An uncertainty principle for star formation – III. The characteristic emission time-scales of star formation rate tracers

Daniel T. Haydon [1], J. M. Diederik Kruijssen [1,2]★ Mélanie Chevance,[1] Alexander P. S. Hygate [1,2]
Mark R. Krumholz [3,4] Andreas Schruba[5] and Steven N. Longmore[6]

[1] *Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12–14, D-69120 Heidelberg, Germany*
[2] *Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[3] *Research School of Astronomy and Astrophysics, Australian National University, Canberra 2611, ACT, Australia*
[4] *Centre of Excellence for Astronomy in Three Dimensions (ASTRO-3D), Australia*
[5] *Max-Planck Institut für Extraterrestrische Physik, Giessenbachstraße 1, D-85748 Garching, Germany*
[6] *Astrophysics Research Institute, Liverpool John Moores University, IC2, Liverpool Science Park, 146 Brownlow Hill, Liverpool L3 5RF, UK*

ABSTRACT
We recently presented a new statistical method to constrain the physics of star formation and feedback on the cloud scale by reconstructing the underlying evolutionary timeline. However, by itself this new method only recovers the *relative* durations of different evolutionary phases. To enable observational applications, it therefore requires knowledge of an absolute 'reference time-scale' to convert relative time-scales into absolute values. The logical choice for this reference time-scale is the duration over which the star formation rate (SFR) tracer is visible because it can be characterized using stellar population synthesis (SPS) models. In this paper, we calibrate this reference time-scale using synthetic emission maps of several SFR tracers, generated by combining the output from a hydrodynamical disc galaxy simulation with the SPS model SLUG2. We apply our statistical method to obtain self-consistent measurements of each tracer's reference time-scale. These include Hα and 12 ultraviolet (UV) filters (from *GALEX*, *Swift*, and *HST*), which cover a wavelength range 150–350 nm. At solar metallicity, the measured reference time-scales of Hα are $4.32^{+0.09}_{-0.23}$ Myr with continuum subtraction, and 6–16 Myr without, where the time-scale increases with filter width. For the UV filters we find 17–33 Myr, nearly monotonically increasing with wavelength. The characteristic time-scale decreases towards higher metallicities, as well as to lower star formation rate surface densities, owing to stellar initial mass function sampling effects. We provide fitting functions for the reference time-scale as a function of metallicity, filter width, or wavelength, to enable observational applications of our statistical method across a wide variety of galaxies.

Key words: H II regions – galaxies: evolution – galaxies: ISM – galaxies: star formation – galaxies: stellar content.

1 INTRODUCTION

It is a challenging problem in astrophysics to characterize the time-scales over which astrophysical processes take place, because most of these processes take much longer than a human lifetime. For this reason, star formation studies have struggled to identify the physical processes governing the evolution of molecular clouds and star-forming regions, which requires knowledge of the underlying time-scales (e.g. Dobbs et al. 2014; Krumholz 2014; Chevance et al. 2020b). Traditionally, the age of the stellar population has been used as a 'reference time-scale' to infer how long other phases of the star formation lifecycle take, such as the molecular cloud lifetime and the time over which gas and young stars are associated (e.g. Leisawitz, Bash & Thaddeus 1989; Elmegreen 2000; Kawamura et al. 2009). However, the time-scales measured to date span a wide range of durations (e.g. Scoville, Solomon & Sanders 1979; Koda et al. 2009; Meidt et al. 2015), which is largely caused by the heterogeneity of the

methods used (see the discussion in the methods section of Kruijssen et al. 2019). Clearly, a systematic framework is needed for placing the lifecycle of molecular clouds, star formation, and feedback on an absolute, empirically determined evolutionary timeline.

In Kruijssen & Longmore (2014), we put forward a new statistical method, titled 'an uncertainty principle for star formation' (hereafter KL14 principle), formalized in the HEISENBERG code (Kruijssen et al. 2018). This analysis method enables the use of pairs of high-resolution emission maps tracing successive phases of the evolutionary cycle between cloud evolution, star formation, and feedback (e.g. CO tracing molecular gas and Hα tracing recent star formation) to infer the relative durations of these phases on the cloud scale. Specifically, this method can be used to measure the relative duration of the 'cloud lifetime' and the relative duration over which a molecular cloud is disrupted by stellar feedback, compared to the characteristic 'H II region lifetime'.

To turn these resulting *relative* durations into an absolute timeline, it is critical to normalize the timeline using a known 'reference time-scale'. Without this reference time-scale, observational applications of the KL14 principle cannot be used to obtain meaningful constraints

★ E-mail: kruijssen@uni-heidelberg.de

on the cloud lifecycle. The most direct way of providing a reference time-scale is by following the example of previous work in this area and characterizing the time-scales of star formation rate (SFR) tracers. Indeed, our new methodology is preceded (and partially inspired) by a wide variety of literature aiming to observationally characterize the cloud lifecycle, which all used the approximate lifetimes of H II regions or the ages of young stellar clusters as a reference time-scale (e.g. Blitz et al. 2006; Kawamura et al. 2009; Miura et al. 2012; Corbelli et al. 2017). In this paper, we calibrate this approach for use in observational applications of the KL14 principle. Doing so requires a controlled experiment, in which the duration of one phase is known exactly and the duration of the other phase is measured from its emission map using HEISENBERG. This requires the use of galaxy simulations rather than real galaxies. Since the reference time-scale likely depends on metallicity and is affected by the sampling of the stellar initial mass function (IMF), the experiment must also be repeated for different tracers, such as H α and various ultraviolet (UV) filters, metallicities, and degrees of IMF sampling.

Even though it is the goal of this paper to define the tracer time-scales within the context of the KL14 principle, measuring the characteristic emission time-scales of SFR tracers is also important in other contexts. For instance, this characteristic time-scale is an indicator of the duration for which photoionizing feedback can act on the surrounding interstellar medium. Deriving absolute SFRs from observed line or broadband emission flux also requires calibration by the time-scale over which a young stellar population emits at that given wavelength. The emission in different wavebands is dominated by different types of stars (Hao et al. 2011) and so it is possible to determine relationships between the luminosity at a given wavelength and the SFR based on the lifetimes of these stars (e.g. Calzetti et al. 2007; Hao et al. 2011; Murphy et al. 2011; Kennicutt & Evans 2012). For example, producing H α emission requires ionizing photons from very massive stars, which have lifetimes <10 Myr (Leitherer et al. 1999; Murphy et al. 2011), such that the emission itself should fade on a characteristic time-scale that is of a similar magnitude.

We emphasize that by describing the duration of SFR tracer emission with a single time-scale, we do not assume that the SFR tracer emission sharply drops at some particular age. Instead, we describe the gradual fading of emission in terms of a single time-scale that is meaningful in the context of our statistical method. Conceptually, this can be regarded as analogous (but not equal) to the e-folding time of an exponential decay, or the time at which 50 per cent of the total H α luminosity that will be produced by a young stellar population has been emitted. This defines a single number for the time-scale of emission for that population. Since the emission fades gradually, it depends on the physical context what definition of a single time-scale is correct. Therefore, our goal is not to provide a general definition of the time-scale of SFR tracer emission, but to provide the right definition of an SFR tracer time-scale for use in observational applications of the KL14 principle.

Previous work attempting to derive characteristic time-scales for different SFR tracers has revealed that the major problem obstructing a conclusive measurement is that there is no obvious definition of the time-scale that should be adopted. Instead, there exists a range of possible definitions, such as a 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 time-scale (Kennicutt & Evans 2012; Leroy et al. 2012). In view of this strong dependence on the precise definition of the reference time-scale, we opt to use a self-consistent approach for determining the SFR tracer time-scale; that is, we measure the emission time-scales of SFR tracers by applying the

KL14 principle itself to synthetic emission maps, which have been generated by combining the output from a hydrodynamical disc galaxy simulation with the stellar population synthesis (SPS) model SLUG2 (da Silva, Fumagalli & Krumholz 2012, 2014; Krumholz et al. 2015). The reference time-scales obtained this way critically enable observational applications of the KL14 principle, which provides measurements of a wide variety of physical quantities as part of a single analysis (Kruijssen et al. 2018), such as the molecular cloud lifetime, the time-scale for cloud destruction by feedback, the separation length between independent star-forming regions, the integrated cloud-scale star formation efficiency, and the feedback outflow velocity. The method has also been extended to also provide physically motivated measurements of the diffuse gas fraction (e.g. of molecular, atomic, or ionized gas; Hygate et al. 2019). For the first applications of the method measuring these quantities across nearby star-forming galaxies, we refer the reader to Kruijssen et al. (2019) and Chevance et al. (2020c).

The structure of this paper is as follows. In Section 2, we summarize the KL14 principle and the practical application of the associated HEISENBERG code. We outline our approach for constraining the characteristic time-scales of different SFR tracers with well-sampled IMFs in Section 3. For solar metallicity, the resulting time-scales are presented in Section 4. In Section 5, we demonstrate how the time-scales depend on metallicity. In Section 6, we demonstrate the effects of incomplete IMF sampling, which is expected to change the results in environments of low SFR surface density. In Section 7, we carry out a brief test of the obtained time-scales, by comparing them to observations of H α and UV emission in NGC 300. Finally, we summarize the results and present our conclusions in Section 8.

2 UNCERTAINTY PRINCIPLE FOR STAR FORMATION

The analysis presented in this paper is based on the KL14 principle and its specific realization in the HEISENBERG code (Kruijssen et al. 2018). The statistical method presented by Kruijssen & Longmore (2014) and Kruijssen et al. (2018) enables the characterization of the cloud-scale physics of star formation and feedback in a systematic way, based on the spatial distribution of emission in pairs of maps tracing particular evolutionary stages of the star formation process. We briefly describe the concept of the method and summarize its specific application in this work, aimed at constraining the characteristic time-scales of SFR tracers.

The KL14 principle has been primarily used to determine the lifetime of giant molecular clouds in nearby galaxies, by comparing the spatial distributions of the emission of a molecular gas tracer (e.g. CO) and SFR tracer (e.g. H α). The first applications of this method (Kruijssen et al. 2019; Hygate et al. 2020; Chevance et al. 2020a, c; Ward et al. 2020a; Zabel et al. 2020) demonstrate that clouds are highly dynamic and, once they host unembedded massive stars, are dispersed on time-scales shorter than the typical supernova delay time, implying that early feedback by photoionization or stellar winds plays a critical role in disrupting molecular clouds. However, the applicability of the method is not restricted to maps of molecular gas or young stellar emission. Depending on the combination of tracers used, it is possible to constrain the durations of different stages of the star formation timeline using HEISENBERG, such as the atomic cloud lifetime (Ward et al. 2020b) or the duration of the embedded phase of star formation (Kim et al. 2020).

In basic terms, the KL14 principle represents a galaxy as a collection of independent (e.g. star-forming) regions, where each region is evolving along its timeline independently of its neighbouring regions.

The number of regions that are emitting in each of the two tracers (with the possibility that some regions are in a transition phase and emit in both tracers) is roughly related to the duration of that phase – the shorter the duration of a phase, the less likely it is to observe a region in that phase.

Most importantly in the context of this paper, this method fundamentally derives the duration of one phase *relative* to another one. The duration of one of the two phases must therefore be known a priori in order to derive *absolute* time-scales. In practice, this 'reference time-scale' can often be associated with the duration of the emission of the SFR tracer, owing to the absolute clock provided by stellar evolution (e.g. Leitherer et al. 1999, 2014). In this paper, we use this clock to measure the reference time-scale of SFR tracers and thereby provide a calibration of the evolutionary timeline between molecular clouds, star formation, and feedback that can be measured observationally using HEISENBERG. Throughout this work, we call 'reference map' the emission map associated with the phase of known duration, and we refer to this duration as the 'reference time-scale'.

We briefly summarize 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 & Longmore (2014), where it is shown that the gas-to-young stellar flux ratio changes relative to the galactic average when focusing apertures on peaks of gas or young stellar emission, and 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:

(i) We use two emission maps of the same galaxy, tracing two successive evolutionary phases (hereafter phase 1 and phase 2). The phase 2 map is used as the reference map, with its duration equal to the reference time-scale, t_R. In this work, we construct a map of star particles within a given age range as the phase 2 map of known duration (see Section 3) and provide a synthetic SFR tracer map as phase 1, of which we measure the duration (see below).

(ii) Each map is convolved using a top hat kernel for a series of aperture sizes (ranging from ∼ the cloud scale to ∼ the galaxy scale).

(iii) For each of these convolution scales, apertures of corresponding sizes are placed on the emission peaks of both convolved maps. The enclosed phase-1-to-phase-2 flux ratio (F_{12}) within these apertures is then measured in units of the galactic average flux ratio ($F_{12,gal}$). On the small scales, the measured flux ratio in apertures centred on phase 1 peaks (respectively, phase 2 peaks) deviates from the galactic average value, as visible in Fig. 1.

(iv) The models describing the shapes of the two branches of this 'tuning fork diagram' (see Kruijssen et al. 2018, equations 81 and 82) are fitted to the measurements in order to constrain the following three free parameters: the typical separation length between identified peaks (λ), the relative temporal overlap between the two phases (t_{over}/t_R), and the relative duration of phase 1 (t_1/t_R).[1] The absolute time-scales t_1 and t_{over} require a reference time-scale t_R.

The reference time-scale is a key ingredient for retrieving the absolute time-scales t_1 and t_{over}. In this paper, we use the above method to calibrate the duration of emission for a variety of SFR tracers, which can then be used as the reference time-scale in
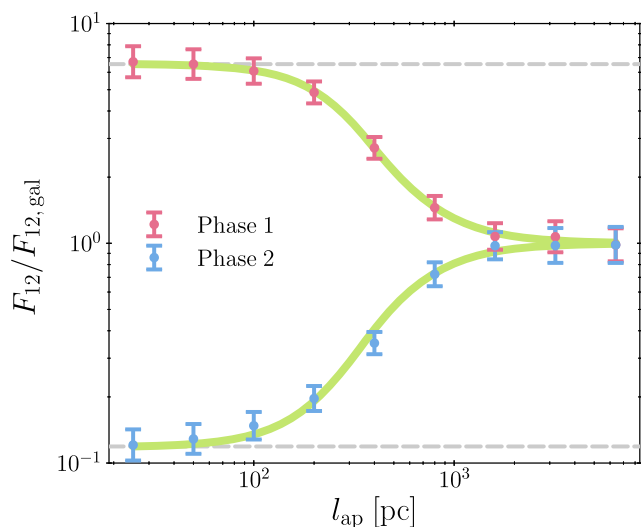

Figure 1. Example tuning fork diagram produced by the HEISENBERG code. The figure shows the relative change of the phase-1-to-phase-2 flux ratio (F_{12}) when placing apertures of size l_{ap} at the locations of emission peaks in the phase 1 map (top branch, in red) or the phase 2 map (bottom branch, in blue), compared to the galactic-scale phase-1-to-phase-2 flux ratio ($F_{12,gal}$) as a function l_{ap}. The best-fitting model is indicated by the two green curves.

observational applications. In Section 3, we describe how we use a simulated galaxy to create reference maps from stars within a known age bin (the duration of which is used as t_R), and synthetic SFR tracer emission maps. By adopting one of the synthetic SFR tracer emission maps as the phase 1 map and a reference map with a known age interval as the phase 2 map, we have (by definition) two maps that trace successive phases of an evolutionary timeline. We then use HEISENBERG to constrain their relative lifetimes as well as the absolute duration of phase 1, t_1.

The above approach implicitly assumes that the method is commutative and transitive, i.e. we can use stars in a known age range (A) to calibrate an SFR tracer time-scale (B), which then is used in observational applications to characterize e.g. the molecular cloud lifetime (C), so that effectively the known age range (A) is used to calibrate the cloud lifetime (C). Both the commutativity (Kruijssen et al. 2018) and transitivity (Ward et al. 2020b) of the method have been demonstrated in other papers, which justifies its use in this work. The fact that the method is transitive and commutative follows somewhat trivially from its ability to predict correct lifetimes, as demonstrated by Kruijssen et al. (2018). If the method correctly predicts a time-scale ratio t_A/t_B, then it must also correctly predict the time-scale ratio t_B/t_A. Likewise, if the method correctly predicts the time-scale ratios t_A/t_B and t_B/t_C, then it must also correctly predict the time-scale ratio t_A/t_C.

3 METHOD FOR CALCULATING THE CHARACTERISTIC EMISSION TIME-SCALES OF SFR TRACERS FOR A FULLY SAMPLED IMF

We present here the steps we take to find the characteristic time-scales for Hα and UV SFR tracers (see Table 1 for details) using synthetic emission maps and the HEISENBERG code. As we described in Section 2, 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

---

[1] These quantities represent flux-weighted averages across the population of emission peaks (see sections 3.2.9 and 3.2.11 of Kruijssen et al. 2018 and equation 1 of Kruijssen et al. 2019). In addition, λ describes the typical separation length in the vicinity of peaks, not the area-averaged value across an entire galaxy, making it relatively insensitive to morphological features such as spiral arms (Kruijssen et al. 2018, 2019; Chevance et al. 2020c).

Table 1. We derive characteristic time-scales for the SFR tracers listed here.

Telescope	Instrument	Filter	$\overline{\lambda}_w$ [nm]
(a) The UV filters we consider. $\overline{\lambda}_w$ is the response-weighted mean wavelength of the filter. The normalized filter response curves are presented in Section 4.			
GALEX		FUV	153.9
GALEX		NUV	231.6
Swift	UVOT	*M*2	225.6
Swift	UVOT	*W*1	261.7
Swift	UVOT	*W*2	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 3.3 for details.
H α + *W*	A narrow band filter including H α and the continuum as defined in equation (3). The total filter width is indicated by *W*; we consider $W = \{10, 20, 40, 80, 160\}$ Å.

means that if we provide HEISENBERG with a galaxy map of one of the SFR tracers (e.g. H α) along with a reference map, HEISENBERG can provide us with the time-scale associated with that SFR tracer. This approach to measure the SFR tracer time-scales ensures that the obtained reference time-scales are self-consistent within the context of our method. After all, the SFR tracer will be applied as the reference time-scale in future observational applications of HEISENBERG.

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 of this paper, 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, by using stellar particles of a specified age range. The SFR tracer maps are generated using an SPS model. This approach allows us to additionally quantify the effects of metallicity (see Section 5) and IMF sampling (see Section 6) on the SFR tracer time-scale. In turn, this will facilitate observational applications of HEISENBERG to a variety of galactic environments.

We discuss the adopted galaxy simulation in Section 3.1, the method for generating the reference maps in Section 3.2, and the method for generating the synthetic SFR tracer maps in Section 3.3.

3.1 Galaxy simulation

The results in this paper 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, Di Matteo & Hernquist (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 9×10^5 $M_\odot$ and the star and gas particle types both have a mass of 2.7×10^3 $M_\odot$. This gives us a 9×10^{11} $M_\odot$ halo, 1.05×10^{10} $M_\odot$ disc (60 per cent in stars and 40 per cent in gas), and 2.7×10^8 $M_\odot$ bulge.

We then evolve the initial conditions for 2.2 Gyr using the smoothed particle hydrodynamics (SPH) 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 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 $M_\odot$ yr^{-1} with a stellar mass of 6.6×10^9 $M_\odot$ and a total cold gas mass of 4.2×10^9 $M_\odot$. These macroscopic galaxy properties are consistent with those of the observed nearby galaxy population, as it resides on the star formation main sequence (Saintonge et al. 2017; Catinella et al. 2018), with a normal total gas depletion time (Bigiel et al. 2008; Leroy et al. 2013). In addition, Fig. 2 shows a stellar reference map (Section 3.2) and a synthetic H α − map (Section 3.3) of this galaxy, demonstrating that its morphology is similar to that of nearby flocculent spirals like M33 and NGC 300.

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 et al. 2019; Chevance et al. 2020b). 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 (1) the birth sites of star particles approximately conform to a galaxy-like morphology and (2) the formation of young stellar populations proceeds approximately instantaneously.

As shown by the images in Fig. 2, the former of the above conditions is indeed achieved by the simulation used here. The additional condition that the stellar population within a cloud forms approximately instantaneously is important, because we aim to measure the SFR tracer emission time-scale for a simple stellar population with a single age. In observational applications of the method, it is certainly possible that the young stellar population generating the SFR tracer emission has a non-zero age spread. For this reason, observational applications include an 'overlap' time-scale that is added to the SFR tracer emission time-scale. This overlap time-scale represents the time for which gas and SFR tracer emission coexist, and is assumed to roughly correspond to the age spread of the stellar population. As a result, the total emission time-scale of the SFR tracer in observations is the sum of the time-scale measured in this paper *and* the overlap time-scale that is observationally inferred with HEISENBERG. As discussed in Kruijssen et al. (2018, Section 4.3.3), the first 1–2 star particles to form in a simulated cloud typically destroy it, such that the cloud-
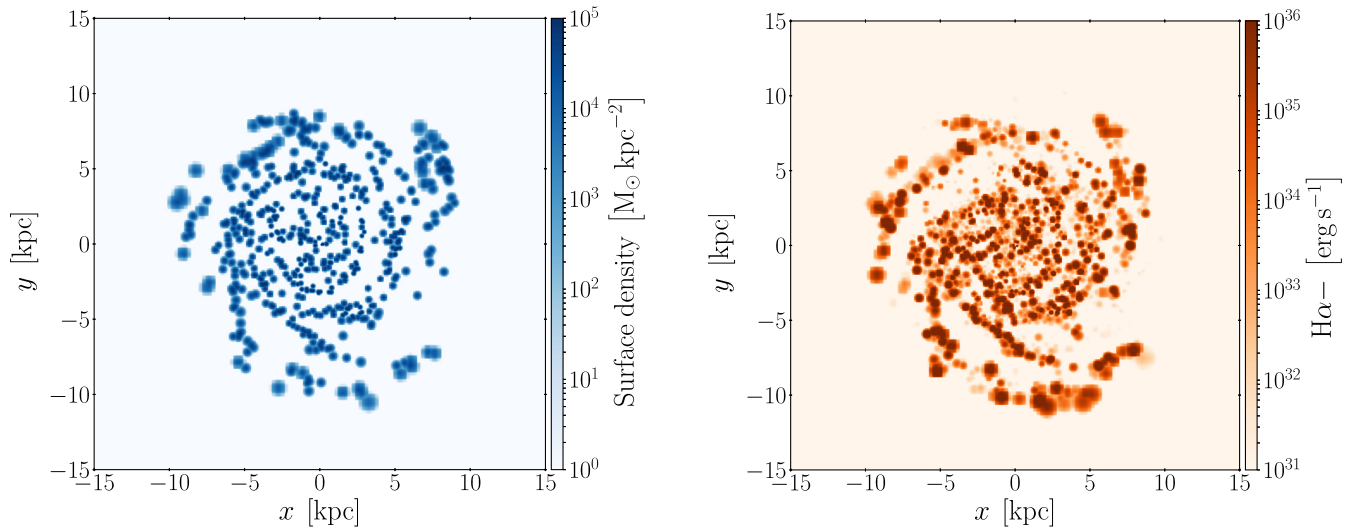
Figure 2. Example maps we use as input for the HEISENBERG code. *Left:* A reference map generated using the mass surface density of star particles in the age range 10–15 Myr, implying a reference time-scale of 5 Myr in this example. See Section 3.2 for details. *Right:* A synthetic Hα emission map without the continuum (H$\alpha-$) generated by performing SPS on the simulated galaxy. See Section 3.3 for details.

scale star formation in the simulation is effectively instantaneous, as desired.

In summary, the simulation satisfies the above requirements for reliably constraining SFR tracer time-scales. In turn, this will enable observational applications of our method that themselves will motivate a future generation of star formation and feedback models, capable of describing cloud-scale evolutionary cycling in galaxy simulations (see Kruijssen et al. 2018 and Fujimoto et al. 2019 for a discussion).

3.2 Generation of the 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 paper, it is used to calibrate the characteristic time-scale of the synthetic SFR tracer emission maps.

In our experiments aimed at measuring the SFR tracer emission time-scales, we need to know the reference time-scale 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 bin. This age bin is chosen to cover a relatively narrow and young age interval that is similar in duration to (but temporally offset from) the stellar age interval at which the SFR tracer emission is generated (see below). The width of this age bin acts as the reference time-scale, t_R. We smoothen the selected star particles using a Wendland C^4 kernel (Dehnen & Aly 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).

As long as the galaxy-average physical conditions do not evolve significantly, and the galaxy is large enough to contain a statistical sample of star-forming regions ($\gtrsim$35, see Kruijssen et al. 2018), the above approach provides a reliable reference time-scale. In that case, the number of star-forming regions in a given age interval is simply proportional to the width of that age interval, irrespectively of the absolute age. This is exactly the setup that we require. The reference maps considered in this work contain well over 35 regions (see Fig. 2 and Kruijssen et al. 2018) and do not experience any significant macroscopic evolution over the time-scales considered.

In principle, we have a free choice over the age bin we use. However, for the best results and the most realistic set-up there are a few restrictions. In Section 2, 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 6) of the evolutionary timeline. This defines the lower limit of the stellar age bin used to generate the reference map:

$$t_M \gtrsim t_{E,0}. \tag{1}$$

This maximizes the diagnostic power of the fit by minimizing the temporal overlap between both phases, thereby avoiding strongly flattened tuning fork diagrams (see Fig. 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 $t_M \approx t_{E,0}$.

Kruijssen et al. (2018) show that HEISENBERG provides the most accurate measurement of the underlying time-scales if the duration associated to both of the input maps is similar (within a factor of 10, but ideally within a factor of 4). This finding sets the preferred width of the age bin:

$$t_R \approx t_{E,0}. \tag{2}$$

This maximizes the diagnostic power of the fit by favouring similar durations of both phases, thereby avoiding strongly asymmetric tuning fork diagrams (see Fig. 1). To summarize the above definitions, Fig. 3 shows 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 time-scale, we investigate the dependence on the choice of stellar age bin 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 2. These are guided by the range of possible characteristic time-scales for Hα and far-ultraviolet (FUV) emission found in Leroy et al. (2012). Leroy et al. use the results of STARBURST99 calculations to determine a characteristic time-scale using several methods: a luminosity-weighted average time,
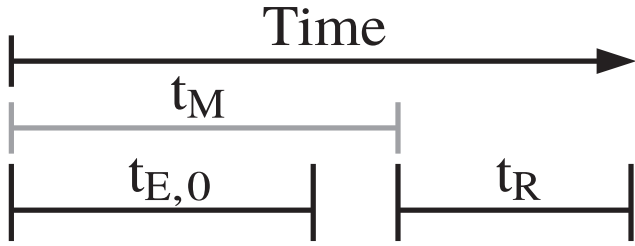
Figure 3. Schematic diagram showing how the different time-scales we define within the paper are related. Time starts at the birth of the star particle. The emission map shows the particles formed within a time-scale $t_{E,0}$ prior to the simulation snapshot, where $t_{E,0}$ represents the characteristic time-scale 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 bin and therefore reference time-scale. 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 time-scale ($t_{E,0}$), 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.

Table 2. We create input reference maps from the star particles that fall within a particular age bin. The age bin, 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 paper for t_M and t_R when defining these age bins. This results in a 9×9 array of reference maps (see Fig. 4). See Section 3.2 for more details.

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

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.

3.3 Generation of the 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) contains no information about the expected emission spectrum. We therefore use SLUG2 (da Silva et al. 2012, 2014; Krumholz et al. 2015), 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 1.

With the SLUG2 model, we predict the expected rest-frame emission spectrum for every star particle within the simulation.[2] 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. SLUG2 then converts the full combined emission spectrum into a single luminosity value for each of the SFR tracers in Table 1 using filter response curves. These single luminosity values are what we assign to our star particles when we produce our synthetic rest-frame emission maps. We use the same smoothing procedure as we described in Section 3.2. 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, will have an age spread. An example of a synthetic H$\alpha-$ map is shown in Fig. 2.

The adopted UV response filters are all included by default in SLUG2 (see Krumholz et al. 2015 for more details). The Hα SFR tracers, however, require different steps. For H$\alpha-$ we use the hydrogen-ionizing photon emission, $Q(\mathrm{H}^0)$, directly[3] and for H$\alpha+W$ we define the narrow band filter, $\mathcal{F}_{\mathrm{H}\alpha+W}$, as

$$\mathcal{F}_{\mathrm{H}\alpha+W} = \begin{cases} 1 & 6562 - \dfrac{W}{2}\ \text{Å} \leq \lambda \leq 6562 + \dfrac{W}{2}\ \text{Å} \\ 0 & \text{Otherwise} \end{cases}. \quad (3)$$

The emission spectrum produced by SLUG2, includes the Hα emission line but does not calculate the underlying absorption feature from the stellar continuum. In Appendix A we use STARBURST99 simulations to investigate when the absorption can no longer be neglected. We find that for the time-scales we are considering the absorption is negligible.

For the analysis in Section 4, 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-stochastically[4] (i.e. we use a well-sampled IMF) and no foreground extinction is applied. The surrounding material has a hydrogen number density of $10^2\,\mathrm{cm}^{-3}$. We assume that only 73 per cent of the ionizing photons are reprocessed into nebular emission, which is consistent with the estimate from McKee & Williams (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 ionized emission map (e.g. Hα), 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

$$t_2 = t_o + t_i, \quad (4)$$

where t_o is the duration of the second phase that overlaps with the first, and t_i the duration that is independent. The characteristic time-scales we define in this paper 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 time-scale. To facilitate this, the HEISENBERG code enables the user to specify if the reference time-scale includes or excludes this overlap phase (see section 3.2.1 Kruijssen et al. 2018).

---

[2]We note that the age binning described in Section 3.2 is not used for the emission maps.

[3]A true Hα luminosity can be calculated from $Q(\mathrm{H}^0)$ using da 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.

[4]In Section 6, we will use the stochastic IMF sampling mode of SLUG2 to investigate its effect on the inferred SFR tracer time-scales.

Hα−	$t_R:1$	$t_R:3$	$t_R:5$	$t_R:7$	$t_R:10$	$t_R:15$	$t_R:20$	$t_R:25$	$t_R:30$
$t_M:1$	$2.88^{+0.52}_{-0.34}$	$2.98^{+0.20}_{-0.24}$	$3.09^{+0.27}_{-0.19}$	$3.78^{+0.34}_{-0.43}$	$3.41^{+0.42}_{-0.19}$	$3.52^{+0.18}_{-0.18}$	$4.22^{+0.12}_{-0.23}$	$4.71^{+0.16}_{-0.24}$	$4.79^{+0.14}_{-0.30}$
$t_M:3$	$2.83^{+0.41}_{-0.26}$	$3.39^{+0.34}_{-0.29}$	$4.19^{+0.22}_{-0.21}$	$4.50^{+0.24}_{-0.25}$	$3.82^{+0.25}_{-0.19}$	$4.20^{+0.24}_{-0.23}$	$4.70^{+0.29}_{-0.23}$	$5.76^{+0.33}_{-0.27}$	$6.29^{+0.37}_{-0.33}$
$t_M:5$	$2.58^{+1.09}_{-0.26}$	$4.74^{+1.08}_{-0.36}$	$6.09^{+1.46}_{-0.76}$	$4.81^{+0.72}_{-0.36}$	$4.79^{+0.83}_{-0.36}$	$5.40^{+1.00}_{-0.59}$	$4.66^{+0.78}_{-0.26}$	$7.22^{+1.13}_{-0.83}$	$7.42^{+1.02}_{-0.81}$
$t_M:7$	$13.41^{+7.69}_{-3.34}$	$10.50^{+3.42}_{-1.78}$	$4.74^{+0.96}_{-0.71}$	$4.24^{+0.78}_{-0.47}$	$6.30^{+0.89}_{-0.79}$	$6.82^{+0.83}_{-0.86}$	$7.30^{+1.07}_{-0.86}$	$6.87^{+1.07}_{-0.76}$	$6.12^{+1.08}_{-0.67}$
$t_M:10$	$1.95^{+0.33}_{-0.30}$	$2.35^{+0.49}_{-0.26}$	$4.19^{+0.55}_{-0.60}$	$4.69^{+0.40}_{-0.67}$	$5.56^{+0.47}_{-0.72}$	$5.99^{+0.69}_{-0.73}$	$6.58^{+0.81}_{-0.82}$	$5.98^{+0.91}_{-0.71}$	$6.32^{+1.02}_{-0.76}$
$t_M:15$	$2.55^{+0.38}_{-0.34}$	$3.23^{+0.39}_{-0.35}$	$4.48^{+0.49}_{-0.49}$	$4.97^{+0.50}_{-0.51}$	$6.17^{+0.61}_{-0.74}$	$9.07^{+0.83}_{-1.11}$	$7.26^{+0.76}_{-0.86}$	$7.84^{+0.86}_{-0.92}$	$9.68^{+1.06}_{-1.22}$
$t_M:20$	$3.62^{+0.77}_{-0.62}$	$4.25^{+0.63}_{-0.54}$	$5.66^{+0.85}_{-0.77}$	$6.01^{+0.74}_{-0.80}$	$7.49^{+0.92}_{-0.99}$	$6.32^{+0.96}_{-0.89}$	$7.01^{+1.07}_{-0.92}$	$7.29^{+1.21}_{-1.03}$	$5.53^{+1.03}_{-0.59}$
$t_M:25$	$3.88^{+1.27}_{-0.70}$	$6.21^{+0.93}_{-1.10}$	$6.44^{+0.68}_{-1.00}$	$7.14^{+1.16}_{-1.03}$	$6.52^{+0.85}_{-0.82}$	$7.35^{+0.83}_{-0.90}$	$8.13^{+0.98}_{-1.05}$	$6.58^{+1.09}_{-0.77}$	$5.60^{+0.92}_{-0.61}$
$t_M:30$	$4.01^{+1.82}_{-0.69}$	$5.18^{+1.27}_{-0.77}$	$5.30^{+0.94}_{-0.75}$	$5.11^{+0.66}_{-0.67}$	$6.33^{+0.88}_{-0.74}$	$6.59^{+0.83}_{-0.83}$	$7.35^{+1.00}_{-0.85}$	$5.78^{+0.77}_{-0.65}$	$5.41^{+0.79}_{-0.62}$

Weighting — 0.14 / 0.12 / 0.10 / 0.08 / 0.06 / 0.04 / 0.02

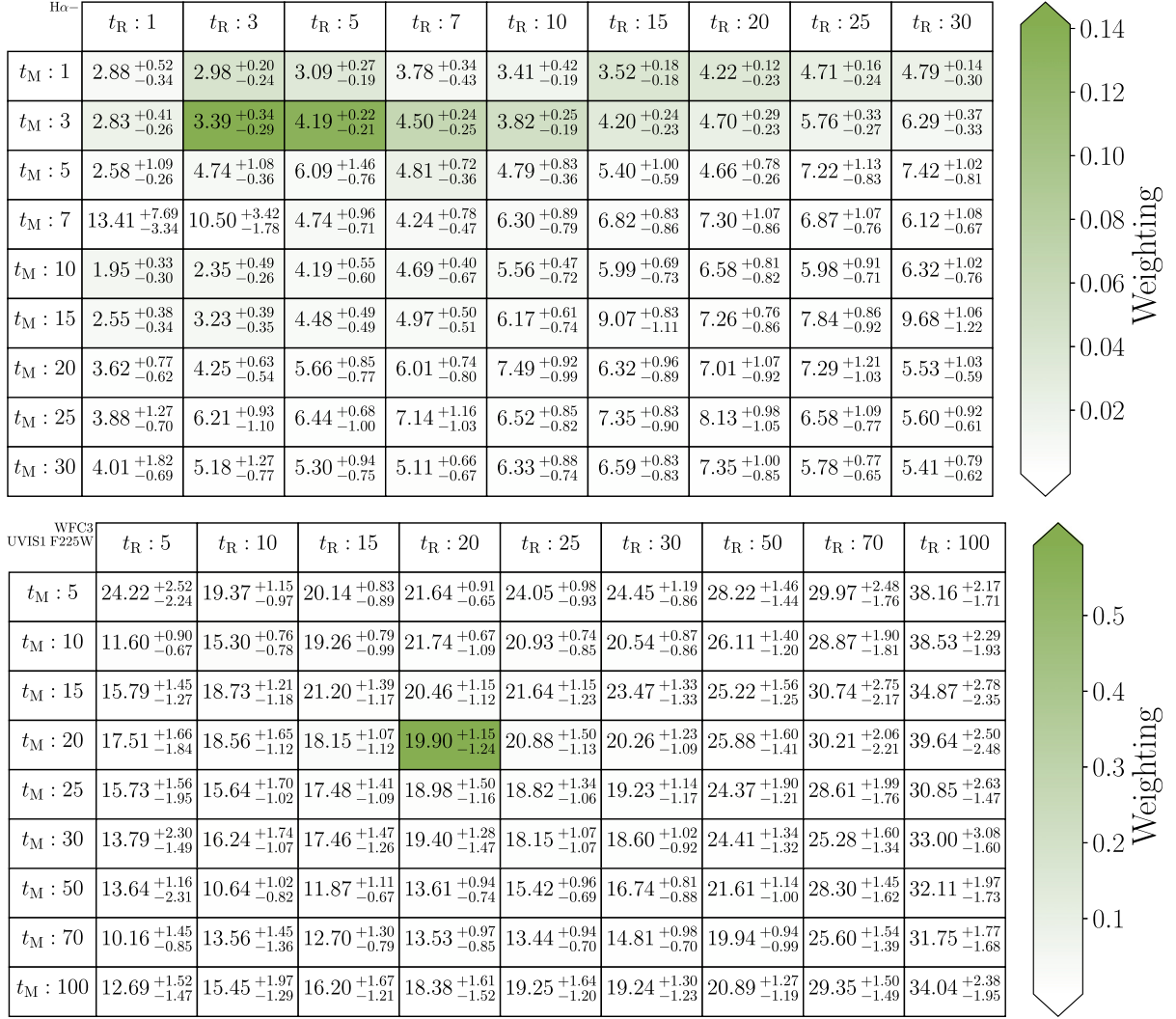| WFC3 UVIS1 F225W	$t_R:5$	$t_R:10$	$t_R:15$	$t_R:20$	$t_R:25$	$t_R:30$	$t_R:50$	$t_R:70$	$t_R:100$
$t_M:5$	$24.22^{+2.52}_{-2.24}$	$19.37^{+1.15}_{-0.97}$	$20.14^{+0.83}_{-0.89}$	$21.64^{+0.91}_{-0.65}$	$24.05^{+0.98}_{-0.93}$	$24.45^{+1.19}_{-0.86}$	$28.22^{+1.46}_{-1.44}$	$29.97^{+2.48}_{-1.76}$	$38.16^{+2.17}_{-1.71}$
$t_M:10$	$11.60^{+0.90}_{-0.67}$	$15.30^{+0.76}_{-0.78}$	$19.26^{+0.79}_{-0.99}$	$21.74^{+0.67}_{-1.09}$	$20.93^{+0.74}_{-0.85}$	$20.54^{+0.87}_{-0.86}$	$26.11^{+1.40}_{-1.20}$	$28.87^{+1.90}_{-1.81}$	$38.53^{+2.29}_{-1.93}$
$t_M:15$	$15.79^{+1.45}_{-1.27}$	$18.73^{+1.21}_{-1.18}$	$21.20^{+1.39}_{-1.17}$	$20.46^{+1.15}_{-1.12}$	$21.64^{+1.15}_{-1.23}$	$23.47^{+1.33}_{-1.33}$	$25.22^{+1.56}_{-1.25}$	$30.74^{+2.75}_{-2.17}$	$34.87^{+2.78}_{-2.35}$
$t_M:20$	$17.51^{+1.66}_{-1.84}$	$18.56^{+1.65}_{-1.12}$	$18.15^{+1.07}_{-1.12}$	$19.90^{+1.15}_{-1.24}$	$20.88^{+1.50}_{-1.13}$	$20.26^{+1.23}_{-1.09}$	$25.88^{+1.60}_{-1.41}$	$30.21^{+2.06}_{-2.21}$	$39.64^{+2.50}_{-2.48}$
$t_M:25$	$15.73^{+1.56}_{-1.95}$	$15.64^{+1.70}_{-1.02}$	$17.48^{+1.41}_{-1.09}$	$18.98^{+1.50}_{-1.16}$	$18.82^{+1.34}_{-1.06}$	$19.23^{+1.14}_{-1.17}$	$24.37^{+1.90}_{-1.21}$	$28.61^{+1.99}_{-1.76}$	$30.85^{+2.63}_{-1.47}$
$t_M:30$	$13.79^{+2.30}_{-1.49}$	$16.24^{+1.74}_{-1.07}$	$17.46^{+1.47}_{-1.26}$	$19.40^{+1.28}_{-1.47}$	$18.15^{+1.07}_{-1.07}$	$18.60^{+1.02}_{-0.92}$	$24.41^{+1.34}_{-1.32}$	$25.28^{+1.60}_{-1.34}$	$33.00^{+3.08}_{-1.60}$
$t_M:50$	$13.64^{+1.16}_{-2.31}$	$10.64^{+1.02}_{-0.82}$	$11.87^{+1.11}_{-0.67}$	$13.61^{+0.94}_{-0.74}$	$15.42^{+0.96}_{-0.69}$	$16.74^{+0.81}_{-0.88}$	$21.61^{+1.14}_{-1.00}$	$28.30^{+1.45}_{-1.62}$	$32.11^{+1.97}_{-1.73}$
$t_M:70$	$10.16^{+1.45}_{-0.85}$	$13.56^{+1.45}_{-1.36}$	$12.70^{+1.30}_{-0.79}$	$13.53^{+0.97}_{-0.85}$	$13.44^{+0.94}_{-0.70}$	$14.81^{+0.98}_{-0.70}$	$19.94^{+0.94}_{-0.99}$	$25.60^{+1.54}_{-1.39}$	$31.75^{+1.77}_{-1.68}$
$t_M:100$	$12.69^{+1.52}_{-1.47}$	$15.45^{+1.97}_{-1.29}$	$16.20^{+1.67}_{-1.21}$	$18.38^{+1.61}_{-1.52}$	$19.25^{+1.64}_{-1.20}$	$19.24^{+1.30}_{-1.23}$	$20.89^{+1.27}_{-1.19}$	$29.35^{+1.50}_{-1.49}$	$34.04^{+2.38}_{-1.95}$

Weighting — 0.5 / 0.4 / 0.3 / 0.2 / 0.1

Figure 4. Two examples showing the range of characteristic time-scale values (and associated uncertainties) as determined using HEISENBERG for different reference maps. *Top:* Hα emission excluding the continuum (Hα −). *Bottom:* UV emission (WFC3 UVIS *F225W*). The reference maps are characterized by the age bin 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 bin. The colour coding is based on the weighting, $\mathcal{W}$, used when calculating the weighted average. All values within the tables are given in Myr.

In addition, it is desirable to exclude extinction for two further reasons. First, 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 extincted 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 future work, we aim to consider extinction using galaxy simulations covering a range of gas surface densities (Haydon et al. 2020).

maps show the star particles in a chosen age interval. Since there is some freedom in choosing this interval, we measure the SFR tracer time-scale for a wide variety of age intervals rather than picking a single one. Appendix B demonstrates this improves the accuracy of the measurements relative to using a single age interval. Mathematically, we change the values of t_M and t_R, which define the age interval as $t_M \leq \text{Age} \leq t_M + t_R$. This approach generates an array of SFR tracer time-scales, spanned by t_M and t_R. We now first describe how we reduce these 'time-scale arrays' (see Fig. 4 for examples) into a single characteristic time-scale for each SFR tracer. To achieve this, we use the complete probability density function (PDF) of each measured time-scale, which is provided by HEISENBERG.

4 CHARACTERISTIC TIME-SCALES FOR A FULLY SAMPLED IMF AT SOLAR METALLICITY

We constrain the characteristic time-scales for several SFR tracers by applying the HEISENBERG code to the synthetic SFR tracer maps and reference maps described in Section 3. The reference

4.1 Mathematical procedure

In principle, HEISENBERG allows any pair of values to be used for generating the reference map, but the accuracy of the method is highest when the durations of both evolutionary phases ($t_{E,0}$ and

t_R) are similar (Kruijssen et al. 2018). In that case, the tuning fork diagram of Fig. 1 is symmetric, allowing the method to retrieve the underlying time-scales with a precision of better than 30 per cent. We additionally prefer numerical experiments with a minimal time offset between the reference map and the SFR tracer, to mimic the close evolutionary correspondence between molecular gas and (massive) star formation. We therefore also prefer solutions in which the temporal offset of the reference phase (t_M) is similar to the duration of the SFR tracer ($t_{E,0}$). Since we do not know the latter a priori, it is not possible to choose an optimal pair of t_R and t_M in advance.

In order to condense an array of SFR tracer time-scale measurements (e.g. Fig. 4) into a single characteristic time-scale, we must design a simple weighting scheme that accounts for the above behaviour of HEISENBERG. This scheme should favour solutions with $t_M \sim t_R \sim t_{E,0}$. Additionally, it should weigh each measurement by the inverse-square of its uncertainty, as is common when calculating the average across a sample of measurements.

First, we weight each measurement simply by the inverse of its geometric distance from t_R and t_M in logarithmic space:

$$\mathcal{W}_{ij}^{d} = \left\{ \left[\log_{10}\left(\frac{t_{ij}}{t_{M,i}} \right) \right]^2 + \left[\log_{10}\left(\frac{t_{ij}}{t_{R,j}} \right) \right]^2 \right\}^{-\frac{1}{2}}. \quad (5)$$

This is the simplest possible approach to performing a proximity-based weighting. It favours more strongly elements that satisfy the criteria we describe in equations (1) and (2) (i.e. the closer t_{ij} is to $t_{M,i}$ and $t_{R,j}$, the better).

Secondly, we weight each measurement by the inverse-square of its uncertainty, which accounts for asymmetric uncertainties by taking the mean of the lower and upper uncertainty:[5]

$$\mathcal{W}_{ij}^{u} = \left(\frac{\sigma_{ij}^{-} + \sigma_{ij}^{+}}{2} \right)^{-2}. \quad (6)$$

The weights $\mathcal{W}^d$ and $\mathcal{W}^u$ are then combined and normalized as

$$\mathcal{W}_{ij} = \frac{\mathcal{W}_{ij}^{d} \mathcal{W}_{ij}^{u}}{\sum_{ij} \mathcal{W}_{ij}^{d} \mathcal{W}_{ij}^{u}}. \quad (7)$$

The above weighting scheme appropriately combines all measurements in the time-scale array into a single SFR tracer time-scale. However, it does not propagate the individual measurement uncertainties into the final measurement. To accomplish this, we adopt a simple Monte Carlo approach. We produce 10^6 realizations of the time-scale array, where the value of each element of each realization of the time-scale array has been randomly sampled from its associated PDF. For each of the 10^6 realizations of the time-scale array we calculate the weighted mean according to equations (5)–(7). This process results in 10^6 characteristic time-scales, from which we take the median to define the characteristic time-scale and the 16th and 84th percentiles to define the uncertainties. These uncertainties contain both the uncertainties on the individual measurements and the systematic uncertainty associated with choosing a combination of t_M and t_R.

With the above procedure, we condense the array of time-scales into a single number that most strongly weighs the values that are the most accurate (based on Kruijssen et al. 2018) and those that have the smallest measurement uncertainties. In practice, we find

that the typical standard deviation of all SFR tracer time-scales is ∼0.15 dex, over a dynamical range of 1.5 dex in t_M and t_R. This demonstrates that the inferred time-scales are not extremely sensitive to the choice of reference map, but the full array of reference maps does allow us to optimize the accuracy of the SFR tracer time-scales. In Appendix B, we demonstrate that our weighting scheme indeed performs better than the approach of choosing a single measurement by simply minimizing the difference in geometric distance between $t_{E,0}$, t_M, and t_R (or maximizing $\mathcal{W}_{ij}^{d}$).

4.2 Measured SFR tracer time-scales

When applying HEISENBERG to the pairs of reference and SFR tracer maps, we use the default input parameters specified in tables 1 and 2 of Kruijssen et al. (2018). The only exceptions are as follows. We set `tstar_incl = 1`, to indicate that the reference time-scale (i.e. the width of the age bin) also includes the overlapping phase.[6] 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 $l_{ap,\,min} = 25$ pc and a number of $N_{ap} = 17$ apertures, to produce 17 logarithmically spaced aperture diameters from 25 to 6400 pc.

In Fig. 4, we present the time-scale arrays obtained for H α − and WFC3 UVIS $F225W$ SFR tracers. These time-scale arrays only serve as examples, since the elements show the output of HEISENBERG and are not from the 10^6 Monte Carlo realizations. Fig. 5 shows the PDFs of the defined characteristic time-scale for H α − and WFC3 UVIS $F225W$, as obtained by applying the weighting scheme described in Section 4.1 to 10^6 different Monte Carlo realizations of Fig. 4.

Reassuringly, the best-fitting numbers in Fig. 4 show relatively limited cell-to-cell variation, which means that stochastic SFR variations in the simulation (which would affect different age bins differently) do not strongly affect our results. This is to be expected – Kruijssen et al. (2018) show that the SFR varies by less than 30 per cent over the typical age intervals considered. This may not always hold in observational applications of our methodology, and should therefore not be applied to systems with strongly varying SFRs, such as galaxy mergers or other starbursts (also see sections 4.2.4.3 and 4.4 of Kruijssen et al. 2018).

Table 3 lists the characteristic time-scales and associated uncertainties obtained for each of the different SFR tracers. The complete set of SFR tracer time-scales spans a range of 4–34 Myr. We use these measurements to define age bins that we will use to generate reference maps when investigating the impact of metallicity and IMF sampling in later sections. These are listed in the final column of the table and are calculated as $t_{E,0} \leq$ Age $\leq 2t_{E,0}$.

Fig. 6 shows the UV-based SFR tracer time-scales as a function of the response curve-weighted mean wavelength, $\bar{\lambda}_w$. The figure shows both are closely correlated, such that similar response-weighted mean wavelengths give similar characteristic time-scales. We perform a weighted least-squares minimization to obtain a relation between $\bar{\lambda}_w$

---

[5]Using the average of the lower and upper uncertainty is not technically correct; however, the methods as suggested by Barlow (2003) would have little impact on the final result and so are neglected.

[6]At first glance, this may seem to contradict the discussion in Section 3.3, but this is not the case. In Section 3.3, we explain that the characteristic time-scales of the SFR tracers we define do not include the overlap phase; and so, when using the characteristic time-scales we present here, one should use `tstar_incl = 0`. The analysis we perform to define the characteristic time-scales, uses a reference map produced from star particles in a specific age bin. The width of this age bin is used as the reference time-scale and is the total duration of that phase: this includes any overlap.
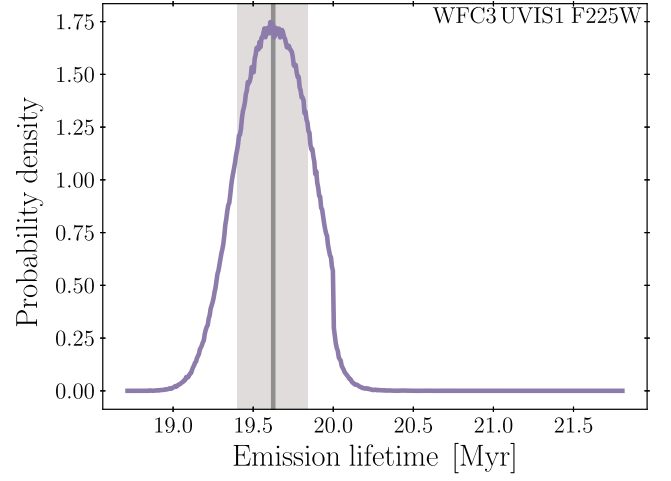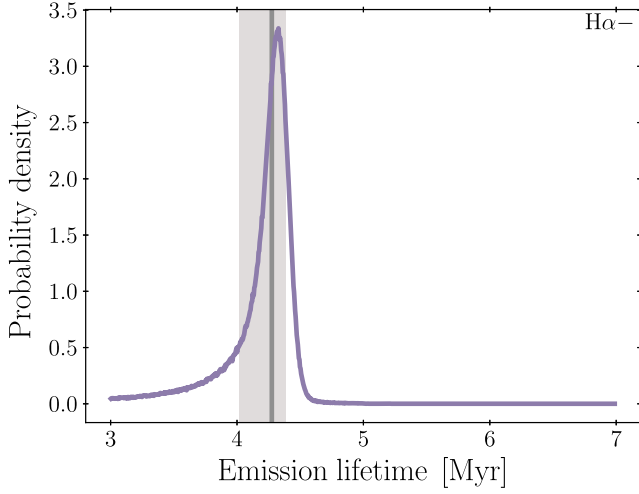
Figure 5. Two examples showing the PDFs associated with the defined characteristic time-scales. *Left:* Hα emission excluding the continuum (Hα −). *Right:* UV emission (WFC3 UVIS *F225W*). The vertical line shows the selected time-scale (the median of the distribution); the shaded region, the uncertainty defined by the 16th and 84th percentiles.

Table 3. The characteristic time-scales, $t_{E,0}$, obtained for the different SFR tracers (see Table 1 for details) and the corresponding age bins ($t_{E,0} \leq$ Age $\leq 2t_{E,0}$) for producing reference maps in later sections of this paper. These results are for a well-sampled IMF at solar metallicity. The filter order is in increasing filter width (*W*) for Hα + and increasing response-weighted mean wavelength ($\bar{\lambda}_w$) for UV. This table is an extract of Table D1, which includes the characteristic time-scales and age bins for different stellar metallicities ($Z/Z_\odot = 0.05$–2).

	$t_{E,0}$ [Myr]	Age bin [Myr]
Hα −	$4.3^{+0.1}_{-0.3}$	4.3–8.6
Hα + 10 Å	$5.6^{+0.2}_{-0.1}$	5.6–11.1
Hα + 20 Å	$7.3^{+0.4}_{-0.2}$	7.3–14.6
Hα + 40 Å	$9.3^{+0.2}_{-0.3}$	9.3–18.6
Hα + 80 Å	$10.7^{+0.2}_{-0.2}$	10.7–21.4
Hα + 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

and the UV characteristic time-scale, $t_{E,0}^{UV}$:

$$t_{E,0}^{UV} \text{ [Myr]} = \left(3.00^{+0.29}_{-0.31}\right) \left(\frac{\bar{\lambda}_w}{225 \text{ nm}}\right)^{\left(4.34^{+0.24}_{-0.20}\right)} + \left(16.42^{+0.36}_{-0.31}\right). \quad (8)$$

The uncertainties on the parameters are calculated using a Monte Carlo approach. This analytical expression enables finding the characteristic emission time-scales for UV filters other than the specific ones that we have considered here.
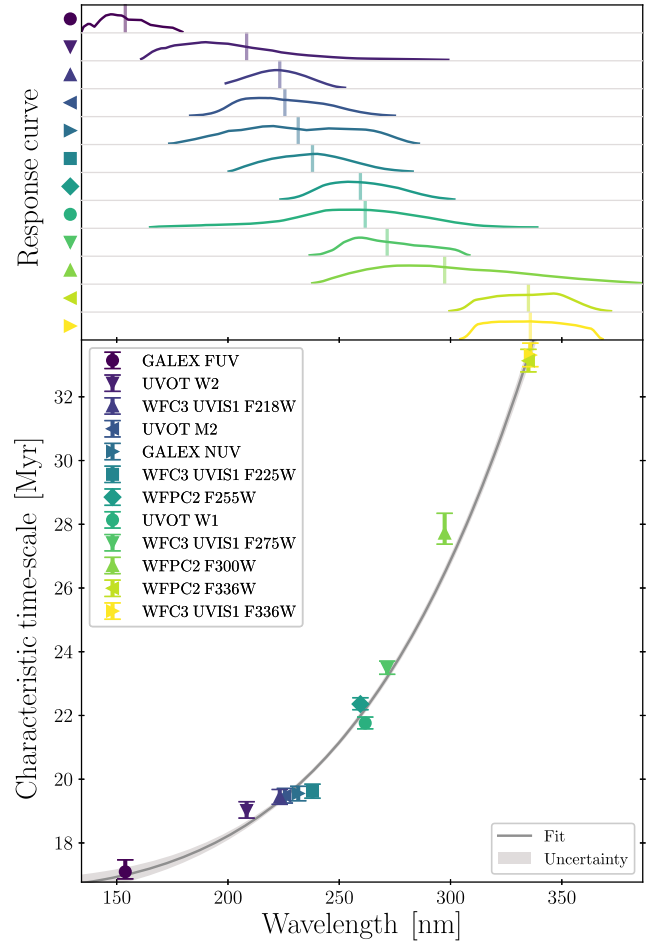
Figure 6. *Top:* The normalized response curves of the UV filters considered in this paper (also see Table 1a). The vertical lines indicate the response-weighted mean wavelengths, $\bar{\lambda}_w$. *Bottom:* Characteristic emission time-scales for UV filters as a function of response-weighted mean wavelength. The grey curve shows the fit described in equation (8) and the shaded regions indicate the associated uncertainty.
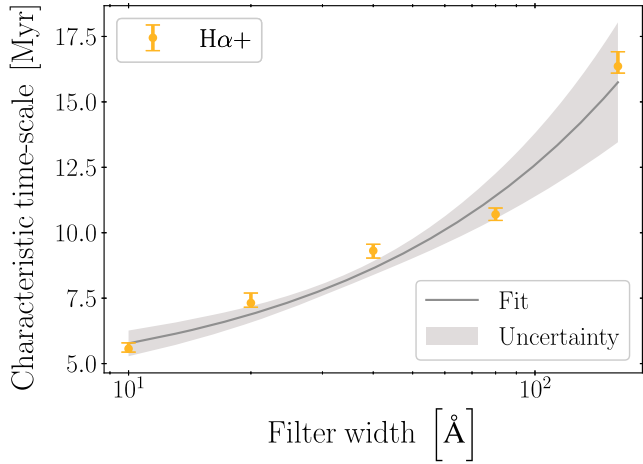
Figure 7. Characteristic emission time-scales for H α + filters as a function of the filter width. The grey curve shows the fit described in equation (9) and the shaded region indicates the associated uncertainty.

Similarly, we derive a relation between the H α + characteristic time-scale and filter width:

$$t_{E,0}^{H\alpha+}[\text{Myr}] = \left(4.8^{+1.3}_{-1.3}\right)\left(\frac{W}{40\,\text{Å}}\right)^{\left(0.65^{+0.20}_{-0.13}\right)} + \left(3.8^{+1.1}_{-1.1}\right). \quad (9)$$

This relation is compared to the measurements in Fig. 7. Note that the increase in characteristic time-scale with filter width is not due to a change in the H α emission, but results from a change in flux from the long-lived continuum emission.

The characteristic time-scales that we recover (4.3–16.4 Myr for H α; 17.1–33.3 Myr for UV) fall within the ranges often quoted in the literature (2–10 and 10–50 Myr, respectively, Kennicutt & Evans 2012; Leroy et al. 2012). In part, the large variation in literature values reflects the broad range of criteria used for defining the characteristic time-scale of an SFR tracer. With the approach taken in this paper, we have remedied this problem for future observational applications of the KL14 principle by adopting a specific definition of the SFR tracer time-scale.

The heterogeneous situation in the previous literature is nicely illustrated by Leroy et al. (2012), who present a table of characteristic time-scales for H α and FUV (at 150 nm). Multiple time-scales are listed for each SFR tracer; these time-scales are defined by the duration required to reach a given percentage (50 or 95 per cent) of the cumulative emission, or of the emission intensity at 1 Myr. These choices are arbitrary, but it is reasonable to ask whether any single percentage of the 1 Myr intensity or cumulative emission can be defined that would correspond to our measured SFR tracer time-scales. To verify this, we take the emission evolution from Leroy et al. (2012, fig. 1) and find which percentages correspond to the characteristic emission time-scales we determine for H α − and *GALEX* FUV. We list these percentage limits in Table 4. We find that no single percentage limit corresponds to the measured time-scales.[7] As there is no consistent limit, the characteristic time-scale for each SFR tracer must be determined individually. This further validates the approach taken in this study.

[7]This 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 5 for more details). There is also no consistent percentage for a single tracer across the metallicity range.

Table 4. Percentages of the emission intensity relative to its instantaneous value at 1 Myr and of its cumulative value over 100 Myr, evaluated at the characteristic time-scales of the SFR tracers presented in Table 3, based on fig. 1 of Leroy et al. (2012).

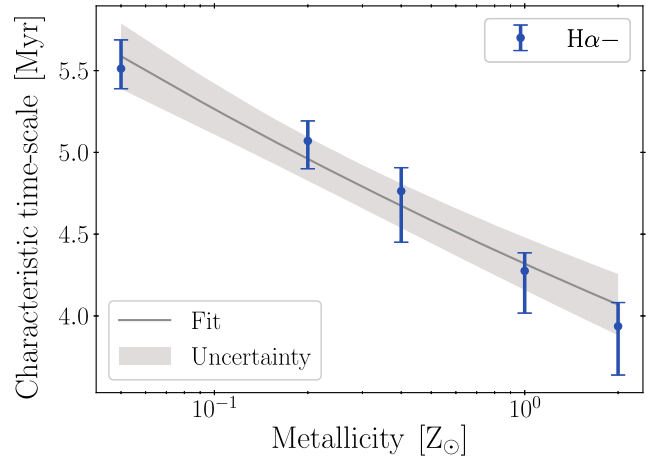|	H α − [per cent]	FUV[a] [per cent]
Per cent of intensity at 1 Myr	$19.4^{+2.8}_{-0.7}$	$8.6^{+0.1}_{-0.2}$
Per cent of cumulative emission	$92.4^{+0.6}_{-1.5}$	$76.5^{+0.3}_{-0.2}$

[a]GALEX FUV.

Figure 8. The relation between metallicity and characteristic time-scale for a well-sampled IMF for H α − filters. The grey curve gives the fits described by equation (10). The shaded region indicates the associated uncertainty.

In summary, our SFR tracer time-scales fall in the range of commonly reported literature values. These do not correspond to any fixed percentage of the initial or cumulative emission in each tracer. For this reason, each SFR tracer time-scale must be determined individually using the presented method. We provide analytic functions (see equations 8 and 9) relating the characteristic emission time-scales for UV and H α + filters to their filter properties, allowing our results to be extended to any UV or H α + filter.

5 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 characteristic emission time-scales of SFR tracers. In order to facilitate observational applications of the KL14 principle to the broadest possible range of galaxies, we therefore quantify how the SFR tracer time-scales depend on metallicity. In this section, we repeat the experiments performed in Section 4 but this time we produce synthetic SFR tracer emission maps using evolutionary tracks of metallicities $Z/Z_\odot = 0.05, 0.20, 0.40, 2.00$ (Schaller et al. 1992; Charbonnel et al. 1993; Schaerer et al. 1993a, b).

In Appendix D, we list the resulting characteristic time-scales for a well-sampled IMF for all metallicities (also including the solar metallicity results from Table 3) and the age bins we select for producing reference maps. We show the $Z − t_{E,0}$ relation in Fig. 8 for H α −, in Fig. 9 for H α + filters, and in Fig. 10 for the UV filters. For all tracers, we find that the characteristic time-scale decreases
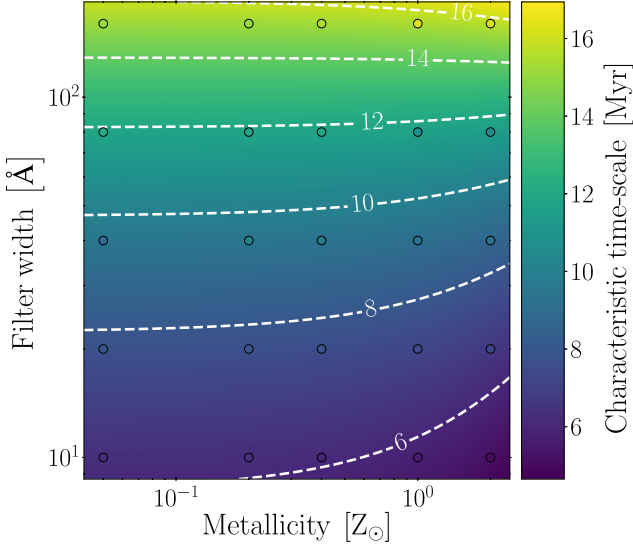
Figure 9. The surface described by equation (11), which relates the metallicity and filter width, W, of an $H\alpha +$ filter to the associated characteristic time-scale for a well sampled IMF. The data points show the measurements coloured using the same colour bar. The surface fits best when it matches the colour of the data points.
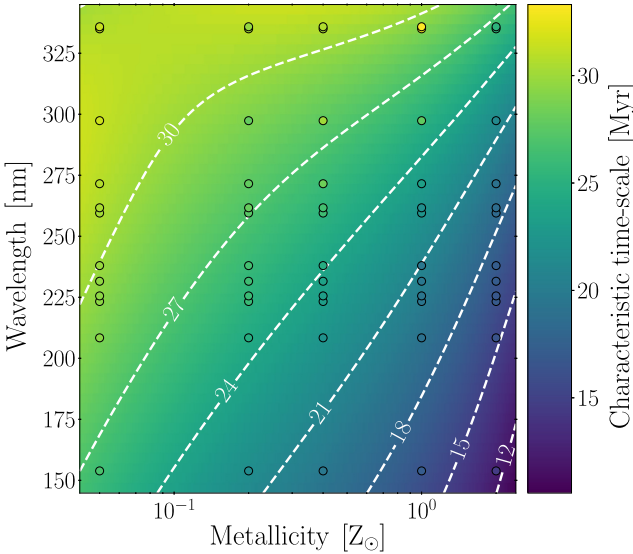

Figure 10. The surface described by equation (12), which relates the metallicity and response-weighted mean wavelength, $\overline{\lambda}_w$, of a UV SFR tracer to the associated characteristic time-scale for a well sampled IMF. The data points show the measurements coloured using the same colour bar. The surface fits best when it matches the colour of the data points.

with metallicity. We also include empirical fits described by

$$t_{E,0}^{H\alpha-}[\text{Myr}] = \left(4.32_{-0.23}^{+0.09}\right) \left(\frac{Z}{Z_\odot}\right)^{\left(-0.086_{-0.023}^{+0.010}\right)}, \quad (10)$$

for $H\alpha -$,

$$t_{E,0}^{H\alpha+}[\text{Myr}] = \left(8.98_{-0.50}^{+0.40}\right) W_0^{\left(0.265_{-0.051}^{+0.028}\right)} + \left(0.23_{-0.11}^{+0.15}\right) Z_0 W_0$$
$$- \left(0.66_{-0.19}^{+0.12}\right) Z_0 + \left(0.55_{-0.29}^{+0.46}\right) W_0, \quad (11)$$

for $H\alpha +$, and

$$t_{E,0}^{UV}[\text{Myr}] = -\left(0.40_{-0.16}^{+0.11}\right) Z_1\lambda_1 + \left(4.5_{-0.9}^{+1.3}\right) Z_1 + \left(0.70_{-0.18}^{+0.26}\right) \lambda_1$$
$$- \left(3.11_{-0.13}^{+0.14}\right) Z_0 + \left(10.98_{-0.48}^{+0.46}\right) \lambda_0 + \left(7.6_{-1.6}^{+1.2}\right), \quad (12)$$

for the UV filters, where

$$Z_0 \equiv \frac{Z}{Z_\odot}; \quad \lambda_0 \equiv \frac{\overline{\lambda}_w}{225\,\text{nm}}; \quad W_0 \equiv \frac{W}{40\,\text{Å}};$$
$$Z_1 \equiv Z_0^{\left(-0.313_{-0.048}^{+0.051}\right)}; \quad \lambda_1 \equiv \lambda_0^{\left(6.52_{-0.71}^{+0.73}\right)}. \quad (13)$$

As before, we determine the free parameters using a weighted least-squares minimization and the uncertainties through Monte Carlo methods. With these relations, it is straightforward to recover the characteristic time-scale for any combination of metallicity and filter properties, without needing to repeat the analysis of this paper.

Fig. 8 shows that the characteristic time-scales of $H\alpha -$ change by less than 2 Myr over the metallicity range $[0.05\,Z_\odot, 2\,Z_\odot]$. The ranges of characteristic time-scales (3.9–5.5 Myr for $H\alpha -$) fall within the range of literature values (1.7–10 Myr, Kennicutt & Evans 2012; Leroy et al. 2012).

In Section 4, we describe a curve which relates the filter width, W, to the characteristic time-scale of $H\alpha +$ filters, $t_{E,0}^{H\alpha+}$, at solar metallicity. Equation (11) now extends this relation to include different metallicities, producing a surface in $(t_{E,0}^{H\alpha+}, W, Z)$ space. As mentioned in Section 4, the measured $H\alpha +$ time-scales reside at the higher end of (and partially exceeds) the literature range of $H\alpha$ time-scales, because wider filters include more of the long-lived continuum emission.

Analogously to $t_{E,0}^{H\alpha+}$, we can extend the relation given for $t_{E,0}^{UV}$ as a function of response-weighted mean wavelength (equation 8) to also include metallicity. We obtain a good fit, with the strongest deviations arising at long ($\lambda_w > 290$ nm) wavelengths. For UV filters at these wavelengths, we recommend interpolating the data points (provided in Appendix D) rather than adopting equation (12). The range of characteristic time-scales found for the UV filters (14.5–33.3 Myr) again fall within the range quoted in literature (10–100 Myr, Kennicutt & Evans 2012; Leroy et al. 2012), but tend 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 time-scales are naturally biased to the ages of regions from which most UV photons emerge.

In Appendix C, we additionally provide figures showing the one-dimensional projections of the data and fits. These show slices across the two-dimensional distributions shown in Figs 9 and 10 for the $H\alpha +$ and UV filters, respectively. We include a projection for each metallicity and each equivalent width or response-weighted mean wavelength and use the fits given in equations (11) and (12). These additional figures enable a direct assessment of how well the fits describe the characteristic time-scales at a specific metallicity or wavelength.

In summary, we see that the characteristic time-scales decrease with increasing metallicity. Observational applications of HEISENBERG should therefore use an SFR tracer time-scale appropriate for the metallicity of the observed region. We define empirical relations between the SFR tracer time-scale and the metallicity (for $H\alpha -$, equation 10) and the filter properties (for $H\alpha +$ and UV filters, equations 11 and 12). For $H\alpha +$ and UV SFR tracers, these relations enable the definition of time-scales even for filters that are not explicitly considered here.

6 THE EFFECTS OF IMF SAMPLING

In the previous sections, we determine the characteristic time-scales 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 stochastic IMF sampling on the characteristic time-scales of the SFR tracers, in particular for low-mass star-forming regions.

We describe in Section 2 how the abundance of regions in each input map reflects the duration associated with that map. If the IMF is not well-sampled, the SFR tracer flux is reduced. In the extreme case, a region might not be able to form stars of sufficient mass to produce the SFR tracer emission at all, and would thus be invisible in that tracer. These effects are particularly important for the $H\alpha \pm$ 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 IMF becomes less well-sampled, the effective characteristic time-scales of the various tracers will decrease, most strongly affecting $H\alpha$. In this section, we show how incomplete IMF sampling affects the inferred SFR tracer time-scales.

6.1 Method for finding the characteristic time-scales for a stochastically sampled IMF

We adapt the method we present in Section 3 to investigate the effects of a stochastically sampled IMF. In Appendix E, we derive how we expect the characteristic time-scales to change as a result of incomplete IMF sampling using a purely analytical approach. Specifically, we consider the probability of finding stars within the star-forming region that are sufficiently massive to generate the desired SFR tracer emission. In this section, we test experimentally if we recover the same behaviour as described there.

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 bins specified in Table 3. However, for the SFR tracer emission maps, we first scale the star particle masses by some mass scaling factor, F_m, before SLUG2 – now using it stochastic IMF sampling module – predicts the expected emission. The values of F_m range from 0.01 to 100, where a lower mass scaling factor means the IMF will be less well sampled. We then use HEISENBERG to determine the characteristic time-scale, as in Section 4. The characteristic time-scale we associate to each mass scaling factor is the average of three characteristic time-scales determined from three independently generated stochastic realizations of the synthetic emission maps. This accounts for the spread in time-scales that results from the stochastic nature of the synthetic emission maps.

We aim to determine the relative change of the SFR tracer time-scale as a function of IMF sampling. To express the latter, we define a characteristic, average star-forming region mass, $\overline{M}_r$, as

$$\overline{M}_r = \Sigma_{SFR} \times \tau \times \pi \left(\frac{\lambda}{2}\right)^2, \tag{14}$$

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

$$\Sigma_{SFR} = \frac{\sum m_i}{t_{E,0} \pi r^2} \times F_m, \tag{15}$$

where $\sum m_i$ is the total mass of all the star particles that fall within the age bin appropriate for the filter, i.e. $0 \leq \text{Age} \leq t_{E,0}$ (see Appendix D for the values of $t_{E,0}$), which is then scaled by the mass scaling factor F_m, $t_{E,0}$ is the width of that age bin, and r is the radius of the galaxy being studied (for our simulated galaxy $r = 10$ kpc, as determined from a visual inspection of the synthetic emission maps).

In equation (15), 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 (15) should account for a non-uniform spatial distribution of star-forming regions across the galaxy by including a factor $\mathcal{E}_{star,glob}$, which indicates the ratio of the mass surface density[8] on a size scale of λ to its area average across the map (see section 3.2.9 Kruijssen et al. 2018 for more details).

By introducing a 'mass scaling factor', F_m, we are able to test experimentally how the characteristic time-scale of different SFR tracers change as a smooth function of IMF sampling. We will use these experimental results to see if we observe the behaviour predicted in Appendix E.

6.2 SFR tracer time-scales for a stochastically sampled IMF

We now present the results of our experiments to test how the characteristic time-scales of $H\alpha$ and several UV SFR tracers change as a result of incomplete IMF sampling. In Fig. 11, we present the solar-metallicity results for $H\alpha$ − and WFC3 UVIS $F225W$ as examples of how the characteristic time-scales change as a function of the average mass of an independent star-forming region, $\overline{M}_r$. The quantity $\overline{M}_r$ characterizes (chiefly through Σ_{SFR}, see equation 14) how well the IMF is sampled: lower values of $\overline{M}_r$ result in a more stochastically sampled IMF. Each data point[9] in the two left-hand panels of Fig. 11 corresponds to a different mass scaling factor, F_m. The quantity shown on the vertical axis, $t_E/t_{E,0}$, is the factor by which the measured characteristic time-scale is reduced relative to the time-scale obtained for a well-sampled IMF (as listed in Appendix D), as a result of incomplete IMF sampling at small region masses or low SFR surface densities.

We describe the relation between reduction factor, $t_E/t_{E,0}$, and $\overline{M}_r$ through the probability $P(N \geq N_{min})$ as derived in Appendix E. The purple curves in left-hand panels of Fig. 11 indicate the best-fitting form of this function, obtained by varying the minimum stellar mass contributing to the SFR tracer emission (M_{min}) and the number of such stars required (N_{min}). We constrain the values for these two free parameters using a brute-force approach: we calculate the value of χ_{red}^2 (accounting for the uncertainties on both the abscissa and the ordinate, see Orear 1982) for a range of $N_{min} = 1$–4 and $M_{min} = 0$–$120 \, M_\odot$, and use the minimum χ_{red}^2 to define the best-fitting parameter values. The right-hand panels of Fig. 11 show the dependence of χ_{red}^2 on N_{min} and M_{min} for the two example filters. The best fits are always obtained for $N_{min} = 1$.

---

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

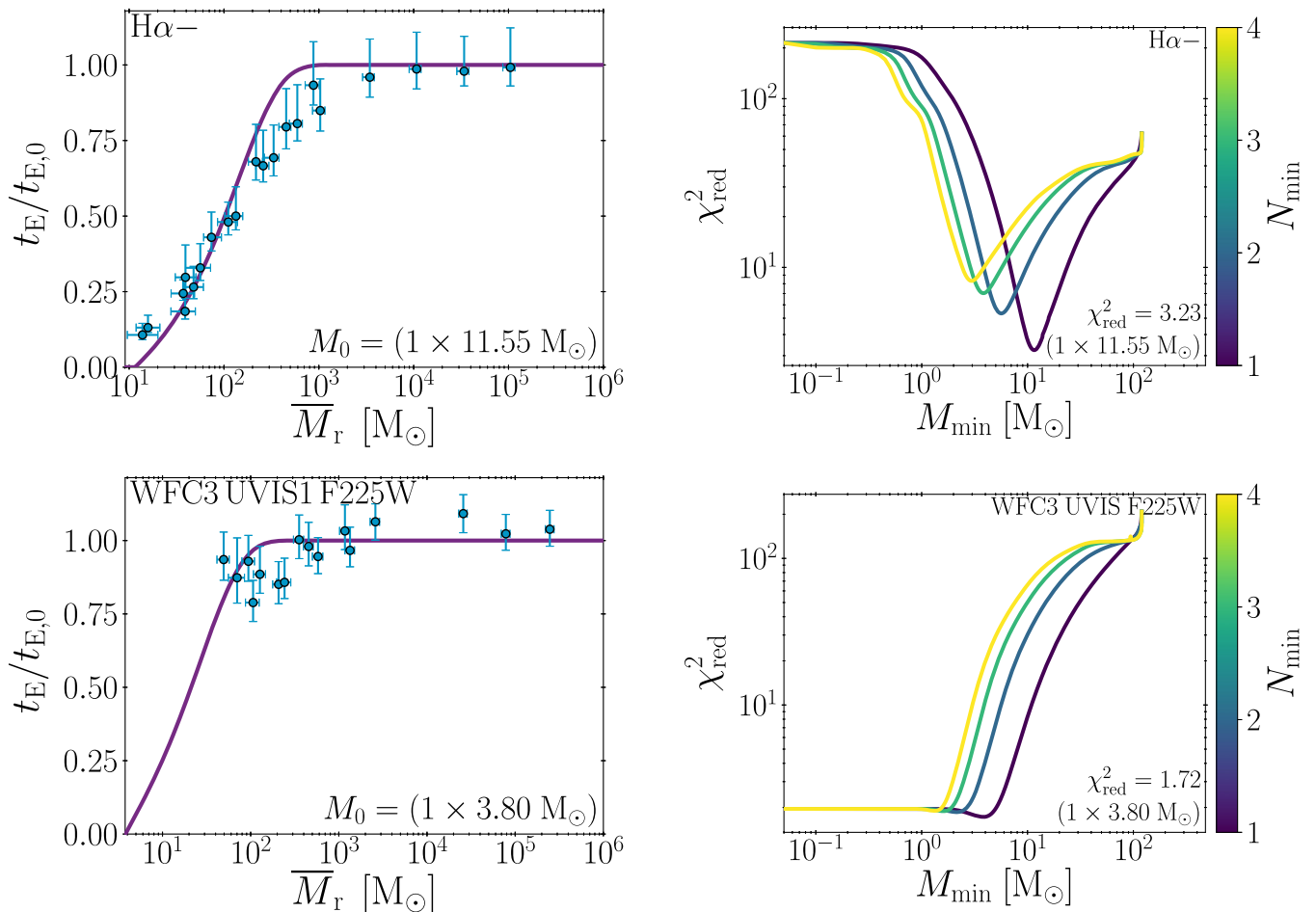[9]For details on the error calculation on $\overline{M}_r$, see Appendix F.

Figure 11. *Top row:* H α emission excluding the continuum (H α −). *Bottom row:* UV emission (WFC3 UVIS *F225W*). *Left column:* Change of the characteristic time-scale of the SFR tracer, relative to the characteristic time-scale 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 stochastically sampled IMFs at solar metallicity. For comparison, the purple curve shows the best-fitting analytical model from Appendix E. At low region masses, the characteristic time-scales decrease due to the incomplete sampling of the IMF. *Right column:* Change of $\chi^2_{\rm red}$ with minimum stellar mass, $M_{\rm min}$, and the minimum number of stars of that mass, $N_{\rm min}$. The minimum $\chi^2_{\rm red}$ found is indicated in the bottom right with the best-fitting model parameters ($N_{\rm min} \times M_{\rm min}$).

When fitting for the two free parameters, we reject data points that exceed the time-scale obtained for a well-sampled IMF (i.e. $t_{\rm E}/t_{\rm E,0} > 1$) by more than 1σ. This is to remedy an issue at low-mass scaling factors (typically $F_{\rm m} \approx 0.01$), where we find that the emission from the continuum can dominate over the SFR tracer when using the H α + filter. This results in characteristic time-scales that describe the long-lived continuum emission and therefore can be orders of magnitude higher than $t_{E,0}$. The above data selection criterion also affects the UV filters at low-mass scaling factors ($F_{\rm m} \approx 0.01$), where the emission becomes dominated by UV-faint, low-mass stars. As a result, the turn-off from $t_{\rm E}/t_{\rm E,0} = 1$ is generally poorly sampled (see Fig. 11 for an example), which means that we cannot reliably distinguish between different $N_{\rm min}$ and $M_{\rm min}$. Therefore, we conclude that UV emission is not significantly affected by IMF sampling.

Table 5 lists the best-fitting values of $M_{\rm min}$ ($N_{\rm min} = 1$ in all cases) for the full range of metallicities ($Z/Z_\odot = 0.05, 0.20, 0.40, 1.00, 2.00$) for the H α filters. We find no unambiguous relation between $M_{\rm min}$ and metallicity or filter width. However, smaller filter widths generally have higher $M_{\rm min}$. Higher values of $M_{\rm min}$ imply higher star-forming region masses below which IMF sampling affects the SFR tracer time-scale.

Table 5. The functional form of the conversion factor, $P(N \geq N_{\rm min})$, between the characteristic time-scale measured for a well-sampled IMF and a stochastically sampled IMF has two parameters, $N_{\rm min}$ and $M_{\rm min}$. We use $N_{\rm min} = 1$ and show here the values of $M_{\rm min}$.

	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

We combine the best-fitting values in Table 5 with the expressions for $P(N \geq N_{\rm min})$ provided in Appendix E to predict the region masses below which incomplete IMF sampling affects the SFR tracer time-scales. For H α −(+), this range is $\overline{M}_r \gtrsim 600$–$800(200$–$900)$ $M_\odot$. For a region separation length of $\lambda = 200$ pc and a total timeline duration of $\tau = 20$ Myr (typical for nearby star-forming galaxies, see Chevance et al. 2020c), these characteristic region mass limits

correspond to $\Sigma_{SFR} \gtrsim$ (1.0–1.3) $\times$ 10^{-3} $M_{\odot}$ yr^{-1} kpc^{-2} for Hα − and $\Sigma_{SFR} \gtrsim$ (0.3–1.4) $\times$ 10^{-3} $M_{\odot}$ yr^{-1} kpc^{-2} for Hα +.

Fig. 11 demonstrates that it is important to consider the effects of IMF sampling at low SFR surface densities, when constraining the characteristic time-scale for the Hα $\pm$ filters. This is because at low SFR surface densities, the massive stars required to produce Hα emission are not always present. If we ignore this fact, the Hα $\pm$ characteristic time-scales will be overestimated; as a result, the evolutionary time-line would be incorrectly calibrated and the time-scales obtained 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 Appendix E accurately describes how the characteristic time-scale of Hα $\pm$ changes due to incomplete IMF sampling. This means that observational applications of the KL14 principle can use the expressions provided in equations (14) and (E6)–(E9) to derive an SFR tracer time-scale corrected for IMF sampling. For the UV tracers, however, we find that the characteristic time-scales are mostly insensitive to the effects of incomplete IMF sampling.

7 COMPARISON TO OBSERVATIONS

This paper predicts the characteristic emission time-scales for SFR tracers, this can only be done using a galaxy simulation because it requires a reference map of which the duration of emission is known exactly. For the simulation, we do this by constructing an artificial reference map occupied by the star particles within a known age range and therefore known reference time-scale. It is not possible to do this for observed galaxies. However, it is possible to test whether the observed *ratio* between the emission time-scales of two different SFR tracers is consistent with our predictions. In Kruijssen et al. (2019), we used our predicted Hα reference time-scale at the half-solar metallicity of NGC 300 ($t_{E,0}^{H\alpha} = 4.59 \pm 0.14$ Myr) to measure a CO cloud lifetime of $10.8^{+2.1}_{-1.7}$ Myr. We can now use the measured CO cloud lifetime as a reference time-scale in an experiment combining the CO map (now acting as the reference map) with a UV emission map. This allows us to test if the resulting UV emission time-scale is consistent with our prediction for the UV reference time-scale.

As a first test of the accuracy of our inferred time-scales, we combine the GALEX FUV map of NGC 300 with the CO data presented in Kruijssen et al. (2019). For the CO map, we use the identical experiment setup as in Kruijssen et al. (2019), adopting the same set of emission peaks identified there and removing diffuse emission in the same way. For the FUV map, the emission peaks on which the apertures are placed 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 λ (Hygate et al. 2019). Other than these details, we apply the default analysis described in Kruijssen et al. (2018, 2019). The resulting tuning fork diagram (also see Fig. 1) is shown in Fig. 12. We obtain a good fit, with a UV emission time-scale of $t_{obs}^{FUV} = 23.1^{+5.9}_{-3.5}$ Myr. Given the half-solar metallicity of NGC 300, this should be compared to the reference time-scale predicted by equation (12) for $Z = 0.5$ $Z_{\odot}$, which is $t_{E,0}^{UV} = 19.2 \pm 2.0$ Myr.

Fundamentally, this experiment expresses the reference time-scale for *GALEX* FUV emission in units of the reference time-scale of continuum-subtracted Hα emission (using CO as an intermediate step). This is the case because in both the CO-Hα experiment of Kruijssen et al. (2019) and the CO-FUV experiment carried out
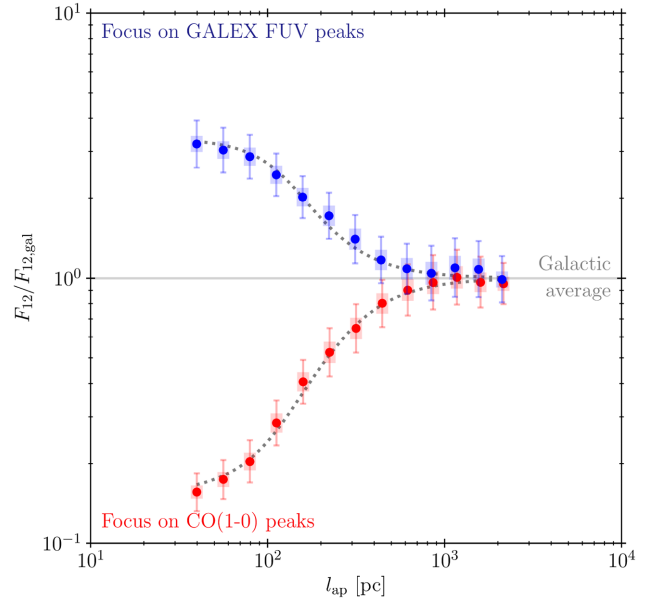

Figure 12. Tuning fork diagram obtained for the combination of CO(1–0) (Phase 2) and GALEX FUV (Phase 1) emission of NGC 300 (see Kruijssen et al. 2019). The symbols show the relation between the phase-1-to-phase-2 flux ratio (F_{12}) calculated at the locations of the emission peaks, relative to the galactic-scale phase-1-to-phase-2 flux ratio ($F_{12, gal}$) 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.

here, we have only measured the time-scale ratios. We should thus compare the observed and predicted $t^{FUV}/t^{H\alpha}$ ratio. We measure $t_{obs}^{FUV}/t_{obs}^{H\alpha} = 5.0^{+1.3}_{-0.8}$, whereas the calibration of this paper predicts $t_{E,0}^{FUV}/t_{E,0}^{H\alpha} = 4.2 \pm 0.5$. These values agree to within the uncertainties (at 0.9σ, or ∼20 per cent), which acts as a first demonstration that the reference time-scales derived in this work are consistent with observations.

Our method yields a measurement of the FUV-to-Hα time-scale *ratio*. This means that an arbitrary scaling of both the Hα and FUV characteristic time-scales, for either the predicted or observed ratios, would also result in agreement. However, the absolute time-scales 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 Hα emission after 1.7 Myr, and 95 per cent after 4.7 Myr. Our characteristic Hα time-scale at solar metallicity of 4.3 Myr falls within this range. The same applies for our GALEX FUV time-scale 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 Hα and UV observations of nearby galaxies will enable a more comprehensive test of the presented time-scales.

8 CONCLUSIONS

We have applied a new statistical method (the HEISENBERG code that uses the 'uncertainty principle for star formation', Kruijssen & Longmore 2014; Kruijssen et al. 2018) to constrain the characteristic emission time-scales of SFR tracers, i.e. the durations over which Hα and UV emission emerge from coeval stellar populations,

specifically within the 'uncertainty principle' formalism. We expect these time-scales to be critical in a variety of future studies. First, observational applications of HEISENBERG will enable the empirical characterization of the cloud lifecycle across a wide range of galactic environments, by measuring e.g. the molecular cloud lifetime and the time-scale for cloud destruction by feedback. However, in order to lead to physically meaningful constraints, these applications require the use of a known 'reference time-scale' for turning the measured relative time-scales into absolute ones. This reference time-scale is provided by the SFR tracer time-scales obtained in this work. Secondly, the emission time-scales obtained here and their dependence on metallicity and filter properties provide a helpful point of reference for studies of photoionization feedback and UV heating.

To obtain the SFR tracer emission time-scales, we generate synthetic SFR tracer emission maps of a simulated near-$L^\star$ isolated flocculent spiral galaxy using the stochastic SPS code SLUG2 (da Silva et al. 2012, 2014; Krumholz et al. 2015). We then apply HEISENBERG to combinations of these synthetic maps and an independent set of 'reference maps', which show the star particles from the simulation in specific, known age bins. With this approach, we self-consistently measure the characteristic time-scales for Hα 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 characteristic time-scales for Hα with continuum subtraction to be $4.3^{+0.1}_{-0.3}$ Myr, and 5.6–16.4 Myr without. For the UV filters, the reference time-scale 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 characteristic time-scales increases, to 3.9–5.5 Myr for Hα with continuum subtraction and 5.1–16.4 Myr without, as well as 14.5–33.3 Myr for the UV filters. We define empirical power-law relations that provide the characteristic time-scale as a function of metallicity (equations 10–12). These empirical relations include the response-weighted mean wavelength ($\overline{\lambda}_w$) for UV filters and the filter width (W) for the H$\alpha+$ filters. These dependences enable the use of a single expression to determine the characteristic time-scale for all UV and H$\alpha \pm$ SFR filters for a given combination of filter properties and the metallicity of the environment.

We also investigate the effects of a stochastically sampled IMF on the characteristic time-scales. Incomplete IMF sampling is found to affect the obtained characteristic emission time-scales in low-Σ_{SFR} galaxies. We quantify this dependence by stochastically sampling from the IMF prior to generating the synthetic SFR tracer emission maps and then measuring the characteristic time-scales with HEISENBERG. We use a Chabrier (2005) IMF to calculate the probability, P, of forming at least N_{min} stars of mass M_{min} or higher given a star-forming region mass $\overline{M}_r$. We then demonstrate that this probability is a good predictor for the ratio between the characteristic time-scale for a stochastically sampled IMF, t_E, and that of a well-sampled IMF, $t_{E,0}$. As a result, we obtain a relation between $t_E/t_{E,0}$ and the characteristic mass of independent star-forming regions, $\overline{M}_r$. Given an SFR surface density (from which the characteristic region mass can be derived), this relation quantifies the relative change of the SFR tracer time-scale due to IMF sampling as a function of the galactic environment.

For UV tracers, the impact of IMF sampling on the characteristic time-scale 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 characteristic time-scales of Hα emission. At low SFR surface densities, the Hα emission time-scale is suppressed due to IMF sampling effects. Depending on the metallicity and on whether the continuum emission has been subtracted, the characteristic time-scale for a well-sampled IMF can be used for $\overline{M}_r \gtrsim 200$–900 M$_\odot$, which for a region separation length of $\lambda = 200$ pc and a total timeline duration of $\tau = 30$ Myr corresponds to $\Sigma_{SFR} \gtrsim (0.3$–$1.3) \times 10^{-3}$ M$_\odot$ yr^{-1} kpc^{-2}. However, at lower region masses or SFR surface densities, the Hα reference time-scale must be corrected to account for the effects of IMF sampling. We derive fitting functions describing the change of the Hα time-scales as a function of the average independent star-forming region mass, $\overline{M}_r$, as parametrized by the minimum stellar mass required for Hα emission, M_{min}, which we tabulate as a function of metallicity (equations 14, 15 and E6–E9 as well as Table 5).

Even though we have arrived at the above reference time-scales by carrying out a set of numerical experiments using a galaxy simulation, and one could thus 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). The measurements carried out in this work require a physically motivated correlation of positions and ages of star particles. The critical goal was to carry out these measurements self-consistently within the framework of our method and thus enabling its future observational applications. In principle, this measurement could have been performed using maps of randomly generated distributions of regions: fundamentally, we have only characterized 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 accuracy of the results is demonstrated by a first comparison to observations of Hα and *GALEX* FUV emission in the nearby galaxy NGC 300 (Section 7), which shows that the time-scales predicted by this work are consistent with the observed time-scales.

In summary, we have measured the characteristic emission time-scales of SFR tracers within the 'uncertainty principle' formalism, as a function of metallicity and (for UV and 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 affect the time-scales of Hα and UV emission, and provides important constraints on the duration of photoionization feedback and UV heating. The reference time-scales derived in this work enable observational applications of the 'uncertainty principle for star formation', in which they are used to turn the relative durations of evolutionary phases into an absolute timeline. Specifically, we expect that the fitting functions provided in equations (10)–(12) and equations (E7)–(E9) will have great practical use because they enable the straightforward calculation of the reference time-scales as a function of metallicity, UV filter wavelength, and SFR surface density. Indeed, the first applications of this method have already used these equations to infer the time-scales driving cloud evolution, star formation, and feedback (Kruijssen et al. 2019; Chevance et al. 2020a, c; Hygate et al. 2020; Ward et al. 2020b, as well as section 7 of this paper). In view of the variety of recent and upcoming applications of this method, the time-scales presented in this work represent an essential ingredient towards empirically constraining the physics driving molecular cloud lifecycle.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Barlow R., 2003, in Lyons L., Mount R., Reitmeyer R., eds, Statistical Problems in Particle Physics, Astrophysics, and Cosmology. p. 250

Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846

Blitz L., Fukui Y., Kawamura A., Leroy A., Mizuno N., Rosolowsky E., 2006, Protostars and Planets V. University of Arizona Press, Tucson, p. 81

Calzetti D. et al., 2007, ApJ, 666, 870

Catinella B. et al., 2018, MNRAS, 476, 875

Chabrier G., 2005, in Corbelli E., Palla F., Zinnecker H., eds, Astrophysics and Space Science Library, Vol. 327, The Initial Mass Function 50 Years Later. Springer-Verlag, Berlin, p. 41

Charbonnel C., Meynet G., Maeder A., Schaller G., Schaerer D., 1993, A&AS, 101, 415

Chevance M. et al., 2020a, MNRAS, submitted

Chevance M. et al., 2020b, Space Sci. Rev., 216, 50

Chevance M. et al., 2020c, MNRAS, 493, 2872

Corbelli E. et al., 2017, A&A, 601, A146

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

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

Dehnen W., Aly H., 2012, MNRAS, 425, 1068

Dobbs C. L. et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI, University of Arizona Press, Tuscon, p. 3

Elmegreen B. G., 2000, ApJ, 530, 277

Fujimoto Y., Chevance M., Haydon D. T., Krumholz M. R., Kruijssen J. M. D., 2019, MNRAS, 487, 1717

Hao C.-N., Kennicutt R. C., Johnson B. D., Calzetti D., Dale D. A., Moustakas J., 2011, ApJ, 741, 124

Haydon D. T., Fujimoto Y., Chevance M., Kruijssen J. M. D., Krumholz M. R., Longmore S. N., 2020, MNRAS

Hopkins P. F. et al., 2018, MNRAS, 480, 800

Hughes I. G., Hase T. P. A., 2010, Measurements and their Uncertainties: A Practical Guide to Modern Error Analysis. Oxford Univ. Press, New York, NY

Hu C.-Y., Naab T., Walch S., Moster B. P., Oser L., 2014, MNRAS, 443, 1173

Hygate A. P. S., Kruijssen J. M. D., Chevance M., Schruba A., Haydon D. T., Longmore S. N., 2019, MNRAS, 488, 2800

Hygate A. P. S., Kruijssen J. M. D., Chevance M., Walter F., Schruba A., Kim J. J., Haydon D. T., Longmore S. N., 2020, MNRAS submitted

James P. A., Shane N. S., Knapen J. H., Etherton J., Percival S. M., 2005, A&A, 429, 851

Katz N., 1992, ApJ, 391, 502

Kawamura A. et al., 2009, ApJS, 184, 1

Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531

Kim J. J. et al., 2020, MNRAS, submitted

Koda J. et al., 2009, ApJ, 700, L132

Kroupa P., 2001, MNRAS, 322, 231

Kruijssen J. M. D., Longmore S. N., 2014, MNRAS, 439, 3239

Kruijssen J. M. D., Schruba A., Hygate A. P. S., Hu C.-Y., Haydon D. T., Longmore S. N., 2018, MNRAS, 479, 1866

Kruijssen J. M. D. et al., 2019, Nature, 569, 519

Krumholz M. R., 2014, Phys. Rep., 539, 49

Krumholz M. R., Fumagalli M., da Silva R. L., Rendahl T., Parra J., 2015, MNRAS, 452, 1447

Leisawitz D., Bash F. N., Thaddeus P., 1989, ApJS, 70, 731

Leitherer C. et al., 1999, ApJS, 123, 3

Leitherer C., Ekström S., Meynet G., Schaerer D., Agienko K. B., Levesque E. M., 2014, ApJS, 212, 14

Leroy A. K. et al., 2012, AJ, 144, 3

Leroy A. K. et al., 2013, AJ, 146, 19

McKee C. F., Williams J. P., 1997, ApJ, 476, 144

Meidt S. E. et al., 2015, ApJ, 806, 72

Miura R. E. et al., 2012, ApJ, 761, 37

Murphy E. J. et al., 2011, ApJ, 737, 67

Orear J., 1982, Am. J. Phys., 50, 912

Saintonge A. et al., 2017, ApJS, 233, 22

Schaerer D., Meynet G., Maeder A., Schaller G., 1993a, A&AS, 98, 523

Schaerer D., Charbonnel C., Meynet G., Maeder A., Schaller G., 1993b, A&AS, 102, 339

Schaller G., Schaerer D., Meynet G., Maeder A., 1992, A&AS, 96, 269

Scoville N. Z., Solomon P. M., Sanders D. B., 1979, in Burton W. B., ed., Proc. IAU Symp. Vol. 84, The Large-Scale Characteristics of the Galaxy. Kluwer, Dordrecht, p. 277

Springel V., 2005, MNRAS, 364, 1105

Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776

Vázquez G. A., Leitherer C., 2005, ApJ, 621, 695

Ward J. L., Chevance M., Kruijssen J. M. D., Kim J. J., Hygate A. P. S., Schruba A., Longmore S. N., 2020a, MNRAS, submitted

Ward J. L., Chevance M., Kruijssen J. M. D., Hygate A. P. S., Schruba A., Longmore S. N., 2020b, MNRAS, 497, 2286

Zabel N. et al., 2020, MNRAS, 496, 2155

APPENDIX A: Hα 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 (da Silva et al. 2012, 2014; Krumholz et al. 2015). SLUG2 then calculates the predicted emission spectrum for each particle, to which we apply UV and Hα + filters (Hα − comes directly from the hydrogen-ionizing photon emission). However, the emission spectrum that SLUG2 produces does not include the underlying Hα absorption from the stellar continuum. In this appendix, we use STARBURST99 (Leitherer et al. 1999; Vázquez & Leitherer 2005) to investigate when the Hα 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 Hα 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
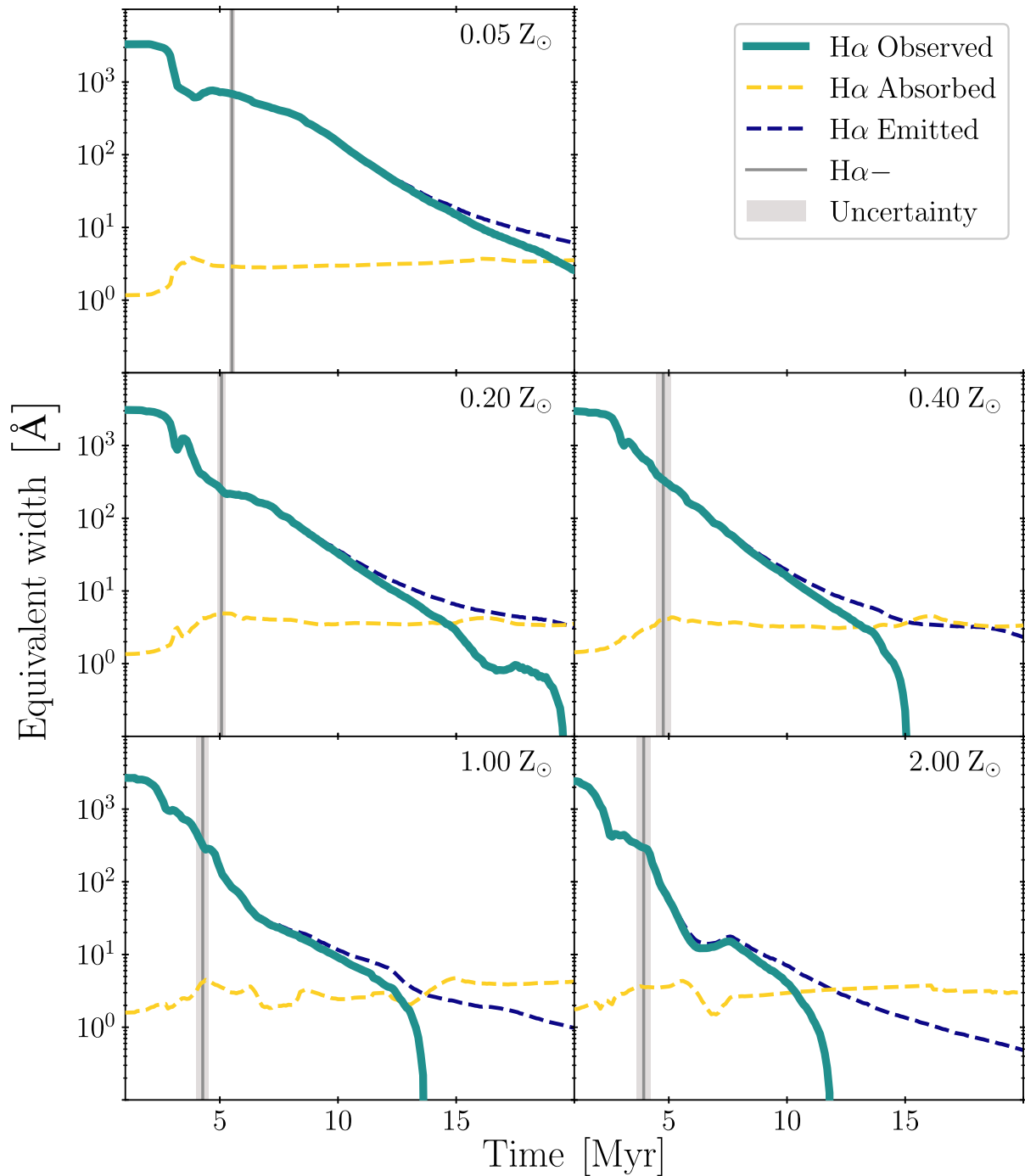
Figure A1. The results of STARBURST99 simulations for an instantaneous burst of star formation at 0 Myr. We show the change in equivalent width of the H α absorption and emission feature; we also include the difference between the two equivalent widths (H α observed). The H α − time-scale is marked for comparison.

the absorption feature (Voigt profile) of the high-resolution spectral data in the wavelength range 6482 Å $\leq \lambda \leq$ 6642 Å.

In Fig. A1, 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 α − time-scales we measure (see Table D1) which are also marked in Fig. A1.

We can see from Fig. A1 that the H α time-scales we are considering fall comfortably within the emission-dominant regime

and conclude that the absorption feature can safely be neglected for our analysis.

APPENDIX B: INFLUENCE OF THE WEIGHTING SCHEME ON THE RESULTS

In Section 4, we adopt a simple weighting scheme to combine the measured emission time-scales for different choices of t_M and t_R (see Fig. 4) into a single value. Because our method performs best when $t_{E,0} \sim t_M \sim t_R$ (Kruijssen et al. 2018), this scheme calculates the geometric distance in logarithmic space between the measured
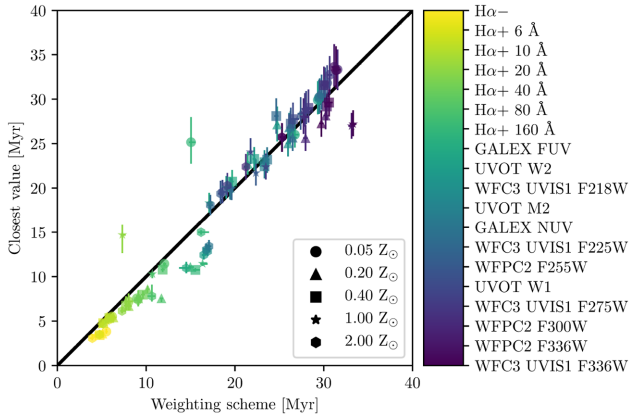
Figure B1. Comparison of the time-scales measured using the weighting scheme adopted in this paper (x-axis) and the time-scales that would have been obtained by taking the value of $t_{E,0}$ that is closest to t_M and t_R (y-axis). The data points are colour coded by filter (see the colour bar), with the shape denoting the metallicity. This figure shows that both methods are in rough agreement, but that there exist clear outliers. For these cases, i.e. when both methods disagree, we demonstrate in the text and in Fig. B2 that the weighting scheme provides the most accurate result.

time-scale and t_M and t_R, and then defines the weight of each cell as its inverse ($\mathcal{W}_{ij}^d$, equation 5). As is common practice, we also weight each cell by the inverse square of the uncertainty ($\mathcal{W}_{ij}^u$, equation 6). A much simpler approach would be to simply adopt the cell for which the measured time-scale is closest to t_M and t_R. In this appendix, we briefly demonstrate that this simple approach generally provides similar results to the adopted weighting scheme, but is more sensitive to stochasticity and outliers. This motivates the use of the adopted weighting scheme.

In Fig. B1, we show the comparison of the emission time-scales obtained with both of the above approaches. As the figure shows, both approaches agree to within ≲20 per cent, which largely validates our approach. However, there are also clear outliers, with deviations up

to a factor of 2. To investigate what causes these outliers, we show the full table of measurements for one of them (H α + 20 Å) in Fig. B2. This figure shows a fragmented weighting landscape, with two local minima around $t_{E,0} = 6-7$ Myr and $t_{E,0} \approx 15$ Myr. Due to the extreme proximity of the second of these two minima to the corresponding t_M and t_R, it has a high weight, even thought the first minimum is the correct answer. This is illustrated by (1) the fact that the surrounding cells have high weights too (i.e. it is not isolated) and (2) the fact that the first minimum is more consistent with the measurements for other metallicities and filter widths listed in Appendix D. Had we simply used the cell for which $t_{E,0}$ is closest to t_M and t_R, we would have selected the second minimum. By contrast, the adopted weighting scheme ends up with $t_{E,0} = 7.3^{+0.4}_{-0.2}$ Myr. This is very close to the correct answer found near the first minimum and shows that the second minimum has a minor effect on the final result when using the adopted weighting scheme.

In principle, we could have adopted a different functional form for equation (5). However, this would necessarily have been more ad hoc than a simple inverse dependence on the geometric distance in logarithmic space used here. We therefore prefer to adopt the simplest approach that avoids a strong sensitivity to outliers, which we have done in Section 4.

APPENDIX C: COMPARISON OF THE ANALYTICAL TIME-SCALE FITS TO THE MEASUREMENTS

In Figs C1 and C2, we compare the best-fitting analytical functions for predicting the H α + and UV SFR tracer time-scales (equations 11 and 12) to the measurements obtained with Heisenberg. These one-dimensional projections of Figs 9 and 10 are useful for identifying parts of parameter space in which the analytical fits do not describe the measurements well, i.e. at long UV wavelengths and solar metallicity (which is the largest discrepancy found across the metallicity-wavelength space considered). In such cases, it might be advisable to use the measurement rather than the analytical expression.

Hα+ 20 Å	$t_R : 1$	$t_R : 3$	$t_R : 5$	$t_R : 7$	$t_R : 10$	$t_R : 15$	$t_R : 20$	$t_R : 25$	$t_R : 30$
$t_M : 1$	$4.38^{+0.63}_{-0.69}$	$6.75^{+0.19}_{-2.20}$	$5.84^{+0.43}_{-0.26}$	$6.71^{+0.46}_{-0.40}$	$6.57^{+0.65}_{-0.43}$	$5.82^{+0.93}_{-0.56}$	$6.03^{+0.80}_{-0.28}$	$6.55^{+1.05}_{-0.29}$	$6.60^{+0.44}_{-0.32}$
$t_M : 3$	$4.90^{+0.68}_{-0.48}$	$5.64^{+0.46}_{-0.45}$	$6.80^{+0.45}_{-0.35}$	$7.16^{+0.41}_{-0.34}$	$6.18^{+0.33}_{-0.31}$	$6.15^{+0.39}_{-0.29}$	$6.64^{+0.35}_{-0.27}$	$7.92^{+0.56}_{-0.34}$	$8.62^{+0.51}_{-0.39}$
$t_M : 5$	$3.94^{+1.49}_{-0.42}$	$7.42^{+1.37}_{-0.59}$	$8.92^{+1.65}_{-0.90}$	$7.36^{+0.98}_{-0.51}$	$7.23^{+0.98}_{-0.48}$	$7.66^{+1.08}_{-0.71}$	$6.73^{+0.85}_{-0.42}$	$10.29^{+1.51}_{-1.21}$	$10.32^{+1.58}_{-1.13}$
$t_M : 7$	$21.04^{+8.49}_{-5.17}$	$15.23^{+5.28}_{-2.67}$	$6.62^{+1.24}_{-0.71}$	$5.97^{+1.45}_{-0.67}$	$8.72^{+1.46}_{-1.07}$	$9.69^{+1.56}_{-1.39}$	$10.36^{+1.91}_{-1.48}$	$10.01^{+1.62}_{-1.43}$	$8.81^{+1.65}_{-1.03}$
$t_M : 10$	$3.20^{+0.48}_{-0.58}$	$3.57^{+0.67}_{-0.42}$	$6.19^{+0.81}_{-0.90}$	$7.02^{+0.56}_{-1.06}$	$7.96^{+0.76}_{-1.02}$	$8.70^{+1.17}_{-1.09}$	$9.65^{+1.37}_{-1.16}$	$8.69^{+1.53}_{-0.97}$	$9.14^{+1.67}_{-1.02}$
$t_M : 15$	$4.81^{+0.64}_{-0.80}$	$5.45^{+0.61}_{-0.65}$	$7.18^{+0.71}_{-0.82}$	$8.69^{+0.75}_{-1.02}$	$10.02^{+0.85}_{-1.37}$	$14.70^{+1.16}_{-2.05}$	$10.67^{+1.64}_{-1.34}$	$11.30^{+1.73}_{-1.39}$	$13.90^{+1.96}_{-1.88}$
$t_M : 20$	$6.50^{+1.51}_{-1.42}$	$6.41^{+1.81}_{-1.06}$	$8.84^{+2.01}_{-1.50}$	$9.29^{+1.82}_{-1.52}$	$11.42^{+2.10}_{-1.82}$	$9.22^{+1.61}_{-1.13}$	$10.38^{+1.74}_{-1.31}$	$10.78^{+1.84}_{-1.37}$	$8.60^{+1.51}_{-1.03}$
$t_M : 25$	$5.71^{+2.93}_{-1.03}$	$9.28^{+2.44}_{-1.61}$	$10.24^{+1.54}_{-1.69}$	$10.49^{+2.16}_{-1.49}$	$9.79^{+1.73}_{-1.26}$	$10.84^{+1.63}_{-1.37}$	$11.98^{+1.69}_{-1.55}$	$10.01^{+1.73}_{-1.28}$	$8.81^{+1.48}_{-1.14}$
$t_M : 30$	$5.23^{+3.16}_{-0.84}$	$7.53^{+2.32}_{-1.27}$	$7.61^{+1.64}_{-1.16}$	$7.56^{+1.20}_{-1.07}$	$9.42^{+1.56}_{-1.18}$	$9.89^{+1.32}_{-1.27}$	$10.88^{+1.63}_{-1.35}$	$8.93^{+1.19}_{-1.07}$	$8.07^{+1.27}_{-0.88}$

Figure B2. Example showing the same table as in Fig. 4, but this time for H α including the continuum (H α +) with a 20 Å filter. As before, the reference maps are characterized by the age bin 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 bin. The colour coding is based on the weighting, $\mathcal{W}$, used when calculating the weighted average. All values within the tables are given in Myr. This figure shows that the simple approach of taking the measured value that is closest to t_M and t_R is sensitive to local minima (which would in this case overestimate the time-scale by a factor of two, see the offset star in Fig. B1), whereas the weighting scheme adopted in this work avoids this problem by calculating a weighted average of all cells (see the text for details).
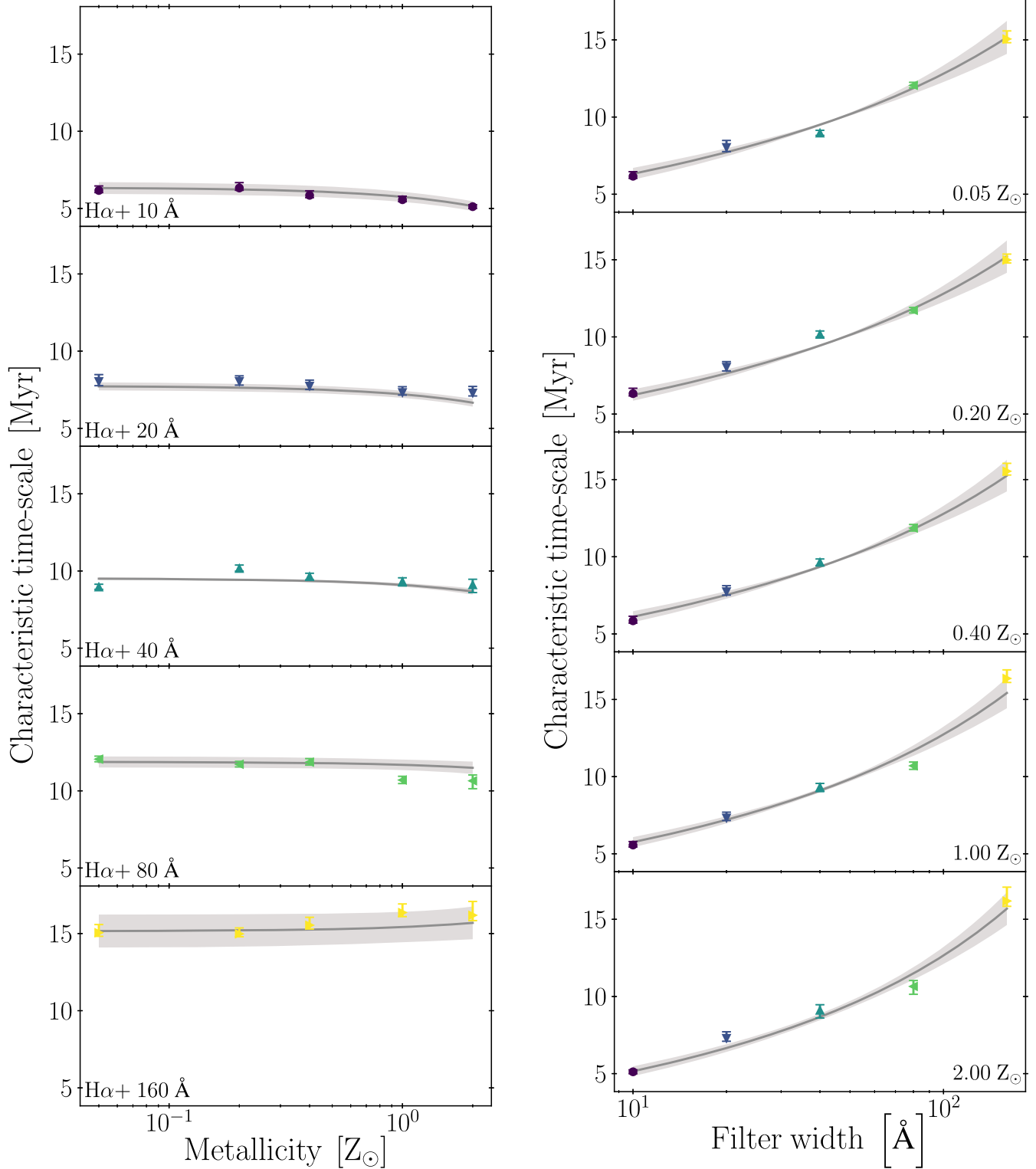
Figure C1. Characteristic time-scales of Hα + filters as a function of metallicity (left) and filter width (right). The symbols show the results of applying the HEISENBERG code to synthetic Hα + maps. The grey curve shows the fit from equation (11) and the shaded region indicates the associated uncertainty. The symbols and colours in the right-hand panels correspond to those used on the left.

Figure C2. Characteristic time-scales 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 (12) and the shaded region indicates the associated uncertainty. The symbols and colours in the right-hand panels correspond to those used on the left.
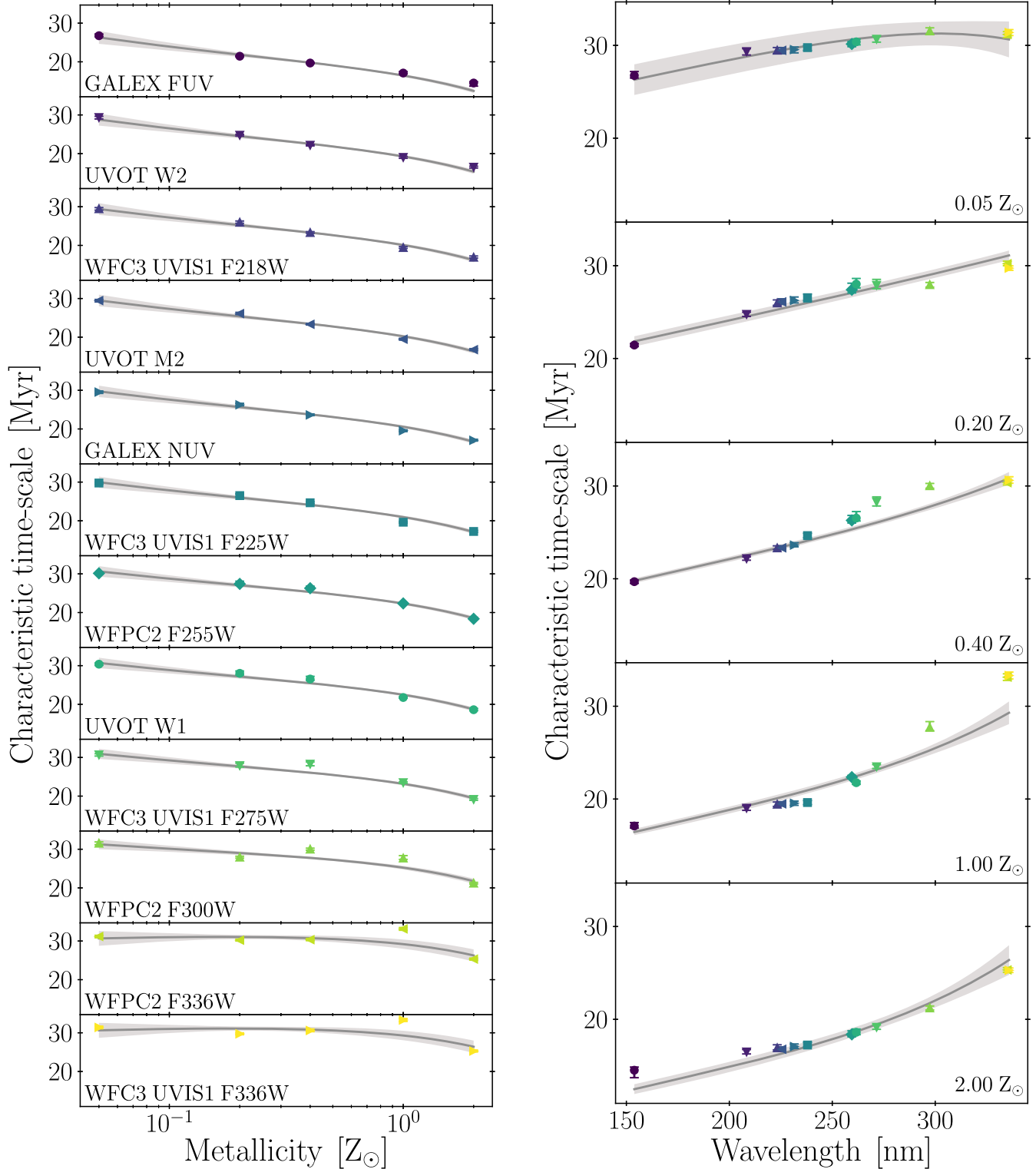
APPENDIX D: COMPLETE SET OF SFR TRACER EMISSION TIME-SCALES

In Table D1, we list the complete set of SFR tracer emission time-scales constrained in this paper. This contains the characteristic time-scales of Hα $\pm$ and all 12 UV filters, for the five different metallicities $Z/Z_\odot = 0.05, 0.20, 0.40, 1.00, 2.00$. In addition, we include the age intervals that we adopt to define the stellar reference maps used when measuring the SFR tracer time-scales with the HEISENBERG code. For more details on the calculations, see Section 4.

Table D1. A summary of all the characteristic time-scales and corresponding age bins (for producing reference maps in Sections 5 and 6), for the different star formation rate tracers (see Table 1 for details). These results are for a well-sampled IMF. The filter order is in increasing filter width (*W*) for Hα + and increasing response-weighted mean wavelength ($\bar{\lambda}_{\mathrm{w}}$) for UV. Table 3 is included in these tables under the 1.00 Z$_\odot$ heading.

	0.05 Z$_\odot$	0.20 Z$_\odot$	0.40 Z$_\odot$	1.00 Z$_\odot$	2.00 Z$_\odot$
		(a) Characteristic time-scales, $t_{\mathrm{E},0}$ [Myr].			
Hα −	$5.5^{+0.2}_{-0.1}$	$5.1^{+0.1}_{-0.2}$	$4.8^{+0.1}_{-0.3}$	$4.3^{+0.1}_{-0.3}$	$3.9^{+0.1}_{-0.3}$
Hα + 10 Å	$6.2^{+0.3}_{-0.1}$	$6.3^{+0.3}_{-0.1}$	$5.9^{+0.3}_{-0.1}$	$5.6^{+0.2}_{-0.1}$	$5.1^{+0.1}_{-0.1}$
Hα + 20 Å	$8.0^{+0.5}_{-0.3}$	$8.0^{+0.4}_{-0.2}$	$7.7^{+0.4}_{-0.2}$	$7.3^{+0.4}_{-0.2}$	$7.3^{+0.4}_{-0.2}$
Hα + 40 Å	$9.0^{+0.2}_{-0.2}$	$10.2^{+0.2}_{-0.2}$	$9.6^{+0.2}_{-0.2}$	$9.3^{+0.2}_{-0.3}$	$9.1^{+0.4}_{-0.5}$
Hα + 80 Å	$12.1^{+0.2}_{-0.2}$	$11.7^{+0.2}_{-0.2}$	$11.9^{+0.2}_{-0.2}$	$10.7^{+0.2}_{-0.2}$	$10.7^{+0.4}_{-0.5}$
Hα + 160 Å	$15.1^{+0.5}_{-0.2}$	$15.0^{+0.4}_{-0.2}$	$15.5^{+0.5}_{-0.2}$	$16.4^{+0.6}_{-0.3}$	$16.2^{+0.9}_{-0.3}$
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	$29.5^{+0.3}_{-0.3}$	$26.3^{+0.4}_{-0.3}$	$23.6^{+0.3}_{-0.3}$	$19.6^{+0.2}_{-0.2}$	$17.1^{+0.3}_{-0.2}$
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*	$31.4^{+0.3}_{-0.3}$	$29.7^{+0.2}_{-0.2}$	$30.6^{+0.4}_{-0.3}$	$33.3^{+0.4}_{-0.4}$	$25.3^{+0.2}_{-0.2}$
		(b) Age bins, $t_{\mathrm{E},0} \leq$ Age $\leq 2t_{\mathrm{E},0}$ [Myr].			
Hα −	5.5–11.0	5.1–10.1	4.8–9.5	4.3–8.6	3.9–7.9
Hα + 10 Å	6.2–12.3	6.3–12.7	5.9–11.7	5.6–11.1	5.1–10.2
Hα + 20 Å	8.0–16.1	8.0–16.1	7.7–15.5	7.3–14.6	7.3–14.6
Hα + 40 Å	9.0–17.9	10.2–20.3	9.6–19.3	9.3–18.6	9.1–18.2
Hα + 80 Å	12.1–24.1	11.7–23.4	11.9–23.7	10.7–21.4	10.7–21.3
Hα + 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

APPENDIX E: THEORETICAL EXPECTATIONS FOR THE EFFECT OF INCOMPLETE IMF SAMPLING

Here, we predict the relationship between how well the IMF is sampled and the characteristic time-scale of the SFR tracer. As mentioned in Section 6, the characteristic time-scale of the SFR tracer is related to the number of star-forming regions in the emission map. We therefore estimate the relative change of the effective SFR tracer time-scale 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 characterize 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.

We can calculate the probability, P, 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 $M_{\mathrm{min}} \leq M \leq M_{\mathrm{r}}$. 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.
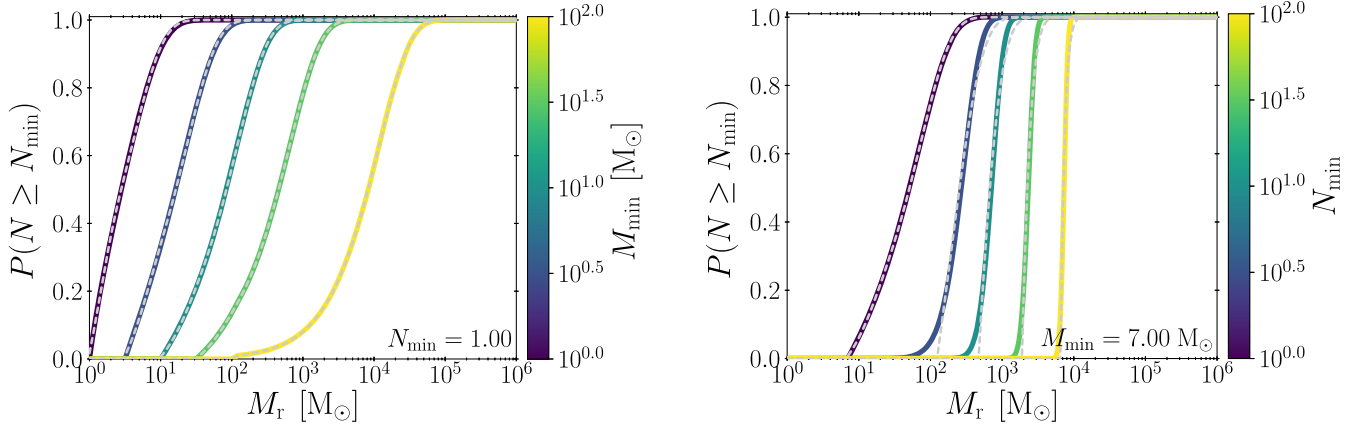
Figure E1. Curves showing how the probability of forming at least $N_{\min}$ stars of mass $M_{\min}$ or higher changes with star-forming region mass, M_r. The grey dashed lines indicate the approximate fit to the full calculation. See Appendix E for details and Table E1 for fit parameter values. *Left:* Constant $N_{\min}$. *Right:* Constant $M_{\min}$.

The binomial distribution gives the probability of k successful stars:

$$P(N = k) = \frac{N_\star!}{k!(N_\star - k)!} p^k (1 - p)^{N_\star - k}. \tag{E1}$$

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

$$P(N \geq N_{\min}) = 1 - P(N < N_{\min}) \tag{E2}$$

$$= 1 - \sum_{k=0}^{N_{\min}-1} P(N = k). \tag{E3}$$

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

$$p = \nu \int_{M_{\min}}^{M_r} \frac{dn}{dm} dm \; ; \; N_\star = \mu \int_0^{M_r} \frac{dn}{dm} dm. \tag{E4}$$

The normalization constants ν and μ are evaluated through

$$1 = \nu \int_0^{M_r} \frac{dn}{dm} dm \; ; \; M_r = \mu \int_0^{M_r} m \frac{dn}{dm} dm. \tag{E5}$$

In order to convert the probability value, $P(N \geq N_{\min})$, into an estimate for the characteristic time-scale, t_E, we assume a Chabrier (2005) IMF and use the characteristic time-scales we find for a fully sampled IMF, $t_{E,0}$, (see Appendix D) in the following equation:

$$t_E = t_{E,0} \times P(N \geq N_{\min}). \tag{E6}$$

In Fig. E1, 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 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 $N_{\min}$ and $M_{\min}$); the same effect is observed for $N_{\min}$ but less pronounced. Higher $N_{\min}$ 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 Fig. E1 have a complex analytical form, therefore we provide a four-parameter function that approximates these curves. These approximations are also included in Fig. E1 as dotted grey

Table E1. Values for the free parameters, a_i and b_i, in the analytical models presented in Fig. E1 and described by equations (E7)–(E9).

$N_{\min}$	$M_{\min}$	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

lines. The following set of equations describe the form of the approximation:

$$M_0 = N_{\min} \times M_{\min}, \tag{E7}$$

$$f(M_r) = 1 + a_1 \exp\left(b_1 \left[\frac{M_r}{M_0}\right]\right) + a_2 \exp\left(b_2 \left[\frac{M_r}{M_0}\right]^2\right), \tag{E8}$$

$$P(N \geq N_{\min}) \approx \begin{cases} 0 & f(M_r) \leq 0 \\ f(M_r) & 0 < f(M_r) < 1 \\ 1 & f(M_r) \geq 1 \end{cases}, \tag{E9}$$

where a_i, b_i for $i = \{1, 2\}$ are four parameters that we determine through least-squares minimization. We present the parameter values for all the approximate curves displayed in Fig. E1 in Table E1. For intermediate values of $M_{\min}$, these best-fitting parameters can be interpolated as a function of $\log(M_{\min})$. The approximate expression gives an almost identical fit in the cases where $N_{\min} = 1$ (see Fig. E1) but for higher values of $N_{\min}$ the approximation does not perform as well. Fortunately, as we show in Section 6, we only need to consider the case of $N_{\min} = 1$.

We now have a description of how the characteristic time-scale of SFR tracers in a star-forming region with a stochastically sampled IMF, t_E, is related to the characteristic time-scale determined when the IMF is well sampled, $t_{E,0}$, through a probability distribution function, $P(N \geq N_{\min})$. The IMF and two free parameters, $N_{\min}$ and $M_{\min}$, characterize the form of $P(N \geq N_{\min})$. We note that the analytical expression for the time-scale correction factor $P(N \geq N_{\min})$

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 characteristic SFR time-scale.

APPENDIX F: ERROR PROPAGATION

In Section 6.1, we calculate the average independent star-forming region mass as

$$\overline{M}_r = \Sigma_{SFR} \times (t_{emi} + t_{ref} - t_{over}) \times \pi \left(\frac{\lambda}{2}\right)^2. \tag{F1}$$

This equation uses the SFR surface density, Σ_{SFR}, and the duration of the reference map, t_{ref}, 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_{emi}; and the duration of the overlap between the emission and reference phases, t_{over}. We note that equations (F1) and (14) are equivalent through the definition

$$\tau \equiv t_{emi} + t_{ref} - t_{over}. \tag{F2}$$

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. $f(x_1, \ldots, x_N)$ is given by (Hughes & Hase 2010)

$$\sigma_f^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \rho_{ij} \sigma_i \sigma_j, \tag{F3}$$

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 (F3), we define the following:

$$\kappa \equiv \Sigma_{SFR} \frac{\pi}{4}, \tag{F4}$$

$$x_1 \equiv \lambda, \tag{F5}$$

$$x_2 \equiv t_{emi}, \tag{F6}$$

$$x_3 \equiv t_{over} \tag{F7}$$

and equation (F1) becomes

$$\overline{M}_r = \kappa \tau x_1^2 \tag{F8}$$

$$= \kappa (x_2 + t_{ref} - x_3) x_1^2. \tag{F9}$$

We note that Σ_{SFR} and t_{ref} 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

$$\frac{\partial}{\partial x_1} \overline{M}_r = 2\kappa \tau x_1, \tag{F10}$$

$$\frac{\partial}{\partial x_2} \overline{M}_r = \kappa x_1^2, \tag{F11}$$

$$\frac{\partial}{\partial x_3} \overline{M}_r = -\kappa x_1^2. \tag{F12}$$

Combining equation (F10) with equation (F3) we find the expression for the uncertainty on $\overline{M}_r$, $\sigma_{\overline{M}_r}$

$$\left[\frac{\sigma_{\overline{M}_r}}{\overline{M}_r}\right]^2 = \frac{4\sigma_1^2}{x_1^2} + \frac{(\sigma_2^2 + \sigma_3^2 - 2\rho_{23}\sigma_2\sigma_3)}{\tau^2} + \frac{4(\rho_{12}\sigma_1\sigma_2 - \rho_{13}\sigma_1\sigma_3)}{\tau x_1}. \tag{F13}$$

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 characteristic time-scales of SFR tracers.

This paper has been typeset from a TEX/LATEX file prepared by the author.