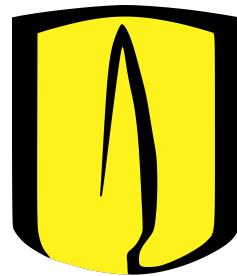


# Observational evidence of star formation stochasticity in the CALIFA dataset

Nicolás Romero Díaz

201127499

Advisor: Jaime Forero-Romero



Para mi familia, quien me ha apoyado siempre en toda circunstancia.  
Para mis amigos, siempre indispensables.

Jaime E. Forero-Romero  
Advisor

---

Signature

---

Date

Nicolás Romero Díaz  
Student

---

Signature

---

Date

## Abstract

It is generally acknowledged that the total luminosity of a galaxy in the UV region of the electromagnetic spectrum and other emission line intensities, e.g.  $H_\alpha$ , are proportional to that galaxy's star formation rate (SFR). However, computational simulations suggest that in regions with low SFR the influence of stochastic phenomena can break this proportionality. When considering these effects, theoretical findings indicate that UV luminosities and emission line intensities are scattered around a mean value as opposed to a deterministic, unique expected value. In this work we analyze data from the public data-releases of the Calar Alto Legacy Integral Area (CALIFA) Survey in order to determine whether these effects can be observed. We find some marginal evidence of the presence of these stochastic phenomena for the  $H_\alpha/H_\beta$  intensity ratio and the equivalent width (EW) of OII. We parametrize the distributions found in the data and discuss the course of action for future works. These include correction for extinction effects and expanding our data sample to other surveys.

## Resumen

La luminosidad total de una galaxia en la región UV y de otras líneas de emisión, e.g  $H_\alpha$ , es generalmente aceptada como una medida proporcional de la tasa de formación estelar (SFR) de aquella galaxia. Sin embargo, se han encontrado indicios en simulaciones de que en regiones con baja SFR la influencia de procesos estocásticos puede romper esta proporcionalidad. Al considerar estos efectos, resultados teóricos indican que tanto las luminosidades en UV como las intensidades de otras líneas de emisión son dispersadas alrededor de un valor medio, dejando de ser ya un valor determinístico y único. En este trabajo analizamos datos de las publicaciones abiertas de la encuesta Calar Alto Legacy Integral Field Area (CALIFA) Survey para poder determinar si estos efectos son observados. Encontramos evidencia marginal de la presencia de estos fenómenos para la razón de intensidades  $H_\alpha/H_\beta$  y en el ancho equivalente (EW) del OII. Parametrizamos las distribuciones encontradas en los datos y discutimos un curso de acción para trabajos futuros. Entre ellos está incluir correcciones de extinción galáctica y aumentar nuestra muestra de datos a otras encuestas.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Theoretical framework</b>	<b>7</b>
2.1	The IMF . . . . .	7
2.2	Stochasticity in star formation . . . . .	8
2.3	Measuring stochasticity . . . . .	11
2.3.1	The Balmer decrement . . . . .	11
<b>3</b>	<b>The CALIFA survey</b>	<b>14</b>
3.1	The CALIFA datacubes . . . . .	15
3.2	The Pipe3D outputs . . . . .	15
3.3	Analysis of the stellar population . . . . .	16
3.4	Analysis of strong emission lines . . . . .	17
<b>4</b>	<b>Results</b>	<b>18</b>
<b>5</b>	<b>Discussion</b>	<b>29</b>
5.1	The $H_\alpha/H_\beta$ ratio . . . . .	29
5.2	The OII EW . . . . .	30
<b>6</b>	<b>Conclusions</b>	<b>32</b>
6.1	Further work . . . . .	33

# Chapter 1

## Introduction

Star formation processes turn galactic gas clouds into stars. These processes are probabilistic by nature and can be parametrized by two components: the initial mass function (IMF) and the cluster mass function (CMF).

The IMF outlines the relative abundance of stars of a certain mass to be formed, and is represented by  $\phi(m)$ . For the most part, the relevant characteristics of the IMF are described by two pieces of information. (*i*) The mass interval, describing the available range of masses possible for the stars being generated ( $m_{min} - m_{max}$ ); (*ii*) the relative abundance of stars with different stellar masses, taken as  $\phi(m) \propto m^{-\gamma}$ . When  $\gamma > 0$  less massive stars are more likely to be formed[1]. The second parameter arises from the fact that star formation is mainly a cluster phenomenon [2] characterized by the cluster mass function,  $\psi(M_{ecl})$ . This CMF describes the number of clusters in a region where stars are being produced, with  $M_{ecl}$  the cluster mass [1].

The sampling of both the IMF and CMF is influenced by the star formation rate (SFR) of a particular system. For a region of high SFR the sampling of the IMF is relatively simple, as opposed to regions of low SFR where the sampling of the IMF and CMF becomes a different statistical problem. Authors in [3] have proposed that in these areas, stochastic processes have a greater influence over the statistical nature of both the IMF and CMF. This statistical foundation of the mass functions has several distinct effects that influence different spectral properties of gas clouds surrounding star forming localities. In particular, stochasticity appears to cause the ratio of H <sub>$\alpha$</sub>  to Far Ultra Violet (FUV) luminosity to fluctuate around a mean value [3].

Using the public SLUG code (Stochastically Light Up Galaxies), Forero-Romero and Dijkstra [1] found that this fluctuation around a mean value

also takes place when comparing the ratio of Lyman alpha ( $Ly\alpha$ ) flux to the FUV photons. This relation is a well known spectral quantity called the equivalent width (EW), which characterizes the  $Ly\alpha$  emission line strength. In another work, Mas-Ribas et.al [4] also observed fluctuation of the  $Ly\alpha$  luminosity when performing stochastic IMF sampling.

In chapter two we discuss the influence of the IMF, the appearance of stochasticity and its effects on spectral properties and a review of computational predictions as well as how spectral data can be affected by phenomena in the interstellar medium (i.e. galactic extinction).

In order to search for this stochastic effects, we will use data published by the Calar Alto Legacy Integral Field Area Survey (CALIFA), reported in detail by Sánchez et. al. [5][6]. A brief description of the CALIFA data is presented in chapter three, as well as an overall description of the survey. This chapter will also including a general outline of the Pipe3D analysis pipeline developed by the CALIFA team of Sánchez et. al. [7]

Representative results are included in chapter four.

In chapter five we present a discussion of the results including quantification of stochasticity. Finally, we present our conclusions in chapter six.

## Chapter 2

# Theoretical framework

### 2.1 The IMF

Results obtained in the field of stellar astrophysics are fundamentally affected by the selection of the IMF. The IMF is an empirical formula that illustrates the probability with which a star of a certain mass is found. Since stars of different masses ionize surrounding gas clouds in varying degrees, we must observe the choice of IMF when making an interpretation of our results. We present a breakdown of the two most widely used IMFs.

The IMF of any stellar cluster is believed to be relatively invariant from one group of stars to the other [8]. This stellar parameter quantifies  $N(m)dm$ , the number of stars with masses between  $m$  and  $m + dm$ , represented commonly in literature as  $\xi(m)\Delta m$ . Even though it is an empirical function, there are two models that are relevant to this work, the Salpeter IMF[9] and the Kroupa IMF [10].

The Salpeter IMF is:

$$\xi(m\Delta m) = \xi_0 \left( \frac{m}{M_\odot} \right)^\alpha \left( \frac{\Delta m}{M_\odot} \right) \quad (2.1)$$

And the Kroupa IMF:

$$\xi(m) = m^{-\alpha} \quad (2.2)$$

The difference between both arising from the distribution of the more massive stars, since Salpeter takes the  $\alpha$  parameter as independent of the stellar mass  $m$  with a constant value of  $\alpha = -2.35$ . On the other hand, Kroupa adjusts this parameter according to different ranges of stellar masses, maintaining  $\alpha = 2.35$  for stellar masses  $m > 0.5M_\odot$ ;  $\alpha = 1.3$  for mass ranges of  $0.08M_\odot < m < 0.5M_\odot$  and  $\alpha = 0.3$  for  $m < 0.08M_\odot$ . This causes

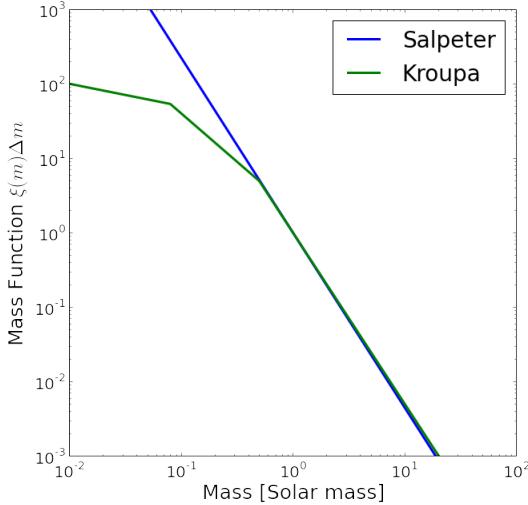


Figure 2.1: Visualization of the Salpeter IMF and Kroupa IMF

the number of massive stars in the Kroupa IMF to be lower than the one predicted by the Salpeter IMF.

The differences of these IMFs can be seen in figure 2.1. We can observe the segmented Kroupa IMF curve corresponding to threshold values of the stellar mass. In the article that inspired this work [1], a Salpeter IMF is selected, while a Kroupa IMF used in [3] with which whom we will be going back and forth regularly in this work. Despite the different characterizations of stellar populations, results in [3] are not significantly influenced by the selection of the IMF.

## 2.2 Stochasticity in star formation

In a region where stars are being produced, these stars will ionize gas clouds adjacent to them. The emission spectra from these clouds can tell us a great deal about the stellar population [3]. One quantity that is measured here is the equivalent width (EW). It is defined, for the purposes of this work, as:

$$\text{EW} \equiv \frac{L_{Ly\alpha}}{L_{\lambda,UV}} \quad (2.3)$$

Where  $L_{Ly\alpha}$  is the  $Ly\alpha$  luminosity and  $L_{\lambda,UV}$  is the flux intensity redwards of the line.

Less massive stars ionize these gas clouds in lesser proportion than their more massive counterparts, making the gas emission spectra dimmer in comparison. In order to separate low SFR regions from higher SFR regions we analyze the  $H_\alpha$  emission line intensity. This line is an indicator of SFR since it is produced by gas ionized by young, hot stars [11]. Greater  $H_\alpha$  intensity corresponds to larger SFR and lower intensity to lesser SFR following the tendency represented in the following formula[12]:

$$\text{SFR}(\text{M}_\odot \text{yr}^{-1}) = 7.9 \times 10^{-42} L(H_\alpha)(\text{ergs s}^{-1}) \quad (2.4)$$

with  $L(H_\alpha)$  the luminosity of the emission line. The SFR is given in units of solar masses per year ( $\text{M}_\odot \text{yr}^{-1}$ ).

In simulations presented in [1], [3], stochastic effects causes the value of different line intensity ratios to fluctuate, with regions of low SFR being most affected. We can illustrate star formation stochasticity with an example: a weighted die that is more likely to cast a lower number. If the die is thrown many times then we will have a good sampling of the probability distribution. On the contrary, if the die is thrown only a few times, the probability distribution will not be sampled accurately.

In this analogy the lower numbers are the less massive stars and the die that is thrown many times is a region of high SFR. When a star forming locality has a high SFR, sampling the IMF is straightforward. If a region presents a low SFR the probability of finding massive stars decreases [3] and thus the IMF is not completely sampled. Since the relation of stellar luminosity and stellar mass is highly nonlinear[13], having an IMF that is not completely sampled makes the sequence in which stars fill a particular population's mass very significant in terms of luminosity output. In this case, the relation between the total mass and the luminosity of a stellar population is no longer deterministic [14]

There is also finite sampling in time since some evolutionary stellar phases of massive stars are brief. This is significant since the ionization luminosity has important variations in different phases [3].

Furthermore, the fact that most stars are formed in clusters gives rise to two additional effects: (i) the maximum stellar mass can not be greater than the cluster mass  $M_{\text{ecl}}$ , and (ii) at low SFR, a star forming region is not producing stars at a constant rate, instead, its output consists of a series of star formation bursts, with each burst corresponding to the formation of a new stellar cluster [3].

*Finite samplings in both mass and time, conditions upon maximum mass and irregular bursts of star formation are the motivation for searching for stochasticity in these low SFR regions.*

In [3], Forero-Romero et. al. found that when stochastic effects are taken into consideration the measured EW varied from  $\sim \text{EW}_0/4$  to  $\sim 3 \times \text{EW}_0$ , with  $\text{EW}_0$  the expected value for the EW. The authors describe the behavior of this effect as a SFR dependent probability density function (PDF) at a given SFR as a double power law of the form:

$$P(\mathcal{M}|\text{SFR}) = P_0 \left[ \left( \frac{\mathcal{M}}{\mathcal{M}_0} \right)^{-\alpha} + \left( \frac{\mathcal{M}}{\mathcal{M}_0} \right)^\gamma \right]^{-1} \quad (2.5)$$

with  $\mathcal{M}$  a dimensionless parameter defined as:

$$\mathcal{M} \equiv \frac{\text{EW}}{\text{EW}_0} \quad (2.6)$$

and  $P_0$  is a normalization constant and  $\alpha$  and  $\gamma$  are free parameters that depend on the given SFR.

In the same work, the authors assume that the totality of star formation happens in clusters ( $f_c = 1$ ). This reduces the fluctuation of the emission line flux by approximately an order of magnitude [15]. The authors obtain this answer considering a range of stellar masses from  $0.08\text{M}_\odot$  to  $120\text{M}_\odot$  and a cluster mass interval spanning values of  $20\text{--}10^7\text{M}_\odot$ , assuming a Salpeter IMF.

This double power law is valid for values of  $P(\mathcal{M}|\text{SFR}) > 10^{-2}\text{ M}_\odot\text{yr}^{-1}$  due to the limited number of points available to perform a thorough fit of the data. This PDF is visualized in figure 2.2 for multiple star formation rates. In this figure, it is clear that as SFR increases, the peak of the function becomes more pronounced. Ideally, if no external effects are present, this plot would consist of a delta function centered around  $\mathcal{M} = 1$  [1].

In a separate work, Mas-Ribas et. al. [4] also found fluctuation of the  $Ly_\alpha$  and  $\text{HeII}\lambda 1640$  luminosities when performing stochastic IMF sampling.

In absence of stochasticity, these theoretical results shown in figure 2.3 should consist of a thin line along the expected  $Ly_\alpha$  and  $\text{HeII}\lambda 1640$  luminosities [4]. Therefore, when stochasticity is present in simulations we see a distribution of luminosity around a mean value. This simulations were performed making a stochastic sampling of the IMF and taking periods of star formation bursts of  $1000\text{M}_\odot$  per Myr [4].

As we can see, stochastic effects in simulations cause fluctuation of multiple spectral parameters, including the  $Ly_\alpha$  EW and luminosity and the  $H_\alpha$  to FUV ratio. In this work, we look at the emission spectra of gas clouds

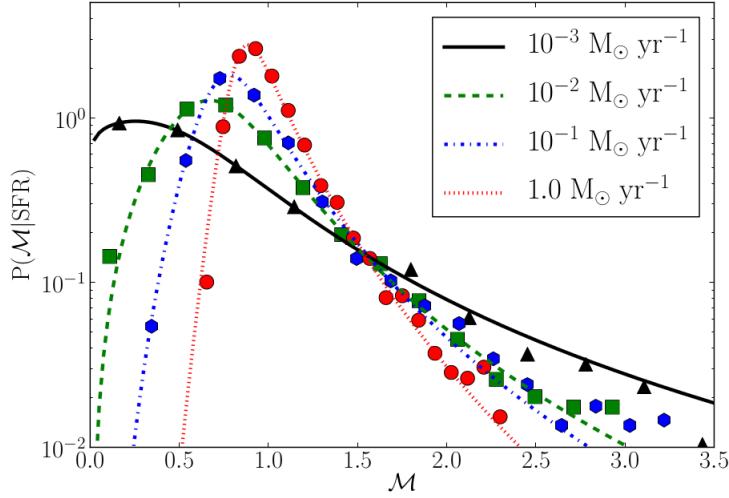


Figure 2.2: Visualization of equation 2.5 for different SFR. As we see, the distribution becomes narrower as SFR takes a larger value. Taken from [1]

and study the  $H_\alpha/H_\beta$  ratio for fluctuations around a mean value as found in simulations by [3].

Expanding on conclusions from [3] a study of the EW of the OII line is also presented. According to their theoretical results, stochasticity should produce a similar distribution in other emission lines. This means that performing our analysis and comparing the values of OII EW in regions of low SFR with values in regions of SFR should yield analogous results.

### 2.3 Measuring stochasticity

When measuring stochasticity it is important to take galactic extinction into account. Since the predicted effects of stochasticity is the fluctuation of spectral parameters, any effects that influence observational data need to be studied and quantified.

In this section the phenomenon of galactic extinction, also known as either the Balmer decrement or interstellar reddening, is detailed.

#### 2.3.1 The Balmer decrement

The Balmer decrement is a phenomenon that occurs when electromagnetic radiation passes through a dust cloud in the interstellar medium. These dust particles are micron sized, which causes a greater absorption of bluer

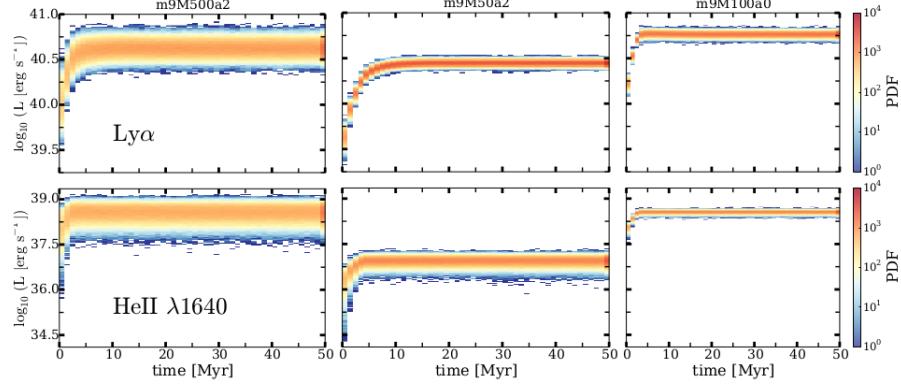


Figure 2.3: Fluctuation of the  $Ly\alpha$  and  $HeII\lambda 1640$  luminosity for a stochastically sampled IMF. When stochastic effects are taken into consideration, fluctuations in luminosity appear for different mass intervals.

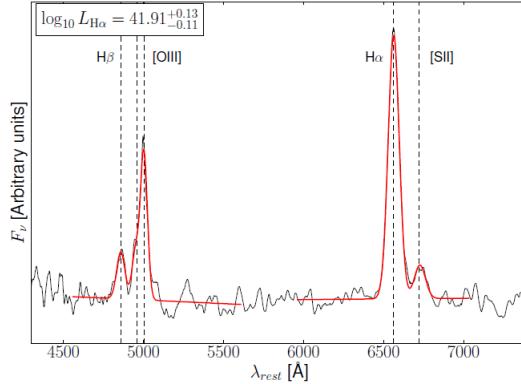


Figure 2.4: Difference in  $H_\alpha$  and  $H_\beta$  intensities. Taken from [16]

light as it is more likely to collide with a small particle. This means that when we observe light that has encountered intergalactic dust, it will appear redder.

When we analyze the  $H_\alpha/H_\beta$  ratio this absorption should be taken into account, since the  $H_\beta$  emission line has a shorter wavelength than  $H_\alpha$  and will more likely collide with dust particles and affect our measured value of  $H_\alpha/H_\beta$ . This attenuation of  $H_\beta$  means that when checking for the value of  $H_\alpha/H_\beta$  we will observe a greater number since  $H_\beta$  has become dimmer in comparison to  $H_\alpha$ . The difference of intensities for these lines can be seen in figure 2.4 [16].

This effect means that not every variation of the observed spectra is caused solely by stochasticity. In fact, to properly quantify the presence of

stochasticity, or lack thereof, we must isolate our data from any external effects presented by outside influence on the data used.

## Chapter 3

# The CALIFA survey

In this work, we use data from the CALIFA survey, which is a comprehensive wide field integral field unit (IFU) survey carried out in the Calar Alto observatory in Almería, Spain. It is operated jointly with the Max Plank Institut für Astronomie and the Instituto de Astrofísica de Andalucía. Observations began in June of 2010, with the survey collecting data from over six hundred galaxies in two configurations, the V500 and the V1200 set ups [5].

The V500 and V1200 configurations cover different wavelength intervals and have different resolution, with the V500 datacube having a resolution of  $\lambda/\Delta\lambda \sim 850$  and covering emission lines between 3745 – 7500 Å, while the V1200 configuration has a higher resolution of  $\lambda/\Delta\lambda \sim 1650$  and compiles wavelengths from 3700 – 4800 Å[5] [6].

The reasoning for selecting CALIFA survey data is simple. As seen in figure 3.1, CALIFA data does not consist of a single integrated spectra, instead, multiple optic fibers were pointed at a single galaxy [5] [6]. For previous surveys having a single measure of a spectra meant that for an entire galaxy there was only one reported value of total SFR. Since stochasticity theoretically appears in low SFR regions, the ability to focus on distinct regions that have lower SFR opens the possibility of searching for this stochastic effects.

In this work, we will be using “face on” galaxies from the CALIFA catalog published in [7]. These type of galaxies are optimal for a complete scanning and allow us to measure the spectra of individual regions within a galaxy. We did a manual selection of the datacubes in [7] to select galaxies whose galactic planes are perpendicular to our field of view.

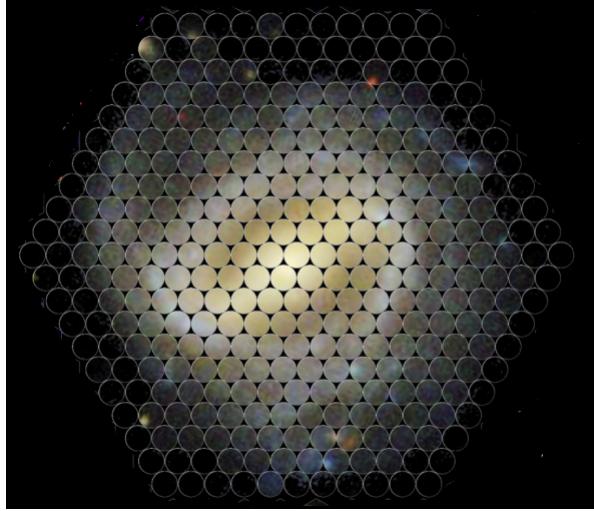


Figure 3.1: [http://califaserv.caha.es/CALIFA/DATA/Figs/CALIFA\\_HexRGB.jpg](http://califaserv.caha.es/CALIFA/DATA/Figs/CALIFA_HexRGB.jpg)

### 3.1 The CALIFA datacubes

The CALIFA datacubes consist of a three dimensional array, where the  $x$  and  $y$  coordinates indicate the right ascension and declination of the target, and the  $z$  axis contains information about flux intensity and wavelength of that pixel along with the error associated with those fluxes, a mask to cover bad pixels and the covariance weight of the error propagation [5] [6].

In this work, we will use the lower resolution V500 datacubes published in [7] which have been subjected to the Pipe3D analysis pipeline [7]. These datacubes contain a great deal of information, including that of intermediate steps of the fitting procedure.

### 3.2 The Pipe3D outputs

The Pipe3D pipeline proposed by Sánchez et.al [7] consist of a series of steps followed in order to organize and successfully provide a fit for the spectra in the CALIFA datacubes. Since the original data consists of a superposition of spectra of gas clouds in interstellar environments as well as the underlying star population, these two spectra have to be treated using several procedures. It is also important to make a distinction between weak and strong emission lines since fitting procedures are different for each. Since we will be studying  $H_{\alpha}$ ,  $H_{\beta}$  and OII lines, we will not include the fitting

procedure of weak emission lines.

### 3.3 Analysis of the stellar population

By applying an algorithm known as “CS-binning algorithm”, the CALIFA team has produced a set of spectral data with which they are able to study the spectra of the stellar population. This algorithm consists of dividing the datacubes into two dimensional cells called “spatial bins” and subsequently taking the mean value of the pixels included, masking pixels with bad values. Once the averaged spectra is accounted for, the CALIFA team applied the Pipe3D analysis pipeline, which is able to fit the observed data via an iterative procedure [7]. When applying this procedure, there appears an associated error for these new binned spectra since bad pixel data can adversely modify the resultant dataproduct. Pipe3D provides the associated error of each spaxel as part of the original datacube. The ultimate goal of this analysis is to be able to obtain an optimum depiction of the stellar population and subtract it from the datacubes, leaving the remaining data of the spectra associated exclusively with the emission lines of the gas clouds in interstellar regions of the galaxy. The results obtained by CALIFA are public and presented in [7].

After this binning procedure, several maps are created:

- Row Stack Spectra (RSS)
- A position table for each galaxies binned cube, corresponding to the indices of the spatial bins from brightest to dimmest.
- Two intensity maps at the wavelength range of the V-band, with one map corresponding to the intensity before the binning procedure and the other one to the binned data.

Once this results were obtained, the stellar population was fitted with a Simple Stellar Population (SSP) template and thus the systemic velocity, the velocity dispersion and dust attenuation could be derived.

The final step consists of fitting the strong emission lines and subtracting them from the datacube obtained at this stage. Then, the procedure is repeated until a certain condition of  $\chi^2$  is met. The fitting of strong emission lines is presented in the following section. The results obtained by CALIFA are public and presented in [7].

### 3.4 Analysis of strong emission lines

A Gaussian fit, though not sophisticated, is valid for fitting strong emission lines. The CALIFA team divides the observed wavelengths into four groups:

- $[O]_{II}\lambda 3727$  ( $3700 - 3750$ ) Å
- $H_\beta, [O_{III}]\lambda 4959$  and  $[O_{III}]\lambda 5707$  ( $4800 - 5050$ ) Å
- $[N_{II}]\lambda 6548, H_\alpha$  and  $[N_{II}]\lambda 6583$  ( $6530 - 6630$ ) Å
- $[S_{II}]\lambda 6717, [S_{II}]\lambda 6731$  ( $6680 - 6770$ ) Å

since we will not be taking into account gas-rich major merger galaxies, AGN nuclei nor galaxies that intercept in our observational foreground, this gaussian fit is valid for our work.

Before assigning a gaussian function to a particular emission line, authors in [7] do a first guess of kinematic properties based on the expected  $H_\alpha$  wavelength of the radiation. They take into account that particular galaxies redshift and by making a parabolic approximation to the centroid of the emission line.

Finally, the emission line is fitted by using a small range of systematic velocities centered around the initial guess. This procedure results in a set of maps of an emission line pure cube which includes different parameters for each emission line [7]. Once again, the results obtained by CALIFA are public and presented in [7].

# Chapter 4

## Results

For representative results of the analysis applied we will present four galaxies: (*i*) IC4566; (*ii*) NGC0001; (*iii*) NGC0036 (*iv*) UGC09476.

We begin by visualizing each galaxy's OII,  $H_{\alpha}$  and  $H_{\beta}$  intensities shown in figures 4.1 to 4.4. Then, we plot the  $H_{\alpha}/H_{\beta}$  ratio to visualize the values measured. Taking the logarithm in base ten gives us a clearer view of smaller variations in the values of this ratio. Since stochasticity would theoretically alter the values of this quotient, we will present the  $\log H_{\alpha}$  vs.  $\log H_{\alpha}/H_{\beta}$  distribution in more detail in figure 4.5. By visualizing the  $\log H_{\alpha}$  values plotted against the  $\log H_{\alpha}/H_{\beta}$  we can more easily identify regions of low SFR.

Making a histogram from the data in figure 4.5, we obtain figure 4.6. We separate the  $H_{\alpha}$  data in three regions. A bright region, a mid-range region and a faint region which correspond to high, mid and low SFR values as per equation 2.4. The values of SFR chosen as threshold values are: *i*)  $2.99 \times 10^{-2} M_{\odot} \text{yr}^{-1}$ , *ii*)  $1.89 \times 10^{-1} M_{\odot} \text{yr}^{-1}$ . This means that the dim region of the  $H_{\alpha}$  spectra will have SFR values lower than *i*), the mid region will have SFR values between *i*) and *ii*) and the bright region will have SFR larger than *ii*).

The bright region is plotted in red, the mid-range spectra is green and the faint spectra in blue. The distinction arises from the fact that stochastic effects appear in simulations when there is low SFR and therefore have a dimmer spectra. The segmented vertical line indicates the theoretical value for the  $H_{\alpha}/H_{\beta}$  ratio, which is 2.85<sup>1</sup> and whose value is theoretically affected by stochasticity.

---

<sup>1</sup>Osterbrock, Astrophysics of Planetary Nebulae and Active Galactic Nuclei, University Science Books, 1989

After plotting the data in figure 4.6 we perform a fit of this results and obtain a probability density function for values of  $H_\alpha/H_\beta$  across all regions of the spectra. The resulting curves are plotted for our four representative galaxies (figure 4.7). Estimations of the free parameters obtained for several other galaxies are plotted one against the other to facilitate comparison. These free parameters are the ones present in equation 2.5, which are  $P_0$ ,  $\mathcal{M}_0$ ,  $\alpha$  and  $\gamma$ .

**IC4566**

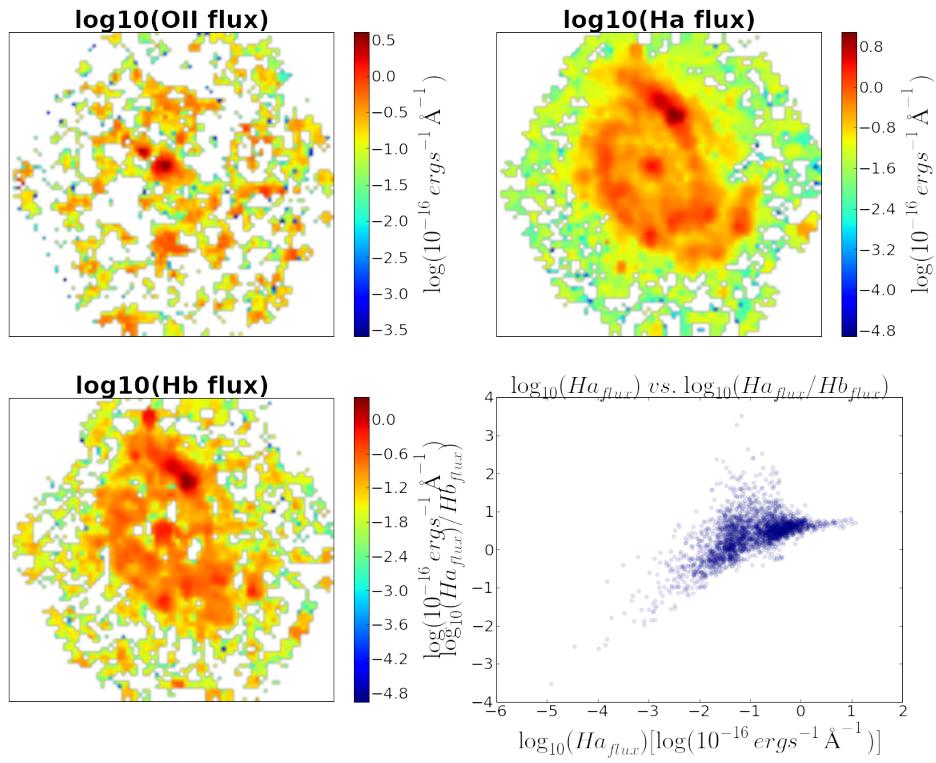


Figure 4.1: a) OII flux, b) H $\alpha$  flux, c) H $\beta$  flux, d)  $\log(H\alpha)$  vs.  $\log(H\alpha/H\beta)$

## NGC0001

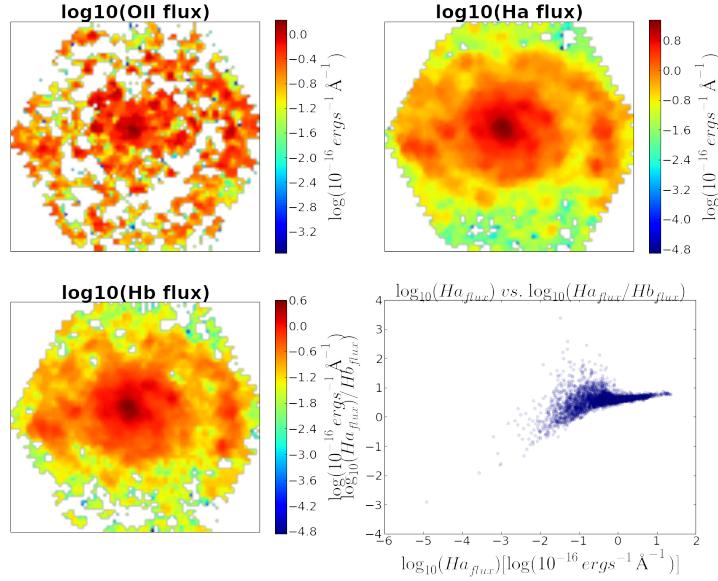


Figure 4.2: a) OII flux, b) H $_{\alpha}$  flux, c) H $_{\beta}$  flux, d) log(H $_{\alpha}$ ) vs. log(H $_{\alpha}$ /H $_{\beta}$ )

## NGC0036

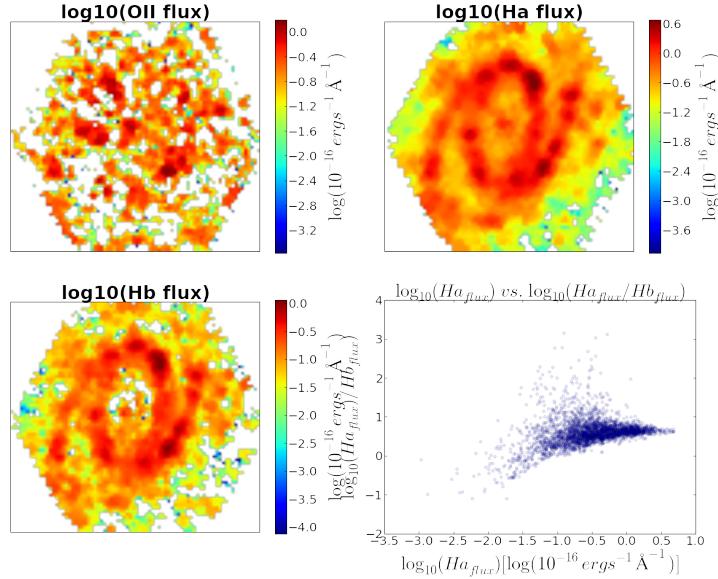


Figure 4.3: a) OII flux, b) H $_{\alpha}$  flux, c) H $_{\beta}$  flux, d) log(H $_{\alpha}$ ) vs. log(H $_{\alpha}$ /H $_{\beta}$ )

## UGC09476

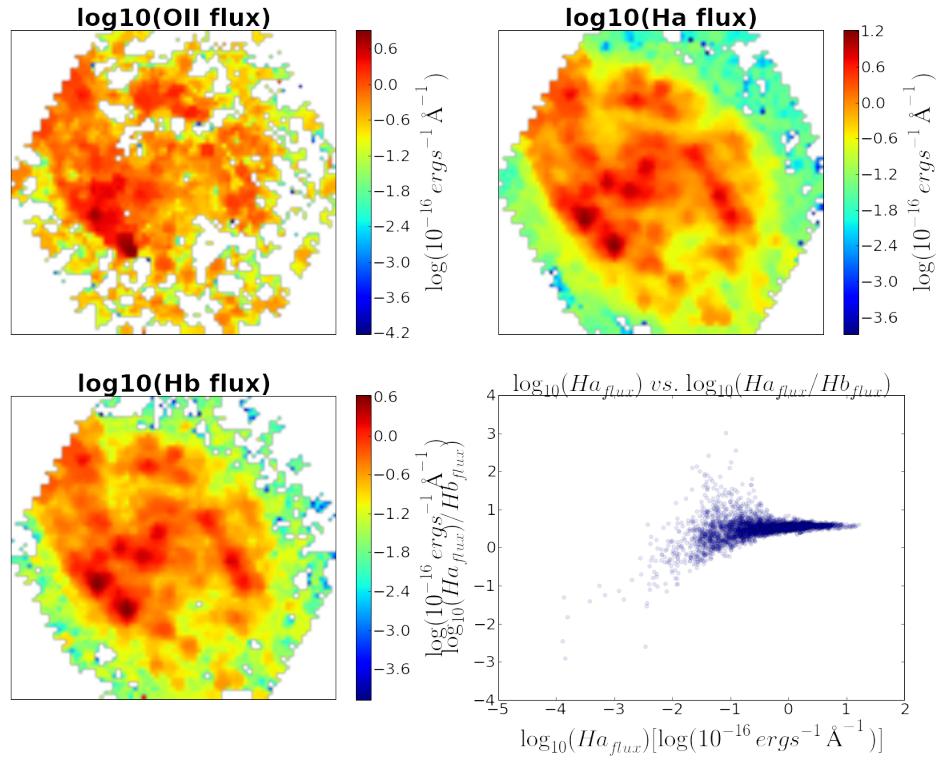


Figure 4.4: a) OII flux, b)  $\text{H}_\alpha$  flux, c)  $\text{H}_\beta$  flux, d)  $\log(\text{H}_\alpha)$  vs.  $\log(\text{H}_\alpha/\text{H}_\beta)$

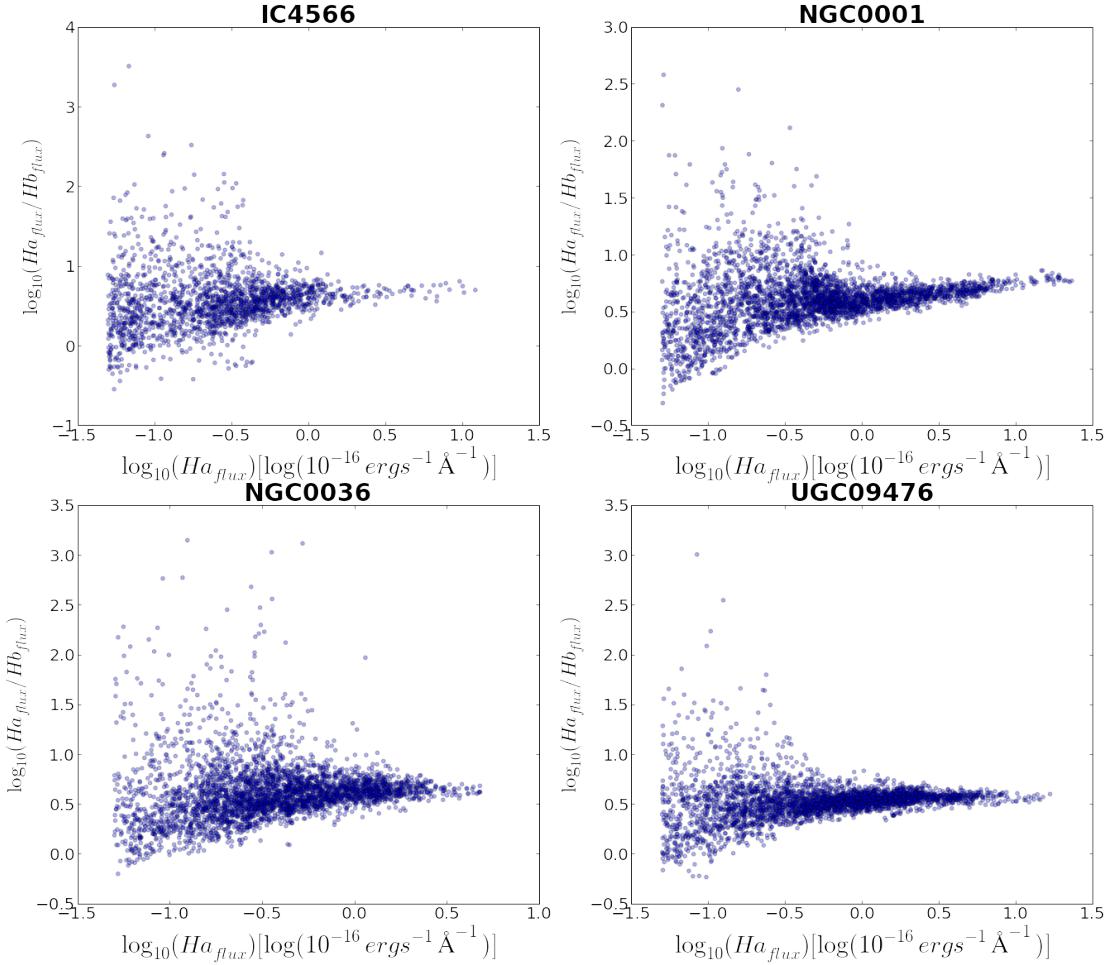


Figure 4.5: Distribution of  $\log H_\alpha / H_\beta$  values for our representative galaxies graphed on the  $y$  axis. The  $H_\alpha$  intensity (i.e. the SFR) is depicted on the  $x$  axis.

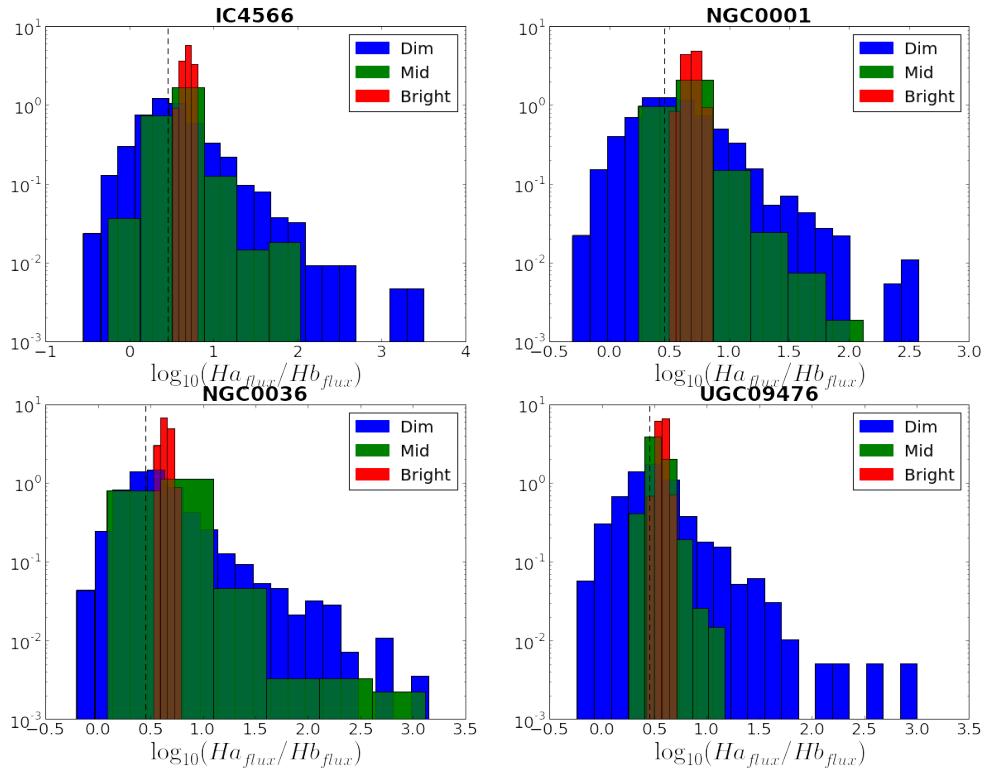


Figure 4.6: Distribution of  $\log H\alpha/H\beta$  values for our representative galaxies plotted as histograms. The blue, green and red histograms correspond to the faint, mid range and bright regions of the  $H\alpha$  spectra respectively. The segmented line lies on the theoretical value of the ratio.

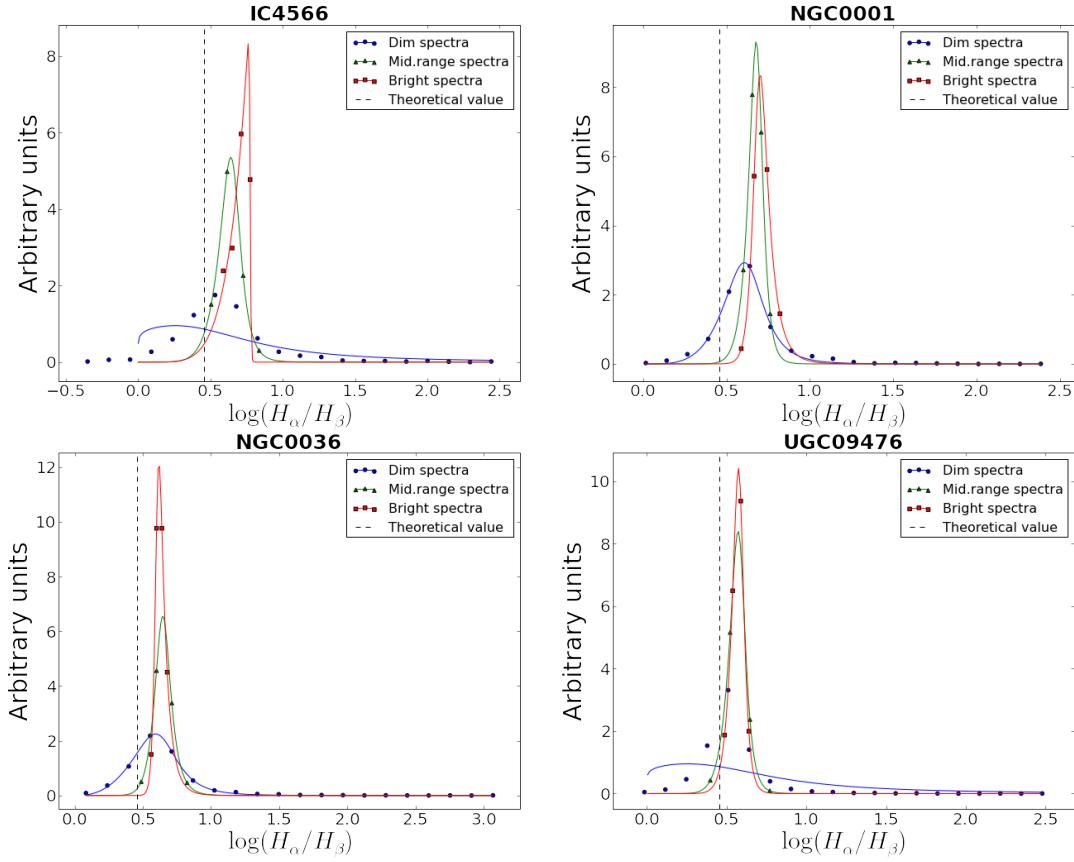


Figure 4.7: Distribution of  $H_\alpha/H_\beta$  values, fitted as a double power law of the form shown in equation 2.5. The segmented line indicates the logarithm in base ten of the theoretical value of the  $H_\alpha/H_\beta$  ratio.

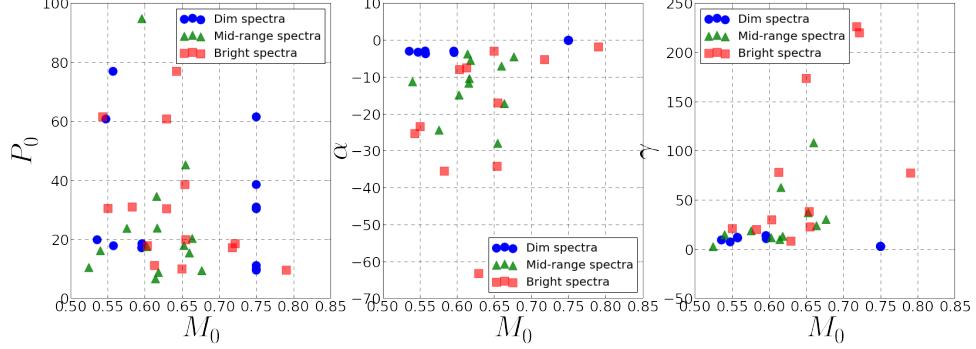


Figure 4.8: Comparison of values of the free parameters: a)  $M_0$  vs.  $P_0$ ; b)  $M_0$  vs.  $\alpha$ ; c)  $M_0$  vs.  $\gamma$

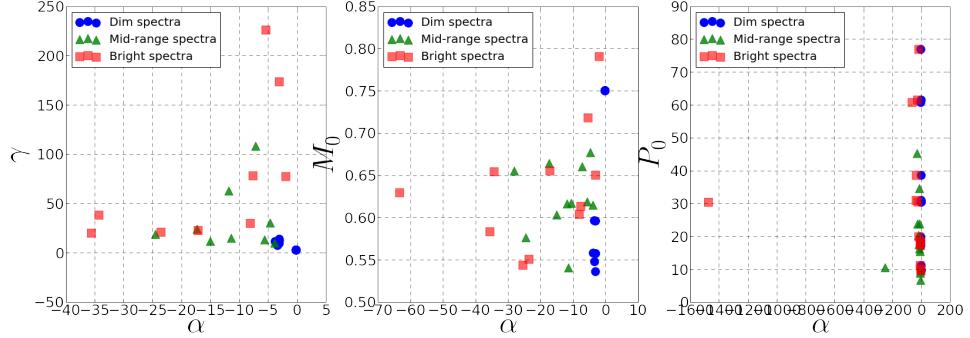


Figure 4.9: Comparison of values of the free parameters: a)  $\alpha$  vs.  $\gamma$ ; b)  $\alpha$  vs.  $M_0$ ; c)  $\alpha$  vs.  $P_0$

We also compute the distribution for the OII EW. Once again, we separate our EW data into bright  $H_\alpha$ , mid  $H_\alpha$  and faint  $H_\alpha$ . This separation of EW data follows the same criteria as the one performed on the  $H_\alpha/H_\beta$  ratio.

A visualization of the OII EW maps is given in figure 4.10 along with a histogram of the data. Results for the representative galaxies are displayed in figure 4.11.

In order to have a clear understanding of the behavior of our data, we plot the values of the variance and skewness of the OII EW data in figure 4.12. Once again separating the values in the three categories of  $H_\alpha$  intensity, spanning the dim, the mid-range and the bright regions of the spectra.

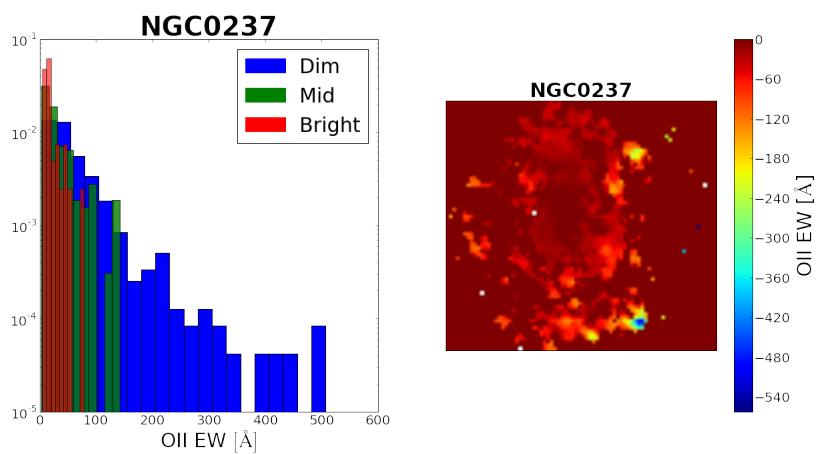


Figure 4.10: a) A histogram showcasing the distribution of values for the OII EW Map of the OII EW in NGC0237. b) Map of the OII EW in NGC0237.

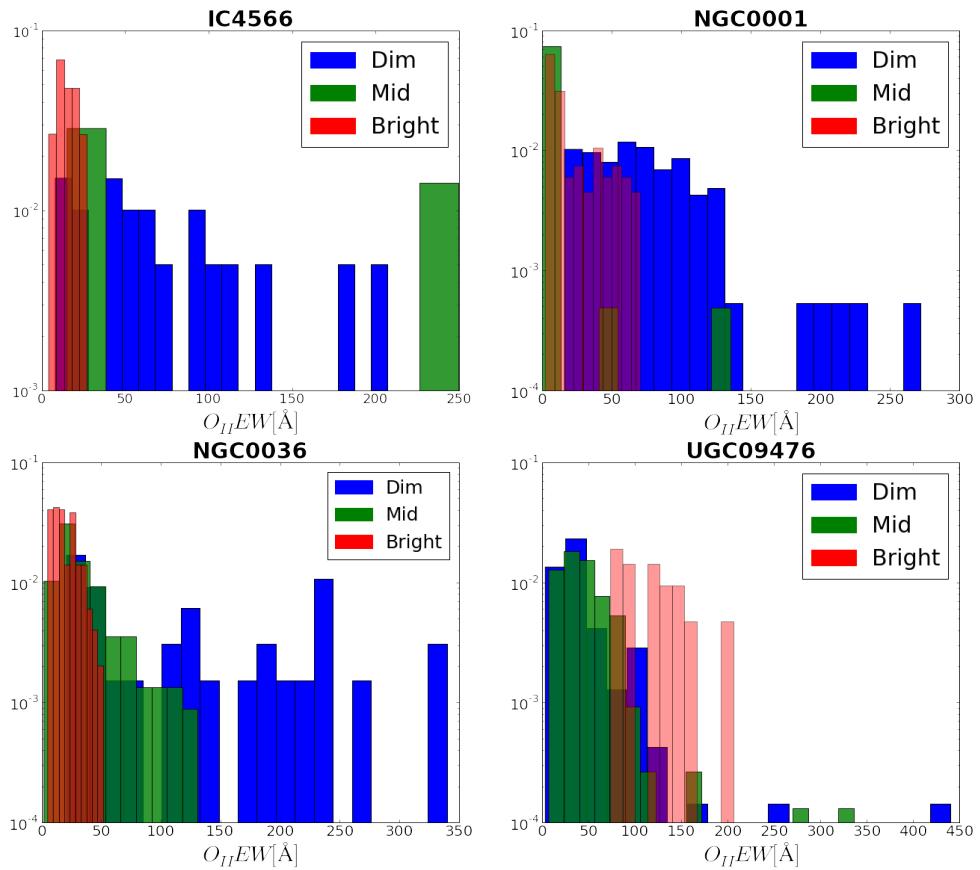


Figure 4.11: Distribution of the OII EW. The blue, green and red histograms correspond to the faint, mid and bright regions of the  $H_{\alpha}$  spectra respectively.

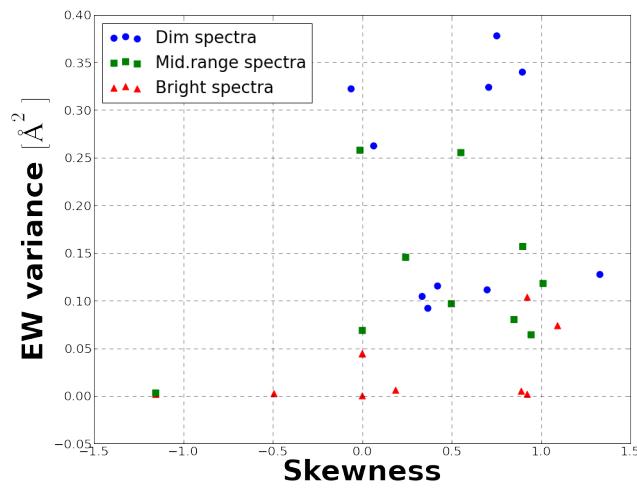


Figure 4.12: Variance versus skewness for 13 of the galaxies analyzed.

# Chapter 5

## Discussion

### 5.1 The $H_{\alpha}/H_{\beta}$ ratio

According to [3], the  $H_{\alpha}/H_{\beta}$  ratio fluctuates if stochasticity is present. Looking at the trimmed data in chapter four, figure 4.5, we can see that at a lower value of  $H_{\alpha}$ , the  $H_{\alpha}/H_{\beta}$  quotient starts becoming more disperse. This is consistent with the computational predictions since the  $H_{\alpha}$  intensity is an indicator of star formation productivity, with higher  $H_{\alpha}$  intensity corresponding to a higher SFR and a lower intensity to a lower SFR [11].

Once we visualize this distributions in figure 4.6, we can also recognize a distribution around the theoretical value of this ratio. In fact, the bright region of the  $H_{\alpha}/H_{\beta}$  spectrum presents a more narrow distribution while the dimmer area presents a broader distribution.

In [1], Forero-Romero expands upon Fumagalli's [3] conclusions and proposes that this fluctuation should be present on other emission line ratios. We find looking at the histograms in figures 4.11 that distributions are also present in the OII EW. These distributions appear to behave in the same manner as the  $H_{\alpha}/H_{\beta}$  ratio, with higher values presenting a lesser variation and lower values being more broadly scattered.

Even though we have found evidence of fluctuation of  $H_{\alpha}/H_{\beta}$  and EW, we must be careful to consider this as evidence of stochasticity since interstellar reddening can cause similar effects. These reddening effects show that not all fluctuations of the  $H_{\alpha}/H_{\beta}$  are due to stochasticity. Despite this, we can see from the histograms in figure 4.6 that this effect does not fully explain the distributions observed, as it only causes the values of the  $H_{\alpha}/H_{\beta}$  ratio to move to the right. We can observe that in our histograms (figure 4.6) there are values of  $H_{\alpha}/H_{\beta}$  that lie to the left of the theoretical value and are

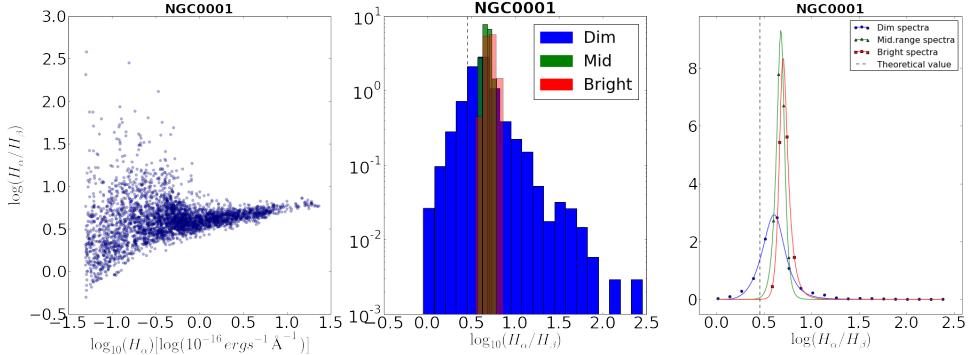


Figure 5.1: Results for NGC0001 including: a) Trimmed data of  $\log H_\alpha$  vs.  $\log H_\alpha/H_\beta$ , b) Histogram of the data and c) Fitting of the probability distribution

not explained by this phenomenon. In particular, the majority of the data that lies to the left hand side of the theoretical value belongs to the  $H_\alpha/H_\beta$  that belongs to the dim region of the  $H_\alpha$  spectra. On the contrary, most of the mid-range and bright regions of the  $H_\alpha$  spectra lie to the right of the theoretical value.

When analyzing the fitted curves for the histograms, we can see that at a higher SFR, the peaks become narrower and higher. This is also expected from the results found in [1]. In their results, the absence of stochasticity meant that the probability density function  $P(\mathcal{M}|\text{SFR})$  presented in 2.5 would become a delta function centered around  $\mathcal{M} = 1$ .

This dispersion of values for the line ratio seems to be consistent with the theoretical prediction, but before drawing any conclusions we must treat the data to minimize external influences.

## 5.2 The OII EW

When analyzing the histograms in figure 4.11 we can observe that the histogram depicting the dim region of the spectra tends to be more broadly distributed than its more intense counterparts. The green bars indicating the regions of higher SFR are relatively more clustered the left side of the graph, and the bright region of the spectra presents the less distribution. This results seem to hint that stochasticity may be involved in the measured EW, but a correction for external effects must be performed.

Looking at figure 4.12 we can see that most of the galaxies presented have a positive skew, which seems to be in agreement with equation 2.5 which is clearly positively skewed.

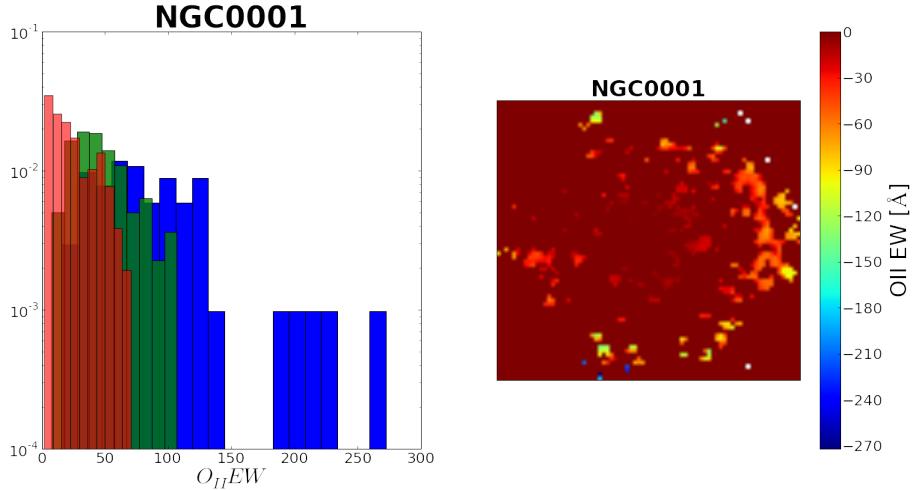


Figure 5.2: a) A histogram of the OII EW data. b) Map of the OII EW values in NGC0001.

Despite encouraging signs found in the OII EW, in some galaxies that were measured in the CALIFA survey there seems to be a considerable amount of low signal to noise. The data in these was somewhat scarce, as it spanned only a few regions of the galaxy and sometimes had few values of EW which generated poor quality histograms of the data.

Yet, a great deal of galaxies presented good data with which more work must be done in order to be able to name stochasticity as a candidate for the fluctuating effects of EW for multiple emission lines. For example, in the galaxy NGC0001 that we have presented in detail in this work, we can see in figure 5.2 that even though the data is not particularly abundant, the EW of OII covers a range of values sufficiently large to work with.

# Chapter 6

## Conclusions

In this work we have set out to detect stochastic influence in the form of the fluctuation of several spectral parameters such the  $H_{\alpha}/H_{\beta}$  ratio and the EW of the OII emission line. The method by which this was done was to discriminate between regions of high SFR and low SFR by way of that particular region's  $H_{\alpha}$  intensity. Having identified these regions, we proceeded to illustrate the distributions of the values of the spectral parameters of interest to this work. Once we obtained these distributions, a fit was performed, following the results obtained in [1].

We have found that fluctuations of both the  $H_{\alpha}/H_{\beta}$  ratio and EW of OII are indeed present in observational data.

Despite the influence of interstellar dust effects which were not accounted for in this work, we can say that the fluctuations detected can not be fully explained by extinction effects. These extinction effects only produce fluctuation of parameters in one direction, that is, it causes the observed  $H_{\alpha}/H_{\beta}$  ratio to take larger values than its theoretical value. Furthermore, high extinction values are usually associated for gas-dense regions with high SFR and would therefore not be influenced by stochastic effects.

At first glance, it appears that stochasticity is a good candidate to explain the values of this  $H_{\alpha}/H_{\beta}$  quotient that are measured to be of lesser value of the theoretical ratio although further work is required.

The fact that some fluctuation was also detected on the OII EW may also indicate that, as presented in [1], fluctuation of spectral quantities is also found in other emission line ratios, despite [1] and [3] focusing on the  $Ly_{\alpha}$  EW and the  $H_{\alpha}$  to FUV ratio.

A successful characterization of the behavior of the distribution of values of  $H_{\alpha}/H_{\beta}$  was obtained based on the computational predictions published

in [1], including the fact that this distribution appears to be dependent on the SFR as predictions stated.

## 6.1 Further work

Further work on the theoretical behavior of the distribution of EW at different SFRs is needed. Expanding on the conclusions derived in [1] we can more reliably quantify the effects of stochasticity.

In order to more be able to suggest stochasticity as a significant influence in regions of low SFR we must perform a correction for dust on our data. Applying this correction will allow us to analyze our results with minimal interference from external sources. To do this we must use an extinction map in order to quantify the influence of interstellar reddening and then perform our analysis on the treated data and determine if stochasticity is observed.

The fitting procedure returned the initial guesses of the free parameters on several occasions, which indicates the necessity to obtain a more educated guess for the first iteration of this procedure. Initial guesses were taken from computational results presented in [1], though that work focused solely on the  $Ly\alpha$  emission line.

A more in depth analysis of the OII EW must be done, although it will not include the same number of galaxies as the  $H_\alpha/H_\beta$  ratio since some data is missing or is not suitable for performing any analysis. Moreover, a more detailed description of the expected behavior of the OII EW dispersion is required.

Expanding our dataset should also provide us with further statistical soundness for any future work. The MaNGA survey presents an opportunity to apply our analysis to nearly 10000 galaxies[17].

# Bibliography

- [1] J. E. Forero-Romero and M. Dijkstra. “Effects of star-formation stochasticity on the Ly $\alpha$  and Lyman continuum emission from dwarf galaxies during reionization”. In: 428 (Jan. 2013), pp. 2163–2170. DOI: [10.1093/mnras/sts177](https://doi.org/10.1093/mnras/sts177). arXiv: [1206.0726](https://arxiv.org/abs/1206.0726).
- [2] C. J. Lada and E. A. Lada. “Embedded Clusters in Molecular Clouds”. In: 41 (2003), pp. 57–115. DOI: [10.1146/annurev.astro.41.011802.094844](https://doi.org/10.1146/annurev.astro.41.011802.094844). eprint: [astro-ph/0301540](https://arxiv.org/abs/astro-ph/0301540).
- [3] M. Fumagalli, R. L. da Silva, and M. R. Krumholz. “Stochastic Star Formation and a (Nearly) Uniform Stellar Initial Mass Function”. In: 741, L26 (Nov. 2011), p. L26. DOI: [10.1088/2041-8205/741/2/L26](https://doi.org/10.1088/2041-8205/741/2/L26). arXiv: [1105.6101](https://arxiv.org/abs/1105.6101).
- [4] L. Mas-Ribas, M. Dijkstra, and J. E. Forero-Romero. “Boosting Ly $\alpha$  and HeII 1640A Line Fluxes from Pop III Galaxies: Stochastic IMF Sampling and Departures from Case-B”. In: *ArXiv e-prints* (Sept. 2016). arXiv: [1609.02150](https://arxiv.org/abs/1609.02150).
- [5] S. F. Sánchez et al. “CALIFA, the Calar Alto Legacy Integral Field Area survey. I. Survey presentation”. In: 538, A8 (Feb. 2012), A8. DOI: [10.1051/0004-6361/201117353](https://doi.org/10.1051/0004-6361/201117353). arXiv: [1111.0962](https://arxiv.org/abs/1111.0962).
- [6] J. Sánchez Almeida et al. “Qualitative Interpretation of Galaxy Spectra”. In: 756, 163 (Sept. 2012), p. 163. DOI: [10.1088/0004-637X/756/2/163](https://doi.org/10.1088/0004-637X/756/2/163). arXiv: [1207.3928](https://arxiv.org/abs/1207.3928).
- [7] S. F. Sánchez et al. “Pipe3D, a pipeline to analyze Integral Field Spectroscopy Data: II. Analysis sequence and CALIFA dataproduccts”. In: 52 (Apr. 2016), pp. 171–220. arXiv: [1602.01830](https://arxiv.org/abs/1602.01830) [[astro-ph.IM](#)].
- [8] M. Geha et al. “The Stellar Initial Mass Function of Ultra-faint Dwarf Galaxies: Evidence for IMF Variations with Galactic Environment”. In: 771, 29 (July 2013), p. 29. DOI: [10.1088/0004-637X/771/1/29](https://doi.org/10.1088/0004-637X/771/1/29). arXiv: [1304.7769](https://arxiv.org/abs/1304.7769).

- [9] E. E. Salpeter. “The Luminosity Function and Stellar Evolution.” In: 121 (Jan. 1955), p. 161. DOI: 10.1086/145971.
- [10] P. Kroupa. “On the variation of the initial mass function”. In: 322 (Apr. 2001), pp. 231–246. DOI: 10.1046/j.1365-8711.2001.04022.x. eprint: astro-ph/0009005.
- [11] L. J. Kewley et al. “The H $\alpha$  and Infrared Star Formation Rates for the Nearby Field Galaxy Survey”. In: 124 (Dec. 2002), pp. 3135–3143. DOI: 10.1086/344487. eprint: astro-ph/0208508.
- [12] R. C. Kennicutt Jr. “Star Formation in Galaxies Along the Hubble Sequence”. In: 36 (1998), pp. 189–232. DOI: 10.1146/annurev.astro.36.1.189. eprint: astro-ph/9807187.
- [13] JE Andrews et al. “AN INITIAL MASS FUNCTION STUDY OF THE DWARF STARBURST GALAXY NGC 4214Based on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with program GO-11360.” In: *The Astrophysical Journal* 767.1 (2013), p. 51.
- [14] R. L. da Silva, M. Fumagalli, and M. R. Krumholz. “SLUG - Stochastically Lighting Up Galaxies - II. Quantifying the effects of stochasticity on star formation rate indicators”. In: 444 (Nov. 2014), pp. 3275–3287. DOI: 10.1093/mnras/stu1688. arXiv: 1403.4605.
- [15] Robert L. da Silva, Michele Fumagalli, and Mark Krumholz. “SLUGStochastically Lighting Up Galaxies. I. Methods and Validating Tests”. In: *The Astrophysical Journal* 745.2 (2012), p. 145. URL: <http://stacks.iop.org/0004-637X/745/i=2/a=145>.
- [16] A. Domnguez et al. “Dust Extinction from Balmer Decrement of Star-forming Galaxies at 0.75  $z$  1.5 with Hubble Space Telescope/Wide-Field-Camera 3 Spectroscopy from the WFC3 Infrared Spectroscopic Parallel Survey”. In: *The Astrophysical Journal* 763.2 (2013), p. 145. URL: <http://stacks.iop.org/0004-637X/763/i=2/a=145>.
- [17] K. Bundy et al. “Overview of the SDSS-IV MaNGA Survey: Mapping nearby Galaxies at Apache Point Observatory”. In: 798, 7 (Jan. 2015), p. 7. DOI: 10.1088/0004-637X/798/1/7. arXiv: 1412.1482.