

# LZIFU: an emission-line fitting toolkit for integral field spectroscopy data

I-Ting Ho<sup>1,2</sup> • Anne M. Medling<sup>2</sup> • Brent Groves<sup>2</sup> •  
 Jeffrey A. Rich<sup>3,4</sup> • David S. N. Rupke<sup>5</sup> • Elise Hampton<sup>2</sup> •  
 Lisa J. Kewley<sup>1,2</sup> • Joss Bland-Hawthorn<sup>6</sup> •  
 Scott M. Croom<sup>6,7</sup> • Samuel Richards<sup>6,7,8</sup> •  
 Adam L. Schaefer<sup>6,7,8</sup> • Rob Sharp<sup>2,7</sup> • Sarah M. Sweet<sup>2</sup>

**Abstract** We present LZIFU (LaZy-IFU), an IDL toolkit for fitting multiple emission lines simultaneously in integral field spectroscopy (IFS) data. LZIFU is useful for the investigation of the dynamical, physical and chemical properties of gas in galaxies. LZIFU has already been applied to many world-class IFS instruments and large IFS surveys, including the Wide Field Spectrograph, the new Multi Unit Spectroscopic Explorer (MUSE), the Calar Alto Legacy Integral Field Area (CALIFA) survey, the Sydney-Australian-

astronomical-observatory Multi-object Integral-field spectrograph (SAMI) Galaxy Survey. Here we describe in detail the structure of the toolkit, and how the line fluxes and flux uncertainties are determined, including the possibility of having multiple distinct kinematic components. We quantify the performance of LZIFU, demonstrating its accuracy and robustness. We also show examples of applying LZIFU to CALIFA and SAMI data to construct emission line and kinematic maps, and investigate complex, skewed line profiles presented in IFS data. The code is made available to the astronomy community through github. LZIFU will be further developed over time to other IFS instruments, and to provide even more accurate line and uncertainty estimates.

I-Ting Ho  
 itho@ifa.hawaii.edu  
 Anne M. Medling  
 Brent Groves  
 Jeffrey A. Rich  
 David S. N. Rupke  
 Elise Hampton  
 Lisa J. Kewley  
 Joss Bland-Hawthorn  
 Scott M. Croom  
 Samuel Richards  
 Adam L. Schaefer  
 Rob Sharp  
 Sarah M. Sweet

<sup>1</sup>Institute for Astronomy, University of Hawaii, 2680 Woodlawn Dr., Honolulu, HI 96822, USA

<sup>2</sup>Research School of Astronomy and Astrophysics, Australian National University, Cotter Rd., Weston ACT 2611, Australia

<sup>3</sup>Infrared Processing and Analysis Center, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA

<sup>4</sup>Observatories of the Carnegie Institution of Washington, 813 Santa Barbara St., Pasadena, CA 91101, USA

<sup>5</sup>Department of Physics, Rhodes College, Memphis, TN 38112, USA

<sup>6</sup>Sydney Institute for Astronomy, School of Physics, University of Sydney, NSW 2006, Australia

<sup>7</sup>ARC Centre of Excellence for All-sky Astrophysics (CAASTRO)

<sup>8</sup>Australian Astronomical Observatory, PO Box 915, North Ryde NSW 1670, Australia

## 1 Introduction

Galaxy emission-line spectroscopy has always been a powerful tool for the analysis of the dynamical, physical and chemical properties of galaxies. Traditionally, spectroscopy of galaxies has been obtained by dispersing the light either across a slit (sacrificing one spatial dimension) or from a fibre (producing a single integrated spectrum). Active development of modern integral field spectroscopy (IFS) has made capturing 3-dimensional structures of galaxies very efficient, revolutionising the way we observe and study galaxies.

The complex and perhaps stochastic nature of different physical processes governing galaxy evolution has inspired large galaxy surveys. In recent decades, large fibre and slit spectroscopic surveys such as the Sloan Digital Sky Survey (SDSS; York et al. 2000), the 2dF Galaxy Redshift Survey (Colless 1999), and the Deep Extragalactic Evolutionary Probe 2 survey (DEEP2; Davis et al. 2003) have drastically improved our understanding of the global (unresolved) properties of galaxy populations at different epochs of the Universe. Integral field spectroscopic surveys have recently become feasible, providing access simultaneously to both spectral and kinematic information of

large numbers of galaxies. Two pioneering IFS surveys, the SAURON survey (Bacon et al. 2001) and its extension the ATLAS<sup>3D</sup> survey (Cappellari et al. 2011), studied about 260 early type galaxies in the local Universe ( $z < 0.01$ ). Surveys targeting both the blue and red galaxy populations, such as the Calar Alto Legacy Integral Field Area (CALIFA) survey (Sánchez et al. 2012), the Sydney-Australian-astronomical-observatory Multi-object Integral-field spectrograph (SAMI) Galaxy Survey (Croom et al. 2012; Bryant et al. 2015), and the Mapping Nearby Galaxies at Apache Point Observatory survey (MaNGA; Bundy et al. 2015), are currently underway. These IFS surveys will provide critical information to bridge the knowledge gaps resulting from the limited spatial and kinematic information delivered by previous single-fibre and slit spectroscopic surveys.

The sample sizes and data flows of these modern IFS surveys are substantial. With each data cube containing typically one to two thousand spectra, the CALIFA survey plans to observe about 600 galaxies in the local Universe ( $0.005 < z < 0.03$ ); the SAMI Galaxy Survey will reach a sample size of 3,400 galaxies at  $z < 0.12$ ; and the MaNGA survey will build up a sample of 10,000 galaxies at a similar redshift to the SAMI Galaxy Survey. Future surveys using high-multiplex integral field unit (IFU) instrument such as HECTOR on the Anglo-Australian Telescope will observe on the order of 100,000 galaxies (Lawrence et al. 2012; Bland-Hawthorn 2015). Current and forthcoming wide-field IFU instruments are also delivering large quantity of high quality data, such as the Wide Field Spectrograph (WiFeS) on the Australian National University 2.3-m telescope (Dopita et al. 2007, 2010), the new Multi Unit Spectroscopic Explorer (MUSE) on the Very Large Telescope (Bacon et al. 2010), the SITELE instrument on the Canada France Hawaii Telescope (Grandmont et al. 2012), the Keck Cosmic Web Imager at the W. M. Keck Observatory (Martin et al. 2010; Morrissey et al. 2012).

Significant efforts have been placed in developing corresponding tools for analysing large volume of spectroscopic data. The stellar continuum contains valuable information about the stellar kinematics, chemistry and star formation history of galaxies. Packages such as the STELLAR Content via Maximum A Posteriori (STECMAP; Ocvirk et al. 2006) package, the penalized pixel-fitting (PPXF; Cappellari and Emsellem 2004) routine and the STARLIGHT package (Cid Fernandes et al. 2005) can perform spectral template fitting and extract various stellar properties. For investigating gas physics, the emission lines fitting tools such as the Gas AND Absorption Line Fitting code (GANDALF; Sarzi et al. 2006), the FIT3D package (Sánchez et al. 2006, 2007; and the successor PIPE3D; Sánchez et al. 2016a,b), and the Peak ANA-

lysis utility (PAN<sup>1</sup>; Dimeo 2005) are commonly adopted to measure emission line fluxes and kinematics.

As the spectral resolution of the instruments continue to improve, the intrinsic non-Gaussian line profile complicates the emission line analysis. When the spectral resolution is high ( $R > 3000$ ), galaxies with active gas dynamics, such as winds, outflows or AGN, usually present skewed line profiles that require fitting multiple, assumed Gaussian, components to separate the different kinematic components overlapping in the line-of-sight direction (also referred as “spectral decomposition”). Performing spectral decomposition on large datasets is non-trivial as significant human input is usually required. Here, we present our emission line fitting pipeline LaZy-IFU<sup>2</sup> (LZIFU; written in the Interactive Data Language [IDL]), which is designed to eliminate the need for individual treatment of each of many thousands of spectra across an IFS galaxy survey (such as CALIFA, SAMI or MaNGA).

The main objective of LZIFU is to extract 2-dimensional emission line flux maps and kinematic maps useful for investigating gas physics in galaxies. LZIFU has already been adopted in various publications using data from multiple instruments and surveys, including MUSE (Kreckel et al. 2016), SAMI (e.g. Ho et al. 2014; Richards et al. 2014; Allen et al. 2015b; Ho et al. 2016), CALIFA (Davies et al. 2014; Ho et al. 2015), WiFeS (Ho et al. 2015; Dopita et al. 2015a,b; Vogt et al. 2015; Medling et al. 2015), and SPIRAL on the Anglo-Australian Telescope (McElroy et al. 2015). The following characteristics were considered carefully while developing LZIFU. First, the pipeline must perform spectral decomposition automatically without needing repeated human instructions. Second, the pipeline needs to be scriptable for batch reduction, such that when necessary the same results can be reproduced by re-executing the same scripts. Third, the pipeline must be flexible and generalised so that data from most modern IFS instruments can be accepted without major restructuring of the inputs. Finally, the calculation speed must be optimised and the pipeline has to take advantage of parallel processing because of 1) the huge data flow from multiplexed IFS surveys, and 2) the possibility of fitting the same datasets multiple times for various experimental purposes.

The focus of this paper is to present the core structure of LZIFU (Section 2), and examine the errors produced by the pipeline (Section 4). We also show examples of applying LZIFU on the CALIFA survey and SAMI Galaxy survey (Section 3). Finally, the code will be continuously

<sup>1</sup>PAN was subsequently adapted and modified by Mark Westmoquette for astronomical requirements. See <http://ifs.wikidot.com/pan>.

<sup>2</sup>The framework of the LZIFU stems from UHSPECFIT, a tool developed at the University of Hawai’i and employed in several previous spectroscopic studies on gas abundances and outflows (e.g., Zahid and Bresolin 2011; Rupke et al. 2010; Rich et al. 2010, 2012; Rupke and Veilleux 2011).

maintained and made available to the public through github (<https://github.com/hoiting/LZIFU/releases>). We discuss future plans for the code in Section 5.

## 2 LZIFU: The spectral fitting toolkit

### 2.1 Overview

To arrive at 2D maps of line fluxes, velocity and velocity dispersion, LZIFU first removes the continuum before modelling user-assigned emission line(s) on a spaxel-to-spaxel basis. If tailored continuum models already exist, users have the option of directly subtracting the continuum by feeding LZIFU the continuum models in Flexible Image Transport System (FITS) format. The subsequent emission line fitting follows the Levenberg-Marquardt least-square method to find the most probable models (with maximum likelihood) describing the emission line spectra. Each emission line can be modelled by up to 3 Gaussians describing (potentially) different kinematic components. The final products delivered by LZIFU are continuum cubes, emission line cubes, emission line flux (and corresponding error) maps, and kinematic (and corresponding error) maps stored in multi-extension FITS files.

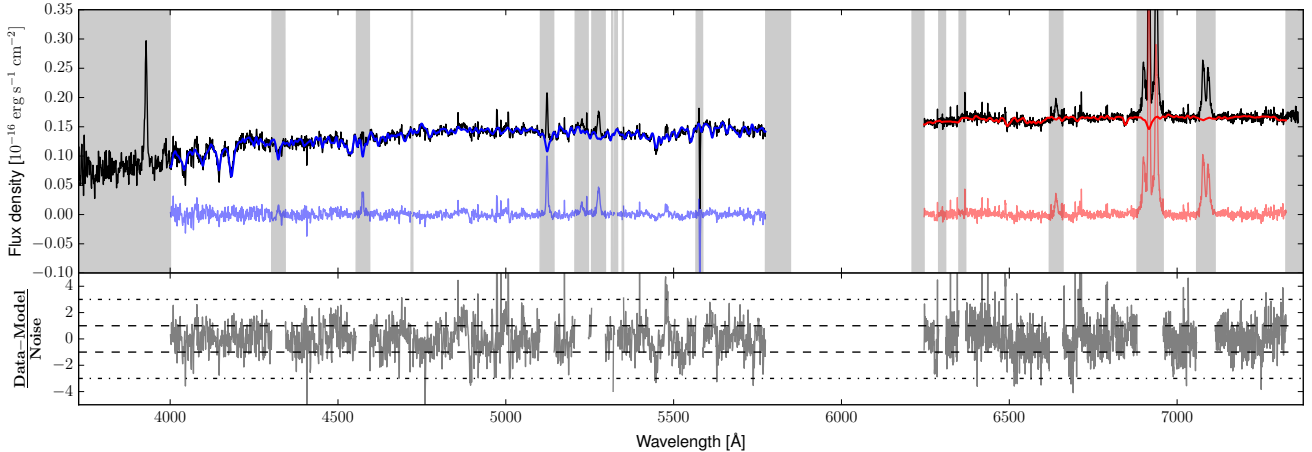
For historical reasons, LZIFU was originally designed for two-sided IFS data with each object having one blue and one red data cube. The two data cubes can have different spectral resolutions, but are required to cover *non-overlapping* spectral ranges. Such an instrumental setup is common in instrument designs and large area IFS surveys because one can achieve a trade-off between spectral coverage and spectral resolution, given that the numbers of CCD pixels are always limited. To generalise the application of LZIFU, the pipeline was modified later to accept one-sided IFS data by disabling procedures related to the blue data. Below, we elaborate on the continuum fitting and emission line fitting procedures based on two-sided data.

### 2.2 Continuum fitting

When pre-determined continuum models are not provided by the users, LZIFU models the continuum using the penalised pixel-fitting routine (PPXF; Cappellari and Emsellem 2004) that performs fits the underlying absorption continuum using a series of input spectral templates from stars or modeled simple stellar populations (SSPs) convolved with a parameterized velocity distribution. The PPXF routine is wrapped in LZIFU as the default continuum fitting method. In our implementation, the data and spectral templates are first aligned and rectified to the same spectral characteristics (i.e. wavelength coverage, spectral resolution, and channel width) before fitting the continuum. A combined spectrum

(of the blue and red data) is formed for each spaxel by convolving the data to a common spectral resolution, and re-sampling the data onto a common spectral grid. The data cube with poorer spectral resolution determines the spectral resolution and channel width of the combined spectrum. Various SSP templates collected from the literature are included in LZIFU as IDL .sav files, so the users can directly select the preferred library of SSP models. The selected SSP templates are redshifted, spectrally trimmed, and spectrally convolved to match the combined spectrum. To fit the underlying absorption-line spectrum, PPXF compares linear combinations of the SSP models with the combined spectrum in a least-square sense, during which the stellar velocity dispersion, stellar velocity, and reddening are constrained simultaneously. Channels contaminated by night sky emission lines and nebular emission lines from the galaxy are masked out prior to the fitting. Poorly-subtracted sky emission lines, with other defect channels, can be masked by providing external masks that specify the wavelength intervals to ignore. Users are also required to specify the emission lines and the width around the emission lines that should be excluded from the continuum fit. Our custom implementation of PPXF allows the users to control critical PPXF keywords directly from a LZIFU setup file. Other hardwired functionalities of PPXF can be adapted for different applications by modifying the LZIFU source code. After the best solution of spectral fitting is found, the continuum models are reconstructed separately for both sides of the data at their native resolutions. The advantage of this implementation is twofold: we utilise the largest possible spectral coverage to constrain the spectral fitting solution, and the reconstructed continuum models retain the original spectral resolution of the data.

Systematic errors in SSP models, residual calibration errors, and potential power law continuum from non-stellar components often cause spectral fitting routines to fail to achieve a perfect description of a spectrum, which would be characterised by a reduced- $\chi^2$  ( $\chi_\nu^2$ ) of approximately 1. To account for these systematic errors and possible non-stellar contributions, a polynomial term can be implemented. In PPXF, additive or multiplicative Legendre polynomials can be included and fit simultaneously with the spectral templates. These options are also maintained and passed on to PPXF. In some situations, the users may wish not to fit polynomials simultaneously with the spectral templates to avoid the continuum fit becoming highly degenerate. In these cases, the continuum subtracted-spectra may not be flat, which can affect the subsequent line flux measurements. To further flatten the continuum-subtracted spectra, LZIFU provides an extra option of fitting Legendre polynomials separately to the continuum-subtracted blue and red spectra.



**Fig. 1** An example of PPXF continuum fitting using data from the SAMI Galaxy Survey. Channels not included in the continuum fitting are masked by grey bands (i.e., bad pixels, cosmic rays, poorly-subtracted sky emission lines, and nebular emission lines). Continuum subtracted data (lower blue and red lines) are subsequently used to perform emission line fitting.

Here, the least-square fitting is performed using the BVLS (bounded-value least-square) algorithm developed by [Lawson and Hanson \(1974\)](#) and implemented in IDL by Michele Cappellari.

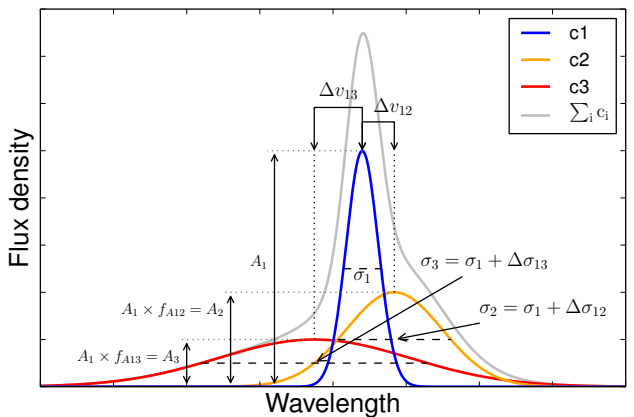
The principal objective of the custom implementation of PPXF is to correct for stellar absorption features affecting predominately the Balmer lines, and to remove the stellar continuum, such that gas physics can be derived from fitting emission lines to continuum-free spectra. The goal is not to constrain stellar parameters such as the stellar population, age and metallicity, which are known to be highly degenerate and require careful investigation of numerous local minima in the  $\chi^2$  space (e.g., [Cid Fernandes et al. 2014](#)).

In Figure 1, we show an example of the continuum fit of one spectrum from the SAMI Galaxy Survey. The blue and red data have  $1\sigma$  spectral resolutions of 1.15 and 0.72 Å, respectively; and we use the SSP spectral libraries constructed by [González Delgado et al. \(2005\)](#) with an additional Legendre polynomial of up to 12 order of Legendre polynomials to fit the continuum. Channels affected by bad pixels, cosmic rays, strong sky lines, or nebular emission lines are masked by grey bands, and are not considered in the continuum fit.

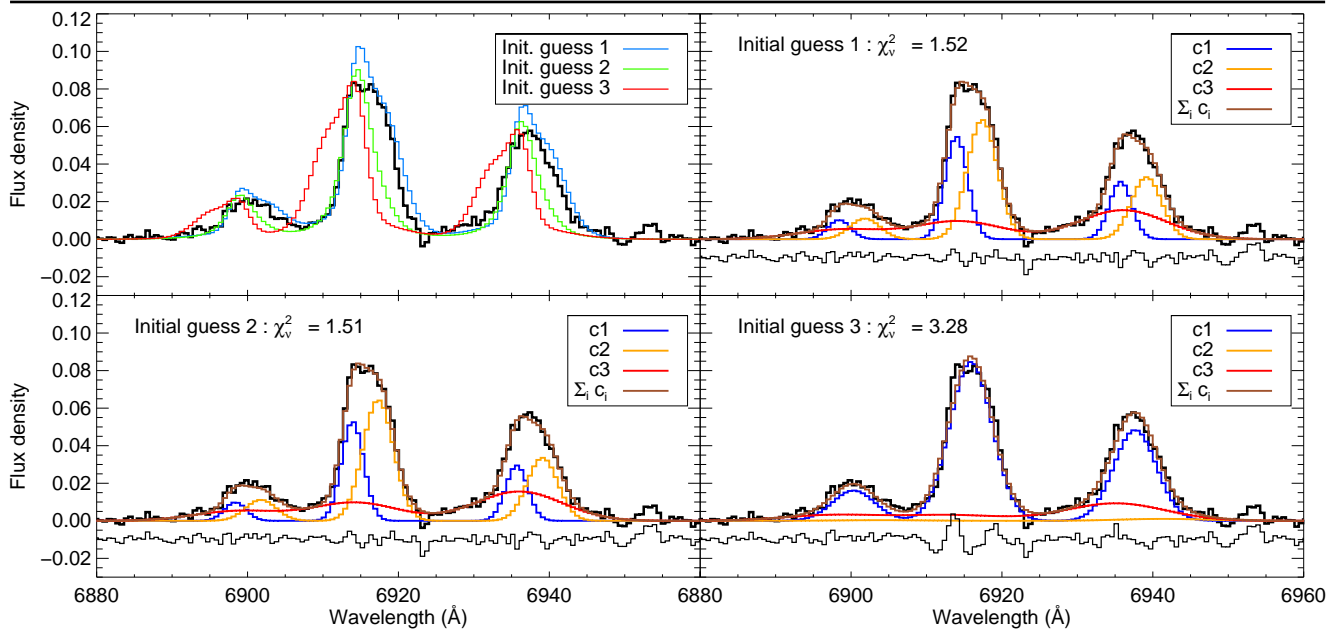
### 2.3 Emission line fitting

Emission lines are fit in the continuum subtracted spectra. The lines are assumed to be gaussian in shape and are fit as Gaussians using the Levenberg-Marquardt least-square method implemented in IDL (MPFIT; [Markwardt 2009](#)). Users have the option of fitting the emission lines using multiple Gaussian components, with this currently limited to a maximum of 3 components. All the lines are fit simultaneously with each kinematic component constrained to share

the same velocity and velocity dispersion. When more than 1 component is fit, LZIFU sorts and groups the fitting results based on either velocity dispersions, velocities or line fluxes, and produces 2D maps of fluxes, velocity, and velocity dispersions separately for different components. Which reference value is adopted to group and sort the different components is determined by the users, and we encourage the users to consider carefully what sorting methods are best for their specific science goals. In the rest of the paper, we sort and group the components based on velocity dispersion. That is: the first component ( $c1$ ) is the Gaussian fit with the



**Fig. 2** A schematic illustrating the definition of the two sets of parameters ( $f_{A12}$ ,  $\Delta v_{12}$ ,  $\Delta\sigma_{12}$ ) and ( $f_{A13}$ ,  $\Delta v_{13}$ ,  $\Delta\sigma_{13}$ ) controlling the initial guesses of the second  $c2$  (intermediate) and the third  $c3$  (broad) kinematic component relative to the first (narrow) kinematic component  $c1$ .



**Fig. 3** Demonstration of the importance of using multiple initial guesses to reject local minima. The data (thick black lines; from [Ho et al. 2014](#)) are fit with three different initial guesses to perform 3-component fitting to the [N II]  $\lambda\lambda 6548,83$  and H $\alpha$  lines. The three different initial guesses are shown in the upper right panel (see Section 2.3.1). The best-fits are shown in the other three panels, with the resulting reduced- $\chi^2$  ( $\chi_v^2$ ) labeled in each panel. The first and the second initial guesses yield very similar fits and  $\chi_v^2$ , but the third initial guess arrives a very different solution with much worse  $\chi_v^2$  (clearly visible from the residuals shown as thin black lines).

narrowest velocity dispersion, and the second ( $c2$ ) and third ( $c3$ ) components have increasing velocity dispersions.

### 2.3.1 Establishing initial guesses

Establishing proper initial guesses for the model parameters is critical when using the Levenberg-Marquardt least-square algorithm, because the initial guesses serve as starting points for the algorithm to explore the  $n$ -dimensional  $\chi^2$  space along its negative gradients.

In LZIFU, initial guesses are established automatically by means of an internal algorithm (for the first component  $c1$ ) and several external parameters determined by the user (for the second  $c2$  and third  $c3$  components). The internal algorithm searches for peak S/N in the spectrum to determine the central wavelengths and amplitudes of the first kinematic components. The redshift of the galaxy input by the user allows LZIFU to estimate the rough locations of the emission lines. The locations are updated if sensible stellar velocity and velocity dispersion can be obtained from the stellar continuum fit. Channels around  $\pm 300$  km s $^{-1}$  from the fiducial location of each emission line are inspected, and the channel with the highest signal-to-noise ratio (S/N) determines the amplitude guess of the first Gaussian component. When fitting multiple emission lines, the line with the best S/N anchors the initial wavelength guess of the first Gaussian component. The width of the first Gaussian component is

provided by the user. For the SAMI Galaxy Survey data we typically adopt a width of 50 km s $^{-1}$ .

When fitting more than one component, how and where to place the second (and third) kinematic components are determined by a set of parameters specified by the user. We use three parameters ( $f_A, \Delta v, \Delta \sigma$ ) to describe the relationship between the second (or third) Gaussian component(s) and the first component. Figure 2 illustrates the definitions of the parameters. The two sets of parameters ( $f_{A12}, \Delta v_{12}, \Delta \sigma_{12}$ ) and ( $f_{A13}, \Delta v_{13}, \Delta \sigma_{13}$ ) control the initial guesses of the second and third component, respectively. The combined profile (grey curve in Figure 2) is normalised to the peak value of the data before proceeding to solve the least-square problem.

Fitting multiple components can sometimes be sensitive to the choice of initial guesses, particularly when the S/N is poor or the spectrum is only marginally resolved. As a result, LZIFU allows multiple initial guesses to be generated by providing arrays of possible  $f_A, \Delta v$  and  $\Delta \sigma$  values. All possible combinations of initial guesses are solved for least-square solutions with MPFIT, and the fit with the best minimum  $\chi_v^2$  is kept as the final solution.

We demonstrate the importance of using multiple initial guesses to reject local minima in Figure 3. Different initial guesses are adopted to model the [N II]  $\lambda\lambda 6548,83$  and H $\alpha$  lines. The spectrum comes from a galaxy observed by the SAMI Galaxy Survey and presented in [Ho et al. \(2014\)](#). In



the upper-left panel, three different initial guesses (color-coded lines) are generated. The second components are characterised by  $f_{A12} = 1$ ,  $\Delta\sigma_{12} = 60 \text{ km s}^{-1}$ , and  $\Delta v_{12}$  with three possible values of 150, 30, and  $-150 \text{ km s}^{-1}$ . The initial guesses of the third component are all the same of  $f_{A13} = 0.2$ ,  $\Delta\sigma_{13} = 200 \text{ km s}^{-1}$ , and  $\Delta v_{13} = -10 \text{ km s}^{-1}$ . The first component has an initial velocity dispersion of  $50 \text{ km s}^{-1}$ . After applying the Levenberg-Marquardt least-square algorithm, the first and second initial guesses arrive at solutions virtually indistinguishable with very similar  $\chi^2_\nu$  of 1.51 and 1.52, respectively. The third initial guess, however, arrives at a very different solution with much higher  $\chi^2_\nu$  of 3.28. A careful visual inspection of neighbouring spaxels reveals that there are indeed three separate kinematic components in this galaxy, but as the first and second kinematic components are spectrally close to each other in this spaxel, fitting the two narrow peaks with a single Gaussian yields a local minimum (the bottom right panel). This local minimum can be rejected by fitting with different initial guesses.

Our implementation of multiple initial guesses has several advantages. Empirical understanding of the physical characteristics of the second and third components can be incorporated directly into guiding the fits by providing proper sets of parameters ( $f_A$ ,  $\Delta v$  and  $\Delta\sigma$ ). In principle, the  $n$ -dimensional  $\chi^2$  space will be explored thoroughly if a chain of initial guesses is carefully chosen. Our algorithm trades computational expense against sensitivity to local minima. This makes the analysis of extended data sets tractable.

### 2.3.2 Optional refit with smoothed initial guesses

Optional refits are possible after fitting the data with the default initial guesses. In the refitting process, results from the first-pass fit to the full data cube are spatially median-smoothed to produce initial guesses to refit the data. The refitting process can be repeated multiple times.

The reasoning behind the refitting process is that flux, velocity and velocity dispersion usually vary smoothly across the spatial dimensions, a direct result of the intrinsic properties of galaxies and the finite spatial resolution of the data. Therefore, the best fits of neighbouring spaxels contain information useful for establishing a proper initial guess. Our experience shows that the refitting process is useful for rejecting some bad results caused by ill-chosen initial guesses in the previous fits, and those poor fits triggered by the presence of local defects in the spectra (poor sky-line subtraction, uncleaned cosmic ray residuals, etc.) which force the initial guess solution into a local minima of limited relevance.

## 2.4 Output

The final products delivered by LZIFU are stored in multi-extension FITS files. For a more detailed description of the

output data structure, the readers are directed to the readme file included in the code release package. In brief, LZIFU generates 3-dimensional model cubes and 2-dimensional maps from the fitting. The model cubes include both the continuum models and emission line models. All the model cubes have the same spatial and spectral dimensions as the input data cubes. These model cubes are not only useful for visualising the fits, but also for removing emission lines or continuum from the data cubes (i.e. for generating line-free or continuum-free data cubes). Emission line fluxes, velocities, and velocity dispersions (and corresponding errors) are stored in 2-dimensional maps. These 2-dimensional maps have the same spatial dimensions as the input data cubes. These maps are most useful for subsequent scientific analysis, e.g. converting line fluxes to star formation rates or extinction, producing emission line ratio maps, fitting disk models to the velocity field, etc.

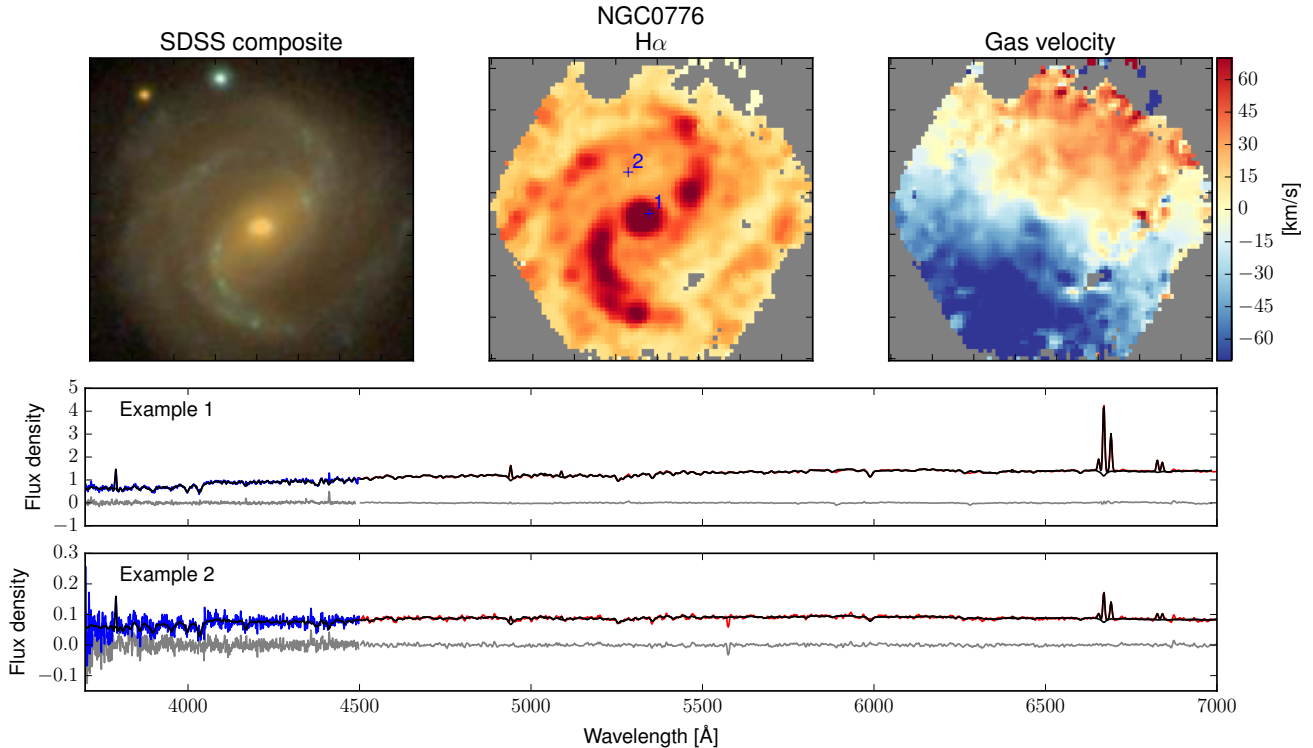
## 3 Applications on Survey Data

We present three examples of flux and kinematic maps generated by LZIFU using public data from the CALIFA survey and the SAMI Galaxy Survey. The examples are chosen to demonstrate the use of LZIFU in different types of data and galaxies. The simple single component analysis is useful for dynamically stable systems or when the spectral resolution is insufficient to resolve the kinematic structures. The more complicated double and triple component analyses are required when the line profiles are skewed due to either complex gas kinematics or beam smearing.

Although the examples below make use of CALIFA and SAMI data, LZIFU is not limited to these two surveys and can be adopted for any IFS data with similar characteristics, i.e. spectral coverage and resolution. Indeed, LZIFU has already been applied to data from multiple IFU instruments for various science cases related to gas physics. Data from the WiFeS instrument on the Australian National University 2.3-m telescope have been tested extensively with LZIFU (Ho et al. 2015; Dopita et al. 2015a,b; Vogt et al. 2015; Medling et al. 2015). Recently, LZIFU has also been adopted to analyse data from MUSE (Kreckel et al. 2016; Juneau et al. in preparation) and SPIRAL on the Anglo-Australian Telescope (McElroy et al. 2015). The reader is referred to corresponding publications for more examples.

### 3.1 Simple single component analysis

We demonstrate 1-component fitting using data from the CALIFA survey (i.e., the first data release; Husemann et al. 2013). Figure 4 shows the NGC0776 SDSS  $g,r,i$  colour composite image, H $\alpha$  map, gas velocity map, and two example spectra from the integral field data. Here, we adopt



**Fig. 4** Demonstration of applying a 1-component fit to NGC0776 using the first data release of the CALIFA survey. The SDSS  $g,r,i$  composite image, LZIFU  $H\alpha$  and gas velocity images are shown left to right in the top row. The spectra and corresponding best-fit (continuum + line) models of the two example spaxels marked in the  $H\alpha$  map are shown in the bottom two panels. We use the V1200 data (blue lines) at wavelengths smaller than  $4500\text{\AA}$  and the V500 data (red lines) at wavelengths larger than  $4500\text{\AA}$ .

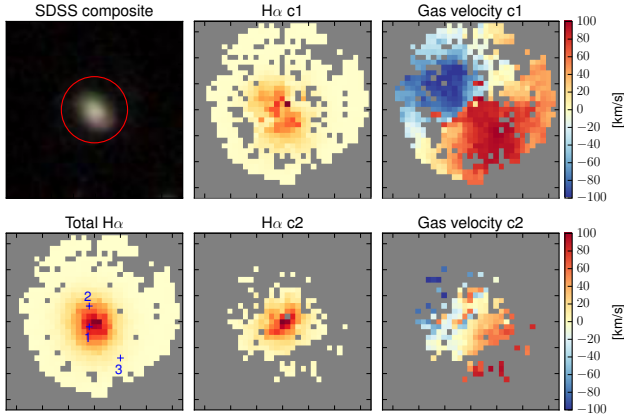
the MIUSCAT SSP libraries (Vazdekis et al. 2012) of solar metallicity to model the continuum. After subtracting the continuum, the line profiles appear to be simple across the entire galaxy and therefore only single component Gaussians are required to model the emission lines. The emission line maps and kinematic maps delivered by LZIFU are directly ready for various studies such as gas dynamics and chemical abundance.

### 3.2 Double component fitting and beam smearing

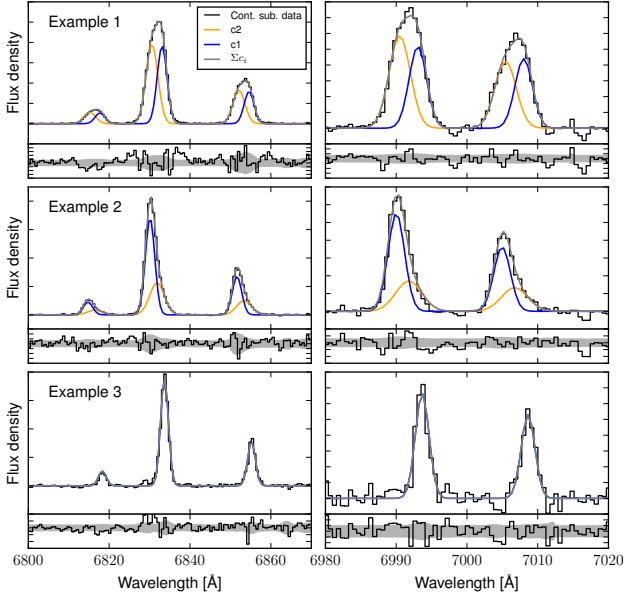
Multiple-component fitting is sometimes required when the spectral resolution is high enough to resolve the intrinsic line profiles. Fitting multiple Gaussian components to an emission line can more accurately constrain its total flux (than fitting a single component Gaussian), and shed light on the possible complex dynamics of the gas traced by the emission line. The number of components required to properly describe the line profile depends on the spectral resolution of the instrument, the signal-to-noise of the data, and the gas dynamics. Typically, one performs 1, 2, and 3-component fitting on every spaxel, and uses both statistical and empirical tests to determine a posteriori the most appropriate numbers of components required to describe the data. The num-

ber of components required frequently changes from spaxel to spaxel within a single galaxy.

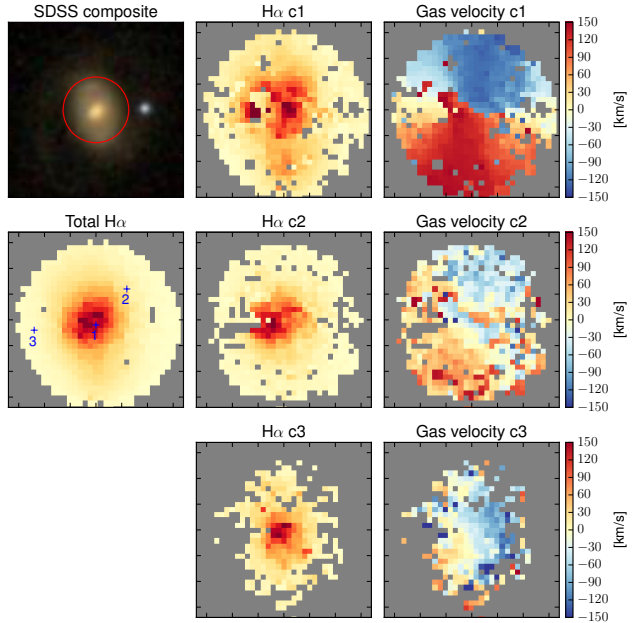
In Figures 5 and 6, we show an example of spectral decomposition using data from the early data release of the SAMI Galaxy Survey (Allen et al. 2015a). After subtracting the continuum, the galaxy (GAMA ID: 594906) presents skewed line profiles changing with position in the galaxy (Figure 6). We perform 1, 2, and 3-component fitting, and determine the number of components required based on the likelihood ratio test and empirical constraints described in detail in Appendix A. We show flux and velocity maps of the first and second components in Figure 5. Only a few spaxels require fitting the third component so we do not show the corresponding maps. In Figure 5, the first component presents a regular rotation pattern tracing the galactic disk. The second component shows a velocity gradient in the same sense as the first component. Both components have similar emission line ratios. We believe that the skewed line profiles are a direct result of beam smearing, which is known to induce non-Gaussian line profiles particularly at the centre of the galaxy where the velocity gradient is steep (e.g., Green et al. 2014). Such non-Gaussian line profiles will not present in high spatial resolution or low spectral resolution observations.



**Fig. 5** Demonstration of multiple component fitting for a galaxy from the early data release of the SAMI Galaxy Survey (Allen et al. 2015a; GAMA ID: 594906). The SDSS  $g, r, i$  image in the upper left panel shows the circular field of view of the SAMI instrument (red circle;  $15''$  in diameter). The  $H\alpha$  flux and gas velocity maps of the first (narrowest component)  $c1$  and the second component  $c2$  are shown in the middle and right panels. The sum of  $H\alpha$  for all components is shown in the bottom left panel (total  $H\alpha$ ). Very few spaxels in this galaxy require the third component so the corresponding maps are not shown. The fits of the three spaxels marked in the total  $H\alpha$  map are shown in Figure 6.



**Fig. 6** Examples of the  $[N II] \lambda\lambda 6548, 83 + H\alpha$  (left column) and  $[S II] \lambda\lambda 6716, 31$  (right column) fits of the three spaxels marked in Figure 5 (in lower left ‘total  $H\alpha$ ’ panel). The continuum-subtracted data and best-fit models are shown as solid lines. In the lower plot of each panel, we show the residuals as black lines, and the  $\pm 1\sigma$  measurement errors as grey shading.

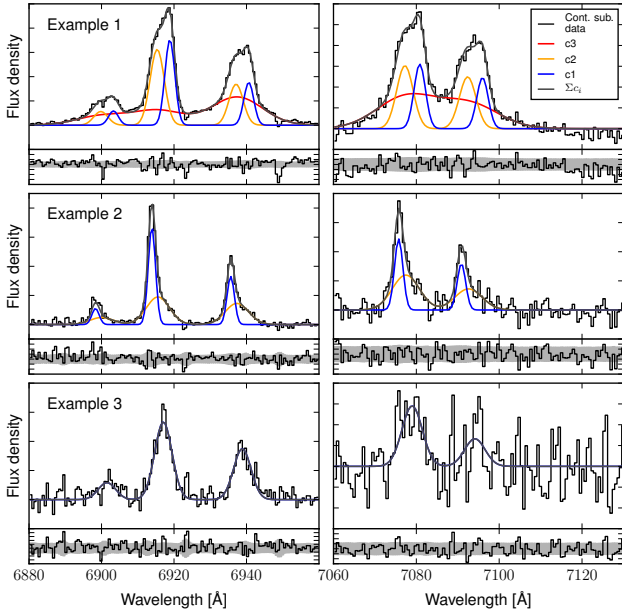


**Fig. 7** Demonstration of multiple component fitting for a normal star-forming galaxy presented in Ho et al. (2014, GAMA ID: 209807). The SDSS  $g, r, i$  image in the upper left panel shows the circular field of view of the SAMI instrument (red circle;  $15''$  in diameter). The  $H\alpha$  flux and gas velocity maps of the first (narrowest component)  $c1$ , the second component  $c2$ , and the third component  $c3$  are shown in the middle and right panels. The sum of  $H\alpha$  for all components is shown in the middle left panel (total  $H\alpha$ ). The fits of the three spaxels marked in the total  $H\alpha$  map are shown in Figure 6.

### 3.3 Triple component fitting and kinematics

In more complex, dynamically active systems such as galaxies hosting galactic winds, AGNs, or mergers, more components are required to capture the activities of the gas. We present an example of multiple-component fitting by Ho et al. (2014) using data from the SAMI Galaxy Survey. As in GAMA 594906, Ho et al. (2014) performed 1, 2, and 3-component fitting, and determined the number of components required based on the likelihood ratio test. The normal star-forming galaxy (GAMA ID: 209807) analysed by Ho et al. (2014) hosts large scale galactic winds. Figure 7 shows the  $H\alpha$  maps and velocity maps of the three different kinematic components, and Figure 8 shows some example spectra requiring different numbers of components. Similarly, the velocity field of the first component shows a regular rotation pattern tracing the galactic disk. Ho et al. (2014) showed that the first component has line ratio consistent with photoionisation originating from star forming regions on the disk. The third component has line ratios consistent with pure shock excitation, indicating the presence of fast winds driven by a central starburst. The second kinematic





**Fig. 8** Examples of the [N II]  $\lambda\lambda 6548,83 + \text{H}\alpha$  (left column) and [S II]  $\lambda\lambda 6716,31$  (right column) fits of the three spaxels marked in Figure 7 (in middle left ‘total  $\text{H}\alpha$ ’ panel). The continuum-subtracted data and best-fit models are shown as solid lines. In the lower plot of each panel, we show the residuals as black lines, and the  $\pm 1\sigma$  measurement errors as grey shading.

component is excited by both photoionisation and shock excitation. The clear velocity gradient of the third component nearly aligned with the minor axis of the galaxy (bottom-right and top-left panel of Figure 7) traces the large scale, bipolar galactic winds in the galaxy.

#### 4 Error Analysis with Monte Carlo Simulations

LZIFU reports  $1\sigma$  errors of the measured fluxes, velocities and velocity dispersions of the emission lines calculated with the Levenberg-Marquardt least-square method from MPFIT. To investigate the reliability of these quantities, we perform simple Monte Carlo (MC) simulations. In these simulations, we create different MC realisations (i.e. mock data cubes) by injecting Gaussian noise into model cubes based on the variance of the data. For each test galaxy (selected from the SAMI Galaxy Survey), 500 sets of mock data cubes are generated, and the mock data cubes are each fit with LZIFU. Each fit yields measurements of flux, velocity, and velocity dispersion maps and their corresponding error maps. To quantify the reliability of LZIFU errors, we compare the  $1\sigma$  spread of the 500 measurements to the median of their errors.

Two types of simulations are performed. Firstly, we inject noise into the best-fit emission line models to test only

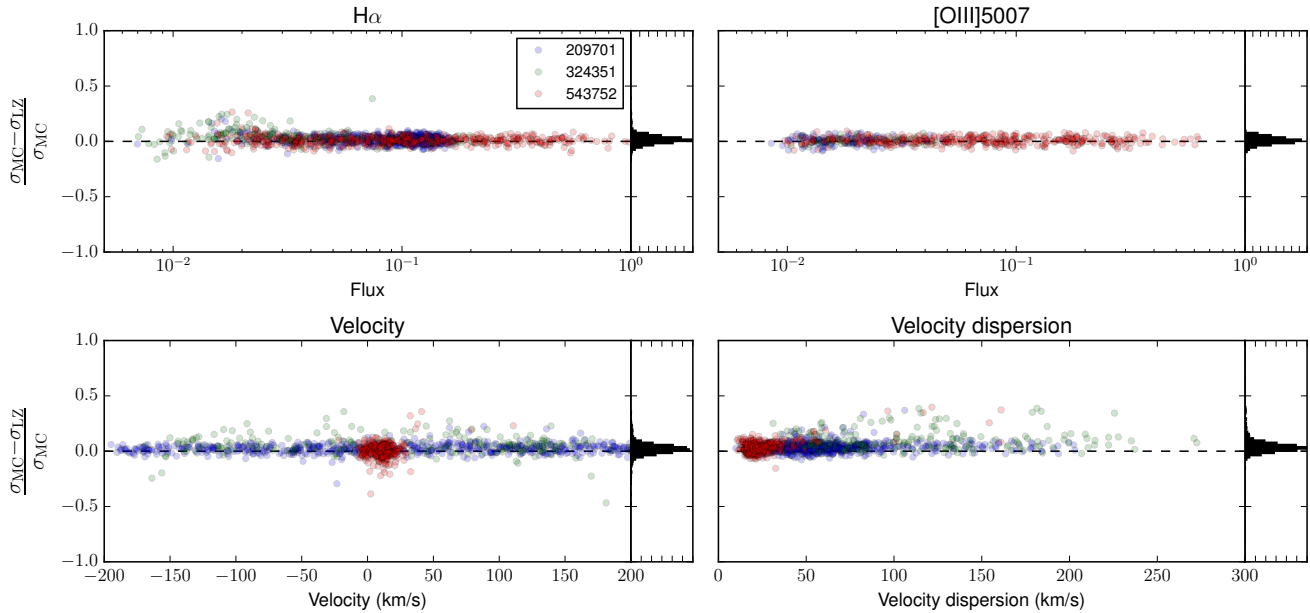
the emission line fitting codes. The mock data cubes are continuum-free so no continuum subtraction is performed. Secondly, we inject noise into cubes of emission models plus continuum models. The purpose of this test is to explore errors caused by modelling and subtracting the continuum. The goal of these simulations is to study whether LZIFU can faithfully propagate the random errors in mock data cubes to the final measured quantities.

##### 4.1 Line-fitting simulations

In Figure 9, we compare the errors derived from MC simulations ( $\sigma_{MC}$ ) to the errors reported by LZIFU ( $\sigma_{LZ}$ ). In these MC simulations, we fit 1-component models to three SAMI galaxies, and we derive  $\sigma_{MC}$  using resistant estimates of the dispersions of the distributions (using the ROBUST\_SIGMA routine in IDL). For  $\sigma_{LZ}$ , we take the median errors of the 500 MC simulations. We do not include the error in the median calculation when the S/N of velocity dispersion is less than 3. In Figure 9, we show spaxels with  $\text{H}\alpha$  flux S/N > 3 in the  $\text{H}\alpha$ , velocity and velocity dispersion panels, and spaxels with [O III]  $\lambda 5007$  flux S/N > 3 in the [O III]  $\lambda 5007$  panel. Figure 9 demonstrates that the flux, velocity and velocity dispersion errors reported by LZIFU agree well with the errors derived from our MC simulations. Typically, the differences between  $\sigma_{LZ}$  and  $\sigma_{MC}$  are consistent with zero (i.e., median  $\pm$  standard deviation of  $2 \pm 4\%$ ,  $1 \pm 3\%$ ,  $3 \pm 6\%$ , and  $4 \pm 6\%$  for  $\text{H}\alpha$ , [O III]  $\lambda 5007$ , velocity and velocity dispersion, respectively). The results of this test indicate that the line-fitting codes faithfully propagate errors in the data cubes to the final measurement errors.

In Figure 10, similar comparisons are conducted for a 2-component fit using the SAMI wind galaxy studied by Ho et al. (2014). Only spaxels requiring 2-component fits determined by the authors are considered. We find that the flux, velocity and velocity dispersion errors reported by LZIFU are good representations of the true errors derived from MC simulations. On average, the differences between  $\sigma_{LZ}$  and  $\sigma_{MC}$  for  $\text{H}\alpha$ , [O III]  $\lambda 5007$ , velocity, and velocity dispersion are  $13 \pm 13\%$ ,  $2 \pm 9\%$ ,  $9 \pm 12\%$ , and  $17 \pm 11\%$  (median  $\pm$  standard deviation), respectively. The differences are larger than those in the 1-component cases with LZIFU typically underestimating the errors by 10% to 20%.

The differences between  $\sigma_{MC}$  and  $\sigma_{LZ}$  arise from the assumptions involved in deriving errors of the fit parameters using the least-square technique. The Levenberg-Marquardt algorithm approximates the  $\chi^2$  surface at minimum by an  $n$ -dimensional quadratic function (see e.g. Bevington and Robinson 1992). This approximation is a result of the Taylor expansion at the minimum  $\chi^2$  where the first order term is zero; the second order term then becomes important for evaluating the increase in  $\chi^2$ . With this assumption, fast computation of the errors of the fit parameters becomes possible because only the second derivative of the  $\chi^2$  surface



**Fig. 9** Comparison between errors reported by LZIFU ( $\sigma_{LZ}$ ) and errors derived from Monte Carlo simulations ( $\sigma_{MC}$ ). The simulations are performed using 1-component fitting assuming the data have no continuum. Details about the simulations are provided in Section 4.1. Different color points correspond to three different galaxies selected from the SAMI Galaxy Survey, with their GAMA IDs shown in the legend. The fractional differences between  $\sigma_{LZ}$  and  $\sigma_{MC}$  are  $2 \pm 4\%$ ,  $1 \pm 3\%$ ,  $3 \pm 6\%$ , and  $4 \pm 6\%$  (median  $\pm$  standard deviation) for  $H\alpha$ ,  $[O\text{III}]\lambda 5007$ , velocity and velocity dispersion, respectively.

at its minimum is required. The second derivative is directly linked to the Jacobian of the model, which is trivial to calculate. While the approximation works well in linear models, the non-linearity of our Gaussian line model can cause the assumption to break down. For example, the Jacobian matrix of velocity dispersion approaches zero at zero velocity dispersion, implying that the  $\chi^2$  surface approaches a flat surface and the  $1\sigma$  error of velocity dispersion is infinity. When a Jacobian approaches zero, higher ( $>$  third) order Taylor terms become important. Unfortunately, higher order terms are non-trivial to calculate. Exploring the  $\chi^2$  space by random walk using techniques such as the MCMC method should be adopted if precise estimates of errors (i.e. better than  $\sim 10\%$ ) are required. The Levenberg-Marquardt technique adopted here provides errors accurate to a few tens percent level in a computationally-economical way.

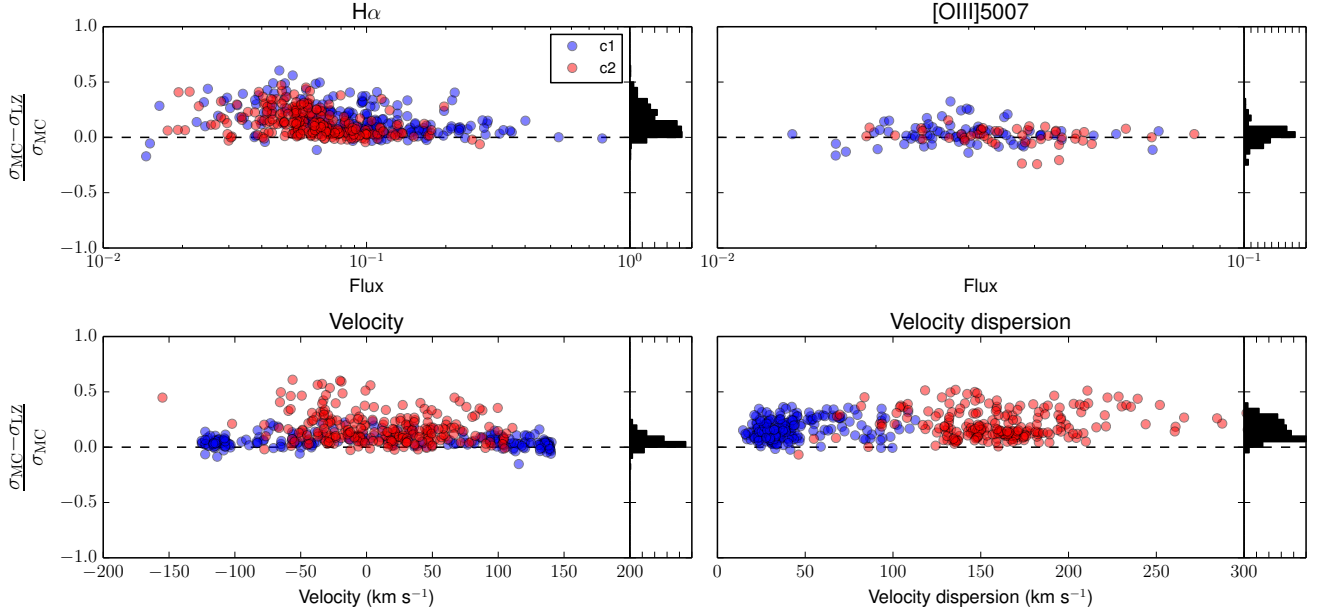
#### 4.2 Continuum- and line-fitting simulations

While the line-fitting algorithm can robustly estimate the flux, velocity, and velocity dispersion errors, these errors do not contain errors of modelling (and subtracting) the continuum. To investigate the impact of modelling continuum on the measured emission line fluxes, we perform MC simulations by injecting noise into the best-fit models that comprise both continuum and emission line models. We first fit

the real data from three SAMI galaxies to obtain their best-fit continuum and line models. For the PPXF continuum fit, we adopt the theoretical SSP libraries from [González Delgado et al. \(2005\)](#). After noise is injected into the models to produce mock data cubes, the same stellar libraries are used to fit the realisations.

In the top row of Figure 11, we compare the errors of the line fluxes as in the line-fitting simulations. The fractional differences between  $\sigma_{MC}$  and  $\sigma_{LZ}$  are  $7 \pm 4\%$ ,  $6 \pm 4\%$ , and  $2 \pm 3\%$  (median  $\pm$  standard deviation) for  $H\alpha$ ,  $H\beta$  and  $[O\text{III}]\lambda 5007$ , respectively. Comparing these results with those performed without considering the continuum (see Figure 9), the fractional difference of errors for  $[O\text{III}]\lambda 5007$  is comparable to the line-fitting simulations, but those for  $H\alpha$  and  $H\beta$  (i.e., Balmer lines) are about a factor of 2 – 3 larger.

The fundamental reason behind this discrepancy is that the errors in the best-fit continuum models are unknown, so the errors are not propagated to the continuum-subtracted spectrum. Essentially, the best-fit continuum models are assumed to be noise-free. While this assumption could be true for some emission lines, those lines at similar wavelengths to strong stellar absorption features can be affected by the continuum errors. As equivalent widths of Balmer lines are strong functions of stellar age (and weakly dependent on metallicity), when a different set of solutions (age and metallicity) is derived from the SSP fit to each realisa-



**Fig. 10** Comparison between errors reported by LZIFU ( $\sigma_{LZ}$ ) and errors derived from Monte Carlo simulations ( $\sigma_{MC}$ ). The simulations are performed on spaxels in [Ho et al. \(2014\)](#) that require 2-component fitting. There is no continuum in the simulations. Details are provided in Section 4.1. On average, the fractional differences between  $\sigma_{LZ}$  and  $\sigma_{MC}$  for  $H\alpha$ ,  $[O III] \lambda 5007$ , velocity, and velocity dispersion are  $13 \pm 13\%$ ,  $2 \pm 9\%$ ,  $9 \pm 12\%$ , and  $17 \pm 11\%$  (median  $\pm$  standard deviation), respectively.

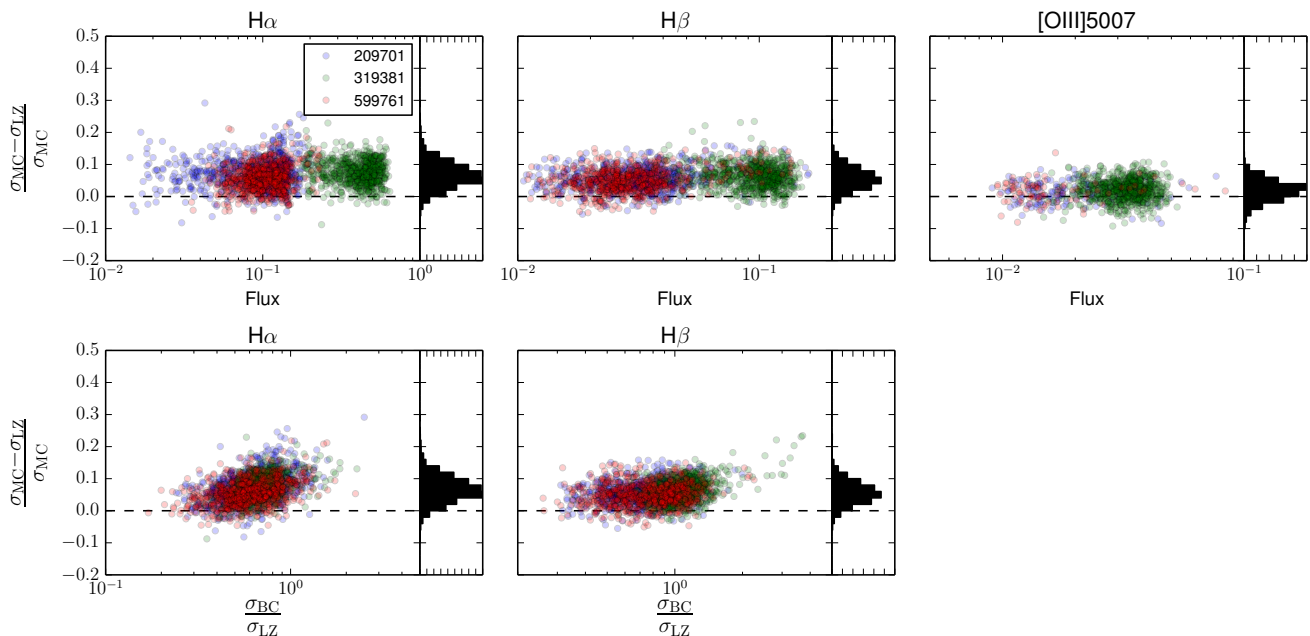
tion, different Balmer corrections (for absorption of Balmer lines from stellar atmosphere) to the emission-line result in different Balmer emission-line fluxes.

To confirm the role that stellar Balmer correction plays in the line flux errors, we compare the fractional differences between  $\sigma_{MC}$  and  $\sigma_{LZ}$  to the importance of errors in the Balmer correction relative to the line flux errors (i.e.  $\sigma_{BC}/\sigma_{LZ}$ ) as shown in the bottom row of Figure 11. Here, we define  $\sigma_{BC}$  as the standard deviation of the Balmer corrections in the 500 MC simulations. The Balmer corrections are calculated over the on-line/off-line windows defined in [González Delgado et al. \(2005\)](#). When  $\sigma_{BC}/\sigma_{LZ}$  is large, the differences of Balmer correction in different realisations are substantial compared to the nominal flux errors ( $\sigma_{LZ}$ ) so one would expect that the nominal flux errors ( $\sigma_{LZ}$ ) underestimate the real errors ( $\sigma_{MC}$ ). We observe this behaviour in the bottom row of Figure 11. Both  $H\alpha$  and  $H\beta$  show positive correlations between the fractional differences in errors (y-axis) and the importance of Balmer correction (x-axis), confirming that the continuum fitting largely causes the discrepancies in the errors.

Obtaining proper errors for the continuum models has been a long-standing problem in spectral fitting (e.g., [Koleva et al. 2008, 2009](#); [Tojeiro et al. 2007](#); [MacArthur et al. 2009](#); [Walcher et al. 2011](#); [Yoachim et al. 2012](#); [Cid Fernandes et al. 2014](#)). The difficulties come from the fact that continuum fitting is a non-linear multi-variable least-square problem with typically multiple local minima  $\chi^2$ . Quantify-

ing the errors requires performing MC simulations that can be computationally very expensive and perhaps only feasible on small-scale simulations applied to a handful of galaxies. In the context of constraining the contamination from Balmer correction, the degree of contamination is likely to depend on the spectral coverage and the spectral resolution of the data, because those factors determines the accuracy of the stellar ages and stellar metallicities from SSP fitting.

In all our MC simulations, the underlying models of the mock data are known a priori so we are only studying the propagation of statistical errors. We did not consider systematic errors, such as non-Gaussian line profiles and various uncertainties associated with the synthesis of the SSP spectral models (e.g., stellar evolutionary track, binary star, TP-AGB star, etc.), and therefore the discrepancies between nominal and real errors are lower limits. Systematic errors can be important in many applications of spectral fitting, particularly the errors between different SSP models. [Cid Fernandes et al. \(2014\)](#) analysed the uncertainties of stellar mass, age, metallicity, and extinction derived with data from the CALIFA survey using the STARLIGHT spectral fitting package. They found that the dominant uncertainties come from the choice of SSP models. For emission line fitting, the equivalent widths of Balmer lines between different models can disagree on the level of a few tens of percent (see figure 1 in [Groves et al. 2012](#)), which can cause systematic errors on the weak, high order Balmer lines. These errors are particularly important when the continuum is strong rel-



**Fig. 11** Comparison between errors reported by LZIFU ( $\sigma_{LZ}$ ) and errors derived from Monte Carlo simulations ( $\sigma_{MC}$ ). The simulations include both line and continuum fitting. The emission lines are 1-component and we assume no systematic errors in the continuum models. Details about the simulations are provided in Section 4.2. Different color points correspond to different galaxies from the SAMI Galaxy Survey, as shown in the legend their GAMA IDs. In the top three panels, we compare the differences between  $\sigma_{MC}$  and  $\sigma_{LZ}$  with line fluxes; the fractional differences between  $\sigma_{MC}$  and  $\sigma_{LZ}$  are  $7 \pm 4\%$ ,  $6 \pm 4\%$ , and  $2 \pm 3\%$  (median  $\pm$  standard deviation) for H $\alpha$ , H $\beta$  and [O III]  $\lambda$ 5007, respectively. In the bottom two panels, we compare the fractional differences between  $\sigma_{MC}$  and  $\sigma_{LZ}$  to the importance of errors in Balmer correction relative to line flux errors, i.e.  $\sigma_{BC}/\sigma_{LZ}$ . The positive correlations demonstrate that continuum fitting could impact the Balmer line errors when  $\sigma_{BC}/\sigma_{LZ}$  is large.

ative to the emission lines. Dedicated studies are required to explore the different systematic effects involved in fitting the continuum.

## 5 Summary and Conclusion

We have presented LZIFU, an IDL toolkit for fitting multiple emission lines and constructing emission line flux maps and kinematic maps from IFS data. We outlined the structure of LZIFU, and described in detail how the code performs spectral fitting and decomposition. We have also conducted simulations to examine the errors estimated by LZIFU and discussed its limitations.

We have demonstrated how LZIFU can be adopted to analyse data from the CALIFA survey and the SAMI Galaxy Survey. In some applications, single component fitting is adequate to capture the dominant kinematic component (typically from H II regions tracing disk rotation) and can produce flux and kinematic maps useful for various studies of gas physics. In cases where the line profiles are more sophisticated due to either active physical environments (e.g., AGN or galactic wind) or observational effects (e.g., beam

smearing), multiple component fitting can better constrain the total line fluxes and provide more insight into the various physical processes. Although only examples from CALIFA and SAMI were presented in the paper, LZIFU is by no means limited to these two datasets. Data from world-class IFS instruments with distinct structures (i.e. fibre-based, image-splitting), sizes and spatial resolutions have already been processed by LZIFU, including MUSE, WiFeS and SPIRAL. The LZIFU products and scientific results extracted from these can be found in Ho et al. (2015); Dopita et al. (2015b); McElroy et al. (2015); Vogt et al. (2015); Kreckel et al. (2016). Further data from these instruments, and surveys from other IFS instruments are currently being analysed, with a wealth of scientific results from LZIFU products expected to be published in the coming years.

While this paper outlines the official release version of LZIFU, future improvements to the pipeline will be implemented. For example, in a soon-to-be available upgrade we will include the option of fitting binned data. Spatially binning data can significantly improve the detection of faint emission lines at large galactic radii. Different binning schemes such as contour binning (Sanders 2006) and Voronoi tessellations (Cappellari and Copin 2003) have es-

established the usefulness of binning imaging and IFS data. On longer timescales, we plan to incorporate a full Bayesian analysis such that the parameter space can be explored more thoroughly and more accurate errors can be reported. It is also possible to analyse mock 3D data cubes from numerical simulations parallel to observational data cubes to directly compare similar parameter maps.

### **Acknowledgement**

We thank the referee for constructive comments that improve the quality of this work. LJK gratefully acknowledges the support of an ARC Future Fellowship, and ARC Discovery Project DP130103925. SMC acknowledges the support of an Australian Research Council Future Fellowship (FT100100457). The SAMI Galaxy Survey is based on observations made at the Anglo-Australian Telescope. The Sydney-AAO Multi-object Integral field spectrograph was developed jointly by the University of Sydney and the Australian Astronomical Observatory. The SAMI input catalogue is based on data taken from the Sloan Digital Sky Survey, the GAMA Survey and the VST ATLAS Survey. The SAMI Galaxy Survey is funded by the Australian Research Council Centre of Excellence for All-sky Astrophysics, through project number CE110001020, and other participating institutions. The SAMI Galaxy Survey website is <http://sami-survey.org/>. This study makes use of the data provided by the Calar Alto Legacy Integral Field Area (CALIFA) survey (<http://califa.caha.es/>). Based on observations collected at the Centro Astronómico Hispano Alemán (CAHA) at Calar Alto, operated jointly by the Max-Planck-Institut für Astronomie and the Instituto de Astrofísica de Andalucía (CSIC).



## A Selecting the number of components required

The number of components required to describe a given spectrum is determined by three factors: 1) the spectral resolution of the instrument, 2) the intrinsic kinematic structure of the line emitting gas, and 3) the S/N of the data. The standard statistical test for model comparison (for the frequentist) is the likelihood ratio test (LRT; see also Section 4.2 in Ho et al. 2014). In spectral decomposition where the different models are nested, the (natural) logarithmic maximum likelihood ratio of the two models ( $n$ - and  $[n+1]$ -component models),

$$\Lambda = -2 \ln \frac{\max(L_n)}{\max(L_{n+1})} = \chi_n^2 - \chi_{n+1}^2, \quad (\text{A1})$$

is an objective gauge of how much improvement in maximum likelihood,  $\max(L_n)$  and  $\max(L_{n+1})$ , the more sophisticated model can offer. Here,  $\chi_n^2$  and  $\chi_{n+1}^2$  are  $\chi^2$  values of the best fit models with  $\nu_n$  and  $\nu_{n+1}$  degrees of freedom, respectively.  $\Lambda$  follows a  $\chi^2$  distribution of  $(\nu_n - \nu_{n+1})$  degrees of freedom, and therefore the null hypothesis that the  $n$ -component model is better than the  $(n+1)$ -component model can be tested by comparing the measured  $\Lambda$  with the critical  $\Lambda$  corresponding to the probability  $p$ -value.

Another common statistical test for model comparison is the  $F$ -test using the  $F$ -distribution. An  $F$ -distribution is formed when one takes the ratio of two random variates,  $U_1$  and  $U_2$ , that are  $\chi^2$  distributed. That is,

$$X \equiv \frac{U_1/\nu_1}{U_2/\nu_2} \quad (\text{A2})$$

follows a  $F$ -distribution of  $(\nu_1, \nu_2)$  degree of freedom when 1)  $U_1$  and  $U_2$  are  $\chi^2$  distributed with  $\nu_1$  and  $\nu_2$  degrees of freedom, respectively; and 2)  $U_1$  and  $U_2$  are *independent*. The  $F$ -test applied in many astrophysical experiments uses the fact that when  $\Lambda$  follows a  $\chi^2$  distribution of  $(\nu_n - \nu_{n+1})$  degrees of freedom, and the index  $\tilde{F}$ , defined as

$$\tilde{F} \equiv \frac{(\chi_n^2 - \chi_{n+1}^2)/(\nu_n - \nu_{n+1})}{\chi_{n+1}^2/\nu_{n+1}} = \frac{\Lambda/(\nu_n - \nu_{n+1})}{\chi_{n+1}^2/\nu_{n+1}}, \quad (\text{A3})$$

follows a  $F$ -distribution of  $(\nu_n - \nu_{n+1}, \nu_{n+1})$  degrees of freedom. This is because in the denominator,  $\chi_{n+1}^2$  also follows a  $\chi^2$  distribution of  $\nu_{n+1}$  degrees of freedom. Once the distribution of  $\tilde{F}$  is predicted, a statistical significance  $p$ -value can be calculated and applied as in the LRT. It is not obvious that the numerator and denominator are independent in the case of nested models, particularly with multiple-Gaussian models for emission line fitting.

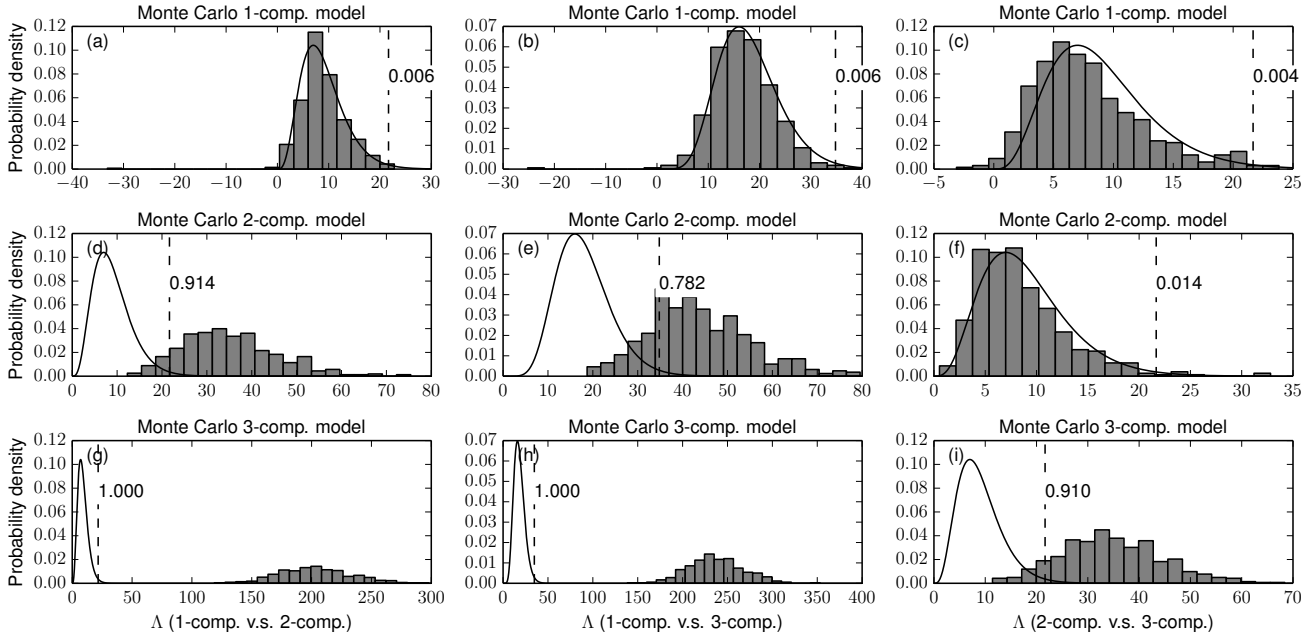
Protassov et al. (2002) point out that the use of the LRT and  $F$ -test are not statistically justified in many line-fitting applications because of the boundary conditions of non-negative line fluxes imposed on the models. The likelihood

ratio therefore does not necessarily follow the same asymptotic behaviour predicted by the  $\chi^2$  distribution. To test whether the LRT and  $F$ -test can be used on our spectral decomposition, we perform simple Monte Carlo simulations to probe the real distribution of  $\Lambda$  and  $\tilde{F}$  for our application. The simulations are designed to study *only* the statistical aspects of the problem. We first randomly select three spaxels from the SAMI galaxy presented in (Ho et al. 2014; GAMA ID: 209807), one in each region requiring different numbers of components (see their figure 4). We inject noise into the best-fit emission line models based on the variance of the real data. The three perturbed spectra are fed to LZIFU to perform 1, 2, and 3-component fits in the same way as processing real data. Since the fake spectra are already continuum-free, we do not perform continuum fitting and subtraction. Each spectrum is perturbed 500 times and fit 1,500 times (i.e. 1, 2, and 3-component fit); and we record the  $\chi^2$  and degrees of freedom of each fit. We compare the distributions of  $\Lambda$  and  $\tilde{F}$  from the Monte Carlo simulations to the expected  $\chi^2$  and  $F$ -distributions. Given that we know a priori the number of components required, we then assess how well the statistical tests perform.

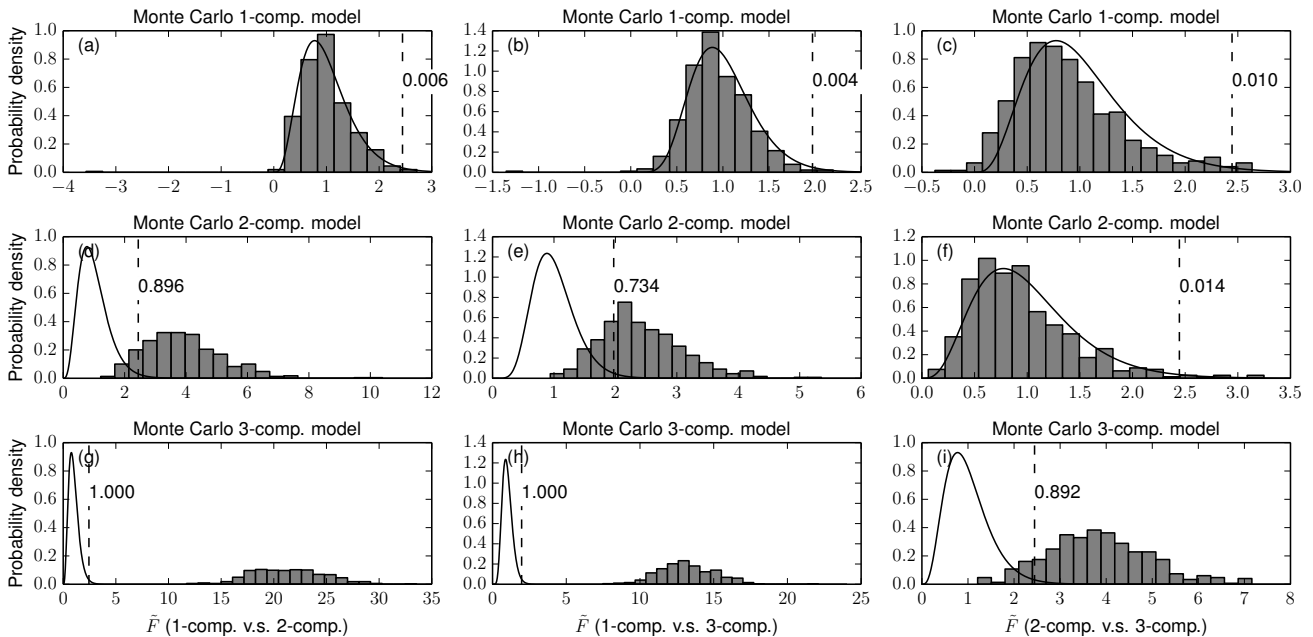
Figure 12 shows the probability density distributions of  $\Lambda$  of the models being 1-component (first row), 2-component (second row), and 3-component (third row). Both the distribution of the 500 Monte Carlo realisations and the theoretical  $\chi^2$  distribution are shown. The vertical dashed lines indicate the positions of the  $p$ -value of 0.01 determined from the theoretical distributions. The number labeled next to each dashed line indicates the actual  $p$ -value determined from the Monte Carlo results. In other words, the dashed lines mark the critical  $\Lambda$  below which the more sophisticated model should be rejected at a significance level of 0.01, whereas numbers indicate those derived from the actual distributions determined from the simulations. Similar comparisons for the  $F$ -test are shown in Figure 13.

In panels (a), (b), (d), (f), (h) and (i) of Figures 12 and 13, the model fits with the real numbers of components (“real answers”) are involved in the comparisons. These distributions demonstrate that both the LRT and the  $F$ -test are appropriate tests for spectral decomposition. At the significance cut of 0.01, panels (a), (b), and (f) give comparable significant levels ( $\approx 0.004$ – $0.014$ ) at the left tails of the distributions. Panel (d) and (i) give higher false classification rates ( $\approx 10\%$ ) at the right tails of the actual distributions (rather than 1%), and panel (h) shows a perfect classification rate. The results between LRT and  $F$ -test are consistent.

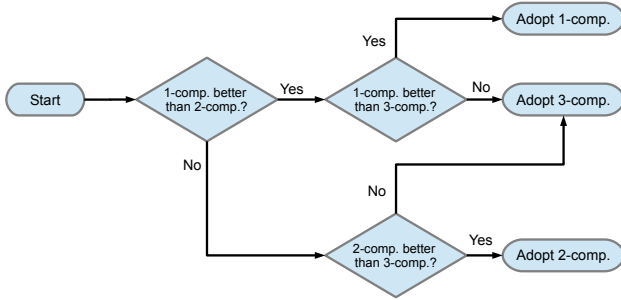
In panels (c), (e) and (g) of Figures 12 and 13, the model fits with the real numbers of components are not involved in the comparisons, i.e. one uses the wrong models to test the data. Situations like this are unavoidable since one does not know a priori the true numbers of components. It is worth pointing out that in (g), where the underlying model (“real



**Fig. 12** Distributions of  $\Lambda$  (see Equation A1) from Monte Carlo realisations (histograms) and idealised  $\chi^2$  distributions (curves) predicted by the likelihood ratio test. The different panels compare the different underlying models (top to bottom: true numbers of components are 1, 2, and 3) with 500  $\Lambda$  computed from fitting each realisation with 1, 2, and 3-component models (as labeled on the x-axes). The vertical dashed lines indicate the positions of  $p$ -value of 0.01 determined from the theoretical distributions, and the number labeled next to each dashed line indicates the actual  $p$ -value determined from the Monte Carlo realisations. In the 1-component case (top row), the simpler model is preferred in each case. In the 2-component case (middle row), the distributions favour models more complex than 1-component (left + centre panels), but simpler than 3 (right panel). In the 3-component spaxel (bottom row), the more complex model is preferred.



**Fig. 13** Same as Figure 12 but for  $F$ -test. The curves are  $F$ -distributions from the  $F$ -test (Equation A3).



**Fig. 14** Flowchart for classifying the number of components required for a spectrum. Each spectrum is fit with 1, 2, and 3-component Gaussians, and decisions (diamonds) are made by performing the likelihood ratio test or  $F$ -test.

answer”) has 3 components, the statistical tests strongly prefer the more sophisticated 2-component fits. In (c), where the underlying model has only 1 component, the statistical tests strongly prefer the simpler 2-component fits. In (e), where the underlying model has 2 components and the statistical tests are choosing between 1-component and 3-component, the theoretical  $p$ -value does not provide useful assessment.

To quantify the overall performance of the statistical tests, we use the flowchart in Figure 14. For a given spectrum, we first choose between the 1-component and 2-component fits, and then compare the preferred fit with the 3-component fit. We classify the 500 Monte Carlo realisations on the three spaxels and we find that, with the LRT, the successful rates for classifying 1-component, 2-component, and 3-component fits are 99%, 90% and 91%, respectively. With the  $F$ -test, similar results are found of 99%, 89% and 89%, respectively.

Although the Monte Carlo results imply that these statistical tests provide robust classification, in practice it is extremely difficult to have accurate estimates of the variance, which means the  $\chi^2$  values may be problematic. Apart from the difficulties in propagating Poisson noise stringently from raw data to reduced data cubes, continuum modelling always carries some statistical and systematic uncertainties that are difficult to quantify and propagate. Strong sky lines could also cause the wrong estimate of variance in sky-dominated channels and/or strong residuals due to imperfect sky subtraction. These factors limit the reliability of these statistical tests, and therefore these tests should be used with great care. Additional means of quality control are always recommended, such as visual inspection and consistency checks of physical parameters.

Our experience shows that in the regime where residuals from systematic effects are smaller than the noise levels, the statistical tests classify spectra in good agreement with human judgement. However, in the regime where the

noise levels are much lower than systematic effects, more sophisticated models are always preferred by the statistical tests and the classifications may not be physical. For example, when the S/N is excellent and the surrounding channels are slightly positive in the continuum-subtracted spectrum due to errors in the continuum fit, an additional low amplitude, broad kinematic component is always preferred by the statistics. The small positive residuals contribute significantly to  $\chi^2$  due to the low noise levels, but the additional broad component usually does not carry significant physical meaning. Adding empirical, physically motivated constraints to the decision metrics can help alleviate the problem (see Hampton et al. in preparation for using machine learning to determine the number of components). In the examples shown in Figures 5 and 6, we adopt the flowchart in Figure 14 using the LRT and additionally require that, for more sophisticated models to be selected, 1) the peak flux densities of the broad kinematic components ( $c2$ ,  $c3$ ) has to exceed at least 15% of the narrow component ( $c1$ ) in  $H\alpha$ , and 2) the broadest kinematic component ( $c3$ ) cannot have less  $H\alpha$  flux than the narrow component ( $c1$ ).

## References

- Allen, J.T., Croom, S.M., Konstantopoulos, I.S., Bryant, J.J., Sharp, R., Cecil, G.N., Fogarty, L.M.R., Foster, C., Green, A.W., Ho, I.-T., Owers, M.S., Schaefer, A.L., Scott, N., Bauer, A.E., Baldry, I., Barnes, L.A., Bland-Hawthorn, J., Bloom, J.V., Brough, S., Colless, M., Cortese, L., Couch, W.J., Drinkwater, M.J., Driver, S.P., Goodwin, M., Gunawardhana, M.L.P., Hampton, E.J., Hopkins, A.M., Kewley, L.J., Lawrence, J.S., Leon-Saval, S.G., Liske, J., López-Sánchez, Á.R., Lorente, N.P.F., McElroy, R., Medling, A.M., Mould, J., Norberg, P., Parker, Q.A., Power, C., Pracy, M.B., Richards, S.N., Robotham, A.S.G., Sweet, S.M., Taylor, E.N., Thomas, A.D., Tonini, C., Walcher, C.J.: *Mon. Not. R. Astron. Soc.* **446**, 1567 (2015a). 1407.6068. doi:10.1093/mnras/stu2057
- Allen, J.T., Schaefer, A.L., Scott, N., Fogarty, L.M.R., Ho, I.-T., Medling, A.M., Leslie, S.K., Bland-Hawthorn, J., Bryant, J.J., Croom, S.M., Goodwin, M., Green, A.W., Konstantopoulos, I.S., Lawrence, J.S., Owers, M.S., Richards, S.N., Sharp, R.: *Mon. Not. R. Astron. Soc.* **451**, 2780 (2015b). 1505.03872. doi:10.1093/mnras/stv1121
- Bacon, R., Copin, Y., Monnet, G., Miller, B.W., Allington-Smith, J.R., Bureau, M., Carollo, C.M., Davies, R.L., Emsellem, E., Kuntschner, H., Peletier, R.F., Verolme, E.K., de Zeeuw, P.T.: *Mon. Not. R. Astron. Soc.* **326**, 23 (2001). astro-ph/0103451. doi:10.1046/j.1365-8711.2001.04612.x
- Bacon, R., Accardo, M., Adjali, L., Anwand, H., Bauer, S., Biswas, I., Blaizot, J., Boudon, D., Brau-Nogue, S., Brinchmann, J., Caillier, P., Capolani, L., Carollo, C.M., Contini, T., Couderc, P., Daguísé, E., Deiries, S., Delabre, B., Dreizler, S., Dubois, J., Dupieux, M., Dupuy, C., Emsellem, E., Fechner, T., Fleischmann, A., François, M., Gallou, G., Gharsa, T., Glindemann, A., Gojak, D., Guiderdoni, B., Hansali, G., Hahn, T., Jarno, A., Kelz, A., Koehler, C., Kosmalski, J., Laurent, F., Le Floch, M., Lilly, S.J., Lizon, J.-L., Loupias, M., Manescau, A., Monstein, C., Nicklas, H., Olaya, J.-C., Pares, L., Pasquini, L., Pécontal-Rousset, A., Pelló, R., Petit, C., Popow, E., Reiss, R., Remillieux, A., Renault, E., Roth, M., Rupprecht, G., Serre, D., Schaye, J., Soucaïl, G., Steinmetz, M., Streicher, O., Stuijk, R., Valentin, H., Vernet, J., Weilbacher, P., Wisotzki, L., Yerle, N.: In: *Ground-based and Airborne Instrumentation for Astronomy III*. Proc. SPIE, vol. 7735, p. 773508 (2010). doi:10.1117/12.856027
- Bevington, P.R., Robinson, D.K.: *Data Reduction and Error Analysis for the Physical Sciences*, (1992)
- Bland-Hawthorn, J.: In: Ziegler, B.L., Combes, F., Dannerbauer, H., Verdugo, M. (eds.) *IAU Symposium*. IAU Symposium, vol. 309, p. 21 (2015). 1410.3838. doi:10.1017/S1743921314009247
- Bryant, J.J., Owers, M.S., Robotham, A.S.G., Croom, S.M., Driver, S.P., Drinkwater, M.J., Lorente, N.P.F., Cortese, L., Scott, N., Colless, M., Schaefer, A., Taylor, E.N., Konstantopoulos, I.S., Allen, J.T., Baldry, I., Barnes, L., Bauer, A.E., Bland-Hawthorn, J., Bloom, J.V., Brooks, A.M., Brough, S., Cecil, G., Couch, W., Croton, D., Davies, R., Ellis, S., Fogarty, L.M.R., Foster, C., Glazebrook, K., Goodwin, M., Green, A., Gunawardhana, M.L., Hampton, E., Ho, I.-T., Hopkins, A.M., Kewley, L., Lawrence, J.S., Leon-Saval, S.G., Leslie, S., McElroy, R., Lewis, G., Liske, J., López-Sánchez, Á.R., Mahajan, S., Medling, A.M., Metcalfe, N., Meyer, M., Mould, J., Obreschkow, D., O’Toole, S., Pracy, M., Richards, S.N., Shanks, T., Sharp, R., Sweet, S.M., Thomas, A.D., Tonini, C., Walcher, C.J.: *Mon. Not. R. Astron. Soc.* **447**, 2857 (2015). 1407.7335. doi:10.1093/mnras/stu2635
- Bundy, K., Bershad, M.A., Law, D.R., Yan, R., Drory, N., MacDonald, N., Wake, D.A., Cherinka, B., Sánchez-Gallego, J.R., Weijmans, A.-M., Thomas, D., Tremonti, C., Masters, K., Coccato, L., Diamond-Stanic, A.M., Aragón-Salamanca, A., Avila-Reese, V., Badenes, C., Falcón-Barroso, J., Belfiore, F., Bizyaev, D., Blanc, G.A., Bland-Hawthorn, J., Blanton, M.R., Brownstein, J.R., Byler, N., Cappellari, M., Conroy, C., Dutton, A.A., Emsellem, E., Etherington, J., Frinchaboy, P.M., Fu, H., Gunn, J.E., Harding, P., Johnston, E.J., Kauffmann, G., Kinemuchi, K., Klaene, M.A., Knapen, J.H., Leauthaud, A., Li, C., Lin, L., Maiolino, R., Malanushenko, V., Malanushenko, E., Mao, S., Maraston, C., McDermid, R.M., Merrifield, M.R., Nichol, R.C., Oravetz, D., Pan, K., Parejko, J.K., Sanchez, S.F., Schlegel, D., Simmons, A., Steele, O., Steinmetz, M., Thanjavur, K., Thompson, B.A., Tinker, J.L., van den Bosch, R.C.E., Westfall, K.B., Wilkinson, D., Wright, S., Xiao, T., Zhang, K.: *Astrophys. J.* **798**, 7 (2015). 1412.1482. doi:10.1088/0004-637X/798/1/7
- Cappellari, M., Copin, Y.: *Mon. Not. R. Astron. Soc.* **342**, 345 (2003). arXiv:astro-ph/0302262. doi:10.1046/j.1365-8711.2003.06541.x
- Cappellari, M., Emsellem, E.: *Publ. Astron. Soc. Pac.* **116**, 138 (2004). arXiv:astro-ph/0312201. doi:10.1086/381875
- 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.: *Mon. Not. R. Astron. Soc.* **413**, 813 (2011). 1012.1551. doi:10.1111/j.1365-2966.2010.18174.x
- Cid Fernandes, R., Mateus, A., Sodré, L., Stasińska, G., Gomes, J.M.: *Mon. Not. R. Astron. Soc.* **358**, 363 (2005). arXiv:astro-ph/0412481. doi:10.1111/j.1365-2966.2005.08752.x
- 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.: *Astron. Astrophys.* **561**, 130 (2014). 1307.0562. doi:10.1051/0004-6361/201321692
- Colless, M.: In: Efstathiou, G., *et al.*(eds.) *Large-Scale Structure in the Universe*, p. 105 (1999)
- 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.: *Mon. Not. R. Astron. Soc.* **421**, 872 (2012). 1112.3367. doi:10.1111/j.1365-2966.2011.20365.x
- Davies, R.L., Kewley, L.J., Ho, I.-T., Dopita, M.A.: *Mon. Not. R. Astron. Soc.* **444**, 3961 (2014). 1408.5888. doi:10.1093/mnras/stu1740
- Davis, M., Faber, S.M., Newman, J., Phillips, A.C., Ellis, R.S., Steidel, C.C., Conselice, C., Coil, A.L., Finkbeiner, D.P., Koo, D.C., Guhathakurta, P., Weiner, B., Schiavon, R., Willmer, C., Kaiser, N., Luppino, G.A., Wirth, G., Connolly, A., Eisenhardt, P., Cooper, M., Gerke, B.: In: Guhathakurta, P. (ed.) *Discoveries and Research Prospects from 6- to 10-Meter-Class*

- Telescopes II. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 4834, p. 161 (2003). [astro-ph/0209419](https://doi.org/10.1117/12.457897). doi:10.1117/12.457897
- Dimeo, R.: PAN User Guide, <ftp://ftp.ncnr.nist.gov/pub/staff/dimeo/pandoc.pdf/> (2005)
- Dopita, M.A., Shastri, P., Davies, R., Kewley, L., Hampton, E., Scharwächter, J., Sutherland, R., Kharb, P., Jose, J., Bhatt, H., Ramya, S., Jin, C., Banfield, J., Zaw, I., Juneau, S., James, B., Srivastava, S.: *Astrophys. J. Suppl. Ser.* **217**, 12 (2015a). 1501.02022. doi:10.1088/0067-0049/217/1/12
- Dopita, M.A., Ho, I.-T., Dressel, L.L., Sutherland, R., Kewley, L., Davies, R., Hampton, E., Shastri, P., Kharb, P., Jose, J., Bhatt, H., Ramya, S., Scharwächter, J., Jin, C., Banfield, J., Zaw, I., James, B., Juneau, S., Srivastava, S.: *Astrophys. J.* **801**, 42 (2015b). 1501.02507. doi:10.1088/0004-637X/801/1/42
- Dopita, M., Hart, J., McGregor, P., Oates, P., Bloxham, G., Jones, D.: *Astrophys. Space Sci.* **310**, 255 (2007). 0705.0287. doi:10.1007/s10509-007-9510-z
- Dopita, M., Rhee, J., Farage, C., McGregor, P., Bloxham, G., Green, A., Roberts, B., Neilson, J., Wilson, G., Young, P., Firth, P., Busarello, G., Merluzzi, P.: *Astrophys. Space Sci.* **327**, 245 (2010). 1002.4472. doi:10.1007/s10509-010-0335-9
- González Delgado, R.M., Cerviño, M., Martins, L.P., Leitherer, C., Hauschildt, P.H.: *Mon. Not. R. Astron. Soc.* **357**, 945 (2005). [arXiv:astro-ph/0501204](https://arxiv.org/abs/astro-ph/0501204). doi:10.1111/j.1365-2966.2005.08692.x
- Grandmont, F., Drissen, L., Mandar, J., Thibault, S., Baril, M.: In: *Ground-based and Airborne Instrumentation for Astronomy IV*. Proc. SPIE, vol. 8446, p. 84460 (2012). doi:10.1117/12.926782
- Green, A.W., Glazebrook, K., McGregor, P.J., Damjanov, I., Wisnioski, E., Abraham, R.G., Colless, M., Sharp, R.G., Crain, R.A., Poole, G.B., McCarthy, P.J.: *Mon. Not. R. Astron. Soc.* **437**, 1070 (2014). 1310.6082. doi:10.1093/mnras/stt1882
- Groves, B., Brinchmann, J., Walcher, C.J.: *Mon. Not. R. Astron. Soc.* **419**, 1402 (2012). doi:10.1111/j.1365-2966.2011.19796.x
- Ho, I.-T., Kewley, L.J., Dopita, M.A., Medling, A.M., Allen, J.T., Bland-Hawthorn, J., Bloom, J.V., Bryant, J.J., Croom, S.M., Fogarty, L.M.R., Goodwin, M., Green, A.W., Konstantopoulos, I.S., Lawrence, J.S., López-Sánchez, Á.R., Owers, M.S., Richards, S., Sharp, R.: *Mon. Not. R. Astron. Soc.* **444**, 3894 (2014). 1407.2411. doi:10.1093/mnras/stu1653
- Ho, I.-T., Kudritzki, R.-P., Kewley, L.J., Zahid, H.J., Dopita, M.A., Bresolin, F., Rupke, D.S.N.: *Mon. Not. R. Astron. Soc.* **448**, 2030 (2015). 1501.02668. doi:10.1093/mnras/stv067
- Ho, I.-T., Medling, A.M., Bland-Hawthorn, J., Groves, B., Kewley, L.J., Kobayashi, C., Dopita, M.A., Leslie, S.K., Sharp, R., Allen, J.T., Bourne, N., Bryant, J.J., Cortese, L., Croom, S.M., Dunne, L., Fogarty, L.M.R., Goodwin, M., Green, A.W., Konstantopoulos, I.S., Lawrence, J.S., Lorente, N.P.F., Owers, M.S., Richards, S., Sweet, S.M., Tesconi, E., Valiante, E.: *Mon. Not. R. Astron. Soc.* **457**, 1257 (2016). 1601.02022. doi:10.1093/mnras/stw017
- Husmann, B., Jahnke, K., Sánchez, S.F., Barrado, D., Bekeraït, S., Bomans, D.J., Castillo-Morales, A., Catalán-Torrecilla, C., Cid Fernandes, R., Falcón-Barroso, J., García-Benito, R., González Delgado, R.M., Iglesias-Páramo, J., Johnson, B.D., Kupko, D., López-Fernández, R., Lyubenova, M., Marino, R.A., Mast, D., Miskolczi, A., Monreal-Ibero, A., Gil de Paz, A., Pérez, E., Pérez, I., Rosales-Ortega, F.F., Ruiz-Lara, T., Schilling, U., van de Ven, G., Walcher, J., Alves, J., de Amorim, A.L., Backsmann, N., Barrera-Ballesteros, J.K., Bland-Hawthorn, J., Cortijo, C., Dettmar, R.-J., Demleitner, M., Díaz, A.I., Enke, H., Florido, E., Flores, H., Galbany, L., Gallazzi, A., García-Lorenzo, B., Gomes, J.M., Gruel, N., Haines, T., Holmes, L., Jungwiert, B., Kalinova, V., Kehrig, C., Kennicutt, R.C., Klar, J., Lehnert, M.D., López-Sánchez, Á.R., de Lorenzo-Cáceres, A., Mármod-Queraltó, E., Márquez, I., Méndez-Abreu, J., Mollá, M., del Olmo, A., Meidt, S.E., Papaderos, P., Puschig, J., Quirrenbach, A., Roth, M.M., Sánchez-Blázquez, P., Spekkens, K., Singh, R., Stanishev, V., Trager, S.C., Vilchez, J.M., Wild, V., Wisotzki, L., Zibetti, S., Ziegler, B.: *Astron. Astrophys.* **549**, 87 (2013). 1210.8150. doi:10.1051/0004-6361/201220582
- Koleva, M., Prugniel, P., Ocvirk, P., Le Borgne, D., Soubiran, C.: *Mon. Not. R. Astron. Soc.* **385**, 1998 (2008). 0801.0871. doi:10.1111/j.1365-2966.2008.12908.x
- Koleva, M., Prugniel, P., Bouchard, A., Wu, Y.: *Astron. Astrophys.* **501**, 1269 (2009). 0903.2979. doi:10.1051/0004-6361/200811467
- Kreckel, K., Blanc, G.A., Schinnerer, E., Groves, B., Adamo, A., Hughes, A., Meidt, S.: *ArXiv e-prints* (2016). 1603.08009
- Lawrence, J., Bland-Hawthorn, J., Bryant, J., Brzeski, J., Colless, M., Croom, S., Gers, L., Gilbert, J., Gillingham, P., Goodwin, M., Heijmans, J., Horton, A., Ireland, M., Miziarski, S., Saunders, W., Smith, G.: In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 8446, September 2012. doi:10.1117/12.925260
- Lawson, C.L., Hanson, R.J.: *Solving Least Squares Problems*, (1974)
- MacArthur, L.A., González, J.J., Courteau, S.: *Mon. Not. R. Astron. Soc.* **395**, 28 (2009). 0901.4135. doi:10.1111/j.1365-2966.2009.14519.x
- Markwardt, C.B.: In: *Bohlender, D.A., Durand, D., Dowler, P. (eds.) Astronomical Data Analysis Software and Systems XVIII*. Astronomical Society of the Pacific Conference Series, vol. 411, p. 251 (2009). 0902.2850
- Martin, C., Moore, A., Morrissey, P., Matuszewski, M., Rahman, S., Adkins, S., Epps, H.: In: *Ground-based and Airborne Instrumentation for Astronomy III*. Proc. SPIE, vol. 7735, p. 77350 (2010). doi:10.1117/12.858227
- McElroy, R., Croom, S.M., Pracy, M., Sharp, R., Ho, I.-T., Medling, A.M.: *Mon. Not. R. Astron. Soc.* **446**, 2186 (2015). 1410.6552. doi:10.1093/mnras/stu2224
- Medling, A.M., U, V., Rich, J.A., Kewley, L.J., Armus, L., Dopita, M.A., Max, C.E., Sanders, D., Sutherland, R.: *Mon. Not. R. Astron. Soc.* **448**, 2301 (2015). 1501.07289. doi:10.1093/mnras/stv081
- Morrissey, P., Matuszewski, M., Martin, C., Moore, A., Adkins, S., Epps, H., Bartos, R., Cabak, J., Cowley, D., Davis, J., Delacroix, A., Fucik, J., Hilliard, D., James, E., Kaye, S., Lingner, N., Neill, J.D., Pistor, C., Phillips, D., Rockosi, C., Weber, B.: In: *Ground-based and Airborne Instrumentation for Astronomy IV*. Proc. SPIE, vol. 8446, p. 844613 (2012). doi:10.1117/12.924729
- Ocvirk, P., Pichon, C., Lançon, A., Thiébaud, E.: *Mon. Not. R. Astron. Soc.* **365**, 46 (2006). [arXiv:astro-ph/0505209](https://arxiv.org/abs/astro-ph/0505209). doi:10.1111/j.1365-2966.2005.09182.x
- Protassov, R., van Dyk, D.A., Connors, A., Kashyap, V.L., Siemiginowska, A.: *Astrophys. J.* **571**, 545 (2002). [astro-ph/0201547](https://doi.org/10.1086/339856). doi:10.1086/339856



- Rich, J.A., Dopita, M.A., Kewley, L.J., Rupke, D.S.N.: *Astrophys. J.* **721**, 505 (2010). 1007.3495. doi:10.1088/0004-637X/721/1/505
- Rich, J.A., Torrey, P., Kewley, L.J., Dopita, M.A., Rupke, D.S.N.: *Astrophys. J.* **753**, 5 (2012). 1204.5520. doi:10.1088/0004-637X/753/1/5
- Richards, S.N., Schaefer, A.L., López-Sánchez, Á.R., Croom, S.M., Bryant, J.J., Sweet, S.M., Konstantopoulos, I.S., Allen, J.T., Bland-Hawthorn, J., Bloom, J.V., Brough, S., Fogarty, L.M.R., Goodwin, M., Green, A.W., Ho, I.-T., Kewley, L.J., Koribalski, B.S., Lawrence, J.S., Owers, M.S., Sadler, E.M., Sharp, R.: *Mon. Not. R. Astron. Soc.* **445**, 1104 (2014). 1409.4495. doi:10.1093/mnras/stu1820
- Rupke, D.S.N., Veilleux, S.: *Astrophys. J. Lett.* **729**, 27 (2011). 1102.4349. doi:10.1088/2041-8205/729/2/L27
- Rupke, D.S.N., Kewley, L.J., Chien, L.-H.: *Astrophys. J.* **723**, 1255 (2010). 1009.0761. doi:10.1088/0004-637X/723/2/1255
- Sánchez, S.F., García-Lorenzo, B., Jahnke, K., Mediavilla, E., González-Serrano, J.I., Christensen, L., Wisotzki, L.: *New Astron. Rev.* **49**, 501 (2006). doi:10.1016/j.newar.2005.10.004
- Sánchez, S.F., Cardiel, N., Verheijen, M.A.W., Pedraz, S., Covone, G.: *Mon. Not. R. Astron. Soc.* **376**, 125 (2007). *astro-ph/0611660*. doi:10.1111/j.1365-2966.2007.11335.x
- Sánchez, S.F., Kennicutt, R.C., Gil de Paz, A., van de Ven, G., Vílchez, 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., Mármol-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.: *Astron. Astrophys.* **538**, 8 (2012). 1111.0962. doi:10.1051/0004-6361/201117353
- 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., Ascasi-bar, Y., Barrera-Ballesteros, J.: *Rev. Mexicana Astron. Astrofis.* **52**, 21 (2016a). 1509.08552
- 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., Ascasi-bar, 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.: *Rev. Mexicana Astron. Astrofis.* **52**, 171 (2016b). 1602.01830
- Sanders, J.S.: *Mon. Not. R. Astron. Soc.* **371**, 829 (2006). *astro-ph/0606528*. doi:10.1111/j.1365-2966.2006.10716.x
- Sarzi, M., Falcón-Barroso, J., Davies, R.L., Bacon, R., Bureau, M., Cappellari, M., de Zeeuw, P.T., Emsellem, E., Fathi, K., Krajnović, D., Kuntschner, H., McDermid, R.M., Peletier, R.F.: *Mon. Not. R. Astron. Soc.* **366**, 1151 (2006). arXiv:astro-ph/0511307. doi:10.1111/j.1365-2966.2005.09839.x
- Tojeiro, R., Heavens, A.F., Jimenez, R., Panter, B.: *Mon. Not. R. Astron. Soc.* **381**, 1252 (2007). 0704.0941. doi:10.1111/j.1365-2966.2007.12323.x
- Vazdekis, A., Ricciardelli, E., Cenarro, A.J., Rivero-González, J.G., Díaz-García, L.A., Falcón-Barroso, J.: *Mon. Not. R. Astron. Soc.* **424**, 157 (2012). 1205.5496. doi:10.1111/j.1365-2966.2012.21179.x
- Vogt, F.P.A., Dopita, M.A., Borthakur, S., Verdes-Montenegro, L., Heckman, T.M., Yun, M.S., Chambers, K.C.: *Mon. Not. R. Astron. Soc.* **450**, 2593 (2015). 1504.03337. doi:10.1093/mnras/stv749
- Walcher, J., Groves, B., Budavári, T., Dale, D.: *Astrophys. Space Sci.* **331**, 1 (2011). 1008.0395. doi:10.1007/s10509-010-0458-z
- Yoachim, P., Roškar, R., Debattista, V.P.: *Astrophys. J.* **752**, 97 (2012). 1204.0026. doi:10.1088/0004-637X/752/2/97
- York, D.G., Adelman, J., Anderson, J.E. Jr., 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., Dombek, 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., Prosapio, 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: *Astron. J.* **120**, 1579 (2000). *astro-ph/0006396*. doi:10.1086/301513
- Zahid, H.J., Bresolin, F.: *Astron. J.* **141**, 192 (2011). doi:10.1088/0004-6256/141/6/192