



**RHODES UNIVERSITY**  
*Where leaders learn*

RHODES UNIVERSITY

MASTERS THESIS

---

# A Bayesian Approach to Tilted-Ring Modelling of Galaxies

---

**Author:**

Eric Kamau MAINA

**Supervisors:**

Dr Gyula I. G. JÓZSA

Dr Michelle LOCHNER

Prof. Oleg SMIRNOV

*A thesis submitted in fulfilment of the requirements  
for the degree of Master of Science  
in Astrophysics*

*in the*

Centre for Radio Astronomy Techniques & Technologies  
Department of Physics and Electronics

March 5, 2020

## Declaration of Authorship

I, [Eric Kamau MAINA](#), declare that this thesis titled, “A Bayesian Approach to Tilted-Ring Modelling of Galaxies” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

---

Date:

---

RHODES UNIVERSITY

# *Abstract*

Faculty of Science  
Department of Physics and Electronics

Master of Science

## **A Bayesian Approach to Tilted-Ring Modelling of Galaxies**

by **Eric Kamau MAINA**

The orbits of neutral hydrogen (H I) gas found in most disk galaxies are circular and also exhibit long-lived warps at large radii where the restoring gravitational forces of the inner disk become weak (Spekkens & Giovanelli 2006). These warps make the tilted-ring model an ideal choice for galaxy parametrisation. Analysis software utilizing the tilted-ring-model can be grouped into two and three-dimensional based software. Józsa et al. (2007b) demonstrated that three dimensional based software is better suited for galaxy parametrisation because it is affected by the effect of beam smearing only by increasing the uncertainty of parameters but not with the notorious systematic effects observed for two-dimensional fitting techniques.

TiRiFiC, The Tilted Ring Fitting Code (Józsa et al. 2007b), is a software to construct parameterised models of high-resolution data cubes of rotating galaxies. It uses the tilted-ring model, and with that, a combination of some parameters such as surface brightness, position angle, rotation velocity and inclination, to describe galaxies. TiRiFiC works by directly fitting tilted-ring models to spectroscopic data cubes and hence is not affected by beam smearing or line-of-site-effects, e.g. strong warps.

Because of that, the method is unavoidable as an analytic method in future H I surveys. In the current implementation, though, there are several drawbacks. The implemented optimisers search for local solutions in parameter space only, do not quantify correlations between parameters and cannot find errors of single parameters. In theory, these drawbacks can be overcome by using Bayesian statistics, implemented in Multinest (Feroz et al. 2008), as it allows for sampling a posterior distribution irrespective of its multimodal nature resulting in parameter samples that correspond to the maximum in the posterior distribution. These parameter samples can be used as well to quantify correlations and find errors of single parameters. Since this method employs Bayesian statistics, it also allows the user to leverage any prior information they may have on parameter values.

## *Acknowledgements*

I would like to thank Prof. Oleg Smirnov and SARA0/NRF for funding my masters, and Dr. Okengo for informing me of this funding. I also thank Dr. Gyula Józsa and Dr. Michelle Lochner for their support and mentorship through my thesis.

Special thanks to Ulrich Mbou Sob, Siyanda Matika, Alexander Akoto-danso, Cyndie Russeeawon, Lexy Andati, Isabella Rammala and other members of our group for their input as well as proofreading my work and sending back useful comments.

For friends who stuck with me through thick and thin, I would like to thank Noluvuyo Vuyo Matiwane, Alexander Akoto-danso, Cyndie Russeeawon, Lexy Andati and Portia Phetogo, as well as my family and friends for their support and encouragement. And finally, I thank God for sustaining me this far as it would have not been possible without him.



# Contents

<b>Declaration of Authorship</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Astronomical and Theoretical Background . . . . .	1
1.1.1 Galactic neutral atomic hydrogen H I . . . . .	3
1.1.2 Structure and mass of H I in galaxies . . . . .	6
1.2 Fundamentals of radio observations . . . . .	9
1.2.1 Source intensity . . . . .	9
1.2.2 Doppler shift effect . . . . .	11
1.2.3 Single dish antennas . . . . .	12
1.2.4 Interferometers . . . . .	14
1.3 H I Spectroscopic data cube . . . . .	15
1.3.1 Moment maps . . . . .	16
1.4 Galaxy parametrisation . . . . .	18
1.4.1 Deriving galaxy kinematic parameters . . . . .	19
1.4.2 Velocity measure by spectral line Doppler broadening . . . . .	19
1.4.3 The tilted-ring-model . . . . .	21
1.4.4 Two dimensional fitting . . . . .	22
1.4.5 Three-dimensional based fitting . . . . .	26
1.4.6 Distance measure by the Tully-Fisher Relation . . . . .	27
<b>2 Bayesian statistics</b>	<b>30</b>
2.1 Probability . . . . .	30
2.1.1 Conditional Probability . . . . .	31
2.1.2 Bayesian Statistics . . . . .	31
2.1.3 Ellipsoidal nested sampling . . . . .	33
2.1.4 MultiNest algorithm background . . . . .	34
<b>3 TiRiFiC-MultiNest</b>	<b>40</b>
3.1 TiRiFiC . . . . .	40
3.1.1 TiRiFiC parameters . . . . .	41
3.1.2 TiRiFiC Chi-Square evaluation . . . . .	42

3.2	Fitting Algorithms in TiRiFiC . . . . .	43
3.2.1	Golden selection method . . . . .	43
3.2.2	The Simplex Algorithm . . . . .	44
3.2.3	Pswarm minimizer . . . . .	46
3.2.4	Selecting TiRiFiC minimizer . . . . .	46
3.3	TiRiFiC-MultiNest General description . . . . .	47
3.3.1	Cube layout details . . . . .	50
3.3.2	Parallelisation . . . . .	52
<b>4</b>	<b>Tests and Results</b>	<b>53</b>
4.1	Testing . . . . .	53
4.1.1	Reducing the prior range using SoFiA . . . . .	53
4.1.2	SoFiA best-fit parameter values . . . . .	55
4.1.3	Fits file data preparation . . . . .	55
4.1.4	Error calculation and correlation plots . . . . .	56
4.1.5	Experiment 1 . . . . .	56
4.1.6	Experiment 2 . . . . .	59
4.1.7	Experiment 3 . . . . .	61
4.1.8	Experiment 4 . . . . .	63
4.1.9	Experiment 5 . . . . .	66
4.1.10	Experiment 6 . . . . .	69
4.2	Further Attempts at reducing the number of likelihood evaluations . . . . .	71
4.2.1	Experiment 7 . . . . .	71
4.2.2	Experiment 8 . . . . .	75
4.2.3	Experiment 9 . . . . .	78
4.2.4	Experiment 10 . . . . .	80
4.2.5	Experimental runs summary . . . . .	82
4.2.6	Test with real galaxy NGC 3198 . . . . .	85
4.2.7	Test with real galaxy NGC 4062 . . . . .	86
<b>5</b>	<b>Summary and Conclusion</b>	<b>89</b>
5.1	Implementation Summary . . . . .	89
5.2	Conclusion . . . . .	94
<b>A</b>	<b>Software Documentation</b>	<b>95</b>
A.1	TiRiFiC Usage . . . . .	95
A.2	Multinest Example . . . . .	95
A.3	Parameters set for SoFiA run . . . . .	96
A.4	Minimizing Prior Values Experiment 1 . . . . .	96
A.5	Minimizing Prior Values Experiment 2 . . . . .	97
A.6	Minimizing Prior Values Experiment 3 . . . . .	98
A.7	Minimizing Prior Values Experiment 4 . . . . .	99
A.8	Minimizing Prior Values Experiment 5 . . . . .	100

A.9 Minimizing Prior Values Experiment 6 . . . . .	105
A.10 Minimizing Prior Values Experiment 7 . . . . .	106
A.11 Minimizing Prior Values Experiment 8 . . . . .	113
A.12 Minimizing Prior Values Experiment 9 . . . . .	113
A.13 Minimizing Prior Values Experiment 10 . . . . .	114
A.14 Minimizing Prior Values Test with real galaxy N3198 . . . . .	121
A.15 Minimizing Prior Values Test with real galaxy N4062 . . . . .	122

# List of Figures

1.1	Gravitational lensing effect . . . . .	2
1.2	H I and optical distribution of NGC 3109 . . . . .	4
1.3	spin-flip transition . . . . .	4
1.4	Warp in galaxy disk of NGC 2862 . . . . .	8
1.5	Gravitational tidal influence on H I content of M 81 . . . . .	9
1.6	Figure illustrating source intensity . . . . .	10
1.7	Doppler shift illustration . . . . .	11
1.8	Single dish antenna . . . . .	12
1.9	Antenna primary beam pattern . . . . .	13
1.10	A spectroscopic data cube . . . . .	15
1.11	Moment maps . . . . .	16
1.12	Intensity weighted mean of a skewed and noisy velocity profile . . . . .	17
1.13	Galaxy position angle . . . . .	19
1.14	Velocity profile of thermal motion in a gas cloud . . . . .	20
1.15	The tilted-ring-model . . . . .	22
1.16	Global H I profile of a galaxy observed face-on . . . . .	24
1.17	Global H I profile of a galaxy with high inclination angle . . . . .	24
1.18	Solid body global H I profile . . . . .	25
1.19	moment 1 and 2 with high and low resolution observation . . . . .	26
1.20	The Tully-Fisher relation . . . . .	28
2.1	Ellipsoidal sampling . . . . .	34
2.2	MultiNest ellipsoidal decomposition . . . . .	37
2.3	Ellipsoidal bounds groups . . . . .	38
3.1	TiRiFiC flow chart . . . . .	42
3.2	The Golden Section minimization (Press et al. 2007). . . . .	43
3.3	The simplex method . . . . .	45
3.4	Multinest flow chart diagram . . . . .	48
3.5	TiRiFiC-MultiNest flowchart diagram with a cube of length 19 . . . . .	51
3.6	Parallel TiRiFiC-Multines . . . . .	52
4.1	MultiNest cube of length 7 . . . . .	57
4.2	Experiment 1 Rotation velocity . . . . .	58
4.3	MultiNest cube of length 13 . . . . .	59
4.4	Experiment 2 Rotation velocity . . . . .	60

4.5	Experiment 2 TiRiFiC fit Vs MultiNest fit with 13 parameters . . . . .	61
4.6	Experiment 3 Rotation velocity . . . . .	62
4.7	Experiment 3 TiRiFiC fit Vs MultiNest fit with 13 parameters . . . . .	63
4.8	PV slice of NGC 4062 . . . . .	64
4.9	Experiment 4 rotation velocity . . . . .	65
4.10	Experiment 4 observed cube Vs MultiNest fit with 19 parameters . . . . .	66
4.11	MultiNest cube of length 19 . . . . .	67
4.12	Experiment 5 rotation velocity . . . . .	68
4.13	Experiment 5 TiRiFiC fit Vs MultiNest fit with 19 parameters . . . . .	69
4.14	Experiment 6 rotation velocity . . . . .	70
4.15	Experiment 6 TiRiFiC fit Vs MultiNest fit with 19 parameters . . . . .	71
4.16	Experiment 7a fitted parameters . . . . .	72
4.17	Experiment 7a TiRiFiC fit Vs MultiNest fit with 19 parameters . . . . .	73
4.18	Experiment 7b fitted parameters . . . . .	74
4.19	Experiment 7b TiRiFiC fit Vs MultiNest fit with 19 parameters . . . . .	75
4.20	MultiNest cube of length 31 . . . . .	76
4.21	Experiment 8 fitted parameters . . . . .	77
4.22	Experiment 7b TiRiFiC fit Vs MultiNest fit with 31 parameters . . . . .	78
4.23	Experiment 9 fitted parameters . . . . .	79
4.24	Experiment 9 TiRiFiC fit Vs MultiNest fit with 31 parameters . . . . .	80
4.25	Experiment 10 fitted parameters . . . . .	81
4.26	Experiment 7b TiRiFiC fit Vs MultiNest fit with 31 parameters . . . . .	82
4.27	Test with real galaxy NGC 3198 . . . . .	85
4.28	Observed galaxy NGC 3198 Vs TiRiFiC-MultiNest fit with 31 parameters . . . . .	86
4.29	Test with real galaxy NGC 4062 . . . . .	87
4.30	Observed galaxy NGC 4062 Vs TiRiFiC-MultiNest fit with 31 parameters . . . . .	88
5.1	Experiment 5 vs Experiment 7 . . . . .	91
5.2	correlation between rotation velocity and Inclination angle . . . . .	93
A.1	Position angle and inclination angle correlation Experiment 5 . . . . .	101
A.2	Position angle and surface brightness correlation Experiment 5. . . . .	102
A.3	Rotation velocity and position angle correlation Experiment 5. . . . .	103
A.4	Rotation velocity and surface brightness correlation Experiment 5. . . . .	104
A.5	Rotation velocity and inclination angle correlation Experiment 5 . . . . .	105
A.6	Rotation velocity and inclination angle correlation Experiment 7 . . . . .	108
A.7	Rotation velocity and position angle correlation Experiment 7 . . . . .	109
A.8	Rotation velocity and surface brightness correlation Experiment 7 . . . . .	110
A.9	Position angle and inclination angle correlation Experiment 7 . . . . .	111
A.10	Position angle and surface brightness correlation Experiment 7 . . . . .	112
A.11	Rotation velocity and Inclination angle correlation Experiment 10 . . . . .	116
A.12	Rotation velocity and position angle correlation Experiment 10 . . . . .	117
A.13	Rotation velocity and surface brightness correlation Experiment 10 . . . . .	118

A.14 Position angle and inclination angle correlation Experiment 10 . . . . .	119
A.15 Position angle and surface brightness correlation Experiment 10 . . . . .	120

# List of Tables

4.1	Comparison between SOFIA and TIRIFIC-Multienst . . . . .	57
4.2	Experiment runs summary . . . . .	84

*Dedicated to my parents*



# Chapter 1

## Introduction

This Chapter is aimed at giving examples of ways in which TiRiFiC may be used to advance the current knowledge base in astronomy as well as give the user a broad overview in astronomy.

### 1.1 Astronomical and Theoretical Background

Before the 20th century, our stellar system, the Milky Way, was perceived to be the entire universe, with extended sources thought to be stars in their early stages of formation. However, [Hubble \(1925\)](#) used a Cepheid variable to find the distance to the Andromeda galaxy and concluded that Andromeda existed outside the Milky Way galaxy building on Shapley's findings on the Milky Way galaxy's diameter.

Since then, a huge number of galaxies were observed and were later broadly classified into elliptical galaxies and spiral galaxies. Elliptical galaxies were observed to have old stars and less interstellar gas, and dust. They also have a higher velocity dispersion ratio to rotation velocity whereas spirals have a disk and a central bulge. The disk serves as the potential star formation region, while the bulge consists of older stars.

Rotation velocity studies of spiral galaxies have been made, an example being [Freeman \(1970\)](#). [Freeman \(1970\)](#) compared the neutral hydrogen (H I) observations of NGC 300 and M 33 disk galaxies with the rotation curves expected after applying Newton's laws to the visible component of a galaxy. He noticed that these rotation curves did not match but were higher towards the edge than the model predicted given only the visible light. With this, he concluded that there must be an additional mass that remained undetected, at least as massive as the detected galaxy, and with a distribution that was different from the observed exponential distribution of the optical galaxy.

These observations contributed to today's commonly accepted picture of our universe that everything we see consists of about 5% whereas about 95% of the universe remains hidden, and this can also be split into 68% dark energy and 27% dark matter.

Two forms of dark matter may contribute to the missing mass ([Schneider 2006](#)). Astrophysical dark matter consists of massive, compact objects such as brown dwarfs, white dwarfs and black holes whereas particle dark matter consists of elementary particles that have not been detected yet.

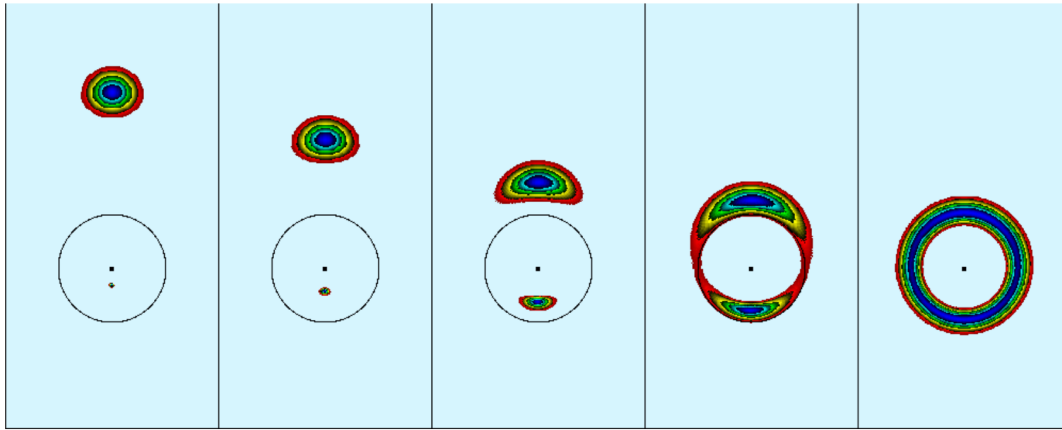


FIGURE 1.1: This is an image of a circular source shown in color, that is lensed by a gravitationally compact object. (Image by [Wambsganss & Wambsganss 1998](#))

Astrophysical dark matter, due to its compact nature and mass, is expected to gravitationally lens objects observed behind them, as shown in Fig 1.1. Using this effect, the Massive Compact Halo (MACHO) and the Extremely Red Objects (EROS) surveys in 1990 began to search for them in the direction of the Magellanic clouds, whereas the Optical Galactic Lensing Experiment (OGLE) surveyed in the direction of the galactic bulge.

After 1993, the MACHO survey found 20 microlensing events in the direction of the Magellanic clouds, and some tens of thousands in the direction of the galactic bulge ([Alcock et al. 1995a](#)). These results could be explained as resulting from the bar at the centre of the Milky Way. Statistics of the lensing effects towards the Magellanic clouds could be explained if 20% of the mass of the dark matter halo consisted of MACHOs with a characteristic mass of about 0.5 solar masses.

The EROS survey used an observation technique slightly different from that of MACHO, and as a result, they observed fewer microlensing events ([Beaulieu et al. 1998](#)). Neither EROS nor OGLE was able to reproduce the high count of microlensing effects detected by MACHO, and OGLE ([Alcock et al. 1995b](#); [Udalski et al. 1994](#)) put an upper bound to massive compact objects in the halo at 2% of its total mass.

The MACHOS, EROS and OGLE surveys ruled out astronomical objects as candidates for dark matter within the Milky Way halo.

Dark matter by definition does not interact with the surrounding material electromagnetically which makes it difficult to observe directly. Therefore, in considering particle dark matter as an alternative hypothesis, the particle candidate should be electrically neutral, highly stable or with a lifetime longer than the age of the universe, and contain mass to interact with the surrounding medium as observed gravitationally. Neutrons and neutrinos are neutral particles in the standard model. However, neutrons are baryonic subatomic particles that are unstable, and their number density is well constrained by the Big Bang Nucleosynthesis (BBN). If neutrinos constituted dark matter, they would be hot dark matter, leading to a different large-scale structure in the universe that would be different from what we currently observe ([Schneider 2006](#)).

Since no particle could describe dark matter sufficiently, there was a need to search for a particle unknown to the standard model, which is electrically neutral, has a finite mass and an extremely long lifetime and interacts only weakly. Such Weakly Interacting Massive Particles (WIMPS), are the most promising dark matter candidates and are ten to a hundred times more massive than a proton ([Schneider 2006](#)).

However, the Modification of the Newtonian Dynamics (MOND) serves as an alternative to the Dark Matter model that has not been excluded as a theory yet ([Milgrom 1983](#)). MOND explains the observed rotation velocities in the outer regions of galaxies. It considers gravitational force on material in the outer regions to be proportional to the square of the materials centripetal acceleration, or varies inversely with radius.

Continued observations of motions of galaxies are required to trace the distribution of dark matter in- and around galaxies.

### 1.1.1 Galactic neutral atomic hydrogen H I

Galaxy motion manifests itself in the form of rotation, dispersion of material within the galaxy, as well as the movement of the galaxy relative to earth. These forms of motion can be influenced by dark matter and thereby aid in determining the distribution of dark matter within galaxies. In particular, a galaxy's rotation velocity may be influenced by the dark matter halo encompassing it ([Begeman 1989](#)). To study this halo, astronomers need a tracer of galaxy motion that extends far into the outskirts of such a galaxy. Galactic neutral atomic hydrogen (H I) is known to extend beyond the optical radius of galaxies and therefore, it is used as a tracer for galaxy rotational motion and gives more details on its distribution in the outskirts (see [Fig 1.2](#)).

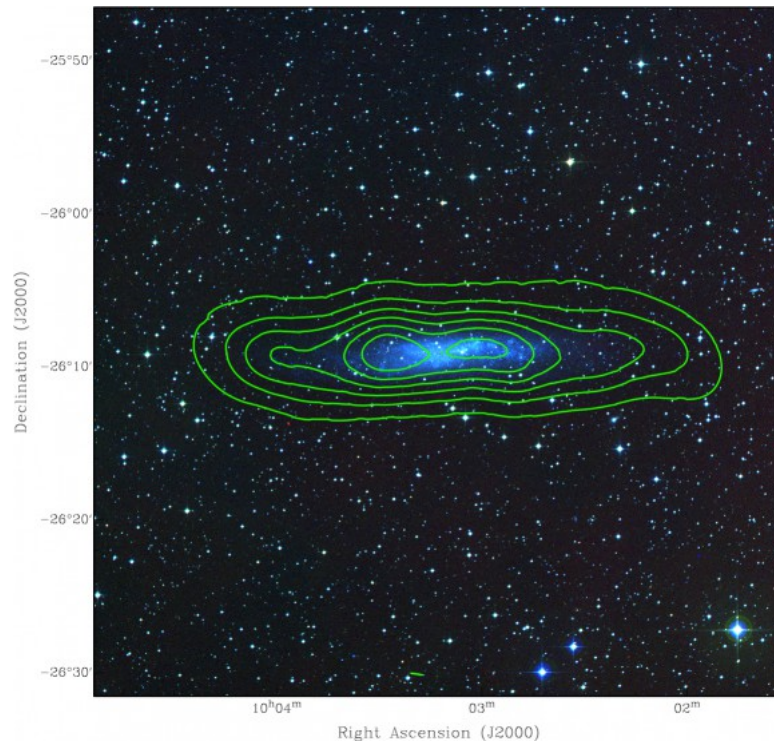


FIGURE 1.2: Three-color image of NGC 3109 from the Digital Sky Survey overlaid by green contour representing H I. It can be seen that the H I extends beyond the size of the optically visible galaxy several times. (Image by SKA/KAT-7)

Hydrogen forms about 90 % of the observable universe, followed by helium at about 9 % and the remainder constituting of heavier trace elements. In the cold interstellar medium where high energy photons barely reach in sufficient intensities and densities are low, hydrogen atoms exist in a neutral state, and since few photons would excite these atoms, one usually finds them at the ground state. However, the 1s ground state can be split into two hyperfine levels, parallel and anti-parallel spin between the electron and nucleus of the atom (see Fig 1.3). This occurs as a result of the electromagnetic interaction between an electron's and proton's spins.

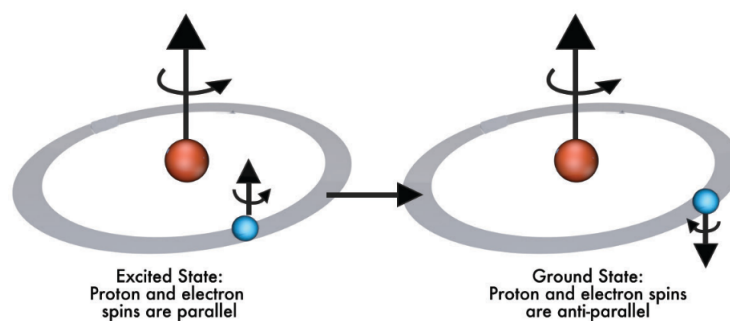


FIGURE 1.3: Hydrogen spin-flip transition occurs with a low probability of 50% within 11 million years resulting in a change in the energy state of the hydrogen atom (Image by Lockman 2017)

The parallel spin state occurs at a higher energy state as compared to the anti-parallel spin state, providing an opportunity for higher energy parallel state atoms to transition back to a lower energy state through a spin-flip transition. This transition would correspond to an energy difference of  $5.9 \times 10^{-6}$  eV, which would be released in the form of radiation corresponding to the rest frequency of about 21.106 cm or in terms of wavelength, that is 1420.4057 MHz. Such transitions occur at a very low probability with the radiative half-life  $\tau_{1/2}$  given as

$$\tau_{1/2} = A_{10}^{-1} \approx 3.5 \times 10^{14} \text{ s} \approx 11 \text{ million years} \quad (1.1)$$

and has an Einstein emission coefficient of  $A_{10} \approx 2.85 \times 10^{-15} \text{ s}^{-1}$  (Furlanetto et al. 2006).

This implies that observing such 21 cm line emission should be incredibly rare. However, hydrogen is by far the most abundant element in the universe. The number of hydrogen atoms observed along the line of sight in the direction of high concentrations of neutral hydrogen such as a galaxy, also called the column density of neutral atomic hydrogen, is in sufficient quantities for there to be atoms at the excited state and consequently, transitioning frequent enough for them to be detected on earth. This column density  $N_{\text{H}}$  ( $\text{cm}^{-2}$ ) can be quantified as

$$N_{\text{H}} \equiv \int_{\text{los}} \eta_{\text{H}}(S) ds, \quad (1.2)$$

where  $\eta_{\text{H}}(S)$  is the number of hydrogen atoms per  $\text{cm}^3$  and  $N_{\text{H}}$  can be determined by integrating along the direction of the line of sight  $s$ .

Opacity  $\tau$  is a quantity that enables us to determine what fraction of the emitted radiation makes it to the observer without obstructions. This quantity is related to the column density of an isothermal H I cloud by

$$N_{\text{H}} = 1.823 \times 10^{18} \int_{\nu} T_{\text{s}} \tau(\nu) d\nu, \quad (1.3)$$

where the spin temperature  $T_{\text{s}}$  is assumed to be constant along the line of sight and the opacity  $\tau(\nu)$  is integrated along the velocity axis instead of frequency axis (see Verschuur et al. 1974).

The opacity  $\tau$  is also related to brightness temperature  $T_{\text{B}}$  by (Verschuur et al. 1974)

$$T_{\text{B}\nu} = T_{\text{s}}(1 - \exp^{-\tau_{\nu}}). \quad (1.4)$$

And by substituting Equation 1.4 into Equation 1.3 gives

$$N_{\text{H}} = -(1.823 \times 10^{18}) T_{\text{s}} \int_{\text{v}} \ln \left[ 1 - \frac{T_{\text{B}}}{T_{\text{S}}} \right] \text{d}v. \quad (1.5)$$

If the cloud is optically thick, then  $T_{\text{B}} = T_{\text{S}}$  making the logarithm term in Equation 1.5 to be undefined. In this situation, the column density of the cloud can not be obtained as only the front of the cloud is visible. However, if the cloud is relatively transparent to the 21 cm radiation such that its opacity  $\tau \ll 1$ , then Equation 1.5 evaluates to Equation 1.6 (Irwin 2007).

$$\left( \frac{N_{\text{H}}}{\text{cm}^{-2}} \right) \approx 1.82 \times 10^{18} \int \left[ \frac{T_{\text{b}}(v)}{\text{K}} \right] \text{d} \left( \frac{v}{\text{km s}^{-1}} \right) \quad (1.6)$$

The spin temperature  $T_{\text{s}}$  used in Equation 1.3 can be described by the relative populations between the excited parallel state and the de-excited anti-parallel state,  $n_1$  and  $n_0$  respectively, using the Boltzmann distribution

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} \exp \left( \frac{E_{21\text{cm}}}{k_{\text{b}} T_{\text{s}}} \right), \quad (1.7)$$

where  $g_1 = 3$  and  $g_0 = 1$  are statistical weights forming  $g_1/g_0 = 3$ ,  $E_{21\text{cm}}$  the difference in energy between the two states  $E = 5.9 \times 10^{-6}$  eV,  $T_{\text{s}}$  the spin temperature and  $K_{\text{b}}$  the Boltzmann constant.

### 1.1.2 Structure and mass of H I in galaxies

When a galaxy is optically thin, all 21 cm radiation emitted by the H I mass, irrespective of the depth of H I, can make it to an observer without being obscured. In such a case, integrating this total emission gives the amount of H I mass in the galaxy (see J. J. Condon & S. M. Ransom 2016)

$$\left( \frac{M_{\text{H}}}{M_{\odot}} \right) \approx 2.36 \times 10^5 \left( \frac{D}{\text{Mpc}} \right)^2 \int \left[ \frac{S(v)}{\text{Jy}} \right] \left( \frac{\text{d}v}{\text{km s}^{-1}} \right), \quad (1.8)$$

where  $M_{\text{H}}$  is the total H I mass relative to the mass of the sun  $M_{\odot}$ ,  $D$  is the distance to the H I mass in megaparsec (Mpc),  $\int S(v) \text{d}v$  is the total flux<sup>1</sup> integrated over velocity. Assuming that the distribution of mass within a galaxy were spherically symmetric, the gravitational pull at any point  $r$  from the centre of the galaxy would be as a result of the enclosed mass  $M(r)$ , therefore,

---

<sup>1</sup>flux is discussed in Subsection 1.2.1

$$GM/r^2 = v^2/r \quad (1.9)$$

$$v^2 = GM/r \quad (1.10)$$

where  $v = v_r/\sin(i)$  is the tangential velocity at a particular radius  $r$ , and  $i$  is the inclination of the galaxy disk which can also be estimated from  $\cos(i) = b/a$  where  $b$  is the minor axis and  $a$  the major axis of the galaxy. This approximation is only valid when a galaxy is assumed to be an infinitely thin circular disk. From this, the total mass of a galaxy dominated by a spherical mass distribution becomes Equation 1.11 (J. J. Condon & S. M. Ransom 2016)

$$\left(\frac{M}{M_\odot}\right) \approx 2.3 \times 10^5 \left(\frac{v}{\text{km s}^{-1}}\right)^2 \left(\frac{r}{\text{kpc}}\right) \approx 2.3 \times 10^5 \left[\frac{(v_r/\sin(i))}{\text{km s}^{-1}}\right]^2 \left(\frac{r}{\text{kpc}}\right). \quad (1.11)$$

However, this mass may be structured within a galaxy in such a way that it is observed to form a warp in the H I distribution of NGC 5907 as has been shown by R. Sancisi (1976). Rogstad et al. (1974) also showed that galaxies typically have a warp starting at  $r_{25}$  from the galactic centre and can be approximated by a straight line at  $r_{25} \leq r \leq r_{26.5}$  as shown in Fig 1.4. Beyond this point, the neutral hydrogen is usually loosely wound and bound to the galaxy. This warp in galaxies originates from the interaction between the H I content of the galaxy that is loosely attached, and its environment (Józsa et al. 2007b; Briggs 1990). The Milky Way itself has been observed to possess such warps in its H I content (KERR et al. 1957).

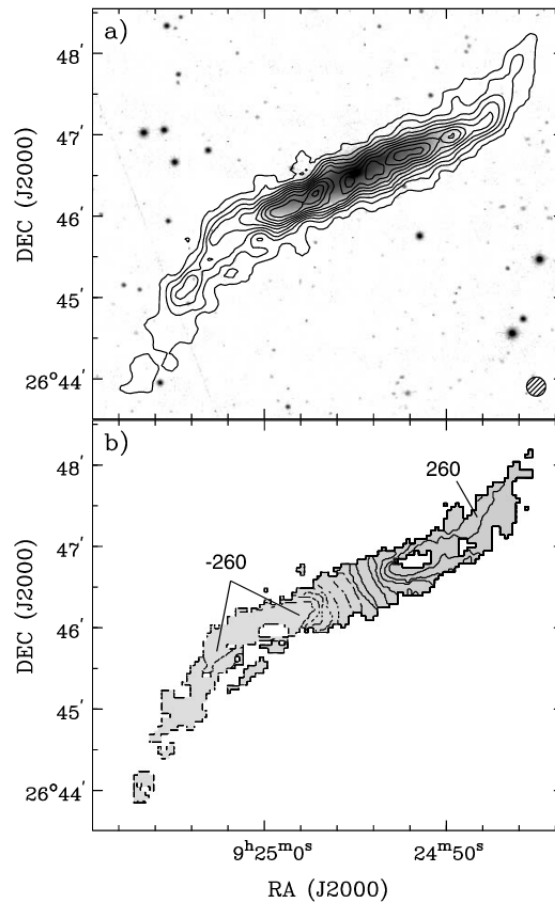


FIGURE 1.4: H I for Galaxy NGC 2862 plotted as a contour overplotted over I-band image. The contours, hence H I, shows a warp in the galaxy disk (Image from [Spekkens & Giovanelli 2006](#))



The vertical profiles of H I disks in galaxies can be modelled as a Gaussian with full width half maximum values constrained at 500 to 700 pc (Peters et al. 2017) where more massive galaxies have a thinner vertical profile than less massive ones, therefore serving as a measure of a galaxy’s gravitational potential. A galaxy’s gravitational potential seeks to constraint H I within its disk closer to the central plane of the galaxy. This mass of neutral hydrogen increases in thickness the further out we get from the centre of the galaxy as self-gravity decreases (Olling 1996).

H I can also act as a tracer to environmental effects such as gravitational tidal influences in and around galaxies (Yun 1999) thereby providing an opportunity to learn about the galaxy’s interaction history with its environment and learn about its evolution. The outer layers of H I in a galaxy may be loosely bound to the host galaxy, resulting in tidal interactions disrupting their structure as shown in Fig 1.5.

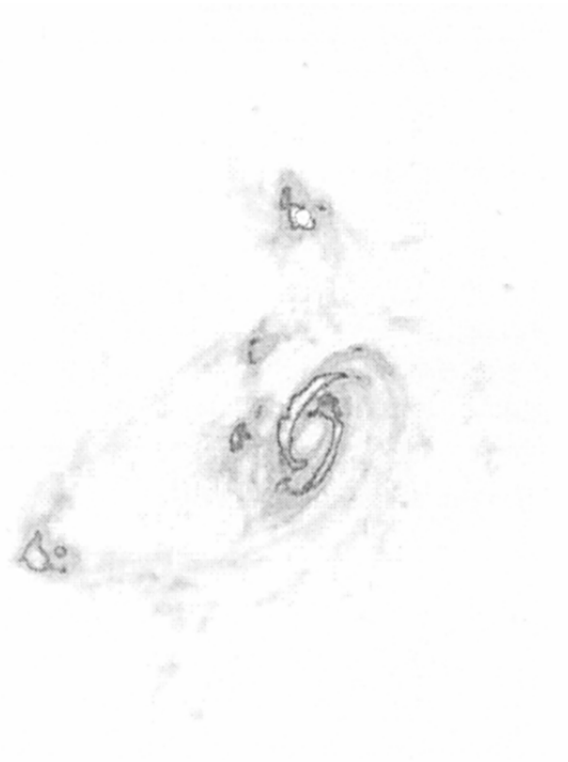


FIGURE 1.5: Gravitational tidal influence on H I content of M 81 (Yun 1999)

## 1.2 Fundamentals of radio observations

### 1.2.1 Source intensity

In astronomy, we strive to measure signals emitted by a source in the sky. The source intensity is thus defined as the amount of energy coming from a unit solid angle and passing through a given unit area, which is perpendicular to the direction of radiation,

per unit time.

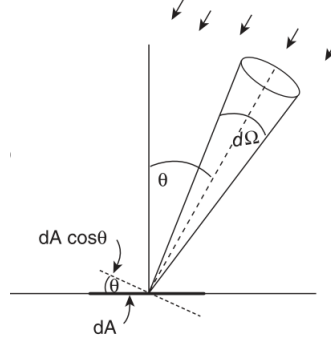


FIGURE 1.6: In the figure above,  $dA$  is the component of area whose perpendicular is towards the zenith,  $d\Omega$  represents the extent of the source being observed in terms of solid angle,  $\theta$  the direction towards the source. The arrows show radiation coming from the source. (Image by [Irwin 2007](#))

A receiving element or telescope will only receive an amount of radiation proportional to its collecting area. Therefore, this makes the specific intensity a quantity of interest, which is defined as the intensity per unit area. We denote specific intensity as  $I_\nu$ . It is measured in units of  $\text{W m}^2 \text{Hz}^{-1} \text{sr}^{-1}$ .

If we consider a source subtending a solid angle  $d\Omega$  in the direction  $\theta$ , radiating at a bandwidth  $d\nu$  per unit time  $dt$  onto a receiving element area  $dA$ , then the energy received is given by Fig 1.6 (see [Irwin 2007](#))

$$dE = I_\nu \cos(\theta) d\nu d\Omega dA dt, \quad (1.12)$$

where  $\theta$  is the direction towards the source.

A detector pointed towards the source direction  $\theta$  that extends a solid angle  $\Omega$  would measure flux given by (see [Irwin 2007](#))

$$S_\nu = \int_{\Omega_s} I_\nu \cos(\theta) d\Omega. \quad (1.13)$$

Therefore, the flux density  $S_\nu$  and a galaxy's specific intensity  $I_\nu$  are related by Equation 1.14 (see [Longair 2011](#))

$$I_\nu = \frac{S_\nu}{\Delta\Omega} \quad (1.14)$$

However, extraterrestrial sources usually have very weak flux densities leading to the introduction of another unit used in radio Astronomy called the Jansky (Jy), named after Karl Jansky ([Jansky 1933](#)) who was the first to detect extraterrestrial radio signals, where

$$1 \text{ Jy} = 10^{-26} \text{ W m}^{-2} \text{ Hz}^{-1}. \quad (1.15)$$

In the case of thermal radiation, the source-specific intensity,  $I_\nu$ , and the brightness temperature  $T_b$  are related through the Plank's formula

$$I_\nu = \frac{2h\nu^3}{c^2} \frac{1}{\exp(\frac{h\nu}{kT_b}) - 1}, \quad (1.16)$$

where  $h$  is the Plank's constant,  $k$  the Boltzmann's constant,  $\nu$  the frequency and  $c$  the speed of light. When the Rayleigh-Jeans relation  $h\nu \ll kT_b$  holds, Equation 1.16 becomes

$$I_\nu = \frac{2kT_b}{\lambda^2} \implies T_b = \frac{I_\nu \lambda^2}{2k}, \quad (1.17)$$

where  $\lambda$  is the wavelength of the radiation and  $T_b$ , the true temperature of a source (Longair 2011).

### 1.2.2 Doppler shift effect

Given a source S, such as a mass of neutral hydrogen gas, moving towards point A emitting radiation at wavelength  $\lambda_0$ , an observer at A observes this radiation at a shorter wavelength  $\lambda_1$ , corresponding to a higher frequency. An observer at B observes this same radiation at a longer wavelength  $\lambda_2$  which corresponds to a lower frequency (see Fig 1.7). This shift in wavelength and frequency depending on the relative motion of the observer and source is referred to as Doppler shift.

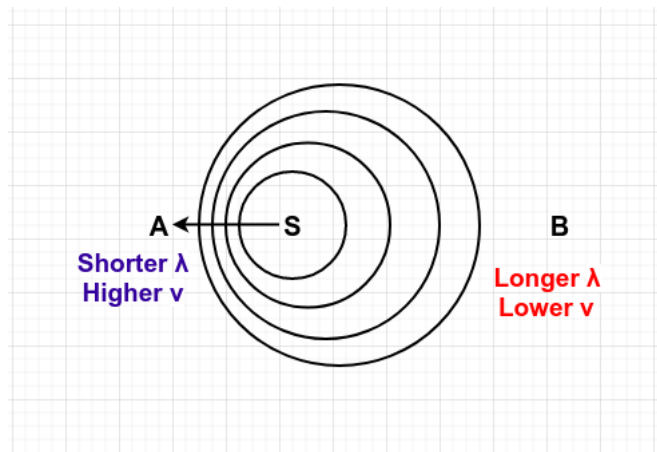


FIGURE 1.7: Motion of object towards the left results in the frequency of radiation at the front of the object increasing while reducing at the back of the object.

For a source moving radially away from or towards an observer, the relationship between the rest wavelength  $\lambda_0$  and observed wavelength  $\lambda$  is given by

$$\lambda = \lambda_0 \left( \frac{v_s}{c} + 1 \right), \quad (1.18)$$

where  $v_s$  is the systemic velocity of the source.

Once incoming radiation from such a source makes it to the ground, it can be observed by astronomers using different means, from dipole antennas to parabolic dishes and interferometers.

### 1.2.3 Single dish antennas

A single dish antenna is an antenna consisting of a parabolic reflector and a receiver placed at the focus of the telescope. The receiver is usually placed either in front of the parabolic dish, at the primary focus, or the base of the dish as shown in Fig 1.9

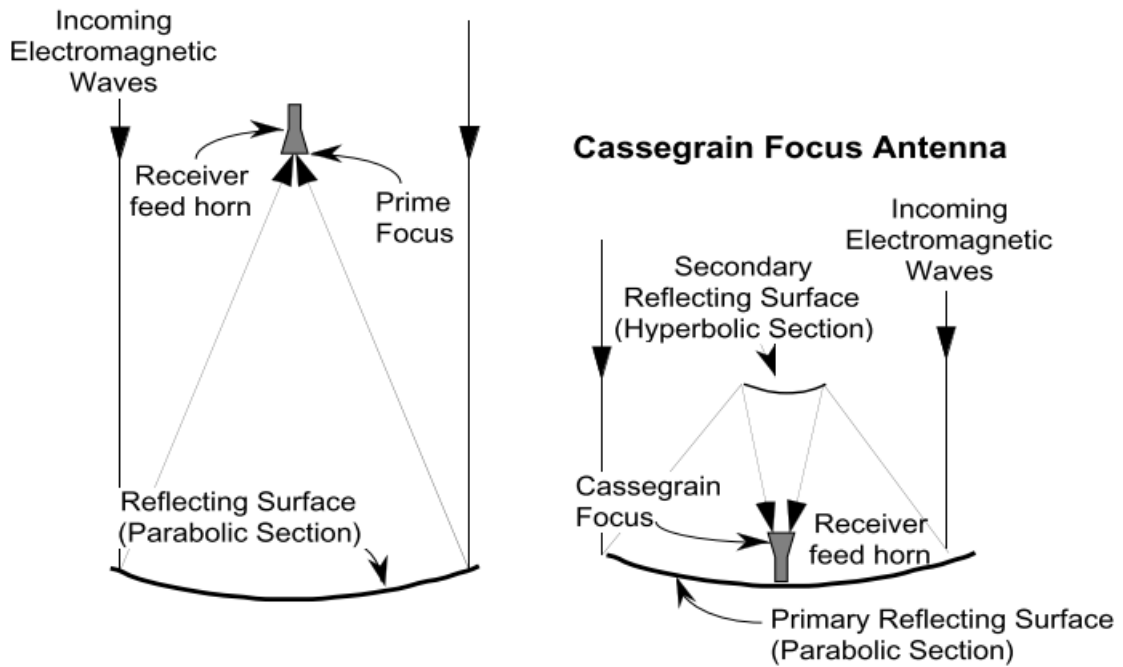


FIGURE 1.8: Left: A prime focus antenna where the receiver is relatively light. Right: A cassegrain focus antenna with the receiver at the base of the dish due to weight of the receiver. (Image by [Miller 1998](#))

Within the receiver, the signal is captured, converted into current, amplified and mixed to a frequency convenient for further signal processing. In the end, power is measured which is the voltage squared. Since signals coming from extraterrestrial sources usually cover a broad range of frequencies, the detector either averages over a

range of frequencies or gives a measurement of power for specific frequencies depending on its type.

Antennas, especially parabolic antennas, are not equally sensitive to radiation coming from all directions but rather have a primary beam pattern which describes its sensitivity per direction (see Fig 1.9).

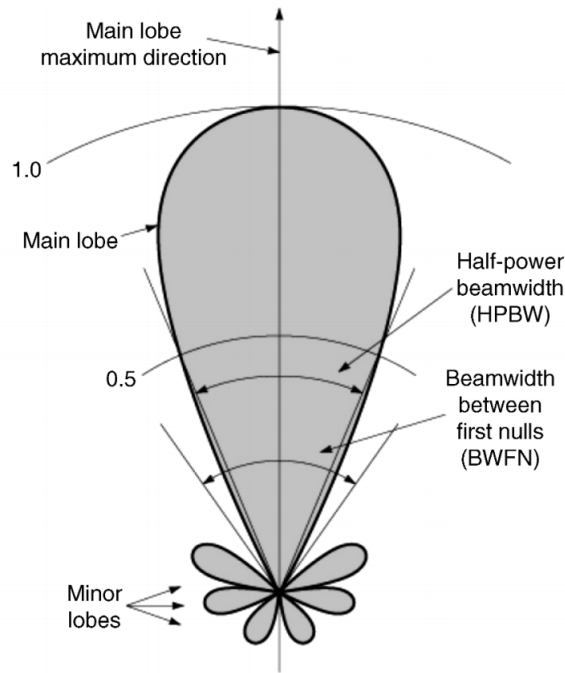


FIGURE 1.9: An illustration of an antenna's primary beam pattern with sidelobes. (Image by [Chryssomallis & Christodoulou 2015](#))

The primary beam pattern usually has sidelobes which are responsible for ground radiation leaking into the receiver contributing to data corruption. The main beam can be characterised by its Half-power beamwidth (HPBW) which is a measure of the field of view of the telescope. It is defined as the angular distance between two opposite points in an antenna's beam where the power is half the maximum value (see Fig 1.9).

The HPBW of a single-dish telescope is also measure of its resolution. Another way to determine a telescope's resolution is the smallest angular scale at which the telescope can distinguish structures. For a single dish antenna with diameter,  $D$  observing at wavelength  $\lambda$ , its resolution  $\theta'$  is given as ([Irwin 2007](#))

$$\theta' \approx \frac{\lambda}{D}. \quad (1.19)$$

Equation 1.19 implies that optical astronomy enjoys higher resolution compared to

radio astronomy for the same observing aperture diameter, due to the relatively small wavelength of the visible light of about 380 nm to 740 nm compared to radio wavelength of about 1 mm to 30 m.

As an example, a single dish telescope with a diameter of 100 m and observing the 21 cm radiation can resolve sources that are larger than 7 arcminute while an optical telescope with an observing aperture of 100 m and observing incoming light of about 500 nm can resolve sources that are larger than about 0.001 arcseconds.

Nevertheless, single-dish telescopes are still useful for studying global H I properties as they are sensitive instruments but they poorly resolve details in H I hence are not preferred for kinematic studies of all but the closest galaxies.

#### 1.2.4 Interferometers

To achieve high resolution, an alternative means to single-dish telescopes quickly gained popularity from 1946 called interferometry. This method involves using multiple telescopes that are spaced out at different positions with different telescope pairs made to work as single units called interferometers (Thompson et al. 1986). Interferometers have a point-source response determined by combining multiple baseline pairs called a synthesized beam.

The output measurements from an interferometer are called visibilities. Visibilities are represented in a frame called the UV frame defined from antenna position and with  $(u, v, w)$  coordinates. The  $u$  and  $v$  axes are in the antennas plane while the  $w$  axis points towards the direction of the source from the antennas plane. The sky is represented by  $(l, m, n = \sqrt{1 - l^2 - m^2})$  coordinates where  $l, m$  and  $n$  point in the direction of the  $u, v$  and  $w$  unit vectors.

The visibilities are related to the sky brightness by the Van Cittert-Zernike theorem (see Thompson et al. 1986) as shown in the equation 1.20.

$$V(u, v, w) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(l, m) I(l, m) \exp^{-2\pi i [ul + vm + w(\sqrt{1 - l^2 - m^2} - 1)]} \frac{dl dm}{\sqrt{1 - l^2 - m^2}} \quad (1.20)$$

Where  $V(u, v, w)$  represents the measured visibilities in the co-ordinate  $(u, v, w)$ ,  $A(l, m)$  representing the antenna response in different directions and  $I(l, m)$  represents the sky brightness.

For simplicity, let's set  $A(l, m)$  to 1. Assuming a coplanar array, where measurements are made in the plane that is normal to the direction to the source, then we set  $w = 0$

and Equation 1.20 becomes

$$V(u, v, w) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(l, m) I(l, m) \exp^{-2\pi i[ul+vm]} \frac{dl dm}{\sqrt{1-l^2-m^2}}. \quad (1.21)$$

Observing a very small patch of the sky, we set  $\sqrt{1-l^2-m^2}-1 \approx 1$  which results to

$$V(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(l, m) \exp^{-2\pi i[ul+vm]} dl dm. \quad (1.22)$$

Under the above-mentioned assumptions, the resulting visibilities are related to the true sky brightness through a Fourier transform relationship. The radio brightness  $I(l, m)$  can, therefore, be recovered by taking an inverse Fourier transform of the visibilities as

$$I(l, m) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V(u, v) \exp^{2\pi i[ul+vm]} du dv. \quad (1.23)$$

### 1.3 H I Spectroscopic data cube

H I in galaxies is observed in the 21 cm signal and stored as visibilities after correlating the output voltages as discussed. These visibilities are later converted to data cubes by imaging, with the x, y and z-axis being the right ascension, declination and velocity or frequency axis as shown in Fig 1.10.

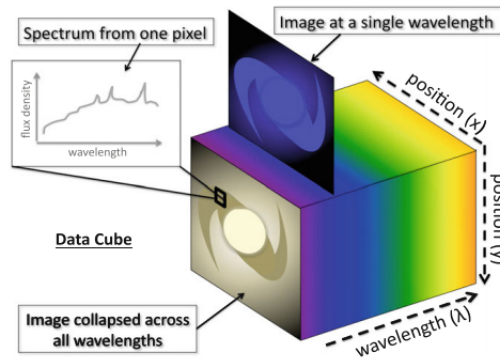


FIGURE 1.10: A spectroscopic data cube with the wavelength axis, right ascension, x and Declination, y. (Image by Harrison 2016)

In the regions covered by the galaxy, every pixel, observed in the z-axis, gives the spectral line profile or the velocity profile. This velocity profile provides information about the velocity distribution of the galaxy along one dimension.

Single dish telescopes are usually used to observe the global properties of H I as discussed. Given sufficient spectral resolution of the telescope, they can be used to determine the velocity profile of the galaxy. Interferometric observations enable the user to also spatially resolve more distant galaxies, which makes them the preferred instrument where the resolved kinematics of a galaxy is of interest.

### 1.3.1 Moment maps

Many useful quantities can be derived from spectroscopic data cubes that may give insight into the structure and kinematics of a given galaxy. The zeroth moment or moment 0 maps can be created by summing the product of the channel width  $\Delta V$  and  $S_i$ , the intensity for the  $i^{\text{th}}$  channel in the units of  $\text{Jy beam}^{-1}$ , across the velocity axis as shown in Equation 1.24

$$I_{\text{H I}} = \sum_i S_i \times \Delta V, \quad (1.24)$$

Where  $I_{\text{H I}}$  represents the value of a single pixel in a moment 0 map. Equation 1.24 is used to form the integrated H I map with units of  $\text{Jy beam}^{-1} \text{ km s}^{-1}$ .

Figure 1.11 shows this integrated H I emission map or moment map, and is useful to show the total quantity of neutral hydrogen, assuming the galaxy is optically thin to H I.

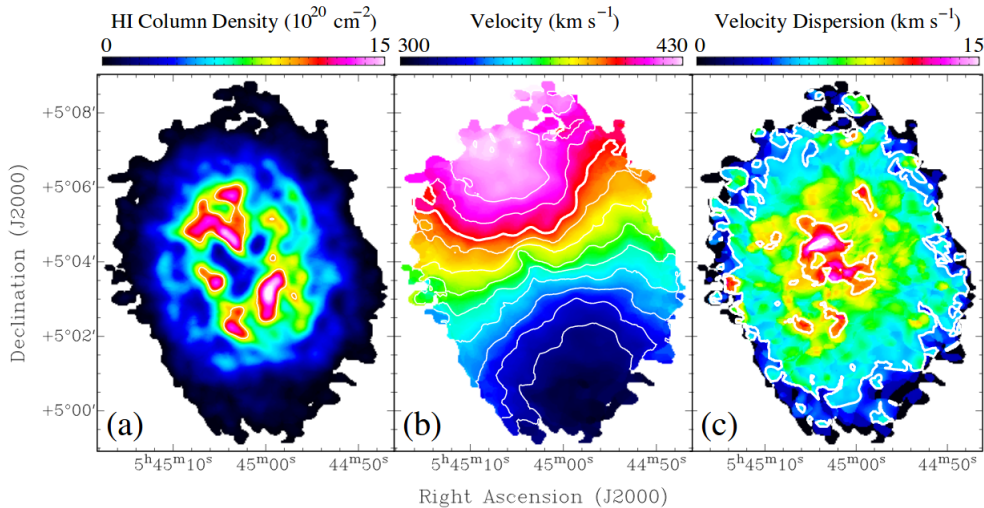


FIGURE 1.11: a) represents the total H I column density. b) represents the intensity weighted velocity c) represents the intensity weighted velocity dispersion. (Image by [Frusciante et al. 2012](#))

From the data cube, velocity fields may be extracted using a variety of techniques. One technique involves selecting the peak velocity in the spectral axis for every pixel



in the spatial coordinates. However, when the velocity profile is noisy, the selected maximum velocity may be merely an artefact of the noise and therefore, incorrect. One way around this is calculating the intensity weighted mean (IWM) of every pixel in the spatial axes and use them to make velocity fields.

The IWM is calculated as

$$\langle V \rangle = \frac{\sum_i S_i \times V_i}{\sum_i S_i}, \quad (1.25)$$

where  $\langle V \rangle$  is the intensity-weighted average velocity (see Fig 1.11 (b)).

This means of computing intensity weighted mean, however, may suffer from systematic errors (Swaters 1999). In the case where the velocity profile is skewed, the calculated IWM will be offset from the actual maximum value towards the skewed side (see Fig 1.12).

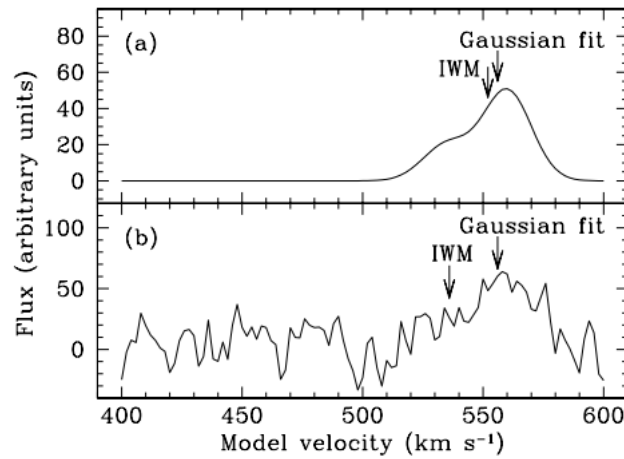


FIGURE 1.12: The intensity weighted mean, IWM, is compared to the gaussian fit of the same velocity profile in the velocity axis. In both the skewed, a), and noisy, b) velocity profiles, Gaussian fit is more accurate than IWM and determines the correct peak velocity value. (Image by Swaters 1999)

However, when a Gaussian is fitted to the same velocity profile, the gaussian's peak coincides better with the peak of the velocity profile, as compared to the Intensity Weighted Mean (IWM) calculation (Swaters 1999).

The intensity weighted mean technique may also be affected by noise in the velocity profile. When the velocity profile is ridden with noise such that the signal to noise ratio is low, IWM techniques will still derive a value, as shown in Fig 1.12 (b). However, this value may not be correct and may be far off from the actual value. Fitting

a Gaussian profile allows for a much more accurate determination of the peak value of the velocity profile.

These systematic errors appear when using IWM, especially in low signal to noise regions of the data cube. However, if the signal to noise is considerably high, and the velocity profile follows a simple profile structure, then the differences between the Gaussian fitted velocity profile and the IWM derived values become small.

A Gaussian fit works well only for certain cases and is superseded by Gauss-Hermite curve fitting in the case where the velocity profile is skewed or has elongated tails.

Moment 2 maps are used to give an impression of the total velocity dispersion of the galaxy, which can also be seen in the velocity width in the velocity profile of the galaxy. This quantity is calculated from

$$\sigma = \sqrt{\frac{\sum_i S_i \times (V_i - \langle V \rangle)^2}{\sum_i S_i}}, \quad (1.26)$$

where  $\sigma$  is the velocity dispersion, as shown in the Fig 1.11 (c).

## 1.4 Galaxy parametrisation

Some properties uniquely identify a galaxy such as brightness, orientation and kinematics. The orientation and position of a galaxy form its geometrical characteristics through which it may be classified. A galaxy observed face-on is said to have an inclination angle of zero, whereas a galaxy observed edge-on is said to have an inclination angle of 90°. Position angle is another measure that is ascribed to a galaxy. It refers to the angle between a galaxy's major axis and the North, measured counterclockwise as shown in Fig 1.13.

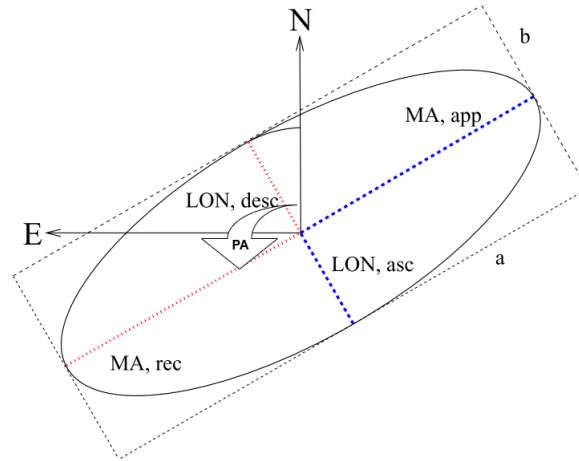


FIGURE 1.13: The figure represents Position angle as measured in a galaxy from the North counter-clockwise (Image by [Józsa et al. 2007b](#)).

#### 1.4.1 Deriving galaxy kinematic parameters

One can determine a galaxy's recession velocity from the spectrum obtained at the center of a galaxy's velocity field. Recession velocity refers to the velocity at which a galaxy is moving relative to us, the observer. [Hubble \(1929\)](#) was able to show that the universe is expanding by observing the redshift of galaxies with distance. Hence, in general, galaxies are moving away from us. However, there are exceptions to this rule. An example of such an exception is the Andromeda galaxy. The Andromeda galaxy has been observed to be on a collision course with our galaxy and should happen in about 5.86 billion years ([van der Marel et al. 2012](#)).

The spectra of a galaxy can be used to determine the velocity dispersion of the galaxy as well as the galaxy's rotation velocity. Velocity dispersion refers to the random motion that is observed in galaxies, whereas the rotation velocity refers to the tangential velocity of the material orbiting about the centre of a galaxy. The rotation velocity is usually measured from a galaxy's centre in steps to its edge, where the signal to noise ratio becomes too low that reliable measurements can no longer be made.

#### 1.4.2 Velocity measure by spectral line Doppler broadening

Ideally, an observed spectral line, such as that of H I, should have a finite width due to quantum mechanical effects. This finite line width is attributed to the quantum mechanical effect of knowing both the position and energy of an electron simultaneously. This uncertainty allows for a margin of error in determining its energy that manifests as the line width whose line profile that has a Lorentzian shape, determined by the line shape function.

However, the line width is usually more extensive than that determined by quantum mechanical limits. This line-width broadening effect is brought about by a variety of processes such as Doppler broadening as well as pressure broadening.

Doppler shift broadening happens when the individual particles that are emitting the spectral line are all moving. This motion may be either thermal, tidal flows, or may even be systematic such as rotation, contraction or expansion as well as movement related to shock-waves from events such as a supernova. The Doppler shift effect depends on the motion of each particle, resulting in different doppler shifts that collectively combine resulting in line broadening. An example of thermal line broadening is shown in Fig 1.14.

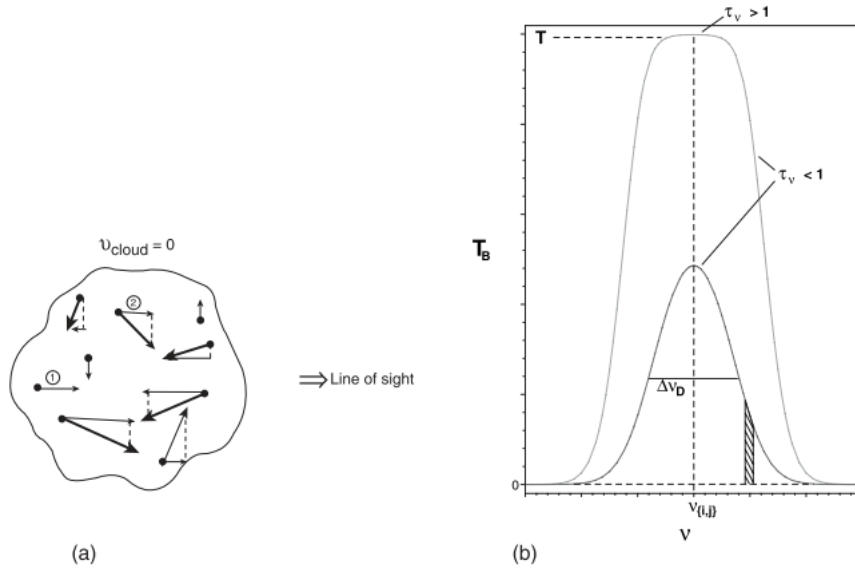


FIGURE 1.14: Thermal motion in a gas cloud at rest results in Gaussian-shaped spectral line. a) represents such a gas, where bold arrows represent velocity vectors of particles within the gas whereas thin arrows represent components of these velocity vectors, in the line of sight, that contribute to spectral line broadening. b) The marked off area represent particles with the same radial velocity, while the larger curve represents emission from an optically thick H I cloud that is denser. The smaller curve in (b) represents emission from an optically thin H I cloud. (Irwin 2007)

When a galaxy's size is comparable to the beam size such that it becomes encompassed inside the beam, the galaxy's rotation serves to increase the Doppler line broadening effect, in such a way that the difference between the approaching and receding sides of a galaxy define the width of the spectral line. The same can be said when there is an expanding shell of gas, from a supernova explosion, caught within the beam. Therefore, the expansion velocity of the galaxy can be determined by taking half the spectral line width.

The other form of broadening, pressure broadening, result from collisions happening between particles emitting the line spectra. These collisions serve to ‘interrupt’ the spectra bringing about a change in the amplitude and phase of the emitted radiation. With time, this results in a broader frequency response hence this effect is called collision broadening.

### 1.4.3 The tilted-ring-model

Wide-field H I surveys such as WALLABY (Duffy et al. 2012) will provide observations of a large number of resolved galaxies. These can be used to derive kinematic models in the hope to constrain the evolution of both dark and baryonic matter distribution in and around galaxies as well as to study environmental effects. On H I scales, the orbits in most disk galaxies are circular, but they also exhibit long-lived warps at large radii where the restoring gravitational forces of the inner disk become weak. These warps make the tilted-ring model an ideal choice for galaxy parametrisation.

The tilted-ring model was first introduced by Rogstad et al. (1974) to explain the peculiar velocity fields observed in Messier 83 by simulating its H I structure using a model consisting of concentric, circular rings which are mutually inclined and rotating at different velocities. By implementing this model, Rogstad et al. (1974) were able to show that M83 most likely has a warped disk. Since then, the tilted-ring model has become a standard tool to parametrise the H I disk of galaxies.

The tilted-ring model is based on the assumption that matter is distributed in a disk, and the majority motion is rotation. It also assumes that the rotation velocity varies only with radius. These assumptions approximately hold for spiral galaxies without a bar at their centre in contrast to stars in elliptical galaxies where the majority motion is random velocity dispersion and not confined to a disk.

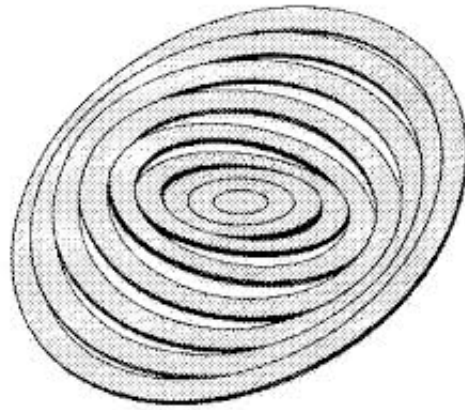


FIGURE 1.15: The tilted-ring-model as a set of concentric rings rotating at different rotation velocities and inclination angles (Image by [Rogstad et al. 1974](#))

There are two approaches in making a tilted-ring model, the two- and three-dimensional approach. The first approach involves forming a velocity field, and later fitting a tilted-ring to this kinematic map as is being done in 2DBAT. The second approach entails making a parameterised data-cube model, which is a simulation of reality, and then comparing this directly to the observations. The second approach forms the basis of three-dimensional fitting techniques, an example being TiRiFiC.

Typically, most tilted-ring algorithms only search for a local minimum, and full statistical error estimation is not done. The Fully automated TiRiFiC (FAT), as an automation wrapper for TiRiFiC, uses the deviation from a smoothed model, while in 2DBat the attempt being made is to use Bayesian statistics to derive errors, but this suffers from the fact that the errors of points in the velocity fields are ill-determined.

There is, therefore, need for the three-dimensional based fitting algorithm, TiRiFiC, to be able to find the correlation between parameter values, estimate errors in parameter values as well as find the global minimum in the Chi-Squared value. These requirements form the basis of this thesis as it documents an attempt to meet them by using a Bayesian inference engine, Multinest.

#### 1.4.4 Two dimensional fitting

Using the methods discussed in 1.3.1, a velocity field map can be made. Each point within the velocity map can be represented as the line of sight velocity  $V_{\text{los}}(x, y)$  at each point  $(x, y)$  in a rectangular coordinate system. The line of sight velocity represents the velocity an observer observes in the line of sight towards the direction of a galaxy. This line of sight velocity can be decomposed into harmonic terms, and the first approximation which ignores any non-rotational motion besides the recessional

velocity is described as

$$V_{\text{los}}(x, y) = V_{\text{sys}} + V_{\text{rot}}(r) \cos(\theta) \sin(i) \quad (1.27)$$

while including expansion velocity,  $V_{\text{exp}}$ , leads to

$$V_{\text{los}}(x, y) = V_{\text{sys}} + \sin(i)(V_{\text{rot}}(r) \cos(\theta) + V_{\text{exp}}(r) \sin(\theta)) \quad (1.28)$$

where  $V_{\text{sys}}$  is systemic velocity representing the line-of-sight velocity of the centre of the galaxy,  $\theta$  represents the  $(x, y)$  pixel coordinate of the galaxy given by

$$\begin{aligned} \cos(\theta) &= \frac{-(x - x_c) \sin \phi + (y - y_c) \cos \phi}{r} \\ \sin(\theta) &= \frac{-(x - x_c) \cos \phi - (y - y_c) \sin \phi}{r \cos i} \end{aligned} \quad (1.29)$$

Equation 1.29 represents the conversion from  $(x, y)$  pixel coordinates to polar coordinates given by  $(r, \theta)$ . In polar coordinates,  $r$  is the distance from the centre,  $(x_c, y_c)$ , and  $\theta$ , the azimuthal angle measured counter-clockwise from the major axis in the disk's plane.  $\phi$  represents the position angle of the galaxy and is measured from the north to the semi-major axis in the counter-clockwise direction, and  $i$  represents the inclination angle of the galaxy.

Along the minor axis of the galaxy, the component of the rotation velocity  $V_{\text{rot}}$  along the line of sight is zero resulting to the observed velocity or the line of sight velocity  $V_{\text{los}}$  being minimal and dominated by non-circular motion (Schmidt et al. 2016), bars (Sellwood & Sanchez 2009) and warps (Józsa et al. 2007b).

From Equation 1.28 when the inclination angle, which is the angle between the galaxy disk's normal and the observer is zero, the component of rotation velocity along the line of sight goes to zero, and so does the component of the expansion velocity,  $V_{\text{exp}}$ , along the line of sight. The line-of-sight velocity,  $V_{\text{los}}$ , ends up representing the systemic velocity,  $V_{\text{sys}}$ , and the velocity dispersion width. The velocity profile of such a galaxy that is face-on looks like Gaussian with the mean representing the systemic velocity and the width representing the average velocity dispersion in the galaxy as shown in Fig 1.16

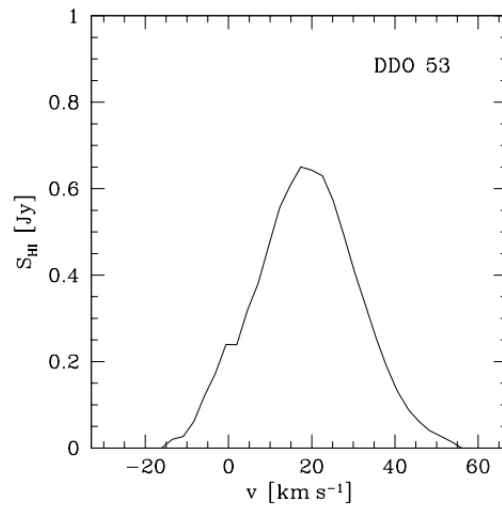


FIGURE 1.16: Global H I profile of galaxy DDO 53 as observed face on or very low inclination angle. (Image by [Walter et al. 2008b](#))

As the inclination angle increases, the contribution from rotation velocity to the measured line of sight velocity also increases, leading to the development of a double-horned velocity profile, with one horn representing the approaching velocity and the opposite horn representing the receding side (see Fig 1.17).

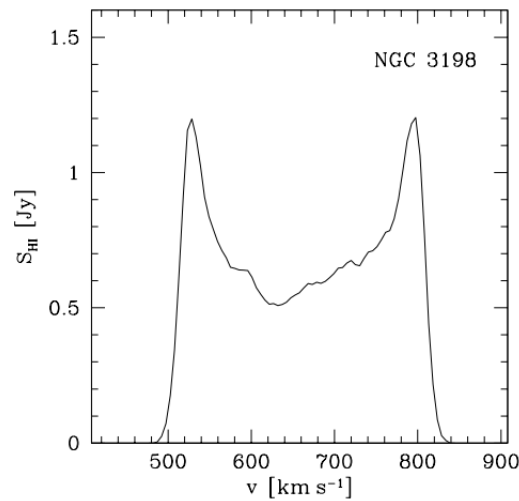


FIGURE 1.17: Global H I profile of galaxy NGC 3198 as observed edge-on or very high inclination angle. (Image by [Walter et al. 2008b](#))

However, bodies with a solid body rotation curve do not have this double-horned profile but rather a profile as shown in Fig 1.18.



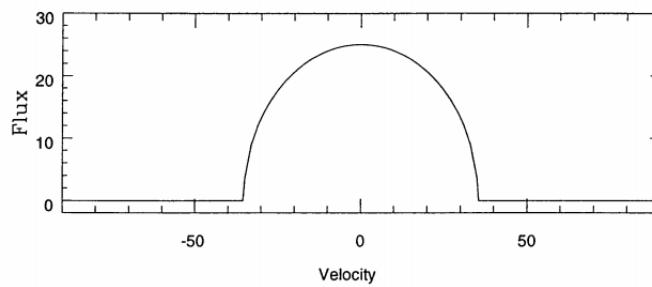


FIGURE 1.18: Demonstration of a solid body global H I profile. (Image by [Skillman 1996](#))

Once a velocity field map has been derived, a Chi-Squared fit is done with a tilted-ring model to derive both the kinematics and orientation of the model.

Two dimensional based tilted-ring fitting techniques benefit from their low computational cost while deriving kinematic parameters but at the same time, are susceptible to the effects of beam smearing (see Fig 1.19). Projection effects such as observing a galaxy that is completely edge-on also do influence the accuracy of these kinematic parameters determined. Such projection effects can be noted where the line of sight crosses the galaxy more than once like in galaxies observed to have a thick disk as well as those observed with extremely high inclination angles such that they appear near edge-on as well as outer flares.

Some publicly available software that takes advantage of this 2-dimensional fitting technique are available, and these include but not limited to ROTCUR ([Begeman 1989](#)) and 2DBAT ([Oh et al. 2018](#)).

[Di Teodoro & Fraternali \(2015\)](#) made a comparison between low-resolution H I galaxy velocity field maps, also called H I moment maps, observed with a single dish and High-resolution H I moment maps seen by an interferometer and found that 2-dimensional galaxy parametrisation approaches were not suitable for low-resolution single-dish observations due to beam smearing. Beam smearing affects the data by making the velocity gradients, towards the centre of the galaxy, a little bit flatter while at the same time, converting bits of the rotation velocity to high-velocity dispersion or larger velocity width as shown in Fig 1.19.

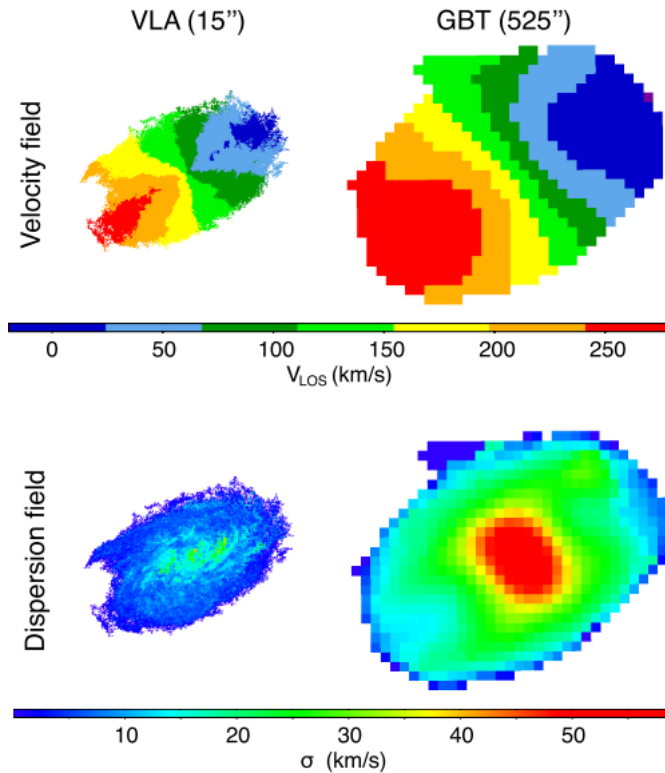


FIGURE 1.19: The figure above shows the velocity dispersion, bottom, vs velocity fields, top, for single dish telescope, Right GBT, and interferometer VLA, left. (Image by [Di Teodoro & Fraternali 2015](#))

Figure 1.19 shows that the central regions of the velocity field of the low-resolution data mimic that of solid body rotation. In comparison, the velocity dispersion of the low-resolution data increases by a factor of 3 to 4 throughout the whole disk.

This, therefore, calls out for either the use of three-dimensional techniques to analyse low-resolution single dish data and observing galaxies at high resolution using interferometers.

#### 1.4.5 Three-dimensional based fitting

H I galaxy observations made by interferometers are typically imaged into data cubes with right ascension, declination and velocity axes. Deriving parameters that describe a galaxy from such a cube directly can be done through the three-dimensional based fitting. In general, 3D fitting algorithms take in a set of supplied parameters that describe the tilted-ring model's geometric and kinematic parameters. From this, they construct a simulated data cube. This data cube is convolved with a beam, approximated as Gaussian with the major and minor axes similar to the observations, and later compared to the model using Chi-Square. A variety of fitting techniques and algorithms are then applied, such as the Golden Section algorithm in TiRiFiC ([Józsa et al. 2007b](#)), to minimise this Chi-Squared value. In the process, a new model data

cube is constructed for every iteration of the fit, where each cube consists of millions of point sources with independent position and motion parameters, making a total of 6 parameters per point source.

With the higher flexibility afforded by the method and a larger number of parameters to be fitted comes the computational cost for the fit to be complete. Unlike the two-dimensional based fitting techniques, three-dimensional fitting does not have an analytic function to fit for but instead, rely on Monte-Carlo techniques. These techniques make 3D fitting algorithms complex as well as susceptible to convergence to a local minimum in the Chi-Square plane. Another drawback to such fitting techniques is that they are not well equipped to handle inhomogeneities in the gas distribution (Józsa et al. 2007b).

Such a fitting technique has advantages over both the two dimensional and one dimensional techniques as it is less affected by both projection effects as well as beam smearing effects. The advantage comes about as a result of performing a direct Chi-Squared fit of the tilted-ring model to the spectroscopic data cube, where parameters describing the ring model are derived. Since this process does not involve the construction of velocity fields, where data compression happens resulting in a loss of information, it becomes a more reliable fitting method.

This method also leads to the possibility for modelling kinematic asymmetries, warps in galaxies even those that are nearly edge-on or face on, as well as gas residing away from the plane of the galaxy.

#### 1.4.6 Distance measure by the Tully-Fisher Relation

Cepheid variables can be used to measure the distance to a galaxy. However, these cepheid variables become too faint to be measured past distances of 20 Megaparsecs (Mpc). This makes it necessary to use a different technique to measure the distance to a given spiral galaxy.

There exist a correlation between the intrinsic luminosity of a spiral galaxy and its rotation velocity, which is known as the Tully-Fisher relation (see Fig 1.20). Using this relation, one can determine the intrinsic luminosity of a disk dominated galaxy by measuring its rotation velocity and using the apparent magnitude of the galaxy, be able to determine the distance to a given galaxy.

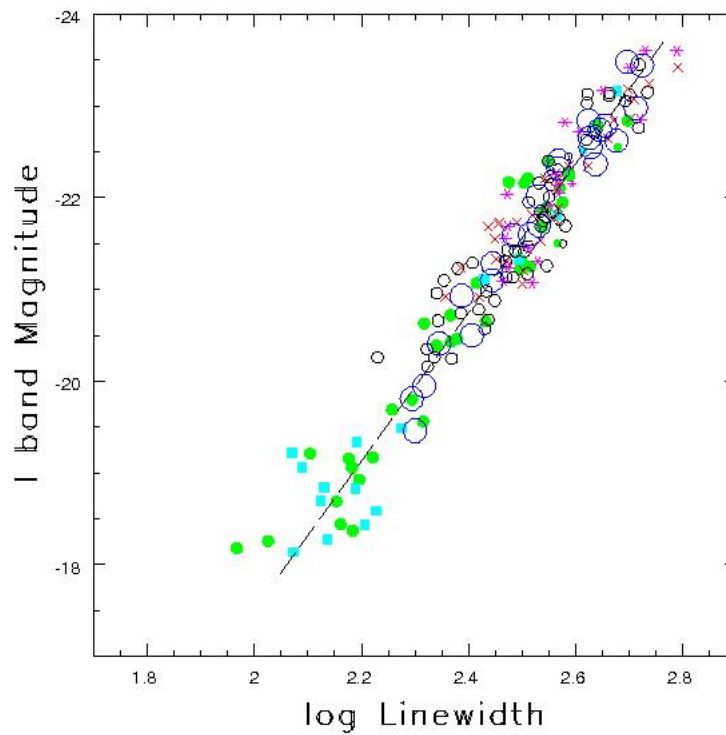


FIGURE 1.20: The Tully-Fisher relation. Galaxies from 5 separate clusters were used to make this figure, the small symbols of various shapes, as well as galaxies that were relatively close to us with independently known distances to each of them, the open symbols. The y-axis represents the absolute magnitude in the  $I^{\text{th}}$  band centred on 820 nm whereas the x-axis represents the logarithm of the H I linewidth. This linewidth is approximately twice the rotation velocity of the galaxy.

(Image by [Tully 2007](#))

Tully-fisher relation gives the relationship between the rotation velocity of a disk dominated galaxy  $V$ , and its intrinsic luminosity  $L$  as

$$L \propto V^4. \tag{1.30}$$

## Chapter 2

# Bayesian statistics

### 2.1 Probability

There are two views on probability theory. The frequentist view of probability and the Bayesian outlook. In the frequentist point of view, the probability of an event occurring depends on how frequent the event has occurred in the past out of a given number of trials. This observed frequency of occurrence may be from experimentation or observations. In contrast to this view, the Bayesian approach can be used even in situations where there is limited number of trials and hence limited information. This approach involves formulating the "degree of belief" of an event occurring based on background knowledge. The "degree of belief" can be represented as a probability distribution and event as a random variable. The Bayesian approach allows for updating this "degree of belief" or distribution with the addition of new data or information.

An important attribute of Bayesian statistics used in this thesis is its inherent ability to treat parameter values as random variables. In parametrising a given data set, the objective of Bayesian statistics is to determine the posterior distribution of the "best-fit" parameter values. However, when approaching the same parametrisation problem using the frequentist approach, the objective shifts to determining the maximum likelihood values of the fitted parameter values. Therefore, in the Bayesian point of view, parameter values are treated as random variables.

A random variable  $\theta$  is a function whose values are a result of a random phenomenon. The random phenomena, and consequently, the random variable, may result from an underlying distribution function that is referred to as the probability distribution function (pdf). The pdf of  $\theta$  is defined as a function that gives the relative likelihood of obtaining an instance of the random variable  $\theta$ .  $\theta$ , as a random variable, can either be continuous or discrete giving rise to the probability density function or probability mass function respectively.

### 2.1.1 Conditional Probability

In formulating Bayesian statistics, the first step is to take a look at conditional probability. In conditional probability,  $A$  given  $B$  is the conditional probability that  $A$  happens given that  $B$  has already happened and some given information  $I$  is true. This conditional probability is denoted by  $P(A|B, I)$ . The joint probability, or conjoined probability, of  $A$  and  $B$ , given by  $P(A, B|I)$  is the probability of  $A$  and  $B$  happening together where

$$P(A, B|I) = P(A|I) P(B|I) \quad (2.1)$$

if  $A$  and  $B$  are independent, and

$$P(A, B|I) = P(A|B, I) P(B, I) \quad (2.2)$$

if  $A$  and  $B$  are dependent.

### 2.1.2 Bayesian Statistics

To determine the probability of  $B$  alone from the joint probability  $P(A, B|I)$ , a sum over all possible outcomes of  $A$  is carried out in a process called marginalization.

$$P(B|I) = \sum_A P(A, B|I). \quad (2.3)$$

The resulting quantity  $P(B|I)$  becomes the marginal probability of  $B$  (refer to Equation 2.3). From the product rule of dependent parameters  $P(A, B|I) = P(A|B, I) P(B|I)$ , inverting the order of  $A$  and  $B$  gives

$$P(B, A|I) = P(B|A, I) P(A|I). \quad (2.4)$$

Equating  $P(A, B|I) = P(B, A|I)$  results to

$$P(A|B, I) P(B|I) = P(B|A, I) P(A|I) \quad (2.5)$$

and finally forms the Bayesian Equation 2.6

$$P(A|B, I) = \frac{P(B|A, I) P(A|I)}{P(B|I)} \quad (2.6)$$

Replacing  $A$  with parameter  $\theta$  and  $B$  with the observed data  $D$ , Equation 2.6 becomes

$$P(\theta|D, I) = \frac{P(D|\theta, I) P(\theta, I)}{P(D, I)}, \quad (2.7)$$

where  $P(\theta|D, I)$  is the posterior distribution,  $P(D|\theta, I) = L(\theta)$  the likelihood,  $P(\theta, I) = \pi(\theta)$  the prior and finally  $P(D, I)$  the evidence. Bayesian statistics enables inferences of parameter  $\theta$ , for a given data  $D$  with background knowledge  $I$  to be obtained as shown in Equation 2.7

The posterior distribution  $P(\theta|D, I)$  of  $\theta$  represents all the possible values the random variable  $\theta$  could take, given the observed data  $D$  and background information  $I$ . This random variable's initial distribution is described by the prior  $P(\theta, I)$ . The Prior also represents any background information one has on the parameter of interest  $\theta$  and can be from experience or scientific theory. New data from experimental runs are acquired and used to find the most probable values for  $\theta$ , that is, the probability of getting the parameter  $\theta$  given the observed data  $P(D|\theta, I)$  and this is called the likelihood. Taking this newly found data represented in the likelihood, and updating the prior results in the un-normalised posterior given as

$$P(\theta|D, I) \propto L(\theta|I) \pi(\theta|I). \quad (2.8)$$

With the addition of more data, the posterior eventually converges to an objective inference on the parameter  $\theta$ . However, in the case where there is not enough data, the resulting posterior may be significantly influenced by the prior. Therefore, assuming different priors would result in different posterior samples, and this leads to the conclusion that the information gained is not informative enough to completely override the prior state of knowledge.

One particularly useful property of Bayesian statistics that is applied to this thesis is the ability to utilize prior information on parameter values. The parameter values that need to be optimized are constrained to physically meaningful values before optimization starts. This is particularly useful because it makes the minimizing algorithm implemented, MultiNest (Feroz et al. 2008), spend as little time as possible in search of the most optimum parameters in the given parameter space, by avoiding the space that would otherwise be physically meaningless.

Another useful property of Bayesian statistics that is needed for this work is the fact that the parameter of interest is described as a probability distribution, that is, the posterior distribution. This posterior gives parameter samples which are needed to quantify correlations between parameters as well as determine an error estimate.

It is also worth noting that the evidence, given by Equation 2.9, is used as a normalising constant in Equation 2.7.



$$z = P(D, I) = \sum_{\theta} P(D|\theta, I) P(\theta, I) \quad (2.9)$$

The evidence is particularly useful for model selection problems where it is used to select the best model among other competing models. It automatically implements Occam's razor principle, where models that are simpler and more compact are preferred over more complex models unless the latter can explain the data better. Selecting between two models,  $I_1$  and  $I_2$  requires evaluating the ratio of their relative probability distributions, given the data  $D$  as shown in Equation 2.10

$$\frac{Pr(I_1|D)}{Pr(I_0|D)} = \frac{Pr(D|I_1)Pr(I_1)}{Pr(D|I_0)Pr(I_0)} = \frac{Z_1}{Z_0} \frac{Pr(I_1)}{Pr(I_0)}, \quad (2.10)$$

where the prior probability of the two models,  $Pr(I_1)/Pr(I_0)$  can be set to one.

Numerically deriving the evidence can be a daunting task. However, some algorithms have been developed to sample the posterior of the Bayesian equation, as well as derive the evidence for model selection purposes. Such algorithms will be discussed in the following sections.

### 2.1.3 Ellipsoidal nested sampling

In the process of sampling, a new point is drawn from a region with higher and higher likelihood hence satisfying the condition that  $L_i > L$  where  $L$  is the likelihood of the point last sampled. Sampling from the prior with this constraint faces the challenge of the acceptance rate steadily decreasing as the prior volume decreases.

Ellipsoidal nested sampling attempts to solve this problem by drawing a multidimensional ellipsoidal bound with its likelihood equal to the likelihood of the lowest point that is,  $L_{\text{bound}} = L$  (see 2.1). However, the ellipsoidal bound is usually enlarged slightly by some predefined factor to ensure it encloses the entire isolikelihood contour of the lowest point. Samples are later drawn from this ellipsoidal bound until a sample is found that satisfies the condition  $L_i > L$  as above. When this multidimensional ellipsoidal bound coincides with the exact iso-likelihood contour, the acceptance rate becomes one, meaning that all points sampled are accepted or the first point sampled is used for further sampling.

Ellipsoidal nested sampling works well for gauss-like modes with only one peak. However, it encounters a challenge when it has to evaluate multiple modes. Its efficiency consequently goes down with time as the sampling rate reduces due to the limited regions satisfying the constraint  $L_i > L$ . This dip in sampling efficiency implies that the probability of getting such a point that satisfies the likelihood constraint goes

down, and therefore the algorithm has to go through numerous likelihood evaluations.

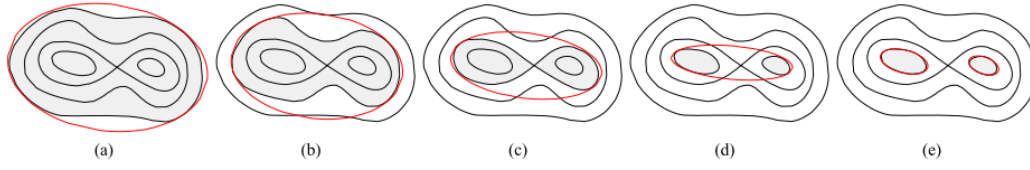


FIGURE 2.1: The ellipsoidal bound in a) ensures high efficiency due to a high sampling rate. But as the likelihood value increases, b) through d), the sampling rate decrease owing to separate modes within a single ellipsoidal bound. e) experiences an increase in the sampling rate as the two modes are separated. (Image by [Feroz et al. 2008](#))

[Shaw J R et al. \(2007\)](#) suggested a technique that would allow for the sampling efficiency to be substantially improved by identifying individual modes and constructing ellipsoidal bounds around these modes. However, this technique, among others, still faces the challenges of handling pronounced curving degeneracies. One can ‘picture’ such degeneracies as multidimensional bananas.

Additionally, Ellipsoidal nested sampling still performs poorly especially in regions where the ellipsoidal bounds overlap, resulting in lower sampling efficiency and this issue is addressed in the MultiNest algorithm.

#### 2.1.4 MultiNest algorithm background

MultiNest ([Feroz et al. 2008](#)) implements Bayesian statistics within it and aims at determining the evidence value in a model which may be used for model comparison. MultiNest also generates samples for the fitted-for parameters which may be used to reveal the underlying distribution of the sampled parameter values and perhaps be used as an error estimate.

To determine the evidence and also generate these samples, MultiNest starts as a  $D$  dimensional unit hypercube, where  $D$  is the number of parameters being fitted for where each parameter varies from 0 to 1. Transforming a point  $u$  in the hypercube to a physical parameter  $\theta$  requires that

$$\int \pi(\theta_1, \theta_2, \dots, \theta_D) d\theta_1 d\theta_2 \dots d\theta_D = \int du_1 du_2 \dots du_D. \quad (2.11)$$

However, should the prior be separable, then it is described by Equation 2.12 such that

$$\pi(\theta_1, \theta_2, \dots, \theta_D) = \pi_1(\theta_1) \pi_2(\theta_2) \dots \pi_D(\theta_D). \quad (2.12)$$

Combining equations 2.11 and 2.12 gives

$$\pi_j(\theta_j)d\theta_j = du_j \quad (2.13)$$

$$u_j = \int_{-\infty}^{\theta_j} \pi_j(\theta'_j)d\theta'_j. \quad (2.14)$$

Calculating Equation 2.14 may be computationally expensive. However, there are ways around this computational overhead. For instance, if the prior function for the parameter values is uniform, converting between the parameter space and the prior volume becomes trivial. Also, one can decide to recast the problem such that a new likelihood is defined where  $L'(\Theta) \equiv L(\Theta)\pi(\Theta)$  and the prior  $\pi(\Theta) \equiv \text{constant}$ . This method, however, may be inefficient, and it does not use the true prior in the sampling process. Whereas other problems allow for Equation 2.14 to be calculated analytically or through the use of look-up tables in an attempt to reduce the computational overhead.

Once a unit hypercube has been uniformly populated with points, an expectation maximisation scheme is applied to cluster or decompose these points into the optimal number of subclusters as described below.

Starting from a set  $S$  bounded by an ellipsoid  $E$  with a volume  $V(E)$  and containing  $N$  active points in the unit hypercube, the set  $S = \{u_1, u_2, \dots, u_N\}$ , is partitioned into  $k$  clusters,  $\{S_k\}_{k=1}^K$ , where  $K \geq 1$ . From this, the bounding ellipse  $E_k$  for these subclusters is defined by Equation 2.15

$$E_k = \{u \in \mathbb{R}^D \mid u^T(f_k C_k)^{-1}u \leq 1\}, \quad (2.15)$$

where

$$C_k = \frac{1}{n_k} \sum_{j=1}^{n_k} (u_j - \mu_k)(u_j - \mu_k)^T \quad (2.16)$$

is the covariance matrix of  $S_k$  and  $\mu_k = \sum_{j=1}^{n_k} u_j$  is the center of mass. These ellipsoids  $E_k$  enclose a subcluster whose volume  $V(E_k) \propto \sqrt{\det(f_k C_k)}$ . If we define another quantity,  $F(S)$  such that

$$F(S) \equiv \frac{1}{V(S)} \sum_{k=1}^K V(E_k), \quad (2.17)$$

minimising this quantity within the constraint  $F(S) \geq 1$  results in an optimal decomposition of the initially sampled region into  $k$  ellipsoids. After every iteration, the remaining volume of the entire prior region  $V(S)$  is updated as  $V(S) = \exp(-i/N)$ .

It is usually computationally expensive to calculate the ellipsoidal volume due to the number of eigenvector evaluations needed. Consequently, MultiNest doesn't actively partition the live points at each iteration but rather scales the ellipsoids mentioned above. However, as the number of iterations increase and the algorithm moves towards regions of higher likelihoods, the previously calculated ellipsoids become less optimal, and the active points are repartitioned again.

In essence, the MultiNest algorithm builds upon the idea of ellipsoidal sampling, with the benefit of correctly accounting for curving degeneracies while also working well for simpler cases. MultiNest identifies isolikelihood contours and attempts to enclose them with ellipsoids. It does so under the constraint that the sum of their volume exceeds that of the prior volume  $X$ .

However, these ellipsoidal bounds may not enclose the entire isolikelihood contour because the ellipsoidal approximation to a region in the prior space may be incorrect. As a result, it is usually desirable to enlarge the ellipsoidal bound, such that the region bounded by the ellipsoids is larger than that of the prior volume. Enlargening the ellipsoidal bound is achieved by setting  $X/e$  as the minimum volume, where  $X$  is the prior volume, and  $e$  is the efficiency, making  $1/e$  an enlargement factor. Setting  $e > 1$  makes the prior to be undersampled, and this may be useful for someone interested in sampling parameter space only and not interested in calculating the likelihood value.

Placing ellipsoidal bounds on isolikelihood contours leaves MultiNest with a set of ellipsoids that possibly overlap with each other, from which a live point with the lowest likelihood is drawn, removed and replaced by a new point from the enclosed volume with higher likelihood value, correctly accounting for any overlaps present. This replaced point becomes inactive but is however stored in MultiNest despite not playing any role.

An example of this ellipsoidal sampling is shown in Fig 2.2. MultiNest algorithm breaks up a posterior approximating a Gaussian into relatively few ellipsoids. In contrast to this, it breaks a mode with a pronounced curving degeneracy into a large number of overlapping ellipsoids as shown in Fig 2.2

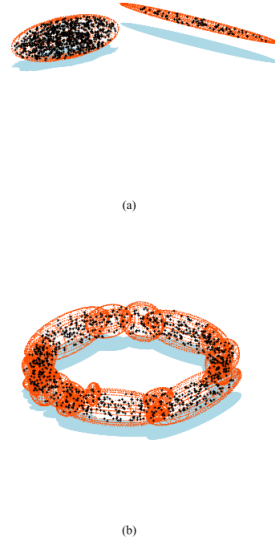


FIGURE 2.2: a) and b) illustrates how ellipsoidal decomposition is used to sample different distributions. a) has two non-intersecting ellipsoids while b) is a torus sampled by many ellipsoids. (Image by [Feroz et al. 2008](#))

Once an optimal number of ellipsoidal bounds has been recovered, at iteration  $i$ , a given ellipsoid is chosen with a given probability

$$p_k = V(E_k)/V_{\text{tot}}, \quad (2.18)$$

corresponding to the ratio of its volume to the total volume  $V_{\text{tot}}$ , where  $V_{\text{tot}} = \sum_{k=1}^K V(E_k)$ . Later, samples are selected from within the ellipsoid with a uniform probability distribution until one is found which satisfies the condition  $L > L_i$  where  $L_i$  is the lowest likelihood value of a point within the ellipsoidal bound.

For highly multimodal problems, a large number of active points are required to ensure that all the modes are detected. However, this will slow down the algorithm. To handle this, MultiNest takes advantage of the fact that the number of active points should decrease with increasing likelihood. This reduction in the number of active points occurs because the number of these multiple modes decreases with increasing likelihood, an analogue being lower hills flooded by water as the water levels rise.

It is usually necessary to identify which live point belongs to which mode and also determine the total number of modes as described in the procedure below.

At the very beginning, all active points are placed in one group, say  $G1$ . However, these points are later split into  $k$  different ellipsoids  $E_k$  and  $k$  subsets,  $S_k$  in the process

described above. From these subsets  $S_k$ , one is chosen at random, where its members become the members of a new set  $\tau$  and the associated ellipsoid becomes a member of a set of ellipsoids  $\epsilon$ . Later, the chosen ellipsoid  $E_k$  is checked for intersection with other ellipsoids. If some ellipsoids are found to have intersected with this ellipsoid, they are added to the set of ellipsoids and their corresponding members added to the new set  $\tau$ . This merger happens until all ellipsoids are evaluated for an intersection. Once this process completes, the points in the set  $\tau$  are set as the new group  $G_1$  assuming all ellipsoids are found to intersect with each other. However, we may find an ellipsoid that is isolated, and these isolated ellipsoids may intersect with each other forming altogether a different set of ellipsoids  $\epsilon$  and a different set of live points  $\tau$ . In this particular situation, two groups are formed,  $G_2$  and  $G_3$ , where  $G_1$  is ‘left behind’ as a group that does not have active points.

This group formation process occurs to separate modes as the algorithm finds points with higher and higher likelihood values, thereby splitting into more and more groups depending on the number of sub-clusters and intersections found. The groups that survive this process of splitting are later promoted to mode category. Groups with no live points would still be called modes at this time because such groups were not split. In an attempt to minimise the computational cost, there is an assumption that is applied in the algorithm, and that is the prior volume, and consequently, ellipsoidal bounds only grow smaller and smaller with increasing likelihood values. Consequently, ellipsoidal values that have been found not to intersect at any point in time are not checked for an intersection at any later stage in the iteration process. Fig 2.3 illustrates the procedure described in this Section but for a two-dimensional uniform prior distribution.

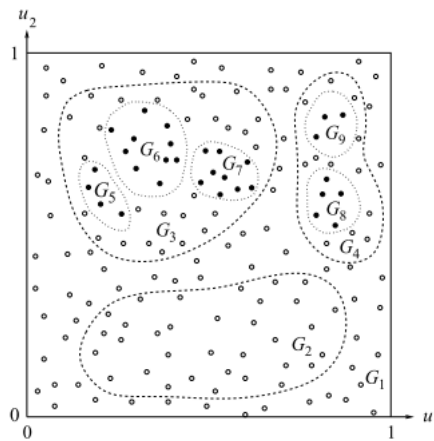


FIGURE 2.3: The dotted larger ellipsoidal bounds represent old groups, that have now been separated into newer groups, smaller ellipsoids enclosing solid dots which represent current live points (Image by [Feroz et al. 2008.](#))

In Fig 2.3, the solid dots represent active points while inactive points are represented

by open circles. At the current iteration  $I = i_2$ , group  $G_2$  becomes inactive, as it drowns at iteration  $i_1$  while group  $G_3$  splits into groups  $G_5$ ,  $G_6$ ,  $G_7$  and group  $G_4$  into  $G_8$  and  $G_9$ . If these groups make it through successive iterations, they become promoted into modes.

It was advised by [Feroz et al. \(2008\)](#) that a large number of points ensure that all the modes have points and are therefore resolved, resulting in correct evidence evaluation. The maximum efficiency parameter,  $e$ , controls the sampling volume at each iteration, which is the sum of the volume of the ellipsoids enclosing the active set. For parameter estimation,  $e$  is recommended to be set to 1 whereas, for evidence evaluation, it should be set to 0.3.

There are challenges that come with MultiNest. Firstly, the ellipsoidal approximation to the true isolikelihood contour may not be perfect as a result of the misalignment between the ellipsoidal bound and the true isolikelihood contour. This misalignment results in a portion of the ellipsoid that is outside the isolikelihood contour. While sampling a point from within such an ellipsoid, the acceptance rate becomes less than unity and points lying in the region are declined. This situation worsens exponentially as the number of dimensions grows, as a result of an increase in the ellipsoidal volume. This exponential increase in the number of likelihood evaluations with dimensions becomes a potential problem as this work attempts to fit a high number of dimensions, typically  $\geq 19$  parameters aimed for, after fixing other parameters hence minimising parameter flexibility.

Secondly, there may also be a region where ellipsoids intersect, while at the same time, the region lies inside the true isolikelihood contour. A point selected from this region will be accepted only with a certain probability  $1/n_e$  as discussed above. These two shortcomings make the sampling efficiency of points for astronomical and particle physics problems to be about 20-30 %, despite setting the enlargement factor or efficiency,  $e$ , as discussed.

## Chapter 3

# TiRiFiC-MultiNest

### 3.1 TiRiFiC

TiRiFiC ([Józsa et al. 2007a](#)) is software that directly fits galaxy models to spectroscopic data cubes without the intermediate step of constructing velocity fields, hence, a three-dimensional fitting technique. The software generates model data cubes from the tilted-ring parameterisation of a rotating disk, which is automatically adjusted to reach an optimum fit via a Chi-Square minimisation method to an observed data cube. A direct fit to the data cube is implemented to overcome fitting challenges associated with velocity fields and to ultimately provide a tool for galaxy parameterisation in large spectroscopic sky surveys.

TiRiFiC’s model has sets of parameters at user-defined radii, a ring or a set of nodes, plus a set of global parameters. A model is calculated by interpolating these rings to form sub-rings with a width specified by the user. All parameters within a sub-ring stay constant, making these sub-rings represent the actual rings in the model. Monte-Carlo integration is then used to model each sub-ring, resulting in a list of point sources depending on the parameters for the specific sub-ring. These point sources are distributed in the Cartesian phase space with equal total flux and well-determined position.

The point sources are later gridded onto a data cube identical to the input data cube. This gridding process assumes that the material through which the flux transverses is optically thin where the flux density, or intensity after gridding and convolution, along the line-of-sight is the sum of the single sources of emission. The final step involves artificially observing the gridded point source cloud by convolving it with an instrumental function.

TiRiFiC is sensitive to very faint structures because it fits many data points in a data cube simultaneously. Therefore it can be used to derive tilted-ring models significantly extending in radius beyond those obtained from a velocity field. The software is also able to parametrise H I disks of galaxies which are heavily warped and have been intersected by the line-of-sight or seen edge-on, and those whose projected centre



of rotation been shifted significantly.

### 3.1.1 TiRiFiC parameters

A set of parameters that describe the model include:

- Radius from the centre of the model, for each ring, given in units of arcsec.
- The radius dependent inclination and position angle which determines the orientation of the model once constructed by TiRiFiC, given in degrees, and allowed to vary with radius to account for warps in the galaxy observed.
- The radius dependent surface brightness distribution given in  $\text{Jy km s}^{-1} \text{ arcsec}^{-2}$ .
- The central coordinates, given as right ascension and declination and allowed to vary from ring to ring.
- The radius dependent systemic velocity in  $\text{km s}^{-1}$ .
- The radially dependent velocity dispersion in  $\text{km s}^{-1}$ .
- The constant total flux of a point source.

TiRiFiC takes these parameters and constructs a model with the procedure described in Fig 3.1, and compares this simulated model to an actual observation by Chi-Square minimisation technique.

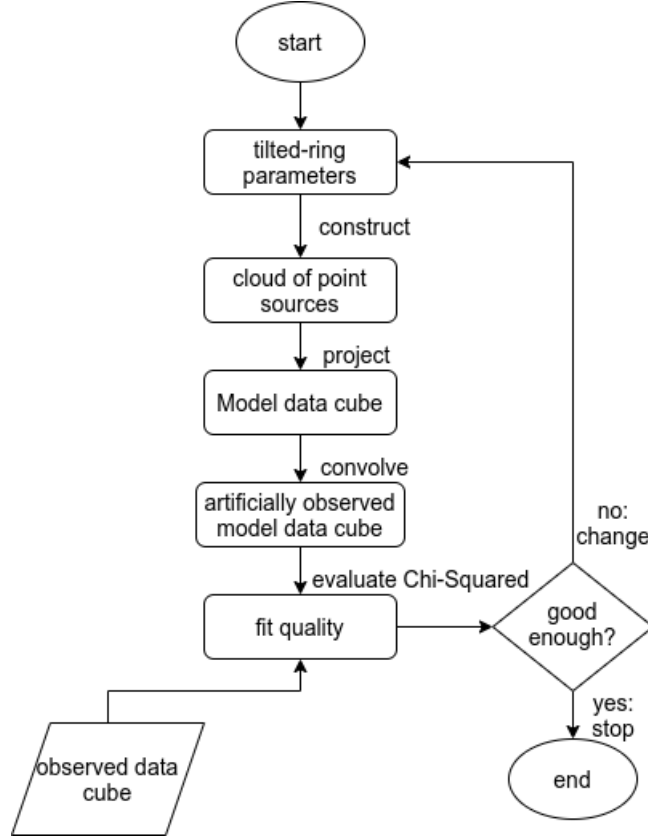


FIGURE 3.1: Flow chart diagram detailing how TiRiFiC works.

### 3.1.2 TiRiFiC Chi-Square evaluation

The convolution step in TiRiFiC serves to reduce the resolution of the model cube to that of the observed galaxy. The fastest means of convolving and the most computationally friendly way is using Fast Fourier Transform (FFT) convolution, where the FFTW library (Frigo & Johnson 2005) is used within TiRiFiC (Józsa et al. 2007b).

Once convolved, TiRiFiC compares the two images, simulated and observed, using a Chi-Square minimisation technique as shown in Equation 3.1

$$\chi^2 = \sum_k \frac{(M_k - O_k)^2}{\sigma_k^2} = \sum_k \frac{(M_k - O_k)^2}{w_k}, \quad (3.1)$$

where  $k$  runs across all pixel values,  $M$  represents the simulated and  $O$  the observed cube, whereas  $\sigma_k$  represents pixel noise in the original cube. The weight parameter,  $w_k$  can be given by Equation 3.2

$$w(k) = \sigma_{\text{rms}}^2 + (\sigma_k^q)^2, \quad (3.2)$$

provided that the quantisation noise,  $\sigma_k^q$  has been correctly accounted for. It is given by Equation 3.3

$$w_k = \frac{\sigma_{\text{rms}}^2 \cdot W^2 + (\sigma_k^q)^2}{W^2}, \quad (3.3)$$

where  $w_k$  is evaluated for each pixel. Setting the  $W$  parameter to  $\infty$  means the weight map has a constant pixel value  $RMS^2$ . However, setting  $W$  to 1 makes it possible to calculate the noise correctly.

To produce the model cube, a Monte-Carlo sampling technique is used.

A minimising algorithm, the Golden Section algorithm, has been applied to find the minimum in the Chi-Square, as described in Section 3.2. This algorithm has the benefits of not strictly requiring the Chi-Square to be normalised, and therefore, the rms in the image does not contribute a lot to the minimisation process. However, the algorithm finds only local modes and is unable to find global modes in a multimodal Chi-Square environment.

Concerning fitting parameter values, TiRiFiC affords the user an option for fitting parameters with multiple radial points as a single value. It also has the option for splitting the fit into multiple radial nodes, or rings, fitting each ring independent of all others.

## 3.2 Fitting Algorithms in TiRiFiC

### 3.2.1 Golden selection method

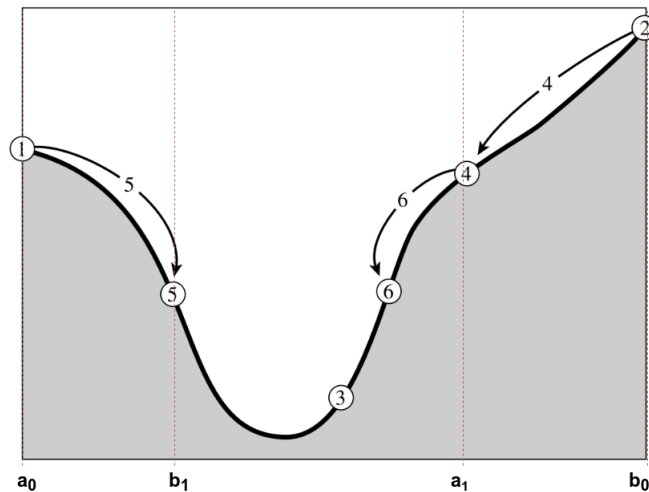


FIGURE 3.2: The Golden Section minimization (Press et al. 2007).

Figure 3.2 is used to illustrate how the Golden Section method works. Taking a function and defining intervals  $a_0$  and  $b_0$  such that they enclose a single minimum, the

second selection of points  $a_1$  and  $b_1$  is made such that the ratio of

$$\frac{(a_1 - a_0)}{(b_0 - a_0)} = \frac{(b_0 - b_1)}{(b_0 - a_0)} = 0.618, \quad (3.4)$$

which is the golden ratio. Later on comparing  $f(a_1)$  and  $f(b_1)$  if  $f(a_1) > f(b_1)$ , then the interval  $a_1$  to  $b_0$  is discarded making  $a_0$  and  $a_1$  the new interval range from which the whole process is repeated. This happens for minimizing case. However, for maximizing, if  $f(a_1) > f(b_1)$ , then the range  $a_0$  to  $b_1$  is discarded and the new interval becomes  $b_1$  to  $b_0$  from where the process is repeated until the algorithm converges either to a maximum or minimum depending on the objective (Press et al. 2007).

In TiRiFiC, this algorithm is selected using the parameter FITMOD and setting it to 1. When all parameters have been iterated over, and a minimum in each has been found in one cycle, this forms a loop in TiRiFiC. The step width is varied by changing the search direction or multiplication with a factor defined as  $(1 - w)/w \approx 2$  where  $w = (3 - \sqrt{5})/2$ , when enlarging the step width or  $w/(1 - w) \approx 0.62$  when reducing the step width. The key words DELSTART, DELEND and MODERATE control the step width.

The stopping condition is defined such that the minimising algorithm will be satisfied and stop only if the maximum number of loops has been reached, only applicable to FITMODE=1, and the parameter change is less than a predetermined value given by SATDELT if using FITMODE=1 or SIZE if using FITMODE=2.

### 3.2.2 The Simplex Algorithm

A simplex is a geometrical object with  $N+1$  points for  $N$  dimensions. A good example of such an object is a triangle in two dimensions or a multidimensional triangle in many dimensions. A simplex also has vertices, faces joined by the vertices.

The downhill simplex method is only concerned with non-degenerate simplexes, that is, those that enclose an  $N$ -dimensional volume. Taking any point as the origin, the other  $N$  points define vectors pointing towards these  $N$  points.

For multidimensional minimisation, the simplex algorithm is given an  $N$ -vector of points as an initial guess or starting point, and from this, it proceeds to make its way downhill through many complex topologies in  $N$ -dimensional space, until it encounters a local minimum. In the process, the simplex method makes a series of steps, where each step takes a point within the simplex, where its function evaluation is largest, to the opposite facet of the simplex where its function evaluation becomes low. Whenever possible, the simplex expands taking large steps, where the steps are

called reflections, in one direction until it reaches the valley floor where it contracts transversally, minimising its volume up to an allowed threshold as shown in Fig 3.3

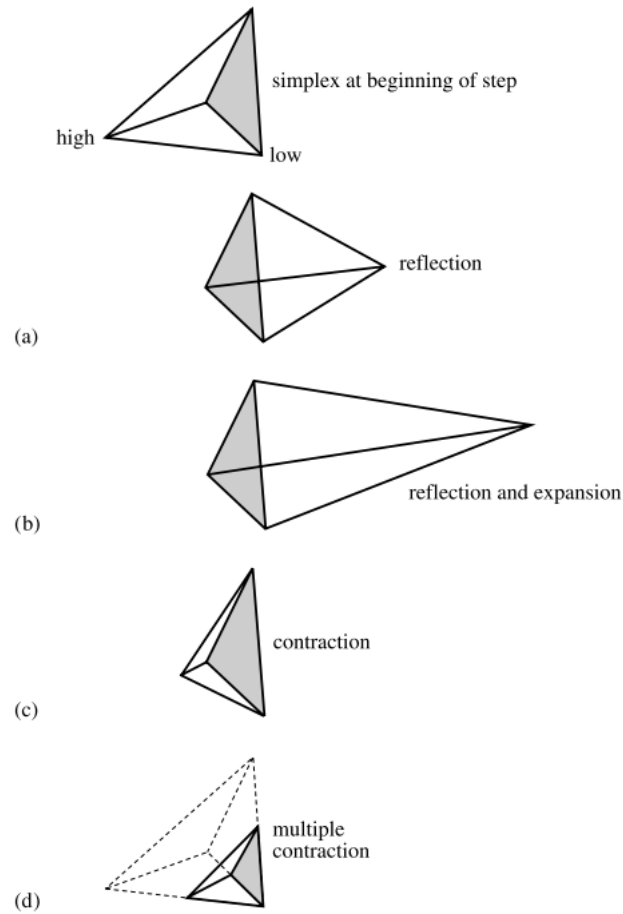


FIGURE 3.3: The simplex method follows a series of steps as shown. a) is a reflection away from the high point, b) is the reflection and expansion away from the high point, c) is a contraction along one dimension from the high point, and d) is a contraction along all dimensions towards the low point. (Press et al. 1992)

It is generally a good practice to restart a multidimensional fit of the simplex algorithm where it claims to have found a minimum. This is because it may be ‘fooled’ when one of the steps did not work.

TiRiFiC uses the version of the simplex algorithm implemented in the GSL library function and allows a user to select the simplex algorithm by setting `FITMODE = 3` in its input. The size of the simplex within the algorithm is given as the rms of the distances between the barycenter and the points or vertices of the simplex.

### 3.2.3 Pswarm minimizer

The Pswarm (Vaz & Vicente 2007), or particle swarm minimiser, is built upon the concept of a swarm of particles where each particle has a given velocity that defines its trajectory. Assuming the current particle  $i$  is located at position  $x^i(t)$  at time  $t$ , at time  $t+1$ , the particle shifts to position  $x^i(t+1)$  where this new position is computed as Equation 3.5

$$x^i(t+1) = x^i(t) + v^i(t+1) \text{ for } i = 1, \dots, s, \quad (3.5)$$

where  $i$  runs from 1 to population size  $s$ . This particle at the new position has an associated velocity vector given by Equation 3.6

$$v_j^i(t+1) = \tau(t) v_j^i(t) + \mu \omega_{1j}(t) (y_j^i(t) - x_j^i(t)) + v \omega_{2j}(t) (\hat{y}_j(t) - x_j^i(t)), \quad (3.6)$$

where  $j$  runs from 1 to  $n$ .  $i(t)$  is the weighting factor,  $\mu$  and  $v$  are positive real parameters or the cognition and social parameters,  $\omega_{1j}$ , and  $\omega_{2j}$  are randomly drawn from the uniform distribution within the range (0,1). While  $y^i(t)$  is the position with the best objective function found so far.

The algorithm is a hybrid of both pattern search methods and a swarm search step. Pattern search methods usually achieve convergence for the local minimum. However, certain parameter choices can enable the pattern method to jump from one local minimum to another that is hopefully better.

The idea behind the partial swarm search method is to start with an initial population and apply one step of particle swarm at each step. Whenever the search step succeeds, the swarm step searches the local neighbourhood of the minimum. However, when the search step fails, a polling step is applied to the best point out of all particles, and a local search is applied to the poll centred at this point.

TiRiFiC implements the c version of the code as part of its minimisers. However, it does not use the parameters ITESTART, ITENEED, SATDELT, DELSTART, DELEND and MODERATE, but instead, some new parameters are used as described in the TiRiFiC website.

### 3.2.4 Selecting TiRiFiC minimizer

Currently, TiRiFiC is running with an option of selecting among the three minimisers as described above. It runs some model evaluations per loop in the minimisation process until a stopping criterion is met. The details about loops and model evaluations vary from minimiser to minimiser, and the user has the option of specifying the total

number of loops they wish in the input. The user can choose the total number of iterations using the keyword `MAXITER` while `CALLITE` is another keyword which specifies how many models are formulated at each iteration. However, this is only applicable to `FITMODE = 2`. Implemented minimisers in `TiRiFiC` require a search direction. The search direction is supplied using `MODERATE`, `DELSTART` and `DELLEND`, where `DELSTART` needs the user to provide a starting step width at the beginning of the first loop.

### 3.3 `TiRiFiC-MultiNest` General description

When a galaxy is fitted with `TiRiFiC`, the first check usually is to compare the observed and simulated cube visually. After visual comparison, a more rigorous inspection is done on the different output parameters, i.e. the variation of rotation velocity, inclination, position angle and surface brightness with radius.

There may be deviations on these output parameters from what is expected from theory. These deviations result from parameters associated with the minimum Chi-Square, which is a local minimum instead of a global minimum.

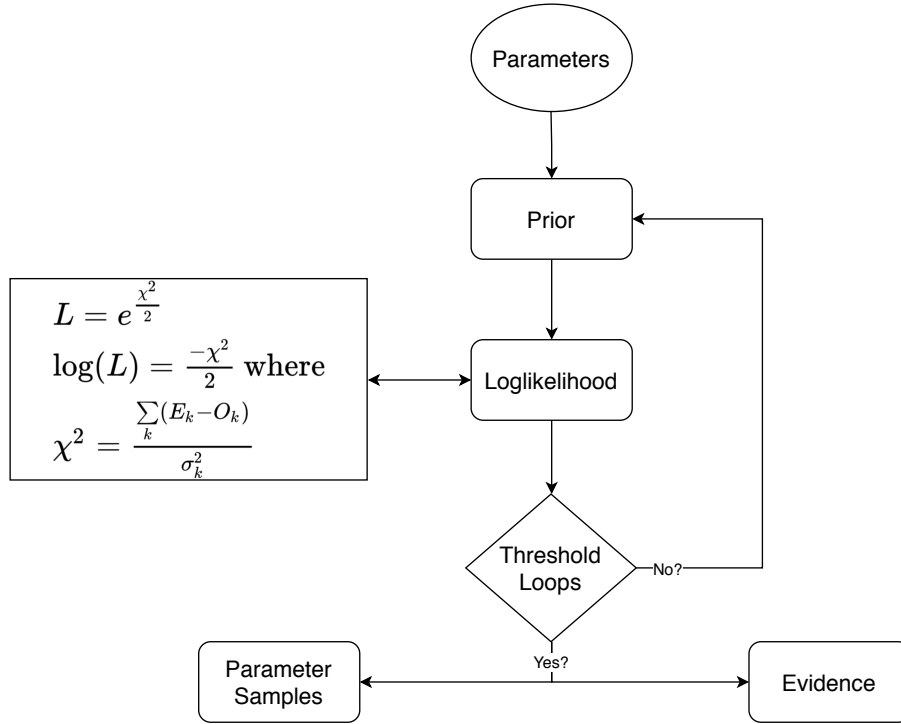
It is usually a challenge for fitting algorithms, including the currently implemented `TiRiFiC`'s minimisers, to find an optimal value in the multimodal Chi-Square environment. Monte-Carlo based fitting algorithms are prone to get stuck in a local minimum resulting in wrongly derived parameter values. A means to combat this in `TiRiFiC` is to adjust parameter values after each `TiRiFiC` fit and re-run the fit. This ensures that parameters conform to what has been found previously and peer-reviewed.

However, the problem of multimodal chi-square minimisation can be solved using Bayesian statistics. This solution would involve transforming the Chi-Square problem into that of maximum likelihood. Priors on parameter values can then be used to determine the posterior distribution. It is this posterior distribution that would be sampled by Bayesian-based samplers, where a maximum in the posterior would correspond to the best fit parameter values.

An ideal Bayesian sampler would need to have the ability to sample from a multimodal distribution and return parameters corresponding to the optimal mode. The sampling process would ensure a cluster of points per parameter values. These points would be used to determine errors of single parameters and quantify correlations between parameter values. The Bayesian approach using ellipsoidal sampling techniques implemented in `MultiNest` offers an opportunity to determine parameters corresponding to the optimal mode (see Chapter 2.1.3). `MultiNest` was chosen over other samplers

due to its ability to sample multimodal distributions that may have curving degeneracies, as well as its ease of use as it does not have tuning parameters that would need to be adjusted for every galaxy.

MultiNest requires a prior and likelihood function. The prior function is executed first, followed by the log-likelihood function for every iteration as illustrated in the flow chart diagram in Fig 3.4.



Implementing MultiNest as a sampler for TiRiFiC requires an interpretation of parameters described in TiRiFiC and MultiNest. These parameters would have to be translated into prior values and supplied to MultiNest configuration file. A uniform prior was used, requiring parameter ranges to be specified in MultiNest configuration file. In the usual sense, an analytical likelihood function is fitted for by MultiNest,



and consequently, finding optimal parameter values.

However, TiRiFiC has no analytic function describing the tilted-ring-model but instead employs Monte-Carlo simulation to make a new model and uses Chi-Square minimisation. It is this Chi-Square value from TiRiFiC that is translated into log-likelihood value and supplied to MultiNest during fitting. MultiNest calculates the evidence value and supplies posterior samples. The resultant posterior values are the best fit parameter values for TiRiFiC with a cluster of points around each of them. These clusters of points were used to show the correlation between parameter values as well as quantify uncertainties in the form of error bars. The assumption made was that the central limit theorem was obeyed on convergence.

To combine TiRiFiC with MultiNest, TiRiFiC has to provide the Chi-Square value as shown in Fig 3.5. The process acts as a cycle. The user starts by selecting which TiRiFiC parameters they will fit as well as the number of radial nodes  $n$  needed for each parameter (see Fig 3.5). They later proceed to set the prior ranges or boundary values of these parameters. In specifying the boundary values, the user may decide to provide a distribution corresponding to any of these parameters and the default uniform distribution. The default uniform distribution allows a user to supply the maximally allowed parameter ranges.

Once these parameters are set, and the algorithm starts running, TiRiFiC is loaded into memory waiting for a change in its configuration file. The algorithm moves onto applying prior constraints to these parameter values in such a way that each group of  $n$  radial nodes is subject to the same prior values where  $n$  is the number of radial nodes. This process results in new parameter values generated within the constraints of the prior values. Later, these parameter values, corresponding to TiRiFiC's parameters, are rearranged in a way that will be understood by MultiNest as an  $N$ -dimensional array called a cube.

Parameter values within the cube are written onto TiRiFiC's configuration file triggering TiRiFiC to generate only one model to calculate the Chi-Square value from the cube parameters. This single run results in TiRiFiC using the parameters, written to the configuration file, to model a simulated galaxy and from it, calculate the corresponding Chi-Square value which is passed on back to the algorithm. The algorithm calculates the log-likelihood value and later generates the evidence and checks for convergence. If convergence is not reached, the whole process is started by generating cube values of length  $N = n \times 4 + 3$ , whereby the algorithm has the same prior constraints (refer to Fig 3.5). This process continues until convergence is reached where the best-fit parameters, as well as the weighted parameters and other relevant files, are written to a folder.

### 3.3.1 Cube layout details

In an attempt to reduce the number of likelihood evaluations undertaken by TiRiFiC-MultiNest, rotation velocity (VROT), inclination angle (INCL), position angle (PA), surface brightness (SBR), right ascension (XPOS), declination (YPOS) and recession velocity (VSYS) were chosen to be fitted for, as they are the most important parameters without which you do not have a tilted-ring-model. However, it is possible for TiRiFiC-MultiNest to solve for more parameters than mentioned above. The cube parameter in MultiNest represents an array of numbers interpreted and arranged as shown in Fig 3.5 with VROT in  $\text{km s}^{-1}$ , INCL in ( $^{\circ}$ ), PA in ( $^{\circ}$ ), SBR in  $\text{Jy km s}^{-1} \text{ arcsec}^2$ , XPOS in ( $^{\circ}$ ), YPOS in ( $^{\circ}$ ) and VSYS in  $\text{km s}^{-1}$  following each other successively. The first VROT value at radial node zero belongs to the first index value, zero, whereas the last variable, VSYS, was set to a single variable at the last index,  $(n \times 4 + 3) - 1$ , where  $n$  is the number of radial nodes (see Fig 3.5).

The cube's length is given by  $N = n \times 4 + 3$  and greatly influences MultiNest's execution time as it's length increases. Since it's length is connected to the number of radius points along the major axis of the galaxy, and hence, the total number of parameters being simultaneously fitted, the convergence speed is greatly influenced by the number of points being fitted for. This increase in the execution time happens because of the misalignment of the ellipsoidal bound, within MultiNest, that results in an exponential increase in volume with increasing dimensions. Points that fall below the lowest likelihood contour but inside the imperfect ellipsoidal bound reduce the sampling efficiency as sample points drawn from this 'extra' volume are rejected slowing down the algorithm. The full TiRiFiC-MultiNest algorithm is shown in Fig 3.5.

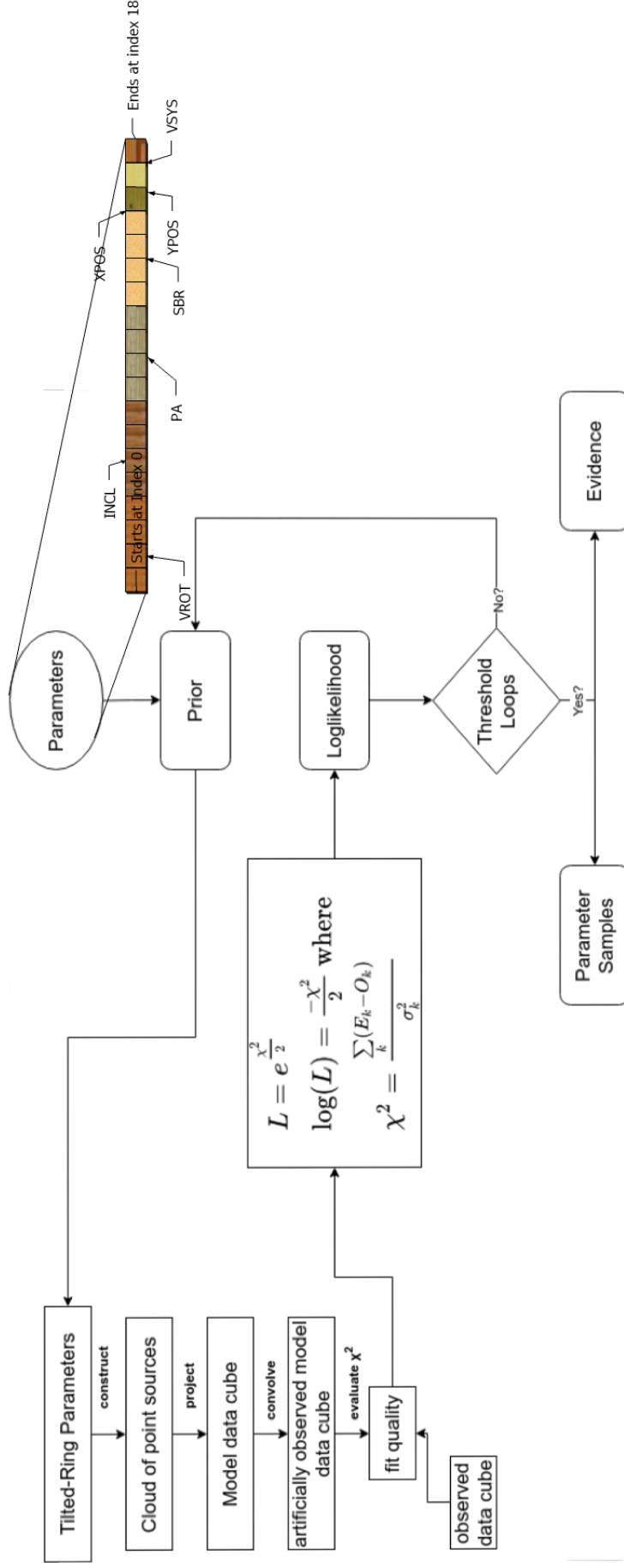


FIGURE 3.5: TiRiFiC-MultiNest Flowchart Diagram including how the cube parameter was divided and fed to MultiNest.

Consequently, the experiments done in the thesis used  $n = 4$  and  $n = 7$  radial nodes. The fitted points for each parameter value can later be interpolated to get more data points per parameter value.

### 3.3.2 Parallelisation

MultiNest is MPI parallelised and handles the entire process from spawning multiple processes to combining them at the end. In the TiRiFiC-MultiNest algorithm, TiRiFiC supplies the likelihood value to MultiNest. Consequently, TiRiFiC-MultiNest is also parallelised, where multiple TiRiFiC instances supply the loglikelihood values needed by the different instances of MultiNest as shown in Fig 3.6.

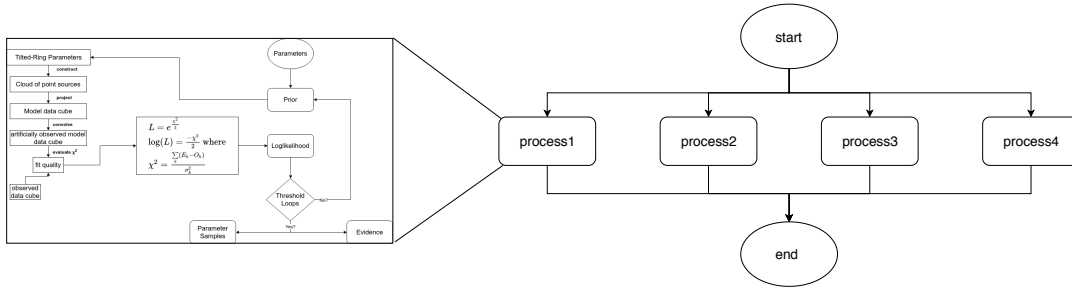


FIGURE 3.6: TiRiFiC-MultiNest parallelization process.

In details, the following sequence is followed.

First,  $m$  copies of MultiNest are determined and launched. Next,  $m$  copies of TiRiFiC's folder containing TiRiFiC's configuration file and observed cube are made where a suffix number is attached to each copy of the folder name differentiating them from each other. The suffix number represents the rank number of the parallelization process. For example, tirific0, tirific1 and so on.

Later, each instance of MultiNest starts a corresponding instance of TiRiFiC using the Linux *screen* command and appends to it a suffix representing the rank number. Each instance of TiRiFiC works on a correspondingly named copy of TiRiFiC folder. For example, TiRiFiC0 works from the folder tirific0 and so on.

Finally, each copy runs as described above (see Fig 3.5), determining its prior values and getting its likelihood value from the corresponding TiRiFiC launched, which is different from all the other copies. Once the fitting is done, MultiNest handles all the merging and eventually, outputs the best fit parameter values.

## Chapter 4

# Tests and Results

### 4.1 Testing

The galaxy NGC 4062 was observed with the Westerbork Synthesis Radio Telescope (WSRT) in the scope of the HALOGAS project (Heald et al. 2011). Experiments 1 through 10 were made based on this observation. A simulated cube with a known parametrisation based on the observation of NGC 4062 was used to simulate a realistic galaxy. Additionally, a final fit was performed on the original data cube of a different galaxy, NGC 3198.

#### 4.1.1 Reducing the prior range using SoFiA

As mentioned previously, the dimensionality of the problem, and hence the number of parameters to be fitted for increases with an increase in the radial node number as well as the total number of TiRiFiC parameters to be fitted for and the sampling volume to be covered, which depends on the prior supplied. Consequently, there is a need for reducing the prior volume to be sampled by reducing the parameter constraints or ranges.

One attempt at this is to run a fit for the galaxy using SoFiA (Serra et al. 2015) in the first stage to derive the best-fit parameters and use them as the central values from which a new range is defined. SoFiA is a software used to detect and parametrize sources detected in three-dimensional spectral line cubes. To run this fit, SoFiA was set using the parameters described in Appendix A.3 Parameters set for SoFiA run.

The inclination value was later roughly estimated according to Equation 4.1 using the best-fit values from SoFiA,

$$\text{incl} = \arccos\left(\frac{\text{ellmin} - \text{beamsize}}{\text{ellmaj} - \text{beamsize}}\right), \quad (4.1)$$

where

$$\text{beamsize} = \frac{\text{beamsizemaj} + \text{beamsizemin}}{2}, \quad (4.2)$$

and  $\text{ellmin}$  and  $\text{ellmaj}$  are the fitted ellipse's major and minor axis values, while  $\text{beamsizemin}$  and  $\text{beamsizemaj}$  are the beam's minor and major axes in pixels, as determined from the header of the fits file and converting the fits file's degrees units to pixels units. This derived inclination value is used to find rotation velocity by applying Equation 4.3

$$\text{vrot} = \frac{\text{channelwidth} \times w_{20}}{2 \sin(\text{incl})} \quad (4.3)$$

where the channel width is the width of a single channel in km/s determined from the fits file header and  $w_{20}$  parameter in *channels* from SoFiA.  $\text{Vrot}$  is in  $\text{km s}^{-1}$  and the inclination angle in degrees.

SoFiA also gives the parameter  $f_{\text{int}}$ , in  $\text{Jy Beam}^{-1}$ , representing the integrated flux derived from Busy Function fit. The parameter  $f_{\text{int}}$  was converted to  $\text{Jy km s}^{-1}$  using Equation 4.4

$$F(\text{Jy km s}^{-1}) = \frac{f_{\text{int}}(\text{Jy beam}^{-1}) \times \Delta x \times \Delta y \times \Delta v}{1.13309 \times \text{HPBWmin} \times \text{HPWBmaj}}, \quad (4.4)$$

where  $\text{HPBWmin}$  is the Smaller Half-Power-Beam-Width ( $\text{bmaj}$ ),  $\text{HPWBmaj}$  is the larger Half-Power-Beam-Width ( $\text{bmin}$ ),  $\Delta x$  is the Pixel size in the right ascension direction ( $|\text{cdelt1}|$ ), and  $\Delta y$  is the Pixel size in the declination direction ( $|\text{cdelt2}|$ ).

To convert from  $\text{Jy km s}^{-1}$  to  $\text{Jy km s}^{-1} \text{ arcsec}^{-2}$ ,  $\text{ellmaj}$  and  $\text{ellmin}$  values were converted from pixel to arcsec units using header information. The galaxy area was assumed to approximate an ellipse and calculated as shown in Equation 4.5.

$$\text{GalaxyArea}(\text{arcsec}^{-2}) = \text{ellmin}(\text{arcsec}) \times \text{ellmaj}(\text{arcsec}) \times \pi. \quad (4.5)$$

The value  $F(\text{Jy km s}^{-1})$  was divided by the galaxy area,  $\text{GalaxyArea}(\text{arcsec}^{-2})$  to give  $\text{SBR}(\text{Jy km s}^{-1} \text{ arcsec}^{-2})$  as shown in Equation 4.6.

$$\text{SBR}(\text{Jy km s}^{-1} \text{ arcsec}^{-2}) = \frac{F(\text{Jy km s}^{-1})}{\text{GalaxyArea}(\text{arcsec}^{-2})} \quad (4.6)$$

#### 4.1.2 SoFiA best-fit parameter values

On running SoFiA, the first best-fit parameter values were recovered from the data cube and later,  $f_{\text{int}}$  parameter was converted to sbr according to the procedure described in Subsection 4.1.1 resulting in the values shown in Table 4.2.

Once the best fit values were determined, the prior ranges were later made from these values as follows

$$\text{VROT max} = \text{VROT best-fit value} \times 3.00 \text{ km s}^{-1}$$

$$\text{VROT min} = 0$$

$$\text{SBR max} = \text{SBR best-fit value} \times 10.00 \text{ Jy beam}^{-1}$$

$$\text{SBR min} = 0$$

$$\text{INCL best-fit value} \pm 20.00^\circ$$

$$\text{PA best-fit value} \pm 20.00^\circ$$

$$\text{XPOS best-fit value} \pm 0.01^\circ$$

$$\text{YPOS best-fit value} \pm 0.01^\circ$$

$$\text{VSYS best-fit value} \pm 20.00 \text{ km s}^{-1}$$

#### 4.1.3 Fits file data preparation

For the first data cube, using NGC 4602 as a rough template, a simulated data cube was made to act as a proof of concept, that the algorithm can determine the global minimum in the Chi-Square value and hence the corresponding best-fit parameters.

This simulated cube was made by editing TiRiFiC's configuration file, the deffile, with parameters VROT, PA, SBR, XPOS, YPOS and VSYS and setting the number of radial nodes to four, the number of disks to one, loops to zero and the error parameter RMS to  $0.99 \text{ mJy beam}^{-1}$ . Later, TiRiFiC was run using this configuration file resulting in an output cube and a corresponding new configuration file exactly describing this new cube. This new cube was used in experiments 1 through 5 and 7 that had one and four radial nodes.

Experiments 6, 8 to 10 were made in such a way that TiRiFiC's parameters VROT, PA, SBR, XPOS, YPOS and VSYS had seven nodes each, with the number of disks set to one and the error parameter, RMS, set to  $0.19 \text{ mJy beam}^{-1}$  which was derived from the datacube by calculating its mean absolute deviation.

When it came to testing with the real galaxy NGC 3198 instead of using a simulation, there was a need to reduce the size of this cube so that the Chi-Square calculation time for every iteration could also be reduced, and also so that the galaxy could comfortably fit into memory as the original cube was 500 MB in size. The first step in reducing this galaxy's size was to convolve it with a Gaussian beam, representing low-resolution observation, then later binning the convolved cube in such a way that there are at least 3 pixels per HPBW. This task was accomplished using the Groningen Imaging Processing SYstem (GIPSY) ([van der Hulst et al. 1992](#)) tasks SMOOTH and REPROJ. The final reduced data cube had a size of 1.8 MB.

#### 4.1.4 Error calculation and correlation plots

TiRiFiC-MultiNest algorithm samples the posterior distribution resulting in samples generated for each TiRiFiC parameter value per radial node such that VROT for radius one has samples with its best fit value, which is different from VROT for radius two or any other parameter specified within TiRiFiC's configuration file. On convergence, these samples are used to derive the marginalised standard deviation and the error bars shown are at a level of  $3 \times \sigma$  where  $\sigma$  is the standard deviation of each TiRiFiC parameter per radial node.

It is from these parameter samples that the correlation plots are also derived, allowing identification of any correlations between TiRiFiC's parameters.

#### 4.1.5 Experiment 1

Experiment 1 was carried out using a real data cube for NGC 4062. TiRiFiC's VROT, PA, INCL, SBR, XPOS, YPOS and VSYS parameters were set as constant values across the four radial nodes hence effectively fitting a flat disk. MultiNest's cube parameter used a single copy of each parameter in each radial node resulting in a cube parameter of length seven as shown in [Fig 4.1](#)



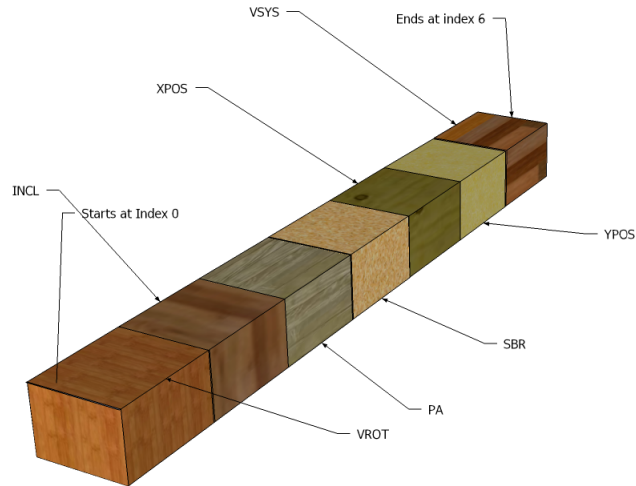


FIGURE 4.1: Cube parameter as seen and used by MultiNest.

The prior values applied on these parameters are shaded grey in 4.2 and can also be found in Appendix A.4 [Minimizing Prior Values Experiment 1](#), were prepared as described in Section 4.1.1 above. Care was taken to ensure that the positional parameters that constrain the dimensions of the galaxy within the data cube, namely, XPOS, YPOS and VSYS, were given realistic constraints. This constraint meant that their maximum and minimum values were observed not to exceed the dimensions of the data cube used, and this condition was applied to all experiments that followed.

This experiment gave the best fit value for every parameter as shown in Table 4.2

TABLE 4.1: Comparison between SOFIA and TIRIFIC-Multienst

Software	VROT	INCL	PA	SBR	XPOS	YPOS	VSYS
TiR-Mult	151.8	72.6	100	0.0002	181.016	31.895	764
SoFiA	174	72	100	0.00015	181.017	31.895	767

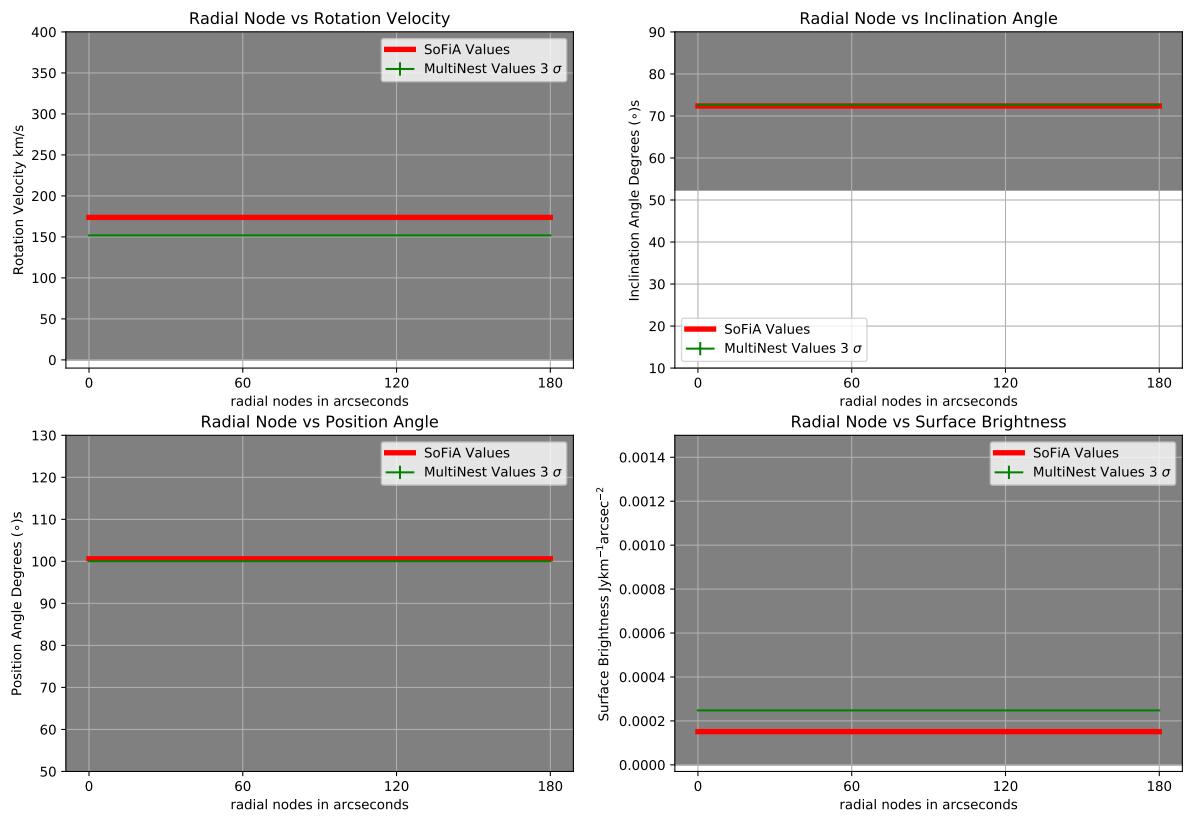


FIGURE 4.2: The Figures above represent the rotation velocity, inclination angle, position angle and surface brightness fits. Green line represents TiRiFiC-MultiNest fit whereas red line represents SoFiA fit. Error bars are shown at  $3\sigma$  but very small.

#### 4.1.6 Experiment 2

Experiment 2 used a simulated data cube of NGC 4062 (see 4.1.3) to fit for more parameters, 13 in total. Rotation velocity (VROT) and surface brightness (SBR) were fitted for with their radius points allowed to vary independently across the four radial nodes, whereas all other parameters were set to constant values and also fitted for. This led to a cube of length 13 as shown in Fig 4.3.

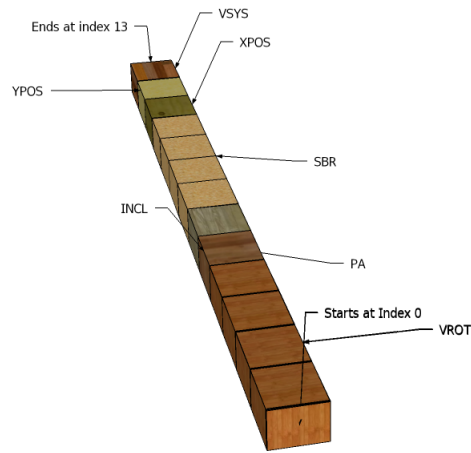


FIGURE 4.3: MultiNest cube of length 13, where TiRiFiC parameters PA, INCL, XPOS, YPOS, VSYS, are taken as constants across the four radial nodes.

The prior ranges, shaded grey (see Fig 4.4) and shown in Appendix A.5 [Minimizing Prior Values Experiment 2](#), were reduced as compared to Experiment 1 in an attempt to reduce the prior sampling volume for MultiNest.

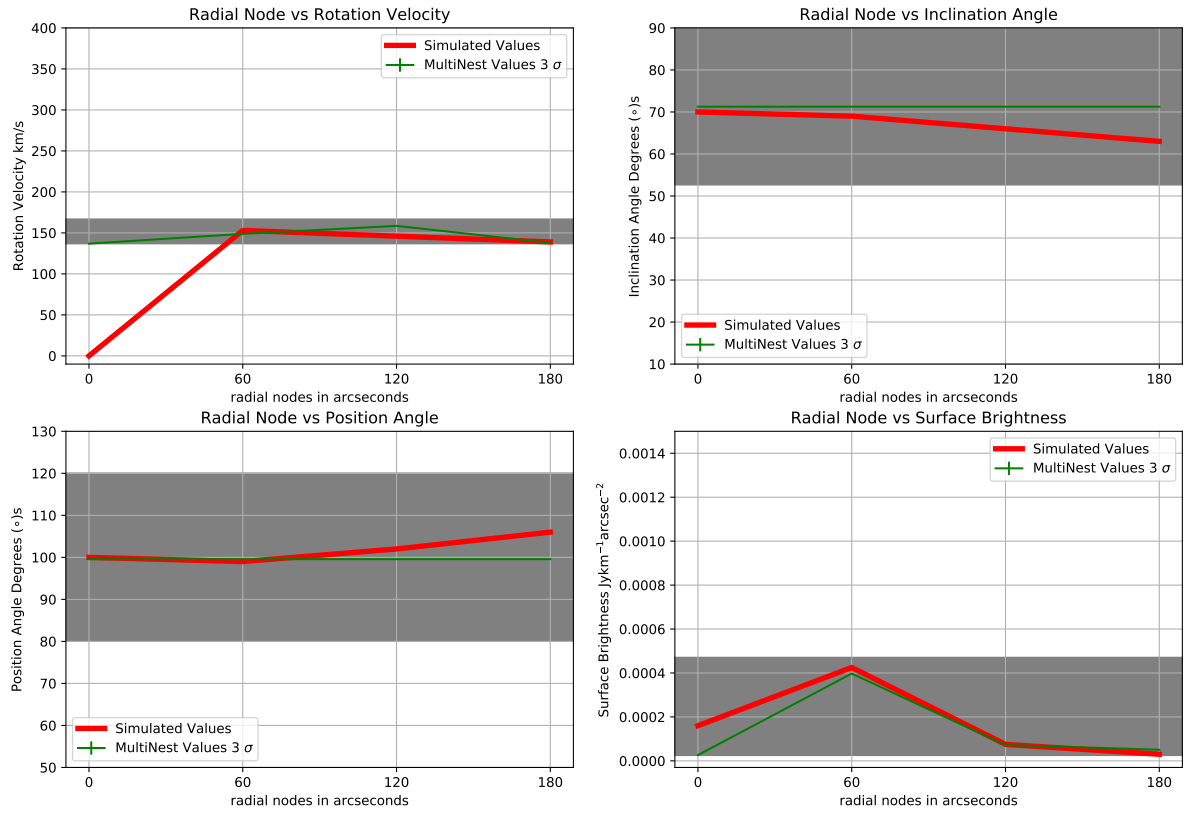


FIGURE 4.4: The Figures above represent the rotation velocity, inclination angle, position angle and surface brightness fits. Green line represents TiRiFiC-MultiNest fit whereas red line represents TiRiFiC fit. Error bars are shown at  $3\sigma$  but very small.

As can be seen in Fig 4.4, an incorrect fit was made. This was attributed to having prior values that were not inclusive of the actual parameter values at some of the radial nodes, which shows the importance of setting prior values that are inclusive of the actual values. To be specific, VROT parameter had a prior range of 169 to 136  $\text{km s}^{-1}$  which excluded the first radial node that had been set to a much lower value.

For visual comparison between the simulated and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.5 was made using SlicerAstro (Punzo et al. 2017).

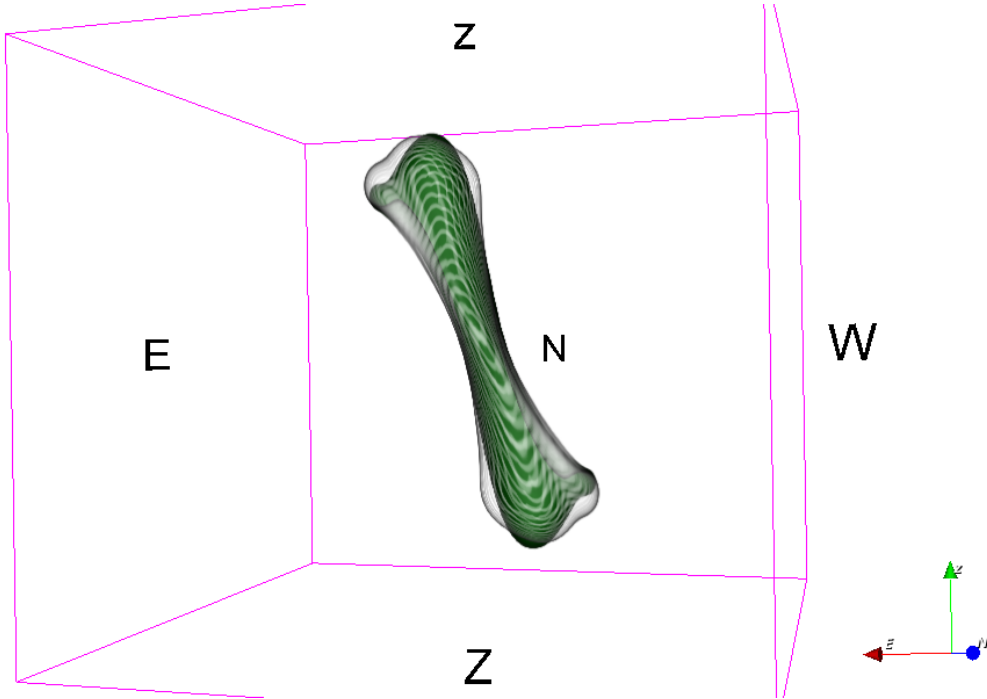


FIGURE 4.5: Comparison between NGC 4062 TiRiFiC simulated cube, green, and TiRiFiC-MultiNest best fit parameter values, white

#### 4.1.7 Experiment 3

Taking lessons from Experiment 2, the next test sets the prior values to their maximally allowed values as shown in 4.6 and also shown in A.6 Minimizing Prior Values Experiment 3, ensuring that no parameter will be left outside the prior value and result in an incorrect fit. Experiment 3 also used a simulated data cube of galaxy NGC 4062 as described in subsection 4.1.3.

The SBR and VROT parameters were set to vary, at four radial nodes, while the rest were set to a constant value, as shown in Fig 4.3. This resulted in the following best-fit plots

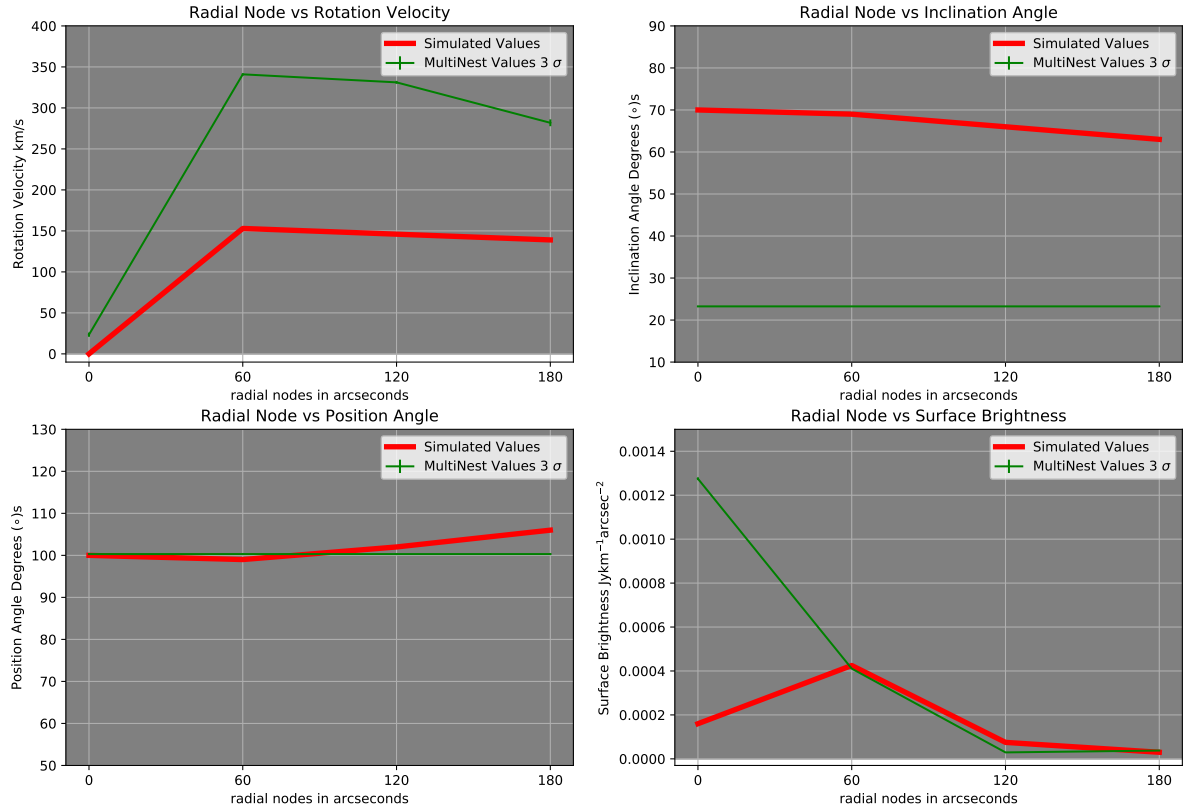


FIGURE 4.6: The Figures above represent the rotation velocity, inclination angle, position angle and surface brightness fits. Green line represents TiRiFiC-MultiNest fit whereas red line represents TiRiFiC fit. Error bars are shown at  $3\sigma$  but very small.

Parameter plots in Fig 4.6 hint at the possibility of the fit going wrong because of not varying PA and INCL with radius. The reason for the obvious failure of this fit could not be established. While in a long-slit experiment we would see the observed degeneracy between inclination and rotation velocity this would not be expected for our applied fitting technique.

For visual comparison between the simulated and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.7 was made using SlicerAstro (Punzo et al. 2017).

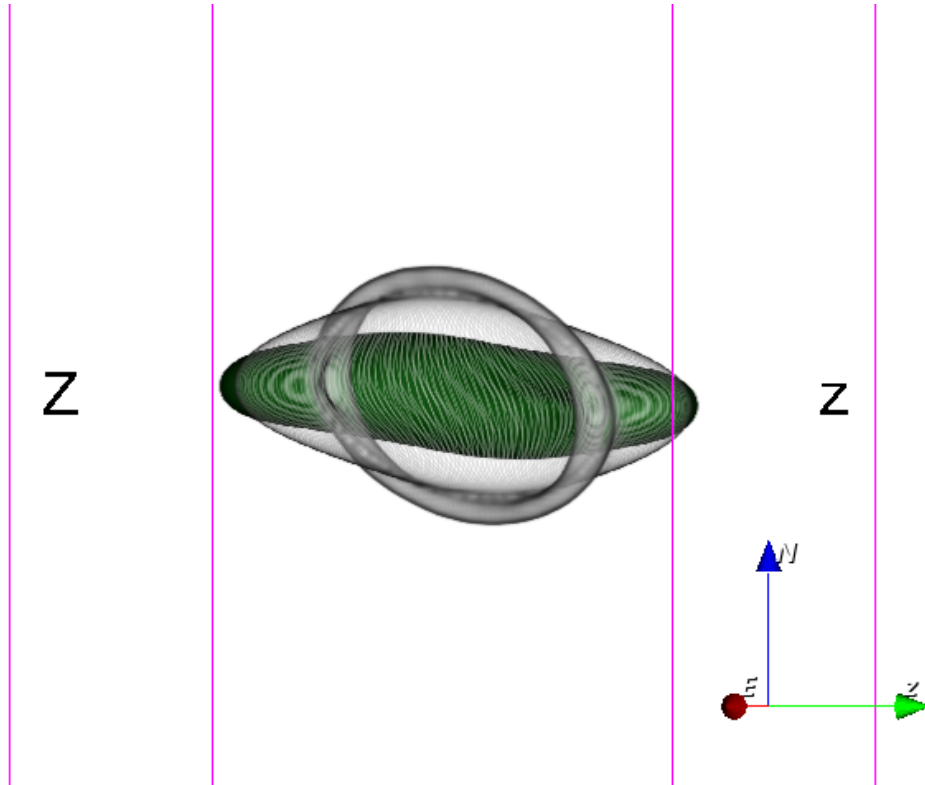


FIGURE 4.7: Comparison between NGC 4062 TiRiFiC simulated cube, green, and TiRiFiC-MultiNest best fit parameter values, white.

#### 4.1.8 Experiment 4

Later, a real data cube of galaxy NGC 4062 was fitted for using both TiRiFiC and TiRiFiC-MultiNest. The rotation velocity and surface brightness parameters were varied across the four radial nodes whereas all the other parameters were kept constant. Prior values were set as described in Appendix A.7 Minimizing Prior Values Experiment 4 and shaded grey in Fig 4.9. In determining the priors for the rotation velocity from SoFiA, the minimum in rotation velocity was set to zero whereas the maximum in rotation velocity was determined from SoFiA. SoFiA's best fit  $v_{\text{rot}}$  value of  $\sim 174$ , calculated as described in Equation 4.3, was first doubled then incremented by a small value,  $50 \text{ km s}^{-1}$ , in such a way that the final result,  $347.6 + 50 \text{ km s}^{-1}$ ,

would not exceed the velocity range spanned by the position velocity diagram shown in Fig 4.8.

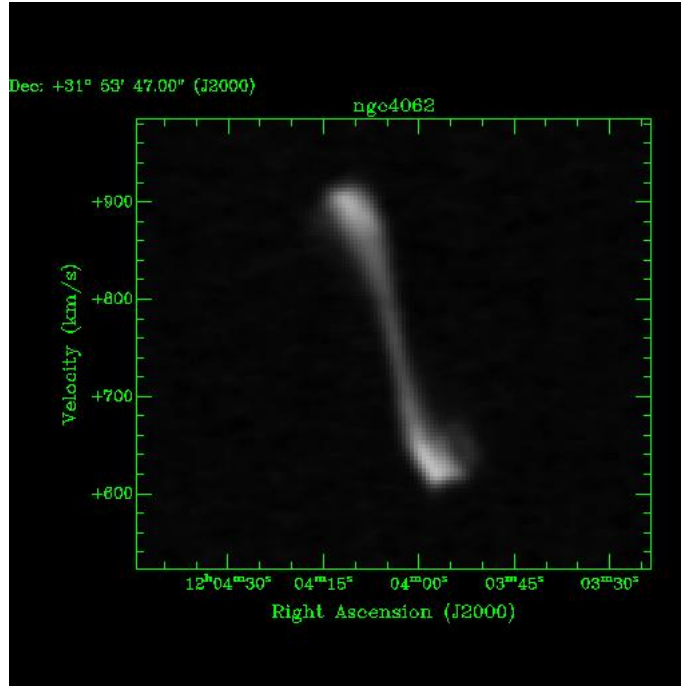


FIGURE 4.8: The range spanned by the PV slice of the galaxy shows that it was less than  $400\text{km s}^{-1}$ . From this, the observed rotation velocity of either the approaching or receding side of the galaxy should be less than  $200\text{km s}^{-1} / \sin(\textit{inclination})$  whereas the observed velocity range was less than  $400\text{km s}^{-1} / \sin(\textit{inclination})$

Other parameters had their prior values set as detailed in Appendix [A.7 Minimizing Prior Values Experiment 4](#), and the run gave plots shown in Fig 4.9.



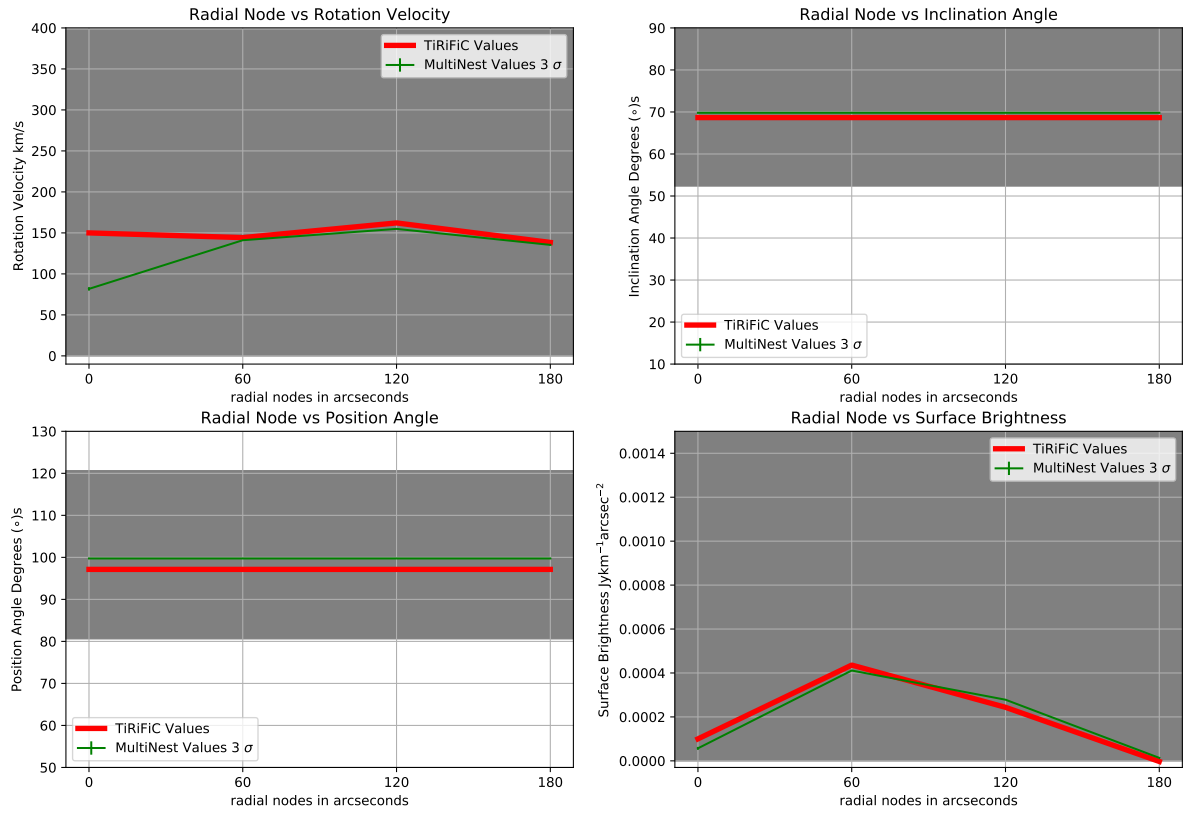


FIGURE 4.9: The Figures above represent the rotation velocity and surface brightness fits. Green is TiRiFiC-MultiNest fitted values whereas red is TiRiFiC fitted values for a real galaxy NGC 4062. Error bars are shown at  $3\sigma$  but very small.

For visual comparison between the actual and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.10 was made using SlicerAstro (Punzo et al. 2017).

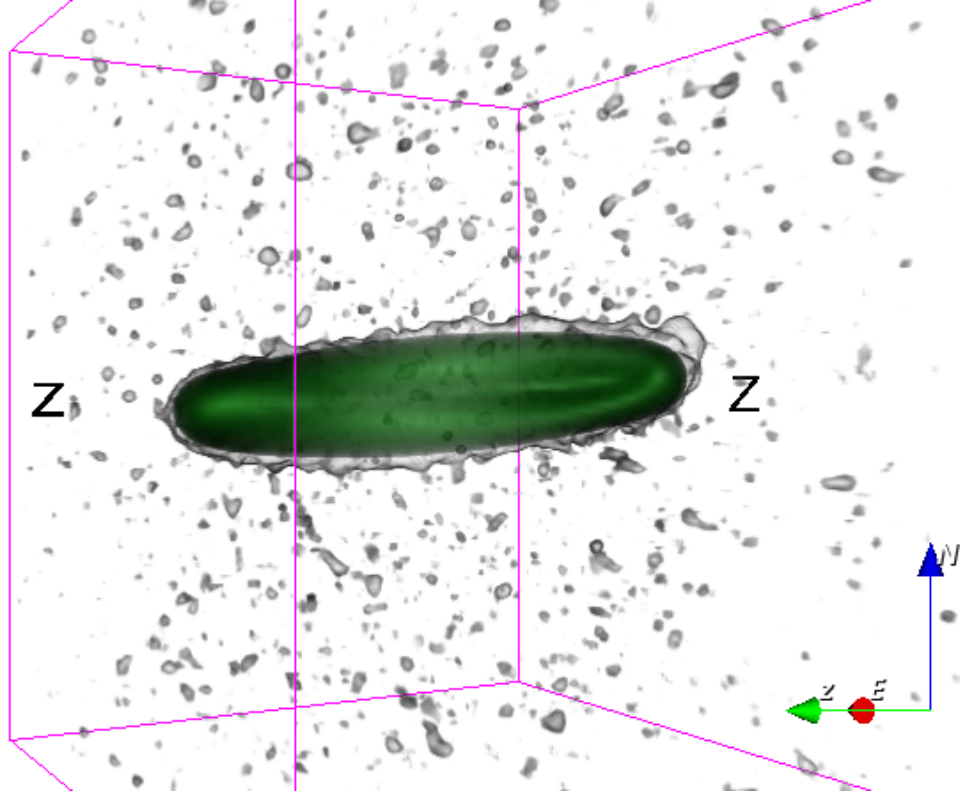


FIGURE 4.10: Comparison between NGC 4062 observed cube, white, and TiRiFiC-MultiNest best fit parameter values, green.

#### 4.1.9 Experiment 5

The simulated data cube for galaxy NGC 4062 was used in this experiment (see 4.1.3). Prior values as derived and used in Experiment 4 above were also used for this experiment and parameters for all radial nodes were set to vary, apart from position parameters, XPOS, YPOS and VSYS. Consequently, the length of the cube parameter became  $n \times 4 + 3$  where  $n$  was the number of rings or nodes or radial points to use, which, in this experiment, was 4, giving a total of 19 parameters. The first four parameters described VROT, then PA for the second four, followed by INCL and SBR. The other parameters, namely XPOS, YPOS and VSYS assumed that all the rings representing the galaxy shared a common centre in the data cube hence these were the only parameters that were left as single variables, therefore, duplicated four times to represent the four radial nodes. The cube is as shown in Fig 4.11

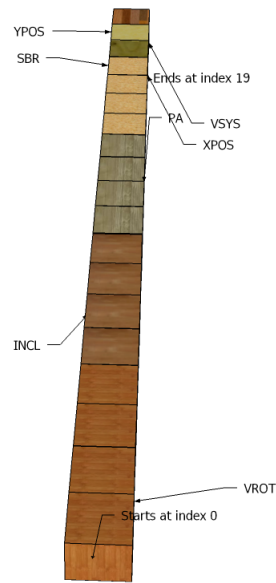


FIGURE 4.11: The cube parameter with VROT, INCL, PA and SBR set to vary with 4 nodes whereas XPOS, YPOS and VSYS are set to a constant value, which translates to one value in the cube array that is duplicated 4 times per parameter value.

The prior values are shaded grey in Fig 4.12 and also detailed in Appendix A.8 [Minimizing Prior Values Experiment 5](#) were used with the runs producing the following best fit values

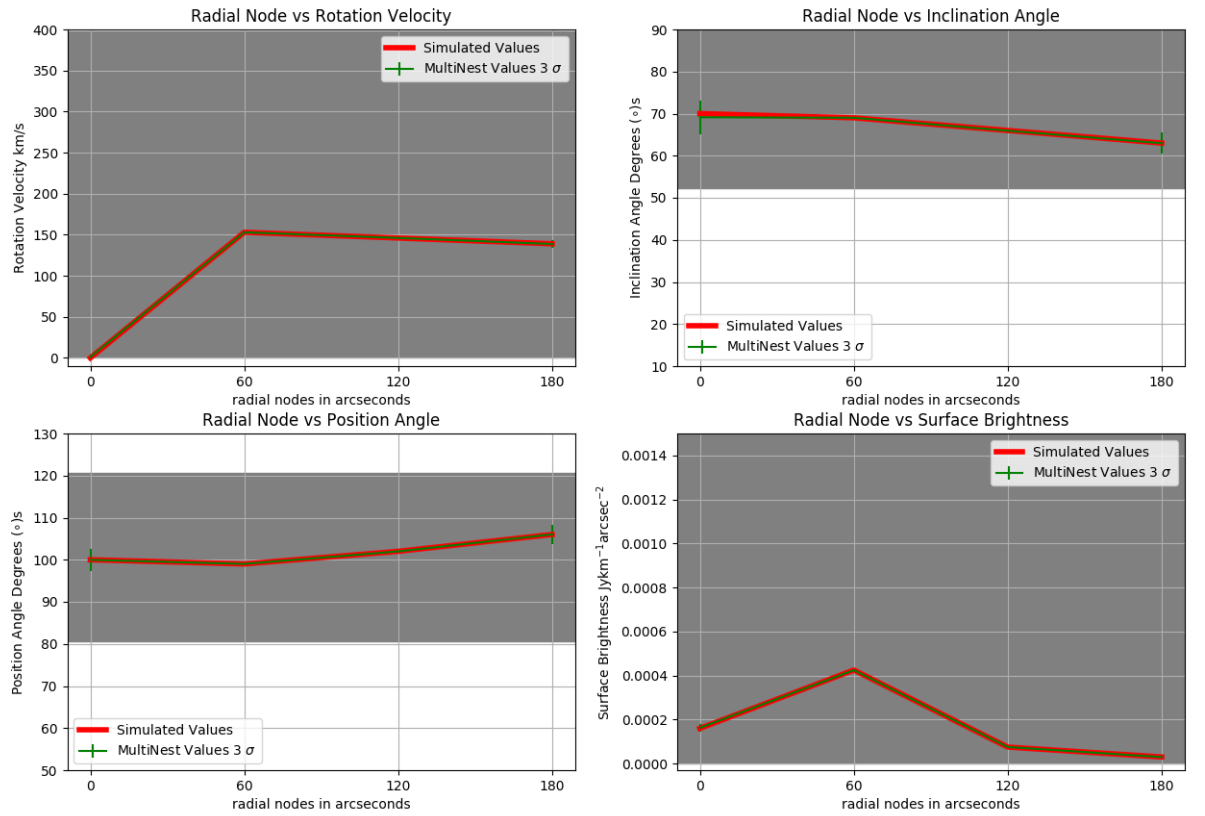


FIGURE 4.12: The Figures above represent the rotation velocity, inclination angle, position angle and surface brightness fits. Green line represents TiRiFiC-MultiNest fit whereas red line represents TiRiFiC fit. Error bars are shown at  $3\sigma$  but very small.

For visual comparison between the simulated and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.13 was made using SlicerAstro (Punzo et al. 2017).

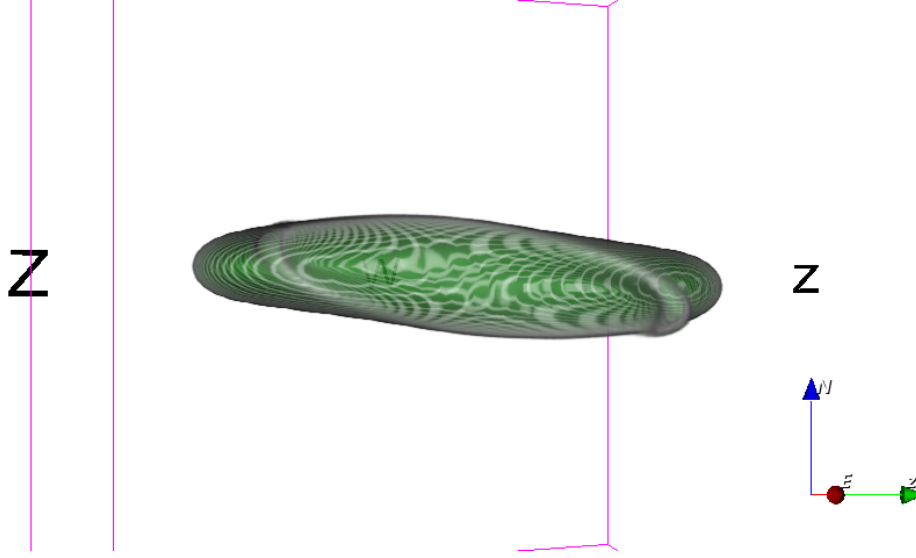


FIGURE 4.13: Comparison between NGC 4062 TiRiFiC simulated cube, green, and TiRiFiC-MultiNest best fit parameter values, white.

From this experiment, it was observed that having prior values general enough to include the actual parameters describing a galaxy resulted in good fits. This experiment also showed that setting the four parameters' radial nodes free resulted in a much better fit as compared to fixing the radial nodes of some parameters as done in Experiments 3 and 4 above.

#### 4.1.10 Experiment 6

Experiment 6 fitted for seven radial nodes instead of four using a real observed data cube. TiRiFiC and TiRiFiC-MultiNest were both used in the fitting process and their results compared as red and green lines in the plots respectively. The first step used VROT and SBR as free parameters with seven radial nodes whereas the rest of the parameters were set to a constant value resulting in a cube of length  $n \times 2 + 5$  where  $n$  is seven which becomes 19.

To obtain the prior values used, the following procedure was used on the best fit values derived from SOFIA

$$\text{SBR max} = \text{SBR best-fit value} \times 10.00 \text{ Jy km s}^{-1}$$

$$\text{SBR min} = 0.00 \text{ Jy km s}^{-1}$$

$$\text{INCL best-fit value} \pm 20.00^\circ$$

$$\text{PA best-fit value} \pm 20.00^\circ$$

XPOS best-fit value  $\pm 0.01^\circ$

YPOS best-fit value  $\pm 0.01^\circ$

VSYS best-fit value  $\pm 20.00 \text{ km s}^{-1}$

Whereas the VROT value was taken directly from SoFiA best fit value and added to it  $50 \text{ km s}^{-1}$  as described in Experiment 4. This resulted in the prior values detailed in Appendix A.9 Minimizing Prior Values Experiment 6 and also shaded grey in the parameter plots below.

And on running, produced the following best fit results

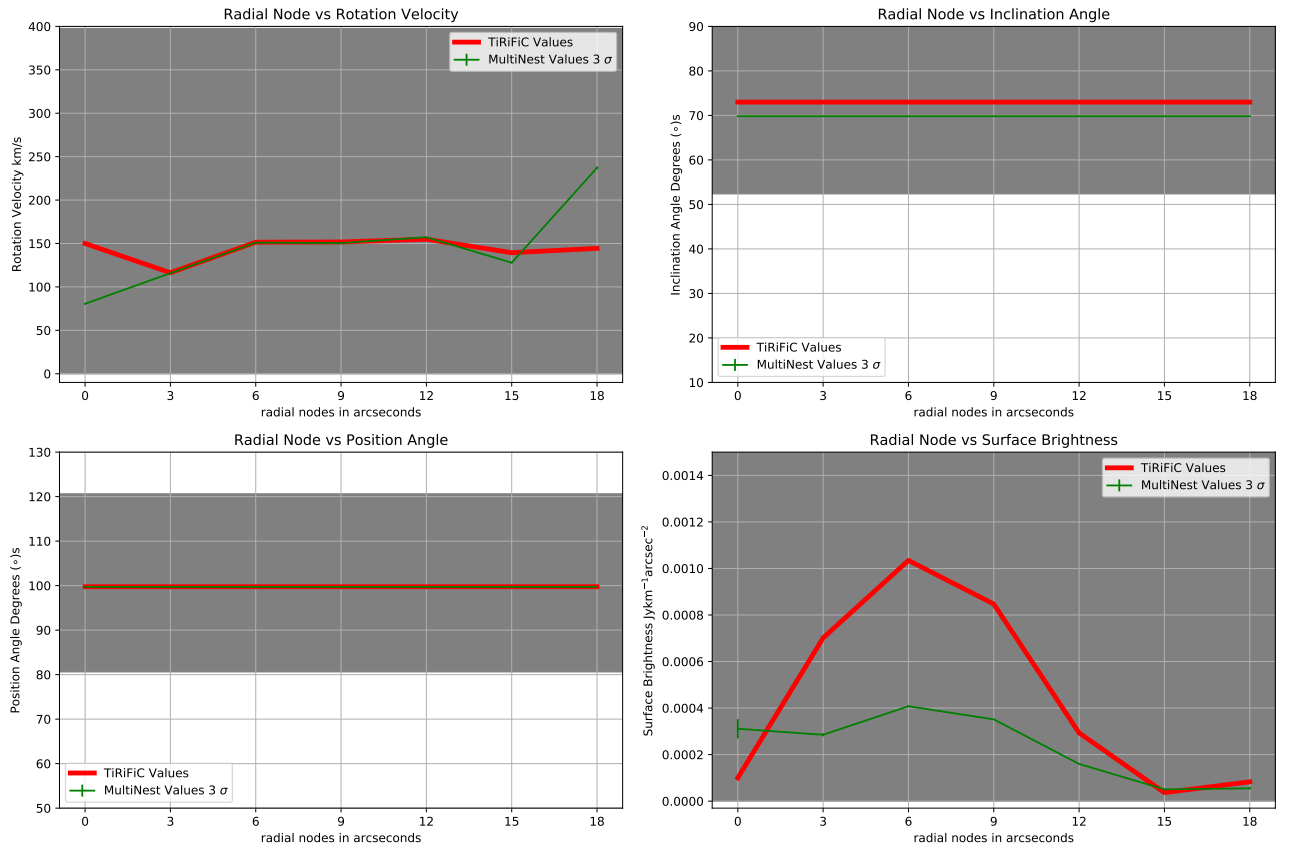


FIGURE 4.14: The Figures above represent the rotation velocity and surface brightness fits. Green line represents TiRiFiC-MultiNest fit whereas red line represents TiRiFiC fitted values for a real galaxy NGC 4062. Error bars are shown at  $3\sigma$  but very small.

It was noted that there was a considerable increase in the amount of time needed for MultiNest to reach convergence,  $\sim 5\,500\,000$  likelihood iterations, parallelized with number of processors at 5, as compared to Experiment 5 above which converged at  $\sim 760\,000$  likelihood iterations, also using 5 processors. It was also noted that the fits produced were not accurate. The suspected cause for this inaccurate fit was the use of the default number of live points, 500. This suspicion is tested in the subsequent

experiments.

For visual comparison between the observed and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.15 was made using SlicerAstro (Punzo et al. 2017).

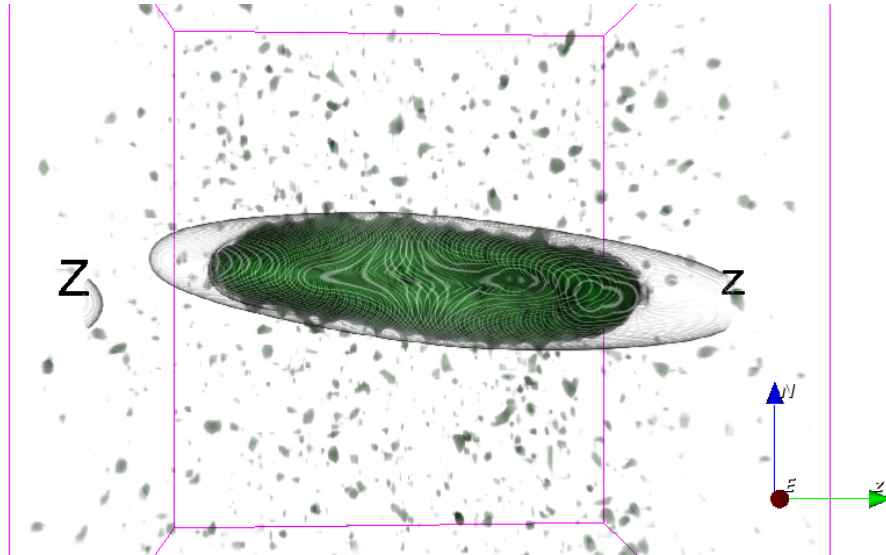


FIGURE 4.15: Comparison between NGC 4062 TiRiFiC simulated cube, green, and TiRiFiC-MultiNest best fit parameter values, white.

## 4.2 Further Attempts at reducing the number of likelihood evaluations

### 4.2.1 Experiment 7

While attempting to reduce the number of likelihood evaluations for the runs, the same simulated data cube as well as MultiNest cube structure defined in Fig 4.11 was used, but this time, the evidence tolerance was increased to 1000, constant efficiency mode set to true and sampling efficiency set to 1 as shown in Appendix A.10 [Minimizing Prior Values Experiment 7](#)

The same algorithm was run using the same prior values as shaded grey in the parameter plots below. This run resulted in the following best fit values, reducing the number of likelihood evaluations from  $\sim 760\,000$  in Experiment 5 to  $\sim 250\,000$  in the current Experiment 7.

Setting these parameters as described resulted in the fits shown in Fig 4.16

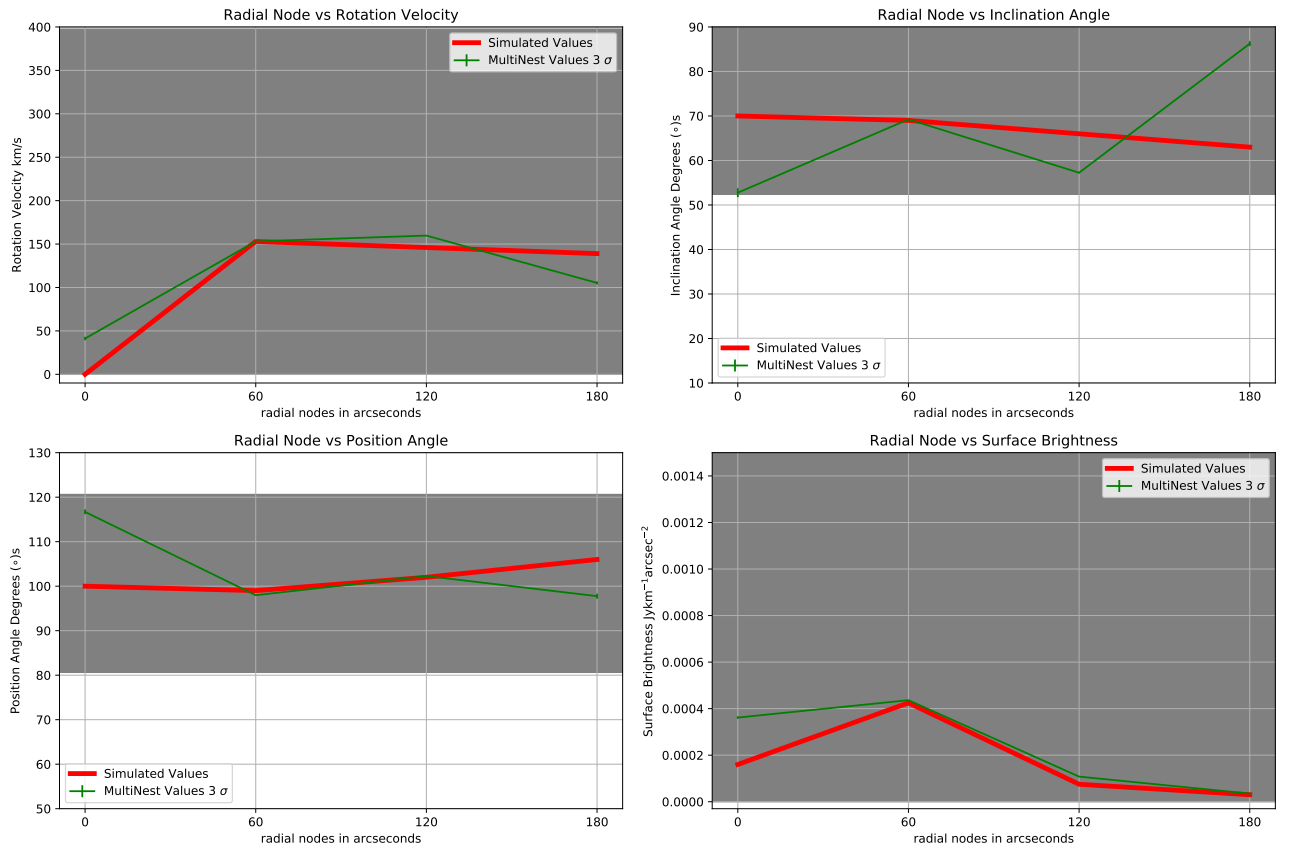


FIGURE 4.16: Comparison between simulated parameter values, red, Vs MultiNest fitted values, green, for TiRiFiC-MultiNest fit with sampling efficiency = 1; const efficiency mode = True; evidence tol = 1 000. Best fit plot for rotation velocity. Error bars are shown at  $3\sigma$ .



For visual comparison between the simulated and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.17 was made using SlicerAstro ([Punzo et al. 2017](#)).

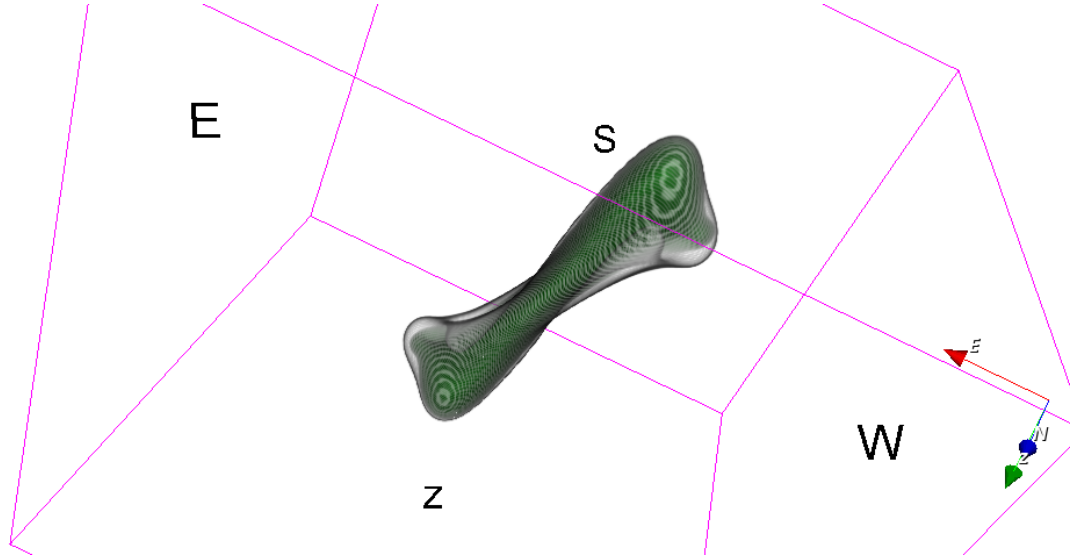


FIGURE 4.17: Comparison between NGC 4062 TiRiFiC simulated cube, green, and TiRiFiC-MultiNest best fit parameter values, white.

From this run, it was observed that with the evidence tolerance parameter set so high, the fit was not good enough. One possibility for these wrong results was missing the maximum likelihood mode.

To test if this was an effect of a mode being missed as a result of 'low resolution' in the number of live points, the number of live points were increased from 500 to 5 000, and the experiment run again to give the following result

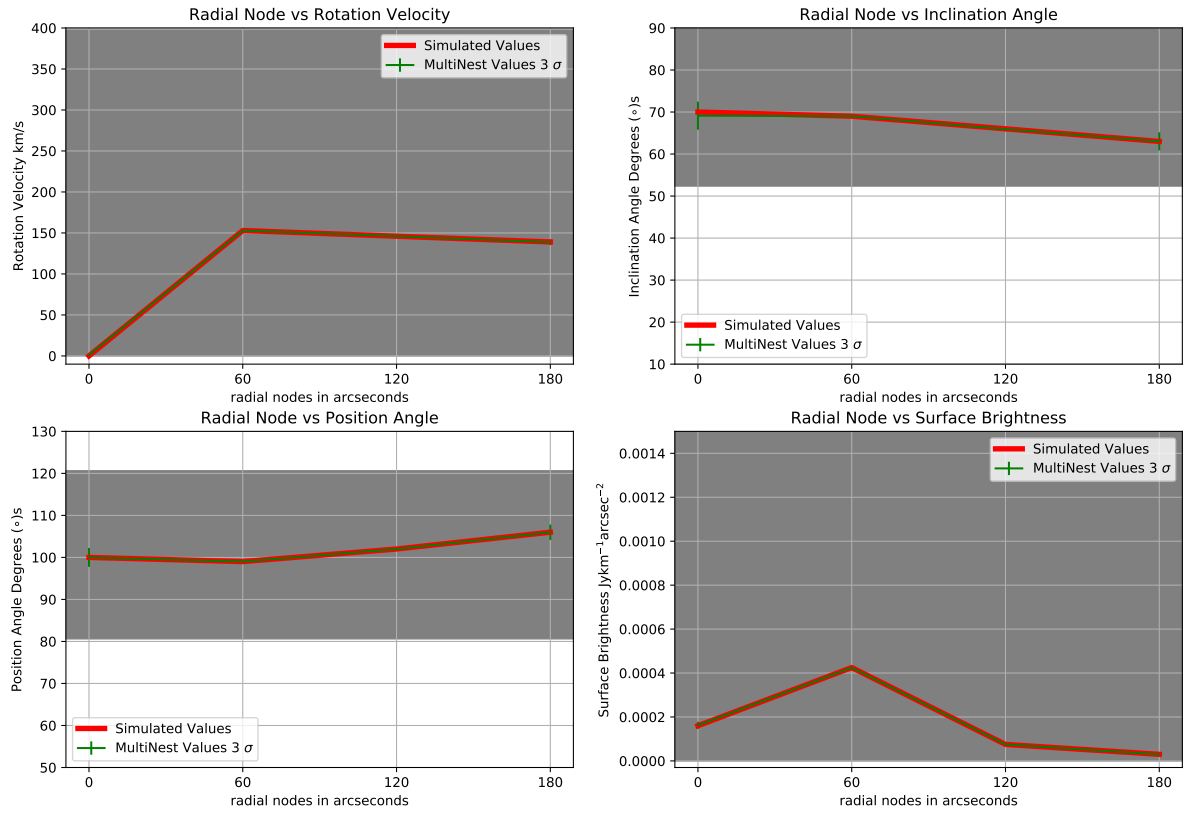


FIGURE 4.18: Comparison between simulated parameter values, red, Vs MultiNest fitted values, green, for TiRiFiC-MultiNest fit with sampling efficiency = 1; const efficiency mode = True; evidence tol = 1 000; live points=5 000. Best fit plot for rotation velocity. Error bars are shown at  $3\sigma$ .

Parameter plots in Fig 4.18 confirm that the correct mode was being missed. Comparing increasing the number of live points and setting the parameters as described in the beginning of this Experiment 7 to the previously run Experiment 5, there was a decrease in the number of likelihood evaluations from  $\sim 760\,000$  in Experiment 5 to  $\sim 720\,000$  in the current Experiment 7.

For visual comparison between the simulated and TiRiFiC-MultiNest derived data cube representing these best fit parameters shown in the Figures above, the Fig 4.19 was made using SlicerAstro (Punzo et al. 2017).

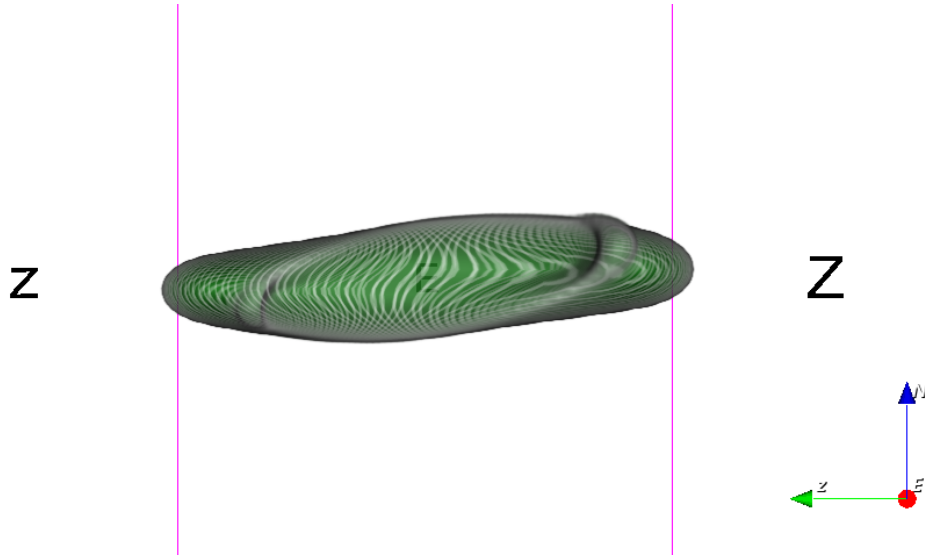


FIGURE 4.19: Comparison between NGC 4062 TiRiFiC simulated cube, green, and TiRiFiC-MultiNest best fit parameter values, white.

#### 4.2.2 Experiment 8

The next test involved running all parameters as free parameters with seven radial nodes for the simulated galaxy NGC 4062, apart from XPOS, YPOS and VSYS as shown in the cube Fig 4.20. Parameters used to construct the simulated data cube were derived from a fit to a real data cube of the same galaxy, and are represented as red lines in the plots.

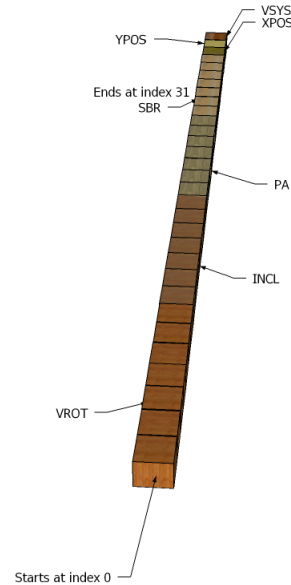


FIGURE 4.20: The cube parameter with VROT, INCL, PA and SBR set to vary with 7 nodes whereas XPOS, YPOS and VSYS are set to a constant value, which translates to one value in the cube array that is duplicated seven times per parameter value.

To begin with, a simulated data cube was made by setting parameter values in TiRiFiC's deffile this time with seven radii nodes. This simulated cube was later fitted for using SoFiA according to the procedure described in Section 4.1.1 above producing the prior values detailed in Appendix A.11 [Minimizing Prior Values Experiment 8](#) and also shaded grey in the parameter plots below.

Since this experiment involved more parameter values, 31 in total, that would be fitted simultaneously by MultiNest, an attempt was made to reduce the number of likelihood iterations by setting the following parameters.

```
sampling efficiency = 1
const efficiency mode = True
evidence tol = 1 000
n live points = 500
```

The number of live points was kept as low as possible in an attempt to reduce the number of likelihood evaluations. The new approach was also taken to divide the resultant log-likelihood value by 50 to further reduce the number of likelihood evaluations. On running the test, parameter plots in Fig 4.21 were derived.

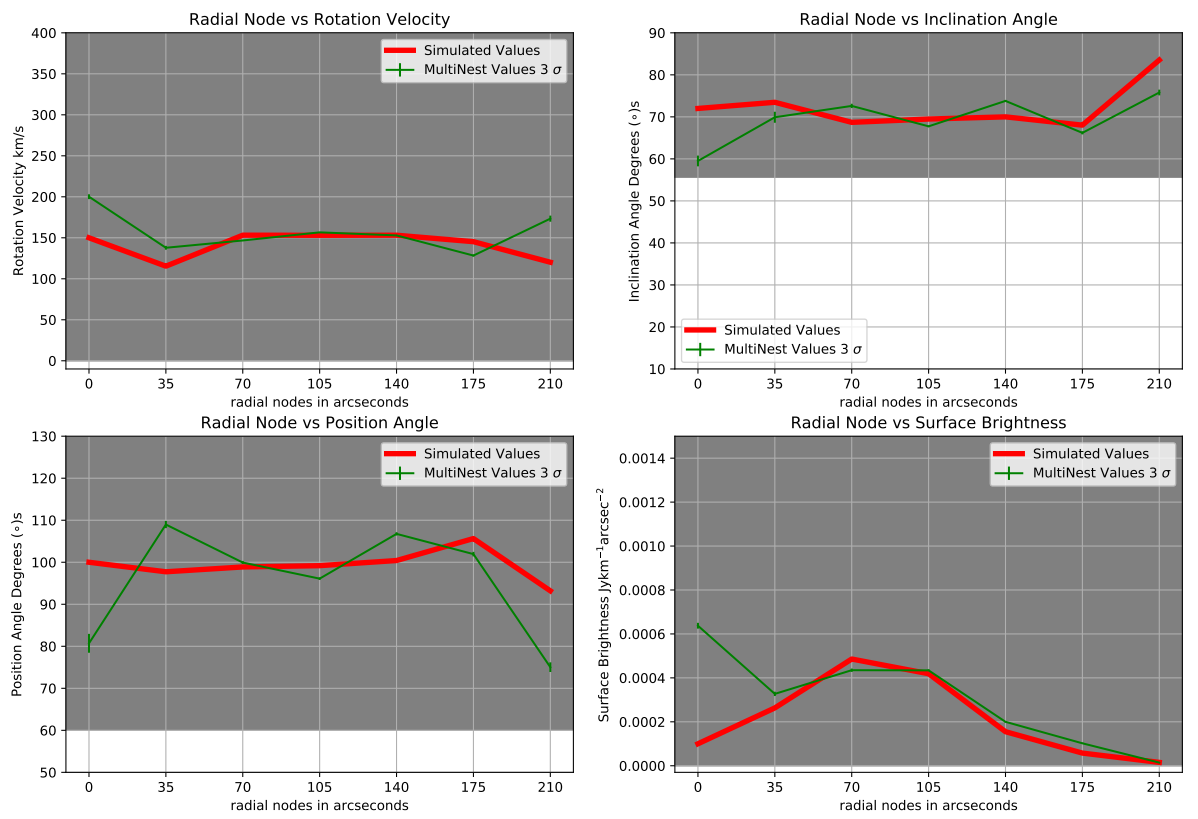


FIGURE 4.21: Best fit plot for rotation velocity, inclination angle, position angle and surface brightness. Error bars are shown at  $3\sigma$ .

For visual comparison between the simulated and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.22 was made using SlicerAstro (Punzo et al. 2017).

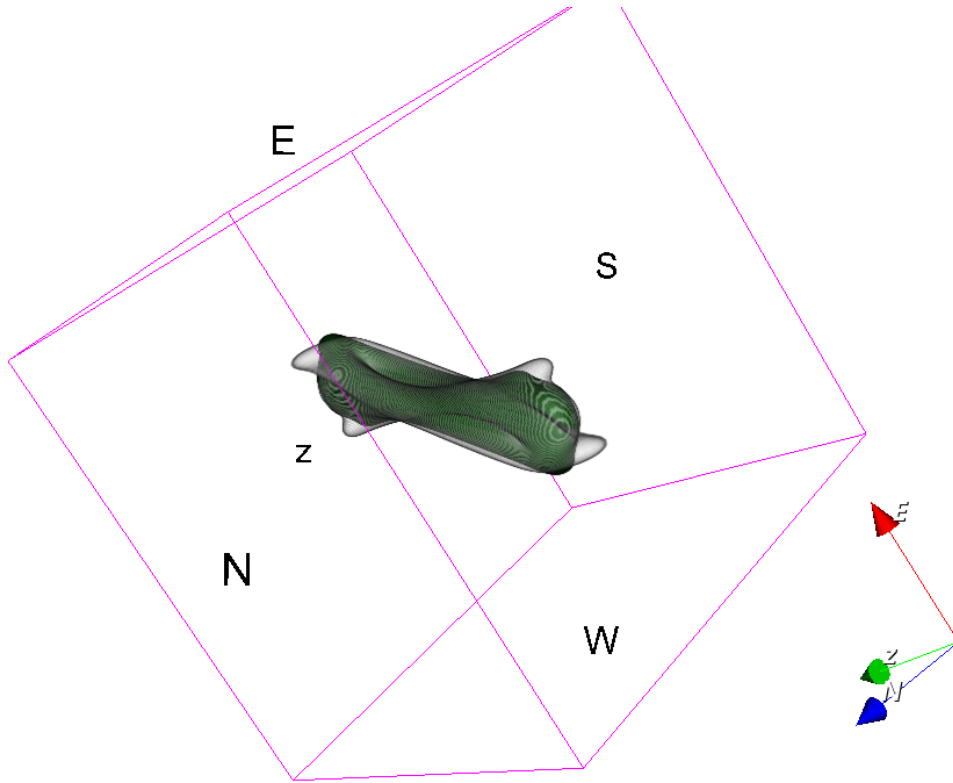


FIGURE 4.22: Comparison between NGC 4062 TiRiFiC simulated cube, green, and TiRiFiC-MultiNest best fit parameter values, white.

The fits showed bad results, but with Experiment 7 as proof of live points resolution playing a major role in finding the correct mode, the next Experiments 9 and 10 increased their number of live points in an attempt to derive correct fits by ensuring MultiNest has enough resolution to find the maximum mode, despite increased cube length.

### 4.2.3 Experiment 9

The next experiment took Experiment 8 and modified it such that there were 2000 live points. To compensate for an increased number of live points as well as more parameters to fit for, 31 in total as Experiment 8, a larger log-likelihood dividing constant of 500 was used. The same prior values were also used and this resulted in the best fit plot as shown in Fig 4.23

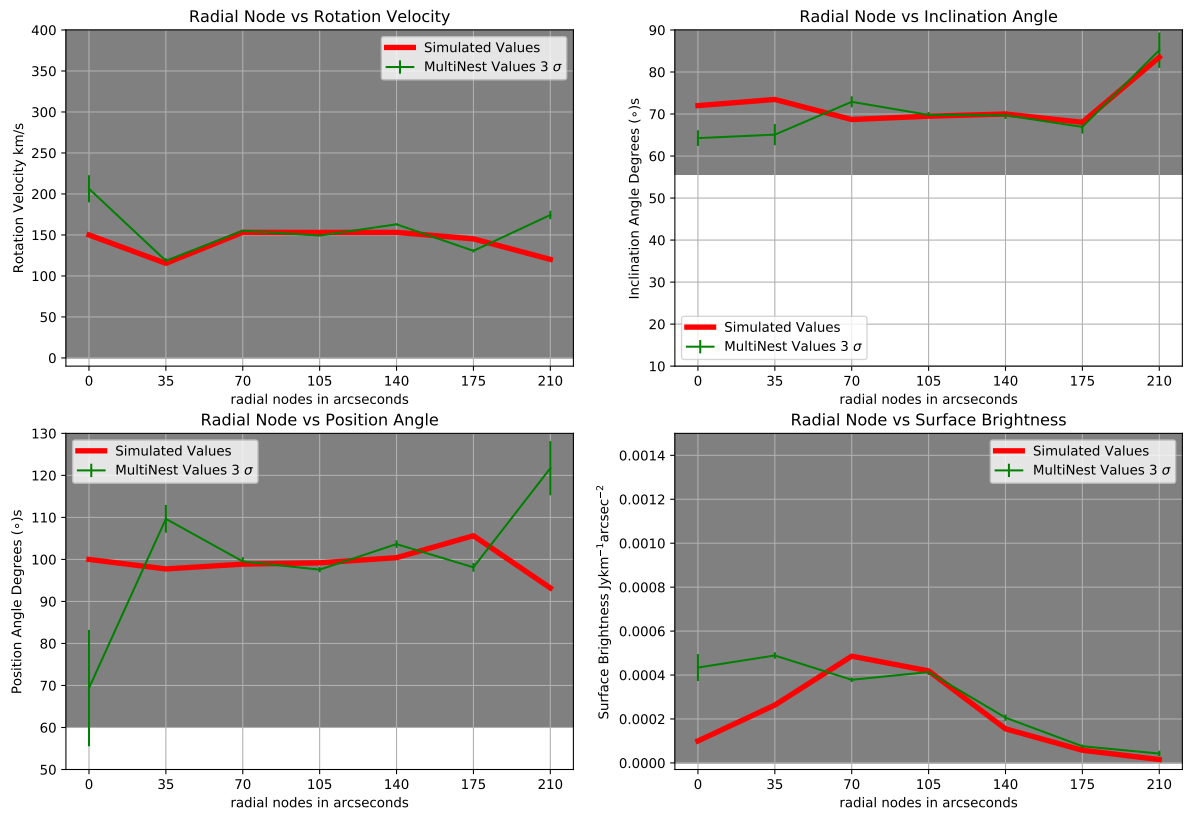


FIGURE 4.23: Best fit plot for rotation velocity, inclination angle, position angle and surface brightness. Simulated data cube parameters derived from a fit to a real data cube, and are represented as red lines whereas green lines are TiRfiC-MultiNest fits. Error bars are shown at  $3\sigma$ .

These plots were much better than those in Experiment 8 but were still inaccurate. This inaccuracy hinted at the fact that 2000 live points were still not enough owing to a larger cube length that was being fitted for in MultiNest, resulting in the next Experiment 10.

For visual comparison between the simulated and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.24 was made using SlicerAstro ([Punzo et al. 2017](#)).

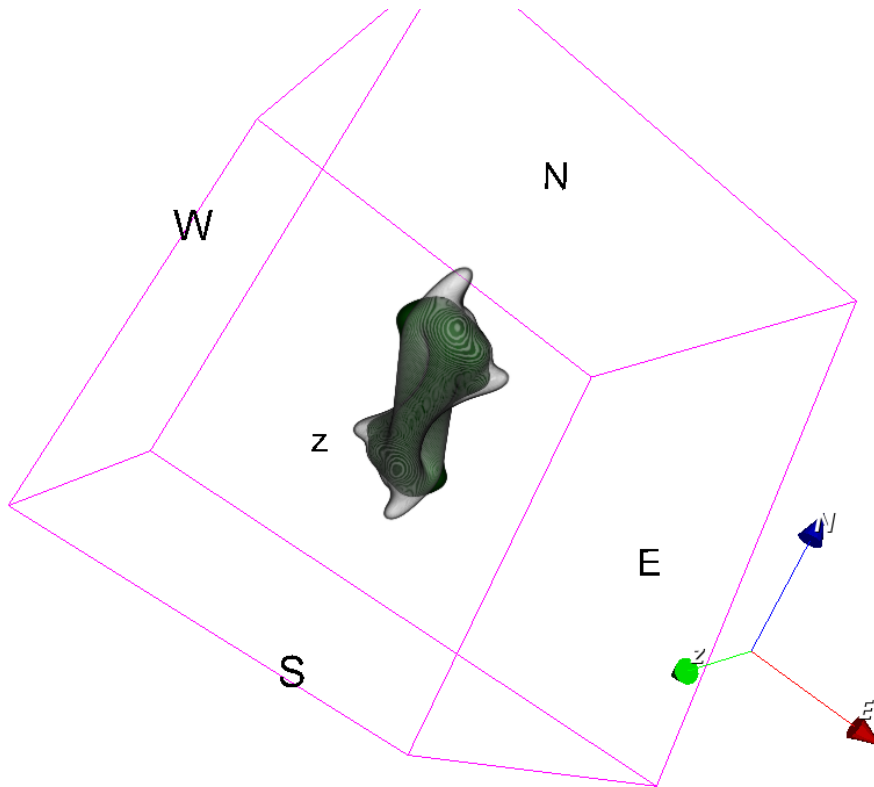


FIGURE 4.24: Comparison between NGC 4062 TiRiFiC simulated cube, green, and TiRiFiC-MultiNest best fit parameter values, white.

#### 4.2.4 Experiment 10

The suspected cause for the inaccuracy experienced in Experiment 9 was the number of live points being insufficient and hence missing the maximum likelihood mode as the live point resolution was not enough. To further test this, Experiment 9 was modified by adding 3000 more live points making a total of 5000 live points, but this time with the log-likelihood divided by 2000 in an attempt to compensate for an increased number of live points and this resulted in the following best fit values.



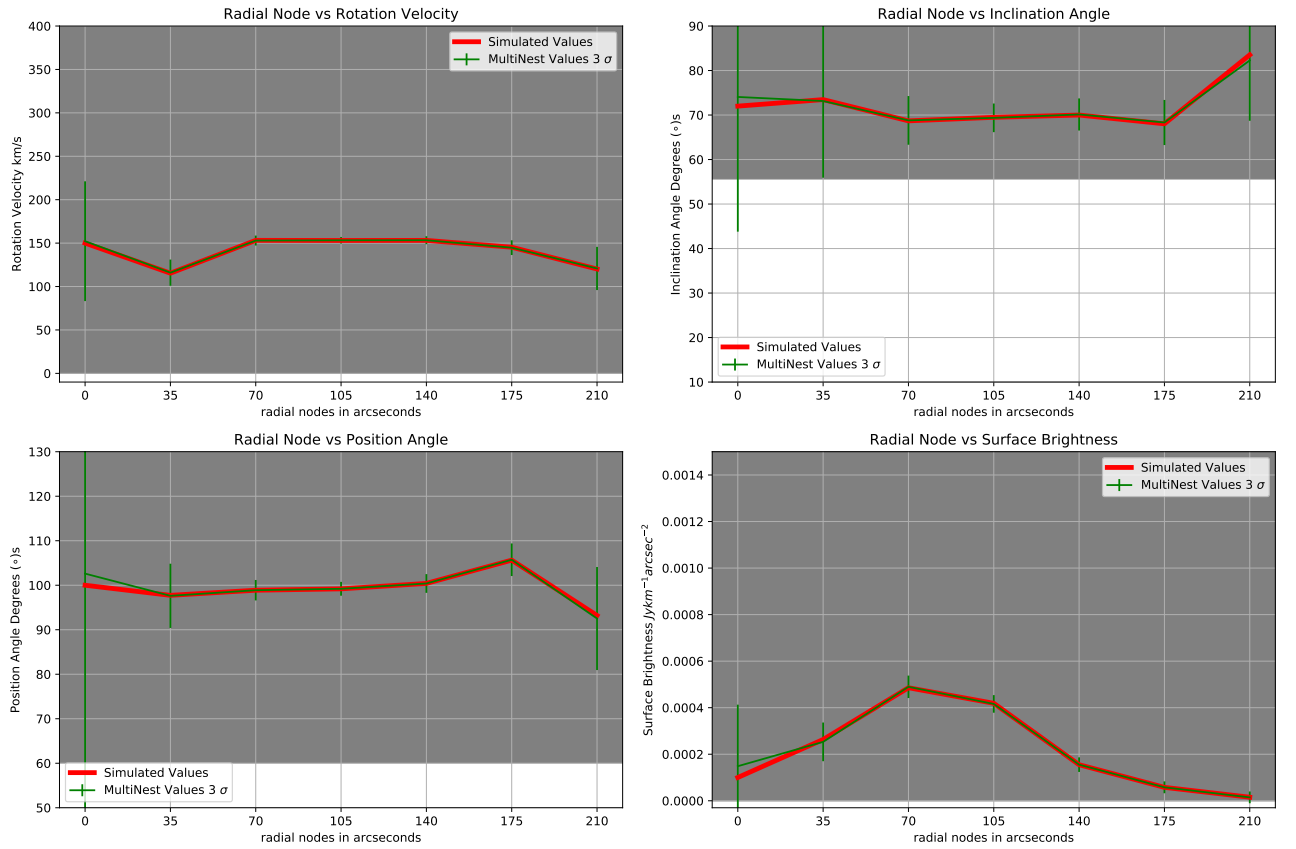


FIGURE 4.25: Best fit plot for rotation velocity, inclination angle, position angle and surface brightness. Simulated data cube parameters derived from a fit to a real data cube, and are represented as red lines whereas green lines are TiRfIC-MultiNest fits. Error bars are shown at  $3\sigma$ .

With an increased number of live points, MultiNest was able to sample the entire multimodal distribution and end up with the best fitting parameter value.

For visual comparison between the simulated and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.26 was made using SlicerAstro (Punzo et al. 2017).

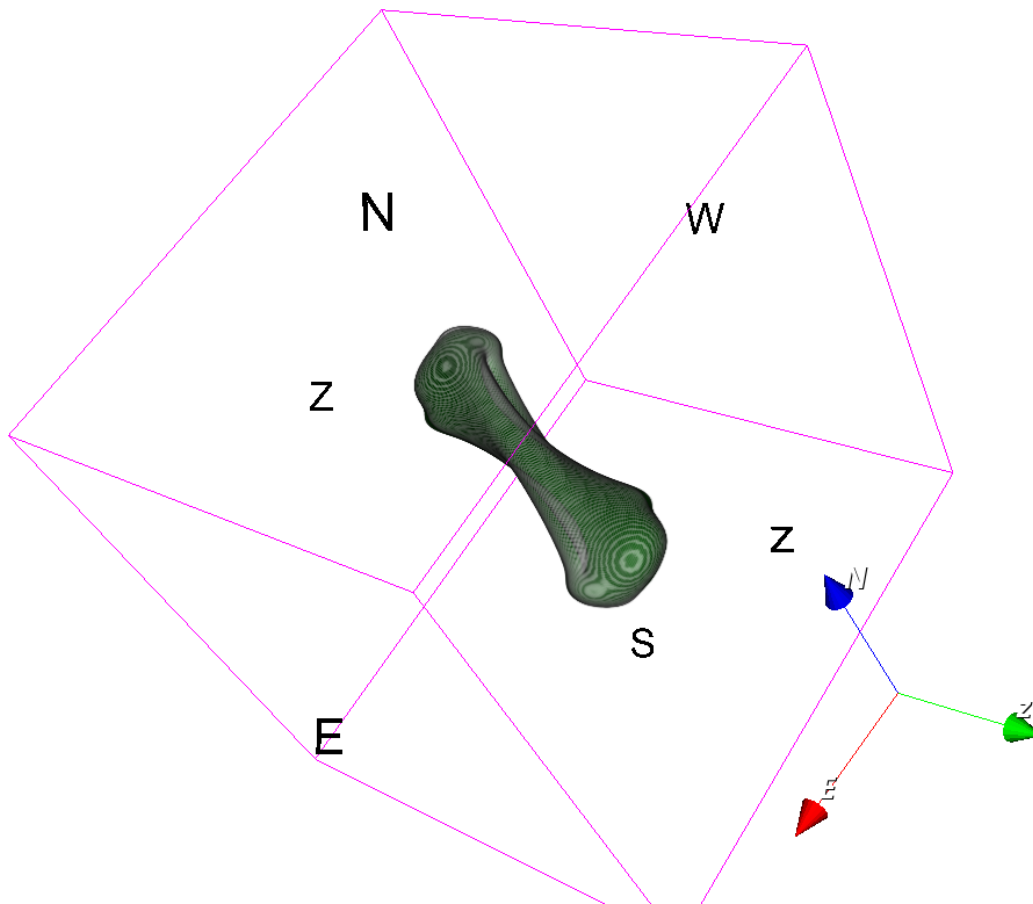


FIGURE 4.26: Comparison between NGC 4062 TiRiFiC simulated cube, green, and TiRiFiC-MultiNest best fit parameter values, white.

#### 4.2.5 Experimental runs summary

The lessons learned from the experimental runs detailed above are as follows.

- The number of live points used in MultiNest has a direct influence on the fit obtained. This is because the number of live points determines the 'resolution' of the sampling region. The higher the number of live points, the higher the resolution and consequently, the higher the chance that the optimal mode in the log-likelihood distribution will be found.

- The higher the number of parameters fitted for, the more the number of likelihood evaluations needed before MultiNest converges.
- Attempts at reducing the number of likelihood evaluations have been made to increase the fitting speed. This was done by increasing the evidence tolerance, setting efficiency to 1 and setting the constant efficiency mode to True as well as dividing the log-likelihood value by a constant value in extreme cases.

Most experimental runs were also parallelized by setting a certain number of Nodes as detailed in Table [4.2](#).

TABLE 4.2: Experiment runs summary

Exp No	Nodes	TiRiFiC Parameters	Cube length	Live Points	Evidence tol	Sampling eff	loglike dividing const	Likelihood iterations
1	1	0	7	35	1	0.5	1	14 868
2	4	2	13	500	1 000	0.5	1	82 799
3	4	2	13	500	1 000	0.5	1	156 193
4	4	2	13	500	1 000	0.5	1	379 179
5	4	4	19	500	1 000	0.5	1	759 179
6	7	2	19	500	1 000	0.5	1	5 502 746
71	4	4	19	500	1 000	1.0	1	249 544
72	4	4	19	5000	1 000	1.0	1	719 621
8	7	4	31	500	1 000	1.0	100	108 929
9	7	4	31	2000	1 000	1.0	500	385 713
10	7	4	31	5000	1 000	1.0	2 000	942 901

#### 4.2.6 Test with real galaxy NGC 3198

The THINGS galaxy NGC 3198 ([Walter et al. 2008a](#)) was reduced in size as described in Section 4.1.3. SoFiA was then used to derive the best-fit parameters, from the reduced data cube, using the procedure described in Subsection 4.1.1 from which the prior values shown in Appendix A.14 [Minimizing Prior Values Test with real galaxy N3198](#) were derived. Later, TiRiFiC-MultiNest algorithm was used to fit the reduced NGC 3198 galaxy cube using only four radial nodes, making a cube length of 19. This reduced and fitted THINGS datacube was compared to a fit by [Gentile et al. \(2013\)](#), made to a HALOGAS datacube of the same galaxy.

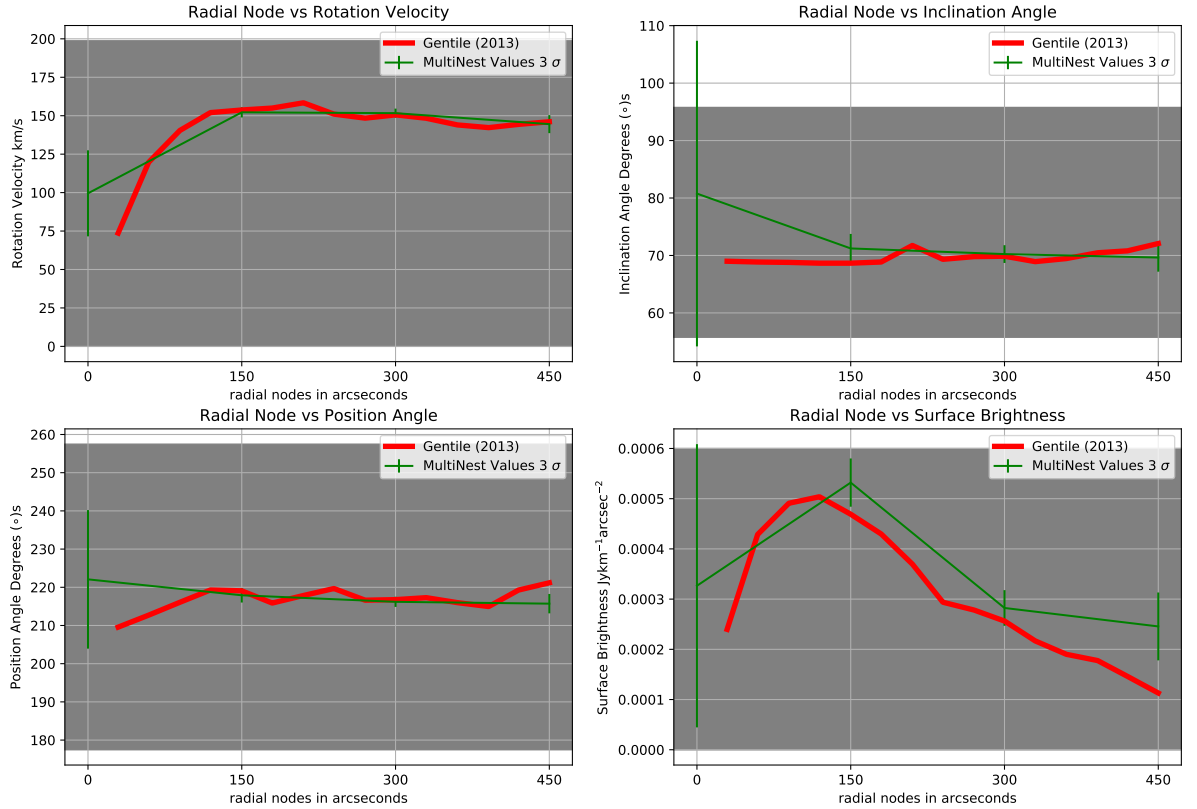


FIGURE 4.27: Best fit plot for rotation velocity, inclination angle, position angle and surface brightness. Error bars are shown at  $3\sigma$ .

As can be seen in Fig 4.27, the TiRiFiC-MultiNest fit was compared to [Gentile et al. \(2013\)](#) parametrisation.

For visual comparison between the real and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.28 was

made using SlicerAstro ([Punzo et al. 2017](#)).

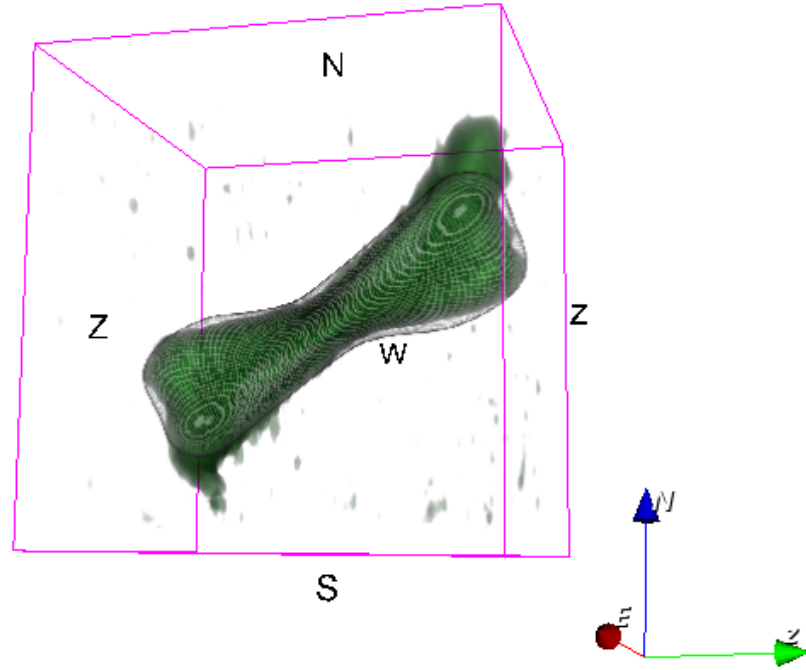


FIGURE 4.28: Comparison between NGC 3198 observed cube, green, and TiRiFiC-MultiNest best fit parameter values, white.

#### 4.2.7 Test with real galaxy NGC 4062

Galaxy NGC 4062 ([Heald et al. 2011](#)) was fit for using both TiRiFiC and TiRiFiC-MultiNest resulting in the red and green fits, shown in Fig 4.29. These fits were made using 1 disk and 7 radial nodes as described in 4.20 making the total cube length become 31. SoFiA was used to derive the best-fit parameters using instructions detailed in Subsection 4.1.1. From these best-fit parameters, the prior values for NGC 4062 were derived as shown in Appendix A.15 [Minimizing Prior Values Test with real galaxy N4062](#). Running this resulted in the following best-fit plots.

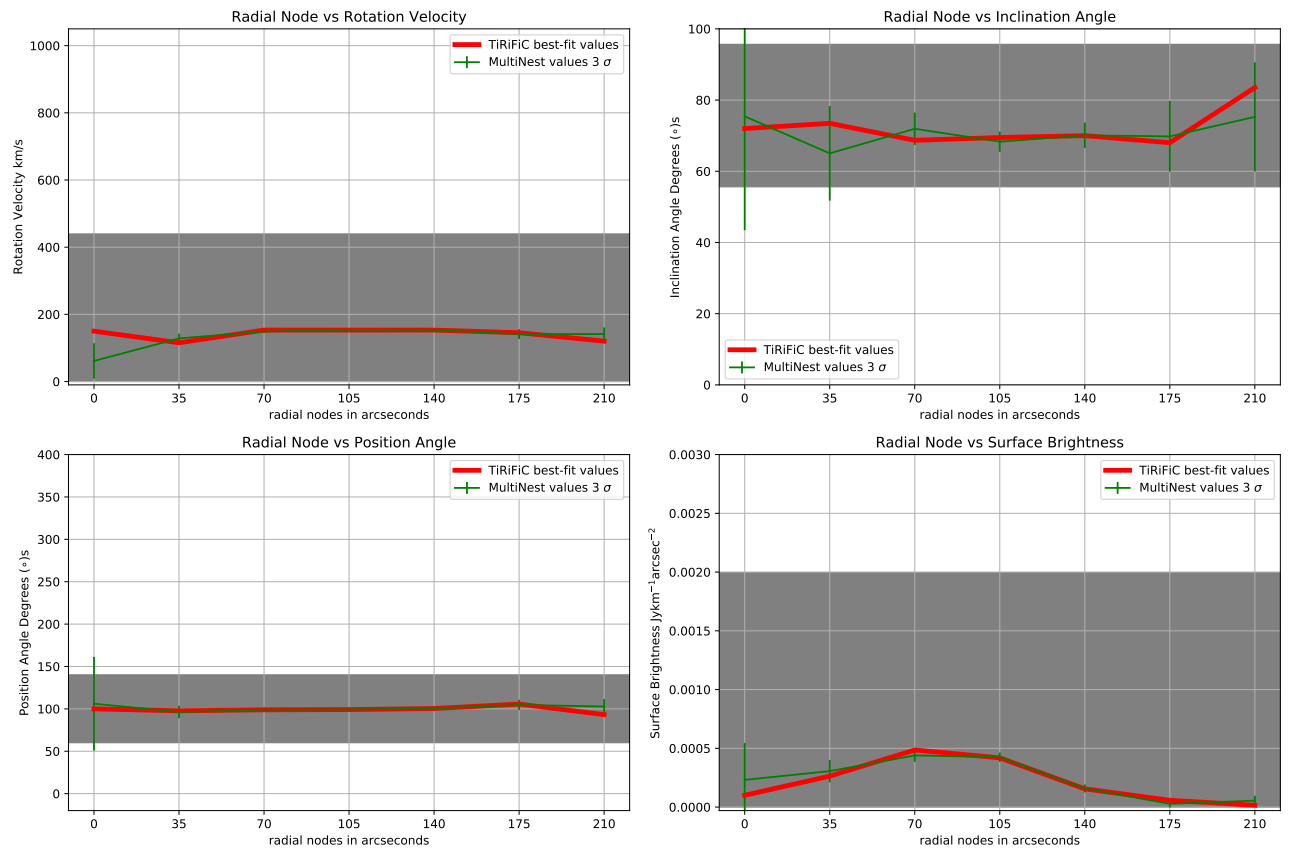


FIGURE 4.29: Best fit plot for rotation velocity, inclination angle, position angle and surface brightness. Error bars are shown at  $3\sigma$ .

For visual comparison between the real and TiRiFiC-MultiNest derived data cube representing these best-fit parameters shown in the Figures above, the Fig 4.30 was made using SlicerAstro ([Punzo et al. 2017](#)).

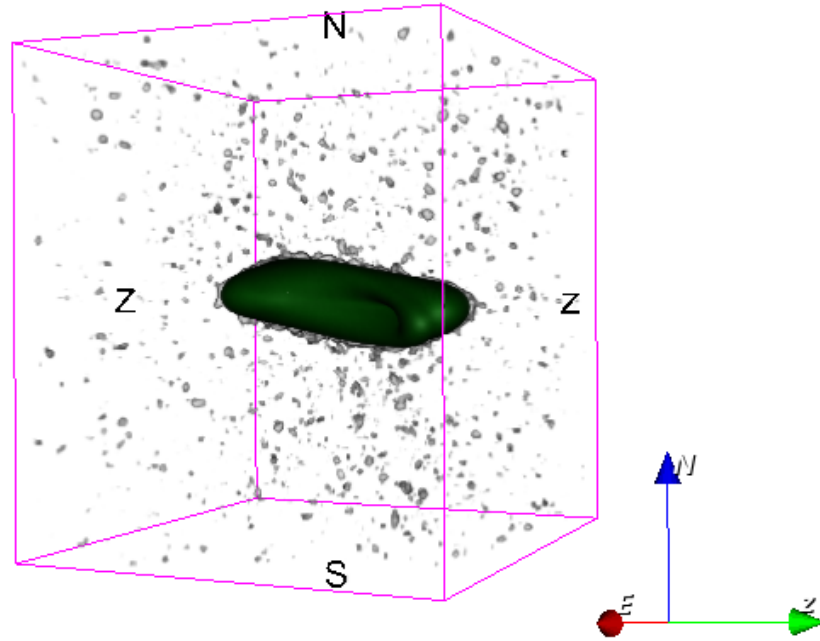


FIGURE 4.30: Comparison between NGC 4062 observed cube, white, and TiRiFiC-MultiNest best fit parameter values, green.



## Chapter 5

# Summary and Conclusion

### 5.1 Implementation Summary

TiRiFiC is a HI galaxy parametrisation software that does a direct fit to data cubes using Chi-Square minimisation. However, the currently implemented minimising algorithms do not have a way of quantifying errors or estimating the correlation between parameter values. These minimising algorithms may also get stuck in a local mode when the Chi-Square is multimodal, leading to incorrect parameters being returned by the software. To navigate this multimodal Chi-Square space, a Bayesian-based algorithm, MultiNest, has been implemented in combination with TiRiFiC and acts as a minimising algorithm for TiRiFiC. The Bayesian-based sampler, MultiNest, also enables estimation of parameter values and correlations between parameters to be found.

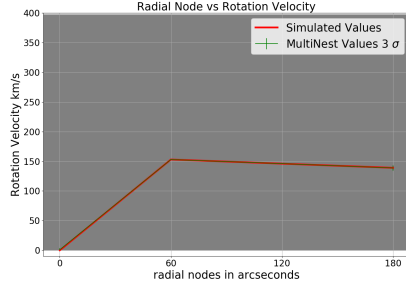
Since TiRiFiC-MultiNest algorithm is Bayesian-based, it takes a default flat prior which define the possible ranges of TiRiFiC's parameter values. The likelihood is calculated from TiRiFiC's  $\chi^2$  value and later converted to log-likelihood value with every likelihood evaluation from MultiNest.

A Python wrapper around TiRiFiC was written to work together with the python based Bayesian inference engine, PyMultiNest ([Buchner & Johannes 2016](#)), which is a python wrapper around the c code, MultiNest. Together, the TiRiFiC-MultiNest combination is parallelised to speed up the fitting process

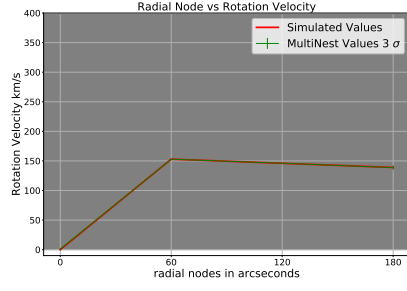
To test the implementation, a simulated observation, i.e. a data cube with known parameter values using TiRiFiC, was made. The simulated data cube was then fitted using TiRiFiC-MultiNest. It was shown that the fit results agree with the simulated parameters (to within two standard deviations ( $\sigma_{\text{rms}}$ ) while providing plausible errors and showing correlations between parameter values. It was also noted that in the event a bad fit is recovered, it is advisable to try increasing the number of live points which would result in a better fit. This evidence was supported by Experiment 7, which showed the benefits of having a sufficient number of live points. A fit to a real galaxy NGC3198 VLA observation was also made.

It was also observed that by dividing the log-likelihood by a large constant value, the number of likelihood evaluations decreased, making testing and or fitting for a galaxy to happen in a much shorter timeframe. A similar reduction in MultiNest running time was observed by increasing the evidence tolerance and setting the constant efficiency mode parameter to true. However, this also required an increase in the number of live points to derive the best fit value.

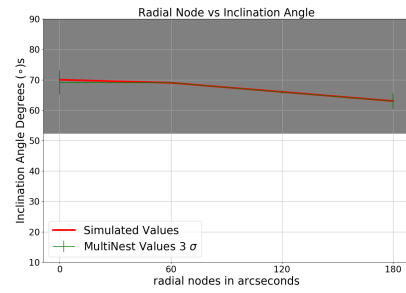
A comparison between the non-accelerated fit Experiment 5 and the accelerated fit Experiment 7 is made.



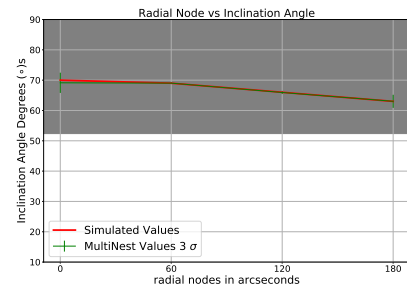
(A) rotation velocity Exp 5



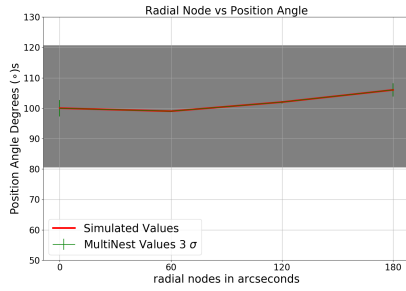
(B) rotation velocity Exp 7



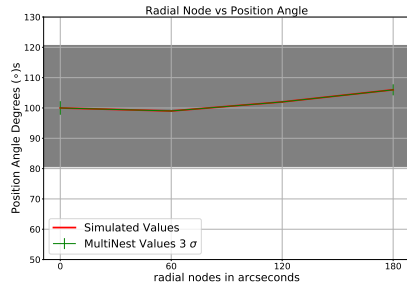
(C) Inclination angle Exp 5



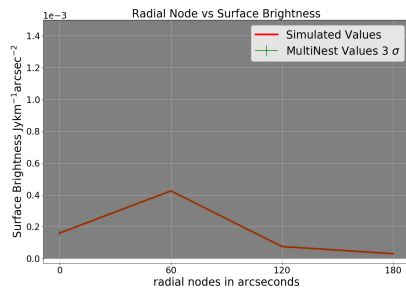
(D) Inclination angle Exp 7



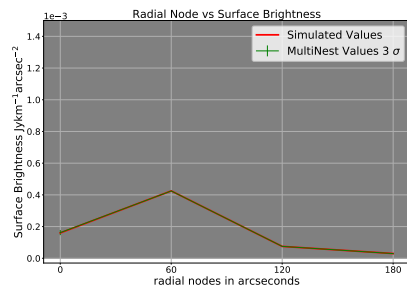
(E) position angle Exp 5



(F) position angle Exp 7



(G) surface brightness Exp 5



(H) surface brightness Exp 7

FIGURE 5.1: Experiment 5 had the parameters sampling efficiency = 1; const efficiency mode = True; evidence tol = 1 000; live points=500, while Experiment 7 had the parameter sampling efficiency = 1; const efficiency mode = True; evidence tol = 1 000; live points=5 000

From Experimental runs 5, the non-accelerated, and 7, the accelerated version, there seem to be a reduction in the error magnitude with acceleration as shown in the Figs 5.1a to 5.1h above. This seems counter-intuitive and implies that the acceleration attempts made influence the spread of posterior samples coming from MultiNest and consequently, the error values obtained. The cause of this reduction in error values with acceleration will be investigated as part of the future work.

The sampling efficiency parameter was set to 1 in experiments 7 through 10 as advised by Feroz et al. (2008) to reduce the number of likelihood evaluations by roughly sampling the posterior as the primary interest is the parameter samples and not necessarily the evidence value.

The constant efficiency mode was set to *True* in an attempt to maintain the sampling efficiency as close to unity as possible as specified by MultiNest README instructions.

Having a high evidence tolerance value had the effect of reducing the number of likelihood evaluations despite a required increase in the number of live points needed to resolve the posterior.

Correlation plots between TiRiFiC parameters were also recovered and are shown in Appendix A. An example of such correlation plots is shown in Fig 5.2 below.

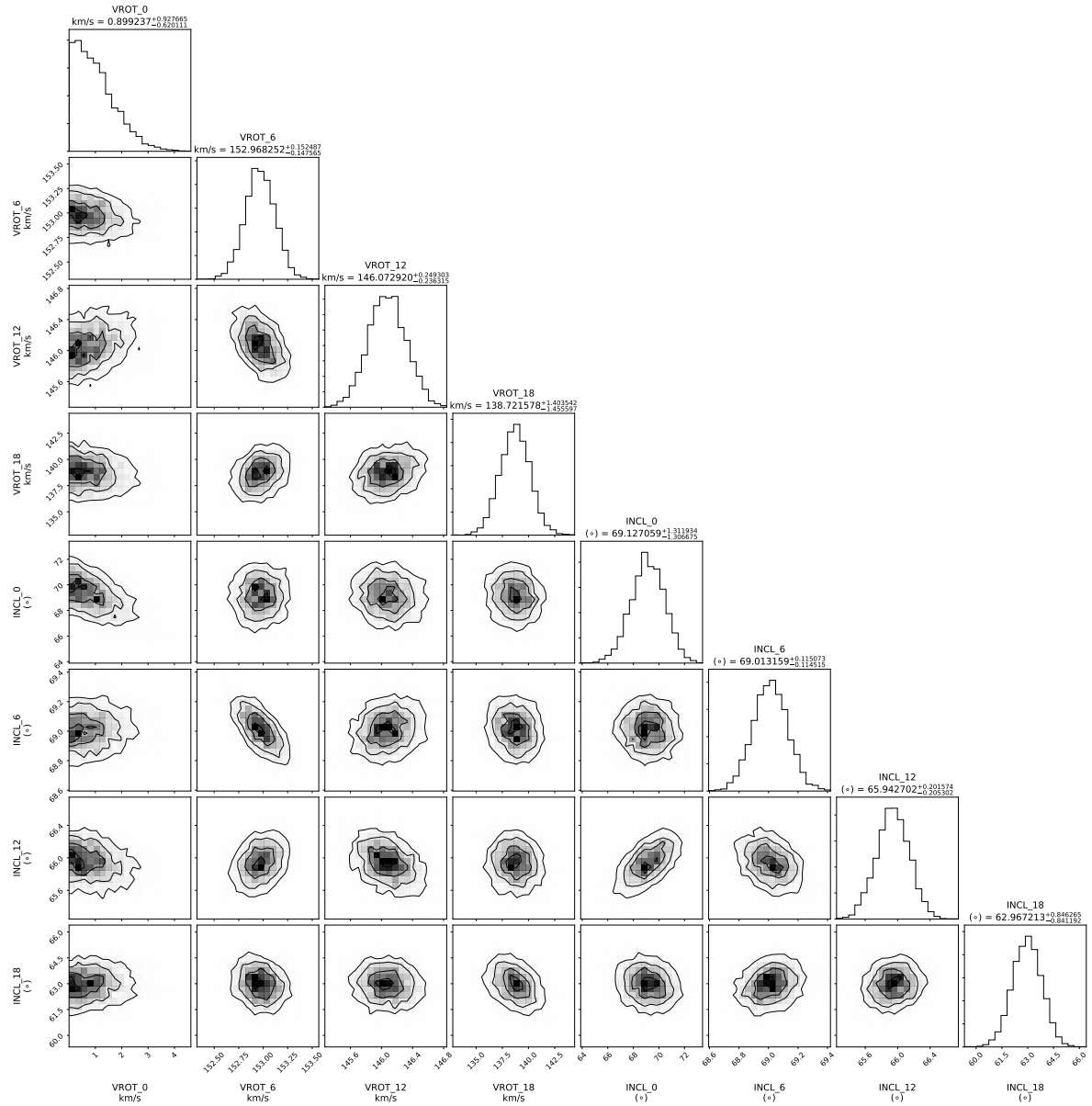


FIGURE 5.2: Plots showing the correlation between rotation velocity and Inclination angle.

From Fig 5.2, it can be seen that the rotation velocity is negatively correlated to the inclination angle at the same radial node as was shown by Begeman (1989). Interestingly, the plots show also that some parameters are correlated to neighbouring data points of the same parameter type. Prominently in Fig 5.2 this is the case for the rotation velocities VROT at radii 6 and 12. It has often been remarked (Józsa, priv. comm.) that TiRiFiC tends to settle on a sawtooth pattern for some parameters, i.e. inclination and rotation velocity. These plots show the reason for this, that is, correlation between parameters of the same type. Apparently, it is possible to balance the deviation of a parameter from the optimal fit through implementing a deviation in the opposite direction of a neighbouring parameter of the same type (e.g. two adjacent rotation velocities). Within the scope of the experiments performed in this thesis, there was no minimum issues encountered, that would manifest as two islands of solutions in the chi-squared.

## 5.2 Conclusion

TiRiFiC-MultiNest at its current state, can reach a very good result when fitting a data cube resulting from a galaxy observation. However, its fitting time increases exponentially with an increase in parameters to be fitted for.

The next step for TiRiFiC-MultiNest algorithm involves determining the point of fitting breakdown and influence in error magnitude of parameter values as the evidence tolerance is increased.

TiRiFiC-MultiNest would also need to take into account the correlation between pixel values. It is also important to note that a potential barrier to this technique of Chi-Square minimization is that the calculated value of the Chi-Square is dependent on the size of the cube.

## Appendix A

# Software Documentation

### A.1 TiRiFiC Usage

In n4565 hi.def file, The variables used are as follows.

**PROGRESSLOG**: The file name used to record tirific's calculated chisquared after every loop.

**INSET**: The input fits file.

**OUTSET**: The output fits file.

**NDISKS= 2**: We are fitting 2 disks.

**NUR= 11**: Represents all 11 radial nodes.

**LOOPS=4**: Total number of loops to run.

**VARY= ! VROT 1:11 VROT 2 1:11**: Vary was set to only vary rotation velocity, varying VROT and VROT 2 as a group for each and every of the 11 parameters.

To run, on the terminal, execute the following command after ensuring n4565 hi.fits and n4565 hi.def are on the same folder as your terminal path.

***tirific deffile=n4565 hi.def***

A model datacube *n4565 hi out01.fits* is produced along with ps file inside the same directory.

### A.2 Multinest Example

An example problem solved with Multinest is as follows:

The file example data.txt contains data whose first column is x values, the second y values and the third the error on each data point.

The function used to fit is of this form:

$$y = a \times x \times \sin(b \times x + c)$$

Multinest was used to find the best fitting parameters  $a$ ,  $b$  and  $c$  and their uncertainties.

### A.3 Parameters set for SoFiA run

SoFiA was used to get the best-fit parameters for the galaxy, with the following parameters

Source finding

Smooth + clip finder

Enable: Yes

Threshold - 4.0

Edge mode: Constant

RMS mode: Gaussian fit to negative fluxes

Kernel units: Pixels

Kernels: "[[ 0, 0, 0,'b'],[ 0, 0, 3,'b'],[ 0, 0, 7,'b'],[ 0, 0, 15,'b'],[ 3, 3, 0,'b'],[ 3, 3, 3,'b'],[ 3, 3, 7,'b'],[ 3, 3, 15,'b'],[ 6, 6, 0,'b'],[ 6, 6, 3,'b'],[ 6, 6, 7,'b'],[ 6, 6, 15,'b']]"

Merging

Merging of Detectors

Enable: yes

Radius X, Radius Y, Radius Z = 3

Min size X, Min size Y, Min size Z = 4

Parametrization

Source parametrization

Enable: Yes

Optimise mask(Ellipse): yes

Fit Busy Function: yes

Reliability calculation: selected

### A.4 Minimizing Prior Values Experiment 1

Prior values used for this experiment are as follows:

$v_{\text{rot max}} = 1042.8 \text{ km s}^{-1}$

$v_{\text{rot min}} = 0.0 \text{ km s}^{-1}$

$\text{incl max} = 92.4^\circ$

$\text{incl min} = 52.4^\circ$

$\text{PA max} = 200.0^\circ$

$\text{PA min} = 0.0^\circ$



$\text{SBR max} = 0.002 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$   
 $\text{SBR min} = 0.0 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$   
 $\text{XPOS max} = 181.02671^\circ$   
 $\text{XPOS min} = 181.00671^\circ$   
 $\text{YPOS max} = 31.905^\circ$   
 $\text{YPOS min} = 31.885^\circ$   
 $\text{VSYS max} = 787.31 \text{ km s}^{-1}$   
 $\text{VSYS min} = 747.31 \text{ km s}^{-1}$

The following parameters were also set

Multinest Parameters:

$\text{n live points} = 35$   
 $\text{evidence tolerance} = 0.5$   
 $\text{sampling efficiency} = 1.0$   
 $\text{const efficiency mode} = \text{True}$

TiRiFiC Parameters

$\text{CFLUX} = 1\text{e-}04$

$\text{Number of multinest processes run} = 5$

$\text{Number of likelihood evaluations} = 14868$

## A.5 Minimizing Prior Values Experiment 2

The prior ranges were adjusted as follows

$\text{VROT1, VROT2 best-fit value from the run before} \pm 15$

$\text{SBR1, SBR2 best-fit value from the run before} \pm 90$

$\text{INCL (best-fit value from run before)} \pm 20$

$\text{PA (best-fit value from run before)} \pm 20$

$\text{XPOS (best-fit value from run before)} \pm 0.005$

$\text{YPOS (best-fit value from run before)} \pm 0.005$

$\text{VSYS (best-fit value from run before)} \pm 5$

Prior values used for this experiment are as follows.

$\text{vrot max} = 166.83 \text{ km s}^{-1}$

$\text{vrot min} = 136.83 \text{ km s}^{-1}$

$\text{incl max} = 92.65^\circ$

$\text{incl min} = 52.65^\circ$

PA max = 120.04 °  
 PA min = 80.04 °  
 SBR max = 0.00047 Jy km s<sup>-1</sup> arcsec<sup>-2</sup>  
 SBR min = 2.5e - 05 Jy km s<sup>-1</sup> arcsec<sup>-2</sup>  
 XPOS max = 181.02 °  
 XPOS min = 181.01 °  
 YPOS max = 31.90 °  
 YPOS min = 31.89 °  
 VSYS max = 769.16 km s<sup>-1</sup>  
 VSYS min = 759.16 km s<sup>-1</sup>

The following parameters were also set

Multinest Parameters:

n live points = 500  
 evidence tolerance = 0.5  
 sampling efficiency = 1000.0  
 const efficiency mode=False

TiRiFiC Parameters

CFLUX= 2e-06

Number of multinest processes run = 5  
 Number of likelihood evaluations = 82799

## A.6 Minimizing Prior Values Experiment 3

Prior values used for this experiment are as follows.

vrot max = 800 km s<sup>-1</sup>  
 vrot min = 0 km s<sup>-1</sup>  
 incl max = 90 °  
 incl min = 0 °  
 PA max = 359 °  
 PA min = 0 °  
 SBR max = 0.002 Jy km s<sup>-1</sup> arcsec<sup>-2</sup>  
 SBR min = 0 Jy km s<sup>-1</sup> arcsec<sup>-2</sup>  
 XPOS max = 181.021 °  
 XPOS min = 181.011 °  
 YPOS max = 31.901 °  
 YPOS min = 31.891 °

VSYS max = 769.164 km s<sup>-1</sup>

VSYS min = 759.164 km s<sup>-1</sup>

The following parameters were also set

Multinest Parameters:

n live points = 500

evidence tolerance = 0.5

sampling efficiency = 1000.0

const efficiency mode=False

TiRiFiC Parameters

CFLUX= 2e-06

Number of multinest processes run = 5

Number of likelihood evaluations = 156193

## A.7 Minimizing Prior Values Experiment 4

Prior values used for this experiment are as follows.

SBR best-fit value  $\times 10$

INCL best-fit value  $\pm 20$

PA best-fit value  $\pm 20$

XPOS best-fit value  $\pm 0.01$

YPOS best-fit value  $\pm 0.01$

VSYS best-fit value  $\pm 20$

vrot max = 397.584 km s<sup>-1</sup>

vrot min = 0.000 km s<sup>-1</sup>

incl max = 92.377 °

incl min = 52.377 °

PA max = 120.619 °

PA min = 80.619 °

SBR max = 0.002 Jy km s<sup>-1</sup> arcsec<sup>-2</sup>

SBR min = 0.0 Jy km s<sup>-1</sup> arcsec<sup>-2</sup>

XPOS max = 181.02671 °

XPOS min = 181.00671 °

YPOS max = 31.90526 °

YPOS min = 31.88526 °

VSYS max = 787.3097 km s<sup>-1</sup>

VSYS min = 747.3097 km s<sup>-1</sup>

The following parameters were also set

Multinest Parameters:

n live points = 500

evidence tolerance = 0.5

sampling efficiency = 1000.0

const efficiency mode=False

TiRiFiC Parameters

CFLUX= 2e-06

Number of multinest processes run = 5

Number of likelihood evaluations = 379179

## A.8 Minimizing Prior Values Experiment 5

Prior values used for this experiment are as follows.

vrot max = 397.584 km s<sup>-1</sup>

vrot min = 0.0 km s<sup>-1</sup>

incl max = 92.377 °

incl min = 52.377 °

PA max = 120.619 °

PA min = 80.619 °

SBR max = 0.0021 Jy km s<sup>-1</sup> arcsec<sup>-2</sup>

SBR min = 0.0 Jy km s<sup>-1</sup> arcsec<sup>-2</sup>

XPOS max = 181.02671 °

XPOS min = 181.00671 °

YPOS max = 31.90526 °

YPOS min = 31.88526 °

VSYS max = 787.3097 km s<sup>-1</sup>

VSYS min = 747.3097 km s<sup>-1</sup>

The following parameters were also set

Multinest Parameters:

n live points = 500

evidence tolerance = 0.5

sampling efficiency = 1000.0

const efficiency mode=False

TiRiFiC Parameters

CFLUX= 2e-06

Number of multinest processes run = 5

Number of likelihood evaluations = 759179

And their corresponding correlation plots

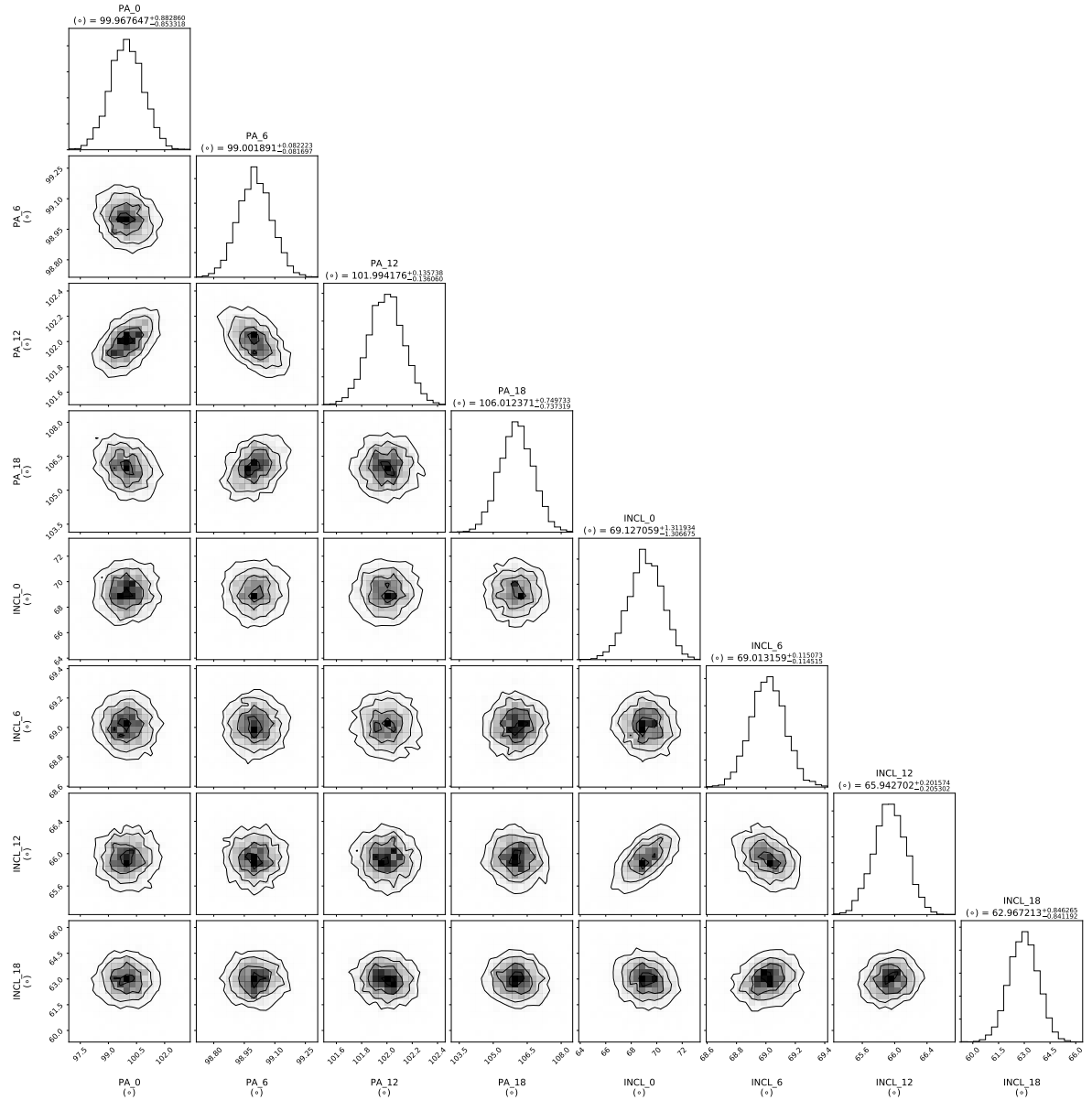


FIGURE A.1: Plots showing the correlation between position angle and inclination angle.

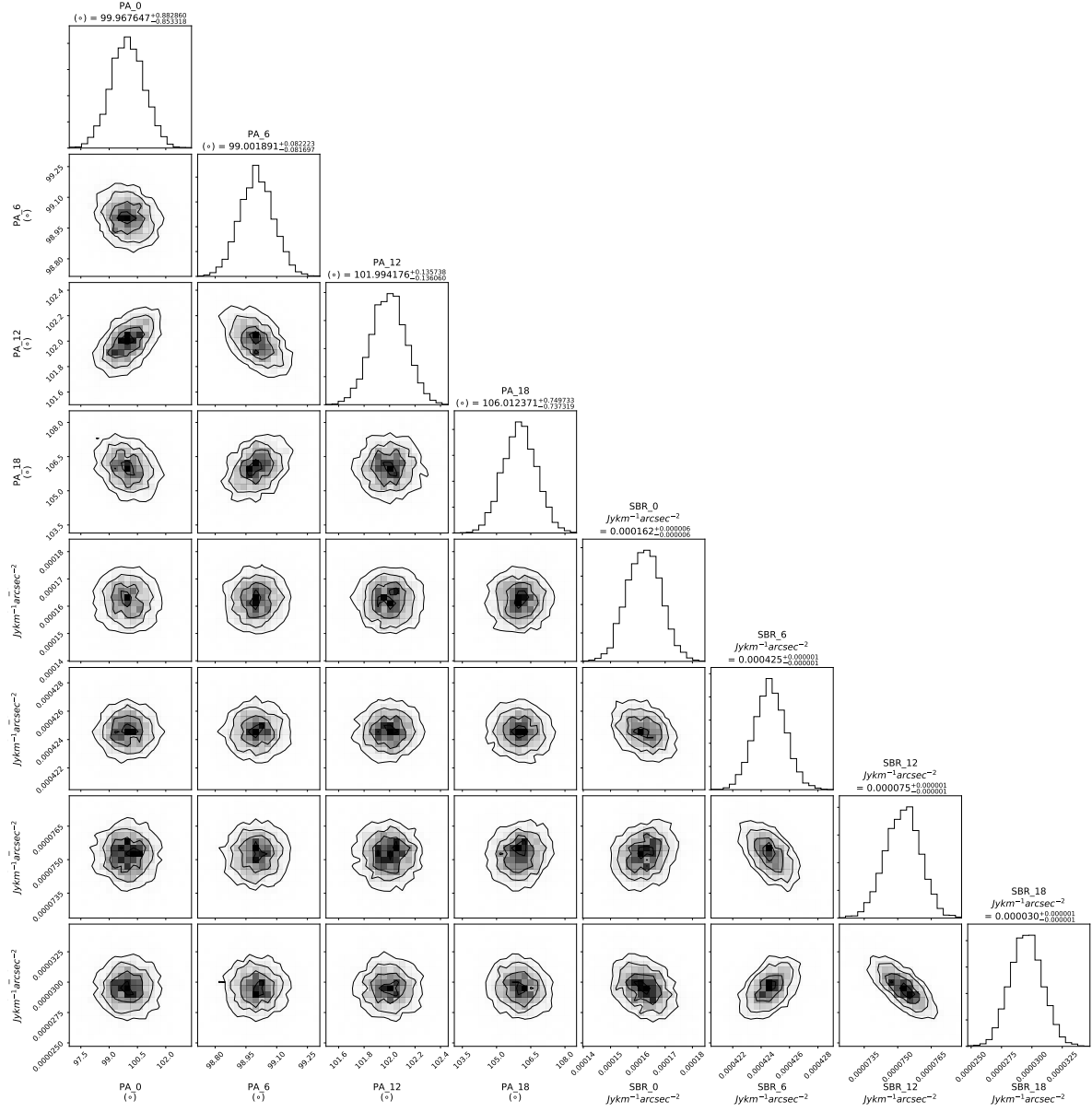


FIGURE A.2: Plots showing the correlation between position angle and surface brightness.

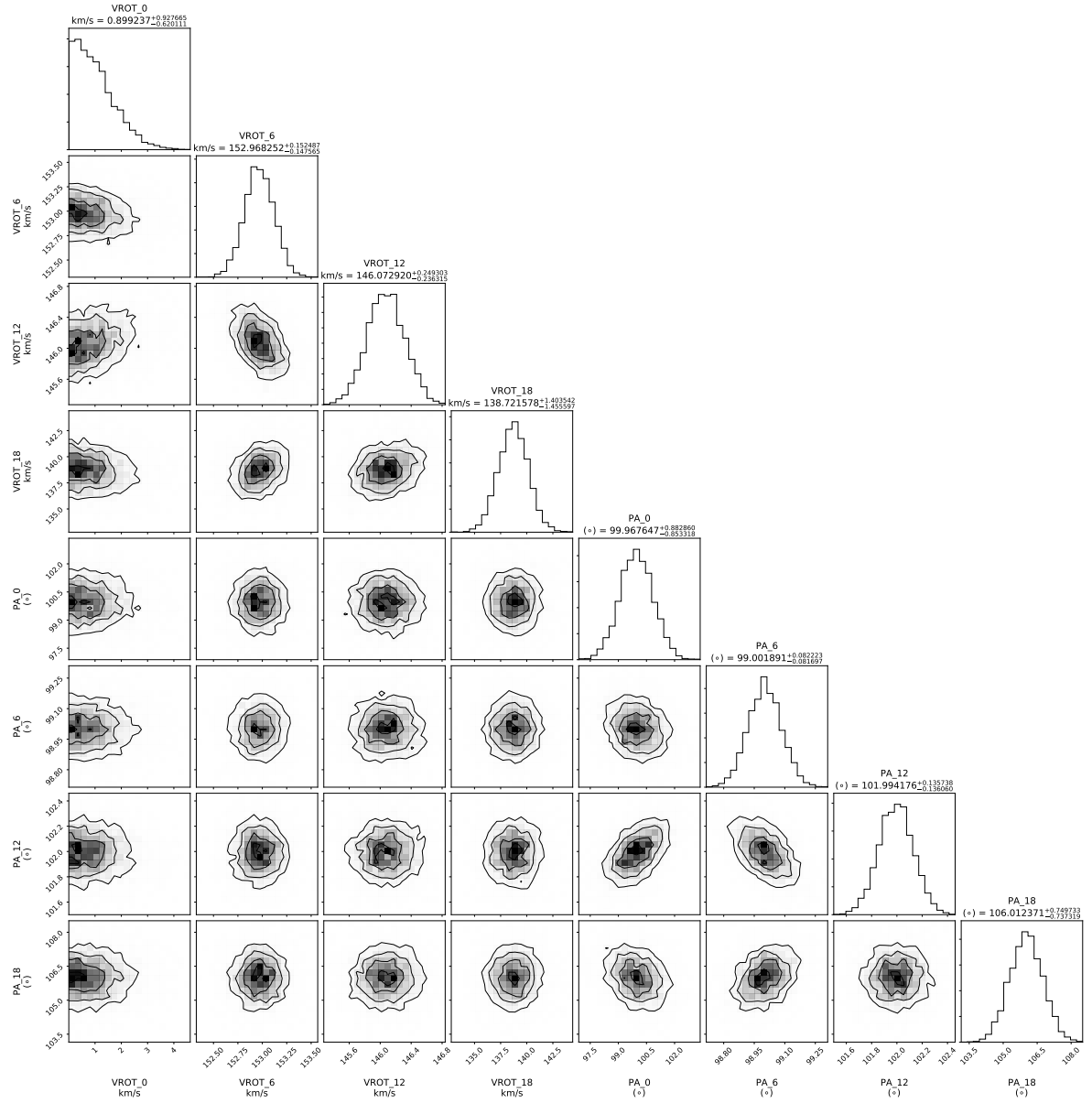


FIGURE A.3: Plots showing the correlation between rotation velocity and position angle.

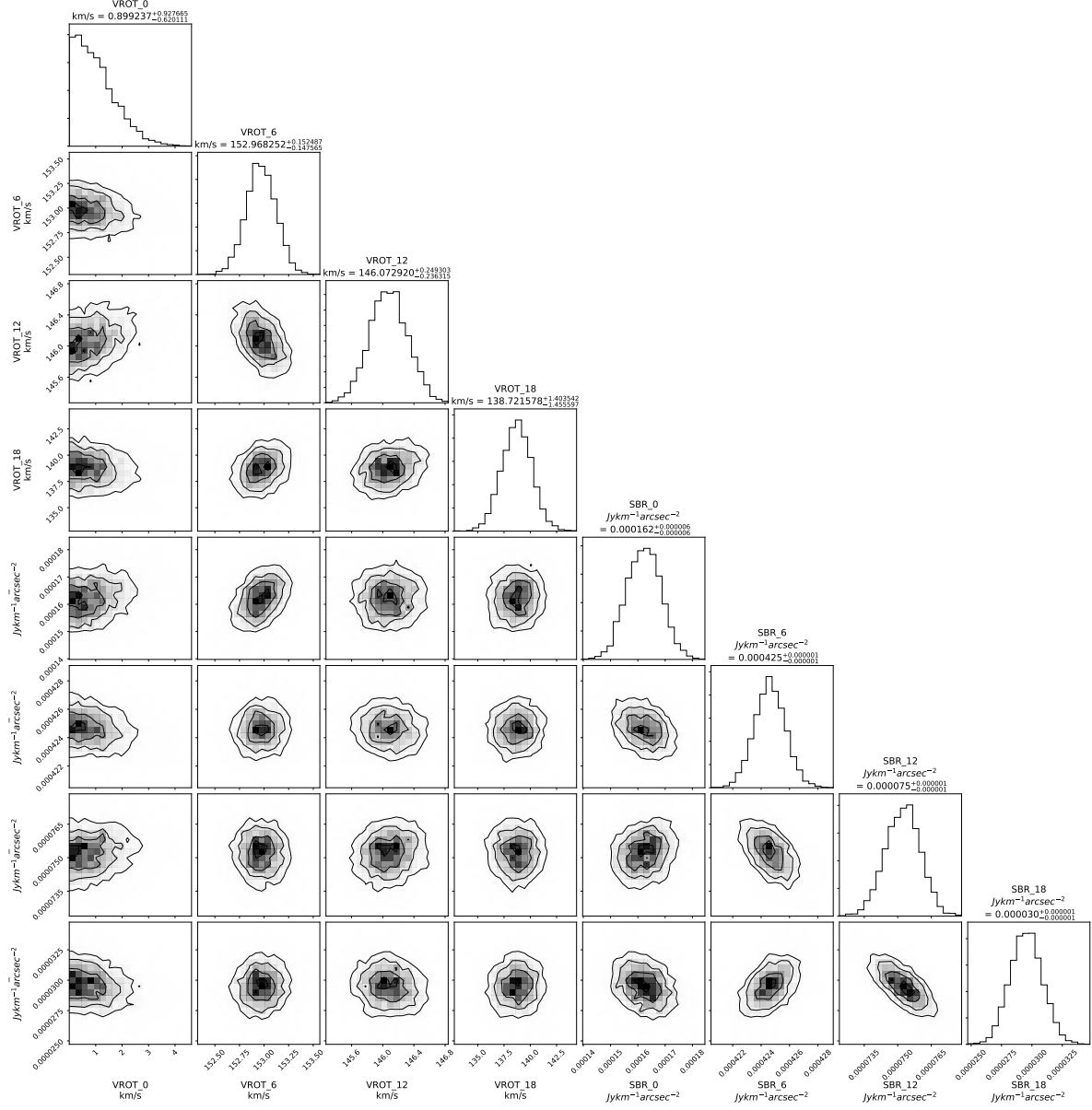


FIGURE A.4: Plots showing the correlation between rotation velocity and surface brightness.



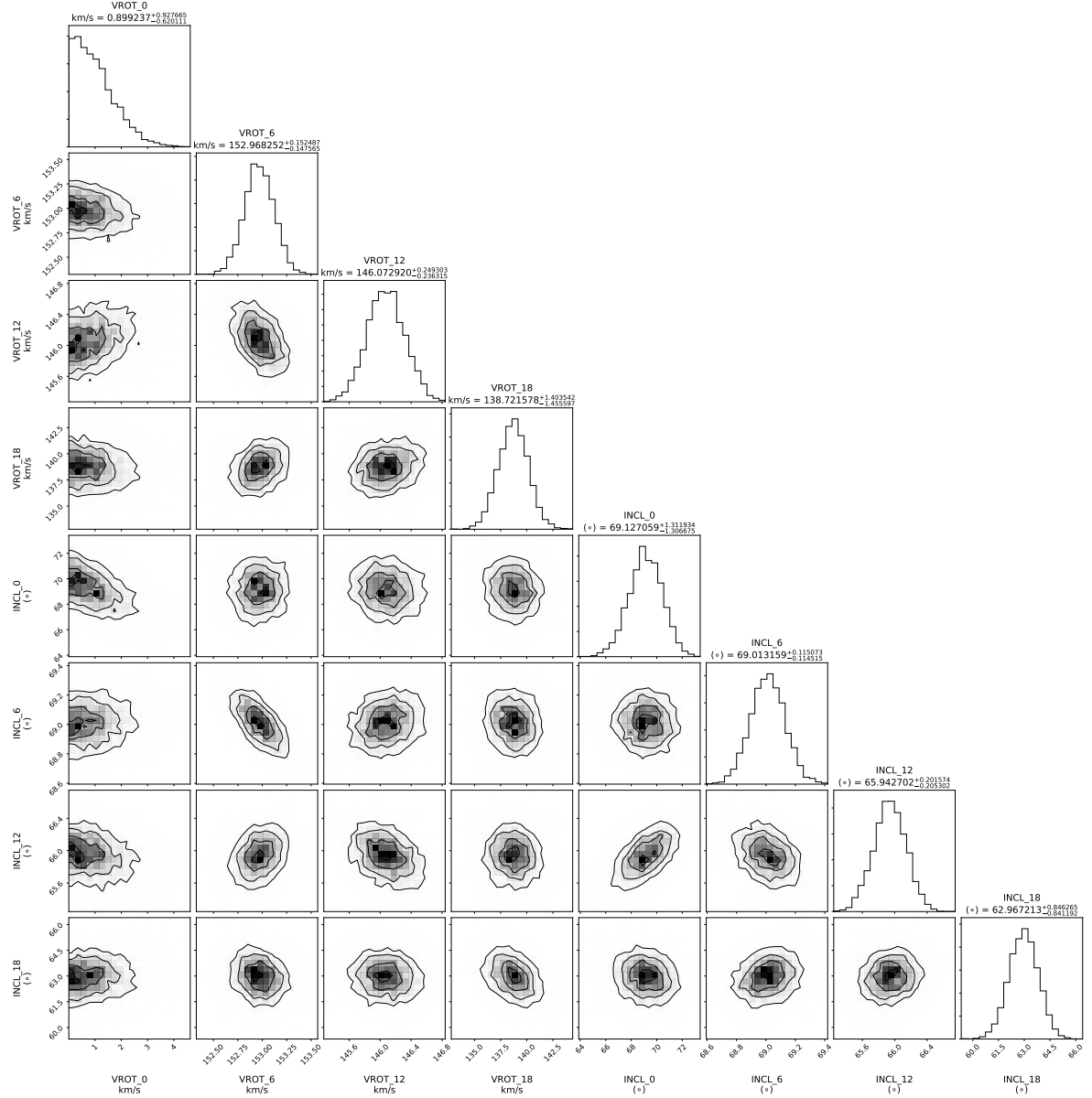


FIGURE A.5: Plots showing the correlation between rotation velocity and inclination angle.

## A.9 Minimizing Prior Values Experiment 6

Prior values used for this experiment are as follows.

$$\text{vrot max} = 397.584 \text{ km s}^{-1}$$

$$\text{vrot min} = 0.0 \text{ km s}^{-1}$$

$$\text{incl max} = 92.376836365$$

$$\text{inclmin} = 52.376836365$$

$$\text{PA max} = 120.6192$$

$$\text{PAmin} = 80.6192$$

$$\text{SBR max} = 0.0021 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$$

SBR min =  $0.0 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$   
 XPOS max =  $181.02671^\circ$   
 XPOS min =  $181.00671^\circ$   
 YPOS max =  $31.90526^\circ$   
 YPOS min =  $31.88526^\circ$   
 VSYS max =  $787.3097 \text{ km s}^{-1}$   
 VSYS min =  $747.3097 \text{ km s}^{-1}$

The following parameters were also set

Multinest Parameters:

n live points = 500  
 evidence tolerance = 0.5  
 sampling efficiency = 1000.0  
 const efficiency mode=False

TiRiFiC Parameters

CFLUX=  $2\text{e-}06$

Number of multinest processes run = 5  
 Number of likelihood evaluations = 5502746

## A.10 Minimizing Prior Values Experiment 7

Prior values used for this experiment are as follows.

vrot max =  $397.585 \text{ km s}^{-1}$   
 vrot min =  $0.0 \text{ km s}^{-1}$   
 incl max =  $92.377^\circ$   
 incl min =  $52.377^\circ$   
 PA max =  $120.619^\circ$   
 PA min =  $80.619^\circ$   
 SBR max =  $0.0021 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$   
 SBR min =  $0.0 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$   
 XPOS max =  $181.02671^\circ$   
 XPOS min =  $181.00671^\circ$   
 YPOS max =  $31.90526^\circ$   
 YPOS min =  $31.88526^\circ$   
 VSYS max =  $787.3097 \text{ km s}^{-1}$   
 VSYS min =  $747.3097 \text{ km s}^{-1}$

The following parameters were also set

Multinest Parameters:

n live points = 500

evidence tolerance = 1000

sampling efficiency = 1

const efficiency mode=True

TiRiFiC Parameters

CFLUX= 2e-06

For 500 live points

Number of multinest processes run = 5

Number of likelihood evaluations = 249544

For 5000 live points

Number of multinest processes run = 5

Number of likelihood evaluations = 719621

These 5000 live points had the following corner plot figures

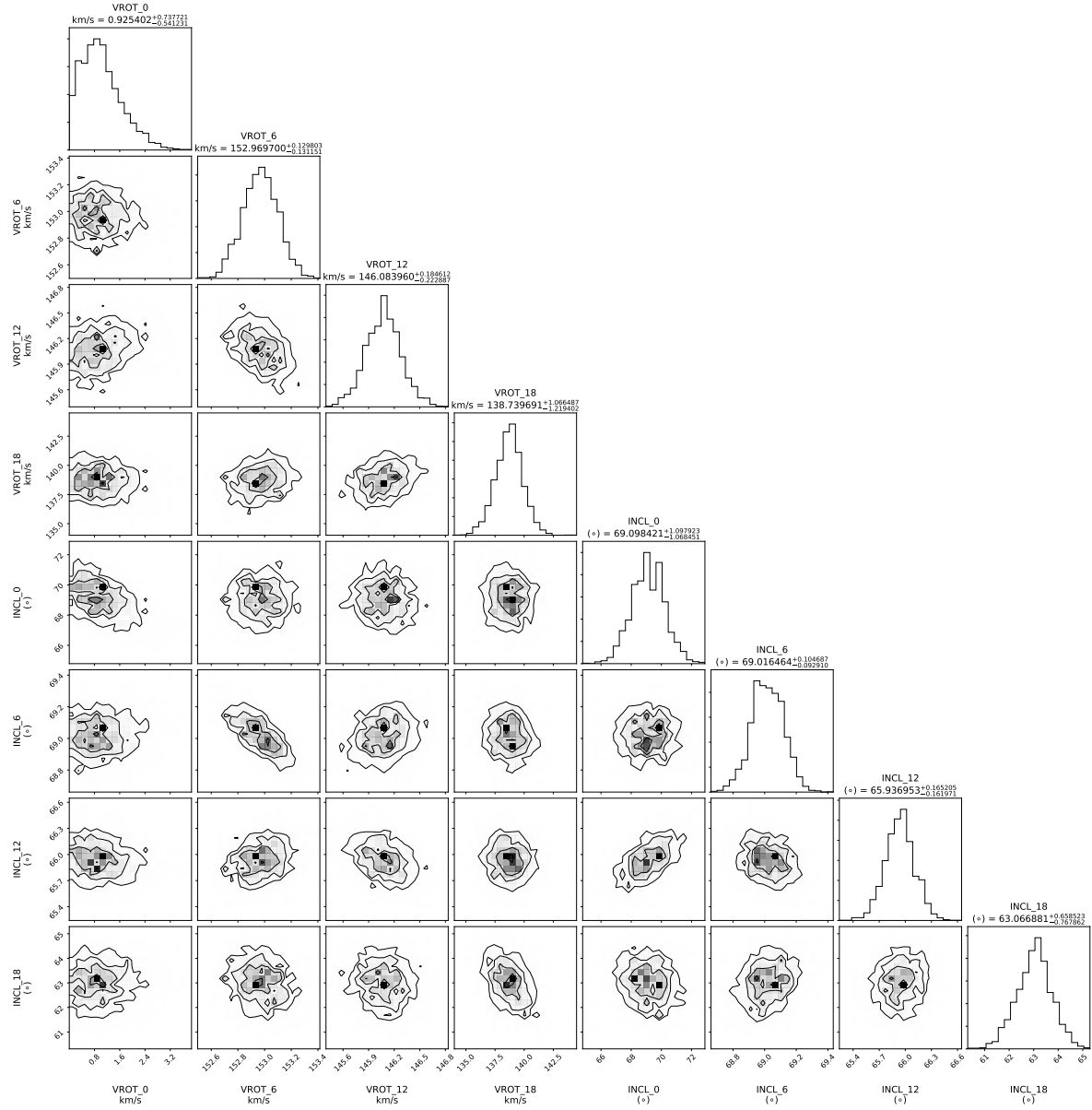


FIGURE A.6: Plots showing the correlation between Rotation velocity and Inclination angle.

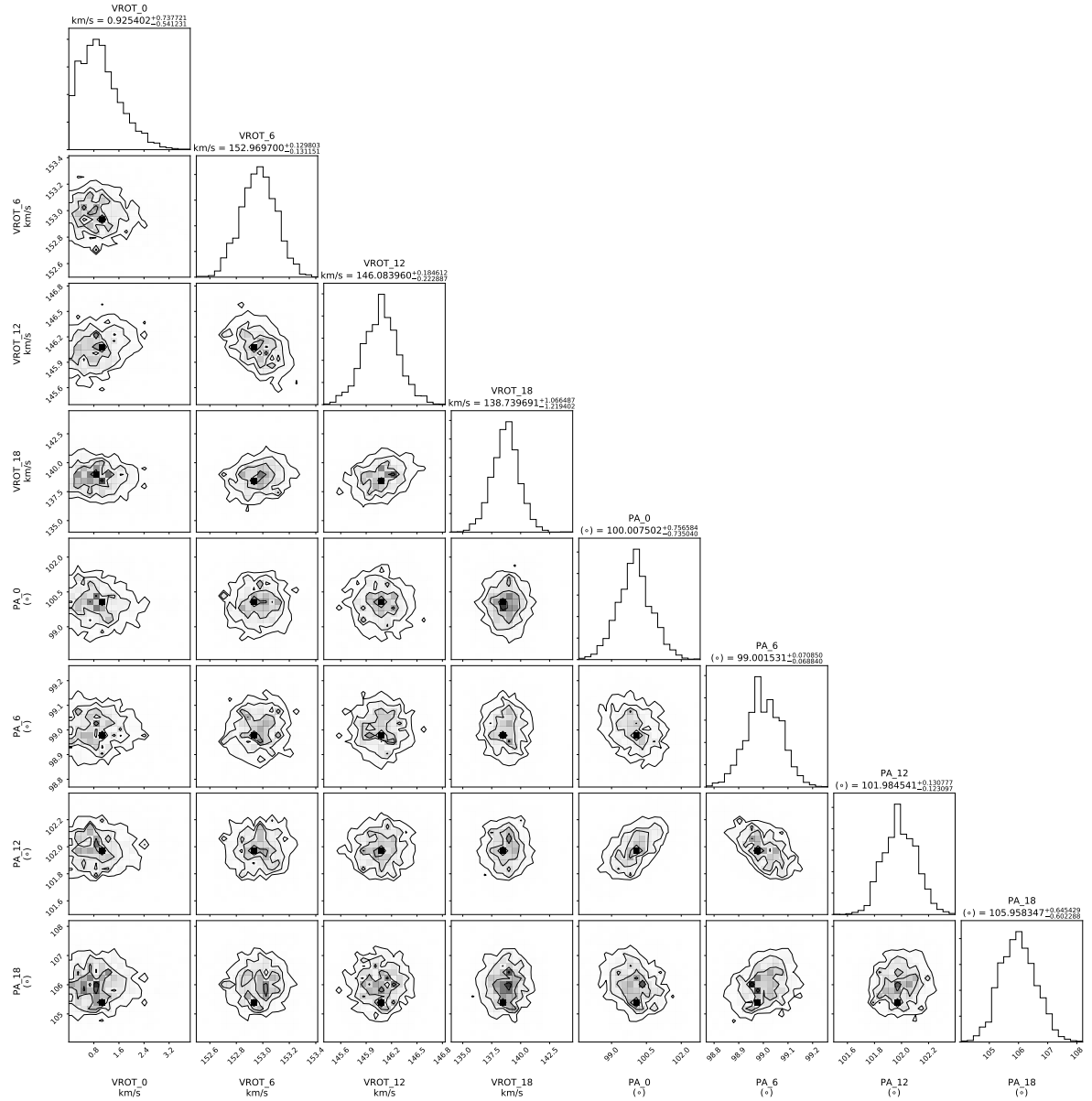


FIGURE A.7: Plots showing the correlation between rotation velocity and position angle.

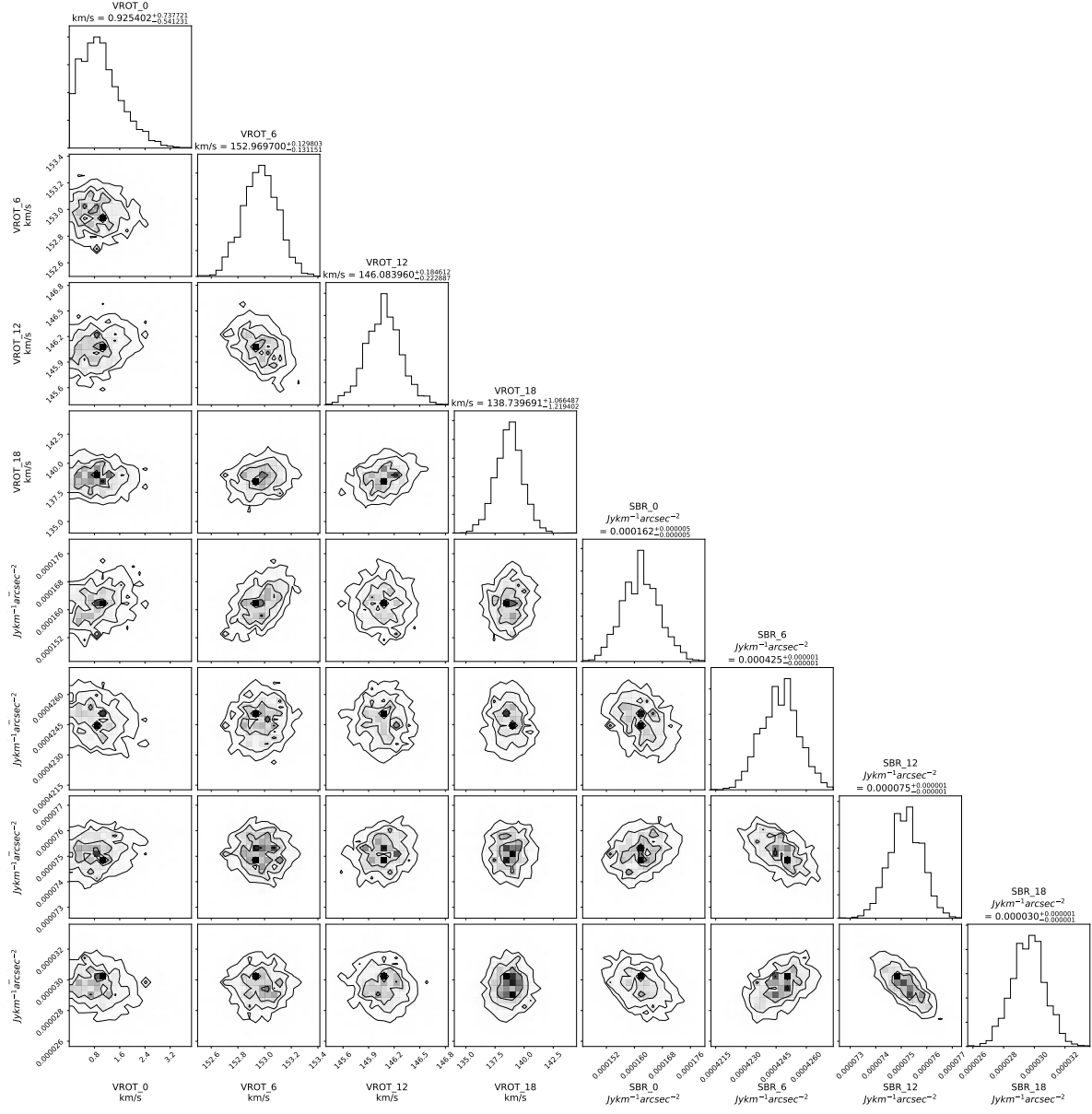


FIGURE A.8: Plots showing the correlation between rotation velocity and surface brightness.

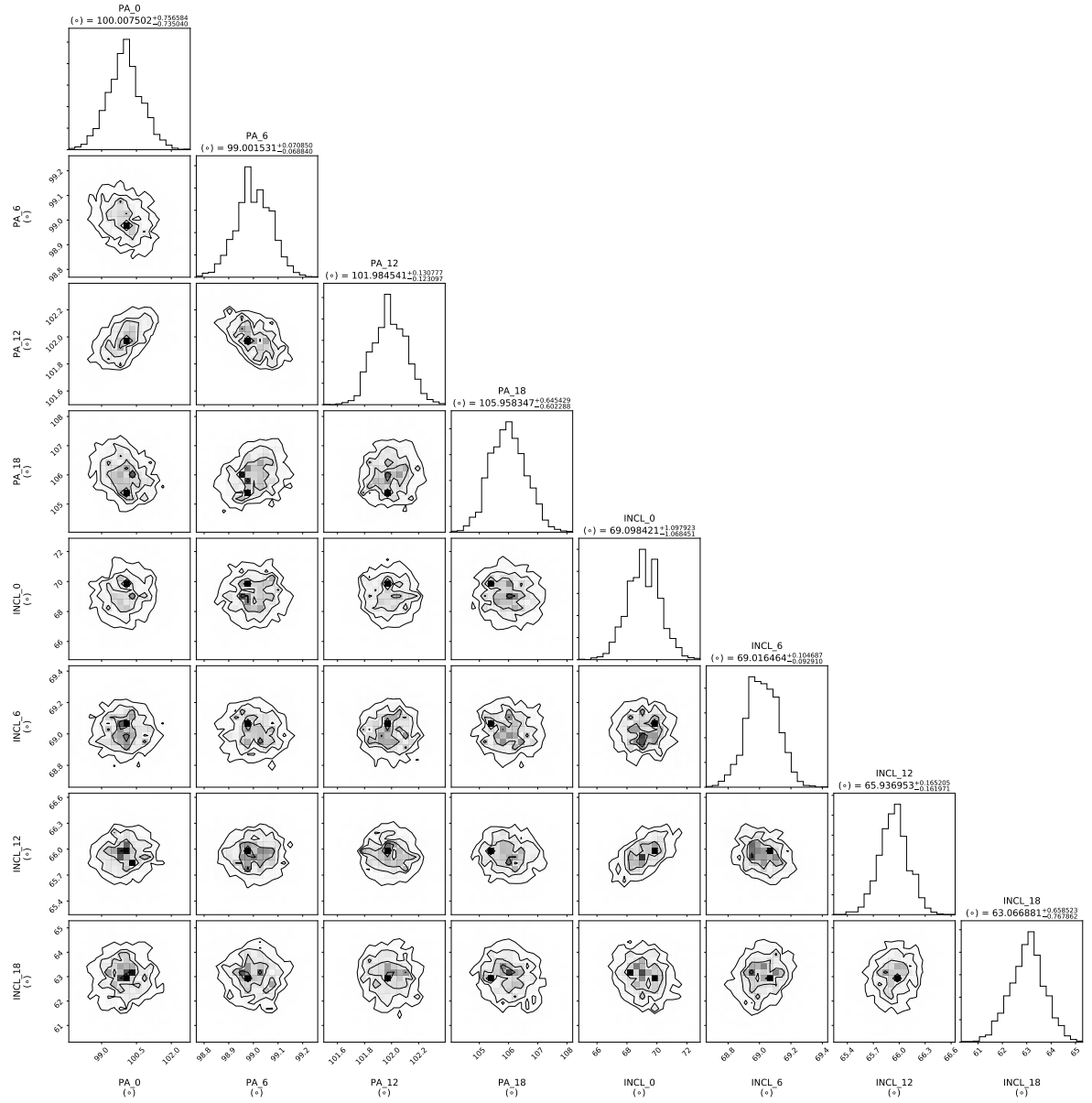


FIGURE A.9: Plots showing the correlation between position angle and inclination angle.

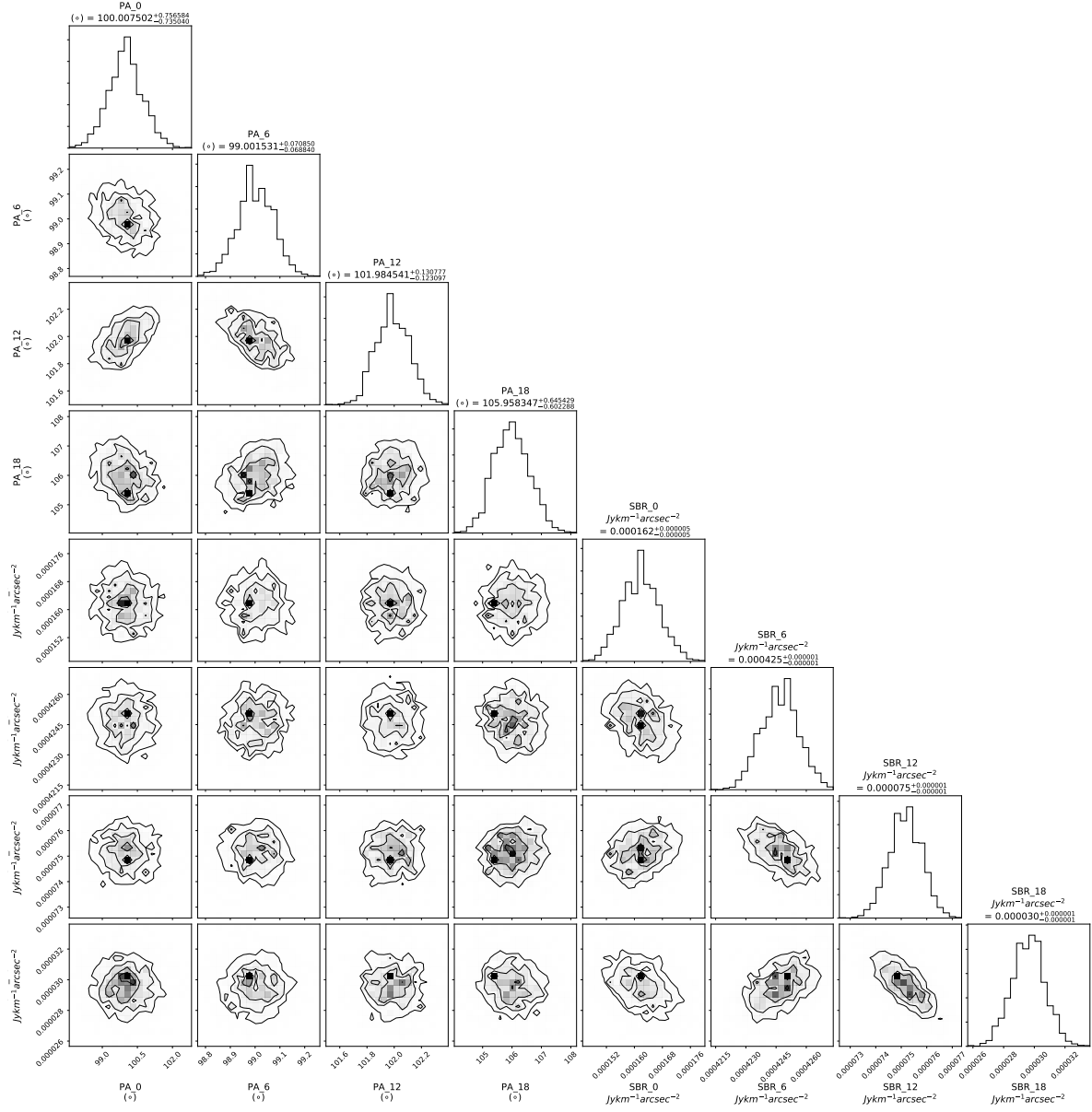


FIGURE A.10: Plots showing the correlation between position angle and Surface brightness.



## A.11 Minimizing Prior Values Experiment 8

Prior values used for this experiment are as follows.

$\text{vrot max} = 439.415 \text{ km s}^{-1}$   
 $\text{vrot min} = 0.0 \text{ km s}^{-1}$   
 $\text{incl max} = 95.6^\circ$   
 $\text{incl min} = 55.6^\circ$   
 $\text{PA max} = 140.1^\circ$   
 $\text{PA min} = 60.1^\circ$   
 $\text{SBR max} = 0.002 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$   
 $\text{SBR min} = 0.0 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$   
 $\text{XPOS max} = 181.026^\circ$   
 $\text{XPOS min} = 181.006^\circ$   
 $\text{YPOS max} = 31.90^\circ$   
 $\text{YPOS min} = 31.89^\circ$   
 $\text{VSYS max} = 784.7 \text{ km s}^{-1}$   
 $\text{VSYS min} = 744.7 \text{ km s}^{-1}$

The following parameters were also set

Multinest Parameters

$\text{n live points} = 500$   
 $\text{evidence tolerance} = 1000$   
 $\text{sampling efficiency} = 1$   
 $\text{const efficiency mode} = \text{True}$

$\text{log-likelihood dividing constant} = 100$

TiRiFiC Parameters

$\text{CFLUX} = 2\text{e-}06$

$\text{Number of multinest processes run} = 5$

$\text{Number of likelihood evaluations} = 108929$

## A.12 Minimizing Prior Values Experiment 9

Prior values used for this experiment are as follows.

$\text{vrot max} = 439.4 \text{ km s}^{-1}$   
 $\text{vrot min} = 0.0 \text{ km s}^{-1}$

incl max =  $95.6^\circ$   
 incl min =  $55.6^\circ$   
 PA max =  $140.1^\circ$   
 PA min =  $60.1^\circ$   
 SBR max =  $0.002 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$   
 SBR min =  $0.0 \text{ Jy km s}^{-1} \text{ arcsec}^{-2}$   
 XPOS max =  $181.026^\circ$   
 XPOS min =  $181.006^\circ$   
 YPOS max =  $31.905^\circ$   
 YPOS min =  $31.885^\circ$   
 VSYS max =  $784.7 \text{ km s}^{-1}$   
 VSYS min =  $744.7 \text{ km s}^{-1}$

The following parameters were also set

Multinest Parameters

n live points = 2000  
 evidence tolerance = 1000  
 sampling efficiency = 1  
 const efficiency mode=True

log-likelihood dividing constant = 500

TiRiFiC Parameters

CFLUX= 2e-06

Number of multinest processes run = 5

Number of likelihood evaluations = 385,713

## A.13 Minimizing Prior Values Experiment 10

Prior values used for this experiment are as follows.

vrot max =  $439.4 \text{ km s}^{-1}$   
 vrot min =  $0.0 \text{ km s}^{-1}$   
 incl max =  $95.6^\circ$   
 incl min =  $55.6^\circ$   
 PA max =  $140.1^\circ$   
 PA min =  $60.1^\circ$   
 SBR max =  $0.002 \text{ Jy km s}^{-1}$   
 SBR min =  $0.0 \text{ Jy km s}^{-1}$

XPOS max = 181.026 °

XPOS min = 181.006 °

YPOS max = 31.905 °

YPOS min = 31.885 °

VSYS max = 784.7 km s<sup>-1</sup>

VSYS min = 744.7 km s<sup>-1</sup>

The following parameters were also set

Multinest Parameters

n live points = 5000

evidence tolerance = 1000

sampling efficiency = 1

const efficiency mode=True

log-likelihood dividing constant = 2000

TiRiFiC Parameters

CFLUX= 2e-06

Number of multinest processes run = 5

Number of likelihood evaluations = 890,314

These 5000 live points had the following corner plot figures

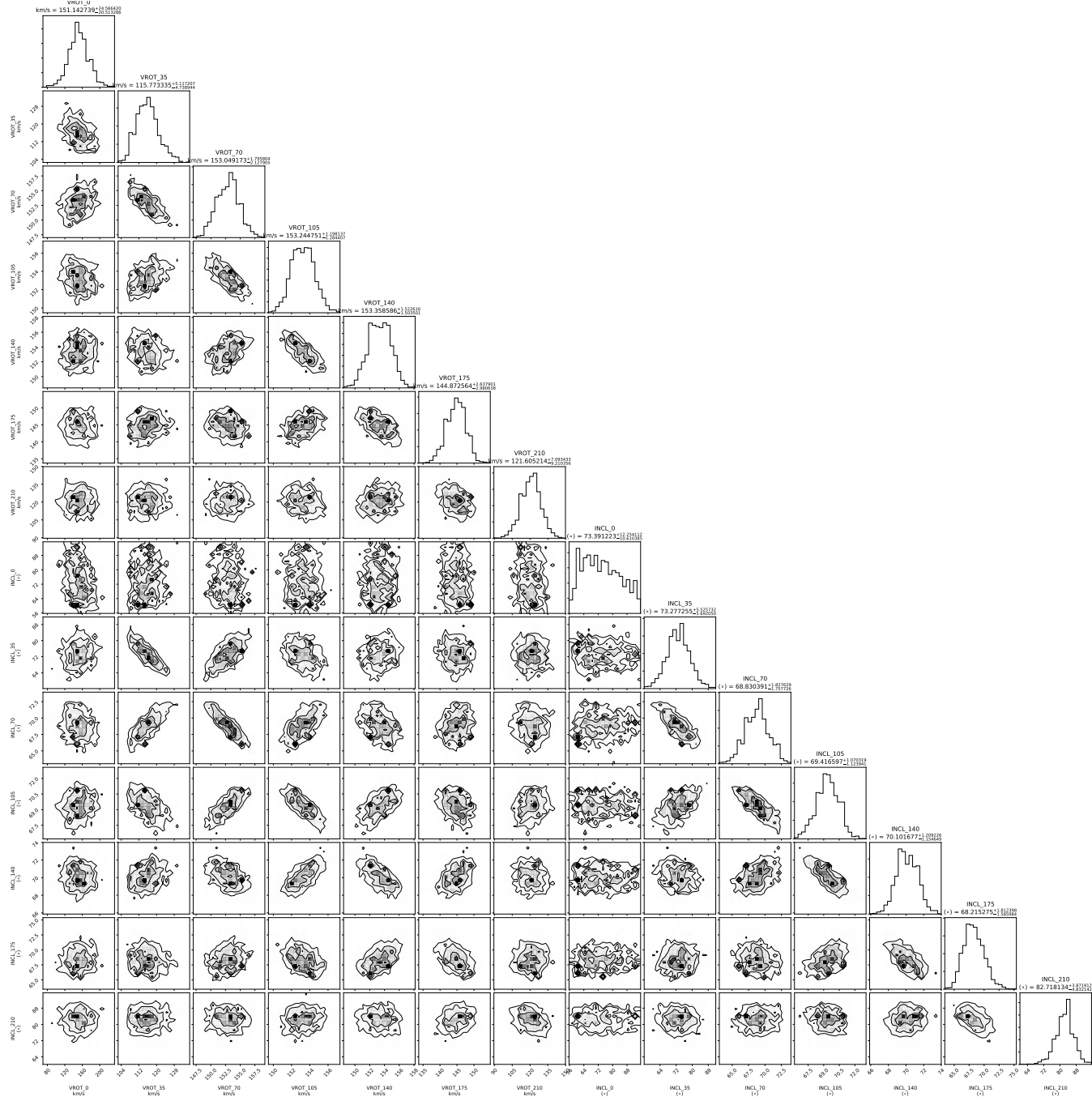


FIGURE A.11: Plots showing the correlation between rotation velocity and inclination angle. The inclination angle at radius 0 is spread over the entire prior range and hence unreliable.

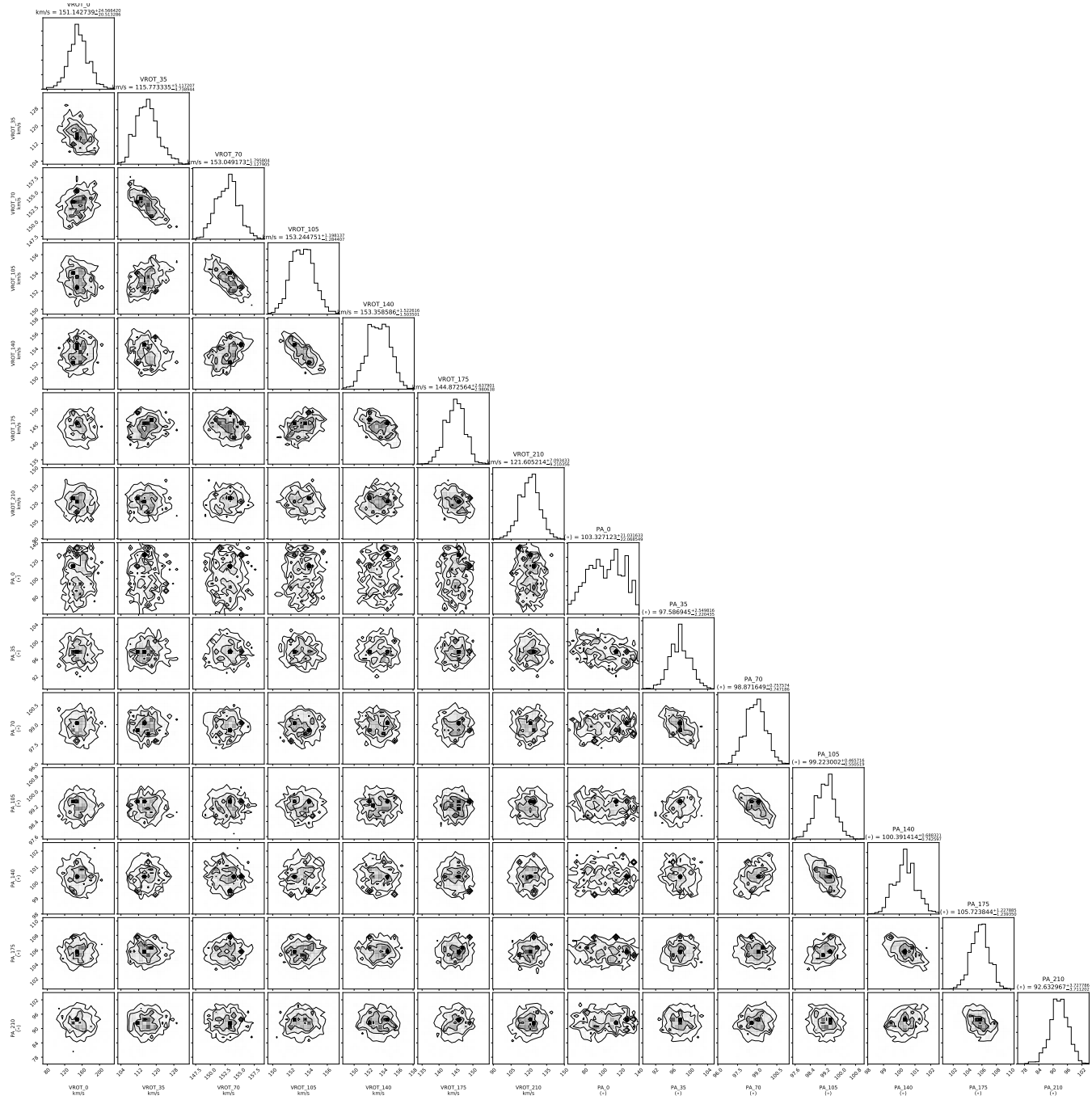


FIGURE A.12: Plots showing the correlation between rotation velocity and position angle. The position angle at radius 0 is spread over the entire prior range and hence unreliable.

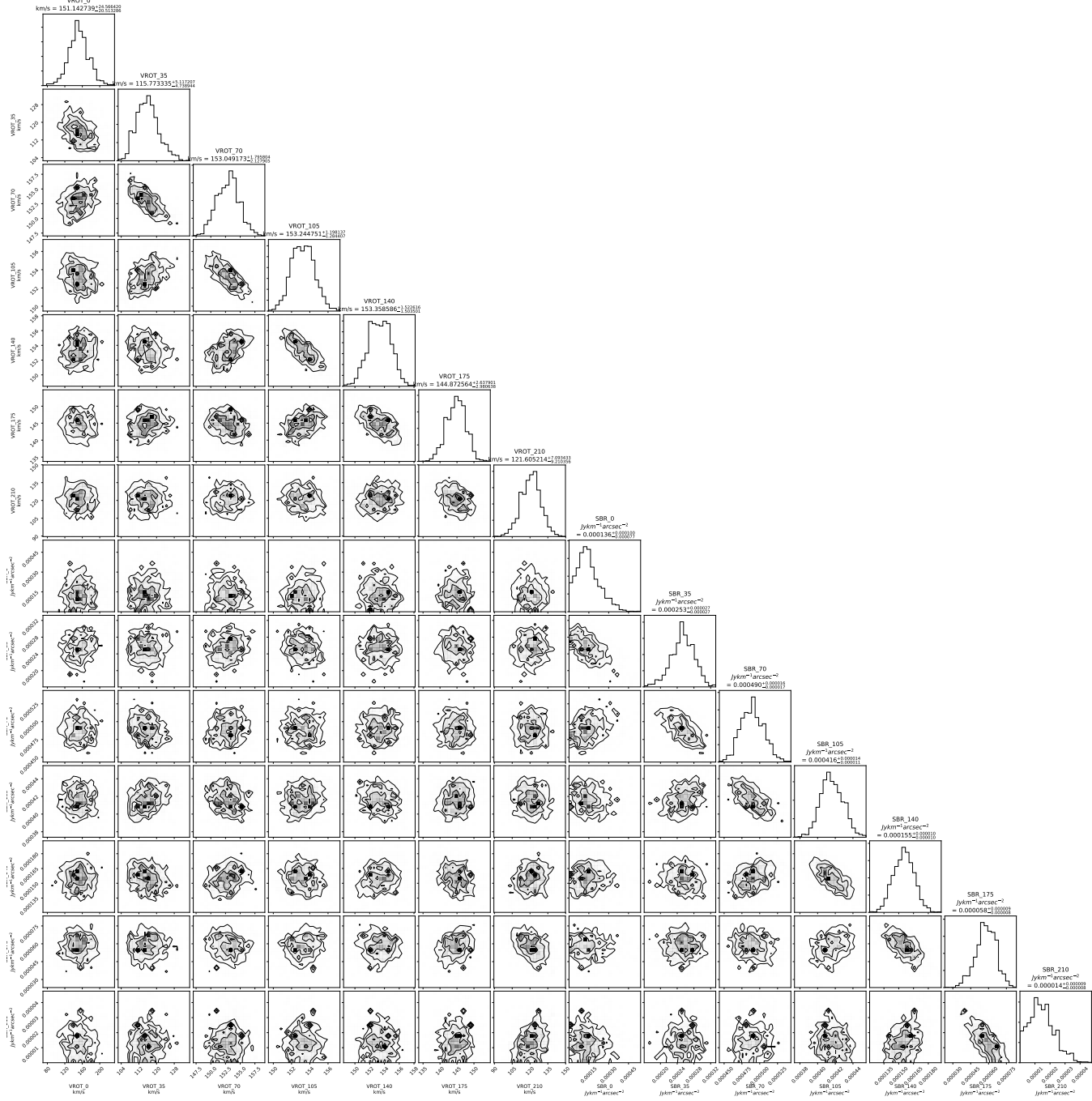


FIGURE A.13: Plots showing the correlation between rotation velocity and surface brightness.

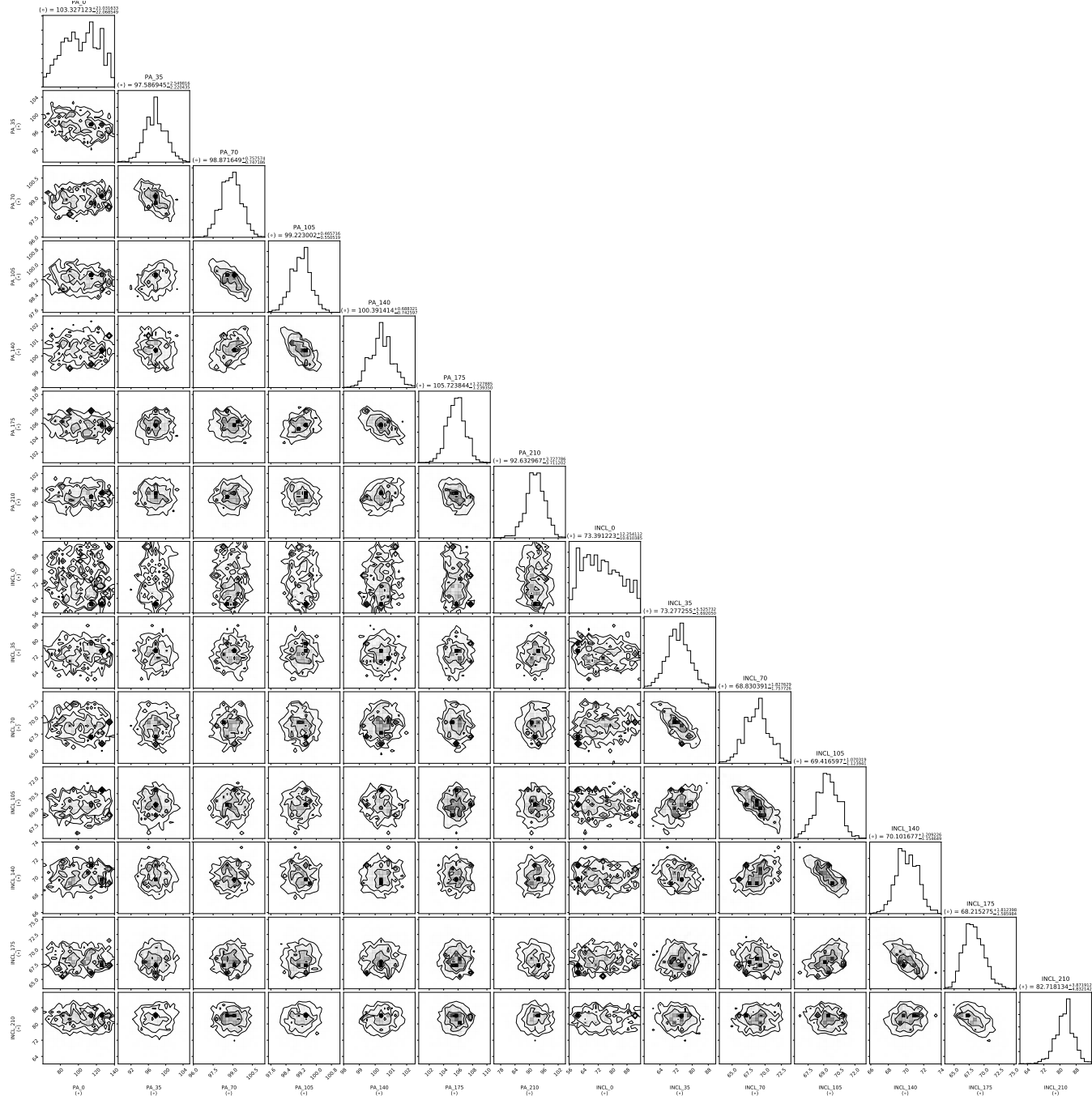


FIGURE A.14: Plots showing the correlation between position angle and inclination angle. The position angle and inclination angle at radius 0 is spread over the entire prior range and hence unreliable.

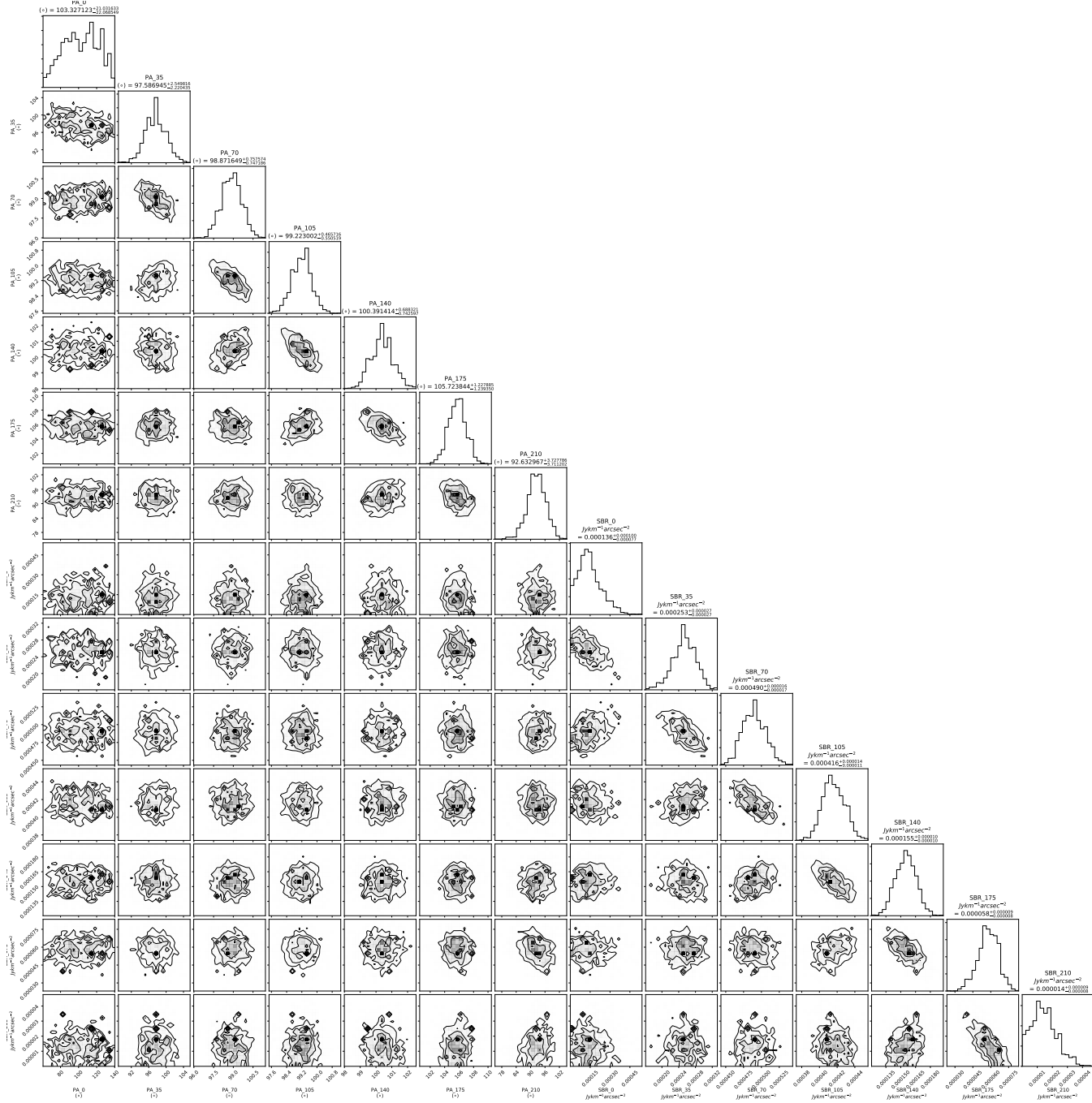


FIGURE A.15: Plots showing the correlation between position angle and surface brightness. The position angle at radius 0 is spread over the entire prior range and hence unreliable.



## A.14 Minimizing Prior Values Test with real galaxy N3198

Prior values used for this experiment are as follows.

```

var par = 4.0
vrot max = 198.7 km s-1
vrot min = 0.0 km s-1
incl max = 95.8 °
incl min = 55.8 °
PA max = 257.5 °
PA min = 177.5 °
SBR max = 0.00009 Jy km s-1 arcsec-2
SBR min = 0.0 Jy km s-1 arcsec-2
XPOS max = 154.986 °
XPOS min = 154.966 °
YPOS max = 45.562 °
YPOS min = 45.542 °
VSYS max = 783.8 km s-1
VSYS min = 583.8 km s-1

```

The following parameters were also set

Multinest Parameters

```

n live points = 5000
evidence tolerance = 1000
sampling efficiency = 1
const efficiency mode=True

```

log-likelihood dividing constant = 2000

TiRiFiC Parameters

```

CFLUX= 1e-04
SBR max = 0.0006
SBR min = 0.0

```

Number of multinest processes run = 5

Number of likelihood evaluations = 463,799

## A.15 Minimizing Prior Values Test with real galaxy N4062

Prior values used for this experiment are as follows.

```

var par = 7.0
vrot max = 439.4 km s-1
vrot min = 0.0 km s-1
incl max = 95.6 °
incl min = 55.6 °
PA max = 140.1 °
PA min = 60.1 °
SBR max = 0.002 Jy km s-1 arcsec-2
SBR min = 0.0 Jy km s-1 arcsec-2
XPOS max = 181.026 °
XPOS min = 181.006 °
YPOS max = 31.905 °
YPOS min = 31.885 °
VSYS max = 784.7 km s-1
VSYS min = 744.7 km s-1

```

The following parameters were also set

Multinest Parameters

```

n live points = 5000
evidence tolerance = 1000
sampling efficiency = 1
const efficiency mode=True

```

log-likelihood dividing constant = 2000

TiRiFiC Parameters

CFLUX= 2e-06

Number of multinest processes run = 5

Number of likelihood evaluations = 961,324

# Bibliography

- Alcock, C., Allsman, R. A., Axelrod, T. S., et al. 1995a, *The Astrophysical Journal*, 461, 84
- Alcock, C., Allsman, R. A., Axelrod, T. S., et al. 1995b, *Physical Review Letters*, 74, 2867
- Beaulieu, J. P., Lamers, H. J. G. L. M., & de Wit, W. J. 1998
- Begeman, K. 1989, PhD thesis, , Kapteyn Institute, (1987)
- Briggs, F. H. 1990, *The Astrophysical Journal*, 352, 15
- Buchner, J. & Johannes. 2016, *Astrophysics Source Code Library*, record ascl:1606.005
- Chrysomallis, M. & Christodoulou, C. 2015, *Wiley Encyclopedia of Electrical and Electronics Engineering*, 1
- Di Teodoro, E. & Fraternali, F. 2015, *Monthly Notices of the Royal Astronomical Society*, 451, 3021
- Duffy, A. R., Meyer, M. J., Staveley-Smith, L., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 426, 3385
- Feroz, F., Hobson, M. P., & Bridges, M. 2008, *Monthly Notices of the Royal Astronomical Society*, 398, 1601
- Freeman, K. C. 1970, *OApJ.. 160. .811F The Astrophysical Journal*, 160, 811
- Frigo, M. & Johnson, S. 2005, *Proceedings of the IEEE*, 93, 216
- Frusciante, N., Salucci, P., Vernieri, D., Cannon, J. M., & Elson, E. C. 2012, *Proceedings of the International Astronomical Union*, 5, 751
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *Physics Reports*, 433, 181
- Gentile, G., Józsa, G. I. G., Serra, P., et al. 2013, *Astronomy and Astrophysics*, 125, 1
- Harrison, C. M. 2016 (Springer, Cham), 37–46
- Heald, G., Józsa, G., Serra, P., et al. 2011, *A&A*, 526
- Hubble, E. 1929, *Proceedings of the National Academy of Sciences of the United States of America*, 15, 168

- Hubble, E. P. 1925, *The Observatory*, Vol. 48, p. 139-142 (1925), 48, 139
- Irwin, J. a. 2007, *Astrophysics: Decoding the Cosmos*, 1–456
- J. J. Condon & S. M. Ransom. 2016, *Essential Radio Astronomy Course*
- Jansky, K. 1933, *Proceedings of the IRE*, 21, 1387
- Józsa, G., Niemczyk, C., Klein, U., & Oosterloo, T. 2007a, *New Astronomy Reviews*, 51, 116
- Józsa, G. I. G., Kenn, F., Klein, U., & Oosterloo, T. A. 2007b, *Astronomy & Astrophysics*, 468, 731
- KERR, F. J., HINDMAN, J. V., & CARPENTER, M. S. 1957, *Nature*, 180, 677
- Lockman, F. J. 2017, *Radio Astronomy: Observing the Invisible Universe (THE GREAT COURSES)*, 273
- Longair, M. S. 2011, *High energy astrophysics*, 58–65
- Milgrom, M. 1983, A MODIFICATION OF THE NEWTONIAN DYNAMICS AS A POSSIBLE ALTERNATIVE TO THE HIDDEN MASS HYPOTHESIS 1, Tech. rep.
- Miller, D. F. 1998, *Jpl D13835*, 108, 20713
- Oh, S.-H. H., Staveley-Smith, L., Spekkens, K., Kamphuis, P., & Koribalski, B. S. 2018, *Monthly Notices of the Royal Astronomical Society*, 473, 3256
- Olling, R. P. 1996, *The Astronomical Journal*, 112, 457
- Peters, S. P. C., van der Kruit, P. C., Allen, R. J., & Freeman, K. C. 2017, *Monthly Notices of the Royal Astronomical Society*, 464, 2
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes in C ++ The Art of Scientific Computing Second Edition*, Tech. rep.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007 (Cambridge University Press), 23–42
- Punzo, D., van der Hulst, J. M., Roerdink, J. B. T. M., Fillion-Robin, J. C., & Yu, L. 2017
- R. Sancisi. 1976, *Astronomy and Astrophysics*, Vol. 53, 159
- Rogstad, D. H., Lockhart, I. A., Lockart, I. A., & Wright, M. C. H. 1974, *The Astrophysical Journal*, 193, 309
- Schmidt, T. M., Bigiel, F., Klessen, R. S., & de Blok, W. J. G. 2016, *Monthly Notices of the Royal Astronomical Society*, 457, 2642

- Schneider, P. 2006 (Springer-Verlag Berlin Heidelberg), 510–510
- Sellwood, J. A. & Sanchez, R. Z. 2009, *Monthly Notices of the Royal Astronomical Society*, 404, 1733
- Serra, P., Westmeier, T., Giese, N., et al. 2015, *Monthly Notices of the Royal Astronomical Society*, 448, 1922
- Shaw J R, Bridges, M., & Hobson, M. P. 2007, 7, 1
- Skillman, E. D. 1996
- Spekkens, K. & Giovanelli, R. 2006, *The Astronomical Journal*, 132, 1426
- Swaters, R. A. 1999, PhD thesis, University of Groningen
- Thompson, A. R., Moran, J. M., & Swenson, G. W. 1986, *Interferometry and synthesis in radio astronomy* (Wiley New York et al.)
- Tully, R. 2007, *Scholarpedia*, 2, 4485
- Udalski, D., Szymanski, M., Kaluzny, J., et al. 1994, *The Astrophysical Journal*, 426, L69
- van der Hulst, J. M., Terlouw, J. P., Begeman, K. G., Zwitter, W., & Roelfsema, P. R. 1992, *Astronomical Data Analysis Software and Systems I*, A.S.P. Conference Series, Vol. 25, 1992, Diana M. Worrall, Chris Biemesderfer, and Jeannette Barnes, eds., p. 131.
- van der Marel, R. P., Besla, G., Cox, T. J., Sohn, S. T., & Anderson, J. 2012, *Astrophysical Journal*, 753, 9
- Vaz, A. I. F. & Vicente, L. N. 2007, *Journal of Global Optimization*, 39, 197
- Verschuur, G. L., Kellermann, K. I., & van Brunt, V. 1974, *Galactic and Extra-Galactic Radio Astronomy*, ed. G. L. Verschuur & K. I. Kellermann (Berlin, Heidelberg: Springer Berlin Heidelberg), 127
- Walter, F., Brinks, E., De Blok, W. J. G., et al. 2008a, *The Astronomical Journal*, 136, 2563
- Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008b, *Astronomical Journal*, 136, 2563
- Wambsganss, J. & Wambsganss, J. 1998, *Gravitational Lensing in Astronomy*, 1, 12
- Yun, M. 1999, *Symposium - International Astronomical Union*, 186, 81