

# **Exame de qualificação de doutorado: Poeira, metalicidade e formação estelar: estimando fração de gás.**

Eduardo Alberto Duarte Lacerda

Orientador:

Prof. Dr. Roberto Cid Fernandes Jr.

Coorientadora:

Rosa M. Gonzalez Delgado

• • •

Universidade Federal de Santa Catarina  
Centro de Ciências Físicas e Matemáticas  
Programa de Pós-Graduação em Física

Florianópolis (SC) – 7 de janeiro de 2018



## Resumo

Neste trabalho estamos estudando a conversão do coeficiente de extinção por poeira em densidade superficial de massa de gás em discos de galáxias espirais. Para tal finalidade estamos utilizando uma amostra de 184 galáxias do projeto CALIFA, um *survey* astronômico de espectroscopia de campo integrado (*Integral Field Spectroscopy - IFS*). Utilizando as medidas de linhas de emissão provenientes dos espectros residuais da síntese de populações estelares, calculamos o coeficiente de extinção através do decremento de Balmer ( $\tau_V^{\text{neb}}$ ) de maneira a compará-lo com aquele proveniente da síntese ( $\tau_V^*$ ). Construímos um objeto em `PYTHON` que organiza essas medidas, facilitando a utilização desses resultados juntamente com os resultados da síntese.

Selecionamos apenas zonas pertencentes aos discos de galáxias e que possuam formação estelar, juntamente com um controle de qualidade. Neste processo, o corte mais brusco em nossa amostra é devido a baixa relação sinal-ruído da linha de ( $S/N < 3$ ) em 91142 zonas. Nossa amostra final engloba 16479 zonas reamostradas em  $\sim 4500$  anéis radiais (elípticos) em uma unidade natural da galáxia (raio no qual abarca metade da luz - HLR).

Em paralelo estamos trabalhando em duas frontes. Na primeira estamos buscando o sentido real de  $\tau_V^*$  utilizando um modelo proposto juntamente com  $\tau_V^{\text{neb}}$  de maneira que possamos encontrar os coeficientes diretamente ligados ao meio interestelar e às nuvens formadoras de estrelas (regiões H II). A segunda, de certa forma, depende da primeira, pois estamos fazendo a conversão de poeira em gás, portanto  $\tau_V$  influencia diretamente nosso resultado. Nessa etapa investigamos relação entre formação estelar e poeira, de maneira a reproduzir uma relação como a de Kennicutt-Schmidt (KS), encontrando uma correlação entre ambos. Através dos resíduos desta relação (que neste trabalho chamamos de *pseudo* KS) vimos que algumas propriedades estão correlacionadas com o espalhamento desta relação, na qual, a mais forte parece ser com a densidade superficial de massa estelar ( $\mu_*$ ). Também utilizamos tanto  $\tau_V^*$  quanto  $\tau_V^{\text{neb}}$  para calcular  $\Sigma_{\text{gas}}$ , elemento que nos falta para calcular frações de gás, ingrediente muito importante na evolução química de galáxias.



# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	O todo e as partes . . . . .	2
1.2	Gás ionizado difuso (DIG) . . . . .	3
1.2.1	Primeiras detecções . . . . .	3
1.2.2	Fonte de ionização do DIG . . . . .	4
1.3	O GAS-UFSC e o IAA-CSIC . . . . .	4
1.3.1	Artigo - CALIFA, the Calar Alto Legacy Field Area survey IV. Third public data release. . . . .	5
1.4	Este trabalho . . . . .	5
<b>A</b>	<b>Artigos publicados</b>	<b>7</b>
A.1	Diffuse ionized gas in galaxies across the Hubble sequence at the CALIFA resolution . . . . .	7
A.2	CALIFA, the Calar Alto Legacy Integral Field Area survey. IV. Third public data release . . . . .	23
<b>Referências Bibliográficas</b>		<b>I</b>



# **Lista de Figuras**



# Capítulo 1

## Introdução

A única forma empírica de estudarmos galáxias é através da luz emitida pelos seus constituintes. Mais precisamente, das imagens e da distribuição espectral de energia (SED<sup>1</sup>) que chegam até nossos telescópios, em terra ou no espaço. Diferentes componentes e eventos os modificam, nos possibilitando a busca de padrões e criação de modelos que se propõem a explicar sua constituição, formação e dinâmica. Atualmente, existem diversos projetos astronômicos de levantamento de informações ou mapeamento de regiões do céu, chamados de *surveys*, formando uma rede de gigantescos bancos de dados de imagens, espectros e metainformação. Com diferentes faixas espectrais (desde raios- $\gamma$  até micro-ondas), diferentes fontes de dados (spectros de galáxias integradas, espectroscopia de campo, imagens, monitoramento temporal de eventos) e diferentes objetivos, os *surveys* astronômicos permeiam por diferentes fenômenos astrofísicos. Através dessa criação e difusão em massa de informações, nossa forma de enxergar o mundo vem se tornando cada vez mais acurada quanto ao Universo. Além de estarem formando um imenso legado de informações para futuros astrofísicos, são basilares para o desenvolvimento de novas ideias e resolução dos desafios atuais da área. Neste capítulo faço uma introdução do atual cenário com um breve resumo dos avanços que nosso grupo de astrofísica (GAS-UFSC) têm obtido através desses levantamentos, cenário no qual esta tese está inserida.

---

<sup>1</sup>*Spectral energy distribution* - quantidade de energia em cada comprimento de onda.

## 1.1 O todo e as partes

Galáxias são formadas por uma complexa mistura de gás, poeira, estrelas e matéria escura, distribuídas em discos, bulbos e halos. Os primeiros grandes levantamentos de dados especiais (SDSS<sup>2</sup>, York et al. 2000; COSMOS<sup>3</sup>, Scoville et al. 2007; ALHAMBRA<sup>4</sup>, Moles et al. 2008; são alguns exemplos) tratavam galáxias como uma fonte puntual de energia, conhecido com espectroscopia de uma fibra (*single-fiber spectroscopy*<sup>5</sup>). Apesar dessa limitação, muito se aprendeu (e ainda se aprende) sobre a formação e evolução das galáxias. Exemplos incluem a conexão entre o poder do AGN<sup>6</sup> e as populações estelares (Kauffmann et al. 2003); a relação entre a taxa de formação estelar (SFR<sup>7</sup>) e a massa estelar das galáxias (Brinchmann et al. 2004); a relação massa-metalicidade (MZR<sup>8</sup>; Tremonti et al. 2004); a evolução química e a história de formação estelar das galáxias(Cid Fernandes et al. 2007; Asari et al. 2007); relação massa estelar-metalicidade (Vale Asari et al. 2009); e mais importantes para o escopo desta tese, a revelação de uma imensa e esquecida população de galáxias aposentadas ionizadas por HOLMES<sup>9</sup> (Stasińska et al. 2008; Cid Fernandes et al. 2010, 2011). Entretanto, esse é um dos mais significantes problemas desse tipo de *survey*.

Podemos perceber que qualquer propriedade que varie em função da posição dentro da galáxia será erroneamente estimada quando há apenas um espectro representando a galáxia inteira. Quando estimamos propriedades referentes a diferentes regimes de ionização na galáxia, como a metalicidade nebulosa por exemplo, também temos problemas. Nesse caso, devemos levar em conta apenas os fótons gerados nas regiões de formação estelar (SF<sup>10</sup>), isolando-os daqueles que vêm de outros regimes nebulosos, como o gás ionizado difuso (DIG<sup>11</sup>), fotoionização pelo núcleo ativo ou estrelas velhas. Dessa forma, para um estudo mais preciso das propriedades derivadas dos espectros integrados e, por consequência, do viés causado por construção, um melhor entendimento desses efeitos se faz necessário.

Um grande passo nessa direção foi dado com a criação dos *surveys* de espectroscopia de campo (IFS<sup>12</sup>). Através da IFS podemos desvencilhar essa mistura de partes distintas, pois

---

<sup>2</sup>Sloan Digital Sky Survey

<sup>3</sup>Cosmic Evolution Survey

<sup>4</sup>Advanced Large Homogeneous Area Medium-Band Redshift Astronomical survey

<sup>5</sup>Um espectro formado apenas de uma determinada área de abertura observada.

<sup>6</sup>Active galactic nucleus

<sup>7</sup>Star-formation rate

<sup>8</sup>Mass-metallicity relation

<sup>9</sup>Hot low-mass evolved stars, estrelas quentes de baixa massa em alto estado de evolução.

<sup>10</sup>Star-forming

<sup>11</sup>Diffuse ionized gas

<sup>12</sup>Integral field spectroscopy

nessa técnica de observação temos espectros para cada parte da galáxia. Assim, para cada par espacial ( $x, y$ ) temos uma dimensão espectral  $\lambda$ . Quanto menores sejam os pixels cobrindo uma mesma área (melhor resolução espacial) e melhor resolução espectral cobrindo um maior intervalo de comprimento de onda teremos uma melhor definição da localização e da assinatura espectral de cada uma das partes. Diversos *surveys* IFS já estão finalizados e com seus dados disponíveis publicamente (CALIFA<sup>13</sup> DR3<sup>14</sup>, Sánchez et al. 2016; PINGS<sup>15</sup>, Rosales-Ortega et al. 2010), outros ainda estão em fase de observação e com alguns dados já disponíveis (MaNGA<sup>16</sup> SDSS-IV DR13, Albareti et al. 2017; SAMI<sup>17</sup> DR1, Green et al. 2017). Com o desenvolvimento de novos equipamentos como o MUSE<sup>18</sup> e o SITELLE<sup>19</sup> poderemos estudar galáxias e suas interações com ainda mais detalhes.

Nessa direção, este trabalho usa os dados de IFS do CALIFA *survey* para estudar a importância e a caracterização do DIG em diferentes regiões de galáxias de todos os tipos morfológicos, que resultou no artigo no Apêndice A.1 (Lacerda et al. 2018). A completa cobertura de galáxias com diferentes morfologias e diferentes inclinações faz do CALIFA um ideal para esse tipo de estudo, mesmo sabendo que a resolução espacial não nos permite uma descrição detalhada das diferentes componentes do meio interestelar (ISM<sup>20</sup>). IFS com melhor resolução já existem (Sánchez et al. 2015; Vogt et al. 2017; Rousseau-Nepton et al. 2017), mas como cobrem tão poucos objetos não podemos usá-los para um estudo mais geral.

## 1.2 Gás ionizado difuso (DIG)

### 1.2.1 Primeiras detecções

O DIG foi detectado pela primeira vez no disco Galáctico através de linhas de emissão fracas fora de regiões H II<sup>21</sup> clássicas (Reynolds 1971). Observações de galáxias espirais de lado através de imageamento em H $\alpha$  (Dettmar 1990; Hoopes et al. 1996, 1999) mostraram

<sup>13</sup>Calar Alto Legacy Integral Field Area survey

<sup>14</sup>Data-release 3

<sup>15</sup>PPAK IFS Nearby Galaxies survey

<sup>16</sup>Mapping nearby Galaxies at Apache Point Observatory

<sup>17</sup>equipamento e *survey* são homônimos - Sydney-AAO Multi-object Integral-field spectrograph

<sup>18</sup>The Multi Unit Spectroscopic Explorer - <https://www.eso.org/sci/facilities/develop/instruments/muse.html>

<sup>19</sup>Spectromètre Imageur à Transformée de Fourier pour l'Etude en Long et en Large de raies d'Emission - <http://cfht.hawaii.edu/Instruments/Sitelle/>

<sup>20</sup>Interstellar medium

<sup>21</sup>Regiões formadoras de estrelas; são formadas por imensas nuvens de gás molecular, originado pelo esfriamento de gás do meio interestelar, que se fragmentam formando estruturas menores e cada vez mais densas.

a existência de DIG a grandes distâncias do plano galáctico. [Oey et al. \(2007\)](#), estudando 109 galáxias do SINGS<sup>22</sup>, chegaram a conclusão que emissão difusa em H $\alpha$  está presente em galáxias de todos os tipos morfológicos e representa  $\sim 60\%$  da emissão total em H $\alpha$ , independentemente do tipo morfológico ou da SFR total.

### 1.2.2 Fonte de ionização do DIG

Fótons de estrelas massivas do tipo OB escapando das regiões H II é a fonte de ionização mais comumente adotada para explicar o DIG (veja o review em [Haffner et al. 2009](#)). Entretanto, razões de linhas como [N II]/H $\alpha$ , [S II]/H $\alpha$ , and [O III]/H $\beta$  crescem com a altura em relação ao plano galáctico, fazendo com que seja necessário a inclusão de fontes adicionais (ou alternativas) de ionização.

## 1.3 O GAS-UFSC e o IAA-CSIC

Nos últimos anos nosso grupo de Astrofísica (GAS-UFSC) aqui na Universidade Federal de Santa Catarina vem trabalhando com dados de diversos *surveys*. Nosso grupo foi pioneiro no estudo das propriedades físicas das populações estelares de aproximadamente um milhão de galáxias do *SDSS* através do projeto SEAGal/STARLIGHT<sup>23</sup> publicando diversos artigos importantes e amplamente citados (e.g., [Cid Fernandes et al. 2005](#); [Mateus et al. 2006](#); [Stasińska et al. 2006](#); [Asari et al. 2007](#); [Stasińska et al. 2008](#); [Cid Fernandes et al. 2011](#)).

Desde antes e durante esse trabalho, participamos de um projeto entre nosso grupo de populações estelares aqui no GAS com pesquisadores do Instituto de Astrofísica de Andalucía (IAA), na cidade de Granada, Comunidade autônoma de Andalucía, ao sul da Espanha. Esse instituto pertence ao *Consejo Superior de Investigaciones Científicas* (CSIC), o maior órgão público (estatal) de pesquisas científicas na Espanha, e o terceiro maior da Europa. Conta com pesquisadores participantes do CALIFA *survey*, funcionando como centro físico do projeto. A pesquisadora Rosa M. González Delgado, coorientadora deste trabalho, uma das principais líderes do projeto e que também atuou como Pesquisadora Visitante Especial (PVE-CsF) aqui na UFSC; Rubén García Benito, que faz parte do grupo de redução dos dados do *survey*; e Enrique Pérez, do grupo de populações estelares, já trabalham em nossa parceria e possuem conhecimento e domínio das técnicas exploradas por nosso projeto, além de participarem ati-

---

<sup>22</sup>*Spitzer Infrared Nearby Galaxies Survey*

<sup>23</sup><http://starlight.ufsc.br>

vamente do desenvolvimento do CALIFA. Durante os últimos cinco anos nosso grupo de populações estelares no CALIFA publicou diversos artigos (e.g., Pérez et al. 2013; González Delgado et al. 2014b,a, 2015, 2016; de Amorim et al. 2017; González Delgado et al. 2017; García-Benito et al. 2017; Lacerda et al. 2017). Paralelamente participamos de diversos congressos e conferências publicando nossos resultados. Detalhes técnicos e comparações entre surveys IFS podem ser encontrados em de Amorim (2015).

### 1.3.1 Artigo - CALIFA, the Calar Alto Legacy Field Area survey IV. Third public data release.

Houveram três lançamentos públicos dos dados armazenados pelo CALIFA durante o decorrer do projeto, finalizando com o DR3<sup>24</sup> (Sánchez et al. 2016, que também pode ser encontrado no Apêndice A.2 deste trabalho). Este DR conta com espectros de 667 galáxias ( $\sim 1,5$  milhões de espectros) com tipos morfológicos cobrindo toda a classificação de Hubble e redshifts variando entre 0.005 e 0.03 (distâncias de 20 a 130 Mpc).

Nosso grupo de populações estelares se encarregou de escrever alguns programas de análise e gerar as imagens nas quais nos embasamos para avaliar a qualidade da síntese de populações estelares com o STARLIGHT para distintas posições em cada galáxia, além de realizarmos uma comparação com os resultados para a amostra do DR1 (Husemann et al. 2013). e do DR2 García-Benito et al. 2015. Durante esse trabalho notamos que os resíduos reduziram sensivelmente desde a última versão. Esta mesma análise nos ajudou a melhorar a máscara de remoção de linhas telúricas<sup>25</sup> dos espectros. Também verificamos que os erros relacionados aos espectros observados possuem uma distribuição muito próxima a uma gaussiana.

## 1.4 Este trabalho

Essa tese é um apanhado de alguns dos trabalhos no qual participei durante o tempo do doutorado, com foco principal no artigo sobre a natureza das linhas de emissão das regiões das galáxias do CALIFA, separando aquelas SF daquelas melhores caracterizadas como DIG. No Cap. ?? apresento a amostra de galáxias utilizadas neste trabalho. O artigo principal será dividido em dois capítulos. A apresentação do método de caracterização está no Cap. ??, e no próximo (Cap. ??) fazemos a discussão sobre esse método de classificação de regiões e a

---

<sup>24</sup>Data-release 3

<sup>25</sup>Linhas provenientes de fenômenos que ocorrem na Terra.

análise das regiões classificadas. Por fim, demais artigos e trabalhos são apresentados no Cap. ??.

## **Apêndice A**

### **Artigos publicados**

#### **A.1 Diffuse ionized gas in galaxies across the Hubble sequence at the CALIFA resolution**

Artigo por [Lacerda et al. \(2018\)](#) ([10.1093/mnras/stx3022](https://doi.org/10.1093/mnras/stx3022)). Também disponível em *pre-print* ([arXiv:1711.07844](https://arxiv.org/abs/1711.07844)).



# Diffuse ionized gas in galaxies across the Hubble sequence at the CALIFA resolution

E. A. D. Lacerda<sup>1\*</sup>, R. Cid Fernandes<sup>1</sup>, G. S. Couto<sup>1</sup>, G. Stasińska<sup>2</sup>, R. García-Benito<sup>3</sup>, N. Vale Asari<sup>1</sup>, E. Pérez<sup>3</sup>, R. M. González Delgado<sup>3</sup>, S. F. Sánchez<sup>4</sup>, A. L. de Amorim<sup>1</sup>

<sup>1</sup>Departamento de Física - CFM - Universidade Federal de Santa Catarina, P.O. Box 476, 88040-900. Florianópolis, SC, Brazil

<sup>2</sup>LUTH, Observatoire de Paris, CNRS, Université Paris Diderot; Place Jules Janssen 92190 Meudon, France

<sup>3</sup>Instituto de Astrofísica de Andalucía (CSIC), P.O. Box 3004, 18080 Granada, Spain

<sup>4</sup>Instituto de Astronomía, Universidad Nacional Autónoma de México, A. P. 70-264, C.P. 04510, México, D.F., México

20 November 2017

## ABSTRACT

We use spatially resolved spectroscopy from the Calar Alto Legacy Integral Field Area (CALIFA) survey to study the nature of the line emitting gas in galaxies of different Hubble types, focusing on the separation of star-forming (SF) regions from those better characterized as diffuse ionized gas (DIG). The diagnosis is carried out in terms of the equivalent width of H $\alpha$  ( $W_{\text{H}\alpha}$ ). Three nebular regimes are identified. Regions where  $W_{\text{H}\alpha} < 3 \text{ \AA}$  define what we call the hDIG, the component of the DIG where photoionization is dominated by hot, low-mass, evolved stars. Regions where  $W_{\text{H}\alpha} > 14 \text{ \AA}$  trace SF complexes.  $W_{\text{H}\alpha}$  values in the intermediate 3–14  $\text{\AA}$  range reflect a mixed regime (mDIG) where more than one process contributes. This three-tier scheme is inspired both by theoretical and empirical considerations. Its application to CALIFA galaxies of different types and inclinations leads to the following results: (i) the hDIG component is prevalent throughout ellipticals and S0's as well as in bulges, and explains the strongly bimodal distribution of  $W_{\text{H}\alpha}$  both among and within galaxies. (ii) Early-type spirals have some hDIG in their discs, but this component becomes progressively less relevant for later Hubble types. (iii) hDIG emission is also present above and below galactic discs, as seen in several edge-on spirals in our sample. (iv) The SF/mDIG proportion grows steadily from early- to late-type spirals, and from inner to outer radii. (v) Besides circumventing basic inconsistencies in conventional DIG/SF separation criteria based on the H $\alpha$  surface brightness, our  $W_{\text{H}\alpha}$ -based method produces results in agreement with a classical excitation diagram analysis.

**Key words:** surveys – galaxies: ISM – ISM: general

## 1 INTRODUCTION

Much of what we know about galaxies comes from their optical emission-lines. Measurements of lines like H $\alpha$ , H $\beta$ , [O III]  $\lambda 5007$ , [N II]  $\lambda 6584$  and others allow properties like star formation rates, dust attenuation, gas-phase metallicity, and the presence and intensity of an active nucleus to be estimated through well-known recipes. This approach of converting emission-line data on to astrophysically valuable information has been extensively explored in the context of large surveys like the Sloan Digital Sky Survey (York et al. 2000), producing important results in the quest for a better understanding of galaxy physics. Examples include the relation between the star formation rate and galaxy stellar

mass (Brinchmann et al. 2004), the mass–metallicity relation (Tremonti et al. 2004), connections between AGN power and stellar populations (Kauffmann et al. 2003), and between gas-phase metallicities and star formation histories (Cid Fernandes et al. 2007), to name but a few.

A fundamental caveat of integrated-light (single-fiber) spectroscopy of galaxies is that it effectively treats as a single point source what is in reality a complex system, with physically and structurally different components. Disc and bulge, arm and inter-arm, H II regions and diffuse gas, dusty and clean regions, young and old stars are all mixed in a total spectrum where the parts are no longer recognizable. This simplified view of galaxies is bound to affect estimates of their properties. For instance, when estimating nebular metallicities, one would in principle need to count only the emission-line photons produced in star-forming (SF) re-

\* E-mail: lacerda@astro.ufsc.br

gions, isolating them from those coming from other nebular regimes, like the diffuse ionized gas (DIG), or photoionization by an active nucleus or old stars. Similarly, estimates of the dust attenuation and star formation rate from a single spectrum forcefully neglect variations in dust content across the face of galaxies. A proper assessment of these effects is important to better understand the precise meaning of (and potential biases in) properties derived from integrated-light data.

As a starting step towards this goal, this paper takes advantage of integral field spectroscopy from the CALIFA survey (Sánchez et al. 2016) to study the importance and characteristics of the DIG in different regions of galaxies of all types. The complete coverage in terms of Hubble types and inclinations makes CALIFA ideal for a first general exploration of the DIG in galaxies of the local Universe, even though its spatial resolution does not allow for a detailed description of the different interstellar medium (ISM) components. Imaging spectroscopy of galaxies with higher spatial resolution is developing (e.g., Sánchez et al. 2015; Vogt et al. 2017; Rousseau-Nepton et al. 2017) but the data available so far cover too few objects for an overall study as the one presented here.

The DIG was first detected in the Galactic disc through faint emission-lines outside the classical H II regions (Reynolds 1971). Observations of edge-on spiral galaxies using deep H $\alpha$  imaging (Dettmar 1990; Hoopes et al. 1996, 1999) showed the existence of DIG not only in the vicinity of H II regions but also at large distances above galaxy planes. By studying 109 H I selected galaxies in the SINGS survey, Oey et al. (2007) came to the conclusion that diffuse H $\alpha$  emission is present in galaxies of all types and represents about 60 per cent of the total H $\alpha$  emission, irrespective of the galaxy Hubble type or total star formation rate.

Radiation from massive OB stars leaking out from H II regions is a commonly advocated ionization source for the DIG (see review by Haffner et al. 2009). However, the existence of additional/alternative sources is required by the increase of [N II]/H $\alpha$ , [S II]/H $\alpha$ , and [O III]/H $\beta$  with galactic height found in many galaxies.

Such features cannot be reproduced with models of photoionization by hot, massive stars, even taking into account the hardening of the ionizing radiation due to intervening absorption (Hoopes & Walterbos 2003). The most commonly invoked sources of additional ionization/heating are shocks (Collins & Rand 2001a), turbulent mixing layers (Slavin et al. 1993; Binette et al. 2009), magnetic reconnection, cosmic rays or photoelectric emission from small grains (Reynolds et al. 2001), and hot low-mass evolved stars (HOLMES; Flores-Fajardo et al. 2011). HOLMES have also been invoked as the ionization source of the weak emission-lines in retired galaxies (Stasińska et al. 2008 and Cid Fernandes et al. 2011, hereafter CF11). These are systems that have stopped forming stars and are ionized by their hot, old stellar populations, producing LINER-like emission-line ratios, a phenomenon that is common in both ellipticals and in bulges of spiral galaxies (Sarzi et al. 2010; Gomes et al. 2016a; Belfiore et al. 2016).

Regardless of what powers the DIG, its nebular regime differs from that in H II regions, with lower densities, lower ionization parameters, and higher electron temperatures.

Failure to account for its contribution may thus lead to biases in the derived properties of galaxies.

In the literature, DIG and SF regions are separated on the basis of the H $\alpha$  surface brightness,  $\Sigma_{\text{H}\alpha}$ . This is a natural criterion, since the H $\alpha$  surface brightness is directly related to the density of the ionized gas. Zhang et al. (2017), for instance, use  $\Sigma_{\text{H}\alpha} > 10^{39} \text{ erg s}^{-1} \text{kpc}^{-2}$  to ‘select reliable H II-dominated spaxels’. In other studies, the criterion is not a simple  $\Sigma_{\text{H}\alpha}$  threshold, but still based on  $\Sigma_{\text{H}\alpha}$  (see discussions in Zurita et al. 2000, Oey et al. 2007, and Vogt et al. 2017). This intuitively valid approach is however not fully adequate. As argued in this paper, separating SF and DIG regions on the basis of  $\Sigma_{\text{H}\alpha}$  is conceptually incorrect, and may lead to inconsistent results under certain circumstances. Furthermore,  $\Sigma_{\text{H}\alpha}$  gives no clue as to the nature of the DIG emission.

These drawbacks are solved with a simple diagnostic based on the equivalent width of H $\alpha$  ( $W_{\text{H}\alpha}$ ). As shown in this paper,  $W_{\text{H}\alpha}$  correctly tracks the qualitative differences inherent to the SF and DIG regimes, and is also capable of identifying the component of the DIG corresponding to gas predominantly ionized by HOLMES, the hDIG, in an observationally simple and physically sound fashion. The nature of the remaining (neither HOLMES nor SF-dominated) DIG emission is most probably a mixture of processes, and will hereafter be dubbed mDIG, where the ‘m’ stands for ‘mixed’.

This paper is organized as follows. Section 2 describes the data and processing steps. Section 3 presents our  $W_{\text{H}\alpha}$ -based method to separate DIG from SF spaxels, and to distinguish the mDIG and hDIG components. Section 4 applies the method to CALIFA data to study hDIG/mDIG/SF fractions and the nature of extraplanar DIG emission in edge-on systems. Comparisons with  $\Sigma_{\text{H}\alpha}$ -based methods and a classical excitation diagram analysis are also presented. Finally, Section 5 highlights our main findings.

## 2 DATA

The spatially resolved 3650–6850 Å spectra covering the whole optical extent of relatively nearby galaxies gathered by CALIFA (Sánchez et al. 2012, 2016; Husemann et al. 2013; García-Benito et al. 2015) provide a suitable data set to address the issues raised above. We use the COMBO data cubes, obtained by merging the observations with the two grisms used by CALIFA. The spectral resolution is 6 Å in full width at half-maximum, the field of view is slightly over 1 arcmin<sup>2</sup>, and the spaxel size is 1 × 1 arcsec<sup>2</sup>, but the spatial resolution is ∼2.5 arcsec. At the distance of our sources (20–123 Mpc), this corresponds to 0.2–1.5 kpc (0.8 kpc on the median).

Our working sample contains 391 galaxies, morphologically distributed as follows: 57 ellipticals, 47 S0–S0a, 62 Sa–Sab, 67 Sb, 70 Sbc, and 88 Sc or later. These same six bins in galaxy morphology will be used to evaluate how the SF, mDIG, and hDIG components vary across the Hubble sequence. Morphologically distorted systems (such as those studied by Wild et al. 2014; Barrera-Ballesteros et al. 2015a,b; Cortijo-Ferrero et al. 2017a,b) were discarded from the analysis. Besides its diversity in Hubble types, the sample also covers inclination angles from edge-on to face-on.

All data cubes were processed through the PyCASSO<sup>1</sup> pipeline described in Cid Fernandes et al. (2013) and de Amorim et al. (2017). Briefly, after masking artefacts, foreground sources, and low signal-to-noise ( $SN$ ) regions, the data cubes are binned into Voronoi zones gauged to reach an  $SN$  of 20 or more in the continuum around 5635 Å. Our sample contains 307958 zones ( $\sim 800$  per galaxy). These zone spectra are then processed through the STARLIGHT code (Cid Fernandes et al. 2005), obtaining a model  $M_\lambda$  for the stellar continuum. Previous papers using this pipeline concentrated on the analysis of the spatially resolved properties of the stellar populations as deduced from the spectral fits (Pérez et al. 2013; González Delgado et al. 2014, 2015, 2016, 2017). Instead, this paper focuses on emission-line properties.

Emission-line fluxes are measured with the SHERPA IFU line fitting software (SHIFU; García-Benito et al. 2017), based on CIAO’s SHERPA package (Freeman et al. 2001; Doe et al. 2007) by fitting single Gaussians to the  $R_\lambda = O_\lambda - M_\lambda$  residual spectra obtained after subtracting the STARLIGHT fit ( $M_\lambda$ ) from the observed spectrum ( $O_\lambda$ ). These can be delicate measurements in the cases where lines are weak, as it is often the case with H $\beta$ . This study deals almost exclusively with H $\alpha$ , whose flux is much less affected by uncertainties. Indeed, the median  $SN_{H\alpha}$  is 16 over all our zones, and in only 5 per cent of the cases  $SN_{H\alpha} < 1$ .

Fig. 1 shows SDSS stamps along with our CALIFA-based  $\Sigma_{H\alpha}$  and  $W_{H\alpha}$  maps for a selection of galaxies in our sample. Dotted ellipses mark the distance to the nucleus ( $R$ ), measured along the major axis of ellipses defined on the basis of the spatial moments of the continuum flux in the (rest-frame)  $5635 \pm 45$  Å continuum. We measure  $R$  in units of the half-light radius (HLR), measured in the same continuum window around  $\lambda = 5635$  Å. As illustrated by the studies of Pérez et al. (2013), Sánchez et al. (2014), and González Delgado et al. (2016), the HLR is a convenient metric for comparing galaxies of different sizes. For our galaxies,  $HLR = 3.9 \pm 1.7$  kpc (mean  $\pm$  dispersion). In spirals, one may generally associate  $R > 1$  HLR with the disc and  $R < 0.5$  HLR with the bulge. The meaning of  $R$  becomes ambiguous for highly inclined systems, a limitation that does not affect our main results.

The constant patches in the example  $\Sigma_{H\alpha}$  and  $W_{H\alpha}$  maps in Fig. 1 correspond to the Voronoi zones used to ensure reliable continuum spectra to be processed through STARLIGHT. As seen in the figure, these are always located in the fainter, outer parts of our galaxies. Within  $R < 1$ , HLR over 97 per cent of the individual spaxels have  $SN > 20$ , and so no spatial binning is performed. Of the 307958 spectra in our sample, 274534 (89 per cent) correspond to single spaxels. The remaining ones correspond to Voronoi zones containing six spaxels in the median.

Part of the analysis that follows is based on the statistics of  $W_{H\alpha}$  among the spectra in our sample. The variable size of our extractions introduces some distortion in these statistics, since a large zone counts the same as one containing a single spaxel. The effects of this distortion will be discussed below, but we anticipate that they do not affect the main results reported in this paper.

<sup>1</sup> Python CALIFA STARLIGHT Synthesis Organizer

### 3 $W_{H\alpha}$ AS A DIAGNOSTIC OF THE NEBULAR REGIME: SFc, mDIG, AND hDIG

We wish to devise a method to characterize the dominant nebular regime in the various galaxy regions, one that is able to distinguish SF- from DIG-dominated zones, and also to differentiate DIG components. The ulterior goal of establishing such a taxonomical scheme is to provide a basis for other studies. With a solid classification scheme, one can, for instance, perform a comparative study of the physical properties (say, dust content or stellar populations) associated with these components of the ISM, or evaluate possible biases resulting from the mix of different nebular regimes in single-fibre (one spectrum per galaxy) observations.

Given that our spatial resolution of  $\sim 0.8$  kpc exceeds the physical scale of H II regions (even giant ones like 30 Dor or NGC 604, which are typically 0.1–0.3 kpc across, e.g.; González Delgado & Pérez 2000), SF regions in our data are bound to contain diffuse gas emission. Naturally, this mixture is even stronger in regions spatially binned into Voronoi zones. To convey this relevant fact in our nomenclature, we will hereafter use the term star-forming complexes (SFc) to refer to regions that contain H II regions, but inevitably mixed with DIG emission in our data. A more precise definition is that our SFc are the zones with a larger SF/DIG ratio. Note that our SFc are not even necessarily *dominated* by star formation, but simply contain a good proportion of SF-powered line emission. This relative, ranking-based scale is implicit in all that follows.

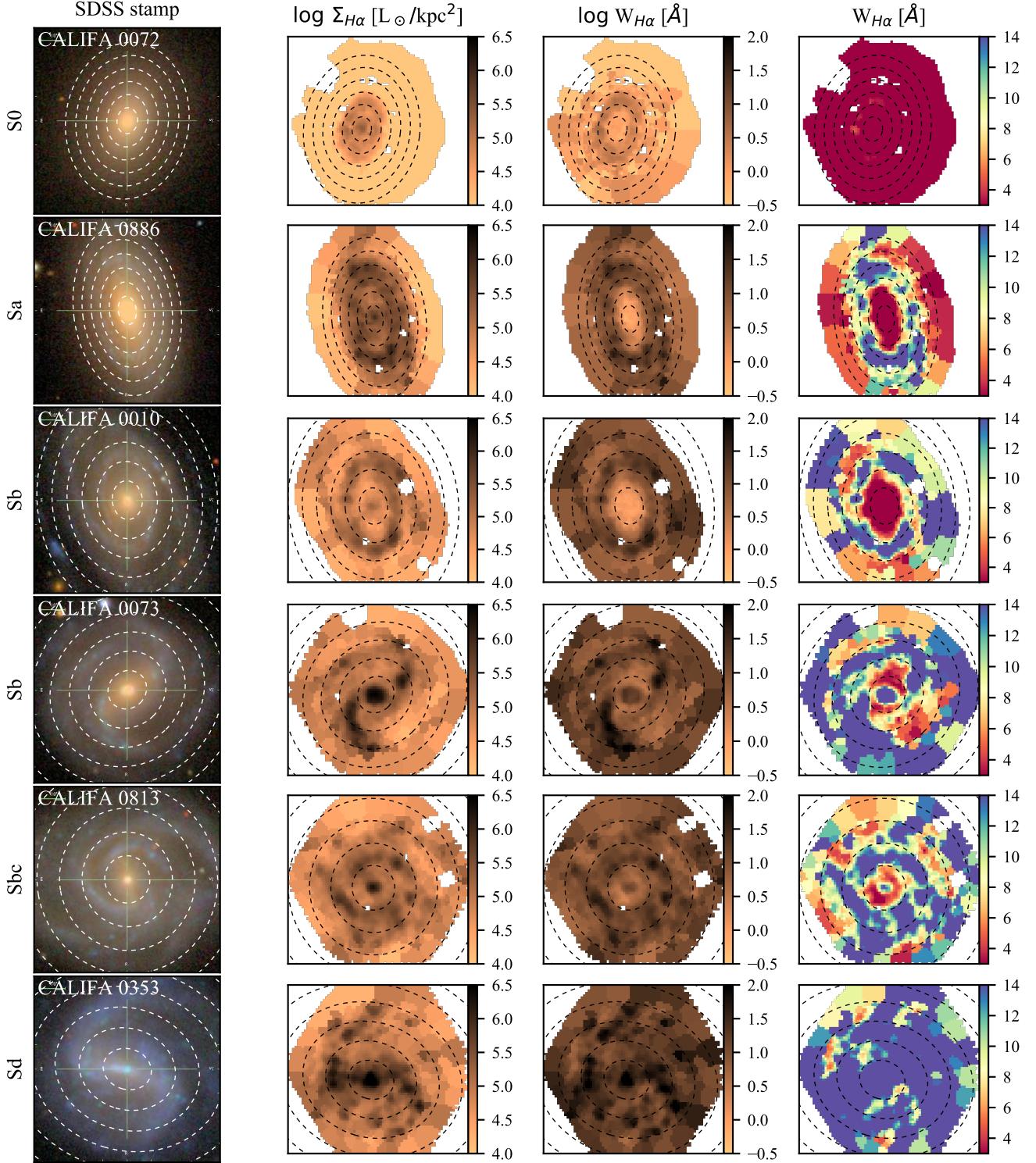
#### 3.1 Rationale for a $W_{H\alpha}$ -based SF/DIG separator

Previous work favours the use of the H $\alpha$  surface brightness ( $\Sigma_{H\alpha}$ ) as a means to separate the DIG from SF regions. For instance, in a recent study, Zhang et al. (2017) claim that, for MaNGA Bundy et al. (2015) data, a  $\Sigma_{H\alpha} > \Sigma_{H\alpha}^{\text{SF,min}} = 10^{39}$  erg s $^{-1}$  kpc $^{-2}$  criterion selects reliable SF-dominated spaxels.

We prefer to distinguish SF from DIG on the basis of the equivalent width of H $\alpha$ . A simple thought experiment suffices to demonstrate that  $W_{H\alpha}$  offers a more suitable way to distinguish DIG from SF regions than  $\Sigma_{H\alpha}$ .

Consider combining the emission of two identical DIG-dominated cubic volume elements, each of area  $A$  and emitting a flux  $F_{H\alpha} = A \times \Sigma_{H\alpha}$ . As is appropriate for diffuse gas, suppose also the medium is optically thin to the H $\alpha$  photons, so that one sees the whole volume. Clearly, this DIG + DIG operation should not alter the identification of the nebular regime in the combined element, which should still be identified as DIG. Seen side by side, the joint surface brightness of the two elements will also be  $\Sigma_{H\alpha}$  (twice the flux over twice the area). If, however, the two cubes were to be placed along the same line of sight, one would see twice the H $\alpha$  flux over the same area  $A$ , and hence a doubling of the surface brightness. A  $\Sigma_{H\alpha}$ -based criterion will thus lead to a DIG + DIG = SF logical inconsistency if the original elements are brighter than half the chosen  $\Sigma_{H\alpha}^{\text{SF,min}}$  threshold. A  $W_{H\alpha}$ -based criterion, on the other hand, would lead to a consistent DIG classification independent of how the merged element is seen, as the final equivalent width is the same as that of the original elements.

As shown in Section 4.3, the difference between  $\Sigma_{H\alpha}$  –



**Figure 1.** SDSS stamps,  $\Sigma_{H\alpha}$ , and  $W_{H\alpha}$  maps for example galaxies in CALIFA. Images on the right show  $W_{H\alpha}$  maps saturated at 3 and 14 Å, highlighting the proposed classification of hDIG, mDIG, and SFc. Dashed elliptical rings mark radial distances to the nucleus of  $R = 0.5, 1.0, 1.5, \dots$  in units of the galaxy's half-light radius (HLR). Empty patches mask foreground sources and other artefacts.

and  $W_{\text{H}\alpha}$ -based criteria is particularly important in bulges, whose geometrically long path lengths may lead to  $\Sigma_{\text{H}\alpha} > \Sigma_{\text{H}\alpha}^{\text{SF,min}}$  even in the absence of SF.

A further and independent argument in favour of a  $W_{\text{H}\alpha}$ -based criterion is that properties such as colours, stellar mass density, gas content, and others show a radial dependence in galaxies, so a criterion based on a fixed threshold is likely not appropriate for all the regions of a galaxy. In particular, in the case where the DIG emission is powered by HOLMES, the approximately constant ratio of ionizing photons per unit mass of old stellar populations implies a  $W_{\text{H}\alpha}$  of the order of 1 Å (Binette et al. 1994; CF11; Belfiore et al. 2016), independently of the total line fluxes involved, so that one may well have  $\text{H}\alpha$ -bright HOLMES-powered regions (the hDIG) mistakenly classified as SF with a  $\Sigma_{\text{H}\alpha}$ -based criterion. Conversely, faint SF regions may be misclassified as DIG because of a low  $\Sigma_{\text{H}\alpha}$ .

In both arguments, it is the extensive nature of  $\Sigma_{\text{H}\alpha}$  that makes it prone to confusing DIG and SF regions.  $W_{\text{H}\alpha}$ , in contrast, is an intensive property<sup>2</sup>, and one that passes the two consistency tests posed above.

### 3.2 The observed distribution of $W_{\text{H}\alpha}$ and the hDIG component

Having made the case for a  $W_{\text{H}\alpha}$ -based characterization of DIG regions, this section presents empirical evidence that guides the implementation of  $W_{\text{H}\alpha}$  thresholds to separate SF from DIG, and to identify the HOLMES-dominated component.

The observed distribution of  $W_{\text{H}\alpha}$  among and within galaxies offers valuable insight on this issue. Fig. 2 shows  $W_{\text{H}\alpha}$  histograms for  $\sim 300k$  zones from the 391 CALIFA data cubes in our sample. The top panels show the result for the full sample, while the other rows split the sample by morphological types, from ellipticals in the second row to Sc and later in the bottom. Histograms in the left column count all zones, whereas columns to the right show results for different radial regions:  $R < 0.5$ ,  $0.5\text{--}1$ , and  $> 1$  HLR.

The global distribution (top-left panel) is strongly bimodal, with a low  $W_{\text{H}\alpha}$ -population peaking at  $W_{\text{H}\alpha} \sim 1$  Å and a higher one at  $\sim 14$  Å. This behaviour is remarkably similar to that seen in SDSS galaxies (Bamford et al. 2008; CF11). Previous spatially resolved studies based on both CALIFA (Morisset et al. 2016) and MaNGA data (Belfiore et al. 2016, 2017) have also identified this bimodality in  $W_{\text{H}\alpha}$ .

The relative amplitudes of the two peaks in the  $W_{\text{H}\alpha}$  distribution are sensitive to the spatial binning scheme. Without Voronoi binning, and restricting the analysis to the

<sup>2</sup> The qualifications ‘intensive’ and ‘extensive’ are used here in analogy with their thermodynamic connotation: intensive properties do not depend on the size or mass of a system (the column projected in a spaxel, in our case), whereas extensive properties are additive. The juxtaposition, either in 3D or just in projection, of two identical volume elements, each emitting a line flux  $F$  and a continuum flux density  $C$  over an area  $A$ , results in a  $W = 2F/2C$  equivalent width identical to that of the individual elements, so that  $W$  is an intensive property. On the other hand, one now has twice the flux exiting the same area, so that  $\Sigma = 2F/A$  equals the sum of the individual  $\Sigma$ ’s. In this sense, surface brightness behaves as an extensive property.

inner 2 HLR to eliminate the noisiest spectra, we find that the high- $W_{\text{H}\alpha}$  population at  $R > 1$  HLR increases by about a factor of 3 with respect to that seen in the top right panel of Fig. 2. This increase comes almost exclusively from galaxies of Hubble type Sb or later. The low- $W_{\text{H}\alpha}$  peak, on the other hand, changes by just  $\sim 20$  per cent. The bimodality, however, is preserved. Indeed, two-Gaussians fits to the pixel and zone-based distributions identify similar components, with peaks centred at  $\sim 1$  and 14 Å in both cases.

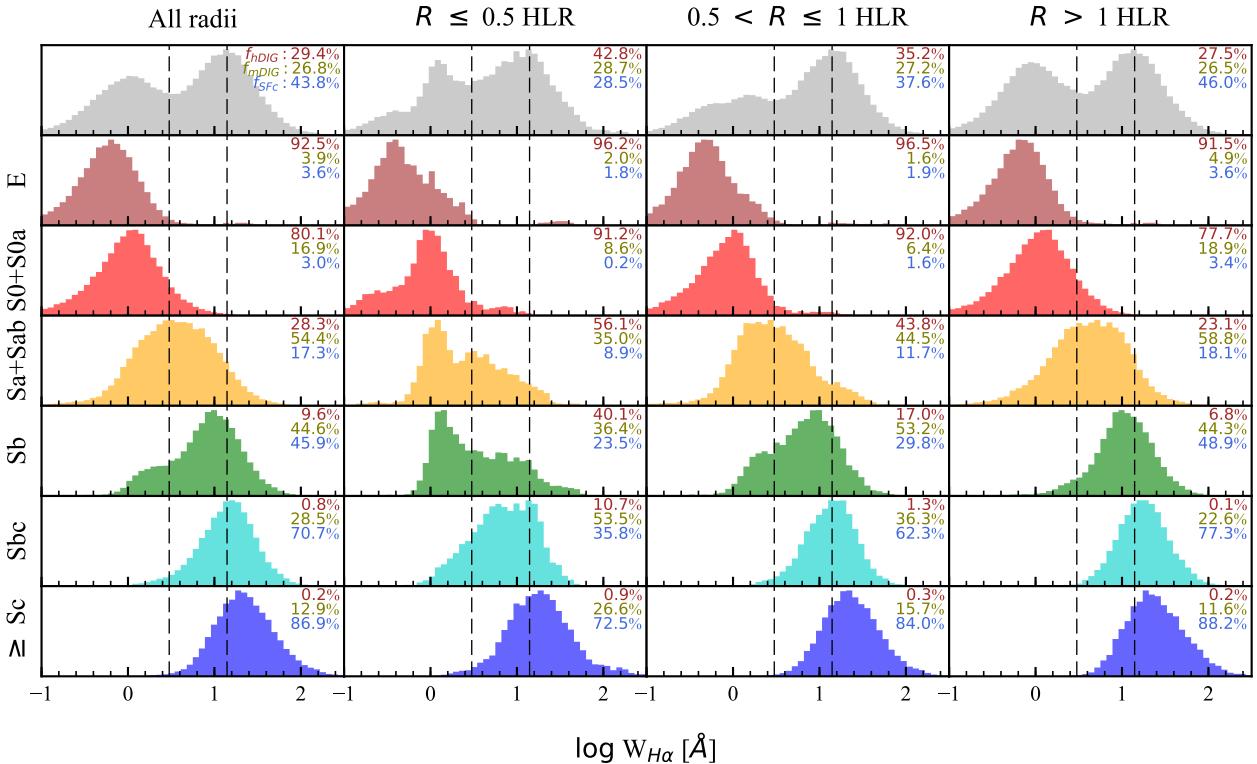
We interpret the low- $W_{\text{H}\alpha}$  population as representing gas photoionized by HOLMES. To test this, we followed the methodology in CF11 by computing the ratio  $\xi$  between the observed  $\text{H}\alpha$  luminosity and that predicted from the ionizing photons produced by populations older than  $10^8$  yr (inferred from the STARLIGHT analysis). Since the stellar population models used in our STARLIGHT fits (González Delgado et al. 2005; Vazdekis et al. 2010) do not extend to  $h\nu > 13.6$  eV, we have borrowed the number of ionizing photons from Bruzual & Charlot (2003) for a Salpeter initial mass function and Girardi et al. (2000) tracks. As discussed in CF11, different sets of models lead to systematic differences of 0.2–0.5 dex in the predicted ionizing fluxes. Fig. 3 shows  $\xi$  as a function of  $W_{\text{H}\alpha}$ , with histograms colour-coded by our hDIG/mDIG/SFc classification. We find that  $\xi$  is indeed of the order of 1 for zones in the low- $W_{\text{H}\alpha}$  peak. Hence, despite all uncertainties involved in this computation (CF11; Belfiore et al. 2016; Morisset et al. 2016), the end result corroborates the interpretation that HOLMES are responsible for the low- $W_{\text{H}\alpha}$  population.

The correspondence of this interpretation with the concept of retired galaxy put forth by Stasińska et al. (2008) is evident. These are systems that have stopped forming stars long ago and whose ionizing photon budget is dominated by hot post-asymptotic giant branch stars and white dwarfs, leading to  $W_{\text{H}\alpha}$  values of the order of  $\sim 1$  Å. Furthermore, the minimum seen at  $W_{\text{H}\alpha} \sim 3$  Å coincides with the threshold proposed by CF11 to distinguish retired galaxies from those where SF or AGN activity dominates the line emission. We thus claim that  $W_{\text{H}\alpha} < 3$  Å regions should be treated as HOLMES-ionized gas, the hDIG, a sub-type of DIG emission.

The breakdown of the  $W_{\text{H}\alpha}$  distribution by Hubble type in Fig. 2 suggests that the bimodality is always present, but the proportion of the low- to high- $W_{\text{H}\alpha}$  populations shifts with morphology: early-type galaxies are overwhelmingly dominated by values around the  $\sim 1$  Å peak, well within the hDIG regime, while in late-type spirals it is the higher  $W_{\text{H}\alpha}$  population that dominates.

The radial dependence of the  $W_{\text{H}\alpha}$  distribution offers further insight. The second to last columns in Fig. 2 show that the hDIG population is evenly spread in  $R$  for early-type galaxies, confirming earlier CALIFA-based studies by Kehrig et al. (2012), Singh et al. (2013), and Gomes et al. (2016b), as well as the MaNGA-based analysis by Belfiore et al. (2016, 2017). Among Sb and later type spirals, on the other hand, the hDIG is evidently concentrated in the central regions of galaxies. To put this in numbers, 82 per cent of the  $W_{\text{H}\alpha} < 3$  Å points in the 225 Sb or later type galaxies are located within  $R < 1$  HLR.

We interpret this higher incidence of hDIG zones in the central regions of galaxies as a corollary of the prevalence of old stellar populations in galaxy bulges. The ionizing photon



**Figure 2.** Distribution of  $W_{\text{H}\alpha}$  among 307 958 zones of 391 CALIFA galaxies. Different rows show the breakdown of this distribution by Hubble type, from ellipticals (second row) to Sc and later (bottom). Results for the whole sample are shown in the top row. Histograms in the left-hand panels count all zones, while the others select different ranges in radius: the inner 0.5 HLR (second column),  $R = 0.5\text{--}1$  HLR (third), and outwards of 1 HLR (fourth). Vertical dashed lines mark the hDIG/mDIG and mDIG/SF frontiers at 3 and 14 Å respectively. Labels on the top right of each panel list the fraction of the  $\text{H}\alpha$  flux associated with the hDIG, mDIG, and SF components (averaged over galaxies in each panel).

budget in these retired bulges is dominated by HOLMES, as any relevant contribution from other sources would raise their observed  $W_{\text{H}\alpha}$  to larger values.

Conversely, the low incidence of  $W_{\text{H}\alpha} < 3$  Å zones at large  $R$  in spirals indicates that pure hDIG emission is not a statistically relevant component of the DIG that permeates the space between SF regions in galaxy disks. This is not meant as a general conclusion, as Fig. 2 itself suggests that HOLMES may explain a substantial part of the disc emission in Sa–Sab galaxies. Among Sb and later type systems, however, hDIG-dominated disc zones are rare.

The motivation to introduce the hDIG category is thus firmly rooted on both observational and theoretical arguments. This well-understood component of the DIG becomes dominant whenever old stellar populations are the most relevant source of ionizing photons.

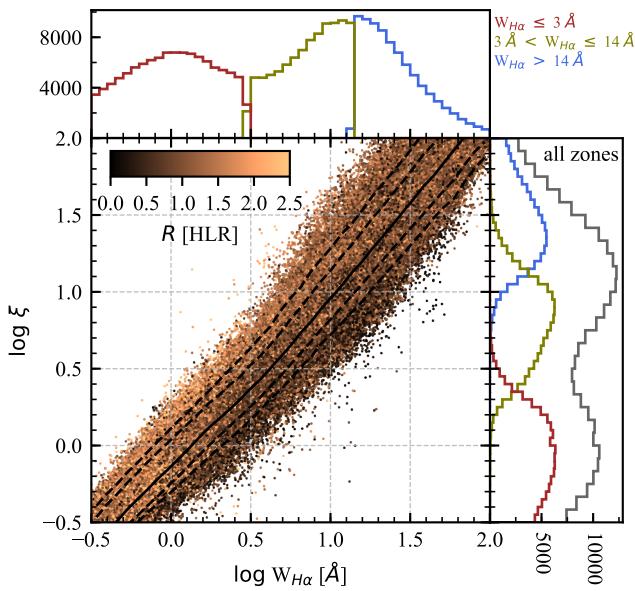
We close this section by noting that experiments were carried out to investigate to which extent galaxy inclination affects the  $W_{\text{H}\alpha}$  distributions depicted in Fig. 2. This was done by first eliminating E and S0's and then splitting the sample into bins in minor-to-major axis ratio  $b/a$  (measured as explained in de Amorim et al. 2017). No major effect is found. The only noteworthy trend identified is that for zones within  $R < 0.5$  HLR, the histograms tend to shift

towards smaller  $W_{\text{H}\alpha}$  (by  $\sim 0.2\text{--}0.3$  dex on the median) as one goes from edge-on to face-on viewing angles, a tendency that is understood as a simple projection effect. While face-on views of these inner regions sample the bulge, with its characteristic hDIG emission, as the inclination increases, parts of the disc and get projected upon the bulge, resulting in a mixture of hDIG and SF zones. Thus, as for the Voronoi binning effects discussed above, inclination effects do not erase the fundamental dichotomy between these two nebular regimes.

### 3.3 Identification of hDIG, mDIG, and SFc components

Unlike the low- $W_{\text{H}\alpha}$  regions, which can be safely associated with an hDIG regime, zones belonging to the high- $W_{\text{H}\alpha}$  population cannot be unequivocally tagged as SFc. Sure enough, SFc are among those with  $W_{\text{H}\alpha} > 3$  Å, but this population includes other processes too. In particular, diffuse gas powered by ionizing radiation leaking from H II regions is also part of this population, with a ratio of ionizing photons per unit stellar mass leading to  $W_{\text{H}\alpha}$  values above those typical of the HOLMES-dominated regime.

Though the whole  $W_{\text{H}\alpha} > 3$  Å population ultimately



**Figure 3.** Ratio between the observed H $\alpha$  luminosity and that predicted from populations older than  $10^8$  yr ( $\xi$ ) as a function of  $W_{H\alpha}$ . Points are colour-coded by the radial distance to the nucleus (in units of HLR). The histograms of  $\xi$  and  $W_{H\alpha}$  are colour-coded as hDIG/mDIG/SFc (red/yellow/blue), showing that low- $W_{H\alpha}$  regions are compatible with ionization by HOLMES ( $\xi \sim 1$ ).

comprises a mixture of regimes, it is useful to sub-divide it into mDIG and SFc classes in order to identify zones where SF is comparatively more important. There is no conspicuous boundary that cleanly separates SFc from mDIG in terms of  $W_{H\alpha}$ , however. Indeed, the continuous, unimodal behaviour of the  $W_{H\alpha} > 3$  Å population in Fig. 2 is not suggestive of sub-populations, but of a continuous distribution. In the absence of a clear-cut criterion, we place the mDIG/SFc division at  $W_{H\alpha} = 14$  Å, coinciding with the peak of the high  $W_{H\alpha}$  population in the histograms in Fig. 2.

Our final hDIG/mDIG/SFc classification scheme thus becomes

- hDIG:  $W_{H\alpha} < 3$ ,
- mDIG:  $3 < W_{H\alpha} < 14$ ,
- SFc:  $W_{H\alpha} > 14$ .

The reader should take note of a marked conceptual asymmetry in these definitions. While the mDIG/hDIG frontier at 3 Å is based on a firm theoretical understanding of the nature of the hDIG population, fully corroborated by the bimodal distribution of  $W_{H\alpha}$ , nothing of the sort can be said about the mDIG/SFc division. All that can be said about regions above the  $W_{H\alpha} = 14$  Å limit is that they have a higher proportion of SFc than those below it. This scheme should thus be used with the understanding that our mDIG regions may well host some star formation, and that  $W_{H\alpha} > 14$  Å does not isolate pure SF regions. Bona fide giant H II regions, those that are the basis of any emission-line study of galaxies, have  $W_{H\alpha}$  an order of magnitude larger (McCall et al. 1985; Garnett & Shields 1987; Kennicutt &

Garnett 1996; Luridiana & Peimbert 2001; Bresolin et al. 2004), but these are heavily diluted at our resolution.

The rightmost panels in Fig. 1 show  $W_{H\alpha}$  maps with a colour scheme designed to saturate at  $< 3$  and  $> 14$  Å. hDIG regions are thus depicted in red and SFc ones in blue, with intermediate colours used to trace the 3–14 Å mDIG range. The S0 galaxy at the top of the plot exemplifies the dominance of the hDIG component amongst early-type galaxies, as previously inferred from their  $W_{H\alpha}$  histograms in Fig. 2. The prevalence of this same component in bulges is also illustrated in this figure, particularly in the cases of CALIFA 0886 (NGC 7311) and 0010 (NGC 0036). As expected, SFc become increasingly important as one moves down the Hubble sequence (top to bottom in Figs. 1 and 2).

The reader is referred to Sánchez et al. (2015) for an example of a  $W_{H\alpha}$  map obtained under much higher spatial resolution. Their MUSE-image of the spiral galaxy NGC 6754 shows a plethora of SFc, embedded in a smoother medium with mDIG-like  $W_{H\alpha}$  values that permeates the whole disc.

## 4 DISCUSSION

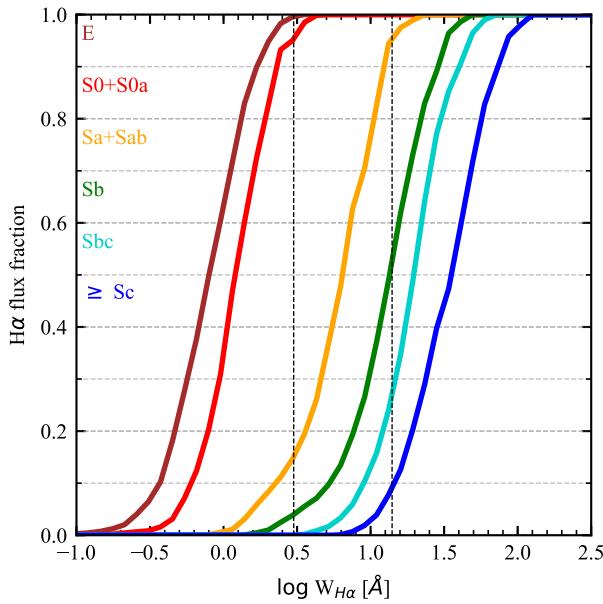
The theoretically and empirically inspired set of criteria to identify hDIG, mDIG, and SFc in galaxies serves a variety of purposes. In this section, we apply them to our CALIFA data with the specific goals of (i) estimating the relative strengths of hDIG, mDIG, and SFc components in galaxies across the Hubble sequence (Section 4.1), (ii) studying the nature of extraplanar diffuse line emission in edge-on systems (Section 4.2), (iii) comparing results obtained with our method with those derived with a  $\Sigma_{H\alpha}$ -based SF/DIG separation scheme (Section 4.3), (iv) investigating the possibility of differentiating SF and DIG regimes with density-sensitive line ratios (Section 4.4), (v) testing the consistency of our criteria with a classical diagnostic diagram analysis (Section 4.5), and (vi) investigating the mDIG mixture (Section 4.6). We close with a discussion of caveats involved (Section 4.7).

### 4.1 The relative strengths of the hDIG, mDIG, and SFc components

One of the questions that can be addressed with the classification proposed above is what are the relative strengths of the hDIG, mDIG, and SFc components, and how these proportions vary as a function of Hubble type. This issue bears on the interpretation of properties derived through spatially unresolved spectroscopic data, where these components cannot be separated, like for galaxies at high redshifts, for example.

A simple and observationally relevant way to quantify this is to compute the fractional contribution of each component to the total H $\alpha$  flux of a galaxy. For the galaxies in Fig. 1, for instance, these fractions range from  $(f_{\text{hDIG}}, f_{\text{mDIG}}, f_{\text{SFc}}) = (87, 13, 0)$  for the S0 galaxy CALIFA 0072, to  $(5.5, 47, 47.5)$  for the Sb galaxy CALIFA 0010, and  $(0.3, 46.1, 52.6)$  for CALIFA 0813, an Sbc. This steady progression along the Hubble sequence reflects the tendencies seen in Fig. 2, where the values of  $(f_{\text{hDIG}}, f_{\text{mDIG}}, f_{\text{SFc}})$  are given in each panel, along with the  $W_{H\alpha}$  histogram, for different radial regions and morphological types.

Fig. 4 presents these fractions for the whole sample in a



**Figure 4.** Cumulative fraction of the total galaxy H $\alpha$  flux coming from regions with  $W_{\text{H}\alpha}$  smaller than a given value. The plot shows the median curves obtained for galaxies in six Hubble types.

more elaborate way. For each galaxy, we compute the cumulative fraction ( $f$ ) of the total H $\alpha$  flux coming from zones with  $W_{\text{H}\alpha}$  smaller than a given value. The resulting  $f(<W_{\text{H}\alpha})$  growth curves yield not only the ( $f_{\text{hDIG}}$ ,  $f_{\text{mDIG}}$ ,  $f_{\text{SFc}}$ ) fractions, but a more continuous depiction of the makeup of a galaxy's H $\alpha$  output in terms of  $W_{\text{H}\alpha}$ . The figure shows the median curves obtained for each of our six bins in Hubble type. Vertical dashed lines mark the hDIG/mDIG and mDIG/SF $c$  frontiers at 3 and 14 Å respectively.

The steady progression from early to late types confirms the expectation from their  $W_{\text{H}\alpha}$  distributions (Fig. 2), and allows us to quantify the relative importance of each component to the total H $\alpha$  flux. Ellipticals and S0's have nearly all of their H $\alpha$  in the hDIG phase ( $W_{\text{H}\alpha} < 3 \text{ \AA}$ ). Among Sa–Sab systems, this component accounts for  $\sim 14\%$  of H $\alpha$ , with mDIG emission contributing most of the remaining flux. From Sb onwards, the SF $c$  fraction is 50 per cent or larger. Note that these fractions reflect the median behaviour. Naturally, there is substantial scatter from galaxy to galaxy even for a fixed morphological class.

DIG fractional contribution to H $\alpha$  fluxes has been estimated in several previous studies based on narrow-band ([H $\alpha$ ] + [N II]) imaging (Ferguson et al. 1996; Zurita et al. 2000; Thilker et al. 2002; Oey et al. 2007), with results varying substantially due to differences in the methodology to separate the DIG emission. The largest study to date is that by Oey et al. (2007), who estimate a mean DIG fraction of  $59 \pm 19\%$  among 109 galaxies in the SINGG survey (Meurer et al. 2006). For our sample and definitions, we find a very similar DIG (mDIG + hDIG) fraction of 56 per cent, but a larger dispersion ( $\pm 38$  per cent). Unlike in our study (Fig. 4), they find no evidence for a correlation with Hubble type, a difference that may be due to sample selection criteria and methodology to define DIG/SF.

## 4.2 Edge-on systems: the nature of extraplanar line emission

Edge-on galaxies are important for the study of the DIG, as they exhibit a systematic behaviour of emission-line properties with height above the galactic discs (Tüllmann & Dettmar 2000; Otte et al. 2002; Jones et al. 2017). The prototype galaxy is NGC 891, which has been extensively studied at all wavelengths (Rand 1998; Hodges-Kluck & Bregman 2013; Seon et al. 2014; Hughes et al. 2015). Many studies have emphasized that the observed properties of the extraplanar DIG in edge-on galaxies cannot be purely due to Lyman photons escaping from disc HII regions. A variety of suggestions have been put forward: dissipation of turbulence (Minter & Spangler 1997), magnetic reconnection (Raymond 1992), shocks (Collins & Rand 2001b), cosmic rays, photoelectric heating from interstellar dust grains (Weingartner & Draine 2001), and Lyman photons from old stars (Flores-Fajardo et al. 2011).

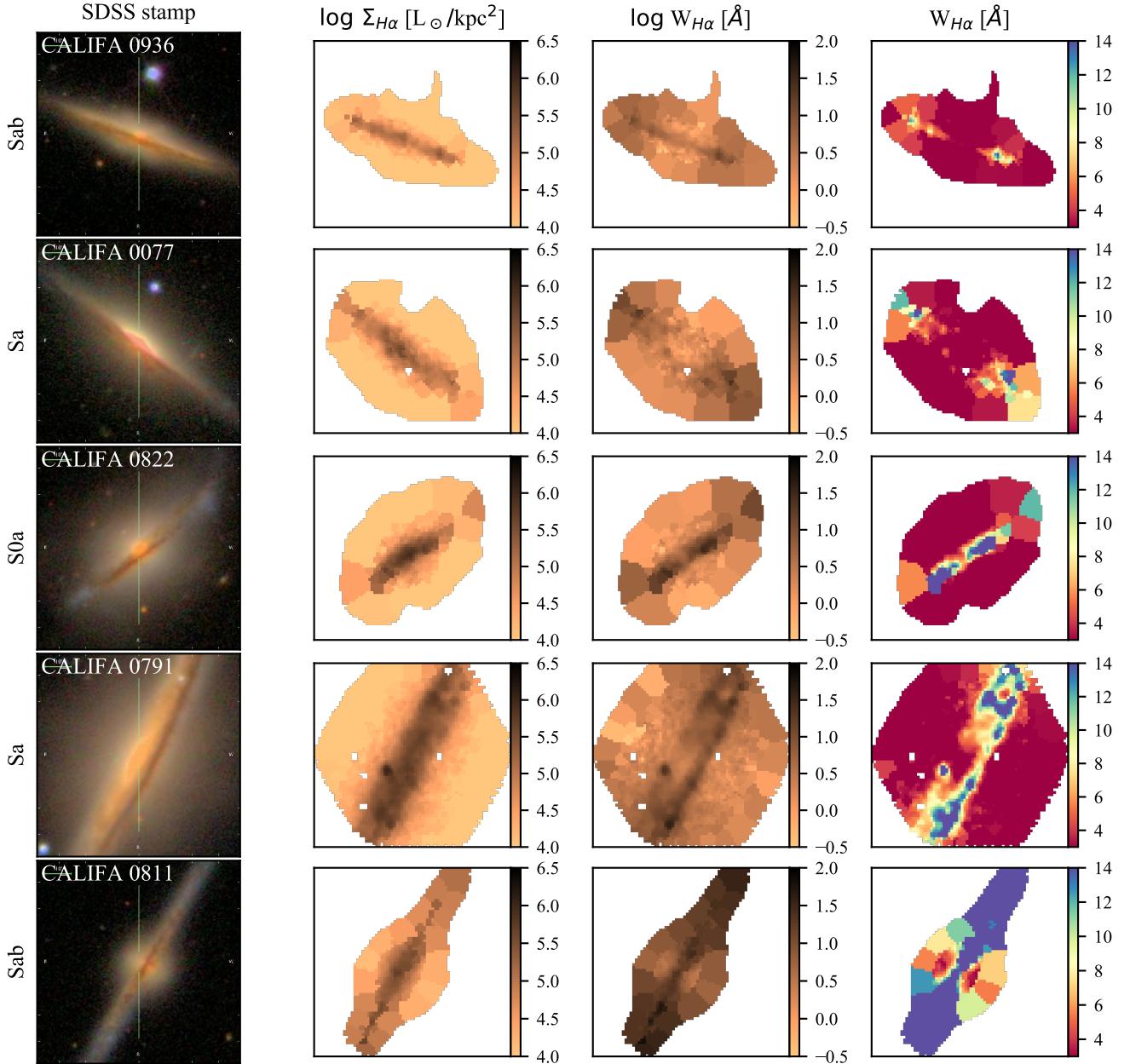
CALIFA data can bring new insight into this problem. Fig. 5 shows SDSS images and our CALIFA H $\alpha$  maps for five edge-on galaxies in our sample.<sup>3</sup> The layout is as in Fig. 1. The top four galaxies show a very similar structure in their  $W_{\text{H}\alpha}$  maps, with SF $c$  emission concentrated in the disc, where it is surrounded by mDIG. Above and below the plane, however, essentially all emission is hDIG-like. This provides strong support for the scenario of Flores-Fajardo et al. (2011) where the ionization becomes dominated by HOLMES at large galactic latitudes. Maps of standard diagnostic line ratios reinforce this conclusion, as illustrated in the MaNGA-based studies by Belfiore et al. (2016) and Zhang et al. (2017).

For the first time, we can relate this extraplanar DIG emission to the underlying stellar population. We do this by using the STARLIGHT results to compute the ratio  $\xi$  (see Section 3.2). For hDIG regions in the top four edge-on galaxies in Fig. 5, we find a median  $\xi$  value of 1.5, and an interquartile range from 1.1 to 1.9. Given the factor of  $\sim 2\text{--}3$  uncertainty in this estimate (CF11), the main conclusion here is that  $\xi$  is of the order of 1, and thus that the old stellar populations in these regions produce enough  $h\nu > 13.6 \text{ eV}$  photons to explain the observed extraplanar H $\alpha$  emission.

Face-on views of these same galaxies would project the extraplanar hDIG on top of a predominantly SF $c$  + mDIG disc. For a constant H $\alpha$  emissivity, the ratio of face-on to edge-on  $\Sigma_{\text{H}\alpha}$  equals the ratio  $h/r$  between height and radius of the extraplanar hDIG layer. In the edge-on examples in Fig. 5, the extraplanar hDIG has  $\Sigma_{\text{H}\alpha}$  of the order of a few times  $10^4 L_\odot \text{ kpc}^{-2}$ . For  $h \sim r$ , this is also the expected face-on surface brightness of this component. This is much smaller than the  $\Sigma_{\text{H}\alpha}$  of SF $c$  in the  $\sim$  face-on systems shown in Fig. 1, so that the projected extraplanar emission has a negligible effect. Some mDIG regions, however, have  $\Sigma_{\text{H}\alpha}$  values not much larger than  $10^4 L_\odot \text{ kpc}^{-2}$ , and thus could carry a non-negligible contribution from extra-planar hDIG.

The galaxy in the bottom row (CALIFA 0811, UGC

<sup>3</sup> While CALIFA 0077, 0936, and 0811 are very nearly edge-on, CALIFA 0822 and 0791 have inclinations of  $\sim 59^\circ$  and  $65^\circ$  respectively. The latter two galaxies are nevertheless still useful in this analysis as long as one considers locations far from the disc plane.

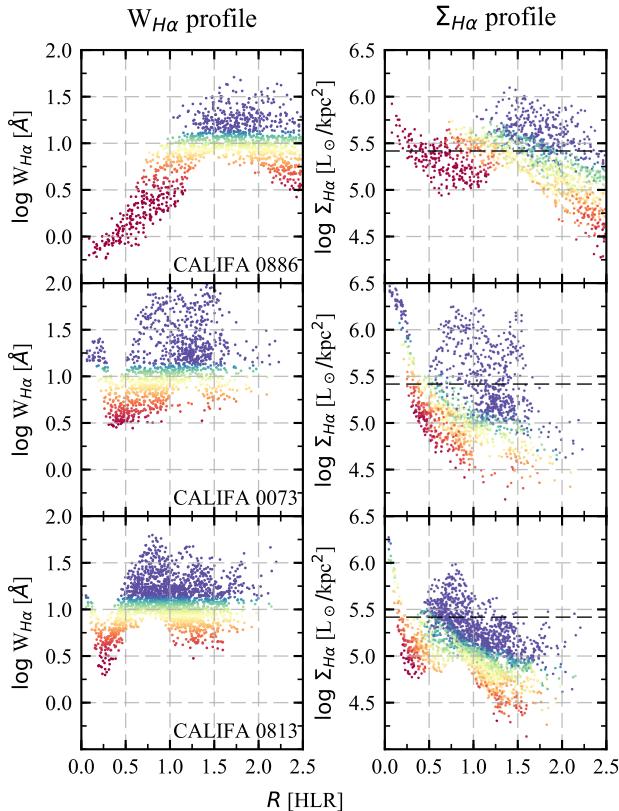


**Figure 5.** As Fig. 1, but for edge-on galaxies.

10043) differs from the others in Fig. 5 in having much more SFc along its disc, as well as SFc-like extra-planar emission over the disc, and a bipolar cone of intermediate  $W_{H\alpha}$  values centred in the nucleus. This galaxy has been recently studied by López-Cobá et al. (2017), who find the emission-line ratios and kinematics along the biconical structure to be consistent with a galactic wind powered by a central SF event. This combination of shock ionization and the widespread SF along its disc explains why there are so little signs of hDIG emission in this galaxy, although it is curious that  $W_{H\alpha}$  does drop close to hDIG values in the inner parts of the bicone.

#### 4.3 Comparison with $\Sigma_{H\alpha}$ -based SF/DIG separation schemes

Despite its conceptual advantages insofar as distinguishing different nebular regimes is concerned,  $W_{H\alpha}$  contains  $\Sigma_{H\alpha}$  in its numerator, so one may think that criteria based on these two variables may end up producing similar results. The maps in Fig. 1 show that structures like SF arms are indeed similarly traced by  $W_{H\alpha}$  and  $\Sigma_{H\alpha}$ , but others are not. Most notably,  $\Sigma_{H\alpha}$  always peaks in the central regions of galaxies while, for early-type galaxies,  $W_{H\alpha}$  shows clear dips.



**Figure 6.**  $W_{\text{H}\alpha}$  and  $\Sigma_{\text{H}\alpha}$  profiles for three of the example galaxies in Fig. 1. Points are coloured according to  $W_{\text{H}\alpha}$ . The dashed lines in the right-hand panels mark  $\Sigma_{\text{H}\alpha} = 10^{39} \text{ erg s}^{-1} \text{ kpc}^{-2}$ .

Fig. 6 examines this issue by means of radial profiles for CALIFA 0886, 0073, and 0813, three of the examples in Fig. 1. (For more examples of  $W_{\text{H}\alpha}$  profiles see Papaderos et al. 2013; Belfiore et al. 2016, 2017; Gomes et al. 2016b; González Delgado et al. 2016.) The left- and right-hand panels plot  $W_{\text{H}\alpha}$  and  $\Sigma_{\text{H}\alpha}$  against  $R$ , respectively. Both are coloured by  $W_{\text{H}\alpha}$  using the same colour scheme of previous plots.

CALIFA 0886 is a good example of the many galaxies that show low- $W_{\text{H}\alpha}$ , hDIG-dominated emission in their central regions, yet a peak in  $\Sigma_{\text{H}\alpha}$ . The reason why these retired bulges appear brighter than the surrounding disc is the much higher concentration of HOLMES in the bulge. This geometrically enhanced  $\text{H}\alpha$  emission can be erroneously attributed to SF with an SF/DIG criterion based on  $\Sigma_{\text{H}\alpha}$ . As seen in the top-right panel, the inner regions of this galaxy cross the  $\Sigma_{\text{H}\alpha}^{\text{SF,min}} = 10^{39} \text{ erg s}^{-1} \text{ kpc}^{-2} = 2.6 \times 10^5 L_{\odot} \text{ kpc}^{-2}$  threshold claimed to ‘select reliable H II region dominated spaxels’ according to Zhang et al. (2017). Yet, their  $W_{\text{H}\alpha}$  of  $\sim 1 \text{ \AA}$  are firmly in the hDIG regime. The  $W_{\text{H}\alpha}$  criterion thus correctly identifies the bulge of this and other galaxies as retired, while a  $\Sigma_{\text{H}\alpha}$  criterion would interpret them as bright, SF regions.

Throughout most of the disc of CALIFA 0886, the Zhang et al. (2017)  $\Sigma_{\text{H}\alpha}$ -based SF/DIG criterion agrees with the nebular regime identified via  $W_{\text{H}\alpha}$ . This agreement is only partial in CALIFA 0073 (central panels in Fig. 6), where we find more disc SF regions with  $W_{\text{H}\alpha}$  than with

$\Sigma_{\text{H}\alpha}$ , and even more so in CALIFA 0813 (bottom panels), where most  $W_{\text{H}\alpha} > 14 \text{ \AA}$  zones fall below the  $\Sigma_{\text{H}\alpha}^{\text{SF,min}}$  threshold. These differences stem from the contrasting radial behaviours of  $\Sigma_{\text{H}\alpha}$  and  $W_{\text{H}\alpha}$ . For a fixed  $W_{\text{H}\alpha}$ , the highest  $\Sigma_{\text{H}\alpha}$ ’s tend to be located in central regions, while, in contrast, for a fixed  $\Sigma_{\text{H}\alpha}$ , the largest values of  $W_{\text{H}\alpha}$  are generally found in the outskirts. Indeed, as seen in the examples in Fig. 6,  $\Sigma_{\text{H}\alpha}$  tends to decrease outwards while  $W_{\text{H}\alpha}$  remains roughly constant, both with large dispersions at any given  $R$  in the disc. About 37 per cent of our SFc spaxels have  $\Sigma_{\text{H}\alpha} < 10^{39} \text{ erg s}^{-1} \text{ kpc}^{-2}$ . On average, these faint SF regions are at  $R = 1.3$  HLR from the centre.

In summary, compared to the hDIG/mDIG/SFc separation criteria proposed in this paper, a  $\Sigma_{\text{H}\alpha}$ -based criterion tends to overestimate the population of SF regions at low  $R$  and underestimate it at large  $R$ . More worryingly, as already mentioned,  $\Sigma_{\text{H}\alpha}$  by itself cannot identify the hDIG component, the main source of  $\text{H}\alpha$  emission in old spheroids. In fact, we have seen that retired bulges may be mistaken for their very opposite when bright enough to exceed  $\Sigma_{\text{H}\alpha}^{\text{SF,min}}$ .

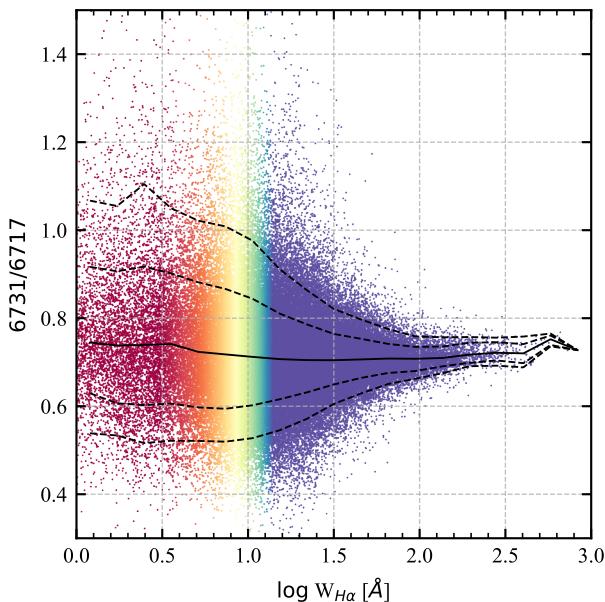
Previous CALIFA-based studies by Kehrig et al. (2012), Singh et al. (2013), and Gomes et al. (2016b) also find  $\Sigma_{\text{H}\alpha}$  values in excess of the  $\Sigma_{\text{H}\alpha}^{\text{SF,min}}$  limit of Zhang et al. (2017) in the inner regions of early-type galaxies, where the absence of young stars is beyond dispute (see also Sarzi et al. 2010 for results based on the SAURON survey). These examples are the observational realization of the DIG + DIG = SF conceptual inconsistency pointed out in Section 3.1. The  $W_{\text{H}\alpha}$ -based scheme presented in this paper solves this problem by extending to spatially resolved data the concept of retired galaxy proposed by Stasińska et al. (2008) and CF11 in the context of integrated galaxy spectra.

#### 4.4 Can the DIG be detected with density-sensitive line ratios?

The electron densities in the Milky Way DIG obtained from dispersion measures and H I column densities towards pulsars at known distances are typically less than  $10^{-1} \text{ cm}^{-3}$  (Berkhuijsen & Fletcher 2008), orders of magnitude smaller than those of H II regions. One could then expect that the density-sensitive  $[\text{S II}] \lambda\lambda 6731, 6716$  line ratios in our data would indicate smaller densities in the DIG than in SFc, and show a tendency with  $W_{\text{H}\alpha}$ . Fig. 7 plots the  $[\text{S II}] 6731/6716$  flux ratio as a function of  $W_{\text{H}\alpha}$  (only zones where  $SN > 3$  in the  $[\text{S II}]$  ratio are plotted). The median and the  $\pm 1\sigma$  and  $2\sigma$  percentiles curves are overplotted. We find no trend of  $[\text{S II}] 6731/6716$  with  $W_{\text{H}\alpha}$ . The increase in scatter towards lower  $W_{\text{H}\alpha}$  is consistent with the corresponding decrease in the  $SN$  of the lines. In essence,  $[\text{S II}] 6731/6716$ , in regions where this ratio can be safely measured, is everywhere consistent with its low density limit of 0.7. Of course, this figure cannot say anything about regions where  $SN < 3$ .

The reason why the  $[\text{S II}]$  line ratio is not substantially different in the DIG zones where it can be measured and in the SFc is twofold: first, at our resolution, SFc contain a significant amount of diffuse gas. Secondly, the  $[\text{S II}]$  line ratio is not sensitive to densities below  $\sim 50 \text{ cm}^{-3}$ .

One could expect to obtain a different picture if using a line doublet sensitive to lower electron densities, like the far-infrared doublet  $[\text{N II}] \lambda\lambda 205, 125 \mu\text{m}$ . Recent observations have allowed one to map this ratio in the Milky Way and



**Figure 7.**  $[\text{S II}]\lambda\lambda 6731/6716$  flux ratio for 111760 zones in our sample where this ratio has  $SN > 3$ . Points are coloured by  $W_{\text{H}\alpha}$  as in previous figures. The solid line traces the median curve, and dashed lines show the  $1\sigma$  and  $2\sigma$  equivalent percentiles.

several galaxies (Goldsmith et al. 2015; Herrera-Camus et al. 2016). The derived densities range between 1 and  $300 \text{ cm}^{-3}$ , with a median value of  $30 \text{ cm}^{-3}$ , and do not get anywhere close to the DIG densities obtained from pulsar measurements.

Classical density estimators are thus not able to detect the DIG, at least not at the resolution of CALIFA and similar surveys. However, it is not clear whether they would do a better job at higher spatial resolution, since, as noted by Rubin (1989) in another context, density inhomogeneities severely hamper any quantitative interpretation of such density-sensitive line ratios.

#### 4.5 $W_{\text{H}\alpha}$ and the BPT diagram

The different ionization and heating conditions in hDIG, mDIG, and SFc should lead to different collisional-to-recombination line flux ratios, and thus to different loci on diagnostic diagrams like  $[\text{O III}]\lambda 5007/\text{H}\beta$  versus  $[\text{N II}]\lambda 6584/\text{H}\alpha$ . This famous BPT diagram (after Baldwin, Phillips & Terlevich 1981) is widely used to separate SF galaxies from those where harder ionizing sources contribute significantly to the ionization of the gas. It is thus legitimate to ask whether our  $W_{\text{H}\alpha}$ -based scheme is consistent with this independent way of characterizing the nebular regime.

Fig. 8 shows BPT diagrams obtained from all zones in our sample where  $SN > 3$  in all four lines involved. The top-left panel shows results for the entire sample, colouring points according to  $W_{\text{H}\alpha}$  following the scheme used for the right-hand panels in Fig. 1, which saturates at  $> 14 \text{ \AA}$  (blue) and  $< 3 \text{ \AA}$  (red). The curves correspond to the demarcation lines proposed by Stasińska et al. (2006, S06), Kauffmann

et al. (2003, K03), and Kewley et al. (2001, K01) – see CF11 for a discussion of the meaning of these curves.

The strong correspondence between  $W_{\text{H}\alpha}$  and the BPT coordinates is evident to the eye, as previously noted by Morisset et al. (2016) for CALIFA data and Belfiore et al. (2016) for MaNGA. The left wing is predominantly populated by SFc, while hDIG regions populate the tip of the right wing, with mDIG regions in between. This correspondence is further confirmed in the right-hand panels in Fig. 8, which separate SFc, mDIG, and hDIG zones according to our criterion, with points coloured by their radial distance to the nucleus. Deviant points are concentrated in the inner regions of galaxies (red–orange points), where some high- $W_{\text{H}\alpha}$  SFc or mDIG zones intrude into a region in the BPT diagram otherwise dominated by hDIG emission. These outliers come from AGN in our sample, as discussed in Section 4.7.

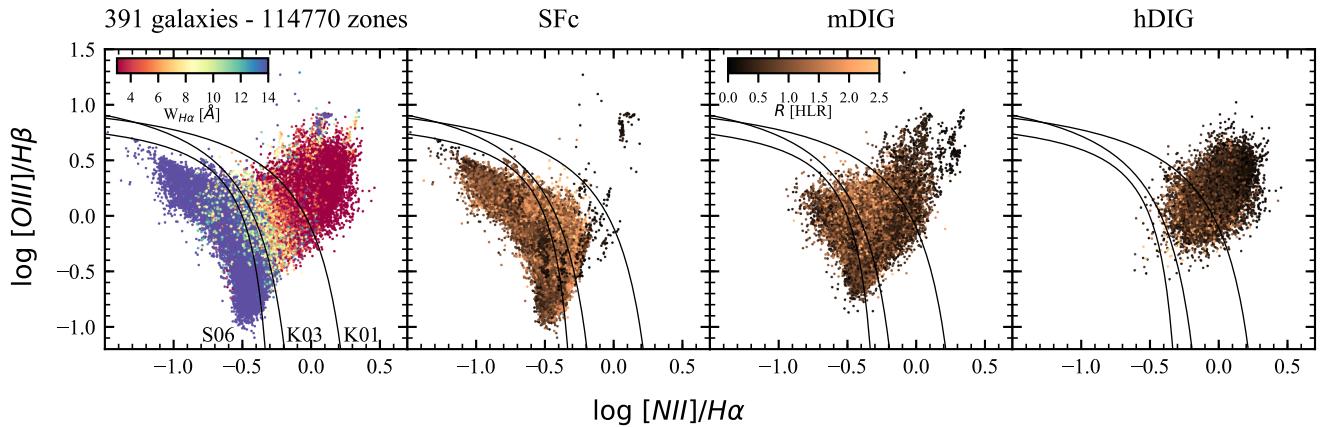
Fig. 8 thus makes a visually compelling case for our  $W_{\text{H}\alpha}$ -based SFc/mDIG/hDIG separation scheme. Translating this impression to numbers, and restricting the analysis to  $R > 1 \text{ HLR}$ , 58 (92) per cent of our  $W_{\text{H}\alpha} > 14 \text{ \AA}$  zones are in the SF region of the BPT diagram according to the S06 (K03) criterion. The relatively large number of zones trespassing the S06 line is not surprising if one recalls that this curve was derived on the basis of photoionization models designed to establish the boundaries of *pure* SF regions in the BPT plane. As reiterated several times in this paper, at the resolution of CALIFA, our SFc are nowhere near pure H II regions, but include plenty of diffuse emission, which enhances both line ratios in the BPT diagram.

We thus conclude that our SFc/mDIG/hDIG separation scheme leads to diagnostic line ratios that are qualitatively consistent with what one expects for these nebular regimes. In conjunction with the conceptual and empirical arguments presented in Section 3, this adds further strength to our methodology.

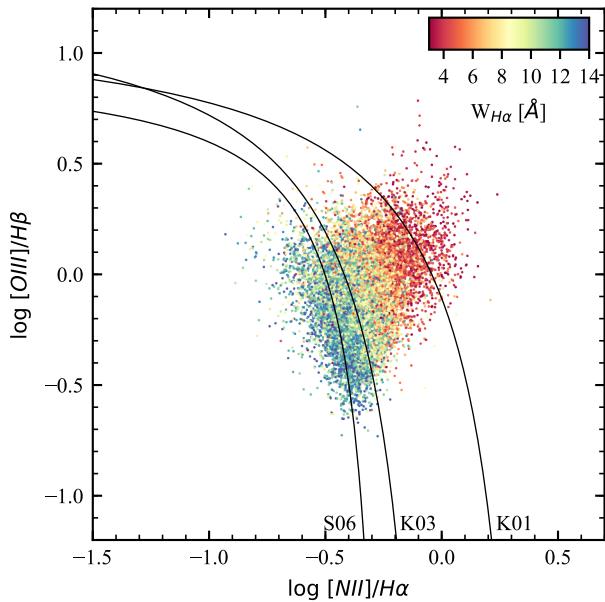
#### 4.6 The mDIG as an SF+hDIG mixture

Fig. 8 shows that mDIG zones straddle the region between the classical SF wing and the loci of hDIG zones in the BPT diagram, a behaviour that lends itself to an interpretation in terms of a mixture of SFc and hDIG emission. This behaviour is better appreciated in Fig. 9, which shows the mDIG points coloured according to  $W_{\text{H}\alpha}$ . To mitigate potential contamination by AGN-affected zones, only points at  $R > 1 \text{ HLR}$  ( $= 5.3 \text{ kpc}$  on average) are shown.

The same steady progression in  $W_{\text{H}\alpha}$  along the right wing seen in the top-left panel of Fig. 8 is also seen in Fig. 9, suggesting that the mDIG component may be described as a mixture of SFc and hDIG emission. That the mDIG is interpretable as a mixture of processes is of course just what one would expect. The large population of mDIG zones approaching the  $W_{\text{H}\alpha} = 14 \text{ \AA}$  threshold and whose BPT coordinates overlap with those of SFc probably correspond to the leakage scenario. Above the SF wing in the BPT diagram, the main ionization process is likely still due to young, massive stars, but the influence of the heating due to HOLMES increases gradually as  $W_{\text{H}\alpha}$  decreases. When  $W_{\text{H}\alpha}$  approaches the  $< 3 \text{ \AA}$  hDIG regime, the ionizing field of HOLMES starts dominating the photoionization budget too.



**Figure 8.** BPT diagrams for our sample. The left-hand panel shows the full sample, with colours coding for  $W_{\text{H}\alpha}$ , as indicated. Other panels split the sample into SFc ( $W_{\text{H}\alpha} > 14 \text{ \AA}$ ), mDIG ( $W_{\text{H}\alpha} = 3-14 \text{ \AA}$ ), and hDIG ( $W_{\text{H}\alpha} < 3 \text{ \AA}$ ) regions, colouring points according to their radial distance  $R$  (in HLR units). In all cases, only zones with  $SN > 3$  in all lines are plotted. Dividing curves come from (from left to right) Stasińska et al. (2006, S06), Kauffmann et al. (2003, K03), and Kewley et al. (2001, K01).



**Figure 9.** BPT diagram for mDIG regions only (i.e., those with  $W_{\text{H}\alpha}$  in the 3–14 Å range), coloured according to  $W_{\text{H}\alpha}$ , and excluding zones inwards of  $R = 1$  HLR.

#### 4.7 AGN and other caveats

Our whole hDIG/mDIG/SFc classification scheme ignores other mechanisms of line production, most notably AGN. AGN are found in the central parts of some galaxies and recognized as such through the BPT diagram. For example, the cluster of blue ( $W_{\text{H}\alpha} > 14 \text{ \AA}$ ) points at the tip of the right wing of the BPT in Fig. 8, at coordinates  $\sim (0.2, 0.8)$  come from the central regions of CALIFA 0897 (UGC12348), a known type 2 Seyfert (Cusumano et al. 2010; Asmus et al. 2014). Other  $W_{\text{H}\alpha} > 3 \text{ \AA}$  outliers in BPT loci otherwise

dominated by hDIG zones also tend to be located at small  $R$  (reddish points in the central panels of Fig. 8).

AGN may also power line emission well outside the nucleus (up to distances as large as 20 kpc; Veilleux et al. 2003). These are the so-called extended emission-line regions (EELRs) or ionization cones. They can be due either to photoionization by X-ray photons leaving the AGN with a small opening angle or to an interaction between radio jets and the galaxy ISM producing strong shocks (Wilson 1996). However, in the framework of the present study, which is to evaluate the importance of the DIG in galaxies and pinpoint its different regimes, EELRs in Seyfert galaxies are a secondary issue, as they affect only specific zones of galaxies with a well-defined AGN – and perhaps not all of them. Understanding the EELRs is a topic in itself, which indeed can be tackled with sensitive 3D spectroscopy, and some recent studies already started doing so (e.g., Dopita et al. 2014), but it is outside the scope of the present paper.

Another line-producing process neglected in this study is shocks. In the case of the galactic wind in CALIFA 0811, shown in Fig. 5, we find  $W_{\text{H}\alpha} = 3-12 \text{ \AA}$  in the shocked region, i.e., mDIG-like values. Again,  $W_{\text{H}\alpha}$  by itself cannot identify the shock origin of the nebular emission, though it at least tells that photoionization by HOLMES is not an energetically feasible explanation. Only a detailed study of the geometry, line ratios, and kinematics can fully reveal the processes governing line emission in objects like this (Kreckel et al. 2014; Beirão et al. 2015; López-Cobá et al. 2017).

Because of their relative rarity and spatial constraints, these processes do not affect much the hDIG/mDIG/SFc statistics found in this study. They should nevertheless be taken into consideration in studies of individual sources.

Finally, a word of caution about the so-called composite region in the BPT diagram, commonly defined as the zone below the K01 line and above the K03 or S06 line. Though this region is usually thought to correspond to SF + AGN mixtures, AGN and hDIG have indistinguishable line ratios. It is therefore not a priori clear what this composite region is composed of!

The way to break this degeneracy is through  $W_{\text{H}\alpha}$ . Given that (unlike AGN or SF) old stars are everywhere in galaxies, one should understand the hDIG regime as a floor level of ionization, one that is only energetically relevant when no other source is. Whenever the stellar continuum around H $\alpha$  is dominated by old stars (which is the case even in SF regions at kpc resolution), the direct scaling between the optical continuum and the ionizing fluxes of the corresponding HOLMES population leads to  $W_{\text{H}\alpha} \sim 1\text{--}2 \text{ \AA}$  according to current evolutionary population synthesis models (see CF11 and references therein). Thus, composite spectra where  $W_{\text{H}\alpha}$  approaches this limit likely represent an SFc + hDIG mixture. Conversely, as long as  $W_{\text{H}\alpha}$  is above the hDIG range, an SF + AGN mixture is more plausible.

The Davies et al. (2014) study illustrates this point. Using the CALIFA data cubes of two type 2 Seyferts (NGC 2410 and NGC 6394) and two other more ambiguous (Seyfert–LINER) cases (IC 0540 and NGC 6762), they identify approximately one-dimensional distributions in the BPT and other diagnostic diagrams, suggestive of SF + AGN mixing sequences. They however verified that in NGC 6762 the majority ( $> 90\%$ ) of the spaxels have  $W_{\text{H}\alpha} < 3 \text{ \AA}$ , so that the contribution of HOLMES cannot be ignored. NGC 2410 and NGC 6394, on the other hand, have central  $W_{\text{H}\alpha}$  values well above the hDIG range, which makes the SF + AGN interpretation plausible. IC0540 has  $W_{\text{H}\alpha}$  values in the mDIG-to-hDIG range, so that the interpretation is less clear, although they favour an SF + AGN scenario. The bottom line here is  $W_{\text{H}\alpha}$  should be taken into account when studying composite systems in order to avoid confusing hDIG and AGN effects.

## 5 SUMMARY

We have used data cubes from the CALIFA survey to investigate the origin of the line emission in over 300k zones of 391 galaxies across the Hubble sequence. Studies based on integrated galaxy data like the SDSS give great emphasis on determining whether or not a galaxy hosts an AGN. In a sample of spaxel spectra like ours, however, a more relevant question is whether the ionization is locally dominated by photons arising from massive stars (either in bona fide H II regions or in diffuse regions surrounding them) or by photons arising from old stellar populations. This has been the focus of our study.

We have shown that the commonly adopted method to separate DIG from SF regions on the basis of the surface brightness of H $\alpha$  ( $\Sigma_{\text{H}\alpha}$ ) is conceptually flawed, and that the H $\alpha$  equivalent width ( $W_{\text{H}\alpha}$ ) is more suitable to distinguish the qualitative differences between these regimes. Moreover, and perhaps more importantly,  $W_{\text{H}\alpha}$  further allows us to confidently identify the cases where the line emission is predominantly powered by HOLMES, an omnipresent population that provides a floor level of ionizing radiation in galaxies.

The observed distribution of  $W_{\text{H}\alpha}$  within and among galaxies was used to propose a classification scheme. Zones where  $W_{\text{H}\alpha} < 3 \text{ \AA}$  are attributed to a HOLMES-ionized gas (hDIG) component, responsible for the  $W_{\text{H}\alpha} \sim 1 \text{ \AA}$  peak in the strongly bimodal distribution of  $W_{\text{H}\alpha}$ . This observational definition of hDIG is identical to the one for retired

galaxies. In the absence of any (astro)physically motivated argument, we have defined the concept of SFc as regions where  $W_{\text{H}\alpha} > 14 \text{ \AA}$ , the mode of the high- $W_{\text{H}\alpha}$  population in our sample. Regions where  $W_{\text{H}\alpha}$  falls in the 3–14  $\text{\AA}$  intermediate range are tagged as having a mixed nature (mDIG).

Some of the main results obtained with this empirically and theoretically motivated scheme are as follows.

(i) In agreement with their predominantly old stellar populations, the hDIG is the main nebular regime in early-type galaxies (E and S0) and in bulges.

(ii) The extraplanar emission in edge-on spirals is also of hDIG type, vindicating the scenario put forward and elaborated by Flores-Fajardo et al. (2011). In face-on systems, this extraplanar hDIG makes a negligible contribution when projected over SFc, but a potentially relevant one when projected on to regions classified as mDIG.

(iii) A  $\Sigma_{\text{H}\alpha}$ -based SF/DIG separation scheme tends to classify hDIG-dominated retired bulges as SF, an inconsistency that is ultimately due to the extensive nature of  $\Sigma_{\text{H}\alpha}$ . In addition, faint SF regions tend to be misclassified as DIG with a  $\Sigma_{\text{H}\alpha}$  criterion.

(iv) The hDIG, mDIG, and SFc contributions to the H $\alpha$  luminosity vary in a systematic way along the Hubble sequence, ranging from (100, 0, 0) per cent in ellipticals and S0's to (9, 60, 31) per cent in Sa–Sb's and (0, 13, 87) per cent in later types.

(v) SFc and hDIG regions occupy characteristic loci on the BPT diagram. mDIG regions form a continuous sequence between SFc and hDIG line ratios, indicative of an mDIG = SFc + hDIG mixture.

## ACKNOWLEDGEMENTS

EADL, RCF, GS, and NVA acknowledge the support from the CAPES CsF-PVE project 88881.068116/2014-01. RGD acknowledges the support of CNPq (Brazil) through Programa Ciéncia sem Fronteiras (401452/2012-3). CALIFA is the first legacy survey carried out at Calar Alto. The CALIFA collaboration would like to thank the IAA-CSIC and MPIA-MPG as major partners of the observatory, and CAHA itself, for the unique access to telescope time and support in manpower and infrastructures. We also thank the CAHA staff for the dedication to this project. Support from the Spanish Ministerio de Economía y Competitividad, through projects AYA2016-77846-P, AYA2014-57490-P, AYA2010-15081, and Junta de Andalucía FQ1580.

## REFERENCES

- Asmus D., Hönig S. F., Gandhi P., Smette A., Duschl W. J., 2014, *MNRAS*, **439**, 1648
- Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, **93**, 5
- Bamford S. P., Rojas A. L., Nichol R. C., Miller C. J., Wasserman L., Genovese C. R., Freeman P. E., 2008, *MNRAS*, **391**, 607
- Barrera-Ballesteros J. K., et al., 2015a, *A&A*, **579**, A45
- Barrera-Ballesteros J. K., et al., 2015b, *A&A*, **582**, A21
- Beirão P., et al., 2015, *MNRAS*, **451**, 2640
- Belfiore F., et al., 2016, *MNRAS*, **461**, 3111
- Belfiore F., et al., 2017, *MNRAS*, **466**, 2570
- Berkhuijsen E. M., Fletcher A., 2008, *MNRAS*, **390**, L19

- Binette L., Magris C. G., Stasińska G., Bruzual A. G., 1994, *A&A*, **292**, 13
- Binette L., Drissen L., Ubeda L., Raga A. C., Robert C., Krongold Y., 2009, *A&A*, **500**, 817
- Bresolin F., Garnett D. R., Kennicutt Jr. R. C., 2004, *ApJ*, **615**, 228
- Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, **351**, 1151
- Bruzual G., Charlot S., 2003, *MNRAS*, **344**, 1000
- Bundy K., et al., 2015, *ApJ*, **798**, 7
- Cid Fernandes R., Mateus A., Sodré L., Stasińska G., Gomes J. M., 2005, *MNRAS*, **358**, 363
- Cid Fernandes R., Asari N. V., Sodré L., Stasińska G., Mateus A., Torres-Papaqui J. P., Schoenell W., 2007, *MNRAS*, **375**, L16
- Cid Fernandes R., Stasińska G., Mateus A., Vale Asari N., 2011, *MNRAS*, **413**, 1687
- Cid Fernandes R., et al., 2013, *A&A*, **557**, A86
- Collins J. A., Rand R. J., 2001a, *ApJ*, **551**, 57
- Collins J. A., Rand R. J., 2001b, *ApJ*, **551**, 57
- Cortijo-Ferrero C., et al., 2017a, preprint, ([arXiv:1706.01896](https://arxiv.org/abs/1706.01896))
- Cortijo-Ferrero C., et al., 2017b, *MNRAS*, **467**, 3898
- Cusumano G., et al., 2010, *A&A*, **524**, A64
- Davies R. L., Kewley L. J., Ho I.-T., Dopita M. A., 2014, *MNRAS*, **444**, 3961
- Dettmar R.-J., 1990, *A&A*, **232**, L15
- Doe S., et al., 2007, in Shaw R. A., Hill F., Bell D. J., eds, Astronomical Society of the Pacific Conference Series Vol. 376, Astronomical Data Analysis Software and Systems XVI. p. 543
- Dopita M. A., et al., 2014, *A&A*, **566**, A41
- Ferguson A. M. N., Wyse R. F. G., Gallagher III J. S., Hunter D. A., 1996, *AJ*, **111**, 2265
- Flores-Fajardo N., Morisset C., Stasińska G., Binette L., 2011, *MNRAS*, **415**, 2182
- Freeman P., Doe S., Siemiginowska A., 2001, in Starck J.-L., Murtagh F. D., eds, Proc. SPIE Vol. 4477, Astronomical Data Analysis. pp 76–87 ([arXiv:astro-ph/0108426](https://arxiv.org/abs/astro-ph/0108426)), doi:10.1117/12.447161
- García-Benito R., et al., 2015, *A&A*, **576**, A135
- García-Benito R., et al., 2017, preprint, ([arXiv:1709.00413](https://arxiv.org/abs/1709.00413))
- Garnett D. R., Shields G. A., 1987, *ApJ*, **317**, 82
- Girardi L., Bressan A., Bertelli G., Chiosi C., 2000, *A&AS*, **141**, 371
- Goldsmith P. F., Yıldız U. A., Langer W. D., Pineda J. L., 2015, *ApJ*, **814**, 133
- Gomes J. M., et al., 2016a, *A&A*, **586**, A22
- Gomes J. M., et al., 2016b, *A&A*, **588**, A68
- González Delgado R. M., Pérez E., 2000, *MNRAS*, **317**, 64
- González Delgado R. M., Cerviño M., Martins L. P., Leitherer C., Hauschildt P. H., 2005, *MNRAS*, **357**, 945
- González Delgado R. M., et al., 2014, *ApJ*, **791**, L16
- González Delgado R. M., et al., 2015, *A&A*, **581**, A103
- González Delgado R. M., et al., 2016, *A&A*, **590**, A44
- González Delgado R. M., et al., 2017, *A&A*, **607**, A128
- Haffner L. M., et al., 2009, *Reviews of Modern Physics*, **81**, 969
- Herrera-Camus R., et al., 2016, *ApJ*, **826**, 175
- Hodges-Kluck E. J., Bregman J. N., 2013, *ApJ*, **762**, 12
- Hoopes C. G., Walterbos R. A. M., 2003, *ApJ*, **586**, 902
- 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
- Hughes T. M., et al., 2015, *A&A*, **575**, A17
- Husemann B., et al., 2013, *A&A*, **549**, A87
- Jones A., et al., 2017, *A&A*, **599**, A141
- Kauffmann G., et al., 2003, *MNRAS*, **341**, 54
- Kehrig C., et al., 2012, *A&A*, **540**, A11
- Kennicutt Jr. R. C., Garnett D. R., 1996, *ApJ*, **456**, 504
- Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, **556**, 121
- Kreckel K., et al., 2014, *ApJ*, **790**, 26
- López-Cobá C., et al., 2017, *MNRAS*, **467**, 4951
- Luridiana V., Peimbert M., 2001, *ApJ*, **553**, 633
- McCall M. L., Rybski P. M., Shields G. A., 1985, *ApJS*, **57**, 1
- Meurer G. R., et al., 2006, *ApJS*, **165**, 307
- Minter A. H., Spangler S. R., 1997, *ApJ*, **485**, 182
- Morisset C., et al., 2016, *A&A*, **594**, A37
- Oey M. S., et al., 2007, *ApJ*, **661**, 801
- Otte B., Gallagher III J. S., Reynolds R. J., 2002, *ApJ*, **572**, 823
- Papaderos P., et al., 2013, *A&A*, **555**, L1
- Pérez E., et al., 2013, *ApJ*, **764**, L1
- Rand R. J., 1998, *ApJ*, **501**, 137
- Raymond J. C., 1992, *ApJ*, **384**, 502
- Reynolds R. J., 1971, PhD thesis, THE UNIVERSITY OF WISCONSIN - MADISON.
- Reynolds R. J., Sterling N. C., Haffner L. M., Tufte S. L., 2001, *ApJ*, **548**, L221
- Rousseau-Nepton L., Robert C., Drissen L., Martin R. P., Martin T., 2017, preprint, ([arXiv:1704.05121](https://arxiv.org/abs/1704.05121))
- Rubin R. H., 1989, *ApJS*, **69**, 897
- Sánchez S. F., et al., 2012, *A&A*, **538**, A8
- Sánchez S. F., et al., 2014, *A&A*, **563**, A49
- Sánchez S. F., et al., 2015, *A&A*, **573**, A105
- Sánchez S. F., et al., 2016, *A&A*, **594**, A36
- Sarzi M., et al., 2010, *MNRAS*, **402**, 2187
- Seon K.-i., Witt A. N., Shinn J.-h., Kim I.-j., 2014, *ApJ*, **785**, L18
- Singh R., et al., 2013, *A&A*, **558**, A43
- Slavin J. D., Shull J. M., Begelman M. C., 1993, *ApJ*, **407**, 83
- Stasińska G., Cid Fernandes R., Mateus A., Sodré L., Asari N. V., 2006, *MNRAS*, **371**, 972
- Stasińska G., et al., 2008, *MNRAS*, **391**, L29
- Thilker D. A., Walterbos R. A. M., Braun R., Hoopes C. G., 2002, *AJ*, **124**, 3118
- Tremonti C. A., et al., 2004, *ApJ*, **613**, 898
- Tüllmann R., Dettmar R.-J., 2000, *A&A*, **362**, 119
- Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, *MNRAS*, **404**, 1639
- Veilleux S., Shopbell P. L., Rupke D. S., Bland-Hawthorn J., Cecil G., 2003, *AJ*, **126**, 2185
- Vogt F. P. A., Pérez E., Dopita M. A., Verdes-Montenegro L., Borthakur S., 2017, *A&A*, **601**, A61
- Weingartner J. C., Draine B. T., 2001, *ApJ*, **563**, 842
- Wild V., et al., 2014, *A&A*, **567**, A132
- Wilson A. S., 1996, *Vistas in Astronomy*, **40**, 63
- York D. G., et al., 2000, *AJ*, **120**, 1579
- Zhang K., et al., 2017, *MNRAS*, **466**, 3217
- Zurita A., Rozas M., Beckman J. E., 2000, *A&A*, **363**, 9
- de Amorim A. L., et al., 2017, *MNRAS*, **471**, 3727

This paper has been typeset from a TeX/LaTeX file prepared by the author.

## **A.2 CALIFA, the Calar Alto Legacy Integral Field Area survey. IV. Third public data release**

Artigo por [Sánchez et al. \(2016\)](#) (doi:[10.1051/0004-6361/201628661](https://doi.org/10.1051/0004-6361/201628661)). Também disponível em *pre-print* ([arXiv:1604.02289](#)).



# CALIFA, the Calar Alto Legacy Integral Field Area survey

## IV. Third public data release<sup>★,★★</sup>

S. F. Sánchez<sup>1</sup>, R. García-Benito<sup>2</sup>, S. Zibetti<sup>3</sup>, C. J. Walcher<sup>4</sup>, B. Husemann<sup>5</sup>, M. A. Mendoza<sup>2</sup>, L. Galbany<sup>6,7</sup>, J. Falcón-Barroso<sup>8,10</sup>, D. Mast<sup>11,12</sup>, J. Aceituno<sup>13</sup>, J. A. L. Aguerri<sup>8</sup>, J. Alves<sup>9</sup>, A. L. Amorim<sup>14</sup>, Y. Ascasibar<sup>20</sup>, D. Barrado-Navascués<sup>13,15</sup>, J. Barrera-Ballesteros<sup>16</sup>, S. Bekeraite<sup>4</sup>, J. Bland-Hawthorn<sup>17</sup>, M. Cano Díaz<sup>1</sup>, R. Cid Fernandes<sup>14</sup>, O. Cavichia<sup>18</sup>, C. Cortijo<sup>2</sup>, H. Dannerbauer<sup>8</sup>, M. Demleitner<sup>19</sup>, A. Díaz<sup>20</sup>, R. J. Dettmar<sup>21</sup>, A. de Lorenzo-Cáceres<sup>22,23</sup>, A. del Olmo<sup>2</sup>, A. Galazzi<sup>3</sup>, B. García-Lorenzo<sup>8</sup>, A. Gil de Paz<sup>24</sup>, R. González Delgado<sup>2</sup>, L. Holmes<sup>25</sup>, J. Iglesias-Páramo<sup>2</sup>, C. Kehrig<sup>2</sup>, A. Kelz<sup>4</sup>, R. C. Kennicutt<sup>26</sup>, B. Kleemann<sup>21</sup>, E. A. D. Lacerda<sup>14</sup>, R. López Fernández<sup>2</sup>, A. R. López Sánchez<sup>17</sup>, M. Lyubenova<sup>27</sup>, R. Marino<sup>28</sup>, I. Márquez<sup>2</sup>, J. Mendez-Abreu<sup>22</sup>, M. Mollá<sup>29</sup>, A. Monreal-Ibero<sup>30</sup>, R. Ortega Minakata<sup>31</sup>, J. P. Torres-Papaqui<sup>32</sup>, E. Pérez<sup>2</sup>, F. F. Rosales-Ortega<sup>33</sup>, M. M. Roth<sup>4</sup>, P. Sánchez-Blázquez<sup>20,34</sup>, U. Schilling<sup>21</sup>, K. Spekkens<sup>25</sup>, N. Vale Asari<sup>14</sup>, R. C. E. van den Bosch<sup>35</sup>, G. van de Ven<sup>35</sup>, J. M. Vilchez<sup>2</sup>, V. Wild<sup>22</sup>, L. Wisotzki<sup>4</sup>, A. Yıldırım<sup>35</sup>, and B. Ziegler<sup>9</sup>

(Affiliations can be found after the references)

Received 7 April 2016 / Accepted 17 June 2016

### ABSTRACT

This paper describes the third public data release (DR3) of the Calar Alto Legacy Integral Field Area (CALIFA) survey. Science-grade quality data for 667 galaxies are made public, including the 200 galaxies of the second public data release (DR2). Data were obtained with the integral-field spectrograph PMAS/PPak mounted on the 3.5 m telescope at the Calar Alto Observatory. Three different spectral setups are available: i) a low-resolution V500 setup covering the wavelength range 3745–7500 Å (4240–7140 Å unvignetted) with a spectral resolution of 6.0 Å (FWHM) for 646 galaxies, ii) a medium-resolution V1200 setup covering the wavelength range 3650–4840 Å (3650–4620 Å unvignetted) with a spectral resolution of 2.3 Å (FWHM) for 484 galaxies, and iii) the combination of the cubes from both setups (called COMBO) with a spectral resolution of 6.0 Å and a wavelength range between 3700–7500 Å (3700–7140 Å unvignetted) for 446 galaxies. The Main Sample, selected and observed according to the CALIFA survey strategy covers a redshift range between 0.005 and 0.03, spans the color-magnitude diagram and probes a wide range of stellar masses, ionization conditions, and morphological types. The Extension Sample covers several types of galaxies that are rare in the overall galaxy population and are therefore not numerous or absent in the CALIFA Main Sample. All the cubes in the data release were processed using the latest pipeline, which includes improved versions of the calibration frames and an even further improved image reconstruction quality. In total, the third data release contains 1576 datacubes, including ~1.5 million independent spectra.

**Key words.** galaxies: evolution – galaxies: general – techniques: spectroscopic – surveys

### 1. Introduction

The advent of large imaging surveys, complemented in some cases by single-fiber spectroscopy (e.g., Sloan Digital Sky Survey, SDSS; Galaxy and Mass Assembly survey, GAMA, York et al. 2000; Driver et al. 2009), has opened important new avenues for our understanding of galaxy evolution. However, one of the most significant limitations of these surveys is that they do not provide resolved spectroscopic information. Galaxies have long been known to be spatially extended objects, with observed properties that vary across their optical extents (e.g., Hubble 1926, 1936). Many of these properties vary coherently as a function of position relative to the galaxy center, and radial gradients have been studied for decades (e.g., Pagel & Edmunds 1981; Peletier 1989). Characterizing galaxies by assigning global values therefore oversimplifies their true

complexity, and determining the spatially resolved properties of galaxies is fundamental to understanding the evolutionary processes that have shaped them. At the same time, evidence for a diversity in galaxy evolutionary histories as a function of mass and environment implies that statistically significant samples over large fractions of the sky are needed to understand the underlying physical mechanisms at work.

Thus, the logical next step for pushing beyond multiband imaging surveys (that provide detailed spatial information and limited spectral information) or single-aperture spectroscopic surveys (that sample only limited galaxy regions) is an Integral Field Spectroscopy (IFS) survey over a representative and statistically significant sample of galaxies. With this aim we undertook the Calar Alto Legacy Integral Field Area (CALIFA) survey in 2010 (Sánchez et al. 2012a) to obtain spatially resolved spectra for ~600 galaxies in the local Universe.

CALIFA was the first survey using imaging spectroscopy that was designed to provide a public dataset of a sample of galaxies representative of the Local Universe, i.e., this survey was foreseen to be a legacy survey from the outset. CALIFA thus aimed at extending the pioneering Spectrographic

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

\*\* The spectra are available at <http://califa.caha.es/DR3>

Areal Unit for Research on Optical Nebulae (SAURON) and Atlas3D surveys (Cappellari et al. 2011) to all galaxy types and larger wavelength coverage. The next generation of IFS surveys are already in progress – Mapping Nearby Galaxies at Apache Point Observatory survey (MaNGA, Bundy et al. 2015) and Sydney Australian-Astronomical-Observatory Multi-object Integral-Field-Spectrograph survey(SAMI, Croom et al. 2012). It is beyond the scope of this article to make a detailed comparison between these surveys and CALIFA, a topic that has already been addressed in previous articles (e.g., Sánchez 2015). Briefly, we note that MaNGA and SAMI will supplant all previous surveys in terms of number of objects. They have adopted a multiplexing scheme that allows us to observe several objects simultaneously. This increases the efficiency of collecting data in terms of number of objects, which was their main goal (10 000 objects, for MaNGA, and 3600 objects, for SAMI). CALIFA and Atlas3D use a single-IFU mode, which limits the number of objects to be observed simultaneously. On the other hand, the CALIFA and Atlas3D surveys observe a larger number of spectra per object, and offer better physical spatial sampling. One obvious manifestation of this is the very similar total number of spectra obtained by CALIFA, MaNGA, and SAMI, compared to the very different total number of objects to be observed.

The sample selections are also quite different. The MaNGA and SAMI samples cover a wider range of redshifts, from  $z \sim 0.001$  to  $z \sim 0.16$ . Because the full width at half maximum (FWHM) of the point spread function (PSF) is very similar in the three surveys this implies a wide range of physical resolutions. Galaxies are thus sampled over a wider range of masses, but at different cosmological distances. On the contrary, the redshift range of the CALIFA and Atlas3D samples are rather small (Cappellari et al. 2011; Walcher et al. 2014), and therefore they present a better and more uniform physical sampling, making them optimal to study spatially resolved structures in galaxies (at  $\sim 1$  kpc resolution).

Another important difference is the coverage of the different IFU surveys in terms of the optical extension of the galaxies. CALIFA observations cover most of the optical extension (beyond  $\sim 2.5 r_e$ ) by construction, while Atlas3D reaches between 1– $1.5 r_e$  on average. MaNGA comprises two main samples, where the goal is to reach either  $1.5 r_e$  ( $\sim 70\%$  of the targets) or  $2.5 r_e$  ( $\sim 20\%$ ), and it hardly samples the outer regions for most of the galaxies (Ibarra-Mendel et al. 2016). Finally, SAMI, with the smallest FoV of all IFU surveys (16''/D) covers around  $\sim 1 r_e$  of the galaxies (Bryant et al. 2015).

In summary the data provided by the CALIFA DR3 presented here occupy a niche which ensures high spatial resolution and good spatial coverage simultaneously, however, at the price of a smaller sample in comparison with currently ongoing surveys like MaNGA and SAMI. The CALIFA collaboration has addressed many different science cases using the information provided by these data, all of them focused on understanding the main properties of galaxies in the local Universe and the evolutionary processes that have shaped them: i) new techniques have been developed to understand the spatially resolved star formation histories (SFH) of galaxies (Cid Fernandes et al. 2013, 2014; López Fernández et al. 2016). Clear evidence that mass-assembly in typical galaxies progresses inside-out (Pérez et al. 2013). The SFH and chemical enrichment of bulges and early-type galaxies are fundamentally related to their total stellar mass, while for disk galaxies they are more closely related to the local stellar mass density (González Delgado et al. 2014b,a, 2015). Negative age gradients indicate that quenching is progressing outwards in massive galaxies (González Delgado et al. 2015),

and age and metallicity gradients suggest that galaxy bars have not significantly altered the SFHs of spirals (Sánchez-Blázquez et al. 2014). Finally, we explored spatially resolved stellar populations and star formation across the Hubble sequence (González Delgado et al. 2015, 2016), and how mergers influence the assembly of blue ellipticals (Haines et al. 2015). ii) We studied the origin of the low intensity, LINER-like, ionized gas in galaxies. These regions are clearly not related to recent star-formation activity, or to AGN activity (Singh et al. 2013). They are most probably related to post-AGB ionization in many cases (Kehrig et al. 2012; Papaderos et al. 2013; Gomes et al. 2016). iii) We explored aperture and resolution effects affecting larger single-fiber (like SDSS) and IFS surveys (like MaNGA and SAMI Bundy et al. 2015; Croom et al. 2012). We explored the effects of signal dilution in IFS data obtained for higher redshift galaxies in different gas and stellar population properties (Mast et al. 2014), and proposed a new empirical aperture correction for SDSS data (Iglesias-Páramo et al. 2013, 2016). We also compared average stellar and ionized gas properties with spatially resolved ones (e.g. González Delgado et al. 2014b, 2015); iv) we studied the kinematics of the ionized gas (García-Lorenzo et al. 2015), the effects of bars on the kinematics of galaxies (Barrera-Ballesteros et al. 2014; Holmes et al. 2015), the effects of the interaction stage on the kinematic signatures (Barrera-Ballesteros et al. 2015a), and measured the bar pattern speeds in late-type galaxies (Aguerri et al. 2015). v) We extended measurements of the angular momentum of galaxies to previously unexplored ranges of morphology and ellipticity (e.g., Falcón-Barroso et al. 2015) and proposed a new dynamical classification scheme for galaxies Kalinova et al. (2015). The stellar dynamics together with detail analysis of the stellar populations revealed a tight relation between the initial mass function (IMF) and the local metallicity (Martín-Navarro et al. 2015) and allowed us to dynamically constrain the shape of the IMF in early-type galaxies (Lyubenova et al. 2016). vi) We explored in detail the impact of galaxy interactions on the enhancement of star-formation rates (SFR) and the ignition of galactic outflows (Wild et al. 2014; Barrera-Ballesteros et al. 2015b). vii) We studied the nature of the progenitors of SNe (Galbany et al. 2014); viii) we explored star formation indicators for extended objects and the suitability of H $\alpha$  as an SFR tracer (Catalán-Torrecilla et al. 2015), as well as the spatially resolved SFR density across the Hubble sequence (González Delgado et al. 2016). ix) We studied oxygen abundance gradients in the gas, developing new calibrators (Marino et al. 2013), finding a characteristic shape (Sánchez et al. 2014; Sánchez-Menguiano et al. 2016), and a weak dependence of the profile truncations on the gradient properties (Marino et al. 2016) and the stellar populations (Ruiz-Lara et al. 2016). x) We explored the mass–metallicity relation for both the stellar populations (González Delgado et al. 2014a) and the gas oxygen abundance (Sánchez et al. 2013). We could not confirm a secondary relation between the SFR and the Metallicity (Sánchez et al. 2015). xi) Finally, we found that many of the global scaling relations such as the star formation main sequence or the mass–metallicity relation are mirrored by local relations that hold on a scale of  $\sim 1$  kpc (e.g., Sánchez et al. 2013; Cano-Díaz et al. 2016).

CALIFA was designed as a legacy survey, and therefore we have distributed the data in successive data releases (DR) as the number of observed objects has increased and the processing pipeline has improved (DR1 and DR2, Husemann et al. 2013; García-Benito et al. 2015, respectively). These publicly accessible data have already allowed the exploration of several different scientific avenues not addressed by the collaboration (e.g.,

Holwerda & Keel 2013; De Geyter et al. 2014; Martínez-García et al. 2014; Davies et al. 2014; Roche et al. 2015; Ho et al. 2015). In this article we present the third and final data release (DR3) comprising all observations with good quality. We distribute 1576 datacubes corresponding to 667 galaxies, 646 of them observed with the V500 setup, 484 observed with the V1200 setup and 446 combined (COMBO) cubes.

The properties of the galaxies in the DR3 sample are described in Sect. 2. We describe the observing strategy and setup (Sect. 3), processing (Sect. 4), structure (Sect. 5), and data (Sect. 6), which comprise essential information for any scientific analysis of the distributed CALIFA data. Several interfaces to access the CALIFA DR3 data are explained in Sect. 7.

## 2. The CALIFA DR3 sample

There are two fundamentally different samples of galaxies in the CALIFA DR3: (1) Galaxies that were targeted because they were part of the CALIFA mother sample that is fully described and characterized in Walcher et al. (2014). This sample is called the CALIFA Main Sample. (2) The CALIFA Extension Sample, which is a heterogeneous set of galaxies observed for various reasons as part of different ancillary science projects within the CALIFA collaboration. The DR3 release is the combination of the Main Sample and the Extension Sample.

### 2.1. Main sample

The Main Sample contains all galaxies for which cubes are released in the CALIFA DR3 and that have been drawn from the CALIFA mother sample.

#### 2.1.1. Mother sample – a recap

The CALIFA mother sample consists of 937 galaxies drawn from SDSS DR7. In earlier papers Walcher et al. (in particular 2014) we quoted the number 939, because we included the galaxies NGC 5947 and NGC 4676B. However, these galaxies did not formally satisfy the selection criteria of the mother sample but were added by hand. We therefore now shifted them to the Extension Sample, where they have a natural place.

The main criteria for the target selection of the mother sample are fully described in Walcher et al. (2014). Briefly for completeness, they are:

- angular isophotal diameter  $45'' < \text{iso}A_r < 79.2''$ ;
- redshift range  $0.005 < z < 0.03$ ;
- Galactic latitude  $|b| > 20^\circ$ ;
- flux limit  $\text{petroMag}_r < 20$ ;
- declination  $\delta > 7^\circ$ .

The lower redshift limit was imposed so that the mother sample would not be dominated by dwarf galaxies. CALIFA thus has a natural lower completeness limit in mass. The upper redshift limit was imposed to keep relevant spectral features observable with a fixed instrumental setup. This limits the total volume probed by the CALIFA sample to roughly  $10^6 \text{ Mpc}^3$ . Because massive galaxies are very rare, this volume effectively sets the upper mass boundary of the CALIFA sample (and not the diameter selection). The 95% completeness limits of the mother sample are studied in detail in Walcher et al. (2014) and are as follows:  $-19 > M_r > -23.1$  in luminosity and  $10^{9.7}$  and  $10^{11.4} M_\odot$  in stellar mass (with a Chabrier IMF, Chabrier 2003).

In Walcher et al. (2014) we showed that the mother sample has well understood properties. In particular, the diameter selection can be translated into  $V_{\max}$  values according to the formalism of Schmidt (1968). This allows us to construct the luminosity function from the mother sample and to show that it agrees with the standard literature determination of the luminosity function within the limits imposed by the sample size. Indeed, while the selection effects are understood and can be corrected within our completeness limits, the finite sample size of CALIFA still implies that some galaxy classes have less representatives within the sample. Specifically for the CALIFA Main Sample, the statistics are best for galaxies with stellar masses around  $10^{10.8} M_\odot$ .

### 2.1.2. Definition of the main sample

Galaxies were selected for observation from the mother sample randomly, i.e., based only on visibility. We can therefore assume that the Main Sample is a random subset of the mother sample. Below we will proceed to verify this assumption. We base our verification on the same galaxy physical properties studied in Walcher et al. (2014). For DR3 we have re-compiled the catalogues of physical properties for two reasons: (1) we introduced a few bug fixes in column names or with single numbers in the catalogues. (2) We computed new properties based on SDSS Petrosian magnitudes to allow for comparison with the Extension Sample. All tabulated properties are available for all galaxies of the mother sample, i.e., by definition for all main sample galaxies in the release.

The Main Sample as used below contains all galaxies with either a V500 and/or a V1200 cube released in this data release and no quality control flags that mark them as unusable (see Sect. 6.4). The number of galaxies in the main sample is 542.

## 2.2. Extension sample

The extension sample consists of an inhomogeneous collection of galaxies observed in the CALIFA setup, but not following the same sample selection criteria of the mother sample. This means in particular that  $V_{\max}$  values cannot be computed for the extension sample. On the other hand, extension sample galaxies have mostly been selected to cover galaxies that are intrinsically rare and thus not found in the CALIFA Main Sample. They thus provide useful benchmarks for such rare galaxies.

The CALIFA setup is used in the observations of the extension sample galaxies, i.e., the same gratings, grating angles, exposure times and observing strategy. However, many extension programs did not select galaxies from the SDSS DR7 imaging survey. This imaging is needed to ensure good photometric calibration (Sect. 4.1). Thus, for all extension programs, the additional selection criteria of being in the SDSS footprint and of fulfilling the quality control (QC) requirements (Sect. 6) are imposed.

### 2.2.1. Dwarf galaxies

The extension program on dwarf galaxies is led by García-Benito. The galaxies in this program have CALIFA IDs between 1000 and 1999. The project aims to observe a statistically meaningful sample of dwarf galaxies in the local Universe ( $9 < D < 40 \text{ Mpc}$ ). The sample was selected to be a magnitude-limited sample of local field galaxies observed with SDSS and

covering a similar observed magnitude range as the CALIFA main sample.

The following selection criteria were used: i) the size of the galaxy (optical diameters) fill the PPak FoV, i.e.,  $29.7'' < \text{iso}A_r < 79.2''$ ; ii) the ratio of the minor to major axis  $\text{iso}B_r/\text{iso}A_r > 0.6$ , so that the galaxies are found to be almost face on; iii)  $M_r > -18$  iv)  $z > 0.002$ .

The input sample contained a total of 82 objects of which 34 were observed and included in the CALIFA DR3.

## 2.2.2. Pairs and interacting galaxies

The extension program on pairs and interacting galaxies is led by Barrera-Ballesteros. The galaxies in this program have CALIFA IDs between 2000 and 2999. The project aims to complete the IFU data for those pairs where only one companion galaxy was included in the CALIFA mother sample. The sample was selected to include companions of a CALIFA main sample galaxy with: i) a projected distance of 160 kpc; ii) a systemic velocity difference of less than  $300 \text{ km s}^{-1}$ ; ii) an  $r$ -band magnitude difference of less than 2 mag. For details of the CALIFA pair selection see [Barrera-Ballesteros et al. \(2015b\)](#).

This sample also includes the second galaxy of the “mice” (NGC 4676B) as ID number 2999. This galaxy was observed in the time allocated to the main survey since it seemed worthwhile to cover the full “mice” merger. These data have been published in ([Wild et al. 2014](#)).

In total there are 29 galaxies from this program in the CALIFA DR3.

## 2.2.3. Low- and high-mass early-type galaxies

The extension program on early-type galaxies (ETGs) is led by van de Ven, Lyubenova & Meidt. The galaxies in this program have IDs between 3000 and 3999. This extension contains are three subprograms with the following scientific aims: a) studying the dark matter content of low-mass ETGs; b) constraining the IMF in high-mass ETGs; and c) testing fading scenarios for low-mass early-type spirals.

The selection criteria for the low-mass ETGs were: a)  $9.4 < \log(M_*/M_\odot) < 10.4$ ; b)  $35'' < \text{iso}A_r < 79.2''$ ; c) declination  $\delta > 7^\circ$  and  $75^\circ < \text{RA} < 300^\circ$ ; and d) concentration  $r_{90}/r_{50} > 2.8$ ; visual inspection to remove non-ETGs. The selection criteria for the high-mass ETGs were: a)  $\log(M_*/M_\odot) > 11.4$ ; b)  $35'' < \text{iso}A_r < 79.2''$ ; c) declination  $\delta > 7^\circ$  and  $75^\circ < \text{Right Ascension RA} < 300^\circ$ ; d) concentration  $r_{90}/r_{50} > 2.8$ ; and e) visual inspection to remove non-ETGs. The selection criteria for low-mass early-type spirals were: a)  $9.0 < \log(M_*/M_\odot) < 10.0$ ; b)  $30'' < \text{iso}A_r < 80''$ ; c) declination  $\delta > 7^\circ$ ; d) inclination =  $\text{acos}(1-\text{iso}B_r/\text{iso}A_r) < 80^\circ$ ; e) morphological types S0, Sa and Sb; and f) preference to those with literature HI observations.

For all three subsamples the number of available galaxies was then reduced further by visibility at the scheduled time of observation. There are 36 galaxies from this program in the CALIFA DR3.

## 2.2.4. Pilot studies

Those galaxies with IDs in the range 4000–4999 were extracted from CALIFA pilot studies and are fully described in [Mármol-Queraltó et al. \(2011\)](#). Most of the pilot study galaxies were observed with either the V300 or the V600 gratings and can therefore not be included in this homogeneous data release. The pilot

studies targeted spiral galaxies with low inclinations to study the properties of the HII regions, as described in [Sánchez et al. \(2012b\)](#).

Note that the galaxy with ID number 4034, NGC 5947, is a galaxy from the pilot study sample that was observed as part of the main sample as described in [Walcher et al. \(2014\)](#). There are 3 galaxies from this program in the CALIFA DR3, including NGC 5947. Those are only ones included in the Pilot Sample that were observed with the final CALIFA configuration and are part of the SDSS imaging survey (a pre-requisite for being distributed in this data release). All the three fulfill the quality control criteria.

## 2.2.5. Supernova environments

The extension program on core collapse supernova (CCSN) environments is led by Galbany. The galaxies in this program have IDs between 5000 and 5999. The project aims to get imaging spectroscopy for low-mass galaxies that hosted Type Ib, Ic and II supernovae with available light curves. This sample increases the completeness of the SN host galaxy sample from the CALIFA Survey presented in [Galbany et al. \(2014\)](#) and [Galbany et al. \(2016\)](#), which has a deficit of CCSNe in low-mass ( $\lesssim 10^{10} M_\odot$ ) galaxies.

Objects were selected from the Asiago SN catalog<sup>1</sup> ([Barbon et al. 1999](#)) following these criteria: a) SN projected galactocentric distance lower than 40 arcsec, in order to cover the local SN environment; b) systemic velocity of the galaxy lower than  $9000 \text{ km s}^{-1}$  ( $\sim z < 0.03$ ); c)  $\log D_{25}$  (decimal logarithm of the apparent 25 mag/arcmin<sup>2</sup> isophotal diameter) lower than 1.12, which corresponds to galactic radius lower than 40 arcsec; d) declination  $\delta > 0^\circ$ ; and e) SN light curve publicly or privately available.

The input sample contains a total of 49 objects of which 14 were observed and included in the CALIFA DR3.

## 2.2.6. Compact early-type galaxies

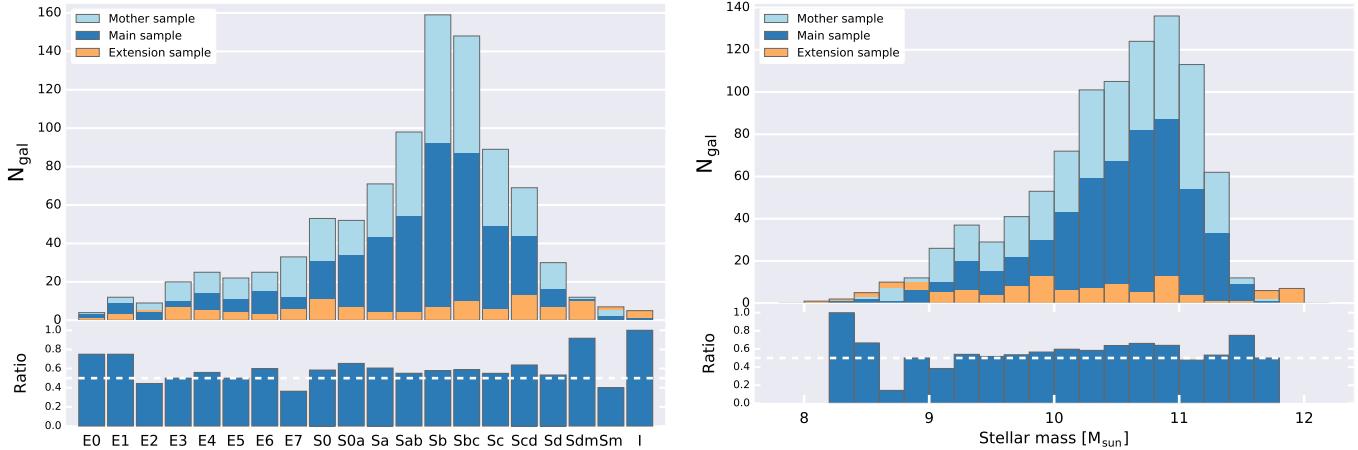
The extension program on compact early-type galaxies is led by Yıldırım and van den Bosch and includes galaxies with CALIFA IDs between 9000 and 9999. The scientific aim is the characterization of extremely compact host galaxies of supermassive black holes. Galaxies were selected from the HETMGS (*Hobby-Eberly Telescope Massive Galaxy Survey*, [van den Bosch et al. 2015](#)) to have: i) large black hole spheres of influence; ii) 2MASS half-light radii smaller than 2 kpc; iii) very high central density as measured by the velocity dispersion over central resolution element ( $\sigma/1''^2$ ). Scientific results and further descriptions of the sample can be found in [Yıldırım et al. \(2015, 2016\)](#). In those papers, the data are deeper than the standard CALIFA depth, as the exposure times are four times longer. For the sake of uniformity of the DR3, we limit ourselves to the standard CALIFA depth, i.e., exposure time. Observations are taken with the V500 low-resolution setup only.

Only 7 of the 17 galaxies are included in this data release as the remainder lack SDSS imaging needed for the CALIFA pipeline.

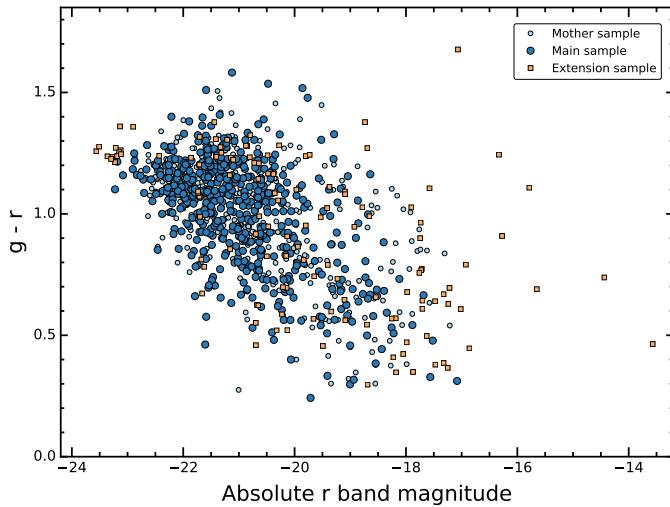
## 2.2.7. Other extension programs

There are two galaxies in DR3 from extension programs that yielded only one released cube, and it is therefore not useful to

<sup>1</sup> <http://graspa.oapd.inaf.it/asnc.html>



**Fig. 1.** Left panel: histograms of visual morphological classification in the DR3 samples. Right panel: stellar mass histogram. The lower portion of each panel shows the ratio between the main sample and the mother sample.



**Fig. 2.** Distribution of DR3 galaxies in the color magnitude diagram.

describe these programs in detail. The galaxy with CALIFA ID number 7001 is CGCG263-044, classified as Sb in NED, is relatively edge-on and classified as a Type 2 AGN. The galaxy with CALIFA ID number 8000, NGC 426, is a massive galaxy, classified as a cD.

### 2.3. Properties of the released galaxies

The physical properties each galaxy in each sample, including its name, CALIFA ID, coordinates, redshift, photometry, morphology, and stellar mass are available from the DR3 web page.

As clearly seen in Fig. 1, the morphological and stellar mass distributions of the main sample are consistent with those of the mother sample. While larger differences are seen at low stellar masses, this follows from the low number of galaxies overall in this mass range. This test therefore implies that the selection of main sample galaxies from the mother sample using target visibility preserves the mother sample statistics.

Figure 2 shows the distribution of galaxies in the color-magnitude diagram. Again, the good coverage of the color-magnitude diagram found for the mother sample is retained for the main sample. The Extension Sample *by design* covers those

regions that were not included in the mother sample, i.e., the bright end of the red sequence and dwarf galaxies.

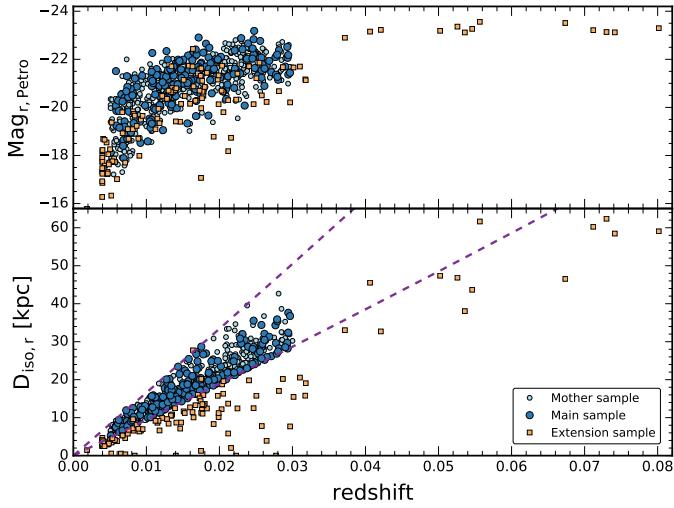
In Walcher et al. (2014) we reported on the average spatial coverage of the mother sample galaxies by the IFU as a fraction of the Petrosian half light radius (SDSS pipeline quantity PetroR<sub>50</sub>)  $r_{50}$  as computed from the growth curve photometry. About 97% of all galaxies are covered out to at least  $2 \times r_{50}$ . This statement holds for the main sample as well. Because growth curve photometry is not available for the Extension Sample, we refer here to the SDSS Petrosian half light radius, which we will denote as  $r_e^P$ . The average spatial coverage in terms of  $r_{50}$  is 4.2 for the main sample and 7.9 for the Extension Sample, with the broader coverage of the Extension Sample being driven by the inclusion of dwarf galaxies therein. All Extension Sample galaxies are covered out to the SDSS isophotal major axis, as is the case for the main sample.

Figure 3 shows the distribution of galaxies from the main sample and Extension Sample in the redshift-magnitude and redshift-size diagrams. Clearly, galaxies slightly larger than the CALIFA size limit are so rare that they are under-represented within the CALIFA volume *independently of their size*. Finding them requires an extension to higher redshift. Galaxies smaller than the CALIFA size limit are abundant in the local Universe. These dwarf galaxies were not included in the mother sample *by design* to avoid swamping the main sample with them. The dwarf galaxies in the Extension Sample somewhat helps to circumvent this self-imposed limitation. Still, a dedicated dwarf imaging spectroscopy survey of similar size to CALIFA is missing in the literature.

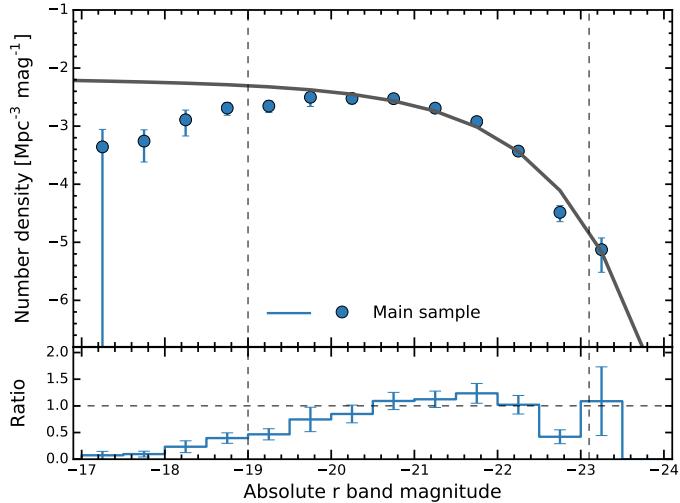
### 2.4. Luminosity functions of the main sample

In Walcher et al. (2014) we established that the luminosity function of the CALIFA mother sample compares well with the reference SDSS sample of Blanton et al. (2005). We now investigate whether the same statement can be made for the main sample. For all technical details on how the luminosity functions are obtained and for the derivation of the CALIFA mother sample completeness limit at  $M_r \approx -19$  we refer the reader to Walcher et al. (2014).

The only difference in the derivation of the luminosity function for the main sample is that the available volume has been made smaller by a random subselection from the mother sample. To correct for this, one needs to multiply the  $V_{\max}$  values of



**Fig. 3.** Distribution of DR3 galaxies in the redshift-magnitude (upper panel) and redshift-size (lower panel) diagrams.



**Fig. 4.** Top panel:  $r$ -band luminosity functions for the SDSS reference sample from Blanton et al. (2005; solid line) and the DR3 main sample. Error bars represent Poissonian uncertainties. Bottom panel: ratio between the main sample and SDSS.

the main sample galaxies with a ratio of the number of galaxies in the sample in use (e.g., 542 in the case of the full main sample) divided by the number of galaxies in the mother sample (937). The precondition for this simple procedure is that the sub-sample in use can be considered a random subset of the mother sample. We compared the mass and morphology distributions of the main sample and the mother sample using a  $\chi^2$  test and found that the probability that they were drawn from the same distributions are 98% and 97%, respectively. Thus we conclude that the main sample is a random subselection from the mother sample.

Figure 4 shows the  $r$ -band luminosity function of the DR3 main sample compared to the Schechter function derived in Blanton et al. (2005). Within the statistical uncertainties, the DR3 main sample reproduces the standard luminosity function. There are two points at the limits of our completeness range (at  $-19.25$  and  $-22.75$  in absolute magnitude) that seem to deviate more than the  $1\sigma$  range of their errorbars. It is not possible to formally decide whether this indicates a (small) issue with our

completeness or whether this is just as expected from statistics ( $\sim 30\%$  of points should lie outside the  $1\sigma$  errorbar). In any case, a formal  $\chi^2$  test indicates a probability of 99.9% that the two functions are identical. We thus conclude from this section that the main sample is representative of the galaxy population within the same limits indicated in Walcher et al. (2014). This is an important result that allows the use of CALIFA data to construct galaxy distribution functions in all scientifically useful parameters. One should keep in mind, however, that not all mass ranges are equally well sampled statistically.

### 3. Observing strategy and setup overview

The observing strategy and setup of the CALIFA survey were described in detail in Sánchez et al. (2012a) and have not changed during the survey's six year duration. For the sake of completeness, we provide here a very brief summary.

All galaxies were observed using PMAS (Roth et al. 2005) in the PPak configuration (Verheijen et al. 2004; Kelz et al. 2006). The PPak science bundle was created to cover a wide area on sky following the requirements of the Disk Mass Survey (Bershady et al. 2010), and is now a common-user instrument. The PPak Integral Field Unit (IFU) has a Field of View (FoV) of  $74'' \times 64''$  and it contains a total of 382 fibers, distributed in three different groups. 331 “science” fibers are arranged in a hexagonal grid, with each fiber projecting to  $2''/7$  in diameter on the sky. The fiber-to-fiber distance is  $3''/2$ , which yields a total filling factor of  $\sim 60\%$ . An additional set of 36 fibers located in a ring with radius  $72''$  measure the surrounding sky background. Finally, 15 fibers are connected to the calibration unit.

The goal of CALIFA was to observe every sample galaxy using two different overlapping setups. The red low-resolution setup (V500;  $R \sim 850$ ) covers the wavelength range 3745–7500 Å. The spectra on the CCD are affected by internal vignetting within the spectrograph, giving an unvignetted range of 4240–7140 Å. The blue mid-resolution setup (V1200;  $R \sim 1650$ ) covers the range 3400–4840 Å, with an unvignetted range of 3650–4620 Å. The resolutions quoted are those at the overlapping wavelength range ( $\lambda \sim 4500$  Å). However, since observing in the V1200-setup was more time consuming and required more restrictive weather conditions, not all galaxies were finally observed in both setups. For those that were observed in both setups, the quality control rejected a larger fraction.

A three pointing dithering scheme was used for each object in order to reach a filling factor of 100% across the entire FoV of the science fibers. The dither comprises the following pattern of offsets:  $(0, 0)$ ,  $(-5.22, -4.53)$ , and  $(-5.22, 4.53)$ , in arcsec with respect to the nominal position of the galaxy. This pattern comprises a jump of two inter-fiber distances instead of one to avoid sampling the same area on sky with the same fiber. We thus minimize the effects of low transmission fibers in the final dataset.

The exposure times per pointing were fixed, and have been selected to match the signal-to-noise (S/N) requirements of Sánchez et al. (2012a). We carried out V1200 observations during dark nights with an exposure time of 1800 s per pointing (split into 2 or 3 individual exposures). We obtained V500 observations during gray nights with 900 s per pointing.

For the observations of the main sample, target galaxies were randomly selected from the mother sample. The strategy for observing extension sample galaxies was more varied, and depended on the extension program. Altogether, 685 galaxies were observed during the last 6 yr, making use of 237 clear equivalent nights distributed between the 6th of June 2010 and the 16th of

April 2016. The number 237 is obtained from 176 totally clear nights, plus observations distributed in partially clear nights or among other service programs. Data cubes for 667 of these are in DR3, following strict quality control procedures.

In the following section, we describe the improvements to the CALIFA data reduction pipeline used to produce the DR3 data.

## 4. Data processing and error propagation

### 4.1. Overview of the reduction scheme

The CALIFA data reduction was performed by a semi-automatic pipeline that follows the procedures for the reduction of fiber-fed IFS data described in Sánchez (2006). The CALIFA data reduction pipeline was first (until V1.2) developed in Perl and using R3D<sup>2</sup>. It was then rewritten using a Python-based core in the Py3D package (V1.3c and V1.5 Husemann et al. 2013; García-Benito et al. 2015). The current pipeline version, V2.2, has now fully replaced the original scheme and uses Python for the full process.

The reduction process comprised the following steps: i) the four different FITS files created by the amplifiers of the detector were re-arranged into a single frame, which is then bias subtracted. Cosmic rays were removed and masked according to Husemann et al. (2012); ii) relative offsets in the tracing due to flexure were estimated with respect to the continuum and arc-lamp calibration frames and the wavelength solution was applied to each individual science frame; iii) the stray-light map was reconstructed using the gaps between the fiber traces<sup>3</sup> and subtracted from the calibration and science exposures; iv) an optimal extraction algorithm (Horne 1986) was used to extract the spectra based on measurements of the positions and widths derived from the continuum lamp. The optimal extraction relies on a good characterization of the shape of the spectra along the cross-dispersion axis. In theory it is required to make a measurement considering both the Gaussian core and the Lorentzian wings. However, in practical terms the adopted procedure to subtract the stray-light removes most of the Lorentzian wings, and therefore a pure Gaussian function is a good representation of the shape of the spectra; v) the extracted flux for each pixel in the dispersion direction was stored in a row-stacked-spectrum (RSS) file; vi) the spectra were resampled to a linear grid in wavelength using the wavelength solution and resolution obtained from the HeHgCd calibration lamp exposure taking for each pointing, taking into account possible flexure offsets in the dispersion axis within each pointing; vii) in the same step, the spectral resolutions were homogenized to a common value along the dispersion axis using an adaptive Gaussian convolution kernel.

The Poisson shot noise, the read-out noise, and bad pixel masks were propagated in the reduction process. For the wavelength solution, errors were analytically propagated during the Gaussian convolution and a Monte Carlo approach was used to estimate the noise vector after the spline resampling of the spectra. Fiber-to-fiber transmission throughput was corrected using an RSS master fiberflat created from sky exposures taken during twilight on all survey observing nights.

Flux calibration was performed using the procedure described in García-Benito et al. (2015). In essence we performed a

dedicated calibration program, re-observing two dozen CALIFA ETGs chosen as secondary calibrators and a set of the standard stars with the PMAS Lens-Array (LArr; Husemann et al., in prep.). This IFU covers a continuous  $16'' \times 16''$  FoV which produces highly accurate spectrophotometric spectra. Comparing the photometry of the calibrated LArr data with aperture-matched SDSS photometry in the  $g$  and  $r$ -bands, the absolute spectrophotometric accuracy of our standard galaxies is  $<0.03$  mag.

During the last 4 years of the survey (2012–2015) we were observing the ETG calibration galaxies regularly, and updating the master sensitivity curve of the instrument/setup created as described in García-Benito et al. (2015). We adopted that sensitivity curve to perform the spectrophotometric calibration for DR3. For each particular pointing the flux calibration was performed by correcting for the atmospheric extinction using the mean observatory curve presented by Sánchez et al. (2007), and using the extinction ( $A_V$ ) provided by the Calar Alto Virtual EXtinction monitor (CAVEX) measured at the moment of the observations. When the CAVEX was not operating ( $\sim 15\%$  of the time) we adopted the average extinction at the observatory ( $A_V \sim 0.15$  mag). Most of the remaining systematic effects in the spectrophotometric uncertainty for CALIFA are driven by the uncertainties in the wavelength-dependent atmospheric extinction at the time of each observation, that is not properly monitored at the observatory.

The science spectra corresponding to the three dithered exposures were combined into a single frame of 993 spectra, following the registration procedure described in García-Benito et al. (2015). In summary, we computed the flux corresponding to the 331 apertures of the fibers for each pointing from sky-subtracted SDSS DR7 images in the bands covering the wavelength of our observation. The apertures were shifted in right ascension and declination over a search box around the nominal coordinates of the pointing and the best registration was found on the basis of a  $\chi^2$  comparison between the SDSS aperture-matched fluxes and those derived from the RSS spectra themselves. This provided us with accurate astrometry for each pointing (with a typical error of  $\sim 0.2''$ ), and a flux recalibration pointing by pointing. This recalibration anchors the absolute CALIFA spectrophotometry to that of the SDSS DR7.

After the Galactic extinction correction (Schlegel et al. 1998; Cardelli et al. 1989), the RSS was ready for the spatial rearranging of the fibers and creation of the datacube. We used a flux-conserving inverse-distance weighting scheme to reconstruct a spatial image with a sampling of  $1''$  using the same parameters as described in García-Benito et al. (2015). This scheme is now adopted by other IFU surveys such as MaNGA (Bundy et al. 2015), as described by Law et al. (in prep.). First, we reconstructed the datacube and estimated the differential atmospheric refraction (DAR) offset. In a second step, we reconstructed the cube again but shifting the position of the fiber at each wavelength against the regular grid according to the DAR offset measured in the first reconstruction. This two-stage iteration avoids one resampling step, which is important for accurate error propagation.

### 4.2. Improvements to the CALIFA data reduction scheme

The main improvements to the pipeline when going from V1.5 (DR2) to V2.2 (DR3) are: i) a new version of the high-level scripting code that handles the night-by-night reduction; ii) an improved version of the COMBO datacubes, i.e., the cubes created by combining the datasets from the V500 and V1200 setups

<sup>2</sup> <http://www.astroscu.unam.mx/~sfsanchez/r3d/>

<sup>3</sup> See Kelz et al. (2006) for a description of the spatial arrangement of the fibers on the pseudo-slit and CCD.

(already used in different science studies, e.g. [González Delgado et al. 2015, 2016](#)); iii) fully automatic verification of the registration and astrometry process; iv) a second-order correction of the datacube reconstruction based on a comparison with the SDSS images.

As indicated before, the new high-level routines that handle the reduction are now written in Python 2.7. Like in the previous version the low-level routines of the reduction are based on the Py3D package, and therefore the reduction sequence and detailed routines have not changed since V1.5 ([García-Benito et al. 2015](#)). The use of Python for the high-level routines provides three basic improvements: i) portability, i.e., the pipeline can be easily installed on any workstation, which is more user friendly; ii) maintenance, i.e., future updates to the pipeline will be easier; and iii) speed, thanks to the multi-processing module. Basically all objects of the same type (science objects, calibration stars, etc.) observed during a night can be processed in parallel.

The pipeline creates a new version of the COMBO cubes by taking into account the data from both the V500 and V1200 setups in the overlapping areas of the spectrum. These COMBO cubes are created to solve the problem of vignetting affecting both setups at the edge of their spectral wavelength ranges (Sect. 4.2 of [García-Benito et al. 2015](#)). The COMBO cubes span an unvignetted wavelength range of 3700–7140 Å.

In the previous version (not distributed in DR2 and DR3), the COMBO cubes were created by glueing the data from both datasets together at a cut-off wavelength. Specifically, the resolution of the V1200 cube was degraded to that of the V500 cube and the cubes were recentered. Then, the resolution-matched V1200 cube was used below a cut-off wavelength of 4500 Å. Above that wavelength the V500 cube was used. In order to avoid low-level spectrophotometric mismatches the V1200 spectra were re-scaled to the V500 ones by the average of the ratio between both datasets in the overlapping wavelength range.

The COMBO cubes in DR3 were created by using the spectra from both setups in the overlapping region simultaneously. Like in the previous case, the V1200 data were spatially recentered, flux rescaled and degraded in resolution to match the V500 data. Then the COMBO cube was created by averaging the spectra corresponding to each dataset in the overlap region, weighted by the inverse of the cube error. For the remaining wavelength range the COMBO cube consists of only the V1200 or the V500 spectra. This new procedure improves the S/N in the overlap region of the COMBO cubes.

The current spatial registration scheme is the same as the one described in [García-Benito et al. \(2015\)](#). It is known that this registration process fails in some cases, particularly in low surface brightness and/or edge-on galaxies or in the presence of bright foreground field stars. These failures happened more frequently in the V1200 setup, given its lower S/N on average compared to that of the V500 setup. The current V2.2 pipeline automatically discards the registration procedure when the minimum  $\chi^2$  is higher than a given threshold. Then it only applies a global flux re-scaling as described in [Husemann et al. \(2013\)](#), relying on the nominal offsets provided by the telescope for the World Coordinate System (WCS [Greisen & Calabretta 2002](#)). A boolean header keyword (REGISTER) is added to the datacube to indicate whether the cube has been fully registered or not. In [García-Benito et al. \(2015\)](#) we found that the astrometric solution has an accuracy better than 3'' for ~93% of the targets. We repeated the analysis for the new dataset and we find that there is a better precision in our astrometry, with a standard deviation in the offset with respect to the SDSS one of ~0.6'' in both RA and Dec. However, we have a systematic offset of -0.6'' in both

directions (e.g.,  $R_{\text{CALIFA}} - R_{\text{SDSS}}$ ), whose source is still unclear.

Finally, a second-order correction was applied to the CALIFA datacubes to match their spectrophotometry as much as possible to that of the SDSS images. This procedure followed a two-step process. First, a second-order correction to the WCS astrometry of the cubes was obtained. Sky-subtracted SDSS DR7  $g$ - (for the V1200) and  $r$ - (for the V500) band images were down-sampled to 1''/pixel. The corresponding images using the proper filter response curve were created from the CALIFA datacubes. The synthetic CALIFA image and SDSS image were registered using a Discrete Fourier Transform and the offsets are updated in the CALIFA image and cube headers. We then convolved the SDSS image with a sequence of circular Moffat kernels,

$$I(r) = I_0 \left[ 1 + \left( \frac{r}{\alpha} \right)^2 \right]^{-\beta}, \quad (1)$$

varying  $\alpha$  and  $\beta$ , i.e., effectively varying the Full Width Half Maximum (FWHM), which depends on both parameters,

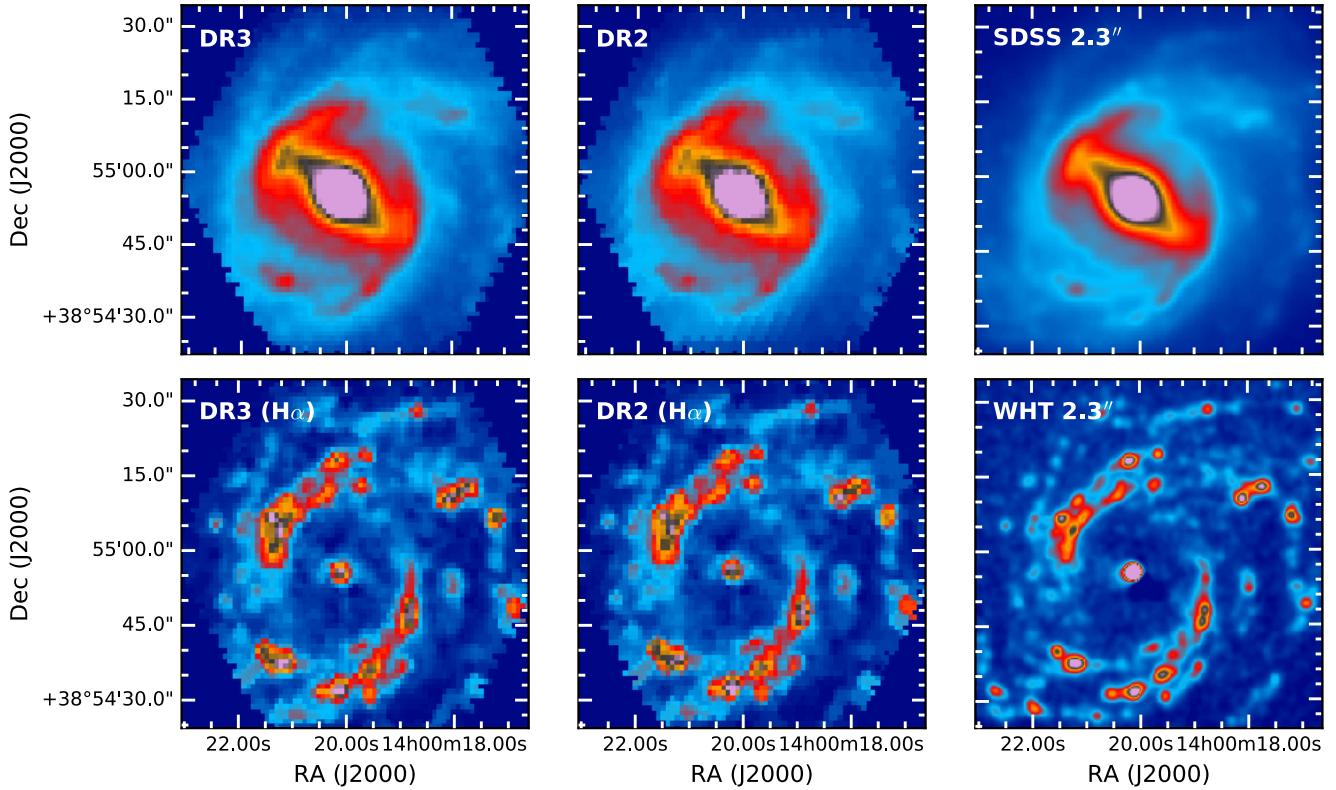
$$FWHM = 2\alpha \sqrt{2^{\frac{1}{\beta}} - 1}. \quad (2)$$

Each convolved SDSS image was then compared to the corresponding reconstructed CALIFA images and the best parameters are selected by  $\chi^2$  minimization. This procedure provides the differential spatial resolution between the SDSS images and the CALIFA cubes. Taking into account the FWHM of the Point Spread Function (PSF) for the SDSS images (that on average is ~1.5'') we obtained a first order estimate of the FWHM of the CALIFA PSF for each cube. This is stored in the header keyword FWHM in units of arcsec.

Once the convolved SDSS image that best reproduces the one reconstructed from CALIFA was obtained, we computed their ratio. This ratio, called *SDSSflat*, is a 2D map with a mean value of one and a scatter of a few percent across the FoV. The final correction was applied by multiplying the fluxes and variances of the data with this 2D map, changing the photometric absolute scale in each spaxel, without affecting the shape of the spectra. The flux level of the integrated spectrum for each datacube changed by less than a few percent both in absolute and relative terms (from blue to red), due to the different relative contribution of each individual spaxel to the sum.

The *SDSSflat* is stored in an additional extension in the FITS files named **FLAT**. In some cases, in particular when there are registration issues and/or very bright field-stars, the procedure fails. This was easily identified during the QC process since the distribution of values within *SDSSflat* was not symmetric, was not centered on ~1, and its application to the cube modified the shape of the integrated spectrum by more than the expected spectrophotometric accuracy (~3%). In those cases we have preferred not to apply the correction. Whether this correction is applied or not is indicated by a header keyword (FLAT\_SDSS), that is set to **true** or **false**.

Figure 5 illustrates the improvements and similarities between V2.2 (DR3) and V1.5 (DR2) of the data reduction by comparing i) the  $g$ -band images extracted from both datacubes of the galaxy NGC 5406 (ID=648) with the corresponding resolution-matched  $g$ -band SDSS image, and ii) the  $H\alpha$  intensity map extracted from both datacubes using the Pipe3D pipeline ([Sánchez et al. 2016](#)), together with a narrow-band image centered on the same emission line. This last image was taken with the 4.2 m *William Herschel* Telescope (Roque de los Muchachos Observatory, La Palma, Spain) using the AUXCAM detector



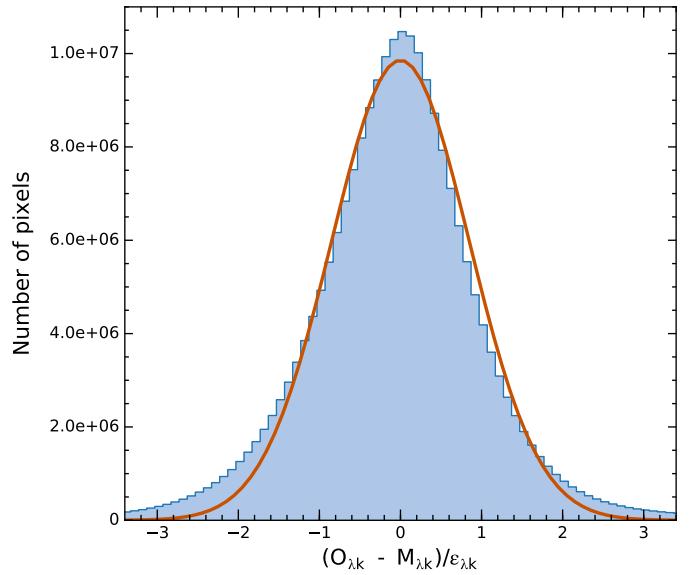
**Fig. 5.** Top panel: comparison between  $g$ -band images extracted from CALIFA datacubes of the galaxy NGC 5406 (ID=684) created using the V2.2 (DR3) and V1.5 (DR2) pipelines, together with the corresponding SDSS image convolved with a Gaussian function to match the spatial resolution of the CALIFA datacubes. Bottom panel: similar comparison between  $H\alpha$  images extracted from the same datacubes using the Pipe3D pipeline, and a resolution-matched  $H\alpha$  image taken with the 4.2 m *William Herschel Telescope* (Roque de los Muchachos Observatory, La Palma, Spain), using the AUXCAM detector (Sánchez-Menguiano et al., in prep.). The FoV has been reduced to match that of CALIFA.

(Sánchez-Menguiano et al., in prep.). The effects of the *SDSSflat* are visible through the absence of patchy structures in the broadband images from V2.2, evident in the V1.5 image. Finally, the spatial resolution is not affected, as clearly seen in the similarities between the three  $H\alpha$  images.

#### 4.3. Characterization of spatially correlated noise

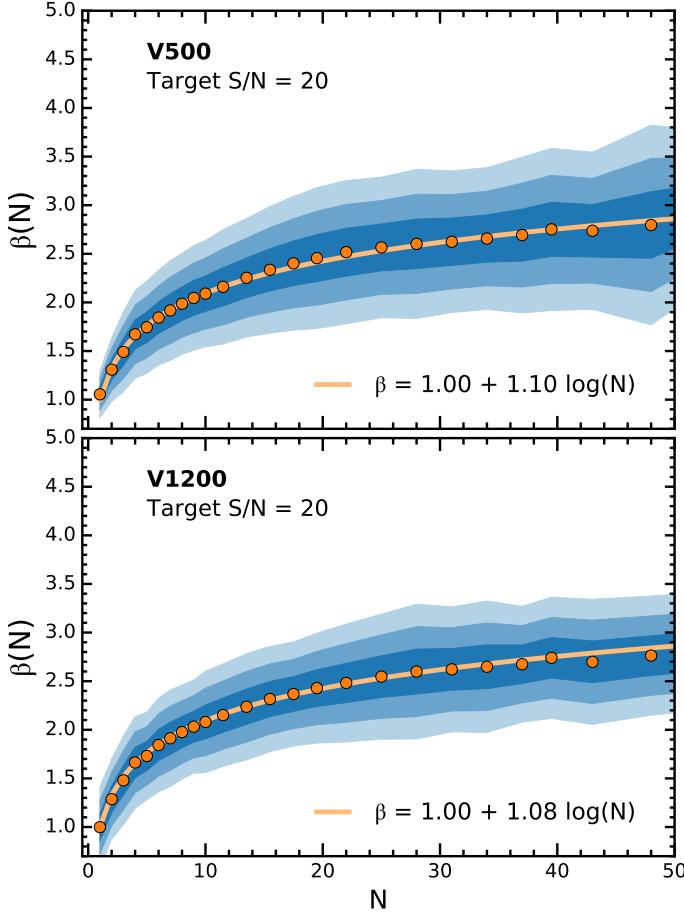
Like in the case of V1.3c and V1.5 the interpolation procedure used to obtain a regular grid implies that the output pixels in the final datacube are not independent of each other. The Gaussian interpolation method distributes the flux from a given fiber to several pixels, which are combined with neighboring pixels within a certain radius (see Sect. 4 of García-Benito et al. 2015). This causes the noise in the adjacent pixels to be spatially correlated. Recall that even in the case that there is no interpolation of the RSS files, all spectra are correlated at some level due to their projection on the detector. This correlation is stronger in adjacent spectra at the detector level, that are not necessarily adjacent in the focal plane of the telescope (Kelz et al. 2006). This correlation implies that a measurement of the noise in a stacked spectrum of  $N$  pixels will be underestimated. Characterizing this effect is essential for estimating the statistical errors when spectra in datacubes are coadded.

Following Husemann et al. (2013) and García-Benito et al. (2015) we checked that the error spectra derived from the pipeline for individual spaxels are reliable. Spectral fitting analysis can provide an approximate assessment of the accuracy of the error spectra. In Fig. 6 we update Fig. 9 of



**Fig. 6.** Histogram of the reduced residuals  $(O_{\lambda,k} - M_{\lambda,k})/\epsilon_{\lambda,k}$  for all wavelengths, all bins ( $k$ ) and all galaxies in DR3 (433400381 points in total). The orange line shows the best Gaussian fit to the sample.

Husemann et al. (2013) and Fig. 10 of García-Benito et al. (2015) for DR3 data. The plot shows the histogram of residuals, i.e., the difference between the observed ( $O_\lambda$ ) and synthetic ( $M_\lambda$ ) spectra obtained with Pipe3D in units of the corresponding error



**Fig. 7.** Noise correlation ratio  $\beta$  (ratio of the real estimated error to the analytically propagated error) as a function of number of spaxels per bin for all the V500 (top panel) and V1200 (bottom panel) data of DR3 at a target S/N of 20. Shaded areas mark the  $1\sigma$ ,  $2\sigma$ , and  $3\sigma$  levels. The orange lines represent the best fitting logarithmic function with a slope  $\alpha = 1.10$  and  $\alpha = 1.08$ , respectively.

$\epsilon_\lambda$  (details of the fitting procedures can be found in Sect. 6.6). The distribution is very well described by a Gaussian centered at 0.01 with  $\sigma = 0.83$ , only slightly lower than expected if residuals are purely due to uncorrelated noise.

The correlated noise can be taken into account by providing the spatial covariance (Sharp et al. 2015). However, like in DR1 and DR2, a more practical approach consists of using the datacubes to introduce the noise correlation ratio as a function of the number of pixels  $\beta(N)$ .  $\beta$  is the ratio of the “real” or measured error to the analytically propagated error of the binned spectra as a function of bin size. To calculate  $\beta$  We used the Voronoi adaptive binning method (implemented for optical IFS data by Cappellari & Copin 2003) with a target S/N of 20 to obtain a sample of coadded spaxels covering different numbers of spaxels. We removed individual spaxels with  $S/N < 5$  from the analysis, and coadded bins with areas larger than 60 spaxels. The “real” noise was then obtained from the detrended standard deviation in certain defined wavelength windows (see Sect. 6.7). The results obtained for all DR3 datacubes, shown in Fig. 7, can be described well by the following logarithmic function:

$$\beta(N) = 1 + \alpha \log N, \quad (3)$$

where  $N$  is the number of spaxels per bin.

The values for the slope  $\alpha$  are very similar in both setups, with a value of  $<1.10$  for V500 and 1.08 for V1200, with errors

in the estimation of the slope around 0.01. Both of them agree well with the observed distribution within one sigma. The slope is lower than the DR1 value (mean  $\sim 1.4$ ), but very similar to the value reported for DR2. This indicates that the noise in DR3 and DR2 datacubes is less correlated than that in the DR1 datacubes. This is expected since we changed the parameters in the interpolation and the registration procedure from V1.3c to V1.5 and V2.2. Detailed instructions on how to estimate the coadded error spectrum are given in Appendix A of García-Benito et al. (2015).

## 5. CALIFA data format and characteristics

The CALIFA data are stored and distributed as 3D data cubes in the standard binary FITS format. Each FITS file consists of several Header Data Units (HDUs). These HDUs contain, in order within each FITS file, (1) the measured flux densities, corrected for Galactic extinction as described in Sánchez et al. (2012a), in units of  $10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$  (primary datacube); (2) the associated errors; (3) the error weighting factors; (4) the bad pixel flags; (5) the fiber coverage; and (6) the SDSSflat correction to the interpolation scheme (see also Table 1). This last HDU was absent in DR1 and DR2, as explained in Sect. 4.2. The remaining extensions were explained in detail in Husemann et al. (2013) and García-Benito et al. (2015).

The first two axes of the cubes correspond to the spatial dimension along right ascension and declination with a  $1'' \times 1''$  sampling. The third dimension represents the wavelength and is linearly sampled. Table 2 summarizes the dimensions of each datacube ( $N_\alpha$ ,  $N_\delta$ , and  $N_\lambda$ ), as well as the spectral sampling ( $d_\lambda$ ) and resolution ( $\delta_\lambda$ ).

### 5.1. Error and weight datacubes

The  $1\sigma$  uncertainty of each pixel as formally propagated by the pipeline can be found in the first FITS extension. Section 4.3 discusses the accuracy of the uncertainties and their correlation properties. This is important when CALIFA data are spatially binned, and an empirical function is provided to account for the correlation effect. The second FITS extension (ERRWEIGHT) stores the error scaling factor for each pixel in the limiting case that all valid spaxels of the cube would be coadded (see also Appendix of García-Benito et al. 2015). In the case of bad pixels, we assign an error value that is roughly ten orders of magnitude higher than the typical value.

### 5.2. Bad pixel datacubes

Bad pixel datacubes are stored in the third FITS extension (BADPIX). This information, in combination with the error vector, is essential to properly account for potential problems in each of the pixels. Pixels with flag = 1 report the absence of sufficient information in the raw data due to cosmic rays, bad CCD columns, or the effect of vignetting. They comprise a 4.2% of the total spaxels in the final datacubes. The vignetting effect imprints a characteristic inhomogeneous pattern across the FoV on the bad pixels vector. More details can be found in Fig. 11 of Husemann et al. (2013). These bad pixels have been interpolated over and we strongly suggest not to use them for any science analysis.

Finally, the uncovered corners of the hexagonal PPak FoV are filled with zeros and flagged as bad pixels for consistency.

**Table 1.** CALIFA FITS file structure.

HDU	Extension name	Format	Content
0	Primary	32-bit float	Flux density in units of $10^{-16}$ erg s $^{-1}$ cm $^{-2}$ Å $^{-1}$
1	ERROR	32-bit float	$1\sigma$ error on the flux density
2	ERRWEIGHT	32-bit float	error weighting factor
3	BADPIX	8-bit integer	Bad pixel flags (1 = bad, 0 = good)
4	FIBCOVER	8-bit integer	Number of fibers used to fill each spaxel
5	FLAT	32-bit float	SDSSflat correction to the interpolation

**Table 2.** Dimension and sampling of CALIFA datacubes.

Setup	$N_\alpha^a$	$N_\delta^a$	$N_\lambda^a$	$\lambda_{\text{start}}^b$	$\lambda_{\text{end}}^c$	$d_\lambda^d$	$\delta_\lambda^e$
V500	78	73	1877	3749 Å	7501 Å	2.0 Å	6.0 Å
V1200	78	73	1701	3650 Å	4840 Å	0.7 Å	2.3 Å
COMB	78	73	1901	3701 Å	7501 Å	2.0 Å	6.0 Å

**Notes.** <sup>(a)</sup> Number of pixels in each dimension. <sup>(b)</sup> Wavelength of the first pixel in wavelength direction. <sup>(c)</sup> Wavelength of the last pixel in wavelength direction. <sup>(d)</sup> Wavelength sampling per pixel. <sup>(e)</sup> Homogenized spectral resolution (FWHM) over the entire wavelength range.

The residuals of bright night-sky emission lines are not flagged as bad pixels.

### 5.3. Fiber coverage datacubes

The inverse-distance weighting that is used to reconstruct the data cubes means that several fibers contribute to each spaxel for most of the spaxels. As explained in García-Benito et al. (2015) we minimized the maximum distance of fibers that can contribute to the flux of a given spaxel to improve spatial resolution. However, the number of contributing fibres for spaxels at the edge of the hexagon defined by the dither pattern is lower than for spaxels inside the hexagon. As a compromise between improved spatial resolution and avoidance of information loss in the outer parts of the hexagon since pipeline V1.5, we also reduced the minimum number of fibers that contribute to each spaxel to one. This extension contains the information on the number of fibers used to compute the flux in each spaxel.

### 5.4. SDSS flat-fielding image

Pipeline V2.2 has introduced a second-order correction to the interpolation scheme that preserves the spectrophotometry at the spatial resolution of our data (see Sect. 4.2). The final correction is a multiplicative term that is stored as a 2D image/map in a new HDU (FLAT). The correction can be easily undone by the user by dividing the Primary and the ERROR datacubes by the content of this HDU without altering any of the other properties of the data cubes. This HDU is present only for those galaxies where the SDSSflat correction is applied.

### 5.5. FITS header information

The FITS header contains the standard keywords that encode the information required to transform the pixel-space coordinates into sky and wavelength-space coordinates, following the WCS. Each CALIFA datacube contains the full FITS header information of all raw frames from which it was created. Information regarding observing and instrumental conditions such as sky brightness, flexure offsets, Galactic extinction or approximate

limiting magnitude is also kept in the FITS header of each datacube. See Sect. 4.3 and Table 4 of Husemann et al. (2013) and Sect. 5.4 of García-Benito et al. (2015) for nomenclature and a summary of the main header keywords and their meaning.

The most important new keyword added in DR3 datacubes is “FLAT\_SDSS”, which takes a boolean value. It indicates whether or not the SDSS flat correction has been applied in Sect. 4.2. This information is also included in the electronic tables describing the quality of the data.

## 6. Data quality

The present third and final CALIFA data release (DR3) provides science-ready data for a sample of 646 galaxies observed in the V500 setup, 484 in V1200, and 446 combined “COMBO” cubes. As for the previous data releases, all datacubes have been checked according to a QC protocol. The DR3 QC protocol is similar in many respects to the DR2-QC. However, some modifications were introduced, which are highlighted in this section. The end products of the QC procedure are tables of flags that indicate the quality of the released data products: the observing conditions (denoted by the OBS prefix), the instrumental performance and effectiveness of the data reduction (RED), and the accuracy and quality of the final calibrated data products (CAL). QC assessments are based on measured parameters extracted by the pipeline at different stages of the reduction procedure and on visual checks of spatially-integrated spectra and wavelength-integrated synthetic images.

Each flag can have one of the following values:

- –1 = undefined;
- 0 = good quality – OK;
- 1 = minor issues that do not significantly affect the quality – WARNING;
- 2 = significant issues affecting the quality – BAD.

Flags depending on measured parameters are assigned by checking against thresholds, as detailed below and summarized in Tables 3–5. Below each flag name, in the table we also report the percentage of cubes with OK, WARNING, BAD, and undefined value, in order to provide a quantitative assessment of the impact of each flag on the QC. The thresholds are determined starting from the actual parameter distribution, so as to flag obvious outliers, by comparison with the nominal quality requirements of the survey, and by checking the impact of exceeding such thresholds on the accuracy of the wavelength and spectrophotometric calibration. Visual checks for each datacube are performed by three members of the collaboration. The median of the three independent assessments is taken as the corresponding QC flag. For flags that combine visual classifications and measured parameters, partial flags are created independently and the worst value is retained as the final flag. The tables of QC flags, along with the relevant QC parameters, are available on the DR3 website.

In naming the QC parameters, we adopt the following convention: the first part is the category prefix (OBS, RED or CAL), followed by a measured parameter, and sometimes a final suffix indicating the statistics applied to combine the parameter as measured in different observations/pointings/fibers (i.e., MEAN, MIN, MAX, RMS).

The QC of the V500 and V1200 setups is based on the same set of parameters and visual checks, except for the parameters and flags related to the spectrophotometric comparison with the SDSS (see below and Sect. 6.3), which only applies to the V500 setup. The COMBO cubes inherit all the flags from the corresponding V500 and V1200 “parent” cubes. In addition they are visually inspected to make sure that the combination process did not introduce any defects or artifacts (see Sect. 6.3).

In the following subsections, we describe the QCs in each of the above-mentioned categories. For any practical use, the definition of the flags are summarized in Tables 3–5. Parameters that either are involved in the determination of the QC flags or that can be useful for independent QC assessment by the user are released as “QCpars” tables available on the CALIFA ftp site, along with a short description of the parameters.

### 6.1. Quality of the observing conditions (OBS)

Three quantities contribute to determine the quality of the observing conditions of the CALIFA data: the airmass, the sky brightness, and the atmospheric extinction. Contrary to DR2, guided by our increased experience, in DR3 observing conditions never raise a BAD flag: in other words, poor observing conditions alone do not imply bad or unsuitable data, but just raise a WARNING.

Note that, as in DR2, we do not consider the seeing in the set of observing condition parameters to set flags, although it is included in the released “QCpars” tables. The reason for this is that, as already pointed out in García-Benito et al. (2015), given the sampling of the fibers on the plane of the sky and the resampling process, the resolution and depth of the CALIFA cubes are largely insensitive to the seeing during the observations.

For the airmass, we consider the average and the maximum airmass of the observations over all contributing pointings (OBS\_AIRMASS\_MEAN and OBS\_AIRMASS\_MAX) and its rms (OBS\_AIRMASS\_RMS). For each of these quantities, we defined a warning threshold, which is more restrictive in V1200 than in V500 due to the more demanding observing conditions for V1200 (see Tables 3 and 4). The combined FLAG\_OBS\_AM results in a WARNING as long as any of the three quantities exceeds the threshold.

The V-band surface brightness of the sky during the observations (SKYMAG) may limit the depth of the observations and the accuracy of the sky subtraction and therefore contributes in defining the quality of the observing conditions. As for DR2, the quantity SKYMAG is measured in each pointing from the sky spectrum obtained from the 36 sky fibers<sup>4</sup>. The mean and the rms of SKYMAG over all pointings are considered to define the corresponding flags. As for the airmass, stricter requirements are applied to V1200 data than to the V500 data. Note also that the thresholds are different from those adopted in DR2, as a result of our improved understanding of the impact of the sky brightness on our data.

The transparency of the sky during each pointing (EXT) is obtained from the monitored V-band extinction at the time of the observation. Large extinctions on average, a large maximum

extinction or a large rms variation across the pointings (indicating inhomogeneous observing conditions) set a WARNING flag, according to the thresholds reported in Tables 3 and 4.

### 6.2. Quality of the instrumental/data reduction performance (RED)

We assess the quality of the instrumental and data reduction performance based on four different properties as measured on the reduced data *before* combining them into the final datacube: STRaylight, spectral DISPERSION, cross dispersion CDISP, and the residuals from the subtraction of bright skylines (namely, the 5577 Å O<sub>2</sub> line in the V500 setup and the 4358 Å HgI in the V1200 setup). Moreover, we consider the limiting surface brightness corresponding to a 3- $\sigma$  detection per spaxel and spectral resolution element measured on the final datacube. Additionally we check that the final datacube does not present a pathological fraction of bad pixels, i.e. pixels characterized by large errors (5 times larger than the absolute value of the flux). Thresholds on this fraction RED\_FRAC\_BIGERR are given in order to raise a WARNING or a BAD FLAG\_RED\_ERRSPEC flag.

The straylight is an additive contribution to the raw spectra that must be removed in the data reduction process. Although the pipeline takes care of the straylight subtraction (see Appendix A.3 of Husemann et al. 2013), we have found that significant residuals that affect the final quality of the data are left whenever a frame presents high mean levels of straylight (MEANSTRAYLIGHT), as well as high maximum values (MAXSTRAYLIGHT) and large rms (RMSSTRAYLIGHT). Tables 3 and 4 report the thresholds above which a WARNING or a BAD FLAG\_RED\_STRAYLIGHT flag is set, respectively, for the three quantities in any of the contributing 2D frames (as indicated by the \_MAX suffix attached to each quantity). Note that we have modified the thresholds adopted in the DR2 to less strict values, based on the larger statistics now available and our greater experience with the data. Specifically, while the original thresholds were set based on the distributions of the parameters and the corresponding percentiles in order to filter out clear outliers, in this release we have analyzed the actual correspondence of WARNING and BAD flags to real problems in the cubes, thus realizing that the requirements on the straylight for an acceptable reduction could be relaxed.

The light from each fiber is dispersed in the wavelength direction with a given spectral DISPERSION along a trace with a finite width or cross-dispersion FWHM (CDISP). Significant departures of these two quantities from the nominal target values raise a FLAG\_RED\_DISP and a FLAG\_RED\_CDISP flag, respectively. This is done by checking the mean values (RED\_DISP\_MEAN, RED\_CDISP\_MEAN), the rms (RED\_DISP\_RMS, RED\_CDISP\_RMS), and the maximum values (RED\_DISP\_MAX, RED\_CDISP\_MAX) against the thresholds provided in Tables 3 and 4 (see footnote 10 in García-Benito et al. 2015, for more details about these quantities). Again, note that the thresholds have been modified with respect to DR2 in order to optimize the effectiveness of the flags. Specifically, we have relaxed the requirements on the maximum values and rms, which could be strongly affected by a few low-quality spaxels, even if the cube has a generally good quality.

The accuracy of the sky subtraction is quantified by the minimum and the maximum over all pointings of the average (over all fibers) flux residual of a bright skyline within an individual pointing (RED\_RES4358\_MIN and RED\_RES4358\_MAX, and RED\_RES5577\_MIN and RED\_RES5577\_MAX for the V1200

<sup>4</sup> See Appendix A.8 of Husemann et al. (2013).

**Table 3.** Definition of CALIFA DR3 quality control flags for the V500 data.

QC flag	QC parameters involved	WARNING condition(s)	BAD condition(s)	Flag definition
FLAG_OBS_AM [98.1%, 1.9%, 0.0%, 0.0%]	OBS_AIRMASS_MEAN OBS_AIRMASS_MAX OBS_AIRMASS_RMS	>2.0 >2.5 >0.15	... ... ...	Worst of the three parameters
FLAG_OBS_SKY MAG [95.5%, 4.5%, 0.0%, 0.0%]	OBS_SKY MAG_MEAN OBS_SKY MAG_RMS	<19.5 mag <sub>V</sub> arcsec <sup>-2</sup> >0.1	... ...	Worst of the two parameters
FLAG_OBS_EXT [65.2%, 4.0%, 0.0%, 30.8%]	OBS_EXT_MEAN OBS_EXT_MAX OBS_EXT_RMS	>0.30 mag >0.35 >0.10	... ... ...	Worst of the three parameters
FLAG_RED_STRAYLIGHT [84.2%, 1.7%, 3.9%, 10.2%]	RED_MEANSTRAYLIGHT_MAX RED_MAXSTRAYLIGHT_MAX RED_RMSSTRAYLIGHT_MAX	>50 counts >75 >15	>100 >150 >30	Worst of the three parameters
FLAG_RED_DISP [75.1%, 19.0%, 0.0%, 5.9%]	RED_DISP_MEAN RED_DISP_MAX RED_DISP_RMS	>5.5 Å (FWHM) >10.0 >0.5	>10 ... ...	Worst of the three parameters
FLAG_RED_CDISP [84.1%, 10.2%, 0.0%, 5.7%]	RED_CDISP_MEAN RED_CDISP_MAX RED_CDISP_RMS	>3.0 pixels (FWHM) ≥ 4.0 >0.25	... ... ...	Worst of the three parameters
FLAG_RED_SKYLINES [66.1%, 23.8%, 0.0%, 10.1%]	RED_RES5577_MIN RED_RES5577_MAX RED_RMSRES5577_MAX	<-0.1 counts >0.1 >1.0	... ... ...	Worst of the three parameters
FLAG_RED_LIMSB [86.5%, 9.9%, 2.9%, 0.6%]	RED_LIMSB	<23.0 mag <sub>AB</sub> arcsec <sup>-2</sup>	<22.5	
FLAG_RED_ERRSPEC [98.6%, 0.0%, 1.2%, 0.2%]	RED_FRAC_BIGERR	>0.4	>0.6	
FLAG_CAL_SPECPHOTO [67.6%, 21.8%, 10.5%, 0.0%]	CAL_QFLUX_G CAL_QFLUX_R CAL_QFLUX_RMS	>0.06 dex <-0.06 dex >0.06 dex <-0.06 dex >0.1	>0.097 dex <-0.097 dex >0.097 dex <-0.097 dex >0.2	Worst of the three parameters combined with visual checks on the 30"-integrated spectrum: spectral shape and comparison with SDSS photometry
FLAG_CAL_WL [92.1%, 7.4%, 0.5%, 0.0%]	CAL_RMSVELMEAN	>2.0 km s <sup>-1</sup>	>5.0	
FLAG_CAL_IMGQUAL [84.8%, 14.9%, 0.0%, 0.3%]				Visual checks on synthetic broad-band image
FLAG_CAL_SPECQUAL [94.6%, 5.4%, 0.0%, 0.0%]				Visual checks on 30"-aperture integrated spectrum
FLAG_CAL_FLATSDSS [47.1%, 43.8%, 0.3%, 8.8%]				Visual checks on SDSSflat map, and effective SDSSflat response from 30"-aperture integrated spectrum -1 if SDSSflat not applied
FLAG_CAL_REGISTRATION [52.6%, 23.4%, 0.0%, 24.0%]				Visual checks on synthetic broad-band image, SDSS footprint, and $\chi^2$ surface plots -1 if registration relative to SDSS not applied

**Notes.** Numbers in square brackets provide the percentage of released cubes with OK, WARNING, BAD, and undefined value, respectively.

and the V500 setup, respectively). We also consider the maximum over all pointings of the rms residuals (over all fibers in an individual pointing), RED\_RMSRES4358\_MAX and RED\_RMSRES5577\_MAX. Large average residuals (in absolute value) are indications of systematic bias in the sky subtraction, while large rms is a symptom of localized failures or noisy data. In these cases, the FLAG\_RED\_SKYLINES is set.

The FLAG\_RED\_LIMSB flag is used to classify the quality of datacubes based on the  $3\sigma$  continuum flux density detection limit per interpolated 1 arcsec<sup>2</sup>-spaxel and spectral resolution element. See Sect. 6.7 for a definition of the wavelength

range used to derive this quantity. Thresholds are provided in AB-magnitudes over the spectral window used for the flux integration and have been tuned slightly with respect to DR2.

### 6.3. Quality of the calibrated data products (CAL)

This part of the QC deals with the final calibrated datacubes and, although similar in many respects, has been significantly modified with respect to DR2.

A series of checks and flags are common to both the V500 and V1200 setups, namely those concerning the quality of the

**Table 4.** Definition of CALIFA DR3 quality control flags for the V1200 data.

QC flag	QC parameters involved	WARNING condition(s)	BAD condition(s)	Flag definition
FLAG_OBS_AM [89.3%, 10.7%, 0.0%, 0.0%]	OBS_AIRMASS_MEAN OBS_AIRMASS_MAX OBS_AIRMASS_RMS	>1.5 >2.0 >0.15	... ... ...	Worst of the three parameters
FLAG_OBS_SKY MAG [87.6%, 12.4%, 0.0%, 0.0%]	OBS_SKY MAG_MEAN OBS_SKY MAG_RMS	<21.5 mag <sub>V</sub> arcsec <sup>-2</sup> >0.1	... ...	Worst of the two parameters
FLAG_OBS_EXT [61.4%, 1.9%, 0.0%, 36.8%]	OBS_EXT_MEAN OBS_EXT_MAX OBS_EXT_RMS	>0.30 mag >0.35 >0.10	... ... ...	Worst of the three parameters
FLAG_RED_STRAYLIGHT [76.9%, 13.4%, 5.6%, 4.1%]	RED_MEANSTRAYLIGHT_MAX RED_MAXSTRAYLIGHT_MAX RED_RMSSTRAYLIGHT_MAX	>15 counts >20 >1.5	>30 >50 >3.0	Worst of the three parameters
FLAG_RED_DISP [78.1%, 19.2%, 0.0%, 2.7%]	RED_DISP_MEAN RED_DISP_MAX RED_DISP_RMS	>2.3 Å (FWHM) >3.0 >0.2	>2.5 ... ...	Worst of the three parameters
FLAG_RED_CDISP [68.0%, 29.3%, 0.0%, 2.7%]	RED_CDISP_MEAN	>3.0 pixels (FWHM)	...	
FLAG_RED_SKYLINES [62.6%, 33.3%, 0.0%, 4.1%]	RED_RES4358_MIN RED_RES4358_MAX RED_RMSRES4358_MAX	<-0.1 counts >0.1 >0.7	... ... ...	Worst of the three parameters
FLAG_RED_LIMSB [93.8%, 4.3%, 1.9%, 0.0%]	RED_LIMSB	<22.5 mag <sub>AB</sub> arcsec <sup>-2</sup>	<22.0	
FLAG_RED_ERRSPEC [99.8%, 0.2%, 0.0%, 0.0%]	RED_FRAC_BIGERR	>0.4	>0.6	
FLAG_CAL_WL [98.1%, 1.7%, 0.0%, 0.2%]	CAL_RMSVELMEAN	>1.0 km s <sup>-1</sup>	>2.0	
FLAG_CAL_IMGQUAL [88.8%, 9.7%, 0.0%, 1.4%]				Visual checks on synthetic broad-band image
FLAG_CAL_SPECQUAL [91.9%, 7.6%, 0.0%, 0.4%]				Visual checks on 30''-aperture integrated spectrum
FLAG_CAL_FLATSDSS [55.6%, 36.0%, 0.0%, 8.5%]				Visual checks on SDSSflat map, and effective SDSSflat response from 30''-aperture integrated spectrum -1 if SDSSflat not applied
FLAG_CAL_REGISTRATION [33.1%, 35.5%, 0.2%, 31.2%]				Visual checks on synthetic broad-band image, SDSS footprint, and χ <sup>2</sup> surface plots -1 if registration relative to SDSS not applied

**Notes.** Numbers in square brackets provide the percentage of released cubes with OK, WARNING, BAD, and undefined value, respectively.

synthetic image (FLAG\_CAL\_IMGQUAL), of the integrated 30''-aperture spectra (FLAG\_CAL\_SPECQUAL), the wavelength accuracy and stability (FLAG\_CAL\_WL), and, when applicable, the quality of the “SDSS flat-field” (see below and Sect. 5.4) and of the registration on the plane of sky relative to SDSS imaging (see below and García-Benito et al. 2015). For V500 only, in addition we perform checks on the spectrophotometric accuracy that determine the FLAG\_CAL\_SPECPHOTO flag. Additional checks are visually performed on the synthetic images and the 30''-aperture spectra for COMBO cubes. Wavelength accuracy and stability (FLAG\_CAL\_WL) are performed on COMBO cubes as well.

Visual checks on the reconstructed synthetic images in *V*-band (V500 and COMB) and *B*-band (V1200 and COMB) are encoded in the FLAG\_CAL\_IMGQUAL flag and take into account the quality of the images in absolute terms and in comparison to the corresponding *g*-band SDSS image. Reasons to raise a WARNING or a BAD flag are: holes, artefacts of any kind, irregular

background, ghosts, evident noise patterns, duplicate/offset images of the same sources, significantly elongated PSF. A BAD FLAG\_CAL\_IMGQUAL flag implies that the datacube is not suitable for any scientific use and occurs whenever multiple images are spotted, indicating a serious problem in the (relative) registration of the pointings, or when a substantial fraction of the image is affected by anomalous background subtraction. Noise patterns or background problems over a limited region are usually flagged as WARNING.

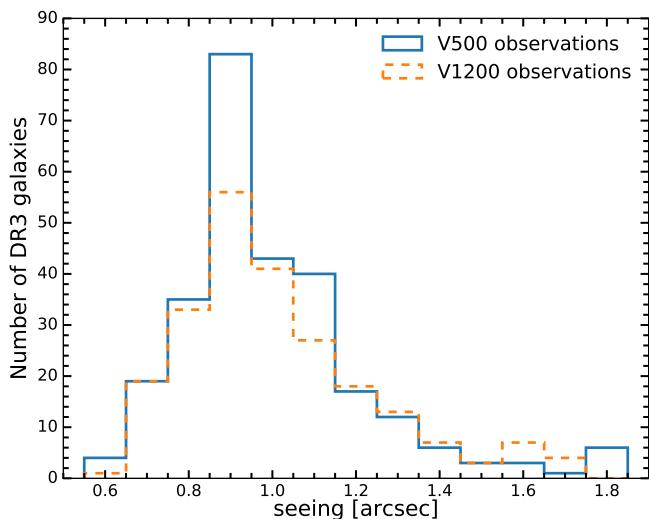
Visual checks on the 30''-aperture spectra are meant to flag anomalies in the overall spectral shape, such as bumps, drops, anomalously low S/N (possibly in limited spectral ranges), and are encoded into the FLAG\_CAL\_SPECQUAL flag. A BAD classification in this flag implies that the cube is not useful for any science and therefore is not released.

Whenever possible, pipeline V2.2 renormalizes the spectra in each spaxel to match the photometric fluxes derived from the

**Table 5.** Definition of CALIFA DR3 quality control flags *additional* for the COMB data.

QC flag	QC parameters involved	WARNING condition(s)	BAD condition(s)	Flag definition
FLAG_CAL_IMGQUAL [76.0%, 24.0%, 0.0%, 0.0%]				Visual checks on synthetic broad-band image
FLAG_CAL_SPECQUAL [96.6%, 3.4%, 0.0%, 0.0%]				Visual checks on 30"-aperture integrated spectrum
FLAG_CAL_V1200V500 [79.6%, 20.4%, 0.0%, 0.0%]				Visual checks on the match of the 30"-aperture integrated spectra in V500, V1200 and resulting COMB
FLAG_CAL_WL	CAL_RMSVELMEAN	>2.0 km s <sup>-1</sup>	>5.0	
		[97.8%, 2.2%, 0.0%, 0.0%]		

**Notes.** Numbers in square brackets provide the percentage of released cubes with OK, WARNING, BAD, and undefined value, respectively.



**Fig. 8.** Distribution of the seeing during the CALIFA observations as measured by the automatic Differential Image Motion Monitor (DIMM, Aceituno 2004).

co-registered SDSS images, by applying the SDSSflat correction. The map of the SDSS flat-fielding function is visually inspected, as well as are the synthetic images and the 30"-aperture integrated spectra before and after applying the SDSSflat. Since the reduced data cubes are already flux-calibrated before applying the SDSSflat, the correction applied in this step must be small, a few percent at most. Large correction factors and/or strong spatial variations, possibly due to bad CALIFA versus SDSS co-registration (spatial offsets or badly matched spatial resolution) and resulting in large differences between the integrated spectra before and after the SDSSflat correction, are initially marked with a BAD FLAG\_CAL\_FLATSDSS. We have then re-checked all such cases and investigated if the problem is due to an independent failure in the observations/data reduction or if it is due to the SDSSflat itself: in the latter case the datacube is re-reduced without the SDSSflat correction and the FLAG\_CAL\_FLATSDSS is assigned undefined (-1). There are cases where the problem is judged as not amendable and therefore a BAD FLAG\_CAL\_FLATSDSS is retained. If the problem is flagged as WARNING the SDSSflat correction is retained.

Pipeline V2.2 by default attempts to register the different pointings relative to the SDSS imaging before combining them. The outcome of the process is checked in the QC protocol by visually inspecting i) the footprints of the real CALIFA fibers and of the simulated fibers on the SDSS images based on the

matched astrometric solution and ii) the  $\chi^2$  surfaces that describe the goodness of the match between SDSS and CALIFA as a function of the 2D spatial offsets. These checks result in the FLAG\_CAL\_REGISTRATION flag. If a BAD condition occurs (i.e. obvious mismatch in the pointings relative to SDSS), the datacube is inspected again and re-reduced with the registration based on relative telescope offsets as in the V1.3 pipeline. Section 4.2 and García-Benito et al. (2015) contain more information about the main differences between the two methods. Should this re-reduction produce acceptable results in terms of image and spectral quality, the FLAG\_CAL\_REGISTRATION is assigned undefined (-1) and the datacube is released, otherwise the FLAG\_CAL\_IMGQUAL and/or FLAG\_CAL\_SPECQUAL flags are assigned BAD and the datacube is rejected.

The QC protocol foresees a posteriori flux calibration checks for the V500 setup only, which determine the flag FLAG\_CAL\_SPECPHOTO. A first quantitative check relies on the visual inspection of the 30"-integrated spectra, whereby the *g*- and *r*-band magnitudes derived from SDSS images integrated over the same area are transformed to flux densities and overplotted on the CALIFA spectrum: significant offsets between the SDSS points and the spectrum raise the flag. The flag is also raised in case of visual checks revealing anomalous spectral shapes (bumps, drops etc.). Finally and more quantitatively the flux ratios in *g*- and *r*-band of the different pointings relative to SDSS are considered: deviations in the mean and/or the rms of the flux ratios over the different pointings by more than given thresholds raise the corresponding flag. The flag FLAG\_CAL\_SPECPHOTO eventually reports the worst classification extracted from all these checks.

In order to check the stability of the wavelength calibration over the full spectral range we performed the same measurements as in DR1 and DR2, as described in Sect. 5.3 of Husemann et al. (2013; see also García-Benito et al. 2015): for each galaxy and setup, the spectra within 5" of the center of the galaxy are integrated and the systemic velocity is estimated first for the full spectrum and then for 3 (4) independent spectral ranges in the V1200 (V500) setup. The rms of these values with respect to the systemic velocity from the full spectrum (CAL\_RMSVELMEAN) is an estimate of the stability of the wavelength calibration across the wavelength range and is used to set the corresponding quality flag FLAG\_CAL\_WL.

#### 6.4. Overall quality assessment

The flags described in the previous sections allow any potential user to select samples that are most suitable for her/his science goals, using ad hoc selection criteria. However, we identify a set

of key flags for which a BAD classification implies unusable data in any respect: cubes with a BAD in either FLAG\_CAL\_IMGQUAL, FLAG\_CAL\_SPECQUAL, FLAG\_CAL\_WL or FLAG\_RED\_DISP are therefore excluded from DR3. The sample of galaxies with a released datacube in either one or both of the V500 and V1200 setups satisfying these criteria has been defined as the CALIFA DR3 in Sect. 2.

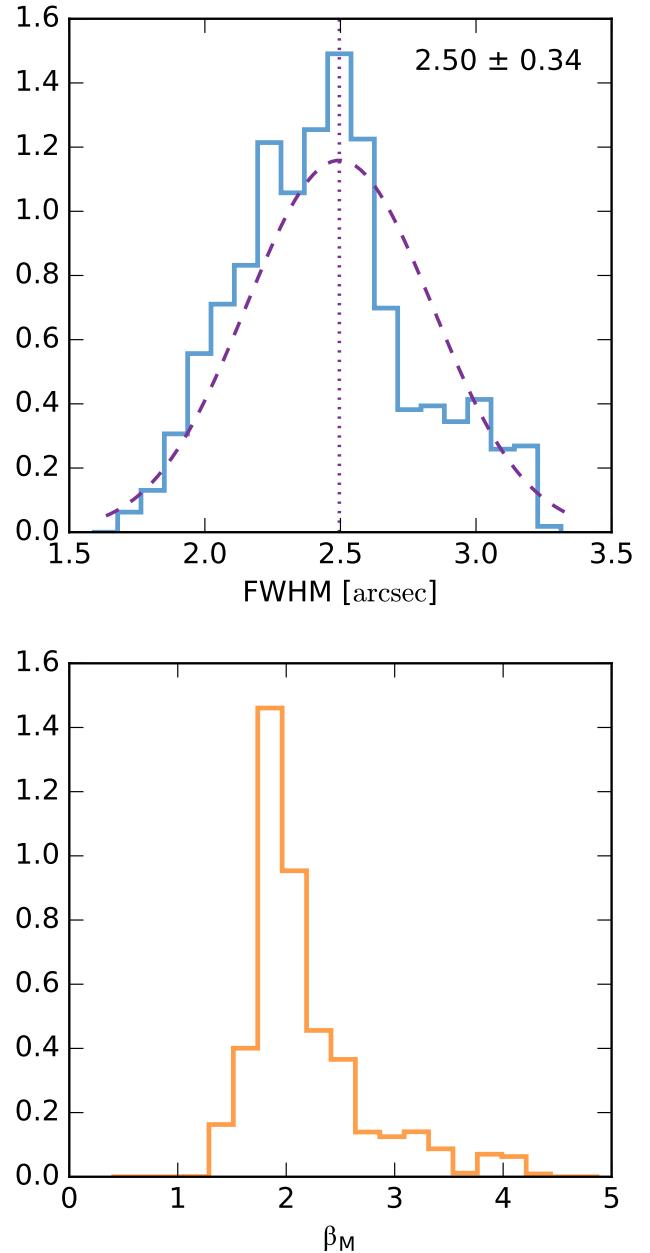
If one wishes to be more strict, one could restrict all flags to WARNING at most (value  $\leq 1$ ). Such a selection produces a sample we have called the high-quality sample (HQ sample) in section, containing 332 galaxies. An even more stringent restriction would additionally require perfect quality (0 value) in the key flags (FLAG\_CAL\_IMGQUAL, FLAG\_CAL\_SPECQUAL, FLAG\_CAL\_WL, FLAG\_RED\_DISP). Such a selection would produce a sample of 124 galaxies with the highest quality data. That sample will be limited in size and therefore of less scientific use, but could still be used as a reference sample for making sure that no data imperfections affect scientific conclusions derived from a specific method or paper.

### 6.5. Seeing and spatial resolution

The average atmospheric seeing conditions during the CALIFA observations were derived from the measurements acquired by the Differential Image Motion Monitor (DIMM, Aceituno 2004), which operates fully automatically at the Calar Alto observatory during the night. The DIMM has different operational constraints from the 3.5 m telescope (humidity lower than 80% and wind speed less than  $12 \text{ m s}^{-1}$ ). Seeing information is thus not available for every CALIFA observation, but the overall seeing distribution is not expected to be very different (see footnote 12 of García-Benito et al. 2015).

Figure 8 shows the DIMM seeing distribution for the DR3 sample, which has a median value of  $1.^{\prime}0$  FWHM, and therefore atmospheric seeing is not a limiting factor in the spatial resolution of the CALIFA cubes. Therefore, the final spatial resolution of the CALIFA data is mainly set by fiber size and the dithering and interpolation scheme.

We used the following approach to measure the PSF in the datacubes. Since January 2012 standard stars were observed using the same dithering pattern adopted for the science observations for both setups. Only a fraction of the nights had weather conditions good enough to acquire a calibration star using this scheme, yielding a total of 182 stars observed using the dithering scheme. We reduced these data using the same procedure described before for the science objects. The PSF can be measured very precisely because the calibration stars have a very high S/N. We took images based on slices of width  $400 \text{ \AA}$  in wavelength from the datacubes for each of these stars. For each of these images, we fitted a 2D Moffat profile (see Eq. (1)) using the software IMFIT (Erwin 2015). Figure 9 shows the normalized distributions of FWHM and  $\beta_M$  parameters of the Moffat profile, weighted by the likelihood of the fit, for all wavelengths and all stars. Counts are normalized to form a probability density so the integral of the histogram is 1. The fits do not show any significant wavelength dependence in any of those parameters. We obtained a mean value and  $1\sigma$  scatter of the FWHM as  $2.50 \pm 0.34$  arcsec. The distribution of  $\beta_M$  is asymmetric, so a better estimate of its value is the weighted median, which gives  $\beta_M = 2.15$ . The ellipticity ( $1 - b/a$ , with  $a$  and  $b$  being the semimajor and semiminor axes, respectively) is also measured, with mean value and  $1\sigma$  scatter of  $0.08 \pm 0.06$ . Given the uncertainties, this value means the PSF can be considered effectively

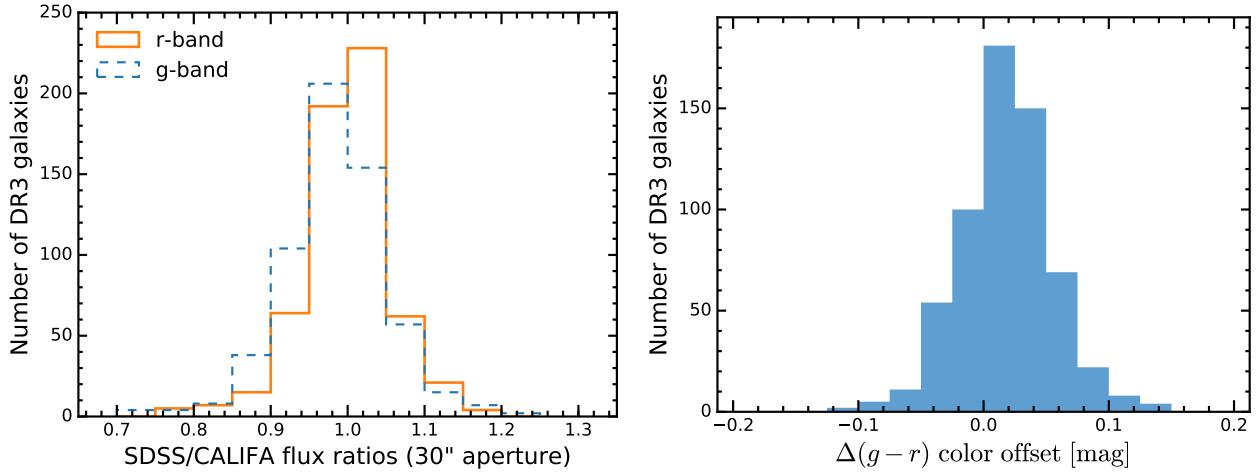


**Fig. 9.** Normalized distributions of the integrated area of the PSF FWHM (top panel) and  $\beta_M$  (bottom panel) parameters of an elliptical 2D Moffat profile fitted to 63 calibration stars, weighted by the likelihood of the fit. In the top panel the normal probability density function is marked with a dashed line and the dotted line indicates the mean value of the distribution.

axisymmetric. The uncertainties in these measurements correspond to  $1\sigma$  of the distributions. Note that the distribution is broader than that reported for DR2 (García-Benito et al. 2015) because in that release we discarded galaxies observed under observing conditions with high seeing.

### 6.6. Spectrophotometric accuracy

As described in Sect. 4.1 the registration scheme of the pipeline rescales the datacubes to the absolute flux level of the SDSS DR7 broad-band photometry, using the  $r$ -band image for the V500



**Fig. 10.** *Left panel:* distribution of the  $30''$  aperture photometry scale factor between the SDSS DR7 images and recalibrated CALIFA data. We compare the photometry only for the  $g$ - and  $r$ -bands, which are both entirely covered by the V500 wavelength range. *Right panel:* distribution of the corresponding color offset between the SDSS DR7 images and the synthetic CALIFA broadband images.

setup and the  $g$ -band image for the V1200 setup. On the other hand, for the COMBO cubes the V1200 data are finally matched to the V500 data. These procedures, together with the recalibrated sensitivity curve (see Sect. 4.2), and the updated calibration frames (master skyflats, master bias...) improves the spectrophotometric calibration of DR3 relative to DR1 and DR2.

This is clearly shown in Fig. 10. As part of the CALIFA pipeline V2.2, a  $30''$  diameter photometric aperture in  $r$  and  $g$  is measured both in the SDSS DR7 images and in the equivalent synthetic CALIFA broadband images. The mean SDSS/CALIFA  $g$ - and  $r$ -band ratios in DR3 and their scatter are  $0.99 \pm 0.09$  and  $1.00 \pm 0.08$ , respectively. In the *right panel* of Fig. 10 the distribution in  $\Delta(g - r)$  color difference between the SDSS and CALIFA data shows that the spectrophotometric accuracy over the wavelength range is better than 4%, with a median value of  $0.02 \pm 0.04$ .

We use spectral fitting methods to make an independent estimate of the spectrophotometric accuracy, following Husemann et al. (2013) and García-Benito et al. (2015). We repeat a similar experiment for the DR3 datacubes, but in this case we use the results from the fitting performed by Pipe3D (Sánchez et al. 2016). Results are shown in Fig. 11. The top panel shows in blue the mean spectrum of 251313 spatial bins of 446 galaxies included in the DR3 COMBO distribution with  $S/N > 15$  in the continuum at  $\sim 5635 \text{ \AA}$  and good quality spectral fitting. The average is taken after normalizing each spectrum by its median flux in the  $5635 \pm 45 \text{ \AA}$  window. The mean synthetic spectrum (overplotted orange line) as well as the mean residual (at the bottom of the upper panel, purple line) are also plotted. The bottom panel zooms in on the residual spectrum.

The layout of Fig. 11 is similar to Fig. 13 in Cid Fernandes et al. (2014) or Fig. 18 in García-Benito et al. (2015), to which it should be compared. Focusing on the middle panel, one sees that from  $\sim 5000 \text{ \AA}$  to the red the residuals are very similar, including the humps around  $5800 \text{ \AA}$  associated with the imperfect removal of telluric features. Toward the blue however, the reduction pipeline leads to smaller residuals than that of version V1.3c, with characteristics very similar to those of version V1.5.

In addition to the previous test, we also performed an independent estimation of the accuracy of the spectrophotometric calibration by comparing the flux-calibrated spectra of the stars observed using the three dithering procedures (Sect. 6.5) with

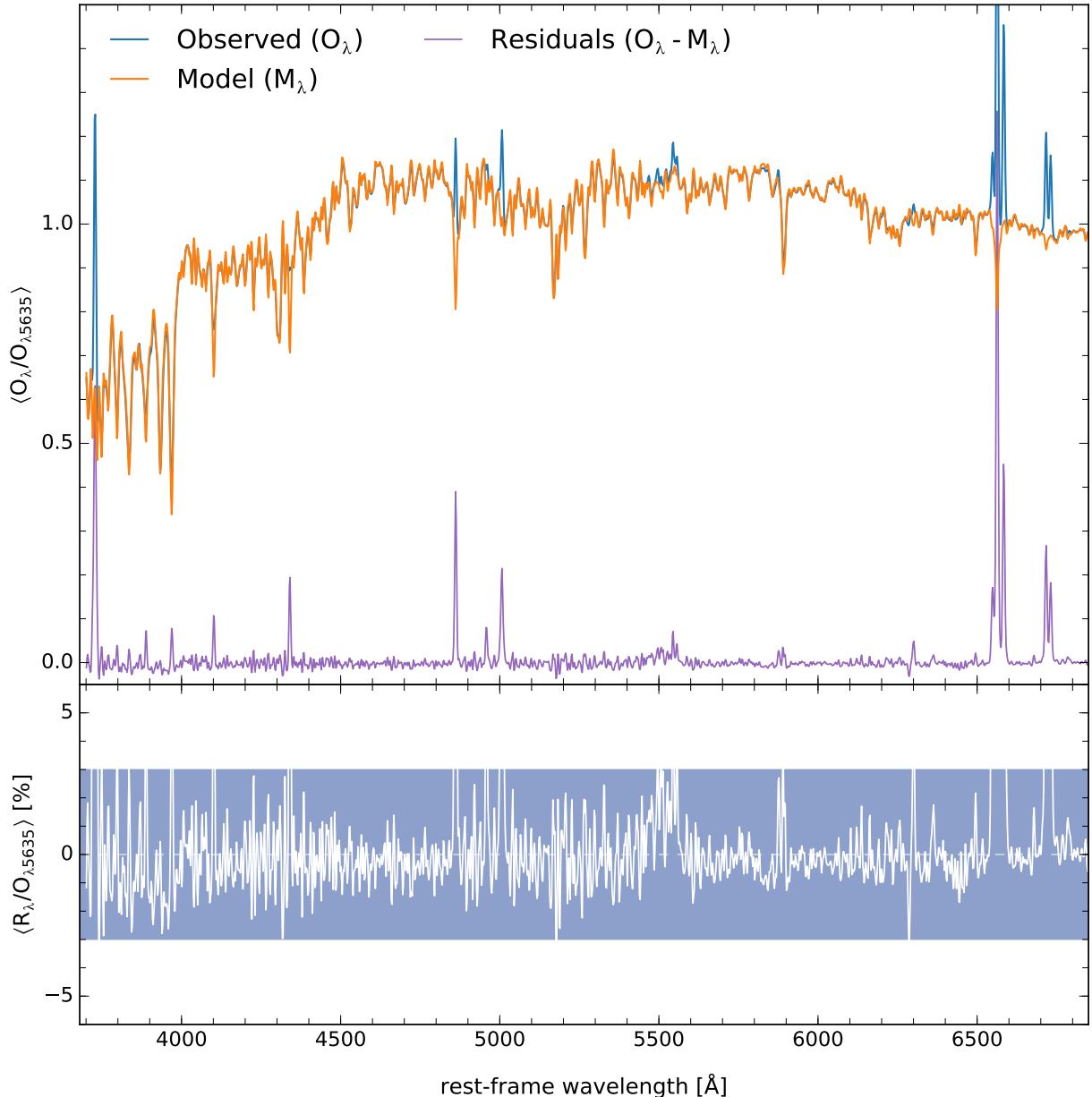
their published spectra. Since all those stars are spectrophotometric standards, they have high-S/N and good quality published spectra.

We performed the same spectrophotometric calibration for these stars than the one applied to the science cubes, using the same sensitivity curve and atmospheric extinction (Sect. 4.2). Then we extracted a  $30''$  aperture spectrum over the DAR-corrected datacube corresponding to each calibration star. Finally we compared those spectra with the published ones, deriving a ratio of  $0.99 \pm 0.10$ , as can be seen in Fig. 12. Thus, the absolute spectrophotometric accuracy is around  $\sim 10\%$ , a result that was anticipated by García-Benito et al. (2015). This was the reason why we decided not to use the calibration stars to derive the sensitivity curve. Once corrected for the absolute spectrophotometric offset, the average spectra derived for each calibration star agree with the published ones within  $\sim 3.4\%$  from blue to red. Figure 13 shows the comparison between the derived spectrum of the spectrophotometric standard star Hz44 and the published one (Oke 1990), showing a high degree of agreement.

### 6.7. Limiting sensitivity and signal-to-noise

To assess the depth of the data, we estimate the  $3\sigma$  continuum flux density detection limit per interpolated  $1 \text{ arcsec}^2$ -spaxel and spectral resolution element for the faintest regions. Figure 14 shows the limiting continuum sensitivity of the spectrophotometrically recalibrated CALIFA cubes. The depth is plotted against the average  $S/N$  per  $1 \text{ arcsec}^2$  and spectral resolution element within an elliptical annulus of  $\pm 1''$  around the galaxies'  $r$ -band half-light semimajor axis (HLR), with position angle (PA) and radius values taken from Walcher et al. (2014) when available or directly from the datacube. A narrow wavelength window at  $4480$ – $4520 \text{ \AA}$  for the V1200 and at  $5590$ – $5680 \text{ \AA}$  for the V500 is used to estimate both values. Specifically, the signal (also used for the surface brightness limit) is computed as the median value in the defined wavelength intervals, while the noise is the detrended standard deviation in the same windows. These small windows are nearly free of stellar absorption features or emission lines. The  $3\sigma$  continuum flux density detection limit per spaxel and spectral resolution element<sup>5</sup> for the V1200

<sup>5</sup> We note that this is a continuum flux density. See Note 5 of Husemann et al. (2013).



**Fig. 11.** Statistics of the spectral residuals. *Top panel:* the mean normalized spectrum of 251313 bins from 446 galaxies. The mean Pipe3D fit is overplotted in orange, while the mean residual is plotted at the bottom of the panel (purple). *Bottom panel:* zoom of the residual spectrum. The shaded rectangle encompasses the  $\pm 3\%$  area.

data ( $I_{3\sigma} = 3.0 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ arcsec}^{-2}$  in the median at 4500 Å) is a factor of  $\sim 2\text{--}3$  brighter than for the V500 data ( $I_{3\sigma} = 1.3 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ arcsec}^{-2}$  in the median at 5635 Å) mainly because of the difference in spectral resolution. These continuum sensitivities can be transformed into equivalent limiting broadband surface brightnesses of 23.0 mag arcsec $^{-2}$  in the  $g$ -band for the V1200 data and 23.4 mag arcsec $^{-2}$  in the  $r$ -band for the V500 data. The variance of the sky brightness on each night might be one of the main factors causing the difference in the limiting continuum sensitivity. Dust attenuation, transparency of the night, and other atmospheric conditions might also affect the depth achievable at fixed exposure times.

The limiting sensitivity is also a measure of the noise due to observing conditions and thus it correlates mildly with the S/N at one HLR. The mean S/N in the continuum per 1 arcsec $^2$  and

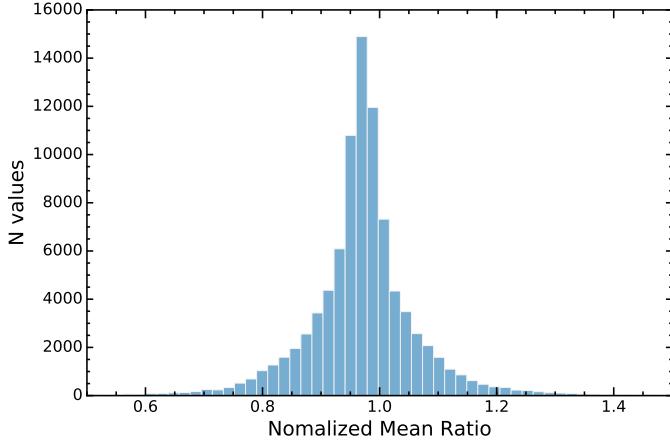
spectral resolution element at the HLR along the semimajor axis for all objects is  $\sim 9.4$  for the V1200 setup, while it is  $\sim 21.2$  for the V500 setup. Thus, we achieve a  $S/N \gtrapprox 10$  at 1 HLR for a significant number of the objects for the V500 setup ( $\sim 85\%$ ) and even for the V1200 setup ( $\sim 40\%$ ).

## 7. Access to the CALIFA DR3 data

### 7.1. The CALIFA DR3 search and retrieval tool

The public data are distributed through the CALIFA DR3 web page<sup>6</sup>. A simple web form interface, already in use for DR1 and DR2, allows the user to select data of a particular target galaxy,

<sup>6</sup> <http://califa.caha.es/DR3>



**Fig. 12.** Histogram of the differences between flux intensities for the published spectra for the spectrophotometric calibration stars described in Sect. 6.5 and the corresponding ones derived by CALIFA.

or a sub-sample of objects within some constraints on observing conditions or galaxy properties. Among the selection parameters, we include the instrument setup, galaxy name and coordinates, redshift,  $g$ -band magnitudes, Hubble type, bar strength and whether or not it is a clearly merging system.

If any CALIFA datasets are available given the search parameters, they are listed in the search results and can be selected for download. The download process requests a target directory on the local machine to store the data, after the downloading option is selected. The CALIFA data are delivered as fully reduced datacubes in FITS format separately for each of the three configurations: V500, V1200 and COMBO. Each DR3 datacube is uniquely identified by its file name, `GALNAME.V1200.rscube.fits.gz`, `GALNAME.V500.rscube.fits.gz` and `GALNAME.COMB.rscube.fits.gz` for the V1200, V500 and COMBO configurations respectively, where `GALNAME` is the CALIFA ID number listed in the electronically distributed tables.

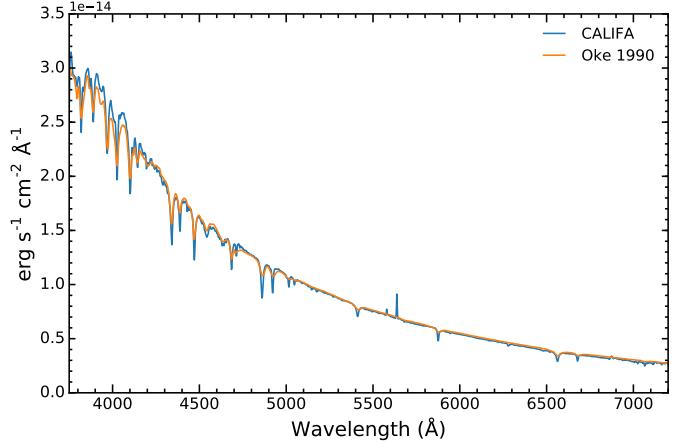
All the QC tables discussed throughout this article are also distributed in CSV and FITS-table formats on the DR3 webpage. In addition, we distribute the tables discussed in Walcher et al. (2014) and in Sect. 2 regarding the characterization of the Main and extension samples, using similar formats. These tables could be useful for further science exploitation of the datacubes.

## 7.2. Virtual observatory services

Just like the previous data releases, the CALIFA DR3 also interfaces with Virtual Observatory (VO) facilities. At release time or shortly thereafter, the datacubes will be made available through the Obscore data model, and as database tables of voxels in the TAP service `ivo://dc.gavo.dc/tap`<sup>7</sup>. The service also contains tables of objects and QC parameters. These and further services can also be found by searching for CALIFA DR3 with any registry client.

## 8. Summary

In this article we have presented the main characteristics of the Third public data release, DR3, of the Calar Alto Legacy



**Fig. 13.** Comparison between the published spectrum for the spectrophotometric standard star Hz44 and the average of the 15 spectra derived by CALIFA for the same star.

Integral Field Area (CALIFA) survey. DR3 comprises 667 galaxies (1576 datacubes) containing information from more than 1.5 million independent spectra, covering a wide range of masses, morphological types, and colors. The released datacubes correspond to two different sample of galaxies: i) the main sample, a randomly selected subsample of the CALIFA mother sample comprising 529 galaxies and representative of galaxies in the local Universe and ii) an Extension Sample comprising a heterogeneous collection of galaxies observed with the same setup that adds objects that are rare in the overall galaxy population and therefore not numerous or absent in the mother sample. The CALIFA DR3 provides science-grade and quality-checked integral-field spectroscopic data to the community<sup>8</sup>.

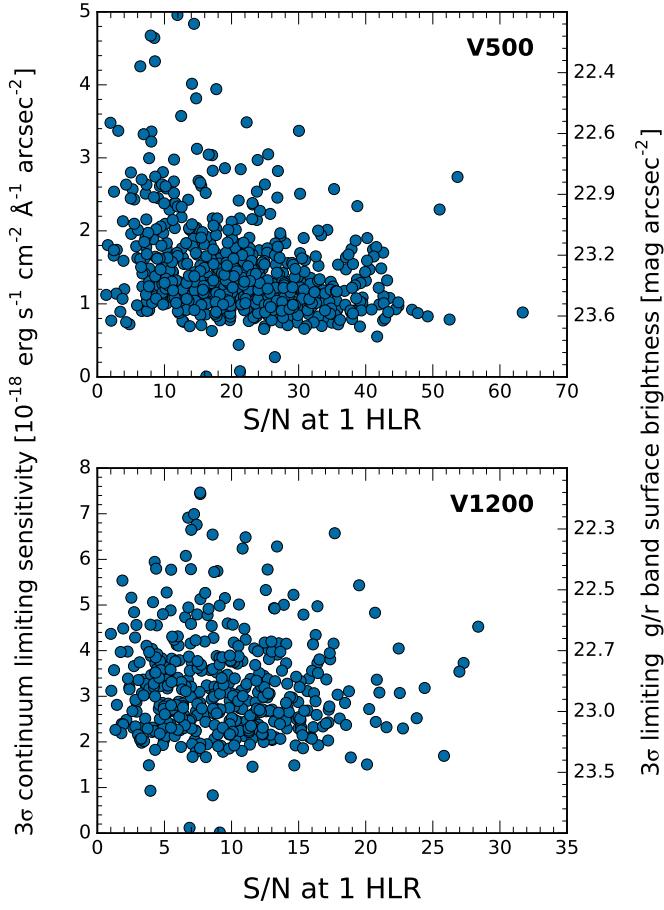
We reduced the data using a new version of the pipeline (V2.2), which slightly improves the quality of the data in terms of: i) reliability of the spatial registration process; ii) the homogeneity in the data reduction; and iii) the quality of the image reconstruction. We described in detail the main quality parameters analyzed in the validation process, which are provided to users with complete tables to select the most suitable objects for their science cases.

Compared with other ongoing major surveys, CALIFA offers a similar projected spatial resolution. The PSF of the datacubes has a mean value of  $\sim 2.5''$  (Sect. 6.5), similar to those reported SAMI (Sharp et al. 2015) and MaNGA (Law et al., in prep.). However, CALIFA galaxies are observed at lower redshift and with a physically larger IFU, thus providing better spatial coverage and resolution, as well as better overall S/N. CALIFA thus offers a highly competitive compromise for analyzing the spatially resolved structures in galaxies. The penalty for this wider coverage is the lower number of galaxies observed (6 times lower than what is projected for SAMI and 15 times lower than the goals of MaNGA), and a lower spectral resolution over the full wavelength range.

While the CALIFA data distributed in this final DR have already been used for a variety of science applications, the potential for new scientific exploitation of the data is still very high. As CALIFA has been conceived as a legacy survey, we sincerely hope that the data will be useful to the community in years to come.

<sup>7</sup> Accessible at <http://dc.gavo.dc/tap>

<sup>8</sup> <http://califa.caha.es/DR3>



**Fig. 14.** Limiting  $3\sigma$  continuum sensitivity per spaxel and spectral resolution element as a function of the average continuum S/N at the half-light radius (HLR). The corresponding broadband surface brightness limits in  $r$  (V500) and  $g$  (V1200) are indicated on the right  $y$ -axis. The limiting continuum sensitivity and the S/N were computed from the median signal and noise in the wavelength region 4480–4520 Å and 5590–5680 Å for the V1200 and V500 data, respectively.

**Acknowledgements.** CALIFA is the first legacy survey being performed at Calar Alto. The CALIFA collaboration would like to thank the IAA-CSIC and MPIA-MPG as major partners of the observatory, and CAHA itself, for the unique access to telescope time and support in manpower and infrastructures. The CALIFA collaboration thanks also the CAHA staff for the dedication to this project. We thank the anonymous referee for his/her help in improving this article. SFS thanks the director of CEFCA, M. Moles, for his sincere support to this project. SFS thanks the CONACYT-125180 and DGAPA-IA100815 projects for providing him support in this study. R.G.B., R.G.D., and E.P. are supported by grants AYA2014-57490-P and JA-FQM-2828. SZ is supported by the EU Marie Curie Integration Grant “SteMaGE” No. PCIG12-GA-2012-326466 (Call Identifier: FP7-PEOPLE-2012 CIG). J. F.-B. from grant AYA2013-48226-C3-1-P from the Spanish Ministry of Economy and Competitiveness (MINECO), as well as from the FP7 Marie Curie Actions of the European Commission, via the Initial Training Network DAGAL under REA grant agreement 289313. B.G.-L. acknowledges financial support by the Spanish MINECO under grants AYA2013-41656-P and AYA2015-68217-P. Support for L.G. is provided by the Ministry of Economy, Development, and Tourism’s Millennium Science Initiative through grant IC12009, awarded to The Millennium Institute of Astrophysics, MAS. L.G. also acknowledges support by CONICYT through FONDECYT grant 3140566, and AYA2013-42227-P from the Spanish Ministerio de Ciencia e Innovación and TIC 114 and PO08-TIC-3531 from Junta de Andalucía. AG acknowledges support from the FP7/2007–2013 under grant agreement no. 267251 (AstroFI). RAM was funded by the Spanish programme of International Campus of Excellence Moncloa (CEI). JMA acknowledges support from the European Research Council Starting Grant (SEDMorph; P.I. V. Wild). I.M. and A.d.O. acknowledge the support by the projects AYA2010-15196 from the Spanish Ministerio de Ciencia e Innovación and TIC 114 and PO08-TIC-3531 from Junta de

Andalucía. AMI acknowledges support from Agence Nationale de la Recherche through the STILISM project (ANR-12-BS05-0016-02). M.M. acknowledges financial support from AYA2010-21887-C04-02 from the Ministerio de Economía y Competitividad. PSB acknowledges support from the Ramón y Cajal program, grant ATA2010-21322-C03-02 from the Spanish Ministry of Economy and Competitiveness (MINECO). C.J.W. acknowledges support through the Marie Curie Career Integration Grant 303912. V.W. acknowledges support from the European Research Council Starting Grant (SEDMorph P.I. V. Wild) and European Career Re-integration Grant (Phiz-Ev P.I. V. Wild). YA acknowledges financial support from the *Ramón y Cajal* programme (RyC-2011-09461) and project AYA2013-47742-C4-3-P, both managed by the *Ministerio de Economía y Competitividad*, as well as the “Study of Emission-Line Galaxies with Integral-Field Spectroscopy” (SELGIFS) programme, funded by the EU (FP7-PEOPLE-2013-IRSES-612701) within the Marie-Sklodowska-Curie Actions scheme. ROM acknowledges support from CAPES (Brazil) through a PDJ fellowship from project 88881.030413/2013-01, program CSF-PVE.

## References

- Aceituno, J. 2004, Calar Alto Newsletter No. 8, <http://www.caha.es/newsletter/news04b/Aceituno/Newsletter.html>
- Aguerri, J. A. L., Méndez-Abreu, J., Falcón-Barroso, J., et al. 2015, *A&A*, **576**, A102
- Barbon, R., Buondi, V., Cappellaro, E., & Turatto, M. 1999, *A&AS*, **139**, 531
- Barrera-Ballesteros, J. K., Falcón-Barroso, J., García-Lorenzo, B., et al. 2014, *A&A*, **568**, A70
- Barrera-Ballesteros, J. K., García-Lorenzo, B., Falcón-Barroso, J., et al. 2015a, *A&A*, **582**, A21
- Barrera-Ballesteros, J. K., Sánchez, S. F., García-Lorenzo, B., et al. 2015b, *A&A*, **579**, A45
- Bershady, M. A., Verheijen, M. A. W., Swaters, R. A., et al. 2010, *ApJ*, **716**, 198
- Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, *AJ*, **129**, 2562
- Bryant, J. J., Owers, M. S., Robotham, A. S. G., et al. 2015, *MNRAS*, **447**, 2857
- Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, *ApJ*, **798**, 7
- Cano-Díaz, M., Sánchez, S. F., Zibetti, S., et al. 2016, *ApJ*, **821**, L26
- Cappellari, M., & Copin, Y. 2003, *MNRAS*, **342**, 345
- Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, *MNRAS*, **413**, 813
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, **345**, 245
- Catalán-Torrecilla, C., Gil de Paz, A., Castillo-Morales, A., et al. 2015, *A&A*, **584**, A87
- Chabrier, G. 2003, *PASP*, **115**, 763
- 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
- Croom, S. M., Lawrence, J. S., Bland-Hawthorn, J., et al. 2012, *MNRAS*, **421**, 872
- Davies, R. L., Kewley, L. J., Ho, I.-T., & Dopita, M. A. 2014, *MNRAS*, **444**, 3961
- De Geyter, G., Baes, M., Camps, P., et al. 2014, *MNRAS*, **441**, 869
- Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, *Astron. Geophys.*, **50**, 5.12
- Erwin, P. 2015, *ApJ*, **799**, 226
- Falcón-Barroso, J., Lyubenova, M., & van de Ven, G. 2015, in IAU Symp. 311, eds. M. Cappellari, & S. Courteau, 78
- Galbany, L., Stanishev, V., Mourão, A. M., et al. 2014, *A&A*, **572**, A38
- Galbany, L., Stanishev, V., Mourão, A. M., et al. 2016, *A&A*, **591**, A48
- García-Benito, R., Zibetti, S., Sánchez, S. F., et al. 2015, *A&A*, **576**, A135
- García-Lorenzo, B., Márquez, I., Barrera-Ballesteros, J. K., et al. 2015, *A&A*, **573**, A59
- Gomes, J. M., Papaderos, P., Kehrig, C., et al. 2016, *A&A*, **588**, A68
- González Delgado, R. M., Cid Fernandes, R., García-Benito, R., et al. 2014a, *ApJ*, **791**, L16
- González Delgado, R. M., Pérez, E., Cid Fernandes, R., & et al. 2014b, *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
- Greisen, E. W., & Calabretta, M. R. 2002, *A&A*, **395**, 1061
- Haines, T., McIntosh, D. H., Sánchez, S. F., Tremonti, C., & Rudnick, G. 2015, *MNRAS*, **451**, 433
- Ho, I.-T., Kudritzki, R.-P., Kewley, L. J., et al. 2015, *MNRAS*, **448**, 2030
- Holmes, L., Spekkens, K., Sánchez, S. F., et al. 2015, *MNRAS*, **451**, 4397
- Holwerda, B. W., & Keel, W. C. 2013, *A&A*, **556**, A42
- Horne, K. 1986, *PASP*, **98**, 609
- Hubble, E. 1936, *ApJ*, **84**, 517
- Hubble, E. P. 1926, *ApJ*, 63

- Husemann, B., Kamann, S., Sandin, C., et al. 2012, [A&A](#), **545**, A137  
 Husemann, B., Jahnke, K., Sánchez, S. F., et al. 2013, [A&A](#), **549**, A87  
 Ibarra-Mendel, H. J., Sánchez, S. F., Avila-Reese, V., et al. 2016, MNRAS, accepted [[arXiv:1609.01304](#)]  
 Iglesias-Páramo, J., Vilchez, J. M., Galbany, L., et al. 2013, [A&A](#), **553**, L7  
 Iglesias-Páramo, J., Vilchez, J. M., Rosales-Ortega, F. F., et al. 2016, ApJ, 826, 71  
 Kalinova, V., Colombo, D., Rosolowsky, E., et al. 2015, MNRAS, submitted [[arXiv:1509.03352](#)]  
 Kehrig, C., Monreal-Ibero, A., Papaderos, P., et al. 2012, [A&A](#), **540**, A11  
 Kelz, A., Verheijen, M. A. W., Roth, M. M., et al. 2006, [PASP](#), **118**, 129  
 López Fernández, R., Cid Fernandes, R., González Delgado, R. M., et al. 2016, [MNRAS](#), **458**, 184  
 Lyubenova, M., Martín-Navarro, I., van de Ven, G., et al. 2016, MNRAS, submitted [[arXiv:1606.07448](#)]  
 Marino, R. A., Rosales-Ortega, F. F., Sánchez, S. F., et al. 2013, [A&A](#), **559**, A114  
 Marino, R. A., Gil de Paz, A., Sánchez, S. F., et al. 2016, [A&A](#), **585**, A47  
 Mármol-Queraltó, E., Sánchez, S. F., Marino, R. A., et al. 2011, [A&A](#), **534**, A8  
 Martín-Navarro, I., Vazdekis, A., La Barbera, F., et al. 2015, [ApJ](#), **806**, L31  
 Martínez-García, E. E., Puerari, I., Rosales-Ortega, F. F., et al. 2014, [ApJ](#), **793**, L19  
 Mast, D., Rosales-Ortega, F. F., Sánchez, S. F., et al. 2014, [A&A](#), **561**, A129  
 Oke, J. B. 1990, [AJ](#), **99**, 1621  
 Pagel, B. E. J., & Edmunds, M. G. 1981, [ARA&A](#), **19**, 77  
 Papaderos, P., Gomes, J. M., Vilchez, J. M., et al. 2013, [A&A](#), **555**, L1  
 Peletier, R. F. 1989, Ph.D. Thesis, University of Groningen, The Netherlands  
 Pérez, E., Cid Fernandes, R., González Delgado, R. M., et al. 2013, [ApJ](#), **764**, L1  
 Roche, N., Humphrey, A., Gomes, J. M., et al. 2015, [MNRAS](#), **453**, 2349  
 Roth, M. M., Kelz, A., Fechner, T., et al. 2005, [PASP](#), **117**, 620  
 Ruiz-Lara, T., Pérez, I., Florido, E., et al. 2016, [MNRAS](#), **456**, L35  
 Sánchez, S., Sánchez-Menguiano, L., Marino, R., et al. 2015, [Galaxies](#), **3**, 164  
 Sánchez, S. F. 2006, [Astron. Nachr.](#), **327**, 850  
 Sánchez, S. F. 2015, in Galaxies in 3D across the Universe, eds. B. L. Ziegler, F. Combes, H. Dannerbauer, & M. Verdugo, IAU Symp., 309, 85  
 Sánchez, S. F., Cardiel, N., Verheijen, M. A. W., Pedraz, S., & Covone, G. 2007, [MNRAS](#), **376**, 125  
 Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012a, [A&A](#), **538**, A8 (S12)  
 Sánchez, S. F., Rosales-Ortega, F. F., Marino, R. A., et al. 2012b, [A&A](#), **546**, A2  
 Sánchez, S. F., Rosales-Ortega, F. F., Jungwiert, B., et al. 2013, [A&A](#), **554**, A58  
 Sánchez, S. F., Rosales-Ortega, F. F., Iglesias-Páramo, J., et al. 2014, [A&A](#), **563**, A49  
 Sánchez, S. F., Pérez, E., Sánchez-Blázquez, P., et al. 2016, [Rev. Mex. Astron. Astrofis.](#), **52**, 171  
 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  
 Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, [ApJ](#), **500**, 525  
 Schmidt, M. 1968, [ApJ](#), **151**, 393  
 Sharp, R., Allen, J. T., Fogarty, L. M. R., et al. 2015, [MNRAS](#), **446**, 1551  
 Singh, R., van de Ven, G., Jahnke, K., Lyubenova, M., & et al. 2013, [A&A](#), **558**, A43  
 van den Bosch, R. C. E., Gebhardt, K., Gültekin, K., Yıldırım, A., & Walsh, J. L. 2015, [ApJS](#), **218**, 10  
 Verheijen, M. A. W., Bershady, M. A., Andersen, D. R., et al. 2004, [Astron. Nachr.](#), **325**, 151  
 Walcher, C. J., Wisotzki, L., Bekeraïté, S., et al. 2014, [A&A](#), **569**, A1 (W14)  
 Wild, V., Rosales-Ortega, F., Falcón-Barroso, J., et al. 2014, [A&A](#), **567**, A132  
 Yıldırım, A., van den Bosch, R. C. E., van de Ven, G., et al. 2016, [MNRAS](#), **456**, 538  
 Yıldırım, A., van den Bosch, R. C. E., van de Ven, G., et al. 2015, [MNRAS](#), **452**, 1792  
 York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, [AJ](#), **120**, 1579
- 
- <sup>1</sup> Instituto de Astronomía, Universidad Nacional Autónoma de México, A.P. 70-264, 04510, México, D.F.  
 e-mail: [sfsanchez@astro.unam.mx](mailto:sfsanchez@astro.unam.mx)
- <sup>2</sup> Instituto de Astrofísica de Andalucía (IAA/CSIC), Glorieta de la Astronomía s/n Aptdo. 3004, 18080 Granada, Spain,
- <sup>3</sup> INAF-Osservatorio Astrofisico di Arcetri – Largo Enrico Fermi, 5, 50125 Firenze, Italy
- <sup>4</sup> Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
- <sup>5</sup> European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching b. München, Germany
- <sup>6</sup> Millennium Institute of Astrophysics, Universidad de Chile, Santiago, Chile
- <sup>7</sup> Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile
- <sup>8</sup> Instituto de Astrofísica de Canarias, Vía Láctea s/n, 38205 La Laguna, Tenerife, Spain
- <sup>9</sup> University of Vienna, Department of Astrophysics, Türkenschanzstr. 17, 1180 Vienna, Austria
- <sup>10</sup> Departamento de Astrofísica, Universidad de La Laguna, 38205 La Laguna, Tenerife, Spain
- <sup>11</sup> Observatorio Astronómico, Laprida 854, X5000BGR, Córdoba, Argentina
- <sup>12</sup> Consejo de Investigaciones Científicas y Técnicas de la República Argentina, Avda. Rivadavia 1917, C1033AAJ CABA, Argentina
- <sup>13</sup> Centro Astronómico Hispano Alemán de Calar Alto (CSIC-MPG), C/ Jesús Durbán Remón 2–2, 4004 Almería, Spain
- <sup>14</sup> Departamento de Física, Universidade Federal de Santa Catarina, PO Box 476, 88040-900 Florianópolis, SC, Brazil
- <sup>15</sup> Centro de Astrobiología (CSIC-INTA), Depto. Astrofísica, ESAC Campus, 28691 Villanueva de la Cañada, Madrid, Spain
- <sup>16</sup> Department of Physics & Astronomy, Johns Hopkins University, Bloomberg Center, 3400 N. Charles St., Baltimore, MD 21218, USA
- <sup>17</sup> Sydney Institute for Astronomy, School of Physics, University of Sydney, NSW 2006, Australia
- <sup>18</sup> Instituto de Física e Química, Universidade Federal de Itajubá, Av. BPS, 1303, 37500-903 Itajubá-MG, Brazil
- <sup>19</sup> Universität Heidelberg, Zentrum für Astronomie, Astronomisches Rechen-Institut, Mönchhofstraße 12–14, 69120 Heidelberg, Germany
- <sup>20</sup> Departamento de Física Teórica, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain
- <sup>21</sup> Astronomisches Institut, Ruhr-Universität Bochum, Universitätsstr. 150, 44801 Bochum, Germany
- <sup>22</sup> School of Physics and Astronomy, University of St. Andrews, SUPA, North Haugh, KY16 9SS, St. Andrews, UK
- <sup>23</sup> Departamento de Física Teórica y del Cosmos, University of Granada, Facultad de Ciencias (Edificio Mecenas), 18071 Granada, Spain
- <sup>24</sup> Departamento de Astrofísica y CC. de la Atmósfera, Universidad Complutense de Madrid, 28040 Madrid, Spain
- <sup>25</sup> Department of Physics, Royal Military College of Canada, PO box 17000, Station Forces, Kingston, Ontario, Canada, K7K 7B4
- <sup>26</sup> Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
- <sup>27</sup> Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, The Netherlands
- <sup>28</sup> Institut für Astronomie (ETH), Wolfgang-Pauli-Str. 27, 8093 Zürich, Switzerland
- <sup>29</sup> CIEMAT, Avda. Complutense 40, 28040 Madrid, Spain
- <sup>30</sup> GEPI, Observatoire de Paris, CNRS, Université Paris Diderot, Place Jules Janssen, 92190 Meudon, France
- <sup>31</sup> Observatorio do Valongo, Universidade Federal do Rio de Janeiro, Ladeira do Pedro Antônio 43, Saude, 20080-090 Rio de Janeiro, Brazil
- <sup>32</sup> Departamento de Astronomía, Universidad de Guanajuato, Apartado Postal 144, 36000 Guanajuato, Guanajuato, Mexico
- <sup>33</sup> Instituto Nacional de Astrofísica, Óptica y Electrónica, Luis E. Erro 1, 72840 Tonantzintla, Puebla, Mexico
- <sup>34</sup> Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile
- <sup>35</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany



# Referências Bibliográficas

- Albareti, F. D., Allende Prieto, C., Almeida, A., Anders, F., Anderson, S., Andrews, B. H., Aragón-Salamanca, A., Argudo-Fernández, M. et al. 2017, ApJS, 233, 25
- Asari, N. V., Cid Fernandes, R., Stasińska, G., Torres-Papaqui, J. P., Mateus, A., Sodré, L., Schoenell, W., & Gomes, J. M. 2007, MNRAS, 381, 263
- Brinchmann, J., Charlot, S., White, S. D. M., Tremonti, C., Kauffmann, G., Heckman, T., & Brinkmann, J. 2004, MNRAS, 351, 1151
- Cid Fernandes, R., Asari, N. V., Sodré, L., Stasińska, G., Mateus, A., Torres-Papaqui, J. P., & Schoenell, W. 2007, MNRAS, 375, L16
- Cid Fernandes, R., Mateus, A., Sodré, L., Stasińska, G., & Gomes, J. M. 2005, MNRAS, 358, 363
- Cid Fernandes, R., Stasińska, G., Mateus, A., & Vale Asari, N. 2011, MNRAS, 413, 1687
- Cid Fernandes, R., Stasińska, G., Schlickmann, M. S., Mateus, A., Vale Asari, N., Schoenell, W., & Sodré, L. 2010, MNRAS, 403, 1036
- de Amorim, A. L. 2015, PhD thesis, Curso de Pós-Graduação em Física da Universidade Federal de Santa Catarina
- de Amorim, A. L., García-Benito, R., Cid Fernandes, R., Cortijo-Ferrero, C., González Delgado, R. M., Lacerda, E. A. D., López Fernández, R., Pérez, E. et al. 2017, MNRAS, 471, 3727
- Dettmar, R.-J. 1990, A&A, 232, L15
- García-Benito, R., González Delgado, R. M., Pérez, E., Cid Fernandes, R., Cortijo-Ferrero, C., López Fernández, R., de Amorim, A. L., Lacerda, E. A. D. et al. 2017, ArXiv e-prints

- García-Benito, R., Zibetti, S., Sánchez, S. F., Husemann, B., de Amorim, A. L., Castillo-Morales, A., Cid Fernandes, R., Ellis, S. C. et al. 2015, *A&A*, 576, A135
- González Delgado, R. M., Cid Fernandes, R., García-Benito, R., Pérez, E., de Amorim, A. L., Cortijo-Ferrero, C., Lacerda, E. A. D., López Fernández, R. et al. 2014a, *ApJ*, 791, L16
- González Delgado, R. M., Cid Fernandes, R., Pérez, E., García-Benito, R., López Fernández, R., Lacerda, E. A. D., Cortijo-Ferrero, C., de Amorim, A. L. et al. 2016, *A&A*, 590, A44
- González Delgado, R. M., García-Benito, R., Pérez, E., Cid Fernandes, R., de Amorim, A. L., Cortijo-Ferrero, C., Lacerda, E. A. D., López Fernández, R. et al. 2015, *A&A*, 581, A103
- González Delgado, R. M., Pérez, E., Cid Fernandes, R., García-Benito, R., de Amorim, A. L., Sánchez, S. F., Husemann, B., Cortijo-Ferrero, C. et al. 2014b, *A&A*, 562, A47
- González Delgado, R. M., Pérez, E., Cid Fernandes, R., García-Benito, R., López Fernández, R., Vale Asari, N., Cortijo-Ferrero, C., de Amorim, A. L. et al. 2017, *A&A*, 607, A128
- Green, A. W., Croom, S. M., Scott, N., Cortese, L., Medling, A. M., D'Eugenio, F., Bryant, J. J., Bland-Hawthorn, J. et al. 2017, ArXiv e-prints
- Haffner, L. M., Dettmar, R.-J., Beckman, J. E., Wood, K., Slavin, J. D., Giannanco, C., Madsen, G. J., Zurita, A. et al. 2009, *Reviews of Modern Physics*, 81, 969
- 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
- Husemann, B., Jahnke, K., Sánchez, S. F., Barrado, D., Bekerait\*error\*é, S., Bomans, D. J., Castillo-Morales, A., Catalán-Torrecilla, C. et al. 2013, *A&A*, 549, A87
- Kauffmann, G., Heckman, T. M., White, S. D. M., Charlot, S., Tremonti, C., Peng, E. W., Seibert, M., Brinkmann, J. et al. 2003, *MNRAS*, 341, 54
- Lacerda, E. A. D., Cid Fernandes, R., Couto, G. S., Stasińska, G., García-Benito, R., Vale Asari, N., Perez, E., Gonzalez Delgado, R. M. et al. 2017, ArXiv e-prints
- Lacerda, E. A. D., Cid Fernandes, R., Couto, G. S., Stasińska, G., García-Benito, R., Vale Asari, N., Pérez, E., González Delgado, R. M. et al. 2018, *MNRAS*, 474, 3727
- Mateus, A., Sodré, L., Cid Fernandes, R., Stasińska, G., Schoenell, W., & Gomes, J. M. 2006, *MNRAS*, 370, 721

- Moles, M., Benítez, N., Aguerri, J. A. L., Alfaro, E. J., Broadhurst, T., Cabrera-Caño, J., Castander, F. J., Cepa, J. et al. 2008, AJ, 136, 1325
- Oey, M. S., Meurer, G. R., Yelda, S., Furst, E. J., Caballero-Nieves, S. M., Hanish, D. J., Levesque, E. M., Thilker, D. A. et al. 2007, ApJ, 661, 801
- Pérez, E., Cid Fernandes, R., González Delgado, R. M., García-Benito, R., Sánchez, S. F., Husemann, B., Mast, D., Rodón, J. R. et al. 2013, ApJ, 764, L1
- Reynolds, R. J. 1971, PhD thesis, THE UNIVERSITY OF WISCONSIN - MADISON.
- 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
- Rousseau-Nepton, L., Robert, C., Drissen, L., Martin, R. P., & Martin, T. 2017, ArXiv e-prints
- Sánchez, S. F., Galbany, L., Pérez, E., Sánchez-Blázquez, P., Falcón-Barroso, J., Rosales-Ortega, F. F., Sánchez-Menguiano, L., Marino, R. et al. 2015, A&A, 573, A105
- Sánchez, S. F., García-Benito, R., Zibetti, S., Walcher, C. J., Husemann, B., Mendoza, M. A., Galbany, L., Falcón-Barroso, J. et al. 2016, A&A, 594, A36
- Scoville, N., Aussel, H., Brusa, M., Capak, P., Carollo, C. M., Elvis, M., Giavalisco, M., Guzzo, L. et al. 2007, ApJS, 172, 1
- Stasińska, G., Cid Fernandes, R., Mateus, A., Sodré, L., & Asari, N. V. 2006, MNRAS, 371, 972
- Stasińska, G., Vale Asari, N., Cid Fernandes, R., Gomes, J. M., Schlickmann, M., Mateus, A., Schoenell, W., Sodré, Jr., L. et al. 2008, MNRAS, 391, L29
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., Brinchmann, J., Charlot, S., White, S. D. M., Seibert, M., Peng, E. W. et al. 2004, ApJ, 613, 898
- Vale Asari, N., Stasińska, G., Cid Fernandes, R., Gomes, J. M., Schlickmann, M., Mateus, A., & Schoenell, W. 2009, MNRAS, 396, L71
- Vogt, F. P. A., Pérez, E., Dopita, M. A., Verdes-Montenegro, L., & Borthakur, S. 2017, A&A, 601, A61
- York, D. G., Adelman, J., Anderson, Jr., J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouse, R. et al. 2000, AJ, 120, 1579