

Dissertation

**Visualising the Synthetic
Universe**

Daniel Thomas Haydon

Universität Heidelberg

2020

DISSERTATION

submitted to the
COMBINED FACULTIES OF THE NATURAL SCIENCES AND
MATHEMATICS
of the
RUPERTO-CAROLA-UNIVERSITY OF HEIDELBERG, GERMANY
for the degree of
DOCTOR OF NATURAL SCIENCES

Put forward by

Daniel Thomas Haydon

born in: Blackburn, United Kingdom

Oral examination: Monday 3rd February, 2020

Visualising the Synthetic Universe

Referees: Dr. J. M. Diederik Kruijssen
Prof. Dr. Simon C. O. Glover

Abstract



STAR formation cannot truly be understood from observational data alone; only with simulations is it possible to assemble the complete picture. Observations guide the physics we build into our simulations, yet the impact of different star formation and feedback models can only be investigated with simulations. Synthetic observations allow us to make a realistic comparison to true observations as well as teach us about the emission tracers we depend upon. Through coupling the stellar population synthesis code `SLUG2` to galaxy simulations, we can generate synthetic star formation rate tracer maps. These maps assume different stellar metallicities, star formation rate surface densities, and suffer from varied amounts of extinction. This allows us to explore and constrain the environmental effects on the characteristic emission lifetimes — the duration for which a tracer is visible. With these emission lifetimes and in conjunction with a new statistical method, the ‘uncertainty principle for star formation’, constraints can be placed upon the durations of different evolutionary phases of the star formation process, allowing us to better understand the physics of star formation and feedback on sub-galactic scales. Studying the interstellar medium can also reveal information about stellar feedback: the gas density structure is altered as a result of the injected energy, momentum, and matter. Surveys of the CO emission in galaxies can tell us how the properties of this medium have evolved over cosmic time. Using `DESPOTIC` to model CO line emission of gas found within the IllustrisTNG50 cosmological simulation, we produce an equivalent synthetic survey. This synthetic survey can be used as a basis for comparison and predictor of observational trends.

Zusammenfassung



STERNENTSTEHUNG kann anhand von Beobachtungen nicht komplett verstanden werden. Deshalb ist man auf Simulationen angewiesen, um dieses Bild zu vervollständigen. Beobachtungsdaten diktieren die Physik, die in Simulationen eingebaut wird, aber die Auswirkungen von verschiedener Sternentstehungs- und Rückkopplungsmodelle können ausschließlich durch Simulationen erkundet werden. Synthetische Beobachtungen erlauben einen realistischen Vergleich von Simulationen zu echten Beobachtungen und geben Auskunft über die Emission der Indikatoren von verschiedenen Evolutionsphasen des Sternentstehungsprozesses. Durch die Kopplung des Sternpopulationssynthese-Codes SLUG2 und Simulationen von Galaxien können wir synthetische Karten bilden, welche die Emission der Indikatoren der Sternentstehungsrate anzeigen. Diese Karten nehmen verschiedene Sternmetallizitäten und Oberflächendichten der Sternentstehungsrate an. Außerdem erleiden sie unterschiedlich viel Extinktion. Mit diesen Karten erforschen und beschränken wir die Umweltauswirkungen auf die typische Lebensdauer der Emission — die Zeitspanne, während der Emission messbar ist. Die Lebensdauern der Emission zusammen mit der neuen statistischen Methode „uncertainty principle for star formation“¹ können die Dauer verschiedener Evolutionsphasen des Sternentstehungsprozess beschränken. Dadurch erreichen wir ein besseres Verständnis der Physik der Sternentstehung und der Rückkopplung in den subgalaktischen Maßstäben. Durch das Studium des interstellaren Mediums gewinnt man auch Informationen über Rückkopplung: Die Dichtestruktur des Gases wird durch den Zustrom von Energie, Impuls und Materie verändert. Eine Durchmusterung nach CO-Emissionslinien in Galaxien (d.h. eine CO-Karte von einem Teil des Himmels) kann zeigen, wie die Eigenschaften des Mediums über die kosmische Zeit entwickelt haben. Mit DESPOTIC modellieren wir CO-Emissionslinien von Gas, das in der kosmologischen Simulation IllustrisTNG50 gefunden wird, und herstellen entsprechende synthetische CO-Karten des Simulationsvolumens. Diese synthetischen Karten können sowohl Vergleichsgrundlage als auch Prädiktor für empirischen Tendenzen verwendet werden.

¹ „Unbestimmtheitsrelation für Sternentstehung“

Contents

Abstract	v
Zusammenfassung	vii
Contents	ix
List of Figures	xi
List of Tables	xv
List of Terms and Abbreviations	xvii
1 Introduction	1
1.1 Cosmology	1
1.2 Galaxies	11
1.3 Stars	19
1.4 This work	31
2 Uncertainty principle for star formation	39
3 Emission lifetimes of SFR tracers	47
3.1 Measuring the emission lifetime of SFR tracers	49
3.1.1 Galaxy simulation	51
3.1.2 Generation of reference maps	53
3.1.3 Generation of synthetic emission maps	56
3.2 A well sampled IMF at solar metallicity	59
3.3 The effects of metallicity	69
3.4 The effects of IMF sampling	79
3.4.1 Theoretical expectation	79
3.4.2 SFR tracer lifetimes for a stochastically sampled IMF	84
3.5 Comparison to observations	89
3.6 Conclusions	92
3.A $H\alpha$ Absorption and Emission Features	94
3.B Error Propagation	97
Acknowledgements	98

4	The influence of dust extinction on SFR tracer lifetimes	99
4.1	Measuring the effects of extinction	100
4.1.1	Analysis framework	100
4.1.2	Galaxy simulation	101
4.1.3	Generation of synthetic emission maps	104
4.2	Extinction	110
4.3	Sensitivity limit	117
4.4	Conclusions	121
	Acknowledgements	123
5	The synthetic ALMA UDF	125
5.1	Cosmological simulation	127
5.2	Building a synthetic light cone	128
5.3	Synthetic CO emission	138
5.3.1	Model cloud	139
5.3.2	Simulation input	142
5.3.3	Predicting CO emission with <code>DESPOTIC</code>	143
5.3.4	Lookup table	146
5.4	Synthetic CO observations	156
5.5	Conclusions	165
	Acknowledgements	167
6	Conclusion	169
	Acknowledgements	175
	List of publications	177
	Author	177
	Co-author	177
	Bibliography	179

List of Figures

1.1	BBN reaction network	4
1.2	Primordial light element abundances	5
1.3	The CMB as seen by the Planck space telescope	7
1.4	The CMB angular power spectrum	7
1.5	Evidence for the Baryon Acoustic Peak	8
1.6	Large scale structure within the Universe	10
1.7	Schematic diagram of galaxy structure	12
1.8	The Hubble sequence galaxy classification scheme	12
1.9	Example spiral galaxies: NGC13000 and NGC4414	15
1.10	Bimodal galaxy distribution in the colour-mass diagram	17
1.11	The galactic gas cycle	18
1.12	Schematic view of star formation	20
1.13	Observational tracers of molecular cloud structure	23
1.14	Internal stellar structure	25
1.15	Stellar evolutionary stages	26
1.16	Schematic Hertzsprung-Russell Diagram	29
1.17	Evolutionary tracks of solar metallicity stars	30
1.18	The Kennicutt-Schmidt relation based on measurements from multiple galaxies	32
1.19	The Kennicutt-Schmidt relation based on measurements of NGC5194 at different spatial resolutions	33
1.20	The effects of stellar feedback on galaxy morphology	36
2.1	Steps of the HEISENBERG code	41
2.2	Example tuning fork diagram produced by the HEISENBERG code	42
3.1	Properties of the simulated galaxy in the context of the nearby galaxy population	52
3.2	Example reference and synthetic $H\alpha$ emission maps used as input for the HEISENBERG code	54
3.3	Schematic diagram showing the relation between the different timescales we define in this chapter	55
3.4	Example emission lifetime measurements for the range of different reference map parameters	62
3.5	Example PDF for defined emission lifetimes	63

3.6	The relation between UV emission lifetimes and response-weighted mean wavelengths of the UV filters	66
3.7	The relation between H α (with continuum) emission lifetimes and filter width	67
3.8	The change in H α and FUV emission with time	68
3.9	The relation between the H α (without continuum) emission lifetime and metallicity	71
3.10	The relations between H α (with continuum) emission lifetimes and metallicity	72
3.11	The relations between UV emission lifetimes, metallicity, and response-weighted mean wavelength	73
3.12	The surface relating H α (with continuum) emission lifetimes, metallicity, and filter width	76
3.13	The surface relating UV emission lifetimes, metallicity, and response-weighted mean wavelength	78
3.14	Sample probability curves as a function of the star-forming region mass	81
3.15	The change in emission lifetime as a function of IMF sampling	86
3.16	Tuning fork diagram for NGC300	91
3.17	Comparison of the equivalent width of H α absorption and emission	96
4.1	Co-spatial H α and CO emission	103
4.2	Distribution of column densities for a face-on, simulated galaxy	106
4.3	Synthetic H α (without continuum) emission maps with and without extinction	107
4.4	The relative duration of synthetic emission maps compared to spatially transformed versions of themselves	109
4.5	Change in H α (without continuum) emission lifetime as a function of column density	112
4.6	Change in H α (with continuum) emission lifetimes as a function of column density	113
4.7	Change in UV emission lifetimes as a function of column density	114
4.8	The PDF of H α (without continuum) emission and the probability of emission falling within the peak-identifying range .	115
4.9	Change in H α (with continuum) emission timescales as a function of sensitivity limit	118
4.10	Change in H α (without continuum) emission timescale as a function of sensitivity limit	119
4.11	Change in UV emission timescales as a function of sensitivity limit	120

5.1	Schematic diagram illustrating the quantities used to calculate the distance between two points	132
5.2	The structure of a synthetic light cone	134
5.3	Consequences of poor light cone construction	136
5.4	Schematic diagram illustrating the two-stage selection process of simulation objects	137
5.5	Illustration of how kernel density estimation works to construct a distribution	148
5.6	An example of a poor selection in lookup table coordinate using inverse transformation sampling	150
5.7	Lookup table coordinate selection	151
5.8	A schematic showing the construction of a selection table . .	153
5.9	Projections of the 8D selection table	155
5.10	A synthetic CO emission map, at 0.01'' resolution, based on a light cone constructed from the IllustrisTNG50 simulation .	158
5.11	Maps showing the spatial distribution of input quantities used for interpolation	159
5.12	A synthetic CO emission map, at 1'' resolution, based on a light cone constructed from the IllustrisTNG50 simulation .	160
5.13	The redshift dependence of CO emission	163
5.14	The CO luminosity function at different redshifts	164

List of Tables

3.1	Description and properties of the SFR tracers considered . .	50
3.2	Parameter values used in creating reference maps	57
3.3	Measured emission lifetimes and corresponding age ranges .	65
3.4	Percentages of the emission intensity corresponding to measured emission lifetimes	65
3.5	Measured emission lifetimes and corresponding age ranges .	74
3.6	Sample parameter values for the approximate curves to the full analytical forms	83
3.7	Parameter values for the analytical function describing the change in $H\alpha$ emission lifetime with IMF sampling	87
5.1	Input parameters for DESPOTIC	140

List of Terms and Abbreviations

Acronyms

AGN	active galactic nucleus
BBN	big bang nucleosynthesis
CDF	cumulative distribution function
CDM	cold dark matter
CMB	cosmic microwave background
FoF	friends-of-friends
FUV	far-ultraviolet
FWHM	full width half maximum
GMC	giant molecular cloud
IGM	intergalactic medium
IMF	initial mass function
IR	infrared
ISM	interstellar medium
ISRF	interstellar radiation field
NUV	near-ultraviolet
PDF	probability density function
SED	spectral energy distribution
SFR	star formation rate
SPH	smoothed-particle hydrodynamics
SPS	stellar population synthesis
UDF	ultra-deep field
UV	ultraviolet
XDF	extreme deep field
YSO	young stellar object

Codes

- DESPOTIC** ‘Derive the Energetics and SPectra of Optically Thick Interstellar Clouds’. A chemical, thermal, and statistical equilibrium code to calculate the state of interstellar clouds and their observable emission (see Krumholz 2014)
- HEISENBERG** Code based on the KL14 principle method (see Kruijssen et al. 2018). Described in Chapter 2^[p. 39]
- SLUG2** ‘Stochastically Lighting Up Galaxies’. A stochastic SPS code which predicts the photometric properties of simple stellar population (see Silva et al. 2012, 2014; Krumholz et al. 2015)
- STARBURST99** Code modelling photometric properties of star-forming galaxies (see Leitherer et al. 1999; Vázquez et al. 2005; Leitherer et al. 2010, 2014)

Instruments

- UVOT** The Swift Ultra-Violet/Optical Telescope
- WFC3** The HST Wide Field Camera 3
- WFPC2** The HST Wide Field and Planetary Camera 2

Methods

- KL14 principle** A new statistical method, titled ‘an uncertainty principle for star formation’ (see Kruijssen et al. 2014). Described in Chapter 2^[p. 39]

Notation

- $\tilde{\square}$ Comoving value
- $\overline{\square}$ Mean value

Observatories

ALMA	Atacama Large Millimeter/submillimeter Array
GALEX	GALaxy Evolution eXplorer
HST	Hubble Space Telescope
Swift	Neil Gehrels Swift Observatory
WISE	Wide-field Infrared Survey Explorer

Simulations

FIRE	Feedback In Realistic Environments (see Hopkins et al. 2014)
IllustrisTNG	The Next Generation Illustris. Magnetohydrodynamical cosmological simulations of galaxy formation (see Weinberger et al. 2017 ; Pillepich, Springel et al. 2018)
OWLS	OverWhelmingly Large Simulations (see Schaye et al. 2010)
SLICC	Simulating the LifeCycle of molecular Clouds (see Walch et al. 2015)

Surveys

ASPECS	ALMA SPECTroscopic Survey in the Hubble UDF (see Walter et al. 2016)
SDSS	Sloan Digital Sky Survey (see York et al. 2000)
xCOLDGASS	eXtended CO Legacy Data base for the GALEX Arecibo SDSS Survey (see Saintonge et al. 2017)
xGASS	eXtended GALEX Arecibo SDSS Survey (see Catinella et al. 2018)

Introduction

The aim of this thesis is to improve our understanding of star formation; while we cannot claim to answer all the questions, we at least provide some bricks to pave the path to a more complete understanding.

Star formation is a product of the environment and the evolutionary steps which came before. Therefore, in the following three sections we describe the background to our work starting from cosmological scales in [Section 1.1](#), down to galactic scales in [Section 1.2](#)^[p. 11], and eventually the stars and star forming regions themselves in [Section 1.3](#)^[p. 19]. We finally describe the context for our work in [Section 1.4](#)^[p. 31].

1.1 Cosmology

Under the model of a homogeneous and isotropic Universe, in which the theory of general relativity applies, the cosmological scale factor a (parametrising the relative expansion of the Universe) is related to the energy content of the Universe through the Friedmann equation

$$\left(\frac{1}{a} \frac{da}{dt}\right)^2 = \frac{8\pi G}{3} \rho - \frac{Kc^2}{a^2}, \quad \text{Equation 1.1}$$

where G is the gravitational constant; c the speed of light in a vacuum; $K = \{+1, 0, -1\}$ specifies the geometry of the Universe as having positive, no, or negative curvature; and ρ is the total energy density in matter, radiation, and dark energy. This equation can be simplified to

$$1 = \Omega + \Omega_K, \quad \text{Equation 1.2}$$

through the use of the following definitions

$$H \equiv \frac{1}{a} \frac{da}{dt}; \quad \Omega \equiv \frac{8\pi G}{3H^2} \rho; \quad \Omega_K \equiv -\frac{Kc^2}{a^2 H^2}, \quad \text{Equation 1.3}$$

where H is the (time-dependent) Hubble parameter.

In the Λ CDM model, the Standard Model of Cosmology, the Universe is taken to have a flat geometry ($\Omega_K = 0$) and contributions to the energy density in the form of

1 Introduction

matter, Ω_m baryonic matter, cold dark matter (CDM)

radiation, Ω_r photons, three active neutrino species¹

dark energy, Ω_Λ as vacuum energy (often expressed through the cosmological constant, Λ).

For use in [Chapter 5](#)^[p. 125], we express [Equation 1.1](#)^[p. 1] in terms of the present-day (denoted by the subscript '0') energy density components and Hubble parameter

Equation 1.4
$$H = H_0 E(a) ,$$

where we define the expansion factor, $E(a)$, as

Equation 1.5
$$E(a) = \sqrt{\Omega_{r,0} a^{-4} + \Omega_{m,0} a^{-3} + \Omega_{\Lambda,0} a^{-2} + \Omega_{K,0}} ,$$

which can also be expressed in terms of redshift, z , using

Equation 1.6
$$a = \frac{1}{1+z} .$$

This model works under the Big Bang paradigm, where the Universe is thought to have begun in an infinitely hot and infinitely dense state before undergoing expansion and cooling. There are six free parameters in this model: $\Omega_b h^2$, $\Omega_c h^2$, H_0 , τ , n_s , and A_s , which we briefly describe here.

$\Omega_b h^2$, $\Omega_c h^2$ The baryonic density, Ω_b , and CDM density, Ω_c , scaled by the dimensionless Hubble parameter

Equation 1.7
$$h \equiv \frac{H}{100 \text{ km s}^{-1} \text{ Mpc}^{-1}} ,$$

such that $\Omega_i h^2 \propto \rho_i$.

H_0 The present day Hubble parameter, H_0 , characterises the expansion rate and age of the Universe.

τ The optical depth to reionisation, τ , is important for defining the early Universe ionisation history and given as ([Griffiths et al. 1999](#))

Equation 1.8
$$\tau = \int_0^{z_{\text{reion}}} \sigma_T n_e(z) \frac{c dz}{(1+z) H(z)} ,$$

where σ_T the Thompson scattering cross-section, n_e the electron number volume density, and z_{reion} the redshift of reionisation.

¹Despite being massive particles, neutrinos behaved like radiation in the hot, early Universe.

n_s, A_s The primordial scalar density perturbations, \mathcal{P} , are described by the scalar spectral index, n_s , and the scalar power spectrum amplitude, A_s , (Planck Collaboration 2014):

$$\mathcal{P}(k) \propto A_s k^{n_s-1},$$

Equation 1.9

where k is the wave number of the perturbation. These perturbations are thought to be the seeds of the large scale structure we see in the Universe. Sometimes this is expressed through the mass dispersion within a sphere of radius $8 h^{-1}$ Mpc, known as σ_8 .

The Universe, over the past ~ 14 Gyr, has evolved through several key stages in order to reach its current state; however, only some of these stages have left measurable signatures in the Universe. Anything that occurred within a Planck time $\sim 10^{-44}$ s of the beginning of the Universe (Planck Epoch), cannot be described by modern Physics and is speculated to require a quantum description of gravity. It is theorised that between $\sim 10^{-33}$ and $\sim 10^{-32}$ s the Universe underwent a phase of exponential expansion and rapid cooling (Inflationary Epoch). By $\sim 10^{-6}$ s the four fundamental forces — gravity, strong nuclear force, weak nuclear force, and electromagnetic force — have separated and are now distinct. Within ~ 1 s, the Universe has cooled sufficiently that quarks remain bound as hadrons — the preferential creation of matter over antimatter is still not understood (the matter-antimatter asymmetry problem).

We now reach the stage of big bang nucleosynthesis (BBN) lasting from around ~ 1 to 10^3 s. This is the epoch in which light element nuclei (beyond Hydrogen) are synthesised, defining the chemical composition of the material from which the first stars are created. The ability to form Deuterium (${}^2\text{H}$, also D) is the critical step in forming heavier nuclei; this is made evident by the chemical reaction network shown in Figure 1.1^[p.4]. Deuterium production is also important in stopping the gradual decay of neutrons into protons (half life 10.61 ± 0.16 minutes, Christensen et al. 1972) but can only occur once the Universe has cooled sufficiently that the photodissociation rate of Deuterium is low. BBN produces Deuterium, Helium-3, Helium-4, and Lithium-7 in sufficient quantities as to be detectable. In Figure 1.2^[p.5], we demonstrate how abundance measurements and independent measurements of $\Omega_b h^2$ can place constraints on our model of the Universe as well as test for consistency. The expansion of the Universe eventually puts a halt to BBN, as the density becomes so low that nuclei seldom interact.

From the end of BBN to the time of recombination (~ 380 kyr), photons and baryons are strongly coupled. This coupling is mediated by electrons through Thompson and Coulomb scattering (Hu et al. 2002). Denser regions within the plasma collapse under gravity, causing a rise in temperature. This in turn leads

1 Introduction

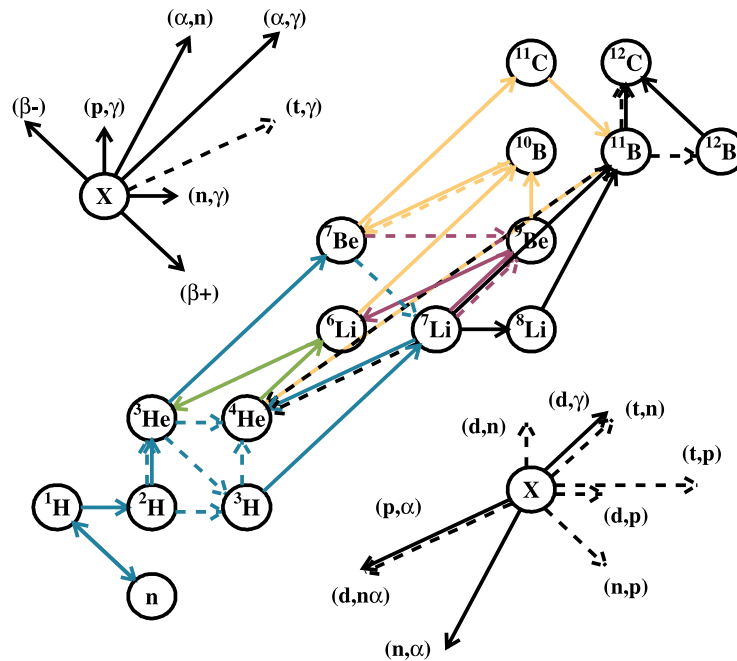
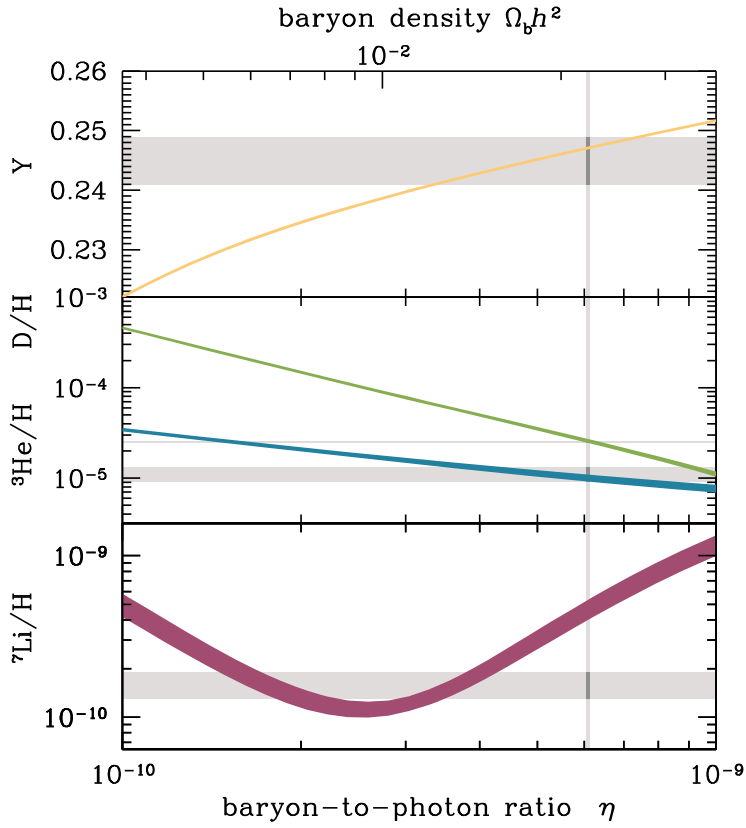


Figure 1.1

A reduced BBN chemical reaction network showing the key reactions for ^2H , $^3,^4\text{He}$ and ^7Li (in blue), ^6Li (in green), ^9Be (in red), $^{10,11}\text{B}$ (in yellow), and CNO production (in black). The two \textcircled{X} , define the additional (reactants, products) of the reaction indicated by the arrow style and direction. It is clear that the possibility of forming elements beyond Hydrogen (^1H). depends on the crucial step of synthesising Deuterium (^2H).

Image Credit: Coc et al. (2012, Figure2), reproduced with permission © AAS



The predicted change in primordial abundances relative to Hydrogen, of top: Helium-4 (by mass), middle: Deuterium and Helium-3 (by number), bottom: Lithium-7 (by number), with the number ratio of baryons to photons, η . This ratio can also be used to calculate $\Omega_b h^2 = 10^{10} \eta / 273.9$ (Steigman 2006). The coloured bands indicate the 1σ uncertainties in the models. We have added the horizontal grey bands to indicate observational constraints: $Y = 0.2449 \pm 0.0040$ (Aver et al. 2015), measurements made from observations of Hydrogen and Helium emission lines in low-metallicity H II regions; $D/H = (2.53 \pm 0.04) \times 10^{-5}$ (Cooke et al. 2014), measurements made from absorption line spectra of high redshift quasars (this is a lower limit as stars readily destroy Deuterium); ${}^3\text{He}/H = (1.1 \pm 0.2) \times 10^{-5}$ (Bania et al. 2002), measurements of H II regions in the Milky Way; ${}^7\text{Li}/H = (1.6 \pm 0.3) \times 10^{-10}$ (Cyburt et al. 2016), measurements made from the atmospheres of metal-poor stars. The vertical grey band shows the (Planck Collaboration 2016) value $\Omega_b h^2 = 0.02230 \pm 0.00014$. The discrepancy between Lithium-7 models and observations is a known and unsolved problem (Cyburt et al. 2008; Fields 2011).

Figure 1.2

Image Credit: Cyburt et al. (2016, Figure1), adapted with permission © American Physical Society

1 Introduction

to an increase in the radiation pressure which eventually causes the plasma to expand again. The result of this is a series of oscillatory modes in the density of the plasma. Eventually, as a consequence of the expanding and cooling Universe, electrons can no longer remain free and bind with atomic nuclei for the first time to form neutral atoms (Epoch of Recombination). Photons, now unable to interact with the bound electrons, are free to propagate throughout the Universe — what we now observe as the cosmic microwave background (CMB).

The CMB has a black body spectrum with the characteristic temperature of $T_0 = 2.718 \pm 0.021$ K (Planck Collaboration 2016) but at the time of recombination (redshift $z_{\text{re}} \approx 10^3$, Planck Collaboration 2016) the temperature would be a factor of $1 + z_{\text{re}}$ higher. In Figure 1.3^[p. 7], we show the temperature fluctuations of the CMB as seen by the Planck space telescope. These temperature fluctuations reflect the underlying density fluctuations at the time of recombination. Lower temperatures correspond to higher densities as a result of the (non-integrated) Sachs–Wolfe effect (White et al. 1997), where photons leaving denser regions at the time of recombination suffer a greater gravitational redshift compared to those of less dense regions.

Figure 1.3^[p. 7] can be decomposed into spherical harmonics, Y_ℓ^m , allowing us to describe the temperature fluctuations, ΔT , statistically through the angular power spectrum, C_ℓ , (Liddle 2003):

Equation 1.10

$$C_\ell \equiv \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2,$$

where the coefficients, $a_{\ell m}$, are define through

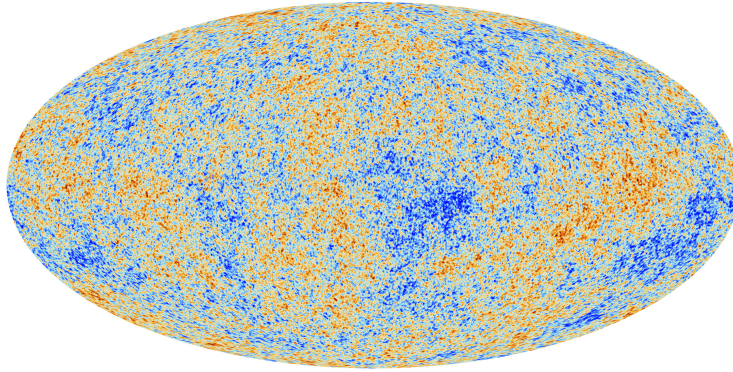
Equation 1.11

$$\frac{\Delta T}{T_0} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_\ell^m.$$

where the multipole number, ℓ , is approximately related to the angular size of the perturbation, ψ : $\ell \approx 180^\circ/\psi$. We present the power spectrum in Figure 1.4^[p. 7].

The shape of the power spectrum is sensitive to many of the Λ CDM model parameters, placing very tight observational constraints. Some peaks within the spectrum have special significance, for example, the location of the first peak at $\ell \approx 200$ indicates a Universe with a flat geometry ($\Omega_{\text{K}} = 0.000 \pm 0.005$ Planck Collaboration 2016). These peaks are the result of the oscillating density perturbations within the baryon-photon plasm, which we described above, and form (approximately) a harmonic series. When recombination

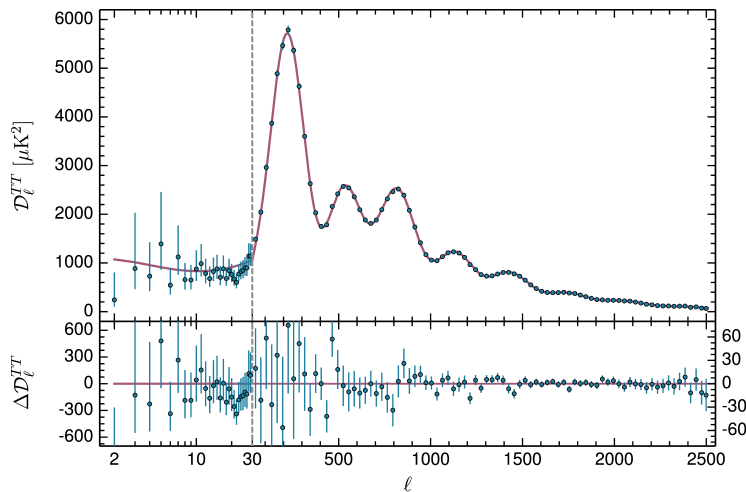
1.1 Cosmology



The temperature fluctuations in the CMB as measured by the Planck space telescope. These fluctuations are of the order $\pm 10^{-5}$ K and are a result of the underlying density fluctuations at the time of recombination.

Figure 1.3

Image Credit: ESA / Planck Collaboration (2013, Figure1)



The top panel shows the scaled angular power spectrum, $\mathcal{D}_\ell^{\text{TT}}$, of the CMB, shown in Figure 1.3. The scaled angular power spectrum is related to the power spectrum C_ℓ defined in Equation 1.10^[p. 6] by $\mathcal{D}_\ell^{\text{TT}} \equiv \ell(\ell + 1)C_\ell/2\pi$. The bottom panel shows the residuals of the data (blue points) from the model (red curve). Error bars show 1σ uncertainties. The dashed line at $\ell = 30$ marks a break in the horizontal axis and a change in the vertical scale used in the lower panel. $\ell = 0$ (the monopole — corresponding to the black body temperature) and $\ell = 1$ (the dipole — corresponding to the Doppler shift in temperature due to our motion relative to the CMB) are excluded because these are not related to the density perturbations of the early Universe.

Figure 1.4

Image Credit: Planck Collaboration (2016, Figure 1), reproduced with permission © ESO

1 Introduction

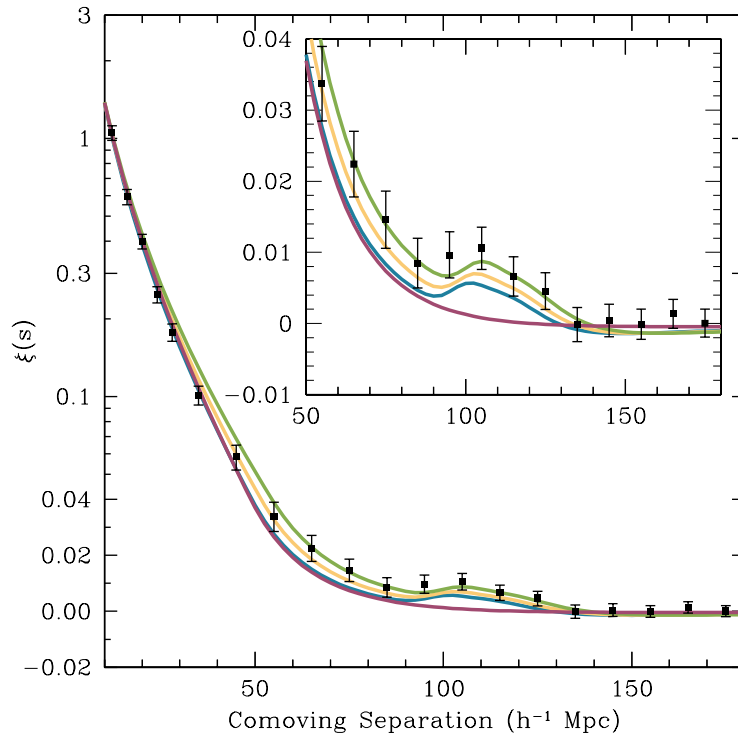


Figure 1.5

The two-point correlation function, $\xi(s)$, based on SDSS Luminous Red Galaxies data. The insert shows a zoom-in view, focusing on the Baryon Acoustic Peak at $\sim 100 \text{ Mpc } h^{-1}$. Different coloured curves refer to different parameter values for the Λ CDM model: $\Omega_m h^2 = 0.12$ (green, top), $\Omega_m h^2 = 0.13$ (yellow), and $\Omega_m h^2 = 0.14$ (blue), all with $\Omega_b h^2 = 0.024$ and $n_s = 0.98$. The bottom red curve ($\Omega_m h^2 = 0.105$) is for a pure CDM model, which is discounted by these observations. The existence of the Baryon Acoustic Peak shows that there is a preferred separation scale between galaxies.

Image Credit: Eisenstein et al. (2005, Figure2), reproduced with permission © AAS

occurred, all oscillations were halted;² the most visible of the oscillatory modes (i.e. the greatest density, and therefore temperature difference) are those at the extrema of their oscillation. Since these modes were all in the same phase at the same time, they must be described by some underlying fundamental frequency. The wavelength of this fundamental frequency is the ‘sound horizon’: the maximum distance an acoustic wave could travel within the plasma by the time of recombination. These density perturbations seeded the observed structure in our Universe and as a result have (some) recollection of the sound horizon. In [Figure 1.5](#)^[p. 8], we show the two-point correlation function calculated for a large number of galaxies. This gives a measure of the probability of finding two galaxies at a given separation length. The data reveal that there is indeed a preferred separation length between galaxies, resulting in a bump within the distribution known as the Baryon Acoustic Peak.

The large scale structure within the Universe becomes more pronounced with time as over-densities gravitationally attract more material. This produces a network of connected filaments known as the ‘cosmic web’. In [Figure 1.6](#)^[p. 10], we show an example of this cosmic web from the Illustris cosmological simulation (Genel et al. 2014; Vogelsberger, Genel, Springel, Torrey, Sijacki, Xu, Snyder, Bird et al. 2014; Vogelsberger, Genel, Springel, Torrey, Sijacki, Xu, Snyder, Nelson et al. 2014; Sijacki et al. 2015). Since within cosmological simulations we can view the dark matter, this network of filaments is clearly visible. These filaments are also reproduced in baryonic matter; however, they are somewhat more diffuse. In the real Universe, the vast majority of gas in the intergalactic medium (IGM) — the material which exists between galaxies and from which galaxies form — is thought to be within these filaments (Umehata et al. 2019). Despite this, the first observations of gas filaments are only just being made (see Umehata et al. 2019) since dark matter and diffuse gas emit no/little detectable emission. However, the nodes (where filaments intersect) are typically the locations at which galaxies form and evolve.

The oldest detected galaxy thus far is at $z = 11.1$ (Oesch et al. 2016); however, after the Epoch of Recombination, the neutral medium present within the Universe became ionised again (Epoch of Reionisation). This is thought to have begun as early as $z \approx 14$ (ending around $z \approx 6$) (Fan et al. 2006) and so the first luminous objects must have existed by ~ 1 Gyr. Once the Universe became neutral again, the Universe entered its present day stage of evolution.

²We note that recombination was not instantaneous, which dampens the effect which we describe here.

1 Introduction

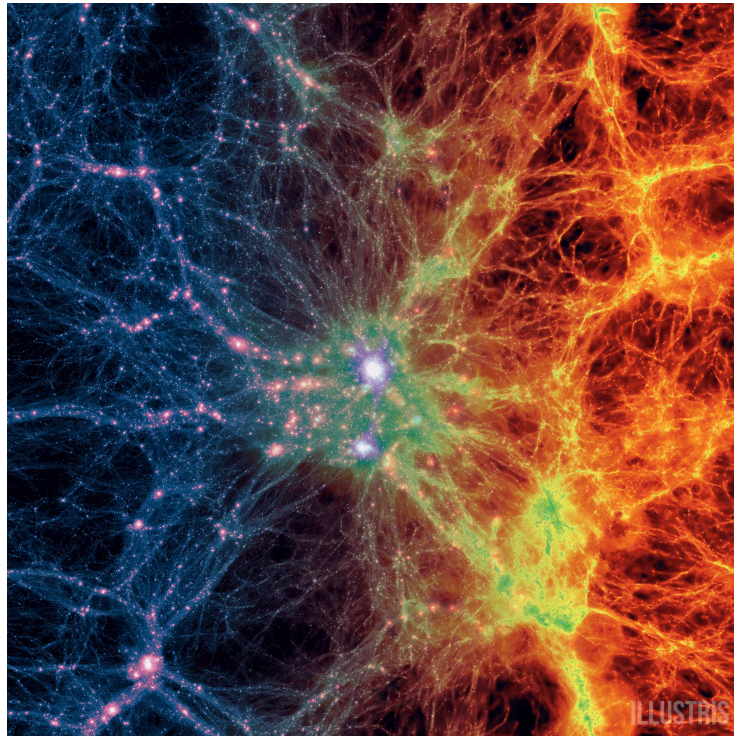


Figure 1.6

A large scale projection of the Illustris volume at $z = 0$ showing the filamentary structure within the Universe. From left to right the image transitions from showing the dark matter density to the gas density. The gas density at the site of nodes, where filaments intersect, is higher and this is typically where galaxies form.

Image Credit: Illustris Collaboration (2018)

1.2 Galaxies

Galaxies, often considered the building blocks of the Universe, are gravitationally bound structures containing stars, gas, dust, dark matter (although galaxies without have been found, Dokkum et al. 2018), and typically a central supermassive black hole (for galaxies with a stellar mass $\gtrsim 10^{10.3} M_{\odot}$, Naab et al. 2017). In Figure 1.7^[p. 12], we present a schematic showing the structure of key components of the two main galaxy types: elliptical and disc. The ‘visible’ part of the galaxy (the part we typically refer to as the galaxy) is where the majority of the galaxy’s stars lie. In elliptical galaxies, the visible part is ellipsoidal and mostly featureless, whereas disc galaxies have two main components: a disc and a central bulge. In some cases the disc is comprised of multiple parts; our own galaxy is thought to have a ‘thin disc’ enveloped by a ‘thick disc’ (although this is still debated Hayden et al. 2017). For both galaxy types this visible part is embedded within a stellar and dark matter halo. The stellar halo contains a population of (typically) older, more metal-poor stars with unordered, elliptical orbits. This region also contains globular clusters: gravitational bound, spherical, and centrally concentrated collections of stars. Globular clusters are long lived structures and sensitive to the properties of their natal environment. This means that they can be used to construct the assembly history of the galaxy in which they are now observed (Kruijssen, Pfeffer et al. 2019).

Galaxies are usually categorised by their morphology based on their appearance in the optical wavelength range. Many classification systems exist (e.g. De Vaucouleurs); a common system is the Hubble Sequence depicted in Figure 1.8^[p. 12], which shows three galaxy types. This system does, however, neglect irregular galaxy, which do not have any clear definable structure, and other features such as rings or an active galactic nucleus (AGN). We summarise the galaxy types included in the Hubble Sequence here.

Elliptical, E_n Ellipticals tend to have an older stellar population. These stars have unordered motion and so there is no disc structure within the galaxy. There is also little gas from which new stars can be produced. The classification number n is calculated as $n = 10(1 - b/a)$, where a and b are the *observed* major and minor axis of the ellipsis, respectively. The higher the value of n the more elliptical the galaxy *appears*. It is not too surprising that no physical properties of elliptical galaxies correlate well with this classification number (Kormendy et al. 1989): the galaxy’s apparent ellipticity from our perspective has no bearing on the galaxy’s inherent properties. However, relations between the galaxy’s luminosity, L , and stellar velocity dispersion, σ , exist through the

1 Introduction

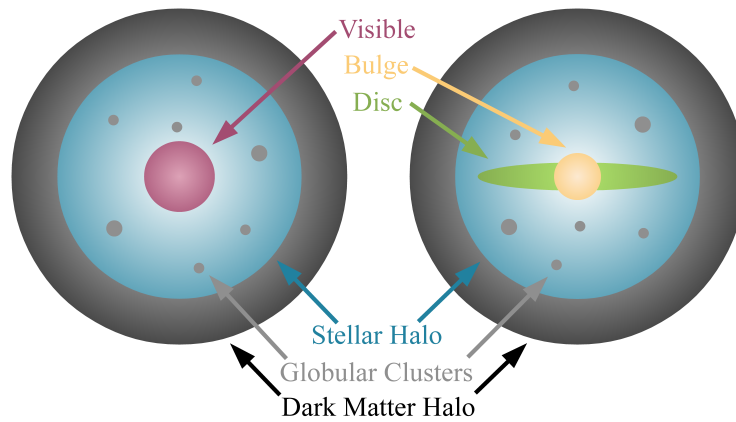


Figure 1.7

Schematic diagram showing the components of an elliptical (left) and disc (right) galaxy. Both galaxy types are situated within a stellar halo and dark matter halo. The stellar halo also contains globular clusters that orbit the galaxy. Elliptical galaxies have only one visible component, whereas disc galaxies are comprised of a disc and bulge.

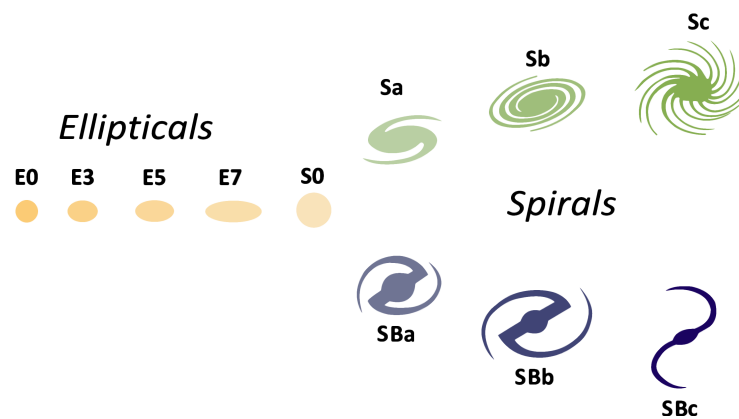


Figure 1.8

The 'Hubble sequence' galaxy classification scheme showing the three galaxy types: ellipticals, En ; lenticular, $S0$; and spirals, $S(B)$ which are subdivided into without (with) a bar. Elliptical galaxies become more oblate with increasing n . Spiral galaxies from 'a' to 'c' have more open spiral arms and smaller bulges. This classification scheme excludes irregular galaxies and other features such as rings.

Image Credit: NASA / ESA (1999), adapted by DTH

empirical Faber-Jackson relation: $L \propto \sigma^\gamma$, where $\gamma \approx 4$ (Faber et al. 1976). This relation is considered to be a projection in mass-velocity space of the Fundamental Plane which also includes a dependence on the half-light radius, R_e : $\sigma \propto L^a R_e^b$, where $a \approx 0.8$ and $b \approx -0.8$ (Cappellari et al. 2013; Desmond et al. 2017). The surface brightness of elliptical galaxies is described by a Sérsic profile characterised through R_e (Blanton et al. 2009).

Lenticular, S0 Lenticular galaxies are something of a transition state between elliptical and spiral galaxies. They have an old stellar population like ellipticals; however, there is a disc, within which their motion is confined. Despite having a disc, there are no visible spiral arms.

Spiral, S Spiral galaxies are subdivided into those with a bar (SB) and those without (S, or sometimes SA). They have a young stellar population with ordered orbits confined to the disc. However, in the bulge the orbits are more disordered. The sub-classifications from ‘a’ to ‘c’ indicate more open spiral arms and a decreasing bulge size. In analogy with the Faber-Jackson relation, there is also a relation between the spiral galaxy’s luminosity and rotational velocity, V_{rot} . This empirical relation, known as the Tully-Fisher relation, is $L \propto V_{\text{rot}}^\alpha$, where $\alpha = 3.64$ (Tully et al. 1977; Torres-Flores et al. 2011). The fact that lenticular galaxies also follow the Tully-Fisher relation (although offset from spirals) supports the idea that they are evolved spiral galaxies (Bedregal et al. 2006). Typically spirals have a surface brightness with an exponential profile for the disc and Sérsic profile for the bulge (Blanton et al. 2009). The disc profile is characterised by the ‘scale height’ (h_z), which is related to the vertical velocity dispersion of the stars (Kruit et al. 2011). The disc is also where the majority of the galaxy’s dust is found. Spiral arms are complicated and long-lived structures found in the galaxy disc; they are often understood through density waves (Shu 2016), with an angular velocity characterised by the pattern speed, Ω_p . In the density wave model, material moves in and out of the spiral arms; however, there is a build up in the high density region. Gas entering this high density region can be triggered to form stars, resulting in a region of young stars at the leading edge of the arm (Martínez-García et al. 2009) and perpetuating the spiral arm’s existence. This model explains the appearance of ‘grand design’ spirals, where the spiral arms are clearly defined. For ‘flocculent’ galaxies, where spiral arm structure is less well defined, the appearance is better described by self-propagating star formation (Gerola et al. 1978). In this case, the star formation can trigger further star formation in the local region as a result of feedback. Two example galaxies with flocculent and grand design spiral arms are shown in Figure 1.9^[p. 15]. One of these galaxies

1 Introduction

also shows a prominent bar, from which the spiral arms emerge. Bars are thought to be short-lived structures and formed as a result of density waves (Bournaud et al. 2002).

In order to study large populations of galaxies in a statistical way, we need to define distributions based on observable quantities. The distribution of galaxy luminosity (without regard for morphology) is often used. The number of galaxies, n , for a given luminosity, L , is often described by the Schechter luminosity function (Schechter 1976):

$$\text{Equation 1.12} \quad \frac{dn}{dL} = \frac{n_\star}{L_\star} \left(\frac{L}{L_\star} \right)^\alpha \exp \left(-\frac{L}{L_\star} \right),$$

where n_\star is a normalising factor giving the mean number density of bright galaxies, L_\star is the characteristic galaxy luminosity, and α is the slope of the faint end of the distribution. Galaxies with luminosities $\lesssim L_\star/100$ are usually considered dwarf galaxies. Since luminosity is affected by distance, it is often more convenient to express Equation 1.12 in terms of magnitudes, M :

$$\text{Equation 1.13} \quad M - M_\star = -2.5 \log_{10} \left(\frac{L}{L_\star} \right),$$

which gives

$$\text{Equation 1.14} \quad \frac{dn}{dM} = 0.4 \ln(10) n_\star \left(10^{0.4(M_\star - M)} \right)^{\alpha+1} \exp \left(-10^{0.4(M_\star - M)} \right).$$

The values of the free parameters in this distribution, as found by Driver et al. (2003), are $M_\star - 5 \log_{10}(h_{0.68}) = -20.47 \pm 0.2$ mag, $\alpha = -1.1 \pm 0.1$, and $n_\star = 0.006 \pm 0.001 h_{0.68}^3 \text{ Mpc}^3$, where $h_{0.68} \equiv H_0/68 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (cf. Equation 1.7^[p. 2]).

Comparisons of the observed luminosity function can be made with cosmological simulations. In cosmological simulations the number of galaxies can be predicted based on the number of dark matter halos. The predictions overestimate the number of galaxies at both low (the ‘missing satellite problem’) and high luminosities. However, the inclusion of baryonic physics and feedback mechanisms is thought to alleviate if not solve the discrepancy (Benson et al. 2003; Bower et al. 2006; Dashyan et al. 2019).

The exact sequence of events which form a galaxy are still unknown. The original model for galaxy formation was the gravitational collapse of a primordial gas cloud into a single galaxy, known as monolithic collapse. This was the case for both elliptical galaxies (Larson 1975) and disc galaxies (Eggen et al. 1962); the differentiation into the two types was a matter of

1.2 Galaxies



Top: An image of NGC1300, a barred grand-design spiral galaxy, which is classified as SBbc in the Hubble sequence (Vaucouleurs et al. 1991). Bottom: An image of NGC4414, a flocculent spiral galaxy, which is classified as Sc in the Hubble sequence (Vaucouleurs et al. 1991).

Figure 1.9

Image Credit: top — Hubble Heritage Team / ESA / NASA (2016)

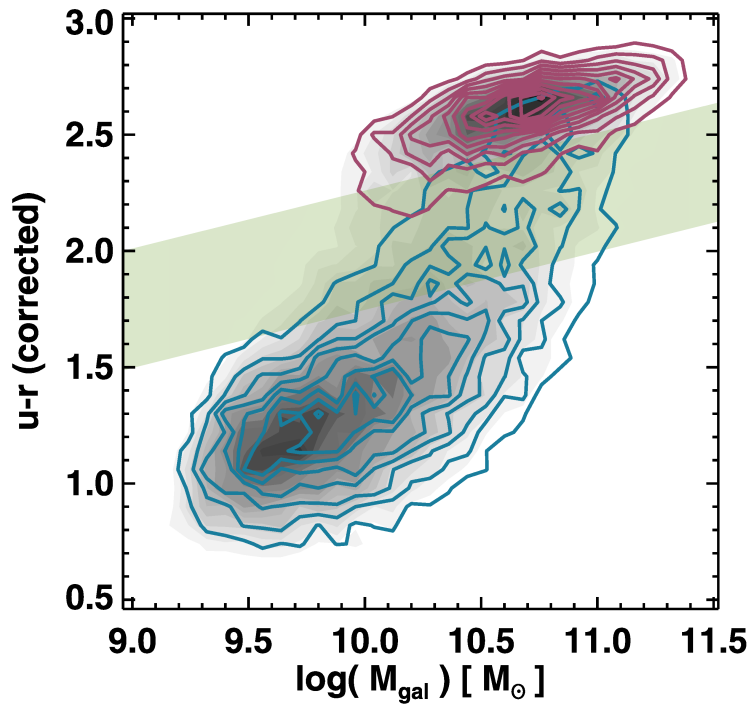
bottom — NASA / ESA / Freedman, W. / Hubble Heritage Team / SDSS (2016)

1 Introduction

rotation (Sandage et al. 1970). If the system was rotating, this rotation would prevent collapse in the radial direction but not in the vertical, resulting in a disc. However, it was later suggested that elliptical galaxies could be the product of galaxy interactions or mergers (Toomre et al. 1972). Under the Λ CDM model, a popular galaxy formation scenario is the hierarchical growth of structure, where structure builds up from the gradual accumulation of objects into progressively larger units (e.g. small dark matter clumps into dark matter halos). In this scenario, it is believed that discs are formed first; elliptical galaxies and the bulges of spirals form through subsequent mergers of these discs (Baugh et al. 1996). This is supported by the idea that elliptical galaxies usually have an old stellar population.

The evolutionary state of a galaxy is typically based on its stellar population and categorised through its location in colour-mass diagrams. An example of such a diagram is shown in [Figure 1.10](#)^[p. 17]. It is clear from galaxy surveys that the distribution in this space is bimodal (Baldry et al. 2006; Alatalo et al. 2014; Schawinski et al. 2014): ‘blue’ galaxies have a young stellar population and are actively star forming, whereas ‘red’ (and dead) galaxies have an old population of stars and are no longer actively forming new ones. This bimodality aligns itself closely with galaxy type; as indicated by the contours in [Figure 1.10](#)^[p. 17]. As galaxies evolve, they transition between these two states through what is known as the ‘green valley’. We indicate this in [Figure 1.10](#)^[p. 17] with the shaded green band. The passage through the green valley does not necessarily occur at the same speed for all galaxies and is thought to primarily go in the direction of blue to red; however, in some instances galaxies can ‘come back from the dead’ (Schawinski et al. 2014; Wright et al. 2019); that is, start forming stars again. How ‘alive’ a galaxy is depends on the galaxy’s ability to form stars and therefore the availability of the raw material, gas.

The gas cycle in galaxies is complex. We present an illustration of this cycle in [Figure 1.11](#)^[p. 18]. Gas can be brought into the galaxy through filamentary accretion from the IGM and removed through galactic outflows as a result of AGN activity. Some of this outflowing material can be recaptured and brought back into the galaxy. This material becomes part of the interstellar medium (ISM). Of course, stars forming from the ISM are a means of locking away gas; although, some of this is eventually returned during a star’s life cycle. The quenching of star formation (where the transition through the green valley begins) can be the result of many mechanisms and is still uncertain (Darvish et al. 2018). Quenching could be a result of the active removal of gas (e.g. stripping, Poggianti et al. 2017) or the prevention of new material entering the system (e.g. strangulation, Peng et al. 2015). These mechanisms might initially enhance star formation but ultimately bring about the demise of the



A dust corrected $u-r$ colour-mass diagram for WISE-detected Galaxy Zoo galaxies (Schawinski et al. 2014). The red contours (top) show the location of elliptical galaxies and the blue contours (bottom) of disc galaxies. The green band shows the ‘green valley’ which divides the bimodal distribution of galaxies (shown in grey scale). This region of the parameter space is where galaxies in transition between the two states are located.

Figure 1.10

Image Credit: Alatalo et al. (2014, Figure 1), reproduced with permission © AAS

1 Introduction

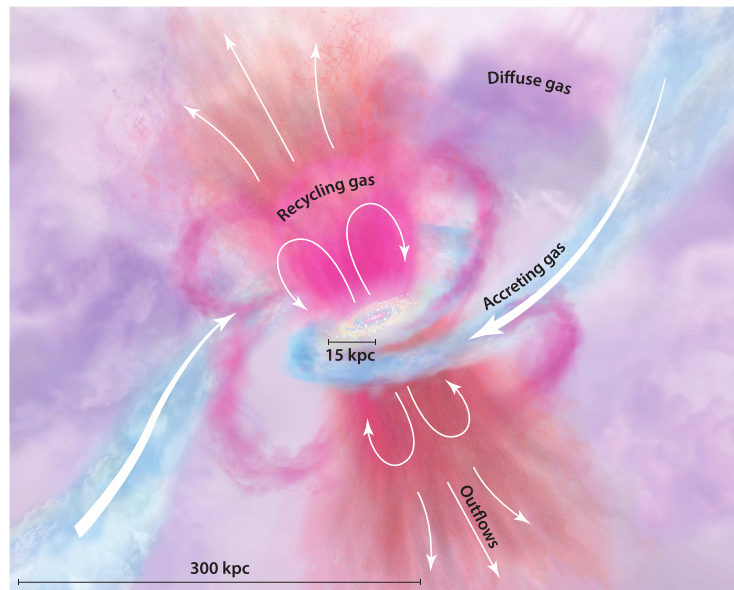


Figure 1.11

A visual representation of the gas cycle in a spiral galaxy. New gas is brought into the galaxy from the IGM through filamentary accretion (blue). This material once in the galaxy becomes part of the ISM, from which stars are born and into which they return material through feedback processes (see [Section 1.3](#)^[19]). Outflows from the galaxy, as a result of AGN activity remove this gas from the galaxy (orange). Some of this outflowing material does not escape from the galaxy and is recycled (pink). Over time, the diffuse gas halo (purple) becomes a mix of all these sources.

Image Credit: Tumlinson et al. (2017, Figure 1), reproduced with permission © Annual Reviews

galaxy as the gas reservoir is rapidly drained. For example, galaxy mergers can bring in new material (Boselli et al. 2006; Schawinski et al. 2014) and the stripping of gas can cause its compression and subsequent star formation (Bekki et al. 2003). Feedback from AGN (Fabian 2012) as well as stars themselves (Hopkins et al. 2014), can also halt further star formation.

Galaxies do not always evolve in isolation and can be part of a group of galaxies (typically ~ 50 members) or cluster of galaxies (typically hundreds to thousands of members). The distinction between a group and clusters is not always clear cut: they are both collections of gravitationally bound galaxies. However, clusters are more massive and cover a much larger volume than groups. Work by Paul et al. (2017), using cosmological hydrodynamic + N-body simulations, found a break point at a cluster mass of $\sim 8 \times 10^{13} M_{\odot}$ in the relation for all studied quantities including X-ray luminosity and baryon fraction; this suggests a formal distinction between the two.

It has been observed that galaxies within clusters are more likely to be elliptical; the reverse holds true for isolated galaxies. In fact, the fraction of spiral galaxies decreases with increasing increasing local galaxy density (Dressler 1980; Houghton 2015). This relation is known as the morphology-density relation. It is still unclear if this relation is governed by initial conditions or later environmental effects. The local environment does not just affect galaxies; it has also been shown that star formation, although occurring on scales much smaller than that of the galaxy, is also sensitive to the galaxy's local environment (Kauffmann et al. 2004).

1.3 Stars

Star formation begins in giant molecular clouds (GMCs) typically located in the discs of spiral galaxies. These GMCs are formed out of the material found in the galaxies ISM. In Figure 1.12^[p.20], we present a schematic of the steps from gas cloud, to protostar, to pre-main-sequence star, and finally (if sufficiently massive) to main-sequence star; we give further details here. Labels (e.g. a, b) in the following text refer to those in Figure 1.12^[p.20]. Protostars and pre-main-sequence stars are collectively known as young stellar objects (YSOs). A classification of YSOs based on the spectral index of their infrared (IR) spectral energy distribution (SED) was introduced by Lada (1987):

$$\alpha = \frac{d \log_{10} (\lambda F_{\lambda})}{d\lambda},$$

Equation 1.15

where λ is the wavelength, and F_{λ} the flux density. The value range of α for each of the classes (0–3), depends on the specifics of the scheme being used.

1 Introduction

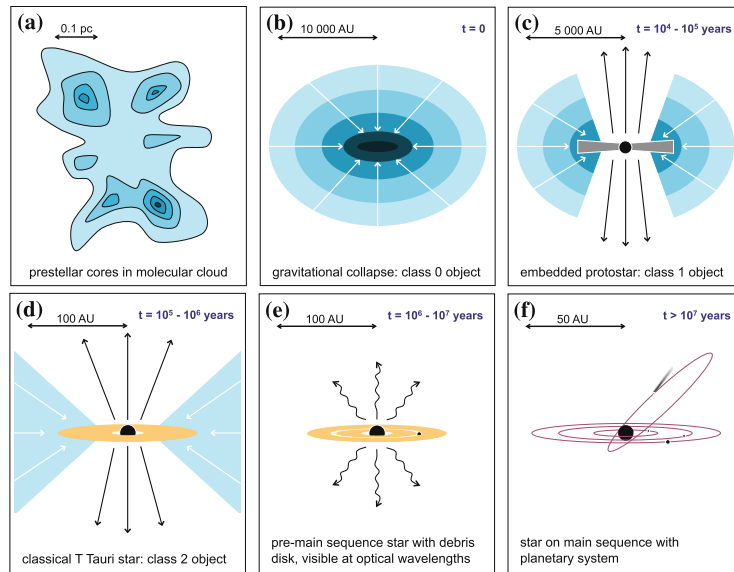


Figure 1.12

A schematic showing the sequence of events that turn molecular clouds into stars. Overdensities within the molecular cloud (prestellar cores) are where the star formation process begins (a). These prestellar cores begin to collapse under their own gravity (b). The in-falling material produces a protostellar disc surrounding the protostar (c). Outflows perpendicular to the disc start to remove the envelope in which the protostar is embedded, which eventually reveals the protostar (d). By the time the gas is removed, a pre-main-sequence star remains; however, still surrounded by a protoplanetary disc (e). Eventually the disc is cleared and a main-sequence star (possibly) with a planetary system is what remains.

Image Credit: Klessen et al. (2016, Figure 23), reproduced with permission © Springer Nature

For mid-IR, Greene et al. (1994) give the ranges Class 0, no detection; Class 1 (protostar), $\alpha > 0.3$; Class 2 ('classic' T Tauri), $-0.3 > \alpha \geq -1.6$; and Class 3 ('weak' T Tauri), $\alpha < -1.6$.

The internal structure of a GMC is not uniform. Substructures form within the GMC as a result of turbulence, producing regions of higher density known as prestellar cores (a). The subsequent collapse of these cores due to gravity into protostars is subject to various supportive mechanisms (b). The turbulent motion of the gas itself is also a supportive mechanism. The first theoretical approach to the criteria for collapse was introduced by Jeans (1902, 1928). The Jeans mass, M_J , gives the mass of a gas cloud with uniform density ρ and temperature T that can no longer be supported by internal gas pressure (Bodenheimer 2011):

$$M_J = \left(\frac{4}{3} \pi \rho \right)^{-1/2} \left(\frac{5 RT}{2 \mu G} \right)^{3/2} \quad \text{Equation 1.16}$$

where R is the gas constant, and μ is the mean molecular mass of the gas. However, internal gas pressure is not the only means to prevent collapse. Approximately spherical regions within a rotating disc can gain additional support against collapse from shearing force. This has been characterised by the Toore Q parameter (Toomre 1964):

$$Q = \frac{\kappa c_s}{\pi G \Sigma_{\text{gas}}} , \quad \text{Equation 1.17}$$

where κ is epicyclic frequency, c_s the sound speed, and Σ_{gas} the gas surface density. Regions are stable against collapse if $Q > 1$. Another source of support is from magnetic fields, but this is highly directional. That is, collapse could still be possible in the direction parallel to the magnetic field lines. Analogous to the Jeans mass, M_J , the critical mass M_ϕ is given as (Mouschovias et al. 1976):

$$M_\phi = \frac{0.53}{3\pi} \phi \sqrt{\frac{5}{G}} , \quad \text{Equation 1.18}$$

where ϕ is the magnetic field flux. Nevertheless, even intrinsically stable cores can be triggered to collapse by external events such as shock waves or cloud-cloud collisions.

Once star formation is triggered, the protostar starts to remove the surrounding envelope of material through outflows, whilst also accreting mass from the protostellar disc (c). The protostar becomes detectable in the IR and has a 'Class 1' SED. Accretion onto the protostar continues (as well as outflows) revealing more of the protostellar disc and causing a shift in the shape of the

1 Introduction

SED, now Class 2 SED (d). Emission lines, particularly $H\alpha$ as well as a high abundance of rare, light elements such as Lithium, which are easily destroyed, are key signatures of these protostars (Bodenheimer 2011) — often called (classic) T Tauri stars named after the prototype star T Tauri (discovered by Hind 1864). With further evolution these features become less pronounced as the rate of accretion decreases the pre-main sequence star is now classed as weak T Tauri/Class 3 SED (e). The star only enters the main sequence once the core begins Hydrogen burning within its core (f).

The study of the gas from which stars form (i.e. the ISM) is an important part of understanding star formation. The gas within galaxies is predominately Hydrogen; however, this is not easily observed. This is because high temperatures are required to excite the rotational energy transitions as a result of the molecule's low mass (Kennicutt et al. 2012). Instead other molecules must be used to trace the cloud structure and dynamics. The effectiveness of these different tracers depends on the properties of the gas — chiefly the density and temperature. The amount of emission from these different tracers can therefore inform us about the conditions of the gas present. In Figure 1.13^[p.23], we show the range in density and temperature for which different tracers operate best.

A commonly used tracer for observing molecular gas is CO, as it has strong and easily observable emission lines in the millimetre wavelength regime. The molecule can only survive in dense gas regions, where it is protected from photodissociation. However, these are the regions which are typically suited to star formation. The emission from different transitions in energy state are denoted by the initial and final total angular momentum, J , which is related to the total angular momentum quantum number, j :

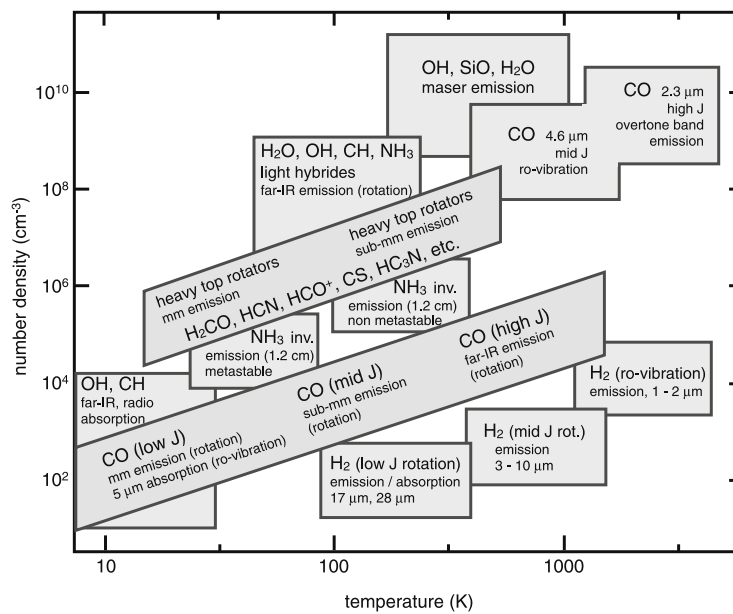
Equation 1.19

$$J^2 = \hbar^2 j(j + 1) ,$$

where \hbar is the Plank constant. To get a measure of the amount of molecular Hydrogen, work has been done to measure the CO-to- H_2 conversion factor, X_{CO} .³ This conversion factor has been estimated for the Milky Way to be $X_{CO} = 2 \times 10^2 \text{ cm}^{-2} (\text{K km s}^{-1})^{-1}$ but there is still uncertainty in how X_{CO} changes with environment (Bolatto et al. 2013).

The ISM, from which stars form, is often modelled as having distinct phases. We summarise these phases here in order of decreasing temperature, T , and increasing Hydrogen number density, n_H . We also include the fraction of the volume occupied by gas in this phase, f_v , and the predominant 'state' of Hydrogen. Further details can be found in Draine (2011), amongst other texts.

³This conversion factor is sometimes referred to as α_{CO} which is usually given in units of $M_\odot (\text{K km s}^{-1} \text{ pc}^2)^{-1}$.



The emission tracers that are commonly used to observe gas cloud structure and dynamics. Regions of the figure mark the temperature and density ranges of the gas, for which different molecules are best suited. CO is of particular use as its emission is suitable for gas densities and temperatures associated with star forming gas.

Figure 1.13

Image Credit: Klessen et al. (2016, Figure 8), reproduced with permission © Springer Nature

1 Introduction

Hot ionised medium ($T \gtrsim 10^{5.5}$ K, $n_{\text{H}} \approx 4 \times 10^{-3}$ cm $^{-3}$, $f_{\text{v}} \approx 0.5$, H II)

Gas shock-heated and collisionally ionised by blast waves from supernova explosions. Typically filling the volume above and below the disc. This gas cools on \sim Myr timescale through adiabatic expansion and X-ray emission.

Warm ionised medium ($T \approx 10^4$ K, $n_{\text{H}} \approx 0.3\text{--}10^4$ cm $^{-3}$, $f_{\text{v}} \approx 0.1$, H II)

Gas ionised by ultraviolet (UV) photons from young massive stars and cools through line emission. This gas can be observed by optical line emission (such as H α emission) and falls in the mid-plane of disc. This gas phase is also associated with recent star formation (H II region) and the ejected material from Red Giants (planetary nebular).

Warm neutral medium ($T \approx 5 \times 10^3$ K, $n_{\text{H}} \approx 0.6$ cm $^{-3}$, $f_{\text{v}} \approx 0.4$, H I)

Gas heated by photoelectrons from dust grains. This gas occupies a large fraction of the disc's volume and can be observed by the 21 cm line.

Cold neutral medium ($T \approx 100$ K, $n_{\text{H}} \approx 30$ cm $^{-3}$, $f_{\text{v}} \approx 0.01$, H I)

Gas heated by photoelectrons from dust grains and cooled through fine structure line emission. Usually observed by the 21 cm line and is typically associated with clouds within the disc.

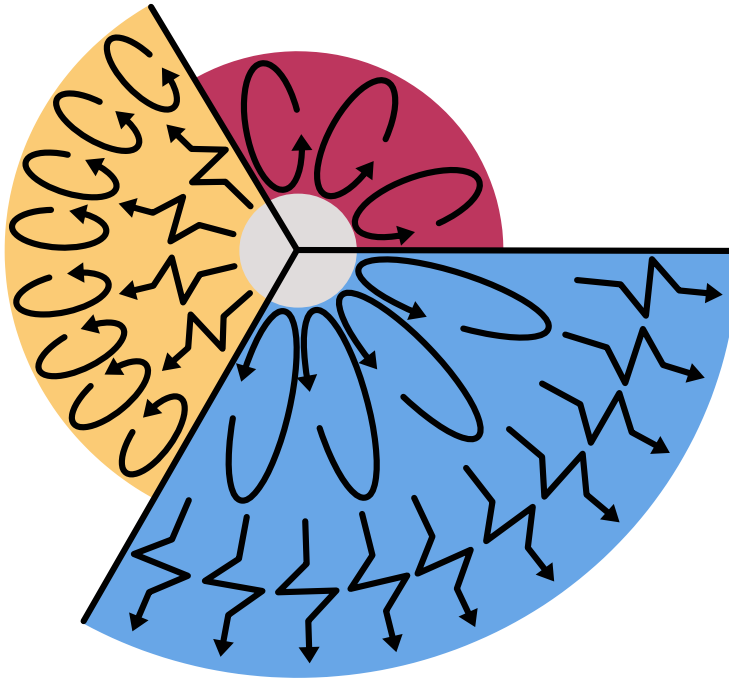
Diffuse molecular gas ($T \approx 50$ K, $n_{\text{H}} \approx 100$ cm $^{-3}$, $f_{\text{v}} \approx 10^{-3}$, H $_2$)

Similar to the cold neutral medium; however, now of sufficient density to shield the interior of the cloud from ionisation (referred to as self-shielding). These densities are also sufficient to allow for the existence of CO (and its associated emission).

Dense molecular gas ($T \approx 10\text{--}50$ K, $n_{\text{H}} \approx 10^3\text{--}10^6$ cm $^{-3}$, $f_{\text{v}} \approx 10^{-4}$, H $_2$)

This gas is found in gravitationally bound clouds and still emitting in CO. This gas can also be traced through far-IR emission from dust. This is also the site of star formation.

Once a star has formed, its evolution is determined predominately by its mass; that is, the sequence of phases as well as the speed at which the star evolves. More massive stars are much shorter lived than less massive ones: a greater outward pressure is required to support the star against collapse and so the consumption of Hydrogen is faster. The luminosity of a star, L , is tightly related to the mass, M , of the star (known as the Mass–luminosity relation): $L \propto M^\alpha$ (Kuiper 1938). The means of internal energy transportation is also mass dependent, shown in [Figure 1.14](#)^[p.25].



The cross section of a star for different masses, M : $M < 0.5 M_{\odot}$ in red, $0.5 M_{\odot} < M < 1.5 M_{\odot}$ in yellow, and $M > 1.5 M_{\odot}$ in blue. The central region (in grey) is the core, where nuclear fusion of Hydrogen occurs. In low mass stars (red) the transportation of energy is primarily conducted through convection. As the mass of the star increases a radiative zone forms in the central region and the convective envelope moves outwards (yellow). Eventually the convective envelope is removed completely and a convective core begins to grow (blue). Unlike radiative zones, convective zones allow the mixing of material, which can aid the fusion process by bringing new fuel into the core. Convection typically occurs when the temperature gradient is sufficiently large; this could be as a result of high energy production or having a high opacity. This is more formally given by the Schwarzschild criterion for stability against convection (Schwarzschild 1958). (The schematic is not to scale.)

Figure 1.14

1 Introduction

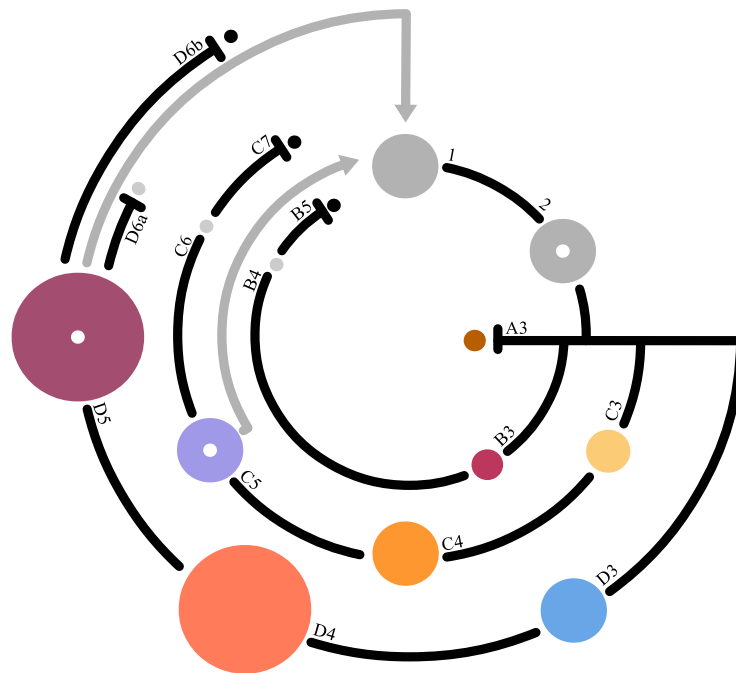


Figure 1.15

The evolutionary stages of a star are primarily dependent on the mass of the star. The black 'rings' indicate the evolutionary sequence, where outer rings are for higher mass stars. The grey arrows indicate stages which return material back to the ISM. Bars mark the end of a sequence and the material that is locked away. See text for further details.

We present a schematic of the possible evolutionary sequences of a star in [Figure 1.15](#)^[p. 26] and give details here. The labels we use in the following text refer to those in [Figure 1.15](#)^[p. 26]. As described above, gas clouds form in the ISM (1) and evolve into protostars (2).

Stars with a mass $< 0.08 M_{\odot}$ are typically referred to as Brown Dwarfs (A3). Brown Dwarfs are of insufficient mass to begin Hydrogen burning and so do not enter the main sequence. However, they are thought to sustain Deuterium and Lithium burning.

With a mass $< 0.20 M_{\odot}$ (Laughlin et al. 1997), the star is referred to as a Red Dwarf (B3). This is the smallest star (by mass) that enters the main sequence, fusing Hydrogen into Helium through the Proton-Proton chain. With such a low mass, the life span of a Red Dwarf is calculated to be longer than the current age of the Universe. However, it is thought that they will slowly collapse into a White Dwarf (B4). Little to no fusion processes take place within the White Dwarf and the emission seen is that from the stored thermal energy. This White Dwarf will eventually cool into a Black Dwarf (B5).

Stellar masses $< 8 M_{\odot}$ result in Low-Mass (or Sun-Like) stars (C3). In lower mass stars, Hydrogen burning is conducted primarily through the Proton-Proton Chain. Higher mass stars use the CNO-cycle, where Carbon, Nitrogen and Oxygen are used as catalysts to ultimately perform the same conversion of Hydrogen into Helium. High-Mass stars $> 8 M_{\odot}$, undergo much the same process; however, they leave the main sequence more quickly (D3). These are stars with an OB spectral class. As the Hydrogen becomes depleted the outward radiation pressure decreases and the core succumbs to gravity. The contraction releases energy and ‘shell-burning’ — Hydrogen fusion in a shell around the core — begins. This new release of energy causes an expansion of the outer layers. The star has left the main sequence to become a Red Giant (C4) or Red Supergiant (D4). Stars remain in this phase as long as they are able to produce sufficient energy to counteract gravitational collapse. More massive stars are able to burn heavier elements (up to Iron) in order to release energy; however, these are also the stars which require greater support against gravitational collapse and so consume fuel more quickly.

Low-Mass stars will blow out the outer layers of the star to produce a metal-rich planetary nebula (C5) and central White Dwarf (C6), which will cool into a Black Dwarf (C7). The ejected material is recycled back into the ISM.

High-Mass stars which can no longer support themselves against gravity through nuclear fusion, will undergo core collapse. It is thought that the inner core is compressed into neutrons by means of electron capture. The innermost regions, once maximally compressed, are thought to rebound. This produces

1 Introduction

an outward flowing shock wave, which is supported by the energy release from the collapse and escaping electron neutrinos. This ultimately results in a Type II supernova (D5). The ejected material, containing metals (even some more massive than Iron), is again recycled back into the the ISM. If the star had an initial mass of $< 40 M_{\odot}$, the remnant of the explosion will be a Neutron Star supported by neutron degeneracy pressure (D6a); otherwise, the core will collapse below the Schwarzschild radius to form a Black Hole (D6b).

Stars are usually categorised by their location within the Hertzsprung–Russell diagram (stellar luminosity against its temperature) or equivalently Colour–Magnitude diagram (apparent magnitude against its colour). An example Hertzsprung–Russell diagram is shown in Figure 1.16^[p. 29]. The path a star takes within the Hertzsprung–Russell diagram as it evolves is referred to as the (stellar) evolutionary track. The evolution of stars is understood sufficiently that model evolutionary tracks can be produced. In Figure 1.17^[p. 30], we show the theoretical evolutionary tracks from the Geneva solar metallicity models. These models can be used to age an observed stellar population and play an important role in stellar population synthesis (SPS) modelling (e.g. in the SLUG2 and STARBURST99 codes which we use in Chapters 3 and 4^[pp. 47 & 99]). However, for SPS modelling one additional component is required: the stellar initial mass function (IMF). The IMF is an empirical relation that describes the number of stars, n , of mass m (upon entering the main sequence) for a given population of stars. Since stellar mass governs many aspects of a star, the IMF is an important diagnostic tool. Observations generally support the notion that the IMF is universal: that is, independent of environment and unchanging over cosmic time (Bastian et al. 2010). This relation was first quantified by Salpeter (1955) through the power law relation:

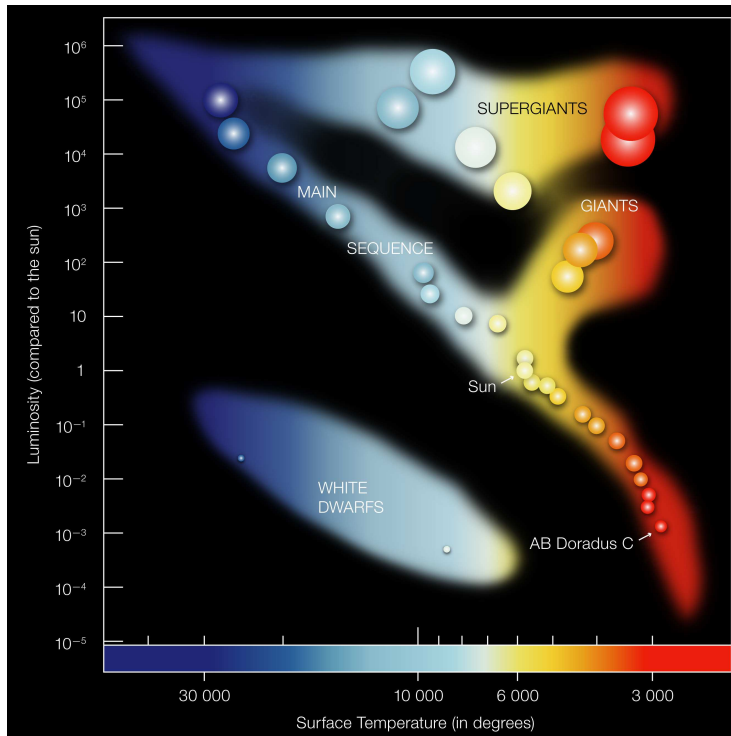
$$\text{Equation 1.20} \quad \frac{dn}{d \log_{10}(m)} = 0.03 \left(\frac{m}{M_{\odot}} \right)^{-1.35}.$$

Other IMFs have since been defined but this is to change the behaviour of the distribution for low stellar masses $< 1 M_{\odot}$. For example the Chabrier (2005) ISM is given as

$$\text{Equation 1.21} \quad \frac{dn}{d \log_{10}(m)} = \begin{cases} 0.093 \exp \left[-\frac{\log_{10}^2(m/0.2)}{2 \times 0.55^2} \right] & m \leq 1 M_{\odot} \\ 0.041 m^{-1.35 \pm 0.3} & m \geq 1 M_{\odot} \end{cases}.$$

With these two components it is possible to model the evolution of a stellar population; this includes, for example, the emission that would be observed and the expected number of Type II supernovae.

1.3 Stars



A schematic Hertzsprung-Russell Diagram, showing the location of different stellar types in the parameter space. The diagonal region marks the main sequence, where stars are producing energy through Hydrogen fusion. Stars evolve off (not along) the main sequence into Giants, Supergiants and/or White Dwarfs (see text for details). The location of the Sun has been marked, as well as AB Doradus C. AB Doradus C is one of the lowest mass stars observed thus far: $M = 0.090 \pm 0.003 M_{\odot}$ (Nielsen et al. 2005).

Figure 1.16

Image Credit: ESO (2007)

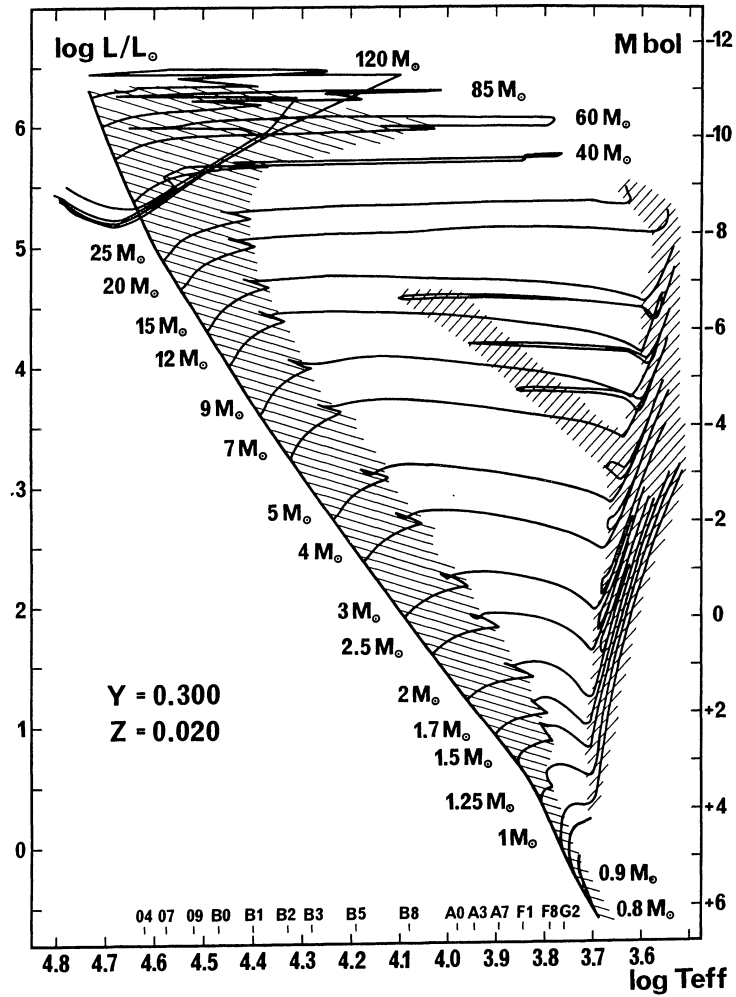


Figure 1.17

The Geneva stellar evolutionary tracks for solar metallicity stars, $Z = 0.020$, with initial Helium mass fraction $Y = 0.300$. The black lines are the theoretical models of how a star of the indicated mass moves within the Hertzsprung–Russell diagram. The left vertical axis is the luminosity (in units of solar luminosity L_{\odot}) and on the right is the bolometric magnitude. The inner horizontal axis marks the spectral class and the outer axis the temperature.

Image Credit: Schaller et al. (1992, Figure 1), reproduced with permission © ESO

Stellar emission and evolution is predictable enough, that some emission is directly linked to the star formation rate (SFR) (Kennicutt et al. 2012). For example, massive stars are ideal for measuring the instantaneous SFR as they are short lived. These massive stars produce far-ultraviolet (FUV) and near-ultraviolet (NUV) emission, which are common SFR tracers. This UV emission can also ionise the surrounding gas to produce H α emission or heat surrounding dust which re-radiates as 24 μm emission (useful for detecting embedded star formation). Since these stars are short lived there are tight constraints on the duration for which this emission can be observed (Hao et al. 2011; Kennicutt et al. 2012).

1.4 This work

For many years, the SFR surface density, Σ_{SFR} , of a galaxy has been related to the gas surface density, Σ_{gas} , through the Kennicutt-Schmidt relation (Schmidt 1959; Kennicutt 1998)

$$\Sigma_{\text{SFR}} \propto \Sigma_{\text{gas}}^N .$$

Equation 1.22

This relation is shown in Figure 1.18^[p.32], with the power-law index $N = 1.40$ (Kennicutt 1998). This is a purely empirical relation, meaning that (although observed) nothing about the nature of the interaction between the SFR and ISM is actually described. For example, star formation recipes used within simulations that do not recover this relation are most likely incorrect; of the many recipes that do, this relation has no means of distinguishing which are correct. Even quantities that must factor into this relation are lost to us: the efficiency of star formation (i.e. how much of the gas is converted into stars) and how quickly it proceeds are degenerate quantities that this relation cannot disentangle.

The Kennicutt-Schmidt relation is based on measurements of whole galaxies or large components thereof. More recent observations, with improved spatial resolution, are able to recreate this Kennicutt-Schmidt relation based on the measurements of a single galaxy. Work by Bigiel et al. (2008) (amongst others) shows that as the spatial resolution improves the scatter around this relation increases dramatically (see Figure 1.19^[p.33]). This is because small scale observations start to capture the different evolutionary stages of the star formation process, whereas galactic scale measurements give an ensemble average. That is, the small scale observations detect the spatial decorrelation of the gas and stars which occurs as a result of star formation. The study of this decorrelation can lead to an improved understanding of the star formation

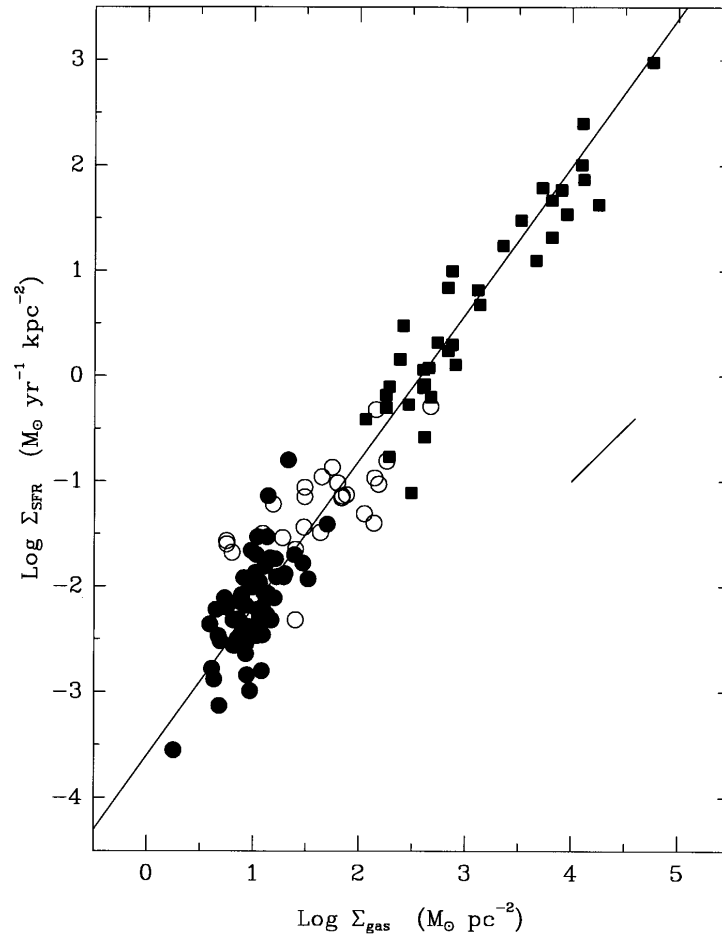
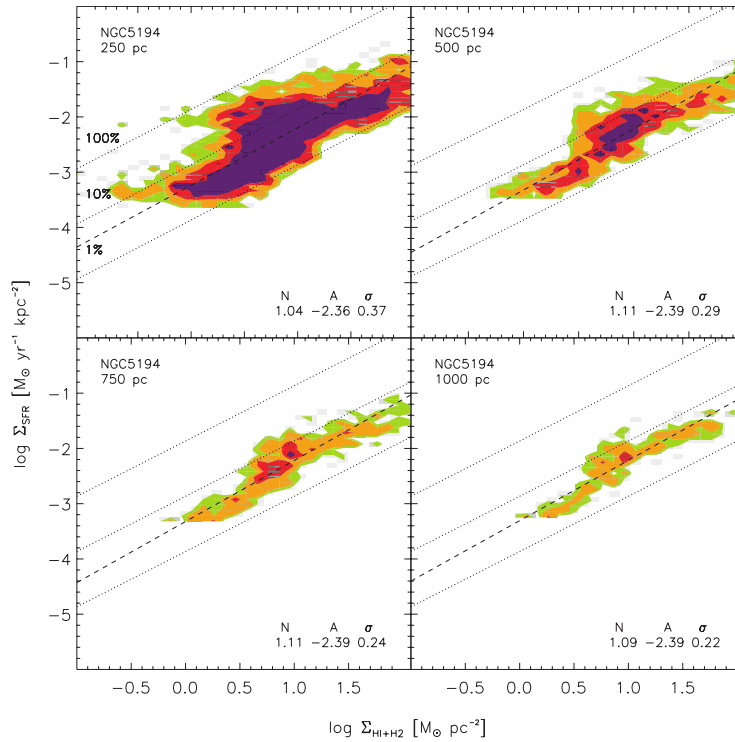


Figure 1.18

The relation between the SFR surface density, Σ_{SFR} , and gas surface density, Σ_{gas} . The solid black line marks the Kennicutt-Schmidt relation with a power-law index of $N = 1.4$. The filled squares are measurements of starburst galaxies, filled circles the discs of normal spiral galaxies, and open circles the centres of normal spiral galaxies.

Image Credit: Kennicutt (1998, Figure 6), reproduced with permission © AAS



Re-creation of the Kennicutt-Schmidt relation based on observations of NGC5194 at four different spatial resolutions: 250, 500, 750, 1000 pc. Although the trend remains largely the same for the four different resolutions, the scatter (σ) increases dramatically with improved resolution. The proportionality constant, A , and power-law index, N , for the Kennicutt-Schmidt relation are indicated in each panel and shown as the dashed line. Dotted lines indicate the level of SFR needed to consume 1, 10, and 100 per cent of the gas reservoir in 10^8 yr.

Image Credit: Bigiel et al. (2008, Figure 6), reproduced with permission © AAS

Figure 1.19

1 Introduction

process. Tamburro et al. (2008) looked at 21 cm emission (tracing neutral Hydrogen) and 24 μm emission (tracing embedded star formation) within spiral arms. Knowing the pattern speed of the spiral arms and the galaxy's rotation curve, the angular offset between emission peaks in the two tracers was used to place a time constraint of 1–4 Myr on the evolutionary sequence. Kawamura et al. (2009) used catalogues of GMCs and young stellar clusters in the Large Magellanic Cloud to determine the number of clouds with and without signs of star formation. Assuming that the number of clouds falling into each category is proportional to the duration in that phase (as well as adopting a duration of 10 Myr for the young stellar clusters), a GMC lifetime of about 20–30 Myr was found. A refinement of this method has been introduced by Kruijssen et al. (2014, 2018). This new statistical method, the ‘uncertainty principle for star formation’ (which we abbreviate to ‘KL14 principle’), constrains evolutionary timescales without the need of spiral arms or catalogues of clouds and stellar clusters. In fact, this method has the benefit of being agnostic to the definition of ‘cloud’ and ‘cluster’. We use this method extensively in Chapters 3 and 4^[pp. 47 & 99] and so we leave a detailed description of the method and associated code, HEISENBERG, for Chapter 2^[p. 39]. However, to summarise, the KL14 principle can operate on arbitrary emission maps tracing different phases of the evolutionary process to constrain the duration of each phase. This is the intent of all these methods: to address one of the fundamental challenges in studying star formation, characterising the durations of the underlying physical processes and as a result constrain the likely active mechanisms driving the evolution.

The KL14 principle does have one key requirement: the need of a ‘reference timescale’. As was the case for Kawamura et al. (2009), who used a reference of 10 Myr, without a reference timescale the KL14 principle can only provide relative durations of evolutionary phases. Gas clouds do not possess an inherent property from which to measure their lifetime. In fact the measurements of cloud lifetimes that have been made vary by up to two orders of magnitude. Work by Scoville et al. (1979) and Koda et al. (2009), looking at interarm molecular clouds in the Milky Way and M51, concluded that GMCs must have lifetimes comparable to the interarm-crossing time ~ 100 Myr. Looking at star forming regions that still contained molecular gas, Elmegreen (2000) and Hartmann et al. (2001) concluded that the absence of post-T-Tauri stars suggests that molecular clouds cannot be long lived, at most ~ 10 Myr. Kruijssen et al. (2015) placed constraints on cloud lifetimes based on the orbital dynamics of molecular clouds in the Central Molecular Zone of the Milky Way, finding lifetimes of the order ~ 1 Myr. Jeffreson et al. (2018) took a theoretical approach, considering the effectiveness of various supportive and destructive

mechanisms of molecular clouds, showing that there exists a strong dependence on galactocentric radius but lifetimes are typically between 10–50 Myr. On the other hand, as hinted at in previous sections and as has been done before, the emission lifetime of SFR tracers can be used as a reference timescale. This is because of the predictable and consistent nature of stellar emission. Although, the emission lifetime of a SFR tracer is not always clearly defined. In [Chapter 3](#)^[p.47], we present our work on measuring the emission lifetime of H α emission as well as 12 UV filters (from GALEX, Swift, and HST). We also test the dependence of metallicity and the impact of low-mass star forming regions (resulting in a stochastically sampled IMF) on these emission lifetimes. In [Chapter 4](#)^[p.99], we place constraints on the amount by which extinction can act to alter these emission lifetimes. This would allow observational applications of the KL14 principle in instances where extinction cannot be corrected for. With these reference timescales in hand, applications of the KL14 principle can inform us about the next generation of star formation and feedback models, by constraining (amongst other quantities) molecular cloud lifetimes and the time it takes feedback to clear out the surrounding gas.

Another, complementary, approach to understanding star formation is through the use of simulations. It has been shown that different star formation and feedback recipes can drastically alter the appearance and gas density of simulated galaxies. [Hopkins et al. \(2013\)](#) explored the effects of different star formation criteria using isolated disc galaxy simulations. These criteria determine when gas within the simulation is permitted to form stars; this includes density and temperature thresholds as well as requiring the cloud to be locally self-gravitating, Jeans unstable, or molecular gas. The results show that comparing total SFR or star formation efficiency cannot distinguish these models; however, the spatial distribution of star formation within a galaxy can. The impact of different stellar feedback models has also been investigated. For example, [Haas et al. \(2013\)](#) compared galaxies in cosmological simulations from the OWLS project. Each simulation ran with a different feedback model. These models include changes in the type of feedback from supernovae (kinetic, thermal, or none), the wind velocity for kinetic feedback models, and the wind mass loading. Additional factors, such as the cooling of gas through metal-lines and AGN feedback were also considered. In [Figure 1.20](#)^[p.36] we present results from [Haas et al. \(2013\)](#) showing how these different feedback models impact the galaxy morphology and gas distribution.

Simulation projects that focus specifically on the interaction between stars and gas within galaxies have existed for some time. For example, the FIRE project has a suite of cosmological zoom-in simulations. These zoom-in simulations, allow the study of galaxy formation within the context of a

1 Introduction

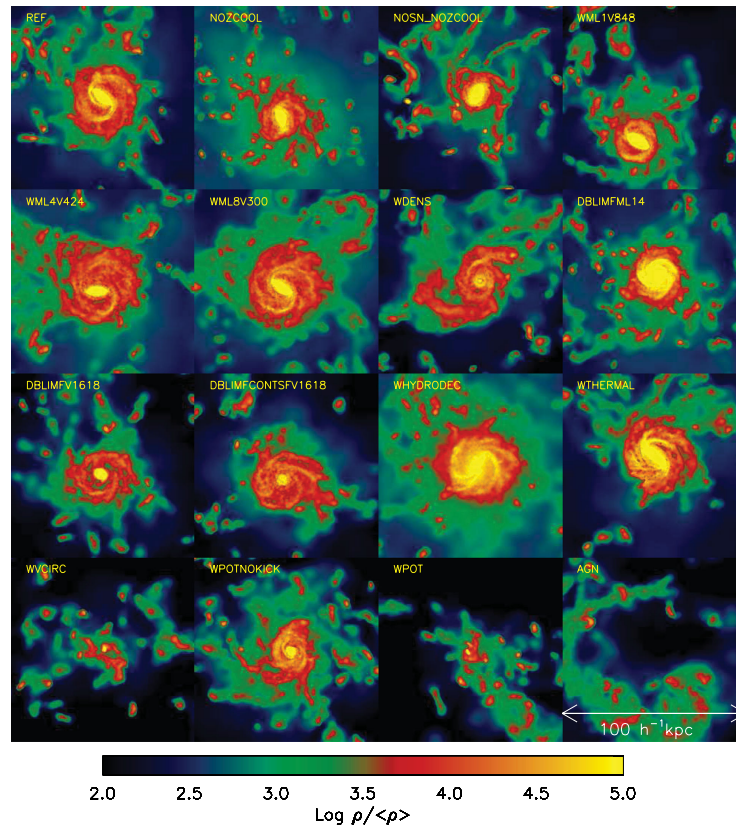


Figure 1.20

A sample of galaxies in a $10^{12.5} M_{\odot}$ halo from 16 different simulations from the OWLS project at redshift $z = 2$. Each simulation has a different implemented stellar feedback model (the name of the model is given in the top left of each panel). Details of the models can be found in Haas et al. (2013, Table 1), but include changes to the supernova feedback, the wind velocity of kinetic feedback models and wind mass loading.

Image Credit: Haas et al. (2013, Figure 2), reproduced with permission © OUP

cosmological environment but at a higher resolution than standard cosmological simulations. Although, only for a selection of galaxies, rather than a full galaxy population. The simulations can resolve the formation of GMCs and the multi-phase ISM. The energy, momentum, mass, and metals that stellar feedback returns to the ISM are specifically tracked. Results presented in Hopkins et al. (2012) from these simulations, with a range of different feedback models, indicate that ISM structure is more sensitive to the physics of stellar feedback than of the global disc. With such simulations, there is always a trade-off between simulation resolution and the size of the simulated volume. The SLICC project takes a different approach by looking at a vertical slice through the disc of a galaxy, rather than simulating the galaxy as a whole (i.e. choosing to reduce the simulated volume and so gain on the resolution). The aim of this project differs somewhat from FIRE, in that it studies the small-scale structure of the ISM and the life cycle of molecular clouds; however, it still investigates the effects of feedback. Results show that feedback mechanisms can compound on one another: early feedback mechanisms such as stellar winds and ionising radiation can disperse molecular clouds, leading to environments more susceptible to the effects of supernovae. On the opposite end of the scale is the new and unique IllustrisTNG50 simulation (Nelson et al. 2019; Pillepich et al. 2019). A full cosmological simulation of a 51.7^3 cMpc³ volume⁴ with a median spatial resolution of around ~ 100 – 140 pc (over cosmic time). The benefit of this is a representative cosmological volume with a resolution comparable to that of some zoom-in simulations. This means that detailed studies of galactic properties can be conducted on a statistically representative sample.

In Chapter 5^[p. 125], we use the IllustrisTNG50 simulation to produce synthetic observations of the gas within these galaxies. With the properties of the gas, including chemical abundance information, we model the expected CO line emission using the DESPOTIC code. These synthetic CO observations capture galaxies that fall within a $2.5' \times 2.5'$ field of view over the redshift range 1–4. With these observations we can study, for example, the redshift dependence of CO emission. Such synthetic observations can be used in comparisons with real observations as well as predictors of expected trends. In comparing to real observations, these synthetic observations can also function as tests of the simulations' ability to reproduce appropriate ISM densities and indirectly the effectiveness of the implemented star formation and feedback models.

⁴cMpc = comoving megaparsec

Uncertainty principle for star formation

This text in this chapter is based on sections from

D. T. Haydon et al. (Oct. 2018). ‘An uncertainty principle for star formation – III. The characteristic emission time-scales of star formation rate tracers’. *MNRAS Submitted*. arXiv: 1810.10897 [astro-ph.GA]

D. T. Haydon et al. (Aug. 2019). ‘An uncertainty principle for star formation – V. The influence of dust extinction on star formation rate tracer lifetimes’. *MNRAS Submitted*

The method described here has been developed by Kruijssen et al. (2014, 2018)

The work presented in [Chapters 3](#) and [4](#)^[pp.47 & 99] relies heavily on the KL14 principle and the code based on this method, HEISENBERG. It would therefore be informative to summarise this method here.

The goal of the KL14 principle is to constrain the cloud-scale physics of star formation and feedback using images of galaxies mapped in tracers that track different stages of the star formation process. By determining the duration of successive phases of an evolutionary process, it is possible to constrain the physical processes involved. The method was illustrated in Kruijssen et al. (2014, 2018) using the example of a molecular gas phase followed by a young stellar phase; we also adopt this example here. If we provide HEISENBERG with a galaxy mapped in an appropriate gas tracer (e.g. CO) and SFR tracer (e.g. H α), HEISENBERG can determine, among other quantities, the lifetime of molecular clouds within the observed galaxy. Measuring these timescales helps constrain the pertinent physical processes driving the cloud life cycle. For instance, a comparison of the molecular cloud lifetime to their free-fall time may indicate whether clouds are dynamic entities or exist in a quasi-equilibrium state. Furthermore, a comparison of the coexistence timescale of molecular gas and SFR tracer emission may indicate whether clouds are destroyed by supernova feedback or if early feedback mechanisms

2 Uncertainty principle for star formation

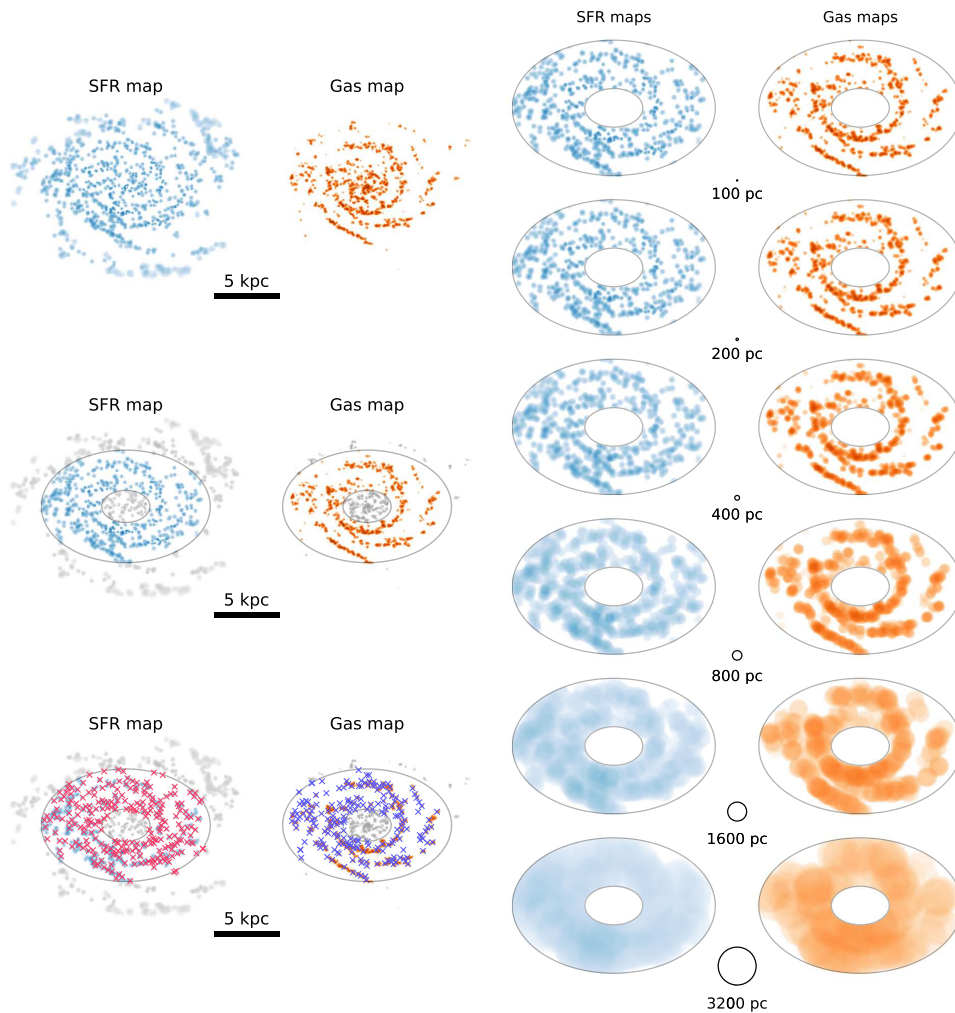
are needed. The power of this approach is obvious: the first applications of the method (Chevance and author 2019; Chevance, Kruijssen et al. 2019; Hygate et al. 2019; Kruijssen, Schrubba, Chevance et al. 2019) demonstrate that clouds are highly dynamic and are dispersed on timescales too short for supernovae, requiring early feedback by photoionisation. However, the applicability of the method is not restricted to maps of molecular gas or young stellar emission. Depending on the combination of tracers one uses, HEISENBERG can constrain the durations of different stages of the star formation timeline.

In basic terms, the KL14 principle represents the galaxy as a collection of independent regions, where each region is evolving along its timeline independently of the neighbouring regions. The number of regions that are emitting in each of the two tracers (with the possibility that the region is in a transition phase and so is emitting in both) is related to the duration of that phase: the shorter the duration of a phase, the less likely that a region would be observed in that phase. The HEISENBERG code quantifies statistically how this behaviour sets the flux ratio of the two phases as a function of position and spatial scale within a galaxy. The goal of HEISENBERG is to determine the duration of the evolutionary phases.

Most importantly in the context of our work, the method can only return *relative* durations of the evolutionary phases. Unless the duration of one of the two phases is known a priori, the relative timescales cannot be converted into an absolute evolutionary timeline. If, however, the duration of an evolutionary phase is known, this can be used as a ‘reference timescale’ to calibrate the timeline. Fortunately, this reference timescale can be associated with the SFR tracer owing to the absolute clock provided by stellar evolution (e.g. Leitherer et al. 1999, 2014). The aim of the work in Chapters 3 and 4^[pp. 47 & 99] is to measure the emission lifetime of SFR tracers under different environmental conditions; this then provides a reference timescale and a means to calibrate the evolutionary timeline. Throughout the work, we use the terms ‘reference map’ and ‘reference timescale’ to refer to the emission map associated to the phase of known duration and its duration, respectively.

We outline here the procedure used by HEISENBERG and refer the reader to Kruijssen et al. (2018) for the specific details. The method relies on the findings of Kruijssen et al. (2014), where it is shown that the flux ratio of the two phases changes relative to the galactic average when focusing apertures on gas or young stellar emission peaks. It also shows that this relative change is a direct function of the underlying evolutionary timeline describing how gas is converted into stars on the cloud scale. The procedure is as follows.

1. The user provides HEISENBERG with an input file and two emission maps of the same galaxy capturing two successive phases of an evolutionary



In each panel, we show the different steps used by the HEISENBERG code to analyse input emission maps. The emission maps shown here are created from the simulated galaxy which we use in [Chapter 3](#)^[p. 47]. Further details are given in the text. Top left: Two input emission maps provided by the user, which capture two stages of an evolutionary process. Middle left: A mask is applied to both emission maps to exclude the centre and outskirts of the galaxy. Bottom left: Emission peaks in each map are identified (red and blue crosses). Right: Emission maps are convolved on different scales.

Figure 2.1

Image Credit: Kruijssen et al. (2018, Figure 4–7), reproduced with permission © OUP

2 Uncertainty principle for star formation

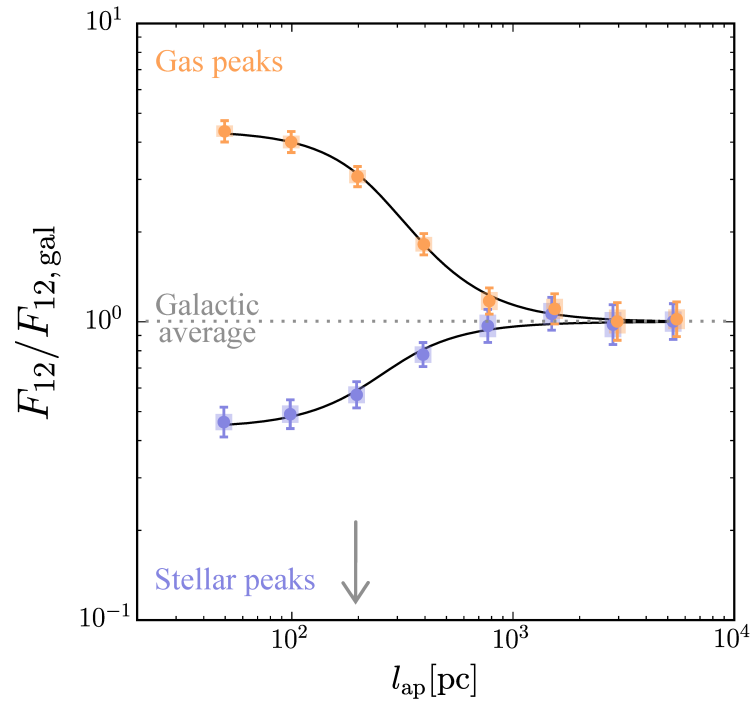


Figure 2.2

Example tuning fork diagram produced by the HEISENBERG code for the galaxy shown in Figure 2.1^[p. 41]. The data points show the relation between the Phase-1-to-Phase-2 flux ratio (F_{12}) relative to the galactic-scale Phase-1-to-Phase-2 flux ratio ($F_{12, \text{gal}}$) as a function of aperture size, l_{ap} . The two branches are a result of placing apertures at the location of Phase 1 (gas) emission peaks (top branch) or of Phase 2 (star) emission peaks (bottom branch). The model is indicated by the solid black curves. The error bars indicate the 1σ uncertainty on each individual data point, whereas the shaded areas indicate the effective 1σ uncertainty range that accounts for the covariance between the data points. The grey arrow marks λ , the typical separation length between identified peaks. As the aperture size increases, the calculated flux ratio tends to the galactic average (i.e. $F_{12}/F_{12, \text{gal}}$ tends to unity).

Image Credit: Kruijssen et al. (2018, Figure 12), adapted with permission © OUP

process. In keeping with the example above, this could be a CO and H α emission map to trace the gas and stellar phases, respectively. See the top left panel of [Figure 2.1](#)^[p.41]. We note here that HEISENBERG always uses the second phase emission map as the reference map: this is usually the SFR tracer map, for which a reference timescale can be defined, t_{star} .

2. Regions of the galaxy that the user wishes to exclude from the analysis, such as the centre or outskirts of the galaxy, are masked. See the middle left panel of [Figure 2.1](#)^[p.41].
3. The HEISENBERG code identifies the location of all the emission peaks present in both input maps within the range $\left[10^{(E_i^{\text{max}} - \Delta \log_{10} \mathcal{F}_i)}, 10^{E_i^{\text{max}}}\right]$, where $i = \{\text{gas}, \text{star}\}$, E_i^{max} is the decadic logarithm of the maximum emission in map i , and $\Delta \log_{10} \mathcal{F}_i$ defines the depth of the range in map i (a parameter specified by the user in the input file). This depth (approximately) corresponds to the dynamic range of the map, from the brightest emission peak to the noise floor. See the bottom left panel of [Figure 2.1](#)^[p.41].
4. Each map is convolved using a top hat kernel for a range of aperture sizes (as specified in the input file). See the right-hand panel of [Figure 2.1](#)^[p.41].
5. For each pair of convolved maps, HEISENBERG places apertures — with diameters equal to the convolution scale — centred on these identified emission peaks. The total flux across all apertures located at gas, and then stellar emission peaks is calculated for both emission maps. This gives four measurements, F_e^p : the total gas/stellar flux ($e = \{\text{gas}, \text{star}\}$) at the location of gas/stellar peaks ($p = \{\text{gas}, \text{star}\}$).
6. The Phase-1-to-Phase-2 flux ratios $F_{12}^p \equiv F_{\text{gas}}^p / F_{\text{star}}^p$, in units of the galactic average Phase-1-to-Phase-2 flux ratio ($F_{12, \text{gal}}$), are calculated as a function of aperture size. This produces a ‘tuning fork diagram’ (see [Figure 2.2](#)^[p.42]), where each branch corresponds to the flux ratio when focusing on the locations of emission peaks in the gas ($p = \text{gas}$) and stellar ($p = \text{star}$) maps.
7. The models which describe the shapes of the two branches (see Kruijssen et al. 2018, Equations 81 and 82) have three free parameters: the typical separation length between independent regions (λ , defining the aperture size at which the branches separate), the relative temporal overlap between the two phases ($t_{\text{over}}/t_{\text{star}}$, setting the convergence points at

2 Uncertainty principle for star formation

small aperture sizes), and the relative duration of the gas phase ($t_{\text{gas}}/t_{\text{star}}$, setting the vertical degree of asymmetry of the diagram).¹

8. To find the model parameters that best describe the data, HEISENBERG carries out a reduced- χ^2 fit.
9. The reference timescale specified in the input file (t_{star}) is used to recover t_{over} and t_{gas} as absolute timescales.

The procedure described above assumes that the reference timescale (i.e. t_{star}) is known; however, this is not true without calibrating the timescale. Without this reference timescale, HEISENBERG can only determine relative lifetimes (i.e. $t_{\text{over}}/t_{\text{star}}$ and $t_{\text{gas}}/t_{\text{star}}$).

In Chapter 3^[p.47], we use the above method to constrain the emission lifetime (t_{E}) for a variety of SFR tracers such that they can be used as reference timescales in observational applications of HEISENBERG. The use of HEISENBERG to measure the SFR tracer lifetime contrasts with normal observational applications (as described above), in which the SFR tracer lifetime is used as a *reference* timescale to calibrate the timeline. The fact that we now seek to *measure* the SFR tracer lifetime requires us to find a new independent way of defining a reference timescale. Using simulated galaxies, we can create reference maps from stars particles within a known age bin; the width of this age bin defines the reference timescale (t_{R}), which can be used to calibrate the emission lifetime of the synthetic SFR tracer maps. For Chapter 3^[p.47], we make notational transformations from the example used here to reflect the new role of the SFR tracer map: ‘gas’ \rightarrow ‘E’ (e.g. $t_{\text{gas}} \rightarrow t_{\text{E}}$) and ‘star’ \rightarrow ‘R’ (e.g. $t_{\text{star}} \rightarrow t_{\text{R}}$).

In Chapter 4^[p.99] we ascertain the affects of extinction on the emission lifetimes we measure in Chapter 3^[p.47]. We do this using the above method to measure R_{\bullet} : the factor by which the duration associated to an SFR tracer emission map without extinction (t_{\circ}) changes when including extinction (t_{\bullet}).

¹These quantities represent flux-weighted averages across the population of emission peaks (see Kruijssen et al. 2018; Kruijssen, Schrubba, Chevance et al. 2019, Section 3.2.9 and 3.2.11, Equation 1). In addition, we note that the region separation length, λ , marks the size scale on which a temporal de-correlation arises between the two tracers, reflecting that they represent different evolutionary phases of the same process. As a result, this length scale is a local quantity that is similar (but not equal) to the nearest neighbour distance and describes the typical separation length in the vicinity of peaks, not the area-averaged value across an entire galaxy. This is important because the degree of clustering may vary strongly across a galaxy (e.g. due to morphological features such as spiral arms). Since λ is a local quantity, it is relatively insensitive to such features (see discussions in Kruijssen et al. 2018; Chevance, Kruijssen et al. 2019; Kruijssen, Schrubba, Chevance et al. 2019).

This again requires the use of simulated galaxies; however, this is in order to produce emission maps both with and without extinction and not to create reference maps. For [Chapter 4](#)^[p. 99], we again make notational transformations from the example used here: ‘gas’ \rightarrow ‘●’ (e.g. $t_{\text{gas}} \rightarrow t_{\bullet}$) and ‘star’ \rightarrow ‘○’ (e.g. $t_{\text{star}} \rightarrow t_{\circ}$). After this transformation, it can be seen that one of the free quantities of the model, namely $t_{\text{gas}}/t_{\text{star}}$, is in fact R_{\bullet} .

The emission lifetime of star formation rate tracers

This chapter is based on

D. T. Haydon et al. (Oct. 2018). ‘An uncertainty principle for star formation – III. The characteristic emission time-scales of star formation rate tracers’. *MNRAS Submitted*. arXiv: 1810.10897 [astro-ph.GA]

In [Chapter 2](#)^[p. 39], we described the KL14 principle: a statistical method to constrain the physics of star formation and feedback on the cloud scale by reconstructing the underlying evolutionary timeline. However, this method is reliant on a ‘reference timescale’ in order to recover absolute durations of the different evolutionary phases. This reference timescale is necessary for observational applications of the method: without it, the method cannot be used to obtain meaningful constraints on, for example, the molecular cloud lifetime, the timescale for cloud destruction by feedback, the separation length between independent star-forming regions, the integrated cloud-scale star formation efficiency, the cloud-scale mass loading factor, and the feedback outflow velocity. Following the example of previous work in this area, the most direct way of providing a reference timescale is through the emission lifetime of SFR tracers. This is because stellar populations provide a natural clock dictated by the physics of stellar evolution. For example, only high mass stars are capable of radiating ionising photons of sufficient energy to produce H α emission. Since these stars have typical lifetimes < 10 Myr (Leitherer et al. 1999; Murphy et al. 2011), the H α emission itself should fade on a timescale that is of a similar magnitude. Different stars dominate the emission in different wavebands and so different wavebands will have emission lifetimes, which reflect the lifetime of the dominant stars (Hao et al. 2011; Kennicutt et al. 2012).

Previous work attempting to derive emission lifetimes for different SFR tracers has revealed that the major problem obstructing a conclusive measurement is that there is no obvious definition of the timescale which we should adopt. Instead, there exists a range of possible definitions, such as a

3 Emission lifetimes of SFR tracers

luminosity-weighted mean, a percentage intensity change, or a percentage of the cumulative emission. The choice of definition can result in differences of up to an order of magnitude in timescale (Kennicutt et al. 2012; Leroy et al. 2012). In view of this strong dependence on the precise definition of the reference timescale, we opt to use a self-consistent approach for determining the SFR tracer timescale; that is, we measure the emission timescales of SFR tracers by applying the KL14 principle itself. Since the amount of observable emission is likely to fade out gradually rather than halt abruptly, the ‘correct’ definition of the emission lifetime depends on the physical context for which the timescale is intended. By using the KL14 principle to define the timescale, we will constrain the emission lifetime in a way that is meaningful in the context of the KL14 principle.

Even though it is the goal of this work to define the emission lifetime of SFR tracers within the context of the KL14 principle, measuring this lifetime is also important in other contexts. For instance, this emission lifetime constrains the duration for which photoionising feedback can act on the surrounding ISM. Deriving absolute SFRs from observed line or broadband emission flux also requires the lifetime for which a young stellar population emits at that given wavelength.

In this chapter, we constrain emission lifetimes for H α emission (with and without continuum subtraction) as well as 12 UV filters (from GALEX, Swift, and HST), which cover a wavelength range 150–350 nm. We do so using HEISENBERG; however, this requires a controlled experiment, in which the absolute duration of one phase is known exactly in order to determine the absolute duration of the emission lifetimes. This therefore necessitates the use of galaxy simulations rather than real galaxies. Since the simulations do not have information about SFR tracer emission, we generate synthetic emission maps by combining the output from the hydrodynamical disc galaxy simulation with the SPS model SLUG2. This also allows us to test the impact of environmental factors, such as metallicity and the sampling of the stellar IMF, upon which the emission lifetimes will likely depend, and not just the different tracers.

The structure of this chapter is as follows. We outline our approach for constraining the emission lifetime of different SFR tracers with a well-sampled IMF in Section 3.1^[p. 49]. In Section 3.2^[p. 59] we present the results for solar metallicity environments (and well-sampled IMF). In Section 3.3^[p. 69], we demonstrate how the timescales depend on metallicity. In Section 3.4^[p. 79], we demonstrate the effects of incomplete IMF sampling, which is expected to change the results in environments of low SFR surface density. In Section 3.5^[p. 89], we carry out a brief test of the obtained timescales, by comparing

3.1 Measuring the emission lifetime of SFR tracers

them to observations of $H\alpha$ and FUV emission in NGC300. This is the nearby flocculent spiral galaxy upon which the first observational application of KL14 principle was performed and the timescales derived in this chapter used (for details see Kruijssen, Schrubba, Chevance et al. 2019). Finally, we summarise the results and present our conclusions in [Section 3.6](#)^[p.92].

3.1 Measuring the emission lifetime of SFR tracers

We present here the steps we take to find the emission lifetime for $H\alpha$ and UV SFR tracers (see [Table 3.1](#)^[p.50] for details) using synthetic emission maps and the HEISENBERG code. As we described in [Chapter 2](#)^[p.39], HEISENBERG can determine the duration of the first input map from the second by using the latter as a reference map (i.e. the map showing the evolutionary phase of known duration). This means that if we provide HEISENBERG with a galaxy map of one of the SFR tracers (e.g. $H\alpha$) along with a reference map, HEISENBERG can provide us with the emission lifetime associated with that SFR tracer. This approach to measure the SFR tracer lifetimes ensures that the obtained reference timescales are self-consistent within the context of our method. After all, in future observational applications of HEISENBERG, the SFR tracer map will be used as a reference map.

We generate both the SFR tracer maps and the reference maps using a numerical simulation of a flocculent spiral galaxy. Fundamentally, we only require some (preferably physically-motivated) correlation of positions and ages of star particles to carry out the experiments we perform in this chapter, implying that we could have used any (e.g. randomly-generated) distribution of points or Gaussian-like regions. However, the use of a galaxy simulation is more physically appropriate, as it contains some imprint of galactic morphology and the positional correlation of star formation events as a result of self-gravity and stellar feedback. Using a galaxy simulation still carries the advantage that we have complete control over the duration of the reference map — created by using stellar particles of a specified age range. The SFR tracer maps are generated using a SPS model. This approach allows us to additionally quantify the effects of metallicity (see [Section 3.3](#)^[p.69]) and IMF sampling (see [Section 3.4](#)^[p.79]) on the SFR tracer lifetimes. In turn, this will facilitate observational applications of HEISENBERG for a variety of galactic environments.

We discuss the adopted galaxy simulation in [Section 3.1.1](#)^[p.51], the method for generating the reference maps in [Section 3.1.2](#)^[p.53], and the method for generating the synthetic SFR tracer maps in [Section 3.1.3](#)^[p.56].

3 Emission lifetimes of SFR tracers

Table 3.1

We derive emission lifetimes for the SFR tracers listed here.

a: The UV filters we consider. $\bar{\lambda}_w$ is the response-weighted mean wavelength of the filter. The normalised filter response curves are presented in [Figure 3.6](#)^[p.66].

Telescope	Instrument	Filter	$\bar{\lambda}_w$ [nm]
GALEX		FUV	153.9
GALEX		NUV	231.6
Swift	UVOT	M2	225.6
Swift	UVOT	W1	261.7
Swift	UVOT	W2	208.4
HST	WFC3	UVIS1 F218W	223.3
HST	WFC3	UVIS1 F225W	238.0
HST	WFC3	UVIS1 F275W	271.5
HST	WFC3	UVIS1 F336W	335.8
HST	WFPC2	F255W	259.5
HST	WFPC2	F300W	297.4
HST	WFPC2	F336W	335.0

b: The H α filters we consider.

Filter	Details
H α -	H α emission with continuum subtraction. This is not a true filter but a direct measurement of the Hydrogen-ionizing photon emission, see Section 3.1.3 ^[p.56] for details.
H α + W	A narrow band filter including H α and the continuum as defined in Equation 3.3 ^[p.58] . The total filter width is indicated by W ; we consider $W = \{10, 20, 40, 80, 160\}$ Å.

3.1.1 Galaxy simulation

The results in this chapter are based on the ‘high-resolution’ simulated galaxy from Kruijssen et al. (2018). We set up the initial conditions for this galaxy using the methods described in Springel et al. (2005). The simulation has a total of 4.95×10^6 particles: 1×10^6 in the dark matter halo, 2.31×10^6 in the stellar disc, 1.54×10^6 in the gas disc, and 1×10^5 in the bulge. The dark matter halo particles have a mass of $9 \times 10^5 M_\odot$ and both the star and gas particle types have a mass of $2.7 \times 10^3 M_\odot$. This gives us a $9 \times 10^{11} M_\odot$ halo, $1.05 \times 10^{10} M_\odot$ disc (60 per cent in stars and 40 per cent in gas), and $2.7 \times 10^8 M_\odot$ bulge.

We then evolve the initial conditions for 2.2 Gyr using the SPS code P-GADGET-3 (last described by Springel 2005), which makes use of the SPHGAL hydrodynamics solver. SPHGAL was implemented by Hu et al. (2014) in order to overcome many of the numerical issues associated with traditional smoothed-particle hydrodynamics (SPH). To be considered for star formation, gas particles require temperatures less than 1.2×10^4 K and Hydrogen particle densities more than 0.5 cm^{-3} . Stars are formed from eligible gas particles stochastically according to the method described in Katz (1992). Supernova explosions return mass, momentum, and thermal energy back to the ISM; these are distributed using a kernel weighting to the 10 nearest gas particles. The result of the simulation is a near- L_\star isolated, flocculent spiral galaxy, forming stars at a rate of roughly $0.3 M_\odot \text{ yr}^{-1}$. Figure 3.1^[p.52] places the macroscopic properties of the galaxy in the context of the observed nearby galaxy population. This shows that the simulated galaxy resides on the main sequence and has a normal total gas depletion time. In addition, Figure 3.2^[p.54] shows a stellar reference map (see Section 3.1.2^[p.53]) and a synthetic $\text{H}\alpha$ -map (see Section 3.1.3^[p.56]) of this galaxy, demonstrating that its morphology is similar to that of nearby flocculent spirals like M33 and NGC300.

The star formation and feedback prescriptions used in the simulation are certainly inadequate to describe the cloud-scale physics governing the evolutionary cycling between molecular gas, star formation, and feedback within galaxies (see e.g. Hopkins et al. 2018; Kruijssen, Schrubba, Chevance et al. 2019). However, this is not a concern in the context of the problem at hand. The goal of this work is not to accurately model cloud-scale star formation and feedback. Instead, we aim to determine how quickly SFR tracer emission fades after the formation of a young stellar population, and to do so self-consistently in the context of the KL14 principle. This can be achieved with any simulation in which the birth sites of star particles approximately conform to a galaxy-like morphology. As shown by the images in Figure 3.2^[p.54], this is indeed achieved by the simulation used here. This

3 Emission lifetimes of SFR tracers

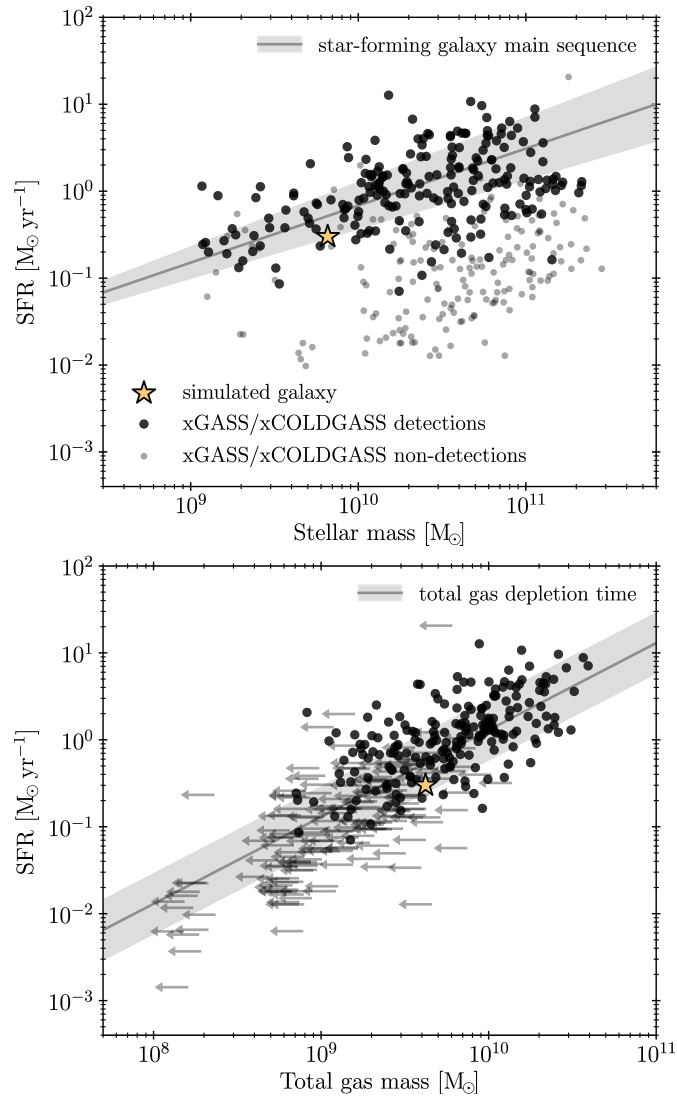


Figure 3.1

Properties of the simulated galaxy in the context of the nearby galaxy population. The panels show the SFR as a function of stellar mass (top) and total gas mass (bottom), both for the simulated galaxy and nearby galaxies from the xCOLDGASS and xGASS surveys. In the bottom panel, the arrows indicate upper limits for galaxies that are non-detections in either or both of xCOLDGASS and xGASS. For the non-detections of molecular and atomic gas, we use 3σ and 5σ upper limits, respectively. The solid lines represent the star-forming galaxy main sequence (top panel), the mean total gas depletion time of the detections (bottom panel), with the 1σ scatter shown in grey.

3.1 Measuring the emission lifetime of SFR tracers

enables the necessary SFR tracer lifetimes to be derived and therefore the reference timescales. These in turn, will enable observational applications of our method that themselves will motivate a future generation of star formation and feedback models that are capable of describing cloud-scale evolutionary cycling in galaxy simulations (for further discussion see Kruijssen et al. 2018; Fujimoto et al. 2019).

3.1.2 Generation of reference maps

The role the reference map plays in the HEISENBERG code is to calibrate the absolute evolutionary timeline of the star formation process. In the context of this chapter, it is used to calibrate the emission lifetime of SFR tracer emission.

In our experiments aimed at measuring the SFR tracer lifetimes, we need to know the reference timescale exactly. For this reason, we use simulated rather than real galaxies. We produce reference maps from the simulation by generating mass surface density maps of the star particles in a specific age range. The width of this age range acts as the reference timescale, t_R . We smooth the selected star particles using a Wendland C^4 kernel (Dehnen et al. 2012) (the same kernel SPHGAL introduces into P-GADGET-3) over the 200 nearest neighbouring particles; this produces a realistic reference map (i.e. not a map of point particles).

In principle, we have a free choice over the age range we use. However, for the best results and the most realistic set-up there are a few restrictions. In Chapter 2^[p. 39], we note that HEISENBERG is designed such that the reference map corresponds to the second phase of the evolutionary timeline. To avoid any overlap between the evolutionary phases, the minimum age of the star particles used in the reference map (t_M) must therefore be at least the duration of the first (SFR tracer emission) phase ($t_{E,0}$, we include the subscript ‘0’ to indicate that this is for a well sampled IMF: this distinction is necessary in Section 3.4^[p. 79]) of the evolutionary timeline. This defines the lower limit of the stellar age range used to generate the reference map:

$$t_M \gtrsim t_{E,0} .$$

Equation 3.1

At the same time, it is undesirable to select a value of t_M much larger than the galactic dynamical time because groups of star particles formed in the same clouds may have dispersed. We therefore prefer using $t_M \approx t_{E,0}$.

It was show by Kruijssen et al. (2018) that HEISENBERG provides the most accurate measurement of the underlying timescales if the duration associated to both of the input maps is of a similar order of magnitude (i.e. within a factor of 10, but ideally within a factor of 4). When the durations are similar, the tuning

3 Emission lifetimes of SFR tracers

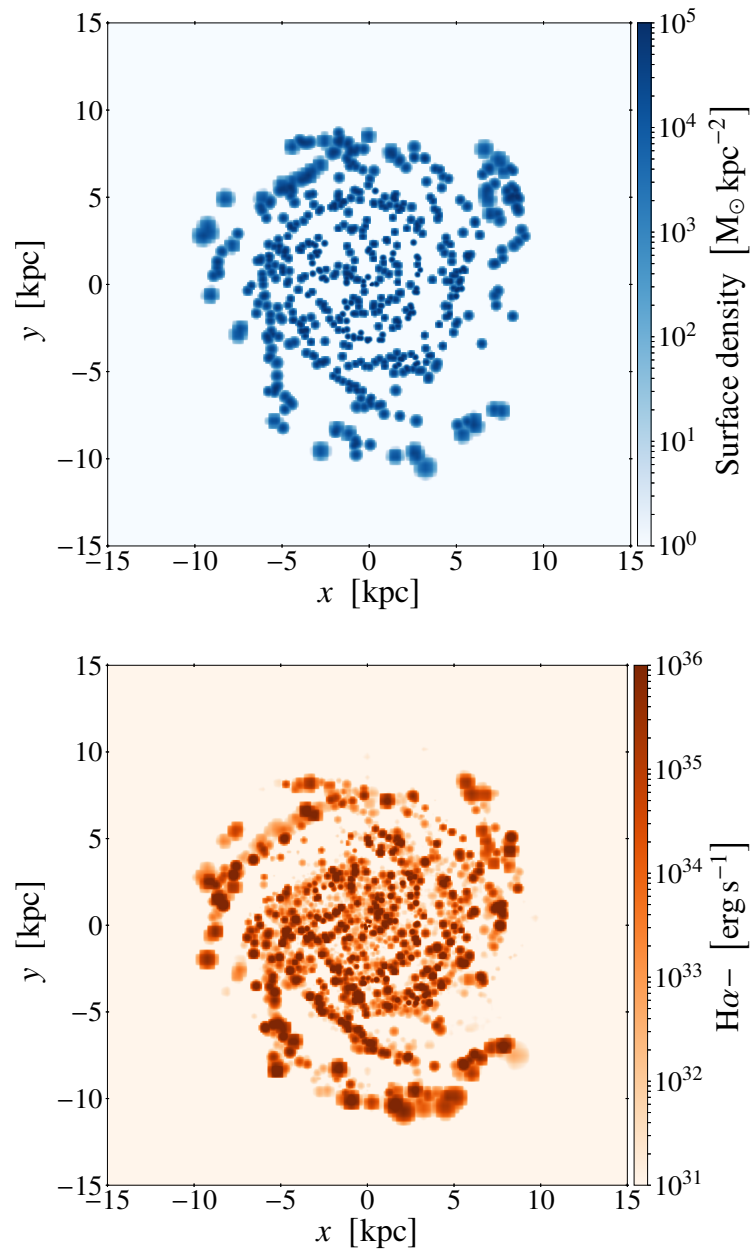
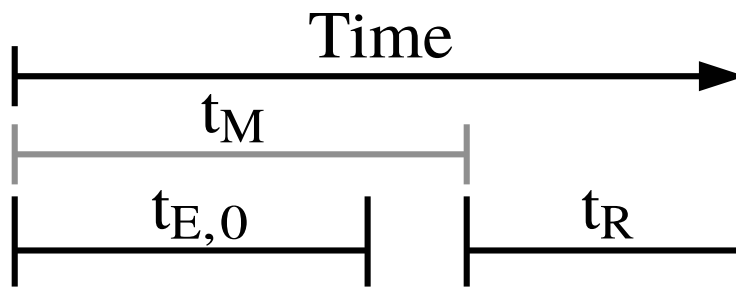


Figure 3.2

Example maps we use as input for the HEISENBERG code. Top: A reference map generated using the mass surface density of star particles in the age range 10–15 Myr, implying a reference timescale of 5 Myr in this example. See [Section 3.1.2](#)^[p.53] for details. Bottom: A synthetic $H\alpha-$ emission map without the continuum ($H\alpha-$) generated by performing SPS on the simulated galaxy. See [Section 3.1.3](#)^[p.56] for details.

3.1 Measuring the emission lifetime of SFR tracers



Schematic diagram showing how the different timescales we define within this chapter are related. Time starts at the birth of the star particle. The emission map shows the particles formed within a timescale $t_{E,0}$ prior to the simulation snapshot, where $t_{E,0}$ represents the emission lifetime of the SFR tracer. The time over which the reference map runs is defined by t_M and t_R , where t_M sets the minimum age of the star particles used to create the reference map and t_R defines the width of the age range and therefore the reference timescale. The structure of the HEISENBERG code is such that, the duration of second evolutionary phase is used to calibrate the duration of the first. This means that to calibrate the SFR tracer lifetime ($t_{E,0}$), the SFR tracer map must be the first evolutionary phase (i.e. $t_{E,0} \leq t_M$); this is unlike observational applications, where it is usually the second.

Figure 3.3

3 Emission lifetimes of SFR tracers

fork diagram (Figure 2.2^[p. 42]) becomes more symmetrical and the method can retrieve the underlying timescales with a precision of about 30 per cent. This finding sets the preferred width of the age range:

Equation 3.2

$$t_R \approx t_{E,0} .$$

In Figure 3.3^[p. 55], we show a schematic timeline of how t_M , t_R , and $t_{E,0}$ are related.

To quantify (and avoid) any systematic biases of the measured SFR tracer lifetimes, we investigate the dependence on the choice of stellar age range used to generate the reference map. In practice, this means we vary the values of t_M and t_R . We present the range of values we use for t_M and t_R in Table 3.2^[p. 57]. These are guided by the range of possible emission lifetimes for H α and FUV emission found in Leroy et al. (2012). Using the results of STARBURST99 calculations, Leroy et al. (2012) determine an emission lifetime using several methods: a luminosity-weighted average time, as well as the times at which the tracer emission reaches a particular limit in terms of the total cumulative emission or its instantaneous intensity.

3.1.3 Generation of synthetic emission maps

In order to perform our analysis, we need to produce synthetic emission maps of each SFR tracer. The simulation that we base this work on (see Section 3.1.1^[p. 51]) contains no information about the expected emission spectrum. We therefore use SLUG2, a stochastic SPS code, to take the age and mass of the star particles and predict the associated emission for the filters specified in Table 3.1^[p. 50].

With the SLUG2 model, we predict the expected rest-frame emission spectrum for every star particle within the simulation.¹ The code first samples an IMF to construct a simple stellar population of total mass matching that of the star particle and then uses stellar evolution tracks along with the age of the star particle to determine the combined emission of this simple stellar population. The SLUG2 code then converts the full combined emission spectrum into a single luminosity value for each of the SFR tracers in Table 3.1^[p. 50] using filter response curves. These single luminosity values are what we assign to our star particles when we produce our synthetic emission maps. We use the same smoothing procedure as we described in Section 3.1.2^[p. 53]. This means that, even though our star particles are treated as simple stellar populations, the star-forming regions themselves, which are a collection of multiple particles,

¹We note that the age selection process described in Section 3.1.2^[p. 53] is not used for the emission maps.

3.1 Measuring the emission lifetime of SFR tracers

We create input reference maps from the star particles that fall within a particular age range. The age range, for a given reference map, is defined through $t_M \leq \text{Age} \leq t_M + t_R$. We show here all the values used in this chapter for t_M and t_R when defining these age ranges. This results in a 9×9 array of reference maps (for example [Figure 3.4](#)^[p. 62]). See [Section 3.1.2](#)^[p. 53] for more details.

Table 3.2

Emission Type	t_M and t_R [Myr]								
H α	1	3	5	7	10	15	20	25	30
UV	5	10	15	20	25	30	50	70	100

3 Emission lifetimes of SFR tracers

will have an age spread. An example of a synthetic $H\alpha$ - map is shown in [Figure 3.2](#)^[p. 54].

The adopted UV response filters are all included by default in SLUG2 (for more details see [Krumholz et al. 2015](#)). The $H\alpha$ SFR tracers, however, require different steps. For $H\alpha$ - we use the Hydrogen-ionizing photon emission, $Q(H^0)$, directly² and for $H\alpha+W$ we define the narrow band filter, $\mathcal{F}_{H\alpha+W}$, as

Equation 3.3

$$\mathcal{F}_{H\alpha+W} = \begin{cases} 1 & 6562 - \frac{W}{2} \text{ \AA} \leq \lambda \leq 6562 + \frac{W}{2} \text{ \AA} \\ 0 & \text{Otherwise} \end{cases},$$

where λ here is the wavelength of the emission. The emission spectrum produced by SLUG2, includes the $H\alpha$ emission line but does not calculate the underlying absorption feature from the stellar continuum. In [Section 3.A](#)^[p. 94] we use STARBURST99 simulations to investigate when the absorption can no longer be neglected. We find that, for the timescales we are considering, the absorption is negligible.

For the analysis in [Section 3.2](#)^[p. 59], we use a [Chabrier \(2005\)](#) IMF with Geneva solar-metallicity evolutionary tracks ([Schaller et al. 1992](#)) and STARBURST99 spectral synthesis. The SLUG2 model samples the IMF non-stochastically³ (i.e. we use a well sampled IMF) and no foreground extinction is applied. The surrounding material has a Hydrogen number density of 10^2 cm^{-3} . We assume that only 73 per cent of the ionising photons are reprocessed into nebular emission, which is consistent with the estimate from [McKee et al. \(1997\)](#); this could be because those photons are absorbed by circumstellar dust, or because they escape outside the observational aperture (the observational effects of these two possibilities are indistinguishable).

We choose to produce our synthetic emission maps without extinction for a number of reasons. In observational applications of the KL14 principle, there is often some overlap between the first and second phases of the evolutionary timeline. For instance, when applying the method to a molecular gas map (e.g. CO) and an ionised emission map (e.g. $H\alpha$), there will be some non-zero time for which both tracers coexist. When a region resides in this ‘overlap’ phase, the star-forming region may be partially embedded in dust and gas; during this phase the region suffers the most from extinction. We can therefore define the duration of this second phase, t_2 , as

Equation 3.4

$$t_2 = t_0 + t_1,$$

²A true $H\alpha$ luminosity can be calculated from $Q(H^0)$ using ([Silva et al. 2014](#), Equation 2); however, using the required scaling factor will not change the results we recover here (see [Kruijssen et al. 2018](#)) and so the conversion is unnecessary.

³In [Section 3.4](#)^[p. 79], we will use the stochastic IMF sampling mode of SLUG2 to investigate its effect on the inferred SFR tracer lifetimes.

3.2 A well sampled IMF at solar metallicity

where t_o is the duration of the second phase that overlaps with the first, and t_i the duration that is independent. The emission lifetimes we define in this chapter are for this independent part, t_i , of the second phase. This is where the region is no longer embedded in dust and gas and therefore not suffering from significant extinction. We motivate this by the notion that molecular gas correlates with star formation: as long as CO emission is present, star formation is likely to be ongoing. The ‘clock’ defined by the SFR tracer lifetime only starts when the last massive stars have formed. This does mean that the application of HEISENBERG to tracers other than CO may require a different definition of the reference timescale. To facilitate this, the HEISENBERG code enables the user to specify if the reference timescale includes or excludes this overlap phase (see Kruijssen et al. 2018, Section 3.2.1).

In addition, it is desirable to exclude extinction for two further reasons. Firstly, the effects of extinction can, in most cases, be significantly reduced if not completely corrected for (e.g. James et al. 2005), meaning in practice the input maps provided to HEISENBERG can be corrected for extinction. Secondly, if we perform our analysis with extinguished maps, the results would no longer be generally applicable and would only apply to galaxies that suffer from the same amount of extinction. Our current approach therefore enables constructing a ‘universal’ baseline of extinction-corrected SFR tracer lifetimes. In Chapter 4^[p.99], we consider the effects of extinction using galaxy simulations covering a range of gas surface densities.

3.2 Measuring the emission lifetime of SFR tracers for a fully sampled IMF at solar metallicity

We constrain the emission lifetime for several SFR tracers using the HEISENBERG code. As input to the code, we use synthetic SFR tracer maps and reference maps as described in Section 3.1^[p.49]. The reference maps are characterised by the ages of the star particles included within them. We investigate how the chosen age range affects the emission lifetime measured by changing the values of t_M and t_R which define the age range: $t_M \leq \text{Age} \leq t_M + t_R$. This approach does not (initially) result in a single emission lifetime for each SFR tracer, but an array spanned by t_M and t_R .

The reason we choose to measure the emission lifetime for the an array of t_M and t_R rather than just a single pair is for accuracy. As already noted in Section 3.1.2^[p.53], the accuracy of the measurement HEISENBERG makes depends on temporal range and offset of the reference map. The best

3 Emission lifetimes of SFR tracers

measurements are made when the duration of the evolutionary phases (i.e. $t_{E,0}$ and t_R) are similar. We also prefer the temporal offset between the reference and SFR tracer maps to not be too large, so as to mimic close evolutionary phases (such as between molecular gas and massive star formation). Since we do not know $t_{E,0}$ in advance (this is the quantity which we aim to constrain with this work), it is not possible to choose an optimal pair of t_M and t_R . We therefore resort to systematically measuring the emission lifetime using reference maps generated for all combinations of t_M and t_R to produce an array of emission lifetimes. However, we do find that the typical standard deviation of all calculated SFR tracer lifetimes is of the order 0.15 dex (in logarithmic space), despite the fact that both t_M and t_R vary by 1.5 dex. This demonstrates that the inferred lifetimes are not extremely sensitive to the choice of reference map, but using the full array of reference maps does allow us to optimise the accuracy of the SFR tracer lifetimes.

In what follows, we describe how we reduce these ‘lifetime arrays’ (see [Figure 3.4](#)^[p. 62] for examples) into a single emission lifetime for each SFR tracer. However, we first note that HEISENBERG not only outputs the measured lifetime but also its probability density function (PDF). This means that each element of the lifetime array has an associated PDF, which we make use of in the process of defining a single emission lifetime.

To define a emission lifetime for each SFR tracer from the array of lifetimes, we produce 10^6 realisations of the lifetime array, where the value of each element of each realisation of the lifetime array has been randomly sampled from its associated PDF. For each of the 10^6 realisations of the lifetime array we calculate the weighted mean of the array. The weighting scheme we use is one that favours solutions where the uncertainties are small and t_M , t_R , and $t_{E,0}$ are the most similar. The weighting for an element in the array with lifetime t_{ij} is given by \mathcal{W}_{ij} :

$$\text{Equation 3.5} \quad \mathcal{W}_{ij} = \frac{\mathcal{W}_{ij}^d \mathcal{W}_{ij}^u}{\sum_{ij} \mathcal{W}_{ij}^d \mathcal{W}_{ij}^u},$$

where the summation is a normalisation factor. The \mathcal{W}^d weighting favours more strongly elements that satisfy the criteria we describe in [Equations 3.1](#) and [3.2](#)^[pp. 53 & 56] (i.e. the closer t_{ij} is to $t_{M,i}$ and $t_{R,j}$ the better):

$$\text{Equation 3.6} \quad \mathcal{W}_{ij}^d = \left\{ \left[\log_{10} \left(\frac{t_{ij}}{t_{M,i}} \right) \right]^2 + \left[\log_{10} \left(\frac{t_{ij}}{t_{R,j}} \right) \right]^2 \right\}^{-\frac{1}{2}},$$

which is simply the geometric distance in logarithmic space. The \mathcal{W}^u weighting takes the lower and upper uncertainty (σ_{ij}^- and σ_{ij}^+ respectively) on the lifetime

3.2 A well sampled IMF at solar metallicity

t_{ij} (as calculated by HEISENBERG) into consideration:⁴

$$\mathcal{W}_{ij}^u = \left(\frac{\sigma_{ij}^- + \sigma_{ij}^+}{2} \right)^{-2} .$$

Equation 3.7

This process results in 10^6 emission lifetimes, from which we can produce a PDF. We define the emission lifetime from the median of the PDF and the uncertainties from the 16th and 84th percentiles (see Figure 3.5^[p. 63] for examples).

With the above procedure, we condense the array of lifetimes into a single number that most strongly reflects the most accurate measurements (based on Kruijssen et al. 2018) and those that have the smallest measurement uncertainties. The uncertainties on this single emission lifetime are calculated using the PDF of the complete set of 10^6 values and so they contain both the uncertainties on the individual measurements and the systematic uncertainty associated with choosing a combination of t_M and t_R .

When applying HEISENBERG to the pairs of reference and SFR tracer maps, we use the default input parameters specified in Kruijssen et al. (2018, Tables 1 and 2). The only exceptions are as follows. We set `tstar_incl` = 1, to indicate that the reference timescale (i.e. the width of the age range) also includes the overlapping phase.⁵ As we are not making any cuts in galactocentric radius, we also set `cut_radius` = 0. Finally, we define the range of aperture sizes using a minimum aperture size of $l_{\text{ap, min}} = 25$ pc and a number of $N_{\text{ap}} = 17$ apertures, to produce 17 logarithmically-spaced aperture diameters from 25–6400 pc.

In Figure 3.4^[p. 62], we present, as examples, two lifetime arrays obtained for H α - and WFC3 UVIS F225W SFR tracers. These lifetime arrays only serve as examples, since the elements show the best-fitting values returned by HEISENBERG and are not from the 10^6 realisations. We also show, in Figure 3.5^[p. 63], the PDFs associated to 10^6 realisations, which we use to defined the emission lifetimes for H α - and WFC3 UVIS F225W.

⁴Using the average of the lower and upper uncertainty is not technically correct; however, the methods as suggested by Barlow (2003) to account for asymmetric uncertainties would have little impact on the final result and so are neglected.

⁵This is not in contradiction with what we discuss in Section 3.1.3^[p. 56]. In Section 3.1.3^[p. 56], we explain that the emission lifetimes of the SFR tracers we define do not include the overlap phase; and so, when using the emission lifetimes we present here, one should use `tstar_incl` = 0. The analysis we perform to define the emission lifetimes, uses a reference map produced from star particles in a specific age range. The width of this age range is used as the reference timescale and is the total duration of that phase: this includes any overlap.

3 Emission lifetimes of SFR tracers

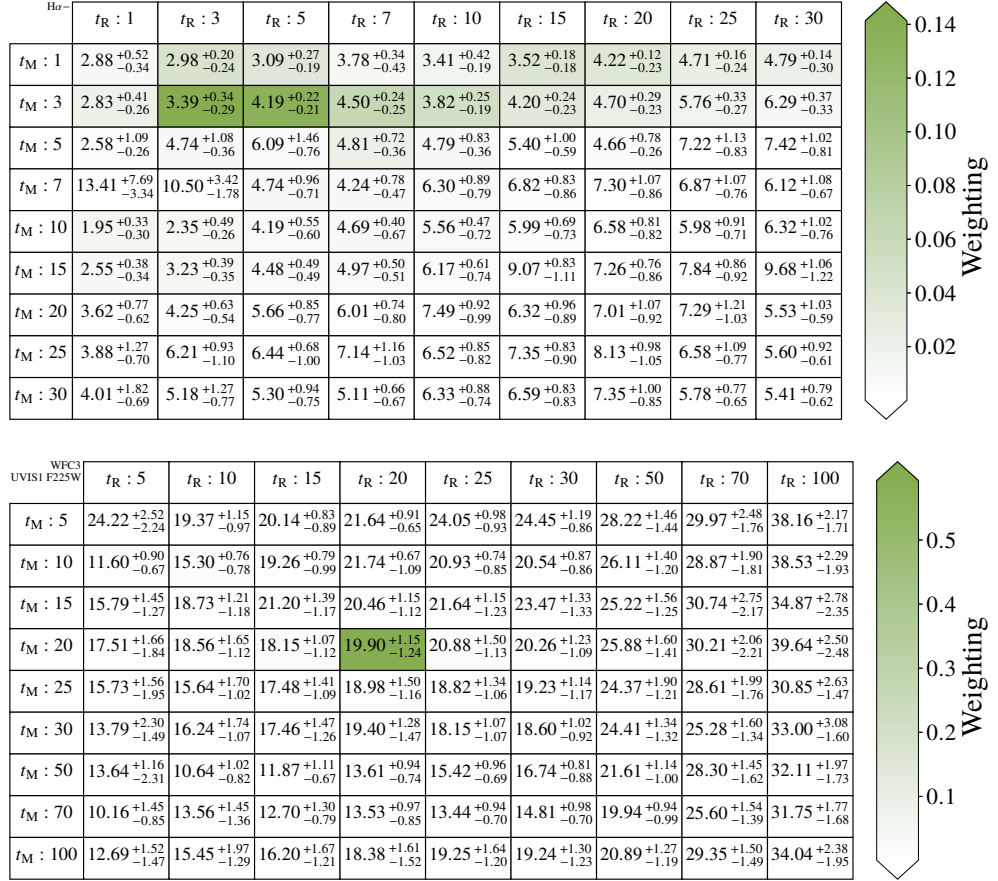
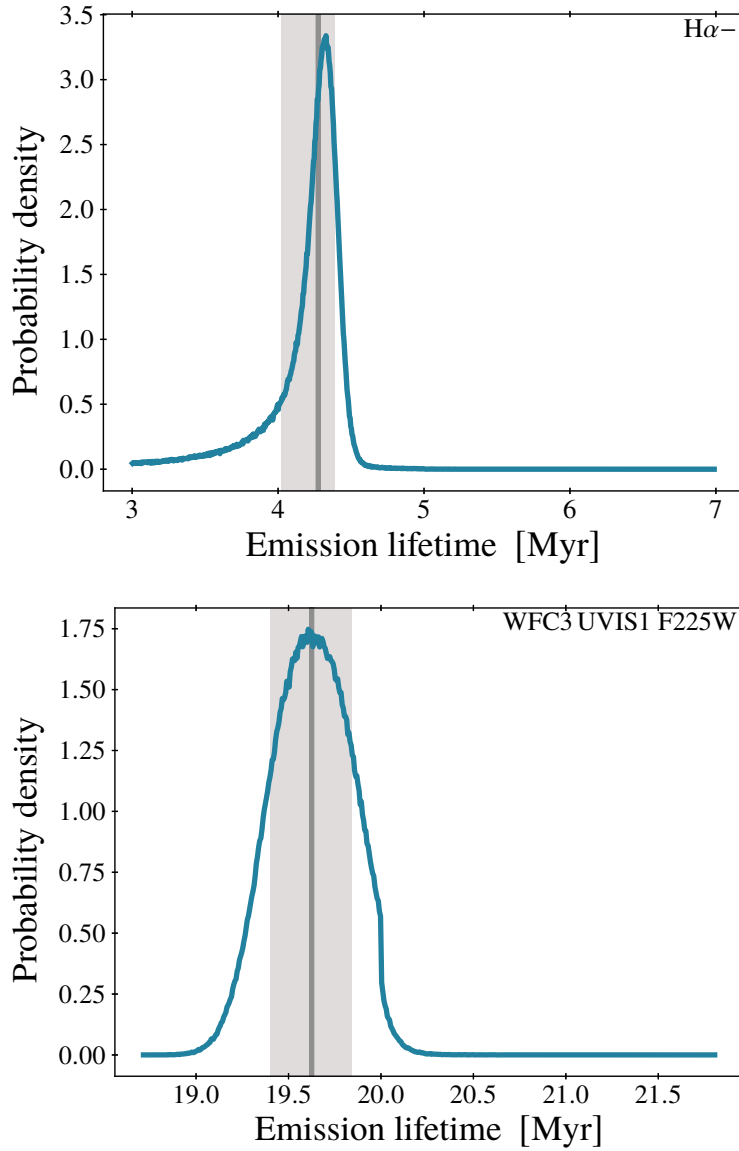


Figure 3.4

Two examples showing the range of emission lifetimes (and associated uncertainties) as determined using HEISENBERG for different reference maps. Top: H α emission excluding the continuum (H α -). Bottom: UV emission (WFC3 UVIS1 F225W). The reference maps are characterised by the age range used to select the star particles which are included in the reference map. t_M denotes the minimum age of the star particles and t_R the width of the age range. The colour-coding is based on the weighting, \mathcal{W} , used when calculating the weighted average (see Equation 3.5^[p.60]). All values within the tables are given in Myr.

3.2 A well sampled IMF at solar metallicity



Two examples showing the PDFs associated to the defined emission lifetimes. Top: H α emission excluding the continuum (H α -). Bottom: UV emission (WFC3 UVIS F225W). The vertical line shows the selected lifetime (the median of the distribution); the shaded region, the uncertainty defined by the 16th and 84th percentiles.

Figure 3.5

3 Emission lifetimes of SFR tracers

We can use the defined emission lifetimes, $t_{E,0}$, to specify the age range we use for creating our reference maps in later sections of this chapter:

Equation 3.8
$$t_{E,0} \leq \text{Age} \leq 2t_{E,0} ,$$

which complies with the conditions specified in [Equations 3.1 and 3.2](#)^[pp. 53 & 56].

In [Table 3.3](#)^[p. 65], we present the emission lifetimes and associated uncertainties we constrain for each of the different SFR tracer filters. We also include the age ranges that we will use for the reference maps in later sections of this chapter; we calculated these using the expression given in [Equation 3.8](#).

In the upper panel of [Figure 3.6](#)^[p. 66], we show the response curves for the UV filters and also mark the response-weighted mean wavelength, $\bar{\lambda}_w$. Visualising the response curves helps to explain the range of measured emission lifetimes. In the lower panel of [Figure 3.6](#)^[p. 66], we show $\bar{\lambda}_w$ against the emission lifetime of the young stellar emission in each filter. The figure shows that similar response-weighted mean wavelengths give similar emission lifetimes. In addition, we see that there is a tight relation between the SFR tracer lifetime and wavelength. We describe this relationship between $\bar{\lambda}_w$ and the UV emission lifetime, $t_{E,0}^{\text{UV}}$, by

Equation 3.9
$$t_{E,0}^{\text{UV}} \text{ [Myr]} = \left(3.00_{-0.31}^{+0.29} \right) \left(\frac{\bar{\lambda}_w}{225 \text{ nm}} \right)^{\left(4.34_{-0.20}^{+0.24} \right)} + \left(16.42_{-0.31}^{+0.36} \right) .$$

This relation is obtained by performing a weighted least-squares minimisation. The uncertainties on the parameter values are calculated using a Monte Carlo approach. With this relationship it is possible to find the emission lifetimes for UV filters that we have not explicitly considered here.

Similarly, we derive a relation between the H α + emission lifetime and filter width, W , to make it possible to find emission lifetimes for intermediate W :

Equation 3.10
$$t_{E,0}^{\text{H}\alpha+} \text{ [Myr]} = \left(4.8_{-1.3}^{+1.3} \right) \left(\frac{W}{40 \text{ \AA}} \right)^{\left(0.65_{-0.13}^{+0.20} \right)} + \left(3.8_{-1.1}^{+1.1} \right) .$$

We present this relation in [Figure 3.7](#)^[p. 67]. We note that the increase in emission lifetime with filter width is not due to a change in the H α emission but the result of including more of the long-lived continuum emission.

The UV emission lifetimes that we recover (17.1–33.3 Myr) are within the ranges often quoted in the literature Kennicutt et al. (10–100 Myr, [2012](#)) and Leroy et al. ([2012](#)). The large variation in literature values is due to the fact that there is no single method or distinct set of criteria which one should use in order to constrain the emission lifetime of an SFR tracer. With the approach

3.2 A well sampled IMF at solar metallicity

The emission lifetimes, $t_{E,0}$, obtained for the different SFR tracers (see Table 3.1^[p.50] for details) and the corresponding age ranges ($t_{E,0} \leq \text{Age} \leq 2t_{E,0}$) for producing reference maps in later sections of this chapter. These results are for a well sampled IMF at solar metallicity. The filter order is in increasing filter width (W) for H α + and increasing response-weighted mean wavelength ($\bar{\lambda}_w$) for UV. This table is an extract of Table 3.5^[p.74], which includes the emission lifetimes and age ranges for different stellar metallicities ($Z/Z_\odot = 0.05-2$).

Table 3.3

	$t_{E,0}$ [Myr]	Age range [Myr]
H α -	$4.3^{+0.1}_{-0.3}$	4.3–8.6
H α + 10 Å	$5.6^{+0.2}_{-0.1}$	5.6–11.1
H α + 20 Å	$7.3^{+0.4}_{-0.2}$	7.3–14.6
H α + 40 Å	$9.3^{+0.2}_{-0.3}$	9.3–18.6
H α + 80 Å	$10.7^{+0.2}_{-0.2}$	10.7–21.4
H α + 160 Å	$16.4^{+0.6}_{-0.3}$	16.4–32.7
GALEX FUV	$17.1^{+0.4}_{-0.2}$	17.1–34.2
UVOT W2	$19.0^{+0.3}_{-0.2}$	19.0–38.0
WFC3 UVIS1 F218W	$19.4^{+0.2}_{-0.2}$	19.4–38.9
UVOT M2	$19.5^{+0.2}_{-0.2}$	19.5–39.0
GALEX NUV	$19.6^{+0.2}_{-0.2}$	19.6–39.1
WFC3 UVIS1 F225W	$19.6^{+0.2}_{-0.2}$	19.6–39.3
WFPC2 F255W	$22.4^{+0.2}_{-0.2}$	22.4–44.7
UVOT W1	$21.8^{+0.2}_{-0.2}$	21.8–43.5
WFC3 UVIS1 F275W	$23.5^{+0.2}_{-0.2}$	23.5–47.0
WFPC2 F300W	$27.7^{+0.6}_{-0.3}$	27.7–55.4
WFPC2 F336W	$33.1^{+0.4}_{-0.3}$	33.1–66.3
WFC3 UVIS1 F336W	$33.3^{+0.4}_{-0.4}$	33.3–66.6

Percentages of the emission intensity relative to its instantaneous value at 1 Myr and of its cumulative value over 100 Myr, evaluated at the emission lifetimes of the SFR tracers presented in Table 3.3, based on Figure 3.8^[p.68].

Table 3.4

	H α - [%]	FUV ^a [%]
% of intensity at 1 Myr	$19.4^{+2.8}_{-0.7}$	$8.6^{+0.1}_{-0.2}$
% of cumulative emission	$92.4^{+0.6}_{-1.5}$	$76.5^{+0.3}_{-0.2}$

^a GALEX FUV

3 Emission lifetimes of SFR tracers

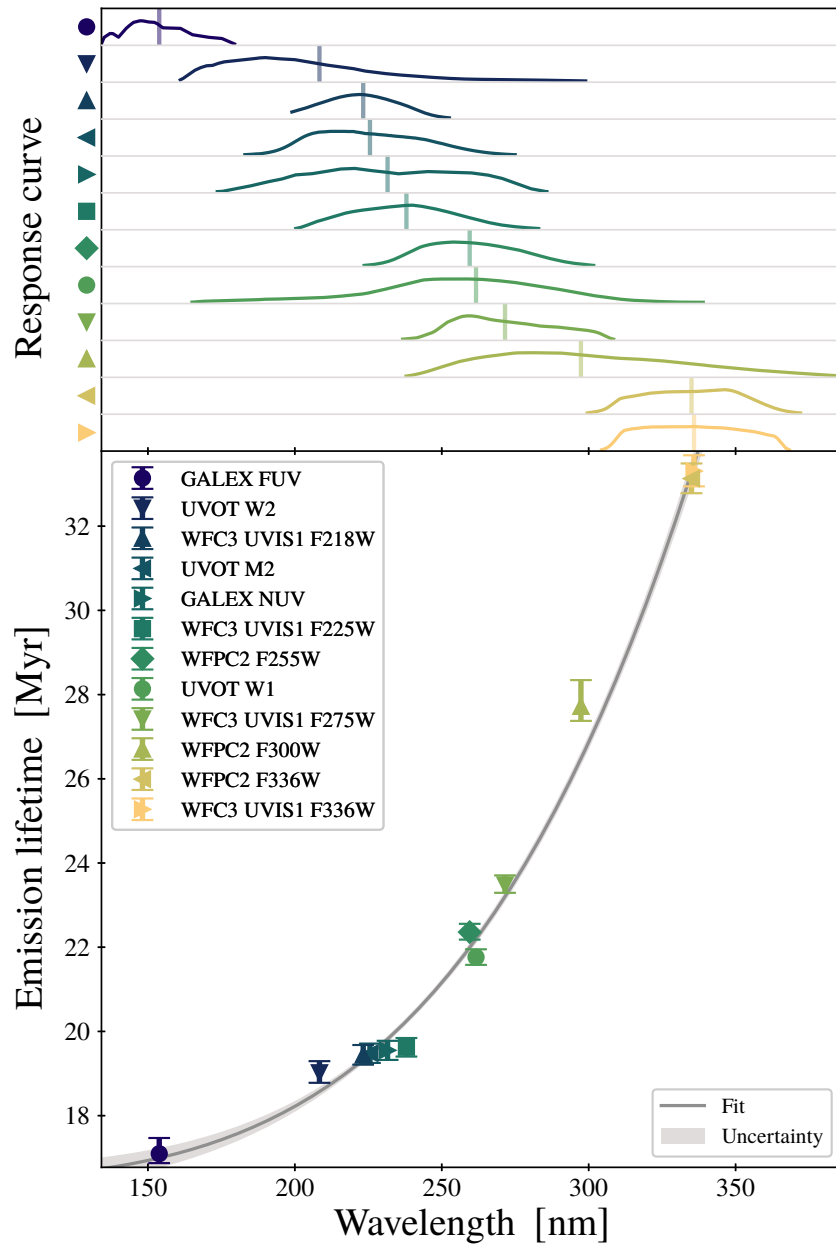
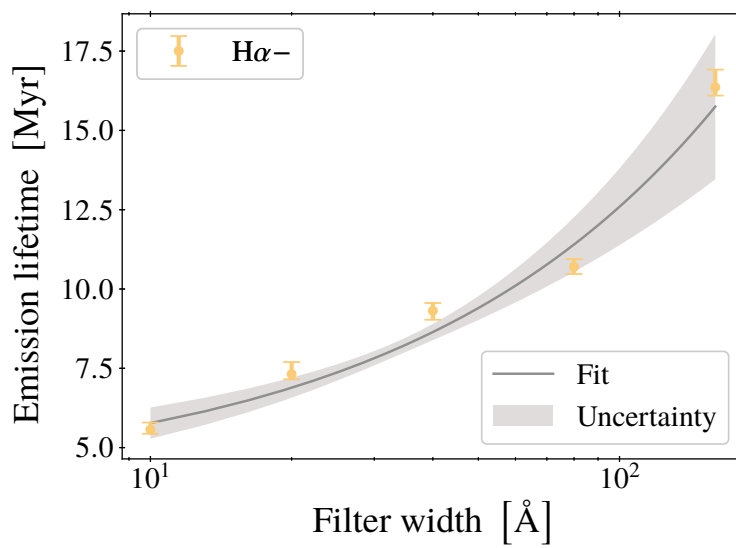


Figure 3.6 Top: The normalised response curves of the UV filters considered in this chapter (also see Table 3.1a^[p.50]). The vertical lines indicate the response-weighted mean wavelengths, $\bar{\lambda}_w$. Bottom: emission lifetimes for UV filters as a function of response-weighted mean wavelength. The grey curve shows the fit described in Equation 3.9^[p.64] and the shaded regions indicates the associated uncertainty.

3.2 A well sampled IMF at solar metallicity



Emission lifetimes for H α + filters as a function of the filter width. The grey curve shows the fit described in Equation 3.10^[p. 64] and the shaded region indicates the associated uncertainty.

Figure 3.7

3 Emission lifetimes of SFR tracers

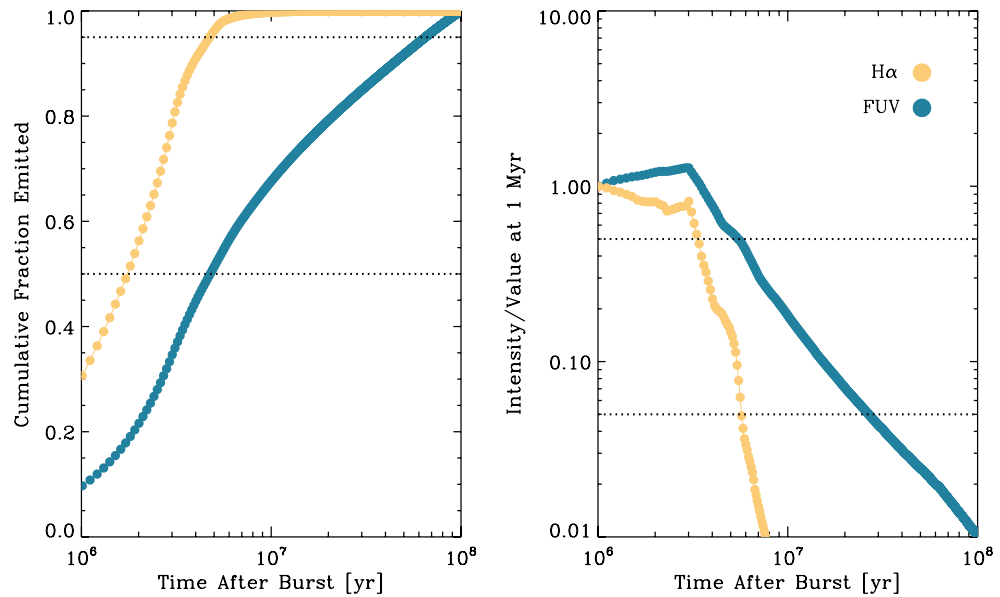


Figure 3.8

The two panels show how the cumulative emission (left) and instantaneous intensity (right) of H α and FUV emission changes with time, as calculated using `STARBURST99` for an instantaneous burst of star formation. Horizontal, dashed lines mark possible (and arbitrarily selected) percentage limits which could be used to define an emission lifetime. It is clear, especially for FUV emission, that the choice in limit could change the lifetime by up to an order of magnitude. Using these results and the lifetimes in [Table 3.3](#)^[p. 65], we reverse the process (i.e. find percentage limits from timescales) to determine if a percentage limit can be defined. The results are presented in [Table 3.4](#)^[p. 65].

Image Credit: Leroy et al. (2012, Figure 1), reproduced with permission © AAS

3.3 The effects of metallicity

taken in this chapter, we have remedied this problem for future observational applications of the KL14 principle.

A table of emission lifetimes is presented in Leroy et al. (2012) for H α and FUV (at 150 nm). Multiple lifetimes are listed for each SFR tracer; these lifetimes are defined by the duration required to reach a given percentage limit of the cumulative emission or of the emission intensity relative to the intensity at 1 Myr. We take Leroy et al. (2012, Figure 1) (reproduced here in Figure 3.8^[p. 68]) and reverse the process to find the percentage that corresponds to the emission lifetimes we determine for H α - and GALEX FUV. We list these percentage limits in Table 3.4^[p. 65].

The percentage limits used in Leroy et al. (2012) are arbitrarily chosen and so we make this conversion into percentage limits to determine if a single percentage of the 1 Myr intensity or cumulative emission can be defined that would correspond to SFR tracer lifetimes that we measure. From the results in Table 3.4^[p. 65], we can see that no single percentage limit can be defined.⁶ As there is no consistent limit, the emission lifetime for each SFR tracer must be determined individually.

In summary, we see that the emission lifetimes fall in the range of commonly reported literature values. By comparing these measurements to the time-evolution of the SFR tracer intensities, we find that no fixed percentage of the initial or cumulative emission is capable of matching the obtained emission lifetimes. For this reason, each SFR tracer lifetime must be determined individually using the presented method. However, we do find empirical functions (see Equations 3.9 and 3.10^[p. 64]) relating the emission lifetimes for UV and H α + filters to their filter properties. These relations predict the emission lifetimes for UV and H α + filters we have not considered here.

3.3 The effects of metallicity

So far, we have only considered stellar populations of solar metallicity; however, it is well-known that the metallicity affects stellar lifetimes (e.g. Leitherer et al. 1999) and thus the emission lifetimes of SFR tracers. In order to facilitate observational applications of the KL14 principle for the broadest possible range of galaxies, we therefore quantify how the SFR tracer lifetimes depend on metallicity. In this section, we repeat the experiments performed in Section 3.2^[p. 59] but this time we produce synthetic SFR tracer emission maps using evolutionary tracks of metallicities $Z/Z_{\odot} = 0.05, 0.20, 0.40, 2.00$

⁶This also holds when we perform the analysis for the other metallicities considered: $Z/Z_{\odot} = 0.05, 0.20, 0.40, 2.00$ (see Section 3.3 for more details). There is also no consistent percentage for a single tracer across the metallicity range.

3 Emission lifetimes of SFR tracers

(Schaller et al. 1992; Charbonnel et al. 1993; Schaerer, Meynet et al. 1993; Schaerer, Charbonnel et al. 1993).

In Table 3.5^[p. 74], we list the emission lifetimes for a well sampled IMF for all metallicities (also including the solar metallicity results from Table 3.3^[p. 65]) and the age ranges we select for producing reference maps. We see that as the metallicity increases, the emission lifetime decreases; we show this $Z - t_{E,0}$ relation in Figure 3.9^[p. 71] for $H\alpha^-$, Figure 3.10^[p. 72] for $H\alpha^+$ filters, and in Figure 3.11^[p. 73] for the UV filters. We also include empirical fits described respectively by

$$\text{Equation 3.11} \quad t_{E,0}^{H\alpha^-} \text{ [Myr]} = \left(4.32_{-0.23}^{+0.09}\right) \left(\frac{Z}{Z_{\odot}}\right)^{\left(-0.086_{-0.023}^{+0.010}\right)},$$

$$\begin{aligned} \text{Equation 3.12} \quad t_{E,0}^{H\alpha^+} \text{ [Myr]} &= \left(8.98_{-0.50}^{+0.40}\right) W_0^{\left(0.265_{-0.051}^{+0.028}\right)} \\ &+ \left(0.23_{-0.11}^{+0.15}\right) Z_0 W_0 \\ &- \left(0.66_{-0.19}^{+0.12}\right) Z_0 + \left(0.55_{-0.29}^{+0.46}\right) W_0, \end{aligned}$$

and

$$\begin{aligned} \text{Equation 3.13} \quad t_{E,0}^{UV} \text{ [Myr]} &= -\left(0.40_{-0.16}^{+0.11}\right) Z_1 \lambda_1 \\ &+ \left(4.5_{-0.9}^{+1.3}\right) Z_1 + \left(0.70_{-0.18}^{+0.26}\right) \lambda_1 \\ &- \left(3.11_{-0.13}^{+0.14}\right) Z_0 + \left(10.98_{-0.48}^{+0.46}\right) \lambda_0 \\ &+ \left(7.6_{-1.6}^{+1.2}\right), \end{aligned}$$

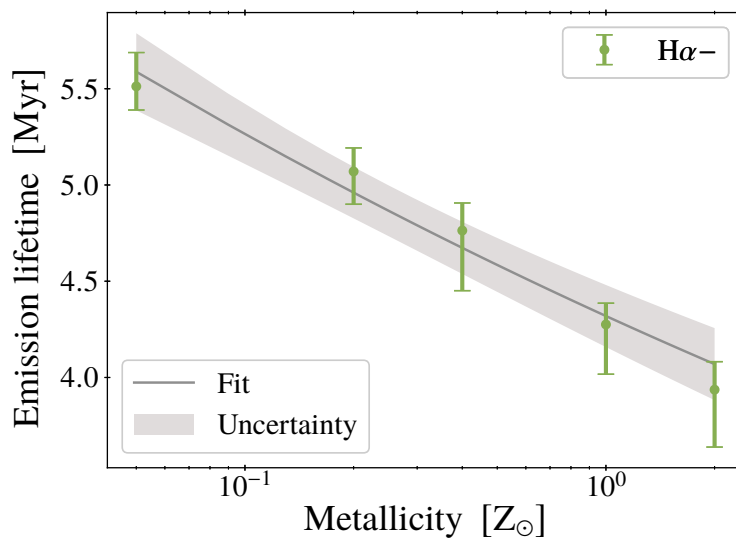
where

$$\begin{aligned} \text{Equation 3.14} \quad Z_0 &\equiv \frac{Z}{Z_{\odot}}; & \lambda_0 &\equiv \frac{\bar{\lambda}_w}{225 \text{ nm}}; & W_0 &\equiv \frac{W}{40 \text{ \AA}}; \\ Z_1 &\equiv Z_0^{\left(-0.313_{-0.048}^{+0.051}\right)}; & \lambda_1 &\equiv \lambda_0^{\left(6.52_{-0.71}^{+0.73}\right)}. \end{aligned}$$

As before, we determine the free parameters using a weighted least-squares minimisation and the uncertainties through Monte Carlo methods. With these relations, it is straight forward to recover the emission lifetime given the metallicity and filter properties without having to repeat the analysis performed here.

Figure 3.9^[p. 71] shows that the emission lifetime of $H\alpha^-$ changes by less than 2 Myr over the metallicity range $[0.05 Z_{\odot}, 2 Z_{\odot}]$. The ranges of emission

3.3 The effects of metallicity



The relation between metallicity and emission lifetime for a well-sampled IMF for H α - filters. The grey curve gives the fits described by Equation 3.11^[p. 70]. The shaded region indicates the associated uncertainty.

Figure 3.9

3 Emission lifetimes of SFR tracers

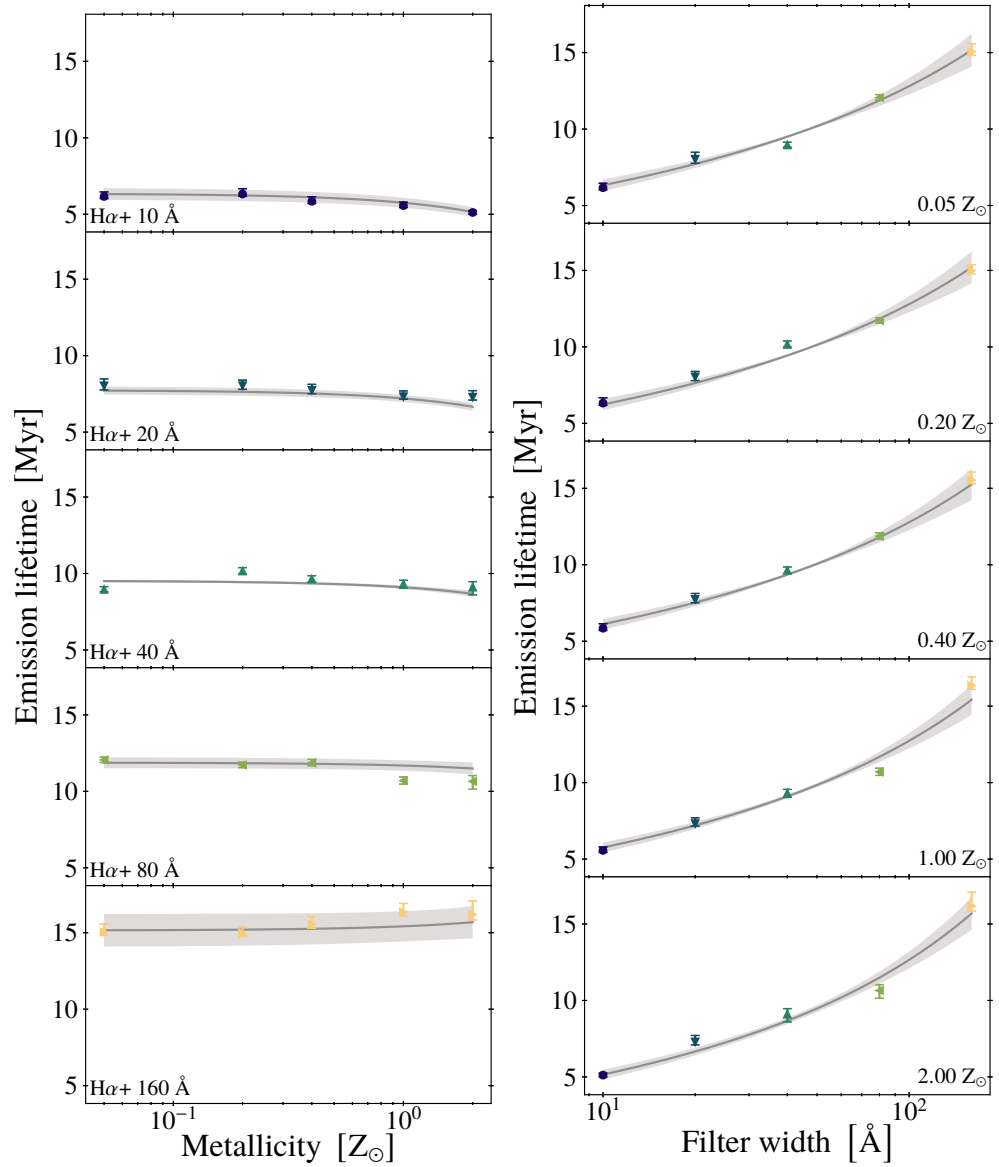
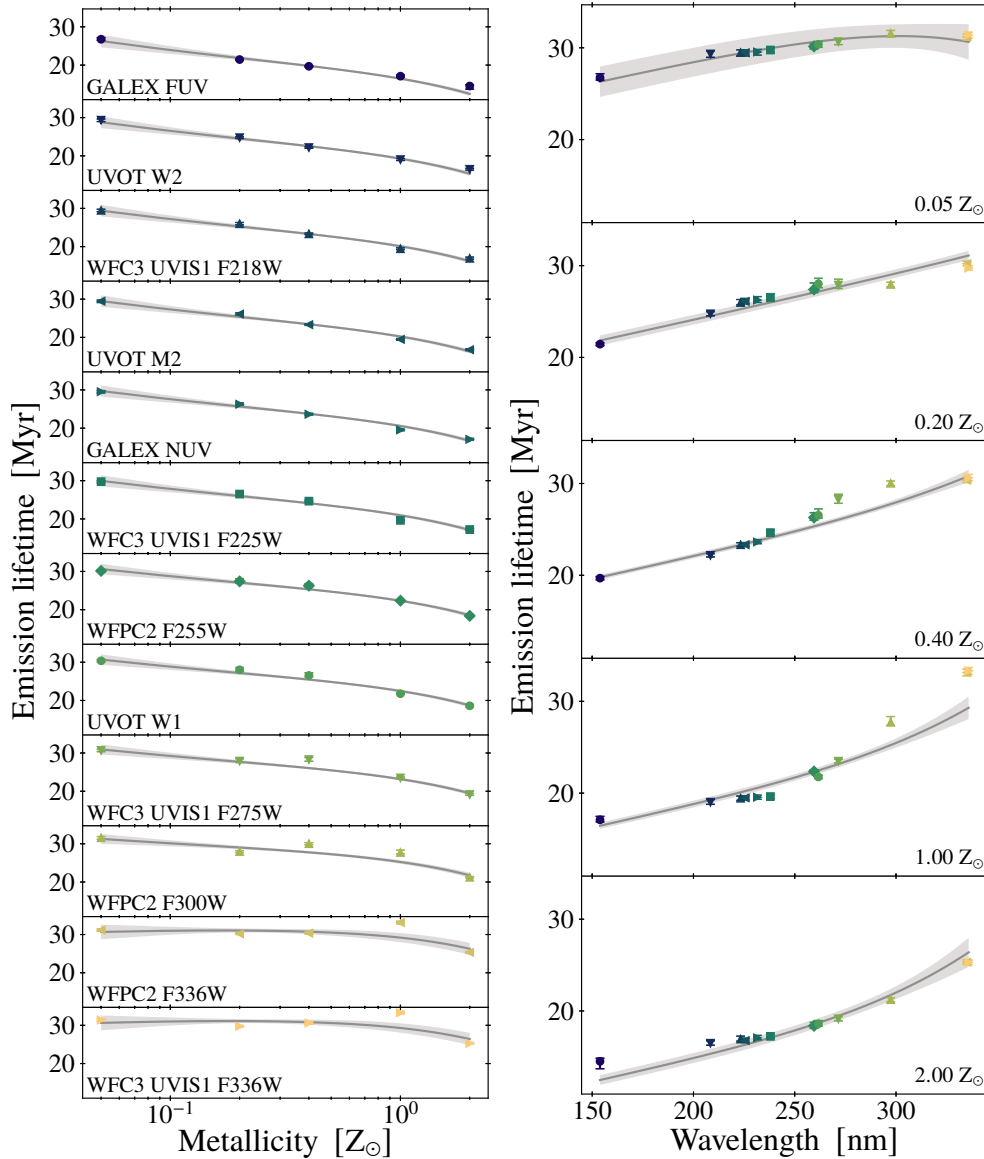


Figure 3.10

Emission lifetimes of $H\alpha+$ filters as a function of metallicity (left) and filter width (right). The symbols show the results of applying the HEISENBERG code to synthetic $H\alpha+$ maps. The grey curve shows the fit from Equation 3.12^[p.70] and the shaded region indicates the associated uncertainty. The symbols and colours in the right-hand panels correspond to those used on the left.

3.3 The effects of metallicity



Emission lifetimes of UV SFR tracers as a function of metallicity (left) and wavelength (right). The symbols show the results of applying the HEISENBERG code to synthetic UV maps. The grey curve shows the fit from Equation 3.13^[p. 70] and the shaded region indicates the associated uncertainty. The symbols and colours in the right-hand panels correspond to those used on the left.

Figure 3.11

3 Emission lifetimes of SFR tracers

Table 3.5

A summary of all the emission lifetimes and corresponding age ranges (for producing reference maps in later sections of this chapter), for the different SFR tracers (see Table 3.1^[p.50] for details). These results are for a well sampled IMF. The filter order is in increasing filter width (W) for $H\alpha+$ and increasing response-weighted mean wavelength ($\bar{\lambda}_w$) for UV. Table 3.3^[p.65] is included in these tables under the $1.00 Z_\odot$ heading.

a: Emission lifetimes, $t_{E,0}$.

	$0.05 Z_\odot$	$0.20 Z_\odot$	$0.40 Z_\odot$	$1.00 Z_\odot$	$2.00 Z_\odot$
$H\alpha-$	$5.5^{+0.2}_{-0.1}$	$5.1^{+0.1}_{-0.2}$	$4.8^{+0.1}_{-0.3}$	$4.3^{+0.1}_{-0.3}$	$3.9^{+0.1}_{-0.3}$
$H\alpha+ 10 \text{ \AA}$	$6.2^{+0.3}_{-0.1}$	$6.3^{+0.3}_{-0.1}$	$5.9^{+0.3}_{-0.1}$	$5.6^{+0.2}_{-0.1}$	$5.1^{+0.1}_{-0.1}$
$H\alpha+ 20 \text{ \AA}$	$8.0^{+0.5}_{-0.3}$	$8.0^{+0.4}_{-0.2}$	$7.7^{+0.4}_{-0.2}$	$7.3^{+0.4}_{-0.2}$	$7.3^{+0.4}_{-0.2}$
$H\alpha+ 40 \text{ \AA}$	$9.0^{+0.2}_{-0.2}$	$10.2^{+0.2}_{-0.2}$	$9.6^{+0.2}_{-0.2}$	$9.3^{+0.2}_{-0.3}$	$9.1^{+0.4}_{-0.5}$
$H\alpha+ 80 \text{ \AA}$	$12.1^{+0.2}_{-0.2}$	$11.7^{+0.2}_{-0.2}$	$11.9^{+0.2}_{-0.2}$	$10.7^{+0.2}_{-0.2}$	$10.7^{+0.4}_{-0.5}$
$H\alpha+ 160 \text{ \AA}$	$15.1^{+0.5}_{-0.2}$	$15.0^{+0.4}_{-0.2}$	$15.5^{+0.5}_{-0.2}$	$16.4^{+0.6}_{-0.3}$	$16.2^{+0.9}_{-0.3}$
GALEX FUV	$26.7^{+0.4}_{-0.3}$	$21.4^{+0.2}_{-0.2}$	$19.7^{+0.2}_{-0.2}$	$17.1^{+0.4}_{-0.2}$	$14.5^{+0.3}_{-0.8}$
UVOT W2	$29.3^{+0.3}_{-0.4}$	$24.8^{+0.2}_{-0.2}$	$22.2^{+0.2}_{-0.2}$	$19.0^{+0.3}_{-0.2}$	$16.5^{+0.3}_{-0.2}$
WFC3 UVIS1 F218W	$29.5^{+0.3}_{-0.4}$	$26.0^{+0.3}_{-0.3}$	$23.3^{+0.2}_{-0.2}$	$19.4^{+0.2}_{-0.2}$	$16.9^{+0.3}_{-0.2}$
UVOT M2	$29.5^{+0.3}_{-0.4}$	$26.1^{+0.3}_{-0.3}$	$23.3^{+0.2}_{-0.2}$	$19.5^{+0.2}_{-0.2}$	$16.8^{+0.3}_{-0.2}$
GALEX NUV	$29.5^{+0.3}_{-0.3}$	$26.3^{+0.4}_{-0.3}$	$23.6^{+0.3}_{-0.2}$	$19.6^{+0.2}_{-0.2}$	$17.1^{+0.3}_{-0.2}$
WFC3 UVIS1 F225W	$29.8^{+0.2}_{-0.3}$	$26.5^{+0.4}_{-0.3}$	$24.7^{+0.3}_{-0.2}$	$19.6^{+0.2}_{-0.2}$	$17.2^{+0.3}_{-0.2}$
WFPC2 F255W	$30.1^{+0.3}_{-0.3}$	$27.4^{+0.7}_{-0.3}$	$26.3^{+0.5}_{-0.3}$	$22.4^{+0.2}_{-0.2}$	$18.4^{+0.4}_{-0.3}$
UVOT W1	$30.4^{+0.3}_{-0.3}$	$28.0^{+0.6}_{-0.4}$	$26.5^{+0.7}_{-0.3}$	$21.8^{+0.2}_{-0.2}$	$18.6^{+0.4}_{-0.3}$
WFC3 UVIS1 F275W	$30.7^{+0.3}_{-0.3}$	$27.9^{+0.6}_{-0.4}$	$28.3^{+0.6}_{-0.4}$	$23.5^{+0.2}_{-0.2}$	$19.1^{+0.2}_{-0.2}$
WFPC2 F300W	$31.6^{+0.3}_{-0.4}$	$28.0^{+0.2}_{-0.3}$	$30.0^{+0.2}_{-0.3}$	$27.7^{+0.6}_{-0.3}$	$21.2^{+0.2}_{-0.2}$
WFPC2 F336W	$31.2^{+0.3}_{-0.3}$	$30.2^{+0.3}_{-0.2}$	$30.3^{+0.3}_{-0.2}$	$33.1^{+0.4}_{-0.3}$	$25.3^{+0.3}_{-0.2}$
WFC3 UVIS1 F336W	$31.4^{+0.3}_{-0.3}$	$29.7^{+0.2}_{-0.2}$	$30.6^{+0.4}_{-0.3}$	$33.3^{+0.4}_{-0.4}$	$25.3^{+0.2}_{-0.2}$

3.3 The effects of metallicity

b: Age ranges, $t_{E,0} \leq \text{Age} \leq 2t_{E,0}$.

	0.05 Z_{\odot}	0.20 Z_{\odot}	0.40 Z_{\odot}	1.00 Z_{\odot}	2.00 Z_{\odot}
H α -	5.5–11.0	5.1–10.1	4.8–9.5	4.3–8.6	3.9–7.9
H α + 10 Å	6.2–12.3	6.3–12.7	5.9–11.7	5.6–11.1	5.1–10.2
H α + 20 Å	8.0–16.1	8.0–16.1	7.7–15.5	7.3–14.6	7.3–14.6
H α + 40 Å	9.0–17.9	10.2–20.3	9.6–19.3	9.3–18.6	9.1–18.2
H α + 80 Å	12.1–24.1	11.7–23.4	11.9–23.7	10.7–21.4	10.7–21.3
H α + 160 Å	15.1–30.1	15.0–30.0	15.5–31.1	16.4–32.7	16.2–32.4
GALEX FUV	26.7–53.5	21.4–42.9	19.7–39.4	17.1–34.2	14.5–29.0
UVOT W2	29.3–58.7	24.8–49.5	22.2–44.3	19.0–38.0	16.5–33.0
WFC3 UVIS1 F218W	29.5–59.0	26.0–52.0	23.3–46.7	19.4–38.9	16.9–33.9
UVOT M2	29.5–58.9	26.1–52.2	23.3–46.7	19.5–39.0	16.8–33.5
GALEX NUV	29.5–59.1	26.3–52.5	23.6–47.3	19.6–39.1	17.1–34.1
WFC3 UVIS1 F225W	29.8–59.5	26.5–53.0	24.7–49.3	19.6–39.3	17.2–34.5
WFPC2 F255W	30.1–60.3	27.4–54.8	26.3–52.6	22.4–44.7	18.4–36.8
UVOT W1	30.4–60.8	28.0–56.0	26.5–53.1	21.8–43.5	18.6–37.2
WFC3 UVIS1 F275W	30.7–61.3	27.9–55.8	28.3–56.5	23.5–47.0	19.1–38.3
WFPC2 F300W	31.6–63.1	28.0–55.9	30.0–60.0	27.7–55.4	21.2–42.5
WFPC2 F336W	31.2–62.3	30.2–60.4	30.3–60.7	33.1–66.3	25.3–50.6
WFC3 UVIS1 F336W	31.4–62.8	29.7–59.5	30.6–61.2	33.3–66.6	25.3–50.6

3 Emission lifetimes of SFR tracers

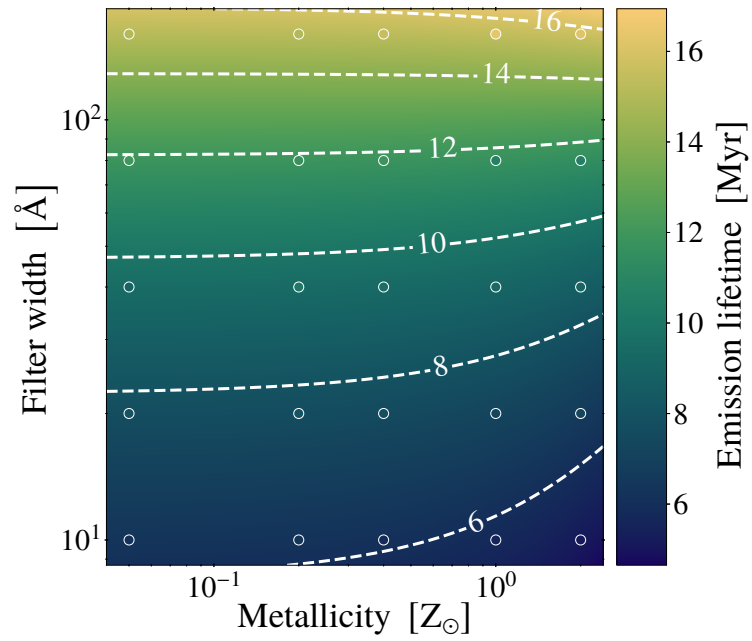


Figure 3.12 The surface described by Equation 3.12^[p. 70], which relates the metallicity and filter width, W , of a $H\alpha+$ filter to the associated emission lifetime for a well sampled IMF. The data points show the measurements coloured using the same colour bar. The surface fits best when it matches the colour of the data points.

3.3 The effects of metallicity

lifetimes (3.9–5.5 Myr for H α –) fall within the range of literature values (1.7–10 Myr, Kennicutt et al. 2012; Leroy et al. 2012).

In Section 3.2^[p.59], we describe a curve which relates the filter width, W , to the emission lifetime of H α + filters, $t_{E,0}^{H\alpha+}$, at solar metallicity. Equation 3.12^[p.70], now extends this relation to include different metallicities to produce a surface in $(t_{E,0}^{H\alpha+}, W, Z)$ space. As mentioned in Section 3.2^[p.59], the H α + emission lifetimes are at the higher end of the literature range (if not exceeding), which is due to including more of the long-lived continuum emission. In Figure 3.10^[p.72], we show the data for constant metallicity and constant width along with the fit described by the $(t_{E,0}^{H\alpha+}, W, Z)$ surface given in Equation 3.12^[p.70] and shown in Figure 3.12^[p.76]. These figures enable a direct assessment of how well the surface describes the emission lifetimes as a function of the metallicity and wavelength.

As we did for $t_{E,0}^{H\alpha+}$, we can extend the relation given by Equation 3.9^[p.64], which describes $t_{E,0}^{UV}$ as a function of response-weighted mean wavelength, $\bar{\lambda}_w$, to also include metallicity. In Figure 3.11^[p.73], we show the data for constant metallicity and constant wavelength along with the fit described by the $(t_{E,0}^{UV}, \bar{\lambda}_w, Z)$ surface given in Equation 3.13^[p.70] and shown in Figure 3.13^[p.78]. The strongest deviations from the fit arise at long ($\lambda_w > 290$ nm) wavelengths. For UV filters at these wavelengths, we recommend interpolating the data points (provided in Table 3.5^[p.74]) rather than adopting Equation 3.13^[p.70]. The range of emission lifetimes found for the UV filters (14.5–33.3 Myr) again fall within the range quoted in literature (10–100 Myr, Kennicutt et al. 2012; Leroy et al. 2012); however, they are towards the low end of this range. This is a direct result of the fact that the UV emission from star-forming regions fades with time, and the measured lifetimes are naturally biased to the ages of regions from which most UV photons emerge.

In summary, we see that the emission lifetimes decrease with increasing metallicity. Observational applications of HEISENBERG should therefore use an SFR tracer lifetime appropriate for the metallicity of the observed region. We define empirical relations between the SFR tracer emission lifetime and the metallicity (for H α –, Equation 3.11^[p.70]) and (for H α + and UV filters, Equations 3.12 and 3.13^[p.70]) the filter properties. For H α + and UV SFR tracers, these relations enable the definition of emission lifetimes even for filters that are not explicitly considered here.

3 Emission lifetimes of SFR tracers

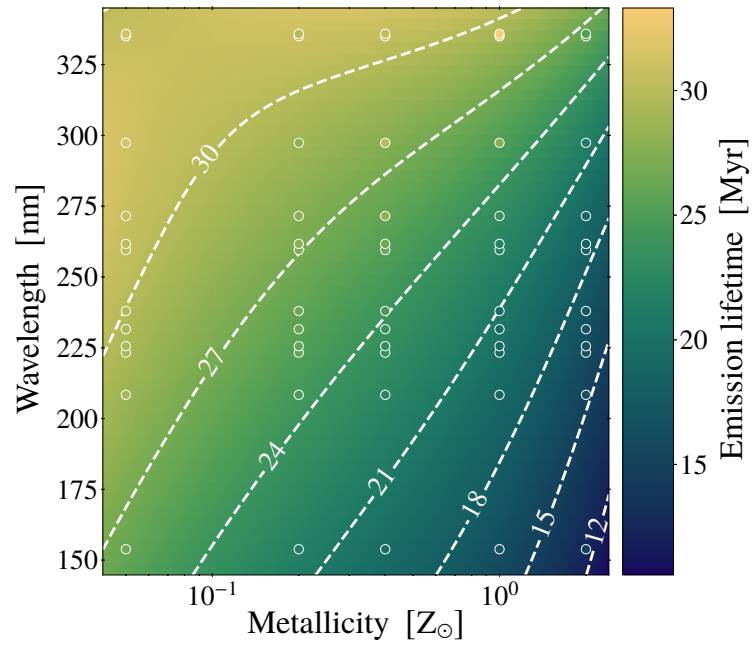


Figure 3.13

The surface described by Equation 3.13^[p. 70], which relates the metallicity and response-weighted mean wavelength, $\bar{\lambda}_w$, of a UV SFR tracer to the associated emission lifetimes for a well sampled IMF. The data points show the measurements coloured using the same colour bar. The surface fits best when it matches the colour of the data points.

3.4 The effects of IMF sampling

In the previous sections, we determine the emission lifetimes of SFR tracers using synthetic emission maps where SLUG2 fully samples the IMF. In observational applications of the KL14 principle, there is no guarantee (or requirement from HEISENBERG) that the regions under consideration have a well sampled IMF. It is therefore important to investigate the impacts of incomplete sampling of the IMF (i.e. a stochastically sampled IMF) on the emission lifetimes of the SFR tracers, in particular for low-mass star forming regions.

We describe in [Chapter 2](#)^[p.39] how the abundance of identified regions in each input map reflects the duration associated to that map. Since star formation is traced using emission from massive stars, an SFR tracer emission map consisting of star-forming regions where the IMF is not well sampled will contain fewer identifiable regions than one with a well sampled IMF. This apparent decrease in the number of star-forming regions is due to the fact that not every region will be able to form stars of sufficient mass to produce the SFR tracer emission. This effect will be particularly important for the $H\alpha$ filters, as $H\alpha$ emission requires high-mass stars ($> 8 M_{\odot}$) and is dominated by stars of even higher masses. We therefore expect that as the sampling of the IMF becomes more incomplete, the effective emission lifetime of the various tracers will decrease, most strongly affecting $H\alpha$.

In [Section 3.4.1](#), we explain how we expect the emission lifetimes to change as a result of incomplete IMF sampling based on purely analytical considerations. In [Section 3.4.2](#)^[p.84], we show how these expectations can be tested experimentally and present the results of these tests.

3.4.1 Theoretical expectation

We quantify the relationship between how well the IMF is sampled and the emission lifetime of the SFR tracer. This is related to the number of star-forming regions that HEISENBERG can identify in the emission map. We therefore estimate the relative change of the effective SFR tracer lifetime as the fraction of star-forming regions that do contain sufficiently massive stars to emit in the tracer of interest. This approach will be tested below. In practice, this means we need to estimate how many stars, N_{\min} , of at least some minimum mass, M_{\min} , are expected to form within a star-forming region of mass M_r . We consider M_{\min} to characterise the stellar mass at which the SFR emission becomes noticeable and not the mass contributing the most. The mass of the star-forming region, M_r , can then act as a proxy for how well the IMF is sampled: smaller values of M_r will result in a region with an IMF that is less well sampled.

3 Emission lifetimes of SFR tracers

We can calculate the probability of producing a minimum number of stars N_{\min} of at least some minimum mass M_{\min} in a given star-forming region through a Bernoulli (i.e. binomial) trial. If the region can produce a sufficient number of stars of sufficient mass, then the region is identifiable in the SFR tracer. Therefore, in our binomial trial, we define a ‘success’ as producing a star of mass M which satisfies the condition of $M_{\min} \leq M \leq M_r$. The probability of success is given by p , N_{\star} is the total number of stars within the star-forming region, and N counts the number of ‘successful’ stars.

The binomial distribution gives the probability of k successful stars as

$$\text{Equation 3.15} \quad P(N = k) = \frac{N_{\star}!}{k! (N_{\star} - k)!} p^k (1 - p)^{N_{\star} - k} .$$

The probability that we wish to calculate (at least N_{\min} stars of a mass of M_{\min} or higher) is given by

$$\begin{aligned} \text{Equation 3.16} \quad P(N \geq N_{\min}) &= 1 - P(N < N_{\min}) \\ &= 1 - \sum_{k=0}^{N_{\min}-1} P(N = k) . \end{aligned}$$

The IMF, dn/dm , describes the distribution of mass amongst the stars within a star-forming region; this means we can use the IMF to determine the values of p and N_{\star} and therefore to calculate $P(N = k)$. In a star-forming region with a well-sampled IMF, p is the fraction of stars that satisfy the condition $M_{\min} \leq M \leq M_r$ and N_{\star} is the total number of stars within the region:

$$\text{Equation 3.17} \quad p = \nu \int_{M_{\min}}^{M_r} \frac{dn}{dm} dm ; \quad N_{\star} = \mu \int_0^{M_r} \frac{dn}{dm} dm .$$

The normalisation constants ν and μ are evaluated through

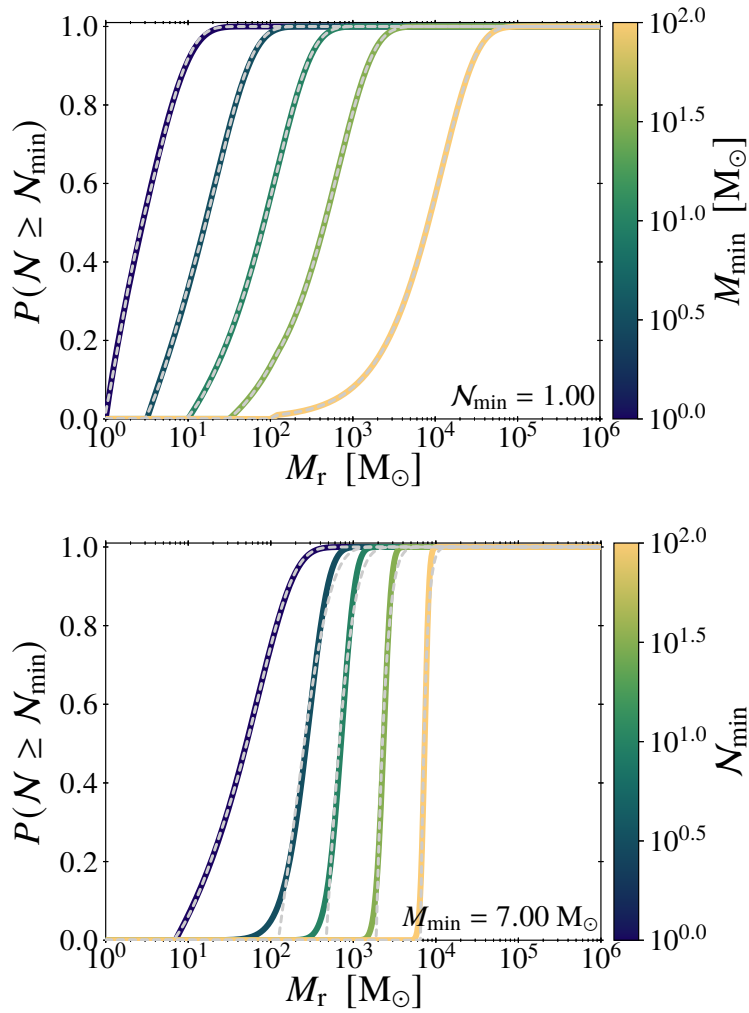
$$\text{Equation 3.18} \quad 1 = \nu \int_0^{M_r} \frac{dn}{dm} dm ; \quad M_r = \mu \int_0^{M_r} m \frac{dn}{dm} dm .$$

In order to convert the probability value, $P(N \geq N_{\min})$, into an estimate for the emission lifetime, t_E , we assume a Chabrier (2005) IMF and use the emission lifetimes we find for a fully sampled IMF, $t_{E,0}$, (see Table 3.5^[p.74]) in the following equation

$$\text{Equation 3.19} \quad t_E = t_{E,0} \times P(N \geq N_{\min}) .$$

In Figure 3.14^[p.81], we show how the form of the probability curve $P(N \geq N_{\min})$ changes for different values of N_{\min} and M_{\min} . Increasing

3.4 The effects of IMF sampling



Curves showing how the probability of forming at least N_{\min} stars of mass M_{\min} or higher changes with star-forming region mass, M_r . The grey dashed lines indicate the approximate fit to the full calculation. See [Section 3.4.1](#)^[p. 79] for details and [Table 3.6](#)^[p. 83] for fit parameter values. Top: Constant N_{\min} . Bottom: Constant M_{\min} .

Figure 3.14

3 Emission lifetimes of SFR tracers

the value of M_{\min} increases the star-forming region mass required to reach a given probability of forming enough sufficiently massive stars (set by \mathcal{N}_{\min} and M_{\min}); the same effect is observed for \mathcal{N}_{\min} but less pronounced. Higher \mathcal{N}_{\min} also affects the probability of forming enough sufficiently massive stars by increasing the rate of change of probability with changing star-forming region mass.

The curves in Figure 3.14^[p.81] have a complex analytical form, therefore we provide a four parameter function that approximates these curves. These approximations are also included in Figure 3.14^[p.81] as dotted grey lines. The following set of equations describe the form of the approximation,

Equation 3.20

$$M_0 = \mathcal{N}_{\min} \times M_{\min} ,$$

Equation 3.21

$$f(M_r) = 1 + a_1 \exp\left(b_1 \left[\frac{M_r}{M_0}\right]\right) + a_2 \exp\left(b_2 \left[\frac{M_r}{M_0}\right]^2\right) ,$$

Equation 3.22

$$P(\mathcal{N} \geq \mathcal{N}_{\min}) \approx \begin{cases} 0 & f(M_r) \leq 0 \\ f(M_r) & 0 < f(M_r) < 1 \\ 1 & f(M_r) \geq 1 \end{cases} ,$$

where a_i, b_i for $i = \{1, 2\}$ are four parameters that we determine through least-squares minimisation.

We present the parameter values for all the approximate curves displayed in Figure 3.14^[p.81] in Table 3.6^[p.83]. For intermediate values of M_{\min} , these best-fitting parameters can be interpolated as a function of $\log_{10}(M_{\min})$. The approximate expression gives an almost identical fit in the cases where $\mathcal{N}_{\min} = 1$ (see Figure 3.14^[p.81]) but for higher values of \mathcal{N}_{\min} the approximation does not perform as well. Fortunately, as we will show below, we only need to consider the case of $\mathcal{N}_{\min} = 1$.

We now have a description of how the emission lifetime of SFR tracers in a star-forming region with a stochastically sampled IMF, t_E , is related to the emission lifetime determined when the IMF is well sampled, $t_{E,0}$; that is, through the probability distribution function, $P(\mathcal{N} \geq \mathcal{N}_{\min})$. The IMF and two free parameters, \mathcal{N}_{\min} and M_{\min} , characterise the form of $P(\mathcal{N} \geq \mathcal{N}_{\min})$. We note that the analytical expression for the correction factor $P(\mathcal{N} \geq \mathcal{N}_{\min})$ does not carry an explicit metallicity dependence. We therefore apply the same theoretical framework for all metallicities, allowing us to combine the effects of both metallicity and IMF sampling on the SFR tracer emission lifetime.

3.4 The effects of IMF sampling

Values for the free parameters, a_i and b_i , in the analytical models presented in Figure 3.14^[p.81] and described by Equations 3.20 to 3.22^[p.82]

Table 3.6

\mathcal{N}_{\min}	M_{\min}	a_1	b_1	a_2	b_2
1.000	1.000	-1.001	-0.251	-0.384	-0.632
1.000	3.162	-1.016	-0.137	-0.146	-0.338
1.000	10.000	-1.021	-0.085	-0.071	-0.235
1.000	31.623	-1.000	-0.048	-0.056	-0.233
1.000	100.000	-1.000	-0.008	-2.393	-5.405
1.000	7.000	-1.020	-0.098	-0.088	-0.257
3.162	7.000	-1.930	-0.118	-0.966	-1.555
10.000	7.000	-4.816	-0.235	-0.987	-1.638
31.623	7.000	-38.108	-0.426	-1.016	-0.891
100.000	7.000	-1000.000	-0.743	-1.016	-0.889

3.4.2 Measuring the emission lifetime of SFR tracers for a stochastically sampled IMF

We adapt the method we present in [Section 3.1](#)^[p. 49] to investigate the effects of a stochastically sampled IMF and thus test experimentally if we recover the same behaviour as described in [Section 3.4.1](#)^[p. 79]. We create the reference maps in the same way as before: the reference maps are mass surface density maps of the star particles within the age ranges specified in [Table 3.5](#)^[p. 74]. The emission maps, however, undergo one additional step. As mentioned previously, we can use the mass of the star-forming region, M_r , as a proxy for how well the IMF is sampled; in this case M_r is the mass of the star particles. We therefore scale these star particle masses by some mass scaling factor, F_m , before SLUG2 predicts the expected emission, this time using its stochastic IMF sampling module. The values of F_m range from 0.01–100, where a lower mass scaling factor means the IMF will be less well sampled. We then use HEISENBERG to determine the emission lifetime, as in [Section 3.2](#)^[p. 59]. The emission lifetime we associate to each mass scaling factor is the average of three lifetimes determined from three independently generated stochastic realisations of the synthetic emission maps. This accounts for the spread in lifetimes that results due to the stochastic nature in which the synthetic emission maps are produced.

To relate the relative change of the SFR tracer lifetime due to IMF sampling to observables, we define an average star-forming region mass, \overline{M}_r , as

$$\text{Equation 3.23} \quad \overline{M}_r = \Sigma_{\text{SFR}} \times \tau \times \pi \left(\frac{\lambda}{2} \right)^2,$$

which uses the SFR surface density, Σ_{SFR} , and quantities that HEISENBERG measures: the total duration of the evolutionary timeline, τ , and the typical separation length of independent star-forming regions, λ , (for details see [Kruijssen et al. 2018](#)).

At a fixed total duration of the evolutionary timeline and region separation length, the degree of IMF sampling is controlled by Σ_{SFR} . We calculate the value of Σ_{SFR} as

$$\text{Equation 3.24} \quad \Sigma_{\text{SFR}} = \frac{\sum_i m_i}{t_{\text{E},0} \pi r^2} \times F_m,$$

where $\sum_i m_i$ is the total mass of all the star particles that fall within the age range appropriate for the filter, i.e. $0 \leq \text{Age} \leq t_{\text{E},0}$ (see [Table 3.5](#)^[p. 74] for the values of $t_{\text{E},0}$), which is then scaled by the mass scaling factor F_m , $t_{\text{E},0}$ is the width of that age range, and r is the radius of the galaxy being studied (for our

3.4 The effects of IMF sampling

simulated galaxy $r = 10$ kpc, as determined from a visual inspection of the synthetic emission maps).

In Equation 3.24^[p. 84], we consider Σ_{SFR} as the galaxy average SFR surface density. If there are no strong large-scale morphological features, as is the case here, this galaxy average SFR surface density is appropriate to use in the calculation of \overline{M}_r . Otherwise, the expression in Equation 3.24^[p. 84] should be updated to include a factor of $\mathcal{E}_{\text{star, glob}}$. This factor is the ratio of the mass surface density⁷ on a size scale of λ to the (reference) map average value; therefore, it accounts for a non-uniform spatial distribution of star-forming regions across the galaxy (for more details see Kruijssen et al. 2018, Section 3.2.9).

By introducing a ‘mass scaling factor’, F_m , we are able to test experimentally how the lifetime of different SFR tracers change when the IMF becomes less well sampled. We will use the experimental results to see if we observe the behaviour predicted in Section 3.4.1^[p. 79].

In Figure 3.15^[p. 86], we present the solar-metallicity results for H α - and WFC3 UVIS F225W as examples of how the emission lifetimes change as a function of the average mass of an independent star-forming region, \overline{M}_r , where lower values of \overline{M}_r correspond to a more stochastically sampled IMF. Each data point⁸ in the two left-hand panels of Figure 3.15^[p. 86] corresponds to a different mass scaling factor, F_m , where the emission lifetime is calculated from the average of the three emission lifetimes determined from the three independently generated synthetic emission maps. The quantity shown on the vertical axis, $t_E/t_{E,0}$, is the factor by which the measured emission lifetime is reduced, compared to the emission lifetime for a well sampled IMF (see Table 3.5^[p. 74] for $t_{E,0}$ values). This is due to incomplete IMF sampling at small region masses or low SFR surface densities. We describe the relation between this conversion factor, $t_E/t_{E,0}$, and \overline{M}_r through $P(\mathcal{N} \geq \mathcal{N}_{\text{min}})$ as in Equation 3.19^[p. 80] (see Section 3.4.1^[p. 79]).

The red curves in Figure 3.15^[p. 86] indicate the best-fitting form of the model $P(\mathcal{N} \geq \mathcal{N}_{\text{min}})$. The analytical form that describes $P(\mathcal{N} \geq \mathcal{N}_{\text{min}})$ has two free parameters \mathcal{N}_{min} and M_{min} . We constrain the values for these two free parameters using a brute-force approach: we calculate the value of χ_{red}^2 for a range of \mathcal{N}_{min} and M_{min} and use the minimum χ_{red}^2 to indicate the best-fitting parameter values. We use the method described in Orear (1982) to determine χ^2 when uncertainties are present on both the abscissa and the ordinate. We

⁷The quantity $\mathcal{E}_{\text{star, glob}}$ represents a mass surface density ratio because the reference maps are mass surface density maps. In typical observational applications, $\mathcal{E}_{\text{star, glob}}$ would be a flux density ratio.

⁸For details of the error calculation on \overline{M}_r , see Section 3.B^[p. 97].

3 Emission lifetimes of SFR tracers

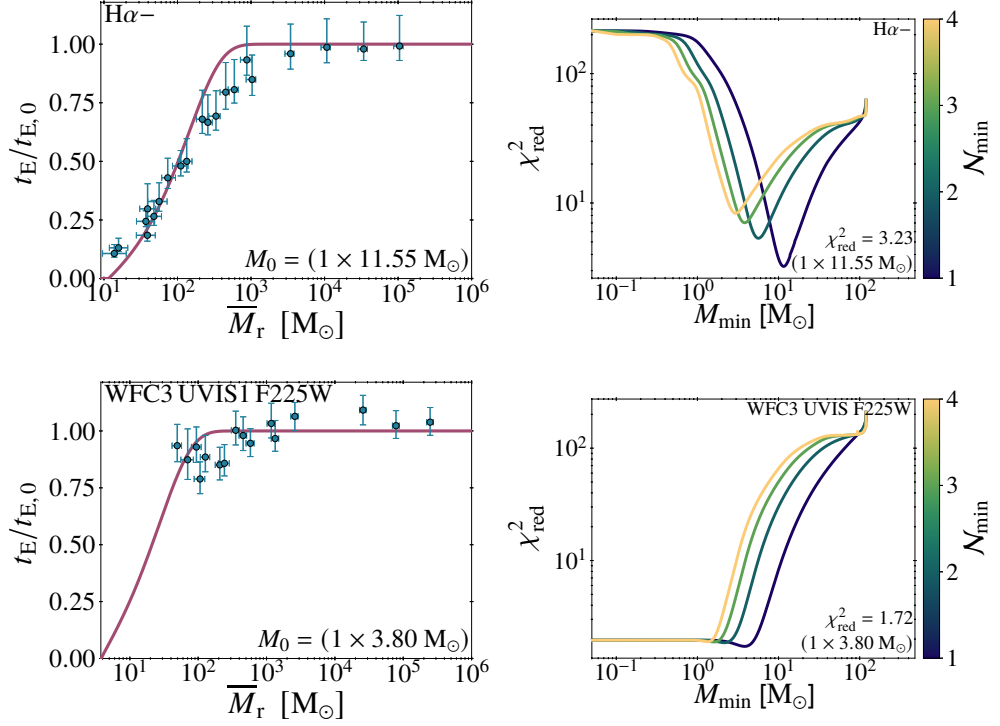


Figure 3.15

Top row: H α emission excluding the continuum (H α -). Bottom row: UV emission (WFC3 UVIS1 F225W). Left column: Change of the emission lifetime of SFR tracers, relative to the emission lifetime we determine from a well sampled IMF, as a function of the average independent star-forming region mass, \overline{M}_r . The data points show the results of the experiments in which we apply HEISENBERG to synthetic SFR tracer maps with a stochastically sampled IMF at solar metallicity. For comparison, the red curve shows the best-fitting analytical model from Section 3.4.1^[p. 79]. At low region masses, the emission lifetimes decrease due to the incomplete sampling of the IMF. Right column: Change of χ_{red}^2 with minimum stellar mass, M_{min} , and the minimum number of stars of that mass, \mathcal{N}_{min} . The minimum χ_{red}^2 found is indicated in the bottom right with the best-fitting model parameters ($\mathcal{N}_{\text{min}} \times M_{\text{min}}$).

3.4 The effects of IMF sampling

The functional form of the conversion factor, $P(\mathcal{N} \geq \mathcal{N}_{\min})$, between the emission lifetime measured for a well-sampled IMF and a stochastically sampled IMF has two parameters, \mathcal{N}_{\min} and M_{\min} . We use $\mathcal{N}_{\min} = 1$ and show here the best fitting values of M_{\min} .

Table 3.7

	0.05 Z_{\odot}	0.20 Z_{\odot}	0.40 Z_{\odot}	1.00 Z_{\odot}	2.00 Z_{\odot}
H α -	11.50	11.95	13.00	11.55	10.45
H α + 10 Å	10.75	10.05	12.25	13.90	12.00
H α + 20 Å	12.10	12.05	12.40	10.35	9.95
H α + 40 Å	9.65	9.45	7.95	8.35	10.45
H α + 80 Å	9.20	8.70	8.85	6.25	8.00
H α + 160 Å	10.20	5.20	7.80	8.60	9.35

3 Emission lifetimes of SFR tracers

consider N_{\min} values ranging from 1 to 4 in integer steps and M_{\min} , from 0 to $120 M_{\odot}$ in steps of $0.05 M_{\odot}$. In the right-hand panels of [Figure 3.15](#)^[p. 86], we show the dependence of χ_{red}^2 on N_{\min} and M_{\min} for the two example filters. [Table 3.7](#)^[p. 87] lists the best-fitting values of M_{\min} ($N_{\min} = 1$ in all cases) for the full range of metallicities ($Z/Z_{\odot} = 0.05, 0.20, 0.40, 1.00, 2.00$) for the $\text{H}\alpha$ filters.

The data points used in the fitting process (and included in [Figure 3.15](#)^[p. 86]) are those that survive a selection cut: we rejected data points for which the calculated emission lifetime exceeds the lifetime for a well sampled IMF by more than 1σ (i.e. $t_{\text{E}}/t_{\text{E},0} > 1$). We choose to remove the data that do not satisfy this criterion because they indicate contamination by the continuum emission from (low-mass) stars. For low mass scaling factors (typically $F_{\text{m}} \approx 0.01$), the emission from the continuum dominates over the SFR tracer. This results in emission lifetimes that describe the long-lived continuum emission and therefore can be orders of magnitude higher than $t_{\text{E},0}$.

For the UV filters, we find that data points associated with low mass scaling factors ($F_{\text{m}} \approx 0.01$) are excluded by our data selection criterion. This results in the turn off from $t_{\text{E}}/t_{\text{E},0} = 1$ being very poorly sampled (see [Figure 3.15](#)^[p. 86] for an example). This means that we cannot reliably distinguish between different N_{\min} and M_{\min} ; therefore, we conclude that UV emission is not significantly affected by IMF sampling and exclude the UV M_{\min} values from [Table 3.7](#)^[p. 87].

[Table 3.7](#)^[p. 87] shows that there is no clear relation between M_{\min} and metallicity or filter width, but generally speaking, smaller filter widths have higher M_{\min} . Higher values of M_{\min} imply higher star-forming region masses below which IMF sampling cannot be neglected (i.e. where $P(N \geq N_{\min}) < 1$). For $\text{H}\alpha-$ (+), M_{\min} ranges from 10.45–13.00 (5.20–13.90) M_{\odot} . From these M_{\min} ranges, we obtain region masses below which incomplete IMF sampling affects the SFR tracer lifetimes, that is, these are the lowest region masses according to [Equation 3.23](#)^[p. 84] for which IMF sampling can be neglected. For $\text{H}\alpha-$ (+), this range is $\overline{M}_{\text{r}} \gtrsim 600\text{--}800$ (200–900) M_{\odot} . For a region separation length of $\lambda = 200$ pc and a total timeline duration of $\tau = 30$ Myr, these characteristic region mass limits correspond to $\Sigma_{\text{SFR}} \gtrsim (6\text{--}9) \times 10^{-4} M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ for $\text{H}\alpha-$ and $\Sigma_{\text{SFR}} \gtrsim (2\text{--}10) \times 10^{-4} M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ for $\text{H}\alpha+$.

[Figure 3.15](#)^[p. 86] demonstrates that it is important to consider the effects of IMF sampling at low SFR surface densities, when constraining the emission lifetime for the $\text{H}\alpha\pm$ filters. This is because at low SFR surface densities, the massive stars required to produce $\text{H}\alpha$ emission are not always present. If we ignore this fact, the $\text{H}\alpha\pm$ emission lifetime will be overestimated; as a result, the evolutionary timeline would be incorrectly calibrated and the lifetimes obtained

3.5 Comparison to observations

with HEISENBERG would also be overestimated. The agreement between the results of these experiments and the theoretical model also demonstrate that the IMF sampling theory presented in Section 3.4.1^[p. 79] accurately describes how the emission lifetime of $H\alpha\pm$ changes due to incomplete IMF sampling. This means that observational applications of the KL14 principle can use the expressions provided in Equations 3.19 to 3.23^[pp. 80, 82 & 84] to derive an SFR emission lifetime corrected for IMF sampling. For the UV tracers, however, the emission lifetimes are mostly insensitive to the effects of incomplete IMF sampling and so these effects can be largely ignored. The constancy of the UV emission lifetime over the range of \overline{M}_r is a result of stars being able to produce UV emission at low masses. This means that the emission will almost always be present in star-forming regions.

In summary, we have shown that the effects of IMF sampling can have a considerable impact on the measured lifetimes of the $H\alpha\pm$ filters. The change of emission lifetime with star-forming region mass \overline{M}_r is well-described by the analytical description of IMF sampling from Section 3.4.1^[p. 79]. This analytical description provides the correction factors the SFR tracer lifetimes require to enable applications of HEISENBERG in regions of low SFR surface densities. By contrast, the emission lifetime of UV tracers is found to be largely insensitive to IMF sampling, implying that no such correction factor is needed. We reiterate that all SFR tracers do still carry an important dependence on the metallicity that should always be accounted for.

3.5 Comparison to observations⁹

Thus far, we have made predictions of the emission lifetime for SFR tracer using galaxy simulations. This was necessary because the KL14 principle method requires a reference map as input, for which the duration of the emission is known exactly. Constructing a reference map for a simulation is straightforward: we produce a map populated by the star particles within a known age range and so the reference timescale is known. For observed galaxies, this is not possible. However, it is possible to test whether the observed *ratio* between the emission lifetimes of two different SFR tracers is consistent with our predictions. In Kruijssen, Schrubba, Chevance et al. (2019), we used our predicted $H\alpha$ emission lifetime at the half-solar metallicity of NGC300 ($t_{E,0}^{H\alpha} = 4.59 \pm 0.14$ Myr) to measure a CO cloud lifetime of $10.8^{+2.1}_{-1.7}$ Myr. We can now use this CO cloud lifetime as a reference timescale in an experiment combining the CO map (now acting as the reference map) with a FUV emission map. This allows us to test

⁹This section was contributed by collaborators Mélanie Chevance and Diederik Kruijssen

3 Emission lifetimes of SFR tracers

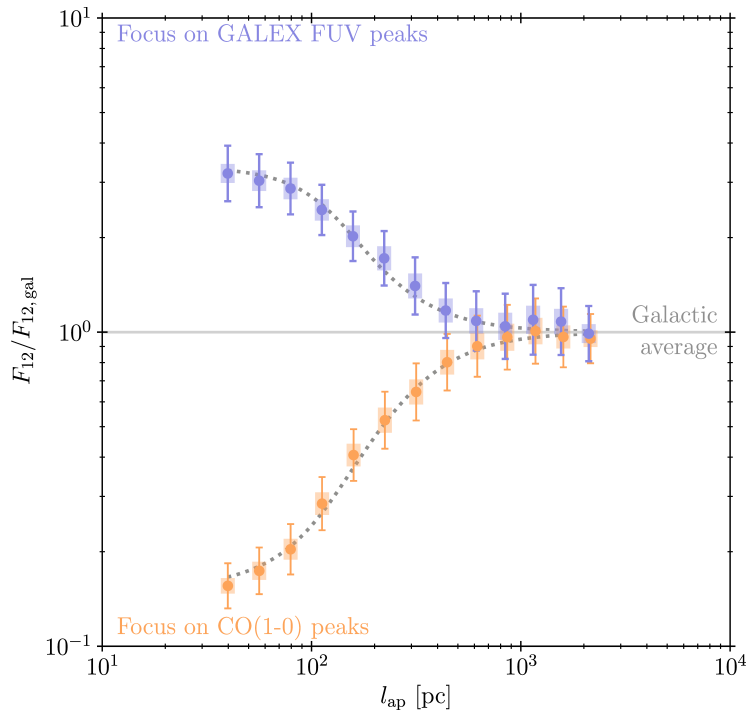
if the resulting FUV emission lifetime is consistent with our prediction for the FUV emission lifetime.

We conduct the first test for the accuracy of our inferred lifetime as follows. We combine a GALEX FUV map of NGC300 with the CO data presented in Kruijssen, Schrubba, Chevance et al. (2019). For the CO map, we use the identical experiment setup as in Kruijssen, Schrubba, Chevance et al. (2019): we adopting the same set of identified emission peaks and remove diffuse emission in the same way. For the FUV map, the emission peaks are identified over a flux range of 1.3 dex below the brightest peak in the map, using flux contours at intervals of 0.35 dex to separate adjacent peaks. In addition, we remove the DC offset from the FUV map by filtering it with a high-pass Gaussian filter in Fourier space on a size scale $> 1000\lambda$. Other than these details, we apply the default analysis described in Kruijssen et al. (2018) and Kruijssen, Schrubba, Chevance et al. (2019). The resulting tuning fork diagram (also see Figure 2.2^[p. 42]) is shown in Figure 3.16^[p. 91]. We obtain a good fit, with a FUV emission lifetime of $t_{\text{obs}}^{\text{FUV}} = 23.1^{+5.9}_{-3.5}$ Myr. Given the half-solar metallicity of NGC300, this should be compared to the emission lifetime predicted by Equation 3.13^[p. 70] for $Z = 0.5 Z_{\odot}$, which is $t_{\text{E},0}^{\text{UV}} = 19.2 \pm 2.0$ Myr.

Fundamentally, this experiment expresses the lifetime for GALEX FUV emission in units of the reference timescale of continuum-subtracted $\text{H}\alpha$ emission (using CO as an intermediate step). This is the case because in both the CO- $\text{H}\alpha$ experiment (Kruijssen, Schrubba, Chevance et al. 2019) and the CO-FUV experiment carried out here, we have only measured the ratio of the emission lifetimes. We should thus compare the observed and predicted $t^{\text{FUV}}/t^{\text{H}\alpha}$ ratio. We measure $t_{\text{obs}}^{\text{FUV}}/t_{\text{obs}}^{\text{H}\alpha} = 5.0^{+1.3}_{-0.8}$, whereas the calibration of this chapter predicts $t_{\text{E},0}^{\text{FUV}}/t_{\text{E},0}^{\text{H}\alpha} = 4.2 \pm 0.5$. These values agree to within the uncertainties (at 0.9σ , or ~ 20 per cent), which acts as a first demonstration that the emission lifetimes derived in this work are consistent with observations.

Our method yields a measurement of the FUV-to- $\text{H}\alpha$ emission lifetime *ratio*. This means that an arbitrary scaling of both the $\text{H}\alpha$ and FUV emission lifetimes, for either the predicted or observed ratios, would also result in agreement. However, the absolute lifetimes individually must also still be physical. A comparison of the above numbers to other measurements in the literature shows that they fall within the range of expected values. For instance, Leroy et al. (2012) find that a young stellar population has emitted 50 per cent of its $\text{H}\alpha$ emission after 1.7 Myr, and 95 per cent after 4.7 Myr. Our characteristic $\text{H}\alpha$ emission lifetime at solar metallicity of 4.3 Myr falls within this range. The same applies for our GALEX FUV emission lifetime of 17.1 Myr, which falls within the time interval at which 50–95 per cent of the cumulative flux has been emitted (4.8–65 Myr, see Leroy et al. 2012). Future work combining

3.5 Comparison to observations



Tuning fork diagram obtained for the combination of CO(1 \rightarrow 0) (Phase 2) and GALEX FUV (Phase 1) emission of NGC300 (see Kruijssen, Schruba, Chevance et al. 2019). The data points show the relation between the Phase-1-to-Phase-2 flux ratio (F_{12}) calculated at the locations of the emission peaks, relative to the galactic-scale Phase-1-to-Phase-2 flux ratio ($F_{12, \text{gal}}$) as a function of aperture size, l_{ap} . The error bars indicate the 1σ uncertainty on each individual data point, whereas the shaded areas indicate the effective 1σ uncertainty range that accounts for the covariance between the data points. The best-fitting model is indicated by the two dotted curves.

Figure 3.16

H α and UV observations of nearby galaxies will enable a more comprehensive test of the presented lifetimes.

3.6 Conclusions

In this chapter, we have used the HEISENBERG code to constrain the emission lifetimes of SFR tracers, i.e. the durations over which H α and UV emission emerges from coeval stellar populations. We expect these lifetimes to be critical in a variety of future studies. Firstly, observational applications of HEISENBERG will enable the empirical characterisation of the cloud life cycle across a wide range of galactic environments, by measuring e.g. the molecular cloud lifetime and the timescale for cloud destruction by feedback. However, in order to lead to physically meaningful constraints, these applications require the use of a known ‘reference timescale’ for turning the measured relative timescales into absolute ones. This reference timescale is provided by the SFR tracer lifetimes obtained in this chapter. Secondly, the emission lifetimes obtained here and their dependence on metallicity and filter properties provide a helpful point of reference for studies of photoionisation feedback and UV heating.

To obtain the SFR tracer lifetimes, we generate synthetic SFR tracer emission maps of a simulated near- L_{\star} , isolated flocculent spiral galaxy using the stochastic SPS code SLUG2. We then apply HEISENBERG to combinations of these synthetic emission maps and an independent set of ‘reference maps’, which is populated with the star particles from the simulation in specific, known age ranges. With this approach, we self-consistently measure the emission lifetimes for H α emission (with and without continuum subtraction), as well as 12 different UV filters.

For stellar populations at solar metallicity and with a fully sampled IMF we find the emission lifetimes for H α - (+) to be $4.3^{+0.1}_{-0.3}$ Myr (5.6–16.4 Myr). For the UV filters, the emission lifetime falls in the range 17.1–33.3 Myr, and nearly monotonically increases with wavelength. When considering stellar populations with different metallicities ($Z/Z_{\odot} = \{0.05, 0.20, 0.40, 1.00, 2.00\}$) the range of emission lifetimes increases, to 3.9–5.5 Myr (5.1–16.4 Myr) for H α - (+) and 14.5–33.3 Myr for the UV filters. We define empirical power-law relations that provide the emission lifetime as a function of metallicity (Equations 3.11 to 3.13^[p. 70]). These empirical relations include the response-weighted mean wavelength, $\bar{\lambda}_w$, for UV filters and the filter width, W , for the H α + filters. These dependences enable the use of a single expression to determine the emission lifetime for all UV and H α \pm SFR filters for a given combination of filter properties and the metallicity of the environment.

3.6 Conclusions

We also investigate the effects of a stochastically sampled IMF on the emission lifetimes. Incomplete IMF sampling is found to affect the obtained emission lifetimes in low- Σ_{SFR} galaxies. We first predict, then demonstrate, that $t_{\text{E}}/t_{\text{E},0}$, the factor by which the emission lifetime changes as a result of IMF sampling effects, is given by the probability of forming at least N_{min} stars of mass M_{min} or higher in a given a star-forming region of mass \overline{M}_{r} . Given an SFR surface density (from which the characteristic region mass can be derived), this relation quantifies the relative change of the SFR tracer lifetime due to IMF sampling as a function of the galactic environment.

For UV tracers, the impact of IMF sampling on the emission lifetime is minimal (< 30 per cent) and can therefore be ignored (this applies to all metallicities). However, incomplete IMF sampling has a significant effect on the lifetime of $\text{H}\alpha$ emission. Depending on the metallicity and on whether the continuum emission has been subtracted, the emission lifetime for a well sampled IMF can be used for $\overline{M}_{\text{r}} \gtrsim 200\text{--}900 M_{\odot}$, which for a region separation length of $\lambda = 200$ pc and a total timeline duration of $\tau = 30$ Myr corresponds to $\Sigma_{\text{SFR}} \gtrsim (2\text{--}10) \times 10^{-4} M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$. However, for lower region masses (SFR surface densities), the $\text{H}\alpha$ emission lifetime must be corrected to account for the effects of IMF sampling.

We have arrived at the above emission lifetimes by carrying out a set of numerical experiments using galaxy simulations and so one could argue that the results are model-dependent. We reiterate that the results are not expected to be sensitive to the details of the baryonic physics in the simulations (see discussion in [Section 3.1](#)^[p. 49]). In principle, these measurements could have been performed using maps of randomly-generated distributions of regions: fundamentally, we have only characterised how quickly young stellar emission fades in the adopted SPS model. However, the main advantage of using a galaxy simulation is that it generates a distribution with a physically reasonable imprint of galactic morphology and the positional correlation of star formation events by self-gravity and stellar feedback. The critical goal of the work in this chapter is to measure the emission lifetimes self-consistently within the framework of our method and thus enabling its future observational applications. The accuracy of the results is demonstrated by a first comparison to observations of $\text{H}\alpha$ and GALEX FUV emission in the nearby galaxy NGC300 ([Section 3.5](#)^[p. 89]), which shows that the emission lifetime predicted by this work are consistent with the observed lifetimes.

In summary, we have measured the emission lifetimes of SFR tracers as a function of metallicity and (for UV and $\text{H}\alpha$ +) filter properties, as well as their sensitivity to IMF sampling, which effectively expresses their dependence on the SFR surface density. This spans the range of key environmental factors that

3 Emission lifetimes of SFR tracers

affect the lifetimes of $H\alpha$ and UV emission, and provides important constraints on the duration of photoionisation feedback and UV heating. The emission lifetimes derived in this work enable observational applications of the KL14 principle, in which they are used to turn the relative durations of evolutionary phases into an absolute timeline. The first applications of this method have already used the results we find here to infer the durations associated to cloud evolution, star formation, and feedback (Chevance and author 2019; Chevance, Kruijssen et al. 2019; Hygate et al. 2019; Kruijssen, Schrubba, Chevance et al. 2019; Ward et al. 2019, as well as Section 3.5^[p. 89] of this paper). In view of the variety of upcoming applications of this method, the emission lifetimes presented in this work thus represent an essential ingredient towards empirically constraining the physics driving molecular cloud life cycle.

3.A $H\alpha$ Absorption and Emission Features

We produce synthetic emission maps by passing the age and mass information of all the star particles from our simulation to SLUG2. The SLUG2 code then calculates the predicted emission spectrum for each particle, to which we apply UV and $H\alpha+$ filters ($H\alpha-$ comes directly from the Hydrogen-ionizing photon emission). However, the emission spectrum that SLUG2 produces does not include the underlying $H\alpha$ absorption from the stellar continuum. In this section, we use STARBURST99 to investigate when the $H\alpha$ absorption feature can no longer be neglected.

We ran STARBURST99 for an instantaneous burst of star formation for the five standard Geneva evolutionary tracks using a Kroupa (2001) IMF and output the data in 0.1 Myr time steps for 20 Myr. We otherwise used the default settings.

The equivalent width of the $H\alpha$ emission is taken directly from the STARBURST99 output files. To determine the equivalent width of the absorption feature, we model the continuum (straight line) and the absorption feature (Voigt profile, Olver et al. 2010) of the high resolution spectral data in the wavelength range $6482 \text{ \AA} \leq \lambda \leq 6642 \text{ \AA}$.

In Figure 3.17^[p. 96], we show the change in the equivalent width of the absorption and emission feature over time; the change in the difference between the two equivalent widths is also included. We see that the emission feature is dominant up to at least 10 Myr and longer for the lower metallicities; this is at least 5 Myr longer than the $H\alpha-$ lifetimes we measure (see Table 3.5^[p. 74]) which are also marked in Figure 3.17^[p. 96].

We can see from Figure 3.17^[p. 96] that the $H\alpha$ lifetimes we are considering fall comfortably within the emission-dominant regime and conclude that the

3.A $H\alpha$ Absorption and Emission Features

absorption feature can safely be neglected for our analysis.

3 Emission lifetimes of SFR tracers

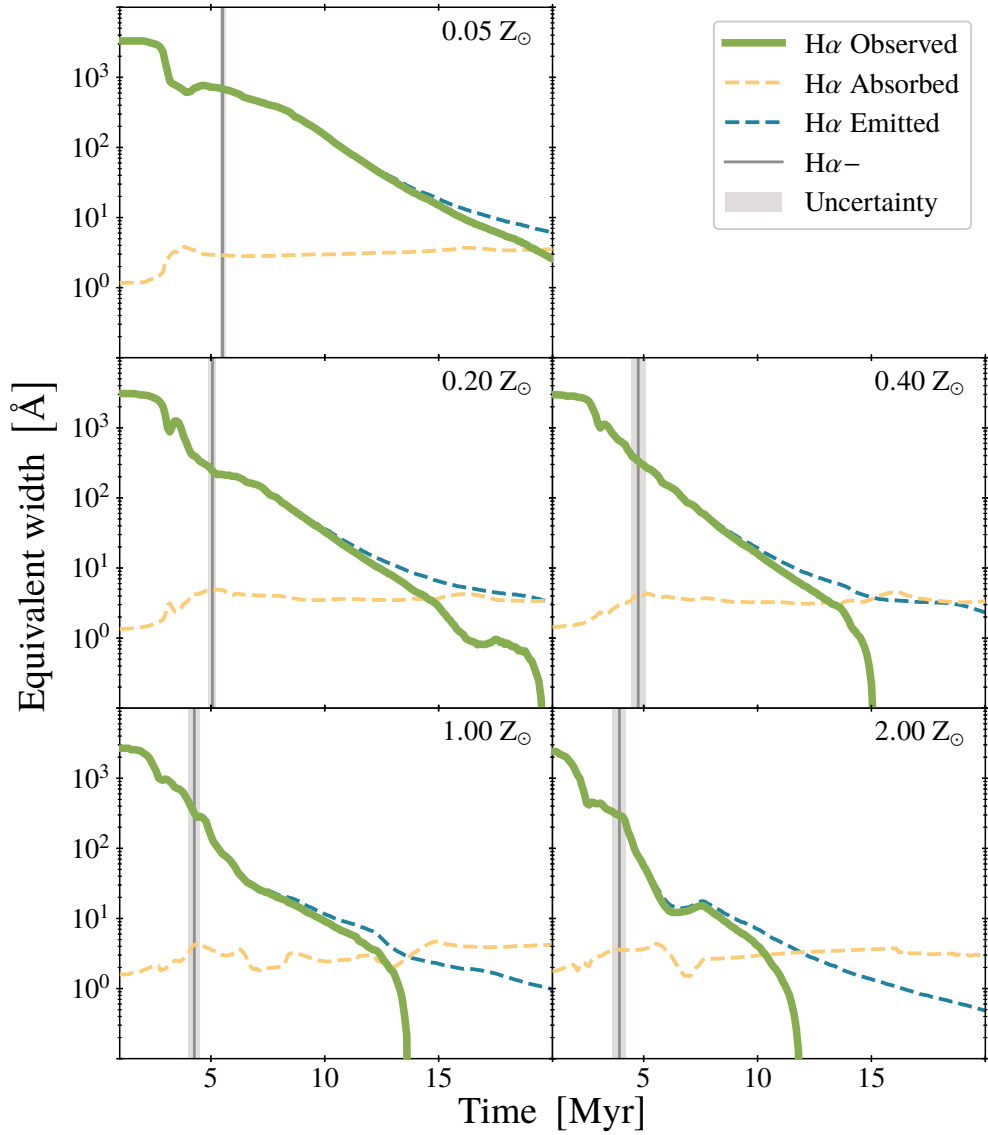


Figure 3.17

The results of `STARBURST99` simulations for an instantaneous burst of star formation at 0 Myr. We show the change in equivalent width of the H α absorption and emission feature. We also include the difference between the two equivalent widths (H α Observed). The H α - emission lifetime is marked for comparison (see [Table 3.5](#)^[p. 74]).

3.B Error Propagation

In [Section 3.4.2](#)^[p. 84], we calculate the characteristic mass on an independent star-forming region as

$$\overline{M}_r = \Sigma_{\text{SFR}} \times (t_E + t_R - t_O) \times \pi \left(\frac{\lambda}{2} \right)^2 . \quad \text{Equation 3.25}$$

This equation uses the SFR surface density, Σ_{SFR} , and the duration of the reference map, t_R , along with quantities that the HEISENBERG code measures: the typical separation length of independent star-forming regions, λ ; the duration of the emission map, t_E ; and the duration of the overlap between the emission and reference phases, t_O . We note that [Equation 3.25](#) and [Equation 3.23](#)^[p. 84] are equivalent through the definition

$$\tau \equiv t_E + t_R - t_O . \quad \text{Equation 3.26}$$

Here we describe how we propagate the uncertainties on these quantities into an uncertainty on the characteristic region mass \overline{M}_r .

To calculate the uncertainty on \overline{M}_r we start with the general expression: the uncertainty on a quantity f , σ_f , which is a function of N variables i.e. $f(x_1, \dots, x_N)$ is given by ([Hughes et al. 2010](#))

$$\sigma_f^2 = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \rho_{ij} \sigma_i \sigma_j , \quad \text{Equation 3.27}$$

where σ_i represents the uncertainty on variable x_i and ρ_{ij} represents correlation coefficients between variable x_i and x_j (where $\rho_{ii} = 1$ and $\rho_{ij} = \rho_{ji}$). In order to simplify our expressions and to use the same notation as in [Equation 3.27](#), we define the following

$$\kappa \equiv \Sigma_{\text{SFR}} \frac{\pi}{4} ; \quad x_1 \equiv \lambda ; \quad x_2 \equiv t_E ; \quad x_3 \equiv t_O \quad \text{Equation 3.28}$$

and [Equation 3.25](#) becomes

$$\begin{aligned} \overline{M}_r &= \kappa \tau x_1^2 \\ &= \kappa (x_2 + t_R - x_3) x_1^2 . \end{aligned} \quad \text{Equation 3.29}$$

We note that Σ_{SFR} and t_R are considered to be without error and do not need to be included as variables. The derivatives we need in order to calculate $\sigma_{\overline{M}_r}$ are

$$\frac{\partial}{\partial x_1} \overline{M}_r = 2\kappa \tau x_1 ; \quad \frac{\partial}{\partial x_2} \overline{M}_r = \kappa x_1^2 ; \quad \frac{\partial}{\partial x_3} \overline{M}_r = -\kappa x_1^2 . \quad \text{Equation 3.30}$$

3 Emission lifetimes of SFR tracers

Combining Equation 3.30^[p. 97] with Equation 3.27^[p. 97] we find the expression for the uncertainty on \overline{M}_r , $\sigma_{\overline{M}_r}$:

Equation 3.31

$$\left[\frac{\sigma_{\overline{M}_r}}{\overline{M}_r} \right]^2 = \frac{4\sigma_1^2}{x_1^2} + \frac{(\sigma_2^2 + \sigma_3^2 - 2\rho_{23}\sigma_2\sigma_3)}{\tau^2} + \frac{4(\rho_{12}\sigma_1\sigma_2 - \rho_{13}\sigma_1\sigma_3)}{\tau x_1} .$$

With this expression, we can take into account the associated uncertainty on the value of \overline{M}_r as part of our error analysis and χ^2 calculations when investigating the effects of incomplete IMF sampling on the emission lifetimes of SFR tracers (see Section 3.4^[p. 79]).

Acknowledgements

This research was undertaken with the assistance of resources from the Baden-Württemberg high-performance computing (bwHPC) cluster: bwForCluster MLS&WISO Production, funded by the German Research Foundation (DFG) through grant INST 35/1134-1 FUGG, and bwForCluster BinAC, through INST 37/935-1 FUGG.

The influence of dust extinction on the lifetime of star formation rate tracers

This chapter is based on

D. T. Haydon et al. (Aug. 2019). ‘An uncertainty principle for star formation – V. The influence of dust extinction on star formation rate tracer lifetimes’. *MNRAS Submitted*

In the previous chapter, we constrained the emission lifetimes over which different SFR tracers emit their light as a function of the SFR surface density and metallicity. We obtained these emission lifetimes for several tracers of recent star formation: 12 UV filters (from GALEX, Swift, and HST) covering a wavelength range 150–350 nm, as well as $H\alpha$ with and without continuum subtraction. However, in [Chapter 3](#)^[p.47] we omitted the effects of extinction on SFR tracer lifetimes. This was partially for the sake of convenience, but also because the effects of extinction can, in most cases, be significantly reduced if not completely corrected for (e.g. James et al. 2005). However, correcting for extinction is not always possible or practical. In particular, for many galaxies we have access to only a single $H\alpha$ image or a single UV band, and thus common extinction correction methods such as the Balmer decrement (Berman 1936) or the UV spectral slope (Calzetti et al. 1994) may not be available. For this reason, it is helpful to have an alternative approach available, which we investigate in this chapter. Since understanding the affects of extinction would allow us to make estimates for gas evolution timescales from gas maps combined with maps of star formation tracers, even when we are not able to make an explicit extinction correction to the tracer maps.

The central idea of our work is that extinction does not fundamentally change the underlying emission lifetime, but it does act to reduce the amount of observed emission; this results in a different (effective) emission lifetime. We can calibrate the emission lifetimes associated with extincted tracers using much the same approach that we used in [Chapter 3](#)^[p.47] to calibrate the unextincted ones. That is, using the KL14 principle with synthetic emission maps. We

4 The influence of dust extinction on SFR tracer lifetimes

generate the synthetic emission maps from a high-resolution hydrodynamical simulation of an isolated, Milky-Way-like disc galaxy (a different simulation to that described in [Section 3.1.1](#)^[p.51]). The stellar feedback included in this simulation is inefficient compared to observations, implying that it represents a limiting case in which the duration of embedded star formation (and the corresponding effect of extinction) is overestimated.

The structure of this chapter is as follows. In [Section 4.1](#), we summarise the key input quantities for the HEISENBERG code, describe the details of the simulations, and the procedure followed to generate synthetic emission maps. In [Section 4.2](#)^[p.110], we investigate and discuss how reference timescales are affected as a result of extinction and in [Section 4.3](#)^[p.117], how this can be altered further by changing sensitivity limits. We summarise our findings in [Section 4.4](#)^[p.121].

4.1 Measuring the effects of extinction

In [Chapter 3](#)^[p.47], we presented measurements of the emission lifetimes for 18 SFR tracer filters. These measurements were, however, based on extinction-free emission maps, giving us the extinction-free emission lifetime. In this chapter we want to understand how extinction can alter the measured emission lifetime. Here we outline the method we use to constrain R_{\bullet} : the factor by which the duration associated to an SFR tracer emission map without extinction (t_{\circ}) changes when including extinction (t_{\bullet}). We state this more explicitly:

Equation 4.1

$$R_{\bullet} = \frac{t_{\bullet}}{t_{\circ}} .$$

We constrain R_{\bullet} for the same 18 SFR filters we previously considered and summarised in [Table 3.1](#)^[p.50].

In [Section 4.1.1](#), we give an overview of key input quantities for the HEISENBERG code, which we use to measure R_{\bullet} directly. Our measurements of R_{\bullet} are based on the simulated galaxy we describe in [Section 4.1.2](#)^[p.101]. From this simulated galaxy, we produce two groups of synthetic emission maps: emission maps with extinction and emission maps without (detailed in [Section 4.1.3](#)^[p.104]). The pair of synthetic SFR tracer maps (i.e. extinguished and unextinguished) are used as input for HEISENBERG.

4.1.1 Analysis framework

The analysis we perform here makes use of the HEISENBERG code, which we describe in detail in [Chapter 2](#)^[p.39]. In this section, we highlight the key

4.1 Measuring the effects of extinction

aspects of HEISENBERG that are relevant to this chapter, along with the input values that differ from the default.

We noted in [Chapter 2](#)^[p.39] that HEISENBERG constrains three quantities, one of which is $t_{\text{gas}}/t_{\text{star}}$. Where the value of t_{star} would be used as reference timescale to recover t_{gas} in absolute terms. However, using the notational transformations described at the end of [Chapter 2](#)^[p.39] ('gas' \rightarrow '•' and 'star' \rightarrow '◦'), we can see that this quantity is in fact R_{\bullet} (cf. [Equation 4.1](#)^[p.100]). This means we do not need to create reference maps or concern ourselves with reference timescales as was necessary in [Chapter 3](#)^[p.47]: we want to measure relative durations.

When using HEISENBERG we use the default input parameters that are listed in Kruijssen et al. (2018, Tables 1 and 2); with the following exceptions. The set-up of the galaxy simulation necessitates the use of cuts in galactocentric radius: we use $R_{\text{min}} = 3$ kpc and $R_{\text{max}} = 11$ kpc (and set the `cut_radius` flag to 1). We also change the range of aperture sizes: we use a minimum aperture size of $l_{\text{ap, min}} = 25$ pc and have $N_{\text{ap}} = 17$ apertures. This results in 17 logarithmically spaced aperture diameters from 25–6400 pc. We set parameters which allows use to measure R_{\bullet} directly (`tstariso` = 1, `tstar_incl` = 1). We also use $\Delta \log_{10} \mathcal{F}_{\text{star}} = \Delta \log_{10} \mathcal{F}_{\text{gas}} \equiv \Delta = 2$.¹ The $\Delta \log_{10} \mathcal{F}_i$ are introduced in [Chapter 2](#)^[p.39] and describe (approximately) the dynamic range of the emission in map i starting from the brightest emission peak and going down to the noise floor. While $\Delta = 2$ dex is often appropriate for observational applications (e.g. Chevance, Kruijssen et al. 2019; Kruijssen, Schrubba, Chevance et al. 2019), as well as for the simulation here, we would like to emphasize that this is not always the case.

4.1.2 Galaxy simulation

The work presented in this chapter uses the high-resolution hydrodynamical simulation of an isolated Milky-Way-like disc galaxy described in Fujimoto et al. (2018). We use the adaptive mesh refinement code ENZO (Bryan et al. 2014) to simulate a 128 kpc box with cell sizes ranging from 7.8125 pc to 31.25 pc over seven levels of refinement during the initial stages of the simulation. The initial conditions of the simulation match those of Tasker et al. (2009). We have a gas disc in a static background potential which accounts for both dark matter and the stellar disc. The axisymmetric background potential has a logarithmic form with a constant circular velocity of 200 km s^{-1} at large galactocentric radii ($r > 2$ kpc). The gas disc has an initial mass of

¹This is the default value, as given in Kruijssen et al. (2018, Table 2); however, Δ plays a role in [Sections 4.2](#) and [4.3](#)^[pp. 110 & 117] and so we explicitly state it here.

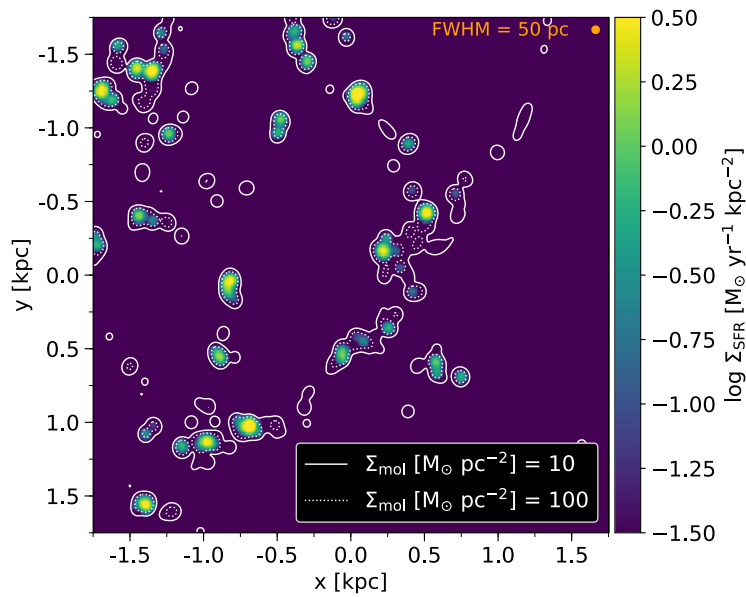
4 *The influence of dust extinction on SFR tracer lifetimes*

$8.6 \times 10^9 M_{\odot}$ and a density profile divided into three regions defined by a constant value of the Toomre Q parameter (Toomre 1964): $Q = 1$ between galactic radii of 2–13 kpc, $Q = 20$ for 0–2 kpc and 13–14 kpc, and beyond 14 kpc is a static, very low density medium. The stellar bulge is not modelled and its effects are included implicitly through the potential. The simulation initially evolves over 730 Myr, wherein the allowed maximum resolution is gradually increased up to the highest resolution of 7.5 pc, and the galaxy settles into an equilibrium state (for details of our refinement strategy, see Fujimoto et al. 2018). We then evolve the simulation further; for the analysis here, we take the 850 Myr snapshot.

Star particles, always with an initial mass of $300 M_{\odot}$, form probabilistically according to a gas density threshold; however, this is only permitted between galactocentric radii of 2–14 kpc. Each star particle represents a simple stellar population, which we model using the SLUG2 stochastic SPS code. The stellar population associated to the star particle is created by sampling a Chabrier (2005) IMF continuously up to a mass of $9 M_{\odot}$ and then stochastically for more massive stars using the POISSON sampling method described in Krumholz et al. (2015, Appendix A). The stellar population evolves during the simulation in accordance with the Padova stellar evolution tracks at solar metallicity (Girardi et al. 2000). Stellar atmospheres are calculated with STARBURST99 spectral synthesis models. The stellar populations associated to the star particles are important for the implemented stellar feedback models, as they determine the ionising luminosity (for photoionisation calculations) as well as the timing of supernova explosions. In order to conserve memory, these stellar populations only remain with the star particle up to the stellar age of 40 Myr.

In Fujimoto et al. (2019), we show that the simulation is consistent with most observational constraints on scales $\gtrsim 100$ pc (e.g. the simulation reproduces the global- and kpc-scale Kennicutt-Schmidt relations for both molecular and total gas; and matches the observed decomposition of the interstellar medium into warm neutral, cold neutral, and molecular phases). However, on smaller ($\lesssim 100$ pc) scales, it can be seen that the implemented feedback mechanisms are insufficient to disperse the surrounding gas. When comparing the synthetic SFR tracer ($H\alpha$) and dense gas ($\text{CO } J = 1 \rightarrow 0$) emission maps of the simulated galaxy, it is clear that most (if not all) $H\alpha$ emitting regions are co-spatial with CO emission. This is demonstrated in Figure 4.1^[p. 103]. Observationally, this would result in a galaxy with a high amount of extinction and a long (overlap) phase where $H\alpha$ and CO are co-spatial. This simulation therefore gives us the opportunity to constrain the emission lifetime of SFR tracers in a system with a high amount of extinction; this complements the work from Chapter 3^[p. 47], which excluded the effects of extinction. This

4.1 Measuring the effects of extinction



A zoom-in of the simulated galaxy described in [Section 4.1.2^{\[p.101\]}](#), showing how SFR tracer emission is always associated with CO emission. The colour map shows the SFR surface density, Σ_{SFR} . The contours show the CO (1 \rightarrow 0) line emission converted into a molecular gas surface density, Σ_{mol} . The maps used to produce this image were convolved using a Gaussian kernel with a FWHM of 50 pc (shown in the top right).

Figure 4.1

Image Credit: Fujimoto et al. (2019, Figure 10), reproduced with permission © OUP

4 The influence of dust extinction on SFR tracer lifetimes

also means that the results we find here should be considered as bounding limits. We expect that real observations would not be subjected to this level of extinction as typically the stellar feedback disrupts the parent cloud in a few Myr (as suggested by the decorrelation of $H\alpha$ and CO on the cloud scale, see Chevance, Kruijssen et al. 2019; Kruijssen, Schrubba, Chevance et al. 2019). Therefore, real observations would have R_{\bullet} values that fall somewhere within the range spanned by the results presented here and $R_{\bullet} = 1$.

4.1.3 Generation of synthetic emission maps

We describe here how we create the synthetic SFR tracer emission maps of the simulated galaxy. In principle, the synthetic emission maps could be created using the SPS associated to each of the star particles within the simulation. For two reasons, we instead choose to run the SLUG2 model as a post-production process. Firstly, for consistency with Chapter 3^[p. 47], we wish to use the Geneva solar-metallicity evolutionary tracks (Schaller et al. 1992) instead of the Padova tracks used within the simulation. Secondly, as described in Section 4.1.2^[p. 101], only star particles up to 40 Myr have an associated stellar population; this would impose an age cut in the emission maps at 40 Myr. The artificial age cut in the emission map would have little impact on the $H\alpha\pm$ filters since their emission lifetimes are short; however, for the WFC3 UVIS1 F336W filter, with $t_{E,0} = 33.3$ Myr (Table 3.3^[p. 65]), this age cut could be problematic.

For the post-production SLUG2 simulations, we use Geneva solar-metallicity evolutionary tracks (Schaller et al. 1992) and STARBURST99 spectral synthesis. For the emission spectrum, we also include the contributions of the surrounding nebular material, which has a Hydrogen number density of 10^2 cm^{-3} and reprocesses 73 per cent of the ionising photons into nebular emission. The 27 per cent ‘loss’ of ionising photons (either absorbed by circumstellar dust, or scattered outside the observational aperture) is consistent with the estimate from McKee et al. (1997). When calculating the extinction, we use the Milky Way extinction curve that comes included with SLUG2.

For consistency, we use the same population that was generated in the simulation, instead of creating a new stellar population for each star particle. This is achieved by extracting the stellar population from the snapshot file in which the star particle first appears (this is always before the stellar population reaches the age of 1 kyr). We do this for all the star particles present in the 850 Myr snapshot which first appear in or after the 730 Myr snapshot. This will introduce a new 120 Myr age cut into the emission map; however, this is sufficiently long as to not impact the results we present here.

4.1 Measuring the effects of extinction

We assign to each star particle a visual extinction (in magnitudes), A_V , which SLUG2 uses to compute an extinguished emission spectrum. To calculate A_V for each star particle we use the ray tracing feature in the Python library YT (Turk et al. 2011). The ‘ray’ passes between the star particle located at $(x_\star, y_\star, z_\star)$ and the edge of the simulated box $(x_\star, y_\star, z_{\text{edge}})$; that is, we are looking at the galaxy face on. For a given cell in the simulation that the ray passes through, i , we recover the gas mass density, $\rho_{\text{g},i}$, as well as the distance travelled through the cell, d_i . From this information we calculate the Hydrogen nuclei number column density, N_{H} , as

$$N_{\text{H}} = \frac{f_{\text{H}} \sum_i \rho_{\text{g},i} d_i}{m_{\text{H}}}, \quad \text{Equation 4.2}$$

where m_{H} is the mass of a Hydrogen nucleus, and $f_{\text{H}} = 0.76$ is the primordial Hydrogen mass fraction. We calculate A_V from N_{H} as (Safranek-Shrader et al. 2017)

$$A_V = N_{\text{H}} \sigma_{\text{d},V}, \quad \text{Equation 4.3}$$

where $\sigma_{\text{d},V} = 5.3 \times 10^{-22} \text{ cm}^2$ (Draine et al. 1996) is the total dust cross-section in the V -band per Hydrogen nucleus.

The full combined photometric spectrum of the stellar population (with and without extinction) is passed through the response profile of the filters of interest (see Table 3.1^[p.50]) to recover a single luminosity for each of filters; this value is what we assign to the star particle. As before, we use the UV response filters that are included by default in SLUG2 (Krumholz et al. 2015, for more details see). We define the narrow band $\text{H}\alpha$ filters that are without continuum subtraction ($\text{H}\alpha+W$) in the same way as in Chapter 3^[p.47] (Equation 3.3^[p.58]). For $\text{H}\alpha$ with continuum subtraction ($\text{H}\alpha-$), we convert the Hydrogen-ionizing photon emission into a true $\text{H}\alpha$ luminosity using Silva et al. (2014, Equation 2).²

In Figure 4.2^[p.106] we show the distribution of column densities (and visual extinction) that is associated to the star particles used in creating the synthetic emission maps (i.e. stars surviving the 120 Myr age cut). In the figure, we also mark a ‘galactic column density’ found to be $16.7 \text{ M}_{\odot} \text{ pc}^{-2}$ (corresponding to $A_V = 0.84 \text{ mag}$). This is calculated from the average gas surface density of the 850 Myr snapshot (without an age cut), $\bar{\Sigma}_{\text{g}}$, between the galactocentric radii 3–11 kpc by using Equation 4.2 with $\bar{\Sigma}_{\text{g}}$ in place of the summation term.

²In Chapter 3^[p.47], we (correctly) commented that this conversion was not necessary; however, in order to calculate extinguished $\text{H}\alpha$ emission we must explicitly calculate the $\text{H}\alpha$ luminosity.

4 The influence of dust extinction on SFR tracer lifetimes

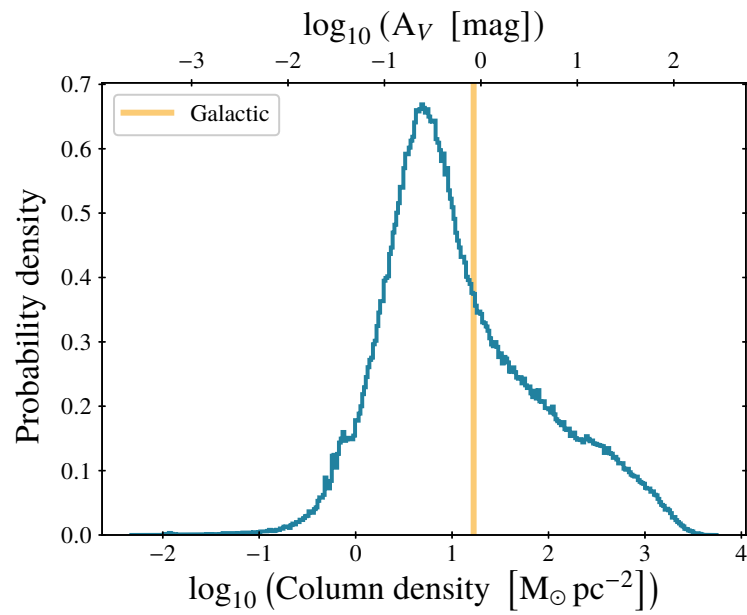
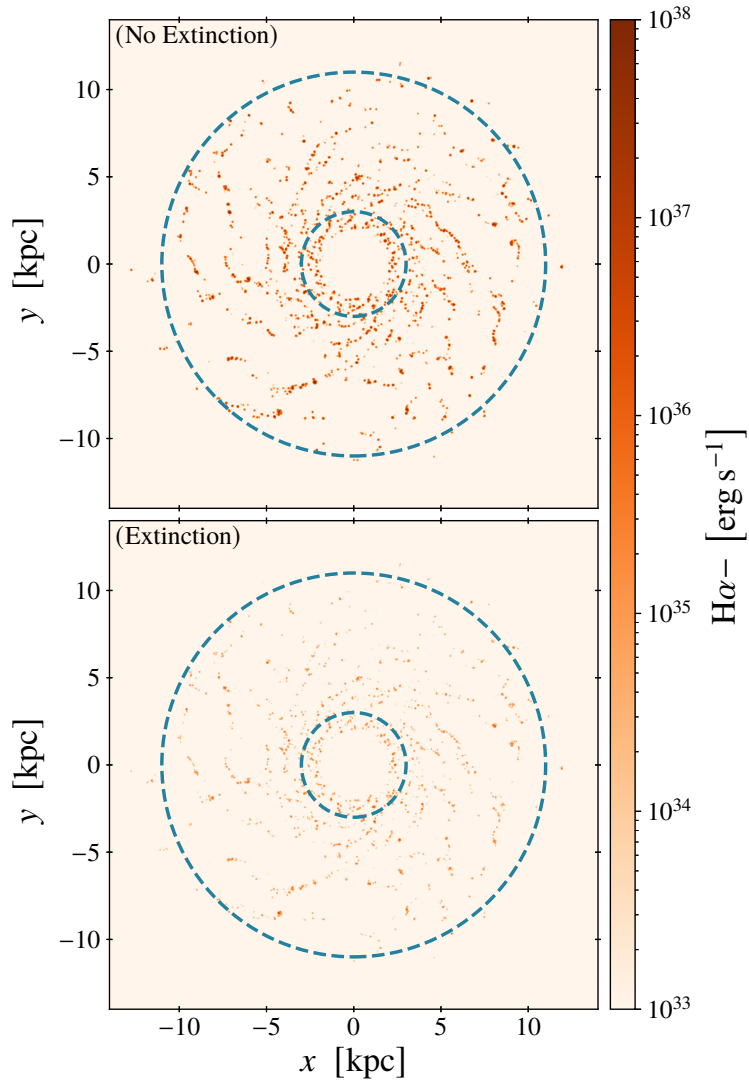


Figure 4.2 A histogram showing the distribution of column densities and visual extinctions (calculated using Equations 4.2 and 4.3^[p. 105]) associated to the star particles used in creating the synthetic emission maps. The vertical line indicates the galactic average column density ($16.7 M_{\odot} \text{ pc}^{-2}$) / visual extinction (0.84 mag). This is calculated from the average gas surface density between the galactocentric radii 3–11 kpc (as marked in Figure 4.3^[p. 107]).

4.1 Measuring the effects of extinction



A comparison between synthetic emission maps without (top) and with (bottom) extinction for continuum subtracted H α emission (H α -). The blue dashed annuli, with inner radius of 3 kpc and outer radius of 11 kpc, indicate the region of the galaxy we use for our analysis.

Figure 4.3

4 The influence of dust extinction on SFR tracer lifetimes

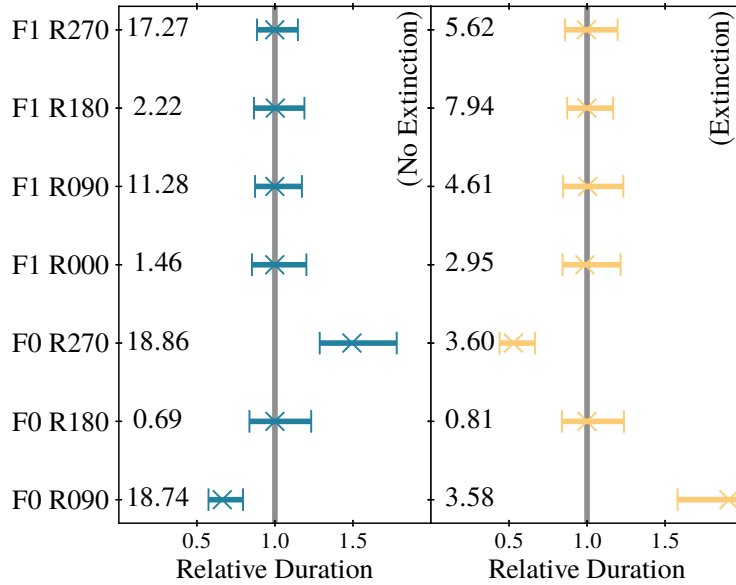
In [Figure 4.3](#)^[p. 107], we present example synthetic H α – emission maps of the simulated galaxy, which are used in our analysis. The two presented maps allow us to see how extinction reduces the amount of emission present in the image. The total emission in the extinguished map is 2.1 per cent of the emission in the unextinguished map. The figure also indicates the region of the galaxy we will use in our analysis with the HEISENBERG code. We exclude the inner 3 kpc and beyond 11 kpc of the galaxy due to the way in which the galaxy simulation was set-up (see [Section 4.1.2](#)^[p. 101]).

The two groups of emission maps we produce here are related to each other. Excluding the effects of extinction, they are in fact identical; that is, all emitting regions are present in both maps. This violates the first requirement listed in [Kruijssen et al. \(2018, Section 4.4\)](#), which details the conditions under which HEISENBERG should be used. The HEISENBERG code assumes that the pair of maps analysed traces two different phases of an underlying evolutionary process. With this assumption, it is expected that within the galaxy many emission peaks in one emission map will be independent of the other and only some regions will be present in both maps (the ‘overlap’ phase). We circumvent this aspect of the method by spatially transforming (through rotation and/or reflection) one emission map relative to the other. This transformation will spatially offset the emitting regions and give the impression of an evolutionary sequence. In performing the transformation, we forfeit information about two quantities constrained by HEISENBERG (the spatial separation of regions and the duration of the overlap phase) whilst retaining the ability to measure the relative duration of one phase to the other: this is unaffected by the relative positions of regions ([Kruijssen et al. 2014, 2018](#), see the discussion in).

Relying on the fact that the relative duration of an emission map to itself is always unity, we can test to find which transformation works best given the specific structure and morphology of the adopted galaxy simulation. Specifically, we test for which transformation HEISENBERG returns a relative duration consistent with unity and a good model fit, with reduced- $\chi^2 \approx 1$, when applied to a given map and the spatially transformation version of itself. We consider either no reflection ‘F0’ or a left-right reflection ‘F1’ and four rotations $\{0^\circ, 90^\circ, 180^\circ, 270^\circ\}$ denoted as ‘RX’, where X is the angle of rotation. This gives the seven possible unique transformations (excluding no transformation, ‘F0 R0’).

[Figure 4.4](#)^[p. 109] shows the relative durations we measure using HEISENBERG when applying each of the seven spatial transformations to the H α – emission maps. For all of the SFR tracers we consider, we find that using a left-right reflection always gives the expected relative duration of unity. Of the ‘F1’ series, the best reduced- χ^2 is for zero rotation. By contrast, a purely rotational

4.1 Measuring the effects of extinction



Results showing the measured relative duration of an emission map when compared to a spatially transformed version of itself. Left: $H\alpha$ - emission map without extinction. Right: $H\alpha$ - emission map with extinction. The vertical axis labels, 'FX RY', indicate the spatial transformation applied to one of the two input maps: X indicates if a left-right reflection was used, and Y gives the angle of rotation. Note that 'F0 R0' (no transformation) is excluded as it violates the conditions given in Kruijssen et al. (2018, Section 4.4). The relative duration of an emission map to itself should be unity and so we have marked this value with a vertical grey line. The reduced- χ^2 value of the model fit is also reported alongside each measurement of the relative duration. The optimal reduced- χ^2 value is unity. The error bars indicate the uncertainties on the measurements derived from the χ^2 fit carried out by HEISENBERG.

Figure 4.4

4 The influence of dust extinction on SFR tracer lifetimes

transformation without reflection ('F0') does not always return a relative duration of unity. This is likely caused by the symmetric nature of the spiral arms, which causes large parts of them to overlap at rotations of 90° and 270° . In all the experiments that follow, we will always apply the 'F1 R0' transformation to the extinguished emission map.

4.2 Extinction

By using the HEISENBERG code on two emission maps, we quantify the relative duration of one map to the other. For example, if we have a gas tracer map (e.g. CO) and an SFR tracer map (e.g. $H\alpha$), we can measure the relative duration of the CO map to the $H\alpha$ map. With the known lifetime of $H\alpha$ emission, we can then recover an absolute duration for the CO emission. However, extinction reduces the overall emission within an emission map and possibly reduces the number of identifiable star-forming regions. The relative duration of the same CO emission map to an *extinguished* $H\alpha$ map will differ and so should the emission lifetime we associate to the extinguished $H\alpha$ emission map. Here we quantify by how much the emission lifetimes we associate to various SFR tracers differ when dealing with extinguished emission. As previously mentioned, the galaxy we are using for this analysis has heavily extinguished emission due to inefficient feedback (Fujimoto et al. 2019): most, if not all, star forming regions are still embedded in gas. This means that the differences in emission lifetimes we would expect for real galaxies would be less extreme than those we find here: the results here represent a limiting case, illustrating the maximum effect induced by extinction.

We provide the HEISENBERG code with an emission map and a spatially transformed ('F1 R0') extinguished emission map and recover R_\bullet , the relative duration of the extinguished emission map compared to the unextinguished map. This is the factor by which the emission lifetime of the SFR tracer changes as a result of extinction. In order to see how the emission lifetimes change as a function of extinction we repeat the experiment with different extinguished emission maps. We produce these different extinguished emission maps by modifying Equation 4.2^[p. 105] to include a density scaling factor, F_d , such that $N'_H \equiv F_d N_H$. This scaling factor multiplies the gas density used to calculate the A_V associated to the star particle, which SLUG2 subsequently uses to calculate the extinguished emission. We choose scaling factors in the range $[10^{-3}, 10^1]$ in 0.25 logarithmic steps. The native simulation column density is at $F_d = 1$.

In this work, we only consider solar metallicity. Any change in metallicity is expected to linearly translate into an analogous change of the dust column density, such that the metallicity change can be absorbed into the column

density scaling factor, F_d , defined here. As such, the dependence on column density quantified here also reflects the expected dependence on metallicity at fixed gas surface density.

Likewise, our analysis of the synthetic emission maps always assumes a face-on orientation. However, our results can still be extended to observations of galaxies that are not face on. The observed average gas column density increases as the inclination angle at which a galaxy is viewed increases. In order to correct for an inclination angle i , this means that the column density scaling factor F_d should be scaled by a factor of $(\cos i)^{-1/2}$. In doing so, we note that there are limits on the inclination angle for which HEISENBERG still provides reliable measurements (ideally $i \lesssim 70^\circ$, see Kruijssen et al. 2018, Section 4.3.7)).

In Figures 4.5 to 4.7^[pp. 112–114] we show how R_\bullet changes with density scaling factor, F_d , for the H α –, H α +, and UV filters. Most importantly, we find that extinction does not affect the H α tracer emission lifetimes at gas surface densities $\Sigma_{\text{gas}} \lesssim 20 \text{ M}_\odot \text{ pc}^{-2}$. The UV tracer emission lifetimes start to become affected at gas surface densities $1 \text{ M}_\odot \text{ pc}^{-2} \lesssim \Sigma_{\text{gas}} \lesssim 10 \text{ M}_\odot \text{ pc}^{-2}$. At the extremes of F_d (or equivalently, column density), the observed behaviour is as expected. As the column density tends towards zero, R_\bullet tends towards unity. This is because, as the column density decreases, the amount of extinction is reduced, which results in an extinguished emission map more like the emission map without extinction. As the column density increases, the amount of extinction increases and R_\bullet tends towards zero as more emission is removed from the map. In practice, the results from HEISENBERG become unreliable once fewer than 35 regions are identified and so R_\bullet will instead tend towards some small non-zero value. The grey-shaded data points in Figure 4.7^[p. 114] (not present in Figures 4.5 and 4.6^[pp. 112 & 113]) indicate these unreliable results, which are typically found for $F_d \gtrsim 10$ (or $\Sigma_{\text{gas}} \gtrsim 200 \text{ M}_\odot \text{ pc}^{-2}$).

One might expect that the value of R_\bullet should always be in the range $[0, 1]$ (i.e. a shorter emission lifetime), as extinction removes emission from the image. However, the results in Figures 4.5 to 4.7^[pp. 112–114] show that for $F_d \lesssim 1$ (or $\Sigma_{\text{gas}} \lesssim 10 \text{ M}_\odot \text{ pc}^{-2}$), R_\bullet typically falls in the range $[1, 1.5]$. We use the results in Figure 4.8^[p. 115] to explain this behaviour. In the top panel of Figure 4.8^[p. 115], we show the upper end of the emission PDF. The area under this curve gives the probability of any point in the emission map having a given amount of emission. We can see that as F_d increases the distribution moves towards the left, i.e. lower emission values become more likely. This shows that extinction removes emission from the map as a whole. The overall shift in the emission distribution has minimal effect on the results from HEISENBERG: absolute emission values are not relevant to the measurement of R_\bullet . The

4 The influence of dust extinction on SFR tracer lifetimes

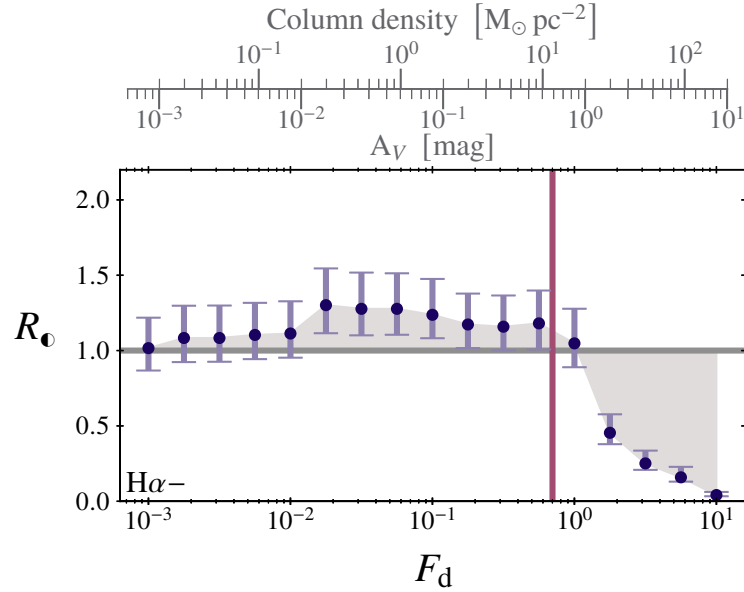
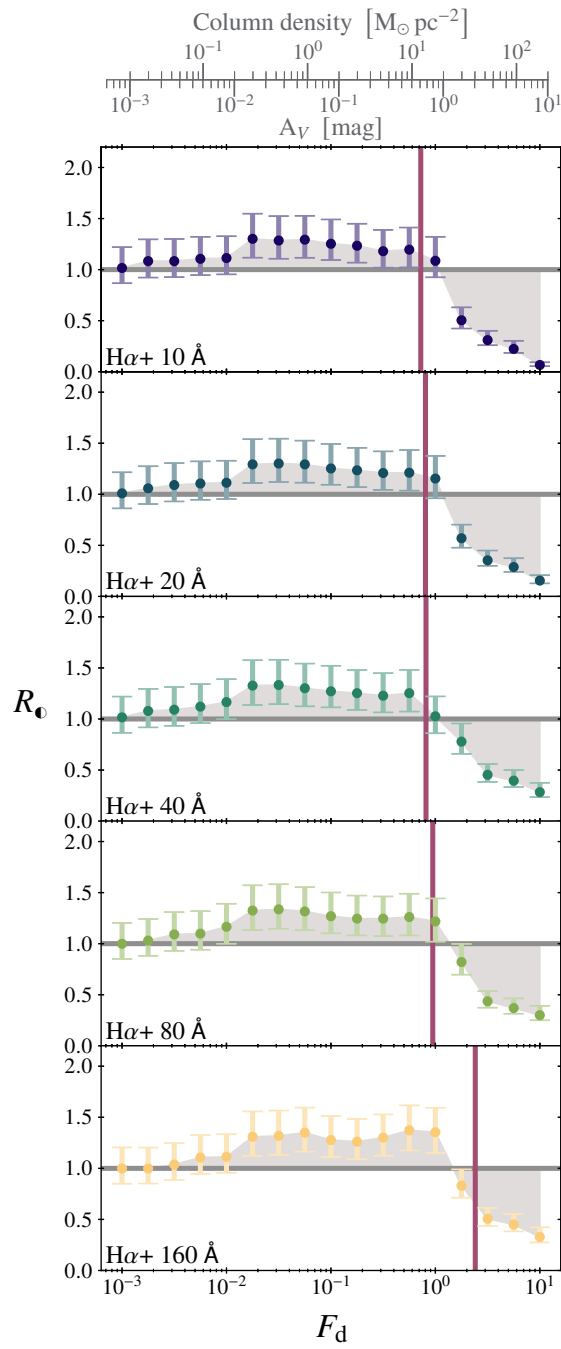


Figure 4.5

Relative amount by which the continuum-subtracted H α (H α -) emission lifetime changes due to extinction. Shown is the relative duration associated to the extinguished compared to the unextinguished SFR tracer emission map, R_\bullet , as a function of the column density scaling factor, F_d . A scale factor of unity indicates column densities taken directly from the simulation. The horizontal grey line (at $R_\bullet = 1$) indicates the point where the duration associated to each of the two emission maps is the same. The vertical red line indicates the approximate location where we expect R_\bullet to transition from $R_\bullet > 1$ to $R_\bullet < 1$ based on the flux density distribution of the maps (see the text and [Figure 4.8](#)^[p. 115] for details). The top grey axis converts F_d into a column density and visual extinction A_V based on a column density equal to the average gas surface density, $\bar{\Sigma}_g = 16.7 M_\odot \text{pc}^{-2}$. Since our experiments probe the maximum effects of extinction, we expect to find R_\bullet for real galaxies to lie within the grey-shaded region (i.e. between the unity line and the data points shown).

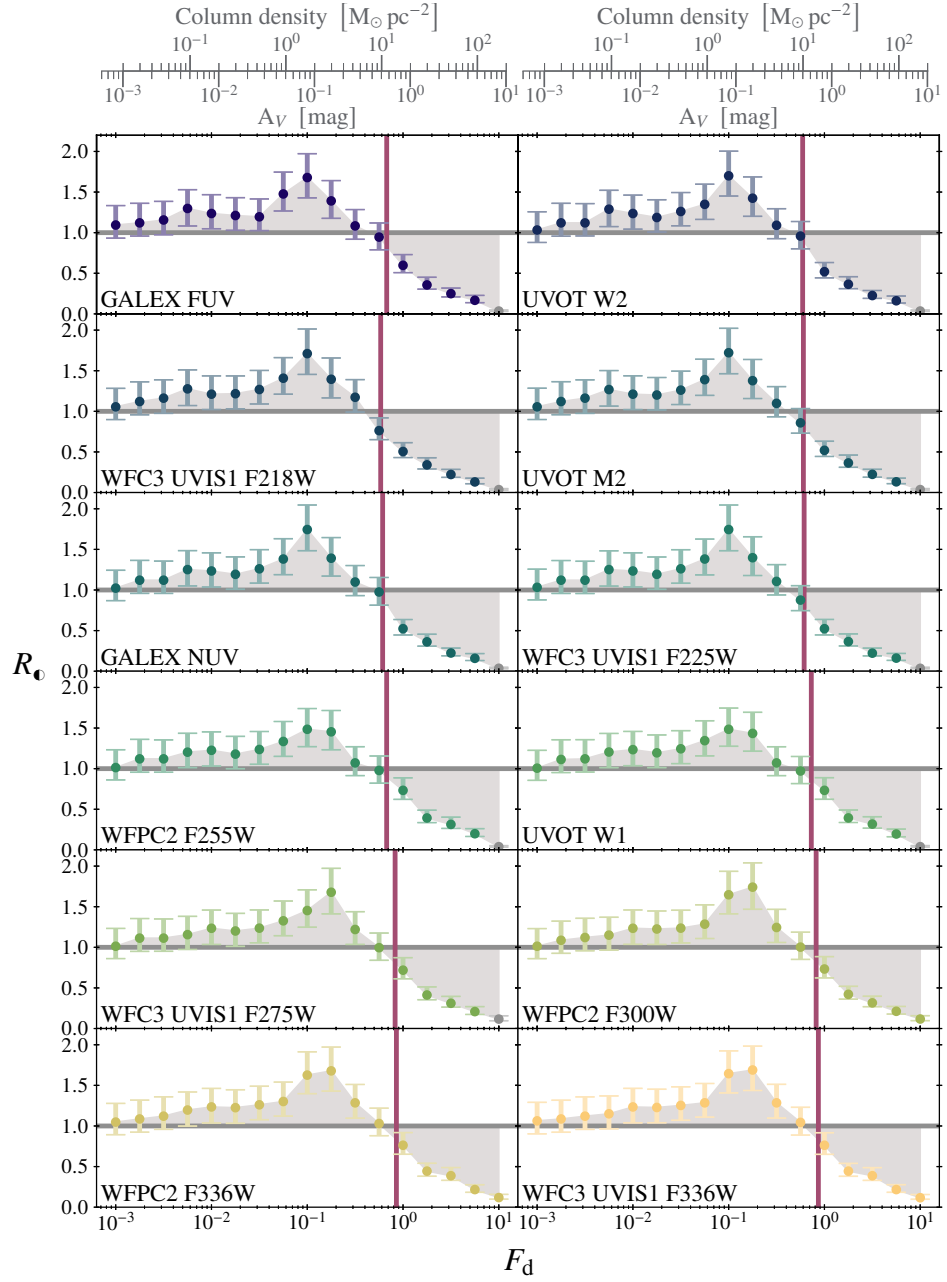
4.2 Extinction



Same as Figure 4.5^[p.112] for $H\alpha$ filters without continuum subtraction ($H\alpha+$). The filter width is indicated in the bottom left corner of each panel.

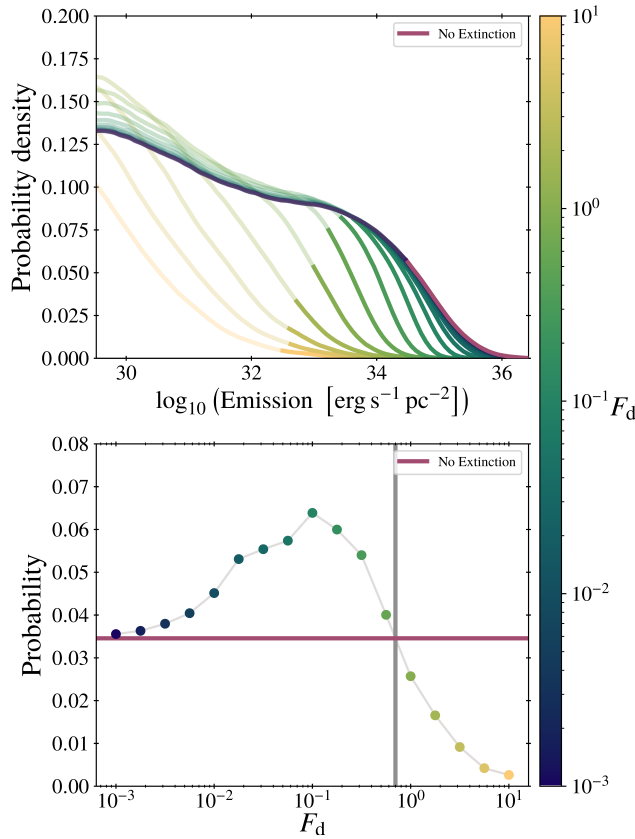
Figure 4.6

4 The influence of dust extinction on SFR tracer lifetimes



Same as Figure 4.5^[p. 112] for UV emission filters. The panels are ordered from left to right, top to bottom, by the response-weighted mean wavelength. The filter name is indicated in the bottom left corner of each panel. The grey-shaded data points have fewer than 35 identified emission peaks in the extinguished emission image and are therefore unreliable (see Kruijssen et al. 2018, Section 4.4).

4.2 Extinction



Top: The high-emission end of the $\text{H}\alpha$ – PDF of emission in emission maps with different levels of extinction. The colour of the line indicates the density scaling factor, F_d , used in creating the emission map. Higher values of F_d correspond to a higher amount of extinction. The native amount of extinction occurs at $F_d = 1$ (or $\Sigma_{\text{gas}} = 16.7 \text{ M}_{\odot} \text{ pc}^{-2}$). The red line indicates the results for the emission map without extinction. The non-transparent line segments indicate the range of the distribution that is within $\Delta = 2 \text{ dex}$ (see Section 4.1.1^[p. 100]) of the maximum. Bottom: The probability (or fraction) of emission falling within Δ of the maximum (i.e. area under the non-transparent line segments in the top panel). The vertical grey line indicates the approximate F_d value for the transition between $P_{\bullet} > P_{\circ}$ and $P_{\bullet} < P_{\circ}$, where P_{\bullet} is the probability of finding emission within Δ of the maximum in the emission map with extinction and P_{\circ} without (shown as the horizontal red line). We see that, at low levels of extinction ($F_d \lesssim 1$), a larger fraction of the emission resides in the top Δ of the distribution, as a result of only the brightest emission peaks being affected; that is, those peaks which formed at the highest gas column densities. By contrast, at high levels of extinction ($F_d \gtrsim 1$), all emission peaks are affected (resulting in the suppression of the ‘knee’ in the distribution in the top panel) and a smaller fraction of emission resides in the top Δ of the distribution.

Figure 4.8

4 The influence of dust extinction on SFR tracer lifetimes

HEISENBERG code identifies emission peaks based on contours that start from the maximum emission value in the image and go down for a certain maximal range, $\Delta = 2$ dex (see Section 4.1.1^[p. 100] for details). This means that the fraction of the total emission contained in this interval Δ is what impacts R_{\bullet} the most.

The non-transparent line segments in Figure 4.8^[p. 115] show the part of the distribution within Δ of the maximum emission value present in the emission map. It is clear that extinction alters the shape of the emission PDF, but what is of particular note is the gradients of these Δ line segments. When extinction is initially introduced (small F_d), the gradients of these segments become steeper than that without extinction. This shows that extinction has preferentially removed emission from the highest emission peaks, because the brightest emission peaks form at the highest gas column densities. Since the gradients are steeper, the area under the curve is larger (for the same change in emission, i.e. Δ) and, as shown in the bottom panel of Figure 4.8^[p. 115], the probability of finding emission within Δ of the maximum for the extinguished emission map (P_{\bullet}) is higher than for the unextinguished emission map (P_{\circ}). In constraining R_{\bullet} , HEISENBERG will then detect more emission peaks within Δ of the maximum compared to the unextinguished image. As a result, the duration associated with the extinguished emission map is likely to be longer, thus causing $R_{\bullet} > 1$. Using linear interpolation, we find the F_d that corresponds to $P_{\bullet} = P_{\circ}$ for each of the SFR tracers and mark this transition point in Figures 4.5 to 4.7^[pp. 112–114] (with a vertical red line). Indeed, we see that for the vast majority of SFR tracers, and within the uncertainties of the measurements, $R_{\bullet} \gtrsim 1$ before this transition line and $R_{\bullet} \lesssim 1$ after, confirming our understanding of the behaviour seen in Figures 4.5 to 4.7^[pp. 112–114].

In relation to observational applications, the results presented here represent the maximum extent by which the SFR reference timescale can be altered by extinction. With future simulations, that have feedback mechanisms to remove gas more effectively from the stellar birth environments, the values of R_{\bullet} that we measure will become more representative of real observations. However, different galaxies suffer from different levels of extinction, implying that deriving a single, uniform calibration of the reference timescale under the influence of extinction remains challenging. If observed SFR tracer maps cannot be corrected for extinction, our results provide the extreme factor relative to unity by which the SFR tracer emission timescale should be corrected.

For our simulated galaxy, we find that R_{\bullet} and extinction do not have a simple monotonic relation. Extinction can both increase and decrease the duration of the SFR tracer emission timescale. When comparing the results for the $H\alpha$ filters Figures 4.5 and 4.6^[pp. 112 & 113] and UV filters Figure 4.7^[p. 114], we

can see that the extremes in R_{\bullet} are greater for UV than they are for $H\alpha$. This means that UV filters are affected by extinction more strongly than $H\alpha$. For real galaxies, which suffer from less extinction than permitted by the inefficient feedback model used in the simulation, not all star forming regions are still embedded in gas and R_{\bullet} must deviate from unity less than what we find here.

4.3 Sensitivity limit

Real-Universe observations are affected by sensitivity limits, which result from a combination of the telescope’s intrinsic sensitivity and as part of the post-processing pipeline. The emission maps we have created so far recover all emission and do not have a sensitivity limit. Since extinction reduces the overall emission in an emission map, a sensitivity limit will more greatly affect the extinguished emission maps compared to the unextinguished emission maps. This will likely lead to lower SFR tracer emission lifetimes (or values of R_{\bullet}).

To investigate how R_{\bullet} changes with the sensitivity limit, we create new emission maps (both extinguished and unextinguished) by applying a flux density cut to our existing ($F_d = 1$) emission maps. The flux density cut is applied by setting to zero all the pixels in the emission map that are below the sensitivity limit. For each emission filter, the sensitivity limits are chosen as $\left[10^{(E_{\circ}^{\max}-5)}, 10^{(E_{\circ}^{\max}-2)}\right]$ in 0.25 logarithmic steps, where E_{\circ}^{\max} is the decadic logarithm of the maximum emission in the *unextinguished* emission map. Choosing more extreme limits would be of little use. A larger upper limit would reduce the amount of emission within the extinguished map to the extent that we could no longer use HEISENBERG reliably, due to fewer than 35 regions being identified (see Kruijssen et al. 2018, Section 4.4). A smaller lower limit would not add any new information, since R_{\bullet} will saturate at the values seen at $F_d = 1$ in Figures 4.5 to 4.7^[pp. 112–114]. When running these experiments, we adopt the same sensitivity limits for both the extinguished and unextinguished emission maps.

In Figures 4.9 to 4.11^[pp. 118–120], we show how R_{\bullet} changes with sensitivity limit, for the $H\alpha-$, $H\alpha+$, and UV filters. The figures show the behaviour that we expected: with a high sensitivity limit the value of R_{\bullet} is low (~ 0.1 – 0.2) and increases (up to the saturation level) as the sensitivity limit decreases. The saturation level (the R_{\bullet} values seen for $F_d = 1$ in Section 4.2^[p. 110]) is recovered for sensitivity limits up to the flux level down to which HEISENBERG identifies emission peaks (about 10^{33} erg s⁻¹ pc⁻² for $H\alpha-$ at the native column density of the simulation, i.e. $F_d = 1$). We also see that sensitivity limits which are too high ($\gtrsim 10^{34}$ erg s⁻¹ pc⁻² for $H\alpha-$) remove too much emission and leave

4 The influence of dust extinction on SFR tracer lifetimes

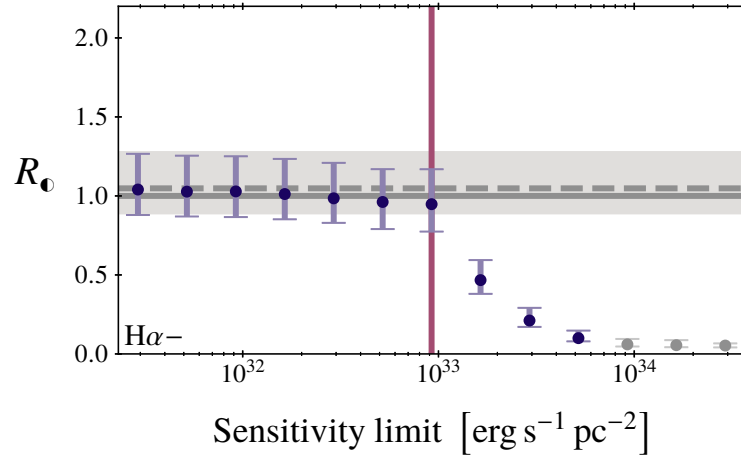
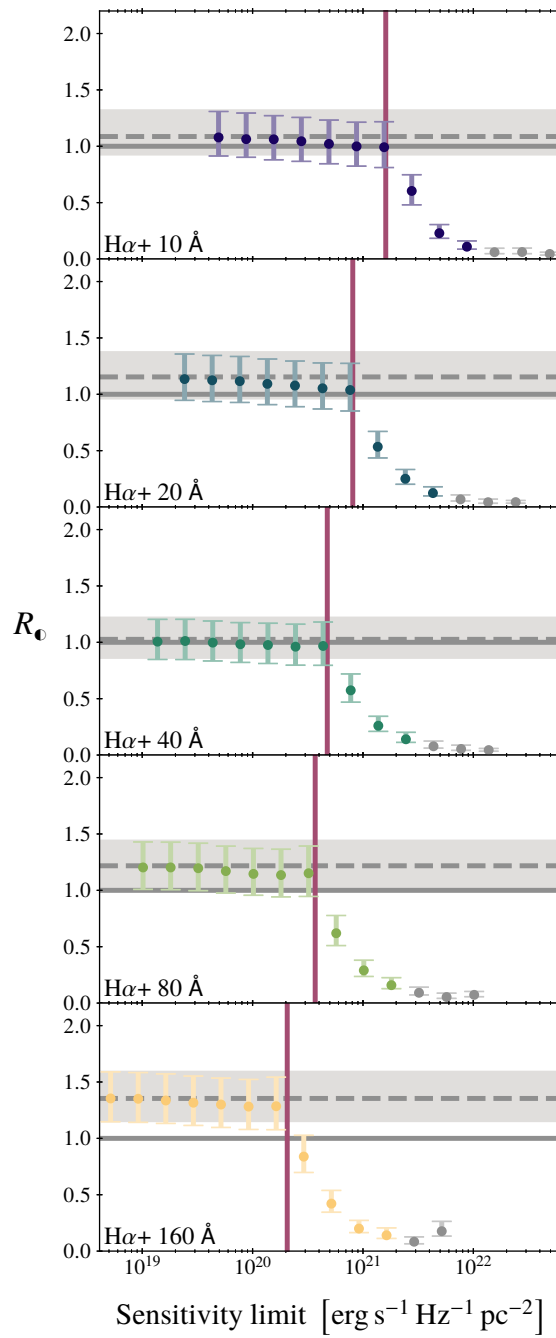


Figure 4.9

Relative amount by which the continuum-subtracted H α (H α -) emission lifetime changes due to extinction and a non-zero sensitivity limit. Shown is the relative duration associated to the extinguished ($F_d = 1$, see [Section 4.2](#)^[p. 110]) compared to the unextinguished SFR tracer emission map, R_\bullet , as a function of the sensitivity limit. The solid horizontal grey line (at $R_\bullet = 1$) indicates the point where the timescale associated to each of the two emission maps is the same. The dashed horizontal grey line indicates the value of R_\bullet without a sensitivity limit applied (see [Figure 4.5](#)^[p. 112]), representing the expected saturation level. The grey-shaded region marks the uncertainty on this value. The vertical red line indicates the lowest contour level used by HEISENBERG to identify emission peaks. The grey-shaded data points have fewer than 35 identified emission peaks in the extinguished emission image and are therefore unreliable (see Kruijssen et al. [2018](#), Section 4.4).

4.3 Sensitivity limit



Same as Figure 4.9^[p. 118] for $H\alpha$ filters without continuum subtraction ($H\alpha+$). The filter width is indicated in the bottom left corner of each panel.

Figure 4.10

4 The influence of dust extinction on SFR tracer lifetimes

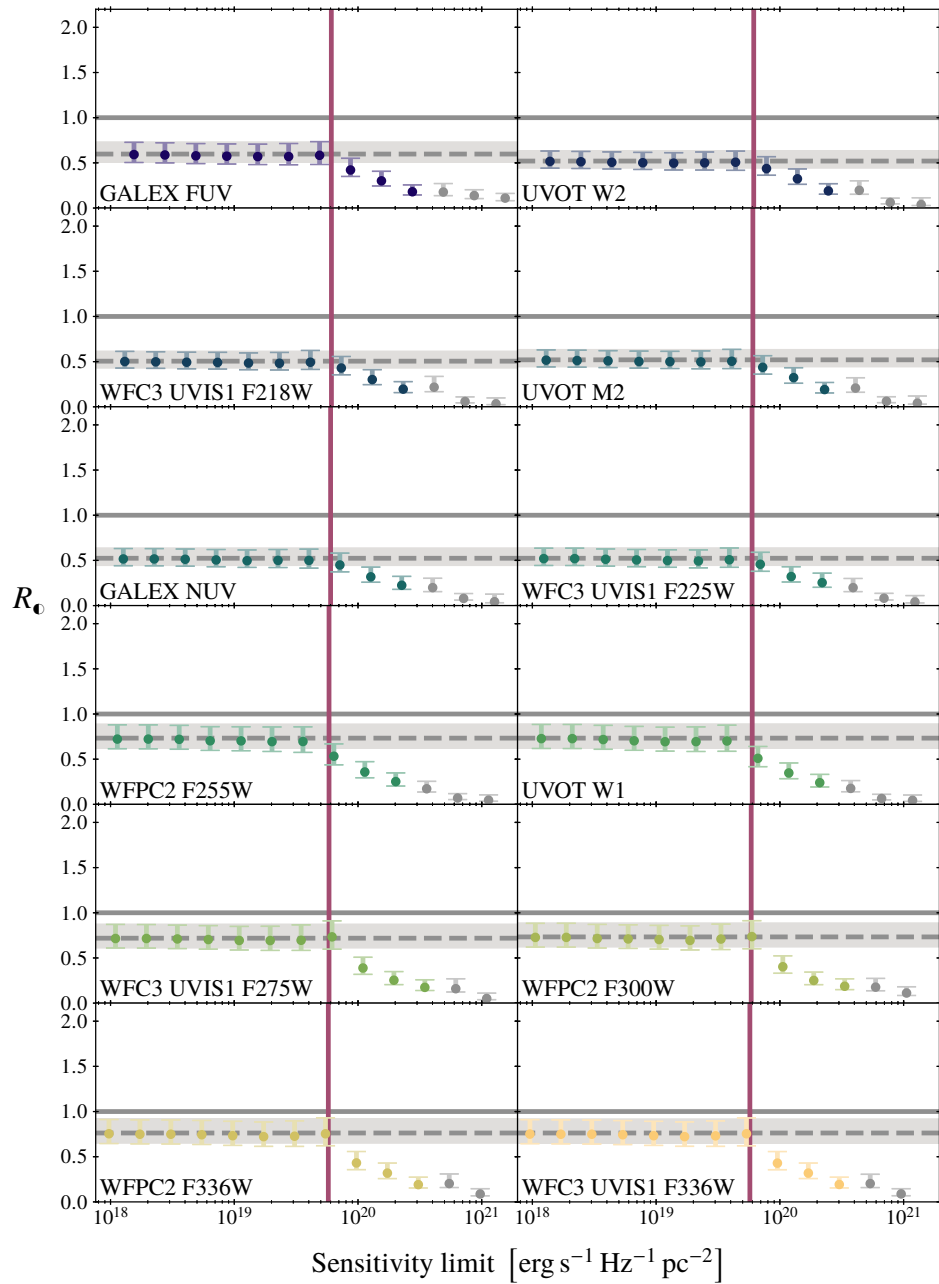


Figure 4.11 Same as Figure 4.9^[p. 118] for UV emission filters. The panels are ordered from left to right, top to bottom, by the filter-weighted mean wavelength. The filter name is indicated in the bottom left corner of each panel.

an insufficient number of peaks for the application of our analysis (grey data points).

The reported values of R_{\bullet} , as a function of the sensitivity limit, are dependent on the results in the previous section, where we varied the column density scaling factor. If we had chosen to use a different density scaling factor, F_d , these changes would also manifest themselves here. The most obvious way in which F_d affects the behaviour of R_{\bullet} as a function of the sensitivity limit is the value at which the measurements would saturate, because this simply matches the corresponding F_d measurement in [Section 4.2](#)^[p. 110]. The sensitivity limit at which the saturation level is reached is less obvious. In [Figures 4.9 to 4.11](#)^[pp. 118–120], we have included a line marking $10^{(E_{\bullet}^{\max} - \Delta)}$ where E_{\bullet}^{\max} is the decadic logarithm of the maximum emission in the *extincted* emission map, and $\Delta = 2$ (see [Section 4.1.1](#)^[p. 100] for details). This line is the lowest contour level used in our experiments by HEISENBERG to identify emission peaks (across both input maps). The lowest contour level will become lower as F_d becomes higher (corresponding to more extinction). We expect, and observe, that R_{\bullet} saturates if the sensitivity limit is lower than the lowest contour level used by HEISENBERG, allowing the same number of peaks to be identified. Above this level, emission is being removed that could be identified as emission peaks by HEISENBERG and so reduces the apparent duration of the emission map.

In summary, the results of this section show that if the sensitivity limit of the observations is lower than the lowest flux level that would be used by HEISENBERG to identify emission peaks (typically chosen to be $\Delta = 2$ dex below the maximum; although, this is not always the case: for example values used in observations see [Chevance, Kruijssen et al. 2019](#); [Hygate et al. 2019](#); [Kruijssen, Schrubba, Chevance et al. 2019](#)), R_{\bullet} will be unchanged relative to what would be measured without a sensitivity limit. Above this limit, the measured SFR tracer emission timescale is reduced: the sensitivity limit further compounds the effects of extinction.

4.4 Conclusions

Using a KL14 principle, we determine how the apparent emission lifetime of SFR tracers changes as a result of dust extinction. In the previous chapter, [Chapter 3](#)^[p. 47], we characterised how the durations over which H α and UV emission emerges from coeval stellar populations changes as a function of the metallicity and the SFR surface density (to emulate incomplete IMF sampling); however, we excluded the effects of extinction. In this chapter, we include the effects of extinction. Extinction does not change the underlying

4 *The influence of dust extinction on SFR tracer lifetimes*

emission lifetime of the emission, but we demonstrate that it changes the flux density distribution of the observed emission. As a result, the apparent duration of the emission timescale measured with the HEISENBERG code is changed and therefore the reference timescale required for calibrating the evolutionary timeline of observational applications of the HEISENBERG code. We quantify this change through R_{\bullet} , the relative duration associated to a SFR tracer emission map affected by extinction relative to the duration associated to the SFR tracer emission map without extinction.

For the analysis in this chapter, we used a Milky-Way-like disc galaxy simulation (Fujimoto et al. 2018, 2019) from which we generate synthetic SFR tracer emission maps. This ENZO simulation has the advantage that on large scales ($\gtrsim 100$ pc) it agrees with observational constraints (e.g. the global- and kpc-scale Kennicutt-Schmidt relations and the phases of the interstellar medium); however, as we demonstrate in Fujimoto et al. (2019), the implemented feedback mechanisms are insufficient to disperse molecular clouds surrounding sites of star formation, in clear contradiction with observations (e.g. Chevance, Kruijssen et al. 2019; Kruijssen, Schrubba, Chevance et al. 2019). This means that this simulation acts as a limiting case of the effects of extinction: the magnitude of the effects we find in this chapter represent the most extreme case. In real-Universe observational studies, the effects of extinction must be less. We are therefore able to use the results here to indicate when extinction could have an impact on our measurements for real observations and the maximum extent that they could be affected. For conditions where we find that extinction does not affect the SFR tracer reference timescale, observational applications are firmly ruled out to be affected by extinction.

We produce synthetic emission maps of 18 SFR tracer filters: 12 different UV filters, 5 $H\alpha$ filters with continuum ($H\alpha+$) and a continuum subtracted $H\alpha$ filter ($H\alpha-$). To generate our synthetic emission maps, we use the properties of the stars within the simulation along with the stochastic SPS code SLUG2 to determine the emission spectrum associated to each of the stars present. The amount of extinction affecting each star comes from the properties of the gas present within the simulation that falls within the line of sight of the star. This information, along with a Milky Way extinction curve, is used by SLUG2 to produce the extinguished synthetic emission maps. In this work we only consider solar metallicity and a fully sampled IMF (see below).

Our experiments show that extinction affects R_{\bullet} non-monotonically. At low levels of extinction, corresponding to mean galactic surface densities $\Sigma_{\text{gas}} \lesssim 20 M_{\odot} \text{pc}^{-2}$, R_{\bullet} is close to but slightly larger than unity. This is because only light from the brightest star-forming regions, which are in

the regions of highest extinction, suffers any significant attenuation. This compresses the range of fluxes present in the extinguished map relative to the unextinguished one, increasing the total number of distinct regions within a fixed dynamic range. This effect is interpreted by HEISENBERG’s statistical analysis as a slight increase in the duration of the extinguished map compared to the unextinguished one. As the extinction increases and most star-forming regions begin to suffer significant attenuation, the effect reverses and R_{\bullet} becomes smaller than unity by as much as a factor of ten in the most extreme case.

We also investigate how a sensitivity limit on our observations could impact R_{\bullet} . The critical value at which the sensitivity limit has an impact on R_{\bullet} depends on the lowest flux level used by HEISENBERG to identify emission peaks, which is typically chosen to be 2 dex below the maximum flux level in the emission map (see e.g. Kruijssen et al. 2018; Chevance, Kruijssen et al. 2019; Kruijssen, Schrubba, Chevance et al. 2019). If the sensitivity limit is lower than this threshold, R_{\bullet} is unchanged from the results we recover if there was no sensitivity limit. For sensitivity limits above the threshold, R_{\bullet} can decrease to as low as ~ 0.1 – 0.2 .

In summary, we have measured the relative change in the emission lifetime for SFR tracers ($H\alpha$ and UV) as a result of extinction and how this might be compounded by the non-zero sensitivity limits of observational data. We find that extinction does not affect the SFR tracer emission timescales at gas surface densities $\Sigma_{\text{gas}} \lesssim 20 \text{ M}_{\odot} \text{ pc}^{-2}$. At higher surface densities, the SFR tracer emission timescales may be reduced (relative to the values presented in Chapter 3^[p. 47]) by a factor as low as $R_{\bullet} = 0.1$. However, the values presented in this work represent extreme limits. The ineffectiveness of stellar feedback at dispersing molecular clouds in the simulation used here has enabled us to set limits on the impact of feedback on SFR tracer timescales. The resulting correction factors are critical for informing observational efforts characterising the molecular cloud life cycle using the HEISENBERG code. Future calibrations using simulations with improved feedback physics will be able to further improve the limits provided here.

Acknowledgements

This work was aided by the Australia-Germany Joint Research Cooperation Scheme (UA-DAAD, grant number 57387355). This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI Australia), an NCRIS enabled capability supported by the Australian Government.

The synthetic ALMA ultra-deep field

In this chapter, we change our focus away from SFR tracer emission and towards CO emission. However, our aim to improve our understanding of the processes governing star formation has not altered. Instead of looking at the emission produced by stars, we consider the emission that traces dense molecular gas in the ISM. The ISM plays a fundamental role in star formation, as this is the material from which stars are born. Through the life cycle of a star, the star returns material and energy back to the ISM through feedback mechanisms such as photoionisation, stellar winds, and supernovae. This exchange of material and energy acts to link cloud scale physics to the galactic scales.

Typically the study of the ISM is performed galaxy by galaxy, meaning that there are few sky surveys tracing the gas content of a population of galaxies. With such surveys, it would be possible to quantify the molecular gas over cosmic time. The results would also be free from the biases imparted by pre-selection criteria (as is the case for targeted observations). Now that observational facilities such as ALMA are available, sky surveys of the molecular gas are now possible. One such survey is ASPECS: an ALMA survey focusing on the Hubble ultra-deep field (UDF). One of the goals of ASPECS is the study of the molecular Hydrogen content of galaxies, using CO emission as a proxy. The pilot phase, covering a region of $\sim 1 \text{ arcmin}^2$, has now been completed (Decarli et al. 2016; Walter et al. 2016). The ASPECS Large Program extends this region to cover $\sim 4.6 \text{ arcmin}^2$, encompassing most of the Hubble extreme deep field (XDF) (Decarli et al. 2019; González-López et al. 2019).

In this chapter, we aim to create a synthetic equivalent of ASPECS, that is a synthetic ALMA UDF. Synthetic observations have many uses, such as testing the effectiveness of the simulations and implemented theoretical models as well as highlighting limitations in the observations (e.g. the impact of sensitivity limits, or cosmic variance) and ultimately to improve our understanding of the relevant physics. To create these observations, we require two key components: a cosmological simulation and the means to predict the CO line emission from the simulated gas properties (CO line emission is rarely calculated during the

simulation).

The simulation required for this project should have both a cosmologically representative volume (to produce a realistic galaxy population) and a high spatial resolution (to recover ISM structure). There is, however, always a compromise between these two requirements. Simulations typically fall into one of two regimes: ‘zoom’ or ‘box’. In the ‘box’ regime (e.g. Illustris, EAGLE (Schaye et al. 2015)), the focus is on large scale properties. There is a large galaxy population but the resolution is coarse. In the ‘zoom’ regime (e.g. FIRE, Auriga (Grand et al. 2017)), the focus is on a single galaxy within the context of a local cosmological environment. The selected galaxy is simulated at a high resolution while the surrounding region is left at a much coarser resolution. The IllustrisTNG50 simulation is the first of a new class of cosmological simulations, which aim to overcome the compromise between volume and resolution; with a simulated volume of 51.7^3 cMpc^3 and a median radius of $\sim 100\text{--}140 \text{ pc}$ for star-forming gas cells (over cosmic time).¹ The IllustrisTNG simulations are a series of gravity+magnetohydrodynamical simulations based on the moving mesh code AREPO (Springel 2010). The simulations, starting from a redshift of 127 and evolved until the present day, track the coupled dynamics of dark and baryonic matter. The key goals of the simulations are to study the physical processes that drive galaxy formation and evolution; this also necessitates effective star formation, evolution, and feedback models.

When it comes to predicting CO line emission from gas properties, there are many available options and each with its own strengths and weaknesses. Some codes, such as RADMC-3D (Dullemond et al. 2012), perform exact radiative transfer in 3D and can be used on arbitrarily defined gas density distributions. However, this comes at the expense of computational time. We choose to use DESPOTIC. This code predicts line emission in cool atomic and molecular gas environments using an escape probability formalism. It is also limited to 1D problems. With these restrictions, DESPOTIC is much faster than RADMC-3D and much more appropriate when working with a large parameter space. We would also benefit little from the detailed calculations of RADMC-3D, since the gas within the simulation will lack detailed internal structure.

The synthetic CO observations we produce are for a $2.5' \times 2.5'$ field of view over the redshift range 1–4. In order to produce these observations from the IllustrisTNG50 simulation, we must first construct a synthetic light cone. That is, we must select the galaxies (more accurately the gas) that fall within a volume corresponding to the field of view and redshift range. This volume is

¹IllustrisTNG50 is in fact the name for a series of simulations each with a different resolution. We will use ‘IllustrisTNG50’ to refer to what is officially IllustrisTNG50–1, which has the highest resolution of the series.

distributed over the simulation snapshots and care must be taken to ensure that inherent redshift of the galaxy agrees with the galaxy’s ‘placement’ within the light cone. For this purpose we have developed the code `LIGHTCONE`. With the contents of the light cone selected, `DESPOTIC` can convert gas properties into CO emission. These synthetic observations, can then be used to construct diagnostics such as a CO luminosity function.

The structure of this chapter is as follows. In [Section 5.1](#) we describe the simulation upon which we base our synthetic observations. We describe the process of extracting data from the simulation to form synthetic light cones in [Section 5.2](#)^[p. 128]. In [Section 5.3](#)^[p. 138] we give details on how we use the properties of the gas in the simulation to predict the CO emission. The synthetic CO observations we construct are discussed in [Section 5.4](#)^[p. 156]. Finally, in [Section 5.5](#)^[p. 165], we summarise the results and present our conclusions.

5.1 Cosmological simulation

In this chapter we make use of the IllustrisTNG50 simulation. This is a 51.7^3 cMpc^3 cosmological volume simulated with the code `AREPO`. IllustrisTNG50 is one of a set of simulations, with companions IllustrisTNG100 and IllustrisTNG300 which have the volumes $\sim 100^3 \text{ cMpc}^3$ and $\sim 300^3 \text{ cMpc}^3$, respectively (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich, Nelson, Hernquist et al. 2018; Springel et al. 2018). Unlike IllustrisTNG50, which has unprecedented spatial resolution given the volume simulated, these two larger volumes fall within the standard ‘box’ regime. These simulations are designed to complement each other, allowing the study of different aspects of galaxy formation and evolution. All of the IllustrisTNG simulations, barring resolution, are in effect the same. That is, the physical aspects of the model, both parameter values and code, are the same across all three volumes. The standard Λ CDM is used; the cosmological parameters are consistent with Planck Collaboration (2016): $\Omega_m = 0.31$, $\Omega_b = 0.0486$, $\Omega_\Lambda = 0.69$, $h_0 = 0.677$, $\sigma_8 = 0.8159$, $n_s = 0.97$.

Despite the relatively small volume, IllustrisTNG50 has a representative galaxy population. At $z = 0$, the population includes one object of mass exceeding $1.82 \times 10^{14} M_\odot$, ~ 60 massive galaxies residing in halo masses of $M > 10^{12.5} M_\odot$, ~ 130 Milky-Way-mass galaxies, and thousands of dwarf galaxies. The simulation itself started with 2160^3 gas particles and an equal number of dark matter particles. The target mass resolution in star particles and gas particles/cells is $5.7 \times 10^4 M_\odot h^{-1}$ and $3.1 \times 10^5 M_\odot h^{-1}$ in dark matter particles. At $z = 0$, the minimum gas radius was just 6 pc, with a mean of 6.2 kpc. The mean radius in star forming gas was 140 pc.

5 The synthetic ALMA UDF

The IllustrisTNG simulations build upon its predecessor, Illustris. Along with numerical improvements, the most notable change is the inclusion of magnetic fields: the evolution of CDM and gas is governed by the coupled equation of gravity and magnetohydrodynamics. In addition to this, there have been improvements in the galaxy formation model. These improvements focused on the formation, growth, and feedback of black holes; star-formation driven galactic winds; and stellar evolution and the chemical enrichment of gas.

Gas cooling occurs radiatively in the presence of a UV background radiation field, which is redshift dependent and spatially uniform; corrections for self-shielding are also taken into account. Following the Springel et al. (2003) model, gas within the simulation is eligible for star formation once reaching a Hydrogen number density threshold of $n_{\text{H}} \simeq 0.1 \text{ cm}^{-3}$. The stars form stochastically following the Kennicutt-Schmidt relation and assuming a Chabrier (2003) IMF. The two-phase ISM model is not resolved but is assumed. Stars return mass and metals to the ambient ISM via supernovae (type Ia and II) and asymptotic giant branch. Nine elements are traced in the simulation (H, He, C, N, O, Ne, Mg, Si, and Fe); although their state (e.g. ionised, molecular) is not.

Accompanying each snapshot is a group catalogue. Halos are identified using the friends-of-friends (FoF) algorithm (Davis et al. 1985), with linking length of $b = 0.2$. The algorithm is performed on dark matter particles, other particle types belong to the same group as their nearest dark matter particle. Subhalos (i.e. galaxies) are found within the FoF halo using the SUBFIND algorithm (Springel et al. 2001).

5.2 Building a synthetic light cone

In this section, we describe how we generate synthetic light cones from the IllustrisTNG50 simulation using the code LIGHTCONE, which we have developed for this purpose. We briefly describe the implementation of the code but we dedicate most of the text to the logic and procedure used.

The LIGHTCONE code uses a command line interface to run the key steps required to produce a synthetic light cone from a cosmological simulation; these steps we describe in the following paragraphs. The code is developed for Python (Rossum et al. 1991) (version 3.7²), with the majority of the code written in Python. The exceptions to this are the two *Python* classes that handle cosmological calculations and the selection process of simulation data

²<https://docs.python.org/3.7/index.html>

5.2 Building a synthetic light cone

(the most important and used aspects of the code), which are written entirely in C (Kernighan et al. 1988) using the C API for Python (CPython). These classes also make use of the GNU Scientific Library³ (GSL) (Galassi et al. 2009) for integration and root-finding. For array handling, we use NUMPY⁴ (Oliphant 2006) through both its Python and C APIs. Where possible the code can run in parallel using MPI⁵ implemented through the Python library MPI4PY⁶ (Dalcín et al. 2005, 2008, 2011). The light cone data is written out in the HDF5⁷ file format using the Python library H5PY⁸ (Collette 2013); more specifically version 2.9, which introduced ‘Virtual Datasets’. The way in which we construct a complete light cone from its parts (described below) was influenced by the implementation of virtual datasets — a single dataset constructed out of multiple existing datasets through links and so avoiding the need to copy data.

We now describe the steps involved to construct a synthetic light cone. It is first necessary to determine the properties of the simulation from which we build the light cone: the cosmology used, the volume simulated, and the redshifts of the available snapshots all play vital roles in what follows. The cosmology implemented in the simulation is used in all cosmological calculations, such as (and most importantly) distance calculations. The available snapshots inform us about the redshift range, $z_{\ell}^{\text{range}} = [z_{\ell}^{\text{min}}, z_{\ell}^{\text{max}}]$, for which we can build a light cone, ℓ , and also determine how far we must deviate from the snapshot’s inherent redshift when constructing a light cone. The width, \tilde{W} , and height, \tilde{H} , of the simulated volume limits the field of view: the field of view cannot exceed \tilde{W} or \tilde{H} over the entire redshift range of the light cone. The depth, \tilde{D} , of the simulated volume determines how many times a single snapshot must be used to cover its assigned redshift range, which is again determined by the redshifts of available snapshots. We clarify here, that for the light cone (and its parts), we refer to the three spatial axes as width, height, and depth: this is more intuitive than the standard (x, y, z) , which we reserve for describing the location of objects within the simulated volume.

We produce light cones from the IllustrisTNG50 simulation (discussed in Section 5.1^[p. 127]) and summarise again the key properties of the simulation when building a light cone. The cosmology used by IllustrisTNG50 is the standard Λ CDM model with the values of cosmological parameters taken from Planck Collaboration (2016). The simulated comoving-volume is

³<http://www.gnu.org/software/gsl>

⁴<https://numpy.org>

⁵<https://www.mpi-forum.org>

⁶<https://mpi4py.readthedocs.io>

⁷<http://www.hdfgroup.org/HDF5>

⁸<http://www.h5py.org>

5 The synthetic ALMA UDF

$\tilde{V} = 51.7^3 \text{ cMpc}^3$. There are 100 snapshots with redshift range 0.0–20.0; 30 of which are in the redshift range 1.0–4.0.

The choice of redshift range and field of view we make for all our light cones conform (approximately) to the specifications of the ASPECS survey (more specifically Decarli et al. 2016). That is, we use a redshift range $z_\ell^{\text{range}} = [1, 4]$ and field of view $\phi = (\phi_w, \phi_h) = (2.5', 2.5')$.

Now that the properties of the simulation and light cone have been defined, we describe the steps we take to produce a synthetic light cone. We first, assign the redshift range, z_j^{range} , for which a given snapshot, j , can be used. We aim to minimise how much we deviate from the snapshot's inherent redshift by assigning redshift ranges that are the midpoints between consecutive snapshots:

Equation 5.1

$$\begin{aligned} z_j^{\text{range}} &= [z_j^{\text{min}}, z_j^{\text{max}}] \\ &= \left[\frac{z_j + z_{j-1}}{2}, \frac{z_j + z_{j+1}}{2} \right]. \end{aligned}$$

We now need to define the volumes we must extract from the snapshots that together form a single light cone; we refer to an extracted volume as a ‘cut-out’. The total depth required of a given snapshot, \tilde{d}_j , to get from z_j^{min} to z_j^{max} is given by

Equation 5.2

$$\tilde{d}_j = \frac{c}{H_0} \int_{z_j^{\text{min}}}^{z_j^{\text{max}}} \frac{1}{E(z)} dz,$$

where $E(z)$ is the expansion factor as defined in Equation 1.5^[p. 2]. It is often the case that \tilde{d}_j exceeds the depth of the simulated volume, \tilde{D} , and so the recurrence of a given snapshot, R_j , is:

Equation 5.3

$$R_j = \left\lceil \frac{\tilde{d}_j}{\tilde{D}} \right\rceil.$$

This means that for snapshot j , we require R_j cut-outs. We define the depth of a given cut-out for a given snapshot, $\tilde{d}_{j,c}$, as

Equation 5.4

$$\tilde{d}_{j,c} = \frac{\tilde{d}_j}{R_j}.$$

It is also important to calculate what redshift range, $z_{j,c}^{\text{range}}$, is associated with the cut-out:

Equation 5.5

$$\tilde{d}_{j,c} = \frac{c}{H_0} \int_{z_{j,c}^{\text{min}}}^{z_{j,c}^{\text{max}}} \frac{1}{E(z)} dz,$$

5.2 Building a synthetic light cone

which is subject to the conditions

$$\tilde{d}_j \stackrel{!}{=} \sum_{c=1}^{R_j} \tilde{d}_{j,c}; \quad z_{j,c=i}^{\min} \equiv z_{j,c=i-1}^{\max}; \quad z_{j,c=1}^{\min} \equiv z_j^{\min}; \quad z_{j,c=R_j}^{\max} \equiv z_j^{\max}. \quad \text{Equation 5.6}$$

The width, $\tilde{w}_{j,c}$, and height, $\tilde{h}_{j,c}$, of the cut-out are both calculated using the same method and depend on the redshift and field of view. Since the width and height are redshift dependent, at present we concern ourselves only with the maximum width, $\tilde{w}_{j,c}^{\max}$, and height, $\tilde{h}_{j,c}^{\max}$; we account for the variation with redshift in later steps. The maximum comoving width and height occur at the upper limit of the cut-out redshift range (i.e. $z = z_{j,c}^{\max}$) and so this is the redshift we use in the calculations below.

We start off in the general case, which we illustrate in [Figure 5.1](#)^[p.132]. The observer, at Point O, is at redshift $z_O = 0$, Point A is at redshift z_A , and Point B is at redshift z_B . On the sky, the angle between Point A and Point B is θ . The distances to Point A and to Point B from Point O are given by

$$\tilde{r}_A = \frac{c}{H_0} \int_{z_O}^{z_A} \frac{1}{E(z)} dz; \quad \tilde{r}_B = \frac{c}{H_0} \int_{z_O}^{z_B} \frac{1}{E(z)} dz. \quad \text{Equation 5.7}$$

The distance between Point A and Point B, \tilde{r}_{AB} , is calculated using (Peacock 1999):

$$\begin{aligned} S_K^2(\tilde{r}_{AB}) &= S_K^2(\tilde{r}_A) C_K^2(\tilde{r}_B)^2 \\ &+ S_K^2(\tilde{r}_B)^2 C_K^2(\tilde{r}_A) \\ &+ K S_K^2(\tilde{r}_A) S_K^2(\tilde{r}_B) \sin^2(\theta) \\ &- 2 S_K(\tilde{r}_A) S_K(\tilde{r}_B) C_K(\tilde{r}_A) C_K(\tilde{r}_B) \cos(\theta), \end{aligned} \quad \text{Equation 5.8}$$

where K is the curvature introduced in [Equation 1.1](#)^[p.1], S_K is defined as

$$S_K(\tilde{u}) = \begin{cases} |K|^{-\frac{1}{2}} \sin\left(\tilde{u} |K|^{\frac{1}{2}}\right) & K > 0 \\ \tilde{u} & K = 0 \\ |K|^{-\frac{1}{2}} \sinh\left(\tilde{u} |K|^{\frac{1}{2}}\right) & K < 0 \end{cases}, \quad \text{Equation 5.9}$$

and the cosine-like equivalent (Peacock 1999), C_K , as

$$\begin{aligned} C_K(\tilde{u}) &\equiv \sqrt{1 - \text{sgn}(K) |K| S_K^2(\tilde{u})} \\ &= \begin{cases} \cos\left(\tilde{u} |K|^{\frac{1}{2}}\right) & K > 0 \\ 1 & K = 0 \\ \cosh\left(\tilde{u} |K|^{\frac{1}{2}}\right) & K < 0 \end{cases}. \end{aligned} \quad \text{Equation 5.10}$$

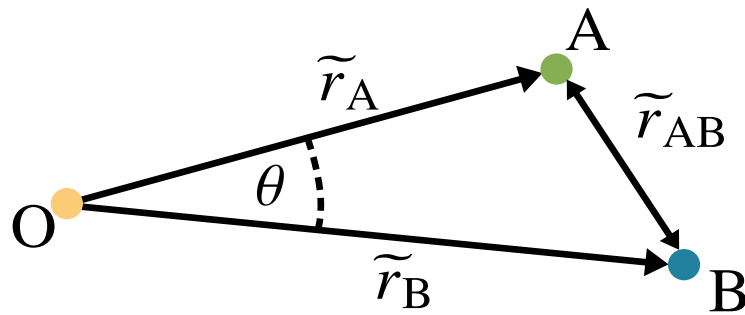


Figure 5.1

A schematic diagram of the setup described in [Section 5.2](#)^[p.128], showing the location of the observer O relative to two distant points A and B. The two points A and B are, from the observer's perspective, separated by an angle θ on the sky. The points are, however, at different redshifts and therefore distances from the observer (\tilde{r}_A and \tilde{r}_B). The comoving separation between the two points is calculated from \tilde{r}_A , \tilde{r}_B , and θ using [Equation 5.8](#)^[p.131].

5.2 Building a synthetic light cone

When calculating $\tilde{w}_{j,e}^{\max}$ and $\tilde{h}_{j,e}^{\max}$, the Points A and B are the edges of our field of view. This means that the two points are at the same redshift $z_A = z_B \equiv z_{j,e}^{\max}$ and therefore at the same distance $\tilde{r}_A = \tilde{r}_B \equiv \tilde{r}_{j,e}^{\max}$. With this simplification, Equation 5.8^[p. 131] becomes

$$\begin{aligned} S_K^2(\tilde{l}_{j,e}^{\max}) &= 2S_K^2(\tilde{r}_{j,e}^{\max}) C_K^2(\tilde{r}_{j,e}^{\max}) \\ &\quad + K S_K^4(\tilde{r}_{j,e}^{\max}) \sin^2(\theta) \\ &\quad - 2S_K^2(\tilde{r}_{j,e}^{\max}) C_K^2(\tilde{r}_{j,e}^{\max}) \cos(\theta) , \end{aligned} \tag{Equation 5.11}$$

where $(\tilde{l}_{j,e}^{\max}, \theta)$ are stand-in variables for $(\tilde{w}_{j,e}^{\max}, \phi_w)$ and $(\tilde{h}_{j,e}^{\max}, \phi_h)$. We now have the defining properties of the light cone. In Figure 5.2^[p. 134], we show the structure of the light cone as formed from individual cut-outs.

With the structure of the light cone defined, we describe the steps we take to extract the data from the simulation. A given cut-out is governed by the properties represented in Figure 5.2^[p. 134] but is defined by two additional quantities: the location of its centre, $C'_{j,e}$, and the arbitrary rotation of the simulated volume, $R'_{j,e}$, (the necessity of this quantity is described below). We mark these with a 'prime', to indicate this is just one realisation of a cut-out: for any given cut-out with $\tilde{w}_{j,e}^{\max}$, $\tilde{h}_{j,e}^{\max}$, and $z_{j,e}^{\text{range}}$ (or equivalently $\tilde{d}_{j,e}$), there are innumerable unique combinations of $C'_{j,e}$ and $R'_{j,e}$. The centre of the cut-out is randomly selected from a uniform distribution $\mathcal{U}(a, b)$, which is defined such that all values are equally probable within the range $[a, b)$. We explicitly define $C'_{j,e}$ as

$$\begin{aligned} C'_{j,e} &= (C'_{j,e,w}, C'_{j,e,h}, C'_{j,e,d}) \\ &= \left(\mathcal{U}\left(\frac{\tilde{w}_{j,e}^{\max}}{2}, \tilde{W} - \frac{\tilde{w}_{j,e}^{\max}}{2}\right), \mathcal{U}\left(\frac{\tilde{h}_{j,e}^{\max}}{2}, \tilde{H} - \frac{\tilde{h}_{j,e}^{\max}}{2}\right), \mathcal{U}\left(\frac{\tilde{d}_{j,e}}{2}, \tilde{D} - \frac{\tilde{d}_{j,e}}{2}\right) \right) , \end{aligned} \tag{Equation 5.12}$$

where the factors of a half are to ensure that the entire cut-out falls within the simulated volume. The second quantity, $R'_{j,e}$, defines the random rotation of the simulated volume along each axis. For practical reasons, the rotation along a given axis i is randomly selected from a choice of rotations that are multiples of 90° : $\psi_i \in \{0^\circ, 90^\circ, 180^\circ, 270^\circ\}$. We therefore define $R'_{j,e}$ as

$$R'_{j,e} = (\psi_x, \psi_y, \psi_z) . \tag{Equation 5.13}$$

Rotating the volume and randomly selecting the location of the cut-out centre are a means to counteract the fact that we are repeatedly viewing the same volume at different times of its evolution. This can result in the appearance of the same objects and structures within a light cone at different redshifts.

5 The synthetic ALMA UDF

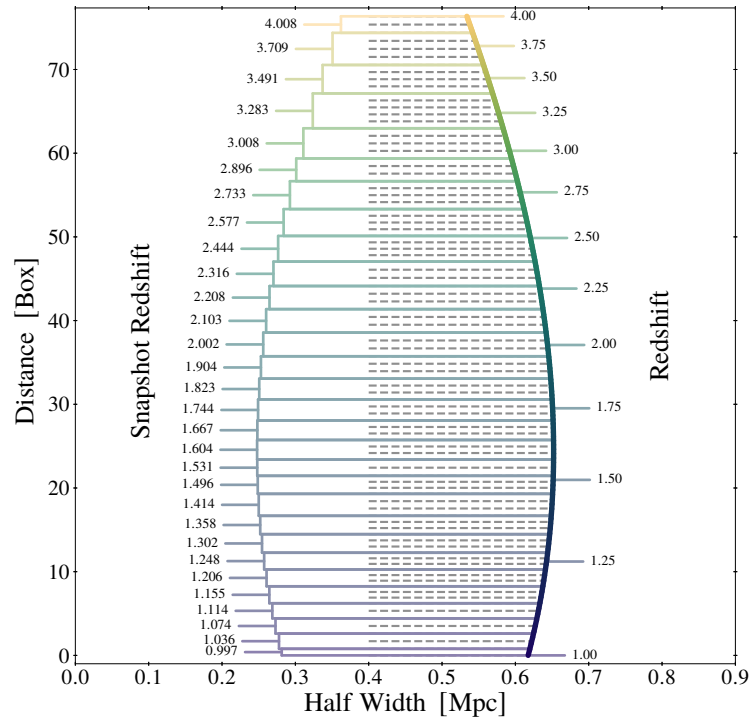


Figure 5.2

In terms of simulation snapshots and individual cut-outs, we illustrate how the synthetic light cone is structured. The horizontal axis is the physical width from the centre of the light cone (half the total width). The vertical axis is the distance from the front of the light cone in units of the depth of the simulated volume (51.7 cMpc). The curve, coloured by redshift, shows how the physical width changes with distance. The fact that this width does not monotonically increase with distance is a result of living in an expanding Universe. At greater distances, we see the Universe when it was younger and therefore smaller. To the left of the curve, we indicate the distance (or equivalently, redshift) that is covered by each of the 30 snapshots. We also indicate (dashed grey lines) the divisions between the different cut-outs of the same snapshot. This makes it clear to see the depth of each cut-out and also how many cut-outs are required of a given snapshot to cover the assigned redshift range.

5.2 Building a synthetic light cone

In the worst case, where snapshots are simply stacked, this can result in a streaming effect; this has been demonstrated in Blaizot et al. (2005, Figure 1 — reproduced in Figure 5.3^[p. 136]). Through this randomisation, we should reduce the chance of the same objects appearing multiple times within a light cone. This would ideally mean no repeated objects but, at the very least, should obfuscate any repetitions that do occur.

With the cut-outs fully defined, the next step is to select the data from the simulation. That is to say, we identify all the objects (in our case, gas cells) that are within the defined cut-out and store their angular position and redshift, both of which must be calculated. We use Figure 5.4^[p. 137] to help illustrate this two-step process.

For each gas cell we first determine if its position (once rotated by $R'_{j,e}$) falls within the cuboid of width, $\tilde{w}_{j,e}^{\max}$, height, $\tilde{h}_{j,e}^{\max}$, and depth, $\tilde{d}_{j,e}$, with its centre at $C'_{j,e}$ (grey dashed cuboid). If not within this volume, the cell is excluded and we move to the next cell. If the gas cell *is* within this volume (green circle), we calculate the more computational expensive quantities: the cell redshift, z_{cell} , and the angular position (ω_w , ω_h). The redshift is given by

$$\Delta \tilde{d} = \frac{c}{H_0} \int_{z_{j,e}^{\min}}^{z_{\text{cell}}} \frac{1}{E(z)} dz, \quad \text{Equation 5.14}$$

where $\Delta \tilde{d}$ is the distance to the cell from the front of the cut-out volume.

The angular position of the cell, which we calculate using its distance from the centre ($\Delta \tilde{w}$, $\Delta \tilde{h}$) and the cell redshift, z_{cell} , is given as follows. Starting with Equation 5.11^[p. 133] and using the identity $\cos^2(\theta) + \sin^2(\theta) = 1$, we get the following equation

$$\begin{aligned} 0 = & \left[K S_K^4(\tilde{r}_{\text{cell}}) \right] \cos^2(\theta) \\ & + \left[2 S_K^2(\tilde{r}_{\text{cell}}) C_K^2(\tilde{r}_{\text{cell}}) \right] \cos(\theta) \\ & + \left[S_K^2(\Delta \tilde{l}) - 2 S_K^2(\tilde{r}_{\text{cell}}) C_K^2(\tilde{r}_{\text{cell}}) - K S_K^4(\tilde{r}_{\text{cell}}) \right], \end{aligned} \quad \text{Equation 5.15}$$

where

$$\tilde{r}_{\text{cell}} = \frac{c}{H_0} \int_0^{z_{\text{cell}}} \frac{1}{E(z)} dz, \quad \text{Equation 5.16}$$

and $(\Delta \tilde{l}, \theta)$ are stand-in variables for $(\Delta \tilde{w}, \omega_w)$ and $(\Delta \tilde{h}, \omega_h)$. Temporarily re-writing Equation 5.15 as $0 = A \cos^2(\theta) + B \cos(\theta) + C$, we can solve for $\cos(\theta)$ using the positive solution of quadratic formula or through simple

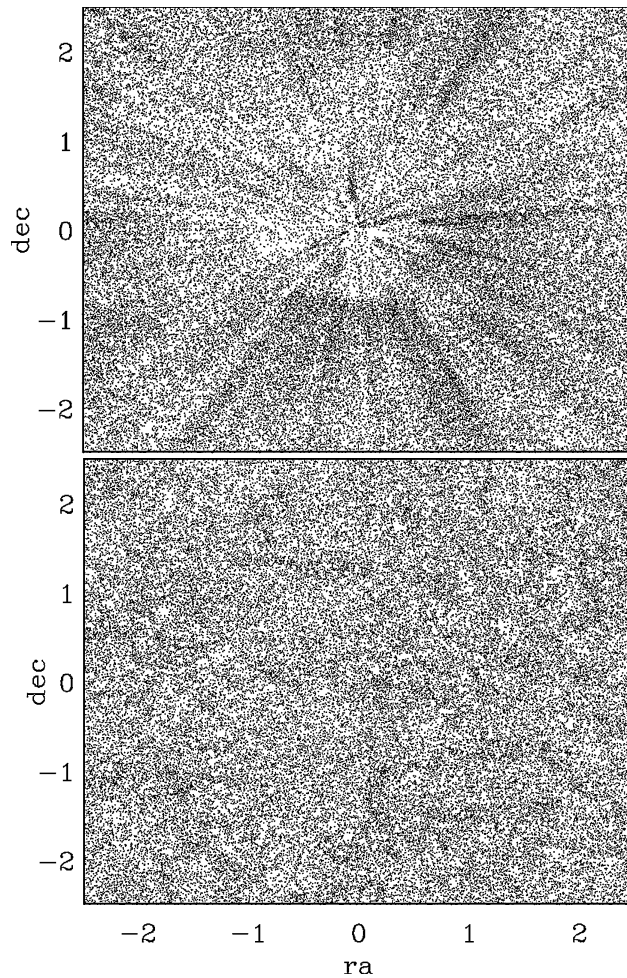
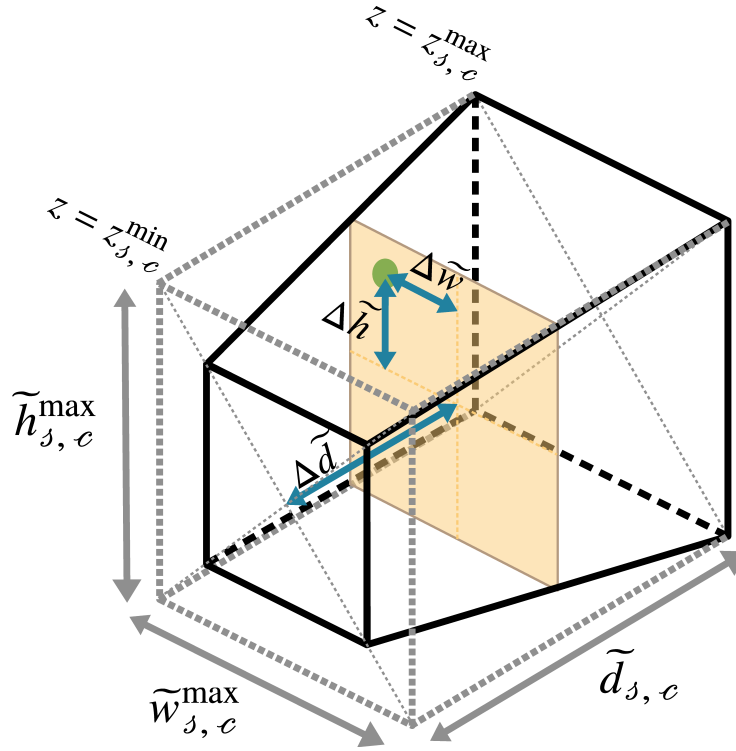


Figure 5.3

Two projections of the same mock galaxy catalogue for a $5^\circ \times 5^\circ$ field of view. These projections highlight how poor light cone construction can produce an unrealistic results. Top: The streaming effect that occurs when snapshots are stacked without randomisation. This is the result of the same objects appearing multiple times in the very similar configurations and at very similar locations. Bottom: Randomisation of the snapshot orientation removes the streaming effect to produce a more realistic image.

Image Credit: Blaizot et al. (2005, Figure 1), reproduced with permission © OUP

5.2 Building a synthetic light cone



We illustrate here the two steps taken to identify if an object in the simulation snapshot falls within the light cone's field of view. The first selection cut is defined by the grey dashed cuboid. This cuboid has the dimensions given by the maximum width ($\tilde{w}_{s,c}^{\max}$), maximum height ($\tilde{h}_{s,c}^{\max}$), and depth ($\tilde{d}_{s,c}$) of the cut-out. Objects from the simulation that fall outside the cuboid do not fall in light cone's field of view. The second selection cut accounts for the change in width and height with distance (over the redshift range $[z_{s,c}^{\min}, z_{s,c}^{\max}]$). The black square frustum gives the true shape of the cut-out. Objects that fall within the cone are in the field of view. The green circle is one of the simulation objects that falls within the field of view. The two steps are taken because calculating the object's redshift from the distance to the front of the cut-out, $\Delta\tilde{d}$, and its angular position from the horizontal and vertical offset from the centre, respectively $\Delta\tilde{w}$ and $\Delta\tilde{h}$, is computationally expensive.

Figure 5.4

rearrangement:

$$\text{Equation 5.17} \quad \cos(\theta) = \begin{cases} -\frac{B + \sqrt{B^2 - 4AC}}{2A} & A \neq 0 \\ -\frac{C}{B} & A = 0 \end{cases}.$$

The angular positions can now be used to perform the final cut, which accounts for the variation of $\tilde{w}_{j,c}$ and $\tilde{h}_{j,c}$ with redshift. This is shown as the black square frustum in [Figure 5.4](#)^[p. 137]. If both the conditions

$$\text{Equation 5.18} \quad \omega_w \leq \frac{\phi_w}{2}; \quad \omega_h \leq \frac{\phi_h}{2}$$

are satisfied the location (ω_w, ω_h, z) of the gas cell is saved, otherwise the cell is excluded. These saved gas cell locations form the cut-out data.

The final step to construct a light cone is to piece the cut-outs together. Given the light cone specifications and simulation properties, we require $\sum_{j=1}^{30} R_j = 91$ cut-outs to form a complete light cone (see [Figure 5.2](#)^[p. 134]). We also define ten realisations (i.e. ten $C'_{j,c}$ and $R'_{j,c}$) for each of the required cut-outs. This means that when constructing a light cone, we can select one cut-out of ten for each of the required redshift ranges. We therefore have the data to form 10^{91} unique, but not independent, light cones. For the light cones we produce, the selection of cut-outs is done randomly. This is also where we utilise the ‘Virtual Datasets’ feature of HDF5 files: a complete light cone is simply a concatenation of selected cut-outs. Instead of copying the data we can link to the existing cut-out datasets. The result is that we can produce hundreds of light cones very cheaply as the required data already exists and no copy is required.

5.3 Synthetic CO emission

The simulation we use for our synthetic CO observations does not contain emission information. We therefore model the gas cells in the simulation as individual GMCs using `DESPOTIC`. In [Section 5.3.1](#)^[p. 139], we describe how `DESPOTIC` models these clouds in order to predict the CO emission. We also include the input quantities that we use to describe the cloud and its environment. We explain in [Section 5.3.2](#)^[p. 142] how simulation-based quantities are calculated for use with `DESPOTIC`. The key steps taken by `DESPOTIC` are outlined in [Section 5.3.3](#)^[p. 143]. For particle purposes, it is not possible to run `DESPOTIC` on all the gas cell data; in [Section 5.3.4](#)^[p. 146] we describe the steps to produce a lookup table allowing us to interpolate CO emission from gas cell properties.

5.3.1 Model cloud

In order to produce synthetic ALMA UDF maps we need to convert the properties of gas cells present in the light cone into CO emission data. For this purpose, we use `DESPOTIC` (a chemical, thermal, and statistical equilibrium code), which calculates the state of interstellar clouds and predicts the (CO) line emission that would be observed. Through describing the inputs we give to `DESPOTIC`, as well as an overview of the code, the caveats of the model will be made apparent. In the description that follows (and [Table 5.1](#)^[p. 140]), the value of quantities marked by † come from the simulation: we describe explicitly how we recover these values once their use has been described. A summary of `DESPOTIC` input quantities, along with their initial values and a brief description, are provided in [Table 5.1](#)^[p. 140] with further details given below. The values in [Table 5.1](#)^[p. 140] follow those seen in Krumholz (2014, Table 1 — MilkyWayGMC) with the exception of ζ where we use the more conservative value used in Narayanan et al. (2017, Table 1).

Physical In `DESPOTIC` a single cloud is modelled as N_{zone} independent, nested spheres (zones); the properties of each zone are uniform across the zone. Each zone i , where $i = [0, N_{\text{zone}})$, is characterised by its column density of Hydrogen nuclei $N_{\text{H},i}$:

$$N_{\text{H},i} = \frac{i + 0.5}{N_{\text{zone}}} \left(\frac{3}{4} N_{\text{H}} \right), \quad \text{Equation 5.19}$$

where N_{H} is the mean Hydrogen number column density[†]. We use $N_{\text{zone}} = 16$, this was shown by Narayanan et al. (2017, Appendix A) to be sufficient for convergence. We note that for all other properties the same initial values are used for all the zones of a single cloud.

Gas The gas (and dust) temperature is set to an initial value of 10 K. This is a typical temperature for the dense molecular gas phase of the ISM and the phase that would be detected with CO emission (Draine 2011). Ideally the temperature information would come from the simulation; however, gas temperatures within cosmological simulations are usually significantly greater ($\sim 10^4$ K) than what is required for dense molecular gas to exist. The IllustrisTNG50 simulation has an effective cooling floor of $\sim 10^4$ K due a lack of metal fine-structure line cooling (Nelson et al. 2019).

[†]This quantity is derived from simulation data.

5 The synthetic ALMA UDF

Table 5.1

The input parameter we use to run `DESPTIC` for predicting CO emission. We include a brief description of each of the quantities and give further details in the text (Section 5.3.1^[p. 139]). Some quantities cannot be fully specified as their value comes from the simulation and are marked with †.

Quantity	Value	Unit	Description
Physical			
N_{zone}	16	—	Number of radial zones
Gas			
T_{g}	10	K	Gas temperature
Dust			
T_{d}	10	K	Dust temperature
β_{d}	2	—	IR spectra index
Z_{d}	$1 \times Z'^{\dagger}$	—	Normalised dust abundance
α_{gd}	$3.2 \times 10^{-34} \times Z'^{\dagger}$	$\text{erg cm}^3 \text{ K}^{-3/2}$	Gas-dust collision coupling coefficient
$\sigma_{\text{d},10}$	$2 \times 10^{-26} \times Z'^{\dagger}$	$\text{cm}^2 \text{ H}^{-1}$	Cross section to 10 K thermal radiation
$\sigma_{\text{d,PE}}$	$1 \times 10^{-21} \times Z'^{\dagger}$	$\text{cm}^2 \text{ H}^{-1}$	Cross section to 8–13.6 eV photons
$\sigma_{\text{d,ISRF}}$	$3 \times 10^{-21} \times Z'^{\dagger}$	$\text{cm}^2 \text{ H}^{-1}$	Cross section to average ISRF
Radiation			
$T_{\text{rad,dust}}$	0	K	IR radiation temperature
T_{CMB}	$2.73 \times (1 + z^{\dagger})$	K	CMB temperature
ζ	$0.1 \times 10^{-16} \times \Sigma'_{\text{SFR}}{}^{\dagger}$	$\text{s}^{-1} \text{ H}^{-1}$	Ionisation rate
χ	$1 \times \Sigma'_{\text{SFR}}{}^{\dagger}$	—	Normalised ISRF strength
Chemistry			
N_{H}	†	cm^{-2}	Hydrogen number column density
n_{H}	†	cm^{-3}	Hydrogen number volume density
r_{op}	0.25	—	Ortho-to-para- H_2 ratio
x_{H_2}	0.5	—	Molecular Hydrogen abundance
$x_{\text{H I}}$	0	—	Atomic Hydrogen abundance
$x_{\text{H II}}$	0	—	Ionised Hydrogen abundance
x_{e^-}	0	—	Free electrons abundance
x_{He}	†	—	Helium abundance
x_{C}	†	—	Carbon abundance
x_{O}	†	—	Oxygen abundance
x_{M}	$2 \times 10^{-7} \times Z'^{\dagger}$	—	Other metals abundance

5.3 Synthetic CO emission

Dust We scale most dust quantities linearly with metallicity[†], Z ; this scaling is expressed through Z' , the normalised metallicity:

$$Z' = \frac{Z}{1 Z_{\odot}} .$$

Equation 5.20

Scaled quantities include Z_d , the total dust abundance normalised to the Milky Way value; dust cross-sections per Hydrogen nucleus, $\sigma_{d,i}$, used in calculating heating and cooling rates for different mechanisms of both dust *and* gas; and α_{gd} , which characterises the collision rate between dust and gas. The quantities that are *not* scaled with metallicity are the dust temperature, T_d , set equal to the gas temperature T_g ; and β_d , the spectral index for IR radiation.

Radiation The clouds experience a background radiation field. This radiation field includes the IR radiation experienced by dust, characterised by the temperature $T_{\text{rad,dust}}$; and the CMB, characterised by the temperature T_{CMB} , which is scaled according to the redshift[†], z , of the cloud. There are two quantities which we scale with the SFR surface density[†], Σ_{SFR} , expressed through the normalised SFR surface density, Σ'_{SFR} :

$$\Sigma'_{\text{SFR}} = \frac{\Sigma_{\text{SFR}}}{\Sigma_{\text{SFR,MW}}} ,$$

Equation 5.21

where $\Sigma_{\text{SFR,MW}} = 0.001 \text{ M}_{\odot} \text{ yr}^{-1} \text{ pc}^{-2}$ (Bonatto et al. 2011). These quantities are the ionisation rate per Hydrogen nucleus caused by interactions with cosmic rays and hard x-rays, ζ ; and the interstellar radiation field (ISRF) accounting for FUV radiation normalised to the Milky Way value, χ . The procedure of scaling with SFR surface density follows that used by Popping et al. (2019), which found better agreement with observations when scaling with surface density than the galactic SFR.

Chemistry The chemical composition of the cloud is specified in terms of abundance relative to Hydrogen nuclei, x_i . We only use a scaling relation for one of the abundances: the collective abundance of all other remaining metals, x_M , where we scale the default value (2×10^{-7}) by the (normalised) metallicity. For molecular Hydrogen it is also important (for specific heat calculations, see Krumholz 2014, Appendix A) to specify r_{op} , the ratio of ortho-Hydrogen (nuclei spinning in the same direction) to para-Hydrogen (nuclei spinning in opposite directions).

5.3.2 Simulation input

In [Section 5.3.1](#)^[p. 139] we listed the quantities used as input for `DESPOTIC` to predict CO emission; however, not all values can be specified explicitly as they originate from the properties of the gas cells within the simulation (marked with †). The base quantities that we use from the simulation to calculate input values are the gas mass, M_g ; (comoving) gas density, $\tilde{\rho}_g$; SFR; gas metallicity, Z ; and gas metal fractions, $f_i \equiv M_i/M_g$ for $i = \{\text{H, He, C, O}\}$, where M_i is the mass in species i .

Not all of the listed quantities are in a form required by `DESPOTIC` and so have to be expressed in a more useful manner. Before we can calculate the quantities we need, we must make decisions regarding the shape and density structure of the gas cell (treated as a single cloud). In the simulation, the shape of a gas cell is irregular since it is simply the shape of the moving mesh cell. The gas cells also lack any internal density structure. We therefore model the cloud as a uniform-density sphere. The radius of the cloud, R_c , is calculated from (proper gas density) ρ_g and M_g :

$$\text{Equation 5.22} \quad R_c = \sqrt[3]{\frac{3M_g}{4\pi\rho_g}} .$$

We use this radius to define the area A_c and volume V_c of the cloud:

$$\text{Equation 5.23} \quad A_c = \pi R_c^2 ; \quad V_c = \frac{4}{3}\pi R_c^3 .$$

With the shape of the cloud defined we can now calculate the Hydrogen number column density, N_H , and the Hydrogen number volume density, n_H . First, we calculate the Hydrogen surface density, Σ_H , and Hydrogen volume density, ρ_H ,

$$\text{Equation 5.24} \quad \Sigma_H = \frac{M_g f_H}{A_c} ; \quad \rho_H = \frac{M_g f_H}{V_c} \equiv \rho_g f_H ,$$

where $M_g f_H$ gives the total mass in Hydrogen. We then determine the two number densities using

$$\text{Equation 5.25} \quad N_H = \frac{\Sigma_H}{m_H} ; \quad n_H = \frac{\rho_H}{m_H} ,$$

where m_H is the atomic mass of Hydrogen.

Since we now have an expression for the area of the cloud, we can also convert SFR into a SFR surface density, Σ_{SFR} ,

$$\text{Equation 5.26} \quad \Sigma_{\text{SFR}} = \frac{\text{SFR}}{A_c} .$$

5.3 Synthetic CO emission

The abundance information `DESPOTIC` requires, x_i , is the abundance of nuclei relative to Hydrogen nuclei, more formally given as,

$$x_i = \frac{\mathcal{N}_i}{\mathcal{N}_H}, \quad \text{Equation 5.27}$$

where \mathcal{N}_i is the number of species i nuclei, and \mathcal{N}_H the number of Hydrogen nuclei. We can define \mathcal{N}_i in terms of f_i as follows,

$$\mathcal{N}_i = \frac{M_i}{m_i} = \left(\frac{M_i}{M_g} \right) \frac{M_g}{m_i} = f_i \frac{M_g}{m_i}, \quad \text{Equation 5.28}$$

where m_i is the atomic mass of species i . Following these definitions, we can now write x_i in terms of f_i :

$$x_i = \frac{f_i}{m_i} \frac{m_H}{f_H}. \quad \text{Equation 5.29}$$

The gas metal fraction data is unlike the other quantities we use from the simulation: not every snapshot stores abundance data. Each `IllustrisTNG` snapshot is one of two types: ‘full’ or ‘mini’. The mini snapshots only store a subset of the particle field data compared to full snapshots. Of the 30 snapshots required to produce a light cone covering $z = 1-4$, only five are full snapshots and so only these five snapshots contain abundance data. As mentioned in [Section 5.1](#)^[p.127], each snapshot also comes with a group catalogue, this contains the properties of halos and subhalos found within the current snapshot. The properties stored include the gas metal fraction of the halo/subhalo as a whole. This means that even for the mini snapshots we can still recover abundance data. We use the following procedure to assign abundance data to a given gas cell. If the gas cell is a member of a subhalo, then the subhalo abundances are assigned to the gas cell. If the gas cell is not a member of a subhalo but is a member of a halo, then the halo abundances are assigned to the gas cell. If the gas cell is neither a member of a subhalo nor halo then the gas cell is excluded. Since our aim is to study the ISM through synthetic CO emission, any gas cell that does not belong to a subhalo/halo is most probably not part of the ISM and so can be safely excluded.

We finally note here that in the above list of ‘base quantities’ we have excluded redshift, z . This is because the redshift of a gas cell does not come from the simulation; instead it is determined by its position in the light cone.

5.3.3 Predicting CO emission with `DESPOTIC`

There are three main stages taken by `DESPOTIC`, for each zone within a cloud, in order to calculate line emission: chemical, thermal, and statistical. Firstly,

5 The synthetic ALMA UDF

the chemical network is run to convergence at constant temperature. Then the temperature is calculated with fixed level population of each species. Finally, statistical equilibrium within the level population is achieved. These three stages must be iterated until there is no change (within a tolerance threshold) of the chemical abundance, temperature or level populations.

Chemical The chemical network we select is NL99_SG: a reduced Carbon-Oxygen network developed by Nelson et al. (1999) extended with a non-equilibrium Hydrogen chemical network and additional Carbon and Oxygen chemistry from Glover et al. (2007, 2012). Hydrogen self-shielding is calculated using an approximate analytical function given by Draine et al. (1996). The reactions and rates used by this network are given in Narayanan et al. (2017, Table 2).

Thermal In order to calculate the thermal state of a zone, we need to consider the relevant heating Γ_i , cooling Λ_i , and energy exchange Ψ_i mechanisms of both gas and dust. The rate of change of the energy in gas per Hydrogen nucleus, e_g , is

$$\begin{aligned} \frac{de_g}{dt} = & \Gamma_{\text{ion}}(\zeta, n_{\text{H}}, r_{\text{op}}, x_{\text{H}_2}, x_{\text{H I}}, x_{\text{e}^-}) \\ & + \Gamma_{\text{PE}}(Z_{\text{d}}, \sigma_{\text{d,PE}}, \chi, N_{\text{H}}) \\ & - \Lambda_{\text{line}}(T_{\text{g}}, T_{\text{CMB}}, n_{\text{H}}, x_p) \\ & + \Psi_{\text{gd}}(T_{\text{g}}, T_{\text{d}}, \alpha_{\text{gd}}, n_{\text{H}}) , \end{aligned} \tag{Equation 5.30}$$

and in dust, e_d ,

$$\begin{aligned} \frac{de_d}{dt} = & \Gamma_{\text{ISRF}}(Z_{\text{d}}, \sigma_{\text{d,ISRF}}, \chi, N_{\text{H}}) \\ & + \Gamma_{\text{d,CMB}}(\beta_{\text{d}}, \sigma_{\text{d,10}}, T_{\text{CMB}}) \\ & + \Gamma_{\text{d,IR}}(\beta_{\text{d}}, \sigma_{\text{d,10}}, T_{\text{rad,dust}}) \\ & + \Gamma_{\text{d,line}}(\Lambda_{\text{line}}, \beta_{\text{d}}, \sigma_{\text{d,10}}, N_{\text{H}}) \\ & - \Lambda_{\text{d}}(T_{\text{d}}, \beta_{\text{d}}, \sigma_{\text{d,10}}, N_{\text{H}}) \\ & - \Psi_{\text{gd}}(T_{\text{g}}, T_{\text{d}}, \alpha_{\text{gd}}, n_{\text{H}}) . \end{aligned} \tag{Equation 5.31}$$

The full expression used for each term along with any assumptions or simplifications is provided in Krumholz (2014, Appendix B1 and B2); here we give a brief description of each term and have indicated which of the quantities given in Table 5.1^[p. 140] the mechanism depends on.

Heating terms:

5.3 Synthetic CO emission

- Γ_{ion} — Ionisation heating occurs when cosmic rays or hard x-rays ionise gas and the resulting free electrons then thermalise with the gas
- Γ_{PE} — Photoelectric heating occurs when FUV photons (8–13.6 eV) eject electrons from dust grains (the photoelectric effect) which then thermalise with the gas
- Γ_{ISRF} — ISRF heating results when dust absorbs radiation from the ISRF
- $\Gamma_{\text{d,CMB}}$ — CMB heating results when dust absorbs radiation from the CMB
- $\Gamma_{\text{d,IR}}$ — IR heating results when dust absorbs radiation from the background IR field.
- $\Gamma_{\text{d,line}}$ — Line heating occurs when dust reabsorbs line emission.

Cooling terms:

- Λ_{line} — Line cooling occurs when the gas emits emission line photons which then escape the cloud
- Λ_{d} — IR cooling occurs when the dust emits thermal continuum radiation.

Energy transfer terms:

- Ψ_{gd} — Dust-gas energy exchange results as a collision between the gas and dust grains (using the sign convention that positive values correspond to heating of the gas).

Statistical Calculating the level population of each species is important for determining the amount of line cooling, Λ_{line} , and more critically the amount of emission expected from a given species for a given transition line. The fraction of a species in a given quantum state depends on the properties of the cloud/zone, background CMB radiation, and the abundance of collision partners, x_p (e.g. x_{H_2} , x_{He} , etc.). Atomic and molecular data is downloaded automatically by DESPOTIC from the Leiden Atomic and Molecular Database (Schöier et al. 2005). This data includes information regarding transition energies, Einstein A coefficients (rates of spontaneous radiation), and collisional rate coefficients. There are two regimes under which level populations are calculated: optically thin and optically thick clouds. In both cases, it is under the condition of a state of statistical equilibrium. When optically thin, there is a balance between transitions into and out of each level. When optically thick, the effects of

the emitted radiation — now trapped inside the cloud — also need to be taken into account. This is handled by `DESPOTIC` using the escape probability approximation. In which, a volume-averaged probability of the radiation escaping the cloud can be calculated. This approximation requires additional terms when calculating the balance between transitions into and out of each level; full details are given in Krumholz (2014, Appendix B3).

5.3.4 Lookup table

In principle it would be possible to run `DESPOTIC` on each gas cell that is present within a given light cone; however, in practice this would be prohibitively slow. Instead we choose to produce a lookup table: a pre-calculated table of CO luminosity for a given range of input values, which we then use for interpolation. Once the lookup table is produced, this will dramatically speed up the production of synthetic CO emission maps but at the expense of accuracy. How accurate the interpolated CO predictions are will depend on the choice and number of input values (i.e. the coordinates in parameter space for which we explicitly calculate the CO luminosity). In what follows, we describe how the lookup table we produce can be optimised for our data.

We first make clear to what we are referring when talking about ‘our data’. We consider only `IllustrisTNG50` snapshots that are required to produce a light cone spanning the redshifts 1–4: these are the 30 snapshots that have a redshift from 1.0 to 4.0, inclusive. As already discussed in [Section 5.3.2](#)^[p. 142], we use the data of gas cells (that are members of a halo or subhalo) required to calculate the seven input quantities: Σ_{H} , ρ_{H} , Σ_{SFR} , Z , x_{C} , x_{O} , and x_{He} . The use of Z and Σ_{SFR} as scaling quantities (and only as scaling quantities) means that a value of zero would be problematic. In the process of optimising the lookup table, we ignore instances of $Z = 0$ and $\Sigma_{\text{SFR}} = 0$ but not the gas cell itself (i.e. we still retain the remaining quantities). When using the lookup table for interpolation, gas cells with $Z = 0$ or $\Sigma_{\text{SFR}} = 0$, will be capped to the lowest value of Z or Σ_{SFR} rather than being excluded.

As we have already noted in [Section 5.2](#)^[p. 128], snapshots are often used multiple times in order to span the required redshift range; in the description that follows, each time a snapshot is used we consider this as a new snapshot. This means we effectively have 91 snapshots (see [Section 5.2](#)^[p. 128] for details). These 91 snapshots are all treated in the same way, resulting in some data being counted or sampled multiple times. In the following text, ‘snapshot’ refers to one of the 91 and not of the 30 *unique* snapshots.

To produce an optimised lookup table we must know the distribution of each quantity so that we can preferentially sample the parameter space where most

5.3 Synthetic CO emission

of the data lies. Ideally we would produce a distribution of each quantity for all of the data; however, the amount of data we are using makes this infeasible. Instead we take the following steps.

We first find the global limits of all our data for the each of the seven input quantities. The quantities that are calculated from the gas density (Σ_{H} , ρ_{H} , and Σ_{SFR}) also need to be converted from comoving to proper lengths and so when finding the minimum and maximum of these quantities we need to consider the minimum and maximum redshift for which the gas cell could be placed within the light cone. These global limits define the bounds of the parameter space for which we can interpolate CO emission data.

We then use kernel density estimation to produce an estimated PDF for each of the seven input quantities for each snapshot file (in IllustrisTNG50 a single snapshot is made up of 680 files), which are then combined to give an estimated PDF for all the data. The limits of the input quantities span several order of magnitude and so we construct the estimated PDF in decadic logarithmic space. For comoving quantities, we use the midpoint-redshift of the associated snapshot redshift range to convert into proper units. With kernel density estimation, we have the means to estimate the PDF of a quantity in a non-parametric way. The kernel density estimator, $\widehat{\mathcal{F}}$, of the n data points, x_i , with an unknown underlying distribution, \mathcal{F} , is formally given as (Parzen 1962)

$$\widehat{\mathcal{F}}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad \text{Equation 5.32}$$

where K is the kernel, and h the bandwidth. The kernel is a smooth function that gives the shape to each data point, which when combined produce the overall distribution of the data. See Figure 5.5^[p. 148] for an example of how $\widehat{\mathcal{F}}$ is constructed. We select a Gaussian kernel

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right). \quad \text{Equation 5.33}$$

The bandwidth determines by how much the final distribution is smoothed and plays a strong role in the appearance of the final distribution (see Figure 5.5^[p. 148]). There are standard methods for estimating an appropriate bandwidth; these tend to result in an over- rather than under-smoothed distribution but this is sufficient for our purpose. We select the bandwidth, h , based on Scott's rule (Scott 1992):

$$h = n^{\frac{-1}{d+4}}, \quad \text{Equation 5.34}$$

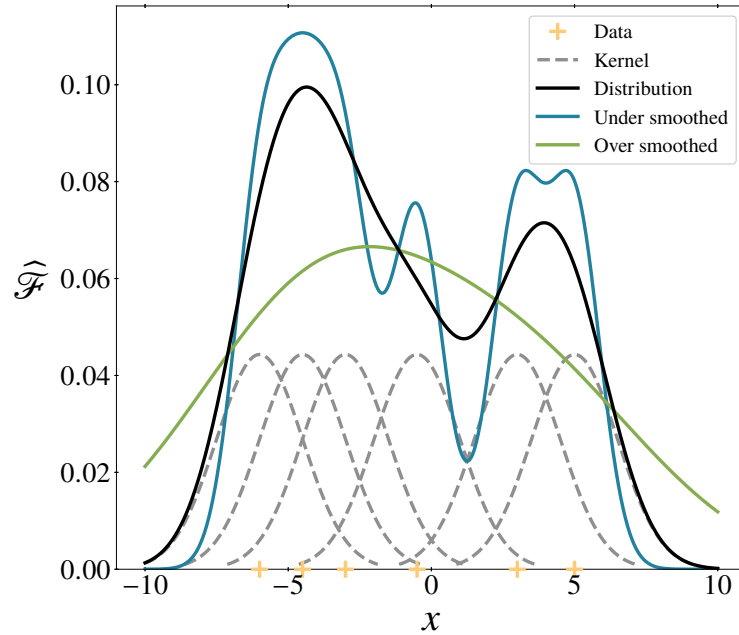


Figure 5.5

A demonstration of how kernel density estimation is used to convert data points into a distribution. Along the horizontal axis, we indicate with yellow crosses the six data points from which we wish to create a distribution. This is an unrealistically low number of data points but sufficient for demonstrating the kernel density estimation process. We use a Gaussian kernel (Equation 5.33^[p. 147]) to construct a distribution for each data point (grey dashed curves). By summing these individual distributions, we can construct the overall distribution of the data points (black curve). The appearance of the kernel depends on its functional form and the bandwidth. If the bandwidth is too small, the individual distributions become highly peaked and the final distribution can contain features that do not represent the data. This results in an under-smoothed distribution (blue curve). If, however, the bandwidth is too large, the individual distributions become very broad and the final distribution obfuscates features in the data. This results in an over-smoothed distribution (green curve).

5.3 Synthetic CO emission

where d is the number of dimensions of the distribution ($d = 1$), and as before n the number of data points. Since creating a distribution for each of the input quantities using all the data from just a single snapshot file is still impractical, we instead randomly sample 10^5 gas cells for each input quantity. This sample size is typically less than 1 per cent of the total number of available gas cells for any given snapshot file. With the sample size defined, $n = 10^5$, we calculate the bandwidth using Equation 5.34^[p. 147] to be $h = 0.1$.

The final estimated distribution, $\widehat{\mathcal{F}}_q$, for a given quantity q is calculated by performing a weighted sum of all the estimated distributions, $\widehat{\mathcal{F}}_{q,f}$, for each file, f :

$$\widehat{\mathcal{F}}_q(x) = \sum_f \left[\kappa_f \left(\int_{\min(q)}^{\max(q)} \widehat{\mathcal{F}}_{q,f}(u) \, du \right)^{-1} \right] \widehat{\mathcal{F}}_{q,f}(x) ,$$

Equation 5.35

where the integration over the global limits of the quantity q normalises $\widehat{\mathcal{F}}_{q,f}$ and κ_f weights the contribution of $\widehat{\mathcal{F}}_{q,f}$ to the sum. This weighting, κ_f , is the fraction of the total number of available gas cells across all snapshots and all files that are present in the current file, f . We use this weighting so that an estimated distribution representing more of the population of gas cells contributes more strongly to the final estimated distribution.

Now that we have the PDF for each of our input quantities, we can use them to define the coordinates in our parameter space for which we explicitly calculate the CO luminosity. That is to say, we define the ‘rows’ and ‘columns’ of our 8D (seven above mentioned input quantities and redshift) lookup table. We choose to select 20 values for each quantity that span the global limits of that quantity; this results in a lookup table with 20^8 elements. Without knowledge of the input quantity PDFs, these 20 values would have been uniformly distributed (in logarithmic space) between the global limits; the parameter space would have been sampled by a uniform grid. This would mean that the distance between interpolation points would be uniform and that all the data would be subject to the same inaccuracies as a result of interpolation. We instead use the quantity PDFs to sample more frequently where more of the data lies; the parameter space is sampled by a non-uniform grid. Now that the spacing of interpolation points depends on the number of data in that region of parameter space, the popular regions have a higher number of interpolation points and so are subject to smaller inaccuracies as a result of interpolation when compared to less-popular regions. That is, we minimise the impact of interpolation for where the majority of data lie. The above discussion excludes redshift, which is uniformly sampled (in linear-space), since the gas cells do not have an inherent redshift: this depends on their location within the light cone.

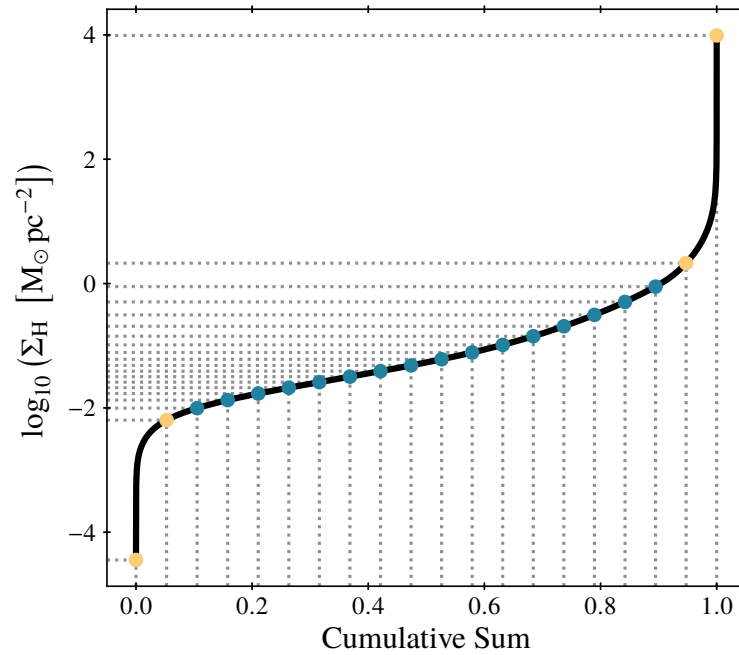
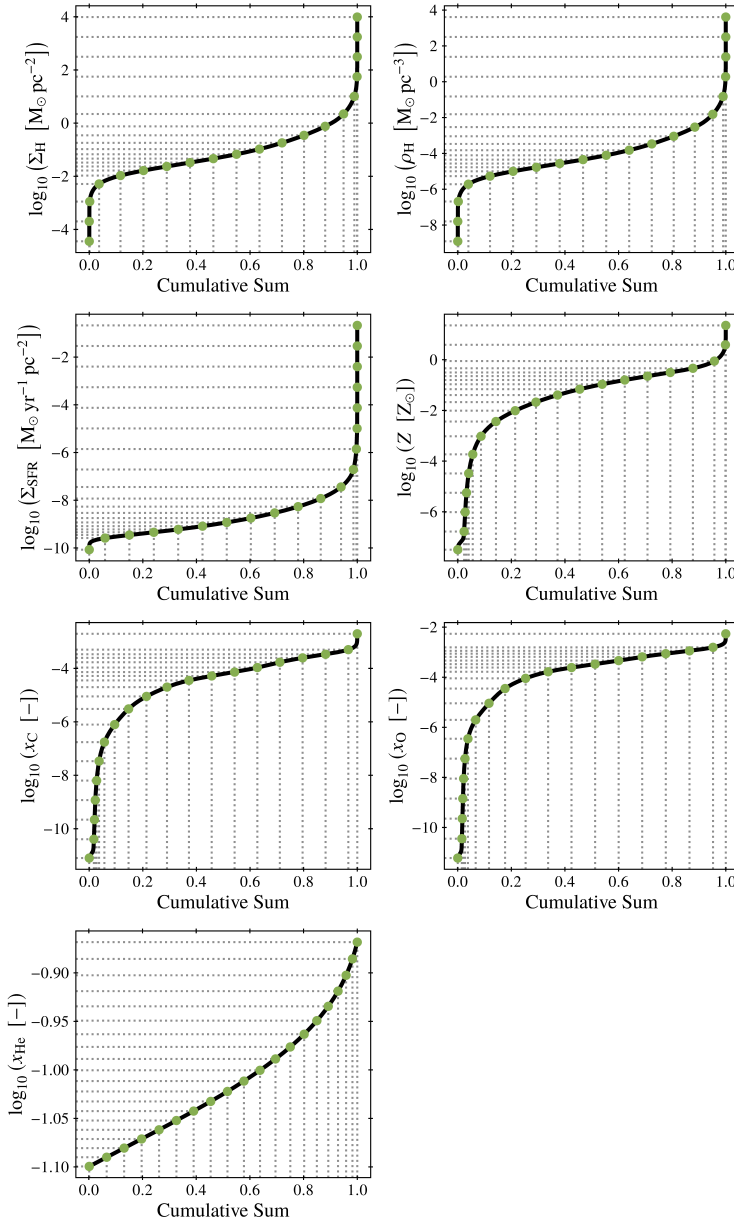


Figure 5.6

We demonstrate here, how using inverse transformation sampling can result in poor coordinate selection at the edges of the distribution. To specify coordinates for the lookup table (e.g. values of Σ_{H}) using inverse transformation sampling, we select 20 uniformly-spaced points in the range $[0, 1]$ (vertical grey dashed lines) and pass them through the inverse of the CDF (black solid curve). The corresponding values of Σ_{H} (horizontal grey dashed lines) are the selected coordinates. The coordinates are distributed such that they are closer together in regions where many parameter values lie, as intended. However, at the tails of the distribution (vertical line segments at zero and one) the selection is poor because several orders of magnitude are covered by just two points (marked in yellow). This would likely produce inaccurate results when interpolating between these two values.

5.3 Synthetic CO emission



We show the coordinate selection for the seven simulation-based quantities used in the lookup table. The remaining quantity, redshift, is excluded as this is not an inherent property of the gas cells and therefore a distribution cannot be created. The coordinates are selected based on uniformly spaced points along the inverse of the CDFs. The top left panel is the same as that seen in Figure 5.6^[p. 150]; however, the coordinate choice is now improved.

Figure 5.7

5 The synthetic ALMA UDF

We initially intended to use inverse transformation sampling to sample the quantity PDFs to select the coordinates of the lookup table. This is performed by first selecting 20 uniformly-spaced points in the range $[0, 1]$. These points are then passed through the inverse of the cumulative distribution functions (CDFs) (which also span the range $[0, 1]$). The result is 20 coordinate values for each quantity that are spaced such that the coordinates are concentrated in regions where more of the data lies. Using inverse transformation sampling fulfils our aims; however, coordinate selection is poor in the tails of the PDF and could result in a difference between consecutive coordinates of several orders of magnitude. In [Figure 5.6](#)^[p. 150], we demonstrate how this method results in poor coordinate selection for Σ_{H} . We instead chose to select points that are equally spaced along the CDFs, which alleviates the problem of poor coordinate selection and still retains a higher sampling frequency in regions where more of the data lies. We show the coordinate selection for the input quantities (excluding redshift) in [Figure 5.7](#)^[p. 151].

In principle, the above process would be sufficient to produce a useable lookup table; however, it is wasteful since there are many regions of the parameter space that are unphysical and many of the quantities are correlated. For example, it would be very improbable that a region would have a high metallicity but low Carbon, Oxygen, or Helium abundance. The density quantities are also strongly correlated, since they are derived from the same quantities. We instead make an effort to find the regions of the lookup table that would be used and only calculate CO luminosities for these regions.

In [Figure 5.8](#)^[p. 153], we illustrate the process we take to find the required regions of the lookup table for a simplified 2D version. In the figure, [Table 1](#) lists all the available gas cells that could possibly be used within a light cone. The first column, listing quantity-A values, represents quantities like Z and x_i : these are quantities that have a fixed value. The second column, listing quantity-B values, represents the comoving quantities (Σ_{H} , ρ_{H} , and Σ_{SFR}): these quantities could have a range of possible values which is bounded by the minimum and maximum redshift the gas cell could be placed in a light cone. It is therefore necessary to consider B-like quantities as spanning a range.

The second table, [Table 2](#), illustrates the simplified 2D ‘selection’ table. The selection table has the same shape as the lookup table but its purpose is to count the number of gas cells that are in this region of the parameter space. The rows give the quantity-A coordinates and the columns the quantity-B coordinates. These coordinates are like those selected in [Figure 5.7](#)^[p. 151].

Starting with the first gas cell in [Table 1](#), we take the following steps. The quantity A has a single value (21.4) and so we find the two coordinates between which the value falls (21–22). These are the two coordinates which would be

5.3 Synthetic CO emission

Table 1

A	B	
	Min	Max
21.4	11.7	13.3
22.7	10.1	11.8
24.0	13.6	15.0
23.7	12.5	14.3
20.8	10.2	13.1

Table 2

A \ B	10	11	12	13	14	15
20						
21		+1	+1	+1	+1	
22	+1	+2	+2	+1	+1	
23	+1	+1	+1			
24				+1	+1	+1
25						

A schematic showing the construction of a selection table, described in [Section 5.3.4](#)^[p. 146]. In Table 1 is listed all the available gas cells. In this example, gas cells have two relevant quantities, A and B. Quantity A is redshift independent; however, quantity B is not and so there is a minimum and maximum value. The selection table is shown in Table 2. The colours of the regions within Table 2 correspond to the colours of the rows in Table 1. These coloured regions are the regions of the parameter space in which a given gas cell falls. These cells are incremented by one ('+1'). Regions that overlap result in a higher count (e.g. '+2').

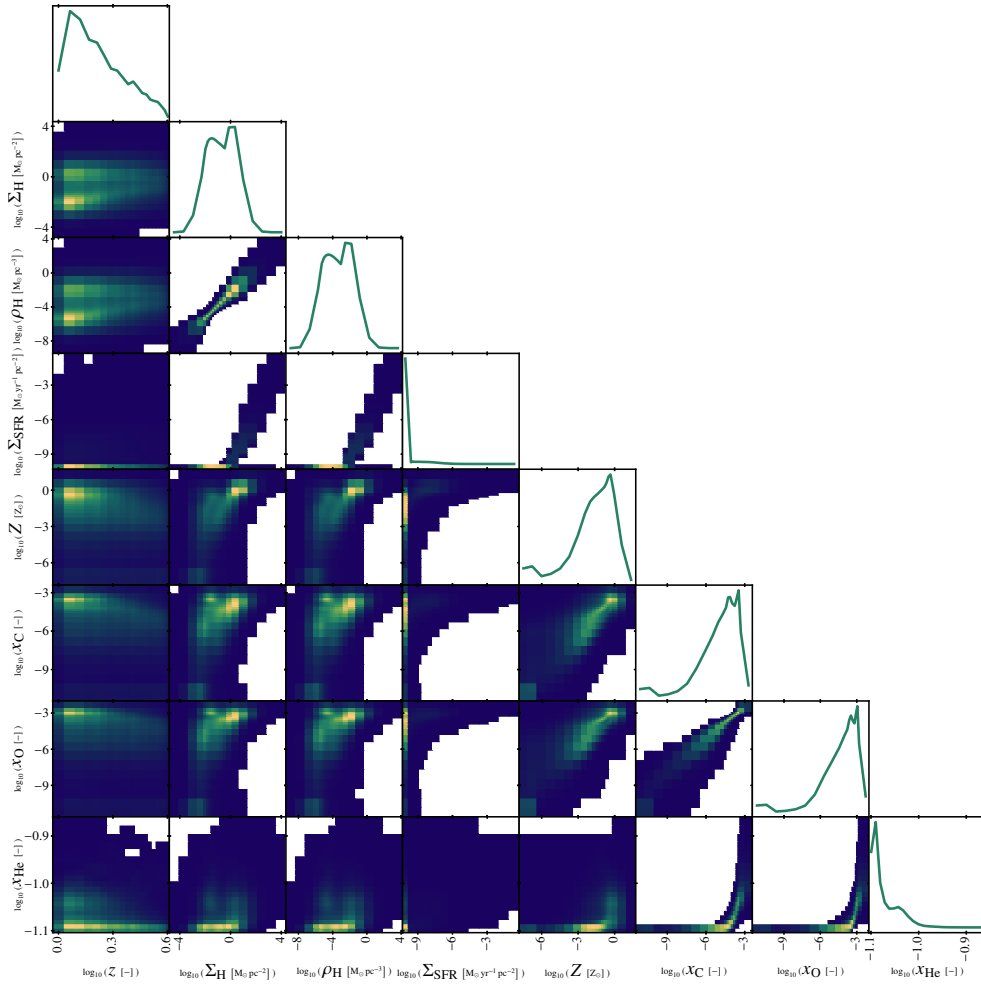
Figure 5.8

5 The synthetic ALMA UDF

used when interpolating along the A axis. The quantity B has a minimum (11.7) and maximum (13.3) value and so the process we use for A must be performed twice: (11–12) and (13–14), respectively. However, B depends on redshift and could take any value between the minimum and maximum; the required lookup table elements must therefore span the full range (11–14), since all the coordinates in this range could be required when interpolating along the B axis. We can now specify the corners of the rectangle that define the region of the lookup table that could be required for interpolation when considering this gas cell: (21,11), (21,14), (22,11), (22,14). We add one to all the elements within this region of the selection table. This process is repeated for every possible gas cell within the data. In the unlikely event that a quantity value matches the coordinate, we simply select that coordinate (as demonstrated in the third row of Table 1).

Of course, our lookup table has to account for eight quantities and so the region of the lookup table required by any given gas cell is an 8D cuboid with 2^8 corners. In Figure 5.9^[p. 155], we show a ‘triangle plot’ (an array of normalised 1D and 2D projections) of our 8D selection table. The 1D distributions along the diagonal show an approximate PDF of each quantity calculated for all the data. The most peculiar of these distributions is seen for the SFR surface density, Σ_{SFR} , where most of the gas cells are at $\sim 10^{-10} \text{ M}_{\odot} \text{ yr}^{-1} \text{ pc}^{-2}$. This is the effect of having previously excluded $\Sigma_{\text{SFR}} = 0 \text{ M}_{\odot} \text{ yr}^{-1} \text{ pc}^{-2}$ and now being capped to the lowest value, which we discussed at the beginning of Section 5.3.4^[p. 146]. The 2D projections demonstrate the correlations between quantities as well as highlighting regions of the parameter space where no gas cells lie. Another feature highlighted by the 2D plots are the varying grid sizes, which is a result of the coordinate selection discussed above (see Figure 5.7^[p. 151]). After constructing this selection table, we find that of all the 20^8 elements within the lookup table, only 0.06 per cent of the table would be used.

5.3 Synthetic CO emission



The figure shows projections of the 8D selection table calculated from IllustrisTNG50 data. This table is the true version of the schematic shown in Figure 5.8^[p. 153]. The diagonal shows the approximate PDF of each quantity based on all the data. Each of the 2D projection shows a normalised distribution, in each plane of the 8D space. The projection plots reveal correlations between related quantities (i.e. ρ_H , Σ_H , and Σ_{SFR} and Z , and the x_i) and also show regions of the parameter space that are not used (in white). The cell sizes in the 2D projections are not uniform, which is a result of the coordinate selection shown in Figure 5.7^[p. 151].

Figure 5.9

5.4 Synthetic CO observations

We present here the results of our synthetic CO observations based on a single synthetic light cone. This light cone contains the gas found within a volume corresponding to a $2.5' \times 2.5'$ field of view and redshift range 1–4. This volume was pieced together from smaller volumes extracted from the IllustrisTNG50 simulation snapshot files. We use the properties of the selected gas to predict the expected CO emission through the interpolation of a lookup table which we created using the code `DESPOTIC`.

In [Figure 5.10](#)^[p. 158], we present a high angular resolution CO ($1 \rightarrow 0$) emission map of the light cone on which we focus our analysis. This image was created at an angular resolution of $0.005'' \times 0.005''$ and smoothed using a Gaussian filter with a FWHM of the target resolution $0.01'' \times 0.01''$. This creates a detailed image showing the distribution of the CO emission in the galaxy population but also within individual galaxies themselves. The detail that we see in the galaxies is due to the high spatial resolution of the IllustrisTNG50 simulation, showing that the simulation is indeed able to recover galactic structure (e.g. spiral arms).

We calculate the CO emission shown in [Figure 5.10](#)^[p. 158] by interpolating a lookup table whose construction we described in [Section 5.3.4](#)^[p. 146]. This interpolation is based on eight input quantities: redshift, z ; Hydrogen surface density, Σ_{H} ; Hydrogen volume density, ρ_{H} ; SFR surface density, Σ_{SFR} ; metallicity, Z ; and the Carbon, Oxygen, and Helium abundances, x_{C} , x_{O} , and x_{He} . In [Figure 5.11](#)^[p. 159], we present the line-of-sight average of these quantities at the locations where we also observe CO ($1 \rightarrow 0$) emission, allowing one to make a direct comparison with [Figure 5.10](#)^[p. 158]. The quantity maps we present in [Figure 5.11](#)^[p. 159] are created at the target resolution ($0.01'' \times 0.01''$) and do not go through the same smoothing process as we described for the synthetic CO map. This prevents the introduction of numerical artefacts when smoothing with zero-value pixels: these pixels do not have a physical meaning in this context. Since we do not apply the smoothing process, the maps (under close inspection) appear pixelated in a way not seen in [Figure 5.10](#)^[p. 158]. These property maps provide a better understanding of the synthetic CO emission map. Using the redshift map, we can see if neighbouring galaxies (in the 2D projection) have matching redshifts or not, allowing us to determine if the galaxies are interacting or simply in the line of sight. A detailed look at the Σ_{H} , ρ_{H} , Σ_{SFR} , and Z maps, reveals variations of these quantities within the individual galaxies. Visually, the variation is as expected: in interarm regions these quantities have lower values than in the spiral arms, conforming to our understanding that spiral arms have higher gas densities and are the birth locations of most stars. The chemical abundance

5.4 Synthetic CO observations

maps do not show variations within a galaxy. The uniformity of the chemical composition within a galaxy is unrealistic and a direct consequence of the procedure we use to assign chemical abundances to individual gas cells. We describe this procedure in [Section 5.3.2](#)^[p. 142], explaining that not all of the IllustrisTNG50 simulation snapshots save chemical abundance information. Therefore, gas cells inherit the chemical abundance of its parent galaxy, which is saved in galaxy catalogues accompanying every snapshot. This also means that these maps can effectively be used to locate individual galaxies identified in the galaxy catalogues.

With current observational facilities, the angular resolution used to create [Figure 5.10](#)^[p. 158] is not possible. The ASPECS survey has an approximate angular resolution of $1'' \times 1''$, and so we recreate [Figure 5.10](#)^[p. 158] with a more realistic resolution in [Figure 5.12](#)^[p. 160] (these two figures are aligned for easy comparison). We again use the smoothing process described above, starting with an inherent resolution of $0.1'' \times 0.1''$ and smoothing to $1'' \times 1''$. As expected, all detail of galaxy structure is lost and we are left with amorphous spots; each spot corresponding, approximately, to an individual galaxy.

The synthetic CO emission map we present in [Figure 5.12](#)^[p. 160] is informative, in that we can see the raw data, but unuseful for a basis of comparison. The data contained within the image must be characterised in a statistical way. In the following, we describe the initial results from our analysis of this synthetic light cone.

In [Figure 5.13](#)^[p. 163], we show how the total CO luminosity density changes as a function of redshift for several CO transitions. This figure is constructed in the following way. First, we define the redshift ranges into which we divide the light cone. The limits of redshift range i , $[z_i^{\min}, z_i^{\max}]$, are defined through

$$\Delta \tilde{D} = \frac{c}{H_0} \int_{z_i^{\min}}^{z_i^{\max}} \frac{1}{E(z)} dz, \quad \text{Equation 5.36}$$

where we select the change in depth, $\Delta \tilde{D}$, to be 500 cMpc. The measurement made from this redshift range is assigned to the midpoint redshift, z_i^{mid} , given as

$$z_i^{\text{mid}} = \frac{z_i^{\max} + z_i^{\min}}{2}. \quad \text{Equation 5.37}$$

The choice of $\Delta \tilde{D}$ is a compromise between minimising the deviation from z_i^{mid} and having a representative volume to avoid local fluctuations. The next redshift range begins at z_{i+1}^{\min} calculated through

$$\Delta \tilde{O} = \frac{c}{H_0} \int_{z_i^{\min}}^{z_{i+1}^{\min}} \frac{1}{E(z)} dz, \quad \text{Equation 5.38}$$

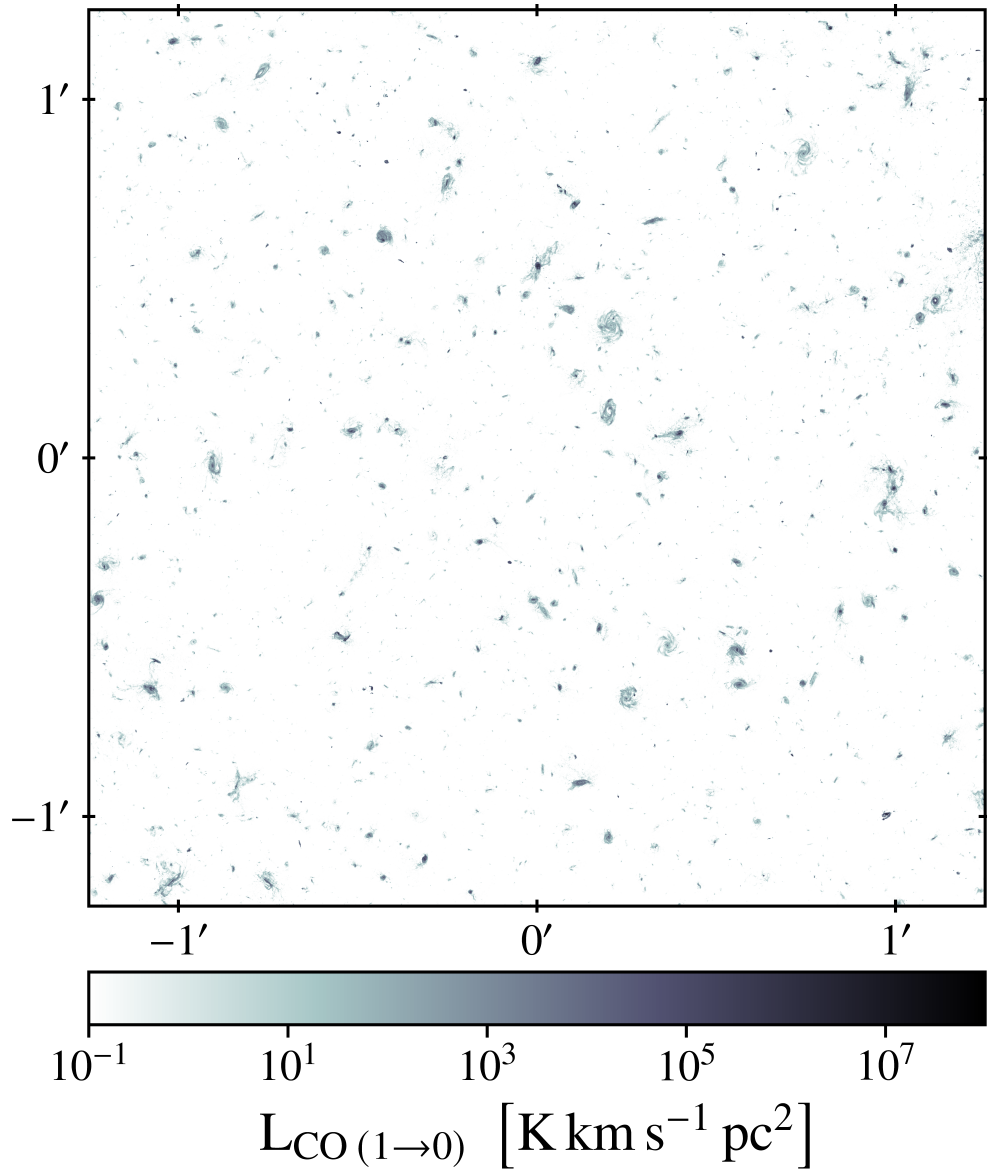
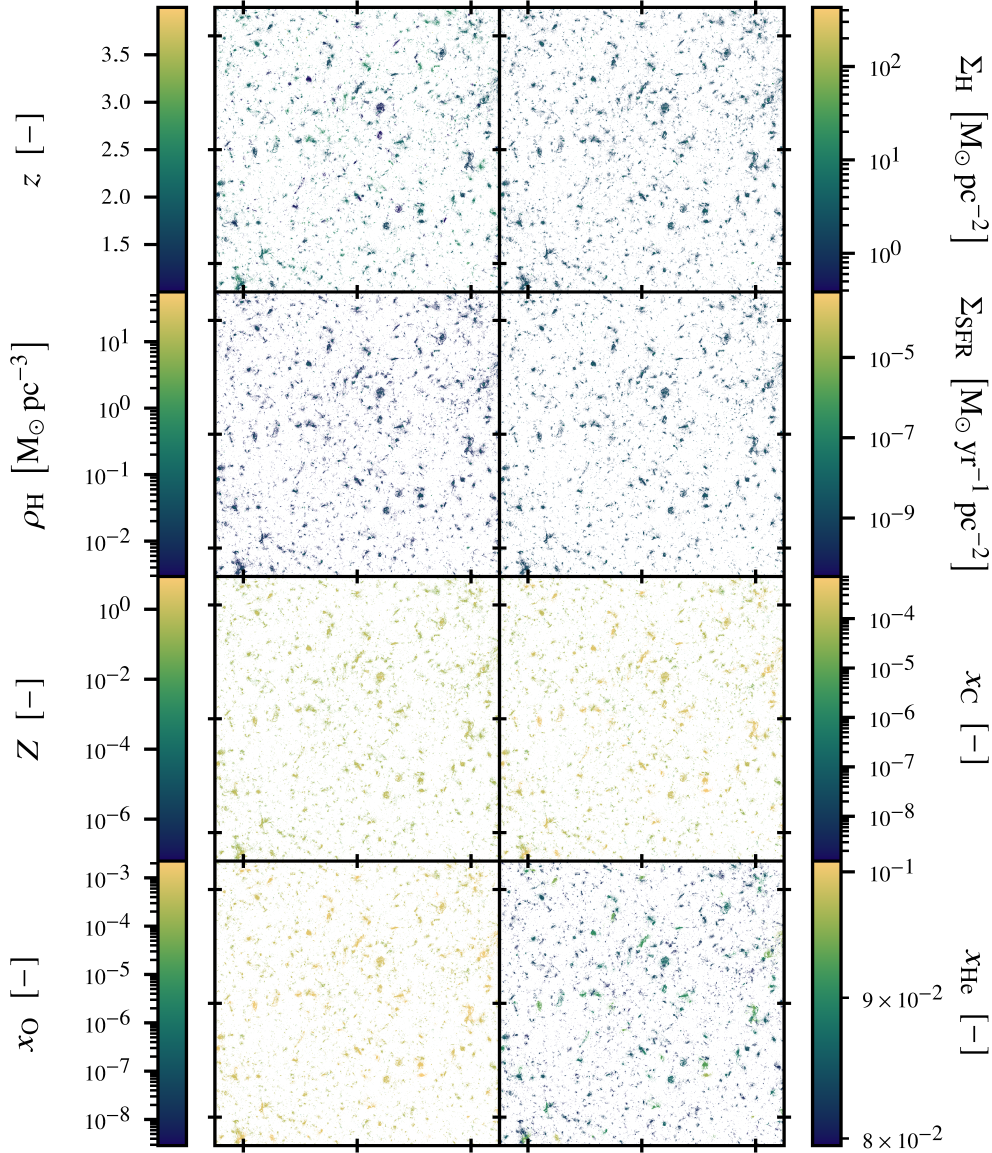


Figure 5.10

A synthetic CO(1 → 0) emission map based on a light cone constructed from the IllustrisTNG50 simulation. We show the $2.5' \times 2.5'$ field of view at $0.01'' \times 0.01''$ resolution. At this resolution, we can see the distribution of CO emission (or equivalently molecular Hydrogen) within galaxies and across the galaxy population. This image also demonstrates the high spatial resolution that makes the IllustrisTNG50 simulation unique.

5.4 Synthetic CO observations



Maps showing the line-of-sight average of the eight quantities used as input for interpolating CO emission. These maps have an angular resolution of $0.01'' \times 0.01''$ and only show data where there is CO ($1 \rightarrow 0$) emission (as seen in Figure 5.10^[p.158]). The redshift map, z , allows us to see which galaxies are interacting (i.e. at the same redshift) and which are at a similar location on the sky. In the Hydrogen surface density (Σ_{H}), Hydrogen volume density (ρ_{H}), SFR surface density (Σ_{SFR}), and metallicity (Z) maps we can see the distribution of these quantities within the individual galaxies. As we described in Section 5.3.2^[p.142], the chemical composition assigned to an individual gas cell is that of the subhalo; the consequence of this can be seen in the chemical abundance maps, x_i , where each galaxy appears as a uniform colour.

Figure 5.11

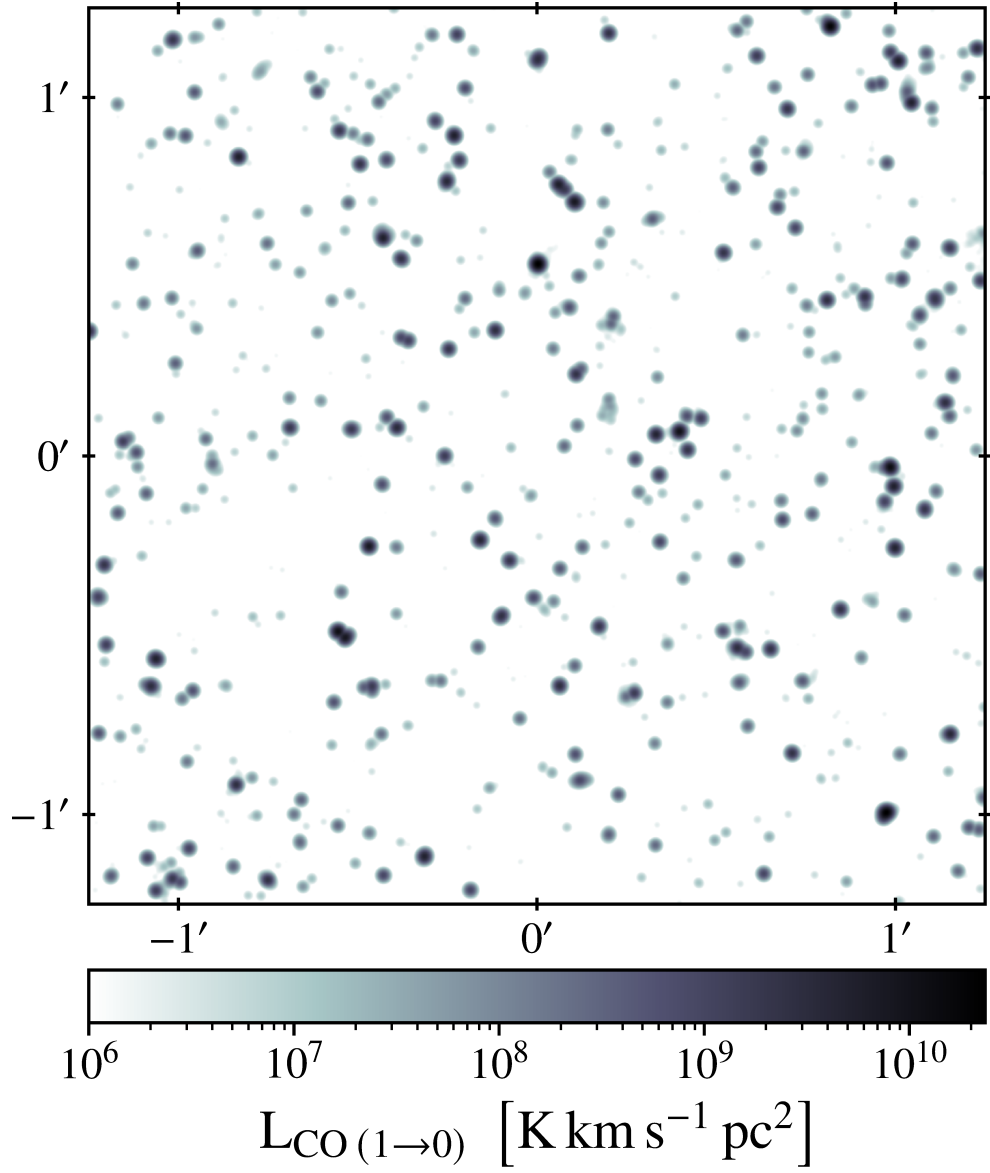


Figure 5.12 We present here the same image as shown in [Figure 5.10](#)^[p. 158]; however, at the lower resolution of $1'' \times 1''$. This resolution is comparable to that of real observations and the resolution we use in our analysis.

5.4 Synthetic CO observations

where the offset, $\Delta\tilde{O}$, is selected to be 50 cMpc. The smaller $\Delta\tilde{O}$ becomes the greater the number of redshift ranges, n , we define and the more measurements we make — starting at $z_1^{\min} = 1$ and ending at $z_n^{\max} = 4$. Having $\Delta\tilde{O} < \Delta\tilde{D}$ means that consecutive ranges are not independent and have an overlap of $\Delta\tilde{D} - \Delta\tilde{O} = 450$ cMpc.

The CO luminosity density, J_{CO} , in redshift range i is calculated using

$$J_{\text{CO}}(z_i^{\min}, z_i^{\max}, \phi) = \frac{A_p(z_i^{\text{mid}}) \times \sum_{i,j} E_{ij}}{V(z_i^{\min}, z_i^{\max}, \phi)}. \quad \text{Equation 5.39}$$

The summation term gives the total amount of CO emission found within the volume V ; however, this is not the sum of individual gas cells but of pixels in the emission map, E . This emission map is the projection of the contents of the light cone that falls within the redshift range onto a $0.1'' \times 0.1''$ resolution grid (with field of view $\phi = 2.5' \times 2.5'$), which is then smoothed using a Gaussian filter to a resolution of $1'' \times 1''$ — the same process used to generate [Figure 5.12](#)^[p. 160]. We perform this smoothing procedure, as it is necessary to calculate the area over which this emission is emitted. The area is calculated using the area per pixel, A_p , which is determined for the midpoint redshift.

Since the CO ($1 \rightarrow 0$) emission can be converted in to a molecular Hydrogen mass through a simple conversion factor, $\alpha_{\text{CO}} = 3.6 M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1}$ (Daddi et al. 2010), we also include the change in molecular Hydrogen density as a function of redshift in [Figure 5.13](#)^[p. 163].

For this light cone the global maximum in CO ($1 \rightarrow 0$) emission (and equivalently ρ_{H_2}) occurs around redshift 1.5. Although a second peak in emission occurs around 2.6, which is where higher CO transitions appear to peak. There is a drop off in emission for all the presented CO transitions at approximately redshift 3.0. The specific features of the relations shown in [Figure 5.13](#)^[p. 163] are particular to the light cone which we analyse. An improved measure of the relation can be recovered by averaging the $J_{\text{CO}}-z$ relations calculated from hundreds of different light cones. This averaging of the relation would also place constraints on the variation in the relation which could be expected from real observations. A sufficient number of real observations to quantify this variation in the relation is unlikely to ever exist and so can only be characterised from synthetic observations. With the expected variation in the $J_{\text{CO}}-z$ relation quantified, a more meaningful comparison to the ASPECS results could be performed. However, from a qualitative comparison with our single light cone, we can see that the ASPECS measurements recover approximately two orders of magnitude less emission for a given redshift from what we present. This discrepancy can be easily

5 The synthetic ALMA UDF

understood when one considers that our synthetic observations are without extinction and sensitivity limits. The overall shape of the distribution does broadly agree.

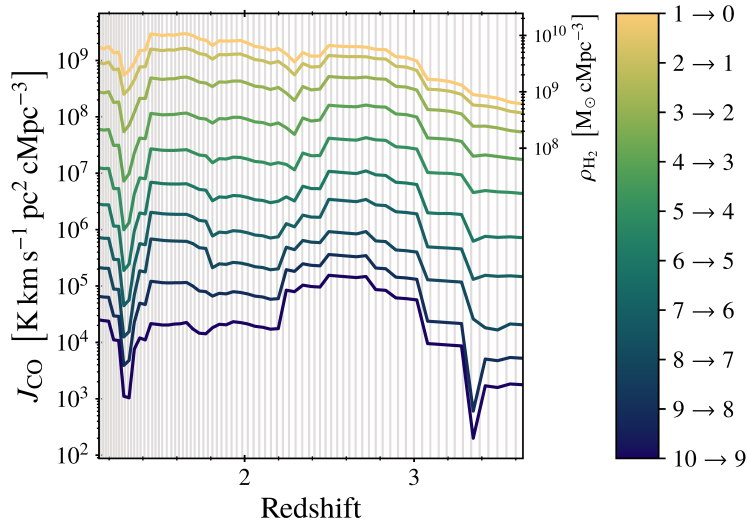
In [Figure 5.14](#)^[p. 164] we present the CO luminosity function, Φ_{CO} , for the first four CO transitions and for a range of different redshifts. Again we include the conversion of CO (1 \rightarrow 0) emission into a molecular Hydrogen mass, which gives us the distribution of molecular Hydrogen mass content. To construct these luminosity functions, we first produce a galaxy catalogue for our light cone containing the galaxy's total CO emission (L_{CO}) and redshift. The redshift of a given galaxy, i , is calculated as the mean redshift of all the gas associated with that galaxy, \bar{z}_i . The galaxy to which a given gas cell is assigned is determined from the original IllustrisTNG50 galaxy catalogue. The CO emission for galaxy i is calculated as $A_p(\bar{z}_i) \times \sum_{i,j} E_{ij}$. As before, E is a $1'' \times 1''$ resolution emission map; however, this map contains just a single galaxy. The way in which we define the redshift ranges (to produce luminosity functions at different redshifts) is as before, using [Equations 5.36](#) and [5.37](#)^[p. 157], but with $\Delta\tilde{D} = 1000$ cMpc and $\Delta\tilde{O} = 250$ cMpc. We made these choices to ensure a sufficiently large population of galaxies within the redshift range and to avoid overcrowding of the figure.

The figure shows that there is indeed a redshift dependence on the luminosity function. The function changes in two key ways. The first way is in the number density: we find that the luminosity function sits higher in the parameter space for higher redshifts. This conforms with what we observed in [Figure 5.13](#)^[p. 163], where there is, overall, more CO emission at higher redshifts. The second change is the point at which the sharp drop off in emission occurs. Although this drop off is approximately between $L_{\text{CO}} = 10^{10} - 10^{11}$ K km s⁻¹ pc² for all transitions and all redshifts, it is clear that lower redshifts are at the lower end of this range.

When comparing with luminosity functions produced from the ASPECS observations, we find that our light cone contains galaxies with a higher (up to an order of magnitude) total CO luminosity than the brightest galaxies in ASPECS. The inclusion of extinction and sensitivity limits would introduce a horizontal shift of the luminosity functions to lower L_{CO} , reducing this discrepancy. Only through an ensemble average of light cones and the inclusion of emission-reducing mechanisms could we perform a true comparison to the observations. However, ASPECS finds a variation of ~ 1 dex in the luminosity at which the drop off occurs (as determined by the L_{\star} parameter of the Schechter luminosity functions they constrain); we also commented on this finding in the previous paragraph.

From the results we present here we can conclude that we observe higher

5.4 Synthetic CO observations



We show the redshift dependence of the total CO emission for the first ten CO emission lines (indicated on the colour bar). This data is based on the synthetic light cone shown in [Figure 5.12](#)^[p. 160]. The vertical axis on the right-hand side converts the CO (1 → 0) emission into a molecular Hydrogen mass through the conversion factor $\alpha_{\text{CO}} = 3.6 M_{\odot} (\text{K km s}^{-1} \text{pc}^2)^{-1}$ (Daddi et al. 2010). We note that this conversion is only valid for the CO (1 → 0) curve. The vertical grey lines indicate the redshift boundaries of the individual cut-out volumes, which are also seen in [Figure 5.2](#)^[p. 134].

Figure 5.13

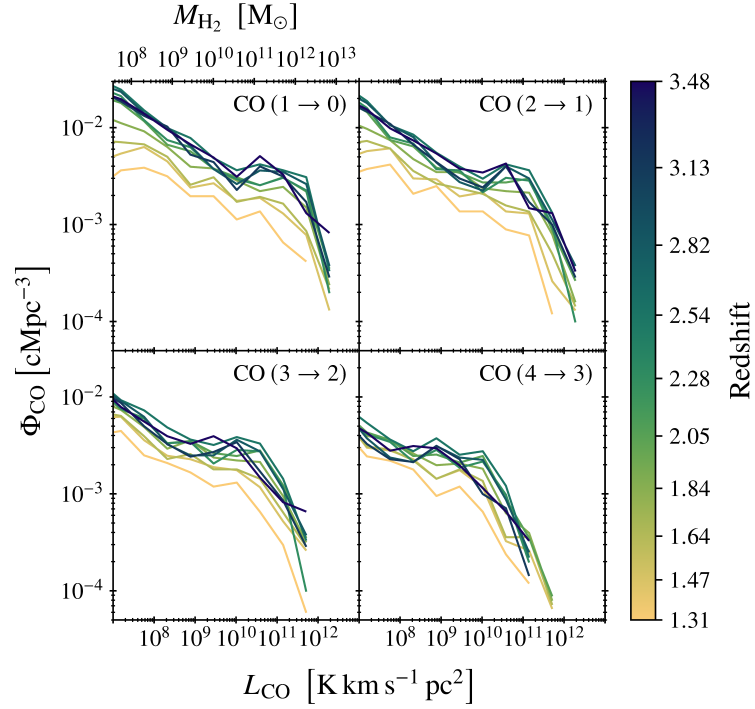


Figure 5.14

We show the luminosity function for the first four CO emission lines (indicated in each panel) based on the synthetic light cone shown in [Figure 5.12](#)^[p.160]. The multiple curves in each panel correspond to different redshift bins. The midpoint redshift of each bin has been used to colour the curve and is explicitly marked on the colour bar. The CO (1 → 0) emission can be converted into a molecular Hydrogen mass through the conversion factor $\alpha_{\text{CO}} = 3.6 M_{\odot} (\text{K km s}^{-1} \text{pc}^2)^{-1}$ (Daddi et al. 2010). We have added an additional axis on the CO (1 → 0) panel to show the molecular Hydrogen mass.

levels of CO emission when compared to the observations of ASPECS. This in turn leads to higher amounts of predicted molecular Hydrogen. The results we present here can be improved upon in two ways. The first is to produce statistical relations based on the ensemble average of hundreds of light cones: this would smooth out any features that are specific to a given light cone and also constrain the expected fluctuations in the relations as a result of field-to-field variation. The second improvement is to include effects experienced by observations. That is, to include the effects of extinction and of sensitivity limits. These ‘observational corrections’, would give a better basis for comparison but could also reveal the extent to which the ASPECS survey is altered by these effects. With these improvements, comparisons to the observations would reveal if the synthetic CO observations we produce agree with real observations. Discrepancies between the two could stem from many sources: the IllustrisTNG50 models could indeed produce incorrect gas density distributions, the assumptions and scaling relations we use to predict CO emission could be inadequate, or there are additional mechanisms effecting the ASPECS observations that have not been corrected for. Future work from both synthetic and true observations, will in time uncover the answers.

5.5 Conclusions

In this chapter we create and study synthetic CO observations. This emission traces the dense gas of the ISM, the medium from which stars form and into which stellar feedback injects energy, momentum and matter. This medium is also sensitive to different stellar feedback models and so by studying the CO emission of a cosmologically representative volume and over cosmic time, we can come to learn more about the relevant physical mechanisms of stellar feedback.

The synthetic observations we produce are based on the IllustrisTNG50 simulation — a gravity+magnetohydrodynamical cosmological simulation based on the moving mesh code AREPO. This simulation is the first in a new class of cosmological simulations, aiming to overcome the traditional compromise between simulation volume and resolution. IllustrisTNG50 simulated a volume of 51.7^3 cMpc^3 and has a median radius of $\sim 100\text{--}140 \text{ pc}$ in star-forming gas cells (over cosmic time). The resolution is sufficient to recover detailed galaxy structure including spiral arms, bulges, and bars.

The observations we produce are designed to mimic those of ASPECS. This is an ALMA sky survey studying the CO emission in the Hubble XDF, an area covering $\sim 4.6 \text{ arcmin}^2$. The redshift range captured by this survey is from 1 to 4. To construct observations with these specifications we must first construct a

5 *The synthetic ALMA UDF*

synthetic light cone. We developed the code `LIGHTCONE` for this purpose. The code pieces together volumes cut out of the simulation snapshots to produce a complete volume corresponding to a $2.5' \times 2.5'$ field of view and redshift range 1–4. The code performs the following steps:

1. each snapshot is assigned a redshift range within the light cone which minimises the deviation from the snapshots inherent redshift
2. the ‘cut-out volume’ is defined for each snapshot with the height and width corresponding to the field of view for the given redshift range and a depth that spans the redshift range (i.e. a square frustum in comoving coordinates)
3. gas cells that fall within this cut-out volume are selected for use in the light cone.

In the final step, care is taken to reduce possible artifices that result from imaging the same volume at different stages of its evolution: the code randomly rotates the simulation volume about its three axes and the cut-out volume is randomly centred.

The CO emission expected from the selected gas cells is predicted using `DESPOTIC` — a chemical, thermal, and statistical equilibrium code which calculates the state of interstellar clouds and their observable emission. From the simulation we provide `DESPOTIC` with the Hydrogen volume and surface density; the local SFR surface density, for scaling an assumed background radiation field; the abundance of Helium, Carbon, and Oxygen; and the metallicity, for scaling assumed dust properties. The position within the light cone provides the redshift. In principle, we could run `DESPOTIC` on all the selected gas cells, but in practice this would be computationally prohibitive. We instead produce an optimised lookup table. This lookup table is an 8D array, one dimension for each of the properties listed above, with 20 coordinates along each dimension. The choice in coordinates is such that they are closer together in regions of the parameter space where more of the gas cells lie, giving more accurate results for the majority of gas cells. A lookup table with 20^8 elements could not be generated in full; however, correlations between the various quantities mean that many regions of the table would never be used (e.g. low Hydrogen volume density with high Hydrogen surface density). In fact, we find that at most only 0.06 per cent of the table would ever be used. Using the properties of the gas cells that form the light cone, we interpolate the lookup table to recover the CO emission.

We present the results based on a single synthetic light cone. A high resolution image of the field of view reveals the detailed galaxy structure, as

mapped out by CO emission. This detailed galaxy structure is visible because of the high spatial resolution of the IllustrisTNG50 simulations. Unfortunately, true observations would be unable to see this detail and all further analysis is carried out on $1'' \times 1''$ resolution images: this is a resolution comparable to that of ASPECS. We show both the variation of total CO emission as a function of redshift (for different CO transitions) and CO luminosity functions (for different CO transitions and at different redshifts). These relations were also converted into molecular Hydrogen mass content through an α_{CO} conversion factor. The forms of these distributions are not unreasonable; however, qualitative comparisons to the ASPECS observations show that we over predict the amount of emission and therefore molecular Hydrogen content. This discrepancy could be reduced by the inclusion of sensitivity limits and the effects of extinction. The best comparisons will be made based on statistical relations calculated from the average of an ensemble of light cones. The relations calculated in this way would be free from the particular features of a given light cone and give the range within which variation could be expected. Future work, based on the improved statistical relations could highlight possible inadequacies in the IllustrisTNG50 models, our assumptions in prediction CO emission, or in the real observations themselves.

Acknowledgements

We thank Annalisa Pillepich, Roberto Decarli, Fabian Walter, Gergö Popping, and Desika Narayanan for helpful discussions in the early stages of this project. We are grateful to the IllustrisTNG Collaboration for granting us access to the IllustrisTNG50 simulation data. This research was undertaken with resources from the ISAAC supercomputer at the Max Planck Computing and Data Facility (MPCDF).

Conclusion

The aim of this work was to improve our understanding of star formation. This is, of course, a massive undertaking that cannot be achieved in a single thesis; however, though a collective effort of many research groups and collaborations across the world, for many decades past and for many decades to come, this is a goal that will be achieved. Therefore, in this work we present our findings, if not complete answers, that will in time help towards this objective.

In [Chapter 3](#)^[p. 47] and [Chapter 4](#)^[p. 99], we study the nature of SFR tracer emission lifetimes, which in combination with the KL14 principle, can place constraints on the evolutionary timeline of star formation. This includes constraints on molecular cloud lifetimes, the duration of cloud destruction by feedback, and the integrated cloud-scale star formation efficiency.

The KL14 principle (with associated code, HEISENBERG), is a new statistical method which can measure the relative durations of different evolutionary phases of the star formation process from emission maps tracing these phases. This is achieved by quantifying the spatial decoration between the emission tracers and the abundance of emission peaks in each tracer. The method is, however, reliant on a reference timescale to convert these relative durations into an absolute evolutionary timeline. Stellar emission is closely tied to stellar mass, which also dictates stellar evolution. This means that SFR tracer emission effectively has a built-in clock governed by the physics of stellar evolution.

Although it is widely accepted that SFR tracer emission has a lifetime, a concrete definition for the lifetime is not. The work we present in [Chapter 3](#)^[p. 47] uses the KL14 principle itself with synthetic SFR tracer maps and reference maps (a map with a known age) to define these SFR tracer lifetimes. The benefit of a method that uses the KL14 principle is that the lifetimes will be self-consistent within the framework of the KL14 principle. That is, we will provide reference timescales that are appropriate for observational applications of the KL14 principle.

The synthetic SFR tracer maps we produce are based on a near- L_{\star} , isolated, flocculent spiral galaxy simulated using P-GADGET-3. Using the stochastic SPS code SLUG2 and the properties of the star particles within the simulation, we produce synthetic emission maps of $H\alpha$ emission (with and without continuum subtraction) as well as 12 UV filters (from GALEX, Swift, and HST). With

6 Conclusion

SLUG2, we have the freedom to choose the stellar evolutionary tracks which are used in defining the SPS assigned to each star particle and therefore its emission. This allows us to explore the effects of metallicity on the emission lifetimes. We find that with decreasing metallicity the emission lifetime increases. For convenience we also provide analytic functions that can be used to calculate the emission lifetime based on the metallicity and properties of the emission filter. Another important environmental factor we test is the effects of low SFR surface densities. This results in a stochastically sampled IMF. We find that only the emission lifetime of $H\alpha$ is altered, as this emission is only produced by the most massive stars. The reduction in the measured emission lifetime with decreasing SFR surface densities can be related to the probability of finding a sufficiently massive star given the mass of the star forming region.

The emission lifetimes we provide here have already been used in the first observational applications of the KL14 principle (Chevance and author 2019; Chevance, Kruijssen et al. 2019; Hygate et al. 2019; Kruijssen, Schrubba, Chevance et al. 2019; Ward et al. 2019), demonstrating their usefulness and effectiveness in constraining the physics driving molecular cloud life cycle. However, one may have concerns regarding the accuracy of the lifetimes we provide. Further work on emission lifetimes could be done to test the possible impact of using different stellar evolutionary tracks. Repeating the experiments on different types of galaxy simulations and of different galaxy types, could also aid in the improved understanding of the lifetimes of SFR tracer emission. The greatest benefit would, however, come from further companions to observational data. In particular, observations where other means of constraining a reference timescale exist. In [Chapter 3](#)^[p.47] we present the first and (currently) only possible observational test of these emission lifetimes and are encouraged by the results.

In [Chapter 4](#)^[p.99] we addressed the possible effects of dust extinction on the emission lifetimes we found in [Chapter 3](#)^[p.47]. Although extinction does not alter the underlying emission lifetime, it can act to alter the emission lifetime that we measure. We use a similar approach to the method presented in [Chapter 3](#)^[p.47], but perform the analysis on a Milky-Way-like disc galaxy simulation. From previous work we found that, although on large scales the simulation agrees with observational constraints, the implemented feedback mechanisms are insufficient to disperse molecular clouds surrounding sites of star formation. This means we can use the simulation as a limiting case; this is the maximum extent in which emission lifetimes could be altered by extinction.

At high levels of extinction, where the mean galactic surface density is $\Sigma_{\text{gas}} \gtrsim 20 M_{\odot} \text{pc}^{-2}$, we find (as expected) that with increasing levels of

extinction the measured emission lifetimes decrease. However, low levels of extinction can in fact act to increase the measured emission lifetime. Our results show that with low levels of extinction, emission in the simulated galaxy is preferentially removed from the brightest star-forming regions. This change in the emission distribution, resulting in HEISENBERG’s statistical analysis to interpret the extinguished emission map as having an increased duration with respect to the unextinguished emission map. We also find that non-zero sensitivity limits in the observational data can act to compound the effects of extinction, reducing the measured emission lifetime further.

The work in [Chapter 4](#)^[p.99] constrains the extreme limits on the effects of extinction on the emission lifetimes. This is important knowledge for understanding the extent to which measurements made by KL14 principle could be affected. However, performing a similar analysis based on simulations with improved feedback physics would help to better constrain these limits. Ideally observational tests based on galaxies that could be reliably corrected for extinction would help to validate our findings.

In [Chapter 5](#)^[p.125] we create and study synthetic CO observations. The ISM plays an important role when it comes to star formation and feedback. Stars born from the dense molecular gas within the ISM later return energy, mass and momentum back to the ISM through stellar feedback. The study of the molecular gas within the ISM, through CO observations, is key to understanding the physics of these processes.

Instead of studying a single galaxy, we create synthetic observations for a $2.5' \times 2.5'$ field of view and spanning a redshift range 1–4. This is to mimic the real observations performed by the ASPECS project, an ALMA survey of gas (though CO emission) in the Hubble UDF. To create our synthetic counterpart, we use the cosmological simulation IllustrisTNG50. IllustrisTNG50 is a gravity+magnetohydrodynamical cosmological simulation based on the moving mesh code AREPO and a new class of simulations designed to overcome the traditional compromise between simulation volume and resolution. The simulation volume is 51.7^3 cMpc^3 and has a median radius of $\sim 100\text{--}140 \text{ pc}$ in star-forming gas cells.

The simulation snapshots cannot be used directly to form our observations: we must first produce a synthetic light cone. For this purpose we developed the code LIGHTCONE. This code stitches together small volumes from the simulation snapshot that together form a complete volume that has a (redshift dependent) height and width corresponding to the field of view and a depth that spans the required redshift range. The properties of the gas that fall within these smaller volumes are used in combination with a lookup table to interpolate the expected CO emission.

We create the lookup table using `DESPOTIC`, a code developed to calculate the state of interstellar clouds and their observable emission. The table is 8D, one dimension for each of the key properties: Hydrogen volume and surface density, SFR surface density, abundance of Helium, Carbon, and Oxygen, metallicity, and redshift. This table is optimised for the IllustrisTNG50, by selecting coordinates within the parameter space that reflect the distribution of the properties and by ignoring unphysical regions of the parameter space.

Based on a single light cone, we present results of the $2.5' \times 2.5'$ field of view showing the distribution of CO emission within a galaxy population; the high spatial resolution of IllustrisTNG50 also allows us to see the distribution of CO emission within the galaxies themselves. However, the angular resolution required to see this level of detail cannot be matched by observations and so the analysis of our synthetic CO observations is carried out on $1'' \times 1''$ resolution images, which matches the angular resolution of ASPECS. Initial results show that we over predict the amount of CO emission; although, the inclusion of the effects of extinction and a sensitivity limit would decrease this discrepancy. The statistical relations we present, showing the change in the total CO emission with redshift and CO luminosity functions could be improved by averaging the relations from hundreds of different light cones. Through this averaging we would remove features specific to a given light cone and would be able to characterise the limits within which we would expect true observations to fall as a result of field-to-field variation. Through these improved relations, we could conduct meaningful comparisons between synthetic and real observations, informing us about the adequacy of the models used within IllustrisTNG50 and our assumptions in modelling CO emission.

Through the study of synthetic SFR tracer observations we have provided insight into the nature of the emission lifetimes of these tracers. These lifetimes can help constrain the physics driving the molecular cloud life cycle. Through the study of synthetic CO observations, we have quantified the distribution of CO emission over cosmic time. These will inform us about ISM properties and provide an understanding of the key stellar feedback mechanisms.

Acknowledgements

I would first like to acknowledge the International Max Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg (IMPRS-HD) who aided in finding my position here at Heidelberg, provided information on settling in Germany, and supplemented my PhD with additional courses, summer schools, and retreats. I also acknowledge the Astronomisches Rechen-Institut (ARI) for hosting me and thank all those who work to help the smooth running of the institute.

I would like to thank my advisor, Diederik Kruijssen, who has guided and supported me over the years. You have helped to improve the quality of my research and allowed me to develop as a scientist. Your efforts have also given me opportunities and resources, that would have been greatly missed. To the group (in order of appearance) Alexander Hygate, Marta Reina-Campos, Sarah Jeffreson, Mélanie Chevance, Jacob Ward, Benjamin Keller, Jindra Gensior, Sebastian Trujillo-Gomez, Maya Petkova, and Jaeyeon Kim, I express my gratitude for the lively and supportive working environment we have created together as the group has grown. I greatly appreciate our discussions as well as your advice, insights, and help. This group has always been about more than just work and so for the dinners, game nights, Weinwanderungen, Wine and Cheese evenings, and more I am truly grateful.

I would like to thank Mark Krumholz and his group for hosting me during my time spent at The Australian National University. In particular, I appreciate Yusuk Fujimoto, Arpita Roy, and Lucia Armillotta for making my stay at Mount Stromlo an enjoyable one. I also thank Mark for informative discussions and input into the research I have presented here.

To Kseniia Sysoliatina, once office mate and now good friend, I thank you for distracting me with concerts, plays, and weekend walks. You helped me during these stressful times.

A special thanks to past and present colleagues, Peter Zeidler, Carolin Wittmann, Faiezeh Shabani, Zdeněk Prudil, and Michael Hanke.

Ich möchte mich bei Carsten Eggersgluß und Michael Zeitler ganz herzlich dafür bedanken, dass ihr mir einen perfekten Anfang meines Lebens in Deutschland gegeben habt. Vom Anfang habt ihr mir geholfen und mich unterstützt. Die Ausflüge, das Essen, der Wein und die Gesellschaft waren immer fantastisch. An zwei der besten Leute, die ich je kennengelernt habe, vielen Dank!

Lastly, I thank my family for the years of support before, during, and undoubtedly after this PhD: I could not have done this without you.

List of publications

Author

- D. T. Haydon, J. M. D. Kruijssen, A. P. S. Hygate, A. Schrubba, M. R. Krumholz, M. Chevance and S. N. Longmore (Oct. 2018). ‘An uncertainty principle for star formation – III. The characteristic emission time-scales of star formation rate tracers’. *MNRAS Submitted*. arXiv: [1810.10897 \[astro-ph.GA\]](#).
- D. T. Haydon, Y. Fujimoto, M. Chevance, J. M. D. Kruijssen, M. R. Krumholz and S. N. Longmore (Aug. 2019). ‘An uncertainty principle for star formation – V. The influence of dust extinction on star formation rate tracer lifetimes’. *MNRAS Submitted*.

Co-author

- M. Chevance, J. M. D. Kruijssen, A. P. S. Hygate, A. Schrubba, S. N. Longmore, B. Groves, J. D. Henshaw, C. N. Herrera, A. Hughes, S. M. R. Jeffreson, P. Lang, A. K. Leroy, S. E. Meidt, J. Pety, A. Razza, E. Rosolowsky, E. Schinnerer, F. Bigiel, G. A. Blanc, E. Emsellem, C. M. Faesi, S. C. O. Glover, D. T. Haydon, I.-T. Ho, K. Kreckel, J. C. Lee, D. Liu, M. Querejeta, T. Saito, J. Sun, A. Usero and D. Utomo (Dec. 2019). ‘The lifecycle of molecular clouds in nearby star-forming disc galaxies’. *MNRAS Accepted*. arXiv: [1911.03479 \[astro-ph.GA\]](#).
- Y. Fujimoto, M. Chevance, D. T. Haydon, M. R. Krumholz and J. M. D. Kruijssen (Aug. 2019). ‘A fundamental test for stellar feedback recipes in galaxy simulations’. *MNRAS* 487.2, pp. 1717–1728. arXiv: [1905.09839 \[astro-ph.GA\]](#).
- A. P. S. Hygate, J. M. D. Kruijssen, M. Chevance, A. Schrubba, D. T. Haydon and S. N. Longmore (Sept. 2019). ‘An uncertainty principle for star formation - IV. On the nature and filtering of diffuse emission’. *MNRAS* 488.2, pp. 2800–2824. arXiv: [1810.11405 \[astro-ph.GA\]](#).
- A. P. S. Hygate, J. M. D. Kruijssen, M. Chevance, F. Walter, J. J. Kim, J. L. Ward, A. Schrubba, D. T. Haydon and S. N. Longmore (Oct. 2019). ‘The molecular cloud lifecycle in M33’. *MNRAS Submitted*.

- J. M. D. Kruijssen, A. Schruba, A. P. S. Hygate, C.-Y. Hu, D. T. Haydon and S. N. Longmore (Sept. 2018). ‘An uncertainty principle for star formation - II. A new method for characterizing the cloud-scale physics of star formation and feedback across cosmic history’. *MNRAS* 479.2, pp. 1866–1952. arXiv: [1805.00012 \[astro-ph.GA\]](#).
- J. M. D. Kruijssen, A. Schruba, M. Chevance, S. N. Longmore, A. P. S. Hygate, D. T. Haydon, A. F. McLeod, J. J. Dalcanton, L. J. Tacconi and E. F. van Dishoeck (May 2019). ‘Fast and inefficient star formation due to short-lived molecular clouds and rapid feedback’. *Nature* 569.7757, pp. 519–522. arXiv: [1905.08801 \[astro-ph.GA\]](#).

Bibliography

- Alatalo, K. et al. (Oct. 2014). ‘Catching Quenching Galaxies: The Nature of the WISE Infrared Transition Zone’. *ApJ* 794.1, L13, p. L13. arXiv: [1409.2489 \[astro-ph.GA\]](#) (cit. on pp. 16, 17).
- Aver, E. et al. (July 2015). ‘The effects of He I λ 10830 on helium abundance determinations’. *J. Cosmology Astropart. Phys.* 2015.7, 011, p. 011. arXiv: [1503.08146 \[astro-ph.CO\]](#) (cit. on p. 5).
- Baldry, I. K. et al. (Dec. 2006). ‘Galaxy bimodality versus stellar mass and environment’. *MNRAS* 373.2, pp. 469–483. arXiv: [astro-ph/0607648 \[astro-ph\]](#) (cit. on p. 16).
- Bania, T. M. et al. (Jan. 2002). ‘The cosmological density of baryons from observations of $^3\text{He}^+$ in the Milky Way’. *Nature* 415.6867, pp. 54–57. (Cit. on p. 5).
- Barlow, R. (2003). ‘Asymmetric Errors’. *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*. Ed. by L. Lyons et al., p. 250. arXiv: [physics/0401042 \[physics.data-an\]](#) (cit. on p. 61).
- Bastian, N. et al. (Sept. 2010). ‘A Universal Stellar Initial Mass Function? A Critical Look at Variations’. *ARA&A* 48, pp. 339–389. arXiv: [1001.2965 \[astro-ph.GA\]](#) (cit. on p. 28).
- Baugh, C. M. et al. (Sept. 1996). ‘Faint galaxy counts as a function of morphological type in a hierarchical merger model’. *MNRAS* 282.1, pp. L27–L32. arXiv: [astro-ph/9607056 \[astro-ph\]](#) (cit. on p. 16).
- Bedregal, A. G. et al. (Dec. 2006). ‘The Tully-Fisher relation for S0 galaxies’. *MNRAS* 373.3, pp. 1125–1140. arXiv: [astro-ph/0609076 \[astro-ph\]](#) (cit. on p. 13).
- Bekki, K. et al. (Oct. 2003). ‘Starbursts from the Strong Compression of Galactic Molecular Clouds due to the High Pressure of the Intracluster Medium’. *ApJ* 596.1, pp. L13–L16. arXiv: [astro-ph/0308264 \[astro-ph\]](#) (cit. on p. 19).
- Benson, A. J. et al. (Dec. 2003). ‘What Shapes the Luminosity Function of Galaxies?’ *ApJ* 599.1, pp. 38–49. arXiv: [astro-ph/0302450 \[astro-ph\]](#) (cit. on p. 14).
- Berman, L. (June 1936). ‘The effect of space reddening on the Balmer decrement in planetary nebulae’. *MNRAS* 96, p. 890. ADS: [1936MNRAS...96..890B](#) (cit. on p. 99).

- Bigiel, F. et al. (Dec. 2008). ‘The Star Formation Law in Nearby Galaxies on Sub-Kpc Scales’. *AJ* 136.6, pp. 2846–2871. arXiv: [0810.2541 \[astro-ph\]](#) (cit. on pp. 31, 33).
- Blaizot, J. et al. (June 2005). ‘MoMaF: the Mock Map Facility’. *MNRAS* 360.1, pp. 159–175. arXiv: [astro-ph/0309305 \[astro-ph\]](#) (cit. on pp. 135, 136).
- Blanton, M. R. et al. (Sept. 2009). ‘Physical Properties and Environments of Nearby Galaxies’. *ARA&A* 47.1, pp. 159–210. arXiv: [0908.3017 \[astro-ph.GA\]](#) (cit. on p. 13).
- Bodenheimer, P. (2011). *Principles of Star Formation*. Astronomy and Astrophysics Library. Springer Berlin Heidelberg. ISBN: 9783642150630. DOI: [10.1007/978-3-642-15063-0](#) (cit. on pp. 21, 22).
- Bolatto, A. D. et al. (Aug. 2013). ‘The CO-to-H₂ Conversion Factor’. *ARA&A* 51.1, pp. 207–268. arXiv: [1301.3498 \[astro-ph.GA\]](#) (cit. on p. 22).
- Bonatto, C. et al. (Aug. 2011). ‘Constraining the star formation rate in the solar neighbourhood with star clusters’. *MNRAS* 415.3, pp. 2827–2831. arXiv: [1104.2182 \[astro-ph.GA\]](#) (cit. on p. 141).
- Boselli, A. et al. (Apr. 2006). ‘Environmental Effects on Late-Type Galaxies in Nearby Clusters’. *PASP* 118.842, pp. 517–559. arXiv: [astro-ph/0601108 \[astro-ph\]](#) (cit. on p. 19).
- Bournaud, F. et al. (Sept. 2002). ‘Gas accretion on spiral galaxies: Bar formation and renewal’. *A&A* 392, pp. 83–102. arXiv: [astro-ph/0206273 \[astro-ph\]](#) (cit. on p. 14).
- Bower, R. G. et al. (Aug. 2006). ‘Breaking the hierarchy of galaxy formation’. *MNRAS* 370.2, pp. 645–655. arXiv: [astro-ph/0511338 \[astro-ph\]](#) (cit. on p. 14).
- Bryan, G. L. et al. (Apr. 2014). ‘ENZO: An Adaptive Mesh Refinement Code for Astrophysics’. *ApJS* 211.2, 19, p. 19. arXiv: [1307.2265 \[astro-ph.IM\]](#) (cit. on p. 101).
- Calzetti, D. et al. (July 1994). ‘Dust Extinction of the Stellar Continua in Starburst Galaxies: The Ultraviolet and Optical Extinction Law’. *ApJ* 429, p. 582. ADS: [1994ApJ...429..582C](#) (cit. on p. 99).
- Cappellari, M. et al. (July 2013). ‘The ATLAS^{3D} project - XV. Benchmark for early-type galaxies scaling relations from 260 dynamical models: mass-to-light ratio, dark matter, Fundamental Plane and Mass Plane’. *MNRAS* 432.3, pp. 1709–1741. arXiv: [1208.3522 \[astro-ph.CO\]](#) (cit. on p. 13).
- Catinella, B. et al. (May 2018). ‘xGASS: total cold gas scaling relations and molecular-to-atomic gas ratios of galaxies in the local Universe’. *MNRAS* 476.1, pp. 875–895. arXiv: [1802.02373 \[astro-ph.GA\]](#) (cit. on p. xix).

- Chabrier, G. (July 2003). ‘Galactic Stellar and Substellar Initial Mass Function’. *PASP* 115.809, pp. 763–795. arXiv: [astro-ph/0304382 \[astro-ph\]](#) (cit. on p. 128).
- Chabrier, G. (Jan. 2005). ‘The Initial Mass Function: From Salpeter 1955 to 2005’. *The Initial Mass Function 50 Years Later*. Ed. by E. Corbelli et al. Vol. 327. Astrophysics and Space Science Library. Springer, Dordrecht, p. 41. arXiv: [astro-ph/0409465 \[astro-ph\]](#) (cit. on pp. 28, 58, 80, 102).
- Charbonnel, C. et al. (Oct. 1993). ‘Grids of stellar models. III. From 0.8 to 120 Msolar at Z=0.004’. *A&AS* 101, p. 415. ADS: [1993A&AS..101..415C](#) (cit. on p. 70).
- Chevance, M. and author (2019). ‘The stellar feedback mechanisms driving molecular cloud dispersal in nearby galaxies’. *To be submitted* (cit. on pp. 40, 94, 170).
- Chevance, M., J. M. D. Kruijssen et al. (Dec. 2019). ‘The lifecycle of molecular clouds in nearby star-forming disc galaxies’. *MNRAS Accepted*. arXiv: [1911.03479 \[astro-ph.GA\]](#) (cit. on pp. 40, 44, 94, 101, 104, 121–123, 170).
- Christensen, C. J. et al. (Apr. 1972). ‘Free-Neutron Beta-Decay Half-Life’. *Phys. Rev. D* 5 (7), pp. 1628–1640. (Cit. on p. 3).
- Coc, A. et al. (Jan. 2012). ‘Standard Big Bang Nucleosynthesis up to CNO with an Improved Extended Nuclear Network’. *ApJ* 744.2, 158, p. 158. arXiv: [1107.1117 \[astro-ph.CO\]](#) (cit. on p. 4).
- Collette, A. (2013). *Python and HDF5*. O’Reilly. ISBN: 9781449367831 (cit. on p. 129).
- Cooke, R. J. et al. (Jan. 2014). ‘Precision Measures of the Primordial Abundance of Deuterium’. *ApJ* 781.1, 31, p. 31. arXiv: [1308.3240 \[astro-ph.CO\]](#) (cit. on p. 5).
- Cyburt, R. H. et al. (Nov. 2008). ‘An update on the big bang nucleosynthesis prediction for ${}^7\text{Li}$: the problem worsens’. *J. Cosmology Astropart. Phys.* 2008.11, 012, p. 012. arXiv: [0808.2818 \[astro-ph\]](#) (cit. on p. 5).
- Cyburt, R. H. et al. (Jan. 2016). ‘Big bang nucleosynthesis: Present status’. *Reviews of Modern Physics* 88.1, 015004, p. 015004. arXiv: [1505.01076 \[astro-ph.CO\]](#) (cit. on p. 5).
- Daddi, E. et al. (Apr. 2010). ‘Very High Gas Fractions and Extended Gas Reservoirs in $z = 1.5$ Disk Galaxies’. *ApJ* 713.1, pp. 686–707. arXiv: [0911.2776 \[astro-ph.CO\]](#) (cit. on pp. 161, 163, 164).
- Dalcín, L. et al. (Sept. 2005). ‘MPI for Python’. *Journal of Parallel and Distributed Computing* 65.9, pp. 1108–1115. (Cit. on p. 129).

- Dalcín, L. et al. (May 2008). ‘MPI for Python: Performance improvements and MPI-2 extensions’. *Journal of Parallel and Distributed Computing* 68.5, pp. 655–662. (Cit. on p. 129).
- Dalcín, L. et al. (Sept. 2011). ‘Parallel distributed computing using Python’. *Advances in Water Resources* 34.9, pp. 1124–1139. (Cit. on p. 129).
- Darvish, B. et al. (Feb. 2018). ‘Quenching or Bursting: The Role of Stellar Mass, Environment, and Specific Star Formation Rate to $z \sim 1$ ’. *ApJ* 853.2, 155, p. 155. arXiv: [1801.02618 \[astro-ph.GA\]](#) (cit. on p. 16).
- Dashyan, G. et al. (Aug. 2019). ‘AGN-driven quenching of satellite galaxies’. *MNRAS* 487.4, pp. 5889–5901. arXiv: [1906.07431 \[astro-ph.GA\]](#) (cit. on p. 14).
- Davis, M. et al. (May 1985). ‘The evolution of large-scale structure in a universe dominated by cold dark matter’. *ApJ* 292, pp. 371–394. (Cit. on p. 128).
- Decarli, R. et al. (Dec. 2016). ‘ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: CO Luminosity Functions and the Evolution of the Cosmic Density of Molecular Gas’. *ApJ* 833.1, 69, p. 69. arXiv: [1607.06770 \[astro-ph.GA\]](#) (cit. on pp. 125, 130).
- Decarli, R. et al. (Sept. 2019). ‘The ALMA Spectroscopic Survey in the HUDF: CO Luminosity Functions and the Molecular Gas Content of Galaxies through Cosmic History’. *ApJ* 882.2, 138, p. 138. arXiv: [1903.09164 \[astro-ph.GA\]](#) (cit. on p. 125).
- Dehnen, W. et al. (Sept. 2012). ‘Improving convergence in smoothed particle hydrodynamics simulations without pairing instability’. *MNRAS* 425.2, pp. 1068–1082. arXiv: [1204.2471 \[astro-ph.IM\]](#) (cit. on p. 53).
- Desmond, H. et al. (Feb. 2017). ‘The Faber-Jackson relation and Fundamental Plane from halo abundance matching’. *MNRAS* 465.1, pp. 820–833. arXiv: [1604.04670 \[astro-ph.GA\]](#) (cit. on p. 13).
- Dokkum, P. van et al. (Mar. 2018). ‘A galaxy lacking dark matter’. *Nature* 555.7698, pp. 629–632. arXiv: [1803.10237 \[astro-ph.GA\]](#) (cit. on p. 11).
- Draine, B. T. et al. (Sept. 1996). ‘Structure of Stationary Photodissociation Fronts’. *ApJ* 468, p. 269. arXiv: [astro-ph/9603032 \[astro-ph\]](#) (cit. on pp. 105, 144).
- Draine, B. T. (2011). *Physics of the interstellar and intergalactic medium*. eng. Princeton series in astrophysics. Formerly CIP Uk. - Includes bibliographical references and index. Princeton [u.a.]: Princeton Univ. Press, XVIII, 540, [8] S. ISBN: 9780691122137 (cit. on pp. 22, 139).
- Dressler, A. (Mar. 1980). ‘Galaxy morphology in rich clusters: implications for the formation and evolution of galaxies.’ *ApJ* 236, pp. 351–365. (Cit. on p. 19).

- Driver, S. et al. (Jan. 2003). ‘The Global, Local and Cluster Galaxy Luminosity Function’. *Ap&SS* 285.1, pp. 175–184. arXiv: [astro-ph/0212520](#) [[astro-ph](#)] (cit. on p. 14).
- Dullemond, C. P. et al. (Feb. 2012). *RADMC-3D: A multi-purpose radiative transfer tool*. ascl: [1202.015](#) (cit. on p. 126).
- Eggen, O. J. et al. (Nov. 1962). ‘Evidence from the motions of old stars that the Galaxy collapsed.’ *ApJ* 136, p. 748. (Cit. on p. 14).
- Eisenstein, D. J. et al. (Nov. 2005). ‘Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies’. *ApJ* 633.2, pp. 560–574. arXiv: [astro-ph/0501171](#) [[astro-ph](#)] (cit. on p. 8).
- Elmegreen, B. G. (Feb. 2000). ‘Star Formation in a Crossing Time’. *ApJ* 530.1, pp. 277–281. arXiv: [astro-ph/9911172](#) [[astro-ph](#)] (cit. on p. 34).
- ESA / Planck Collaboration (Mar. 2013). *Cosmic Microwave Background seen by Planck*. URL: <https://sci.esa.int/s/WLGmGdw> (visited on 09/11/2019) (cit. on p. 7).
- ESO (2007). *Hertzsprung-Russell Diagram*. URL: <https://www.eso.org/public/images/eso0728c/> (visited on 10/11/2019) (cit. on p. 29).
- Faber, S. M. et al. (Mar. 1976). ‘Velocity dispersions and mass-to-light ratios for elliptical galaxies.’ *ApJ* 204, pp. 668–683. (Cit. on p. 13).
- Fabian, A. C. (Sept. 2012). ‘Observational Evidence of Active Galactic Nuclei Feedback’. *ARA&A* 50, pp. 455–489. arXiv: [1204.4114](#) [[astro-ph.CO](#)] (cit. on p. 19).
- Fan, X. et al. (Sept. 2006). ‘Observational Constraints on Cosmic Reionization’. *ARA&A* 44.1, pp. 415–462. arXiv: [astro-ph/0602375](#) [[astro-ph](#)] (cit. on p. 9).
- Fields, B. D. (Nov. 2011). ‘The Primordial Lithium Problem’. *Annual Review of Nuclear and Particle Science* 61.1, pp. 47–68. arXiv: [1203.3551](#) [[astro-ph.CO](#)] (cit. on p. 5).
- Fujimoto, Y. et al. (Nov. 2018). ‘Short-lived radioisotopes in meteorites from Galactic-scale correlated star formation’. *MNRAS* 480.3, pp. 4025–4039. arXiv: [1802.08695](#) [[astro-ph.GA](#)] (cit. on pp. 101, 102, 122).
- Fujimoto, Y. et al. (Aug. 2019). ‘A fundamental test for stellar feedback recipes in galaxy simulations’. *MNRAS* 487.2, pp. 1717–1728. arXiv: [1905.09839](#) [[astro-ph.GA](#)] (cit. on pp. 53, 102, 103, 110, 122).
- Galassi, M. et al. (Jan. 2009). *GNU Scientific Library Reference Manual*. Ed. by B. Gough. 3rd. Network Theory Ltd. ISBN: 9780954612078 (cit. on p. 129).
- Genel, S. et al. (Nov. 2014). ‘Introducing the Illustris project: the evolution of galaxy populations across cosmic time’. *MNRAS* 445.1, pp. 175–200. arXiv: [1405.3749](#) [[astro-ph.CO](#)] (cit. on p. 9).

- Gerola, H. et al. (July 1978). ‘Stochastic star formation and spiral structure of galaxies.’ *ApJ* 223, pp. 129–139. (Cit. on p. 13).
- Girardi, L. et al. (Feb. 2000). ‘Evolutionary tracks and isochrones for low- and intermediate-mass stars: From 0.15 to 7 M_{sun} , and from $Z=0.0004$ to 0.03’. *A&AS* 141, pp. 371–383. arXiv: [astro-ph/9910164](#) [[astro-ph](#)] (cit. on p. 102).
- Glover, S. C. O. et al. (Apr. 2007). ‘Simulating the Formation of Molecular Clouds. I. Slow Formation by Gravitational Collapse from Static Initial Conditions’. *ApJS* 169.2, pp. 239–268. arXiv: [astro-ph/0605120](#) [[astro-ph](#)] (cit. on p. 144).
- Glover, S. C. O. et al. (Mar. 2012). ‘Approximations for modelling CO chemistry in giant molecular clouds: a comparison of approaches’. *MNRAS* 421.1, pp. 116–131. arXiv: [1102.0670](#) [[astro-ph.GA](#)] (cit. on p. 144).
- González-López, J. et al. (Mar. 2019). ‘The ALMA Spectroscopic Survey in the HUDF: CO emission lines and 3 mm continuum sources’. *arXiv e-prints*, arXiv:1903.09161, arXiv:1903.09161. arXiv: [1903.09161](#) [[astro-ph.GA](#)] (cit. on p. 125).
- Grand, R. J. J. et al. (May 2017). ‘The Auriga Project: the properties and formation mechanisms of disc galaxies across cosmic time’. *MNRAS* 467.1, pp. 179–207. arXiv: [1610.01159](#) [[astro-ph.GA](#)] (cit. on p. 126).
- Greene, T. P. et al. (Oct. 1994). ‘Further Mid-Infrared Study of the rho Ophiuchi Cloud Young Stellar Population: Luminosities and Masses of Pre-Main-Sequence Stars’. *ApJ* 434, p. 614. (Cit. on p. 21).
- Griffiths, L. M. et al. (Sept. 1999). ‘Cosmic microwave background constraints on the epoch of reionization’. *MNRAS* 308.3, pp. 854–862. arXiv: [astro-ph/9812125](#) [[astro-ph](#)] (cit. on p. 2).
- Haas, M. R. et al. (Nov. 2013). ‘Physical properties of simulated galaxy populations at $z = 2$ - I. Effect of metal-line cooling and feedback from star formation and AGN’. *MNRAS* 435.4, pp. 2931–2954. arXiv: [1211.1021](#) [[astro-ph.CO](#)] (cit. on pp. 35, 36).
- Hao, C.-N. et al. (Nov. 2011). ‘Dust-corrected Star Formation Rates of Galaxies. II. Combinations of Ultraviolet and Infrared Tracers’. *ApJ* 741.2, 124, p. 124. arXiv: [1108.2837](#) [[astro-ph.CO](#)] (cit. on pp. 31, 47).
- Hartmann, L. et al. (Dec. 2001). ‘Rapid Formation of Molecular Clouds and Stars in the Solar Neighborhood’. *ApJ* 562.2, pp. 852–868. arXiv: [astro-ph/0108023](#) [[astro-ph](#)] (cit. on p. 34).
- Hayden, M. R. et al. (Nov. 2017). ‘The AMBRE project: The thick thin disk and thin thick disk of the Milky Way’. *A&A* 608, L1, p. L1. arXiv: [1712.02358](#) [[astro-ph.GA](#)] (cit. on p. 11).

- Haydon, D. T. et al. (Oct. 2018). ‘An uncertainty principle for star formation – III. The characteristic emission time-scales of star formation rate tracers’. *MNRAS Submitted*. arXiv: [1810.10897 \[astro-ph.GA\]](#) (cit. on pp. 39, 47).
- Haydon, D. T. et al. (Aug. 2019). ‘An uncertainty principle for star formation – V. The influence of dust extinction on star formation rate tracer lifetimes’. *MNRAS Submitted* (cit. on pp. 39, 99).
- Hind, J. R. (Jan. 1864). ‘Note on the Variable Nebula in Taurus’. *MNRAS* 24, p. 65 (cit. on p. 22).
- Hopkins, P. F. et al. (Apr. 2012). ‘The structure of the interstellar medium of star-forming galaxies’. *MNRAS* 421.4, pp. 3488–3521. arXiv: [1110.4636 \[astro-ph.CO\]](#) (cit. on p. 37).
- Hopkins, P. F. et al. (July 2013). ‘The meaning and consequences of star formation criteria in galaxy models with resolved stellar feedback’. *MNRAS* 432.4, pp. 2647–2653. arXiv: [1303.0285 \[astro-ph.CO\]](#) (cit. on p. 35).
- Hopkins, P. F. et al. (Nov. 2014). ‘Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation’. *MNRAS* 445.1, pp. 581–603. arXiv: [1311.2073 \[astro-ph.CO\]](#) (cit. on pp. xix, 19).
- Hopkins, P. F. et al. (Oct. 2018). ‘FIRE-2 simulations: physics versus numerics in galaxy formation’. *MNRAS* 480.1, pp. 800–863. arXiv: [1702.06148 \[astro-ph.GA\]](#) (cit. on p. 51).
- Houghton, R. C. W. (Aug. 2015). ‘Revisiting the original morphology-density relation’. *MNRAS* 451.4, pp. 3427–3436. arXiv: [1505.04788 \[astro-ph.GA\]](#) (cit. on p. 19).
- Hu, C.-Y. et al. (Sept. 2014). ‘SPHGal: smoothed particle hydrodynamics with improved accuracy for galaxy simulations’. *MNRAS* 443.2, pp. 1173–1191. arXiv: [1402.1788 \[astro-ph.CO\]](#) (cit. on p. 51).
- Hu, W. et al. (Jan. 2002). ‘Cosmic Microwave Background Anisotropies’. *ARA&A* 40, pp. 171–216. arXiv: [astro-ph/0110414 \[astro-ph\]](#) (cit. on p. 3).
- Hubble Heritage Team / ESA / NASA (Jan. 2016). *NGC1300*. URL: <https://apod.nasa.gov/apod/ap160109.html> (visited on 09/11/2019) (cit. on p. 15).
- Hughes, I. G. et al. (2010). *Measurements and their uncertainties : a practical guide to modern error analysis*. Oxford : New York, NY: Oxford University Press (cit. on p. 97).
- Hygate, A. P. S. et al. (Oct. 2019). ‘The molecular cloud lifecycle in M33’. *MNRAS Submitted* (cit. on pp. 40, 94, 121, 170).

- Illustris Collaboration (2018). *Illustris Media*. URL: <https://www.illustris-project.org/media> (visited on 09/11/2019) (cit. on p. 10).
- James, P. A. et al. (Jan. 2005). ‘The H α Galaxy Survey. II. Extinction and [NII] corrections to H α fluxes’. *A&A* 429, pp. 851–867. arXiv: [astro-ph/0410386](https://arxiv.org/abs/astro-ph/0410386) [[astro-ph](#)] (cit. on pp. 59, 99).
- Jeans, J. H. (Jan. 1902). ‘The Stability of a Spherical Nebula’. *Philosophical Transactions of the Royal Society of London Series A* 199, pp. 1–53. (Cit. on p. 21).
- Jeans, J. H. (1928). ‘Astronomy and Cosmogony’. *Journal of Philosophical Studies* 3.12, pp. 533–535 (cit. on p. 21).
- Jefferson, S. M. R. et al. (May 2018). ‘A general theory for the lifetimes of giant molecular clouds under the influence of galactic dynamics’. *MNRAS* 476.3, pp. 3688–3715. arXiv: [1803.01850](https://arxiv.org/abs/1803.01850) [[astro-ph.GA](#)] (cit. on p. 34).
- Katz, N. (June 1992). ‘Dissipational Galaxy Formation. II. Effects of Star Formation’. *ApJ* 391, p. 502. ADS: [1992ApJ...391..502K](https://ui.adsabs.org/1992ApJ...391..502K) (cit. on p. 51).
- Kauffmann, G. et al. (Sept. 2004). ‘The environmental dependence of the relations between stellar mass, structure, star formation and nuclear activity in galaxies’. *MNRAS* 353.3, pp. 713–731. arXiv: [astro-ph/0402030](https://arxiv.org/abs/astro-ph/0402030) [[astro-ph](#)] (cit. on p. 19).
- Kawamura, A. et al. (Sept. 2009). ‘The Second Survey of the Molecular Clouds in the Large Magellanic Cloud by NANTEN. II. Star Formation’. *ApJS* 184.1, pp. 1–17. arXiv: [0908.1168](https://arxiv.org/abs/0908.1168) [[astro-ph.CO](#)] (cit. on p. 34).
- Kennicutt Robert C., J. (May 1998). ‘The Global Schmidt Law in Star-forming Galaxies’. *ApJ* 498.2, pp. 541–552. arXiv: [astro-ph/9712213](https://arxiv.org/abs/astro-ph/9712213) [[astro-ph](#)] (cit. on pp. 31, 32).
- Kennicutt, R. C. et al. (Sept. 2012). ‘Star Formation in the Milky Way and Nearby Galaxies’. *ARA&A* 50, pp. 531–608. arXiv: [1204.3552](https://arxiv.org/abs/1204.3552) [[astro-ph.GA](#)] (cit. on pp. 22, 31, 47, 48, 64, 77).
- Kernighan, B. W. et al. (Mar. 1988). *The C Programming Language*. Prentice-Hall software series. Prentice Hall. ISBN: 9780131103627 (cit. on p. 129).
- Klessen, R. S. et al. (2016). ‘Physical Processes in the Interstellar Medium’. *Star formation in galaxy evolution: connecting numerical models to reality*. Ed. by Y. Revaz et al. Springer. ISBN: 9783662478899 (cit. on pp. 20, 23).
- Koda, J. et al. (Aug. 2009). ‘Dynamically Driven Evolution of the Interstellar Medium in M51’. *ApJ* 700.2, pp. L132–L136. arXiv: [0907.1656](https://arxiv.org/abs/0907.1656) [[astro-ph.CO](#)] (cit. on p. 34).
- Kormendy, J. et al. (Jan. 1989). ‘Surface photometry and the structure of elliptical galaxies.’ *ARA&A* 27, pp. 235–277. (Cit. on p. 11).

- Kroupa, P. (Apr. 2001). ‘On the variation of the initial mass function’. *MNRAS* 322.2, pp. 231–246. arXiv: [astro-ph/0009005 \[astro-ph\]](#) (cit. on p. 94).
- Kruijssen, J. M. D. et al. (Apr. 2014). ‘An uncertainty principle for star formation - I. Why galactic star formation relations break down below a certain spatial scale’. *MNRAS* 439.4, pp. 3239–3252. arXiv: [1401.4459 \[astro-ph.GA\]](#) (cit. on pp. xviii, 34, 39, 40, 108).
- Kruijssen, J. M. D. et al. (Feb. 2015). ‘The dynamical evolution of molecular clouds near the Galactic Centre - I. Orbital structure and evolutionary timeline’. *MNRAS* 447.2, pp. 1059–1079. arXiv: [1412.0664 \[astro-ph.GA\]](#) (cit. on p. 34).
- Kruijssen, J. M. D. et al. (Sept. 2018). ‘An uncertainty principle for star formation - II. A new method for characterizing the cloud-scale physics of star formation and feedback across cosmic history’. *MNRAS* 479.2, pp. 1866–1952. arXiv: [1805.00012 \[astro-ph.GA\]](#) (cit. on pp. xviii, 34, 39–44, 51, 53, 58, 59, 61, 84, 85, 90, 101, 108, 109, 111, 114, 117, 118, 123).
- Kruijssen, J. M. D., A. Schrubba, M. Chevance et al. (May 2019). ‘Fast and inefficient star formation due to short-lived molecular clouds and rapid feedback’. *Nature* 569.7757, pp. 519–522. arXiv: [1905.08801 \[astro-ph.GA\]](#) (cit. on pp. 40, 44, 49, 51, 89–91, 94, 101, 104, 121–123, 170).
- Kruijssen, J. M. D., J. L. Pfeffer et al. (July 2019). ‘The formation and assembly history of the Milky Way revealed by its globular cluster population’. *MNRAS* 486.3, pp. 3180–3202. arXiv: [1806.05680 \[astro-ph.GA\]](#) (cit. on p. 11).
- Kruit, P. C. van der et al. (Sept. 2011). ‘Galaxy Disks’. *ARA&A* 49.1, pp. 301–371. arXiv: [1101.1771 \[astro-ph.GA\]](#) (cit. on p. 13).
- Krumholz, M. R. (Jan. 2014). ‘DESPOTIC - a new software library to Derive the Energetics and SPectra of Optically Thick Interstellar Clouds’. *MNRAS* 437.2, pp. 1662–1680. arXiv: [1304.2404 \[astro-ph.IM\]](#) (cit. on pp. xviii, 139, 141, 144, 146).
- Krumholz, M. R. et al. (Sept. 2015). ‘SLUG - stochastically lighting up galaxies - III. A suite of tools for simulated photometry, spectroscopy, and Bayesian inference with stochastic stellar populations’. *MNRAS* 452.2, pp. 1447–1467. arXiv: [1502.05408 \[astro-ph.GA\]](#) (cit. on pp. xviii, 58, 102, 105).
- Kuiper, G. P. (Nov. 1938). ‘The Empirical Mass-Luminosity Relation.’ *ApJ* 88, p. 472. (Cit. on p. 24).
- Lada, C. J. (Jan. 1987). ‘Star formation: from OB associations to protostars’. *Star Forming Regions*. Ed. by M. Peimbert et al. Vol. 115. IAU Symposium, p. 1 (cit. on p. 19).
- Larson, R. B. (Dec. 1975). ‘Models for the formation of elliptical galaxies.’ *MNRAS* 173, pp. 671–699. (Cit. on p. 14).

- Laughlin, G. et al. (June 1997). ‘The End of the Main Sequence’. *ApJ* 482.1, pp. 420–432. (Cit. on p. 27).
- Leitherer, C. et al. (July 1999). ‘Starburst99: Synthesis Models for Galaxies with Active Star Formation’. *ApJS* 123.1, pp. 3–40. arXiv: [astro-ph/9902334](#) [[astro-ph](#)] (cit. on pp. xviii, 40, 47, 69).
- Leitherer, C. et al. (Aug. 2010). ‘A Library of Theoretical Ultraviolet Spectra of Massive, Hot Stars for Evolutionary Synthesis’. *ApJS* 189.2, pp. 309–335. arXiv: [1006.5624](#) [[astro-ph.SR](#)] (cit. on p. xviii).
- Leitherer, C. et al. (May 2014). ‘The Effects of Stellar Rotation. II. A Comprehensive Set of Starburst99 Models’. *ApJS* 212.1, 14, p. 14. arXiv: [1403.5444](#) [[astro-ph.GA](#)] (cit. on pp. xviii, 40).
- Leroy, A. K. et al. (July 2012). ‘Estimating the Star Formation Rate at 1 kpc Scales in nearby Galaxies’. *AJ* 144.1, 3, p. 3. arXiv: [1202.2873](#) [[astro-ph.CO](#)] (cit. on pp. 48, 56, 64, 68, 69, 77, 90).
- Liddle, A. (2003). *An introduction to modern cosmology*. 2nd ed. Chichester, UK: Wiley. ISBN: 9780470848357 (cit. on p. 6).
- Marinacci, F. et al. (Nov. 2018). ‘First results from the IllustrisTNG simulations: radio haloes and magnetic fields’. *MNRAS* 480.4, pp. 5113–5139. arXiv: [1707.03396](#) [[astro-ph.CO](#)] (cit. on p. 127).
- Martínez-García, E. E. et al. (Mar. 2009). ‘Spiral Density Wave Triggering of Star Formation in SA and SAB Galaxies’. *ApJ* 694.1, pp. 512–545. arXiv: [0812.3647](#) [[astro-ph](#)] (cit. on p. 13).
- McKee, C. F. et al. (Feb. 1997). ‘The Luminosity Function of OB Associations in the Galaxy’. *ApJ* 476.1, pp. 144–165. (Cit. on pp. 58, 104).
- Mouschovias, T. C. et al. (Dec. 1976). ‘Note on the collapse of magnetic interstellar clouds.’ *ApJ* 210, p. 326. (Cit. on p. 21).
- Murphy, E. J. et al. (Aug. 2011). ‘Calibrating Extinction-free Star Formation Rate Diagnostics with 33 GHz Free-free Emission in NGC 6946’. *ApJ* 737.2, 67, p. 67. arXiv: [1105.4877](#) [[astro-ph.CO](#)] (cit. on p. 47).
- Naab, T. et al. (Aug. 2017). ‘Theoretical Challenges in Galaxy Formation’. *ARA&A* 55.1, pp. 59–109. arXiv: [1612.06891](#) [[astro-ph.GA](#)] (cit. on p. 11).
- Naiman, J. P. et al. (June 2018). ‘First results from the IllustrisTNG simulations: a tale of two elements - chemical evolution of magnesium and europium’. *MNRAS* 477.1, pp. 1206–1224. arXiv: [1707.03401](#) [[astro-ph.GA](#)] (cit. on p. 127).
- Narayanan, D. et al. (May 2017). ‘A physical model for the [C II]-FIR deficit in luminous galaxies’. *MNRAS* 467.1, pp. 50–67. arXiv: [1601.05803](#) [[astro-ph.GA](#)] (cit. on pp. 139, 144).

- NASA / ESA (Oct. 1999). *The Hubble tuning fork*. URL: <https://www.spacetelescope.org/images/heic9902o/> (visited on 09/11/2019) (cit. on p. 12).
- NASA / ESA / Freedman, W. / Hubble Heritage Team / SDSS (Nov. 2016). URL: <https://apod.nasa.gov/apod/ap161120.html> (visited on 09/11/2019) (cit. on p. 15).
- Nelson, D. et al. (Mar. 2018). ‘First results from the IllustrisTNG simulations: the galaxy colour bimodality’. *MNRAS* 475.1, pp. 624–647. arXiv: [1707.03395](https://arxiv.org/abs/1707.03395) [astro-ph.GA] (cit. on p. 127).
- Nelson, D. et al. (Aug. 2019). ‘First Results from the TNG50 Simulation: Galactic outflows driven by supernovae and black hole feedback’. *MNRAS*, p. 2010. arXiv: [1902.05554](https://arxiv.org/abs/1902.05554) [astro-ph.GA] (cit. on pp. 37, 139).
- Nelson, R. P. et al. (Oct. 1999). ‘On the Stability and Evolution of Isolated BOK Globules’. *ApJ* 524.2, pp. 923–946. (Cit. on p. 144).
- Nielsen, E. L. et al. (Dec. 2005). ‘AB Doradus C: age, spectral type, orbit, and comparison to evolutionary models’. *Astronomische Nachrichten* 326.10, pp. 1033–1039. arXiv: [astro-ph/0509400](https://arxiv.org/abs/astro-ph/0509400) [astro-ph] (cit. on p. 29).
- Oesch, P. A. et al. (Mar. 2016). ‘A Remarkably Luminous Galaxy at $z=11.1$ Measured with Hubble Space Telescope Grism Spectroscopy’. *ApJ* 819.2, 129, p. 129. arXiv: [1603.00461](https://arxiv.org/abs/1603.00461) [astro-ph.GA] (cit. on p. 9).
- Oliphant, T. E. (2006). *A guide to NumPy*. Vol. 1. Trelgol Publishing USA (cit. on p. 129).
- Olver, F. W. J. et al., eds. (2010). *NIST handbook of mathematical functions*. 1. publ. Cambridge University Press. ISBN: 9780521192255 (cit. on p. 94).
- Orear, J. (Oct. 1982). ‘Least squares when both variables have uncertainties’. *American Journal of Physics* 50.10, pp. 912–916. (Cit. on p. 85).
- Parzen, E. (Sept. 1962). ‘On Estimation of a Probability Density Function and Mode’. *Ann. Math. Statist.* 33.3, pp. 1065–1076. (Cit. on p. 147).
- Paul, S. et al. (Oct. 2017). ‘Understanding ‘galaxy groups’ as a unique structure in the universe’. *MNRAS* 471.1, pp. 2–11. arXiv: [1706.01916](https://arxiv.org/abs/1706.01916) [astro-ph.CO] (cit. on p. 19).
- Peacock, J. A. (1999). *Cosmological Physics*. Cambridge Astrophysics. Cambridge University Press. ISBN: 9780521422703 (cit. on p. 131).
- Peng, Y. et al. (May 2015). ‘Strangulation as the primary mechanism for shutting down star formation in galaxies’. *Nature* 521.7551, pp. 192–195. arXiv: [1505.03143](https://arxiv.org/abs/1505.03143) [astro-ph.GA] (cit. on p. 16).
- Pillepich, A., V. Springel et al. (Jan. 2018). ‘Simulating galaxy formation with the IllustrisTNG model’. *MNRAS* 473.3, pp. 4077–4106. arXiv: [1703.02970](https://arxiv.org/abs/1703.02970) [astro-ph.GA] (cit. on p. xix).

- Pillepich, A., D. Nelson, L. Hernquist et al. (Mar. 2018). ‘First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies’. *MNRAS* 475.1, pp. 648–675. arXiv: [1707.03406 \[astro-ph.GA\]](#) (cit. on p. 127).
- Pillepich, A. et al. (Dec. 2019). ‘First results from the TNG50 simulation: the evolution of stellar and gaseous discs across cosmic time’. *MNRAS* 490.3, pp. 3196–3233. arXiv: [1902.05553 \[astro-ph.GA\]](#) (cit. on p. 37).
- Planck Collaboration (Nov. 2014). ‘Planck 2013 results. XVI. Cosmological parameters’. *A&A* 571, A16, A16. arXiv: [1303.5076 \[astro-ph.CO\]](#) (cit. on p. 3).
- Planck Collaboration (Sept. 2016). ‘Planck 2015 results. XIII. Cosmological parameters’. *A&A* 594, A13, A13. arXiv: [1502.01589 \[astro-ph.CO\]](#) (cit. on pp. 5–7, 127, 129).
- Poggianti, B. M. et al. (July 2017). ‘GASP. I. Gas Stripping Phenomena in Galaxies with MUSE’. *ApJ* 844.1, 48, p. 48. arXiv: [1704.05086 \[astro-ph.GA\]](#) (cit. on p. 16).
- Popping, G. et al. (Feb. 2019). ‘The art of modelling CO, [C I], and [C II] in cosmological galaxy formation models’. *MNRAS* 482.4, pp. 4906–4932. arXiv: [1805.11093 \[astro-ph.GA\]](#) (cit. on p. 141).
- Rossum, G. van et al. (Dec. 1991). ‘Interactively testing remote servers using the Python programming language’. *CWI Quarterly* 4.4, pp. 283–304 (cit. on p. 128).
- Sandage, A. et al. (June 1970). ‘The Intrinsic Flattening of e, so, and Spiral Galaxies as Related to Galaxy Formation and Evolution’. *ApJ* 160, p. 831. (Cit. on p. 16).
- Safranek-Shrader, C. et al. (Feb. 2017). ‘Chemistry and radiative shielding in star-forming galactic discs’. *MNRAS* 465.1, pp. 885–905. arXiv: [1605.07618 \[astro-ph.GA\]](#) (cit. on p. 105).
- Saintonge, A. et al. (Dec. 2017). ‘xCOLD GASS: The Complete IRAM 30 m Legacy Survey of Molecular Gas for Galaxy Evolution Studies’. *ApJS* 233.2, 22, p. 22. arXiv: [1710.02157 \[astro-ph.GA\]](#) (cit. on p. xix).
- Salpeter, E. E. (Jan. 1955). ‘The Luminosity Function and Stellar Evolution.’ *ApJ* 121, p. 161. (Cit. on p. 28).
- Schaerer, D., G. Meynet et al. (May 1993). ‘Grids of stellar models. II. From 0.8 to 120 M_{solar} at Z=0.008’. *A&AS* 98, p. 523. ADS: [1993A&AS...98..523S](#) (cit. on p. 70).
- Schaerer, D., C. Charbonnel et al. (Dec. 1993). ‘Grids of stellar models. IV. From 0.8 to 120 M_{solar} at Z=0.040’. *A&AS* 102, p. 339. ADS: [1993A&AS..102..339S](#) (cit. on p. 70).

- Schaller, G. et al. (Dec. 1992). ‘New grids of stellar models from 0.8 to 120 M_{solar} at $Z=0.020$ and $Z=0.001$ ’. *A&AS* 96, p. 269. ADS: [1992A&AS...96..269S](#) (cit. on pp. 30, 58, 70, 104).
- Schawinski, K. et al. (May 2014). ‘The green valley is a red herring: Galaxy Zoo reveals two evolutionary pathways towards quenching of star formation in early- and late-type galaxies’. *MNRAS* 440.1, pp. 889–907. arXiv: [1402.4814 \[astro-ph.GA\]](#) (cit. on pp. 16, 17, 19).
- Schaye, J. et al. (Mar. 2010). ‘The physics driving the cosmic star formation history’. *MNRAS* 402.3, pp. 1536–1560. arXiv: [0909.5196 \[astro-ph.CO\]](#) (cit. on p. xix).
- Schaye, J. et al. (Jan. 2015). ‘The EAGLE project: simulating the evolution and assembly of galaxies and their environments’. *MNRAS* 446.1, pp. 521–554. arXiv: [1407.7040 \[astro-ph.GA\]](#) (cit. on p. 126).
- Schechter, P. (Jan. 1976). ‘An analytic expression for the luminosity function for galaxies.’ *ApJ* 203, pp. 297–306. (Cit. on p. 14).
- Schmidt, M. (Mar. 1959). ‘The Rate of Star Formation.’ *ApJ* 129, p. 243. (Cit. on p. 31).
- Schöier, F. L. et al. (Mar. 2005). ‘An atomic and molecular database for analysis of submillimetre line observations’. *A&A* 432.1, pp. 369–379. arXiv: [astro-ph/0411110 \[astro-ph\]](#) (cit. on p. 145).
- Schwarzschild, M. (1958). *Structure and Evolution of the Stars*. Princeton University Press. ISBN: 9780486614793 (cit. on p. 25).
- Scott, D. W. (1992). *Multivariate Density Estimation. Theory, Practice, and Visualization*. Wiley Series in Probability and Statistics. Wiley. ISBN: 9780471547709. DOI: [10.1002/9780470316849](#) (cit. on p. 147).
- Scoville, N. Z. et al. (Jan. 1979). ‘CO Observations of Spiral Structure and the Lifetime of Giant Molecular Clouds’. *The Large-Scale Characteristics of the Galaxy*. Ed. by W. B. Burton. Vol. 84. IAU Symposium, p. 277. ADS: [1979IAUS...84..277S](#) (cit. on p. 34).
- Shu, F. H. (Sept. 2016). ‘Six Decades of Spiral Density Wave Theory’. *ARA&A* 54, pp. 667–724. (Cit. on p. 13).
- Sijacki, D. et al. (Sept. 2015). ‘The Illustris simulation: the evolving population of black holes across cosmic time’. *MNRAS* 452.1, pp. 575–596. arXiv: [1408.6842 \[astro-ph.GA\]](#) (cit. on p. 9).
- Silva, R. L. da et al. (Feb. 2012). ‘SLUG - Stochastically Lighting Up Galaxies. I. Methods and Validating Tests’. *ApJ* 745.2, 145, p. 145. arXiv: [1106.3072 \[astro-ph.IM\]](#) (cit. on p. xviii).
- Silva, R. L. da et al. (Nov. 2014). ‘SLUG - Stochastically Lighting Up Galaxies - II. Quantifying the effects of stochasticity on star formation rate indicators’.

- MNRAS* 444.4, pp. 3275–3287. arXiv: [1403.4605 \[astro-ph.GA\]](#) (cit. on pp. xviii, 58, 105).
- Springel, V. et al. (Dec. 2001). ‘Populating a cluster of galaxies - I. Results at $z=0$ ’. *MNRAS* 328.3, pp. 726–750. arXiv: [astro-ph/0012055 \[astro-ph\]](#) (cit. on p. 128).
- Springel, V. et al. (Feb. 2003). ‘Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation’. *MNRAS* 339.2, pp. 289–311. arXiv: [astro-ph/0206393 \[astro-ph\]](#) (cit. on p. 128).
- Springel, V. et al. (Aug. 2005). ‘Modelling feedback from stars and black holes in galaxy mergers’. *MNRAS* 361.3, pp. 776–794. arXiv: [astro-ph/0411108 \[astro-ph\]](#) (cit. on p. 51).
- Springel, V. (Dec. 2005). ‘The cosmological simulation code GADGET-2’. *MNRAS* 364.4, pp. 1105–1134. arXiv: [astro-ph/0505010 \[astro-ph\]](#) (cit. on p. 51).
- Springel, V. (Jan. 2010). ‘E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh’. *MNRAS* 401.2, pp. 791–851. arXiv: [0901.4107 \[astro-ph.CO\]](#) (cit. on p. 126).
- Springel, V. et al. (Mar. 2018). ‘First results from the IllustrisTNG simulations: matter and galaxy clustering’. *MNRAS* 475.1, pp. 676–698. arXiv: [1707.03397 \[astro-ph.GA\]](#) (cit. on p. 127).
- Steigman, G. (Oct. 2006). ‘The cosmological evolution of the average mass per baryon’. *J. Cosmology Astropart. Phys.* 2006.10, 016, p. 016. arXiv: [astro-ph/0606206 \[astro-ph\]](#) (cit. on p. 5).
- Tamburro, D. et al. (Dec. 2008). ‘Geometrically Derived Timescales for Star Formation in Spiral Galaxies’. *AJ* 136.6, pp. 2872–2885. arXiv: [0810.2391 \[astro-ph\]](#) (cit. on p. 34).
- Tasker, E. J. et al. (July 2009). ‘Star Formation in Disk Galaxies. I. Formation and Evolution of Giant Molecular Clouds via Gravitational Instability and Cloud Collisions’. *ApJ* 700.1, pp. 358–375. arXiv: [0811.0207 \[astro-ph\]](#) (cit. on p. 101).
- Toomre, A. (May 1964). ‘On the gravitational stability of a disk of stars.’ *ApJ* 139, pp. 1217–1238. ADS: [1964ApJ...139.1217T](#) (cit. on pp. 21, 102).
- Toomre, A. et al. (Dec. 1972). ‘Galactic Bridges and Tails’. *ApJ* 178, pp. 623–666. (Cit. on p. 16).
- Torres-Flores, S. et al. (Sept. 2011). ‘GHASP: an $H\alpha$ kinematic survey of spiral and irregular galaxies - IX. The near-infrared, stellar and baryonic Tully-Fisher relations’. *MNRAS* 416.3, pp. 1936–1948. arXiv: [1106.0505 \[astro-ph.CO\]](#) (cit. on p. 13).
- Tully, R. B. et al. (Feb. 1977). ‘A new method of determining distance to galaxies.’ *A&A* 500, pp. 105–117 (cit. on p. 13).

- Tumlinson, J. et al. (Aug. 2017). ‘The Circumgalactic Medium’. *ARA&A* 55.1, pp. 389–432. arXiv: [1709.09180 \[astro-ph.GA\]](#) (cit. on p. 18).
- Turk, M. J. et al. (Jan. 2011). ‘yt: A Multi-code Analysis Toolkit for Astrophysical Simulation Data’. *ApJS* 192.1, 9, p. 9. arXiv: [1011.3514 \[astro-ph.IM\]](#) (cit. on p. 105).
- Umehata, H. et al. (Oct. 2019). ‘Gas filaments of the cosmic web located around active galaxies in a protocluster’. *Science* 366.6461, pp. 97–100. arXiv: [1910.01324 \[astro-ph.GA\]](#) (cit. on p. 9).
- Vaucouleurs, G. de et al. (1991). *Third Reference Catalogue of Bright Galaxies*. Springer, New York, NY. ISBN: 9781475743654. DOI: [10.1007/978-1-4757-4363-0](#) (cit. on p. 15).
- Vázquez, G. A. et al. (Mar. 2005). ‘Optimization of Starburst99 for Intermediate-Age and Old Stellar Populations’. *ApJ* 621.2, pp. 695–717. arXiv: [astro-ph/0412491 \[astro-ph\]](#) (cit. on p. xviii).
- Vogelsberger, M., S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, S. Bird et al. (May 2014). ‘Properties of galaxies reproduced by a hydrodynamic simulation’. *Nature* 509.7499, pp. 177–182. arXiv: [1405.1418 \[astro-ph.CO\]](#) (cit. on p. 9).
- Vogelsberger, M., S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson et al. (Oct. 2014). ‘Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe’. *MNRAS* 444.2, pp. 1518–1547. arXiv: [1405.2921 \[astro-ph.CO\]](#) (cit. on p. 9).
- Walch, S. et al. (Nov. 2015). ‘The SILCC (Simulating the LifeCycle of molecular Clouds) project - I. Chemical evolution of the supernova-driven ISM’. *MNRAS* 454.1, pp. 238–268. arXiv: [1412.2749 \[astro-ph.GA\]](#) (cit. on p. xix).
- Walter, F. et al. (Dec. 2016). ‘ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: Survey Description’. *ApJ* 833.1, 67, p. 67. arXiv: [1607.06768 \[astro-ph.GA\]](#) (cit. on pp. xix, 125).
- Ward, J. L. et al. (2019). ‘Towards a multi-tracer timeline of star formation in the LMC I: Deriving the lifetimes of H I clouds’. *To be submitted* (cit. on pp. 94, 170).
- Weinberger, R. et al. (Mar. 2017). ‘Simulating galaxy formation with black hole driven thermal and kinetic feedback’. *MNRAS* 465.3, pp. 3291–3308. arXiv: [1607.03486 \[astro-ph.GA\]](#) (cit. on p. xix).
- White, M. et al. (May 1997). ‘The Sachs-Wolfe effect.’ *A&A* 321, pp. 8–9. arXiv: [astro-ph/9609105 \[astro-ph\]](#) (cit. on p. 6).
- Wright, A. C. et al. (Jan. 2019). ‘Reignition of star formation in dwarf galaxies’. *MNRAS* 482.1, pp. 1176–1189. arXiv: [1802.03019 \[astro-ph.GA\]](#) (cit. on p. 16).

York, D. G. et al. (Sept. 2000). ‘The Sloan Digital Sky Survey: Technical Summary’. *AJ* 120.3, pp. 1579–1587. arXiv: [astro-ph/0006396](https://arxiv.org/abs/astro-ph/0006396) [[astro-ph](#)] (cit. on p. [xix](#)).

