



Universidad Nacional Autónoma de México

POSGRADO EN ASTROFÍSICA  
INSTITUTO DE ASTRONOMÍA

INTERPRETACIÓN FÍSICA DE LAS REGIONES HII DEL PROYECTO  
CALIFA BASADA EN MODELOS DE FOTOIONIZACIÓN

TESIS

PARA OPTAR POR EL GRADO DE:  
DOCTOR EN CIENCIAS (ASTROFÍSICA)

PRESENTA:  
CARLOS CRISPÍN ESPINOSA PONCE

**Directores de Tesis:**  
DR. SEBASTIÁN FRANCISCO SÁNCHEZ SÁNCHEZ  
DR. CHRISTOPHE MORISSET  
INSTITUTO DE ASTRONOMÍA

CIUDAD DE MÉXICO, OCTUBRE 2022



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



*Para todas aquellas personas que he conocido  
y que influyeron directa o indirectamente  
para que llegara a este punto.  
Gracias a todos ustedes...*





# Declaración de autoría

Yo, M. en C. Carlos Crispín **Espinosa Ponce**, declaro que esta tesis llamada: Interpretación física de las regiones H II del proyecto CALIFA basada en modelos de fotoionización y el trabajo que se presenta en ella es de mi autoría. Yo confirmo que:

- Este trabajo se realizó en sus totalidad mientras se presentaba una candidatura para un título de investigación en esta universidad.
- En los casos en que alguna parte de esta tesis haya sido presentada previamente para un título o cualquier otra calificación en esta universidad o cualquier otra institución, esto ha sido claramente establecido.
- En los casos en que he consultado la obra publicada de otros, esto siempre se cita claramente.
- Donde he citado el trabajo de otros, siempre se da la fuente. Con la excepción de tales citas, esta tesis es enteramente mi propio trabajo.
- He agradecido todas las fuentes principales de ayuda.
- Cuando la tesis se basa en el trabajo realizado por mí mismo conjuntamente con otros, he dejado claro exactamente lo que se ha hecho por otros y lo que yo mismo he reportado.

---

**Carlos Crispín Espinosa Ponce**



# Agradecimientos

Este trabajo no hubiera sido posible sin la ayuda de tanta gente que sería imposible listarla toda en este espacio. Muchas gracias a toda la gente que ha pasado por mi vida, especialmente en esta época tan interesante que nos ha tocado vivir.

Una mención especial a los sínodos: Dra. Almudena Zurita, Dra. Isabel Pérez, Dra. Aida Nava, Dr. Guillermo Blanc y Dr. Antonio Peimbert. Muchas gracias por tomarse el tiempo de revisar este trabajo. Su retroalimentación ha sido invaluable para la mejora de este trabajo.

Agradezco a los proyectos de PAEP, DGAPA/PAPIIT: IN100519, IA101217, IG100622, 107215, IN101220, y CONACyT: CB285080, FC-2016-01-1916 y CB2015-254132 por brindarme los apoyos necesarios para asistir a escuelas y congresos que permitieron desarrollar y complementar mi aprendizaje. Además, tuve la oportunidad de presentar mi trabajo en foros internacionales así como discutir mi trabajo y temas relacionados con otros investigadores.

Gracias al Instituto de Astronomía, sede CU por brindarnos el espacio y las herramientas necesarias para concluir mis estudios así como su apoyo durante la emergencia sanitaria. Gracias a Bertha y Heike por estar al tanto de nuestras dudas y su apoyo para los trámites necesarios.

Gracias a mi comité tutor: Dra. Leticia Carigi, Dra. Aida Nava y Dra. Irene Cruz por sus comentarios sobre mi trabajo y avances semestre a semestre. Por siempre estar al tanto de mi trabajo y su apoyo durante los últimos años de mi doctorado.

A mis tutores, Dr Sebastián Sánchez y Dr Christophe Morisset, por acompañarme durante todo este viaje. Por su apoyo y guía durante todo este tiempo, en especial durante los últimos años. Sin sus enseñanzas, consejos y palabras de apoyo, esto nunca hubiera sido posible.



# Índice general

<b>Dedicatoria</b>	<b>III</b>
<b>Declaración de autoría</b>	<b>V</b>
<b>Agradecimientos</b>	<b>VII</b>
<b>1. Introducción</b>	<b>1</b>
1.1. El medio interestelar . . . . .	1
1.1.1. El medio interestelar gaseoso . . . . .	3
1.2. Regiones HII . . . . .	4
1.2.1. Equilibrio de ionización . . . . .	6
1.2.2. Equilibrio térmico . . . . .	7
1.2.3. Temperatura y densidad de las regiones HII . . . . .	12
1.2.4. Composición química . . . . .	13
1.3. Datos para el estudio de las regiones HII . . . . .	17
1.3.1. Espectroscopía de campo integral . . . . .	18
1.3.2. CALIFA . . . . .	20
1.4. Análisis con PIPE3D . . . . .	24
1.4.1. Análisis del espectro central . . . . .	24
1.4.2. Teselado de los datos . . . . .	25
1.4.3. Análisis de poblaciones estelares . . . . .	26
1.4.4. Análisis de líneas de emisión . . . . .	27
1.5. Modelos de Fotoionización . . . . .	28
1.5.1. Parámetros de los modelos de fotoionización . . . . .	29
1.5.2. Modelos de fotoionización para el estudio de regiones HII . . . . .	31
1.6. Propiedades globales de las regiones HII en galaxias . . . . .	32
1.7. Objetivo del trabajo . . . . .	34
<b>2. Catálogo de regiones HII</b>	<b>35</b>
<b>3. Propiedades de las Regiones HII</b>	<b>63</b>
<b>4. Abundancias relativas de O/Fe</b>	<b>93</b>
<b>5. Exploración con Modelos de Fotoionización</b>	<b>101</b>
5.1. Estudio de las regiones HII con modelos de fotoionización . . . . .	101
5.1.1. Relación entre O/H, U y N/O . . . . .	102
5.1.2. Modelos de fotoionización para el estudio de regiones HII . . . . .	103

---

5.1.3. Redes Neuronales Artificiales . . . . .	103
5.2. Metodología . . . . .	105
5.2.1. Base de datos observacional . . . . .	105
5.2.2. Base de datos de modelos de fotoionización . . . . .	106
5.2.3. Entrenamiento de la ANN . . . . .	107
5.2.4. Ajuste de los datos observacionales . . . . .	113
5.3. Resultados . . . . .	117
5.4. Conclusiones . . . . .	119
<b>6. Discusión, conclusiones y trabajo a futuro</b>	<b>125</b>

# Capítulo 1

## Introducción

### 1.1. El medio interestelar

El medio interestelar (ISM por sus siglas en inglés) es un término usado para referirse a la materia y la energía que existe entre las estrellas en las galaxias (Osterbrock & Ferland, 2006; Draine, 2011). El principal componente del ISM es la materia en forma de gas que puede encontrarse en estado neutro e ionizado, y el polvo (Spitzer, 1978). La energía que se encuentra en el ISM es la radiación, rayos cósmicos, partículas de alta energía, campos magnéticos y la energía cinética del gas (principalmente turbulencia). No se considera parte del ISM la materia o energía oscura.

El gas interestelar está compuesto principalmente por hidrógeno ( $\sim 91\%$  por número de átomos) y se puede encontrar en su forma atómica, molecular o ionizada dependiendo de su temperatura y densidad. El helio es el segundo elemento más abundante en el ISM, el cual constituye el  $\sim 8.9\%$  en número. Otros elementos más pesados que el hidrógeno y el helio, como el C, N, O, Ne y Fe, y las moléculas, contribuyen solamente con el  $\sim 0.1\%$  de todo el gas. Por motivos históricos, dentro del lenguaje técnico astronómico, los elementos más pesados que el Helio se les llama *metales*.

El polvo interestelar proviene del material eyectado de las diversas etapas de la evolución de las estrellas. El material es lanzado hacia el ISM en su forma gaseosa y protogranos de polvo formados en las atmósferas de supergigantes rojas. Los granos de polvo pueden formarse a partir de dos precursores: i) los protogranos de polvo antes mencionados y ii) pequeños granos de polvo resultado de la fractura de granos grandes de polvo. Estas *semillas* acretan material por diversos procesos químicos, como hidrogenación, desabsorción térmica y fotólisis UV (Spitzer, 1978). El material del que se forman los granos de polvo está relacionado directa o indirectamente<sup>1</sup> con la depleción de los elementos pesados del gas del ISM. El polvo tiene un papel importante en el ISM dado que es el causante de la atenuación en las observaciones y es el principal redistribuidor térmico en el medio interestelar. Sin embargo, mientras el 99% de masa del ISM es gas, solamente el 1% es polvo (Boulanger et al., 2000; Klessen & Glover, 2016).

El ISM tiene un papel fundamental en la formación y evolución de las galaxias. En éste se encuentra toda la materia de la cual se forman las estrellas. Durante la vida de éstas, el ISM recibe su energía y, al morir las estrellas, es enriquecido por los elementos formados durante todo el proceso, p. ej. cadena PP, ciclo CNO, proceso-s y proceso-r. Además, al existir

---

<sup>1</sup>La depleción está relacionada indirectamente con la formación de polvo dado que los precursores de los granos de polvo contienen moléculas que de otra manera terminarían formando parte del gas del ISM



este complejo intercambio de energía y materia entre las estrellas y el medio interestelar, las propiedades del ISM se conectan con las propiedades de las estrellas y su proceso de formación (Matteucci, 2021; Sánchez et al., 2021a).

### Evolución química en el medio interestelar

El elemento más abundante en el universo es el hidrógeno,  $\sim 73\%$  por masa y más del  $90\%$  por número de átomos actualmente. Se utiliza el símbolo  $\mathbf{X}$  para indicar la fracción de masa del hidrógeno respecto a otros elementos. El segundo elemento más común actualmente en el universo es el helio, el cual conforma el  $\sim 26\%$  por masa en el universo observable. La mayoría del He se formó en los primeros 10 minutos después del Big Bang y otro poco es el resultado de la fusión nuclear del hidrógeno en las estrellas. Se representa con  $\mathbf{Y}$  la fracción de masa del helio. Los elementos más pesados que el helio, como C, O, N, etc, son llamados metales. La metalicidad, representada con la letra  $\mathbf{Z}$ , es definida como la fracción de masa de los elementos pesados con respecto a todos los elementos, de tal manera que  $X + Y + Z = 1$ .

La metalicidad es de interés para estudios de estrellas, cúmulos globulares, galaxias, cúmulos de galaxias, nubes moleculares, etc. Después de la nucleosíntesis primordial, el universo tenía un número casi nulo de metales. Con el tiempo, los elementos más pesados fueron sintetizados. La metalicidad depende de la edad, la historia o el origen de los objetos.

Los metales se forman por procesos termonucleares en el núcleo de las estrellas, durante las explosiones de supernova y en la fusión de objetos compactos, como estrellas de neutrones. Los elementos se eyectan al medio interestelar a diferentes escalas de tiempo debido a los diferentes procesos y etapas de la vida de una estrella. La mayoría de los metales permanecen en el núcleo de las estrellas y, al final de la vida de estas, terminarán en agujeros negros, enanas blancas y estrellas de neutrones (Kobayashi et al., 2020). Otra fracción de los metales es expulsado hacia el ISM por explosiones de supernova y vientos estelares enriqueciendo el medio para la siguiente generación de estrellas. Otros procesos como el proceso lento de captura de neutrones (proceso s) y el proceso rápido de captura de neutrones (proceso r) son los responsables de la producción de los elementos más pesados que el hierro.

La producción y enriquecimiento de metales de algunos elementos está directamente relacionada con el proceso de formación estelar. En particular los elementos  $\alpha$  (p. ej. O, Mg, Ca, Si, etc.), que se crean a partir de un núcleo y una partícula *alfa*<sup>2</sup>, son producidos en estrellas masivas. Estos metales se eyectan al ISM cuando las estrellas explotan como supernova por el colapso de su núcleo (*CCSNe* por sus siglas en inglés Core-Collapse SuperNovae). Por el contrario, los elementos del pico de hierro (*iron peak elements*), que son aquellos elementos con un número atómico cercano al del hierro (p. ej Ti, V, Cr, Mn, Fe), son producidos dentro de las estrellas y también en supernovas de sistemas binarios (Tipo 1a). Este tipo de elementos no están conectados con procesos de formación estelar recientes. Es por esto que las abundancias relativas de los dos grupos nos proporciona información sobre supernovas pasadas y la historia de formación estelar.

Uno de los elementos  $\alpha$  más importantes es el O dado que está ligado a la formación estelar y es el elemento más abundante después del H y He. Se produce en estrellas masivas así como en estrellas de masa intermedia. La causa principal del enriquecimiento de O en el ISM se debe a las CCSNe, aunque puede seguir aumentando lentamente debido a la contribución de estrellas AGB (siglas de *Asymptotic Giant Branch*). Esto último no pasa para otros elementos  $\alpha$  como Mg, Ca y Si, dado que no se producen en estrellas AGB de baja masa. Esto hace

<sup>2</sup>Un núcleo de He, dos protones y dos neutrones

que la abundancia de oxígeno sea un indicador de la evolución química de la galaxia (p. ej. [Burbidge et al., 1957](#); [Busso et al., 1999](#)).

### 1.1.1. El medio interestelar gaseoso

El medio interestelar en estado gaseoso se puede encontrar en su forma atómica, molecular y ionizada. Entre estas formas, el gas ionizado es el que ocupa la mayor parte del volumen del ISM. Sin embargo, la mayoría de la masa está concentrada en nubes densas dominadas por gas atómico neutro o gas molecular.

El gas del ISM puede ser descrito en términos de las distintas fases, que dependen de su estado térmico y químico. El *gas atómico neutro* se encuentra en dos fases: una fría y densa llamada *medio frío y neutro* (CNM por sus siglas en inglés) y otra difusa y caliente llamada *medio tibio y neutro* (WNM por sus siglas en inglés). Los primeros modelos para el CNM y WNM fueron propuestos por [Field et al. \(1969\)](#), basándose en el equilibrio térmico. Las propiedades de estos dos tipos de gas están bien definidas en el equilibrio térmico y de presión ([Kulkarni & Heiles, 1987](#); [Wolfire et al., 1995](#); [Savage & Sembach, 1996](#); [Cox, 2005](#)): el WNM tiene una alta temperatura ( $T \sim 8000$  K con una densidad  $n \sim 0.4 \text{ cm}^{-3}$ ) mientras que CNM tiene una baja temperatura ( $T \sim 80$  K con  $n \sim 40 \text{ cm}^{-3}$ ). El gas con una temperatura intermedia es inestable térmicamente. Puede enfriarse y hacerse más denso hasta combinarse con el CNM o calentarse y volverse menos denso hasta combinarse con el WNM ([Vázquez-Semadeni, 2012](#); [Ballesteros-Paredes et al., 2020](#)). Ambos tipos de gas atómico neutro, CNM y WNM, se estudian usando la transición de 21 cm del hidrógeno atómico neutro y líneas de absorción en el rango ultravioleta (UV) y óptico del espectro ([Heiles & Troland, 2003a,b](#)). Sin embargo, derivar las propiedades del gas neutro del ISM es muy complicado, la línea de 21 cm es complicada de observar dada la interferencia que puede producir la ionosfera, emisión de sincrotrón y emisión libre-libre. Por otra parte, las otras líneas de emisión usadas se ven fuertemente afectadas por la extinción por polvo, además de que no es nada sencillo separar la contribución de las dos fases.

La fase molecular del ISM está distribuida en nubes moleculares discretas (MC por sus siglas en inglés) de diferentes masas y tamaños (p. ej. [Blitz et al., 2007](#)). Las MC se pueden formar debido a la condensación de una nube difusa bajo el efecto de la gravedad o por presión externa (p. ej. [Elmegreen, 1990](#)). Las MC tienen bajas temperaturas ( $\sim 10 - 20$  K) con un tamaño típico que puede ir entre  $\sim 1$  pc a  $\sim 200$  pc (p. ej. [Miville-Deschênes et al., 2017](#)). Se catalogan de acuerdo a su tamaño y su masa como: i) nubes moleculares gigantes (GMCs por sus siglas en inglés) que tienen masas arriba de  $10^5 M_{\odot}$  y un tamaño de  $\gtrsim 30$  pc pudiendo alcanzar entre  $100 - 200$  pc, ii) nubes moleculares (MCs por sus siglas en inglés) con masas entre  $10^2$  a  $10^4 M_{\odot}$  con tamaños entre  $10 - 20$  pc, iii) grumos con masas entre  $10 - 10^2 M_{\odot}$  y un tamaño de algunos parsec, y iv) núcleos densos con masas por debajo de  $10 M_{\odot}$  y tamaños de  $0.1$  pc o menos (p. ej. [Blitz, 1993](#); [Heyer & Dame, 2015](#)). Mientras que las GMCs y MCs se encuentran rodeadas por un medio atómico tibio ( $T \sim 8000$  K), los grumos se encuentran dentro de las MCs, y los núcleos densos dentro de estos grumos. Las estructuras más densas siempre ocupan una muy pequeña fracción del volumen de las estructuras más grandes. Las MC presentan una estructura con filamentos los cuales juegan un rol importante en el proceso de formación estelar (p. ej. [André et al., 2010](#); [Hill et al., 2011](#); [Wang et al., 2015](#)).

La densidad en las MC es lo suficientemente grande ( $n \sim 10^2 - > 10^5 \text{ cm}^{-3}$  a  $T \sim 10 - 70$  K) para que se forme hidrógeno molecular ( $\text{H}_2$ ) en su interior. La mayoría de las moléculas de  $\text{H}_2$  se forman en la superficie de los granos de polvo ([Tielens, 2010](#); [Draine, 2011](#)). El hidrógeno molecular puede ser rápidamente fotodisociado por fotones UV ( $E > 11.2$  eV). Sin

embargo, las altas densidades hacen que la tasa de fotodisociación del  $\text{H}_2$  disminuya. También la extinción por polvo puede contribuir a disminuir esta tasa de fotodisociación (Ballesteros-Paredes et al., 2020). Las MCs se pueden encontrar en grandes agrupamientos (con  $10^6 M_\odot$ ) y tamaños de  $\sim 100$  pc como en pequeños grupos (con  $50 M_\odot$  con tamaños de  $\sim 1$  pc). Las nubes moleculares son una fracción importante de la masa total del ISM, del 30 al 60 %. Sin embargo, solo ocupan el 1 % del volumen total del medio interestelar (debido a su densidad).

Finalmente, las regiones más calientes en el ISM son las que están compuestas de gas ionizado. La fase más energética del medio interestelar es llamado el medio caliente e ionizado (HIM por sus siglas en inglés; Spitzer, 1956). El HIM es un gas caliente e ionizado que es el resultado de choques rápidos producidos, entre otros, por explosiones de supernova o vientos estelares. Estos procesos pueden dar lugar a grandes burbujas de gas caliente e ionizado ( $T \sim 10^5 - 10^7$  K con una  $n \sim 10^{-3} \text{ cm}^{-3}$ ) donde la ionización es producto de la colisión del medio con los electrones. El HIM presenta usualmente emisión en rayos X. Presenta líneas de absorción en el lejano UV y un continuo debido a radiación de frenado (también llamado *Bremsstrahlung* o radiación Libre-Libre). La última fase, el medio ionizado tibio (WIM por sus siglas en inglés), es el gas que ha sido ionizado por fotones UV que son producidos por diferentes procesos (estrellas calientes, AGNs, etc). El WIM tiene una temperatura y densidad características de  $T \sim 10^4$  K y  $n \sim 10^{-1} \text{ cm}^{-3}$ , respectivamente. Entre los diferentes objetos con este tipo de gas están las nebulosas planetarias, remanentes de supernovas, regiones H II, gas difuso ionizado, etc. Entre los más relevantes están: (i) el gas difuso ionizado que es el resultado de la ionización del medio difuso que está presente a lo largo de toda la extensión de las galaxias; (ii) las regiones H II que son gas ionizado y caliente alrededor de estrellas tipo OB. Ambos tipos producen líneas de recombinación y líneas prohibidas que pueden usarse para el estudio de sus propiedades físicas (p ej. Kumari et al., 2019; Marino et al., 2013; Kewley et al., 2019).

Todas las fases descritas anteriormente constituyen en su conjunto el medio interestelar. Las mismas pueden interactuar unas con otras, regulando el complejo sistema que conforma la evolución de las galaxias. Entre todas las fases del ISM gaseoso, las regiones H II se han estudiado ampliamente para entender la evolución química de las galaxias y los procesos de formación estelar.

## 1.2. Regiones HII

Las regiones H II son nebulosas discretas que envuelven un cúmulo de estrellas recién formadas. El gas es fotoionizado por la radiación de estrellas calientes y jóvenes tipo O y B (Strömgren, 1939; Osterbrock & Ferland, 2006). Este tipo de estrellas producen fotones lo suficientemente energéticos para ionizar al hidrógeno ( $> 13.6$  eV). Es por ello que a este proceso se le conoce como fotoionización. Además, dentro del gas se dan colisiones termalizadas de los iones de hidrógeno con otros iones y electrones. Como resultado, la temperatura final de la nebulosa ( $T_e$ ) puede estar en el rango de 7,000 K a 15,000 K con densidades ( $n_e$ ) de  $\sim 10^2 \text{ cm}^{-3}$ .

Hay dos procesos físicos que dominan la física de las regiones H II:

- Equilibrio de ionización
- Equilibrio térmico

Sin embargo, hay otros mecanismos que afectan a las regiones H II como: choques, vientos



Figura 1.1: Región H II 30 Doradus. Esta nebulosa se encuentra en la vecindad de la Vía Láctea en la llamada Gran Nube de Magallanes. La imagen es una composición de datos en tres bandas: óptico (verde, Hubble, NASA/STScI), infrarrojo (rojo, Spitzer, NASA/JPL/PSU/L) y rayos X (azul, Chandra, NASA/CXC/PSU/L). En la imagen, el color verde indica estrellas masivas que ionizan el gas circundante. La emisión infrarroja (IR), en color rojo, representa la radiación del gas y el polvo de las nubes moleculares. Finalmente, los rayos X en color azul muestran el gas a alta temperatura que, en este caso, se debe a frentes de choque generados por la gran actividad estelar de alta energía. Imagen tomada de la página de [Observatorio Chandra X-ray](#).



de salida (*outflows*), vientos estelares y frentes de ionización y fotodisociación<sup>3</sup>.

### 1.2.1. Equilibrio de ionización

Dentro de las regiones H II se producen dos procesos muy importantes: i) fotoionización y ii) recombinación. La fotoionización se da cuando colisiona un fotón con un electrón que se encuentra ligado a un átomo o ion. Si la energía del fotón es lo suficientemente alta, el electrón se desligará del átomo (o ion) al cual pertenece absorbiendo el fotón. Se le llama potencial de ionización a la energía mínima necesaria para arrancar un electrón del átomo. Este potencial es diferente para cada átomo e ion (y para cada nivel electrónico de cada átomo). Por ejemplo, para el hidrógeno es de 13.6 eV (ver Figura 1.3) mientras que para el helio neutro es de 24.6 eV. Suponiendo una nebulosa formada únicamente de hidrógeno con una densidad  $n_{\text{HI}}$ , podemos calcular la tasa volumétrica de ionización como:

$$n_{\text{HI}} \int_{\nu_0}^{\infty} \frac{4\pi J_{\nu}}{h\nu} a_{\text{H}}(\nu) d\nu, \quad (1.1)$$

donde la integral toma en cuenta todos los fotones con  $h\nu \geq h\nu_0$ , donde  $h\nu_0$  es el potencial de ionización de HI, 13.6 eV,  $J_{\nu}$  es la intensidad media de la radiación ionizante (en unidades de energía por unidad de área, tiempo, ángulo sólido e intervalo de frecuencia),  $h$  es la constante de Planck, y  $a_{\text{H}}(\nu)$  es la sección eficaz de fotoionización para el hidrógeno (donde la sección eficaz es la probabilidad de que se produzca el proceso de ionización por unidad de intensidad radiante).

La recombinación se produce cuando un ion captura un electrón libre. Tomando en cuenta la misma nebulosa formada de puro hidrógeno, podemos calcular la tasa volumétrica de recombinación,  $\Lambda_{\text{rec}}(T)$ , como:

$$\Lambda_{\text{rec}}(T) = n_e n_p \alpha(\text{H}, T), \quad (1.2)$$

donde  $n_e$  y  $n_p$  son las densidades de electrones y protones, y  $\alpha(\text{H}, T)$  es el coeficiente de recombinación radiativa (que es la tasa de recombinaciones por unidad de volumen para el ion y temperatura dada).

El equilibrio de ionización es el balance entre la ionización (Ec. 1.1) y recombinación de electrones e iones (Ec. 1.2):

$$n_{\text{HI}} \int_{\nu_0}^{\infty} \frac{4\pi J_{\nu}}{h\nu} a_{\text{H}}(\nu) d\nu = n_e n_p \alpha(\text{H}, T), \quad (1.3)$$

la cual determina la estructura de la nebulosa y la distribución espacial de los iones de los elementos en la zona ionizada.

De la Ec. 1.3 se puede deducir a qué distancia de la fuente ionizante el gas de la nebulosa pasa a estar en estado neutro, es decir gas no ionizado (suponiendo una geometría esférica). A este punto de transición se le llama *Radio de Strömgren*,  $R_S$  (Osterbrock & Ferland, 2006). Dentro de este radio, todos los fotones ionizantes se absorben. Este radio se define como:

$$Q(\text{HI}) = \frac{4\pi}{3} R_S^3 n_{\text{H}}^2 \alpha_{\text{B}}(\text{H}, T), \quad (1.4)$$

<sup>3</sup>Un frente de ionización es la frontera que separa el gas ionizado del gas neutro. Cuando una región H II se encuentra rodeada de gas molecular, bajo ciertas condiciones puede existir un frente de fotodisociación delante del frente de ionización. Los fotones con una energía  $11.2\text{ eV} < E < 13.6\text{ eV}$  que escapen de la nebulosa serán absorbidos por el  $\text{H}_2$  y solo el  $\sim 15\%$  de las absorciones será seguida por una fotodisociación (Draine, 2011)

donde  $Q(\text{HI})$  es el número de fotones ionizantes provenientes de la fuente por unidad de tiempo y  $\alpha_{\text{B}}(\text{H}, T)$  es el coeficiente de recombinación radiativa para el caso B (ver § 1.2.2).

Todos estos resultados son para el caso ideal donde la nebulosa se compone puramente de hidrógeno. Si bien una nebulosa está compuesta en su mayoría de hidrógeno, también contiene helio y otros metales (como oxígeno, nitrógeno, azufre, entre otros). Estos metales pueden contribuir en el equilibrio de fotoionización con otros procesos, por ejemplo intercambio de carga, recombinación dielectrónica y auto-ionización.

### 1.2.2. Equilibrio térmico

Las regiones ionizadas están sometidas a una interacción térmica entre procesos de calentamiento y enfriamiento. El calentamiento es una ganancia de energía por elemento de volumen, mientras que el enfriamiento es una pérdida de energía. La energía térmica final del sistema resulta del balance entre el *calentamiento derivado de la fotoionización* y el *enfriamiento derivado de la recombinación* y otras pérdidas radiativas (principalmente por emisión de líneas y radiación libre-libre).

Además de estos mecanismos de enfriamiento y calentamiento, el polvo tiene una contribución en el balance térmico de la nebulosa. El polvo absorbe radiación UV-óptico proveniente de las estrellas y se reemite como un continuo térmico en el rango del infrarrojo (Draine, 2003). También puede ionizarse y contribuir al calentamiento por efecto foto-eléctrico<sup>4</sup> (p ej. Draine, 1978; Weingartner & Draine, 2001).

#### Calentamiento por fotoionización

Como se mencionó en la Sec. 1.2.1, la fotoionización juega un papel importante en la estructura de la nebulosa. Desde un punto de vista térmico, la principal fuente de calentamiento de una región ionizada es precisamente este proceso. Aunque todos los elementos presentes en el gas son afectados por la fotoionización, el H y He son los que más participan en la producción de foto-electrones que calientan debido a su mayor abundancia. Otra fuente de calentamiento pueden ser choques (Peimbert et al., 1991), calentamiento fotoeléctrico por polvo (Draine, 1978) y rayos cósmicos (Peimbert & Goldsmith, 1972).

Suponiendo una nebulosa constituida de puro hidrógeno, la tasa de calentamiento neto por fotoionización viene dada por:

$$G(\text{HI}) = n_{\text{HI}} \int_{\nu_0}^{\infty} \frac{4\pi J_{\nu}}{h\nu} h(\nu - \nu_0) a_{\text{H}}(\nu) d\nu \text{ [erg cm}^{-3} \text{ s}^{-1}\text{]}. \quad (1.5)$$

La expresión define la entrada de energía por unidad de volumen y de tiempo. Además toma en cuenta el número de fotones ionizantes provenientes de la fuente ionizante,  $n_{\text{HI}}$ , la energía inyectada por fotoionización y la sección eficaz de fotoionización,  $a_{\text{H}}(\nu)$ .

#### Procesos de enfriamiento

El enfriamiento en las regiones H II tiene tres fuentes principales de enfriamiento: (I) Recombinación, (II) emisión de líneas colisionales, y (III) emisión por radiación de frenado.

---

<sup>4</sup>Es la emisión de electrones debido a la transferencia de energía entre radiación electromagnética y un material, comúnmente metales.

## Recombinación

La recombinación produce líneas de emisión (RLs por sus siglas en inglés) y un continuo debido un proceso llamado *Libre-Ligado*. La recombinación se produce cuando un electrón libre es capturado por un ion en un estado de energía diferente al nivel base. Posteriormente, el electrón decae a niveles de menor energía donde se libera un fotón con una energía, y una frecuencia, correspondiente a la diferencia de energía entre ambos niveles. Por otra parte, el proceso Libre-Ligado se refiere a cuando un electrón libre, con cualquier energía cinética dentro de la distribución termalizada considerada, es capturado por un ion a cierto nivel energético. Esto resultará en la emisión de un fotón con una energía igual a la diferencia de energía entre la energía del electrón libre antes de ser capturado y la energía del nivel en el cual es capturado. Si se tiene una distribución de electrones libres con diferentes energías cinéticas, entonces se producirá una emisión de continuo.

Las líneas de recombinación más importantes son las generadas por la recombinación de iones de hidrógeno con electrones libres. El hidrógeno recombinado resultante se encontrará en un estado excitado, por lo que se produce un proceso de cascada donde el electrón decae a algún estado excitado inferior hasta que alcanza el estado base dando lugar a la emisión de un fotón por cada transición electrónica. Además, también se emiten RLs debidas a los elementos pesados. Sin embargo, éstas líneas son más débiles debido a que la abundancia de estos elementos es menor.

El proceso de recombinación del hidrógeno, así como las longitudes de onda de las líneas asociadas a estas transiciones, se han estudiado a detalle (p. ej. Lyman, 1906, 1914; Balmer, 1885; Paschen, 1908). En el átomo de hidrógeno, las transiciones de un electrón de un nivel  $n \geq 2$ <sup>5</sup> a  $n = 1$  se les conoce como serie de Lyman y las transiciones de un nivel  $n \geq 3$  a  $n = 2$  se les conoce como serie de Balmer. Las líneas correspondientes a esta última pueden verse en la Fig. 1.2. Los cocientes relativos de las RLs presentan una dependencia con la profundidad óptica de la nebulosa, siendo las más afectadas las líneas de la serie de Lyman (al tener las longitudes de onda más cortas, correspondientes al UV cercano, ver Figura 1.3). Hay dos casos de interés en función de la profundidad óptica de la nebulosa: cuando esta es ópticamente delgada (Caso A de recombinación) y cuando es ópticamente gruesa (Caso B de recombinación). En general, un electrón puede recombinarse directamente al estado base (emitiendo un fotón con una energía  $> 13.6$  eV) o a un estado excitado ( $n \geq 2$ ) de un átomo de hidrógeno ionizado. En el segundo caso, el electrón alcanza el estado base por transiciones electrónicas entre niveles (y produciendo líneas de emisión). Cuando el fotón emitido tiene una energía  $> 13.6$  eV, este puede ionizar un átomo de H. En el caso A de recombinación se toman en cuenta las transiciones que caen al estado base o a un estado excitado ( $n \geq 2$ ). En este último caso, el átomo de H alcanzará el estado base por emitir líneas espectrales (y posiblemente por la emisión de un continuo debido a la transición  $2s$  a  $1s$ ). La consecuencia de este caso A es que todas los fotones correspondientes a las serie de Lyman pueden escapar de la nebulosa. Una nebulosa ópticamente gruesa (caso B de recombinación) se considera como el caso más realista para una región H II. En este último caso se consideran que ningún fotón con una energía  $> 13.6$  eV escapa de la nebulosa dado que es reabsorbido inmediatamente, por lo que no se toman en cuenta las recombinaciones de electrones que caen en el estado base. (Osterbrock & Ferland, 2006). En el caso B de recombinación las líneas correspondientes a la serie de Lyman no se observan, pero se pueden detectar las líneas de la serie de Balmer (las cuales tienen longitudes de onda que corresponden mayormente al rango visible, ver Figura 1.2).

<sup>5</sup>Donde  $n$  es el número cuántico principal

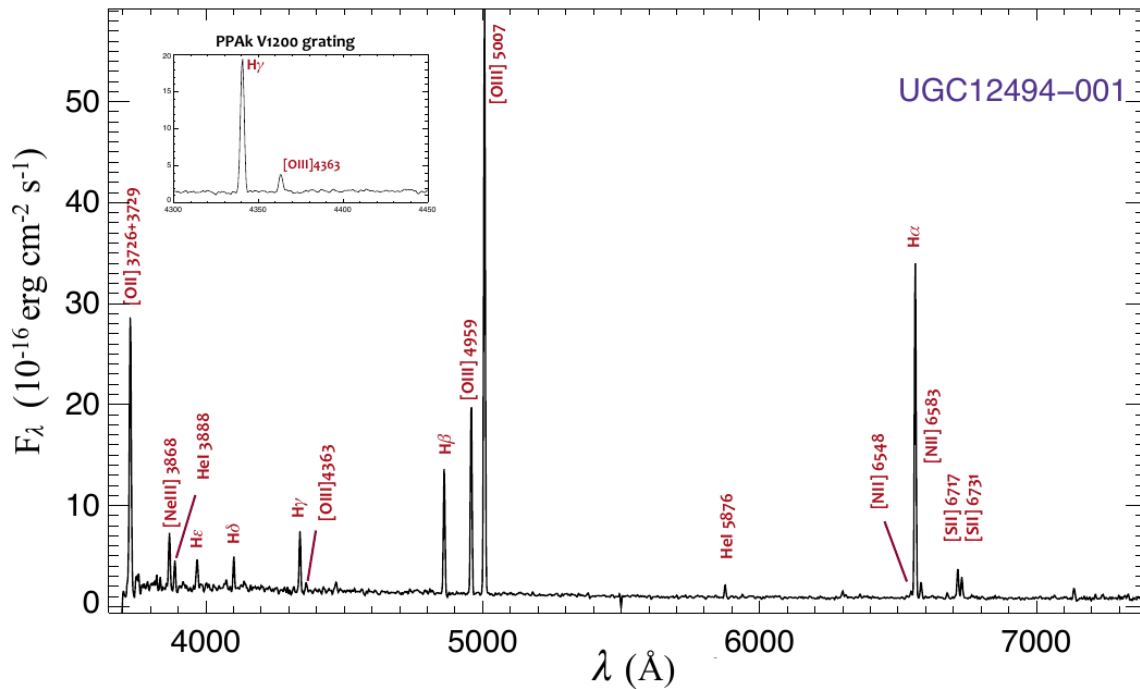


Figura 1.2: Espectro típico de una región H II identificada en la galaxia UGC 12494 observada dentro del proyecto CALIFA. El espectro principal fue tomado con la configuración V500 y la del cuadro pequeño con la configuración V1200 (ver sección 1.3.2) en el rango óptico del espectro. Las líneas más intensas son identificadas con su respectivo ion y longitud de onda. Figura extraída de [Marino et al. \(2013\)](#).



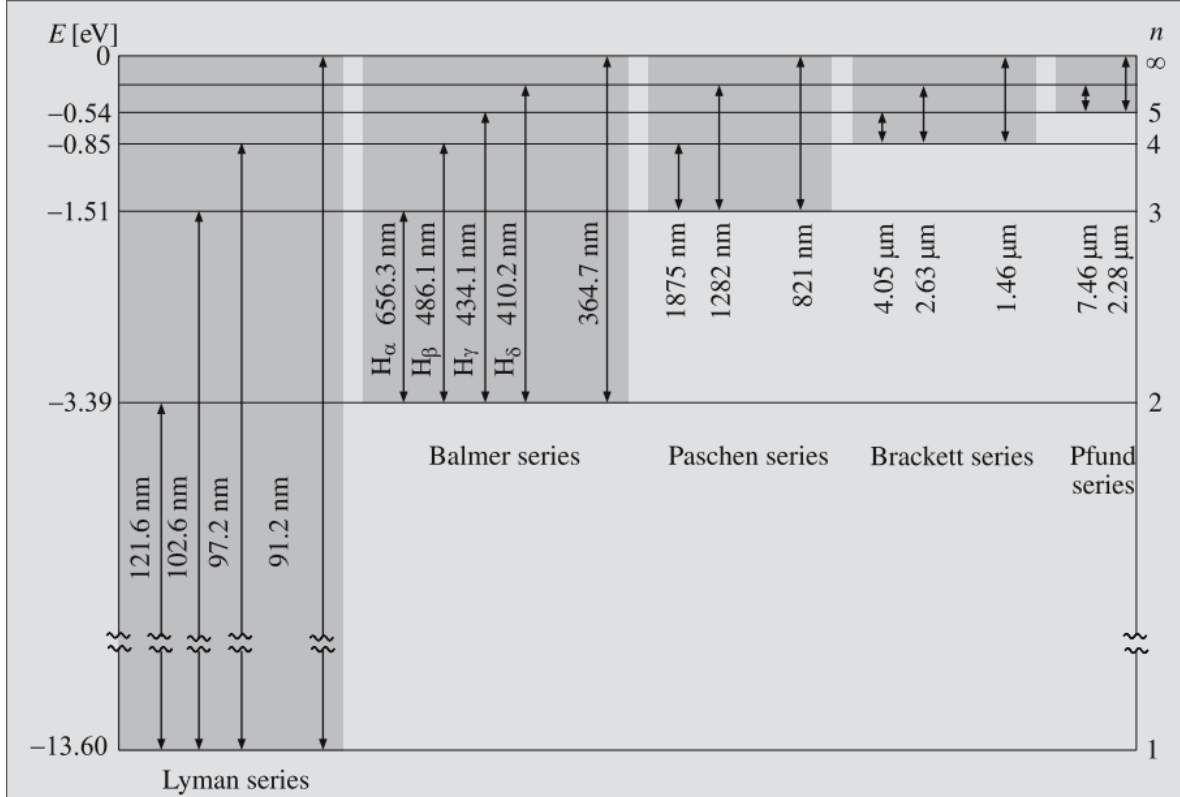


Figura 1.3: Transiciones para el hidrógeno. Se marcan las diferentes series que presenta este elemento. La longitud de onda correspondiente se marca en nm. La energía base se marca a  $-13.6$  eV en la parte inferior del diagrama. Figura extraída de [Karttunen et al. \(2017\)](#).

La pérdida de energía por recombinación para el hidrógeno está dada por:

$$L_R = n_e n_p k T \beta(H, T) [\text{erg cm}^{-3} \text{s}^{-1}], \quad (1.6)$$

donde  $kT\beta(H, T)$ <sup>6</sup> es el enfriamiento por unidad de volumen para recombinación, y  $\beta(H, T)$  es el coeficiente de recombinación del hidrógeno a una temperatura  $T$ .

Si la nebulosa está compuesta por más elementos, se define una expresión análoga a la Ec. 1.6 para cada uno de ellos.

### Excitación colisional

Las líneas de excitación colisional (CELs, por sus siglas en inglés) son la prueba observacional de la presencia de metales en regiones H II. Las CELs son el resultado de colisiones inelásticas entre los electrones y los iones de los metales. Esta colisión provoca que un electrón ligado al ion pueda alcanzar un estado excitado. Cuando el sistema (ion o átomo) se desexcita y vuelve un estado de menor energía el sobrante de energía es emitido como un fotón.

Aunque los iones metálicos presentan una baja abundancia, las CELs de algunos elementos son muy intensas, por ejemplo O ([O III]  $\lambda 5007$ ), N ([N II]  $\lambda 6583$ ) o S ([S II]  $\lambda 6731$ ). Esto es debido a la poca energía que se requiere para excitar estas líneas (se puede ver gráficamente

<sup>6</sup> $k$  es la constante de Boltzmann.

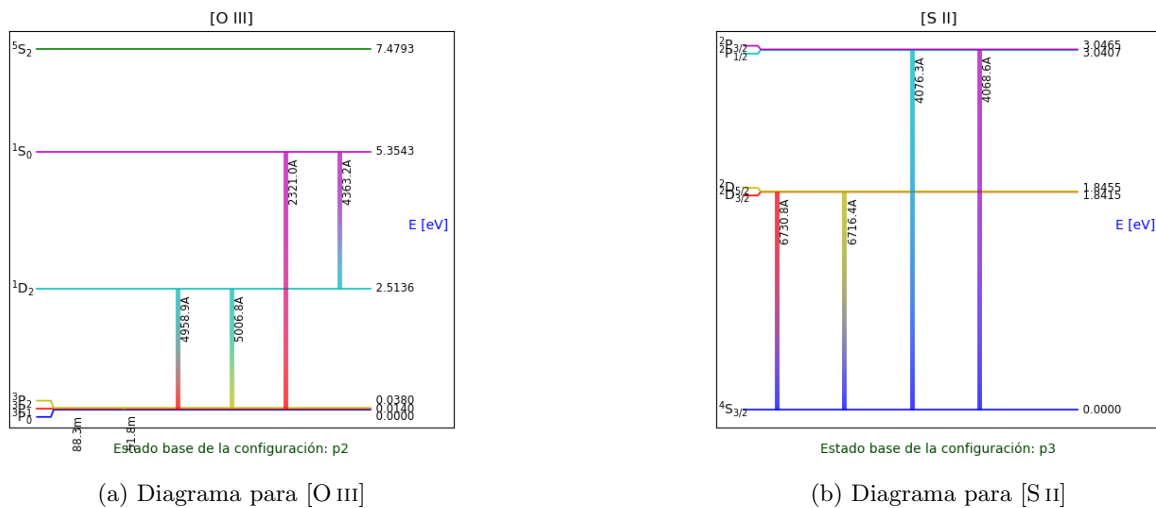


Figura 1.4: Diagramas de Grotian para [O III] y [S II]

en la Fig. 1.4a). Por otro lado, el flujo de las CELs dependen de la  $T_e$  y  $N_e$  de la nebulosa, por ejemplo las líneas del ion  $O^{++}$ : [O III]  $\lambda 4363$ , [O III]  $\lambda 4959$  y [O III]  $\lambda 5007$ .

Para un átomo o ion con  $N$  niveles, la tasa de enfriamiento total por radiación por excitaciones colisionales viene dada por la suma de cada una de las tasas de enfriamiento correspondientes a cada una de las líneas de emisión producidas (Osterbrock & Ferland, 2006):

$$L_C = \sum_{j=1}^N n_j \sum_{k < j} A_{jk} h \nu_{jk}, [\text{erg cm}^{-3} \text{s}^{-1}], \quad (1.7)$$

donde  $A_{jk}$  es la tasa de transición radiativa entre el nivel energético  $j$  y el nivel  $k$ .

Para poder calcular el enfriamiento por CELs se tiene que tener en cuenta que existe un equilibrio de todas las excitaciones/desexcitaciones colisionales y desexcitaciones radiativas que pueblan cierto nivel energético  $j$  contra todas las excitaciones/desexcitaciones colisionales y desexcitaciones radiativas que despueblan el mismo nivel  $j$  para cada uno de los elementos. Este equilibrio depende de la densidad electrónica,  $n_e$ , dando a lugar a dos límites: baja ( $n_e \rightarrow 0$ ) y alta densidad ( $n_e \rightarrow \infty$ ).

En el límite de baja densidad, después de cualquier colisión se produce una desexcitación radiativa y todas las desexcitaciones colisionales son despreciables. Mientras que en el límite de alta densidad, las desexcitaciones colisionales dominan completamente a las radiativas. Esto es debido básicamente al balance entre la probabilidad de que se produzca una colisión (que es baja en el primer caso y alta en el segundo) y la probabilidad de que las líneas se desexciten por radiación (que es la misma en ambos casos). A la densidad de separación de ambos límites se le llama densidad crítica ( $n_c(i)$ ), y su valor dependerá del ion, nivel y la temperatura. Por ejemplo, el ion [O II], para una  $T_e = 10,000$  K se tiene que para  ${}^2D_{5/2}^0$ ,  $n_c = 1.207 \times 10^3 \text{ cm}^{-3}$ , y para  ${}^2D_{3/2}^0$ ,  $n_c = 4.093 \times 10^4 \text{ cm}^{-3}$  (Osterbrock & Ferland, 2006).

La eficiencia del enfriamiento debido a CELs será mayor a altas densidades. En general, una CEL tendrá su mayor contribución de enfriamiento si la densidad es o esta cerca de la densidad crítica.

## Radiación Libre-Libre

El enfriamiento por radiación *libre-libre* se produce al interactuar dos partículas cargadas. En este caso, generalmente un electrón cambia su aceleración al encontrarse con el campo eléctrico de un protón o un núcleo atómico. El cambio de velocidad del electrón (frenado) genera la emisión de un fotón con una energía igual a la energía cinética perdida. Los electrones pueden encontrarse con los núcleos de cualquier elemento o con otros electrones que estén presentes en la nebulosa. El elemento que presenta una mayor contribución a la producción de este tipo de radiación es el hidrógeno seguido por el helio.

La pérdida de energía por este tipo de radiación para un ion con carga nuclear  $Z$  está dada por:

$$L_{ff}(Z) = \frac{32\pi e^6 Z^2}{3^{3/2} h m_e c^3} \left( \frac{2\pi kT}{m_e} \right)^{1/2} n_e n_+ g_{ff}, [\text{erg cm}^{-3} \text{ s}^{-1}], \quad (1.8)$$

donde  $n_+$  es la densidad de iones en número con carga nuclear  $Z$  en el gas,  $m_e$  y  $e$  la masa y carga del electrón y  $g_{ff}$  el factor de Gaunt para radiación libre-libre. Este último es un factor multiplicativo que describe la separación de los resultados de la física cuántica y clásica para absorción o emisión continua (Gaunt, 1930), y varía muy lentamente con la temperatura y con la densidad (Osterbrock & Ferland, 2006). Para condiciones en las nebulosas puede variar entre 1.0 a 1.5.

Este tipo de enfriamiento es el más ineficiente de todos. Sin embargo contribuye lo suficiente para disminuir un poco la temperatura de la nebulosa.

## Condición de equilibrio

La nebulosa estará en equilibrio térmico si la ganancia de energía es igual a la pérdida de esta:

$$G = L_R + L_C + L_{ff}. \quad (1.9)$$

La evaluación correcta de esta ecuación implica que tenemos que conocer qué iones están presentes en cada punto de la nebulosa. El proceso dominante de pérdida de energía son las líneas de excitación colisional (Osterbrock & Ferland, 2006).

El polvo puede agregar ciertos efectos además de modificar la radiación. Los granos de polvo pueden depletar a algunos iones importantes para el enfriamiento, esto puede afectar el enfriamiento general de la nebulosa.

### 1.2.3. Temperatura y densidad de las regiones III

La temperatura y densidad nebulares (o electrónicas) se entienden como la temperatura cinética promedio y densidad de los electrones libres, respectivamente. Sin embargo, puede haber variaciones de temperatura y de densidad entre diferentes zonas del gas ionizado (Peimbert, 1967; Garnett, 1992). Se necesita realizar una estimación precisa de la temperatura y densidad dado que otras propiedades físicas dependen de éstas.

La temperatura y densidad se pueden determinar con ciertos cocientes de líneas de emisión adecuados. Algunas razones de CELs son más sensibles a estas propiedades físicas (ver Fig. 1.5). Por ejemplo, ciertos cocientes de líneas dependen de la temperatura al tener diferentes energías de excitación. La temperatura nebular se puede determinar con el cociente de líneas de emisión generadas por el mismo ion pero provenientes de dos niveles atómicos con una

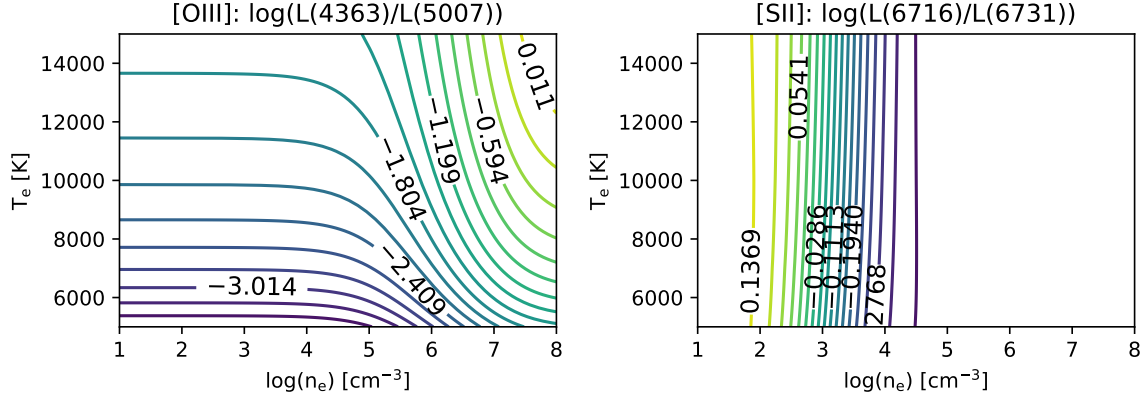


Figura 1.5: Rejilla de emisividad de los cocientes de líneas de emisión: (*Panel Izq*)  $\log([\text{O III}]\lambda 4363/5007)$  y (*Panel Der*)  $\log([\text{S II}]\lambda 6716/6731)$  para un rango de  $T_e$  desde 5,000 a 15,000 K y un rango de  $n_e$  desde  $10$  a  $10^8 \text{ cm}^{-3}$ . Generado con PYNEB (Luridiana et al., 2015). Las regiones de alta densidad ( $> 10^5 \text{ cm}^{-3}$ ) tienen cociente de emisividad constante independientemente de la densidad (esto se puede observar en el panel derecho).

diferencia clara de energía entre ambos (esto es, líneas de emisión que se observan a longitudes de onda con una separación amplia, ver Fig. 1.4a, Osterbrock & Ferland, 2006). Las líneas más usadas en el rango óptico son las líneas de O,  $[\text{O III}]\lambda\lambda(4959)/5007$  y líneas de N,  $[\text{N II}]\lambda\lambda 5755/(6548+84)$ .

La densidad electrónica puede ser determinada con cocientes de CELs emitidas por niveles con una energía de excitación similar pero con una diferente probabilidad de transición radiativa (esto es, dobletes de líneas con longitudes de onda muy similares, ver Fig. 1.4b). Esto hará que los cocientes de línea solamente dependan de la densidad. En el rango óptico, los cocientes ampliamente usados son los de las líneas de O,  $[\text{O II}]\lambda\lambda 3729/3726$ , y S,  $[\text{S II}]\lambda\lambda 6716/6731$ . La dependencia con la densidad de ambos cocientes de línea ha sido estudiado extensivamente desde hace mucho tiempo (p. ej. Aller et al., 1949; Osterbrock, 1960) para estimar densidades de  $10^{2-3} \text{ cm}^{-3}$ . Esto se debe cada línea de emisión tiene su propia tasa de transición debido a que tienen densidades críticas diferentes, por lo tanto el cociente de líneas varía con la densidad (Wei, 1988).

#### 1.2.4. Composición química

El gas del que están compuestas las regiones H II es el mismo con el cual se formaron las estrellas que las ionizan. Por lo tanto, la composición química de las regiones H II nos provee información sobre la generación de estrellas recién formadas y sobre el medio interestelar actual.

Como ya se mencionó anteriormente, los elementos más pesados que el helio son llamados metales. Se denomina metalicidad a la fracción de la masa o número de átomos de estos elementos pesados. Es importante resaltar que el oxígeno constituye el  $\sim 50\%$  (por número de átomos y de la masa) de todos los elementos pesados en el universo local. Podemos calcular las abundancias de los elementos pesados por medio de las líneas de emisión que producen las nebulosas, tanto las RLs y las CELs.

La intensidad de las líneas de recombinación es proporcional a la abundancia del ion emisor. Es por esto que las RLs más intensas son aquellas que provienen de los iones más abundantes,

por ejemplo, H y de He (e.g. [Bresolin et al., 2016](#)). Sin embargo, las RLs de los elementos más pesados son mucho más débiles y, en general, no se tiene la señal-a-ruido necesaria para determinar la abundancia (al menos con el tipo de datos que trabajaremos en esta tesis). Para trabajar con RLs y calcular abundancias se necesitan observaciones profundas con los telescopios de mayor apertura existentes y altas resoluciones espectrales. Por otro lado, todas las RLs presentan una dependencia similar con  $n_e$  y  $T_e$ . Esto es una gran ventaja porque una medida errónea de estos parámetros no afectará a la determinación de la abundancia. Sin embargo, solo se pueden usar cuando es posible su observación.

Por otra parte, la intensidad de una CEL presenta una alta dependencia con la  $T_e$  y en menor medida con la  $n_e$  por lo que es importante tener una buena medición de estos parámetros para usar este tipo de líneas para la determinación de abundancias. Dado que la intensidad es proporcional a la abundancia del ion que emite la línea, las CELs son ideales para determinar las abundancias de diversos elementos siempre que se cuente con una medición precisa de la  $T_e$ .

## Método directo

El método directo se basa en calcular todas las abundancias iónicas de un elemento dado, tal que la abundancia total de ese elemento se calcula como la suma de éstas:

$$\frac{n(X)}{n(\text{H})} = \sum_i \frac{n(X^{+i})}{n(\text{H}^+)}. \quad (1.10)$$

El hidrógeno es el elemento más abundante en el ISM. En las regiones fotoionizadas, donde la mayoría de los átomos están ionizados, el ion más abundante es  $\text{H}^+$ . Por esta razón, las abundancias se miden de forma relativa a  $n(\text{H}^+)$ .

De acuerdo a la Ec. (1.10), se necesitan todas las abundancias iónicas de un elemento para obtener la abundancia total. Para esto, se necesitaría la información de todos los iones presentes en el gas. Pero las observaciones del objeto de estudio pueden no incluir toda la información necesaria para poder calcular todas estas abundancias iónicas. Esto puede ser debido a que los datos estén restringidos a una parte del espectro o que las líneas de emisión sean muy débiles. En este caso se deben de tomar en cuenta los factores de corrección por ionización (ICF, por sus siglas en inglés). Estos factores toman en cuenta la contribución de aquellos iones que no podamos derivar, de tal forma que:

$$\frac{n(X)}{n(\text{H})} = \sum_{obs} \frac{n(X^{+i})}{n(\text{H}^+)} \times \text{ICF}. \quad (1.11)$$

Los ICFs se calculan usando conjuntos de modelos de fotoionización los cuales, en principio, toman en cuenta toda la física de las regiones fotoionizadas (p. ej. [Delgado-Inglada et al., 2014](#); [Amayo et al., 2020, 2021](#)). También se pueden calcular de acuerdo a potenciales de ionización similares entre algunos iones (p. ej. [Peimbert & Torres-Peimbert, 1971](#)). Sin embargo, de esta manera no se toman en cuenta todos los procesos que se llevan a cabo en un gas ionizado.

Dado que las líneas de recombinación son débiles, generalmente se usan CELs para calcular las abundancias iónicas. Para esto, se necesita una medida precisa de la temperatura electrónica. Como ya se ha mencionado en la § 1.2.3, las líneas más usadas para esta tarea son las líneas provenientes de los iones O III, N II y S III. Estas líneas incluyen las líneas aurorales [O III]  $\lambda 4363$ , [N II]  $\lambda 5755$  y [S III]  $\lambda 6312$ . Además, se necesita incluir una estimación de la densidad electrónica, la cual puede ser calculada con los dobletes de [O II] y [S II].

La precisión de este método recae en la calidad de la temperatura estimada. Sin embargo, las líneas aurales son débiles y solamente pueden medirse en objetos cercanos o en pocas regiones con una alta emisión (generalmente regiones de baja metalicidad). Para resolver este problema se han creado calibraciones empíricas las cuales relacionan las abundancias calculadas por el método directo con cocientes de línea de líneas fuertes (e.g. [Pagel et al., 1979](#); [Kewley & Dopita, 2002](#); [Marino et al., 2013](#); [Pilyugin & Grebel, 2016](#)).

Otra desventaja del método directo es cuando se tienen nebulosas ricas en metales. En este tipo de objetos existen gradientes de temperatura importantes debido al enfriamiento altamente eficiente debido a las líneas infrarrojas de [O III]. La consecuencia directa de este proceso es que el cociente [O III]  $\lambda\lambda 4363/5007$  no represente correctamente la temperatura y la abundancia estimada puede ser sobre o sub-estimada ([Stasińska, 2005](#)).

### Método de líneas fuertes

En los estudios basados en grandes bases de datos espectroscópicos de objetos extragalácticos las líneas aurales son difíciles de detectar. Por ejemplo, en el estudio de [Marino et al. \(2013\)](#), de  $\sim 3000$  regiones detectadas, solamente 16 de éstas tenían una medición fidedigna de las líneas aurales de [O III]. Los métodos de líneas fuertes están basados en cocientes de líneas mucho más intensas que las líneas aurales sensibles a la temperatura, siendo así trazadores indirectos de la metalicidad. Sin embargo, estos métodos son susceptibles a variaciones de otras condiciones físicas (p. ej. [Marino et al., 2013](#); [Bresolin, 2019](#); [Kewley et al., 2019](#); [Zurita et al., 2021](#)).

Este tipo de métodos se basa en obtener una calibración comparando cocientes de líneas, llamados indicadores, con la abundancia (generalmente de O o N, ver Fig. 1.6). Esta última puede obtenerse a partir de observaciones con el método directo (p. ej. [Marino et al., 2013](#); [Pilyugin & Grebel, 2016](#)) o usando conjuntos de modelos de fotoionización (p. ej. [McGaugh, 1991](#); [Kewley & Dopita, 2002](#); [Kobulnicky & Kewley, 2004](#); [Morisset et al., 2016](#)).

El primer calibrador de líneas fuertes se basó en el índice R23 ([Pagel et al., 1979](#)), definido como:

$$R23 = \log \left( \frac{[\text{OII}]\lambda 3727 + [\text{OIII}]\lambda 4959, 5007}{\text{H}\beta} \right) \quad (1.12)$$

Esta calibración se ha mejorado de manera empírica ([Pilyugin, 2000](#); [Pérez-Montero & Díaz, 2005](#)) y teórica ([Kewley & Dopita, 2002](#); [Kobulnicky & Kewley, 2004](#)), a lo largo de los años. El índice R23 mide la tasa de enfriamiento total debida a la intensidad de la línea del oxígeno en el óptico. Por lo tanto, este índice es sensible a la temperatura electrónica. Desafortunadamente, este cociente de líneas presenta un comportamiento bi-valuado con la abundancia de oxígeno (ver Figura 1.6, panel superior derecho). Por lo tanto hay dos rangos posibles: una de baja metalicidad y otra de alta metalicidad, y dos valores posibles para el mismo valor de R23. Para distinguir entre éstas dos *ramas* es necesario contar con información adicional.

Otros cocientes de líneas presentan el mismo comportamiento, por ejemplo algunos son  $S23 = ([\text{S II}] \lambda\lambda 6717, 6731 + [\text{S III}] \lambda\lambda 9069, 9532) / \text{H}\beta$ ,  $S2 = [\text{S II}] \lambda\lambda 6717, 6731 / \text{H}\alpha$  (ver Fig. 1.6). Los calibradores basados en un indicador bi-valuado necesitan de mediciones de otros cocientes de línea para seleccionar el rango de metalicidad adecuado. Sin embargo, cuando la abundancia calculada mediante este tipo de calibradores está en la zona de transición entre ambas ramas, las incertidumbres asociadas son muy grandes. Por otro lado, algunos indicadores no presentan este comportamiento bi-valuado (o este es mínimo), por ejemplo

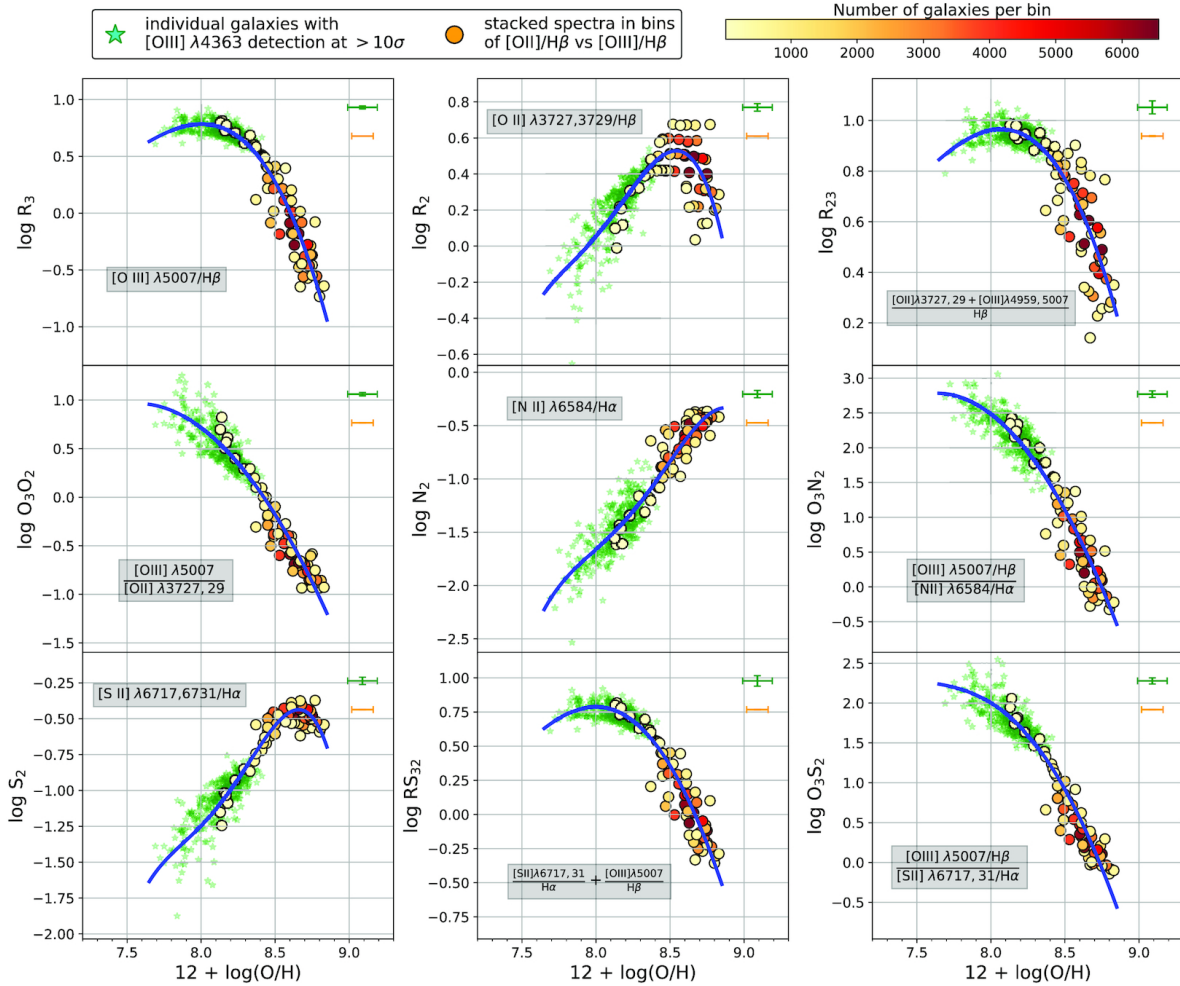


Figura 1.6: Ejemplos de diferentes indicadores usados en los calibradores de líneas fuertes usando datos del SDSS. Los puntos verdes corresponden a galaxias en las cuales las abundancias de oxígeno se derivaron con base en el método directo para galaxias individuales. Los puntos circulares corresponden al mismo tipo de derivaciones obtenidas para espectros apilados (*stacked* en inglés), de galaxias, los cuales individualmente no tenían la señal-a-ruido (S/N) suficiente en las líneas aurales para aplicar el método directo. El apilado y promediado de los mismos permite incrementar esta señal-a-ruido y aplicar el método directo. El código de color representa el número de espectros/galaxias apiladas para obtener cada valor representado. Las curvas azules representan los ajustes polinomiales que definen la calibración por líneas fuertes. Algunos de estos cocientes de líneas fuertes son bi-valuados. De Curti et al. (2020).



$O3N2 = [O III] \lambda 5007 / H\beta / [N II] \lambda 6583 / H\alpha$  (Pettini & Pagel, 2004; Marino et al., 2013) o  $N2O2 = [N II] \lambda \lambda 6548, 6583 / [O II] \lambda 3727$  (Kewley & Dopita, 2002; Bresolin, 2007). Sin embargo, la abundancia estimada por estos calibradores puede ser solo válida en un cierto rango de valores del índice y de la abundancia. En la Fig. 1.6 se muestran diversos indicadores y su dependencia con la abundancia de O para datos del SDSS. Se usaron espectros integrados de galaxias donde se detectaba la línea de  $[O III] \lambda 4363$  y espectros integrados apilados para alcanzar la SNR necesaria. En algunos de estos calibradores vemos claramente que la relación presenta una degeneración. Por otra parte, las relaciones derivadas dependen de los puntos observaciones en los cuales están basados. Por lo tanto, las abundancias derivadas serán válidas solamente en el rango que abarcaban los datos originales.

A pesar de que los calibradores de líneas fuertes son una buena alternativa al método directo, presentan también algunos problemas. Por ejemplo, existen discrepancias entre los resultados de los calibradores empíricos y los calibradores teóricos (los cuales están basados en modelos de fotoionización). Estos últimos predicen una abundancia  $\sim 0.2$  ó  $0.3$  dex más grande que sus contra partes empíricas. La causa de esta discrepancia sigue siendo un problema abierto (p. ej. Stasińska et al., 2006; Peña-Guerrero et al., 2012a).

Además, muchos de estos calibradores de líneas fuertes dependen del parámetro de ionización, y de otras propiedades de las nebulosas. Esta dependencia puede ser eliminada usando una técnica recursiva (Kobulnicky & Kewley, 2004), escogiendo cocientes de líneas menos sensibles a este efecto o usando un nuevo enfoque, por ejemplo, la inferencia bayesiana (Blanc et al., 2015; Vale Asari et al., 2016; Thomas et al., 2018) o utilizando un mayor número de líneas (p. ej. Pérez-Montero, 2014) o aprendizaje automático (*machine learning* en inglés; Ho, 2019). Otros calibradores de líneas fuertes están basados en líneas colisionales que dependen altamente de la temperatura electrónica (López-Sánchez et al., 2012). La solución obvia es usar cocientes de líneas que sean independientes de la temperatura electrónica (Esteban et al., 2004; Peimbert et al., 2005), es decir, usar líneas de recombinación. Desafortunadamente, para poder obtener una observación fidedigna se necesitan grandes telescopios y espectros de alta resolución, como ya se indicó anteriormente (Bresolin et al., 2016).

### 1.3. Datos para el estudio de las regiones HII

Una de las maneras más directas para el estudio de las regiones H II es con el uso de observaciones. La técnica conocida como espectroscopía de rendija larga ha sido usada ampliamente para el estudio de estos objetos (p. ej. Baldwin et al., 1991; Ho et al., 1997; Peña-Guerrero et al., 2012b). Los espectros obtenidos mediante esta técnica nos permiten detectar las principales líneas fuertes como  $[O III] \lambda 5007$ ,  $[N II] \lambda 6548$ ,  $[N II] \lambda 6583$ ,  $[S II] \lambda 6717$ , etc. Debido a esto, este tipo de observaciones se han usado ampliamente para estimar abundancias iónicas y condiciones físicas. Para la medición de líneas débiles, como  $[O III] \lambda 4363$  y  $[N II] \lambda 5755$ , se hace uso de espectros profundos con rendija larga. Desafortunadamente, la mayor desventaja de esta técnica es que no es posible obtener el espectro de cada parte de un objeto extendido.

A finales del siglo XX se empezó a desarrollar una nueva técnica que mejoraría el estudio de objetos extendidos: la espectroscopía de campo integral. Esta técnica permitiría obtener múltiples espectros a lo largo de un campo de visión de dos dimensiones simultáneamente (Kissler-Patig et al., 2007). En esta sección se describirán los datos usados en este trabajo.



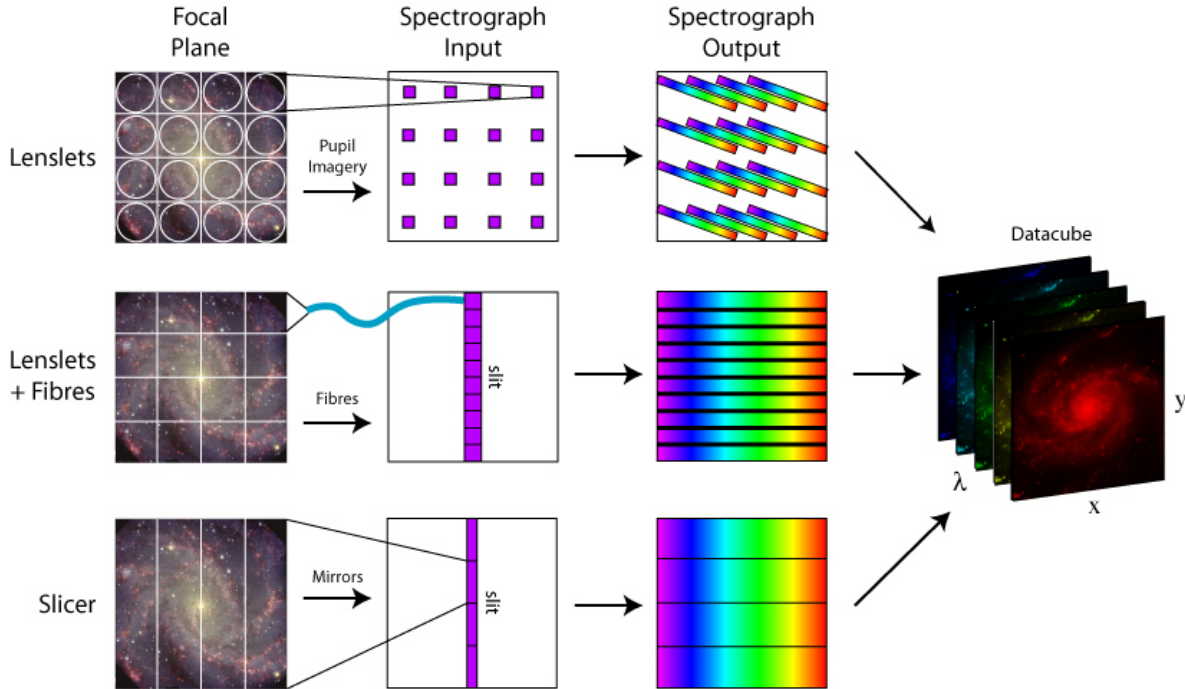


Figura 1.7: Esquema de las diferentes técnicas usadas en la espectroscopía de campo integral. Imagen adaptada de [Allington-Smith et al. \(1998\)](#) por M. Westmoquette.

### 1.3.1. Espectroscopía de campo integral

La principal desventaja de las técnicas tradicionales de observaciones espectroscópicas (como espectroscopía de rendija larga) es la limitación impuesta por la apertura de entrada. Si se tiene un objeto extendido (p. ej. una galaxia) o se requiere observar varios objetos al mismo tiempo, se puede optar por realizar diferentes observaciones. Sin embargo, esto no es eficiente dado que las observaciones no se realizan al mismo tiempo y se ocupa mucho tiempo del telescopio para un solo objeto.

La espectroscopía de campo integral (IFS por las siglas en inglés: *Integral Field Spectroscopy*) soluciona los problemas de las técnicas tradicionales de observación al proveer datos espacialmente resueltos. Esto quiere decir que podemos obtener los espectros de diversos puntos del cielo sobre un campo de visión, continuo o cuasi-continuo (FoV por sus siglas en inglés, *Field of View*) de dos dimensiones al mismo tiempo. Esto nos permite que podamos resolver espacialmente las galaxias y estudiar las variaciones espaciales de las distintas propiedades físicas de las poblaciones estelares y del gas (e. g. [Relaño et al., 2010](#); [García-Benito et al., 2017](#); [Sánchez-Menguiano et al., 2018](#); [Sánchez, 2020](#))

La IFS utiliza un espectrógrafo de campo integral que consiste de dos componentes: un espectrógrafo y una unidad de campo integral (IFU por sus siglas en inglés *integral field unit*). Este último tiene la función de dividir el campo de visión de dos dimensiones en un arreglo de elementos espaciales, llamados *spaxels*, y dirigir el haz de luz de cada uno al espectrógrafo.

Podemos clasificar a los IFUs de acuerdo a la técnica usada para dividir el FoV, distinguiendo tres tipos principales ([Allington-Smith, 2006](#)):

- Arreglo de Lentes: Este tipo de IFU se compone por un arreglo de microlentes el cual divide el plano focal del telescopio (primera fila en Figura 1.7). La luz de cada micro-lente es enfocada en un pequeño punto y enviada al espectrógrafo, del cual obtenemos

su espectro correspondiente. Este IFU es el más simple de los tres y fue desarrollado por primera vez para el instrumento TIGER (Bacon et al., 1995). Desafortunadamente, para evitar una superposición de los espectros provenientes de cada lente, se tiene que limitar a un rango espectral pequeño.

- **Fibras ópticas:** Este tipo de IFU se compone de un arreglo bidimensional de fibras ópticas que conducen la luz al espectrógrafo. El campo de visión puede ser reacomodado en un formato tipo rendija dado que la fibras presentan cierta flexibilidad (segunda fila en Figura 1.7). La luz es redirigida al espectrógrafo y, debido a la reorganización del FoV, los espectros obtenidos no tienen ningún desplazamiento en longitud de onda entre ellos. Esta es una de las técnicas más usadas actualmente, por ejemplo DensePak (Barden & Wade, 1988), GMOS (Allington-Smith et al., 1997), SAMI (Croom et al., 2012), MaNGA (Bundy et al., 2015).
- **Rebanador de imagen:** Se compone de dos espejos, un espejo segmentado en secciones delgadas que manda en una dirección diferente cada *rebanada* de la imagen de entrada; y un segundo espejo que reacomoda las rebanadas tal que se forme una pseudo rendija de forma que la luz reflejada se envía al espectrógrafo (última fila en Figura 1.7). Este tipo de IFUs es sencillo de enfriar permitiendo que se puedan alcanzar longitudes de onda en el infrarrojo donde la emisión de fondo térmico es prominente. Sin embargo, este tipo de instrumentos pueden llegar a ser bastante grandes y costosos. Algunos instrumentos que usan esta técnica son SINFONI (Eisenhauer et al., 2003) y MUSE (Bacon et al., 2010).

Generalmente el producto final del proceso de reducción para estas técnicas es un cubo de datos donde dos ejes son espaciales ( $x$ ,  $y$  o RA y DEC) y el tercer eje corresponde a la longitud de onda. Esto quiere decir que por cada spaxel en el FoV tenemos un espectro, o dicho de otro modo, tenemos una imagen del objeto para cada longitud de onda (Figura 1.8). Esto nos permite estudiar componentes estructurales de las galaxias, tales como discos, brazos espirales, regiones H II, etc.

SAURON (de Zeeuw et al., 2002) fue uno de los primeros muestreos en usar la IFS para observar un grupo de galaxias de tipo temprano,  $\sim 70$  objetos. Este primer estudio se enfocó en estudiar las propiedades cinemáticas de las poblaciones estelares y del gas ionizado, así como la forma intrínseca de las galaxias. El proyecto ATLAS3D (Cappellari et al., 2011) fue el primer gran muestreo de galaxias en usar la técnica de IFS. Con 260 galaxias observadas de tipo temprano, este proyecto fue diseñado para estudiar la cinemática y dinámica estelar global y su relación con su formación y evolución.

Uno de los primeros proyectos diseñados para estudiar las abundancias químicas y las propiedades globales de galaxias de disco fue PINGS (Rosales-Ortega et al., 2010). Su muestra constaba de 17 galaxias de tipo tardío del universo local. Los estudios derivados de estos datos mostraron que las mediciones de las líneas de emisión de una región H II son espacialmente dependientes, apreciándose claramente gradientes de metalicidad decrecientes (Sánchez et al., 2011; Rosales-Ortega et al., 2011). Además, las observaciones especialmente resueltas de regiones H II permitieron encontrar la relación local entre la masa, la metalicidad y la formación estelar (Rosales-Ortega et al., 2012).

Los primeros muestreos que incluyeron cientos de galaxias fueron CALIFA (Sánchez et al., 2012a) y SAMI (Croom et al., 2012), con 667 y 3068 galaxias observadas, respectivamente, en su última publicación de datos. En los últimos años han surgido muestreos mucho más grandes que incluyen la información de más de 10000 galaxias, como MaNGA (Bundy et al.,

2015), y otros que mejoran considerablemente algunas propiedades de los datos, como los que hacen uso del instrumento MUSE (Bacon et al., 2010; López-Cobá et al., 2020; Emsellem et al., 2022) y su alta resolución espacial.

En este trabajo se usaron datos de galaxias del proyecto CALIFA. En la próxima sección se describe las propiedades de estos datos así como su análisis.

### 1.3.2. CALIFA

El proyecto Calar Alto Legacy Integral Field Area (CALIFA) es uno de los muestreos de galaxias más completos hasta la fecha. A diferencia de los estudios anteriores que se centraban en cierto tipo de galaxias, como PINGS o SAURON, la muestra de CALIFA incluye galaxias de cualquier tipo morfológico en el universo local. La selección de la muestra (mencionada más adelante) permite un estudio en toda la extensión de la galaxia. Además, el rango espectral incluye la parte roja y azul del espectro visible (de  $3745 - 7500\text{\AA}$ ), lo cual permite recuperar la información de la componente estelar y estudiar las principales líneas de emisión desde  $[\text{O II}] \lambda 3727$  a  $[\text{S II}] \lambda 6731$ . La línea de  $[\text{O II}] \lambda 3727$  puede observarse debido al redshift de las galaxias.

El muestro CALIFA usa el espectrógrafo *Postdam Multi Aperture* (PMAS por sus siglas en inglés, Roth et al., 2005) en el modo PPaK (*PMAS fiber Package*, Kelz et al., 2006) montando en el telescopio de 3.5m del observatorio de Calar Alto. Este IFU tiene 382 fibras, con un diámetro de  $2.7''$ , divididas en tres grupos (ver Figura 1.9): (I) 331 fibras dispuestas en un arreglo hexagonal con un FoV de  $74'' \times 65''$  usadas para observar las galaxias; (II) 36 fibras divididas en subgrupos de 6 fibras alrededor del arreglo hexagonal para muestrear el cielo; y, (III) 15 fibras que pertenecen a la unidad de calibración. El factor de llenado (área cubierta por las fibras con respecto al campo total cubierto por el haz hexagonal central) del arreglo hexagonal principal es del  $\sim 60\%$ . Para lograr tener una cobertura total del FoV se procedió a realizar tres apuntados, los cuales están separados una distancia menor a la separación entre fibras entre si, siguiendo un patrón triangular. A este proceso se le conoce como *dithering* y tiene como finalidad cubrir los espacios entre las fibras (Sánchez et al., 2007; Rosales-Ortega et al., 2010).

Las observaciones realizadas por CALIFA usaron dos diferentes configuraciones: (I) Configuración de baja resolución (V500) que cubre el rango óptico del espectro ( $3745 - 7500\text{\AA}$ ) con una resolución nominal de  $R = \delta\lambda/\lambda \sim 850$  a  $5000\text{\AA}$  que corresponde una anchura media altura<sup>7</sup> de  $\sim 6\text{\AA}$ ; (II) Configuración de resolución media (V1200) que cubre el rango espectral de  $3400 - 4848\text{\AA}$  con  $R \sim 1650$  a  $4500\text{\AA}$  que corresponde a  $\text{FWHM} \sim 2.7\text{\AA}$ .

La configuración V500 se usa para estudiar las poblaciones estelares y las propiedades del gas ionizado. Por otra parte, la configuración V1200 se usa para estudiar la cinemática de las componentes estelar y gaseosa. Sin embargo, esta última necesita de condiciones atmosféricas específicas y requiere de mayor tiempo de telescopio que la configuración V500. Por esto, no todos los objetos se observaron con ambas configuraciones. A lo largo del estudio descrito en esta tesis se usaron solamente los datos de la configuración V500.

La muestra inicial de CALIFA, el cual llamaremos *muestra madre*, se seleccionó principalmente por el diámetro de las galaxias, de tal manera que se cubriera totalmente la extensión óptica de los objetos (Walcher et al., 2014). Para lograr este objetivo, se escogieron aquellas galaxias del catálogo fotométrico del *Data Release 7* del SDSS Abazajian et al. (2009) que tuvieran un diámetro isofotal angular<sup>8</sup>,  $D_{25}$ , tal que  $45'' < D_{25} < 80''$ . La muestra se confor-

<sup>7</sup>en inglés *Full Width at Half Maximum*, *FWHM*

<sup>8</sup>El diámetro isofotal angular es el eje mayor de la isofota a  $25 \text{ mag/arcsec}^2$  en la banda *r* del SDSS

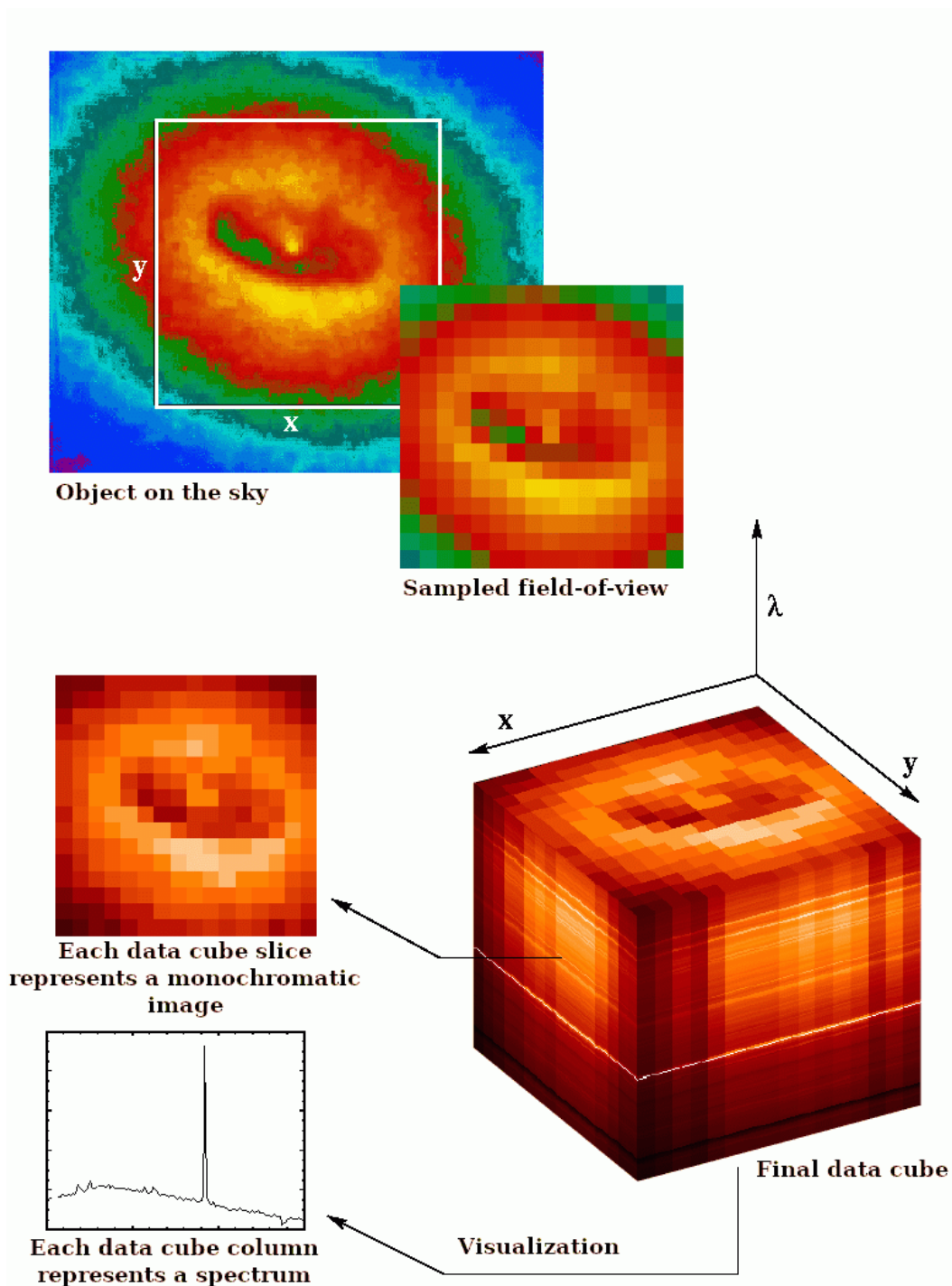


Figura 1.8: El principio de la espectroscopía de campo integral es obtener datos espacialmente resueltos de los objetos. Los cubos de datos finales tienen dos coordenadas espaciales y una coordenada que corresponde a la longitud de onda. Para cada spaxel se tiene un espectro, o para cada longitud de onda se tiene una imagen del objeto. Imagen adaptada de la original hecha por M. M. Roth

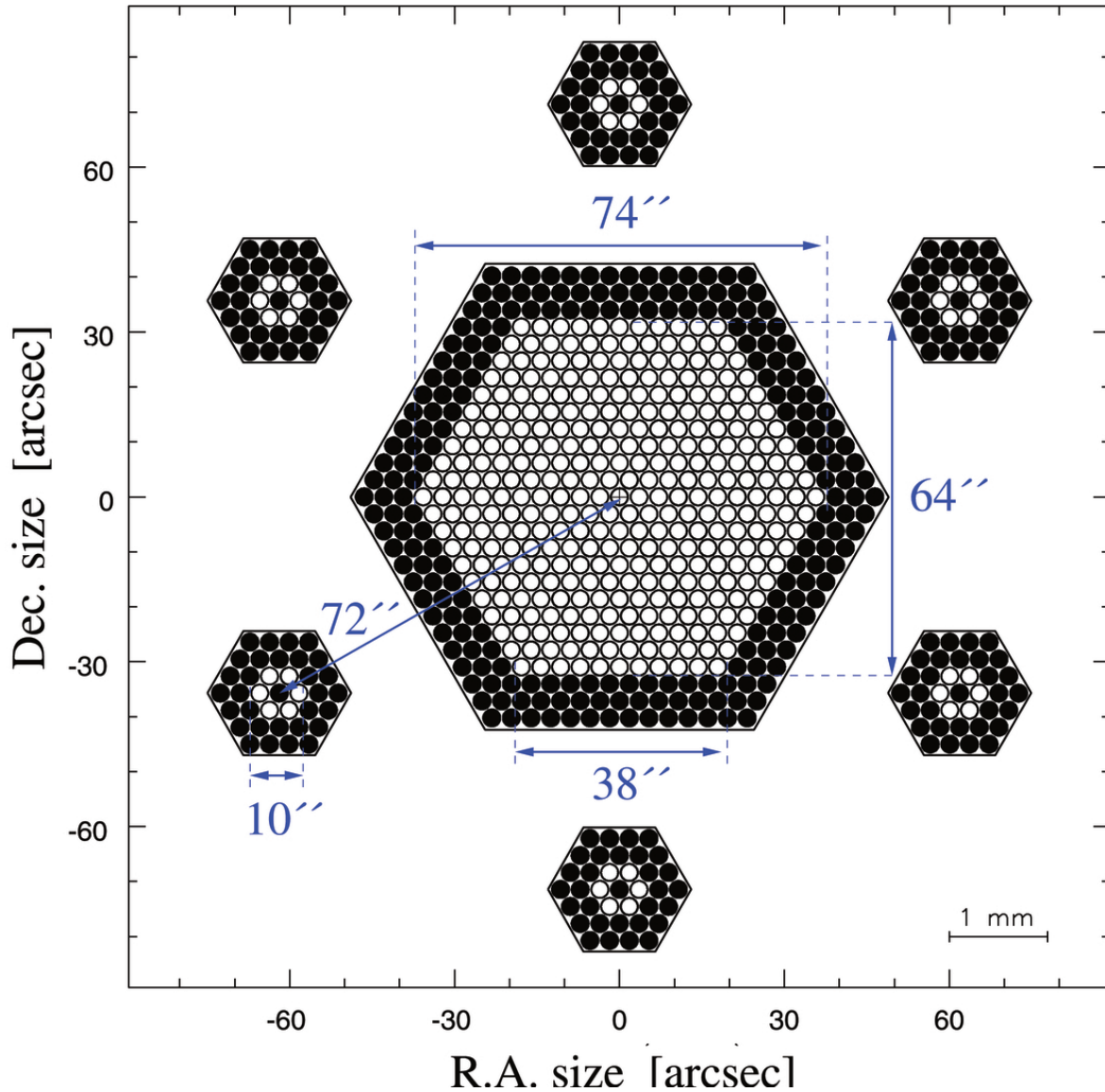


Figura 1.9: Distribución de las 382 fibras de PPaK. El arreglo hexagonal principal tiene 331 fibras. Además, seis grupos de fibras a 72'' del centro que rodean la estructura principal se incluyeron para medir el cielo. Finalmente, el conjunto incluye 15 fibras de calibración no mostradas en la figura. Los círculos blancos representan las fibras activas. Figura extraída de [Kelz et al. \(2006\)](#).



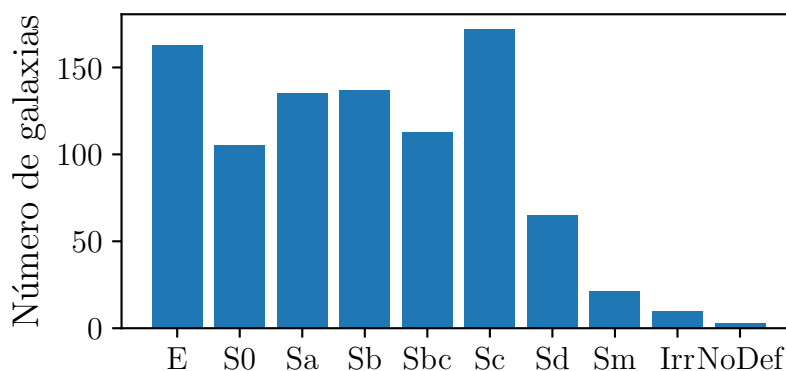


Figura 1.10: Número de galaxias según su tipo morfológico. La muestra eCALIFA contiene información de galaxias tempranas y tardías dentro del universo cercano. La etiqueta *NoDef* se refiere a galaxias que no pudieron ser clasificadas morfológicamente.

ma de galaxias del universo local ( $0.005 < z < 0.03$ ). El último *Data Release* publicado de CALIFA (Sánchez et al., 2016a) contenía un total de 667 galaxias observadas.

La muestra se extendió al incluir la compilación de 232 galaxias donde se hayan producido supernovas recientemente (PISCO, por sus siglas en inglés, *PMAS/PPaK Integral-field Supernova hosts COmpilation*, Galbany et al., 2018). Esta agregó a CALIFA objetos de baja masa, resultando con una masa estelar dentro del rango  $10^8 M_{\odot} < M_{*} < 10^{11.5} M_{\odot}$ . Finalmente, para el presente proyecto, se agregaron las galaxias observadas cuando el proyecto CALIFA se encontraba en fase *piloto*. A esta muestra la llamaremos pCALIFA. La muestra final de galaxias incluyen los tres muestreos anteriormente mencionados con 924 galaxias observadas, a esta muestra le llamaremos eCALIFA.

La muestra final tiene información de galaxias desde tipo temprano hasta tipo tardío. El mayor número de galaxias se encuentra entre los tipos E y Sc, con un número menor de galaxias Sd, SM e irregulares (como se ve en Figura 1.10). Los tipos morfológicos que menos presencia tienen en la muestra son Sm e irregulares. Igualmente, la muestra cubre un gran rango en masas estelares ( $10^8$  a  $10^{11.5}$ , solo las galaxias enanas no se han cubierto), lo cual lo hace una muestra representativa de las galaxias del universo cercano.

La reducción de datos se realizó mediante por códigos semiautomáticos siguiendo todos los procedimientos estándar. La versión usada es la V2.2 desarrollada en Python descrito en Sánchez et al. (2016c). El proceso de reducción incluye los pasos usuales para este tipo de datos (descritos en detalle en Sánchez, 2006): (i) sustracción de BIAS y corrección por CCD-flat; (ii) trazado y extracción de los espectros; (iii) corrección por distorsión y calibración en longitud de onda; (iv) calibración en flujo; (v) corrección por refracción atmosférica diferencial (*DAR*, por sus siglas en ingles); y (vi) re-ordenación y re-muestreo espacial de los datos. El producto final del proceso de reducción de los datos es un cubo de datos con una resolución espacial  $\text{FWHM} \sim 2.5''$  y una spaxel de tamaño  $1''$ . Esto corresponde una resolución física promedio de  $\sim 0.8 \text{ kpc}$  (dado el rango en redshift antes mencionado). En la siguiente sección se incluye un resumen del análisis que se hace a los datos una vez reducidos.

## 1.4. Análisis con pipe3D

El análisis espectral de los datos reducidos se realizó con el programa PIPE3D (Sánchez et al., 2016c). Este es un *dataducto* automatizado, el cual tiene como finalidad la segregación y extracción de las propiedades de las dos principales componentes de un espectro: (1) la componente estelar que domina el continuo y las líneas de absorción, y (2) la componente del gas ionizado a la cual le corresponden las líneas de emisión a ciertas longitudes de onda. El producto final del análisis son un conjunto de mapas 2D de las propiedades estelares y del gas ionizado, los cuales se entregan en cubos de datos, llamados *dataproducs*.

PIPE3D está desarrollado para procesar datos espacialmente resueltos usando la técnica IFS. Además ha sido probado y usado para proyectos como SAMI, CALIFA, MaNGA y MUSE. Originalmente se desarrolló en el lenguaje PERL y está descrito a detalle en Sánchez et al. (2016c). Una nueva versión escrita totalmente en Python, llamada PYPIPE3D ha sido desarrollada y presentada en Lacerda et al. (2022). A continuación se dará una descripción general del procesado que realiza sobre los datos.

### 1.4.1. Análisis del espectro central

El primer paso es el análisis de la región central del cubo de datos. En el caso de los datos de CALIFA se usa una apertura de 5" de diámetro centrada en el spaxel de mayor intensidad en una banda específica. Esta banda depende de las características de las observaciones. Para CALIFA se usa la banda *V* para la configuración V500 y la banda *B* para la configuración V1200.

Sobre este espectro central se realiza un análisis de las poblaciones estelares y líneas de emisión siguiendo los procedimientos descritos en Sánchez et al. (2016b). Como primer paso se realiza una estimación de parámetros cinemáticos. Esto se basa en un ajuste del espectro con una plantilla de poblaciones estelares sintéticas simples (SSP por sus siglas en inglés *Single Stellar Populations*).

Una SSP es un modelo que describe la evolución en el tiempo de la distribución de energía espectral (*SED* por sus siglas en inglés) producido por un brote de formación estelar a partir de un gas de una cierta composición química y a un tiempo dado. Por lo que, una SSP requiere i) la teoría de evolución estelar en forma de isócronas (p ej. Meynet & Maeder, 2000; Marigo et al., 2008); ii) librerías espectrales estelares, que nos proporciona el espectro correspondiente del resultado de la evolución estelar, pueden ser *Teóricas* (p ej. Smith et al., 2002; Martins et al., 2005; Coelho et al., 2007) o *Empíricas* (Sánchez-Blázquez et al., 2006; Le Borgne et al., 2003); y iii) una función inicial de masa (p ej. Salpeter, 1955; Kroupa, 2001; Chabrier, 2003), que nos describe la distribución inicial de masa de una población de estrellas.

El polvo es otro de los ingredientes que debe de tomarse en cuenta en las SSPs. Este oscurece la luz en el rango entre el UV y el infrarrojo cercano (NIR por sus siglas en inglés Calzetti, 2001) y emite en el infrarrojo (IR por sus siglas en inglés Mathis et al., 1977). Estas dos contribuciones se deben de tomar en cuenta al utilizar las SSPs como base para sistemas estelares más complejos.

El espectro observado de una región de una galaxia (o una galaxia entera) puede ser considerado como la suma de los espectros de todas las estrellas que la componen y del gas (considerando el ensanchamiento por la dispersión de velocidades, el oscurecimiento por polvo y con cierto corrimiento al rojo). El espectro observado,  $S_{\text{obs}}$ , idealmente sería la suma de los espectros evolucionados a lo largo del tiempo de todas las poblaciones estelares que se han generado en los diferentes brotes que conforman la historia de formación estelar en una región

(o galaxia), es decir, de las diferentes SSP:

$$S_{i,\text{obs}} = \sum a_j F_{j,i}, \quad (1.13)$$

donde  $a_j$  es el coeficiente la  $j$ -ésima SSP en el espectro observado,  $F_{j,i}$  es el flujo a cierta longitud de onda  $i$  observada de la  $j$ -ésima SSP.  $F_{j,i}$  es el resultado de tomar el flujo de la SSP $_j$  original,  $f_{j,i}$ , convolucionada por la dispersión de velocidades,  $G(\sigma_{vel})$ , tomando en cuenta la extinción por el polvo en la longitud de onda  $i$ ,  $A_i$ , y corrida al rojo cierta longitud de onda,  $\lambda$ :

$$F_{j,i} = f_{j,(i'+\lambda)} * G(\sigma_{vel}) 10^{-0.4A_{i'}}, \quad (1.14)$$

donde  $i = i' + \lambda$  y representa el corrimiento al rojo.

En principio, si se cuenta con una librería de SSP infinita, podríamos mostrar todos los posibles brotes de formación estelar y, por lo tanto, recuperar la historia de enriquecimiento químico (chEH por sus siglas en inglés chemical Enrichment History). Desafortunadamente, esta descomposición no es práctica debido a la complejidad computacional.

En la práctica se utilizan librerías finitas de SSPs, tratando de que sean lo más completas y no redundantes. Sin embargo, este tipo de análisis siempre se verá limitado por la librería utilizada (MacArthur et al., 2004; Cid Fernandes et al., 2014). Algunas de las librerías que pueden encontrarse en la literatura son GSD156 (Cid Fernandes et al., 2013), MILES72 (extraída del proyecto MILES, Sánchez-Blázquez et al., 2006), BC138 (creado con el código GISSEL, Bruzual & Charlot, 2003), MAR136 (creado con el código presentado en Maraston, 2005), etc.

La exploración de los parámetros cinemáticos del espectro central utiliza una plantilla con pocas SSPs para reducir el tiempo de cómputo. Se incluyen una SSP joven y de baja metalicidad ( $\sim 90$  Myr y  $Z/Z_\odot = 0.2$ ) y otra SSP vieja y de alta metalicidad ( $\sim 17.8$  Gyr y  $Z/Z_\odot = 1.5$ ).

Una vez sustraído el espectro modelo de la población estelar ajustada, también se ajustan las líneas de emisión más intensas en el rango óptico, desde [O II]  $\lambda 3727$  hasta [S II]  $\lambda 6731$ . Estas líneas se ajustan en grupos de acuerdo con su longitud de onda asumiendo que tienen propiedades cinemáticas similares.

El análisis de la componente estelar puede llevarse a cabo asumiendo que en el espectro de la región todas las poblaciones estelares comparten la misma velocidad con una dispersión de velocidades siguiendo una función Gaussiana. Se sabe que la distribución de velocidades de las poblaciones estelares no siguen una función puramente Gaussiana (Rix & White, 1992). Sin embargo, para realizar el análisis detallado del espectro que usa otro tipo de funciones son necesarios datos con una alta resolución espectral (p ej. Naab et al., 2014).

#### 1.4.2. Teselado de los datos

El siguiente paso sería analizar todo los spaxels para todo el cubo de datos. Desafortunadamente los datos espacialmente resueltos tienden a tener una variación en S/N bastante grande. En general, el espectro central tiene una  $S/N > 50$  para la mayoría de galaxias en CALIFA (Sánchez et al., 2012a). Sin embargo, la  $S/N$  cae rápidamente hacia las regiones externas. Esto puede causar que la señal-a-ruido sea insuficiente para realizar un análisis de poblaciones estelares (Cappellari & Copin, 2003; Cid Fernandes et al., 2014; Sánchez et al., 2016c). Para solucionar este problema se agregan subgrupos de spaxels, los cuales son promediados localmente para incrementar la  $S/N$  de los spaxels a expensas de perder resolución



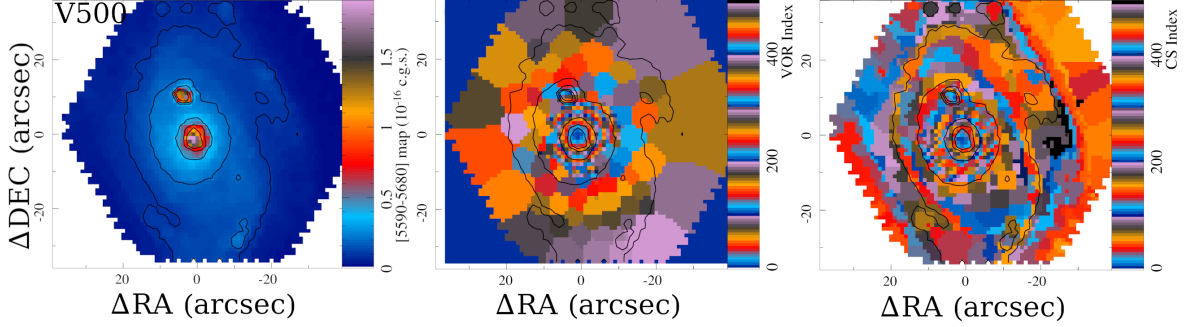


Figura 1.11: Efecto del teselado CS y VORONOI. *Panel Izk*: Mapa de intensidad de una banda angosta resultado de la suma de los flujos en un rango en longitud de onda entre 5590 – 5680Å para NGC 2916 para la configuración V500 de CALIFA. *Panel Med*: Mapa de segmentación obtenido aplicando un teselado de Voronoi. *Panel Der*: Mapa de segmentación obtenido usando el teselado CS. En todos los mapas, los contornos representan el nivel de intensidad iniciando en  $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-1}$  y con un paso de  $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-1}$ . Imagen extraída de [Sánchez et al. \(2016c\)](#).

espacial. A este procedimiento se le conoce como teselado (en inglés *binning*) y a los subgrupos de spaxels se les da el nombre de teselas (en inglés *bins*).

PIPE3D adopta el teselado llamado CS (por sus siglas en inglés *Continuum plus S/N* [Sánchez et al., 2016c](#)). Existen otros métodos de teselado que suelen no preservar la forma del objeto astronómico, al depender de parámetros arbitrarios para cada objeto o combinar componentes estructurales en algunos bins (p. ej. algoritmo *Quadtree* [Samet \(1984\)](#); segmentación isofotal [Papaderos et al. \(2002\)](#); teselado Voronoi ([Cappellari & Copin, 2003](#))). El algoritmo de teselado CS se desarrolló para seguir la distribución de luz del objeto. Este método adopta dos criterios para el teselado: (i) un umbral en S/N en el cual las propiedades de las poblaciones estelares se recuperan con una incertidumbre aceptable ( $\sim 10\%$ ), y (ii) la máxima diferencia de la intensidad del flujo entre spaxels adyacentes. De esta manera se previene que se combinen spaxels de diferentes componentes estructurales de las galaxias (y por lo tanto con diferentes propiedades físicas) en una tesela. En la Figura 1.11 podemos ver un ejemplo de las teselas generadas por el teselado CS y VORONOI aplicado a una banda angosta en un rango de longitud de onda de 5590 – 5680Å. Se puede apreciar la distribución de las teselas es homogénea en el caso del teselado de VORONOI. Mientras que para el caso del teselado de CS se puede apreciar una clara estructura.

### 1.4.3. Análisis de poblaciones estelares

Una vez realizado el teselado espacial de los cubos de datos, se calcula el promedio de todos los espectros por tesela. Los espectros resultantes son acomodados en una tabla de espectros llamada *Row Stacked Spectrum* (RSS).

A cada espectro en el RSS se le realiza un análisis de poblaciones estelares igual al procedimiento descrito anteriormente para la región central. Se derivan los parámetros cinemáticos (la velocidad estelar y la dispersión de velocidades) y la atenuación por polvo con una plantilla de SSPs simple (como la usada para la región central, el resultado puede apreciarse en la Fig. 1.12). Adoptando estos parámetros se realiza una descomposición de las poblaciones estelares usando la biblioteca de SSPs: GSD156, la cual es descrita en detalle en [Cid Fernandes et al. \(2013\)](#). Esta biblioteca está compuesta por SSPs que cubren 39 edades (de 1 Myr hasta 13

Gyr) y 4 metalicidades ( $Z/Z_{\odot} = 0.2, 0.4, 1$  y  $1.5$ ), dando un total de 156 SSPs. Esta biblioteca es una combinación de espectros estelares sintéticos de GRANADA (Martins et al., 2005) para edades  $< 68$  Myr y de la biblioteca de SSPs del proyecto MILES para edades mayores a 68 Myr (Sánchez-Blázquez et al., 2006; Vazdekis et al., 2010).

Una vez obtenido el mejor modelo de poblaciones estelares resultando de la combinación de SSPs, los pesos de dicha descomposición (que indican la fracción de luz correspondiente a cada SSP en la longitud de onda de normalización,  $5500 \text{ \AA}$  dado que es el punto medio de la banda V), se pueden usar para estimar propiedades como edad y metalicidad promedias pesadas en luz y por masa, la historia de formación estelar, el cociente masa-luminosidad, etc.

#### 1.4.4. Análisis de líneas de emisión

A partir del análisis anterior se crea un cubo de datos con los espectros de los modelos de poblaciones estelares para cada spaxel. Con la finalidad de separar la componente estelar de la componente del gas ionizado, se sustrae spaxel por spaxel el espectro del modelo estelar al espectro original. De este modo se obtiene un espectro que solo tiene la contribución del gas ionizado, es decir, que solo presentan líneas de emisión. Sin embargo, este también contendrá los residuos de una sustracción no totalmente perfecta de la componente estelar y el ruido total debido a la sustracción.

La pipeline realiza dos tipos de análisis de las líneas de emisión. Para las líneas mas intensas en el rango óptico se realiza un ajuste con un modelo utilizando una función gaussiana para cada una de ellas. El ajuste gaussiano funciona bastante bien para las líneas de emisión más intensas. Desafortunadamente, para las líneas más débiles, y en general, para todas las líneas de emisión de los cubos de datos, el tiempo de cómputo del ajuste es muy alto. Por lo tanto, para las líneas menos intensas se adoptó un análisis de momentos ponderado por la función gaussiana de la resolución instrumental. La principal ventaja de este método es el tiempo de cómputo es mucho menor.

El análisis de momentos requiere de las líneas de emisión producidas por el gas y del modelo de las poblaciones estelares calculados anteriormente. En resumen, se utilizan la velocidad y dispersión de velocidades de las líneas más intensas,  $H\alpha$ , como punto de partida para calcular estas mismas propiedades de las otras líneas de emisión. Esto da como resultado una estimación de la intensidad, velocidad y dispersión de velocidades de cada línea. Además, se calcula la anchura equivalente (EW por sus siglas en inglés *equivalent width*) de cada línea, dado que se tiene el flujo de la línea emisión y la densidad del flujo del continuo subyacente. Esto se repite 50 veces en un proceso Monte Carlo (MC). Al final, se obtiene una estimación de los todos los parámetros, flujo, velocidad, dispersión de velocidades y EW para cada línea de emisión (ver Fig. 1.12).

El resultado del análisis realizado por PIPE3D son cubos de datos para cada galaxia, los cuales incluyen un mapa por cada propiedad derivada (con su respectivo error) de la población estelar y del gas ionizado. El análisis completo se explica a detalle en Sánchez et al. (2016c, 2018). Es importante remarcar que, además de la información de las líneas de emisión, PIPE3D también nos proporciona información sobre la población estelar: metalicidades y edades (pesadas por luminosidad y por masa), cociente masa-a-luz, y los pesos de la descomposición por cada SSPs. Estos nos permiten derivar la historia de formación estelar.

Los datos procesados por PIPE3D son el punto de partida de los trabajos descritos en los capítulos 2 y 3.

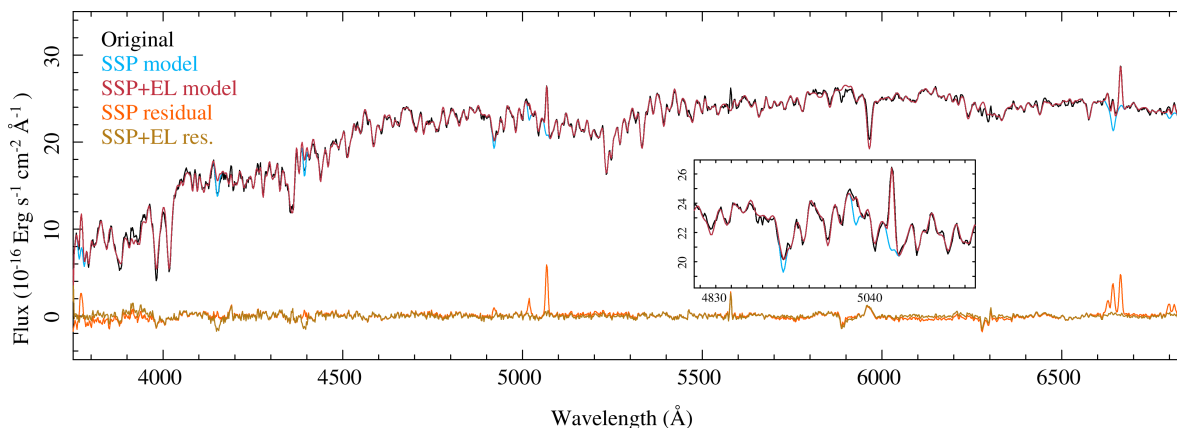


Figura 1.12: Ajuste de la componente estelar y líneas de emisión realizado por PIPE3D para el espectro central de NGC 6754 de los datos de CALIFA en configuración V500. La línea negra representa el espectro original. El mejor modelo de la componente estelar es mostrado en color azul claro. El espectro sin la componente estelar, es decir, puramente con la contribución de las líneas de emisión se muestra como una línea naranja. La línea roja representa el mejor ajuste que incluye la componente estelar y las líneas de emisión. Finalmente, la línea amarilla claro representa el residuo entre el mejor modelo, que incluye la componente estelar y la componente del gas ionizado, y el espectro original. En el recuadro se puede ver en detalle la línea de H $\beta$  y las líneas de O, [O III]  $\lambda$ 4959 y [O III]  $\lambda$ 5007. Extraída de [Sánchez et al. \(2016c\)](#).

## 1.5. Modelos de Fotoionización

Hasta ahora se ha descrito el estudio de las regiones H II desde el enfoque observacional. Sin embargo, como ya se mencionó antes, en algunos casos pueden existir limitaciones con los métodos empíricos (ver Sec. 1.2.3). Otra forma de estudiar el gas ionizado es modelando teóricamente los procesos internos que se producen debido a la fotoionización, es decir, un modelo de fotoionización.

Existen códigos especializados que calculan estos modelos (códigos de fotoionización), entre los que se encuentran: CLOUDY ([Ferland et al., 2013](#)), MAPPINGS ([Sutherland et al., 1993](#)), TITAN ([Dumont et al., 2000](#); [Collin et al., 2004](#)), MOCASSIN ([Ercolano et al., 2003](#)), XSTAR ([Kallman & Bautista, 2001](#)) y otros que no son de acceso público. Actualmente, los dos primeros son los de uso más frecuente.

En general, los códigos de fotoionización resuelven el conjunto de ecuaciones acopladas de (1) transferencia de radiación ionizante, (2) de balance de ionización entre los elementos, y (3) de balance de energía. Estas ecuaciones ya fueron revisadas en las secciones § 1.2.1 y § 1.2.2.

Además de resolver las ecuaciones antes mencionadas, los modelos también deben de tomar en cuenta algunos otros fenómenos que ocurren en un gas ionizado. Entre ellos la inclusión de otros procesos atómicos (transferencia de carga<sup>9</sup> y recombinación dielectrónica<sup>10</sup>, [Kingdon & Ferland \(1995\)](#)), además de ionización por rayos-X y sus efectos (p. ej. [Tarter et al., 1969](#); [Buff & McCray, 1974](#); [Krolik & Kallman, 1984](#)), etc.

<sup>9</sup>Proceso colisional de un ion con otro ion, átomo o molécula donde el primer ion puede ganar un electrón de los segundos.

<sup>10</sup>Proceso en el cual un electrón es capturado por un ion. El exceso de energía del electrón capturado es absorbida por otro electrón presente en el ion tal que este sufra un cambio en su estado energético.

Los datos atómicos (como probabilidades de transición, coeficientes de recombinación, fuerzas de colisión, etc. ver Sec. 1.2) también presentan una relevancia en los modelos de fotoionización. Se han realizado diversas exploraciones para estudiar la variación en propiedades derivadas, como la abundancias de metales, al cambiar los datos atómicos (p. ej. y referencias en [Mendoza, 1983](#); [Butler, 1993](#); [Storey, 1997](#); [Rodríguez & García-Rojas, 2010](#); [Juan de Dios & Rodríguez, 2017](#)). Diversas bases de datos de datos atómicos se han desarrollado y actualizando a través de los años con el fin de mejorar los resultados en diagnósticos espectroscópicos de plasmas astrofísicos (p. ej. [Mendoza et al., 2014, 2021](#); [Del Zanna et al., 2021](#)). También se han desarrollado herramientas para el análisis de las líneas de emisión con diferentes datos atómicos (p. ej. PYNB [Luridiana et al., 2015](#); [Morisset et al., 2020](#)).

### 1.5.1. Parámetros de los modelos de fotoionización

En general, para construir un modelo de fotoionización necesitamos definir (i) la intensidad y distribución espectral del campo de radiación ionizante, (ii) la composición química, (iii) distribución espacial de la densidad de hidrógeno, (iv) si hay polvo y su composición y, por último, (v) la posición del gas con respecto a la fuente ionizante, que se encuentra en el centro en un caso esférico. Como resultado se obtienen (1) las distribuciones espaciales de los distintos iones, así como de la temperatura electrónica,  $T_e$  y (2) las emisiones dadas por los distintos procesos (líneas de emisión de recombinación y de excitación colisional, el continuo libre-libre, la emisión del polvo, entre otros) que puede ser comparadas con las observaciones de un objeto real.

Clásicamente, en una región H II el campo de radiación ionizante se puede originar por una estrella. Sin embargo, en regiones H II gigantes o extragigantes la fuente ionizante es un grupo compacto de estrellas jóvenes. Para calcular el campo de radiación ionizante correspondiente se usan los modelos de síntesis de poblaciones estelares (p. ej. [García Vargas & Díaz, 1994](#); [Leitherer et al., 1999](#); [Bruzual & Charlot, 2003](#); [Leitherer et al., 2014](#); [Mollá et al., 2009](#); [da Silva et al., 2012](#); [Eldridge et al., 2017b](#); [Orozco-Duarte et al., 2022](#)). Estos modelos se componen de SSP (ver Sec. 1.4.1) donde (I) su edad esté dentro de un rango de edades dado por una historia de formación estelar (SFH por sus siglas en inglés), (II) su metalicidad esté dado por una función de distribución de metalicidades que depende con el tiempo, y (III) que tenga la contribución del polvo. Tomando los parámetros adecuados para estos modelos se reproduce el espectro esperado de las poblaciones estelares que ionizan el gas.

La composición química de una nebulosa puede ser bastante compleja. Sin embargo, como se mencionó en la sec. 1.1, el O es el elemento más abundante después del H y He. Por simplicidad, y como primera aproximación, se considera que la abundancia de oxígeno define la metalicidad total de la nebulosa. Para los otros elementos se fijan sus abundancias relativas a valores relativos de los valores solares ([Asplund et al., 2009](#)). Esta aproximación es válida en la vecindad solar, sin embargo, como se discutirá en el Cap. 4 esto puede no ser válido en casos generales.

Además del oxígeno, otros elementos pueden tener relevancia en las nebulosas. Por ejemplo, el nitrógeno, el carbono y elementos refractarios<sup>11</sup> pueden afectar al balance térmico de la nebulosa. El N y C pueden contribuir de manera importante al enfriamiento del gas en abundancias grandes. Por el otro lado, los elementos refractarios pueden estar depletados si hay polvo presente, y por lo tanto, afectar el enfriamiento del gas. Esto puede ser de gran importancia en nebulosas con altas metalicidades ([Henry, 1993](#)).

<sup>11</sup>Elementos con altas temperaturas de condensación,  $> 1400$  K, como el Si, Mg y el Fe, Son los primeros metales en formar moléculas cuando el gas se enfría.

Uno de los parámetros menos explorados es la distribución de densidad. Las observaciones de las regiones H II indican que existen estructuras internas y, por lo tanto, una distribución no homogénea de la densidad (p. ej. [Seaton & Osterbrock, 1957](#); [Osterbrock & Flather, 1959](#)). Esto nos dice que solamente una fracción de la nebulosa está completamente llena de gas ionizado con alta densidad y el resto está casi vacía. Esto define el factor de llenado (*filling factor*,  $\epsilon$ ), el cual nos indica el porcentaje de volumen lleno de gas. En general, en estudios teóricos de regiones H II se toman dos distribuciones de densidad extremas: una esfera homogénea llena de gas y una delgada capa homogénea de gas (p. ej. [Ercolano et al., 2007](#)), con una densidad constante de hidrógeno (típicamente  $n_{\text{H}} = 100 \text{ cm}^{-3}$ ). Otros modelos desarrollados se han centrado en las inhomogeneidades del gas (p. ej. [Cantó et al., 1998](#); [O'Dell et al., 2003](#); [Giammanco et al., 2004](#); [Gahm et al., 2007](#)). En general, estos modelos consideran zonas grumosas dentro de las regiones H II las cuales están ionizadas. Una consecuencia de este tipo de modelos es que las regiones H II pueden albergar una cantidad considerable de gas neutro. Lo que significaría que una fracción de fotones ionizantes escapa de la nebulosa sin ionizar el gas (el cual se le conoce como goteo de fotones, *leaking* en inglés).

Otro de los parámetros que influye en un modelo es el polvo. El principal efecto del polvo es la absorción de fotones ([Mathis, 1986](#)) y su contribución al calentamiento y enfriamiento del gas (ver Sec. 1.2.2, [Baldwin et al., 1991](#)). El modelado a detalle del polvo es complicado, sin embargo debe tenerse en cuenta dado su papel en el balance térmico.

Finalmente, la posición de la estrella con respecto al gas ionizado tiene un efecto en las propiedades observadas del mismo. Este problema puede abordarse de diferentes maneras. La aproximación más simple es considerar una simetría esférica donde la fuente ionizante se localiza en el centro de una esfera de Strömgren (una esfera cuyo radio es igual a  $R_S$ , ver ec. 1.4). Una de las formas de considerar diversas fuentes de ionización (varias estrellas jóvenes) es modelar la región H II como la suma de varias esferas de Strömgren de gas ionizado correspondiente a cada fuente ionizante (p. ej. [Peña, 1986](#)). A pesar de que observacionalmente es complicado obtener información sobre la posición de la fuente ionizante con respecto al gas ionizado, hay algunos objetos donde la posición de las estrellas ionizantes son conocidas.

### Parámetro de ionización

El espacio de parámetros de entrada de un modelo de fotoionización puede ser bastante complejo y presentar degeneraciones con las líneas de emisión y cocientes de estas (p. ej. como la abundancia de O, como se muestra en la Fig. 9 y Fig. 6 de [Kewley & Ellison, 2008](#); [Bresolin, 2011](#), respectivamente). Uno de los parámetros que nos puede ser de utilidad para definir distintas propiedades de un gas ionizado (y reducir el espacio de parámetros) es el parámetro de ionización,  $U$ .

El parámetro de ionización caracteriza el nivel de ionización de una región H II, el cual está definido como el cociente adimensional:

$$U \equiv \frac{Q(\text{H}^0)}{4\pi r^2 n_{\text{H}} c}, \quad (1.15)$$

donde  $r$  es la distancia entre la fuente ionizante y la cara iluminada del gas,  $n_{\text{H}}$  es la densidad del hidrógeno,  $Q(\text{H}^0)$  es el número de fotones ionizantes emitidos por la fuente ionizante por segundo, y  $c$  es la velocidad de la luz.  $U$  nos permite reducir el espacio de parámetros de nuestro modelo de fotoionización dado que incluye tres propiedades usadas para caracterizar un gas ionizado mencionados en la sección anterior: la densidad del gas, la distancia a la estrella ionizante y la cantidad de fotones ionizantes.

Por lo tanto, una región puede ser modelada en principio definiendo tres parámetros: el parámetro de ionización, la metalicidad del gas (o solo la abundancia de O) y la temperatura efectiva de la fuente ionizante. Se retomarán los parámetros de entrada de un modelo de fotoionización en el Capítulo 5.

### 1.5.2. Modelos de fotoionización para el estudio de regiones HII

Los modelos de fotoionización se han usado ampliamente para entender las condiciones físicas internas y como afecta al proceso de ionización y emisión de líneas. Describimos a continuación algunos de sus usos encontrados en la literatura. (1) Han sido de gran ayuda para derivar la composición química de las nebulosas dado que los modelos nos proporcionan la distribución de la temperatura electrónica en el gas ionizado (p. ej. [Kewley & Ellison, 2008](#); [Pérez-Montero, 2014](#); [Blanc et al., 2015](#); [Vale Asari et al., 2016](#); [Thomas et al., 2018](#)). (2) También han sido ampliamente usados para el cálculo de ICFs (p. ej. [Vilchez, 1989](#); [Amayo et al., 2020](#)). (3) Pueden proporcionarnos información sobre la población estelar ionizante por medio de la mitigación del campo de radiación ionizante. Esto se puede trazar por el parámetro de mitigación de la radiación,  $\eta$  (*radiation softness parameter* en inglés, [Vilchez & Pagel, 1988](#)), el cual es calibrado con modelos de fotoionización ([Bresolin et al., 1999](#); [Gruenwald et al., 2002](#)). Todos estos trabajos están basados en redes de cientos o miles de modelos de fotoionización ya sea combinándolos con datos observacionales o no. En general, este tipo de estudios nos permiten estudiar los parámetros que afectan a las intensidades de las líneas emitidas por el gas.

Algunos trabajos están basados en el modelado y estudio de regiones H II individuales, ya sea con un apuntado o varios apuntados en diferentes zonas de la región ionizada (p. ej. [García-Vargas et al., 1997](#); [Luridiana et al., 1999](#); [Relaño et al., 2002](#); [Jamet et al., 2005](#)). Este tipo de trabajos se basan en obtener información observacional de alta calidad que nos permita delimitar el modelo y poder inferir las propiedades del gas ionizado correctamente.

Sin embargo, hay que tener en cuenta que todas las suposiciones hechas sobre los parámetros, los datos atómicos, propiedades estelares, geometría, etc, pueden dar como resultado un modelo que pueda distar de la realidad. Los cocientes de líneas de emisión predichos por los modelos pueden no coincidir con los datos observacionales. Una de las discrepancias más importantes es con el cociente  $[\text{O III}] \lambda\lambda 4363/5007$  que es subestimado comparado con las observaciones. Además, este cociente está relacionado directamente con la temperatura del gas ionizado (p. ej. [Stasińska & Schaerer, 1999](#)). Esto produce que los modelos no reproduzcan la temperatura electrónica de forma correcta en las regiones H II. Por lo tanto, hay una fracción de energía térmica no despreciable que falta en los modelos.

Esta discrepancia ha sido encontrada y explorada no solamente en regiones H II. También puede se encuentra en modelos de nebulosas planetarias, en AGN y galaxias Seyfert (p. ej. [Stasińska, 1984](#); [Binette et al., 1996](#); [Simpson et al., 1996](#); [Peña et al., 1998](#)). Se han propuesto diferentes fuentes de calentamiento para explicar la discrepancia en las temperaturas electrónica. Por ejemplo puede existir calentamiento por choques (p. ej. [Dufour & Buckalew, 1999](#)), calentamiento por mezclado con gas de más alta temperatura proveniente de vientos estelares (p. ej. [Weaver et al., 1977](#)), variaciones de la temperatura dentro del gas ionizado por lo que la temperatura derivada por  $[\text{O III}] \lambda\lambda 4363/5007$  no es representativa (p. ej. [Peimbert, 1967](#); [Mathis, 1995](#)). Este sigue siendo un problema abierto y de suma importancia dado que impacta directamente en nuestro entendimiento sobre el balance de energía de los gases fotoionizados y en las propiedades de las nebulosas, (como la abundancia, [Kewley & Ellison, 2008](#); [Morisset et al., 2016](#)), derivado de las líneas de emisión.



## 1.6. Propiedades globales de las regiones H II en galaxias

Como se ha descrito en las secciones anteriores, el estudio de regiones H II extragalácticas nos provee de la información para el estudio de diversos problemas astrofísicos, por ejemplo entender como es la formación de estrellas masivas, el estudio de los perfiles de abundancias observadas en los discos de las galaxias espirales, y la derivación de relaciones de propiedades globales. Las regiones H II se encuentran principalmente en galaxias de tipo tardío donde hay formación estelar (p. ej. [Ho et al., 1997](#)) aunque pueden encontrarse en menor proporción en galaxias tipo elípticas. Uno de las herramientas más importantes para el estudio de las propiedades de formación estelar en galaxias es la función de luminosidad de regiones H II.

En astronomía, una función de luminosidad nos proporciona el número de estrellas, o galaxias o regiones H II, por intervalo de luminosidad. La luminosidad de H $\alpha$  de una región H II está directamente relacionada con la cantidad de radiación ionizante, por lo que la función de luminosidad de las regiones H II nos permite constreñir la función de masa de las regiones estelares jóvenes ([Santoro et al., 2022](#)). Típicamente, la luminosidad de las regiones H II extragalácticas está entre  $\log(L_{\text{H}\alpha} [\text{erg s}^{-1}]) \sim 35 - 40$  (p. ej. [Kennicutt et al., 1989](#); [Bradley et al., 2006](#); [Santoro et al., 2022](#)). Nebulosas con una luminosidad menor a  $10^{37} \text{ erg s}^{-1}$  comúnmente son ionizadas por una sola estrella tipo OB, mientras que aquellas regiones con una luminosidad en el rango de  $10^{37} - 10^{39} \text{ erg s}^{-1}$  son ionizadas por poblaciones estelares que contienen múltiples estrellas tipo OB. Las regiones con una luminosidad mayor a  $10^{39} \text{ erg s}^{-1}$  se asocian a regiones gigantes o supergigantes que típicamente se encuentran en galaxias tardías (p. ej. [Kennicutt et al., 1989](#); [Elmegreen et al., 1996](#)).

La distribución de tamaños puede ser descrita convencionalmente como una relación exponencial (p. ej. [van den Bergh, 1981](#); [Hodge, 1987](#)), aunque una ley de potencias también puede ser usada (p. ej. [Pleuss et al., 2000](#)). Esta última está relacionada físicamente con la función de luminosidad de regiones H II (p. ej. [Oey et al., 2003](#)). El diámetro de las regiones H II extragalácticas pueden estar en el rango de  $\log(d [\text{pc}]) \sim 1.5 - 3.0$ . Aunque, el tamaño de las regiones H II depende de muchos factores como ha sido remarcado por [Barnes et al. \(2021\)](#).

Las regiones H II nos proporcionan la composición química actual del medio interestelar en las galaxias (p. ej. [Sánchez et al., 2014](#)). Los metales en estas nebulosas son el resultado de la historia de enriquecimiento químico en el lugar donde la nebulosa se ha formado, es decir, donde se han formado las estrellas que ionizan al gas. Por lo tanto, las regiones H II son trazadores de procesos de formación estelar y de la composición química en las galaxias. Esto hace que su análisis sea fundamental para el estudio de la evolución de las galaxias.

Uno de los escenarios más aceptados de formación galáctica es el crecimiento de dentro hacia afuera (p. ej. [Matteucci & Franco, 1989](#); [Molla et al., 1996](#)). Este modelo nos dice que en el proceso de acrecimiento, el gas se concentra en las partes internas de las galaxias, alcanzando altas densidades, dando a lugar a un proceso de formación estelar violento y eficiente, en etapas tempranas de la evolución. Por lo tanto, el gas se reprocesa rápidamente en las zonas centrales de las galaxias. Esto provoca que en la actualidad haya poblaciones viejas y metalmente ricas rodeadas de un entorno de gas de alta metalicidad. Mientras tanto, las zonas externas están conformadas por estrellas jóvenes y metálicamente pobres que se forma a partir de un gas con baja metalicidad. Diversos autores han encontrado evidencia que apoya este escenario usando perfiles de colores estelares (p. ej. [Taylor et al., 2005](#); [Muñoz-Mateos et al., 2007](#)). El estudio de la composición química del gas nos provee de restricciones para este tipo de modelos (p. ej. [Molla et al., 1996](#); [Chiappini et al., 2001](#)). Por otra parte, diversos estudios observacionales han encontrado un gradiente negativo de la abundancia de

oxígeno del gas ionizado a lo largo de las galaxias cercanas (p. ej. Vila-Costas & Edmunds, 1992; Moustakas et al., 2010; Bresolin et al., 2009). Este perfil negativo es una consecuencia directa del escenario dentro-fuera de evolución galáctica.

Por otra parte, las propiedades físicas de las regiones H II también dependen de las propiedades globales de las galaxias (como su masa estelar o morfología). Los gradientes de metalicidad dependen del tipo de galaxia, masa y entorno donde se encuentre la nebulosa y del ambiente donde se encuentre de la galaxia (p. ej. Sánchez et al., 2015b; Belfiore et al., 2017; Espinosa-Ponce et al., 2020; Sánchez, 2020; Franchetto et al., 2021). En las galaxias más masivas y de tipo temprano se tiene una fracción alta de poblaciones viejas mientras que en las galaxias menos masivas y de tipo tardío predominan las poblaciones jóvenes. Esto es una consecuencia de la evolución de las galaxias, ya que las galaxias de tipo tardío y más masivas han formado estrellas antes (Thomas et al., 2010; González Delgado et al., 2014). Dado que las regiones H II son trazadoras de los procesos de formación estelar y de la metalicidad actual, se espera que exista una relación entre la morfología y la masa de las galaxias y de las propiedades de estas nebulosas.

Por lo tanto, para poder explorar la dependencia de las propiedades de las regiones H II y el ambiente en donde se forman se necesita de un catálogo de datos espacialmente resueltos. Esto nos permitirá poder estudiar como cambian las nebulosas y sus propiedades a lo largo de toda la galaxia. Por otra parte, las muestras deben incluir regiones ionizadas hospedadas en diferentes tipos morfológicos de galaxias y de diferentes masas estelares. Como ya se ha mencionado antes, la evolución de las galaxias, y por lo tanto del ambiente donde se forman las nebulosas, dependerá de cada tipo morfológico y masa de las galaxias.

La relación entre las propiedades físicas de las nebulosas, abundancia de oxígeno, cociente de N/O y parámetro de ionización, y las líneas de emisión y cociente de líneas no se entiende completamente en la actualidad. Parte del problema reside en los métodos utilizados para derivar las propiedades físicas. Por ejemplo, para la determinación de abundancia de oxígeno se tienen distintos métodos dependiendo de los datos observados. Cuando se cuenta con observaciones de líneas aurales podemos usar el método directo (p. ej. Peimbert et al., 2017). En caso contrario, podemos usar los calibradores de líneas fuertes (p. ej. Kewley & Dopita, 2002; Pettini & Pagel, 2004; Marino et al., 2013; Pérez-Montero, 2014; Pilyugin & Grebel, 2016; Vale Asari et al., 2016; Ho, 2019). Pero, se sabe que hay una discrepancia (que puede ir desde  $\sim 0.025$  dex a  $\sim 0.3$  dex, p. ej. Yin et al., 2007) entre los dos tipos de abundancias derivadas por causas desconocidas (p. ej. Kewley & Ellison, 2008; Peimbert et al., 2017).

Por otra parte, los calibradores empíricos de líneas fuertes también presentan discrepancias entre ellos mismos. Al compararse los distintos calibradores desarrollados por diversos autores (p. ej. Marino et al., 2013; Pérez-Montero, 2014; Blanc et al., 2015; Pilyugin & Grebel, 2016; Thomas et al., 2018; Ho, 2019; Curti et al., 2020) se ha encontrado que los resultados derivados de cada uno difieren entre sí (p. ej. Kewley & Ellison, 2008; Zurita et al., 2021; Espinosa-Ponce et al., 2022). Esto hace que la elección del calibrador de metalicidad sea importante ya que los resultados derivados dependerán a partir del calibrador (p. ej. relación masa-metalicidad, perfiles radiales, etc.).

El contar con un catálogo extenso, estadísticamente completo y de buena calidad (p. ej. buena resolución espectral, alta S/N en las líneas de emisión necesarias, como  $[\text{O III}]\lambda 3727$  ó  $\text{N II}\lambda 6548$ ) nos garantiza una mejor interpretación de los resultados derivados. Aún así, una muestra con estas características puede ser insuficiente para determinar las causas de las discrepancias en las abundancias y/o metalicidades calculadas (Kewley & Ellison, 2008; Espinosa-Ponce et al., 2022). Por tanto, el uso de modelos de fotoionización es necesario para entender la física del gas ionizado. Una exploración en los parámetros que definen un modelo



de fotoionización (como abundancias, fuente ionizante, distribución de densidad y parámetro de ionización) puede darnos pistas sobre la discrepancia de resultados en las propiedades derivadas, principalmente la metalicidad.

## 1.7. Objetivo del trabajo

El objetivo de este trabajo es el realizar una exploración de las propiedades físicas (abundancia de O, N/O, parámetro de ionización, densidad electrónica, extinción por polvo y  $EW(H\alpha)$ ) de las regiones H II del universo cercano (con un redshift entre 0.005 y 0.03) usando datos observacionales espacialmente resueltos ( $\sim 800$  pc) y modelos de fotoionización. Para esto se generará un catálogo nuevo de regiones H II a partir de las observaciones del catastro de galaxias con datos espacialmente resueltos del proyecto CALIFA. Hasta la fecha, CALIFA tiene una base de datos de 924 galaxias. La muestra abarca muchos tipos de galaxias: galaxias interactuantes, de baja y alta masa estelar, galaxias tempranas y tardías, galaxias donde se han presentado CCSNe, etc (Husemann et al., 2013; García-Benito et al., 2015; Sánchez et al., 2016a; Galbany et al., 2018). Esto la hace una muestra ideal para el estudio de regiones H II bajo diferentes ambientes de formación

Con el catálogo de regiones H II generado se estimarán sus distintas propiedades químicas usando calibradores de líneas fuertes. Esto debido a que las líneas necesarias para usar el método directo tienen una baja SNR en la mayoría de las regiones, por lo tanto los resultados no son fiables. Una vez obtenidas sus propiedades, se compararan con las propiedades estelares y las distintas propiedades globales de las galaxias, masa y tipo morfológico.

Por último, como ya se ha mencionado, los modelos de fotoionización nos permiten explorar las condiciones físicas internas y como afectan las mismas al proceso de ionización y emisión de líneas. Sin embargo, es bien sabido que existe una discrepancia entre los resultados observacionales y teóricos del cálculo de abundancias, principalmente oxígeno (Kewley & Ellison, 2008; Peimbert et al., 2017; Stasińska, 2009). Esta diferencia en los resultados puede deberse a distintas causas, ya sea falta de alguna fuente ionizante, como choques o mezcla con gas ionizado por vientos estelares (Dufour & Buckalew, 1999; Weaver et al., 1977), variaciones en la temperatura del gas (Peimbert, 1967), etc. Para explorar esta discrepancia, se realizará una exploración de la dependencia entre las propiedades físicas,  $12 + \log(O/H)$ ,  $\log(N/O)$  y  $\log(U)$ , y las líneas de emisión de las regiones H II por medio de modelos de fotoionización. Se buscarán los modelos de fotoionización que mejor empaten con los datos observacionales. De esta manera, podremos entender la causa de la discrepancia entre los resultados derivados a partir de modelos teóricos y de las observaciones de las propiedades físicas,  $12 + \log(O/H)$ ,  $\log(N/O)$ , y  $\log(U)$ .

## Capítulo 2

# Catálogo de regiones HII

Durante muchos años, las muestras de datos espectroscópicos de regiones H II han estado limitadas al tener pocas observaciones de estos objetos (p ej. [Peimbert et al., 2005](#); [Valerdi et al., 2021](#)). En general, las muestras de regiones H II son pequeñas pero con espectros de buena calidad. Por tanto, los estudios que usan estas muestras se limitan a nebulosas en cierto tipo de galaxias y con propiedades específicas. Recientemente con los grandes muestreos de galaxias y el desarrollo de espectroscopía de campo integral, las muestras de regiones H II con datos espectroscópicos disponibles crecieron en número pudiendo estudiar a detalle cómo sus propiedades cambian a lo largo de las galaxias y como varían sus propiedades en diferentes tipos de galaxia (p ej. [Marino et al., 2013](#); [Sánchez et al., 2015b](#); [Sánchez-Menguiano et al., 2018](#); [Rousseau-Nepton et al., 2018](#)).

La muestra de galaxias de CALIFA es una de las más completas exploradas con la técnica de espectroscopía de campo integral (Sec. 1.3.2) hasta el momento. Por otro lado, su rango espectral permite estudiar las poblaciones estelares así como las principales propiedades del gas ionizado en el óptico. Además, los datos son espacialmente resueltos lo que nos permite estudiar galaxias en toda su extensión. Un primer catálogo de regiones H II fue publicado en [Sánchez et al. \(2015b\)](#) con cerca de 5000 objetos correspondientes a  $\sim 300$  galaxias.

En este trabajo se presenta el catálogo actualizado de regiones H II para la muestra de CALIFA hasta este momento. Se usó una forma novedosa de seleccionar regiones H II. Nuestra selección se basa solamente en las propiedades que definen una región H II: (I) una estructura grumosa con un alta emisión de  $H\alpha$ , y (II) una población estelar subyacente conformada por estrellas jóvenes. El resultado del proceso de selección usando estos dos supuestos son regiones que son compatibles con la definición de una región H II. Además, la selección es independiente a las líneas de emisión, las cuales están ligadas a las propiedades mismas del gas ionizado (ver Sección 1.2).

Nuestro análisis comprende una corrección por gas difuso ionizado (DIG por sus siglas en inglés: *diffuse ionized gas*). El DIG no se comprende del todo. Se han propuesto cuales son sus posibles fuentes de ionización y cómo se puede identificar (p. ej. [Lacerda et al., 2018](#)). Este gas, al estar presente a lo largo de toda la galaxia, contamina las mediciones de diversos objetos, en especial, las de las regiones H II. En este trabajo agregamos un enfoque para una corrección de la contribución del DIG.

En el estudio se mostró que la distribución de regiones H II seleccionadas con nuestro método se localiza en la zona de formación estelar de los principales diagramas de diagnóstico. Además, se encontró que la posición en estos diagramas presenta una fuerte correlación con el valor de la anchura equivalente de  $H\alpha$  ( $EW(H\alpha)$ ) (resultado que ha sido encontrado antes por otros autores, p. ej. [Sánchez et al., 2015a](#); [Lacerda et al., 2018](#), usando datos de la mues-

tra de galaxias de CALIFA). También se encontró una dependencia entre la posición en los diagramas de diagnóstico y el tipo morfológico de las galaxias y la distancia galactocéntrica. Este resultado está relacionado con la manera en que las galaxias se forman y con la historia de formación estelar, lo que da a lugar a que existan gradientes de metalicidad. Finalmente, se estudió la dependencia entre los dos cocientes de línea del diagrama de diagnóstico más usado,  $[\text{N II}]\lambda 6583/\text{H}\alpha$  y  $[\text{O III}]\lambda 5007/\text{H}\beta$ , y las propiedades físicas de las poblaciones estelares subyacentes, metalicidad y edad. El cociente de la línea de N presenta una mayor correlación con las propiedades estelares que el cociente de la línea de O. Por otra parte, como parte de la exploración de las propiedades del DIG, se encontró que la distribución en el diagrama BPT del gas difuso presenta una correlación con la morfología de la galaxia y la distancia galactocéntrica.

La presentación del catálogo, así como la descripción del método de segregación e identificación de regiones HII fue publicado en un artículo en la revista *Monthly Notices of the Royal Astronomical Society* (MNRAS), el cual se incluye a continuación. La versión electrónica de esta publicación puede ser consultada en la página <https://ui.adsabs.harvard.edu/abs/2020MNRAS.494.1622E>. Igualmente los datos son de acceso público y pueden ser encontrados en [http://ifs.astroscu.unam.mx/CALIFA/HII\\_regions/](http://ifs.astroscu.unam.mx/CALIFA/HII_regions/). El código PYHIIEXPLORER puede ser encontrado en <https://github.com/cespinosa/pyHIIexplorerV2>.

# H II regions in the CALIFA survey: I. catalogue presentation

C. Espinosa-Ponce<sup>1</sup>,<sup>1</sup>★ S. F. Sánchez,<sup>1</sup> C. Morisset,<sup>2</sup> J. K. Barrera-Ballesteros,<sup>1</sup>  
L. Galbany<sup>3</sup>,<sup>3</sup> R. García-Benito<sup>4</sup>,<sup>4</sup> E. A. D. Lacerda<sup>1</sup> and D. Mast<sup>5,6</sup>

<sup>1</sup>Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 70264, 04510 Mexico City, Mexico

<sup>2</sup>Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 106, 22800 Ensenada, B. C., Mexico

<sup>3</sup>Departamento de Física Teórica y del Cosmos, Universidad de Granada, E-18071 Granada, Spain

<sup>4</sup>Instituto de Astrofísica de Andalucía, CSIC, Apartado de correos 3004, E-18080 Granada, Spain

<sup>5</sup>Universidad Nacional Córdoba, Observatorio Astronómico de Córdoba, Laprida 854, (X5000BGQ) Córdoba, Argentina

<sup>6</sup>Consejo de Investigaciones Científicas y Técnicas de la República Argentina, Avda. Rivadavia 1917, C1033AAJ CABA, Argentina

Accepted 2020 March 12. Received 2020 March 12; in original form 2019 November 20

## ABSTRACT

We present a new catalogue of H II regions based on the integral field spectroscopy (IFS) data of the extended CALIFA and PISCO samples. The selection of H II regions was based on two assumptions: a clumpy structure with high contrast of H $\alpha$  emission and an underlying stellar population comprising young stars. The catalogue provides the spectroscopic information of 26 408 individual regions corresponding to 924 galaxies, including the flux intensities and equivalent widths of 51 emission lines covering the wavelength range between 3745 and 7200 Å. To our knowledge, this is the largest catalogue of spectroscopic properties of H II regions. We explore a new approach to decontaminate the emission lines from diffuse ionized gas contribution. This diffuse gas correction was estimated to correct every emission line within the considered spectral range. With the catalogue of H II regions corrected, new demarcation lines are proposed for the classical diagnostic diagrams. Finally, we study the properties of the underlying stellar populations of the H II regions. It was found that there is a direct relationship between the ionization conditions on the nebulae and the properties of stellar populations besides the physical condition on the ionized regions.

**Key words:** techniques: spectroscopic – ISM: general – H II regions – galaxies: ISM – galaxies: star formation.

## 1 INTRODUCTION

Classical H II regions are clouds of ionized gas surrounding young massive stars in which the star formation has recently taken place (<15 Myr). The ultraviolet photons that ionize the surrounding medium are produced in large amounts by OB stars that were formed in the H II regions. Since OB stars only live  $\sim 10$  Myr, the H II regions trace recent star formation processes. These regions possess a broad range of physical size (diameter), from a few parsecs, such as the Orion nebula with  $D \lesssim 8$  pc (Anderson 2014), to hundreds of parsecs, like 30 Doradus with  $D \sim 20$  pc, NGC 604 with  $D \sim 460$  pc, or NGC 5471 with  $D \sim 1$  kpc (Oey et al. 2003; García-Benito et al. 2011). The smaller regions are ionized by a single star or a handful of them, while the largest ones are ionized by a young cluster of thousands of them. Those larger star-forming (SF) complexes are the precursors of the extragalactic giant H II regions found commonly in the discs of spiral galaxies (e.g. Hodge & Kennicutt 1983; Dottori 1987; Dottori & Copetti 1989; Knapen

1998), or in starbursts and blue compact galaxies (Kehrig et al. 2008; López-Sánchez & Esteban 2009; Cairós et al. 2012).

In addition to H II regions, there are other emission-line objects with different main excitation mechanisms: planetary nebulae and objects photoionized by hard radiation fields such as active galactic nucleus (AGN), shock excitation, or even post-AGB stars (e.g. Veilleux & Osterbrock 1987; Binette et al. 1994; Stasińska & Izotov 2003; Binette et al. 2009; Morisset & Georgiev 2009; Flores-Fajardo et al. 2011; Papaderos et al. 2013; Singh et al. 2013). All these objects, including H II regions, could be classified, in principle, based on their emission line ratios. The first diagnostic diagram using the emission line ratios [N II] $\lambda 6583$ /H $\alpha$  versus [O III] $\lambda 5007$ /H $\beta$  was proposed by Baldwin, Phillips & Terlevich (1981, hereafter BPT diagram). Osterbrock (1989) and Veilleux & Osterbrock (1987) proposed a demarcation curve between AGN and SF galaxies and two new diagnostic diagrams.<sup>1</sup> Theoretical

<sup>1</sup>The two more commonly used diagnostic diagrams are [S II] $\lambda\lambda 6717, 6761$ /H $\alpha$  versus [O III] $\lambda 5007$ /H $\beta$  and [O I] $\lambda 6300$ /H $\alpha$  versus [O III] $\lambda 5007$ /H $\beta$ .

★ E-mail: cespino@astro.unam.mx

photoionization models have been used to infer the demarcation line between the different ionization sources.

Dopita et al. (2000) and Kewley et al. (2001) proposed one of the first classification schemes to identify pure galaxies hosting star formation and AGNs, using a combination of stellar population synthesis and photoionization models. Their models assume that the emission-line spectrum depends directly on: (i) the chemical abundances of heavy elements (oxygen being the most important) of the gas phase within an H II region; (ii) the shape of ionizing radiation spectrum; (iii) the electron density; and (iv) the geometrical distribution of gas with respect to the ionizing sources. All geometrical factors are included in the so-called ionization parameter  $q$  (defined as the ratio of ionizing photon density to hydrogen density) or its dimensionless counterpart,  $U$ .<sup>2</sup> They assume a priori that these parameters are independent one from each other and, hence, the models are presented as grids of the four of them. On the other hand, Kauffmann et al. (2003) use the Sloan Digital Sky Survey (SDSS; York et al. 2000) data to propose a different demarcation line to separate ionization by SF and AGNs, based purely on observational properties.

H II regions allow the present chemical composition of the interstellar medium (ISM) to be determined. They have been used for a long time to measure the gas-phase abundance directly at discrete locations in galaxies (e.g. Diaz 1989; Rosales-Ortega et al. 2011; Marino et al. 2013; Sánchez-Menguiano et al. 2016). These measurements provide the necessary information to understand the distribution of chemical abundances across the optical extension of galaxies, to constrain different theories of galactic chemical evolution, to obtain information about the stellar nucleosynthesis in SF galaxies, and to derive the star formation histories of galaxies.

Classical studies of the chemical abundance of extragalactic nebulae are based on the observations of nebular emissions through long-slit or single-aperture spectrographs. Unfortunately, these techniques provided a limited sample of galaxies in detail, including a low number of H II regions per galaxy, in general. Therefore, there was a limited coverage of these regions along the whole galaxy extension (e.g. Bresolin 2019). Conversely, the classical slit-spectroscopy surveys introduce a bias by aperture effects: long-slit observations do not integrate over the full extension of an H II region. Instead, only the brightest region of each nebula is observed, i.e. the areas of higher excitation. Hence, the derived spectra may not be representative of the whole ionization conditions (e.g. Dopita et al. 2014).

Moreover, traditional surveys could present a biased sample of H II regions. Most of these spectroscopic surveys are biased towards high-contrast (thus, equivalent width of emission lines) nebulae located in the outer disc of galaxies. Since late-type galaxies offer the best contrast with their lower surface brightness in the continuum, these types of galaxies were mainly observed by these surveys. Thus, the final samples and conclusions could be biased to the properties of H II regions in a particular type of galaxy. Moreover, in most of the cases, the studies were restricted to regions ionized by younger and/or massive stellar clusters. In general, the central regions of galaxies were excluded, although it is known that they distinguish themselves spectroscopically from those in the disc: inner SF regions present stronger low-ionization forbidden emission

and may populate the right branch of the BPT diagram (Kennicutt, Keel & Blaha 1989; Ho, Filippenko & Sargent 1997; Sánchez et al. 2012b).

Despite these caveats, these studies provide important relations, clear patterns, and scaling laws, e.g. characteristics versus integrated abundances: Moustakas & Kennicutt 2006; effective yield versus luminosity and circular velocity relations (e.g. Garnett 2002); abundance gradients and the effective radius of discs (e.g. Diaz 1989); possible differences in the gas-phase abundance gradients between normal and barred spirals (Martin & Roy 1994; Zaritsky, Kennicutt & Huchra 1994); surface brightness versus metallicity relations, luminosity–metallicity, and mass–metallicity (Lequeux et al. 1979; Skillman, Kennicutt & Hodge 1989; Vila-Costas & Edmunds 1992; Zaritsky et al. 1994; Tremonti et al. 2004), to cite just a few.

In the last years, new observational techniques, such as integral field spectroscopy (IFS) with a large field of view (FoV) and multiobject spectroscopy, allow us to observe hundreds of H II regions per galaxy and a full two-dimensional (2D) coverage of the disc of nearby ones (e.g. Rosales-Ortega et al. 2010; Sánchez et al. 2012a, 2014). This new generation of emission-line surveys provided large catalogues with thousands of objects over an unbiased sample of galaxies, from early- to late-type galaxies (Marino et al. 2016; Sánchez-Menguiano et al. 2016). Moreover, the sampled H II regions are at any galactocentric distances (even at more than  $2 R_e$ ),<sup>3</sup> which allows us to study a wider range of H II regions.

The Calar Alto Legacy Integral Field Area survey (CALIFA; Sánchez et al. 2012a) was one of the pioneering in the massive exploration of H II regions over complete samples of galaxies (e.g. Sánchez et al. 2014). CALIFA provides IFS data for a well-defined and statistically significant sample of galaxies on the nearby Universe ( $z < 0.08$ ). The observations cover the full optical extent (up to  $3-4R_e$ ) of the galaxies of any morphological type and with a spectral range 3745–7200 Å. CALIFA still provides the best physical resolution ( $\sim 0.8$  kpc; Sanchez 2019) compared to other IFS galaxy surveys (e.g. Croom et al. 2012; Bundy et al. 2015) due to its limited redshift range and lower average redshift. However, it is overpassed by IFS observations with MUSE (e.g. López-Coba et al. submitted); none the less, this instrument does not cover wavelengths bluer than 4650 Å. In Sánchez et al. (2012a), the first H II region catalogue derived from the observations of the pilot study of the CALIFA survey was presented (p-CALIFA; Mármol-Queraltó et al. 2011). Besides that, several studies were performed with a larger statistical sample of H II regions based on later catalogue updates: oxygen abundance profiles on face-on spiral galaxies (Marino et al. 2016; Sánchez-Menguiano et al. 2016), global and local  $M-Z$  relations and the dependence on the star formation rate (SFR) and the specific star formation rate (sSFR; Sánchez et al. 2013, 2017), etc. However, there was no public update of the CALIFA H II region catalogue after Sánchez et al. (2012a).

In this study, we present the updated CALIFA catalogue of H II regions for  $\sim 1000$  nearby galaxies observed so far within the framework of this survey (2019 August). We propose new criteria to select H II regions within each galaxy and a new methodology to extract their main spectroscopic information. We explore the distribution of H II regions along the classic diagnostic diagrams. For the first time we do not use those diagrams to classify H II regions. Instead, we present an unbiased exploration of the location

<sup>2</sup>The ionization parameter is defined as  $U = S/(4\pi r^2 n_H c)$ , where  $S$  is the number of ionization photons emitted by a source per unit of time,  $r$  is the distance between the source and the nebula,  $n_H$  is the H density, and  $c$  is the speed of light.

<sup>3</sup>The effective radius,  $R_e$ , of a galaxy is defined as the radius at which half of the total light of the galaxy is emitted.



of these objects across the diagnostic diagrams. Moreover, we propose a new approach to correct the contamination by diffuse ionized gas (DIG). With the final catalogue of H II regions corrected by diffuse gas, we explore how this correction changes the distribution along those diagnostic diagrams. Finally, we show how H II regions distribute among galaxies of different morphologies and at different galactocentric distances, revising patterns with the properties of the underlying stellar population already covered in previous explorations. In addition, we deliver the full catalogue of H II regions,<sup>4</sup> including all the spectroscopic properties extracted by our code. The code itself is also delivered for its use by the community.<sup>5</sup>

The layout of this article is as follows: In Section 2 we present the galaxy sample and data analysed in this study; the performed analysis is described in Section 3, including a summary of the procedure adopted to extract the information of the emission of stellar population and the emission lines (Section 3.1), and a detailed description of the procedure to detect and extract that information for the H II regions within each galaxy (Section 3.2). The final criteria adopted to select the H II regions are described in Section 3.3, and the proposed procedure to de-contaminate from the DIG emission is described in Section 3.4. The main results of our analysis are presented in Section 4, including (i) the main effects of the DIG decontamination (Section 4.1); (ii) the new proposed demarcation lines to distinguish H II regions from other sources of ionization (Section 4.2); (iii) the distribution of H II regions by morphology (Section 4.3) and (iv) by galactocentric distance (Section 4.4); and (v) the revisited relations between the properties of the underlying stellar population and the line ratios in H II regions (Section 4.5). Finally, we show the demarcation line for the other two most used diagnostic diagrams in Appendix A. We describe how to access the catalogue and the code in Appendix C. We show the equivalent width of H  $\alpha$ , EW(H  $\alpha$ ), for the DIG sample in Appendix D.

## 2 SAMPLE AND DATA

We selected the analysed galaxies from the extended CALIFA (eCALIFA), the pilot CALIFA (pCALIFA), and PISCO samples (Sánchez et al. 2012a, 2016c; Galbany et al. 2018). eCALIFA comprises all galaxies observed by the CALIFA survey (Sánchez et al. 2012a), and all the extensions to the original sample observed using the same configuration and almost similar selection criteria (as described in Sánchez et al. 2016c). pCALIFA comprises the galaxies observed by the CALIFA survey in its pilot phase. PMAS/PPak Integral-field Supernova hosts COmpilation (PISCO; Galbany et al. 2018) sample comprises IFS data of 232 supernova host galaxies. The three surveys (eCALIFA, pCALIFA, and PISCO) were observed with the 3.5 m telescope at the Calar Alto Observatory (CAHA). eCALIFA + pCALIFA + PISCO comprises a homogeneous data set in terms of set up, depth, and redshift range. It includes galaxies selected in a similar way (basically matching the FoV of the instrument) observed along the last 9 yr without a fixed observing schedule. We based the current results on the 924 galaxies observed up to 2019 August. We will call them eCALIFA data set for simplicity along this article hereafter.

The details of the CALIFA survey, sample selection, observational strategy, and reduction are explained in Sánchez et al.

(2012a). The details of PISCO sample are explained in Galbany et al. (2018). The whole galaxy sample was observed using the PMAS spectrograph (Roth et al. 2005) in the PPAK configuration (Kelz et al. 2006). The instrument FoV has a hexagonal shape of 74 arcsec  $\times$  64 arcsec. In order to map the full optical extent of the galaxies up to  $\sim 2.5R_e$  within this FoV, the galaxies of mother sample were selected by their diameter (Walcher et al. 2014), a primary criterion adopted for all extended sub-samples. Details of the particular selection criteria for each CALIFA (and PISCO) extension are given in Sánchez et al. (2016c). The complete coverage of the FoV is guaranteed with the observing strategy described in Sánchez et al. (2016c), with a final spatial resolution of FWHM  $\sim 2.4$  arcsec, corresponding to  $\sim 0.8$  kpc at the average redshift of the survey. The spectroscopic resolution and sampled wavelength range for the V500 set-up (the currently adopted in this study) are 3745–7200 Å and  $\lambda/\Delta\lambda \sim 850$ , respectively. This resolution and range are sufficient to analyse the most important ionized gas emission lines from [O II] $\lambda$ 3727 to [S II] $\lambda$ 6731 in the redshift range of our galaxy sample and to deblend and subtract the underlying stellar population (e.g. Kehrig et al. 2012; Sánchez et al. 2012a; Cid Fernandes et al. 2013, 2014; Sánchez et al. 2016a).

The data were reduced using version 2.2 of the CALIFA pipeline. The modifications with respect to the early versions presented in Sánchez et al. (2012a), Husemann et al. (2013), and García-Benito et al. (2015) are described in detail in Sánchez et al. (2016c). In summary, the data fulfil the predicted quality-control requirements with a spectrophotometric accuracy that is better than 5 percent everywhere within the explored wavelength range, both absolute and relative, with a depth that allows us to detect emission lines in individual H II regions as faint as  $\sim 10^{-17}$  erg s $^{-1}$  cm $^{-2}$  and with a signal-to-noise of S/N  $\sim 3$ –5. For the strong emission lines considered in this study, the S/N is well above this limit, and the measurement errors are negligible in most of the cases (as described in Sánchez et al. 2015a). Unavoidably, the S/N of the weak emission lines, e.g. [O III] $\lambda$ 4363, could be less than  $\sim 3$ –5. In all cases, they have been propagated and included in the final error budget.

The final product of the data reduction is a regular-grid data cube, with two spatial dimensions ( $x$  and  $y$  coordinates corresponding to the right ascension and declination of the target, respectively) and one spectral dimension ( $z$  coordinate corresponding to a common step in wavelength). The CALIFA pipeline also provides a proper mask cube of bad pixels, the propagated error cube, and a prescription of how to handle the errors when performing spatial binning (due to covariance between adjacent pixels after image reconstruction). We use as the starting point of our analysis these data cubes, together with the ancillary data described in Walcher et al. (2014).

## 3 ANALYSIS

### 3.1 Stellar population and emission line properties

Spatially resolved stellar population and emission line properties provided by the PIPE3D pipeline (developed in purpose for the analysis of IFS data; Sánchez et al. 2016b) are used in this work. This pipeline has been extensively used in the study of CALIFA (e.g. Sánchez-Menguiano et al. 2016; López-Cobá et al. 2019; Méndez-Abreu, Sánchez & de Lorenzo-Cáceres 2019), MaNGA (e.g. Barrera-Ballesteros et al. 2016, 2017, 2018; Ibarra-Medel et al. 2016; Sánchez et al. 2018), SAMI (e.g. Sánchez et al. 2019), and MUSE (López-Cobá et al. 2017) data sets. PIPE3D adopts FIT3D as the core fitting package (Sánchez et al. 2016a), and the GSD156

<sup>4</sup>[http://ifs.astroscu.unam.mx/CALIFA/HII\\_regions/](http://ifs.astroscu.unam.mx/CALIFA/HII_regions/)

<sup>5</sup><https://github.com/cespinosa/pyHIIexplorerV2>

simple stellar population library in the current implementation. This library that includes 156 templates: 39 stellar ages (from 1 Myr to 14.1 Gyr), and four metallicities ( $Z/Z_{\odot} = 0.2, 0.4, 1, \text{ and } 1.5$ ) has been already used in many different previous publications (e.g. Pérez et al. 2013; González Delgado et al. 2014; Ibarra-Medel et al. 2016; Méndez-Abreu et al. 2019).

Details of the fitting procedure, the adopted dust attenuation curve, and limitations of the processing of the stellar populations are given in Sánchez et al. (2016a, b). We provide here a summary for understanding the nature of the properties explored in this study. First, a spatial binning is performed in each data cube to increase the S/N of the continuum ( $S/N \sim 50$  per  $\text{\AA}$  at  $5000 \text{\AA}$ ) at any location within the FoV. This limit was selected as a compromise between not losing significant spatial information and providing accurate properties of the stellar populations. It is based on the simulations presented in Sánchez et al. (2016a). As shown in Sánchez et al. (2012a, 2016c), a significant fraction of the original spaxels ( $\sim 50$  per cent) have an S/N above this goal, and those remain to unbind. For the remaining ones, the adjacent spaxels are aggregated, co-adding the corresponding spectra, to increase the S/N. However, contrary to other binning schemes (e.g. Voronoi binning, Cappellari & Copin 2003), the adopted one only aggregates those spaxels whose relative flux intensities differ less than 15 per cent from each other at the selected wavelength range ( $5000 \text{\AA}$ ). In this way, the original shape of the galaxy is better preserved, and there is no mixing of nearby regions that correspond to different structures (e.g. arm/inter-arms). The penalty is that the goal S/N is not reached for all the final bins, which in general, have an S/N well above 40 for the CALIFA data (see Ibarra-Medel et al. 2016, for some examples of the procedure). On average,  $\sim 1000$ – $2000$  tessellas (binned regions) are obtained, along with their corresponding spectra for each CALIFA data cube.

After the binning process, the stellar population is fitted for each of the binned spectra following two steps. As the first step, the stellar velocity and velocity dispersion, and the stellar dust attenuation, are derived using a limited sub-set of the stellar library described before. As the second step, a multi single stellar population (SSP) linear fitting is performed, using the full GSD156 library, adopting the kinematics and dust attenuation derived in the previous step. A Monte Carlo iteration is performed for the second step, perturbing the original spectra with their corresponding errors, to provide errors on the estimated stellar parameters. This procedure gives a stellar-population model for each spaxel by re-scaling the model within each bin to the flux intensity of the continuum at the corresponding spaxel (within the considered bin; Cid Fernandes et al. 2013; Sánchez et al. 2016a). Using this model, we estimated the stellar population properties in each spaxel, generating maps of each property. The most relevant parameter considered in this article is the percentage of light and mass (weights) corresponding to each stellar population template within the library, in particular, for the fraction of young and old stellar populations (separated by a specific age limit, as we discuss later). Besides, several properties of the stellar component such as the stellar mass density, light-weighted (LW) and mass-weighted (MW) average stellar age ( $\text{Age}_{\text{LW, MW}}$ ) and metallicity ( $[Z/H]_{\text{LW, MW}}$ ), and the average dust attenuation (together with the kinematic properties already described before) are derived. The maps of each of these properties are packaged as channels (slices) of two data cubes (for convenience), comprising the average stellar properties (SSP cubes) and the weights of the decomposition of the stellar population (SFH cubes), as described in Sánchez et al. (2016b).

In order to derive the properties of the ionized gas emission lines, for each spaxel, the stellar spectrum model (as described above) is subtracted from the original one creating a ‘gas-pure cube’. This cube contains the emission lines together with the noise and the residuals of the stellar-population modelling. The main properties of a set of 51 emission lines are then derived for each spaxel based on a weighted momentum analysis (described in detail in Sánchez et al. 2016b), including for each line the integrated flux intensity, the velocity, velocity dispersion, and the EW. The set of emission lines comprises the most relevant ones detected in the optical spectrum of H II regions (see a complete list in Sánchez et al. 2007), as detected in the Orion nebulae using a similar wavelength range. The final data products of this analysis, described in Sánchez et al. (2016b, 2018), comprise a data cube for each galaxy in which each channel (slice) contains the spatial distribution (map) of each of the derived properties for each analysed emission line (plus the corresponding maps for their errors). These data products (labelled as `fluxelines` cubes), together with the ones storing the properties of the stellar populations, are the primary input for the further analyses performed in this article.<sup>6</sup>

### 3.2 H II detection and extraction

We performed the segregation of H II regions using PYHIIEXPLORER.<sup>7</sup> This code is based on HIIEXPLORER, originally written in PERL (Sánchez et al. 2012b). PYHIIEXPLORER does essentially the same procedures of HIIEXPLORER, re-coding it in PYTHON, which is more commonly used nowadays in astronomy. This makes a easier to distribute, install, (semi-) automatic update, and a major compatibility with other PYTHON codes or packages. We optimized this new version to use all the advantages of data processing that PYTHON offers. Finally, we implement PYHIIEXPLORER to run in parallel so that it can process several galaxies at the same time, depending on the available number of cores.

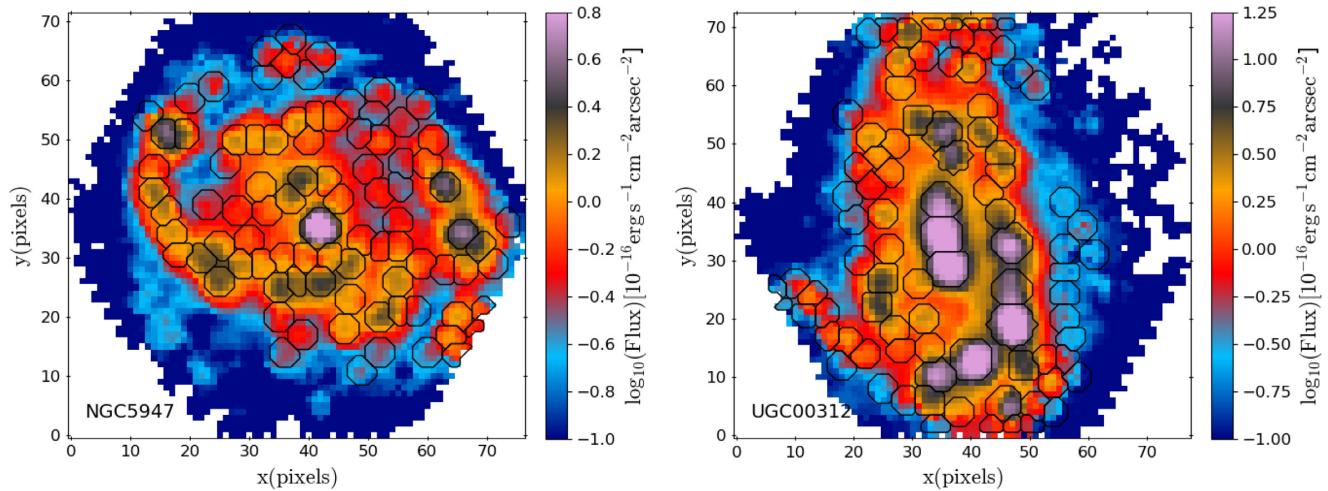
The detection of ionized regions performed by PYHIIEXPLORER is based on two assumptions. (1) H II regions have strong emission lines that are clearly above the continuum emission and the average ionized gas emission across each galaxy. (2) The typical size of H II regions is about a few hundreds of parsecs, which corresponds to a usual projected size of a few arcsec at the distance of our galaxies. These assumptions will define clumpy structures with a high H  $\alpha$  emission line contrast in comparison to the continuum.

Thus, the main input of the algorithm is an H  $\alpha$  emission map (preferentially, as it is the most intense emission line in H II regions in the optical regime). The process to identify ionized regions consists of two steps: (i) the identification of the emission peaks (i.e. local maximum) from the input emission map, and (ii) the aggregation of adjacent pixels to each peak in order to segment the image. For this, the algorithm requires some additional input parameters: (i) a minimum flux intensity threshold for the peak emission; (ii) a minimum relative fraction with respect to the peak emission to keep aggregating nearby pixels; (iii) a maximum distance to the ionized region peak emission to stop the aggregation; and (iv) a minimum absolute intensity threshold in the adjacent pixels to continue the aggregation.

The first, second, and fourth parameters are intensity lower limits, and the third one is a physical-geometry parameter. The peak (local

<sup>6</sup>A subset of these products was distributed for a subsample of the current analysed galaxies in Sánchez et al. (2016b).

<sup>7</sup><https://github.com/cespinosa/pyHIIexplorerV2>



**Figure 1.**  $H\alpha$  emission maps of NGC 5947 and UGC 00312 of the analysed sample; the colour code represents flux intensity. Black contours show the ionized regions found by PYHIEXPLOER for each galaxy.

maximum) emission must be higher than the first parameter to define a new H II region. The algorithm starts looking for a peak (local maximum) and aggregates the adjacent pixels following the described criteria until no further one fulfils them. Then it looks for a new peak emission and iterates the process until no new local maximum is found. Finally, it iterates again over the segregation maps to redistribute pixels to the nearest peak when two or more adjacent H II regions overlap following the previous criteria.

All the parameters described above can be derived by visual inspection of the  $H\alpha$  emission map or with a statistical analysis of the same emission maps (i.e. determining the  $3\sigma$  intensity threshold, the typical  $H\alpha$  flux of an H II region to our galaxy distance, and their typical size or the point-spread-function (PSF), resolution). In this particular implementation, for the PYTHON version, we adopted the same set of parameters introduced for the PERL one (Sánchez et al. 2012a). It allows us to use the same numerical values for parameters in both versions and make a direct comparison between the results.

The output of the algorithm is a FITS file that contains a segmentation map (see Fig. 1) similar in all ways to the one provided by other object detection algorithms like SEXTRACTOR (Bertin & Arnouts 1996). Each ionized region is identified by a unique ID value starting with 1. The pixels that are not associated with any ionized regions have a zero as ID. These areas are regions without emission, or their emission does not fulfil the criteria described before. Besides, a second FITS file is generated containing a masked map where all H II regions identified previously have been masked. The structures of both FITS files are described in Appendix C.

Finally, we extract the flux values of every emission line for each ionized region. For this, we calculate an  $H\alpha$  luminosity-weighted mean with the fluxes in the original cubes (i.e. the data products of PIPE3D described in Section 3.1) for all spaxels with the same ID-index in the derived segmentation map. The final flux values and their associated errors are stored in tables for each ionized region ordered by rows (see Appendix C). Also, we extract the properties and the weights of the decomposition of the underlying stellar population (see Section 3.1) for each ionized region and store it in two other tables with similar structures as the flux tables (see Appendix C).

For this particular study, the input parameters used to identify the ionized regions from the  $H\alpha$  emission maps of all galaxies analysed here are: (i) a flux intensity threshold for each ionized

region peak emission of  $3 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-1}$ , (ii) a minimum relative flux to the peak emission for associated spaxels corresponding to the same H II regions of 5 per cent, (iii) a maximum distance to the location of the peak of 5.5 arcsec, and (iv) an absolute flux intensity threshold of  $0.5 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-1}$  in the adjacent pixels to associate them to the peak emission. All these parameters were selected by a try-and-error procedure on a handful of galaxies, in order to maximize the number of detected H II without including clear diffuse ionized regions, based on the experience of previous studies (e.g. Sánchez et al. 2012b, 2015a; Sánchez-Menguiano et al. 2016, 2018). These parameters were derived by performing a visual inspection and a statistical analysis of the input emission line map, in order to maximize the number of detected ionized regions for the bulk sample of galaxies.

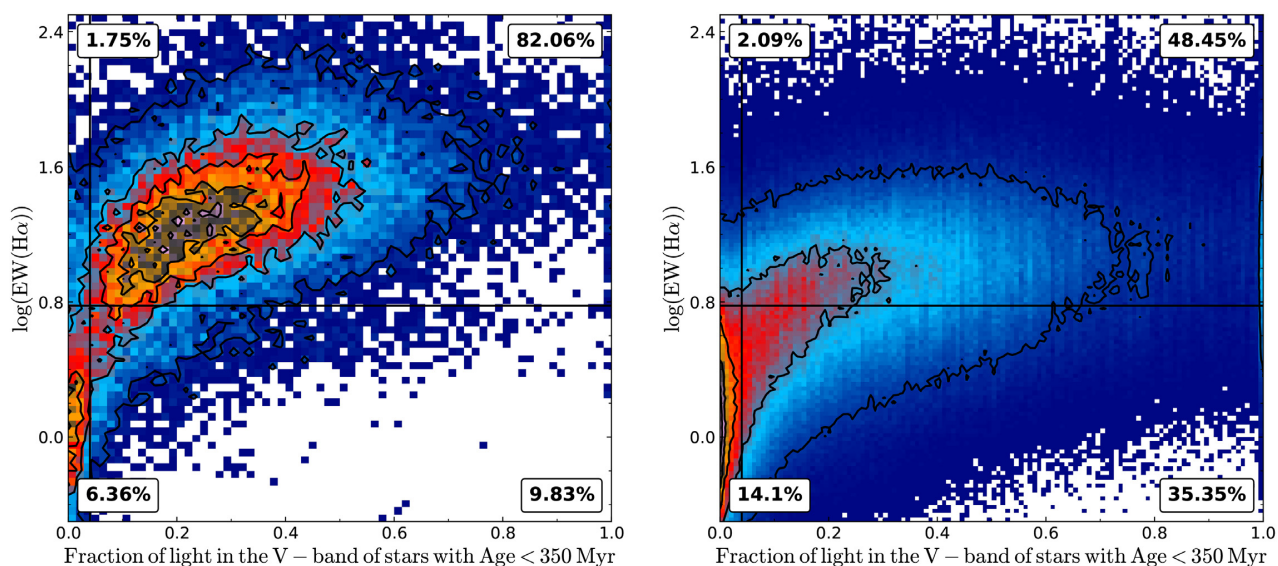
The current catalogue was generated by running PYHIEXPLOER v2 on a computer system with an AMD® Ryzen® 2700x CPU with a 3.7 GHz clock speed. This CPU has 8 cores (and 16 threads) that can directly address 16 GB of RAM. The complete generation of the catalogue running PYHIEXPLOER in parallel mode took  $\sim 25$  min.

### 3.3 H II region catalogue

We apply the procedure outlined in the previous section over 924 data cubes corresponding to a similar number of galaxies observed within the eCALIFA + pCALIFA and PISCO survey. A total of 38 807 clumpy ionized regions were identified. In order of segregate the H II regions, from the clumpy ionized regions sample, we select those that fulfil the following criteria: (i) a minimum  $S/N > 3$  in the  $H\alpha$  flux; (ii) an  $H\alpha/H\beta$  ratio above 2.7, to reject those regions with un-physical Balmer ratios; and finally (iii) a minimum  $EW(H\alpha) = 6 \text{ \AA}$  and (iv) a fraction of young stars to the total luminosity ( $f_y$ ) above of the 4 per cent was applied, following Sánchez et al. (2014).

The cuts in the  $EW(H\alpha)$  and  $f_y$  were selected based on the previous knowledge of the properties of H II regions and after a detailed inspection of the distribution of both parameters (e.g. Sánchez et al. 2015a; Morisset et al. 2016; Lacerda et al. 2018). Fig. 2 shows the distribution of  $EW(H\alpha)$  against the fraction of light in the V band of stars younger than 350 Myr ( $f_y$ ) for (left-hand panel) all clumpy ionized regions selected by the code and (right-hand panel) all spaxels that do not belong to any clumpy ionized





**Figure 2.**  $\text{EW}(\text{H}\alpha)$  versus the fraction of light in the  $V$  band with age lower than 350 Myr for the underlying stellar population for the 38 807 detected ionized regions (*left-hand panel*) and for the pixels excluded by the segregation code (*right-hand panel*), i.e. the diffuse (not clumpy) regions. The light fractions are based on the SSP analysis performed by PIPE3D (Sánchez et al. 2016b). The vertical and horizontal lines (in both the panels) represent the adopted limits for the  $\text{EW}(\text{H}\alpha)$  and fraction of young stars to identify H II regions. The contours (and colour code) show the density distribution of regions (or spaxels). The outermost contour encloses 95 per cent of the regions and each consecutive one encloses the 75, 55, and 35 per cent of the points. Finally, in both the panels we show the percentage of the number of ionized regions (*left-hand panel*) and DIG spaxels (*right-hand panel*) for each subdivision according our limits in  $\text{EW}(\text{H}\alpha)$  and fraction of light.

region (i.e. not selected by the code). The vertical and horizontal lines correspond to the percentage of young stars equals 4 per cent, and the demarcation limits of  $\text{EW}(\text{H}\alpha)=6 \text{ \AA}$ , respectively. These limits were selected as a compromise between preserving the largest number of H II regions and removing the largest number of ionized regions that are not compatible with being an H II region (i.e. lack of enough young stars). Indeed, the actual limits correspond roughly to the location of the contour including 95 per cent of the clumpy regions with young stars (upper right) and with a lack of young stars (bottom left). Those regions that are identified as H II regions (upper right zone in the figure) represent the 67.7 per cent of the total clumpy ionized regions detected by the code.<sup>8</sup>

Applying all the criteria described above, we recover 26 408 H II regions. Unlike other selection methods to obtain H II regions commonly used in the literature, our method is based only on well-known properties of H II regions: a clumpy ionized gas structure with high  $\text{H}\alpha$  emission and a young underlying stellar population. The first criterion is satisfied by construction; it is a fundamental part of the detection algorithm described in Section 3.2. The second one is provided by the adopted  $\text{EW}(\text{H}\alpha)$  and  $f_y$  cuts that ensure that the ionized gas is indeed dominated by a young stellar population ( $f_y$ ) and has a high contrast in the  $\text{H}\alpha$  emission [ $\text{EW}(\text{H}\alpha)$ ].

### 3.4 The diffuse ionized gas

It is well known that galaxies exhibit a low-intensity emission-line spectrum that it is broadly distributed along all their optical extension, known as DIG. Different authors have studied the DIG in the literature (e.g. Reynolds 1991; Rand 1996, 1998; Minter &

Balsler 1997; Minter & Spangler 1997; Zhang et al. 2017; Lacerda et al. 2018). Unfortunately, its physical origin and properties are still not fully understood. It is detected far from the galactic plane as well as close to it (e.g. Flores-Fajardo et al. 2011, Levy et al. 2019), and in galaxies of any morphological type (e.g. Papaderos et al. 2013; Singh et al. 2013). The first detection of DIG was in the areas adjacent to classical H II regions on the Galactic disc (Reynolds 1971); however, later works proved the existence of DIG at large distances above the galaxy planes (e.g. Hoopes, Walterbos & Greenwalt 1996; Hoopes, Walterbos & Rand 1999). Oey et al. (2007) showed that DIG emission is present in galaxies of all types, and it may represent up to 60 per cent of the total  $\text{H}\alpha$  flux intensity (Relaño, Peimbert & Beckman 2002). However, other authors reduce its contribution to  $\sim 4$ –10 per cent when flux intensity is corrected by the dust attenuation (i.e. its contribution to the total luminosity; Sánchez et al. 2012b).

In the literature, the ionizing source of the DIG has been widely studied by different authors, too (e.g. Ferguson et al. 1996; Rand 1998). The radiation from massive OB stars leaked out from H II regions is one of the most widely accepted candidates (Haffner et al. 2009). Another one is ionization by hot low-mass evolved stars (HOLMES; Flores-Fajardo et al. 2011; Stasińska 2012) or post-AGBs stars (Papaderos et al. 2013). Finally, low-velocity shocks could contribute somehow to the DIG ionization (Dopita & Sutherland 1996; Monreal-Ibero, Arribas & Colina 2006; Monreal-Ibero et al. 2010), although its contribution is still not well estimated. Most probably, the ionizing source is different depending on the galaxy type, and on the location within galaxies (Lacerda et al. 2018; Sanchez 2019). In more early-type galaxies and in the retired areas of late-type ones, the two later ionizing sources are maybe the most common ones (e.g. Singh et al. 2013). On the other hand, in more late-type galaxies and in the vicinity of SF areas, the first proposed one is more likely to happens (Zurita, Rozas & Beckman 2000; Relaño et al. 2002; Zurita et al. 2002).

<sup>8</sup>In Appendix B, we compare the distribution of  $[\text{O III}] \lambda 5007/\lambda 4959$  and the  $[\text{N II}] \lambda 6584/\lambda 6548$  flux emission line ratios for the H II regions with the theoretical values.

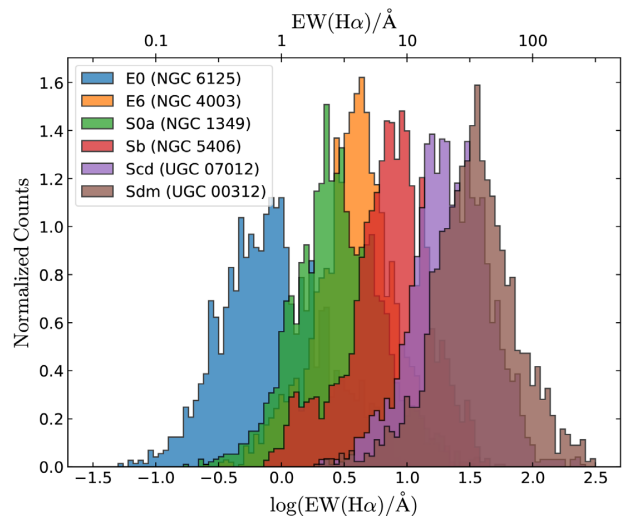
As indicated before, DIG may be present anywhere in galaxies, and in particular, juxtaposed and mixed with the ionization of H II regions. This may contaminate the emission lines detected at that location, and alter the line ratios if the ionization is different than the one produced by nearby young massive stars themselves. The effects of DIG in low-resolution IFS data were explored by Mast et al. (2014) and Zhang et al. (2017), demonstrating that it can critically modify physical derivations like the oxygen abundance (recently addressed by Vale Asari et al. 2019). The physical spatial resolution of the eCALIFA data set is, on average, better than the resolution at which the spatial contamination effect is critical. However, these results highlight the need to attempt a correction of the DIG contamination in the emission line intensities derived for the detected H II regions.

Unfortunately, it is not possible to determine and decouple the contribution of the DIG in the H II regions from the individual spectra themselves: (i) The emissions by the H II regions are in general stronger, and it opaquely the emission of the DIG; (ii) at the spectral resolution of the data, both emissions have similar kinematic properties, precluding any decoupling. A tentative approach would be to estimate this contribution using the DIG emission at the adjacent areas of the detected H II regions and interpolate it. However, there are several potential problems in this approach. First, if the DIG is ionized by photons leaked from the H II regions themselves, then removing its contribution is unnecessary (at a first order), since it should show very similar emission properties, just re-scaled to lower intensity values. Secondly, due to the limited spatial resolution of our data, the wings of the PSF may create a considerable contribution to the adjacent areas. Again, this ionization corresponds to the H II regions and should not be removed. In order to decontaminate the DIG and preserve the contribution due to young stars, we characterize properties and determine the dominant ionizing source of DIG.

### 3.4.1 Characterizing the DIG ionizing source

A commonly adopted procedure to segregate the DIG ionized by old stars and the SF regions is using the  $H\alpha$  surface brightness ( $\Sigma_{H\alpha}$ ) since it is directly related to the density of ionized gas (Oey et al. 2007; Zhang et al. 2017). However, this criterion is valid only for face-on thin galaxies. Due to projection effects and for galaxies with any thickness,  $\Sigma_{H\alpha}$  becomes an extensive quantity, since the light is co-added along the line of sight for the same observed area. This was recently discussed by Lacerda et al. (2018), who proposed a classification scheme based on the EW of  $H\alpha$ . This approach relies on the results by Cid Fernandes et al. (2010) and Sánchez et al. (2014), and on the fact that the  $EW(H\alpha)$  is always an intensive quantity. According to them the ionized gas could be classified in three groups: (i) areas with  $EW(H\alpha) < 3 \text{ \AA}$  correspond to diffuse gas being ionized predominantly by an old stellar population (HOLMES/post-AGBs); (ii) areas with  $EW(H\alpha) > 14 \text{ \AA}$  correspond to the diffuse gas that is ionized by OB stars; and (iii)  $3 \text{ \AA} < EW(H\alpha) < 14 \text{ \AA}$  corresponds to the diffuse gas that is powered by a mixture of both ionization sources.

Fig. 3 illustrates how the distribution of  $EW(H\alpha)$  for the DIG regions (spaxels excluded from the clumpy regions by PYHIIEXPLORER) correlates with respect to morphology (equivalent with the stellar ages). Early-type galaxies (NGC 6125, NGC 1349, and NGC 4003) tend to have low values of EW in comparison to late-type ones (NGC 5406, UGC 07012, and UGC 00312) (González Delgado et al. 2014). This is consistent with results obtained by



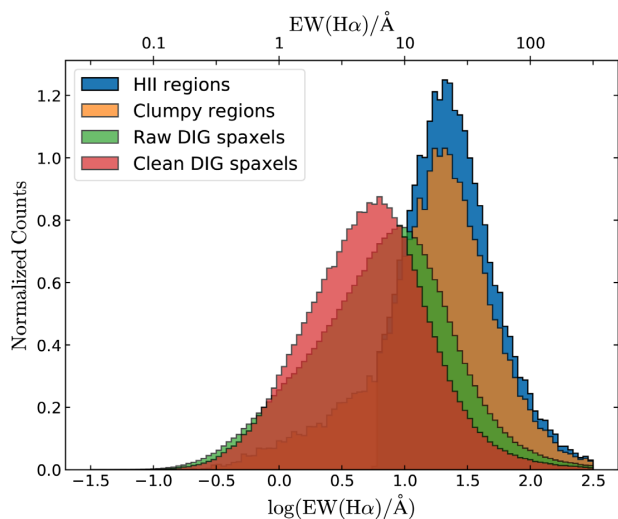
**Figure 3.** Normalized histograms of the  $EW(H\alpha)$  of DIG spaxels along the Hubble sequence. One representative galaxy was chosen for six morphological types: E0 (NGC 6125), E6 (NGC 1349), S0a (NGC 4003), Scd (UGC 07012), Sb (NGC 5406), and Sdm (UGC 00312).

Lacerda et al. (2018). The late-type galaxies show equivalent widths that correspond to ionization due to old stars or leaking/PSF-wings contamination associated with SF areas. Therefore, for early-type galaxies, the associated ionizing sources are predominantly old stars as expected since the star formation is absent in these galaxies. Conversely, for late-type galaxies, both ionizing sources, old and young stars, should be observed. However, the number of spaxels with DIG associated with ionization by young stars should dominate over the number of them being ionized by old stars.<sup>9</sup>

In order to account for the effects of PSF wings, we repeat the selection of DIG areas by enlarging the masked areas around H II regions by one spaxel (i.e. 1 arcsec,  $\sim 700$  pc at the redshift of the sample). However, although this procedure improves somehow the segregation, it does not guarantee a clean separation between the two ionizing sources. We repeat the procedure masking larger areas around the peak intensity of each H II region. The larger the radius, the smaller is the area selected as DIG in each galaxy. Furthermore, the number of galaxies with any area classified as diffuse decreases. Thus, the statistics become poorer. Following this procedure, we are not able to obtain a clear segregation between the different ionizing sources even for very large masking radius (3–4 times the PSF FWHM), when most of the area within the FoV of the data cubes is masked.

As a final approach, we explore the properties of the DIG gas of the sample, considering all galaxies together, trying to derive an average correction that would be applied to each individual spaxel in each data cube. For doing so, we select all the DIG spaxels again for all the galaxies as described in Section 3.4, i.e. by selecting those areas outside the segregation maps used to select the H II-region candidates. The green histogram in Fig. 4 shows the distribution of the  $EW(H\alpha)$  of the diffuse selected in this way (raw DIG). The distribution shows a peak centred on  $\sim 10 \text{ \AA}$ , with an asymmetric shape, more tailed towards the low EW values. This distribution corresponds to that of the mixed DIG or intermediate ionization as described by Lacerda et al. (2018), showing a clear contamination

<sup>9</sup>A further study about the DIG sample and their distribution across the BPT diagram is presented in Appendix D.



**Figure 4.** Normalized histograms of the  $\text{EW}(\text{H}\alpha)$  for the four different subgroups of ionized regions described in the text: (i) the raw DIG spaxels (DIG spaxels for all galaxies); (ii) clean DIG spaxels (DIG spaxels for the clean galaxy sample); (iii) clumpy ionized regions (all those regions identified by PYHIEXPLORER); and finally (iv) H II regions (clumpy ionized regions selected as H II regions; see Section 3.3).

by different sources of ionization. For comparison purposes, the orange histogram in the figure shows the distribution of  $\text{EW}(\text{H}\alpha)$  for the clumpy ionized regions (as selected by PYHIEXPLORER). As expected, it is clearly peaked towards higher EWs ( $\sim 30 \text{ \AA}$ ), although it presents a smooth but clear trend towards low EWs, so this indicates the presence of contamination by ionizing sources not compatible with H II regions.

In an attempt to select just the best suitable spectra to characterize the DIG, we start selecting only those galaxies with high-S/N diffuse emission. First, we exclude those galaxies in which (i) the mean value of half of the emission lines from all DIG spaxels has an SNR value lower than 1.5, and (ii) if any average value of the emission lines presents unrealistic values identified by eye for the DIG emission line fluxes (as a result of a bad fitting or a stellar population subtraction bias, or for regions at the edges of the hexagonal FoV of the data, in most of the cases). The remaining galaxies comprise 30 per cent of the galaxy sample ( $\sim 300$  galaxies,  $\sim 2.3$  million of spaxels). Hereafter, we will refer to the spectra of this galaxy sub-set as the ‘clean DIG’ sample. The effects of this selection are clearly appreciated in the red histogram of Fig. 4. The distribution of  $\text{EW}(\text{H}\alpha)$  for the ‘clean DIG’ sample has a peak now centred on  $\sim 5 \text{ \AA}$ , showing a more symmetric shape. The new values for the EWs are more in agreement with an ionization dominated by old stars. In order to investigate further the true nature of the ionizing source, we explore their distribution in the classical diagnostic diagrams.

Fig. 5 shows the  $[\text{O III}]\lambda 5007/\text{H}\beta$  versus  $[\text{N II}]\lambda 6583/\text{H}\alpha$  diagnostic diagram for all DIG spaxels. The spaxels were segregated by the morphology of their host galaxy in each panel, and the colour represents the mean  $\text{EW}(\text{H}\alpha)$  at each location. The  $\text{EW}(\text{H}\alpha)$  distribution for every group is plotted in an inset within each panel. There is a clear correlation between the location on the BPT diagram and the morphological type of the host galaxy. For early-type galaxies, the line ratios are mostly located in the intermediate area of the diagram (between the Kewley and Kauffmann curves). This region is dominated by a mixture of ionization by HOLMES/post-AGB and young stars. In particular, these line ratios cannot be

explained by SF models (Kewley et al. 2001, 2006; Kauffmann et al. 2003; Stasińska et al. 2006). The distribution for late-type galaxies is more located below the demarcation curves. This zone is dominated by regions that are classically ionized by young stars and leaking photons from H II regions.

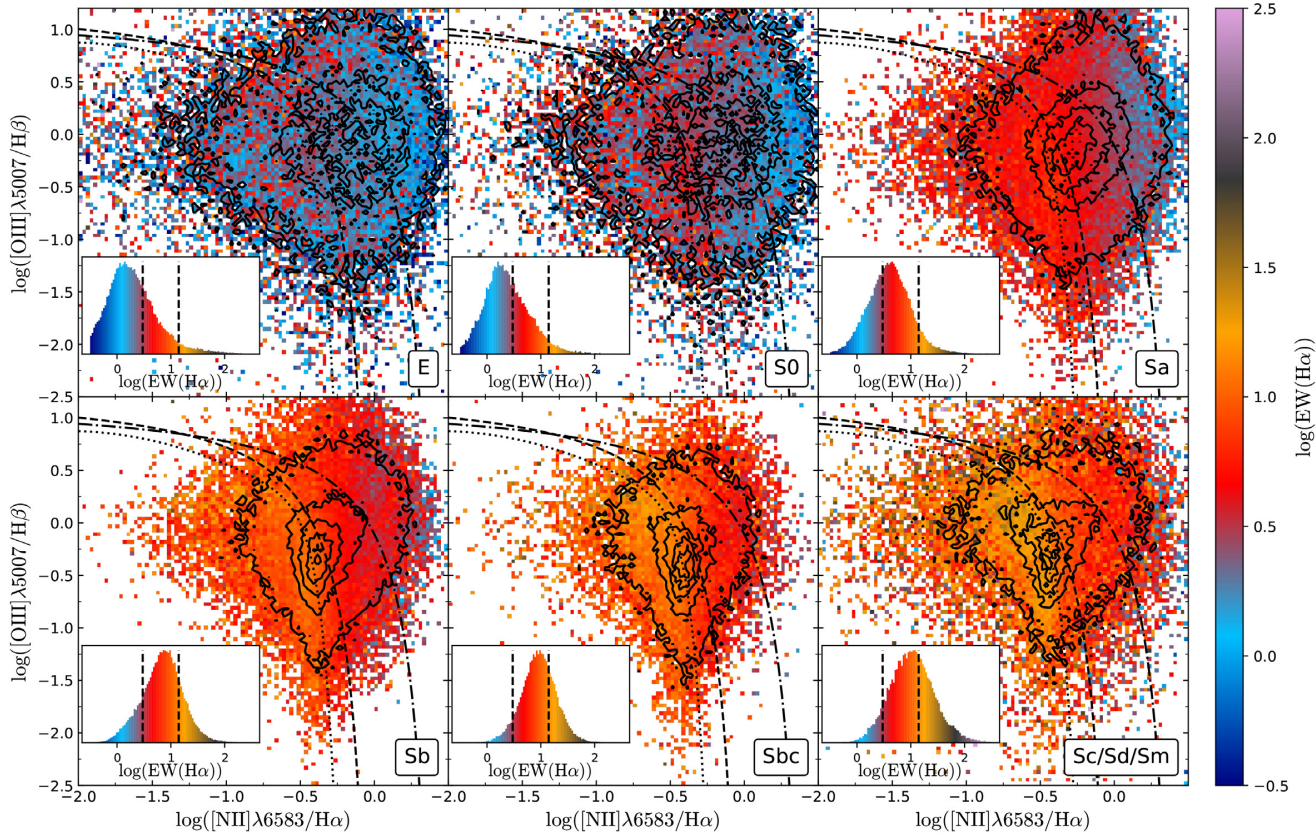
Furthermore, the values of the  $\text{EW}(\text{H}\alpha)$  depend on the morphological type as well. The DIG in early-type galaxies tends to have low values of  $\text{EW}(\text{H}\alpha)$ . Conversely, late-type galaxies show higher values. Therefore, the  $\text{EW}(\text{H}\alpha)$  correlates with the morphology and with the position in the diagnostic diagrams. These results are expected by the correlation between the  $\text{EW}(\text{H}\alpha)$  and the ionizing source, according to Lacerda et al. (2018). However, for low  $\text{EW}(\text{H}\alpha)$  values, there are some spaxels below the Kauffmann’s demarcation line, i.e. they are in the area assumed to correspond to ionization by young stars or post-AGBs (Morisset et al. 2016). This effect is considerably more severe before cleaning the sample based on the criteria indicated before, as shown in Appendix D (Fig. D1), highlighting the need for the applied selection. In addition, there is a clear variation of the  $\text{EW}(\text{H}\alpha)$  values and the location on the BPT diagram along the Hubble sequence, in agreement with Lacerda et al. (2018) and the examples are shown in Fig. 3.

We followed the same reasoning applied to select the H II regions from the clumpy ionized areas (Section 3.3) to select those spaxels within the *clean DIG* sub-set corresponding to ionization due to old stars adopting the inverse criteria. We recall that this criterion is based on the  $\text{EW}(\text{H}\alpha)$  and  $f_y$  (see 3.3). Previously, we selected the regions with  $\text{EW}(\text{H}\alpha) > 6 \text{ \AA}$  and  $f_y > 4$  per cent as areas dominated by ionization due to young stars. On the contrary, to select DIG areas, we use the opposite constraints [ $\text{EW}(\text{H}\alpha) < 6 \text{ \AA}$  and  $f_y < 4$  per cent] as spaxels ionized by old stars. Fig. 2, right-hand panel, shows the distribution of  $\text{EW}(\text{H}\alpha)$  against  $f_y$  for the *clean DIG* spaxels. A direct comparison between the distributions in this diagram for clumpy ionized regions (left-hand panel) and DIG ones (right-hand panel) shows that most of the clumpy regions are already associated with areas consistent with being ionized by young stars. In contrast, most of the DIG ones are concentrated in areas associated with ionization by old stars. We should note here that a fraction below 4 per cent of young stars is still reliable based on the inversion method adopted to derive the stellar population (FIT3D; Sánchez et al. 2016a). Different tests with similar codes indicate that below a 3 per cent the fraction of young stars is very unreliable (e.g. González Delgado et al. 2016), and below this limit, all fractions should be considered just upper limits (e.g. Bitsakis et al. 2019). On the other hand, based on our criteria, the upper right zone of the diagram for the DIG area corresponds, most probably, to contamination by leaking photons from H II regions and the contribution of the PSF wings (which have similar properties in this diagram). The two other remaining quadrants are much less populated, and they are most probably contaminated by low-S/N DIG areas, both in the continuum and the emission line properties, which shifts values from the other two ones that have a more clear physical interpretation.

### 3.4.2 Decontaminating the H II regions by the DIG

In the previous sections, we explore the distribution of the DIG and its nature. We selected a sub-set of galaxies from which it is possible to explore the properties of the DIG and define a procedure to determine which spaxels are ionized by old stars. In this section, we describe how we derive the average properties of the emission





**Figure 5.**  $[\text{O III}]\lambda 5007/\text{H}\beta$  versus  $[\text{N II}]\lambda 6583/\text{H}\alpha$  diagnostic diagrams for all the spaxels selected as *clean DIG*, segregated by the morphology of their respective host galaxies. The colour indicates the  $\text{EW}(\text{H}\alpha)$ . The inset within each panel shows the  $\text{EW}(\text{H}\alpha)$  distribution of the corresponding spaxels. In these inset plots, we show the  $\text{EW}(\text{H}\alpha)=3$  and  $14 \text{ \AA}$  as black dashed vertical lines. The dashed, dot-dashed, and dotted lines represent the Kauffmann et al. (2003), Kewley et al. (2001), and Stasińska et al. (2008) demarcation curves, respectively. The density contours are similar to those presented in Fig. 2.

line ratios for this final sub-set of spaxels (so far, for our data set, the best possible representation of the DIG). Then, we show how we correct the emission line intensities of our catalogue of H II regions from this DIG contamination.

For a  $j$ -th emission line flux of the  $i$ -th H II region, the corrected flux should be

$$F_{\text{corr}}^{i,j} = F_{\text{obs}}^{i,j} - F_{\text{diff}}^{i,j}, \quad (1)$$

where  $F_{\text{obs}}^{i,j}$  is the observed emission line flux and  $F_{\text{diff}}^{i,j}$  is the emission line flux of the DIG to be corrected from. Thus, we need to estimate this later flux intensity.

In order to do so, we estimate the average flux intensity of all DIG regions included in the *clean* sample. First, for each spaxel  $k$ -th within this sample of DIG areas we normalize the flux of each  $j$ -th emission line by the flux of  $\text{H}\alpha$ , defining the line ratio:

$$R_{\text{diff}}^{k,j} = \frac{F_{\text{diff}}^{k,j}}{F_{\text{diff}}^{k,\text{H}\alpha}}. \quad (2)$$

Then, we derive the error-weighted mean value of all these line ratios through all the *clean* sample of DIG spaxels, by using the formula

$$C_j = \frac{\sum_{k=1}^{k=n} w_{k,j} R_{\text{diff}}^{k,j}}{\sum_{k=1}^{k=n} w_{k,j}}, \quad (3)$$

where  $w_{k,j}$  is the square of the inverse of the error of the  $j$ -th emission line for the  $k$ -th spaxel, and  $n$  is the total number of *clean* DIG spaxels.

Once derived  $C_j$ , i.e. the error-weighted mean value of the line ratio with respect to  $\text{H}\alpha$  for the  $j$ -th emission line for the DIG, we can calculate the DIG emission line flux of that line that would correspond to a particular  $i$ -th H II region,  $F_{\text{diff}}^{i,j}$ , using the formula

$$F_{\text{diff}}^{i,j} = F_{\text{diff}}^{i,\text{H}\alpha} \times C_j, \quad (4)$$

where  $F_{\text{diff}}^{i,\text{H}\alpha}$  is the  $\text{H}\alpha$  flux of the diffuse gas. In principle, we do not have a direct way to estimate this flux. However, we can use the definition of the EW, and determine the relation

$$F_{\text{diff}}^{i,\text{H}\alpha} = \text{EW}_{\text{diff}}^{i,\text{H}\alpha} \times \rho^{i,\text{H}\alpha}, \quad (5)$$

where  $\rho^{i,\text{H}\alpha}$  is the continuum flux density at the central wavelength of  $\text{H}\alpha$ . We have direct access to this flux density as a direct data product of the analysis by PIPE3D, which provides a version of the original data cube once we subtracted the estimated contribution of the emission lines. These data cubes are used to estimate the EWs of all the emission lines, as explained in Sánchez et al. (2016b). On the other hand, for the DIG (ionized by HOLMES or post-AGBs), we consider that  $\text{EW}(\text{H}\alpha) \approx 1 \text{ \AA}$ , based on theoretical estimations (Binette et al. 1994; Sarzi et al. 2010; Papaderos et al. 2013; Gomes et al. 2016b). Indeed, this value is in agreement with our results shown in the right-hand panel of Fig. 2. Although the observational mean value is  $\sim 1.4 \text{ \AA}$ , we prefer to use a conservative value,  $\text{EW}(\text{H}\alpha) \approx 1 \text{ \AA}$ , to perform the DIG correction. It is important

to note that this value is not related to the mean value shown in Fig. 4 since this distribution is dominated by the PSF-wings contamination and the photons leaking, as indicated before.

Using equations (5) and (4), and the parameters  $C_j$  estimated for each  $j$ -th emission line, we estimate  $F_{\text{diff}}^{i,j}$ . Finally, we decontaminate the contribution of the DIG for each emission line in each H II region within the catalogue using equation (1). The resulting values are stored in a new catalogue of H II regions corrected by DIG.

### 3.4.3 Caveats to the adopted procedure

We should note here that the adopted approach to derive the DIG contribution and its use to decontaminate line fluxes of H II regions is based on different assumptions described below that may not be entirely valid. We have assumed that the only DIG contribution to correct for is the one produced by old stars (HOLMES/post-AGBs), considering that the diffuse ionization by photon leaking would produce similar emission line ratios. That assumption is far from been fully proved (e.g. Relaño et al. 2012). Furthermore, Vale Asari et al. (2019) recently studied the effects of a decontamination by DIG ionized by photon leaking in the derivation of the oxygen abundance. Using SF galaxies from the MaNGA data set (Bundy et al. 2015), they found that there is a difference of up to 0.1 dex in the oxygen abundance calculated with the N2 index at the high metallicity end and low EW. However, there is no significant change in the oxygen abundances derived with other strong line indices (e.g. O23, O3N2, etc.). This agrees with our perception that this particular DIG does not alter reported line ratios significantly (besides maybe N2, with a small effect in any case).

On the other hand, other contributions, like shock ionization, that could be considerably important at different scales (Veilleux, Cecil & Bland-Hawthorn 2005; López-Cobá et al. 2019) are not considered here. Finally, we considered that all the DIG presents the same line ratios everywhere within a galaxy and for different galaxies. This is a very strong assumption that it is only valid at first order, as we have indeed shown in Fig. 5. The line ratios expected by HOLMES/post-AGB ionization cover a wide range of values (e.g. Gomes et al. 2016b; Morisset et al. 2016). However, our empirical exploration of these line ratios indicates a small degree of variation, almost compatible with the expected distribution due to their errors. In addition, we consider that  $\text{EW}_{\text{diff}}^{\text{H}\alpha}$  is 1 Å everywhere across the optical extension of galaxies and for all galaxies. This assumption, based on theoretical expectations, is empirically verified for the DIG of the earlier-type galaxies (Fig. 5). However, for later morphological types, the observed EWs are slightly larger. We assumed that the contamination by other sources of DIG and the wings of the PSF (from the adjacent H II-region) alter the observed EWs of the DIG in these later-type galaxies. We acknowledge all these caveats to the adopted procedure. However, based on the performed experiments, we consider that this is the best approach we can adopt so far for the considered data set. Being aware of possible improvements in the procedure, we provide with an uncorrected version of the H II catalogue together with the corrected one to allow the community to perform their evaluation on the issue of the DIG contamination.

## 4 RESULTS

### 4.1 Effects of the DIG decontamination

Fig. 6 shows the distribution across the classical BPT diagnostic diagram for the final sample of H II regions before and after applying

the DIG correction (left-top and right-top panels). As can be seen, most of the uncorrected values are located below the maximum starburst envelope demarcation line (Kewley et al. 2001). This result is expected since the criteria to select H II regions are based on the physical properties of this type of nebulae. After DIG correction, there is a change in the emission line fluxes and, consequently, in the respective EWs, fluxes ratios, and the distribution across the diagnostic diagram. The distribution stretches at the edges, covering a slightly wider range of line ratios. However, the average trend is very similar to the contaminated fluxes. Most of the decontaminated values remain in the SF zone in the diagnostic diagram. Besides, the  $\text{EW}(\text{H}\alpha)$  values decrease slightly on average, since the continuum flux density is constant along the DIG correction calculations (see Section 3.4.2), but the flux intensities have decreased slightly too. This effect is stronger for the H II regions with lower  $\text{EW}(\text{H}\alpha)$ .

This result is somewhat expected. Mast et al. (2014) have already shown the effects of the contamination by DIG on the flux ratios used in the explored diagnostic diagrams. Although our correction only takes into account a diffuse mainly powered by old stars and post-AGBs (among other assumptions), our results show the same trend that they obtained with their methodology. Unfortunately, the spatial resolution of our data does not allow us to identify all diffuse gas components accurately and, hence, our data are not fully decontaminated from the complete DIG contribution. However, the proposed correction is enough to produce a small but noticeable change in the distribution of H II regions along the BPT diagram.

Fig. 6 shows the classical BPT diagram for the H II regions sample corrected by DIG with (bottoms panels) a colour code representing the difference between the flux corrected and the flux uncorrected for each line ratio in the diagram.<sup>10</sup> As can be seen, the most significant change in the emission line fluxes is in the edges of the distribution (upper left for  $[\text{N II}]/\text{H}\alpha$  and lower right for  $[\text{O III}]/\text{H}\beta$ ). These results are in agreement with the change of the density points between the original distribution and the decontaminated one (upper panels of the same figure). Also, the error-weighted mean value,  $C_j$  (see Section 3.4.2), for the corresponding line ratios is shown as a white star. The value of  $\Delta[\text{N II}]/\text{H}\alpha$  and  $\Delta[\text{O III}]/\text{H}\beta$  depends on the position of  $C_j$ .

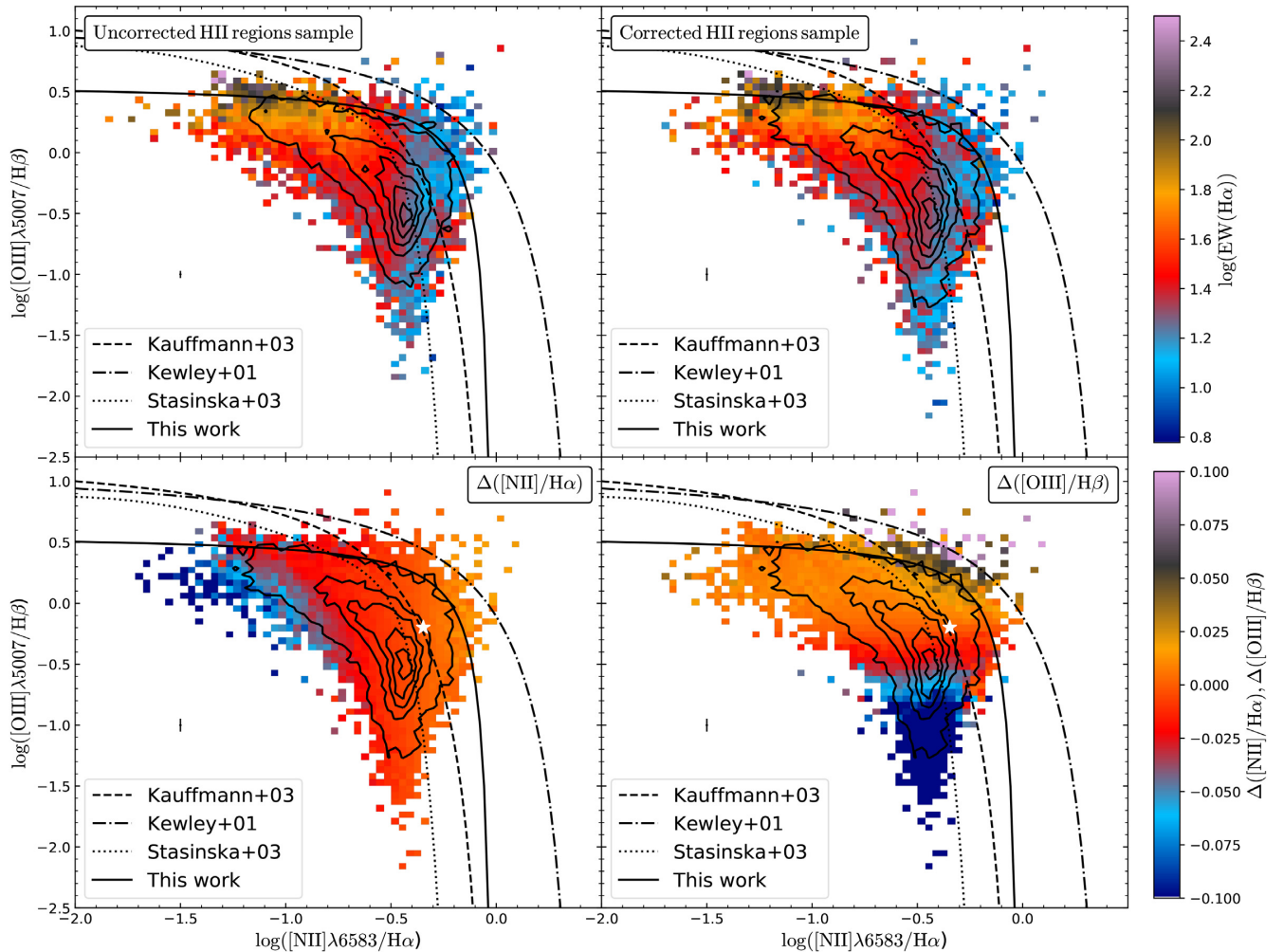
### 4.2 An empirical demarcation line for H II regions

As indicated before, Fig. 6 shows the classic diagnostic diagram for the entire sample of H II regions (corrected and uncorrected by DIG). For comparison purposes, we include in Fig. 6 the demarcation lines defined by Kewley et al. (2001) (dashed-dotted line), Kauffmann et al. (2003) (dashed line), and Stasińska et al. (2006) (dotted line). Since each demarcation line was derived on different assumptions, they do not provide the same classification (what is appreciated in the figure).

The most conservative demarcation line, in terms of minimizing the possible contamination by other ionizing sources different from H II/SF regions, is the one proposed by Stasińska et al. (2006) (dotted line in Fig. 6). This curve, like the one proposed by Kewley et al. (2001) (dot-dashed line in Fig. 6), defines the maximum envelope of line ratios derived by the location of pure SF regions based on photoionization models. The observed differences between both curves are due to the different models adopted in their derivations. Finally, an empirical demarcation line was derived based on a large sample of SF SDSS galaxies (Kauffmann et al. 2003; in Fig. 6).

<sup>10</sup>Other diagnostic diagrams are shown in Appendix A.





**Figure 6.** Distribution of the final sample of H II regions along the BPT diagram. *Left upper panel:* Distribution without the correction of the contamination of the DIG. *Right upper panel:* Distribution after correcting by the DIG contamination. In these panels, the colour code represents the EW of  $H\alpha$ . *Left bottom panel:* Distribution of the corrected H II regions with a colour code representing the difference between the flux corrected and uncorrected for the  $[NII]/H\alpha$  ratio ( $\Delta[NII]/H\alpha$ ). *Right bottom panel:* Distribution of the corrected H II regions with a colour code representing the difference between the flux corrected and uncorrected for the  $[OIII]/H\beta$  ratio ( $\Delta[OIII]/H\beta$ ). In these panels, the white star is the error-weighted mean value for the corresponding line ratio. Notice that with the DIG correction the distribution has spread out slightly at the edges. This can be seen better on the density contours on the upper panel and in the colour gradient in the bottom panels. All lines represent the same demarcation curves described in Fig. 5. Also, the demarcation curve proposed in Section 4.2 is shown as a solid line.

This line was defined by inspecting the BPT diagram visually as the line that better distinguishes the AGN from the SF galaxies. We should advise the reader that none of these demarcation lines should be interpreted as strict boundaries between different ionizing sources. As extensively discussed in previous articles (e.g. Sánchez et al. 2018, 2019), the demarcation lines should be interpreted as the upper limit of line ratios that can be achieved due to ionization by massive/young OB stars (Kewley et al. 2001). However, not all the ionizing regions located below that curve are associated with ionization by young stars per se. As clearly shown in Fig. 5, the DIG produced by old stars can be located below any of those curves, in particular, the one defined by Kewley et al. (2001). The regions between the Kewley and Kauffmann curves have been extensively classified as ‘mixed’ or ‘composite’. It is true that for single aperture spectroscopic data that encompass a large portion of the optical extension of galaxies or for low-resolution IFS data (and/or particular inclinations), these line ratios could be reproduced by the

combination of SF + AGN or SF + post-AGB ionizations (e.g. Mast et al. 2014; Davies et al. 2016). However, uncontaminated ionization by (i) certain types of H II regions, like the N-enhanced/shock mixed ones described by Ho et al. (2014) and Sánchez et al. (2012a); (ii) pure DIG produced by stars of certain ages (e.g. Morisset et al. 2016); (iii) pure low-velocity shocks (e.g. López-Cobá et al. 2019), and even (iv) low-metallicity AGNs (e.g. Stasińska 2017) may be located in that area without requiring a mixed ionization.

In this study, we propose a new purely empirical demarcation line that, like the other quoted ones, should be considered as an upper boundary for the involved line ratios due to ionization by young/massive stars. Based on the distribution of our sample of H II regions, we select the envelop that includes 95 per cent of them as the boundary of ionized regions mainly powered by young stars. Then, by fitting this set of line ratios that defines the limit using the same functional form adopted in the literature (e.g. Kewley et al.

2001; Kauffmann et al. 2003), we obtain the following demarcation lines for the most commonly used diagnostic diagrams:

$$\log([\text{O III}]/\text{H}\beta) < 0.13/(\log([\text{N II}]/\text{H}\alpha) - 0.003) + 0.57 \quad (6)$$

$$\log([\text{O III}]/\text{H}\beta) < 0.04/(\log([\text{S II}]/\text{H}\alpha) + 0.012) + 0.58 \quad (7)$$

$$\log([\text{O III}]/\text{H}\beta) < 0.056/(\log([\text{O I}]/\text{H}\alpha) + 0.40) + 0.61. \quad (8)$$

The curve for  $[\text{O III}]/\text{H}\beta$  versus  $[\text{N II}]/\text{H}\alpha$  (6) is represented by a solid line in Fig. 6. The locations of the proposed demarcation lines for the remaining classical diagnostic diagrams are shown in the corresponding figures in Appendix A (Figs A1 and A2).

It is important to note that these new demarcation lines are based on a sample of H II regions that were selected based on only two assumptions: (i) a clumpy structure with a high-contrast H $\alpha$  emission and (ii) an underlying stellar population comprising young stars (i.e. compatible with the assumed source of ionization).

### 4.3 H II regions along the Hubble sequence

Early explorations on the properties of H II regions have determined that later-type spirals present a larger number of H II regions, distributed along (more) open spiral arms than earlier-type ones (Kennicutt 1988, 1989). Furthermore, Kennicutt et al. (1989) and Ho et al. (1997) showed that H II regions in the central areas of early spirals distinguish themselves spectroscopically from those in the disc by their stronger low-ionization forbidden emission. These results were confirmed by more recent studies using large IFS data sets (Sánchez et al. 2012a, 2015b). In particular, Sánchez et al. (2015b) demonstrated that the ionization properties of H II regions strongly depend on the properties of the underlying stellar population, and, therefore, they present clear trends with the morphology as well. The enlarged H II catalogue presented here allows us to perform a more detailed exploration of this dependence.

Fig. 7 shows the distribution of H II regions across the BPT diagram segregated by the morphology of their respective host galaxies. The top left-hand panel shows the 254 regions found in E and S0 galaxies ( $\sim 2$  per cent of the total sample), corresponding to 38 galaxies. Although early-type galaxies are frequently associated with the absence of SF regions, recent results indicate that a small fraction of them present H II regions (e.g. Gomes et al. 2016a). They could be either remnants of a former disc or the result of a rejuvenation due to the capture of gas-rich galaxies or in falling gas (Gomes et al. 2016a). Table 1 shows the number of H II regions along the different morphological types. Most of the H II regions are located in late-type galaxies, with only  $\sim 7$  per cent of the H II regions hosted by E/S0 galaxies. However, most of the H II regions are not found in later spirals, but rather in Sb/Sbc galaxies, comprising 21/30 per cent of the total sample. On average, the number of H II regions per galaxy increases from earlier to later types up to Sc, spanning from  $\sim 7$  regions in E/S0 galaxies to  $\sim 50$  in Sbc/Sc galaxies. Beyond that, for later spirals, the number decreases slightly. However, we cannot be completely sure if this decline is real or due to resolution effects. So far, the seminal studies by Kennicutt (1988, 1998) indicate that the number of H II regions should rise as later is the host, without any decline. The same articles indicate that H II regions in late-type galaxies are significantly brighter than those of more early types. However, the H II regions are limited by the properties of CALIFA's data. Inevitably, we are losing most of the low-luminosity H II regions. As we will discuss later (see Section 5), in a future work, we will explore the methodology used in this work on data with a higher spatial resolution.

The relation between the distribution of H II regions across the BPT diagram and the morphology of their host galaxy uncovered by Sánchez et al. (2015b) is appreciated in Fig. 7. The few H II regions found in early-type galaxies are located around the central region of the BPT diagram without any significant trend, maybe due to the low number of statistics. The trend becomes more evident when exploring the distributions from early-spirals (Sa) towards later ones, with H II regions shifting along the classical sequence of these objects (e.g. Osterbrock 1989).

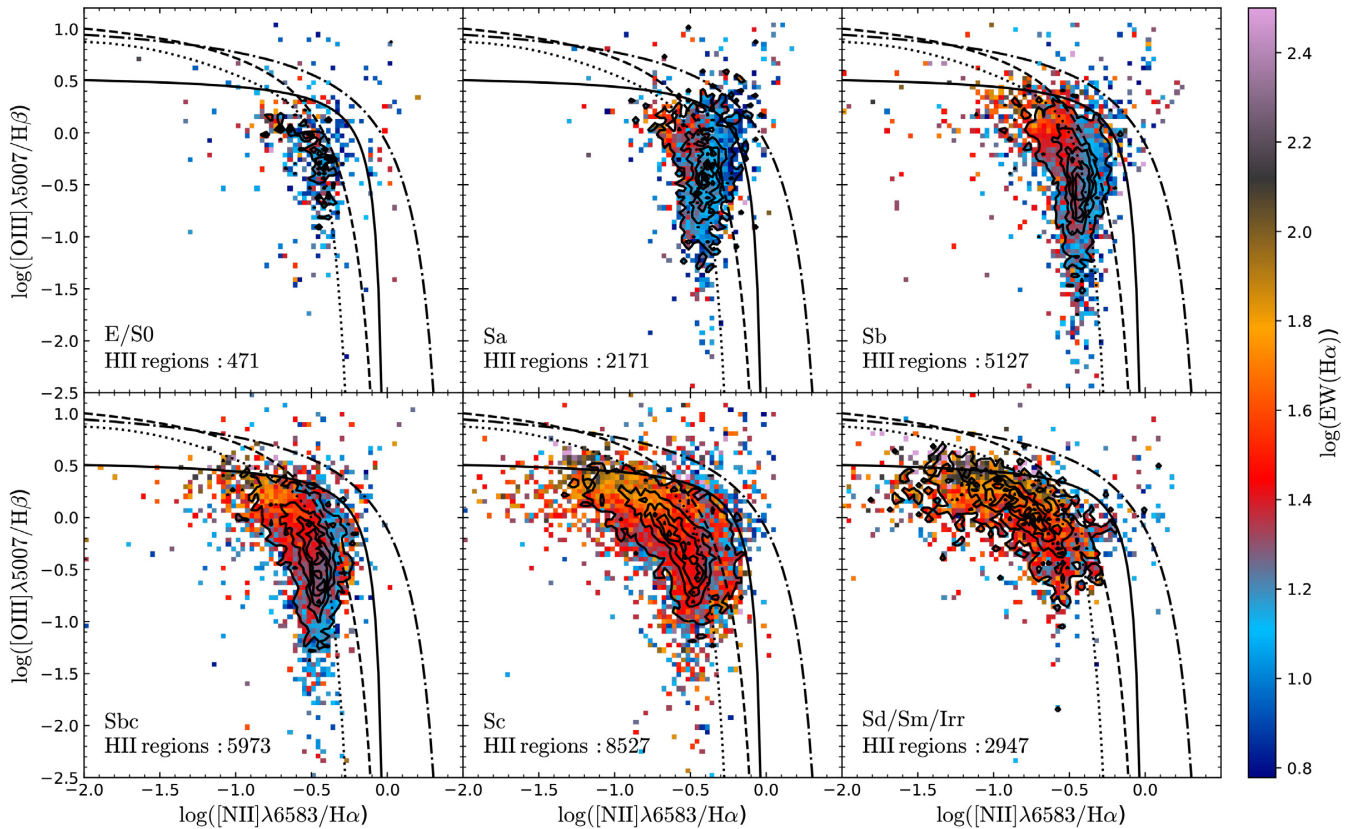
The location in the BPT diagram of an H II region is primarily related to the ionization conditions of the nebulae (e.g. Kewley & Dopita 2002; Morisset et al. 2016). Thus, a priori, it is not expected a connection between this location and the morphological properties of the galaxy that hosts the H II regions, which is what we see precisely in Fig. 7. The low number of H II regions in early-type galaxies, in which SF activity is finished (or almost finished; Stasińska et al. 2008; Blanton & Moustakas 2009), was expected. However, the nature of the observed trend within the BPT diagram with the Hubble sequence, besides the clear increase in the number of H II regions, is less obvious. We can discard strong contamination by DIG as the primary driver for the observed distribution. First, the selection of clumpy ionized regions, together with the implemented cuts (Section 3.3), guarantees that we have indeed selected regions in which OB stars dominate ionization. Secondly, as shown in Section 4.1, this contamination shrinks somehow the range of covered parameters, but it does not blur the observed trends completely. Finally, we have performed a detailed DIG subtraction, so its contribution is minimized.

Assuming that OB stars indeed ionize all the selected regions, the observed trend should be the consequence of variations between these stars (and the surrounding nebulae). For the most late-type galaxies (Sd, Sm, and irregular), the location (upper left end on the diagram) is normally interpreted as the presence of low-metallicity stars with high ionization strengths. Meanwhile, for earlier-spirals (Sa or Sb), the location (lower right end on the diagram) is linked to higher metallicity stars with lower ionization strengths (these results are also found and discussed by Sánchez in press). Finally, in intermediate galaxy types, the H II regions extend within the two regimes. Despite the significant effect of the aging of the ionizing stars (which would impose certain variations within the observed trends), the only obvious conclusion is that we are observing a trend driven by a change in the metallicity, which induces a change in the ionization strength due to an anticorrelation<sup>11</sup> between both parameters (e.g. Dopita et al. 2016; Morisset et al. 2016; Pilyugin & Grebel 2016). If so, the observed trend implies a gradation with morphology in the stellar metallicity (in particular, the oxygen abundance).

### 4.4 H II regions along galactocentric distances

In the previous section, we explore the distribution of the H II regions segregated by the morphological type of their host galaxies across the diagnostic diagrams. In this section, we explore the distribution of the H II regions across the BPT for different galactocentric distances. The distance was directly derived from the H II region catalogue, which comprises the physical properties of the region by their location with respect to the centre of the galaxies, in addition

<sup>11</sup>Recent studies found that in some galaxies there exists a positive correlation between both parameters (e.g. Poetrodjojo et al. 2018; Thomas et al. 2018).



**Figure 7.**  $[\text{O III}]\lambda 5007/\text{H}\beta$  versus  $[\text{N II}]\lambda 6583/\text{H}\alpha$  diagram for the final sample of decontaminated H II regions. The panels (from top-left to bottom-right) comprise the data for E/S0, Sa, Sb, Sbc, Sc, and Sd/Sm/Irr galaxies. In each panel, the colour code represents the mean  $\text{EW}(\text{H}\alpha)$  of H II regions. The density contours are similar to those presented in Fig. 2. Also, the demarcation curve proposed in Section 4.2 is shown as a solid line.

**Table 1.** H II regions by morphology. (1) Morphology of the host galaxies. (2) Number of galaxies. (3) Number of galaxies with detected H II regions. (4) Fraction of galaxies with H II regions detected. (5) Number of H II regions per morphological type of host galaxy. (6) Fraction of H II regions. (7) Average number of H II regions per galaxy.

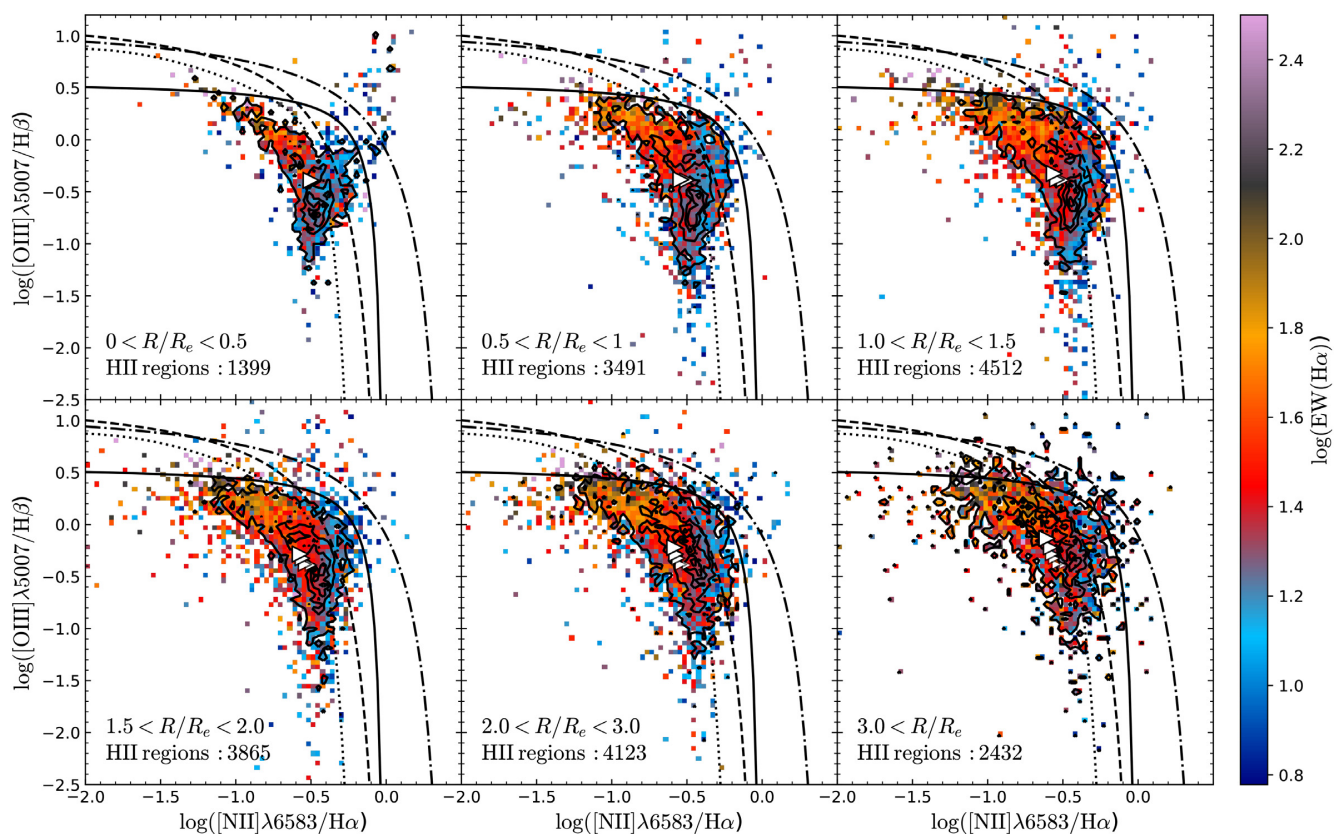
Morph.	$N_{\text{gal}_T}$	$N_{\text{gal}}$	$\%_{\text{gal}}$	$N_{\text{H II}}$	$\%_{\text{H II}}$	$\langle N_{\text{H II}} \rangle$
E	163	44	4.8	190	0.7	4.3
S0	105	46	5.0	415	1.5	9.0
Sa	135	121	13.1	2572	9.5	21.3
Sb	137	132	14.3	5598	20.6	42.4
Sbc	113	112	12.1	6274	23.1	56.0
Sc	172	172	18.6	8917	32.9	51.8
Sd	65	65	7.0	2259	8.3	34.8
Sm	21	21	2.3	633	2.3	30.1
Irr	10	10	1.1	214	0.8	21.4

to many other parameters. Based on these coordinates, we derive the galactocentric distance, correcting them by the inclination of each galaxy and position angle, and normalizing by the effective radius.

Fig. 8 shows the distribution of H II regions across the BPT diagram segregated into different galactocentric distances, ranging from the smallest to the largest distances. The distribution changes clearly with the distance to the centre of the galaxy. Those regions towards the centre ( $R/R_e < 0.5$ ) are mostly located at the right end of the BPT diagram, at a very similar location where early-spirals H II regions are found (Fig. 7). On the other hand, the regions at larger distances ( $R/R_e > 0.5$ ) follow the classical trend for these objects

(e.g. Osterbrock 1989), with a shift towards the upper left range of the distribution as the galactocentric distances increase. This result indicates that the ionization conditions of the H II regions depend on its location in the galaxy. This result is observed with the mean line ratios (white triangles) at different galactocentric distances. Earlier results already found that the classical H II regions are those within the disc of spiral galaxies (e.g. Diaz 1989). However, towards the centre of galaxies, they present stronger low-ionization forbidden emission lines (e.g. Kennicutt et al. 1989; Ho et al. 1997), shifting them towards the right-hand range of the diagram close to the demarcation curve (e.g. Sánchez et al. 2012a). These studies suggested that this may be due to the contamination by an additional





**Figure 8.**  $[\text{O III}]\lambda 5007/\text{H}\beta$  versus  $[\text{N II}]\lambda 6583/\text{H}\alpha$  diagram for the H II regions. The sample was divided into six groups according to their galactocentric distance, normalized to the effective radius: (i)  $R/R_e > 0.5$  (top left hand); (ii)  $0.5 < R/R_e < 1$  (top centre panel); (iii)  $1 < R/R_e < 1.5$  (top right hand); (iv)  $1.5 < R/R_e < 2$  (bottom left hand); (v)  $2 < R/R_e < 3$  (bottom centre panel); and (vi)  $R/R_e > 3$  (bottom right hand). The density contours are similar to those presented in Fig. 2. The colour code indicates the mean  $\text{EW}(\text{H}\alpha)$  value in the respective bin. The white triangles in each panel represent the mean value of the  $[\text{O III}]/\text{H}\beta$  and  $[\text{N II}]/\text{H}\alpha$  ratios. Also, the previous values of each panel are shown in the next panels as transparent white triangles. All lines represent the same demarcation curves described in Fig. 5. Also, the demarcation curve proposed in Section 4.2 is shown as a solid line.

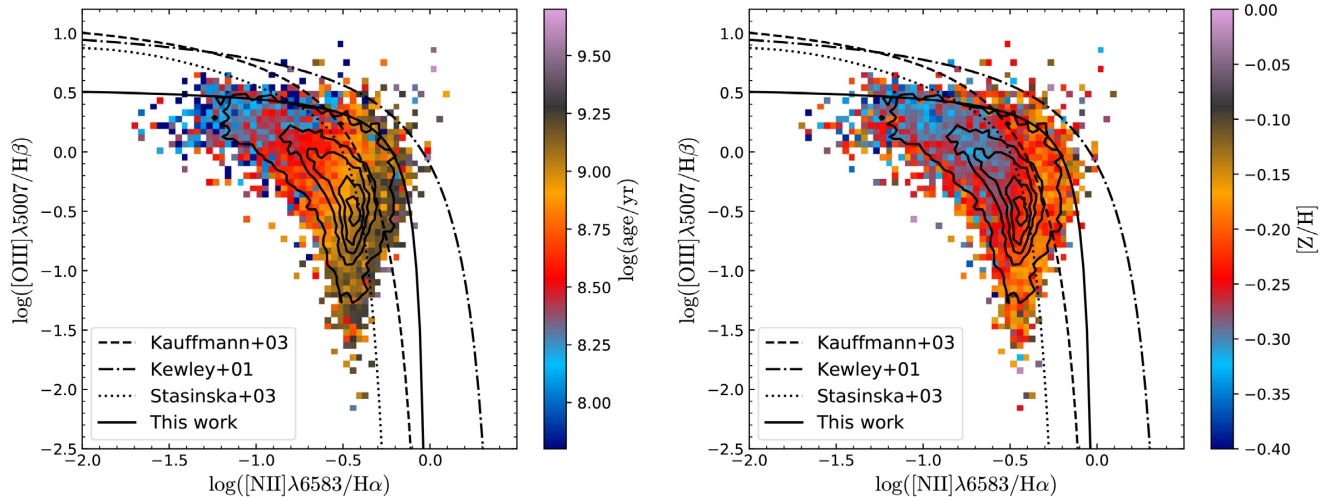
source ionization, as diffuse emission or shocks. However, the data in Fig. 8 are already corrected by the contribution of DIG and the distribution of H II regions remains at the same location in the BPT diagram. In general, we confirm the results by Sánchez et al. (2015b) in this regard, where the observed trends were already reported using a more limited number of H II regions (over a more limited number of galaxies) and without performing a DIG correction.

A plausible explanation for the enhanced low-ionization forbidden line ratios in central H II regions could be the contamination by shocks, as suggested by Ho et al. (1997). The presence of shock contamination in these regions was first suggested by Stauffer (1981) and explored to explain the differences in the disc and central H II regions by other authors (e.g. Peimbert, Torres-Peimbert & Ruiz 1992). Other authors do not report a clear difference between regions found at different galactocentric distances (e.g. Veilleux & Osterbrock 1987). Indeed, although there is a clear trend towards the right-hand regime of the BPT diagram for H II regions in the centre of galaxies, not all central H II regions are located at that extreme end. Therefore, if shocks are the source of the extra ionization, it is not present in all H II regions. A plausible explanation is that remnants of recent supernova explosions associated with the star formation events that ionized the observed region could contaminate the line ratios, deviating them from the expected location due to pure photoionization by OB stars. Under this scenario, our results indicate that central H II regions are more

prompt to this contamination. We will explore this possibility in forthcoming studies. As a first approach, models with shock (like MAPPINGS models; Sutherland & Dopita 2017) could be used to study the contribution of shocks, using, for example, the grid recently published by Alarie & Morisset (2019).

#### 4.5 H II regions and the underlying stellar population

In the previous sections, we confirmed the results by Sánchez et al. (2015b), indicating that the location of an H II region within the BPT diagram is associated with the galactocentric distance and the morphology of the host galaxy. These authors connected the reported trends with the properties of the underlying stellar population (i.e. not the ionizing population) at the location (and host) of the H II regions. It is known that the bulk of the stellar population of the centre of disc galaxies presents similarities with that of early-type galaxies (e.g. González Delgado et al. 2014; Goddard et al. 2017; Sánchez et al. 2018). Moreover, they present radial gradients in both the stellar ages and metallicities (e.g. González Delgado et al. 2014), the sSFR (e.g. González Delgado et al. 2016; Sánchez et al. 2018) and the gas content (e.g. Utomo et al. 2017; Sánchez et al. 2018), and the overall star formation histories (García-Benito et al. 2017). Thus, the ionization conditions seem to be connected somehow with the local stellar evolution within the considered galaxy.



**Figure 9.**  $[\text{O III}]\lambda 5007/\text{H}\beta$  versus  $[\text{N II}]\lambda 6583/\text{H}\alpha$  diagram for the H II regions in our final catalogue. *Left-hand panel:* the colour indicates the average value of the luminosity-weighted age of the underlying stellar population. *Right-hand panel:* the colour indicates the average value of the luminosity-weighted metallicity of the underlying stellar population. The density contours are similar to those presented in Fig. 2. In both the panels, the dotted, dot-dashed, dashed, and solid lines are the demarcation curves used in Fig. 6.

Fig. 9 shows the distribution of H II regions across the BPT diagram with a colour code representing properties of underlying stellar populations. The left-hand panel shows the luminosity-weighted ages and the right-hand panel shows the metallicity of the underlying stellar population. The properties of the underlying stellar populations at the location of each H II region were extracted from the data products provided by PIPE3D, as described in Section 3.1, following Sánchez et al. (2013, 2014). As can be seen in Fig. 9, the location of an H II region in the BPT diagram depends on both stellar parameters. The regions with a younger underlying stellar population are located in the upper left area of the diagram. Otherwise, the regions with an older underlying stellar population are located in the right end of the diagram.

Similarly, the regions with metal-poor underlying stellar populations are found in the upper left area, while the regions with metal-rich underlying stellar populations are mostly located at the right end of the diagram. It is important to remark that we are not talking about properties of the ionizing population. The reported ages and metallicities correspond to the total stellar population at the location of each H II region. This result is consistent with the one obtained by Sánchez et al. (2015b), where similar trends were reported.

The trends along the BPT diagram with morphology and galactocentric distance reported in the previous sections are naturally explained by the distributions shown in Fig. 9. Early-type galaxies (and the centre of all galaxies) are dominated by old and metal-rich stellar populations (González Delgado et al. 2014; Sánchez et al. 2015a; Goddard et al. 2017). In general, and therefore, the H II regions on those galaxies are located at the right end of the distribution for these objects. On the contrary, late-type galaxies (and the outer areas of all galaxies) are mostly dominated by younger and more metal-poor stellar populations. Therefore, the H II regions on those galaxies are located in the left end of the distribution.

#### 4.6 Imprints of the galactic evolution in the ionization

As described in the previous section, the location in the diagnostic diagrams depends on the underlying stellar population properties (age and metallicity). Thus, the BPT diagram exhibits a connection

between the stellar population properties and the line ratios. We explore this connection in here in a more quantitative way.

Fig. 10 shows the distribution of  $[\text{N II}]/\text{H}\alpha$  line ratio (left-hand panel) and the  $[\text{O III}]/\text{H}\beta$  line ratio (right-hand panel) along the luminosity-weighted age of the corresponding underlying stellar population at the location of each H II region. The colour code represents the stellar metallicity. As can be seen in the figure, there exists a trend between the line ratios and the stellar properties. Indeed, the distributions present a correlation coefficient of  $r = 0.50$  (0.63) for low metallicities,  $[Z/H] \sim -0.55$  dex, and  $r = 0.46$  (0.48) for high metallicities,  $[Z/H] \sim 0.15$  dex, for the  $[\text{N II}]/\text{H}\alpha$  ( $[\text{O III}]/\text{H}\beta$ ) line ratio. In general, the older and more metal-rich underlying stellar population is related to a higher (lower) value of the  $[\text{N II}]/\text{H}\alpha$  ( $[\text{O III}]/\text{H}\beta$ ) line ratio.

In order to illustrate more clearly the correlation described above, we averaged the values in bins of 0.1 dex in  $[Z/H]$  and 0.25 dex in  $\log(\text{age})$ . The results are shown in Fig. 10, where every coloured circle is the mean value of the corresponding line ratio and age in each bin for a fixed metallicity (coded by the colour). The relation of the line ratios on both parameters is clear when explored the binned data.

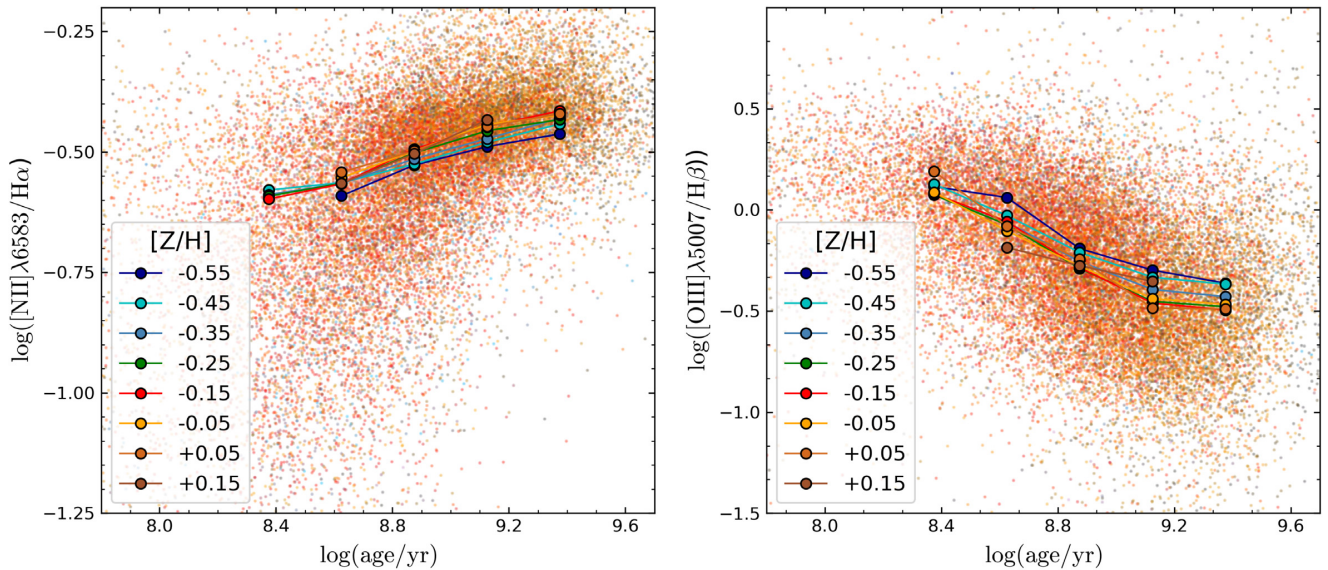
To characterize the reported trends, we fit both line ratios with a linear combination of both properties of the stellar population properties (age and  $[Z/H]$ ). We consider that with the current dispersion adopting a more complex functional form would not provide a better characterization. The best linear regression fitting for the  $[\text{N II}]/\text{H}\alpha$  line ratio was

$$\log\left(\frac{[\text{N II}]}{\text{H}\alpha}\right) = 0.20_{\pm 0.02} \log(\text{age}/\text{yr}) + 0.06_{\pm 0.02} [Z/H] - 2.32_{\pm 0.26}. \quad (9)$$

The standard deviation of the residual of the line ratio, once we subtracted the best-fitting model, is  $\sim 0.166$  dex. Compared with the standard deviation of the original distribution of values ( $\sim 0.194$  dex), the standard deviation was reduced<sup>12</sup> by a factor of  $\sim 52$  per cent. Therefore, the  $[\text{N II}]/\text{H}\alpha$  largely depends on the un-

<sup>12</sup>The difference was calculated by quadratic difference.





**Figure 10.** Distribution of the  $[\text{N II}]\lambda 6583/\text{H}\alpha$  line ratio (left-hand panel) and  $[\text{O III}]\lambda 5007/\text{H}\beta$  (right-hand panel) along the luminosity-weighted age of the underlying stellar population for the H II regions in our catalogue, colour coded by the  $[\text{Z}/\text{H}]$  of the underlying stellar population. The solid circles represent the average line ratios within bins of 0.25 dex in  $\log(\text{age}/\text{yr})$  for a fixed range of metallicities (binned in ranges of 0.1 dex). The average  $[\text{Z}/\text{H}]$  corresponding to each colour is included in the inset of each panel. All circles within the same metallicity range are linked by a solid line and plotted with the same colour (which codifies  $[\text{Z}/\text{H}]$ ).

derlying stellar population. For  $[\text{O III}]/\text{H}\beta$ , the best linear regression fitting with the properties of the underlying stellar population is

$$\log\left(\frac{[\text{O III}]}{\text{H}\beta}\right) = -0.58_{\pm 0.01} \log(\text{age}/\text{yr}) - 0.17_{\pm 0.02} [\text{Z}/\text{H}] + 4.94_{\pm 0.12}. \quad (10)$$

The residual between the line ratio observed and the best fit presents a dispersion of  $\sim 0.348$  dex. Again, compared with the initial dispersion of  $\sim 0.382$  dex, the standard deviation was decreased by a factor of  $\sim 41$  per cent. Like in the case of the previous line ratio, this modelling clearly illustrates the reported relation. However, despite the reduction of the dispersion, the line ratios still present a wide range of values. Other properties influence the observed line ratios besides the underlying stellar properties. In particular, the physical conditions of the nebulae (i.e. electronic density, ionization parameter, metallicity, the geometry of the ionized gas, etc.) must indeed influence the line ratios.

## 5 DISCUSSION AND CONCLUSIONS

In this study, we presented a new catalogue of H II regions extracted from the eCALIFA (+ pCALIFA) + PISCO sample. This new catalogue comprises information of  $\sim 28\,000$  H II regions, including the flux of the most important emission lines in the optical range (between 3745 and 7200 Å) and the corresponding properties of the underlying stellar population. It is important to remark that we based the selection of H II regions on basically two assumptions: (i) they should appear as a clumpy structure in the  $\text{H}\alpha$  emission line map (see Section 3.2) and (ii) the underlying stellar populations should be compatible with the definition of H II region (see Section 3.3).

However, as can be seen in Fig. 1, the identification algorithm is not perfect. We are missing small clumpy regions and, also, we cannot resolve those regions that are too close one from each other. This is a direct consequence of the spatial resolution of the CALIFA data. As noticed by Mast et al. (2014), with a low spatial

resolution, the smallest ionized regions cannot be segregated and they are erroneously assigned to bigger adjacent ones or confused with the diffuse. This is why almost all identified ionized regions have a similar size in our catalogue. Thus, an important fraction of small ionized regions is lost in the final sample. This has a direct effect on the calculation of some properties, like the luminosity function (e.g. Sánchez et al. 2012b), that cannot be well recovered in the low-luminosity regime.

Furthermore, the wings of the PSF limit our ability to segregate among the different ionizing sources. First, they affect our ability to segregate between adjacent H II regions, as the boundaries of each clumpy ionized region are not well defined at the current coarse resolution. Therefore, the segregation algorithm developed to identify the ionized regions could erroneously associate pixels that contain certain emission coming from the PSF wing of any adjacent (brighter) region to a fainter one, modifying the estimated values of the derived parameters. Secondly, due to these wings, the DIG is polluted by the emission of any adjacent H II region too. Finally, H II regions could be contaminated by the emission of other possible ionizing sources (not only DIG but shocks or AGNs) as well. This may somehow affect the physical properties calculated from the emission line fluxes (Mast et al. 2014; Zhang et al. 2017).

To minimize the effects of this contamination, we present a new approach to correct the emission line fluxes of the H II regions by the contribution of the DIG. As we mention in Section 3.4, the DIG emission was calculated by considering all the spaxels in between those selected as being part of an H II region by our code. Again, spatial resolution plays an essential role in the proposed DIG correction. If we could identify exactly the limits of each ionized region (or model its shape), we would be able to identify clearly the DIG and, hence, make a proper correction irrespectively of the source of ionization of this diffuse (hot, old stars and leaking photons). Unfortunately, as indicated before and discussed in Section 3.4.3, we have an important contamination by ionization from the H II regions themselves due to the coarse resolution and

relatively large PSF wings. This contribution does not allow us to distinguish the DIG powered by leaking photons, for instance: The physical properties of this contamination, like  $\text{EW}(\text{H}\alpha)$  and line ratios, are too similar to that of adjacent H II regions. Hence, we have corrected only the DIG powered by old, hot stars, which indeed is the one that may produce the strongest effects in the line ratios. This result is supported by the study of Vale Asari et al. (2019). They proposed a DIG correction where its contribution is calculated using only SF galaxies from the MaNGA data set, i.e. a DIG of higher  $\text{EW}(\text{H}\alpha)$ . There is a small but not significant change between the metallicities calculated by the corrected emission lines and the uncorrected ones for most of the explored calibrators (besides N2, which presents a 0.1 dex offset). Besides, the physical spatial resolution of the data analysed in Vale Asari et al. (2019) is nearly twice worse than the one included here (the FWHM  $\sim 0.8$  kpc for CALIFA and  $\sim 2.5$  kpc for MaNGA), which makes any correction by the DIG more complicated and with a stronger expected effect. To compare the contribution of all DIG types, we required observations with a high spatial resolution. However, the DIG correction present here, although not perfect, is an overall improvement with respect to not applying any correction at all.

It is clear that better spatial resolutions will facilitate the segregation of the clumpy ionized regions. It will improve the detection of smaller (thus, fainter) H II regions, the separation between adjacent regions, the definition of the DIG (and its different sources), and the overall decontamination of the DIG. New generation data sets, like the ones provided by MUSE (e.g. Sánchez-Menguiano et al. 2018) and SITELLE (e.g. Rousseau-Nepton et al. 2018), would require a re-evaluation of the proposed procedure, both for the detection of the H II regions and for the decontamination by the DIG.

A good example of the capabilities of new instruments is the study by Sánchez-Menguiano et al. (2018). They used MUSE observations of 102 galaxies, selected from the *All-weather MUSE Supernova Integral-field Nearby Galaxies* survey (AMUSING; Galbany et al. 2016; López-Cobá, submitted), to analyse the shape of oxygen abundance profiles. For doing so, they obtained a catalogue of H II regions using HIIEXPLORER. The total number of regions detected was 14 345. The spatial resolution of their data, FWHM  $\sim 0.8$  arcsec (corresponding to the natural seeing) compared to the current value of 2.5 arcsec for CALIFA, allows them to identify a number equivalent to  $\sim 65$  per cent of our current catalogue in a number of galaxies that comprises just  $\sim 10$  per cent of the ones analysed here.

On the other hand, SITELLE provides observations where the natural seeing also limits resolution. Using this instrument, Rousseau-Nepton et al. (2018) identified 4285 H II region candidates in a single galaxy (NGC 628). However, the methodology used for the identification of H II regions generates several spurious ionized regions. In comparison, using the PINGS data (of similar projected resolution than CALIFA), and HIIEXPLORER, Sánchez et al. (2012b) recovered 373 H II regions in the same galaxy. Using the same data, we recovered 243 H II regions (from 244 ionized regions detected). The difference between the number of recovered H II regions by Sánchez et al. (2012b) and us is due to the input parameters used in HIIEXPLORER and PYHIIEXPLORER. In Sánchez et al. (2012b), the input parameters are listed.

In comparison, their minimum flux is lower than the value used in our calculations. However, they report that only 282 are H II regions with a good-quality spectrum and good estimation of the  $\text{H}\beta$  flux. Meanwhile, the 243 H II regions that we report in here are the final number after applying the quality check test described in the text. Taking into account these details, we consider that the number of H II

regions found in both the studies is consistent and the differences are clearly justified. Thus, the use of SITELLE represents a theoretical improvement of a factor of 10 in the detection of these ionized sources.

Despite the limitations due to the coarse spatial resolution, there are still significant advantages in the current data set. MUSE does not cover the blue wavelength range of the optical spectrum below  $\lambda < 4650 \text{ \AA}$ . Therefore, fundamental emission lines are missing in the data provided by that instrument for galaxies at low redshifts. Moreover, the information provided for the stellar populations is limited since the critical wavelength range around the 4000  $\text{\AA}$  break is not covered. With SITELLE, the information for the stellar population is even less reliable due to the instrumental technique that affects the features of the continuum strongly. Finally, CALIFA allows us to present the current exploration for a statistically significant sample of galaxies, representative of the population at the nearby universe. Thus, the current H II region catalogue is not (a priori) biased by morphology, mass, or other galaxy properties. However, it is recognized that the possible properties derived from weak emission lines could drive to bias results, e.g. electron temperature derived from  $[\text{O III}]\lambda 4363$ . This is a direct effect of the inherent weakness of these emission lines and/or possible contamination from other sources (e.g. sky emission lines). However, the properties derived for weak emission lines could be calculated well enough from the H II regions.

We study the relationship between the location on the BPT diagram of the distribution of DIG-corrected H II regions and (i) the morphology of the host galaxies; (ii) the galactocentric distances of the H II regions; and (iii) the properties of the underlying stellar population. First, we demonstrate that the distribution of H II regions based on our simple selection criteria is located in the SF area of the diagnostic diagrams (without this being a prerequisite or part of the selection process). We found that the location in the BPT diagram depends strongly on the value of the  $\text{EW}(\text{H}\alpha)$ . This result was already reported in previous studies (e.g. Sánchez et al. 2015a; Lacerda et al. 2018). However, in comparison with them, we demonstrate that the result still holds when using a larger statistical sample of H II regions covering a more extensive range of physical properties and ionization conditions.

We clearly showed that the location in the diagnostic diagram depends on the morphological type of the host galaxy and the galactocentric distance. Early-type galaxies, in general, massive and metal rich, have H II regions located in the lower right end of the diagnostic diagram. Contrarily, in late-type galaxies, less massive and metal poor, the distribution moves towards the upper left zone of the diagrams. Thus, the change of location in the diagram is a consequence of the change of the stellar population properties along the Hubble sequence. González Delgado et al. (2015) already noticed that the metallicity, age, and other properties of stellar populations depend on the morphological type of galaxies. Indeed, they proved that the average stellar metallicity increases with the morphology.

Similarly, we found that the distribution in the BPT diagrams depends on the galactocentric distance. The most straightforward explanation is the same as the one described for the differences reported by morphological type. The centre of galaxies is populated by older and more metal-rich stellar populations, while the outer regions are populated by younger and more metal-poor ones (e.g. González Delgado et al. 2015; Sánchez et al. in press). This radial gradient is also evident in the oxygen abundance, which decreases with the galactocentric distances, too (e.g. Sánchez et al. 2012b, 2014).

This result is related to the inside-out scenario for the formation of galaxies. Other studies (e.g. Pérez et al. 2013; Ibarra-Medel et al. 2016; García-Benito et al. 2017) demonstrated the radial dependence on the star formation history and the time-scales of the gas infall. Therefore, the chemical evolution of the ISM is affected by the star formation history at different distances resulting in the metallicity gradients on galaxies.

All the discussed results confirm that the locations within the BPT diagrams are tightly related to the properties of the underlying population, as already suggested by Sánchez et al. (2015a). This is clearly shown in Fig. 9. Moreover, we demonstrate that the line ratios  $[\text{N II}]\lambda 6583/\text{H}\alpha$  and  $[\text{O III}]\lambda 5007/\text{H}\beta$  present a clear correlation with both the ages and metallicities of the underlying stellar populations (Fig. 10). This relation is stronger for the  $[\text{N II}]\lambda 6583/\text{H}\alpha$  line ratio. This result is somehow expected since this line ratio is widely used as a proxy of the oxygen abundance. The previous generation of stars enriched the chemical composition, oxygen, and other element abundances in the ionized regions. Hence, the oxygen abundance of the ISM naturally correlates with the metallicity of the younger stars, provided that the radial mixing is not strong (Sánchez et al. 2015a; Sánchez-Menguiano et al. 2016).

On the other hand, there is a correlation with the age of the stellar population due to the chemical enrichment processes in galaxies. It is well known that, in disc-dominated galaxies, stars are older and more metal rich in the central regions than the stars in the outer regions of the disc (Pérez et al. 2013; Sánchez-Blázquez et al. 2014; González Delgado et al. 2015). Thus, metallicity and its relation with the differential star formation history within galaxies and at different locations in them (e.g. García-Benito et al. 2017) are the most plausible explanation for the observed trends.

The other line ratio,  $[\text{O III}]\lambda 5007/\text{H}\beta$ , also shows a correlation with both analysed stellar parameters (age and  $[Z/H]$ ). However, since this ratio is less sensitive to oxygen abundance and more sensitive to variations in the geometry, density, and temperature of the ionized nebulae, it presents a weaker correlation than that found for  $[\text{N II}]\lambda 6583/\text{H}\alpha$ .

Although other authors already noticed some of the results discussed above, we present them, for the first time, for a statistically well-defined sample of H II regions in the nearby Universe, several times larger than those analysed previously. We demonstrate that those results do not depend strongly on the possible contamination by DIG since we have applied a correction to that contribution. Moreover, we present a new demarcation line defined with this sample that is consistent with well-known previous ones, although it is totally unbiased regarding the involved line ratios.

The current catalogue of H II regions and the full set of emission lines discussed in this article are publicly available. To our knowledge, this is the largest catalogue of H II region properties, derived from the broadest range of galaxies of different morphologies and masses. With the analysis presented in this study, we have demonstrated that this sample is suitable to perform further explorations on the physical properties of these nebulae (in particular, for those that do not require detailed knowledge of the size and absolute luminosity). In forthcoming studies, we will present the main patterns of physical properties for these H II regions.

## ACKNOWLEDGEMENTS

We thank the referee for the helpful comments that have improved the quality of this article. We are grateful for the support of CONA-

CYT grants CB-285080, FC-2016-01-1916, and CB2015-254132, and funding from the DGAPA/PAPIIT-IA101217, DGAPA/PAPIIT-IN100519, and DGAPA/PAPIIT-107215 (UNAM) projects.

This study uses data provided by the CALIFA survey (<http://califa.caha.es>); it is based on observations collected at the Centro Astronómico Hispano Alemán (CAHA) at Calar Alto, operated jointly by the Max-Planck-Institut für Astronomie and the Instituto de Astrofísica de Andalucía (CSIC).

LG was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 839090.

This research makes use of python (<http://www.python.org>), of Matplotlib (a framework to the creation of scientific plots for Python; Hunter 2007), Astropy (a Python package for astronomy; Astropy Collaboration 2013; Price-Whelan et al. 2018), and SciPy (a Python-based ecosystem of open-source software for mathematics, science, and engineering; Jones et al. 2001).

## REFERENCES

- Alarie A., Morisset C., 2019, *Rev. Mex. Astron. Astrofis.*, 55, 377  
 Anderson L. D., 2014, *Am. Astron. Soc. Meeting Abstr.*, 223, 312.01  
 Astropy Collaboration, 2013, *A&A*, 558, A33  
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5 (BPT)  
 Barrera-Ballesteros J. K. et al., 2016, *MNRAS*, 463, 2513  
 Barrera-Ballesteros J. K. et al., 2018, *ApJ*, 852, 74  
 Barrera-Ballesteros J. K., Sánchez S. F., Heckman T., Blanc G. A., The MaNGA Team, 2017, *ApJ*, 844, 80  
 Bertin E., Arnouts S., 1996, *A&AS*, 117, 393  
 Binette L., Magris C. G., Stasińska G., Bruzual A. G., 1994, *A&A*, 292, 13  
 Binette L., Flores-Fajardo N., Raga A. C., Drissen L., Morisset C., 2009, *ApJ*, 695, 552  
 Bitsakis T. et al., 2019, *MNRAS*, 483, 370  
 Blanton M. R., Moustakas J., 2009, *ARA&A*, 47, 159  
 Bresolin F., 2019, *MNRAS*, 488, 3826  
 Bundy K. et al., 2015, *ApJ*, 798, 7  
 Cairós L. M., Caon N., García Lorenzo B., Kelz A., Roth M., Papaderos P., Streicher O., 2012, *A&A*, 547, A24  
 Cappellari M., Copin Y., 2003, *MNRAS*, 342, 345  
 Cid Fernandes R. et al., 2013, *A&A*, 557, A86  
 Cid Fernandes R. et al., 2014, *A&A*, 561, A130  
 Cid Fernandes R., Stasińska G., Schlickmann M. S., Mateus A., Vale Asari N., Schoenell W., Sodré L., 2010, *MNRAS*, 403, 1036  
 Croom S. M. et al., 2012, *MNRAS*, 421, 872  
 Davies R. L. et al., 2016, *MNRAS*, 462, 1616  
 Diaz A. I., 1989, in Beckman J. E., Pagel B. E. J., eds, *Evolutionary Phenomena in Galaxies*. Cambridge Univ. Press, Cambridge, p. 377  
 Dopita M. A., Sutherland R. S., 1996, *ApJS*, 102, 161  
 Dopita M. A., Kewley L. J., Heisler C. A., Sutherland R. S., 2000, *ApJ*, 542, 224  
 Dopita M. A., Rich J., Vogt F. P. A., Kewley L. J., Ho I.-T., Basurah H. M., Ali A., Amer M. A., 2014, *Ap&SS*, 350, 741  
 Dopita M. A., Kewley L. J., Sutherland R. S., Nicholls D. C., 2016, *Ap&SS*, 361, 61  
 Dottori H. A., 1987, *Rev. Mex. Astron. Astrofis.*, 14, 463  
 Dottori H. A., Copetti M. V. F., 1989, *Rev. Mex. Astron. Astrofis.*, 18, 115  
 Ferguson A. M. N., Wyse R. F. G., Gallagher J. S. III, Hunter D. A., 1996, *AJ*, 111, 2265  
 Flores-Fajardo N., Morisset C., Stasińska G., Binette L., 2011, *MNRAS*, 415, 2182  
 Galbany L. et al., 2016, *MNRAS*, 455, 4087  
 Galbany L. et al., 2018, *ApJ*, 855, 107  
 García-Benito R. et al., 2015, *A&A*, 576, A135  
 García-Benito R. et al., 2017, *A&A*, 608, A27  
 García-Benito R., Pérez E., Díaz Á. I., Maíz Apellániz J., Cerviño M., 2011, *AJ*, 141, 126



- Garnett D. R., 2002, *ApJ*, 581, 1019
- Goddard D. et al., 2017, *MNRAS*, 466, 4731
- Gomes J. M. et al., 2016a, *A&A*, 585, A92
- Gomes J. M. et al., 2016b, *A&A*, 588, A68
- González Delgado R. M. et al., 2014, *A&A*, 562, A47
- González Delgado R. M. et al., 2015, *A&A*, 581, A103
- González Delgado R. M. et al., 2016, *A&A*, 590, A44
- Haffner L. M. et al., 2009, *Rev. Mod. Phys.*, 81, 969
- Hodge P. W., Kennicutt R. C. Jr, 1983, *ApJ*, 267, 563
- Ho I.-T. et al., 2014, *MNRAS*, 444, 3894
- Hoopes C. G., Walterbos R. A. M., Greenwalt B. E., 1996, *AJ*, 112, 1429
- Hoopes C. G., Walterbos R. A. M., Rand R. J., 1999, *ApJ*, 522, 669
- Ho L. C., Filippenko A. V., Sargent W. L. W., 1997, *ApJ*, 487, 579
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Husemann B. et al., 2013, *A&A*, 549, A87
- Ibarra-Medel H. J. et al., 2016, *MNRAS*, 463, 2799
- Jones E. et al., 2001, SciPy: Open Source Scientific Tools for Python. <http://www.scipy.org/>
- Kauffmann G. et al., 2003, *MNRAS*, 346, 1055
- Kehrig C. et al., 2012, *A&A*, 540, A11
- Kehrig C., Vílchez J. M., Sánchez S. F., Telles E., Pérez-Montero E., Martín-Gordón D., 2008, *A&A*, 477, 813
- Kelz A. et al., 2006, *PASP*, 118, 129
- Kennicutt Robert C. J., 1998, *ARA&A*, 36, 189
- Kennicutt R. C. Jr, 1988, *ApJ*, 334, 144
- Kennicutt R. C. Jr, 1989, *ApJ*, 344, 685
- Kennicutt R. C. Jr, Keel W. C., Blaha C. A., 1989, *AJ*, 97, 1022
- Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35
- Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, 556, 121
- Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, *MNRAS*, 372, 961
- Knapen J. H., 1998, *MNRAS*, 297, 255
- Lacerda E. A. D. et al., 2018, *MNRAS*, 474, 3727
- Lequeux J., Peimbert M., Rayo J. F., Serrano A., Torres-Peimbert S., 1979, *A&A*, 80, 155
- Levy R. et al., 2019, *ApJ*, 882, 84
- López-Cobá C. et al., 2017, *ApJ*, 850, L17
- López-Cobá C., Sánchez S. F., Bland-Hawthorn J., Moiseev A. V., Cruz-González I., García-Benito R., Barrera-Ballesteros J. K., Galbany L., 2019, *MNRAS*, 482, 4032
- López-Sánchez A. R., Esteban C., 2009, *A&A*, 508, 615
- Marino R. A. et al., 2013, *A&A*, 559, A114
- Marino R. A. et al., 2016, *A&A*, 585, A47
- Mármol-Queraltó E. et al., 2011, *A&A*, 534, A8
- Martin P., Roy J.-R., 1994, *ApJ*, 424, 599
- Mast D. et al., 2014, *A&A*, 561, A129
- Méndez-Abreu J., Sánchez S. F., de Lorenzo-Cáceres A., 2019, *MNRAS*, 484, 4298
- Minter A. H., Balsa D. S., 1997, *ApJ*, 484, L133
- Minter A. H., Spangler S. R., 1997, *ApJ*, 485, 182
- Monreal-Ibero A., Arribas S., Colina L., 2006, *ApJ*, 637, 138
- Monreal-Ibero A., Arribas S., Colina L., Rodríguez-Zaurín J., Alonso-Herrero A., García-Marín M., 2010, *A&A*, 517, A28
- Morisset C., Georgiev L., 2009, *A&A*, 507, 1517
- Morisset C. et al., 2016, *A&A*, 594, A37
- Moustakas J., Kennicutt R. C. Jr, 2006, *ApJS*, 164, 81
- Oey M. S., Parker J. S., Mikles V. J., Zhang X., 2003, *AJ*, 126, 2317
- Oey M. S. et al., 2007, *ApJ*, 661, 801
- Osterbrock D. E., 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*. Univ. Sci. Books, Herndon, VA, USA
- Papaderos P. et al., 2013, *A&A*, 555, L1
- Peimbert M., Torres-Peimbert S., Ruiz M. T., 1992, *Rev. Mex. Astron. Astrofis.*, 24, 155
- Pérez E. et al., 2013, *ApJ*, 764, L1
- Pilyugin L. S., Grebel E. K., 2016, *MNRAS*, 457, 3678
- Poetrodjojo H. et al., 2018, *MNRAS*, 479, 5235
- Price-Whelan A. M. et al., 2018, *AJ*, 156, 123
- Rand R. J., 1996, *ApJ*, 462, 712
- Rand R. J., 1998, *ApJ*, 501, 137
- Relaño M., Peimbert M., Beckman J., 2002, *ApJ*, 564, 704
- Relaño M., Kennicutt R. C. Jr, Eldridge J. J., Lee J. C., Verley S., 2012, *MNRAS*, 423, 2933
- Reynolds R. J., 1971, PhD thesis. Univ. Wisconsin - Madison
- Reynolds R. J., 1991, in Bloemen H., ed., *Proc. IAU Symp. 144, The Interstellar Disk-Halo Connection in Galaxies*. Kluwer, Dordrecht, p. 67
- Rosales-Ortega F. F., Kennicutt R. C., Sánchez S. F., Díaz A. I., Pasquali A., Johnson B. D., Hao C. N., 2010, *MNRAS*, 405, 735
- Rosales-Ortega F. F., Díaz A. I., Kennicutt R. C., Sánchez S. F., 2011, *MNRAS*, 415, 2439
- Roth M. M. et al., 2005, *PASP*, 117, 620
- Rousseau-Nepton L., Robert C., Martin R. P., Drissen L., Martin T., 2018, *MNRAS*, 477, 4152
- Sánchez-Blázquez P. et al., 2014, *A&A*, 570, A6
- Sánchez-Menguiano L. et al., 2016, *A&A*, 587, A70
- Sánchez-Menguiano L. et al., 2018, *A&A*, 609, A119
- Sánchez S. F. et al., 2012a, *A&A*, 538, A8
- Sánchez S. F. et al., 2012b, *A&A*, 546, A2
- Sánchez S. F. et al., 2013, *A&A*, 554, A58
- Sánchez S. F. et al., 2014, *A&A*, 563, A49
- Sánchez S. F. et al., 2015a, *A&A*, 573, A105
- Sánchez S. F. et al., 2015b, *A&A*, 574, A47
- Sánchez S. F. et al., 2016a, *Rev. Mex. Astron. Astrofis.*, 52, 21
- Sánchez S. F. et al., 2016b, *Rev. Mex. Astron. Astrofis.*, 52, 171
- Sánchez S. F. et al., 2016c, *A&A*, 594, A36
- Sánchez S. F. et al., 2017, *MNRAS*, 469, 2121
- Sánchez S. F. et al., 2018, *Rev. Mex. Astron. Astrofis.*, 54, 217
- Sánchez S. F. et al., 2019, *MNRAS*, 484, 3042
- Sánchez S. F., 2019, preprint ([arXiv:1911.06925](https://arxiv.org/abs/1911.06925))
- Sánchez S. F., Cardiel N., Verheijen M. A. W., Martín-Gordón D., Vilchez J. M., Alves J., 2007, *A&A*, 465, 207
- Sarzi M. et al., 2010, *MNRAS*, 402, 2187
- Singh R. et al., 2013, *A&A*, 558, A43
- Skillman E. D., Kennicutt R. C., Hodge P. W., 1989, *ApJ*, 347, 875
- Stasińska G. et al., 2008, *MNRAS*, 391, L29
- Stasińska G., 2012, in Manchado A., Stanghellini L., Schönberner D., eds, *Proc. IAU Symp. 283, Planetary Nebulae: An Eye to the Future*. Kluwer, Dordrecht, p. 239
- Stasińska G., 2017, in Morisset C., Delgado-Inglada G., García-Rojas J., eds, *Cosmic Feast of the Elements*. Puebla, Mexico, p. 37
- Stasińska G., Izotov Y., 2003, *A&A*, 397, 71
- Stasińska G., Cid Fernandes R., Mateus A., Sodré L., Asari N. V., 2006, *MNRAS*, 371, 972
- Stauffer J. R., 1981, PhD thesis. Univ. California, Berkeley
- Sutherland R. S., Dopita M. A., 2017, *ApJS*, 229, 34
- Thomas A. D., Dopita M. A., Kewley L. J., Groves B. A., Sutherland R. S., Hopkins A. M., Blanc G. A., 2018, *ApJ*, 856, 89
- Tremonti C. A. et al., 2004, *ApJ*, 613, 898
- Utomo D. et al., 2017, *ApJ*, 849, 26
- Vale Asari N., Couto G. S., Cid Fernandes R., Stasińska G., de Amorim A. L., Ruschel-Dutra D., Werle A., Florido T. Z., 2019, *MNRAS*, 489, 4721
- Veilleux S., Osterbrock D. E., 1987, *ApJS*, 63, 295
- Veilleux S., Cecil G., Bland-Hawthorn J., 2005, *ARA&A*, 43, 769
- Vila-Costas M. B., Edmunds M. G., 1992, *MNRAS*, 259, 121
- Walcher C. J. et al., 2014, *A&A*, 569, A1
- York D. G. et al., 2000, *AJ*, 120, 1579
- Zaritsky D., Kennicutt R. C. Jr, Huchra J. P., 1994, *ApJ*, 420, 87
- Zhang K. et al., 2017, *MNRAS*, 466, 3217
- Zurita A., Rozas M., Beckman J. E., 2000, *A&A*, 363, 9
- Zurita A., Beckman J. E., Rozas M., Ryder S., 2002, *A&A*, 386, 801

## APPENDIX A: DEMARCATION LINES AND EFFECTS OF DIG DECONTAMINATION ON OTHER DIAGNOSTIC DIAGRAMS

In Section 4.2, we present new demarcation lines for H II regions for the most frequently used diagnostic diagrams: equations (6), (7), and (8). The upper panels on Figs A1 and A2 show the  $[\text{O III}]\lambda 5006/\text{H}\beta$  versus  $[\text{S II}]\lambda 6716 + 30/\text{H}\alpha$  and  $[\text{O III}]\lambda 5006/\text{H}\beta$  versus  $[\text{O I}]\lambda 6300/\text{H}\alpha$  diagnostic diagrams for the DIG uncorrected and corrected H II region sample, respectively (see Fig. 6). The most frequently used demarcation curves (Kewley et al. 2006) for each diagram, together with our proposed ones, are plotted too.

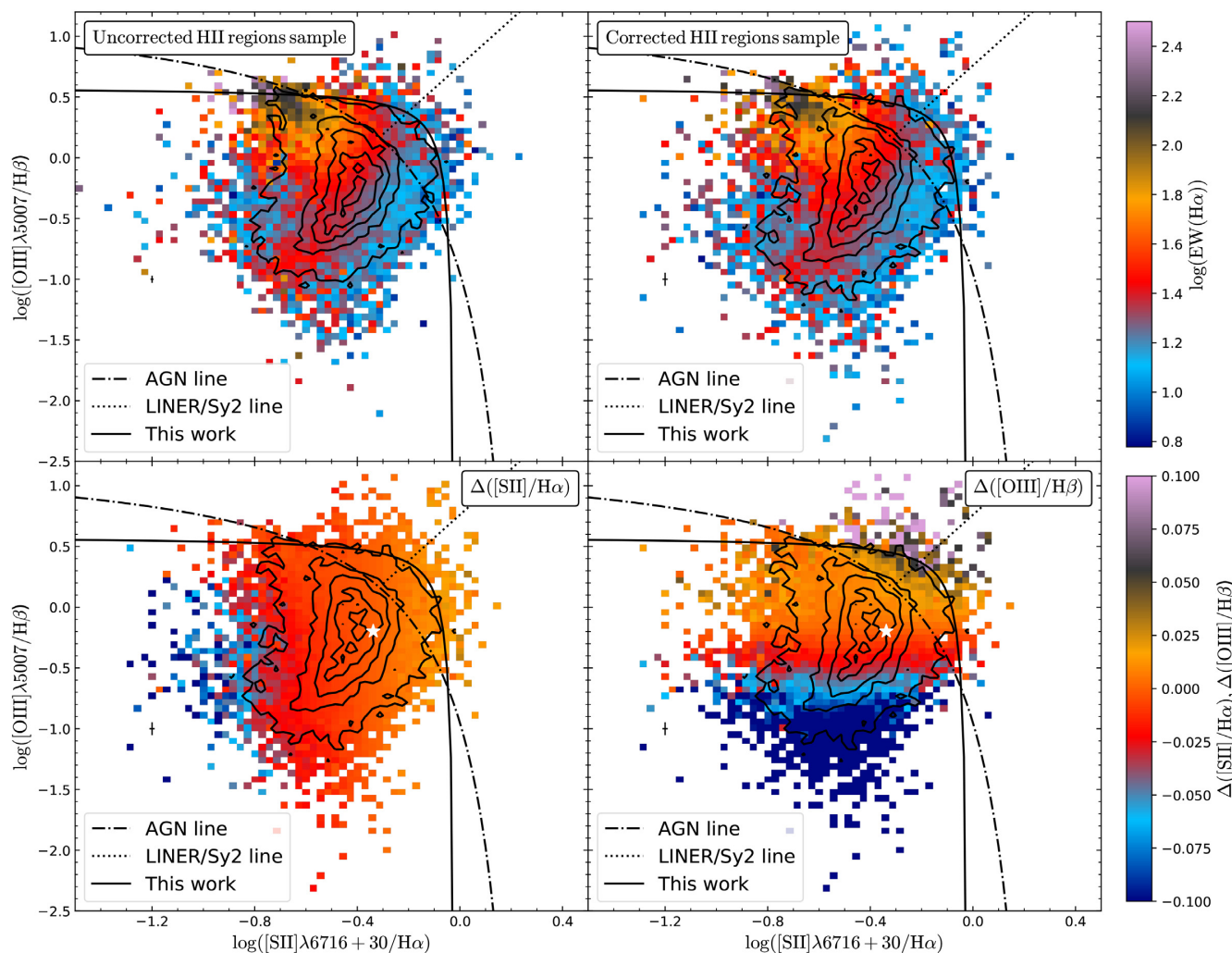
The distribution of regions in the different diagnostic diagrams corresponds to the expected one from previously published results. However, we should note that our selection did not use any of those line ratios, but a rather simple set of assumptions involving the shape of the region and the stellar population properties of the ionized regions (Section 3.3).

Despite the overall agreement with previous selections, there are some differences. In the  $[\text{O III}]$  versus  $[\text{O I}]$  diagram, we observe

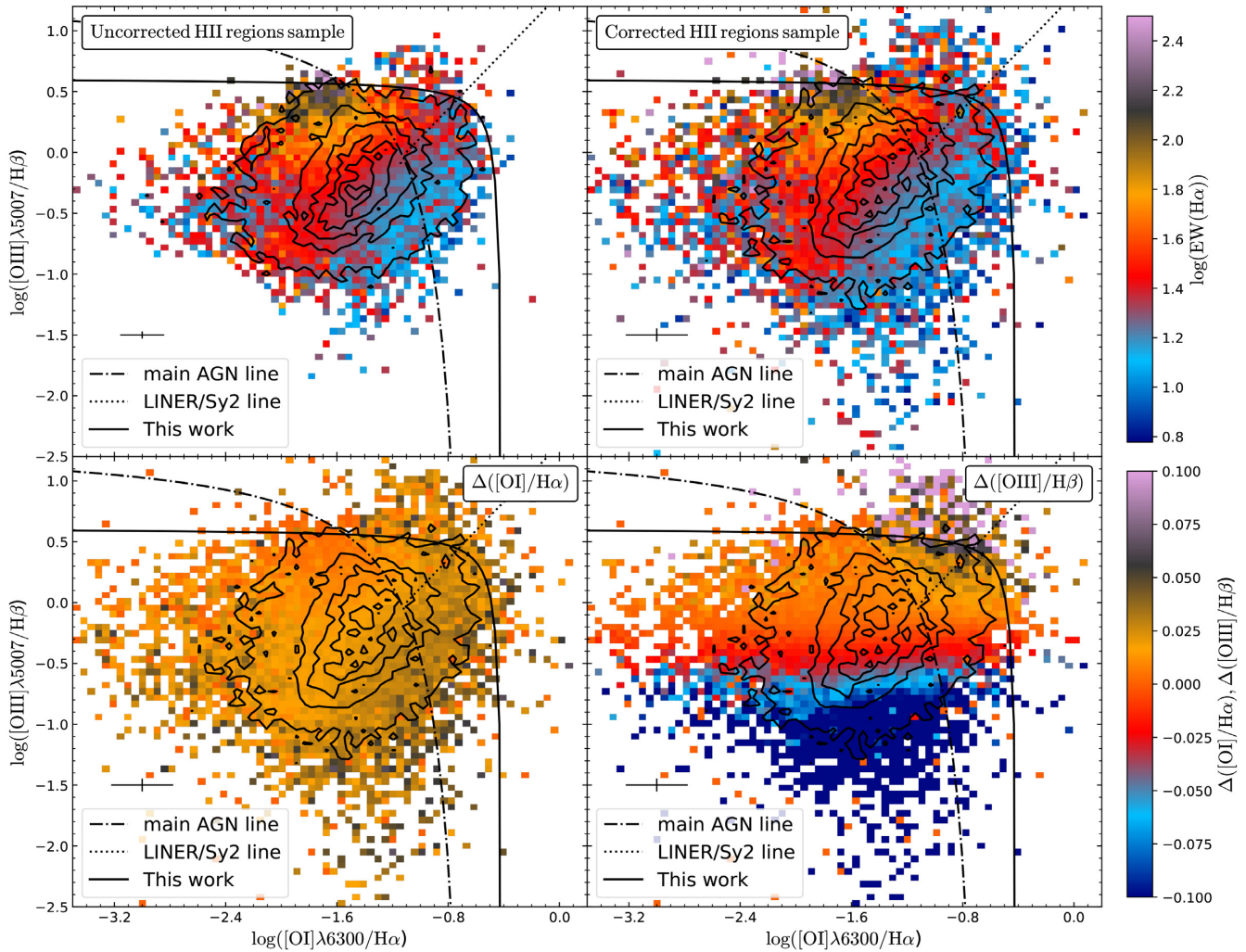
that the  $\sim 14$  per cent of our selected H II regions lie in the AGN zone according to the classical demarcation lines. After the DIG correction, this fraction remains  $\sim 13$  per cent. Thus, the differences with classical selection criteria cannot be attributed to this effect.

In both diagrams, the new demarcation curves present a different shape than the classical ones, and the SF zone defined by the new curves is higher than the classical ones. As we include regions with an  $\text{EW}(\text{H}\alpha) \sim 6 \text{ \AA}$ , our curve could include some regions that have a mixture in the ionizing sources, according to Lacerda et al. (2018). Therefore, the classical lines provide a more conservative classification scheme to identify SF regions; meanwhile, the new curve defines a classification scheme that includes SF regions with the largest contribution of other ionizing sources than the classical ones.

In Section 4.1, we present the effects of DIG correction across the BPT diagram (Fig. 6). The bottom panels on Figs A1 and A2 show the changes between the corrected and uncorrected fluxes across their corresponding diagnostic diagrams ( $[\text{O III}]\lambda 5006/\text{H}\beta$  versus  $[\text{S II}]\lambda 6716 + 30/\text{H}\alpha$  and  $[\text{O III}]\lambda 5006/\text{H}\beta$  versus  $[\text{O I}]\lambda 6300/\text{H}\alpha$ ), respectively. As was mentioned before, there is a clear dependence



**Figure A1.** Distribution of H II regions along the  $[\text{O III}]\lambda 5006/\text{H}\beta$  versus  $[\text{S II}]\lambda 6716 + 30/\text{H}\alpha$  diagnostic diagram. *Left-hand panel:* distribution without DIG correction. *Right-hand panel:* distribution with DIG correction. The colour code indicates EW of H  $\alpha$ . The most frequently used demarcation curves for this diagram are shown (Kewley et al. 2006). The density contours are similar to those presented in Fig. 6.



**Figure A2.** Similar figure as the one presented in Fig. A1 for the  $[\text{O III}]\lambda 5006/\text{H}\beta$  versus  $[\text{O I}]\lambda 6300/\text{H}\alpha$  diagnostic diagram.

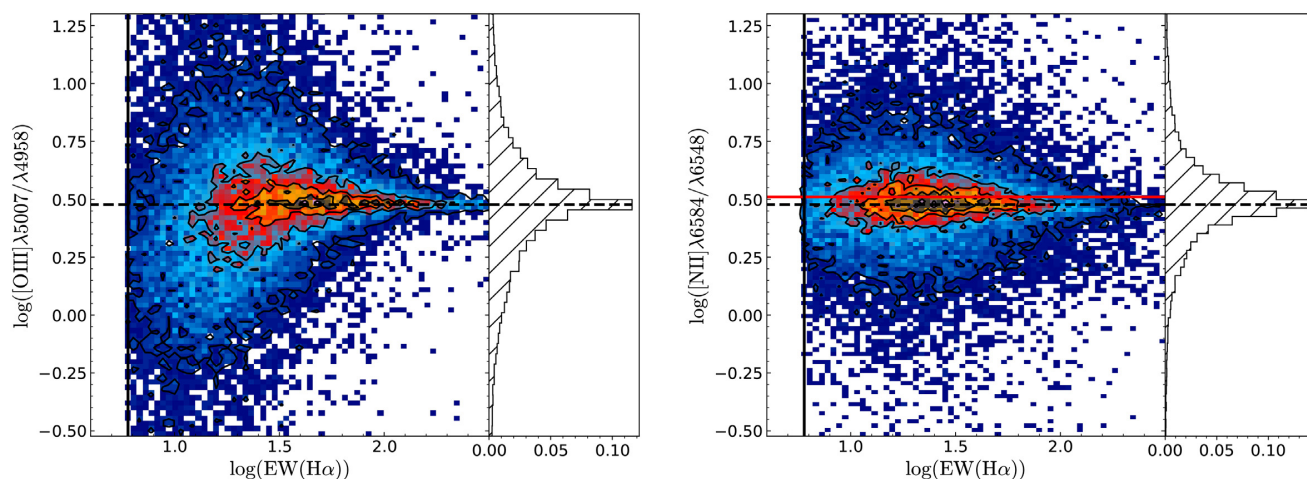
between the location on diagnostic diagrams and the difference of fluxes,  $\Delta([\text{S II}]/\text{H}\alpha)$ ,  $\Delta([\text{O III}]/\text{H}\beta)$ , and  $\Delta([\text{O I}]/\text{H}\alpha)$ . Unfortunately, for  $[\text{O I}]\lambda 6300/\text{H}\alpha$  ratio, the coefficient's order,  $C_j$ , is less than the associate error order; hence, the dependence is not clear in the diagnostic diagram. This issue, as we discussed before, may be improved by using data with a better resolution.

## APPENDIX B: QUALITY OF THE H II REGION SAMPLE

We check the quality of our H II region sample by comparing the observed distribution of the  $[\text{O III}]\lambda 5007/\lambda 4959$  and  $[\text{N II}]\lambda 6584/\lambda 6548$  flux emission line ratios to the theoretical value ( $\sim 3$ ; Osterbrock 1989). Therefore, these line ratios are an excellent indicator of the quality of the spectra of an ionized nebular region (and the subtraction of the underlying stellar population). Fig. B1 shows the density distribution of the considered line ratios as a function of the  $\text{EW}(\text{H}\alpha)$  for the final sample of H II regions (with blue representing the low density values). The black contours

represent the density distribution; the outermost contour encloses 95 percent of the regions and each consecutive one encloses a 20 percent fewer number. In both the panels, the horizontal line represents the theoretical value of the corresponding line ratio and the vertical line represents the  $\text{EW}(\text{H}\alpha)$  threshold that we used to classify the H II regions [ $\text{EW}(\text{H}\alpha) = 6 \text{ \AA}$ ]. As can be seen in the figure, most of the final selected H II regions lie near the theoretical value. However, there is a significant scatter in those regions near the adopted  $\text{EW}(\text{H}\alpha)$  threshold. Only low-S/N (i.e. low contrast at low EW) spaxels present non-physical line ratios, as expected due to the significant errors of the measured fluxes. The actual value used to decontaminate the DIG is represented as a red horizontal line, being more compatible with expected physical value. Unfortunately, several emission lines, like  $[\text{O III}]\lambda 4958$ , have  $\text{S/N} < 1$  even when combining several DIG spaxels. Therefore, not all emission lines (and line ratios) in the spectral range of data could be corrected by DIG in our data set.





**Figure B1.**  $[\text{O III}]\lambda 5007/\lambda 4958$  and  $[\text{N II}]\lambda 6584/\lambda 6548$  ratios as a function of  $\text{EW}(\text{H}\alpha)$  for: (i) H II region sample and (ii) DIG sample (mean value, red diamond point) as described in the text. The red horizontal line represents the weighted mean value used to derive the DIG correction (see Section 3.4). The black horizontal line corresponds to the theoretical value of  $[\text{O III}]$  and  $[\text{N II}]$  line ratio. The vertical line stands for the  $\text{EW}(\text{H}\alpha)$  threshold used to segregate the H II regions from the ionized regions. On the right hand, for both the panels, we show the distribution of each emission line ratio. Both distributions are centred in the expected value.

## APPENDIX C: DESCRIPTION OF THE CATALOGUE

The catalogue of H II regions described along this article, with the full set of parameters derived for each region, is presented in five different files for each galaxy (labelled by its name, GALNAME):

(i) `seg_Ha.GALNAME.fits`: Fits file comprising the segmentation maps indicating the actual selection of H II regions. Each map has the same geometry and WCS of the original  $\text{H}\alpha$  maps used to select the regions. All pixels corresponding to each H II region are labelled with a single ID number that identifies that region on the catalogue (for each galaxy), with IDs starting from one. Pixels labelled with zeroes correspond to areas excluded from the selection of H II regions.

(ii) `mask_Ha.GALNAME.fits`: Fits file comprising a mask of those pixels not considered to be part of any selected H II region.

(iii) `HII.GALNAME.fluxelines.csv`: CSV table including the fluxes of the emission lines extracted for each H II region within the considered galaxy, together with the location

of the region with respect to the centre of the galaxy. Each region corresponds to one row, labelled with the ID included in the `seg_Ha.GALNAME.fits` file and the galaxy name (GALNAME). Details of the content in this file are given in Table C1.

(iv) `HII.GALNAME.SFH.csv`: CSV table including the fraction of light in the  $V$  band of the decomposition of the stellar population currently for each stellar population included in the SSP template adopted by PIPE3D (Sánchez et al. 2016b). Details of the content in this file are given in Table C2.

(v) `HII.GALNAME.SSP.csv`: CSV table including the average properties of the underlying stellar population (age,  $[Z/H]$ , dust attenuation), as described in Sánchez et al. (2016b). Details of the content in this file are given in Table C3.

The full catalogue comprising all these files is stored at: [http://ifs.astroscu.unam.mx/CALIFA/HII\\_regions/](http://ifs.astroscu.unam.mx/CALIFA/HII_regions/) for public access.

**Table C1.** The table has 213 columns; from 10th column to 213th column it comprises the flux and EW data for the 51 emission lines explored in this study (irrespective of their S/N). We have to remark that errors in the WCS are propagated in RA and DEC.

Keyword	Description
HIIREGID	Unique identifier of the ionized region. The generic format for all regions is GALNAME-XX where GALNAME is the name of the host galaxy and XX is the unique ID described in the previous section.
GALNAME	Name of the host galaxy
CALIFAID	ID assigned to each galaxy in the CALIFA survey
X	X centroid of the ionized region in the cube
Y	Y centroid of the ionized region in the cube
RA	Right ascension of the centroid of the ionized region
DEC	Declination of the centroid of the ionized region
Ha_VEL	H $\alpha$ rotational velocity in km s <sup>-1</sup>
Ha_DISP	H $\alpha$ rotational velocity dispersion in km s <sup>-1</sup>
fluxOII3727	Integrated flux of [O II] $\lambda$ 3727 in units of 10 <sup>-16</sup> erg s <sup>-1</sup> cm <sup>-2</sup>
e_fluxOII3727	Estimated error of the integrated flux of [O II] $\lambda$ 3727 in units of 10 <sup>-16</sup> erg s <sup>-1</sup> cm <sup>-2</sup>
EW_OII3727	EW of [O II] $\lambda$ 3727 in Å
e_EWOII3727	Estimated error of EW of [O II] $\lambda$ 3727 in Å
...	...
fluxSII6731	Integrated flux of [S II] $\lambda$ 6731 in units of 10 <sup>-16</sup> erg s <sup>-1</sup> cm <sup>-2</sup>
e_fluxSII6731	Estimated error of the integrated flux of [S II] $\lambda$ 6731 in units of 10 <sup>-16</sup> erg s <sup>-1</sup> cm <sup>-2</sup>
EWSII6731	EW of [S II] $\lambda$ 6731 in Å
e_EWSII6731	Estimated error of EW of [S II] $\lambda$ 6731 in Å

**Table C2.** Description of the HII.GALNAME.SFH.csv files.

Keyword	Description
COLUMN1	HIIREGID
COLUMN2	Luminosity Fraction for age–met 0.0010–0.0037 SSP
COLUMN3	Luminosity Fraction for age–met 0.0010–0.0076 SSP
...	...
COLUMN156	Luminosity Fraction for age–met 7.9433–0.0190 SSP
COLUMN157	Luminosity Fraction for age–met 7.9433–0.0315 SSP
COLUMN158	Luminosity Fraction for age 0.0010 SSP
COLUMN159	Luminosity Fraction for age 0.0030 SSP
...	...
COLUMN195	Luminosity Fraction for age 12.5893 SSP
COLUMN196	Luminosity Fraction for age 14.1254 SSP
COLUMN197	Luminosity Fraction for met 0.0037 SSP
COLUMN198	Luminosity Fraction for met 0.0076 SSP
COLUMN199	Luminosity Fraction for met 0.0190 SSP
COLUMN200	Luminosity Fraction for met 0.0315 SSP
COLUMN201	Error in the Lum. Fract. for age–met 0.0010–0.0037 SSP
...	...
COLUMN399	Error in the Lum. Fract. for met 0.0315 SSP

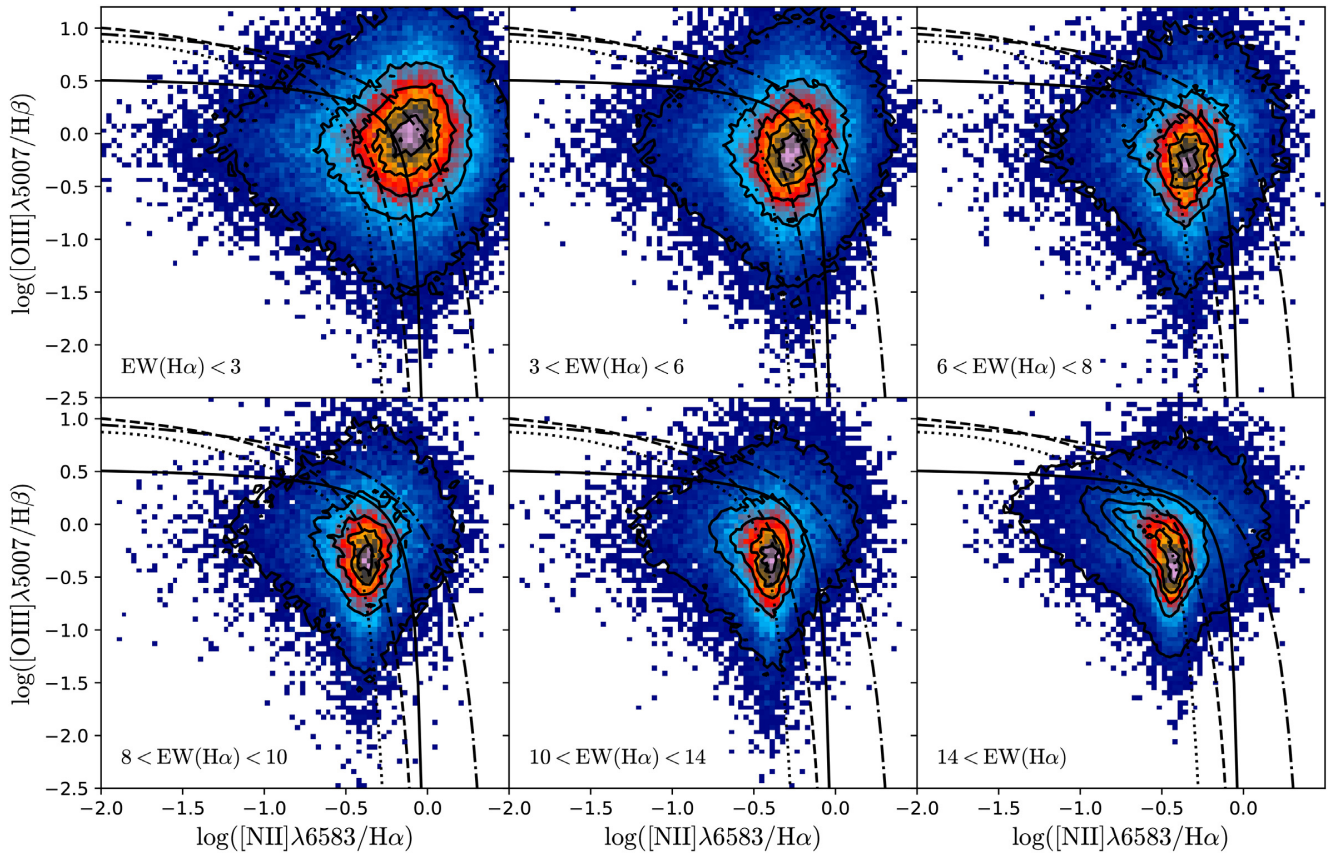
**Table C3.** Description of the `HI I . GALNAME . SSP . csv` files.

Column	Keyword	Description
COLUMN1	HI I REGID	Ionized region ID
COLUMN2	GALNAME	Galaxy name
COLUMN3	CALIFAID	Name of the host galaxy
COLUMN4	X	X centroid of the ionized region in the cube
COLUMN5	Y	Y centroid of the ionized region in the cube
COLUMN6	RA	Right ascension of the centroid of the ionized region
COLUMN7	DEC	Declination of the centroid of the ionized region
COLUMN8	log_age_LW	Luminosity-weighted age of the stellar population in log(age/yr)
COLUMN9	log_age_MW	Mass-weighted age of the stellar population in log(age/yr)
COLUMN10	e_log_age	Error of the age of the stellar population
COLUMN11	ZH_LW	Luminosity-weighted metallicity of the stellar population in log(Z/Z <sub>⊙</sub> )
COLUMN12	ZH_MW	Mass-weighted metallicity of the stellar population in log(Z/Z <sub>⊙</sub> )
COLUMN13	e_ZH	Error metallicity of the stellar population
COLUMN14	AV_ssp	Average dust attenuation of the stellar population ( $A_{V, stars}$ ) in mag
COLUMN15	e_AV_ssp	Error of the average dust attenuation of the stellar population in mag
COLUMN16	vel_ssp	Velocity of the stellar population in km s <sup>-1</sup>
COLUMN17	e_vel_ssp	Error in the velocity of the stellar population
COLUMN18	disp_ssp	Velocity dispersion of the stellar population in km s <sup>-1</sup>
COLUMN19	e_disp_ssp	Error in velocity dispersion of the stellar population
COLUMN20	log_ML	Average mass-to-light ratio of the stellar population in Solar Units
COLUMN21	log_Sigma_Mass	Stellar mass density in M <sub>⊙</sub> arcsec <sup>-2</sup> , not dust corrected
COLUMN22	log_Sigma_Mass_corr	Stellar mass density in M <sub>⊙</sub> arcsec <sup>-2</sup> , dust corrected using the $A_{V, stars}$

#### APPENDIX D: DISTRIBUTION OF THE EW(H $\alpha$ ) OF THE DIG

In Section 3.4.1, the properties of DIG spaxels from all galaxies were discussed. In order to enhance the DIG's characterization, we apply selection criteria to galaxies, as we discussed above. Fig. D1 shows the distribution of *clean* DIG sample on the BPT diagram segregated by the EW(H $\alpha$ ). The code colour represents the point density and the black contours represent the distribution of raw DIG sample. As can be seen in the figure, for low EW values, the spaxels distribution is more located towards the area associated with ionization by the old stellar population. On the other hand, for higher EW values, the distribution is located in the SF area of the diagram.

Along the intermediate range, the spaxel distribution moves from the area associated with ionization by the old stellar populations towards the SF area. This trend is also observed in the distribution of the raw DIG sample (black contours). However, a significant fraction of spaxels lies below the Kauffman's curve for raw DIG sample (black contours) and clean DIG sample. They correspond to spaxels contaminated by ionization by H II regions due to PSF wings. Although, the contribution of these types of spaxels was reduced considerably in the *clean* DIG sample. Unfortunately, the spaxels in the SF zones are still present across all EW values. Hence, we cannot segregate the diffuse gas ionized by leaking of ionizing photons from the H II regions due to the PSF wing contribution.



**Figure D1.** Final distribution of DIG spaxels segregated by mean  $EW(H\alpha)$  along the BPT diagram. The density contours are the same that in Fig. 6 for the raw DIG sample. In both the panels, the dotted, dot-dashed, dashed, and solid lines are the demarcation curves used in Fig. 6.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.



## Capítulo 3

# Propiedades de las Regiones HII

En el capítulo anterior, presentamos el catálogo de regiones H II publicado en [Espinosa-Ponce et al. \(2020\)](#) fue generado a partir de  $\sim 900$  galaxias del proyecto CALIFA con la información de  $\sim 25,000$  objetos individuales. Entre esta información caber destacar los propiedades de 51 líneas de emisión en el rango óptico. En este trabajo exploramos las propiedades físicas de las regiones H II: (1) fracción de población estelar joven, (2) parámetro de ionización, (3) abundancia de oxígeno, (4) cociente de abundancias nitrógeno-oxígeno, (5) extinción por polvo y (6) densidad electrónica. Estudiamos en detalle la dependencia de estas propiedades físicas con el entorno, tanto local (caracterizada por la distancia galactocéntrica), como global (masa y morfología de las galaxias).

Debido a las limitaciones del método directo descritas en la Sec 1.2.4 del Cap 1, se usaron calibradores de líneas fuertes para el cálculo del parámetro de ionización y para la abundancia de oxígeno. En concreto se adoptaron 6 y 24 diferentes calibradores, respectivamente. Todos los valores calculados forman ahora parte del catálogo publicado de libre acceso. Se estudiaron los gradientes radiales de las propiedades mencionadas anteriormente. Dado que la muestra cubre diversos tipos de galaxias anfitrionas, en este estudio se exploró la dependencia de las propiedades y sus distribuciones radiales con la morfología y con la masa estelar de las galaxias. Por otra parte, anteriormente se remarcó la importancia del oxígeno como trazador de procesos de formación estelar, por lo que en este trabajo se exploró su relación con las otras propiedades físicas y con las propiedades de las poblaciones estelares subyacentes.

Como resultado de esta exploración se concluyó que algunas (p. ej. abundancia de O, cociente N/O, parámetro de ionización) de las propiedades de las regiones H II están relacionadas con la evolución estelar y el entorno donde se forman. Además, la mayoría de las propiedades físicas de las regiones H II presentan un gradiente radial que depende de la masa y la morfología de las galaxias anfitrionas. Estos resultados concuerdan con los perfiles radiales antes encontrados por otros autores (p. ej. [González Delgado et al., 2016](#); [Pérez-Montero et al., 2016](#); [Belfiore et al., 2017](#); [Esteban & García-Rojas, 2018](#); [Sánchez et al., 2018](#)). En particular, se encontró que las propiedades físicas y las líneas emitidas por las regiones ionizadas están conectadas con las propiedades de las poblaciones estelares subyacentes. Por lo tanto, esto indica que estas propiedades pueden ser el resultado de la evolución química de las poblaciones estelares que se encuentran en la vecindad donde estas regiones se formaron.

El análisis y resultados se presentaron en un artículo publicado en la revista MNRAS. La versión electrónica puede ser encontrada en la página: <https://ui.adsabs.harvard.edu/abs/2022arXiv220207518E/>. Los datos pueden ser encontrados en [http://ifs.astroscu.unam.mx/CALIFA/HII\\_regions/](http://ifs.astroscu.unam.mx/CALIFA/HII_regions/). El código para el cálculo de las propiedades físicas puede ser consultado en <https://github.com/cespinosa/pyOxy>. Finalmente, en <https://github.com/cespinosa/pyOxy>.

[com/cespinosa/HII\\_regions\\_catalog/](http://com/cespinosa/HII_regions_catalog/) pueden consultarse un ejemplo de los códigos utilizados para analizar los datos, mostrándose como generar la primera figura del artículo.



# H II regions in CALIFA survey: II. The relation between their physical properties and galaxy evolution

C. Espinosa-Ponce<sup>1</sup>,<sup>1</sup>★ S. F. Sánchez,<sup>1</sup> C. Morisset,<sup>2</sup> J. K. Barrera-Ballesteros<sup>1</sup>,<sup>1</sup> L. Galbany<sup>3,4</sup>,<sup>3,4</sup>  
R. García-Benito<sup>5</sup>,<sup>5</sup> E. A. D. Lacerda<sup>6</sup> and D. Mast<sup>6,7</sup>

<sup>1</sup>Universidad Nacional Autónoma de México, Instituto de Astronomía, AP 70-264, CDMX 04510, México

<sup>2</sup>Universidad Nacional Autónoma de México, Instituto de Astronomía, AP 106, Ensenada, 22800 BC, México

<sup>3</sup>Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain

<sup>4</sup>Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain

<sup>5</sup>Instituto de Astrofísica de Andalucía, CSIC, Apartado de correos 3004, E-18080 Granada, Spain

<sup>6</sup>Universidad Nacional Córdoba, Observatorio Astronómico de Córdoba, Córdoba, Argentina

<sup>7</sup>Consejo de Investigaciones Científicas y Técnicas de la República Argentina, Avda. Rivadavia 1917, C1033AAJ CABA, Argentina

Accepted 2022 February 13. Received 2022 February 9; in original form 2021 December 21

## ABSTRACT

We present here the exploration of the physical properties of the sample of H II regions and aggregations of the last H II regions catalogue of the Calar Alto Legacy Integral Field Area (CALIFA) survey. This sample comprises the optical spectroscopic properties of more than  $\sim 26\,000$  ionized regions corresponding to 924 galaxies from the Integral Field Spectroscopy data, including the flux intensity and equivalent widths and the properties of their underlying stellar population. In the current study, we derive a set of physical quantities for all these regions based on those properties, including (i) the fraction of young stars; (ii) the ionization strength (using six different estimations); (iii) the oxygen abundance (using 25 different calibrators); (iv) the nitrogen and nitrogen-to-oxygen abundance; (v) the dust extinction; and (vi) the electron density. Using this data set, we explore how the loci in the classical diagnostic diagrams are connected with those quantities, the radial distributions of these parameters, and the interrelations between themselves and with the properties of the underlying stellar populations. We conclude that many properties of the H II regions are tightly related to the galactic stellar evolution at the location where those regions are observed. Those properties are modulated only as a second-order effect by the properties of the ionizing stars and the ionized nebulae that do not depend on the astrophysical context in which they are formed. Our results highlight the importance of H II regions to explore the chemical evolution in galaxies, clarifying which of their properties can be used as proxies of that evolution.

**Key words:** techniques: spectroscopic – ISM: general – H II regions – galaxies: ISM – galaxies: star formation.

## 1 INTRODUCTION

H II regions are gas nebulae ionized by young and massive, short-lived, OB stars (e.g. Strömgren 1939). Due to their nature, they are tracers of the star formation processes in galaxies, frequently found in late-type/disc galaxies (e.g. Kennicutt & Hodge 1980), although they have been recently observed in some early-type ones (e.g. Gomes et al. 2016). Their nebular emission lines make possible to directly measure the gas-phase abundance at discrete spatial positions in galaxies (e.g. Searle 1971; Peimbert, Torres-Peimbert & Rayo 1978; Kennicutt & Garnett 1996; Kennicutt, Bresolin & Garnett 2003). Hence, they are one of the primary tools for constraining galactic chemical evolution theories (e.g. Carigi, Peimbert & Peimbert 2019), as the present abundances are the consequence of the enrichment history (Kewley, Nicholls & Sutherland 2019; Maiolino & Mannucci 2019; Sánchez 2020). This connection induces relations between their observational properties, the chemical composition, the properties of the underlying stellar

populations, and the global properties of their host galaxies (e.g. Sánchez et al. 2015b; Espinosa-Ponce et al. 2020).

Previous studies have determined essential relationships, galaxy wide patterns, and scaling laws between the chemical abundances in galaxies and their physical properties. Among them, the most relevant ones are the surface brightness–metallicity, mass–metallicity, and luminosity–metallicity relations (Lequeux et al. 1979; Skillman, Kennicutt & Hodge 1989; Vila-Costas & Edmunds 1992; Zaritsky, Kennicutt & Huchra 1994; Tremonti et al. 2004; Barrera-Ballesteros et al. 2018), characteristics versus integrated abundances (Moustakas & Kennicutt 2006), or characteristic shapes of abundance gradients (e.g. Vila-Costas & Edmunds 1992). These relations were derived either for integrated properties of galaxies or using limited samples of H II regions (and for a small number of galaxies), limiting our understanding of the chemical evolution derived from them.

In the last years, the new generation of emission-line surveys based on Integral Field Spectroscopy (IFS) instruments and multi-object spectrometers, covering large fields of view (FoVs), offers catalogues of H II regions with hundreds of objects covering the full optical extension of nearby galaxies (e.g. Rosales-Ortega et al. 2010; Sánchez et al. 2012a, 2014). The H II region catalogues produced by these surveys are based on large samples of galaxies covering a

\* E-mail: [cespinosa@astro.unam.mx](mailto:cespinosa@astro.unam.mx), [carlcespinosa@gmail.com](mailto:carlcespinosa@gmail.com)

wide range of morphological types (Sánchez et al. 2012b, 2013). Significant results have been determined using these catalogues: (i) the existence of a typical abundance gradient (Sánchez et al. 2014, 2015a; Sánchez-Menguiano et al. 2018), (ii) the deviations from this radial abundance gradient induced by spiral arms (Sánchez-Menguiano et al. 2020) and their azimuthal variations (Vogt et al. 2017), (iii) the correlation between the oxygen abundance of H II regions and the stellar mass densities of underlying stellar populations (Rosales-Ortega et al. 2012; Barrera-Ballesteros et al. 2016), (iv) the existence of average patterns in the radial gradient of different physical properties (Sánchez et al. 2012b), (v) the existence of a large number of H II regions contaminated by supernovae remnants (Cid Fernandes et al. 2021), (vi) chemical evolution models at kpc scales (Barrera-Ballesteros et al. 2018), and (vii) the comparison between the pre-SNe feedback and environmental pressure across the optical extension of spiral galaxies (Barnes et al. 2021), among the most important ones.

In Espinosa-Ponce et al. (2020), we presented a new catalogue of H II regions based on the Calar Alto Legacy Integral Field Area (CALIFA; Sánchez et al. 2012a) data. We used the extended sample (hereafter, eCALIFA) that comprises the CALIFA (692 galaxies) and PISCO samples (232 galaxies) (Sánchez et al. 2016b; Galbany et al. 2018). We use the tool PYHIEXPLORER<sup>1</sup> to detect, segregate, and extract the main spectroscopic properties of these H II regions. Then, we built the most extensive and homogeneous catalogue of spectroscopic properties of H II regions in the nearby Universe. This catalogue has allowed us to confirm the correlations between the ionization conditions in H II regions and the properties of the underlying stellar populations. In particular, the relation between  $[\text{O III}]\lambda 5007/\text{H}\beta$  and  $[\text{N II}]\lambda 6583/\text{H}\alpha$  ratios with the age and metallicity of the underlying stellar populations (previously found in Sánchez et al. 2015b). More recently, we have used this catalogue to explore the  $[\alpha/\text{Fe}]$  patterns in galaxies (Sánchez 2020).

In the current study, we derive the main physical properties of the H II regions in this catalogue, including different derivations for the oxygen and nitrogen abundances using a wide set of state-of-the-art calibrators, together with estimations of the ionization strength, electron density, and dust extinction. We explore the distribution of those properties in the classical diagnostic diagrams, connecting them with stellar population properties. Finally, we explore the gradients of all the derived physical properties and characterize them by the mass and morphology of the galaxies.

The structure of the article is as follows: the data explored in this study are described in Section 2, including a brief description of the sample of galaxies, and the procedure to detect and extract the properties of the H II regions; the analysis and results are presented in Section 3, with the derivations of the properties for the H II regions described in Section 3.1. The main trends across the diagnostic diagrams are presented in Section 3.2, while the radial gradients are shown in Section 3.3. Finally, the relations of the oxygen abundance with other physical properties of the H II regions and the properties of the underlying stellar populations are shown in Sections 3.4 and 3.5, respectively; the main conclusions of this study are presented in Section 5.

## 2 DATA

We use the catalogue of spectroscopic properties of H II regions described in the previous section. This catalogue was extracted from

the sample of  $\sim 1000$  galaxies comprising the eCALIFA survey. This sample was primarily diameter-selected to have a characteristic projected size of  $\sim 60$  arcmin (Walcher et al. 2014). It spans through a narrow redshift range centred at  $z \sim 0.015$ , covering a wide range of stellar masses and morphologies. In summary, it is a somehow representative sample of the galaxies in the nearby universe.

The whole data set was observed at the 3.5-m telescope of the Calar Alto Observatory, using the Potsdam Multi Aperture Spectrograph (PMAS; Roth et al. 2005) in the Pmas fibre PAcK configuration (PPAK; Kelz et al. 2006). This configuration provides a sufficient FoV ( $74'' \times 64''$ ) to map the full optical extent of these galaxies. It is possible to map up to 2.5 effective radii due the diameter sample selection, described before. The observing strategy guaranteed a final spatial resolution of full width at half-maximum  $\sim 2.4''$  ( $\sim 0.8$  kpc) at the sample's average redshift (Sánchez et al. 2016b). The current sample was observed using the V500 setup, which provides a spectral resolution of  $\lambda/\Delta\lambda \sim 850$  and a wavelength range (3745–7200 Å). This is sufficient to study the most important ionized gas emission lines in the optical range (from  $[\text{O II}]\lambda 3727$  to  $[\text{S II}]\lambda 6731$ , at the average redshift of the explored galaxies). The wavelength range and resolution are good enough to deblend these emission lines from the underlying stellar population and to derive the main properties of both components (e.g. Kehrig et al. 2012; Sánchez et al. 2012a, 2016a; Cid Fernandes et al. 2013, 2014).

The data were reduced with the CALIFA v2.2 pipeline (which is described in detail in Sánchez et al. 2016b). The reduction comprises the standard procedures for this kind of data outlined in Sánchez (2006). This procedure provides for each observed galaxy a regular-gridded data cube, with two axis recording the spatial dimensions (with a sampling size of 1 arcsec/spaxel), and the remaining one recording the spectral dimension. These data cubes were analysed using the PIPE3D pipeline in order to derive the spatial resolved properties of the ionized gas emission lines and the stellar population. This tool was developed (and it has been used) to analyse different IFS data, e.g. CALIFA (Sánchez-Menguiano et al. 2016; López-Cobá et al. 2019), MaNGA (Ibarra-Medel et al. 2016; Barrera-Ballesteros et al. 2018; Sánchez et al. 2018), MUSE (López-Cobá et al. 2020), and SAMI (Sánchez et al. 2019). The underlying stellar population is fitted using the GSD156 single stellar population library (Cid Fernandes et al. 2013), which comprises 156 templates: 39 stellar ages (from 1 Myr to 14.1 Gyr) and four metallicities ( $Z/Z_{\odot} = 0.2, 0.4, 1, \text{ and } 1.5$ ). The derived stellar spectrum model is then subtracted from the original one for each spaxel to obtain gas-pure data cube that it is then analysed to extract the properties of the emission lines. The final data product of this analysis is a set of maps that retain the world coordinate system of the original cubes comprising the spatial distribution of each particular quantity.

The segregation of H II regions and extraction of the corresponding emission lines information are performed by applying the PYHIEXPLORER tool to the data products provided by PIPE3D. This semiautomatic procedure detects and segregates candidates to H II regions based on three assumptions: (i) H II regions have a strong and peaky  $\text{H}\alpha$  emission that is clearly recognizable above both the continuum emission and the diffuse ionized gas (DIG) emission across the galaxies; (ii) the typical size of an extragalactic H II regions at this cosmological distance is of the order of the size of the point spread function; and (iii) the underlying stellar population at the location of the selected regions has enough young stars to produce the observed ionization, as described in detail in Espinosa-Ponce et al. (2020). The input parameters used to identify the ionized regions are: (i) a flux intensity threshold for each ionized region peak emission of  $3 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-1}$ , (ii) a minimum

<sup>1</sup><https://github.com/cespinosa/pyHIexplorerV2>

relative flux to the peak emission for associated spaxels of the same H II region of 5 per cent, (iii) a maximum distance to the peak spaxel of 5.5 arcsec, and (iv) an absolute flux intensity threshold of  $0.5 \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-1}$  in the adjacent pixels to associate them to the peak spaxel. The initial sample of ionized regions identified by PYHIIEXPLORER includes the information of 38 807 objects from 924 galaxies. The  $\sim 70$  per cent of the total sample corresponds to H II regions, the 6.4 per cent corresponds to DIG powered by hot old low-mass evolved stars (HOLMES), and  $\sim 3$  per cent corresponds to AGN-like (for further details, see Espinosa-Ponce et al. 2020).

The result of this analysis for the full eCALIFA sample is a catalogue that comprises the information of 51 emission lines in the considered wavelength range and the properties of the underlying stellar population for  $\sim 26\,000$  H II regions. However, not all those regions are suitable for exploring the distributions along the diagnostic diagrams and to derive good-quality physical quantities, as we will show in the forthcoming sections. Therefore, we select a sub-sample of regions following the procedures that we will describe in Section 3.1. We use this catalogue as the starting point of the current analysis.

### 3 ANALYSIS AND RESULTS

We focus in this study on the exploration of the physical properties of H II regions, trying to determine connections between the ionization conditions, the nebular emission, and the properties of the underlying stellar populations. Such connection may unveil which properties of H II regions are related to the evolution of different locations in galaxies (and which ones are not).

#### 3.1 Empirical estimation of physical properties of H II regions

The physical properties of H II regions have been frequently explored using the classical diagnostic diagrams (Baldwin, Phillips & Terlevich 1981; Veilleux & Osterbrock 1987; Kewley et al. 2001). It is known that the location in these diagrams depends on the properties of the ionizing source (a stellar population in the case of an H II region), and the internal physical conditions, structure, and chemical composition of the ionized gas (e.g. Evans & Dopita 1985; Dopita & Evans 1986; Kewley et al. 2001; Kauffmann et al. 2003; Sánchez et al. 2015a). In summary, six main parameters define the location in these diagrams: (i) the shape/hardness of the ionizing spectral energy distribution (SED) that in an H II region is defined by the sum of the spectra of the ionizing stars (a cluster of them at the typical resolution of our data); (ii) the mean ionization parameter across the nebula, defined as the Lyman continuum photons with respect to the total amount of hydrogen (defined as  $U$ , see below); (iii) the chemical abundance of different elements in the ionized gas, relative to hydrogen (e.g. O/H, N/H...); (iv) the electron density of the gas (e.g. Sánchez et al. 2015b; Morisset et al. 2016; Kewley et al. 2019), denoted as  $n_e$ ; (v) the dust extinction, characterized by the extinction at the  $V$ -band ( $A_V$ ) and a certain extinction curve; and (vi) the optical depth at 13.6 eV (matter-bounded regions show higher ionization, due to the reduction of the low ionization zone). Following Sánchez et al. (2014), we investigate the main trends across the diagnostic diagrams of different physical properties of the H II regions: fraction of young stars, ionization parameter, oxygen abundance, nitrogen abundance, electron density, and dust extinction.

The fraction of young stars that contributes to the observed spectrum within the aperture at which the H II region (or cluster) is selected as a gauge of how strong is the contribution of the ionizing OB stars. It is also a proxy of how evolved is the area in the galaxy

(or the host galaxy itself), since it weights the relative amount of old stars with respect to the ionizing ones. Furthermore, it is a proxy of the shape/hardness of the ionizing spectrum, since this will be the product of the mix of the young and old ionizing stars (e.g. hot evolved stars, HOLMES, Flores-Fajardo et al. 2011). In Espinosa-Ponce et al. (2020), we defined this parameter as the fraction of light in the  $V$ -band corresponding to stars younger than 350 Myr.

The ionization parameter is defined as  $U(r) = Q(H^0)/4\pi r^2 n_H c$ , where  $Q(H^0)$  is the number of  $H^0$ -ionizing photons emitted by the source per unit of time,  $r$  is the distance between the source and a particular location within the nebula,  $n_H$  is the hydrogen density, and  $c$  is the speed of light. By measuring the ratio between the photons and the hydrogen atom densities,  $U$  gives an insight to the efficiency of the radiation that ionizes the gas. The higher  $U$  is, the easier the radiation can ionize metals. The value of  $U$  is straightforward to determine in theoretical models but difficult to determine from observations. Its mean that value within an H II region is usually estimated using different line ratios, like [O II]/[O III], as photoionization models unveil a connection with this parameter and  $O^+/O^{++}$  (e.g. Díaz et al. 2000). However, the reported connection depends strongly on many other properties of the nebulae, in particular, on its shape (shell or filled), on the shape of the ionizing SED, and on the optical depth of the nebula at 13.6 eV. Despite this issue, different calibrators have been proposed that link this proxy with  $U$  (e.g. Díaz et al. 2000; Dors et al. 2011; Morisset et al. 2016; Kewley et al. 2019). We adopted as fiducial calibrator the one presented by Morisset et al. (2016), which is an update of the relation first proposed by Díaz et al. (2000) and Dors et al. (2011). The [S II]/[S III] line ratio is frequently claimed to get a better estimation of the ionization parameter (Kewley et al. 2019), but the [S III] $\lambda$ 9069 +  $\lambda$ 9532 emission line is not available in the spectral range of our data. Additional estimators of  $U$  have been considered along this article and discussed in Appendix F.

The abundance of a particular element (e.g. oxygen) may be derived in theory by adding up all its ionic abundances. On the one hand, this can be done by measuring the flux of the recombination lines corresponding to the considered element. Unfortunately, most of those lines are extremely weak in the optical range. If those lines are not accessible, it is possible to derive O/H using the direct method; this is by measuring the flux of collisionally excited lines (read Peimbert, Peimbert & Delgado-Inglada 2017; Pérez-Montero 2017, for a review on this topic). For doing so, it is required to derive the electronic temperature and density of the nebulae. The estimation of the temperature requires to derive the flux ratio between a set of auroral to nebular lines, e.g. [O III] $\lambda$ 4363/ $\lambda$ 5007. Then, the O/H is derived computing first certain ionic abundances. This method assumes that the temperature distribution is homogeneous for each considered ion through the ionized extension of the H II region. If this is not the case, it is still possible to apply the so-called  $t^2$  correction (Peimbert 1967). The direct method relies on auroral lines that are brighter than the recombination ones but are very weak too (e.g. [O III] $\lambda$ 4363). In many cases, they are not accessible with the usual depth and resolution of the IFS data of the most recent galaxy surveys (see Marino et al. 2013, as an example).

When neither the recombination nor the auroral lines are detected or accessible, the estimation of the gas-phase abundance relies on the use of known relations between particular strong emission line ratios and the required physical properties to derive the abundance (e.g.  $T_e$ , Alloin et al. 1979; Pilyugin & Mattsson 2011) and/or the abundances themselves (Peimbert et al. 2017). These estimations are called strong-line calibrators, and they can be derived using two approaches: (1) by comparing the ratios of strong emission lines with direct estimation of the oxygen abundances from observational data



sets (based either on the direct method or by the use of collisional lines). Then, a polynomial fitting is performed (e.g. Pilyugin & Mattsson 2011; Marino et al. 2013; Pilyugin & Grebel 2016) or another more complex regression techniques (e.g. machine learning Ho 2019), and (2) using grids of photoionization models to derive the relation between oxygen abundances and emission line ratios (e.g. Kewley & Dopita 2002; Thomas et al. 2018). Unfortunately, the derived oxygen abundance using the two methods shows a discrepancy up to  $\sim 0.1$ – $0.4$  dex (Kewley & Ellison 2008; Blanc et al. 2015). This discrepancy may be due to (i) the presence of electron temperature fluctuations in H II regions (Peimbert 1967), (ii) uncertain ionization corrections factors, (iii) oversimplicity of the photoionization models used to obtain calibrations (Kennicutt et al. 2003), or (iv) high energetic electrons following a non-Maxwellian distribution (e.g. the  $\kappa$  distribution Nicholls, Dopita & Sutherland 2012), among other reasons. The most used strong-line indicators are based on the following emission lines: [OII] $\lambda$ 3727, H  $\beta$ , [OIII] $\lambda$ 5007, H  $\alpha$ , [NII] $\lambda$ 6583, [SII] $\lambda$ 6717 + 31. They comprise N2 ( $\equiv$ [NII]/H  $\alpha$ ), O3 ( $\equiv$ [OIII]/H  $\beta$ ), O2 ( $\equiv$ [OII]/H  $\beta$ ), S2 ( $\equiv$ [SII]/H  $\alpha$ ), the combination among them, like R23 ( $\equiv$ O2+O3), O3N2 ( $\equiv$ O3/N2), N2O2 ( $\equiv$ N2/O2), or N2S2 ( $\equiv$ N2/S2), and more complex combinations, like the  $P$  parameter (e.g. Pilyugin, Vílchez & Thuan 2010). Each one of these line ratios presents different strengths and caveats for the oxygen abundance calculation (e.g. Curti et al. 2020). An additional complication is that many of the adopted emission line ratios present a bi-valuated relation with the oxygen abundance defining upper and lower oxygen branches (e.g. R23 Peimbert et al. 2017).

In this work, we calculate the oxygen abundance with different strong-line calibrators proposed by Kewley & Dopita (2002), Kobulnicky & Kewley (2004), Pettini & Pagel (2004), Tremonti et al. (2004), Pilyugin et al. (2010), Pilyugin & Mattsson (2011), Pilyugin & Grebel (2016), Marino et al. (2013), Ho (2019), and Curti et al. (2020). They comprise a heterogeneous compilation of calibrators, both empirical and theoretical, and using different indicators and mathematical procedures to anchor the abundance. We adopt the Ho (2019) calibrator as the fiducial one along this article, although we provide the measurements of the remaining ones in our published catalogue (see Appendix B).

As we discussed for the oxygen abundance, the nitrogen abundance determination has similar issues. Moreover, the recombination and auroral nitrogen emission lines are even more difficult to detect than those of the oxygen. Nevertheless, there are relations between several strong emission line ratios and the nitrogen abundance that could be used for the abundance determination (i.e. strong-line calibrations). In this work, we estimate the nitrogen abundance (N/H) and nitrogen-to-oxygen (N/O) relative abundance with the calibrations described by Pilyugin & Grebel (2016) and Ho (2019). We adopt this later one as our fiducial one.

The electron density,  $n_e$ , is derived from the ratio of emission lines of the same ion originated from levels with similar energy (e.g. Osterbrock 1989). In those cases, the emission line ratio does not depend strongly on the temperature (Peimbert et al. 2017). In particular, we use the [S II] doublet by solving the equation:

$$\frac{[\text{S II}]\lambda 6717}{[\text{S II}]\lambda 6731} = 1.49 \frac{1 + 3.77x}{1 + 12.8x}, \quad (1)$$

where  $x = 10^{-4} n_e t^{-1/2}$  and  $t$  is the electron temperature in units of  $10^4$  K (McCall, Rybski & Shields 1985). We assume a typical electron temperature of  $T = 10^4$  that corresponds to the usual conditions in H II regions (Osterbrock 1989). We should note that, as we indicate before, the dependence with the  $t_e$  is very weak. Unfortunately, the [S II] doublet ratio is sensitive to electron density

only for a limited range of values (from  $\sim 50 \text{ cm}^{-3}$  to  $\sim 7000 \text{ cm}^{-3}$ ). Nevertheless, this calibrator is reasonably robust to study the main trends explored in this study.

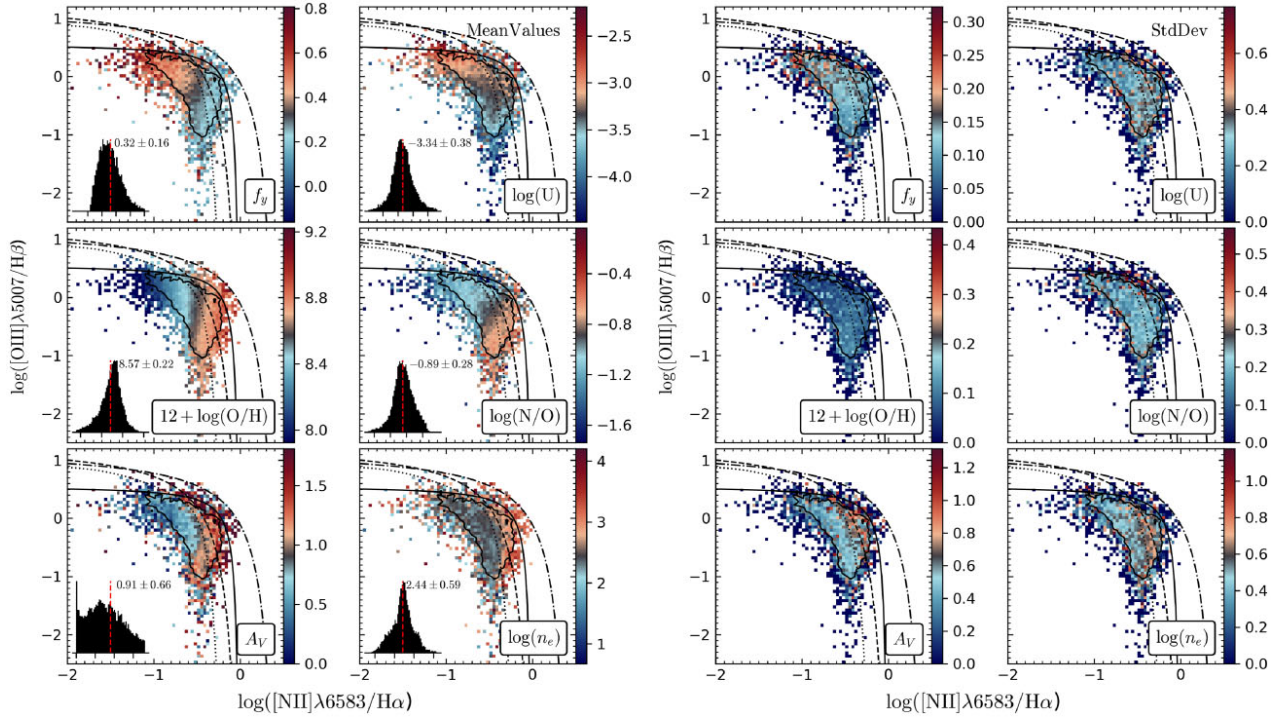
The dust extinction ( $A_v$ ) was estimated using the Balmer line ratio (H  $\alpha$ /H  $\beta$ ). We assume the extinction law by Cardelli, Clayton & Mathis (1989) using a theoretical value for the unobscured line ratio for case B recombination H  $\alpha$ /H  $\beta = 2.86$  (for  $T_e = 10000$  K and  $n_e = 100 \text{ cm}^3$ ) and a specific dust extinction of  $R_v = 3.1$ . Despite the dependence between the H  $\alpha$ /H  $\beta$  and the electron density and the temperature, we consider this variations negligible (Sánchez et al. 2007).

A catalogue with all the physical properties derived for the H II regions is distributed as described in Appendix B. From this catalogue, we apply different filters to ensure the quality of the derived properties: (i) the [O III]/H  $\beta$ , [N II]/H  $\alpha$ , [S II]/H  $\alpha$ , [O I]/H  $\beta$ , and [O II]/H  $\alpha$  line ratios should be positive and finite. Each of these consecutive cuts implies a reduction in the original number of  $\sim 26000$  H II regions by 712, 189, 288, 3326, and 8886 regions, respectively. Therefore, the strongest reduction is due to the presence or absence of [O II] ( $\sim 30$  per cent) in the wavelength range sampled by the data (due to the redshift), followed by the detectability of [O I] ( $\sim 12$  per cent). The final number of selected H II regions after applying these filter is  $\sim 15000$ ; (ii) we select those regions for which we can derived all the quantities enlisted along this subsection (EW(H  $\alpha$ ),  $f_y$ ,  $U$ , O/H, N/O,  $n_e$ , and  $A_v$ ) for the fiducial calibrators. This further filter removes just  $\sim 500$  additional regions; (iii) we exclude those regions above the Kewley et al. (2001) demarcation line in the diagnostic diagrams (that we will show in the forthcoming sections). This filter excludes  $\sim 450$  regions; finally, (iv) we exclude those regions that are hosted by a galaxies with an inclination greater than  $70^\circ$ . This final filter excludes  $\sim 3600$  regions corresponding to 177 galaxies. The final catalogue once applied all the filters comprises  $\sim 11000$  regions. The H II regions sample has a threshold in  $S/N > 3$  for H  $\alpha$  emission line from the initial selection of H II regions (Espinosa-Ponce et al. 2020).

### 3.2 Trends across the line-ratios diagnostic diagrams

In this section, we describe the observed trends of these properties across the emission-line diagnostic diagrams. These diagrams show the distribution of pairs of strong emission-line ratios among the ones described in the previous section to classify their primary ionization source. The most frequently used one, known as the BPT diagram, represents the O3 versus N2 line ratios (Baldwin et al. 1981). Other diagnostic diagrams involving different line ratios were proposed by Veilleux, Shopbell & Miller (2001), and a combination of them with the equivalent width of H  $\alpha$ , EW(H  $\alpha$ ), was first proposed by Cid Fernandes et al. (2010). In all those diagrams, there have been defined a set of demarcation lines that segregate certain regions associated with particular ionization sources (e.g. AGNs/SF Kewley et al. 2001; Kauffmann et al. 2003; Stasińska et al. 2008), although its ability to distinguish between them is now under discussion (e.g. Sánchez et al. 2021a).

Beside the ionizing source, the location of an H II region on this diagram depends on their physical properties (see Kewley et al. 2019; Sánchez et al. 2021a, and reference therein, for a discussion). In the particular case of H II regions, different theoretical works using photoionization models predict the location on the BPT diagrams according to the metallicity and the ionization parameter (e.g. Kewley & Dopita 2002; Morisset et al. 2016). They found a primary trend in which metal-poor H II regions are located in the upper-left zone of the diagram, while metal-rich ones are located



**Figure 1.** Average physical properties (left-hand panel), together with their standard deviation (right-hand panel) for the H II regions at a certain location within the  $[\text{O III}]\lambda 5007/\text{H}\beta$  versus  $[\text{N II}]\lambda 6583/\text{H}\alpha$  diagnostic diagram. In each panel, the colour code represents a physical property described in the text: *top left-hand panel*: fraction of young stars to the total luminosity,  $f_y$ ; *top right-hand panel*: ionization parameter; *middle left-hand panel*: oxygen abundance; *middle right-hand panel*: nitrogen-to-oxygen abundance; *bottom left-hand panel*: dust extinction; and *bottom right-hand panel*: electron density ( $\log(n_e)$ ). In each panel, the scale of the colour map was selected to cover the dynamical range of the physical parameter:  $\pm 3\sigma$  range around the mean value in the left-hand panels, and between zero and  $2\sigma$  in the right-hand panels, for all parameters but  $A_V$ . For this parameter, the dynamical range in the left-hand panel is restricted to 0 and two times the mean value ( $\sim 2$  mag). The inset within each panel shows the distribution of the corresponding property. The contour represents the density distribution enclosing 95 per cent of the regions. For all the panels, the dot–dashed line, dashed line, dotted line, and solid line represent the demarcation lines proposed by Kewley et al. (2001), Kauffmann et al. (2003), Stasińska et al. (2008), and Espinosa-Ponce et al. (2020), respectively.

in the right-bottom zone. Unfortunately, the photoionization models present different degeneracies in this diagnostic diagram between the oxygen abundance, the ionization strength, and the relative nitrogen-to-oxygen abundance (among others). As a result, one location in the diagram can be reproduced by a set of different models corresponding to different physical conditions. From the observational point of view, other trends have been reported. For instance, Sánchez et al. (2015a) described a clear trend in which H II regions located in regions with older stellar populations are more frequently found in the bottom-right part of the diagram, while those with younger stellar populations are found towards the upper-left zone. This trend, clearly connected with the one described for the metallicity, cannot be explained by photoionization models, being a consequence of the linked co-evolution of the stellar populations and their metal content at a certain location in galaxies. Furthermore, other physical properties show additional trends across the diagnostic diagrams. For instance, there is a clear trend with the electron density easily explained by photoionization models (e.g. Kewley et al. 2013). Finally, it was described a trend with dust extinction, despite the fact that the involved line ratios are almost insensitive to dust Sánchez et al. (2015a) and even with the gas velocity dispersion (Law et al. 2021).

Fig. 1 shows the distribution of the physical properties described in Section 3 across the BPT diagram for our sample of H II regions. To generate this figure, each physical property is averaged (left-hand panels) for those regions at the same location within the diagram. Once derived the mean value, we estimate the standard deviation around it (right-hand panels). We first explore the trends with the

fraction of young stars. The H II regions located on the upper-left zone of the diagram show the largest value of  $f_y$ ; meanwhile, the H II regions located on the bottom-right zone had the lowest one. This result agrees with the trends already reported by Sánchez et al. (2015a), highlighting the connection between the properties of the underlying stellar population and the observed line ratios. Thus, H II regions generated in (less) evolved regions of a galaxy, those with (younger) older and (less) more metal-rich stellar populations, are preferentially found in the areas with (higher) lower  $f_y$ . Similar trend is observed in the top right-hand panel of Fig. 1 that shows the same distribution colour-coded by the ionization parameter. The H II regions with a high value of the ionization parameters are located in the upper zone of the diagram, and regions with low values of  $U$  are located in the bottom zone. A connection between  $U$ ,  $f_y$ , and the metal content in H II regions may be the reason for these trends (e.g. Ji & Yan 2021).

The middle panels of Fig. 1 show the same distribution colour-coded by the oxygen abundance (left) and the nitrogen-to-oxygen relative abundance (right). As expected, the metal-poor H II regions are located in the upper-left zone of the diagram, and the metal-rich ones are located in the bottom-right zone. A similar trend is found for the nitrogen–oxygen ratio, as a direct consequence of relation between this quantity and the oxygen abundance (e.g. Matteucci 1986; Vila-Costas & Edmunds 1993; Pilyugin et al. 2012). Thus, the H II regions with the lowest value of N/O ratio are located on the upper-left zone of the diagram, and those with the largest ratio value are located on the bottom-right zone.

The distribution traced by the electron density ( $n_e$ ) is shown in the bottom right-hand panel of Fig. 1. The regions with the highest electron density ( $\sim 10^3 \text{ cm}^{-3}$ ) are found on the right-hand area of the diagram, near to the edge of the classical location of H II regions. Then, it is appreciated a sharp transition to values  $\sim 10^2 \text{ cm}^{-3}$ , distributed along that classical loci of these regions. An increase of  $n_e$  towards the upper right is predicted by photoionization models (e.g. Kewley et al. 2013). However, the described distribution is not that easily replicated based only on these models. We must remark that electron density derived with the [S II] lines ratio is reliable between  $\sim 10^2 \text{ cm}^{-3}$  and  $\sim 10^4 \text{ cm}^{-3}$ . For our current sample, we are in any case in the low-density regime for all H II regions.

Finally, the bottom left-hand panel of Fig. 1 shows the distribution of H II regions across the BPT diagram with a colour code representing the dust extinction. There is a clear trend across the diagnostic diagram: dust extinction is lower in the H II regions located in the upper left-hand region, and it is higher in those regions at the bottom right-hand end and in the upper right-hand end of the diagram. The regions with the largest dust extinction are located in the intermediate area delimited by Kauffmann et al. (2003) and Kewley et al. (2001) curves. Like in the case of some of the distributions discussed before, this result is not predicted by photoionization models that most of the time do not consider this parameter. We should note that the two line ratios involved in this diagram ([N II] and O III) are expected to be insensitive to dust extinction.

The distribution of the same physical properties for a set of additional diagnostic diagrams ([OIII] $\lambda 5007/H\beta$  versus [SII] $\lambda 6716 + 30/H\alpha$  and [OIII] $\lambda 5007/H\beta$  versus [OI] $\lambda 6300/H\alpha$ ) is shown in Appendix C. In all cases, the distributions cannot be reproduced based on predictions from simple photoionization models. Their origin should be related to the properties of the ambient in which the H II regions are formed. In summary, young massive stars born in regions with a large number of old stars (i.e. low  $f_y$ ), would be formed from a gas that has been recycled and enriched by previous stellar generations (i.e. high oxygen abundance and N/O ratio). The gas from which it is formed should also present a larger amount of dust and therefore a stronger dust extinction. More difficult is to interpret the trends described for the ionization parameter and the electron density. A possible explanation for the trend described by  $U$  could be that H II regions of a high metal content gas, ionized by a cluster of stars with a narrow range of ages, may present a lower ionization parameter due to metal blanketing. However, in this case, O/H and  $U$  should present just opposite trends across the diagram, and this is clearly not the case. On the other hand, for the electron density, the results suggest that regions with higher dust extinction (i.e. corresponding to large gas densities: Brinchmann et al. 2004; Barrera-Ballesteros et al. 2020) do have larger  $n_e$  too. This would imply that their internal pressure should be also higher, in agreement with the trends recently reported by Barnes et al. (2021). In this case, since  $U$  depends inversely on the electron density, the trend observed for this parameter could be a combination of the trends reported for both O/H and  $n_e$ . Another possibility would be that for the same ionizing stellar cluster, the ionization  $U$  decreases as  $A_V$  increases, as a pure consequence of the dust extinction decreases the available UV ionizing photons (Ji & Yan 2021). Finally, the observed trend for  $U$  could be a consequence of all those effects combined together. Besides, this result suggests that the diagnostic diagrams do not provide a proper description of the properties of H II regions; therefore, the models need to consider a parameter space with higher dimensions.

The outlined interpretation holds if the described distributions, traced by the average values of the considered parameters, are indeed

representative of the bulk sample of H II regions. In other words, the dispersion associated with the mean value should be enough small compared with the corresponding typical value. As indicated before, the distributions of the standard deviations are shown in the right-hand panels of Fig. 1 for the same physical parameters shown in the left-hand panels. It is appreciated that in all cases, the standard deviation is small with respect to the dynamical range of the data. However, there are differences parameter to parameter. For instance,  $f_y$  and O/H present very low-standard deviations, of the order of  $\sim 0.05\text{--}0.1$  dex, at almost any location within the diagram, what indicates that indeed the mean values describe well the behaviour of the H II regions. Slightly larger dispersions are observed for  $\log(U)$  and N/O ( $\sim 0.2\text{--}0.3$  dex), both of them distributed in the same way along the explored diagrams. Thus, despite the larger spread, for both parameters the mean value seems to describe well the distributions. Much larger scatters are found for  $n_e$  and  $A_V$ , with values around  $\sim 0.6$  dex, for the first parameter, and  $\sim 0.3\text{--}0.8$  mag, for the second one, respectively. Despite the fact that their dynamical ranges are larger, it is clear that the trends reported based on the mean values are just valid at a first order. Furthermore, in these two parameters, a trend is observed in the standard deviation that appears to be smaller at the upper-left area of the diagram than at the bottom-right one. Those regions correspond to areas with relatively larger electron densities and dust extinction (left-hand panel). Therefore, at this location, the reported trends for  $n_e$  and  $A_V$  should not be interpreted just as an increase of both parameters but as an enlargement of the range of values for both of them. Accordingly, the outlined trends between both parameters and other physical quantities (O/H and, in particular,  $U$ ) discussed in the previous paragraph should be revised, considering that they most probably would present larger scatters as  $n_e$  and  $A_V$  increase.

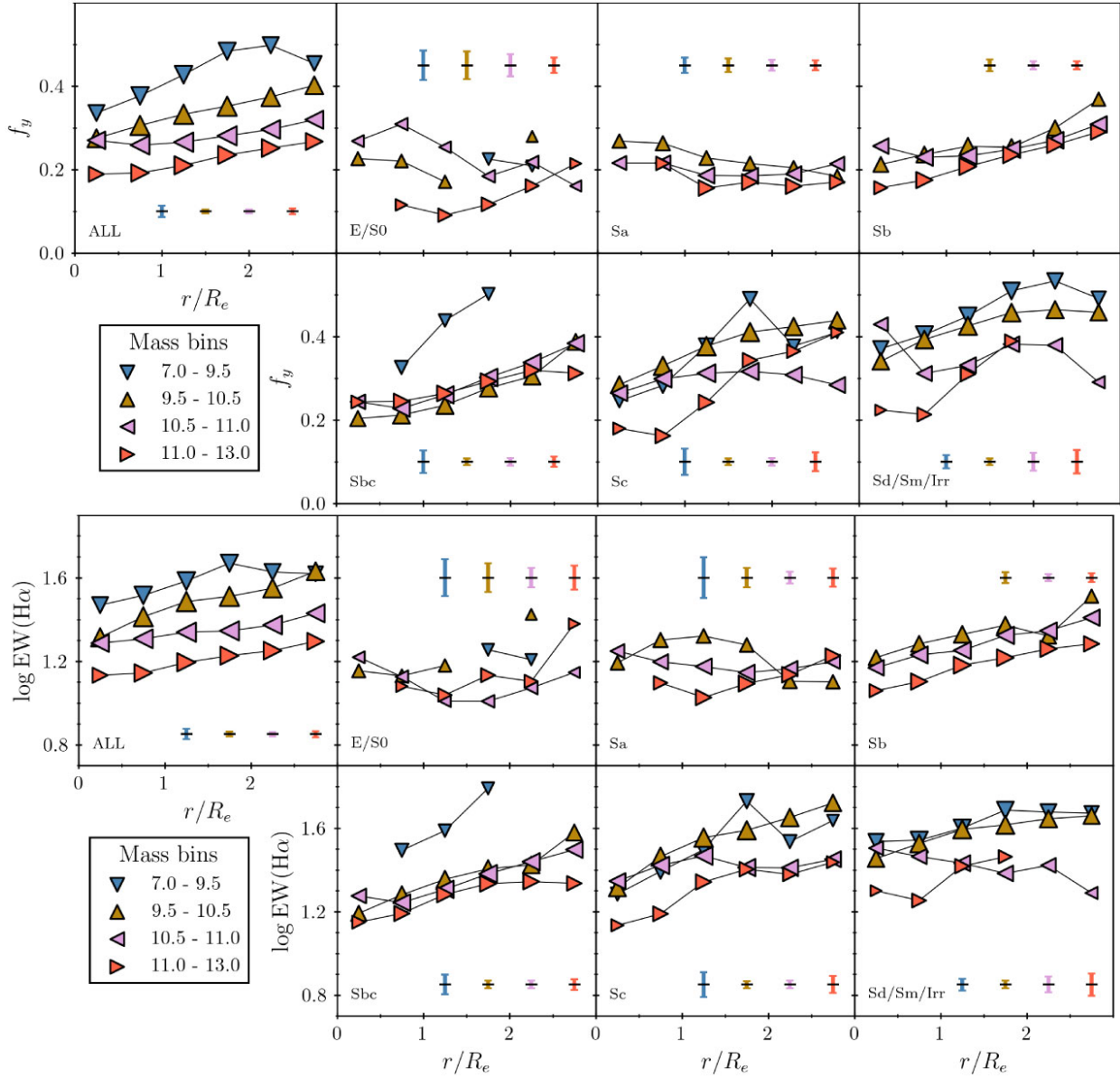
### 3.3 Radial distributions of the physical properties

The exploration of the distribution of the physical properties of the H II regions across the BPT diagram, shown in the previous section, indicates that the location on this diagram and therefore the observed line ratios present clear dependencies with the physical properties. Since many of these physical quantities are tightly connected with the evolution of the stellar populations (in particular  $f_y$ , O/H and N/O but also  $A_V$ ), they should present patterns and variations with the same parameters that affect that evolution: (i) galactocentric distances; (ii) stellar mass; and (iii) morphology of the host galaxies (e.g. García-Benito et al. 2017; Sánchez 2020).

In this section, we present the radial distribution of the properties described in Section 3.1 for galaxies of different stellar mass and morphology. For each H II region, we derive its galactocentric distance, taking into account the effects of inclination. Then, we normalize the distance by effective radius of the corresponding host galaxy ( $R_e$ , derived following Walcher et al. 2014). Then, for each galaxy, we obtain the azimuthal average of each property in radial bins of  $0.5 r/R_e$  (a size that guarantees a sufficient number of H II regions within each bin). Finally, we segregate the galaxies in four bins of stellar mass ( $\log(M_*/M_{\text{sun}} = 7\text{--}9.5, 9.5\text{--}10.5, 10.5\text{--}11.0$ , and  $11.0\text{--}13.0$ ), and six morphological types (E/S0, Sa, Sb, Sbc, Sc, and Sd/Sm/Irr). For each sub-sample, we obtain the average radial distribution of each physical property. Finally, in order to quantify the variation in the scale and gradient of each property, we perform a linear fitting to each average radial distribution.

Figs 2–4 summarize the result of this analysis. For each physical property, we show the average radial profile segregated by the stellar mass of the galaxies for all the sample (first column panels), and





**Figure 2.** Radial distribution of the  $f_y$  (top panels) and  $\text{EW}(\text{H}\alpha)$  (bottom panels) along the galactocentric distance for the H II regions in our catalogue. Each symbol corresponds to the average value of the corresponding parameter in bins of 0.5 effective radius for each stellar mass bin (colour coded) and morphology (different panels) of their host galaxy, with its size being proportional to the number of H II regions in each bin. The solid lines link the different values corresponding to the same stellar mass bin. The average standard error of each parameter in each panel is represented with an error bar.

for each morphological type (right-hand panels). The values of the zero-point ( $c_0$ ), and slopes ( $c_1$ ), derived from the linear regression to each of these radial distributions are presented in Appendix D, listed in Tables D1–D7. The zero-point corresponds to the value of the physical quantity in the central regions of the galaxies while the slope is a gauge of the strength of the radial gradients.

The upper panels of Fig. 2 show the radial profiles for the  $f_y$ . In general, the H II regions in the inner areas of the galaxies show a lower value of  $f_y$  than those regions in the outer ones. In addition, there is a clear trend with the stellar mass. The  $f_y$  is higher for low values of the stellar mass. On the contrary, this fraction is lower for those H II regions located in more massive galaxies. We find a similar trend for  $\text{EW}(\text{H}\alpha)$ , in which radial distributions are shown in the lower panels of Fig. 2. The H II regions with the largest  $\text{EW}(\text{H}\alpha)$  are located in the outer regions of low-mass galaxies, and those with the lowest

values of this parameter are located in the inner regions of high-mass galaxies. These trends have been already explored in previous studies (e.g. González Delgado et al. 2016; Belfiore et al. 2017b; Sánchez et al. 2018). They are the consequences of the higher star formation rates with respect to the already accumulated stellar mass (i.e. higher specific star formation rates, sSFRs) found in the outer regions of low-mass galaxies, in combination with the larger amount of stars formed in earlier star formation episodes in the centre of high-mass galaxies and their lower actual sSFR. In summary, both relations are a consequence of the so-called local-downsizing (Pérez et al. 2013) and the inside-out quenching/ageing (Belfiore et al. 2017b). Those trends with the stellar mass are modulated by the morphology. The few earlier-type galaxies with H II regions show lower values of both  $f_y$  and  $\text{EW}(\text{H}\alpha)$  and a weak negative gradient (i.e. decrease with the galactocentric distance). On the contrary, spiral galaxies present

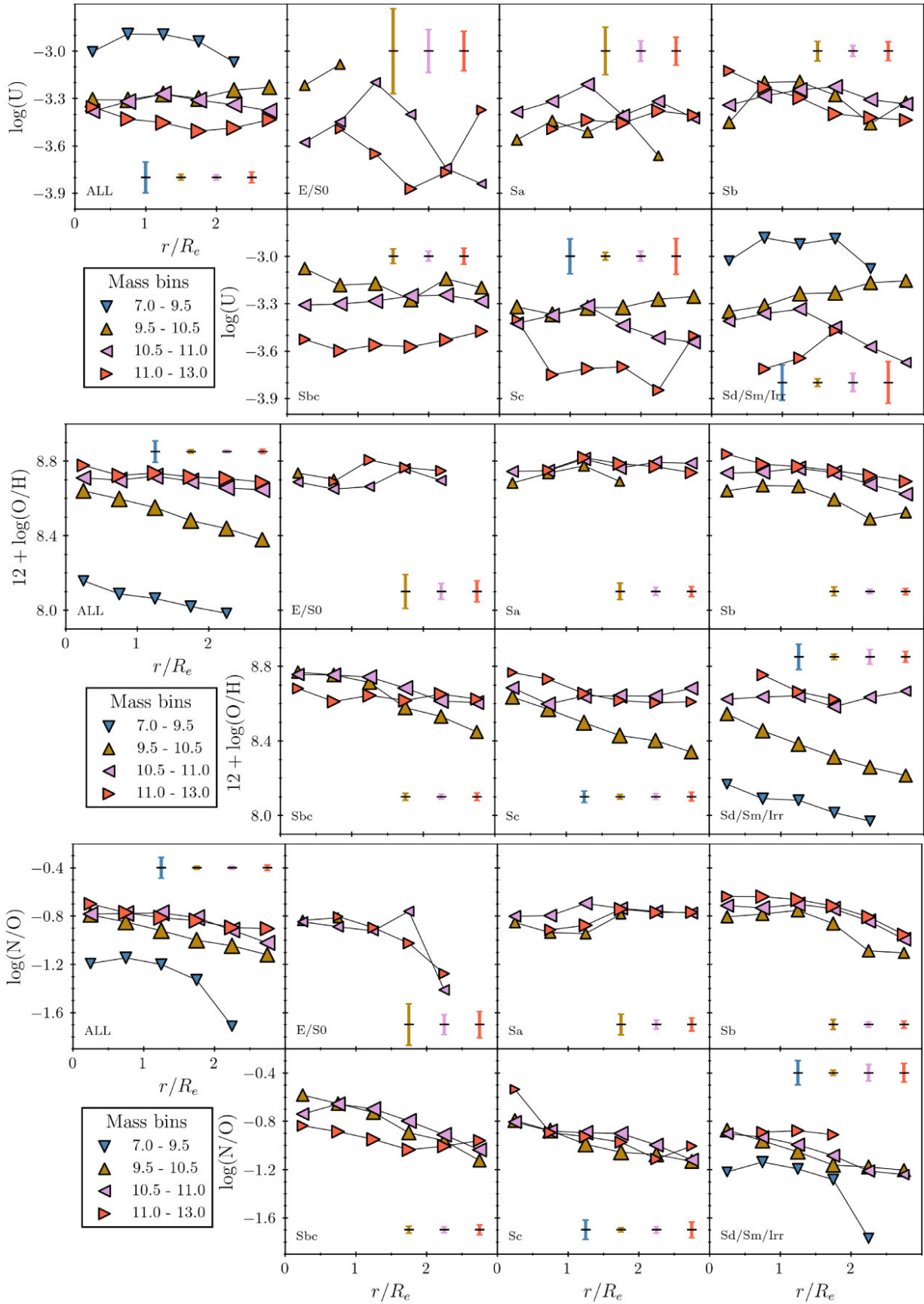
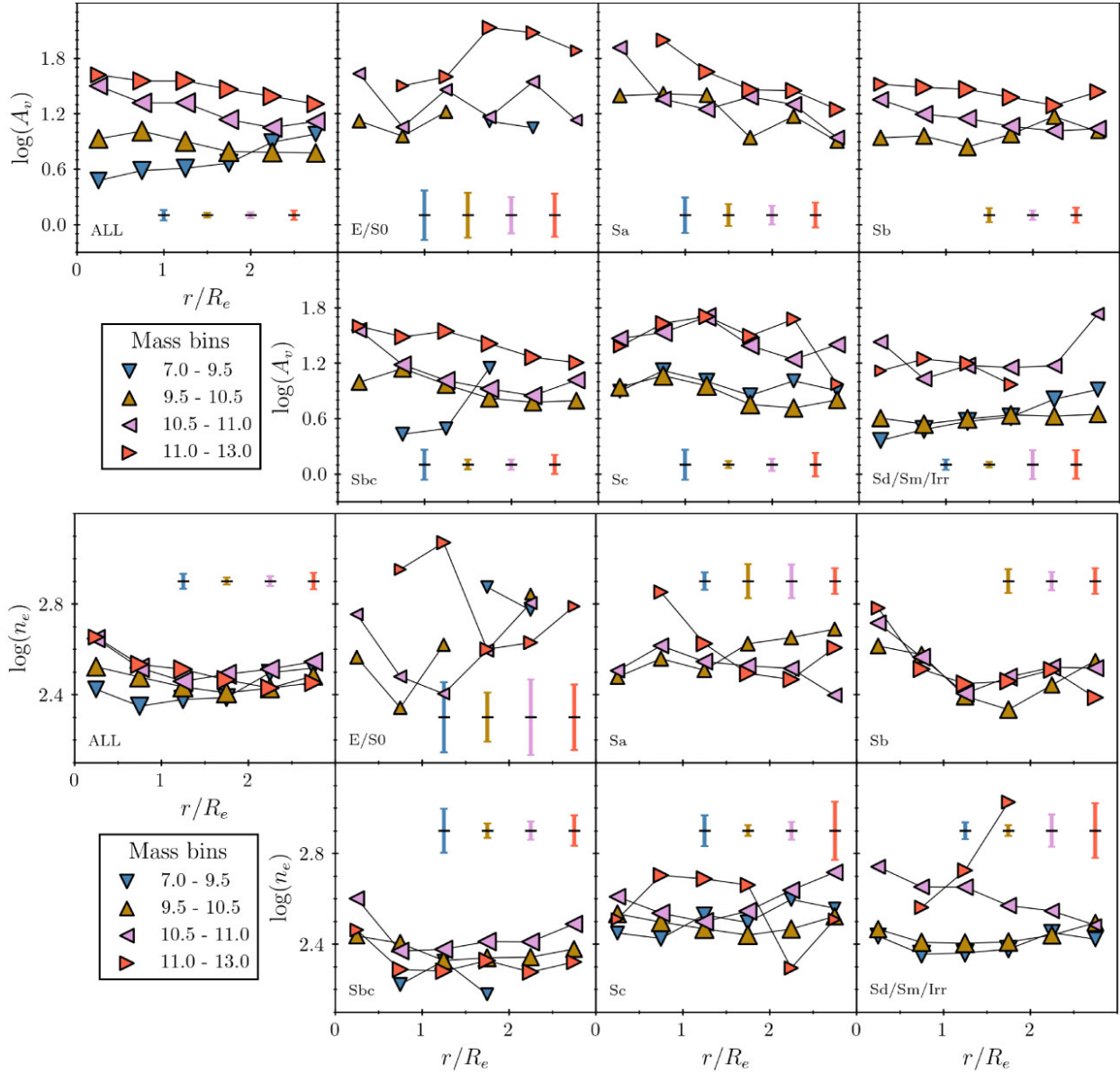


Figure 3. Similar figure as Fig. 2, for the  $\log(U)$  (top panels),  $12 + \log(O/H)$  (middle panels), and  $\log(N/O)$  (bottom panels) parameters.



**Figure 4.** Similar figure as Fig. 2, for the  $n_e$  (top panels), and  $A_V$  (bottom panels) parameters.

higher values and a clear positive gradient. This suggests that the general trend described before is dominated by late-type galaxies, with early-types presenting a slightly different pattern.

The upper panels of Fig. 3 show the radial profiles of the ionization parameter. There is no clear trend of this parameter with the galactocentric distance for any of the explored mass bins. However, the ionization parameter shows a higher value for the lowest stellar mass bin than for the other ones. Although we should note that the mean error associated with this mass bin is the largest one in comparison with other bins. When segregating by morphology, we find a variety of behaviours without a clear pattern. If any, the trend with the stellar mass is much more clear for the very late spiral galaxies (Sd/Sm).

The central panels of Fig. 3 show the radial profiles of the oxygen abundance. There is a clear and well-known trend of this parameter with the galactocentric distance and stellar mass. The H II regions with higher oxygen abundances are located in the inner regions of

the galaxy, and those with lower oxygen abundances are towards the outer zones. Also, there is a clear dependence between oxygen abundance and stellar mass (a consequence of the well-known mass–metallicity relation, e.g. Garnett 2002; Tremonti et al. 2004). Low-mass galaxies generally host the more metal-poor H II regions, and high-mass galaxies host the metal-rich ones. The negative gradient was first described for the Milky Way and galaxies of a similar mass by Searle (1971) and Peimbert et al. (1978), and it has been deeply explored by many previous studies (e.g. Sánchez et al. 2013; Esteban & García-Rojas 2018). The slope of the gradient presents a dependence with the mass as well, with more massive galaxies presenting flatter gradients (as already reported by Belfiore et al. 2017b). Both trends are modulated by the morphology, with early-type galaxies showing an even flatter gradient, with even positive gradients observed (as noticed by Sánchez et al. 2021a). Finally, the range of values for the oxygen abundances is larger in late-type galaxies than in early types. These results speak against a single

characteristic oxygen abundance gradient for all galaxies as proposed by Sánchez et al. (2013), where most of the explored galaxies have similar stellar mass as the Milky Way (for which the range of slopes is rather narrow). Indeed, they are more in agreement with the recent results by Belfiore et al. (2017b), Boardman et al. (2020), and Zinchenko et al. (2021), where it was reported that the slope indeed depends on different properties of the galaxies, including the stellar mass.

A similar trend is found for the radial gradient of the nitrogen-to-oxygen ratio, in which radial profiles are found in the bottom panel of Fig. 3. In average, the H II regions with a high N/O ratio are located in the inner areas of galaxies, and those with lower N/O are located in the outer zones. This result is in agreement with the most recent explorations of those distributions (Pérez-Montero et al. 2016; Belfiore et al. 2017b; Zinchenko et al. 2021). Also, there is a clear dependence of the N/O ratio with the stellar mass. The lower N/O ratios are found on H II regions hosted on low-mass galaxies, and the higher values are on regions in high-mass galaxies. This, again, is a consequence of a global relation between this parameter and the stellar mass (e.g. Schaefer et al. 2020, for a recent study) that can be a consequence of the N/O–O/H relation via the Mass-Metallicity Relation (MZR). A less clear pattern is observed in the slopes of the gradients. For early-type galaxies, they seem to be flatter in average. However, for the more massive ones, it was reported the highest negative gradient of all the sample (although the low number statistic of this particular subsample does not allow us to make a strong statement in this regard).

The top panels of Fig. 4 show the same distributions for the dust extinction. For this parameter, both the stellar mass and the morphology play a clear role. In general, the two more massive bins correspond to the higher values of the dust extinction. In both cases, they present a negative radial gradient (i.e. more dust towards the centre). The highest values are found for the highest mass bin. For the third mass bin, the distribution shows a slight decline with the radius too but less evident than in the other two cases. Finally, for the lowest mass bin, there is a positive gradient, with more dust in outer regions. These trends trace the presence of molecular gas in these galaxies via the metallicity (Brinchmann et al. 2004; Barrera-Ballesteros et al. 2020, 2021b), suggesting that in more massive galaxies, it is clearly concentrated in the central regions. When exploring the dust content by morphology, it is clear that the most massive Sa galaxies, at any galactocentric distances, and E/S0, in the outer regions, are those with the highest dust content. On the other hand, low-mass Sd/Sm/Irr are those with the lowest dust content. Regarding the gradient, clear negative slopes are found for Sa of any mass and for the more massive late-type galaxies, then the slope seems to become shallower and finally positive for the lowest mass latest-type galaxies. The gradient does not present a clear pattern for E/S0 galaxies, maybe due to the limited number of H II regions in these galaxies. If any, they present a flat distribution of  $A_V$ .

Finally, the bottom panels of Fig. 4 show the radial profiles of the electron density. In general, the denser H II regions are located towards the centre of the galaxies. This result is expected since at the central regions of the galaxies, with larger values of the stellar and molecular gas surface densities, the pressure is expected to be larger (e.g. Barrera-Ballesteros et al. 2021a; Barnes et al. 2021). Then, the density declines and rises again in the outer regions. There is a slight trend between the electron density and the stellar mass, with denser regions hosted by high-mass galaxies. Again, this could be a consequence of the effect of the pressure. Indeed, a similar trend is described with the morphology, with earlier galaxies showing the higher values of this quantity. The pattern with the radial trend is

less evident, with slopes covering a wide range of values for any morphology and stellar mass.

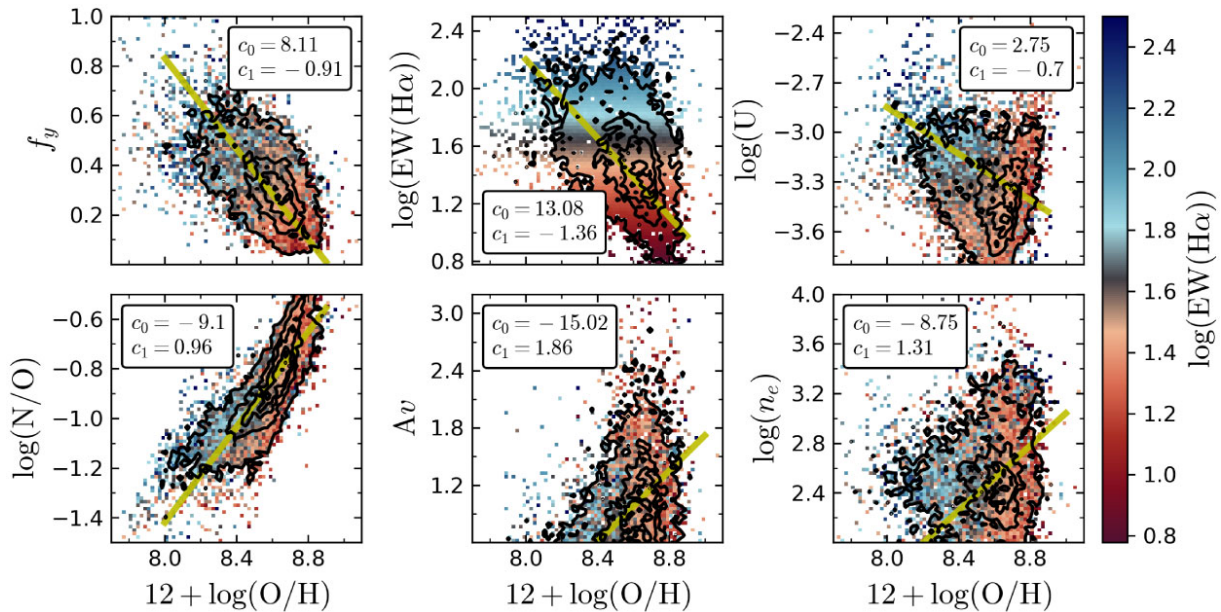
### 3.4 Relation between oxygen abundance and others physical properties of H II regions

So far, we have explored how the physical properties of H II regions distribute along the diagnostic diagrams (Section 3.2 and Appendix C) and along the galactocentric distance (Section 3.3). In this section, we explore the dependence of these parameters among themselves. For doing so, we adopted the oxygen abundance as the primary property, studying how the rest of the parameters correlate (or not) with O/H. We quantify the correlations based on the Pearson correlation coefficient between each pair of explored parameters. To characterize the relation between them, we explore three different linear regressions: (i) fitting Y versus X, (ii) fitting X versus Y, and (iii) deriving the bisector line between both previous relations. We adopt as the best relation the one that minimizes the dispersion of the residuals ( $\sigma$ ). Regardless of the procedure that provides the best fit of the three outlined before, we report the parameters of the relation between the Y- and X-axes, i.e.  $y = c_1x + c_0$  (see Table D8).

Fig. 5 shows the derived distributions for all the parameters described in Section 3.1, where clear trends are appreciated. Regarding  $f_y$  and  $\text{EW}(\text{H}\alpha)$ , they both show a clear decline with the oxygen abundance, with more metal-rich regions corresponding to those having a low fraction of young stars and low values of the equivalent width of  $\text{H}\alpha$ . This is expected based on the distributions described for the BPT diagrams shown in Section 3.2 (and Espinosa-Ponce et al. 2020) and the radial distributions shown in Section 3.3, as well. Regions of high O/H correspond to those located in areas with higher stellar mass density, based on the resolved MZR relation (e.g. Rosales-Ortega et al. 2012; Barrera-Ballesteros et al. 2016), i.e. the regions with the brightest continuum emission. Thus, this is the location where  $f_y$  and  $\text{EW}(\text{H}\alpha)$  is expected to be lower for H II regions ionized by the same kind of stellar clusters with similar  $\text{H}\alpha$  luminosities. In consequence, there is a relative significant correlation between both parameters ( $\rho = -0.58$  and  $-0.46$ , for  $f_y$  and  $\text{EW}(\text{H}\alpha)$  respectively), parametrized by a clear negative slope ( $c_1 = -0.91$  and  $-1.36$ , respectively), and with a small scatter ( $\sigma = 0.12$  and  $0.19$  dex, respectively). Thus, for these two parameters, the reported relations are a pure consequence of the differential stellar/chemical evolution of galaxies in different locations.

The relations of O/H with  $\log(U)$  and the N/O ratio have a different nature. In principle, those relations are not expected to be connected with the properties and evolution of the underlying stellar population in the areas in which the H II regions are formed, like the former two. Early explorations on the properties of H II regions suggest that  $\log(U)$  declines with O/H (e.g. Dopita & Evans 1986), a trend that has been ample adopted in the literature, in particular, when photoionization models are used to derive both parameters (e.g. Thomas et al. 2018). The same trend is found using empirical calibrators of  $\log(U)$  (e.g. Sánchez et al. 2015b). However, using photoionization models, a loose relation is found between both parameters, tighter at high oxygen abundances than at low ones (e.g. Morisset et al. 2016). This trend can be characterized using a linear relation only at first order. Consequently, we found a weaker trend between both parameters as the one described for the other two ones ( $\rho = -0.23$ ), parametrized by a linear relation with a slope of  $c_1 = -0.7$ , and a large standard deviation compared with the dynamical range of  $\log(U)$ ,  $\sigma = 0.2$  dex. Indeed, it is clearly observed that the negative trend is obtained for the fiducial calibrators adopted for both parameters only in the low-abundance regime. At





**Figure 5.** Relation between the oxygen abundances and the others physical properties of H II regions. The oxygen abundance is compared with (i)  $f_y$  (top left); (ii) EW( $H\alpha$ ) (top middle); (iii) ionization parameter (top right); (iv) nitrogen–oxygen ratio (bottom left); (v) dust extinction (bottom middle); and (vi) electron density (bottom right). The contours represent the density distribution, the outermost contour encloses the 90 per cent of the regions, the middle contour encloses the 75 per cent of the regions, and the innermost contour encloses the 60 per cent of the regions.

high abundance, there is an apparent turn-up of  $\log(U)$  that reach high values. Contrary to other distributions explored in this study, we find a strong dependence on the actual adopted calibrators on the observed trends between both parameters. Appendix F includes similar explorations for the different ionization parameter calibrators included in our catalogue of physical properties of H II regions. The expected decline of  $U$  with the abundance is not observed in most of the cases. Only for the  $U$  calibrators reported by Dors et al. (2011) and Morisset et al. (2016), there is a weak trend like the one reported here, with the corresponding up-turn. In other cases, like the values provided by the NB code (Thomas et al. 2018, respectively), the distribution is discrete and does not show a clear trend. In other cases, like HcM (Pérez-Montero 2014), the distribution covers the same range of parameters as the one shown by our fiducial calibrator but with two clear discrete clouds. Finally, for the case of the values reported by the IZI code (Blanc et al. 2015; Mingozzi et al. 2020), the ionization parameter seems to be constant.

The nature of this predicted relation is still broadly unclear (see Ji & Yan 2021, for an extensive discussion). Early explanations involved (i) a correlation between the initial mass function (IMF) and the metallicity, (ii) an effect of the dust absorption, and (iii) an environmental effect in the sense that H II regions that are formed in the inner and outer regions of galaxies present physical differences (e.g. Dopita & Evans 1986). The first of this hypothesis requires that a low metal ISM enhances the upper end of the IMF because of the increased Jeans mass, producing an apparent larger  $\log(U)$ , while metal-rich gas, that cool faster, may result in a lower IMF cutoff and therefore a lower apparent  $\log(U)$ . Recent results have suggested that this IMF-Z relation is indeed present, at least for early-type galaxies (Martín-Navarro et al. 2015). On the other hand, the dust absorption may produce a softer radiation field, producing low-excitation H II regions in dusty environments. If this is the main driver between the  $\log(U)$ –O/H relation, a decrease (increase) is expected of the ionization parameter (oxygen abundance) with the dust extinction. We will get back to that later on in Section 4. Finally,

the environmental scenario suggests that H II regions in the central locations of galaxies are more numerous and they have a smaller average surface brightness (Dopita & Evans 1986). The reason for that could be the enhancement of the cooling and fragmentation induced by the metallicity: molecular clouds in regions of larger metal content would fragment in smaller clouds and collapse to form smaller and lower surface-brightness H II regions. Since both the larger metal content and the highest molecular gas surface density are found in the central regions of galaxies, the  $\log(U)$ –O/H relation would naturally emanate from the two observed trends. However, this was never confirmed. More recently, it has been suggested that blanketing due to metal content may be responsible of this relation (e.g. Sánchez et al. 2015b; Ji & Yan 2021), although it is not clear if this effect may produce a softer or a harder radiation field. All these explanations suggest that  $\log(U)$  is considered as a tracer of the hardness of the ionizing source. However, this is not completely true. The issue is due to the fact that the empirical indicators adopted to trace  $\log(U)$  actually trace the excitation level of the gas that depends on both  $\log(U)$  and the shape of the ionizing source (and also changes if the cloud is matter- or radiation-bounded). In summary, what theory defines as this parameter and what we actually trace may well be two different things. Most probably, there is no simple explanation to this relation, if it exists, and a combination of all the former scenarios is required to describe it. We should note that this correlation is not observed for all calibrators, and using a different combination of them could produce very different results.

The relative nitrogen-to-oxygen abundance presents, as expected, a strong positive correlation with the oxygen abundance ( $\rho = 0.83$ ), as seen in Fig. 5. This distribution is well characterized by a tight ( $\sigma = 0.09$  dex) linear relation of slope  $c_1 = 0.96$  (for the range of abundances covered by the current data set). This well-known relation is a consequence of the differential nature of the oxygen and nitrogen production. Oxygen is a primary element, whose production depends only on physical processes (Matteucci 1986; Maiolino & Mannucci 2019). However, nitrogen has both a primary



and secondary nature. At low metallicity, i.e.  $12 + \log(\text{O}/\text{H}) \lesssim 8$ , it may be primary, and therefore the N/O ratio would be independent of O/H. On the contrary, at high metallicity, i.e.  $12 + \log(\text{O}/\text{H}) \geq 8.2$ , its production depends on the initial abundance of other elements, in particular, oxygen. Thus, in this regime we observe the reported linear relation between both quantities, corresponding to the regime of oxygen abundances covered by our catalogue of H II regions. With the current data, we just appreciate that the linear relation is blended at low metallicity, hinting the reported plateau. This flattening (i.e. constant N/O) is observed at low-O/H H II regions (or galaxies) (e.g. Izotov, Thuan & Guseva 2012; Vincenzo et al. 2016). Recent result has shown that there are secondary trends in this relation, being modulated by the global stellar mass and the local star formation efficiency (SFE, Schaefer et al. 2020), although they could be due to systematics in the adopted calibrators. In our data, a clear segregation of the relation is observed between N/O and O/H as a function of the EW(H $\alpha$ ). This segregation may be connected with the one reported by Schaefer et al. (2020), once considered the relation between EW(H $\alpha$ ) and SFE (e.g. Colombo et al. 2018; Sánchez et al. 2021a). Despite this difference, both results suggest a possible lack of universality of the N/O–O/H relation, or they highlight the effects of the systematics in the adopted calibrators (that were different in both studies).

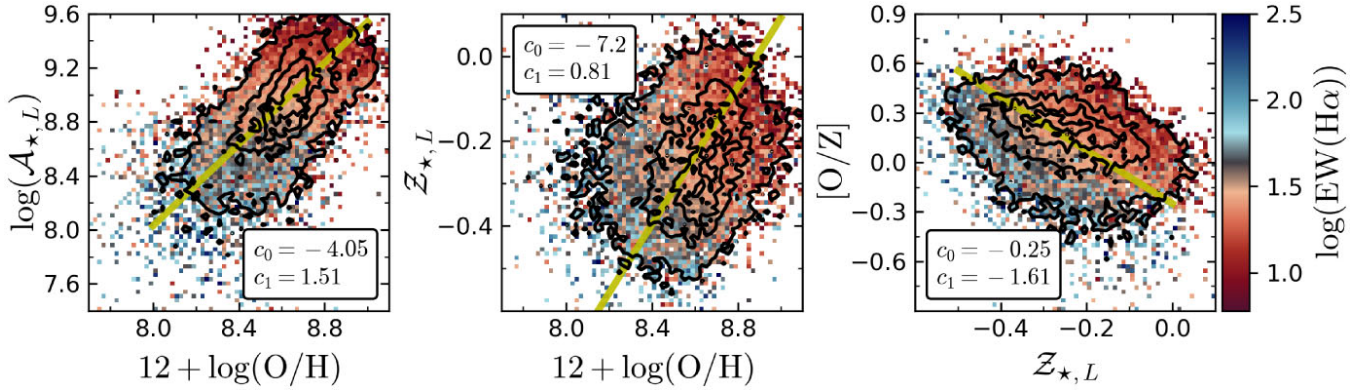
The distribution of the dust extinction of the ionized gas as a function of the oxygen abundance presented in Fig. 5 shows that both parameters do not present a clear correlation ( $c_0 = 0.2$ ). Both parameters present a fuzzy trend that can be characterized by the linear relation with a positive slope of  $c_1 = 1.86$  mag/dex. Consistently, with the weakness of the relation the scatter is considerably larger than the one reported for other relations described between the explored parameters ( $\sigma = 0.36$  mag), apart from the log(U)–O/H one. This weak trend is better appreciated when considering the radial gradients observed for both the oxygen abundance and the dust extinction. Those azimuthal-averaged values for galaxies of the same mass and morphology (Figs 3 and 4) present very similar radial trends that would directly produce a relation between both quantities. Despite this average trend, the dispersion around this relation for individual H II regions is large. Taking into account the direct relation between the dust extinction and the gas column density (Schmidt 1959; Barrera-Ballesteros et al. 2020), this relation may suggest that regions with higher metallicity are those with higher molecular gas surface density, what would be against one of the lines of reasoning by Dopita & Evans (1986) to explain the relation between log(U) and O/H, discussed before. Again, this relation highlights the strong connection between different properties of the H II regions and the local/resolved properties of the location in which they are formed in galaxies. The fact that it is weaker/loose than other ones described here, like the N/O–O/H or the EW(H $\alpha$ )- or  $f_y$ -O/H one, indicates that there is a clear hierarchy between the explored relations. The N/O–O/H relation seems to be more universal (i.e. independent of properties external to the nebulae itself), despite the possible secondary trend reported. On the other hand, the relations with EW(H $\alpha$ ),  $f_y$ , and  $A_V$  seem to be deeply related with the location in the galaxy in which the H II region is formed and the imprints that the evolution and chemical enrichment have left in the properties of the ISM. The fact that the relation with  $A_V$  is the weakest indicates that this relation is strongly modulated by the dependence of both parameters with other physical properties. This is not a total surprise. Based on equation (13) of Brinchmann et al. (2004), the relation between  $A_V$  and O/H should depend at least on the molecular gas surface density and on the dust-to-metal ratio of the ionized gas, two parameters that are clearly non-constant across the extension of a galaxy.

Finally, Fig. 5 shows that the distribution of the electron density is essentially insensitive to the oxygen abundance (bottom-middle panel). Both parameters present the weakest correlation coefficient among the explored ones ( $\rho = 0.16$ ), with a large dispersion around the estimated linear regression ( $\sigma = 0.2$  dex). Contrary to all the remaining parameters, the electron density is clearly unconnected with the oxygen abundance. Furthermore, it seems to be totally insensitive to the described connection with the evolution of the stellar population or the location at which the H II regions are formed, observed in all the properties explored so far.

### 3.5 Relation between the oxygen abundance and physical properties of underlying stellar populations

Most of the relations between the different explored parameters described in Section 3.4 suggest a direct connection between the oxygen abundance of H II and the properties of the surrounding underlying stellar population. Such connection is behind the recently proposed relations between O/H and different local/resolved properties in galaxies, such as the stellar surface mass density (rMZR Rosales-Ortega et al. 2012; Barrera-Ballesteros et al. 2016), the gas fraction, and/or the escape velocity (Barrera-Ballesteros et al. 2018). Based on these results, we explore in this section the relation between the properties of the underlying stellar population and the oxygen abundance for our sample of H II regions.

Fig. 6, left-hand panel, shows the distribution of  $\mathcal{A}_{*,L}$  as a function of O/H. There is a strong correlation between both parameters, with a correlation coefficient of  $\rho = 0.66$ . We estimate the best linear relation between the two parameters by exploring the linear regressions assuming that either the  $\mathcal{A}_{*,L}$  or the O/H is the independent parameter (i.e. fitting Y versus X, or X versus Y), or neither of them (i.e. deriving the bisector line between themselves). The best linear relation for the  $\mathcal{A}_{*,L}$ -O/H distribution results from the fitting of the later (X-axis) versus the former (Y-axis), corresponding to a slope of  $c_1 = 1.51$  (where  $y = c_1x + c_0$ , see Fig. 6), and a standard deviation in the Y-axis of  $\sigma = 0.15$  dex. Since the age of the ionizing stellar populations in an H II region is very similar in all of them, being limited to less than 10 Myr, the reported  $\mathcal{A}_{*,L}$  is essentially tracing the distribution of ages of the underlying stellar population. Therefore, this correlation indicates that the more metal-rich H II regions born in the areas of the galaxies where the stellar population were formed a long time ago (in the cosmological time context). On the contrary, metal-poor H II regions are formed in areas where the stellar population has been formed along a wider range of times, and therefore they present a broader range of ages. Since the oxygen abundance is tightly connected with the amount of stars formed, as indicated before, the reason behind that correlation is that older stellar populations were formed in star formation processes that were stronger and sharper, peaked in earlier cosmological times. This is the basis of local downsizing (e.g. Pérez et al. 2013), a proved scenario in which regions of larger stellar mass density formed the stars earlier and following more abrupt SFHs than regions of smaller stellar mass density (e.g. García-Benito et al. 2017; Sánchez 2020). As a consequence, these locations have both older stellar populations and a higher metal enrichment. Indeed, the distribution of EW(H $\alpha$ ) along this figure follows almost the same distribution of both O/H and  $\mathcal{A}_{*,L}$ , highlighting the fact that regions located in areas with relative low sSFR nowadays (low EW(H $\alpha$ )) are those that have accumulated their stellar mass (and metals) long before, and therefore they have the older stellar populations. On the contrary, areas with relative large sSFR (high EW(H $\alpha$ )) are those that are still actively forming stars, having accumulated their stellar mass (and metals) more recently.



**Figure 6.** Relation between the oxygen abundances and physical properties of underlying stellar populations of H II regions. The oxygen is compared with (i) luminosity-weighted age,  $\mathcal{A}_{*,L}$  (*left-hand panel*), and (ii) luminosity-weighted metallicity,  $\mathcal{Z}_{*,L}$  (*middle panel*) of underlying stellar population. The *right-hand panel* shows the relative oxygen abundance of H II regions with respect to the Z-elements of underlying stellar populations ( $[O/Z]$ ) versus the metallicity of stellar population  $[Z/H]$ . The density contours are the same that are used in Fig. 5.

This scenario implies a connection between the oxygen abundance of the H II regions and the metallicity of the underlying stellar population, since they are the result of the local metal enrichment. Fig. 6, central panel, shows the distribution of these two parameters. As expected, the stellar metallicity,  $\mathcal{Z}_{*,L}$ , presents a very loose trend with O/H ( $\rho = 0.27$ ). Once again, the best linear relation is derived by fitting the oxygen abundance (X-axis) to the  $\mathcal{Z}_{*,L}$  (Y-axis), corresponding to a slope of  $c_1 = 0.81$ , and a standard deviation of  $\sigma = 0.15$  dex. However, the correlation is much weaker than the one found between O/H and  $\mathcal{A}_{*,L}$ . It is worth noticing that  $\mathcal{Z}_{*,L}$  is the average of the metallicities of all the stellar populations in which the observed spectra are decomposed. By construction, it essentially traces iron trapped on the surviving stars. On the other hand, O/H traces the oxygen abundance, i.e. the abundance of an  $\alpha$  element, at the ISM of the H II region. This abundance should be the same as one of the young ionizing stars, recently formed ( $\sim 4$  Myr) from this gas. Therefore, both parameters trace a different type of element at different time-scales and for different families of stars. It is important to recall that iron is mostly produced and expelled to the media by SNIa, being the product of the collapse of binary systems of average massive stars, while oxygen is essentially produced by core-collapsed supernovae produced at the death of very massive stars. Thus,  $\mathcal{Z}_{*,L}$  and O/H may present significant differences in their corresponding chemical enrichment histories (see Maiolino & Mannucci 2019, for a review on the topic). The current LW stellar metallicity depends on both the final stellar mass and the shape of the star formation histories (SFH, e.g. Camps-Fariña et al. 2021). On the other hand, the oxygen abundance follows more tightly the strength rather than the shape of the SFH, being more strongly tied to the final stellar mass (e.g. Lacerda et al. 2020) and the stellar mass surface density (e.g. Gallazzi et al. 2005; González Delgado et al. 2016; Sánchez 2020). The relative weakness of the observed correlation and the large standard deviation suggest that both parameters are connected but without one fully following the other.

Finally, right-hand panel of Fig. 6 shows the distribution of the  $[O/Z]$  ratio along  $\mathcal{Z}_{*,L}$ . This ratio was constructed by subtracting  $\mathcal{Z}_{*,L}$  to the O/H normalized to the solar value (assuming a  $12 + \log(O/H)_{\odot} = 8.69$ ; Asplund et al. 2009), as described in detail in Sánchez et al. (2021b). In this recent letter, we demonstrated that this parameters is a good tracer of the  $\alpha$ -enhancement at the location of H II regions in galaxies. As reported in Sánchez et al. (2021b),  $[O/Z]$  declines with  $\mathcal{Z}_{*,L}$ , following a weak correlation ( $\rho = -0.35$ ) that is better described by the linear regression of the former

parameter (Y-axis) versus the later one (X-axis). This relation shows a slope of  $c_1 = -1.61$ . The dispersion around this trend is similar to the one reported for the relation between O/H and  $\mathcal{Z}_{*,L}$  ( $\sigma = 0.13$  dex), despite the fact that the current correlation is slightly stronger for this one. Furthermore, as explicitly discussed in Sánchez et al. (2021b), the dispersion is modulated by a secondary trend with the  $\mathcal{A}_{*,L}$  of the underlying stellar population that in the current figure is coded by the  $\text{EW}(\text{H}\alpha)$  (based on the correspondence between both parameters discussed before). It is appreciated that a sequence of relations between  $[O/Z]$  and  $\mathcal{Z}_{*,L}$  with a different zero-point (and maybe different slopes) for each value of  $\text{EW}(\text{H}\alpha)$  would produce a significant lower scatter than a single relation for the bulk of H II regions. This is in agreement with the strength of the relation between O/H and  $\mathcal{A}_{*,L}$  reported before that based on the results presented in Sánchez et al. (2021b) and the distribution observed in here is most probably equally modulated by  $[Z/H]$ , highlighting the connection between both parameters. It is worth noticing that these trends are fully compatible with those found for individual stars in the MW (e.g. Hayden et al. 2015) and galaxy-wide integrated populations (e.g. Walcher et al. 2014). The straightforward interpretation of this relation is a consequence of the differential enrichment histories between the  $\alpha$  and iron peak elements (e.g. Hayden et al. 2015; Walcher et al. 2016; Yu et al. 2021). As discussed before, this relation is heavily modulated by the shape of the IMF, the shape of the SFH, or a combination of both, as predicted by the chemical evolution models (e.g. Carigi et al. 2019; Weinberg, Andrews & Freudenberg 2017, and references therein).

We report all parameters present in this section in the Table D9.

## 4 DISCUSSION

Along this study, we have explored the physical properties of the catalogue of  $\sim 26\,000$  H II regions and aggregations extracted from the CALIFA IFS data presented by Espinosa-Ponce et al. (2020). First, we studied how these properties distribute across the diagnostic diagrams (Section 3.2, and Appendix C). As already shown in Espinosa-Ponce et al. (2020), our selection of H II regions, purely based on basic properties of these regions,<sup>2</sup> recovers the well-known distributions along the diagnostic diagrams for these objects. Just

<sup>2</sup>(1) To present a clumpy/peaky shape in the emission line maps; (2) to present a minimum  $f_y$ ; and (3) to present and  $\text{EW}(\text{H}\alpha)$  compatible with that  $f_y$ .

a small fraction of the finally selected regions (<4 percent) are located above the classical demarcation lines proposed by Kewley et al. (2001), based on theoretical considerations and photoionization models. Thus, the proposed selection reproduces the expectations by these models, without requiring a prior selection based on line ratios.

Any set of photoionization models aimed to reproduce the emission line ratios produced by the ionization due to young massive OB stars (e.g. Kewley et al. 2001; Morisset et al. 2016) covers the observed distribution (e.g. Sánchez et al. 2015b). However, it is well known that (i) these models are heavily degenerated, i.e. a set of line ratios (e.g. O3 versus N2) can be reproduced by a different set of physical properties associated with different kind of modelled H II regions (for instance, two values of  $\log(U)$  and  $O/H$ ), and (ii) physically possible photoionization models cover a much wider region of the diagrams than the actually covered by the observed regions. As a consequence, photoionization models, per se, cannot predict univocally the physical properties of these regions, like the oxygen abundance or the ionization parameter, purely based on the line ratios involved in the diagnostic diagrams. However, there are clear trends between the physical parameters and the observed line ratios. In Section 3.2, we show that the actual values of several physical parameters of H II regions are well defined by the location on these diagrams. Although we observed a certain degree of data scattering, in general, it is much lower than the dynamical range of the explored parameter: e.g. Fig. 1 (and also in Figs B1–B3). This result was already reported in Sánchez et al. (2015b), using a smaller sample of H II regions. Furthermore, the connection between different physical properties and the location in the diagnostic diagram are not putative of H II regions. It is observed in other sources of ionization, as a connection between line ratios and the dynamical stage of the underlying stellar population (e.g. Law et al. 2021), the velocity dispersion of the ionized gas (e.g. D’Agostino et al. 2019), or the morphology, the stellar mass, and the galactocentric distance (e.g. Sánchez 2020). Again, none of those trends can be predicted by photoionization models. The reason behind this discrepancy is that photoionization models are physical models of the ionized regions but they do not include the astrophysical context in which the ionization is produced. Thus, given a certain ionizing source (e.g. an OB star), a certain geometry and density distribution of the ionized nebulae, and a certain chemical composition of this nebulae, photoionization models are able to predict a set of line ratios. However, they are unable to determine which type of ionizing source, which type of geometry, or which chemical composition is found at a certain location within a certain type of galaxy. Essentially, they ignore the constrains in which star formation and chemical enrichment took place in galaxies (and within galaxies).

The reported trends along the diagnostic diagrams are not a consequence of the predictions of how the ionization happens. They are the consequence of the spatially resolved evolution of the stellar populations, the consequent chemical enrichment at the location in which H II regions are formed, and maybe the kind of nebulae that can be formed at different locations. Because of this, there are well-defined gradients in all physical properties (e.g.  $O/H$ ,  $N/O$ ,  $A_V$ , as seen in Figs 3 and 4), well described in the literature (e.g. Searle 1971; Peimbert et al. 1978; Vila-Costas & Edmunds 1992; Sánchez et al. 2014; Belfiore et al. 2017b; Esteban & García-Rojas 2018). These gradients have been frequently used to explore the SFH and chemical enrichment histories in other galaxies (e.g. Tissera et al. 2013; Belfiore et al. 2019) and in the Milky Way (e.g. Carigi et al. 2019). However, it is not so frequently acknowledged that they left imprints in the physical properties of H II regions that are fossil records of those processes. This perspective was proposed by

Sánchez et al. (2015b), once the direct connection was observed between the line ratios observed in H II regions and the physical properties of the underlying stellar populations (like the B–V colour; Sánchez et al. 2012a). These results were confirmed by Espinosa-Ponce et al. (2020), where the direct relation was explored between those line ratios and physical properties of the stellar populations, like the  $A_{*,L}$  and  $Z_{*,L}$ .

As indicated before, photoionization models require to be combined with empirical relations between different physical quantities, like the  $N/O$ – $O/H$  and the  $\log(U)$ – $O/H$  relations explored in Section 3.4 (Dopita et al. 2013; Vale Asari et al. 2016) to make predictions on those properties based on the observed line ratios. On the contrary, they may lead to significant errors and misinterpretations. For instance, the average physical properties of those models that reproduce the observed line ratios are not observed previously in the literature or are not in general representative of the real physical properties, if they are heavily degenerated. Indeed, in many cases, they are substantially wrong. The most simple example of this situation is the well-known bi-valuate relation between the R23 line ratio and the oxygen abundance (e.g. Pagel et al. 1979).

The nature of these relations between  $N/O$ ,  $\log(O/H)$ , and  $\log(U)$  is astrophysical as well. As indicated before, the  $N/O$ – $O/H$  relation is a consequence of the primary production of oxygen and the primary and secondary production of nitrogen (Edmunds & Pagel 1978; Alloin et al. 1979; Mollá & Díaz 2005). In particular, the secondary production, which depends on the previous production of oxygen (and carbon), is the mechanism behind the relation above  $12 + \log(O/H) > 8.2$  observed in Fig. 5. The flattening at lower oxygen abundances is due to the primary phases of the production of nitrogen, which is observed when the location in which the H II region is formed has not reached the required oxygen abundance needed to ignite the secondary production in the newly formed stars. Thus, it depends on the location within galaxies, the stellar mass of those galaxies (a gauge of the oxygen abundance), and their morphological type (e.g. Belfiore et al. 2017a, Schaefer et al., in preparation). On the other hand, the nature of the  $\log(U)$ – $O/H$  relation is less clear. Different mechanisms were already proposed in early explorations by Evans & Dopita (1985) and Dopita & Evans (1986), as discussed in Section 3.4. All of them require a connection between this relation and the location in which the H II is formed: (i) a change in the IMF and its connection with the oxygen abundance; (ii) an effect of the dust or the metal blanketing on the hardness of the ionization; and (iii) an environmental bias on which kind of nebulae is formed.

As indicated before, a non-universal or variable IMF and its connection with galaxy (or local) properties is a topic of discussion. Chabrier (2003) reviewed the possible variability of the IMF depending on the environment within a galaxy, although some recent results indicate that this may not be the case (e.g. Wegg, Gerhard & Portail 2017). A possible time evolution from early cosmological times has been proposed too (e.g. Vazdekis et al. 1996), what would explain the differential  $\alpha$ -enhancement in different regions within our Galaxy (e.g. Carigi et al. 2019) or among different galaxies (e.g. Walcher et al. 2014). Since in this period the star formation was dominated by present-day early-type galaxies (e.g. Sánchez et al. 2019), or in the bulge/centre of early spirals (e.g. Li et al. 2015; García-Benito et al. 2019), this time change would imply an environmental dependency of the IMF too. Recent explorations have shown further relations between the shape of the IMF and other galaxy (or regions within galaxies) properties, such as the stellar metallicity (e.g. Martín-Navarro et al. 2015) or the velocity dispersion (McDermid et al. 2015). Those relations may be well interconnected and connected with the time evolution discussed before. In any case,



if this is the cause of the  $\log(U)$ – $O/H$  relation, the connection would be deeply related on how the star formation happens due to how the metallicity enhances the cooling, fragmentation, and collapse of the molecular clouds (Dopita & Evans 1986). Thus, this relation would be a consequence of the enrichment process but only through this physical connection. The possible effect of the dust and/or metal blanketing on the hardness of the radiation has a similar origin, since those regions with stronger metal enrichment would be those suffering this effect more clearly.

The final scenario proposed for the observed  $\log(U)$ – $O/H$  by Dopita & Evans (1986) has a totally different origin. They proposed that H II regions are formed in locations within a galaxy with larger molecular gas densities, i.e. the centre may have a smaller surface brightness. The reason for that could be, one more time, the enhancement of the cooling and fragmentation induced by the metallicity: molecular clouds in regions of larger metal content would fragment in smaller clouds and collapse to form smaller and lower surface-brightness H II regions. Since both the larger metal content and the highest molecular gas surface density are found in the central regions of galaxies, the  $\log(U)$ – $O/H$  relation would naturally emanate from the two observed trends. The local pressure may also play a role in this regard, in addition to the direct effect of the metal content. Radial declines in the pressure are suggested based on the relations proposed by Barrera-Ballesteros et al. (2021a) and the direct estimations reported by Barnes et al. (2021). A high pressure may facilitate the fragmentation and collapse of smaller molecular clouds and the formation of lower surface brightness H II regions. We should not exclude that the  $\log(U)$ – $O/H$  anticorrelation is the consequence of a combined effect of all the reported scenarios that operate at a different level but simultaneously. This could be the reason why the relation is weaker and less well defined as other trends found along this study, due to its multiparametric nature.

Finally, we should acknowledge that not all the properties of the H II regions depend on the star formation and chemical evolution of galaxies. It is obvious that the age of the ionizing population is independent of those processes by definition of these regions (i.e. ionized by young massive OB stars), although their metallicity depends on the evolution. The electron density seems to be decoupled of the evolution processes in galaxies as well, at least to the precision in which we can measure this parameter with our current data set. We should highlight that the average values reported for this parameter ( $n_e \sim 300$ ), although they are similar to the ones reported using IFS data with similar spatial and spectral resolutions (e.g. Sánchez et al. 2012b), are larger than the ones recently reported for similar data with better spatial and spectral resolutions (Barnes et al. 2021). Thus, we need to explore the reported trends with better quality data and other ionic line ratios to understand these discrepancies and confirm (or not) our results.

## 5 CONCLUSIONS

Along this study, we explored the connection between the physical properties of a large catalogue of H II regions extracted from a representative sample of the galaxies in the nearby universe, finding that:

- (i) The location of the H II regions across the classical diagnostic diagrams and, therefore, the values of the line ratios involved on those diagrams are tightly connected with their physical properties and the properties of the underlying stellar populations.
- (ii) This correspondence between observed and physical properties is univocal in most of the cases (within a certain dispersion),

contrary to the predictions of photoionization models when they cover a Cartesian grid in  $O/H$ ,  $N/O$ , age, and  $\log(U)$ . But if additional relations between these parameters are used to restrict the grid, models may recover the observed behaviour.

(iii) Most of the explored physical properties present clear radial gradients with patterns that depend on the mass and morphology of the host galaxies, relations that in many cases have been reported and discussed before in the literature. Once again, those trends are tightly related to those of the underlying stellar populations (e.g.  $f_y$ ).

(iv) The physical properties of the H II regions present clear relations between themselves and with the properties of the underlying stellar populations, which in some cases are more evident when exploring the azimuthal averaged radial patterns discussed before.

In conclusion, we confirm the results from previous explorations suggesting that the astrophysical context in which H II regions are generated is of a fundamental importance in shaping their observed properties. Indeed, most of those properties are the result from the local chemical evolution of the stellar populations at the locations in which those regions are found. In summary, H II regions are a fundamental proxy of the chemical evolution of galaxies that have left clear imprints in their observed properties.

## ACKNOWLEDGEMENTS

We are very grateful with the referee for the comments and suggestions that have improved the quality of this paper.

CM acknowledges support from grant UNAM/PAPIIT – IN101220. We are grateful for the support of the PAPIIT-DGAPA-IG100622 and PAPIIT-DGAPA-IN112620 (UNAM) projects.

RGB acknowledges additional financial support from the State Agency for Research of the Spanish MCIU through the Center of Excellence Severo Ochoa award to the Instituto de Astrofísica de Andalucía (SEV-2017-0709) grants PID2019-109067GB-I00 (MCIU) and P18-FRJ-2595 (Junta de Andalucía).

LG acknowledges financial support from the Spanish Ministerio de Ciencia e Innovación (MCIN), the Agencia Estatal de Investigación (AEI) 10.13039/501100011033, and the European Social Fund (ESF) ‘Investing in your future’ under the 2019 Ramón y Cajal program RYC2019-027683-I and the PID2020-115253GA-I00 HOSTFLOWS project, and from Centro Superior de Investigaciones Científicas (CSIC) under the PIE project 20215AT016.

## DATA AVAILABILITY

The H II region’s catalogue used in this paper is presented in Espinosa-Ponce et al. (2020). The estimated physical properties of H II regions are available online at [http://ifs.astroscu.unam.mx/CALIFA/HII\\_regions/](http://ifs.astroscu.unam.mx/CALIFA/HII_regions/). An example code to read and use the data is available at [https://github.com/cespinosa/HII\\_regions\\_catalog](https://github.com/cespinosa/HII_regions_catalog).

## REFERENCES

- Alloin D., Collin-Souffrin S., Joly M., Vigroux L., 1979, *A&A*, 78, 200
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
- Barnes A. T. et al., 2021, *MNRAS*, 508, 5362
- Barrera-Ballesteros J. K. et al., 2016, *MNRAS*, 463, 2513
- Barrera-Ballesteros J. K. et al., 2018, *ApJ*, 852, 74
- Barrera-Ballesteros J. K. et al., 2020, *MNRAS*, 492, 2651
- Barrera-Ballesteros J. K. et al., 2021a, *MNRAS*, 503, 3643
- Barrera-Ballesteros J. K. et al., 2021b, *ApJ*, 909, 131
- Belfiore F. et al., 2017a, *MNRAS*, 466, 2570

- Belfiore F. et al., 2017b, *MNRAS*, 469, 151
- Belfiore F., Vincenzo F., Maiolino R., Matteucci F., 2019, *MNRAS*, 487, 456
- Blanc G. A., Kewley L., Vogt F. P. A., Dopita M. A., 2015, *ApJ*, 798, 99
- Boardman N. et al., 2020, *MNRAS*, 491, 3672
- Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151
- Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, *ApJ*, 840, 44
- Camps-Fariña A., Sanchez S. F., Lacerda E. A. D., Carigi L., García-Benito R., Mast D., Galbany L., 2021, *MNRAS*, 504, 3478
- Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, 345, 245
- Carigi L., Peimbert M., Peimbert A., 2019, *ApJ*, 873, 107
- Chabrier G., 2003, *PASP*, 115, 763
- Cid Fernandes R., Stasińska G., Schlickmann M. S., Mateus A., Vale Asari N., Schoenell W., Sodré L., 2010, *MNRAS*, 403, 1036
- Cid Fernandes R. et al., 2013, *A&A*, 557, A86
- Cid Fernandes R. et al., 2014, *A&A*, 561, A130
- Cid Fernandes R., Carvalho M. S., Sánchez S. F., de Amorim A., Ruschel-Dutra D., 2021, *MNRAS*, 502, 1386
- Colombo D. et al., 2018, *MNRAS*, 475, 1791
- Curti M., Mannucci F., Cresci G., Maiolino R., 2020, *MNRAS*, 491, 944
- D'Agostino J. J., Kewley L. J., Groves B. A., Medling A., Dopita M. A., Thomas A. D., 2019, *MNRAS*, 485, L38
- Díaz A. I., Castellanos M., Terlevich E., Luisa García-Vargas M., 2000, *MNRAS*, 318, 462
- Dopita M. A., Evans I. N., 1986, *ApJ*, 307, 431
- Dopita M. A., Sutherland R. S., Nicholls D. C., Kewley L. J., Vogt F. P. A., 2013, *ApJS*, 208, 10
- Dors O. L. Jr, Krabbe A., Hägele G. F., Pérez-Montero E., 2011, *MNRAS*, 415, 3616
- Edmunds M. G., Pagel B. E. J., 1978, *MNRAS*, 185, 77P
- Espinosa-Ponce C., Sánchez S. F., Morisset C., Barrera-Ballesteros J. K., Galbany L., García-Benito R., Lacerda E. A. D., Mast D., 2020, *MNRAS*, 494, 1622
- Esteban C., García-Rojas J., 2018, *MNRAS*, 478, 2315
- Evans I. N., Dopita M. A., 1985, *ApJS*, 58, 125
- Flores-Fajardo N., Morisset C., Stasińska G., Binette L., 2011, *MNRAS*, 415, 2182
- Galbany L. et al., 2018, *ApJ*, 855, 107
- Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, *MNRAS*, 362, 41
- García-Benito R. et al., 2017, *A&A*, 608, A27
- García-Benito R., González Delgado R. M., Pérez E., Cid Fernandes R., Sánchez S. F., de Amorim A. L., 2019, *A&A*, 621, A120
- Garnett D. R., 2002, *ApJ*, 581, 1019
- Gomes J. M. et al., 2016, *A&A*, 585, A92
- González Delgado R. M. et al., 2016, *A&A*, 590, A44
- Hayden M. R. et al., 2015, *ApJ*, 808, 132
- Ho I. T., 2019, *MNRAS*, 485, 3569
- Ibarra-Medel H. J. et al., 2016, *MNRAS*, 463, 2799
- Izotov Y. I., Thuan T. X., Guseva N. G., 2012, *A&A*, 546, A122
- Ji X., Yan R., 2021, preprint ([arXiv:2110.00612](https://arxiv.org/abs/2110.00612))
- Kauffmann G. et al., 2003, *MNRAS*, 346, 1055
- Kehrig C. et al., 2012, *A&A*, 540, A11
- Kelz A. et al., 2006, *PASP*, 118, 129
- Kennicutt R. C., Hodge P. W., 1980, *ApJ*, 241, 573
- Kennicutt Robert C. J., Garnett D. R., 1996, *ApJ*, 456, 504
- Kennicutt Robert C. J., Bresolin F., Garnett D. R., 2003, *ApJ*, 591, 801
- Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35
- Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183
- Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, 556, 121
- Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, *MNRAS*, 372, 961
- Kewley L. J., Dopita M. A., Leitherer C., Davé R., Yuan T., Allen M., Groves B., Sutherland R., 2013, *ApJ*, 774, 100
- Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, *ARA&A*, 57, 511
- Kobulnicky H. A., Kewley L. J., 2004, *ApJ*, 617, 240
- Lacerda E. A. D., Sánchez S. F., Cid Fernandes R., López-Cobá C., Espinosa-Ponce C., Galbany L., 2020, *MNRAS*, 492, 3073
- Law D. R. et al., 2021, *ApJ*, 915, 35
- Lequeux J., Peimbert M., Rayo J. F., Serrano A., Torres-Peimbert S., 1979, *A&A*, 80, 155
- Levesque E. M., Kewley L. J., Larson K. L., 2010, *AJ*, 139, 712
- Li C. et al., 2015, *ApJ*, 804, 125
- López-Cobá C., Sánchez S. F., Bland-Hawthorn J., Moiseev A. V., Cruz-González I., García-Benito R., Barrera-Ballesteros J. K., Galbany L., 2019, *MNRAS*, 482, 4032
- López-Cobá C. et al., 2020, *AJ*, 159, 167
- McCall M. L., Rybski P. M., Shields G. A., 1985, *ApJS*, 57, 1
- McDermid R. M. et al., 2015, *MNRAS*, 448, 3484
- Maiolino R., Mannucci F., 2019, *A&AR*, 27, 3
- Marino R. A. et al., 2013, *A&A*, 559, A114
- Martín-Navarro I. et al., 2015, *ApJ*, 806, L31
- Matteucci F., 1986, *MNRAS*, 221, 911
- Mingozzi M. et al., 2020, *A&A*, 636, A42
- Mollá M., Díaz A. I., 2005, *MNRAS*, 358, 521
- Morisset C. et al., 2016, *A&A*, 594, A37
- Moustakas J., Kennicutt R. C. Jr, 2006, *ApJS*, 164, 81
- Nicholls D. C., Dopita M. A., Sutherland R. S., 2012, *ApJ*, 752, 148
- Osterbrock D. E., 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*. University Science Books, Sausalito, CA
- Pagel B. E. J., Edmunds M. G., Blackwell D. E., Chun M. S., Smith G., 1979, *MNRAS*, 189, 95
- Peimbert M., 1967, *ApJ*, 150, 825
- Peimbert M., Torres-Peimbert S., Rayo J. F., 1978, *ApJ*, 220, 516
- Peimbert M., Peimbert A., Delgado-Inglauda G., 2017, *PASP*, 129, 082001
- Pérez E. et al., 2013, *ApJ*, 764, L1
- Pérez-Montero E., 2014, *MNRAS*, 441, 2663
- Pérez-Montero E., 2017, *PASP*, 129, 043001
- Pérez-Montero E. et al., 2016, *A&A*, 595, A62
- Pettini M., Pagel B. E. J., 2004, *MNRAS*, 348, L59
- Pilyugin L. S., Grebel E. K., 2016, *MNRAS*, 457, 3678
- Pilyugin L. S., Mattsson L., 2011, *MNRAS*, 412, 1145
- Pilyugin L. S., Vílchez J. M., Thuan T. X., 2010, *ApJ*, 720, 1738
- Pilyugin L. S., Vílchez J. M., Mattsson L., Thuan T. X., 2012, *MNRAS*, 421, 1624
- Rosales-Ortega F. F., Kennicutt R. C., Sánchez S. F., Díaz A. I., Pasquali A., Johnson B. D., Hao C. N., 2010, *MNRAS*, 405, 735
- Rosales-Ortega F. F., Sánchez S. F., Iglesias-Páramo J., Díaz A. I., Vílchez J. M., Bland-Hawthorn J., Husemann B., Mast D., 2012, *ApJ*, 756, L31
- Roth M. M. et al., 2005, *PASP*, 117, 620
- Sánchez-Menguiano L. et al., 2016, *A&A*, 587, A70
- Sánchez-Menguiano L. et al., 2018, *A&A*, 609, A119
- Sánchez-Menguiano L., Sánchez S. F., Pérez I., Ruiz-Lara T., Galbany L., Anderson J. P., Kuncarayakti H., 2020, *MNRAS*, 492, 4149
- Sánchez S. F., 2006, *Astron. Nachr.*, 327, 850
- Sánchez S. F., 2020, *ARA&A*, 58, 99
- Sánchez S. F., Cardiel N., Verheijen M. A. W., Martín-Gordón D., Vílchez J. M., Alves J., 2007, *A&A*, 465, 207
- Sánchez S. F. et al., 2012a, *A&A*, 538, A8
- Sánchez S. F. et al., 2012b, *A&A*, 546, A2
- Sánchez S. F. et al., 2013, *A&A*, 554, A58
- Sánchez S. F. et al., 2014, *A&A*, 563, A49
- Sánchez S. F. et al., 2015a, *A&A*, 573, A105
- Sánchez S. F. et al., 2015b, *A&A*, 574, A47
- Sánchez S. F. et al., 2016a, *Rev. Mex. Astron. Astrofis.*, 52, 21
- Sánchez S. F. et al., 2016b, *A&A*, 594, A36
- Sánchez S. F. et al., 2018, *Rev. Mex. Astron. Astrofis.*, 54, 217
- Sánchez S. F. et al., 2019, *MNRAS*, 484, 3042
- Sánchez S. F., Walcher C. J., Lopez-Cobá C., Barrera-Ballesteros J. K., Mejía-Narváez A., Espinosa-Ponce C., Camps-Fariña A., 2021a, *Rev. Mex. Astron. Astrofis.*, 57, 3
- Sánchez S. F. et al., 2021b, *A&A*, 652, L10



- Schaefer A. L., Tremonti C., Belfiore F., Pace Z., Bershadsky M. A., Andrews B. H., Drory N., 2020, *ApJ*, 890, L3  
 Schmidt M., 1959, *ApJ*, 129, 243  
 Searle L., 1971, *ApJ*, 168, 327  
 Skillman E. D., Kennicutt R. C., Hodge P. W., 1989, *ApJ*, 347, 875  
 Stasińska G. et al., 2008, *MNRAS*, 391, L29  
 Strömgren B., 1939, *ApJ*, 89, 526  
 Thomas A. D., Dopita M. A., Kewley L. J., Groves B. A., Sutherland R. S., Hopkins A. M., Blanc G. A., 2018, *ApJ*, 856, 89  
 Tissera P. B., Scannapieco C., Beers T. C., Carollo D., 2013, *MNRAS*, 432, 3391  
 Tremonti C. A. et al., 2004, *ApJ*, 613, 898  
 Vale Asari N., Stasińska G., Morisset C., Cid Fernandes R., 2016, *MNRAS*, 460, 1739  
 Vazdekis A., Casuso E., Peletier R. F., Beckman J. E., 1996, *ApJS*, 106, 307  
 Veilleux S., Osterbrock D. E., 1987, *ApJS*, 63, 295  
 Veilleux S., Shopbell P. L., Miller S. T., 2001, *AJ*, 121, 198  
 Vila-Costas M. B., Edmunds M. G., 1992, *MNRAS*, 259, 121  
 Vila-Costas M. B., Edmunds M. G., 1993, *MNRAS*, 265, 199  
 Vincenzo F., Belfiore F., Maiolino R., Matteucci F., Ventura P., 2016, *MNRAS*, 458, 3466  
 Vogt F. P. A., Pérez E., Dopita M. A., Verdes-Montenegro L., Borthakur S., 2017, *A&A*, 601, A61  
 Walcher C. J. et al., 2014, *A&A*, 569, A1  
 Walcher C. J. et al., 2016, *A&A*, 594, A61  
 Wegg C., Gerhard O., Portail M., 2017, *ApJ*, 843, L5  
 Weinberg D. H., Andrews B. H., Freudenburg J., 2017, *ApJ*, 837, 183  
 Yu Z. et al., 2021, *ApJ*, 912, 106  
 Zaritsky D., Kennicutt R. C. Jr, Huchra J. P., 1994, *ApJ*, 420, 87  
 Zinchenko I. A., Vílchez J. M., Pérez-Montero E., Sukhorukov A. V., Sobolenko M., Duarte Puertas S., 2021, *A&A*, 655, A58

## APPENDIX A: DESCRIPTION OF THE CATALOGUE

The catalogue of physical properties of H II regions used along this study is publicly available. It is distributed through the same webpage as the original catalogue of H II regions published

**Table A1.** Description of the distributed files comprising the derived physical properties for the H II regions. Each ionized region is identified by their HIIRGID; therefore, these new tables can be used with the previous ones published by Espinosa-Ponce et al. (2020), which include only the observational properties of the regions.

Column	Key word	Description
COLUMN1	HIIRGID	Ionized region ID
COLUMN2	f <sub>y</sub>	Fraction of young stars to the total luminosity
COLUMN3	dist	Galactocentric distance normalized by $R_e$
COLUMN4	A <sub>v</sub>	Dust extinction
COLUMN5	OH <sub>author</sub> .elines.ftype 1	Oxygen abundance
COLUMN6	OH <sub>author</sub> .elines.ftype 2	Oxygen abundance
⋮	⋮	⋮
COLUMN29	OH <sub>author</sub> .elines.ftype 26	Oxygen abundance
COLUMN30	OH_IZI.value.model 1	Oxygen abundance
⋮	⋮	⋮
COLUMN44	OH_IZI.value.model 15	Oxygen abundance
COLUMN45	U <sub>author</sub> .elines.ftype 1	Ionization parameter (adimensional)
⋮	⋮	⋮
COLUMN52	U <sub>author</sub> .elines.ftype 8	Ionization parameter (adimensional)
COLUMN53	q_IZI.value.model 1	Ionization parameter
⋮	⋮	⋮
COLUMN67	q_IZI.value.model 15	Ionization parameter
COLUMN68	NH <sub>author</sub> .ftype	Nitrogen abundance
COLUMN69	NO_HCm.ftype 1	Nitrogen–oxygen ratio
COLUMN70	NO_HCm.ftype 2	Nitrogen–oxygen ratio
COLUMN71	Ne <sub>Oster</sub> .S	Electronic density ratio
COLUMN72	eA <sub>v</sub>	Error of dust extinction
COLUMN73	eOH <sub>author</sub> .elines.ftype 1	Error of oxygen abundance
⋮	⋮	⋮
COLUMN172	eNe <sub>Oster</sub> .S	Error of electronic density ratio

by Espinosa-Ponce et al. (2020),<sup>3</sup> comprising a set of additional files. Each file, labelled by the name of the host galaxy (HII.GALNAME.phys\_props.fits), contains the full set of parameters derived for each region in a FITS-table format. Each row in the table corresponds to one H II region, labelled with the same ID used in the original H II regions catalogue (HIREGID): the galaxy name (GALNAME) plus the corresponding index of its segmentation map, seg\_Ha.GALNAME.fits.

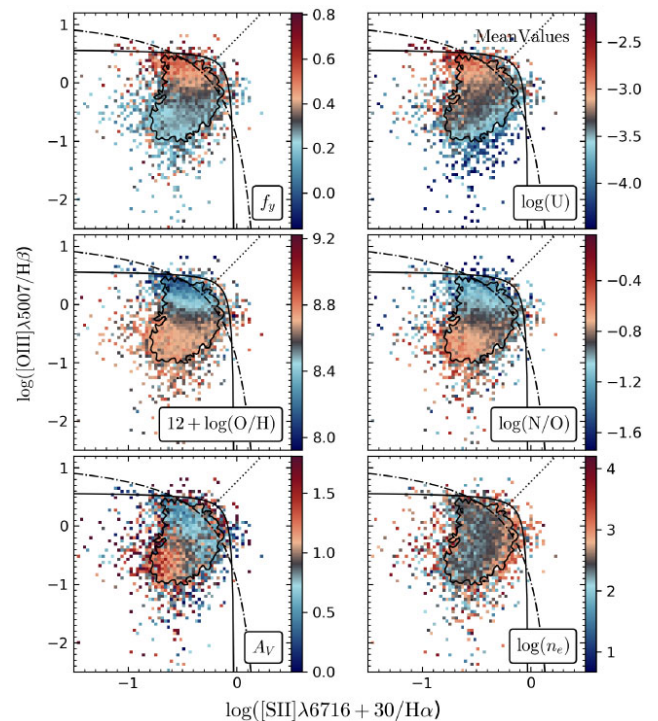
The general structure of each file is shown in Table A1. We include the HIREGID, the fraction of young stars to the total luminosity, the galactocentric distance, and the dust extinction for each object. The next set of properties corresponds to the oxygen abundances, ionization parameters, and nitrogen/nitrogen-to-oxygen abundances derived using several calibrators described in the literature. The key word of each one includes a description of the considered physical property (e.g. OH for O/H) and a reference to the article from which it was taken. Most key words included the line ratios used for the respective calibrations and, if it were the case, the type of calibration (i.e. linear or polynomial fitting).

As indicated before, we include several values of oxygen abundances calculated with different calibrators, all them listed in Table E1. We use the v5.1 of HcM code, the v0.9.9 of NebulaBayes code, and the PYTHON version of IZI code.

For the ionization parameter,  $U$ , we include the results from several studies too: Dors et al. (2011), Morisset et al. (2016, Mor16), Thomas et al. (2018, NB), and Pérez-Montero (2014, HcM). The nitrogen abundance was calculated with the calibration described by Pilyugin & Grebel (2016, Pil16). In the later case, N/O relative abundance was not included; however, it could be calculated with the oxygen and nitrogen abundances described above. In addition, we include the N/O ratio estimated by the HII-CHI-MISTRY code (Pérez-Montero 2014). The electron density described in the Section 3.1 is included in this table too. Finally, we include the complete output of IZI code (Mingozzi et al. 2020) for both the oxygen abundance and ionization parameter,  $q$ . In order to run the IZI code, we use the photoionization models set presented in Dopita et al. (2013, labelled as d13\_k\_20 and d13\_k\_inf), Levesque, Kewley & Larson (2010, labelled as levesque), and Byler et al. (2017, labelled as byler and byler\_CSFR). The description of each of those photoionization grids and their respective results derived by IZI is fully explored in Blanc et al. (2015) and Mingozzi et al. (2020).

## APPENDIX B: PHYSICAL PROPERTIES ON OTHER DIAGNOSTIC DIAGRAMS

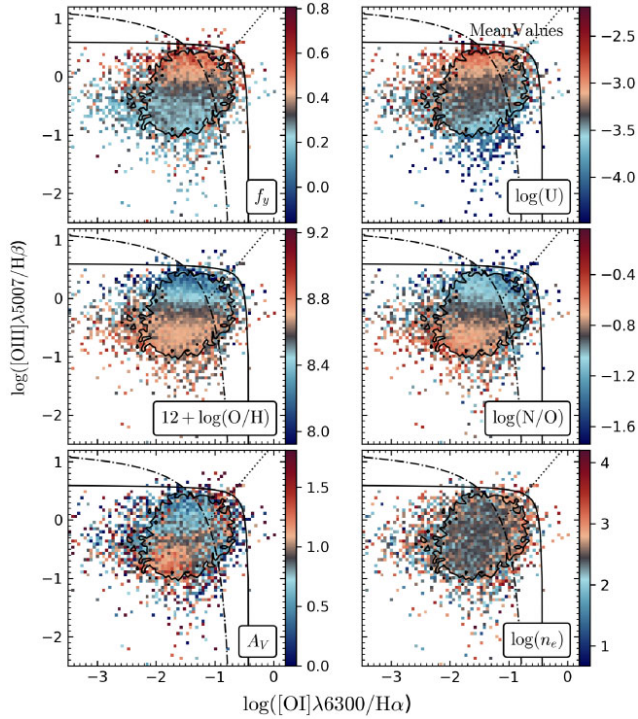
As we show in Section 3.2, the physical properties described in Section 3.1 present evident trends across the BPT diagram. In this appendix, we include the distributions of those properties in other most frequently used diagnostic diagrams: Figs B1 and B2 show the distribution across the  $[\text{OIII}]\lambda 5007/\text{H}\beta$  versus  $[\text{SII}]\lambda 6716 + 30/\text{H}\alpha$  and  $[\text{OIII}]\lambda 5007/\text{H}\beta$  versus  $[\text{OI}]\lambda 6300/\text{H}\alpha$  diagrams, respectively. In addition, we include the  $[\text{OII}]\lambda 3726 + 29/\text{H}\beta$  versus  $[\text{OIII}]\lambda 5007/\text{H}\beta$  diagnostic diagram in Fig. B3. Some trends are clearly highlighted in the observed distributions: (i) H II regions with high  $f_y$  values are always found in the upper range of the three diagrams, with a decreasing gradient along the Y-axis for the first two diagrams, and a diagonal gradient for the third one (i.e. the



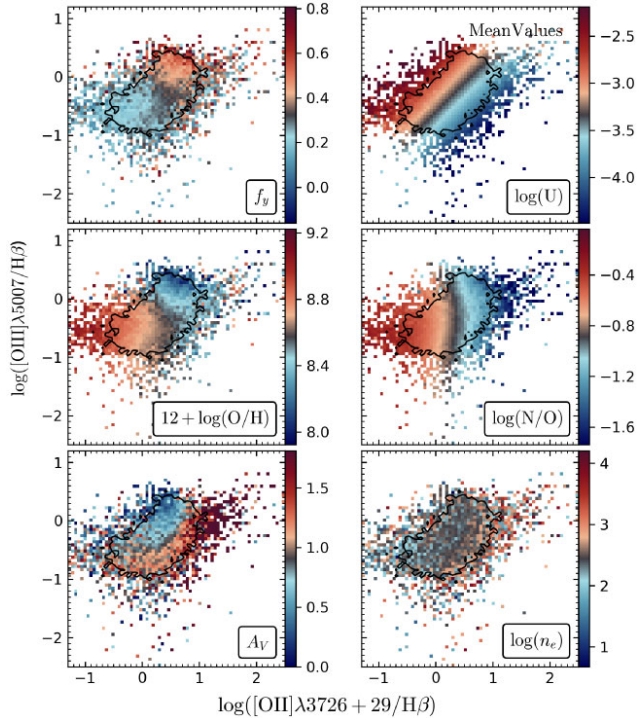
**Figure B1.** Physical properties of the H II regions explored along this study across the  $[\text{OIII}]\lambda 5007/\text{H}\beta$  versus  $[\text{SII}]\lambda 6716 + 30/\text{H}\alpha$  diagnostic diagram. In each panel, the colour code represents one of the physical properties described in the text: *top left-hand panel*: fraction of young stars to the total luminosity,  $f_y$ ; *top right-hand panel*: ionization parameter; *middle left-hand panel*: oxygen abundance; *middle right-hand panel*: nitrogen-to-oxygen abundance; *bottom left-hand panel*: dust extinction; and *bottom right-hand panel*: electron density. The density contours are similar to those shown in Fig. 1. The most frequently used demarcation curves (Kewley et al. 2006) for this diagram are shown. In each panel, the solid line corresponds to the demarcation line proposed by Espinosa-Ponce et al. (2020), the dashed-dotted line to the AGN demarcation line, and the dotted line to the LINER/Sy2 demarcation line.

larger values are at the upper-right end of this diagram); (ii) regions with higher values of  $\log(U)$  are also located in the upper range of the diagrams, with a diagonal distribution from the upper left (high values) to the lower right (low values) in the first two diagrams, and an inverse diagonal distribution from the upper right (high values) to the lower left (low values) for the third diagram; (iii) O/H and N/O present similar trends, as expected due to the strong correlation between both parameters. In the first two diagrams, the gradient is vertical, with both abundances increasing from the upper to the lower values, which in the third diagram the gradient is diagonal (or even horizontal), with abundances declining from the lower-left to the upper-right distribution of this diagrams; (iv) the trends for  $A_V$  and  $n_e$  are the less evident ones. For the dust extinction, there is a clearer diagonal trend in the third diagram, with lower values at the upper-left regime and higher values at the lower-right locations. A similar trend, weaker, but still appreciated, is observed for the  $n_e$  in this diagram too. For the other two diagrams, trends are much weaker or absent. For  $A_V$ , it is observed as a loose trend to lower values in the upper range of the distribution. Finally, no trend is observed for  $n_e$ .

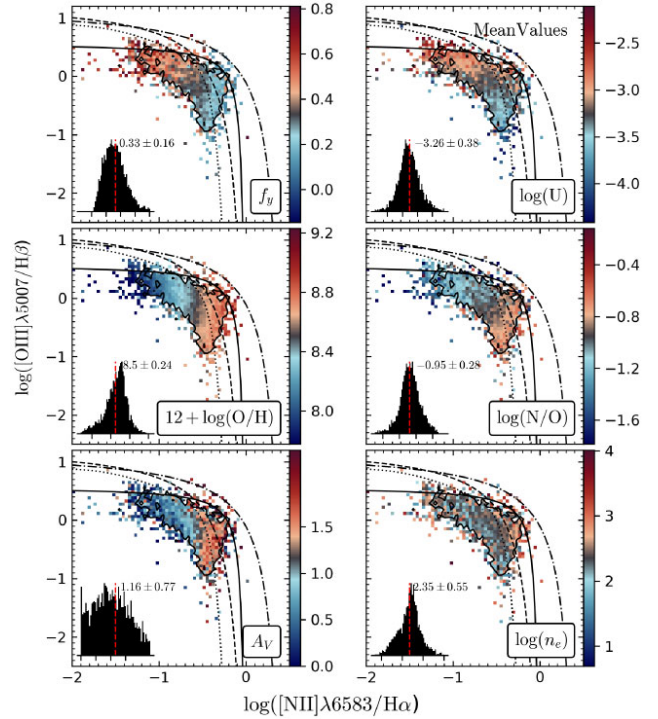
<sup>3</sup>[http://ifs.astroscu.unam.mx/CALIFA/HII\\_regions/](http://ifs.astroscu.unam.mx/CALIFA/HII_regions/)



**Figure B2.** Similar plots as those shown in Fig. C1 for the  $[\text{OIII}]\lambda 5007/\text{H}\beta$  versus  $[\text{OI}]\lambda 6300/\text{H}\alpha$  diagnostic diagram, using the same representation for the parameters, similar colour codes, contours, and meaning for the plotted lines.



**Figure B3.** Similar plots as those shown in Fig. C1 for the  $[\text{OIII}]\lambda 5007/\text{H}\beta$  versus  $[\text{OII}]\lambda 3727/\text{H}\beta$  diagnostic diagram, using the same representation for the parameters, similar colour codes, contours, and meaning for the plotted lines.



**Figure C1.** Similar plots as those shown in Fig. 1 for the  $[\text{OIII}]\lambda 5007/\text{H}\beta$  versus  $[\text{NII}]\lambda 6583/\text{H}\alpha$  diagnostic diagram for the  $\text{H II}$  regions hosted by edge-on galaxies.

## APPENDIX C: PHYSICAL PROPERTIES OF EDGE-ON GALAXIES

In this section, we study how the galaxy inclination affects the trends described for the different physical properties on the diagnostic diagrams. For doing so, we select those  $\text{H II}$  regions in edge-on galaxies, i.e. with an inclination  $>70^\circ$ . There are  $\sim 100$  of these galaxies in which we detect 3600  $\text{H II}$  regions.

As we can see in Fig. C1, we find similar trends for all physical properties as those ones found for the non-edge-on galaxies, shown in Fig. 1. The  $\text{H II}$  regions with high  $f_y$  values are found in the upper-left end of the distribution, with values decreasing towards the lower-right zone of the diagram. The trends of  $\log(U)$ , O/H and N/O present the same trends along the Y-axis. The  $\log(U)$  is higher in those regions on the upper-left of the distribution, and O/H and N/O present a diagonal trend with higher abundances on the upper left of the distribution and lower abundances on the lower right of the diagram. Finally, the previously trends for  $A_V$  and  $n_e$  are found in this sub-sample of the  $\text{H II}$  regions. There is a clear diagonal trend for the dust extinction, with higher values at the lower-right zone of the diagram and lower values at the upper left of the distribution. A weaker trend is observed for the  $n_e$ ; the regions with the highest density are found on the right-hand area of the diagram. The only appreciated difference is that the average dust attenuation is clearly higher for the  $\text{H II}$  regions residing in edge-on galaxies ( $\sim 1.2$  mag versus  $\sim 0.9$  mag). This is in any case expected just due to inclination effects. In summary, the inclusion or exclusion of edge-on galaxies does not affect the main trends described across the diagnostic diagrams.



**APPENDIX D: PHYSICAL PROPERTIES ALONG THE GALACTOCENTRIC DISTANCE: FITTING RESULTS**

Tables D1–D7 show the zero-point ( $c_0$ ) and slopes ( $c_1$ ), and their associated errors, derived from the linear regression to the radial distributions shown in Figs 2–4 and discussed in

Section 3.3. Tables D8 and D9 show the zero-point ( $c_0$ ) and slopes ( $c_1$ ), and their associated errors, derived from the linear regression between the physical properties of H II regions and stellar population, and the oxygen abundance shown in Figs 5 and 6. In addition, we provide with the rms of the residual for each radial distribution once subtracted the best-fitting linear relation.

**Table D1.** Coefficients of the linear fit ( $c_0$  for the zero-point and  $c_1$  for the slope) and their associated errors ( $\sigma_i$ ) for the radial distributions of  $f_y$  along the galactocentric distance shown in Fig. 2. The rms of the residuals is listed for each linear regression.

$f_y M_*$	Parameter	All	E/S0	Sa	Sb	Sbc	Sc	Sd/Sm/Irr
7.0–9.5	$c_0$	0.324	0.284	–	–	0.182	0.218	0.354
	$\sigma_{c_0}$	0.009	0.142	–	–	0.037	0.014	0.011
	$c_1$	0.077	–0.033	–	–	0.194	0.11	0.076
	$\sigma_{c_1}$	0.007	0.073	–	–	0.039	0.014	0.008
	rms <sub>r</sub>	0.0353	~ 0.0	–	–	0.0138	0.0673	0.0306
9.5–10.5	$c_0$	0.268	0.223	0.278	0.203	0.163	0.28	0.351
	$\sigma_{c_0}$	0.004	0.023	0.013	0.009	0.006	0.005	0.007
	$c_1$	0.049	–0.022	–0.035	0.038	0.066	0.068	0.051
	$\sigma_{c_1}$	0.003	0.023	0.008	0.007	0.004	0.004	0.004
	rms <sub>r</sub>	0.003	0.0544	0.0057	0.0263	0.0203	0.0151	0.0186
10.5–11.0	$c_0$	0.243	0.32	0.21	0.21	0.182	0.289	0.353
	$\sigma_{c_0}$	0.004	0.019	0.011	0.008	0.008	0.007	0.019
	$c_1$	0.023	–0.059	–0.01	0.027	0.068	0.009	0.0
	$\sigma_{c_1}$	0.003	0.011	0.007	0.005	0.005	0.005	0.012
	rms <sub>r</sub>	0.0105	0.0275	0.0151	0.0197	0.0192	0.0179	0.0473
11.0–13.0	$c_0$	0.167	0.056	0.169	0.136	0.219	0.081	0.108
	$\sigma_{c_0}$	0.006	0.018	0.014	0.008	0.011	0.015	0.023
	$c_1$	0.037	0.039	–0.002	0.056	0.04	0.13	0.159
	$\sigma_{c_1}$	0.004	0.012	0.007	0.005	0.007	0.012	0.019
	rms <sub>r</sub>	0.0059	0.0279	0.0218	0.0033	0.0103	0.0334	0.0386

**Table D2.** Coefficients of the linear fit ( $c_0$  for the zero-point and  $c_1$  for the slope) and their associated errors ( $\sigma_i$ ) for the radial distributions of EW(H $\alpha$ ) along the galactocentric distance shown in Fig. 2. The rms of the residuals is listed for each linear regression.

log EW(H $\alpha$ ) $M_*$	Parameter	ALL	E/S0	Sa	Sb	Sbc	Sc	Sd/Sm/Irr
7.0–9.5	$c_0$	1.466	1.421	–	–	1.283	1.286	1.496
	$\sigma_{c_0}$	0.02	0.244	–	–	0.079	0.045	0.023
	$c_1$	0.086	–0.095	–	–	0.269	0.169	0.084
	$\sigma_{c_1}$	0.014	0.121	–	–	0.073	0.032	0.016
	rms <sub>r</sub>	0.0438	0.0	–	–	0.0282	0.0954	0.0306
9.5–10.5	$c_0$	1.323	1.092	1.287	1.216	1.18	1.336	1.47
	$\sigma_{c_0}$	0.009	0.047	0.034	0.018	0.016	0.012	0.015
	$c_1$	0.111	0.099	–0.055	0.08	0.128	0.151	0.081
	$\sigma_{c_1}$	0.006	0.048	0.021	0.013	0.01	0.009	0.01
	rms <sub>r</sub>	0.0208	0.0633	0.0736	0.045	0.0283	0.0328	0.023
10.5–11.0	$c_0$	1.269	1.165	1.211	1.153	1.169	1.406	1.51
	$\sigma_{c_0}$	0.009	0.033	0.026	0.015	0.017	0.015	0.034
	$c_1$	0.049	–0.055	–0.022	0.09	0.118	0.013	–0.063
	$\sigma_{c_1}$	0.005	0.022	0.015	0.009	0.01	0.01	0.021
	rms <sub>r</sub>	0.0117	0.076	0.0282	0.0105	0.0319	0.0345	0.0303
11.0–13.0	$c_0$	1.101	1.014	0.924	1.051	1.156	1.127	1.163
	$\sigma_{c_0}$	0.013	0.057	0.033	0.018	0.023	0.033	0.073
	$c_1$	0.069	0.051	0.097	0.091	0.083	0.134	0.179
	$\sigma_{c_1}$	0.007	0.032	0.018	0.01	0.013	0.022	0.055
	rms <sub>r</sub>	0.0082	0.1037	0.0475	0.0141	0.0302	0.0453	0.0537

**Table D3.** Coefficients of the linear fit ( $c_0$  for the zero-point and  $c_1$  for the slope) and their associated errors ( $\sigma_i$ ) for the radial distributions of  $\log(U)$  along the galactocentric distance shown in Fig. 3. The rms of the residuals is listed for each linear regression.

$\log(U) M_*$	Parameter	All	E/S0	Sa	Sb	Sbc	Sc	Sd/Sm/Irr
7.0–9.5	$c_0$	−2.36	–	–	–	–	–	−2.402
	$\sigma_{c_0}$	0.077	–	–	–	–	–	0.083
	$c_1$	0.006	–	–	–	–	–	0.032
	$\sigma_{c_1}$	0.066	–	–	–	–	–	0.073
	rms <sub>r</sub>	0.0863	–	–	–	–	–	0.1019
9.5–10.5	$c_0$	−2.813	−2.767	−3.101	−2.902	−2.606	−2.877	−2.868
	$\sigma_{c_0}$	0.018	0.163	0.07	0.04	0.051	0.025	0.02
	$c_1$	0.035	0.342	0.046	0.08	−0.031	0.051	0.105
	$\sigma_{c_1}$	0.012	0.325	0.091	0.032	0.029	0.016	0.015
	rms <sub>r</sub>	0.0188	0.0	0.1196	0.1547	0.0729	0.0291	0.0178
10.5–11.0	$c_0$	−2.805	−3.033	−2.843	−2.772	−2.782	−2.793	−2.813
	$\sigma_{c_0}$	0.018	0.069	0.047	0.034	0.039	0.033	0.048
	$c_1$	0.0	0.015	0.024	0.01	0.028	−0.087	−0.096
	$\sigma_{c_1}$	0.011	0.073	0.036	0.021	0.021	0.021	0.036
	rms <sub>r</sub>	0.0473	0.2942	0.0858	0.0465	0.031	0.0703	0.098
11.0–13.0	$c_0$	−2.919	−3.035	−3.04	−2.596	−3.122	−3.175	−3.515
	$\sigma_{c_0}$	0.038	0.179	0.11	0.059	0.055	0.118	0.229
	$c_1$	−0.031	−0.095	0.071	−0.147	0.03	−0.053	0.274
	$\sigma_{c_1}$	0.022	0.107	0.058	0.035	0.035	0.074	0.166
	rms <sub>r</sub>	0.0436	0.2406	0.0286	0.0381	0.0353	0.2389	0.0283

**Table D4.** Coefficients of the linear fit ( $c_0$  for the zero-point and  $c_1$  for the slope) and their associated errors ( $\sigma_i$ ) for the radial distributions of  $12 + \log(O/H)$  along the galactocentric distance shown in Fig. 3. The rms of the residuals is listed for each linear regression.

$12 + \log(O/H) M_*$	Parameter	All	E/S0	Sa	Sb	Sbc	Sc	Sd/Sm/Irr
7.0–9.5	$c_0$	8.175	–	–	–	–	–	8.186
	$\sigma_{c_0}$	0.028	–	–	–	–	–	0.035
	$c_1$	−0.094	–	–	–	–	–	−0.099
	$\sigma_{c_1}$	0.03	–	–	–	–	–	0.037
	rms <sub>r</sub>	0.0105	–	–	–	–	–	0.0143
9.5–10.5	$c_0$	8.671	8.75	8.707	8.669	8.85	8.653	8.559
	$\sigma_{c_0}$	0.007	0.066	0.028	0.013	0.016	0.01	0.01
	$c_1$	−0.105	−0.062	−0.011	−0.049	−0.144	−0.119	−0.136
	$\sigma_{c_1}$	0.005	0.129	0.029	0.01	0.01	0.007	0.007
	rms <sub>r</sub>	0.0062	0.0	0.0382	0.0361	0.0288	0.0139	0.0156
10.5–11.0	$c_0$	8.73	8.658	8.749	8.778	8.826	8.631	8.635
	$\sigma_{c_0}$	0.007	0.027	0.017	0.01	0.013	0.013	0.028
	$c_1$	−0.026	0.029	0.021	−0.04	−0.085	0.007	−0.011
	$\sigma_{c_1}$	0.004	0.027	0.011	0.006	0.008	0.008	0.02
	rms <sub>r</sub>	0.0141	0.0307	0.0216	0.0292	0.0282	0.0297	0.0324
11.0–13.0	$c_0$	8.755	8.727	8.818	8.824	8.633	8.748	8.844
	$\sigma_{c_0}$	0.01	0.081	0.031	0.013	0.018	0.023	0.049
	$c_1$	−0.023	0.029	−0.022	−0.047	0.001	−0.068	−0.139
	$\sigma_{c_1}$	0.006	0.055	0.015	0.008	0.011	0.014	0.037
	rms <sub>r</sub>	0.0097	0.0434	0.0202	0.0044	0.0233	0.0354	0.0191



**Table D5.** Coefficients of the linear fit ( $c_0$  for the zero-point and  $c_1$  for the slope) and their associated errors ( $\sigma_i$ ) for the radial distributions of N/O ratio along the galactocentric distance shown in Fig. 3. The rms of the residuals is listed for each linear regression.

$\log(N/O) M_*$	Parameter	All	E/S0	Sa	Sb	Sbc	Sc	Sd/Sm/Irr
7.0–9.5	$c_0$	−1.102	–	–	–	–	–	−1.105
	$\sigma_{c_0}$	0.056	–	–	–	–	–	0.071
	$c_1$	−0.119	–	–	–	–	–	−0.117
	$\sigma_{c_1}$	0.051	–	–	–	–	–	0.062
	rms <sub>r</sub>	0.1606	–	–	–	–	–	0.1893
9.5–10.5	$c_0$	−0.755	−0.85	−0.861	−0.764	−0.5	−0.776	−0.848
	$\sigma_{c_0}$	0.009	0.048	0.034	0.019	0.021	0.012	0.012
	$c_1$	−0.134	0.051	−0.002	−0.066	−0.211	−0.145	−0.156
	$\sigma_{c_1}$	0.007	0.123	0.04	0.017	0.015	0.009	0.01
	rms <sub>r</sub>	0.0074	0.0	0.0566	0.1086	0.0292	0.0301	0.0356
10.5–11.0	$c_0$	−0.725	−0.852	−0.797	−0.667	−0.527	−0.793	−0.858
	$\sigma_{c_0}$	0.011	0.027	0.025	0.016	0.02	0.017	0.038
	$c_1$	−0.068	−0.022	0.039	−0.066	−0.162	−0.086	−0.132
	$\sigma_{c_1}$	0.007	0.039	0.019	0.011	0.013	0.013	0.03
	rms <sub>r</sub>	0.0558	0.213	0.0472	0.0637	0.0837	0.0423	0.022
11.0–13.0	$c_0$	−0.701	−0.544	−0.88	−0.544	−0.824	−0.643	−0.802
	$\sigma_{c_0}$	0.02	0.157	0.067	0.025	0.025	0.056	0.125
	$c_1$	−0.081	−0.308	0.053	−0.113	−0.084	−0.181	−0.052
	$\sigma_{c_1}$	0.012	0.113	0.031	0.016	0.019	0.034	0.084
	rms <sub>r</sub>	0.0129	0.08	0.0482	0.0563	0.0375	0.1048	0.0314

**Table D6.** Coefficients of the linear fit ( $c_0$  for the zero-point and  $c_1$  for the slope) and their associated errors ( $\sigma_i$ ) for the radial distributions of  $n_e$  along the galactocentric distance shown in Fig. 4. The rms of the residuals is listed for each linear regression.

$\log(n_e) M_*$	Parameter	All	E/S0	Sa	Sb	Sbc	Sc	Sd/Sm/Irr
7.0–9.5	$c_0$	2.355	3.12	–	–	2.239	2.405	2.383
	$\sigma_{c_0}$	0.021	0.714	–	–	0.128	0.05	0.024
	$c_1$	0.034	−0.152	–	–	0.001	0.07	0.004
	$\sigma_{c_1}$	0.017	0.347	–	–	0.114	0.036	0.019
	rms <sub>r</sub>	0.049	0.0	–	–	0.0609	0.0431	0.041
9.5–10.5	$c_0$	2.483	2.418	2.438	2.553	2.393	2.498	2.406
	$\sigma_{c_0}$	0.014	0.064	0.068	0.043	0.029	0.02	0.02
	$c_1$	−0.026	0.162	0.086	−0.095	−0.023	−0.015	0.012
	$\sigma_{c_1}$	0.009	0.05	0.041	0.028	0.017	0.013	0.013
	rms <sub>r</sub>	0.0349	0.1272	0.0196	0.1234	0.0301	0.0336	0.0356
10.5–11.0	$c_0$	2.523	2.511	2.603	2.574	2.392	2.468	2.748
	$\sigma_{c_0}$	0.019	0.107	0.069	0.036	0.037	0.033	0.053
	$c_1$	−0.01	0.077	−0.044	−0.047	0.016	0.06	−0.092
	$\sigma_{c_1}$	0.012	0.079	0.041	0.022	0.021	0.021	0.037
	rms <sub>r</sub>	0.0645	0.1595	0.0614	0.1008	0.0893	0.0647	0.0431
11.0–13.0	$c_0$	2.604	3.163	2.756	2.572	2.34	2.906	2.294
	$\sigma_{c_0}$	0.034	0.175	0.082	0.05	0.064	0.112	0.221
	$c_1$	−0.072	−0.222	−0.107	−0.052	−0.025	−0.206	0.37
	$\sigma_{c_1}$	0.019	0.095	0.042	0.03	0.038	0.071	0.173
	rms <sub>r</sub>	0.0339	0.1537	0.1078	0.0979	0.0496	0.1494	0.0503

**Table D7.** Coefficients of the linear fit ( $c_0$  for the zero-point and  $c_1$  for the slope) and their associated errors ( $\sigma_i$ ) for the radial distributions of  $A_v$  along the galactocentric distance shown in Fig. 4. The rms of the residuals is listed for each linear regression.

$\log(A_v) M_*$	Parameter	All	E/S0	Sa	Sb	Sbc	Sc	Sd/Sm/Irr
7.0–9.5	$c_0$	0.42	1.478	–	–	0.166	1.032	0.317
	$\sigma_{c_0}$	0.041	1.527	–	–	0.192	0.141	0.035
	$c_1$	0.172	–0.188	–	–	0.313	–0.048	0.198
	$\sigma_{c_1}$	0.03	0.739	–	–	0.183	0.09	0.028
	rms <sub>r</sub>	0.0564	0.0	–	–	0.2565	0.0925	0.0322
9.5–10.5	$c_0$	1.003	0.892	1.476	0.864	1.135	1.058	0.546
	$\sigma_{c_0}$	0.022	0.154	0.094	0.05	0.05	0.035	0.022
	$c_1$	–0.1	0.249	–0.212	0.043	–0.152	–0.139	0.037
	$\sigma_{c_1}$	0.014	0.186	0.055	0.037	0.029	0.02	0.016
	rms <sub>r</sub>	0.0477	0.5836	0.1325	0.0979	0.0778	0.0863	0.027
10.5–11.0	$c_0$	1.476	1.374	1.616	1.295	1.27	1.591	1.059
	$\sigma_{c_0}$	0.029	0.152	0.086	0.044	0.043	0.051	0.125
	$c_1$	–0.17	–0.079	–0.216	–0.119	–0.177	–0.092	0.076
	$\sigma_{c_1}$	0.017	0.091	0.053	0.027	0.026	0.033	0.083
	rms <sub>r</sub>	0.0635	0.2236	0.1811	0.0464	0.1685	0.1195	0.2395
11.0–13.0	$c_0$	1.685	1.312	2.03	1.548	1.674	1.629	1.3
	$\sigma_{c_0}$	0.049	0.256	0.161	0.077	0.093	0.106	0.155
	$c_1$	–0.129	0.312	–0.291	–0.088	–0.165	–0.071	–0.146
	$\sigma_{c_1}$	0.029	0.151	0.083	0.045	0.056	0.065	0.122
	rms <sub>r</sub>	0.026	0.1908	0.0941	0.0601	0.0483	0.237	0.0971

**Table D8.** Coefficients of the linear fit ( $c_0$  for the zero-point and  $c_1$  for the slope) and their associated errors ( $\sigma_i$ ) for the distributions shown in Fig. 5, where each of the explored physical properties is fitted as a function of the oxygen abundance (i.e.  $\text{param} = c_0 + c_1(12 + \log(O/H))$ ).

Parameter	$f_y$	EW(H $\alpha$ )	$\log(U)$	$\log(N/O)$	$A_v$	$\log(n_e)$
$c_0$	8.11	13.08	2.75	–9.1	–15.02	–8.75
$\sigma_{c_0}$	1.87	4.52	2.39	1.74	8.76	2.87
$c_1$	–0.91	–1.36	–0.7	0.96	1.86	1.31
$\sigma_{c_1}$	0.23	0.63	0.3	0.2	1.38	0.42

**Table D9.** Coefficients of the linear fit ( $c_0$  for the zero-point and  $c_1$  for the slope) and their associated errors ( $\sigma_i$ ) for the distributions shown in Fig. 6, where each of the LW ages, metallicities, and O/Z are fitted as a function of the oxygen abundance (i.e.  $\text{param} = c_0 + c_1(12 + \log(O/H))$ ).

Parameter	$\log(\text{Age/yr})$	[Z/H]	[O/Z]
$c_0$	–4.05	–7.2	–0.25
$\sigma_{c_0}$	1.5	1.58	0.09
$c_1$	1.51	0.81	–1.61
$\sigma_{c_1}$	0.23	0.18	0.45

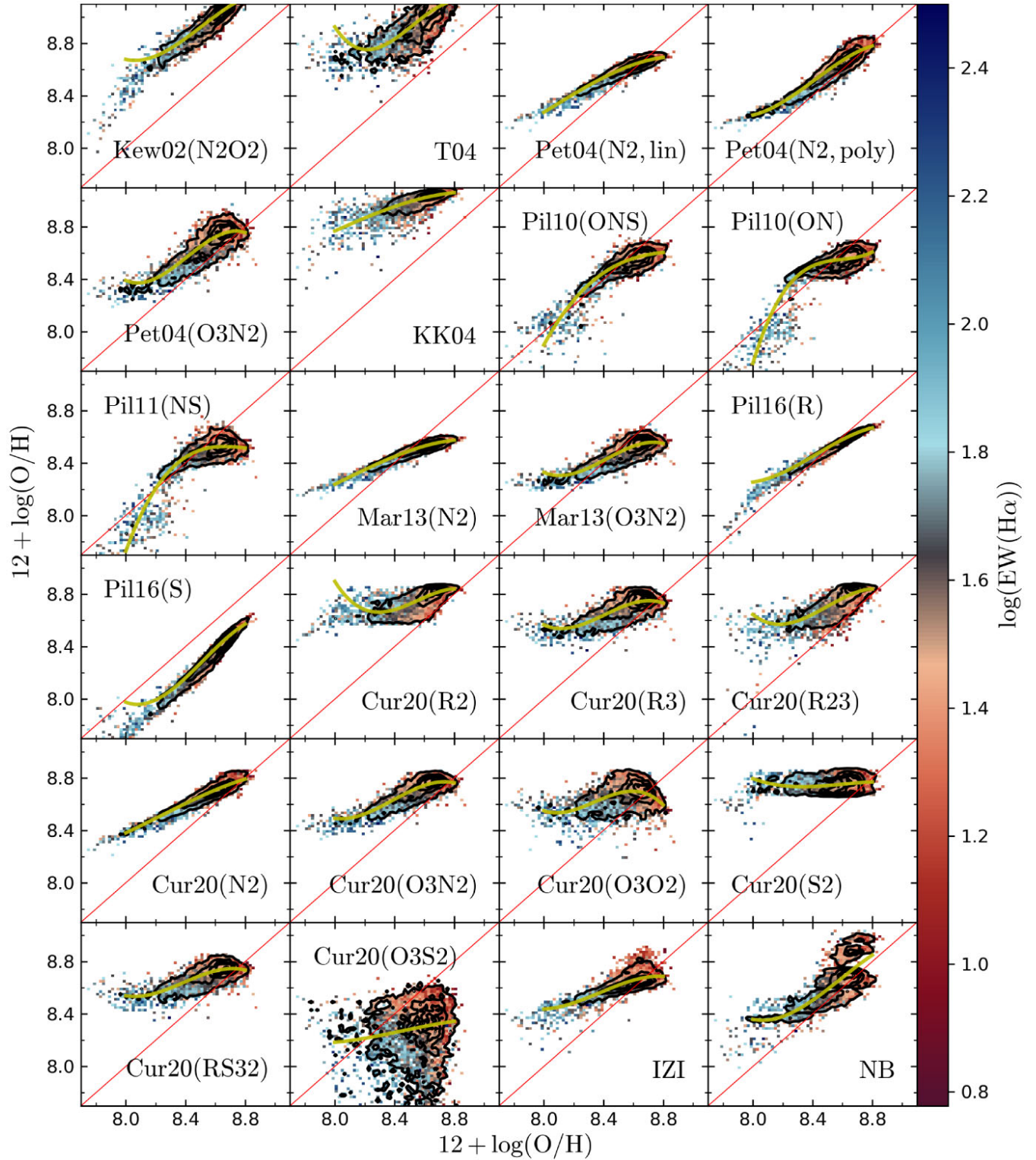
**APPENDIX E: COMPARISON BETWEEN SEVERAL OXYGEN ABUNDANCE CALIBRATORS**

Along this work, we adopt the calibration developed by Ho (2019) for the oxygen abundance as our fiducial one. However, we calculated the oxygen abundance with several calibrations proposed in the literature. The complete list of calibrators included in the final catalogue is listed in Table E1. In this table, we indicate the identification code (ID) adopted to label each calibrator and the emission lines involved. Also, we include if the calibrator is based on observational data (empirical) or photoionization models grid (theoretical). Fig. E1 shows, for each calibrator, the distribution of oxygen abundances for each H II region as a function of the fiducial one. It is appreciated that there are significant differences in the derived values for each adopted

calibrator, although in most of the cases there is monotonic positive trend between the two values. Thus, although many of the results discussed along this article would change quantitatively depending on the adopted O/H calibrator, almost none of them would change from a qualitative point of view. In some cases, the correspondence between both estimations of the abundance is well represented by a first-order polynomial function (e.g. Pil16\_S), while in other cases a higher order polynomial function is required to transform one estimation to the other (e.g. Cur20(O3N2)). In order to provide with the broadest possible transformation among the different calibrators, we fit each distribution using a first- (linear), second-, and third-order polynomial function. The results of this analysis are listed in Tables E2–E4. We include the values and errors of each coefficient of the adopted polynomial together with the standard deviations of the residuals.

**Table E1.** List of oxygen abundances calculated for the H II regions sample.

ID	Emission lines	Calibration type	Reference
OH_Kew02_N2O2	[N II] $\lambda$ 6583/[O II] $\lambda$ 3727	Theoretical	Kewley & Dopita (2002)
OH_T04	[N II] $\lambda$ 6548/H $\alpha$ , $R_{23}$	Empirical	Tremonti et al. (2004)
OH_Pet04_N2_lin	[N II] $\lambda$ 6548/H $\alpha$	Empirical	Pettini & Pagel (2004)
OH_Pet04_N2_poly	[N II] $\lambda$ 6548/H $\alpha$	Empirical	Pettini & Pagel (2004)
OH_Pet04_O3N2	O3N2	Empirical	Pettini & Pagel (2004)
OH_KK04	$R_{23}$ , [O III]/[O II]	Theoretical	Kobulnicky & Kewley (2004)
OH_Pil10_ONS	[N II] $\lambda$ 6583/H $\beta$ , $R_2$ , $R_3$ , $P$ , [S II] $\lambda$ $\lambda$ 6717, 31	Empirical	Pilyugin et al. (2010)
OH_Pil10_ON	[N II] $\lambda$ 6583/H $\beta$ , $R_2$ , $R_3$ , [S II] $\lambda$ $\lambda$ 6717, 31	Empirical	Pilyugin et al. (2010)
OH_Pil11_NS	[N II] $\lambda$ 6583/H $\beta$ , $R_3$ , [S II] $\lambda$ $\lambda$ 6717, 31	Empirical	Pilyugin & Mattsson (2011)
OH_M13_N2	[N II] $\lambda$ 6548/H $\alpha$	Empirical	Marino et al. (2013)
OH_M13_O3N2	O3N2	Empirical	Marino et al. (2013)
OH_Pil16_R	[N II] $\lambda$ 6583/H $\beta$ , $R_2$ , $R_3$	Empirical	Pilyugin & Grebel (2016)
OH_Pil16_S	[N II] $\lambda$ 6583/H $\beta$ , $R_3$ , [S II] $\lambda$ $\lambda$ 6717, 31	Empirical	Pilyugin & Grebel (2016)
OH_Cur20_R2	$R_2$	Empirical	Curti et al. (2020)
OH_Cur20_R3	$R_3$	Empirical	Curti et al. (2020)
OH_Cur20_R23	$R_{23}$	Empirical	Curti et al. (2020)
OH_Cur20_N2	[N II] $\lambda$ 6583/H $\alpha$	Empirical	Curti et al. (2020)
OH_Cur20_O3N2	$R_3$ , [N II] $\lambda$ 6583/H $\alpha$	Empirical	Curti et al. (2020)
OH_Cur20_O3O2	[O III] $\lambda$ 5007, [O II] $\lambda$ 3727 + 29	Empirical	Curti et al. (2020)
OH_Cur20_S2	[S II] $\lambda$ $\lambda$ 6717, 31	Empirical	Curti et al. (2020)
OH_Cur20_RS32	$R_3$ + [S II] $\lambda$ $\lambda$ 6717, 31/H $\alpha$	Empirical	Curti et al. (2020)
OH_Cur20_O3S2	$R_3$ /[S II] $\lambda$ $\lambda$ 6717, 31/H $\alpha$	Empirical	Curti et al. (2020)
OH_Ho	$R_2$ , $R_3$ , [N II] $\lambda$ 6583/H $\beta$ , [S II] $\lambda$ $\lambda$ 6717, 31/H $\alpha$	Empirical	Ho (2019)
OH_NB	[O II] $\lambda$ $\lambda$ 3727, 29, [O III] $\lambda$ 5007, [O I] $\lambda$ 6100, [N II] $\lambda$ 6583, [S II] $\lambda$ $\lambda$ 6717, 31, H $\alpha$ , H $\beta$	Theoretical	Thomas et al. (2018)
OH_HCm_(interp/no_interp)	[O III] $\lambda$ 4959, [O III] $\lambda$ 5007, [N II] $\lambda$ 6584, [S II] $\lambda$ 6725	Theoretical	Pérez-Montero (2014)
OH_IJL_value_models	[O II] $\lambda$ 3727, [O III] $\lambda$ $\lambda$ 4959 + 5007, [N II] $\lambda$ 6583, 48, [S II] $\lambda$ $\lambda$ 6717, 31, H $\alpha$ , H $\beta$	Theoretical	Mingozzi et al. (2020)



**Figure E1.** Distribution of the oxygen abundance derived using each of the adopted calibrators and codes (one for each panel) as a function of the fiducial one (Ho 2019) for our sample of H II regions. We show the best third-order polynomial relation for each calibrator. In each panel, the colour code represents the  $\text{EW}(\text{H}\alpha)$  with a scale ranging between  $\pm 3\sigma$  range around its mean value. The contours represent the density distribution, the outermost contour enclosing 95 percent of the regions, and each consecutive one enclosing 20 percent fewer regions. The red line represents the one-to-one relation and the yellow line corresponds to the fitted third-order polynomial relation between the two parameters.

**Table E2.** Coefficients of the best-fitting linear relation between each different calibrator of the oxygen abundance ( $y$ ) and the fiducial one (Ho 2019) ( $x$ ): i.e.  $y = c_0 + c_1x$ . For each calibrator, we include the values of the coefficients and an estimation their errors together with the standard deviation of the residuals ( $res_{std}$ ).

ID	Kew_02_N202		T04		Pet04_N2_lin		Pet04_N2_poly		Pet04_O3N2		KK04	
$c_0$	2.203	0.194	3.085	0.385	4.6	0.254	2.574	0.16	3.678	0.319	6.432	0.155
$c_1$	0.79	0.023	0.687	0.045	0.467	0.029	0.707	0.019	0.581	0.037	0.3	0.018
$res_{std}$	0.052		0.095		0.029		0.044		0.076		0.051	
	Pil10_ONS		Pil10_ON		Pil11_NS		M13_N2		M13_O3N2		Pil16_R	
$c_0$	2.784	0.954	2.321	1.419	3.77	1.858	5.168	0.096	5.068	0.19	3.567	0.204
$c_1$	0.668	0.111	0.723	0.165	0.548	0.215	0.389	0.011	0.399	0.022	0.581	0.024
$res_{std}$	0.069		0.092		0.103		0.023		0.051		0.028	
	Pil16_S		Cur20_R2		Cur20_R3		Cur20_R23		Cur20_N2		Cur20_O3N2	
$c_0$	-0.478	0.123	5.741	0.348	5.996	0.128	4.908	0.218	4.382	0.131	5.279	0.177
$c_1$	1.029	0.014	0.352	0.04	0.314	0.015	0.447	0.025	0.502	0.015	0.399	0.021
$res_{std}$	0.039		0.062		0.062		0.074		0.031		0.05	
	Cur20_O3O2		Cur20_S2		Cur20_RS32		Cur20_O3S2		IZI		NB	
$c_0$	7.916	0.657	8.245	0.193	6.165	0.123	6.553	0.29	5.793	0.446	2.453	0.641
$c_1$	0.084	0.077	0.059	0.023	0.295	0.014	0.203	0.034	0.331	0.052	0.724	0.075
$res_{std}$	0.089		0.048		0.056		0.208		0.053		0.108	

**Table E3.** Coefficients of the best second-order polynomial relation between each different calibrator of the oxygen abundance ( $y$ ) and the fiducial one (Ho 2019) ( $x$ ): i.e.  $y = c_0 + c_1x + c_2x^2$ . For each calibrator, we include the values of the coefficients and an estimation their errors together with the standard deviation of the residuals ( $res_{std}$ ).

ID	Kew_02_N202		T04		Pet04_N2_lin		Pet04_N2_poly		Pet04_O3N2		KK04	
$c_0$	-12.074	7.935	-5.993	24.386	-27.595	6.793	-16.071	16.597	-53.803	24.601	-10.024	6.401
$c_1$	4.134	1.859	2.775	5.681	8.044	1.584	5.076	3.874	14.068	5.74	4.162	1.494
$c_2$	-0.196	0.109	-0.12	0.331	-0.445	0.092	-0.256	0.226	-0.791	0.335	-0.227	0.087
$res_{std}$	0.051		0.098		0.028		0.047		0.084		0.051	
	Pil10_ONS		Pil10_ON		Pil11_NS		M13_N2		M13_O3N2		Pil16_R	
$c_0$	-70.712	18.553	-107.157	46.237	-121.66	26.435	-18.786	5.615	-30.287	17.76	-21.729	7.248
$c_1$	18.045	4.347	26.635	10.788	30.089	6.152	6.042	1.309	8.71	4.142	6.512	1.687
$c_2$	-1.026	0.255	-1.532	0.629	-1.738	0.358	-0.333	0.076	-0.488	0.241	-0.348	0.098
$res_{std}$	0.065		0.072		0.075		0.023		0.055		0.022	
	Pil16_S		Cur20_R2		Cur20_R3		Cur20_R23		Cur20_N2		Cur20_O3N2	
$c_0$	-16.867	11.752	12.493	13.975	-35.367	21.256	8.742	13.683	-8.763	5.266	-39.595	16.713
$c_1$	4.853	2.735	-1.241	3.253	10.028	4.953	-0.479	3.192	3.595	1.235	10.929	3.894
$c_2$	-0.223	0.159	0.094	0.189	-0.57	0.288	0.056	0.186	-0.182	0.072	-0.617	0.227
$res_{std}$	0.045		0.061		0.068		0.073		0.031		0.056	
	Cur20_O3O2		Cur20_S2		Cur20_RS32		Cur20_O3S2		IZI		NB	
$c_0$	-78.438	24.358	17.234	7.022	-31.068	19.236	3.48	14.839	-17.5	14.588	47.51	10.233
$c_1$	20.416	5.665	-2.062	1.654	9.039	4.488	0.928	3.473	5.789	3.441	-9.96	2.423
$c_2$	-1.196	0.329	0.125	0.097	-0.513	0.262	-0.043	0.203	-0.319	0.203	0.633	0.143
$res_{std}$	0.093		0.049		0.062		0.208		0.052		0.106	

Downloaded from https://academic.oup.com/mnras/article/512/3/3436/6530657 by guest on 02 July 2022



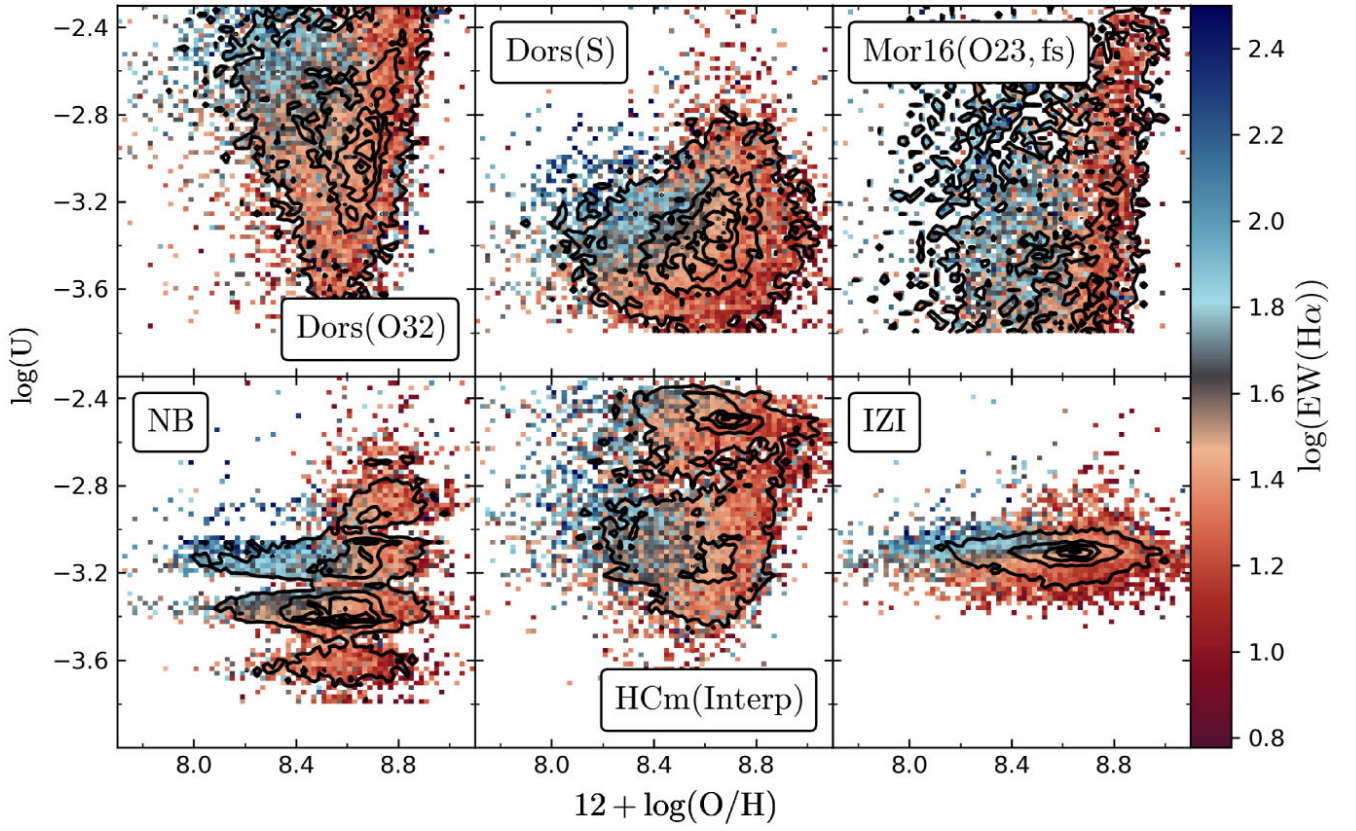
**Table E4.** Coefficients of the best third-order polynomial relation between each different calibrator of the oxygen abundance ( $y$ ) and the fiducial one (Ho 2019) ( $x$ ): i.e.  $y = c_0 + c_1x + c_2x^2 + c_3x^3$ . For each calibrator, we include the values of the coefficients and an estimation of their errors together with the standard deviation of the residuals ( $\text{res}_{\text{std}}$ ).

ID	Kew_02_N202		T04		Pet04_N2_lin		Pet04_N2_poly		Pet04_O3N2		KK04	
$c_0$	898.109	531.542	2007.277	623.214	172.453	148.794	666.123	164.08	1653.529	301.103	90.595	258.552
$c_1$	-315.241	186.699	-705.19	218.547	-62.649	52.166	-236.103	57.772	-588.617	106.048	-32.027	90.371
$c_2$	37.15	21.855	82.843	25.542	7.879	6.095	28.155	6.78	70.101	12.447	4.108	10.527
$c_3$	-1.455	0.853	-3.24	0.995	-0.327	0.237	-1.115	0.265	-2.779	0.487	-0.173	0.409
$\text{res}_{\text{std}}$	0.07		0.102		0.026		0.042		0.073		0.051	
	Pil10_ONS		Pil10_ON		Pil11_NS		M13_N2		M13_O3N2		Pil16_R	
$c_0$	-501.536	740.109	1747.281		940.027		141.663	108.763	1220.416	268.648	569.051	171.187
			-2163.682		-1068.199							
$c_1$	170.276	261.0	758.028	613.19	366.09	330.243	-50.795	38.129	-433.051	94.818	-200.948	59.948
$c_2$	-18.95	30.671	-88.189	71.714	-41.476	38.663	6.375	4.455	51.505	11.153	23.93	6.996
$c_3$	0.703	1.201	3.421	2.795	1.566	1.508	-0.264	0.173	-2.039	0.437	-0.947	0.272
$\text{res}_{\text{std}}$	0.071		0.089		0.085		0.021		0.05		0.035	
	Pil16_S		Cur20_R2		Cur20_R3		Cur20_R23		Cur20_N2		Cur20_O3N2	
$c_0$	1199.71	498.673	1577.692	635.709	1240.384	239.532	1272.809	770.875	136.559	128.678	1009.504	164.634
$c_1$	-421.71	174.78	-551.19	222.466	-440.462	84.675	-447.563	270.272	-47.764	45.502	-359.155	58.08
$c_2$	49.62	20.416	64.487	25.945	52.437	9.976	52.746	31.578	5.866	5.363	42.886	6.828
$c_3$	-1.941	0.795	-2.513	1.008	-2.078	0.392	-2.069	1.23	-0.237	0.211	-1.704	0.268
$\text{res}_{\text{std}}$	0.05		0.085		0.059		0.079		0.031		0.047	
	Cur20_O3O2		Cur20_S2		Cur20_RS32		Cur20_O3S2		IZI		NB	
$c_0$	1478.095	164.607	315.066	356.235	900.092	91.215	141.086	919.541	589.36	865.24	681.624	828.078
$c_1$	-529.588	58.138	-107.002	125.36	-319.792	32.168	-47.615	324.535	-208.984	305.511	-236.714	292.23
$c_2$	63.561	6.844	12.447	14.703	38.181	3.781	5.661	38.175	25.009	35.953	27.648	34.372
$c_3$	-2.541	0.269	-0.482	0.575	-1.517	0.148	-0.223	1.497	-0.995	1.41	-1.072	1.347
$\text{res}_{\text{std}}$	0.08		0.053		0.053		0.208		0.052		0.106	

## APPENDIX F: OXYGEN ABUNDANCE VERSUS IONIZATION PARAMETER

In Section 3.4, we explore the relation between the oxygen abundances and the ionization parameter using our fiducial calibrators for both parameters. However, the observed trend in Fig. 5 depends on the calibration used to estimate O/H and  $\log(U)$ . In the case of O/H, based on the results discussed in Appendix F, the trend would change only quantitatively but not qualitatively. However, this is not the case when adopting a different calibrator of  $\log(U)$ . In Fig. F1, we show the distribution of  $\log(U)$  as a function of our fiducial O/H for a set of different calibrators for the former parameter: (i) two of the calibrators proposed by (Dors et al. 2011, O32 and S2); (ii) the O32 fill sphere (fs) calibrator proposed by Morisset et al. (2016); (iii) the ionization parameters derived from applying the NebulaBayes, HCm,

and IZI codes to our data set (Pérez-Montero 2014; Thomas et al. 2018; Mingozi et al. 2020, respectively). As we can see in the figure, the trend between the two parameters heavily depends on the adopted calibrator. Moreover, the empirical calibrators present a negative trend (weak in all cases), but the theoretical ones (e.g. NebulaBayes and Hcm code) do not have any clear trend, showing a discrete distribution. For HCm code, the distribution covers the same range of parameters as the one shown by our fiducial calibrator. Finally, the IZI code presents a flat/constant distribution. It is clear that the theoretical calibrations depend on the adopted grid of photoionization models. It is beyond the scope of the current exploration to determine the nature of this discrepancy, but our comparison uncovers a strong difference that may affect the interpretation of any exploration using different calibrators.



**Figure F1.** Distribution of the ionization parameter,  $\log(U)$ , using each of the adopted calibrators (one for each panel) as a function of the fiducial oxygen abundance (Ho 2019) for our sample of H II regions. In each panel, the colour code represents the  $EW(H\alpha)$  with a scale ranging between  $\pm 3\sigma$  range around its mean value. The contours represent the density distribution, the outermost contour enclosing 95 per cent of the regions, and each consecutive one enclosing 20 per cent fewer regions.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.

## Capítulo 4

# Abundancias relativas de O/Fe

En el capítulo anterior se resumen los resultados presentados en el artículo [Espinosa-Ponce et al. \(2022\)](#), en el cual se estudió la relación entre las propiedades físicas de las regiones H II y las propiedades de las poblaciones estelares subyacentes. Usando el catálogo descrito en el Cap. II ([Espinosa-Ponce et al., 2020](#)) y los parámetros descritos en el Cap. III, en el presente capítulo abordamos la exploración de la abundancia relativa de oxígeno (en la fase gaseosa) frente a la metalicidad estelar  $[O/Z]$ . Este parámetro se calcula dividiendo la metalicidad de las estrellas pesada por luminosidad,  $Z_{*,L}$ , y la abundancia de oxígeno del gas, O/H, normalizado al valor solar,  $12 + \log(O/H)_{\odot} = 8.69$  ([Asplund et al., 2009](#)).  $[O/Z]$  es un trazador de  $[\alpha/Fe]$ , ya que el O es un elemento  $\alpha$  (ver Sec 1.1), mientras que la metalicidad estelar es la suma tanto de los elementos pico de hierro (como el Fe) y de elementos  $\alpha$  (como el oxígeno). Se encuentra que  $[O/Z]$  decrece con Z, resultado que está de acuerdo con la tendencia descrita para  $[\alpha/Fe]$  y  $[Fe/H]$  cuando se estudian estrellas individuales en nuestra galaxia ([Hayden et al., 2015](#)) y en datos integrados de galaxias de tipo temprano (por ejemplo [Walcher et al., 2015](#)).

El cociente  $[\alpha/Fe]$  es un estimador de la duración de los eventos de formación estelar en galaxias ([Tinsley, 1979](#); [Matteucci & Greggio, 1986](#)). Dado que los elementos  $\alpha$  son producidos en estrellas masivas y jóvenes los cuales son lanzados al ISM por CCSNe (aunque la mayoría de estos elementos se produce durante la explosión misma). Por otro lado, los elementos *iron-peak*, producidos en estrellas de todas las masas, llegan al ISM por medio de supernovas de tipo Ia (SNIa), las cuales son el producto del colapso de estrellas binarias. Estos dos procesos ocurren a diferentes escalas de tiempo, ya que mientras las CCSNe suceden solo durante los procesos de formación estelar, las SNIa tienen lugar a escalas de tiempo necesarias para el colapso de las binarias. En estudios previos se ha encontrado que las diferentes escalas de tiempo en la producción de estos dos grupos de elementos está relacionado así mismo con la historia de formación estelar ([Matteucci, 1992](#); [Matteucci et al., 1999](#)).

En la Vía Láctea, la exploración de  $[\alpha/Fe]$  se basa en la comparación de espectros estelares con modelos teóricos de atmósferas estelares (por ejemplo [Hayden et al., 2015](#)). El cálculo en otras galaxias se basa en comparar los espectros observados de poblaciones estelares con los resultados de códigos de síntesis estelar descritos en el Cap. 1, Sec. 1.4 (por ej. [Walcher et al., 2009](#)). Por esta razón, varios estudios sesgan sus muestras a galaxias del tipo temprano, en especial galaxias elípticas. Los espectros de estas galaxias presentan fuertes absorciones metálicas y casi no hay líneas de emisión provenientes del gas ionizado.

En este trabajo usamos la abundancia de oxígeno como indicador de la abundancia de los elementos  $\alpha$ . A lo largo de los años, se han usado las regiones H II para trazar la metalicidad

actual del ISM (por ejemplo Peimbert et al., 1978; Sánchez et al., 2014; Sánchez-Menguiano et al., 2016, 2018; Esteban & García-Rojas, 2018). Esta metalicidad es el resultado de un proceso de enriquecimiento químico en el punto donde la región H II se formó, tal y como vimos en el capítulo anterior. Por otra parte Z es la suma de todos los elementos, los cuales son el resultado de la suma de elementos  $\alpha$  y *iron-peak*. Los elementos más abundantes de estos dos grupos son el O y el Fe, por lo que en principio Z, o sea la metalicidad, se puede trazar por la abundancia de estos dos elementos. En el caso concreto de las SSPs adoptadas en este estudio, Z se traza a partir de la abundancia de Fe, por lo que ambos parámetros son proporcionales. En resumen, el cociente O/Z es un trazador, salvo un factor de escala, de  $[\alpha/\text{Fe}]$ .

Dada la variedad de galaxias en la muestra de CALIFA, en este estudio se pudo explorar la dependencia del cociente  $[\text{O}/\text{Fe}]$  con la masa estelar y la morfología de las galaxias. Se compararon los resultados con modelos de evolución química basados en (Weinberg et al., 2017). Se exploraron diferentes formas funcionales, todas dependientes del tiempo de retraso en la formación estelar ( $\tau_*$ ) y del tiempo de depleción ( $\tau_{\text{dep}}$ ). Estos modelos reproducen el comportamiento de  $[\alpha/\text{Fe}]$  vs  $[\text{Fe}/\text{H}]$  observado en la Vía Láctea (para mas detalles, ver Weinberg et al., 2017).

Se encontró que el cociente  $[\text{O}/\text{Fe}]$  presenta un comportamiento similar al observado para  $[\alpha/\text{Fe}]$  para galaxias de tipo temprano y la Vía Láctea. Existe una dependencia entre la distribución de  $[\text{O}/\text{Fe}]$  y la masa estelar y morfología de la galaxia. Finalmente, la distribución de datos concuerdan con los resultados de modelos de evolución química descritos anteriormente, de forma que  $\tau_*$  y  $\tau_{\text{dep}}$  dependen tanto de la masa como de la morfología. Esto último se ha comprobado con resultados recientes (Colombo et al., 2018).

Estos resultados se presentaron en una *letter* en la A&A que se incluye a continuación. La versión electrónica puede ser consultada en <https://ui.adsabs.harvard.edu/abs/2021A%26A...652L..10S>.

LETTER TO THE EDITOR

## **$[\alpha/\text{Fe}]$ traced by H II regions from the CALIFA survey**

### **The connection between morphology and chemical abundance patterns**

S. F. Sánchez<sup>1</sup>, C. Espinosa-Ponce<sup>1</sup>, L. Carigi<sup>1</sup>, C. Morisset<sup>2</sup>, J. K. Barrera-Ballesteros<sup>1</sup>, C. J. Walcher<sup>3</sup>,  
R. García-Benito<sup>4</sup>, A. Camps-Fariña<sup>1</sup>, and L. Galbany<sup>5</sup>

<sup>1</sup> Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 70-264, 04510 CDMX, Mexico  
e-mail: [sfsanchez@astro.unam.mx](mailto:sfsanchez@astro.unam.mx)

<sup>2</sup> Universidad Nacional Autónoma de México, Instituto de Astronomía, AP 106, Ensenada 22800, BC, Mexico

<sup>3</sup> Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany

<sup>4</sup> Instituto de Astrofísica de Andalucía (IAA/CSIC), Glorieta de la Astronomía s/n Aptdo. 3004, 18080 Granada, Spain

<sup>5</sup> Departamento de Física Teórica y del Cosmos, Universidad de Granada, 18071 Granada, Spain

Received 30 April 2021 / Accepted 4 August 2021

#### **ABSTRACT**

*Context.* Differential enrichment between  $\alpha$  and Fe-peak elements is known to be strongly connected with the shape of the star formation history (SFH), the star formation efficiency (SFE), the inflow and outflow of material, and even the shape of the initial mass function (IMF). However, beyond the Local Group, detailed explorations are mostly limited to early-type galaxies due to the lack of a good proxy for  $[\alpha/\text{Fe}]$  in late-type ones, limiting our understanding of the chemical enrichment process.

*Aims.* We intend to extend the explorations of  $[\alpha/\text{Fe}]$  to late-type galaxies in order to understand the details of the differential enrichment process.

*Methods.* We compare the gas-phase oxygen abundance with the luminosity-weighted stellar metallicity in an extensive catalogue of  $\sim 25\,000$  H II regions extracted from the Calar Alto Legacy Integral Field Area (CALIFA) survey, an exploration that uses the integral-field spectroscopy of  $\sim 900$  galaxies and covers a wide range of masses and morphologies. This way, we define  $[\text{O}/\text{Fe}]$  as the ratio between both parameters, proposing it as an indirect proxy of the  $[\alpha/\text{Fe}]$  ratio. This procedure is completely different from the one adopted to estimate  $[\alpha/\text{Fe}]$  from high-resolution spectroscopic data for stars in our Galaxy.

*Results.* We illustrate how the  $[\text{O}/\text{Fe}]$  parameter describes the chemical enrichment process in spiral galaxies, finding that: (i) it follows the decreasing pattern with  $[\text{Fe}/\text{H}]$  reported for the  $[\alpha/\text{Fe}]$  ratio and (ii) its absolute scale depends on the stellar mass and the morphology. We reproduce both patterns using two different chemical evolution models, considering that galaxies with different stellar masses and morphologies present either different SFHs, SFEs, and inflow and outflow rates or a different maximum stellar mass cut for the IMF. We will explore the differential chemical enrichment using this new proxy galaxy by galaxy and region by region in further studies.

**Key words.** galaxies: abundances – galaxies: evolution – galaxies: fundamental parameters – ISM: abundances – stars: abundances

## **1. Introduction**

All metals in the Universe are produced by the thermonuclear fusion reactions that are the core engines of stars. For intermediate- and low-mass stars, a fraction of these metals are expelled into the interstellar medium (ISM) during their lifetime as part of the stellar winds and, particularly in their later phases, when stars lose their envelopes. However, most of these metals remain inside the stars and end up in white dwarfs, neutron stars, and black holes (e.g. Kobayashi et al. 2020). The metals that enrich the ISM, traced by the most abundant elements, are produced during supernova explosions. In particular,  $\alpha$  elements (O, Mg, Si, S, Ar, Ca, and Ti; e.g. Matteucci 1992), produced mostly in massive, short-lived stars, are transferred to the ISM when these stars explode as Type-II, Type-Ib, or Type-Ic core-collapse supernovae (e.g. Woosley & Weaver 1995). Therefore, the production and enrichment of these heavy metals is directly associated with the star formation process. On the other hand, iron-peak elements are produced in stars of a wide range of masses, and their ISM enrichment is dominated by Type-Ia supernovae (SNeIa), which are triggered in binary systems

(Kobayashi et al. 2020). These events are not, in principle, connected with the most recent star formation processes (e.g. Galbany et al. 2014), as they happen during the lifetime of a stellar population after a certain delay time (e.g. Walcher et al. 2016; Castrillo et al. 2021) and span a wide time interval.

The most detailed explorations of the differential abundance between these two families of elements and its connection with the evolution of the stellar populations come from the analysis of high- and intermediate-resolution optical and near-infrared (NIR) spectra of hundreds to hundreds of thousands of stars at different locations within our Galaxy (e.g. Hayden et al. 2015; Yu et al. 2021; Nandakumar et al. 2020; Franchini et al. 2021). These analyses rely on the comparison of the full observed spectra, or a particular set of absorption features, with predictions from theoretical stellar atmosphere models (e.g. Hayden et al. 2015) or previously well-labelled empirical observations (e.g. Nandakumar et al. 2020). These previous studies found that (i) disk stars, in particular those in the inner disk, follow a well-defined track between  $[\alpha/\text{Fe}]$  and  $[\text{Fe}/\text{H}]$ , with a slope near  $\sim -0.3$ ; (ii) stars above the thin disk present a clear super-solar  $[\alpha/\text{Fe}]$  enhancement; and (iii) this enhancement is directly



connected with the age of the stellar population (e.g. Yu et al. 2021, Fig. 8). These results agree with the idea that the different timescales for the production of  $\alpha$  and iron-peak elements can be used as a clock for the star-formation history (SFH; Matteucci 1992; Matteucci et al. 1999; Thomas et al. 2005).

Beyond our galaxy, the derivation of the  $[\alpha/\text{Fe}]$  ratio relies on the comparison of the strength of certain absorption features or the full observed spectra for unresolved stellar populations with the predictions by stellar synthesis codes (e.g. Trager et al. 2000; Walcher et al. 2009; Vazdekis et al. 2015). For this reason, most of these explorations are limited to the study of elliptical galaxies, whose spectra present strong metal absorption features and a more limited contamination by ionized gas emission lines (e.g. Conroy & van Dokkum 2012; Walcher et al. 2015). Only very recently has this kind of exploration been extended to late-type galaxies as well (Watson et al. 2021). These studies have found that (i) early-type galaxies exhibit the same negative trend between  $[\alpha/\text{Fe}]$  and  $[\text{Fe}/\text{H}]$  reported for the resolved stellar populations in our galaxy (Walcher et al. 2016), and (ii) the  $[\alpha/\text{Fe}]$  ratio decreases with morphology (Watson et al. 2021). We should note that in none of those cases were the galaxies explored as resolved entities.

In this study we propose a different approach, using the presence of strong ionized gas emission lines to our advantage. Instead of tracing  $[\alpha/\text{Fe}]$  using the information provided by the absorption features in the stellar population, we gauge the  $\alpha$ -element abundance traced by the oxygen abundance ( $\text{O}/\text{H}$ ) in the ISM. Ever since the pioneering studies by Searle (1971), Comte (1975), and Peimbert et al. (1978), H II regions have been used to trace the chemical content in spiral galaxies (e.g. Sánchez et al. 2014), including the Milky Way (MW; e.g. Esteban & García-Rojas 2018). These gaseous clouds trace the current metal content of the ISM, which is assumed to be the same as that of the short-lived OB stars that ionized them (i.e. the most recent generation of stars). However, this metal content is the consequence of the full chemical enrichment history (ChEH) at the location in which the star formation is triggered and, therefore, the H II is created (Sánchez 2020). Due to that, it is possible to combine the oxygen abundance estimates derived from the analysis of the emission lines observed in the optical spectra of H II regions with the iron abundance estimated from the analysis of the underlying stellar spectrum to estimate the  $[\alpha/\text{Fe}]$  ratio, despite the fact that both measurements trace the metal content of two different populations. We demonstrate that this is a useful approach for exploring this elusive property in spiral galaxies, which we will exploit in detail in a forthcoming article (Espinosa-Ponce et al., in prep.). Previous attempts in the same direction (e.g. Lian et al. 2018; Sánchez 2020) compared the two parameters separately, reaching similar conclusions. This Letter is organized as follows. In Sect. 2 we describe our data, presenting the properties of the sample of galaxies and H II regions as well as a summary of the analysis performed on the data to derive the explored physical parameters; in Sect. 3 we present the main results of our analysis, showing the observational trends and comparing them with the expectations obtained by means of chemical evolution models (ChEMs). Finally, we summarize our main conclusions in Sect. 4.

## 2. Data and analysis

We used the catalogue of properties of H II regions and aggregations published by Espinosa-Ponce et al. (2020)<sup>1</sup>. These H II

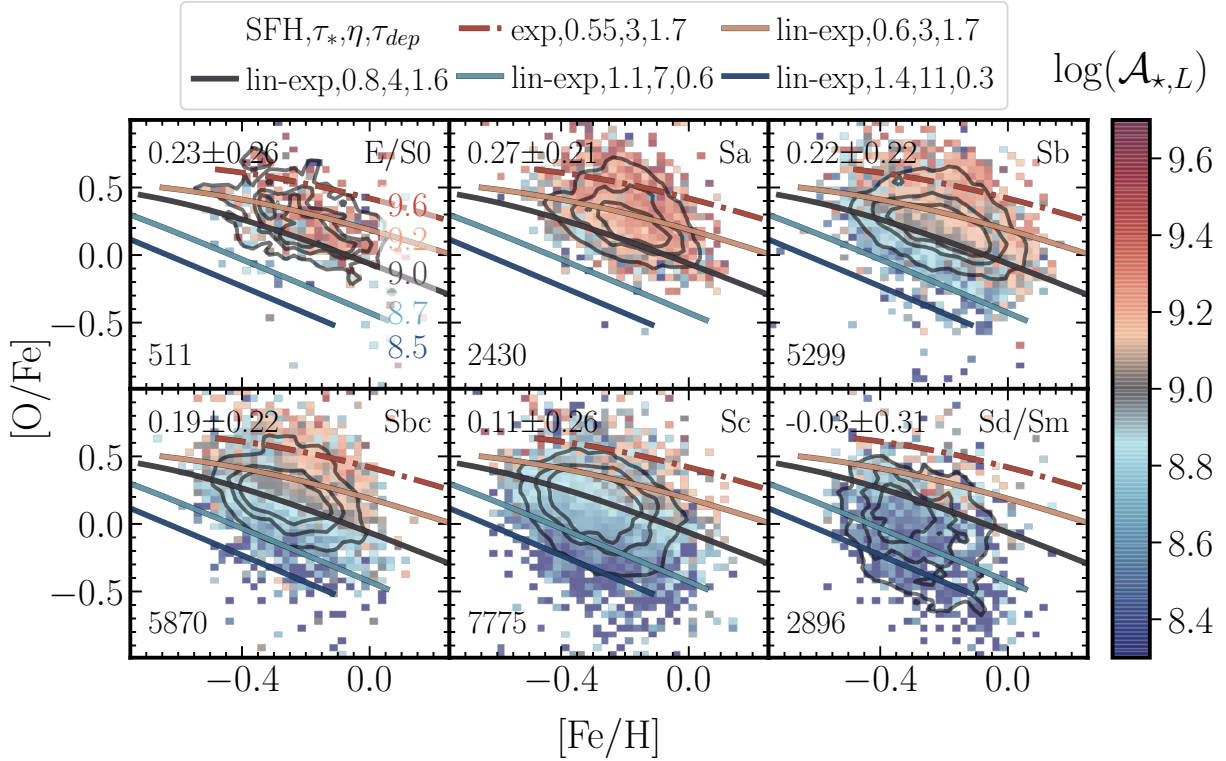
regions, and their spectroscopic properties, were extracted from the data provided by the Calar Alto Legacy Integral Field Area (CALIFA) survey (Sánchez et al. 2012a). This survey explored the full optical extension of a sample of galaxies representative of the bulk population in the nearby Universe ( $<100$  Mpc) (Walcher et al. 2014), using the PPAK Integral Field Unit (Kelz et al. 2006). Further details on the sample, survey strategy, observations, and reduction are provided in Walcher et al. (2014), Sánchez et al. (2016a), and Galbany et al. (2018).

The emission lines and stellar population properties are derived using the broadly tested Pipe3D pipeline (e.g. Sánchez et al. 2016b, 2021; Ibarra-Medel et al. 2019). This pipeline performs a spectral fitting decomposition of the stellar continuum based on a combination of Synthetic Stellar Populations (SSPs) that are convolved and shifted using a Gaussian kernel to account for the line-of-sight velocity distribution and attenuated by an extinction law. Once the best stellar model is obtained for each spectrum in each cube, this model is subtracted from the original spectra, creating a cube with only the ionized emission line component. This cube is then used to derive the main properties of a set of predefined emission lines.

Espinosa-Ponce et al. (2020) extracted a catalogue of H II regions and aggregations from the data products provided by PIPE3D and applied to the CALIFA data cubes. In order to do so, they followed the prescriptions by Sánchez et al. (2012b), selecting clumpy structures in the  $\text{H}\alpha$  intensity map of each galaxy and further selecting from them those compatible with being ionized by a young stellar population, that is to say, with a fraction of young stars and an equivalent width of  $\text{H}\alpha$  high enough to produce the observed ionization. This analysis provided a catalogue that comprises the main properties of the ionized gas emission lines (including the flux intensities of 52 lines) and the main physical parameters of the underlying stellar population, for a total of  $\sim 25\,000$  H II regions across the 924 explored galaxies.

We made use of this catalogue, extracting three main physical properties. The first is the luminosity-weighted age ( $\mathcal{A}_{\star,L}$ ) of the underlying stellar population, which corresponds to the first moment of the age distribution function (ADF) in light, directly provided by the stellar decomposition performed by PIPE3D. This parameter is directly connected with how sharp the SFH is in galaxies, with older stellar populations having SFHs that peak at earlier epochs (e.g. García-Benito et al. 2017). The second is the luminosity-weighted metallicity normalized to the solar value ( $[Z/\text{H}]$  or  $Z_{\star,L}$ ). This is the first moment of the metallicity distribution function (MDF) in light at the same band as the ADF (e.g. Mejía-Narváez et al. 2020). Due to the nature of the adopted SSPs, this parameter is a direct proxy of the luminosity-weighted iron abundance,  $[\text{Fe}/\text{H}]$ , for a solar abundance of  $12 + \log(\text{Fe}/\text{H}) = 7.50$  (Asplund et al. 2009). It is worth noting that we derive  $[Z/\text{H}]$  using SSP models with solar composition (i.e.  $[\alpha/\text{Fe}] = 0$ ). This may create a bias in  $[\text{Fe}/\text{H}]$  towards higher values, which can be as large as  $\sim 30\%$ . However, this bias affects neither our qualitative results nor the reported trends, although it may produce relative offsets in the absolute reported abundances. Hereafter we refer to this parameter as  $[\text{Fe}/\text{H}]$ . Finally, the third is the gas-phase oxygen abundance,  $12 + \log(\text{O}/\text{H})$ , derived using the calibrator presented by Ho (2019). This calibrator uses a set of emission line ratios and associates them with oxygen abundances measured using the direct method, applying a neural-network procedure. We should note that the use of this or another of the calibrators applied to derive the oxygen abundance for the considered catalogue does not qualitatively alter the results presented here (Espinosa-Ponce et al., in prep.). We transformed this value to  $[\text{O}/\text{H}]$  by adopting a

<sup>1</sup> [http://ifs.astroscu.unam.mx/CALIFA/HII\\_regions/](http://ifs.astroscu.unam.mx/CALIFA/HII_regions/)



**Fig. 1.** Distribution of  $[\text{O}/\text{Fe}]$  as a function of  $[\text{Fe}/\text{H}]$  for the H II regions of our sample (segregated by galaxy morphology), colour-coded by the average  $\mathcal{A}_{*,L}$  of the underlying stellar population. Contours represent the density of the plotted regions, encircling 95%, 80%, and 40% of them, respectively, with the total number shown at the bottom left of each panel. The average  $[\text{O}/\text{Fe}]$  value for each morphology is shown at the top left of each panel as well. The solid and dashed-dotted lines are all the same in the six panels. They correspond to the predictions by the five ChEMs described in the text, generated to cover the observed distributions. Each model is defined by the values of a set of input parameters ( $\tau_*$ ,  $\eta$ ,  $\tau_{\text{dep}}$ ) and two shapes of the SFHs (exponential decay, *exp*, and linear rising plus exponential decay, *lin-exp*), indicated in the upper legend. Lines are colour-coded according to the average  $\mathcal{A}_{*,L}$  as well. We note that there is no exact one-to-one correspondence between a single model and the observed distribution for a single morphological bin (i.e. panel).

solar oxygen abundance of  $12 + \log(\text{O}/\text{H})_{\odot} = 8.69$  (Asplund et al. 2009).

### 3. Results

Figure 1 shows the distribution of the relative oxygen-to-iron abundance,  $[\text{O}/\text{Fe}]$ , as a function of  $[\text{Fe}/\text{H}]$  for the catalogue of H II regions explored in this article. The  $[\text{O}/\text{Fe}]$  is constructed from the  $[\text{O}/\text{H}]$  and  $[\text{Fe}/\text{H}]$  parameters described in the previous section. We should recall that while  $[\text{Fe}/\text{H}]$  is an average of the iron abundances of the surviving stars,  $[\text{O}/\text{H}]$  traces the oxygen content of the most recently formed ones (e.g. González Delgado et al. 2014). However,  $[\text{O}/\text{H}]$  is also a consequence of the evolution of the stellar population, which is clearly connected with its age (e.g. Sánchez et al. 2015; Espinosa-Ponce et al. 2020) and, more strongly, with the stellar mass at any look-back time (e.g. Fig. 1 of Maiolino & Mannucci 2019). Furthermore, a fraction of the oxygen is depleted into dust in H II regions, and therefore the  $[\text{O}/\text{H}]$  (and  $[\text{O}/\text{Fe}]$ ) reported here may be  $\sim 0.08$ – $0.12$  dex lower than if they were both measured in stars (Peimbert & Peimbert 2010). Thus, the  $[\text{O}/\text{Fe}]$  shown here should not be interpreted as the  $\alpha$  enhancement of either the average population (luminosity or mass weighted) at the location of the H II region or of a particular subset of the population. However, as we demonstrate below, it still retains valuable information about the  $\alpha$ -enhancement process along the cosmological evolution of the explored populations.

Each panel of Fig. 1 shows the distribution for the H II regions located in galaxies of different morphological types. The observed distributions in each panel exhibit similar trends. In all cases,  $[\text{O}/\text{Fe}]$  decreases as  $[\text{Fe}/\text{H}]$  increases, following a linear trend with a slope of  $\sim -0.7$ . This trend is less defined, showing a broader dispersion, for the later morphological types (Sc and Sd). The described pattern is observed in all panels, despite the different number of H II regions in each of them (a consequence of the morphological distribution of the CALIFA galaxies, which peaks at Sb/Sbc, and the lower number of H II regions in earlier morphological types; e.g. Lacerda et al. 2020; Espinosa-Ponce et al. 2020). Furthermore, the average  $[\text{O}/\text{Fe}]$  (top-left label of each panel in Fig. 1) declines for galaxies with later morphologies. However, the  $[\text{O}/\text{Fe}]$  peak is not reached by the earlier morphological types (E/S0,  $[\text{O}/\text{Fe}] \sim 0.23$  dex), but it is for the Sa ones ( $[\text{O}/\text{Fe}] \sim 0.27$  dex). We should note that the difference between these two average values is small compared to their corresponding standard deviations ( $\sim 0.26$ – $0.21$  dex). For later morphological types, this value progressively declines with morphology.

The decline in  $[\text{O}/\text{Fe}]$  as a function of  $[\text{Fe}/\text{H}]$  is qualitatively similar to that of the stars in the MW (e.g. Yates et al. 2013; Griffith et al. 2021), not only for the oxygen abundance (e.g. Franchini et al. 2021) but also for any other  $\alpha$  elements. However, the quantities reported in those studies and the ones presented here have a different nature: the abundance of individual

stars (for the MW) versus the average properties of two different populations (in this study). To understand the nature of this decline and to demonstrate that we are tracing the same patterns, we colour-coded the distributions in Fig. 1 by the  $\mathcal{A}_{\star,L}$  of the underlying stellar population. New patterns emerge from this exploration: (i) the decline in [O/Fe] as a function of [Fe/H] is much better defined for a fixed  $\mathcal{A}_{\star,L}$  and morphology (the nature of this relation is discussed below); and (ii) the broadening and shift of the observed distribution with morphology is a consequence of the strong dependence of [O/Fe] with  $\mathcal{A}_{\star,L}$ . Early-type (late-type) galaxies have older (younger) stellar populations in general, with very little (strong) variation in the observed SFHs and therefore a narrow (broad) range of  $\mathcal{A}_{\star,L}$  values (Ibarra-Medel et al. 2016; García-Benito et al. 2017). This scenario fits with the observations for all morphological bins for early and late spirals (Sa to Sd/Sm). However, we find that E/S0 may present a slightly lower [O/Fe] than Sa galaxies (although the difference may not be statistically significant). If true, this could be due to a bias as our exploration is restricted to those galaxies with an observed H II region. For E/S0 galaxies, there may be scarce (Espinosa-Ponce et al. 2020) and peculiar objects: galaxies with either an SFH more extended than the bulk population of early-type ones or suffering a rejuvenation by the capture of pristine gas (e.g. Gomes et al. 2016). Nevertheless, despite this possible bias, the result fits with the proposed connection between [O/Fe] and  $\mathcal{A}_{\star,L}$  since, at least for the CALIFA sample, the peak in ages for the stellar population is found for the Sa and not for the E/S0 galaxies (e.g. Fig. 9 of García-Benito et al. 2017).

Similar trends between  $[\alpha/\text{Fe}]$ , [Fe/H], and the age of the stellar populations have been reported when exploring the  $\alpha$  enhancement for resolved stars in the MW (e.g. Hayden et al. 2015; Yu et al. 2021; Nandakumar et al. 2020; Franchini et al. 2021), for different compilations of early-type galaxies (Walcher et al. 2015, 2016), and when comparing MW-like galaxies with simulations (e.g. Calura & Menci 2009; Yates et al. 2012). In some cases the quoted studies reported a weak decline in the metallicity with the stellar age (e.g. Fig. 2 in Walcher et al. 2016) that is not seen in our results due to a different definition of the parameters (Coelho et al. 2007)<sup>2</sup>. Finally, the stars in the disk of the MW present a smoother relation of [O/Fe] with [Fe/H] than that reported here (Nandakumar et al. 2020).

We explore now if the observed distributions are physically compatible with our current understanding of the chemical evolution in galaxies. To do so we made use of the ChEM from Weinberg et al. (2017)<sup>3</sup>, an analytic solution for the one-zone (fully mixed) model that incorporates a realistic delay time distribution for SNeIa. In this way, it is possible to track the separate evolution of  $\alpha$  and the iron-peak elements. The model requires us to define the shape of the SFH (instead of the infall rate as other ChEMs do; e.g. Carigi et al. 2019), by selecting either an exponential (exp,  $e^{-t/\tau_{\star}}$ ) or a linear-exponential (lin-exp,  $te^{-t/\tau_{\star}}$ ) functional form, both parametrized by the star formation time delay ( $\tau_{\star}$ ). In addition, it requires us to define the depletion time ( $\tau_{\text{dep}}$ ) and the mass-loading factor ( $\eta$ ), parameters that define the rate at which gas is transformed into stars and the amount of leaked metals from the system as a function of the star formation rate (e.g. Barrera-Ballesteros et al. 2018), respectively. This model has been successfully used to reproduce the distribution of  $[\alpha/\text{Fe}]$  versus [Fe/H] observed in the MW (Weinberg et al. 2017).

Using this code, we generated a set of five ChEMs without attempting to fit them with the observations or to obtain a perfect

one-to-one matching with the individual distributions reported for each morphological bin. We intended to reproduce the bulk observed distributions shown in Fig. 1 using a reasonable set of input parameters: (i) one model uses the exp functional form to reproduce SFHs of earlier-type galaxies (or the centre of early spirals) and the remaining models a lin-exp one that matches better with the SFHs of late-type galaxies (following López Fernández et al. 2018; García-Benito et al. 2019); (ii) a different value of  $\tau_{\star}$  was adopted for each of the five models, spanning from 0.55 to 1.4 Gyr, to reproduce the range of observed  $\mathcal{A}_{\star,L}$ ; (iii) four different values for  $\eta$  were adopted, ranging between 3 and 11 and covering the expected values for galaxies of different morphologies and stellar masses, based on the relation between  $\eta$  and  $M_{\star}$  from Leethochawalit et al. (2019) and the morphology-mass relation in the CALIFA sample (e.g. Lacerda et al. 2020); finally, (iv) we adopted four different values of  $t_{\text{dep}}$  to reproduce the dependence of this parameter on the morphological type described by Colombo et al. (2018), using the values tabulated by Sánchez et al. (2021). The top inset of Fig. 1 indicates the parameters adopted for the five ChEMs. For each model we derived the oxygen abundance of the last generation of stars, at each look-back time, which would correspond to our observed gas phase [O/H], and the luminosity-weighted [Fe/H] for the surviving stars formed before that time (i.e. older ages). To derive this latter parameter we were required to average the [Fe/H] weighted by the SFH, transforming the mass to light using the same M/L ratio at each age adopted by PIPE3D and considering the mass loss at each look-back time. The resulting tracks for each model, covering the last 2 Gyr, are represented in Fig. 1 as dot-dashed or solid lines (for each of the selected SFHs) and colour-coded by the current  $\mathcal{A}_{\star,L}$ .

The simple ChEMs described above reproduce the main observed trends between [O/Fe] and [Fe/H]. In particular, (i) the average decline in [O/Fe] with increasing [Fe/H], (ii) the range of [O/Fe] values, and (iii) its decline with the  $\mathcal{A}_{\star,L}$ , which induces the trends with morphology, are well reproduced. It is worth noting that we did not fit or adjust the input parameters of the model to reproduce the observed abundance patterns:  $\eta$  and  $\tau_{\text{dep}}$  were extracted from the literature (Leethochawalit et al. 2019; Colombo et al. 2018; Sánchez et al. 2021), and the shape of the SFHs and  $\tau_{\star}$  were selected to reproduce just the average  $\mathcal{A}_{\star,L}$ . Based on this modelling, massive and earlier-type galaxies present stronger [O/Fe] enhancement, covering a narrower range of values and with a better defined trend of this parameter with [Fe/H] as a consequence of shorter and sharper SFHs, a low diversity in the SFH, lower star formation efficiency (SFE) values, and a better ability to retain metals. On the contrary, less massive and later-type galaxies present weaker [O/Fe] enhancements for the opposite reasons: longer and steady SFHs, a larger variety of SFHs, higher values of SFE, and a lower ability to retain metals. All these trends agree with our current understanding of galaxy evolution as well as the most recent results reported in the literature regarding the SFH, ChEM, MDFs, and their relation with galaxy morphology, mass, and location within galaxies (e.g. García-Benito et al. 2017; González Delgado et al. 2017; Ibarra-Medel et al. 2016; Mejía-Narváez et al. 2020; Camps-Fariña et al. 2021; Sánchez 2020).

We should note that this is not the only interpretation possible for the observed distributions. For instance, using a more elaborated ChEM that assumes no outflows (Carigi et al. 2019, 2020), it is possible to reproduce all the observed trends by adopting a variable initial mass function (IMF). The model assumes a certain gas infall rate as the basic regulator for the star formation. Changing this rate and adopting a Salpeter (1955)

<sup>2</sup> In particular,  $[Z/H] = [\text{Fe}/H] + 0.75[\alpha/\text{Fe}]$  instead of  $[Z/H] = [\text{Fe}/H]$ .

<sup>3</sup> <https://github.com/jobovy/kimmy>



IMF with different mass ends ( $M_{\text{cut}} = 20\text{--}120 M_{\odot}$ ), it is possible to reproduce the observed trends. The possibility of a variable IMF is a topic of current discussion (Kroupa 2001), without a general consensus of what the main driver is (e.g. Martín-Navarro et al. 2015). Based on this model, the observed trends would be the consequence of the ability of more massive galaxies to form more massive stars (higher  $M_{\text{cut}}$ ) than less massive ones (lower  $M_{\text{cut}}$ ; Gunawardhana et al. 2011; Fernández-Alvar et al. 2018).

The comparison with these two particular ChEMs indicates that the observed trends (which are our primary results) are compatible with plausible scenarios and interpretations. Those scenarios are not unique, incompatible, or exclusive. Some of the observed trends were already predicted by the ChEMs presented in Matteucci (2003) as a consequence of the change in the SFH between early- and late-type galaxies. However, a change in the IMF does not exclude a possible change in the shape of the SFH or a different SFE. Finally, many of the parameters involved in those models are strongly degenerated, and modifying any of them may produce concordant results. However, these would not change the described observed trends, though the detailed interpretation may be different.

#### 4. Conclusions

We present for the first time an exploration of the  $\alpha$  enhancement in galaxies based on the comparison of the gas-phase oxygen abundance ( $[\text{O}/\text{H}]$ ) with the stellar metallicity ( $[\text{Z}/\text{H}]$ ) for a large sample of H II regions and aggregations extracted for a representative sample of galaxies in the nearby Universe. From this exploration we show that:

1. The  $[\text{O}/\text{Fe}]$  presents a decline with  $[\text{Fe}/\text{H}]$  similar to those observed for  $[\alpha/\text{Fe}]$  in the MW, in early-type galaxies, and described in simulations.
2. The zero point (slope) of this relation (i.e. the absolute scale of  $[\text{O}/\text{Fe}]$ ) presents a strong (mild) dependence on both the stellar mass and morphology of the galaxy, in agreement with early scenarios, ChEMs, simulations, and recent results.
3. We reproduce both trends using ChEMs by assuming that either (i) the SFH, SFE, and  $\eta$  or (ii) the high-mass cutoff IMF increases with the stellar mass of galaxies, in agreement with previous results.

We have presented in this study the most relevant results of our ongoing exploration. In a forthcoming article, Espinosa-Ponce et al. (in prep.), we will provide further details on the modelling, exploring the reported trends galaxy by galaxy and in different regions within them.

*Acknowledgements.* We thank the referee for his/her comments and suggestions. SFS and J.B.-B. thanks CONACYT for grants CB-285080 and FC-2016-01-1916, and PAPIIT-DGAPA-IN100519 (UNAM) project. C.M. thanks UNAM/PAPIIT-IN101220. J.B.-B. thanks IA-100420 (DGAPA-PAPIIT, UNAM) and CONACYT grant CF19-39578 support. L.G. thanks M.S.-Curie grant 839090. R.G.B. acknowledges support from grants SEV-2017-0709 and P18-FRJ-2595. This study uses data provided by the Calar Alto Legacy Integral Field Area (CALIFA) survey (<http://califa.caha.es/>), observed at the Calar Alto Observatory.

#### References

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481  
 Barrera-Ballesteros, J. K., Heckman, T., Sánchez, S. F., et al. 2018, *ApJ*, 852, 74  
 Calura, F., & Menci, N. 2009, *MNRAS*, 400, 1347  
 Camps-Fariña, A., Sanchez, S. F., Lacerda, E. A. D., et al. 2021, *MNRAS*, 504, 3478  
 Carigi, L., Peimbert, M., & Peimbert, A. 2019, *ApJ*, 873, 107  
 Carigi, L., Peimbert, A., Peimbert, M., & Delgado-Inglada, G. 2020, *Rev. Mex. Astron. Astrofis.*, 56, 235

Castrillo, A., Ascasibar, Y., Galbany, L., et al. 2021, *MNRAS*, 501, 3122  
 Coelho, P., Bruzual, G., Charlot, S., et al. 2007, *MNRAS*, 382, 498  
 Colombo, D., Kalinova, V., Utomo, D., et al. 2018, *MNRAS*, 475, 1791  
 Comte, G. 1975, *A&A*, 39, 197  
 Conroy, C., & van Dokkum, P. G. 2012, *ApJ*, 760, 71  
 Espinosa-Ponce, C., Sánchez, S. F., Morisset, C., et al. 2020, *MNRAS*, 494, 1622  
 Esteban, C., & García-Rojas, J. 2018, *MNRAS*, 478, 2315  
 Fernández-Alvar, E., Carigi, L., Schuster, W. J., et al. 2018, *ApJ*, 852, 50  
 Franchini, M., Morossi, C., Di Marcantonio, P., et al. 2021, *AJ*, 161, 9  
 Galbany, L., Stanishev, V., Mourão, A. M., et al. 2014, *A&A*, 572, A38  
 Galbany, L., Anderson, J. P., Sánchez, S. F., et al. 2018, *ApJ*, 855, 107  
 García-Benito, R., González Delgado, R. M., Pérez, E., et al. 2017, *A&A*, 608, A27  
 García-Benito, R., González Delgado, R. M., Pérez, E., et al. 2019, *A&A*, 621, A120  
 Gomes, J. M., Papaderos, P., Vílchez, J. M., et al. 2016, *A&A*, 585, A92  
 González Delgado, R. M., Cid Fernandes, R., García-Benito, R., et al. 2014, *ApJ*, 791, L16  
 González Delgado, R. M., Pérez, E., Cid Fernandes, R., et al. 2017, *A&A*, 607, A128  
 Griffith, E., Weinberg, D. H., Johnson, J. A., et al. 2021, *ApJ*, 909, 77  
 Gunawardhana, M. L. P., Hopkins, A. M., Sharp, R. G., et al. 2011, *MNRAS*, 415, 1647  
 Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, *ApJ*, 808, 132  
 Ho, I. T. 2019, *MNRAS*, 485, 3569  
 Ibarra-Medel, H. J., Sánchez, S. F., Avila-Reese, V., et al. 2016, *MNRAS*, 463, 2799  
 Ibarra-Medel, H. J., Avila-Reese, V., Sánchez, S. F., González-Samaniego, A., & Rodríguez-Puebla, A. 2019, *MNRAS*, 483, 4525  
 Kelz, A., Verheijen, M. A. W., Roth, M. M., et al. 2006, *PASP*, 118, 129  
 Kobayashi, C., Karakas, A. I., & Lugaro, M. 2020, *ApJ*, 900, 179  
 Kroupa, P. 2001, *MNRAS*, 322, 231  
 Lacerda, E. A. D., Sánchez, S. F., Cid Fernandes, R., et al. 2020, *MNRAS*, 492, 3073  
 Leethochawalit, N., Kirby, E. N., Ellis, R. S., Moran, S. M., & Treu, T. 2019, *ApJ*, 885, 100  
 Lian, J., Thomas, D., Maraston, C., et al. 2018, *MNRAS*, 476, 3883  
 López Fernández, R., González Delgado, R. M., Pérez, E., et al. 2018, *A&A*, 615, A27  
 Maiolino, R., & Mannucci, F. 2019, *A&ARv*, 27, 3  
 Martín-Navarro, I., Vazdekis, A., La Barbera, F., et al. 2015, *ApJ*, 806, L31  
 Matteucci, F. 1992, *Mem. Soc. Astron. It.*, 63, 301  
 Matteucci, F. 2003, *Ap&SS*, 284, 539  
 Matteucci, F., Romano, D., & Molaro, P. 1999, *A&A*, 341, 458  
 Mejía-Narváez, A., Sánchez, S. F., Lacerda, E. A. D., et al. 2020, *MNRAS*, 499, 4838  
 Nandakumar, G., Hayden, M. R., Sharma, S., et al. 2020, *MNRAS*, submitted [arXiv:2011.02783]  
 Peimbert, A., & Peimbert, M. 2010, *ApJ*, 724, 791  
 Peimbert, M., Torres-Peimbert, S., & Rayo, J. F. 1978, *ApJ*, 220, 516  
 Salpeter, E. E. 1955, *ApJ*, 121, 161  
 Sánchez, S. F. 2020, *ARA&A*, 58, 99  
 Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012a, *A&A*, 538, A8  
 Sánchez, S. F., Rosales-Ortega, F. F., Marino, R. A., et al. 2012b, *A&A*, 546, A2  
 Sánchez, S. F., Rosales-Ortega, F. F., Iglesias-Páramo, J., et al. 2014, *A&A*, 563, A49  
 Sánchez, S. F., Pérez, E., Rosales-Ortega, F. F., et al. 2015, *A&A*, 574, A47  
 Sánchez, S. F., García-Benito, R., Zibetti, S., et al. 2016a, *A&A*, 594, A36  
 Sánchez, S. F., Pérez, E., Sánchez-Blázquez, P., et al. 2016b, *Rev. Mex. Astron. Astrofis.*, 52, 171  
 Sánchez, S. F., Walcher, C. J., Lopez-Cobá, C., et al. 2021, *Rev. Mex. Astron. Astrofis.*, 57, 3  
 Searle, L. 1971, *ApJ*, 168, 327  
 Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, *ApJ*, 621, 673  
 Trager, S. C., Faber, S. M., Worthey, G., & González, J. J. 2000, *AJ*, 119, 1645  
 Vazdekis, A., Coelho, P., Cassisi, S., et al. 2015, *MNRAS*, 449, 1177  
 Walcher, C. J., Coelho, P., Gallazzi, A., & Charlot, S. 2009, *MNRAS*, 398, L44  
 Walcher, C. J., Wisotzki, L., Bekeraité, S., et al. 2014, *A&A*, 569, A1  
 Walcher, C. J., Coelho, P. R. T., Gallazzi, A., et al. 2015, *A&A*, 582, A46  
 Walcher, C. J., Yates, R. M., Minchev, I., et al. 2016, *A&A*, 594, A61  
 Watson, P. J., Davies, R. L., Brough, S., et al. 2021, *ArXiv e-prints* [arXiv:2106.01928]  
 Weinberg, D. H., Andrews, B. H., & Freudenburg, J. 2017, *ApJ*, 837, 183  
 Wosley, S. E., & Weaver, T. A. 1995, *ApJS*, 101, 181  
 Yates, R. M., Kauffmann, G., & Guo, Q. 2012, *MNRAS*, 422, 215  
 Yates, R. M., Henriques, B., Thomas, P. A., et al. 2013, *MNRAS*, 435, 3500  
 Yu, Z., Li, J., Chen, B., et al. 2021, *ApJ*, 912, 106





## Capítulo 5

# Exploración con Modelos de Fotoionización

### Resumen

El estudio las propiedades físicas de las regiones H II y su relación con las líneas de emisión se realiza con modelos de fotoionización. Para esto, se requiere la creación de un conjunto de modelos cuyos parámetros de entrada, las propiedades físicas del gas ionizado, cubran de la mejor manera posible el espacio de parámetros. Por tanto, se requiere un gran número de modelos para tomar en cuenta la mayor cantidad de combinaciones posibles entre los parámetros de entrada. Sin embargo, esto conlleva un tiempo de computo considerable para la generación de todo el conjunto de modelos. En este capítulo se describe el estudio de los parámetros físicos, como  $12 + \log(\text{O}/\text{H})$ ,  $\log(\text{N}/\text{O})$  y  $\log(U)$ , con modelos de fotoionización y técnicas de aprendizaje automatizado. Se explora la posibilidad de, dados unos parámetros de entrada, generar resultados congruentes entre los modelos de fotoionización y una red neuronal artificial. Utilizando una base de modelos existente, se entrenó una red neuronal artificial que produce resultados comparables a los producidos por un código de fotoionización en un menor tiempo. Se encontró que es posible utilizar una red neuronal artificial para producir modelos de fotoionización en un menor tiempo. Aunque este resultado está limitado a que la red neuronal sea entrenada adecuadamente. Esto forma parte de un estudio en proceso que actualmente se está llevando a cabo.

### 5.1. Estudio de las regiones HII con modelos de fotoionización

En las secciones anteriores se han estudiado las regiones H II y sus propiedades físicas de forma empírica. Sin embargo, como se mencionó en el Cap. 1 se pueden estudiar teóricamente estas nebulosas con ayuda de los modelos de fotoionización. Usando estos modelos podemos entender cómo las propiedades físicas del gas ionizado afectan su estructura de ionización y, por lo tanto, como se producen las líneas de emisión que se observan.

En el Cap 1 Sec. 1.5.2 se presentó un preámbulo de como los modelos de fotoionización se han usado ampliamente para el estudio de gas ionizado. Con ayuda de conjuntos de modelos se puede explorar el espacio de parámetros físicos y su conexión con las observables (por ejemplo razones de intensidades de las líneas de emisión). En cada red de modelos se toma una distribución discreta para cada uno de los parámetros de entrada de un modelo de fotoionización (p. ej. abundancia de elementos, densidad de H, etc. ver Sec. 1.5).

El número de modelos en las redes para el estudio de regiones H II están limitados por el tiempo de cómputo. Cada modelo tarda en completarse entre 2 a 5 minutos. Por lo que, una exploración exhaustiva de las propiedades físicas resulta poco práctica. Por lo tanto se debe escoger un número adecuado de modelos que muestree el espacio de parámetros de la forma más completa posible.

Básicamente, un modelo de fotoionización toma una serie de parámetros físicos de entrada (p. ej. abundancia de los elementos, densidad de H, etc) y a partir de ellos se generan un conjunto de datos observables de salida. El proceso que toma más tiempo de cómputo en un modelo de fotoionización es resolver todas las ecuaciones que describen la física interna del gas ionizado. Por tanto, para poder crear un conjunto de modelos lo suficientemente grande que abarque todo el espacio de parámetros se necesita, de alguna manera, reducir el tiempo requerido por el proceso que convierte los datos de entrada a los datos de salida. Para esta tarea podemos usar técnicas de inteligencia artificial.

### 5.1.1. Relación entre O/H, U y N/O

Como se ha mencionado anteriormente, una de las propiedades físicas más importantes de las regiones H II es la abundancia de oxígeno. Este elemento juega un papel fundamental en el entendimiento de la evolución química de las galaxias por diversas razones. Como se ha mencionado anteriormente, es un elemento muy abundante y sus líneas pueden ser observadas con relativa facilidad. Por lo que existen buenos datos espectrofotométricos donde están disponibles las intensidades de línea necesarias para calcular la abundancia de oxígeno. Además, se conoce de forma detallada el origen del oxígeno, siendo casi completamente producido en estrellas masivas y eyectado al ISM por medio de las explosiones de supernova tipo II (Chiappini et al., 2003). Por otra parte, puede darse una sobreproducción de O en estrellas AGB poco metálicas y de baja masa (entre  $\sim 1 - 3M_{\odot}$ , García-Hernández et al., 2016). Por otra parte, elementos como C y N pueden ser producidos por estrellas de todas las masas, lo cual hace que su origen sea más complejo. El nitrógeno puede depender del hidrógeno y helio original en una estrella, ya sea por una producción directa o en alguna etapa de nucleosíntesis (a este caso se le llama producción primaria). Sin embargo, el nitrógeno puede depender de un enriquecimiento previo de C y O (en este caso se le conoce como producción secundaria). Una herramienta fundamental que nos permitirá obtener información sobre la producción de N es el cociente N/O. Además, este cociente depende de la historia de formación estelar (Mollá et al., 2006) y de flujos de gas, es decir, caída de gas, de gas prístino y de los vientos (Köppen & Hensler, 2005; Belfiore et al., 2016).

Hay una relación aparentemente universal (pero con una dispersión considerable) entre N/O y O/H ampliamente estudiada (p. ej. Izotov & Thuan, 1999; Henry et al., 2000; Belfiore et al., 2017; Schaefer et al., 2020). El cociente permanece constante a bajas metalicidades o en las primeras etapas de la evolución química de una galaxia. Esto implicaría que el N es primario en este régimen. Conforme la abundancia de O crece, el cociente N/O crece rápidamente debido al propio enriquecimiento de elementos secundarios. Esto nos indica que la producción del nitrógeno está dominada por un proceso primario a bajas metalicidades mientras que a metalicidades altas es dominado por procesos secundarios. Algunas de las relaciones estudiadas se muestran en el panel derecho de la Fig. 5.1. Muchas de las relaciones estudiadas indican que a valores bajos de  $12 + \log(\text{O}/\text{H})$  ( $\lesssim 8$ ), se tiene una valor constante del cociente N/O, mientras para valores altos de  $12 + \log(\text{O}/\text{H})$ , el cociente N/O crece rápidamente.

Otra de las propiedades que modifican las condiciones físicas dentro del gas ionizado es el parámetro de ionización,  $\log(U)$  (descrito en el Cap. 1). Diversas exploraciones han encontrado

que  $\log(U)$  decrece con la abundancia de oxígeno, (p. ej. [Dopita & Evans, 1986](#)). Este tipo de dependencia se ha usado en la literatura con modelos de fotoionización (p. ej. [Thomas et al., 2018](#)). Empíricamente se ha encontrado la misma tendencia usando calibradores de líneas fuertes para  $\log(U)$  (p. ej. [Sánchez et al., 2015b](#)). Sin embargo, el decremento no se observa siempre. Con algunos calibradores, por ejemplo con el calibrador NebulaBayes ([Thomas et al., 2018](#)) no se observa una tendencia clara (ver Fig. F1 en [Espinosa-Ponce et al., 2022](#)). El origen de la relación entre  $\log(U)$  y  $\log(N/O)$  no está clara ([Ji & Yan, 2022](#)), algunas de las posibles explicaciones son: (i) una relación entre la función de masa inicial y la metalicidad, (ii) un efecto de la absorción del polvo y (iii) un efecto del ambiente donde se formó la región H II ([Dopita & Evans, 1986](#)).

Las relaciones entre  $O/H$ ,  $\log(U)$ , y  $\log(N/O)$  tienen una diferente naturaleza y cada una de ellas son importantes en la física del gas ionizado. En el trabajo presentado aquí proponemos estudiar el efecto de las relaciones  $N/O$  vs  $O/H$  y  $\log(U)$  vs  $O/H$  sobre las líneas de emisión que pueden ser observadas y comparar con una muestra observacional. Esto con la finalidad de explorar cuáles son las tendencias que mejor ajustan los datos observados y obtener restricciones sobre la forma de las dos relaciones basadas en un mejor ajuste de las observaciones. Para realizar una correcta exploración, necesitamos generar un gran número de modelos variando los parámetros que definen las dos relaciones. Dado que el tiempo de cómputo supone una restricción de los códigos de fotoionización, hemos decidido utilizar un programa basado en el uso de métodos de inteligencia artificial (IA, por sus siglas en inglés) el cual es considerablemente más rápido en lugar del código de fotoionización Cloudy. Esto nos permitirá explorar el espacio de parámetros de una manera más fluida y eficiente. Sin embargo, esto no implica que dejemos de usar completamente los códigos de fotoionización. Esto se explicará en las secciones posteriores.

### 5.1.2. Modelos de fotoionización para el estudio de regiones HII

La inteligencia artificial nació alrededor de 1950 con el fin de hacer que las computadoras “pensaran”. Dos grandes ramas de la inteligencia artificial son el aprendizaje automatizado (ML, por sus siglas en inglés de *machine learning*) y el aprendizaje profundo (DL, por sus siglas en inglés de *deep learning*). Hay dos tipos de aprendizaje: I) supervisado, y II) no supervisado. Hay dos áreas donde el aprendizaje supervisado es útil: en problemas de clasificación y problemas de regresión.

El uso de técnicas de aprendizaje automatizado para estudios astrofísicos ha ido en aumento rápidamente (p. ej. [Kim & Brunner, 2017](#); [Tuccillo et al., 2018](#); [Ho, 2019](#)).

### 5.1.3. Redes Neuronales Artificiales

Una de las técnicas fundamentales en los métodos de ML son las redes neuronales artificiales (ANN por sus siglas en inglés *Artificial Neural Networks*). Una ANN es un modelo computacional inspirado en el cerebro humano, el cual establece las relaciones de un conjunto de datos basado en un conjunto de algoritmos matemáticos.

Una red neuronal artificial se conforma de un conjunto interconectado de *neuronas*. Una neurona (también llamados nodos) es un modelo de su contraparte biológica que convierte las entradas de la neurona ( $x_1, x_2, \dots, x_n$ ), un vector de pesos correspondientes a cada una de las entradas ( $w_1, w_2, \dots, w_n$ ) y una constante ( $b$ ), a un valor de salida (un esquema se muestra en la Fig. 5.2). Los pesos  $w_i$  y la constante  $b$  se obtienen a través de un entrenamiento previo de la neurona con datos conocidos. La salida se calcula como la suma pesada de las entradas más

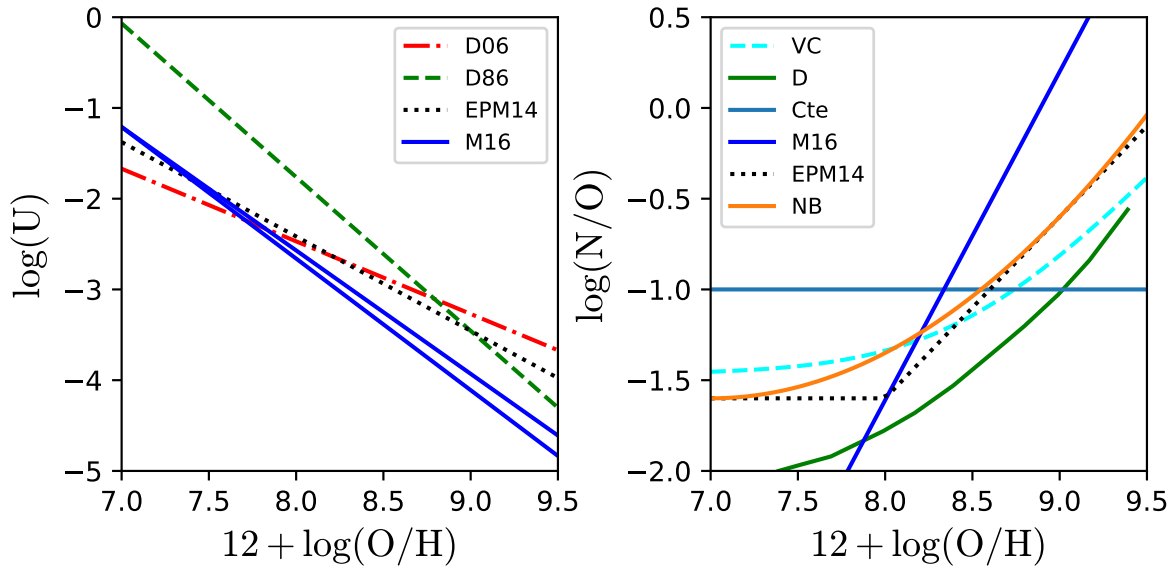


Figura 5.1: Relaciones propuestas en la literatura entre  $\log(U)$ ,  $12 + \log(O/H)$ , y  $\log(N/O)$ . *Panel Izq*: Relación entre  $\log(U)$  y  $12 + \log(O/H)$ . Se muestran las relaciones de [Dopita & Evans \(1986\)](#), línea verde discontinua), [Dopita et al. \(2006\)](#), línea roja punto-guión). *Panel Der*: Relación entre  $\log(N/O)$  y  $12 + \log(O/H)$ . Se muestran las relaciones de [Dopita et al. \(2013\)](#), línea verde), [Vila-Costas & Edmunds \(1993\)](#), línea azul claro discontinua), [Thomas et al. \(2018\)](#), línea naranja) y una relación constante como línea azul. En ambos paneles se muestran las relaciones encontradas por [Morisset et al. \(2016\)](#), línea azul marino) y [Pérez-Montero \(2014\)](#), línea negra punteada).

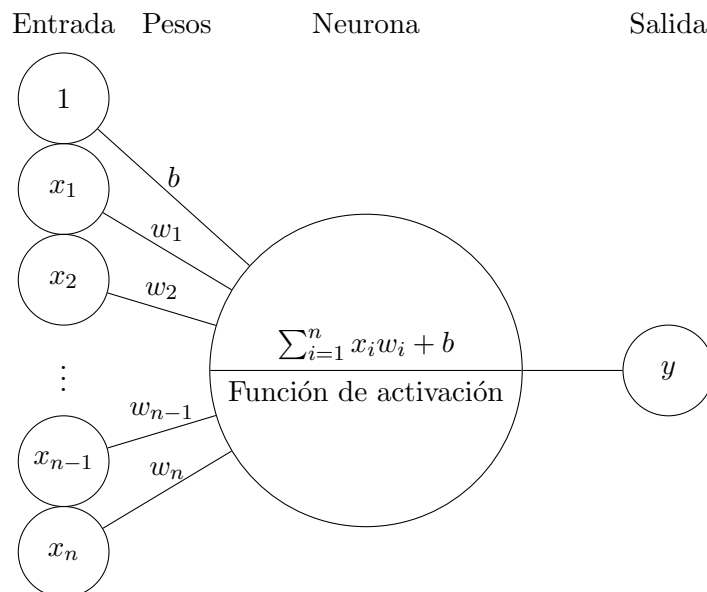


Figura 5.2: Esquema del modelo de una neurona para su uso en una red neuronal artificial. A la izquierda se indican las entradas y sus pesos correspondientes. La salida  $y$  es el resultado de la suma pesada de las entradas procesada por la función de activación.

la constante  $b$ . El resultado puede ser cualquier valor numérico entre 0 a infinito. Para limitar la respuesta del nodo y normalizar el valor resultante de la suma, se aplica una función de activación, la cual puede lineal o no lineal. En este caso el valor numérico resultante está entre 0 ó -1 y 1. La salida de una neurona artificial puede ser la entrada de otro nodo dependiendo de la estructura de la ANN propuesta (Aggarwal, 2018).

La función de activación debe de escogerse con cuidado. Dependiendo de esta función, el modelo puede ajustar mejor o peor, dado que de ella depende la calidad del aprendizaje de los datos (p. ej. Szandala, 2021). Hay diversos escenarios complejos donde una ANN se puede usar, pueden adoptarse funciones de activación lineales y/o no lineales. Entre las funciones más usadas están (ver Fig. 5.3): una función lineal en la parte positiva y nula si su entrada es negativa, una función sigmoide, una función tanh, etc.

La ANN usada en este trabajo fue implementada en el código AI4NEB (Morisset, *et al.* (*in prep.*)<sup>1</sup>. Este software nos permite entrenar una ANN y usarla como regresor fácilmente. Un regresor es la parte funcional de una regresión. Es decir, la función que nos proporciona la relación entre variables independientes y variables dependientes. Estas técnicas son usadas en *machine learning* para obtener un modelo predictivo en el cual un algoritmo es usado para predecir una salida continua.

## 5.2. Metodología

### 5.2.1. Base de datos observacional

Para la exploración de las propiedades físicas de las regiones H II con modelos de fotoionización se necesita una base de datos observacionales para poder contrastar ambos resultados. Los datos observacionales deben incluir la información de las líneas de emisión más intensas

<sup>1</sup>Alojado en su página de Github: <https://github.com/Morisset/AI4neb>



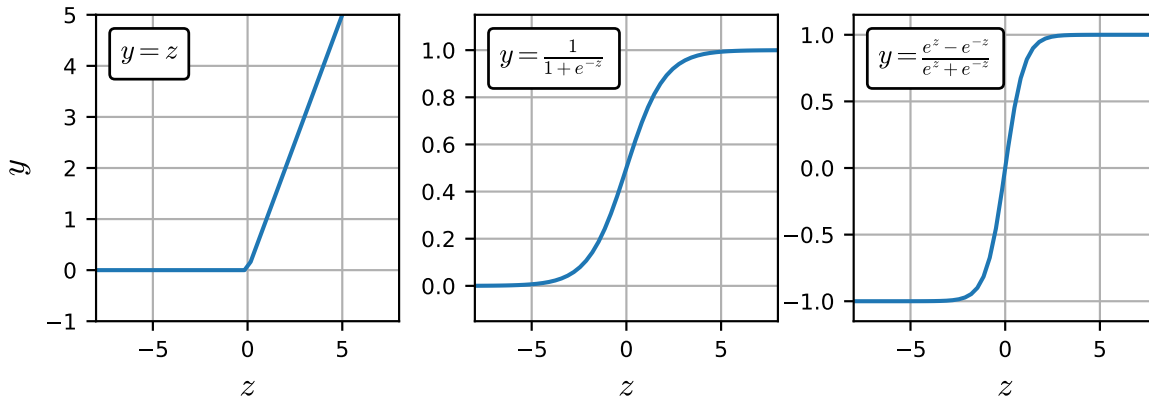


Figura 5.3: Ejemplos de funciones de activación. Panel *Izq.* Función Lineal; Panel *Med.* Función tanh; Panel *Der.* Función Sigmoide.

y de líneas aurales (sensibles a la  $T_e$ ). Esto nos permitirá calcular sus propiedades, como la abundancia de oxígeno. Como ya se mencionó anteriormente (ver Sec. 5.1.1), la abundancia de oxígeno esta estrechamente ligada con la estructura y estado de ionización de la nebulosa.

Se escogió la muestra observacional recopilada por [Marino et al. \(2013\)](#) con la cual vamos a comparar los modelos. Se escogió esta muestra porque todas las regiones H II cuentan con la abundancia de oxígeno calculada por medio del método directo (ver Sec. 1.2.4 del Cap. 1). Además, cuenta con la información espectroscópica de las líneas de emisión más intensas como [O III]  $\lambda 5007$ , [N II]  $\lambda 6583$ , [N II]  $\lambda 5755$  y [O II]  $\lambda 3726+29$  y de la línea auroral [O III]  $\lambda 4363$ .

En la Fig. 5.4 se muestran los datos observacionales en diferentes diagramas de diagnóstico. Podemos ver que en el diagrama BPT (panel arriba-izquierda) las regiones caen en la zona de formación estelar definida por las líneas de demarcación usuales ([Kewley et al., 2001](#); [Kauffmann et al., 2003](#)), como era de esperar. Además hemos incluido las líneas de [Stasińska et al. \(2008\)](#) y [Espinosa-Ponce et al. \(2020\)](#) las cuales son más restrictivas para delimitar la zona de regiones de formación estelar. Vemos que la distribución de la muestra de datos observacional coincide con el área definida por estas dos últimas curvas.

### 5.2.2. Base de datos de modelos de fotoionización

La base de datos de modelos de fotoionización escogida para entrenar la ANN será la base de libre acceso 3MDB (Mexican Million Models database [Morisset et al., 2015](#); [Alarie & Morisset, 2019](#)). Esta base de datos es un esfuerzo por recopilar las redes de modelos completas de diversos estudios con un formato sencillo de utilizar, usando el protocolo MYSQL. La base de datos comprende los modelos utilizados en los trabajos de [Flores-Fajardo et al. \(2011\)](#), [Cid Fernandes et al. \(2013\)](#), [Delgado-Inglada et al. \(2014\)](#), [Pérez-Montero \(2014\)](#), [Morisset et al. \(2016\)](#) y [Vale Asari et al. \(2016\)](#). Todos los conjuntos de modelos antes mencionados fueron creados en su totalidad con el código de fotoionización CLOUDY ([Ferland et al., 2017](#)). Recientemente fue agregado un conjunto de modelos a la base de datos constituido por modelos con choques radiativos generado con el código de fotoionización y de choques MAPPINGS V ([Sutherland & Dopita, 2017](#)).

De todos los conjuntos de modelos disponibles en 3MDB, en este trabajo se usarán los modelos correspondientes al trabajo del método BOND para determinación de abundancias (descrito en detalle en [Vale Asari et al., 2016](#)). Se escogió este conjunto de modelos dado que son modelos de regiones H II calculados con la última versión publicada de CLOUDY ([Ferland](#)

et al., 2017). Estos modelos suponen que una región H II es una nebulosa ionizada por un brote instantáneo de formación estelar. Los principales parámetros que definen al gas ionizado son la abundancia de oxígeno y de nitrógeno, el parámetro de ionización y la edad del brote estelar. Para la abundancia de oxígeno se tomaron 15 valores en un rango de 6.6 dex a 9.4 dex (en escala de  $12 + \log(\text{O}/\text{H})$ ). Para la abundancia de nitrógeno se tomaron 9 valores en el rango de  $-2$  dex a  $0$  dex (en escala de  $\log(\text{N}/\text{O})$ ). Para el parámetro de ionización fueron tomados 13 valores en un rango de  $-4$  dex a  $-1$  dex (en escala de  $\log(U)$ ). Tomamos dos casos extremos de distribuciones de densidad para nuestros modelos: i) el caso de una esfera uniforme llena de gas y ii) el caso de un delgado cascarón uniforme esférico lleno de gas. Por último, la SED (ver Cap. 1.4.1) se obtiene por el código de síntesis de poblaciones POPSTAR (Mollá et al., 2009) con una función inicial de masa tipo Chabrier (2003). La edad del brote estelar toma valores de 1 a 6 Myr con un intervalo de 1 Myr. Este intervalo corresponde a la edad de las poblaciones estelares que, por definición, ioniza al gas en una región H II. Para el gas ionizado se aplica el valor por defecto en CLOUDY al usar la opción *ism*, el cual es una mezcla típica de silicatos y carbonatos. Para tomar en cuenta otras contribuciones de polvo, como el correspondiente al gas neutro, se aplica un factor siguiendo el trabajo de Stasińska et al. (2015). Finalmente, para tomar en cuenta el espesor óptico de las regiones modeladas se agrega un parámetro llamado *HbFrac* que es la fracción de  $\text{H}\beta$  relativo a un modelo de una nebulosa limitada por radiación.<sup>2</sup> Por lo que si  $\text{HbFrac} < 0.9$ , solo tomaríamos el  $< 90\%$  de la nebulosa en nuestro modelo, por lo que existiría el escape de fotones, por tanto sería una nebulosa limitada por densidad/masa. Para valores  $\text{HbFrac} > 0.9$  se toma como si fuera una nebulosa limitada por radiación.

### 5.2.3. Entrenamiento de la ANN

La red neuronal se compone de cinco capas: una capa constituida por los parámetros de entrada, una capa de salida y tres capas ocultas. Cada una de las capas ocultas está compuesta por 50 neuronas con una tangente hiperbólica como función de activación.

Los parámetros de entrada de nuestros modelos de fotoionización (y de la ANN) son esencialmente los mismos que los utilizados para los modelos de BOND antes descritos (ver Tabla 5.1), estos son la abundancia de O, el cociente N/O y el parámetro de ionización,  $U$ . Para tomar en cuenta las dos distribuciones de densidad definidas en los modelos de BOND, definiremos el factor de morfología,  $f_r = R_{in}/R_{Str}$ , donde  $R_{in}$  es el radio interno de la nebulosa (distancia entre la fuente ionizante y el gas) y  $R_{Str}$  es el radio de Strömgren si la nebulosa fuera una esfera llena de gas. Finalmente, tomamos 5 edades de la población ionizante entre 1 Myr a 5 Myr, en intervalos de 1 Myr.

Los parámetros de salida de nuestros modelos y de la misma ANN, corresponden a líneas de emisión normalizadas por  $\text{H}\beta$  (mostrados en la Tabla 5.1). Recordemos que contamos con toda la información que nos proveen los modelos de fotoionización. De todo este conjunto de datos, escogimos las líneas de emisión correspondientes a los principales diagramas de diagnóstico (ver Fig 5.4):  $[\text{O II}] \lambda 3727/\text{H}\beta$ ,  $[\text{O III}] \lambda 5007/\text{H}\beta$ ,  $[\text{N II}] \lambda 6584/\text{H}\beta$  y los cocientes de línea:  $[\text{O III}] \lambda \lambda 4363/5007$  y  $[\text{N II}] \lambda \lambda 5755/6584$ . La muestra final con la cual se va a entrenar la ANN está compuesta de los valores logarítmicos de estas razones de línea. Finalmente, hemos agregado como salida el cociente de las fracciones iónicas de  $\text{N}^+$  y  $\text{O}^+$ .

Para verificar que los resultados derivados por la ANN son equiparables a los obtenidos con

---

<sup>2</sup>Una nebulosa de emisión es limitada por radiación si los fotones ionizantes no son suficientes para ionizar el gas circundante. Por otro lado, una nebulosa es limitada por masa si la estrella asociada ionizada todo el gas circundante.

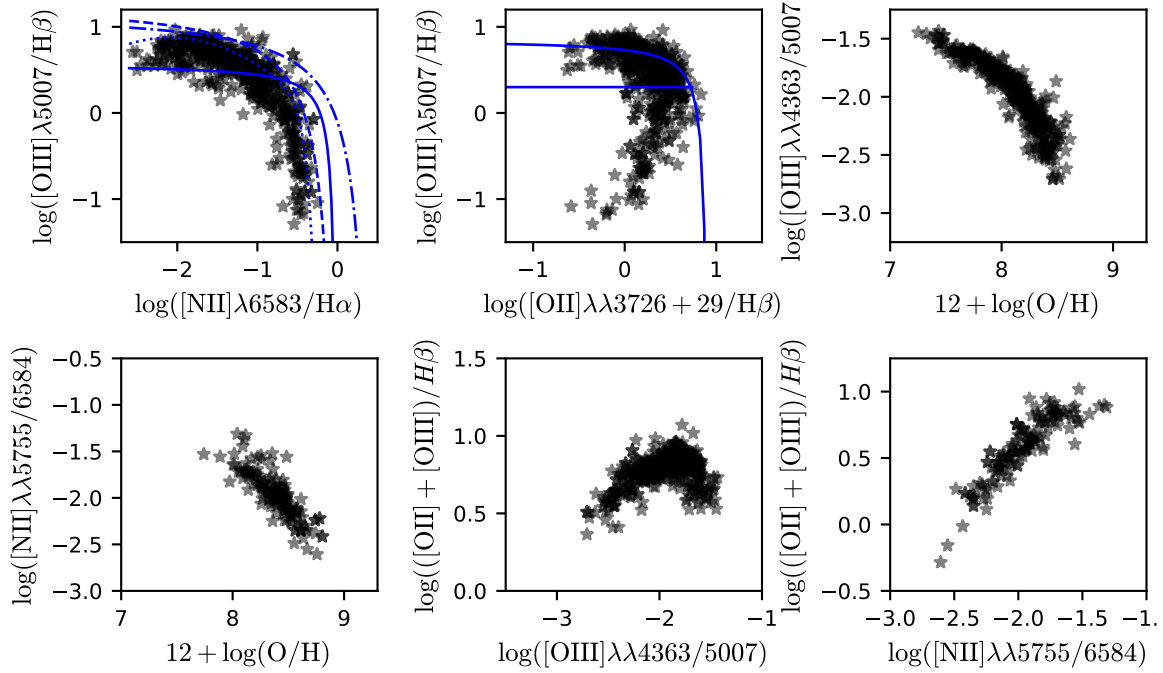


Figura 5.4: Distribución de los datos observacionales de la muestra adoptada de regiones H II en diferentes diagramas de diagnóstico. *Panel Sup-Izq*: diagrama BPT; *Panel Sup-Med*: diagrama  $[\text{O III}] \lambda 5007/\text{H}\beta$  vs  $[\text{O II}] \lambda\lambda 3726+29/\text{H}\beta$ ; *Panel Sup-Der*:  $[\text{O III}] \lambda\lambda 4363/5007$  vs abundancia de O (en unidades de  $12+\log(\text{O}/\text{H})$ ); *Panel Inf-Izq*:  $[\text{N II}] \lambda\lambda 5755/6584$  vs abundancia de O (en unidades de  $12+\log(\text{O}/\text{H})$ ); *Panel Inf-Med*: diagrama  $([\text{O II}] \lambda\lambda 3726, 3729+[\text{O III}] \lambda 5007)/\text{H}\beta$  vs  $[\text{O III}] \lambda\lambda 4363/5007$ ; *Panel Inf-Der*: diagrama  $([\text{O II}] \lambda\lambda 3726, 3729+[\text{O III}] \lambda 5007)/\text{H}\beta$  vs  $[\text{N II}] \lambda\lambda 5755/6584$ . En el diagrama BPT se muestran las líneas de demarcación típicas para separar la zona de SF y AGNs (línea punteada [Kewley et al., 2001](#)) y (línea discontinua [Kauffmann et al., 2003](#)) además de la línea dada por [Espinosa-Ponce et al. \(2020\)](#), (línea continua). En el panel Sup-Med se muestran las líneas de demarcación, una que separa las zonas de SF y AGN en este diagrama (línea continua, [Lamareille, 2010](#)) y otra que separa a las regiones SF y aquellas que es una combinación de SF y AGN (línea discontinua, [Lamareille, 2010](#)).

Parámetros de entrada	Descripción
O/H	Abundancia de oxígeno (en escala de $12 + \log(\text{O}/\text{H})$ )
N/O	Cociente de la abundancia de N y O (en escala de $\log(\text{N}/\text{O})$ )
U	Parámetro de ionización (en escala de $\log(\text{U})$ )
fr	Factor morfológico
edad	Edad de la población estelar ionizante (en escala de Myr)
HbFrac	Fracción de H $\beta$ relativa a un modelo de una nebulosa limitada por radiación
Parámetros de salida	Descripción
O3	[O III] $\lambda 5007/\text{H}\beta$
N2	[N II] $\lambda 6583/\text{H}\beta$
O2	[O II] $\lambda 3727/\text{H}\beta$
rO3	[O III] $\lambda\lambda 4363/5007$
rN2	[N II] $\lambda\lambda 5755/6583$
Np-Op	Cociente de las fracciones iónicas de N $^+$ y O $^+$

Tabla 5.1: Resumen de los parámetros de entrada y de salida de nuestros modelos de fotoionización y, por consiguiente, de la ANN.

modelos de fotoionización se realizó la siguiente prueba. Se entrenó la ANN usando solamente el 80 % de la muestra de modelos de fotoionización antes descrita. El otro 20 % de la muestra se utilizó para comparar los resultados dados por la red neuronal entrenada y los propios modelos. Es decir, parte de la muestra original se usa para probar los resultados de la red neuronal.

En la Fig. 5.5 se muestra la diferencia entre los valores de las observables de los modelos de fotoionización y el resultado obtenidos de la ANN entrenada. Las observables que se muestran son: O3 ([O III]  $\lambda 5007/\text{H}\beta$ ), N2 ([N II]  $\lambda 6583/\text{H}\beta$ ), O2 ([O II]  $\lambda 3727/\text{H}\beta$ ), rO3 ([O III]  $\lambda\lambda 4363/5007$ ), rN2 ([N II]  $\lambda\lambda 5755/6583$ ), y finalmente, Np-Op (cociente entre las fracciones iónicas de N $^+$  y O $^+$ ). Podemos observar que para cada uno la distribución de la diferencia están centrados en cero. Asimismo la desviación estándar de cada distribución es del orden de 0.03 dex. Por lo tanto, los resultados de la ANN son congruentes con los resultados de los modelos de fotoionización y, por tanto podemos sustituir el código CLOUDY en el cálculo de los modelos de fotoionización. Es decir, si tomamos los códigos de fotoionización como un programa que toman valores de entrada y generan valores de salida, una ANN puede sustituir este programa tomando los mismos valores de entrada y generando valores de salida similares. Esto será cierto siempre y cuando la red de modelos usada para entrenar la ANN cubra de manera suficiente el espacio de parámetros de entrada y de salida.

Posteriormente a comprobar la fiabilidad de nuestra ANN entrenamos nuevamente. Esta vez, la nueva red neuronal se ha entrenado con el 100 % de la muestra de la base de datos de modelos. Sin embargo la ANN utilizada para para generar la Fig. 5.5 podría haber funcionado igual. Los parámetros de entrada y de salida son iguales a los descritos anteriormente. Como ya se mencionó en el párrafo anterior, con una ANN propiamente entrenada se puede generar un modelo de fotoionización, al cual llamaremos modelo ANN. En este caso, podemos generar un modelo ANN con esta red neuronal dado una valor de  $12 + \log(\text{O}/\text{H})$ ,  $\log(\text{U})$ , y  $\log(\text{N}/\text{O})$  para una edad,  $fr$  y  $\text{Hbfrac}$  dados. Por lo tanto, podemos producir un conjunto de modelos siguiendo relaciones específicas entre  $12 + \log(\text{O}/\text{H})$  y  $\log(\text{U})$  y entre  $12 + \log(\text{O}/\text{H})$  y  $\log(\text{N}/\text{O})$ . Para todos los modelos ANN generados a partir de este momento, los parámetros libres de

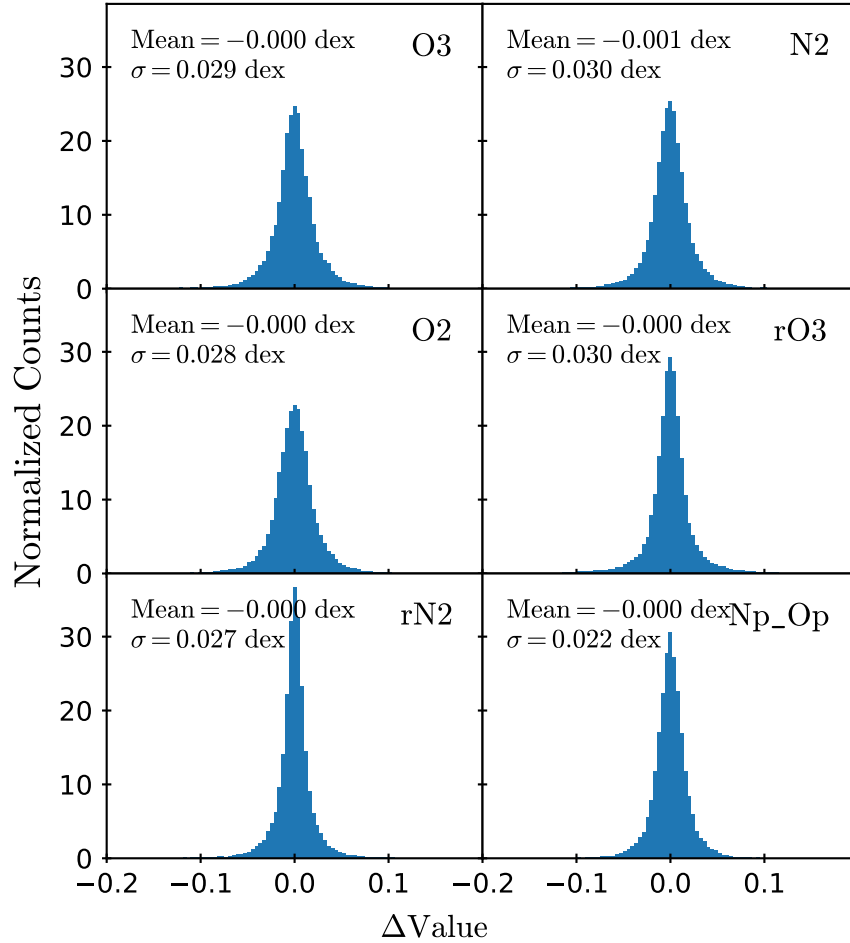


Figura 5.5: Distribución de la diferencia entre los resultados de los modelos de fotoionización y los resultados obtenidos para el 20% de la muestra original de modelos por medio de la ANN entrenada para diferentes parámetros de salida. *Panel Sup-Izq*: O3 ( $[\text{O III}] \lambda 5007/\text{H}\beta$ ), *Panel Sup-Der*: N2 ( $[\text{N II}] \lambda 6583/\text{H}\beta$ ), *Panel Med-Izq*: O2 ( $[\text{O II}] \lambda 3727/\text{H}\beta$ ), *Panel Med-Der*: rO3 ( $[\text{O III}] \lambda \lambda 4363/5007$ ), *Panel Inf-Izq*: rN2 ( $[\text{N II}] \lambda \lambda 5755/6583$ ), y finalmente, *Panel Inf-Der l*: Np\_Op (cociente entre las fracciones iónicas de  $\text{N}^+$  y  $\text{O}^+$ )



nuestros modelos serán  $12 + \log(\text{O}/\text{H})$ ,  $\log(U)$  y  $\log(\text{N}/\text{O})$ . Para  $f_r$  tomaremos solo dos posibles valores posibles: i)  $f_r = 0.03$ , es decir,  $R_{Str} \sim 33R_{in}$  que corresponde a una esfera llena de gas y ii)  $f_r = 3.0$ , es decir,  $R_{Str} = \frac{1}{3}R_{in}$  que corresponde a un cascarón de gas. En este último caso, recordemos que  $R_{Str}$  para una esfera llena de gas, sin embargo, al tener un radio interno grande, el  $R_{Str}$  de una nebulosa cualquiera cambia. El parámetro HbFrac se fijó a 1, es decir, se utilizarán aquellos modelos de las regiones H II donde las nebulosas están limitadas por radiación. Por último, la edad de los brotes estelares que ionizan el gas será de 1 a 5 Myr con una separación de 1 Myr.

Se realizó una primera prueba tomando relaciones conocidas de O/H,  $U$ , y N/O. Tomamos un gradiente negativo entre O/H y  $U$  (en específico, usamos la expresión derivada por [Morisset et al., 2016](#)). Para la relación entre O/H y el cociente N/O elegimos una relación compuesta, inicialmente con un valor constante seguido de un gradiente positivo (en este caso, usamos la relación derivada por [Pérez-Montero, 2014](#)). Ambas relaciones se escogieron por lo discutido en la Sec 5.1.1. Se generaron 100 modelos ANN en total siguiendo las relaciones antes mencionadas entre dos límites para O:  $6.6 \leq 12 + \log(\text{O}/\text{H}) \leq 9.4$ . Por consiguiente, esto también impone límites para  $U$  y N/O dado que ambas dependen de O. Cabe señalar que estos 100 modelos son para un valor de  $f_r$  y una edad de la población estelar ionizante fija. En total, se producen 1000 modelos que corresponden a todos los valores posibles de la edad de la fuente ionizante (de 1 a 5 Myr) y a ambos valores de  $f_r$ .

Podemos ver el resultado del *modelo de referencia* descrito en el párrafo anterior en la Fig. 5.6. Se pueden ver en el panel superiores izquierdo y central las relaciones utilizadas de O/H, N/O y  $U$  para generar el conjunto de modelos ANN. Se muestran los modelos de fotoionización de la base 3MdB (puntos negros) que se encuentran cerca de las relaciones usadas. Esto muestra que existen modelos dentro de la base que siguen de cerca las leyes propuestas. También nos permite verificar que los modelos ANN producidos siguen la ley que queremos usar. En el panel superior derecho se muestra el  $\text{ICF}(\text{N}^+/\text{O}^+)$ . Este valor se ha estudiado ampliamente en la literatura (p. ej. [Garnett, 1990](#); [Pilyugin et al., 2003](#); [Izotov et al., 2004](#); [Nava et al., 2006](#); [Amayo et al., 2021](#)) y usualmente aproximado<sup>3</sup> como  $\text{ICF}(\text{N}^+/\text{O}^+) \approx 1$ , por que  $\text{N}/\text{O} \sim \text{N}^+/\text{O}^+$ . En nuestro modelo de referencia, vemos que el valor del ICF oscila en torno a 1. En los otros paneles se muestra el conjunto de modelos ANN producidos comparados con la distribución de los datos observacionales (ver Sec. 5.2.1) en diversos diagramas de diagnóstico (mismos presentados en la Fig. 5.4). De forma cualitativa podemos ver que los modelos ANN coinciden con los datos observacionales bastante bien. Podemos observar que no hay una diferencia significativa entre ambas geometrías modeladas de las regiones H II (esfera llena o vacía). Por otro lado, la dependencia con la edad de la población ionizante es evidente en algunos diagramas, por ejemplo en el diagrama BPT,  $[\text{O III}]\lambda 5007/\text{H}\beta$  vs  $[\text{N II}]\lambda 6583/\text{H}\alpha$ . Siendo las edades de 5 Myr las que se separan de las distribuciones principales.

Dado que nos interesa más encontrar un conjunto de modelos ANN que puedan reproducir mejor que nuestro modelo ANN de referencia, necesitamos una manera de cuantificar la calidad de nuestros modelos. Con esta finalidad, haremos que nuestros modelos *ajusten* la distribución de los datos observacionales.

<sup>3</sup>Esto es debido a lo similares que son los potenciales de ionización de los dos iones.

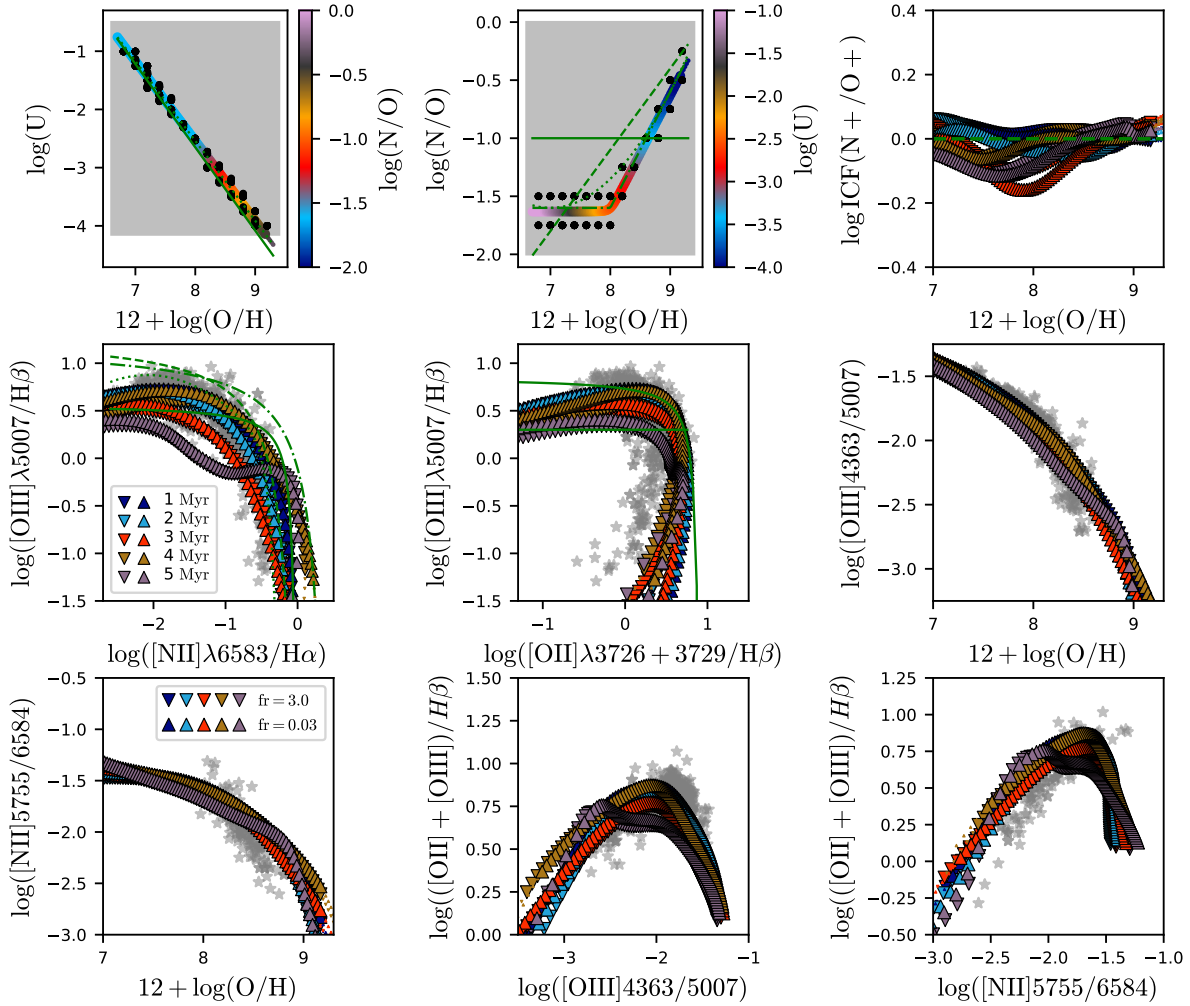


Figura 5.6: Resultado de modelos ANN con relaciones específicas de  $12 + \log(O/H)$ ,  $\log(N/O)$  y  $\log(U)$ . Además se muestra la distribución de los modelos en distintos diagramas. *Panel Sup-Izq* se muestra la relación entre  $12 + \log(O/H)$  y  $\log(N/O)$  usada en este modelo de referencia. La escala de colores representa los valores de  $\log(N/O)$ . Además se muestran como línea sólida y punteada de color verde las relaciones propuestas por [Morisset et al. \(2016\)](#). *Panel Sup-Med* se muestra la relación entre  $12 + \log(O/H)$  y  $\log(N/O)$ . El código de color representa los valores de  $\log(U)$ . Se muestran distintas relaciones propuestas en la literatura: como línea sólida se muestra una relación constante. Como una línea punteada se muestra la relación propuesta por [Thomas et al. \(2018\)](#), como línea discontinua se muestra la relación propuesta por [Morisset et al. \(2016\)](#) y la línea punto-guión es la relación propuesta por [Pérez-Montero et al. \(2016\)](#). En ambos paneles se muestran los modelos de fotoionización que corresponden a las relaciones usadas. *Panel Sup-Der* se muestra el  $ICF(N^+/O^+)$  de los modelos. La línea verde discontinua indica el valor 0. *Panel Med-Izq*: diagrama BPT; *Panel Med-Med*: diagrama  $[O III] \lambda 5007/H\beta$  vs  $[O II] \lambda \lambda 3726 + 29/H\beta$ ; *Panel Med-Der*:  $[O III] \lambda \lambda 4363/5007$  vs abundancia de O (en unidades de  $12 + \log(O/H)$ ); *Panel Inf-Izq*:  $[N II] \lambda \lambda 5755/6584$  vs abundancia de O (en unidades de  $12 + \log(O/H)$ ); *Panel Inf-Med*: diagrama  $([O II] \lambda \lambda 3726, 3729 + [O III] \lambda 5007)/H\beta$  vs  $[O III] \lambda \lambda 4363/5007$ ; *Panel Inf-Der*: diagrama  $([O II] \lambda \lambda 3726, 3729 + [O III] \lambda 5007)/H\beta$  vs  $[N II] \lambda \lambda 5755/6584$ . En todos los paneles donde aparecen los modelos producidos por la ANN se muestran las diferentes edades de las poblaciones ionizantes con un diferente color. Igualmente se muestran las dos diferentes distribuciones de densidad (esfera llena de gas,  $f_r = 0.03$  y un para un cascaron de gas,  $f_r = 3.0$ ) con diferentes símbolos en el panel Inf-Izq. .

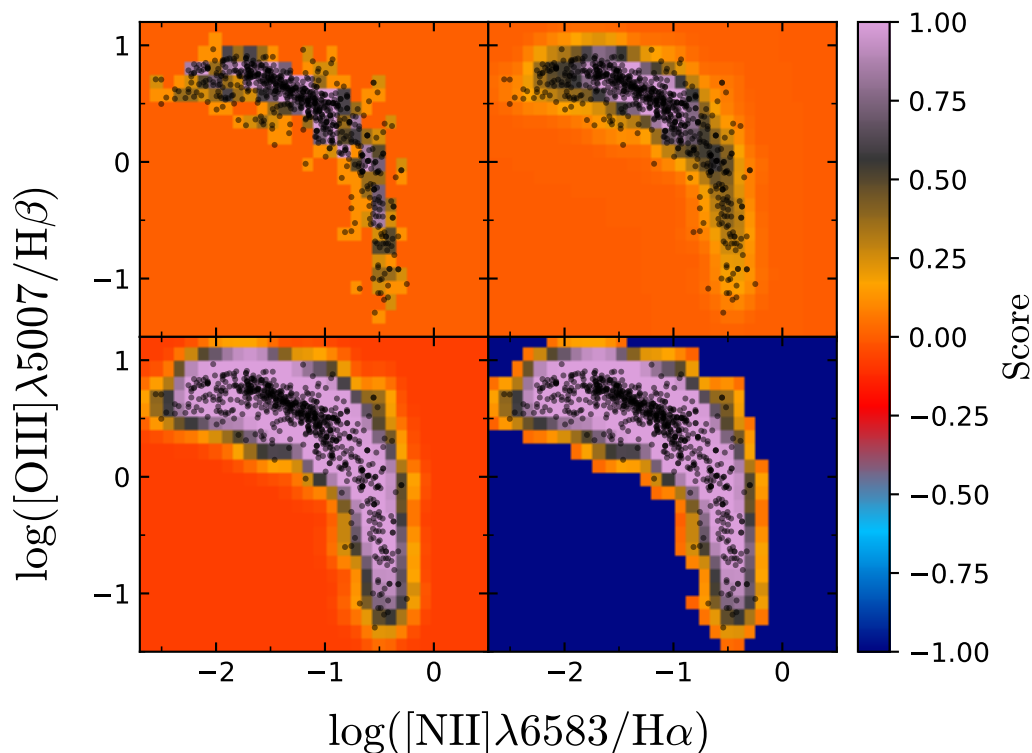


Figura 5.7: Esquema del cálculo de los mapas de densidad. En cada panel se muestran los diferentes mapas obtenidos en cada paso descrito en el texto. *Panel Sup-Izq*: Histograma obtenido a partir de los datos observacionales. *Panel Sup-Der*: Histograma convolucionado con una función gaussiana. *Panel Inf-Izq*: Histograma convolucionado con una función de error (ERF). *Panel Inf-Der*: Histograma final donde se han enmascarado todos los valores menor a 0. El mapa a utilizar en el cálculo del parámetro de verosimilitud es mostrado en el panel inf. der.

#### 5.2.4. Ajuste de los datos observacionales

##### Calidad de los modelos obtenidos

Como se puede observar en la Fig. 5.6, las observaciones presentan una distribución definida en cada diagrama de diagnóstico. Para cuantificar la calidad de nuestros modelos ANN, mediremos la similitud de su distribución con los datos observacionales. Como primer paso, se calculó el mapa de densidad de los datos observacionales de cada diagrama de diagnóstico. A cada modelo ANN se le asignará un valor, (*score*), que dependerá directamente de la densidad de puntos en cada diagrama.

Para esto se calculó un histograma en dos dimensiones para un diagrama de diagnóstico dado (ver panel sup-izq de la Fig. 5.7) con una rejilla de  $30 \times 23$ . Se eligió este número de elementos por ser el mínimo número de elementos con los que se recupera con suficiente detalle la distribución de cada diagrama. El histograma obtenido a partir de los datos observacionales compone una zona bien definida donde se encuentran la mayoría de puntos. Sin embargo, en los extremos de la distribución principal se tiene saltos muy abruptos en el histograma. Esto puede causar que el mapa de densidad sea muy restrictivo para nuestros modelos ANN generados. Para hacer más suave la transición en estas zonas del mapa de densidad, el histograma fue

convolucionado con una función gaussiana (panel sup-der en la Fig. 5.7). Sin embargo, la distribución obtenido es muy estrecha y nos interesa tener la mayor contribución posible por cada punto observador. Para esto se convolucionó el último mapa obtenido con una función de error (ERF). Esto provocará que se obtenga un mayor contraste entre la zona con datos y la zona donde no hay o se tenga muy pocos datos. El efecto resultante es que la distribución cubre la mayor área posible basado en donde hay la mayor cantidad de puntos. Finalmente, todos aquellos valores que sean  $\leq 0$  se fijan a  $-1$ . Esto es para que, si algunos modelos caen fuera de la distribución definida por los datos observacionales, contribuyan de manera negativa. Se puede apreciar cada uno de los pasos anteriormente descritos en la Fig. 5.7.

Utilizando este método, se calculó un mapa de densidad para cada uno de los diagramas de diagnóstico definidos. Como se puede ver en la Fig. 5.8, cada mapa de densidad tiene un valor máximo donde se concentra la mayoría de los puntos, mientras tanto, el valor decrece rápidamente cuando se aleja de la distribución principal.

Para cada diagrama de diagnóstico se definió el parámetro de verosimilitud,  $f_{\text{diag}}$ , que cuantifica la calidad del ajuste de un conjunto de modelos ANN con los datos observacionales en un diagrama dado. Para esto se hace una *pseudo* mapa de densidad de los modelos ANN. Se crea una rejilla del mismo tamaño a la creada para las observaciones, salvo que en cada segmento solo podrá contener valores 0 y 1. El 1 representa que al menos existe un modelo ANN, mientras que el 0 representa que no existe ningún modelo. Se define un mapa de densidad de cada diagrama de los modelos para todas las combinaciones de valores  $fr$  y edades de las poblaciones ionizantes. Finalmente, se convoluciona con una función gaussiana por la misma razón que se hizo para el mapa de densidad de las observaciones. Por lo que, para un diagrama dado,  $f_{\text{diag}}$  se calcula como:

$$f_{\text{diag}} = \sum_{fr, age} m_{\text{diag}} * Gauss \quad (5.1)$$

donde  $m_{\text{diag}}$  es el mapa de densidad para los modelos antes mencionado y *Gauss* es una función gaussiana.

Por lo tanto, el parámetro de verosimilitud total,  $f_{\text{total}}$ , de un conjunto de modelos ANN tomando en cuenta todos los diagramas de diagnóstico se calcula como

$$f_{\text{total}} = \sum_{\text{diag}} w_{\text{diag}} f_{\text{diag}}, \quad (5.2)$$

$f_{\text{diag}}$  corresponde al parámetro estimado para cada diagrama de diagnóstico,  $w_{\text{diag}}$  corresponde a un peso único que se asigna a cada diagrama. Este valor nos permite darle un mayor peso a un diagrama de diagnóstico en específico, por ejemplo el diagrama BPT.

Este parámetro se calcula para cada conjunto de modelos ANN generados. Sin embargo, el espacio de parámetros es lo suficientemente complejo para que existan millones de posibles combinaciones que den resultados parecidos. Para encontrar el conjunto de modelos que maximice el parámetro  $f_{\text{total}}$  se utilizará un método genético.

## Método genético

En este punto, nuestro problema se reduce a una optimización matemática de las propiedades físicas de las regiones H II que maximizan  $f_{\text{total}}$ . Cabe señalar que nosotros no buscamos los valores de O/H, N/O y  $U$  que reproduzcan las observaciones. En este caso nos interesan los parámetros que definen las relaciones entre O/H y N/O y  $U$ . Por ejemplo, como se mencionó

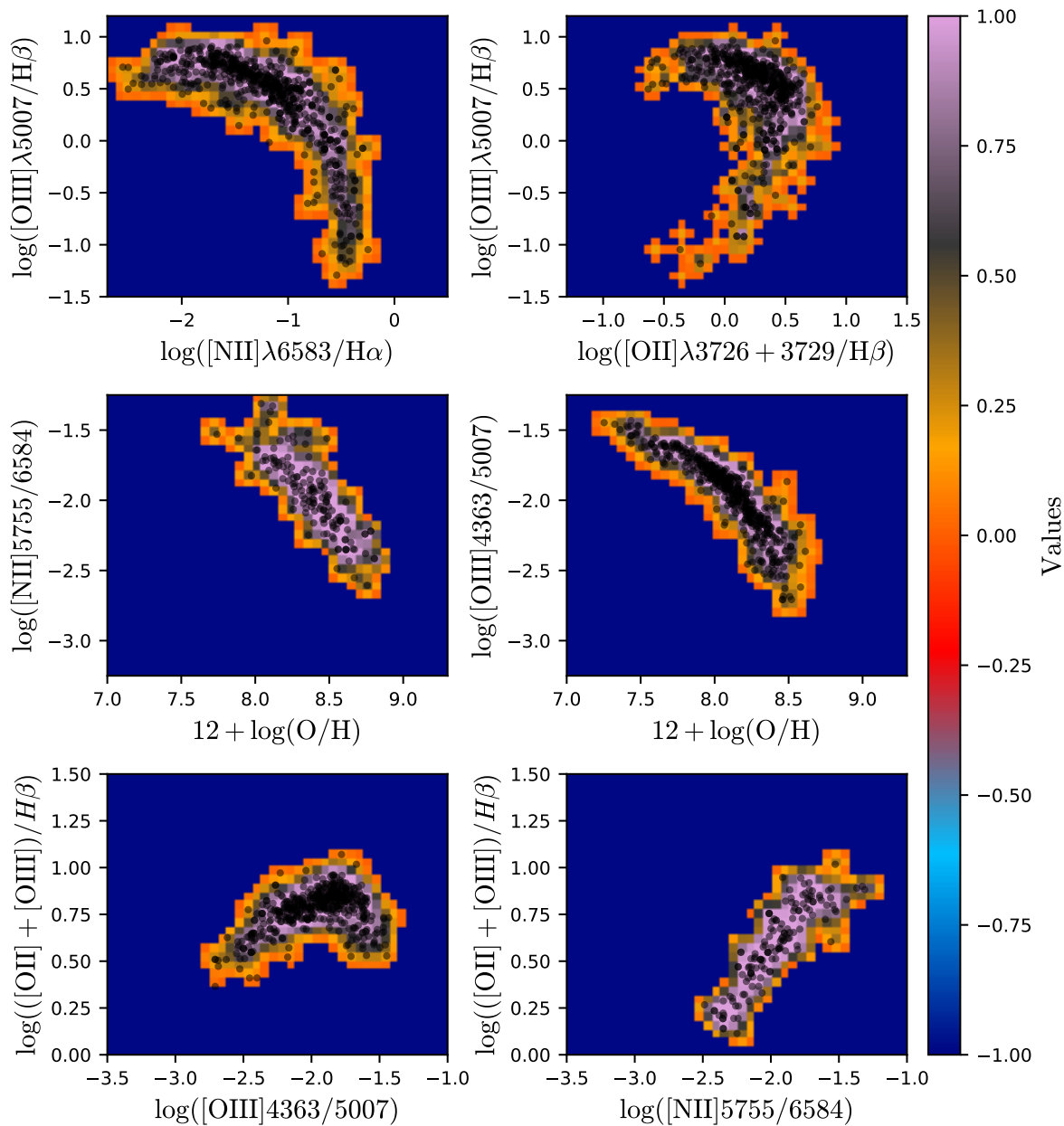


Figura 5.8: Mapas de densidad finales para cada uno de los diagramas de diagnóstico. El código de color representa el valor en cada bin.



en la Sec. 5.1.1, comúnmente se observa que entre O/H y  $U$  se tiene un gradiente negativo con la forma funcional

$$\log(U) = a_0 + a_1 \times (12 + \log(\text{O}/\text{H})), \quad \text{donde } a_1 < 0. \quad (5.3)$$

mientras que entre N/O y  $U$  se observa una relación compuesta tal que:

$$\log(\text{N}/\text{O}) = \begin{cases} a'_0, & \text{si } 12 + \log(\text{O}/\text{H}) < a'_1 \\ a'_2 + a'_3 \times (12 + \log(\text{O}/\text{H})), & \text{si } 12 + \log(\text{O}/\text{H}) > a'_1. \end{cases} \quad (5.4)$$

Si tomamos las relaciones anteriores, nuestro problema se reduce a ajustar las constantes  $a_0$ ,  $a_1$ ,  $a'_0$ ,  $a'_1$ ,  $a'_2$  y  $a'_3$  que den como resultado el máximo valor de  $f_{\text{diag}}$ . Dado que el espacio está altamente degenerado, diversos métodos de optimización, como MCMC, pueden obtener una solución local. Por tanto, para esta tarea haremos uso del método genético originalmente basado de cierta manera en la evolución biológica, los algoritmos genéticos se basan en tres hechos de la evolución: i) selección natural, ii) herencia y iii) variación (Charbonneau, 1995; Cantó et al., 2009). En este trabajo usaremos una versión simple del código genético que nos permitirá obtener las mejores soluciones posibles. El proceso general de este tipo de métodos es mostrado en la Fig. 5.9. Todo inicia a partir de una muestra inicial de modelos<sup>4</sup> con ciertos parámetros que muestren homogéneamente o no el espacio de parámetros de una forma completa. Se eligen los mejores modelos de esta muestra con base en un valor al parámetro de verosimilitud. En general, este parámetro mide la calidad del modelo, su valor es función de los parámetros de un modelo y su calculo dependerá del problema en cuestión. Se seleccionan los  $n$  mejores modelos y, a partir de cada uno de ellos, se generan  $k$  modelos descendientes. Estos nuevos modelos pueden generarse de dos maneras diferentes dependiendo de operador genético utilizado (García-Martínez et al., 2018):

- **Recombinación:** En general este tipo de operador genético se basa en la combinación de las propiedades de los modelos progenitores. Esta recombinación de propiedades puede realizarse de forma homogénea o inhomogénea o de manera uniforme.
- **Mutación:** Este tipo de operador genético se basa en cambiar las propiedades de los modelos. La propiedad que se modifica es seleccionada de forma arbitraria. A su vez, la modificación puede seguir cierta distribución de probabilidad, como una distribución normal, centrada en el valor original de la propiedad.

Cuando se generan los modelos descendientes, pueden usarse uno o ambos operadores genéticos descritos anteriormente. El uso de un operador u otro, así como la implementación del operador en sí dependerá del problema que se está resolviendo. Finalmente, con los modelos progenitores y descendientes se tiene una nueva muestra de modelos con la que se repetirán los mismos pasos. El proceso se itera hasta que, idealmente, el parámetro de verosimilitud de los modelos no presente variación alguna. A cada iteración se le da el nombre de generación, haciendo referencia al proceso biológico asociado.

A continuación describiremos el método genético que hemos usado para nuestro problema. El primer paso es construir la muestra inicial de conjuntos de modelos ANN. Aunque el número de modelos ANN de la muestra inicial puede ser variable, el número de modelos ANN por defecto es 200. En este caso buscamos los parámetros que definen las formas funcionales de

<sup>4</sup>En este caso, la palabra modelos no hace referencia a modelos de fotoionización, si no a los modelos parametrizados por las ecuaciones 5.3 y 5.4.

las relaciones entre O/H y N/O y  $U$ . Por lo tanto, un modelo ANN inicial  $i$ ,  $\text{mod}_0^i$ , es definido por los parámetros que establecen las relaciones de las propiedades físicas y que generan un conjunto de modelos de fotoionización. Para generar el conjunto de modelos ANN iniciales se tiene que realizar un muestreo del espacio de parámetros lo más completo posible. Es decir, tratar que los parámetros iniciales cubran totalmente el espacio de parámetros. Sin embargo, el espacio que definen las constantes  $a_n$  y  $a'_n$  es desconocido, complejo y carece de sentido astrofísico (usando las Ec. 5.3 y 5.4, por ejemplo). Por lo tanto, definiremos estas constantes con base en puntos,  $(12 + \log(\text{O}/\text{H}))_i, \log(\text{N}/\text{O})_i$  y  $(12 + \log(\text{O}/\text{H}))_i, \log(U)_i$  (dado que con estos podemos calcular los coeficientes de los polinomios de las Ecs. 5.3 y 5.4), que definen las relaciones en los diagramas  $12 + \log(\text{O}/\text{H})$  vs.  $\log(U)$  y  $12 + \log(\text{O}/\text{H})$  vs.  $\log(\text{N}/\text{O})$  (ver paneles superiores de la Fig. 5.6). Dependiendo del tipo de relación que se proponga, el número de puntos puede variar.

Con la muestra inicial de modelos ANN,  $\text{mod}_0^i$ , se calcula para cada uno el parámetro  $f_{\text{total}}$  correspondiente (como se explicó en la Sec. 5.2.4). De estos modelos ANN se seleccionan los 25 mejores modelos ANN que tomarán el papel de progenitores. De cada uno de ellos se generaran 50 modelos ANN descendientes usando el operador genético de mutación. En este caso, se modifican los puntos  $(12 + \log(\text{O}/\text{H}))_i, \log(\text{N}/\text{O})_i$  y  $(12 + \log(\text{O}/\text{H}))_i, \log(U)_i$  sumando un número entre -1 y 1 siguiendo una distribución uniforme. Este número decrece con cada generación que se calcula de acuerdo a un coeficiente (0.8 en nuestro caso). Esto tiene como objetivo que cada vez se limite más el área de exploración de los parámetros (bajo el supuesto que el proceso está convergiendo). Esta muestra de modelos ANN descendientes se agrega a la muestra original de modelos y se repite el proceso. Puede variar el número de generaciones que uno desee realizar. Se recomienda que en pruebas iniciales se realicen pocas generaciones,  $\sim 5$ . Se realizaron hasta 20 generaciones en nuestro caso. Sin embargo, esto puede variar dependiendo del problema que se este resolviendo.

### 5.3. Resultados

Siguiendo el método explicado en la sección anterior, se corrieron diversas generaciones para encontrar los mejores parámetros que definieran las relaciones entre  $12 + \log(\text{O}/\text{H})$ ,  $\log(\text{N}/\text{O})$  y  $\log(U)$ , de tal manera que los modelos ANN reprodujeran lo mejor posible las distribuciones de los datos observacionales. Se entrenó la ANN como se discutió anteriormente. Este proceso puede tardar algunas horas en completarse. Generar un modelo ANN con una red ya entrenada es prácticamente instantáneo. Se creó una base inicial de 200 modelos ANN iniciales que muestrean el espacio de parámetros de forma no homogénea. En cada generación se escogen 20 modelos que serán los progenitores de 50 descendientes cada uno. Se dejó correr al código por 20 generaciones. El proceso en total tarda alrededor de un par de horas.

El mejor resultado obtenido a lo largo de las 20 generaciones se muestra en la Fig. 5.10. En este caso hemos restringido que la forma funcional de las relaciones entre  $12 + \log(\text{O}/\text{H})$ ,  $\log(\text{N}/\text{O})$  y  $\log(U)$  siga un resultado ya conocidos. En el caso de  $12 + \log(\text{O}/\text{H})$  y  $\log(U)$  se eligió que siguiera un gradiente negativo, mientras que para el caso de  $12 + \log(\text{O}/\text{H})$  y  $\log(\text{N}/\text{O})$  se eligió que siguiera una función compuesta de una parte constante y posteriormente un gradiente positivo (acorde a los discutido en la Sec. 5.1.1). Se puede ver en la Fig. 5.10 que en la mayoría de los paneles, los modelos reproducen bastante bien la distribución de los datos observacionales. Como ya se había comentado con el modelo de referencia, la diferencia entre ambas distribuciones de densidad es mínima. En algunos diagramas, como en el BPT y O/H vs [O III]  $\lambda\lambda 4363/5007$ , la diferencia es casi imperceptible. Recordemos que estamos

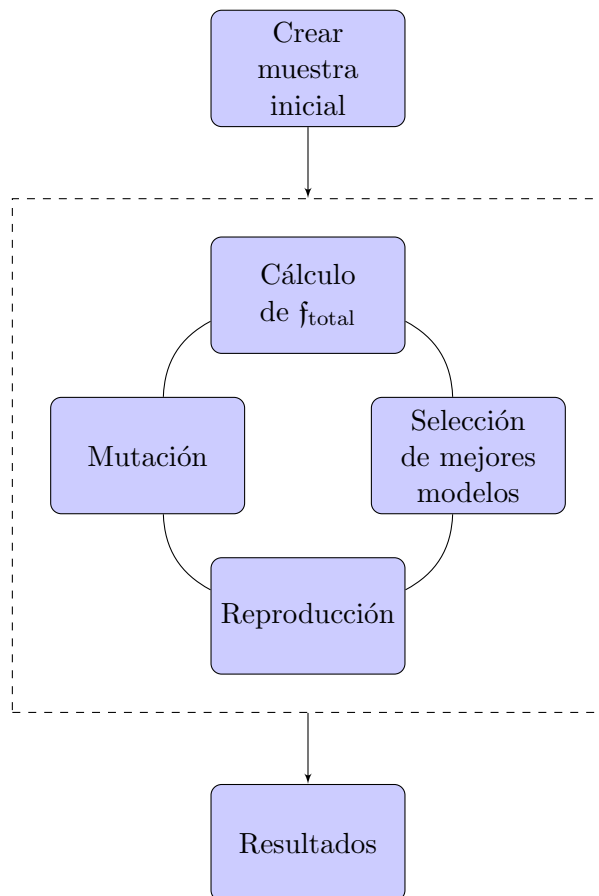


Figura 5.9: Diagrama que muestra los puntos principales de un método genético. A pesar de que se muestra en el diagrama, no siempre es necesario el paso de mutación.

reproduciendo la distribución geométrica de las observaciones. En los otros diagramas se puede observar una diferencia solo en los extremos de los diagramas. Por ejemplo, en el diagrama BPT (Panel Med-Izq), se puede notar una diferencia entre los modelos ANN de esfera llena y esfera vacía de gas en el extremo superior izquierdo del diagrama. Los modelos ANN que representan una esfera vacía son los que tienen un valor más alto el cociente de  $[\text{O III}] \lambda 5007 / \text{H}\beta$ . Esta diferencia también se puede observar en algunos modelos en el extremo inferior derecho del diagrama. Esta diferencia se puede observar también en otros diagramas en el panel central,  $[\text{O II}] \lambda \lambda 3726 + 29 / \text{H}\beta$  vs  $[\text{O III}] \lambda 5007 / \text{H}\beta$ . Se observa el efecto de la edad de la población ionizante. En general, los modelos que tienen una población ionizante con una edad entre 1 y 4 Myr se observa que sus cocientes de línea son similares en todos los diagramas. Sin embargo la edad en que cambia drásticamente la distribución de los modelos es para aquellos con una población ionizante de 5 Myr. Para edades aún mayores, desaparecen todas las estrellas calientes tipo O y para edades mayores a 10 Myr desaparecen las estrellas tipo B, por lo que ya casi no hay fotoionización. Por otro lado, muchas de las estrellas masivas se encuentran en sistemas binarios cercanos. Sin embargo, en este caso no hemos tomado en cuenta los sistemas binarios de modelos de poblaciones estelares (p. ej. [Eldridge et al., 2017a](#)). En general, los cocientes observacionales de línea son menores para esta población ionizante.

Por otro lado, podemos restringir la forma funcional de la relación entre  $12 + \log(\text{O}/\text{H})$  y  $\log(\text{U})$  a otro tipo. Por ejemplo, una relación creciente entre estas dos propiedades. Nuestra metodología nos permite obtener conjunto de modelos ANN correspondientes como se puede ver en la Fig. 5.11. En este caso podemos ver el impacto de la relación entre  $12 + \log(\text{O}/\text{H})$  y  $\log(\text{U})$  y las líneas de emisión obtenidas en los modelos ANN. Se puede observar casi en todos los diagramas que se ajusta sólo cierta parte de la distribución de los datos observacionales. De igual manera, un cambio en la relación entre  $12 + \log(\text{O}/\text{H})$  y  $\log(\text{N}/\text{O})$  puede modificar las líneas de emisión obtenidas. Se requiere una mayor exploración de estas relaciones, la cual se está llevando a cabo actualmente.

## 5.4. Conclusiones

Lo presentado en este capítulo es una revisión del último trabajo realizado en mi doctorado. Los modelos de fotoionización son la herramienta ideal para entender el efecto de las propiedades físicas en las observables de un gas ionizado. El excesivo tiempo de cómputo que puede llegar a tomar el cálculo de un número grande de modelos de fotoionización hace que una exploración completa del espacio de parámetros sea muy costosa (en tiempo de CPU y/o en hardware). En este trabajo hemos propuesto una alternativa a los grandes conjuntos de modelos sustituyendo al código de fotoionización, en este caso CLOUDY, por una red neuronal artificial.

La primera parte del capítulo se centró en justificar que es posible usar una ANN en lugar de CLOUDY. Se mostró que una ANN puede sustituir satisfactoriamente a un código de fotoionización aunque con ciertas precauciones. Para poder realizar este cambio se debe de tomar en cuenta que la base de modelos con la cual se entrenará la ANN debe de muestrear lo mejor posible el espacio de parámetros a estudiar. Los resultados de la red neuronal solo serán válidos dentro del espacio definido por nuestra muestra de entrenamiento. No se puede confiar en los resultados de la ANN fuera del rango definido por la muestra de entrenamiento, es decir, es válida la interpolación y no la extrapolación. Por otro lado, todos los parámetros involucrados en la definición de la red neuronal (desde el número de capas hasta la función de activación de una neurona) se están estudiando actualmente. Algunos de los parámetros

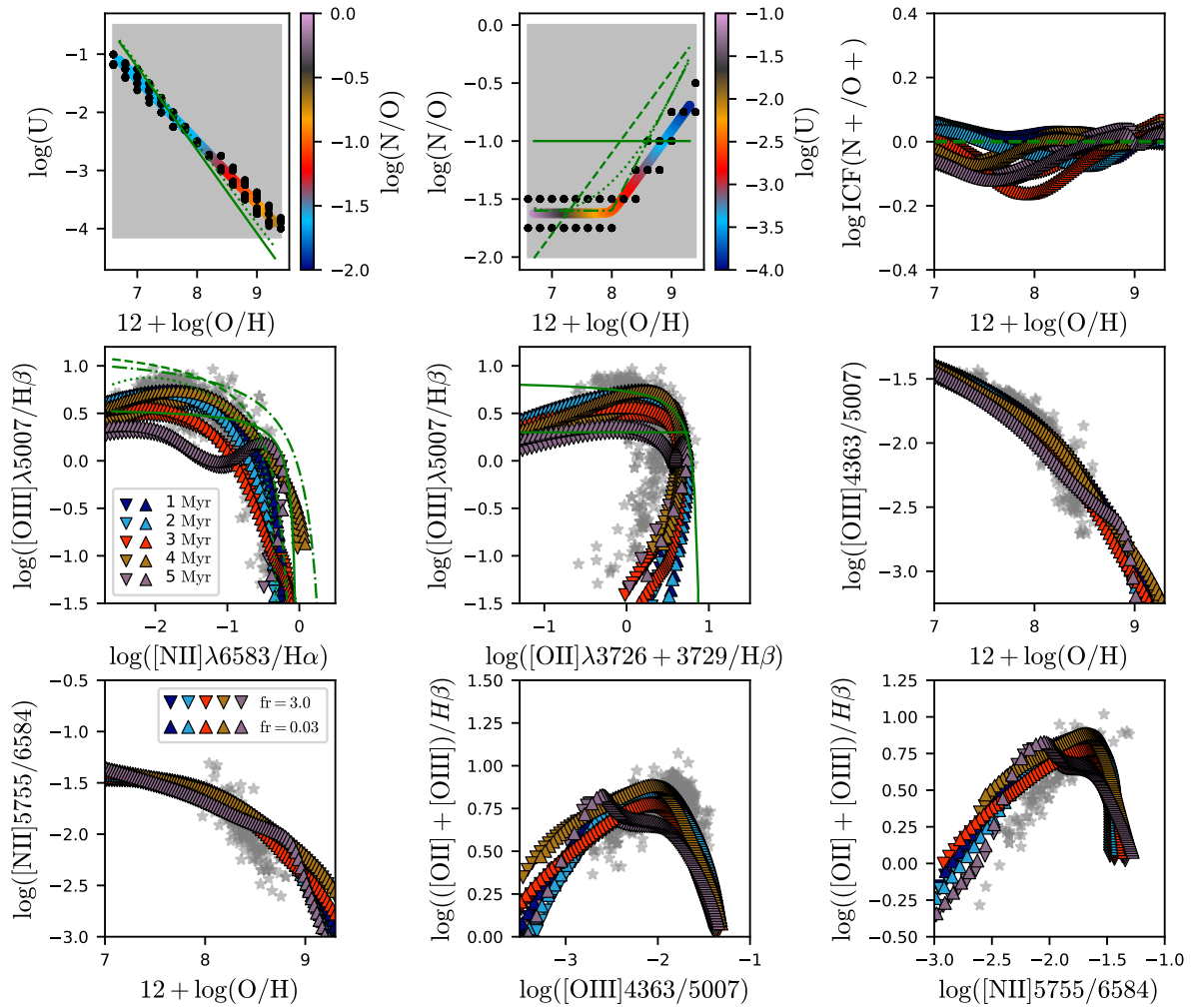


Figura 5.10: Relaciones entre  $12 + \log(O/H)$ ,  $\log(N/O)$  y  $\log(U)$  que se ha usado para definir el conjunto de modelos del mejor resultado obtenido. Además se muestra la distribución de los modelos en distintos diagramas. Esta figura es totalmente análoga a la Fig. 5.6.



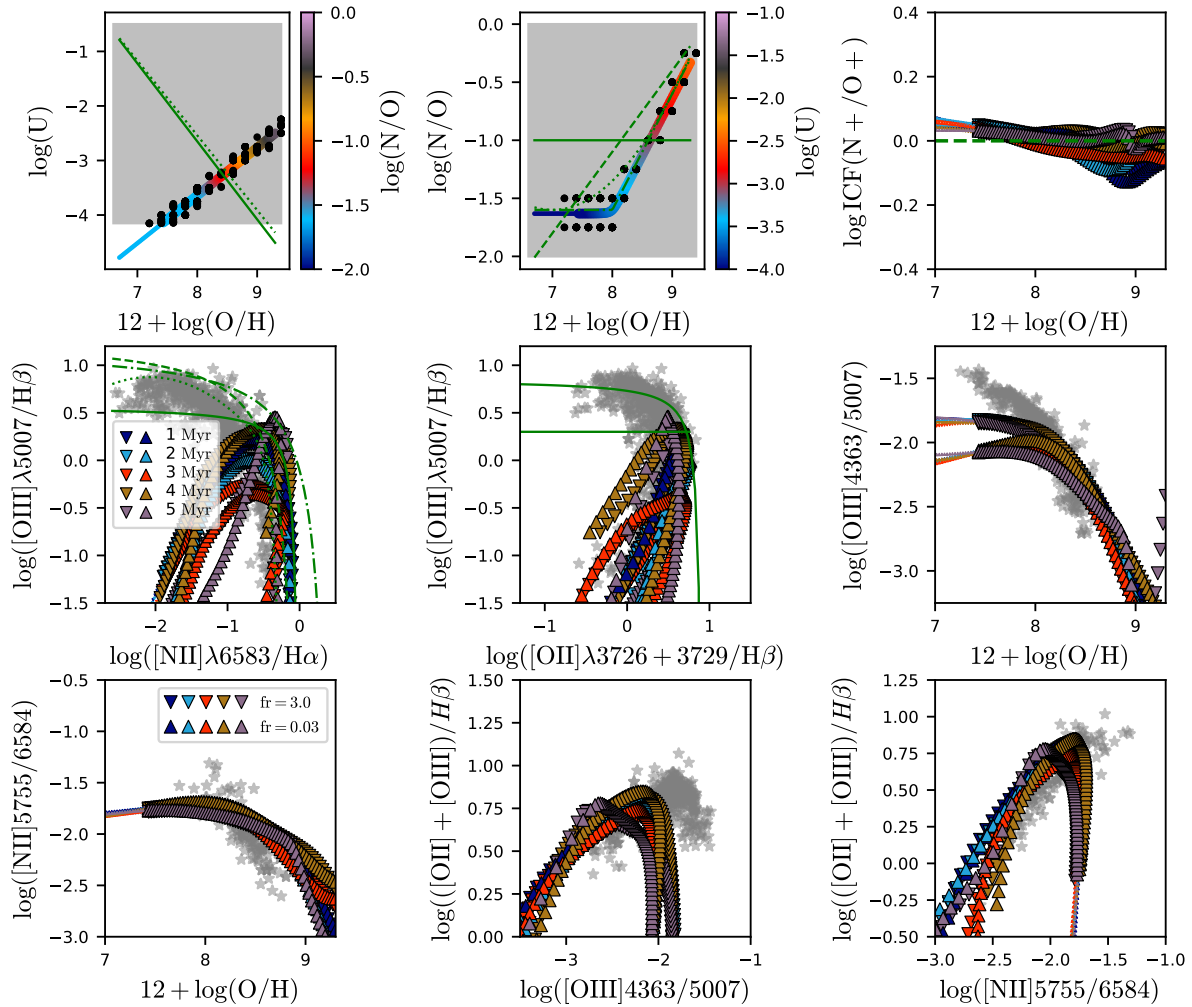


Figura 5.11: Relaciones entre  $12 + \log(O/H)$ ,  $\log(N/O)$  y  $\log(U)$  que se ha usado para definir el conjunto de modelos utilizando otro tipo de relación entre  $12 + \log(O/H)$  y  $\log(U)$ . Además se muestra la distribución de los modelos en distintos diagramas. Los diagramas mostrados, líneas de demarcación y símbolos son análogos a los mostrados en la Fig. 5.6.

utilizados en este trabajo preliminar pueden ser no óptimos. Sin embargo, el estudio de las ANNs es un área en constante cambio actualmente. Pero aun así, hemos mostrado que el uso de las ANN es completamente viable.

El método propuesto para encontrar los parámetros de las relaciones entre  $12 + \log(\text{O}/\text{H})$ ,  $\log(\text{N}/\text{O})$  y  $\log(\text{U})$  presenta una versatilidad bastante grande. En el resultado presentado en la sección anterior, se ha limitado la forma funcional de las relaciones de las propiedades físicas. Esto tiene fundamentos astrofísicos y está acorde a los resultados obtenidos por otros autores (p. ej. [Morisset et al., 2016](#); [Vila-Costas & Edmunds, 1993](#); [Pérez-Montero et al., 2016](#)). Sin embargo, las formas funcionales usadas pueden ser no las correctas para estas relaciones. Por ejemplo, para la relación entre  $12 + \log(\text{O}/\text{H})$  y  $\log(\text{U})$  diversos autores han encontrado un gradiente positivo y que puede depender de muchos factores (p. ej. [Espinosa-Ponce et al., 2022](#); [Ji & Yan, 2022](#)), por lo que podría ser conveniente no imponer una forma funcional que limite el resultado obtenido. Nuestra metodología permite usar cualquier forma funcional definida por el usuario. Aunque esto podría derivar en resultados inesperados, dado que el programa siempre tratará de encontrar el mejor resultado. Esto requiere una mayor exploración de este método ya que es bastante complejo. En pruebas preliminares hemos encontrado que el espacio de parámetros es bastante degenerado donde no todas las soluciones tienen sentido astrofísico. Esto es bien conocido, sin embargo nuestro método nos proporciona una manera rápida de poder explorar todo el espacio de parámetros (definido por la muestra de modelos inicial). Debido a esto, se necesita explorar y restringir ciertos aspectos en nuestro método.

Teniendo en cuenta lo anteriormente dicho, debemos remarcar que aún así los modelos que encontramos caen en la zona de formación estelar de los diagramas de diagnóstico. Además, en nuestra primera aproximación, encontramos que nuestros modelos reproducen bastante bien la distribución de datos observados. Cabe señalar que en algunos diagramas usados que involucran líneas débiles, como  $[\text{O III}]\lambda 4363$ , se tienen menos datos observacionales. Esto podría indicar que también estamos limitados por nuestra muestra observacional, por lo que una muestra más completa podría mejorar nuestros resultados. Por otro lado, para los modelos que hemos definido parece no existir una gran diferencia en las diferentes distribuciones de densidad. Hay que remarcar que este es un resultado preliminar, aunque en algunos trabajos (p. ej. [Morisset et al., 2016](#)) no se han encontrado diferencias significativas entre los resultados derivados utilizando diferentes geometrías en las distribuciones de los datos observacionales.

La edad de las poblaciones ionizantes presenta un efecto significativo en los modelos obtenidos. Recordemos que las regiones H II son ionizadas por poblaciones jóvenes masivas y recién formadas y solo viven durante algunos Myr. Obtenemos que conforme aumenta la edad de la población ionizante, los cocientes de línea cambian. Este es un efecto esperado, pero podría indicar que debe de explorarse el ajuste de las relaciones restringiendo aun más las edades de las poblaciones ionizantes para tener información de su efecto en conjunto sobre las observables.

Por último, los modelos con los cuales se ha entrenado a la red neuronal cubren una parte importante del espacio de parámetros. Sin embargo, algunas propiedades no han sido tomadas en cuenta para este estudio. Por ejemplo, la depleción de metales puede tener un efecto significativo en las observables. En un futuro, se planea estudiar el efecto de la depleción en nuestros resultados. Por otra parte, hemos supuesto que la población ionizante que se encuentra al centro es el resultado de un brote estelar. En nuestro caso se obtiene por el código POPSTAR. Sin embargo, hay otros códigos que se desarrollan basados en distintas hipótesis, como la presencia de estrellas super masivas, o rotando, o en conjuntos binarios (p. ej. [Wofford et al., 2016](#); [Xiao et al., 2018](#); [D'Agostino et al., 2019](#)). Estas suposiciones pueden afectar en las observables obtenidos. Todos estos efectos deberán tenerse en cuenta en estudios

posteriores.



## Capítulo 6

# Discusión, conclusiones y trabajo a futuro

En esta sección se incluye resumen de la discusión y conclusiones incluidas en los artículos publicados hasta la fecha<sup>1</sup>. Igualmente se presenta una discusión breve del resultado del estudio realizado con los modelos de fotoionización y redes neuronales. A lo largo de esta tesis hemos visto cómo los estudios de regiones H II se han beneficiado enormemente con las grandes muestras observacionales de galaxias usando la técnica de espectroscopía de campo integral. El acceso a datos espacialmente resueltos nos permite estudiar las propiedades más importantes de las nebulosas en distintas posiciones dentro de las galaxias. Además, esto nos ha proporcionado la herramienta necesaria para obtener información de la dependencia entre sus propiedades físicas y el lugar donde se formaron estas nebulosas.

El análisis de las distintas componentes del espectro también nos ha proporcionado información que hemos utilizado ampliamente a lo largo de este trabajo. Por una parte, el correcto modelado de la componente estelar de cada espectro nos permite obtener la emisión del gas ionizado sin alguna contribución estelar, la cual ha sido eliminada de la forma óptima. Esto nos permite estudiar directamente las líneas de emisión provenientes del gas ionizado y explorar sus propiedades, tales como composición química, densidad, parámetro de ionización, etc. Por otro lado, la componente estelar es importante para el estudio presentado a lo largo de esta tesis, dado que el gas ionizado y sus propiedades físicas tienen una conexión directa con las poblaciones estelares subyacentes, es decir, de la componente estelar modelada.

En este trabajo se exploraron las propiedades físicas empíricamente de una muestra de un nuevo catálogo de regiones H II generado a partir de los datos del proyecto CALIFA. Este catálogo consta con la información de  $\sim 28,000$  objetos individuales extraídos de una muestra de  $\sim 1000$  galaxias. Esta muestra de galaxias es estadísticamente significativa y representativa del universo local. Incluye todos los tipos morfológicos cubriendo un rango de masas de desde  $\approx 10^7 M_{\odot}$  hasta  $\approx 10^{13} M_{\odot}$ . Por lo que el catálogo no presenta un sesgo con respecto a la morfología, masa u otras propiedades globales de las galaxias. Así mismo cuenta con la información espectroscópica de las líneas de emisión más importantes en el rango óptico (entre 3745 y 7200Å).

La identificación de las regiones H II se realizó utilizando solamente la definición de estas nebulosas: i) estructuras grumosas con una alta emisión en H $\alpha$  y ii) una población estelar joven asociada a la nebulosa. La selección nos permite evitar sesgos asociados a las líneas de emisión que se producen en las regiones ionizadas. Esto es importante dado que estos sesgos

---

<sup>1</sup>Espinosa-Ponce et al. (2020), Sánchez et al. (2021b) y Espinosa-Ponce et al. (2022)



pueden derivar en propiedades y condiciones de ionización no representativos de este tipo de nebulosas. Para identificar la población joven asociada a las regiones H II se ha usado la anchura equivalente de H $\alpha$ ,  $EW(H\alpha) > 6\text{\AA}$  así como la fracción de poblaciones jóvenes,  $f_y$  de las poblaciones estelares subyacentes. Se ha mostrado que  $EW(H\alpha)$  depende de la fuente ionizante (Lacerda et al., 2018), siendo este más alto para poblaciones jóvenes. Sin embargo, este parámetro, junto a la fracción estelar joven, también pueden ser de ayuda para identificar la contribución del gas ionizado difuso (*diffuse ionized gas*, DIG por sus siglas en inglés).

El DIG puede ser ionizado principalmente por fotones que escapan de las regiones H II (*leaking*) y por estrellas viejas calientes por HOLMES. Su naturaleza se puede identificar usando  $EW(H\alpha)$  y  $f_y$ , pero esto dependerá de la calidad de los datos. Para los datos de CALIFA, las alas de la PSF limita nuestra habilidad de segregar entre diferentes fuentes de ionización. Esto es debido a que el límite de las regiones H II no está bien definido por la resolución de los datos. Esto provoca que las propiedades físicas, como  $EW(H\alpha)$  y las correspondientes líneas de emisión del DIG ionizado por *leaking* de fotones sean muy parecidas de las regiones H II adyacentes. Por lo tanto, no es posible con estos datos identificar este tipo de DIG. Por otro lado, el DIG ionizado por HOLMES puede ser identificado con cierta facilidad. La contribución de este gas difuso es mayor en zonas donde no se encuentran regiones H II, por lo que la baja resolución de los datos resulta un problema menor. Podemos caracterizar y calcular la contribución de este DIG que contamina las líneas de emisión de las regiones H II. A pesar de que solo hemos identificado la contaminación de un tipo de DIG, podemos suponer que esta contribución es la que produce los efectos mas fuertes en los cocientes de línea (Vale Asari et al., 2019).

Una mejor resolución espacial puede ser de ayuda para la segregación de regiones ionizadas y una mejor comparación de las contribuciones de todos los tipos de DIG. Como se ha visto en los datos de CALIFA, su resolución (con  $FWHM \sim 2.5''$ ) provoca que muchas de las regiones ionizadas pequeñas no son identificadas. Con datos de mejor resolución, como MUSE (López-Cobá et al., 2020) o SITELLE (Rousseau-Nepton et al., 2018), se podría mejorar la detección de regiones ionizadas pequeñas así como una segregación del DIG (y sus diferentes fuentes) más adecuada. Sin embargo, la mayor resolución provoca que el algoritmo de identificación de regiones ionizadas tenga que ser re-planteado (p ej. Lugo-Aranda et al., 2022).

Hay diversos trabajos que muestran las ventajas de una mejor resolución espacial de los datos. Por ejemplo, los datos de MUSE cuyo  $FWHM \sim 0.8''$  (limitado solo por el *seeing*) fueron usados por Sánchez-Menguiano et al. (2018) para estudiar los perfiles radiales de abundancia de oxígeno. Para esto obtuvieron un catálogo de regiones H II cuyo número de objetos (14 345) representa el  $\sim 65\%$  en número de objetos de nuestro catálogo usando solamente una muestra de galaxias que, en número, representa solo el  $\sim 10\%$  de la muestra que usamos en esta tesis. Por otro lado, Rousseau-Nepton et al. (2018) usaron los datos de SITELLE que cuenta con una resolución limitada por el *seeing*. Identificaron 4285 regiones H II en una sola galaxia (NGC 628). Esta galaxia está incluida en nuestra muestra de galaxias. En la misma nosotros solamente encontramos 243 regiones H II.<sup>2</sup> A pesar de que el método de segregación usado para los datos de SITELLE puede identificar regiones ionizadas no reales, el uso de estos datos supone una mejora en la detección de estas nebulosas.

Estudios más recientes que son parte del proyecto PHANGS (por las siglas de Physics at

<sup>2</sup>Esta galaxia fue analizada por Sánchez et al. (2012b) usando datos de PINGS, los cuales tienen una resolución similar a los datos de CALIFA. Encontraron 373 regiones H II donde 282 eran objetos con un espectro de buena calidad. Se ha utilizado en principio el mismo algoritmo en la identificación de las regiones ionizadas, sin embargo con diferentes parámetros. Esto causa la diferencia entre el número de objetos identificados, siendo los dos valores bastantes similares.

High Angular Resolution in Nearby Galaxies [Emsellem et al., 2022](#)) usan observaciones de MUSE complementándolas con observación de CO (de ALMA: *Atacama Large Millimeter/submillimeter Array*). Esta muestra consta de 19 galaxias espirales. En [Belfiore et al. \(2022\)](#) usan esta muestra para estudiar el DIG y su origen. Dada las características de la muestra, pueden estudiar ambos tipos de DIG asumiendo un modelo para el gas difuso ionizado por fotones que escapan de las regiones H II. En [Barnes et al. \(2022\)](#) usan estos datos para estudiar la relación entre las poblaciones estelares y las regiones H II en la galaxia NGC 1672 donde han identificado cerca de 1581 nebulosas. Por último, [Santoro et al. \(2022\)](#) presentó el catálogo de regiones H II con  $\sim 23,000$  objetos con una resolución de 67 pc a partir de una muestra de 19 galaxias. Estos tres trabajos recientes son el claro ejemplo donde podemos ver la ventaja de datos de mejor resolución espacial.

A pesar de la mejora que supone el uso de estos datos, también presentan desventajas. Por ejemplo, los datos de MUSE no incluyen la parte azul del espectro (solo hasta 4650Å). Esto excluye varias líneas de emisión para objetos a bajo corrimiento al rojo y también perdemos la información de las poblaciones estelares en esta parte del espectro. Mientras la técnica usada en SITELE afecta directamente las propiedades del continuo que pueden ser observadas haciendo que la información de las poblaciones estelares que son recuperadas sea menos confiable.

Con el catálogo de regiones H II generado se estudió la relación directa entre la localización en el diagrama BPT y otros diagramas de diagnóstico y sus propiedades. Tenemos que remarcar que no hemos usado las líneas de emisión de las nebulosas para su identificación. Sin embargo, encontramos que la distribución de las regiones H II se encuentra en la zona de formación estelar de los diagramas de diagnóstico. Solo el 4% de las regiones está por encima de la línea de demarcación propuesta por [Kewley et al. \(2001\)](#), la cual está basada en modelos de fotoionización. Por lo que nuestra selección reproduce lo esperado por los modelos de fotoionización sin hacer ninguna suposición sobre las líneas de emisión.

Se encontró que la posición en los diagramas de diagnóstico, y por lo tanto, de los valores de los cocientes de línea involucrados en estos diagramas, depende de las propiedades físicas y las propiedades de las poblaciones estelares subyacentes. Para diversos parámetros físicos se tiene una tendencia bien definida a lo largo de estos diagramas (p ej para  $EW(H\alpha)$ ,  $O/H$ ,  $N/O$ ,  $U$ , etc), por lo que la posición en estos diagramas dependen de los parámetros que definen la ionización del gas. En el diagrama BPT, las regiones H II en la zona superior izquierda son de baja metalicidad con un alto parámetro de ionización, con una extinción por polvo menor que las regiones en la zona derecha inferior del diagrama. Se han encontrado tendencias similares para muestras con un menor número de regiones y/o galaxias (p ej [Evans & Dopita, 1985](#); [Dopita & Evans, 1986](#); [Bresolin et al., 2012](#); [Sánchez et al., 2015b](#)). Debemos remarcar que algunos calibradores de líneas fuertes pueden inducir algunas de estas tendencias, por ejemplo los cocientes de líneas como  $O3N2$  ([Alloin et al., 1979](#); [Stasińska et al., 2006](#); [Marino et al., 2013](#)) y  $N2$  ([Pettini & Pagel, 2004](#); [Marino et al., 2013](#)) que se usan principalmente para estimar la abundancia de oxígeno.

También se encontró que la posición en estos diagramas de diagnóstico también depende del tipo morfológico de la galaxia y la distancia galactocéntrica. Estas tendencias están relacionadas con la evolución química en las galaxias. La relación con el tipo morfológico se esperaba dada las diferentes propiedades que presentan los diferentes tipos morfológicos que influyen directamente en el gas ionizado. La dependencia con la distancia está relacionado con el escenario dentro-fuera de formación de galaxias. Al centro de las galaxias se tiene una población estelar más vieja y metálica que las estrellas en las regiones exteriores de las mismas. Por tanto existe una dependencia de la historia de formación estelar a diferentes distancias,

lo cual provoca que se encuentre un gradiente de metalicidad del ISM en las galaxias.

Incluso las propiedades físicas de las regiones H II pueden presentar relaciones entre ellas mismas. Se encontraron tendencias claras entre O/H, EW(H $\alpha$ ) y  $f_y$ . Estos resultados se esperaban dado que las regiones con alto O/H son localizadas en áreas con una alta densidad de masa estelar (basado en la relación MZR resuelta [Rosales-Ortega et al., 2012](#); [Barrera-Ballesteros et al., 2016](#)), es decir, son regiones con una alta emisión de continuo. Por tanto esto corresponde a lugares donde se tiene valores bajos de EW(H $\alpha$ ) y  $f_y$  para regiones H II que son ionizadas por el mismo tipo de cúmulos estelares, a la resolución de los datos explorados.

Las relaciones de O/H con log(U) y N/O no se espera que estén conectadas con las propiedades y la evolución de las poblaciones estelares en las zonas en donde las regiones H II se forman. La relación entre  $12 + \log(\text{O}/\text{H})$  y  $\log(\text{N}/\text{O})$  es el resultado de la producción primaria del oxígeno y de la producción primaria y secundaria del nitrógeno. Podemos ver como la relación entre estos dos parámetros es constante cuando el nitrógeno se está produciendo por su mecanismo primario. Por lo tanto, esto quiere decir que en el lugar de formación de la región H II aún no se ha alcanzado la abundancia de oxígeno necesaria para que el nitrógeno se empiece a producir con su mecanismo secundario. Recordemos que esta producción depende directamente del oxígeno producido anteriormente. Para abundancias bajas de oxígeno, el cociente de N/O tendrá un valor constante debido a la producción primaria del nitrógeno. Por tanto, en la zona donde la región H II se ha formado no se ha alcanzado la abundancia de oxígeno necesaria para que se inicie la producción secundaria de nitrógeno. Esto refuerza el gradiente radial para N/O, ya que este dependerá del lugar, el tipo morfológico y la masa de la galaxia donde se ha formado la región H II ([Belfiore et al., 2017](#)). La relación entre  $12 + \log(\text{O}/\text{H})$  y  $\log(\text{U})$  no se entiende por completo ([Ji & Yan, 2022](#)). Para explicar las diferentes causas de esta relación se requiere que exista una conexión entre esta relación y el lugar donde la región H II se ha formado ([Evans & Dopita, 1985](#); [Dopita & Evans, 1986](#)). Esto puede deberse a: i) un sesgo del ambiente en el cual se forman las nebulosas, ii) un cambio en la función inicial de masa y su conexión con la abundancia de oxígeno y iii) un efecto del polvo en la dureza del espectro ionizante.

En la primera posibilidad se propone que en el lugar en donde se formaron las regiones H II existe una nube de gas molecular con una densidad alta, cuyo brillo superficial es menor en el centro ([Dopita & Evans, 1986](#)). Esto puede deberse a que en la nube molecular con alta metalicidad se fragmenta en nubes de menor tamaño y colapsen para formar regiones H II pequeñas de bajo brillo superficial, es decir, es un efecto directo del enfriamiento y fragmentación debido a la metalicidad. Este tipo de gas molecular se puede encontrar en las regiones centrales de las galaxias, donde se puede observar la relación entre log(U)-OH. Por otro lado, la presión puede jugar un papel en esta relación además de la metalicidad. Una alta presión puede favorecer la formación de regiones H II a partir de las nubes moleculares. En diversos estudios se ha visto que la presión decrece con el radio de las galaxias ([Barnes et al., 2021](#); [Barrera-Ballesteros et al., 2021](#)).

Por otro lado, un cambio en la IMF puede tener diversas explicaciones, por ejemplo puede existir una variación en la IMF debido a las distintas componentes de una galaxia ([Chabrier, 2003](#)), aunque este escenario se ha mostrado que no es posible ([Wegg et al., 2017](#)). Otra de las posibles causas es que exista una dependencia con el tiempo de evolución de una galaxia ([Vazdekis et al., 1996](#)). También puede existir una relación entre la forma funcional de la IMF y otras propiedades de las galaxias, como la dispersión de velocidades y la metalicidad estelar ([Martín-Navarro et al., 2015](#); [McDermid et al., 2015](#)). Por tanto la relación entre log(U)-O/H podría estar relacionada con el proceso de enriquecimiento, como la metalicidad aumenta el enfriamiento, fragmentación y colapso de las nubes moleculares. Finalmente el efecto del

polvo o el opacamiento metálico puede tener un origen similar, esto se puede ver en como las regiones con una alta metalicidad son las que sufren una extinción por polvo mayor.

Es posible que la relación entre  $\log(U)$  y  $O/H$  dependa de todos los escenarios que se han mencionado simultáneamente a diferente nivel cada uno. Esta puede ser la razón por la que esta relación sea menos *fuerte* que las relaciones entre otras propiedades. También puede ser la razón en la que en algunos estudios, se encuentre una anticorrelación entre estos dos parámetros (p. ej. [Ji & Yan, 2022](#)).

Por otra parte, estas relaciones son muy importantes cuando se realizan modelos de fotoionización. Los modelos de fotoionización necesitan que se hagan ciertas suposiciones entre las propiedades físicas, como  $N/O-O/H$  y  $\log(U)-O/H$ , de las nebulosas para reproducir los cocientes de líneas debido a la ionización por estrellas jóvenes. Estas relaciones son puramente empíricas, y en caso de no tomarse en cuenta puede llevar a mal interpretaciones de los resultados. Esto se debe a que la relación entre las propiedades físicas y los cocientes de línea observados está altamente degenerada (como es el caso del cociente  $R23^3$ , [Pagel et al., 1979](#)).

La extinción por polvo no presenta una clara relación con la abundancia de O. Además, presenta una dispersión bastante grande en comparación con las otras relaciones encontradas. Sin embargo, si tomamos en cuenta los gradientes radiales de  $A_v$  y  $O/H$ , podemos ver que hay tendencias radiales similares para galaxias de la misma masa y tipo morfológico, lo que podría derivar en una relación entre ambos parámetros. Si tomamos en cuenta la relación entre la extinción por polvo y la densidad de la columna de gas ([Schmidt, 1959](#); [Barrera-Ballesteros et al., 2020](#)), se puede interpretar que las regiones con una alta metalicidad son aquellas donde la densidad superficial de las nubes moleculares es alta, lo cual estaría en contra de una de las posibles explicación para la relación entre  $\log(U)$  y  $O/H$  discutida anteriormente. Esto resalta la fuerte conexión entre las distintas propiedades de una región H II y las propiedades locales de las zonas en donde las nebulosas se formán. También esto puede indicar que hay jerarquía entre las diferentes relaciones exploradas. La relación entre  $O/H$  y  $N/O$  parece ser independiente de las propiedades externas a la nebulosa, por lo que es una relación *universal*. Otras relaciones, como con  $EW(H\alpha)$ ,  $f_y$  y  $A_v$  parecen estar relacionadas con el lugar en la galaxia donde las regiones H II se forman y, por lo tanto, dependen de la evolución y enriquecimiento químico previo del ISM. Otros parámetros, como la  $n_e$ , parece no estar conectada con la abundancia de O. Es decir, es totalmente independiente a las conexiones descritas con la evolución de las poblaciones estelares o del lugar donde se forman las regiones H II.

Las tendencias encontradas a lo largo de los diagramas de diagnóstico son una consecuencia directa de: i) la evolución de las poblaciones estelares espacialmente resueltas, ii) el enriquecimiento químico en el lugar donde las regiones H II se formaron, iii) de los diferentes ambientes donde las regiones H II se forman. Es por esto que se pueden observar gradientes bien definidos de las propiedades físicas, como  $O/H$  y  $N/O$ , que han sido ampliamente estudiados en la literatura (p ej [Peimbert et al., 1978](#); [Sánchez et al., 2014](#); [Belfiore et al., 2017](#); [Esteban & García-Rojas, 2018](#)). Estos gradientes también dependen en las propiedades globales de las galaxias (masa y tipo morfológico). Esto se debe a la relación directa entre las líneas de emisión de las nebulosas y las propiedades estelares subyacentes. Ya que la historia de formación estelar y enriquecimiento químico afecta directamente a las propiedades físicas de las regiones H II ([Sánchez et al., 2015b](#); [Espinosa-Ponce et al., 2020](#)).

Dado que se ha visto que la posición en el diagrama BPT depende de las propiedades globales de las galaxias, podemos ver como los cocientes involucrados:  $[N II]\lambda 6583/H\alpha$  y  $[O III]\lambda 5007/H\beta$ , se relacionan con las propiedades estelares, edad ( $\mathcal{A}_{*,L}$ ) y metalicidad ( $\mathcal{Z}_{*,L}$ ).

---

<sup>3</sup> $([O III] \lambda\lambda 4959, 5007 + [O II] \lambda 3727)/H\beta$

El  $[\text{N II}]\lambda 6583/\text{H}\alpha$  presenta una fuerte relación con estas propiedades dado que se relaciona con la composición química (abundancia de oxígeno) la cual es el resultado del enriquecimiento de metales producidos por las estrellas. Esto debido a que las generaciones anteriores de las estrellas han enriquecido la composición química del medio. Por lo tanto, la abundancia de oxígeno está relacionada con la metalicidad de las estrellas jóvenes (Sánchez et al., 2015a; Sánchez-Menguiano et al., 2016). Mientras que la relación con la edad de las poblaciones estelares puede deberse al proceso de enriquecimiento químico en las galaxias, es decir, por las diferentes poblaciones estelares que existen a lo largo de las galaxias (Sánchez-Blázquez et al., 2014; García-Benito et al., 2017). Por otro lado, el cociente  $[\text{O III}]\lambda 5007/\text{H}\beta$  presenta una relación más débil que el cociente de  $[\text{N II}]$ . Esto es debido a que este cociente depende más de la temperatura y densidad del gas así como de la geometría del mismo y no es tan sensible a cambios en la abundancia de O.

Todas las relaciones encontradas entre las distintas propiedades físicas de las regiones H II con la abundancia de O apuntan a que O/H tiene una conexión directa con las propiedades de las poblaciones estelares subyacentes. Hay diversos trabajos donde se propone que O/H está relacionado con propiedades locales tales como la densidad superficial de masa estelar (Rosales-Ortega et al., 2012; Barrera-Ballesteros et al., 2016) y la fracción de gas (Barrera-Ballesteros et al., 2018). Se encontró que O/H presenta una clara relación con las propiedades de las poblaciones subyacentes: con su edad,  $\mathcal{A}_{*,L}$ , y con su metalicidad  $\mathcal{Z}_{*,L}$ . Este resultado era de esperar dado que se encontró la misma dependencia entre el cociente  $[\text{N II}]\lambda 6583/\text{H}\alpha$  y las mismas propiedades. La relación con la  $\mathcal{A}_{*,L}$  indica que las regiones H II más metálicas se forman en zonas de la galaxia donde la población estelar se formó hace mucho tiempo. Por el otro lado, las regiones menos metálicas se han formado en zonas donde la población estelar se ha formado a lo largo de un rango largo de tiempo, y por lo tanto, tienen un rango amplio en sus edades.

Esto nos indica que debe existir una relación entre la abundancia de oxígeno y la metalicidad de las poblaciones estelares, ya que está conectada con el enriquecimiento químico local. Se encontró que existe esta relación, aunque es menos fuerte que la relación con  $\mathcal{A}_{*,L}$ . Esto puede deberse a que  $\mathcal{Z}_{*,L}$  es el promedio de todas las metalicidades de todas las poblaciones estelares en donde el espectro observado es analizado. Por tanto, por construcción, esta traza el Fe en las estrellas. Por otro lado, el O/H traza la abundancia de un elemento  $\alpha$  en el ISM. En conclusión, esta abundancia debería ser similar a la de las estrellas jóvenes recién formadas que ionizan el gas (González Delgado et al., 2015). Entonces, ambos parámetros trazan diferentes elementos a diferentes escalas de tiempo y para diferentes familias de estrellas. Dando como resultado que  $\mathcal{Z}_{*,L}$  y O/H presenten importantes diferencias en las historias de enriquecimiento química correspondientes (p ej. Maiolino & Mannucci, 2019). Lo débil de esta relación parece indicar que ambos parámetros están conectados pero sin depender totalmente uno del otro. Esto se debe a que  $\mathcal{Z}_{*,L}$ , depende tanto de la masa estelar final y de la forma de la historia de formación estelar (Camps-Fariña et al., 2021). Por otro lado, O/H depende más fuertemente a la de la masa final estelar (Lacerda et al., 2020) y de la densidad superficial de masa estelar (p ej González Delgado et al., 2016; Sánchez, 2020).

Dado que O/H y  $\mathcal{Z}_{*,L}$  están relacionados, se exploró el cociente  $[\text{O}/\text{Z}]$  el cual es la resta entre  $\mathcal{Z}_{*,L}$  y O/H normalizado al valor solar (Asplund et al., 2009). Debemos hacer notar que  $\mathcal{Z}_{*,L}$  es un trazador directo de la abundancia de hierro, estando conectado con  $[\text{Fe}/\text{H}]$  esto es, el cociente normalizado al valor solar de Fe (Asplund et al., 2009). Por lo que con  $[\text{O}/\text{Z}]$  y  $[\text{Fe}/\text{H}]$  puede obtenerse el cociente  $[\text{O}/\text{Fe}]$ , el cual presenta una tendencia similar cuando se compara con  $[\text{Fe}/\text{H}]$  a la que presenta  $[\alpha/\text{Fe}]$  descrita para i) las estrellas en la Vía Láctea (p ej Yates et al., 2013; Hayden et al., 2015; Griffith et al., 2021), ii) en galaxias de tipo temprano



(Walcher et al., 2014, 2015) y iii) en simulaciones de galaxias tipo Vía Láctea (Calura & Menci, 2009; Yates et al., 2012). Debemos de tener en cuenta que  $[\text{O}/\text{Fe}]$  no debe ser interpretado como el exceso de elementos  $\alpha$  (por ej. el oxígeno) respecto a elementos no  $\alpha$  (por ejemplo el hierro)<sup>4</sup> para una población estelar promedio o para un conjunto particular de poblaciones estelares en el lugar donde se formó la región H II. Sin embargo, preserva cierta información sobre el proceso de exceso de elementos  $\alpha$ . Pero debemos recordar que este cociente son las propiedades promedio de dos poblaciones totalmente diferentes: el gas y las estrellas. Por otra parte, el cociente presenta una dependencia con las propiedades de la galaxia, como su masa estelar y su morfología. Su valor depende de cada uno de los tipos morfológicos siendo más grande para los tipos tempranos que para el encontrado para tipos tardíos de galaxias. Al mismo tiempo, presenta una fuerte dependencia con la  $\mathcal{A}_{\star,L}$ . Podemos ver como la relación  $[\text{O}/\text{Fe}]$  y  $[\text{Fe}/\text{H}]$  está mejor definido para un valor fijo de  $\mathcal{A}_{\star,L}$  y la morfología. Además, la dependencia entre los cambios de la distribución y la morfología es una consecuencia de la dependencia entre  $[\text{O}/\text{Fe}]$  con la  $\mathcal{A}_{\star,L}$ . Esto se confirma mediante modelos de evolución química simples donde el tiempo de retraso en la formación estelar ( $\tau_{\star}$ ) y el tiempo de depleción ( $\tau_{\text{dep}}$ ) dependen tanto de la masa como de la morfología. Por tanto parece que existe una  $[\text{O}/\text{Fe}]$  tiene relación con la historia de formación estelar o con la IMF.

Finalmente, a lo largo de esta tesis hemos confirmado que los modelos de fotoionización son una herramienta fundamental para estudiar la relación entre las propiedades físicas y las observables del gas ionizado. Sin embargo, existen ciertos resultados que solo se pueden obtener empíricamente como ya se ha mencionado. Una correcta exploración del espacio parámetros de las propiedades físicas de las regiones H II supone realizar un gran número de modelos, aunque esto está delimitado por el tiempo de computo de los modelos. Por lo que usualmente, se proponen modelos que muestreen el espacio de parámetros de manera representativa aunque limitada.

El desarrollo de nuevas técnicas computacionales nos han permitido disminuir y optimizar el tiempo de computo de ciertos modelos. Sin embargo, las técnicas de aprendizaje automatizado nos brindan la oportunidad de sustituir los códigos de fotoionización convencionales con alternativas tales como las redes neuronales. Se debe entender que no es reemplazo de una herramienta por otra, las redes neuronales tomarán el conjunto de modelos producidos por los códigos de fotoionización y actuarán como regresores. Esto significa que una ANN aprenderá como es la relación entre los parámetros de entrada y de salida de los modelos. Esto permitirá que se puedan generar nuevos resultados a partir de otras entradas sin tener que correr el código de fotoionización, lo cual supone una ventaja ya que el tiempo de cómputo que toma este proceso es muy pequeño comparado con el tiempo que toma correr un modelo de fotoionización.

Sin embargo, la red neuronal y su respectivo entrenamiento debe de hacerse con cuidado. El conjunto de modelos con la cual se entrena debe de representar todo el espacio de parámetros que uno desea estudiar. Ya que los resultado producidos por la ANN sólo serán válidos dentro de los límites definidos por la misma muestra de modelos con la cual se entrenó. Hemos mostrado que los resultados de la ANN entrenada con una muestra representativa del espacio de parámetros son congruentes con la salida directa de modelos de fotoionización. Más aún, hemos producido modelos a partir de las relaciones de  $12 + \log(\text{O}/\text{H})$ ,  $\log(\text{N}/\text{O})$  y  $\log(\text{U})$  lo cual nos permite poder estudiar el efecto de estas propiedades sobre los observables, es decir, líneas de emisión. Comparándola con una muestra observacional, podemos encontrar que relaciones ajustan mejor los datos observacionales. Este problema podría suponer un gran

<sup>4</sup>En inglés se le conoce como  $\alpha$ -enhancement

número de modelos de fotoionización que se necesiten calcular, sin embargo, con el uso de una ANN como regresora no es necesario todo el cálculo de modelos. Esto nos permite explorar un sin fin de posibilidades y parámetros para conciliar los resultados de modelos con los datos observacionales.

Algunas de las exploraciones realizadas a lo largo de esta tesis ha abierto nuevas posibilidades de estudios a futuro. Por ejemplo, la exploración del parámetro  $[O/Fe]$  es un estudio que se encuentra actualmente en proceso. Unido a los principales resultados presentados en [Sánchez et al. \(2021b\)](#), se necesita explorar las tendencias galaxia a galaxia y en diferentes regiones dentro de ella. Además se necesita explorar con modelos de evolución química para estudiar más a fondo la dependencia entre el cociente  $[O/Fe]$  y los parámetros de los modelos. Por otro lado, el estudio presentado usando modelos de fotoionización y redes neuronales está en fase exploratoria. El método descrito en el Cap. 5 produce modelos que nos proporcionan una aproximación a la distribución de datos observacionales. Sin embargo, hay propiedades en los modelos que no se han explorado, por ejemplo la depleción de metales y los efectos de usar otras descripciones de los espectros estelares ionizantes. Además las técnicas de *machine learning* siguen evolucionando y se necesita explorar sus parámetros para garantizar su correcto funcionamiento.

Recapitulando, se ha encontrado como las regiones H II y sus propiedades físicas dependen de las condiciones donde se han formado. Los procesos de evolución química y de las estrellas tienen un impacto en las regiones H II que podemos apreciar en los diagramas de diagnóstico. Esto está conectado directamente con las propiedades globales de las galaxias. A pesar de que estos resultados ya se han reportado antes, esta es la primera vez que se obtienen con una muestra tan grande de regiones H II identificadas de una muestra estadísticamente completa de galaxias para el universo local. Además, se ha podido identificar el DIG en una primera aproximación. Sin embargo, estamos limitados por la propia resolución de nuestros datos. Para las regiones H II esto significa que perdamos un número considerable de nebulosas pequeñas. Por otro lado, con una mejor resolución podríamos identificar ambas componentes del DIG, pudiendo obtener una mejor descontaminación para nuestras regiones. Los modelos de fotoionización siguen siendo la única manera conocida hasta ahora de comprender como las propiedades físicas afectan a las observables de un gas ionizado. Aunque dependen de las relaciones encontradas empíricamente para poder reproducir nebulosas astrofísicamente reales. Con el uso de ANNs hemos podido disminuir el tiempo de computo de este tipo de estudios. Sin embargo, esto requiere una exploración mucho más extensa para tomar todas las variables posibles que puedan afectar en las líneas de emisión observadas. Como se ha mencionado anteriormente, el estudio de las regiones H II es de vital importancia para el entendimiento de la evolución química de las galaxias. Además, hemos visto que las poblaciones estelares están ligadas a las propiedades físicas de las regiones, lo que nos proporciona información sobre como afectan los procesos evolutivos en el gas ionizado. Igualmente, la exploración con modelos de fotoionización representa un avance en el entendimiento del efecto de las propiedades físicas con las observables. Con la ayuda de las ANNs podemos mejorar el tiempo de computo lo cual nos permite generar modelos mucho más rápido acordes al código CLOUDY. Permitiendo una exploración mucho más completa del espacio de parámetros físicos de las regiones H II.

# Bibliografía

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJS*, 182, 543
- Aggarwal, C. C. 2018, *Neural Networks and Deep Learning* (Springer)
- Alarie, A., & Morisset, C. 2019, *Rev. Mexicana Astron. Astrofis.*, 55, 377
- Aller, L. H., Ufford, C. W., & van Vleck, J. H. 1949, *ApJ*, 109, 42
- Allington-Smith, J. 2006, *New A Rev.*, 50, 244
- Allington-Smith, J. R., Content, R., & Haynes, R. 1998, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 3355, *Optical Astronomical Instrumentation*, ed. S. D’Odorico, 196–205
- Allington-Smith, J. R., Content, R., Haynes, R., & Lewis, I. J. 1997, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 2871, *Optical Telescopes of Today and Tomorrow*, ed. A. L. Ardeberg, 1284–1294
- Alloin, D., Collin-Souffrin, S., Joly, M., & Vigroux, L. 1979, *A&A*, 78, 200
- Amayo, A., Delgado-Inglada, G., & García-Rojas, J. 2020, *MNRAS*, 492, 950
- Amayo, A., Delgado-Inglada, G., & Stasińska, G. 2021, *MNRAS*, 505, 2361
- André, P., Men’shchikov, A., Bontemps, S., et al. 2010, *A&A*, 518, L102
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481
- Bacon, R., Adam, G., Baranne, A., et al. 1995, *A&AS*, 113, 347
- Bacon, R., Accardo, M., Adjali, L., et al. 2010, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 7735, *Ground-based and Airborne Instrumentation for Astronomy III*, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 773508
- Baldwin, J. A., Ferland, G. J., Martin, P. G., et al. 1991, *ApJ*, 374, 580
- Ballesteros-Paredes, J., André, P., Hennebelle, P., et al. 2020, *Space Sci. Rev.*, 216, 76
- Balmer, J. J. 1885, *Annalen der Physik*, 261, 80
- Barden, S. C., & Wade, R. A. 1988, in *Astronomical Society of the Pacific Conference Series*, Vol. 3, *Fiber Optics in Astronomy*, ed. S. C. Barden, 113–124
- Barnes, A. T., Glover, S. C. O., Kreckel, K., et al. 2021, *MNRAS*, 508, 5362

- Barnes, A. T., Chandar, R., Kreckel, K., et al. 2022, arXiv e-prints, arXiv:2205.05679
- Barrera-Ballesteros, J. K., Heckman, T. M., Zhu, G. B., et al. 2016, MNRAS, 463, 2513
- Barrera-Ballesteros, J. K., Heckman, T., Sánchez, S. F., et al. 2018, ApJ, 852, 74
- Barrera-Ballesteros, J. K., Utomo, D., Bolatto, A. D., et al. 2020, MNRAS, 492, 2651
- Barrera-Ballesteros, J. K., Sánchez, S. F., Heckman, T., et al. 2021, MNRAS, 503, 3643
- Belfiore, F., Maiolino, R., & Bothwell, M. 2016, MNRAS, 455, 1218
- Belfiore, F., Maiolino, R., Tremonti, C., et al. 2017, MNRAS, 469, 151
- Belfiore, F., Santoro, F., Groves, B., et al. 2022, A&A, 659, A26
- Binette, L., Wilson, A. S., & Storchi-Bergmann, T. 1996, A&A, 312, 365
- Blanc, G. A., Kewley, L., Vogt, F. P. A., & Dopita, M. A. 2015, ApJ, 798, 99
- Blitz, L. 1993, in *Protostars and Planets III*, ed. E. H. Levy & J. I. Lunine, 125
- Blitz, L., Fukui, Y., Kawamura, A., et al. 2007, in *Protostars and Planets V*, ed. B. Reipurth, D. Jewitt, & K. Keil, 81
- Boulanger, F., Cox, P., & Jones, A. P. 2000, in *Infrared Space Astronomy, Today and Tomorrow*, ed. F. Casoli, J. Lequeux, & F. David, Vol. 70, 251
- Bradley, T. R., Knapen, J. H., Beckman, J. E., & Folkes, S. L. 2006, A&A, 459, L13
- Bresolin, F. 2007, ApJ, 656, 186
- . 2011, ApJ, 730, 129
- . 2019, MNRAS, 488, 3826
- Bresolin, F., Kennicutt, Robert C., J., & Garnett, D. R. 1999, ApJ, 510, 104
- Bresolin, F., Kennicutt, R. C., & Ryan-Weber, E. 2012, ApJ, 750, 122
- Bresolin, F., Kudritzki, R.-P., Urbaneja, M. A., et al. 2016, ApJ, 830, 64
- Bresolin, F., Ryan-Weber, E., Kennicutt, R. C., & Goddard, Q. 2009, ApJ, 695, 580
- Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
- Buff, J., & McCray, R. 1974, ApJ, 189, 147
- Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7
- Burbidge, E. M., Burbidge, G. R., Fowler, W. A., & Hoyle, F. 1957, *Rev. Mod. Phys.*, 29, 547
- Busso, M., Gallino, R., & Wasserburg, G. J. 1999, *ARA&A*, 37, 239
- Butler, K. 1993, in *Planetary Nebulae*, ed. R. Weinberger & A. Acker, Vol. 155, 73
- Calura, F., & Menci, N. 2009, MNRAS, 400, 1347

- Calzetti, D. 2001, *PASP*, 113, 1449
- Camps-Fariña, A., Sanchez, S. F., Lacerda, E. A. D., et al. 2021, *MNRAS*, 504, 3478
- Cantó, J., Curiel, S., & Martínez-Gómez, E. 2009, *A&A*, 501, 1259
- Cantó, J., Raga, A., Steffen, W., & Shapiro, P. R. 1998, *ApJ*, 502, 695
- Cappellari, M., & Copin, Y. 2003, *MNRAS*, 342, 345
- Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, *MNRAS*, 413, 813
- Chabrier, G. 2003, *PASP*, 115, 763
- Charbonneau, P. 1995, *ApJS*, 101, 309
- Chiappini, C., Matteucci, F., & Romano, D. 2001, *ApJ*, 554, 1044
- Chiappini, C., Romano, D., & Matteucci, F. 2003, *MNRAS*, 339, 63
- Cid Fernandes, R., Pérez, E., García Benito, R., et al. 2013, *A&A*, 557, A86
- Cid Fernandes, R., González Delgado, R. M., García Benito, R., et al. 2014, *A&A*, 561, A130
- Coelho, P., Bruzual, G., Charlot, S., et al. 2007, *MNRAS*, 382, 498
- Collin, S., Dumont, A. M., & Godet, O. 2004, *A&A*, 419, 877
- Colombo, D., Kalinova, V., Utomo, D., et al. 2018, *MNRAS*, 475, 1791
- Cox, D. P. 2005, *ARA&A*, 43, 337
- Croom, S. M., Lawrence, J. S., Bland-Hawthorn, J., et al. 2012, *MNRAS*, 421, 872
- Curti, M., Mannucci, F., Cresci, G., & Maiolino, R. 2020, *MNRAS*, 491, 944
- da Silva, R. L., Fumagalli, M., & Krumholz, M. 2012, *ApJ*, 745, 145
- D'Agostino, J. J., Kewley, L. J., Groves, B., et al. 2019, *ApJ*, 878, 2
- de Zeeuw, P. T., Bureau, M., Emsellem, E., et al. 2002, *MNRAS*, 329, 513
- Del Zanna, G., Dere, K. P., Young, P. R., & Landi, E. 2021, *ApJ*, 909, 38
- Delgado-Inglada, G., Morisset, C., & Stasińska, G. 2014, *MNRAS*, 440, 536
- Dopita, M. A., & Evans, I. N. 1986, *ApJ*, 307, 431
- Dopita, M. A., Sutherland, R. S., Nicholls, D. C., Kewley, L. J., & Vogt, F. P. A. 2013, *ApJS*, 208, 10
- Dopita, M. A., Fischera, J., Sutherland, R. S., et al. 2006, *ApJ*, 647, 244
- Draine, B. T. 1978, *ApJS*, 36, 595
- . 2003, *ARA&A*, 41, 241
- . 2011, *Physics of the Interstellar and Intergalactic Medium* (Princeton University Press)



- Dufour, R. J., & Buckalew, B. 1999, in *Wolf-Rayet Phenomena in Massive Stars and Starburst Galaxies*, ed. K. A. van der Hucht, G. Koenigsberger, & P. R. J. Eenens, Vol. 193, 350
- Dumont, A. M., Abrassart, A., & Collin, S. 2000, *A&A*, 357, 823
- Eisenhauer, F., Abuter, R., Bickert, K., et al. 2003, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 4841, *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*, ed. M. Iye & A. F. M. Moorwood, 1548–1561
- Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017a, in *The Lives and Death-Throes of Massive Stars*, ed. J. J. Eldridge, J. C. Bray, L. A. S. McClelland, & L. Xiao, Vol. 329, 396–396
- Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017b, *PASA*, 34, e058
- Elmegreen, B. G. 1990, in *Astronomical Society of the Pacific Conference Series*, Vol. 12, *The Evolution of the Interstellar Medium*, ed. L. Blitz, 247–271
- Elmegreen, B. G., Elmegreen, D. M., Salzer, J. J., & Mann, H. 1996, *ApJ*, 467, 579
- Emsellem, E., Schinnerer, E., Santoro, F., et al. 2022, *A&A*, 659, A191
- Ercolano, B., Barlow, M. J., Storey, P. J., & Liu, X. W. 2003, *MNRAS*, 340, 1136
- Ercolano, B., Bastian, N., & Stasińska, G. 2007, *MNRAS*, 379, 945
- Espinosa-Ponce, C., Sánchez, S. F., Morisset, C., et al. 2020, *MNRAS*, 494, 1622
- . 2022, *MNRAS*, 512, 3436
- Esteban, C., & García-Rojas, J. 2018, *MNRAS*, 478, 2315
- Esteban, C., Peimbert, M., García-Rojas, J., et al. 2004, *MNRAS*, 355, 229
- Evans, I. N., & Dopita, M. A. 1985, *ApJS*, 58, 125
- Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, *Rev. Mexicana Astron. Astrofis.*, 49, 137
- Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, *Rev. Mexicana Astron. Astrofis.*, 53, 385
- Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, *ApJ*, 155, L149
- Flores-Fajardo, N., Morisset, C., Stasińska, G., & Binette, L. 2011, *MNRAS*, 415, 2182
- Franchetto, A., Mingozzi, M., Poggianti, B. M., et al. 2021, *ApJ*, 923, 28
- Gahm, G. F., Grenman, T., Fredriksson, S., & Kristen, H. 2007, *AJ*, 133, 1795
- Galbany, L., Anderson, J. P., Sánchez, S. F., et al. 2018, *ApJ*, 855, 107
- García-Benito, R., Zibetti, S., Sánchez, S. F., et al. 2015, *A&A*, 576, A135
- García-Benito, R., González Delgado, R. M., Pérez, E., et al. 2017, *A&A*, 608, A27

- García-Hernández, D. A., Ventura, P., Delgado-Inglada, G., et al. 2016, *MNRAS*, 458, L118
- García-Martínez, C., Rodríguez, F. J., & Lozano, M. 2018, *Genetic Algorithms*, ed. R. Martí, P. M. Pardalos, & M. G. C. Resende (Cham: Springer International Publishing), 431–464
- García Vargas, M. L., & Díaz, A. I. 1994, *ApJS*, 91, 553
- García-Vargas, M. L., González-Delgado, R. M., Pérez, E., et al. 1997, *ApJ*, 478, 112
- Garnett, D. R. 1990, *ApJ*, 363, 142
- . 1992, *AJ*, 103, 1330
- Gaunt, J. A. 1930, *Philosophical Transactions of the Royal Society of London Series A*, 229, 163
- Giammanco, C., Beckman, J. E., Zurita, A., & Relaño, M. 2004, *A&A*, 424, 877
- González Delgado, R. M., Pérez, E., Cid Fernandes, R., et al. 2014, *A&A*, 562, A47
- González Delgado, R. M., García-Benito, R., Pérez, E., et al. 2015, *A&A*, 581, A103
- González Delgado, R. M., Cid Fernandes, R., Pérez, E., et al. 2016, *A&A*, 590, A44
- Griffith, E., Weinberg, D. H., Johnson, J. A., et al. 2021, *ApJ*, 909, 77
- Gruenwald, R., Steigman, G., & Viegas, S. M. 2002, *ApJ*, 567, 931
- Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, *ApJ*, 808, 132
- Heiles, C., & Troland, T. H. 2003a, *ApJS*, 145, 329
- . 2003b, *ApJ*, 586, 1067
- Henry, R. B. C. 1993, *MNRAS*, 261, 306
- Henry, R. B. C., Edmunds, M. G., & Köppen, J. 2000, *ApJ*, 541, 660
- Heyer, M., & Dame, T. M. 2015, *ARA&A*, 53, 583
- Hill, T., Motte, F., Didelon, P., et al. 2011, *A&A*, 533, A94
- Ho, I. T. 2019, *MNRAS*, 485, 3569
- Ho, L. C., Filippenko, A. V., & Sargent, W. L. W. 1997, *ApJ*, 487, 579
- Hodge, P. W. 1987, *PASP*, 99, 915
- Husemann, B., Jahnke, K., Sánchez, S. F., et al. 2013, *A&A*, 549, A87
- Izotov, Y. I., Stasińska, G., Guseva, N. G., & Thuan, T. X. 2004, *A&A*, 415, 87
- Izotov, Y. I., & Thuan, T. X. 1999, *ApJ*, 511, 639
- Jamet, L., Stasińska, G., Pérez, E., González Delgado, R. M., & Vílchez, J. M. 2005, *A&A*, 444, 723

- Ji, X., & Yan, R. 2022, *A&A*, 659, A112
- Juan de Dios, L., & Rodríguez, M. 2017, *MNRAS*, 469, 1036
- Kallman, T., & Bautista, M. 2001, *ApJS*, 133, 221
- Karttunen, H., Kröger, P., Oja, H., Poutanen, M., & Donner, K. J. 2017, *Fundamental Astronomy* (Springer), doi:10.1007/978-3-662-53045-0
- Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, *MNRAS*, 346, 1055
- Kelz, A., Verheijen, M. A. W., Roth, M. M., et al. 2006, *PASP*, 118, 129
- Kennicutt, Robert C., J., Edgar, B. K., & Hodge, P. W. 1989, *ApJ*, 337, 761
- Kewley, L. J., & Dopita, M. A. 2002, *ApJS*, 142, 35
- Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, *ApJ*, 556, 121
- Kewley, L. J., & Ellison, S. L. 2008, *ApJ*, 681, 1183
- Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, *ARA&A*, 57, 511
- Kim, E. J., & Brunner, R. J. 2017, *MNRAS*, 464, 4463
- Kingdon, J. B., & Ferland, G. J. 1995, *ApJ*, 450, 691
- Kissler-Patig, M., Walsh, J. R., & Roth, M. M., eds. 2007, *Science Perspectives for 3D Spectroscopy*
- Klessen, R. S., & Glover, S. C. O. 2016, *Saas-Fee Advanced Course*, 43, 85
- Kobayashi, C., Karakas, A. I., & Lugaro, M. 2020, *ApJ*, 900, 179
- Kobulnicky, H. A., & Kewley, L. J. 2004, *ApJ*, 617, 240
- Köppen, J., & Hensler, G. 2005, *A&A*, 434, 531
- Krolik, J. H., & Kallman, T. R. 1984, *ApJ*, 286, 366
- Kroupa, P. 2001, *MNRAS*, 322, 231
- Kulkarni, S. R., & Heiles, C. 1987, *The Atomic Component*, Vol. 134 (Springer, Dordrecht), 87
- Kumari, N., Maiolino, R., Belfiore, F., & Curti, M. 2019, *MNRAS*, 485, 367
- Lacerda, E. A. D., Sánchez, S. F., Cid Fernandes, R., et al. 2020, *MNRAS*, 492, 3073
- Lacerda, E. A. D., Sánchez, S. F., Mejía-Narváez, A., et al. 2022, arXiv e-prints, arXiv:2202.08027
- Lacerda, E. A. D., Cid Fernandes, R., Couto, G. S., et al. 2018, *MNRAS*, 474, 3727
- Lamareille, F. 2010, *A&A*, 509, A53

- Le Borgne, J. F., Bruzual, G., Pelló, R., et al. 2003, *A&A*, 402, 433
- Leitherer, C., Ekström, S., Meynet, G., et al. 2014, *ApJS*, 212, 14
- Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, *ApJS*, 123, 3
- López-Cobá, C., Sánchez, S. F., Anderson, J. P., et al. 2020, *AJ*, 159, 167
- López-Sánchez, Á. R., Dopita, M. A., Kewley, L. J., et al. 2012, *MNRAS*, 426, 2630
- Lugo-Aranda, A. Z., Sánchez, S. F., Espinosa-Ponce, C., et al. 2022, arXiv e-prints, arXiv:2204.04252
- Luridiana, V., Morisset, C., & Shaw, R. A. 2015, *A&A*, 573, A42
- Luridiana, V., Peimbert, M., & Leitherer, C. 1999, *ApJ*, 527, 110
- Lyman, T. 1906, *ApJ*, 23, 181
- . 1914, *Nature*, 93, 241
- MacArthur, L. A., Courteau, S., Bell, E., & Holtzman, J. A. 2004, *ApJS*, 152, 175
- Maiolino, R., & Mannucci, F. 2019, *A&A Rev.*, 27, 3
- Maraston, C. 2005, *MNRAS*, 362, 799
- Marigo, P., Girardi, L., Bressan, A., et al. 2008, *A&A*, 482, 883
- Marino, R. A., Rosales-Ortega, F. F., Sánchez, S. F., et al. 2013, *A&A*, 559, A114
- Martín-Navarro, I., Vazdekis, A., La Barbera, F., et al. 2015, *ApJ*, 806, L31
- Martins, L. P., González Delgado, R. M., Leitherer, C., Cerviño, M., & Hauschildt, P. 2005, *MNRAS*, 358, 49
- Mathis, J. S. 1986, *PASP*, 98, 995
- Mathis, J. S. 1995, in *Revista Mexicana de Astronomía y Astrofísica Conference Series, Vol. 3, Revista Mexicana de Astronomía y Astrofísica Conference Series*, ed. M. Pena & S. Kurtz, 207
- Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, *ApJ*, 217, 425
- Matteucci, F. 1992, *Mem. Soc. Astron. Italiana*, 63, 301
- . 2021, *A&A Rev.*, 29, 5
- Matteucci, F., & Francois, P. 1989, *MNRAS*, 239, 885
- Matteucci, F., & Greggio, L. 1986, *A&A*, 154, 279
- Matteucci, F., Romano, D., & Molaro, P. 1999, *A&A*, 341, 458
- McDermid, R. M., Alatalo, K., Blitz, L., et al. 2015, *MNRAS*, 448, 3484
- McGaugh, S. S. 1991, *ApJ*, 380, 140

- Mendoza, C. 1983, in *Planetary Nebulae*, ed. L. H. Aller, Vol. 103, 143–172
- Mendoza, C., Boswell, J., Ajoku, D., & Bautista, M. 2014, *Atoms*, 2, 123
- Mendoza, C., Bautista, M. A., Deprince, J., et al. 2021, *Atoms*, 9, 12
- Meynet, G., & Maeder, A. 2000, *A&A*, 361, 101
- Miville-Deschênes, M.-A., Murray, N., & Lee, E. J. 2017, *ApJ*, 834, 57
- Molla, M., Ferrini, F., & Diaz, A. I. 1996, *ApJ*, 466, 668
- Mollá, M., García-Vargas, M. L., & Bressan, A. 2009, *MNRAS*, 398, 451
- Mollá, M., Vílchez, J. M., Gavilán, M., & Díaz, A. I. 2006, *MNRAS*, 372, 1069
- Morisset, C., Delgado-Inglada, G., & Flores-Fajardo, N. 2015, *Rev. Mexicana Astron. Astrofis.*, 51, 103
- Morisset, C., Luridiana, V., García-Rojas, J., et al. 2020, *Atoms*, 8, 66
- Morisset, C., Delgado-Inglada, G., Sánchez, S. F., et al. 2016, *A&A*, 594, A37
- Moustakas, J., Kennicutt, Robert C., J., Tremonti, C. A., et al. 2010, *ApJS*, 190, 233
- Muñoz-Mateos, J. C., Gil de Paz, A., Boissier, S., et al. 2007, *ApJ*, 658, 1006
- Naab, T., Oser, L., Emsellem, E., et al. 2014, *MNRAS*, 444, 3357
- Nava, A., Casebeer, D., Henry, R. B. C., & Jevremovic, D. 2006, *ApJ*, 645, 1076
- O'Dell, C. R., Peimbert, M., & Peimbert, A. 2003, *AJ*, 125, 2590
- Oey, M. S., Parker, J. S., Mikles, V. J., & Zhang, X. 2003, *AJ*, 126, 2317
- Orozco-Duarte, R., Wofford, A., Vidal-García, A., et al. 2022, *MNRAS*, 509, 522
- Osterbrock, D., & Flather, E. 1959, *ApJ*, 129, 26
- Osterbrock, D. E. 1960, *ApJ*, 131, 541
- Osterbrock, D. E., & Ferland, G. J. 2006, *Astrophysics of gaseous nebulae and active galactic nuclei* (University Science Books)
- Pagel, B. E. J., Edmunds, M. G., Blackwell, D. E., Chun, M. S., & Smith, G. 1979, *MNRAS*, 189, 95
- Papaderos, P., Izotov, Y. I., Thuan, T. X., et al. 2002, *A&A*, 393, 461
- Paschen, F. 1908, *Annalen der Physik*, 332, 537
- Peña, M., Stasińska, G., Esteban, C., et al. 1998, *A&A*, 337, 866
- Peña-Guerrero, M. A., Peimbert, A., & Peimbert, M. 2012a, *ApJ*, 756, L14
- Peña-Guerrero, M. A., Peimbert, A., Peimbert, M., & Ruiz, M. T. 2012b, *ApJ*, 746, 115



- Peimbert, A., Peimbert, M., & Ruiz, M. T. 2005, *ApJ*, 634, 1056
- Peimbert, M. 1967, *ApJ*, 150, 825
- Peimbert, M., & Goldsmith, D. W. 1972, *A&A*, 19, 398
- Peimbert, M., Peimbert, A., & Delgado-Inglada, G. 2017, *PASP*, 129, 082001
- Peimbert, M., Sarmiento, A., & Fierro, J. 1991, *PASP*, 103, 815
- Peimbert, M., & Torres-Peimbert, S. 1971, *ApJ*, 168, 413
- Peimbert, M., Torres-Peimbert, S., & Rayo, J. F. 1978, *ApJ*, 220, 516
- Pérez-Montero, E. 2014, *MNRAS*, 441, 2663
- Pérez-Montero, E., & Díaz, A. I. 2005, *MNRAS*, 361, 1063
- Pérez-Montero, E., García-Benito, R., Vílchez, J. M., et al. 2016, *A&A*, 595, A62
- Pettini, M., & Pagel, B. E. J. 2004, *MNRAS*, 348, L59
- Peña, M. 1986, *PASP*, 98, 1061
- Pilyugin, L. S. 2000, *A&A*, 362, 325
- Pilyugin, L. S., & Grebel, E. K. 2016, *MNRAS*, 457, 3678
- Pilyugin, L. S., Thuan, T. X., & Vílchez, J. M. 2003, *A&A*, 397, 487
- Pleuss, P. O., Heller, C. H., & Fricke, K. J. 2000, *A&A*, 361, 913
- Relaño, M., Monreal-Ibero, A., Vílchez, J. M., & Kennicutt, R. C. 2010, *MNRAS*, 402, 1635
- Relaño, M., Peimbert, M., & Beckman, J. 2002, *ApJ*, 564, 704
- Rix, H.-W., & White, S. D. M. 1992, *MNRAS*, 254, 389
- Rodríguez, M., & García-Rojas, J. 2010, *ApJ*, 708, 1551
- Rosales-Ortega, F. F., Díaz, A. I., Kennicutt, R. C., & Sánchez, S. F. 2011, *MNRAS*, 415, 2439
- Rosales-Ortega, F. F., Kennicutt, R. C., Sánchez, S. F., et al. 2010, *MNRAS*, 405, 735
- Rosales-Ortega, F. F., Sánchez, S. F., Iglesias-Páramo, J., et al. 2012, *ApJ*, 756, L31
- Roth, M. M., Kelz, A., Fechner, T., et al. 2005, *PASP*, 117, 620
- Rousseau-Nepton, L., Robert, C., Martin, R. P., Drissen, L., & Martin, T. 2018, *MNRAS*, 477, 4152
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- Samet, H. 1984, *ACM Comput. Surv.*, 16, 187–260
- Sánchez, S. F. 2006, *Astronomische Nachrichten*, 327, 850

- . 2020, *ARA&A*, 58, 99
- Sánchez, S. F., Cardiel, N., Verheijen, M. A. W., Pedraz, S., & Covone, G. 2007, *MNRAS*, 376, 125
- Sánchez, S. F., Rosales-Ortega, F. F., Kennicutt, R. C., et al. 2011, *MNRAS*, 410, 313
- Sánchez, S. F., Walcher, C. J., Lopez-Cobá, C., et al. 2021a, *Rev. Mexicana Astron. Astrofis.*, 57, 3
- Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012a, *A&A*, 538, A8
- Sánchez, S. F., Rosales-Ortega, F. F., Marino, R. A., et al. 2012b, *A&A*, 546, A2
- Sánchez, S. F., Rosales-Ortega, F. F., Iglesias-Páramo, J., et al. 2014, *A&A*, 563, A49
- Sánchez, S. F., Galbany, L., Pérez, E., et al. 2015a, *A&A*, 573, A105
- Sánchez, S. F., Pérez, E., Rosales-Ortega, F. F., et al. 2015b, *A&A*, 574, A47
- Sánchez, S. F., García-Benito, R., Zibetti, S., et al. 2016a, *A&A*, 594, A36
- Sánchez, S. F., Pérez, E., Sánchez-Blázquez, P., et al. 2016b, *Rev. Mexicana Astron. Astrofis.*, 52, 21
- . 2016c, *Rev. Mexicana Astron. Astrofis.*, 52, 171
- Sánchez, S. F., Avila-Reese, V., Hernandez-Toledo, H., et al. 2018, *Rev. Mexicana Astron. Astrofis.*, 54, 217
- Sánchez, S. F., Espinosa-Ponce, C., Carigi, L., et al. 2021b, *A&A*, 652, L10
- Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, *MNRAS*, 371, 703
- Sánchez-Blázquez, P., Rosales-Ortega, F. F., Méndez-Abreu, J., et al. 2014, *A&A*, 570, A6
- Sánchez-Menguiano, L., Sánchez, S. F., Pérez, I., et al. 2016, *A&A*, 587, A70
- . 2018, *A&A*, 609, A119
- Santoro, F., Kreckel, K., Belfiore, F., et al. 2022, *A&A*, 658, A188
- Savage, B. D., & Sembach, K. R. 1996, *ARA&A*, 34, 279
- Schaefer, A. L., Tremonti, C., Belfiore, F., et al. 2020, *ApJ*, 890, L3
- Schmidt, M. 1959, *ApJ*, 129, 243
- Seaton, M. J., & Osterbrock, D. E. 1957, *ApJ*, 125, 66
- Simpson, C., Ward, M., Clements, D. L., & Rawlings, S. 1996, *MNRAS*, 281, 509
- Smith, L. J., Norris, R. P. F., & Crowther, P. A. 2002, *MNRAS*, 337, 1309
- Spitzer, Lyman, J. 1956, *ApJ*, 124, 20

- Spitzer, L. 1978, *Physical processes in the interstellar medium* (John Wiley & Sons, Ltd), doi:10.1002/9783527617722
- Stasińska, G. 1984, *A&A*, 135, 341
- . 2005, *A&A*, 434, 507
- . 2009, in *The Emission-Line Universe*, ed. J. Cepa, 1
- Stasińska, G., Cid Fernandes, R., Mateus, A., Sodré, L., & Asari, N. V. 2006, *MNRAS*, 371, 972
- Stasińska, G., Izotov, Y., Morisset, C., & Guseva, N. 2015, *A&A*, 576, A83
- Stasińska, G., & Schaerer, D. 1999, *A&A*, 351, 72
- Stasińska, G., Vale Asari, N., Cid Fernandes, R., et al. 2008, *MNRAS*, 391, L29
- Storey, P. J. 1997, in *Planetary Nebulae*, ed. H. J. Habing & H. J. G. L. M. Lamers, Vol. 180, 161
- Strömberg, B. 1939, *ApJ*, 89, 526
- Sutherland, R. S., Bicknell, G. V., & Dopita, M. A. 1993, *ApJ*, 414, 510
- Sutherland, R. S., & Dopita, M. A. 2017, *ApJS*, 229, 34
- Szandala, T. 2021, *Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks* (Singapore: Springer Singapore), 203–224
- Tarter, C. B., Tucker, W. H., & Salpeter, E. E. 1969, *ApJ*, 156, 943
- Taylor, V. A., Jansen, R. A., Windhorst, R. A., Odewahn, S. C., & Hibbard, J. E. 2005, *ApJ*, 630, 784
- Thomas, A. D., Dopita, M. A., Kewley, L. J., et al. 2018, *ApJ*, 856, 89
- Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, *MNRAS*, 404, 1775
- Tielens, A. G. G. M. 2010, *The Physics and Chemistry of the Interstellar Medium*
- Tinsley, B. M. 1979, *ApJ*, 229, 1046
- Tuccillo, D., Huertas-Company, M., Decencière, E., et al. 2018, *MNRAS*, 475, 894
- Vale Asari, N., Couto, G. S., Cid Fernandes, R., et al. 2019, *MNRAS*, 489, 4721
- Vale Asari, N., Stasińska, G., Morisset, C., & Cid Fernandes, R. 2016, *MNRAS*, 460, 1739
- Valerdi, M., Peimbert, A., & Peimbert, M. 2021, *MNRAS*, 505, 3624
- van den Bergh, S. 1981, *AJ*, 86, 1464
- Vazdekis, A., Casuso, E., Peletier, R. F., & Beckman, J. E. 1996, *ApJS*, 106, 307
- Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., et al. 2010, *MNRAS*, 404, 1639

- Vázquez-Semadeni, E. 2012, in EAS Publications Series, Vol. 56, EAS Publications Series, ed. M. A. de Avillez, 39–49
- Vila-Costas, M. B., & Edmunds, M. G. 1992, MNRAS, 259, 121
- . 1993, MNRAS, 265, 199
- Vilchez, J. M. 1989, Ap&SS, 157, 9
- Vilchez, J. M., & Pagel, B. E. J. 1988, MNRAS, 231, 257
- Walcher, C. J., Coelho, P., Gallazzi, A., & Charlot, S. 2009, MNRAS, 398, L44
- Walcher, C. J., Coelho, P. R. T., Gallazzi, A., et al. 2015, A&A, 582, A46
- Walcher, C. J., Wisotzki, L., Bekeraité, S., et al. 2014, A&A, 569, A1
- Wang, K., Testi, L., Ginsburg, A., et al. 2015, MNRAS, 450, 4043
- Weaver, R., McCray, R., Castor, J., Shapiro, P., & Moore, R. 1977, ApJ, 218, 377
- Wegg, C., Gerhard, O., & Portail, M. 2017, ApJ, 843, L5
- Wei, Z. 1988, Astrophysical Letters and Communications, 27, 275
- Weinberg, D. H., Andrews, B. H., & Freudenburg, J. 2017, ApJ, 837, 183
- Weingartner, J. C., & Draine, B. T. 2001, ApJS, 134, 263
- Wofford, A., Charlot, S., Bruzual, G., et al. 2016, MNRAS, 457, 4296
- Wolfire, M. G., Hollenbach, D., McKee, C. F., Tielens, A. G. G. M., & Bakes, E. L. O. 1995, ApJ, 443, 152
- Xiao, L., Stanway, E. R., & Eldridge, J. J. 2018, MNRAS, 477, 904
- Yates, R. M., Henriques, B., Thomas, P. A., et al. 2013, MNRAS, 435, 3500
- Yates, R. M., Kauffmann, G., & Guo, Q. 2012, MNRAS, 422, 215
- Yin, S. Y., Liang, Y. C., Hammer, F., et al. 2007, A&A, 462, 535
- Zurita, A., Florido, E., Bresolin, F., Pérez, I., & Pérez-Montero, E. 2021, MNRAS, 500, 2380