

DISSERTATION

submitted to the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of the
RUPERTO-CAROLA-UNIVERSITY of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by

Fabian SCHEUERMANN

born in Buchen (Odenwald)

Oral examination 18 November 2025

CHARACTERISING
EMISSION LINE NEBULAE
IN NEARBY GALAXIES

Referees

Dr. Kathryn KRECKEL
Priv.-Doz. Dr. Anna PASQUALI

*»The lunatic is in the hall,
The lunatics are in my hall.*

...

I'll see you on the dark side of the moon!«
—Brain Damage, Pink Floyd

Copyright © 2025, Fabian SCHEUERMANN

CHARACTERISING EMISSION LINE NEBULAE IN NEARBY GALAXIES

RUPRECHT-KARLS-UNIVERSITÄT

Zentrum für Astronomie

Astronomisches Rechen-Institut

Mönchhofstraße 12–14

69120 Heidelberg

*»Ohana heißt Familie. Familie heißt,
dass alle zusammenhalten und für einander da sind.«*
—Lilo & Stitch

Danke für eure Hilfe und Unterstützung,
vor allem während der letzten zwei Jahre.

In Gedenken an Erich Zimmermann (1940–2025)

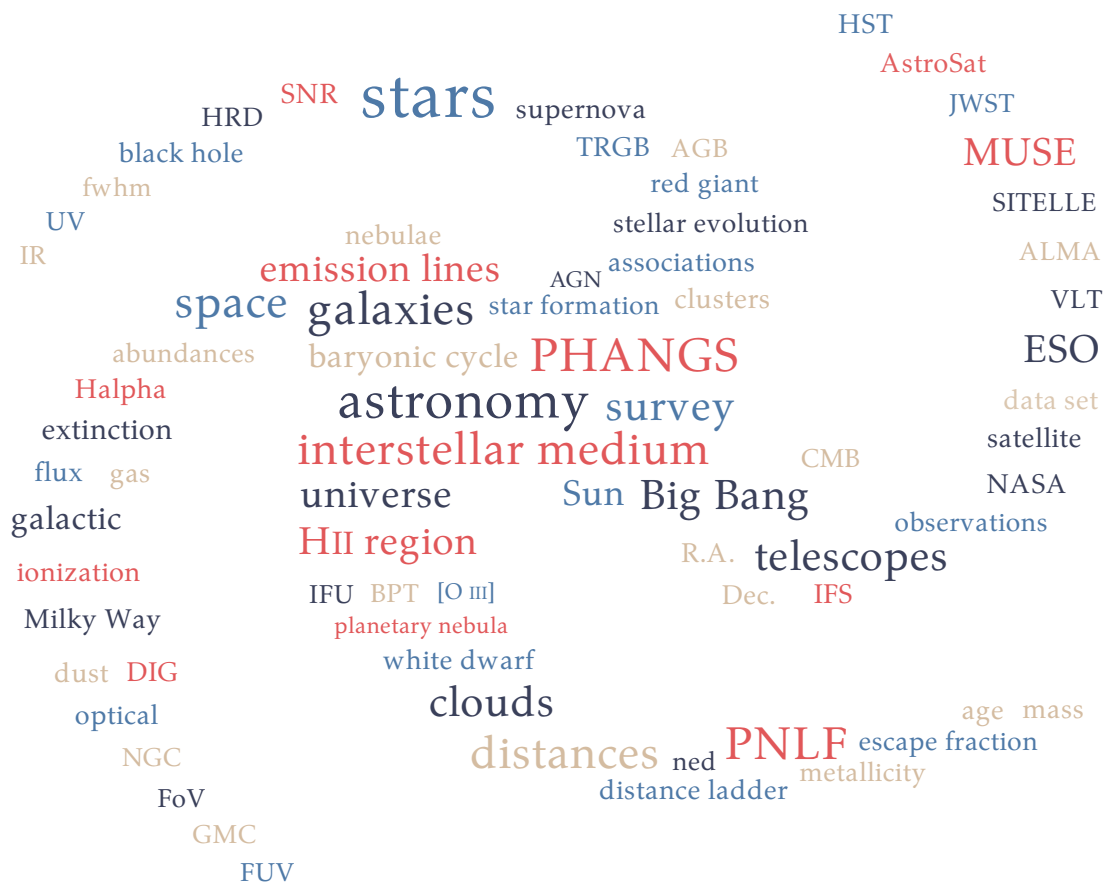
ABSTRACT

The ionized *interstellar medium* (ISM) is a central part of the galactic ecosystem, where the interaction between stars and gas becomes exceptionally vivid. Its main constituents – H II regions, *planetary nebulae* (PNe), and supernova remnants – trace different phases of stellar evolution and provide complementary perspectives on feedback and galactic structure. The emergence of *integral field spectroscopy* has opened new insights and enabled studies of the various components in unprecedented detail. We use a combination of observations – including MUSE spectroscopy and high-resolution *HST* imaging – across a diverse sample of nearby star-forming galaxies to study various aspects of emission line nebulae. In the first part of this thesis, we construct a dedicated PNe catalogue and measure the distance to each galaxy with the *planetary nebula luminosity function* (PNLF). The results highlight the potential of this method, but also its challenges, such as misclassified PNe or internal extinction. The second part links H II regions to the young stellar populations that power them. We identify potential age tracers – such as the equivalent width $\text{EW}(\text{H}\alpha)$, the flux ratio $\text{H}\alpha/\text{FUV}$, and the ionization parameter $\log q$ – that all show a correlation with stellar age. Building on this catalogue, we compare the predicted ionizing photon flux from the stars to the observed flux from the H II region to compute the escape fraction of ionizing photons. Individual regions yield comparatively high escape fractions, though with large variations and significant uncertainties. When combining all nebulae and stars on a galactic scale, the value remains high, but it becomes more robust and consistent across the galaxy sample. It makes it clear that stellar feedback is not locally bound, but influences the galaxies as a whole.

ZUSAMMENFASSUNG

Das ionisierte *interstellare Medium* ist ein zentraler Bestandteil des galaktischen Ökosystems, in dem die Wechselwirkung zwischen Sternen und Gas besonders deutlich zum Vorschein kommt. Seine Hauptbestandteile – H II-Regionen, *planetarische Nebel* (PNe) und Supernovaüberreste – spiegeln verschiedene Phasen der Sternentwicklung wider und eröffnen komplementäre Perspektiven auf Feedback und galaktische Strukturen. Die *Integralfeldspektroskopie* ermöglicht neue Einblicke und erlaubt es, die verschiedenen Komponenten in nie dagewesenen Details zu studieren. Wir verwenden eine Kombination von Beobachtungen – darunter den MUSE Spektrographen und hochaufgelöste Bilder von *HST* – um verschiedene Aspekte von Emissionsnebeln in nahe gelegenen Sternbildenden Galaxien zu untersuchen. Im ersten Teil dieser Arbeit konstruieren wir einen PNe Katalog und ermitteln die Entfernung zu den Galaxien mithilfe der *Leuchtkraftfunktion planetarischer Nebel*. Die Ergebnisse unterstreichen das Potenzial dieser Technik, aber auch ihre Herausforderungen, zu denen beispielsweise falsch klassifizierte PNe oder interne Extinktion gehören. Im zweiten Teil werden H II-Regionen mit den jungen Sternpopulationen, die sie antreiben, in Verbindung gebracht. Wir identifizieren eine Reihe möglicher Altersindikatoren – wie die Äquivalentbreite $EW(H\alpha)$, das Flussverhältnis $H\alpha/FUV$, und den Ionisationsparameter $\log q$ – die alle eine Korrelation mit dem Sternalter aufweisen. Auf Grundlage dieses Katalogs vergleichen wir den ionisierenden Photonenfluss der Sterne mit dem beobachteten Fluss der H II-Regionen, um die Escape Fraction ionisierender Photonen zu berechnen. Die einzelnen Regionen weisen im Durchschnitt vergleichsweise hohe Werte auf, allerdings gibt es große Schwankungen und erhebliche Unsicherheiten. Wenn man alle Nebel und Sterne auf galaktischer Ebene aufsummiert, bleibt der Wert hoch, wird jedoch robuster und konsistenter über die gesamte Stichprobe hinweg. Dies macht deutlich, dass stellares Feedback nicht lokal begrenzt ist, sondern die Galaxien als Ganzes beeinflusst.

484 references.



»Space: the final frontier«
—Captain Kirk

TABLE OF CONTENTS

FRONTMATTER

LIST OF FIGURES	ix
LIST OF TABLES	x
LIST OF ABBREVIATIONS	xi
LIST OF PUBLICATIONS	xiv
LIST OF CONSTANTS	xvi

PROLOGUE

1	MOTIVATION	3
1.1	A brief look into history	4
1.2	About this thesis	5
2	THE INTERSTELLAR MEDIUM	7
2.1	Characteristics of the ISM	7
2.1.1	Element abundances	7
2.1.2	The multi-phase medium	10
2.2	Emission line nebulae	11
2.2.1	H II regions	11
2.2.2	Supernova remnants	12
2.2.3	Planetary nebulae	13
2.3	Baryonic cycle	14
3	MODELS & METHODS	17
3.1	Stellar component	17
3.1.1	Single stars	17
3.1.2	Clusters and associations	21
3.2	Emission line diagnostics	23
3.2.1	Classification	24
3.2.2	Physical properties	25
4	OBSERVATIONS & DATA	29
4.1	Telescopes	29
4.1.1	MUSE	31
4.1.2	<i>HST</i>	32
4.1.3	<i>AstroSat</i>	32
4.1.4	Ancillary telescopes	32
4.2	PHANGS	33
4.2.1	Sample	33
4.2.2	Data sets	34
4.2.3	Data products	36

RESEARCH

5	PNLF DISTANCES FOR 19 GALAXIES OBSERVED BY PHANGS–MUSE	39
5.1	Introduction	40
5.2	Finding planetary nebulae	42
5.2.1	Data	42
5.2.2	Source detection and photometry	44
5.2.3	Comparison with existing studies	46
5.2.4	Emission line diagnostics	47
5.2.5	Overluminous sources	49
5.2.6	Final PN sample	50
5.3	The planetary nebula luminosity function	52
5.3.1	Completeness limit	52
5.3.2	Fitting the data	53
5.4	Results	54
5.5	Discussion	63
5.5.1	Comparison with literature distances	63
5.5.2	Zero point of the PNLF	65
5.5.3	Contamination with supernova remnants	67
5.5.4	Limitations of IFU observations to the PNLF	67
5.6	Conclusions	69
6	THE MUSE VIEW OF THE SCULPTOR GALAXY	71
6.1	The most detailed MUSE mosaic	72
6.2	Planetary nebulae luminosity function in NGC 253	73
6.3	Comparison with literature distances	74
7	DEFINING AN EVOLUTIONARY SEQUENCE	77
7.1	Introduction	78
7.2	Data	80
7.2.1	MUSE H II region catalogue	80
7.2.2	<i>HST</i> stellar association catalogue	82
7.2.3	<i>AstroSat</i>	83
7.3	Matching nebulae and stellar associations	84
7.3.1	Statistics of the matched catalogue	87
7.3.2	The physical nature of H II regions without associations	88
7.3.3	Validating the sample	90
7.4	H II region evolutionary sequence	92
7.4.1	Correlations with SED ages	92
7.4.2	Age trends in the nebular catalogue	94
7.4.3	Understanding the weak correlations with SED age	97
7.4.4	Recommendations on nebular age tracers	98
7.5	Conclusion	100
8	ESCAPE FRACTION OF IONIZING PHOTONS	101
8.1	Introduction	102

8.2	Data and models	103
8.2.1	MUSE H II regions	103
8.2.2	<i>HST</i> stellar associations and star clusters	105
8.2.3	Calculating the number of ionizing photons	107
8.2.4	Matched catalogues	109
8.3	Escape fractions	112
8.3.1	Ionizing photon fluxes for different samples	112
8.3.2	DiStribution of the escape fraction	112
8.3.3	What about negative escape fractions?	114
8.3.4	Photoionization budget of entire galaxies	115
8.3.5	Trends with nebulae properties	118
8.3.6	Comparison with existing studies	119
8.4	Conclusion	121
EPILOGUE		
9	BEYOND THE HORIZON	125
9.1	Improvement of stellar ages	125
9.2	Extended spectrum	125
10	RÉSUMÉ	127
10.1	Catalogues with emission line nebulae	127
10.2	DiStance measurements	128
10.3	Stars that ionize gas	129
10.4	Closing remarks	130
APPENDIX		
A	PLANETARY NEBULA LUMINOSITY FUNCTION	133
A.1	Comparison with literature distances	133
A.2	Extinction correction	141
A.3	Precision of the PNLF	142
B	DEFINING AN EVOLUTIONARY SEQUENCE	145
B.1	Model predictions	145
B.2	Background corrected EW	145
B.3	Density and temperature	146
B.4	Objects in unmatched H II regions	146
B.5	Binned age histograms	148
C	ESCAPE FRACTION OF IONIZING PHOTONS	149
C.1	Ionizing photon flux from different models	149
C.2	Comparison with Teh et. al (2023)	150
C.3	Completeness issues	152
BACKMATTER		
	REFERENCES	155
	ACKNOWLEDGEMENTS	167

LIST OF FIGURES

Figure 1.1	Pale Blue Dot & Hubble Ultra Deep Field.	4
Figure 2.1	Origin of the elements.	9
Figure 2.2	Images of emission line nebulae.	12
Figure 2.3	Evolution of a $2 M_{\odot}$ star in the HRD.	14
Figure 2.4	Baryonic life cycle.	15
Figure 3.1	Hertzsprung–Russell diagram of our neighborhood.	18
Figure 3.2	Solar spectrum vs black-body spectra.	19
Figure 3.3	Stellar evolutionary tracks and isochrones.	21
Figure 3.4	Spectra of emission line nebulae.	24
Figure 3.5	Density from PYNEB.	26
Figure 4.1	Very Large Telescope & Hubble Space Telescope.	31
Figure 4.2	PHANGS sample in the mass star-formation plane.	34
Figure 4.3	PHANGS atlas with the footprints of the different observations.	35
Figure 5.1	Location of the emission lines sources in NGC 0628.	44
Figure 5.2	[OIII] flux comparison with literature.	47
Figure 5.3	Emission line diagnostics.	48
Figure 5.4	RGB images of the overluminous objects.	49
Figure 5.5	Completeness limit for the PNLF.	53
Figure 5.6	Planetary nebula luminosity function.	56
Figure 5.7	Cumulative planetary nebula luminosity function.	57
Figure 5.8	Literature distances comparison for NGC 0628.	59
Figure 5.9	Literature distances comparison by method.	64
Figure 5.10	Zero point of the PNLF.	66
Figure 5.11	Luminosity-specific planetary nebula number.	69
Figure 6.1	MUSE mosaic of the Sculptor galaxy NGC 253.	73
Figure 6.2	Planetary nebula luminosity function for NGC 253.	74
Figure 6.3	Literature distances comparison for NGC 253.	75
Figure 7.1	Postage stamps for the possible overlaps.	85
Figure 7.2	Mass and age distribution of the catalogue.	88
Figure 7.3	H α luminosity function.	89
Figure 7.4	Comparison between nebula and association properties.	91
Figure 7.5	Stellar association ages vs age tracers.	93
Figure 7.6	Comparison between proposed age tracers.	95
Figure 7.7	Age trends in the nebular catalogue.	96
Figure 8.1	Comparison between different population synthesis models.	108

Figure 8.2	Postage stamps for the possible matches. 109
Figure 8.3	Comparison between $Q(H^0)$ and $Q_{H\alpha}$ for individual objects. 112
Figure 8.4	Histogram of the observed escape fractions for the different samples. 113
Figure 8.5	Contribution of external ionizing sources. 115
Figure 8.6	Comparison between $Q(H^0)$ and $Q_{H\alpha}$ for different galaxies. 118
Figure 8.7	Trends between different nebula properties and the escape fraction. 119
Figure 9.1	[OII] line map from SITELLE. 126
Figure A.1	Literature distances comparison for IC 5332. 133
Figure A.2	Literature distances comparison for NGC 1087. 134
Figure A.3	Literature distances comparison for NGC 1300. 134
Figure A.4	Literature distances comparison for NGC 1365. 135
Figure A.5	Literature distances comparison for NGC 1385. 135
Figure A.6	Literature distances comparison for NGC 1433. 135
Figure A.7	Literature distances comparison for NGC 1512. 136
Figure A.8	Literature distances comparison for NGC 1566. 136
Figure A.9	Literature distances comparison for NGC 1672. 136
Figure A.10	Literature distances comparison for NGC 2835. 137
Figure A.11	Literature distances comparison for NGC 3351. 137
Figure A.12	Literature distances comparison for NGC 3627. 138
Figure A.13	Literature distances comparison for NGC 4254. 138
Figure A.14	Literature distances comparison for NGC 4303. 139
Figure A.15	Literature distances comparison for NGC 4321. 139
Figure A.16	Literature distances comparison for NGC 4535. 140
Figure A.17	Literature distances comparison for NGC 5068. 140
Figure A.18	Literature distances comparison for NGC 7496. 140
Figure A.19	Effects of extinction on the PNLf. 141
Figure A.20	Effects of sample size and distance on the PNLf. 143
Figure B.1	Model predictions for $H\alpha$ /FUV and $EW(H\alpha)$ age evolution. 145
Figure B.2	Comparison between density prescriptions. 147
Figure B.3	Unmatched HII regions versus single stars. 147
Figure B.4	Histogram of age for different overlaps. 148
Figure B.5	Histogram of age for different $EW(H\alpha)$. 148
Figure C.1	$Q(H^0)$ as a function of age from different models and metallicities. 149
Figure C.2	Comparison between HII regions from MUSE and SITELLE. 151
Figure C.3	Comparison of the SED fit from CIGALE and SLUG. 152
Figure C.4	Normalized cumulative $Q(H^0)$ as a function of $Q(H^0)$. 153

LIST OF TABLES

Table 2.1	Phases of the ISM. 11
Table 3.1	Stellar classifications. 20
Table 5.1	Galaxy sample for the planetary nebula luminosity function. 43
Table 5.2	Planetary nebulae catalogue. 51
Table 5.3	Measured PNLF distances. 55
Table 7.1	Galaxy sample for the evolutionary sequence. 81
Table 7.2	Columns in matched catalogue. 86
Table 7.3	Spatial correlation between H II regions and stellar associations. 87
Table 8.1	Galaxy sample for the escape fraction. 104
Table 8.2	Sample size of the differently matched catalogues. 111
Table 8.3	Ionizing photon budget for entire galaxies. 117
Table B.1	Intervals for the equivalent width measurement. 146
Table C.1	Codes to compute the ionizing photon flux $Q(\text{H}^0)$. 150

LIST OF ABBREVIATIONS

A&A	Astronomy & Astrophysics
AGN	Active galactic nucleus
AO	Adaptive optics
ACS	Advanced Camera for Surveys
AJ	Astronomical Journal
ApJ	Astrophysical Journal
ARI	Astronomisches Rechen-Institut
AGB	Asymptotic giant branch
ALMA	Atacama Large Millimeter Array
BPT	Baldwin–Phillips–Terlevich
BPASS	Binary Population and Spectral Synthesis
BH	Black hole
CFHT	Canada–France–Hawaii Telescope
CSA	Canadian Space Agency
CNO	Carbon–nitrogen–oxygen cycle
CCD	Charge-coupled device
CIGALE	Code Investigating GALaxy Emission
CNM	Cold neutral medium
CMB	Cosmic microwave background
DAP	Data analysis pipeline
DR	Data release
DIG	Diffuse ionized gas
EW	Equivalent width
ESO	European Southern Observatory
ESA	European Space Agency
FoV	Field of view
FWHM	Full width half maximum
GALEX	<i>Galaxy Evolution Explorer</i>
GMC	Giant molecular cloud
GC	Globular cluster
GSGD	Gravitational Stability Gaseous Disk
HRD	Hertzsprung–Russell diagram
HIM	Hot ionized medium
<i>HST</i>	<i>Hubble Space Telescope</i>
ISRO	Indian Space Research Organisation
IR	Infrared
MIR	Mid-infrared

NIR	Near-infrared
IRAS	<i>Infrared Astronomical Satellite</i>
IMF	Initial mass function
IFS	Integral field spectroscopy
IFU	Integral field unit
IFTS	Imaging Fourier transform spectrometer
ISM	Interstellar medium
JWST	<i>James Webb Space Telescope</i>
JPL	Jet Propulsion Laboratory
LMC	Large Magellanic Cloud
LEGUS	Legacy ExtraGalactic Ultraviolet Survey
LINER	Low ionization nuclear emission region
MaNGA	Mapping Nearby Galaxies at Apache Point Observatory
MLE	Maximum likelihood estimation
MNRAS	Monthly Notices of the Royal Astronomical Society
MUSE	Multi Unit Spectroscopic Explorer
MAD	The MUSE Atlas of Discs
NED	NASA/IPAC Extragalactic Database
NASA	National Aeronautics and Space Administration
NGC	New General Catalogue
NAM	Numerical action method
ORCS	Outils de Réduction de Cubes Spectraux
PHANGS	Physics at High Angular resolution in Nearby Galaxies
PN	Planetary nebula
PNLF	Planetary nebula luminosity function
PSF	Point spread function
PI	Principal investigator
PDF	Probability density function
PP chain	Proton–proton chain
PASP	Publications of the Astronomical Society of the Pacific
RGB	Red giant branch
r-process	Rapid neutron-capture process
s-process	Slow neutron-capture process
SWG	Science working group
S/N	Signal-to-noise
SDSS	Sloan Digital Sky Survey
SMC	Small Magellanic Cloud
STScI	Space Telescope Science Institute
SED	Spectral energy distribution
SITELLE	Spectromètre Imageur à Transformée de Fourier pour l'Etude en Long et en Large de raies d'Emission

SIGNALS	Star formation, ionized gas and nebular abundances legacy survey
SFR	Star formation rate
SLUG	Stochastically Light Up Galaxies
SN	Supernova
SNR	Supernova remnant
SBF	Surface brightness fluctuations
TIMER	Time Inference with MUSE in Extragalactic Rings
TRGB	Tip of the red giant branch
FUV	Far ultraviolet
NUV	Near ultraviolet
UV	Ultraviolet
VLT	Very Large Telescope
WIM	Warm ionized medium
WNM	Warm neutral medium
WD	White dwarf
WFI	Wide Field Imager
UVIT	Ultra-Violet Imaging Telescope
WFC3	Wide Field Camera 3
WR star	Wolf-Rayet star
Z0MGS	$z = 0$ Multi-wavelength Galaxy Synthesis
ZAHB	Zero-age horizontal branch

LIST OF PUBLICATIONS

Articles 7, 16, 20, and 22 are presented in this thesis.

1. Kreckel, K. et al. (inc. Scheuermann, F.), 2020, *MNRAS*, 499, 193, *Measuring the mixing scale of the ISM within nearby spiral galaxies.*
2. Anand, G. S. et al. (inc. Scheuermann, F.), 2021, *MNRAS*, 501, 3621, *Distances to PHANGS galaxies: New tip of the red giant branch measurements and adopted distances.*
3. Leroy, A. K. et al. (inc. Scheuermann, F.), 2021a, *ApJS*, 255, 19, *PHANGS–ALMA Data Processing and Pipeline.*
4. Leroy, A. K. et al. (inc. Scheuermann, F.), 2021b, *ApJS*, 257, 43, *PHANGS–ALMA: Arcsecond CO(2–1) Imaging of Nearby Star-forming Galaxies.*
5. Lee, J. C. et al. (inc. Scheuermann, F.), 2022, *ApJS*, 258, 10, *The PHANGS–HST Survey: Physics at High Angular Resolution in Nearby Galaxies with the Hubble Space Telescope.*
6. Emsellem, E. et al. (inc. Scheuermann, F.), 2022, *A&A*, 659, A191, *The PHANGS–MUSE survey. Probing the chemo-dynamical evolution of disc galaxies.*
7. Scheuermann, F. et al., 2022, *MNRAS*, 511, 6087, *Planetary nebula luminosity function distances for 19 galaxies observed by PHANGS–MUSE.*
8. Barnes, A. T. et al. (inc. Scheuermann, F.), 2022, *A&A*, 662, L6, *Linking stellar populations to H II regions across nearby galaxies. I. Constraining pre-supernova feedback from young clusters in NGC 1672.*
9. Kreckel, K. et al. (inc. Scheuermann, F.), 2022, *A&A*, 667, A16, *A physically motivated »charge-exchange method« for measuring electron temperatures within H II regions.*
10. Egorov, O. V. et al. (inc. Scheuermann, F.), 2023, *ApJL*, 944, L16, *PHANGS–JWST First Results: Destruction of the PAH molecules in H II regions probed by JWST and MUSE.*
11. Lee, J. C. et al. (inc. Scheuermann, F.), 2023, *ApJL*, 944, L17, *The PHANGS–JWST Treasury Survey: Star Formation, Feedback, and Dust Physics at High Angular resolution in Nearby Galaxies.*
12. Watkins, E. J. et al. (inc. Scheuermann, F.), 2023, *ApJL*, 944, L24, *PHANGS–JWST First Results: A statistical view on bubble evolution in NGC 628.*

13. Whitmore, B. C. et al. (inc. Scheuermann, F.), 2023, [MNRAS](#), **520**, 63,
Improving Cluster Age Estimates in PHANGS–HST Galaxies.
14. Groves, B. et al. (inc. Scheuermann, F.), 2023, [MNRAS](#), **520**, 4902,
The PHANGS–MUSE Nebular Catalogue.
15. Congiu, E. et al. (inc. Scheuermann, F.), 2023, [A&A](#), **672**, A148,
PHANGS–MUSE: detection and Bayesian classification of $\sim 40,000$ ionised nebulae in nearby spiral galaxies.
16. Scheuermann, F. et al., 2023, [MNRAS](#), **522**, 2369,
Stellar associations powering H II regions – I. Defining an evolutionary sequence.
17. Watkins, E. J. et al. (inc. Scheuermann, F.), 2023, [A&A](#), **676**, A67,
Quantifying the energetics of molecular superbubbles in PHANGS galaxies.
18. Hannon, S. et al. (inc. Scheuermann, F.), 2023, [MNRAS](#), **526**, 2991,
Star Cluster Classification using Deep Transfer Learning with PHANGS–HST.
19. Whitmore, B. C. et al. (inc. Scheuermann, F.), 2025, [ApJ](#), **982**, 50,
Empirical SED Templates for Star Clusters Observed with HST and JWST: No Strong PAH or IR Dust Emission after 5 Myr.
20. Congiu, E. et al. (inc. Scheuermann, F.), 2025, [A&A](#), **700**, A125,
The MUSE view of the Sculptor galaxy: survey overview and the planetary nebulae luminosity function.
21. Barnes, A. T. et al. (inc. Scheuermann, F.), 2025, [A&A](#), submitted,
The PHANGS-MUSE/HST-H α Nebulae Catalogue: Parsec-Scale Resolved Structure, Physical Conditions, and Stellar Associations across Nearby Galaxies.
22. Scheuermann, F. et al., 2025, [A&A](#), submitted,
Stellar associations powering H II regions – II. Escape fraction of ionizing photons.
23. Razza, A. et al. (inc. Scheuermann, F.), in preparation,
PHANGS-H α : Ground-based Narrow-band Imaging Survey of Nearby Star-Forming Galaxies.
24. Habjan, E. et al. (inc. Scheuermann, F.), in preparation,
H II region Metallicities in Nearby Galaxies with MUSE and SITELLE.

LIST OF CONSTANTS

Speed of light	c	$2.998 \cdot 10^{10}$	cm s^{-1}
Gravitational constant	G	$6.673 \cdot 10^{-8}$	$\text{cm}^3 \text{g}^{-1} \text{s}^{-2}$
Planck constant	h	$6.626 \cdot 10^{-27}$	erg s
Boltzmann constant	k_B	$1.381 \cdot 10^{-16}$	erg K^{-1}
Stefan–Boltzmann constant	σ_{SB}	$5.670 \cdot 10^{-5}$	$\text{erg cm}^{-2} \text{s}^{-1} \text{K}^{-4}$
Wien’s displacement constant	b	$2.898 \cdot 10^{-1}$	cm K
Rydberg constant (hydrogen)	R_{H}	$1.097 \cdot 10^{-5}$	cm^{-1}
Elementary charge	e	$1.602 \cdot 10^{-19}$	C
Electron mass	m_e	$9.109 \cdot 10^{-28}$	g
Proton mass	m_p	$1.673 \cdot 10^{-24}$	g
Astronomical unit	AU	$1.496 \cdot 10^{13}$	cm
Light-year	ly	$9.461 \cdot 10^{17}$	cm
Parsec	pc	$3.086 \cdot 10^{18}$	cm
		206 265	AU
		3.26	ly
Solar mass	M_{\odot}	$1.988 \cdot 10^{33}$	g
Solar luminosity	L_{\odot}	$3.846 \cdot 10^{33}$	erg s^{-1}

Note. Taken from Mohr & Taylor (2000) and Unsöld & Baschek (2002).

PROLOGUE

MOTIVATION

FEW THINGS fascinate us as much as the endless vastness of space. In 1977, NASA launched *Voyager 1* and *Voyager 2* to explore the outer Solar System (Kohlhase & Penzo 1977). It took them over a decade to pass the giant planets and around 35 yr to enter interstellar space (Krimigis et al. 2013; Stone et al. 2013). In 1990, long after its main mission was completed, the spacecraft looked back one last time and took the famous photograph that Sagan (1994) called »Pale Blue Dot«. Shown in Figure 1.1a, it was taken from a distance of 40.5 AU, where the Earth appears as a mere pixel. While the exact duration is debated, it is estimated that it will take another 30 000 yr for the mission to leave the Oort cloud. Yet this is just our immediate neighbourhood, while the universe as a whole extends to far larger, almost inconceivable scales. In the words of a great poet

Jupiter and Saturn are gas giants, whereas Uranus and Neptune are ice giants.

*»But outer Space,
At least this far,
For all the fuss
Of the populace
Stays more popular
Than populous.«*

—Robert FROST (1874–1963)

Visiting any object of interest beyond our own Solar System is therefore not feasible, and research relies heavily on observations. As astronomers, we have the pleasure of studying the diverse facets of the universe and figuring out our position in it by »reaching for the stars«. Fortunately for us, there are many of them to study. It is estimated that the Sun is only one of hundreds of billions of stars in our Milky Way, which, in turn, is just one of many galaxies. The Hubble Ultra-Deep Field in Figure 1.1b shows only a tiny snippet of the sky, but thanks to its staggering exposure time of roughly one million seconds, around 10 000 galaxies are visible (Beckwith et al. 2006). As far as the entire universe is concerned, this number could reach two trillion (Conselice et al. 2016).

Although it has been researched for centuries, the recent improvements in instrumentation have proven to be a giant leap that enables us to look deeper than ever before and understand how the whole structure works. Before we look into the night sky, let us first look back in time and take a short trip through the history of astronomy. Along the way, we will highlight a selection of discoveries that will prove to be important for our work. This provides context for understanding the motivation and purpose of this work and highlights its relation to other subject areas.

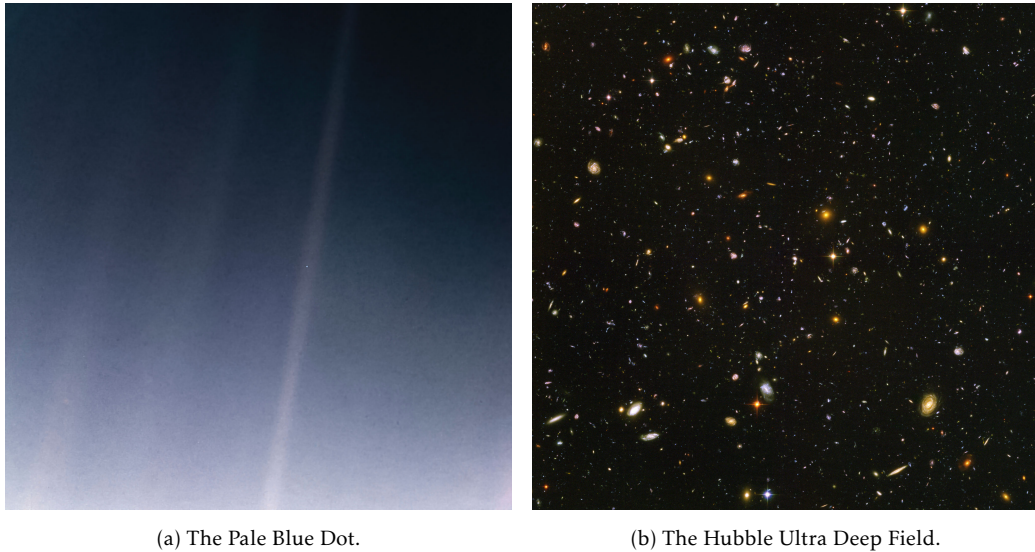


Figure 1.1: *Left*: in the last image taken by Voyager 1, our Earth occupies only a single pixel. Credit: NASA and JPL-Caltech. *Right*: thanks to its extremely long exposure time, this image from *HST* reveals a large number of galaxies in a seemingly empty patch of sky. Credit: NASA, ESA, Beckwith (STScI), and the HUDF Team.

1.1 A BRIEF LOOK INTO HISTORY

The night sky has captivated mankind since ancient times and often played an influential role in culture, impacting a multitude of subjects. This includes *constellations* that were developed all around the world, from Babylon to Egypt, and China. It also covers *celestial navigation*, as practised by the Polynesians, for example, who were guided by the stars when sailing across the wide stretches of the Pacific ocean. To keep track of time, people followed the *solar* or *lunar calendars*. And some people were intrigued to explain things on a more scientific level. Polymaths such as Archimedes, Eratosthenes, Pythagoras, or Hipparchus arose and came up with theories, some of which persist to this day. This includes the magnitudes to describe the brightness, predicting solar eclipses, or the *heliocentric model*. Overall, however, their models were based on observations with their eyes and hence also limited by them. Only the invention of the telescope at the beginning of the 17th century propelled the evolution to more advanced scientific theories. During the *Scientific Revolution*, the work of old masters such as Nicolaus COPERNICUS, Galileo GALILEI, Johannes KEPLER, or Sir Isaac NEWTON provided new insights.

This idea, proposed by Aristarchus of Samos, found little support at the time and the geocentric model dominated for millennia.

When it comes to measuring distance in astronomy, we rely on different techniques that are built upon each other, the so-called *cosmic distance ladder* (Webb 1999). This becomes necessary because all methods are only applicable to a limited distance range. For our galactic neighbourhood, we can use the rotation around the Sun to determine the distance to an object from the parallax angle. That alone proved to be quite challenging due to the tiny angles involved. Bessel (1838) was the first one who managed to apply this

technique to the nearby star 61 Cygni and measured a distance of 3.16 pc. At the time, the result represented an unfathomable distance and raised questions about the scale and structure of our universe. In order to climb the next step, we need another technique and that also took quite a while again. For a *standard candle*, we know the intrinsic luminosity and by comparing it to the observed luminosity, we can deduce the distance. The breakthrough came when Leavitt & Pickering (1912) discovered *Cepheids* as variable stars for which she linked the intrinsic luminosity to the duration of the pulsation. That enlarged our universe once again. This culminated in the *Great Debate* between Shapley & Curtis (1921) which marked a turning point in our understanding. The conclusion was that our own Galaxy is just one of many galaxies and the cosmic distances are even bigger. Further proof for this came when Hubble (1929a,b) measured the distance to the Andromeda Galaxy and provided evidence for the expansion of the universe. Our perspective shifted from a static universe that is limited to the Milky Way, to our current theory that is built around an expanding universe that formed in the *Big Bang* (Lemaître 1931).

Shapley argued that the Milky Way represents the entire universe, while Curtis favoured additional galaxies outside our own.

During the same period, our understanding of the astronomical objects also evolved. Some objects appeared extended and it became clear that not all of them are stars or planets, but their exact nature remained a mystery. Catalogues compiled by Messier (1781) or Herschel (1786) contained a variety of such objects which the latter called »*nebulae and clusters of stars*«. After the previously mentioned Great Debate it became clear that some of those objects were galaxies but others required additional investigations. Of particular relevance there are the advancements in spectral analysis. Prisms have been around for a long time and it was known that they disperse the light into the colours of the rainbow. Building upon that, Wollaston (1802) and Fraunhofer (1817) independently noticed a set of spectral absorption lines in light of the Sun, which were named after the latter. It was Kirchhoff & Bunsen (1860) who first delivered an explanation that linked the lines to chemical elements, enabling us to study the composition of astronomical objects. Huggins & Miller (1864a) applied this technique and noted the difference in spectra between nebulae and galaxies. This led to a better differentiation between different types of celestial objects. This laid part of the foundation for the theory on how stars form that evolved during the 20th century.

The Messier and the NGC are still in use today and will make further appearances in this thesis.

As far as the size and structure of the universe is concerned, our understanding has changed dramatically over the last 200 yr. Astronomy today encompasses a wide range of scales and topics, from the size of planets (or even smaller, such as processes in dust) to star clusters and galaxies, and ultimately to the entire universe. To get the big picture, one has to consider multiple branches that often strongly affect each other.

1.2 ABOUT THIS THESIS

This brings us to the purpose of this thesis. Measurements of astronomical distances have improved tremendously over the years, and thanks to a wide variety of techniques, we are able to map large portions of the universe (Jacoby et al. 1992). Our neighbourhood in the Milky Way has been cartographed by satellites such as *Hipparcos* and *Gaia* (Gaia

Collaboration et al. 2018), and cosmological scales are now determined with unprecedented precision (Riess et al. 2022). However, many galaxies remain poorly constrained, and uncertainties in their distances propagate to derived properties such as luminosity and mass, complicating further analyses. The planetary nebula luminosity function has experienced a revival in recent years, and modern telescopes now allow its application to many galaxies without dedicated observations. We apply this technique to a large sample of nearby galaxies, which also enables us to explore potential weaknesses and assess the reliability of the derived values. In the process, we construct a large catalogue of planetary nebula, which are themselves valuable objects for further study.

In the second part, the goal is to link H II regions to newly formed stars that power them. Despite their close physical relationship, these objects are usually studied independently of each other. By combining them, we can exchange information between the two populations and deepen our overall understanding. H II regions provide measurements such as metallicity, ionization parameter, density, and dust attenuation, while stellar populations yield age, mass, and chemical composition. Of particular importance is the timescale of their coexistence, as the ionized gas will eventually vanish – either because the stars cease to exist, or because their feedback disrupts and disperses the surrounding cloud.

Carrying out such an analysis requires spatially resolved spectroscopy to measure nebular properties across entire galaxies, and high-resolution imaging to resolve the stellar populations that power them (Schinnerer & Leroy 2024). Combining these data sets makes it possible to link ionized gas directly to its stellar sources and to resolve ambiguities that would remain if each were studied in isolation. Equally important is the large sample size provided by PHANGS, which enables us to investigate thousands of regions across dozens of galaxies. Together, these observations offer a uniquely robust foundation for disentangling the interplay between stars and gas. Within this framework, this research addresses the following goals:

- (i) Distinguish the different components of the ionized gas.
- (ii) Derive fundamental properties, such as distance, metallicity, and temperature.
- (iii) Link ionized regions to their sources of ionization to improve our understanding of their evolution.

This thesis is structured as follows: In Chapter 2 we go over the galactic components that we aim to study, i.e. certain parts of ionized gas around stars. Following this, in Chapter 3 we introduce the underlying models and introduce different methods to determine the physical properties that we rely on in our analysis. The underlying data and the PHANGS survey are presented in Chapter 4. Based on this data set, we measure the distance to 19 galaxies in Chapter 5 and the same method is then also applied to the Sculptor Galaxy in Chapter 6. Next we investigate the evolution of H II regions in Chapter 7 and their escape fraction in Chapter 8. Finally, we look for future improvements in Chapter 9 and conclude all important findings in Chapter 10.

THE INTERSTELLAR MEDIUM

AT FIRST GLANCE, the space between stellar systems appears to be empty. That may be the case compared to our conditions on earth, but even a minuscule density adds up to a substantial mass over a huge volume – in which we undoubtedly find ourselves, as we saw earlier. In the case of the Milky Way, it is estimated that around 10 per cent of the baryonic matter is contained in the so-called *interstellar medium* (ISM) and to put it bluntly

»The interstellar medium is anything not in stars.«
—Donald OSTERBROCK (1924–2007)

That may be a bit too simplistic and especially cosmologists may roll their eyes, as less than 5 per cent of the energy content in our universe is present in baryonic form (Planck Collaboration et al. 2020). But when it comes to the galactic ecosystem, this material plays an important role and sets the stage for the formation and evolution of stars. The distribution of the medium is inhomogeneous, as some of it is packed into very dense clumps, while other areas are very diluted. Depending on the ambient temperature and density, it exists in all possible states, from molecular to neutral to ionized.

*The rest is assumed to be
26 per cent dark matter and
69 per cent dark energy.*

The vast majority of the mass is found in the form of gas (99 per cent) while the remaining material is bound in dust grains, ranging in size from 50 to 250 nm (Mathis, Rumpl & Nordsieck 1977; Draine & Lee 1984). That amount of dust might appear negligible, but it contains a significant share of the metals (~ 50 per cent) and plays an important role in the observability and other processes we are dealing with.

In all the following, we refer specifically to spiral and irregular galaxies, unless specifically stated. Much of it does not necessarily apply to elliptical galaxies, which in turn are even more sparsely filled. We start by going over some characteristics of the ISM. This is mostly based on the literature by Ferrière (2001), Cox (2005), Tielens (2005), Draine (2011), Klessen & Glover (2016), and Saintonge (2025), all of which offer an excellent insight to this topic.

2.1 CHARACTERISTICS OF THE ISM

2.1.1 Element abundances

The composition of the baryonic matter is dominated by hydrogen and helium, but despite their low occurrence, heavier elements have a major impact on many processes. The cooling of the ISM is strongly affected by them (Glover & Clark 2014), which in turn is essential for the creation of stars (Krumholz 2012) and impacts their evolution (Hirschi

et al. 2008). In the form of dust, they are an opaque medium that is the main cause of extinction (Gordon et al. 2003; Li et al. 2021). And last but not least, they are an important building block for planets, especially for terrestrial planets (Johnson & Li 2012).

Astronomers call everything heavier than helium a metal.

There are a number of ways to quantify the metallicity. A common way is to simply specify the mass fraction of the elements, where it is common that X represents hydrogen, Y helium, and Z everything heavier. The composition is similar to the Sun and mostly consists of hydrogen $X \approx 70$ percent, followed by helium $Y \approx 28$ percent, and a small amount of metals $Z < 2$ percent. If we ignore the mass and only consider the number of atoms, it becomes even more lopsided, with 90.8 percent in hydrogen and merely 0.12 percent in metals (Ferrière 2001). However, this is by no means the same everywhere and undergoes an important change over time. But before we get into the nitty-gritty details we need to tweak up the notation a little.

When dealing with the individual elements, the numbers become very small and extend over many orders of magnitude. It is therefore handy to use a logarithmic scale and a ratio of the number densities. For stars, it is common to compare this value to the one from the Sun

$$[\text{Fe}/\text{H}] = \log_{10} \left(\frac{N_{\text{Fe}}}{N_{\text{H}}} \right)_{\star} - \log_{10} \left(\frac{N_{\text{Fe}}}{N_{\text{H}}} \right)_{\odot}. \quad (2.1)$$

So in this notation, a value of $[\text{Fe}/\text{H}] = -1$ corresponds to one tenth of the solar iron abundance. For the ISM, an absolute scale is used instead, the so-called *gas phase abundance*

$$12 + \log(\text{O}/\text{H}) = 12 + \log_{10} \left(\frac{N_{\text{O}}}{N_{\text{H}}} \right). \quad (2.2)$$

Oxygen is often used as a reference, as it is the most common element after hydrogen and helium. The rarest elements appear roughly 10^{-12} times as much as hydrogen (Grevesse & Sauval 1998; Asplund et al. 2009), so the plus 12 ensures that even they are positive. But why is there such a big difference in the frequency of the individual elements? To understand the *origin of elements*, a complex interplay of various processes is required (Gibson et al. 2003; José & Iliadis 2011; Kobayashi, Karakas & Lugaro 2020).

Of course, stars play a central role in this, but first we must go back to the beginning – the very beginning. For the two lightest elements, it is fairly certain that they originate from the primordial nucleosynthesis (Gamow 1948a,b). Shortly after the Big Bang, the universe went through a phase of extreme temperatures and densities that were sufficient for nuclear fusion to happen (Wagoner, Fowler & Hoyle 1967; Wagoner 1973). The time frame with suitable conditions depends on the rate of expansion and is estimated to yield $Y = 22$ to 25 percent (Schramm & Turner 1998). Apart from lithium and small traces of beryllium, it was unable to merge heavier elements, so they must have formed somewhere else.

That brings us back to the stars. The ones we observe today were separated into two classes by Baade (1944): *Population I* are metal-rich and found in the spiral arms of galaxies. *Population II* on the other hand are old and metal-poor and reside in the bulge

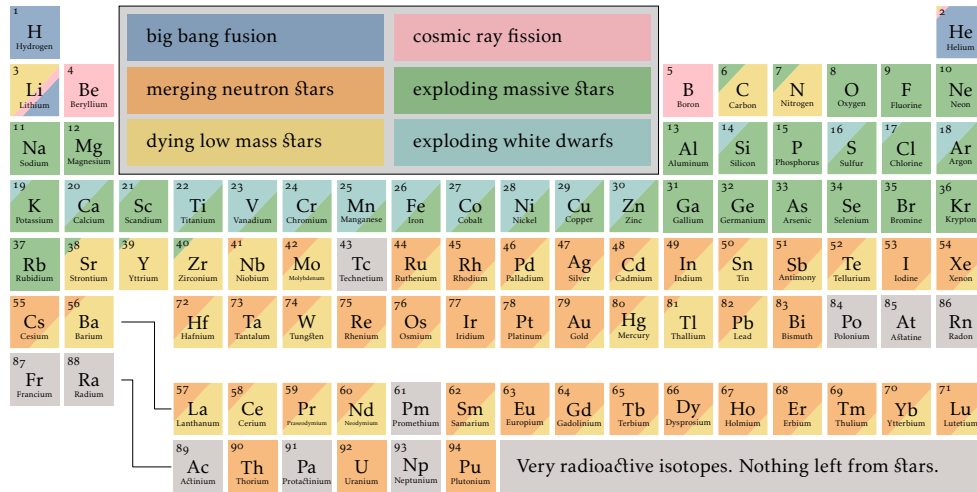


Figure 2.1: The origin of elements in our Solar System. Credit: based on the figure by Johnson, Fields & Thompson (2020).

of the galaxies or *globular clusters* (GCs). The latter formed earlier during a metal poor era. However, even they contain more heavy elements than the primordial nucleosynthesis can explain, so there had to be something beforehand. *Population III* stars are a proposed solution that only existed during a short period in the early universe (Heger & Woosley 2002; Bromm & Larson 2004; Tominaga, Umeda & Nomoto 2007). Their properties are entirely based on simulations and so far, we have not been able to observe them. Observers have been hunting for very metal poor stars for decades, but the lowest we found is around $[\text{Fe}/\text{H}] < -4$ (Wallerstein 1962; Starkenburg et al. 2017).

Although stars get their energy via nuclear fusion (see also Section 3.1.1), most of the heavier elements are created in other ways (Johnson, Fields & Thompson 2020). Fusion inside stars can only build nuclei up to iron, while heavier ones require neutron capture (Burbidge et al. 1957). In this process, an atomic nucleus collides with a free neutron and merges to a heavier isotope, which often undergoes beta decay to transform into a stable element. In the *slow neutron-capture process* (s-process), which occurs mainly in *red giant branch* (RGB) and *asymptotic giant branch* (AGB) stars, unstable nuclei have time to decay between captures, gradually producing heavier elements. The *rapid neutron-capture process* (r-process), in contrast, proceeds in environments with extreme neutron densities, most convincingly in neutron star mergers (Abbott et al. 2017), and produces many of the heaviest nuclei. *Supernovae* (SNe) also play a central role in chemical enrichment: core-collapse supernovae (Type II) mark the deaths of massive stars ($> 8 M_{\odot}$) and eject large amounts of intermediate-mass elements such as oxygen and sodium. Thermonuclear supernovae (Type Ia) occur when a *white dwarf* (WD) in a binary system undergoes runaway nuclear burning, and they are the dominant source of the iron group. Together, these stellar and explosive processes seeded the universe with the diversity of elements observed today. Figure 2.1 sketches a picture for the origin of the elements.

There is great hope that we will finally be able to observe them with JWST.

2.1.2 The multi-phase medium

Although the physical conditions in the ISM vary greatly, much of the gas can be categorized into one of a few distinct phases. These are mostly stable and only change over very long timescales (Tielens 2005). Starting with the neutral gas, Field, Goldsmith & Habing (1969) introduced a two phase model by assuming an equilibrium between the heating from cosmic rays and radiative cooling. There are two thermally stable solutions, the *cold neutral medium* (CNM) with temperatures less than 100 K and the *warm neutral medium* (WNM) with temperatures up to 10^4 K. Gas with an intermediate temperature either cools down or heats up to either stable state, depending on its density. The former has a comparably high density up to 50 cm^{-3} but covers a relatively small volume, while the latter fills a large part of the disk, but at a much lower density of less than 1 cm^{-3} . Together they dominate the overall mass of the ISM and occupy a large chunk of the galactic plane. Given that the gas mostly consists of neutral atomic hydrogen (HI), it is not possible to observe it at optical wavelengths. Ewen & Purcell (1951) were the first who managed to observe it via the 21 cm hyperfine transition line in the radio wavelengths. In the case of the CNM it appears as an absorption line against the background, while for the WNM it is an emission line. This has proven itself as a valuable tool for studying the kinematics or mass within our own Galaxy (Kalberla et al. 2005) and nearby galaxies (Walter et al. 2008).

The air on earth at sea level has a density of 10^{19} cm^{-3} .

Spitzer (1956) proposed a galactic corona in which the disk is embedded in order to keep everything in a pressure equilibrium. By considering the very hot (10^6 K), shock heated material from SNe, McKee & Ostriker (1977) added a third phase, called the *hot ionized medium* (HIM). The particles in this phase are highly ionized and fill up a huge volume at a miniscule density. While this gas will eventually cool down and become WNM, the time it spends around 10^6 K is much longer than it takes to cool down. Therefore, it makes sense to treat this as an additional phase. Out of all the phases, this one has by far the lowest density ($< 10^{-2} \text{ cm}^{-3}$), but at the same time takes up the most space, as it fills large cavities throughout the disc. This phase primarily emits thermal bremsstrahlung (free-free radiation) and highly ionized metal lines, observed mainly in the soft X-ray and ultraviolet (UV) wavelength ranges.

Our Solar System is currently located in the Local Bubble which is part of the HIM (Farhang et al. 2019).

The gas can also be ionized by massive stars and Hoyle & Ellis (1963) proposed another phase which was later called the *warm ionized medium* (WIM) and first observed by Hewish et al. (1968). The temperature is similar to the WNM and this phase accounts for 90 per cent of the ionized gas in terms of mass (Haffner et al. 2009).

What all the phases above have in common is that they stay in a stable state for a long time. But there is more: we are particularly interested in local environments which are subject to constant change. In the coolest and densest parts of the ISM, the hydrogen combines to H_2 and forms *giant molecular clouds* (GMCs). Due to their high density, those regions can become gravitationally unstable and collapse to form stars. The fact that H_2 is a symmetric molecule without a dipole moment makes it very hard to observe at such low temperatures. Instead, emission from the rotational transition of carbon monoxide, $\text{CO}(2-1)$ at 1.3 mm, is used as a tracer (Wilson, Jefferts & Penzias 1970).

Table 2.1: Properties of the different phases of the ISM. The numbers presented here are compiled from various articles (Ferrière 2001; Tielens 2005; Girichidis et al. 2020).

Phase	Temperature K	Density cm^{-3}	Fractional volume	Fractional mass
Molecular clouds	10–20	$10^2 - 10^6$	<1 %	13–20 %
CNM	50–100	20–50	1–5 %	~30 %
WNM	6000–10 000	0.2–0.5	10–20 %	35–38 %
WIM	8000	0.2–0.5	20–50 %	12–14 %
HIM	$10^6 - 10^7$	$10^{-4} - 10^{-2}$	30–70 %	3–4 %
H II regions	8000	$10^2 - 10^4$	<1 %	–

We briefly discussed ionized gas above, where it is distributed over large scales and maintained in this state by a constant flux of ionizing photons. This diffuse component must be distinguished from localized clouds that are ionized by a nearby source for only a limited time. Although these clouds make up the smallest fraction of the ISM, their wealth of emission lines enables detailed study, and they will be discussed in much greater detail in the next section.

An overview of the characteristics of the different phases is shown in Table 2.1. The gas in the ISM spans a wide range of temperature and densities, but occur in certain conditions. It is unclear how distinct the individual phases actually are and this is an active topic in research.

2.2 EMISSION LINE NEBULAE

It is time to turn our attention to the eponymous subject of this work. *Emission line nebulae* are clouds of gas that are predominantly ionized. For them to emerge, there must be a source of high-energy radiation. The emitted photons must have an energy of at least 13.6 eV, in order to ionize hydrogen, the most abundant element. Only a handful of objects are capable of doing this, and those that are of interest to us all represent a different stage in the life of a star. At first glance, they all look very similar, as shown in Figure 2.2, but there are very different processes underlying them.

2.2.1 H II regions

Let us begin, fittingly, with the birth of the stars. When the collapse of a GMC triggers the formation of a new generation, only a small fraction of the cloud is actually turned into stars (Krumholz & Tan 2007; Freundlich 2024). Most of the material remains in gaseous form and is now exposed to the strong radiation of massive OB stars. They are able to ionize their immediate neighbourhood, which we call an H II region. In theory, the radius of the ionized sphere depends only on the density of the gas and the production rate

In the nomenclature, H I refers to neutral and H II denotes singly ionized.

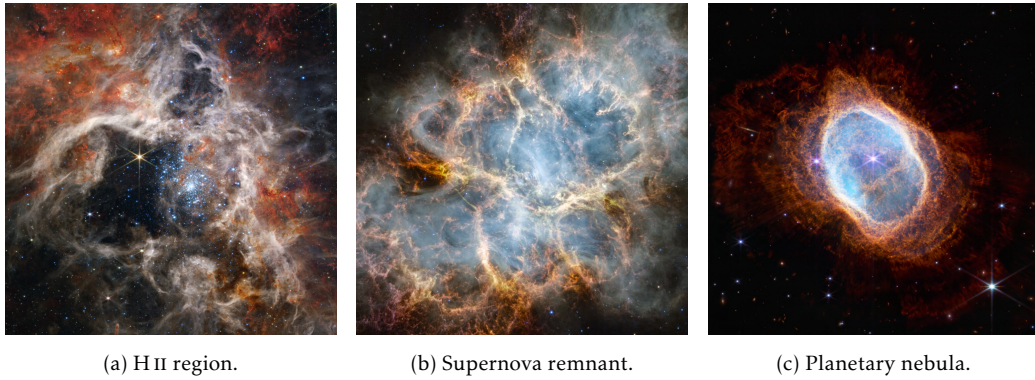


Figure 2.2: Images of the three most common types of emission line nebulae. They show the H II region NGC 2070 (Tarantula Nebula), the SNR NGC 1952 (Crab Nebula) and the PN NGC 3132 (Southern Ring Nebula). All three were observed with *JWST*. Credit: NASA, ESA, CSA, STScI.

of ionizing photons from the star (Strömgren 1939). In reality, the stellar feedback will expand the H II region and push it into the surrounding medium (Spitzer 1978; Raga, Cantó & Rodríguez 2012). In addition, inhomogeneities in the ISM generally lead to non-spherical expansion, so they span a wide range in sizes, from 1 to 100 pc. A famous example for an H II region is displayed in Figure 2.2a.

The conditions in this cloud are perfect for emitting bright emission lines in the optical spectra. For hydrogen, this gives rise to series of recombination lines, whereby the *Balmer series* is best suited for our purposes of studying stars and the ionized gas surrounding them. Other series, such as the *Lyman series* in the UV, are difficult to observe on Earth and are only relevant for very massive, and hence hot stars, while the *Paschen series* in the *infrared* (IR) suits better with colder objects, such as planets. In addition, there are many other, often forbidden emission lines from elements such as oxygen, nitrogen, or sulphur. Together, these lines contain valuable information about the conditions and composition of the cloud.

Part of ionizing radiation may escape and power the *diffuse ionized gas* (DIG; Belfiore et al. 2022). By modelling the stellar source of the radiation and comparing it with the amount retained by the H II region, we try to measure the escape fraction in Chapter 8.

Since their evolution is closely linked to massive stars, they also share their short lifespan. This amounts to a maximum of a few million years, and in Chapter 7 we will try to narrow this down more precisely by connecting the regions with the stars. But sometimes it is even shorter. As the region expands into the ambient ISM, it gets diluted and vanishes. Especially the process described in the next section can accelerate this.

2.2.2 Supernova remnants

The destiny of a star is closely coupled to its mass – the heavier it is, the shorter its life span. Stars heavier than $\sim 8 M_{\odot}$ can reach stages such as carbon or oxygen burning, during which they fuse heavier elements, however, iron marks the end of the line for

all of them. Once fusion ceases, the star can no longer withstand gravitational collapse. What follows is an implosion of the core, which is abruptly halted and reversed into an explosion, known as a core collapse *supernova* (SN). However, not all SNe originate from a massive star. They can also occur when a WD in a binary system accretes mass from its companion and is pushed beyond the *Chandrasekhar limit*, leading to a thermonuclear explosion. As illustrated in Figure 2.1, those processes are responsible for producing many of the heavier elements in the universe.

The origin of the name suggests a new bright star, as opposed to the actual last blaze of a past star.

Such events are rare – only a handful have been recorded in our galactic neighbourhood over the past millennia – but whenever they occur, they attract a lot of attention. A famous example is SN 1572, extensively studied by Tycho BRAHE and others. The light curve of such a SN drops sharply, and the visibility ranges from a few days to several months, depending on the distance. While no longer visible to the naked eye, the ensuing shock wave excites the surrounding gas, causing it to glow. These *supernova remnants* (SNRs) are easily observable thanks to their strong emission lines. The remnant of SN 1054 is shown in Figure 2.2b. The entire process plays a crucial role in galaxy evolution by injecting both energy and heavy elements into the ISM. For the purposes of this thesis, however, SNRs mainly represent a contaminant: their strong emission lines must be identified and distinguished from those of other nebulae in order to avoid biases in our results.

2.2.3 Planetary nebulae

Long after the massive stars have disappeared – sometimes billions of years later – low to intermediate mass stars embark on their final journey. A *planetary nebula* (PN) is a brief stage in their evolution, during which they shine brightly for one last time. The path to get there is sketched out in Figure 2.3 and described in greater detail in Kwok (2000) and Kwitter & Henry (2022).

Those objects are unrelated to planets, hence this is another misnomer.

After the hydrogen burning ceases, the star leaves the main sequence. Its exact evolutionary track depends on the stellar mass: it climbs up the RGB and the AGB, and once the core consists mainly of carbon or oxygen, conditions are no longer suitable to fuse heavier elements. As in the case of SN explosions, the cessation of nuclear burning means that gravity takes the upper hand – but with one important difference. Here the core withstands the enormous pressure and heats up to $\sim 10^8$ K. As the core shrinks, the outer layers expand and form an envelope that is only weakly bound by gravity. Driven by radiation pressure, the strong stellar wind ejects the gas and ionizes it, producing beautiful and irregular morphologies like the one shown in Figure 2.2c.

The whole process happens on a very short time scale of around $\sim 10\,000$ yr (Carroll & Ostlie 2017). Yet because for the large number of stars that end their lives this way, and their relatively high brightness, PNe are plentiful in the sky. Herschel (1786) already noticed those objects in his catalogue and marked them as nebulae, resembling planets, hence the name. Only the spectral analysis by Huggins & Miller (1864b) started to reveal their true nature. Since their central stars are extremely hot and the surrounding gas has very low density, they show strong forbidden-line emission, most prominently [O III] $\lambda 5007$. This makes them easily detectable, and there are thousands of PNe in our Milky Way (Parker et al. 2003). Due to the strong stellar winds, they eject a lot of material

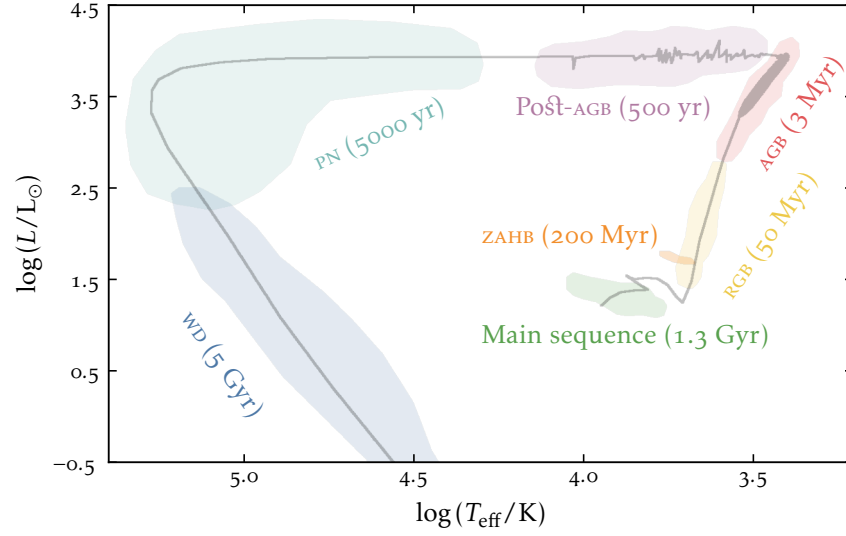


Figure 2.3: Evolution of a $2 M_{\odot}$ star in the *Hertzsprung–Russell diagram* (HRD). The track marks the different branches that the star passes through after it left the main sequence, along with how long it remains at each stage. Credit: this sketch is based on Herwig (2005).

into the ISM. Furthermore, as we will see in Chapter 5, PNe can also serve as valuable tools for measuring the distances to their host galaxies

2.3 BARYONIC CYCLE

In a sense, we have completed the stellar circle of life. The ISM acts as a melting pot where stellar life begins and ends, and where the ingredients are constantly stirred and reshaped. Stars are born from its raw material and give back the very elements that enrich the next generation – through winds, radiation, and violent deaths (Péroux & Howk 2020). Thus, the ISM and the stars are locked in a mutual dependency. This ongoing exchange of matter and energy is known as the *baryonic cycle*, a continuous flow that links the smallest dust grains to the grand evolution of galaxies, as illustrated in Figure 2.4.

The birth of new stars takes place in GMCs, which are the densest areas of the ISM. One can estimate the necessary conditions for star formation by assuming that the gas is in hydrostatic equilibrium. Comparing the gravitational potential to the pressure within the cloud yields a critical mass, above which the cloud will collapse under its own gravitational pull (Beech 2019). It is called the *Jeans Mass*

$$M_J = \left(\frac{5k_B T}{G\mu m_p} \right)^{\frac{3}{2}} \left(\frac{3}{4\pi\rho} \right)^{\frac{1}{2}}, \quad (2.3)$$

and depends on the temperature T and density ρ with the Boltzmann constant k_B , the gravitational constant G , the proton mass m_p , and the mean molecular weight μ . Pertur-

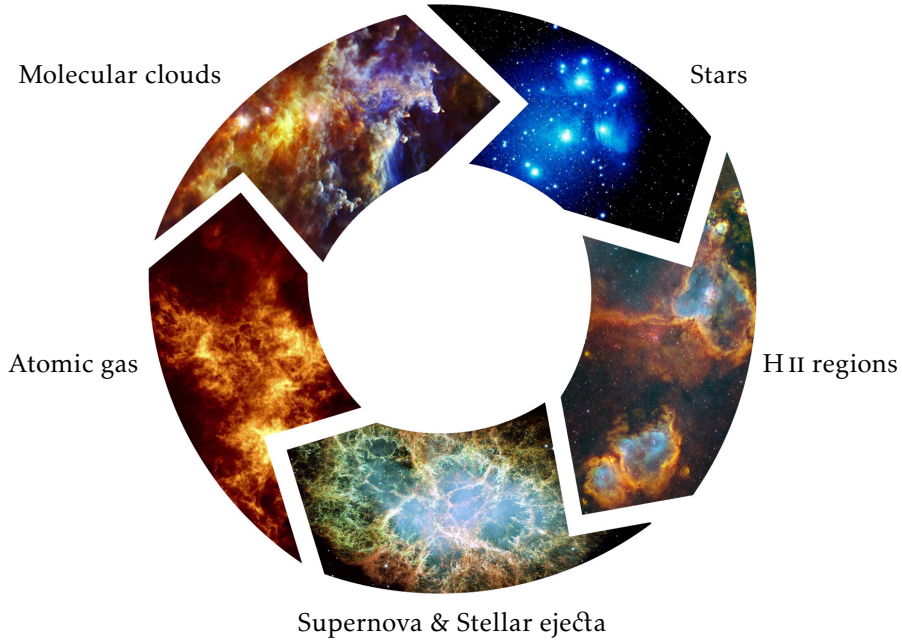


Figure 2.4: The baryonic life cycle. The atomic gas in the lower left panel cools and becomes molecular gas. Out of over densities, stars are born that ionize their surrounding. Enriched by supernovae and stellar winds, the gas cools and becomes atomic again. Credit: F. Santoro.

bations such as shock waves or colliding clouds can be sufficient to trigger the formation of new stars. The newly formed clusters can then ionize the left over gas in the ways described above. These emission line nebulae contribute to the galactic enrichment in metallicity. Their lifespan differs significantly, but eventually they will all cease to exist.

At that point, things calm down again. Without a source of radiation, the dispersed gas recombines to atomic H I (Goldsmith, Li & Krčo 2007). Further cooling is required until the material forms molecules that collect again in GMCs, allowing the whole cycle to repeat. From a galactic perspective, how active this whole process is depends largely on the available gas, as described by Schmidt (1959) and Kennicutt (1989)

$$\Sigma_{\text{SFR}} = \Sigma_{\text{gas}}^N. \quad (2.4)$$

This empirical relation between the *star formation rate* (SFR) and the surface gas density Σ_{gas} , known as the *Kennicutt-Schmidt law*, encapsulates how the baryon cycle regulates the pace of galactic evolution, setting the rhythm by which stars are born, live, and return to their origin.

BEFORE we dive into the data part, we need to establish some basic theories that form the foundation upon which this work is built. Stars are closely intertwined with the ISM, since they arise from it and later give something back. They contribute to the enrichment and even more importantly, enlighten the sky, so without them we would be left in the dark. Therefore, we start by introducing some fundamental stellar properties, followed by their formation in groups before ending with techniques that are helpful to model their impact.

Once we are back from that little detour, we shift our attention to the processes in the ionized gas. Many characteristics of a nebula can be derived from the emission lines in their spectra. After some physical background, we review a number of prescriptions used to measure features such as density, temperature, or metallicity.

3.1 STELLAR COMPONENT

Let us begin with the basics. Stars are often classified based on their spectral class, and sorted into the Morgan–Keenan system that goes O, B, A, F, G, K, and M (Morgan, Keenan & Kellman 1943). A common way to look at them is by scattering their luminosity (or absolute magnitude) against their color (or temperature, but be careful, the axis is inverted). Such a plot is called *Hertzsprung–Russell diagram* (HRD) and an example that contains stars from the Orion Arm is shown in Figure 3.1. Looking at the axes, one can see that their values span orders of magnitudes, but at the same time, there seems to be an underlying structure. Most of them are located along the diagonal which is called the *main sequence*, but there is also a second branch and faint outliers. This representation will prove helpful in understanding the evolution of stars, but also in drawing conclusions about the properties of stellar groups. In the following, all those subjects are condensed to the bare essentials and a more detailed description of stellar astrophysics can be found in books such as Collins (1989) or Beech (2019).

Oh, Be A Fine Girl: Kiss Me!

3.1.1 Single Stars

Not every object that lights up the night sky is a star. The most obvious counterexample are planets, which merely reflect the light. A star is a self-luminous object, and one must be careful to draw a demarcation line to similar objects, such as brown dwarfs (Rebolo, Zapatero Osorio & Martín 1995; Fernie et al. 2025). It is difficult to separate them (Burrows & Ostriker 2014), but since we are primarily interested in massive stars, the distinction is not critical for our purposes.

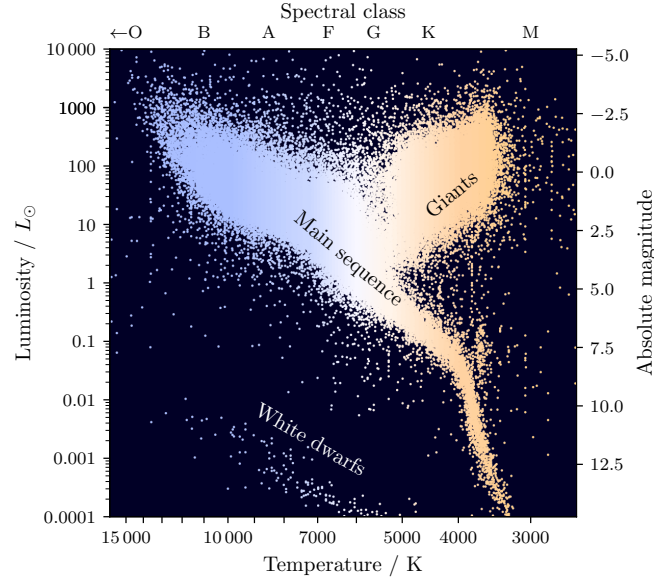


Figure 3.1: *Hertzsprung–Russell diagram* for the stars in our solar neighbourhood ($D < 1000\text{pc}$). Credit: the data used in this figure is from the [HYG database](#) (version 4.1), a compilation of the Hipparcos, Yale Bright Star, and Gliese catalogues, and is licensed under CC BY 4.

For a long time, the source of energy remained a mystery, as neither chemical processes, nor the thermal time scale, also known as the *Kelvin–Helmholtz time scale*, can supply sufficient energy over a large enough timescales (Mayer 1848). A suitable candidate did not appear until Harkins & Wilson (1915) and Perrin (1919) recognized that energy can be produced by nuclear fusion, according to Einstein’s (1905) mass–energy equivalence. Building on this idea, Eddington (1920) proposed that nucleosynthesis is powering the stars. This theory was further elaborated by Weizsäcker (1937, 1938) and Bethe (1939) among others.

For most of its life, a star generates energy by fusing hydrogen into helium. The simplest example is the *proton–proton chain* (PP chain). Two protons fuse to form deuterium, releasing a positrons and an electron neutrino



The deuterium nucleus captures another proton to form helium-3 and a gamma ray



Finally, two helium-3 nuclei fuse to form one helium-4 and two protons



This is just one possible branch of the PP chain and in more massive stars, the *Carbon–nitrogen–oxygen* (CNO) cycle dominates instead. According to $E = mc^2$, the slight loss

These are very easy to rule out, since even on Earth we are aware of events older than the thereby possible lifetime of the Sun.

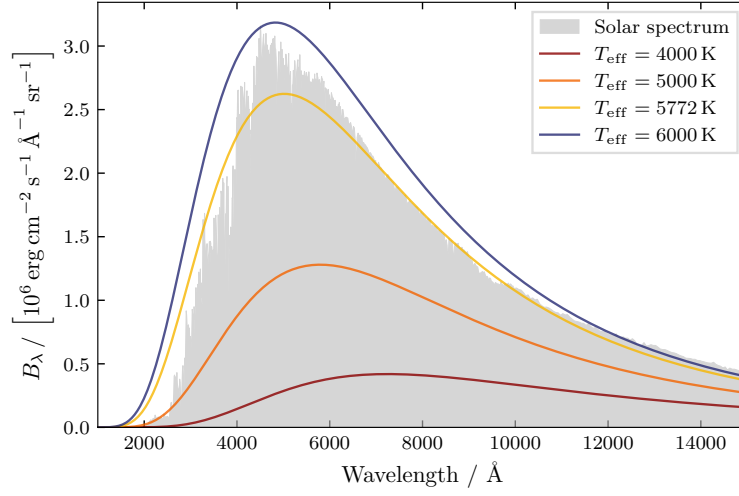


Figure 3.2: The solar spectrum as measured by Wehrli (1985) is displayed in the gray shaded. Overplotted are a number of black-body spectra with different effective temperatures. The assumed photosphere temperature around $T_{\text{eff}} = 5772$ K (Prša et al. 2016) fits the curve quite well, especially at longer wavelengths. However emission lines and at lower temperatures effects such as limb darkening result in a visible differences.

of mass releases a huge amount of energy. The energy produced in the core diffuses outwards and is eventually emitted in radiation that resembles *black-body spectrum*, a term coined by Kirchhoff (1860). The spectrum depends only on the effective temperature T_{eff} and can be described with Planck’s law (1901)

$$B_{\lambda}(\lambda, T_{\text{eff}}) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left(\frac{ch}{\lambda k_B T_{\text{eff}}}\right) - 1}, \quad (3.4)$$

with the Planck constant h , the speed of light c and the Boltzmann constant k_B . The peak of the spectrum, and thus the color at which we perceive it, changes with temperature according to *Wien’s displacement law* (1893)

$$\lambda_{\text{max}} = \frac{b}{T}, \quad (3.5)$$

where b is Wien’s displacement constant. By integrating over all wavelengths, one can derive the *Stefan–Boltzmann law* that describes the total intensity

$$F = \sigma_{\text{SB}} T^4, \quad (3.6)$$

with the Stefan–Boltzmann constant σ_{SB} . It is extremely temperature dependent, as is the luminosity, which can be obtained by integrating the flux over the surface area. Figure 3.2 shows the spectrum of the Sun, overplotted with black-body spectra of different effective temperatures. Due to interaction in the solar atmosphere, the observed spectrum features additional dark lines, which are the topic in Section 3.2.

Table 3.1: Stellar classification with some properties (Habets & Heintze 1981).

Class	Temperature K	Luminosity L_{\odot}	Mass M_{\odot}
O	$\geq 30\,000$	$\geq 30\,000$	≥ 16
B	10 000 – 30 000	25 – 30 000	2.1 – 16
A	7500 – 10 000	5 – 25	1.4 – 2.1
F	6000 – 7500	1.5 – 5	1.04 – 1.4
G	5300 – 6000	0.6 – 1.5	0.8 – 1.04
K	3900 – 5300	0.08 – 0.6	0.45 – 0.8
M	2300 – 3900	≤ 0.08	0.08 – 0.45

As we saw, the temperature is an essential parameter of the star, but what actually determines it in the first place? The fundamental parameter is the stellar mass. More massive stars have a stronger gravity, which compresses the core, raises the pressure and temperature, and drives nuclear fusion at much faster rate. Halm (1911) already noticed this relation and Kuiper (1938) later formulated an exponential dependence

$$\frac{L}{L_{\odot}} = \left(\frac{M}{M_{\odot}} \right)^a. \quad (3.7)$$

The exponent varies between $1 < a < 6$, depending on the mass, but $a = 3.5$ is used for most of the main sequence. This gives us a basic understanding of how stars work. Table 3.1 provides an overview of the individual classes.

Stellar evolutionary models such as the Geneva models (Ekström et al. 2012; Georgy et al. 2012, 2013), the Padova models (Girardi et al. 2000), or MIST (Choi et al. 2016; Dotter 2016) calculate how stars change over time. The models are built on fundamental physics including hydrostatic equilibrium, nuclear fusion rates, energy transport, and mass loss. They differ in the physical assumptions they adopt, for example metallicity, rotation, and binarity, and they predict stellar properties such as luminosity, effective temperature, radius, and emitted spectrum. Based on this, they produce evolutionary tracks that describe how stars of different initial masses evolve as they age. Because they also predict the emergent radiation, the models can be directly compared with observations, making them the essential link between stellar physics and stellar observations.

With that established, we can take a closer look from another perspective at the HRD. The left panel in Figure 3.3 shows models for individual stars. With their help, it is possible to derive the properties of a star based on its position in the diagram. Every star spends most of its life on the main sequence, indicated by the gray stripe, but eventually it will leave this phase. The coloured stripes show the evolution each star takes after the main sequence. This happens once the star has used up all the hydrogen in his core. Heavy stars are able to fuse heavier elements, but they consume their fuel at a much faster rate. While a solar mass star can shine for billions of years, the duration for the

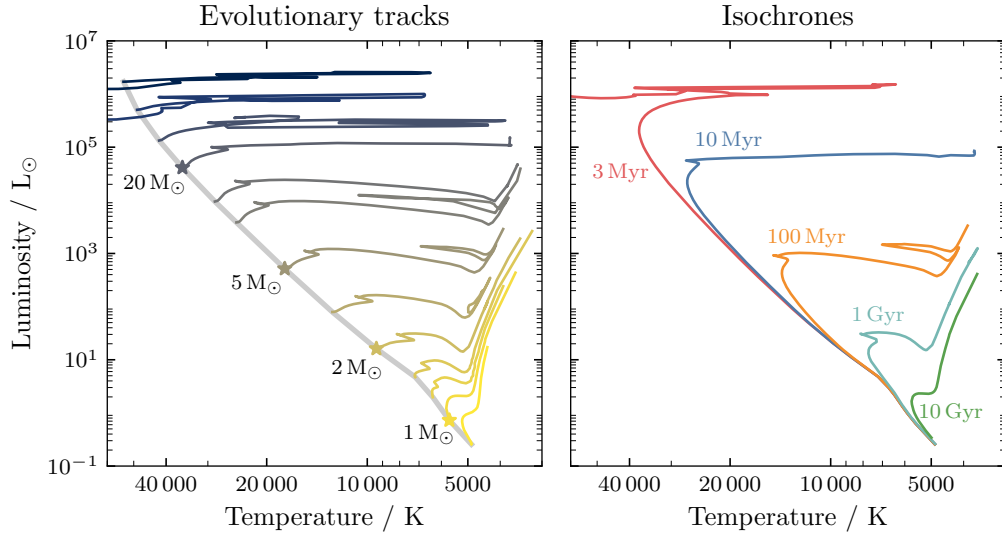


Figure 3.3: Stellar tracks in the HRD, based on the Geneva models (Ekström et al. 2012), shown from two perspectives. *Left*: temporal evolution for stars with different masses. *Right*: isochrones for groups of stars at a certain age.

large ones is reduced to just a few million years. As the core heats up, the outer layer of the star expands and appears cooler. Due to this extreme expansion, they appear brighter despite their cooler temperature as they go through the RGB and AGB. The whole thing culminates as a PN or SN, the remains of which are a WD, a neutron star or a *black hole* (BH).

Understanding how individual stars evolve provides us with the foundation to tackle a bigger question: how do entire stellar populations behave? To answer that, we turn to groups of stars that formed together and whose collective properties we want to derive from our observations.

3.1.2 Clusters and associations

Stars predominantly form in clusters or associations (Lada & Lada 2003; McKee & Ostriker 2007). In nearby galaxies beyond our Local Group, many of them will appear as point-like sources or only be marginally resolved. The formation process happens on a short timescale, allowing us to assume that all stars in the group are coeval and made from the same material. As we saw in Equation 2.3, the Jeans criterion shows that only gas clouds above a certain mass are gravitationally unstable and collapse. Even at relatively low densities and temperatures, this critical mass is substantial, meaning that star formation tends to occur in GMCs rather than in small, diffuse clumps. As a result, most stars form in clusters whose total masses range from 10^2 up to $10^6 M_\odot$.

Though there might be multiple stellar populations (Gratton et al. 2004).

Stars of different masses behave and evolve differently, and so in order to model a group of stars, we have to know the inventory we are dealing with. The distribution of mass of the individual stars within a population is not equal, with less massive stars dominating the overall mass. As Salpeter (1955) noticed, their mass follows a decreasing power law called the *initial mass function* (IMF). In reality, the distribution flattens below $1 M_{\odot}$, as in the prescription from Kroupa (2001)

$$\xi(m) \propto m^{-\alpha} \quad \text{with} \quad \begin{cases} \alpha = 0.3 & \text{for } m/M_{\odot} < 0.08 \\ \alpha = 1.3 & \text{for } 0.08 < m/M_{\odot} < 0.5 \\ \alpha = 2.3 & \text{for } 0.5 < m/M_{\odot}. \end{cases} \quad (3.8)$$

Another description is the more complex model by Chabrier (2003). Above a certain cluster mass, the IMF is fully sampled and all outputs scale linearly with stellar mass. Although the lower mass stars dominate the overall mass, the massive OB stars are responsible for the evolution of the cloud. They produce the largest amount of stellar feedback, and are consequently mainly responsible for the enrichment and the origin of the ionizing radiation.

The next step is to take the models of the individual stars and combine them, weighted by their share of the cluster. Such a track for a given timestamp is called an *isochron*. The right panel in Figure 3.3 shows them at different ages. As a cluster ages, its most massive stars evolve off the main sequence and eventually die, leaving progressively lower-mass stars at the main sequence turn-off. This turn-off point allows us to determine the age of the cluster. The location of this turn-off point – the point where stars begin to leave the main sequence and move toward the giant branches – directly reflects the cluster’s age: the lower the mass of stars at the turn-off, the older the cluster.

A *population synthesis model* then samples the IMF and sums up the contribution from evolutionary tracks of the individual stars. Popular models are STARBURST99 by Leitherer et al. (1999), the one from Bruzual & Charlot (2003), *Binary Population and Spectral Synthesis* (BPASS) from Eldridge & Stanway (2009), or *Stochastically Light Up Galaxies* (SLUG) by da Silva, Fumagalli & Krumholz (2012). The final step is then to use another tool to compare the observed *spectral energy distribution* (SED) to the ones predicted by the models. This can for example be done with *Code Investigating GALaxy Emission* (CIGALE) by Boquien et al. (2019).

We can divide the stellar groups we observe into three categories. The ones associated to nearby H α emission have to be formed recently, as the short life time of massive stars limits how long this can happen. Depending on the conditions where they form, some of them are loosely bound and disperse quickly and we call those *stellar associations*. The ones that are more densely packed and resemble a point source are *star clusters*. For the most massive ones, the gravitational bond can be so strong that they stay together for a very long time – sometimes billions of years. These are known as *globular cluster*, but since at this age they consist solely of low-mass stars that are incapable to ionize gas, they are less relevant to us.

3.2 EMISSION LINE DIAGNOSTICS

The spectra emitted by astronomical objects reveal a wealth of information about their origin. Depending on the nature of the emitter, different types of spectra arise. The most familiar example is probably the spectrum of a star, which resembles a black-body curve. An even more precise example is the *cosmic microwave background* (CMB) – a relic of the epoch when the opaque universe became transparent after the recombination – whose spectrum matches a perfect black-body with extraordinary precision. Beyond thermal emission, objects can also radiate through processes, such as synchrotron radiation or bremsstrahlung, and some sources, such as *active galactic nuclei* (AGNs), often feature a mixture of these mechanisms.

The spectrum of a star is not a perfect black-body, as the radiation must pass the stellar atmosphere, where it interacts with the gas.

Among these possibilities, line spectra are of particular importance to us. The strength and shape of these lines are set by the physical conditions of the gas, such as temperature, density, ionization state, and metallicity. They appear either as missing features in a continuum (*absorption lines*) or as additional lines that exist on their own (*emission lines*). Many of them are *recombination lines* which originate when free electrons recombine with ions and cascade down through discrete atomic energy levels. The wavelengths of these transitions can be calculated using the *Rydberg formula*

$$\frac{1}{\lambda} = R_H \left(\frac{1}{n_{\text{initial}}^2} - \frac{1}{n_{\text{final}}^2} \right), \quad (3.9)$$

where R_H is the Rydberg constant for hydrogen and n_{initial} and n_{final} are the principal quantum numbers of the lower and upper orbits, respectively (Rydberg 1890). This behaviour was later explained with the *Rutherford–Bohr model* from Rutherford (1911) and Bohr (1913). In absorption, photons are removed from the incoming radiation field as electrons are lifted to higher orbits. In emission, atoms with excited electrons release photons at well-defined wavelengths. Classic examples are the Lyman series (transitions to the ground state, $n = 1$) and the Balmer series (transitions to the first excited state, $n = 2$). The Balmer series includes $H\alpha$ ($n = 3$ to 2 at $\lambda = 6562 \text{ \AA}$) and $H\beta$ ($n = 4$ to 2 at $\lambda = 4861 \text{ \AA}$), which are both strong and visible in the optical, making them ideal tracers of ionized gas in nebulae and galaxies.

In contrast to recombination lines, some emission lines are highly improbable under laboratory conditions and are therefore referred to as *forbidden lines*. The associated transitions are extremely slow – in a man-made vacuum, they should collisionally de-excite long before they can emit a photon. In the low-density ISM, collisions are rare, allowing these long-lived states decay radiatively and produce strong spectral features. The historical name has been retained, and for notation, they are written in square brackets, for example [OIII] or [SII].

The bright green lights of the *aurora borealis* are the namesake to the final category we mention here. After its wavelength was measured by Ångström (1869), among others, it was later identified as [OI] λ 5577 (Kragh 2009). This line, along with other *aurora lines* such as [OIII] λ 4363 or [NII] λ 5755, arises from highly excited, metastable states. Because these upper levels are sparsely populated, the corresponding transitions are

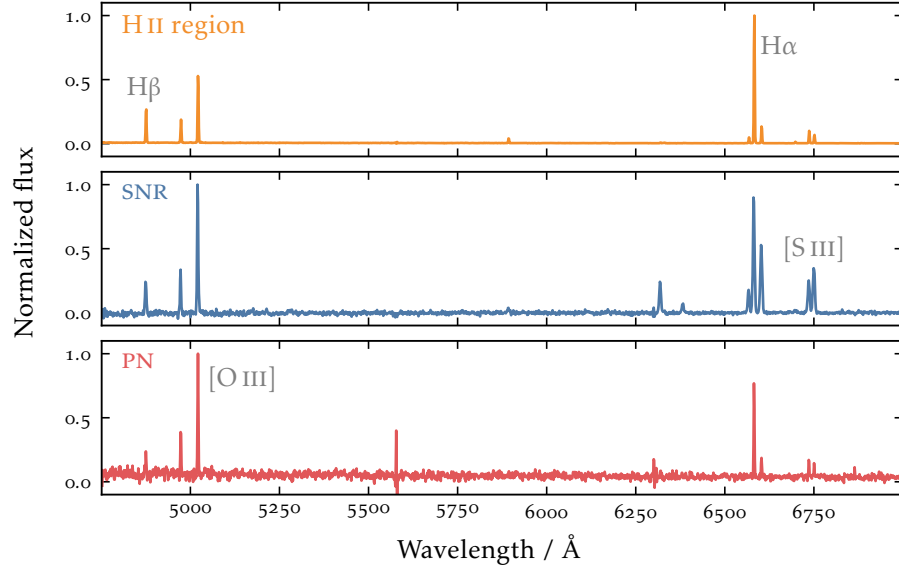


Figure 3.4: Spectra of an H II region, a SNR and a PN. All three are dominated by emission lines such as H α and H β , [O III] λ 5007, or [S II] λ 6717 + [S II] λ 6730, but their ratios vary due to differing physical conditions. The spectra are background subtracted and normalized based on masks from the catalogue by Congiu et al. (2023).

relatively rare, making these lines much fainter than the more common recombination or low-excitation forbidden lines.

All elements possess a wealth of unique emission lines, and the photons they emit can be unambiguously assigned to their element of origin. As the name suggests, the spectra of emission line nebulae are dominated by these atomic transitions. By carefully measuring and comparing these line intensities, we can use them as powerful diagnostics to classify nebulae and to infer the physical properties of the emitting gas.

3.2.1 Classification

When studying nebulae, the first essential step is to determine their type. For galactic nebulae, which are spatially resolved, this can often be done based on their morphology, whereas for extragalactic objects we rely on their spectra. Figure 3.4 depicts example spectra for an H II region, SNR and PN. At first glance, they may appear similar, but each is dominated by a characteristic set of emission lines that reflects the underlying excitation mechanism.

For the brightest nebulae within galaxies, a secure identification is often straightforward: they are usually H II regions, since no single star can rival the ionizing output of a massive cluster of OB-type stars. These regions exhibit strong hydrogen recombination lines, particularly H α and H β , along with additional forbidden lines such as [O III], [N II],

and [SII]. For less massive HII regions, the same lines are present but with lower absolute fluxes, making careful spectral analysis necessary for identification. In PNe, the central ionizing source is a hot post-AGB star. Its ultraviolet radiation is capable of doubly ionizing oxygen, so [OIII] $\lambda 5007$ is often the strongest line in the spectrum, dominating over the Balmer lines in many cases. In contrast, SNRs are powered by shock heating rather than photoionization. Their spectra are characterised by enhanced [SII] $\lambda 6716$ and $\lambda 6731$ relative to $H\alpha$, a feature often used to distinguish them from HII regions. Additionally, the rapid expansion of the shock front broadens the emission lines, providing another spectral signature of their origin.

To classify a large sample of objects, astronomers traditionally use a comparison between different line ratios. These ratios are independent of distance or total luminosity, yet remain sensitive to the physical conditions and excitation mechanisms within the gas. The *Baldwin–Phillips–Terlevich* (BPT) diagram is a standard tool for this purpose (Baldwin, Phillips & Terlevich 1981). By plotting ratios such as [OIII] $\lambda 5007/H\beta$ against [NII] $\lambda 6584/H\alpha$, it is possible to separate star-forming regions, shock-excited nebulae, and AGNs in a systematic and quantitative way. Nowadays, there is a wide range of selection criteria available. Kewley et al. (2001a,b, 2006) and Kauffmann et al. (2003) provide commonly used schemes to separate star formation from AGNs, while Ciardullo et al. (2002) and Herrmann et al. (2008) present criteria to identify PNe.

However, this approach is not without flaws. In practice, the sources occupy continuous regions of the diagram, yet the classification divides them with sharp demarcation lines. Moreover, only a fraction of the available spectral information is typically used, and some lines are so faint that their use is accompanied by a high degree of uncertainty. This incentivises us to refine the criteria and make them more reliable (Congiu et al. 2023).

3.2.2 Physical properties

The strength of the emission lines depends on many physical parameters, which can in turn be inferred from their observed intensities. By comparing observed line intensities to predictions from physical models, it is possible to infer these underlying conditions. Theoretical calibrations are constructed using detailed photoionization simulations. Codes such as CLOUDY (Ferland et al. 2017) or MAPPINGS (Sutherland et al. 2018) compute the spectrum produced under specified conditions, allowing one to predict how different lines respond to changes in temperature, density, and ionization.

Not all diagnostic lines are easily measurable in real observations. Some lines are intrinsically faint, appear only under rare physical conditions, or may fall outside the observed wavelength range. To address these limitations, empirical calibrations are developed by identifying robust correlations between easily observed lines and the desired physical properties. These correlations provide practical tools for deriving physical conditions in large samples of objects, even when faint or auroral lines are unavailable.

In practice, the physical properties of ionized gas can be derived using either theoretical or empirical methods. Tools such as PYNEB (Luridiana, Morisset & Shaw 2015) allow the computation physical properties from observed line intensities, using either approach. In this work, we focus on a selection of diagnostics most relevant to our study. For a

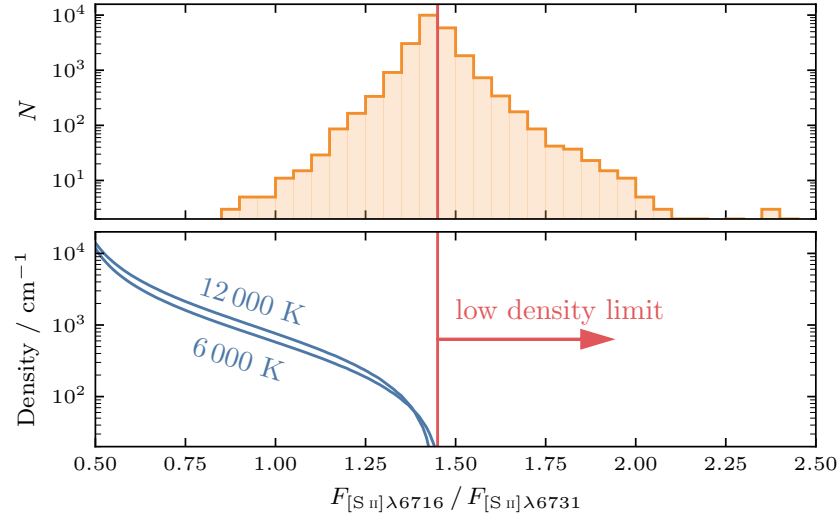


Figure 3.5: Distribution of the $[\text{SIII}]/[\text{SII}]$ ratio in the nebular catalogue by Groves et al. (2023) and the electron density predicted by PYNEB based on the $F_{[\text{SII}]\lambda 6716}/F_{[\text{SII}]\lambda 6731}$ ratio at two different temperatures.

comprehensive overview of the many available prescriptions, we refer to the review by Kewley, Nicholls & Sutherland (2019).

As a starting point, we consider the *electron temperature* T_e . The classical way to measure it is through auroral lines, whose strengths are primarily sensitive to temperature. A common ratio is between $[\text{NII}]\lambda 5755$ and a strong line such as $[\text{NII}]\lambda 6548$ or $[\text{NII}]\lambda 6584$. They are generally faint and only detectable in a handful of H II regions (Ho et al. 2019). In order to determine the temperature of the rest, empirical methods that are based on strong lines are calibrated (Kreckel et al. 2022).

Next, we turn to the *electron density* n_e . It can be derived from the $[\text{SII}]\lambda 6717/[\text{SII}]\lambda 6731$ ratio, but this diagnostic exhibits a weak temperature dependence. In Figure 3.5, we use PYNEB to compute this for a sample of H II regions, finding that roughly half of the nebulae fall in the low-density limit. The result is only moderately sensitive to temperature, but instead of assuming a fixed temperature, one could also set T_e as a free parameter that is jointly determined during the fit.

The *ionization parameter* q is defined as the ratio of the incident ionizing photon flux Φ_{H^0} to the local hydrogen density n_{H} (Kewley & Dopita 2002). For a spherical geometry, it can be expressed as

$$q = \frac{\Phi_{\text{H}^0}}{n_{\text{H}}} = \frac{Q(\text{H}^0)}{4\pi R^2 n_{\text{H}}}, \quad (3.10)$$

where $Q(\text{H}^0)$ is the rate of ionizing photons striking the cloud at a distance R . It is often convenient to use the dimensionless form $q = U/c$. A good proxy is the sulphur ratio $[\text{SIII}]\lambda(9069 + 9532)/[\text{SII}]\lambda(6717 + 6731)$, based on a calibration by Diaz et al. (1991) 5

$$\log U = (-1.684 \pm 0.076) \cdot \log([\text{SIII}]/[\text{SII}]) - (2.986 \pm 0.027). \quad (3.11)$$

With the physical conditions of the gas constrained, the final step is to determine its chemical composition. Measuring metallicity is challenging because line strengths depend not only on the abundance of the element but also on T_e , n_e , and the ionization structure. The most reliable way is with the direct T_e method, which relies on auroral lines such as $[\text{OIII}]\lambda 4363$ or $[\text{NII}]\lambda 5755$ (Osterbrock & Ferland 2006). This limits this method to very few bright HII regions where the auroral lines can be detected, but photoionization models can help to calibrate alternatives.

That brings us to empirical strong line methods. They only use strong nebula lines and are calibrated against the previously mentioned models. Unlike the other diagnostics, metallicity cannot be traced by a single line ratio. Instead, these calibrations combine multiple lines to account for the influence of both excitation and abundance. Two widely applied families are the *R calibration*, which employs N_2 , R_2 and R_3 , and the *S calibration*, which uses N_2 , R_3 and S_2 (Pilyugin & Grebel 2016). The ratios mentioned above are

$$N_2 = I_{[\text{NII}]\lambda 6548 + [\text{NII}]\lambda 6584} / I_{\text{H}\beta \lambda 4861} \quad (3.12)$$

$$R_2 = I_{[\text{OII}]\lambda 3727 + [\text{OII}]\lambda 3729} / I_{\text{H}\beta \lambda 4861} \quad (3.13)$$

$$R_3 = I_{[\text{OIII}]\lambda 4959 + [\text{OIII}]\lambda 5006} / I_{\text{H}\beta \lambda 4861} \quad (3.14)$$

$$S_2 = I_{[\text{SII}]\lambda 6717 + [\text{SII}]\lambda 6730} / I_{\text{H}\beta \lambda 4861}. \quad (3.15)$$

Since $[\text{OII}]\lambda 3727$ – and thus also R_2 – lies often outside the range of the spectrograph, we are limited to the S calibration.

Together, these emission line diagnostics provide a comprehensive view of the physical conditions in ionized gas. Recombination lines trace the amount of ionized material, forbidden and auroral lines allow precise measurements of temperature and density, and combinations of strong lines can yield reliable metallicities even when auroral lines are too faint to detect. By carefully combining these measurements, we can characterize the ionization structure, excitation conditions, and chemical composition of HII regions. This provides fundamental insights into star formation, feedback, and chemical evolution across the galaxies.

IN the previous chapter, we covered the theoretical part, but before we move on to the practical side, we first need the data to which we can apply it. As we will see, the requirements for the observing instruments are high, so that only a few telescopes are capable to deliver them. The days when astronomers gazed at the stars alone are long gone. Nowadays, it takes large organisations such as the *European Southern Observatory* (ESO), the *European Space Agency* (ESA) or the *National Aeronautics and Space Administration* (NASA) to build and operate these sophisticated telescope facilities. And due to the high demand for the available observing time supplied, the application to obtain observation time can be a real competition.

In addition, this work relies on a large sample that has to be observed not just by one, but several telescopes. And once the vast amounts of data have been recorded, they are not yet ready for evaluation, but must first be reduced. For an individual, this represents a daunting task and one would spend a lifetime dealing with it. This is the reason why most of such projects are carried out in well-coordinated groups these days. The bundled knowledge makes it possible to approach the issue from different angles. In this chapter we discuss the telescopes that were used and the collaboration that made this work possible.

4.1 TELESCOPES

We start with a broad overview of the basic parameters of telescopes. There have been continuous improvements over the centuries, but perhaps the most groundbreaking innovation was the switch from analogue to digital images. While photographic plates possess some advantages, such as their resolution, they are less convenient and more complicated to evaluate (Turriziani 2019). An alternative emerged when the *charge-coupled device* (CCD) was conceived by Boyle & Smith (1970) and shortly after used for astronomical observations by Smith (1976).

The *field of view* (FoV) is the angular size of the sky that is covered in a single observation. It is set by the focal length of the telescope and the size of the detector, and the areas we are dealing with are tiny. To illustrate this, our targets of interest are only a few arcmin across while a whole sphere spans a whopping

Our Moon has a solid angle of $\Omega = 758 \text{ arcmin}^2$

$$\Omega = 4\pi \text{ sr} = 4\pi \left(\frac{180}{\pi}\right)^2 \text{ deg}^2 = 148 \cdot 10^6 \text{ arcmin}^2. \quad (4.1)$$

Another parameter is the diameter of the aperture, and it has several implications. The larger the telescope, the more sensitive it is, because the amount of light it collects scales with the square of the diameter. For a long time, refracting telescopes dominated the

scene, but as they grew in size, it became increasingly difficult to construct the lens (Starr 1897). Once we were able to construct reflecting telescopes, they took over.

The size also affects the angular resolution that a telescope can resolve. This is limited by diffraction and aberration and the former originates from the waveform of light. Even a perfect lens will result in circular pattern, called the *Airy disk*, named after Airy (1835). We can use *Rayleigh's criterion* to calculate the angular resolution

$$\theta \approx 1.22 \frac{\lambda}{D}, \quad (4.2)$$

where λ is the observed wavelength and D is the diameter of the telescope (Rayleigh 1879). So in order to observe something at a longer wavelength with the same resolution as at a shorter one, a larger diameter is required. As an example, if we observe in the optical at 500 nm with an 1 m telescope, we have an angular resolution of 0.13 arcsec. At a distance of $d = 15$ Mpc this corresponds to a physical resolution of

$$\Delta l \approx \theta \cdot d = \frac{0.13 \text{ arcsec}}{206265} \cdot 15 \text{ Mpc} = 9.15 \text{ pc}. \quad (4.3)$$

At this point, another issue arises – the aforementioned aberration. Turbulence in the atmosphere disturbs the incoming light and limits the angular resolution to ~ 1 arcsec (Chromey 2010). The easiest way to reduce the seeing is to pick an observation site with as little atmosphere as possible above. This is one reason why most observatories today are located at high altitude in remote mountainous terrain. And this has further advantages. Since electrification in the late 19th century, artificial light causes light pollution, especially near large cities, which makes it much more difficult to observe faint sources (see European Southern Observatory 2025).

Even at high altitudes, atmospheric seeing still degrades the images, but with a technique called *adaptive optics* (AO), this distortion can be minimized (Hippler 2019). To correct for it, a reference point source with a well-defined *point spread function* (PSF) is chosen – either an unresolved guide star or an artificial one created by shooting a laser into the sky. A computer then calculates the atmospheric distortion and adjusts the shape of a deformable mirror accordingly to remove the aberration. This correction is repeated continuously to keep the observed image as sharp as possible.

The atmosphere has an even bigger disadvantage. For most parts of the spectrum, it is opaque and only a small range of wavelengths can make it through. This includes the *optical window*, which allows light visible to the human eye to pass through. Beyond that, only a few other ranges in the IR and radio exist, while others, such as the UV, are mostly absorbed. The only way to get around this is by moving the observation site into space, which is significantly more costly and also entails other limitations.

In the end, different targets have different requirements on the telescope. When picking a suitable telescope, one has to consider the following points:

- (i) the wavelength at which the target is observed.
- (ii) the desired physical resolution.
- (iii) the signal-to-noise (sensitivity and detection limit).

To put this into perspective,
one could resolve an euro
coin at a distance of 38 km.

The ESO is concerned about
the plans to construct an
industrial complex close to
Paranal, which would
worsen the observations.

Next we go over the specific telescopes that are relevant for us. Much of the details are repeated in the chapters where the methods are applied to the data, so this is more of a look at what distinguishes the individual telescopes.

4.1.1 MUSE

Conventional detectors can observe colours or emission lines only one at a time. For this purpose, a filter that only allows the desired wavelengths to pass through is inserted. Since this has to be repeated for each desired filter, the process is protracted and inefficient. One could also point a spectrometer at the object, but this requires that one already knows the position of the target and again, the process has to be repeated for each target.

Integral field spectroscopy (IFS) revolutionized this field. Instead of a spatial 2D image, a 3D data cube is recorded, in which the third dimension represents the spectral axis. To achieve this, the instrument is equipped with an *integral field unit* (IFU) that splits and redirects the incoming light to spectrographs Hagen & Kudenov (2013). In this way, each pixel becomes a »spaxel« that contains a full spectrum.

The *Multi Unit Spectroscopic Explorer* (MUSE) is the most powerful implementation of this technique (Bacon et al. 2010). In 2014, it was mounted at the *Very Large Telescope* (VLT), which itself is located on the summit of Mount Paranal in the Atacama Desert of Chile, as shown in Figure 4.1a. It possesses an 8.2 m mirror which makes it sensitive enough to detect faint objects in distant galaxies. The FoV depicts 1×1 arcmin, sampled at 0.2 arcsec per spaxel, and thus it is capable of recording 90 000 spectra simultaneously. Those cover the wavelengths between 4800 to 9300 Å and have a spectral sampling of 1.24 Å per spaxel.

Other studies that use MUSE to study galaxies are *Generalising Edge-on galaxies and their Chemical bimodalities, Kinematics, and Outflows out to Solar environments* (GECKOS) by van de Sande et al. (2024), *The MUSE Atlas of Discs* (MAD) by den Brok et al. (2020), or *Time Inference with MUSE in Extragalactic Rings* (TIMER) by Gadotti et al. (2019).



(a) Very Large Telescope.



(b) Hubble Space Telescope.

Figure 4.1: *Left*: The VLT consists of four telescopes and MUSE is mounted in the one furthest to the right (Yepun, Evening Star). Credit: ESO, G. Hüdepohl. *Right*: *HST* photographed during the last service mission in 2009. Credit: NASA.

An earlier launch was planned, but the disaster of the Space Shuttle Challenger delayed things.

4.1.2 HST

The launch of the *Hubble Space Telescope* (HST) in 1990 ushered in a new era for astronomy. After some initial problems, it started delivering superb images and continues to do so today. This is partly thanks to multiple service missions that have maintained and upgraded it (Hartnett et al. 2018; Gainor 2020). During the last mission in 2009, the crew of the space shuttle has taken Figure 4.1b. For a long time, it was the most expensive telescope, but it has proven to be a valuable investment for science.

The satellite has a mirror with a diameter of $D = 2.4$ m and houses a number of instruments. The primary ones for wide-field imaging are the *Advanced Camera for Surveys* (ACS) and the *Wide Field Camera 3* (WFC3). The latter has a pixel scale of 0.04 arcsec per pixel and a FoV of 2.7×2.7 arcmin. This makes it possible to observe the majority of the galaxies in our sample with a single pointing. Both instruments offer a large number of broad-band filters (Hathi 2024; Pagul & Rivera 2024), such as F275W (NUV), F336W (U), F438W (B), F555W (V), and F814W (I). In addition, there are some narrow-band filters, such as F658N or F657N for $H\alpha$, and F502N for $[OIII]\lambda 5007$.

Existing studies, such as *Legacy ExtraGalactic Ultraviolet Survey* (LEGUS), have already used this to study the formation and evolution of stars in nearby galaxies (Calzetti et al. 2015).

4.1.3 AstroSat

AstroSat is a space telescope operated by *Indian Space Research Organisation* (ISRO) that provides access to the *Far ultraviolet* (FUV) regime through its *Ultra-Violet Imaging Telescope* (UVIT). With a median angular resolution of 1.4 arcsec, it clearly outperforms all predecessor such as *Galaxy Evolution Explorer* (GALEX), which only resolved down to 4.5 arcsec (Martin et al. 2005). Together with its wide circular FoV spanning 28 arcmin, *AstroSat* is particularly well suited for wide-field mapping of young stellar populations and recent star formation in nearby galaxies (Singh et al. 2014). While its spatial resolution is not sufficient to match the scales probed by *HST*, it offers valuable complementary information when combined with MUSE by directly tracing the FUV emission from hot, massive stars and their surrounding gas.

4.1.4 Ancillary telescopes

Our work is mainly based on the three telescopes presented above. While not used in this thesis, they are number of telescope that deliver data to related topics. Two of them are particularly noteworthy:

We also want to observe the molecular gas, which works best with the CO(2–1) line at 1.3 mm. According to Equation 4.2, a telescope with a diameter of over 300 m is required to get a resolution of 1 arcsec. This is not possible in one piece, and this is where interferometry comes into play. By linking several smaller telescopes together, one can achieve the same resolution as a single giant telescope. *Atacama Large Millimeter Array* (ALMA) is such an interferometer, located on the high plateau of the Atacama Desert

in Chile. It consists of 66 individual antennas that can be repositioned across the site, allowing to adjust the array's configuration depending on the needs of the observation.

The Earth's atmosphere absorbs most infrared light, allowing only a few narrow windows to reach the ground. In order to observe the full range, we have to go to space again. With the launch of the *James Webb Space Telescope* (JWST) in 2021, we can finally observe the galaxies in the *near-infrared* (NIR) and *mid-infrared* (MIR).

Other telescopes worth mentioning are *Wide Field Imager* (WFI) and DuPont which provide $H\alpha$ maps (Razza et al. [in preparation](#)) and among other things, play a central role in the data reduction of MUSE.

4.2 PHANGS

Physics at High Angular resolution in Nearby Galaxies (PHANGS) is an international collaboration that aims to study star formation in nearby galaxies, by utilizing state-of-the-art telescopes (Schinnerer et al. [2019](#)). At the core of this are observations with the four telescopes ALMA, *HST*, JWST, and MUSE that provide a detailed census of GMCs, stellar populations and H II regions. Paired with the large sample size, it enables us to perform an unprecedented study of the galactic structure and star formation cycle. To maximise the output, the observations were tailored to cover these processes as comprehensively as possible. To achieve this, the galaxy sample was carefully selected based on a number of criteria

www.phangs.org

- (i) They need to be observable from the southern hemisphere.
- (ii) Nearby to guarantee a good enough physical resolution.
- (iii) Low inclination to avoid too much extinction.
- (iv) Massive star-forming galaxies.

4.2.1 Sample

The number of observed galaxies has grown over time and continues to do so. The figures below refer to the original survey papers and may no longer be up to date. When referring to the *core sample*, we talk about the 19 galaxies that were observed with all four major telescopes. They lie at distances $5.2 < d/\text{Mpc} < 19.6$, have masses $9.4 < \log(M_\star/M_\odot) < 11.0$ and star formation rates of $-0.56 < \log(\text{SFR}/M_\odot \text{ yr}^{-1}) < 1.23$. The galaxies cover a wide range of the main sequence of star-forming galaxies, though they tend to fall at the more massive end. The sample is shown in Figure 4.2 in front of the $z = 0$ *Multi-wavelength Galaxy Synthesis* (ZOMGS) sample from Leroy et al. ([2019](#)). The footprints of observations of the 19 galaxies with coverage from three facilities are depicted in Figure 4.3. For nearby galaxies, this extensive collection of observations is unique and enables us to connect small scale physical processes to the larger structure of the galaxy.

Of course, just handling such enormous data sets can be challenging. The team's extensive expertise ensures optimal data reduction, delivering finished datasets and saving researchers the time-consuming work of processing the raw data themselves. Furthermore, several derived data products are also made publicly available.

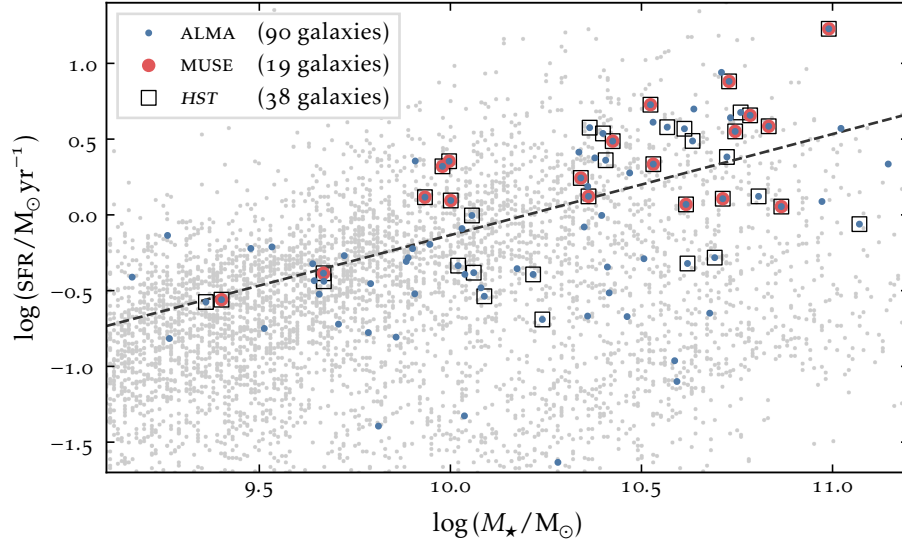


Figure 4.2: PHANGS sample in the mass star-formation plane. The background shows the galaxies in the Z0MGS catalogue by Leroy et al. (2019).

4.2.2 Data sets

Here we go over the different data sets that form the foundations for all further studies. We will proceed chronologically, according to the time when they were observed.

The completion of the observations for the PHANGS–ALMA survey in 2019 marks an important milestone for this collaboration. Including 32 galaxies from archival data, it consists of 90 nearby galaxies in total that form the full PHANGS sample. The technical side of data reduction and processing is described in Leroy et al. (2021a) and the scientific motivation and selection of targets in Leroy et al. (2021b). Since this data is used to study molecular gas, it only marginally touches our work.

Our main workhorse, on the other hand, comes in the form of the PHANGS–MUSE survey. It is a subsample of 19 galaxies that form the core sample (i.e. observations from all four main telescopes are available). The footprints of the observations were designed to match the existing ones from ALMA and cover a large fraction of the galactic disks. Between 3 and 15 pointings were needed per galaxy, which adds up to 168 pointings in total. The full data reduction is described in the survey paper Emsellem et al. (2022) and was performed with the data processing pipeline from Weilbacher et al. (2020). The individual pointings were aligned and mosaicked, whereby H α maps from WFI served as a reference (Razza et al. *in preparation*). Since the individual pointings vary in resolution, all pointings in each galaxy were convolved to single Gaussian PSF. The complex process includes steps, such as removing the sky background and ultimately returns a reduced data cube. These cubes were then processed through the *Data analysis pipeline* (DAP) to extract high-level data products, such as emission line maps and other properties.

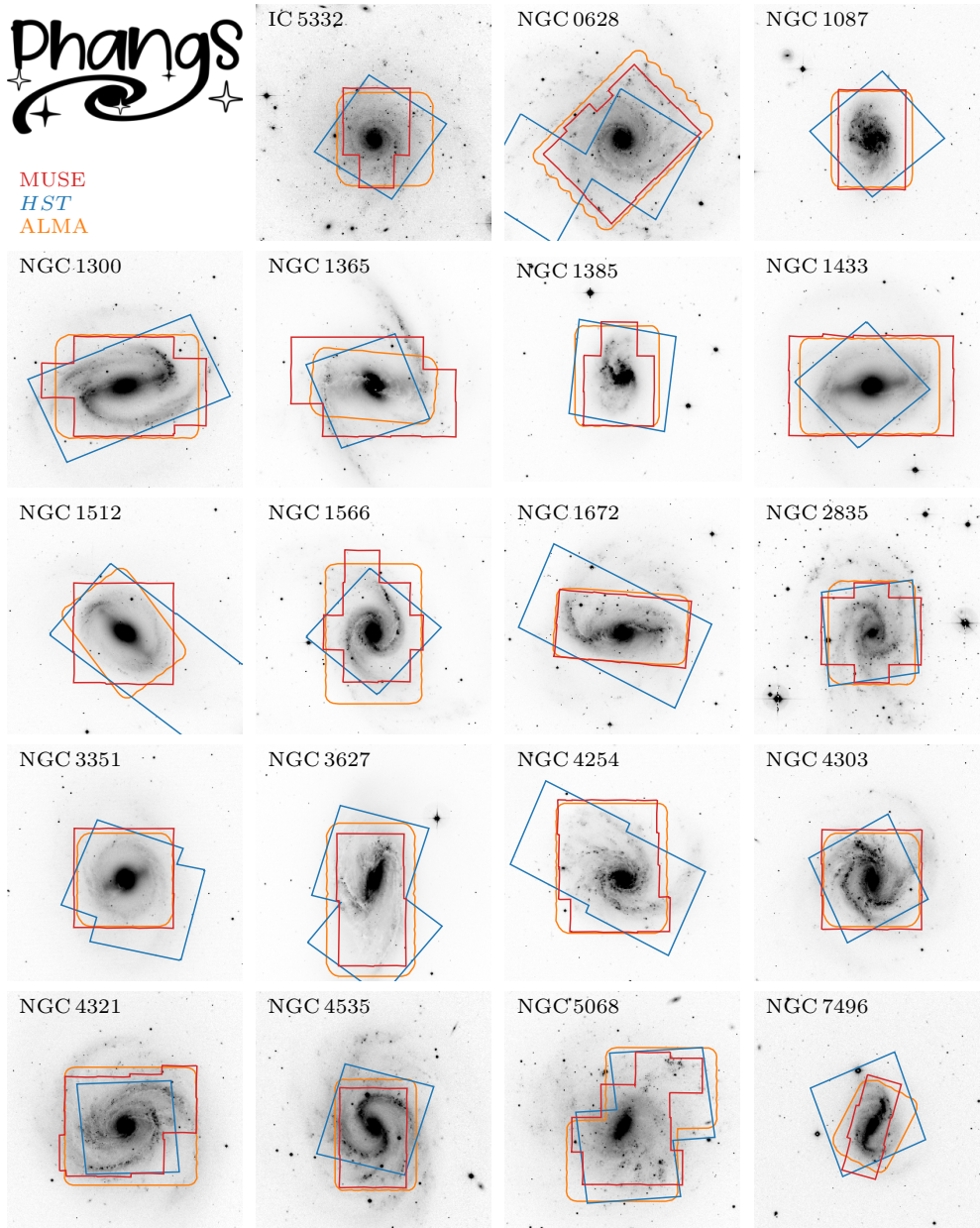


Figure 4.3: PHANGS atlas of the 19 galaxies with a 7×7 arcmin WFI images in the back and the footprints from MUSE in red, *HST* in blue, and ALMA in orange.

This number is about to grow thanks to an approved proposal.

The original PHANGS–HST survey consists of 38 galaxies, where each one was observed with the WFC3 in five bands NUV–U–B–V–I. Some of them were already observed by LEGUS, but only four of them had all required filters. The filters cover the optical spectrum and range from 2400 up to 9500 Å. Parallel imaging was taken with the ACS. For more details, see the survey paper Lee et al. (2022). There is also a PHANGS–HST–H α survey which helps to observe nebulae at an even better resolution than MUSE (Chandar et al. 2025). The galaxies were among the first to be observed with JWST, so we have an additional 8 bands between 2 to 21 μ m. This will help to study young embedded stellar populations and improve our understanding of dust. The survey is described in Lee et al. (2023) and exceeded all expectations and immediately yielded a series of papers with first results.

Not part of the core instruments, but important for our work is the PHANGS–AstroSat survey, described in Hassani et al. (2024). It provides *Near ultraviolet* (NUV) and FUV observations at an resolution of 1.4 arcsec. It includes 31 nearby star-forming galaxies, among them 16 of the 19 MUSE targets in our sample (NGC 1087 and NGC 4303 cannot be observed and NGC 5068 is not yet observed).

4.2.3 Data products

The reduced data sets provide a strong foundation for research, but always starting from scratch is often inefficient. Collaborations enable sharing derived results and catalogues, extending the value of the data. The following overview highlights products relevant to this work.

We start with the essential property on which most studies depend: the distance to the galaxy. Anand et al. (2021) compiled a catalogue of distances, using the *tip of the red giant branch* (TRGB) method from ACS parallel images whenever possible. When unavailable, carefully selected literature values were adopted, including two from Paper I (see Chapter 5). This provides a consistent distance scale for all subsequent work.

Another widely used product is the nebula catalogue by Santoro et al. (2022) and Groves et al. (2023), which provides a key reference for studies of ionized gas, particularly H II regions. A more comprehensive catalogue was compiled by Congiu et al. (2023), while Barnes et al. (submitted) present one based on the PHANGS–HST–H α survey. These resources are complemented by studies deriving physical conditions such as metallicity (Kreckel et al. 2019, 2020; Brazzini et al. 2024), temperatures (Ho et al. 2019; Rickards Vaught et al. 2024), or multiple combined properties (Barnes et al. 2021). Together, they provide valuable insights into the properties of ionized nebulae.

On the stellar side, a series of papers describes the creation of the catalogues for the star clusters by Thilker et al. (2022, 2025) and Maschmann et al. (2024) and stellar associations by Larson et al. (2023). Potential cluster candidates were first identified either with machine-learning methods (Wei et al. 2020; Hannon et al. 2023) or through human classification (Whitmore et al. 2021). Subsequent steps include SED fitting (Turner et al. 2021) and aperture correction (Deger et al. 2022). These techniques continue to be refined, for example to improve age determinations (Whitmore et al. 2023, 2025).

RESEARCH


















»To try to make a model of an atom by studying its spectrum is like trying to make a model of a grand piano by listening to the noise it makes when thrown downstairs. Therefore, making a model of a galaxy by studying its spectrum is like modelling an entire symphony orchestra from the noise it makes when falling downstairs.«

—Kewley et al. (2019)

PNLF DISTANCES FOR 19 GALAXIES OBSERVED BY PHANGS–MUSE

WE PROVIDE new *planetary nebula luminosity function* distances to 19 nearby spiral galaxies that were observed with VLT/MUSE by the PHANGS collaboration. Emission line ratios are used to separate PNe from other bright [OIII] emitting sources like compact SNRs or HII regions. While many studies have used narrowband imaging for this purpose, the detailed spectral line information provided by IFU spectroscopy grants a more robust way of categorising different [OIII] emitters. We investigate the effects of SNR contamination on the *planetary nebula luminosity function* (PNLF) and find that we would fail to classify all objects correctly, when limited to the same data narrowband imaging provides. However, the few misclassified objects usually do not fall on the bright end of the luminosity function, and only in three cases does the distance change by more than 1σ . We find generally good agreement with literature values from other methods. Using metallicity constraints that have also been derived from the same IFU data, we revisit the PNLF zero point calibration. Over a range of $8.34 < 12 + \log(\text{O}/\text{H}) < 8.59$, our sample is consistent with a constant zero point and yields a value of $M^* = -4.542^{+0.103}_{-0.059}$ mag, within 1σ of other literature values. MUSE pushes the limits of PNLF studies and makes galaxies beyond 20 Mpc accessible for this kind of analysis. This approach to the PNLF shows great promise for leveraging existing archival IFU data on nearby galaxies.

CONTRIBUTIONS

This chapter and Appendix A were published in MNRAS as Scheuermann et al. (2022), hereafter referred to as [Paper I](#). The original article was written by Fabian Scheuermann , Kathryn Kreckel , Gagandeep S. Anand , Guillermo A. Blanc , Enrico Congiu , Francesco Santoro , Schuyler D. Van Dyk , Ashley T. Barnes , Frank Bigiel , Simon C. O. Glover , Brent Groves , Ralf S. Klessen , J. M. Diederik Kruijssen , Erik Rosolowsky , Eva Schinnerer , Andreas Schrubba, Elizabeth J. Watkins , and Thomas G. Williams . The idea for this project came from Kathryn Kreckel. I developed the code, produced the figures, and wrote the text. The next six co-authors were part of the paper team and regularly contributed to the work. Before publication, the other co-authors provided additional feedback. This work was conducted in the »Ionized ISM and its Relation to Star Formation« *science working group* (SWG) of PHANGS.

5.1 INTRODUCTION

Observations are the primary way to gather data in astronomy, but not all information is directly accessible. The measurement of some properties, like the intrinsic luminosity or the physical size of galaxies or stars depend on their distances. It is therefore fundamental for our understanding of the universe to know those distances, but measuring them is a delicate task. Cosmic distances span many orders of magnitude, which prohibits the use of a single distance measure and instead necessitates a combination of techniques – the so-called *cosmic distance ladder*. Unfortunately, flaws in the different rungs can propagate into large uncertainties for the most distant objects (Bernal, Verde & Riess 2016; Freedman 2021). It is therefore desirable to find techniques that are accurate and easily applicable to a large number of objects.

For nearby galaxies, the *tip of the red giant branch* (TRGB; e.g. Lee, Freedman & Madore 1993), *Cepheids* (e.g. Freedman et al. 2001), *Type Ia supernovae* (SNe Ia; e.g. Riess, Press & Kirshner 1996), and *surface brightness fluctuations* (SBF; e.g. Tonry et al. 2001) are among the premier techniques used to obtain precise distances, but redundant methods are essential to check for systematic differences. Another method that can achieve similar precision is the PNLF. Ford & Jenner (1978) were among the first to realise that PNe can be utilised to determine the distance to their host galaxy. A PN is the last breath of a dying intermediate mass star (2 to 8 M_{\odot} ; Kwok 2005). During the final stages of its life, the star will expel its outer layers, which will then be ionized by the UV radiation of the exposed hot core. The low densities of the rarefied gas, enables a wealth of collisionally excited emission lines (Osterbrock & Ferland 2006). Even though many stars will go through this phase at the end of their lives, the actual number of PNe that we observe is rather small, because they only last for a few tens of thousands of years (Buzzoni, Arnaboldi & Corradi 2006).

Since the maximum absolute magnitude of all PNe is roughly constant across all galaxies, they can be used as standard candles. Jacoby (1989) and Ciardullo et al. (1989a) formulated an empirical luminosity function that made use of this luminosity cut-off and in doing so, laid the groundwork for how we measure distances with PNe to the present day. Since then, the PNLF has been established as a reliable method for obtaining distances to nearby galaxies up to ~ 20 Mpc, with a precision better than 10 percent (e.g., Ciardullo, Jacoby & Tonry 1993; Ferrarese et al. 2000b; Ciardullo et al. 2002; Feldmeier, Jacoby & Phillips 2007).

The PNLF also plays a special role in the cosmic distance ladder. To the best of our knowledge (though still lacking a good theoretical foundation, Ciardullo 2013), the bright end cutoff of the PNLF is invariant to changes in the local environment or the metallicity and is applicable to both spiral (Feldmeier, Ciardullo & Jacoby 1997) and elliptical galaxies (Ciardullo, Jacoby & Ford 1989b). This makes this method useful for testing and comparing other methods. Beyond their use as a distance indicator, PNe are also used to measure kinematics in the outer halo where measuring stellar kinematics directly is difficult (Hartke et al. 2017).

The greatest challenge in measuring the PNLF distance is to compile a clean catalogue of PNe. Due to their strong $[\text{OIII}]\lambda 5007$ emission (Schönberner et al. 2010), they are easily detectable, but it can be difficult to discriminate them from other nebular objects like HII regions or SNRs. Baldwin et al. (1981) showed that emission line ratios can be used to make this discrimination. To this end, narrowband imaging with a combination of an $[\text{OIII}]$ on-band and a wider off-band filter together with a filter centred around $\text{H}\alpha$, are used to identify PNe.

In recent years, the emergence of powerful IFU spectroscopy instruments has revolutionized the field (Kreckel et al. 2017; Spriggs et al. 2020; Roth et al. 2021; Spriggs et al. 2021). Their FoVs are large enough to detect a sufficient number of PNe in order to measure the PNLF. PNe trace the underlying stellar population and hence we expect many more PNe in the denser central regions. This means that we can potentially find more objects in a smaller FoV. However, the strong stellar continuum in this part of the galaxy also makes it difficult to detect them with narrowband imaging. With the full spectral information from the IFU, we can remove the continuum and search for PNe in regions that were previously inaccessible.

We use data that were observed for PHANGS. This is a collaboration aimed at studying the baryon cycle within galaxies at high spatial resolution, sufficient to isolate and characterise individual molecular clouds and HII regions. One of the pillars of this project are optical IFU spectroscopy observations with MUSE (Bacon et al. 2010) at the VLT. The PHANGS–MUSE sample consists of 19 nearby spiral galaxies that are all roughly face-on, ranging in mass from $\log_{10}(M/M_{\odot}) = 9.4$ to 10.99 and star formation rate from $\log_{10}(\text{SFR}/M_{\odot} \text{ yr}^{-1}) = -0.56$ to 0.88.

This chapter has three objectives. The first aim is to test how susceptible narrowband imaging is to misclassifying SNRs as PNe. Kreckel et al. (2017) found that, in the case of the nearby spiral galaxy NGC 0628, contamination with SNRs can bias the measured distance. Meanwhile, Davis et al. (2018) could not find such problems for M 31 and M 33. The lingering question is whether NGC 0628 was just an anomaly or if this issue compromises other galaxies as well.

The second goal is to quantify the applicability of IFU surveys for PN studies. Past PN studies required special observations that were taken for the sole purpose of finding PNe. Because IFU studies cover the relevant emission lines anyway, this opens up the possibility of measuring the PNLF as a »bonus«. The gain in versatility comes at a price: the FoVs are usually narrower. Here we try to quantify whether IFU surveys can compensate the smaller FoVs with a gain in spectral information and explore the limitations to derive PNLF-based distances from IFU observations.

The last objective is to provide new and precise distance measurements for the 19 nearby galaxies observed by the PHANGS–MUSE survey, some of which did not have good distance estimates before. In Section 5.2, we present the data that is used for this study and describe our process for identifying PNe. In Section 5.3, we fit the observed data to the PNLF and in Section 5.4 and Section 5.5 we present and discuss the results. We conclude in Section 5.6.

5.2 FINDING PLANETARY NEBULAE

We begin by compiling a clean catalogue of PN candidates followed by a second catalogue with potential SNR contaminants that might be misclassified in narrowband studies.

5.2.1 Data

We use the full sample of 19 nearby spiral galaxies (see Table 5.1) that were observed with VLT/MUSE for the PHANGS–MUSE large observing program (PI: Schinnerer; Emsellem et al. 2022). The wide-field mode of MUSE has a FoV of 1×1 arcmin² and only covers part of the targeted galaxies. Between 3 and 15 pointings were taken to sample the bulk of the star-forming disk of each target. The observations cover a wavelength range from 4800 to 9300 Å, with a typical spectral resolution of 2.75 Å. Each pixel covers 0.2×0.2 arcsec² with an average spatial resolution of 0.72 arcsec (67 pc).

The data reduction was performed by the PHANGS–MUSE team, using the MUSE data processing pipeline by Weilbacher et al. (2020), and is described in Emsellem et al. (2022). They produced reduced and mosaicked spectral cubes that form the base for further DAP products which we use for our analysis. This includes emission line maps that are extracted by fitting Gaussian profiles along with the stellar continuum. We use those maps for the H β λ 4861, H α λ 6562, [NII] λ 6583, [SII] λ 6717 and [SII] λ 6731 emission lines. We also use the [OIII] λ 5007 line maps, which are not tied to any of the other lines during the fit, for PN source detection. For our analysis we use the native resolution DAP maps from the public DR1.0.

However, this procedure is not suitable to obtain accurate [OIII] λ 5007 fluxes. [OIII] emission is mostly associated with certain objects like PNe, SNRs and HII regions, and most parts of a galaxy will have little to no emission. By fitting a Gaussian profile in a region without emission, we create a positive bias. To circumvent this issue, we calculate the non-parametric emission-line moment for [OIII] λ 5007 following the procedure employed in the *Mapping Nearby Galaxies at Apache Point Observatory* (MaNGA) data analysis pipeline (Belfiore et al. 2019; Westfall et al. 2019). We use the velocity from the H α line to calculate the expected rest-frame wavelength of the [OIII] line and sum the flux over 17 spectral bins, corresponding to a wavelength range of 4998 to 5018 Å. This range is wide enough to ensure that we always capture the [OIII] line, even in the case that the velocity from the H α is off. In a similar way, the continuum is measured in the range from 4978 to 4998 Å and 5018 to 5038 Å with three sigma clipping and then subtracted from the line flux.

Table 5.1: Properties of the galaxies in the PHANGS–MUSE sample.

Name	Type	R.A. (J2000)	Dec. (J2000)	$(m-M)^a$ mag	i^b deg	PA ^b deg	r_{25} arcmin	$12 + \log(\text{O}/\text{H})^c$	$E(B-V)^d$	FWHM ^e arcsec
IC 5332	SABc	23 ^h 34 ^m 27.49 ^s	−36 ^d 06 ^m 03.89 ^s	29.77 ± 0.10	26.9	74.4	3.03	8.38	0.015	0.72 ^{AO}
NGC 0628	Sc	01 ^h 36 ^m 41.73 ^s	+15 ^d 47 ^m 01.11 ^s	29.96 ± 0.14	8.9	20.7	4.94	8.50	0.062	0.73 ^{AO}
NGC 1087	Sc	02 ^h 46 ^m 25.18 ^s	−00 ^d 29 ^m 55.38 ^s	31.00 ± 0.30	42.9	359.1	1.49	8.37	0.030	0.74 ^{AO}
NGC 1300	Sbc	03 ^h 19 ^m 41.00 ^s	−19 ^d 24 ^m 40.01 ^s	31.39 ± 0.33	31.8	278.0	2.97	8.52	0.026	0.63
NGC 1365	Sb	03 ^h 33 ^m 36.36 ^s	−36 ^d 08 ^m 25.45 ^s	31.46 ± 0.09	55.4	201.1	6.01	8.54	0.018	0.82 ^{AO}
NGC 1385	Sc	03 ^h 37 ^m 28.56 ^s	−24 ^d 30 ^m 04.18 ^s	31.18 ± 0.33	44.0	181.3	1.70	8.40	0.018	0.49
NGC 1433	SBa	03 ^h 42 ^m 01.49 ^s	−47 ^d 13 ^m 18.99 ^s	29.78 ± 0.49	28.6	199.7	3.10	8.55	0.008	0.65
NGC 1512	Sa	04 ^h 03 ^m 54.14 ^s	−43 ^d 20 ^m 55.41 ^s	30.28 ± 0.33	42.5	261.9	4.22	8.56	0.009	0.80 ^{AO}
NGC 1566	SABb	04 ^h 20 ^m 00.38 ^s	−54 ^d 56 ^m 16.84 ^s	31.24 ± 0.25	29.5	214.7	3.61	8.55	0.008	0.64
NGC 1672	Sb	04 ^h 45 ^m 42.49 ^s	−59 ^d 14 ^m 50.13 ^s	31.44 ± 0.33	42.6	134.3	3.08	8.54	0.021	0.72 ^{AO}
NGC 2835	Sc	09 ^h 17 ^m 52.91 ^s	−22 ^d 21 ^m 16.84 ^s	30.44 ± 0.17	41.3	1.0	3.21	8.38	0.089	0.85 ^{AO}
NGC 3351	Sb	10 ^h 43 ^m 57.76 ^s	+11 ^d 42 ^m 13.21 ^s	29.99 ± 0.07	45.1	193.2	3.61	8.59	0.024	0.74 ^{AO}
NGC 3627	Sb	11 ^h 20 ^m 15.00 ^s	+12 ^d 59 ^m 29.40 ^s	30.27 ± 0.09	57.3	173.1	5.14	8.55	0.037	0.77 ^{AO}
NGC 4254	Sc	12 ^h 18 ^m 49.63 ^s	+14 ^d 24 ^m 59.08 ^s	30.59 ± 0.33	34.4	68.1	2.52	8.53	0.035	0.58
NGC 4303	Sbc	12 ^h 21 ^m 54.93 ^s	+04 ^d 28 ^m 25.48 ^s	31.15 ± 0.39	23.5	312.4	3.44	8.56	0.020	0.58
NGC 4321	SABb	12 ^h 22 ^m 54.93 ^s	+15 ^d 49 ^m 20.29 ^s	30.91 ± 0.07	38.5	156.2	3.05	8.56	0.023	0.64
NGC 4535	Sc	12 ^h 34 ^m 20.30 ^s	+08 ^d 11 ^m 52.70 ^s	30.99 ± 0.05	44.7	179.7	4.07	8.55	0.017	0.44
NGC 5068	Sc	13 ^h 18 ^m 54.74 ^s	−21 ^d 02 ^m 19.48 ^s	28.58 ± 0.09	35.7	342.4	3.74	8.34	0.091	0.73 ^{AO}
NGC 7496	Sb	23 ^h 09 ^m 47.29 ^s	−43 ^d 25 ^m 40.26 ^s	31.36 ± 0.33	35.9	193.7	1.67	8.51	0.008	0.79

Notes. Adopted from the PHANGS sample table (v1p6, see Leroy et al. 2021b).

^a Curated by Anand et al. (2021). For NGC 1433 and NGC 1512 the distances are from Sabbi et al. (2018).

^b Inclination and position angle from Lang et al. (2020).

^c At the mean positions of the PNe with gradients from Kreckel et al. (2020) and Santoro et al. (2022).

^d Galactic foreground extinction from Schlafly & Finkbeiner (2011).

^e Average FWHM of the PSF across all pointings.

Galaxies that were observed with adaptive optics are marked with AO.

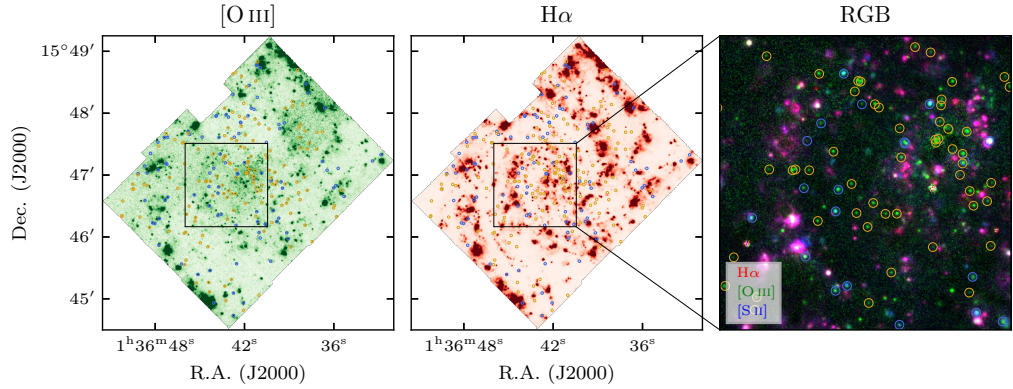


Figure 5.1: Mapping the location of emission lines sources in NGC 628. PNe (marked with gold circles) appear as bright point like objects in the [O III] map (left). Objects that also show significant H α emission (centre) are often other objects like H II regions or SNRs (marked with blue circles). In the composite RGB image (right), PNe stand out as bright green dots.

In order to measure the flux of a point source like a PN, one needs to know the shape of the PSF. For our MUSE data, it is best parameterised by a Moffat profile (Bacon et al. 2017; Fusco et al. 2020)

$$p(r) = A \left[1 + \left(\frac{r^2}{\gamma^2} \right) \right]^{-\alpha}. \quad (5.1)$$

A fixed power index of $\alpha = 2.3$ is used for galaxies with AO observations and $\alpha = 2.8$ for the remaining galaxies (both values were provided by ESO). The core width γ varies from pointing to pointing and is related to the full width half maximum via $\text{FWHM} = 2\gamma\sqrt{2^{1/\alpha} - 1}$. The latter is usually measured from bright point-like objects. However since many pointings in our data do not have such objects, we use a different approach where we measure the *full width half maximum* (FWHM) with the help of a reference image with a known PSF (Emsellem et al. 2022). Unfortunately the derived FWHM can have a rather large uncertainties of up to 0.1 arcsec. The FWHM is measured at a reference wavelength of 6483.58 Å and it decreases linearly with increasing wavelength. This can be approximated with

$$\text{FWHM}(\lambda) = \text{FWHM}_{\text{ref}} - 3 \cdot 10^{-5} \text{ arcsec } \text{\AA}^{-1} (\lambda - 6483.58 \text{ \AA}), \quad (5.2)$$

where FWHM_{ref} is measured at the reference wavelength.

5.2.2 Source detection and photometry

To create an initial catalogue of PNe candidates, we use a four step procedure: First, we search for unresolved objects in the Gaussian fit [O III] line map. Then, we measure their fluxes within the non-parametric moment 0 maps with aperture photometry and subtract

the background from an annulus. Next, we correct for the flux that is lost outside of the aperture and finally apply a correction for Milky Way extinction.

The physical resolution of the observations varies between 25 and 95 pc (based on prior distances from Anand et al. 2021). This means that PNe, which typically have sizes of less than 1 pc (Bojić et al. 2021), will always appear as unresolved objects. For the bright PNe that are of interest to us, the strongest emission line is the forbidden [OIII] $\lambda 5007$ line, so we search for point sources in this line map (see Figure 5.1). Before we start with the source finding, we mask out regions around known stars from *Gaia* DR2 (Gaia Collaboration et al. 2018) to account for foreground objects. Other problematic regions are the centres of galaxies with strong [OIII] emission. This can lead to other objects being mistakenly classified as PNe. To avoid this, we create an elliptical mask and place it at the galaxy centre. We select a semi-major-axis of $0.2 \cdot r_{25}$. This allows us to reliably get rid of most problematic objects. We also mask the parts of the image that have H α fluxes above the 95th percentile of the entire image. This removes some parts of the spiral arms. This problem only concerns some of the galaxies in our sample and we only apply this treatment to the following galaxies: NGC 1300, NGC 1365, NGC 1512, NGC 1566, NGC 1672, NGC 3351, NGC 3627, NGC 4303, NGC 4321, NGC 4535 and NGC 7496. Between 6.5 to 13.6 percent of the FoVs are masked in the process.

The source detection itself is performed with the Python package PHOTUTILS (Bradley et al. 2019) which provides an implementation of the DAOPHOT algorithm from Stetson (1987). Because individual pointings differ slightly in noise and PSF size, we process them individually. For each of them, the noise level is estimated from the standard deviation with iterative three sigma clipping. We then search for objects above a threshold of 3 times this noise level. This constitutes our preliminary catalogue of PN candidates.

The next step is to measure the fluxes, which is done with aperture photometry. A smaller aperture is generally favourable because it only uses the brightest part of the object and hence increases the signal to noise (Howell 1989). However, this requires a good knowledge of the PSF. Because we are not able to get a precise measurement of the PSF in each pointing (see Emsellem et al. 2022, for a detailed discussion of the PSF), those errors would then propagate into the measured fluxes. We therefore use a rather large aperture with a diameter of $2.5 \cdot \text{FWHM}$. Only with this large aperture size does the uncertainty due to the PSF reach a level similar to that of the photometric error (≈ 0.1 mag).

The background is estimated from an annulus, with the inner radius set to $4 \cdot \text{FWHM}$ and the outer radius chosen such that the covered area is five times the area of the aperture, corresponding to $6.9 \cdot \text{FWHM}$. To remove bright contaminants within the annulus, we apply iterative three sigma clipping. The sigma-clipped median of the annulus is then scaled to the size of the central aperture and subtracted. Afterwards, we correct for flux that is lost outside of the aperture. To do this we integrate the PSF from Equation 5.1. This yields the fraction of flux within an aperture of radius r ,

$$f(r) = 1 - \left[1 + \left(\frac{r^2}{\gamma^2} \right) \right]^{1-\alpha}. \quad (5.3)$$

This correction is then applied to all measured fluxes. Compared to the noise in the image, the associated error maps seem to be underestimated. This is confirmed by looking at lines with a fixed theoretical ratio, e.g., $[\text{OIII}]\lambda 4958/[\text{OIII}]\lambda 5007 = 0.35$, $[\text{OI}]\lambda 6363/[\text{OI}]\lambda 6300 = 0.33$ or $[\text{NII}]\lambda 6548/[\text{NII}]\lambda 6583 = 0.34$. By fitting a pair of lines independent from each other and comparing the deviation from the theoretical ratio with the measured uncertainties, we conclude that the uncertainties are underestimated by a factor of 1.38 to 1.87 (with larger deviations at shorter wavelengths). To compensate for that, we increase the measured uncertainties by a factor of 1.67, reflecting the mean correction factor from our tests. The average uncertainty of the $[\text{OIII}]$ magnitudes after this correction is 0.078 mag. To this we add in quadrature the uncertainty that arises from the error of the PSF measurement (estimated to be 0.1 mag on average). This makes up the final value that is reported in Table 5.2 and used throughout the analysis.

Next, all measured fluxes are corrected for Galactic foreground extinction. For the Milky Way foreground we use the Cardelli, Clayton & Mathis (1989) extinction curve with $R_V = 3.1$ and $E(B-V)$ from Schlafly & Finkbeiner (2011, see Table 5.1). We choose not to correct for internal extinction. Neither for circumstellar extinction that is associated with the PN, nor for the dust within the target galaxy. The former is part of the empirical calibration of the luminosity function, and so far only a few studies investigated it (Davis et al. 2018). For the latter it is common practice not to correct for it, based on a number of arguments (Feldmeier et al. 1997). In the case of the Milky Way, PNe have a higher vertical scale height (Bobilev & Bajkova 2017) than dust (Li et al. 2018). If PNe are equally distributed in front of and behind the dust, we will observe about half of the sample with minimal extinction. As long as this subset sufficiently samples the bright end of the PNLF, we will measure the correct distance. Strictly speaking, this implies that we should observe a different shape, composed of two shifted luminosity functions. In Appendix A.2 we show that our algorithm is able to recover the correct distance even when we have a compound luminosity function, consisting of a sample with extinction and one without extinction. On the other hand, even a few bright PNe that are wrongly extinction corrected can significantly reduce the measured distances.

Finally, the $[\text{OIII}]$ fluxes (in $\text{erg s}^{-1} \text{cm}^{-2}$) are converted to apparent magnitudes with the formula by Jacoby (1989)

$$m_{[\text{OIII}]} = -2.5 \cdot \log_{10} I_{[\text{OIII}]} - 13.74. \quad (5.4)$$

5.2.3 Comparison with existing studies

For a handful of galaxies in our sample, the PNLF has been measured in previous studies. This includes NGC 0628 (Herrmann et al. 2008; Kreckel et al. 2017; Roth et al. 2021), NGC 3351 and NGC 3627 (both Ciardullo et al. 2002) and NGC 5068 (Herrmann et al. 2008). To validate our source detection and photometry, we compare the positions and measured fluxes of our PNe candidates when available.

For NGC 0628, we recover around 75 percent of the sources (including the 20 brightest objects). In Figure 5.2, we compare our measured fluxes. While we find good agreement for the $[\text{OIII}]$ magnitudes, with only minor offsets of $\Delta m_{[\text{OIII}]} = 0.03 \pm 0.23$ mag (Herrmann

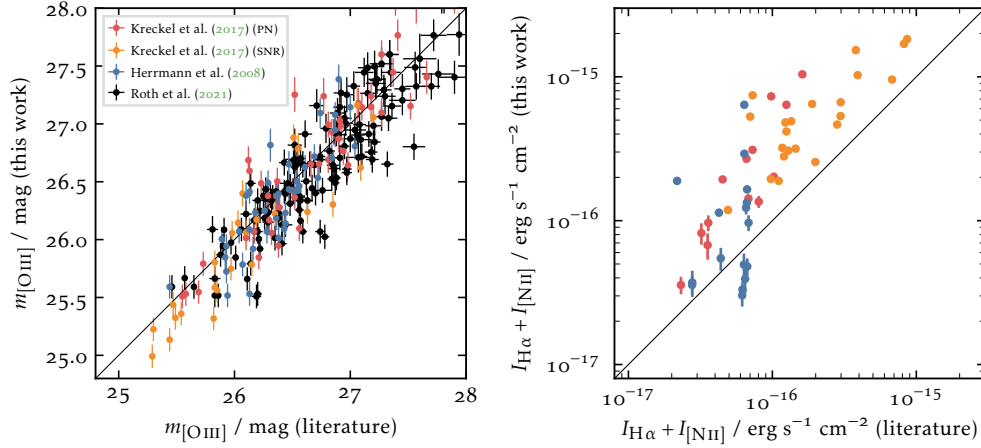


Figure 5.2: *Left*: Comparison of the [OIII] apparent magnitudes in NGC 0628 between our detections and the catalogues from Herrmann et al. (2008), Kreckel et al. (2017), and Roth et al. (2021). The solid line represents equality. *Right*: Comparison of the $H\alpha + [NII]$ fluxes in NGC 0628. Narrowband imaging does not resolve the $H\alpha$ and [NII] lines individually and instead measures a wider part of the spectrum. While not identical, we try to replicate this behaviour by summing the $H\alpha$ and the [NII] emission line. The solid line represents equality.

et al. 2008), $\Delta m_{[OIII]} = 0.01 \pm 0.29 \text{ mag}$ (Kreckel et al. 2017) and $\Delta m_{[OIII]} = 0.08 \pm 0.26 \text{ mag}$ (Roth et al. 2021), the $H\alpha$ fluxes are at odds. Figure 5.2 suggests systematically higher $H\alpha$ fluxes compared to Kreckel et al. (2017). This discrepancy with Kreckel et al. (2017), who used the same data as this work, can be attributed to a different technique for the background subtraction. Kreckel et al. (2017) smoothed the image to 2 arcsec to create a global background map, which is then subtracted from the original image. In our tests, we find that this results in a higher background compared to our annulus background subtraction. The ensuing larger $H\alpha$ fluxes in our analysis can alter the subsequent classification in favour of HII regions.

For the other galaxies, the overlap is sparse. Narrowband studies have difficulties in removing the stellar continuum and therefore focus on the outer regions to find PNe while our observations do not extend far beyond the central region. Therefore, only a handful of objects appear in both samples. NGC 3351 has six and NGC 3627 has one previously detected PNe in our FoV, and we are able to recover them all. Except for the faintest objects, the [OIII] fluxes agree within 0.2 mag, but no $H\alpha$ fluxes are reported. For NGC 5068, two of their PNe fall in our FoV but we are only able to detect one of them.

5.2.4 Emission line diagnostics

The galaxies in the PHANGS-MUSE sample are, by selection, star-forming. This means that the detected [OIII] emission is not exclusively coming from PNe, but can also originate from HII regions or compact SNRs. To eliminate them from our catalogue of PN

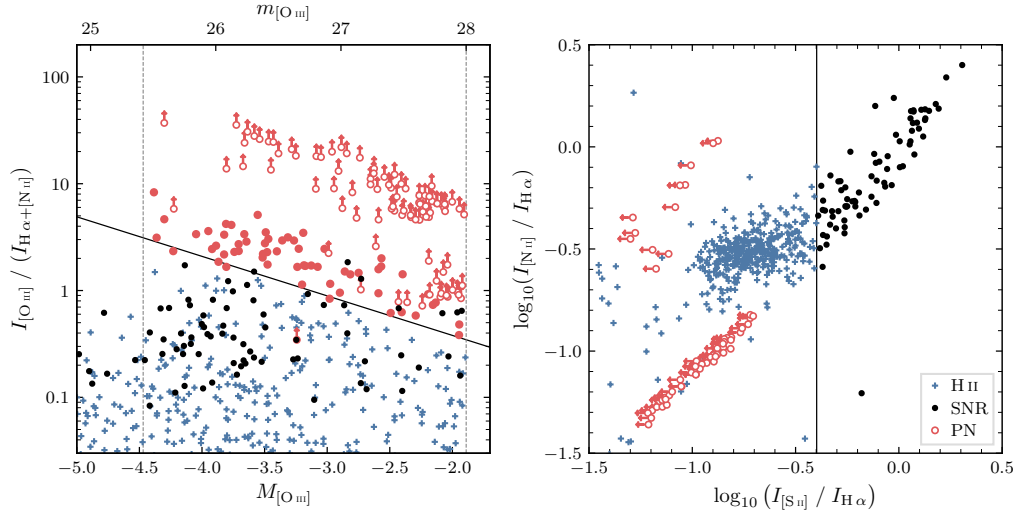


Figure 5.3: Emission line diagnostics for NGC 0628. The left panel shows the criterion from Equation 5.5 that is used to eliminate H II regions from our sample. Open circles indicate when a PN is not detected in H α and the symbol is a lower limit. The right panel shows the criterion from Equation 5.6 that is used to remove SNRs. Open circles indicate that a PN is not detected in [S II] and the symbol is an upper limit.

candidates, we apply a set of *emission line diagnostics*. The origin of the ionizing radiation is different for the three classes of objects (white dwarfs for PNe, O and B stars for H II regions and shocks for SNRs). This yields different characteristics in their spectra which can be used to discriminate them. The star at the centre of a bright PN is much hotter than the star(s) that ionize an H II region. It has a harder spectrum where more doubly-ionized oxygen is produced compared to H α (Shaver et al. 1983) and the ratio [O III]/(H α + [N II]) should therefore be greater than 1.6 (H α and [N II] are taken together because the lines are close and narrow band imaging can not resolve the individual lines). This effect is even more pronounced for luminous PNe which usually show even higher ratios. Therefore, instead of drawing a straight horizontal line, we use the criterion from Ciardullo et al. (2002) who showed that PNe typically fall above a line defined by

$$\log_{10} \frac{I_{[\text{O III}]\lambda 5007}}{I_{\text{H}\alpha + [\text{N II}]\lambda 6583}} > -0.37M_{[\text{O III}]} - 1.16. \quad (5.5)$$

This still leaves possible contamination from compact SNRs. To eliminate them, we use a criterion from Blair & Long (2004). The shock-heated material of SNRs will have higher [S II]/H α ratios with

$$\log_{10} \frac{I_{[\text{S II}]\lambda 6717 + [\text{S II}]\lambda 6731}}{I_{\text{H}\alpha}} > -0.4. \quad (5.6)$$

This cut can also remove some H II regions, but it is unlikely that it removes PNe. The classification based on both criteria is shown in Figure 5.3. Many of our PN candidates

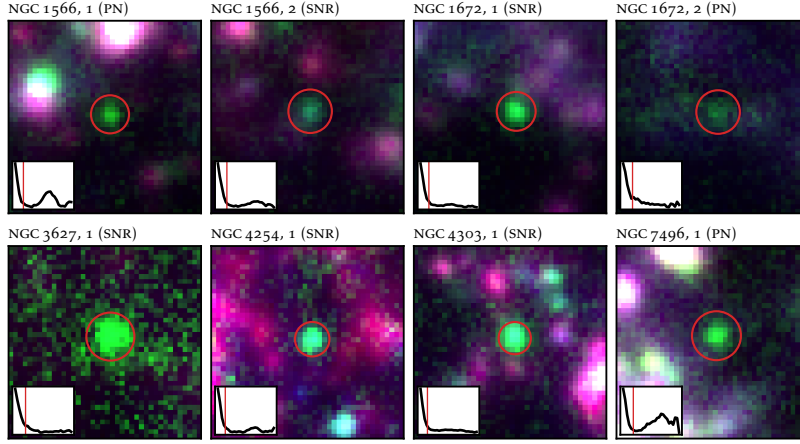


Figure 5.4: In total eight objects were marked and rejected as overluminous objects. Three of them are PNe and five are SNRs. The cutouts show composite RGB images of those objects with $H\alpha$ in red, $[OIII]$ in green and $[SII]$ in blue. The bottom left corner shows the azimuthally averaged radial profile measured from the $[OIII]$ line map. The title contains the name of the galaxy, the »id« in the catalogue (Table 5.2) and the type of the object.

are detected neither in $H\alpha$ nor in $[SII]$ (we call everything below 3σ undetected). Objects that are consistent within the uncertainty of being a PN are retained in the sample.

A last criterion is given by the velocity dispersion of the emission lines. The shell of a PN expands slowly ($\sim 20 \text{ km s}^{-1}$) into the ambient medium (Schönberner et al. 2005) whereas SNRs have much higher expansion velocities ($\sim 100 \text{ km s}^{-1}$, Franchetti et al. 2012). We measure the velocity dispersion from the $[OIII]$ line, averaged over the central pixels of the object, and deconvolve it from the instrumental resolution. We note that, while many objects have quite high signal to noise, the instrumental spectral resolution is much larger ($\sim 50 \text{ km s}^{-1}$) than the intrinsic dispersion and so our deconvolution is sensitive to the shape of the assumed line spread function. Because of this, we are confident in relative velocity dispersions, but have reservations about their absolute values. We therefore decide not to use this criterion for the classification, but only as an additional check. Indeed, SNRs (61 km s^{-1}) have a larger velocity dispersion than PNe (48 km s^{-1}) or H II regions (30 km s^{-1}).

5.2.5 Overluminous sources

The resulting PN catalogue is then visually inspected to exclude any non-point-like objects that slipped through the previous selection process. Most objects that are rejected are removed because their azimuthally averaged radial profile (average flux in annuli of increasing radii, from now on referred to as radial profile) exposes them to be not point-like or they fall very close to another bright object, hampering our ability to measure the flux, but we also have to remove a few objects purely based on their luminosity.

These »overluminous« objects are distinctly brighter than the bright end of the remaining luminosity function (see Section 5.4, Figure 5.6), and hence, cannot be described by the PNLF. They have been found in many PN studies (Longobardi et al. 2013; Hartke et al. 2017; Roth et al. 2021), but their exact origin remains unclear (Jacoby, Ciardullo & Harris 1996). Possible explanations include two PNe that accidentally fall close to each other, a misclassified H II region or a background Ly- α galaxy. If such an object falls close to the bright end of the luminosity function, there is no clear way to eliminate it from the sample and it can lead us to underestimate the distances. However, if a gap exists to the remaining sample, it is easy to exclude such an object because the quality of the resulting fit would be very poor. In total, eight objects are rejected as being overluminous. RGB composites of all eight objects are shown in Figure 5.4. Three of them are PNe and, therefore, also appear in Figure 5.6, while the other five are SNRs that could be classified as PNe.

5.2.6 Final PN sample

Based on these detection and selection criteria, we classify 899 objects as PNe across the 19 PHANGS–MUSE galaxies. A further 150 objects are classified as SNRs, which could contaminate narrowband PNe studies, i.e. they satisfy the criterion from Equation 5.5 but not the one from Equation 5.6. The final catalogue is presented in Table 5.2. The table contains the PNe that are brighter than the completeness limit, and are used in our fit to the PNLF. The position is reported, together with the apparent [O III] magnitude and the line ratios used to make the classification. We also tabulate the SNRs that would be classified as PNe without the [S II] criterion. Special objects like overluminous PNe are flagged with »OL«.

Table 5.2: Planetary nebula identifications and supernova remnant contaminants.

Galaxy	ID	Type	R.A. (J2000)	Dec. (J2000)	$m_{[\text{OIII}]}$ mag	$\log(I_{[\text{OIII}]} / I_{\text{H}\alpha})$	$\log(I_{[\text{NII}]} / I_{\text{H}\alpha})$	$\log(I_{[\text{SII}]} / I_{\text{H}\alpha})$	Notes
NGC 0628	1	PN	01 ^h 36 ^m 43.31 ^s	+15 ^d 47 ^m 18.15 ^s	25.51 ± 0.1	1.1 ± 0.07	-0.83 ± 0.44	-0.71 ± 0.44	
NGC 0628	2	PN	01 ^h 36 ^m 43.35 ^s	+15 ^d 47 ^m 18.35 ^s	25.53 ± 0.1	0.67 ± 0.03	-1.24 ± 0.44	-1.13 ± 0.44	
NGC 0628	3	PN	01 ^h 36 ^m 41.29 ^s	+15 ^d 47 ^m 04.41 ^s	25.59 ± 0.1	1.75 ± 0.43	-0.03 ± 0.61	0.05 ± 0.61	
NGC 0628	4	PN	01 ^h 36 ^m 39.96 ^s	+15 ^d 47 ^m 02.64 ^s	25.59 ± 0.1	0.84 ± 0.04	-1.14 ± 0.44	-1.06 ± 0.44	
NGC 0628	5	PN	01 ^h 36 ^m 41.37 ^s	+15 ^d 46 ^m 58.31 ^s	25.66 ± 0.1	0.54 ± 0.03	-1.24 ± 0.44	-1.15 ± 0.44	
NGC 0628	6	PN	01 ^h 36 ^m 41.37 ^s	+15 ^d 46 ^m 55.67 ^s	25.67 ± 0.1	1.71 ± 0.43	0.9 ± 0.44	0.06 ± 0.61	
NGC 0628	7	SNR	01 ^h 36 ^m 42.48 ^s	+15 ^d 47 ^m 01.48 ^s	25.76 ± 0.1	1.83 ± 0.43	1.58 ± 0.43	1.35 ± 0.43	
NGC 0628	8	PN	01 ^h 36 ^m 43.35 ^s	+15 ^d 48 ^m 04.68 ^s	25.85 ± 0.1	0.7 ± 0.04	-1.12 ± 0.44	-0.98 ± 0.44	
NGC 0628	9	PN	01 ^h 36 ^m 35.82 ^s	+15 ^d 46 ^m 19.60 ^s	25.94 ± 0.1	0.65 ± 0.03	-0.6 ± 0.09	-1.18 ± 0.44	
NGC 0628	10	PN	01 ^h 36 ^m 37.95 ^s	+15 ^d 46 ^m 11.46 ^s	25.97 ± 0.1	0.49 ± 0.02	-0.58 ± 0.06	-0.53 ± 0.07	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
NGC 0628	145	PN	01 ^h 36 ^m 38.39 ^s	+15 ^d 48 ^m 24.20 ^s	27.95 ± 0.20	-0.24 ± 0.08	-1.27 ± 0.44	-1.14 ± 0.44	
NGC 0628	146	PN	01 ^h 36 ^m 36.17 ^s	+15 ^d 47 ^m 06.59 ^s	27.95 ± 0.19	-0.14 ± 0.07	-1.2 ± 0.44	-1.1 ± 0.44	
NGC 0628	147	PN	01 ^h 36 ^m 42.34 ^s	+15 ^d 47 ^m 56.95 ^s	27.95 ± 0.21	0.94 ± 0.44	-0.06 ± 0.61	0.08 ± 0.61	
NGC 0628	148	SNR	01 ^h 36 ^m 40.53 ^s	+15 ^d 47 ^m 31.74 ^s	27.97 ± 0.17	0.14 ± 0.07	0.06 ± 0.06	-0.01 ± 0.07	
NGC 0628	149	PN	01 ^h 36 ^m 42.53 ^s	+15 ^d 47 ^m 33.41 ^s	27.99 ± 0.21	0.89 ± 0.44	-0.05 ± 0.61	0.06 ± 0.61	

Note. OL: rejected as overluminous sources.

5.3 THE PLANETARY NEBULA LUMINOSITY FUNCTION

The *planetary nebula luminosity function* is an empirical relation that describes the number of PNe that we expect to observe at a certain luminosity. Ciardullo et al. (1989a) combined an exponential function from theoretical evolutionary models with a cutoff at the bright end for the following formula:

$$N(M_{[\text{OIII}]}) \propto e^{0.307 M_{[\text{OIII}]}} (1 - e^{3(M^* - M_{[\text{OIII}]})}), \quad (5.7)$$

where M^* is the zero point of the luminosity function. This simple form of the luminosity function has prevailed even though other parametrisations have been proposed. Hartke et al. (2017), for example, treated the slope of the faint end (0.307 in the original model) as a free parameter and it is even possible to construct a completely numerical luminosity function (e.g., Méndez et al. 2001; Teodorescu et al. 2011).

The zero point of the luminosity function, M^* , determined by the luminosity of the brightest PN, has more weight in determining the distance than the functional form itself. This value has to be calibrated from PNe in galaxies with known distances. The original value of $M^* = -4.48$ mag was derived based on the Cepheid distance to M 31 by Ciardullo et al. (1989a). A later study by Ciardullo et al. (2002) used galaxies with distances from the *HST* Key Project (Freedman et al. 2001) to quantify the impact of metallicity on the zero point. They found a modest dependence with an increase at lower metallicities. The metallicity of the galaxies in our sample varies from $12 + \log(\text{O}/\text{H}) = 8.41$ to 8.63 in the centre and $12 + \log(\text{O}/\text{H}) = 8.00$ to 8.62 at r_{25} (Kreckel et al. 2019; Santoro et al. 2022). Since the expected change in M^* is rather small, we decide to adopt a constant value of $M^* = -4.47$ mag as a starting point. In Section 5.5.2, we use our measured distances and compare them to existing distances from the TRGB to review this relation.

5.3.1 Completeness limit

Before we fit our data to the luminosity function, we need to estimate the completeness of our sample. Due to noise in the [OIII] image and blending with the background, we will miss some of the fainter objects and hence, our luminosity function will not be fully sampled at lower luminosities. To estimate the luminosity where this becomes problematic, we inject mock sources, sampled from a theoretical luminosity function, into the [OIII] map and record how many we are able to recover. We never recover 100 percent because the mock sources are randomly placed in the image and sometimes fall onto crowded areas where they are not detectable. Figure 5.5 shows the fraction of recovered objects as a function of their [OIII] magnitude.

One could use this function and »multiply« it with the luminosity function to account for the sources that were missed. Instead, we use an easier approach and fit only the subset of objects that are brighter than a certain threshold. We find that for most galaxies we are able to recover at least 80 percent of the objects brighter than a limiting magnitude of 28 mag. Including or excluding a few objects at the faint end does not alter the measured distance and we therefore adopt a uniform completeness limit of $m_{[\text{OIII}]} = 28$ mag. There are two exceptions, NGC 2835 and NGC 3627. For these two

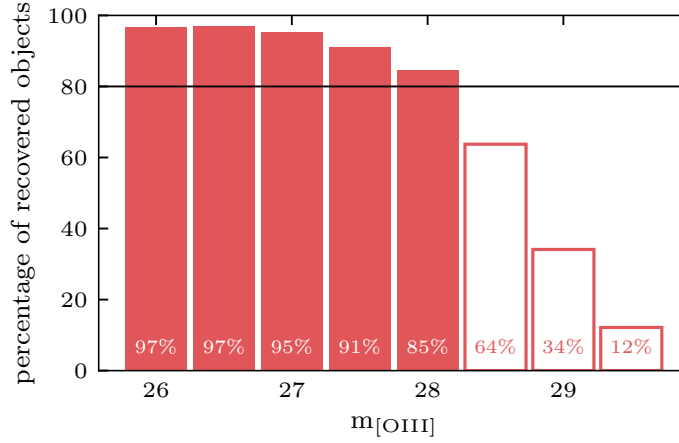


Figure 5.5: Determining the completeness limit for NGC o628. We assume our sample to be complete when at least 80 per cent of the sources are recovered ($m_{[\text{OIII}]} = 28$ mag in this case).

galaxies, the 80 per cent completeness required by us is only recovered for magnitudes brighter than $m_{[\text{OIII}]} = 27.5$ mag.

5.3.2 Fitting the data

We can write the absolute magnitude in terms of the observed apparent magnitude and the distance modulus $\mu = (m - M)$ and fit Equation 5.7 to our observed data. We normalise the function to the number of observed PNe from the root ($= M^* + \mu$) up to the previously determined completeness limit. The fitting itself is done via the method of *maximum likelihood estimation* (MLE, see e.g. Hogg, Bovy & Lang 2010 for a detailed introduction). This alleviates the uncertainties that are introduced when binning the data for a least χ^2 fit. For the MLE, we define the likelihood of a parameter, given some data, as the product of the individual probabilities (or in this case the sum of the log likelihood)

$$\mathcal{L}(\mu) = \sum_i \log p(m_i | \mu). \quad (5.8)$$

The maximum of this function yields the parameters that best describe our observed data. To account for uncertainties in the observed magnitudes, we calculate the probability of observing a PNe of apparent magnitude $m_{[\text{OIII}]}$ as the integral of the luminosity function multiplied by a Gaussian centred around the magnitude of the PN with the uncertainty of the apparent magnitude σ_m as the width

$$p(m_{[\text{OIII}]}) = \int N(m') \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left[-\frac{(m' - m_{[\text{OIII}]})^2}{2\sigma_m^2}\right] dm'. \quad (5.9)$$

Strictly speaking, the uncertainties of the magnitudes are not Gaussian (assuming that the uncertainties of the fluxes are). However for sufficiently high signal to noise (> 10), the

distribution approaches a normal distribution, which is the case for most of our objects. The effect of this procedure is only marginal, but it can take away some weight from the brightest PNe that usually dominate the outcome of the fit. The uncertainty of the fit is taken from the 68 per cent confidence interval of the likelihood function. In Appendix A.3 we look in more detail at the fitting procedure and the reliability of the PNLF.

The classification criterion in Equation 5.5 depends on a prior estimate of the distance, for which we use the distances from Table 5.1 (Anand et al. 2021) as a starting point. If our measured distance deviates significantly from this value, the classification of some objects might change. To account for this, we use an iterative process where we classify and fit repeatedly until the classification and the measured distance does not change anymore.

5.4 RESULTS

In this section, we present an overview of the measured distances and discuss the quality of the individual fits. For each galaxy, we first fit the PNLF to the clean PNe sample and adopt this value as our preferred distance. We then also measure the distance from the catalogue with the potential SNR contaminants to quantify how misclassified SNRs would impact the result. Both values are presented in Table 5.3. Figure 5.6 shows the PNLF for all galaxies in our sample and Figure 5.7 shows their cumulative luminosity function. We use a *Kolmogorov–Smirnov test* to compare the observed sample to a theoretical luminosity function. A high test statistic, D_{\max} (the maximum distance between the empirical distribution function and the cumulative distribution function of the underlying model), and a low p -value rejects the null hypothesis that the observed data were drawn from the theoretical distribution. The quality of the PNLFs varies greatly in our sample. While the closer galaxies are well sampled, the quality of the fit deteriorates with distance. Beyond 15 Mpc, it becomes increasingly difficult to detect a sufficient number of PNe. This corresponds to a distance modulus of $(m - M) = 30.9$ mag, which puts the bright end cutoff at $m_{[\text{O III}]} = 26.4$ mag. With a completeness limit of 28 mag, this means we are only able to observe the upper 1.6 mag of the luminosity function. The reported values for those galaxies should therefore be treated with caution. Figure 5.8 shows a comparison between our measured distances with a compilation of literature distance for NGC 0628. Similar plots for the full sample are shown in Appendix A.1.

IC 5332 is located in the constellation Sculptor. Until recently, the only available distances came from the Tully–Fisher method and they show a large scatter, with the newest reported distance modulus of 29.62 ± 0.40 mag (Tully & Fisher 1988) being significantly larger than the old ones (see Figure A.1). A new TRGB distance from the PHANGS–*HST* program (Anand et al. 2021) finds a distance modulus of 29.77 ± 0.09 mag. This places this galaxy among the closest ones in our sample. Our sample consists of 49 PNe that are brighter than our completeness limit. Two objects were excluded because of their irregular radial profile, but neither of them fall at the bright end, and hence, this does not impact the measured distance. From the remaining 47 PNe we measure a distance modulus of $(m - M) = 29.73^{+0.10}_{-0.20}$ mag ($D = 8.84^{+0.39}_{-0.82}$ Mpc). This is in excellent agreement

Table 5.3: Measured PNLF distances. N_{PN} is the number of detected PNe brighter than our completeness limit and N_{SNR} is the number of SNRs that could be misclassified as PNe. $(m - M)$ and $(m - M)_{\text{SNR}}$ are the distance moduli measured without and with SNR contamination.

Name	N_{PN}	N_{SNR}	$(m - M)$ mag	$(m - M)_{\text{SNR}}$ mag	Distance Mpc
IC 5332	47	16	$29.73^{+0.10}_{-0.20}$	$29.78^{+0.09}_{-0.16}$	$8.84^{+0.39}_{-0.82}$
NGC 0628	139	10	$29.89^{+0.06}_{-0.09}$	$29.90^{+0.06}_{-0.09}$	$9.52^{+0.26}_{-0.41}$
NGC 1087	15	5	$31.05^{+0.10}_{-0.24}$	$31.06^{+0.09}_{-0.20}$	$16.25^{+0.74}_{-1.79}$
NGC 1300	17	3	$32.06^{+0.08}_{-0.12}$	$32.03^{+0.07}_{-0.11}$	$25.77^{+0.90}_{-1.42}$
NGC 1365	29	5	$31.22^{+0.08}_{-0.14}$	$31.19^{+0.07}_{-0.13}$	$17.53^{+0.66}_{-1.16}$
NGC 1385	11	9	$29.96^{+0.14}_{-0.32}$	$30.04^{+0.12}_{-0.28}$	$9.81^{+0.63}_{-1.46}$
NGC 1433	90	6	$31.39^{+0.04}_{-0.07}$	$31.39^{+0.04}_{-0.06}$	$18.94^{+0.39}_{-0.56}$
NGC 1512	43	5	$31.27^{+0.07}_{-0.11}$	$31.29^{+0.06}_{-0.10}$	$17.93^{+0.53}_{-0.88}$
NGC 1566	27	2	$31.13^{+0.08}_{-0.17}$	$31.14^{+0.08}_{-0.16}$	$16.84^{+0.60}_{-1.29}$
NGC 1672	19	2	$30.99^{+0.09}_{-0.23}$	$30.98^{+0.09}_{-0.22}$	$15.80^{+0.68}_{-1.68}$
NGC 2835	27	1	$30.57^{+0.08}_{-0.17}$	$30.57^{+0.08}_{-0.17}$	$13.03^{+0.46}_{-1.04}$
NGC 3351	142	10	$30.36^{+0.06}_{-0.08}$	$30.32^{+0.06}_{-0.08}$	$11.80^{+0.31}_{-0.43}$
NGC 3627	43	7	$30.18^{+0.08}_{-0.15}$	$30.16^{+0.07}_{-0.14}$	$10.88^{+0.39}_{-0.77}$
NGC 4254	42	22	$29.97^{+0.09}_{-0.20}$	$29.91^{+0.08}_{-0.16}$	$9.86^{+0.42}_{-0.91}$
NGC 4303	19	7	$30.65^{+0.10}_{-0.26}$	$30.17^{+0.10}_{-0.25}$	$13.49^{+0.64}_{-1.60}$
NGC 4321	62	11	$31.10^{+0.06}_{-0.10}$	$30.78^{+0.08}_{-0.11}$	$16.62^{+0.46}_{-0.74}$
NGC 4535	53	0	$31.43^{+0.06}_{-0.09}$	$31.43^{+0.06}_{-0.09}$	$19.29^{+0.56}_{-0.82}$
NGC 5068	58	22	$28.46^{+0.11}_{-0.26}$	$28.54^{+0.09}_{-0.21}$	$4.93^{+0.24}_{-0.59}$
NGC 7496	13	2	$31.64^{+0.09}_{-0.19}$	$31.04^{+0.11}_{-0.23}$	$21.31^{+0.89}_{-1.89}$

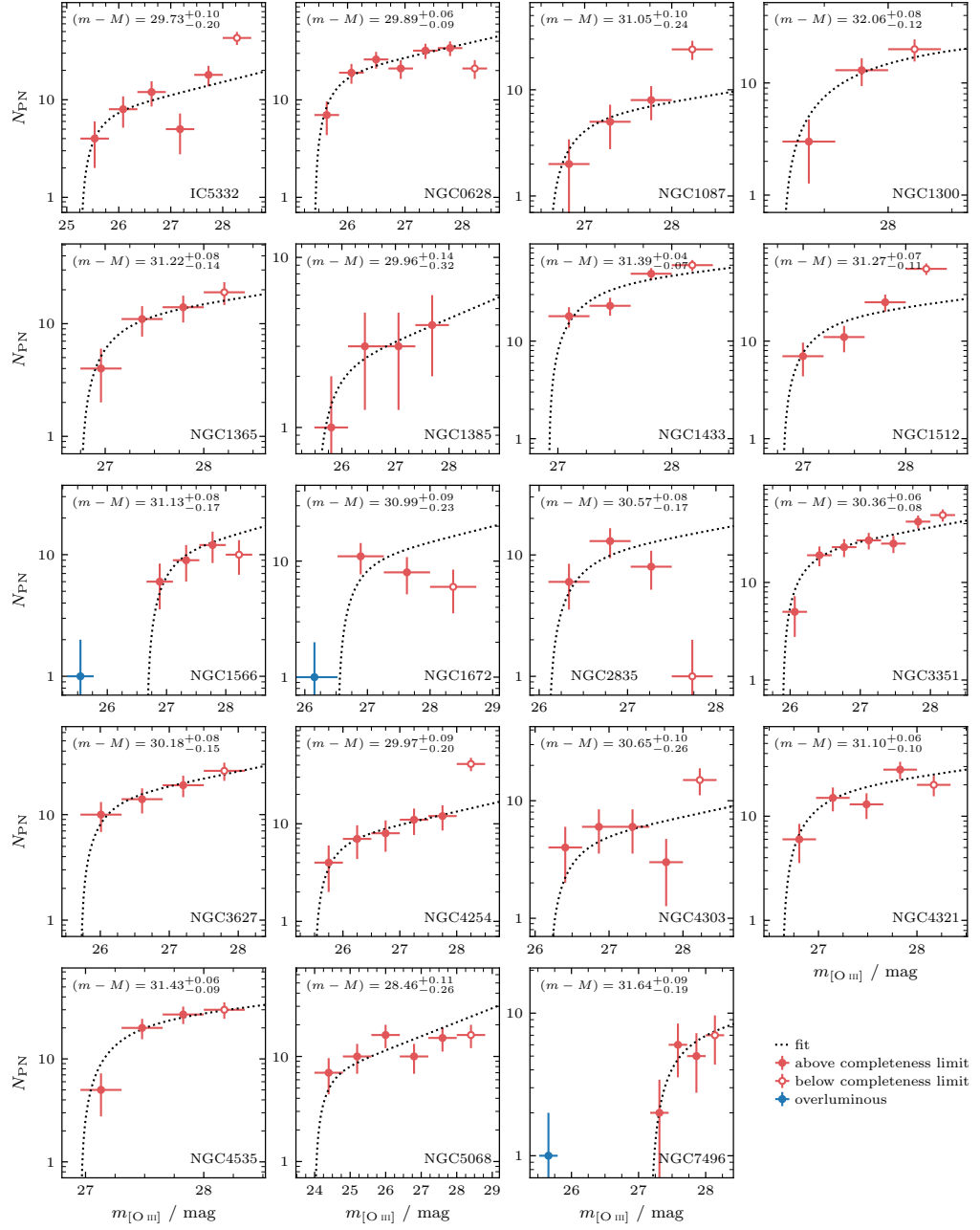


Figure 5.6: PNLF for all 19 galaxies in our sample. The filled circles denote objects brighter than our completeness limit that contribute to the fit. The y -errors are from Poisson statistics and the x -errors are the bin widths. The fit of Equation 5.7 is done via the method of maximum likelihood. This binning is only used to illustrate the result and not used in the fit. In blue are objects that were rejected as overluminous objects.

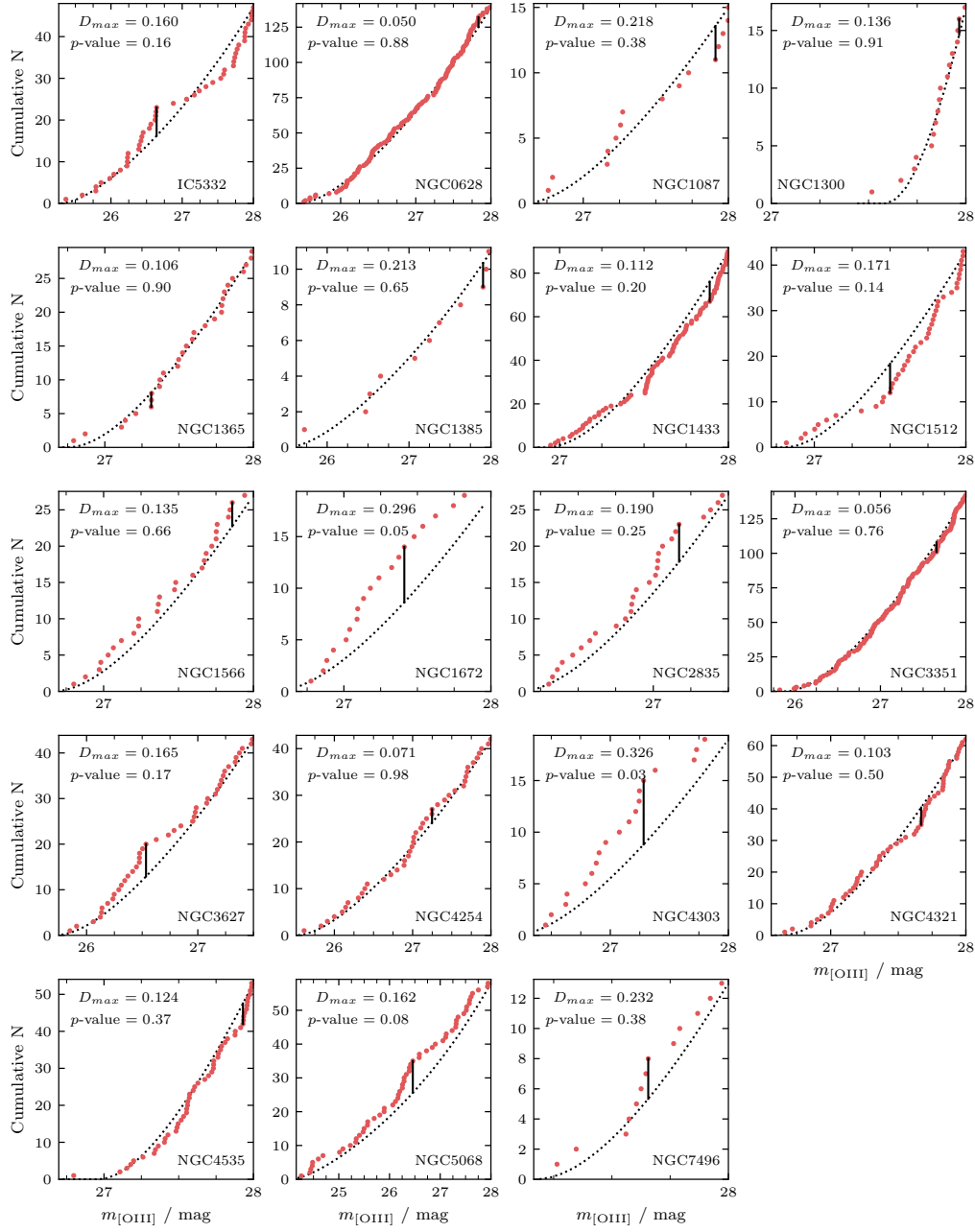


Figure 5.7: Cumulative PNLF for all 19 galaxies in our sample. In the top left corner is the statistic D_{\max} and the p -value from the KS test. The former is indicated with a vertical black bar in the plot (not normalised). The fitted luminosity function is shown with a black dotted line.

with the previously mentioned TRGB value and the latest Tully–Fisher distance. There are 16 SNRs that would be classified as PNe without the [SII] line ratio, but none of them fall at the bright end of the luminosity function and their inclusion does not alter the measured distance significantly.

NGC 0628, also known as M 74, is located in the constellation Pisces. This galaxy has been studied extensively with numerous existing distance measurements from different methods (see Figure 5.8 for a selection). Among them are the aforementioned PNLf studies by Herrmann et al. (2008), Kreckel et al. (2017) and Roth et al. (2021). This makes this galaxy a valuable benchmark to validate our methods. The distance of Kreckel et al. (2017) is based on the same PHANGS-MUSE data, but was limited to the three pointings that were available at the time, and Roth et al. (2021) uses the same raw data but with a different data reduction. We find 139 PNe that provide a well constrained distance modulus of $(m - M) = 29.89^{+0.06}_{-0.09}$ mag ($D = 9.52^{+0.26}_{-0.41}$ Mpc). This is in excellent agreement to the result from Kreckel et al. (2017) of $(m - M) = 29.91^{+0.08}_{-0.13}$ mag, while both Herrmann et al. (2008) and Roth et al. (2021) derived a smaller distance modulus of $(m - M) = 29.67^{+0.06}_{-0.07}$ mag and $(m - M) = 29.76^{+0.03}_{-0.05}$ mag, respectively. The discrepancy was explained by Kreckel et al. (2017) with SNRs that were misclassified as PNe. We are unable to confirm this explanation. Because we measure larger $H\alpha$ fluxes (see also Section 5.2.3), the objects in question are mostly classified as H II regions in our analysis (see Figure 5.3) and, therefore, do not factor into the measured distance. Roth et al. (2021) also noted that the brightest PN in their sample is 0.11 mag brighter than the second brightest PN. While not really overluminous, they re-fit their data without this point and obtain a better fit with a larger distance modulus of $(m - M) = 29.87^{+0.03}_{-0.05}$ mag. Taking the catalogue from Roth et al. (2021), we run our fitting algorithm and measure a larger distance modulus of $(m - M) = 29.97^{+0.02}_{-0.05}$ mag (also using the zero point of $M^* = -4.53$ that they adopted in their paper), indicating that the fitting algorithm and especially the treatment of the uncertainties plays an important role (as Roth et al. (2021) convolved the luminosity function with the photometric error). This galaxy also has a number of TRGB distances (e.g. Jacobs et al. 2009) and they are all in excellent agreement.

NGC 1087: lies in the constellation Cetus. The 15 PNe that are brighter than our completeness limit yield a distance modulus of $(m - M) = 31.05^{+0.10}_{-0.24}$ mag ($D = 16.25^{+0.74}_{-1.79}$ Mpc). This is in agreement with the distances derived from the *Numerical action method* (NAM) or by group affiliation (see Figure A.2).

NGC 1300: is located in the constellation Eridanus and part of the eponymous cluster. Our source detection picks up many extended sources that are associated with very significant $H\alpha$ emission. Even though this galaxy does not have an active nucleus, we mask the central region and spiral arms to get rid of those objects. We measure a distance modulus of $(m - M) = 32.06^{+0.08}_{-0.12}$ mag ($D = 25.77^{+0.90}_{-1.42}$ Mpc). Our value is considerably larger than the existing literature values (see Figure A.3). However the small sample size of only 17 PNe means that this value is more of an upper limit.

NGC 1365: is also known as the *Great Barred Spiral Galaxy* and is part of the Fornax cluster. This galaxy has an active nucleus and many of the detected objects do not have

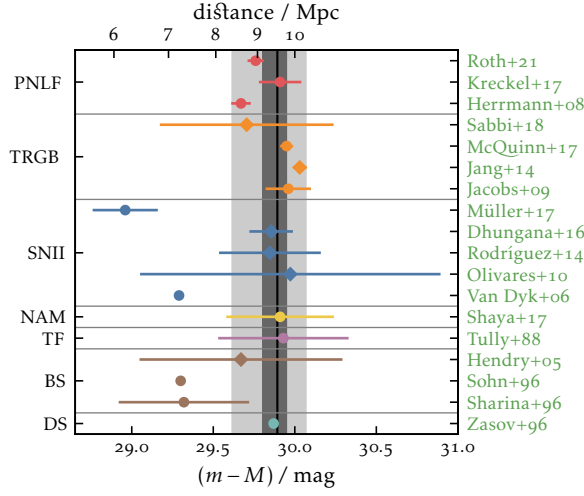


Figure 5.8: Comparison between our measured distance and literature values for NGC 0628 (M 74). The value measured by this study is marked by a black line with the one and three sigma intervals shaded in grey. We measure a distance modulus of $(m - M) = 29.89^{+0.06}_{-0.09}$ mag ($9.52^{+0.26}_{-0.41}$ Mpc). A list of abbreviations and similar figures for the other galaxies can be found in Appendix A.1.

a clean radial profile, but are instead dominated by noise. From 29 PNe we derive a distance modulus of $(m - M) = 31.22^{+0.08}_{-0.14}$ mag ($D = 17.53^{+0.66}_{-1.16}$ Mpc). This galaxy also has a number of other reliable distance estimates (Figure A.4). Both the two most recent Cepheid distances and the TRGB distance agree well with our measured distance, while another TRGB and Cepheid distance indicate a larger distance.

NGC 1385: we only detect 11 PNe that are brighter than our completeness limit. Most of them are not isolated and have other [O III] emitting sources nearby, which makes the background subtraction more challenging. The derived distance modulus of $(m - M) = 29.96^{+0.14}_{-0.32}$ mag ($D = 9.81^{+0.63}_{-1.46}$ Mpc) is therefore subject to a large uncertainty and should be considered an upper limit for the distance. The value is in good agreement with existing Tully–Fisher distances but shows a large discrepancy to the distance derived with the numerical action method (see Figure A.5). 9 SNRs are potentially misclassified as PNe. None of them fall at the bright end and their inclusion has the opposite effect of increasing the measured distance by 0.08 mag.

NGC 1433: also known as *Miltron’s Galaxy*, is a barred spiral galaxy with an active nucleus. We detect 90 PNe that are all isolated with regular radial profile. This is a bit surprising, as the resulting distance modulus of $(m - M) = 31.39^{+0.04}_{-0.07}$ mag ($D = 18.94^{+0.39}_{-0.56}$ Mpc) places it at the far end of our sample where we usually do not find a clean sample. Our measured distance is also significantly larger than all existing literature distances, including a TRGB distance that was measured from LEGUS–*HST* data (Sabbi et al. 2018). We downloaded and reduced the same set of *HST* data, using the technique

Jacoby et al. (2024) find more PNe and measure a further distance of $D = 25.0^{+1.4}_{-1.5}$ Mpc.

Jacoby et al. (2024) measure a consistent distance of $D = 19.2^{+0.7}_{-0.7}$ Mpc.

outlined in Anand et al. (2021), however we are unable to derive a distance based on the color-magnitude diagram we obtain. The data are relatively shallow and a possible explanation for the discrepancy is that Sabbi et al. (2018) measured the tip of the AGB and not the tip of the red giant branch.

*Jacoby et al. (2024) measure
a consistent distance of
 $D = 19.3_{-0.6}^{+0.5}$ Mpc.*

NGC 1512: is a barred spiral galaxy that is interacting with the nearby galaxy NGC 1510. We find 43 PNe, from which we measure $(m - M) = 31.27_{-0.11}^{+0.07}$ mag ($D = 17.93_{-0.88}^{+0.53}$ Mpc). Similar to NGC 1433, our derived distance is considerably larger than all existing literature distances (see Figure A.7), including again the TRGB distance from LEGUS–HST (Sabbi et al. 2018).

NGC 1566: is located in the constellation Dorado. We identify 28 PNe, but one of them is visibly separated from the rest of the luminosity function (by 1 mag) and including it results in a very poor fit. We therefore mark the brightest PN as overluminous and redo the analysis and get a much better fit ($D_{\max} = 0.135$, p -value = 0.66) with a distance modulus of $(m - M) = 31.13_{-0.17}^{+0.08}$ mag ($D = 16.84_{-1.29}^{+0.60}$ Mpc). This is considerably smaller than existing literature values based on the group affiliation (Kourkchi & Tully 2017, see Figure A.8). Five SNRs could be misclassified as PNe. The brightest among them is as bright as the overluminous PN and is therefore also excluded. The four remaining SNRs do not alter the measured distance.

NGC 1672: is a barred spiral galaxy with an active nucleus (Seyfert type 2). There are two objects that are significantly brighter than the rest (one PN and one SNR). Even when we exclude them, the resulting fit is still rather poor with $D_{\max} = 0.296$, p -value = 0.05 and yields a distance modulus of $(m - M) = 30.99_{-0.23}^{+0.09}$ mag ($D = 15.80_{-1.68}^{+0.68}$ Mpc). The measured value falls between the values derived from the NAM and those derived with Tully–Fisher measurements (see Figure A.9). SNRs do not impact the measured distance.

NGC 2835: this is one of two galaxies where our completeness limit estimation yields a smaller value of 27 mag. However, we do not see a drop in the luminosity function until 27.5 mag (see Figure 5.6) and, since the number of detected PNe is already small, we decided to include objects up to 27.5 mag in our sample. From the 27 PNe that match this criterion, we measure a distance modulus of $(m - M) = 30.57_{-0.17}^{+0.08}$ mag (we also measure 30.56 mag from the sample with the smaller completeness limit, $D = 13.03_{-1.04}^{+0.46}$ Mpc). Existing Tully–Fisher distances arrive at a smaller distance, but a recent TRGB distance from Anand et al. (2021) is in good agreement with our value (see Figure A.10).

NGC 3351: also known as M 95, is a barred spiral galaxy located in the Leo constellation that has a characteristic ring. This is one of the few galaxies in our sample with an existing PNLf distance. Ciardullo et al. (2002) measured a distance modulus of $(m - M) = 30.05_{-0.16}^{+0.08}$. We recover the six PNe that fall in our FoV, however two of them are classified differently. We classify their 5th brightest PN as a SNR and their 6th brightest object is classified as an H II region. This galaxy has the second largest number of PN detections in our sample and we find an excellent fit with a distance modulus of $(m - M) = 30.36_{-0.08}^{+0.06}$ mag ($D = 11.80_{-0.43}^{+0.31}$ Mpc). This is significantly larger than the previous PNLf distance and also larger than most of the existing TRGB (e.g., Rizzi et al. 2007; Sabbi et al.

2018) and Cepheid (e.g., Freedman et al. 2001; Saha et al. 2006) distances. There are however also a TRGB distance (Sakai et al. 2004) that is in good agreement and a Cepheid distance (Paturel & Teerikorpi 2006) that is even larger (see Figure A.11).

NGC 3627: also known as M 66, can also be found in the constellation Leo and is part of the Leo Triplet. This is the second galaxy from Ciardullo et al. (2002) that is also in our sample, but only one of the PNe that was detected in that paper falls within our FoV. We recover this object and measure the same $m_{[\text{OIII}]}$ magnitude but classify it as a SNR. This object is the 35th brightest PN in the catalogue of Ciardullo et al. (2002). Therefore, this misclassification will hardly impact the result. Their brighter PNe all fall outside of our FoV and, hence, we can not make any statements about them. In our PN sample, this object would constitute one of the brightest object. Including this object would have a minuscule effect on our result and decrease the measured distance modulus by 0.02 mag. The distance modulus that we measure from the clean PNe sample is $(m - M) = 30.18^{+0.08}_{-0.15}$ mag ($D = 10.88^{+0.39}_{-0.77}$ Mpc). This is in good agreement with the existing PNLf distance of $(m - M) = 29.99^{+0.07}_{-0.08}$ or other recent TRGB distances (e.g. Jacobs et al. 2009; Anand et al. 2021, see Figure A.12). There are 8 potential SNR contaminants, one of which is overluminous. If we exclude this object, the measured distance is not impacted by the other 7 SNRs.

NGC 4254: also known as M 99, is one of four galaxies in our sample that are part of the Virgo cluster. 45 objects constitute our initial PNe catalogue. However, a large number of objects are surrounded by strong H α emission and three objects are rejected because of their irregular radial profile. Despite this, the remaining 42 PNe provided an excellent fit with $D_{\text{max}} = 0.071$, $p\text{-value} = 0.982$ and yield a distance modulus of $(m - M) = 29.97^{+0.09}_{-0.20}$ mag ($D = 9.86^{+0.42}_{-0.91}$ Mpc). This is smaller than all previously published distances. Among them are two distances based on the type II supernova SN1986I (Nugent et al. 2006; Poznanski et al. 2009) and a number of Tully–Fisher estimates. They all indicate a larger distance modulus around $(m - M) = 30.75$ mag (see Figure A.13). The Hubble velocity of this galaxy is 2400 km s^{-1} (Springob et al. 2005) and also implies a significant larger distance. The inclusion of 22 misclassified SNRs does not affect the measured distance (one SNR is removed because it is clearly overluminous).

NGC 4303: also known as M 61, is also part of the Virgo Cluster and an active (Seyfert type 2) barred spiral galaxy. We detect 19 PNe from which we measure $(m - M) = 30.65^{+0.10}_{-0.26}$ mag ($D = 13.49^{+0.64}_{-1.60}$ Mpc), but the fit is not good with $D_{\text{max}} = 0.326$, $p\text{-value} = 0.03$. This galaxy did not have a good distance estimate, only a number of type II supernovae that do not agree on one distance as well as some uncertain Tully–Fisher estimates. Eight SNR could be misclassified as PNe. One of them is clearly overluminous (1.2 mag below the next three objects and 1.8 mag below the rest). The next three brightest objects are also SNRs and they are 0.6 mag brighter than the bright end of the PNLf. While there is a gap in the luminosity function, their inclusion improves the fit ($D_{\text{max}} = 0.243$, $p\text{-value} = 0.227$) and they are therefore retained in the sample. With them, the measured distance is decreased by 0.48 mag.

NGC 4321: also known as M 100 is located in the Virgo cluster. It is a *low ionization nuclear emission region* (LINER) galaxy and we mask out the centre and the spiral arms. We get a decent fit ($D_{\max} = 0.103$, $p\text{-value} = 0.50$) and measure a distance modulus of $(m - M) = 31.10^{+0.06}_{-0.10}$ mag ($D = 16.62^{+0.46}_{-0.74}$ Mpc) from 62 PNe. NGC 4321 has existing literature distances from Cepheids (Freedman et al. 2001; Tammann, Sandage & Reindl 2008) and a recent TRGB distance from Anand et al. (2021). They all agree with our reported value. Out of 11 SNRs, one falls 0.6 mag below the bright end cutoff. With this object, the measured distance modulus decreases to $30.78^{+0.08}_{-0.11}$ mag. If we only include the other 10 SNRs, the distance modulus is very similar to the one derived from the clean PN sample ($31.053^{+0.06}_{-0.09}$ mag). However the fit is not bad enough ($D_{\max} = 0.108$, $p\text{-value} = 0.336$), to exclude this one object as overluminous.

NGC 4535: is the fourth galaxy in our sample that is part of the Virgo Cluster. From 53 PNe we measure $(m - M) = 31.43^{+0.06}_{-0.09}$ mag ($19.29^{+0.56}_{-0.82}$ Mpc). This is in good agreement to the Cepheid distance from Paturel & Teerikorpi (2006) who reported a value of 31.6 mag, but significantly bigger than other Cepheid distances (e.g. Freedman et al. 2001, 30.85 ± 0.05 mag). This is the only galaxy for which we do not find any potentially misclassified SNRs.

Noteworthy is the large range of distances that we obtained for the four Virgo galaxies in our sample (ranging from 9.86 Mpc for NGC 4254 to 19.29 Mpc for NGC 4535). Among all galaxies in the cluster, there are a few published distances that are smaller than our value for NGC 4254, but neither of them is coming from TRGB, Cepheid or PNLF measurements.

NGC 5068: is located in the Virgo constellation. This is the nearest and also the least massive galaxy in our sample. We find 58 PNe and our fit yields a distance modulus of $(m - M) = 28.46^{+0.11}_{-0.26}$ mag ($4.93^{+0.24}_{-0.59}$ Mpc). This is another galaxies with an existing PNLF distance from Herrmann et al. (2008). Unfortunately, only two of their PN are in the central region that is covered by our observations. We only recover one of them and the [OIII] line map does not show any emission at the location of the second one. The data from Herrmann et al. (2008) indicate a larger distance of $28.85^{+0.09}_{-0.16}$ mag. Because of the lower metallicity of this galaxy, they decide to use a fainter M^* and obtain the reported distance of $(m - M) = 28.68^{+0.08}_{-0.18}$. We do not apply such a correction here, as this would even further decrease our measured distance, which is already smaller than the existing PNLF and TRGB distances. Within the uncertainties, our distance is in good agreement to the TRGB distances from Anand et al. (2021) and Karachentsev et al. (2017). While there is a large number of SNRs (22), their inclusion does not alter the measured distance.

NGC 7496: is a barred spiral galaxy with Seyfert type 2 activity, located in the constellation Grus. PHANGS–MUSE covered this galaxies with only 3 pointings. Due to the active core we mask the central part and arms, leading to the detection of only 14 PNe and 2 SNRs that could be misclassified. One PNe was excluded as an overluminous object. Previous distance estimates were all based on the Tully–Fisher method, with a large scatter between the different studies. Our distance modulus of $(m - M) = 31.64^{+0.09}_{-0.19}$ mag ($21.31^{+0.89}_{-1.89}$ Mpc) is in good agreement to NAM distance modulus of $(m - M) = 31.36 \pm 0.33$ mag. Both of the SNRs fall slightly below the bright end and decreases the measured distance by 0.60 mag.

The few data points that we have for this galaxy makes it difficult to judge whether these two points belong to the sample or not.

5.5 DISCUSSION

5.5.1 Comparison with literature distances

To benchmark our results, we compare them to published distances. We have literature distances for all galaxies, for some galaxies from a myriad of different methods (see Appendix A.1). For the four galaxies with existing PNLf distances, we find excellent agreement with less than 1σ difference for NGC 0628 (derived from the same data) and NGC 3627, while the previous PNLf distance to NGC 5068 is slightly larger (2σ) and the previous PNLf distance to NGC 3351 is significantly smaller (3.8σ).

Ten galaxies also have TRGB distances and with the exception of NGC 1433 and NGC 1512, they are in good agreement with our values. Furthermore, we have five galaxies with Cepheid distance. The distances derived with this method often vary between studies and our results always fall within the range of published distances. PNLf distances have historically shown a systematic offset to distances measured from SBF (Ciardullo et al. 2002). However, as there are no elliptical galaxies in the PHANGS sample, we have no SBF distances to compare with. To check for systematic offsets between our PNLf distances and other methods, we plot the difference $\Delta(m - M) = (m - M)_{\text{ref}} - (m - M)_{\text{PHANGS}}$ between a number of commonly used distance indicators and our derived distance modulus in Figure 5.9. The plot shows generally good agreement between the different methods, but there is a slight systematic offset as our PNLf distances are on average larger than the literature distances across all methods. However the offsets are generally small and always smaller than the scatter. The rather large scatter compared to the literature TRGB distances mostly stems from the discrepancies with the distances to NGC 1433 and NGC 1512. In Section 5.4 we provided a possible explanation for this and if we exclude them from the comparison, the TRGB method shows an excellent agreement with $\Delta(m - M) = -0.04 \pm 0.18 \text{ mag}$.

The galaxies that previously lacked a good distance estimate (i.e. from the PNLf, TRGB or Cepheids) are NGC 1087, NGC 1300, NGC 1385, NGC 1672, NGC 4254, NGC 4303 and NGC 7496. Unfortunately, these are also the galaxies with the fewest PN detections, which in general leads to larger errors on the measured distances. However they are usually still smaller than what is available in the literature.

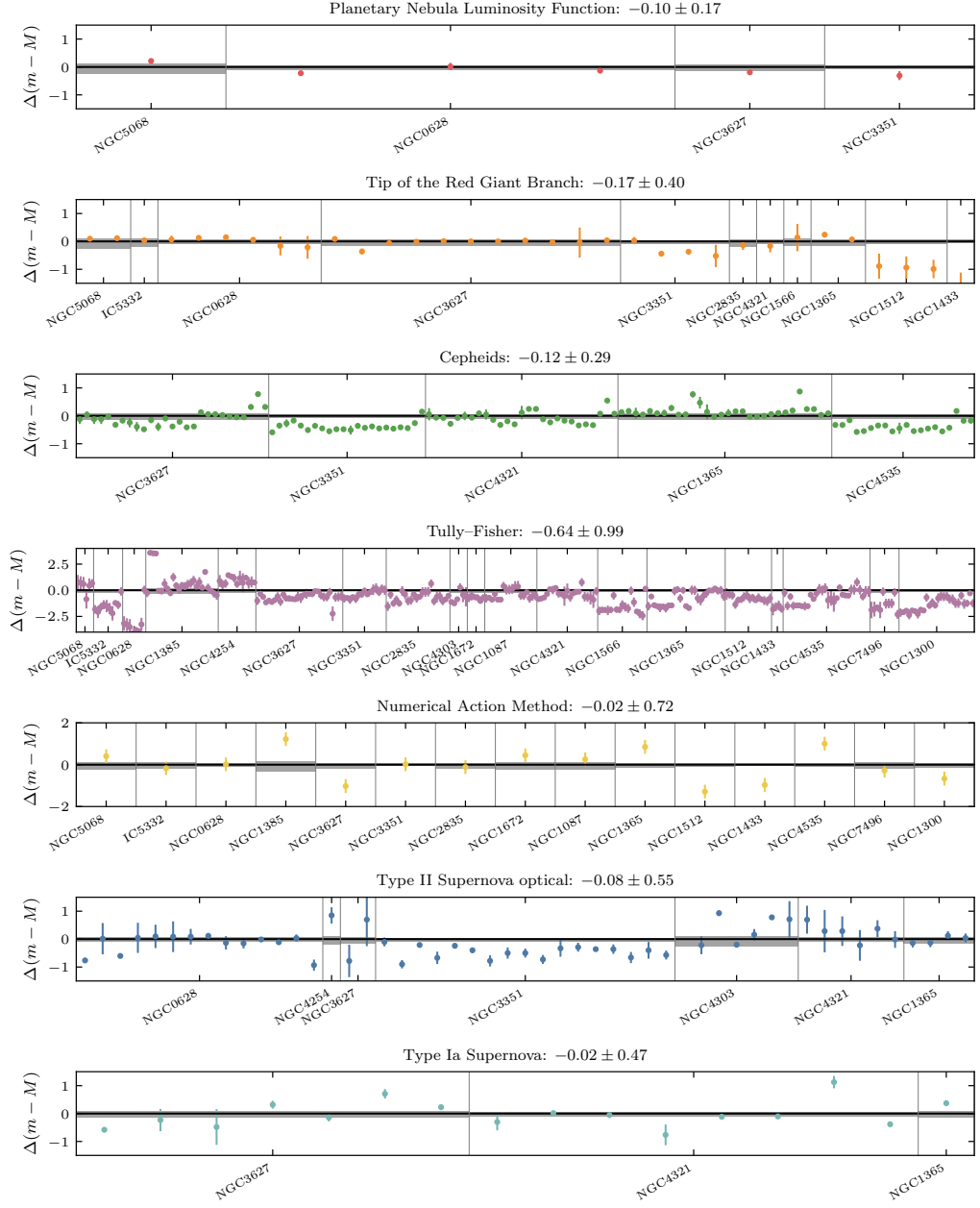


Figure 5.9: Comparison with literature distances, taken from NED. Shown is the difference between our value and the literature $\Delta(m - M) = (m - M)_{\text{ref}} - (m - M)_{\text{PHANGS}}$. The galaxies are sorted by distance and the points for each galaxy are sorted by year of publication. If more than one value was published for a source, each value is shown separately. The uncertainties of our measurement is indicated by the grey shaded area.

5.5.2 Zero point of the PNLF

The single most important input to the PNLF is the zero point. For our analysis, we assume a constant value across all galaxies, but theoretical models (Dopita, Jacoby & Vassiliadis 1992) and observations (Ciardullo et al. 2002) indicate an increase of M^* at low metallicities. The change can be described by a quadratic formula,

$$\Delta M^* = 0.928 [\text{O}/\text{H}]^2 - 0.109 [\text{O}/\text{H}] + 0.004, \quad (5.10)$$

with the solar oxygen abundance $12 + \log(\text{O}/\text{H})_{\odot} = 8.69$ from Asplund et al. (2009). We can compare our measured distances to literature distances to re-examine this relation. For this, we use the distances in Table 5.1 that were curated by Anand et al. (2021). We only use the subsample that is based on the TRGB (8 galaxies). To first approximation the change in the zero point is given by $\Delta M^* \approx (m - M)_{\text{PNLF}} - (m - M)_{\text{ref}}$. We use this as a starting point in our PNLF fit and vary M^* until we reproduce the literature distance.

One distinct advantage of IFU observations is that one can derive the PNLF distance and the metallicity from the same data. For our data, the gas-phase oxygen abundance is measured from H II regions (based on the S-calibration from Pilyugin & Grebel 2016) by Kreckel et al. (2019) and Santoro et al. (2022). We use their linear abundance gradient fits to calculate the metallicity at the mean radius of the PNe (see Table 5.1). In contrast, the metallicities that were used in Ciardullo et al. (2002) were compiled by Ferrarese et al. (2000a) from different literature sources (most of them are based on the strong line method calibration by Dopita & Evans 1986). The galaxies that overlap with our sample (NGC 3351 and NGC 3627) show a large discrepancy of ~ 0.7 dex between the PHANGS and literature metallicity (considerably more than what could be explained by differences in the value for the solar oxygen abundance). More recent abundances from Pilyugin, Grebel & Kniazev (2014) and Toribio San Cipriano et al. (2017) for the *Small Magellanic Cloud* (SMC) and *Large Magellanic Cloud* (LMC) show much better agreement with our sample (see top panel of Figure 5.10). This result shows that such analysis is severely hampered by uncertainties in the metallicity prescriptions and greatly benefits from the homogeneous sample that IFU observations can provide.

Our sample lacks low metallicity galaxies to constrain the decrease at low metallicities and Ciardullo (2012) found that for galaxies with metallicities above the LMC, the zero point is roughly constant. Even though we can not test the metallicity dependence, it is still worth to re-examine the zero point. Fitting a constant line to our own data (see lower panel of Figure 5.10) we find a zero point of $-4.469^{+0.049}_{-0.025}$ mag. The fit is heavily skewed by NGC 3351 where our distance is more than 3σ larger than the TRGB distance. We therefore use a mixture model for outlier pruning (Foreman-Mackey 2014) with which we measure a zero point of $-4.542^{+0.103}_{-0.059}$ mag. This value is in excellent agreement to the value that Ciardullo (2012) derived with the help of TRGB distances. Because our sample only consists of 8 points, one of which is rejected by the outlier pruning, we prefer to use the initial adopted value of $M^* = -4.47$ for the PNLF distances computed in this paper. Given the growing databases of galaxies with both TRGB distances and optical IFU data in the archives, this topic is ripe for future work to revisit this question using larger samples of galaxies.

Adopted from Ciardullo et al. (2002) but with $12 + \log(\text{O}/\text{H})_{\odot}$ from Asplund et al. (2009) instead of Grevesse et al. (1996).

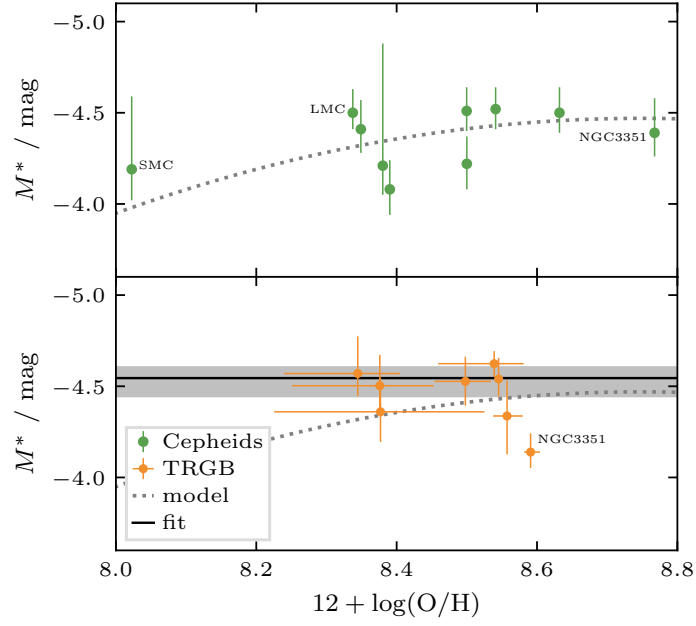


Figure 5.10: Determining the zero point of the PNLF. The points in the top panel are based on Cepheid distances from Ciardullo et al. (2002) and abundances from Pilyugin, Grebel & Kniazev (2014). The points in the lower panel are based on our measured distances in combination with existing TRGB distances from Table 5.1. The metallicity is determined by taking the value of the H II region metallicity gradient (Kreckel et al. 2019; Santoro et al. 2022) at the mean galactic radius of the PNe. The error bars in x -direction represent the minimum and maximum abundance for the PNe in our sample. The dotted curve shows the theoretical dependence from Equation 5.10 and the solid line is a fit with a constant line (with the 1σ interval in grey). Overall our sample is consistent with a constant zero point.

5.5.3 Contamination with supernova remnants

Since Kreckel et al. (2017) found that SNRs that are misclassified as PNe can alter the distances measured from the PNLf, the question stood whether narrowband observations are sufficient for PN studies. In order for SNRs to bias the results, they must meet a couple of criteria. First of all, they must be classified as PN by narrowband studies. As we see in Figure 5.3, most SNRs are in a regime where they would be classified as HII regions without the additional diagnostics from the [SII] line. Second, the contamination must occur at the bright end of the luminosity function. The way the fitting of the PNLf works, only the brightest objects have a noticeable contribution to the outcome. Third, the contaminant must not be too much brighter than the bright end cut-off. When a single object is much brighter than the main sample, it makes the fit worse and is therefore easily excluded as being overluminous. Hence, either a single SNR close to the bright end or a number of SNRs that are fainter than the bright end are needed to change the distance.

In almost all galaxies in our sample, we find at least one SNR that could be classified as a PN without the [SII] criteria. However, only in the case of three galaxies do the misclassified objects have a significant impact the measured distance. This is the case for NGC 4303 (-0.48 mag), NGC 4321 (-0.33 mag) and NGC 7496 (-0.60 mag). In the case of NGC 4303, three SNRs fall 0.5 mag below the cutoff, however they do not worsen the fit, making it difficult to justify their exclusion. NGC 4321 has one SNR that is clearly separated from the rest of the sample, but the goodness of the fit is not impacted. Finally for NGC 7496 the sample is too small to reliably exclude the two slightly overluminous SNRs.

Directly related to this is the challenge in accurately determining the flux emitted by an object. The largest uncertainty arises from background contamination. The [OIII] line fluxes are relatively robust because there is little to no background emission. However, this is not the case for $H\alpha$. The background can have a significant contribution and so correct background subtraction is crucial. As shown in Figure 5.2, different approaches to the background subtraction can lead to vastly different results for the $H\alpha$ fluxes. If not done properly, wrong fluxes can alter the line ratios and hence the classification of objects.

5.5.4 Limitations of IFU observations to the PNLf

The PNLf is rarely applied to galaxies beyond 20 Mpc, since with narrowband imaging, the expected flux from a single PN at larger distances is similar to the background. Recent attempts by Ventimiglia, Arnaboldi & Gerhard (2011), Arnaboldi et al. (2013), or Roth et al. (2021) have pushed the limits of what is possible by using multi-slit imaging spectroscopy or IFU spectroscopy. Here, we make a rough estimate on the maximum distance that can be measured from MUSE with exposure times similar to our observations (43 min). We require a certain number of PN detections (from our experience $\gtrsim 20$) in order to measure a reliable distance. However, it is difficult to predict the number of PNe that one can expect to observe in a given galaxy, as it depends on a number of

parameters. It relies foremost on the underlying stellar population. More stars means more PNe. To compare the number of detected PNe between galaxies, it is useful to define the *luminosity-specific planetary nebula number*,

$$\alpha = \log_{10} \frac{N_{\text{PN}}}{L_{\text{bol}}}, \quad (5.11)$$

where N_{PN} is the number of detected PNe above our completeness limit and L_{bol} is the bolometric luminosity of the surveyed area. This definition is different from the $\alpha_{2.5}$ that is common in the literature (e.g., Hartke et al. 2017) in that we only use the number of PNe that are brighter than our completeness limit and not within a fixed range above the bright end cut off. The bolometric luminosity of the survey area is estimated from a simulated V -band image via a bolometric correction. We use $f_{\lambda} = 363.1 \cdot 10^{-11} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ (Bessell, Castelli & Plez 1998) as the zero point to convert fluxes into magnitudes and then calculate the bolometric luminosity as

$$L_{\text{bol}} = 10^{-0.4(M_V - 4.79)} 10^{-0.4(\text{BC}_V + 0.07)} L_{\odot}, \quad (5.12)$$

with $\text{BC}_V = -0.85 \text{ mag}$ (from Buzzoni et al. 2006). Since we also have stellar mass maps, we could also utilise those instead, but due to its prevalence in the literature, we opt to use L_{bol} . Both yield similar results.

Because all galaxies have the same observed completeness limit (for NGC 2835 and NGC 3627 we extrapolate the number of PNe to match the other galaxies), the part of the PNLf that we observe decreases with distance. This reduces N_{PN} while L_{bol} remains unaffected and, hence, we expect a decrease in α as can be seen in Figure 5.11. As evident in Figure 5.11, compared to NGC 0628, more than half of the galaxies in our sample have fewer PNe than expected. We find up to a factor of 7 difference in the number of PNe. A possible explanation is a decrease of α with metallicity. But our galaxies are all similar and only show small variations in metallicity (0.24 dex). The observed change is too large to be explained by the models of Buzzoni et al. (2006). However, a recent study by Galán-de Anta et al. (2021) also finds comparable differences.

We can use this relation to calculate the farthest distance at which we are able to obtain a sufficient number of PNe to measure a reliable distance. For each galaxy, we use the previously measured L_{bol} to calculate the α that corresponds to $N_{\text{PN}} = 20$. We can then compute the distance that corresponds to this value based on the sampling of the PNLf. Based on this analysis, we find the largest possible distance that could be measured for a galaxy in our sample is $\sim 26 \text{ Mpc}$.

To some degree, the number of detections should also decrease with resolution as fainter objects start to blend in with their background at coarser resolution. However, Kreckel et al. (2017) showed that this only becomes an issue above 100 pc physical resolution. With the typical seeing that we achieve with MUSE (0.60 arcsec with AO and 0.76 arcsec without), this corresponds to a distance of 34.38 Mpc and 27.14 Mpc, respectively, both above what we could measure. This means that this issue should only concern less powerful IFU instruments with a coarser spatial resolution and indeed we do not see any obvious connection between the resolution and the number of detections in our data.

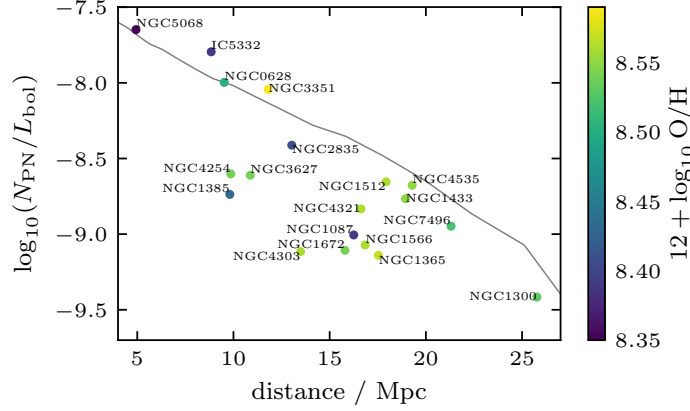


Figure 5.11: Luminosity-specific planetary nebula number as a function of distance. The number of detections decreases with distance because we sample a smaller range of the PNLf. The grey line uses NGC 0628 as a reference and shows the expected decrease of α based on the sampled part of the PNLf.

5.6 CONCLUSIONS

We use VLT/MUSE observations obtained by the PHANGS collaboration to identify 899 PNe across 19 nearby galaxies. We then calculate new PNLf distances for each galaxy (see Table 5.3), 15 of which did not have a PNLf distance before. This significantly increases the number of galaxies with PNLf distances, which previously stood at around 70. Fourteen galaxies achieve an uncertainty better than 10 percent and six galaxies better than 5 percent.

When comparing with other studies, we find good agreement of the [OIII] line fluxes but significant variations in the $H\alpha$ line fluxes. This can be traced back to different techniques for background subtraction (unlike [OIII], $H\alpha$ has a significant background contribution). This highlights the difficulty of the background subtraction which can have a major impact on the classification and, thereby, ultimately also on the measured distance.

The findings of Kreckel et al. (2017) raised questions about the reliability of narrow-band observations for PNe studies, with the possibility of misclassified SNRs biasing the measured distance. While we can not reproduce this issue in the particular case of NGC 0628, we find three other galaxies where SNRs can impact the measured distance. For the other galaxies, the changes are insignificant and smaller than the uncertainties. Most misclassified SNRs fall in the regime of H II regions, and hence, do not contribute to the measured distance. The few SNRs that are classified as PNe often are at the faint end of the luminosity function where their contribution is negligible.
























With metallicities derived from the same MUSE data and TRGB distances from the literature, we revisit the calibration of the PNLf zero point. Our sample is consistent with a constant zero point and we find a value of $-4.542^{+0.103}_{-0.059}$ mag.

Because they do not require dedicated observations, IFU surveys can potentially multiply the number of galaxies with PNLF distances in the near future. A growing number of spectral imaging surveys that utilise VLT/MUSE like MAD (den Brok et al. 2020) or TIMER (Gadotti et al. 2019), or *Canada–France–Hawaii Telescope (CFHT)/Spectromètre Imageur à Transformée de Fourier pour l’Etude en Long et en Large de raies d’Emission* (SITELLE) like *Star formation, ionized gas and nebular abundances legacy survey* (SIGNALS) (Rousseau-Nepton et al. 2019) fill the archives with a plethora of suitable observations that could be harnessed to measure PNLF distances.

THE MUSE VIEW OF THE SCULPTOR GALAXY

THE *Sculptor galaxy* NGC 253, is the southern, massive, star-forming disk galaxy that is closest to the Milky Way. We present a new 103-pointing MUSE mosaic of this galaxy that covers most of its star-forming disk up to $0.75 \times R_{25}$. With an area of $\sim 20 \times 5$ arcmin ($\sim 20 \times 5$ kpc, projected) and a physical resolution of ~ 15 pc, this mosaic constitutes one of the largest integral field spectroscopy surveys with the highest physical resolution of any star-forming galaxy to date. We exploited the mosaic to identify a sample of ~ 500 PNe (the sample is ~ 20 times larger than in previous studies) to build the PNLF and obtain a new estimate of the distance to NGC 253. The value we obtained is 17 per cent higher than the estimates returned by other reliable measurements, which were mainly obtained via the *tip of the red giant branch* method. The PNLF also varies between the centre ($r < 4$ kpc) and the disk of the galaxy. The distance derived from the PNLF of the outer disk is comparable to that of the full sample, while the PNLF of the centre returns a distance that is larger by ~ 0.9 Mpc. Our analysis suggests that extinction related to the dust-rich interstellar medium and edge-on view of the galaxy (the average $E(B - V)$ across the disk is ~ 0.35 mag) plays a major role in explaining both the larger distance recovered from the full PNLF and the difference between the PNLFs in the centre and in the disk.

CONTRIBUTIONS

This chapter uses parts of an article that was published in A&A as Congiu et al. (2025). The original article was written by Enrico Congiu , Fabian Scheuermann , Kathryn Kreckel , Adam K. Leroy , Eric Emsellem , Francesco Belfiore , Johanna Hartke , Gagandeep S. Anand , Oleg Egorov , Brent Groves , Timo Kravtsov , David A. Thilker , Carla Tovo, Frank Bigiel , Guillermo A. Blanc , Alberto Bolatto , Serena A. Cronin , Daniel A. Dale , Rebecca McClain , J. Eduardo Méndez-Delgado , Elias Oakes , Ralf S. Klessen , Eva Schinnerer , and Thomas G. Williams . The data reduction and detection of the PNe were primarily carried out by Enrico Congiu, with contributions from other co-authors. The fit of the PNLF relies on the code I developed in [Paper I](#). The following text is an abridged summary of the findings that are relevant to this thesis. This work was conducted in the »Ionized ISM and its Relation to Star Formation« SWG of PHANGS.

6.1 THE MOST DETAILED MUSE MOSAIC

The galaxies we have dealt with so far sit at distances between 5 and 20 Mpc and have a median spatial resolution of 71 pc using MUSE (Groves et al. 2023). One could improve the resolution by targeting closer galaxies, but doing so comes with a trade-off. The angular size would also grow, requiring a greater number of pointings to cover a representative fraction of the disc. In practice, observers often have to settle on a feasible compromise – either map a small region at high resolution or go for a coarser resolution across the full system. But there is a third option, namely to invest a lot of observing time into a single object. This strategy has been attempted with MUSE, for instance by Della Bruna et al. (2022b), who observed M 83 at a distance of 4.9 Mpc in 26 pointings, and McLeod et al. (2021), who observed NGC 300 at a distance of 2 Mpc in 35 pointings. The present study, however, takes things to another level.

NGC 253 is a massive ($M \sim 4.37 \cdot 10^{10} M_{\odot}$), nearby ($D \sim 3.7$ Mpc), and actively star-forming ($\text{SFR} \sim 5.01 M_{\odot} \text{ yr}^{-1}$) galaxy (Leroy et al. 2021b). With an apparent size of 42×12 arcmin (Jarrett et al. 2019), it is both one of the largest and most prominent galaxies in our immediate neighbourhood. It was observed with MUSE in 103 pointings, with a total exposure time of 51.5 h (PI: Congiu). As of the writing of this thesis it represents the largest contiguous MUSE data set. To put this in context, the other galaxies in the PHANGS sample contain between 3 to 15 pointings and all 19 galaxies together were covered by a total of 168 pointings.

Boyce et al. (2017) observed the LMC with more pointings, but they do not form a continuous image.

The data reduction broadly follows the methods described in Emsellem et al. (2022), with additional refinements in the alignment and sky subtraction. This became necessary due to the large number of exposures and the proximity of the galaxy. Its systemic velocity is only a few hundred km s^{-1} , comparable to the amplitude of the internal rotation. As a result, residual sky features and small alignment errors could more easily interfere with the galaxy signal, requiring extra care in the reduction.

After everything has been stitched together, the final product is an astonishing image with a contiguous field of 5×20 arcmin covering the central star-forming disk, at a spatial resolution of 15 pc. The DAP is used to extract a number of emission line maps, which form the basis for the subsequent analysis. The result is the marvellous mosaic shown in Figure 6.1. The top panel is a three color composite image based on synthetic broadband filters, while the lower panel shows a composition of emission lines from hydrogen, nitrogen, sulphur and oxygen.

Its combination of depth, spatial resolution, and contiguous coverage makes this mosaic an unprecedented resource for connecting small-scale ISM physics to global galaxy structure. Overall, the richness of this data set makes it a genuine treasure trove with many possible scientific applications. A low-hanging fruit – and one that delivers important information for further studies – is the determination of the galaxy’s distance using the PNLf.



Figure 6.1: Three colour composite images of the Sculptor galaxy NGC 253. The mosaics consist of 103 individual MUSE pointings. *Top*: the channels of the RGB composite are created from the MUSE data cube by integrating over synthetic filters. *Bottom*: emission lines like hydrogen, nitrogen, sulphur and oxygen are used for the channels. A closer look reveals many bright purple dots that pop out and are potential PNe. Credits: ESO and Congiu et al. (2025).

6.2 PLANETARY NEBULAE LUMINOSITY FUNCTION IN NGC 253

The first step is once again the compilation of a PNe catalogue. They are commonly detected as bright point sources in the [OIII] emission line map. Detection algorithms such as DAOPHOT (Stetson 1987) are often used to automate this task. However, they require accurate knowledge of the PSF, and many pointings unfortunately lack sufficiently bright stars to measure the FWHM beforehand. For this reason, unlike in Paper I, the detection of the PN is not automated but done by hand. Potential candidates were identified by eye using their strong [OIII] $\lambda 5007$ emission, and contaminants such as HII regions or SNRs were removed after cross-checking them in the $H\alpha$ and [SII] maps. Since this approach is subjective, three co-authors independently performed this task, detecting 832 candidates in total, of which 444 were detected by all three.

Because the photometry also depends on the PSF, the unresolved PNe themselves were then used to determine the FWHM in each pointing. Fluxes were then measured with PHOTUTILS (Bradley et al. 2019) and corrected for local background emission, estimated from an annulus. They were corrected for Galactic extinction, but not for internal extinction. Applying the same emission line diagnostics as in Paper I, 571 candidates were classified as PNe.

Using the same technique as in Paper I returns a consistent distance, but way to many false detections at the faint end.

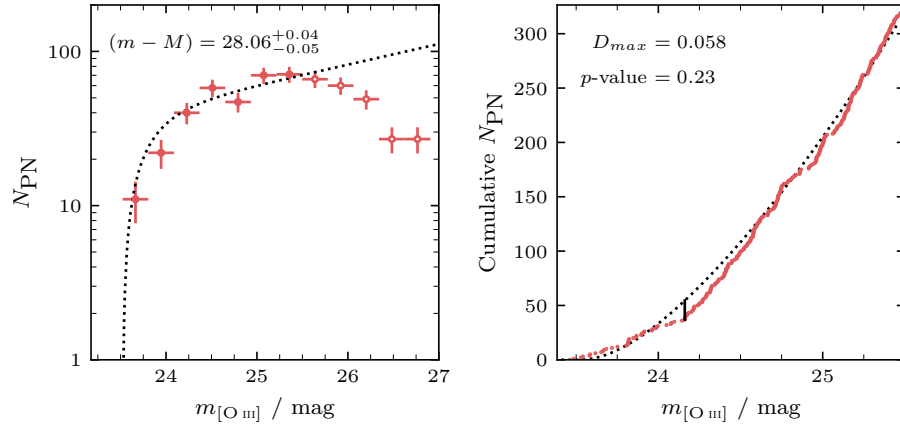


Figure 6.2: *Left*: PNLF for NGC 253. The open circles are fainter than our completeness limit and not used in the fit. The x -errors are the bin widths and the y -errors are from Poisson statistics. *Right*: cumulative PNLF with D_{\max} and the p -value from the KS test.

The completeness of the sample must also be reassessed to account for missing sources at the faint end. We do this by visually comparing the three manually detected samples, all of which begin to drop at 25.5 mag. We adopt this comparatively bright threshold as our completeness limit and still end up with 320 PNe that are usable for the analysis.

The fit of the PNLF from Equation 5.7 is done using the maximum likelihood approach presented in Paper I, but with a slightly lower zero point of $M^* = -4.54$ mag. The result is shown in Figure 6.2 and we find a distance modulus of $(m - M) = 28.06^{+0.04}_{-0.05}$ mag and a distance of $D = 4.10^{+0.07}_{-0.09}$ Mpc.

6.3 COMPARISON WITH LITERATURE DISTANCES

The comparison with other literature distances in Figure 6.3 shows that most recent estimates, such as those from the TRGB method, place NGC 253 between 3.3 to 3.6 Mpc (Newman et al. 2024; Okamoto et al. 2024). This is significantly smaller than our value of $4.10^{+0.07}_{-0.09}$ Mpc. Our measurement falls between two previous PNLF estimates, so we begin our comparison with those studies.

The distance reported by Rekola et al. (2005) was measured from ground-based narrow-band imaging and agrees with most other literature distances. Most of their PNe have a counterpart in our catalogue, however, in our data, their brightest object was classified as an H II region. Redoing our fit, including the misclassified object, yields a distance that is consistent with the value from Rekola et al. (2005). The slight deviations in the measured fluxes between the two catalogues have only a minor impact on the derived distance. The difference in this case is mainly driven by a single misclassified object.

Jacoby et al. (2024) also used MUSE data, but their analysis was limited to two archival central pointings. Their measurement of $D = 5.4^{+0.3}_{-0.6}$ Mpc is significantly larger than all

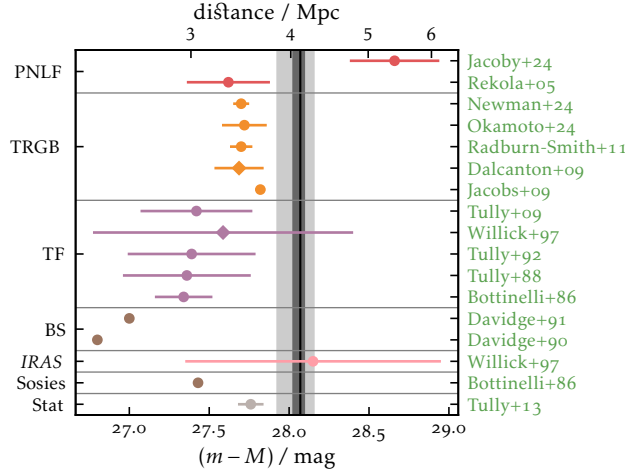


Figure 6.3: Comparison between our measured distance and literature values for NGC 253. The value measured by this study is marked by a black line with the one and three sigma intervals shaded in grey. We measure a distance modulus of $28.06^{+0.04}_{-0.05}$ mag.

other recent distance estimates. Restricting our analysis to only the objects in common with their catalogue raises our distance to 4.95 Mpc, bringing the two results into agreement within 1σ .

Thanks to the large sample size, it was possible to examine the discrepancies in more detail. We divide our sample into three subsamples: the central region within 4 kpc, the middle region between 4 and 8 kpc, and the outer region beyond 8 kpc. Each bin contains enough PNe to redo the analysis independently. The result from the middle and outer bins are slightly smaller but consistent with our overall distance, whereas the central bin yields a significantly larger value. Notably, the central region shows a deficit at the bright end of the luminosity function.

We tested various factors that could cause the discrepancy. Variations in aperture size and metallicity dependence can be ruled out and we identified dust extinction as the main problem. While the flux is corrected for Milky Way extinction, the internal extinction is not taken into account. The usual justification is that the vertical scale height of PNe exceeds that of the dust, so part of the population can be observed without significant attenuation. However, this assumption does not necessarily hold, especially for dusty or inclined galaxies. NGC 253 is highly inclined ($i = 76^\circ$) and therefore particularly susceptible to this effect.


















Without knowing the scale height, it is difficult to accurately correct for extinction. Instead, we model the problem by calculating the distance as a function of $E(B-V)$ and the scale height of the PNe and dust. If the scale height of PNe is lower than the dust one, $E(B-V) \sim 0.3$ mag is already enough to explain the discrepancy. This shows that one must be careful when applying this technique and be aware of the structure and dust content of the galaxy.

STELLAR ASSOCIATIONS POWERING H II REGIONS

I. DEFINING AN EVOLUTIONARY SEQUENCE

CONNECTING the gas in H II regions to the underlying source of the ionizing radiation can help us constrain the physical processes of stellar feedback and how H II regions evolve over time. With PHANGS–MUSE we detect nearly 24 000 H II regions across 19 galaxies and measure the physical properties of the ionized gas (e.g. metallicity, ionization parameter, density). We use catalogues of multi-scale stellar associations from PHANGS–HST to obtain constraints on the age of the ionizing sources. We construct a matched catalogue of 4177 H II regions that are clearly linked to a single ionizing association. A weak anti-correlation is observed between the association ages and the $H\alpha$ equivalent width $EW(H\alpha)$, the $H\alpha$ /FUV flux ratio, and the ionization parameter, $\log q$. As all three are expected to decrease as the stellar population ages, this could indicate that we observe an evolutionary sequence. This interpretation is further supported by correlations between all three properties. Interpreting these as evolutionary tracers, we find younger nebulae to be more attenuated by dust and closer to giant molecular clouds, in line with recent models of feedback-regulated star formation. We also observe strong correlations with the local metallicity variations and all three proposed age tracers, suggestive of star formation preferentially occurring in locations of locally enhanced metallicity. Overall, $EW(H\alpha)$ and $\log q$ show the most consistent trends and appear to be most reliable tracers for the age of an H II region.

CONTRIBUTIONS

This chapter and Appendix B were published in MNRAS as Scheuermann et al. (2023), hereafter referred to as [Paper II](#). The original article was written by Fabian Scheuermann , Kathryn Kreckel , Ashley T. Barnes , Francesco Belfiore , Brent Groves , Stephen Hannon , Janice C. Lee , Rebecca Minsley, Erik Rosolowsky , Frank Bigiel , Guillermo A. Blanc , Médéric Boquien , Daniel A. Dale , Sinan Deger , Oleg Egorov , Eric Emsellem , Simon C. O. Glover , Kathryn Grasha , Hamid Hassani , Sarah M. R. Jeffreson , Ralf S. Klessen , J. M. Diederik Kruijssen , Kirsten L. Larson , Adam K. Leroy , Laura A. Lopez , Hsi-An Pan , Patricia Sánchez-Blázquez , Francesco Santoro , Eva Schinnerer , David A. Thilker , Bradley C. Whitmore , Elizabeth J. Watkins , and Thomas G. Williams . The idea for this project came from Kathryn Kreckel. I developed the code, produced the figures, and wrote the text. The next eight co-authors were part of the paper team and regularly contributed to the work. Before publication, the other co-authors provided additional feedback. This work was conducted in the »Ionized ISM and its Relation to Star Formation« SWG, with additional input from the »Stellar Populations and Clusters« SWG of PHANGS.

7.1 INTRODUCTION

The formation of stars on galactic scales is a continuous cycle in which material from previous generations is recycled into new stars. This so-called *baryon cycle* is regulated by the feedback from massive ($> 8 M_{\odot}$), short-lived stars (Hopkins et al. 2014; Kim & Ostriker 2017). They produce UV radiation that ionizes the surrounding gas, forming H II regions. Stellar winds and supernovae will further deposit energy in the cloud and enrich the gas, but the latter also mark the end of the short life (3 to 30 Myr; Ekström et al. 2012) of the most massive O and B stars. If these feedback mechanisms are strong enough to overcome gravity, they are able to disperse the host cloud (Dale et al. 2014; Rahner et al. 2017; Haid et al. 2018; Kim, Kim & Ostriker 2018; Kruijssen et al. 2019; Chevance et al. 2022), but if not, the H II region will cease to exist once the last B stars explode. The exact ages of H II regions are difficult to pin down, as line emission can vary as a function of multiple local physical conditions (age, but also e.g. metallicity or density). One approach is to determine the ages of underlying star clusters (e.g. Whitmore et al. 2011; Hollyhead et al. 2015; Hannon et al. 2019; Stevance et al. 2020; Hannon et al. 2022), which can be estimated by fitting the observed SED with theoretical models (e.g. Turner et al. 2021). Direct constraints from ionized nebulae on the other hand are rather rare, and existing studies are almost exclusively on kpc scales. One possibility is to use the $H\alpha$ *equivalent width* $EW(H\alpha)$ (Dottori 1981; Copetti, Pastoriza & Dottori 1986; Fernandes, Leão & Lacerda 2003; Levesque & Leitherer 2013) or the $H\alpha/FUV$ ratio. Both fluxes are extensively used, most commonly as tracers for star formation (Lee et al. 2009) or to constrain the initial mass function (Meurer et al. 2009; Hermanowicz, Kennicutt & Eldridge 2013), but they have also been used on cloud scales as age indicators (e.g. Sánchez-Gil et al. 2011; Faesi et al. 2014). $H\alpha$ is only created by the emission of the most massive stars in the stellar population. The stellar continuum in the FUV and underlying $H\alpha$, on the other hand, have a significant contribution from lower mass stars. Hence, once the most massive stars are gone, both $EW(H\alpha)$ and $H\alpha/FUV$ start to decline. Assuming an instantaneous burst of star formation, population synthesis models like STARBURST99 (Leitherer et al. 2014) or BPASS (Eldridge & Stanway 2009), the latter in combination with the photoionization model CLOUDY (Ferland et al. 2017), also predict that after 2 to 3 Myr the ratio decreases monotonically with the age of the cluster.

Another potential age tracer is the *ionization parameter* q , defined as the ratio of the incident ionizing photon flux Φ_{H^0} to the local hydrogen density n_H (Kewley & Dopita 2002). In the case of a spherical geometry, one can write the ionization parameter as

$$q = \frac{\Phi_{H^0}}{n_H} = \frac{Q_{H^0}}{4\pi R^2 n_H}, \quad (7.1)$$

where Q_{H^0} is the rate of ionizing photons striking the cloud at a distance R . Assuming the model of a partially-filled *Strömgren sphere* (Charlot & Longhetti 2001), this becomes

$$q = (\alpha_B \epsilon)^{2/3} \left(\frac{3 Q_{H^0} n_H}{4\pi} \right)^{1/3}, \quad (7.2)$$

where α_B is the case-B hydrogen recombination coefficient and ϵ is the volume-filling factor of the gas. As the stellar population ages, the rate at which it produces ionizing

photons decreases (Smith, Norris & Crowther 2002) and so does the density as the H II region expands (if the shell does not sweep up additional material), leading to a decrease of the ionization parameter over time. Dopita et al. (2006) present a more realistic scenario and also conclude that the ionization parameter decreases as a function of time, with secondary dependencies on the ambient pressure, the cluster mass and the metallicity. Unlike the other two quantities, $\log q$ can not be measured directly, but the $[\text{SIII}]/[\text{SII}]$ ratio is a good proxy for it (Diaz et al. 1991).

The physical conditions of the ISM (e.g. metallicity, ionization parameter, density) regulate future star formation, and can be diagnosed via optical emission line ratios (Kewley et al. 2019). However some ratios can be sensitive to multiple properties, and careful work is needed to break those degeneracies (Kewley & Dopita 2002; Kewley et al. 2019). One particular issue is the relation between the chemical abundance and the ionization parameter. Some studies find an anti-correlation between the two (e.g. Pérez-Montero 2014; Espinosa-Ponce et al. 2022), while others find a positive correlation (e.g. Kreckel et al. 2019; Grasha et al. 2022). This discrepancy is sometimes attributed to the resolution of the observations (Kewley et al. 2019), the type of galaxy that is studied (star-forming or not, Dopita et al. 2014) or to the underlying photoionization models that are used to measure the metallicity and ionization parameter (Ji & Yan 2022).

With an established age sequence, it becomes possible to directly track how conditions in the ISM evolve as a result of stellar feedback processes. However, it is challenging to identify an ionizing source for each H II region and the sample for which we are able to is therefore biased. Direct age constraints from the ionized gas would allow us to include H II regions without an ionizing source and hence study an unbiased populations. With this, it becomes possible to study the evolutionary timescales, using large representative samples of H II regions to probe the different stages in their evolution. Modern IFU spectroscopic surveys enable us to isolate individual H II regions at $< 100 \text{ pc}$ scale and observe thousands of H II regions across individual galaxies. Previous work has typically focused on detailed case studies of individual galaxies (Niederhofer et al. 2016; McLeod et al. 2020; Della Bruna et al. 2021; McLeod et al. 2021). In order to investigate the dependence on galactic properties like environments (morphological features), metallicities or star formation rates, a more comprehensive sample is required. The PHANGS collaboration studies star formation in nearby galaxies, using large samples of GMCs, H II regions and star clusters. It combines observations from ALMA (Leroy et al. 2021b), MUSE (Emsellem et al. 2022), *HST* (Lee et al. 2022) and *JWST* (Lee et al. 2023) for a sample of 19 nearby galaxies. This provides an unprecedented sample that allows us to study the ionized gas in H II regions together with the massive stars that ionize them. We produce for the first time a catalogue of cross-matched ionizing sources and ionized nebulae, well suited for addressing how stellar feedback evolves (empirically) and constraining stellar feedback models.

This chapter is organised as follows: in Section 7.2, we present the data and the existing catalogues that are used in the analysis. In Section 7.3 we match the H II regions to their ionizing sources. In Section 7.4 we establish the H II region evolutionary sequence and discuss our findings and conclude in Section 7.5.

7.2 DATA

We perform a joint analysis of the 19 PHANGS galaxies listed in Table 7.1, combining MUSE optical spectroscopy, *HST* stellar association catalogues, and *AstroSat* FUV imaging. The galaxies in the sample have masses in the range of $9.4 < \log(M_*/M_\odot) < 11.0$ and star formation rates $-0.56 < \log(\text{SFR}/M_\odot \text{ yr}^{-1}) < 1.23$. They contain diverse morphological features like bars, rings, and active galactic nuclei, enabling us to study the impact of a wide variety of parameters.

7.2.1 MUSE H II region catalogue

To trace the ionized gas, we use IFU data observed by MUSE (Bacon et al. 2010) at the VLT in Chile. The PHANGS–MUSE survey (PI: Schinnerer) and the data reduction are described in Emsellem et al. (2022). They produced reduced and mosaicked spectral cubes and additional high-level data products like emission line maps. These maps achieve an average resolution of 0.72 arcsec, corresponding to a spatial resolution between 18.4 and 77.9 pc, depending on the distance to the galaxy.

Santoro et al. (2022) and Groves et al. (2023) utilised these products to create a catalogue of H II regions, which we use in this work and describe briefly below. HIIPHOT (Thilker, Braun & Walterbos 2000) was used on the $\text{H}\alpha$ line maps to define the boundaries of the H II regions based on their surface brightness profile. The integrated spectrum of each region was extracted from the spectral cubes and a number of nebular emission lines were fitted. We do not correct for the contribution of the DIG, as it is very sensitive to the exact H II region boundaries and the majority of our H II regions are bright ($L(\text{H}\alpha) > 10^{37} \text{ erg s}^{-1}$ and $\text{H}\alpha$ surface brightness $> 10^{39} \text{ erg s}^{-1} \text{ kpc}^{-2}$). The fluxes were then corrected for Milky Way extinction with the extinction curve from O’Donnell (1994), adapting $R_V = 3.1$ and $E(B - V)$ from Schlafly & Finkbeiner (2011). For the internal extinction, the colour excess $E(B - V)$ is measured from the *Balmer decrement*, assuming a theoretical ratio of $\text{H}\alpha/\text{H}\beta = 2.86$, and all fluxes are corrected assuming the same extinction curve parameters.

To remove contaminants such as SNRs or PNe, the *Baldwin–Phillips–Terlevich* diagram (Baldwin et al. 1981) is used to classify the objects (based on the demarcation lines by Kewley et al. 2001a; Kauffmann et al. 2003, using the [NII], [SII] and [OI] lines). For each line used, we require a $\text{S/N} > 5$ to be included in the final catalogue. This conservative S/N cut accounts for the possibility that the line flux uncertainties are slightly underestimated (Emsellem et al. 2022). Across 19 galaxies, a total of 31 497 nebulae are detected and 23 736 of them are classified as H II regions.

Next, a number of physical properties are derived. We calculate $\text{EW}(\text{H}\alpha)$ following the procedure outlined in Westfall et al. (2019). The $\text{H}\alpha$ and continuum flux are obtained by directly summing the spectral bins in the data cube. Our H II regions are observed in projection against the stellar disk, which is dominated by light from older stars that are not physically associated with the massive star-forming regions. To correct for this contribution to our $\text{EW}(\text{H}\alpha)$ measurement, we estimate the background contributed by the stellar continuum from a spectrum integrated over an annular mask. These masks

Table 7.1: Galaxy sample with number of HII regions (from Santoro et al. 2022; Groves et al. 2023) and stellar associations (from Larson et al. 2023). This only includes HII regions and associations that are inside the overlapping FoV of *HST* and MUSE. N_{match} is the number of objects in the one-to-one sample (see Section 7.3).

Name	Distance ^a Mpc	Resolution ^b pc	N_{HII}	N_{asc}	N_{match}
IC 5332	9.01 ± 0.41	31.4	608	397	135
NGC 0628	9.84 ± 0.63	34.7	1651	1141	379
NGC 1087	15.85 ± 2.24	56.9	892	487	181
NGC 1300	18.99 ± 2.85	57.9	1147	401	179
NGC 1365	19.57 ± 0.78	77.9	445	489	90
NGC 1385	17.22 ± 2.58	40.9	922	525	131
NGC 1433	18.63 ± 1.86	58.7	729	494	223
NGC 1512	18.83 ± 1.88	72.9	479	521	197
NGC 1566	17.69 ± 2.00	54.9	1448	1582	297
NGC 1672	19.40 ± 2.91	67.7	1047	1247	237
NGC 2835	12.22 ± 0.94	50.5	777	649	192
NGC 3351	9.96 ± 0.33	35.7	769	708	249
NGC 3627	11.32 ± 0.48	42.3	1016	1325	183
NGC 4254	13.10 ± 2.80	36.9	2375	1661	390
NGC 4303	16.99 ± 3.04	47.8	1956	1736	342
NGC 4321	15.21 ± 0.49	47.2	984	1483	332
NGC 4535	15.77 ± 0.37	33.7	1168	469	175
NGC 5068	5.20 ± 0.21	18.4	1405	715	174
NGC 7496	18.72 ± 2.81	71.6	523	265	91
Total			20341	16295	4177

^a From Anand et al. (2021).

^b Based on the average value of all MUSE pointings in that galaxy.

are created by growing the existing nebula masks until they have increased by a factor of three in area, excluding any pixels that fall inside of other nebulae. We then calculate the background corrected equivalent width, $EW(H\alpha)_{\text{corr}}$, by subtracting the background from continuum (but not from the $H\alpha$ line flux, see Appendix B.2 for more details).

The gas phase abundance $12 + \log(O/H)$ is measured with the S -calibration from Pilyugin & Grebel (2016) and the radial abundance gradient is fitted and subtracted from the individual nebulae to get the local metallicity offset $\Delta(O/H)$ (Kreckel et al. 2019). The ionization parameter $q = U/c$ is derived from the $[SIII]/[SII]$ ratio, based on the calibration by Diaz et al. (1991):

$$\log U = -1.684 \cdot \log([SIII]/[SII]) - 2.986, \quad (7.3)$$

where $[SIII]/[SII] = [SIII]\lambda(9069 + 9532)/[SII]\lambda(6717 + 6731)$. We measure $[SIII]\lambda 9069$ directly and assume a fixed atomic ratio of $[SIII]\lambda 9532 = 2.5 \cdot [SIII]\lambda 9069$ (Osterbrock & Ferland 2006).

Finally we use PYNEB (Luridiana et al. 2015) to derive the electron density from the $[SII]\lambda 6731/6717$ ratio. Ideally one would do this in conjunction with the temperature (e.g. from the $[NII]\lambda 5755/6548$ ratio). However, only ~ 4 per cent of the HII regions have a strong enough detection ($S/N > 10$) in the $[NII]\lambda 5755$ line to measure a temperature. We therefore assume a constant temperature of 8000K for all nebulae. This is close to the mean temperature that we measure for the sub-sample that has a detection in the $[NII]\lambda 5755$ line (Kreckel et al. 2022). That being said, the choice of T_e has only minor effects on the derived density (less than 10 percent, see also Appendix B.3 for a more careful analysis). More important is that most of the HII regions are close to the low-density limit. We apply the procedure outlined in Barnes et al. (2021) and do not report densities for the regions that have $[SII]\lambda 6731/6717$ ratios within 3σ of this limit, as we are unable to derive reliable density measurements for them.

7.2.2 HST Stellar association catalogue

The resolution of ground-based observations (~ 1 arcsec, corresponding to an average spatial resolution of ~ 70 pc for the galaxies in our sample) is not sufficient to resolve the star clusters and associations that are the origin of the ionizing radiation and only the *HST* is able to resolve the required ~ 10 pc spatial scales (Ryon et al. 2017) in the optical and UV. The PHANGS-*HST* survey (PI: Lee) covers 38 galaxies, including all galaxies in the PHANGS-MUSE sample, and is described in Lee et al. (2022). The galaxies were observed in 5 bands: F275W (NUV), F336W (*U*), F438W (*B*), F555W (*V*), F814W (*I*).

A number of high-level data products were produced by the PHANGS-*HST* team. First is a catalogue of compact star clusters (Thilker et al. 2022, we use IR₃ in this work). They are identified as sources that are slight more extended than a point source, based on their concentration index. However, the overlap with the HII region catalogue is rather limited (as expected, as most clusters have ages > 10 Myr) and we find that the majority of the ionizing sources are in OB associations (Goddard, Bastian & Kennicutt 2010; Kruijssen 2012; Adamo et al. 2015; Adamo et al. 2020). These larger and loosely

bound structures were identified by Larson et al. (2023) at different spatial scales and we use their multi-scale stellar association catalogue for our analysis. A summary of the procedure is as follows, and also provided in Lee et al. (2022). First, DOLPHOT is used to create a catalogue of point-like sources (using either the NUV or V-band). Then a »tracer image« is generated by smoothing to a fixed resolution (either 8 pc, 16 pc, 32 pc or 64 pc) and a local background correction is applied to each by subtracting the equivalent image generated at four times larger scale. Next a *watershed algorithm* (Beucher & Lantuéjoul 1979; van der Walt et al. 2014) is used to define the boundaries of the stellar associations that are identified at each scale as local enhancements in the background-subtracted tracer image. Fluxes for the five available *HST* filters are measured, using the photometry of detected stars, contained within the association. Age, mass and reddening are derived by fitting theoretical models of a single stellar population (Bruzual & Charlot 2003) to the observed SED with CIGALE (Boquien et al. 2019), assuming a fully sampled Chabrier (2003) IMF. The gas-phase metallicities are close to solar for most galaxies (Groves et al. 2023), suggesting that the assumption of solar metallicity is reasonable for the young stellar populations probed here. Complete details on the SED fitting are provided in Turner et al. 2021.

For this work, we choose to use the 32 pc stellar associations identified from the NUV images. This catalogue contains 16 295 associations that fall within the MUSE FoV. This was done as the NUV is better at tracing the young massive stars that we are interested in, and the 32 pc structures are well matched to our MUSE resolution. While we find a larger number of HII regions that overlap with the 64 pc associations, the percentage with a one-to-one relation is smaller, especially for the more nearby galaxies. It should be noted that using a different scale for the analysis does not impact the results significantly.

7.2.3 *AstroSat*

We used the Indian space telescope *AstroSat* (Singh et al. 2014) to observe our galaxy sample in the FUV (PI: Rosolowsky). The sample and data reduction are described in Hassani et al. (2024). The images have a resolution of 1.4 arcsec with a pixel size of 0.42 arcsec, reasonably well matched to our MUSE resolution. Observations are available for 16 of the 19 MUSE galaxies (NGC 1087 and NGC 4303 can not be observed and for NGC 5068, observing time is awarded but not yet executed). When available, we use the F148W filter, which is the case for all galaxies but NGC 1433, NGC 1512 and NGC 4321, which were observed with the F154W filter. We subtract a uniform foreground that originates from Milky Way and solar system sources and whose average contribution varies between galaxies from 1 to 9 percent. The *AstroSat* images are then reprojected to the MUSE images to measure the FUV fluxes in the spatial masks of the nebula catalogue. The fluxes are corrected for Milky Way and internal extinction, following the procedure described for the HII regions. For the colour excess we use the value derived from the Balmer decrement and assume $E(B - V)_{\text{stellar}} = X \cdot E(B - V)_{\text{Balmer}}$. It is common to use $X = 0.44$ to account for the difference between the colour excess derived from the Balmer decrement and the stellar continuum (Calzetti et al. 2000), reflecting that nebular sources are often more closely associated with dusty star-forming complexes in contrast to the

bulk stellar light (Charlot & Fall 2000). However, we are preferentially targeting stellar continuum emission that we believe to be co-spatial with the nebular emission, and so instead assume $X = 1.0$. In Section 7.3.3 we compare the colour excess derived from the Balmer decrement to the one from the SED fit and find good agreement between the two, validating this choice.

7.3 MATCHING NEBULAE AND STELLAR ASSOCIATIONS

To relate the H II regions to the source of the ionizing radiation, we match the nebula catalogue with the association catalogue. Both have spatial masks, but with different pixel scales (0.2 arcsec per pixel for MUSE and 0.04 arcsec per pixel for *HST*). The coarser resolution of the nebula masks can result in some larger H II regions that are not correctly partitioned (see Barnes et al. 2022). However we would still correctly map them to their ionizing sources. The absolute astrometric accuracy is excellent for both data sets. *HST* achieves 0.1 pixel accuracy while MUSE reaches 0.5 pixel, corresponding to two and a half *HST* pixels. Considering the uncertainties in deriving the exact boundaries of the H II regions, this is negligible when matching the data sets. We reproject the nebula masks to the association masks and determine the fractional overlap as measured in terms of the association or nebula pixel areas. In Figure 7.1 we show examples for the overlap between the H II regions and associations. For each of the stellar associations we define the overlap to be one of the following three:

- isolated: the association does not overlap with any nebula (fractional overlap of 0 percent).
- partial: part of the association overlaps with one or more nebula, but part of it extends beyond it (fractional overlap between 0 and 100 percent).
- contained: the association is fully contained in the nebulae. It can be contained within multiple nebulae (fractional overlap of 100 percent).

For the H II regions we determine the following properties:

- neighbors: number of neighbouring H II regions (i.e. regions that share a common boundary)
- Nassoc: number of stellar associations that overlap with the H II region.

Finally, we construct a catalogue of matched objects where each H II region and stellar association overlap with exactly one stellar association and H II region respectively. It contains 4177 objects that constitute the parent sample that we will be working with and to which we refer to as the one-to-one sample. This catalogue is available in the supplementary online material and incorporates a number of columns from the nebula and association catalogues. Table 7.2 gives an overview of the properties that are included.

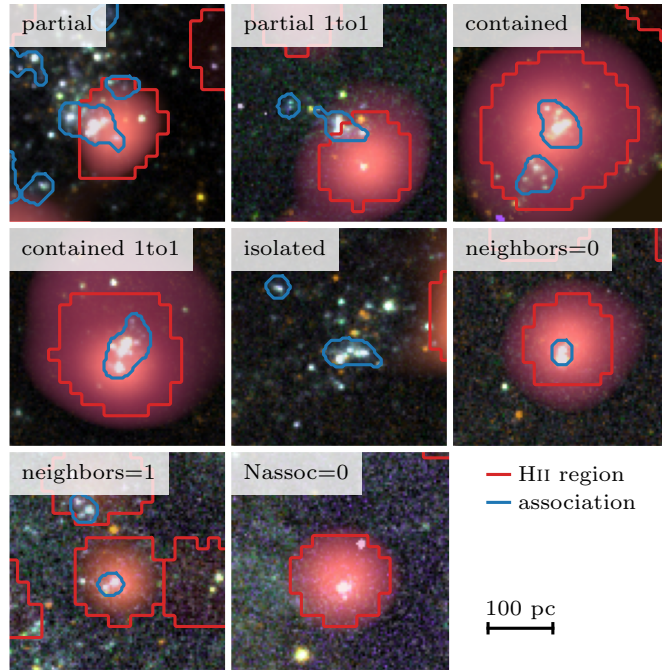


Figure 7.1: Examples for the overlap between the H II regions and stellar associations in NGC 1365. The cutouts show a three colour composite images, based on the 5 available *HST* bands, overlaid with the $H\alpha$ line emission of MUSE in red. The boundaries of the H II regions are shown in red and the stellar associations in blue. Flags that characterise the overlap between the two catalogues are showcased in this figure.

Table 7.2: Columns of the matched catalogue (one-to-one sample).

Column	Description
gal_name	Name of the Galaxy
region_ID	ID of the H II region
assoc_ID	ID of the stellar association
ra_neb	Right ascension of the nebulae
dec_neb	Declination of the nebulae
ra_asc	Right ascension of the stellar association
dec_asc	Declination of the stellar association
overlap_neb	Overlap percentage with stellar association
overlap_asc	Overlap percentage with H II region
overlap	Flag for overlap (isolated, partial, contained)
environment	Galactic environment (e.g. bar,centre,disc)
neighbors	Number of neighbouring H II regions
HA6562_lum [†]	Extinction corrected H α λ 6562 luminosity
EBV_balmer [†]	Colour excess from the Balmer decrement
EBV_stellar [†]	Colour excess from the SED fit
age [†]	Age of the stellar association in Myr
mass [†]	Mass of the stellar association in M_{\odot}
{filter}_flux ^{†‡}	Flux in <i>HST</i> band in mJy
Ha/FUV [†]	H α to FUV flux ratio (extinction corrected)
EW_HA [†]	H α equivalent width in Å
EW_HA_corr [†]	Background corrected EW(H α) in Å
logq [†]	Ionization parameter $\log q$
Delta_met_scal	Local metallicity offset $\Delta(\text{O}/\text{H})$
density [†]	Electron density in cm^{-3}
GMC_sep	Distance to nearest GMC in pc

[†] associated errors are included as *_err.

[‡] for filter in *NUV*, *U*, *B*, *V* and *I*.

Table 7.3: Level of spatial correlation between H II regions and stellar associations. We list the total number and percentage of H II regions (associations) that overlap with N associations (H II regions) respectively.

N	H II regions		Stellar associations	
0	11 687	(57.5 %)	4596	(28.2 %)
1	5673	(27.9 %)	9288	(57.0 %)
2	1702	(8.4 %)	2011	(12.3 %)
3	643	(3.2 %)	319	(2.0 %)
4	291	(1.4 %)	57	(0.3 %)
>5	345	(1.7 %)	24	(0.1 %)

7.3.1 Statistics of the matched catalogue

From the initial nebula catalogue, 20 341 H II regions fall inside the *HST* FoV and 8654 of them (42.5 per cent) overlap with at least one stellar association. Conversely, the association catalogue contains 16 295 associations in the MUSE FoV and 11 699 of them (71.8 per cent) overlap with at least one H II region. Table 7.3 provides a detailed breakdown of how the H II regions and associations overlap.

In cases where a nebula contains multiple associations, it is difficult to assess the contribution of the individual objects and ambiguous to assign a single age, as they are not necessarily coeval (Efremov & Elmegreen 1998). In fact, when we find multiple stellar associations in one H II region, they rarely have the same age. The age spread is usually small (< 10 Myr), but there are also more extreme cases, with ~ 25 per cent having larger age spreads. Therefore, to simplify our analysis, in this paper we require a one-to-one match (this can be with either partial or contained overlap) between the associations and H II regions, leaving us with a sample of 4177 objects in our parent one-to-one sample (see Table 7.1 for a detail breakdown by galaxy). Depending on the application, we also apply further cuts in age, mass or overlap to obtain cleaner samples.

The nebula masks are dense and cover most of the spiral arms and hence it is likely that some of the overlaps are coincidental. To estimate how many of our matches this might concern, we run a test where we rotate the full set of nebula masks by 90 deg around the galaxy centre, before matching with the associations. While previously one third of the H II regions overlapped with an association, this number drops to one sixth. The number of partial overlapped associations decreases only slightly to 81 per cent. The difference is more apparent with the contained sample: we estimate that only around 16 per cent of the contained associations could be chance alignments. This motivates us to only use the contained objects for our analysis, reducing the sample to 1918 H II regions and stellar associations.

Another concern is stochastic sampling of the IMF (Fouesneau & Lançon 2010; Hannon et al. 2019). Below a certain cluster mass, the presence or absence of a single massive star can have a significant impact on the observed colours and hence the derived properties of

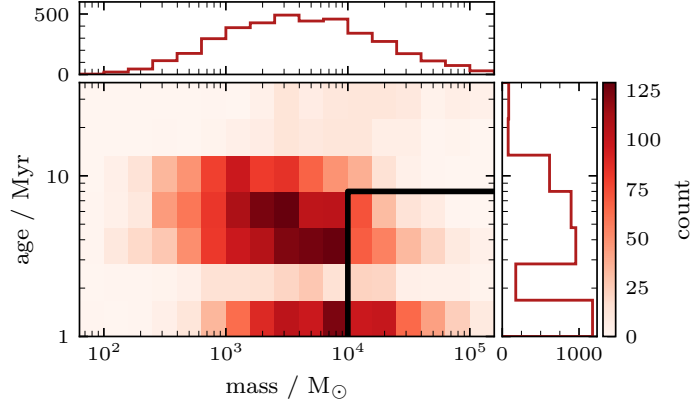


Figure 7.2: Distribution of masses and ages of the stellar associations in the matched catalogue (the one-to-one sample). The black lines mark the cuts that we apply to the sample to ensure a fully sampled IMF (more massive than $> 10^4 M_{\odot}$) and to only include young clusters that should be associated with ionized gas (younger than ≤ 8 Myr). The mass cut leaves us with a sample of 1014 objects and the age cut with 3531. Applying both cuts results in a sample of 756 objects.

the association, as well as on whether it is able to ionize hydrogen and has an associated detectable H II region. We assume that stellar associations that are more massive than $10^4 M_{\odot}$ will fully sample the IMF (da Silva et al. 2012). We also do not expect significant H α emission to be associated with old stellar populations and therefore introduce an age cut at 8 Myr. The distribution of masses and ages of the matched catalogue is shown in Figure 7.2, with both cuts illustrated by black lines. Note that the gap at 2 to 3 Myr is inherited from the stellar association catalogue, and reflects specific features in the stellar population tracks (Larson et al. 2023). In total, 1041 objects pass the mass cut and 3531 pass the age cut. Applying both of these criteria, and further requiring the stellar associations to be contained, leaves 469 objects, constituting our robust sample.

7.3.2 The physical nature of H II regions without associations

As shown in Table 7.3, the majority of our H II regions do not overlap with an association, which raises the question of the origin of their ionizing radiation. We consider several scenarios to explain this discrepancy. First of all, the unmatched H II regions could host deeply embedded, highly extincted stars that we are unable to observe. However their reddening distribution is almost identical to the matched H II regions with a mean of $E(B - V)_{\text{Balmer}} = 0.26$, making it unlikely that we miss many objects due to extinction.

Secondly, we restricted our analysis to the 32 pc scale associations, but there are some variations between the different scales and 20.6 per cent of the unmatched H II regions overlap with an 16 pc or 64 pc association. By design, the stellar association catalogue only includes objects with multiple peaks, excluding other ionizing sources like compact

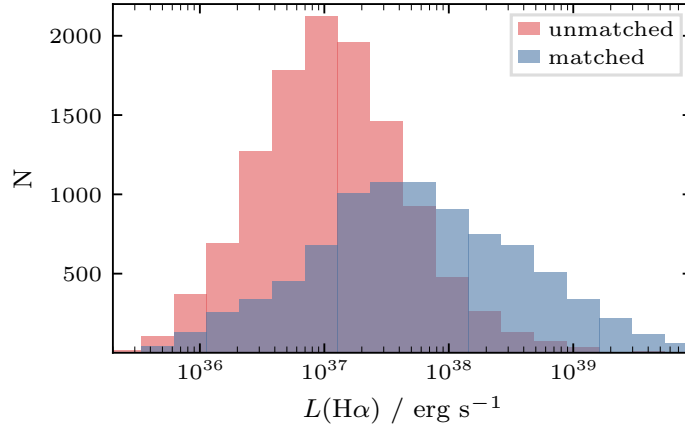


Figure 7.3: $H\alpha$ luminosity function of the H II regions that overlap with one or more associations and those that do not. The matched H II regions are on average a factor 10 brighter than the unmatched ones.

star clusters or individual stars. Looking at the $H\alpha$ luminosity function in Figure 7.3, we see that H II regions that are matched to an association are on average brighter by a factor of 10 compared to those that are unmatched, meaning that the unmatched regions are ionized by fainter sources. We consider the previously mentioned PHANGS-*HST* compact cluster by (Thilker et al. 2022, for example, the unmatched H II region in Figure 7.1 contains an object that is classified as a compact cluster). Using the machine learning classified catalogue (Wei et al. 2020; Hannon et al. 2023), we find that 2569 of all H II regions contain a compact cluster (15.3 per cent, comparable to the 13 per cent reported by Della Bruna et al. 2022a for M 83). The majority of these clusters are already contained in an association and only 6.1 per cent of the previously unmatched H II regions contain a compact cluster. By design, the compact cluster identification pipeline disfavours the selection of looser association-like structures that are young. Considering that we are primarily interested in young and massive clusters, the number of potentially interesting compact clusters is only of the order of a few dozen.

Even after accounting for different scale associations and compact clusters, 43.5 per cent of the H II regions are still without a catalogued ionizing source. However, looking at the *HST* NUV images, we find peaks that are clearly associated with those unmatched H II regions. Their absence in the aforementioned catalogues can have a few reasons. First, the selection of cluster candidates is based on the V band and hence does not necessarily trace the ionizing sources as well as the NUV. Secondly, they are often too compact to be classified as a star cluster according to their concentration index. As for the association catalogue, multiple peaks are required and the objects here are mostly single peaks. Using the point source catalogue, created with DOLPHOT, that forms the base for both the stellar association and compact cluster catalogues (Thilker et al. 2022), we find peaks in most

H II regions. If we require that a peak is detected with a $S/N > 5$ in both the NUV and at least one other filter, around 63.7 percent of the unmatched H II regions contain a peak.

The fainter unmatched H II regions ($L_{H\alpha} \lesssim 10^{37} \text{ erg s}^{-1}$) fall in a regime where they could be ionized by a single O-star (Martins, Schaerer & Hillier 2005). The apparent magnitudes of the DOLPHOT sources inside those unmatched H II regions are also comparable to the luminosity of a single O-star (see also Appendix B.4 with Figure B.3 for a comparison with models). This subsample of H II regions that are potentially ionized by a single star (or a very small star cluster, dominated by a single massive star) is interesting in its own right, however beyond the scope of this paper.

For now, we leave this with the understanding that most of the unmatched H II regions in our catalogue have an ionizing source that is visible in the *HST* data, but that these sources do not meet the criteria for our catalogue of stellar associations and hence are not included in our analysis. Overall, we are able to identify NUV-bright ionizing sources for the vast majority of our sample. After accounting for the different sources of ionizing radiation, only 18.6 percent of the H II regions are left without an ionizing counterpart. This number could be even further decreased by lowering the *signal-to-noise* (S/N) requirement for the DOLPHOT peaks (to 9.9 percent with $S/N > 3$).

7.3.3 Validating the sample

We validate our cross-matched catalogue by comparing physical properties between the nebulae and stellar associations. If the associations are the origin of the ionizing radiation, the $H\alpha$ luminosity of the H II region is expected to be related to the mass of the association. As shown in Figure 7.4, we find a strong correlation between the two. For this comparison we utilise the sub-sample of fully contained associations (1918 objects). There is some scatter, as the $H\alpha$ flux is also dependent on the age of the stellar association and the amount of radiation that escapes the cloud.

We also compare the dust content, as traced by the colour excess $E(B-V)$, derived from the SED fit to the one calculated from the Balmer decrement. We find that for young ages, the two methods are in good agreement, while for older associations, the reddening from the Balmer decrement is systematically higher than that from the SED fit. This could be due to problems with the SED fit, where young reddened clusters are identified by the fitting routine as old and un-reddened (Hannon et al. 2022). The age of some associations ($> 8 \text{ Myr}$) raises some doubts if they are the origin of the ionizing radiation. For most of those objects, the reddening derived from the Balmer decrement is higher than the one derived from the SED fit, suggesting that the age/reddening degeneracy is not correctly resolved and we therefore exclude those objects from further analysis. Alternatively, for young clusters the gas and the stars are closely associated, while they decouple over time (Charlot & Fall 2000), which would also result in lower reddening of the older clusters relative to the nebular emission. Future inclusion of *JWST* bands and refined age determinations will help us distinguish these scenarios (Whitmore et al. 2023).

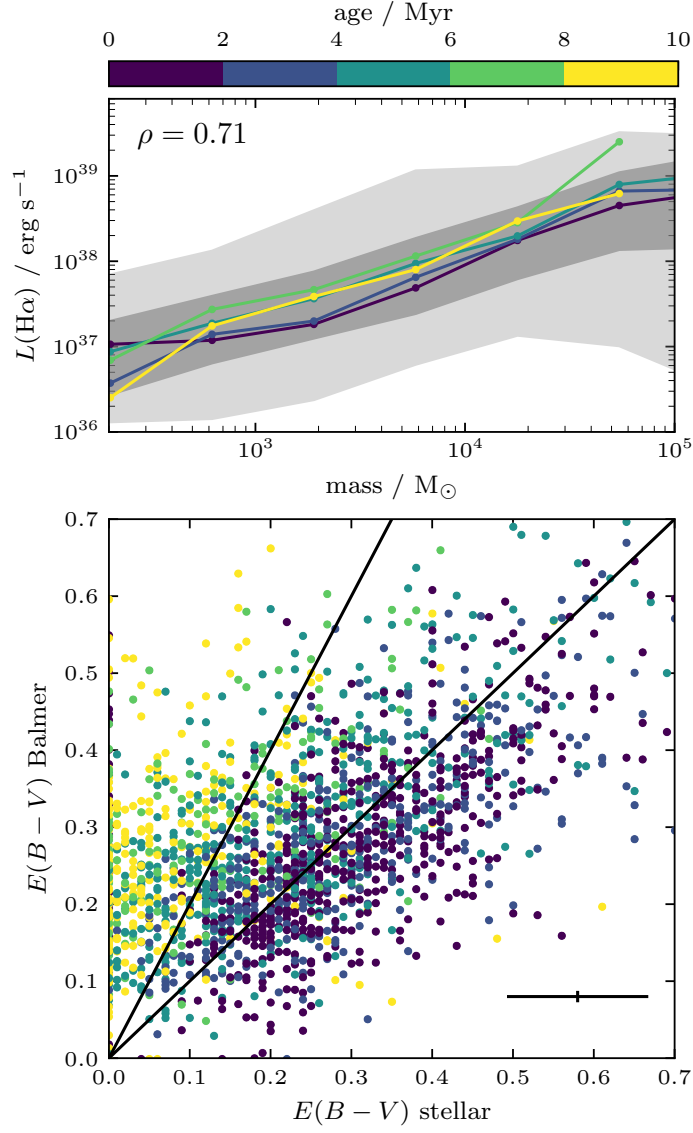


Figure 7.4: Top: comparison between stellar association mass and extinction corrected H α luminosity. The 68 and 98 percentile are shaded in grey. Bottom: Comparison between the stellar $E(B-V)$ derived from the SED fit and the one from the Balmer decrement. The black lines indicate one-to-one and 0.44 to 1 correlation (Calzetti et al. 2000). The average uncertainties are indicated in the bottom right corner. Both panels use the contained objects in the one-to-one sample.

Given the above discussion, we are confident that the stellar associations are clearly linked to the nebulae, enabling us to investigate more deeply evolutionary trends with stellar age.

7.4 H II REGION EVOLUTIONARY SEQUENCE

In this section we use the matched catalogue to establish different nebular properties as proxies for the age and then explore how other H II region properties evolve as the nebula ages.

7.4.1 Correlations with SED ages

In Section 7.1, we introduced $\text{EW}(\text{H}\alpha)$, $\text{H}\alpha/\text{FUV}$ and $\log q$ as potential age tracers. With the matched catalogue, we can assess how well they trace the age of the stellar population that is powering the nebula. Because accurate ages are of utmost importance to this application, we only use the robust sample and also apply further signal-to-noise cuts. For the $\text{H}\alpha/\text{FUV}$ and $\log q$ we only use objects with $\text{S/N} \geq 5$ and for the ages we require that they are at least as large as their uncertainties. This mainly affects $\text{H}\alpha/\text{FUV}$ and choice of a higher S/N would reduce the sample too much. For the $\text{EW}(\text{H}\alpha)$, the largest uncertainty comes from the continuum, which is often indistinguishable from the background. We therefore compute $\text{EW}(\text{H}\alpha)_{\text{corr}}$ and require that the background subtracted continuum has $\text{S/N} \geq 5$ (this cut is used for the corrected and un-corrected equivalent width). In Figure 7.5 we compare the age from the SED fit to all four properties. We observe weak to moderate anti-correlations between the SED ages and all proposed age tracers. To test how robust the trends are, we apply a Monte Carlo approach, where we repeatedly sample our data, based on the associated uncertainties and assuming a normal distribution. The observed trends are robust and persist, with a scatter around 10 per cent, except for the $\text{H}\alpha/\text{FUV}$, which shows a larger scatter of 20 per cent.

The $\text{EW}(\text{H}\alpha)$ drops by 47 per cent within the first 4 Myr. This is caused by the death of the most massive stars and in line with the 42 per cent decrease that is predicted by models like STARBURST99 (Leitherer et al. 2014). However, after that, the observed $\text{EW}(\text{H}\alpha)$ stays roughly constant and even increases slightly towards 8 Myr, compared to the models that predict further decrease to only 6 per cent at the same age. Also conspicuous is the smaller range of values covered by the observations, compared to the models (see Appendix B.1). An easy explanation for this is that we neglected the background. For the stellar continuum around $\text{H}\alpha$, we observe a significant contribution from older stellar populations throughout the galaxy. Accounting for this, the $\text{EW}(\text{H}\alpha)_{\text{corr}}$ increases by a factor of 10, bringing the measured values closer to the model predictions. The correlation with the background corrected equivalent width $\text{EW}(\text{H}\alpha)_{\text{corr}}$ is slightly stronger than the un-corrected one. However it is quite challenging to correctly disentangle the contribution of the background.

The flux ratio of $\text{H}\alpha$ to FUV stays roughly constant for the first 4 Myr, after which it drops to ~ 65 per cent. Again, after that, contrary to the model predictions, the ratio does

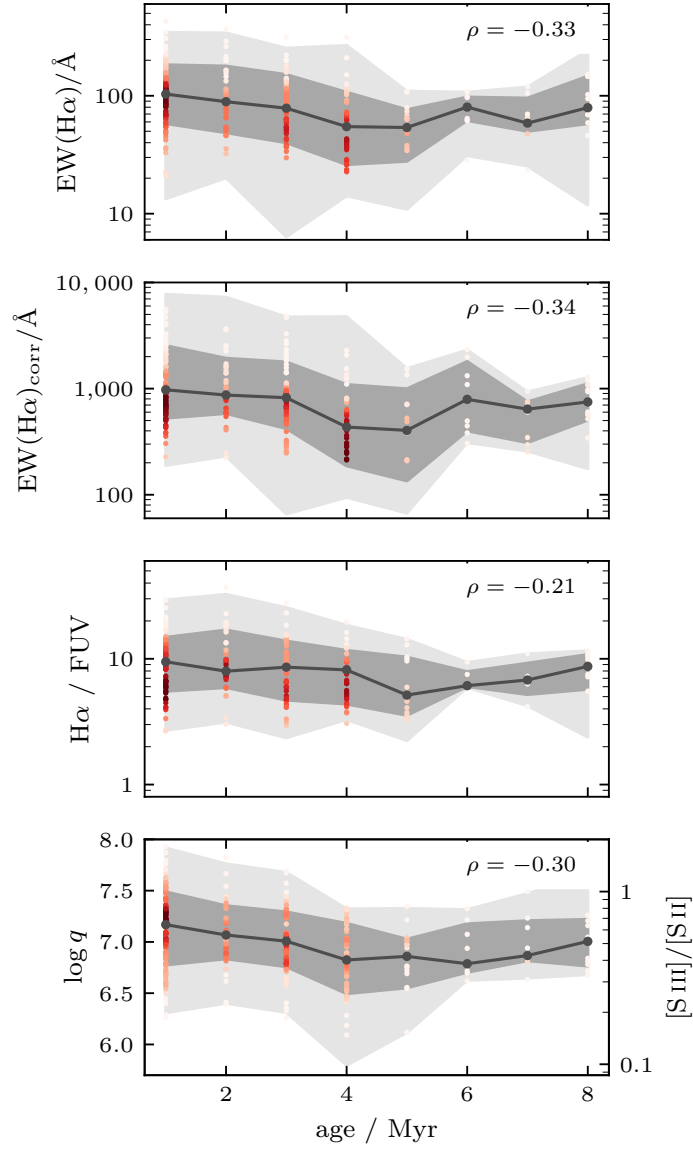


Figure 7.5: Comparison between stellar associations ages with MUSE H II region properties. The age is derived from the SED fit and compared to $\text{EW}(\text{H}\alpha)$, $\text{EW}(\text{H}\alpha)_{\text{corr}}$ (the continuum is corrected for background emission), $\text{H}\alpha/\text{FUV}$ and $\log q$. The binned median is shown with a black line and the 68 and 98 percentile are shaded in grey. The colour of the points indicates the density of the points. The Spearman's rank correlation coefficient is shown in the top right corner. The $\text{EW}(\text{H}\alpha)$, $\text{EW}(\text{H}\alpha)_{\text{corr}}$ and $\log q$ have $p\text{-value} < 10^{-7}$ and the $\text{H}\alpha/\text{FUV}$ $p\text{-value} < 10^{-2}$.

not further decrease. Similar to the equivalent width, the observed values are smaller than the model predictions, although the difference is not quite as large. Applying the same reasoning as with the equivalent width to resolve this discrepancy does not work. In the case of the FUV, we do not observe a global background, but rather other events of previous nearby star formation, which should not bias our measurements.

Lastly, we look at the ionization parameter. For reference we also show $[\text{SIII}]/[\text{SII}]$ based on Equation 7.3. We observe a similar behaviour with an initial drop and a flattening after 4 Myr, corresponding to the decrease in ionizing photon flux and expansion of the nebula. However it should be noted that the range of ionization parameters that we observe only covers the first few Myr in the models (this could be due to systematic offsets between observations and model grids as highlighted in Mingozi et al. 2020).

Given that the uncertainties on the ages are often similar to the ages themselves, we consider a broader statistical analysis by binning our sample by the fiducial age tracers and looking at the median SED ages. We find that associations that overlap with an HII region are significantly younger than those that are isolated. When binning our sample based on the $\text{EW}(\text{H}\alpha)$ we find that the associations with high $\text{EW}(\text{H}\alpha)$ are younger than their counterparts with low $\text{EW}(\text{H}\alpha)$ (see Appendix B.5 for more details). A similar but less pronounced trend can also be observed when binning with the $\text{H}\alpha/\text{FUV}$ and $\log q$.

Even though the observed decrease with SED ages are weaker than expected, $\text{EW}(\text{H}\alpha)$, $\text{EW}(\text{H}\alpha)_{\text{corr}}$, $\text{H}\alpha/\text{FUV}$ and $\log q$ are all consistent in their evolution. In Figure 7.6 we use the full HII region sample of $\sim 20\,000$ HII regions (applying the same signal-to-noise cuts as to the previous sample) to compare the proposed age tracers with each other. Across the entire galaxy sample, spanning a wide range of stellar masses, SFRs and metallicities, we find moderate to strong correlations between all properties, consistent with the idea of an evolutionary sequence.

7.4.2 Age trends in the nebular catalogue

In the previous section, we showed that $\text{EW}(\text{H}\alpha)$, $\text{H}\alpha/\text{FUV}$ and $\log q$ all exhibit anti-correlations with age, albeit weaker than expected. However, for this analysis we were limited to the few hundred HII regions that contained a massive stellar association, severely limiting our ability to study the evolution of the HII regions in terms of different ISM properties. In this section we use the full catalogue of 23 736 HII regions and look for correlations with the proposed age tracers. Figure 7.7 shows trends between $\text{EW}(\text{H}\alpha)$, $\text{EW}(\text{H}\alpha)_{\text{corr}}$, $\text{H}\alpha/\text{FUV}$ and $\log q$ with various nebula properties. All trends have p -values $< 10^{-6}$, with the exception of $\text{H}\alpha/\text{FUV}$ versus δ_{GMC} . Following the same procedure as in Section 7.4.1, we find that the trends are robust when varying the properties within their uncertainties.

First, the metallicity variations $\Delta(\text{O}/\text{H})$ are the difference between the local metallicity and the global metallicity gradient. We recover the same correlation between $\log q$ and the $\Delta(\text{O}/\text{H})$ that were also found by Kreckel et al. (2019, based on the same data) and Grasha et al. (2022). The metallicity offset also shows a correlation with $\text{EW}(\text{H}\alpha)$ and $\text{H}\alpha/\text{FUV}$, indicating that younger regions are more enriched. This suggests that the natal environment, not yet dispersed and associated with the youngest regions, is more metal

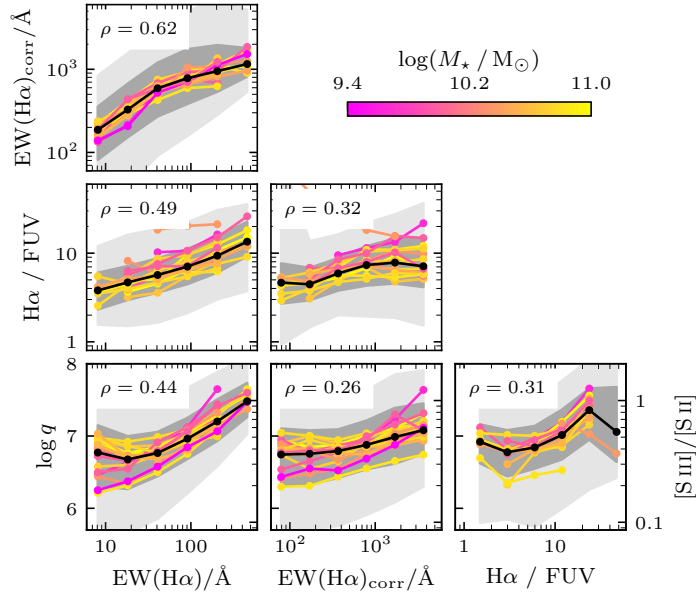


Figure 7.6: Comparison between the proposed age tracers for the full H II region catalogue. The nebulae are grouped by their host galaxy, sorted by stellar mass M_* , with the median of the entire sample indicated by a black line. The 68 and 98 percentile ranges are shaded in grey.

rich. Over time, this material could mix on larger scales (Kreckel et al. 2020), decreasing the metallicity offset.

Next, we observe correlations with the $H\alpha$ luminosity of the H II regions with our age tracers. The strong correlation with $EW(H\alpha)$ is likely due to the significant contamination of the stellar continuum by (unrelated) older stellar populations. Due to this, the variations could be attributed to the smaller variations in both the FUV and the stellar continuum compared to $H\alpha$.

For the electron density of the ionized gas, the majority of our H II regions fall in the low-density limit, making it difficult to robustly measure a density. Only 1375 H II regions are significantly different from the low-density limit which limits our statistical capabilities for individual galaxies. We find no strong correlations with density, and in most tracers we observe weak anti-correlations, contradicting expectations that newly formed star clusters are still embedded in a dense cloud of gas. Focusing on the ionization parameter, we do recover a positive trend, consistent with the simplistic Strömgren sphere assumptions from Equation 7.2.

The extinction on the other hand, traced by the colour excess from the Balmer decrement, shows a correlation with $EW(H\alpha)$ and $\log q$, fitting the picture of a young embedded cluster. However, by contrast, we observe an anti-correlation with $H\alpha/FUV$. It should be noted that especially for the lower mass galaxies, the trends are very flat.

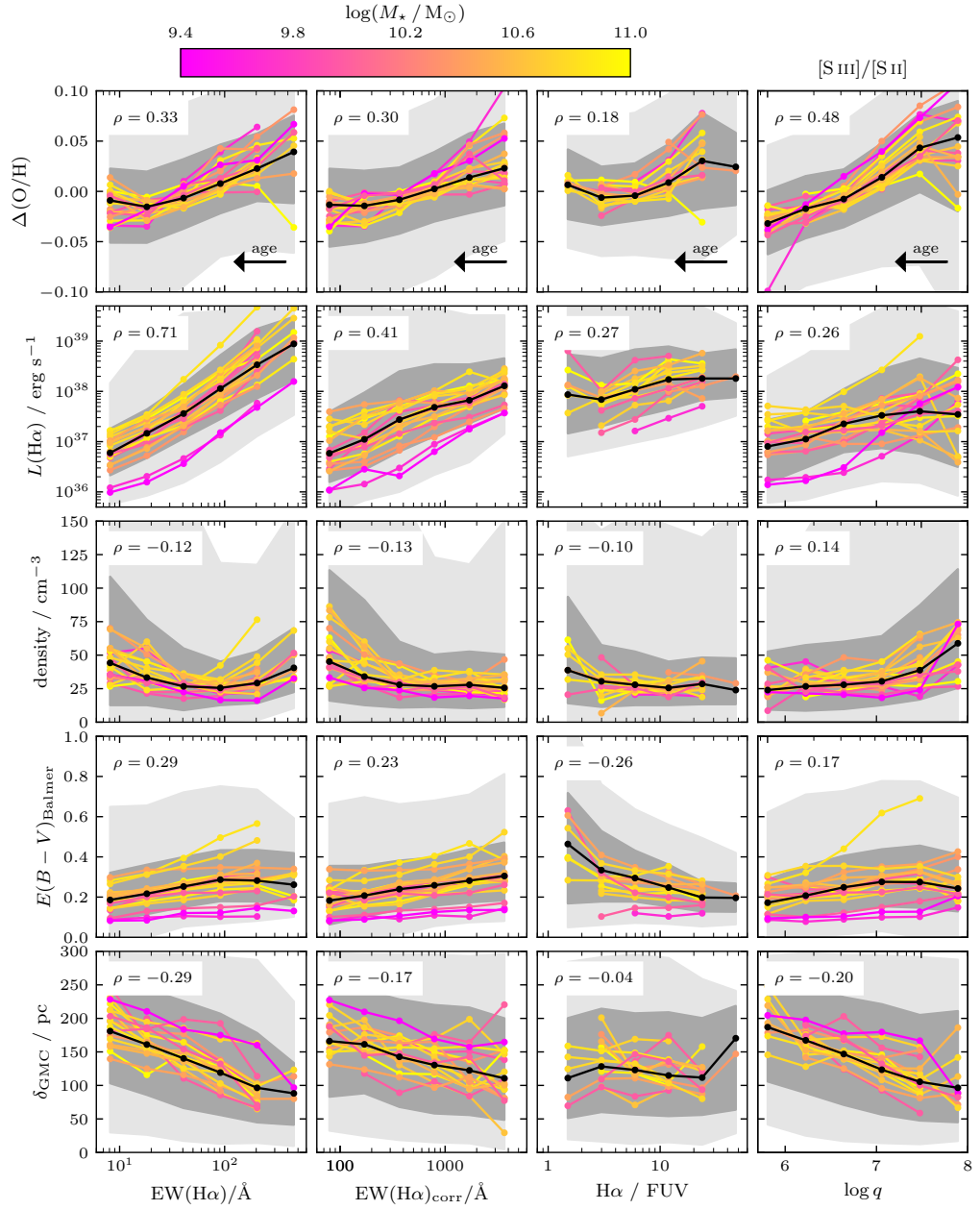


Figure 7.7: Correlations between the physical properties in the H II region catalogue and the proposed age tracers. We show trends with metallicity offset, $\text{H}\alpha$ luminosity, electron density, colour excess and separation to the nearest GMC. The nebulae are grouped by their host galaxies, sorted by their stellar mass M_{\star} , with the average of the entire sample indicated by a black line. The 68 and 98 percentile are shaded in grey.

Finally, we compare with the separation to the nearest GMC. For this we use the GMC catalogue from Rosolowsky et al. (2021) and Hughes et al. in preparation, based on the PHANGS–ALMA survey (Leroy et al. 2021b), and cross-match with the H II region catalogue. Given that the typical intercloud and interregion distances in these catalogues are a few hundred pc (e.g. Kim et al. 2022, Groves et al. 2023, Machado et al. in preparation) we consider separations of this order to be unphysical but closer clouds may indicate a real separation. Therefore, we exclude H II regions with a GMC separation larger than 300 pc from this analysis. We expect the youngest H II regions to still be closely associated to their birth environment and indeed we find an anti-correlation between separation and age.

For this investigation we presented both the observed and corrected $\text{EW}(\text{H}\alpha)$. Overall, these two measures show good correlation with each other (Figure 7.6), and similar qualitative trends in Figure 7.5. For most properties we find stronger correlations with the un-corrected $\text{EW}(\text{H}\alpha)$, which could hint at complications with the background subtraction.

7.4.3 Understanding the weak correlations with SED age

Although we do see some trends of $\text{EW}(\text{H}\alpha)$, $\text{H}\alpha/\text{FUV}$ and $\log q$ with age, they are weaker than predicted by models. There are a number of reasons why this could be the case.

First, looking at Figure 7.2, we find that roughly half of our sample falls in the youngest 1 Myr age bin. Given that our median reported uncertainty in the ages is 1 Myr, the age resolution of the SED fit is not sufficient to probe the evolution of these H II regions or (realistically) changes that happen within the first 1 to 2 Myr. Beyond model limitations, another major issue is that we are using only broad bands with limited UV coverage. This makes it very difficult to achieve higher precision at very young ages. However, the ~ 50 per cent drop in intensity of the nebular age indicators within the first 4 Myr is broadly consistent with models, and suggests that generally the expected age trends are recovered (see Section 7.4).

If we only consider the youngest objects (≤ 2 Myr), we still recover the same trends between $\log q$, $\text{EW}(\text{H}\alpha)$ and $\text{H}\alpha/\text{FUV}$. We note that we also observe the same trends in the older subsample (> 2 Myr), albeit somewhat weaker. This suggests that the evolution is not bound to a fixed or absolute timescale, but can vary greatly between clouds. In many ways, this is unsurprising, and depending on the environment, the H II regions can evolve differently. The timescales over which natal molecular gas clouds are cleared are estimated to be 2 to 3 Myr (Kim et al. 2021a; Chevance et al. 2022), suggesting that significant morphological changes in the local environment are occurring. Depending on the initial density of the cloud and the ionizing cluster mass, this can lead to large variations in the evolution timescale between different H II regions (Kruijssen et al. 2019). In this way, we would still expect to see trends between these nebular age tracers, as they are all reflecting the clearing of the cloud, but the absolute timescales (traced by the age of the underlying stellar cluster) would not necessarily agree across different clouds.

There are also clearly secondary dependencies beyond age, which may or may not significantly impact the trends we explore but are challenging to account for. For example,

it is likely that different stellar population models would give different results for the age and mass of the stellar associations, and this is not something that has been explored. In the case of the ionization parameter, variations will also arise due to changes in metallicity (Dopita et al. 2014; Kreckel et al. 2019; Grasha et al. 2022) and density (Dopita et al. 2006), and smear out the trend with age. However these relations, particularly with metallicity, are poorly understood from observations or modeling, and it is not even clear if we expect a correlation or anti-correlation (Ji & Yan 2022, and references therein).

One key assumption is that we are looking at nebular emission associated with an instantaneous burst of star formation. This assumption is not only important for the SED fit, but if the star formation is extended over a larger period of time, it will also be reflected in the observed flux ratios. The decrease of both $H\alpha/FUV$ and $EW(H\alpha)$ will appear to be slower and much less pronounced, due to the additional underlying stellar continuum emission from the pre-existing older population (Smith et al. 2002; Levesque & Leitherer 2013). Meanwhile, the $H\alpha$ flux and $\log q$ will largely trace more directly the evolution of the latest burst. The exact duration of star formation is actively discussed in the literature, and can vary from very short duration of less than 1 Myr (Povich & Whitney 2010) to multiple episodes of star formation across tens of Myr (Ramachandran et al. 2018). A prominent example here are the two clusters at the centre of 30 Doradus that are thought to have an age spread of a few Myr (Sabbi et al. 2012; Rahner et al. 2018). By studying only a sample where the HII region and stellar association are matched one-to-one, we have attempted to minimise the impact of blending of different stellar populations.

One aspect we have neglected is the impact of leaking radiation, which we know must occur as it is largely responsible for the kpc scale ionization of the DIG (Belfiore et al. 2022). Ionizing photons that escape the cloud result in a lower $H\alpha$ flux measured within the nebula, while the measurement of the FUV and stellar continuum is mostly unaffected. Existing studies measure a wide range of *escape fractions*, ranging from only a few per cent (Pellegrini et al. 2012) all the way up to 60 to 70 per cent (McLeod et al. 2019; Della Bruna et al. 2021). Because the fraction of leaking radiation is likely to change as the cloud ages and dissolves, this can introduce considerable scatter, depending on how much the escape fraction varies between clouds. However this process will only decrease the measured $EW(H\alpha)$ or $H\alpha/FUV$, and hence the true value should form an upper envelope above the data. Escape fractions within the range from 40 to 80 per cent would result in a factor of 5 scatter in $EW(H\alpha)$ at fixed age, roughly consistent with our result.

Overall, a combination of the effects listed above are probably responsible for the observed trends in age being less pronounced than expected. However, we emphasise that our observations are still consistent with an evolutionary sequence.

7.4.4 Recommendations on nebular age tracers

The equivalent width measurement has most commonly been used as an age tracer in the literature, and suffers from relatively few systematics or uncertainties due to extinction corrections and is easy to measure. However, the contribution of stellar light from the galaxy disk is problematic. The stellar continuum emission arising from old stars in the

disk is not physically associated with the nebulae, does not contribute to its ionization, and hence should not compromise its role as an age tracer. As suggested in Figure 7.6, the raw $\text{EW}(\text{H}\alpha)$ appears to correlate strongly with the corrected one $\text{EW}(\text{H}\alpha)_{\text{corr}}$, however robust use of this tracer is only possible if we include S/N cuts on the sample that require an understanding of the background contribution.

On average, the background contribution makes up 90 per cent of the light, such that the uncertainties introduced when converting between observed and galaxy-corrected $\text{EW}(\text{H}\alpha)$ introduce a lot of scatter. This makes this age proxy particularly problematic to use for individual H II regions. As the background correction we currently implement relies heavily on identifying nearby (but not co-spatial) contributions to the continuum stellar light, this could be improved by utilising constraints directly at the location of the H II regions themselves. This should be possible by carrying out physically motivated stellar population fits to the H II region spectrum itself, separating the contribution of older stars and refining this age tracer. However, challenges in identifying robust young stellar templates (Emsellem et al. 2022) make such an analysis beyond the scope of this paper.

The ionization parameter also shows a lot of promise as an evolutionary tracer, however it is strongly dependent on local physical conditions and is perhaps less robust as a direct age tracer. Here we are deriving ionizing parameter based on the $[\text{SIII}]/[\text{SII}]$ ratio, which shows a very strong primary correlation and minimal secondary dependencies in photoionization modeling (Kewley & Dopita 2002). While the wide wavelength separation of the lines means they are susceptible to uncertainties in the extinction correction, the lines are at relatively red wavelengths and therefore attenuation effects are minimised. As $[\text{SIII}]/[\text{SII}]$ can be detected in the largest number (86 per cent) of our nebulae, it also holds the most promise for exploring evolutionary trends statistically.

We deem $\text{H}\alpha/\text{FUV}$ to be the least trustworthy age tracer. We note that the trends with other physical properties measured in the nebulae are weaker with $\text{H}\alpha/\text{FUV}$ than with the other two nebula age tracers. In addition, there are large uncertainties arising from the extinction correction, as $\text{H}\alpha$ and FUV are at vastly different wavelengths. At these physical scales it is not entirely clear at what stage the reddening in the stars begins to deviate from the reddening measured in the gas (via the Balmer decrement) and we find that our choice of X in $E(B - V)_{\text{stellar}} = X \cdot E(B - V)_{\text{Balmer}}$ can alter some of the observed trends. Our comparison with the *HST* SED derived $E(B - V)$ in Figure 7.4 suggests most of the nebulae are best fit by $X = 1$, however this changes at ages roughly older than 5 Myr. Another issue is the sensitivity of the observations: we detect $\text{H}\alpha/\text{FUV}$ in less than 23 per cent of our H II regions with a $\text{S/N} \geq 5$. For these reasons, we are hesitant to recommend $\text{H}\alpha/\text{FUV}$ as an age tracer, but note that it does still encode information about the early evolutionary state of these nebulae.

7.5 CONCLUSION

We combine observations from MUSE, *HST* and *AstroSat* to study H II regions and their ionizing sources together. This enables us to use the SED age from the ionizing stellar associations as clocks and time the evolution of the H II regions.

- (i) We present a catalogue of 4177 H II regions and stellar associations that are clearly matched to each other. This catalogue is well suited for placing empirical constraints on models.
- (ii) We find strong correlations between the masses of the stellar associations and the $H\alpha$ fluxes of the H II regions as well as the colour excess derived from the Balmer decrement and the SED fit.
- (iii) We search for age trends and find weak to moderate correlations with the $EW(H\alpha)$, $H\alpha/FUV$ and $\log q$.
- (iv) All three properties show consistent trends among each other, hinting at an evolutionary sequence.
- (v) We find similar trends with the raw and corrected $EW(H\alpha)$. However the S/N cuts that we apply necessitate an understanding of the background in both cases.
- (vi) Using our nebular age indicators, we find tentative trends for younger H II regions to exhibit higher densities, higher reddening and smaller separations to GMCs. Interestingly, we also find strong correlations with local metallicity, with the youngest H II regions exhibiting locally elevated metallicities.






























The catalogue presented in this paper provides a novel statistical base for further investigations of the interaction between newly formed stars and their natal gas cloud. For example it can be used to validate models like Kang et al. (2022) that use emission line diagnostics of H II regions to predict the properties of the underlying star cluster. It also allows us to study the impact of stellar feedback on the ISM and in a follow up paper, we will use this catalogue to measure escape fractions of the H II regions. The addition of PHANGS-*JWST* observations for all 19 targets (Lee et al. 2023) will further serve to refine our view of the early embedded phase of star formation, completing our view of the nebular evolutionary sequence and baryon cycle in these nearby galaxies.

STELLAR ASSOCIATIONS POWERING H II REGIONS

II. ESCAPE FRACTION OF IONIZING PHOTONS

NEWLY formed stars have a profound impact on their environment by depositing energy and momentum into the surrounding gas. However, only a fraction of the stellar feedback is retained in the cloud and observational constraints are needed to further our understanding of this process. In a sample of 19 nearby galaxies, we match H II regions from PHANGS–MUSE to their ionizing stellar source from PHANGS–HST and measure the percentage of ionizing radiation that is leaking into the surrounding DIG. Based on a catalogue, where each H II region is powered by a single young and massive stellar association, we measure a photon escape fraction of $f_{\text{esc}} = 82^{+12}_{-24}$ percent. Comparable results are obtained when different procedures are used to match the ionized gas to its source. All samples we study contain a substantial fraction of objects (up to 20 percent), where the stellar source is not sufficient to produce the H α flux observed from the nebula. Many of them are probably related to uncertain age estimates, but we also find numerous regions, where a significant fraction of the ionizing photon budget is contributed by stars that reside outside the boundaries of the H II region. This motivates the use of an alternative galaxy-wide approach, in which we include all H II regions and stellar sources, not just the ones that show a clear overlap. When summing up the ionization budget over entire galaxies, we measure slightly lower, but consistent values.

CONTRIBUTIONS

This chapter and Appendix C were submitted to A&A as Scheuermann et al. (submitted), hereafter referred to as Paper III. The original article was written by Fabian Scheuermann , Kathryn Kreckel , Jia Wei Teh , Francesco Belfiore , Brent Groves , Ashley T. Barnes , Médéric Boquien , Melanie Chévanche , Daniel A. Dale , Oleg Egorov , Simon C. O. Glover , Kathryn Grasha , Stephen Hannon , Ralf S. Klessen , Kirsten L. Larson , Janice C. Lee , Fu-Heng Liang , Laura A. Lopez , J. Eduardo Méndez-Delgado , Justus Neumann , Eve Ostriker , Hsi-An Pan , Lise Ramambason , Francesco Santoro , Eva Schinnerer , David A. Thilker , Qiushi Chris Tian , Bradley C. Whitmore , and Thomas G. Williams . The project developed as a spin-off from Paper II. I developed the code, produced the figures, and wrote the text. The next four co-authors were part of the paper team and regularly contributed to the work. Before publication, the other co-authors provided additional feedback. This work was conducted in the »Ionized ISM and its Relation to Star Formation« SWG, with additional input from the »Stellar Populations and Clusters« SWG of PHANGS.

8.1 INTRODUCTION

Stellar feedback plays a crucial role in the evolution of the ISM (Lopez et al. 2014; Klessen & Glover 2016; Krumholz, McKee & Bland-Hawthorn 2019). It shapes the surrounding gas, halts or triggers new star formation, and enriches the gas with metals. However, not all energy and momentum from the stellar feedback is deposited in the H II regions (Hanish et al. 2010). Turbulence and radiative feedback can fragment the cloud (Kakiichi & Gronke 2021), which opens up channels through which some radiation escapes, while part of the stellar feedback is absorbed by gas and dust. This is an important contribution to the DIG (Haffner et al. 2009).

A common way to constrain the leakage is by comparing the number of predicted ionizing photons emitted by individual stars or clusters, $Q(\text{H}^0)$, to the ionizing photon flux inferred from the observed $\text{H}\alpha$ luminosity of an associated H II region, $Q_{\text{H}\alpha}$ (Niederhofer et al. 2016). This yields the escape fraction of ionizing photons which is usually defined as

$$f_{\text{esc}} = \frac{Q(\text{H}^0) - Q_{\text{H}\alpha}}{Q(\text{H}^0)}. \quad (8.1)$$

It is expected to evolve over time (Trebitsch et al. 2017) and depend on various local ISM conditions, such as gas density, metallicity and the star-gas geometry (Ramambason et al. 2022). Therefore, having a large sample is necessary to marginalize over these variations, or explore trends.

For H II regions located within a massive galaxy, any escaping radiation is expected to primarily impact the surrounding galactic ISM and it is clear that within galaxies there are kpc-scale consequences of stellar feedback (Zurita, Rozas & Beckman 2000; Belfiore et al. 2022). Connecting this feedback to the galactic escape fraction, which is vital to understanding cosmic reionization (Paardekooper et al. 2011; Mitra, Ferrara & Choudhury 2013; Japelj et al. 2017; Kado-Fong et al. 2020; Ma et al. 2020; Ramambason et al. 2020; Wang et al. 2021; Chisholm et al. 2022; Flury et al. 2022a,b; Saldana-Lopez et al. 2022), is challenging.

Measurements of f_{esc} with various techniques have resulted in a wide range of estimates. One of the first attempts that applied a method similar to ours, i.e. analysing individual ionizing sources in nearby galaxies, was conducted by McLeod et al. (2020) for NGC 300. Because the galaxy is nearby (2.09 Mpc; Jacobs et al. 2009), they were able to use MUSE to identify individual O-stars as the source of the ionizing radiation. By matching them to two H II regions they measured an escape fraction of 28 ± 6 per cent and 51 ± 1 per cent respectively. Della Bruna et al. (2021) analysed 8 H II regions in NGC 7793 and found a value of $f_{\text{esc}} = 67^{+8}_{-12}$ per cent. In Della Bruna et al. (2022a), they applied their technique to M 83 and found 541 matches between H II regions and young star clusters. But only 13 per cent of them returned a positive f_{esc} , i.e. for most objects, the photons emitted by the stars were insufficient to explain the observed $\text{H}\alpha$ emission. Recently, Teh et al. (2023) studied 42 H II regions in NGC 628 and determined an escape fraction of 9^{+6}_{-6} per cent.

For the LMC, Pellegrini et al. (2012) analysed 14 H II regions and measured 30 to 60 per cent, but due to the method they used, their cited values are lower limits. One of

the most prominent regions in this galaxy is the Tarantula Nebula. Doran et al. (2013) derived that 5 to 10 percent of the radiation should leak from this gigantic star-forming region, but also note, that due to high uncertainties, the value might be as high as 61 percent.

In a sample of 39 metal-poor dwarf galaxies, Ramambason et al. (2022) observed a wide range of escape fractions with values as high as $f_{\text{esc}} = 60$ percent. The same value was found by Egorov et al. (2018), while Gerasimov et al. (2022) reported a range of 55 to 69 percent.

As far as simulations are concerned, considerable progress has been made in recent years. While early studies like Dale, Ercolano & Bonnell (2012, 2013), Howard, Pudritz & Klessen (2017) and Howard et al. (2018), or Rahner et al. (2017) were computed for very specific environments, more recent studies like Kim, Kim & Ostriker (2019, 2021b) or Tacchella et al. (2022) are able to explore a broader range of parameters.

With the availability of large FoV IFU spectroscopy instruments, like MUSE or SITELLE, it is now possible to study the ionized gas using multiple optical emission line diagnostics, isolating individual H II regions while imaging the full star-forming disk for a large sample of nearby galaxies. In combination with space-based imaging we are able to obtain an accurate census of the ionizing sources in galaxies beyond the Local Group. With its combination of MUSE and *HST* imaging, the PHANGS survey is ideally suited to the systematic study of star formation in nearby galaxies (Leroy et al. 2021a,b; Emsellem et al. 2022; Lee et al. 2022, 2023). In this work, we use the large sample of H II regions and stellar associations compiled in Scheuermann et al. 2023 to calculate f_{esc} and search for systematic variations with local physical conditions and star cluster properties.

This chapter is organized as follows: in Section 8.2 we present the catalogues and models that were used in this work. In Section 8.3 we use these data to compute the escape fractions of the H II regions and discuss trends with physical properties. We conclude in Section 8.4.

8.2 DATA AND MODELS

Our sample consists of 19 PHANGS galaxies that were observed with both MUSE and *HST* and is listed in Table 8.1. The matching of the nebulae and ionizing sources is described in detail in Scheuermann et al. (2023), hereafter referred to as Paper II. Below we provide a short summary of the individual data sets and how they are combined to form matched catalogues, as well as the stellar models that are used during the analysis.

8.2.1 MUSE H II regions

The MUSE is a powerful IFU spectrograph that is mounted at the VLT in Chile (Bacon et al. 2010). It possesses a $1 \times 1 \text{ arcmin}^2$ FoV and covers the wavelength range from 4800 to 9300 Å in the nominal mode. The PHANGS–MUSE survey (PI: Schinnerer) has produced wide-field mosaics of 19 nearby spiral galaxies. The analysis and main data products are described in Emsellem et al. (2022).

Table 8.1: The PHANGS galaxy sample with the number of H II regions N_{HII} , stellar associations N_{asc} and clusters N_{clu} .

Name	Distance ^a (Mpc)	N_{HII}	N_{asc}	N_{clu}
IC 5332	9.01 ± 0.41	608	397	193
NGC 0628	9.84 ± 0.63	1651	1119	1365
NGC 1087	15.85 ± 2.24	892	487	926
NGC 1300	18.99 ± 2.85	1147	401	797
NGC 1365	19.57 ± 0.78	445	489	833
NGC 1385	17.22 ± 2.58	922	525	997
NGC 1433	18.63 ± 1.86	729	494	390
NGC 1512	18.83 ± 1.88	479	521	519
NGC 1566	17.69 ± 2.00	1448	1582	2142
NGC 1672	19.40 ± 2.91	1047	1247	1628
NGC 2835	12.22 ± 0.94	777	649	728
NGC 3351	9.96 ± 0.33	769	708	595
NGC 3627	11.32 ± 0.48	1016	1325	2320
NGC 4254	13.10 ± 2.80	2375	1661	3897
NGC 4303	16.99 ± 3.04	1956	1736	3188
NGC 4321	15.21 ± 0.49	984	1483	1921
NGC 4535	15.77 ± 0.37	1168	469	567
NGC 5068	5.20 ± 0.21	1405	715	1137
NGC 7496	18.72 ± 2.81	523	265	385
Total		20 341	16 273	24 529

^aFrom Anand et al. (2021).

Santoro et al. (2022) and Groves et al. (2023) used this dataset to construct an H II region catalogue. In the first step, HIIPHOT (Thilker et al. 2000) was used to draw the boundaries of the emission line nebulae based on their $\text{H}\alpha$ brightness. This initial catalogue is composed of 31 497 nebulae, but still contains other $\text{H}\alpha$ bright objects like PNe or SNRs. They are removed by analyzing diagnostic line ratios in the BPT diagrams (Baldwin et al. 1981; Veilleux & Osterbrock 1987). We require regions to fall below the Kauffmann et al. 2003 diagnostic curve in the $[\text{OIII}]/\text{H}\beta$ versus $[\text{NII}]/\text{H}\alpha$ diagrams, and below the Kewley et al. 2006 diagnostic curve in the $[\text{OIII}]/\text{H}\beta$ versus $[\text{SII}]/\text{H}\alpha$ diagrams, with $\text{S/N} > 5$ in all lines. Where $[\text{OI}]$ is detected with $\text{S/N} > 5$, we further require that regions fall below the Kewley et al. 2006 diagnostic curve in the $[\text{OI}]/\text{H}\beta$ versus $[\text{SII}]/\text{H}\alpha$ diagram. For this work, we only use the objects that fall inside the *HST* coverage, yielding 20 341 H II regions. In addition, properties like metallicity, ionization parameter and density are

derived from the measured emission lines. For a full summary of the catalogue and the derived properties with detailed statistics, see also [Paper II](#).

We are primarily interested in the H II region $H\alpha$ luminosity, which we want to compare to the predicted ionizing photon rate from the stars. The flux measurements for the nebulae are very precise, with a typical uncertainty of less than 1 percent, but the uncertainties arising from the determination of the boundaries of the nebulae are not taken into account here. Especially for the unresolved objects, increasing or decreasing the boundary by just one pixel radially (this corresponds to 5 to 19 pc, depending on the distance to the galaxy) will alter the measured flux on average by ± 40 percent.

Next, we need to correct the measured fluxes for extinction. For the Milky Way, we use the extinction curve from O'Donnell (1994), $E(B - V)$ from Schlafly & Finkbeiner (2011) and $R_V = 3.1$. For the internal extinction, we adopt the same extinction curve and determine $E(B - V)$ from the Balmer decrement, assuming $H\alpha/H\beta = 2.86$. To convert the fluxes to luminosities we use the distances listed in Table 8.1.

In order to compare those values with the emission from the stars, we have to convert the $H\alpha$ luminosity to the number of ionizing photons. This is commonly done with a conversion factor like the one by Hummer & Storey (1987). Assuming Case B recombination, an electron temperature of 10^4 K and a density of 10^2 cm^{-3} (well matched to our sample; Barnes et al. 2021) this yields

$$Q_{H\alpha}/\text{s}^{-1} = 7.31 \cdot 10^{11} L(H\alpha)/\text{erg s}^{-1}. \quad (8.2)$$

It should be noted that the temperatures observed in our H II regions can vary between 5000 and 10000 K (Kreckel et al. 2022). We therefore run INTRAT by Storey & Hummer (1995) to compute the conversion factor for a number of temperatures and densities. A lower temperature of 5000 K would decrease the conversion factor by about 8 percent, and imply that we slightly underestimate the number of ionizing photons (and hence overestimate the escape fraction) for some objects. An increase of the density to 10^3 cm^{-3} does not change the result by more than 0.3 percent.

8.2.2 HST stellar associations and star clusters

Interstellar gas can be ionized by various sources like AGNs, SNRs, or PNe, but for this work we are only interested in the gas that is photoionized by massive stars. However, those can also form in various birth environments, ranging from regular, compact and gravitationally bound star clusters all the way to physically larger, loosely bound stellar associations (Elmegreen 2008; Gouliermis 2018). The stellar catalogues that we use in this work were created using five broadband filters from the PHANGS-*HST* survey (Lee et al. 2022).

Based on this data, stellar association catalogues at different scales (8 to 64 pc) and from different tracer images (NUV and V) were created, as described in Larson et al. (2023). Briefly, they used DOLPHOT to identify tracer stars, which were used to create simplified maps smoothed to various scales, and then used a watershed algorithm to group multiple peaks together that form a single larger structure, or stellar association. As demonstrated

in [Paper II](#), our analysis yields similar results when we use different scales and tracer bands, and we decide to use 32 pc NUV catalogue for our analysis, as it has the largest overlap with our H II region catalogue. This catalogue contains 16 273 associations across 19 galaxies that fall inside the MUSE coverage. Due to an error in the SED fit, some associations in NGC 0628 had masses that were too low. This bug has been fixed but it also caused a slight change in the number of objects compared to [Paper II](#).

In [Paper II](#), we mostly focused on stellar associations because they had better one-to-one matches with the H II regions, enabling us to associate a specific age with individual nebulae. For this study, achieving a one-to-one match is less important and we are able to consider H II regions that contain multiple ionizing sources, as we can simply sum up the ionizing photon flux. Because compact star clusters are easier to identify and model, they were the focus of previous studies constraining H II region escape fractions (e.g. Della Bruna et al. 2021; Teh et al. 2023), and we also include them in this analysis for comparison. For this we use the cluster catalogue generated by Thilker et al. (2022) and Maschmann et al. (2024). The sources are detected with DOLPHOT and DAOSTARFINDER (v2.0, Stetson 1987; Dolphin 2000) and classified into Classes 1, 2, 3 and 4 using machine learning models (Wei et al. 2020; Hannon et al. 2023). Class 1 and 2 objects represent the compact clusters while Class 3 objects are often multi-peaked and are typically referred to as »compact associations« of stars. Class 4 objects are not considered to be star clusters, but instead are classified as artefacts like background galaxies or single stars, and are therefore excluded. The catalogue that we use contains 24 529 star clusters that fall inside the MUSE coverage (6593 of Class 1 or 2 and 17 936 of Class 3).

For both the stellar association and compact star cluster catalogues, the age, mass and reddening were derived by fitting the SED as described in Turner et al. (2021). CIGALE (Boquien et al. 2019) was used with theoretical models of a single stellar population by Bruzual & Charlot (2003), a metallicity of $Z = 0.02$ and assuming a fully sampled Chabrier (2003) IMF. Since the publication of [Paper II](#), the cluster catalogue was updated by Thilker et al. (2025). However, they are broadly consistent and the changes do not alter our results, so for reproducibility we carry out our analysis on the catalogue presented in [Paper II](#).

We note a few caveats on the use of these catalogues to study the youngest stellar populations. A limited colour-space is used to derive the stellar properties, with five *HST* bands used to fit three stellar properties (mass, age and reddening). While this will be refined in the near future with the inclusion of additional *JWST* bands (Lee et al. 2023), the lack of suitable FUV coverage still presents a problem, especially for the youngest clusters. While the galaxies were observed with *GALEX* (Martin et al. 2005) and *AstroSat* (Hassani et al. 2024), the resolution of those instruments are not sufficient for the analysis. Another issue that we cannot address with our current data is the contribution of *Wolf-Rayet stars* (WR stars; Crowther 2007) or other massive single stars (Zastrow, Oey & Pellegrini 2013) to the ionizing photon output.

8.2.3 Calculating the number of ionizing photons

What is missing in the published catalogues is the number of ionizing photons $Q(\text{H}^\circ)$ that is emitted by the stars. This number can depend on a variety of assumptions, such as the adopted stellar atmosphere model (Simón-Díaz & Stasińska 2008), stellar rotation (Levesque et al. 2012), and binarity (Lecroq et al. 2024). In this section we use a number of different codes to compute $Q(\text{H}^\circ)$ as a way of understanding the systematics that go into this calculation.

Generally, the models take a certain IMF to create a grid of stars and then compute the ionizing photon flux as a function of age, based on different stellar models and atmospheres. If the IMF is fully sampled, the output scales linearly with mass. This should be treated with caution, as this is certainly not the case for lower mass cluster (Cerviño & Luridiana 2006). We try to circumvent this problem by applying certain mass cuts, as discussed in Section 8.2.4. To propagate the uncertainties from the ages, we assume a Gaussian age distribution that is clipped at 1 Myr and draw a random sample for which we then compute $Q(\text{H}^\circ)$. Since the flux is expected to scale linearly with mass (for a fully sampled IMF), error propagation is straightforward. The uncertainty of the distance affects the $\text{H}\alpha$ luminosity and the derived cluster mass in the same way and is therefore not included.

For the comparison, we consider results from CIGALE, STARBURST99 (the models without rotation SB99_v00 and those with rotation SB99_v40), and the binarity model from BPASS. An overview of the details and the age-dependency of different models is shown in Appendix C.1. Another popular model is SLUG (da Silva et al. 2012), but instead of discrete values, it treats the individual properties as probability distributions. This makes it difficult to include here and we refer to Appendix C.2 for some simplistic comparisons of our results with what is modelled by SLUG.

BC03

As previously stated, CIGALE (Boquien et al. 2019) was used to fit the SED and derive stellar properties. Therefore, the self-consistent approach is to also use the ionizing photon flux that is computed by this code. The catalogues published by Thilker et al. (2022) and Larson et al. (2023) do not contain this quantity, so we run CIGALE with the same input parameters in `savefluxes` mode to compute the ionizing photon flux as a function of age. We then use the previously fitted ages and masses of each stellar cluster and association to interpolate their value of $Q(\text{H}^\circ)$. The stellar models from Bruzual & Charlot (2003) are used, together with an IMF from Chabrier (2003).

SB99

Another popular stellar population synthesis model is STARBURST99 (Leitherer et al. 1999; Vázquez & Leitherer 2005; Leitherer et al. 2010; Leitherer et al. 2014). As input parameters, we select a single burst of star formation that follows a fully sampled Kroupa (2001) IMF with the Geneva stellar models at solar metallicity (Ekström et al. 2012) and stellar atmospheres from Lejeune, Cuisinier & Buser (1997). The remaining parameters

are left at their default values. After that we use the stellar ages and masses again to interpolate the output to the values in our sample. While it represents more of an extreme case, we also consider a model with rotating stars, spinning at 40 per cent break-up velocity, as a limiting case.

BPASS

Most models track single star evolution, but many stars form in binaries, which impacts their evolution (Sana et al. 2012). Fortunately for us, the difference in ionizing photons is expected to only become significant after 8 Myr (Lecroq et al. 2024), and hence this should not be an issue for our purposes. In order to get a rough estimate on the impact binarity might have on $Q(H^0)$, we include the Binary Population and Spectral Synthesis code BPASS by Eldridge & Stanway (2009) in our comparison.

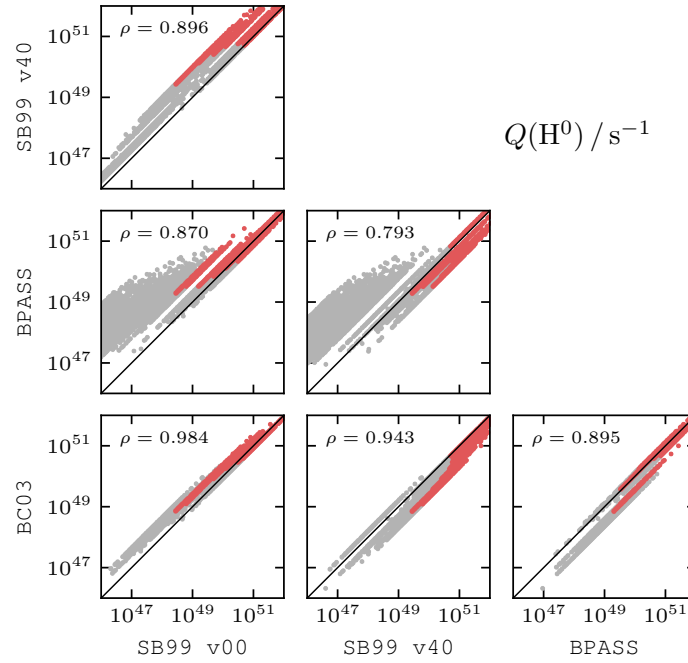


Figure 8.1: Comparison between different population synthesis models. We use a sample with random ages and masses and compute the ionizing photon flux $Q(H^0)$ with BC03, SB99 and BPASS. The grey points are the full sample and the red points show a young and massive subsample (age ≤ 8 Myr and mass $> 10^4 M_{\odot}$). The numbers in the top left corners show the Pearson correlation coefficient.

Comparison between the different models

To gauge the variations between the different models introduced in the previous sections, we compare the fluxes they predict for a random sample. For this, we create 10000

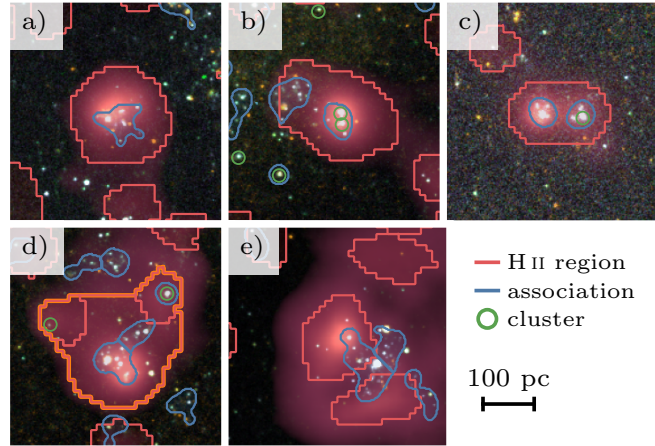


Figure 8.2: Examples for the overlap between H II regions and their ionizing sources. Each cutout shows three-colour composite images from NGC 2835. a) an H II region with a fully contained association; b) an H II region with a fully and a partially contained association and two compact clusters; c) an H II region with two fully contained associations and a compact cluster; d) multiple H II regions that form a single H II region complex with multiple associations and clusters; e) an association that overlaps with two H II regions.

clusters with random ages and masses, selected to be representative of the objects in our real sample. This is just meant to provide some order of magnitude understanding of the impact different processes have on our result.

The comparison of the predicted ionizing photon flux $Q(\text{H}^0)$ is shown as a corner plot in Figure 8.1. Overall we find decent agreement between most models, but there are a few noticeable, systematic differences. The models with rotation predict higher ionizing photon fluxes, as expected from the study of Levesque et al. (2012). As for *bpass*, there are two distinct branches: during the first few Myr the agreement is good, but around 6 Myr the difference between single stars and binaries becomes relevant.

While the models can differ significantly from one another, those differences affect mostly a specific part of the samples, like older or low mass clusters. For the objects that are young and massive, the differences are much smaller and all models return similar values. For the remainder of the paper, we use the STARBURST99 model without rotation as our fiducial model.

8.2.4 Matched catalogues

Based on the one catalogue for the H II regions and the two for the stars (clusters and associations), it is possible to define different matched catalogues. Here we describe the different samples that are used in this paper. Figure 8.2 illustrates examples for possible overlaps and Table 8.2 provides an overview of the different samples and their sizes.

one-to-one sample

The ideal scenario is a single H II region that is powered by a single stellar population, as this enables us to study trends of the escape fraction with different nebula and stellar properties. This is the case for the catalogue that was published in [Paper II](#). It consists of H II regions and stellar associations that have a clear one-to-one match and contains 4169 objects in the full sample. This includes 20.5 per cent of all H II regions and 25.6 per cent of the associations.

For lower mass associations, the presence or absence of individual massive stars can have a significant impact on the measured fluxes (Fouesneau & Lançon 2010; Fouesneau et al. 2012; Hannon et al. 2019). To alleviate stochastic sampling effects of the IMF, we also define a robust one-to-one subsample, consisting of objects that are more massive than $10^4 M_{\odot}$, younger than 8 Myr (without taking uncertainties into account) and are fully contained within the nebulae. This leaves us with 474 objects and constitutes our fiducial sample that we will use as a reference throughout the paper. Panel a) in Figure 8.2 shows an example for such an object.

extended sample

As laid out in [Paper II](#), a considerable number of our detected H II regions (14.7 per cent) overlap with more than one association. In this regard, we find 2981 H II regions that contain 8871 stellar associations, but some of the associations overlap with another H II region, making it difficult to map their ionizing photon flux to the correct nebula. We therefore limit ourselves to the 966 H II regions where the associations do not overlap with another H II region (a total of 2248 associations). Over three-quarters of them contain two associations and only a single one contains more than six associations. We refer to this sample as the extended sample and panel b) and c) in Figure 8.2 show postage stamp cutouts for this sample.

Again we aim to define a robust subsample, however, in this case the definition is not so straightforward. We require that at least one association is young and massive ($> 10^4 M_{\odot}$ and ≤ 8 Myr), and there are 707 of them in the initial catalogue. As for the overlap, we are less concerned about the old or light associations, but if a young and massive one only partially overlaps, we remove the object and are left with 291 objects for our robust extended subsample. In comparison to the one-to-one sample, the H II regions in the extended catalogue are brighter in $H\alpha$ by a factor of 5.

complexes sample

As alluded to in Section 8.2.1, difficulties in drawing the boundaries of the H II regions can lead to large uncertainties for the measured fluxes. This is particularly problematic for the 34 per cent of our H II regions that touch another ionized nebula.

For this sample, we group bordering nebulae together and call them a complex. We find 9434 nebulae in 3019 complexes, but many of them contain ionized nebulae inconsistent with photoionization (e.g. SNRs; PNe), indicating that other processes may contribute to the $H\alpha$ emission. There are also complexes that do not contain an ionizing source while

Table 8.2: The matched samples we define for this paper with the sizes of the robust subsamples in parenthesis.

Name	Sample size		Description
one-to-one	4169	(474)	H II regions with a one-to-one match to a stellar association (from Paper II).
extended	966	(291)	H II regions that contain multiple stellar associations.
complexes	1124	(317)	Multiple H II regions that form a larger complex and contain one or more associations.
cluster	8905	(2223)	H II regions that contain one or more compact clusters.
galaxy	19		All H II regions and associations in the overlapping FoV of a galaxy.

for some the ionizing source is shared with another complex (two complexes that do not touch, but share an association that overlaps with both as shown in panel e) of Figure 8.2). Excluding all of them leaves 1124 complexes in our sample that contain 3722 H II regions and 3923 associations. An example for such an object can be seen in panel d) of Figure 8.2.

For the robust subsample, we apply the same criteria as for the robust extended sample, i.e. at least one fully contained young and massive association and no partially overlapping young and massive ones. This leaves 317 objects in the robust complexes subsample.

cluster sample

In order to directly compare our results with existing studies that focused on star clusters instead of stellar associations (e.g. Della Bruna et al. 2021; Teh et al. 2023), we also construct a sample focused on the overlap of H II regions and star clusters. We use the Class 1, 2 and 3 clusters from Thilker et al. (2022) and find 22632 clusters that overlap with 8905 H II regions. Similar to the extended sample, having multiple clusters inside one H II region is not an issue, and this constitutes our cluster sample. We also apply the same mass and age restrictions for the robust subsample (the overlap criterion does not apply as the clusters are treated as point sources), and find 2223 objects in the robust cluster subsample.

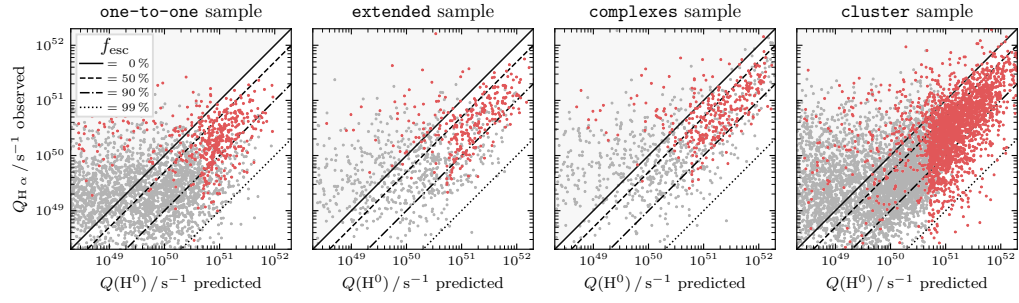


Figure 8.3: We compare the predicted ionizing photon flux $Q(H^0)$ to the observed values $Q_{H\alpha}$ from the H II region. The full samples are shown in grey and the robust subsamples in red. Lines corresponding to different escape fractions are shown for reference.

8.3 ESCAPE FRACTIONS

As is evident from the previous section, there are many assumptions involved in the calculation, but the choice of the stellar population model should only slightly alter the value of $Q(H^0)$ for the robust subsamples. Here we shift our focus to the different samples that were defined. By combining the values from the H II regions to those from the stellar catalogues we estimate the percentage of leaking radiation.

8.3.1 Ionizing photon fluxes for different samples

We start by taking the samples defined in Section 8.2.4 and plot their observed ionizing photon flux from the H II regions, $Q_{H\alpha}$, against the one predicted from the stars, $Q(H^0)$. The result is shown in Figure 8.3 and the robust subsamples are marked in red.

Looking on the one-to-one sample, we find that for more than half of the full sample, the ionizing photon production rate from the associations is not sufficient to explain the observed $H\alpha$ flux. However, most of those objects are either low mass or only partially overlap. Focusing only on the previously defined robust one-to-one subsample, we find that 80 percent of the H II regions can be ionized by their associated stars. The stellar association catalogues created at different physical scales (16 pc or 64 pc, see also Paper II) provide similar results.

For the other samples, we find similar trends, emphasising the importance of selecting robust subsamples. It also becomes clear that the different catalogues probe different regimes in terms of the intensity of the ionizing photon flux. Unsurprisingly, the extended sample and the complexes sample have ionizing photon fluxes that are on average more than four times as large as the one-to-one sample.

8.3.2 Distribution of the escape fraction

With $Q(H^0)$ from the stars and $Q_{H\alpha}$ from the H II regions, we use Equation 8.1 to compute the percentage of escaping ionizing radiation. As already obvious from Figure 8.3, for

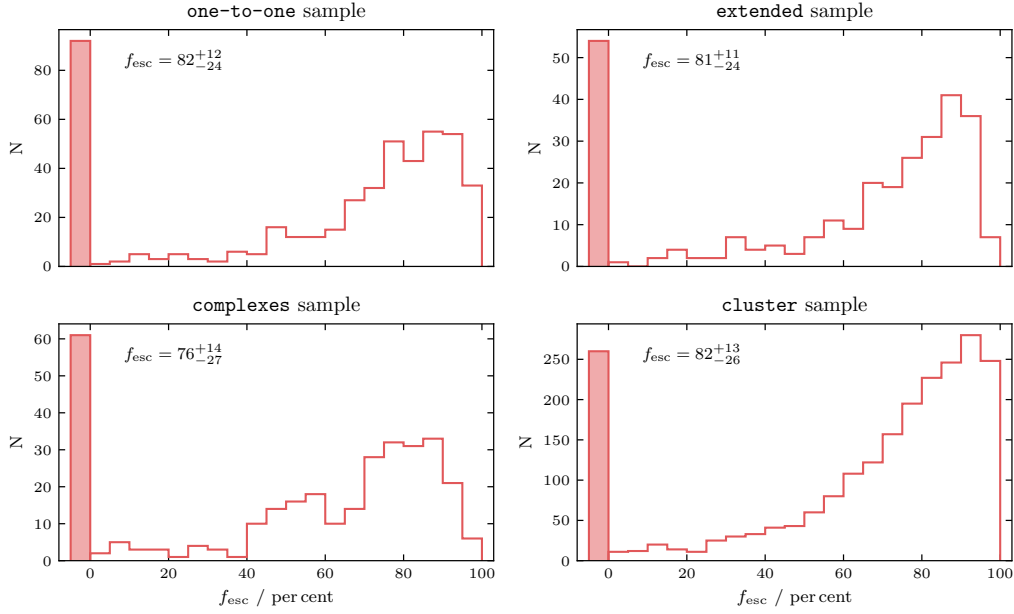


Figure 8.4: Histograms of the observed escape fractions for the different robust subsamples. All regions with negative escape fractions are grouped together in the shaded left bar.

a substantial number of objects, the ionizing photon flux from the stars is not sufficient to ionize their matched H II region. As a result, some objects have a negative escape fractions, which makes it somewhat ambiguous how to determine a representative escape fraction for the sample as a whole. Excluding all regions with negative f_{esc} leads to a higher average, while including values that are clearly unphysical also does not make sense. Our approach is a compromise to account for objects that have a positive escape fraction within their uncertainties. We assume that the escape fraction for each object follows a normal distribution and sum up the individual distributions

$$p(x) = \sum_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_{f_{\text{esc},i}}} \exp\left(-\frac{(\mu_{f_{\text{esc},i}} - x)^2}{2\sigma_{f_{\text{esc},i}}^2}\right). \quad (8.3)$$

We then compute the median (50th percentile) and the 1σ interval (16th and 84th percentile) in the range of valid values (0 to 100 percent). The results are shown as histograms in Figure 8.4 and the majority has an escape fraction larger than 50 percent. Independent of the source of ionization, we measure similar medians, ranging from 76 to 82 percent. The number of objects with negative escape fractions varies between 12 and 19 percent and they are grouped together in the left bar.

8.3.3 What about negative escape fractions?

There is a considerable number of objects with a negative escape fraction, so here we take a closer look at what may cause this. First of all, the values in the robust one-to-one subsample are consistent with $f_{\text{esc}} = 0$ within their uncertainties. Looking at the contributors to the absolute uncertainty, the dominant factor is the age, followed slightly behind by the mass while the error on the $\text{H}\alpha$ flux is mostly negligible. It is therefore not surprising that most objects with a negative escape fraction are 4 Myr or older and have a median uncertainty of 2 Myr which makes them particularly susceptible that their predicted flux is wrong. An overestimation of the age could account for an underestimation of $Q(\text{H}^0)$ and hence resolve most of the negative escape fractions.

Another possibility is that we are missing some ionizing sources all together. To explore this, we visually inspect the NUV images of the 92 objects with negative escape fraction and look for missed peaks inside the H II regions. Only about 10 per cent of them show peaks that were not included by the source detection of the associations.

Along these lines, the cloud might also be ionized by stars that fall outside of the H II region. To follow up on this point, for each H II region, we compile a list of associations that are within 10 arcsec (this corresponds to 250 to 950 pc, measured from boundary to boundary). We then take the flux of each of those nearby associations (excluding associations that are fully contained in another H II region) and scale the estimated ionizing photons that could be absorbed by the H II region, assuming the flux is emitted isotropically. For each H II region, we compare the ionizing photon flux that originates from external sources to the one coming from overlapping sources. The result is shown in Figure 8.5. We find that in the subsample with negative escape fraction the estimated contribution from external ionizing photons is about a factor of 10 more than the internal contribution, which is more than a factor 100 larger than what is found in the subsample with positive escape fraction. However, even after accounting for these external contributions, the majority of the objects with negative escape fraction remain negative. Nevertheless, it shows that the escape fraction is underestimated if external sources are not taken into account.

So far we have mainly focussed on the stellar side, but the $Q_{\text{H}\alpha}$ estimated from the H II regions are also subject to uncertainties. Given that the $\text{H}\alpha$ fluxes are directly measured and have a high S/N, the stated errors are significantly smaller for $Q_{\text{H}\alpha}$ than for $Q(\text{H}^0)$, but this does not include uncertainties due to choices like where we define the boundaries of the regions. As discussed in Section 8.2.1, even a small change in the size can significantly alter the measured flux, on average by 40 per cent. This is especially true for closely clustered H II regions where we might attribute the measured flux to a neighbouring object. In addition, in particular for faint H II regions, the DIG might contribute a majority of $\text{H}\alpha$ emission, though our brighter robust sample should only be impacted at a ~ 20 per cent level (Belfiore et al. 2022; Congiu et al. 2023).

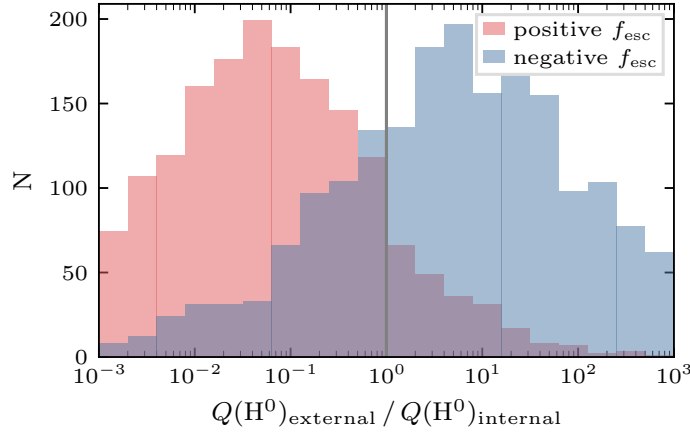


Figure 8.5: Contribution of external ionizing sources. For H II regions with negative escape fractions, the average contribution from nearby associations, not overlapping with the nebula, is far greater than for those with positive escape fractions.

8.3.4 Photoionization budget of entire galaxies

As we previously saw, it is difficult to correctly assign each H II region to their source of ionizing radiation. A careful selection is required in order to include all real sources, while excluding any false ones. Another approach would be to simply consider all objects together in order to avoid problems with the matching. And as we saw in Figure 8.5, sources outside the H II region can have a significant contribution to the ionizing photon budget as well. We therefore perform this analysis on all H II regions and associations in the overlapping coverage of MUSE and *HST*. Since we have no way of excluding H II regions that are ionized by older or low mass associations, this analysis also has to include all associations and not just the robust subsample. One potentially problematic area are the galactic centres, as some galaxies possess active nuclei. We therefore use the environment masks from Querejeta et al. (2021) to mask out the centres of the galaxies (this eliminates 286 H II regions and 690 associations from our sample).

We sum up the observed ionizing photon flux across all H II regions that fall inside the *HST* coverage, and measure $Q_{H\alpha} = 2.3 \cdot 10^{54} \text{ s}^{-1}$ across all galaxies. We then sum up the predicted ionizing photon flux from all associations inside the MUSE coverage, and find $Q(H^0) = 7.2 \cdot 10^{54} \text{ s}^{-1}$, yielding an escape fraction of $f_{\text{esc}} = 68.9 \pm 1.3$ per cent. Similarly, we can sum up the flux from all star clusters, resulting in a value of $Q(H^0) = 7.4 \cdot 10^{54} \text{ s}^{-1}$ and an escape fraction of $f_{\text{esc}} = 69.6 \pm 5.0$ per cent. The values for the individual galaxies are listed in Table 8.3 and displayed in Figure 8.6.

From this approach, the upper end of the luminosity function should be fully sampled for both the associations and H II regions, but there are differences in the sensitivity at the lower end. When looking at the cumulative luminosity function (see also Appendix C.3 for more details), we see that it is dominated by the brightest objects. H II regions fainter

than $10^{38} \text{ erg s}^{-1}$ contain less than 10 percent of the total ionizing photon flux (this is already ten times brighter than the completeness limit estimated by Santoro et al. 2022). This means that completeness should not be an issue for the H II regions. As for the stars, they are even more dominated by the upper end of the luminosity function.

Other minor adjustments to our approach have only a negligible impact on our global result. If we have excluded some real H II regions with the BPT cuts, thereby artificially reducing the observed ionizing photon flux and increasing the escape fraction, we estimate that these only contribute around ~ 10 percent to the total $\text{H}\alpha$ flux in all nebulae. This would only decrease the resulting escape fraction by a few percent. There are also some compact clusters that are not contained inside an association. Including them in our analysis raises the predicted ionizing photon flux only slightly and the escape fraction only increases by ~ 3 percent.

Overall, both our global approach and our resolved approach return comparatively high values of 70 percent and above, providing confidence in our overall result for this large statistical sample of stellar sources and H II regions.

Table 8.3: Ionizing photon budget for entire galaxies. We use the H II regions from Santoro et al. (2022) and Groves et al. (2023) and either the stellar associations (32 pc, NUV) from Larson et al. (2023) or the compact clusters from Thilker et al. (2022).

Name	H II regions	Stellar Associations		Star Cluster		
	$\log(Q_{\text{H}\alpha}/\text{s}^{-1})$	$\log(Q(\text{H}^0)/\text{s}^{-1})$	$f_{\text{esc}} / \text{ per cent}$	$\log(Q(\text{H}^0)/\text{s}^{-1})$	$f_{\text{esc}} / \text{ per cent}$	
IC 5332	51.56 ± 0.06	52.38 ± 0.04	84.9 ± 2.5	52.13 ± 0.36	$73.0 \pm$	22.5
NGC 0628	52.50 ± 0.04	53.29 ± 0.01	83.5 ± 1.6	53.05 ± 0.58	$71.7 \pm$	37.7
NGC 1087	53.11 ± 0.06	53.65 ± 0.04	71.4 ± 4.6	53.55 ± 0.15	$63.7 \pm$	13.0
NGC 1300	52.70 ± 0.04	52.97 ± 0.06	45.5 ± 9.2	53.25 ± 1.98	$71.5 \pm$	130.1
NGC 1365	53.15 ± 0.14	53.43 ± 0.06	47.5 ± 18.4	53.48 ± 0.57	$52.8 \pm$	64.0
NGC 1385	53.34 ± 0.05	53.70 ± 0.04	56.3 ± 6.4	53.61 ± 0.04	$46.8 \pm$	7.3
NGC 1433	52.26 ± 0.04	52.71 ± 0.02	64.0 ± 4.0	53.00 ± 0.14	$81.4 \pm$	7.1
NGC 1512	52.29 ± 0.05	52.62 ± 0.03	53.2 ± 6.3	53.59 ± 0.17	$95.0 \pm$	5.6
NGC 1566	53.44 ± 0.06	53.99 ± 0.02	71.7 ± 4.0	53.85 ± 0.20	$61.4 \pm$	17.8
NGC 1672	53.26 ± 0.05	53.72 ± 0.02	65.5 ± 4.1	53.67 ± 0.42	$61.1 \pm$	37.4
NGC 2835	52.56 ± 0.05	53.14 ± 0.03	74.1 ± 3.6	53.18 ± 0.44	$76.1 \pm$	24.2
NGC 3351	51.99 ± 0.04	52.95 ± 0.05	89.0 ± 1.6	52.72 ± 0.19	$81.5 \pm$	8.3
NGC 3627	53.42 ± 0.04	53.75 ± 0.05	53.9 ± 7.0	53.81 ± 0.29	$59.7 \pm$	27.5
NGC 4254	53.47 ± 0.02	54.06 ± 0.03	74.1 ± 2.2	54.10 ± 0.04	$76.6 \pm$	2.6
NGC 4303	53.60 ± 0.03	54.13 ± 0.02	70.0 ± 2.6	54.17 ± 0.07	$72.9 \pm$	4.9
NGC 4321	53.05 ± 0.04	53.77 ± 0.02	80.8 ± 2.1	53.69 ± 0.05	$76.9 \pm$	3.3
NGC 4535	52.73 ± 0.04	53.27 ± 0.05	71.0 ± 4.2	53.34 ± 0.16	$75.4 \pm$	9.6
NGC 5068	52.20 ± 0.05	52.63 ± 0.04	62.9 ± 5.6	52.67 ± 0.12	$65.7 \pm$	10.8
NGC 7496	52.62 ± 0.06	53.23 ± 0.03	75.8 ± 3.9	53.20 ± 0.05	$74.0 \pm$	4.7
Total	54.36 ± 0.02	54.87 ± 0.01	68.9 ± 1.3	54.88 ± 0.07	$69.6 \pm$	5.0

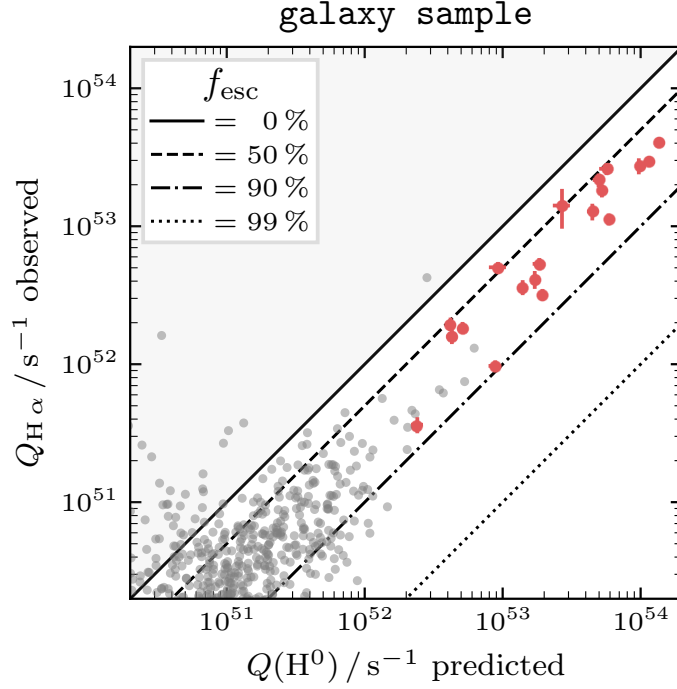


Figure 8.6: Same as Figure 8.3 but with the ionizing photon budget of entire galaxies. The values for the individual galaxies are listed in Table 8.3 and include all H II regions and associations in the overlapping coverage. For reference, the values of individual objects from the robust one-to-one and robust extended sample appear as grey dots in the lower left corner.

8.3.5 Trends with nebulae properties

The escape fraction is expected to change as the cloud evolves (Dayal & Ferrara 2018) and as a function of local conditions. For example, one might expect higher f_{esc} in lower density environments or older H II regions, as the gas is already partially cleared. In Figure 8.7 we present trends with selected properties. For this analysis we use the robust one-to-one subsample. In general, the large scatter in f_{esc} makes it challenging to identify strong correlations.

Contrary to our expectations, we find that H II regions containing younger associations are leakier (similar to the results reported by Teh et al. 2023). As a function of the stellar mass of the associations, the scatter is too large to identify any trends. In terms of the dust reddening, as traced by the Balmer decrement, we observe an anti-correlation, consistent with the picture that a deeply embedded cloud leaks less radiation than a more permeable one. However, using the stellar $E(B - V)$ we do not find a strong correlation. For the ionization parameter, $\log q$, and the local metallicity offset, $\Delta(\text{O}/\text{H})$, we do not

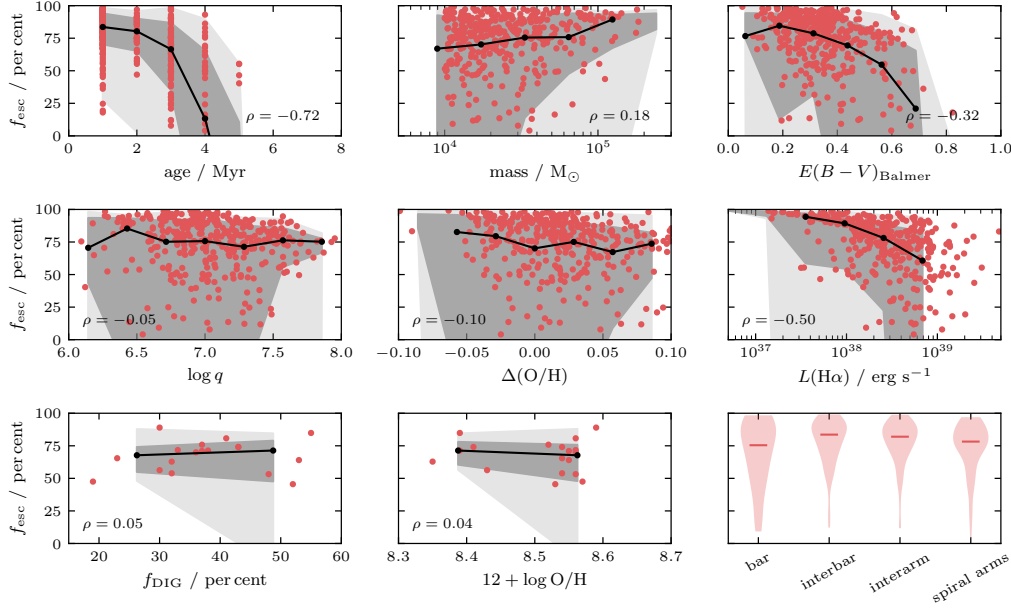


Figure 8.7: Trends between different nebula properties and the escape fraction. Shown are the age and mass of the stellar associations, the reddening $E(B - V)_{\text{Balmer}}$, the ionization parameter $\log q$, the local metallicity offset $\Delta(\text{O}/\text{H})$ of the H II region and the luminosity $L(\text{H}\alpha)$. The last row shows the global escape fraction (see Section 8.3.4) versus the fraction of DIG emission, the global escape fraction versus the average metallicity of the galaxy, and the escape fraction for different environments.

find any correlation. Lastly, for the individual H II regions, we look at the $\text{H}\alpha$ luminosity and find an anti-correlation with the escape fraction, similar to Teh et al. (2023).

Based on our global measurements of f_{esc} , we also search for galaxy wide trends. The escaping photons are thought to ionize the DIG and hence we expect a positive correlation. However, we see no trend between our observed f_{esc} and than the fraction of the DIG f_{DIG} measured by Belfiore et al. (2022), whereby our values are also generally larger. Another property is the global metallicity $12 + \log \text{O}/\text{H}$ (taken from Kreckel et al. 2020), but here too we see no dependency.

Finally, we look at the variations between different galactic environments within each galaxy, considering separately the bar, interbar, interarm and spiral arm environments. While overall they are very similar, with medians ranging from 75 to 83 per cent, the H II regions in the bar have slightly lower escape fractions.

8.3.6 Comparison with existing studies

The large sample of H II regions and stellar sources that we use to measure the escape fraction results in a wide span of values. Looking at their distribution, it becomes evident that it is deceptive to quote a single value. The values published in the literature reinforce

this as they also cover a wide range, depending on the conditions of the studied objects and the techniques used for the analysis.

We therefore focus our comparison on similar studies that also match observed H II regions to their stellar counterparts in nearby galaxies. Most of our measurements of the escape fraction exceeds 60 per cent, which itself falls at the upper end of values present in the literature (9 to 67 per cent) of comparable studies like McLeod et al. (2020) or Della Bruna et al. (2021). Even for independent calculations done in exactly the same galaxy (NGC 0628 by Teh et al. 2023; see Appendix C.2), our approach returns much larger values for f_{esc} . While it is clear that specific modelling and measurement approaches can introduce large scatter, it is not obvious that there is any clear systematic difference from our approach compared to what has been previously done. We aim to highlight here that these modelling uncertainties remain dominant while attempting to quantitatively constrain the escape fraction and below we emphasize a number of issues that stand out.

The large size of our sample prompted us to introduce certain cuts to define a robust subsample where we remove objects we deem unreliable. The stellar component must be fully contained within the H II region, its mass must be larger than $10^4 M_{\odot}$ to ensure that the IMF is fully sampled and its age must be ≤ 8 Myr because the ionized cloud should not exist much longer than a few Myr (Hollyhead et al. 2015; Kim et al. 2021a; Hannon et al. 2022). This procedure is certainly more objective than picking the objects by hand, but it might still be biased towards certain physical parameters. Partly due to this selection, the regions in our sample are rather young with a median age of 2 Myr. Nevertheless, the age distribution seems physically reasonable, as according to Kim et al. (2023), 50 per cent of the LyC photons are already emitted by this age. Around this age, the ionizing photon flux is strongly dependent on the age and hence even a small uncertainty of only 1 Myr can cut the value for $Q(\text{H}^0)$ in half. This is also reflected in the propagated error of f_{esc} , for which this is the biggest contributor.

Then there are also other issues whose uncertainties are difficult to incorporate. As mentioned in Section 8.2.1, a small change to the size of a region can alter the flux considerably and this is not accounted for in the stated error. When manually inspecting some of the outliers, it becomes clear that especially the larger H II regions are often complicated complexes that are difficult to segment into physically meaningful pieces. Furthermore, we do not consider external sources, which can have a major impact on small H II regions as shown in Section 8.3.3.

One aspect that we have not yet addressed is the impact of dust. This is a complex issue, and while some simulations are able to account for the ionizing photons that are directly absorbed by the dust, we are only able to correct the $\text{H}\alpha$ fluxes for reddening. According to Kim et al. (2019) a double-digit percentage might be lost this way. This cannot explain the difference to other observations, which have generally followed the same approach as we have, but it could help explain the higher values that we observe when compared to simulations.

8.4 CONCLUSION

We combine catalogues of H II regions with catalogues of young stellar associations and star clusters to measure the escape fraction across an unprecedentedly large sample of thousands of regions across 19 nearby galaxies. This data sets enables a statistical view of how much ionizing radiation escapes from the local H II region environment out into the galaxy disk.

To perform the calculation of f_{esc} , we compare $Q_{\text{H}\alpha}$, derived from the extinction-corrected H α luminosity, and $Q(\text{H}^0)$, derived from the associated stellar counterpart. After comparing different methods of matching stars and gas, we focus our analysis on regions with a one-to-one match between a single H II region and a single young stellar association. To mitigate the impact of stochastic sampling, we establish a robust subsample of 474 massive ($> 10^4 M_{\odot}$) and young (≤ 8 Myr) stellar associations that are fully contained in the nebula.

- (i) From the robust one-to-one subsample, we measure a comparatively high value of $f_{\text{esc}} = 82^{+12}_{-24}$ percent.
- (ii) We derive similar escape fractions with different crossmatching methods or when using compact clusters instead of the loosely bound associations.
- (iii) The escape fraction depends heavily on the derived stellar parameters and the boundaries of the H II regions.
- (iv) For many H II regions, a significant fraction of the ionizing photon budget comes from stars that reside outside their boundaries.
- (v) If we take all stellar associations and H II regions together, to account for the entire ionizing photon budget in the galaxies, we measure slightly lower but consistent values.
- (vi) We do not identify any pronounced trends between f_{esc} and local or global galaxy properties.

This work demonstrates that a large fraction of ionizing photons can escape their immediate H II region bubble, and be available to power the surrounding pervasive DIG observed in local galaxies (Haffner et al. 2009; Belfiore et al. 2022).

Overall, more work is needed on precise determination of the stellar ionizing photon output before it is possible to establish more concrete connections between changes in f_{esc} and local conditions in the ISM.

We expect that significant progress on parametrising the properties of the stellar clusters will come from the inclusion of additional infrared *JWST* bands, and the use of priors based on the presence of H α emission Whitmore et al. (2025). In addition, the recent availability of new *HST* H α high resolution maps (Chandar et al. 2025) is enabling more detailed cross-matching between ionized gas and stellar populations at matched resolution for the brightest, most compact regions Barnes et al. (submitted). These maps already reveal the ways in which the clustering and hierarchical structuring of young stars make simple one-to-one matching of ionized gas and stars extremely challenging.

EPILOGUE

BEYOND THE HORIZON

ALONGSIDE the new insights that we gained, we also encountered limitations in certain domains. Here we take a brief look at the challenges and opportunities that lie ahead. The following two points are unrelated and represent specific topics that are relevant for the further development of the work presented in this thesis.

9.1 IMPROVEMENT OF STELLAR AGES

When determining the age of a stellar population, there are a number of degeneracies that can alter the result. As shown in Appendix C.2, the stochasticity of the IMF can be a major source of uncertainty. The result depends on the SED model that was used, and to our surprise, there were also inconsistencies for massive clusters, which should not be affected.

Next, the presence of dust in the vicinity of young clusters can cause them to display identical optical colours as old GCs. Therefore, it is necessary to take additional information such as morphology, location or H α emission into account. In the latest version of the cluster catalogue, the age estimates were already improved (Thilker et al. 2025) and further refinements are expected by including additional *JWST* bands (Whitmore et al. 2025). This will provide robust age determinations on cosmic timescales.

Nevertheless, narrowing down the exact age of very young clusters (< 5 Myr) remains an open challenge for the time being. That became apparent in Chapter 8, where even a minor alteration can drastically impact the result. A key reason for this is that the rapid evolution of massive stars during this period is not sufficiently represented in our observed filters. At this point, we sorely miss the absence of UV observations. Preceding FUV telescopes were simply unable to achieve a resolution that is comparable to *HST*. There are plans for future space telescopes such as the *Cosmological Advanced Survey Telescope for Optical and UV Research* (CASTOR; Cote et al. 2019) or the *Ultraviolet Explore* (UVEX; Kulkarni et al. 2021), but we will have to wait a while until they take pictures.

9.2 EXTENDED SPECTRUM

There is a long-standing debate in the literature about which method is the most reliable for determining metallicity. MUSE enables us to study the ionized gas in great detail, but it only covers a wavelength range from 4800 to 9300 Å. So we are missing the [OII] $\lambda 3726, 3729$ doublet, which is commonly used to derive gas-phase oxygen abundances, densities and temperatures (Kewley & Dopita 2002). BlueMUSE will eventually extend the spectrum down to 3500 Å (Richard et al. 2019), but it will still take a few years before that happens.

However, there is already an alternative today: SITELLE is an optical *imaging Fourier transform spectrometer* (IFTS) with a wavelength range of 3500 to 9000 Å (Drissen et al. 2019). It is mounted at the CFHT and has a large FoV of 11×11 arcmin with a resolution similar to MUSE (0.32 arcsec per pixel). The instrument is a Michelson interferometer and a data cube is recorded by moving one of its two mirrors. After applying a Fourier transformation, a spectral cube is obtained. As a result, the spectral resolution depends on the number of images taken and the width of the observed spectrum. For our purposes, the SN1 filter (3650 to 3850 Å) is well equipped to observe the [OII] doublet.

We have a subsample of five galaxies that were already observed with SITELLE (NGC 0628, NGC 2835, NGC 3351, NGC 3627 and NGC 4535) and another five possible targets (NGC 1087, NGC 1300, NGC 4254, NGC 4303, NGC 4321 and NGC 5068). Here we take a preliminary look at the PHANGS–SITELLE survey and the possibilities it opens up. The data reduction is similar to the one described in Rousseau-Nepton et al. (2018) and is performed using the *Outils de Réduction de Cubes Spectraux* (ORCS) package by Martin, Prunet & Drissen (2016, 2018). Even with a spectral resolution of $R = 1300$, the two [OII] lines are not resolved in our data, so we assume a fixed flux ratio of 1.4, corresponding to the low density limit (Sanders et al. 2016).

Our data only includes the
SN1 filter of SITELLE.

A first glimpse at the extracted [OII] $\lambda 3726$ line map is shown in Figure 9.1. Because there is no spectral overlap between our SITELLE data and MUSE, we resort to an external catalogue for comparison. We make use of the catalogue from the TYPHOON survey (Grasha et al. 2022) and measure the corresponding [OII] fluxes from SITELLE and the $H\alpha$ fluxes from MUSE. The comparison reveals consistent flux measurements across all three instruments. The combination of data from different telescopes opens up the possibility to examine prescriptions that were previously unavailable to us. These findings reinforce our plan to examine a wide array of properties such as density, temperature, and metallicity calibrations in more detail (Habjan et al. in preparation).

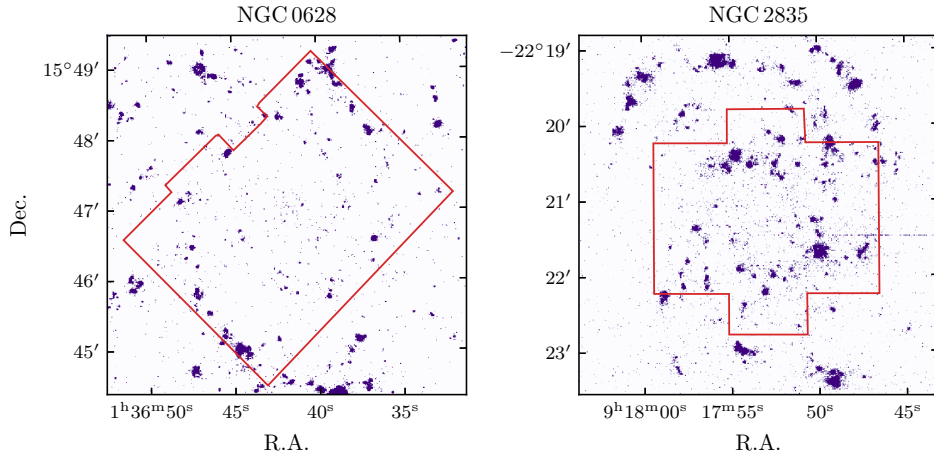


Figure 9.1: Cutouts of the [OII] $\lambda 3726$ line maps from SITELLE with footprints of the MUSE observations.

RÉSUMÉ

NOW that we are nearing the end of this thesis, it is time to reflect on the results. We have studied different components of the ionized ISM with the help of an extensive data set that combines observations from complementary instruments such as IFS and high resolution imaging. This has provided us with a comprehensive view of its structure, evolution, and role in the galaxy. In the following chapter, we will go over the most important findings, the knowledge gained, but also the questions that remain.

10.1 CATALOGUES WITH EMISSION LINE NEBULAE

The detection and classification of different types of emission line nebulae form the basis for all our further analyses. Numerous studies within PHANGS have tackled this topic, each with a different focus. The nebular catalogue by Santoro et al. (2022) and Groves et al. (2023) serves as a starting point for many studies. As the name suggests, it contains various types of nebulae, but its core function is the identification of H II regions.

This work is also part of this group. In order to apply the PNLF in Paper I, we created a dedicated catalogue tailored specifically to this purpose. We found 1049 candidates and classified 899 of them as PNe. It is not surprising that most of these PNe (706 of 899) do not appear in the nebular catalogue, and those that do were previously unclassified.

In contrast, Congiu et al. (2023) aimed to produce a more general catalogue that is as complete as possible and contains all types of nebulae. This was achieved by using multiple line maps for the detection, returning 40920 nebulae, among them a similar number of PNe as in Paper I. However, only 54.1 per cent of the PNe from our catalogue have counterparts in theirs, and many are classified differently. A comparison with Santoro et al. (2022) also revealed significant discrepancies. While both cover a similar amount of H α emission, the segmentation between the individual H II regions differs considerably.

The most recent addition is the SNR catalogue by Li et al. (2024). They detected 2233 potential candidates and classified 1166 of them as robust SNRs. Many candidates overlap with H II regions, which is hardly surprising, given that both are associated with young stars. In addition, Belfiore et al. (2025) concentrated specifically on the classification of nebulae rather than their detection, offering a complementary perspective on the catalogues discussed above.

In conclusion, no existing catalogue provides a complete census of all emission line nebulae and each one serves a distinct purpose. Thus, catalogue selection must be made carefully, depending on the scientific goals, and users should remain aware of potential contamination and classification inconsistencies.

10.2 DISTANCE MEASUREMENTS

In the past decade, distance measurements with the PNLF have experienced a renaissance, to which [Paper I](#) also contributed. Before that, around 70 galaxies had PNLF distances and in [Chapters 5](#) and [6](#) we applied this technique to 20 targets, 15 of which were new. Several of these lie among the most distant systems yet studied, demonstrating how the sensitivity of MUSE extends the limits of this method.

We are not the only ones in taking advantage of this capability. Comparisons with the work of other groups, such as Roth et al. ([2021, 2023](#)), helped us to identify the strengths and weaknesses of this technique. Their most recent work, Jacoby et al. ([2024](#)), offered some points for comparison. For NGC 1433 and NGC 1512, our different implementations yield consistent results. In contrast, our distance to NGC 1385 is significantly smaller, which they attribute to challenges in the background subtraction, as many of our PNe lie in the central region with strong emission. A similar situation occurs for NGC 253, where their analysis is based on a much smaller sample that is limited to the centre.

Thanks to the large data set, we can investigate such inconsistencies more closely. Internal extinction is usually assumed to be negligible because PNe have a higher vertical scale height than dust. However, this assumption can break down in galaxies such as NGC 253, where a starburst-driven outflow increases the vertical extent of the dust. Variations in metallicity among our galaxies (Kreckel et al. [2019](#)) are too small to determine any dependency. At this point, metal-poor elliptical galaxies are needed to test for a possible effect. On the bright side, radial metallicity gradients are typically shallow, so they rarely pose a concern.

But caution is required when selecting objects from a catalogue. Thanks to spectral information, it is now relatively easy to avoid misclassified SNRs that could bias the distance. Overluminous PNe, however, remain a recurring challenge that can skew the results.

It is difficult to use an existing nebular catalogue that was not specifically compiled for this purpose, because PNe can easily be overlooked. We demonstrated in Congiu et al. ([2023](#)) that it is nevertheless possible. However, it should be noted that the aim of this catalogue was to detect as much nebulae as possible. Even there, we usually found barely enough to carry out the measurement, but the results were consistent with our dedicated survey.

We still lack a theoretical explanation of the PNLF, but the wealth of IFU observations enables improved distance measurements to many galaxies. It can be tempting to apply the PNLF to any existing data set, but especially when only the central region is observed, the result can be biased. Our work highlights the importance of carefully selecting targets and of avoiding the blind application of the method to any galaxy.

10.3 STARS THAT IONIZE GAS

The second half of this thesis, comprising [Chapters 7 and 8](#), focused on linking the gas in H II regions to its source of ionized radiation. To this end, we fully exploited the diverse PHANGS data set. In [Paper II](#), we carefully matched the spatial masks of H II regions with those of young stellar associations. We found that 42.5 per cent of the H II regions overlap with one or more associations and 71.8 per cent of the associations overlap with one or more H II regions.

This matching allowed us to estimate the age of the H II region – but with the limited overlap only for a portion of the full sample. We aim to find a potential age tracers that can be directly measured from the data cube. The ones we study are $\text{EW}(\text{H}\alpha)$, $\text{H}\alpha/\text{FUV}$ and $\log q$ and all of them show a moderate anti-correlation with the age. However, there is a clear interdependence between them, supporting their eligibility as age tracers. Based on these, we find that young H II regions have higher densities, metallicities, and extinctions, and over time they move away from their nearest GMC.

[Paper III](#) builds on this catalogue to measure a widely debated property: the escape fraction of ionizing photons from H II regions f_{esc} . This quantity is essential for understanding the feedback processes between stars and their natal clouds. Previous studies were mostly restricted to specific environments due to their small sample sizes, which made it challenging for them to perform more comprehensive analysis. Our diverse catalogues enable us to explore this in far more detail. In addition to the one-to-one H II region–association matches, we constructed additional catalogues to account for overlapping objects, forming larger complexes when multiple associations contribute to a single ionized region.

We predict the number of ionizing photons emitted by the stars $Q(\text{H}^0)$ with the help of population synthesis models, and compare them to the observed ionizing photon flux $Q_{\text{H}\alpha}$. The resulting escape fractions f_{esc} cover the full range, with the majority at the upper end ($f_{\text{esc}} = 82^{+12}_{-24}$ per cent). This contrasts with most previous studies, which found lower values. The largest uncertainty comes from age, as the stars evolve quickly during the first few Myr. Another source of error is drawing the exact boundaries of the H II regions. There is also a substantial number of objects, where the flux of the star cannot explain the values from the cloud. Part of this can be explained by external sources that do not overlap, but are close enough to contribute. The large uncertainties of individual objects have prompted us to look at the whole thing from a galactic perspective. To do this, we sum up the fluxes of all H II regions and stellar associations that fall inside the overlapping FoV. We find a total escape fraction of $f_{\text{esc}} = 68.9 \pm 1.3$ per cent and the fluctuations between the different galaxies are much smaller.

It may be tempting to think of the escape fraction as a uniform property across each galaxy, but our results reveal a more complex picture: it is deeply dependent on underlying assumptions and local conditions. Individual regions tell very different stories, yet when viewed together their contradictions blend into a coherent whole. On galactic scales, the noise fades, and a clearer signal emerges – one that offers a more reliable basis for understanding how stellar feedback spreads into the wider environment.

10.4 CLOSING REMARKS

The work presented here shows that the ionized *interstellar medium* is a window into the life of galaxies. With the versatility of *integral field spectroscopy*, supported by the sharpness of high-resolution imaging, uncertainties can often be resolved by turning to layers of information that had not been considered before. By following the light of *planetary nebulae*, HII regions, and the stars that give them life, we touch on processes that drive the growth, transformation, and renewal of galaxies. These same processes echo across cosmic time, linking the nearby universe we can resolve in exquisite detail to the first galaxies that illuminated the cosmos. In this sense, the small scale stories told here are part of a larger narrative: how stars and gas together shape the evolution of galaxies and, ultimately, the history of the universe.

APPENDIX

PLANETARY NEBULA LUMINOSITY FUNCTION

This chapter contains the appendix for Chapter 5 (Paper I).

A.1 COMPARISON WITH LITERATURE DISTANCES

In Figure A.1 to A.18, we compare our measured distances with literature values. The data for the plots were taken from the *NASA/IPAC Extragalactic Database* (NED).¹ If a source published more than one distance, the mean of the published distances is used with the uncertainties added in quadrature (those values are marked with a diamond). Only the five most recent publications are considered for each method. The NAM distances were obtained by Anand et al. (2021) with the online tool² by Kourkchi et al. (2020) and Shaya et al. (2017).

We use the following abbreviations in the figures: Planetary Nebula Luminosity Function (PNLF), Tip of the Red Giant Branch (TRGB), Type Ia Supernova (SNIa), SNII optical (SNII), Numerical action method (NAM), distance to galaxy group (Group), Tully–Fisher (TF), Tully estimate (TE), Brightest Stars (BS), Gravitational Stability Gaseous Disk (GSGD), Disk Stability (DS), Infrared Astronomical Satellite (*IRAS*), Ring Diameter (RD) and Statistical (Stat).

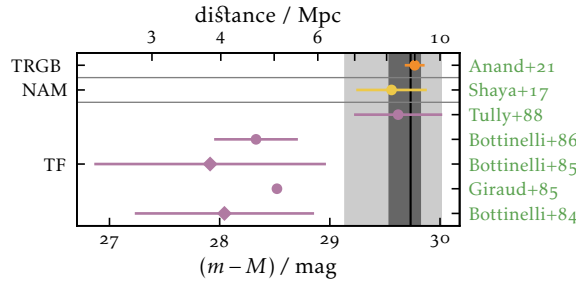


Figure A.1: Comparison with literature distances for IC 5332. We measure a distance modulus of $(m-M) = 29.73^{+0.10}_{-0.20}$ mag and a distance of $D = 8.84^{+0.39}_{-0.82}$ Mpc.

¹<https://ned.ipac.caltech.edu/Library/Distances/>

²<https://edd.ifa.hawaii.edu/NAMcalculator/>

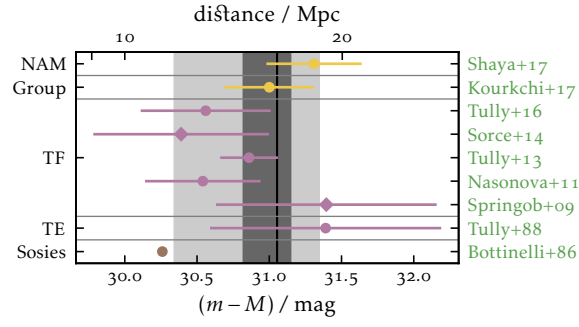


Figure A.2: Comparison with literature distances for NGC 1087. We measure a distance modulus of $(m - M) = 31.05^{+0.10}_{-0.24}$ mag and a distance of $D = 16.25^{+0.74}_{-1.79}$ Mpc.

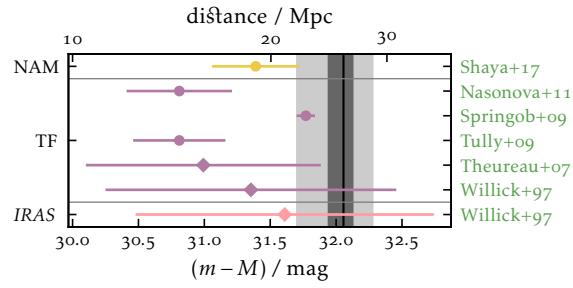


Figure A.3: Comparison with literature distances for NGC 1300. We measure a distance modulus of $(m - M) = 32.06^{+0.08}_{-0.12}$ mag and a distance of $D = 25.77^{+0.90}_{-1.42}$ Mpc.

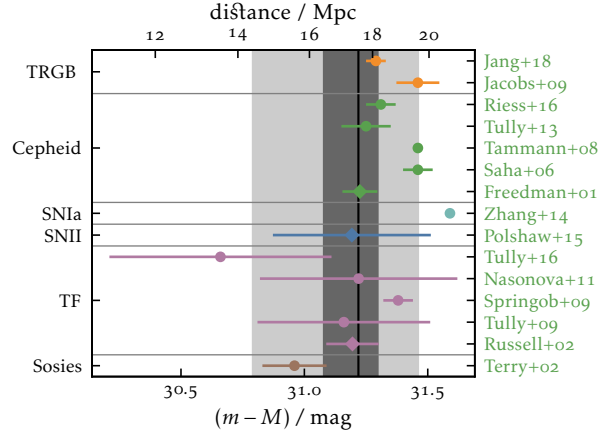


Figure A.4: Comparison with literature distances for NGC 1365. We measure a distance modulus of $(m-M) = 31.22^{+0.08}_{-0.14}$ mag and a distance of $D = 17.53^{+0.66}_{-1.16}$ Mpc.

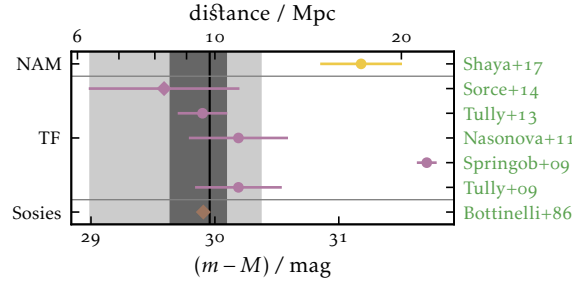


Figure A.5: Comparison with literature distances for NGC 1385. We measure a distance modulus of $(m-M) = 29.96^{+0.14}_{-0.32}$ mag and a distance of $D = 9.81^{+0.63}_{-1.46}$ Mpc.

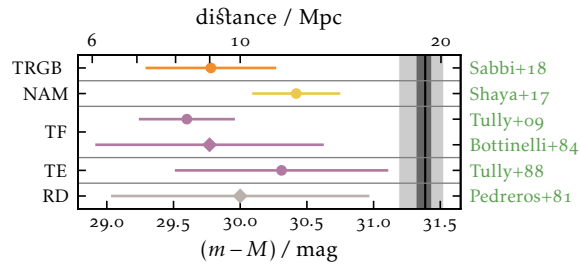


Figure A.6: Comparison with literature distances for NGC 1433. We measure a distance modulus of $(m-M) = 31.39^{+0.04}_{-0.07}$ mag and a distance of $D = 18.94^{+0.39}_{-0.56}$ Mpc.

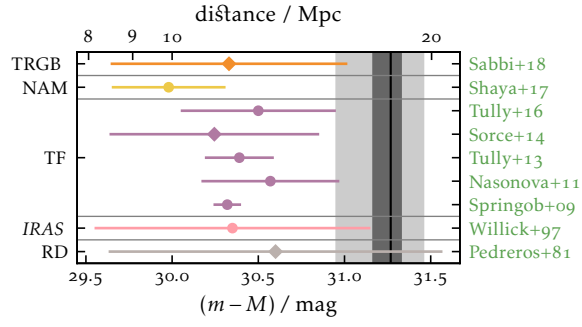


Figure A.7: Comparison with literature distances for NGC 1512. We measure a distance modulus of $(m-M) = 31.27^{+0.07}_{-0.11}$ mag and a distance of $D = 17.93^{+0.53}_{-0.88}$ Mpc.

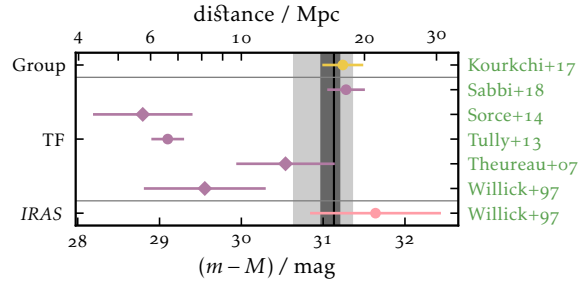


Figure A.8: Comparison with literature distances for NGC 1566. We measure a distance modulus of $(m-M) = 31.13^{+0.08}_{-0.17}$ mag and a distance of $D = 16.84^{+0.60}_{-1.29}$ Mpc.

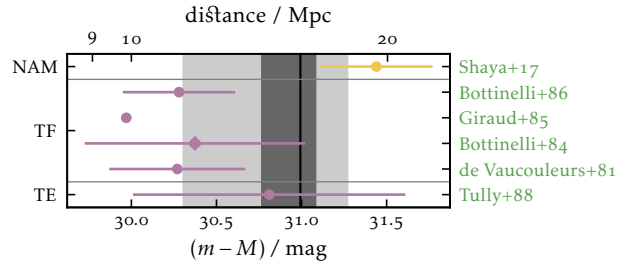


Figure A.9: Comparison with literature distances for NGC 1672. We measure a distance modulus of $(m-M) = 30.99^{+0.09}_{-0.23}$ mag and a distance of $D = 15.80^{+0.68}_{-1.68}$ Mpc.

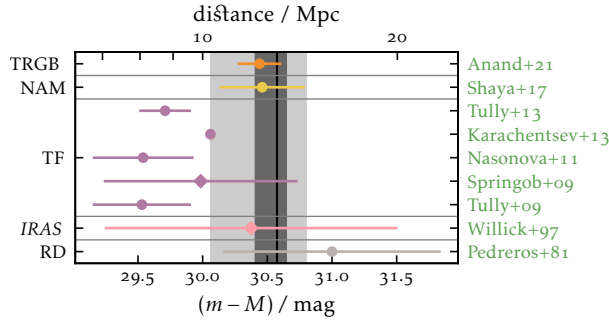


Figure A.10: Comparison with literature distances for NGC 2835. We measure a distance modulus of $(m - M) = 30.57^{+0.08}_{-0.17}$ mag and a distance of $D = 13.03^{+0.46}_{-1.04}$ Mpc.

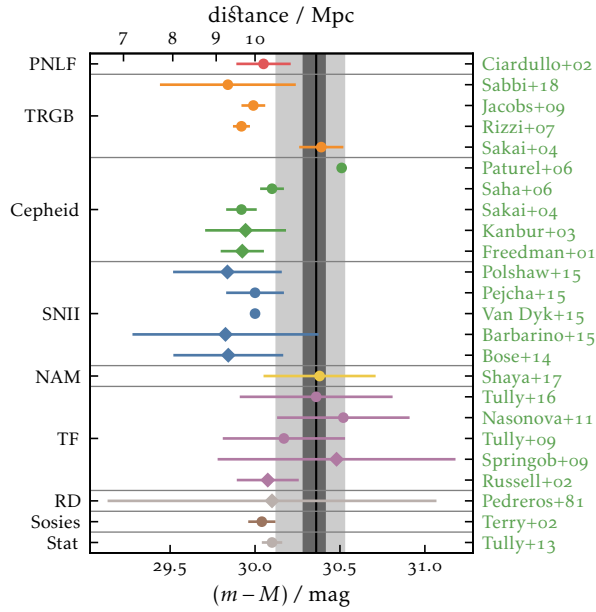


Figure A.11: Comparison with literature distances for NGC 3351 (M 95). We measure a distance modulus of $(m - M) = 30.36^{+0.06}_{-0.08}$ mag and a distance of $D = 11.80^{+0.31}_{-0.43}$ Mpc.

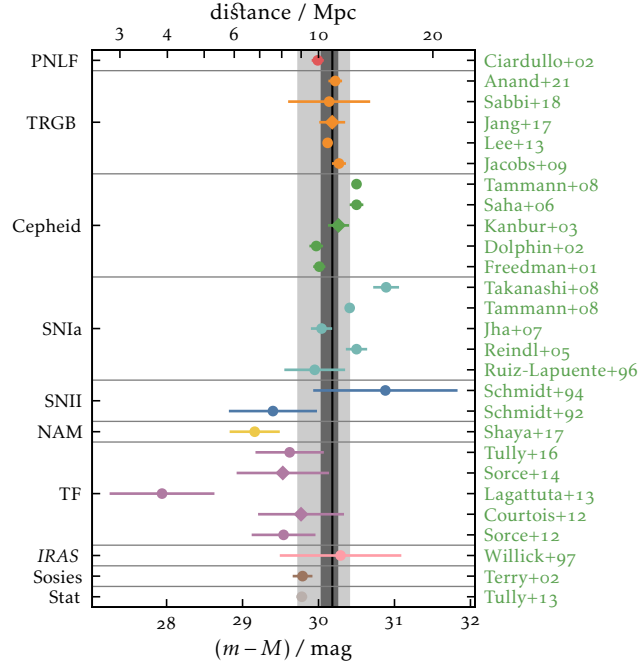


Figure A.12: Comparison with literature distances for NGC 3627 (M66). We measure a distance modulus of $(m - M) = 30.18^{+0.08}_{-0.15}$ mag and a distance of $D = 10.88^{+0.39}_{-0.77}$ Mpc.

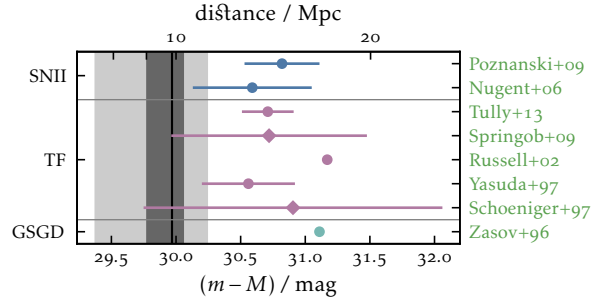


Figure A.13: Comparison with literature distances for NGC 4254 (M99). We measure a distance modulus of $(m - M) = 29.97^{+0.09}_{-0.20}$ mag and a distance of $D = 9.86^{+0.42}_{-0.91}$ Mpc.

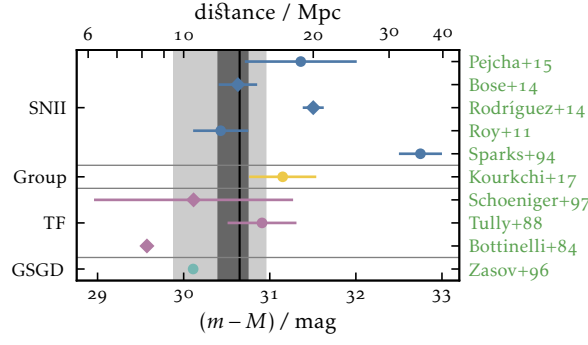


Figure A.14: Comparison with literature distances for NGC 4303 (M 61). We measure a distance modulus of $(m-M) = 30.65^{+0.10}_{-0.26}$ mag and a distance of $D = 13.49^{+0.64}_{-1.60}$ Mpc.

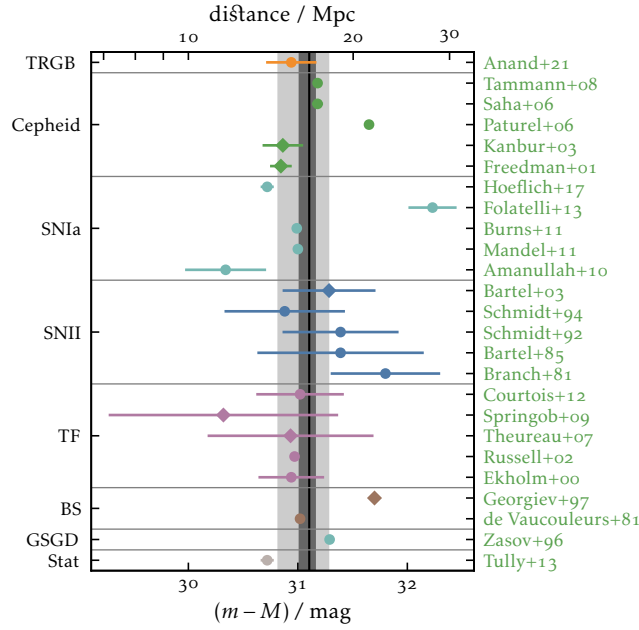


Figure A.15: Comparison with literature distances for NGC 4321 (M 100). We measure a distance modulus of $(m-M) = 31.10^{+0.06}_{-0.10}$ mag and a distance of $D = 16.62^{+0.46}_{-0.74}$ Mpc.

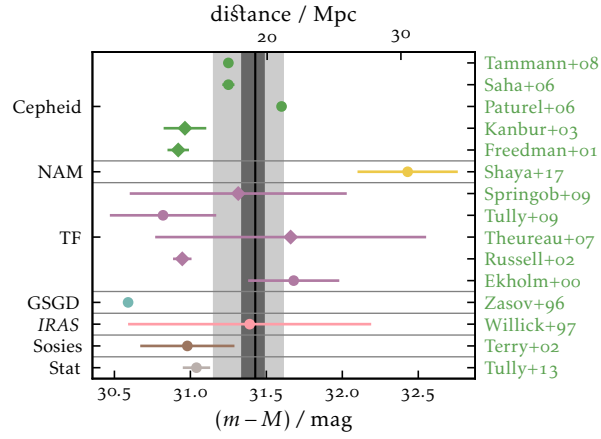


Figure A.16: Comparison with literature distances for NGC 4535. We measure a distance modulus of $(m-M) = 31.43^{+0.06}_{-0.09}$ mag and a distance of $D = 19.29^{+0.56}_{-0.82}$ Mpc.

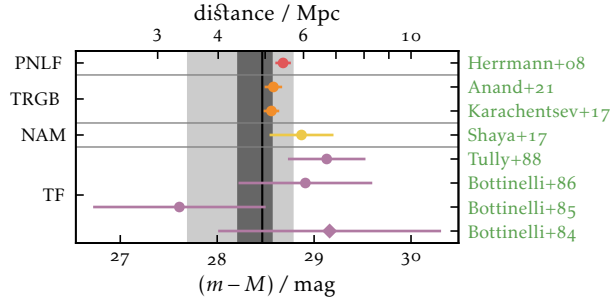


Figure A.17: Comparison with literature distances for NGC 5068. We measure a distance modulus of $(m-M) = 28.46^{+0.11}_{-0.26}$ mag and a distance of $D = 4.93^{+0.24}_{-0.59}$ Mpc.

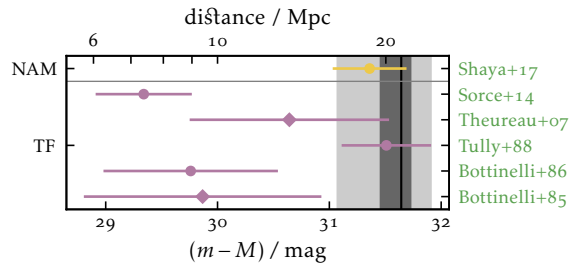


Figure A.18: Comparison with literature distances for NGC 7496. We measure a distance modulus of $(m-M) = 31.64^{+0.09}_{-0.19}$ mag and a distance of $D = 21.31^{+0.89}_{-1.89}$ Mpc.

A.2 EXTINCTION CORRECTION

To quantify if we can ignore the effects of internal extinction (dust within the target galaxy, not the circumstellar extinction of the PN itself), we create a synthetic luminosity function by sampling from Equation 5.7. To half of the sample we randomly add extinction. Then we use our fitting algorithm to measure the distance to the sampled data. Figure A.19 shows the PNLF without extinction and with extinction. As shown in Appendix A.3, the ability to derive a reliable distance depends on the sample size and the range of the PNLF that we sample. For a sample size of 20 PNe, the measured distance is unaffected for $A_{5007} = 0.1$ mag, but can be significantly overestimated for $A_{5007} = 1$ mag. Once the sample gets larger, the impact decreases and for a sample size of 100 PNe, our algorithm is able to derive the correct distances from the compound sample with any A_{5007} .

As Figure A.19 shows, the composition of a luminosity function should bend the observed luminosity function below the fitted function. Since we do not observe this behaviour in our observed sample, this further indicates that our sample is not significantly reddened.

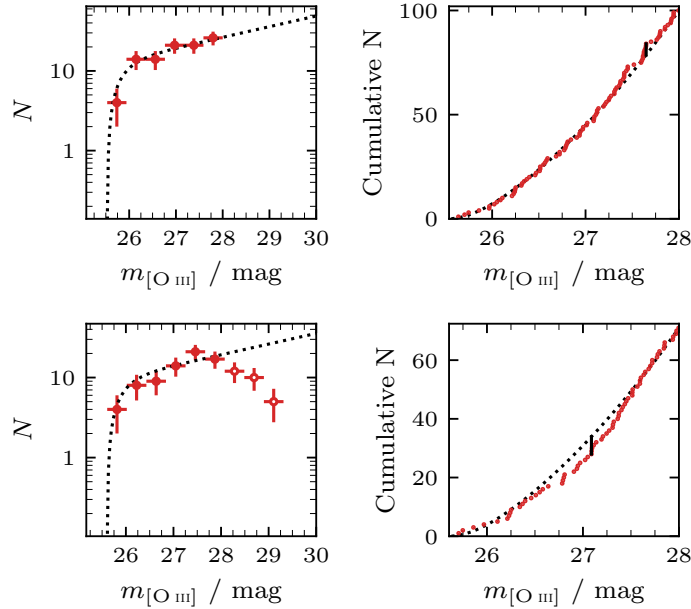


Figure A.19: We synthesis a PNLF with 100 points. Half of the points are randomly chosen and an extinction of $A_{5007} = 1$ mag is applied. The top panels show the full sample without extinction and the bottom panels show the compound sample with the extinction applied. The left panels show the binned PNLF and the right panels show the cumulative PNLF.

A.3 PRECISION OF THE PNLF

The fit of the PNLF is dominated by the brightest PNe. Once we observe a PNe with apparent magnitude $m_{[\text{O III}]}$, the galaxy can not be further away than $(m-M)_{\text{max}} = m_{[\text{O III}]} - M^*$, where M^* is the zero point of the PNLF (it can be slightly further away due to the way we include the uncertainties with Equation 5.9). Due to the exponential tail of the PNLF, the fainter PNe push the function to the right (e.g. towards larger distances). Hence if we only observe the bright end, we usually underestimate the distance.

Jacoby (1997) investigated the effects of sample size by measuring the PNLF from different sized subsets from the PNe detected in M 87. To test the reliability of the PNLF, we do something similar with synthetic data. We vary the sample size from 20 to 150 objects and the completeness limit from 26.5 to 28 mag. For each step we sample 1000 luminosity functions with a fixed distance modulus of $(m-M) = 30$ mag. The later means that the part of the PNLF that we sample varies from 1.0 to 2.5 mag. The result can be seen in Figure A.20. The precision increases with sample size, showing that even from a small sample of 20 PNe, one can achieve a precision better than ~ 5 percent. For a brighter completeness limit, the sample is concentrated at the steep cut off, and hence the distance is very well constrained. For fainter completeness limits, this part is sampled more sparsely which leads to a lower precision of the measured distance.

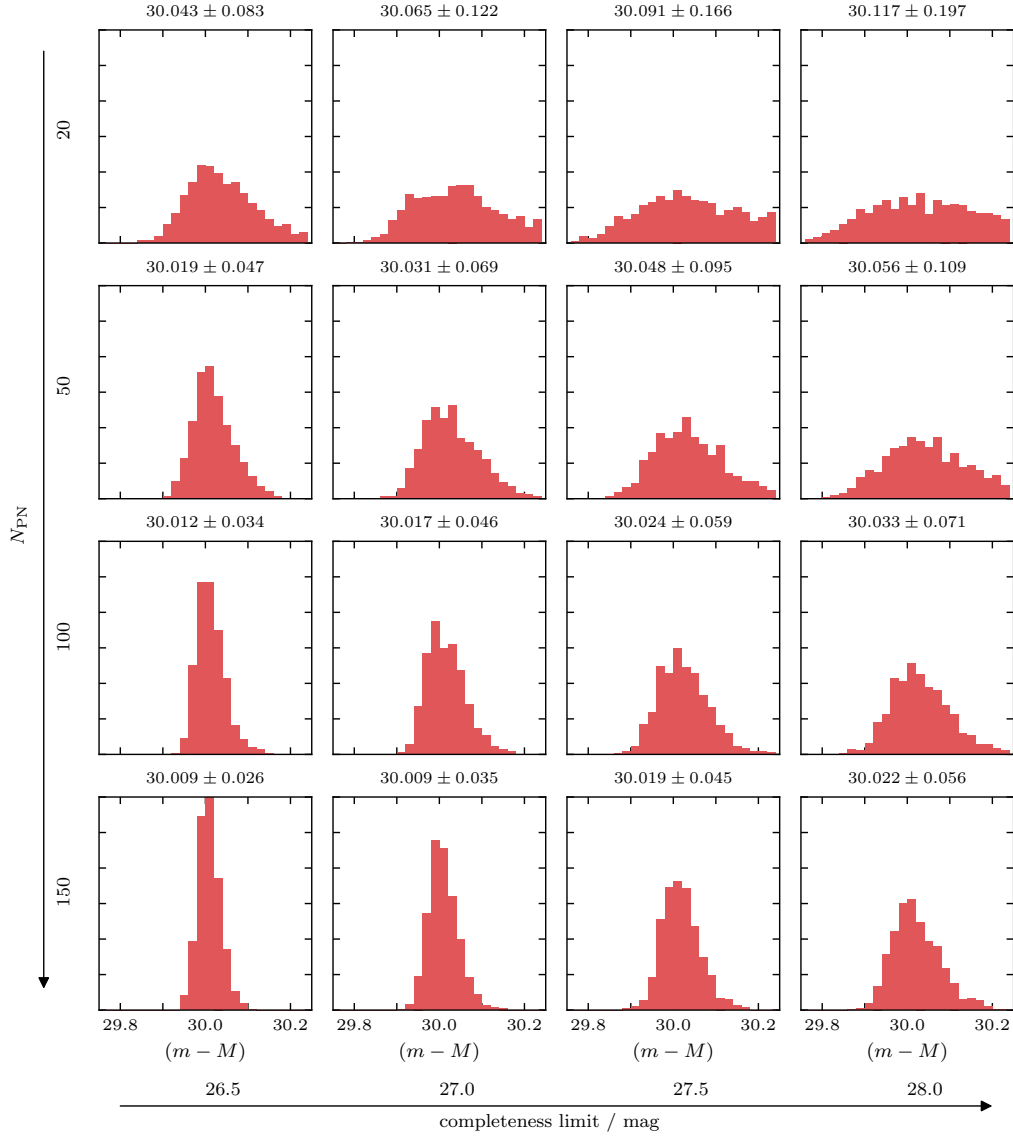


Figure A.20: A demonstration of the precision that can be achieved with the PNLF, depending on the sample size and part of the PNLF that is sampled. We sample the PNLF for a distance modulus of $(m - M) = 30$ mag. The sample size N_{PN} increases from 20, 50, 100 to 150 and the completeness limit increases from 26.5 to 28 mag in 0.5 mag increments (meaning that we sample 1.0 to 2.5 mag of the PNLF). The precision increases with sample size and even a small sample of 20 PNe is enough to achieve a precision around 5 percent. When we only sample a small part of the luminosity function, the sample is concentrated at the bright cut-off and the distance is well constrained. With a fainter completeness limit, the sample is distributed more and the more sparsely sampled cut-off yields a less precise distance.

DEFINING AN EVOLUTIONARY SEQUENCE

This chapter contains the appendix for Chapter 7 (Paper II).

B.1 MODEL PREDICTIONS

In Figure B.1 we show the age evolution of $H\alpha$ /FUV and $EW(H\alpha)$ as predicted by STARBURST99 (Leitherer et al. 2014), following a Kroupa (2001) IMF. Independent of metallicity, we see a plateau for the first ~ 3 Myr after which both the $H\alpha$ /FUV and the $EW(H\alpha)$ decrease almost monotonically.

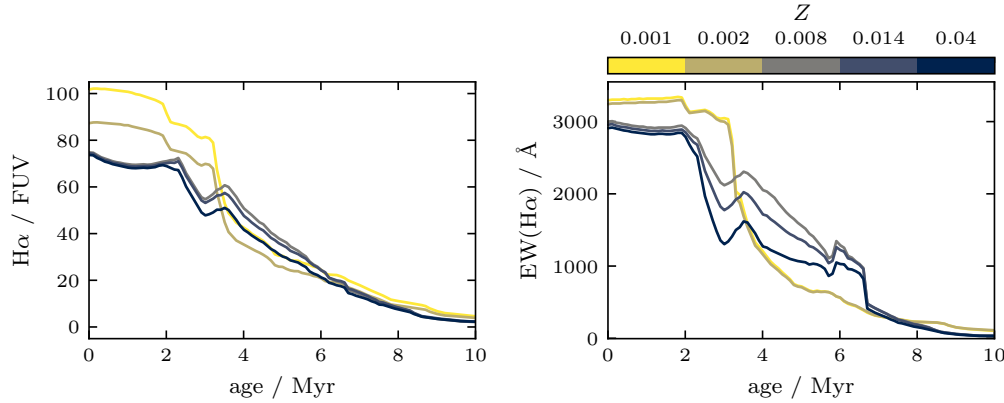


Figure B.1: $H\alpha$ /FUV flux ratio and $EW(H\alpha)$ as a function of cluster age for different metallicities as predicted by STARBURST99.

B.2 BACKGROUND CORRECTED EW

We provide the raw and background corrected $EW(H\alpha)$ and $EW(H\beta)$ for the full nebula catalogue from Groves et al. (2023). The fluxes are measured in the rest frame wavelength intervals listed in Table B.1, following the procedure from Westfall et al. (2019). To account for the contribution of older stars in the stellar disk, we subtracted the background from an annulus with three times the size of the H II region. For roughly one third of the H II regions, the background subtraction is complicated by neighbouring H II regions, and we mask pixels that fall in other nebulae. Because the continuum is relatively smooth, this should not be an issue, except for the 0.5 percent H II regions that are completely surrounded by other nebulae. Those objects are hence excluded from the analysis of the EW.

Table B.1: Intervals for the equivalent width measurement.

	H α	H β
Line	6557.6 to 6571.35 Å	4847.9 to 4876.6 Å
Continuum low	6483.0 to 6513.0 Å	4827.9 to 4847.9 Å
Continuum high	6623.0 to 6653.0 Å	4876.6 to 4891.6 Å

B.3 DENSITY AND TEMPERATURE

As discussed in Section 7.2, we assume a fixed temperature of 8000 K to derive the density. To validate this assumption, we compare the values derived with a fixed temperature versus those fitted simultaneously with the temperature. Auroral lines are faint, and only 840 H II regions are detected with a S/N > 10 in the temperature sensitive [N II] λ 5754 line. For this sub-sample, we use PYNEB (Luridiana et al. 2015) to derive the electron density and the temperature simultaneously from the [S II] λ 6731/6717 and [N II] λ 5755/6548 ratios. We also derive the density with a fixed temperature of 8000 K and compare the two in Figure B.2. As evident in the figure, the small variations in temperature do not affect the measured density significantly. While the majority of these H II regions are significantly different from the low density limit, the sample is only half the size of the full catalogue. Due to the large difference in sample size and the small difference in the derived density we opt to use the densities derived with a fixed temperature in our analysis.

B.4 OBJECTS IN UNMATCHED H II REGIONS

As stated in Section 7.3.2, almost 9000 H II regions do neither contain a stellar association nor a compact star cluster. However, we do find a DOLPHOT peak in 63.7 per cent of them. In Figure B.3, we compare the H α luminosity of those unmatched H II regions with the V-band magnitudes of the DOLPHOT peaks and overplot models for single stars from Martins et al. (2005). We find that most unmatched H II regions fall in a regime where they could be ionized by a single massive star.

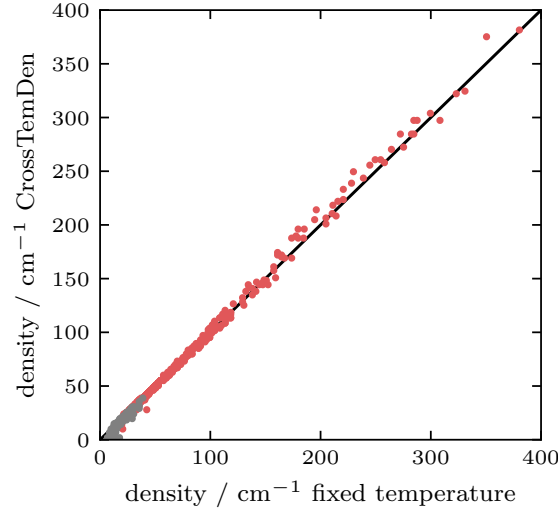


Figure B.2: Comparison between the densities derived at a fixed temperature of 8000 K and those derived by fitting density and temperature simultaneously. The grey points are indistinguishable from the low-density limit.

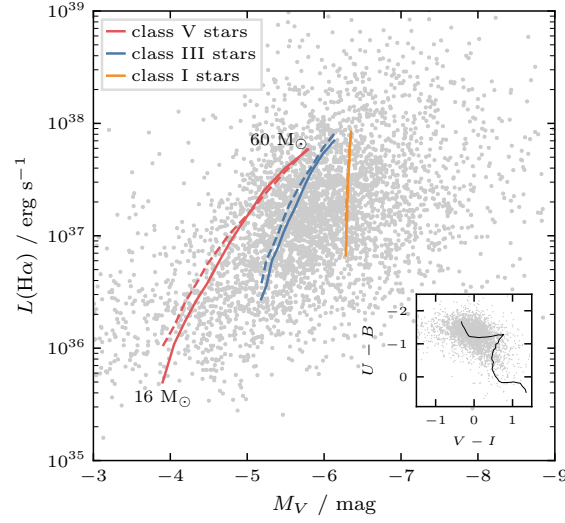


Figure B.3: $H\alpha$ luminosity and absolute V band magnitude for the DOLPHOT peaks that are matched to our previously unmatched H II regions. Overplotted are models for different class O-type stars from Martins, Schaerer & Hillier (2005). The ionizing photon flux is converted to an $H\alpha$ luminosity via $Q(H\alpha) = 7.31 \cdot 10^{11} L(H\alpha) s^{-1}$ (Kennicutt 1998). The solid lines show their theoretical calibration of T_{eff} and the dashed lines the observational calibration. The insert shows a colour-colour diagram of the sample with the track of a single stellar population at solar metallicity for reference (Bruzual & Charlot 2003).

B.5 BINNED AGE HISTOGRAMS

Figure B.4 shows the distribution of stellar association ages for different sub-samples. In Figure B.4 we separate the sample based on the overlap with the H II regions. We find that associations that overlap with an H II region are significantly younger than those that are isolated. In Figure B.5 we separate the sample based on the first, second and third percentile in the un-corrected $\text{EW}(\text{H}\alpha)$. The sample with the highest $\text{EW}(\text{H}\alpha)$ has on average the lowest SED ages.

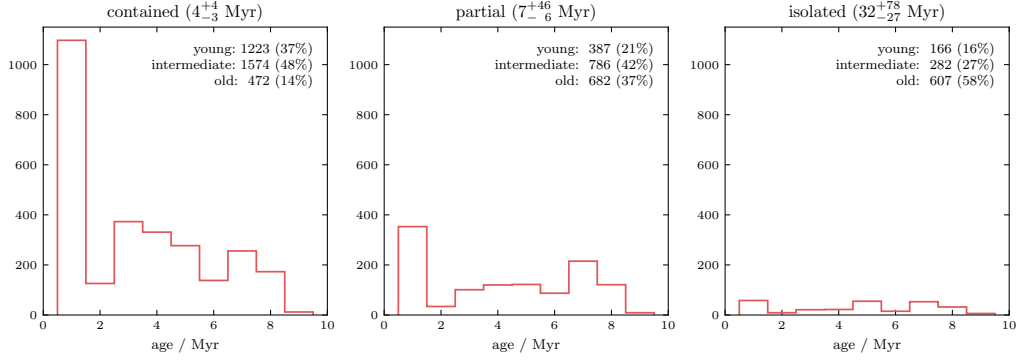


Figure B.4: Age distribution of the associations with isolated, partial and contained overlap (only the first 10 Myr are shown here). For each subsample we show the median age with the uncertainty taken from the 68 percent interval of the data and the number of young (≤ 2 Myr), intermediate (between 2 Myr and 10 Myr) and old (> 10 Myr) associations. We only include the 6027 associations that are more massive than $10^4 M_{\odot}$.

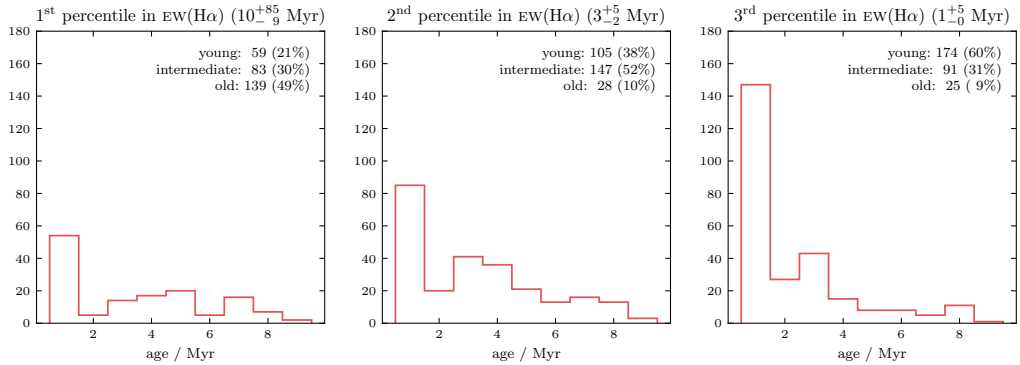


Figure B.5: Age distribution, similar to Figure B.4, but based on the $\text{EW}(\text{H}\alpha)$. We only include associations that are more massive than $10^4 M_{\odot}$ and apply the same S/N cut described in Section 7.4.1, leaving us with 854 objects.

ESCAPE FRACTION OF IONIZING PHOTONS

This chapter contains the appendix for Chapter 8 (Paper III).

C.1 IONIZING PHOTON FLUX FROM DIFFERENT MODELS

In Figure C.1 we compare the ionizing photon flux $Q(\text{H}^0)$ for different models (top panel) and metallicities (lower panel). A detailed listing of the models can be found in Table C.1. These plots demonstrate that, at constant metallicity, all models except the one with rotation are similar for the first few Myr. Around 6 Myr, however, the effects of binarity becomes relevant as its flux decreases much slower than the other models. For the impact of metallicity, we include the low-metallicity models from STARBURST99, based on the stellar model by Georgy et al. (2013). The variation between models that are close to solar metallicity is practically negligible, while the predicted flux is slightly higher at lower metallicities. Given that the flux decreases by an order of magnitude every few million years and the uncertainty of the age is comparatively high, the choice of model is likely to have only a minor influence.

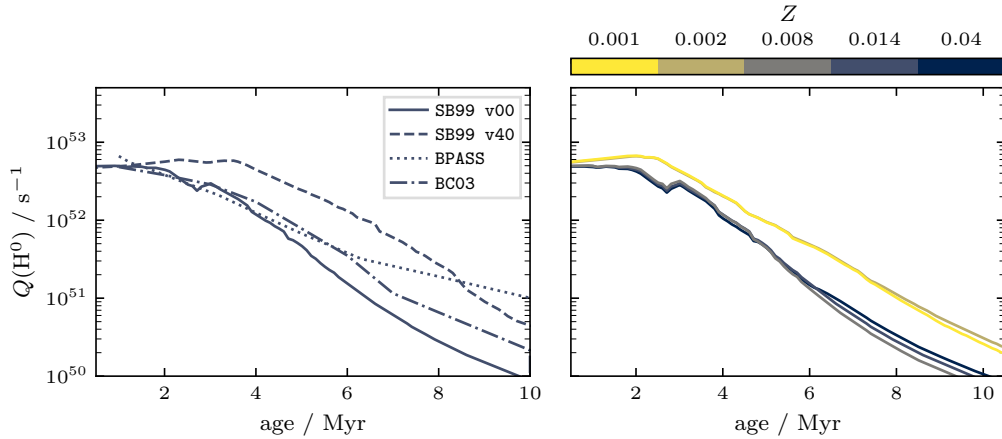


Figure C.1: Ionizing photon flux $Q(\text{H}^0)$ of a star cluster with a mass of $10^6 M_\odot$ as a function of age. The top panel shows the values predicted by different models, all at solar metallicity, while the lower panel shows the SB99 v00 model at various metallicities.

Table C.1: Codes used to compute the ionizing photon flux $Q(\text{H}^0)$.

Name	Code	Stellar model	IMF
BCo3	CIGALE (Boquien et al. 2019)	BCo3 (Bruzual & Charlot 2003)	(Chabrier 2003)
SB99 v00	STARBURST99 (Leitherer et al. 2014)	Geneva without rotation (Ekström et al. 2012)	(Kroupa 2001)
SB99 v40	STARBURST99 (Leitherer et al. 2014)	Geneva with rotation (Ekström et al. 2012)	(Kroupa 2001)
BPASS	BPASS (Eldridge & Stanway 2009)		(Kroupa 2001)
SLUG	SLUG (da Silva, Fumagalli & Krumholz 2012)	Geneva (Ekström et al. 2012)	(Kroupa 2001)

C.2 COMPARISON WITH TEH ET. AL (2023)

The stellar catalogues presented in Section 8.2.2 rely on the assumption that the clusters and associations in our sample are massive enough such that a deterministic relationship between the physical and photometric properties of the stellar populations exists. While we try to ensure this by focusing our analysis on high mass objects from the robust subsamples defined in Section 8.2.4, this means that we are omitting a large part of the sample. Another approach is to employ a stochastic stellar population synthesis code like SLUG (Stochastically Lighting Up Galaxies; da Silva et al. 2012, 2014) to infer a distribution of possible values for $Q(\text{H}^0)$.

A particularly relevant literature result is the study by Teh et al. (2023), who measured the escape fraction of NGC 0628 (which is also in our sample) with the help of SLUG. They combined H II regions from the SIGNALS survey (Rousseau-Nepton et al. 2018) with star clusters from LEGUS (Grasha et al. 2015; Adamo et al. 2017). Based on this, they measured a relatively low value of $f_{\text{esc}} = 9 \pm 6$ percent and a large number of their H II regions have a negative escape fraction. The striking difference to our results asks for a closer comparison.

We start by comparing the H II regions from MUSE to those from SITELLE and 643 of them match within 0.8 arcsec. In Figure C.2 we compare the measured fluxes, and while the ones identified by SIGNALS are more luminous at the fainter end, the ones from PHANGS are brighter for the most luminous objects.

To compare the stellar component, we take the five-band filters from our association catalogue and redo the SED fit with SLUG, following the procedure described in Teh et al. (2023): First, a library of 10^7 stellar populations is generated to account for the

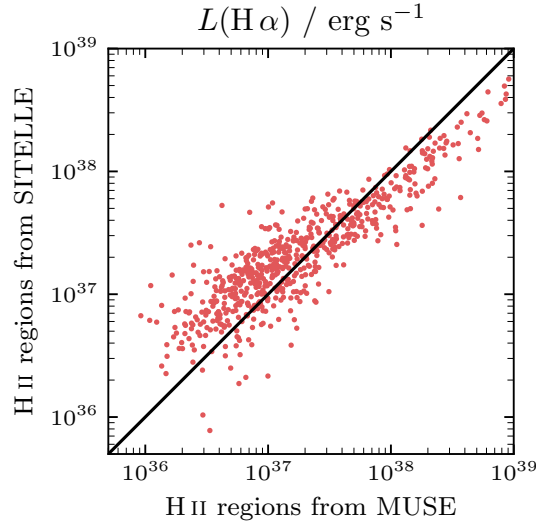


Figure C.2: Comparisons between the H II regions from PHANGS (observed with MUSE Groves et al. 2023) and the ones from SIGNALS (observed with SITELLE Rousseau-Nepton et al. 2019).

effects of stochasticity in the IMF. For this, we adopt the following physical parameters: Solar metallicity, the IMF by Kroupa (2001), Geneva evolutionary tracks (Ekström et al. 2012; Georgy et al. 2012), STARBURST99 stellar atmosphere models (Leitherer et al. 1999; Vázquez & Leitherer 2005), and a Milky Way extinction curve (Cardelli et al. 1989). Then we use a modified version of `cluster_slug` (Della Bruna et al. 2021, 2022a; Teh et al. 2023) to compute the posterior *probability density function* (PDF) of the age, mass, and ionizing luminosity (see Krumholz et al. 2015, for more details). We represent the $Q(H^0)$ values with the median (50th percentile) and the 1σ uncertainty with the 16th–84th percentile of the PDF distribution.

The result is shown in Figure C.3. There are significant deviations in the derived ages, with those from SLUG generally being older. This is particularly problematic in the case of very young stars. While the lower limit of the age distribution is often consistent with being 3 Myr or younger, the median rarely is. Since the majority of the ionizing photons are emitted during this period, the predicted fluxes are generally lower. As far as the masses are concerned, there is a large scatter, but not systematic differences.

It is expected that SLUG would deliver different values at the faint end, due to stochastic sampling. Overall, it predicts systematically lower $Q(H^0)$, even at the bright end, where this should not be an issue. This contributes significantly to the lower escape fractions found by previous studies that are based on SLUG like Della Bruna et al. (2022a) or Teh et al. (2023).

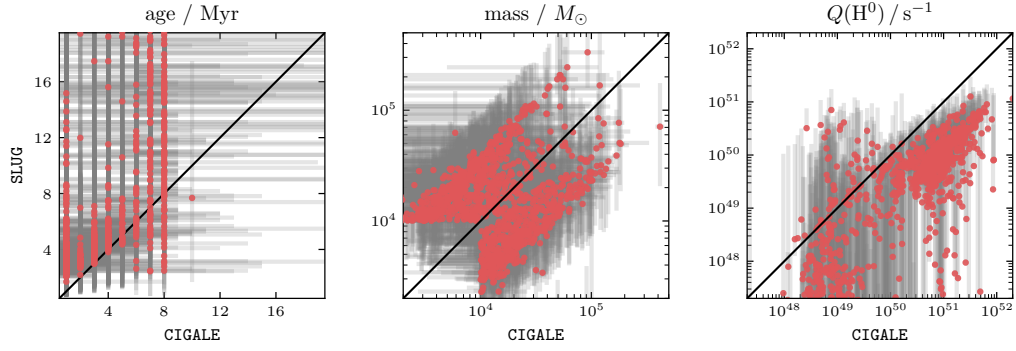


Figure C.3: We ran SLUG on the associations in the matched catalogue. For the comparison we define a robust subsample, where either the CIGALE or the SLUG age is below 8 Myr and the mass predicted by either is above $10^4 M_{\odot}$. The masses show a wide scatter in both directions, but older ages are generally predicted. As a result, the ionizing photon flux is systematically lower than the one from CIGALE.

C.3 COMPLETENESS ISSUES

In Section 8.3.4 we consider the entire ionizing photon budget of each galaxy. In this scenario, different completeness limits for the H II regions and stellar populations could bias those results. We therefore estimate the completeness for both to assess the contribution of missed objects. In Figure C.4 we show the normalised cumulative ionizing photon flux to estimate the contribution of faint H II regions and low mass associations.

In the case of IC 5332, NGC 3351 and NGC 5068, faint H II regions (below the completeness limit from Santoro et al. 2022) contribute more than 10 per cent of the total ionizing photon flux. The potentially missing H α flux could increase the escape fraction, but while the values of the first two galaxies are above the average (see Table 8.3), both also appear on the stellar side, weakening this conclusion.

For IC 5332, NGC 1300, NGC 1433, NGC 1512, NGC 2835 and NGC 3351, the contribution of low mass associations ($< 5000 M_{\odot}$) is larger than 10 per cent. Contrary to the case of the H II regions, the potentially overlooked clusters would lead to a decrease of the escape fraction. While half of galaxies have lower escape fractions, the other half, including the two galaxies mentioned above, have higher values.

For the 12 other galaxies, the completeness limit should not be an issue, and even for the seven discussed above, there is no noticeable impact.

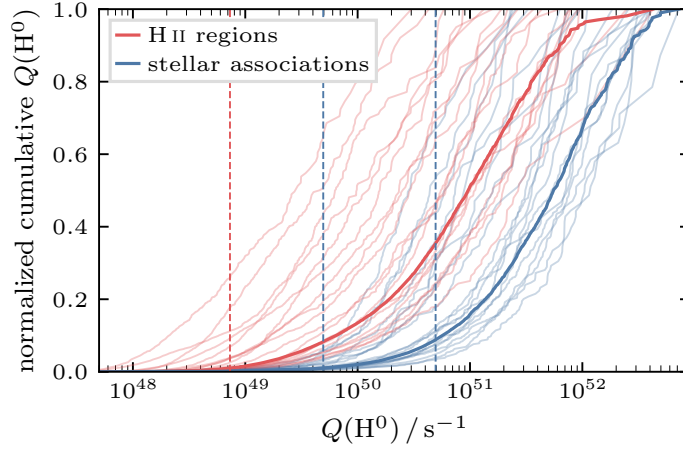


Figure C.4: Normalized cumulative $Q(H^0)$ as a function of $Q(H^0)$ for H II regions and stellar associations. The individual galaxies are shown in pastel and the sum of all galaxies is shown in bold colours. We only show the H II regions and associations in the overlapping coverage that do not fall in the centre (corresponding to the sample used in Section 8.3.4). The H II region completeness limit from Santoro et al. (2022), corresponding to a luminosity of $10^{37} \text{ erg s}^{-1}$ is indicated by the left red dashed line. The two blue lines corresponds to a $10^3 M_{\odot}$ and $10^4 M_{\odot}$ cluster that is 1 Myr old.

REFERENCES

- Abbott, B. P. et al., 2017, *ApJ*, **848**, L12
- Adamo, A. et al., 2015, *MNRAS*, **452**, 246
- Adamo, A. et al., 2017, *ApJ*, **841**, 131
- Adamo, A. et al., 2020, *Space Sci. Rev.*, **216**, 69
- Airy, G. B., 1835, *Trans. Camb. Philos. Soc.*, **5**, 283
- Amanullah, R. et al., 2010, *ApJ*, **716**, 712
- Anand, G. S. et al., 2021, *MNRAS*, **501**, 3621
- Ångström, J. A., 1869, *AnP*, **213**, 161
- Arnaboldi, M., Longobardi, A., Gerhard, O., Okamura, S., 2013, in »*Advancing the Physics of Cosmic Distances*«, Vol. 289, p. 287
- Asplund, M., Grevesse, N., Sauval, A. J., Scott, P., 2009, *ARA&A*, **47**, 481
- Baade, W., 1944, *ApJ*, **100**, 137
- Bacon, R. et al., 2010, in »*Ground-based and Airborne Instrumentation for Astronomy III*«, Vol. 7735, p. 773508
- Bacon, R. et al., 2017, *A&A*, **608**, A1
- Baldwin, J. A., Phillips, M. M., Terlevich, R., 1981, *PASP*, **93**, 5
- Barbarino, C. et al., 2015, *MNRAS*, **448**, 2312
- Barnes, A. T. et al., 2021, *MNRAS*, **508**, 5362
- Barnes, A. T. et al., 2022, *A&A*, **662**, L6
- Barnes, A. T. et al., submitted, »*The PHANGS-MUSE/HST-H α Nebulae Catalogue*«
- Bartel, N. et al., 1985, *Nature*, **318**, 25
- Bartel, N., Bietenholz, M. F., 2003, *ApJ*, **591**, 301
- Beckwith, S. V. W. et al., 2006, *AJ*, **132**, 1729
- Beech, M., 2019, »*Introducing the Stars*«, Springer
- Belfiore, F. et al., 2019, *AJ*, **158**, 160
- Belfiore, F. et al., 2022, *A&A*, **659**, A26
- Belfiore, F. et al., 2025, *A&A*, **694**, A212
- Bernal, J. L., Verde, L., Riess, A. G., 2016, *J. Cosmol. Astropart. Phys.*, **10**, 019
- Bessel, F. W., 1838, *Astron. Nachr.*, **16**, 65
- Bessell, M. S., Castelli, F., Plez, B., 1998, *A&A*, **333**, 231
- Bethe, H. A., 1939, *Phys. Rev.*, **55**, 103
- Beucher, S., Lantuéjoul, C., 1979, in »*International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation*«
- Blair, W. P., Long, K. S., 2004, *ApJS*, **155**, 101
- Bobylev, V. V., Bajkova, A. T., 2017, *Astron. Lett.*, **43**, 304
- Bohr, N., 1913, *Philos. Mag.*, **26**, 857
- Bojčić, I. S. et al., 2021, *MNRAS*, **503**, 2887
- Boquien, M. et al., 2019, *A&A*, **622**, A103
- Bose, S., Kumar, B., 2014, *ApJ*, **782**, 98

- Bottinelli, L., Gouguenheim, L., Paturel, G., de Vaucouleurs, G., 1984, *A&AS*, **56**, 381
- Bottinelli, L., Gouguenheim, L., Paturel, G., de Vaucouleurs, G., 1985, *A&AS*, **59**, 43
- Bottinelli, L., Gouguenheim, L., Paturel, G., Teerikorpi, P., 1986, *A&A*, **156**, 157
- Boyce, H. et al., 2017, *ApJ*, **846**, 14
- Boyle, W., Smith, G., 1970, *Bell System Technical Journal*, **49**, 587
- Bradley, L. et al., 2019, »*Astropy/Photutils: vo.6*«, *Zenodo*
- Branch, D. et al., 1981, *ApJ*, **244**, 780
- Brazzini, M. et al., 2024, *A&A*, **691**, A173
- Bromm, V., Larson, R. B., 2004, *ARA&A*, **42**, 79
- Bruzual, G., Charlot, S., 2003, *MNRAS*, **344**, 1000
- Burbidge, E. M., Burbidge, G. R., Fowler, W. A., Hoyle, F., 1957, *Rev. Mod. Phys.*, **29**, 547
- Burns, C. R. et al., 2011, *AJ*, **141**, 19
- Burrows, A. S., Ostriker, J. P., 2014, *Proc. Natl. Acad. Sci*, **111**, 2409
- Buzzoni, A., Arnaboldi, M., Corradi, R. L. M., 2006, *MNRAS*, **368**, 877
- Calzetti, D. et al., 2000, *ApJ*, **533**, 682
- Calzetti, D. et al., 2015, *AJ*, **149**, 51
- Cardelli, J. A., Clayton, G. C., Mathis, J. S., 1989, *ApJ*, **345**, 245
- Carroll, B. W., Ostriker, D. A., 2017, »*An introduction to modern astrophysics*«, *Cambridge Univ. Press*
- Cerviño, M., Luridiana, V., 2006, *A&A*, **451**, 475
- Chabrier, G., 2003, *PASP*, **115**, 763
- Chandar, R. et al., 2025, *AJ*, **169**, 150
- Charlot, S., Fall, S. M., 2000, *ApJ*, **539**, 718
- Charlot, S., Longhetti, M., 2001, *MNRAS*, **323**, 887
- Chevance, M. et al., 2022, *MNRAS*, **509**, 272
- Chisholm, J. et al., 2022, *MNRAS*, **517**, 5104
- Choi, J. et al., 2016, *ApJ*, **823**, 102
- Chromey, F. R., 2010, »*To Measure the Sky*«, *Cambridge Univ. Press*
- Ciardullo, R., Jacoby, G. H., Ford, H. C., Neill, J. D., 1989a, *ApJ*, **339**, 53
- Ciardullo, R., Jacoby, G. H., Ford, H. C., 1989b, *ApJ*, **344**, 715
- Ciardullo, R., Jacoby, G. H., Tonry, J. L., 1993, *ApJ*, **419**, 479
- Ciardullo, R. et al., 2002, *ApJ*, **577**, 31
- Ciardullo, R., 2012, *Ap&SS*, **341**, 151
- Ciardullo, R., 2013, in »*Advancing the Physics of Cosmic Distances*«, *Vol. 289*, p. 247
- Collins II, G. W., 1989, »*The fundamentals of stellar astrophysics*«, *W.H. Freeman*
- Congiu, E. et al., 2023, *A&A*, **672**, A148
- Congiu, E. et al., 2025, *A&A*, **700**, A125
- Conselice, C. J., Wilkinson, A., Duncan, K., Mortlock, A., 2016, *ApJ*, **830**, 83
- Copetti, M. V. F., Pastoriza, M. G., Dottori, H. A., 1986, *A&A*, **156**, 111
- Cote, P. et al., 2019, in »*Canadian Astronomy Long Range Plan*«, *Vol. 2020*, p. 18
- Courtois, H. M., Tully, R. B., 2012, *ApJ*, **749**, 174
- Cox, D. P., 2005, *ARA&A*, **43**, 337
- Crowther, P. A., 2007, *ARA&A*, **45**, 177

da Silva, R. L., Fumagalli, M., Krumholz, M., 2012, *ApJ*, **745**, 145

da Silva, R. L., Fumagalli, M., Krumholz, M. R., 2014, *MNRAS*, **444**, 3275

Dalcanton, J. J. et al., 2009, *ApJS*, **183**, 67

Dale, J. E., Ercolano, B., Bonnell, I. A., 2012, *MNRAS*, **424**, 377

Dale, J. E., Ercolano, B., Bonnell, I. A., 2013, *MNRAS*, **430**, 234

Dale, J. E., Ngoumou, J., Ercolano, B., Bonnell, I. A., 2014, *MNRAS*, **442**, 694

Davidge, T. J., Pritchett, C. J., 1990, *AJ*, **100**, 102

Davidge, T. J., Le Fevre, O., Clark, C. C., 1991, *ApJ*, **370**, 559

Davis, B. D., Ciardullo, R., Feldmeier, J. J., Jacoby, G. H., 2018, *Res. Notes AAS*, **2**, 32

Dayal, P., Ferrara, A., 2018, *Physics Reports*, **780**, 1

de Vaucouleurs, G. et al., 1981a, *ApJ*, **248**, 408

de Vaucouleurs, G. et al., 1981b, *PASP*, **93**, 36

Deger, S. et al., 2022, *MNRAS*, **510**, 32

Della Bruna, L. et al., 2021, *A&A*, **650**, A103

Della Bruna, L. et al., 2022a, *A&A*, **666**, A29

Della Bruna, L. et al., 2022b, *A&A*, **660**, A77

den Brok, M. et al., 2020, *MNRAS*, **491**, 4089

Dhungana, G. et al., 2016, *ApJ*, **822**, 6

Diaz, A. I. et al., 1991, *MNRAS*, **253**, 245

Dolphin, A. E., Kennicutt Robert C., J., 2002, *AJ*, **123**, 207

Dolphin, A. E., 2000, *PASP*, **112**, 1383

Dopita, M. A., Evans, I. N., 1986, *ApJ*, **307**, 431

Dopita, M. A., Jacoby, G. H., Vassiliadis, E., 1992, *ApJ*, **389**, 27

Dopita, M. A. et al., 2006, *ApJ*, **647**, 244

Dopita, M. A. et al., 2014, *Ap&SS*, **350**, 741

Doran, E. I. et al., 2013, *A&A*, **558**, A134

Dotter, A., 2016, *ApJS*, **222**, 8

Dottori, H. A., 1981, *Ap&SS*, **80**, 267

Draine, B. T., Lee, H. M., 1984, *ApJ*, **285**, 89

Draine, B. T., 2011, »*Physics of the Interstellar and Intergalactic Medium*«, *Princeton Univ. Press*

Drissen, L. et al., 2019, *MNRAS*, **485**, 3930

Eddington, A. S., 1920, *Nature*, **106**, 14

Efremov, Y. N., Elmegreen, B. G., 1998, *MNRAS*, **299**, 588

Egorov, O. V., Lozinskaya, T. A., Moiseev, A. V., Smirnov-Pinchukov, G. V., 2018, *MNRAS*, **478**, 3386

Einstein, A., 1905, *AnP*, **323**, 639

Ekholm, T. et al., 2000, *A&A*, **355**, 835

Ekström, S. et al., 2012, *A&A*, **537**, A146

Eldridge, J. J., Stanway, E. R., 2009, *MNRAS*, **400**, 1019

Elmegreen, B. G., 2008, *ApJ*, **672**, 1006

Emsellem, E. et al., 2022, *A&A*, **659**, A191

Espinosa-Ponce, C. et al., 2022, *MNRAS*, **512**, 3436

- European Southern Observatory, 2025, »*Technical Report on INNA project proposal to SEIA – Executive Summary*«, www.eso.org
- Ewen, H. I., Purcell, E. M., 1951, *Nature*, **168**, 356
- Faesi, C. M. et al., 2014, *ApJ*, **789**, 81
- Farhang, A. et al., 2019, *Nature Astronomy*, **3**, 922
- Feldmeier, J. J., Ciardullo, R., Jacoby, G. H., 1997, *ApJ*, **479**, 231
- Feldmeier, J. J., Jacoby, G. H., Phillips, M. M., 2007, *ApJ*, **657**, 76
- Ferland, G. J. et al., 2017, *Rev. Mex. Astron. Astrofís.*, **53**, 385
- Fernandes, R. C., Leão, J. R. S., Lacerda, R. R., 2003, *MNRAS*, **340**, 29
- Fernie, J. D., Brecher, K., Chaisson, E. J., Aller, L. H., 2025, »*star*«, *Encyclopedia Britannica*
- Ferrarese, L. et al., 2000a, *ApJS*, **128**, 431
- Ferrarese, L. et al., 2000b, *ApJ*, **529**, 745
- Ferrière, K. M., 2001, *Rev. Mod. Phys.*, **73**, 1031
- Field, G. B., Goldsmith, D. W., Habing, H. J., 1969, *ApJ*, **155**, L149
- Flury, S. R. et al., 2022a, *ApJS*, **260**, 1
- Flury, S. R. et al., 2022b, *ApJ*, **930**, 126
- Folatelli, G. et al., 2013, *ApJ*, **773**, 53
- Ford, H. C., Jenner, D. C., 1978, in »*Bulletin of the AAS*«, Vol. 10, p. 665
- Foreman-Mackey, D., 2014, »*Blog Post: Mixture Models*«, *Zenodo*
- Fouesneau, M., Lançon, A., 2010, *A&A*, **521**, A22
- Fouesneau, M., Lançon, A., Chandar, R., Whitmore, B. C., 2012, *ApJ*, **750**, 60
- Franchetti, N. A. et al., 2012, *AJ*, **143**, 85
- Fraunhofer, J., 1817, *AnP*, **56**, 264
- Freedman, W. L. et al., 2001, *ApJ*, **553**, 47
- Freedman, W. L., 2021, *ApJ*, **919**, 16
- Freundlich, J., 2024, *Fundamental Plasma Physics*, **11**, 100059
- Fusco, T. et al., 2020, *A&A*, **635**, A208
- Gadotti, D. A. et al., 2019, *MNRAS*, **482**, 506
- Gaia Collaboration et al., 2018, *A&A*, **616**, A1
- Gainor, C., 2020, »*Not yet imagined: a study of Hubble Space Telescope operations*«, *NASA*
- Galán-de Anta, P. M. et al., 2021, *A&A*, **652**, A109
- Gamow, G., 1948a, *Nature*, **162**, 680
- Gamow, G., 1948b, *Phys. Rev.*, **74**, 505
- Georgiev, T. B., Bilkina, B. I., Dencheva, N. M., 1997, *Astron. Lett.*, **23**, 644
- Georgy, C. et al., 2012, *A&A*, **542**, A29
- Georgy, C. et al., 2013, *A&A*, **558**, A103
- Gerasimov, I. S. et al., 2022, *MNRAS*, **517**, 4968
- Gibson, B. K. et al., 2003, *Publ. Astron. Soc. Australia*, **20**, 401
- Girardi, L., Bressan, A., Bertelli, G., Chiosi, C., 2000, *A&AS*, **141**, 371
- Giraud, E., 1985, *A&A*, **153**, 125
- Girichidis, P. et al., 2020, *Space Sci. Rev.*, **216**, 68
- Glover, S. C. O., Clark, P. C., 2014, *MNRAS*, **437**, 9
- Goddard, Q. E., Bastian, N., Kennicutt, R. C., 2010, *MNRAS*, **405**, 857

Goldsmith, P. F., Li, D., Krčo, M., 2007, *ApJ*, **654**, 273
 Gordon, K. D. et al., 2003, *ApJ*, **594**, 279
 Gouliermis, D. A., 2018, *PASP*, **130**, 072001
 Grasha, K. et al., 2015, *ApJ*, **815**, 93
 Grasha, K. et al., 2022, *ApJ*, **929**, 118
 Gratton, R., Sneden, C., Carretta, E., 2004, *ARA&A*, **42**, 385
 Grevesse, N., Noels, A., Sauval, A. J., 1996, in »Cosmic Abundances«, Vol. 99, p. 117
 Grevesse, N., Sauval, A. J., 1998, *Space Sci. Rev.*, **85**, 161
 Groves, B. et al., 2023, *MNRAS*, **520**, 4902
 Habets, G. M. H. J., Heintze, J. R. W., 1981, *A&AS*, **46**, 193
 Habjan, E. et al., in preparation, »H II region Metallicities in Nearby Galaxies with MUSE and SITELLE«
 Haffner, L. M. et al., 2009, *Rev. Mod. Phys.*, **81**, 969
 Hagen, N., Kudenov, M. W., 2013, *OptEn*, **52**, 090901
 Haid, S. et al., 2018, *MNRAS*, **478**, 4799
 Halm, J., 1911, *MNRAS*, **71**, 638
 Hanish, D. J. et al., 2010, *ApJ*, **725**, 2029
 Hannon, S. et al., 2019, *MNRAS*, **490**, 4648
 Hannon, S. et al., 2022, *MNRAS*, **512**, 1294
 Hannon, S. et al., 2023, *MNRAS*, **526**, 2991
 Harkins, W. D., Wilson, E. D., 1915, *PNAS*, **1**, 276
 Hartke, J. et al., 2017, *A&A*, **603**, A104
 Hartnett, K. et al., 2018, »Hubble: An Overview of the Space Telescope«, NASA's Goddard Space Flight Center
 Hassani, H. et al., 2024, *ApJS*, **271**, 2
 Hathi, N. P., 2024, »ACS Data Handbook v. 13.0«, Baltimore: STScI
 Heger, A., Woosley, S. E., 2002, *ApJ*, **567**, 532
 Hendry, M. A. et al., 2005, *MNRAS*, **359**, 906
 Hermanowicz, M. T., Kennicutt, R. C., Eldridge, J. J., 2013, *MNRAS*, **432**, 3097
 Herrmann, K. A., Ciardullo, R., Feldmeier, J. J., Vinciguerra, M., 2008, *ApJ*, **683**, 630
 Herschel, W., 1786, *Philos. Trans. R. Soc. Lond. Ser. I*, **76**, 457
 Herwig, F., 2005, *ARA&A*, **43**, 435
 Hewish, A. et al., 1968, *Nature*, **217**, 709
 Hippler, S., 2019, *J. Astron. Instrum.*, **8**, 1950001
 Hirschi, R. et al., 2008, in »Massive Stars as Cosmic Engines«, Vol. 250, p. 217
 Ho, I. T. et al., 2019, *ApJ*, **885**, L31
 Hoeflich, P. et al., 2017, *ApJ*, **846**, 58
 Hogg, D. W., Bovy, J., Lang, D., 2010, *arXiv*, 1008.4686
 Hollyhead, K. et al., 2015, *MNRAS*, **449**, 1106
 Hopkins, P. F. et al., 2014, *MNRAS*, **445**, 581
 Howard, C., Pudritz, R., Klessen, R., 2017, *ApJ*, **834**, 40
 Howard, C. S., Pudritz, R. E., Harris, W. E., Klessen, R. S., 2018, *MNRAS*, **475**, 3121
 Howell, S. B., 1989, *PASP*, **101**, 616

Hoyle, F., Ellis, G. R. A., 1963, [Aust. J. Phys.](#), **16**, 1
 Hubble, E. P., 1929a, [PNAS](#), **15**, 168
 Hubble, E. P., 1929b, [ApJ](#), **69**, 103
 Huggins, W., Miller, W. A., 1864a, *Philos. Trans. R. Soc. Lond. Ser. I*, **154**, 413
 Huggins, W., Miller, W. A., 1864b, *Philos. Trans. R. Soc. Lond. Ser. I*, **154**, 437
 Hummer, D. G., Storey, P. J., 1987, [MNRAS](#), **224**, 801
 Jacobs, B. A. et al., 2009, [AJ](#), **138**, 332
 Jacoby, G. H. et al., 1992, [PASP](#), **104**, 599
 Jacoby, G. H., Ciardullo, R., Harris, W. E., 1996, [ApJ](#), **462**, 1
 Jacoby, G. H. et al., 2024, [ApJS](#), **271**, 40
 Jacoby, G. H., 1989, [ApJ](#), **339**, 39
 Jacoby, G. H., 1997, in »*The Extragalactic Distance Scale*«, p. 197
 Jang, I. S., Lee, M. G., 2014, [ApJ](#), **792**, 52
 Jang, I. S., Lee, M. G., 2017, [ApJ](#), **836**, 74
 Jang, I. S. et al., 2018, [ApJ](#), **852**, 60
 Japelj, J. et al., 2017, [MNRAS](#), **468**, 389
 Jarrett, T. H. et al., 2019, [ApJS](#), **245**, 25
 Jha, S., Riess, A. G., Kirshner, R. P., 2007, [ApJ](#), **659**, 122
 Ji, X., Yan, R., 2022, [A&A](#), **659**, A112
 Johnson, J. L., Li, H., 2012, [ApJ](#), **751**, 81
 Johnson, J. A., Fields, B. D., Thompson, T. A., 2020, *Philos. Trans. R. Soc. Lond. Ser. A*, **378**, 20190301
 José, J., Iliadis, C., 2011, [Reports on Progress in Physics](#), **74**, 096901
 Kado-Fong, E., Kim, J.-G., Ostriker, E. C., Kim, C.-G., 2020, [ApJ](#), **897**, 143
 Kakiichi, K., Gronke, M., 2021, [ApJ](#), **908**, 30
 Kalberla, P. M. W. et al., 2005, [A&A](#), **440**, 775
 Kanbur, S. M. et al., 2003, [A&A](#), **411**, 361
 Kang, D. E. et al., 2022, [MNRAS](#), **512**, 617
 Karachentsev, I. D., Makarov, D. I., Kaisina, E. I., 2013, [AJ](#), **145**, 101
 Karachentsev, I. D. et al., 2017, [MNRAS](#), **469**, L113
 Kauffmann, G. et al., 2003, [MNRAS](#), **346**, 1055
 Kennicutt Robert C., J., 1998, [ARA&A](#), **36**, 189
 Kennicutt Jr., R. C., 1989, [ApJ](#), **344**, 685
 Kewley, L. J., Heisler, C. A., Dopita, M. A., Lumsden, S., 2001a, [ApJS](#), **132**, 37
 Kewley, L. J. et al., 2001b, [ApJ](#), **556**, 121
 Kewley, L. J., Dopita, M. A., 2002, [ApJS](#), **142**, 35
 Kewley, L. J., Groves, B., Kauffmann, G., Heckman, T., 2006, [MNRAS](#), **372**, 961
 Kewley, L. J., Nicholls, D. C., Sutherland, R. S., 2019, [ARA&A](#), **57**, 511
 Kim, C.-G., Ostriker, E. C., 2017, [ApJ](#), **846**, 133
 Kim, J.-G., Kim, W.-T., Ostriker, E. C., 2018, [ApJ](#), **859**, 68
 Kim, J.-G., Kim, W.-T., Ostriker, E. C., 2019, [ApJ](#), **883**, 102
 Kim, J. et al., 2021a, [MNRAS](#), **504**, 487
 Kim, J.-G., Ostriker, E. C., Filippova, N., 2021b, [ApJ](#), **911**, 128

Kim, J. et al., 2022, [MNRAS](#), **516**, 3006

Kim, J.-G., Gong, M., Kim, C.-G., Ostriker, E. C., 2023, [ApJS](#), **264**, 10

Kirchhoff, G., Bunsen, R., 1860, [AnP](#), **186**, 161

Kirchhoff, G., 1860, [AnP](#), **185**, 275

Klessen, R. S., Glover, S. C. O., 2016, [Saas-Fee Advanced Course](#), **43**, 85

Kobayashi, C., Karakas, A. I., Lugaro, M., 2020, [ApJ](#), **900**, 179

Kohlhase, C. E., Penzo, P. A., 1977, [Space Sci. Rev.](#), **21**, 77

Kourkchi, E., Tully, R. B., 2017, [ApJ](#), **843**, 16

Kourkchi, E. et al., 2020, [AJ](#), **159**, 67

Kragh, H., 2009, [Astron. Geophys.](#), **50**, 5.25

Kreckel, K. et al., 2017, [ApJ](#), **834**, 174

Kreckel, K. et al., 2019, [ApJ](#), **887**, 80

Kreckel, K. et al., 2020, [MNRAS](#), **499**, 193

Kreckel, K. et al., 2022, [A&A](#), **667**, A16

Krimigis, S. M. et al., 2013, [Science](#), **341**, 144

Kroupa, P., 2001, [MNRAS](#), **322**, 231

Kruijssen, J. M. D. et al., 2019, [Nature](#), **569**, 519

Kruijssen, J. M. D., 2012, [MNRAS](#), **426**, 3008

Krumholz, M. R., Tan, J. C., 2007, [ApJ](#), **654**, 304

Krumholz, M. R. et al., 2015, [ApJ](#), **812**, 147

Krumholz, M. R., McKee, C. F., Bland-Hawthorn, J., 2019, [ARA&A](#), **57**, 227

Krumholz, M. R., 2012, [ApJ](#), **759**, 9

Kuiper, G. P., 1938, [ApJ](#), **88**, 472

Kulkarni, S. R. et al., 2021, [arXiv](#), **2111.15608**

Kwitter, K. B., Henry, R. B. C., 2022, [PASP](#), **134**, 022001

Kwok, S., 2000, »*The Origin and Evolution of Planetary Nebulae*«, [Cambridge Univ. Press](#)

Kwok, S., 2005, [J. Korean Astron. Soc.](#), **38**, 271

Lada, C. J., Lada, E. A., 2003, [ARA&A](#), **41**, 57

Lagattuta, D. J. et al., 2013, [ApJ](#), **771**, 88

Lang, P. et al., 2020, [ApJ](#), **897**, 122

Larson, K. L. et al., 2023, [MNRAS](#), **523**, 6061

Leavitt, H. S., Pickering, E. C., 1912, [Harv. Obs. Circ.](#), **173**, 1

Lecroq, M. et al., 2024, [MNRAS](#), **527**, 9480

Lee, M. G., Freedman, W. L., Madore, B. F., 1993, [ApJ](#), **417**, 553

Lee, J. C. et al., 2009, [ApJ](#), **706**, 599

Lee, M. G., Jang, I. S., 2013, [ApJ](#), **773**, 13

Lee, J. C. et al., 2022, [ApJS](#), **258**, 10

Lee, J. C. et al., 2023, [ApJ](#), **944**, L17

Leitherer, C. et al., 1999, [ApJS](#), **123**, 3

Leitherer, C. et al., 2010, [ApJS](#), **189**, 309

Leitherer, C. et al., 2014, [ApJS](#), **212**, 14

Lejeune, T., Cuisinier, F., Buser, R., 1997, [A&AS](#), **125**, 229

Lemaître, G., 1931, [MNRAS](#), **91**, 483

Leroy, A. K. et al., 2019, [ApJS](#), 244, 24
 Leroy, A. K. et al., 2021a, [ApJS](#), 255, 19
 Leroy, A. K. et al., 2021b, [ApJS](#), 257, 43
 Levesque, E. M. et al., 2012, [ApJ](#), 751, 67
 Levesque, E. M., Leitherer, C., 2013, [ApJ](#), 779, 170
 Li, L. et al., 2018, [ApJ](#), 858, 75
 Li, Q. et al., 2021, [MNRAS](#), 507, 548
 Li, J. et al., 2024, [A&A](#), 690, A161
 Longobardi, A. et al., 2013, [A&A](#), 558, A42
 Lopez, L. A. et al., 2014, [ApJ](#), 795, 121
 Luridiana, V., Morisset, C., Shaw, R. A., 2015, [A&A](#), 573, A42
 Ma, X. et al., 2020, [MNRAS](#), 498, 2001
 Mandel, K. S., Narayan, G., Kirshner, R. P., 2011, [ApJ](#), 731, 120
 Martin, D. C. et al., 2005, [ApJ](#), 619, L1
 Martin, T. B., Prunet, S., Drissen, L., 2016, [MNRAS](#), 463, 4223
 Martin, T. B., Drissen, L., Melchior, A.-L., 2018, [MNRAS](#), 473, 4130
 Martins, F., Schaerer, D., Hillier, D. J., 2005, [A&A](#), 436, 1049
 Maschmann, D. et al., 2024, [ApJS](#), 273, 14
 Mathis, J. S., Ruml, W., Nordsieck, K. H., 1977, [ApJ](#), 217, 425
 Mayer, J. R., 1848, »Beiträge zur Dynamik des Himmels in populärer Darstellung«, [Joh. Alr. Landherr](#)
 McKee, C. F., Ostriker, J. P., 1977, [ApJ](#), 218, 148
 McKee, C. F., Ostriker, E. C., 2007, [ARA&A](#), 45, 565
 McLeod, A. F. et al., 2019, [MNRAS](#), 486, 5263
 McLeod, A. F. et al., 2020, [ApJ](#), 891, 25
 McLeod, A. F. et al., 2021, [MNRAS](#), 508, 5425
 McQuinn, K. B. W. et al., 2017, [AJ](#), 154, 51
 Méndez, R. H. et al., 2001, [ApJ](#), 563, 135
 Messier, C., 1781, Connaissance des Temps ou des Mouvements Célestes, 227
 Meurer, G. R. et al., 2009, [ApJ](#), 695, 765
 Mingozzi, M. et al., 2020, [A&A](#), 636, A42
 Mitra, S., Ferrara, A., Choudhury, T. R., 2013, [MNRAS](#), 428, L1
 Mohr, P. J., Taylor, B. N., 2000, [Rev. Mod. Phys.](#), 72, 351
 Morgan, W. W., Keenan, P. C., Kellman, E., 1943, »An atlas of stellar spectra, with an outline of spectral classification«,
 Müller, T., Prieto, J. L., Pejcha, O., Clocchiatti, A., 2017, [ApJ](#), 841, 127
 Nasonova, O. G., de Freitas Pacheco, J. A., Karachentsev, I. D., 2011, [A&A](#), 532, A104
 Newman, M. J. B. et al., 2024, [ApJ](#), 966, 175
 Niederhofer, F., Hilker, M., Bastian, N., Ercolano, B., 2016, [A&A](#), 592, A47
 Nugent, P. et al., 2006, [ApJ](#), 645, 841
 O'Donnell, J. E., 1994, [ApJ](#), 422, 158
 Okamoto, S. et al., 2024, [ApJ](#), 967, L24
 Olivares E., F. et al., 2010, [ApJ](#), 715, 833

- Osterbrock, D. E., Ferland, G. J., 2006, »*Astrophysics of gaseous nebulae and active galactic nuclei*«, [University Science Books](#)
- Paardekooper, J. P., Pelupessy, F. I., Altay, G., Kruip, C. J. H., 2011, [A&A](#), **530**, [A87](#)
- Pagul, A., Rivera, I., 2024, »*WFC3 Data Handbook v. 6*«, [Baltimore: STScI](#)
- Parker, Q. A. et al., 2003, in »*Planetary Nebulae: Their Evolution and Role in the Universe*«, Vol. 209, [p. 25](#)
- Paturel, G., Teerikorpi, P., 2006, [A&A](#), **452**, [423](#)
- Pedreros, M., Madore, B. F., 1981, [ApJS](#), **45**, [541](#)
- Pejcha, O., Prieto, J. L., 2015, [ApJ](#), **799**, [215](#)
- Pellegrini, E. W. et al., 2012, [ApJ](#), **755**, [40](#)
- Pérez-Montero, E., 2014, [MNRAS](#), **441**, [2663](#)
- Péroux, C., Howk, J. C., 2020, [ARA&A](#), **58**, [363](#)
- Perrin, J., 1919, [Ann. Phys.](#), **9**, [5](#)
- Pilyugin, L. S., Grebel, E. K., Kniazev, A. Y., 2014, [AJ](#), **147**, [131](#)
- Pilyugin, L. S., Grebel, E. K., 2016, [MNRAS](#), **457**, [3678](#)
- Planck, M., 1901, [AnP](#), **309**, [553](#)
- Planck Collaboration et al., 2020, [A&A](#), **641**, [A6](#)
- Polshaw, J. et al., 2015, [A&A](#), **580**, [L15](#)
- Povich, M. S., Whitney, B. A., 2010, [ApJ](#), **714**, [L285](#)
- Poznanski, D. et al., 2009, [ApJ](#), **694**, [1067](#)
- Prša, A. et al., 2016, [AJ](#), **152**, [41](#)
- Querejeta, M. et al., 2021, [A&A](#), **656**, [A133](#)
- Radburn-Smith, D. J. et al., 2011, [ApJS](#), **195**, [18](#)
- Raga, A. C., Cantó, J., Rodríguez, L. F., 2012, [MNRAS](#), **419**, [L39](#)
- Rahner, D., Pellegrini, E. W., Glover, S. C. O., Klessen, R. S., 2017, [MNRAS](#), **470**, [4453](#)
- Rahner, D., Pellegrini, E. W., Glover, S. C. O., Klessen, R. S., 2018, [MNRAS](#), **473**, [L11](#)
- Ramachandran, V. et al., 2018, [A&A](#), **615**, [A40](#)
- Ramambason, L. et al., 2020, [A&A](#), **644**, [A21](#)
- Ramambason, L. et al., 2022, [A&A](#), **667**, [A35](#)
- Rayleigh, L., 1879, [Phil. Mag. Ser. 5](#), **8**, [261](#)
- Razza, A. et al., in preparation, »*PHANGS-H α : Ground-based Narrow-band Imaging Survey of Nearby Star-Forming Galaxies*«
- Rebolo, R., Zapatero Osorio, M. R., Martín, E. L., 1995, [Nature](#), **377**, [129](#)
- Reindl, B., Tammann, G. A., Sandage, A., Saha, A., 2005, [ApJ](#), **624**, [532](#)
- Rekola, R. et al., 2005, [MNRAS](#), **361**, [330](#)
- Richard, J. et al., 2019, [arXiv](#), **1906.01657**
- Rickards Vaught, R. J. et al., 2024, [ApJ](#), **966**, [130](#)
- Riess, A. G., Press, W. H., Kirshner, R. P., 1996, [ApJ](#), **473**, [88](#)
- Riess, A. G. et al., 2016, [ApJ](#), **826**, [56](#)
- Riess, A. G. et al., 2022, [ApJ](#), **934**, [L7](#)
- Rizzi, L. et al., 2007, [ApJ](#), **661**, [815](#)
- Rodríguez, Ó., Clocchiatti, A., Hamuy, M., 2014, [AJ](#), **148**, [107](#)
- Rosolowsky, E. et al., 2021, [MNRAS](#), **502**, [1218](#)

Roth, M. M. et al., 2021, *ApJ*, **916**, 21
 Roth, M. M. et al., 2023, *arXiv*, **2311.14230**
 Rousseau-Nepton, L. et al., 2018, *MNRAS*, **477**, 4152
 Rousseau-Nepton, L. et al., 2019, *MNRAS*, **489**, 5530
 Roy, R. et al., 2011, *ApJ*, **736**, 76
 Ruiz-Lapuente, P., 1996, *ApJ*, **465**, L83
 Russell, D. G., 2002, *ApJ*, **565**, 681
 Rutherford, E., 1911, *Philos. Mag.*, **21**, 669
 Rydberg, J. R., 1890, *Philos. Mag.*, **29**, 331
 Ryon, J. E. et al., 2017, *ApJ*, **841**, 92
 Sabbi, E. et al., 2012, *ApJ*, **754**, L37
 Sabbi, E. et al., 2018, *ApJS*, **235**, 23
 Sagan, C., 1994, *»Pale Blue Dot: A Vision of the Human Future in Space«*, Random House
 Saha, A. et al., 2006, *ApJS*, **165**, 108
 Saintonge, A., 2025, *»Encyclopedia of Astrophysics, The Interstellar Medium«*, Elsevier
 Sakai, S., Ferrarese, L., Kennicutt Robert C., J., Saha, A., 2004, *ApJ*, **608**, 42
 Saldana-Lopez, A. et al., 2022, *A&A*, **663**, A59
 Salpeter, E. E., 1955, *ApJ*, **121**, 161
 Sana, H. et al., 2012, *Science*, **337**, 444
 Sánchez-Gil, M. C. et al., 2011, *MNRAS*, **415**, 753
 Sanders, R. L. et al., 2016, *ApJ*, **816**, 23
 Santoro, F. et al., 2022, *A&A*, **658**, A188
 Scheuermann, F. et al., 2022, *MNRAS*, **511**, 6087
 Scheuermann, F. et al., 2023, *MNRAS*, **522**, 2369
 Scheuermann, F. et al., submitted, *»Escape fraction of ionizing photons«*
 Schinnerer, E. et al., 2019, *The Messenger*, **177**, 36
 Schinnerer, E., Leroy, A. K., 2024, *ARA&A*, **62**, 369
 Schlafly, E. F., Finkbeiner, D. P., 2011, *ApJ*, **737**, 103
 Schmidt, B. P., Kirshner, R. P., Eastman, R. G., 1992, *ApJ*, **395**, 366
 Schmidt, B. P. et al., 1994, *ApJ*, **432**, 42
 Schmidt, M., 1959, *ApJ*, **129**, 243
 Schoeniger, F., Sofue, Y., 1997, *A&A*, **323**, 14
 Schönberner, D. et al., 2005, *A&A*, **431**, 963
 Schönberner, D., Jacob, R., Sandin, C., Steffen, M., 2010, *A&A*, **523**, A86
 Schramm, D. N., Turner, M. S., 1998, *Rev. Mod. Phys.*, **70**, 303
 Shapley, H., Curtis, H. D., 1921, Bulletin of the National Research Council, **2**, 171
 Sharina, M. E., Karachentsev, I. D., Tikhonov, N. A., 1996, *A&AS*, **119**, 499
 Shaver, P. A. et al., 1983, *MNRAS*, **204**, 53
 Shaya, E. J., Tully, R. B., Hoffman, Y., Pomarède, D., 2017, *ApJ*, **850**, 207
 Simón-Díaz, S., Stasińska, G., 2008, *MNRAS*, **389**, 1009
 Singh, K. P. et al., 2014, in *»Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray«*, Vol. **9144**, p. 91441S
 Smith, L. J., Norris, R. P. F., Crowther, P. A., 2002, *MNRAS*, **337**, 1309

Smith, B. A., 1976, in »Charge-Coupled Device Technology and Applications«, p. 135

Sohn, Y.-J., Davidge, T. J., 1996, *AJ*, 111, 2280

Sorce, J. G., Tully, R. B., Courtois, H. M., 2012, *ApJ*, 758, L12

Sorce, J. G. et al., 2014, *MNRAS*, 444, 527

Sparks, W. B., 1994, *ApJ*, 433, 19

Spitzer, L., 1978, »Physical processes in the interstellar medium«, Wiley

Spitzer Jr., L., 1956, *ApJ*, 124, 20

Spriggs, T. W. et al., 2020, *A&A*, 637, A62

Spriggs, T. W. et al., 2021, *A&A*, 653, A167

Springob, C. M., Haynes, M. P., Giovanelli, R., Kent, B. R., 2005, *ApJS*, 160, 149

Springob, C. M. et al., 2009, *ApJS*, 182, 474

Starkenburger, E. et al., 2017, *MNRAS*, 471, 2587

Starr, F., 1897, in »Appletons' Popular Science Monthly«, Vol. L1, p. 784

Stetson, P. B., 1987, *PASP*, 99, 191

Stevance, H. F. et al., 2020, *MNRAS*, 498, 1347

Stone, E. C. et al., 2013, *Science*, 341, 150

Storey, P. J., Hummer, D. G., 1995, *MNRAS*, 272, 41

Strömgren, B., 1939, *ApJ*, 89, 526

Sutherland, R., Dopita, M., Binette, L., Groves, B., 2018, »MAPPINGS V: Astrophysical plasma modeling code«, *Astrophysics Source Code Library*

Tacchella, S. et al., 2022, *MNRAS*, 513, 2904

Takanashi, N., Doi, M., Yasuda, N., 2008, *MNRAS*, 389, 1577

Tammann, G. A., Sandage, A., Reindl, B., 2008, *A&A Rev.*, 15, 289

Teh, J. W. et al., 2023, *MNRAS*, 524, 1191

Teodorescu, A. M. et al., 2011, *ApJ*, 736, 65

Terry, J. N., Paturel, G., Ekholm, T., 2002, *A&A*, 393, 57

Theureau, G. et al., 2007, *A&A*, 465, 71

Thilker, D. A., Braun, R., Walterbos, R. A. M., 2000, *AJ*, 120, 3070

Thilker, D. A. et al., 2022, *MNRAS*, 509, 4094

Thilker, D. A. et al., 2025, *ApJS*, 280, 1

Tielens, A. G. G. M., 2005, »The Physics and Chemistry of the Interstellar Medium«, Cambridge Univ. Press

Tominaga, N., Umeda, H., Nomoto, K., 2007, *ApJ*, 660, 516

Tonry, J. L. et al., 2001, *ApJ*, 546, 681

Toribio San Cipriano, L. et al., 2017, *MNRAS*, 467, 3759

Trebitsch, M. et al., 2017, *MNRAS*, 470, 224

Tully, R. B., Fisher, J. R., 1988, »Catalog of Nearby Galaxies«, Cambridge Univ. Press

Tully, R. B., Shaya, E. J., Pierce, M. J., 1992, *ApJS*, 80, 479

Tully, R. B. et al., 2009, *AJ*, 138, 323

Tully, R. B. et al., 2013, *AJ*, 146, 86

Tully, R. B., Courtois, H. M., Sorce, J. G., 2016, *AJ*, 152, 50

Turner, J. A. et al., 2021, *MNRAS*, 502, 1366

Turiziani, S., 2019, »50 Years of CCDs«, AAS Historical Astronomy Division

- Unsöld, A., Baschek, B., 2002, »*Der neue Kosmos. Einführung in die Astronomie und Astrophysik*«, Springer
- van de Sande, J. et al., 2024, in »*Early Disk-Galaxy Formation from JWST to the Milky Way*«, Vol. 377, p. 27
- van der Walt, S. et al., 2014, PeerJ, 2, e453
- Van Dyk, S. D., Li, W., Filippenko, A. V., 2006, PASP, 118, 351
- Van Dyk, S. D. et al., 2015, ApJ, 806, 195
- Vázquez, G. A., Leitherer, C., 2005, ApJ, 621, 695
- Veilleux, S., Osterbrock, D. E., 1987, ApJS, 63, 295
- Ventimiglia, G., Arnaboldi, M., Gerhard, O., 2011, A&A, 528, A24
- Wagoner, R. V., Fowler, W. A., Hoyle, F., 1967, ApJ, 148, 3
- Wagoner, R. V., 1973, ApJ, 179, 343
- Wallerstein, G., 1962, ApJS, 6, 407
- Walter, F. et al., 2008, AJ, 136, 2563
- Wang, B. et al., 2021, ApJ, 916, 3
- Webb, S., 1999, »*Measuring the Universe*«, Springer
- Wehrli, C., 1985, World Radiation Center, 615, 10
- Wei, W. et al., 2020, MNRAS, 493, 3178
- Weilbacher, P. M. et al., 2020, A&A, 641, A28
- Weizsäcker, C. F. V., 1937, Phys. Z., 38, 176
- Weizsäcker, C. F. V., 1938, Phys. Z., 39, 633
- Weßfall, K. B. et al., 2019, AJ, 158, 231
- Whitmore, B. C. et al., 2011, ApJ, 729, 78
- Whitmore, B. C. et al., 2021, MNRAS, 506, 5294
- Whitmore, B. C. et al., 2023, MNRAS, 520, 63
- Whitmore, B. C. et al., 2025, ApJ, 982, 50
- Wien, W., 1893, AnP, 285, 633
- Willick, J. A. et al., 1997, ApJS, 109, 333
- Wilson, R. W., Jefferts, K. B., Penzias, A. A., 1970, ApJ, 161, L43
- Wollaſton, W. H., 1802, Philos. Trans. R. Soc. Lond. Ser. I, 92, 365
- Yasuda, N., Fukugita, M., Okamura, S., 1997, ApJS, 108, 417
- Zasov, A. V., Bizyaev, D. V., 1996, Astron. Lett., 22, 71
- Zaſtrow, J., Oey, M. S., Pellegrini, E. W., 2013, ApJ, 769, 94
- Zhang, J.-J. et al., 2014, AJ, 148, 1
- Zurita, A., Rozas, M., Beckman, J. E., 2000, A&A, 363, 9

ACKNOWLEDGEMENTS

Congratulations, you made it through the 166 pages of the academic part, so hopefully you will also endure the last one. This thesis marks the end of my ~~decade~~ dozen-year long journey at *Heidelberg University* – and what a joyful ride it has been. This I owe to the people who accompanied me on this voyage, and now all that remains for me is to pay tribute to everyone who made it worthwhile.

First and foremost, I would like to express my deepest gratitude to Kathryn. Thank you for giving me the opportunity to pursue this project, for always being willing to listen and taking the time to answer my questions. Your positive attitude encouraged me to keep going, and your lively manner breathed fresh life into the institute. It truly has been a wonderful adventure to explore the vastness of space under your guidance.

And this whole undertaking was only possible on board PHANGS, an amazing group of people who join forces to broaden our understanding by collecting and analysing a treasure trove of data. I am much obliged for the many instructive discussions with experts from all around the world. Participating in such a fantastic collaboration was a fun and rewarding experience. A special thanks go to my co-authors, and even more so to my paper teams – with a shout-out to Ash, Brent, Enrico, Francesco B., Francesco S., Janice, Jia Wei, and Stephen. It was a pleasure working with Y'all.

This also applies to all the nice colleagues at ARI. As usual in this field, there is a constant coming and going, so I have made many acquaintances over the years. To the current and former members of the *Baryonic Life Cycle Group*, Liz, Oleg & Zhenya, Jing, Ema, Eduardo, Shuverthi, Mar, Natascha, Silvia, Hannah, Michael, and Eric. To my office mates, Zdeněk, Bahar, Guštavo, Payel, Kolja, Jelena, Bruno, and Pranav. To Eva and Simon for being part of my thesis committee, and to Anna, Ralf, and Mathias for being a part of my examination committee. To the administrative staff who work in the background to ensure that everything runs smoothly, to Saskia and her colleagues.

There was an unexpected interruption and admittedly, this thesis finished under rather unusual circumstances. The fact that I am standing here today is owed to many skilled and caring people. I am deeply grateful to everyone who helped me through the last two years, but most of all to my family. Thank you from the bottom of my heart for accompanying me through this difficult time and always looking after me.

That brings us to the end. Now it is my turn to say goodbye, and I do so with a smile in one eye and a tear in the other. I will always fondly remember this time. Fare thee well!



STATEMENT OF AUTHORSHIP

I herewith declare that this thesis was solely composed by myself, unless otherwise acknowledged in the text.

Heidelberg, 8 September 2025

Fabian SCHEUERMANN

*»The time is gone, the song is over,
thought I'd something more to say.«*
—Time, Pink Floyd

COLOPHON

This thesis has been typeset from a \LaTeX file prepared by the author. It is based on the `scrbook` documentclass from `KOMA-script`. `Kp-Fonts` serves as the typeface, with ligatures and `oldstyle` figures lending it a classic and elegant appearance. Bibliographic references are managed with `BibLaTeX` and processed using `Biber`, while abbreviations are organized with the `glossaries-extra` package, and cross-references are styled with `cleveref`. Units are consistently presented with `siunitx` and certain diagrams and graphic elements were created with the assistance of `PGF/TikZ`. A template for this thesis is available here:

<https://github.com/fschmnn/thesis-template>