

# Photoionization models of the CALIFA HII regions

## I. Hybrid models

C. Morisset<sup>1, 10</sup>, G. Delgado-Inglada<sup>1</sup>, S. F. Sánchez<sup>1</sup>, L. Galbany<sup>2, 3</sup>, R. García-Benito<sup>4</sup>, B. Husemann<sup>5</sup>, R. A. Marino<sup>6, 7</sup>, D. Mast<sup>4</sup>, M. M. Roth<sup>8</sup>, and CALIFA collaboration<sup>9</sup>

<sup>1</sup> Instituto de Astronomía, Universidad Nacional Autónoma de México, Apdo. Postal 70264, 04510, México D.F., México

<sup>2</sup> Millennium Institute of Astrophysics, Chile

<sup>3</sup> Departamento de Astronomía, Universidad de Chile, Camino El Observatorio 1515, Las Condes, Santiago, Chile

<sup>4</sup> Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía s/n, Aptdo. 3004, E18080-Granada, Spain

<sup>5</sup> European Southern Observatory (ESO), Karl-Schwarzschild-Str.2, D-85748 Garching b. München, Germany

<sup>6</sup> Department of Physics, Institute for Astronomy, ETH Zürich, CH-8093 Zürich, Switzerland

<sup>7</sup> Departamento Astrofísica, Universidad Complutense de Madrid, Avda. Ciudad Universitaria s/n 28040 - Madrid, Spain

<sup>8</sup> Leibniz Institut für Astrophysik, An der Sternwarte 16, 14482 - Potsdam, Germany

<sup>9</sup> <http://califa.caha.es/>

<sup>10</sup> e-mail: [chris.morisset@gmail.com](mailto:chris.morisset@gmail.com)

ORCID 0000-0001-5801-6724

Received 04 2016; accepted 06 2016

### ABSTRACT

Photoionization models of H II regions require as input a description of the ionizing SED (Spectral Energy Distribution) and of the gas distribution, in terms of ionization parameter  $U$  and chemical abundances (e.g. O/H and N/O). A strong degeneracy exists between the hardness of the SED and  $U$ , which in turn leads to high uncertainties in the determination of the other parameters, including abundances. One way to resolve the degeneracy is to fix one of the parameters using additional information.

For each of the  $\sim 20,000$  sources of the CALIFA H II regions catalog, a grid of photoionization models is computed assuming the ionizing SED being described by the underlying stellar population obtained from spectral synthesis modeling. The ionizing SED is then defined as the sum of various stellar bursts of different ages and metallicities. This solves the degeneracy between the shape of the ionizing SED and  $U$ . The nebular metallicity (associated to O/H) is defined using the classical strong line method O3N2 (which gives to our models the status of "hybrids"). The remaining free parameters are the abundance ratio N/O and the ionization parameter  $U$ , which are determined by looking for the model fitting [N II]/H $\alpha$  and [O III]/H $\beta$ . The models are also selected to fit [O II]/H $\beta$ . This process leads to a set of  $\sim 3,200$  models that reproduce simultaneously the three observations.

We find that the regions associated to young stellar bursts (i.e., ionized by OB stars) suffer leaking of the ionizing photons, the proportion of escaping photons having a median of 80%. The set of photoionization models satisfactorily reproduces the electron temperature derived from the [O III] $\lambda 4363/5007$  line ratio. We determine new relations between the nebular parameters, like the ionization parameter  $U$  and the [O II]/[O III] or [S II]/[S III] line ratios. A new relation between N/O and O/H is obtained, mostly compatible with previous empirical determinations (and not with previous results obtained using photoionization models). A new relation between  $U$  and O/H is also determined.

All the models are publicly available on the Mexican Million Models database 3MdB.

**Key words.** Interstellar medium (ISM), ISM: abundances, HII regions, Galaxies: ISM

## 1. Introduction

Classical H II regions are large, low-density clouds of partially ionized gas in which star formation has recently taken place ( $<15$  Myr). The short-lived blue stars forged in these regions emit large amounts of ultraviolet radiation that ionizes the surrounding gas. They span a wide range of physical scales, from a few parsecs, like the Orion nebula ( $D \sim 8$  pc), or even smaller (Anderson 2014), to hundreds of parsecs, such as 30 Doradus ( $D \sim 20$  pc), NGC 604 ( $D \sim 460$  pc), or NGC 5471 ( $D \sim 1$  kpc) as reported by Oey et al. (2003) and García-Benito et al. (2011). These latter ones are the prototypes of the extragalactic giant H II regions found frequently in the disks of spiral galaxies (e.g. Hodge & Kennicutt 1983; Dottori 1987; Dottori & Copetti 1989; Knapen 1998), or starburst and blue compact galaxies (e.g.

Kehrig et al. 2008; López-Sánchez & Esteban 2009; Cairós et al. 2012).

Baldwin et al. (1981) first proposed the [O III] $\lambda 5007$ /H $\beta$  versus [N II] $\lambda 6584$ /H $\alpha$  diagnostic diagram (now known as the BPT diagram) to separate emission-line objects according to the main gas excitation mechanism: normal H II regions, planetary nebulae, and objects photoionized by a harder radiation field. The latter can be produced by either a power-law continuum from an AGN (Active Galactic Nuclei), shock excitation, planetary nebulae central stars or even post-AGB (Asymptotic Giant Branch) stars (e.g. Binette et al. 1994; Stasińska et al. 2008; Binette et al. 2009; Morisset & Georgiev 2009; Flores-Fajardo et al. 2009; Kehrig et al. 2012; Singh et al. 2013; Papaderos et al. 2013). Veilleux & Osterbrock (1987) and Osterbrock (1989) extended and refined this classification scheme, incorporating new

diagnostic diagrams. Osterbrock (1989) used theoretical photoionization models to infer the demarcation line between star-forming (SF) and AGN galaxies, and added two new diagnostics diagrams that exploit the  $[\text{O I}]/\text{H}\beta$  versus  $[\text{S II}]/\text{H}\alpha$  line ratios. Dopita et al. (2000) and Kewley et al. (2001) combined stellar population synthesis and photoionization models to build the first purely theoretical classification scheme for separating pure AGN from galaxies hosting star formation, and Kauffmann et al. (2003) used SDSS (Sloan Digital Sky Survey York et al. 2000) data to observationally constrain these classifications.

In essence, these models assume that the main factors that control the emission line spectrum are the chemical abundances of the heavy elements in the gas phase within an  $\text{H II}$  region (oxygen being the most important), the shape of the ionizing radiation spectrum, and the geometrical distribution of gas with respect to the ionizing sources. Generally speaking, all the geometrical factors are subsumed into a single factor, the ionization parameter  $q$  (with dimensions  $\text{cm s}^{-1}$ ) or the (dimensionless) ionization parameter  $U = q/c$ . They also assume *a priori* that these parameters are independent, and thus these models are presented as grids of oxygen abundance, ionization parameter, shape of the ionizing spectrum (effective temperature or stellar burst age) and sometimes electron densities.

Most of the present day knowledge of these regions is based on the comparison of the predictions between these photoionization models and the largest accessible databases for the observed properties of  $\text{H II}$  regions. However, in many cases, the samples/catalogs are limited in number (a few hundred of  $\text{H II}$  regions), and/or biased ( $\text{H II}$  hosted by Sc/Sd galaxies, due to the better contrast with the continuum). This has been recently overcome by the advent of large IFU surveys that have provided large catalogs of  $\text{H II}$  regions/aggregations with spectroscopic information (of the order of thousands), over unbiased sample of galaxies (from E to Sds, see e.g. Marino et al. 2016; Sánchez-Menguiano et al. 2016).

This is the case of the CALIFA (Calar Alto Legacy Integral Field spectroscopy Area survey) survey (Sánchez et al. 2012b), that has acquired IFU (Integral Field Unit) data of a sample of  $\sim 600$  galaxies in the Local Universe ( $0.005 < z < 0.03$ ), covering the full optical extension of these galaxies (see Walcher et al. 2014, for more information on the sample). This survey has created one of the largest catalogs of  $\text{H II}$  regions/aggregations, with more than 20,000 ionized regions, with spectroscopic information covering most of the typical emission lines in the optical wavelength range from  $[\text{O II}]\lambda 3727$  to  $[\text{S II}]\lambda 6731$ , and with an accurate spectral modeling and subtraction of the underlying stellar population.

One of the main problems in the determination of the properties of  $\text{H II}$  regions is that some of the parameters that describe these properties act on very similar ways on the observations. This is the case for the softness of the ionizing radiation and the ionization parameter  $U$ . Both change the ionization state of the nebula, in particular the line ratios involving two subsequent ions (e.g.  $[\text{O II}]/[\text{O III}]$ ). The best way to resolve this degeneracy is to find a method to determine one of the two parameters using an alternative observable. Combining the output from spectral synthesis modeling of the CALIFA observations gives us access to the softness of the ionization field, kipping only  $U$  to be determined.

In this article we use this extensive catalog to create an 'ad-hoc' grid of photoionization models, with the properties of the ionizing sources *a priori* provided by the analysis of the stellar populations, in order to understand the physical conditions of these nebulae. We also use an *a priori* determination of the

O/H metallicity indicator (using a strong line method) to only have the ionization parameter  $U$  and the N/O abundance ratio as free parameter. We are actually doing a work similar to Pérez-Montero et al. (2010) but using a much larger set of observations.

The paper is organized as follows: the Sec. 2 describes the CALIFA data set used in this work. The grids of models (meta-grid and ad-hoc models for each region) are described in Sec. 3. The results are presented and discussed in Sec. 4, while the conclusions are drawn in Sec. 5.

## 2. The CALIFA data

The galaxies were selected from the CALIFA observed sample. Since CALIFA is an ongoing survey whose observations are scheduled on a monthly basis (i.e., dark nights), the list of objects increases regularly. The current results are based on the 612 galaxies observed up to June 2014, comprising both galaxies from the CALIFA mother sample and the so-called extended sample (details in Sánchez et al. in prep.). Their main characteristics have already been described in Sánchez et al. (2015).

The details of the survey, sample, observational strategy, and reduction are explained in Sánchez et al. (2012a). All galaxies were observed using PMAS (Roth et al. 2005) in the PPAK configuration (Kelz et al. 2006), covering a hexagonal field of view (FoV) of  $74'' \times 64''$ , which is sufficient to map the full optical extent of the galaxies up to two to three disk effective radii. This is possible because of the diameter selection of the sample (Walcher et al. 2014). The observing strategy guarantees complete coverage of the FoV, with a final spatial resolution of  $\text{FWHM} \sim 2.5''$  (García-Benito et al. 2015), corresponding to  $\sim 1$  kpc at the average redshift of the survey. The sampled wavelength range and spectroscopic resolution ( $3745\text{\AA} - 7500\text{\AA}$ ,  $\lambda/\Delta\lambda \sim 850$ , for the low-resolution setup) are more than sufficient to explore the most prominent ionized gas emission lines from  $[\text{O II}]\lambda 3727$  to  $[\text{S II}]\lambda 6731$  at the redshift of our targets, on one hand, and to deblend and subtract the underlying stellar population, on the other (e.g., Sánchez et al. 2012a; Kehrig et al. 2012; Cid Fernandes et al. 2013, 2014). The dataset was reduced using version 1.5 of the CALIFA pipeline, whose modifications with respect to the ones presented in Sánchez et al. (2012a) and Husemann et al. (2013) are described in detail in García-Benito et al. (2015). In summary, the data fulfill the predicted quality-control requirements with a spectrophotometric accuracy that is better than 5% everywhere within the wavelength range, both absolute and relative with a depth that allows us to detect emission lines in individual  $\text{H II}$  regions as faint as  $\sim 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ , and with a signal-to-noise ratio of  $\text{S/N} \sim 3-5$ . For the emission lines considered in the current study, the S/N is well above this limit, and the measurement errors are negligible in most of the cases. In any case, they have been propagated and included in the final error budget.

The final product of the data reduction is a regular-grid datacube, with  $x$  and  $y$  coordinates that indicate the right ascension and declination of the target, and  $z$  is a common step in wavelength. The CALIFA pipeline also provides the propagated error cube, a proper mask cube of bad pixels, and a prescription of how to handle the errors when performing spatial binning (due to covariance between adjacent pixels after image reconstruction). These datacubes, together with the ancillary data described in Walcher et al. (2014), are the basic starting points of our analysis.

### 2.1. H II regions: detection and extraction

The segregation of H II regions and the extraction of the corresponding spectra is performed using a semi-automatic procedure named HIIEXPLORER<sup>1</sup>. The procedure is based on some basic assumptions: (a) H II regions are peaky and isolated structures with a strong ionized gas emission, which is significantly above the stellar continuum emission and the average ionized gas emission across the galaxy. This is particularly true for H $\alpha$  because (b) H II regions have a typical physical size of about a hundred or a few hundred parsecs (e.g., González Delgado & Perez 1997; Lopez et al. 2011; Oey et al. 2003), which corresponds to a typical projected size of a few arcsec at the distance of the galaxies.

These assumptions are based on the fact that most of the H $\alpha$  luminosity observed in spiral and irregular galaxies is a direct tracer of the ionization of the interstellar medium (ISM) by the ultraviolet (UV) radiation produced by young high-mass OB stars. Since only high-mass, short-lived stars contribute significantly to the integrated ionizing flux, this luminosity is a direct tracer of the current star-formation rate (SFR), independent of the previous star-formation history. Therefore, clumpy structures detected in the H $\alpha$  intensity maps are most probably associated with classical H II regions (i.e., those regions for which the oxygen abundances have been calibrated).

The details of HIIEXPLORER are given in Sánchez et al. (2012b) and Rosales-Ortega et al. (2012). In summary we create a narrow-band image centered on the wavelength of H $\alpha$  at the redshift of the object. Then we run HIIEXPLORER to detect and extract the spectra of each individual H II region, adopting the parameters presented in Sánchez et al. (2014a). The algorithm starts looking for the brightest pixel in the map. Then, the code aggregates the adjacent pixels until all pixels with flux greater than 10% of the peak flux of the region and within 500 pc or 3.5 spaxels from the center have been accumulated. The distance limit takes the typical size of H II regions of a few hundreds of parsecs into account (e.g., González Delgado & Perez 1997; Lopez et al. 2011). Then, the selected region is masked and the code keeps iterating until no peak with flux exceeding the median H $\alpha$  emission flux of the galaxy is left. Mast et al. (2014) studied the loss of resolution in IFS using nearby galaxies observed by PINGS (PPAK ISF Nearby Galaxies Survey, see Rosales-Ortega et al. 2010). Some of these galaxies were simulated at higher redshifts to match the characteristics and resolution of the galaxies observed by the CALIFA survey. Regarding the H II region selection, the authors conclude that at  $z \sim 0.02$ , the H II clumps can contain on average from 1 to 6 of the H II regions obtained from the original data at  $z \sim 0.001$ . Another caveat is that this procedure tends to select regions with similar sizes, although real H II regions actually have different sizes. However, the actual adopted size is close to the FWHM (2.5'') of the CALIFA data for the version of the data reduction we used (García-Benito et al. 2015).

Then, for each individual extracted spectrum we modeled the stellar continuum using FIT3D<sup>2</sup>, a fitting package described in Sánchez et al. (2006b), Sánchez et al. (2011) and Sánchez et al. (2016). This fitting tool performs multiple linear regressions to derive the optimal combination of a single-stellar population (SSP) library over a set of Monte-Carlo realizations of the input spectrum, providing with best fitted set of weights for each population and the corresponding errors. Prior to this analysis the procedure derive the best kinematics and dust attenuation for each fitted spectrum. In this particular study we use the gsd156 tem-

plate library, described in detail by Cid Fernandes et al. (2013). It comprises 156 templates that cover 39 stellar ages (1 Myr to 13 Gyr), and 4 metallicities ( $Z/Z_{\odot} = 0.2, 0.4, 1, \text{ and } 1.5$ ). These templates were extracted from a combination of the synthetic stellar spectra from the GRANADA library (Martins et al. 2005) and the SSP library provided by the MILES project (Sánchez-Blázquez et al. 2006; Vazdekis et al. 2010; Falcón-Barroso et al. 2011). This library has been extensively used within the CALIFA collaboration in different studies (e.g. Pérez et al. 2013; Cid Fernandes et al. 2013; González Delgado et al. 2014). The only difference with respect to these studies is that the spectral resolution of the library was not fixed to the spectral resolution of the CALIFA V500 setup data (FWHM  $\sim 6 \text{ \AA}$ ), to allow its use for datasets with different resolution. As shown in Sánchez et al. (2016) the results are not strongly affected by the selection of a different stellar template. The reliability of the derived parameters for the stellar population using FIT3D was extensively tested against simulations and perturbed data. In particular it was found that it is required a S/N above 50 to break the well-known degeneracies and provide reliable weights for the stellar population when they contribute at least to  $\sim 5\%$  to the total flux in the visible range (e.g. Sánchez et al. 2016, Fig. 9 and 15). Indeed in previous articles (Sánchez et al. 2014b) we already explored the correspondence between the estimated fraction of young stars ( $f_y$ ) and the equivalent width of H $\alpha$  ( $EW_{\alpha}$ ), parameters that show a clear correlation when the  $f_y > 20\%$  and  $EW_{\alpha} > 6 \text{ \AA}$ . This is indeed a good test that supports the reliability of the derived fraction of young stars.

After subtracting the underlying stellar population, the flux intensity of the strong emission lines was extracted for each gas-pure spectrum by fitting a single Gaussian model to each line, resulting in a catalog of the emission-line properties (Sánchez et al. 2012b).

The final catalog comprises the strongest emission line and emission line ratios from [O II] $\lambda 3727$  to [S II] $\lambda 6731$  for 18178 H II regions from 612 galaxies, together with their equivalent widths and the luminosity-weighted ages and metallicities of the underlying stellar population. So far, this is the largest catalog of H II regions and aggregations with spectroscopic information. It is also one of the few catalogs derived for a statistically well-defined sample of galaxies representative of the entire population of galaxies in the local Universe (Walcher et al. 2014).

In this work, we will only use the HI, [N II], [O II], and [O III]<sup>3</sup> emission lines, which are present in almost all the sources. Auroral lines are seen only in some tens of them (see the 16 [O III] $\lambda 4363$  lines used by Marino et al. 2013).

## 3. The grid of models

### 3.1. The meta-grid

For each H II region of each galaxy, we run a grid of 220 photoionization models using the CLOUDY code (Ferland et al. 2013, c13.03) driven by the pyCloudy package<sup>4</sup> (Morisset 2013, 2014). The grids are obtained by varying the mean ionization parameter  $\log(\bar{U})$ <sup>5</sup> which takes 11 values from -4 to -1.5, the abundance ratio  $\log(N/O)$  (5 values from -1.5 to 1.5 around the solar

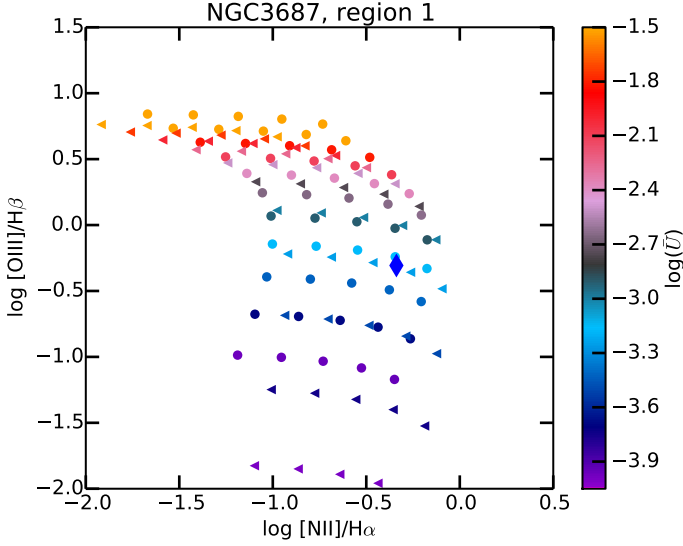
<sup>3</sup> In the following, we will use [N II], [O I], [O II], [O III], [S II], and [S III] for the [N II] $\lambda 6584 \text{ \AA}$ , [O I] $\lambda 6300 \text{ \AA}$ , [O II] $\lambda 3726+29 \text{ \AA}$ , [O III] $\lambda 5007 \text{ \AA}$ , [S II] $\lambda 6716+31 \text{ \AA}$ , and [S III] $\lambda 6312 \text{ \AA}$  lines respectively.

<sup>4</sup> <https://sites.google.com/site/pycloudy/>

<sup>5</sup> The ionization parameter is defined as  $U(r) = Q(H^0)/4\pi r^2 N_H c$ , where  $Q(H^0)$  is the number of ionizing photons emitted by the source per unit of time,  $r$  is the distance between the source and the nebula,

<sup>1</sup> [http://www.caha.es/sanchez/HII\\_explorer/](http://www.caha.es/sanchez/HII_explorer/)

<sup>2</sup> <http://www.astroscu.unam.mx/~sfsanchez/FIT3D/>



**Fig. 1.** Illustration of the modeled BPT diagram  $[\text{O III}]/\text{H}\beta$  vs.  $[\text{N II}]/\text{H}\alpha$  used to interpolate the values of  $\log(\bar{U})$  and N/O for the region 1 of NGC3687. The blue diamond corresponds to the observed values in this  $[\text{O III}]/\text{H}\beta$  vs.  $[\text{N II}]/\text{H}\alpha$  diagram. The circles and triangles correspond to the values of the models obtained with the morphological factor  $fr$  set to 0.03 (filled sphere) and to 3.0 (thin shell) respectively. The colors correspond to the values of  $\log(\bar{U})$  while the different values of N/O lead to models from left to right for increasing N/O.

value), the morphology parameter  $fr$  (0.03 and 3.0, see below) and the nebular metallicity ("Neb" and "Stel", see below). Following Stasińska et al. (2015), the desired value of  $\log(\bar{U})$  is obtained by setting the  $\text{H}^0$ -ionizing photons emission rate  $Q(\text{H}^0)$  to:

$$Q(\text{H}^0) = \frac{4 \times \pi \times c^3 \times \bar{U}^3}{3 \times N_H \times ff^2 \times \alpha_B^2 \times w^3}$$

where  $c$  is the speed of light,  $N_H$  is the hydrogen density (set to  $10 \text{ H}/\text{cm}^3$  for all the models),  $ff$  is the filling factor,  $\alpha_B$  is the effective case B recombination coefficient, and  $w = (1 + fr^3)^{1/3} - fr$ , with the morphology factor  $fr = R_{in}/R_{Str}$  being the ratio between  $R_{in}$ , the inner radius of the nebula, and  $R_{Str}$ , the Strömgren radius of the nebula if it were a full filled sphere. A morphology factor  $fr \gg 1$  ( $w \sim 0$ ) corresponds to a thin shell (e. g. a plan parallel model), while  $fr \sim 0$  ( $w \sim 1$ ) corresponds to a filled sphere. The Strömgren radius of a filled sphere is:

$$R_{Str} = \left[ \frac{3 \times Q(\text{H}^0)}{4 \times \pi \times N_H^2 \times \alpha_B \times ff} \right]^{1/3}$$

The ionizing SED is obtained by summing up individual models from POPSTAR code (Mollá et al. 2009). Each model corresponds to an individual burst of age and metallicity from the gsd156 template and has a weight given by the multi-SSP analysis described before. We use POPSTAR models obtained following the IMF from Salpeter (1955). We checked that using an IMF from Chabrier (2003) does not significantly change our results.

The oxygen abundance of the ionized gas is determined in two ways:

$N_H$  is the hydrogen density and  $c$  is the speed of light. We use the mean value of  $U$  on the volume of the nebula weighted by the electron density, and name  $\log(\bar{U})$  its logarithmic value.

- "Neb": from the nebular  $[\text{O III}]/[\text{N II}]$  line ratios applying the  $\text{O}(\text{O3N2})$  relation determined by Marino et al. (2013, hereafter M13), namely  $12 + \log(\text{O}/\text{H}) = 8.533 - 0.214 \times \text{O3N2}$ , where  $\text{O3N2}$  is  $\log([\text{O III}]/\text{H}\beta)/([\text{N II}]/\text{H}\alpha)$ .
- "Stel": The luminosity-weighted log metallicity of the underlying young stellar population, derived by co-adding the metallicities of the corresponding SSPs within the library multiplied by its contributed fraction of light in the V-band, but only for those SSPs with ages younger than 2 Gyrs, following González Delgado et al. (2014). This set of models will not be used in the following main analysis, as we know that this determination of the nebular metallicity is less reliable, due to the low S/N of the underlying continuum for a fraction of the  $\text{H II}$  regions (Sánchez et al. 2016). Although the continuum has a good S/N ( $\sim 30$ -50) to perform a SSP decomposition, to derive the metallicity of the young stars it is needed to have a similar S/N for only the young component, that may contribute between a 100% and a 20% of the total flux. This can not be granted in general. Therefore, the estimated metallicities would have large errors. The results obtain with these models are discussed in 4.8 and are showed in the Online only Appendix A. This method have been calibrated to be used for  $\text{H II}$  regions and may not apply for regions ionized by old stars.

Notice that given the range of masses of the galaxies considered here ( $10^{9.5}$  to  $10^{12}$  solar masses), we expect the O/H abundance to range from  $10^{-4}$  to  $10^{-3}$ , based on the Mass-Metallicity relation (e.g Sánchez et al. 2013).

The element abundances are following O/H (except N/H which is a free parameter). The abundances relative to O are taken from Asplund et al. (2009).

Dust is included in the model, in the form of the "ism" type defined by Cloudy, with a dust to gas ratio following a broken power law, as in the case  $X_{\text{CO,Z}}$  defined and recommended by Rémy-Ruyer et al. (2014). Following Draine (2011), we apply an additional factor of 2/3 to the final dust to gas ratio used.

The models of this meta-grid (grid of grids) are run in a quick mode (no iteration, no level-2 lines; see Cloudy manual). The ad-hoc models (see next section) do not have these limitations.

Not all the CALIFA regions have been used, we apply a filter to select only the ones where the values of the CALIFA field ratio  $\text{med\_flux}/\text{StdDev}$  are over 10 for the lines of interest (this correspond to the average S/N from blue to red through the full spectral range) and the value of the CALIFA field  $\text{MIN\_CHISQ}$  is over 0.7 (the reduced  $\chi^2$  parameter). Both cuts ensure that the fitting provides reliable results. A total of 397 galaxies have been used, summing up 9181 regions corresponding to 2,019,820 individual photoionization models.

The models have been stored in the unpublished working database associated with 3MdB (Mexican Million Models database, see Morisset et al. 2015).

### 3.2. The ad-hoc models

We use the meta-grid to find the model that reproduces the best the observed line ratios  $[\text{O III}]/\text{H}\beta$  and  $[\text{N II}]/\text{H}\alpha$  for each region. Figure 1 illustrates as an example the results for region 1 of NGC3687. The figure shows the classical BPT diagram,  $[\text{O III}]/\text{H}\beta$  vs.  $[\text{N II}]/\text{H}\alpha$  (Baldwin et al. 1981). The blue diamond corresponds to the observed line ratios whereas the colored circles (triangles) are the results of the grid of models obtained with  $fr$  being 0.03 (3.0), using the SED corresponding to this region

and the nebular metallicity derived from the O3N2 diagnostic (see Sec. 3.1).

We have performed a 2D-interpolation in the BPT diagram to determine the  $\log(\bar{U})$  and the N/O ratio for each of the models. The metallicity method ("Stel" or "Neb", see sec. 3.1) and the geometry ( $f_r$ ) take both two values and thus, we obtained 4 ad-hoc models for each region. As no extrapolation is done in the BPT diagrams, for some regions less than 4 ad-hoc models are performed. The final number of ad-hoc models is 20,793 (from 272 galaxies), from which 10,196 are "Neb" models (obtained using the O3N2 method to determine O/H).

All the 20,793 ad-hoc models are stored in the 3MdB database (Morisset et al. 2015) and accessible for any user under the "CALIFA\_ah" reference (the "ref" field). The "com5" field is used to store the value of the fit for the  $[\text{O II}]/\text{H}\beta$  line ratio, see Sec. 3.3. The "com8" field is used to store the result of the BPT-Population filter, see Sec. 3.6. Once the final release of the CALIFA data will be publicly available, we plan to rerun all the procedure to obtain more data points and to store the corresponding new models in 3MdB under the "CALIFA\_ah2" reference.

### 3.3. The $[\text{O II}]/\text{H}\beta$ filter

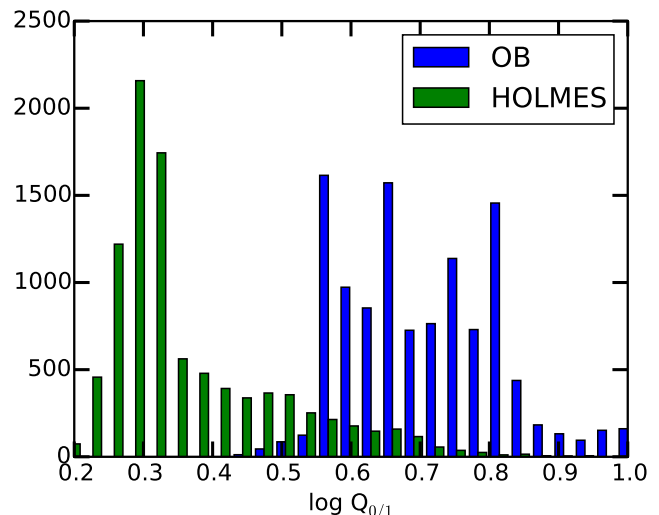
The ad-hoc models have been designed to fit the  $[\text{N II}]/\text{H}\alpha$  and  $[\text{O III}]/\text{H}\beta$  line ratios, but we can also impose that they fit the observed  $[\text{O II}]/\text{H}\beta$  line ratio. We filter the models for which the value of  $\log([\text{O II}]/\text{H}\beta)$  equals the observed value within 0.1 dex. That reduces the number of models by a factor of  $\sim 6$  but provides us with a more realistic reference data set. To apply this filter we have corrected the  $[\text{O II}]$  line intensities from the reddening, using  $\text{H}\alpha/\text{H}\beta = 2.85$  and the (Fitzpatrick 1999) extinction law. The number of ad-hoc models that also fit this  $[\text{O II}]/\text{H}\beta$  filter is 3195, from which 1574 are "Neb" models. In the 3MdB database, one can select the models that fit the  $[\text{O II}]/\text{H}\beta$  line ratio, using the "com5" entry, that contains  $\log([\text{O II}]/\text{H}\beta)_{\text{obs}} - \log([\text{O II}]/\text{H}\beta)_{\text{mod}}$ .

This filter certainly adds some bias in our sample, as only the 1/6 of the models remains.

Note that other emission line ratios could also be used to filter models, but they all involve abundance ratios that are not free parameters in our modeling process (e.g. using  $[\text{S III}]/\text{H}\beta$  depends on the S/H abundance ratio). In other words, not fitting the observed  $[\text{S III}]/\text{H}\beta$  ratio for a given observation may only indicates that the S/H abundance is not correct; but S/H is determined by fixing S/O, we have no way to act on  $[\text{S III}]/\text{H}\beta$  ratio. The fact that S/H may be incorrect has virtually no consequences on the results presented in the following sections, but this would artificially exclude "good" models if we use the  $[\text{S III}]/\text{H}\beta$  ratio as a filter.

### 3.4. Characterizing the ionizing population

From the study by Morisset et al. (2015), we can define limits in an age-metallicity plane for the ionizing stellar populations. Using their Fig. 5, we can determine that, for  $\log(\text{O}/\text{H}) < -3.5$ , OB stars correspond to  $\log(\text{age}/\text{yr}) < 6.8$  and HOLMES correspond to  $\log(\text{age}/\text{yr}) > 8.25$ . For  $\log(\text{O}/\text{H}) > -3.5$ , OB stars correspond to  $\log(\text{age}/\text{yr}) < 6.7$  and HOLMES correspond to  $\log(\text{age}/\text{yr}) > 7.9$ . Using these limits and the decomposition of the spectra on the gsd156 library base, we can define what is the proportion of the ionizing photons coming from OB stars in the total number of ionizing photons  $Q_0$ . This proportion will be named  $f(\text{OB})$  in the rest of the paper. There is a strong correlation between



**Fig. 2.** Histogram of  $Q_{0/1} = Q(\text{H}^0)/Q(\text{He}^0)$  for the OB and HOLMES dominated spectra.

the type of dominant ionizing stellar population depicted by this  $f(\text{OB})$  and the value of  $Q_{0/1} = Q(\text{H}^0)/Q(\text{He}^0)$ , the ratio of the number of photons ionizing  $\text{H}^0$  and  $\text{He}^0$  (a kind of softness parameter). This is illustrated by Fig. 2, where the histograms of  $Q_{0/1}$  obtained for OB stars and HOLMES are compared. There is a clear separation at a value of  $\sim 0.55$ , HOLMES being associated with the lowest values. In the following we will mainly use the  $Q_{0/1}$  ratio to trace the underlying population, keeping in mind that the purple/bluish dots point to HOLMES and the gray/reddish/yellow dots to OB stars for all figures throughout this article.

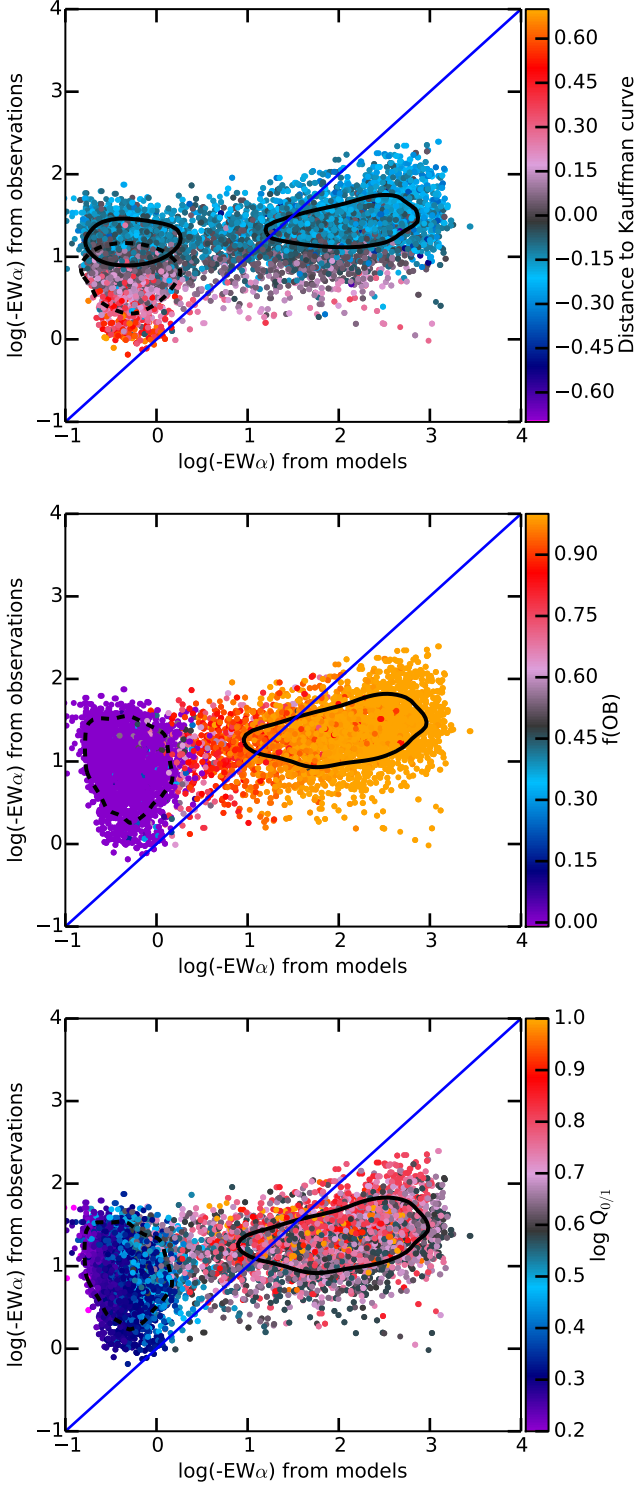
### 3.5. The $\text{H}\alpha$ equivalent width

We can compare the observed  $\text{H}\alpha$  equivalent widths ( $EW_\alpha$ ) with the prediction from the models. We first corrected the observed value from the extinction, using  $\text{H}\alpha/\text{H}\beta = 2.85$  and the Fitzpatrick (1999) extinction law to correct the  $\text{H}\alpha$  flux and using the  $A_V$  from the stellar observations to correct the stellar continuum. It is well known that the stellar continuum is affected by less dust attenuation than the ionizing gas, in general, for star-forming galaxies (Calzetti 2001). The resulting differential correction has a median of  $0.9 \pm 0.1$ , changing virtually nothing in our results, but adding a few aberrant values due to bad observations of the  $\text{H}\alpha/\text{H}\beta$  ratio. Finally we did not apply the correction.

We plot in Figure 3 the comparison between the  $EW_\alpha$  from the observations and those from the computed models. Three color codes are used: the first one in the upper panel is associated to the distance to the Kauffman curve  $K_{\text{dist}}^6$  (upper panel). In the middle panel  $f(\text{OB})$  is used as color code while in the bottom panel, the color code is following  $Q_{0/1}$ . The solid line in the plot represents where both values of  $EW_\alpha$  are equal.

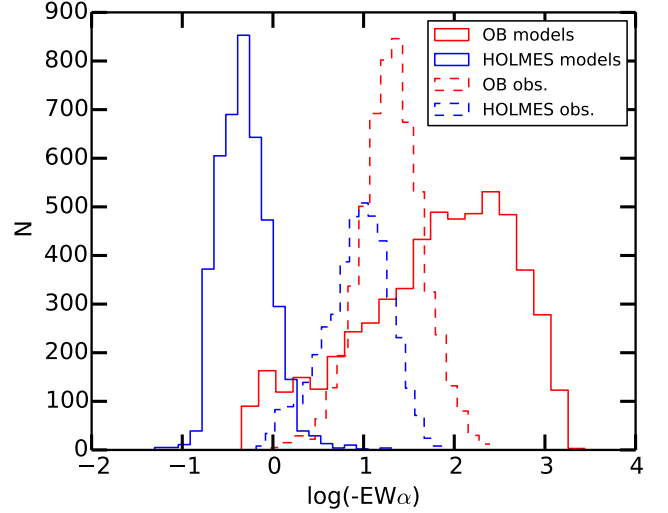
We plot in Figure 4 the histograms of  $EW_\alpha$  for the observation and the models, for the regions ionized by HOLMES (in

<sup>6</sup> The  $K_{\text{dist}}$  parameter is defined by the minimum value of the distance  $D$  between a given point and the Kauffmann et al. (2003) curve, being negative for points below the curve and positive otherwise. The distance between 2 points in the BPT diagram is defined by  $D = \sqrt{\Delta(\log([\text{O III}]/\text{H}\beta))^2 + \Delta(\log([\text{N II}]/\text{H}\alpha))^2}$



**Fig. 3.** Comparison between the  $EW_\alpha$  from the models and from the observations. Colors are representing  $K_{dist}$ , the distance to the Kauffmann et al. (2003) curve (upper panel), the OB stars proportion  $f(OB)$  (middle panel), and the ratio  $Q_{0/1} = Q(He^0)/Q(H^0)$  (lower panel). Solid blue line follows  $y=x$ . In the upper panel, the solid black curves enclose half of the models which are below the Kauffmann curve (negative  $K_{dist}$ ), while the dashed black curve encloses half of the models above the same curve (positive  $K_{dist}$ ). In the middle panel, the solid black contour encloses half of the models ionized by OB stars ( $f(OB) > 0.5$ ), while the dashed black contour encloses half of the models ionized by HOLMES ( $f(OB) < 0.5$ ). In the lower panel, the solid black contour encloses half of the models ionized by OB stars ( $Q_{0/1} > 0.55$ ), while the dashed black contour encloses half of the models ionized by HOLMES ( $Q_{0/1} < 0.55$ ).

Article number, page 6 of 22



**Fig. 4.** Histogram of  $EW_\alpha$  for the models (solid lines) and observations (dashed lines) for star forming regions (red,  $Q_{0/1} > 0.55$ ) and HOLMES ionized regions (blue,  $Q_{0/1} < 0.55$ ).

blue) and by OB stars (in red). These histograms actually correspond to the bottom panel of Figure 3.

To our knowledge, this is the first time the two values of  $EW_\alpha$  (observed and modeled) are compared for such a set of objects where a detailed determination of the underlying stellar population is available.

We can easily see that there is a clear trend in the variations of the colors in the middle and lower panels, indicating that the ratio  $Q_{0/1}$  is a good indicator of the type of stars dominating the ionizing flux (see also Fig. 2 in Sec. 3.4). The color code used for  $Q_{0/1}$  reflect the fact that OB stars (red/orange) are the coolest of our sample, while HOLMES (the hottest) are in violet.

In the upper panel of Fig. 3 we explore how the differences between the observed and modeled  $EW_\alpha$  are related to the  $K_{dist}$  parameter. We can see that most of the regions below the Kauffmann curve ( $K_{dist} < 0$ , blue regions), that correspond to classical HII regions, have observed values of  $EW_\alpha > 10$ , while most of the regions above the same curve (HOLMES, low-ionization nuclear emission-line region, i.e. LINERS etc) have observed values of  $EW_\alpha < 10$ .

The distribution of the source of the ionizing photons described by  $f(OB)$  and  $Q_{0/1}$  are clearly bimodal: there is no example of regions ionized by OB stars in a proportion around 50%. This means that we found a clear separation in the models between classical HII regions and regions ionized by old stars. Fig. 3 also shows the contours enclosing the two different populations highlighted in each panels: the models falling above and under the Kauffmann curve (positive and negative values of  $K_{dist}$  resp., top panel) and the regions ionized by OB stars and by HOLMES (value of  $f(OB)$  and  $Q_{0/1}$ , middle and lower panels respectively). This indicates if these two populations are distinguishable using the values of  $EW_\alpha$ .

The  $EW_\alpha$  is related to the ratio between the number of ionizing photons actually processed by the gas and the number of (ionizing) stars. Both the gas and the stars are supposed to be included in the observed beam. We can see from the Fig. 3 that there is an obvious trend between the observed and modeled  $EW_\alpha$ . The regions lying on the right side of the  $y = x$  line, thus having an observed  $EW_\alpha$  lower than the modeled value, corre-

spond to regions where less photons are ionizing the gas than what is expected in a closed geometry model, leading to what is commonly called "leaking". Stasińska et al. (2001) also proposed the presence of old populations to explain this discrepancy, but we include these old stars in our modeling process. This leaking can be due to matter-bounded regions (in which there is not enough material to be ionized in some directions from the ionizing source point of view) or because of a covering factor less than one (in some directions there is no gas at all), or even a combination of both effects. In both cases, some ionizing photons escape the region. In the case of a covering factor less than one, the photoionization models correctly predict the line ratios (only the absolute fluxes are overestimated, all in the same proportions) contrary to the case of matter-bounded regions. The presence of strong [O I] and [O II] emission lines favors the idea of these H II regions having a covering factor less than one, then validating the use of the photoionization models.

On the other side of the  $y = x$  line (the left one), the  $EW_\alpha$  from the observations is higher than the one from the models. This corresponds to regions where the amount of ionized gas is too much compared to what could be ionized by the observed stellar population. It means that we are missing some of the ionizing sources for these regions. It can actually be the effect of photons coming from sources out of the beam, perhaps escaping from the same regions previously described (located on the right side of the line). It would be very interesting to estimate the number of leaking photons in each galaxy and then see if they are enough to explain the discrepancies in the  $EW_\alpha$  determined from observations and models, but this is out of the scope of this paper.

The difference between the observed and the modeled  $EW_\alpha$  is clearly related to the proportion of OB stars (top and middle panels of Fig. 3) or the hardness of the ionizing radiation (bottom panel). Most of the star forming regions are located on the right side of the  $y = x$  line and correspond to photon leaking. On the other hand, the regions mainly photoionized by HOLMES are only on the left side and correspond to regions ionized by additional sources. The same conclusion can be reached analyzing Fig. 4, where observed star forming regions (red dashed line) have a median  $\log(EW_\alpha)$  of  $\sim 1.3$ , while the corresponding models have a median  $\log(EW_\alpha)$  of  $\sim 1.9$ , leading to a factor of leaking  $\sim 4$  (80% of the photons escape). This median leaking is the same if we use  $EW_\beta$ , as all the equivalent widths are shifted by  $\sim -0.5$  dex. Notice that the distribution of the  $EW_\alpha$  values for the models of H II regions is broader than the distribution of the observed values, reflecting the variety of morphologies leading to leaking factors from 1.0 (no leaking) to some tens (See also Papaderos et al. 2013). When considering regions ionized by HOLMES (blue lines in Fig. 4), we see that the discrepancy between the observed and modeled values of  $EW_\alpha$  goes in the other direction (regions in the left part of the Figs. 3), with  $\log(EW_\alpha)$  median values of  $\sim 1.0$  and  $\sim -0.3$  for observation and models respectively.

Previous studies have analyzed the nature of the LINER-like emission in galaxies (Papaderos et al. 2013; Singh et al. 2013; Sarzi et al. 2010). All of them concluded that the nature of this ionization is most probably due to post-AGBs stars, i.e., HOLMES in our nomenclature. In particular Papaderos et al. (2013) and Gomes et al. (2016) presented a comparison between the observed H $\alpha$  fluxes and the predicted ones based on photoionization models which ionizing source was selected from the analysis of the underlying stellar population for a sample of early-type galaxies. Thus, their analysis is somehow similar to the one presented here. They found that there are two kinds of

galaxies on the basis of this comparison: type-i, for which the observed and predicted fluxes match very well, in general, and type-ii, for which they describe a deficit of observed flux, compatible with a Lyman-continuum leaking, in agreement with the results presented here. We need to recall that by selection procedure adopted we have excluded a substantial fraction of the diffuse regions, that are those ones dominating the type-i ETGs.

We must notice that state-of-the-art population spectral synthesis models are still plagued by significant degeneracies (e.g., the notorious age-metallicity degeneracy) and uncertainties in the best-fitting star formation history. A tiny variation/uncertainty in the mass fraction of young ( $<15$ -20 Myr) ionizing stars (simple stellar population-SSP models in this study) results in a very significant change in the expected value of  $Q_0$ , consequently the Balmer recombination line luminosities. On the other hand, the discrepancy between the observed and modeled  $EW_\alpha$  in case of HOLMES can just come from an underestimation of the ionizing flux from post-AGBs. These kind of short-lived, highly variable period of the evolution of stars is not very well understood, and its inclusion in SSP models is quite recent, being still a topic that present large uncertainties. Even in the case that we derive correctly the fraction of SSPs comprising post-AGBs it may be still the case that the predicted ionizing photon distribution is not totally correct. For all those reasons we prefer not to include those ionized regions in further analysis (see next section), concentrating ourselves in the much better understood regions ionized by young stars. In further studies we will try to improve our analysis: (1) excluding or subtracting the possible contribution of a central AGN, if feasible, and (2) updating as much as possible the SSPs and the ionization models adopted for post-AGB stars.

### 3.6. The BPT-population filter

The regions that correspond to star forming regions in the BPT diagram (negative  $K_{dist}$  values) and that are on the left side of the  $y=x$  line in the Fig. 3.5 are not leading to trustable models, as the ionizing source should be of OB-stars type, and what is obtained from the SSP decomposition is of type HOLMES. We apply another filter to the ad-hoc models to remove objects that have negative  $K_{dist}$  values and that are ionized by old populations. Applying this BPT-population filter to the ad-hoc models already filtered by [O II]/H $\beta$  leads to a set of 2558 models, from which 76% are star forming regions, and 24% are regions ionized by HOLMES and which are over the Kauffmann curve. In the 3MdB database, we set to 1 the value for the "com8" field for the models that fit this filter (and 0 otherwise).

The final 2558 models used in the next section fit the [N II], [O II], and [O III] lines with the following mean value and standard deviation: [N II]/H $\beta$  Model/Obs =  $1.05 \pm 0.08$ , [O II]/H $\beta$  Model/Obs =  $0.98 \pm 0.13$ , and [O III]/H $\beta$  Model/Obs =  $1.04 \pm 0.14$ .

In the following sections, the figures show the star forming regions and the ones ionized by HOLMES. But we select only star forming regions to compute fits to our results, given that we are actually not sure about the pertinence of the HOLMES models (even after applying the filter described above): the missing photons may have a very different distribution, and the O/H abundance have been obtained using the O3N2 ratio, calibrated on H II regions, not on HOLMES-ionized regions.

## 4. Results and discussion

In all the following sections, we will present results obtained using the ad-hoc "Neb" models selected after applying the [OII] filter described in Sec. 3.3 and the BPT-population filter described in Sec. 3.6.

### 4.1. A set of models nearly compatible with the direct method

In Fig. 5, we show the values of two  $T_e$ -diagnostic line ratios as a function of the oxygen abundance (left panels) and the  $([O\text{II}]\lambda 3727 + [O\text{III}]\lambda 5007)/H\beta$  line ratio (right panels) for the photoionization models (only the "Neb" models, colored circles), along with one of the largest compilation of H II regions up-to-date with O/H derived using the direct-method (M13) and for a set of observations used by Marino et al. (2013) (black diamonds). Top and bottom panel are showing the respective  $[O\text{III}]\lambda 4363/5007$  and  $[N\text{II}]\lambda 5755/6584$  diagnostics.

We find a good agreement between the models and the observations in all the panels, for metallicities corresponding to  $12 + \log(O/H) > 8.2$ . Our star forming region models (orange/red/purple models) are falling over the coolest observed regions, pointing to a small underestimation of the electron temperature (but see below). The HOLMES<sup>7</sup>-ionizing models (turquoise/blue) models are falling over the hottest regions. This indicates that the set of models computed for this work, for which the O abundances have been calculated with the O3N2 method, is globally compatible with the determination of O/H using a direct method. We check that the "Stel" models are also falling on the same regions in these diagrams, leading to the conclusion that this behavior is not due to the way we define O/H. Other previous sets of photoionization models systematically show discrepancies between the values used as input for the O abundances and the values determined from the direct method (e. g. López-Sánchez et al. 2012). This is illustrated with MAPPINGS models presented in the Fig. 5 with a grid that shows the model results obtained from the tables electronically published by Dopita et al. (2013). All these O/H values are systematically larger than the ones obtained with the direct method (black diamonds), while the results for our models cover the same region and describe the same trend. If we consider only star forming regions, our models are a little bit too cold, as they reproduce only half of the observed values of  $[O\text{III}]\lambda 4363/5007$ . As far as we know this is the first time that it is possible to reconcile the predictions by photoionization models with at least a significant amount of direct estimations of the abundance and the line ratios.

On the right panels we can see the differences between the models obtained for this work and the Dopita et al. (2013) models obtained with MAPPINGS. While our models show a very good agreement with the observations, we can see in the upper and middle panels that some of the MAPPINGS models fall in a region where no observations are found (they turn around the data points cloud).

The main differences between the two sets of models seems to be the  $T_e$  obtained for a given value of O/H, depicted for example by a difference of  $\approx 0.5$  dex in  $[O\text{III}]\lambda 4363/5007$  at  $\log(O/H) = -3.5$ , MAPPING models being hotter. This is probably the result of a lower heating or a higher cooling in our models. A lower heating can be explained by the use by Dopita et al. (2013) of Starburst99 models (Leitherer et al. 1999) from 2005, which provides a harder radiation field than the newest models.

In contrast, we use in our models the SED predicted by POP-STAR models (Mollá et al. 2009).

A higher cooling can be due to our higher values of N/O for a given O/H (see Sec. 4.5), our lower values of  $\log(\bar{U})$  for a given O/H (see Sec. 4.6) or the fact that we do not consider depletion of some elements. To test the effect of N/O and  $\log(\bar{U})$  on the electron temperature, we extract a subset of models from 3MdB (Morisset et al. 2015). We use the "HII\_CHIm" models (Pérez-Montero 2014) with  $\log(O/H) = -3.5$  and an age of the ionizing stellar cluster of 1 Myr, N/O and  $\log(\bar{U})$  let free. We show in Fig. 6 both effects on the line ratio  $[O\text{III}]\lambda 4363/5007$  of changing N/O (on the x-axis) and  $\log(\bar{U})$  (the color code). We can see that a difference in  $\log(N/O)$  from -0.8 to -1.5 implies a very small difference on the  $T_e$ -diagnostic line ratio ( $\approx 0.05$  dex). Changing  $\log(\bar{U})$  from -2.5 to -3.5 leads also to a very small effect on  $[O\text{III}]\lambda 4363/5007$  ( $\approx 0.05$  dex). To test the effect of the depletion of some elements on the electron temperature, we run two models with 1) the abundances from Asplund et al. (2009) as used in our models, and 2) a depletion of 1 dex for Si and Mg and of 1.5 dex for Fe. The effect on the line ratio  $[O\text{III}]\lambda 4363/5007$  is of increasing it by 0.09 dex, not enough to explain the observed difference of  $\approx 0.5$  dex at  $\log(O/H) \approx -3.5$ , but almost enough to increase the ratio to the region where the observations are.

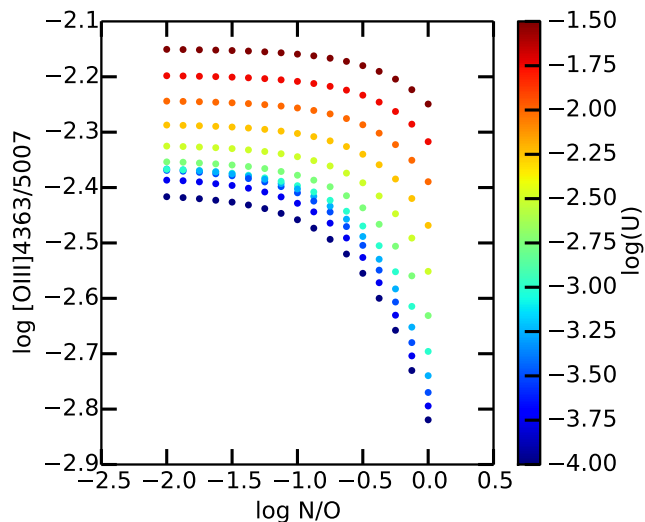
We conclude that the differences observed in Fig. 5 (especially in the upper-left panel) between our grid of models and the models from Dopita et al. (2013) are neither due to the differences in N/O, or the differences in  $\log(\bar{U})$  nor the use or not of depletion. It may reside in the choice of the ionizing SED, if not coming from the code used. In summary, our photoionization models are almost compatible with the electron temperature derived from the direct method at a given O/H, being the first ones in the literature to our knowledge. Our electron temperature is a little bit cool, perhaps pointing to a small lack of depletion of some elements or the presence of some light extra heating process or the presence of a process favoring the emission of high temperature lines (temperature fluctuations *a la* Peimbert (1967),  $\kappa$  distribution *a la* Nicholls et al. (2012)). Indeed, the differences in the selected ionizing source could explain and solve the long standing incompatibility between the direct method and photoionization models, and we are presenting in this work a set of models approaching the observational reality.

### 4.2. The BPT diagram

Figure 7 shows our photoionization models in the classical BPT diagram (Baldwin et al. 1981),  $\log([O\text{III}]/H\beta)$  versus  $\log([N\text{II}]/H\alpha)$ . The curves derived by Kauffmann et al. (2003), Kewley & Dopita (2002) and Stasińska et al. (2006) have been included in the plots for reference. These curves are often used to distinguish between star-forming regions (below the envelope empirically defined by Kauffmann et al. 2003) and AGNs (above the envelope defined by Kewley & Dopita 2002). The color bars located on the right side run from low to high values of the O abundance, the N/O ratio, the  $H\alpha$  equivalent widths (determined from observations and from models), the ionization parameter  $\log(\bar{U})$  and the fraction of OB stars  $f(\text{OB})$  (from upper-left to lower-right panels respectively). We plot here only the results concerning the models where the O/H abundance is determined from the O3N2 relation from M13 ("Neb" models, see Sec. 3.1).

We can see from the middle panels a general trend of the models with lower  $EW_\alpha$  and lower  $f(\text{OB})$  to be located above the Kauffmann et al. (2003) curve, although there are also models with low  $EW_\alpha$  and low  $f(\text{OB})$  below this curve. We can also see that there are no photoionization models with high  $EW_\alpha$  and

<sup>7</sup> HOLMES stands for HOt Low Mass Evolved Stars, see e.g. Flores-Fajardo et al. (2011)



**Fig. 6.** Electron temperature diagnostic line ratios  $[\text{O III}]\lambda 4363/5007$  vs.  $\text{N/O}$ . The color bar is indicating  $\log(\bar{U})$ . The models are extracted from 3MdB and correspond to "HII\_CHIm" models of 1 Myr HII regions with  $\log(\text{O/H}) = -3.5$ .

$f(\text{OB})$  above the Kewley & Dopita (2002) curve. Our models reproduce the observational results obtained by Sánchez et al. (2015): below the curve, the models with higher O abundance and lower U are located in the lower right region whereas those with a low O abundance and high U are located in the upper left corner (Evans & Dopita 1985; Veilleux & Osterbrock 1987; López-Sánchez et al. 2012). Notice that the middle left plot uses only observed data and then show more numerous points. The color separation in the BPT diagram is clearer when using the observed  $EW_\alpha$  (left panel), while the mixing of the colors is more important for the  $EW_\alpha$  from the models (right panel). This is coherent with the results from Fig. 3 described in the Sec. 3.5: the observed  $EW_\alpha$  may appear to be a very good diagnostic for the ionizing population (and to derive what kind of nebular region is observed), but when using the  $EW_\alpha$  from the models, the situation is really less obvious.

In the lower left panel of Fig. 7 we use  $\log(\bar{U})$  as color code. It shows a clear gradient of  $\log(\bar{U})$  decreasing in the decreasing  $[\text{O III}]/\text{H}\beta$ -increasing  $[\text{N II}]/\text{H}\alpha$  direction. In the lower right panel of Fig. 7 we use  $f(\text{OB})$  as color code. As in Fig. 3, the two extreme values of  $f(\text{OB})$  are dominating the distribution and are not very well separated. We can also see some regions where the HOLMES dominate the ionizing SED well inside the star forming region, while almost no OB-stars dominated regions enter the part of the BPT diagram between the Kauffman and the Kewley curves, and not at all above the Kewley curve, in agreement with the definition of these demarcation line.

Figure 8 shows the comparison between different BPT-type diagrams taken from Baldwin et al. (1981); Veilleux & Osterbrock (1987):  $\log([\text{O III}]/\text{H}\beta)$  versus  $\log([\text{N II}]/\text{H}\alpha)$ ,  $\log([\text{O III}]/\text{H}\beta)$  versus  $\log([\text{O II}]/[\text{O III}])$ ,  $\log([\text{N II}]/\text{H}\beta)$  versus  $\log([\text{O II}]/[\text{O III}])$ , and  $\log([\text{O I}]/\text{H}\alpha)$  versus  $\log([\text{O II}]/[\text{O III}])$ . In these plots, we use the  $Q_{0/1}$  ratio as color code, tracing the softness of the ionizing flux.

The  $\log([\text{O III}]/\text{H}\beta)$  versus  $\log([\text{O II}]/[\text{O III}])$  and  $\log([\text{O I}]/\text{H}\alpha)$  versus  $\log([\text{O II}]/[\text{O III}])$  ratios (upper right and lower right panels, respectively) do not depend on the N/O abundance ratio. Both plots exhibit a clear separation between

the regions ionized by OB stars  $Q_{0/1} > 0.55$  and those ionized by HOLMES (otherwise), but note that in the  $\log([\text{O III}]/\text{H}\beta)$  versus  $\log([\text{O II}]/[\text{O III}])$  plot, some regions ionized by HOLMES are mixed with the OB star ionized regions whereas the separation in the  $\log([\text{O I}]/\text{H}\alpha)$  versus  $\log([\text{O II}]/[\text{O III}])$  plot is clearer (in agreement with the results pointed out by Baldwin et al. (1981). One must keep in mind that the values of  $[\text{O I}]/\text{H}\alpha$  used here are pure predictions from the models, not necessarily reproducing the observations.

#### 4.3. WHAN diagram

Figure 9 shows the WHAN diagram for our models. This diagram is based on  $\text{H}\alpha$  and  $[\text{N II}]$  lines and was proposed by Cid Fernandes et al. (2010) to determine the ionizing population. In the left panel we show the  $EW_\alpha$  from the models and in the right panel the ones derived from the observations.

The general trend is that the regions ionized by OB stars have higher  $EW_\alpha$  than the regions ionized by HOLMES (note that the  $EW_\alpha$  from the models cover a wider range of values than the  $EW_\alpha$  derived from the observations).

The separation between the regions ionized by OB stars and those ionized by HOLMES is very clear when using the  $EW_\alpha$  from the models whereas the different type of ionized regions are more mixed in the other panel. Another interesting result from this figure is that there is no correlation between the  $EW_\alpha$  from the models and the  $[\text{N II}]/\text{H}\alpha$  ratios but there is a clear trend between the  $EW_\alpha$  computed from the observations and the  $[\text{N II}]/\text{H}\alpha$  ratios. This indicates that if we only know the  $[\text{N II}]/\text{H}\alpha$  ratio we cannot distinguish between regions ionized by OB stars or HOLMES. Thus, the knowledge of the properties of the underlying stellar population is a fundamental tool to distinguish between both types of regions. This was already discussed by Sanchez et al. (2014). Another interesting result is that while a limit of  $EW_{\alpha\text{mod}} > 3\text{\AA}$  to segregate classical H II regions from HOLMES is well predicted by the models, in practice, an empirical cut of  $EW_\alpha > 6\text{\AA}$  or the demarcation line

$$\log(EW_\alpha) > 2 + 2 \log([\text{N II}]/\text{H}\alpha)$$

seems to separate both ionizing regions better.

#### 4.4. $\log(\bar{U})$ vs. $[\text{O II}]/[\text{O III}]$ and $[\text{S II}]/[\text{S III}]$

The  $[\text{O II}]/[\text{O III}]$  and  $[\text{S II}]/[\text{S III}]$  line ratios has been frequently used as tracers of the ionization strength (Diaz 2001; Kewley & Dopita 2002), based on empirical correlations between these line ratios and this parameter (Dors & Copetti 2003). However, as we have seen along this article, previous results are sometimes based on photoionization models which ionization source was not selected to match any observed constraint. Therefore, it is important to revise the trends between those parameters based on our new set of models. In Figs. 10 and 11, we explore the relations between  $\log(\bar{U})$  and the  $[\text{O II}]/[\text{O III}]$  and  $[\text{S II}]/[\text{S III}]$  line ratio, respectively.

In Fig. 10 we show the first of these relations for each of the geometries considered in our models: thin shell and empty sphere (left and right panel respectively). We overplot the relation determined by (Díaz et al. 2000, hereafter D00). This relation does not match the results of our models, specially the ones corresponding to star forming regions ( $Q_{0/1} > 0.55$ , gray/red points). In the case of thin shell geometry (left panel), there is an underestimation of  $\log(\bar{U})$  by  $\sim 0.25$  dex with a scatter of more

or less the same amount. In the case of a geometrically thick region (right panel), the slope of the D00 relation is not recovered. Our results indicate that there is a steeper relation, although the average value is similar to that predicted by D00. Our fits limited to the regions where  $Q_{0/1} > 0.55$  (star forming regions) leads to:

$$\log(\bar{U}) = -2.74 \pm 0.02 - 1.00 \pm 0.03 \times \log([\text{O II}]/[\text{O III}]) \quad (1)$$

with a standard deviation of 0.14 and

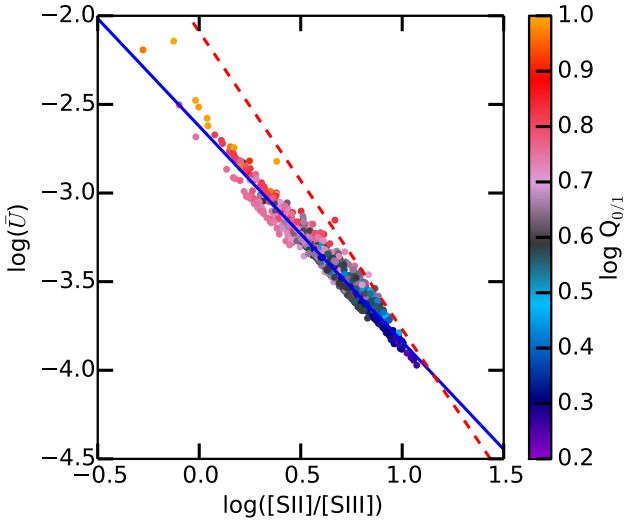
$$\log(\bar{U}) = -2.38 \pm 0.04 - 2.36 \pm 0.10 \times \log([\text{O II}]/[\text{O III}]) \quad (2)$$

with a standard deviation of 0.22, for the left (thin shells) and right (filled spheres) panel respectively.

In Fig. 11 we do not split the two geometries in different plots, as the results are very similar for each of them. A tighter relation is derived in comparison with the one derived for  $[\text{O II}]/[\text{O III}]$ . We overplot the relation determined by Diaz et al. (1991), showing a different slope but similar values at low ionization. The changes in the atomic data (especially for  $[\text{S II}]$ ) since 1991 may explain the differences. The following linear fit, limited to the regions where  $Q_{0/1} > 0.55$  (star forming regions), is obtained:

$$\log(\bar{U}) = -2.62 \pm 0.01 - 1.22 \pm 0.01 \times \log([\text{S II}]/[\text{S III}]) \quad (3)$$

with a standard deviation of 0.06.



**Fig. 11.**  $\log(\bar{U})$  vs.  $[\text{S II}]/[\text{S III}]$  for the models. The colors coding the hardness of the ionizing radiation. As there is no noticeable difference due to the morphology of the region, contrary to the case of  $[\text{O II}]/[\text{O III}]$  shown in Fig. 10, both morphologies are plotted here in the same figure. The blue line corresponds to our linear fit (taking only star forming regions into account i.e.  $Q_{0/1} > 0.55$ ) see text for the values. The red dashed line corresponds to the fit by Diaz et al. (1991).

#### 4.5. N/O vs. O/H

The nucleosynthesis paths for Nitrogen and Oxygen are different. The origin of Nitrogen is both primary, produced from the initial content of Hydrogen, and secondary, produced from the initial content of Carbon and Oxygen created by previous stellar generations (e. g. Matteucci 1986). Therefore, the N/H and

O/H abundance ratios are not supposed to evolved in lockstep. The N/O vs. O/H relation derived from observations is mainly horizontal for  $\log(\text{O}/\text{H}) < -4.1$  where the Nitrogen is of primary production, and almost linear above this, where the secondary production of Nitrogen take place (see e.g. Vila Costas & Edmunds 1993; Thurstun et al. 1996; Chiappini et al. 2003; Liang et al. 2006; Mollá et al. 2006, and references therein). The regions we are using in this work are all above this limit and only allow us to probe the linear part of the relation.

Figure 12 shows the relation between the abundance ratio N/O and the oxygen abundance O/H. Left and right panels use different color code: stellar population as depicted by the softness parameter  $Q_{0/1}$  and  $\log(\bar{U})$  respectively. Remember that the N/O abundance ratio is one of the main results, with  $\log(\bar{U})$ , of the fitting process we applied, see Sec. 3.1.

The difference between the regions ionized by OB stars and by HOLMES is shown in the left panel: at a given metallicity, the N/O ratio is lower in the second group ( $Q_{0/1} < 0.55$ , blue points) than in the classical star forming regions ( $Q_{0/1} > 0.55$ ). This is also the case even if we define the nebular metallicity from the stellar content ("Stel" models, see Sec. 4.8). These differences in the N/O determinations may be an indication that the amount of Nitrogen of the regions ionized by HOLMES is systematically lower than in the star forming regions.

If we concentrate on the star forming regions only, we can determine a linear fit to the relation between N/O and O/H:

$$\log(\text{N}/\text{O}) = -16.09 \pm 0.40 + 1.81 \pm 0.04 \times (12 + \log(\text{O}/\text{H})). \quad (4)$$

with a standard deviation of 0.13.

This fit is valid only on the O/H range available with the data used here, namely between  $8.1 < 12 + \log(\text{O}/\text{H}) < 8.8$ . There a clear color gradient following  $\log(\bar{U})$  in the right panel, pointing to a second order effect of the gas ionization stage on the relation.

We compare this fit to the ones determined by Pilyugin et al. (2012, light green dashed line) and by Dopita et al. (2013, red dotted line). There is a good agreement with the first one, and a clear offset with respect to the second. This result is expected, since Pilyugin et al. (2012) adopted oxygen abundances derived using a direct method, while Dopita et al. (2013) uses a particular set of photoionization models that over-estimate the abundances. Again, the result confirms our previous claim that our photoionization models are compatible with direct method estimations, contrary to some previous results.

#### 4.6. $\log(\bar{U})$ vs. O/H

We plot in Fig. 13 the position of the models in the  $\log(\bar{U})$  vs.  $12 + \log(\text{O}/\text{H})$  diagram. We also added the empirical relations from Dopita & Evans (1986), Dopita et al. (2006) and Pérez-Montero (2014). We can easily see that the  $\log(\bar{U})$  we determined are lower by around 0.7 dex than the values obtained from these relations. Or that our oxygen abundance is between 0.4 and 0.7 dex lower, depending on the relation we consider. This apparent discrepancy is principally due to the way we determined the oxygen abundance (using O3N2 from M13); using the abundance estimator from e. g. Kewley & Dopita (2002) would lead to higher values for O/H and could reconcile the values we obtained with the different relations shown here. See also the discussion of the results obtained with the "Stel" models in Sec. 4.8.

The effect of the ionizing stellar population (used for the color code) is also very clear. The  $\log(\bar{U})$  values for the regions ionized by HOLMES are 0.5 dex lower than for the classical HII regions (ionized by OB stars).

Our fits to the models corresponding to star forming regions ( $Q_{0/1} > 0.55$ ) are:

$$\log(\bar{U}) = 8.79 \pm 0.76 - 1.43 \pm 0.09 \times (12 + \log(O/H)) \quad (5)$$

with a standard deviation of 0.20 and

$$\log(\bar{U}) = 8.42 \pm 2.02 - 1.37 \pm 0.24 \times (12 + \log(O/H)) \quad (6)$$

with a standard deviation of 0.20 for the left (thin shell) and 0.36 for the right (filled sphere) panel respectively. Notice the very high uncertainties and dispersion in the case of filled sphere models.

#### 4.7. The variation of the $\eta$ parameter with the ionizing SED

Following Vilchez & Pagel (1988), we plot in Fig. 14 the values of  $\eta = (O^+/O^{++})/(S^+/S^{++})$  vs.  $S^+/S^{++}$  and  $\eta' = ([O II]\lambda 3727+ / [O III]\lambda 5007+) / ([S II]\lambda 6720+ / [S III]\lambda 9067+)$  vs.  $[S II]\lambda 6720+ / [S III]\lambda 9067+$ . We show that these  $\eta$  and  $\eta'$  depend on the softness of the ionizing SED, represented here by  $Q_{0/1}$  as color code, as already pointed out by Vilchez & Pagel (1988). But we also see here that the  $\eta'$  depends strongly on the geometry too, left panel being the thin models ( $fr = 3.0$ ), and right panels the filled sphere models ( $fr = 0.03$ ). Given the scatter observed in each panel and the relatively high difference between the two geometries, the relation between the  $\eta$ 's and the softness of the ionizing radiation  $Q_{0/1}$  is far from being established.

#### 4.8. Effect of the nebular abundance determination on the results

We redraw all the figures presented in the previous sections, but now using the models obtained with the "Stel" determination of the nebular O/H (see Sec. 3.1). The corresponding figures are available in the online edition appendix A. The main differences between the two sets of models are obviously seen in plots directly involving the nebular metallicity, for example in the upper left panel of Figs. 7/A.1 ("Neb"/"Stel" models resp.) where the color distribution is strongly affected. Notice that the position of the models in all the BPT diagrams are the same, as each model is reproducing the  $[N II]/H\alpha$ ,  $[O II]/H\beta$ , and  $[O III]/H\beta$  line ratio, whatever the way its nebular metallicity is obtained.

The results related to the relation between  $\log(\bar{U})$  and  $[O II]/[O III]$  shown in Figs. 10/A.2 are a little bit affected, but the main conclusion stay unchanged: the relation from Díaz et al. (2000) is not recovered. When using  $[S II]/[S III]$  as in Figs. 11, there is a perfect match between the both sets of models. This leads to the conclusion that the results obtained here, in particular the fit from Eq. 3 (but also 1 and 2), are very robust.

The more important differences are obtained for 12/A.3 and Figs. 13/A.4 showing the relations between  $\log(\bar{U})$  and N/O vs. O/H respectively. Clear relations are not obtained at all when the nebular metallicity is derived from the stellar content ("Stel" models). Nevertheless, those two relations exist and have been observed using other methods. The lack of relation between  $\log(\bar{U})$  and N/O vs. O/H when using the "Stel" models indicates that 1) the "Stel" models are not correct regarding the nebular metallicity, and 2) that these relations obtained with the "Neb" models may be dependent on the way the nebular metallicity is obtained.

## 5. Conclusions

We present in this paper a set of photoionization models based on the CALIFA HII regions catalog. Each model uses as ioniz-

ing SED the combination of POPSTAR stellar population models (Mollá et al. 2009), based on the analysis of the continuum spectra performed by the FIT3D program (Sánchez et al. 2011) for the corresponding region. Each model corresponds to an interpolation in the  $\log(\bar{U})$  vs. N/O parameter space to fit the observation of the  $[N II]/H\alpha$  and  $[O III]/H\beta$  line ratios of an individual HII region. Two different morphologies (filled or empty sphere) as well as two different ways to derive O/H (from O3N2 and from the stellar population) are explored. The fact that O/H is determined by strong line method leads to qualify these models as "hybrid" ones. We finally filter the models by selecting only the ones that also fit of the  $[O II]/H\alpha$  line ratio and excluding the ones that correspond to HOLMES ionizing spectrum and fall in the star forming region of the BPT diagram. We obtain a set of 2,558 models, each one fitting simultaneously the tree line ratios of a given HII region. This incomparable database allows us to explore relations between parameters, with the possibility to take into account the effect of the way the nebular metallicity is defined, or the morphology of the region.

The ionizing stellar population can be divided in two groups: classical OB-stars ionizing star forming regions, and HOLMES resulting from the evolution of old starbursts. Three quarters of the models correspond to star forming regions. We found that the first regions show a difference in the  $H\alpha$  equivalent width between models and observations that can be interpreted as the result of a median leaking of 80% of the photons. On the contrary, for the HOLMES ionized regions we found that the models predict higher values for  $EW_\alpha$  than what is actually observed. This can be understood as missing ionizing photons compared to what would be needed to produce the HI recombination lines.

We show that our models are mainly compatible with the electron temperature derived from observations for a given value of O/H, which was not the case for previously publish grids of models (Dopita et al. 2013, e.g.). We attribute the better match of our models to the use of a detailed ionizing SED obtained from the stellar underlying population for each region.

We derive new relations between  $\log(\bar{U})$  vs.  $[O II]/[O III]$  and  $[S II]/[S III]$ , showing that the first one strongly depends on the morphology of the nebula, while the latest one is a very robust result (it does not depend on the way the nebular abundance is determined).

The relation between N/O and O/H we derive is compatible with Pilyugin et al. (2012, using a method based on observations) and not with Dopita et al. (2013, based on photoionization models). The relation between  $\log(\bar{U})$  and O/H is different from the previous determinations, leading for lower values at a given metallicity. We also conclude that  $\eta'$  is not a good indicator of the softness of the radiation field, as it also strongly depends on the morphology of the region.

All the figures presented in this paper can easily be generated by anyone given that the data are available from 3MdB and that the python codes used to make the models and the figures are available from the 3MdB web page <https://sites.google.com/site/mexicanmillionmodels>.

**Acknowledgements.** CM took advantages of the useful discussions with G. Hägele and his co-workers while invited at La Plata University.

We thank the anonymous referee for her/his careful review of the paper and valuable comments.

We want to thank Yago Ascasibar and Polychronis Papaderos, the internal CALIFA referee for this paper, for their useful comments.

Part of the results presented here have been obtained using computers from the CONACyT project CB2010/153985 and the UNAM-PAPIIT project 107215. The computers "Tychos" (Posgrado en Astrofísica-UNAM, Instituto de Astronomía-UNAM and PNPC-CONACyT) have been used for this research.

GDI gratefully acknowledges support from the Mexican CONACyT grant CB-

2014-241732.

SFS thanks the CONACYT-125180 and DGAPA-IA100815 projects for providing him support in this study.

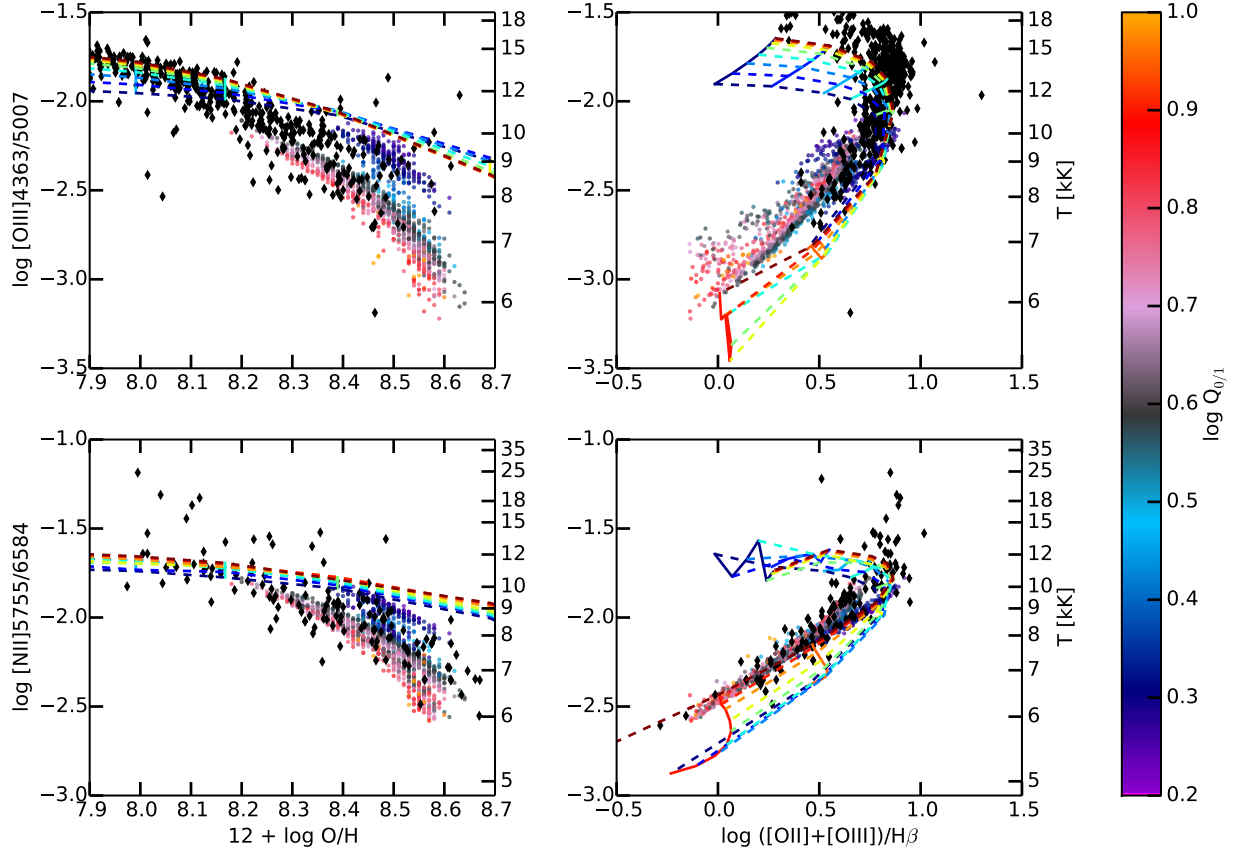
Support for LG is provided by the Ministry of Economy, Development, and Tourism's Millennium Science Initiative through grant IC120009, awarded to The Millennium Institute of Astrophysics, MAS. LG acknowledges support by CONICYT through FONDECYT grant 3140566.

This study uses data provided by the Calar Alto Legacy Integral Field Area (CALIFA) survey (<http://califa.caha.es/>). CALIFA is the first legacy survey performed at Calar Alto. The CALIFA collaboration would like to thank the IAA-CSIC and MPIA-MPG as major partners of the observatory, and CAHA itself, for the unique access to telescope time and support in manpower and infrastructures. The CALIFA collaboration also thanks the CAHA staff for the dedication to this project.

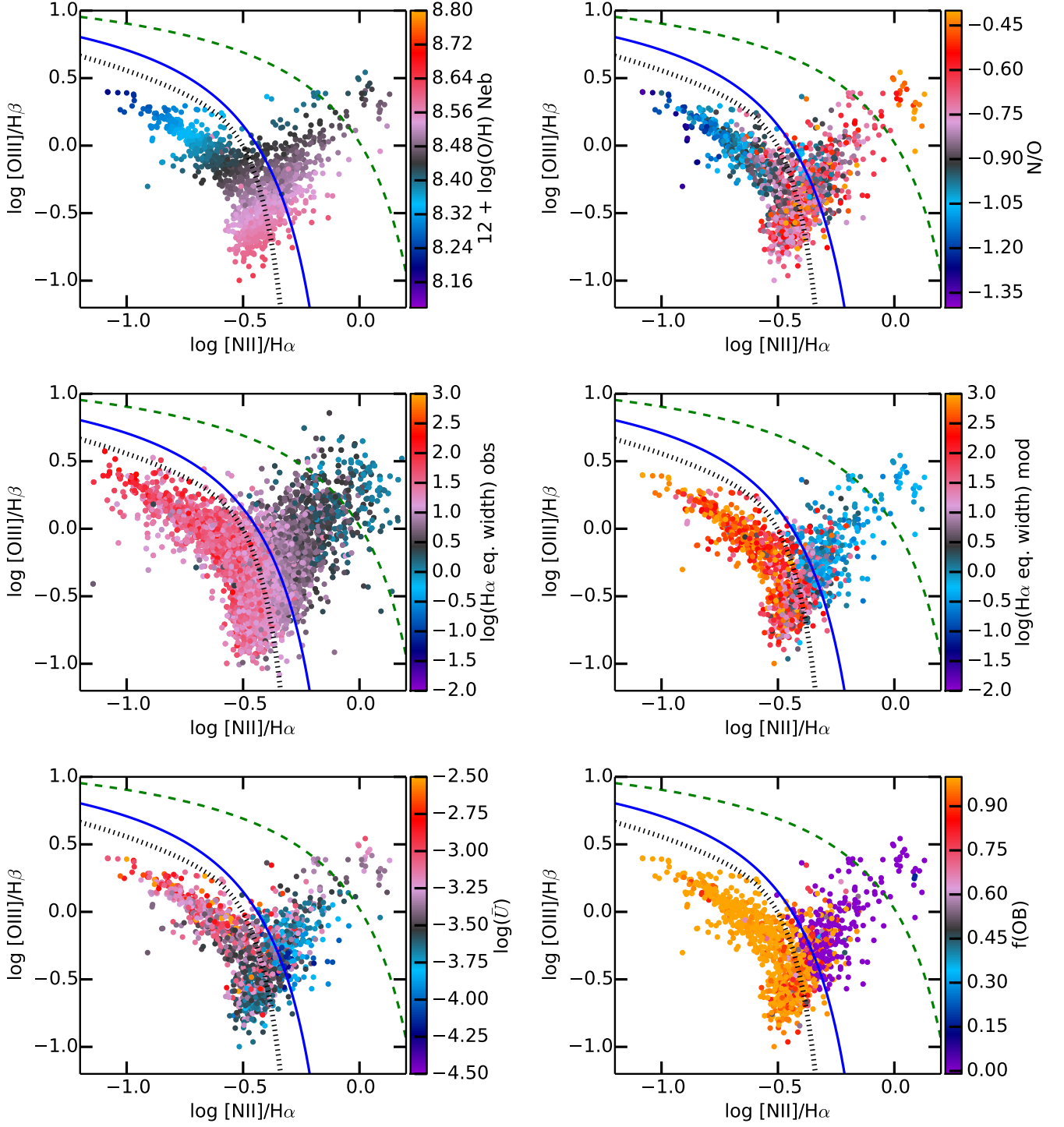
Based on observations collected at the Centro Astronómico Hispano Alemán (CAHA) at Calar Alto, operated jointly by the Max-Planck-Institut für Astronomie and the Instituto de Astrofísica de Andalucía (CSIC).

## References

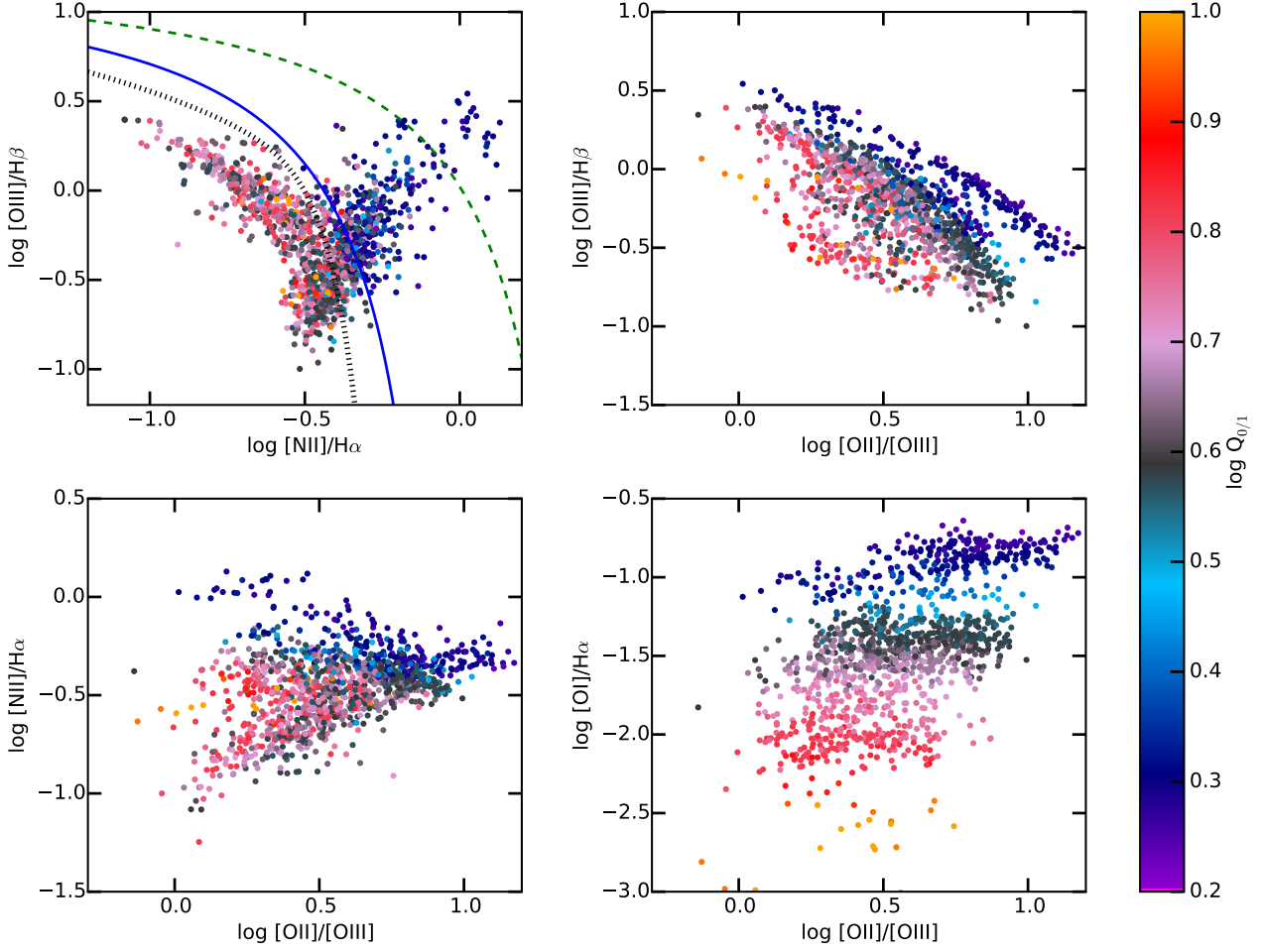
- Anderson, L. D. 2014, in American Astronomical Society Meeting Abstracts, Vol. 223, American Astronomical Society Meeting Abstracts 223, 312.01
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
- Binette, L., Flores-Fajardo, N., Raga, A. C., Drissen, L., & Morisset, C. 2009, *ApJ*, 695, 552
- Binette, L., Magris, C. G., Stasińska, G., & Bruzual, A. G. 1994, *A&A*, 292, 13
- Cairós, L. M., Caon, N., García Lorenzo, B., et al. 2012, *A&A*, 547, A24
- Calzetti, D. 2001, *PASP*, 113, 1449
- Chabrier, G. 2003, *PASP*, 115, 763
- Chiappini, C., Romano, D., & Matteucci, F. 2003, *MNRAS*, 339, 63
- Cid Fernandes, R., González Delgado, R. M., García Benito, R., et al. 2014, *A&A*, 561, A130
- Cid Fernandes, R., Pérez, E., García Benito, R., et al. 2013, *A&A*, 557, A86
- Cid Fernandes, R., Stasińska, G., Schlickmann, M. S., et al. 2010, *MNRAS*, 403, 1036
- Díaz, A. I., Castellanos, M., Terlevich, E., & Luisa García-Vargas, M. 2000, *MNRAS*, 318, 462
- Díaz, A. I., Terlevich, E., Vilchez, J. M., Pagel, B. E. J., & Edmunds, M. G. 1991, *MNRAS*, 253, 245
- Díaz, M. 2001, in *Astronomical Society of the Pacific Conference Series*, Vol. 247, *Spectroscopic Challenges of Photoionized Plasmas*, ed. G. Ferland & D. W. Savin, 227–+
- Dopita, M. A. & Evans, I. N. 1986, *ApJ*, 307, 431
- Dopita, M. A., Fischera, J., Sutherland, R. S., et al. 2006, *ApJ*, 647, 244
- Dopita, M. A., Kewley, L. J., Heisler, C. A., & Sutherland, R. S. 2000, *ApJ*, 542, 224
- Dopita, M. A., Sutherland, R. S., Nicholls, D. C., Kewley, L. J., & Vogt, F. P. A. 2013, *ApJS*, 208, 10
- Dors, Jr., O. L. & Copetti, M. V. F. 2003, *A&A*, 404, 969
- Dottori, H. A. 1987, *Rev. Mexicana Astron. Astrofis.*, 14, 463
- Dottori, H. A. & Copetti, M. V. F. 1989, *Rev. Mexicana Astron. Astrofis.*, 18, 115
- Draine, B. T. 2011, *ApJ*, 732, 100
- Evans, I. N. & Dopita, M. A. 1985, *ApJS*, 58, 125
- Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., et al. 2011, *A&A*, 532, A95
- Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, *Rev. Mexicana Astron. Astrofis.*, 49, 137
- Fitzpatrick, E. L. 1999, *PASP*, 111, 63
- Flores-Fajardo, N., Morisset, C., & Binette, L. 2009, *Revista Mexicana de Astronomía y Astrofísica*, 45, 261
- Flores-Fajardo, N., Morisset, C., Stasińska, G., & Binette, L. 2011, *MNRAS*, 415, 2182
- García-Benito, R., Pérez, E., Díaz, Á. I., Maíz Apellániz, J., & Cerviño, M. 2011, *AJ*, 141, 126
- García-Benito, R., Zibetti, S., Sánchez, S. F., et al. 2015, *A&A*, 576, A135
- Gomes, J. M., Papaderos, P., Kehrig, C., et al. 2016, *A&A*, 588, A68
- González Delgado, R. M. & Perez, E. 1997, *ApJS*, 108, 199
- González Delgado, R. M., Pérez, E., Cid Fernandes, R., et al. 2014, *A&A*, 562, A47
- Hodge, P. W. & Kennicutt, Jr., R. C. 1983, *ApJ*, 267, 563
- Husemann, B., Jahnke, K., Sánchez, S. F., et al. 2013, *A&A*, 549, A87
- Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, *MNRAS*, 346, 1055
- Kehrig, C., Monreal-Ibero, A., Papaderos, P., et al. 2012, *A&A*, 540, A11
- Kehrig, C., Vilchez, J. M., Sánchez, S. F., et al. 2008, *A&A*, 477, 813
- Kelz, A., Verheijen, M. A. W., Roth, M. M., et al. 2006, *PASP*, 118, 129
- Kewley, L. J. & Dopita, M. A. 2002, *ApJS*, 142, 35
- Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, *ApJ*, 556, 121
- Knapen, J. H. 1998, *MNRAS*, 297, 255
- Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, *ApJS*, 123, 3
- Liang, Y. C., Yin, S. Y., Hammer, F., et al. 2006, *ApJ*, 652, 257
- Lopez, L. A., Krumholz, M. R., Bolatto, A. D., Prochaska, J. X., & Ramirez-Ruiz, E. 2011, *ApJ*, 731, 91
- López-Sánchez, Á. R., Dopita, M. A., Kewley, L. J., et al. 2012, *MNRAS*, 426, 2630
- López-Sánchez, Á. R. & Esteban, C. 2009, *A&A*, 508, 615
- Marino, R. A., Gil de Paz, A., Sánchez, S. F., et al. 2016, *A&A*, 585, A47
- Marino, R. A., Rosales-Ortega, F. F., Sánchez, S. F., et al. 2013, *A&A*, 559, A114
- Martins, L. P., González Delgado, R. M., Leitherer, C., Cerviño, M., & Hauschildt, P. 2005, *MNRAS*, 358, 49
- Mast, D., Rosales-Ortega, F. F., Sánchez, S. F., et al. 2014, *A&A*, 561, A129
- Matteucci, F. 1986, *MNRAS*, 221, 911
- Mollá, M., García-Vargas, M. L., & Bressan, A. 2009, *MNRAS*, 398, 451
- Mollá, M., Vilchez, J. M., Gavilán, M., & Díaz, A. I. 2006, *MNRAS*, 372, 1069
- Morisset, C. 2013, *pyCloudy: Tools to manage astronomical Cloudy photoionization code*
- Morisset, C. 2014, in *Asymmetrical Planetary Nebulae VI conference, Proceedings of the conference held 4-8 November, 2013*. Edited by C. Morisset, G. Delgado-Inglada and S. Torres-Peimbert.
- Morisset, C., Delgado-Inglada, G., & Flores-Fajardo, N. 2015, *Rev. Mexicana Astron. Astrofis.*, 51, 103
- Morisset, C. & Georgiev, L. 2009, *A&A*, 507, 1517
- Nicholls, D. C., Dopita, M. A., & Sutherland, R. S. 2012, *ApJ*, 752, 148
- Oey, M. S., Parker, J. S., Mikles, V. J., & Zhang, X. 2003, *AJ*, 126, 2317
- Osterbrock, D. E. 1989, *Astrophysics of gaseous nebulae and active galactic nuclei* (University Science Books)
- Papaderos, P., Gomes, J. M., Vilchez, J. M., et al. 2013, *A&A*, 555, L1
- Peimbert, M. 1967, *ApJ*, 150, 825
- Pérez, E., Cid Fernandes, R., González Delgado, R. M., et al. 2013, *ApJ*, 764, L1
- Pérez-Montero, E. 2014, *MNRAS*, 441, 2663
- Pérez-Montero, E., García-Benito, R., Hägele, G. F., & Díaz, Á. I. 2010, *MNRAS*, 404, 2037
- Pilyugin, L. S., Vilchez, J. M., Mattsson, L., & Thuan, T. X. 2012, *MNRAS*, 421, 1624
- Rémy-Ruyer, A., Madden, S. C., Galliano, F., et al. 2014, *A&A*, 563, A31
- Rosales-Ortega, F. F., Kennicutt, R. C., Sánchez, S. F., et al. 2010, *MNRAS*, 405, 735
- Rosales-Ortega, F. F., Sánchez, S. F., Iglesias-Páramo, J., et al. 2012, *ApJ*, 756, L31
- Roth, M. M., Kelz, A., Fechner, T., et al. 2005, *PASP*, 117, 620
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- Sánchez, S. F., García-Lorenzo, B., Jahnke, K., et al. 2006b, *New A Rev.*, 49, 501
- Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012a, *A&A*, 538, A8
- Sánchez, S. F., Pérez, E., Rosales-Ortega, F. F., et al. 2015, *A&A*, 574, A47
- Sánchez, S. F., Pérez, E., Sánchez-Blázquez, P., et al. 2016, *Rev. Mexicana Astron. Astrofis.*, 52, 21
- Sánchez, S. F., Rosales-Ortega, F. F., Iglesias-Páramo, J., et al. 2014a, *A&A*, 563, A49
- Sánchez, S. F., Rosales-Ortega, F. F., Iglesias-Páramo, J., et al. 2014b, *A&A*, 563, A49
- Sánchez, S. F., Rosales-Ortega, F. F., Jungwiert, B., et al. 2013, *A&A*, 554, A58
- Sánchez, S. F., Rosales-Ortega, F. F., Kennicutt, R. C., et al. 2011, *MNRAS*, 410, 313
- Sánchez, S. F., Rosales-Ortega, F. F., Marino, R. A., et al. 2012b, *A&A*, 546, A2
- Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, *MNRAS*, 371, 703
- Sánchez-Menguiano, L., Sánchez, S. F., Pérez, I., et al. 2016, *A&A*, 587, A70
- Sarzi, M., Shields, J. C., Schawinski, K., et al. 2010, *MNRAS*, 402, 2187
- Singh, R., van de Ven, G., Jahnke, K., et al. 2013, *A&A*, 558, A43
- Stasińska, G., Cid Fernandes, R., Mateus, A., Sodré, L., & Asari, N. V. 2006, *MNRAS*, 371, 972
- Stasińska, G., Izotov, Y., Morisset, C., & Guseva, N. 2015, *A&A*, 576, A83
- Stasińska, G., Schaerer, D., & Leitherer, C. 2001, *A&A*, 370, 1
- Stasińska, G., Vale Asari, N., Cid Fernandes, R., et al. 2008, *MNRAS*, 391, L29
- Thurston, T. G., Edmunds, M. G., & Henry, R. B. C. 1996, *MNRAS*, 283, 990
- Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., et al. 2010, *MNRAS*, 404, 1639
- Veilleux, S. & Osterbrock, D. E. 1987, *ApJS*, 63, 295
- Vila Costas, M. B. & Edmunds, M. G. 1993, *MNRAS*, 265, 199
- Vilchez, J. M. & Pagel, B. E. J. 1988, *MNRAS*, 231, 257
- Walcher, C. J., Wisotzki, L., Bekeraitė, S., et al. 2014, *A&A*, 569, A1
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579



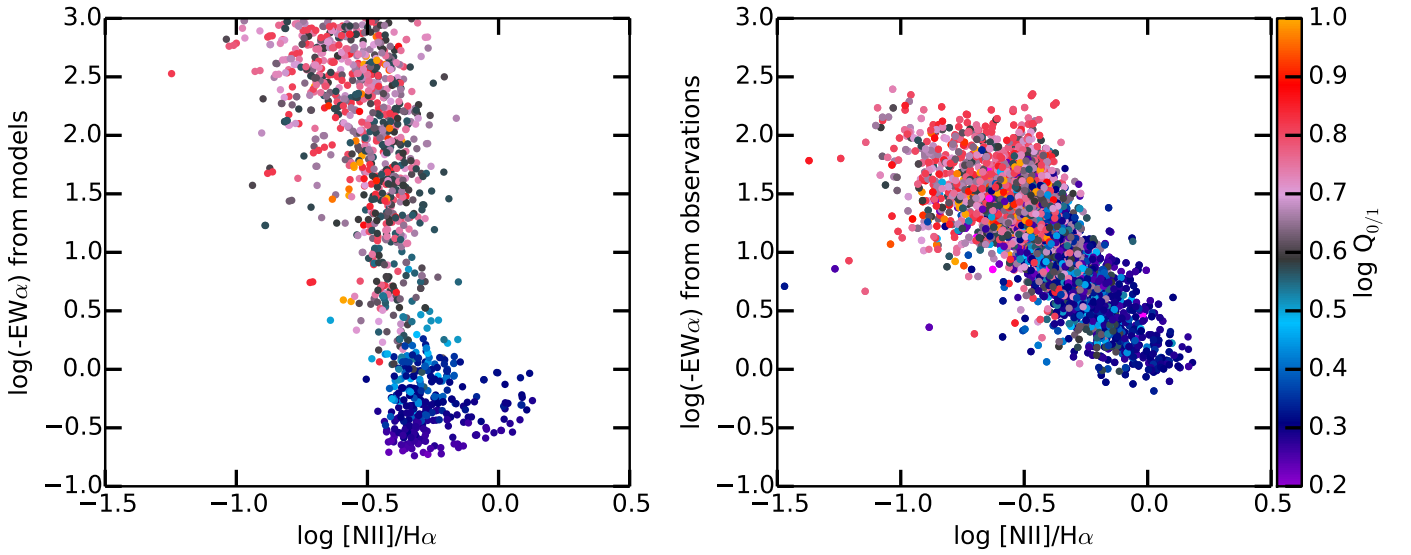
**Fig. 5.** Electron temperature diagnostic line ratios as a function of O/H (left panels) and  $([\text{OII}]\lambda 5007 + [\text{OII}]\lambda 3727)/\text{H}\beta$  (right panels) for our models (colored circles). From top to bottom, the diagnostics are  $[\text{OIII}]\lambda 4363/\lambda 5007$  and  $[\text{NII}]\lambda 5755/\lambda 6584$ . Black diamonds represent the  $T_e$ -based sample of HII regions used by Marino et al. (2013). The dashed-color lines correspond to the grid of models computed by (Dopita et al. 2013). The color bar is following  $Q_{0/1}$ , the softness of the ionizing radiation.



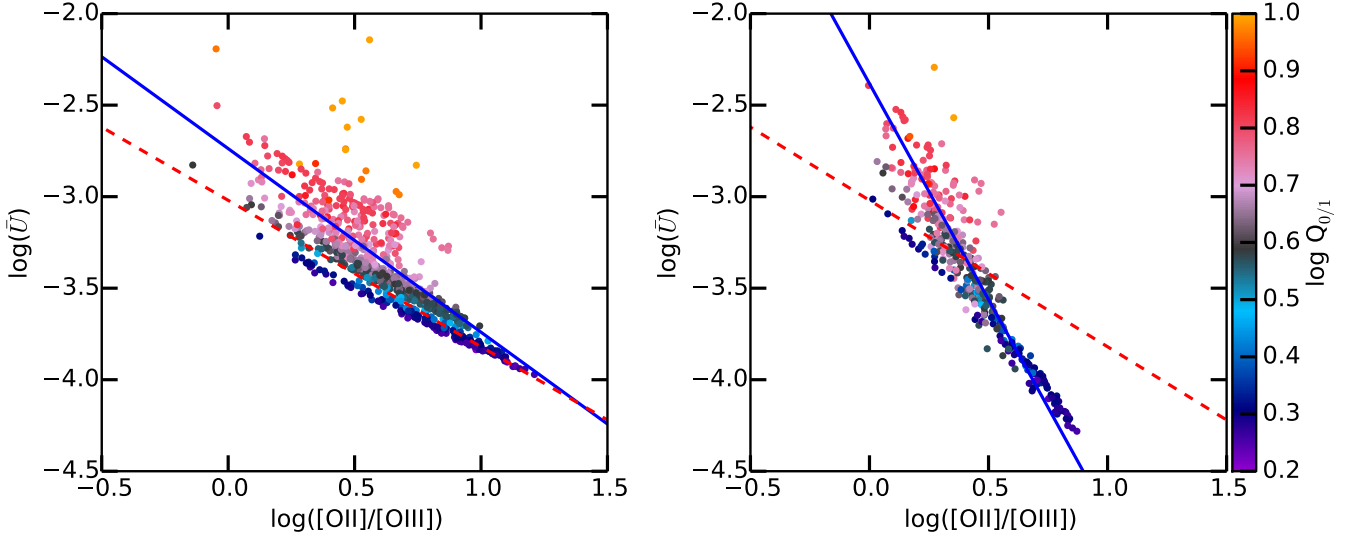
**Fig. 7.** Classical BPT diagrams of the model results. For each panel, the color code is changed according to the description on the right of the corresponding color bar. Upper panels show distributions of the models, with colors related to the chemical abundances: O/H on the left and N/O on the right. The middle panels show the distribution of the  $H\alpha$  equivalent width, determined from the observation on the left and from the models on the right panels respectively. The lower panels show the distribution of the mean ionization parameter  $\log(\bar{U})$  on the left and the proportion of OB stars in the ionizing SED on the right panels respectively. The solid blue line is from Kauffmann et al. (2003), the green dashed line is from Kewley et al. (2001) and the dotted black line is from Stasińska et al. (2006).



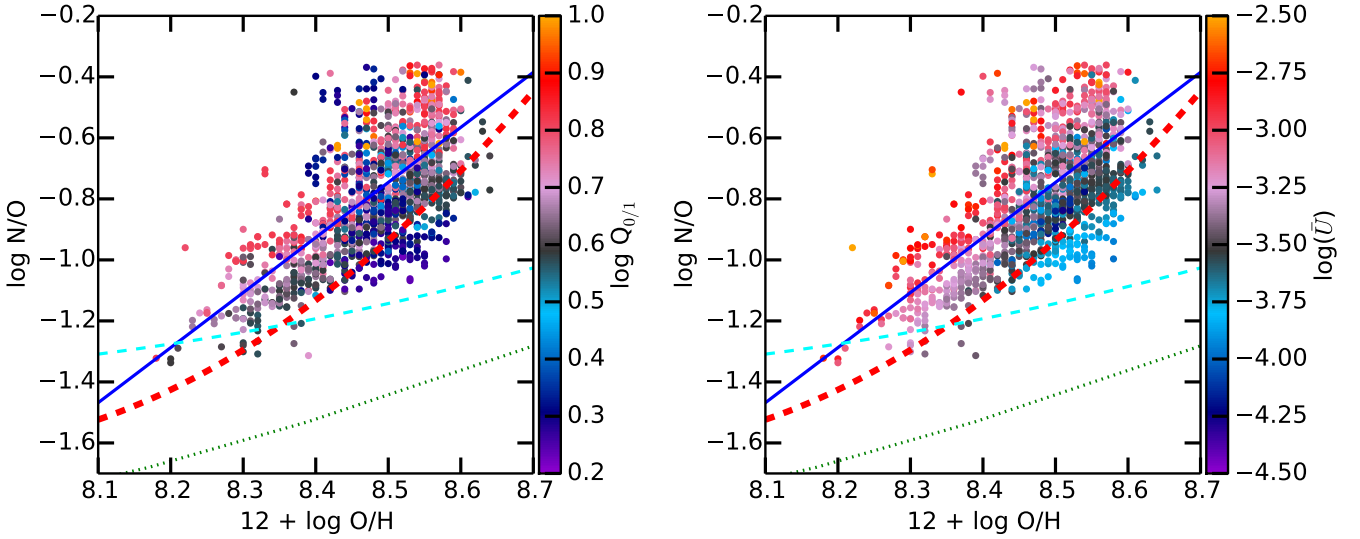
**Fig. 8.** The BPT diagrams inspired by Baldwin et al. (1981). The color code represents the hardness of the radiation  $Q_{0/1}$ . The solid blue line is from Kauffmann et al. (2003), the green dashed line is from Kewley et al. (2001) and the dotted black line is from Stasińska et al. (2006).



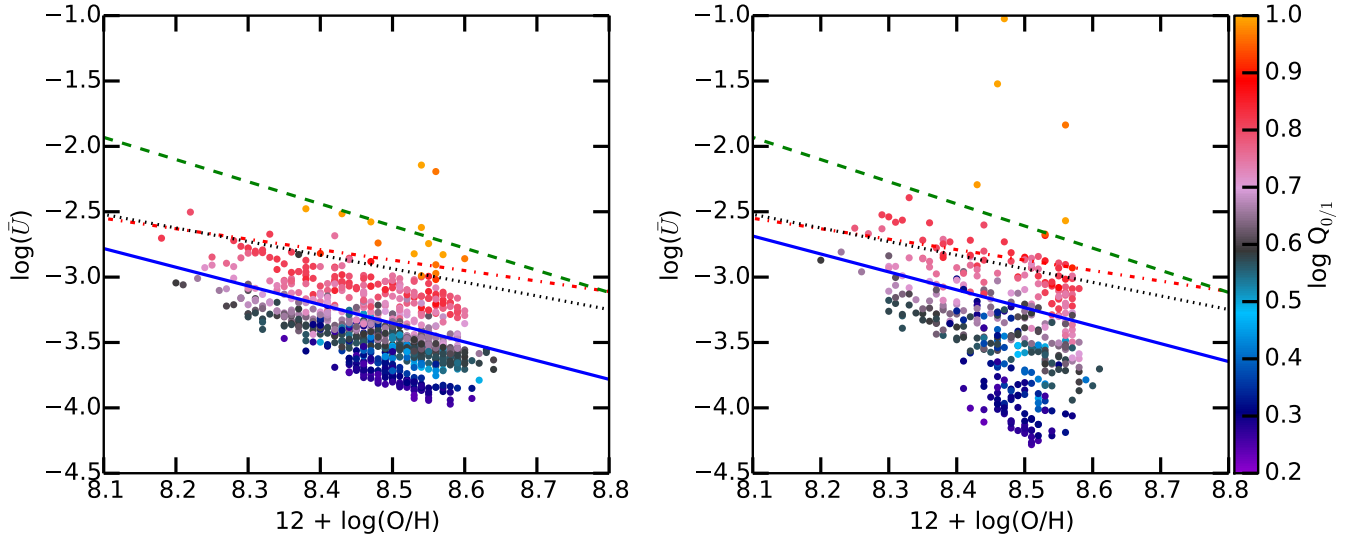
**Fig. 9.** WHAN diagrams: values of  $EW_{\alpha}$  from the models (left panel) and from the observations (right panel) as a function of  $[NII]/H\alpha$ . The color code represents the proportion of OB stars.



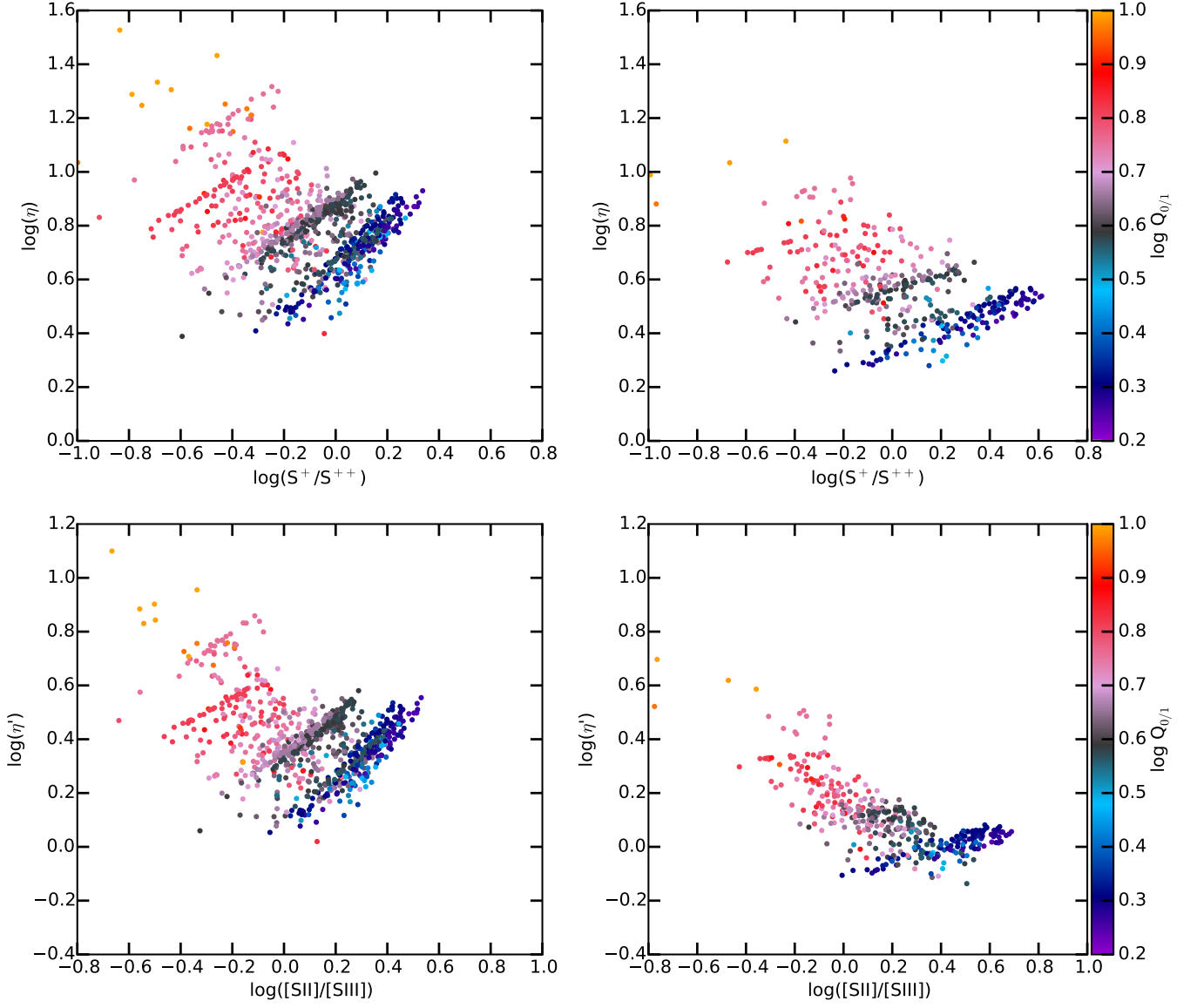
**Fig. 10.**  $\log(\tilde{U})$  vs.  $[\text{O II}]/[\text{O III}]$  for the models. The colors coding the softness of the radiation emitted by the stellar population. Left panel: thin shell models ( $fr = 3.0$ ), right panels: filled sphere models ( $fr = 0.03$ ). The dashed red line is the fit by Díaz et al. (2000). The blue lines correspond to our fit, taking only star forming regions into account ( $Q_{0/1} > 0.55$ ).



**Fig. 12.**  $\text{N/O}$  vs.  $\text{O/H}$ . Left panel: the color is coding the hardness of the ionizing radiation  $Q_{0/1}$ . Right panel: the color is coding the value of  $\log(\tilde{U})$ . The blue line is the fit of the  $Q_{0/1} > 0.55$  regions (gray/red/orange points), see text for the corresponding values. The red dashed line corresponds to the fit by Pilyugin et al. (2012), the cyan dashed line to Vila Costas & Edmunds (1993), and the green dotted line to the fit by Dopita et al. (2013).



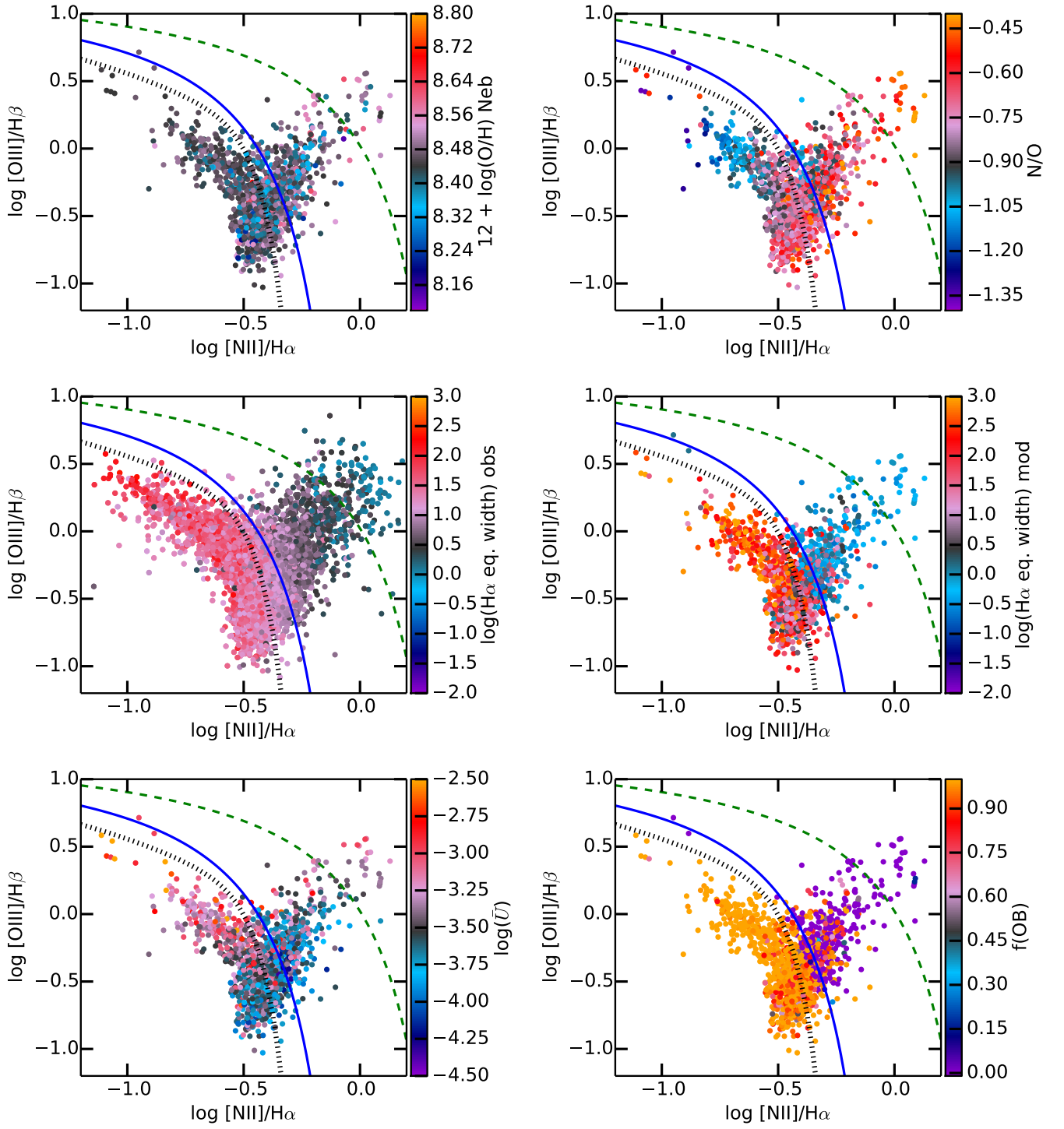
**Fig. 13.**  $\log(\bar{U})$  vs.  $O/H$ . Left panel: thin shell models ( $fr = 3.0$ ), right panel: filled sphere models ( $fr = 0.03$ ). The solid red dot-dashed line corresponds to the relation from Dopita & Evans (1986), the green dashed line to Dopita et al. (2006) and the red dotted line to a fit to the Fig. 3 from Pérez-Montero (2014). The color code is following  $Q_{0/1}$ . Our fit is shown with the blue line, it is obtained considering only the  $Q_{0/1} > 0.55$  regions, see text for the corresponding values.



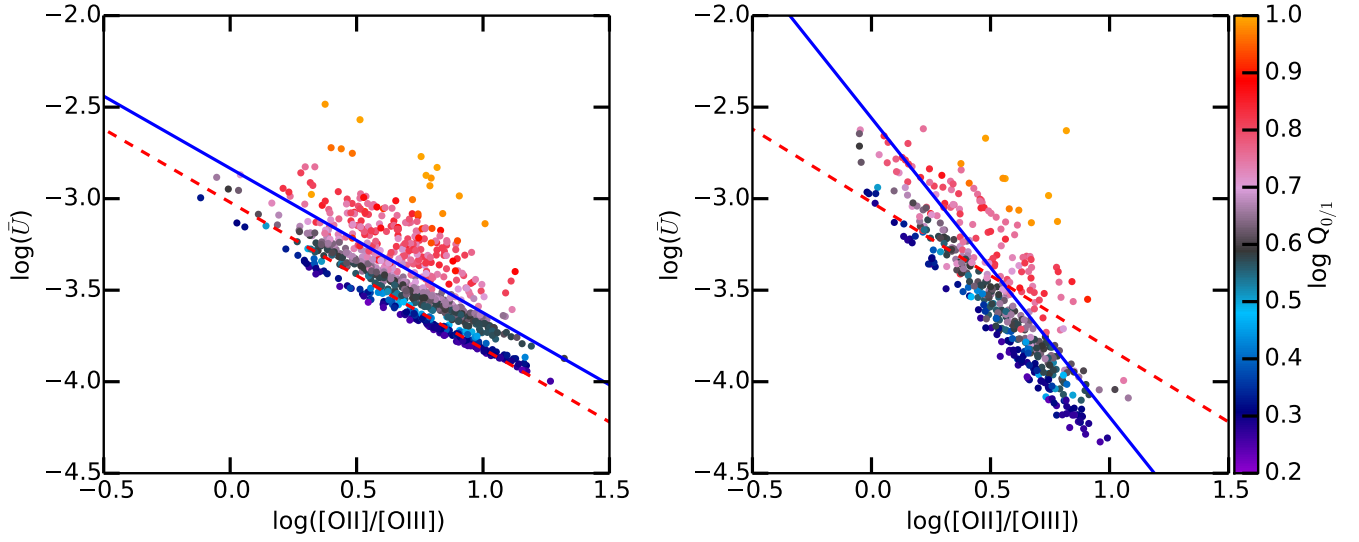
**Fig. 14.** Relation between  $\eta = (O^+/O^{++}) / (S^+/S^{++})$  and  $S^+/S^{++}$  (upper panels) and between  $\eta' = ([OII]\lambda 3727+/[OIII]\lambda 5007+) / ([SII]\lambda 6720+/[SIII]\lambda 9067+)$  and  $[SII]\lambda 6720+/[SIII]\lambda 9067+$  (lower panels). Left panel: thin shell models ( $fr = 3.0$ ), right panel: filled sphere models ( $fr = 0.03$ ). The color code follows the  $Q_{0/1}$  ratio (see text).

## **Appendix A: Results obtained from the "Stel" models**

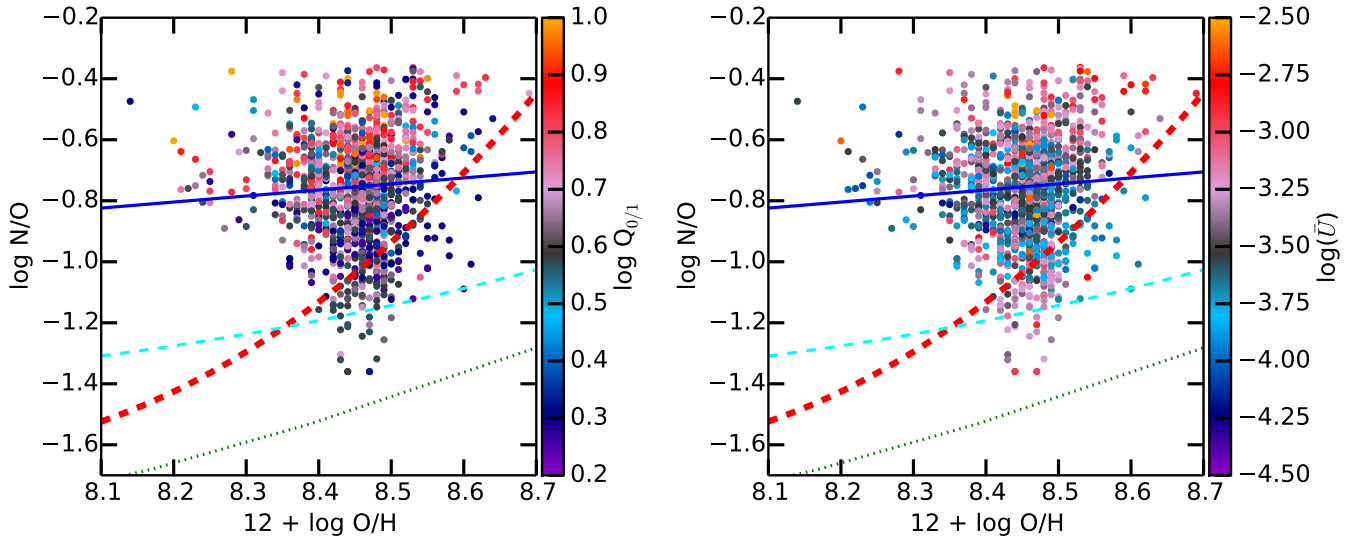
In this on-line appendix, we present the same figures as already shown in the paper, but obtained with the "Stel" abundance determination, instead of the "Neb" one (See Secs. 3.1 and 4.8). We only show the figures significantly different from the "Neb" case.



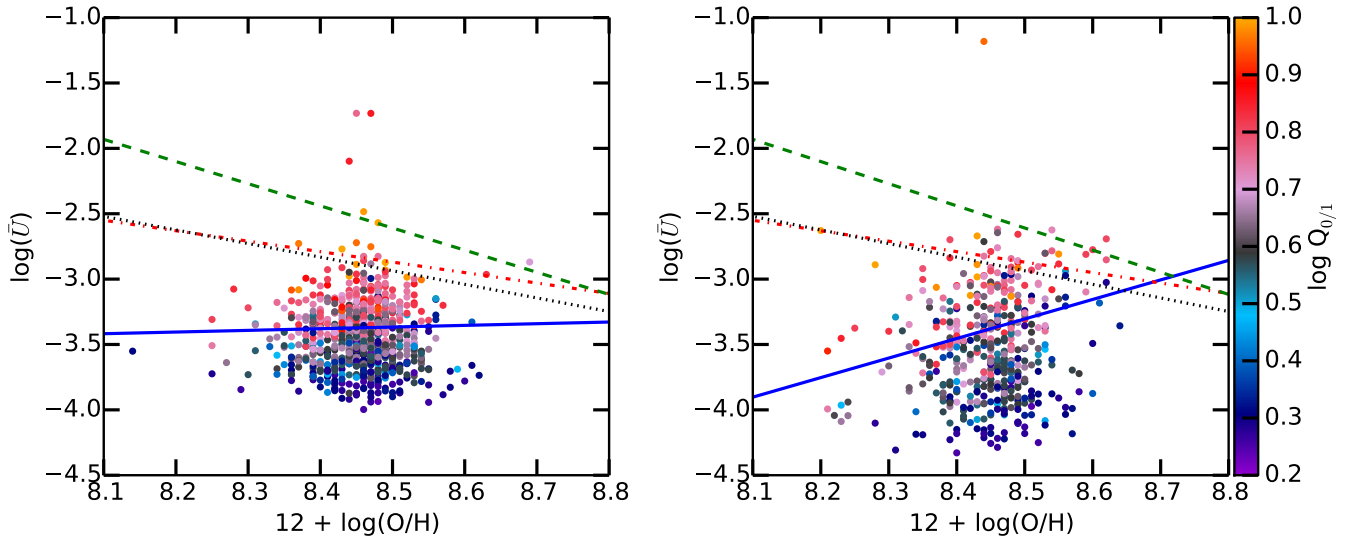
**Fig. A.1.** Same as Fig. 7 but using the "Stel" models.



**Fig. A.2.** Same as Fig. 10 but using the "Stel" models.



**Fig. A.3.** Same as Fig. 12 but using the "Stel" models.



**Fig. A.4.** Same as Fig. 13 but using the "Stel" models.