Observsational evidence of star formation stochasticity in the CALIFA dataset

Nicolás Romero Díaz
201127499
Advisor: Jaime Forero-Romero

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Star formation processes are the mechanisms that turn galactic gas clouds into stars. These processes are probabilistic by nature and are mostly 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 mass 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.

The second parameter arises from the fact that star formation is mainly a cluster phenomenon, 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.

The sampling of these two parameters of star formation is intrinsically related with the star formation rate (SFR) of a particular system. This correlation arises from the fact that less massive stars will dominate the bulk of star formation. Due to this, regions of low SFR will not produce enough stars to account for the less probable massive stars. Fumagali et. al (2011) [1] 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_α flux to Far Ultra Violet (FUV) photons to fluctuate around a mean value [1].

Using the public SLUG code (Stochastically Light Up Galaxies), Forero-

Romero and Dijkstra [2] have found that this fluctuation around a mean value also takes place when comparing the ratio of Lyman alpha (Ly_α) flux to the FUV photons. This relation is a well known spectral quantity called the equivalent width (EW), which characterizes the Ly_α emission line strength. The interest in the EW and its fluctuation arises from the fact that the Ly_α line is a powerful indicator of stellar populations and intergalactic dust and can be a determining factor for classification of galaxies [2].

In order to search for this stochastic effects, we will be using data published by the Calar Alto Legacy Integral Field Area Survey (CALIFA), reported in detail by Sánchez et. al. [3, 4]. A compendious description of the CALIFA data will be presented in chapter two, as well as an overall description of the survey.

The minutiae and handling of CALIFA data is detailed in chapter three, as well as a discussion of the different computational tools used to navigate this data. This chapter will also including a general outline of the Pipe3D analysis pipeline developed by the CALIFA team of Sánchez et. al. [5]

Representative results are included in chapter four.

In chapter five we present a discussion of the results obtained. The detection and quantification of stochasticity is also detailed, as well as how spectral data can be affected by phenomena in the interstellar medium (e.g. galactic extinction, interstellar reddening).Finally, we present our conclusions in the last chapter, a catalog of images, full results and a table consisting of all galaxies researched for this work.

Chapter 2

The CALIFA survey

In this chapter we present an overview of the CALIFA survey, including several generalities concerning the CALIFA data. In particular, we discuss the galaxies observed and the selection criteria of the galaxies which are relevant to this work. In addition, a brief description of the two different configurations available for CALIFA data is presented.

2.1 The CALIFA survey

The CALIFA survey 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 Insititut für Astronomie and the Instituto de Astrofísica de Andalucía. Obervations 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 [3].

In this work, we will be using "face on" galaxies from the CALIFA catalog published in [5]. These type of galaxies are optimal for a complete scanning and allow us to measure the spectra of multiple regions within a galaxy. This implies that a manual selection of the datacubes in [5] must be done for galaxies that meet this criteria.

2.2 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 [3, 4].

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 resolution ${}^{\lambda}/_{\Delta\lambda} \sim 1650$ and measures from $3700 - 4800$ Å.[3, 4]

Chapter 3

The CALIFA dataset

In this chapter, we will discuss the details of the CALIFA data pertinent to this work. We will review the different computational tools that have been applied in order to organize and analyze the different CALIFA datacubes. Specifically, we will discuss the structure of said datacubes, how the `fitsio` library allows us to manipulate `.FITS` files and how the `Pipe3D` pipeline is applied to both the spectra of the stellar population of a galaxy as well as the spectra of giant gas clouds in a galaxy in order to obtain an optimum fit of the different spectra present in the datacubes.

3.1 The fitsio library

In order to successfully extract, manage and analyze the `.FITS` datacubes we will use the `fitsio` library, which allows us to navigate the three dimensional data matrices with ease. This library includes several commands that allows us to take slices from the three dimensional datacubes and thus handle data for each galaxy as if it were regular two dimensional matrix data in Python.

The `.FITS` format also allows each of its data sets to have one keyword, which can be used to reference different sets of data within a single datacube.

3.2 The Pipe3D analysis pipeline

The `Pipe3D` pipeline proposed by Sánchez et.al [5] 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.

In the case of this study, we are interested in fluctuations of the EW of the Ly_α line emitted by gas clouds. For this

3.2.1 Analysis of the stellar population

We are able to study the spectra of the stellar populaion using an algorithm known as "CS-binning algorithm", which consists of dividing the datacube into two dimensional cells called "spatial bins" and subsequently taking the mean value of the pixels included. Once we have the averaged spectra, the `PIPE3D` is able to fit the observed data via an iterative procedure [5]. By applying this procedure, we must be cautious to take into account the associated error of 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.

3.2.2 Analysis of strong and weak emission lines

A Gaussian fit, though not sophisticated, is valid for fitting strong emission lines, since we will not be taking into account gas-rich major merger galaxies, AGN nuclei nor galaxies that intercept in our observational foreground. `PIPE3D` divides the observed wavelengths into four groups (i) $(3700 - 3750)$ Å, (ii) $(4800 - 5050)$ Å, (iii) $(6530 - 6630)$ Å, (iv) $(6680 - 6780)$ Å. Finally, we make a parabolic approximation around the centroid of the emission line, this way we can perform the fitting procedure using a small range of systematic velocities centered around the initial approximation. This procedure results in a set of maps of an *emission line pure cube* which includes different parameters for each emission line [2]

— Hago el de las líneas de emisión débiles? No puedo describirlo superficialmente porque el análisis es muy técnico

Chapter 4

Results

Chapter 5

Discussion

5.1 Measuring stochasticity

5.1.1 The Balmer decrement

Chapter 6

Conclusions

Chapter 7

Apendix