

# pyFIT3D and pyPipe3D - The new version of the Integral Field Spectroscopy data analysis pipeline

Eduardo A. D. Lacerda<sup>c</sup>, S. F. Sánchez<sup>1</sup>, A. Mejía-Narváez<sup>1</sup>, A. Camps-Fariña<sup>1</sup>, C. Espinosa-Ponce<sup>1</sup>, J. K. Barrera-Ballesteros<sup>1</sup>, H. Ibarra-Medel<sup>b</sup>, A. Z. Lugo-Aranda<sup>1</sup>

<sup>a</sup>*Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 70-264, CDMX 04510, México*

<sup>b</sup>*University of Illinois Urbana-Champaign, Department of Astronomy, 1002 W Green St, Urbana, Illinois, 61801, USA*

<sup>c</sup>*Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 70-264, CDMX 04510, México*

## Abstract

We present a new version of the FIT3D and Pipe3D codes, two packages to derive properties of the stellar populations and the ionized emission lines from optical spectroscopy and integral field spectroscopy data respectively. The new codes have been fully transcribed to Python from the original Perl and C versions, modifying the algorithms when needed to make use of the unique capabilities of this language with the main goals of (1) respecting as much as possible the original philosophy of the algorithms, (2) maintaining a full compatibility with the original version in terms of the format of the required input and produced output files, and (3) improving the efficiency and accuracy of the algorithms, and solving known (and newly discovered) bugs. The complete package is freely distributed, with an available repository online. pyFIT3D and pyPipe3D are fully tested with data of the most recent IFS data surveys and compilations (e.g. CALIFA, MaNGA, SAMI and AMUSING++), and confronted with simulations. We describe here the code, its new implementation, its accuracy in recovering the parameters based on simulations, and a showcase of its implementation on a particular dataset.

**Keywords:** galaxies:ISM — techniques: spectroscopy

## 1. Introduction

The large volume of observed data available nowadays in every area of science demands high speed, precise and accurate analysis algorithms. Due to the availability of different data, the analysis packages are increasing their capability and coverage of usability cases. In the exploration of the nearby Universe ( $z < 0.1$ ), the scenario could not be different. Over the last couple of decades, the datasets evolve from a few number of surveys sampling a single value and/or a single spectrum for each galaxy, to an exploding amount of surveys with observations of millions of spectra covering the entire field-of-view of a huge amount of galaxies (2dFRGS: Folkes et al. 1999; SDSS: York et al. 2000; MaNGA: Blanton et al. 2017; CALIFA, Sánchez et al. 2012; SAMI: Croom et al. 2012; AMUSING: Galbany et al. 2016). As a consequence, the new datasets require dedicated pipelines for their data reduction and analysis (e.g. LZIFU, Ho et al. 2016; MaNGA DAP, Westfall et al. 2019; GIST, Bittner et al. 2019).

Pipe3D (Sánchez et al., 2016b, hereafter S16b) is one of these tools, a dedicated pipeline to extract the properties of the stellar populations and emission lines in the optical spectra of Integral Field Spectroscopic data of galaxies. This pipeline uses as basic fitting algorithms the ones provided by FIT3D (Sánchez et al., 2016c, hereafter S16a), a package that can explore the same properties for individual spectra, row-stacked multi-object spectra (or classical long-slit ones), or IFS datacubes. Both tools were coded in PERL (with some rou-

tines coded in C), making use of the numerical and fitting algorithms included in the Perl Data Language (PDL GLAZEBROOK AND ECONOMOU, 1997). Pipe3D has been broadly used in the analysis of IFS data from individual galaxies and large datasets of different IFS galaxy surveys including CALIFA (e.g. Sánchez-Menguiano et al., 2016), MaNGA (Sánchez et al., 2018), MUSE (e.g. López-Cobá et al., 2020), and SAMI (e.g. Sánchez et al., 2019). Its capabilities to recover the properties of the stellar populations and emission lines have been contrasted with hydrodynamical simulations (e.g. Guidi et al., 2018; Ibarra-Medel et al., 2019; van de Sande et al., 2019), and fully compared with the values recovered using other similar tools (e.g. Belfiore et al., 2019; Sánchez et al., 2019).

Despite of its well proved capabilities there are some drawbacks with the current implementation of the code. The most important one is that it is deeply attached to a coding language and a numerical package that are used only by a small fraction of the astronomical community. Furthermore, it uses PGPLOT<sup>1</sup>, a graphics library that is no longer officially supported, and again, with a small number of users nowadays. In addition, new improvements in basic algorithms that frequently appear for other languages (like Python) and numerical packages (like NUMPY or SCIPY) cannot be implemented. For all those reasons the code is difficult to be updated, upgraded and maintained. It is difficult to be installed in new versions of operative systems, and therefore, to be distributed.

<sup>1</sup><https://sites.astro.caltech.edu/~tjp/pgplot/>

In order to solve all those problems, update the code, and maintain it as competitive as possible, we embarked in the transcription of the code to PYTHON, the coding language that is nowadays the one with largest fraction of users among the astronomical community. We present in here the new code, describing its conceptual and coding philosophy, summarizing the main algorithms, and showing its capabilities in the analysis of the stellar populations and emission lines for both individual spectra and IFS data. The structure of the article is as follows: (i) Section 2 describes the dataset adopted as a showcase of the use of the code; (ii) In Sec. 3 we describe the philosophy of the new code; (iii) Sec. 4 comprises the description of the algorithms, with 4.1 describing `pyFIT3D`, including the updates performed (Sec. 4.1, a description of procedures to recover the non-linear parameters (velocity, velocity dispersion and dust attenuation, Sec. 4.1.1) and linear parameters (i.e., decomposition of the stellar population in the adopted single-stellar population library, Sec. 4.1.3), and 4.2 describing the different steps included in `pyPIPE3D` to analyze an individual IFS datacube; (iv) the accuracy in the recovery of the properties of the new code is presented in Sec. 5, first contrasted against simulations for both the stellar population and emission line properties (Sec. 5.1 and Sec. 5.2, respectively), and then against real data (Sec. 6); (v) Finally, the conclusions of this study are presented in Sec. 7.

## 2. Data

The data adopted in this study as a showcase of the code is obtained from the extended Calar Alto Legacy Integral Field Area survey sample (eCALIFA, Sánchez et al. 2016a; Galbany et al. 2018). The survey was built with observations from the 3.5m telescope at the Calar Alto observatory, using PPAK Integral Field Unit (Kelz et al., 2006) of the Potsdam Multi-Aperture Spectrograph (Roth et al., 2005, PMAS), with a covering factor of 60% of a  $74'' \times 64''$  field-of-view (FoV) with 331 fibers of  $2.7''$ . A complete coverage of the FoV is achieved with a three position dithering scheme resulting in a spatial resolution (characterized by the point-spread function, PSF) with a full-width at half-maximum (FWHM) of  $\sim 2.5''$ . The sample of galaxies covered by eCALIFA, corresponding to objects within a narrow range of redshifts around  $z \sim 0.015$ , was primarily selected by diameter (Walcher et al., 2014) in order to fit the optical extension of the galaxies within the FoV of the instrument. We use the V500 observation setup ( $3745 - 7500\text{\AA}$ ,  $\lambda/\Delta\lambda \sim 850$ ) for the dataset analyzed during this study, which guaranties the simultaneous covering of (i) the most relevant spectral features of the stellar populations in the optical regimes and (ii) the more relevant emission lines from `[OII] $\lambda$ 3727` to `[SII] $\lambda$ 6717,31` to explore the properties of the ionized gas. The observing strategy guaranties to sufficient signal-to-noise through the optical extension of the galaxies, making this dataset ideal for the purposes of testing the new version of the code. It is worth noticing that this code, like its predecessor (Pipe3D), has been already applied to other similar datasets, like MaNGA (Sánchez et al., in prep.) or data of IFS data of better spectral resolution (e.g., MUSE).

## 3. Philosophy of the new code

Contrary to the previous version of the code, that was mostly a single-coder package, the new version was developed by a group of people. This requires to adopt an inclusive and easy-to-write programming language. Following the massive growth of python users (e.g. Van Rossum and Drake Jr, 1995) in astronomy, hence the availability of modern, well documented, helpers and packages of analysis, we rewrite the code package entirely from perl (Wall et al., 2000) to python 3 (Van Rossum and Drake, 2009, 3.6 or higher). Furthermore, to rewrite Pipe3D with such an inclusive programming language will encourage more people to participate in the development of the code at a lower level. At the end, we choose not to rename the pipeline and the spectral fitting tool, but to append a prefix “*py*” to original names.

One of the most important aspects of this code refactoring project is the new adopted programming philosophy: to provide modular and reusable standalone pieces of code (such as classes and modules) making easy the production of analysis, data-exploration and interactive<sup>2</sup> scripts. Moreover, the new code is constructed based on well documented public libraries to facilitate users/developers to understand how the analysis is made. In addition to some standard python 3 library modules, we utilize numerical methods from NumPy (Oliphant, 2006) and SciPy (Virtanen et al., 2020), the FITS reading package of astropy (astropy.io.fits; Price-Whelan et al., 2018) together with some plot functionalities from matplotlib (Hunter, 2007) and seaborn (Waskom et al., 2017). Finally, the new package carries a complete code-based documentation and an online repository<sup>3</sup> including handy tools to help the users dealing with the output files from the analysis and interactive scripts with various types of procedures taking advantage of the code reusability.

## 4. New code implementation

As indicated before `pyFIT3D` and `pyPipe3D` are based on `FIT3D/Pipe3D`. We attain the detailed description along the following subsections for the parts that are needed to the understanding of the article and also for those which receive some update. We present the main changes revisiting some details of the spectral fitting code and introducing the updates to the code. Finally, we close this Section with the description of `pyPipe3D` procedure for the analysis of an IFS datacube.

### 4.1. `pyFIT3D`: The spectral fitting code update

The spectra of a galaxy (or a region of a galaxy) are the result of the emitting light from stars and the ionized gas, kinematically shifted due to its dynamical stage and dust attenuated. The idea behind `pyFIT3D` is to decouple the stellar continuum and the ionized emission spectrum in the observed spectral energy density (SED). The fitting process works disentangling the

<sup>2</sup>Such as jupyter-notebooks: <https://jupyter.org/>.

<sup>3</sup><http://gitlab.com/pipe3d>

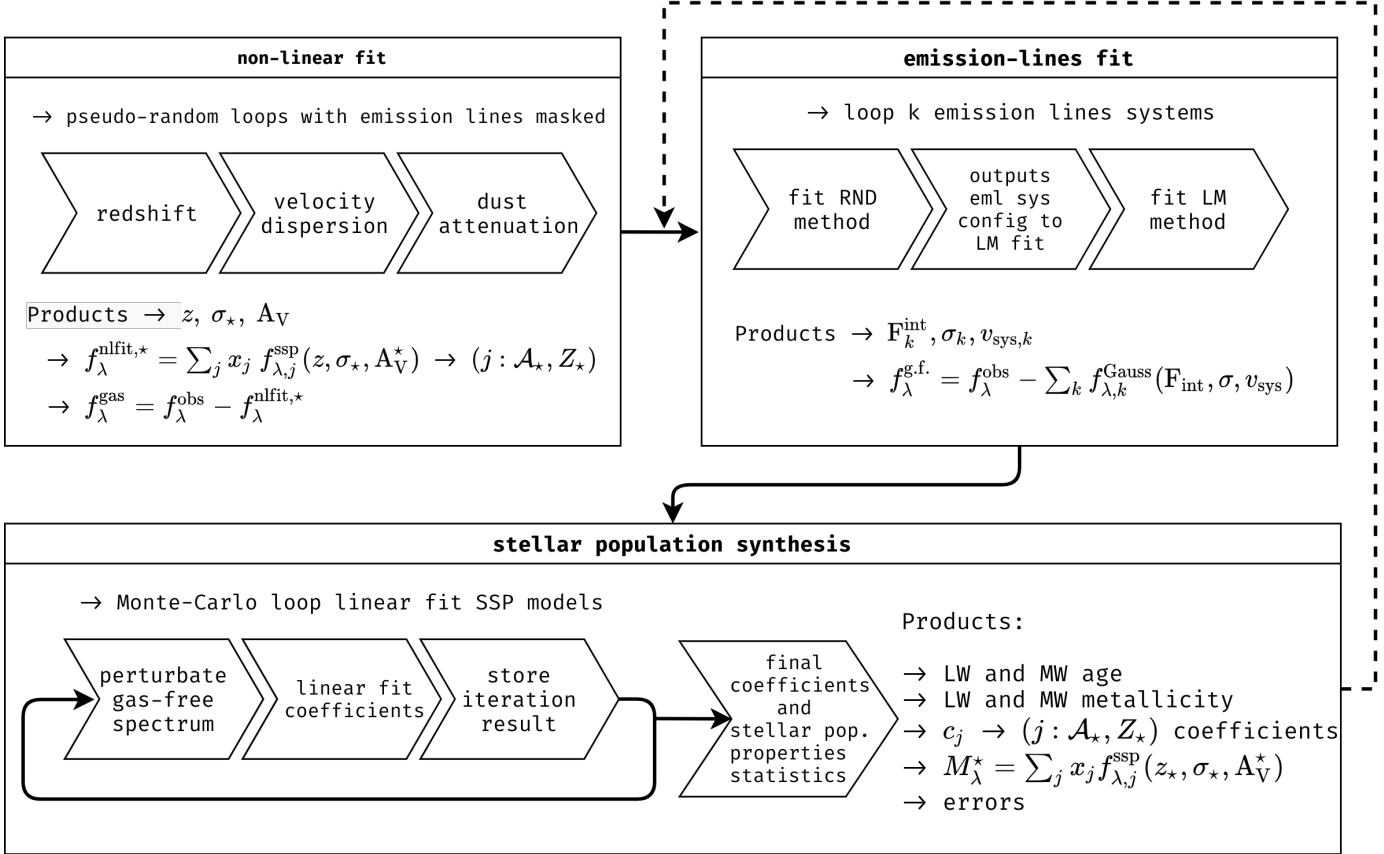


Figure 1: Flow-chart showing the three-step fitting process included in pyFIT3D, together with the output products at every step: non-linear fit (top-left panel), where the systemic velocity/redshift ( $z_*$ ), line-of-sight velocity dispersion ( $\sigma_*$ ) and the dust attenuation ( $A_V^*$ ) are estimated; emission lines fit (top-right panel), where it is created a model for the ionized gas emission lines spectrum; and finally the stellar population synthesis (bottom panel), where the gas-free resultant spectrum is then modeled creating the light population vector ( $x_j$ , i.e., the coefficients of the decomposition of the stellar spectrum in a base of synthetic stellar templates). At the end, the light-weighted (LW) and mass-weighted (MW) age and metallicity are estimated, based on this decomposition, together with associated errors.

stellar and the interstellar medium spectra without a direct assumption of the star-formation and chemical enrichment histories (SFH and ChEH). However, we assume that the individual spectrum comprises an unresolved stellar population. Thus, the physical aperture from which the spectra is obtained is large enough to encircle  $\sim 10^4$  stars, a number sufficient to fully sample an entire Mass Function (without stochastic problems Cerviño and Luridiana, 2004). This way, the stellar population can be modeled by a linear combination of single-stellar populations (SSPs), each one corresponding to the different star-formation burst along the SFH. In comparison to the stellar continuum, we further assume that the nebular continuum is negligible throughout the considered aperture. The flowchart presented in Figure 1 schematize the three-steps process of the spectral analysis described along this Section:

- Non-linear fit step (*upper-left panel*), where it is derived the redshift (systemic velocity), the line-of-sight velocity dispersion (LOSVD) and the dust attenuation ( $z_*$ ,  $\sigma_*$  and  $A_V^*$ ).
- Emission lines fit step (*upper-right panel*), where it is derived the main properties of a pre-defined set of emission

lines.

- Stellar population synthesis step (*bottom panel*), where the stellar spectrum is decomposed in a set of SSPs, deriving its main properties (e.g., light-weighted ages and metallicities).

This sequence is essentially the same as the one adopted by the previous version of the code.

There are other two additional analysis that can be performed by FIT3D (and are an integral part of Pipe3D) which are not included originally in the formal loop of analysis, although they complement it:

- Moment analysis: a non-parametric estimation of emission lines properties.
- Stellar-indices analysis: where it is measured a set of stellar indices, such as the Lick/IDS index system (e.g. Burstein et al., 1984; Worthey, 1994), and D4000 (Bruzual A., 1983).

Details on the procedures and their modifications are given over next subsections.

#### 4.1.1. Step one: non-linear fit

As indicated before, the first group of analysis performed by pyFIT3D is the determination of the stellar spectrum kinematic parameters ( $z_\star$  and  $\sigma_\star$ ) and the dust attenuation ( $A_V^\star$ ). During this step, the parameters are explored once at a time following the sequence shown in upper-left diagram of Figure 1. For each explored parameter pyFIT3D covers a range of values constraint between an user defined *minimum* and *maximum* values (defining an explored *interval*), following a random sampling with an average *step* (defined by the user too). An *input guess* value is also required, which use will be described later on. Optionally, setting up the *step* to zero makes the program to fix the value of this parameter to the *input guess* along the analysis, i.e. the parameter is not determined by pyFIT3D. The kinematic parameters receive a second round of exploration as a refinement of the first one. In this turn, the *minimum*, *maximum* and *step* of exploration are automatically updated for a fine search around the best value found during the first round of the parameter determination.

To estimate the best fitted parameter, for each set of values ( $z_\star, \sigma_\star$  and  $A_V^\star$ ) the program computes a model for the stellar spectrum (following the inversion procedure described below). For doing so, the adopted single/synthetic stellar population (SSP) library is configured to the candidate observed frame (i.e., redshifted and broadened with the explored values of the kinematic properties and dust attenuated based on the assumed extinction law) with a predefined set of emission lines (in the observed-frame) masked together with the strongest night-sky emission lines (in the rest-frame). For the broadening of the spectra due to the velocity dispersion it is assumed a Gaussian profile for the Line-of-sight velocity distribution (LOSVD). We acknowledge that this is just a first order approximation and that effects like the asymmetry of the profiles cannot be recovered using this approximation (e.g. Cappellari et al., 2011). However, for the typical signal-to-noise (S/N) of the current IFS surveys this modeling is sufficient and preferred than a more accurate, but more demanding, modeling of the stellar populations. In addition, we also include the option to input an instrumental dispersion (in Å) to be considered during the broadening of the spectra.

pyFIT3D includes the option to use two SSPs libraries, one for this step and another for the stellar population synthesis at the final part of the program (Section 4.1.3). We recommend the use of a simple set of SSPs for the non-linear fit step for two reasons: (1) speed-up the process (that is very time-consuming) and (2) avoid known degeneracies between the parameters (e.g. stellar metallicity and  $\sigma_\star$ ). Then, the coefficients ( $x_j$ ) of the linear combination of the SSP models which better reproduces the stellar continuum are derived through a recursive weighted least-squares (WLS) inverse matrix method,

$$f_\lambda^{\text{nlfit},\star} = \sum_j^{x_j > 0} x_j f_{\lambda,j}^{\text{ssp}}(z_\star, \sigma_\star, A_V^\star) \rightarrow (j : \mathcal{A}_\star, Z_\star). \quad (1)$$

Where  $f_\lambda^{\text{nlfit},\star}$  is the modeled spectrum,  $f_{\lambda,j}^{\text{ssp}}(z_\star, \sigma_\star, A_V^\star)$  is the shifted, broadened and dust attenuated individual SSP template,

where  $j$  runs through the range of ages ( $\mathcal{A}_\star$ ) and metallicities ( $Z_\star$ ) covered by the considered library. The recursive procedure applied iterates the WLS matrix inversion, rejecting those SSP templates within the library that produce negative coefficients, shrinking the SSP library to include only those ones for which a positive coefficient is derived. The iteration continues until there are no negative coefficients. The adopted merit function is the reduced  $\chi^2$ , calculated as

$$\chi^2 = \frac{1}{n-m} \sum_\lambda^n \left( \frac{f_\lambda^{\text{obs}} - f_\lambda^{\text{nlfit},\star}}{\sigma_\lambda} \right)^2, \quad (2)$$

where  $n$  is the number of pixels of the observed spectrum,  $m$  is the number of templates present in the adopted SSP templates library and  $\sigma_\lambda$  is the error associated with the observed spectrum. In addition to a global input wavelength interval for the analysis performed by pyFIT3D, we include the option to configure an exclusive interval for the derivation of the kinematic parameters ( $z_\star$  and  $\sigma_\star$ ). Although the lower wavelength interval for the determination of  $A_V^\star$  would speed-up the process, it is preferred a large baseline of wavelength to provide with the best estimation of this parameter. At the end, the best-fitted value for each parameter corresponds to that one for which the stellar population model produces the lower value of  $\chi^2$ . Since all three properties are necessary to model the stellar spectrum, if a property is yet to be determined, it adopts the *input guess* value.

Figure 2 illustrates this entire process for the central spectrum (5") of the showcase dataset (NGC 5947, observed by CALIFA). Details on the adopted parameters, masked wavelength intervals and configuration files for the current fitting procedure are given in Appendix Appendix C. The final result depends on the particular parameters included in this configuration files, and in particular on the adopted SSP library (see S16a and Cid Fernandes et al., 2014, for a discussion on the topic). For this particular example, we adopted a sub-set of the gsd156 stellar population library comprising just 12 templates (four ages and three metallicities). The gsd156 library (Cid Fernandes et al., 2013), comprises 156 SSP templates, that sample 39 ages (1 Myr to 14 Gyr, on an almost logarithmic scale), and 4 different metallicities ( $Z/Z_\odot=0.2, 0.4, 1, \text{ and } 1.5$ ), adopting the Salpeter IMF (Salpeter, 1955). We must clarify that the code can run using any suitable SSP library, once transformed to the required input format (described in Appendix Appendix A). All attempts of the modeled spectra (i.e., all the realizations of Eq. 1), residuals and the best model are shown along the first column following the sequence of determination from top to bottom. In each panel, vertical shaded regions represent the masked wavelengths during the procedure. The second column shows the merit curve for each explored parameter.

We highlight here the importance of a good choice of the *interval* explored in the determination of each parameter. The reason is that there may be local minima present in the parameter space exploration, as can be seen in the *upper-right panel* of Figure 2. For this, a selection of the *interval* of parameters based on the known properties of the observed object (e.g., redshift) is required, improving the velocity and reliability of the

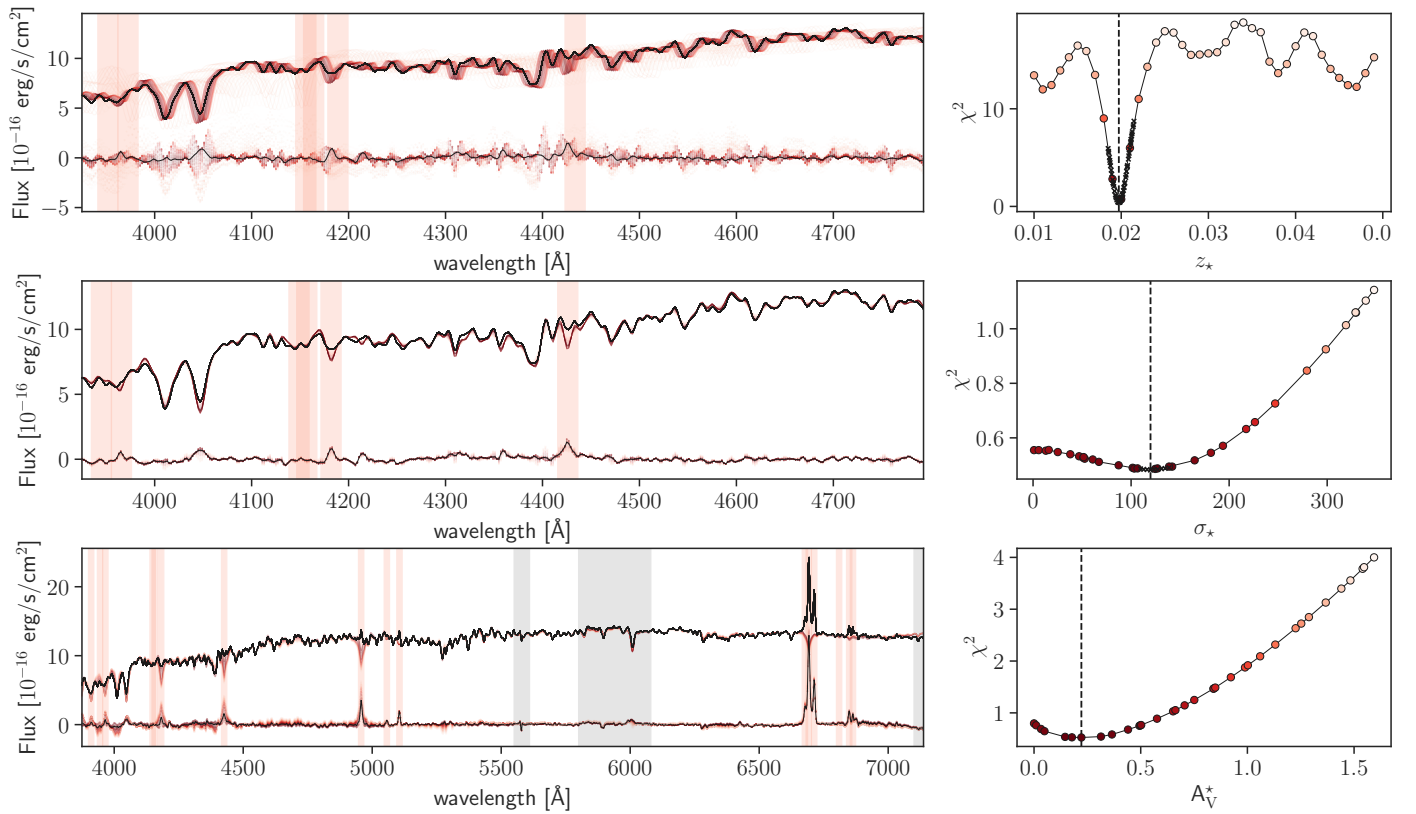


Figure 2: Showcase of the determination of the non-linear parameters for the central (5'' aperture) spectrum of galaxy NGC 5947: Redshift ( $z_*$ , upper-panels), line-of-sight velocity dispersion ( $\sigma_*$ , middle-panels) and dust attenuation ( $A_V^*$ , lower-panels). The rows are ordered from top to bottom in the sequence that the three parameters are determined (as appears in Figure 1). At each row, the left panel shows the observed spectrum (black line) and the different models evaluated during the parameter exploration with associated residuals (colors distributed from light- to dark-red following the range of explored values). The vertical shaded regions indicate the wavelength ranges masked during the process. The right panel shows the merit curve ( $\chi^2$ ) as a function of each parameter. The colored dots follow the parameter value during the exploration of the first loop and, the black crosses, the second step exploration, a fine sweep around the best value found in the first loop. The best value found for each parameter is then indicated as a vertical dashed black line.

program. We will discuss that in Section 4.2.

#### 4.1.2. Step two: Emission lines fit

After the derivation of the best kinematic and the dust attenuation parameters the program produces a first model for the stellar spectrum based on the simple SSPs template (i.e., the one adopted for the non-linear fit step). This stellar spectrum model is then subtracted to the original one to produce a spectrum with just the information of the ionized gas emission lines (plus noise and residuals). The new process of the emission lines fit implemented in pyFIT3D is made through a two-round procedure summarized at the upper-right diagram of Figure 1. Which *systems*<sup>4</sup> the program will fit depends on the configuration made by the user<sup>5</sup>.

For the fit of the emission lines, we have implemented now two different methods that we call RND (for random) and LM

<sup>4</sup>We call *system* to a set of emission lines fitted at same time, i.e., included by the same configuration file.

<sup>5</sup>More details and examples on the configuration files are included in the pyPipe3D webpage

(for Levenberg–Marquardt), both modeling the emission lines as Gaussian profiles:

- **The RND method** mimics the one already implemented in the previous version of the code. The method is constructed based on a Monte-Carlo (MC) loop with a  $\chi^2$  minimization scheme where all emission line models are created in a pseudo-random search of the free parameters. The process of the exploration of these parameters works in a similar way as the one performed by the non-linear fit step (i.e., the range of allowed values is fully covered using a pseudo-random exploration). Again, this is important to avoid local minima inherent to other fitting/minimization procedures (that may be faster in principal). At the end of each MC loop it is evaluated the best fitted  $\chi^2$  and the intervals of exploration are narrowed around the best fitted parameters (following a similar process to a Markov chain MC inference). Once the program provides with a first estimation of the parameters that define each emission line, the algorithm performs a second fit with the **LM method**.
- **The LM method** is the implementation of the *Levenberg–*

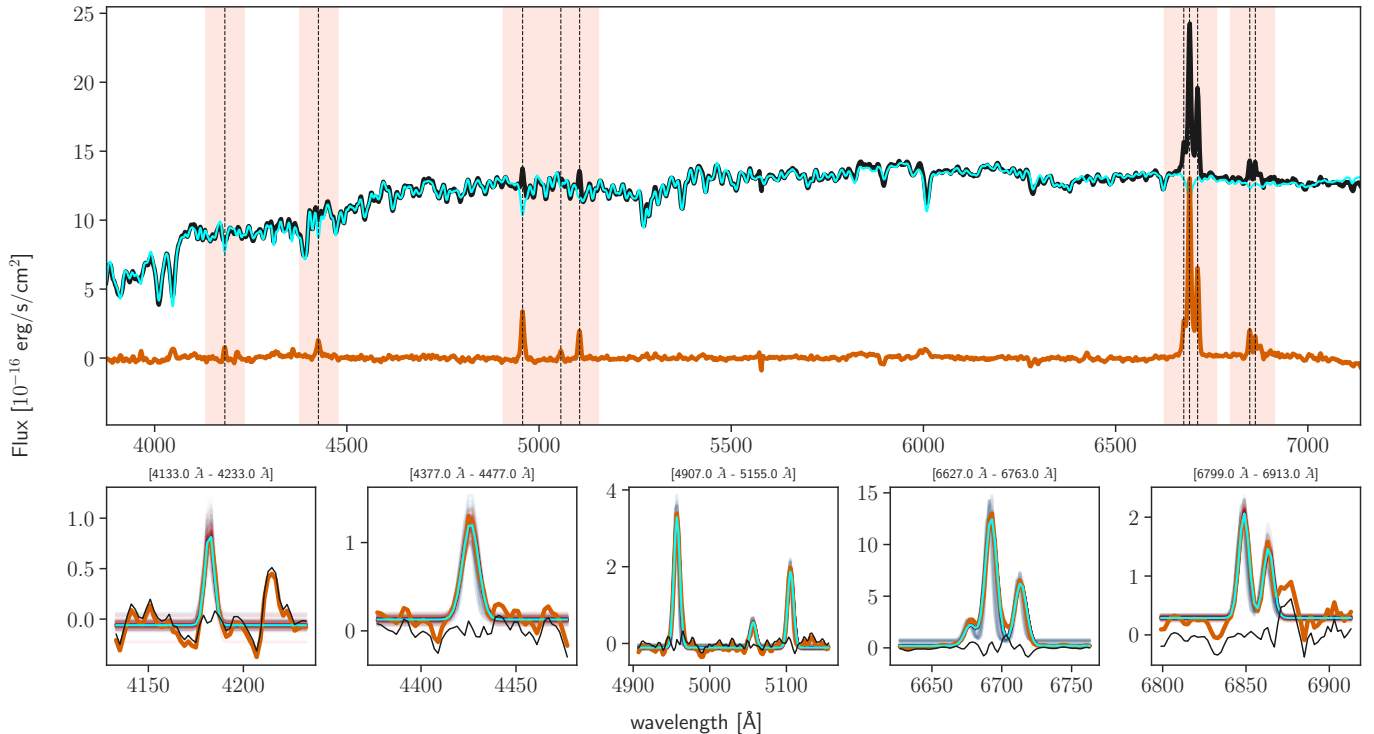


Figure 3: Example of the exploration of the properties of the ionized gas emission lines for the central spectrum of NGC 5947. Top panel shows the observed spectrum (black) together with the best fitted model of the stellar population spectrum provided by the 1st non-linear step of the procedure (cyan). The residual spectrum, once subtracted the best model to the original spectrum, is shown in red. This spectrum comprises the ionized gas emission lines to be fitted (together with the residual of the stellar population fitting plus noise). Light-red shaded vertical regions highlights the intervals selected for fitting each emission lines *system*. Bottom panels shows a zoom on those wavelength intervals. In each panel the red line corresponds to the same residual spectrum shown in the top panel. This time, the cyan corresponds to the best fitted model for the set of emission lines in each wavelength regime and the residual of the fit is shown as a thin black line. The shaded area highlights all the models explored by the MC method implemented in the fitting procedure. `pyFIT3D` derives the intensity, the velocity and the velocity dispersion of all implicated emission lines.

*Marquardt* minimization, which has good precision and is very fast. However, this method, like many other minimization algorithms, may present issues with the accuracy due to the presence of local minima (e.g., right panels of Figure 2). Moreover, if the intervals for the exploration of each parameter are not well defined, the method could diverge. As indicated before, the RND method has a very robust accuracy. However, it is slower and its precision is affected on low S/N regime. The combination of both methods produce optimal results in terms of accuracy and precision. This is the reason why we include the second round of analysis in the emission lines fit step of `pyFIT3D`. At the end, the final values for the explored parameters are derived using the LM method and the associated errors are estimated based on a MC iteration.

The Figure 3 exemplifies this process for our example dataset. In this example we show the fit of five *systems* encompassing ten emission lines (bottom row, from left to right): H $\delta$ ; H $\gamma$ ; [O III] $\lambda\lambda$ 4958,5007 + H $\beta$ ; [N II] $\lambda\lambda$ 6548,6583 + H $\alpha$ ; [S II] $\lambda\lambda$ 6731,6716) through the **RND+LM method**. The set of models for different emission lines are combined to derive an ionized gas emission spectrum ( $\sum_k f_{\lambda,k}^{\text{Gauss}}$ ). This is then subtracted from the observed one, producing a gas-free spectrum (plus residuals and noise), which will be used as the input spec-

trum for the stellar population synthesis.

#### 4.1.3. Step three: Stellar population analysis

The last step of the fitting procedure included in `pyFIT3D` comprises the synthesis of the stellar population. This process works by solving the linear decomposition of the gas-free spectrum ( $f_{\lambda}^{\text{g.f.}}$ ) between the different SSP templates of the adopted library. The decomposition is done using a MC method over perturbed realizations of the input spectrum. First, all SSP templates are shifted to the observed frame and dust attenuated using the kinematics parameters and  $A_V^*$  derived in the first step (Section 4.1.1). Then, the average values of the set of  $x_j$  parameters derived for each MC iteration are recovered as the final weights/coefficients of the stellar population model, with the standard deviation being adopted as the  $1\sigma$  error of each parameter. Figure 4 outlines this part of the procedure for our adopted example dataset. For this particular example we use the full `gsd156` stellar population library, that has been extensively used in many different studies (e.g. Pérez et al., 2013; González Delgado et al., 2014; Ibarra-Medel et al., 2016; Sánchez et al., 2018, 2021). Once again, we must remind that the code can run using any SSP library with the proper format.

Let  $F_{\lambda}^i$  be the perturbed spectrum for the  $i^{\text{th}}$  MC iteration,  $R_{\lambda}^i$  the noise factor, corresponding to a random value following a

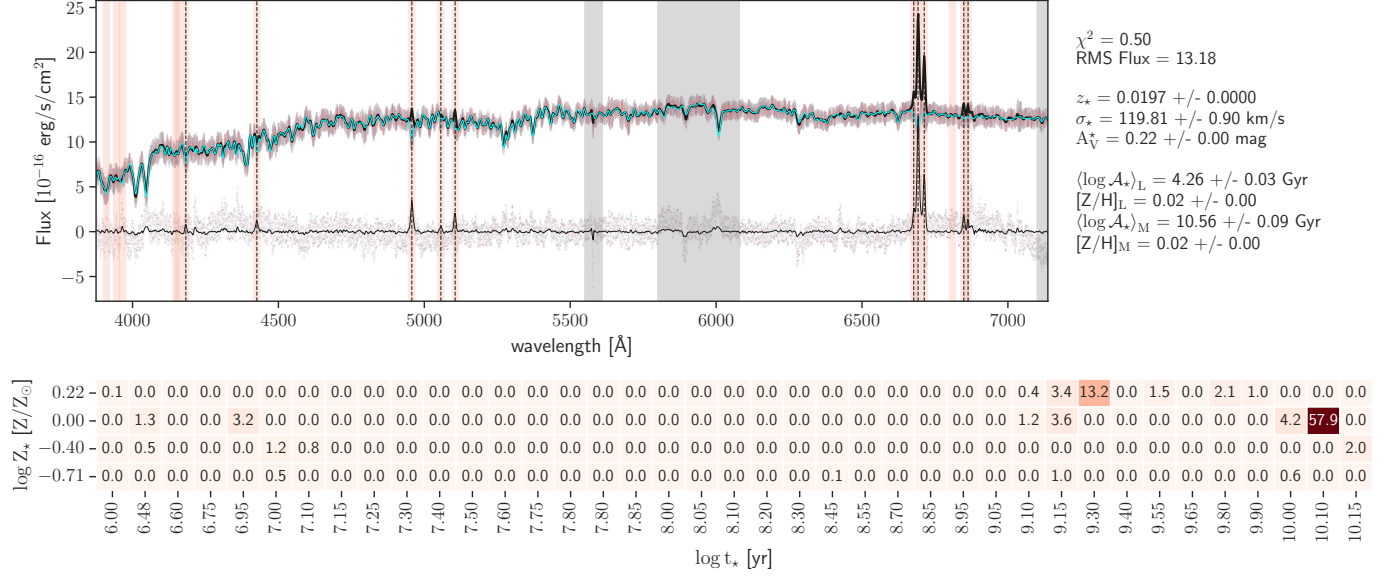


Figure 4: Example of the stellar population synthesis procedure for the central spectrum of NGC 5947. The upper panel shows the spectral fit (left) together with the values of the main properties derived by the synthesis (right). The observed spectrum (dark-thick line) is overplotted with the best model for the stellar spectrum (cyan line) with the final residual spectrum drawn as a black-thin line at the bottom. All the  $F_\lambda^k$  MC realizations are represented by the hatched gray area around the observed spectrum. The noise range (as in Eq. 3) and the residual of the k-realizations have been amplified by a factor of five to be better appreciated. The vertical shaded regions correspond to the masked areas due to intrinsic emission lines (red) and telluric ones (gray). Vertical dashed lines highlight the fitted emission lines (represented in upper-right panel of Figure 1 and in Figure 3). The bottom panel shows the fraction of light in the selected wavelength range ( $5500 \pm 45 \text{ \AA}$ ) that corresponds to each SSP within the template (i.e., the coefficients of the stellar decomposition), color coded in red.

(-1, 1) clipped Gaussian distribution, and  $\sigma_\lambda$  being the noise spectrum:

$$F_\lambda^i = f_\lambda^{g,f} + R_\lambda^i \sigma_\lambda. \quad (3)$$

The left panel of Figure 4 shows the set of  $F_\lambda^i$  (gray lines) with the noise factor ( $R_\lambda^i \sigma_\lambda$ ). In order to visualize better the differences between the  $i$  iterations the noise has been multiplied by 5 (this factor was introduced just for visualization purposes, it is not included as such in the code). The reduced  $\chi^2$  is calculated for each MC iteration as:

$$\chi_i^2 = \frac{1}{n-m} \sum_\lambda \left( \frac{F_\lambda^i - M_{\star,\lambda}^i}{\sigma_\lambda} \right)^2, \quad (4)$$

where  $n$  is the number of pixels in the observed spectrum,  $m$  is the number of SSP models and  $M_{\star,\lambda}^i$  is the modeled spectrum at the  $i^{\text{th}}$  MC realization,

$$M_{\star,\lambda}^i = \sum_j^{x_j > 0} x_j^i f_{\lambda,j}^{\text{SSP}}(z_*, \sigma_*, A_V^*), \quad (5)$$

and  $x_j^i$  is the coefficient of the  $j^{\text{th}}$  template of the adopted SSP library at the  $i^{\text{th}}$  MC realization. The final modeled stellar spectrum and associated uncertainties are given by the mean,

$$M_{\star,\lambda} = \text{mean}(M_{\star,\lambda}^{1..i}), \quad (6)$$

and standard deviation,

$$\sigma(M_{\star,\lambda}) = \text{stddev}(M_{\star,\lambda}^{1..i}), \quad (7)$$

of the individual  $i$  iterations, respectively. Finally, the coefficients of the best fitted multi-SSP model are given by the mean,

$$x_j(\mathcal{A}_*, Z_*) = \text{mean}[x_j^{1..i}(\mathcal{A}_*, Z_*)], \quad (8)$$

and their uncertainties by their standard deviation,

$$\sigma[x_j(\mathcal{A}_*, Z_*)] = \text{stddev}[x_j^{1..i}(\mathcal{A}_*, Z_*)]. \quad (9)$$

We should note that these final coefficients provided by the algorithm correspond to the fraction of light that each SSP template contributes to the observed spectrum at a certain wavelength,

$$c_j = \frac{x_j}{f_{\lambda_{\text{norm}},j}^{\text{SSP}}}. \quad (10)$$

For the optical wavelength range  $\lambda_{\text{norm}}$  is usually selected to be  $5500 \text{ \AA}$ .

Once we derive the coefficients of the decomposition of the stellar population in the SSP templates, it is possible to derive different average properties. In particular, pyFIT3D estimates the light-weighted (LW) stellar age,

$$\langle \log \mathcal{A}_* \rangle_L = \sum_j^m c_j \log \mathcal{A}_{\star,j}, \quad (11)$$

and metallicity,

$$[Z/H]_L = \sum_j^m c_j \log Z_{\star,j}. \quad (12)$$

The light-weighted averages highlight the contribution of the young stellar populations, i.e., recent star forming events, since

they contribute more to the light relative to their old counterparts. On the other hand, mass averaged values are less sensitive to the presence of young stellar populations, highlighting the contribution of the old stellar ones. Mass-weighted values are also calculated defining the mass fraction contributed by each SSP model as

$$\mu_j = c_j \Upsilon_{\lambda_{\text{norm}},j}, \quad (13)$$

where  $\Upsilon_{\lambda_{\text{norm}},j}$  is the stellar mass-to-light (i.e. M/L) associated with the  $j^{\text{th}}$  SSP model. So, the mass-weighted (MW) mean stellar age and metallicity are calculated as

$$\langle \log \mathcal{A}_\star \rangle_M = \frac{\sum_j^m \mu_j \log \mathcal{A}_{\star,j}}{\sum_j^m \mu_j} \quad (14)$$

and

$$[Z/H]_M = \frac{\sum_j^m \mu_j \log Z_{\star,j}}{\sum_j^m \mu_j}. \quad (15)$$

#### 4.1.4. Moment analysis

After the description of the three steps present on the default procedure of analysis of pyFIT3D (Figure 1), we will move to revisit a second method of analysis of the ionized gas spectrum implemented on pyFIT3D. We present in Section 4.1.2 the emission line parameters derived through the mixed **RND+LM method** procedure. However, there are more emission lines than the strongest and frequently observed (and/or analyzed) ones, like those presented in the aforementioned Section. The parameters of the emission lines distributed by pyPipe3D comprise those for more than 50 emission lines and the most majority of them weak ones. It is not practical to perform a Gaussian fit for all of them because it would be very time consuming. Considering this, we implement a different scheme to extract the main properties of these emission lines.

This method is based on a direct estimation of the flux intensity, velocity, velocity dispersion and equivalent width (EW) of the emission lines from the moment analysis of the ionized spectrum. For the operation of the method it is needed (i) a list of central wavelengths of emission lines to be fitted (in the rest-frame), (ii) the ionized gas spectrum (once subtracted the stellar population model), (iii) the stellar population model spectrum and (iv) an error spectrum. In addition, an input guess estimation of the gas velocity (in km/s) and velocity dispersion (including the instrumental dispersion, in Å) are also required. For this initial guess we adopt the output of the H $\alpha$  emission analysis with the method described in Section 4.1.2.

For each emission line, the algorithm defines a wavelength range at which it is going to perform the analysis. This range is centered in the wavelength of the emission line provided by the list (described before), shifted by the input guess velocity, and it covers a wavelength interval of  $\pm 7.5$  times the input velocity dispersion around this central wavelength. In this wavelength range the code calculates the first three statistical moments, that corresponds to the flux intensity, the central velocity and the velocity dispersion of the considered emission line. This moment estimation is performed weighting the fluxes using a Gaussian Kernel that adopted as width the input velocity dispersion. In this way, the possible contributions of adjacent

(or even blended) emission lines are minimized, giving, at the same time, more weight to the regions of the emission lines with higher fluxes (and signal-to-noise values). Furthermore, the EW of the emission line is calculated by dividing the estimated integrated flux by the flux density of the underlying continuum. For doing so, the program estimates the flux density of the continuum from the stellar population model spectrum as the average flux density inside two wide (30 Å) side wavelength bands centered at  $\pm 60$  Å from the central one. The size of the bands is chosen to mitigate contribution from stellar absorption features, although this is not perfect and some effects are always impossible to avoid. Finally, the process is iterated for each emission line following a MC loop to derive the errors of the four estimated parameters.

#### 4.1.5. Stellar-indices analysis

The second additional analysis performed by pyFIT3D different than the modeling of the stellar population and the emission lines is the estimation of the equivalent widths of a predefined set of stellar indices. This is part of a classical technique to derive properties of stellar populations in galaxies from the measurement of determined absorption line strength indices, such as the Lick/IDS system (e.g. Burstein et al., 1984; Faber et al., 1985; Worthey, 1994). Those indices are sensitive to the age and metallicity of the stellar populations and provide model-independent information complementary to that described in Section 4.1.3 (fit of the stellar spectrum with multi-SSP templates).

The adopted algorithm of measurement follows the prescription (formulas and steps) implemented in INDEXF (Cardiel et al., 2003). Briefly, the algorithm estimates the EW for each of the stellar indices from the stellar population model spectrum using the same procedure to derive the EW of the emission lines described in Section 4.1.3. In order to perform the analysis over a set of MC realizations, the program also considers the residuals from the stellar population synthesis as a guess of the noise pattern.

The current version of the code includes the measurement of eight stellar indices (H $\delta$ , H $\delta$ mod, H $\gamma$ , H $\beta$ , Mgb, Fe5270, Fe5335 and D4000). We note that D4000 is not an EW, but a color parameter, that corresponds to the ratio between the flux density observed at wavelengths bluer and redder than the 4000Å. In this particular case we use the definition based on the flux density in wavelengths (D4000 $_l$ ) instead of the most frequently adopted one that uses the flux density in frequencies (D4000 $_v$ ) for this parameter. This parameter is frequently defined as B4000 and it corresponds to D4000 $_v$ /1.1619 (e.g. Gorgas et al., 1999). The list of adopted bandwidths for each parameter is indicated in S16b, Table 2. For the calculation of the equivalent widths (and color parameters), the bandwidths are shifted to the observed-frame using the velocity estimated from the analysis of the kinematics described in Section 4.1.1. At the end, the program derives the standard deviation of the measurements of each index along the MC simulation.



## 4.2. *pyPipe3D: revisiting the procedure*

So far we have described the treatment performed to analyze a single spectrum using `pyFIT3D`. Based on this fitting tool we develop a pipeline to provide with the spatially resolved properties of the stellar populations and emission lines from an IFS cube. This tool, named `Pipe3D`, was previously described in S16b. Here, we implemented the new `python` version of the code, and, based on the aforementioned philosophy described in Section 3, we create a single script, `pyPipe3D`, that performs all the steps of this pipeline. To describe this new script we revisit briefly the steps implemented in this pipeline highlighting the differences and updates with respect to the previous version.

`Pipe3D` was thought to be a generalist tool, able to analyze data from any IFS galaxy survey (or individual observations). However, each survey (and instrument reduction package) provides the data in a slightly different format. In order to apply `Pipe3D` (and `pyPipe3D`), it is needed to do some pre-processing on the data, which is not part of the pipeline itself. The goal of this pre-processing is to adapt any IFS data to the format expected by the pipeline. At the end, this format is a FITS file with at least two extensions (three preferred). The first extension should be a cube containing in each spaxel the observed spectrum (i.e., the flux intensity), with a regular linear step in the  $z$ -axis sampling the wavelength range. The spectra are assumed to be corrected for Galactic extinction. The other two extensions should have the same format (i.e., the same 3D world coordinate system, WCS). The second extension contains the  $1\sigma$  level of the error, estimated by the reduction process. Finally, the third extension should contain a 3D mask, including all the possible defects of the CCDs, cosmic rays masks, etc.

The analysis performed by `pyPipe3D` begins with the extraction and the fitting of the central spectrum ( $5''$  aperture) of the observed source, using the analysis described from Section 4.1.1 to 4.1.3. This first analysis helps the subsequent ones, recovering the redshift and velocity dispersion of the central region and allowing to automatically select the guess and range of values to explore for each non-linear parameter (Section 4.1.1).

The surface brightness of galaxies decreases as a function of the galactocentric distance, with the S/N following the same behavior, degrading the reliability of the analysis of the stellar continuum in the outer regions (e.g. Cappellari and Copin 2003; Cid Fernandes et al. 2013, 2014; S16a). `pyPipe3D` overcomes this problem with a process of spatial binning called Continuum Segmentation binning, or CS-binning. This process, fully described in S16b, combines the criteria adopted in other segmentation/binning schemes, the isophotal (Papaderos et al., 2002, 2013) and the Voronoi binning (Cappellari and Copin, 2003) schemes. The first method aggregates adjacent spaxels only if they belong to the same isophote (i.e., if their flux intensities differ less than a certain pre-defined percentage) and their distance is lower than a maximum value. By construction, this procedure preserves at a certain stage the shape of the original light distribution. However it does not guarantee to reach an optical S/N after binning, and the selection of the parameters is somehow arbitrary. On the other hand, the Voronoi binning uses as a unique criterium to aggregate adjacent spaxels

the S/N. The procedure provides with the minimum number of voxels that increases the S/N above a desired threshold defined by the user. In this way, it aggregates spaxels without any consideration to their physical properties (e.g. it can aggregate arm and inter-arm regions in the same tessella/spatial bin).

The CS-binning combines both criteria in a procedure that uses both a goal S/N and a maximum difference in the flux intensity to aggregate adjacent spaxels. In summary, the procedure starts by selecting the un-binned spaxel with the largest flux intensity in a certain band (5590–5680 Å). If this spaxel already fulfills the requirement in S/N (i.e., if it is larger than the desired goal), the spaxel is selected as a bin itself. If not, the algorithm looks for all the adjacent spaxels for which the flux intensity differs less than a certain fractional threshold with respect to this flux intensity of the considered spaxel. If this criteria is fulfilled then each adjacent spaxel is aggregated to the bin, and the S/N and average flux are reevaluated. If the S/N is now above the threshold, the bin is defined and the procedure starts looking for a new bin (i.e., looking for the highest flux spaxel within the unbinned ones). If not, further adjacent spaxels are incorporated into the tessella until one of the two criteria are not fulfilled. The full procedure is finished when no un-binned spaxels are found above a threshold in S/N (normally  $1\sigma$ ). Then, the code provides with (i) a RSS with each row comprising the average spectrum within each of the final tessellas, (ii) a map of the fractional contribution of each spaxel to this spectrum (the so-called dezonification map Cid Fernandes et al. 2013; S16b) and (iii) a segmentation map in which is stored the unique ID of each tessella, for a future identification of the average spectra with their location in the sky.

Figure 5 exemplifies this process for the CALIFA V500-datacube of galaxy NGC 2916. The three maps are: the flux intensity of V-band; the S/N map at this band and the CS-binned map. The average spectrum and the spectra of the binned spaxels of tessella 184 are shown at the bottom panel (this tessella was arbitrary chosen within those located near one effective radius from the center of the galaxy). This binned region is highlighted in all three maps. The S/N map shows the patchy pattern due to the three-pointing dithering scheme adopted by CALIFA (and other IFS surveys, like MaNGA or SAMI)<sup>6</sup>. The smooth distribution of the CS-binned map (top-right panel in Figure 5) shows that this spatial binning essentially removes this effect, conserving the shape of the light distribution of the galaxy. Our experiments have shown that when using other binning schemes, in particular the Voronoi binning, the light distribution of the galaxy is not preserved as well as using the adopted CS-binning. We need to recall the reader that this former binning was developed to explore the central regions of early type galaxies, which usually have a smooth light distribution without significant features (Cappellari and Copin, 2003).

<sup>6</sup>The final datacubes in these surveys are the result of combining at least three dithered pointings with fibers larger than the final selected spaxel. For this reason certain spaxels comprise the information of three fibers, while adjacent spaxels comprise the information of only two or just one of the original fibers. Thus, for the same flux intensity the S/N is not homogeneous, maintaining the triangular structure of the adopted dithering scheme.

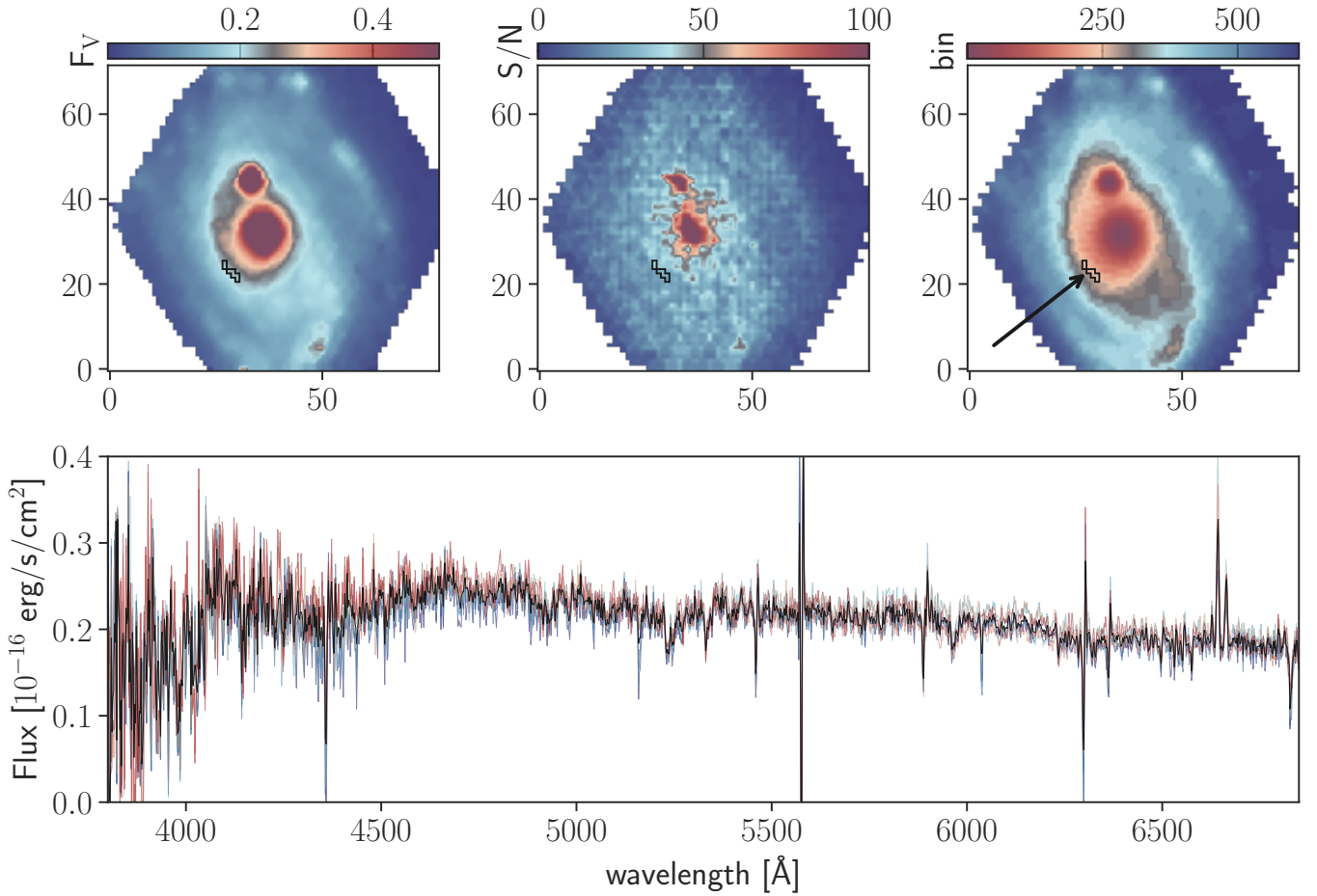


Figure 5: Example of the CS binning performed by pyFIT3D for the IFS datacube of the galaxy NGC 2916, observed by CALIFA. *Top panels:* From left to right, (i) map of the mean observed flux intensity with the wavelength range 5590-5680 Å, (ii) S/N map at this wavelength interval and (iii) the spatial bin (tessella) identification number (from 0 to 614 in this case). An arbitrary binned region (184) at approximately one effective radius of the center is highlighted on all three maps. The bottom panel shows the average spectrum (black) together with the individual spectra (red) within this tessella.

S16b already performed a direct comparison of the three aforementioned spatial binning procedures for the same dataset confirming this result.

At this point, pyPipe3D run the pyFIT3D algorithm (described in Sections from 4.1.1 to 4.1.3) for each individual spectrum of the CS-binned dataset. This process provides with all the parameters summarized in Section 4.1 at each location within the FoV of the IFS data. This way, for each parameter is generated a map (2D distribution), in which each derived parameter is associated with the location of the CS-bin defined by the segmentation map described before. Finally, all the maps with the distribution of properties are integrated into a set of cubes (3D arrays) for more convenient and simple distribution. These cubes will be described later on. In addition, pyPipe3D provides with an additional cube with the spectral information provided by pyFIT3D, comprising (i) the original spectra, (ii) the best model for the stellar populations, (iii) the best model combining both the stellar populations and the emission lines models, and (iv) three residual spectra: (a) once subtracted the combined model, (b) once subtracted the best stellar popula-

tion model (i.e., gas spectra), and (c) once subtracted the model spectra for the emission lines (i.e., emission-line cleaned spectra). This product has a cube structure since it comprises at each slice the RSS including these six spectra for each tessella, with the z-axis running for the number of spatial bins.

The pyPipe3D procedure continues with an analysis of the stellar indices, following the procedure described in Sec. 4.1.5. This exploration was performed for each individual spectrum corresponding to each tessella once subtracted the best model for the emission lines, removing their possible contribution/pollution to the estimation of the stellar indices. Like in the case of the previously described procedure the result of this analysis is a set of maps, one for each analyzed index (and another one for its corresponding error), with a spatial structure that follows the applied tessellation (i.e., all spaxels within the same spatial bin have the same estimated stellar index, by construction). Following the same storing philosophy all those maps are integrated into a single dataproducs cube for a more simple distribution of the results.

Finally, pyPipe3D re-analyze of the emission lines adopt-

ing two different procedures, the parametric fitting described in Sec. 4.1.2, and the moment analysis described in Sec. 4.1.4. Contrary to the previous explorations described above these two new analysis are performed spaxel-by-spaxel, not using the CS-binned RSS spectra. The reason behind that is that the spatial distribution of the continuum, that was the basis of the adopted tessellation procedure, does not necessarily corresponds to that of the ionized gas (which may present both a smooth component associated with the continuum and clumpy of filamentary structures Sánchez et al., 2021). Prior to any analysis it is constructed a GAS-pure datacube, i.e., a cube in which the contribution of the stellar population is removed. For doing so the best fitted stellar population model provided by the pyFIT3D algorithms for each tessella is re-scaled to the flux intensity of each spaxel within this tessella using the dezonification map. A further polynomial correction is applied to remove the possible color effects between the model (resulting from the fitting to the average spectra) and the individual spectrum. This matched stellar-population model spectrum is then removed to the original one creating a spectrum that contains the emission by the ionized gas (i.e., the emission lines), plus noise and residuals. This GAS-pure cube is also provided as a dataproduct of the pyPipe3D analysis.

The final analysis explores the strongest emission lines in the optical wavelength range ([O II] $\lambda$ 3727, H $\beta$ , [O III] $\lambda$ 4959, [O III] $\lambda$ 5007, [N II] $\lambda$ 6548, H $\alpha$ , [N II] $\lambda$ 6584, [S II] $\lambda$ 6716, [S II] $\lambda$ 6731, in the current implementation of the code) using the parametric fitting (assuming a Gaussian profile for each emission line), and the full set of strong and weak emission lines described before, for the non-parametric moment analysis, spaxel by spaxel using the GAS-pure datacube. Once more, like in the previous cases, each of these procedures provides with a map with the spatial distribution for each of the derived properties (flux, velocity...) and each of the emission lines, and their corresponding errors. For convenience, once more, all those maps are rearranged into two different dataproduct cubes, one for each of the procedures.

In summary, pyPipe3D provides with five different dataproduct cubes, that can be stored as individual FITS files, or as extensions of the same Pipe3D FITS file. Those dataproducts are tagged as:

- **SSP**: Main parameters derived from the analysis of the stellar populations, including the LW, and MW ages, metallicities, dust attenuation and stellar kinematics properties, derived for each tessella;
- **SFH**: Weights of the decomposition of the stellar population for the adopted SSP templates library (e.g., values in Eq. 10), derived for each tessella. It can be used to derive the spatial resolved star-formation and chemical enrichment histories of the galaxies and the LW and MW properties included in the **SSP** dataproducts;
- **INDICES**: Set of stellar absorption indices derived for each tessella once subtracted the emission line contribution;
- **ELINES**: Flux intensities for the strongest emission lines in the optical wavelength range, together with the kinematics

properties of H $\alpha$ , derived based on a Gaussian fitting of each emission line, spaxel-by-spaxel;

- **FLUX\_ELINES**: Main parameters of a set of strong and weak emission lines (~50) derived using a weighted momentum analysis based on the kinematics of H $\alpha$ , derived spaxel-by-spaxel. It includes the flux intensity, equivalent width, velocity and velocity dispersion, and the corresponding errors for the different analyzed emission lines.

## 5. Accuracy of the fitting code

As indicated before, the full code has been transcribed and completely re-coded to python with a set of modifications. Therefore, it is required to evaluate the accuracy and the precision of all the recovered parameters. In order to do so we run the code on a set of simulated and real spectra.

### 5.1. Testing the stellar population analysis

We generate a set of 2000 realistic simulated stellar spectra by adopting the set of parametric SFH and ChEH recently published by Mejía-Narváez et al. (2020), Appendix B. These evolutionary histories were designed to reproduce the wide range of M/L,  $\mathcal{A}_\star$  and metallicity distribution functions observed in the most recent IFS galaxy surveys such as CALIFA or MaNGA (e.g. Sánchez et al., 2018; Lacerda et al., 2020). The considered SFH and ChEH were transformed to age and metallicity distribution functions in stellar masses, re-sampled to the grid of values covered by the adopted gsd156 SSP library. Once we estimate the amount of stellar mass that corresponds to each SSP template (e.g., values described in Eq. 13), this fraction is transformed to light based on the corresponding  $Y_{\lambda_{\text{norm}}}$  (i.e., coefficients described in Eq. 10). Then, applying Eq. 5, a simulated stellar spectrum is created corresponding to a particular SFH and ChEH. Finally, the set of simulated spectra are (i) re-sampled to  $2 \text{ \AA}/\text{pix}^7$ , (ii) shifted to a specific redshift; (iii) convolved with a Gaussian function of width  $\sigma_\star$  to replicate the LoSVD of the stars and (iv) attenuated by the dust extinction value obtained applying the extinction law by Cardelli et al. (1989). Hence, for each simulation we adopt a set of non-linear parameters randomly selected from a flat distribution within pre-defined intervals: (i) from 0.005 to 0.05 for  $z_\star$ ; (ii) from 75 and 350 km/s for the  $\sigma_\star$ ; and (iii) from 0 to 1.6 mag for the  $A_V^\star$ . These parameters were selected to emulate the typical values for nearby galaxies.

Finally, we add white noise to the simulated spectra following a Gaussian distribution in order to simulate the range of S/N values reported in the most recent IFS galaxy surveys (1- $\sigma$  of the input have S/N between 50 and 100, and all simulated data with S/N values ranging between 20 and 300). We do not include other types of noise pattern in the simulated spectra, limiting our exploration to situations in which problems like spectrophotometric calibration, errors in sky subtraction, CCD defects, etc, are negligible. Furthermore, no emission lines have

<sup>7</sup>to mimic the sampling of the CALIFA V500 dataset

been included in these simulations. However, it is well known that a well-defined set of masks for the emission lines (e.g. Sten- cel, 1977; Matsuoka et al., 2007; Schröder et al., 2009) is essential for a good determination of the stellar parameters in real data.

We adopted the same configuration files, initial *guess* parameters and *intervals* for the analysis of all the simulated spectra. For the initial *guess* we use the average values covered by our simulations and as *intervals* the simulated range of values. For these simulations we use the same stellar population library to simulate the spectra and to model them. Thus, the `gsd156` library is adopted for the analysis of the stellar populations and the `gsd12` for the non-linear step. So far we want to explore how well the code reproduces well-known and controlled parameters. For a discussion on the effects of fitting a simulation generated with a particular stellar library using other libraries we refer the user to Cid Fernandes et al. (2014).

### 5.1.1. Recovery of simulated parameters

Figure 6 shows the comparison between the input (I) and recovered (R) values for the following stellar parameters:  $z_*$ ,  $\sigma_*$ ,  $A_V^*$ ,  $\langle \log \mathcal{A}_* \rangle_L$  (Eq. 11),  $[Z/H]_L$  (Eq. 12) and the stellar mass-to-light ratio ( $\Upsilon_V$ ), based on the simulations described before. The offset between the input and recovered parameters and the corresponding standard deviations have been included in each panel of the figure. It is clearly appreciated that the all parameters are very well recovered.

Among the non-linear parameters, the redshift is the one recovered with the most precision and accuracy by `pyFIT3D`. This is due to the procedure applied on the determination of the non-linear parameters. We have to remember that there is no curve fit of the parameter, i.e. the method relies on the lowest value of Eq. 2, which is primarily affected by the match between the central wavelengths of the stellar absorption lines. On the other hand, the  $\sigma_*$  is the non-linear parameter recovered with less accuracy ( $\sim 2$  km/s) and precision ( $\sim 22$  km/s). Although `pyFIT3D` has not been developed optimized for kinematic analysis of the stellar population, the values are recovered with great accuracy and without any significant bias related to the input values. However, the precision is somehow affected. The extraction of the LOSVD is a degenerated problem and both the adopted SSP library and the LOSVD profile chosen (in `pyFIT3D` case, a Gaussian one) may affect it (e.g. Bender, 1990; Statler, 1995; Cappellari and Emsellem, 2004; Ocvirk et al., 2006; Falcón-Barroso et al., 2011; Falcón-Barroso and Martig, 2021). For the last parameter determined during the non-linear fit step, the dust attenuation, there is a good agreement between input and recovered values. However, the determination of  $A_V^*$  is another degenerated problem, which affects and is affected by the determination of other parameters. We will discuss this matter forward along this Section.

The remaining stellar parameters are then derived using the non-linear parameters already determined and fixed, following the procedure described in Section 4.1.3. From this analysis we recover the coefficients of the decomposition of the stellar population in the considered SSP library, from which we derive average properties, both LW and MW. Along the bottom row of

Figure 6 we evaluate the capability of `pyFIT3D` to recover these properties. The LW mean age is the best recovered stellar property, with an accuracy of  $\sim -0.01$  dex and a precision of  $\sim 0.05$  dex. On the other hand, the metallicity is recovered slightly worse, with no appreciable offset and a precision of  $\sim 0.06$  dex, on average. Contrary to the age, the metallicity presents a negative bias (recover lower values) at the high metallicity range, and a positive bias (recover higher values) at low metallicity. Usually the age is the main source of variance between the SSP templates with the metallicity responsible for higher order effects (e.g. Ronen et al., 1999). Thus, the age is the parameter that it is a priori easier to distinguish between different stellar populations, what explain why it is better recovered than the metallicity. For the stellar  $\Upsilon_V$  we see a good agreement between the input and recovered values, with an average accuracy of  $\sim 0.01$  dex and a precision of  $\sim 0.04$  dex. Like in the case of the metallicity we appreciate a positive offset (larger recovered values), in the range of larger  $\Upsilon_V$  values ( $> 0.75$  dex). However, this bias is never larger than a few percent.

Table 1: Stellar population simulation: statistics of the residuals for a typical S/N of  $\sim 60$ .

$\Delta(\text{par})$	mean $\pm$ stdev
$v_*$	$13.71 \pm 21.23$ km/s
$\sigma_*$	$4.94 \pm 20.32$ km/s
$A_V^*$	$0.01 \pm 0.04$ mag
$\langle \log \mathcal{A}_* \rangle_L$	$-0.01 \pm 0.04$ dex
$[Z/H]_L$	$0.01 \pm 0.07$ dex
$\Upsilon_V$	$0.01 \pm 0.04$ dex

Results of the mean  $\pm$  the standard deviation of the distributions.

### 5.1.2. Interdependence on the ability of estimating the parameters

Until now we characterized the capability of `pyFIT3D` to recover a set of input parameters. Now we proceed to characterize how the accuracy and precision in the determination of one parameter affect other ones, i.e., the bias in the derivation of those parameters. Figure 7 shows the pair-plot for the distributions of the difference between the recovered and input values ( $\Delta\text{par} = \text{par}_R - \text{par}_I$ ), for the set of parameters shown in Figure 6 (discussed in the previous Section). In addition, we include a first column with the input S/N of each run, illustrating the precision and accuracy of the fit accordingly to the input goodness of data, i.e., the relative residual offsets as a function of the S/N. In each panel is represented a pair of parameters, with the diagonal representing the corresponding kernel density estimate for the matching parameter. The median and standard deviations of the distributions included in this diagonal are the same as the ones shown in Figure 6, being included only for completeness. The plots involving  $\Delta(z)$  have been deliberately removed since the redshift presents almost no bias, according to the results shown in the previous Section, and for this particular case it is appreciated no correlation at all with any of the explored  $\Delta\text{par}$ .

Examining the first column of Figure 7 we see no correlation between the residuals and the input S/N, what tunes the

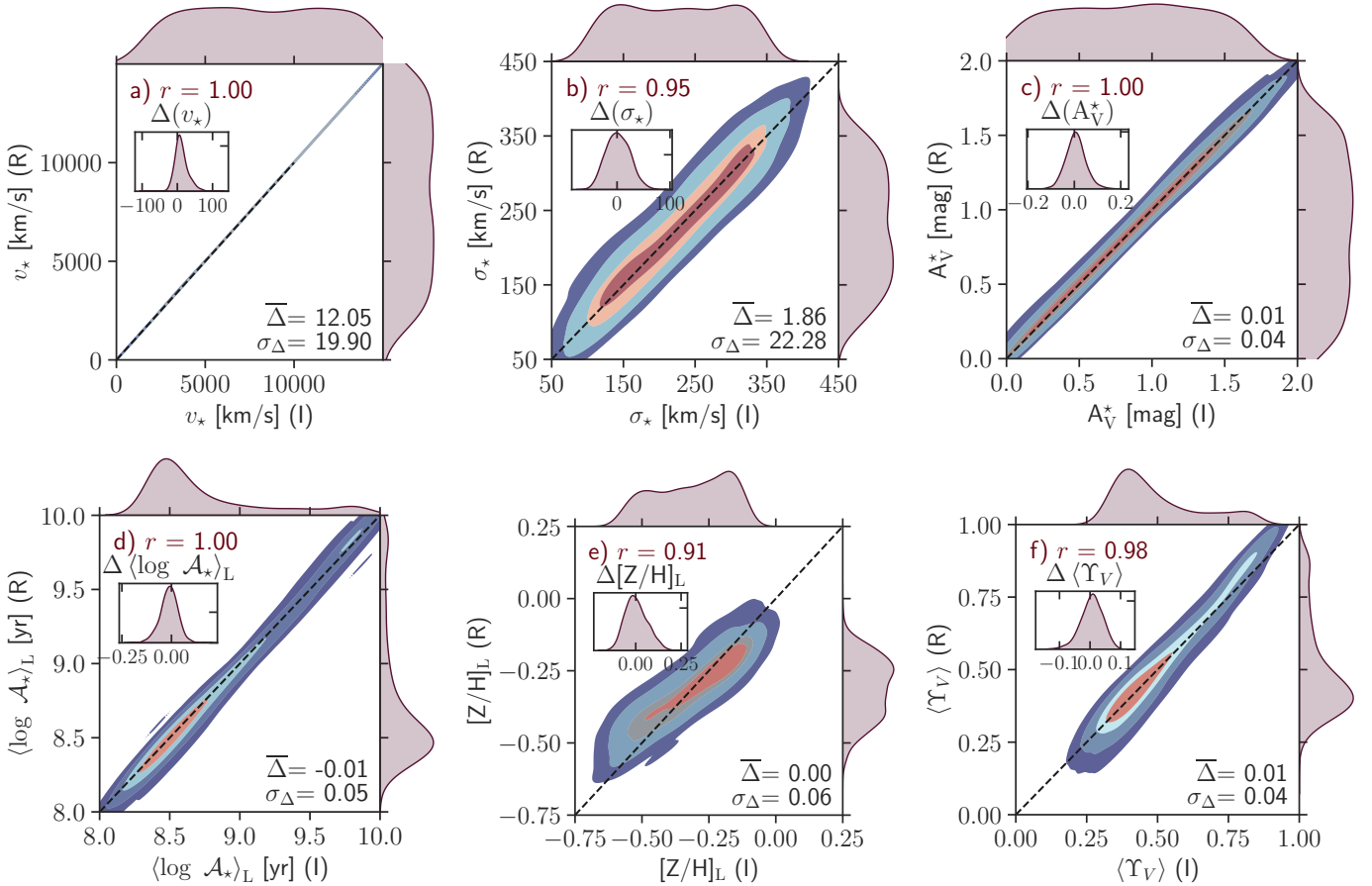


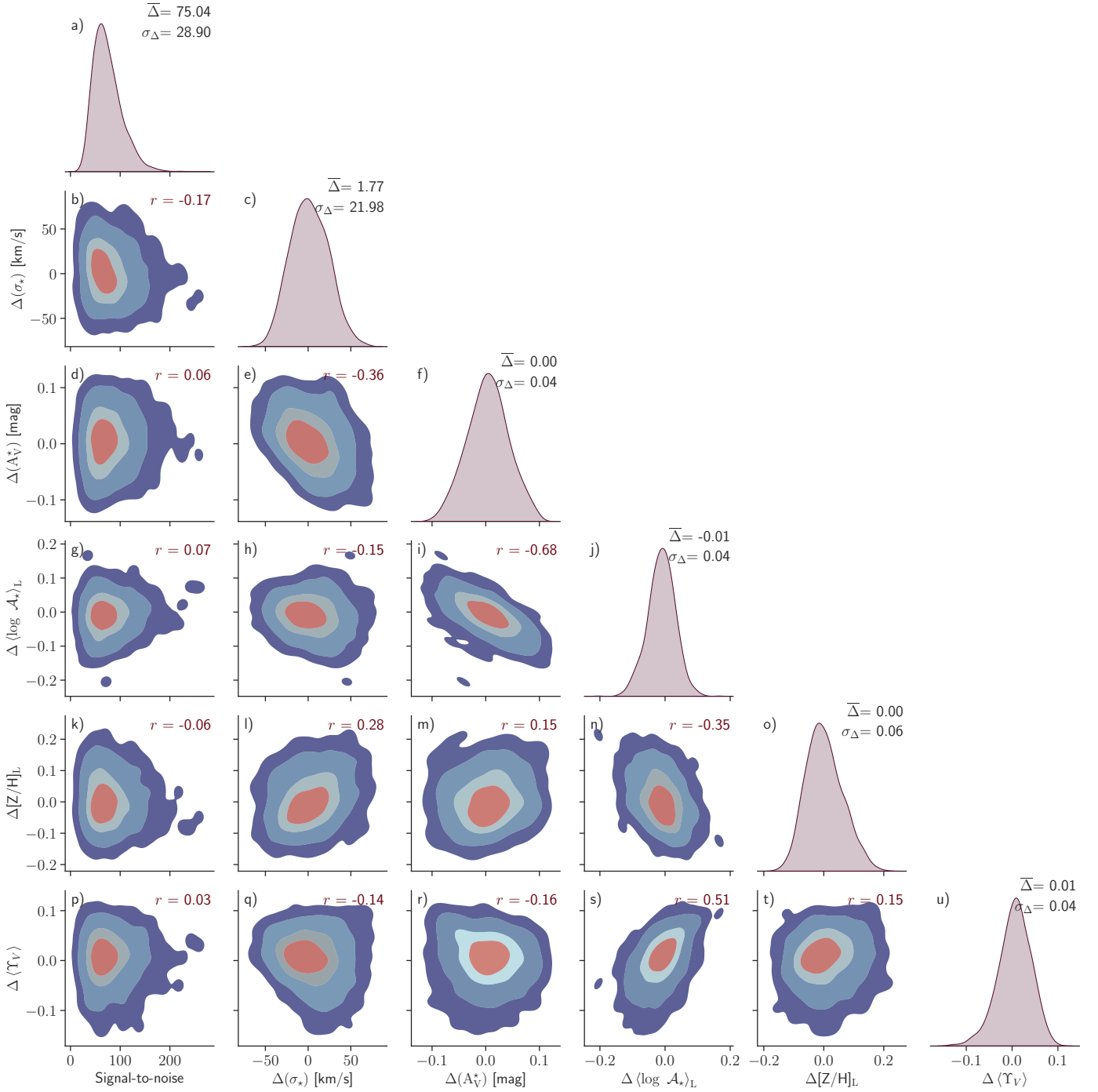
Figure 6: Stellar parameters recovered by pyFIT3D (R) as a function of the input ones (I) for a set of 2000 simulations with an average S/N $\sim$ 60. Upper panels correspond, from left to right, to the non-linear parameters: redshift ( $z_*$ ), line-of-sight velocity dispersion ( $\sigma_*$ ) and  $A_V^*$ . Lower panels correspond, from left to right, to the parameters recovered by the stellar population analysis: LW age ( $\langle \log \mathcal{A}_* \rangle_L$ ), LW metallicity ( $[Z/H]_L$ ) and the stellar mass-to-light ratio ( $\Upsilon_V$ ). Each panel shows the Pearson correlation coefficient ( $r$ ), the mean and the standard deviation for the distribution, with each contour encircling a 1, 2, 4 and 6 standard deviations of the kernel density estimate (KDE) of points. We warn the reader that KDE method introduces extrapolation effects in the distributions. Most of data points used to generate distributions are within  $2\text{-}\sigma$ . The projected KDE of each parameter has been included at the upper and right borders of each panel. The inset in each panel shows the KDE of the difference between both parameters, i.e.,  $\Delta(\text{par})$

code accuracy. We also note that the results precision is improving with the increasing of the input S/N. We show in Table 5.1.1 the statistics (mean and standard deviation) of the explored residuals for a typical S/N of  $\sim$ 60. Conversely, there are clear correlations that highlight the interdependence in the estimation of the different explored parameters, in agreement with the results found in the literature using different stellar synthesis codes (e.g. Cid Fernandes et al., 2005; Sánchez-Blázquez et al., 2011; Cid Fernandes et al., 2014). The strongest correlation is found between the  $\Delta A_V^*$  and  $\Delta \langle \log \mathcal{A}_* \rangle_L$  ( $r = -0.71$ ), indicating that an over-estimation of the dust attenuation produces an under-estimation of the LW age of the stellar population (and the other way around). This degeneracy is somehow expected since the dust attenuation and the age produce similar effects in the shape of the spectra (i.e., making them redder). Despite of its strength we need to highlight that the effect involves a maximum variation of  $\sim 0.1$  dex in the  $\mathcal{A}_*$  for an error of  $\sim 0.2$  mag in  $A_V$ . The second stronger correlation is that of the  $\Delta \langle \log \mathcal{A}_* \rangle_L$  with the  $\Upsilon_V$  ( $r = 0.51$ ), that illustrates the direct connection between both parameters. As indicated before, the

strongest variance in the SSPs is due to the  $\mathcal{A}_*$ , changing essentially the  $\Upsilon$  ratio. Since older stellar populations have larger  $\Upsilon$  ratios than younger ones, the observed positive relationship between the two residuals is naturally understood.

The third strongest correlation ( $r = -0.37$ ) corresponds to the well known degeneracy between age and metallicity: an increase (decrease) in the fraction of old and metal rich stars have a similar effect, i.e. the spectra becomes redder (bluer). Furthermore, intrinsic uncertainties in the stellar evolution models, in the understanding of the stellar evolution and stellar contributions for the light in different wavelength ranges depending on the age and metallicity contributes to this degeneracy too (e.g. Worthey, 1994; Cid Fernandes et al., 2005; Sánchez-Blázquez et al., 2011). The effect in the residuals of both parameters is of the same order as the one reported by (Cid Fernandes et al., 2014) when using the STARLIGHT code for a similar SSP library. In that exploration, it was found that the use of different libraries may affect this degeneracy, but without completely removing it.

An almost similar correlation is found between the  $\Delta \sigma_*$  and



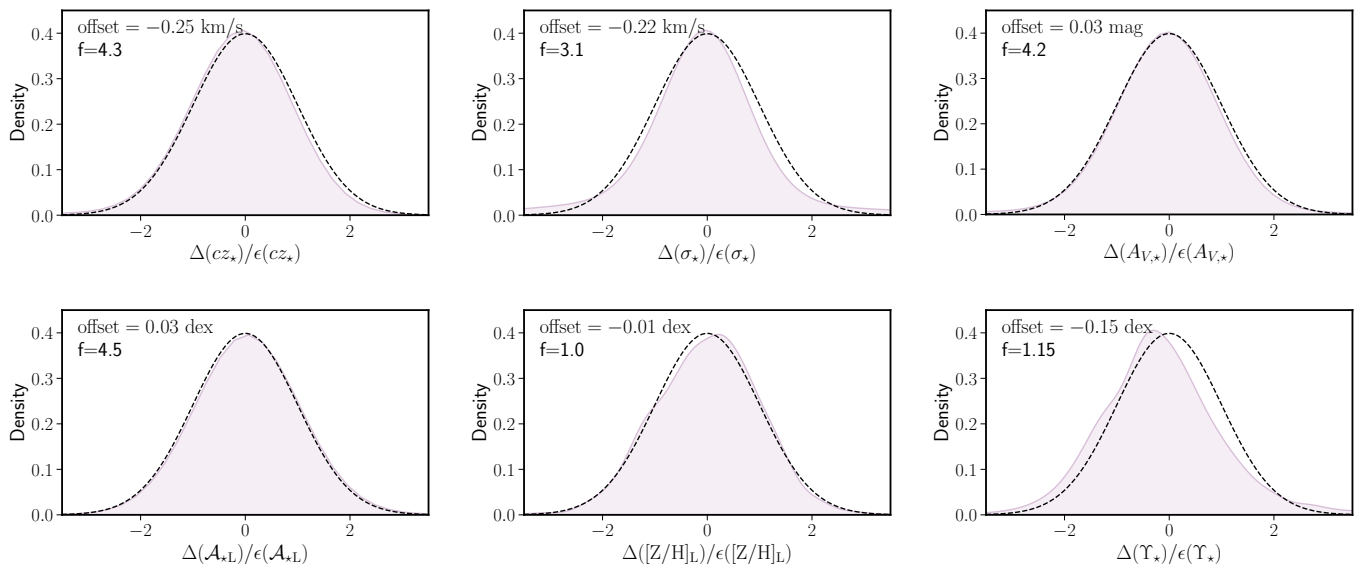


Figure 8: Density distribution (shaded curve) of the uncertainty in the derivation of the explored stellar parameters,  $\Delta(\text{par})$  (i.e., the estimated value minus the input/simulated one), divided by the error estimated by pyFIT3D,  $\epsilon(\text{par})$ . Top-panels show the distribution for the non-linear parameters ( $z_*$ ,  $\sigma_*$  and  $A_V^*$ ), while bottom panels the mean properties of the stellar populations ( $\langle \log \mathcal{A}_* \rangle_L$ ,  $[Z/H]_L$  and  $\Upsilon$ ). The dotted line in each panel corresponds to a Gaussian function centered in zero with a standard deviation of one, representing the expected distribution if the estimated errors were fully representative of the real uncertainties in a statistically way.  $\epsilon(\text{par})$  has been multiplied by a factor  $f$  (shown in each panel) to match the observed and expected distributions as much as possible. For comparison purposes we also added the offset in  $\Delta(\text{par})$ , in the original units of each parameter, to illustrate the nature of the observed shift in some panels (i.e., a systematic bias in the derivation of the parameters, not related with the estimation of the errors).

$\Delta A_V^*$  ( $r = -0.35$ ), in the sense than an over-estimation of the velocity dispersion produce an under-estimation of the dust-attenuation. Since a change in the velocity dispersion does not introduce a direct change in the shape of the observed spectrum (the primary property sensitive to the dust attenuation), this degeneracy should be induced by other interdependence among the explored parameters. The more suitable candidate is the effect that the velocity dispersion has on the derivation of the stellar metallicity: a deeper absorption feature can be obtained by either increasing the velocity dispersion or the metallicity. This leads to a correlation between the bias/errors when trying to derive both parameters at the same time (e.g. Koleva et al., 2008; Sánchez-Blázquez et al., 2011). We indeed find a very weak trend between  $\Delta \sigma_*$  and  $\Delta [Z/H]_L$  ( $r = 0.28$ ), an effect that via the stronger age-metallicity and age-dust degeneracy may induce the observed trend.

All the remaining parameters do not present any clear trend among their residuals, what suggests that their derivation is essentially independent: (i)  $\sigma_*$  vs  $\Upsilon_V$  and  $A_V^*$ ; (ii)  $A_V^*$  vs  $\Upsilon_V$ ; and (iii)  $[Z/H]_L$  vs  $\Upsilon_V$ .

### 5.1.3. Accuracy of the estimated uncertainties

One of the goals of pyFIT3D is to give a good estimation of the uncertainties for each derived parameter. The stellar population decomposition in pyFIT3D comprises an intrinsic MC method over perturbed realizations of the input spectrum. This process enable us to estimate the uncertainties of the estimated coefficients of the stellar decomposition in the considered SSP library (Equation 9). These uncertainties are propagated to estimate the errors in the averaged properties of the stellar pop-

ulations (such as the LW and MW ages and metallicities). On the other hand, the uncertainties estimated in the exploration of non-linear parameters (Sec. 4.1.1) are derived from the analysis of the  $\chi^2$  curves (Figure 2, right columns), assuming as a minimum error the average sampling of the explored range. This latter estimation is valid when the parameters are independent one each-other and the errors have a normal Gaussian distribution, what it is not true in many cases (as discussed in the previous Section). In general, there is no guarantee a priori that the estimated errors are representative of the real ones.

In order to determine how the errors estimated by pyFIT3D for each parameter,  $\epsilon(\text{par})$ , are representative of the real ones we compare them with the value of the offsets between the output and input values,  $\Delta(\text{par})$ , for our simulated dataset. As indicated before, in the case of an accurate estimation of the parameters (see Sec. 5.1.1) this residual is a good estimation of the precision and therefore a realistic tracer of the error. We should note that uncertainties and errors should match only in an statistical sense. A direct comparison between both parameters is presented in Appendix Appendix B. Figure 8 shows the distribution of the ratio between both parameters, where the error is scaled by a factor  $f$  (i.e.,  $\Delta(\text{par})/(f\epsilon(\text{par}))$ ). In the ideal case in which our estimated errors are a good representation of the real uncertainties these distributions should follow a perfect Gaussian function (shown in Figure 8), centered in zero (with any offset measuring the accuracy of the estimation of the error), and with a standard deviation of one (with any difference measuring the precision of the estimation of the error). The mean value of  $\Delta(\text{par})$  is shown in each panel, showing that pyFIT3D provides an accurate estimation of the parameters, as discussed

in the previous Section. In addition we include in each figure the factor  $f$  introduced to match the observed and expected distributions, i.e., the correction that has to be applied to the reported errors to be fully representative of the real uncertainties. It is noticed that pyFIT3D underestimates the errors by a factor  $\sim 3$ – $4$  for several parameters ( $z_*$ ,  $\sigma_*$ ,  $A_{V,*}$  and  $\langle \log \mathcal{A}_* \rangle_L$ ), although in other cases the error is of the same order of the uncertainty ( $[Z/H]_L$  and  $\Upsilon_V$ ).

We are not completely aware of the nature of the reported differences between the estimated and the real errors for certain parameters. However, we consider that a combination of reasons may produce them. The most relevant one is that the estimation of the errors relies on the assumption that the derivation of the parameters is fully independent, and that the errors are propagated linearly from the uncertainties. Both assumptions are not fully valid, based on the results described in Sec. 5.1.2, where we reported clear correlations in the residuals between the estimated and input values for different parameters. The results of this analysis force us to make a revision on the currently adopted procedure to propagate the errors. However, although the nominal errors reported by the code present clear offsets, by applying the reported correction factors the code is able to provide with reliable and representative values of the real uncertainties.

## 5.2. Recovering the properties of the emission lines

Along this Section we estimate the accuracy and precision on the recovery of the properties of the emission lines based on the fitting procedure implemented in pyFIT3D. For doing so, we simulate 10000 spectra including the  $[\text{N II}]\lambda\lambda 6548, 6583 + \text{H}\alpha$  emission line system (i.e. three different emission lines), fixing their integrated flux intensities to 22, 66 and 100 respectively (in arbitrary units). Then, for each simulation we consider a different velocity and velocity dispersion for the entire system (i.e., same kinematic properties for the three lines), randomly chosen within the following predefined ranges: 5500 – 6500 km/s and 2.5 – 6.5 Å ( $\sigma_v \approx 100$  – 400 km/s) respectively. The simulated spectra covers the wavelength range between 6500 – 7000 Å (rest-frame), repeating the sampling of 2 Å/pix employed in the stellar analysis simulation. Each emission line is simulated adopting a Gaussian profile, with the corresponding flux intensities listed before, shifted by the systemic velocity and broadened by the velocity dispersion corresponding to each simulation. Subsequently, we add different levels of noise, assuming a Poissonian distribution, to cover a wide range of S/N ratios between  $\sim 0.1$  and  $\sim 1500$ . We define here the S/N as the ratio between the peak flux-intensity of the emission line and the  $1\sigma$  level of the simulated white noise. For a conversion to the conventional S/N of the integrated flux, a factor of 3 should be applied to the reported S/N values ( $S/N_{\text{int}} \approx 3 S/N_{\text{peak}}$ ). Finally, all simulated spectra were fitted using the **RND+LM** procedure described in Section 4.1.2. In addition, we repeat the fit to the simulated spectra using only the **RND** method in order to test the improvements from the introduction of a second step in the emission lines fit procedure employed by pyFIT3D in comparison with the previous version of FIT3D.

Table 2: Emission lines simulation: statistics of the relative residuals segregated by S/N

S/N	$\Delta\text{Flux}/\text{Flux}_{\text{in}}$	$\Delta\sigma/\sigma_{\text{in}}$	$10 \times \Delta v_{\text{sys}}/v_{\text{sys,in}}^1$
< 3	$0.00 \pm 0.16$	$-0.04 \pm 0.12$	$-0.01 \pm 0.16$
3 – 10	$-0.01 \pm 0.09$	$-0.01 \pm 0.06$	$0.00 \pm 0.07$
10 – 100	$-0.01 \pm 0.08$	$-0.01 \pm 0.05$	$0.00 \pm 0.04$
> 100	$-0.01 \pm 0.07$	$-0.01 \pm 0.04$	$0.00 \pm 0.03$

Results of the mean  $\pm$  the standard deviation of the distributions.

- (1)  $\Delta v_{\text{sys}}$  statistics is multiplied<sup>2</sup> by 10 to show it in the same scale of the other two parameters.
- (2) We did not include the multiplication factor to the data points during the statistical calculation avoiding possible systematic bias.

For this simulation we use the same input configuration for all rounds (see Appendix Appendix C for more details on the adopted configuration). The values for the *guess* (*interval*) adopted for the systemic velocity and the velocity dispersion were 6000 (5500 – –6500) km/s and 2.7 (2.5 – –6.5) Å respectively. However, the *intervals* are rewritten by pyFIT3D taking into account the shift introduced in the simulated spectrum due to the systemic velocity. All emission lines in the modeled system are set to have the same kinematic parameters (systemic velocity and velocity dispersion). In addition, we set a physical link between the integrated flux of each  $[\text{N II}]\lambda\lambda 6548, 6583$  line as  $[\text{N II}]\lambda 6548 = (1/3)F_{[\text{N II}]\lambda 6584}$ . On the other hand, no link was imposed between  $[\text{N II}]\lambda 6584$  and  $\text{H}\alpha$ , i.e. both are free parameters on the simulation. By not fixing these values, the simulation can cover different sources of ionization (e.g. AGNs, H II regions, post-AGBs; Kewley et al. 2001; Kauffmann et al. 2003; Stasińska et al. 2006; Cid Fernandes et al. 2010).

### 5.2.1. Bias on the estimation of the parameters for the emission lines

Figure 9 gives an overview of the results of this exploration showing a pair-plot with the relative residuals ( $\Delta\text{par}/\text{par}_{\text{in}}$ ) of the different parameters of the simulated emission lines system together with the input S/N of each simulation run. Table 5.2 summarizes the results showing the mean and the standard deviation of the distributions of the relative residual offsets for four different S/N ranges. Figure 10 shows four examples of the analysis (one for each S/N range of Table 5.2).

Exploring the first column of Figure 9 we see no correlation between the input S/N and the relative residuals (i.e., accuracy of the results), with a clear improvement of the precision of the fitting with the S/N, confirmed by the residuals at each S/N range at Figure 10. Even when the S/N is lower than 3, where the systematic effects dominate (like errors in the stellar population subtraction and noise patterns not considered in this simulation), the precision on the flux intensity is not worse than  $\approx 16\%$  ( $\approx 12\%$  for the velocity dispersion). Furthermore, pyFIT3D provides with an almost constant precision and accuracy above  $S/N > 10$  for all parameters. However, the precision for those parameters does not get better than  $\approx 4\%$  even in the high S/N cases considered in this simulation. In the case of



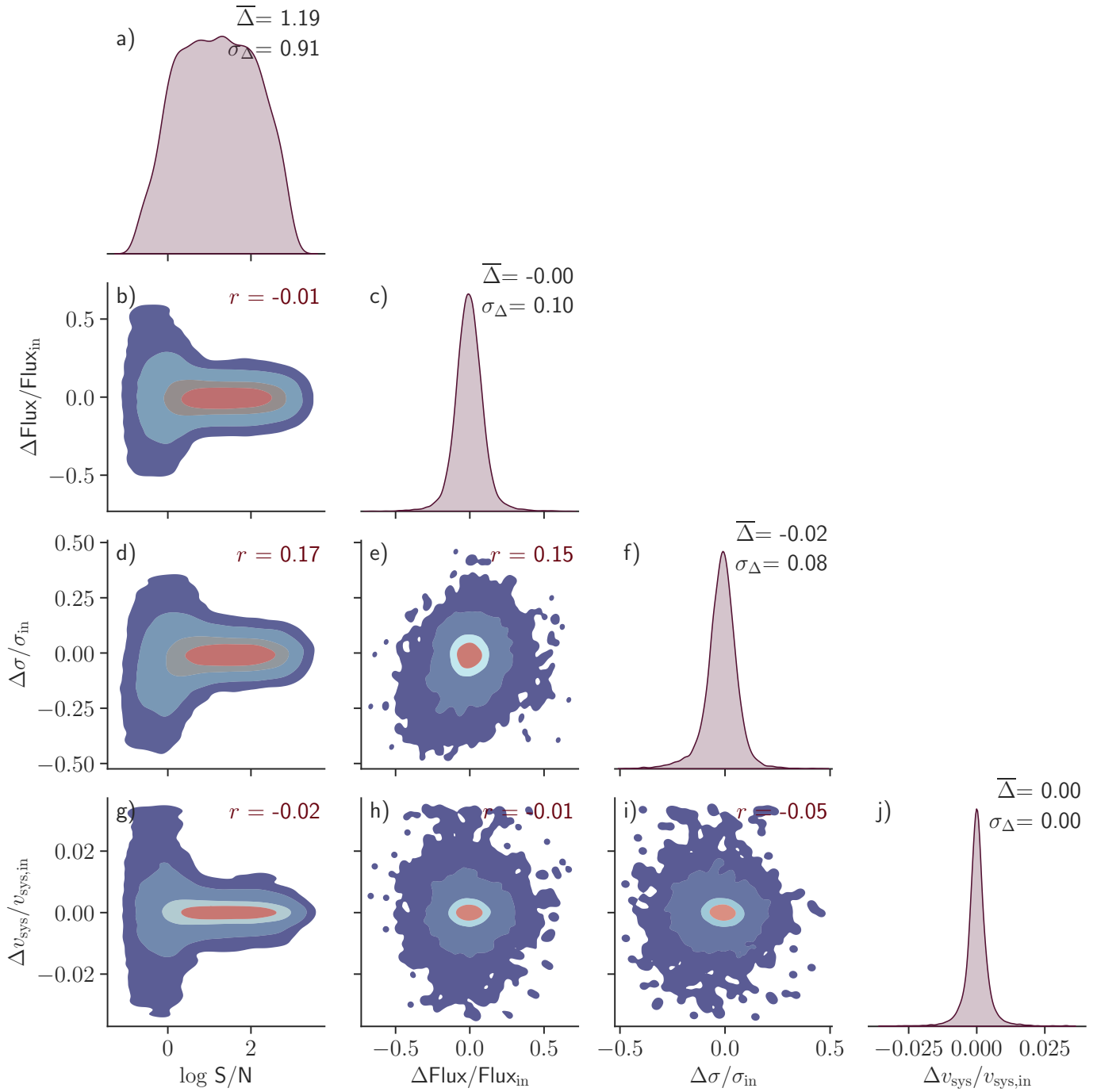


Figure 9: Pair-plot with the relative residuals for each explored properties of the  $[\text{N II}]\lambda\lambda 6548, 6583 + \text{H}\alpha$  simulated emission lines and their peak flux-intensity  $S/N$  using the new double step fitting method (RND + LM) implemented in pyFIT3D. Contours, density histograms, and legends in each panel follow the same scheme presented on Figure 7.

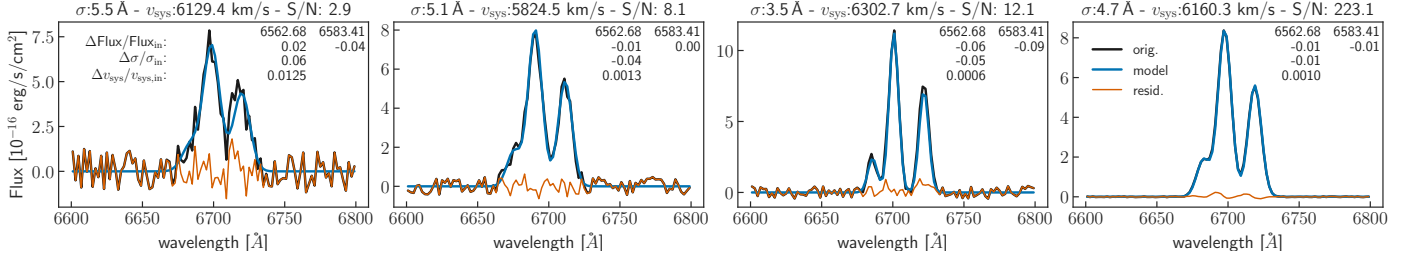


Figure 10: Four examples of the simulated spectra analysis, one for each S/N interval with statistics presented in Table 5.2. At each panel we show the original spectrum (black), the modeled spectrum (blue) and the residual spectrum (red). We also show the input parameters, peak S/N and the relative residuals of the simulation analysis.

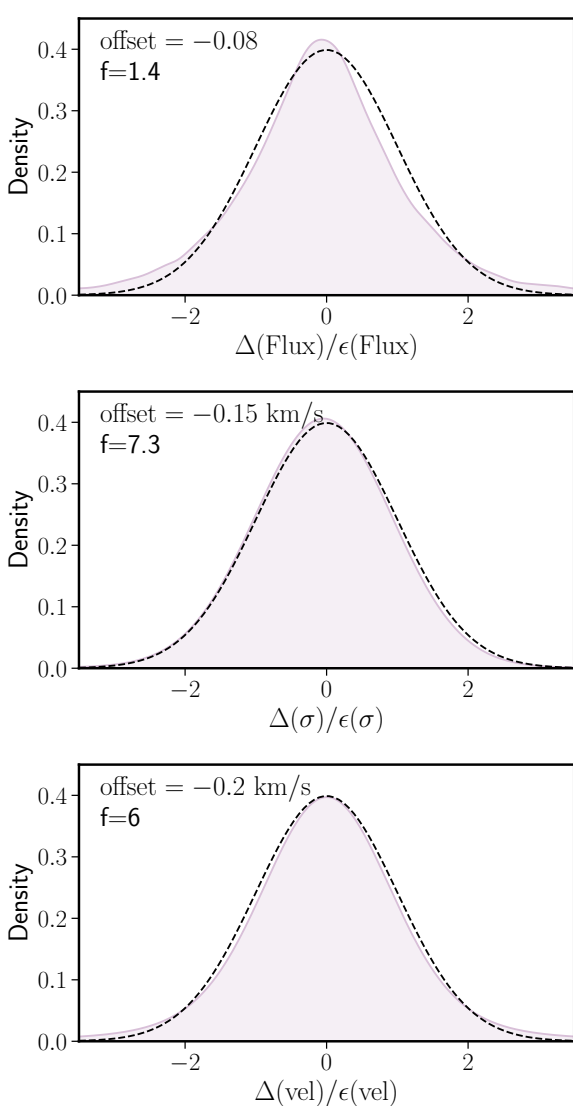


Figure 11: Similar figure as Figure 8 for the properties derived from the analysis of the simulated emission lines. From top-to-bottom we show the distribution of  $\Delta(\text{par})/\epsilon(\text{par})$  for the flux intensity, the velocity dispersion, and the systemic velocity.

the systemic velocity, the program does recover this parameter with a precision and an accuracy  $\approx 10\%$  worse than the other two. Although the accuracy on the flux and the velocity deter-

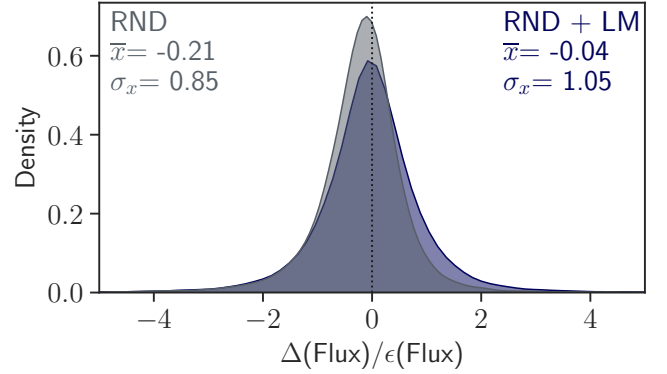


Figure 12: Density distributions of the uncertainty in the derivation of the flux,  $\Delta(\text{Flux})$ , divided by the error estimated by pyFIT3D,  $\epsilon(\text{Flux})$ , for two rounds of fittings of 10000 simulated spectra comprising the [N II]  $\lambda\lambda 6548, 6583 + \text{H}\alpha$  emission lines system. For the first round (light-red) it was adopted the **RND method**, while for the second one (dark-red), it was adopted the combined **RND+LM methods**. Upper-left and right corners present the mean and standard deviation of each distribution.

mination is less affected when compared to the sigma one, the uncertainties begin to suffer a considerable increase at the very low S/N regime ( $S/N < 1$ ), with offsets on the recovered values around 20% on the flux and the velocity, and above 15% for the sigma determination.

The estimation of the integrated flux depends directly on the velocity dispersion, so, in principal, any bias in the determination of one could affect both the accuracy and the precision of the determination of the other. However, we see no correlation between the relative residual offsets, i.e. the determination of the parameters do not present any interdependence bias. The bulk of the distribution rests within  $\pm 0.1 \Delta \text{par}/\text{par}_{\text{in}}$  in both axis for all three parameters (panels *e*, *h* and *i* of Figure 9).

Repeating the experiment using only the fit with the **RND method** increases the values of the standard deviations of the relative residuals with  $S/N > 3$ , that in this case shows similar distributions as the simulations with  $S/N < 3$  (in agreement with the results presented in S16 using the former version of the code). This tell us that the inclusion of a second round of fit with the **LM method** increases the precision of the estimated parameters. Other notable result is that the correlation coefficient between  $\Delta\sigma/\sigma_{\text{in}}$  and  $\Delta\text{Flux}/\text{Flux}_{\text{in}}$  increases from 0.15 to 0.26. Although this is a mild increase, it is a significant change. This time we could conclude that the inclusion of the

**LM method** as a second step of the emission line fit process also helps to decrease the Flux- $\sigma$  bias, already found for the previous version of the code (S16).

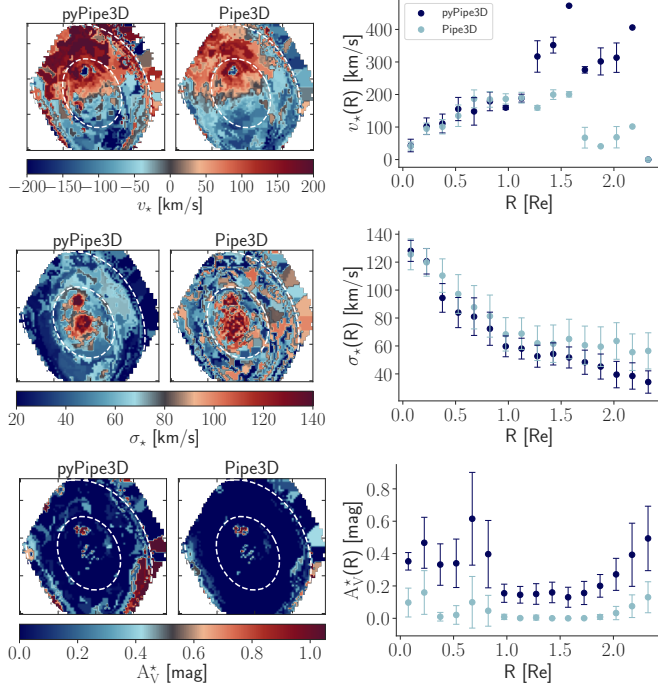


Figure 13: Comparison of the spatial distribution (left panels) and radial profiles (right panels) for the non-linear parameters derived by pyPipe3D and Pipe3D for the NGC 2916 galaxy observed by CALIFA:  $v_{\text{sys}}$  (top panels, in absolute values),  $\sigma_{\star}$  (middle panels),  $A_V^{\star}$  (bottom panels). The white ellipses in the maps represents the location of one and two effective radius (Re). The errors in the left panels represent the mean and  $1\sigma$  interval of the azimuthally averaged parameters.

### 5.2.2. Accuracy of the estimated uncertainties of the emission line properties

We repeat here the test of the hypothesis that the errors estimated by pyFIT3D could represent the real uncertainty in the recovery of each parameter (e.g., Sec. 5.1.3, Figure 8), this time for the emission line analysis. Like in the case of the stellar populations the comparison between the uncertainties and the errors should be done in an statistical way. In Appendix B we discuss the individual comparison between both parameters. Figure 11 shows the distribution of  $\Delta(\text{par})/\epsilon(\text{par})$  for the explored properties of the emission lines (flux, velocity dispersion and systemic velocity) based on the simulations described in the previous section (i.e., the equivalent to Figure 8, but for the emission line properties). We first notice that there is no significant bias (offset) between the recovered and simulated values for each of the three explored parameters, in agreement with the results shown in Figure 9. Due to that the observed distributions are well center in zero. Like in the case of the errors estimated for the stellar population properties we introduce a correction factor ( $f$ ), to match the observed distribution with the expected one (a Gaussian function centered in zero with a  $\sigma = 1$ ). As a result we appreciate that the estimated errors are

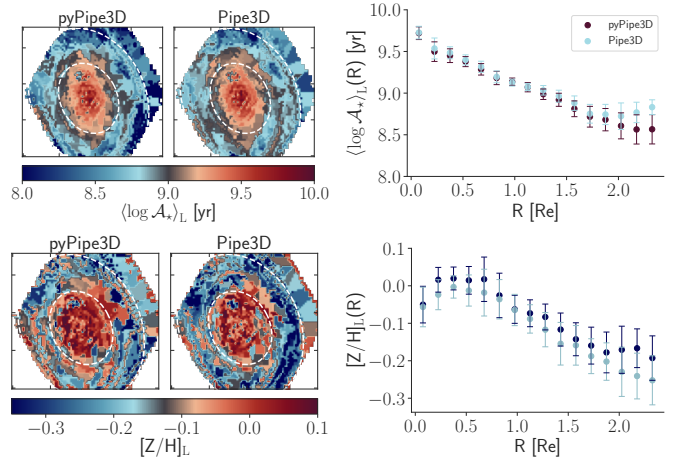


Figure 14: Similar comparison as the one shown in Figure 13 for the  $(\log \mathcal{A}_{\star})_L$  and  $[Z/H]_L$ , using the same nomenclature.

of the same order of the real uncertainties for the emission line fluxes ( $f \sim 1$ ), however, for the kinematics parameters there is a clear underestimation of the errors ( $f \sim 6-7$ ). The observed and expected distributions are very similar once applied this correction factor. Therefore the corrected errors can be used as good estimations of the real uncertainties.

Finally, we repeat this experiment using the two methods implemented in pyFIT3D to explore the emission lines, the **RND method**, already present in FIT3D, and the new **RND+LN method**. Figure 12 shows the distribution of  $\Delta(\text{par})/\epsilon(\text{par})$  for two different fitting rounds (one for each method) over a set of 10,000 simulated spectra including the emission line system described before ( $[\text{N II}]\lambda\lambda 6548, 6583 + \text{H}\alpha$ ). There is a remarkable good agreement between the distributions of the real and estimated errors for the second method, with the distribution being well centered in -0.04 dex with a standard deviation of 1.05. It is clear that the concomitant work of both methods produces a better result than using the **RND method** only. For this previous method the distribution presents a bias of -0.21 dex and narrower distribution ( $\sigma = 0.85$ ), what suggest that there is a bias in the recovered values and an overestimation of the errors. In summary, combining the **RND+LM methods** improves the precision and accuracy of the code.

## 6. pyPipe3D applied on real data

The experiments in the previous sections, using simulated data, with well controlled inputs allowed us to demonstrate that pyFIT3D yield reliable results, characterizing the uncertainties associated to the derived parameters. However, it does not exemplify the full applicability of pyPipe3D, what is only possible when using it over real IFS data. As a showcase of the use of the pipeline (Sec. 4.2) we reanalyze the V500-setup data for the galaxy NGC 2916, provided by the CALIFA survey. This is a an Sbc galaxy with  $\log(M_{\star}/M_{\odot}) \sim 10.8$  and redshift  $\sim 0.0122$  ( $v_{\star}^{\text{cen}} = 3669$  km/s), which comprises a wide range of stellar populations and emission line properties within its optical extension. For this reason was selected as the showcase example

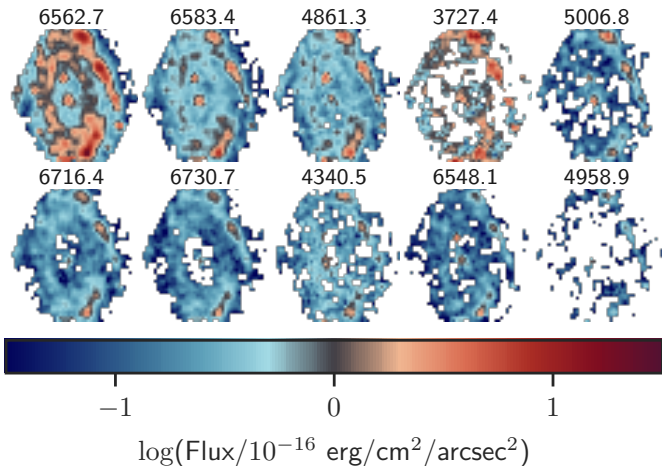


Figure 15: Flux intensity maps for the ten strongest emission lines observed in the galaxy NGC 2916, sorted by the mean flux. The spiral shape of the galaxy is very clearly traced by these emission line intensity maps. As the mean flux decreases, the noise is more evident as a patchy pattern in the images.

for the presentation of the previous version of the code (S16b). This way we will be able to compare the results provided by the two versions of the code. Considering that the previous version has been extensively used with several datasets, confronted with hydrodynamical simulations and compared with other fitting tool (Sec. 1), this comparison will allow us to estimate the quality of the analysis performed by the new version too.

Furthermore it will allow us to compare the speed of the two processes. Using the same CPU (Intel Xeon Gold 6130 (64) @ 2.101GHz), for a single-core, for a computer with 128Gb of RAM, the full analysis of the NGC 2916 datacube takes a total of 12602s (~3.5h) for pyPipe3D. The same process took 29859s (~8.3h) to Pipe3D. Thus, the new procedure, even including a second round in the analysis of the emission lines (i.e., the **RND+LM** method), is ~2.4 times faster. Similar numbers are derived when fitting a single spectrum (i.e., when running the pyFIT3D routine): using the same hardware configuration the new code requires 15s while the old one consumes 34s, thus, ~2.3 times faster.

### 6.1. Stellar population analysis

Figure 13 shows the spatial distributions and radial profiles (in units of the effective radius,  $R_e$ ) of the properties estimated during the non-linear analysis (Sec. 4.1), using both pyPipe3D and Pipe3D, for the considered dataset. The radial profiles were calculated up to 2.5  $R_e$ , adopting an azimuthal average of the properties inside elliptical concentric bins of 0.15  $R_e$  width. For practical reasons, we choose to show the velocity map instead of the redshift (i.e.,  $v_\star = cz_\star$ ), although both parameters are interchangeable. No spatial masking was applied to the data since our goal is to compare the output produced by pyPipe3D with that from Pipe3D regardless the quality of potential issues with the original data. For this reason the location of a foreground star at about 8'' North-East of the center of the galaxy is clearly

shown as a location of anomalous values in all maps. The left map is produced by first while the right one by latter.

In general, the new version of the code presents maps and profiles qualitatively similar to those of the previous one, and also with those previously published by other studies using different techniques (e.g. Cid Fernandes et al., 2013; Sánchez et al., 2016b). Regarding the velocity profile, the old distribution (and radial profile) presents a non realistic drop in the outer regions (above 1.5 $R_e$ ). Despite the fact that this region already corresponds to an area of low S/N, it is evident that this effect is not seen in the velocity maps provided by new version of the code. Thus, we consider that it provides with a better determination of the velocity in the outskirts of the galaxy. Furthermore, the spatial distribution of  $\sigma_\star$  is clearly smoother, more realistic, with a slightly steeper profile towards the center. However, the differences clearly seen along the map are not propagated to the radial profile, which is very similar for both versions of the code. The only difference is that the new version of the code provides somehow lower values of the velocity dispersion in the outer regions (from ~60 km/s to ~30km/s in the last radial bins). The main quantitative difference is found in the clear systematic offset in  $A_V^\star$ . For this parameter the values derived by pyPipe3D are in general ~1.2 larger than the ones derived by the old version of the code. This difference is most probably due to the different implementation of the currently adopted extinction curve. Pipe3D implemented our own coded algorithm that reproduces the polynomials published by Cardelli et al. (1989), while the new version uses the `EXTINCTION` module implemented in `PYTHON`. In addition we found an issue with the normalization wavelength at which the dust attenuation was derived for FIT3D in the re-coding of pyFIT3D. We consider that the two combined effects explain this difference.

The maps and profiles of the light-weighted average properties of the stellar populations, age and metallicity, are shown in Figure 14. NGC 2916 is an Sbc galaxy and it follows the same age and metallicity profiles reported for galaxies with the same morphology and stellar mass (Cid Fernandes et al., 2013; González Delgado et al., 2015; Mejía-Narváez et al., 2020; Sánchez, 2020). Both properties present negative gradients along the radius with the metallicity suffering a flattening or drop towards the inner regions ( $R < 0.5 R_e$ ). For these parameters the two versions of the code provide with very similar spatial and radial distributions of the explored parameters, despite of the mild/small differences reported in the non-linear parameters discussed before.

### 6.2. Emission lines analysis

For the emission lines properties pyPipe3D implements two different procedures, as discussed in Section 4. Figure 15 show the maps of the flux intensity for the ten strongest emission lines, ordered by the mean flux intensity of the map, calculated by the moment analysis (see 4.1.4). Since the analysis implemented in pyPipe3D is not restricted to a wavelength range by construction, this list of emission lines is part of a configuration file and can be altered by the final user. Actually, the pipeline code has been tested including the derivation of the parameters of more than 200 emission lines with no significant increase

over the total time of the analysis. As expected, the maps of the weakest emission lines present a noisy pattern due to low S/N, on top of the patchy structure intrinsic to emission line maps (due to the nature of the ionization, that it is not homogeneous across the optical extension of galaxies).

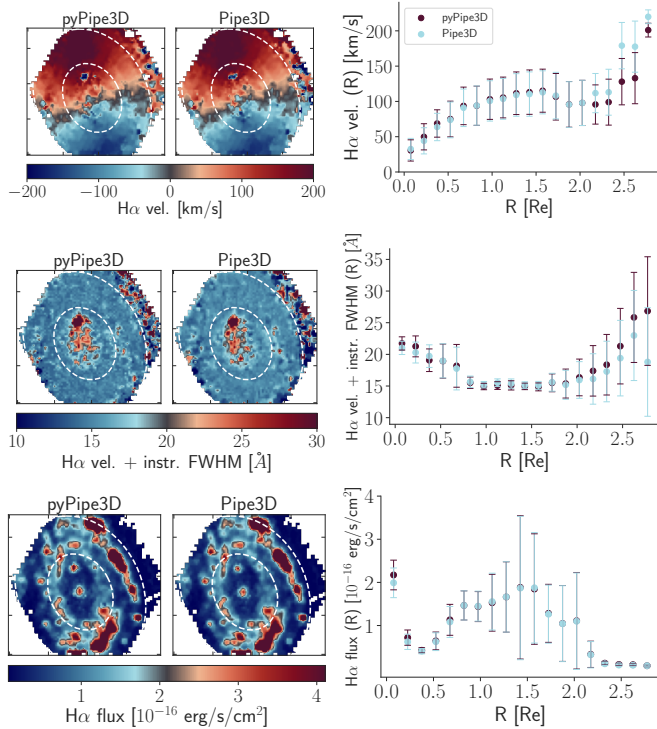


Figure 16: Comparison of the spatial distribution (left panels) and radial profiles (right panels) for the parameters derived by the moment analysis for the  $H\alpha$  emission line by pyPipe3D and Pipe3D for the NGC 2916 galaxy observed by CALIFA: velocity (1st row panels, in absolute values), velocity dispersion (2nd row panels), flux intensity (3rd row panels), and EW( $H\alpha$ ) (4rd row panels). White ellipses in the maps, symbols and error-bars in the radial plots follow the same scheme adopted in Figure 13.

Like in the case of the stellar populations we rely on the well proved abilities of the former version of the code, Pipe3D, to extract reliable properties for the emission lines. Therefore we compare the results provided by pyPipe3D with those provided by Pipe3D in order to estimate the quality of the derived quantities. Figure 16 shows the comparison between the spatial distribution and radial profiles of the velocity, velocity dispersion, flux intensity and EW for the  $H\alpha$  emission line derived by both codes using the moment analysis, following the same schemes adopted in Figure 13 and 14. Like in those cases no masking is applied to highlight the behavior of both codes even under non optical conditions (low-S/N, presence of foreground sources). There is a remarkable agreement, with even less differences than in the case of the properties of the stellar populations. In summary, the new code provides with a total consistent derivation of the properties of the emission lines for the explored procedure.

As mentioned before, pyPipe3D also derives the properties of the strongest emission lines adopting a Gaussian parametrization of the emission line, using the **RND+LM**

**method** (Sec. 4.1.2). Figure 17 shows the comparison between the parameters derived for the  $H\alpha$  emission line using both procedures, moment analysis and Gaussian parametrization, including the flux intensity, velocity and velocity dispersion (in this case, converted to FWHM). To increase the number statistics as much as possible we use the full dataset provided by the eCALIFA survey (867 galaxies Lacerda et al., 2020), with comprises more than 4 million analyzed spaxels/spectra in total. The same Figure was presented in S16b just of one galaxy (NGC 2916), using the former version of the code (i.e. Pipe3D). The conclusions can be directly extrapolated to the case of the entire survey analysis. The flux intensity is the parameter with presents the lowest offset between both methods of analysis ( $\Delta F_{H\alpha} = -0.04 \pm 0.55 \cdot 10^{-16} \text{ erg/s/cm}^2$ ). The kinematic parameters present a non-negligible bias ( $\Delta v_{H\alpha} = 46.4 \pm 59.8 \text{ km/s}$  and  $\Delta \sigma_{H\alpha} = -0.78 \pm 1.61 \text{ \AA}$ ). A priori it is expected that the weighted-moment algorithm works very well for emission lines which are not blended, while the Gaussian parametrization should recover the parameters better in those cases (if the S/N is high enough). We included in the last panel two lines delimiting the maximal dispersion where the  $H\alpha$  lines are not blended with [N II] $\lambda\lambda 6548, 6583$  doublet (corresponding to a FWHM of  $10 \text{ \AA}$ ). Selecting only those spaxels/spectra under this limit in the velocity dispersion, the offset and the dispersion for  $\Delta \sigma_{H\alpha}$  decreases significantly to a value of  $-0.38 \pm 0.65 \text{ \AA}$ .

## 7. Summary and conclusions

Along this article we have presented the new implementation of Pipe3D IFS analysis pipeline and FIT3D, the fitting tools adopted by the pipeline, fully transcribed from perl to python. The new pipeline conserves all the implemented tools and functionality from previous versions with a complete documentation available online<sup>8</sup>. Moreover, the entire project comprises a new documentation, with examples, of how to use it, to facilitate the final user to understand how the analysis is made and also to build pieces of analysis themselves.

This article is based on the assurance of the quality of the results already yielded by the Pipe3D pipeline over the last five years (e.g. Sánchez et al., 2016a; Cano-Díaz et al., 2016; Barrera-Ballesteros et al., 2016; Sánchez-Menguiano et al., 2017; Sánchez et al., 2018; López-Cobá et al., 2019; Sánchez, 2020). For this reason we maintained the main fitting sequence and the basic philosophy of the algorithms. We described the new coding philosophy and the main steps of the pyFIT3D fitting procedure, aimed to decouple the emission by stellar populations and the ionized gas and analyze both of them. We explain how this procedure is performed in three main steps: (i) the analysis of the non-linear parameters of the stellar population (i.e.,  $z_*$ ,  $\sigma_*$  and  $A_V^*$ ), (ii) the analysis and removal of the emission lines and (iii) the modeling of the stellar spectrum by a multi-SSP decomposition. In addition we describe further explorations that can be performed on both the stellar component (indices analysis) and the emission lines (moment analysis).

<sup>8</sup><http://ifs.astro.unam.mx/pyPipe3D>

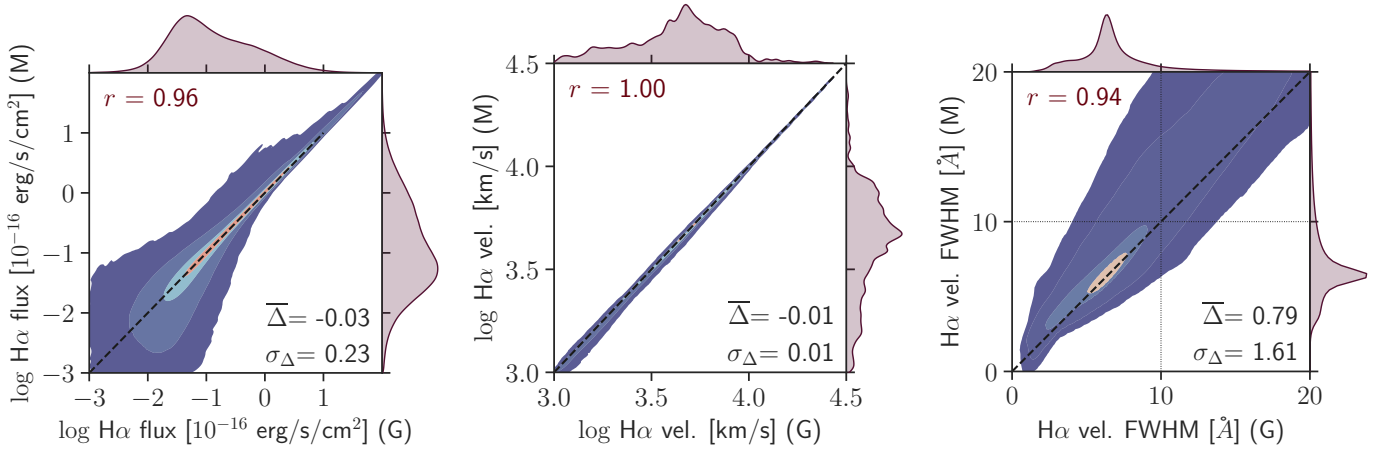


Figure 17: Comparison of the estimated parameters (flux intensity, velocity and velocity dispersion) of the  $H\alpha$  emission line derived by both methods, Gaussian models (G) versus weighted-moment analysis (M), implemented in `pyPipe3D`, for the full sample of eCALIFA 867 datacubes (Lacerda et al., 2020). Both dashed lines in the rightmost panel corresponds to a FWHM of  $10 \text{ \AA}$  a value below which the  $H\alpha$  emission line is not significantly blended with the  $[\text{N II}]\lambda\lambda 6548, 6583$  doublet (for the spectral resolution of this data). Contours and density histograms in each panel follow the same scheme presented on Figure 7.

All these different pieces of analysis were integrated into the final `pyPipe3D` pipeline, aimed to analyze automatically IFS observations on a single galaxy, extracting once more the properties of the stellar populations and the ionized gas. We described along this article the different procedures performed by this pipeline, which algorithms from `pyFIT3D` are adopted (highlighting the differences with respect to the previous version of the code), and described briefly the main dataproducs delivered. In order to determine the accuracy and precision in the derivation of the parameters, and to compare with previous results, we applied our new code to an extensive set of simulated spectra and to real data.

The main conclusions of this study are the following ones:

- The new version of the code implements all the algorithms and analysis tools of the previous version, respecting the input and output format of the data and products. However, the modules and tools under which the entire package is built are designed to be reused, facilitating also the work under exploration environments such as IPython (Pérez and Granger, 2007). Furthermore, taking advantage of the reusability of the code, the main scripts of `pyFIT3D` and `pyPipe3D` are distributed with new implementations, accepting new input arguments and output logs.
- Although the fitting philosophy adopted by `pyFIT3D` has not change, the code runs significantly faster allowing us to include a second round on the determination of the properties of the emission lines `pyFIT3D`. In addition, this change in the implementation of the code improves the accuracy and the precision of the estimated parameters and uncertainties, as by simulations.
- In general the code present a good accuracy and precision in the recovery of the explored parameters. The best recovered stellar parameters are the redshift ( $\Delta v \sim 10 \text{ km/s}$ ), and the worst ones are the velocity dispersion ( $\Delta\sigma_{\star} \sim 22$

km/s) and the metallicity ( $\Delta[Z/H]_{\text{L}} [0.6 \text{ dex}]$ ). The remaining stellar parameters are recovered with a precision better than 0.05 dex, and an even better accuracy.

- Our results confirm that the derivation of the different stellar parameters is not fully independent one-each-other. The strongest correlations are found between the uncertainties of the age and those of the dust, M/L and metallicity, and between the metallicity and the velocity dispersion, but not among themselves. Of them, the only clearly strong correlations are two first ones (age-dust and age-M/L,  $r > 0.5$ ), with the remaining ones present a much lower correlation ( $r \sim 0.3$ ). These correlations are inherent to the method, emerging from the intrinsic uncertainties of the stellar population models. These results are qualitatively and quantitatively similar to those found in the literature.
- The uncertainties estimated by `pyFIT3D` for the properties of the stellar populations are of the same order of the real errors, with mild underestimation. However, there is no one-to-one correspondence between the estimated and real uncertainties.
- Regarding the emission lines, we found that the new implemented **RND+LM** method recovers the simulated parameters with a much better precision ( $\sim 1\%$ ) and accuracy ( $\sim 3\text{-}9\%$ ) than the previously implemented **RND**, when the  $S/N > 3$ . Furthermore, we have not found any significant interdependence in the determination of the different parameters. Finally, the estimated errors are of the order of the real ones, with a slight overestimation of a  $\sim 5\%$  in the error of the flux, and a  $1\sigma$  agreement for the errors in the kinematics parameters.
- The comparison of the results derived from the current version of the code and those obtained with its previous implementation on real data shows that there is a qualitative

and quantitative agreement in the derived parameters for both the stellar populations and the emission lines. The main difference is found in the derivation of  $A_V^*$ , what we attribute to a bug discovered in the normalization wavelength in the previous version.

In summary, the new implementation of the pipeline could reproduce any of the studies already performed along the years using the previous version of the code and it is prepared for the analysis of IFS data from any of the present surveys and for those planned to be released over the next years. Since the new code is substantially faster, studies that require explorations using different input libraries, larger libraries or just over larger datasets are now feasible. Future upgrades already under-development will include (i) the modifications required to explore of the  $[\alpha/\text{Fe}]$ , (ii) a differential treatment of the dust attenuation depending on the properties of the stellar populations, (iii) a spectro-kinematic decomposition of the different stellar component, and (iv) the inclusion of more complex models for the emission lines (e.g., Voigt models or models including asymmetries in the profiles).

## Acknowledgements

We acknowledge the support of the PAPIIT-DGAPA IN100519 and IG100622 projects. J.B-B acknowledges support from the grant IA101522 (DGAPA-PAPIIT, UNAM) and funding from the CONACYT grant CF19-39578. EADL and SFS thank to Laura Sánchez-Menguiano for her valuable commentaries and time spent analyzing the products of beta versions of pyPipe3D and pyFIT3D improving the present versions. EADL thanks to Carlos Henrique Mantovani Sammartin de Abreu for the art design of the pyPipe3D logo. This study uses data provided by the extended Calar Alto Legacy Integral Field Area (eCALIFA) survey (<http://califa.caha.es/>). This research has not been possible without the support from the UNAM and the DGAPA Post Doctoral Scholarship Program.

## Appendix A. FIT3D format for the SSP templates

As indicated before pyFIT3D (and pyPipe3D) can use, in principle, any stellar or SSP library to fit the stellar component of the analyzed spectra. However, it requires that these library adopts a certain format. The actual format is a row-stacked spectra (RSS) FITS file, in which each row corresponds to one of the components to be fitted (i.e., each  $j$  component  $f_{\lambda,j}^{\text{SSP}}$  in Equations 1 and 5). So far the code is prepared to analyze spectra which wavelength solution is linear, i.e., each spectral pixel (X-axis of the RSS file) corresponds to a certain wavelength with a fixed wavelength interval in Å between adjacent pixels. It also expects that the SSP templates are in this format. Therefore, the SSP FITS file should include in the header the reference pixel at which the starting wavelength is defined (CRPIX1), the value of this wavelength in Å (CRVAL1), and the constant step in wavelength in the spectral axis (CDELTA1). The SSP templates should be normalized

to one at a certain wavelength within the interval covered by the spectra (WAVENORM). The program requires the flux in solar units at this wavelength to be stored in the header with an entry for each individual SSP (NORM $j$ ), in order to provide with a reliable  $Y_V$  and stellar mass. In addition, to provide with reliable estimations of the  $\langle \log \mathcal{A}_\star \rangle_L$  and  $[Z/H]_L$  (and  $\langle \log \mathcal{A}_\star \rangle_M$  and  $[Z/H]_M$ , too), it is required an entry in the header for each individual SSP (NAME $j$ ) following the format `spec_ssp_AGEGyr_zMET.spec`, where **AGE** corresponds to the age in Gyr of the SSP and **MET** corresponds to the decimals of the metallicity (i.e.,  $Z = 0.1\text{MET}$ ). For instance `spec_ssp_06.0000Gyr_z0021.spec` corresponds to an SSP with an age of 6 Gyr and a metallicity of 0.0021<sup>9</sup>

## Appendix B. Relation between errors and uncertainties

Figure B.18 shows the distribution of the estimated errors for the stellar populations,  $\epsilon(\text{par})$ , as a function of the absolute value of the uncertainty  $|\Delta(\text{par})|$ , for properties discussed in Section 5.1. Despite of the good statistical correspondence between both parameters once applied the correction factor discussed in Sec. 5.1.3, there is no significant correlation between both parameters. Only the errors for the redshift and  $\langle \log \mathcal{A}_\star \rangle_L$  present some degree of correlation.

In many cases the distribution is near or on-top of the one-to-one relation ( $A_V^*$ ,  $\langle \log \mathcal{A}_\star \rangle_L$ ,  $[Z/H]_L$  and  $Y_V$ ), what indicates that the average errors estimated by pyFIT3D traced the real ones. However, already discussed in Sec. 5.1.3, and shown in Figure 11 the errors are clearly underestimated for different properties. In the case of the kinematics ones both kinematic, the underestimation of is considerably larger than in the case of the velocity dispersion. In addition, the redshift presents a tiny bimodality as a consequence of the cases when it does not find a better solution than the input *guess*. In these cases pyFIT3D assumes the *guess* as the redshift value and the input *interval* as the error. Those cases correspond to less than a 1% of the total number of simulations. The same happens with the velocity dispersion but the substantial scatter on  $\epsilon(\sigma_\star)$  does not allow us to visualize it in the Figure B.18. On the other hand, for the dust attenuation, the center of the  $\epsilon \times |\Delta|$  distribution shows a noteworthy relation. However, the estimated errors cover a wide range of values, and therefore, as indicated before, there is no significant correlation between both parameters. The errors of  $\langle \log \mathcal{A}_\star \rangle_L$  present a clear underestimation, covering a narrower range of values than the real ones. However, as we described before, this is the parameter for which the estimated errors present the strongest correlation with the real uncertainties. This is not the case for the  $[Z/H]_L$  and  $Y_V$ , where the estimated errors cover a much narrower range of parameters than the real ones. In both cases there is a clear over-estimation of the errors. In summary, there is no direct correspondence between the real uncertainties and the errors estimated by the code for each individual simulation, and the agreement between both parameters is valid on an statistical basis, as discussed in Sec. 5.1.3.

<sup>9</sup>A few examples of the format have been included in the following webpage: <http://ifs.astrosu.unam.mx/pyPipe3D/templates/>

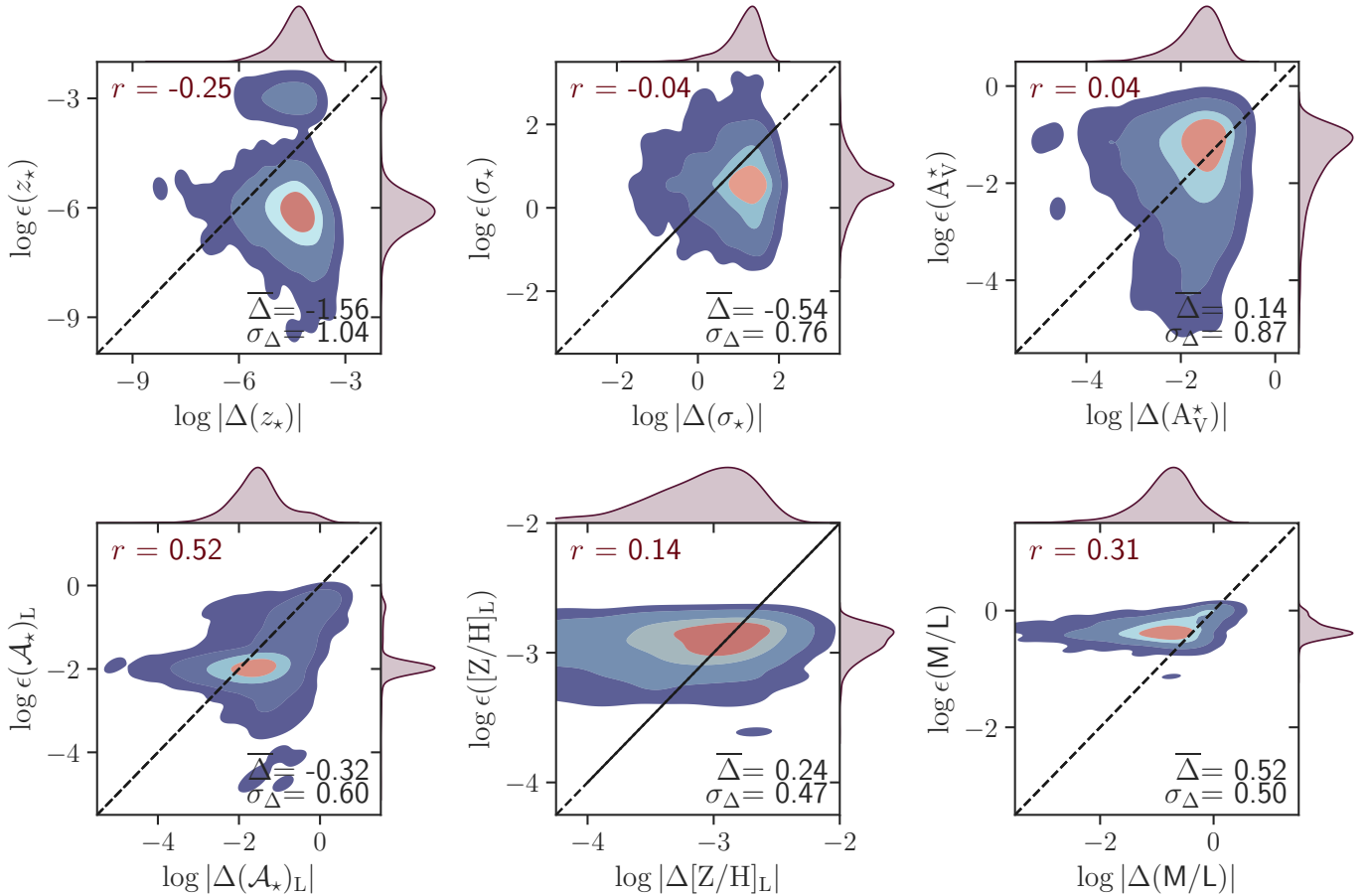


Figure B.18: Distribution of the estimated error,  $\epsilon_{\text{par}}$ , versus the absolute value of the difference between the estimated and input value,  $\Delta_{\text{par}}$ , for each of the stellar parameters shown in Figure 6. Contours and density histograms in each panel follow the same scheme presented on Figure 7. Each panel includes the correlation coefficient between the represented parameters (up-left legend), together with the average and standard deviation of the difference between them (bottom-right legend).

Figure B.19 shows the distribution of the relative value of the estimated errors,  $\epsilon(\text{par})/\text{par}_{\text{in}}$ , as a function of this tracer of the uncertainties, i.e.  $|\Delta_{\text{par}}/\text{par}_{\text{in}}|$ , for the simulated emission lines discussed in Sec. 5.2. As already discussed in Sec. 5.2.2 the three parameters present a mild underestimation of the errors, what results in  $\epsilon(\text{par})/\text{par}_{\text{in}}$  covering a much narrower range of values than  $|\Delta_{\text{par}}/\text{par}_{\text{in}}|$ . As expected, there is no clear correlation between the uncertainties and errors ( $r < 0.3$  in all cases), as both parameters should be compared only in an statistical sense.

### Appendix C. Parameters adopted for the showcase fitting example

For the description of the fitting procedure of pyFIT3D in Section 4.1 we analyze the central spectrum (5" aperture) of the galaxy NGC 5947, extracted from the data provided by the CALIFA V500 IFS datacube as a showcase example. The chosen input values for the non-linear parameters are (guess/step/min/max):

- $z_*$ : 0.01/0.001/0.003/0.05;

- $\sigma_*$ : 30/10/75/375 km/s;

- $A_V^*$ : 0.15/0.05/0/1.8 mag.

All domains are chosen covering up the real observed intervals for CALIFA data. The kinematic stellar parameters are estimated using the wavelength interval from 3800 to 4700 Å because this spectral region is dominated by stronger stellar absorption features. All subsequent analysis use the wavelength interval from 3800 to 7000 Å.

For the emission lines analysis, we choose to fit five different *systems* with ten different emission lines: H $\delta$ ; H $\gamma$ ; [O III] $\lambda\lambda$ 4958,5007 + H $\beta$ ; [N II] $\lambda\lambda$ 6548,6583 + H $\alpha$  and [S II] $\lambda\lambda$ 6731,6716, with the procedure fitting one *system* at a time. The kinematic is selected to be the same for all emission lines, assuming a physical scenery where they are formed at same conditions. We should note that this is not a limitation of the method, just the choice done for the current analysis. For the [O III] $\lambda\lambda$ 4958,5007 + H $\beta$  and emission line *system* we tie the integrated flux of each [O III] $\lambda\lambda$ 4958,5007 line as  $F_{[\text{O III}]\lambda 5007} = (1/3)F_{[\text{O III}]\lambda 4959}$ . Following the same hypothesis, for the [N II] $\lambda\lambda$ 6548,6583 + H $\alpha$  *system* we link the integrated flux of the [N II] $\lambda\lambda$ 6548,6583 line as  $F_{[\text{N II}]\lambda 6584} = (1/3)F_{[\text{N II}]\lambda 6548}$ .



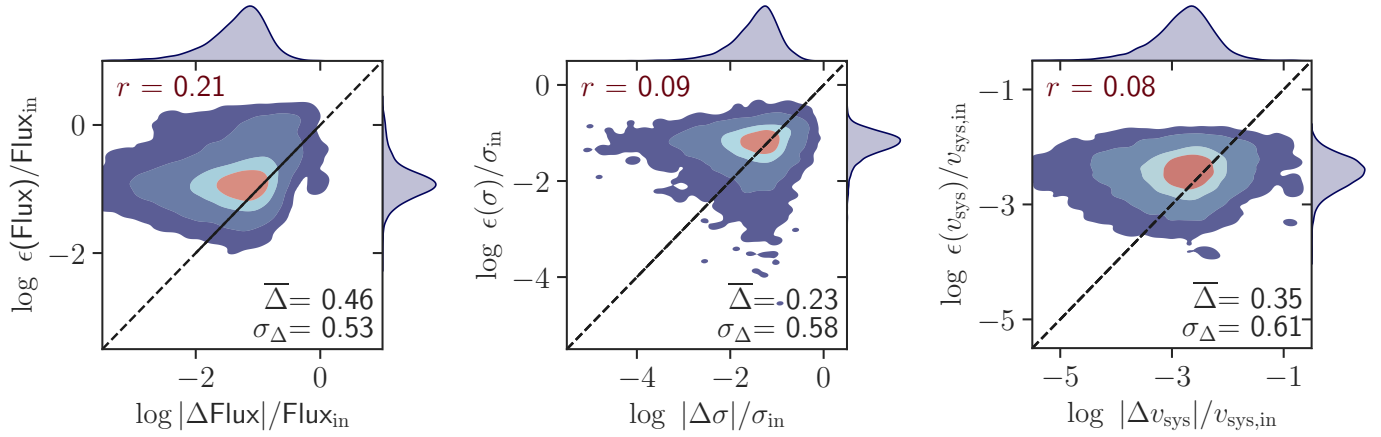


Figure B.19: Distribution of the errors estimated by pyFIT3D as a function of the absolute value of the relative difference between the derived and input values for each of the explored parameters of the simulated emission lines: flux intensity (left), velocity dispersion (middle) and systemic velocity (right). Contours and density histograms in each panel follow the same scheme presented on Figure 7.

An example of a simple configuration file of an emission lines *system* can be found in the pyPipe3D webpage. The *RND method* is configured to proceed 20 MC loops for the search of the kinematic parameters and the integrated flux. This procedure is repeated five times, with the parameters intervals narrowed around the best fitted parameters, as described in Sec. 4.1.2.

The adopted SSP template library was the *gsd156*, precisely described in (Cid Fernandes et al., 2013). It includes templates extracted from two projects: (a) synthetic stellar spectra from the GRANADA library (Martins et al., 2005) for the ages < 65 Myr and (b) the SSP library of MILES (Sánchez-Blázquez et al., 2006; Vazdekis et al., 2010; Falcón-Barroso et al., 2011). The full spectral library is composed of 39 stellar ages (from 1 Myr to 13 Gyr) and 4 metallicities ( $Z/Z_{\odot} = 0.2, 0.4, 1$  and  $1.5$ ). Again, this is not a limitation of the procedure, and indeed we have tested the code using several different SSP libraries. The results of that comparison will be presented elsewhere. For the non-linear fit of the showcase example we use a limited version of the *gsd156* library including only 12 SSPs (*gsd12*: 3 ages,  $t_{\star} = 0.001, 0.5$  and  $14.1$  Gyr; and the same 4 metallicities) for the reasons mentioned in Sec. 4.1.1. The actual configuration file required to run the pyFIT3D script on this spectrum can be found again in the pyPipe3D webpage.

## References

Barrera-Ballesteros, J.K., Heckman, T.M., Zhu, G.B., Zakamska, N.L., Sánchez, S.F., Law, D., Wake, D., Green, J.E., Bizyaev, D., Oravetz, D., Simmons, A., Malanushenko, E., Pan, K., Roman Lopes, A., Lane, R.R., 2016. Do galaxy global relationships emerge from local ones? The SDSS IV MaNGA surface mass density-metallicity relation. *Mon. Not. R. Astron. Soc.* 463, 2513–2522. doi:10.1093/mnras/stw1984, arXiv:1609.01740.

Belfiore, F., Westfall, K.B., Schaefer, A., Cappellari, M., Ji, X., Bershady, M.A., Tremonti, C., Law, D.R., Yan, R., Bundy, K., Shetty, S., Drory,

N., Thomas, D., Emsellem, E., Sánchez, S.F., 2019. The Data Analysis Pipeline for the SDSS-IV MaNGA IFU Galaxy Survey: Emission-line Modeling. *Astron. J.* 158, 160. doi:10.3847/1538-3881/ab3e4e, arXiv:1901.00866.

Bender, R., 1990. Unraveling the kinematics of early-type galaxies. Presentation of a new method and its application to NGC 4621. *Astronom. Astrophys.* 229, 441–451.

Bittner, A., Falcón-Barroso, J., Nedelchev, B., Dorta, A., Gadotti, D.A., Sarzi, M., Molaiezhad, A., Iodice, E., Rosado-Belza, D., de Lorenzo-Cáceres, A., Fragkoudi, F., Galán-de Anta, P.M., Husemann, B., Méndez-Abreu, J., Neumann, J., Pinna, F., Querejeta, M., Sánchez-Blázquez, P., Seidel, M.K., 2019. The GIST pipeline: A multi-purpose tool for the analysis and visualisation of (integral-field) spectroscopic data. *Astronom. Astrophys.* 628, A117. doi:10.1051/0004-6361/201935829, arXiv:1906.04746.

Blanton, M.R., Bershady, M.A., Abolfathi, B., Albareti, F.D., Allende Prieto, C., Almeida, A., Alonso-García, J., Anders, F., Anderson, S.F., Andrews, B., Aquino-Ortiz, E., Aragón-Salamanca, A., Argudo-Fernández, M., Armengaud, E., Aubourg, E., Avila-Reese, V., Badenes, C., Bailey, S., Barger, K.A., Barrera-Ballesteros, J., Bartosz, C., Bates, D., Baumgarten, F., Bautista, J., Beaton, R., Beers, T.C., Belfiore, F., Bender, C.F., Berlind, A.A., Bernardi, M., Beutler, F., Bird, J.C., Bizyaev, D., Blanc, G.A., Blomqvist, M., Bolton, A.S., Boquien, M., Borrissova, J., van den Bosch, R., Bovy, J., Brandt, W.N., Brinkmann, J., Brownstein, J.R., Bundy, K., Burgasser, A.J., Burtin, E., Busca, N.G., Cappellari, M., Delgado Carigi, M.L., Carlberg, J.K., Carnero Rosell, A., Carrera, R., Chanover, N.J., Cherinka, B., Cheung, E., Gómez Maqueo Chew, Y., Chiappini, C., Choi, P.D., Chojnowski, D., Chuang, C.H., Chung, H., Cirolini, R.F., Clerc, N., Cohen, R.E., Comparat, J., da Costa, L., Cousinou, M.C., Covey, K., Crane, J.D., Croft, R.A.C., Cruz-Gonzalez, I., Garrido Cuadra, D., Cunha, K., Damke, G.J., Darling, J., Davies, R., Dawson, K., de la Macorra, A., Dell’Agli, F., De Lee, N., Delubac, T., Di Mille, F., Diamond-Stanic, A., Cano-Díaz, M., Donor, J., Downes, J.J., Drory, N., du Mas des Bourboux, H., Duckworth, C.J., Dwelly, T., Dyer, J., Ebelke, G., Eigenbrot, A.D., Eisenstein, D.J., Emsellem, E., Eracleous, M., Escoffier, S., Evans, M.L., Fan, X., Fernández-Alvar, E., Fernández-Trincado, J.G., Feuillet, D.K., Finoguenov, A., Fleming, S.W., Font-Ribera, A., Fredrickson, A., Freisclad, G., Frinchaboy, P.M., Fuentes, C.E., Galbany, L., García-Díaz, R., García-Hernández, D.A., Gaulme, P., Geisler, D., Gelfand, J.D., Gil-Marín, H., Gillespie, B.A., Goddard, D., Gonzalez-Perez, V., Grabowski, K., Green, P.J., Grier, C.J., Gunn, J.E., Guo, H., Guy, J., Hagen, A., Hahn, C., Hall, M., Harding, P., Hasselquist, S., Hawley, S.L., Hearty, F., Gonzalez Hernández, J.I., Ho, S., Hogg, D.W., Holley-Bockelmann, K., Holtzman, J.A., Holzer, P.H., Huehnerhoff, J., Hutchinson, T.A., Hwang, H.S., Ibarra-Medel, H.J., da Silva Ilha, G., Ivans, I.I., Ivory, K., Jackson, K., Jensen, T.W., Johnson, J.A., Jones, A., Jönsson, H., Jullo, E., Kamble, V., Kinemuchi, K., Kirkby, D., Kitaura, F.S., Klaene, M., Knapp, G.R., Kneib, J.P., Kollmeier, J.A., Lacerna, I., Lane, R.R., Lang, D., Law, D.R., Lazarz, D., Lee, Y., Le Goff, J.M., Liang,

<sup>9</sup>[http://ifs.astroscu.unam.mx/pyPipe3D/pipeline\\_example\\_README.html](http://ifs.astroscu.unam.mx/pyPipe3D/pipeline_example_README.html)

- F.H., Li, C., Li, H., Lian, J., Lima, M., Lin, L., Lin, Y.T., Bertran de Lis, S., Liu, C., de Icaza Lizaola, M.A.C., Long, D., Lucatello, S., Lundgren, B., MacDonald, N.K., Deconto Machado, A., MacLeod, C.L., Mahadevan, S., Geimba Maia, M.A., Maiolino, R., Majewski, S.R., Malanushenko, E., Malanushenko, V., Machado, A., Mao, S., Maraston, C., Marques-Chaves, R., Masseron, T., Masters, K.L., McBride, C.K., McDermid, R.M., McGrath, B., McGreer, I.D., Medina Peña, N., Melendez, M., Merloni, A., Merrifield, M.R., Meszaros, S., Meza, A., Minchev, I., Minniti, D., Miyaji, T., More, S., Mulchaey, J., Müller-Sánchez, F., Muna, D., Munoz, R.R., Myers, A.D., Nair, P., Nandra, K., Correa do Nascimento, J., Negrete, A., Ness, M., Newman, J.A., Nichol, R.C., Nidever, D.L., Nitschelm, C., Ntelis, P., O'Connell, J.E., Oelkers, R.J., Oravetz, A., Oravetz, D., Pace, Z., Padilla, N., Palanque-Delabrouille, N., Alonso Palicio, P., Pan, K., Parejko, J.K., Parikh, T., Pâris, I., Park, C., Patten, A.Y., Peirani, S., Pellejero-Ibanez, M., Penny, S., Percival, W.J., Perez-Fournon, I., Petitjean, P., Pieri, M.M., Pinsonneault, M., Pisani, A., Poleski, R., Prada, F., Prakash, A., Queiroz, A.B.d.A., Raddick, M.J., Raichoor, A., Barboza Rembold, S., Richstein, H., Riffel, R.A., Riffel, R., Rix, H.W., Robin, A.C., Rockosi, C.M., Rodríguez-Torres, S., Roman-Lopes, A., Román-Zúñiga, C., Rosado, M., Ross, A.J., Rossi, G., Ruan, J., Ruggeri, R., Rykoff, E.S., Salazar-Albornoz, S., Salvato, M., Sánchez, A.G., Aguado, D.S., Sánchez-Gallego, J.R., Santana, F.A., Santiago, B.X., Sayres, C., Schiavon, R.P., da Silva Schimoia, J., Schlafly, E.F., Schlegel, D.J., Schneider, D.P., Schultheis, M., Schuster, W.J., Schwobe, A., Seo, H.J., Shao, Z., Shen, S., Shetrone, M., Shull, M., Simon, J.D., Skinner, D., Skrutskie, M.F., Slosar, A., Smith, V.V., Sobek, J.S., Sobreira, F., Somers, G., Souto, D., Stark, D.V., Stassun, K., Stauffer, F., Steinmetz, M., Storch-Bergmann, T., Streblynska, A., Stringfellow, G.S., Suárez, G., Sun, J., Suzuki, N., Szegedi, L., Taghizadeh-Popp, M., Tang, B., Tao, C., Tayar, J., Tembe, M., Teske, J., Thakar, A.R., Thomas, D., Thompson, B.A., Tinker, J.L., Tissera, P., Tojeiro, R., Hernandez Toledo, H., de la Torre, S., Tremonti, C., Troup, N.W., Valenzuela, O., Martinez Valpuesta, I., Vargas-González, J., Vargas-Magaña, M., Vazquez, J.A., Villanova, S., Vivek, M., Vogt, N., Wake, D., Walterbos, R., Wang, Y., Weaver, B.A., Weijmans, A.M., Weinberg, D.H., Westfall, K.B., Whelan, D.G., Wild, V., Wilson, J., Wood-Vasey, W.M., Wylezalek, D., Xiao, T., Yan, R., Yang, M., Ybarra, J.E., Yèche, C., Zakamska, N., Zamora, O., Zarrouk, P., Zasowski, G., Zhang, K., Zhao, G.B., Zheng, Z., Zheng, Z., Zhou, X., Zhou, Z.M., Zhu, G.B., Zoccali, M., Zou, H., 2017. Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe. *Astron. J.* 154, 28. doi:10.3847/1538-3881/aa7567, arXiv:1703.00052.
- Bruzual, A. G., 1983. Spectral evolution of galaxies. I. Early-type systems. *Astrophys. J.* 273, 105–127. doi:10.1086/161352.
- Burstein, D., Faber, S.M., Gaskell, C.M., Krumm, N., 1984. Old stellar populations. I. A spectroscopic comparison of galactic globular clusters, M 31 globular clusters, and elliptical galaxies. *Astrophys. J.* 287, 586–609. doi:10.1086/162718.
- Cano-Díaz, M., Sánchez, S.F., Zibetti, S., Ascasibar, Y., Bland-Hawthorn, J., Ziegler, B., González Delgado, R.M., Walcher, C.J., García-Benito, R., Mast, D., Mendoza-Pérez, M.A., Falcón-Barroso, J., Galbany, L., Husemann, B., Kehrig, C., Marino, R.A., Sánchez-Blázquez, P., López-Cobá, C., López-Sánchez, Á.R., Vilchez, J.M., 2016. Spatially Resolved Star Formation Main Sequence of Galaxies in the CALIFA Survey. *Astrophys. J. Lett.* 821, L26. doi:10.3847/2041-8205/821/2/L26, arXiv:1602.02770.
- Cappellari, M., Copin, Y., 2003. Adaptive spatial binning of integral-field spectroscopic data using Voronoi tessellations. *Mon. Not. R. Astron. Soc.* 342, 345–354. doi:10.1046/j.1365-8711.2003.06541.x, arXiv:astro-ph/0302262.
- Cappellari, M., Emsellem, E., 2004. Parametric Recovery of Line-of-Sight Velocity Distributions from Absorption-Line Spectra of Galaxies via Penalized Likelihood. *Publ. Astron. Soc. Pac.* 116, 138–147. doi:10.1086/381875, arXiv:astro-ph/0312201.
- Cappellari, M., Emsellem, E., Krajnović, D., McDermid, R.M., Scott, N., Verdoes Kleijn, G.A., Young, L.M., Alatalo, K., Bacon, R., Blitz, L., Bois, M., Bournaud, F., Bureau, M., Davies, R.L., Davis, T.A., de Zeeuw, P.T., Duc, P.A., Khochfar, S., Kuntschner, H., Lablanche, P.Y., Morganti, R., Naab, T., Oosterloo, T., Sarzi, M., Serra, P., Weijmans, A.M., 2011. The ATLAS<sup>3D</sup> project - I. A volume-limited sample of 260 nearby early-type galaxies: science goals and selection criteria. *Mon. Not. R. Astron. Soc.* 413, 813–836. doi:10.1111/j.1365-2966.2010.18174.x, arXiv:1012.1551.
- Cardelli, J.A., Clayton, G.C., Mathis, J.S., 1989. The relationship between infrared, optical, and ultraviolet extinction. *Astrophys. J.* 345, 245–256. doi:10.1086/167900.
- Cardiel, N., Gorgas, J., Sánchez-Blázquez, P., Cenarro, A.J., Pedraz, S., Bruzual, G., Klement, J., 2003. Using spectroscopic data to disentangle stellar population properties. *Astronom. Astrophys.* 409, 511–522. doi:10.1051/0004-6361:20031096, arXiv:astro-ph/0306560.
- Cerviño, M., Luridiana, V., 2004. Physical limits to the validity of synthesis models. The Lowest Luminosity Limit. *Astronom. Astrophys.* 413, 145–157. doi:10.1051/0004-6361:20031454, arXiv:astro-ph/0304061.
- Cid Fernandes, R., González Delgado, R.M., García Benito, R., Pérez, E., de Amorim, A.L., Sánchez, S.F., Husemann, B., Falcón Barroso, J., López-Fernández, R., Sánchez-Blázquez, P., Vale Asari, N., Vazdekis, A., Walcher, C.J., Mast, D., 2014. Resolving galaxies in time and space. II. Uncertainties in the spectral synthesis of datacubes. *Astronom. Astrophys.* 561, A130. doi:10.1051/0004-6361/201321692, arXiv:1307.0562.
- Cid Fernandes, R., Mateus, A., Sodr , L., Stasińska, G., Gomes, J.M., 2005. Semi-empirical analysis of Sloan Digital Sky Survey galaxies - I. Spectral synthesis method. *Mon. Not. R. Astron. Soc.* 358, 363–378. doi:10.1111/j.1365-2966.2005.08752.x, arXiv:astro-ph/0412481.
- Cid Fernandes, R., Pérez, E., García Benito, R., González Delgado, R.M., de Amorim, A.L., Sánchez, S.F., Husemann, B., Falcón Barroso, J., Sánchez-Blázquez, P., Walcher, C.J., Mast, D., 2013. Resolving galaxies in time and space. I. Applying STARLIGHT to CALIFA datacubes. *Astronom. Astrophys.* 557, A86. doi:10.1051/0004-6361/201220616, arXiv:1304.5788.
- Cid Fernandes, R., Stasińska, G., Schlickmann, M.S., Mateus, A., Vale Asari, N., Schoenell, W., Sodr , L., 2010. Alternative diagnostic diagrams and the ‘forgotten’ population of weak line galaxies in the SDSS. *Mon. Not. R. Astron. Soc.* 403, 1036–1053. doi:10.1111/j.1365-2966.2009.16185.x, arXiv:0912.1643.
- Croom, S.M., Lawrence, J.S., Bland-Hawthorn, J., Bryant, J.J., Fogarty, L., Richards, S., Goodwin, M., Farrell, T., Miziarski, S., Heald, R., Jones, D.H., Lee, S., Colless, M., Brough, S., Hopkins, A.M., Bauer, A.E., Birchall, M.N., Ellis, S., Horton, A., Leon-Saval, S., Lewis, G., López-Sánchez, Á.R., Min, S.S., Trinh, C., Trowland, H., 2012. The Sydney-AAO Multi-object Integral field spectrograph. *Mon. Not. R. Astron. Soc.* 421, 872–893. doi:10.1111/j.1365-2966.2011.20365.x, arXiv:1112.3367.
- Faber, S.M., Friel, E.D., Burstein, D., Gaskell, C.M., 1985. Old stellar populations. II. an analysis of K-giant spectra. *Astrophys. J. Suppl. S* 57, 711–741. doi:10.1086/191024.
- Falcón-Barroso, J., Martig, M., 2021. BAYES-LOSVD: A Bayesian framework for non-parametric extraction of the line-of-sight velocity distribution of galaxies. *Astronom. Astrophys.* 646, A31. doi:10.1051/0004-6361/202039624, arXiv:2011.12023.
- Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., Ricciardelli, E., Cardiel, N., Cenarro, A.J., Gorgas, J., Peletier, R.F., 2011. An updated MILES stellar library and stellar population models. *Astronom. Astrophys.* 532, A95. doi:10.1051/0004-6361/201116842, arXiv:1107.2303.
- Folkes, S., Ronen, S., Price, I., Lahav, O., Colless, M., Maddox, S., Deeley, K., Glazebrook, K., Bland-Hawthorn, J., Cannon, R., Cole, S., Collins, C., Couch, W., Driver, S.P., Dalton, G., Efstathiou, G., Ellis, R.S., Frenk, C.S., Kaiser, N., Lewis, I., Lumsden, S., Peacock, J., Peterson, B.A., Sutherland, W., Taylor, K., 1999. The 2dF Galaxy Redshift Survey: spectral types and luminosity functions. *Mon. Not. R. Astron. Soc.* 308, 459–472. doi:10.1046/j.1365-8711.1999.02721.x, arXiv:astro-ph/9903456.
- Galbany, L., Anderson, J.P., Rosales-Ortega, F.F., Kuncarayakti, H., Krühler, T., Sánchez, S.F., Falcón-Barroso, J., Pérez, E., Maureira, J.C., Hamuy, M., González-Gaitán, S., Förster, F., Moral, V., 2016. Characterizing the environments of supernovae with MUSE. *Mon. Not. R. Astron. Soc.* 455, 4087–4099. doi:10.1093/mnras/stv2620, arXiv:1511.01495.
- Galbany, L., Anderson, J.P., Sánchez, S.F., Kuncarayakti, H., Pedraz, S., González-Gaitán, S., Stanishev, V., Domínguez, I., Moreno-Raya, M.E., Wood-Vasey, W.M., Mourão, A.M., Ponder, K.A., Badenes, C., Mollá, M., López-Sánchez, A.R., Rosales-Ortega, F.F., Vilchez, J.M., García-Benito, R., Marino, R.A., 2018. PISCO: The PMAS/PPak Integral-field Supernova Hosts Compilation. *Astrophys. J.* 855, 107. doi:10.3847/1538-4357/aaaf20, arXiv:1802.01589.
- Glazebrook, M., Economou, F., 1997. PDL: The Perl Data Language. *Perl Journal* 5. doi:10.3847/1538-4365/aacb24, arXiv:1997.0001.
- González Delgado, R.M., García-Benito, R., Pérez, E., Cid Fernandes, R., de Amorim, A.L., Cortijo-Ferrero, C., Lacerda, E.A.D., López Fernández, R., Vale-Asari, N., Sánchez, S.F., Mollá, M., Ruiz-Lara, T., Sánchez-Blázquez,

- P., Walcher, C.J., Alves, J., Aguerri, J.A.L., Bekeraité, S., Bland-Hawthorn, J., Galbany, L., Gallazzi, A., Husemann, B., Iglesias-Páramo, J., Kalinova, V., López-Sánchez, A.R., Marino, R.A., Márquez, I., Masegosa, J., Mast, D., Méndez-Abreu, J., Mendoza, A., del Olmo, A., Pérez, I., Quirrenbach, A., Zibetti, S., 2015. The CALIFA survey across the Hubble sequence. Spatially resolved stellar population properties in galaxies. *Astronom. Astrophys.* 581, A103. doi:10.1051/0004-6361/201525938, arXiv:1506.04157.
- González Delgado, R.M., Pérez, E., Cid Fernandes, R., et al., 2014. The star formation history of CALIFA galaxies: Radial structures. *Astronom. Astrophys.* 562, A47. doi:10.1051/0004-6361/201322011, arXiv:1310.5517.
- Gorgas, J., Cardiel, N., Pedraz, S., González, J.J., 1999. Empirical calibration of the lambda 4000 Å break. *Astron. Astrophys. Sup.* 139, 29–41. doi:10.1051/aas:1999375, arXiv:astro-ph/9905264.
- Guidi, G., Casado, J., Ascasibar, Y., García-Benito, R., Galbany, L., Sánchez-Blázquez, P., Sánchez, S.F., Rosales-Ortega, F.F., Scannapieco, C., 2018. The SELGIFS data challenge: generating synthetic observations of CALIFA galaxies from hydrodynamical simulations. *Mon. Not. R. Astron. Soc.* 479, 917–931. doi:10.1093/mnras/sty1480, arXiv:1610.07620.
- Ho, I.T., Medling, A.M., Groves, B., Rich, J.A., Rupke, D.S.N., Hampton, E., Kewley, L.J., Bland-Hawthorn, J., Croom, S.M., Richards, S., Schaefer, A.L., Sharp, R., Sweet, S.M., 2016. LZIFU: an emission-line fitting toolkit for integral field spectroscopy data. *Astrophysics and Space Science* 361, 280. doi:10.1007/s10509-016-2865-2, arXiv:1607.06561.
- Hunter, J.D., 2007. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering* 9, 90–95. doi:10.1109/MCSE.2007.55.
- Ibarra-Medel, H.J., Avila-Reese, V., Sánchez, S.F., González-Samaniego, A., Rodríguez-Puebla, A., 2019. Optical integral field spectroscopy observations applied to simulated galaxies: testing the fossil record method. *Mon. Not. R. Astron. Soc.* 483, 4525–4550. doi:10.1093/mnras/sty3256, arXiv:1811.04856.
- Ibarra-Medel, H.J., Sánchez, S.F., Avila-Reese, V., Hernández-Toledo, H.M., González, J.J., Drory, N., Bundy, K., Bizyaev, D., Cano-Díaz, M., Malanushenko, E., Pan, K., Roman-Lopes, A., Thomas, D., 2016. SDSS IV MaNGA: the global and local stellar mass assembly histories of galaxies. *Mon. Not. R. Astron. Soc.* 463, 2799–2818. doi:10.1093/mnras/stw2126, arXiv:1609.01304.
- Kauffmann, G., Heckman, T.M., Tremonti, C., Brinchmann, J., Charlot, S., White, S.D.M., Ridgway, S.E., Brinkmann, J., Fukugita, M., Hall, P.B., Ivezić, Ž., Richards, G.T., Schneider, D.P., 2003. The host galaxies of active galactic nuclei. *Mon. Not. R. Astron. Soc.* 346, 1055–1077. doi:10.1111/j.1365-2966.2003.07154.x, arXiv:astro-ph/0304239.
- Kelz, A., Verheijen, M.A.W., Roth, M.M., Bauer, S.M., Becker, T., Paschke, J., Popow, E., Sánchez, S.F., Laux, U., 2006. PMAS: The Potsdam Multi-Aperture Spectrophotometer. II. The Wide Integral Field Unit PPak. *Publications of the Astronomical Society of the Pacific* 118, 129–145. doi:10.1086/497455, arXiv:astro-ph/0512557.
- Kewley, L.J., Dopita, M.A., Sutherland, R.S., Heisler, C.A., Trevena, J., 2001. Theoretical Modeling of Starburst Galaxies. *Astrophys. J.* 556, 121–140. doi:10.1086/321545, arXiv:astro-ph/0106324.
- Koleva, M., Prugniel, P., Ocvirk, P., Le Borgne, D., Soubiran, C., 2008. Spectroscopic ages and metallicities of stellar populations: validation of full spectrum fitting. *Mon. Not. R. Astron. Soc.* 385, 1998–2010. doi:10.1111/j.1365-2966.2008.12908.x, arXiv:0801.0871.
- Lacerda, E.A.D., Sánchez, S.F., Cid Fernandes, R., López-Cobá, C., Espinosa-Ponce, C., Galbany, L., 2020. Galaxies hosting an active galactic nucleus: a view from the CALIFA survey. *Mon. Not. R. Astron. Soc.* 492, 3073–3090. doi:10.1093/mnras/staa008, arXiv:2001.00099.
- López-Cobá, C., Sánchez, S.F., Anderson, J.P., Cruz-González, I., Galbany, L., Ruiz-Lara, T., Barrera-Ballesteros, J.K., Prieto, J.L., Kuncarayakti, H., 2020. The AMUSING++ Nearby Galaxy Compilation: I. Full Sample Characterization and Galactic-Scale Outflows Selection. arXiv e-prints, arXiv:2002.09328, arXiv:2002.09328.
- López-Cobá, C., Sánchez, S.F., Bland-Hawthorn, J., Moiseev, A.V., Cruz-González, I., García-Benito, R., Barrera-Ballesteros, J.K., Galbany, L., 2019. Systematic study of outflows in the Local Universe using CALIFA: I. Sample selection and main properties. *Mon. Not. R. Astron. Soc.* 482, 4032–4056. doi:10.1093/mnras/sty2960, arXiv:1811.01253.
- Martins, L.P., González Delgado, R.M., Leitherer, C., Cerviño, M., Hauschildt, P., 2005. A high-resolution stellar library for evolutionary population synthesis. *Mon. Not. R. Astron. Soc.* 358, 49–65. doi:10.1111/j.1365-2966.2005.08703.x, arXiv:astro-ph/0501225.
- Matsuoka, Y., Oyabu, S., Tsuzuki, Y., Kawara, K., 2007. Observations of O I and Ca II Emission Lines in Quasars: Implications for the Site of Fe II Line Emission. *Astrophys. J.* 663, 781–798. doi:10.1086/518399, arXiv:astro-ph/0703659.
- Mejía-Narváez, A., Sánchez, S.F., Lacerda, E.A.D., Carigi, L., Galbany, L., Husemann, B., García-Benito, R., 2020. The stellar metallicity distribution function of galaxies in the CALIFA survey. *Mon. Not. R. Astron. Soc.* 499, 4838–4853. doi:10.1093/mnras/staa3094, arXiv:2009.13712.
- Ocvirk, P., Pichon, C., Lançon, A., Thiébaud, E., 2006. STECKMAP: STEllar Content and Kinematics from high resolution galactic spectra via Maximum A Posteriori. *Mon. Not. R. Astron. Soc.* 365, 74–84. doi:10.1111/j.1365-2966.2005.09323.x, arXiv:astro-ph/0507002.
- Oliphant, T.E., 2006. A guide to NumPy. volume 1. Trelgol Publishing USA.
- Papaderos, P., Gomes, J.M., Vilchez, J.M., Kehrig, C., Lehnert, M.D., Ziegler, B., Sánchez, S.F., Husemann, B., Monreal-Ibero, A., García-Benito, R., Bland-Hawthorn, J., Cortijo-Ferrero, C., de Lorenzo-Cáceres, A., del Olmo, A., Falcón-Barroso, J., Galbany, L., Iglesias-Páramo, J., López-Sánchez, Á.R., Márquez, I., Mollá, M., Mast, D., van de Ven, G., Wisotzki, L., 2013. Nebular emission and the Lyman continuum photon escape fraction in CALIFA early-type galaxies. *Astronom. Astrophys.* 555, L1. doi:10.1051/0004-6361/201321681, arXiv:1306.2338.
- Papaderos, P., Izotov, Y.I., Thuan, T.X., Noeske, K.G., Fricke, K.J., Guiseva, N.G., Green, R.F., 2002. The blue compact dwarf galaxy I Zw 18: A comparative study of its low-surface-brightness component. *Astronom. Astrophys.* 393, 461–483. doi:10.1051/0004-6361:20021023, arXiv:astro-ph/0207314.
- Pérez, E., Cid Fernandes, R., González Delgado, R.M., García-Benito, R., Sánchez, S.F., Husemann, B., Mast, D., Rodón, J.R., Kupko, D., Bockmann, N., de Amorim, A.L., van de Ven, G., Walcher, J., Wisotzki, L., Cortijo-Ferrero, C., CALIFA Collaboration, 2013. The Evolution of Galaxies Resolved in Space and Time: A View of Inside-out Growth from the CALIFA Survey. *Astrophys. J. Lett.* 764, L1. doi:10.1088/2041-8205/764/1/L1, arXiv:1301.1679.
- Pérez, F., Granger, B.E., 2007. IPython: a system for interactive scientific computing. *Computing in Science and Engineering* 9, 21–29. URL: <https://ipython.org>, doi:10.1109/MCSE.2007.53.
- Price-Whelan, A.M., Sipőcz, B., Günther, H., Lim, P., Crawford, S., Conwell, S., Shupe, D., Craig, M., Dencheva, N., Ginsburg, A., et al., 2018. The astropy project: Building an open-science project and status of the v2.0 core package. *The Astronomical Journal* 156, 123.
- Ronen, S., Aragon-Salamanca, A., Lahav, O., 1999. Principal component analysis of synthetic galaxy spectra. *Mon. Not. R. Astron. Soc.* 303, 284–296. doi:10.1046/j.1365-8711.1999.02222.x, arXiv:astro-ph/9805130.
- Roth, M.M., Kelz, A., Fechner, T., Hahn, T., Bauer, S.M., Becker, T., Böhm, P., Christensen, L., Dionies, F., Paschke, J., Popow, E., Wolter, D., Scholl, J., Laux, U., Altmann, W., 2005. PMAS: The Potsdam Multi-Aperture Spectrophotometer. I. Design, Manufacture, and Performance. *Publications of the Astronomical Society of the Pacific* 117, 620–642. doi:10.1086/429877, arXiv:astro-ph/0502581.
- Salpeter, E.E., 1955. The luminosity function and stellar evolution. *ApJ* 121, 161. URL: [http://adsabs.harvard.edu/cgi-bin/nph-data\\_query?bibcode=1955ApJ...121..161S&link\\_type=ABSTRACT](http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=1955ApJ...121..161S&link_type=ABSTRACT), doi:10.1086/145971.
- Sánchez, S.F., 2020. Spatially Resolved Spectroscopic Properties of Low-Redshift Star-Forming Galaxies. *Annu. Rev. Astron. Astrophys.* 58, 99–155. doi:10.1146/annurev-astro-012120-013326, arXiv:1911.06925.
- Sánchez, S.F., Avila-Reese, V., Hernandez-Toledo, H., Cortes-Suárez, E., Rodríguez-Puebla, A., Ibarra-Medel, H., Cano-Díaz, M., Barrera-Ballesteros, J.K., Negrete, C.A., Calette, A.R., de Lorenzo-Cáceres, A., Ortega-Minakata, R.A., Aquino, E., Valenzuela, O., Clemente, J.C., Storch-Bergmann, T., Riffel, R., Schimoia, J., Riffel, R.A., Rembold, S.B., Brownstein, J.R., Pan, K., Yates, R., Mallmann, N., Bitsakis, T., 2018. SDSS IV MaNGA - Properties of AGN Host Galaxies. *Rev. Mex. Astron. Astr.* 54, 217–260. arXiv:1709.05438.
- Sánchez, S.F., Barrera-Ballesteros, J.K., López-Cobá, C., Brough, S., Bryant, J.J., Bland-Hawthorn, J., Croom, S.M., van de Sande, J., Cortese, L., Goodwin, M., Lawrence, J.S., López-Sánchez, A.R., Sweet, S.M., Owers, M.S., Richards, S.N., Walcher, C.J., 2019. The SAMI galaxy survey: exploring the gas-phase mass-metallicity relation. *Mon. Not. R. Astron. Soc.* 484,

- 3042–3070. doi:10.1093/mnras/stz019, arXiv:1812.11263.
- Sánchez, S.F., García-Benito, R., Zibetti, S., Walcher, C.J., Husemann, B., Mendoza, M.A., Galbany, L., Falcón-Barroso, J., Mast, D., Aceituno, J., Aguerri, J.A.L., Alves, J., Amorim, A.L., Ascasibar, Y., Barrado-Navascues, D., Barrera-Ballesteros, J., Bekeraité, S., Bland-Hawthorn, J., Cano Díaz, M., Cid Fernandes, R., Cavichia, O., Cortijo, C., Dannerbauer, H., Demleitner, M., Díaz, A., Dettmar, R.J., de Lorenzo-Cáceres, A., del Olmo, A., Galazzi, A., García-Lorenzo, B., Gil de Paz, A., González Delgado, R., Holmes, L., Iglésias-Páramo, J., Kehrig, C., Kelz, A., Kennicutt, R.C., Kleemann, B., Lacerda, E.A.D., López Fernández, R., López Sánchez, A.R., Lyubenova, M., Marino, R., Márquez, I., Mendez-Abreu, J., Mollá, M., Monreal-Ibero, A., Ortega Minakata, R., Torres-Papaqui, J.P., Pérez, E., Rosales-Ortega, F.F., Roth, M.M., Sánchez-Blázquez, P., Schilling, U., Spekkens, K., Vale Asari, N., van den Bosch, R.C.E., van de Ven, G., Vilchez, J.M., Wild, V., Wisotzki, L., Yıldırım, A., Ziegler, B., 2016a. CALIFA, the Calar Alto Legacy Integral Field Area survey. IV. Third public data release. *Astronom. Astrophys.* 594, A36. doi:10.1051/0004-6361/201628661, arXiv:1604.02289.
- Sánchez, S.F., Kennicutt, R.C., Gil de Paz, A., van de Ven, G., Vilchez, J.M., Wisotzki, L., Walcher, C.J., Mast, D., Aguerri, J.A.L., Albiol-Pérez, S., Alonso-Herrero, A., Alves, J., Bakos, J., Bartáková, T., Bland-Hawthorn, J., Boselli, A., Bomans, D.J., Castillo-Morales, A., Cortijo-Ferrero, C., de Lorenzo-Cáceres, A., Del Olmo, A., Dettmar, R.J., Díaz, A., Ellis, S., Falcón-Barroso, J., Flores, H., Gallazzi, A., García-Lorenzo, B., González Delgado, R., Gruel, N., Haines, T., Hao, C., Husemann, B., Iglésias-Páramo, J., Jahnke, K., Johnson, B., Jungwiert, B., Kalinova, V., Kehrig, C., Kupko, D., López-Sánchez, Á.R., Lyubenova, M., Marino, R.A., Marmol-Queraltó, E., Márquez, I., Masegosa, J., Meidt, S., Mendez-Abreu, J., Monreal-Ibero, A., Montijo, C., Mourão, A.M., Palacios-Navarro, G., Papaderos, P., Pasquali, A., Peletier, R., Pérez, E., Pérez, I., Quirrenbach, A., Relaño, M., Rosales-Ortega, F.F., Roth, M.M., Ruiz-Lara, T., Sánchez-Blázquez, P., Sengupta, C., Singh, R., Stanishev, V., Trager, S.C., Vazdekis, A., Viironen, K., Wild, V., Zibetti, S., Ziegler, B., 2012. CALIFA, the Calar Alto Legacy Integral Field Area survey. I. Survey presentation. *Astronom. Astrophys.* 538, A8. doi:10.1051/0004-6361/201117353, arXiv:1111.0962.
- Sánchez, S.F., Pérez, E., Sánchez-Blázquez, P., García-Benito, R., Ibarra-Mede, H.J., González, J.J., Rosales-Ortega, F.F., Sánchez-Menguiano, L., Ascasibar, Y., Bitsakis, T., Law, D., Cano-Díaz, M., López-Cobá, C., Marino, R.A., Gil de Paz, A., López-Sánchez, A.R., Barrera-Ballesteros, J., Galbany, L., Mast, D., Abril-Melgarejo, V., Roman-Lopes, A., 2016b. Pipe3D, a pipeline to analyze Integral Field Spectroscopy Data: II. Analysis sequence and CALIFA dataproducts. *Rev. Mex. Astron. Astr.* 52, 171–220. arXiv:1602.01830.
- Sánchez, S.F., Pérez, E., Sánchez-Blázquez, P., González, J.J., Rosales-Ortega, F.F., Cano-Díaz, M., López-Cobá, C., Marino, R.A., Gil de Paz, A., Mollá, M., López-Sánchez, A.R., Ascasibar, Y., Barrera-Ballesteros, J., 2016c. Pipe3D, a pipeline to analyze Integral Field Spectroscopy Data: I. New fitting philosophy of FIT3D. *Rev. Mex. Astron. Astr.* 52, 21–53. arXiv:1509.08552.
- Sánchez, S.F., Walcher, C.J., Lopez-Cobá, C., Barrera-Ballesteros, J.K., Mejía-Narváez, A., Espinosa-Ponce, C., Camps-Fariña, A., 2021. From Global to Spatially Resolved in Low-Redshift Galaxies. *Rev. Mex. Astron. Astr.* 57, 3–38. doi:10.22201/ia.01851101p.2021.57.01.01, arXiv:2009.00424.
- Sánchez-Blázquez, P., Ocvirk, P., Gibson, B.K., Pérez, I., Peletier, R.F., 2011. Star formation history of barred disc galaxies. *Mon. Not. R. Astron. Soc.* 415, 709–731. doi:10.1111/j.1365-2966.2011.18749.x, arXiv:1103.3796.
- Sánchez-Blázquez, P., Peletier, R.F., Jiménez-Vicente, J., Cardiel, N., Cenarro, A.J., Falcón-Barroso, J., Gorgas, J., Selam, S., Vazdekis, A., 2006. Medium-resolution Isaac Newton Telescope library of empirical spectra. *Mon. Not. R. Astron. Soc.* 371, 703–718. doi:10.1111/j.1365-2966.2006.10699.x, arXiv:astro-ph/0607009.
- Sánchez-Menguiano, L., Sánchez, S.F., Pérez, I., Debattista, V.P., Ruiz-Lara, T., Florido, E., Cavichia, O., Galbany, L., Marino, R.A., Mast, D., Sánchez-Blázquez, P., Méndez-Abreu, J., de Lorenzo-Cáceres, A., Catalán-Torrecilla, C., Cano-Díaz, M., Márquez, I., McIntosh, D.H., Ascasibar, Y., García-Benito, R., González Delgado, R.M., Kehrig, C., López-Sánchez, Á.R., Mollá, M., Bland-Hawthorn, J., Walcher, C.J., Costantin, L., 2017. Arm and interarm abundance gradients in CALIFA spiral galaxies. *Astronom. Astrophys.* 603, A113. doi:10.1051/0004-6361/201630062, arXiv:1705.05733.
- Sánchez-Menguiano, L., Sánchez, S.F., Pérez, I., García-Benito, R., Husemann, B., Mast, D., Mendoza, A., Ruiz-Lara, T., Ascasibar, Y., Bland-Hawthorn, J., Cavichia, O., Díaz, A.I., Florido, E., Galbany, L., González Delgado, R.M., Kehrig, C., Marino, R.A., Márquez, I., Masegosa, J., Méndez-Abreu, J., Mollá, M., Del Olmo, A., Pérez, E., Sánchez-Blázquez, P., Stanishev, V., Walcher, C.J., López-Sánchez, Á.R., Califa Collaboration, 2016. Shape of the oxygen abundance profiles in CALIFA face-on spiral galaxies. *Astronom. Astrophys.* 587, A70. doi:10.1051/0004-6361/201527450, arXiv:1601.01542.
- Schröder, C., Reiners, A., Schmitt, J.H.M.M., 2009. Ca II HK emission in rapidly rotating stars. Evidence for an onset of the solar-type dynamo. *Astronom. Astrophys.* 493, 1099–1107. doi:10.1051/0004-6361:200810377.
- Stasińska, G., Cid Fernandes, R., Mateus, A., Sodré, L., Asari, N.V., 2006. Semi-empirical analysis of Sloan Digital Sky Survey galaxies - III. How to distinguish AGN hosts. *Mon. Not. R. Astron. Soc.* 371, 972–982. doi:10.1111/j.1365-2966.2006.10732.x, arXiv:astro-ph/0606724.
- Statler, T., 1995. Bringing the Cross-Correlation Method Up to Date. *Astron. J.* 109, 1371. doi:10.1086/117369.
- Stencel, R.E., 1977. Emission lines in the wings of Ca II H and K. II. Stellar observations - dependence of line width on luminosity and related topics. *Astrophys. J.* 215, 176–187. doi:10.1086/155345.
- van de Sande, J., Lagos, C.D.P., Welker, C., Bland-Hawthorn, J., Schulze, F., Remus, R.S., Bahé, Y., Brough, S., Bryant, J.J., Cortese, L., Croom, S.M., Devriendt, J., Dubois, Y., Goodwin, M., Konstantopoulos, I.S., Lawrence, J.S., Medling, A.M., Pichon, C., Richards, S.N., Sanchez, S.F., Scott, N., Sweet, S.M., 2019. The SAMI Galaxy Survey: comparing 3D spectroscopic observations with galaxies from cosmological hydrodynamical simulations. *Mon. Not. R. Astron. Soc.* 484, 869–891. doi:10.1093/mnras/sty3506, arXiv:1810.10542.
- Van Rossum, G., Drake, F.L., 2009. Python 3 Reference Manual. CreateSpace, Scotts Valley, CA.
- Van Rossum, G., Drake Jr, F.L., 1995. Python tutorial. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands.
- Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., Cenarro, A.J., Beasley, M.A., Cardiel, N., Gorgas, J., Peletier, R.F., 2010. Evolutionary stellar population synthesis with MILES - I. The base models and a new line index system. *Mon. Not. R. Astron. Soc.* 404, 1639–1671. doi:10.1111/j.1365-2966.2010.16407.x, arXiv:1004.4439.
- Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, İ., Feng, Y., Moore, E.W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., Contributors, S., 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17, 261–272. doi:https://doi.org/10.1038/s41592-019-0686-2.
- Walcher, C.J., Wisotzki, L., Bekeraité, S., Husemann, B., Iglesias-Páramo, J., Backmann, N., Barrera Ballesteros, J., Catalán-Torrecilla, C., Cortijo, C., del Olmo, A., García Lorenzo, B., Falcón-Barroso, J., Jilkova, L., Kalinova, V., Mast, D., Marino, R.A., Méndez-Abreu, J., Pasquali, A., Sánchez, S.F., Trager, S., Zibetti, S., Aguerri, J.A.L., Alves, J., Bland-Hawthorn, J., Boselli, A., Castillo Morales, A., Cid Fernandes, R., Flores, H., Galbany, L., Gallazzi, A., García-Benito, R., Gil de Paz, A., González-Delgado, R.M., Jahnke, K., Jungwiert, B., Kehrig, C., Lyubenova, M., Márquez Perez, I., Masegosa, J., Monreal Ibero, A., Pérez, E., Quirrenbach, A., Rosales-Ortega, F.F., Roth, M.M., Sanchez-Blázquez, P., Spekkens, K., Tundo, E., van de Ven, G., Verheijen, M.A.W., Vilchez, J.V., Ziegler, B., 2014. CALIFA: a diameter-selected sample for an integral field spectroscopy galaxy survey. *Astronom. Astrophys.* 569, A1 (W14). doi:10.1051/0004-6361/201424198, arXiv:1407.2939.
- Wall, L., Christiansen, T., Orwant, J., 2000. Programming perl. ” O’Reilly Media, Inc.”.
- Waskom, M., Botvinnik, O., O’Kane, D., Hobson, P., Lukauskas, S., Gempferline, D.C., Augspurger, T., Halchenko, Y., Cole, J.B., Warmenhoven, J., de Ruiter, J., Pye, C., Hoyer, S., Vanderplas, J., Villalba, S., Kunter, G., Quintero, E., Bachant, P., Martin, M., Meyer, K., Miles, A., Ram, Y., Yarkoni, T., Williams, M.L., Evans, C., Fitzgerald, C., Brian, Fonnesbeck, C., Lee, A., Qalieh, A., 2017. mwaskom/seaborn: v0.8.1 (septem-

- ber 2017). URL: <https://doi.org/10.5281/zenodo.883859>, doi:10.5281/zenodo.883859.
- Westfall, K.B., Cappellari, M., Bershady, M.A., Bundy, K., Belfiore, F., Ji, X., Law, D.R., Schaefer, A., Shetty, S., Tremonti, C.A., Yan, R., Andrews, B.H., Brownstein, J.R., Cherinka, B., Coccato, L., Drory, N., Maraston, C., Parikh, T., Sánchez-Gallego, J.R., Thomas, D., Weijmans, A.M., Barrera-Ballesteros, J., Du, C., Goddard, D., Li, N., Masters, K., Ibarra Medel, H.J., Sánchez, S.F., Yang, M., Zheng, Z., Zhou, S., 2019. The Data Analysis Pipeline for the SDSS-IV MaNGA IFU Galaxy Survey: Overview. *Astron. J.* 158, 231. doi:10.3847/1538-3881/ab44a2, [arXiv:1901.00856](https://arxiv.org/abs/1901.00856).
- Worthey, G., 1994. Comprehensive Stellar Population Models and the Disentanglement of Age and Metallicity Effects. *Astrophys. J. Suppl. S* 95, 107. doi:10.1086/192096.
- York, D.G., Adelman, J., Anderson, Jr., J.E., Anderson, S.F., Annis, J., Bahcall, N.A., Bakken, J.A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W.N., Bracker, S., Briegel, C., Briggs, J.W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Carr, M.A., Castander, F.J., Chen, B., Colestock, P.L., Connolly, A.J., Crocker, J.H., Csabai, I., Czarapata, P.C., Davis, J.E., Doi, M., Dombeck, T., Eisenstein, D., Ellman, N., Elms, B.R., Evans, M.L., Fan, X., Federwitz, G.R., Fiscelli, L., Friedman, S., Frieman, J.A., Fukugita, M., Gillespie, B., Gunn, J.E., Gurbani, V.K., de Haas, E., Haldeman, M., Harris, F.H., Hayes, J., Heckman, T.M., Hennessy, G.S., Hindsley, R.B., Holm, S., Holmgren, D.J., Huang, C.h., Hull, C., Husby, D., Ichikawa, S.I., Ichikawa, T., Ivezić, Ž., Kent, S., Kim, R.S.J., Kinney, E., Klaene, M., Kleinman, A.N., Kleinman, S., Knapp, G.R., Korienek, J., Kron, R.G., Kunszt, P.Z., Lamb, D.Q., Lee, B., Leger, R.F., Limmongkol, S., Lindenmeyer, C., Long, D.C., Loomis, C., Loveday, J., Lucinio, R., Lupton, R.H., MacKinnon, B., Mannery, E.J., Mantsch, P.M., Margon, B., McGehee, P., McKay, T.A., Meiksin, A., Merelli, A., Monet, D.G., Munn, J.A., Narayanan, V.K., Nash, T., Neilsen, E., Neswold, R., Newberg, H.J., Nichol, R.C., Nicinski, T., Nonino, M., Okada, N., Okamura, S., Ostriker, J.P., Owen, R., Pauls, A.G., Peoples, J., Peterson, R.L., Petravick, D., Pier, J.R., Pope, A., Pordes, R., Protopapio, A., Rechenmacher, R., Quinn, T.R., Richards, G.T., Richmond, M.W., Rivetta, C.H., Rockosi, C.M., Ruthmansdorfer, K., Sandford, D., Schlegel, D.J., Schneider, D.P., Sekiguchi, M., Sergey, G., Shimasaku, K., Siegmund, W.A., Smee, S., Smith, J.A., Snedden, S., Stone, R., Stoughton, C., Strauss, M.A., Stubbs, C., SubbaRao, M., Szalay, A.S., Szapudi, I., Szokoly, G.P., Thakar, A.R., Tremonti, C., Tucker, D.L., Uomoto, A., Vanden Berk, D., Vogeley, M.S., Waddell, P., Wang, S.i., Watanabe, M., Weinberg, D.H., Yanny, B., Yasuda, N., SDSS Collaboration, 2000. The Sloan Digital Sky Survey: Technical Summary. *Astron. J.* 120, 1579–1587. doi:10.1086/301513, [arXiv:astro-ph/0006396](https://arxiv.org/abs/astro-ph/0006396).