmological observations, provides strong indirect evidence for its existence.

In this project, we choose the **Navarro-Frenk-White (NFW) profile** to model the underlying Dark Matter distribution. Our aim is to find the best fitting parameters for this profile using the **SPARC catalogue** of the galaxy rotation curve data. We have used **Stochastic Gradient Descent (SGD)** optimization technique to fit these parameters.

1.1 The SPARC Catalogue

We use the rotational curve observational data from the Spitzer Photometry & Accurate Rotation Curves (SPARC) catalogue [3]. This catalogue contains data of 175 disk galaxies.

Chapter 2

The Navarro-Frenk-White profile (NFW)

2.1 Introduction

The **Navarro-Frenk-White profile** is a mathematical model describing the distribution of dark matter in galaxies and galaxy clusters. This profile was proposed by Julio Navarro, Carlos Frenk, and Simon White in 1995 [4] as a way to better explain the observed density distribution of dark matter in simulations of cosmic structure formation.

The NFW profile describes the density of Dark Matter as a function of radius from the center of a halo (a gravitationally bound region of matter that can include galaxies and galaxy clusters). The profile is characterized by an inner cusp and an outer power-law decline. The NFW density profile is given by the formula:

$$\rho(r) = \frac{\rho_0}{\frac{r}{r_s}(1 + \frac{r}{r_s})^2} \tag{2.1}$$

where $\rho(r)$ is the dark matter density at a radius 'r', ρ_0 is a **characteristic density**, r_s is the **scale radius** that determines the transition between the inner cusp and the outer power-law decline.

The NFW profile is one of the several models of Dark Matter distributons has been found to be a good fit to the distribution of Dark Matter in simulations of structure formation, particularly in the context of cold dark matter (CDM) models. It has been used to model the dark matter halos surrounding galaxies and galaxy clusters, helping to understand the large-scale structure of the universe.

2.2 Modifications to the Profile

In theory, the parameters governing the NFW profile, namely the scale-radius r_s and the characteristic density ρ_0 should be independent. However, several N-body simulations have claimed relationships between them. These relationship is parameterized by a parameter called the **concentration parameter**:

$$c = \frac{r_{200}}{r_s} \tag{2.2}$$

Here, r_{200} is the a radius from the center of the galaxy which encloses a sphere of overall density of $200\rho_{crit}$. (Here ρ_{cirt} represents the critical density of the universe, which is a function of the Hubble's constant value at the galaxy's red-shift). The mass inside this radius is denoted as M_{200} and can thus be related as:

$$M_{200} = \frac{4\pi}{3} 200 \rho_{crit} r_{200}^3 \tag{2.3}$$

Inverting this relationship, we can write:

$$\rho_s = \frac{200}{3} \frac{c^3 \rho_{crit}}{\ln(1+c) - \frac{c}{1+c}} \tag{2.4}$$

and

$$r_s = \frac{1}{c} \left(\frac{3M_{200}}{4\pi(200\rho_{crit})} \right)^{1/3} \tag{2.5}$$

This leads to a relationship between c and M_{200}:

$$c(M_{200}) = 10^{.905} \left(\frac{M_{200}}{10^{12} h^{-1} M_\odot} \right)^{-.101} \tag{2.6}$$

Here, h is the dimensionless Hubble constant. We use the values $\rho_{crit} = 143.84 M_\odot/kpc^3$ and $h =$]0.671 [5].

Chapter 3

Mathematical Formulation

We consider the following components to evaluate the kinematics of disk galaxies : gas, disk, bulge and dark matter halo. The physical quantity we are modelling is the circular velocity in the galactic plane V_c which is related to the total gravitational potential Ψ as:

$$V_c^2 = r\frac{d\Psi}{dr} \tag{3.1}$$

Since the total gravitational potential is the sum of the potential due to all the components, it is given by:

$$\Psi = \Phi_{gas} + \Phi_{disk} + \Phi_{bulge} + \Phi_{NFW} \tag{3.2}$$

This also allows us to write the total circular velocity in terms of each component. Thus,

$$V_c^2 = V_{gas}^2 + \gamma_{*D}V_{disk}^2(r) + \gamma_{*B}V_{bulge}^2(r) + V_{NFW}^2(r) \tag{3.3}$$

Here, γ_{*D} and γ_{*B} refer to the stellar-mass-to-light ratio for the disk and bulge respectively.

3.1 Optimizing the Parameters

Assuming that the observed rotation curve data follows a Gaussian distribution, we can build a likelihood of the set of free parameters p_j =

$\{\gamma_{*D}, \gamma_{*B}, M_{200}\}$ for each galaxy with N observational points and σ_i data error as:

$$\mathcal{L}_j(p_j) = (2\pi)^{-N/2}\left\{\prod_{i=1}^{N}\frac{1}{\sigma_i^{-1}}\right\}\exp\left\{-.5\sum_{i=1}^{N}\left(\frac{V_{obs,j}(r_i) - V_c(r_i, p_j)}{\sigma_i}\right)^2\right\} \tag{3.4}$$

The observational radial velocity $V_{obs,j}$ is provided in the SPARC catalogue.

To constrain the parameter with respect to an entire set of galaxies, we must consider an overall likelihood $\mathcal{L}$. Since each galaxy's likelihood is independent of the other, the overall likelihood is simply the product of likelihood of each galaxy. This leaves us with an overall likelihood function of:

$$\mathcal{L}(\mathbf{p}) = \prod_{j=1}^{N_g}\mathcal{L}_j(p_j) \tag{3.5}$$

According to Bayes theorem, the posterior distribution is proportional to the prior times the likelihood. We add the constraints that: $.3 < \gamma_{*D}, \gamma_{*B} < .8$. A wide range of the other parameters is considered: $10^9 < \frac{M_{200}}{M_o dot} < 10^{14}$, where λ_0 is the mean value among the smallest observable radii when N_g galaxies are considered. The lower limit ensures non-divergence when $\lambda \to 0$.

The SPARC catalogue contains 175 galaxies, hence a complete analysis would require 384 free parameters: $175(\gamma_{*D}, M_{200}) + 32\gamma_{*B}$ (many galaxies do not show a bulge). The main paper refered to in this project so far [2] uses the MCMC optimization method. Thus, they decided to analyze 4 random sets of 10 galaxies each. For two sets (B and D) we have then 23 free parameters each, while for the other two sets we have 21 parameters each. The calculation of one set demands roughly one day of computation on a machine with 4 CPUs and 16 GB of RAM. We have used the same grouping and results in our project. The exact dataset from [2] used is given in Fig. 3.1.

Set	Galaxy	Best-fit values			$\chi^2_{\rm red}$
		$\Upsilon_{*\rm D}$	$\Upsilon_{*\rm B}$	$M_{200}(10^{11}M_\odot)$	
A	F568V1	$0.60^{+0.20}_{-0.10}$	-	$2.91^{+0.61}_{-0.82}$	0.35
A	NGC0024	$0.79^{+0.01}_{-0.01}$	-	$1.63^{+0.21}_{-0.28}$	1.68
A	NGC2683	$0.64^{+0.04}_{-0.04}$	$0.52^{+0.15}_{-0.21}$	$3.79^{+0.64}_{-0.81}$	1.37
A	NGC2915	$0.32^{+0.01}_{-0.02}$	-	$0.76^{+0.10}_{-0.14}$	0.98
A	NGC3198	$0.40^{+0.04}_{-0.05}$	-	$4.30^{+0.27}_{-0.29}$	1.31
A	NGC3521	$0.49^{+0.01}_{-0.02}$	-	$12.00^{+2.22}_{-2.81}$	0.37
A	NGC3769	$0.33^{+0.02}_{-0.03}$	-	$1.90^{+0.25}_{-0.32}$	0.75
A	NGC3893	$0.46^{+0.04}_{-0.04}$	-	$8.64^{+2.58}_{-2.22}$	1.26
A	NGC3949	$0.36^{+0.03}_{-0.05}$	-	$8.85^{+4.37}_{-5.81}$	0.45
A	NGC3953	$0.62^{+0.07}_{-0.07}$	-	$3.39^{+1.64}_{-2.82}$	0.73
B	NGC3992	$0.74^{+0.05}_{-0.03}$	-	$14.47^{+2.06}_{-2.11}$	0.88
B	NGC4051	$0.40^{+0.05}_{-0.10}$	-	$2.32^{+1.23}_{-1.64}$	1.27
B	NGC4088	$0.31^{+0.01}_{-0.01}$	-	$3.62^{+0.65}_{-0.76}$	1.09
B	NGC4100	$0.67^{+0.03}_{-0.03}$	-	$4.70^{+0.76}_{-0.83}$	1.20
B	NGC4138	$0.69^{+0.09}_{-0.05}$	$0.53^{+0.10}_{-0.21}$	$3.18^{+1.00}_{-1.44}$	2.67
B	NGC4157	$0.35^{+0.02}_{-0.03}$	$0.45^{+0.09}_{-0.15}$	$7.44^{+1.24}_{-1.38}$	0.76
B	NGC4183	$0.49^{+0.09}_{-0.14}$	-	$1.32^{+0.22}_{-0.25}$	0.19
B	NGC4559	$0.31^{+0.01}_{-0.01}$	-	$1.85^{+0.22}_{-0.22}$	0.43
B	NGC5005	$0.43^{+0.06}_{-0.11}$	$0.50^{+0.07}_{-0.08}$	$52.94^{+32.80}_{-49.57}$	0.08
B	NGC6503	$0.45^{+0.02}_{-0.03}$	-	$1.99^{+0.20}_{-0.22}$	1.91
C	UGC06983	$0.51^{+0.11}_{-0.16}$	-	$1.66^{+0.29}_{-0.36}$	0.69
C	UGC07261	$0.53^{+0.12}_{-0.21}$	-	$0.41^{+0.11}_{-0.14}$	0.17
C	UGC07690	$0.68^{+0.11}_{-0.06}$	-	$0.13^{+0.05}_{-0.05}$	0.89
C	UGC07866	$0.38^{+0.06}_{-0.08}$	-	$0.02^{+0.07}_{-0.02}$	2.52
C	UGC08490	$0.78^{+0.02}_{-0.01}$	-	$0.60^{+0.08}_{-0.10}$	0.78
C	UGC08550	$0.49^{+0.08}_{-0.17}$	-	$0.18^{+0.04}_{-0.04}$	1.02
C	UGC08699	$0.71^{+0.05}_{-0.05}$	$0.67^{+0.03}_{-0.05}$	$6.86^{+1.16}_{-1.34}$	0.86
C	UGC09992	$0.43^{+0.10}_{-0.13}$	-	$0.03^{+0.03}_{-0.03}$	1.98
C	UGC10310	$0.53^{+0.10}_{-0.22}$	-	$0.28^{+0.06}_{-0.08}$	1.25
C	UGC12506	$0.78^{+0.02}_{-0.01}$	-	$17.77^{+2.53}_{-2.22}$	1.22
D	NGC7331	$0.32^{+0.01}_{-0.01}$	$0.49^{+0.08}_{-0.18}$	$20.56^{+0.86}_{-0.79}$	0.87
D	NGC7793	$0.41^{+0.05}_{-0.05}$	-	$1.01^{+0.21}_{-0.24}$	0.95
D	NGC7814	$0.76^{+0.04}_{-0.03}$	$0.60^{+0.03}_{-0.03}$	$21.39^{+2.07}_{-2.09}$	0.82
D	UGC02259	$0.72^{+0.08}_{-0.05}$	-	$0.75^{+0.09}_{-0.11}$	2.84
D	UGC03546	$0.55^{+0.04}_{-0.04}$	$0.38^{+0.04}_{-0.04}$	$9.33^{+0.66}_{-0.64}$	1.05
D	UGC06446	$0.50^{+0.09}_{-0.19}$	-	$0.56^{+0.08}_{-0.09}$	0.25
D	UGC06930	$0.40^{+0.06}_{-0.10}$	-	$1.19^{+0.21}_{-0.20}$	0.62
D	UGC06983	$0.40^{+0.05}_{-0.09}$	-	$1.65^{+0.21}_{-0.23}$	0.67
D	UGC07261	$0.49^{+0.09}_{-0.18}$	-	$0.34^{+0.08}_{-0.10}$	0.11
D	UGC07690	$0.66^{+0.13}_{-0.07}$	-	$0.10^{+0.03}_{-0.04}$	0.72

Figure 3.1: The maximum likelihood estimation for the parameters of each galaxy data point as found in [2]

Chapter 4

Stochastic Gradient Descent(SGD)

Taking inspiration from [1], we used the python Machine Learning framework **PyTorch** [6] and used its autograd module to perform Stochastic Gradient Descent(SGD) [7].

Gradient descent is an optimization technique widely used in Machine Learning (ML) and Numerical Optimization. Its primary objective is to minimize a cost function by iteratively adjusting the parameters of a model or system. The key idea behind Gradient Descent is to move in the direction of steepest descent of the cost function. In each iteration, the algorithm calculates the gradient of the cost function with respect to the parameters, indicating the direction of the fastest increase. It then updates the parameters by taking a step proportional to the negative of the gradient. The size of the step is controlled by a parameter called the learning rate. This iterative process continues until convergence, where the algorithm reaches a minimum or a point where further changes in parameters yield negligible improvements.

Given a function F to be minimized, it will implement the formula:

$$\mathbf{p}^{new} = \mathbf{p}^{old} - \gamma \nabla_{\mathbf{p}} F(\mathbf{p}^{old}) \tag{4.1}$$

where $\mathbf{p}$ stands for the independent variables of the function F and γ is the parameter called the learning rate, whose aim is to regulate the "length" of the steps.

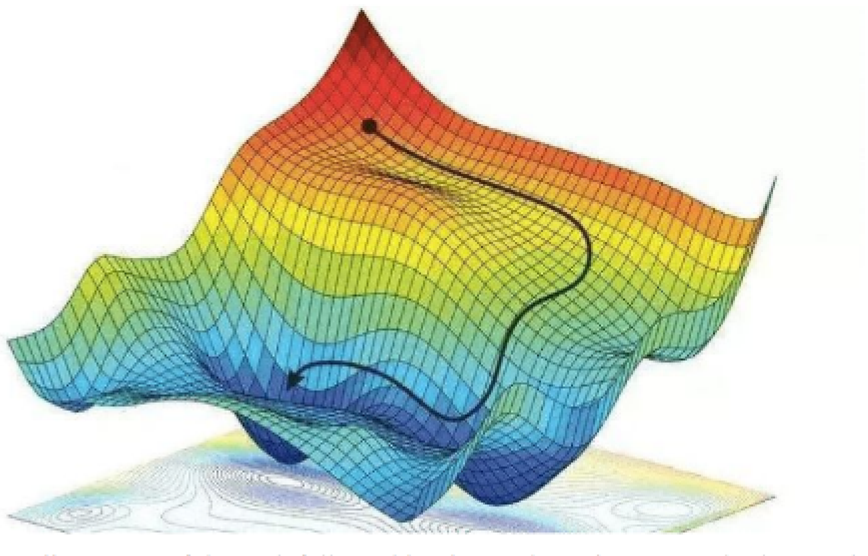

Figure 4.1: Illustration of the path followed by the gradient descent algorithm to reach the minimum of a 2-D function. Taken from https://easyai.tech/en/ai-definition/gradient-descent

Gradient Descent is a foundational optimization technique and forms the basis for more advanced algorithms, playing a crucial role in training machine learning models by minimizing the error or loss function. The negative-likelihood loss used in this project is also a well-studied loss function in Machine Learning literature.

4.1 Stochastic Gradient Descent

Stochastic Gradient Descent is a variant of Gradient Descent. In standard Gradient Descent, the entire training dataset is used to compute the gradient of the cost function, and then the parameters are updated.

In Stochastic Gradient Descent, the optimization is performed using only a subset (or even just one) randomly chosen training sample to compute the gradient and update the parameters. This random sampling introduces "stochasticity" in the optimization process, and it can lead to faster convergence and is particularly useful when dealing with large datasets. It also helps escape from local minima.

4.2 Results

After about 100,000-200,000 (based on galaxy group) iterations, which take 10-15 minutes, the negative log-likelihood saturates to a minimum as shown in Fig. 4.2. The fuzzy looking oscillation happens at the end since the model has reached very close to the best local value and taking even a minor step leads to overshooting to the opposite side of minimum. Since the changes being observed at this stage are negligible to compared to the variations caused by the errors in data measurement, hence it can be considered safe to stop the optimization loop.

We plot the data-points, the fit found using SGD and the MCMC fits given in [2] (ref Fig.4.3 to Fig. 4.6). The plots clearly show us that the new found parameters obtained using SGD give us a significant improvement over the pre-existing values found in the literature using just MCMC.

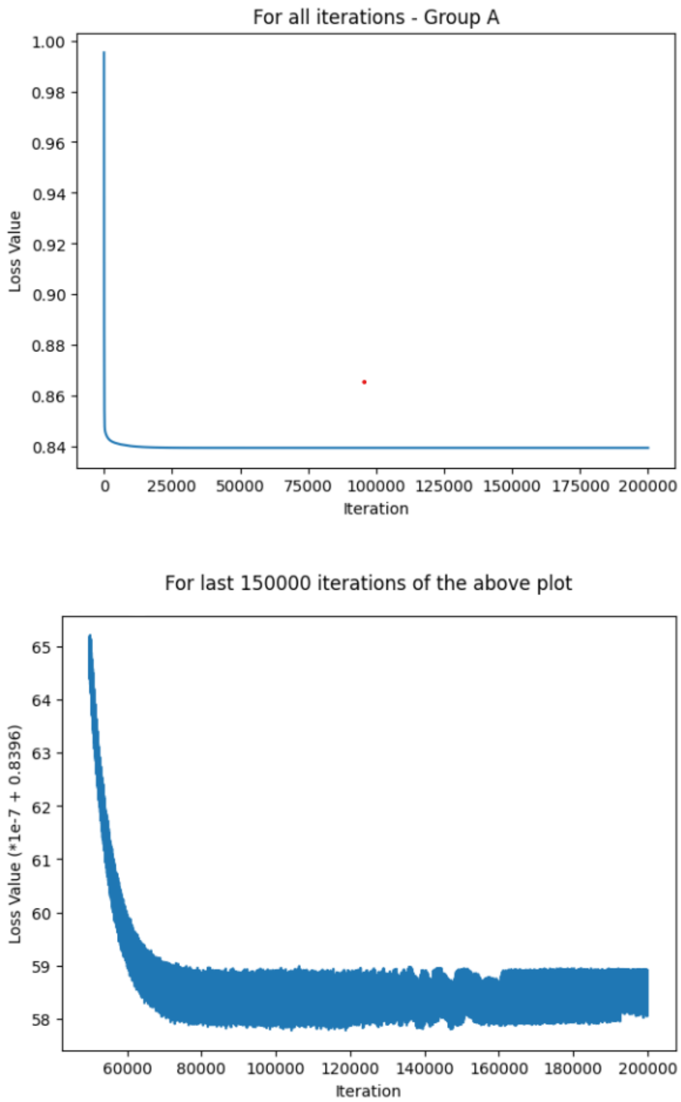

Figure 4.2: All the iterations of the gradient descent algorithm

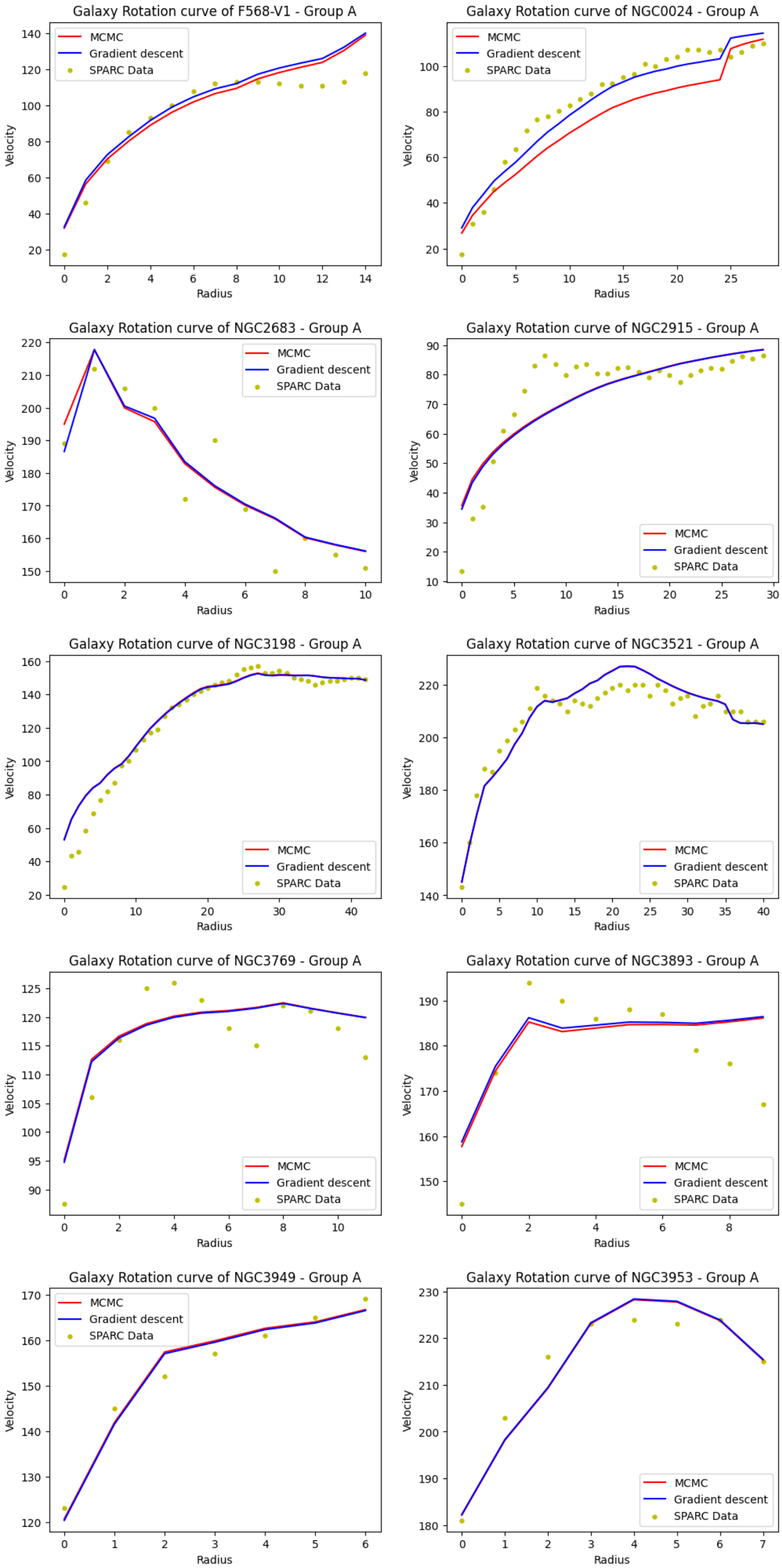

Figure 4.3:
R vs. V Rotation curves comparing Data Points, model found using MCMC and model found using SGD - For group A

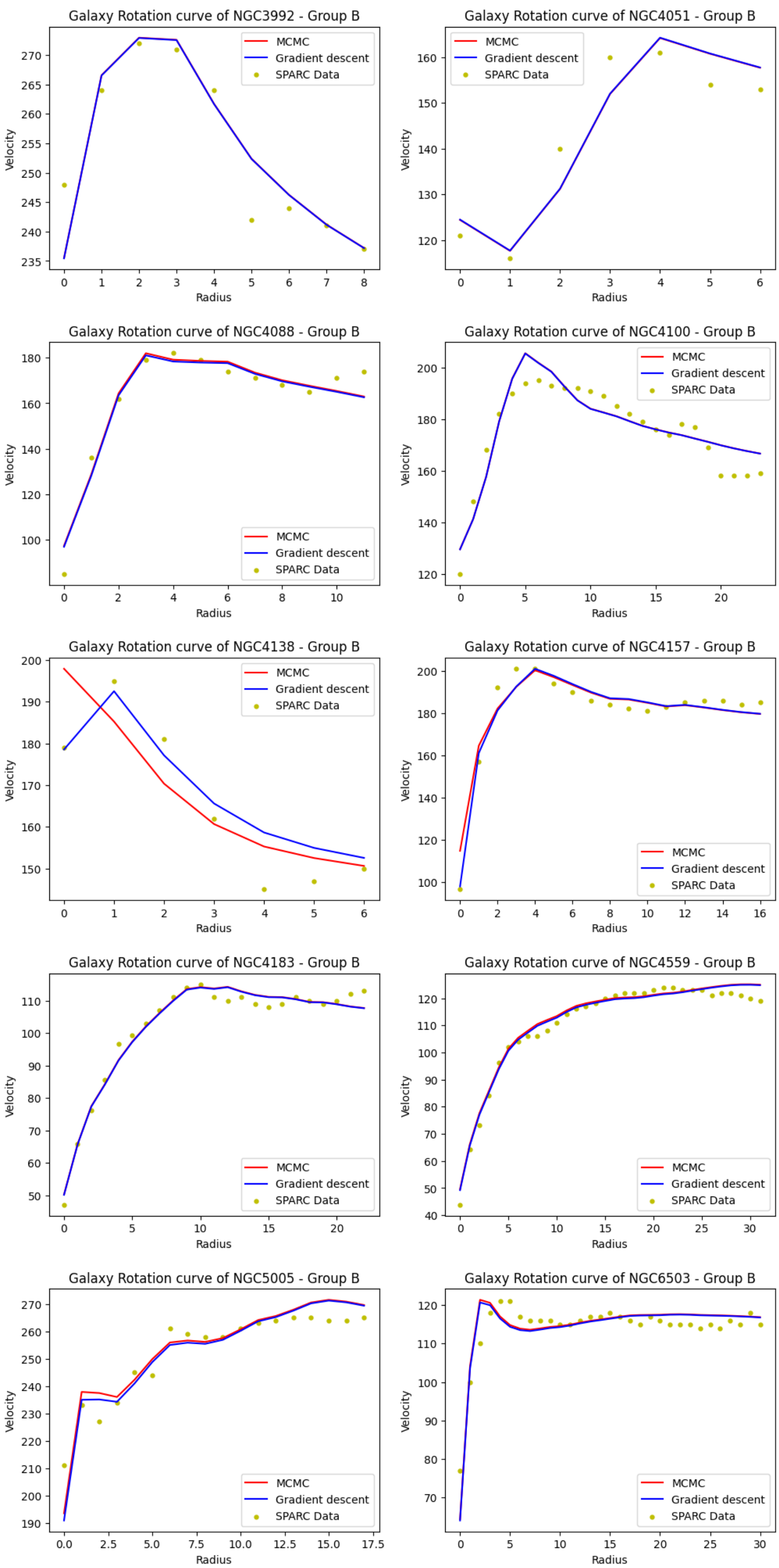

Figure 4.4:
R vs. V Rotation curves comparing Data Points, model found using MCMC and model found using SGD - For group B

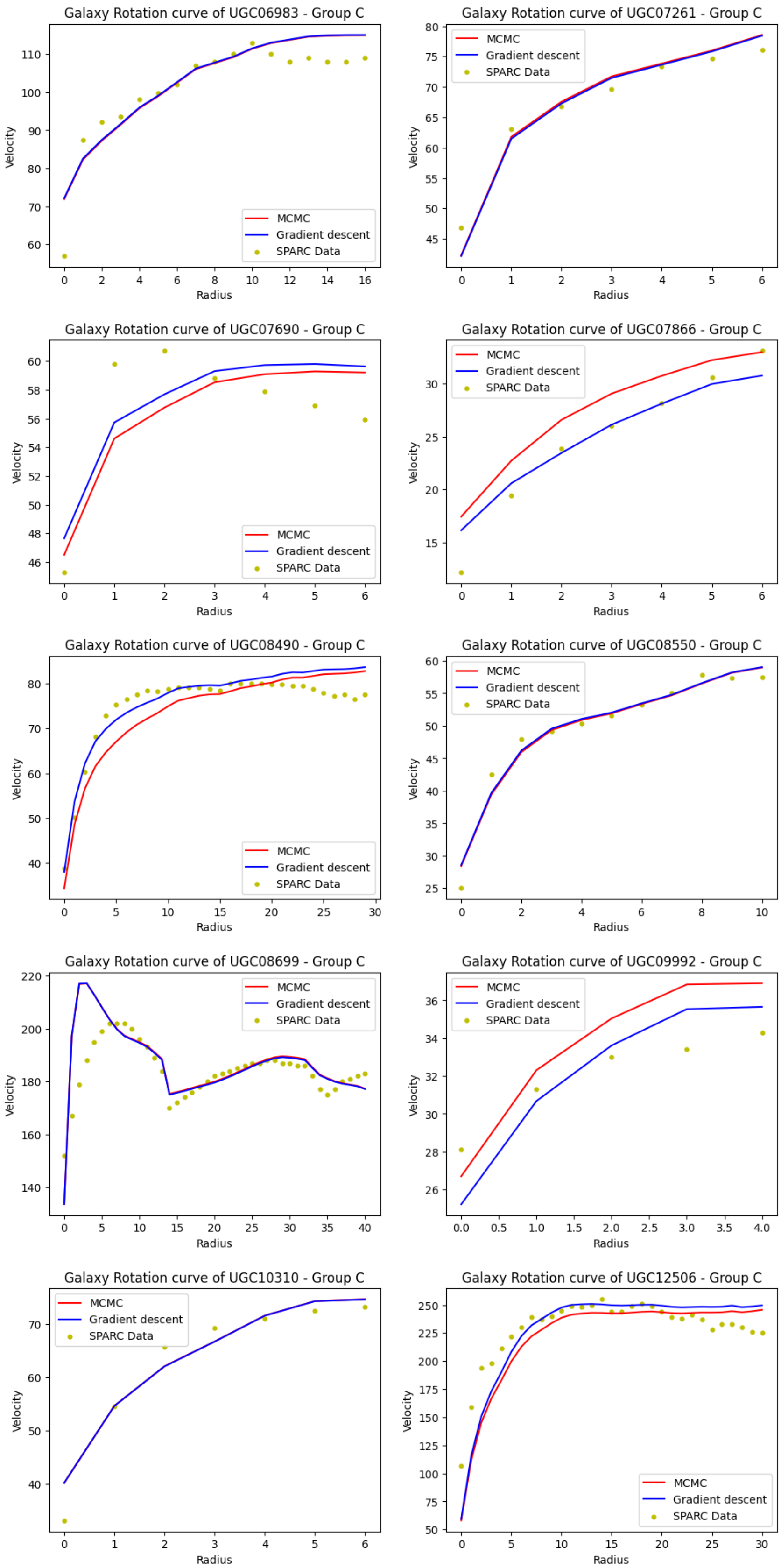

Figure 4.5:
R vs. V Rotation curves comparing Data Points, model found using MCMC and model found using SGD - For group C

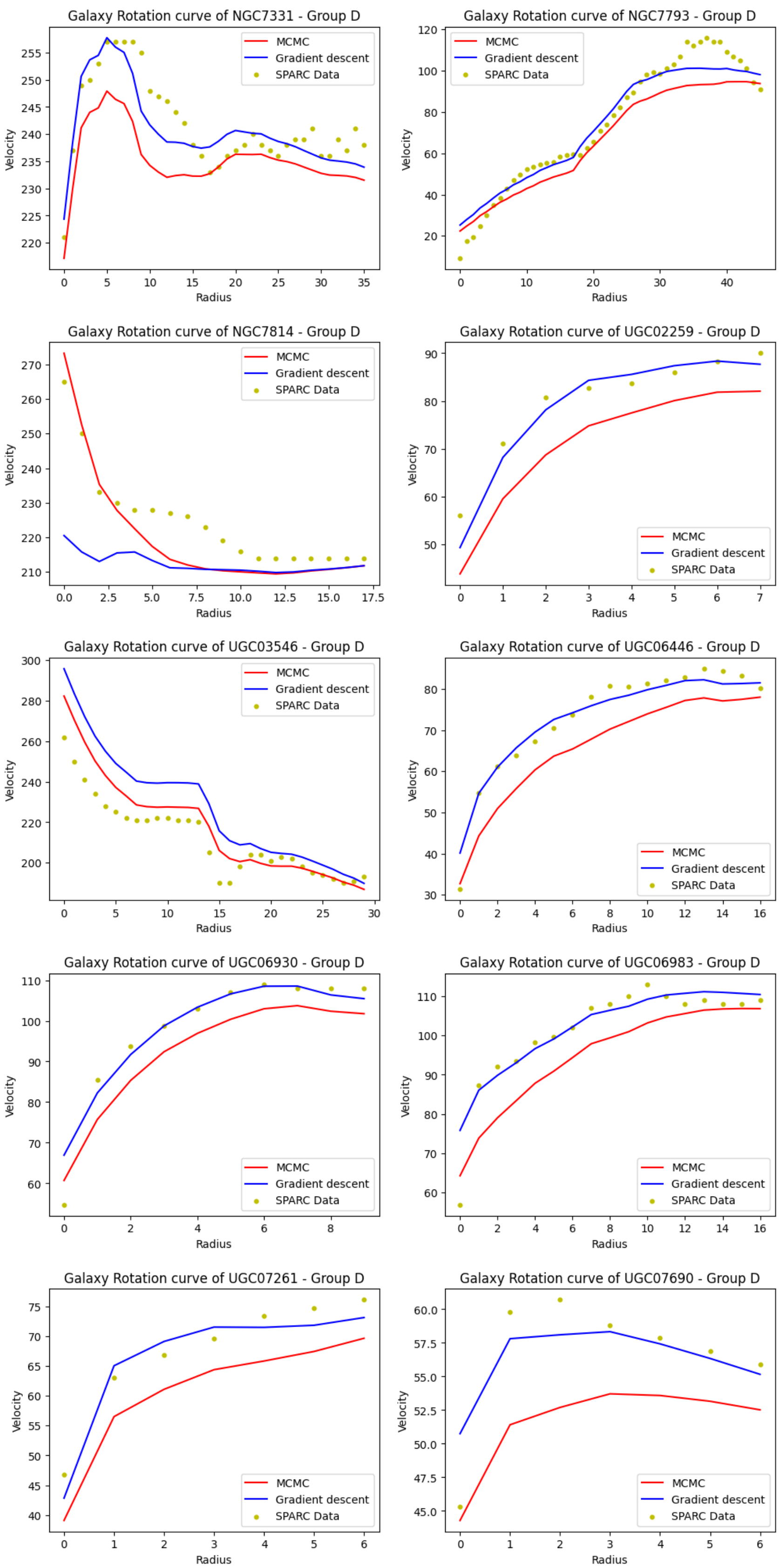

Figure 4.6:
R vs. V Rotation curves comparing Data Points, model found using MCMC and model found using SGD - For group D

Chapter 5

Future of the Project

In the future of this project we would like to do the following:

- The Amendola paper also additionally incorporates a modified Newtonian gravity on top of the dark matter model. We could not include that as its expression contains the exponential integral function defined as

$$Ei(x) = -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt \tag{5.1}$$

  which is not supported by PyTorch as of November 2023. This is something we would like to implement.

- We can try getting the same, or even better, results from scratch using some other optimization algorithm that's less computationally intensive.

Bibliography

[1] C. R. Argüelles and S. Collazo, "Galaxy rotation curve fitting using machine learning tools," *Universe*, vol. 9, p. 372, Aug 2023.

[2] Á . de Almeida, L. Amendola, and V. Niro, "Galaxy rotation curves in modified gravity models," *Journal of Cosmology and Astroparticle Physics*, vol. 2018, pp. 012–012, aug 2018.

[3] F. Lelli, S. S. McGaugh, and J. M. Schombert, "SPARC: MASS MODELS FOR 175 DISK GALAXIES WITH iSPITZER/i PHOTOMETRY AND ACCURATE ROTATION CURVES," *The Astronomical Journal*, vol. 152, p. 157, nov 2016.

[4] J. F. Navarro, C. S. Frenk, and S. D. M. White, "The structure of cold dark matter halos," *The Astrophysical Journal*, vol. 462, p. 563, may 1996.

[5] P. Collaboration, "Planck 2015 results. xiii. cosmological parameters," *Astronomy &amp Astrophysics*, vol. 594, p. A13, sep 2016.

[6] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, pp. 8024–8035, Curran Associates, Inc., 2019.

[7] S. Ruder, "An overview of gradient descent optimization algorithms," *CoRR*, vol. abs/1609.04747, 2016.