















PHANGS-MUSE: Detection and Bayesian classification of ~40 000 ionised nebulae in nearby spiral galaxies[★]

Enrico Congiu^{1,2}, Guillermo A. Blanc^{3,1}, Francesco Belfiore⁴, Francesco Santoro⁵, Fabian Scheuermann⁶, Kathryn Kreckel⁶, Eric Emsellem^{7,8}, Brent Groves⁹, Hsi-An Pan¹⁰, Frank Bigiel¹¹, Daniel A. Dale¹², Simon C. O. Glover¹³, Kathryn Grasha^{14,15}, Oleg V. Egorov⁶, Adam Leroy^{16,17}, Eva Schinnerer⁵, Elizabeth J. Watkins⁶, and Thomas G. Williams⁵

¹ Departamento de Astronomía, Universidad de Chile, Camino del Observatorio 1515, Las Condes, Santiago, Chile
e-mail: econgiu@das.uchile.cl

² European Southern Observatory (ESO), Alonso de Córdova 3107, Casilla 19, Santiago 19001, Chile

³ Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA

⁴ INAF – Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50157 Firenze, Italy

⁵ Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

⁶ Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12–14, 69120 Heidelberg, Germany

⁷ European Southern Observatory, Karl-Schwarzschild Straße 2, 85748 Garching bei München, Germany

⁸ Univ. Lyon, Univ. Lyon1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, 69230 Saint-Genis-Laval, France

⁹ International Centre for Radio Astronomy Research, University of Western Australia, 7 Fairway, Crawley, WA 6009, Australia

¹⁰ Department of Physics, Tamkang University, No. 151, Yingzhan Road, Tamsui District, New Taipei City 251301, Taiwan

¹¹ Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany

¹² Department of Physics & Astronomy, University of Wyoming, Laramie, WY 82070, USA

¹³ Institut für Theoretische Astrophysik, Zentrum für Astronomie der Universität Heidelberg, Albert-Ueberle-Strasse 2, 69120 Heidelberg, Germany

¹⁴ Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

¹⁵ ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia

¹⁶ Department of Astronomy, The Ohio State University, 140 West 18th Avenue, Columbus, OH 43210, USA

¹⁷ Center for Cosmology and Astroparticle Physics, 191 West Woodruff Avenue, Columbus, OH 43210, USA

Received 5 October 2022 / Accepted 22 January 2023

ABSTRACT

In this work, we present a new catalogue of >40 000 ionised nebulae distributed across the 19 galaxies observed by the PHANGS-MUSE survey. The nebulae have been classified using a new model-comparison-based algorithm that exploits the odds ratio principle to assign a probabilistic classification to each nebula in the sample. The resulting catalogue is the largest catalogue containing complete spectral and spatial information for a variety of ionised nebulae available so far in the literature. We developed this new algorithm to address some of the main limitations of the traditional classification criteria, such as their binarity, the sharpness of the involved limits, and the limited amount of data they rely on for the classification. The analysis of the catalogue shows that the algorithm performs well when selecting H II regions. In fact, we can recover their luminosity function, and its properties are in line with what is available in the literature. We also identify a rather significant population of shock-ionised regions (mostly composed of supernova remnants), which is an order of magnitude larger than any other homogeneous catalogue of supernova remnants currently available in the literature. The number of supernova remnants we identify per galaxy is in line with results in our Galaxy and in other very nearby sources. However, limitations in the source detection algorithm result in an incomplete sample of planetary nebulae, even though their classification seems robust. Finally, we demonstrate how applying a correction for the contribution of the diffuse ionised gas to the nebulae's spectra is essential to obtain a robust classification of the objects and how a correct measurement of the extinction using diffuse-ionised-gas-corrected line fluxes prompts the use of a higher theoretical $H\alpha/H\beta$ ratio (3.03) than what is commonly used when recovering the $E(B - V)$ via the Balmer decrement technique in massive star-forming galaxies.

Key words. galaxies: ISM – H II regions – planetary nebulae: general – ISM: supernova remnants – catalogs

* The catalogue of nebulae is only available at the CDS via anonymous ftp to cdsarc.cds.unistra.fr (130.79.128.5) or via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/672/A148>. The catalogue, together with the segmentation maps, is also available through the CADZ via <http://dx.doi.org/10.11570/23.0006>

1. Introduction

The interstellar medium (ISM) is one of the main components of galaxies, and its study is critical to understanding their formation and evolution. The vast majority of the ISM is in the form of gas, whose properties span a wide range of physical conditions, mostly depending on the gas phase (atomic, molecular, ionised) (e.g. [Ferrière 2001](#); [Haffner et al. 2009](#); [Cox 2005](#)). Each gas phase is related to specific properties of the host galaxy, and has a different role in its structure and evolution. In particular, the warm phase of the ISM, composed of partially ionised gas with temperatures around 8000–10 000 K, (e.g. [Osterbrock & Ferland 2006](#)), is probably one of the most studied components of the ISM since it produces an emission-line-rich spectrum in the optical band. While only a relatively small fraction of the mass of the ISM is in this form ($\sim 15\text{--}20\%$ in the Milky Way; [Ferrière 2001](#)), its study is crucial in understanding many properties of galaxies (e.g. kinematics, chemical composition and enrichment, and massive star formation). Part of this ionised gas is concentrated in nebulae connected to specific ionisation mechanisms, which are processes that provide energy to the gas and ionise it. The vast majority of these nebulae can be classified into three classes: H II regions, planetary nebulae (PNe), and supernova remnants (SNRs). The first class of nebulae, H II regions, are clouds of gas associated with regions of recent massive star formation. The young, massive and hot stars produced in these regions emit large quantities of ultraviolet photons with energies capable of ionising the surrounding hydrogen (13.6 eV). These nebulae are typically characterised by an electron temperature of the order of $\sim 10\,000$ K, relatively low electron density ($10\text{--}10^3\text{ cm}^{-3}$), and sizes varying between ~ 1 and 100 pc depending on the type, number, and distribution of the stars powering the nebula (e.g. [Maciel 2013](#)). Since they are strictly connected with the star formation process, H II regions are primarily found in star-forming galaxies, where they account for the majority of the observed nebulae. They are a precious tool for investigating many aspects of the star formation process and the properties of the ionised gas, such as the initial abundances of the gas from which stars are currently being produced (e.g. [Stasińska 2004](#)).

The second class of sources, PNe, are again clouds of gas photo-ionised by a central ionising source. This time, however, each nebula is ionised by a single star, typically an extremely hot (50 000–300 000 K; [Maciel 2013](#)) white dwarf. Those harder spectral sources produce high ionisation lines, which are significantly stronger than in H II regions. In particular, PNe are bright in the $[\text{O III}]\lambda 5007$ line, which is used to identify them in extragalactic environments. On the other hand, the low luminosity of the white dwarfs leads to relatively small ionised regions (e.g. < 1 pc for bright PNe; [Acker et al. 1992](#)). Planetary nebulae play an essential role in studying galaxies. For example, their luminosity function can be used to measure the distance of relatively nearby galaxies (e.g. [Jacoby 1989](#); [Ford et al. 1996](#); [Ciardullo et al. 2002](#); [Rekola et al. 2005](#); [Herrmann et al. 2008](#); [Kreckel et al. 2017](#); [Scheuermann et al. 2022](#)) and it is, therefore, a key step of the distance ladder. They are also widely used as test particles to trace galaxies' gravitational potential and dark matter content via dynamical modelling (e.g. [Hui et al. 1995](#); [Arnaboldi et al. 1996, 1998](#); [Douglas et al. 2002](#)). Their physics is also similar to that of H II regions, so it is relatively easy to use their spectra to study, for example, the temperature, density and chemical composition of the gas ([Stasińska 2004](#)). However, PNe trace gas that has been processed by the central star in its previous life stages, while H II regions typically trace the unprocessed gas from which they were recently formed.

Finally, SNRs, as their name says, are the result of a supernova explosion interacting with the ISM. During a supernova explosion, a large amount of material processed by the exploding star is ejected into the ISM, where it starts expanding at velocities of the order of one to ten thousands of km s^{-1} . This creates powerful shock waves that heat up the gas to extremely high temperatures ($\sim 10^7\text{--}10^8$ K). In these nebulae, the main ionisation mechanism of the gas is not photo-ionisation from a central source, as in H II regions and PNe, but collisional ionisation by radiative shocks. Supernova remnants are characterised by strong synchrotron continuum emission in the radio band, but their optical spectra are relatively similar to those of other nebulae. The main differences reside in the line ratios that involve low ionisation lines (i.e. $[\text{S II}]\lambda\lambda 6717, 6731$, $[\text{N II}]\lambda 6584$), which are typical shock tracers, and the kinematics of the gas since the high expansion velocities are reflected in the profile and properties of the observed emission lines (e.g. [Maciel 2013](#)). These nebulae mostly trace the positions where supernova explosions occurred in the past. Knowing where, how and when supernovae exploded is fundamental to understanding the energetics and chemical evolution of the ISM since these episodes are one of its main sources of feedback and chemical enrichment.

In this work, we take advantage of the data from one of the main surveys carried out by the Physics at High Angular Resolution in Nearby Galaxies¹ (PHANGS) collaboration, the PHANGS-MUSE survey ([Emsellem et al. 2022](#)), to develop a new automatic algorithm that satisfies all these criteria. In Sect. 2, we give an overview of the traditional methods used to classify ionised nebulae in the literature. In Sect. 3, we describe the data we used for this work. Section 4 focuses on the creation of the nebula catalogue we use to develop the classification algorithm. In Sect. 5, we describe the classification process extensively, while in Sect. 6 we compare its result with the classification obtained by applying traditional classification criteria. In Sect. 7, we discuss some of the main properties of the catalogue and how it performs in some common applications. Finally, in Sect. 8 we provide a final summary of this work.

2. Traditional classification of nebulae

Identifying and correctly classifying nebulae is fundamental to measuring their properties and understanding their nature and origin. In our Galaxy and in a few nearby sources like the Magellanic Clouds, classifying these objects is relatively easy. Most of the nebulae we observe in these systems can be easily resolved in multiple different bands (e.g. optical, radio) with the current instrumentation. Thus, we can identify the ionising source powering them with reasonable confidence and study their spatially resolved properties. When we move to other galaxies, even to the closest ones (e.g. M31, M33), we soon lose both the ability to resolve most nebulae (except for the largest star-forming regions) and to perform their multiwavelength characterisation (most of them can be detected only in the optical and infrared bands). This makes the classification of single nebulae extremely challenging. Nevertheless, several classification methods have been developed in the literature that leverage the main differences in the optical spectra of the nebulae (e.g. line ratios or luminosities) to define simple criteria that can be applied efficiently to large samples of nebulae.

Traditionally, the search for ionised nebulae has been performed using narrow-band images, that is, images acquired with filters characterised by a narrow pass-band ($< 100\text{ \AA}$) centred on

¹ www.phangs.org

specific emission lines (e.g. [Jacoby 1989](#); [Ciardullo et al. 2002](#); [Kennicutt et al. 2003, 2008](#)). This method allows one to cover a large field of view, arcminutes or even degrees in diameter depending on the instrument, and analyse many nebulae simultaneously. However, it is highly time-consuming because of the inherent long exposure time needed to get good signal-to-noise in narrow-band images, and the large overheads caused by the acquisition of off-band images needed to remove the contribution of the underlying stellar continuum. Consequently, only a small subset of lines is typically observed when studying ionised nebulae, the two or three lines where the specific type of nebulae the authors are interested in is brighter. This strongly influenced the characteristics of many classification criteria commonly used in the literature.

2.1. Planetary nebulae

Originally, extragalactic PNe were identified by blinking photographic plates acquired with different filters. One of these filters was centred on one or more lines ('on-band'), typically $[\text{O III}]\lambda 5007$, while the other one was observing a nearby region of the spectrum free of emission lines ('off-band'). Planetary nebulae would appear as unresolved objects in the on-band images but would be undetected in the off-band images (e.g. [Baade 1955](#); [Ford et al. 1973](#)). One of the first attempts to create simple, objective, criteria to classify ionised nebulae using their spectral properties is the one from [Sabbadin et al. \(1977\)](#). Using data from galactic nebulae, they define several regions in two different diagnostic diagrams that could be used to identify the nature of the considered nebulae (their Figs. 1 and 2). This criterion, revised recently by [Riesgo & López \(2006\)](#), has been used for decades, particularly for identifying PNe. These diagrams, however, require detecting a relatively large number of emission lines ($[\text{S II}]\lambda\lambda 6717, 6731$, $\text{H}\alpha$, $[\text{N II}]\lambda 6584$) which makes this method more suitable for studying bright galactic nebulae than faint extragalactic sources, especially when using narrow-band images. Therefore, the blinking technique continued to be used for many years until [Ciardullo et al. \(2002\)](#) finally developed a more objective way to separate PNe from other contaminants. While they still use the blinking technique to identify the first sample of nebulae, the new method could be used independently as long as precise measurements of the $[\text{O III}]\lambda 5007$ and $\text{H}\alpha$ lines are available for all the objects. Specifically, they consider as PNe all those sources that have $R = [\text{O III}]\lambda 5007 / (\text{H}\alpha + [\text{N II}]\lambda 6584) > 1.6^2$. Later, [Herrmann et al. \(2008\)](#) defined a new relation that identifies the parameter space in the R vs. $[\text{O III}]\lambda 5007$ absolute magnitude diagram where PNe are supposed to be located based on the work of [Ciardullo et al. \(2002\)](#). Although, it must be noted that this criterion is less restrictive than [Ciardullo et al. \(2002\)](#)'s one, and it could produce samples with rather high contamination. Despite this, it is one of the most commonly used criteria to identify PNe (e.g. [Kreckel et al. 2017](#); [Roth et al. 2021](#); [Galán-de Anta et al. 2021](#); [Scheuermann et al. 2022](#)) together with a more refined version of the on- and off-band blinking (e.g. [Bhattacharya et al. 2019](#); [Hartke et al. 2020](#)). Other classification criteria based on the BPT diagrams have been developed in the literature, for example [Frew & Parker \(2010\)](#), but even though they require flux measurements for more emission lines to be applied (which makes them more suitable for bright, galactic nebulae) they still do not ensure an exact classification (e.g. [Roth et al. 2021](#)).

² The authors argue that the proper ratio should be $R > 2$ but their $\text{H}\alpha$ images were not deep enough to reach such high ratios.

2.2. Supernova remnants

The vast majority of Galactic SNRs were historically detected via their radio emission (e.g., [Bolton et al. 1949](#); [Mills 1952](#); [Bennett 1963](#)). At optical wavelengths, they appear as faint and diffuse objects, and the heavy dust extinction of the Galactic disk makes them even more difficult to observe ([Green 1984](#); [Magnier et al. 1995](#)). However, the first searches for SNRs outside the Milky Way still focus on the radio emission to identify them. [Mathewson & Clarke \(1972\)](#) were the first who tried to define an optical criterion to isolate SNRs from H II regions, by noticing that, typically, SNRs have a $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha$ ratio which is relatively close to 1, while the $\text{H}\alpha$ line observed in H II regions is more or less an order of magnitude brighter than the $[\text{S II}]\lambda\lambda 6717, 6731$ doublet. While this is a poor quantitative criterion, it provided the framework for the development of more refined criteria in future works. The first quantitative definition of the $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha$ criterion comes from [D'odorico et al. \(1978\)](#). Based on the analysis of line ratios reported in the literature for both SNRs and H II regions, they classify as SNRs those nebulae with $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha \geq 0.6$ and as candidate SNRs other nebulae with $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha \geq 0.4$. A few years later, [D'odorico et al. \(1980\)](#) finally set $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha \geq 0.4$ as the criterion to distinguish between SNRs and H II regions in extragalactic environments.

Despite being more than 40 yr old, this classification criterion is still widely used for identifying extragalactic SNRs (e.g. [Long et al. 2010](#); [Lee & Lee 2014a](#); [Moumen et al. 2019](#)), but it is often associated with other criteria requiring, for example, a shell-like morphology of the nebulae or the absence of an obvious ionising source (i.e. a blue, hot star). The already mentioned [Frew & Parker \(2010\)](#) and [Sabbadin et al. \(1977\)](#) diagrams can be used to identify also SNRs, but since they require spectroscopic observations or more narrow-band imaging to produce the required diagnostic ratios, most of the search for SNRs performed in extragalactic environments still prefer to use the simpler $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha$ criterion.

Other narrow-band friendly classification criteria have been proposed in the literature. For example, [Fesen et al. \(1985\)](#) proposed new criteria based on the $[\text{O I}]\lambda 6300/\text{H}\beta$ or $[\text{O II}]\lambda 3727/\text{H}\beta$ ratios while [Boeshaar et al. \(1980\)](#) one based on the $[\text{S II}]\lambda\lambda 6717, 6731/[\text{Ar III}]\lambda 7136$ line. More recently, [Kopsacheili et al. \(2020\)](#) used shock models from [Allen et al. \(2008\)](#) and starburst models from [Kewley et al. \(2001\)](#) and [Levesque et al. \(2010\)](#) to study the best approach to the identification of SNRs. They show that while a multi-ratio approach outperforms single-ratio approaches, the $[\text{O I}]\lambda 6300/\text{H}\alpha$ is the single-ratio criterion that best discriminates between SNRs and other nebulae, while the $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha$ generally does not perform as well. Therefore, while some of these criteria seem to be more effective in discriminating SNRs from other regions than the $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha$ ratio, probably the faintness of the required lines ($[\text{O I}]\lambda 6300$, $[\text{Ar III}]$) or the need for precise extinction correction has prevented them from becoming particularly popular.

With the diffusion of integral field spectroscopy, however, it is becoming more clear that simple narrow-band imaging is not sufficient to correctly identify SNRs but that only a careful evaluation of the spectrum can produce reliable results. For example, [Long et al. \(2022\)](#) observationally confirm [Kopsacheili et al. \(2020\)](#) results, showing that while objects with $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha > 0.4$ are typically SNRs, a relatively large fraction of objects with properties compatible with those of SNRs (e.g. enhanced $[\text{O I}]\lambda 6300/\text{H}\alpha$, $[\text{N II}]\lambda 6584/\text{H}\alpha$,

[S III] λ 9068/[S II] λ 6717, 6731, broadened lines) are characterised by a [S II] λ 6717, 6731/H α < 0.4 and would be missed by this traditional criterion. On the other hand, some more complex diagrams involving several lines and the line velocity dispersion have been developed in the literature (e.g. [Davies et al. 2017](#); [D’Agostino et al. 2019a,b](#)), and confirmed that indeed, the velocity dispersion is critical in correctly identifying shock-ionised regions. However, most of these methods focus on identifying shock-ionised regions connected to galactic outflows or to the interaction between AGN jets and the ISM. Consequently, some of the criteria used are unsuitable for identifying SNRs. In particular, SNRs are expected to have enhanced gas velocity dispersion with respect to PNe and H II regions, but there is no consensus around an exact cut that can be used to distinguish between the different classes of nebulae because the observed velocity dispersion depends on the SNR age. The older the SNR is, the lower the expected velocity dispersion. High velocity dispersion objects are rarely observed in nearby galaxies (e.g. [Long et al. 2022](#); [Winkler et al. 2021](#)), so applying cuts at velocity dispersion $\sim 100 \text{ km s}^{-1}$ as in [Davies et al. \(2017\)](#) to identify SNRs would return only this young population of objects.

2.3. H II regions

H II regions are the most common nebulae in star-forming galaxies, with hundreds or thousands of nebulae detected in single galaxies (e.g. [Knapen 1998](#); [Bradley et al. 2006](#); [Azimlu et al. 2011](#); [Rousseau-Nepton et al. 2018](#); [Santoro et al. 2022](#), hereafter [S22](#)). They can span a wide range of sizes and luminosities, and they are typically both brighter and larger than most PNe and SNRs. Their size and brightness make them relatively easy to observe even in an extragalactic environment, and since their blended emission still appears as an H II region, they can also be observed in relatively distant objects where it is not possible to resolve the single nebulae (e.g. [Espinosa-Ponce et al. 2020](#)).

An emission-line-rich spectrum is characteristic of H II regions, with H α typically being the brightest line ([Osterbrock & Ferland 2006](#)). As a consequence, H α emission became one of the main H II region tracers, both in our galaxy (e.g. [Sharpless 1953, 1959](#); [Gum 1955](#)) and outside (e.g. [Baade & Arp 1964](#); [Hodge 1969](#); [Pellet et al. 1978](#)). Other detection techniques based on their radio or infrared emission exist and have been applied to search for H II regions in the Milky Way (e.g. [Reifenstein et al. 1970](#); [Lockman 1989](#); [Kuchar & Clark 1997](#); [Anderson et al. 2014](#)), but H α is still the main H II region tracer in other galaxies (e.g. [Azimlu et al. 2011](#); [Espinosa-Ponce et al. 2020](#); [S22](#)).

While H II regions are bright and easy to detect, for several reasons, no simple criterion based on a single line ratio has been defined yet for their identification. First of all, if we consider data with similar depth, H II regions dominate in numbers over PNe and SNRs in star-forming galaxies. For example, [Hodge et al. \(1999\)](#) found 2338 H II regions in M 33, while [Lee & Lee \(2014b\)](#) found only 199 SNR and [Magrini et al. \(2000\)](#) 134 PNe. Secondly, PNe emit most of their radiation through the [O III] λ 5007 line, with H α a factor of ~ 2 fainter than [O III] λ 5007, so they are relatively difficult to detect in the H α narrow-band images typically used to identify H II regions. Finally, both PNe and SNRs are typically fainter than most H II regions (see the different luminosity functions, e.g. [Lee & Lee 2014a](#); [Ciardullo et al. 2002](#); [Kennicutt et al. 1989](#)). They contaminate only the faint end of the H II region luminosity function (which is typically ignored when fitting and studying

this quantity; [S22](#)), and they are easily excluded from spectroscopically characterised samples of H II regions. So for most applications, cleaning H II region catalogues is not as essential as cleaning samples of SNRs or PNe.

Since H II regions can be relatively bright, it is possible to acquire high-quality spectra of these nebulae also in more distant galaxies. Therefore, it is possible to characterise the properties of these regions rather easily and confirm their classification, for example, by building traditional diagnostic diagrams (also known as BPT diagrams; [Baldwin et al. 1981](#); [Veilleux & Osterbrock 1987](#)) to identify the ionising mechanism of the nebulae. The relations from [Kewley et al. \(2001\)](#) and [Kauffmann et al. \(2003\)](#) are used to separate star-forming regions from other types of nebulae in these diagrams. These relations can be applied to spectra of both integrated sources (i.e. galaxies) and single nebulae. With the advent of integral field spectroscopy, this technique has been used to identify H II regions in several integral field surveys of galaxies ([López-Hernández et al. 2013](#); [Sánchez-Menguiano et al. 2018](#); [Kreckel et al. 2019](#); [Della Bruna et al. 2020](#); [Espinosa-Ponce et al. 2020](#); [Grasha et al. 2022](#); [S22](#)).

2.4. The need for a new classification scheme

As emphasised in the previous section, while such classification criteria have been used for decades, including in recent surveys (e.g. [Moumen et al. 2019](#); [Scheuermann et al. 2022](#); [S22](#)), they suffer from significant limitations.

The first thing to notice is that most of the criteria are defined to select a specific class of nebulae. This is a consequence of using narrow-band images to identify nebulae. Acquiring these images can be time-consuming and with large amounts of overheads, and it encourages the definition of criteria aimed at selecting a specific type of nebulae (requiring the observation of a small number of emission lines), more than classifying them comparatively. Some criteria that try to classify all the three classes of nebulae simultaneously exist (e.g. [Sabbadin et al. 1977](#); [Frew & Parker 2010](#)), but the need for a relatively large number of narrow-band images or spectroscopic follow-ups probably limited their popularity and detailed quantification of their performance.

Another characteristic of all criteria is that they are based on sharp limits, which can fail to correctly classify nebulae with properties close to the limits. For example, [Kopsacheili et al. \(2020\)](#) showed how there is an overlap between shock-ionised and photo-ionised regions and that the traditional [S II] λ 6717, 6731/H α criterion can both misclassify photo-ionised regions as shocks and vice versa. Similarly, the PNe classification criterion from [Herrmann et al. \(2008\)](#) and [Ciardullo et al. \(2002\)](#) identifies the region of a specific parameter space where PNe lives, but it does not exclude the presence of other types of nebulae. This also creates degeneracies between different classifications that must be solved in other ways. Also, faint regions with low S/N detections in the important lines can be easily misclassified by these sharp criteria, since relatively small errors in the flux measurements can result in different classifications. For this reasons, these criteria often require dedicated complex analysis or human interaction (e.g. morphological classification of barely resolved regions, association with faint ionising sources, etc.) to perform as expected, and are not suitable for the automatic analysis of a large number of sources such as the ones produced by modern integral field surveys (thousands or tens of thousands, e.g. [Rousseau-Nepton et al. 2018](#); [Sánchez-Menguiano et al. 2018](#); [Kreckel et al. 2019](#)). While, for obvious reasons, there is no overlap in the criteria

that can classify different types of nebulae at the same time (e.g. Sabbadin et al. 1977; Frew & Parker 2010), they still suffer from the problem of having sharp limits, and not being able to correctly classify nebulae at the edge of the classification criteria (Roth et al. 2021).

Finally, in the last decade, the advent of new integral field instruments capable of acquiring spatially resolved spectra of large fields of view, such as the Multi Unit Spectroscopic Explorer (MUSE, Bacon et al. 2010) and the Spectromètre Imageur à Transformée de Fourier pour l'Etude en Long et en Large de raies d'Emission (SITELE, Grandmont et al. 2012) changed our view of the ISM in nearby galaxies completely. With these instruments, it is possible to simultaneously observe properties of the nebulae that would require both imaging (position, size, morphology) and spectroscopic (multiple line fluxes, kinematics) observations. Despite this wealth of information, the classification of the nebulae is still primarily performed using the traditional classification methods (e.g. Sánchez-Menguiano et al. 2018; Moumen et al. 2019; Espinosa-Ponce et al. 2020; Della Bruna et al. 2020; Rhea et al. 2021; Scheuermann et al. 2022; S22) focusing on a single class of nebulae at a time.

An ideal classification algorithm should have a few key characteristics. First, it should be probabilistic. That is, it should return a probability associated with each classification. It should then take advantage of the wealth of information recovered from integral field units (IFU) data while maintaining a certain degree of flexibility (e.g. it should work with incomplete datasets). Finally, the algorithm should be automated and objective. This is essential to work on the large influx of data that is being produced by modern IFU surveys without the need for any human intervention.

3. Data

In this paper, we analyse 19 galaxies observed with MUSE (Bacon et al. 2010) as part of the PHANGS-MUSE ESO Large Program (P.I. Schinnerer; Emsellem et al. 2022). The galaxies in the sample are all nearby ($D < 20$ Mpc), star-forming and mildly inclined ($i < 75^\circ$), to minimise the effect of extinction and blending. Each galaxy was observed with a variable number of pointings (from 3 to 15), to map a significant fraction of its star-forming disk. All the targets were observed in wide field mode (WFM). All the data have been reduced using `pymusepipe`³, a python wrapper of the ESO data reduction pipeline (Weilbacher et al. 2020), specifically developed for the reduction of PHANGS data. Then, the reduced data have been combined, to produce a single mosaicked datacube for each target galaxy. In this work, we use a homogenised version of the mosaics, with constant point-spread-function both spatially and as a function of wavelength. All the details about the sample selection, the data reduction, convolution, and much more can be found in Emsellem et al. (2022).

The fully reduced data are then processed by the data analysis pipeline (DAP) to extract high-level data products, such as continuum-subtracted cubes, two-dimensional maps of line properties (flux, velocity, velocity dispersion), stellar mass, age, etc. In particular, the DAP is a python software based on the `gist` code (Bittner et al. 2019) that allows us to extract high-level data products (e.g. fluxes, kinematics) from the emission lines and stellar continuum observed in our data. The data analysis takes place in several steps, which are extensively described in Emsellem et al. (2022).

³ <https://github.com/emsellem/pymusepipe>

To compile our new catalogue of nebulae, however, we did not use the DAP emission-line maps because, by design, the DAP cannot return negative flux values when fitting emission lines. As a consequence, performing the fitting of a line where no emission is present results in positively biased fluxes which can simulate the emission of faint regions. These artefacts then can be identified by the segmentation algorithm when trying to identify faint objects. To minimise this issue we integrated the flux over a fixed wavelength window in a continuum-subtracted cube to create channel-integrated flux maps. While the continuum subtracted cubes are still somewhat connected with the Gaussian line fitting, especially in the Balmer lines where both the absorption and emission features are fitted simultaneously, this approach should depend much less on the results of the actual fit especially for lines not overlapping with strong absorption features like the [O III] λ 5007 line and the [S II] λ 6717, 6731 doublet.

We used the $H\alpha$ velocity maps produced by the DAP to estimate the reference wavelength of each line in every single spaxel. Since $H\alpha$ is detected in the vast majority of the observed area (Emsellem et al. 2022), the kinematics of the line is reliable enough for our goal, particularly considering that we define a large extraction window (500 km s^{-1}) centred on each line to allow for eventual differences in the line kinematics. In the case of the [S II] λ 6717, 6731 doublet, we position the extraction window at the average wavelength of the two lines, and we increase its size to extract both of them simultaneously to avoid cross-contamination of the line maps that may arise by their vicinity. Finally, for each line channel-integrated flux map, we also create an error map by summing in quadrature the errors extracted from the original datacubes in the considered wavelength range. In Sect. 4, we describe how we used these maps to produce the detection map we use to create the catalogue.

4. Catalogue creation

The first step towards the creation of our catalogue of classified nebulae is identifying all the nebulae we detect in our data. We use CLUMPFIND (Williams et al. 1994) to perform the first identification and to create a tentative segmentation map, which is then processed by a custom algorithm that rejects spurious detections and refines the segmentation.

4.1. CLUMPFIND analysis

CLUMPFIND is an algorithm developed by Williams et al. (1994) to detect emission regions in a variety of astronomical data. It uses a contour-based approach first to identify peaks, and then to connect pixels inside the same contour to the nearest peak in a similar fashion as what is done by the human eye. The algorithm is efficient at identifying peaks, especially in crowded areas, but it tends to overestimate sizes and to have a significant fraction of spurious detections caused by noise spikes. This motivates a post-processing of the CLUMPFIND output via a dedicated algorithm (see Sect. 4.2).

To proceed with the detection, CLUMPFIND needs two inputs: an image, and a configuration file containing all the segmentation parameters, including the definition of the contours that will be used during the analysis. To maximise the number of detected regions, we run the algorithm on a detection image, created by combining the [O III] λ 5007, $H\alpha$ and [S II] λ 6717, 6731 line channel-integrated flux maps described in Sect. 3. This particular combination of lines accounts for the fact that some of the nebulae we are looking for can be particularly bright in some lines while undetected in the others. The final detection image

is a weighted average of the three channel-integrated flux maps described above, where the weight of each pixel in each line flux map is defined by the square of its signal-to-noise ratio (S/N), that is the ratio between the line flux and its error according to the channel-integrated flux and error maps produced in the previous step. This allows us to boost the detection of regions emitting only in a sub-sample of the considered lines. We also extract the error map associated with the detection image by summing in quadrature the emission-line flux error maps weighted by the same weights used for the creation of the detection image, as indicated by standard error propagation techniques.

The second step is to define the contours that CLUMPFIND will consider for the segmentation process. To do so, we measure the average background level and the standard deviation of the detection image in areas without the evident presence of structured emission (e.g. nebulae, filaments). This is used to set up a lower limit (background + standard deviation) that defines the faintest contour considered by the algorithm. From this lower limit (reported in Table 1), we define logarithmically and evenly spaced contours to allow a good sampling of the faintest regions, without crowding the brightest peaks with too many contours. At this point, we run CLUMPFIND on all the galaxies of the sample. We use the version of the algorithm implemented in the CUPID⁴ software package (Currie et al. 2014; Berry et al. 2007), in legacy mode (to behave like the original CLUMPFIND IDL algorithm from Williams et al. 1994), and with default options⁵.

4.2. Post-processing and catalogue creation

We run CLUMPFIND on all the 19 galaxies of the sample, producing the first tentative segmentation maps and region catalogues. However, because of the aggressive CLUMPFIND settings, this first version of the catalogues still contains a significant number of spurious detections. Moreover, most of the detected regions are overgrown, because, by construction, the algorithm assigns every pixel inside the lowest contour to one of the regions. To address these issues, we developed a post-processing algorithm that cleans the spurious detections from the CLUMPFIND catalogues and assigns a more realistic size to the final regions.

The algorithm uses CLUMPFIND catalogues and segmentation maps to create one-dimensional surface-brightness profiles for each region centred on their CLUMPFIND peak, with these profiles then fitted with a Gaussian superimposed on a constant background. Most regions do not have a Gaussian profile, since they are resolved. However, a Gaussian fit allows us to extract information like a first guess for its FWHM that will be used later in the rejection criteria. The background is an estimate of the local contribution of the diffuse ionised gas (DIG) and the contamination from nearby nebulae.

At this point, the algorithm estimates the new boundaries of the regions. It first subtracts from each region's pixels the constant background estimated by the fit. Then, it computes the cumulative sum of the remaining flux for each region and identifies the isophote containing 90% of this flux of the considered nebula, that then becomes the new region boundary (Fig. 1b). We also pass the new regions through an algorithm that fills "holes" in the regions created by the new definition of the boundaries.

Finally, several criteria are evaluated to reduce the number of false detections included in the final catalogue. First, the region's total S/N measured from the detection image and associated

error map is evaluated, and regions with $S/N < 5$ are rejected. Then, all regions with an area < 12 pix or $FWHM < 1$ pix are rejected as spurious detections. Regions where the surface brightness cumulative distribution does not grow monotonically as a function of distance from the peak are also considered spurious detections and rejected. Finally, we reject all regions whose peak flux is $< 5 \times rms$, where the rms is the standard deviation of the detection image background as computed in Sect. 4.1, all regions with a $FWHM > 50\%$ of the size of the one-dimensional profile (typically diffuse, filamentary regions) and all regions that overlap with the stellar masks as defined in Emsellem et al. (2022).

All these criteria, including the 90% flux limit used to redefine the regions boundaries, have been carefully tuned by eye, by comparing the final segmentation map of each galaxy with the detection image and with a colour image obtained associating each channel-integrated line map to a different colour (blue for the $[O\ III]\lambda 5007$ line, green for $H\alpha$, red for the $[S\ II]\lambda\lambda 6717, 6731$ doublet). The colour information significantly simplifies distinguishing real emission from spurious detections. All the applied limits are the result of a compromise between rejecting the fake detections returned by CLUMPFIND, while retaining as many regions that the human eye can pick up as possible. While these criteria to define the boundaries of the regions are somewhat arbitrary, currently no clear agreement exists in the literature on how best to define the boundary of an ionised gas nebula in external galaxies. The flattening of the $H\alpha$ surface brightness profile of each nebula is used to define the boundary of the regions by common segmentation algorithms like HIIPhot (Thilker et al. 2000), while other works experiment with using the $[S\ II]\lambda\lambda 6717, 6731/H\alpha$ ratio (e.g. Kreckel et al. 2016; Della Bruna et al. 2020) or the $[S\ II]\lambda\lambda 6717, 6731/[O\ III]\lambda 5007$ ratio (Pellegrini et al. 2012) to distinguish between H II regions and the DIG. Moreover, for density-bounded regions a clear boundary does not simply exist.

The output of the detection code is a segmentation map (see Fig. 1a) and a catalogue with basic information, such as the position of the region peak, the flux-weighted centre, the area (in pixels) covered by the region, its maximum radius, and its circularised radius. The total number of detected nebulae across our sample of 19 galaxies is 40920. On average, we detect ~ 2150 nebulae per galaxy, with a minimum of 853 in NGC 7496, which is one of the less massive targets and the galaxy with the smallest number of observed pointings (3) and a maximum of 4101 in NGC 628, one of the galaxies with the largest covered area. In Table 1 we report, in detail, the number of detected nebulae for all the observed galaxies.

We use the segmentation map to extract the one-dimensional spectrum of each nebula, which we then process with the DAP to recover the fluxes of the brightest lines that are available within the MUSE wavelength range and the gas kinematics. The lines included in the fit, together with their wavelength and the ID used in the catalogue, are listed in Table 2. The emission lines are divided into two different groups. The first group includes the hydrogen emission lines ($H\alpha$ and $H\beta$), and the second group includes all the remaining lines. The kinematics of the lines is tied within each group (to the $H\alpha$ line and the $[N\ II]\lambda 6584$ line, respectively) to improve the fit of the faintest ones, as described in Emsellem et al. (2022). Differently from what was done in that work, we do not separate the high and low ionisation metal lines but include them in the same group. In particular, we noticed that leaving free the kinematics of the $[O\ III]\lambda 5007$ line results in unrealistic velocity dispersions, too large with respect to the profile of the line observed in the spectra, for a significant

⁴ <http://starlink.eao.hawaii.edu/starlink/CUPID>

⁵ For the default values see the CLUMPFIND documentation

Table 1. Properties of the examined galaxies and the results of the detection algorithm.

Name	RA(J2000) hh:mm:ss.s	Dec(J2000) dd:mm:dd	Distance ^(a) Mpc	Scale pc arcsec ⁻¹	Stellar mass ^(b) log(M_{\odot})	SFR ^(b) M_{\odot} yr ⁻¹	Area ^(c) kpc ²	Min ^(d)	All ^(e)	H α ^(e)	OIII ^(e)	BPT ^(e)
IC 5332	23:34:27.5	-36:06:04	8.18	40	9.45	0.31	34	38	972	862	569	407
NGC 628	01:36:41.7	+15:47:01	9.84	48	10.24	1.73	89	62	3502	2813	1594	1067
NGC 1087	02:46:25.2	-00:29:55	15.85	77	9.88	1.26	126	22	1168	965	740	590
NGC 1300	03:19:41.0	-19:24:40	18.99	92	10.57	1.07	356	21	1707	1566	1061	806
NGC 1365	03:33:36.4	-36:08:25	19.57	95	10.82	16.35	409	20	2351	1784	1341	901
NGC 1385	03:37:28.6	-24:30:04	17.22	83	9.94	2.01	101	29	1138	905	726	577
NGC 1433	03:42:01.5	-47:13:19	12.11	59	10.41	0.42	426	17	2768	2451	1651	939
NGC 1512	04:03:54.1	-43:20:55	17.13	83	10.58	0.79	266	11	1137	947	743	437
NGC 1566	04:20:00.4	-54:56:17	17.69	85	10.66	4.36	208	46	3274	2444	1765	1145
NGC 1672	04:45:42.5	-59:14:50	19.40	94	10.64	7.39	250	41	1953	1470	1030	768
NGC 2835	09:17:52.9	-22:21:17	12.38	60	9.85	1.18	87	45	1489	1122	858	666
NGC 3351	10:43:57.8	+11:42:13	9.96	48	10.27	1.17	73	31	1820	1565	973	526
NGC 3627	11:20:15.0	+12:59:29	11.32	55	10.73	3.66	85	39	2105	1483	1024	590
NGC 4254	12:18:49.6	+14:24:59	13.0	63	10.30	2.97	169	21	3536	2635	2028	1466
NGC 4303	12:21:54.9	+04:28:25	16.99	82	10.64	5.14	214	40	3756	2697	1683	1231
NGC 4321	12:22:54.9	+15:49:20	15.21	74	10.71	3.48	191	27	2858	2016	1265	814
NGC 4535	12:34:20.3	+08:11:53	15.77	76	10.49	2.04	124	52	2551	2145	1158	715
NGC 5068	13:18:54.7	-21:02:19	5.16	25	9.36	0.28	23	55	1974	1612	1219	955
NGC 7496	23:09:47.3	-43:25:40	18.72	91	9.84	2.08	92	33	861	728	468	377
Total									40920	32210	21896	14977

Notes. ^(a)Distances from [Anand et al. \(2021\)](#). ^(b)From [Leroy et al. \(2021\)](#). ^(c)Projected area covered by the survey from [Belfiore et al. \(2022\)](#). ^(d)Threshold used by CLUMPFIND to create the first segmentation map in 10^{-20} erg/(cm² s Å). ^(e)Samples as defined in Sect. 6.

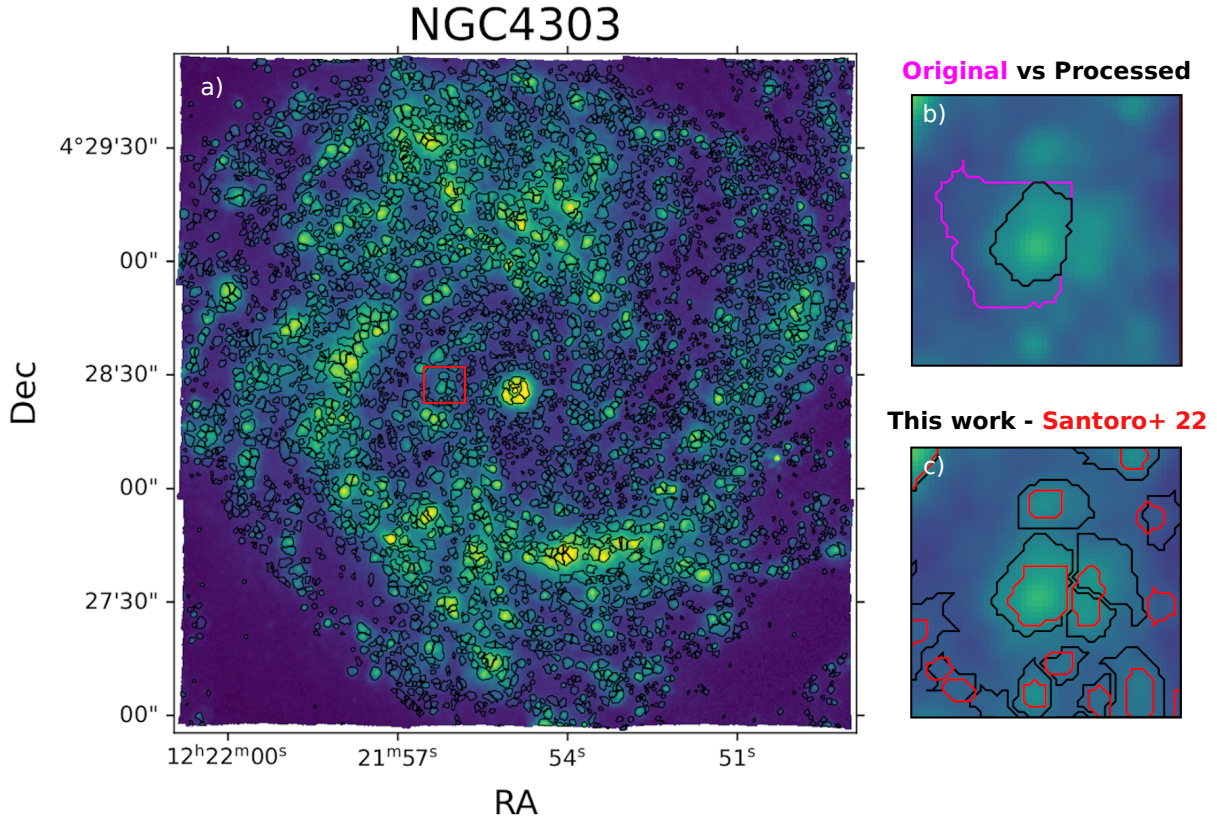


Fig. 1. NGC4303 segmentation map. Panel a: Map of the S/N-weighted combination of the [O III] λ 5007, H α and [S II] λ 6717, 6731 lines (our detection image) of NGC 4303 used for the identification of the nebulae with the final segmentation map (after the post-processing), shown in black contours. The red box highlights the area shown in panels b and c. Panel b: Zoom-in on a small region of the galaxy showing the difference between the original CLUMPFIND region boundaries (in magenta) and those that result after the post-processing described in Sect. 4.2 (in black, only one region is shown). Panel c: Zoom-in on the same area of panel b showing the difference between our new segmentation map (black) and that produced by S22 (red). As is clear in this panel we find that our regions are larger on average, but also the segmentation for some regions is different.

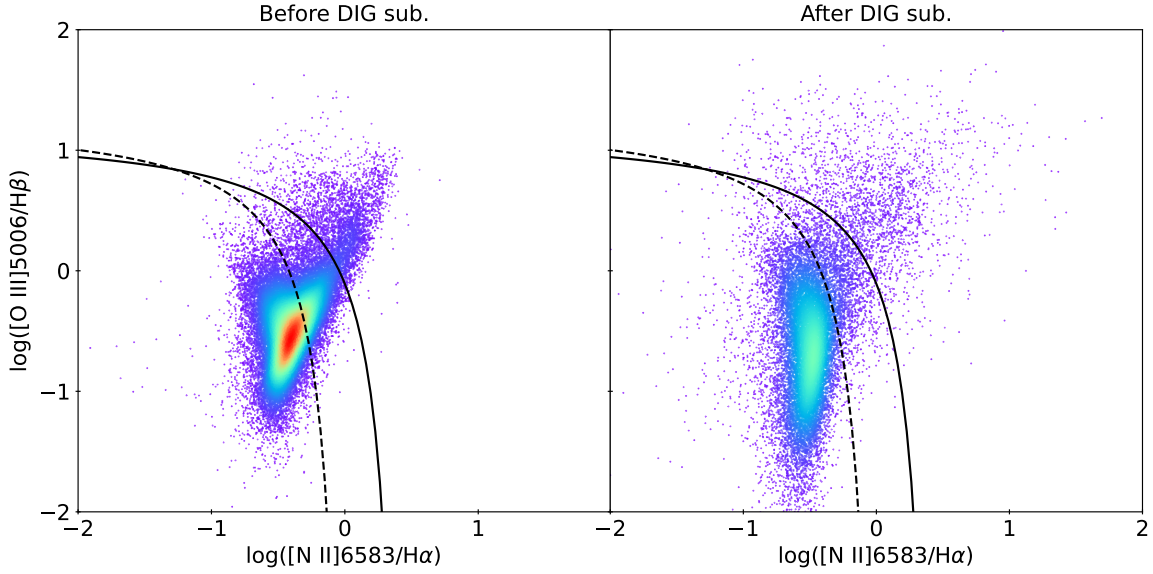


Fig. 2. $\log([\text{N II}]/\text{H}\alpha)$ vs. $\log([\text{O III}]/\text{H}\beta)$ diagram of the nebulae before (left) and after (right) the correction for DIG contamination. Only nebulae where $\text{H}\alpha$, $\text{H}\beta$, $[\text{O III}]\lambda 5007$ and $[\text{N II}]\lambda 6584$ are detected at 3σ after the DIG correction are plotted in the right panel. The colour map qualitatively shows how the points are concentrated. The solid and dashed black lines are the relation from Kewley et al. (2001) and Kauffmann et al. (2003) which separate star-forming regions from regions ionised by other ionising sources.

Table 2. Emission lines included in the catalogue.

Name	Wavelength (Å)	ID
$\text{H}\beta$	4861.35	HB4861
$[\text{O III}]\lambda 5007$	5006.84	OIII5006
$\text{H}\alpha$	6562.79	HA6562
$[\text{N II}]\lambda 6584$	6583.45	NI6584
$[\text{S II}]\lambda 6717$	6716.44	SII6716
$[\text{S II}]\lambda 6731$	6730.82	SII6730
$[\text{O I}]\lambda 6300$	6300.30	OI6300
$[\text{S III}]\lambda 9068$	9068.60	SIII9068

Notes. Wavelengths are recovered from the National Institute of Standards and Technology (NIST) https://physics.nist.gov/PhysRefData/ASD/lines_form.html

number of faint regions. Fixing the kinematic of the $[\text{O III}]\lambda 5007$ to the other metal lines does not significantly influence the fit in high S/N spectra, but it improves the fit in low S/N spectra.

4.3. DIG correction

In all galaxies, there is a warm, low-density component of ionised gas, called DIG, which is distributed across the whole galactic disk. While the ionisation source of this gas is still a matter of debate (e.g. Belfiore et al. 2022), it has been proven that the DIG can contribute a significant fraction to the total line emission in some galaxies (between 10 and 60% of the total $\text{H}\alpha$ emission of a galaxy, Thilker et al. 2002; Oey et al. 2007; Blanc et al. 2009).

This diffuse emission has a larger scale height with respect to the distribution of ionised nebulae (Reynolds 1997) and it can contaminate our integrated regions due to projection effects. Consequently, when measuring the emission lines flux in nebulae such as H II regions, they will be contaminated by the DIG emission. The level of contamination depends on several factors, for example, the relative strength of the nebula emission with respect to the DIG. The brighter the nebula is with respect to the DIG, the less important the contamination is, and vice versa.

The main problem is that the DIG spectrum is characterised by enhanced low-ionisation line emission (e.g. Reynolds 1985; Hoopes et al. 1996; Otte & Dettmar 1999) with respect to the spectrum of other nebulae like H II regions and PNe. As a consequence, the DIG emission can significantly alter the line ratios of the observed nebulae resulting in the likely misclassification of DIG-dominated regions if its contamination is not taken into account during the classification procedure.

To correct the spectra of our sources for DIG emission, we use the following procedure. First, for each region, we use a Gaussian filter with a standard deviation equal to the circularised radius of the nebula to increase the size of the nebula itself. Then, the area covered by the original nebula is masked, and we recover an annulus whose size depends on the circularised radius of the original nebula. We further mask any overlap between this annulus and the neighbouring regions. The algorithm extracts the DIG spectrum associated with the region using the annulus and feeds it to the same spectral extracting algorithm used to measure the line fluxes of the nebulae. Finally, the correction is applied by subtracting the estimated DIG flux of each line from the regions' flux, after applying a rescaling factor to consider the difference in size between the nebula and the associated annulus. In some cases, the estimated DIG emission is larger than the flux emitted by the considered region. This happens mostly for faint regions in crowded locations, where the DIG varies strongly. These lines are considered undetected, and when needed in the following analysis their flux is substituted by an upper limit after conducting a formal error propagation of the nebula and DIG emission line fluxes.

Figure 2 shows a comparison between the $[\text{O III}]\lambda 5007/\text{H}\beta$ vs. $[\text{N II}]\lambda 6584/\text{H}\alpha$ diagnostic diagram before and after applying the DIG correction. It is possible to see how the distribution of the points significantly changes after applying a DIG correction to our catalogue of nebulae. Before subtracting the DIG contribution, the points are relatively clustered close to the centre of the plot, following the distribution typically observed in diagnostic diagrams when comparing the properties of different types of sources (e.g. Kewley et al. 2001), with a branch of nebulae extending from the core of the distribution towards the

upper right (LINER-like) region of the diagram. After the DIG correction, the distribution of points appears to be less concentrated, with a larger scatter. The bulk of the points seem to follow a sequence that runs parallel to the delimitation line and is consistent with a dominant H II region population. Only regions where all lines necessary to compute the diagnostic ratios ($H\alpha$, $H\beta$, $[O III]\lambda 5007$, $[N II]\lambda 6584$) are detected at a 3σ level after the DIG correction ($\sim 46\%$ of the sample) are plotted in the right panel, so the scatter is real, and not only a consequence of low S/N regions. Something interesting to notice is that the prominent structure of points connecting the star-forming region (on the bottom/left of the [Kauffmann et al. 2003](#) line) to the region of the diagram typically associated with AGN, LINER and shock emission is not as prominent after the DIG correction. This is the typical location where DIG-dominated regions are found (e.g. Fig. 15 of [Belfiore et al. 2022](#)). Its absence, therefore, means that we are successfully removing DIG contamination from the spectra of the nebulae, and it highlights the importance of this step prior to the classification of individual nebulae.

5. Classification

Here, we develop a new classification algorithm based on comparing the regions' morphological and spectral characteristics with models of nebular properties using the principle of the odds ratio. This new algorithm is designed to overcome most of the issues of the traditional classification criteria: (1) it does a comparative classification, (2) it associates a probability to each classification, (3) it consistently considers non-detections and upper limits, (4) it can take advantage of the plethora of morphological and spectroscopic information that can be recovered from modern high spatial resolution integral field spectroscopic datasets.

5.1. Odds ratio

Our classification algorithm is based on the concept of the *odds ratio*. According to Bayesian statistics, the odds ratio is defined as:

$$O_{ij} = \frac{p(M_i, |I) p(D, |M_i, I)}{p(M_j, |I) p(D, |M_j, I)}, \quad (1)$$

where $p(M_i, |I)$ are the priors on the model i and $p(D, |M_i, I)$ is the global likelihood of the data D given the model i ([Gregory 2005](#)), also called the 'evidence'. When comparing two models to the same data and given some priors on them, if we can estimate the evidence of each model, Eq. (1) gives us a way to determine which one is better describing the data. The odds ratio also works when comparing more than two models. In this case, the probability of each model can be recovered with the following equation:

$$p(M_i, |D, I) = \frac{O_{i1}}{\sum_{j=1}^{N_{\text{mod}}} O_{ij}}, \quad (2)$$

where $p(M_i, |D, I)$ is the probability of model i given the data and the priors on the data, and O_{i1} is the odds ratio between model i and model 1, which is assumed as a reference.

The main problem with this approach is the computation of the evidence. Standard model fitting techniques like Markov-chain Monte Carlo (MCMC) methods are extremely powerful when used for parameter estimation, but they are not designed for evidence estimation. An alternative method, that has been specifically developed to perform well both for parameter estimation

and for evidence estimation, is 'nested sampling' ([Skilling 2004, 2006](#)). Based on the concept of 'typical sets', this algorithm estimates the evidence and the posterior simultaneously by integrating the prior in nested shells of constant likelihood.

5.2. Models

To classify our sources, we want to compare their spectra to different families of models, each one of them associated with a different class of nebulae. As explained in Sect. 1, almost all the emission-line regions belong to one of three classes of sources: H II regions, PNe and SNRs. To perform our classification, we need to define models corresponding to each class.

We use two-parameter grids of standard photo-ionisation or shock models, that adequately represent different families of objects. While these models suffer from limitations associated with the simplistic assumptions they adopt (e.g. simple geometries, homogeneity in the gas distribution, reduction of extra secondary parameters), they are sufficient for our purposes as we are not interested in inferring the physical properties of individual nebulae at this stage, but instead just assigning a classification. Moreover, the larger the model grid, the more computationally expensive is the fitting procedure.

To represent H II regions, we use the models from [Pérez-Montero \(2014\)](#). These models are a small grid developed to study chemical abundance in H II regions using direct methods. The grid depends on a total of three parameters (ionisation parameter, metallicity, N/O ratio), and there are two versions of it, one where the C abundance is linked to the O abundance and the second where the C is connected to the N one. In this work, we use the grid with the C abundance tied to the O abundance, and a fixed $\log(N/O) = -0.75$. We selected these values by comparing the grids with the BPT diagrams ([Baldwin et al. 1981](#)) and identifying the grid that was best covering the H II regions area of the BPT diagrams. Table 3 reports a summary of the selected grid parameters.

Planetary nebulae are represented by the models from [Delgado-Inglada et al. \(2014\)](#). These are a large grid of models that have been produced to study the ionisation correction factor of several ions in PNe. The grid depends on eight parameters (spectral energy distribution shape, density law, metallicity, dust depletion, temperature of the ionising source, hydrogen density, stellar luminosity and inner radius of the nebula). As done for H II regions, we consider only a sub-sample of the full grid of models. The details of the parameters of the selected grid are reported in Table 3. We selected the more realistic models produced with stellar atmospheres from [Rauch \(2003\)](#) and including dust depletion. Abundances are the default CLOUDY solar abundances since they are the only ones with which the models are provided. For the inner radius of the nebula and the gas density, we selected the intermediate values between the set provided in the grids. This left as our free parameters the temperature and luminosity of the ionising star.

Finally, we select the fast shock models by [Allen et al. \(2008\)](#) to represent SNRs. These are a set of simulations of shocks produced using the MAPPINGS photo-ionisation and shock modelling code typically used to study SNRs (e.g. [Dopita et al. 2010; Sabin et al. 2013; Blair et al. 2014; Micelotta et al. 2016; Kopsacheili et al. 2020](#)). As we did for the H II regions and PNe grids, we selected a sub-sample of the available models representing standard conditions found in SNRs (e.g. [Dopita et al. 2010](#)). One problem with these models is that they have been extensively used in the literature to also describe the emission of the narrow-line region and extended narrow-line region of AGN

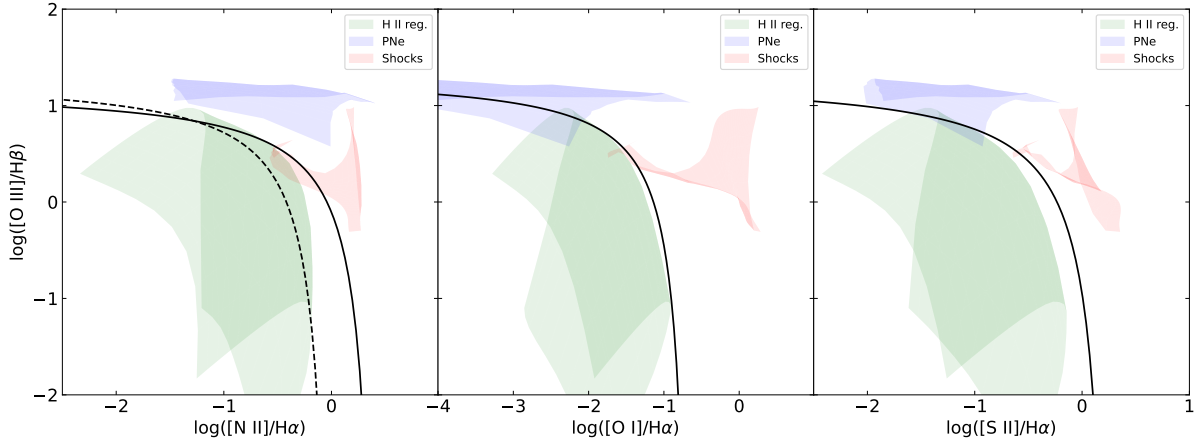


Fig. 3. Line ratios from the models used for classifying the nebulae in the catalogue plotted in the traditional diagnostic diagrams from Baldwin et al. (1981). Black solid lines represent the relations from Kewley et al. (2001) which separates star-forming regions from other kinds of sources. The black dashed line represents the limit for pure star formation identified by Kauffmann et al. (2003).

Table 3. Summary of the model grids used by the classification algorithm.

H II region	
Reference	Pérez-Montero (2014)
Free parameters	Ionisation par., $12 + \log(\text{O}/\text{H})$
C tied to	O
$\log(\text{N}/\text{O})$	-0.75
Planetary Nebulae	
Reference	Delgado-Inglada et al. (2014)
Free parameters	Luminosity, temperature
Spectral energy distribution	Rauch (TR)
Dust depletion	Yes
Density	3000 cm^{-3}
Inner radius	$1 \times 10^{17} \text{ cm}$
Model set	Matter bounded (M40)
Abundances	8.94 ($12 + \log(\text{O}/\text{H})$)
Shock	
Reference	Allen et al. (2008)
Free parameters	Magnetic field, shock velocity
Model set	shock w/o precursor
Pre-shock density	1 cm^{-3}
Abundances	Solar (Anders & Grevesse 1989)

Notes. For each grid, we report the reference, the free parameters defining the grids, and the fixed parameter that we used to reduce the dimension of the grid.

(e.g. Allen et al. 2008; Schlesinger et al. 2009; Holt et al. 2009; Sarzi et al. 2010, and many others). While most of our galaxies are purely star-forming galaxies, a few objects are known to host an AGN in their nucleus (e.g. NGC 1365). So, it might be possible that some of the regions classified using these models are AGN ionised regions and not real SNRs. For this reason, we decided to identify the nebulae classified by these models as shocks or shock-ionised nebulae, and we leave to Sect. 7.7 a discussion of how other types of nebulae are contaminating our sample of SNRs.

The Delgado-Inglada et al. (2014) and Pérez-Montero (2014) models have been downloaded from the Mexican Million Model

database⁶ (3MdB). They are a rerun of the original grids performed with Cloudy v13 (Ferland et al. 2013) and Cloudy v17.01 (Ferland et al. 2017) respectively. The Allen et al. (2008) models are downloaded from the shock section of the 3MdB⁷, and they are a rerun of the original grid with MAPPINGS V (Sutherland et al. 2018). Figure 3 shows how the models cover the traditional diagnostic diagrams for classifying ionised nebulae (Baldwin et al. 1981).

5.3. Model fitting with IZI

The classification process happens in two steps. First, we compare the observations to the different grids and compute the evidence using the nested sampling algorithm. Second, we use the evidence to compute the odds ratio and the probability for each region to belong to a specific class via Eqs. (1) and (2).

To compare the regions to the models and compute the evidence, we modified the python version of IZI (Blanc et al. 2015) developed by Mingozi et al. (2020)⁸ to allow for the computation of the evidence required for the model comparison process.

The original IZI code was designed to estimate the metallicity and ionisation parameter of H II regions by doing a full likelihood calculation over the parameter space, given a photoionisation model grid and a set of observed extinction-corrected line fluxes and their associated errors. Mingozi et al. (2020) added the possibility to use a more efficient Markov chain Monte Carlo (MCMC) algorithm instead of the full likelihood calculation and to fit the amount of dust extinction, modelled as a foreground screen with a fixed extinction law and parameterised by a reddening $E(B - V)$ parameter following the original idea from Brinchmann et al. (2004). While MCMC is not ideal for computing evidences, IZI provides all the infrastructure we need to compare observations to models, including a self-consistent treatment of errors and upper limits.

To estimate the evidence, however, we changed the core fitting method from MCMC to nested sampling. In particular, we use the implementation provided by the dynesty⁹ python package (Speagle 2020) which can simultaneously estimate both the

⁶ <https://sites.google.com/site/mexicanmillionmodels>

⁷ <http://3mdb.astro.unam.mx:3686>

⁸ https://github.com/francbelf/python_izi

⁹ <https://dynesty.readthedocs.io>

evidence and the posterior PDF, allowing the software to be used for parameter estimation. Another advantage of *dynesty* is that it can automatically decide when the algorithm converges based on objective parameters, eliminating the uncertainties connected to the long time that an MCMC could take to reach convergence and to the arbitrary selection of a proper burn-in phase. However, nested sampling must fully sample the prior volume, resulting in a loss of efficiency. The change from MCMC to nested sampling required further modifications to the code, mainly related to how nested sampling (and *dynesty* in particular) treats priors with respect to other commonly used MCMC-based algorithms. Finally, we include the possibility of using model grids that depend on whatever couple of parameters, not only on the metallicity and ionisation parameter as in the original software. This code version will not be released to the public yet because we are planning a complete refactoring of IZI, which will include these and other updates (e.g. the possibility to use n-dimensional grids). The final version of the code will be released as a self-consistent pip package, together with an upcoming publication focused on the physical properties of the nebulae classified in this work.

5.4. Comparison with the models

Our updated version of IZI compares the properties of each region to the prediction of the grids of models previously described and to prior information regarding the expected morphology and kinematics of different types of nebulae, to compute the evidence associated with each considered class. In our case, we consider the principal lines used in the traditional diagnostic diagrams ($H\alpha$, $H\beta$, $[O\ III]\lambda 5007$, $[O\ I]\lambda 6300$, $[N\ II]\lambda 6584$, $[S\ II]\lambda 6717$ and $[S\ II]\lambda 6731$; Baldwin et al. 1981; Veilleux & Osterbrock 1987). Both the observed and model fluxes are normalised to $H\beta$. The observed line fluxes are corrected for Galactic extinction using the Cardelli et al. (1989) extinction law, but not for internal extinction. The internal reddening is self-consistently estimated by IZI when matching the model to the data, by using a standard Cardelli et al. (1989) reddening law, as revised by O'Donnell (1994), with $R_V = 3.1$. As a consequence, for each family of models used during the classification, we estimate a separate value of the reddening ($E(B - V)$). A comparison between these estimates and the one recovered via the Balmer decrement is reported in Sect. 7.1.

We classify all the nebulae of the catalogue, without applying any cut related to the S/N of any line. We expect most of the regions to be detected with a $S/N > 3$ only in a few emission lines (mostly $H\alpha$), especially after the DIG correction. Moreover, sources like PNe are expected to be mostly detected in $[O\ III]\lambda 5007$, and a S/N cut would result in losing a significant fraction of these objects. All lines with $S/N < 3$ are considered as upper limits (with the associated error being considered as the upper limit), which IZI can treat appropriately during the likelihood and evidence calculations.

Line fluxes, however, are not the only quantities relevant for the classification of ionised nebulae. In particular, PNe are expected to be compact objects, with diameters of the order of 1 pc (e.g. Acker et al. 1992; Bojičić et al. 2021) and, therefore, unresolved in all PHANGS galaxies. Similarly, SNR can reach sizes of the order of 100 pc (Asvarov 2014). They are larger than PNe but smaller than many H II regions. Another quantity that is relevant to the classification is the velocity dispersion of the emission lines. H II regions and PNe are usually characterised by low-velocity dispersion (e.g. Richer 2006; Hajian et al. 2007; Schönberner et al. 2014; Della Bruna et al. 2020, 2021; Spriggs

et al. 2020) and unresolved in medium-resolution data such as the one provided by MUSE. On the other hand, SNRs are characterised by more complex kinematics, and they can show higher velocity dispersion (Points et al. 2019). We include the information on the size of the nebula and velocity dispersion of the gas as a prior during the calculation of the evidence. In particular, we require PNe to be unresolved and with a velocity dispersion $< 25\text{ km s}^{-1}$ (Richer 2006), shock ionised regions to be smaller than 100 pc, and H II regions to have a velocity dispersion $< 24\text{ km s}^{-1}$ (Della Bruna et al. 2020, 2021). The prior is uniform below the limit and has an exponential decay above it.

Once all of the regions have been compared to the three grids of models and the evidence has been computed, Eq. (2) was used to recover the probability for each region to belong to one of the three classes of sources we are considering in this work. The final classification of the nebulae is tied to these probabilities, but it is not simply determined by the class of nebulae with the highest probability. Only those sources with a probability $\geq 90\%$ in one of the three classes (H II regions, PNe and Shocks) are considered ‘classified’. Those sources where all probabilities are $< 90\%$ are labelled as ‘ambiguous’. This decision is driven by the fact that the odds ratio paradigm for model comparison is definitive (i.e. produce reliable classifications) only when the results are overwhelmingly in favour of one class with respect to all the others. Moreover, we expect the presence of regions where it was not possible to correctly deblend the contribution of nearby nebulae with different classification (e.g. H II regions with embedded or nearby SNRs). If both nebulae have similar brightness, that is one of the nebulae is not overwhelmingly brighter with respect to the other, we expect them to fall into the ambiguous class. Finally, there is a sub-sample of sources where the software did not converge in any of the three comparisons and, therefore, no probability has been computed. These sources are labelled as ‘unclassified’ and mainly consist of nebulae where no line was significantly detected. Table 4 summarises all the information included in the catalogue which is available through *vizier* and the Canadian Astronomy Data Centre (CDAC)¹⁰.

6. Analysis

In this section, we analyse the results of the detection and classification processes. We start by investigating how the distribution of the objects in the different classes changes as we perform different cuts to the sample to understand how different selections could bias the classification of a catalogue (Sect. 6.1). Then, we apply to the same catalogue some of the most common traditional classification criteria (Sect. 6.2), and we compare the results of the two different classification approaches to investigate similarities and differences (Sect. 6.3).

6.1. Reliability of the classification

As introduced in Sect. 5.4, we run the classification algorithm indiscriminately on all the regions included in our catalogue, independently of the number of detected lines. However, a lack of detection in certain lines can influence the final classification and how they distribute across the different classes. To study how the number of detected lines affects the classification of the nebulae, we create three sub-samples by requiring 3σ detections for different subsets of emission lines. Therefore, in the following analysis, we consider four sub-samples of nebulae: (a) full sample – no requirements on line detection significance; (b) the

¹⁰ <http://dx.doi.org/10.11570/23.0006>

Table 4. Summary of the information included in the released catalogue of classified nebulae.

Identification	
index	Unique identification number of the region
gal_name	Name of the galaxy hosting the region
region_ID	Identification number of the region within the galaxy
DIST	Distance of the galaxy in Mpc ^(a)
Position and morphology	
ra	Right ascension of the H α luminosity-weighted centre of the region
dec	Declination of the H α luminosity-weighted centre of the region
deproj_dist	Deprojected distance from the nucleus of the galaxy (in arcsec) ^(a)
deproj_phi	Deprojected position angle (0–2 π) ^(a)
region_area	Area of the region in pixels ^(b)
region_circ_rad	Circularised radius of the region in arcsec
Spectral properties	
lineid should be replaced with the line ID reported in Table 2	
lineid_FLUX	Measured line flux in 10 ⁻²⁰ erg cm ⁻² s ⁻¹
lineid_FLUX_ERR	Measured line flux error in 10 ⁻²⁰ erg cm ⁻² s ⁻¹
lineid_SNR	S/N of the line
lineid_FLUX_DIG	DIG corrected line flux in 10 ⁻²⁰ erg cm ⁻² s ⁻¹
lineid_FLUX_ERR_DIG	DIG corrected line flux error in 10 ⁻²⁰ erg cm ⁻² s ⁻¹
lineid_SNR_DIG	S/N of the line after the DIG correction
L_HA	H α luminosity in erg s ⁻¹
L_HA_DIG	DIG corrected H α luminosity in erg s ⁻¹
Kinematics	
The kinematics is reported only for the H α and [N II] λ 6584 lines.	
lineid_SIGMA	Velocity dispersion of the line in km s ⁻¹ (c)
lineid_SIGMA_ERR	Error on the velocity dispersion of the line in km s ⁻¹
lineid_VEL	Velocity of the line in km s ⁻¹
lineid_VEL_ERR	Error on the velocity of the line in km s ⁻¹
Classification	
z_hii	Evidence of the H II region models
z_pn	Evidence of the PNe models
z_sh	Evidence of the shock models
p_hii	Probability of the H II region models
p_pn	Probability of the PNe models
p_sh	Probability of the shock models
c	Final classification
Extinction	
ebv	Balmer decrement $E(B - V)$ from measured fluxes
ebv_dig	Balmer decrement $E(B - V)$ from DIG corrected fluxes
ebv_sh	$E(B - V)$ estimated by IZI using shock models
ebv_hii	$E(B - V)$ estimated by IZI using H II region models
ebv_pn	$E(B - V)$ estimated by IZI using PNe models

Notes. ^(a)Distances from Anand et al. (2021), ^(b)pixel size: 0.2 arcsec pix⁻¹, ^(c)corrected for instrumental broadening.

H α sample requires a 3 σ detection of H α ; (c) the OIII sample requires a 3 σ detection of the [O III] λ 5007 line; (d) the BPT sample requires a 3 σ detection of all the lines required to build BPT diagrams ([O III] λ 5007, H β , H α , [N II] λ 6584, [S II] λ 6717, [S II] λ 6731, [O I] λ 6300). Each sample's total number of regions is reported in Table 5.

The BPT sample is the most restrictive one. By design, it includes only the brightest and best-detected regions. The

[O III] λ 5007 line is relatively faint, so most of the regions in the OIII sub-sample are detected in several other lines (91% of the regions detected in [O III] λ 5007 are also detected in H α). As a consequence, this cut is expected to be the second most restrictive one. Finally, H α is the brightest line of the spectrum for the vast majority of ionised nebulae, so many objects in our catalogue are detected in H α even after removing the DIG contribution. Figure 4 shows the diagnostic diagrams of

Table 5. Distribution of the regions in the classes defined in Sect. 5.3 as a function of the considered sample.

Class	All	H α	[O III]	BPT
H II	29 986 (0.73)	23 939 (0.74)	14 313 (0.65)	10 993 (0.73)
PNe	796 (0.02)	535 (0.02)	775 (0.04)	32 (0.00)
Sh.	6336 (0.15)	4454 (0.14)	4181 (0.19)	2356 (0.16)
Amb.	3722 (0.09)	3222 (0.10)	2556 (0.12)	1546 (0.10)
Un.	80 (0.00)	60 (0.00)	71 (0.00)	47 (0.00)
Total	40 920	32 210	21 896	14 977

Notes. Next to the absolute number, we report the fraction of nebulae in that class with respect to the total.

the full catalogue and compares the classification of each region with their expected classification according to their location in the diagrams. Similar diagrams for the other sub-samples are reported in Figs. 5–7.

The full catalogue contains 40 920 regions, with the vast majority of them (29 986, 73% of the sample) classified as H II regions. As expected, these nebulae are mostly located in the diagnostic diagrams under (or close to) the Kewley et al. (2001) lines (Fig. 4, top row). However, there is a significant amount of points located in other regions of the diagrams (25%). The bulk of these ‘misplaced’ H II regions (96%) actually corresponds to objects that lack detections in one or more emission lines, and therefore are plotted as limits in the BPT diagrams. Comparing Fig. 4 with Fig. 5, we can see that the vast majority of these points (~6000 nebulae) disappear from the diagrams, meaning that their H α was not clearly detected. Removal of these nebulae with no significant H α detection gives 23 939 H II regions (or 74% of the H α sample).

The number of H II regions further decreases if we require the detection of the [O III] λ 5007 line, a relatively strict requirement. In fact, the number of nebulae reduces to 14 313 (65%). If we apply the final cut and look at the BPT sample, the number of H II regions continues to decrease, but they still represent the vast majority of the nebulae (10 993, 73% of the BPT sample). These are the better-detected sources, and they should have the more reliable classification. In fact, ~98% of them lie below the Kewley et al. (2001) lines, where we expect star-forming regions to be.

The second most common class of nebulae are shock-ionised ones (6336 regions, 15% of the sample). The diagnostic diagrams (Fig. 4, middle row) show that most of the regions classified as shocks overlap with the selected models, as expected. There are some points sparsely distributed across the diagrams, but, similarly to what is observed with H II regions, most of them are characterised by limits and, therefore, have a less reliable classification. When we require the detection of the H α line we recover only 4454 nebulae (14%), while requiring the detection of the [O III] λ 5007 line we get 4181 nebulae (19%). Finally, in the BPT sub-sample we have a smaller number of shock ionised regions (2356, 16%).

Planetary nebulae is the class with the lowest number of associated regions (796, 2%). However, they seem to be relatively well clustered at the top of the diagnostic diagrams. By requiring the detection of H α the number of nebulae slightly decreases (535, 2%), while, when we force the detection of the [O III] λ 5007 line, we practically recover the full sample (775, 4%). Finally, if we further strengthen the constraints (requiring the detection of all the BPT lines), only a handful of nebulae

are classified as PNe (32, <1%). This confirms that most PNe are faint objects detected only in a few lines. One of those lines, however, must be the [O III] λ 5007, since this is typically the brightest line in PNe spectra due to the hard ionising radiation field emitted by their central stars.

The number of sources in the ambiguous class varies from 3722 (9%) in the full catalogue to 1546 (10%) in the BPT sample. The ambiguity for most of these objects is between an H II region or a shock classification, with most of them having higher probabilities of being H II regions. Only a minority of them show a different combination of non-zero probabilities. In fact, the most prominent cloud of points is located close to the region of the diagnostic diagrams where the H II region and shock models overlap (Fig. 3 and Figs. 4–7, fourth row).

Finally, we have a few unclassified objects in the catalogue (80). This is a negligible amount of regions (<1%), and they are mostly located in the top right corner of the diagrams, in a region not covered by any grid of models. The percentage of regions in this class grows slightly when moving from the full sample to the BPT sample (from 0.1 of the full sample to 0.3% of the BPT sample) but it always stays a negligible amount with respect to the total number of considered regions. The reason why these regions are not classified is still not clear. While there might be some exotic objects in this class (e.g. background galaxies or AGN), most of them look like ordinary nebulae in a visual inspection.

6.2. Traditional classification criteria

As introduced in Sect. 1, our classification algorithm is only the most recent one in a list of methods that have been developed in the literature. The main novelty of the new algorithm is its use of models for a comparative approach to the classification of resolved nebulae, the assumption of a single ionisation mechanism for each nebula, the automatic inclusion of size and velocity dispersion in the procedure without any need for human interaction, and its probabilistic nature. A comparison between the new classification and the classification that would have been obtained applying the traditional classification algorithms is fundamental to understanding the differences between the approaches and their reliability.

To traditionally classify H II regions, we apply the criteria based on the BPT diagrams used by S22 and Groves et al. (2023) on a similar catalogue. More explicitly, we classify as H II region all the nebulae that are under the Kauffmann et al. (2003) line in the [O III] λ 5007/H β vs. [N II] λ 6584/H α diagram and under the Kewley et al. (2001) lines in the other two diagrams. All three criteria must be satisfied simultaneously for a region to be classified as an H II region. For PNe we apply the method developed by Ciardullo et al. (2002) and Herrmann et al. (2008) based on the relative strength of the [O III] λ 5007 line with respect to H α . All the nebulae classified as PNe must satisfy Eq. (3), where $M_{[\text{O III}]}$ is the [O III] λ 5007 absolute magnitude of the nebula computed starting from the apparent magnitude calibrated with Eq. (4) from Jacoby (1989) and the distances reported in Anand et al. (2021).

$$\log \frac{[\text{O III}]5006}{\text{H}\alpha + [\text{N II}]6583} > -0.37 \times M_{[\text{O III}]} - 1.16, \quad (3)$$

$$m_{[\text{O III}]} = -2.5 \log_{10} (I_{[\text{O III}]}) - 13.74. \quad (4)$$

We also require the nebulae to be unresolved, that is their apparent size should be smaller than the size of the PSF.

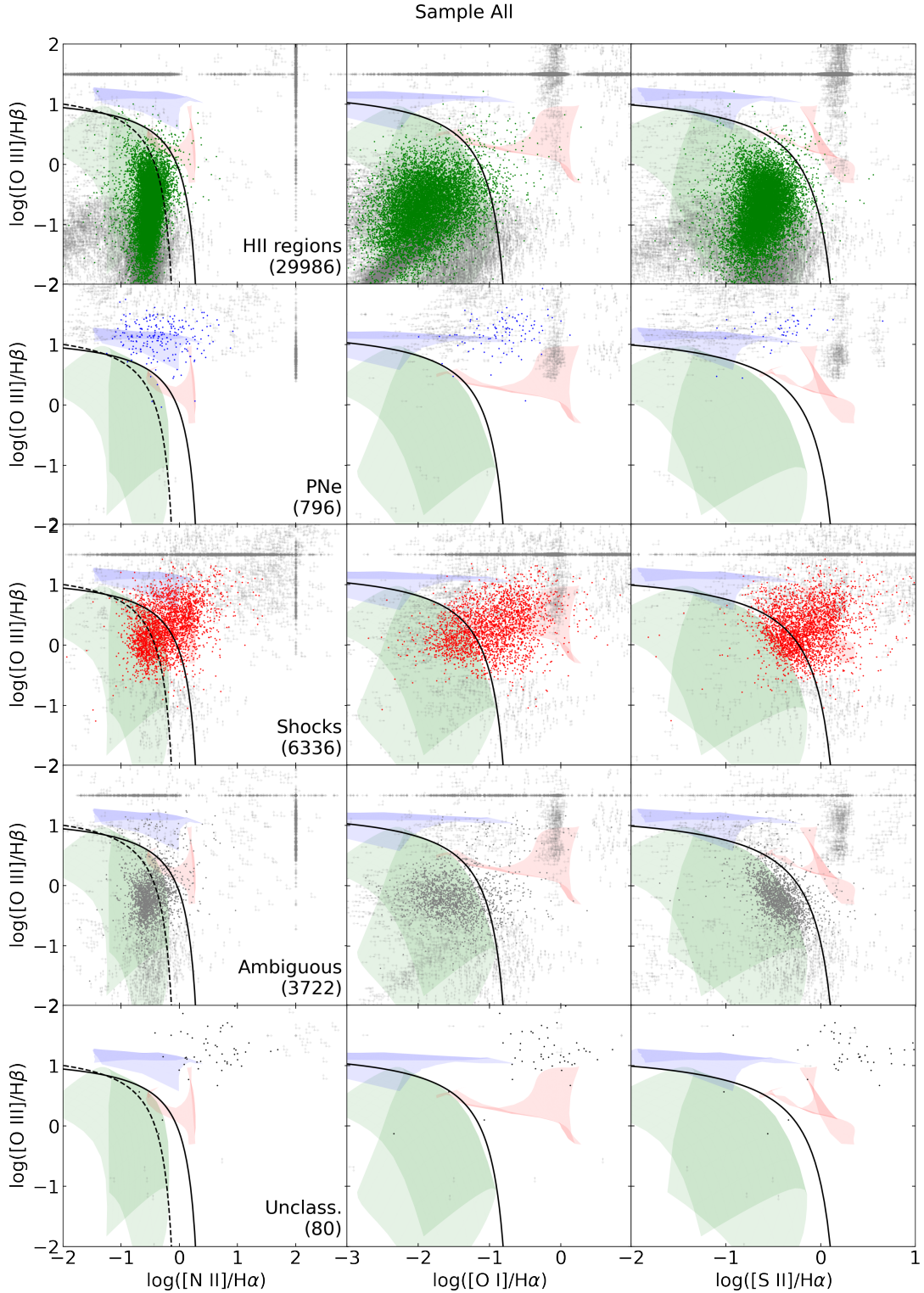


Fig. 4. Diagnostic diagrams from Baldwin et al. (1981) and Veilleux & Osterbrock (1987) for the full sample. Lines and shaded areas are defined as in Fig. 3. In each row only, from top to bottom, we show: H II regions (green), PNe (blue), shocks (red), ambiguous (grey) and unclassified (black) objects. Light grey points in each plot represent upper or lower limits.

Lastly, we classify as shocks (or SNRs) all the regions with $[S II]\lambda\lambda 6717, 6731/H\alpha > 0.4$, the most common criterion for identifying SNRs using narrow-band images (D’Odorico et al. 1980). If a region shows one or more limits among its ratios, we classify the object if its ratios and limits are consistent with a specific class.

Also in this case, we include the unclassified or ambiguously classified classes. The first class includes all those objects that are not picked up by any criterion, while we consider ambiguously classified all those sources that are selected by more than one criterion. We show in Table 6 the distribution of the regions classified with the traditional methods, while Figs. A.1–A.4

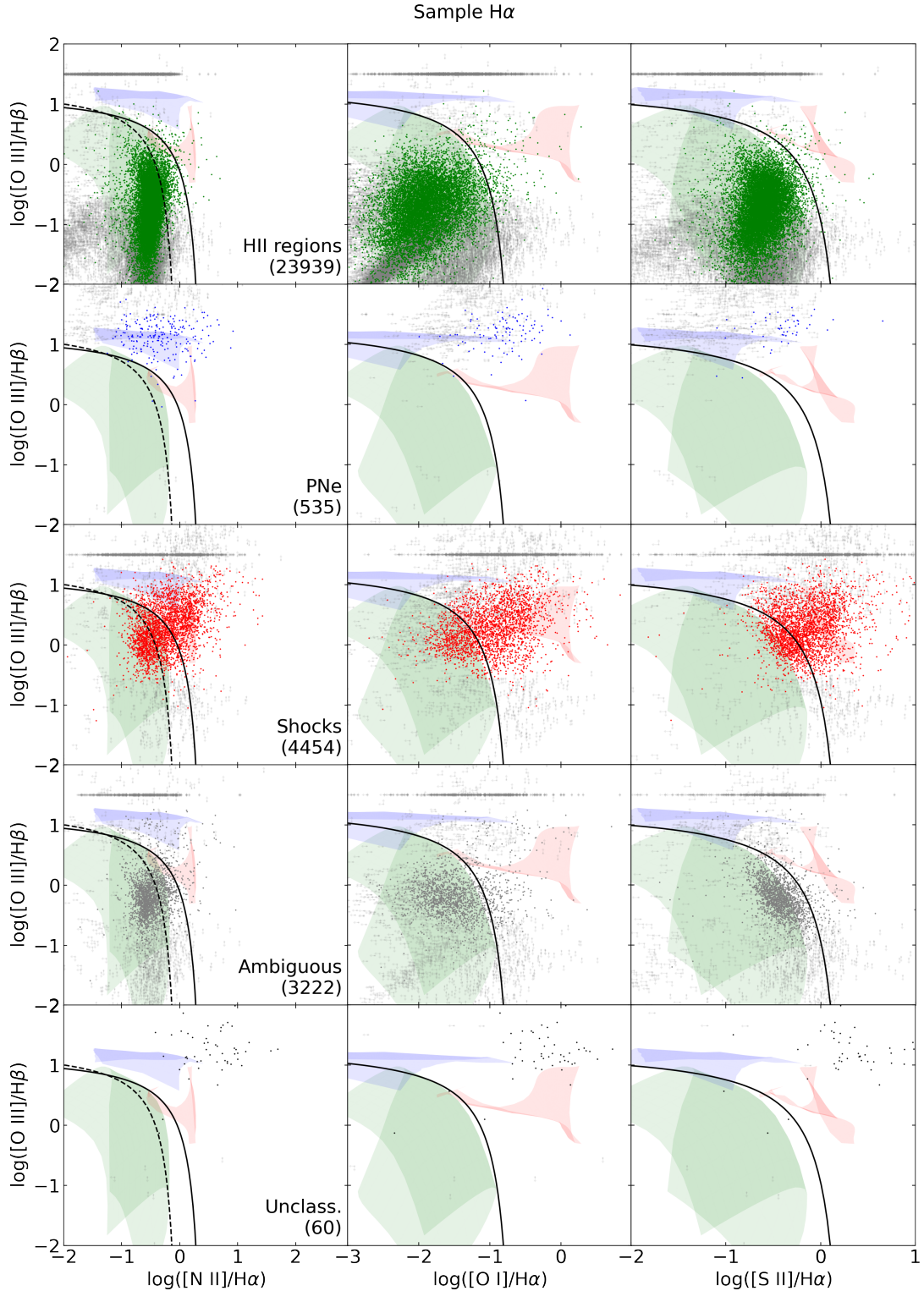


Fig. 5. Diagnostic diagrams from Baldwin et al. (1981) and Veilleux & Osterbrock (1987) for the H α sample. Colours and symbols are as in Fig. 4.

show the same diagnostic diagrams presented for our original classification method.

By comparing the BPT diagrams in Figs. 4–7 with those in Figs. A.1–A.4 it is possible to see how the distribution of traditionally classified nebulae changes significantly with respect to what we obtain with the model-comparison-based classification algorithm. In the traditional paradigm, the shock class collects most of the poorly detected regions. This happens because the

traditional criteria do not correctly consider the non-detection of one or more lines needed for the classification. Restricting the catalogue, the number of shocks decreases significantly, to reach a number comparable to what is classified by our algorithm. Many points classified with the traditional criteria overlap with the models we used to represent shocks in the model-comparison-based classification. Still, the traditional criterion tends to classify as shock regions with significantly lower

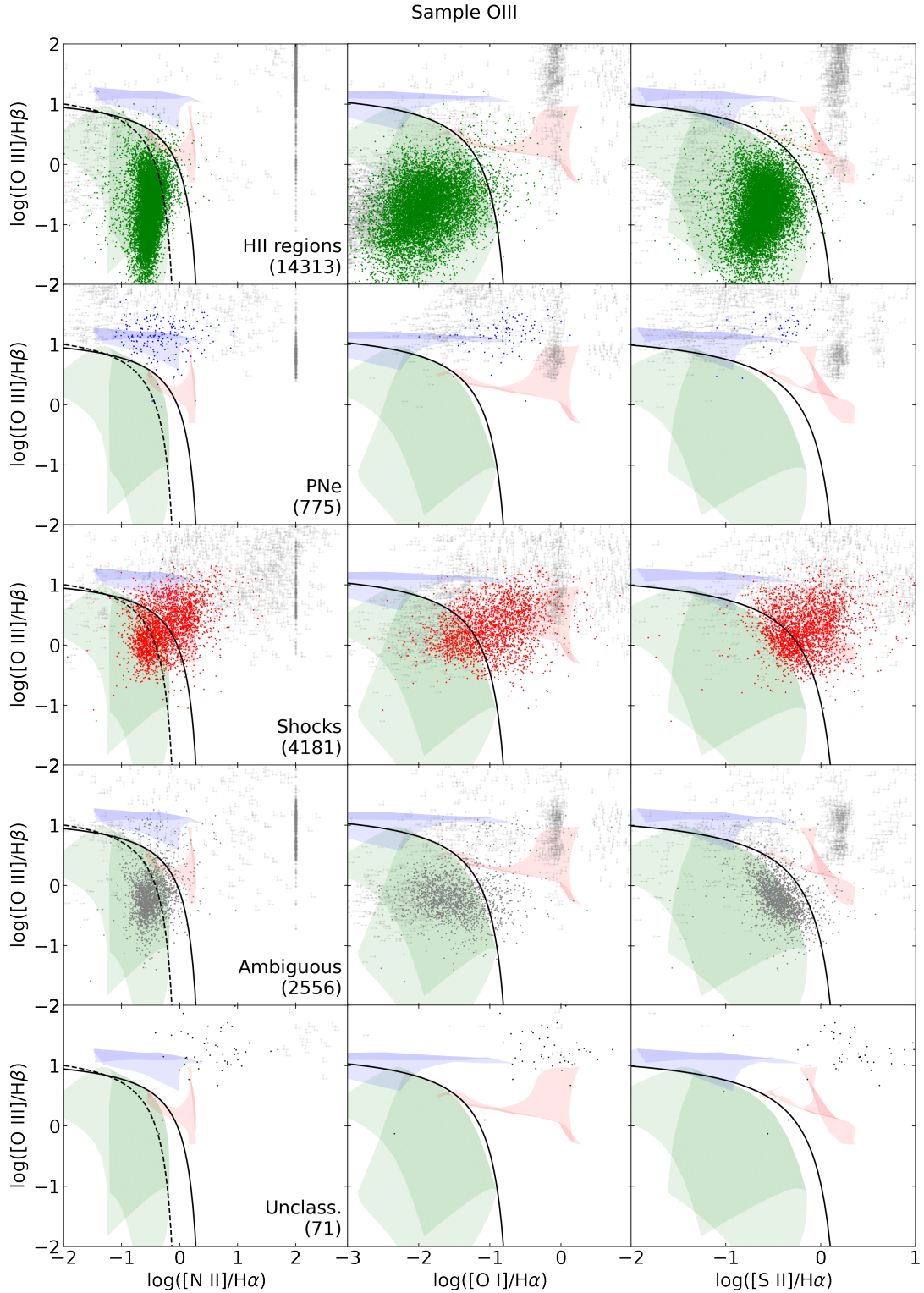


Fig. 6. Diagnostic diagrams from Baldwin et al. (1981) and Veilleux & Osterbrock (1987) for the OIII sample. Colours and symbols are as in Fig. 4.

$[\text{O III}]\lambda 5007/\text{H}\alpha$ than our automatic classifier and it extends into the region of the diagnostic diagrams typically occupied by H II regions. This results in a high fraction of ambiguously classified nebulae, which is a factor ~ 2 times larger with respect to the results of the model-comparison-based algorithm (15–20% vs 10% of the total sample of nebulae) and becomes comparable (although still 10–20% higher) only when applying very strict restrictions (i.e. in the [O III] or BPT sample).

Finally, the sharp boundary in the $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha$ implies that the selection misses a large fraction of regions with $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha$ below the cut, but that, according to our algorithm, are better modelled by shock models than H II regions or PNe, as already shown by Kopsacheili et al. (2020) and Long et al. (2022). This highlights the need for a more refined classification method not based on simple, sharp criteria, but able to assign probabilities to ambiguous regions.

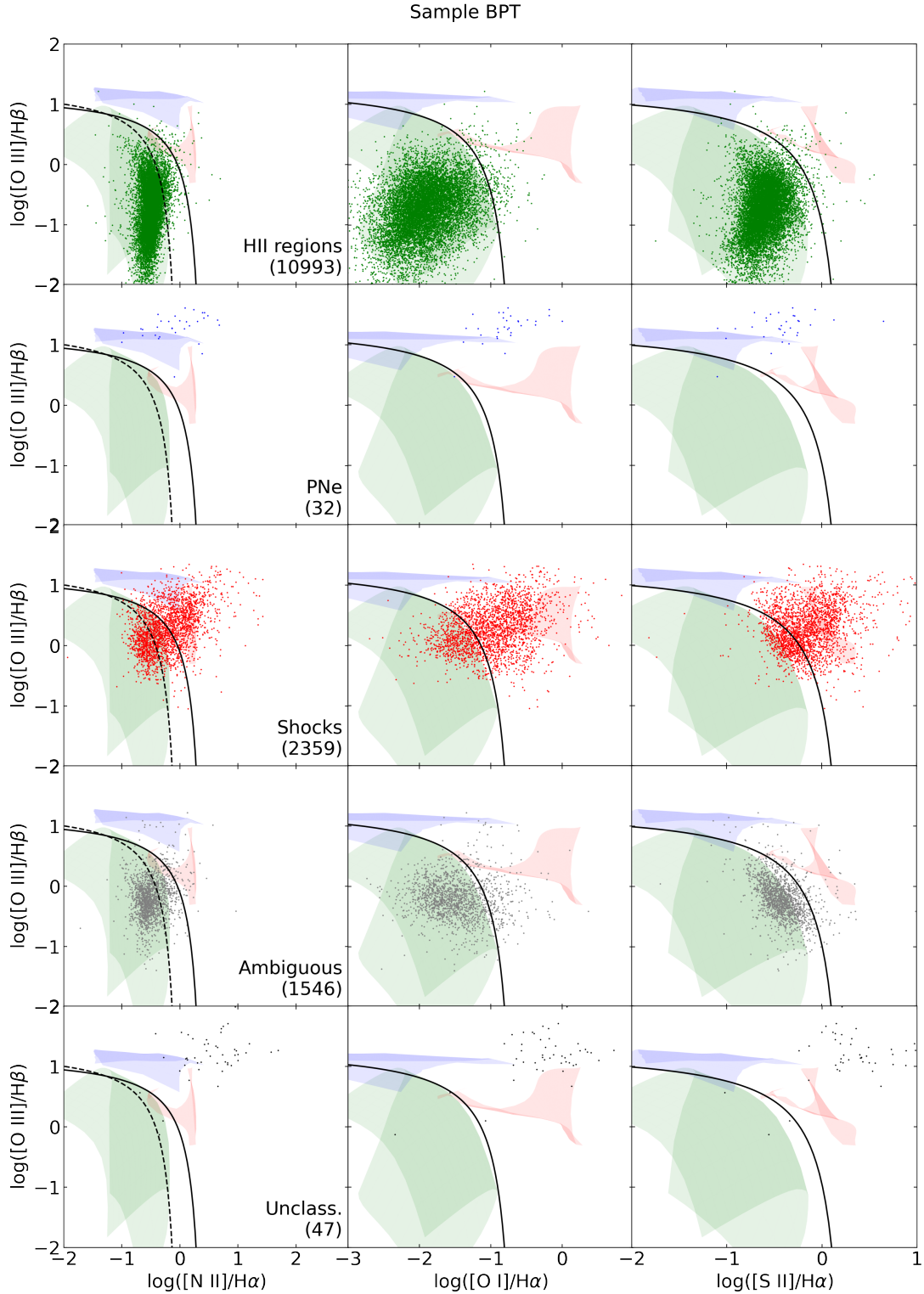


Fig. 7. Diagnostic diagrams from Baldwin et al. (1981) and Veilleux & Osterbrock (1987) for the BPT sample. Colours and symbols are as in Fig. 4.

The traditional classification criteria also produce many more unclassified regions (a factor of 20–30 higher) with respect to what we find with the new algorithm. This shows how there are regions of the parameter space of the properties of nebulae that are not covered by any criteria. By looking at the distribution of the unclassified nebulae on the diagnostic diagrams, it is clear how most of them would be considered as H II regions by using the Kewley et al. (2001) line in the $[\text{N II}]\lambda 6584/\text{H}\alpha$

diagram instead of the Kauffmann et al. (2003) one or by our new classification algorithm.

Most of the nebulae are classified as H II regions also by the traditional criteria (20 540, 50%), but they are only $\sim 70\%$ of the H II regions classified by our algorithm in the full catalogue. However, we have already discussed in Sect. 6.1 how the $\text{H}\alpha$ sample might provide a more reasonable comparison. If we compare the results for the $\text{H}\alpha$ sample, the numbers are much more

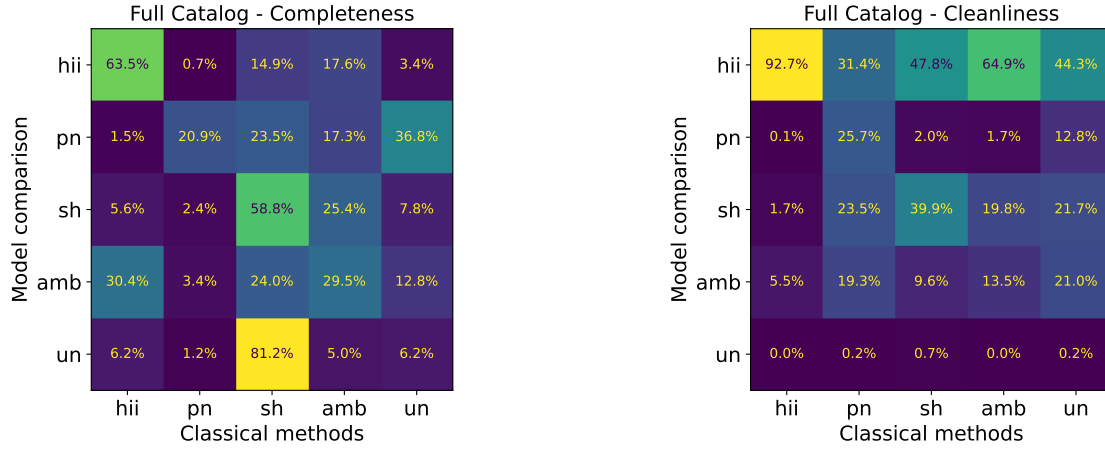


Fig. 8. Confusion matrices comparing the results of the model-comparison-based classification and traditional nebulae classification techniques for the full sample. The five classes included in the matrices are described in Sects. 6.1 and 6.2 for the new and traditional classifier respectively. The left matrix reports the completeness of the traditionally classified sample in the diagonal, while the misclassification of each different class outside the diagonal. The right matrix, instead, shows how clean the traditionally classified sample is on the diagonal, and how much is contaminated by objects belonging to different classes outside of the diagonal.

Table 6. Distribution of nebulae in the different classes according to the traditional classification criteria.

Class	All	H α	[OIII]	BPT
H II	20 540 (0.50)	20 540 (0.64)	12 622 (0.58)	10 338 (0.69)
PNe	647 (0.02)	647 (0.02)	456 (0.02)	46 (0.00)
Sh.	9328 (0.23)	4123 (0.13)	4110 (0.19)	2211 (0.15)
Amb.	8124 (0.20)	4619 (0.14)	2992 (0.14)	1584 (0.11)
Un.	2281 (0.06)	2281 (0.07)	1716 (0.08)	798 (0.05)
Total	40 920	31 110	21 896	14 977

Notes. Next to the absolute number, we report the fraction of nebulae in that class with respect to the total.

similar (20 540, 64%, vs 23 939, 81%), even though they are still significantly lower. However, we must consider that many of the H II regions classified by the model-comparison-based algorithm are classified as shocks, ambiguous and not classified at all in the traditional paradigm, as discussed above.

Finally, the traditional criteria identify a similar number of PNe with respect to our new algorithm (647, 2%). Their distribution in the diagnostic diagrams matches relatively well the area covered by the models we are using for the classification, except for the [O I] λ 6300/H α diagram. There, the nebulae show a significantly higher [O I] λ 6300/H α with respect to the models. This problem is shared with the nebulae classified by the new algorithm. This means that either we overestimate the brightness of the [O I] λ 6300 line in our measurements, or that the models are not representative of the full population of PNe, which is possible, since we are not using the complete set of PNe models. There is also a relatively large cloud of points with low [O III] λ 5007/H β that overlap, at least partially, with the H II regions models. This could be a consequence of the fact that the traditional criterion does not exclude regions that are not PNe, but only select regions that are compatible with being PNe candidates (Ciardullo et al. 2002; Herrmann et al. 2008).

6.3. IZI model comparison vs. traditional criteria

In Figs. 8–11, we show a direct comparison between the results of our classification algorithm and the traditional criteria for

each different sub-sample. In each figure, the left matrix is normalised with respect to the IZI model-comparison classification, so it represents the fraction of the nebulae classified in a particular class by the model-comparison-based algorithm that falls on a different class in the traditional classification scheme. Basically, these matrices show how the traditional classification criteria misclassify the nebulae, assuming that the IZI classification is correct. The right matrix, on the other hand, shows the opposite. This can be interpreted as a way of evaluating how contaminated the samples of traditionally classified nebulae are, assuming, once again, that the IZI model-comparison-based classification is correct. Looking at the left matrices, it is possible to see that the traditional criteria agree with the IZI classification for most H II regions and shocks. In all the different samples, at least 60% of H II regions and 55% of shock regions keep their classification using the traditional criteria. In the BPT sample, the percentage increases to 86% for H II regions and 65% for shocks.

The same is not true for PNe. Typically, around 20–30% of the PNe classified by the new algorithm keep their classification using the traditional criteria. This test clearly shows that the two approaches are classifying different objects as PNe. The reason why this happens is still not clear. The traditional criterion mostly relies on precise measurement of the [O III] λ 5007 and H α fluxes and their ratio. It is possible to see from the literature on this topic that most of the effort when classifying PNe is dedicated to precisely measuring the line fluxes and their biases. On the other hand, our new algorithm relies on comparing line ratios and other properties with models, which should allow us to perform the classification in less ideal conditions.

The unclassified and ambiguous classes produced by the two methods are not directly comparable, since they are defined differently. However, it is interesting to see that the vast majority of the nebulae that our algorithm cannot classify are identified as shocks by the traditional classifier (~80–90%).

The matrices on the right side of Figs. 8–11 highlight other interesting effects. While the sample of traditionally classified H II regions seems to be quite clean, independently of which sub-sample we are considering, the same cannot be said of all the other classes. Both PNe and shocks show significant contamination, with up to 50% them being classified

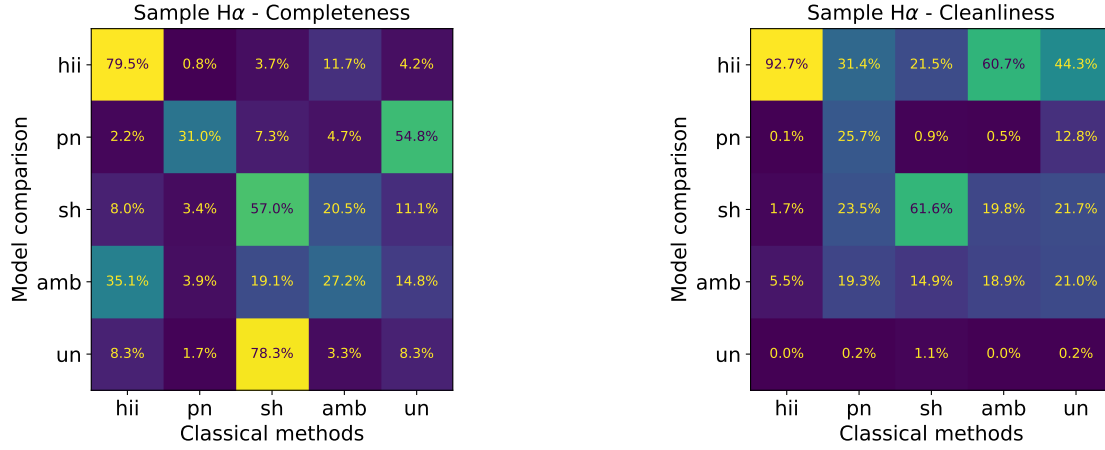


Fig. 9. Confusion matrices comparing the results of the new, model comparison-based classification and traditional nebulae classification techniques for the H α sample. Caption as in Fig. 8.

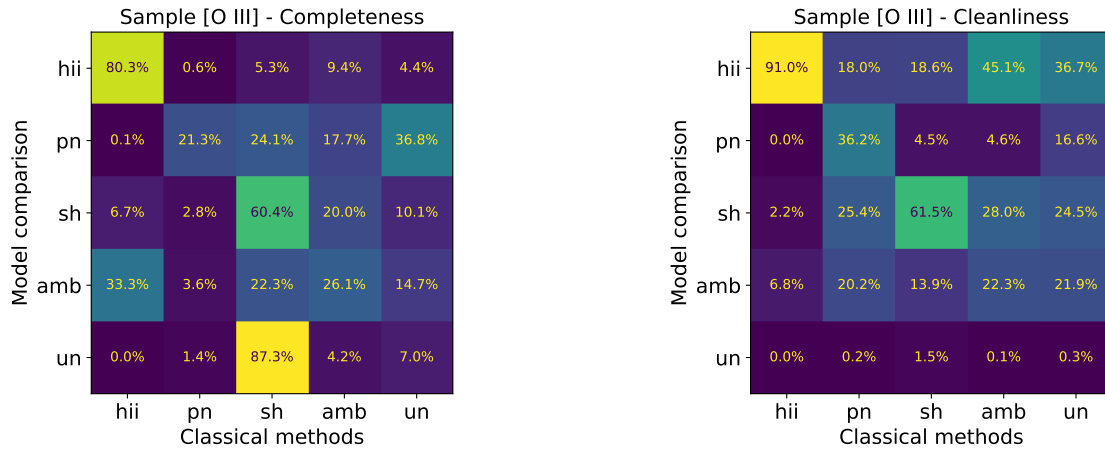


Fig. 10. Confusion matrices comparing the results of the new, model comparison-based classification and traditional nebulae classification techniques for the H α sample. Caption as in Fig. 8.

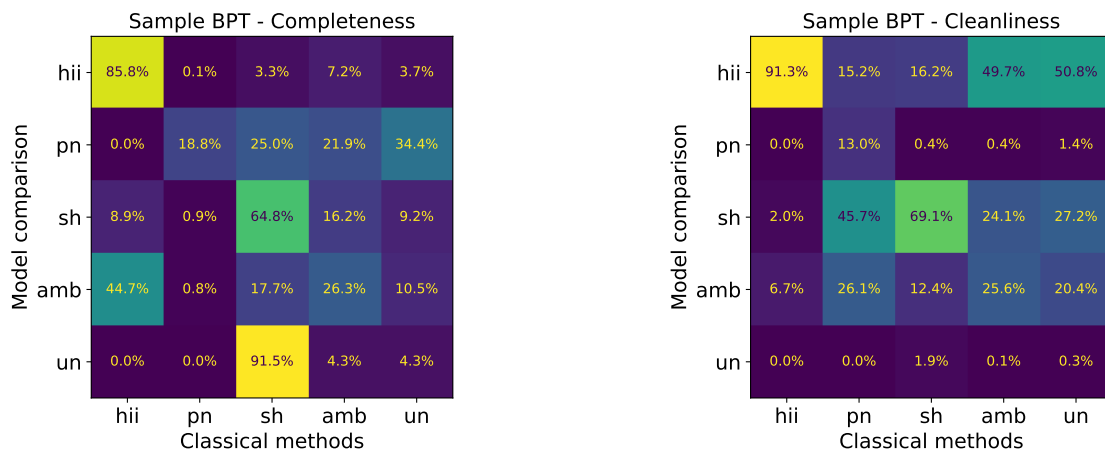


Fig. 11. Confusion matrices comparing the results of the new, model comparison-based classification and traditional nebulae classification techniques for the H α sample. Caption as in Fig. 8.

as H II region by the model-comparison-based algorithm. The situation improves significantly by moving from the full sample to the more restrictive sub-samples, especially for shocks. If we consider only the BPT sub-sample, almost 70% of the nebulae traditionally classified as shocks are considered shocks

also by our algorithm. In PNe the results are not as good, but we can still see some improvement. Significantly, the best results for PNe are found using the OIII sample and not the BPT one, as expected since the number of PNe in the BPT sample is negligible. Finally, the matrices also show that more

E(B-V) comparison

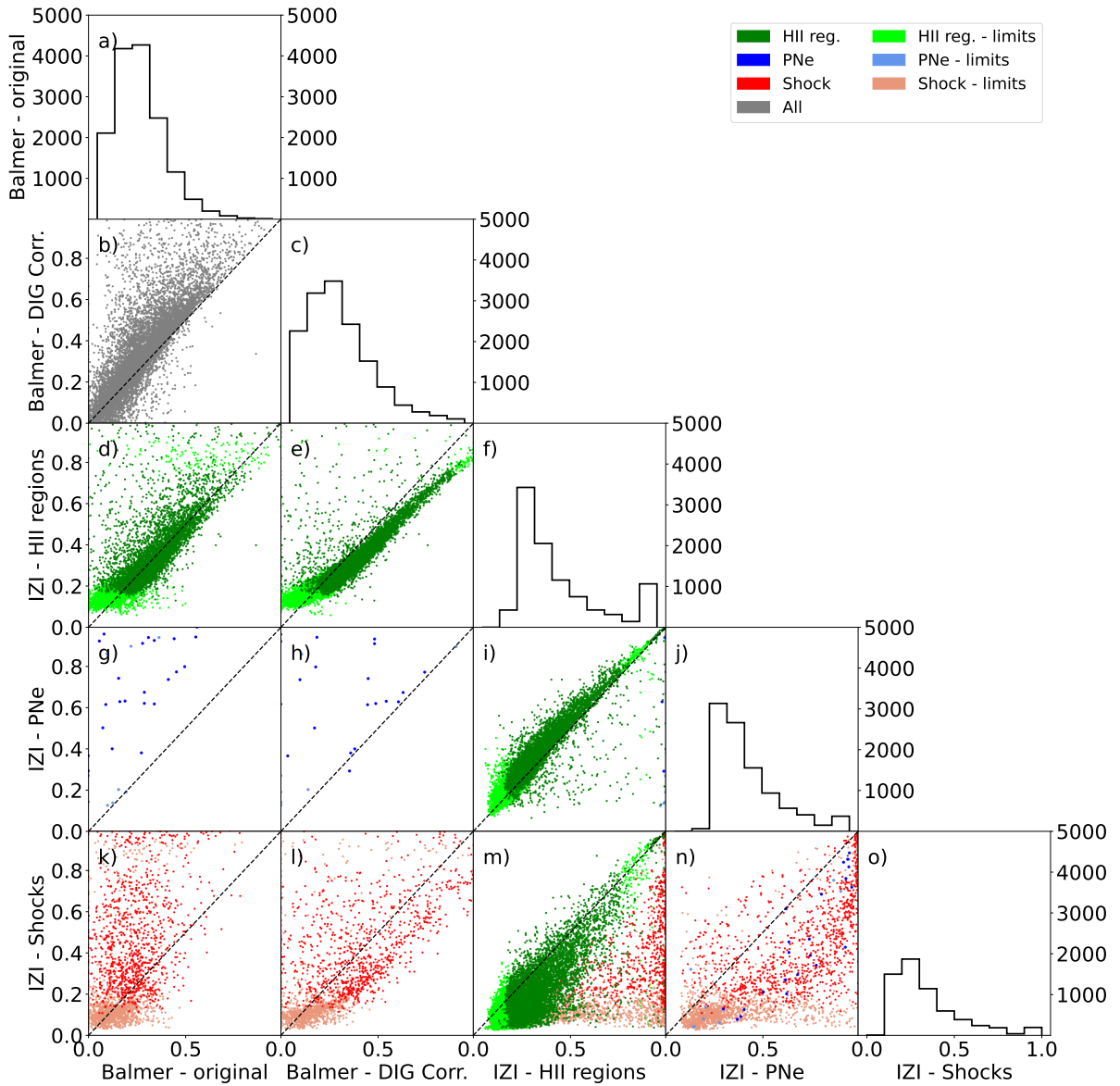


Fig. 12. Corner plot comparing the reddening value ($E(B - V)$) measured directly from the data (before and after the correction for DIG contribution) with the values obtained by IZI using the three different grids of models used in the classification algorithm (H II regions, PNe and Shocks). Only the regions belonging to the BPT sample are considered in this plot (see Sect. 6.1). In each plot, the coloured points represent the nebulae we could classify with our algorithm, following the usual convention. Only the points belonging to the grids considered by the specific panel are coloured. The plots on the matrix's diagonal show regions' distribution as a function of $E(B - V)$ for each method.

than half of the ambiguous and unclassified regions found using traditional criteria are H II regions according to the model-comparison-based algorithm.

7. Discussion

7.1. Extinction estimates

Having a reliable estimate of the reddening in each region is essential when measuring quantities involving lines with wide wavelength separations (e.g. the ionisation parameter from the $[S\ III]\lambda 9068/[S\ II]\lambda\lambda 6717, 6731$ or $[O\ III]\lambda 5007/[O\ II]\lambda 3727$ ratios) or intrinsic extinction-corrected luminosities (e.g. star formation rate from $H\alpha$). The reddening is typically measured by comparing the observed $H\alpha/H\beta$ ratio with its theoretical value for a nebula with a temperature of 10 000 K, low-density limit

and Case B conditions ($H\alpha/H\beta = 2.86$, Osterbrock & Ferland 2006). We use this technique, as implemented in the `pyneb` python package (Luridiana et al. 2015), to estimate the colour excess ($E(B - V)$) in all nebulae where both $H\alpha$ and $H\beta$ are detected at better than 3σ , and we compute it both before and after the DIG correction. The $E(B - V)$ is a reddening tracer since it expresses how red an object is with respect to the expectation, and it is one of the free parameters considered by IZI when comparing the nebulae properties to the models. IZI does not assume a fixed $H\alpha/H\beta$ to compute the $E(B - V)$ but instead samples the parameter space to identify the reddening producing the best agreement between the observations and the models. This is done for each grid of models, so, on top of the values obtained directly from the Balmer decrement ratio, we have three further measurements of the $E(B - V)$ for each nebula, each one

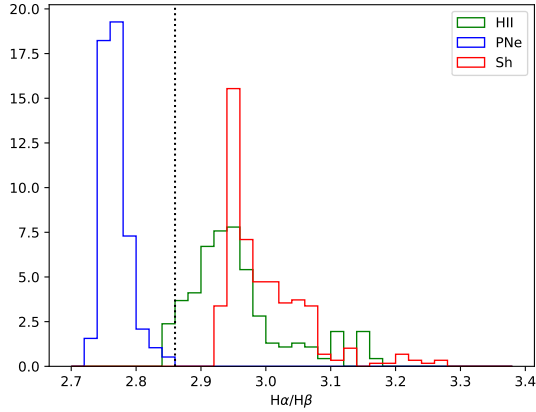


Fig. 13. Histogram showing the intrinsic $H\alpha/H\beta$ ratio for the different grids of models used in the classification procedure. Histograms are coloured following the usual convention for H II regions, PNe and shocks. The vertical dashed line represents the theoretical value of the $H\alpha/H\beta$ ratio for Case B conditions and a temperature of 10^4 K, typically assumed when recovering the extinction from the observed $H\alpha/H\beta$ ratio in H II regions.

associated with the expectation value from the marginalised posterior probability distribution function in $E(B - V)$ as determined by IZI for each class (H II regions, PNe, and shocks). For values of $E(B - V)$ which are too close to the edge of the prior space (and in particular close to 0), IZI returns a limit.

In this section, we compare measurements of $E(B - V)$ obtained with the different methods to understand differences and similarities. We can make this comparison because all the different measurements are based on the same reddening law, that is the Cardelli et al. (1989) law as revised by O’Donnell (1994) with $R(V) = 3.1$. Figure 12 shows a corner plot comparing all the $E(B - V)$ we measured. We only show nebulae belonging to the BPT sub-sample because we are confident that their $H\alpha$ and $H\beta$ are well detected, and that the IZI results should be the most reliable since it can use the full set of lines when comparing the observations to the models. In the figure, the regions for which IZI return an upper limit in $E(B - V)$, are shown in a lighter colour and are omitted from the following discussion.

We first analyse how the DIG correction influences the reddening correction. Figure 12b shows the comparison between the $E(B - V)$ measured using the original and the DIG corrected Balmer decrement. The two quantities show a well-defined relation, but there is a consistent scatter. In general, the $E(B - V)$ measured after the DIG correction is higher than what was measured before. This is consistent with a screen model for the dust distribution (e.g. Tomičić et al. 2017), which assumes that the DIG is distributed in a disk with a larger scale height than the dust and the other nebulae, mostly located in a thinner disk. As a consequence, the DIG is less attenuated and results in the lower $E(B - V)$ generally measured before the DIG correction.

Panels d, e, g, h, k, l, of Fig. 12 compare the IZI $E(B - V)$ to the Balmer decrement $E(B - V)$ measurements. Here it is possible to see that the situation is more complex. Panels d and e compare the IZI and Balmer decrement $E(B - V)$ for H II regions, which are the vast majority of the sample. The two plots show two different behaviours. When comparing IZI $E(B - V)$ to the non-DIG corrected Balmer decrement $E(B - V)$ (panel d), we obtain a relation that follows relatively well the 1-to-1 relation, albeit with a large scatter. Moving to the Balmer

decrement $E(B - V)$ obtained using the DIG-corrected flux, we have the opposite situation. The scatter is much less, but we are not so close to the 1-to-1 relation (it is important to remember that we run IZI using the DIG-corrected line fluxes). In particular, the Balmer ratio systematically overestimates the extinction with respect to IZI, and it seems like the difference between the two relations increases when moving towards higher $E(B - V)$. The most likely reason for this behaviour is that when measuring the extinction using the Balmer decrement, we assume a fixed theoretical $H\alpha/H\beta$ ratio for all nebulae (2.86). However, the underlying gas conditions assumed when using this value are not always satisfied. In particular, the $H\alpha/H\beta$ depends on the gas temperature, which can change significantly from cloud to cloud. The lower the temperature, the higher the expected $H\alpha/H\beta$ ratio (Osterbrock & Ferland 2006). For example, assuming a temperature of 5000 K ¹¹ the theoretical $H\alpha/H\beta$ ratio is 3.04 and results in a decrease of the $E(B - V)$ of ~ 0.06 mag. Consequently, using a theoretical $H\alpha/H\beta$ ratio that does not correspond to the actual conditions of the clouds results in a wrong extinction estimate.

The H II regions models we use for the classification are not parameterised in terms of electron temperature. Their temperature profiles are set by the thermal equilibrium given by the cooling and heating processes incorporated in the CLOUDY calculations, which depend mainly on the chemical abundance, the electron density and the shape of the ionising spectrum. Therefore models with different metallicities and ionisation parameters will have different temperature profiles and different intrinsic $H\alpha/H\beta$ ratios, as shown in Fig. 13. Only a small fraction of the H II region models have a $H\alpha/H\beta$ ratio close to 2.86. The vast majority have higher values, indicating, on average, lower electron temperatures.

Selecting the correct theoretical value of the $H\alpha/H\beta$ ratio is, however, not trivial. In high S/N spectra, where it is possible to detect the auroral lines needed to measure the gas temperature, it is possible to apply an iterative process, to constrain the temperature (and the $H\alpha/H\beta$ ratio) and the $E(B - V)$ simultaneously. However, such high-quality spectra are not always available and assuming a single theoretical value for all nebulae is often necessary. By minimising the offset between the relation we observe in Fig. 12e and the 1-to-1 relation, we can identify a new theoretical value that should result, on average, in better extinction estimates. This is just a zero-order correction, but it should still produce significant improvements. A detailed analysis to understand the origin of the higher order trends and how they should be considered is outside the scope of this paper, and we leave it to a future work. The new ratio resulting from the minimisation is $H\alpha/H\beta = 3.03$ (Fig. 14), which corresponds to a temperature of $\sim 5200\text{ K}$, slightly lower than what is typically observed in nearby extragalactic H II regions (e.g. Ho et al. 2019; Berg et al. 2020; Kreckel et al. 2022). We recommend using this value instead of the classic $H\alpha/H\beta = 2.86$ when applying the Balmer decrement to H II regions in massive star-forming galaxies having nearly solar metallicity. We also notice that this new approach works only if DIG-corrected fluxes are used to calculate the extinction. If uncorrected fluxes are used, the dilution provided by the DIG contribution counteracts the effect of using the wrong theoretical ratio (Fig. 12d). It is interesting how the observed agreement, and supposedly good performance of non-DIG corrected Balmer decrement extinction correction, seems to arise from a coincidence, where two different biases (underestimation

¹¹ All the computations involving the $H\alpha/H\beta$ ratio and the gas temperature are performed via pyneb.

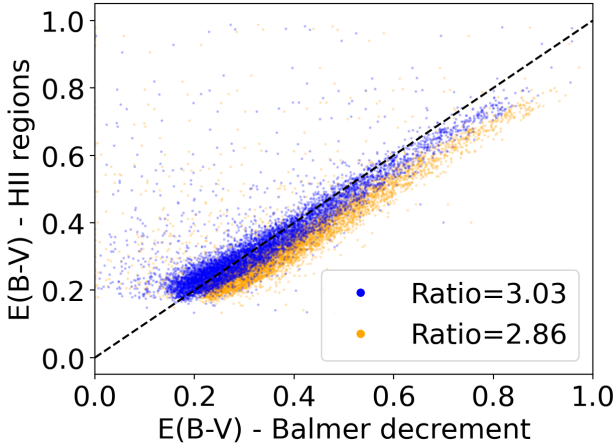


Fig. 14. Relation between the $E(B - V)$ measured via the DIG corrected Balmer decrement and by IZI when using the traditional $H\alpha/H\beta = 2.86$ (orange points) and the new recommended value $H\alpha/H\beta = 3.03$ (blue).

of the extinction due to DIG contamination and overestimation of the extinction due to adopting the wrong intrinsic Balmer ratio) have similar magnitudes and act in opposite directions.

Figures 12k,l, indicate a similar, yet different, scenario for shock ionised regions. Panel k shows that the $E(B - V)$ measured by IZI using the shock models and that measured via the Balmer decrement using the original fluxes are not in good agreement. IZI almost always finds much higher extinctions with respect to the Balmer decrement. There seems to be a very steep relation between the two quantities, but the scatter is definitely high, and this relation's origin is unclear. If we consider the DIG-corrected Balmer decrement, the picture changes. There is a large scatter, and the smaller statistics prevent us from performing the same test we did for H II regions, but it is possible to see a similar scenario. The main difference is that the offset from the 1-to-1 relation is larger. This is expected. Shocks are a much more complex situation than photo-ionisation regarding the mechanism and the parameters that can influence the observed spectrum of a nebula, and Fig. 13 shows that the $H\alpha/H\beta$ of the models we are using are much higher than 2.86.

Figures 12g,h show the same plots again but for PNe. Although the number of PNe in this sub-sample is only 32, so we do not have sufficient statistics to perform any analysis. We notice, however, that IZI recovers mostly high extinction for these objects, higher than what is found via the Balmer decrement, with the DIG-corrected Balmer decrement returning values that are more similar to the IZI ones. Planetary nebulae are known for having higher electron temperature and densities with respect to H II regions, therefore $H\alpha/H\beta = 2.86$ is typically not a good approximation for these objects. As it is also possible to see from Fig. 13 a lower reference ratio should be used. Once again, in Galactic or nearby objects where it is possible to obtain high S/N spectra, the extinction has typically been measured assuming the appropriate $H\alpha/H\beta$ for the set of conditions found in the specific nebula (e.g. Meatheringham & Dopita 1991a,b; Walsh et al. 2016).

Finally, Figs. 12i,m,n compare the $E(B - V)$ measured by IZI when applying two different models to the same nebulae to test how a misclassification would affect the measurement of the extinction. For H II regions misclassified as PNe, the final $E(B - V)$ would be slightly overestimated. In the opposite case, a PNe misclassified as an H II region, the $E(B - V)$ would be heavily underestimated. This is in line with the distribution of

the $H\alpha/H\beta$ ratios of the models shown in Fig. 13. When comparing shocks and H II regions (panel m), the $E(B - V)$ of the shock ionised regions would be overestimated when misclassifying them as H II regions. For most of the H II regions, on the other hand, the impact of misclassification is much more limited, and would affect principally the nebulae with low extinction. Concluding, panel n compares shocks and PNe. While the statistics for PNe is still small, we can see that for most shocks, a misclassification would result in a larger measured $E(B - V)$.

7.2. Physical properties of the nebulae

Thanks to the complete set of spectral information included in our nebulae catalogue, we can compute some regions' physical properties and study how their distribution changes depending on the selection criteria and the classification scheme. This could help us identify biases introduced by different approaches to classifying and selecting sources when investigating ionised nebulae. Some additional information could be recovered by comparing the values directly measured from the data with those obtained by IZI as we did for the extinction. However, we decided not to perform this comparison since the set of models we are using for the classification is limited, and it has not been chosen to obtain precise and reliable estimates of the nebulae' physical parameters. A similar analysis based on the advanced photo-ionisation and shock models will be the subject of a follow-up paper.

For H II regions, we computed metallicities, following the S-calibrations by Pilyugin & Grebel (2016), the ionisation parameter from Diaz et al. (1991, $\log(u)^{12}$), size (in pc) and $H\alpha$ luminosity. For PNe, we show the $H\alpha$ luminosity and the $[O III]\lambda 5007$ absolute magnitude, the latter a critical quantity in estimating the distance of nearby galaxies via the PNe luminosity function (see Sect. 7.6 for a discussion on this specific quantity). Finally, for the shock-ionised regions, we show their $H\alpha$ luminosity and the $H\alpha$ and $[N II]\lambda 6584$ velocity dispersion, which represent the kinematics of the two groups of lines defined in Sect. 4.2.

All the quantities follow unimodal distributions, with a relatively narrow spread around a well-defined peak, for the H II regions classified by both approaches (Fig. 15). The only exception is the ionisation parameter, which shows a very sharp peak at high values. The peak slowly disappears when moving to more restrictive samples suggesting that the feature is related to regions where the lines needed for the measurements are not well detected (i.e. the $[S III]\lambda 9068$ line). The figure also shows that when applying more restrictive cuts the distribution of metallicities changes, moving towards slightly higher metallicities. Also, the shape of the size and $H\alpha$ luminosity distribution changes significantly. As expected, more restrictive cuts mostly reject fainter and smaller regions, a sign that the most luminous regions are typically the largest ones. As expected by the agreement between both classification paradigms when identifying H II regions, the properties of these nebulae are distributed similarly in the two samples (top and bottom panels of Fig. 15).

Since the size of the PNe sample does not change much between the different sub-samples when applying the new,

¹² This calibration requires both the $[S III]\lambda 9068$ and $[S III]\lambda 9532$, but only the former falls inside MUSE wavelength range. However, these two emission lines belong to the same doublet, and the following relation ties their flux: $I(9532) = 2.47 \times I(9068)$ (Osterbrock & Ferland 2006).

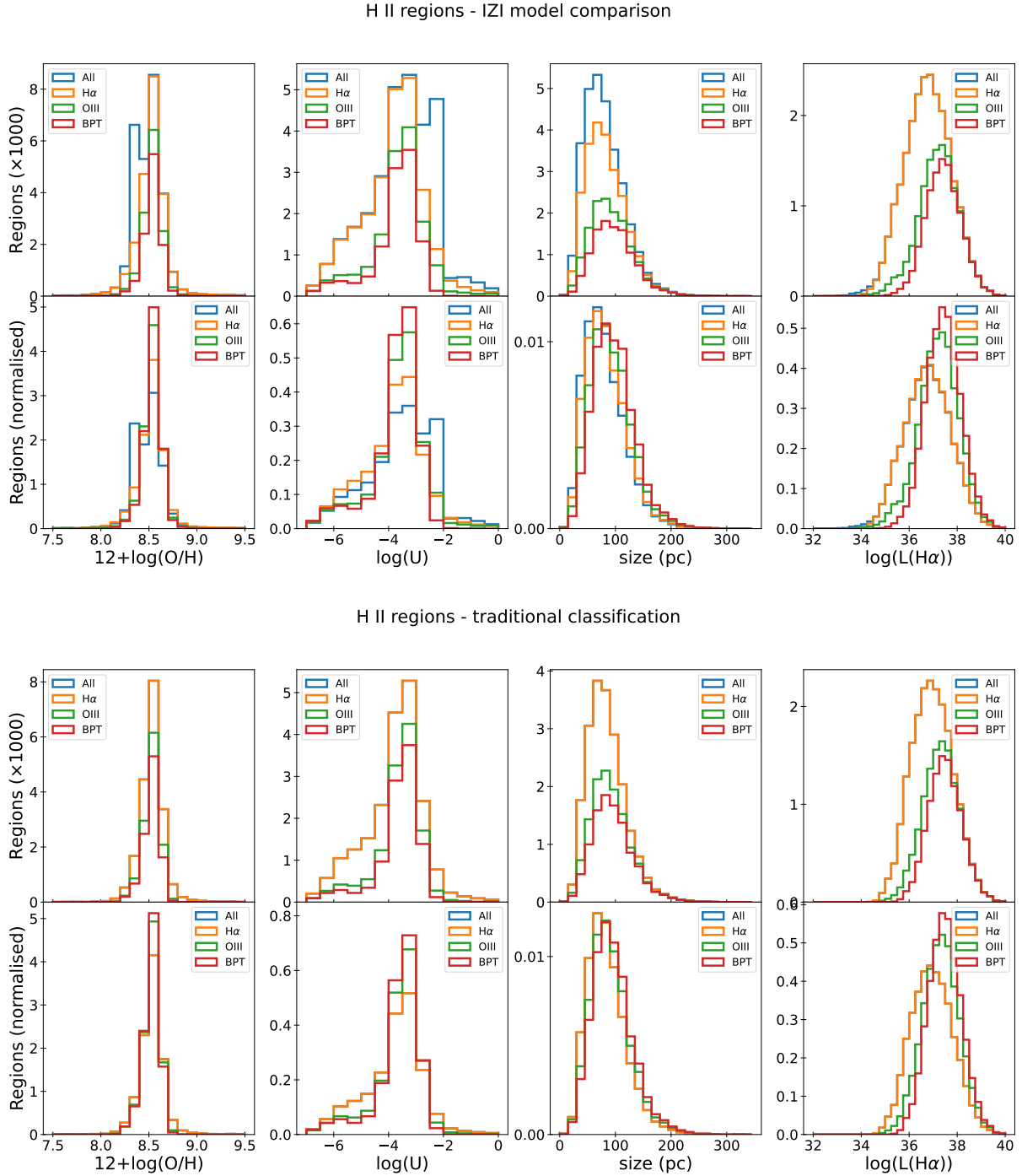


Fig. 15. Physical properties of H II regions. Top panel: histogram showing how the distribution of the main measured properties of H II regions changes as a function of the considered sample. From the left to the right the following quantities are shown: S-calibration metallicities (Pilyugin & Grebel 2016), ionisation parameter (Diaz et al. 1991), circularised radius, H α luminosity. The top row shows the absolute distribution for each quantity, while the bottom row shows the normalised distribution. Bottom panel: same histograms as shown in the top panel, but for the H II regions classified using the traditional classification criteria. Since there is no change in the number of H II regions in the full catalogue and the H α sample, the two histograms overlap, and the blue line representing the full catalogue is not visible.

model-comparison-based classification algorithm, the H α luminosity and $M_{[\text{O III}]}$ distributions for these objects (Fig. 16, top panel) is relatively independent of the considered sub-sample. The only exception is when we consider the BPT sub-sample. It contains only a handful of objects, which are typically among the most luminous ones, as expected. The situation changes significantly when considering the traditionally classified PNe (Fig. 16, bottom panels). First of all, the H α distribution ends

sharply at the bright end, as a consequence of the criterion used for the classification (Eq. (3)). The PNe classified by the model-comparison-based algorithm seem to reach higher H α luminosities (up to $\sim 10^{37}$ erg s $^{-1}$) and they do not show such a sharp cut. Also the distribution of $M_{[\text{O III}]}$ changes significantly between the PNe identified by the two approaches. The sample classified by the model-comparison-based algorithm shows a relatively symmetric and narrow peak between -2 and -4 mag. The

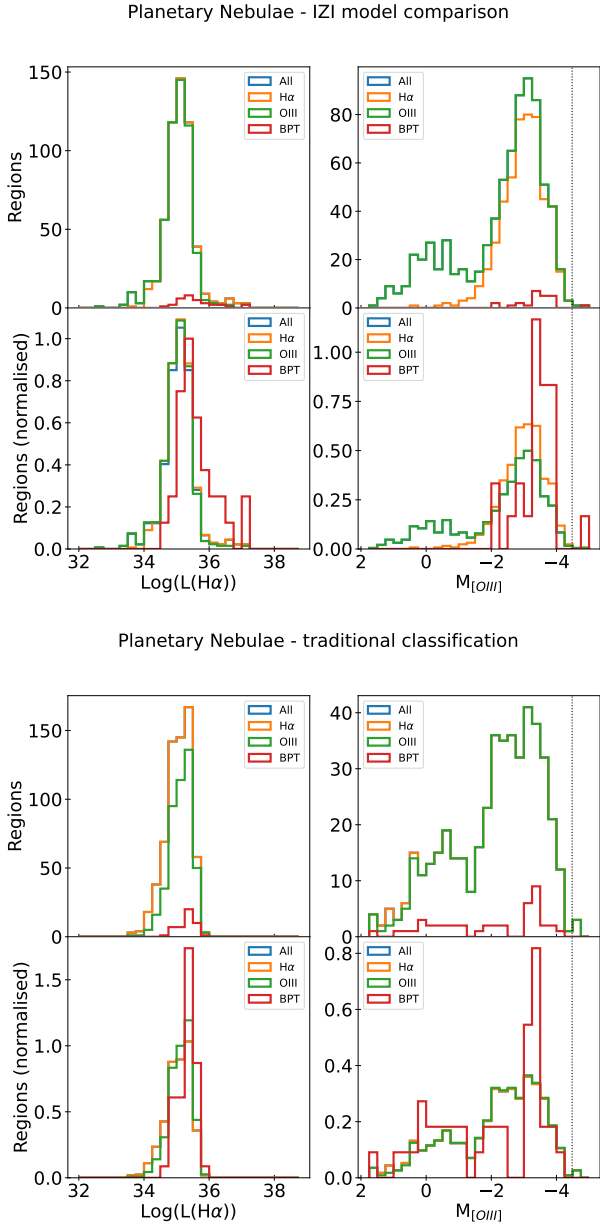


Fig. 16. Physical properties of PNe. Top panel: histogram showing how the distribution of the main measured properties of PNe changes as a function of the considered sample. From the left to the right, the following quantities are shown: $H\alpha$ luminosity, $[O\ III]\lambda 5007$ absolute magnitude (adapted from Eq. (4) using distances from Table 1). The top row shows the absolute distribution for each quantity, while the bottom row shows the normalised distribution. Bottom panel: same histograms as shown in the top panel, but for the PNe classified using the traditional classification criteria.

traditionally classified sample shows a similar peak, but much wider, and it decreases more slowly when moving towards fainter nebulae. This histogram also shows the presence of a second peak of objects with relatively low $[O\ III]\lambda 5007$ emission. It is likely dominated by contamination from H II regions and therefore does not affect the knee of the luminosity function and the inferred distance so much. All these differences highlight the discrepancies between the two classification approaches explored in this work. Finally, in both cases, there are a few objects which surpass the theoretical limit in absolute magnitude defined by Ciardullo et al. (1989; dotted lines in Fig. 16). Such findings are

not uncommon, even though the nature of these objects is not clear (Jacoby et al. 1996; Scheuermann et al. 2022).

Lastly, we analyse the properties of the shock-ionised regions, which show unimodal distributions in all the considered quantities and samples (Fig. 17, top and bottom panel). Both velocity dispersion distributions have consistent shapes in all the considered sub-samples. Only the number of objects increases when relaxing the selection criteria. The peak of the distribution is located at a relatively low velocity dispersion, but both $H\alpha$ and $[N\ II]\lambda 6584$ show a prominent high-velocity tail. The $H\alpha$ luminosity distribution of the regions does not show anything peculiar, but, as expected, when restricting the criteria we are removing nebulae from the faint end of the distribution. Similarly to what happens with the H II regions, there are no significant differences between the samples selected by the two classification approaches.

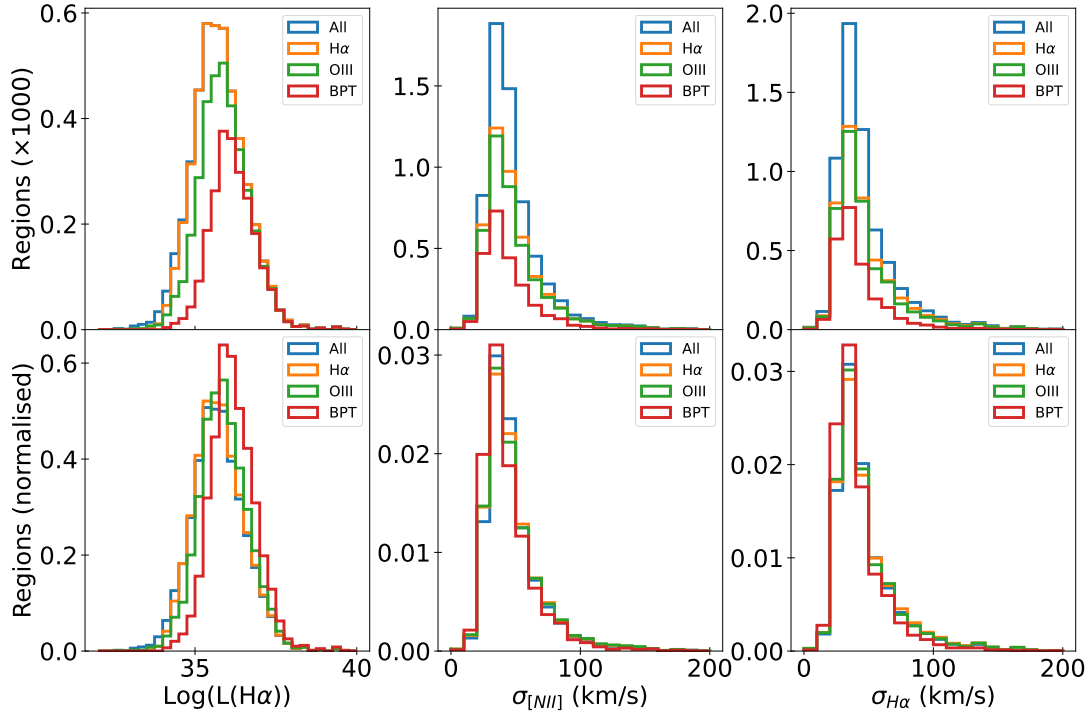
7.3. Comparison with Santoro et al. (2022)

S22 recently used the PHANGS-MUSE data to compile a nebulae catalogue to study the H II regions luminosity function (LF). While based on the same data, S22 aimed to produce a very clean sample of H II regions, resulting in several differences between their catalogue and the one presented in this work. Firstly, the S22 catalogue is based only on the $H\alpha$ maps produced by the DAP, boosting the detection of H II regions but reducing the possibility of detecting other classes of nebulae. Secondly, the catalogue has been created using HIIPHOT (Thilker et al. 2000), a common source detection algorithm specifically developed to detect H II regions in two-dimensional images. HIIPHOT detects nebulae by looking for morphologies typical of H II regions in narrow-band images. This contributes to the reduction of false detections (e.g. noise spikes), but it can also prevent the detection of regions whose morphology is not expected by the software. Finally, S22 clean the catalogue by applying the H II regions selection criteria based on the BPT diagrams we presented in Sect. 6.2 when investigating the LF properties. For a detailed view of how the catalogue was created and its general properties, we refer the reader to S22. We compare the two catalogues in the following, highlighting similarities and differences.

The most noticeable difference between them is in the number of detected regions. Our new catalogue contains 40 920 nebulae, while S22 detect 31 497 (including those not classified as H II regions). This discrepancy can be attributed to the approaches used for nebulae detection. We run our detection algorithm on a combination of different line maps (see Sect. 4). This allows us to identify regions that are not detected in $H\alpha$ and favours the detection of faint sources that are easily missed by the S22 approach.

To investigate how well the two compilations of nebulae overlap, we matched the two catalogues by comparing their associated segmentation maps. For each pair of regions (one for each catalogue), we computed the fraction of overlapping pixels with respect to the area covered by each region. We consider a match only if both fractions are 0.5 or larger. We went for this approach since the diverse segmentation approaches prevent us to rely on the linear distance between the regions' centres to match them, as commonly done with point source catalogues. We find matches for 10509 regions, roughly a third of the original S22 catalogue and a fourth of our catalogue, which seems to suggest that the two catalogues detect entirely different sources and that just a small fraction of them is in common. However, a comparison between the respective segmentation maps shows that, on average, $\sim 86\%$ of the area covered by S22 catalogue is also covered

Shocks - IZI model comparison



Shocks - traditional classification

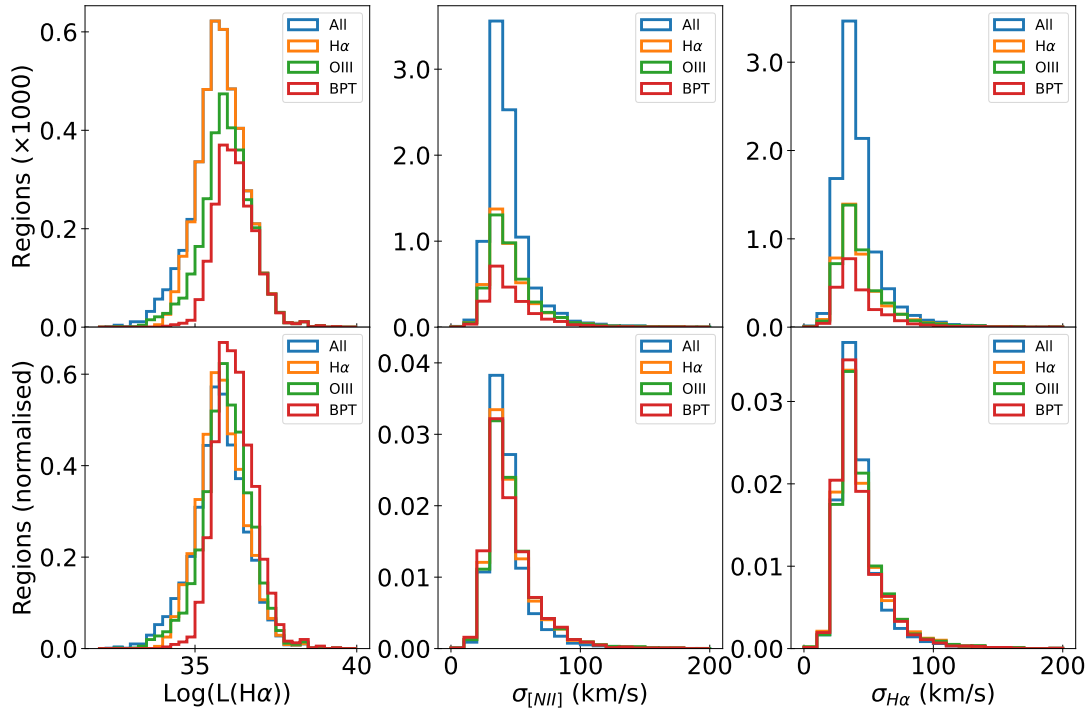


Fig. 17. Physical properties of shock-ionised regions. Top panel: histogram showing how the distribution of the main measured properties of SNRs changes as a function of the considered sample. The following quantities are shown from the left to the right: $\text{H}\alpha$ luminosity, $[\text{N II}]\lambda 6584$ velocity dispersion and $\text{H}\alpha$ velocity dispersion. The top row shows the absolute distribution for each quantity, while the bottom row shows the normalised distribution. Bottom panel: same histograms as shown in the top panel, but for the SNRs classified using the traditional classification criteria.

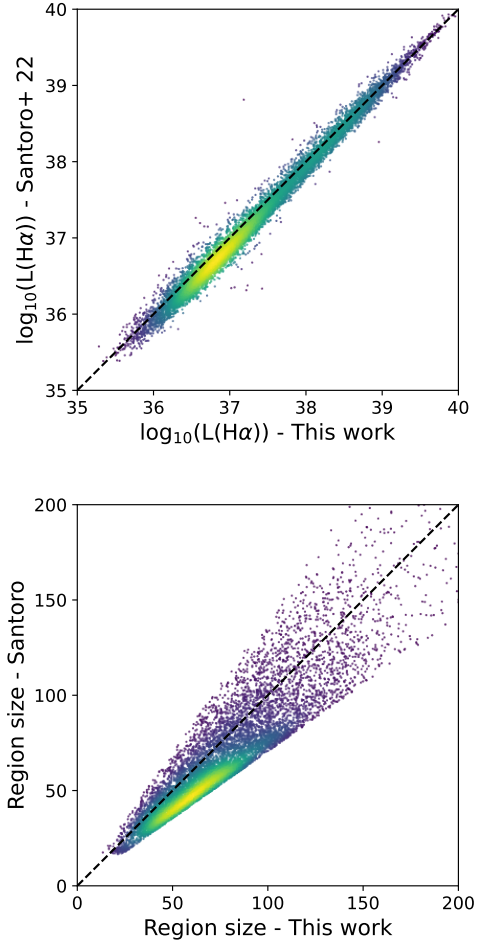
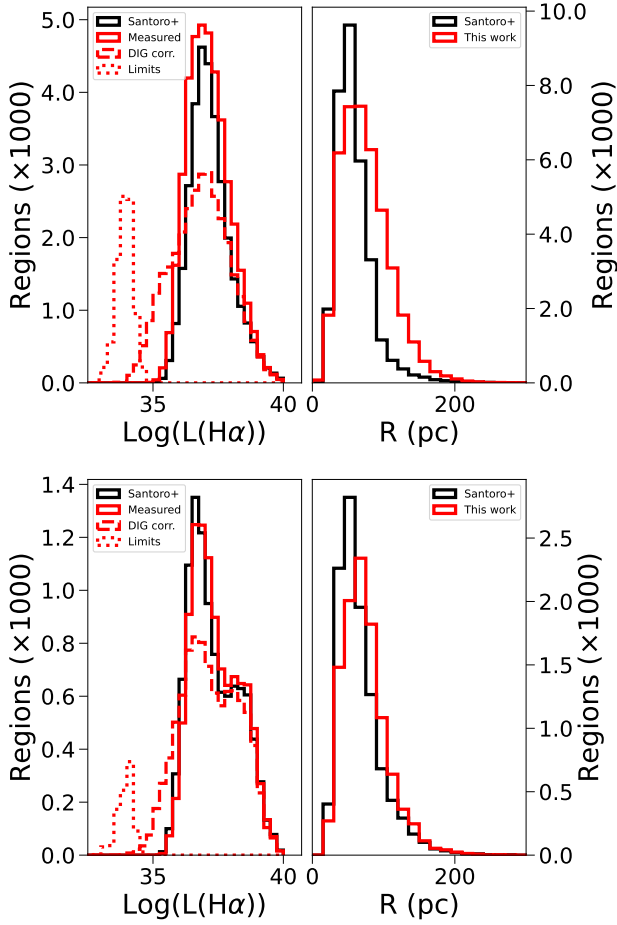


Fig. 18. Comparison with S22 catalogue. Top: histograms comparing the main properties of our catalogue to the ones of the nebulae catalogue published in S22. The left panel shows the $H\alpha$ luminosity distribution of the regions in the S22 catalogue (black), and that of our regions before the correction for DIG contribution (red). The dashed histogram reports the $H\alpha$ luminosity distribution of our sample after the DIG correction, and the dotted one shows the upper limits for those regions where $H\alpha$ is not detected after the DIG subtraction. On the right panel, we show the region size distribution in black for S22 and in red for this work. For both catalogues, the reported size is the circularised radius of the regions. Bottom: similar histograms comparing the properties of the matched regions. Symbols and colours are the same as in the top panel.

by our catalogue and that we recover a similar fraction of the total observed $H\alpha$ emission ($\sim 64\%$ for S22 vs. $\sim 70\%$ for our catalogue). On the other hand, only $\sim 43\%$ of our segmentation maps is covered by S22's ones.

These simple tests paint a scenario where both catalogues mainly contain the same nebulae (similar fraction of recovered $H\alpha$), but the detection algorithms produce significantly different segmentation maps (Fig. 1c). In particular, our catalogue contains more regions, which are also generally larger, even though this results, on average, only in a marginal increase of recovered $H\alpha$ flux. Figure 18 (top panel), where we compare the luminosity distribution (extinction corrected¹³) and the size distribution of the regions in both catalogues clearly highlights these trends. Considering the black and red (solid) histograms in the left panel of Fig. 18 (top), it is possible to see that the two distributions mostly match at the high-luminosity end but that moving towards

Fig. 19. Scatter plots showing a direct comparison between the properties of the regions that are present in both catalogues. The top plot shows a comparison between the $H\alpha$ luminosity, while the bottom one compares the circularised radii of the nebulae. The dashed lines in both plots represent a one-to-one relation. Colours visually represent the density of points in a specific plot area.

fainter sources our catalogue detects significantly more regions in each luminosity bin. However, the faint regions contribute less to the total amount of detected $H\alpha$ than the bright ones, resulting in similar $H\alpha$ flux detection fractions. Finally, the right panel of Fig. 18 (top) shows how the size of the regions in our catalogue is typically larger than in S22. All these properties make associating regions unequivocally quite challenging, explaining the low number of matches we find. If we look only at the matching regions (Fig. 18, bottom panel), it is possible to see that the $H\alpha$ luminosity distributions are much more consistent. In contrast, the size distribution still shows that our regions typically have slightly larger sizes, confirmed by the direct comparison presented in Fig. 19. The figure shows that the nebulae from S22 are $\sim 15\%$ less bright than the ones in our catalogue. The size distribution is much more spread, with the regions in our catalogue showing, on average, circularised radii $\sim 19\%$ larger than the corresponding nebula in S22.

Finally, Fig. 18 also shows how the luminosity distribution of our regions changes when we correct the $H\alpha$ luminosity for the DIG contribution (dashed and dotted red histograms). While the top end of the distribution is not affected by the correction, generally, the shape of the distribution widens, and the peak moves towards fainter luminosities. A significant number of regions ($\sim 20\%$ for the full catalogue, $\sim 10\%$ for the

¹³ We use the $E(B - V)$ measured via the Balmer decrement with theoretical $H\alpha/H\beta = 3.03$.

matched regions) are no longer detected in $H\alpha$ after the DIG correction.

7.4. $H\text{ II}$ regions luminosity function

The $H\text{ II}$ regions LF is a fundamental tool for studying the properties of star-forming galaxies. Typically, it is described by a simple power law with a slope of ~ -2 (Elmegreen & Falgarone 1996). Variations of the steepness are often connected to both global properties of the host galaxy and local parameters that can influence the star formation process. It is possible to build the $H\text{ II}$ regions LF using a wide variety of tracers, from the UV emission (e.g. Cook et al. 2016) to the radio or infrared emission (e.g. Mascoop et al. 2021). However, the most commonly used tracer is their $H\alpha$ luminosity, since $H\alpha$ is a bright line, easy to observe with ground-based facilities.

While it is outside of the scope of this paper to discuss the $H\text{ II}$ regions LF in the PHANGS-MUSE sample accurately, a topic already discussed in detail in S22, reproducing it with our catalogue is an interesting test to confirm if our approach is a viable option for these kinds of studies. As a first step, to obtain results that are directly comparable to S22 study, we recover a new sub-sample of nebulae from our catalogue following steps similar to what they describe in their work. Firstly, we select only nebulae classified as $H\text{ II}$ regions by the model-comparison-based algorithm. Secondly, we select only nebulae with 3σ detections of both $H\alpha$ and $H\beta$ after the DIG subtraction, and we correct our luminosities using the DIG corrected Balmer decrement $E(B-V)$. To measure the $E(B-V)$, we assume a theoretical $H\alpha/H\beta$ ratio of 3.03, following Sect. 7.1. Finally, S22 exclude the regions close to the centre of the galaxies where it is difficult to deblend the contribution of very bright and nearby regions. Following their method, we use the environmental masks from Querejeta et al. (2021) to reject these regions, which are only a minor fraction of the sample. At the end of the selection process, we retain 22 947 regions, distributed across the full sample of galaxies. This means that ~ 6000 $H\text{ II}$ regions have been rejected in total, the vast majority by requesting a 3σ detection of both $H\alpha$ and $H\beta$.

We build and fit the LF following the procedure described by S22 based on the work by Clauset et al. (2009) and the `powerlaw` python package (Alstott et al. 2014), which uses a maximum-likelihood estimation method in combination with the Kolmogorov–Smirnov statistics to recover the slope α and the low luminosity cut of the LF (L_{\min} , the minimum luminosity to consider for the fit) simultaneously. In the following, we parametrise the LF as $\propto L^{-\alpha}$, so that α is always positive. The results of this analysis are shown in Table 7 (under the ‘‘Custom’’ columns) and Fig. 20.

First of all, the slopes of the LF are consistently between 1.50 and 1.80, with an average of 1.63 and a standard deviation of 0.07, in agreement with previous results in the literature (e.g. Kennicutt et al. 1989; Elmegreen & Salzer 1999) and with S22 in particular. Also considering every single galaxy, the agreement between the two works is remarkable. Most of our slopes agree within 1σ with that measured in S22, with only a few being consistent only at a 3σ level. These objects are those with the steepest slopes in S22, which can almost be considered outliers. Also, the agreement between the actual LFs is excellent, as it is possible to see from Fig. 15. The only difference is that our LF includes a larger number of regions at the low-luminosity end, which means that we are recovering a larger number of faint $H\text{ II}$ regions.

We also tested how the properties of the LF change when not applying the previously mentioned cuts and when considering directly the $H\text{ II}$ regions identified in the full catalogue and the other sub-samples described in Sect. 6.1. Table 7 shows that most of the changes to the slope and L_{\min} are negligible when moving to the $H\alpha$ sample, and the largest ones are well within 1σ . This is expected, since most of the selection applied by S22 cuts relies on the detection of the $H\alpha$ and $H\beta$ and, therefore, rejects basically the same regions. If we use the full sample of nebulae, the poorly detected regions affect the slope, which becomes flatter. When we move to the more restrictive samples (the OIII and the BPT samples), things start to change significantly. Table 7 shows how L_{\min} and α move towards higher luminosities after applying the more stringent cuts. Therefore, the importance of the bright end of the luminosity function in the fit increases and also α moves towards higher values (steeper LF).

From this analysis, we can conclude that our approach to detecting nebulae, combined with the model-comparison-based classification algorithm, produces similar, if not better, results than more traditional approaches when investigating the $H\text{ II}$ regions LF. In particular, it seems that with our approach, we can sample the low luminosity end of the LF better, and measure more consistent slopes across the galaxies of the full PHANGS-MUSE sample with respect to what is found by S22. We also see that not sampling well enough the low-luminosity end of the LF results in higher values of α .

7.5. Comparison with Scheuermann et al. (2022)

Scheuermann et al. (2022) performed a detailed search for PNe in the 19 PHANGS-MUSE galaxies to measure the galaxies’ distance via the PNe luminosity function (PNLF, Jacoby 1989; Ciardullo et al. 2002). They used traditional techniques (point-source detection, accurate aperture photometry) to identify and classify nebulae starting from [O III] $\lambda 5007$ maps at native resolution (see Spriggs et al. 2020; Roth et al. 2021; Scheuermann et al. 2022, for a detailed description of the procedures). The search resulted in a catalogue of 1049 sources, 899 of which have been confirmed as PNe while the remaining 150 have been classified as SNR. The number is comparable to the number of nebulae we classify as PNe using our model-comparison-based algorithm in the full sample (796, Table 5).

To understand if there is a direct correspondence between the PNe identified by the two approaches, we match the two catalogues following the procedure developed in Sect. 7.3. Since ‘segmentation maps’ are not provided along with the Scheuermann et al. (2022) catalogue, we start by building them by positioning at the coordinates of each nebula a circular mask representing the area used in the work to perform aperture photometry. Then, we proceed as in Sect. 7.3 but requiring less overlap to identify matches, since we expect our nebulae to be much more extended than the point-like sources in Scheuermann et al. (2022) catalogue. In particular, we consider a match if one of the nebulae in Scheuermann et al. (2022) shares at least 50% of its area with one of our nebulae and 10% of our nebula is shared with Scheuermann et al. (2022) nebula, or the overlap is higher than 90% in any direction.

Matching the catalogue results in 567 (54.1%) of the Scheuermann et al. (2022) regions having a counterpart in our own catalogue, among which 232 have a ‘‘perfect match’’ (overlap >0.6 on both directions). The low yield seems to arise principally from the different approaches to nebulae detection. Scheuermann et al. (2022) performed a detailed search for this kind of object by directly examining the [O III] $\lambda 5007$ emission

Table 7. Values of the slope of the LF and L_{\min} recovered by fitting a simple power-law to the data.

Galaxy	α						L_{\min}					
	Santoro	Custom	Full	H α	OIII	BPT	Santoro	Custom	Full	H α	OIII	BPT
IC 5332	1.82±0.13	1.72±0.11	1.66±0.09	1.71±0.10	1.82±0.13	1.90±0.14	36.25±0.19	36.25±0.03	36.12±0.04	36.23±0.03	36.62±0.04	36.79±0.05
NGC 0628	1.71±0.10	1.60±0.08	1.50±0.08	1.58±0.09	1.70±0.10	1.76±0.11	36.63±0.08	36.62±0.03	36.28±0.03	36.56±0.03	37.08±0.03	37.29±0.03
NGC 1087	1.71±0.10	1.58±0.09	1.46±0.09	1.57±0.09	1.64±0.09	1.77±0.12	37.49±0.10	37.41±0.06	37.00±0.06	37.38±0.05	37.69±0.06	37.92±0.05
NGC 1300	1.77±0.12	1.67±0.10	1.64±0.09	1.65±0.09	1.73±0.11	1.78±0.12	37.08±0.25	37.04±0.02	36.98±0.03	37.04±0.02	37.38±0.03	37.52±0.03
NGC 1365	1.74±0.11	1.57±0.09	1.45±0.08	1.52±0.08	1.57±0.09	1.65±0.09	37.82±0.20	37.38±0.03	36.99±0.04	37.36±0.03	37.65±0.05	37.88±0.04
NGC 1385	1.52±0.08	1.49±0.08	1.38±0.08	1.47±0.09	1.57±0.09	1.64±0.09	37.29±0.17	37.49±0.06	36.90±0.08	37.42±0.06	37.82±0.06	38.05±0.07
NGC 1433	1.96±0.16	1.71±0.10	1.65±0.09	1.69±0.10	1.75±0.11	1.86±0.13	37.32±0.14	36.84±0.02	36.72±0.02	36.82±0.02	37.13±0.03	37.38±0.02
NGC 1512	2.04±0.17	1.76±0.11	1.61±0.08	1.70±0.10	1.73±0.11	1.80±0.12	37.41±0.10	37.06±0.03	36.86±0.04	37.05±0.03	37.25±0.04	37.48±0.04
NGC 1566	1.57±0.09	1.54±0.08	1.44±0.08	1.53±0.08	1.58±0.09	1.67±0.10	37.14±0.14	37.21±0.03	36.74±0.04	37.15±0.02	37.51±0.04	37.78±0.03
NGC 1672	1.67±0.10	1.63±0.09	1.46±0.08	1.58±0.09	1.68±0.10	1.70±0.10	37.55±0.12	37.49±0.03	37.01±0.04	37.45±0.04	37.78±0.03	37.93±0.03
NGC 2835	1.76±0.11	1.61±0.08	1.46±0.08	1.59±0.08	1.76±0.11	1.78±0.12	37.08±0.10	36.98±0.04	36.45±0.06	36.94±0.04	37.37±0.03	37.50±0.04
NGC 3351	1.98±0.16	1.76±0.11	1.64±0.09	1.72±0.11	1.69±0.10	1.76±0.11	36.84±0.13	36.56±0.02	36.38±0.04	36.53±0.03	36.73±0.05	37.03±0.04
NGC 3627	1.57±0.09	1.55±0.08	1.38±0.08	1.53±0.08	1.59±0.09	1.64±0.09	37.43±0.23	37.47±0.04	36.74±0.06	37.39±0.05	37.74±0.04	38.03±0.06
NGC 4254	1.61±0.08	1.64±0.09	1.47±0.09	1.63±0.09	1.71±0.10	1.78±0.12	37.36±0.12	37.53±0.02	37.04±0.04	37.50±0.03	37.73±0.03	37.89±0.03
NGC 4303	1.56±0.09	1.60±0.08	1.42±0.08	1.57±0.09	1.69±0.10	1.74±0.11	37.26±0.15	37.49±0.03	36.88±0.04	37.43±0.03	37.88±0.03	38.04±0.03
NGC 4321	1.95±0.15	1.67±0.10	1.49±0.08	1.62±0.08	1.71±0.10	1.79±0.12	37.95±0.27	37.42±0.03	36.96±0.04	37.37±0.02	37.72±0.04	37.90±0.04
NGC 4535	1.63±0.09	1.60±0.08	1.52±0.08	1.58±0.09	1.69±0.10	1.75±0.11	36.79±0.09	36.85±0.03	36.58±0.04	36.79±0.03	37.30±0.03	37.54±0.04
NGC 5068	1.58±0.09	1.59±0.08	1.52±0.08	1.58±0.09	1.66±0.09	1.75±0.11	35.92±0.06	36.20±0.04	35.85±0.05	36.18±0.04	36.57±0.04	36.81±0.05
NGC 7496	1.69±0.10	1.61±0.08	1.56±0.09	1.60±0.08	1.69±0.10	1.77±0.12	37.16±0.09	37.09±0.04	36.91±0.04	37.08±0.04	37.52±0.06	37.68±0.05

Notes. The results recovered by S22 are reported in the ‘Santoro’ column as a reference.

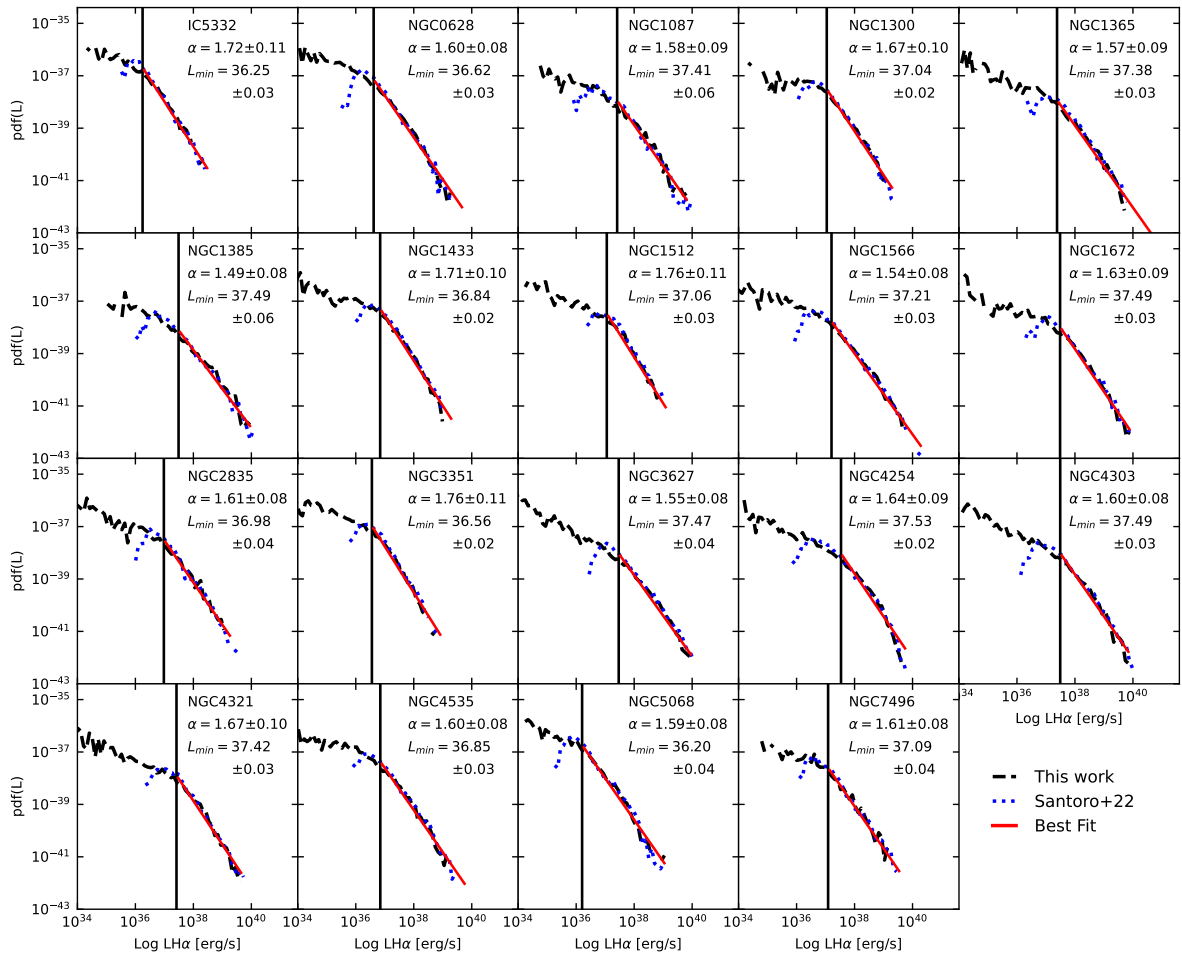


Fig. 20. H II regions LF for each galaxy in the PHANGS-MUSE sample. Only objects satisfying the selection described in Sect. 7.4 have been used to compute the LF. The black dashed lines represent the observed LF, the blue dotted lines represent the LF from S22, the red line is the fit obtained from our data, and finally, the black vertical line is the low luminosity cut (L_{\min}). Only regions with $L > L_{\min}$ are considered when fitting the LF.

line map, significantly reducing crowding from other types of nebulae not particularly bright in [O III] λ 5007. On the other hand, we produced a detection map by combining multiple line maps (Sect. 4). This can blend the emission of nebulae in different lines, making the detection of faint PNe in crowded areas more challenging or even impossible (Fig. 21). On top of this, from a visual inspection, we estimate that $\sim 10\%$ of the regions detected by Scheuermann et al. (2022) are not visible in the [O III] λ 5007 homogenised maps we are using in this work (Fig. 21, right panel). This could either result from the convolution, which spreads out the emission of faint nebulae, making them undetectable, or they are false detections.

We now compare how the nebulae included in both catalogues are classified by the two different approaches. Table 8 shows that half of the sample ($\sim 47\%$) of Scheuermann et al. (2022) PNe retains the same classification in our compilation. Most of the remaining nebulae are either considered H II regions (21.5%) or shocks (18.6%) by the model-comparison-based algorithm. A lower fraction of objects fall in the ambiguous category (12.5%), and only one object is unclassified. The agreement between the model-comparison-based algorithm and the Scheuermann et al. (2022) classification for PNe is much better than what we see between our model-comparison-based classification and our application of the traditional criterion for PNe. This suggests that the careful analysis performed by Scheuermann et al. (2022) to ensure an accurate measurement of the [O III] λ 5007 fluxes and the point-like nature of the emission is essential for the traditional classification criterion to perform well. However, the disagreement between the two approaches is still substantial. The two methods do not agree for more than 50% of the sample of PNe identified by Scheuermann et al. (2022). On the other hand, $\sim 73\%$ of Scheuermann et al. (2022) SNRs are classified as shocks by our model-comparison-based algorithm.

In summary, on one side, the detection images we are using in this work prevent the detection of faint nebulae like PNe in crowded areas since their emission is blended with that of other regions bright in other lines. On the other side, the two classification methods seem to not agree on identifying what is a PN, even though the agreement between the two paradigms is better than what is seen in Sect. 6.2. This discrepancy does not necessarily mean that our new classification criterion is failing, since the nebulae it classifies as PNe are located in the regions of the diagnostic diagrams where they are expected to be, and a visual inspection of the data reveals that these objects look as expected: unresolved [O III] λ 5007 bright nebulae.

7.6. Planetary nebulae luminosity function

Extragalactic PNe are often used for measuring the distance of their host via the PNLF (Jacoby 1989; Ciardullo et al. 2002). Although the sample of PNe recovered in this work is different from that compiled in Scheuermann et al. (2022), the nebulae we identify typically show the characteristics of PNe. Therefore we can build the PNLF, estimate the distance of the galaxies in our sample and compare our results with the measurement from Scheuermann et al. (2022) to evaluate if our approach is a viable option for these kinds of studies.

To reproduce the PNLF, we followed the procedure described in Scheuermann et al. (2022) to enable a direct comparison between the results. We use the Eq. (4) (Jacoby 1989) and the distance modulus definition (Eq. (5)) to convert the nebulae DIG corrected, line flux ($I_{[\text{OIII}]}$) in an absolute magnitude. The

Table 8. Comparison between Scheuermann et al. (2022) PNe and SNR sample and our catalogue.

Class	PNe	SNR
H II regions	109 (21.6)	4 (6.5)
PNe	238 (47.1)	6 (9.6)
Shocks	94 (18.6)	45 (72.6)
Ambiguous	63 (12.5)	4 (6.5)
Unclassified	1 (0.2)	3 (4.8)
Undetected	394	88
Total	899	150

Notes. Each column shows how the objects in each of the two Scheuermann et al. (2022) samples are classified by our classification algorithm. Both the absolute number and the percentages (in brackets) are shown.

Table 9. Estimated distances of the galaxies where we could fit the PNLF compared with the results from Scheuermann et al. (2022).

Galaxy	N	μ mag	D Mpc	μ_{Sc22} mag	D_{Sc22} Mpc
IC 5332	12	$30.19^{+0.04}_{-0.26}$	$10.94^{+0.20}_{-1.33}$	$29.73^{+0.10}_{-0.20}$	$8.84^{+0.39}_{-0.82}$
NGC 0628	40	$29.87^{+0.02}_{-0.16}$	$9.43^{+0.08}_{-0.69}$	$29.89^{+0.06}_{-0.09}$	$9.52^{+0.26}_{-0.41}$
NGC 1433	13	$31.38^{+0.02}_{-0.13}$	$18.84^{+0.13}_{-1.16}$	$31.39^{+0.04}_{-0.07}$	$18.94^{+0.39}_{-0.56}$
NGC 1566	14	$31.49^{+0.03}_{-0.19}$	$19.85^{+0.26}_{-1.77}$	$31.13^{+0.08}_{-0.17}$	$16.84^{+0.60}_{-1.29}$
NGC 2835	16	$30.54^{+0.05}_{-0.24}$	$12.81^{+0.29}_{-1.44}$	$30.57^{+0.08}_{-0.17}$	$13.03^{+0.46}_{-1.04}$
NGC 3351	52	$30.40^{+0.04}_{-0.15}$	$12.00^{+0.25}_{-0.84}$	$30.36^{+0.06}_{-0.08}$	$11.80^{+0.31}_{-0.43}$
NGC 3627	16	$30.59^{+0.06}_{-0.24}$	$13.12^{+0.34}_{-1.45}$	$30.18^{+0.08}_{-0.15}$	$10.88^{+0.39}_{-0.77}$
NGC 4303	14	$30.85^{+0.04}_{-0.27}$	$14.82^{+0.25}_{-1.85}$	$30.65^{+0.10}_{-0.26}$	$13.49^{+0.64}_{-1.60}$
NGC 4535	16	$31.75^{+0.03}_{-0.14}$	$22.41^{+0.30}_{-1.47}$	$31.43^{+0.06}_{-0.09}$	$19.29^{+0.56}_{-0.82}$
NGC 5068	16	$28.57^{+0.10}_{-0.33}$	$5.17^{+0.24}_{-0.78}$	$28.46^{+0.11}_{-0.26}$	$4.93^{+0.24}_{-0.59}$

Notes. The columns report: the name of the galaxy, the number of PNe available for the fit, the distance modulus recovered with error, the distance of the galaxy in Mpc, the distance modulus recovered with error from Scheuermann et al. (2022), the distance of the galaxy in Mpc from Scheuermann et al. (2022).

fluxes are corrected for Milky Way extinction, but not for internal extinction.

$$\mu = m - M = 5 \log_{10}(D) - 5. \quad (5)$$

At this point, we fit the PNLF with the procedure described in Scheuermann et al. (2022) and using the following functional form from Ciardullo et al. (1989):

$$N(M_{[\text{OIII}]}) \propto e^{0.307M_{[\text{OIII}]}} \left(1 - e^{3(M^* - M_{[\text{OIII}]})}\right). \quad (6)$$

M^* is the zero-point of the luminosity function, and we assume it to be -4.47 .

The results are presented in Fig. 22 while in Table 9 we directly compare our findings with the results from Scheuermann et al. (2022). Only 10 out of 19 galaxies contain enough bright PNe to fit the PNLF. In all cases, our measurements are within 3σ from what is reported by Scheuermann et al. (2022), with six galaxies being consistent within 1σ . We notice, however, that we generally measure larger distances.

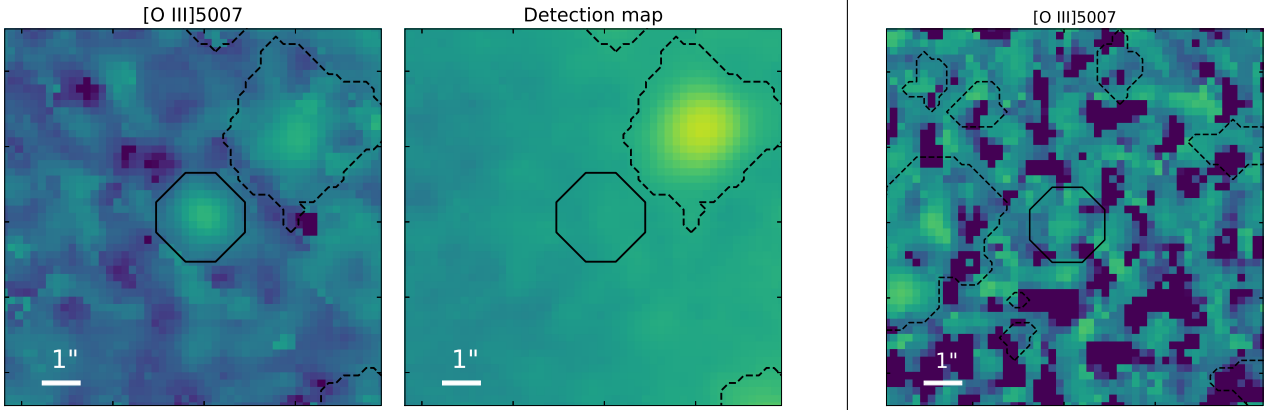


Fig. 21. Examples of nebulae in [Scheuermann et al. \(2022\)](#) not included in our catalogue. Left: detail of NGC 1365 (region 10 in [Scheuermann et al. 2022](#)) in the [O III] λ 5007 emission line map (left) and our detection map (right). In both plots, the solid black contours show the position of the nebulae detected in [Scheuermann et al. \(2022\)](#) catalogue, while the dashed contours show the position of the nebulae in our catalogue. This figure shows how the nebula, clearly detected in the single line map, is not detectable in our detection map since it is blended with the relatively high background. 1'' in NGC 1365 corresponds to ~ 100 pc. Right: detail of the [O III] λ 5007 emission line map of IC 5332 (region 56 in [Scheuermann et al. 2022](#)). The solid black contours show the position of one of the nebulae detected in [Scheuermann et al. \(2022\)](#) catalogue, while the dashed contours show the position of the nebulae in our catalogue. 1'' in IC 5332 corresponds to ~ 45 pc.

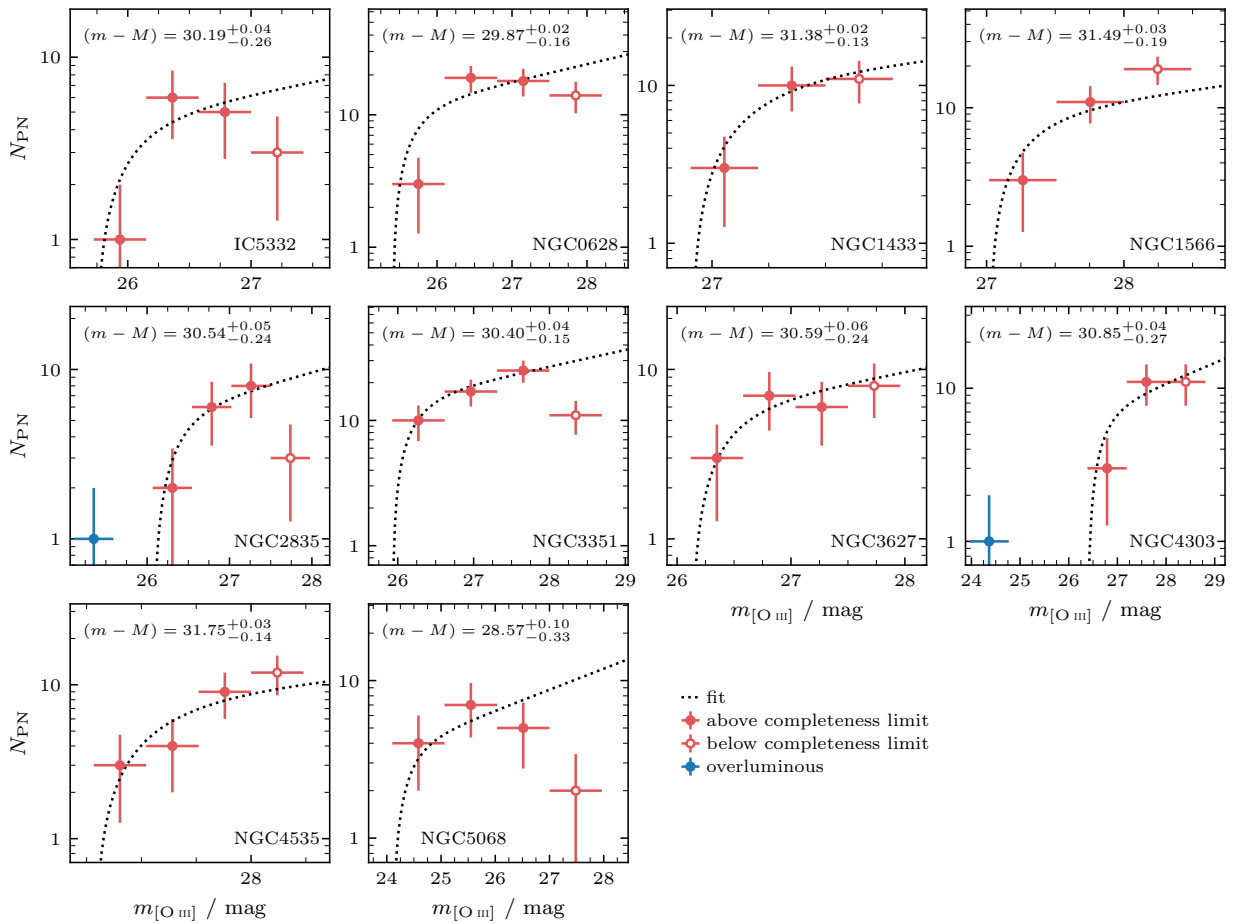


Fig. 22. PNLF for the galaxies where it was possible to perform the fit. The points with errors represent the observed PNLF, while the dotted line is the result of the fit. Only the points coloured in solid orange have been used in the fitting routine. The empty points are PNe below the completeness limit, while blue points represent discarded overluminous objects.

We also compare our distances with the work from [Anand et al. \(2021\)](#), which provides a compilation of distances to PHANGS galaxies based on different methods (mostly tip of the red giant branch and Cepheids). Also in this case we see a similar pattern, with most of our measurements being consistent within 3σ with their results. The exception is NGC 4535, where the Cepheid-based distance reported in [Anand et al. \(2021\)](#) is significantly lower also compared to [Scheuermann et al. \(2022\)](#) results.

In summary, we can estimate a distance for just half of the galaxies in our sample, but our results generally agree with what is found in the literature and by [Scheuermann et al. \(2022\)](#) in particular. The problem arises from the low number of bright PNe detected, and, therefore, from their segmentation more than their classification, as discussed in Sect. 7.5. While running a dedicated search for PNe is better, our approach is still useful to clean a sample of PNe from other sources. Applying the segmentation algorithm only to the $[\text{O III}]\lambda 5007$ line map before applying the classification algorithm might produce better results.

7.7. Shocks and supernova remnants

So far, we have identified all of the nebulae that can be described by the [Allen et al. \(2008\)](#) models as shocks and not as SNRs since we expect this category to be somewhat contaminated by the presence of AGN ionised regions. However, this contamination can only occur if the galaxy hosts an AGN. Four galaxies in our sample are known to host an AGN (NGC 1365, NGC 1512, NGC 1566 and NGC 1672, e.g. [Véron-Cetty & Véron 2006](#); [Belfiore et al. 2022](#)), but only NGC 1365 is known to host an extended region of AGN-ionised gas (also known as ionisation cones, e.g. [Storchi-Bergmann & Bonatto 1991](#); [Venturi et al. 2018](#)). The other galaxies show some extended AGN- or LIER-like (low ionisation emission region) emission outside the nucleus, but the low equivalent width of $\text{H}\alpha$ seems to suggest that these are likely regions filled by DIG, which is ionised by the hard radiation produced by hot low-mass evolved stars (HOLMES, [Belfiore et al. 2022](#)). Such behaviour is also observed in a few other galaxies with no known AGN in their nucleus (NGC 1300, NGC 1433, NGC 3351, NGC 3627). The AGN-SNR degeneracy problem can be significant in NGC 1365 and in the very central regions of the other three known AGN, but it should not produce significant contamination in all the other objects. In fact, the top panel of Fig. 23 shows how there is an over-density of bright shock-ionised nebulae along the ionisation cones of the galaxy, which are easily visible in the $[\text{O III}]\lambda 5007$ line map shown in the figure, while no ionisation cones nor over-densities of shocks nebulae are observed in another AGN, for example NGC 1672. We can conclude that while there is contamination by a certain number of AGN-ionised regions, it is limited to NGC 1365 and the nucleus of a few other galaxies.

Unfortunately, SNRs (and shocks) are the least well-studied class of sources in the extragalactic environment. There are only a handful of SNRs catalogues in galaxies outside the Local Group, and none of them covers any of the galaxies included in the PHANGS-MUSE survey. As a consequence, we cannot perform the same analysis we did for PNe and H II regions to understand how well our segmentation and classification algorithms are performing. However, we can compare the number of SNRs we find in our galaxies with what is available in the literature. We consider five recent catalogues, each one focused on a different, nearby, main sequence galaxy: M31 ([Lee & Lee 2014a](#)), M33 ([Lee & Lee 2014b](#)), M83 ([Long et al. 2022](#)), NGC 3344 ([Moumen et al. 2019](#)) and NGC 4030 ([Cid Fernandes](#)

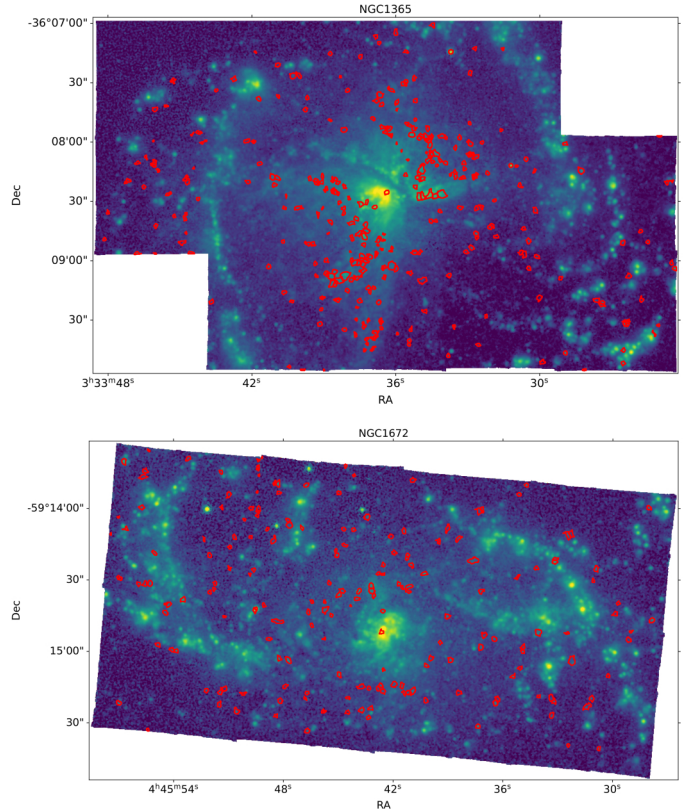


Fig. 23. Location of the regions classified as shocks in NGC 1365 (top) and NGC 1672 (bottom) on top of $[\text{O III}]\lambda 5007$ line maps.

[et al. 2021](#)). The first two used narrow-band images to identify SNR candidates using the $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha$ ratio together with some additional criteria based on their expected shell-like morphology and on the presence or absence of blue stars at the candidate position (e.g. [Lee & Lee 2014a,b](#)). [Long et al. \(2022\)](#) uses emission-line maps from the MUSE mosaic produced by [Della Bruna et al. \(2022\)](#) and other narrow-band images from Magellan and HST to identify SNR candidates, that are then classified using a dedicated analysis of their spectra acquired by MUSE or other instruments (e.g. GMOS). [Moumen et al. \(2019\)](#) used SITELE, the imaging Fourier Transform spectrograph of the Canada-France-Hawaii Telescope (CFHT), to obtain emission-line maps of the galaxy in multiple emission lines. They selected the SNRs candidates using a combination of criteria: $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha > 0.4$, high S/N for the two $[\text{S II}]\lambda\lambda 6717, 6731$ lines, size of the region < 120 pc and finally they require concentrated regions based on their pseudo-Voigt profile. Finally, [Cid Fernandes et al. \(2021\)](#) used a principal component analysis on a sample of nebulae recovered via the analysis of MUSE data, combined with the $[\text{S II}]\lambda\lambda 6717, 6731/\text{H}\alpha$ vs. $[\text{N II}]\lambda 6584/\text{H}\alpha$ diagram from [Kopsacheili et al. \(2020\)](#), to identify a sample of 59 SNRs. Following the described criteria, [Lee & Lee \(2014a\)](#) report 156 SNR in M31, [Lee & Lee \(2014b\)](#) 199 SNR in M33, [Long et al. \(2022\)](#) 366 in M83, [Moumen et al. \(2019\)](#) 129 in NGC 3344¹⁴ and [Cid Fernandes et al. \(2021\)](#) 59 in NGC 4030. In Table 10, we show the number of SNR for each galaxy in our sample. It is possible to see that we find,

¹⁴ We are considering the total number of SNR candidates identified by [Moumen et al. \(2019\)](#). However, they further distinguish between “confirmed” SNRs (42), “probable” SNRs (45) and “less likely” SNRs (42).

Table 10. SNRs included in our sample per each galaxy.

Galaxy	Full	H α	OIII	BPT
IC 5332	184	142	139	101
NGC 0628	344	197	194	111
NGC 1087	131	98	97	73
NGC 1300	117	98	88	45
NGC 1385	136	97	104	77
NGC 1433	349	276	254	119
NGC 1512	121	89	93	33
NGC 1566	496	302	324	156
NGC 1672	210	128	123	74
NGC 2835	286	214	204	157
NGC 3351	205	149	140	67
NGC 3627	349	206	197	83
NGC 4254	361	207	226	115
NGC 4303	424	213	196	99
NGC 4321	289	134	142	56
NGC 4535	250	160	162	73
NGC 5068	464	383	367	304
NGC 7496	78	55	55	38

Notes. NGC 1365 is not shown here, since we expect its sample to be heavily contaminated.

on average, 271 SNR per galaxy, with a maximum of 496 in NGC 1566 and a minimum of 78 in NGC 7496. This is perfectly in line with what is reported by the literature, especially considering the different observation techniques, depth of the data, different surveyed areas and detection methods. These numbers are also comparable to the 294 known Galactic SNRs (Green 2019), most of which have been identified through their radio emission. If we consider only the BPT sub-sample, which should contain only the most reliable detections, we detect ~ 100 SNRs per galaxy on average, still on the same order of magnitude as what is found in the literature. We can also compare the distribution of the H α luminosity for our sample (Fig. 17) of SNRs with that reported by Lee & Lee (2014b) for their SNR candidates in M31 and M33. Both distributions peak around 10^{36} erg s $^{-1}$ and they have a similar shape, even though it is clear that we have much better statistics. Also, our data are significantly deeper, since we have a significant amount of nebulae below 10^{35} erg s $^{-1}$ even when we apply the more restrictive cuts, while they barely report any below this limit. We also have a certain number of brighter nebulae, which can be connected to multiple aspects. On one hand, the sample size definitely increases the possibility of finding brighter objects. On the other hand, the different methods used to define the size of the nebulae, especially considering that they are observing much closer galaxies, with higher spatial resolutions, and the background subtraction, can influence the measurement of the H α luminosity.

Finally, we must note that, in a few objects, the spatial distribution of the shock classified regions is peculiar. The atlas presented in Appendix B, shows that in some galaxies (e.g. NGC 3627) these nebulae fill the interarm region while in other objects (e.g. NGC 1433, NGC 3351) there is a peculiar over-density of shock classified regions just outside of the galaxies' bulge. These seem regions permeated by the DIG emission, which shows line ratios typical of LI(N)ERS (Belfiore et al. 2022) and similar to what is observed in SNRs and shocks. It is possible, therefore, that these circumnuclear or interarm regions are not SNRs, but particularly bright DIG clouds and that they

can constitute a significant source of contamination for the shock sample. We also have to consider that other mechanisms can produce shocks in the ISM (e.g. bars: Athanassoula 1992, nuclear spirals: Kim & Elmegreen 2017), so it is possible that some of the nebulae are shock ionised by these kinds of processes. For example, Fig. B.4 clearly shows some spiral structures outside of the circumnuclear ring where the vast majority of the nebulae are classified as shocks.

While it is clear that this analysis cannot give us a picture as complete as we have for the other types of nebulae, it still suggests that we are recovering a sample of SNRs which is reasonably comparable with what is found by other works in the literature, both in size of the sample (for each galaxy) and H α luminosity distribution, even though there might be some source of contamination (bright DIG clouds in particular). A detailed analysis of our shock or SNRs sample will definitely be interesting since the literature on these objects is relatively scarce, but it is outside the scope of this work. An ongoing work will construct tailored SNR catalogues using classical methods (Li et al., in prep) and will be well suited for comparison with our catalogue of shock-dominated objects.

8. Conclusions

This work presents a new catalogue of ionised nebulae in nearby galaxies, containing spectral (line fluxes, kinematics) and spatial (position, shape) information for over 40 000 nebulae distributed across the 19 galaxies of the PHANGS-MUSE sample. The nebulae have been classified using a new approach based on comparing their observed properties with grids of models representing the three main classes of ionised nebulae found in galaxies: H II regions, PNe and shock-ionised nebulae. This makes our catalogue the largest compilation of nebulae with a reliable classification and a uniform set of information (independent from their classification) available in the literature to date.

The algorithm exploits the principle of the odds ratio to assign to each nebula the probability it belongs to each one of the three considered classes. Thanks to its comparative and probabilistic nature, the algorithm overcomes some of the main limitations characterising the criteria traditionally used in the literature to classify nebulae, such as their binarity, the sharp limits used for the classification and the limited amount of information that they rely on to perform the classification. Given the large volume of data on ionised nebulae that will be produced in the near future by instruments like MUSE and surveys like the Local Volume Mapper survey, this algorithm is an important step toward more complex and robust automatic approaches to the classification of any type of nebulae.

The analysis of the catalogue and the comparison between the results of the model-comparison-based classification with what can be obtained from the traditional criteria reveals that:

- The DIG emission can significantly alter the observed line ratios of the nebulae, particularly in the faintest ones where the DIG emission is dominant. A DIG correction is therefore essential to classify the nebulae correctly;
- The model-comparison-based classification algorithm works remarkably well for H II regions, as shown by the analysis of the H II regions LF in Sect. 7.4. Also, the agreement between the traditional criteria and the new approach is good for this class of objects, but the model-comparison-based algorithm can identify and classify H II regions in areas of the BPT diagrams that are typically ignored by the traditional criteria or in competition with other classes of nebulae;

- A comparison between the $E(B - V)$ measured via the Balmer decrement and that estimated by IZI during the evidence estimation for H II regions revealed that the extinction is systematically overestimated by ~ 0.05 mag when applying the Balmer decrement technique to DIG corrected fluxes. The origin of this discrepancy arises from the theoretical $H\alpha/H\beta$ ratio assumed during the computation. The typical value for standard conditions (low-density limit, $T_e \sim 10\,000$ K, Case B) $H\alpha/H\beta = 2.86$ is not representative of the true $H\alpha/H\beta$ ratio according to the model grid we are considering. A theoretical ratio of 3.03 (Case B conditions, low-density limit, $T_e \sim 5200$ K) seems to minimise the offset between the $E(B - V)$ measured by IZI and that recovered via the Balmer decrement;
- We also observe that the agreement, and good performance of non-DIG corrected Balmer decrement extinction correction, seems to arise from a coincidence. The underestimation of the extinction due to DIG contamination and its overestimation due to adopting the wrong intrinsic Balmer ratio, produces biases of similar magnitudes but act in opposite directions, cancelling each other;
- The sample of PNe we recover seems to be incomplete when compared with other works in the literature studying the same galaxies (Scheuermann et al. 2022). However, the issue seems to be mostly related to limitations of the source detection method, while the classification of the detected objects appears robust;
- For 10 out of 19 galaxies, we were able to fit the PNLF, and the results are comparable with the literature, and Scheuermann et al. (2022) in particular;
- We identify a rather large population of nebulae classified as shocks (10–15% of the sample). Despite for a few contaminants, like AGN-ionised nebulae in a few galaxies of the sample bright DIG clouds, and dynamically shocked regions, we expect this to be the largest compilation of extragalactic SNRs available in the literature to date. This catalogue will open the possibility for a systematic and detailed study of this interesting, but still relatively poorly understood population of objects.

Concluding, we showed in this work that our model-comparison-based algorithm is an invaluable tool that can significantly simplify and objectify the classification and characterisation of large samples of ionised nebulae. Moreover, with its number and variety of nebulae, our catalogue opens the doors to a wide variety of further studies, but also to the possibility to develop even more automatised, machine learning-powered, identification and classification algorithms.

Acknowledgements. The authors of the paper want to thank the anonymous referee for their helpful suggestions. Based on observations collected at the European Southern Observatory under ESO programmes 1100.B-0651, 095.C-0473, and 094.C-0623 (PHANGS-MUSE; PI: Schinnerer), as well as 094.B-0321 (MAGNUM; PI: Marconi), 099.B-0242, 0100.B-0116, 098.B-0551 (MAD; PI: Carollo) and 097.B-0640 (TIMER; PI: Gadotti). Science-level MUSE mosaicked datacubes and high-level analysis products are provided via the ESO archive phase 3 interface (https://archive.eso.org/scienceportal/home?data_collection=PHANGS). A full description of the first PHANGS data release is presented in Emsellem et al. (2022). This work was carried out as part of the PHANGS collaboration. E.C. and G.A.B. acknowledge support from ANID Basal projects ACE210002 and FB210003. K.K., F.S., and O.E. gratefully acknowledge funding from the German Research Foundation (DFG) in the form of an Emmy Noether Research Group (grant number KR4598/2-1, PI Kreckel). H.A.P. acknowledges support by the National Science and Technology Council of Taiwan under grant 110-2112-M-032-020-MY3. S.C.O.G. acknowledges funding from the European Research Council via the ERC Synergy Grant “ECOGAL – Understanding our Galactic ecosystem: From the disk of the Milky Way to the formation sites of stars and planets” (project ID 855130) and from

the Heidelberg Cluster of Excellence (EXC 2181 - 390900948) “STRUCTURES: A unifying approach to emergent phenomena in the physical world, mathematics, and complex data”, funded by the German Excellence Strategy. K.K., E.J.W. and S.C.O.G. acknowledge support from the Deutsche Forschungsgemeinschaft (DFG) via the Collaborative Research Center (SFB 881 – 138713538) “The Milky Way System” (subprojects A1, B1, B2 and B8, P1 and P2). E.S. acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 694343). F.B. acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No.726384/Empire) The Starlink software (Currie et al. 2014) is currently supported by the East Asian Observatory. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration 2013, 2018), Matplotlib (Hunter 2007), NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), Scikit-learn (Pedregosa et al. 2011).

References

- Acker, A., Marcout, J., Ochsenbein, F., et al. 1992, *The Strasbourg-ESO Catalogue of Galactic Planetary Nebulae, Parts I, II* (Garching: European Southern Observatory)
- Allen, M. G., Groves, B. A., Dopita, M. A., Sutherland, R. S., & Kewley, L. J. 2008, *ApJS*, **178**, 20
- Alstott, J., Bullmore, E., & Plenz, D. 2014, *PLoS One*, **9**, e85777
- Anand, G. S., Lee, J. C., Van Dyk, S. D., et al. 2021, *MNRAS*, **501**, 3621
- Anders, E., & Grevesse, N. 1989, *Geochim. Cosmochim. Acta*, **53**, 197
- Anderson, L. D., Bania, T. M., Balsler, D. S., et al. 2014, *ApJS*, **212**, 1
- Arnaboldi, M., Freeman, K. C., Mendez, R. H., et al. 1996, *ApJ*, **472**, 145
- Arnaboldi, M., Freeman, K. C., Gerhard, O., et al. 1998, *ApJ*, **507**, 759
- Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, **558**, A33
- Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, **156**, 123
- Asvarov, A. I. 2014, *A&A*, **561**, A70
- Athanassoula, E. 1992, *MNRAS*, **259**, 345
- Azimlu, M., Marciniak, R., & Barnby, P. 2011, *AJ*, **142**, 139
- Baade, W. 1955, *AJ*, **60**, 151
- Baade, W., & Arp, H. 1964, *ApJ*, **139**, 1027
- Bacon, R., Accardo, M., Adjali, L., et al. 2010, *SPIE Conf. Ser.*, **7735**, 773508
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, **93**, 5
- Belfiore, F., Santoro, F., Groves, B., et al. 2022, *A&A*, **659**, A26
- Bennett, A. S. 1963, *MNRAS*, **127**, 3
- Berg, D. A., Pogge, R. W., Skillman, E. D., et al. 2020, *ApJ*, **893**, 96
- Berry, D. S., Reinhold, K., Jenness, T., & Economou, F. 2007, *ASP Conf. Ser.*, **376**, 425
- Bhattacharya, S., Arnaboldi, M., Hartke, J., et al. 2019, *A&A*, **624**, A132
- Bittner, A., Falcón-Barroso, J., Nedelchev, B., et al. 2019, *A&A*, **628**, A117
- Blair, W. P., Chander, R., Dopita, M. A., et al. 2014, *ApJ*, **788**, 55
- Blanc, G. A., Heiderman, A., Gebhardt, K., Evans, Neal J., I., & Adams, J. 2009, *ApJ*, **704**, 842
- Blanc, G. A., Kewley, L., Vogt, F. P. A., & Dopita, M. A. 2015, *ApJ*, **798**, 99
- Boeshaar, G. O., Boeshaar, P. C., Czyzak, S. J., Aller, L. H., & Lasker, B. M. 1980, *Ap&SS*, **68**, 335
- Bojičić, I. S., Filipović, M. D., Urošević, D., Parker, Q. A., & Galvin, T. J. 2021, *MNRAS*, **503**, 2887
- Bolton, J. G., Stanley, G. J., & Slee, O. B. 1949, *Nature*, **164**, 101
- Bradley, T. R., Knapen, J. H., Beckman, J. E., & Folkes, S. L. 2006, *A&A*, **459**, L13
- Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, **351**, 1151
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, **345**, 245
- Ciardullo, R., Jacoby, G. H., Ford, H. C., & Neill, J. D. 1989, *ApJ*, **339**, 53
- Ciardullo, R., Feldmeier, J. J., Jacoby, G. H., et al. 2002, *ApJ*, **577**, 31
- Cid Fernandes, R., Carvalho, M. S., Sánchez, S. F., de Amorim, A., & Ruschel-Dutra, D. 2021, *MNRAS*, **502**, 1386
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. 2009, *SIAM Rev.*, **51**, 661
- Cook, D. O., Dale, D. A., Lee, J. C., et al. 2016, *MNRAS*, **462**, 3766
- Cox, D. P. 2005, *ARA&A*, **43**, 337
- Currie, M. J., Berry, D. S., Jenness, T., et al. 2014, *ASP Conf. Ser.*, **485**, 391
- D’Agostino, J. J., Kewley, L. J., Groves, B. A., et al. 2019a, *MNRAS*, **485**, L38
- D’Agostino, J. J., Kewley, L. J., Groves, B. A., et al. 2019b, *MNRAS*, **487**, 4153
- Davies, R. L., Groves, B., Kewley, L. J., et al. 2017, *MNRAS*, **470**, 4974
- Delgado-Inglada, G., Morisset, C., & Stasińska, G. 2014, *MNRAS*, **440**, 536
- Della Bruna, L., Adamo, A., Bik, A., et al. 2020, *A&A*, **635**, A134
- Della Bruna, L., Adamo, A., Bik, A., et al. 2021, *A&A*, **652**, A6
- Della Bruna, L., Adamo, A., Amram, P., et al. 2022, *A&A*, **660**, A77
- Diaz, A. I., Terlevich, E., Vilchez, J. M., Pagel, B. E. J., & Edmunds, M. G. 1991, *MNRAS*, **253**, 245
- D’odorico, S., Benvenuti, P., & Sabbadin, F. 1978, *A&A*, **63**, 63

- D'Odorico, S., Dopita, M. A., & Benvenuti, P. 1980, *A&AS*, **40**, 67
- Dopita, M. A., Blair, W. P., Long, K. S., et al. 2010, *ApJ*, **710**, 964
- Douglas, N. G., Arnaboldi, M., Freeman, K. C., et al. 2002, *PASP*, **114**, 1234
- Elmegreen, B. G., & Falgarone, E. 1996, *ApJ*, **471**, 816
- Elmegreen, B. M., & Salzer, J. J. 1999, *AJ*, **117**, 764
- Emsellem, E., Schinnerer, E., Santoro, F., et al. 2022, *A&A*, **659**, A191
- Espinosa-Ponce, C., Sánchez, S. F., Morisset, C., et al. 2020, *MNRAS*, **494**, 1622
- Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, *Rev. Mex. Astron. Astrofis.*, **49**, 137
- Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, *Rev. Mex. Astron. Astrofis.*, **53**, 385
- Ferrière, K. M. 2001, *Rev. Mod. Phys.*, **73**, 1031
- Fesen, R. A., Blair, W. P., & Kirshner, R. P. 1985, *ApJ*, **292**, 29
- Ford, H. C., Jenner, D. C., & Epps, H. W. 1973, *ApJ*, **183**, L73
- Ford, H. C., Hui, X., Ciardullo, R., Jacoby, G. H., & Freeman, K. C. 1996, *ApJ*, **458**, 455
- Frew, D. J., & Parker, Q. A. 2010, *PASA*, **27**, 129
- Galán-de Anta, P. M., Sarzi, M., Spriggs, T. W., et al. 2021, *A&A*, **652**, A109
- Grandmont, F., Drissen, L., Mandar, J., Thibault, S., & Baril, M. 2012, *SPIE Conf. Ser.*, **8446**, 84460U
- Grasha, K., Chen, H., Battisti, A. J., et al. 2022, *ApJ*, **929**, 118
- Green, D. A. 1984, *MNRAS*, **209**, 449
- Green, D. A. 2019, *J. Astrophys. Astron.*, **40**, 36
- Gregory, P. C. 2005, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with 'Mathematica' Support* (Cambridge University Press)
- Groves, B., Kreckel, K., Santoro, F., et al. 2023, *MNRAS*, **520**, 4902
- Gum, C. S. 1955, *MmRAS*, **67**, 155
- Haffner, L. M., Dettmar, R. J., Beckman, J. E., et al. 2009, *Rev. Mod. Phys.*, **81**, 969
- Hajian, A. R., Movit, S. M., Trofimov, D., et al. 2007, *ApJS*, **169**, 289
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, **585**, 357
- Hartke, J., Arnaboldi, M., Gerhard, O., et al. 2020, *A&A*, **642**, A46
- Herrmann, K. A., Ciardullo, R., Feldmeier, J. J., & Vinciguerra, M. 2008, *ApJ*, **683**, 630
- Ho, I. T., Kreckel, K., Meidt, S. E., et al. 2019, *ApJ*, **885**, L31
- Hodge, P. W. 1969, *ApJS*, **18**, 73
- Hodge, P. W., Balsley, J., Wyder, T. K., & Skelton, B. P. 1999, *PASP*, **111**, 685
- Holt, J., Tadhunter, C. N., & Morganti, R. 2009, *MNRAS*, **400**, 589
- Hoopes, C. G., Walterbos, R. A. M., & Greenwalt, B. E. 1996, *AJ*, **112**, 1429
- Hui, X., Ford, H. C., Freeman, K. C., & Dopita, M. A. 1995, *ApJ*, **449**, 592
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
- Jacoby, G. H. 1989, *ApJ*, **339**, 39
- Jacoby, G. H., Ciardullo, R., & Harris, W. E. 1996, *ApJ*, **462**, 1
- Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, *MNRAS*, **346**, 1055
- Kennicutt, Robert C., J., Edgar, B. K., & Hodge, P. W. 1989, *ApJ*, **337**, 761
- Kennicutt, Robert C., J., Armus, L., Bendo, G., et al. 2003, *PASP*, **115**, 928
- Kennicutt, Robert C., J., Lee, J. C., Funes, J. G., et al. 2008, *ApJS*, **178**, 247
- Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, *ApJ*, **556**, 121
- Kim, W.-T., & Elmegreen, B. G. 2017, *ApJ*, **841**, L4
- Knapen, J. H. 1998, *MNRAS*, **297**, 255
- Kopsacheili, M., Zezas, A., & Leonidaki, I. 2020, *MNRAS*, **491**, 889
- Kreckel, K., Blanc, G. A., Schinnerer, E., et al. 2016, *ApJ*, **827**, 103
- Kreckel, K., Groves, B., Bigiel, F., et al. 2017, *ApJ*, **834**, 174
- Kreckel, K., Ho, I. T., Blanc, G. A., et al. 2019, *ApJ*, **887**, 80
- Kreckel, K., Egorov, O., Belfiore, F., et al. 2022, *A&A*, **667**, A16
- Kuchar, T. A., & Clark, F. O. 1997, *ApJ*, **488**, 224
- Lee, J. H., & Lee, M. G. 2014a, *ApJ*, **786**, 130
- Lee, J. H., & Lee, M. G. 2014b, *ApJ*, **793**, 134
- Leroy, A. K., Schinnerer, E., Hughes, A., et al. 2021, *ApJS*, **257**, 43
- Levesque, E. M., Kewley, L. J., & Larson, K. L. 2010, *AJ*, **139**, 712
- Lockman, F. J. 1989, *ApJS*, **71**, 469
- Long, K. S., Blair, W. P., Winkler, P. F., et al. 2010, *ApJS*, **187**, 495
- Long, K. S., Blair, W. P., Winkler, P. F., et al. 2022, *ApJ*, **929**, 144
- López-Hernández, J., Terlevich, E., Terlevich, R., et al. 2013, *MNRAS*, **430**, 472
- Luridiana, V., Morisset, C., & Shaw, R. A. 2015, *A&A*, **573**, A42
- Maciel, W. J. 2013, *Astrophysics of the Interstellar Medium* (New York, NY: Springer)
- Magnier, E. A., Prins, S., van Paradijs, J., et al. 1995, *A&AS*, **114**, 215
- Magrini, L., Corradi, R. L. M., Mampaso, A., & Perinotto, M. 2000, *A&A*, **355**, 713
- Mascoop, J. L., Anderson, L. D., Wenger, T. V., et al. 2021, *ApJ*, **910**, 159
- Mathewson, D. S., & Clarke, J. N. 1972, *ApJ*, **178**, L105
- Meatheringham, S. J., & Dopita, M. A. 1991a, *ApJS*, **75**, 407
- Meatheringham, S. J., & Dopita, M. A. 1991b, *ApJS*, **76**, 1085
- Micelotta, E. R., Dwek, E., & Slavín, J. D. 2016, *A&A*, **590**, A65
- Mills, B. Y. 1952, *Aust. J. Sci. Res. A Phys. Sci.*, **5**, 266
- Mingozzi, M., Belfiore, F., Cresci, G., et al. 2020, *A&A*, **636**, A42
- Moumen, I., Robert, C., Devost, D., et al. 2019, *MNRAS*, **488**, 803
- O'Donnell, J. E. 1994, *ApJ*, **422**, 158
- Oey, M. S., Meurer, G. R., Yelda, S., et al. 2007, *ApJ*, **661**, 801
- Osterbrock, D. E., & Ferland, G. J. 2006, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei* (University Science Books)
- Otte, B., & Dettmar, R. J. 1999, *A&A*, **343**, 705
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *J. Mach. Learn. Res.*, **12**, 2825
- Pellegrini, E. W., Oey, M. S., Winkler, P. F., et al. 2012, *ApJ*, **755**, 40
- Pellet, A., Astier, N., Viale, A., et al. 1978, *A&AS*, **31**, 439
- Pérez-Montero, E. 2014, *MNRAS*, **441**, 2663
- Pilyugin, L. S., & Grebel, E. K. 2016, *MNRAS*, **457**, 3678
- Points, S. D., Long, K. S., Winkler, P. F., & Blair, W. P. 2019, *ApJ*, **887**, 66
- Querejeta, M., Schinnerer, E., Meidt, S., et al. 2021, *A&A*, **656**, A133
- Rauch, T. 2003, *A&A*, **403**, 709
- Reifenstein, E. C., Wilson, T. L., Burke, B. F., Mezger, P. G., & Altenhoff, W. J. 1970, *A&A*, **4**, 357
- Rekola, R., Richer, M. G., McCall, M. L., et al. 2005, *MNRAS*, **361**, 330
- Reynolds, R. J. 1985, *ApJ*, **294**, 256
- Reynolds, R. J. 1997, in *The Physics of Galactic Halos*, eds. H. Lesch, R.-J. Dettmar, U. Mebold, & R. Schlickeiser (Wiley), 57
- Rhea, C., Rousseau-Nepton, L., Prunet, S., et al. 2021, *ApJ*, **910**, 129
- Richer, M. G. 2006, *IAU Symp.*, **234**, 317
- Riesgo, H., & López, J. A. 2006, *Rev. Mex. Astron. Astrofis.*, **42**, 47
- Roth, M. M., Jacoby, G. H., Ciardullo, R., et al. 2021, *ApJ*, **916**, 21
- Rousseau-Nepton, L., Robert, C., Martin, R. P., Drissen, L., & Martin, T. 2018, *MNRAS*, **477**, 4152
- Sabbadin, F., Minello, S., & Bianchini, A. 1977, *A&A*, **60**, 147
- Sabin, L., Parker, Q. A., Contreras, M. E., et al. 2013, *MNRAS*, **431**, 279
- Sánchez-Menguiano, L., Sánchez, S. F., Pérez, I., et al. 2018, *A&A*, **609**, A119
- Santoro, F., Kreckel, K., Belfiore, F., et al. 2022, *A&A*, **658**, A188
- Sarzi, M., Shields, J. C., Schawinski, K., et al. 2010, *MNRAS*, **402**, 2187
- Scheuermann, F., Kreckel, K., Anand, G. S., et al. 2022, *MNRAS*, **511**, 6087
- Schlesinger, K., Pogge, R. W., Martini, P., Shields, J. C., & Fields, D. 2009, *ApJ*, **699**, 857
- Schönberner, D., Jacob, R., Lehmann, H., et al. 2014, *Astron. Nachr.*, **335**, 378
- Sharpless, S. 1953, *ApJ*, **118**, 362
- Sharpless, S. 1959, *ApJS*, **4**, 257
- Skilling, J. 2004, *AIP Conf. Proc.*, **735**, 395
- Skilling, J. 2006, *Bayesian Anal.*, **1**, 833
- Speagle, J. S. 2020, *MNRAS*, **493**, 3132
- Spriggs, T. W., Sarzi, M., Napiwotzki, R., et al. 2020, *A&A*, **637**, A62
- Stasińska, G. 2004, in *Cosmochemistry. The Melting Pot of the Elements*, eds. C. Esteban, R. García López, A. Herrero, & F. Sánchez (Cambridge University Press), 115
- Storchi-Bergmann, T., & Bonatto, C. J. 1991, *MNRAS*, **250**, 138
- Sutherland, R., Dopita, M., Binette, L., & Groves, B. 2018, *Astrophysics Source Code Library* [record ascl:1807.005]
- Thilker, D. A., Braun, R., & Walterbos, R. A. M. 2000, *AJ*, **120**, 3070
- Thilker, D. A., Walterbos, R. A. M., Braun, R., & Hoopes, C. G. 2002, *AJ*, **124**, 3118
- Tomčić, N., Kreckel, K., Groves, B., et al. 2017, *ApJ*, **844**, 155
- Veilleux, S., & Osterbrock, D. E. 1987, *ApJS*, **63**, 295
- Venturi, G., Nardini, E., Marconi, A., et al. 2018, *A&A*, **619**, A74
- Véron-Cetty, M. P., & Véron, P. 2006, *A&A*, **455**, 773
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nat. Methods*, **17**, 261
- Walsh, J. R., Monreal-Ibero, A., Barlow, M. J., et al. 2016, *A&A*, **588**, A106
- Weilbacher, P. M., Palsa, R., Streicher, O., et al. 2020, *A&A*, **641**, A28
- Williams, J. P., de Geus, E. J., & Blitz, L. 1994, *ApJ*, **428**, 693
- Winkler, P. F., Coffin, S. C., Blair, W. P., Long, K. S., & Kuntz, K. D. 2021, *ApJ*, **908**, 80

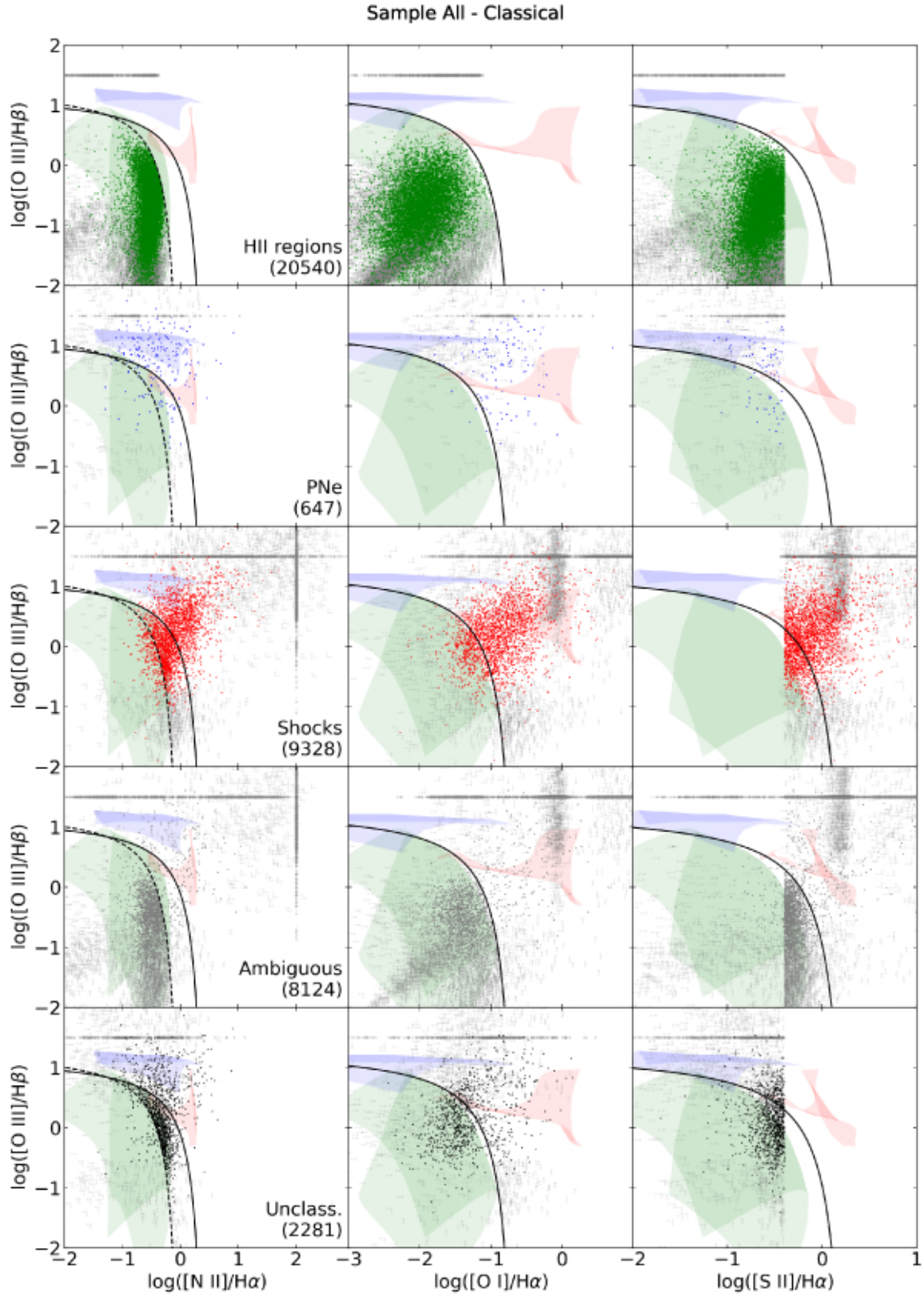
Appendix A: Diagnostic diagrams with traditional classification


Fig. A.1. Diagnostic diagrams from Baldwin et al. (1981) and Veilleux & Osterbrock (1987) for the full sample. Lines and shaded areas are defined as in Fig. 3. Here, the regions are classified according to the traditional classification criteria. In each row only, from top to bottom, we show: H II regions (green), PNe (blue), shocks (red), ambiguous (grey) and unclassified (black) objects.

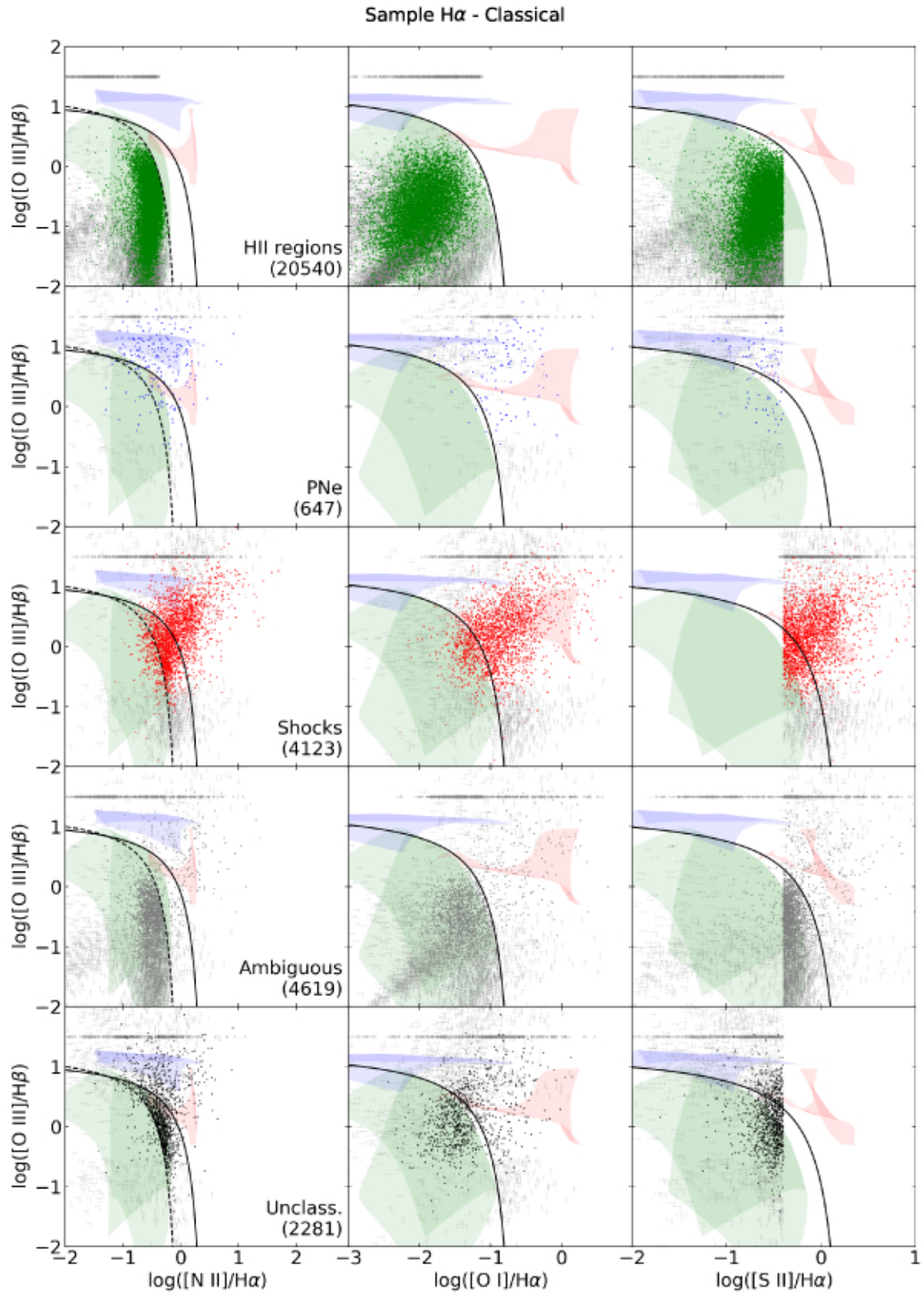


Fig. A.2. Diagnostic diagrams from Baldwin et al. (1981) and Veilleux & Osterbrock (1987) for the H α sample. Colours and symbols are as in Fig. A.1.

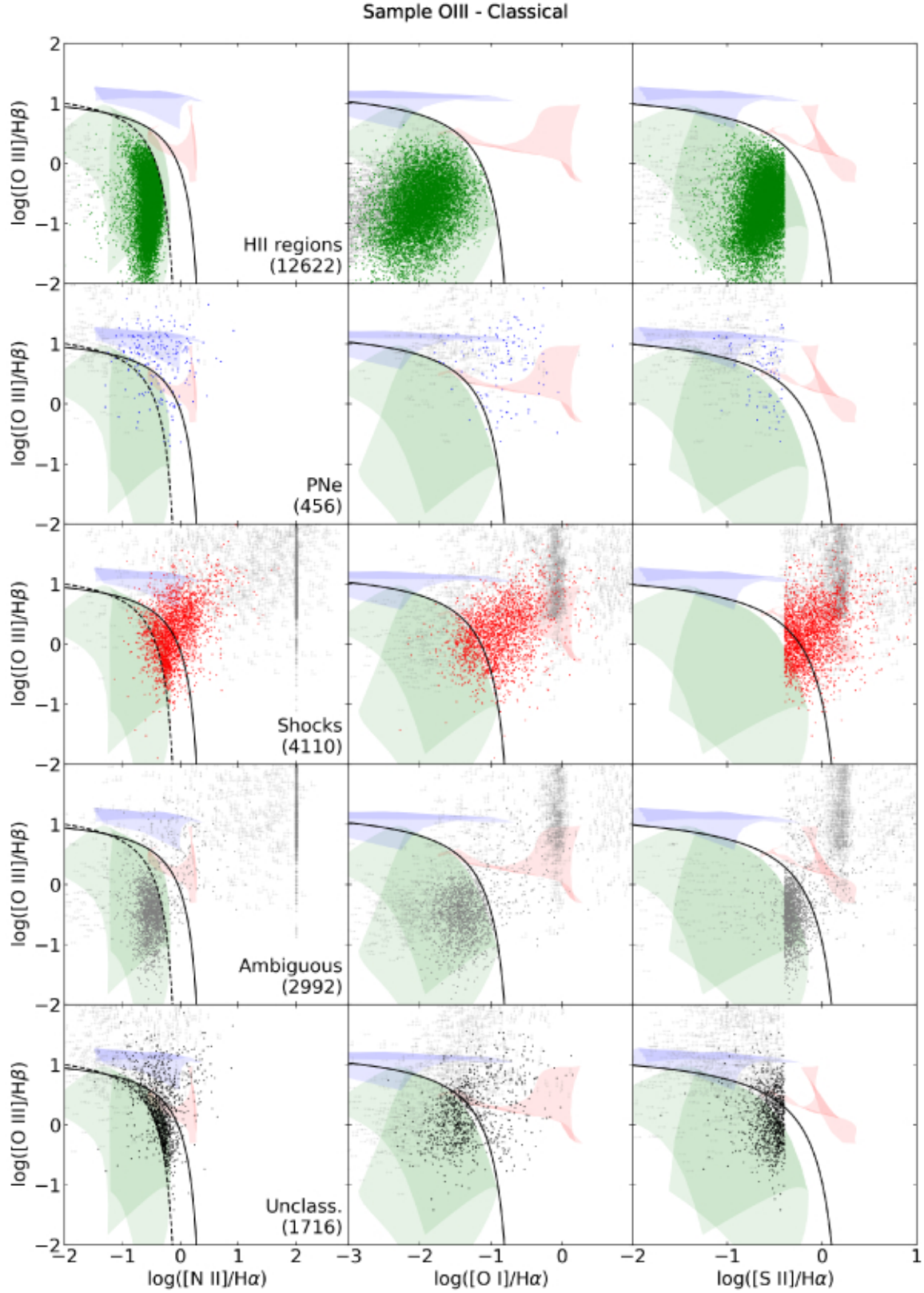


Fig. A.3. Diagnostic diagrams from Baldwin et al. (1981) and Veilleux & Osterbrock (1987) for the OIII sample. Colours and symbols are as in Fig. A.1.

Sample BPT - Classical

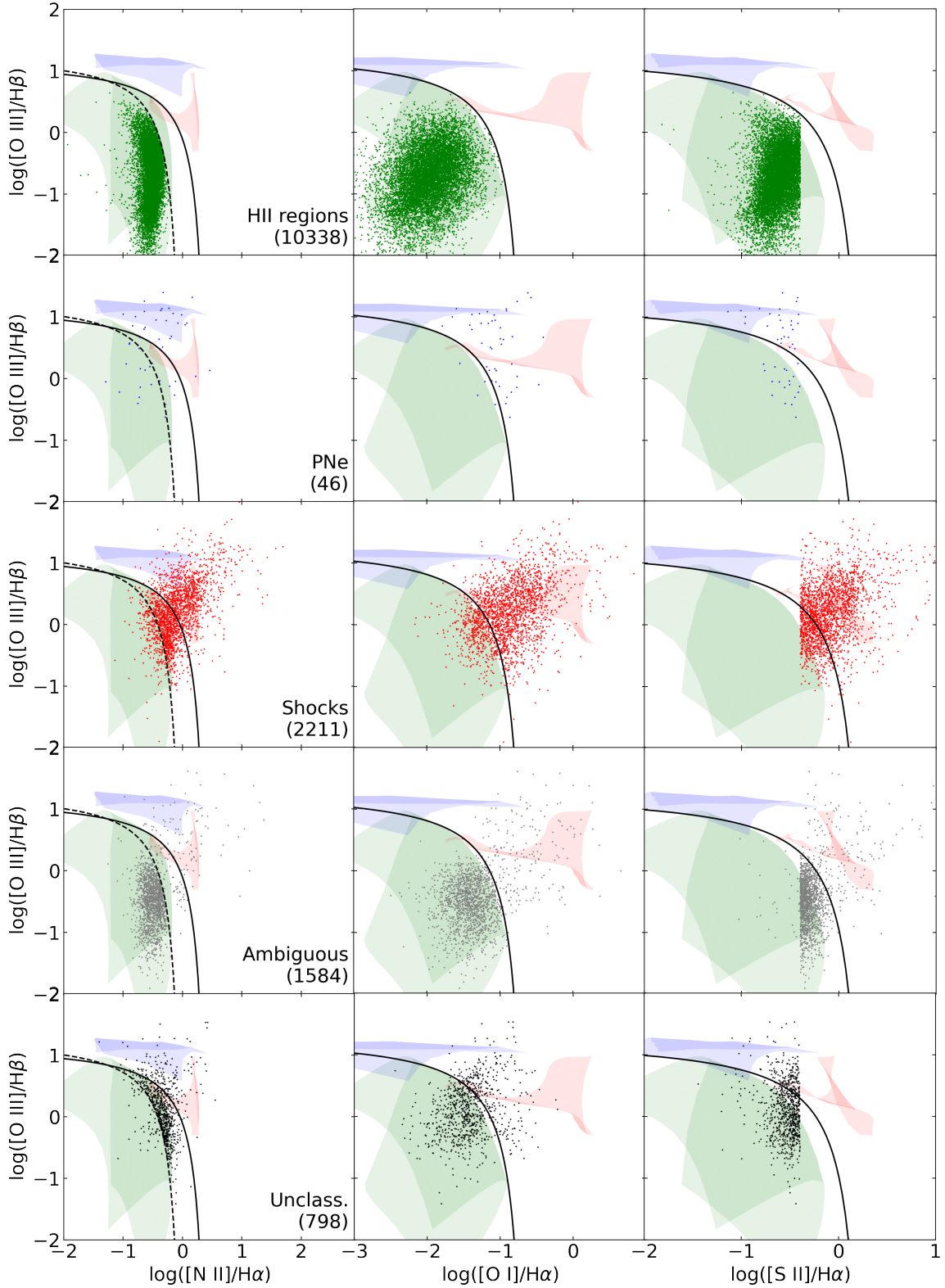


Fig. A.4. Diagnostic diagrams from Baldwin et al. (1981) and Veilleux & Osterbrock (1987) for the BPT sample. Colours and symbols are as in Fig. A.1.

Appendix B: Atlas of nebulae

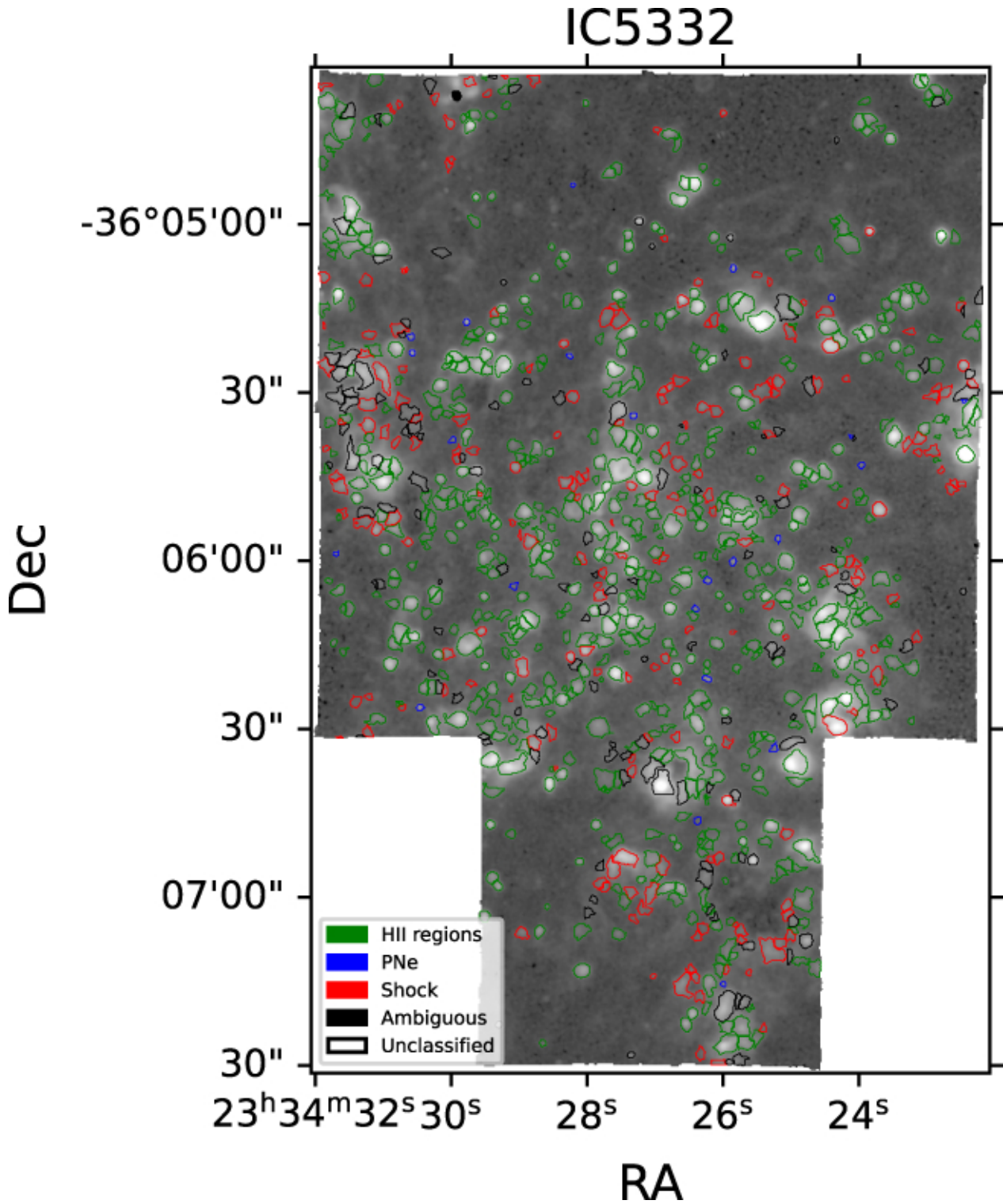


Fig. B.1. Boundaries of the nebulae detected in IC 5332 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

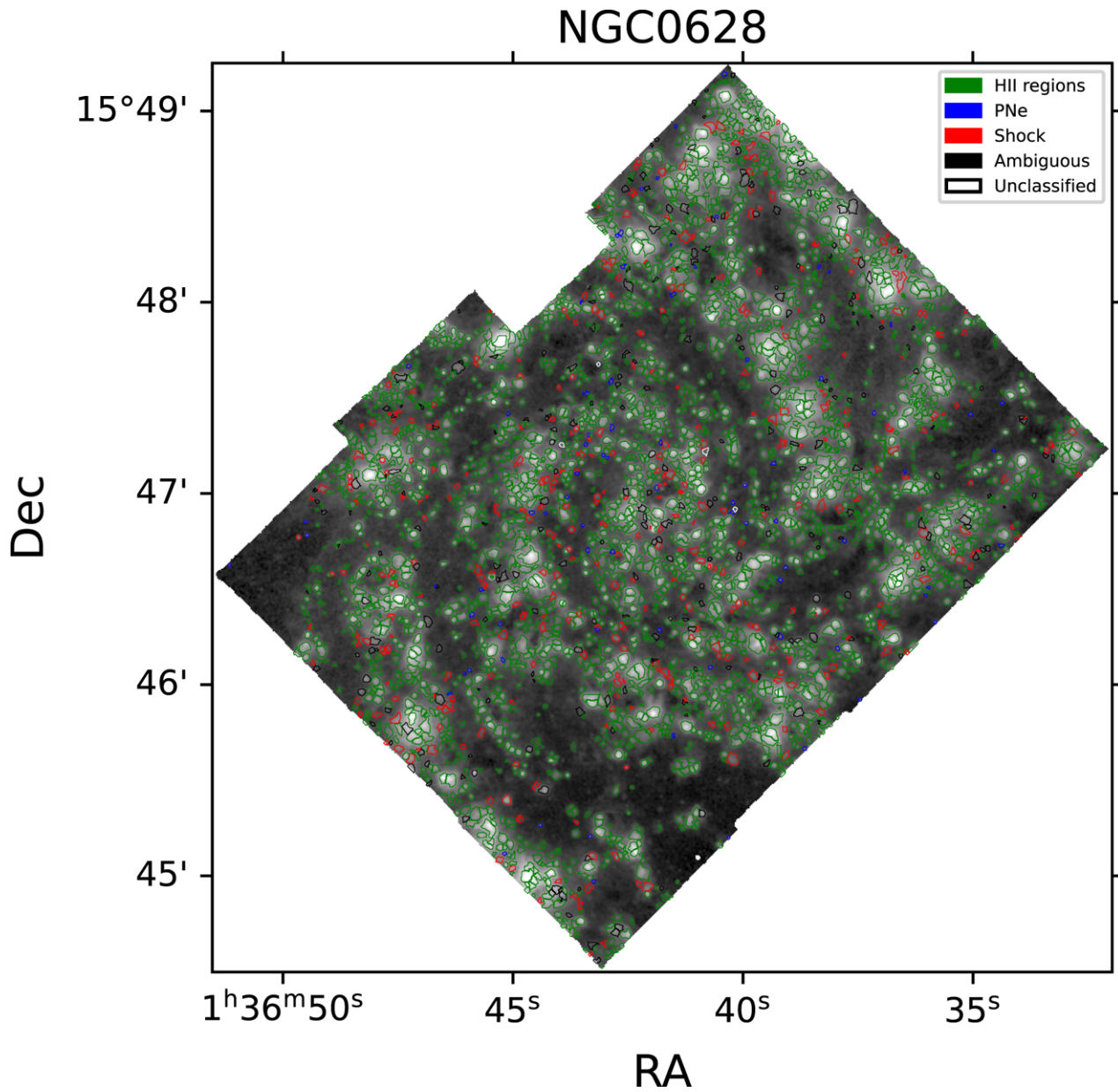


Fig. B.2. Boundaries of the nebulae detected in NGC 0628 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

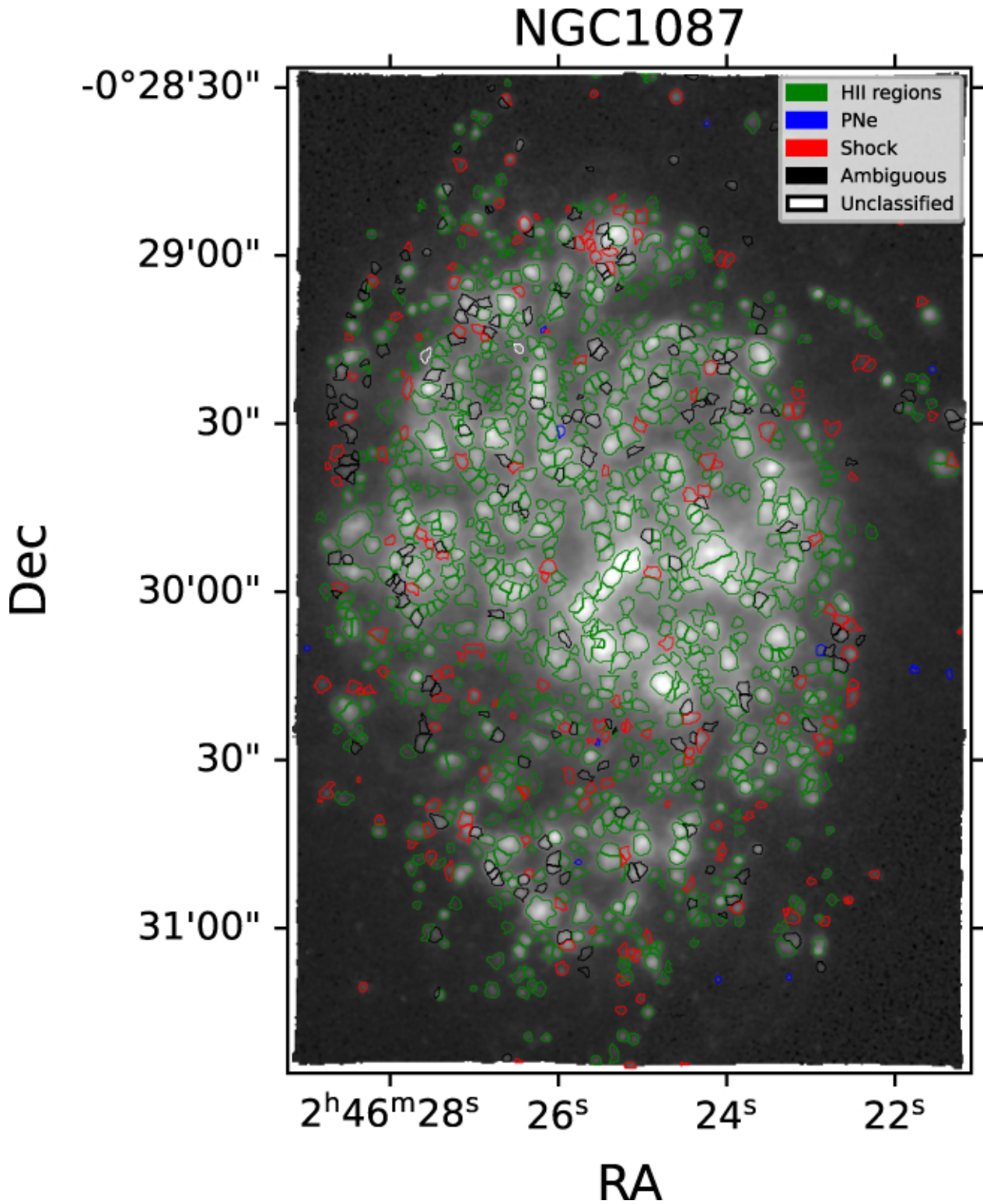


Fig. B.3. Boundaries of the nebulae detected in NGC 1087 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

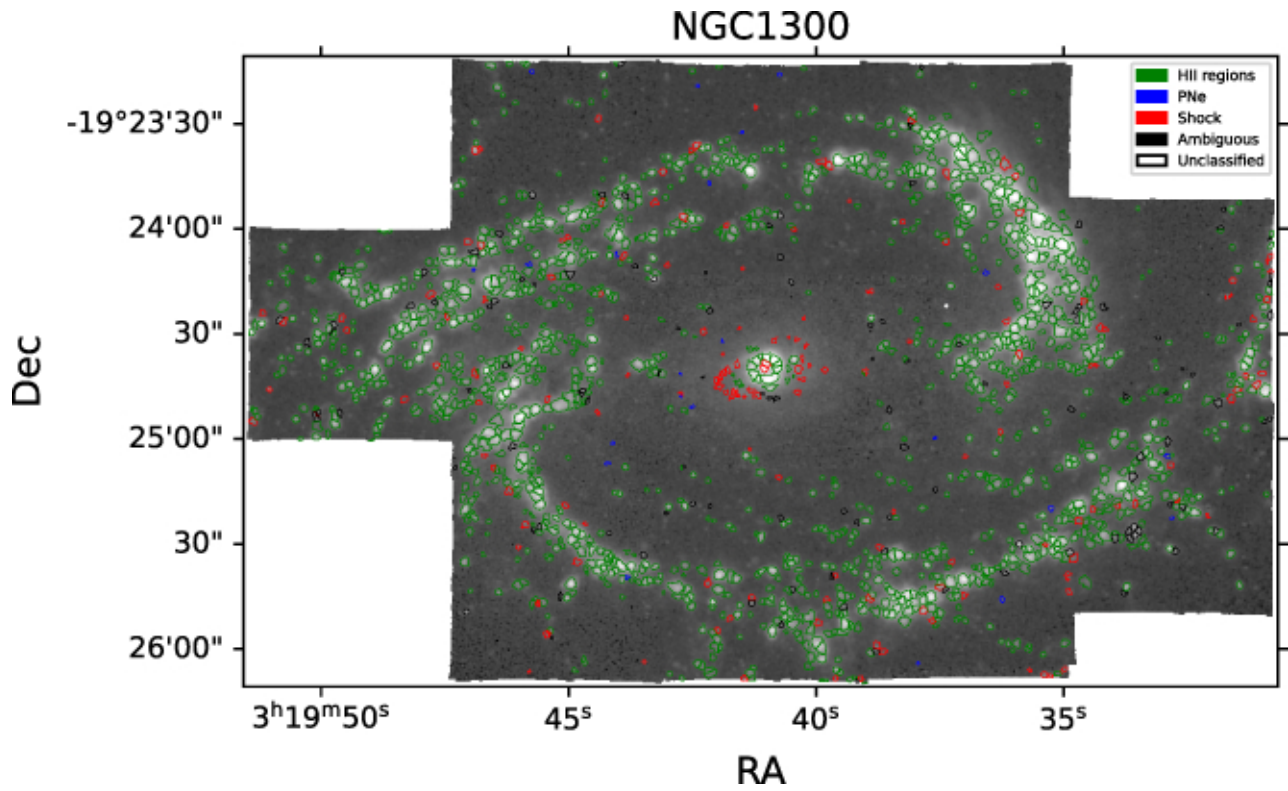


Fig. B.4. Boundaries of the nebulae detected in NGC 1300 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

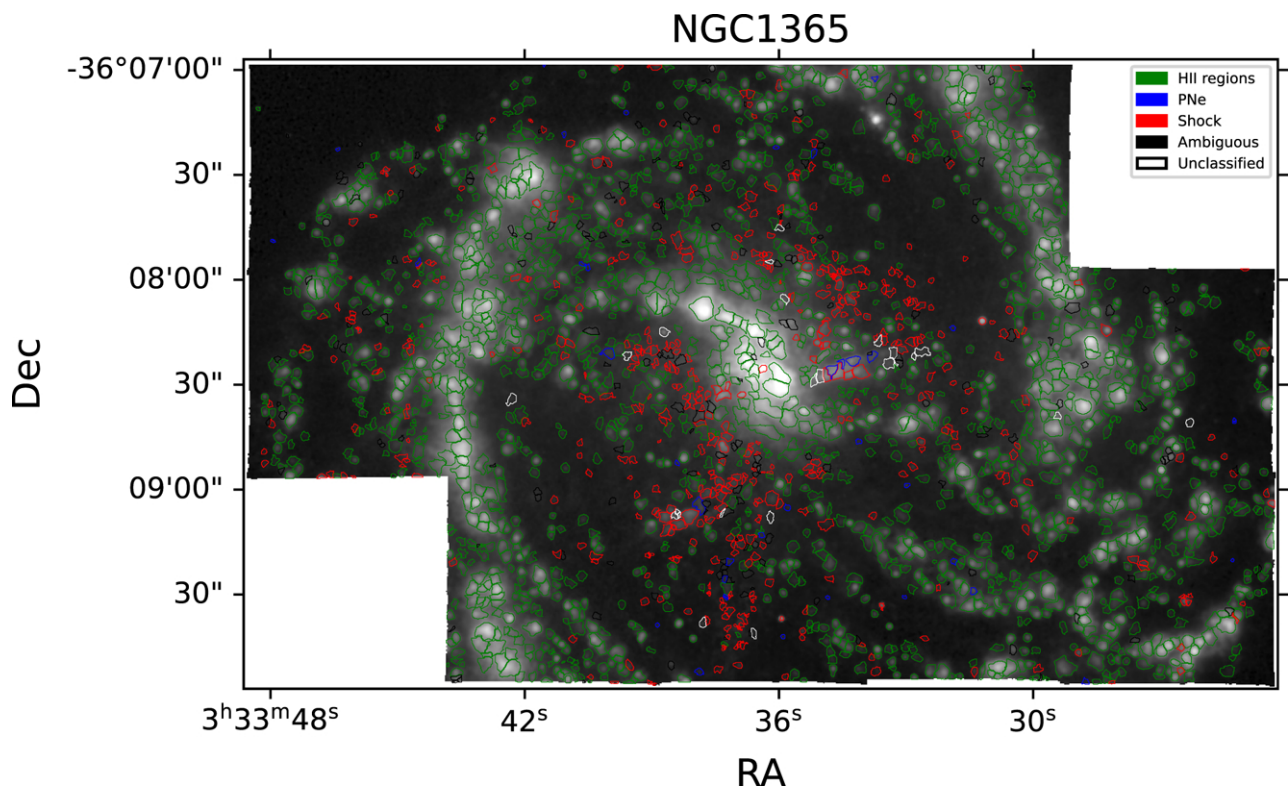


Fig. B.5. Boundaries of the nebulae detected in NGC 1365 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

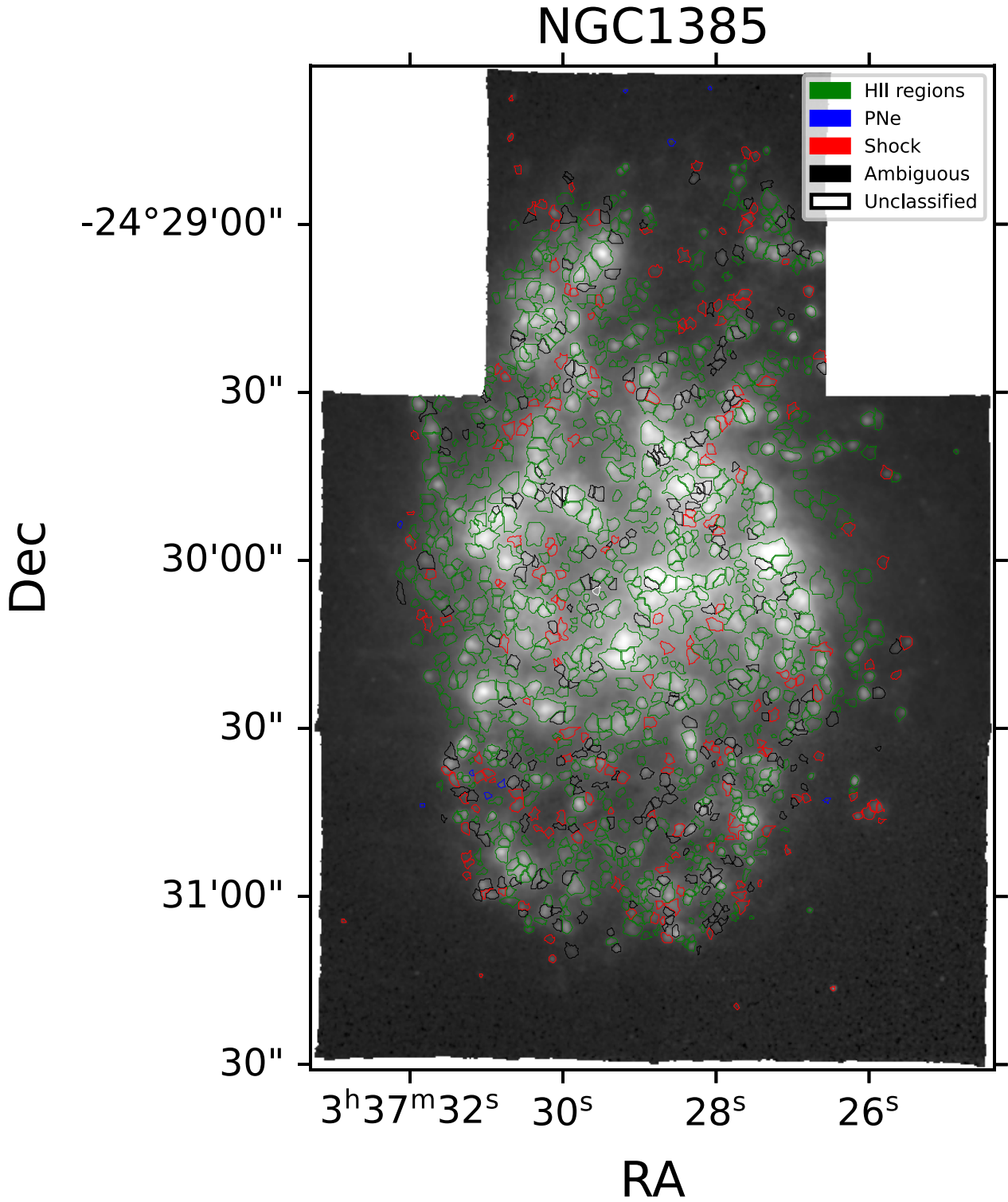


Fig. B.6. Boundaries of the nebulae detected in NGC 1385 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

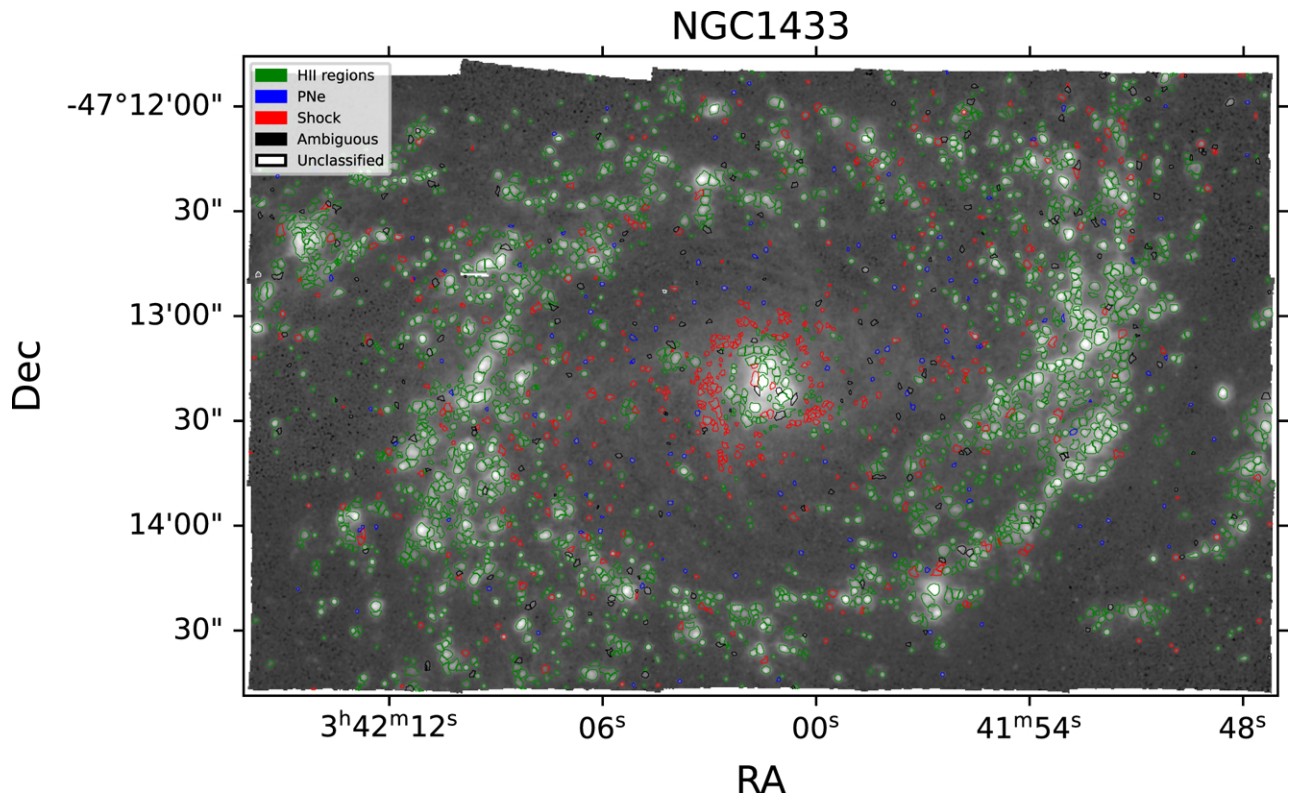


Fig. B.7. Boundaries of the nebulae detected in NGC 1433 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

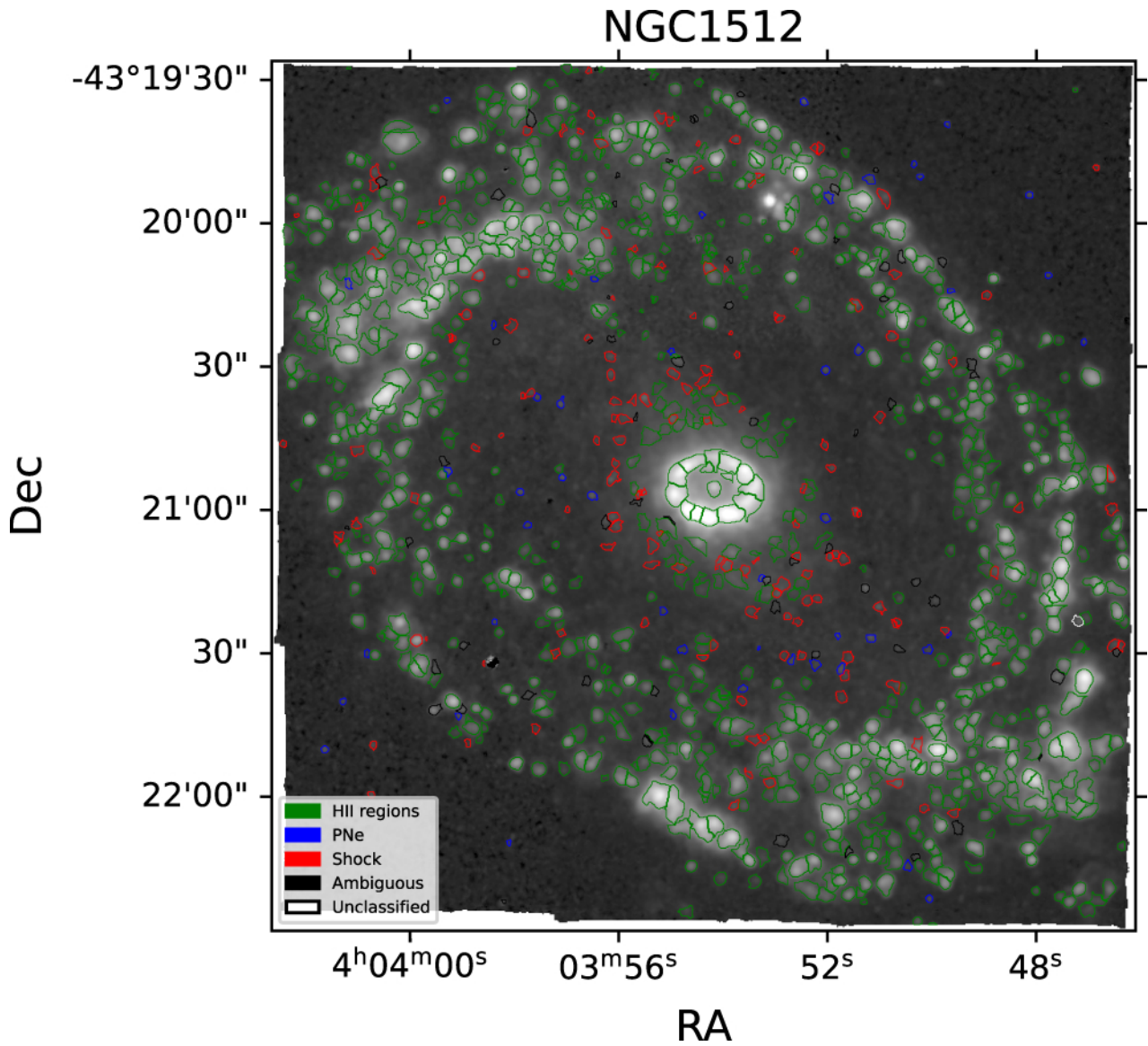


Fig. B.8. Boundaries of the nebulae detected in NGC 1512 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

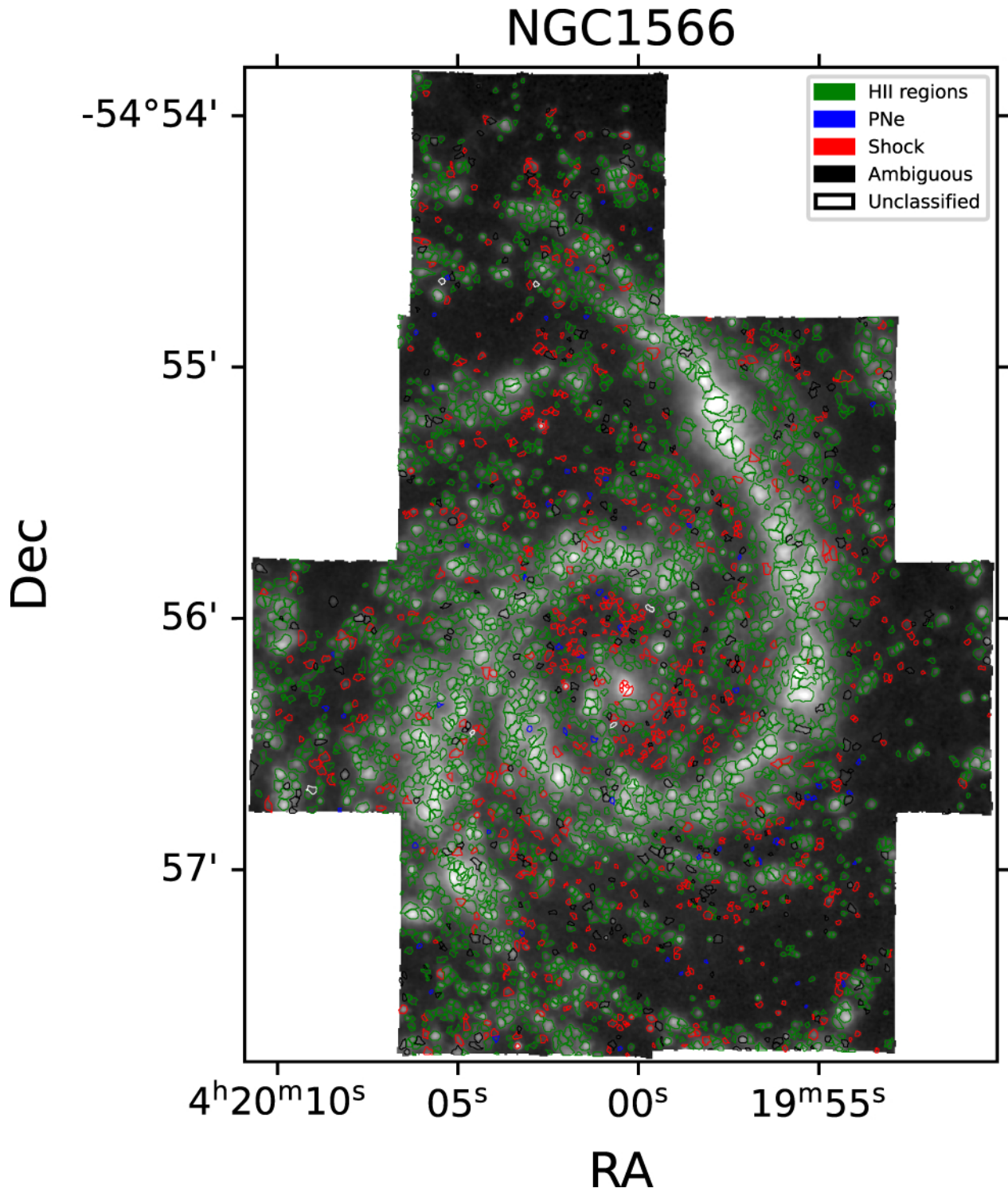


Fig. B.9. Boundaries of the nebulae detected in NGC 1566 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

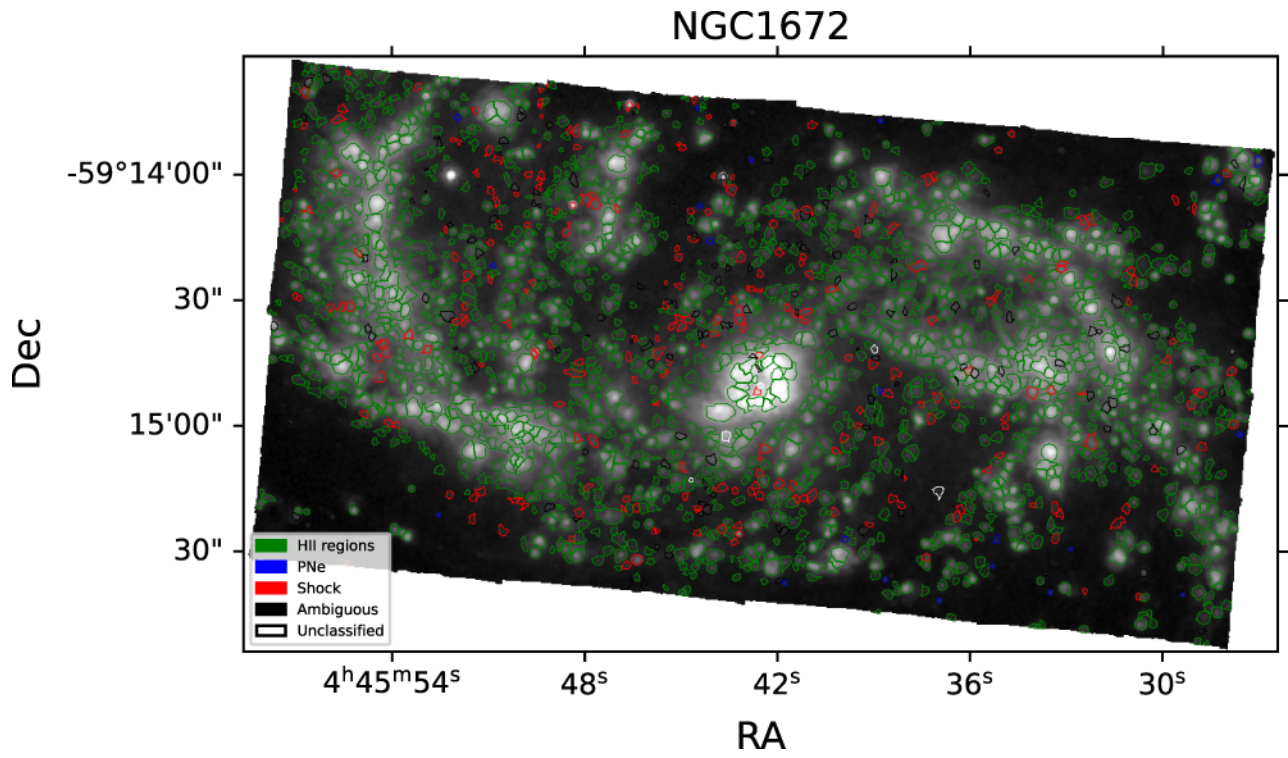


Fig. B.10. Boundaries of the nebulae detected in NGC 1672 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

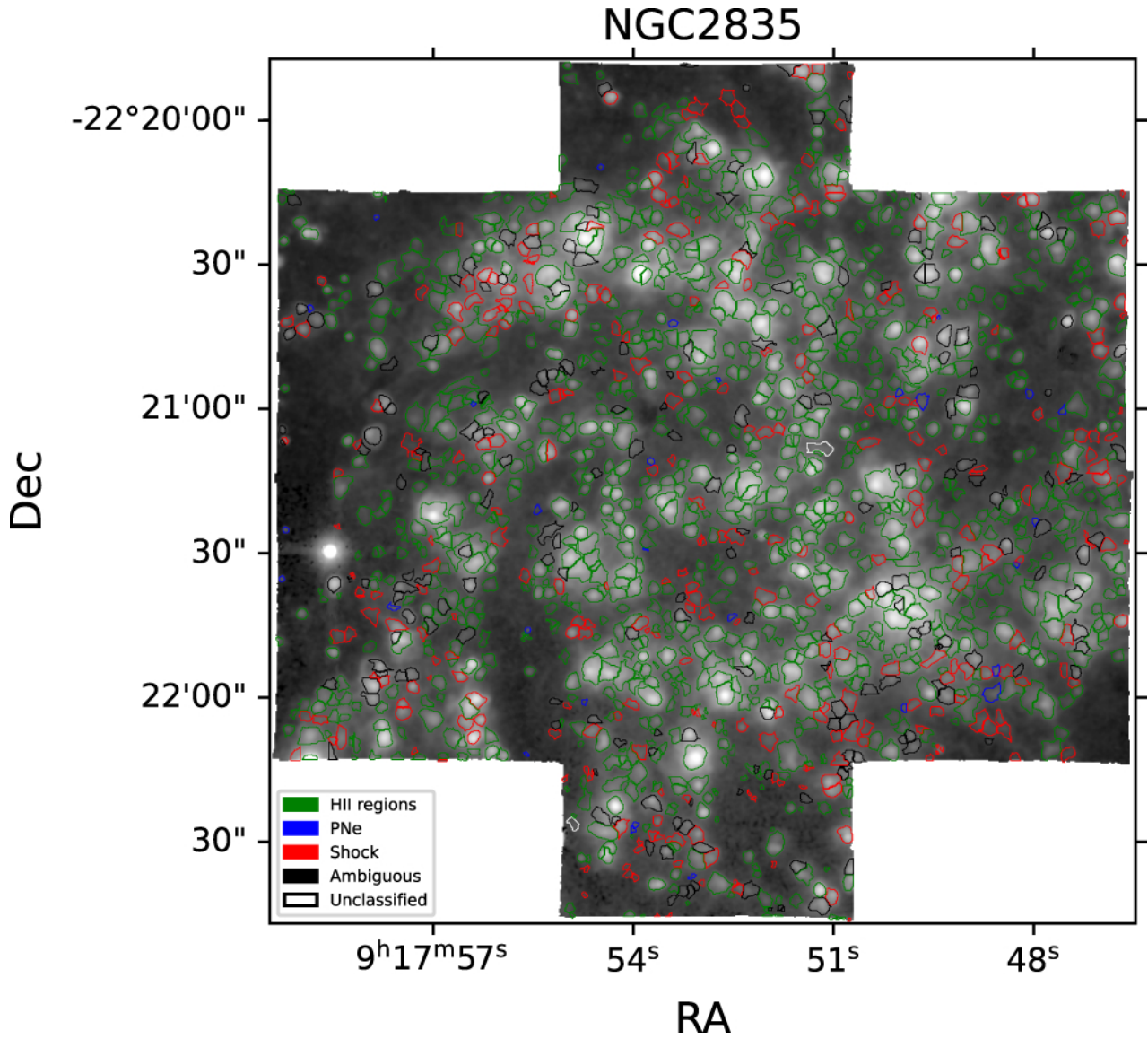


Fig. B.11. Boundaries of the nebulae detected in NGC 2835 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

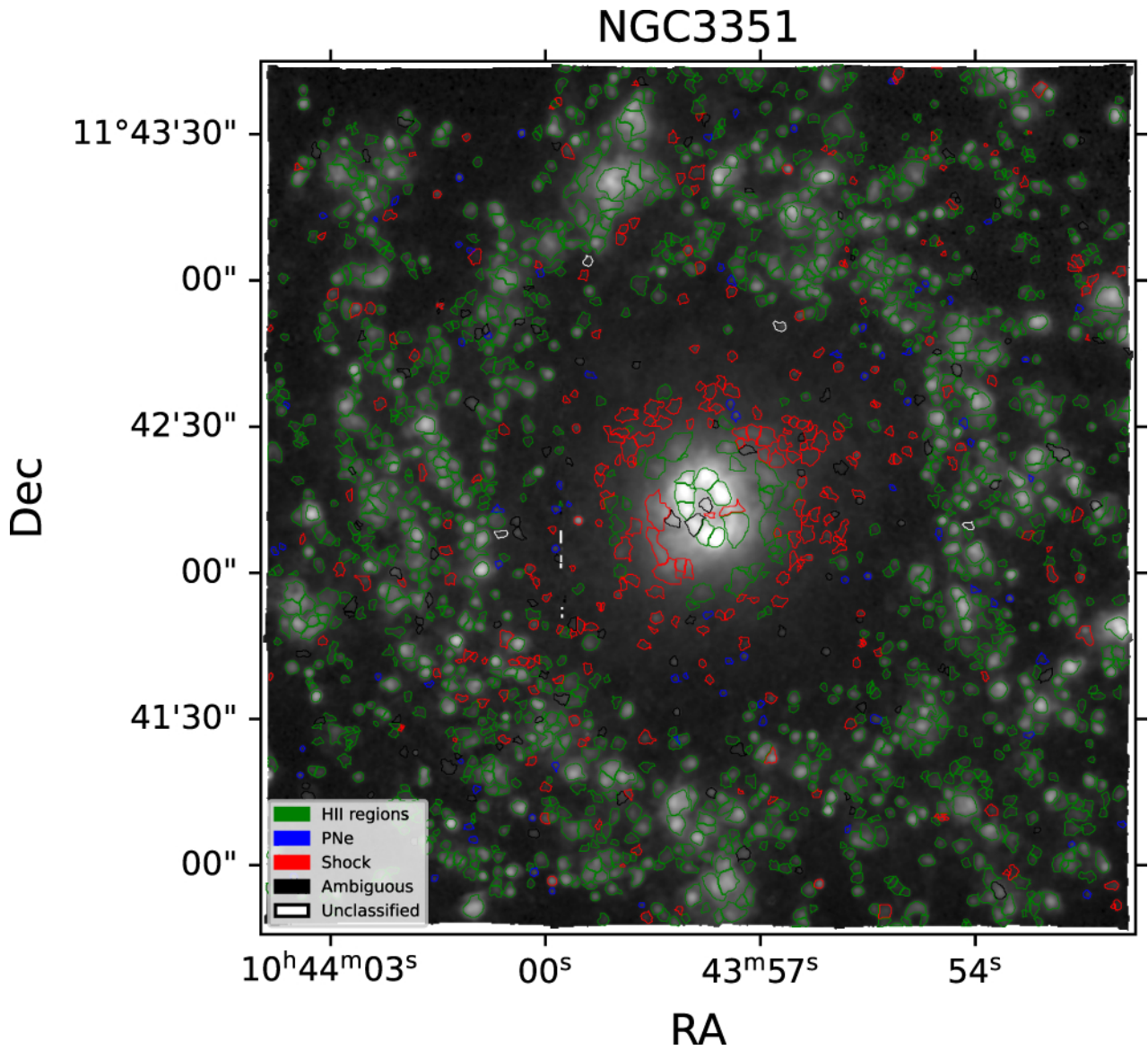


Fig. B.12. Boundaries of the nebulae detected in NGC 3351 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

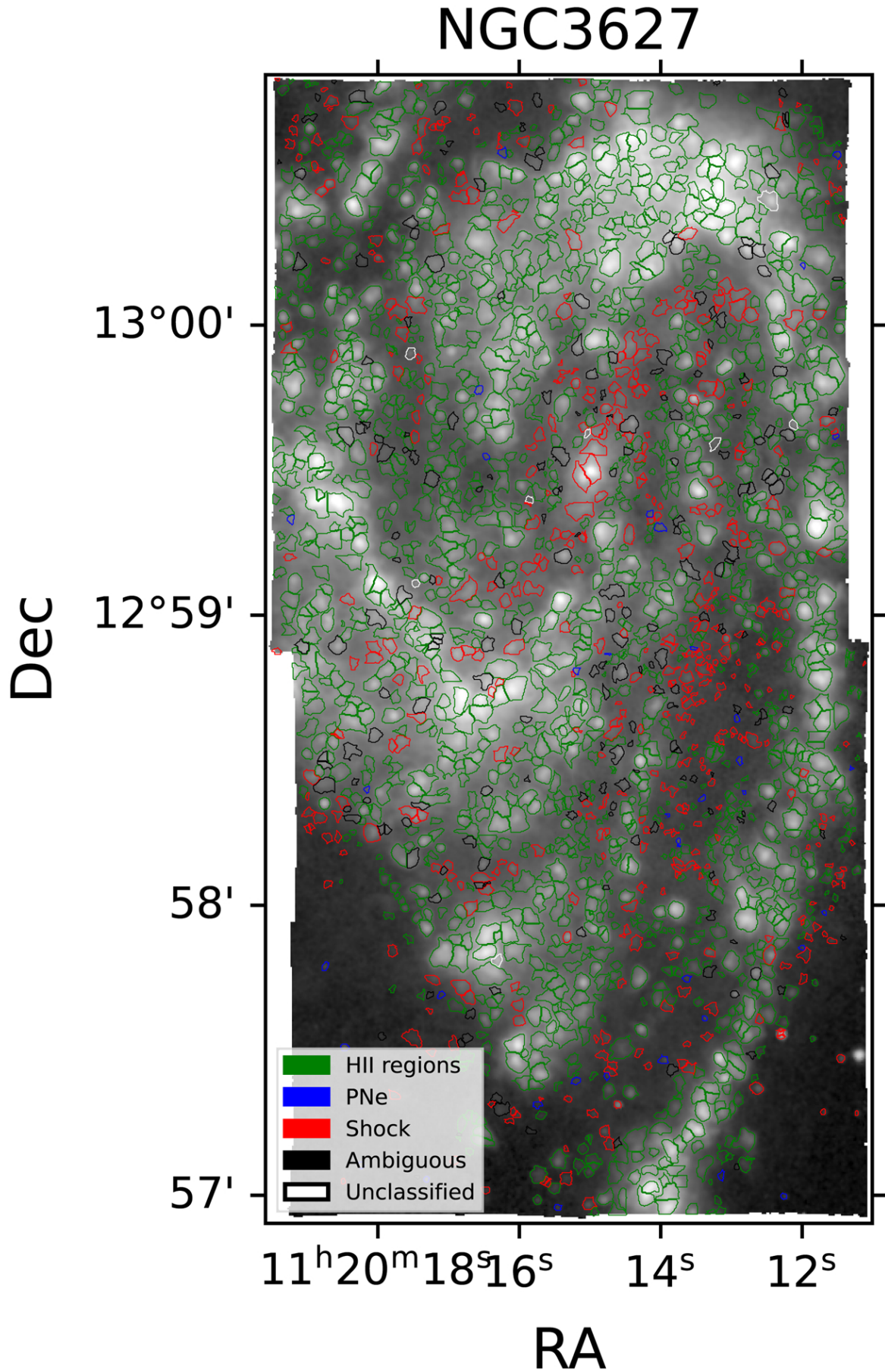


Fig. B.13. Boundaries of the nebulae detected in NGC 3627 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

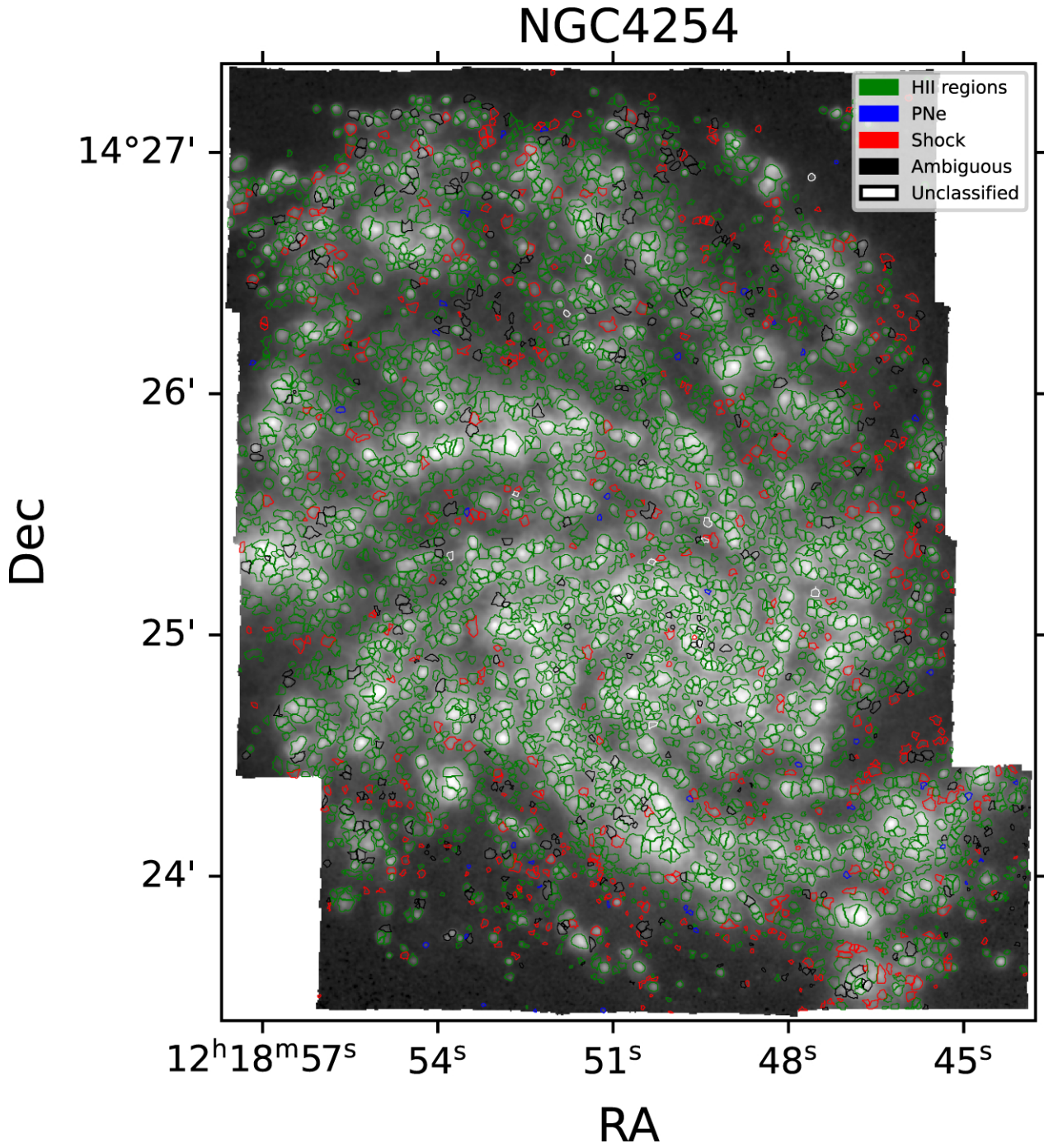


Fig. B.14. Boundaries of the nebulae detected in NGC 4254 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

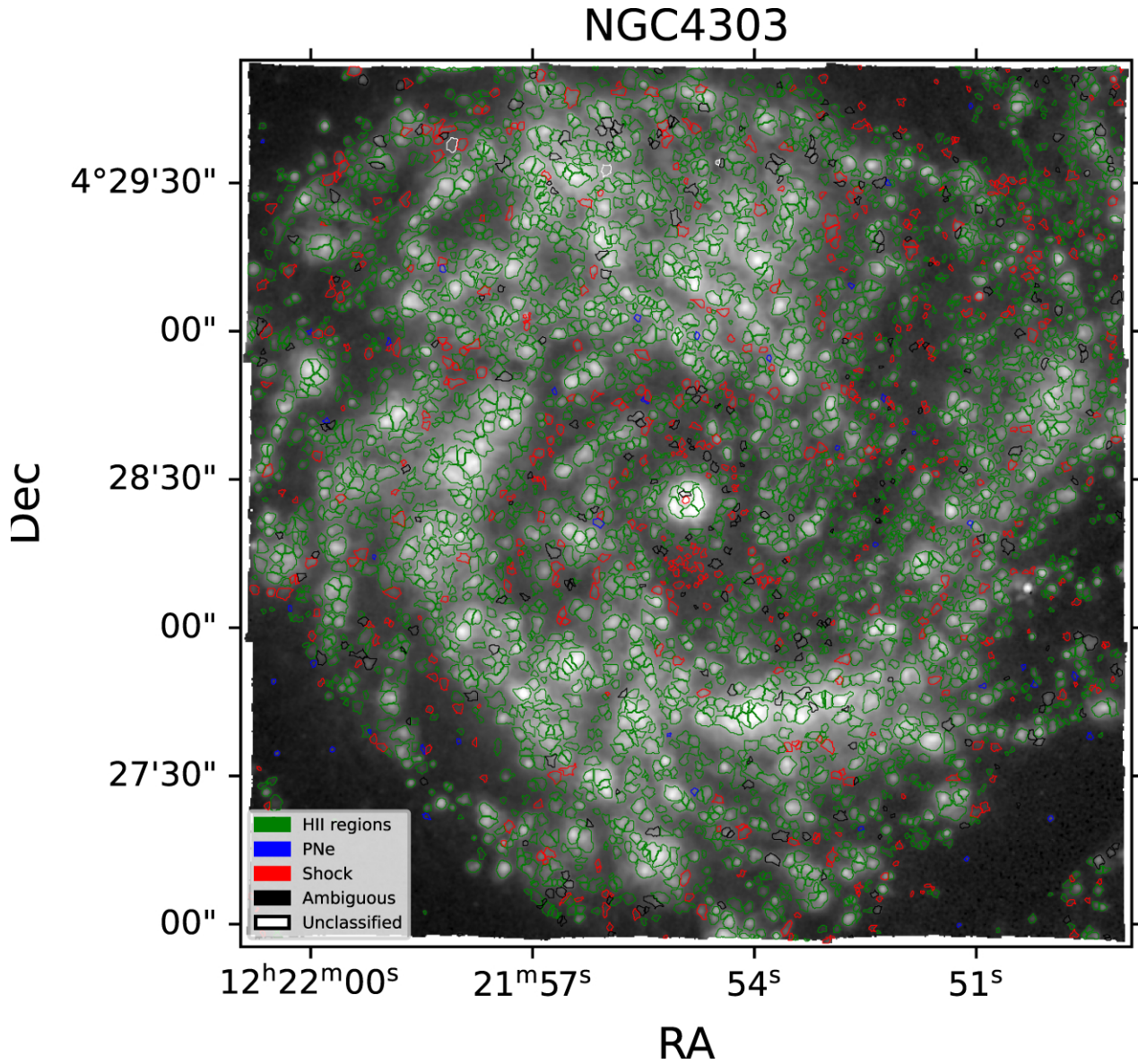


Fig. B.15. Boundaries of the nebulae detected in NGC 4303 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

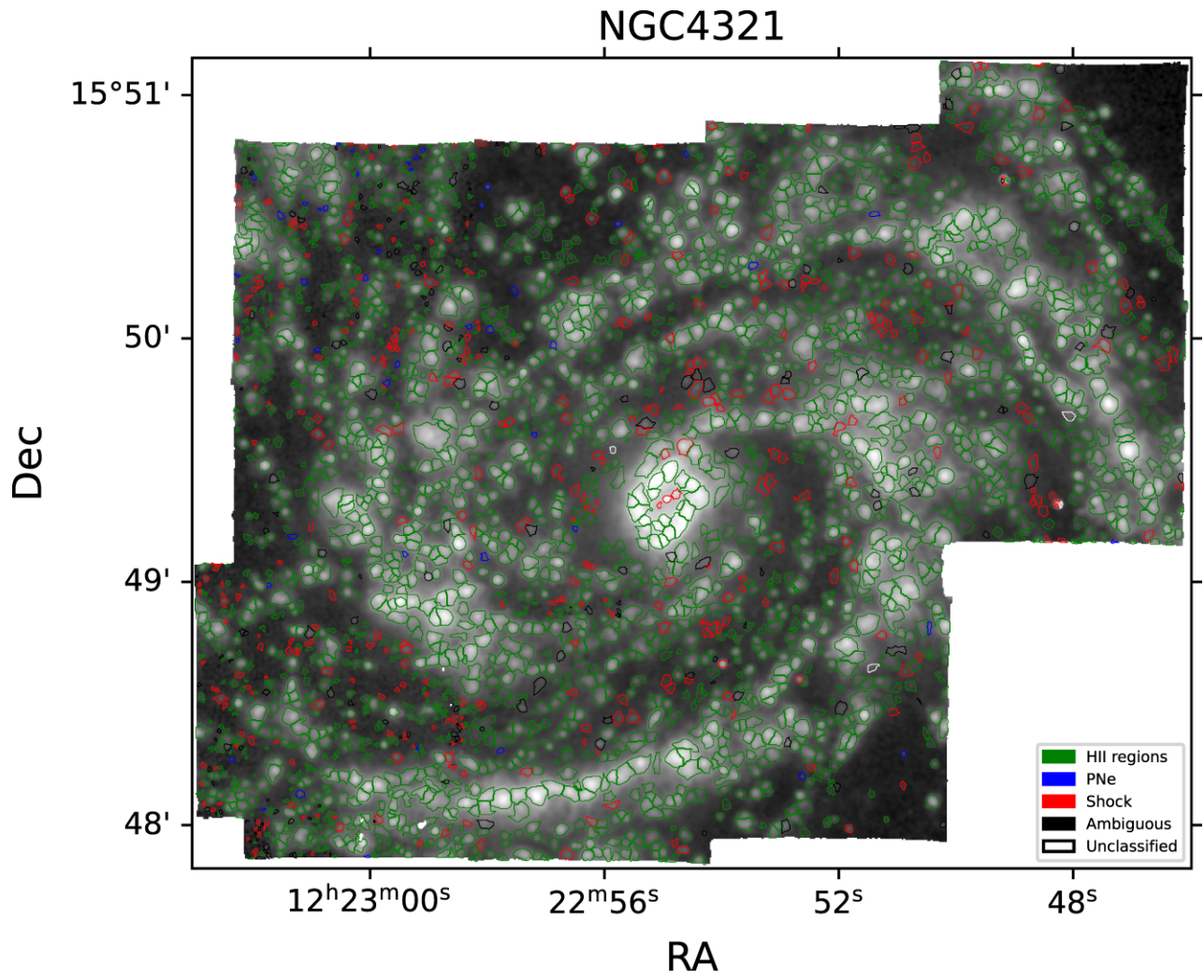


Fig. B.16. Boundaries of the nebulae detected in NGC 4321 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

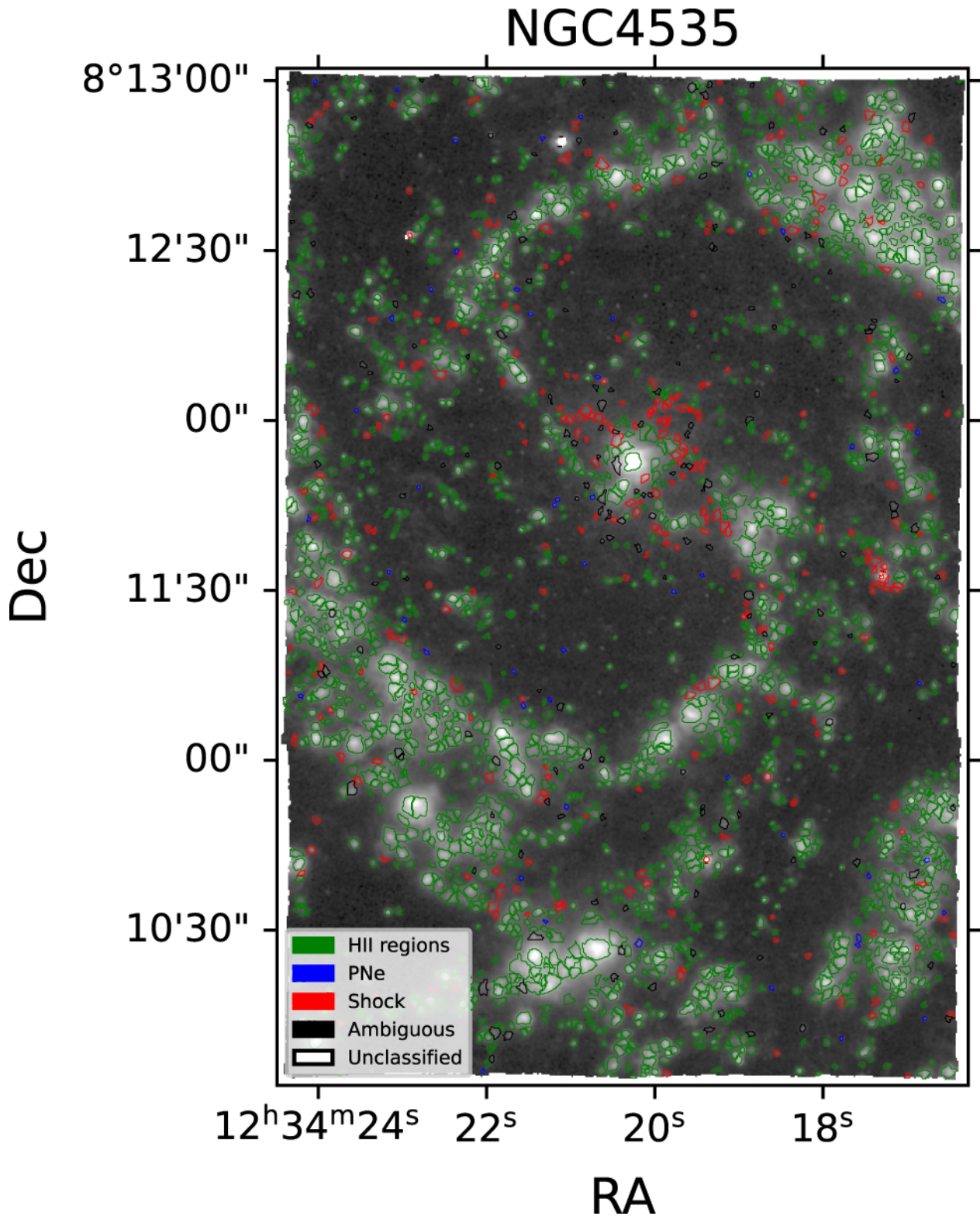


Fig. B.17. Boundaries of the nebulae detected in NGC 4535 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

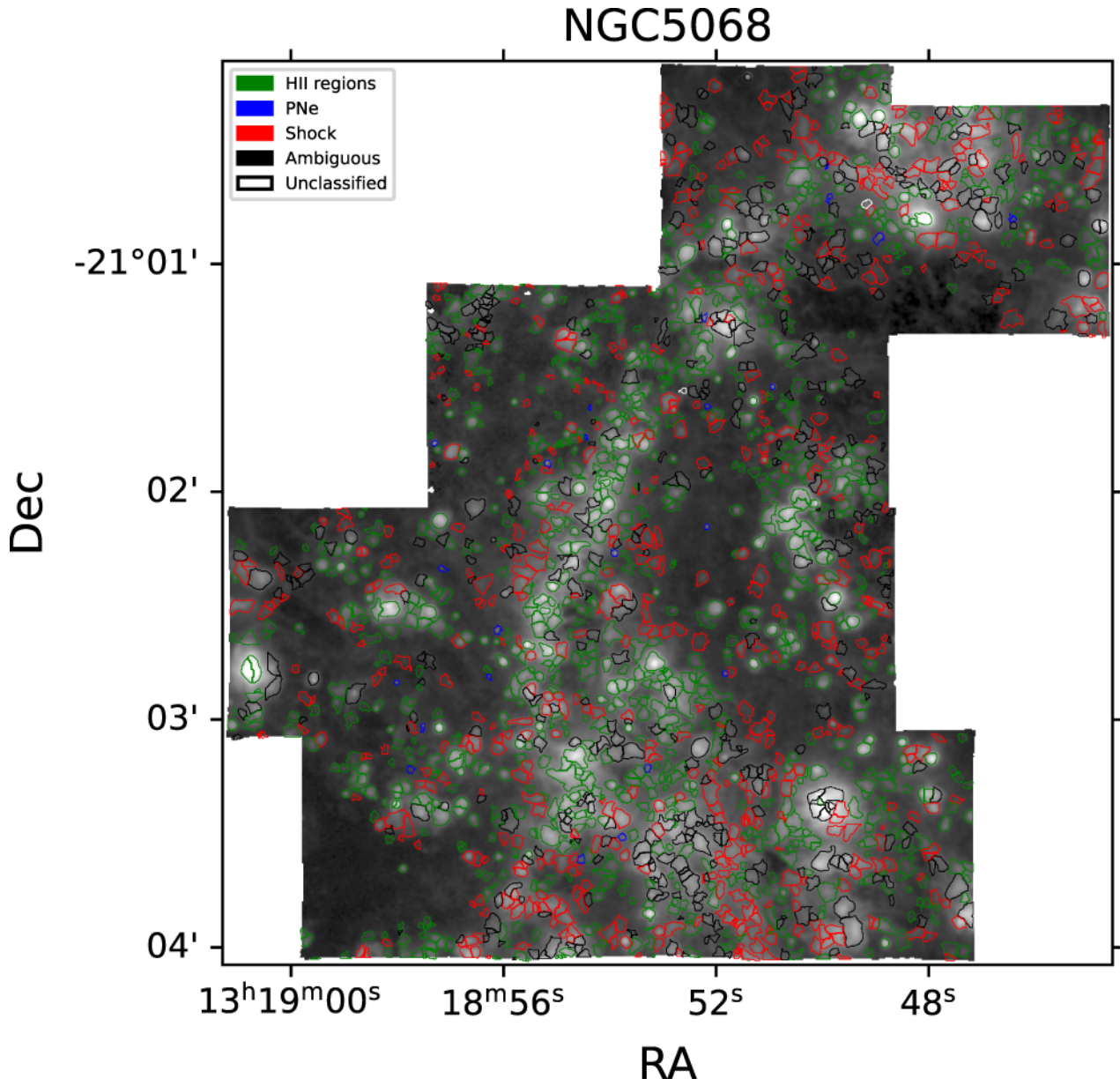


Fig. B.18. Boundaries of the nebulae detected in NGC 5068 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.

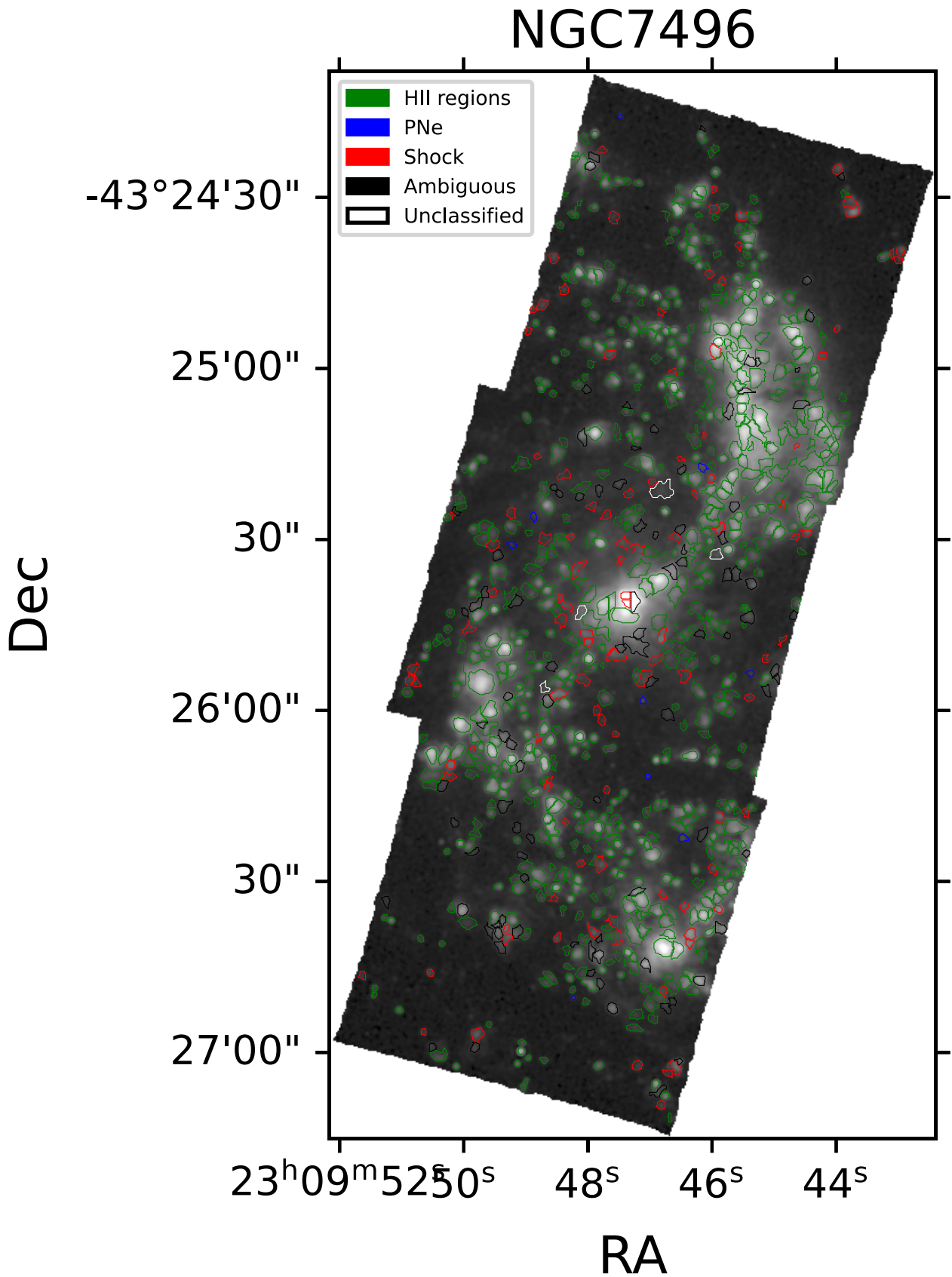


Fig. B.19. Boundaries of the nebulae detected in NGC 7496 superimposed on the detection map for the galaxy. The colour of the contour represents the classification of the nebula according to our model-comparison-based algorithm.