

The relation between stars and gas in distant galaxies

Jacky Cao

Level 4 Project, MPhys Physics

Supervisor: Dr. Mark Swinbank

Department of Physics, Durham University

Submitted: 27th December 2018

Optical and spectroscopic observations of galaxies provide a way to understand the internal dynamics of the stars and gas within it. Data for the Hubble Ultra Deep Field (HUDF) from the Hubble Space Telescope and the Multi-Unit Spectroscopic Explorer (MUSE) on the VLT provides a potential sample of more than 10,000 galaxies. The data-set was reduced by considering objects with redshifts $z > 0.3$ and with V-band magnitudes brighter than 25.0 mag, the ~ 30 remaining galaxies were extracted and then analysed. Different regions of the galaxy spectra were fitted for. The [OII] emission region could be separated and fitted with a Gaussian doublet to represent the gas dynamics. Applying a Penalized Pixel-Fitting method (pPXF) to the entire spectrum could quantify the stellar dynamics, however various issues were encountered in this initial research which prevented the successful operation of the pPXF fitting routine. Future work would be based around isolating the issues and then extending the routine to consider variations to the spectra.

Contents

1. Introduction

- a. Galactic classification
- b. Galactic stellar formation
- c. Galactic data
 - 1. HUDF
 - 2. MUSE
- d. Project Aims

2. Analysis

- a. Cube extraction
- b. Line fittings and pPXF
 - 1. Gaussian doublet fitting
 - 2. pPXF fitting
 - 3. Voigt fitting

3. Conclusions

Acknowledgments

References

1. INTRODUCTION

- 1 Amongst the different types of cosmic
2 structure within our universe, galaxies can
3 be described as the most unique and diverse.
4 With each containing countless numbers of
4 stars and vast amounts of gas, dust, and
5 dark matter (Carroll and Ostlie 2007), it
5 would be surprising if these various inner
objects were not connected in any way.
5 Through observational astronomy the in-
5 ternal structure of galaxies and the motions
6 of the internal objects can be studied and
7 understood. With approximately $(2.0^{+0.7}_{-0.6})$
8 $\times 10^{12}$ galaxies in the universe up to $z = 8$
9 which in principle could be observed (Con-
9 selice et al. 2016), there is definitely not a
9 lack of choice. What is important is how
9 these objects are observed and how the col-
lected data is later analysed.
9 But it is always significant to understand
the basics, so developing a comprehension
of what a galaxy generally is and how they
can be defined and placed into different cat-
egories. Once an appreciation is built for
galactic classification, the intricacies of mo-
tions and inter-relationships can be explored
further.

a. Galactic classification

As stated previously, a galaxy can be quite broadly defined as a collection of gas, dust, stars and dark matter. But if a large enough sample was observed then one would begin to see that they can be grouped and classified together.

The most general categorisation is called the *Hubble Sequence* or the *Hubble Tuning Fork* (Carroll and Ostlie 2007). Developed by Edwin Hubble in 1926, galaxies can be roughly divided into ellipticals, spirals and irregulars (Fig. 1). With early Hubble type ellipticals along the horizontal handle, then the two prongs contain normal and barred spirals (later Hubble types), and irregulars as the third category. What can be seen from the Tuning Fork is a summarised view of the main galaxy types, however in reality there are more than the 11 named.

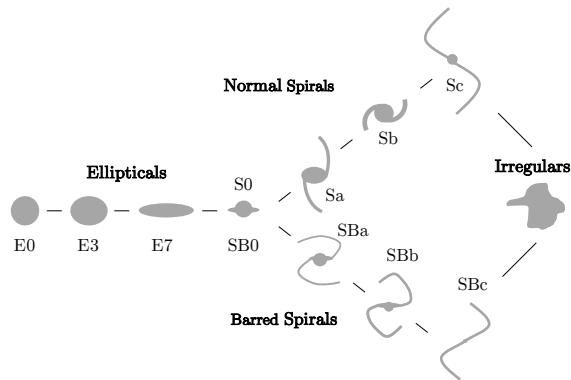


Figure 1: The Hubble Sequence displays the different morphologies of galaxies, they can be classified into three general groups: ellipticals, spirals, and irregulars. The former two can be broken up further, and from the diagram one can see an example pictogram and the respective classification name.

(*This diagram has been adapted from An Introduction to Modern Astrophysics by Carroll and Ostlie (2007).*)

The sequence itself does not show the evolution of the galaxies, rather it provides a way to view the different potential morphologies. So then, what does each grouping from the sequence actually represent?

Starting with the most broad, irregulars are objects which do not fall into the two

main galaxy types (ellipticals and spirals). In fact they themselves can be split into two sub-categories depending on if a particular galaxy could be seen to have structure, such as spiral arms (Carroll and Ostlie 2007). If they did then they would be Irr I galaxies, and if they appeared to be extremely disorganised then Irr II. Generally, irregulars are not particularly large, their diameters typically range from 1 to 10 kpc, and they have an absolute B-band magnitude of -13 to -20 mag.

This classification by appearance can easily be extended to ellipticals and spirals, and their individual component composition can be studied and explored as well.

With ellipticals, they span from being virtually spherical (E0) to highly flattened (E7) collections of objects and material (Moore and Rees 2011). Through observations of their stars, it can be seen that the majority of them are old and red types. This may be attributed to the amount of gas used in the initial stages of galactic stellar formation. If a larger proportion was used up initially then current observations would show the stellar-birthrate in ellipticals to be low (Carroll and Ostlie 2007). This could additionally explain the lack of disks which can otherwise be seen in spiral galaxies, if there is not enough surrounding material then the disk and arms features would not be able to form.

Exploring spirals, they can be described as being composed of a central nucleus with a surrounding disk of material. This disk contains denser regions which forces material to collapse and coalesce to form additional features. One of the characteristics leads to a supplemental classification, and that is whether the nucleus has a “bar” running through it. Observing and comparing the two different types, one finds that barred spirals appear to be more elongated along the bar axis than their non-barred counterparts.

The other main attribute for spiral galaxies are the protruding arms surrounding the main bulge (Carroll and Ostlie 2007). In the later type spirals (Sc and SBC) they have

arms which are more loosely wound than the earlier types (Sa and SBa) (Moore and Rees 2011). It can be seen implied that with more available “raw” gas and dust (Carroll and Ostlie 2007), spirals in their various forms could be assumed to have an overall younger stellar population.

It is within the arms of spiral galaxies that new stars are found to be created, however stellar birth is definitely not limited to just these regions. Inside the spiral structure the gravitational field allows for angular momentum to be transported outwards. Older and less massive stars in the galaxy produce a gravitational field which eventually leads to the shocking of interstellar gas (Binney 2016). As a result the density of the gas in the arms increase and certain regions then collapse to form new, young, blue, and massive stars.

There is therefore a range in the stellar population age of spiral galaxies. The spiral arms contain young stars whilst the central nucleus is akin to that of elliptical galaxies, where they have an older population of stars and fewer new stars are being created (Binney 2016, Carroll and Ostlie 2007).

These then are the general characteristics of galaxies, and whilst they can be classified based on their appearance, it is through studying their stellar populations that our understanding of them will improve.

b. Galactic stellar formation

One could argue that the main backbone in any galaxy lies with the gas and dust, as without them there would be no stars and galaxies would simply not exist. There are many models which aim to provide an explanation for galactic evolution, from an initial proto-nebulae to potential structure. It would be difficult to explain them all, but it would be beneficial to address some issues which would have to be answered by a comprehensive theory.

One such problem is deciding if an initial nebula collapses in free-fall or whether it goes through a slow and dissipative col-

lapse (Carroll and Ostlie 2007). If the time taken to cool the nebula significantly (cooling timescale) is much less than the time taken for free-fall, then the cloud would not be pressure supported and the collapse would be a rapid free-fall. On the other hand, if the cooling time exceeds the free-fall time then a rapid collapse cannot occur as the gas cannot radiate its energy away fast enough, so the gravitational potential energy released from collapse will heat the nebula adiabatically.

This is important to understand as from the heated and collapsed gas, the first stars can then be produced. It is not surprising then that a theory for galactic evolution should also be able to explain the rate of star formation. This rate can be described as a stellar birthrate function, $B(M, t)$, and be expressed as,

$$B(M, t)dMdt = \psi(t)\xi(M)dMdt, \quad (1)$$

where $\psi(t)$ is the star formation rate (SFR), $\xi(M)$ is the initial mass function (IMF), M is the stellar mass, and t is the time (Carroll and Ostlie 2007).

Describing what each term physically represents, $B(M, t)$ is the number of stars per unit volume with masses between M and $M + dM$ which are formed out of the interstellar medium (ISM) between time t and $t + dt$, $\psi(t)$ is the rate per unit volume at which mass in the ISM is converted into stars, and $\xi(M)$ is the relative number of stars which form in each mass interval (Carroll and Ostlie 2007).

Various problems arise when several researchers provide different assumptions for the different terms in the equation (Carroll and Ostlie 2007). Some say that the SFR is time-independent, whilst others describe it as exponentially decreasing with time, and a few even argue that the SFR is proportional to a power of the surface mass density of the galactic disk. Then take the IMF, there is disagreement on the exact form it takes, some model it to be a power-law as a function of mass, but it is not clear if it also varies with time or location.

Through contentious debate and research, astronomers can continue to build these different models and theories which attempt to explain galactic evolution. Whilst on the other hand, the evolution of the stars themselves are known and understood to a greater degree.

In a brief overview: stars are born from an interstellar medium where the gas and dust coalesce, collapse and heat-up until a protostar is created. Depending on the initial mass of material used, the star can then take various paths which will affect its subsequent future (McCoy 2012). To view the different stages of stellar evolution, Hayashi tracks and the Hertzsprung-Russell diagram can be paired together (Carroll and Ostlie 2007). In doing so, one would find that the majority of stars lie on the main-sequence, including our own Sun.

Type	Colour and notable lines
O	Hot blue-white (HeII, HeI)
B	Hot blue-white (HeI, HI)
A	White (Balmer, CaII)
F	Yellow-white (CaII, FeI, Cr,I)
G	Yellow (CaII, FeI, neutral metals)
K	Cool orange (CaII H and K, metals)
M	Cool red (TiO, VO, neutral metals)
L	Very cool, dark red, infrared (Molecular absorption, CrH, FeH, H ₂ O, CO, Na, K, Rb, Cs)
T	Coolest, infrared (CH ₄ , CO)

Table I: The Harvard Spectral Classification for stellar objects. Stars are defined by their colour and from the features seen from their spectra. Provided in the table are the spectral type, the associated colour plus the main emission or absorption lines to consider.
(Table adapted from *An Introduction to Modern Astrophysics* by Carroll and Ostlie (2007).)

The classification of stars can be made with the *Harvard Spectral Classification* where stars are classed by their spectral types (Table I). It essentially orders stars

by their observed temperature and their internal composition. If a star's spectra was observed, it could be identified by the absorption and emission features found in the sequence.

This is therefore one of ways to study the stellar population of a galaxy as the signatures of different spectral types can be matched up in the spectra of a particular galaxy.

It is clear then that observational data would be required for own study of galaxies and the internal motions of the objects within them.

c. Galactic data

For an investigation of galaxies, one would have to collect data optically and also spectroscopically. The former would be useful as it would allow for a quick identification of the galaxy based on its appearance, this could then be used to infer whether a galaxy would be worth studying. Spectroscopic data on the other hand would be important to collect as a variety of information could be obtained such as the type of stars within a galaxy, the types of gas in surrounding clouds, and also the redshift z .

Astronomers on a daily basis employ an assortment of telescopes and spectrographs, but there are just a few particular instruments which are worth discussing as they have produced remarkable datasets and results.

1. HUDF

The Hubble Space Telescope (HST) has been operating since April 1990 and for the duration of that time it has collected over 50 terabytes of data (McCoy 2012). From observing nebulae to attempts to detect extra-solar planets, HST has been used in a variety of searches. However, one of the most impressive and insightful uses of HST was the collection of the long-exposures taken in 1996 to produce the Hubble Deep Field (HDF).

Through a combination of its visible and infrared cameras, the HDF was produced by piecing together different datasets which were collected over a period of 10 consecutive days (McCoy 2012, Williams et al. 1996). The four-band, 0.5 million s exposure revealed sources with redshifts $z > 1$ which would have been difficult to view from ground-based observations (Beckwith et al. 2006).

It is therefore unsurprising that another attempt was made in 2011 which resulted in the Hubble Ultra Deep Field (HUDF). This survey utilised 1 million s of exposure time on an 11 arcmin² area in the southern sky. Viewed again with four different filters, this survey resulted in at least 10,000 objects being detected in the field (Beckwith et al. 2006). With a large majority being galaxies, the HUDF would be a prime dataset for a study of them as one would have a large range of candidates spanning a large proportion of the universe's history.

2. MUSE

In addition to optical observations, to truly study the HUDF objects, spectroscopic data must also be employed. One particular instrument which was used to acquire data on the HUDF region is the Multi-Unit Spectroscopic Explorer (MUSE) mounted to the Nasmyth platform on the Very Large Telescope (VLT) (Henault et al. 2003).

The MUSE Hubble Ultra Deep Field Survey was performed over a two year time-frame from September 2014 to February 2017 (Bacon et al. 2017). It made use of 137 h of observation time and covers 90% of the HUDF region. The 3D composite data cube mosaic made use of nine fields and represents the deepest spectroscopic survey ever performed. Both HST and MUSE HUDFs can be optically compared together, the latter is achieved by collapsing the cube axis containing the spectroscopic data into three components representing (R,G,B). Looking at both images (Figure 2), the HST data

appears to have more resolution whilst the MUSE objects are smeared out. However this is of no particular issue as the main use for MUSE is the manipulation of the spectroscopic data which in turn will help us inform the optical images.

Through a coupling of both, the dynamics of stars and gas in galaxies can therefore be studied. The main analysis arising from the manipulation and comparison of different component areas of the spectra.

d. Project Aims

This paper discusses the study undertaken to understand the dynamics between the gas and stars in galaxies. Data extraction is performed on a MUSE data cube, the sample is reduced, and various fittings are then applied to the data set.

In Section 2, the methodology of the galaxy extraction from the MUSE cube is discussed, as well as the subsequent analysis performed by applying different fitting routines: [OII] Gaussian doublet fittings, Voigt profile fittings, and fitting with a penalised pixel-fitting method.

2. ANALYSIS

The initial goal in the analysis was to reduce down the MUSE data cube into individual galaxy spectra. These could then be collapsed to produce a single galaxy integrated spectrum to represent the dynamics of the entire galaxy. Subsequently different fitting routines could be applied which would allow for a comparison of the gas to the stars in a galaxy.

a. Cube extraction

To begin with, galaxies were extracted from the 3D MUSE cube using *SExtractor*; an automated routine which detects, de-blends, measures and classifies sources (Bertin and Arnouts 1996). Employed in the astronomical software pack-

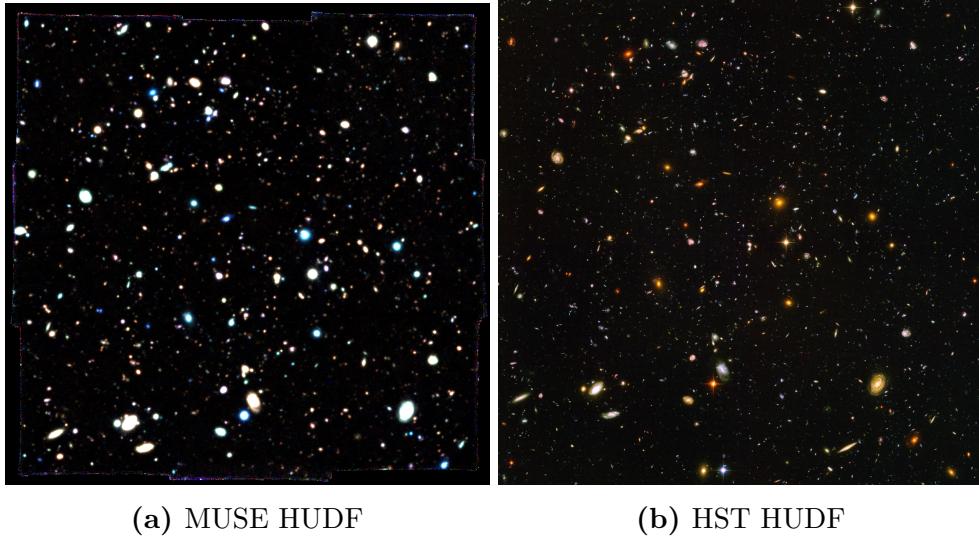


Figure 2: (a) A colour image created from the MUSE spectroscopic data of the HUDF. The wavelength range was split into three equal regions and then collapsed to create three bands (R, G, B). A final colour image was produced by combining these separate frames together. (b) The optical HUDF as captured by the Advanced Camera for Surveys instrument on the Hubble Space Telescope (NASA et al. 2018).

age *GAIA* (Currie et al. 2013), every potential galactic source was retrieved along with a “segmentation map” which discerns between multiple galaxies if several are close together.

The final objects from SExtractor were compared and merged with a catalogue which contained every object identified in the HST HUDF frame. The latter would contain information such as initial guesses for redshifts, the magnitudes of the objects in the different wavelength filters, and various flux values for different emission and absorption lines.

With every extracted galaxy, the signal-to-noise (S/N) can be calculated with the flux values from the catalogue this can then be plotted against the V-band magnitude (Figure 3(a)). This should show that as the S/N increases, the magnitude of the object should increase as well. A similar trend should also be found if the S/N was measured using regions of the galaxy-integrated spectrum (Figure 3(b)). The “signal” calculated by taking the median of the region and the “noise” as the standard deviation. With both methods of finding the S/N, the two plots produced agree and therefore further

analysis should in theory produce sensible results.

b. Line fittings and pPXF

After extracting the individual galactic objects from the MUSE cube, the spectroscopic data could then be manipulated and analysed using different fitting routines: (i) [OII] doublet fittings, (ii) *pPXF* absorption and emission line fittings, and (iii) Voigt profile fittings.

It would be quite non-sensical to use the entire catalogue of objects as there are more than 10,000 and the majority are not high enough in S/N for spectra evaluation. Therefore several sample reducing methods were applied: (1) galaxies were sorted by redshift and then objects below $z < 0.3$ ignored, (2) sample was then re-sorted by V-band magnitude and just the top 300 brightest objects extracted, and finally (3) the final set was found by considering galaxies with a magnitude brighter than 25.0 mag.

With this reduced sample in hand, the model fitting analysis could then be performed.

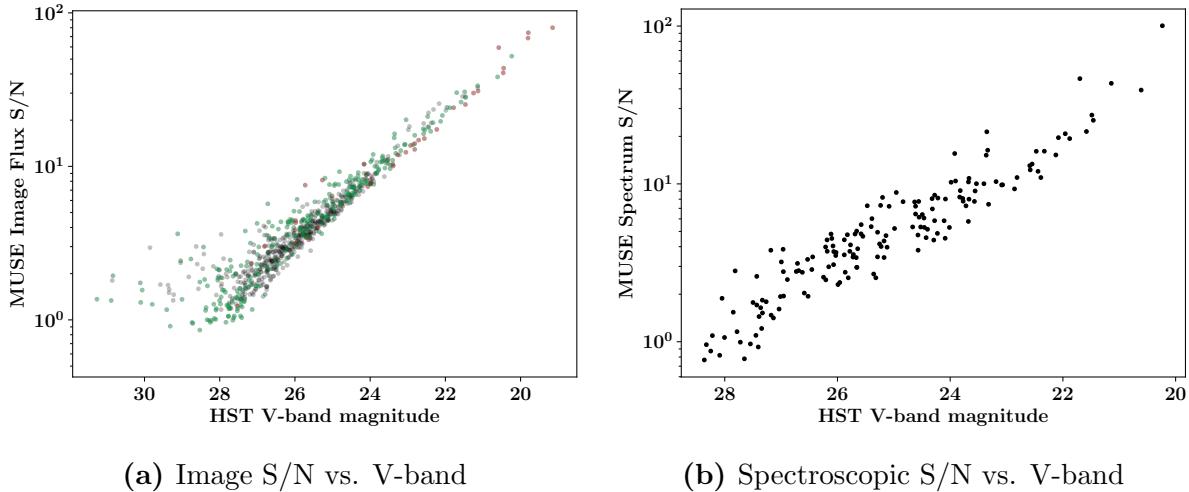


Figure 3: Signal-to-noise (S/N) versus V-band magnitude plots for the MUSE galaxy objects. The V-band magnitudes are from the HST catalogue. (a) S/N is from the collapsed MUSE cube image. The red points represent those with redshifts $z < 0.3$, and the green points are a chosen sample of 300 as defined by the SExtractor probability that they are not stars. (b) S/N used here is from the region $3700 \times (1+z)$ Å to $4500 \times (1+z)$ Å in the galaxy integrated spectra. Objects in this sample are those defined as the “green points” in Figure 3(a).

1. Gaussian doublet fitting

Through collapsing the entire galaxy into a single spectrum, one would find the [OII] doublet emission line to be the most prominent feature due to its large amplitude when compared to other spectral lines (Figure 5). This therefore allows it to be a useful tool for estimating the SFR, which in turn can be used as an indication of the evolutionary state of the galaxy (Maddox 2018).

For every doublet fitting performed, the region $3600 \times (1+z)$ Å to $3750 \times (1+z)$ Å was extracted to isolate the peaks from other absorption or emission features. The peaks themselves have rest wavelengths 3727.092 Å and 3729.875 Å (Abazajian et al. 2004). The following equation was applied as the model,

$$f_{\text{[OII]}}(x) = cx + \frac{I_1}{\sigma\sqrt{2\pi}} \exp -\frac{(x - L_1)^2}{2\sigma^2} + \frac{I_2}{\sigma\sqrt{2\pi}} \exp -\frac{(x - L_2)^2}{2\sigma^2}, \quad (2)$$

where x are the redshifted wavelengths, c is a constant, I_1 and I_2 are scaling factors, L_1 and L_2 are the [OII] peaks multiplied by

$(1+z)$, and $\sigma = \sqrt{\sigma_{\text{gal}}^2 + \sigma_{\text{inst}}^2}$ the quadrature sum of the fixed instrumental resolution σ_{inst} and the deconvolved width of the doublet σ_{gal} .

After a successful fitting, the following variables would be outputted: c , I_1 , I_2 , z , and σ_{gal} . The latter two would become the most useful in later research.

The redshifts along with the [OII] flux $f_{\text{[OII]}}(x)$ can be used to calculate the luminosity for the doublet, the latter can then be plotted against the redshifts (Figure 4). What can be found is a general agreement that the galaxies follow the luminosity-redshift relation, and thus the calculated parameters are sensible so they can be applied to more advanced analysis routines.

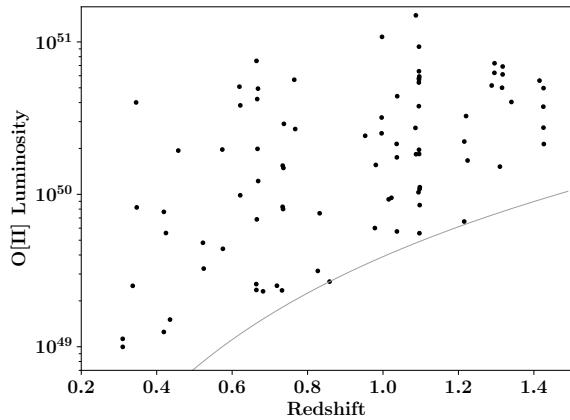


Figure 4: [OII] luminosity plotted against redshift. Variables used for each galaxy were obtained from the fitting of the galaxy-integrated spectra with a model (Equation 2). The line represents the lower-limit of the sample as it is using the galaxy with the smallest flux. All galaxies in the set are above this line and approximately follow the luminosity-redshift trend.

2. pPXF fitting

With the gas dynamics of galaxies represented through the width of the [OII] doublet, a comparison to the galactic stars would require the width of the stellar spectra features. To obtain this standard deviation value a Penalized Pixel-Fitting method (pPXF) can be applied, Cappellari (2017) describes the method and implementation in the form of a Python library, *pPXF*.

A reasonable initial test would be to pass an entire galaxy spectrum to the kinematics fitter and then examine the output to check if it seems reasonable. As well as a final model spectrum (Figure 5), pPXF also outputs its optimal variables and amongst which includes a velocity dispersion. However for these values and results to comprehensible, one must be aware of what initial parameters are passed to the fitting routine.

An estimate of the redshift can be obtained from the [OII] doublet modelling, the instrumental resolution can be found by fitting a single Gaussian profile over the peak

at $\sim 6300 \text{ \AA}$ in a spectrum of the sky noise, and different template sets can be passed to the routine.

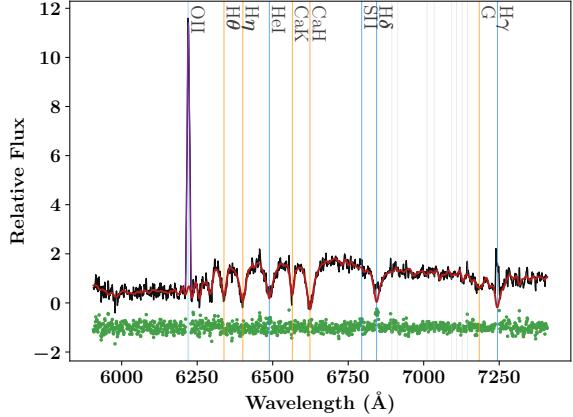


Figure 5: A pPXF fitting for a galaxy spectra. The galaxy data (black) and the best-fitting model (red) were normalised in the flux to avoid numerical errors during the fitting routine. The emission (blue lines), absorption (orange lines), and iron (grey lines) features have been labelled for easier recognition. The residuals between the data and model are provided by the green scatter points below the spectra. pPXF appears to produce good fits from manipulating various stellar templates.

As a synthetic test of pPXF, 10 galaxy spectra were each perturbed 300 times by sequentially adding random Gaussian noise. During each perturbation the spectra was passed to pPXF and all associated outputs were saved and recorded. This allows for the S/N to be plotted against the modulus difference between the optimal line widths and the perturbed ones (Figure 6). One finds that as the S/N decreases, the difference (or fractional error) should increase dramatically as there would not be enough resolution for emission or absorption features to be discerned.

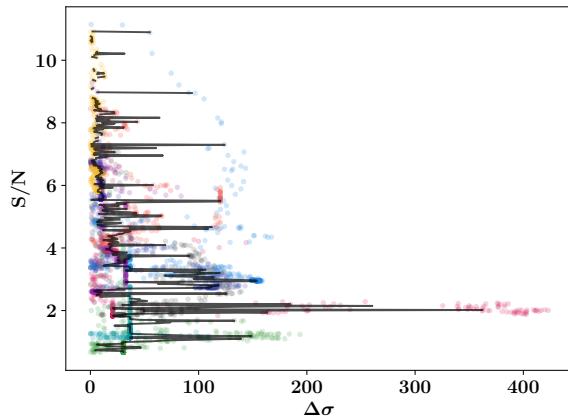


Figure 6: Signal-to-noise versus the modulus difference between the optimal pPXF velocity dispersion and the values from the perturbed models. As the S/N decreases, the difference increases manifold which implies that high S/N galaxies should be used as they would have smaller fractional uncertainty.

3. Voigt fitting

sadfasd

3. CONCLUSIONS

In conclusion, through extensive data and statistical analysis it can be said that the dynamics of stars and gas in galaxies are ... (?)

Acknowledgments

The author would like to thank Dr. M. Swinbank and Alfie Tiley for their continual help and support throughout the project period, without which, the project would have been experimentally grounded.

References

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A. et al. (2004), ‘Algorithms - Emission and absorption line fitting - SDSS DR7’, <http://classic.sdss.org/dr7/>

- [algorithms/specclinefits.html](#). [Online; accessed 26th December 2018].
- Bacon, R., Conseil, S., Mary, D. et al. (2017), ‘The MUSE Hubble Ultra Deep Field Survey’, *Astronomy and Astrophysics* **608**.
- Beckwith, S. V. W., Stiavelli, M., Koekecker, A. M. et al. (2006), ‘The Hubble Ultra Deep Field’, *The Astronomical Journal* **132**, 1729–1755.
- Bertin, E. and Arnouts, S. (1996), ‘SExtractor: Software for source extraction’, *Astronomy and Astrophysics Supplement* **117**, 393–404.
- Binney, J. (2016), *Astrophysics: A Very Short Introduction*, Vol. 1, Oxford Very Short Introductions.
- Cappellari, M. (2017), ‘Improving the full spectrum fitting method: accurate convolution with Gauss-Hermite functions’, *MNRAS* **466**, 798–811.
- Carroll, B. W. and Ostlie, D. A. (2007), *An Introduction to Modern Astrophysics*, 2nd edn, Pearson.
- Conselice, C. J., Wilkinson, A., Duncan, K. and Mortlock, A. (2016), ‘The Evolution of Galaxy Number Density at $z < 8$ and its Implications’, *The Astrophysical Journal* **830**(83), 17.
- Currie, M. J. et al. (2013), ‘Starlink Software in 2013’, *Astronomical Society of the Pacific Conference Series* **485**.
- Henault, F., Bacon, R., Bonneville, C. et al. (2003), ‘MUSE: a second-generation integral-field spectrograph for the VLT’, *Proc.SPIE* **4841**.
- Maddox, N. (2018), ‘[O II] as a proxy for star formation in AGN host galaxies: beware of extended emission line regions’, *Monthly Notices of the Royal Astronomical Society* **480**, 5023–5210.
- McCoy, J. F. (2012), *Space Sciences*, Vol. 2, 2nd edn, Macmillan Reference USA.
- Moore, P. and Rees, R. (2011), *Patrick Moore’s Data Book of Astronomy*, Cambridge University Press, pp. 357 – 362.
- NASA, ESA, Beckwith, S. and HUDF Team (2018), ‘Hubble Ultra Deep Field’, <https://svs.gsfc.nasa.gov/30946>. [Online; accessed 5th November 2018].

Williams, R. E., Blacker, B., Dickinson, M.
et al. (1996), ‘The Hubble Deep Field: Ob-
servations, Data Reduction, and Galaxy
Photometry’, The Astronomical Journal
112(4).