bond: Bayesian Oxygen and Nitrogen abundance Determinations in giant H ii regions using strong and semistrong lines

N. Vale Asari, G. Stasińska, C. Morisset, R. Cid Fernandes

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **460,** 1739–1757 (2016)
Advance Access publication 2016 April 23

doi:10.1093/mnras/stw971

BOND: Bayesian Oxygen and Nitrogen abundance Determinations in giant H II regions using strong and semistrong lines

N. Vale Asari,[1⋆] G. Stasińska,[2] C. Morisset[3] and R. Cid Fernandes[1]

[1]*Departamento de Física – CFM, Universidade Federal de Santa Catarina, C.P. 476, 88040-900, Florianópolis, SC, Brazil*
[2]*LUTH, Observatoire de Paris, PSL Research University, CNRS, Université Paris Diderot, Sorbonne Paris Cité, 5 place Jules Janssen, F-92195 Meudon, France*
[3]*Instituto de Astronomía, Universidad Nacional Autónoma de México, Apdo. Postal 70264, México D.F., 04510 México*

Accepted 2016 April 20. Received 2016 April 20; in original form 2016 January 26

ABSTRACT

We present the Bayesian oxygen and nitrogen abundance determinations (BOND) method. BOND is a Bayesian code (available at: http://bond.ufsc.br) to simultaneously derive oxygen and nitrogen abundances in giant H II regions. It compares observed emission lines to a grid of photoionization models without assuming any relation between O/H and N/O. Our grid spans a wide range in O/H, N/O and ionization parameter U, and covers different starburst ages and nebular geometries. Varying starburst ages accounts for variations in the ionizing radiation field hardness, which arise due to the ageing of H II regions or the stochastic sampling of the initial mass function. All previous approaches assume a strict relation between the ionizing field and metallicity. The other novelty is extracting information on the nebular physics from semistrong emission lines. While strong lines ratios alone ([O III]/Hβ, [O II]/Hβ and [N II]/Hβ) lead to multiple O/H solutions, the simultaneous use of [Ar III]/[Ne III] allows one to decide whether an H II region is of high or low metallicity. Adding He I/Hβ pins down the hardness of the radiation field. We apply our method to H II regions and blue compact dwarf galaxies, and find that the resulting N/O versus O/H relation is as scattered as the one obtained from the temperature-based method. As in previous strong-line methods calibrated on photoionization models, the BOND O/H values are generally higher than temperature-based ones, which might indicate the presence of temperature fluctuations or kappa distributions in real nebulae, or a too soft ionizing radiation field in the models.

Key words: methods: data analysis – ISM: abundances – H II regions – galaxies: abundances.

1 INTRODUCTION

Thanks to their conspicuous emission lines, giant H II regions are used as indicators of the chemical composition of the interstellar medium in galaxies, and have permitted important advances in our understanding of the chemical evolution of galaxies (see e.g. Esteban et al. 2004 and references therein). While the so-called temperature-based abundance determinations, which require the measurement of weak auroral lines to measure the electron temperatures, are commonly considered the most reliable, strong-line methods have become increasingly popular since the pioneering studies by Pagel et al. (1979) and Alloin et al. (1979) because they can also be applied for distant galaxies.

Strong-line methods involve some restrictions, though: they assume that giant H II regions form a one (or two) parameter(s) family and they need to be calibrated. Calibration can be done via a subsample of objects with temperature-based abundances or using a grid of photoionization models. The first method is potentially biased, since calibration samples are likely to have different properties than the samples one wishes to study. In particular, they are biased against objects having intrinsically weak auroral lines. Calibrations based on grids of photoionization models do not have this problem (assuming the models cover all the combinations of important parameters that are encountered in nature and are realistic enough).

A large number of calibrations have been proposed. If we label the methods by their abundance indicators adopting the notation O3N2 for [O III]/[N II],[1] O23 for ([O II] + [O III]$_S$)/Hβ, N2Ha

⋆ E-mail: natalia@astro.ufsc.br

[1] In the entire paper, the notations [N II], [O II], [O III], [Ne III], [S II], [S III], [Ar III], [Ar IV], and He I stand for [N II]λ6584, [O II]λ3726 + λ3729, [O III]λ5007, [Ne III]λ3869, [S II]λ6716 + λ6731, [S III]λ9069, [Ar III]λ7135, [Ar IV]λ4711 + λ4740, and He Iλ5876, respectively, [O III]$_S$ for [O III]λ4959 + λ5007, and [N II]$_S$ for [N II]λ6548 + λ6584 (where the subscript S is used to denote a sum of emission lines).

for [N II]/Hα, etc., the most popular ones are: O23 (Pagel et al. 1979), O3N2 (Alloin et al. 1979), O23–O3O2 (McGaugh 1991), N2Ha (Storchi-Bergmann, Calzetti & Kinney 1994), S23 (Vilchez & Esteban 1996). All these methods with their numerous calibrations (here we have quoted the pioneering ones) give very different outcomes (see e.g. Kewley & Ellison 2008 for a comparison of the results).

Apart from this problem of leading to discrepant results, strong-line methods face two important issues. One is that factors other than just the metallicity and the ionization parameter influence the strength of the strong lines emitted by giant H II regions. This potentially leads to incorrect inferences when applying the methods to compare different samples (see Stasińska 2010). The second issue is that there are two regimes where the intensities of the strong oxygen lines used for abundance determinations have the same value with respect to Hβ: the low-metallicity and the high-metallicity regimes. To resolve this bimodality, one uses the intensity of the strong nitrogen line since, in the astrophysical context, the N/O ratio is a function of metallicity. This procedure, a priori reasonable, is however not totally secure since the frontier between the two regimes is fuzzy. For example, McGaugh (1994) adopts $\log$ [N II]/[O II] > -1 while Kewley & Ellison (2008) adopt > -1.2 for the high-metallicity regime. The difference between these two values may appear insignificant but can lead to somewhat different conclusions on metallicity trends within and among galaxies. In addition, this procedure does not allow one to pin down objects with pathological N/O ratios – which may be particularly interesting for unveiling peculiarities in the star-forming histories of galaxies (Mollá & Gavilán 2010). Apart from peculiar N/O ratios, the N/O versus O/H relation may be different at high redshifts, which would systematically bias metallicity measurements based on the local N/O versus O/H relation for high-redshift galaxies. Of course, methods using directly N/O as an indicator of the oxygen abundance (e.g. the [N II]/Hα method proposed by Storchi-Bergmann et al. 1994 or the [N II]/[O II] method proposed by Kewley & Dopita 2002) present the same drawback.

In this paper, we show that using the semistrong lines [Ne III], [Ar III] and He Iλ5876, in conjunction with the classical strong lines, it is possible to estimate with reasonable accuracy both the oxygen and the nitrogen abundance in giant H II regions without any prior assumption on the N/O ratio and without the implicit priors of classical strong-line methods regarding the ionizing radiation field. The intensities of these semistrong lines have been listed in many papers reporting on deep spectroscopy of giant H II regions, so those lines must be present in the spectra for which only the most common strong lines have had their intensities published.

We construct a finely meshed grid of photoionization models varying not only O/H and the ionization parameter as has been done before (McGaugh 1991; Kewley & Dopita 2002; Blanc et al. 2015), but also N/O (like Pérez-Montero 2014). We use this grid to estimate the abundances of O and N in giant H II regions by means of standard Bayesian inference methods (like Blanc et al. 2015). Unlike Blanc et al. (2015), we do not assume an N/O versus O/H relation, and explicitly explore variations in N/O. The main novelties of our approach are that we consider variations in the hardness of the ionizing field, and that we extract information from semistrong emission lines in addition to the commonly used strong lines.

The paper is organized as follows. In Section 2, we present the spectroscopic data we have collected from the literature to develop and test our method. In Section 3, we show two extreme versions of the N/O versus O/H diagram obtained from these data using a temperature-based method and using the strong-line method of Pilyugin, Vílchez & Thuan (2010). In Section 4, we present our grid of photoionization models, built using the code CLOUDY (Ferland et al. 2013). In Section 5, we present our method, and, in Section 6, we show our results for the N/O versus O/H diagram. In Section 7, we provide a summary and elaborate on future directions of work. Three appendices complement this paper. The first one presents a realistic sample of fake sources constructed by selecting model nebulae from our grid. The second one describes a few tests using these fake sources. The third one compares the abundances derived by Bayesian oxygen and nitrogen abundance determinations (BOND) with those obtained by other published methods on the same set of observational data.

2 THE OBSERVATIONAL DATA BASE

2.1 Giant H II regions in spiral galaxies

Data on giant H II regions in spiral galaxies were gathered from recent medium-resolution high-quality observational studies, mostly with very large telescopes (Keck, VLT) whose high S/N allowed the measurement of auroral lines in at least part of the observed samples. Apart from the large data base from van Zee et al. (1998), all the other works involve Bresolin as first or second author, which guarantees a certain homogeneity in the treatment of the data. The following sources were used (the letters correspond the reference labels in Table 1):

(a) Bresolin et al. (2005);
(b) Bresolin, Garnett & Kennicutt (2004);
(c) Kennicutt, Bresolin & Garnett (2003);
(d) van Zee et al. (1998);
(g) Bresolin et al. (2009b);
(i) Bresolin (2007);
(j) Bresolin et al. (2010);
(k) Li, Bresolin & Kennicutt (2013);
(l) Zurita & Bresolin (2012);
(m) Bresolin, Kennicutt & Ryan-Weber (2012);
(n) Goddard et al. (2011);
(p) Bresolin et al. (2009a).

All these sources give the line fluxes corrected for extinction and the associated uncertainties. When the intensity of Hα was not given, it was assumed to be equal to 2.86 times that of Hβ. Many sources list only the intensities of lines that are used in the classical strong-lines methods. i.e. [O III], [O II] and [N II]. For papers giving tables with the fluxes of all the lines seen in the spectra, we roughly estimated the upper limits for the intensities of the lines that were not detected, e.g. [Ar III] or [O III]λ4363. Upper limits are estimated by taking twice the lowest uncertainty in measured lines. This, of course, is a stopgap solution in absence of any direct information from the observers.

2.2 Blue compact galaxies

To increase the number of objects at low metallicities, we use the sample of blue compact galaxies with high-quality spectra which were used by Izotov, Thuan & Stasińska (2007, reference labelled z in Table 1) to derive the pre-galactic helium abundance. The abundances of O and N have been recomputed in exactly the same way as for the giant H II regions in spiral galaxies.

Table 1. A sample of the data table available for download at http://bond.ufsc.br. Line fluxes, uncertainties and upper limits (F, eF and limF) are in units of Hβ, and references are labelled in column r as in Sections 2.1 and 2.2.

id	r	name	F3727	eF3727	F3869	eF3869	...	F7135	eF7135	limF4363	limF5755
001	a	NGC 1232 02	3.9100	0.3300	–	–	...	0.0610	0.0090	0.0140	0.0140
002	a	NGC 1232 03	3.3300	0.2500	0.2530	0.0350	...	0.0720	0.0130	0.0260	0.0260
003	a	NGC 1232 04	2.0700	0.1700	0.2910	0.0240	...	0.0820	0.0090	0.0038	0.0038
004	a	NGC 1232 05	2.5300	0.1600	0.0480	0.0060	...	0.0650	0.0050	0.0022	–
...	...	...	...	...	...	...	...	...	...	...	...
705	z	HSS1809+6612	2.5600	0.1013	0.3150	0.0177	...	–	–	–	–
706	z	Mrk259	2.3900	0.0615	0.3230	0.0099	...	0.0820	0.0033	–	–
707	z	SBS1428	1.8800	0.0459	0.1870	0.0059	...	0.0600	0.0022	–	–
708	z	S1657+575	2.1300	0.0834	0.2430	0.0157	...	0.0740	0.0083	–	–

2.3 Subsamples

For the needs of this study, we merge the two samples described above, and then constitute several subsamples.

(i) Sample A is constructed from the entire merged sample by selecting all the objects with [O II], [O III] and [N II] available. It contains 708 objects. The line intensities and associated uncertainties are reported in Table 1, available for download from http://bond.ufsc.br.

(ii) Sample T is the subsample of sample A with available temperature measurements from [O III]λ4363/5007 and/or [N II]λ5755/6584; it contains 261 objects.

(iii) Sample B is the subsample fulfilling the minimum requirements for the use of the BOND method, i.e. with available fluxes for [O II], [O III] and [N II] and for the semistrong lines [Ne III], [Ar III] and He I. It contains 156 objects.

3 THE OBSERVED N/O VERSUS O/H DIAGRAM

3.1 Computation of temperature-based O and N abundances

The abundances of O and N were recomputed in a homogeneous way using 5-level atoms for O^+, $O^{++}$and N^+. The sources for the collision strengths and transition probabilities are the following.[2] For O II: Kisielius et al. (2009) and Zeippen (1982); for O III: Aggarwal & Keenan (1999), Galavis, Mendoza & Zeippen (1997), and Storey & Zeippen (2000); for N II: Tayal (2011) and Galavis et al. (1997). The electron densities were computed from the [S II]λ6731/6717 ratio (when available) using the atomic data from Tayal & Zatsarinny (2010) and Mendoza & Zeippen (1983). When the [S II]λ6731/6717 ratio was not available, it was assumed that the electron density is equal to 100 cm^{-3}.

The ionic abundances were computed with a two-zone electron temperature scheme. The temperature derived from [O III]λ4363/5007 was used for O^{++} and the temperature derived from [N II]λ5755/6584 was used for O^+ and N^+. When one of the two line ratios was missing, the following classical relation from Garnett (1992) was used:

$$T_{[\mathrm{N\,II}]} = T_{[\mathrm{O\,II}]} = 0.70 \times T_{[\mathrm{O\,III}]} + 3000\mathrm{K}. \quad (1)$$

[2] Note that the atomic data used to compute the O and N abundances are the same as the ones entering in the version of CLOUDY used to compute our grid of models.

The O and N abundances were obtained using the classical assumption that oxygen is only in the form of O^+ and O^{++} in the H II region and that $N/O = N^+/O^+$.

The uncertainties were estimated by a Monte Carlo procedure using the uncertainties on the observed line fluxes as described in detail in Stasińska et al. (2013). Uncertainties due to possible deviations from equation (1) as well from the $N/O = N^+/O^+$ equation were not taken into account in the Monte Carlo procedure.

3.2 Comparison of N/O versus O/H diagrams

Fig. 1 compares the N/O versus O/H diagram using two different methods for the abundance determination. In the left-hand panel, we used the temperature-based method as described above. In the right-hand panel, we considered the ON method from Pilyugin et al. (2010), which is based on the [O II]/Hβ, [O III]$_S$/Hβ, [N II]$_S$/Hβ, and [S II]/Hβ emission line ratios. Both panels show exactly the same 232 objects: from our sample T, we select only those objects where [S II] has been measured, which is necessary for the ON method. This is a strong-line method calibrated on a sample of H II regions with available temperature-based abundances. One can see that the two panels of Fig. 1 look very different, with the left one showing significantly more dispersion than the right one. Note that panel (a) shows some points which are really far away from the main trend, while their associated uncertainties are small.

Which of the two diagrams is closer to reality? Temperature-based methods are often considered the most reliable. However, this assertion must be tempered by several considerations. Temperature-based methods assume relations between some parameters (like $T_{[\mathrm{N\,II}]}$ and $T_{[\mathrm{O\,III}]}$ or N/O and N^+/O^+), whereas, in fact, some dispersion is expected (see Appendix A). They are also strongly dependent on errors in the intensities of the weak lines that serve to determine the temperatures. At the highest metallicities, important temperature gradients inside the H II regions may bias the abundance results, as shown by Stasińska (2005). Shocks may contribute to the intensities of the auroral lines, and falsify the results on abundances. Finally, if the electron velocities in the ionized gas are not Maxwellian but rather follow a κ distribution as suggested by Nicholls, Dopita & Sutherland (2012), classical temperature-based methods will result in underestimated abundances with respect to hydrogen. Because of all these reasons, it is not unreasonable to think that part of the scatter observed in Fig. 1 may be artificial. On the other hand, the very tight relation between N/O and O/H seen in Fig. 1 right-hand panel may be unreal, since the formulae developed by Pilyugin et al. (2010) tend to strongly tighten any preexisting correlation (see Appendix C, Fig. C5).

In what follows, we develop a new method to derive oxygen and nitrogen abundances in giant H II regions which is much less

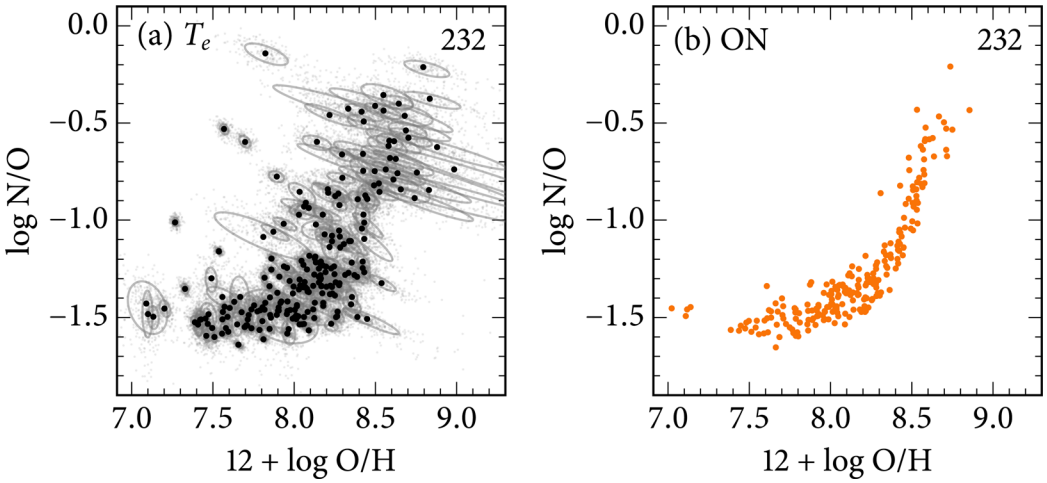

Figure 1. The observed N/O versus O/H diagram for the objects of our sample T where [S II] has been measured. (a) Diagram derived from the temperature-based method. The large dots represent the nominal solution, the small background dots the 200 Monte Carlo realizations for each object, and the ellipses the covariances. (b) Diagram calculated by the ON method by Pilyugin et al. (2010) for the same objects.

affected by the intensities of auroral lines than the temperature-based methods and, unlike previous strong-line methods, does not involve any assumption on the N/O ratio.

4 THE MODEL GRID

4.1 Definition of the grid

Because we want to avoid any biases in our method, we need to construct a grid in which we vary all the determinant parameters. If we view a giant H II region as a nebula powered by an instantaneous burst of star formation, the main parameters for our problem are the oxygen and nitrogen abundances, the mean ionization parameter and the age of the burst. The density distribution may also have a certain importance.

Using CLOUDY 13.03 (Ferland et al. 2013), we constructed a grid of models defined as follows.

(i) The oxygen abundance on the scale of 12 + log O/H goes from 6.6 to 9.4 in steps of 0.2 dex (15 values). The abundances of all the heavy elements except nitrogen and carbon are taken proportional to that of oxygen, as in Stasińska et al. (2015). The helium abundance varies with the oxygen abundances as in Stasińska et al. (2015).

(ii) The N/O ratio takes the logarithmic values −2, −1.5, −1, −0.5, 0. The abundance of carbon is linked to that of nitrogen by $\log \mathrm{C/H} = 0.48 + \log \mathrm{N/H}$.

(iii) Dust is included in the models, being related to the oxygen abundance in exactly the same way as in Stasińska et al. (2015), following the works of Rémy-Ruyer et al. (2014) and Draine (2011).

(iv) The mean *input* ionization parameter, defined by equation (4) of Stasińska et al. (2015), takes the logarithmic values −1, −1.5, −2, −2.5, −3, −3.5, −4. Note that the real mean ionization parameter of the *computed* model is somewhat different from the input value, since it depends on the electron temperature and on the partial absorption of the ionizing photons by dust (see fig. B2 of Stasińska et al. 2015). In the remaining of the paper, we denote this mean *input* ionization parameter as U.

(v) The starburst age takes the values of 1, 2, 3, 4, 5 and 6 Myr. The spectral energy distribution (SED) of the ionizing radiation is obtained from the population synthesis code PopStar (Mollá, García-Vargas & Bressan 2009) for a Chabrier (2003) stellar initial mass function and for the appropriate metallicity, obtained by interpolation.

(vi) In order to assess the effect that geometry might have, we consider two density distributions. One is a filled sphere of density $n = 100\,\mathrm{cm}^{-3}$, the other is a thin spherical shell of same density. Roughly, the first scenario can correspond to a relatively young H II region and the second to an evolved one. Clearly, these choices are very simplistic and the role of the density distribution should be explored further. The mathematical definitions of the shell and filled sphere are detailed in section 4.1 of Stasińska et al. (2015).

All the models are computed from the inner boundary until the ratio of ionized hydrogen to total hydrogen density falls below 0.02.

4.2 Some characteristics of the grid

Fig. 2 shows the entire grid of computed models in two planes frequently used to contemplate observations or models of H II regions. One is the [O III]/Hβ versus [N II]/Hα plane often used for excitation diagnostics (commonly called the BPT diagram after Baldwin, Phillips & Terlevich 1981), and the other is [O III]λ5007 + 4959/[O II]λ3727 versus ([O III]λ5007 + 4959 + [O II]λ3727)/Hβ introduced by McGaugh (1991) to derive the oxygen abundance and to which we will refer as the McG diagram. The models for filled spheres are displayed in the left-hand column, while the models for shells are displayed in the right-hand column. Each row of panels corresponds to a given starburst age, increasing downwards. Fig. 3 is a zoom in the BPT and McG planes for the 1 Myr filled sphere subgrid, which serves to better illustrate the coloured lines and points drawn in Fig. 2. Models joined with full coloured lines have the same O/H and same N/O (the colour is defined by the value of O/H and runs from purple to red following the rainbow colours as O/H increases), while models joined with thin grey lines have the same input value of the mean ionization parameter. In all the panels, the model curves are superimposed on the observational points.

The first thing we can notice is that the entire grid appears to cover most of the observational points in these two planes, which is what we were looking for, i.e. the fact that the observational points are difficult to see in the figure is a feature, and not a flaw. However, in the McG diagram, a small proportion of objects appears slightly to the right of the grid, at any of the ages considered,

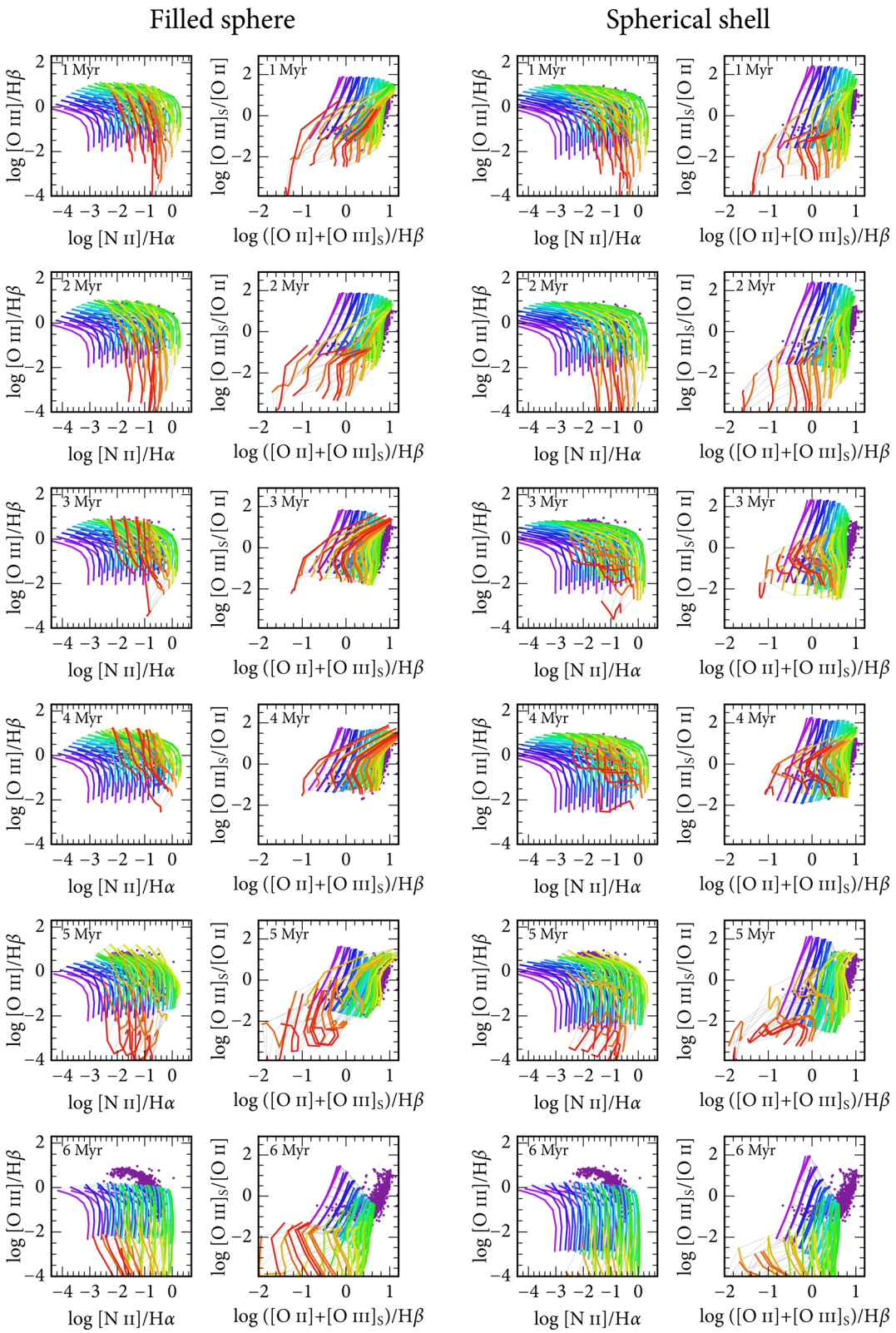

Figure 2. The model grid on [O III]/Hβ versus [N II]/Hα (BPT), and [O III]$_S$/[O II] versus ([O III]$_S$ +[O II])/Hβ diagrams for various ages (1–6 Myr) and geometries (filled sphere or empty spherical shell, indicated at the top of the figure). The aim of the postage-stamp size panels is twofold. First, one can see at a glance how the different parameters change the shape of the model subgrids. Secondly, the model grid is overplotted on the observational data, which allows one to judge how well the grid covers the observational points. The fact that most points are hidden behind the grid is thus a fact to be celebrated. The colour-coding of grid lines is detailed in Fig. 3, which zooms into two panels. A colour version of this figure is available in the electronic edition.

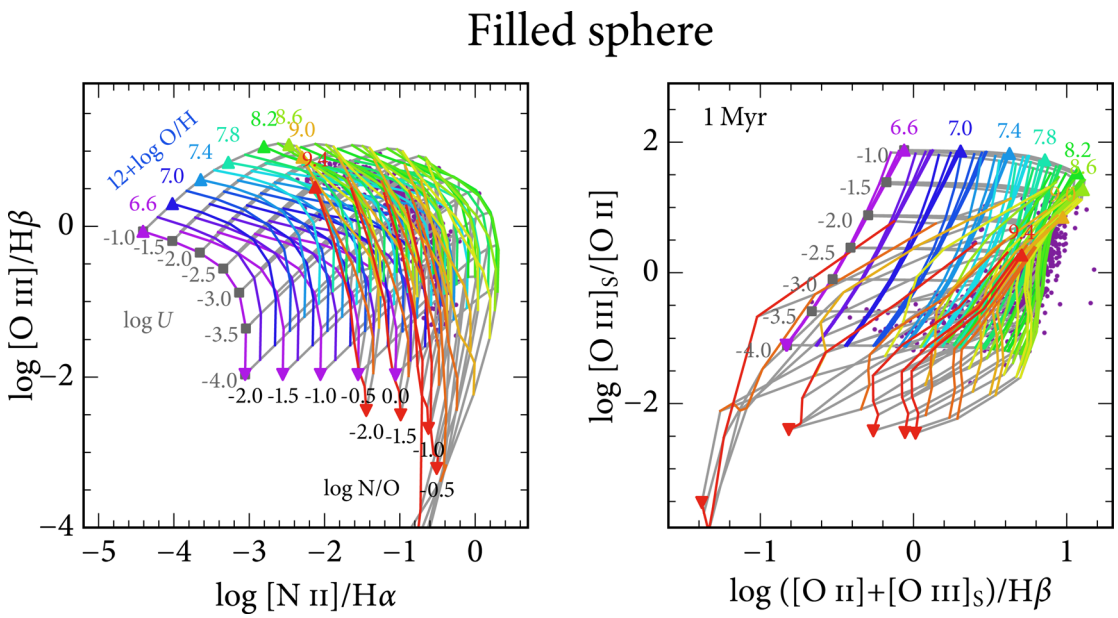

Figure 3. Zoom into the top-left panels of Fig. 2. The lines mark the region covered by a subset of our grid of photoionization models (those with starburst ages of 1 Myr and a filled sphere nebular geometry). The small dots mark our observational data, which are mostly hidden behind the grid. Triangles pointing up, down, and squares mark a few values of O/H, N/O and U, respectively. The colours of the grid lines change with the value of O/H, going from red for the largest values to purple for the lowest one, following the order of the rainbow. Lines of the same colour have the same O/H but different N/O. Models with the same ionization parameter U are linked by a grey line. Readers interested in exploring our grid in detail can zoom in the pdf version of the paper or can download an electronic file with our grid of models from the BOND website (http://bond.ufsc.br) or the 3MdB data base (https://sites.google.com/site/mexicanmillionmodels/).

meaning that, in this region, the electron temperature computed by the models is probably lower than in real H II regions. We have explored several possibilities to reduce the problem by playing with dust and abundance ratios and density, but we did not succeed. Anyway, the discrepancy is much smaller than in the studies of McGaugh (1994) and Dopita et al. (2013, especially for their grid with a κ distribution of electrons). We think that the discrepancy we find is due to our models still not reproducing exactly real objects rather than to observational errors. Nevertheless, the magnitude of the problem is sufficiently small to warrant our further use of the grid for abundance determinations. Indeed, we find that excluding objects that fall off grid does not change our results.

The two different density distributions (filled sphere and empty shell) produce only slight apparent differences in the grid but, as we will see in Appendix B, this is sufficient to affect the O/H and N/O ratios by up to 0.05 dex. More realistic density distributions, such as a core-halo density distribution or a constant pressure distribution, may have a larger impact.

The variations in N/O obviously have an impact on the BPT diagram but they also affect the McG diagram at high metallicities, since they affect the cooling rates. In other words, if the N/O ratio is abnormally high, this would bias the O/H derived from strong-line methods not accounting for a possible scatter in N/O.

We also note that, at the highest metallicities, the curves of equal chemical composition become ill-behaved. This is because, at such metallicities, most of the cooling occurs through the infrared fine-structure lines whose intensity is not very sensitive to temperature. Therefore, a small change in the physical conditions of the gas may alter the electron temperature considerably, which, in turn, strongly affects the intensity of the [O III]λ5007 line. This means that, for $12 + \log \mathrm{O/H}$ greater than, say, 9.2, the real error in the abundances derived from optical lines is probably larger than can be estimated from our grid.

Fig. 2 shows that the starburst age modifies the shape of the model grid, especially in the McG diagram, so that assuming the same age for all the H II regions will produce significant errors in the oxygen abundance determination. What actually changes from one age to another is the 'hardness' of the ionizing radiation field, i.e. its capacity of heating the surrounding medium by photoionization.

The hardness can be viewed as the ratio $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$, where $Q(\mathrm{He}^0)$ is the number of photons above 24.6 eV and $Q(\mathrm{H}^0)$ is the number of photons above 13.6 eV. Fig. 4 shows the variations of $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$ as a function of time for the six PopStar metallicities. Generally, the ionizing radiation field softens as metallicity increases. However, during the Wolf–Rayet phase, the radiation field hardens and this effect is higher at high metallicity. As a result, the radiation field is the hardest at the highest metallicities and at ages around 3–5 Myr. This implies that for these ages the $([\mathrm{O\,III}]_\mathrm{S} + [\mathrm{O\,II}])/\mathrm{H}\beta$ ratio can reach quite high values at high metallicities.

In reality, the process of star formation may not be instantaneous, as in the PopStar models, but extend over a certain time. In practice, what is important for the line intensities is not so much the age of the ionizing stellar population or the regime of star formation, but rather the hardness of the resulting ionizing radiation field. Our models with different ages should thus be viewed as models for SEDs of different hardness.

Fig. 5 shows the variations of $([\mathrm{O\,III}]_\mathrm{S} + [\mathrm{O\,II}])/\mathrm{H}\beta$ with respect to O/H for models with $\log U = -2$, the curves representing the sequences of models being coloured according to the value of $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$. It is clear that for $12 + \log \mathrm{O/H}$ larger than, say, 7.5, the primary metallicity indicator, $([\mathrm{O\,III}]_\mathrm{S} + [\mathrm{O\,II}])/\mathrm{H}\beta$, is strongly dependent on $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$, reinforcing the relevance of considering this effect in the modelling.

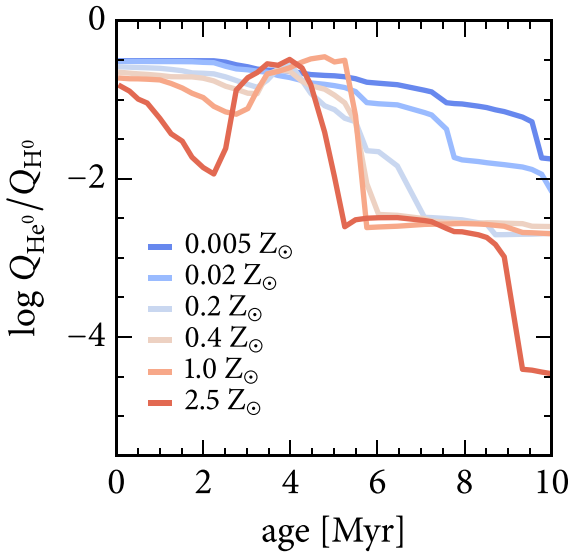

Figure 4. The values of $Q(He^0)/Q(H^0)$ as a function of time for the PopStar models at different metallicities. The hardness of the radiation field in an H II region tends to decrease with age, but the Wolf–Rayet phases wreaks havoc on this simple view.

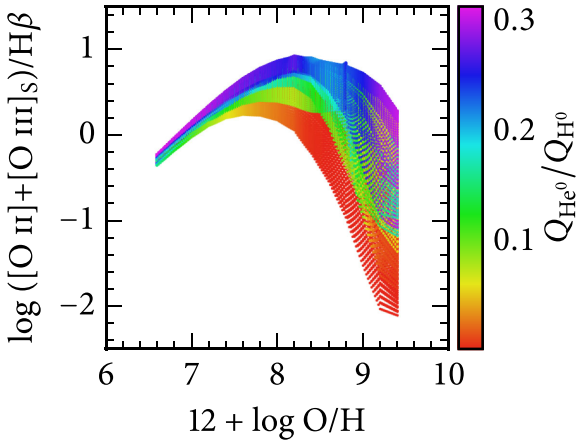

Figure 5. $([O\,III]_S + [O\,II])/H\beta$ versus O/H for $\log U = -2$ and all the values of N/O, ages and density structures in our grid coloured by $Q(He^0)/Q(H^0)$. In order to see the trends better, we have created a very fine interpolated grid for this figure (sampled 0.02, 0.05, 0.05 dex in O/H, N/O and U, respectively). Note that, for high metallicities, the $([O\,III]_S + [O\,II])/H\beta$ ratio does not pin down the value of the O/H abundance. A colour version of this figure is available in the electronic edition.

5 DESCRIPTION OF THE BOND METHOD

Our grid of models spans a wide range of physical parameters: N/O, O/H, U, the hardness of the radiation field, and the density profile of the nebula. Although we are only interested in inferring the nitrogen and oxygen abundances, we need to constrain the other nuisance parameters. This section explains our choices of observational constraints and the formalism for our BOND method.

5.1 Observational constraints

5.1.1 Uncovering N/O and O/H

Our first set of observational constraints are the extinction-corrected line ratios $\log [N\,II]/H\beta$, $\log [O\,II]/H\beta$ and $\log [O\,III]/H\beta$. The formal assumption is that those logarithmic line ratios are Gaussianly distributed and independent. Using line luminosities instead of line ratios is meaningless for our models, since the models are not defined by luminosities but by ionization parameters.

The physical reasoning behind using this set of line ratios is that $[N\,II]/[O\,II]$, $([O\,III] + [O\,II])/H\beta$, and $[O\,III]/[O\,II]$ are proxies for N/O, O/H and U, respectively. Since we are using only strong lines, there is no constraint on the electron temperature. That way, because $([O\,III] + [O\,II])/H\beta$ versus O/H is bivalued, this first set of constraints finds bivalued solutions for O/H.

We believe that using carefully chosen emission lines is better than using all information available for this problem. We therefore choose not to include [S II] for two reasons. [S II] comes from the outskirts of the nebula, which do not coincide with the region where the other strong lines are produced. Any density structure will change [S II] in relation to the other lines. Secondly, the S/O ratio in H II regions could be subject to variations due to different production sites of S and O (e.g. it has recently be proposed by Delgado-Inglada et al. 2015 that intermediate-mass stars could contribute to the global oxygen budget in galaxies) and to different dust-depletion schemes.

5.1.2 Eliminating the bimodality

As explained in the introduction, using nitrogen to break the $([O\,III] + [O\,II])/H\beta$ degeneracy with O/H is not satisfactory since the relation between the N/O ratio and O/H is likely dispersed. It is better to use a physical argument that does not depend on astrophysical conditions, like one based on the electron temperature, which will be low in the high-abundance regime and high in the low-abundance one.

We need a line ratio that is easy to observe, and that depends strongly on the electron temperature and weakly on ionization conditions and abundance ratio. The [Ar III]/[Ne III] ratio fulfils these requirements. The [Ar III] and [Ne III] lines have different excitation thresholds (1.7 and 3.2 eV, respectively), so different dependencies on the electron temperature. Argon and neon are two primary elements; in addition, they are both rare gases not suspected of dust depletion, so their abundance ratio is expected to be constant. The [Ar III] and [Ne III] lines do not arise exactly from the same zone, the ionization potential of Ne^{++} being higher than that of Ar^{++}, but we can use the [O III]/[O II] ratio to figure out what the ionization is.

Fig. 6 shows [Ar III]/[Ne III] as a function of [O III]/[O II] for our subsample of models corresponding to an age of 2 Myr and a filled sphere nebula. The colours indicate the abundance regime: blue corresponds to the low-metallicity branch in the $([O\,III] + [O\,II])/H\beta$ versus O/H diagram, and red to the high-metallicity branch. We can see how using the [Ar III]/[Ne III] ratio in conjunction with the [O III]/[O II] one separates the two branches. In practice, we use $\log [Ar\,III]/H\beta$ and $\log [Ne\,III]/H\beta$ separately as our constraints because they are closer to an ideal Gaussian distribution than [Ar III]/[Ne III]. Also, because the grid was not designed to closely explain [Ar III]/[Ne III], we add an extra noise e in quadrature to [Ar III]/Hβ and [Ne III]/Hβ and marginalize it away, as explained in Section 5.3.1.

The [Ar III]/[Ne III] as a function of [O III]/[O II] for all starburst ages and scenarios is a bit fuzzier, showing some superposition of high- and low-metallicity grid points. For those cases, it is useful to include an extra constraint to exclude the grid points on the wrong O/H branch. We find that using the [O III]λ4363 and [N II]λ5755 upper limits as additional criteria for the electron temperature

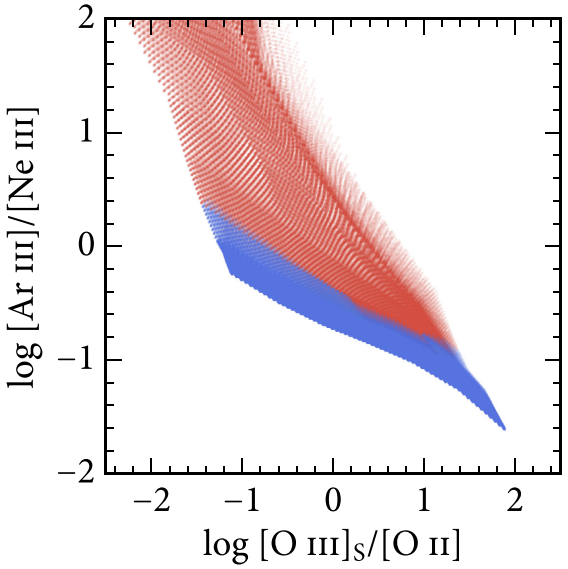

Figure 6. [Ar III]λ7135/[Ne III]λ3869 versus [O III]$_S$/[O II] for age 2 Myr and filled sphere of our grid (the same used in Fig. 5). For each value of N/O and U, we tag a model as being in the high or low-O/H branch from the ([O III] + [O II])/Hβ versus O/H diagram. We colour the low and high branches in blue and red, respectively. Note how the two branches are almost cleanly separated in this diagram, with very little overlap of blue and red points. A colour version of this figure is available in the electronic edition.

improves our solutions. The reasoning for using upper limits is that, if the [O III]λ4363 (or [N II]λ5755) line is not observed, and the observational upper limit for its intensity is below the expected value in the low-metallicity regime, this implies that we are in the high-metallicity regime. We include those upper limits in our inference as discussed in Section 5.3.2. We check that this does not force our solutions to match the observed [O III]λ4363 or [N II]λ5755.

5.1.3 *Characterising the radiation field*

Line ratios such as ([O III]/[O II])/([S III]/[S II]) (Vilchez & Pagel 1988) or ([Ar IV]/[Ar III])/([O III]/[O II]) (Stasińska et al. 2015) can be used to estimate the mean effective temperature of the radiation field. However, [Ar IV] is often too weak to be measured in giant H II regions. ([O III]/[O II])/([S III]/[S II]) does depend on U (see Stasińska et al. 2015), but, because it depends on [S II], it suffers from the problem that it is affected by the density distribution in real H II regions mentioned above (Section 5.1.1).

Another potential indicator is He Iλ5876/Hβ which, as long as helium is not fully ionized in the H II region, is dependent on the SED of the ionizing stars. The He I/Hβ ratio, however, depends on the metallicity, since the dependence of the He I and Hβ line emissivities with electron temperature is not the same. Fig. 7 shows He I/Hβ as a function of O/H in our grid of models, coloured according to $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$ with the same colour scale as in Fig. 5. We see that, with the information on O/H given by the other lines used in BOND, He I/Hβ allows one to estimate $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$ up to a value of ~ 0.2. This is not entirely satisfactory, because Fig. 5 shows that at the highest metallicities, the value of ([O III]$_S$ +[O II])/Hβ depends on $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$ also at values larger than 0.2. In this paper, we restrict ourselves to using He I/Hβ to characterize the hardness of the radiation field. In future works, when large data bases of giant H II regions with fully described deep spectra become available, it will be possible to add information on [Ar IV]/[Ar III]. We also allow

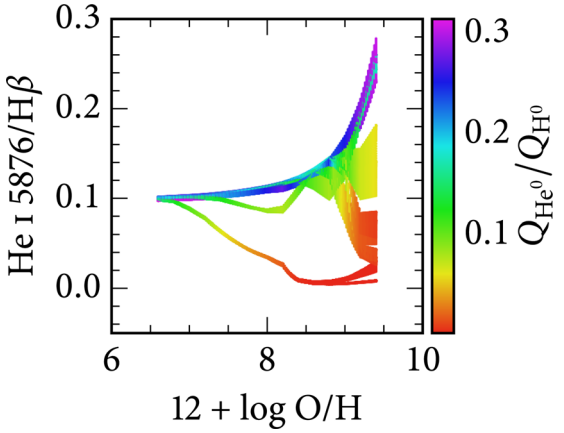

Figure 7. O/H versus He Iλ5876/Hβ coloured by $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$ for $\log U = -2$ and all the values of N/O, ages and density structures in our grid (the same used in Fig. 5). For a given O/H, He Iλ5876/Hβ is a good proxy of $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$. For the highest $Q(\mathrm{He}^0)/Q(\mathrm{H}^0)$, however, He Iλ5876/H$\beta$ does not change much (blue and purple points all fall on the same region). The sparsity of the stellar ages in our grid shows up as the large uncovered parts of this diagram. A colour version of this figure is available in the electronic edition.

an extra noise e and integrate it out for $\log$ He I/Hβ as described in Section 5.3.1.

5.2 The probabilistic formalism

We aim to find the oxygen and nitrogen abundances of an H II region by comparing its observed lines O to our grid of models. A given observed line is characterized by its intensity and uncertainty (o_j, σ_j), and an H II region by its $j = 1 \ldots J$ emission lines:

$$O \equiv \{O_j\} = \{o_j, \sigma_j\}, \tag{2}$$

where the curly braces define a set of values spanning the rightmost index (i.e. j for the equation above). Each model M in our grid is defined by its $i = 1 \ldots I$ model parameters m_i, and generates a set of computed line intensities $\{c_j\}$:

$$M \equiv \{\{m_i\}; \{c_j\}\}. \tag{3}$$

Our grid of models spans a wide range of values not only for our two parameters of interest (oxygen and nitrogen abundances), but also for the ionization parameter, starburst ages and nebular geometry. While the latter play an important role in the photoionization modelling of an H II region, we do not wish to infer them. From a pragmatic point of view, a Bayesian formalism offers a framework to marginalize away those nuisance parameters by simply integrating them out. This comes with the cost of writing down the posterior probability so that the dimensions of our probability function still make physical sense after those integrals are performed (see e.g. Hogg 2012). The posterior probability density function (PDF) for a model M, given the observed data O and any other relevant background information B is

$$p(M|O, B) = \mathcal{N}\, p(M|B)\; p(O|M, B), \tag{4}$$

where PDFs are written as p, and $\mathcal{N}$ is a normalization constant so that the posterior integrates to unity over all the parameter space. The PDFs on the right-hand side of the equation are the prior probability of the model parameters, and the likelihood of observing O assuming M and B are true. In what follows, we will discuss our generative model for the likelihood and our choice of the prior.

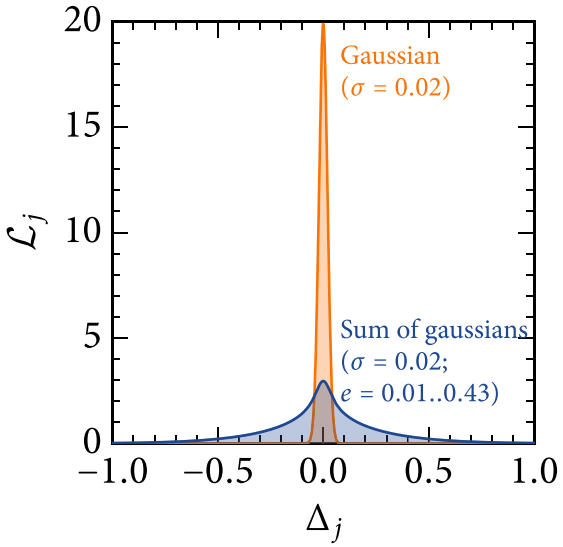

Figure 8. Comparison between a Gaussian, in orange, centred in Δ_j and dispersion $\sigma_j = 0.02$ dex (the typical uncertainty of semistrong-line ratios [Ar III]/Hβ, [Ne III]/Hβ and log He Iλ5876/Hβ in our sample B), and a sum of Gaussians, in blue, with dispersions $(\sigma_j^2 + e^2)^{1/2}$ for e between 0.01 and 0.43 dex (corresponding to 2–100 per cent of the measured line intensity). Both functions are scaled so that their areas are equal to unity.

5.3 Generative model for data points

5.3.1 Partially marginalized likelihoods

We assume Gaussian uncertainties for our constraints, which are the logarithmic line fluxes with respect to Hβ. This is a good approximation given that we consider high S/N observations, so that Hβ is very well determined and the line ratios noise uncertainties must deviate very little from a Gaussian distribution. The likelihood of observing an emission line $O_j = (o_j, \sigma_j)$ given the model M and the background information B, plus an extra source of noise with dispersion e is

$$p(O_j|e, M, B) = \mathcal{N}' \frac{1}{\sqrt{\sigma_j^2 + e^2}} \exp\left[-\frac{(c_j - o_j)^2}{2(\sigma_j^2 + e^2)}\right], \quad (5)$$

where $\mathcal{N}'$ is a normalization constant.

The term e was introduced in order to account for deviations in emission line ratios not contemplated in our models. For instance, we use [Ar III]/Hβ and [Ne III]/Hβ as constraints, but our models were not meant to reproduce the argon and neon abundance dispersions in nature. For constraints involving strong lines ([N II], [O II], [O III]), we simply set $e = 0$. For constraints based on semistrong lines ([Ar III], [Ne III], He I), we consider e in an interval of 0.01–0.43 dex (i.e. for an extra noise from 2 to 100 per cent of the measured line intensity). Since we are not interested in e, we integrate it out to calculate the marginalized likelihood for O_j:

$$\mathcal{L}_j = p(O_j|M, B) = \mathcal{N}'' \int p(e|M, B)\, p(O_j|e, M, B)\, \mathrm{d}e, \quad (6)$$

where $\mathcal{N}''$ is yet another normalization constant. In practice, we calculate $\mathcal{L}_j$ as a sum of Gaussians for logarithmically spaced values of e. Using only 20 Gaussians for this sum guarantees that the numerical integral with log -spaced e is equivalent down to $\sim 10^{-4}$ to a numerical integral with linear spacing of 10^{-6} dex.

Fig. 8 compares a Gaussian with $\sigma_j = 0.02$ dex (a typical value for the semistrong lines in our sample) centred in $\Delta_j = c_j - o_j$ and $e = 0$ to a sum of Gaussians with variances $\sigma_j^2 + e^2$, with e varying from 0.01 to 0.43 dex. The marginalized likelihood for the sum of Gaussians is much broader than the likelihood for a single Gaussian, but note that it eventually drops to zero. This is an effect brought about by the term $(\sigma_j^2 + e^2)^{-1/2}$ in equation (5), which penalizes very large values of e. Therefore, although we allow e to vary, the likelihood is still shaped by the data. What happens in practice is that in our model grid we are probing regions of the emission-line space that are far from the nominal observed measurement.

5.3.2 Treating weak lines

We do not constrain weak line intensities such as [O III]λ4363 or [N II]λ5755, but we use their upper limits as an additional temperature constraint. The upper limit u_j of the weak line j is defined as the 2σ detection limit for a given spectrum. The likelihood for weak lines is a step function that masks out models whose computed emission lines c_j are above the upper limit u_j:

$$p(O_j|M, B) \propto \begin{cases} 1 \text{ if } c_j \leq u_j, \\ 0 \text{ if } c_j > u_j. \end{cases} \quad (7)$$

This constraint is useful when the [Ar III]/[Ne III] ratio alone is unable to distinguish the low- and high-temperature branches. Imposing an upper limit on [O III]λ4363 or [N II]λ5755 flags out the solutions for which the temperatures are too high.

This constraint is peculiar, since it is only available for undetected [O III]λ4363 or [N II]λ5755 lines. Increasing the quality of the observations means that [O III]λ4363 and/or [N II]λ5755 would be detected and we would not be able to use this constraint any longer. In order to avoid the asymmetry of having this constraint applied for some objects and not others, we impose our upper limit criteria even for sources where [O III]λ4363 and/or [N II]λ5755 are detected. This procedure also guarantees that higher S/N data are not penalized. We thus assume that the upper limit for a detected line is its intensity plus its 2σ uncertainty, which again helps clear out solutions with too high a temperature.

5.3.3 Taking all constraints into account

Assuming that the observed line intensities O_j are independent, the likelihood for all observed line intensities O for a given model M is

$$p(O|M, B) = \prod_j p(O_j|M, B). \quad (8)$$

We usually write this down as $\ln p(O|M, B) = \sum_j \ln p(O_j|M, B)$. Expressing the likelihood in natural logarithmic highlights two points. First, we emphasize that our code adds up instead of multiplying values to minimize numerical errors. Secondly, we see that, in the case of a fixed extra noise source (constant e), the likelihood reduces to $\ln p = -0.5 \sum_j (c_j - o_j)^2/(\sigma_j^2 + e^2) \equiv -0.5\chi^2$ apart from a constant of proportionality. We warn, however, that the familiar χ^2 minimization should be looked with suspicion when applied to abundance determinations. First, when strong and weak lines are all fitted at the same time, weak lines are penalized for having larger uncertainties. Albeit formally correct, this lessens the importance of weak lines while they may carry important information, for example, [O III]λ4363 or [N II]λ5755, which pin down the electron temperature. Secondly, the χ^2 can compensate one badly fitted line with one that is extremely well fitted. The correct way to

fit photoionization models would be to fit each line within an appropriate error-bar, which is not ensured by calculating the likelihood by the χ^2. To some extent, our method is immune to this problem because we only use strong and semistrong lines as constraints that strongly shape the likelihood, while weak lines are only used as upper limit measurements.

5.4 Adaptive octree grids

The missing piece to calculate the posterior is the prior PDF. Our background knowledge B (hence our prior) is encoded in the sampling of our model grid. We follow the reasoning by Blanc et al. (2015) and assume a flat logarithmic prior on O/H, N/O and U. This is equivalent to a Jeffrey's prior for a Gaussian distribution with fixed standard deviation. The age of the ionizing source is linearly sampled, and we have two nebular geometries (a filled sphere and a spherical shell); we assume a flat prior for those.

Our original grid is finely meshed in O/H (0.2 dex), but coarse in all other parameters (0.5 dex in U and N/O, 6 starburst ages from 1 to 6 Myr and two nebular geometries). The emission line space is consequently sparsely sampled. When uncertainties in the data are much smaller than the distance between grid models, very few models will be near the observed data. Creating a finer grid can mitigate this problem. Running a sufficient number of photoionization models to fill in the ionizing source ages and nebula density structures adequately would, however, be unnecessarily time-expensive. Interpolating our grid solves the grid sparsity problem quickly and is a good approximation, since the emission line intensities vary smoothly with those parameters once the initial grid is dense enough. We thus interpolate our original grid in $\log \mathrm{O/H}$, $\log U$ and $\log N/O$, but not in starburst age and geometries, which would be dangerous and meaningless, respectively. A finer grid in the latter parameters would require running more photoionization models, a time-consuming task both for generating the grid and running the BOND code. We choose to keep ages and geometries fixed, which pop up as discontinuities and multimodal solutions in our posterior PDFs (take a sneak peek at e.g. the 'islands' of solutions in Fig. 11).

We create a different interpolated grid adapted to each object using an octree sampling algorithm. Octree grids are usually applied to sample a Cartesian 3D space, and are used extensively in video games, computer graphics, hydrodynamics simulations, and Monte Carlo radiative transfer codes (e.g. Saftly et al. 2013 and references therein).

We start off with a grid containing 226, 548 models separated by 0.1 dex in O/H, U and N/O. Starburst ages and nebular geometries are kept fixed. Each grid point represents a cell of volume $\mathrm{d}V = \mathrm{d}(\log \mathrm{O/H})\,\mathrm{d}(\log \mathrm{N/O})\,\mathrm{d}(\log U)$. For each object, we calculate the posterior PDF for all grid points, and the contribution $\mathrm{d}P = p(M|O, B)\,\mathrm{d}V$ of each grid cell to the total probability.

After this first run, we remove grid cells which contribute too little to P to speed up the calculations (the default option is to remove grid points for which $\mathrm{d}P < 10^{-20}$ considering each age and geometry scenario separately). Grid cells where $\mathrm{d}P \geq 10^{-4}$ are subdivided into eight subcells, with each subcell corresponding half the size of the parent cell in O/H, U and N/O. The 10^{-4} threshold was chosen as a compromise between the precision of the posterior PDF and the computing time. Smaller thresholds create very large octree grids, whilst our nominal solutions (i.e. the posterior summaries, see Section 5.5) for $\log \mathrm{O/H}$, $\log \mathrm{N/O}$ change by less than 0.02 dex. We recalculate the posterior with the new octree grid, subdivide the cells where needed, and reiterate until there is no remaining cell with $\mathrm{d}P$ above the threshold. This procedure creates a grid that is finer in the parameter space region where the posterior probability is higher.

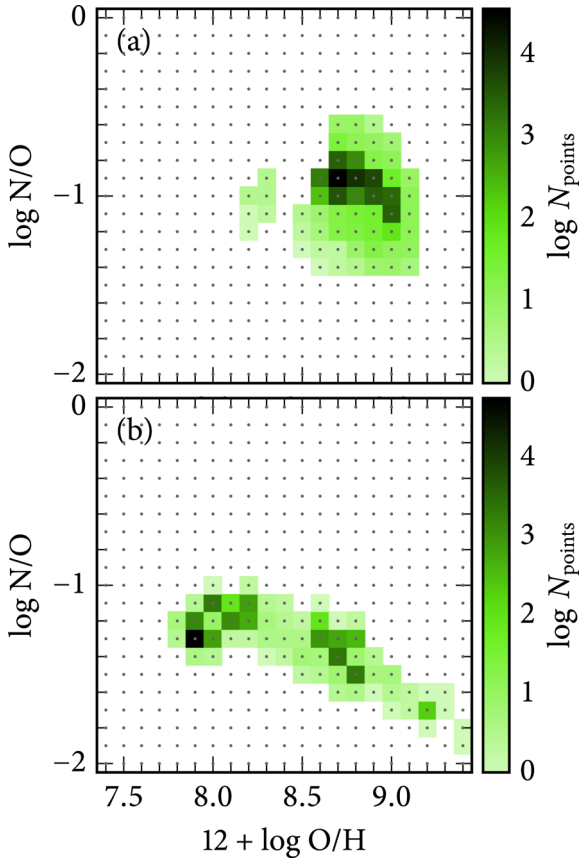

Figure 9. Original grid cell centres (dots) and the final number of octree grid cells (green scale) for two sources in our sample. Panel (a) is for an H II region in NGC 1232 (slit number 5 from Bresolin et al. 2005) and panel (b) for the blue compact dwarf HS0837+4717.

Fig. 9 shows the final octree grid for two objects in sample B, compressed in the N/O versus O/H space. The dots are the centres of the original 0.1 dex-sampled grid cells, and the green scale represents the final number of subcells. Swathes of white space stand for grid cells that have been removed due to contributing too little to the final probability. For our sample B, the median number of cells is ∼65 000 (for a minimum and a maximum of 31 604 and 79 555), and all the grid cells usually go below the $\mathrm{d}P$ threshold after six iterations, which yields subcells $8^{-6} = 1/262,144$ times smaller than the original cell (i.e. which span 0.001 5625 dex in O/H, N/O and U). The average time to run BOND for one source with the octree sampling algorithm in a 1-core 1.7 GHz CPU is 20 s.

5.5 Summarizing the posterior PDF

Having calculated the full posterior $p(M|O, B)$, we can integrate out all parameters we are not interested in and leave out the PDF as a function of only the oxygen and nitrogen abundances. The joint posterior PDF for N/O and O/H (joint PDF for short) is calculated as

$$p(\{z, n\}|O, B) = \sum_i p(\{z_i = z, n_i = n, u_i, t_i, g_i\}|O, B)\,\Delta u_i, \quad (9)$$

where z, n, u, t, g are the model input parameters log O/H, log N/O, log U, age and geometry, respectively, and the subscript i tags each model in the grid. The sum is made over all models with the same values of log O/H and log N/O, and Δu_i takes into account the variable octree cell size in log U. The expectation values for a model input parameter m_i or a computed emission line c_j are given, respectively, by

$$E(\{m_i\}|O, B) = \mathcal{N} \sum_i \{m_i\} \times p(\{z_i = z, n_i = n, u_i, t_i, g_i\}|O, B)\, \Delta u_i, \quad (10)$$

$$E(\{c_j\}|O, B) = \mathcal{N} \sum_i \{c_j\} \times p(\{z_i = z, n_i = n, u_i, t_i, g_i\}|O, B)\, \Delta u_i. \quad (11)$$

The BOND code computes different summaries for the joint PDF: the maximum a posteriori (MAP, i.e. the point of highest probability of the joint PDF), the central point of the credible regions (i.e. the regions on the N/O versus O/H plane of highest probability) that encompass 5, 50, 68 and 95 per cent of the joint PDF, plus its covariance ellipses (scaled so that its area is the same as that of the credible region).

We also calculate the marginalized posterior PDF for several parameters and emission lines. The marginalized posterior PDFs are summarized by their average, median, and mode (i.e. the peak), plus their dispersion, and the extremes of their 50, 68 and 95 per cent equal-tailed (i.e. calculated from the percentiles) and highest density intervals.

For the sake of clarity, in what follows, we show our results in three descriptions only: the joint PDF, the MAP plus the 68 per cent credibility ellipse, and the marginalized median plus the 68 equal-tailed interval. Section 6.2 discusses the differences in the summaries of the posterior PDF.

5.6 A worked example

Fig. 10 shows an example of the influence of each set of constraints on the PDFs for an H II region in NGC 1232. Each row of plots shows the effect of cumulatively adding a set of constraints to the joint PDF of N/O and O/H and to the PDF of the emission line ratios included in the likelihood calculation ([N II]/Hβ, [O II]/Hβ, [O III]/Hβ, [Ar III]/Hβ, [Ne III]/Hβ and He I/Hβ). The joint PDF is represented by a blue-scale map. The grey bands mark $\pm1\sigma$ of the temperature-based values for O/H and N/O, and the purple bands delimit $\pm1\sigma$ of the observed line ratios.

The top row shows the effect of using [N II]/Hβ, [O II]/Hβ, [O III]/Hβ. A range of solutions pops up in the N/O versus O/H plane, spread out in the high- and low-metallicity regimes. The second row adds the information on the upper limit detection of [O III]λ4363 and [N II]λ5755. This selects the high-metallicity solutions. The third row shows how [Ar III]/[Ne III] weighs the four different 'islands' of solutions. Those islands correspond to different starburst ages and geometries and are a consequence of the sparsity of our grid in those parameters. The last row shows which solutions are favoured by He I/Hβ, which will be those with the right hardness for the ionizing source. For the [Ar III]/Hβ, [Ne III]/Hβ and He I/Hβ ratios, we add an extra noise source and then integrate it out, to account for a dispersion in the Ar/Ne and He/H abundance ratios in real nebulae with respect to the values adopted in the model grid.

This example shows that there are three solutions of about the same probability in the N/O versus O/H plane. Fig. 11 is a zoom of the final N/O versus O/H plane, showing the joint and the marginalized PDFs for N/O and O/H. The dots and lines above the marginalized PDFs are the median and the interval between the 16 and 84 percentiles of the marginalized PDF for each age and geometry combination. Those solutions are displaced from the plot axes for clarity. Spherical shells and filled spheres appear as open and filled circles, respectively, and ages are colour-coded in the rainbow order with 1 Myr ages farther and 6 Myr closer to the axes.

To compare to the Monte Carlo temperature-based realizations, we draw its covariance ellipse on the joint PDF panel, and its ±1, 2 and 3σ as grey bands on the marginalized PDFs panels. While the method has not completely eliminated the multimodal nature of the solutions, three out of the four islands of probability on the joint PDF plane are compatible with the temperature-based solution (i.e. inside its covariance ellipse). The multimodal nature of the solutions draws attention to the discreteness of our age grid, as marked by the coloured dots and lines at the edges of the figure. The important message from this plot is that assuming different ionizing fields one finds different values for O/H, and, while He I/Hβ helps pinpointing the right ionizing field, we may still end up with a range of acceptable solutions. This should improve when we have better constraints for the stellar radiation field.

6 RESULTS

6.1 The BOND N/O versus O/H diagram

We apply our BOND method to sample B, which contains 156 objects. Table 2 shows a sample of the summaries for the posterior PDF available to download from http://bond.ufsc.br.

Fig. 12 shows the N/O versus O/H diagram obtained with BOND for those objects. The blue points in both panels are the MAP values for each object. Panel (a) shows the superposition of the joint PDFs for all objects, while panel (b) shows the 68 per cent confidence ellipses. Note that some points fall a long way from their ellipses; this is an evidence of multimodal solutions.

We find that our N/O versus O/H diagram is much more dispersed than the one obtained by the ON method of Pilyugin et al. (2010), indicating that in nature this diagram is not a tight sequence. Points are rather spread out like when using temperature-based methods. Naturally part of the spread may be due to imperfections in our models. For instance, we use a simplistic prescription for the nebular density profile, which may not be realistic enough.

Because we do not impose any a priori solution for the N/O versus O/H behaviour, unlike the tight correlation from the grid by Blanc et al. (2015) or the model selection by Pérez-Montero (2014), we find outliers in the N/O versus O/H plane. In Fig. 12, there are at least two objects with low O/H and high N/O in both panels, marked as large red points in panel (b) (SBS0335-052E and 0837+4717, id numbers 631 and 700 in Table 2). Those objects are likely interesting from a chemical evolution perspective.

6.2 The O/H ratio if one is not interested in N/O

Given that we do not get rid of multimodal solutions, it is to be expected that different ways to summarize the posterior (see Section 5.5) result in different solutions for O/H and N/O. In theory, if one is not interested in N/O, one would be better advised to use the mode, mean or median of the O/H PDF marginalized over N/O, and

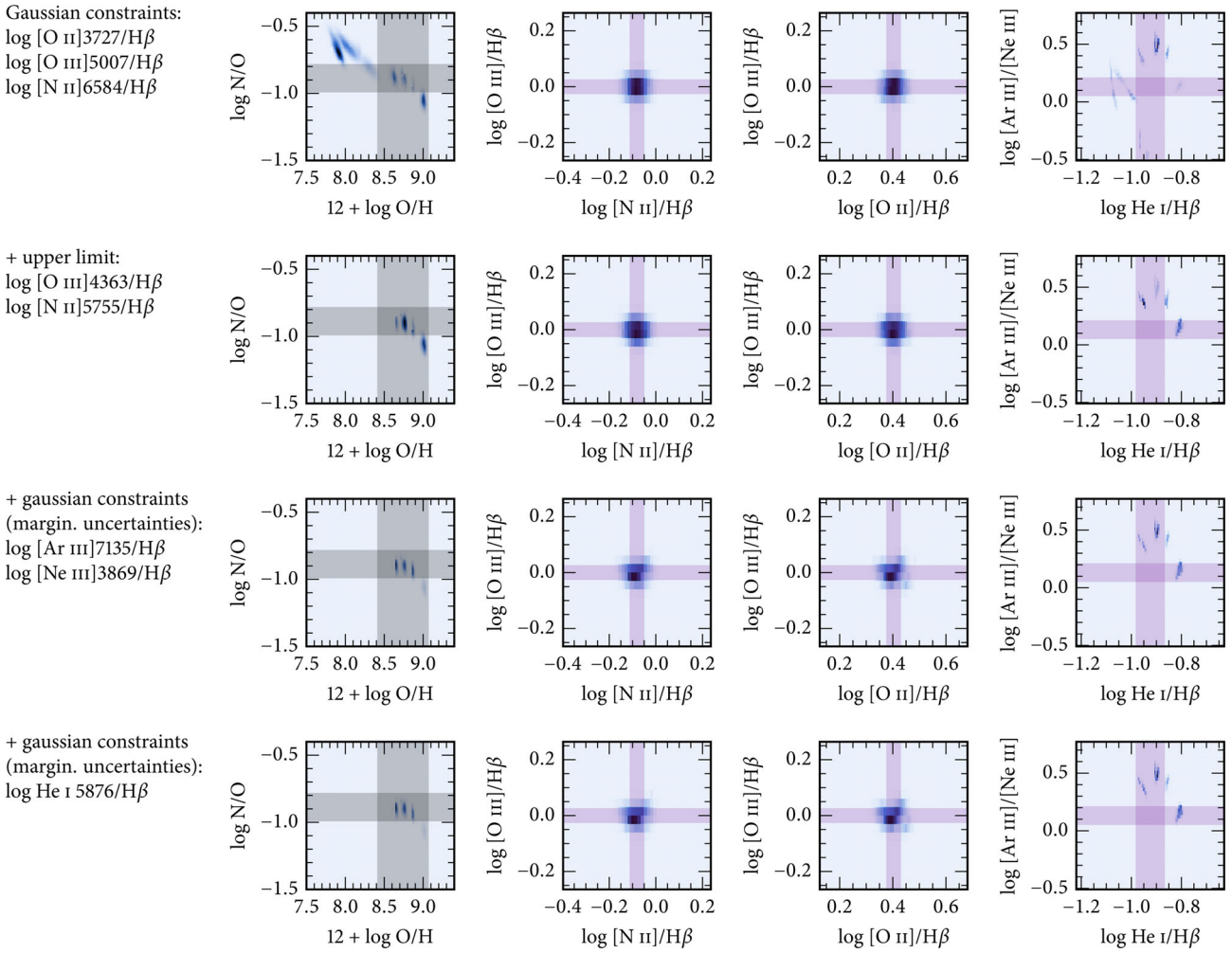

Figure 10. Example of the BOND solution for an H II region in NGC 1232 (slit number 5 from Bresolin et al. 2005). The panels from left to right show the joint posterior PDFs in blue for N/O versus O/H, [O III]/Hβ versus [N II]/Hβ, [O III]/Hβ versus [O II]/Hβ, and [Ar III]/[Ne III] versus He I/Hβ. The grey bands delimit $\pm 1\sigma$ of the temperature-based values for O/H and N/O, and the purple bands $\pm 1\sigma$ of the observed line ratios. Each row shows the effect of *cumulatively* adding another set of observational constraints. The top row shows the effect of a χ^2 likelihood for [O II]/Hβ, [O III]/Hβ and [N II]/Hβ. The second row applies the upper limit for [O III]λ4363/Hβ and/or [N II]λ5755/Hβ, which selects the solutions on the high-metallicity branch. The third row imposes the constraint on [Ar III]/Hβ and [Ne III]/Hβ, that further pins down the metallicity solution. Finally, the last row shows the effect of adding He I/Hb, which selects all possible ionization sources. Note that the N/O versus O/H PDF is multipeaked, which means that there is a family of acceptable solutions in our grid (affecting mainly O/H). Those islands of solutions are a consequence of the discreteness of the starburst ages and nebular geometries in our grid.

those are not expected to be exactly the same as the descriptions for the joint PDF.

Fig. 13 compares the O/H from MAP value of the joint N/O versus O/H PDF to the median of the marginalized O/H PDF. For most objects, those two solutions agree to within 0.1 dex: there is no bias between those two nominal solutions (the average difference is <0.02 dex) and the dispersion in small (0.06 dex). Other descriptions of the joint and marginalized PDFs may have a better or worse agreement, but the important point is to always test a few of those descriptions plus their credibility regions.

6.3 Comparison to the direct method

Fig. 14 shows a comparison of BOND to temperature-based results for the objects in sample B that have a direct temperature measure. The small black circles are the temperature-based results and the large coloured dots are the BOND results. The latter are colour-coded according to the difference between log [O III]λ4363/[O III]λ5007 in the models and the observed value (blue for extreme positive values to red for extreme negative values). The BOND and temperature-based results for the same H II region are joined by a solid line.

We see that the BOND results generally migrate towards higher values of O/H and lower values of N/O. The colour-coding shows that, for the majority of the points, the BOND models have lower temperatures than the observations. This explains why the BOND oxygen abundances are higher than the temperature-based ones (typically by 0.2–0.4 dex). Concomitantly, the BOND N/O ratios are smaller since the emissivity of the [N II] line is less dependent on the temperature than that of the [O II] line, which has a higher excitation threshold.

This problem is not unique either to our grid of models or to our code. Pérez-Montero (2014) only obtains O/H values that are in agreement with the direct method when [O III]λ4363 is fitted (see his fig. 2). His method gives a huge weight to [O III]λ4363/Hβ (see his equation 19); in other words, it becomes essentially a temperature-based method. Blanc et al. (2015) compared the results

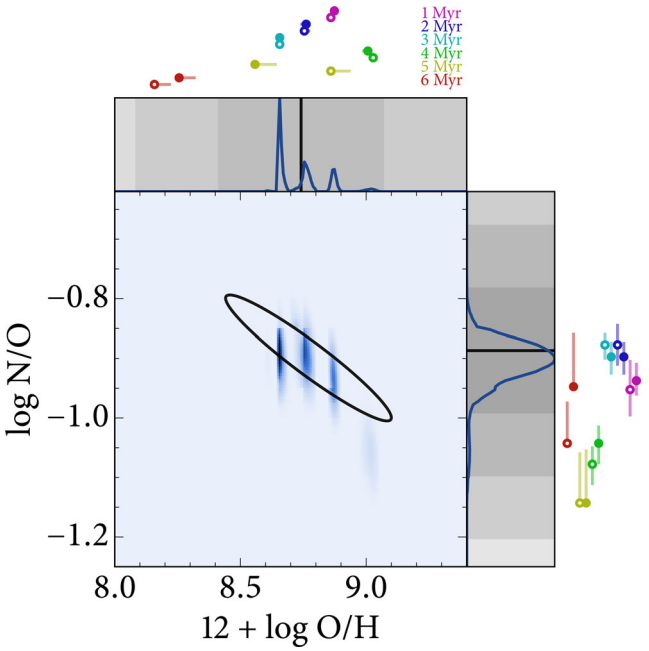

Figure 11. A zoom-in on the final N/O versus O/H panel for the same object in Fig. 10, plus the marginalized PDFs for O/H and N/O. This plot highlights how the discreteness of starburst ages in our grid of models leads to multipeaked solutions. We would expect a smoother PDF if the ages were more finely sampled. Above the marginalized PDFs, we mark the median (points) and the interval between the 16 and 84 percentiles (lines) of the marginalized PDF for each age and geometry combination. Starburst ages are ordered from 1 to 6 Myr from top to bottom and colour-coded with a rainbow palette, while spherical shells and filled spheres are represented by open and filled circles, respectively. The four islands of solutions on the joint PDF become a broad PDF in N/O and a multipeaked PDF in O/H. To compare to the temperature-based method, we plot its covariance ellipse from the Monte Carlo realizations on the joint PDF, and mark the nominal temperature-based solution and ± 1, 2 and 3σ as grey bands on the marginalized PDFs. The nominal temperature-based solution is marked as a black line on the marginalized PDF panels.

from their code IZI using several photoionization models (Kewley et al. 2001; Levesque, Kewley & Larson 2010; Dopita et al. 2013), and found offsets of −0.07 to 0.32 dex with respect to recombination lines, which translate into 0.17–0.56 dex offsets with respect to the temperature-based method.

It is a well-known problem that collisionally excited lines, when using temperature-based methods, lead to lower oxygen abundances than recombination lines (see e.g. García-Rojas & Esteban 2007), typically by 0.2–0.3 dex in H II regions. One explanation for this abundance discrepancy problem could be that, in real nebulae, strong temperature fluctuations (Peimbert 1967; more important than the temperature gradients arising in classical photoionization models) or a κ distribution of the free electron velocities (Nicholls et al. 2012) boost the [O III]λ4363 line, a fact which is not taken into account in the classical temperature-based method nor in the grid of photoionization models we use.

The simplicity of the direct method can be a double-edged sword: although powerful and straightforward to apply, it might be holding too simplistic assumptions as to the production of [O III]λ4363.

7 SUMMARY, DISCUSSION AND FUTURE DIRECTIONS

BOND determines nitrogen and oxygen gas-phase abundances by using strong and semistrong lines and comparing them to a grid of photoionization models in a Bayesian framework. The code is written in python and its source is publicly available at http://bond.ufsc.br. The grid of models presented here is included in the 3MdB data base (Morisset, Delgado-Inglada & Flores-Fajardo 2015, see https://sites.google.com/site/mexicanmillionmodels/) under the reference 'BOND'. The Bayesian posterior probability calculated by BOND stands on two pillars: our grid of models and our choice of observational constraints (from which we calculate our likelihoods). We discuss each of these in turn.

The ideal grid of models should be all-encompassing and able to describe any emission-line object found in nature. Creating such a grid would be a daunting and nearly impossible task; therefore we have crafted a set of models that covers enough physical parameters of H II regions not to be plagued by the usual preconceptions that go into making these grids. Our models span a wide range in N/O and O/H, without imposing any relation between N/O and O/H. The only model grid that has so far taken this approach is the one by Pérez-Montero (2014). Unlike his method, we leave the starburst age and the nebular density structure as free parameters. Finally, a crucial step forward in our approach is taking into account the importance of the hardness of the ionization field. All model grids in the literature consider only a single type of ionizing sources. If the ionization field of the H II regions differs from those in the models, the O/H obtained will be strongly biased (see the discussion in Appendix B). The hardness of the ionization source may vary due to a few reasons. In local galaxies, the main effects will be the ageing of the stellar populations in H II regions, and the stochastic sampling of the stellar initial mass function (e.g. Cerviño et al. 2013,

Table 2. A sample of the posterior PDF summaries for sample B available for download at http://bond.ufsc.br. For log O/H and log N/O, we report the maximum a posteriori (jmod); the centre, dispersion, covariance term and scaling to construct the 68 per cent credibility ellipse (jc68 cen, sig, cov, scale); and the marginalized median (mmed) and 68 per cent equal-tailed interval (mp68 low, mp68 upp). See Section 5.5 for details on how those terms have been defined.

id	name	log O/H jmod	log N/O jmod	log O/H jc68 cen	log O/H jc68 sig	...	log O/H mmed	log N/O mmed	log O/H mp68 low	log O/H mp68 upp	...
002	NGC 1232 03	−3.7168	−1.0797	−3.6198	0.0800	...	−3.6359	−1.0949	−3.7195	−3.5279	...
003	NGC 1232 04	−2.8035	−0.9974	−2.8047	0.0095	...	−2.8041	−0.9897	−2.8180	−2.7937	...
004	NGC 1232 05	−3.3440	−0.8872	−3.2526	0.0807	...	−3.2655	−0.9034	−3.3454	−3.1340	...
009	NGC 1232 10	−2.9860	−0.8719	−2.9871	0.0663	...	−3.0012	−0.8691	−3.0258	−2.9600	...
...	...	...	...	...	...	...	...	...	...	...	...
704	F1629+205	−3.7860	−1.3184	−3.7247	0.0693	...	−3.7279	−1.3333	−3.7926	−3.6240	...
706	Mrk259	−3.8633	−1.2782	−3.7037	0.1904	...	−3.7960	−1.2626	−3.8678	−3.4639	...
707	SBS1428	−3.7087	−1.1127	−3.7678	0.1419	...	−3.7127	−1.1080	−3.9198	−3.5843	...
708	S1657+575	−3.8281	−0.8720	−3.8408	0.1062	...	−3.8013	−0.8999	−3.9260	−3.6710	...

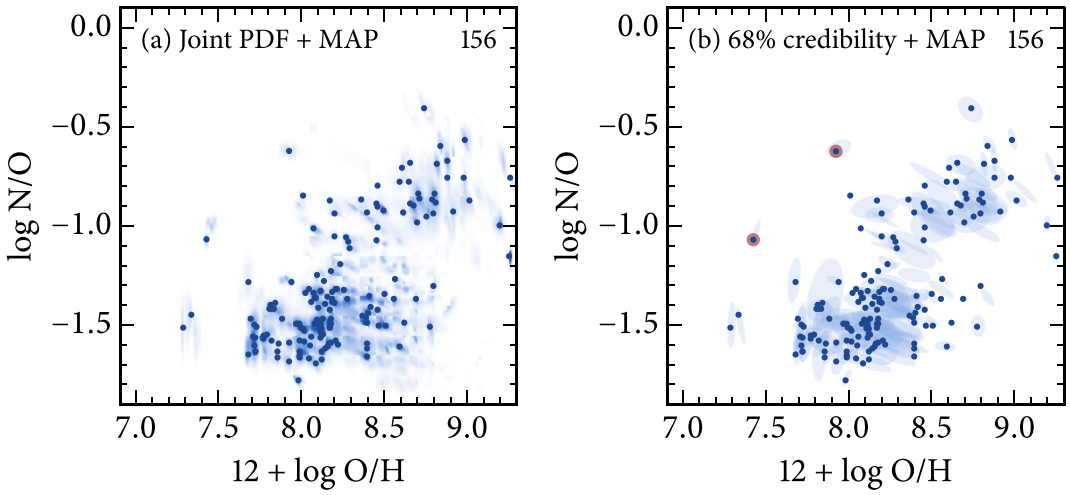

Figure 12. Blue points on both panels are the maximum a posteriori (MAP) for each object in sample B. The blue scale on each map is a different description of our results. (a) Superposition of the N/O versus O/H joint posterior PDFs for all objects. The blue gradient is linear and scaled such than the joint PDF for each object integrates to unity over this diagram. (b) The ellipses are the covariances of the 68 per cent credibility region, and are scaled to cover the same area as the credibility region. Strong multimodal solutions appear as MAP points far-flung from their elliptical companions. Large red dots mark the two outliers on the N/O versus O/H plane (blue compact dwarves SBS0335-052E and 0837+4717).

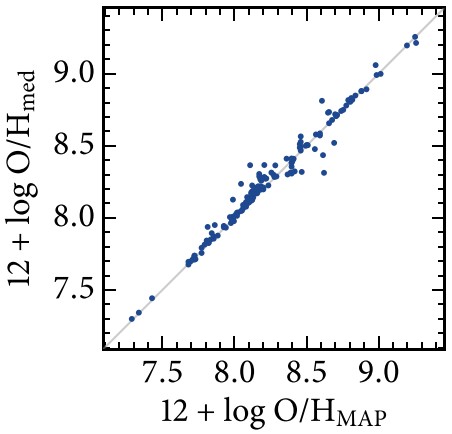

Figure 13. Comparison between the maximum a posteriori O/H and the marginalized median O/H. The identity line is drawn in grey.

and references therein). Our model grid attacks this by using simple stellar populations of different ages for the ionizing sources to account for variations in the ionizing field.

Constructing the grid with great care is not enough. Given the strengths and limitations of our set of models, we need to critically assess which theoretical predictions should be trusted and which observational constraints should go into our fitting procedure. We have set out to infer O/H and N/O, but also to constrain the nuisance parameters U, the correct O/H bimodality branch, and the hardness of the ionizing field. The strong lines [O III]/Hβ, [O II]/Hβ and [N II]/Hβ constrain O/H, N/O and U. To pin down the correct O/H branch, we use an upper limit in [O III]λ4363 or [N II]λ5755 when at hand, and the ratio of the semistrong lines [Ar III]/[Ne III], which depends mostly on the electron temperature (modulo the ionizing structure of the nebula, already constrained by [O III]/[O II]). Indeed [Ar III] and [Ne III] have different excitation thresholds while Ar/Ne in the gas phase is expected to be constant, since both argon and neon are primary elements and inert. Lastly, another semistrong line comes to rescue: He I/Hβ helps constrain the mean effective temperature of the ionizing radiation field.

Unlike several authors (Dopita et al. 2013; Pérez-Montero 2014; Blanc et al. 2015), we do not use the [S II]λ6716 + λ6731 line in our procedure, first because this line is emitted in the outskirts of H II regions so that its intensity in relation with [O II] or [N II] is dependent on the detailed density structure of the nebulae, second because there might be an intrinsic scatter in the S/O ratio due to stellar nucleosynthesis and/or to depletion effects not yet fully documented.

For a set of giant H II regions and blue compact dwarf galaxies, we have calculated their gas-phase N/O and O/H abundances and compared them to the ones obtained by the temperature-based method and by the Pilyugin et al. (2010) ON method. We find that the N/O versus O/H relation obtained by BOND is as scattered as the one obtained by the temperature-based method, and that the very tight relation obtained with the Pilyugin method is a consequence of that method itself.

We also note that, when using the BOND method on objects which have direct temperature measurements, we systematically obtain lower values of [O III]λ4363/[O III]λ5007 than observed and higher values of the oxygen abundance than with the temperature-based method. This discrepancy has been seen in many other strong-line methods calibrated on photoionization models and might point to too soft an ionizing SED in the models. This is in line with the fact that Stasińska et al. (2015), using a subset of our models, found that the SEDs are not hard enough to produce the observed [Ar IV]/[Ar III] line ratio. It might also indicate that the density distributions of our models are too simplistic to represent real H II regions. Alternatively, it might be a sign that important temperature fluctuations of unknown origin or a κ distribution of electron velocities are present in real H II regions and lead to an overestimate of the temperature indicated by [O III]λ4363/[O III]λ5007. Such an explanation has been suggested by several authors and would at the same time help resolve the famous abundance discrepancy problem (see e.g. García-Rojas & Esteban 2007; Nicholls et al. 2012; Dopita et al. 2013). Since we do not try reproduce the [O III]λ4363 line, nitrogen and oxygen

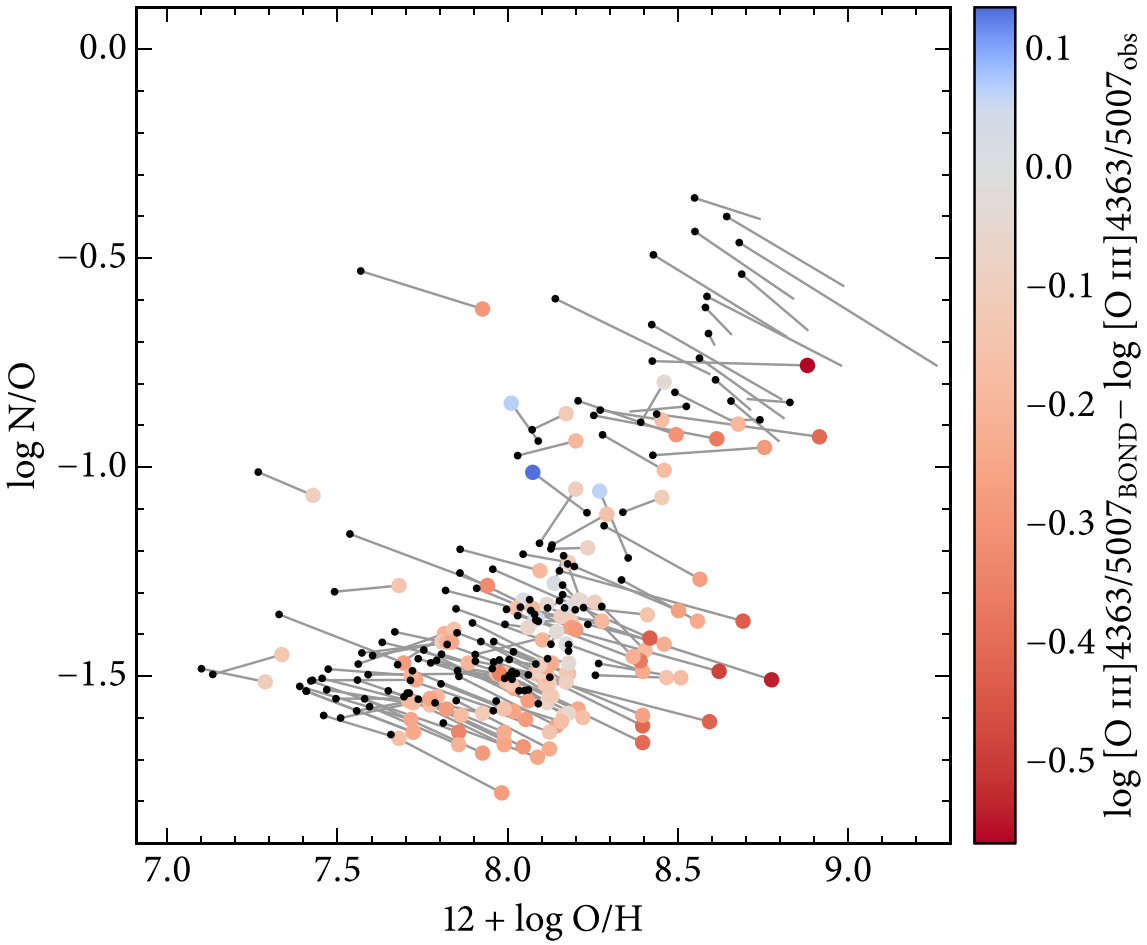

Figure 14. Comparison of N/O versus O/H as obtained with the temperature-based method (small black circles) and with the BOND method (large coloured circles). Solutions for the same object are joined by a line. The large circles are coloured with respect to the difference between [O III]λ4363/[O III]λ5007 for the fitting models and the measured [O III]λ4363/[O III]λ5007. Lines lacking a coloured counterpart indicate there is no measured [O III]λ4363 (i.e. the direct method used [N II]λ5755/[N II]λ6584 to measure the electron temperature). Note how points tend to shift to higher O/H with BOND, and that the temperature in the models tend to be smaller than the one inferred from the auroral lines. A colour version of this figure is available in the electronic edition.

abundances inferred by BOND might be more accurate estimates than those of the temperature-based method. Nevertheless, the accuracy of BOND abundances strongly relies on how well the photoionization model grid represents real objects.

We have shown that BOND, when applying our extensive grid of photoionization models to a well-chosen set of strong and semistrong lines, allows one to obtain O/H and N/O simultaneously, getting rid of the spectre of bimodality without recourse to empirical oxygen and nitrogen abundance correlations. Our method is very easily extendable and can accommodate many improvements in the future. In spite of many issues still to resolve in the determination of nebular abundances, we hope that BOND does offer a quantum of solace.

ACKNOWLEDGEMENTS

We are thankful to Robert Kennicutt and the referee, Guillermo Blanc, for their insightful comments, which led us to rephrase a number of issues to improve clarity. We also thank Guillermo Blanc for the kindness of running IZI for us in the refereeing process. GS and NVA acknowledge the support from the CAPES CsF–PVE project 88881.068116/2014-01. NVA acknowledges the NEBULATOM school held in Choroni, Venezuela, in 2013, and the support and hospitality for short term visits of the LUTH, at Observatoire de Paris. GS and RCF acknowledge the support from the CAPES-COFECUB Te 585/07 project. The grid of models has been run on computers from the CONACyT/CB2010:153985, UNAM-PAPIIT-IN107215 and UNAM Posgrado de Astrofísica projects. This research made use of Astropy, a community-developed core PYTHON package for Astronomy; MATPLOTLIB, a PYTHON library for publication quality graphics; SCIPY; NUMPY; h5py; and hickle.

REFERENCES

Aggarwal K. M., Keenan F. P., 1999, ApJS, 123, 311
Alloin D., Collin-Souffrin S., Joly M., Vigroux L., 1979, A&A, 78, 200
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Blanc G. A., Kewley L., Vogt F. P. A., Dopita M. A., 2015, ApJ, 798, 99
Bresolin F., 2007, ApJ, 656, 186
Bresolin F., Garnett D. R., Kennicutt R. C., 2004, ApJ, 615, 228
Bresolin F., Schaerer D., González Delgado R. M., Stasińska G., 2005, A&A, 441, 981
Bresolin F., Ryan-Weber E., Kennicutt R. C., Goddard Q., 2009a, ApJ, 695, 580
Bresolin F., Gieren W., Kudritzki R.-P., Pietrzyński G., Urbaneja M. A., Carraro G., 2009b, ApJ, 700, 309

Bresolin F., Stasińska G., Vílchez J. M., Simon J. D., Rosolowsky E., 2010, MNRAS, 404, 1679
Bresolin F., Kennicutt R. C., Ryan-Weber E., 2012, ApJ, 750, 122
Cerviño M., Román-Zúñiga C., Luridiana V., Bayo A., Sánchez N., Pérez E., 2013, A&A, 553, A31
Chabrier G., 2003, PASP, 115, 763
Delgado-Inglada G., Rodríguez M., Peimbert M., Stasińska G., Morisset C., 2015, MNRAS, 449, 1797
Dopita M. A., Sutherland R. S., Nicholls D. C., Kewley L. J., Vogt F. P. A., 2013, ApJS, 208, 10
Draine B. T., 2011, ApJ, 732, 100
Esteban C., García López R., Herrero A., Sánchez F., eds, 2004, Cosmochemistry. The Melting Pot of the Elements. Cambridge Univ. Press, Cambridge, p. 298
Ferland G. J. et al., 2013, Rev. Mex. Astron. Astrofis., 49, 137
Galavis M. E., Mendoza C., Zeippen C. J., 1997, A&AS, 123, 159
García-Rojas J., Esteban C., 2007, ApJ, 670, 457
Garnett D. R., 1992, AJ, 103, 1330
Goddard Q. E., Bresolin F., Kennicutt R. C., Ryan-Weber E. V., Rosales-Ortega F. F., 2011, MNRAS, 412, 1246
Hogg D. W., 2012, preprint (arXiv:1205.4446)
Izotov Y. I., Thuan T. X., Stasińska G., 2007, ApJ, 662, 15
Kennicutt R. C., Bresolin F., Garnett D. R., 2003, ApJ, 591, 801
Kewley L. J., Dopita M. A., 2002, ApJS, 142, 35
Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183
Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, ApJ, 556, 121
Kisielius R., Storey P. J., Ferland G. J., Keenan F. P., 2009, MNRAS, 397, 903
Levesque E. M., Kewley L. J., Larson K. L., 2010, AJ, 139, 712
Li Y., Bresolin F., Kennicutt R. C., 2013, ApJ, 766, 17
McGaugh S. S., 1991, ApJ, 380, 140
McGaugh S. S., 1994, ApJ, 426, 135
Maiolino R. et al., 2008, A&A, 488, 463
Mendoza C., Zeippen C. J., 1983, MNRAS, 202, 981
Mollá M., Gavilán M., 2010, Mem. Soc. Astron. Ital., 81, 992
Mollá M., García-Vargas M. L., Bressan A., 2009, MNRAS, 398, 451
Morisset C., Delgado-Inglada G., Flores-Fajardo N., 2015, Rev. Mex. Astron. Astrofis., 51, 103
Nicholls D. C., Dopita M. A., Sutherland R. S., 2012, ApJ, 752, 148
Pagel B. E. J., Edmunds M. G., Blackwell D. E., Chun M. S., Smith G., 1979, MNRAS, 189, 95
Peimbert M., 1967, ApJ, 150, 825
Pérez-Montero E., 2014, MNRAS, 441, 2663
Pilyugin L. S., Vílchez J. M., Thuan T. X., 2010, ApJ, 720, 1738
Pilyugin L. S., Vílchez J. M., Mattsson L., Thuan T. X., 2012, MNRAS, 421, 1624
Rémy-Ruyer A. et al., 2014, A&A, 563, A31
Saftly W., Camps P., Baes M., Gordon K. D., Vandewoude S., Rahimi A., Stalevski M., 2013, A&A, 554, A10
Stasińska G., 2005, A&A, 434, 507
Stasińska G., 2010, in Bruzual G., Charlot S., eds, Proc. IAU Symp. 262, Stellar Populations – Planning for the Next Decade. Kluwer, Dordrecht, p. 93
Stasińska G., Cid Fernandes R., Mateus A., Sodré L., Asari N. V., 2006, MNRAS, 371, 972
Stasińska G., Peña M., Bresolin F., Tsamis Y. G., 2013, A&A, 552, A12
Stasińska G., Izotov Y., Morisset C., Guseva N., 2015, A&A, 576, A83
Storchi-Bergmann T., Calzetti D., Kinney A. L., 1994, ApJ, 429, 572
Storey P. J., Zeippen C. J., 2000, MNRAS, 312, 813
Tayal S. S., 2011, ApJS, 195, 12
Tayal S. S., Zatsarinny O., 2010, ApJS, 188, 32
van Zee L., Salzer J. J., Haynes M. P., O'Donoghue A. A., Balonek T. J., 1998, AJ, 116, 2805
Vilchez J. M., Esteban C., 1996, MNRAS, 280, 720
Vilchez J. M., Pagel B. E. J., 1988, MNRAS, 231, 257
Zeippen C. J., 1982, MNRAS, 198, 111
Zurita A., Bresolin F., 2012, MNRAS, 427, 1463

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table 1. Data table for sample A. Line fluxes, uncertainties and upper limits (F, eF and limF) are in units of Hβ, and references are labelled in column r as in Sections 2.1 and 2.2.
Table 2. Posterior PDF summaries for sample B. For log O/H and log N/O, we report the maximum a posteriori (jmod); the centre, dispersion, covariance term and scaling to construct the 68 per cent credibility ellipse (jc68 cen, sig, cov, scale); and the marginalized median (mmed) and 68 per cent equal-tailed interval (mp68 low, mp68 upp) (http://www.mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stw971/-/DC1).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX A: A FAKE SAMPLE FOR TESTS

To assess some of the recipes used in temperature-based methods and to test the BOND method on objects with known abundances, we construct a 'fake' sample by selecting from our grid of models a subsample that roughly follows the expected properties of our observational sample.

To limit the size of our fake sample, we have created an interpolated grid having a resolution of 0.1 dex in O/H and N/O, and 0.5 dex in U. We then perturb each cell point in the grid with uniform noise (setting its maximum amplitude to be the size of the cell) in those three input parameters, so that the fake sources are not superimposed in our plots. Our 'fake' sources are then chosen to fall roughly in the same loci as observed data, as Fig. A1 shows. We select models in the vicinity of the observed N/O and O/H relation as expressed by equation (2) of Pilyugin et al. (2012), around the U versus O/H relation found by Pérez-Montero (2014), and below the Stasińska et al. (2006) line delimiting pure H II regions in the BPT diagram. For a given age and geometry, we have around ∼350–400 fake sources (except for the 6 Myr scenarios, which fail to cover a large part of the observational data and thus have ∼160 fake sources).

Fig. A2(a) shows the relation between the temperatures in the high and in the low-excitation zones and Fig. A2(b) shows the relation between the ionic fractions of $N^+$and O^+. The temperatures and ionic fractions come directly from the photoionization models for our fake source sample. The continuous lines indicate the relations we used in the temperature-based method. We see that they represent well the trends shown by the models. We also see that the models show some dispersion about these lines. For the temperature the dispersion is of 600 K, while, for the logarithm of ionic abundances, it is of 0.06 dex.

APPENDIX B: TESTS OF THE ACCURACY OF THE BOND METHOD

Here we run a suite of tests fitting models with models, using the same code and the same assumptions as for the sources in sample B. The aim of this exercise is twofold. First, we show that our method works when the input and the output are the same, which is the zeroth test of reliability of any method. Secondly, we check how

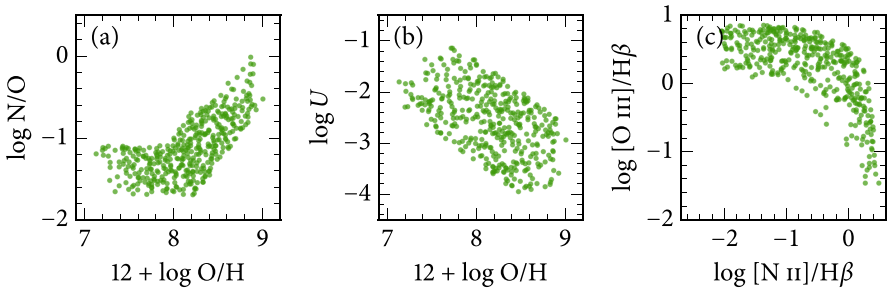

Figure A1. Fake sources chosen from the grid with age 2 Myr and filled sphere. The grid is finer than our original grid, but not as fine as the interpolated one we use as our initial octree grid. The panels show our selection of grid models on the N/O versus O/H, U versus O/H and BPT planes.

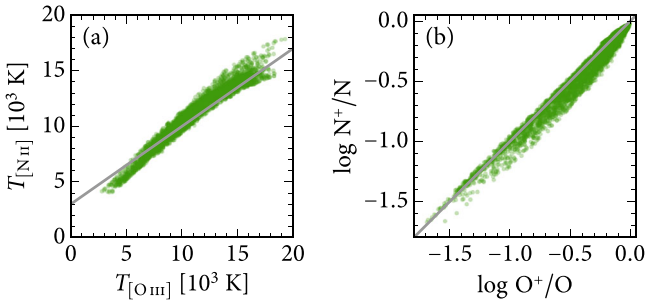

Figure A2. (a) Comparison between the temperatures in the high- and low-excitation zones for the fake sources. The grey line is the classical relation from Garnett (1992) given by equation (1). (b) Comparison between the ionic fractions of $N^+$and O^+. The one-to-one line is in grey.

the different ionizing source ages and density structures affect our results, since this is the main novelty of our model grid.

The model grid considered is the octree sampled grid. For the tests in this section, we select subgrids of single ages and density structures to highlight the effects of those parameters. For all the tests presented here, we assume that the uncertainties in the intensity ratios [O III]/Hβ, [O II]/Hβ, [N II]/Hβ, [Ar III]/Hβ, [Ne III]/Hβ, and He I/Hβ are of 10 per cent.

The results of our tests are shown in Fig. B1, where the different rows correspond to different choices of age and geometry. The first row shows 2-Myr starburst and spherical shell fake sources modelled with a grid of the same age and geometry. For O/H and N/O, we show the difference Δ between the output and the input as a function of the input parameter. The results shown are for the MAP values. N/O and O/H are well recovered (the dispersion is 0.007 dex).

The second and third rows of Fig. B1 show the effect of using the wrong density structure and age, respectively. The second row shows the 2 Myr filled sphere fake sources fitted with 2 Myr spherical shell models. The residuals for O/H and N/O are very dispersed (0.02 and 0.06 dex, respectively) and slightly biased (0.006 and 0.02 dex). The third row shows the effect of using the wrong hardness for the ionization source. Here we have 4 Myr fake sources fitted with a 2 Myr grid, both modelled as filled spheres. This is a very worrying scenario: O/H is underestimated by $\sim$0.1 dex (and up to 0.4 dex) for the high-metallicity branch, while N/O is slightly overestimated (0.04 dex) and rather dispersed (0.05 dex).

The last row shows how our 2 Myr filled sphere fake sources are modelled using our entire grid, i.e. without any a priori knowledge of the geometry and the age of the ionizing source. The results are quite encouraging, and the code seems to choose the right age and geometry combination; or, at least in practice, the right O/H and N/O solution. O/H and N/O are recovered to within better than 0.05 dex (0.02 dex of dispersion).

APPENDIX C: COMPARISON WITH OTHER STRONG-LINE METHODS

Here we compare our BOND method to several other strong-line methods. Fig. C1 shows the comparison to the O/H measured by McGaugh (1991), using equations (A1) and (A2) from Kewley & Ellison (2008). The two panels show the effect of choosing different criteria to separate the low and high-metallicity solutions: on the right, we use $\log$ [N II]/[O II] $= -1$ (as in McGaugh 1994), and on the left -1.2 (as in Kewley & Ellison 2008). The separation between the two branches is fuzzy, and the effect of choosing a slightly different frontier is seen when comparing one panel to the other. Focusing on the comparison of McGaugh (1991) to BOND, we see a good agreement between the results from BOND and those from the much simpler McGaugh recipe, but there are important differences for a non-negligible number of sources at high metallicities. There are, as expected, huge differences around 12 + log O/H = 8.5, where the O23 ratio is insensitive to metallicity while BOND is aided by using the [Ar III]/[Ne III] ratio, which steadily increases with increasing metallicity.

McGaugh (1991) was the first to take into account the effect of the ionization parameter when measuring abundances. Fig. C2 shows the comparison of BOND results with those using a simple O23 calibration (we have used the one by Maiolino et al. 2008 as an example). The systematics with this simple O23 calibrator are of the order of 0.2–0.5 dex over the whole O/H range.

Fig. C3 compares the BOND results obtained with our full grid of models to those from the IZI code by Blanc et al. (2015) with their default grid (Levesque et al. 2010, 6 Myr constant star formation). Blanc et al. (2015), like McGaugh (1991), assume a N/O versus O/H relation and consider a unique family of ionizing stellar energy distributions. Panel (a) shows the comparison between the values of O/H derived by BOND and IZI for the 151 objects in sample B (see Section 2.3) which have the IZI quality flags `npeakZ` and `limZ` equal to one (this removes only five objects from sample B). For $\log$ O/H $\lesssim 8.4$, IZI metallicities are systematically larger than BOND by 0.1–0.2 dex. For $\log$ O/H $\gtrsim 8.4$, the O/H from BOND can be 0.3–0.8 dex larger IZI for a few objects, while for other objects, the codes agree quite well (differing by $\lesssim$0.02 dex). Panel (b) of Fig. C3 shows the values of N/O versus O/H derived by BOND linked by a straight line to the values obtained by Blanc et al. for the same objects (actually, Blanc et al. determines only O/H, since the N/O values lie on the relation assumed by them). We see that some objects are actually quite far from the tight N/O versus O/H relation assumed, and that for those objects the O/H values derived by BOND

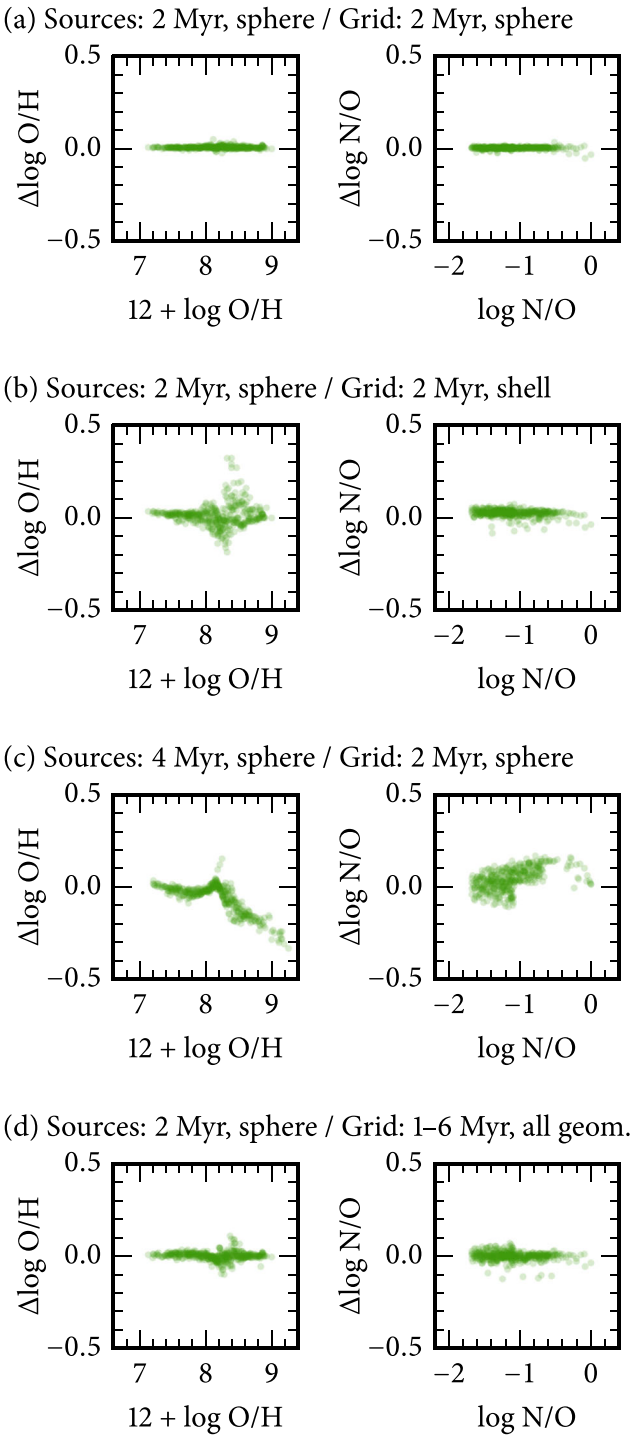

Figure B1. Results from modelling the grid fake sources with BOND. The panels show the residual parameters (output minus input) for O/H and N/O. The top row shows that we recover O/H and N/O quite well when we fit sources from the 2 Myr filled sphere grid with a grid of the same age and geometry. The second row swaps the filled sphere by a shell in the fitting grid. The third row fits 4 Myr fake sources with a grid of the same geometry but 2 Myr ionizing sources. We see that O/H is highly biased for the high-O/H solutions. The last row shows 2 Myr filled spheres fitted with our entire grid. The results for O/H are very good, with a dispersion of only 0.02 dex.

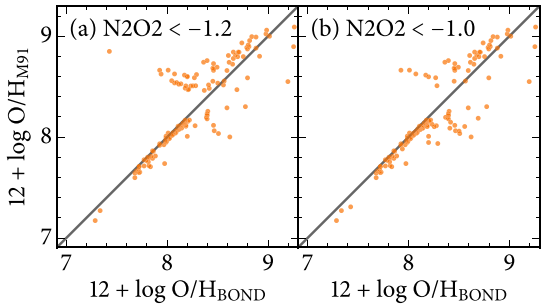

Figure C1. Comparison between O/H from BOND and from McGaugh (1991) for our sample B. The panel on the right shows O/H from McGaugh by choosing the O/H according to $\log [N\,II]/[O\,III] = -1$ (McGaugh 1994), and on the left according to $\log [N\,II]/[O\,II] = -1.2$ (Kewley & Ellison 2008).

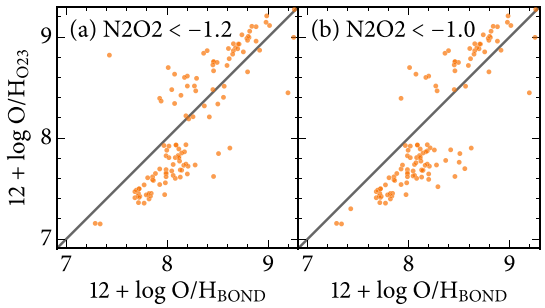

Figure C2. As Fig. C1, but for O/H derived with the O23 calibration from Maiolino et al. (2008).

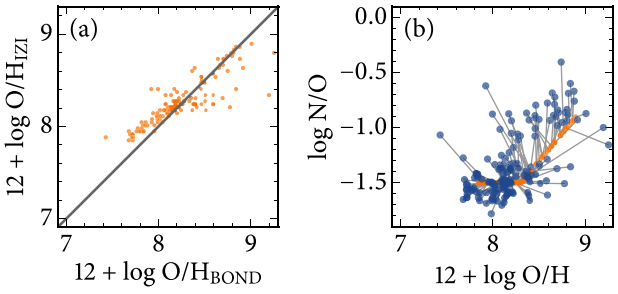

Figure C3. How the assumption of a relation between N/O and O/H affects the derived O/H. Panel (a) compares the values of O/H obtained by BOND to those by IZI (Blanc et al. 2015) with their default grid, which assumes a relation between N/O and O/H. We show the results for 151 objects in our sample B that have also been flagged as having reliable results by IZI. Panel (b) shows N/O as a function of O/H for BOND and IZI. The results using the Blanc et al. (2015) code and grid are the small orange points, and those with BOND are large blue points. Results for the same object are linked by a grey line.

differ substantially from those derived by Blanc et al. This illustrates that, for a number of objects (which are not the majority but are not known a priori) it is necessary to simultaneously derive N/O and O/H to obtain a reliable oxygen abundance.

Fig. C4 shows the comparison between BOND and the HII-CHI-MISTRY code by Pérez-Montero (2014) for our sample B. We do not show objects that the Pérez-Montero (2014) code flags as bad, i.e. when his `grid` output is either 2 or 3. From the original 156 objects in sample B, we are left with 129 objects. We use the version 1.2 of his code. The figure shows N/O versus O/H as obtained by the Pérez-Montero (2014) code on the left, and the comparison of

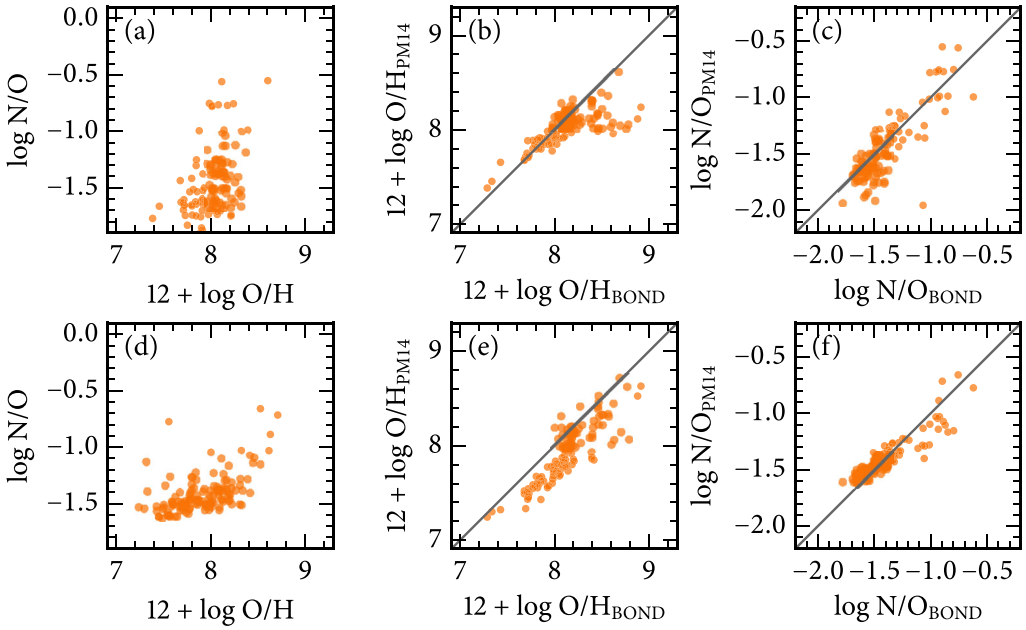

Figure C4. Results for applying the HII-CHI-MISTRY code by Pérez-Montero (2014) to our sample B. Top and bottom rows, respectively, show results excluding or including [O III]λ4363 from the HII-CHI-MISTRY fits. From left to right, the N/O versus O/H diagram, and the comparison between the O/H and N/O values to BOND.

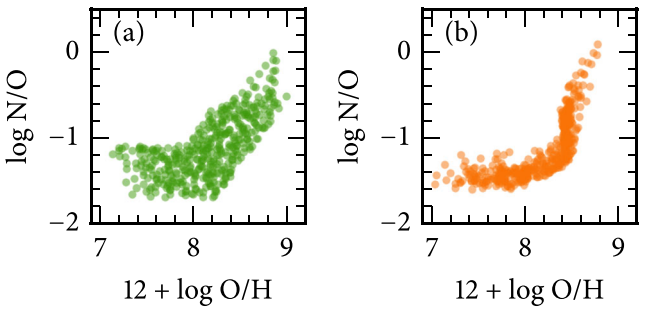

Figure C5. (a) The original N/O versus O/H diagram for our fake sources, and (b) the N/O versus O/H diagram for our fake sources as calculated by the ON calibration from Pilyugin et al. (2010).

the O/H and N/O values to ones obtained by BOND on the middle and right-hand panels. The two rows correspond to two different runs of the Pérez-Montero (2014) code. On the top row, we have withheld the [O III]λ4363 line from the code, to see how it would behave using only strong lines. This is not the recommended way of running the Pérez-Montero (2014) code, but this exercise shows that it cannot be used as a strong-line method. The bottom row shows the results from the Pérez-Montero (2014) code when asking it to fit the [O III]λ4363 line as well (with a very strong weight as resulting from his equation 19). Note that the Pérez-Montero (2014) code finds systematically lower values of O/H than the BOND method, exactly like the temperature-based method. The N/O is also rather scattered.

Finally, we use the fake sample of Fig. A1 to show how the ON strong-line method of Pilyugin et al. (2010) biases the abundance results in the N/O versus O/H diagram. As seen in Fig. C5, the ON calibration from Pilyugin et al. (2010) has considerably squeezed the broad input relation.

This paper has been typeset from a TEX/LATEX file prepared by the author.