MNRAS **464,** 3315–3330 (2017)
Advance Access publication 2016 October 8
doi:10.1093/mnras/stw2580

Radiative transfer meets Bayesian statistics: where does a galaxy's [C II] emission come from?

G. Accurso,[1⋆] A. Saintonge,[1⋆] T. G. Bisbas[2,3] and S. Viti[1]

[1]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[2]*Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse 1, D-85748 Garching, Germany*
[3]*Department of Astronomy, University of Florida, Gainesville, FL 32611, USA*

Accepted 2016 October 6. Received 2016 October 3; in original form 2016 July 12; Editorial Decision 2016 October 5

ABSTRACT

The [C II] 158 μm emission line can arise in all phases of the interstellar medium (ISM), therefore being able to disentangle the different contributions is an important yet unresolved problem when undertaking galaxy-wide, integrated [C II] observations. We present a new multiphase 3D radiative transfer interface that couples STARBURST99, a stellar spectrophotometric code, with the photoionization and astrochemistry codes MOCASSIN and 3D-PDR. We model entire star-forming regions, including the ionized, atomic, and molecular phases of the ISM, and apply a Bayesian inference methodology to parametrize how the fraction of the [C II] emission originating from molecular regions, $f_{[\mathrm{C\,II}],\mathrm{mol}}$, varies as a function of typical integrated properties of galaxies in the local Universe. The main parameters responsible for the variations of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ are specific star formation rate (SSFR), gas phase metallicity, H II region electron number density (n_e), and dust mass fraction. For example, $f_{[\mathrm{C\,II}],\mathrm{mol}}$ can increase from 60 to 80 per cent when either n_e increases from $10^{1.5}$ to $10^{2.5}$ cm^{-3}, or SSFR decreases from $10^{-9.6}$ to $10^{-10.6}$ yr^{-1}. Our model predicts for the Milky Way that $f_{[\mathrm{C\,II}],\mathrm{mol}} = 75.8 \pm 5.9$ per cent, in agreement with the measured value of 75 per cent. When applying the new prescription to a complete sample of galaxies from the Herschel Reference Survey, we find that anywhere from 60 to 80 per cent of the total integrated [C II] emission arises from molecular regions.

Key words: astrochemistry – ISM: molecules – photodissociation region (PDR) – ISM: structure – infrared: galaxies – infrared: ISM.

1 INTRODUCTION

The [C II] 158 μm emission of singly ionized carbon is one of the strongest cooling lines of the interstellar medium (ISM), and can carry up to a few per cent of the total far-infrared (FIR) energy emitted from galaxies. It correlates with the total molecular gas mass of galaxies ($M_{\mathrm{H_2}}$), as measured from ^{12}CO emission, and with the total star formation rate (SFR) inferred from FIR luminosity. For these reasons, [C II] is an important and widely used tracer of massive star formation in galaxies at both low and high redshifts (Stacey et al. 2010; de Looze et al. 2011).

Ionized carbon (C^+) can be found throughout the ISM, from photodissociation regions (PDRs) to diffuse ionized and atomic regions, owing to the fact that carbon has a first ionization potential of 11.3 eV, lower than that of hydrogen (Kaufman et al. 1999; Kauffmann et al. 2003). While [C II] originates in good part from PDRs, explaining the correlation with $M_{\mathrm{H_2}}$ and SFR, observations have shown that a non-negligible fraction of the emission can originate from the ionized and diffuse atomic gas components where massive star formation does not occur (Heiles et al. 1994; Langer et al. 2010; Vasta et al. 2010; Kramer et al. 2013; Pineda et al. 2013).

The CO molecule, as a tracer of the cold molecular phase of the ISM, suffers from the opposite problem: in low-metallicity environments, the CO molecule can be photodissociated by UV radiation while H_2 self-shields and survives, resulting in the presence of molecular gas that is missed by CO observations. For example, Velusamy et al. (2010) find that ∼25 per cent of the molecular gas in the Milky Way may be CO-dark in such a way. If it were possible to discern the contribution of the different ISM phases to the total [C II] emission, then the combination of CO and [C II] measurements would increase significantly the accuracy of $M_{\mathrm{H_2}}$ calculations (Mookerjea et al. 2016). There is also a new interest in using [C II] as a probe of the ISM in $z > 5$ galaxies and up to well into the epoch of reionization; such studies are now made increasingly possible with facilities such as ALMA and NOEMA (e.g. Ota et al. 2014; Riechers et al. 2014; Maiolino et al. 2015). In this context also, disentangling the contributions from the different phases of the ISM is of significant importance, yet this problem remains unsolved.

⋆ E-mail: gioacchino.accurso.13@ucl.ac.uk (GA); a.saintonge@ucl.ac.uk (AS)

Solving this problem requires one of two things: either high spatial resolution observations of several FIR lines such as [N II]122,205 μm and [O I]63,145 μm in addition to [C II], or a self-consistent model of the ISM on galaxy-wide scales, including PDRs, ionized and neutral diffuse regions. Since such detailed observations are only available for a handful of very nearby galaxies, we focus here on the modelling approach. Numerical codes treating PDRs have been around for decades and have now grown into complex models capable of solving the thermal balance equations and chemical reactions occurring within these regions (Viti 2013). Some codes have aimed to include all the small-scale physics to describe the chemical and thermal processes at work in the gas and grains, while others focus on treating the gas–grain chemistry while approximating other processes. Various codes treating one-dimensional PDRs have been developed in the past and only recently three-dimensional codes have emerged that can treat PDRs of an arbitrary density distribution (Bisbas et al. 2012). Furthermore, calculations typically treat the ionized and PDR regions separately (Röllig et al. 2006), which is problematic for emission lines such as [C II] that originate in both these phases of the ISM and can lead to overestimations of line intensities and incorrect interpretations of the physical conditions in the ISM (e.g. hydrogen column density and incident ionization field).

Photoionization codes, used to model the H II regions, have likewise been around for several decades now. They typically work by solving the equations of radiative transfer while making assumptions concerning spherical symmetry. The earliest H II region models contained the basic physics of ionization, recombination of hydrogen and helium, thermal balance, and the emission of photons from the nebula (Flower 1968), with subsequent codes having seen the addition of other important physical processes such as charge exchange and dielectric recombination, and the consideration of a wider range of ions. More recently, three-dimensional codes have been developed to handle varying geometries using a Monte Carlo (MC) approach to solve the 3D equations of radiative transport (Wood, Mathis & Ercolano 2004).

Although numerical models for the individual components of the ISM are aplenty, codes which can simulate all aforementioned phases of the ISM consistently are not so common. One very successful example is CLOUDY (Ferland et al. 2013), a plasma simulation that models the ionization, chemical and thermal states of the gas that may be exposed to an external radiation field coming from a nearby heating sources such as star clusters. The code works by predicting the spectrum from this non-equilibrium gas and simulating its level populations as well as its ionization, molecular and thermal states, over a wide range of densities and temperatures. However, CLOUDY is intrinsically a 1D code, and involves assumptions concerning the thermal balance in PDRs. Both of these limitations should be lifted to accurately simulate the entire ISM of star-forming regions and galaxies. Indeed, it has been shown that results from 1D and 3D simulations vary depending on the specific physical conditions (Ercolano et al. 2012). An example of a successful attempt at building such a 3D, multiphase radiative transfer code is TORUS-3DPDR (Bisbas et al. 2015), a hydrodynamics and MC radiative transfer code. TORUS-3DPDR does not use the complexity of MOCASSIN to calculate particular photoionization calculations and therefore would not be as accurate.

In this paper, we present a new modelling interface that combines self-consistently state-of the-art astrochemistry and photoionization codes: 3D-PDR, a three-dimensional code for treating PDRs and molecular regions, MOCASSIN, a full 3D MC photoionization code, and STARBURST99, a stellar population synthesis code. This integrated code is used to simulate entire star-forming regions, including the ionized, neutral and molecular phases of the ISM, with the aim of parametrizing how the fraction of the total [C II] emission originating from molecular regions, $f_{[\mathrm{C\,II}],\mathrm{mol}}$, varies as a function of the physical conditions in the ISM. A Bayesian inference technique is used to solve this complex multiparameter problem, allowing us to derive a series of prescriptions to calculate the contributions of the different phases of the ISM to the total integrated [C II] emission in extragalactic sources.

In Sections 2 and 3, we provide further technical details concerning the codes and the modelling strategy, while the choice of input parameters is explained in Section 4. The results of the modelling and of the Bayesian analysis to produce scaling relations for $f_{[\mathrm{C\,II}],\mathrm{mol}}$ are presented in Sections 5 and 6, respectively. These prescriptions are validated through comparisons with observations in Sections 6.5. Finally, we conclude in Section 7 with a concise presentation of our new prescriptions to infer $f_{[\mathrm{C\,II}],\mathrm{mol}}$ from galaxy-integrated quantities.

Throughout this paper, we use a standard flat Λcolddarkmatter cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and the initial mass function (IMF) from Chabrier (2003).

2 OVERVIEW OF NUMERICAL CODES

We here briefly describe each of the codes used within this work; for more detailed explanations, see the references provided for each code.

2.1 The STARBURST99 code

STARBURST99 (Leitherer et al. 1999, 2010; Vázquez & Leitherer 2005; Conroy 2013), hereafter SB99, is a multipurpose evolutionary synthesis code that models the spectrophotometric properties of unresolved stellar populations, and makes predictions for various observables, such as spectral energy distributions (SEDs), supernovae rates and mass loss rates. The code simulates a population of stars based on an input metallicity and stellar IMF, and evolves them across the Hertzprung–Russell (HR) diagram. The code creates a grid covering the HR diagram and integrates over it with weights assigned according to properties of the stellar population to produce the integrated spectrum for the population. SB99 can thus provide the relation between the stellar mass (M_t) and luminosity (L_t of the stellar population at any time interval; see Conroy (2013) for a comprehensive review of the synthesis technique.

2.2 The MOCASSIN code

The numerical code MOCASSIN [MOnte CArlo SimulationS of Ionized Nebulae; Ercolano et al. (2003)] is a three-dimensional MC radiative transfer code that operates on a non-uniform Cartesian grid. It was originally intended as a tool to construct realistic gas models of planetary nebulae, but has since evolved to incorporate dust radiative transfer (Ercolano, Barlow & Storey 2005) and is now used to simulate ionized gas emission on galaxy-wide scales (Karczewski et al. 2013). Photoionization calculations are performed using an iterative MC photon energy packet propagating routine, based on the methods presented by Lucy (1999). Photons are emitted from the ionizing source in random but isotropic directions, and propagate for a path length, l, determined by a randomly selected optical depth (Harries & Howarth 1997). An abundance file is used as an input, providing the chemical abundance of each species, along with an input SED and files specifying the dust properties, opacities,

cross-sections, etc., while the dust-to-gas ratio is also specified as an input parameter. Given these, the code self-consistently solves the radiative transfer equations and calculates the gas and dust temperatures, ionization degree, and the overall emergent SED of the full dust, gas and stars network.

2.3 The 3D-PDR code

The 3D-PDR code (Bisbas et al. 2012) is a three-dimensional astrochemistry code that simulates PDRs of arbitrary density distribution. It solves the chemistry and the thermal balance self-consistently in each computational element of a given cloud and uses the chemical model features of Bell et al. (2006). Like MOCASSIN, the code has been used in various extragalactic applications such as modelling molecular line intensities in NGC 4038 (Bisbas et al. 2014) and neutral carbon mapping (Offner et al. 2014).

3D-PDR uses a ray-tracing scheme based on the HEALPIX (Górski et al. 2005) package that calculates properties along a given line of sight. This allows for the quick calculation of (a) column densities of species along a particular direction, (b) the attenuation of the Draine field in the PDR and (c) the propagation of the FIR/submm line emission out of the PDR.

The reaction rates within the chemical network are taken from the UMIST 2012 chemical network data base in McElroy et al. (2013). Extinction within the cloud is calculated assuming a grain size of 0.1 μm, albedo of 0.7 and a mean photon scattering by grains of $g = 0.9$. Emission and fine structure lines are calculated using the escape probability method of de Jong, Boland & Dalgarno (1980) and non-LTE level populations determined from the collisional rate coefficients explained in Bisbas et al. (2012).

Moreover, for the H_2 and CO photodissociation rates, the code adopts the treatments of Lee et al. (1996) and van Dishoeck & Black (1988). To account for the shielding of C I, the code uses the treatment of Kamp & Bertoldi (2000) in order to estimate the photoionization rate of carbon. The rate of molecular hydrogen formation on dust grains is calculated using the treatment of Cazaux & Tielens (2004) while the thermally averaged sticking coefficient of hydrogen atoms on dust grains is taken from Hollenbach & McKee (1979). The dust temperature at each point in the density distribution is calculated using the treatment of Hollenbach, Takahashi & Tielens (1991) to account for the grain heating due to the incident far-UV (FUV) photons. Finally, the code also handles varying gas-phase metallicities. The grain surface H_2 formation rate of de Jong (1977) is adopted and scales linearly with metallicity, while the dust-to-gas ratio also scales linearly with metallicity, taking a standard value of 10^{-2} at solar metallicity, following the prescription by Leroy et al. (2011). The dust and polycyclic aromatic hydrocarbon (PAH) photoelectric heating also scales with metallicity.

3 SELF-CONSISTENT COUPLING METHOD

To self-consistently couple all three of the above-mentioned codes, it is important to use as many outputs from one model as inputs for subsequent models, ensuring consistency within the full simulation. This coupling technique was first attempted in Vasta (2010) to model carbon and oxygen emission in nearby galaxies. We describe here how each code is numerically coupled to the other models, with a detailed discussion of the specific input parameters and their values presented in Section 4. Spherical symmetry is assumed throughout this paper, in all phases of the ISM and for all the simulated 3D clouds. Although MOCASSIN and 3D-PDR are both fully capable of handling non-uniform densities, this will not be used here.

3.1 Coupling STARBURST99 to MOCASSIN

A stellar radiation density field, coming from the stellar population within our simulated star-forming regions, is created using SB99. From this output SB99 stellar spectrum, the luminosity, temperature and number of ionizing photons of the source are calculated; these quantities are then used as input parameters for the 3D photoionization code MOCASSIN. In this way, the radiation field is coupled with the photoionization in the H II region.

3.2 Coupling MOCASSIN to 3D-PDR

Henney et al. (2005) showed how dynamical processes, such as gas flows and thermal gas pressure, link the H II and PDR regions, which are simulated here by MOCASSIN and 3D-PDR, respectively. The physical properties of the PDR are a consequence of the transport of gas, dust and radiation through the ionized region, while the converse is also true. Other than in the case of a very fortuitous choice of initial conditions, simply matching boundary conditions between quantities such as temperature and density, of the two regions, can lead to discontinuities in the thermal and dynamical pressure across the two ISM phases. Only by physically coupling the two regions is it possible to get an accurate representation of the multiphase ISM; this modelling philosophy is at the heart of this work as we aim to self-consistently calculate the temperature, ionization and density at the face of PDR regions.

The output of the MOCASSIN code is the SED of the ionized gas, dust and stars emerging from the H II region, along with the flux in the most important FIR fine structure emission lines from the ionized gas, such as [C II]. We calculate the strength of the radiation field, G_0, at the ionization front between the H II and neutral gas regions by integrating the MOCASSIN SED in the FUV range between 912 and 2400 Å, which is the classical Draine field definition (Draine 1978). This value of G_0 is used as an input into 3D-PDR. We also need to ensure that MOCASSIN only simulates the gas up to the edge of the ionized region and does not leak over into the neutral regions (which 3D-PDR will simulate). To this end, we first calculate the outer radius of the ionized cloud simply by running MOCASSIN to a very large radius, and then inspecting at which radius the ionized hydrogen abundance is less than 10 per cent, ensuring that the end of the ionized region has been reached; this is taken to be the outer radius of the ionized part of the star-forming region. The inner radius of the PDR is equal to the outer radius of the ionized region, ensuring that the PDR is adjacent to the ionized region (see Section 4.2 for further details).

We further link the H II and PDR regions by assuming constant total pressure at the interface, ensuring that the temperature and density of the gas and dust between the two regions are self-consistently calculated. Constant pressure is an approximation to the actual flow, which has been assumed by previous authors such as Carral et al. (1994), Abel et al. (2005) and Ferland et al. (2013). Pressure originates from the stellar continuum and internally generated light, gas pressure, and sometimes from turbulence, ram, and magnetic pressures when appropriate (Baldwin et al. 1991; Henney et al. 2005). We are only interested in terms that change across the boundary and hence focus on the gas and dust terms as all others remain constant. The pressure terms of interest originate from internal radiation, from the gas and dust, and the thermodynamic gas pressure. We assume that the gas and dust emission is optically thick as we are only interested in solving the equation at the boundary between the two regions. Therefore, we set up an equation of pressure

balance as follows:

$$n_{\mathrm{H\,II}} k_{\mathrm{b}} T_{\mathrm{H\,II}}^{\mathrm{gas}} + \frac{\epsilon_{\mathrm{gas}}^{\mathrm{H\,II}} \sigma_{\mathrm{b}} T_{\mathrm{H\,II}}^{4,\mathrm{gas}}}{c} + \frac{\epsilon_{\mathrm{dust}}^{\mathrm{H\,II}} \sigma_{\mathrm{b}} T_{\mathrm{H\,II}}^{4,\mathrm{dust}}}{c} = n_{\mathrm{PDR}} k_{\mathrm{b}} T_{\mathrm{PDR}}^{\mathrm{gas}} + \frac{\epsilon_{\mathrm{gas}}^{\mathrm{PDR}} \sigma_{\mathrm{b}} T_{\mathrm{PDR}}^{4,\mathrm{gas}}}{c} + \frac{\epsilon_{\mathrm{dust}}^{\mathrm{PDR}} \sigma_{\mathrm{b}} T_{\mathrm{PDR}}^{4,\mathrm{dust}}}{c}, \quad (1)$$

where $n_{\mathrm{H\,II}}$ and n_{PDR} are the electron number densities in the H II and PDR region, respectively, k_{b} is the Boltzmann constant, σ_{b} is the Stefan Boltzmann constant, while ϵ_{gas} and ϵ_{dust} are the emissivities of the gas and dust species. $T_{\mathrm{H\,II}}^{\mathrm{gas}}$ and $T_{\mathrm{H\,II}}^{\mathrm{dust}}$ are the gas and dust temperatures at the edge of the H II region, as calculated from the MOCASSIN output. Since $n_{\mathrm{H\,II}}$ is one of the input parameters of our code, the above equation needs to be solved for n_{PDR}, $T_{\mathrm{PDR}}^{\mathrm{gas}}$ and $T_{\mathrm{PDR}}^{\mathrm{dust}}$.

Motivated by dust temperature continuity across the two regions obtained in other self-consistent calculations such as CLOUDY (Ferland et al. 2013) and TORUS-3DPDR (Bisbas et al. 2015), we set

$$T_{\mathrm{H\,II}}^{\mathrm{dust}} = T_{\mathrm{PDR}}^{\mathrm{dust}}. \quad (2)$$

We further assume the same dust species in both regions, leading to the cancellation of the terms describing radiation pressure from dust emission in equation (1). The equation of pressure balance solely for the gas remains, such that

$$n_{\mathrm{H\,II}} k_{\mathrm{b}} T_{\mathrm{H\,II}}^{\mathrm{gas}} + \frac{\epsilon_{\mathrm{gas}}^{\mathrm{H\,II}} \sigma_{\mathrm{b}} T_{\mathrm{H\,II}}^{4,\mathrm{gas}}}{c} = n_{\mathrm{PDR}} k_{\mathrm{b}} T_{\mathrm{PDR}}^{\mathrm{gas}} + \frac{\epsilon_{\mathrm{gas}}^{\mathrm{PDR}} \sigma_{\mathrm{b}} T_{\mathrm{PDR}}^{4,\mathrm{gas}}}{c}. \quad (3)$$

Due to the low emissivity of gas, we make the approximation that the radiation pressure caused by photons emitted from the gas can be ignored and so this term is set to zero. Therefore, to set the conditions in the PDR, caused by the ionized region, we are left to solve

$$n_{\mathrm{H\,II}} T_{\mathrm{H\,II}}^{\mathrm{gas}} = n_{\mathrm{PDR}} T_{\mathrm{PDR}}^{\mathrm{gas}}. \quad (4)$$

The temperature at the surface of the PDR is dependent on its hydrogen number density, n_{PDR}; therefore the above equation can be solved using a Newton–Raphson numerical method to provide the value of the hydrogen number density of the PDR given the conditions in the H II region,[1] ensuring self-consistency.

Overall our method leads to dust temperature continuity between the two regions, and also ensures gas pressure equilibrium. The resulting temperature and density profiles of the gas and dust across the two regions are consistent with profiles obtained in CLOUDY (Ferland et al. 2013) and TORUS-3DPDR (Bisbas et al. 2015).

4 MODEL PARAMETER SPACE

The main purpose of this work is to provide a prescription to calculate the fraction of the total integrated [C II] emission of a galaxy emanating from the molecular phase of the ISM, $f_{[\mathrm{C\,II}],\mathrm{mol}}$, from typical extragalactic observables such as stellar mass, specific star formation rate (SSFR) and metallicity. Therefore, we are interested in running our self-consistent modelling interface over an input parameter space corresponding to meaningful observables on galaxy-wide scales. To this end, parameters that do not correspond to galactic observables shall be kept constant and typical values shall be used and taken from the literature. Before presenting the results of the modelling, we describe in this section the parameters chosen for the different codes, and the final seven input parameters that are required to run the full coupled model.

4.1 Stellar population parameters in STARBURST99

We simulate the total stellar SED as originating from a single ionizing source (even though it physically originates from multiple sources) for simplicity and ease. Karczewski et al. (2013) have compared the outputs of MOCASSIN when the ionizing flux is produced by a single source or by 100 sources distributed uniformly within a sphere of 0.2 kpc. They find the results to vary only at small radii; outside of the inner radius, the gas effectively sees a point source. As we are here integrating far out of the cloud, this simplifying assumption will not affect our results.

The specific SED produced by SB99 depends on assumptions made regarding the star formation history, IMF and metallicity of the stellar population. The code allows input stellar metallicities of 0.02, 0.2, 1 or 2.5 $\mathrm{Z}_{\odot}$, and either periods of constant star formation or instantaneous bursts. Although the stellar metallicity is not a variable parameter in this work, it is undoubtedly correlated with the gas-phase metallicity, which will be a variable parameter in our models. Within these constraints, we build the star formation histories of our model galaxies as follows: (1) a first instantaneous burst of star formation at approximately the Hubble time, with the total stellar mass produced in this burst given as an input parameter and a stellar metallicity of 0.02 $\mathrm{Z}_{\odot}$ as the metal content of the early universe is negligible; (2) a period of quiescence followed by a secondary burst, with the age of this second burst another input parameter; and (3) a period of constant star formation until the present day, with this rate of star formation a third input parameter. Due to metal enrichment of the gas from previous supernovae, we set the stellar metallicity of phases (2) and (3) to be the available input parameter greater than the input gas-phase metallicity; e.g. for a gas-phase metallicity of 0.65 $\mathrm{Z}_{\odot}$, we would set the stellar metallicity to 1 $\mathrm{Z}_{\odot}$. The mass lost to supernovae and stellar winds in the initial burst is calculated and fed back into the secondary burst, ergo the input stellar mass parameter for phase (2), ensuring that the total stellar mass of the star-forming region (one of the input parameters) is successfully produced by the present day. We do this in keeping with the current paradigm that star formation is regulated by outflows from stellar winds and supernovae (Davé, Oppenheimer & Finlator 2011; Davé, Finlator & Oppenheimer 2012; Walch 2016).

SB99 only allows the use of a piecewise power-law IMF, and so we build one that closely matches the IMF of Chabrier (2003). We use exponents of 2.3 and 1.3 for IMF boundaries of $1.0 < M_{\mathrm{i}}/\mathrm{M}_{\odot} < 100.0$ and $0.1 < M_{\mathrm{i}}/\mathrm{M}_{\odot} < 1.0$, respectively. The latter boundary exponent choice is equivalent to the IMF of Kroupa (2001), which is approximately equal to that of Chabrier (2003) as noted by Speagle et al. (2014). When running SB99, we use the Padova stellar evolution tracks, detailed in Fagotto et al. (1994), with thermally pulsating asymptotic giant branch stars and Pauldrach and Hiller model atmospheres. Fig. 1 shows an example of a stellar SED produced by SB99 for such a star formation history, with the contribution of the three different phases also shown separately.

4.2 Ionized region parameters in MOCASSIN

The MOCCASIN simulations employ 3D spherically symmetric geometry, with the ionizing source at the centre of the clouds. The inner

[1] We iteratively solve equation (4) up to a 1 per cent accuracy level.

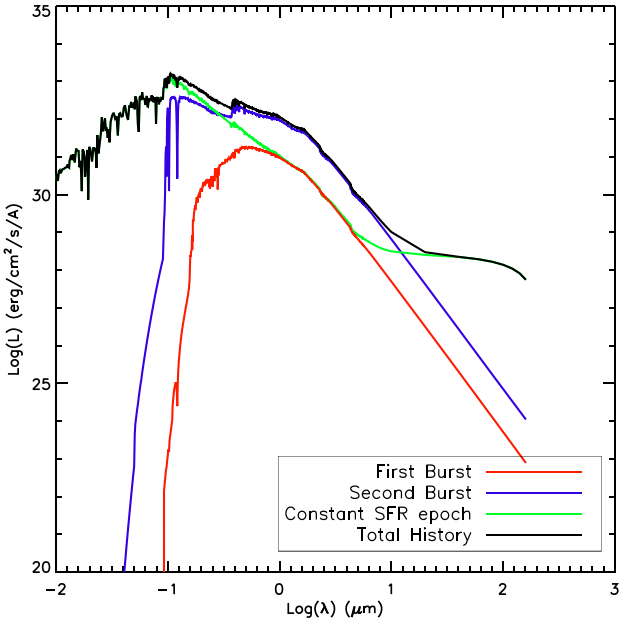

Figure 1. Example stellar SEDs from the population of stars created throughout the whole star formation history of our clouds. The instantaneous burst and constant SFR epoch both compete to dominate the UV part of the spectrum.

radius of the ionized gas region is set to as close to zero as computationally possible. To determine the outer radius of the ionized gas region, we run MOCASSIN up to a large outer radius and calculate the radius at which the ionized hydrogen fraction drops below 10 per cent (this is the effective Strömgren sphere radius). We then take this calculated radius value and re-run MOCASSIN but now setting the outer radius to this calculated radius. Hydrodynamical effects such as turbulence, shocks and magnetic fields are ignored in our simulations.

The input stellar spectrum, the source luminosity and temperature are all taken from SB99. MOCASSIN calculates from them the number of ionizing photons per second, $Q_{\rm PHOT}$. The 3D grid used for the simulations has $15 \times 15 \times 15$ resolution elements so as not to be too computationally expensive, while also ensuring no loss of detail via blending across cells. The two input parameters we are free to vary are the electron number density of the H II region, which takes typical values from $10^{1.5}$ to 10^{3} cm^{-3}, and the gas-phase metallicity, which we vary between 0.2 and 1.1Z$_{\odot}$. The metallicity determines the dust-to-gas ratio input as, to ensure consistency with 3D-PDR, we use the prescription by Leroy et al. (2011, as discussed in Section 2.3). We use different grain properties for H II and PDR regions due to the different physical conditions found in these regions. Within MOCASSIN, we use the standard silicate dust properties detailed in Draine & Lee (1984). For the PDR region, we use a mixture of silicates + PAHS + graphite, due to the higher column densities, with graphite grains being the dominant dust species.

4.3 PDR parameters in 3D-PDR

For 3D-PDR, we consider a spherically symmetric shell of uniform density neutral and molecular gas, surrounding the ionized region. The inner radius of the PDR region is therefore the outer radius of the ionized region as calculated with MOCASSIN, and the outer radius (and corresponding $A_{\rm v}$) is set by the dust mass fraction, $M_{\rm dust}/M_{t}$; we integrate out to a radius that is set by the dust mass budget available. We define the molecular region as the region where more than 1 per cent of hydrogen is in molecular form, marking the beginning of the CO-dark phase. Geometrical dilution effects of the UV field are taken into account to obtain accurate 3D results. We assume a standard turbulent velocity of 1.5 km s^{-1}, while the hydrogen number density is self-consistently calculated[2] (see Section 3.2). Therefore, the two input parameters we are free to vary for the PDR regions are the cosmic ray ionization rate and the dust mass fraction. The gas-phase metallicity is taken to be the same as the value selected for the ionized region. Abundances of all metals scale linearly with metallicity, and Table 1 summarizes the initial chemical abundances used in both MOCASSIN and in 3D-PDR at solar metallicity. We use identical chemical abundances between the codes to maintain self-consistency and take the abundances at solar metallicity from Cormier et al. (2012).

Table 1. Gas-phase elemental abundances used in MOCASSIN and 3D-PDR, relative to total hydrogen number density, at solar metallicity. All these elements, except hydrogen and helium, which are primordial in origin, scale linearly with metallicity.

Species	Gas-phase abundance
He/H	0.1
O/H	4.9×10^{-4}
N/H	6.9×10^{-5}
Ne/H	1.1×10^{-4}
S/H	8.1×10^{-6}
Ar/H	1.9×10^{-6}
C/H	3.6×10^{-4}
Si/H	4.8×10^{-6}
Mg/H	4.0×10^{-5}
Fe/H	3.6×10^{-6}

4.4 Summary of input parameters

Our objective is to provide a prescription for variations of $f_{\rm [C\,II],mol}$ in galaxy-integrated observations that can be applied to unresolved galaxy-wide observations. Therefore, we use as input parameters quantities that are motivated by galaxy-wide observations, where possible. Choices for all the input parameters of our multiphase ISM code are justified here and summarized in Table 2.

(i) *Stellar mass* – We let the stellar mass of our simulated star-forming regions vary from 10^{2} to 10^{4} M$_{\odot}$. These values are typical of star-forming regions within the Milky Way (Wright et al. 2010).

(ii) *Age of the secondary burst* – Since our choice of star formation histories is meant to reproduce a broad range of possible integrated population ages, we choose to probe a wide range for the time since the secondary burst, spanning over 1.5 dex from 10^{2} to $10^{3.5}$ Myr.

(iii) *Specific star formation rate* – Deep multiwavelength extragalactic surveys have revealed a tight correlation between SFR and stellar mass for star-forming galaxies (e.g. Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007). This correlation is well-established in the local universe and up to $z \sim 3$ (e.g. Brinchmann et al. 2004; Peng et al. 2010; Reddy et al. 2012). We choose SSFRs in the range of $10^{-11.5}$–$10^{-9.5}$ yr^{-1}, typical of main-sequence galaxies in the local universe, using the stellar mass of the star-forming regions for normalization.

[2] This turbulence contributes a negligible amount to the total pressure between the two regions, and hence is excluded in equation (1).

Table 2. Variable input parameters used in the fully coupled multiphase code.

Input parameter	Minimum value	Maximum value	Number of Variations	Model
Gas-phase metallicity	0.20 $Z_\odot$	1.1 $Z_\odot$	4	MOCASSIN, 3D-PDR, SB99
Stellar mass of the cloud	10^2 $M_\odot$	10^4 $M_\odot$	3	SB99
Stellar population age	10^2 Myr	$10^{3.0}$ Myr	3	SB99
H II region electron number density	$10^{1.5}$ cm^{-3}	$10^{3.0}$ cm^{-3}	4	MOCASSIN
Cosmic ray ionization rate	10^{-17} s^{-1}	10^{-14} s^{-1}	4	3D-PDR
Dust mass fraction	10^{-4}	10^{-2}	5	3D-PDR
Specific star formation rate	$10^{-11.5}$ yr^{-1}	$10^{-9.5}$ yr^{-1}	3	SB99

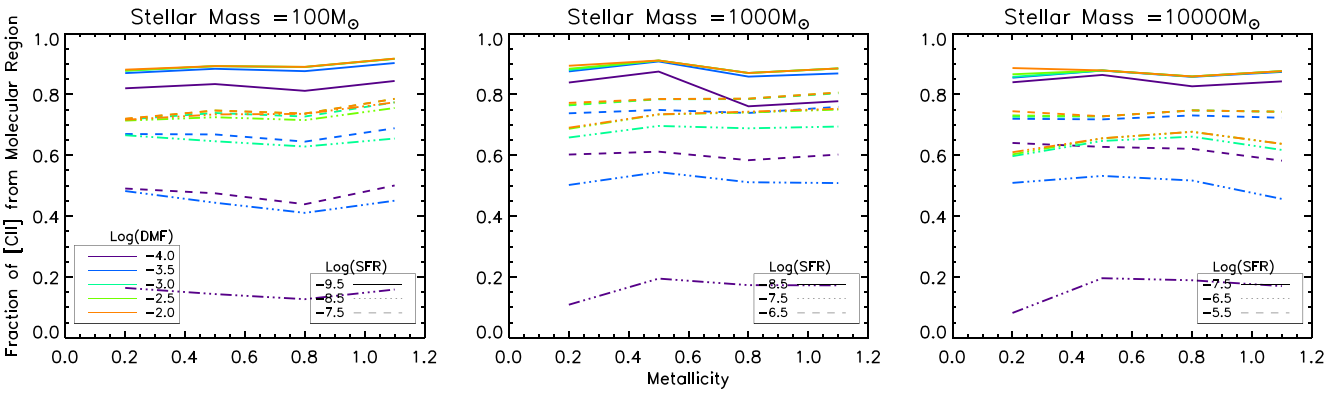

Figure 2. Variations of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ as a function of metallicity for three different stellar mass bins. In each panel, the relation is shown for different values of the dust mass fraction (different colours) and SFR (different linestyles). The hydrogen number density, cosmic ray ionisation rate and age of the secondary burst of star formation are kept constant.

(iv) *Gas-phase metallicity* – We use the mass–metallicity relation from Tremonti et al. (2004) to guide this choice, and adopt a metallicity range of 0.2–1.1 $Z_\odot$ to reproduce conditions in local universe star-forming galaxies with $M_* > 10^9$ $M_\odot$.

(v) *Electron number density of the H II region* – We choose to vary the hydrogen number density between $10^{1.5}$ and $10^{3.0}$ cm^{-3} based on the values calculated by Hunt & Hirashita (2009) for extragalactic H II regions.

(vi) *Cosmic ray ionization rate* – In the local universe, this is known to be roughly 10^{-17}–10^{-16} s^{-1} (Dalgarno 2006; Cummings et al. 2015), but values can be larger by up to three orders of magnitude in galaxies with very large SFRs such as local ULIRGs and high-redshift star-forming galaxies (Papadopoulos 2010). We therefore explore a range of cosmic ray ionization rates ranging from 10^{-17} to 10^{-14} s^{-1} to allow us to also explore conditions typical of $z \sim 2$ galaxies.

(vii) *Dust mass fraction* – We run models with the dust mass fraction (M_{dust}/M_t) varying between 10^{-4} and 10^{-2} based on the scaling relation between dust mass fraction and stellar mass derived from the galaxies in the Herschel Reference Survey (HRS; Boselli et al. 2010).

5 NUMERICAL RESULTS

To cover the full parameter space summarized in Table 2, we simulated a total of 8640 individual star-forming regions. Of those, 8016 clouds fully converged to a solution achieving thermal equilibrium. The main quantity of interest in the context of this study is $f_{[\mathrm{C\,II}],\mathrm{mol}}$, the fraction of the total [C II] emission originating from the molecular regions. A quantitative analysis of the data requires a multidimensional hierarchical Bayesian inference method, which will be performed in Section 6.1. For now, we qualitatively investigate the dependence of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ on the input parameters of the model and explain the physics behind the trends that emerge.

5.1 Stellar mass and star formation rate

We first show in Fig. 2 how $f_{[\mathrm{C\,II}],\mathrm{mol}}$ varies as a function of metallicity for three different values of stellar mass. At fixed mass and metallicity, the figure also shows the impact of a varying dust mass fraction and SFR. In this example, the hydrogen number density, the cosmic ray ionization rate and the age of the secondary burst of star formation are kept fixed. The figure shows that $f_{[\mathrm{C\,II}],\mathrm{mol}}$ does not vary significantly with either stellar mass or metallicity. This is as expected because, as the stellar mass decreases, the number of ionizing photons also decreases, reducing the overall size of the cloud. However, the relative sizes and densities of the H II to PDR regions will not change. By scaling down the stellar mass at the centre of each cloud, we have simply scaled down the size of the cloud while maintaining the same physical structure throughout each cloud.

Fig. 2 also shows that at fixed stellar mass, metallicity and dust mass fraction, an increase in SFR corresponds to a decrease in $f_{[\mathrm{C\,II}],\mathrm{mol}}$. This is because, at fixed dust mass, an increase in the SFR leads to an increase in the radius of the H II regions as more photoionizing UV photons are available. More [C II] will therefore arise from the ionized regions versus the molecular regions, and hence $f_{[\mathrm{C\,II}],\mathrm{mol}}$ decreases.

5.2 Age of secondary burst

The variations caused by the age of the secondary burst are of a similar nature to the SFR. Star formation histories that involve a younger secondary burst provide more photoionizing UV photons. Hence, the younger the age of the secondary burst, the more [C II]

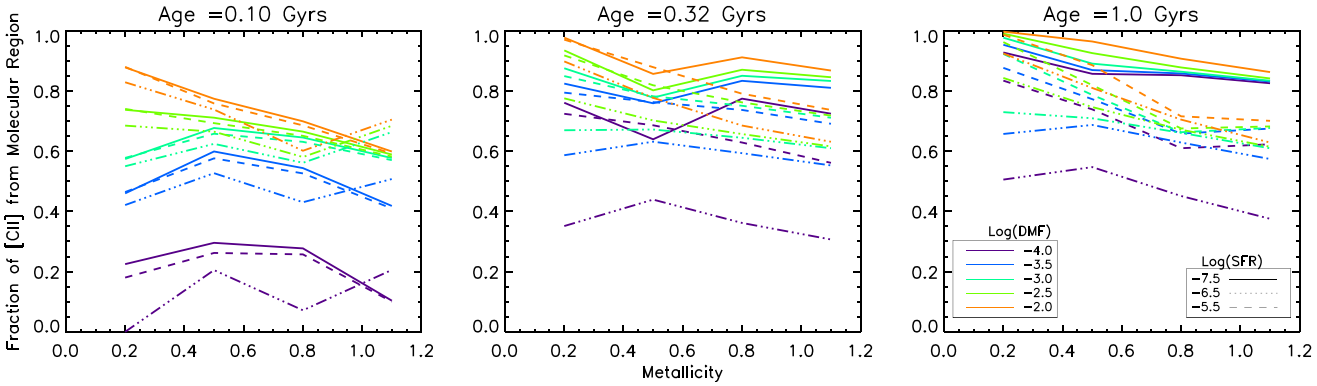

Figure 3. Variations of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ as a function of metallicity for three different ages for the secondary burst of star formation. In each panel, the relation is shown for different values of the dust mass fraction (different colours) and SFR (different linestyles). The hydrogen number density, cosmic ray ionization rate and stellar mass are kept constant.

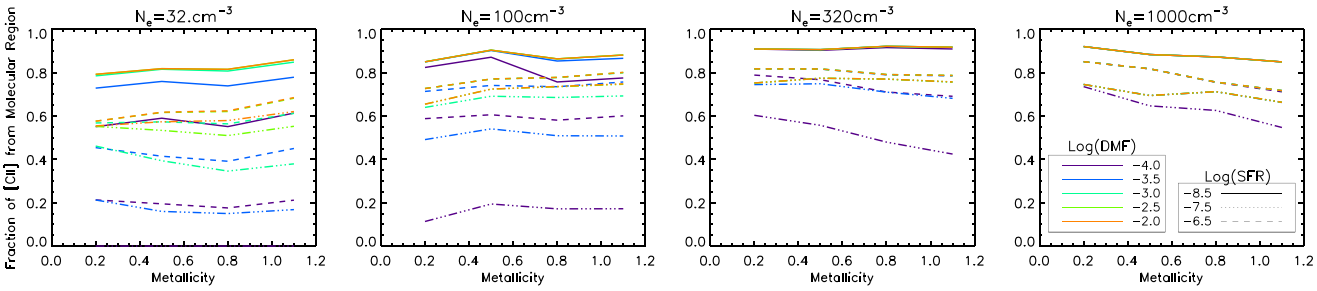

Figure 4. Variations of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ as a function of metallicity for four different density bins. In each panel, the relation is shown for different values of the dust mass fraction (different colours) and SFR (different linestyles). The stellar mass, cosmic ray ionization rate and age of the secondary burst of star formation are kept constant.

will emerge from the H II regions, as the Strömgren sphere radius increases. This can be seen in Fig. 3, with the mean value of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ increasing as the age of the secondary burst increases from 0.1 to 0.32 to 1.0 Gyr. This effect is less pronounced than that caused by variations in SFR, because the majority of UV photons are produced by the low-level star formation happening at the present time rather than by the secondary burst (see Fig. 1). Interestingly though, the time since the last burst of star formation none the less has a detectable effect on $f_{[\mathrm{C\,II}],\mathrm{mol}}$ variations.

5.3 Gas-phase metallicity

Metallicity is responsible for variations in a more complex manner, with two main effects competing for dominance. To investigate these two processes, we refer to Figs 2 and 3.

One could naively expect that by decreasing the amount of metals available throughout the whole system and that the abundance of carbon in the ionized, neutral and molecular regions would decrease in equal measure and hence no variations of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ should be seen because of metallicity (a similar argument to the lack of variations caused by stellar mass). However, within the ionized regions, the cooling rate is a function of metallicity; a decrease in metallicity leads to a lower cooling rate, and therefore, an increase in the size of the Strömgren sphere. Hence, from this first effect (the Cooling Rate effect, hereafter), we can expect that by decreasing metallicity there will follow a decrease of $f_{[\mathrm{C\,II}],\mathrm{mol}}$.

However, a second, more dominant effect, is the well-known photodissociation of CO into ionized carbon (The Photodissociation effect, hereafter). In low-metallicity environments, FUV radiation penetrates further into the clouds leading to an enhanced abundance of ionized carbon in the molecular regions. The Photodissociation effect therefore has the opposite effect of increasing $f_{[\mathrm{C\,II}],\mathrm{mol}}$ as metallicity decreases.

In different parts of the parameter space, the Cooling Rate and the Photodissociation effects cancel each other out, leading to negligible variations of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ as a function of metallicity, as seen in Figs 2 and 3. Under other circumstances, the Cooling Rate or the Photodissociation effect dominates, leading to positive or negative slopes in the $f_{[\mathrm{C\,II}],\mathrm{mol}}$–$Z$ relation, respectively.

5.4 Electron number density of the ionized region

An increase in the density of the ionized region (n_{e}) leads to no change in the [C II] emission from this region as we have already reached the critical density for collisions with electrons, of $\sim$50 cm^{-3} (Goldsmith et al. 2012), which dominate in the H II region. However, due to equilibrium, this leads to an increase in the density of the PDR allowing for an increase in the [C II] emission from the molecular region, and correspondingly, an increase of $f_{[\mathrm{C\,II}],\mathrm{mol}}$. This continues until we reach the critical density for collisions with hydrogen in the PDR, of $\sim 10^{3.5}$ cm^{-3} (Goldsmith et al. 2012), which dominate in the neutral ISM phases, at which point $f_{[\mathrm{C\,II}],\mathrm{mol}}$ remains roughly constant. We can see these variations caused by the electron number density in the ionized regions in Fig. 4. We keep the cosmic ray ionization rate, stellar mass and age of the secondary burst constant and see that increasing n_{e} leads to an increase of $f_{[\mathrm{C\,II}],\mathrm{mol}}$, in all dust mass fraction and SFR bins.

The variations caused by density are linked with those of metallicity. In lower density environments, which have larger ionized

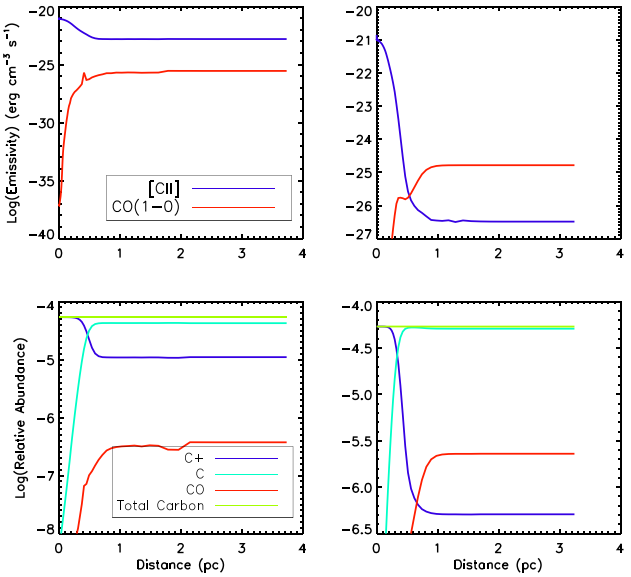

Figure 5. Examples of a warm (left) and cold (right) clouds. The top row shows for each example cloud how the emissivity of the different carbon phases varies with radius in the cloud, while the bottom two shows the relative abundance of these species. In warm clouds, even within PDRs, the emission of ionized carbon always remains larger than that of CO(1–0), while in cold clouds the CO(1–0) molecular phase dominates.

regions, the metallicity variations are either flat or have a positive gradient implying that the Cooling Rate effect, detailed above, is more dominant. However, as density increases, the slope of the $f_{[\mathrm{C\,II}],\mathrm{mol}}$–$Z$ relation changes as the Photodissociation effect begins to dominate. Understanding quantitatively how, and when, these effects dominate follows in Section 6.1.

5.5 Cosmic ray ionization rate

The cosmic ray ionization rate input parameter is used only in the molecular and neutral regions (i.e. the PDR) where the heating function increases with the cosmic ray ionization rate. This creates higher temperatures in the PDR regions, which leads an increase in the [C II] PDR emission, as this line is a major coolant of the gas, leading to an increase in $f_{[\mathrm{C\,II}],\mathrm{mol}}$.

5.6 Dust mass fraction

As discussed above, the dust mass fraction effectively controls the total size of our clouds and determines how far into the PDR we integrate up to. In Fig. 5, we show for two different example clouds how the emissivity and carbonaceous species abundances relative to hydrogen vary as a function of the radius in the neutral and molecular regions. Under certain conditions, 'warm clouds' are simulated such that the temperature of the PDR never reaches 10 K, the freeze-out temperature of hydrogen on to dust grains. This occurs for clouds with low PDR densities, high cosmic rays ionization rates and high gas-phase metallicities. In these warm clouds, the emissivity of [C II] always remains larger than that of CO (top-left panel). This is because of the warmer conditions, and because the relative abundance of ionized carbon is also always larger than its molecular counterpart (bottom-left panel). Therefore, by increasing the dust mass fraction, we are able to retrieve more [C II] the deeper we integrate, and so $f_{[\mathrm{C\,II}],\mathrm{mol}}$ increases with the dust mass fraction.

However, in 'colder' conditions, where the temperature of the PDR eventually reaches 10 K, the emissivity of ionized carbon decreases deep into the molecular regions, where molecular emission begins to dominate (top-right panel). Therefore, increasing the dust mass fraction (i.e. integrating further into the cloud) does not affect the relative emission of [C II]. This can also be seen in the abundance profile (bottom-right panel), where the molecular carbon abundance now dominates deep into the clouds. Hence, in cases such as these, increasing the dust mass fraction will only increase $f_{[\mathrm{C\,II}],\mathrm{mol}}$ up to a point before no more [C II] is obtained and increasing dust mass fraction makes no difference.

We find that 6513, out of our 8016, clouds harbour these 'warmer' conditions with [C II] emission dominating CO(1–0). The emergence of these two groups of clouds has been physically observed in the Galactic Plane (Langer et al. 2014), where 557/1804 clouds observed were detected in [C II] with no CO. Our fraction of warm clouds to cold clouds defer to the observations because our parameter space is not representative of the Galactic plane, as explained above.

6 APPLICATIONS TO GALAXY-WIDE OBSERVATIONS

The qualitative discussion in Section 5 was sufficient to understand the physics underpinning the variations of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ between our different simulated star-forming regions. In this section, we make the jump from these individual star-forming regions to the ISM of entire galaxies. Ideally, we would want to build a model for the ISM of a whole galaxy by appropriately summing up a number of our individual simulated clouds. To do this, we could start from observations of the molecular cloud mass function (e.g. Wright et al. 2010; Wong et al. 2011; Colombo et al. 2014; Gusev et al. 2016); however, it is still highly debated whether there is a universal cloud mass function that is applicable to all galaxies, or whether the properties of clouds depend on other global physical parameters and therefore vary from galaxy to galaxy (see, e.g. Hughes et al. 2013). Given this uncertainty and as a first step, we propose here a simpler alternative method to predict how $f_{[\mathrm{C\,II}],\mathrm{mol}}$ varies as a function of integrated galaxy properties using our simulated clouds. We make the assumption that the physical conditions found in each of our clouds, for a given set of input parameters, can represent the average physical conditions found on galaxy-wide scales for galaxies with similar physical properties. Under this assumption, a whole galaxy can be considered to be built up from an appropriate number of identical star-forming regions.

6.1 Bayesian inference

We now want (a) to parametrize an analytic prescription for how $f_{[\mathrm{C\,II}],\mathrm{mol}}$ varies as a function of our model parameters for extragalactic observations on galaxy-wide scales, and (b) to determine the minimum number of parameters needed to provide a statistically robust fit to our data. We therefore use a Bayesian inference method to find the best-fitting relations and the minimum number of parameters required. Bayesian inference fitting methods have been successfully employed in several, wide ranging, astrophysical scenarios from the derivation of the extinction law in the Perseus molecular cloud (Foster et al. 2013) and Type Ia supernova light-curve analysis (Mandel, Narayan & Kirshner 2011) to the extragalactic Kennicutt–Schmidt relation (Shetty, Kelly & Bigiel 2013) and the formation and evolution of interstellar ice (Makrymallis &

Viti 2014). For a more in-depth explanation of the Bayesian regression fitting method, we refer the reader to Kelly (2007) and restrict ourselves here to the basic concepts. Our 3D radiative transfer methodology provides a complete model for how [C II] varies as a function of the seven input parameters of the coupled code. However, the radiative transfer modelling is highly non-linear and complex, so we explore how well a polynomial fit can describe the outputs from the coupled 3D radiative transfer simulation, and what is the optimal number of parameters for this fit. This is done by evaluating the posterior probability of the simulated data, $y_{\rm RT\ Model}$, given the polynomial fit, denoted $y_{\rm QF}$. We assume that the measurement uncertainties, associated with each of our fits, are normally distributed; therefore, $y_{{\rm QF},i}$ is a random variable distributed as

$$y_{{\rm QF},i} = \mathcal{N}(y_{{\rm RT\ Model},i}, \sigma^2_{y_{{\rm QF},i}}), \tag{5}$$

where $\sigma_{y_{{\rm QF},i}}$ is the measurement uncertainty associated with the polynomial fit $y_{{\rm QF},i}$ on the ith model, which will be an additional parameter that we need to fit. For simplicity, we assume that all $\sigma_{y_{{\rm QF},i}}$ are equal to the same value, $\sigma_{y_{\rm QF}}$.

Under the assumption of the normal distribution in equation (5), the probability of obtaining a certain polynomial fit, given the output of the numerical modelling, combined with the fitted uncertainties and the weighting factors is

$$P(y_{{\rm QF},i}|y_{{\rm RT\ Model},i}, \sigma_{y_{\rm QF}}) = \frac{\sqrt{g_i}}{\sqrt{2\pi\sigma^2_{y_{\rm QF}}}} \times \exp\left(-\frac{g_i(y_{{\rm RT\ Model},i} - y_{{\rm QF},i})^2}{2\sigma^2_{y_{\rm QF}}}\right), \tag{6}$$

where g_i is the dimensionless statistical weighting for each cloud. As described in Section 6.3, weights are assigned to each of the simulated clouds to take into account how likely they are to reproduce ISM conditions typical of local galaxies.

The next assumption to make is that all our radiative transfer-simulated data points are independent, which is perfectly reasonable as we ran through each point in parameter space regardless of the other parameters. Under this assumption, all the individual probabilities can be multiplied to produce the likelihood. By taking the log-likelihood, the product returns back to a sum, so

$$\mathcal{L} = -\frac{N}{2}\ln(2\pi) - N\ln(\sigma_{y_{\rm QF}}) - \sum_{i=1}^{N}\left(\frac{g_i(y_{{\rm RT\ Model},i} - y_{{\rm QF},i})^2}{2\sigma^2_{y_{\rm QF}}}\right) + \sum_{i=1}^{N}\left(\frac{\ln g_i}{2}\right). \tag{7}$$

. Maximizing this log-likelihood for the polynomial fit parameters and the associated error, $\sigma_{y_{\rm QF}}$, will provide us with the best-fitting analytical expression alongside the 1σ error of the parametrization.

To compare likelihoods from models with different numbers of free parameters, we use two different methodologies. First, we employ the Akaike Information Criterion (AIC; Akaike 1981):

$${\rm AIC} = -2\mathcal{L} + 2p + \frac{2p(p+1)}{N-p-1}, \tag{8}$$

where p is the number of free parameters and N is the sample size. The best model and the optimal number of free parameter is found by minimizing the AIC. We also calculate the Bayesian Information Criterion (BIC; Schwarz 1978):

$$BIC = -2\mathcal{L} + p\log(N) \tag{9}$$

and compare the results of both tests to ensure that our results are not dependent on the choice of the information criterion used.

6.2 Sampling methods and quadratic models

A direct solution for the posterior probability distribution is computationally expensive and so, to efficiently and effectively sample the full parameter space, we use the well-tested PYTHON implementation of the affine-invariant ensemble sampler for Markov chain Monte Carlo called EMCEE[3] (Goodman & Weare 2010).

Given the saturation effect that may occur when [C II] is mainly emitted from the molecular regions (i.e. when $f_{\rm [C\,II],mol}$ approaches 1) and the low number of bins in our parameter space, we only fit quadratic polynomials to our data, including all second-order cross-terms when multiple parameters are involved, e.g. for three parameters we would use

$$y_{{\rm QF},i} = \alpha_1 + \alpha_2 x_{1,i} + \alpha_3 x^2_{1,i} + \alpha_4 x_{2,i} + \alpha_5 x^2_{2,i} + \alpha_6 x_{3,i} + \alpha_7 x^2_{3,i} + \alpha_8 x_{1,i}x_{2,i} + \alpha_9 x_{1,i}x_{3,i} + \alpha_{10} x_{2,i}x_{3,i}. \tag{10}$$

We also fit for $\sigma_{y_{\rm QF}}$ and therefore have $2(\epsilon+1) + {}^{\epsilon}C_2$ free parameters to constrain, where ϵ is the number of different variables in our fits.[4] This number can range from 1 to 4 as we focus on the four input parameters of the coupled code, which are also commonly available extragalactic observables. These are the gas-phase metallicity, the electron number density of the H II regions, the SSFR and the dust mass fraction.

6.3 Statistical weighting calculation

The simulated clouds, which we now assume represent average physical conditions on galaxy-wide scales, fill up a very large parameter space representing a large range of possible physical conditions. In which parts of this parameter space do galaxies actually lie? Which simulated clouds therefore represent average physical conditions in local galaxies? To account for this, we calculate a weighting factor for each cloud based on how likely it is to be representative of a local galaxy.

To determine these weighting factors, we make use of the HRS (Cortese et al. 2012), a statistically complete K-band-selected sample of galaxies located between 15 and 25 Mpc (Boselli et al. 2010). We retrieve dust masses, stellar masses and SFRs from HRS catalogues (Boselli et al. 2013; Cortese et al. 2014; Boselli et al. 2015). From these, we can directly infer for 112 HRS galaxies two of the input parameters of our coupled model: dust mass fractions and SSFRs. Another input parameter, the electron number densities in the ionized regions, is calculated from the [S II] line intensity ratio $R = [{\rm S\,II}]\lambda6716/\lambda6731$ using the prescription of Sanders et al. (2016):

$$n_{\rm e} = \sqrt{T_{\rm e}}\left(\frac{1.4498 - R}{0.1595R - 0.0688}\right), \tag{11}$$

where $T_{\rm e}$ is the electron temperature and assumed to be a standard 10^4 K, typical for H II regions. Similar temperature assumptions have been made previously in Ho et al. (2014) and Sanders et al. (2016). These electron densities are equivalent to hydrogen number densities in the H II regions in the range of 10–1000 cm^{-3}. Comparison with the input parameters for the coupled multiphase code given in Table 2 confirms that we have sampled the appropriate ranges to reproduce conditions typical of local galaxies.

[3] An example of the code can be found at http://dan.iel.fm/emcee/current/

[4] This applies when ϵ is greater or equal to 2. For one variable we have four free parameters.

We bin the HRS data to match the sampling used in the radiative transfer modelling (shown in Table. 2), i.e. for the above three parameters (dust mass fractions, SSFR and n_e) we bin the 112 HRS galaxies into 60 bins ($5 \times 4 \times 3$). It is possible to include metallicity in the binning; however, this would restrict and reduce the sample size further. If we did include metallicity, our sample would now shrink to 69 objects, and we would now have 120 bins, meaning that our weighting function would be comb-like leading to erroneous results as the number of bins exceed the sample size. A variant of the Freedman & Diaconis (1981) rule states that the number of bins must be less than the sample size, which is why we do not include metallicity into the weighting. We use the PYTHON *N*-dimensional histogram routine, *histogramdd*, to calculate the normalized weighting for each of our simulated clouds.

From here on, we limit our sample to clouds with cosmic ray ionization rates equal to the average Milky Way value (10^{-17} s^{-1}), as it is unlikely to vary much from this value for any of the HRS galaxies, which are local normal star-forming and quiescent galaxies. Higher cosmic ray rates ($\sim 10^3 \times$ Milky Way value) are found in ULIRGs and galaxies with more enhanced star formation (Kelly et al. 2015); hence, we keep this fixed at the average Milky Way value for now. Even though the HRS sample does not have measured cosmic ray ionization rates, based on their position in the SFR–M_* plane, and the fact that they are local galaxies, we can be sure that they all have a value approximately equal to that of the average Milky Way cosmic ray ionization rate. Therefore, we set to 0 the statistical weight of any cloud simulated with a cosmic ray ionization rate higher than that of the Milky Way.

We also present in Appendix A the results if we limit our sample to clouds with cosmic ray ionization rates equal to ten times the average Milky Way value (10^{-16} s^{-1}). We find identical results when using the two different cosmic ray ionization rates; using the average Milky Way value (10^{-17} s^{-1}) does not affect our results for local universe galaxies.

6.4 Statistical results

Using the Bayesian formalism and statistical weights described above, we fit the simulated values of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ as a function of the four key observables (density, dust mass fraction, SSFR and metallicity), allowing the number of these parameters used in any one fit to vary between 1 and 4. As the number of free parameters increases, the quality of the fit improves as shown by both the AIC and BIC (Fig. 6). We hereby present several novel prescriptions for the fraction of [C II] emission emerging from molecular regions on galaxy-wide scales simply involving dust mass fraction, H II region electron number density, SSFR and metallicity. The full analytical prescription, according to the AIC and BIC, is the one involving all four galaxy parameters, namely

$$\begin{aligned} f_{[\mathrm{C\,II}],\mathrm{mol}} = {} & -4.405 + 0.133\frac{Z}{Z_\odot} - 0.172\frac{Z}{Z_\odot}^2 + 1.448\rho \\ & - 0.206\rho^2 + 0.814\phi - 0.050\phi^2 - 0.818\psi - 0.032\psi^2 \\ & - 0.063\frac{Z}{Z_\odot}\rho + 0.003\frac{Z}{Z_\odot}\phi - 0.027\frac{Z}{Z_\odot}\psi \\ & - 0.222\rho\phi + 0.098\rho\psi + 0.050\phi\psi, \end{aligned} \quad (12)$$

where $\frac{Z}{Z_\odot}$ is the metallicity, $\rho = \log n_e$, $\phi = \log \frac{M_{\mathrm{dust}}}{M_*}$ and $\psi = \log(\mathrm{SSFR})$. The 1σ error derived from the fitting is $\sigma_{f_{[\mathrm{C\,II}],\mathrm{mol}}} = 0.0597$ (a unit-less quantity as it is a relative fraction). Furthermore, we also present the best three-, two- and one-parameter prescriptions as we understand acquiring all the necessary data to use equation (12) may be a challenge. The three-parameter prescription includes only dust mass fraction, n_e and SSFR:

$$\begin{aligned} f_{[\mathrm{C\,II}],\mathrm{mol}} = {} & -3.92 + 1.50\rho - 0.209\rho^2 + 0.471\phi - 0.072\phi^2 \\ & - 0.628\psi - 0.018\psi^2 - 0.227\rho\phi + 0.106\rho\psi + 0.027\phi\psi. \end{aligned} \quad (13)$$

The 1σ error in this case is $\sigma_{f_{[\mathrm{C\,II}],\mathrm{mol}}} = 0.061$. The two-parameter prescription does away with the dust mass fraction and therefore simplifies as

$$\begin{aligned} f_{[\mathrm{C\,II}],\mathrm{mol}} = {} & -5.63 + 1.31\rho - 0.17\rho^2 - 0.87\psi \\ & - 0.034\psi^2 + 0.046\psi\rho \end{aligned} \quad (14)$$

with an error of $\sigma_{f_{[\mathrm{C\,II}],\mathrm{mol}}} = 0.064$. The best-fitting one-parameter prescription involves only n_e and has $\sigma_{f_{[\mathrm{C\,II}],\mathrm{mol}}} = 0.069$:

$$f_{[\mathrm{C\,II}],\mathrm{mol}} = -0.556 + 1.087\rho - 0.219\rho^2. \quad (15)$$

As we will show in the next section, there is also a trend between $f_{[\mathrm{C\,II}],\mathrm{mol}}$ and SSFR. As this latter quantity is typically more readily available to extragalactic observers than n_e, we also perform this one-parameter fit even though it is not formally selected by the AIC and BIC. This alternative one-parameter prescription, with an associated error of $\sigma_{f_{[\mathrm{C\,II}],\mathrm{mol}}} = 0.072$, is

$$f_{[\mathrm{C\,II}],\mathrm{mol}} = -6.224 - 1.235\psi - 0.0543\psi^2. \quad (16)$$

In the following section, we test and compare these five prescriptions, and then advise on the best relation to use to estimate $f_{[\mathrm{C\,II}],\mathrm{mol}}$ for individual galaxies in Section 7.

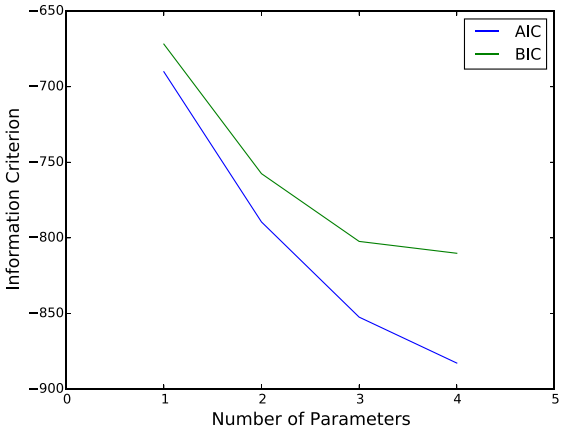

Figure 6. We provide a plot for the variation of the Akaike Information Criterion and the Bayesian Information Criteria in blue and green, respectively. It can be seen how, although they give different absolute numerical values (due to their different analytic expressions), they reach a minimum at four parameters.

6.5 Validation of the $f_{[\mathrm{C\,II}],\mathrm{mol}}$ prescriptions and example applications

Measurement of the [C II] fraction emerging from molecular regions in extragalactic objects are uncommon at best, making validating our prescription for $f_{[\mathrm{C\,II}],\mathrm{mol}}$ against a large and complete galaxy sample impossible. However, Pineda et al. (2013) have measured $f_{[\mathrm{C\,II}],\mathrm{mol}}$ across the Milky Way as part of the GOT C$^+$ survey. Assuming a dust mass for the Milky Way of $10^{7.7}$ M$_\odot$ (Pierce-Price et al. 2000), an SFR of 1.65 M$_\odot$ yr^{-1} and a total stellar mass of

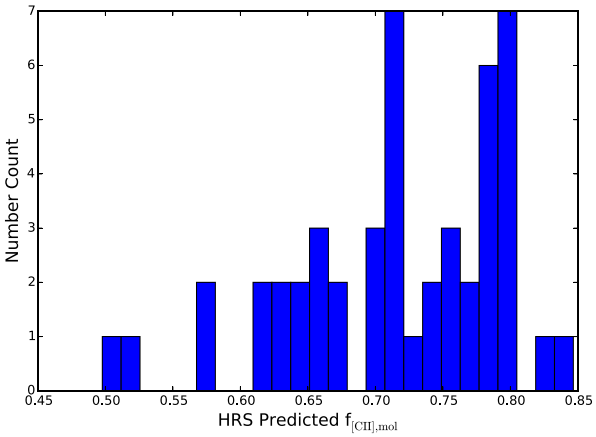

Figure 7. By applying our prescription to the HRS sample, we find that the majority of the galaxies have 60–80 per cent of their total integrated [C II] emission arising from molecular regions. Due to the completeness of the sample it implies that on galaxy-wide scales, in the local universe, 60–80 per cent of a galaxy's [C II] emission will originate from molecular regions.

$10^{10.78}$ $M_\odot$ (Licquia & Newman 2014), a metallicity of 1 $Z_\odot$ and finally an electron number density of 100 cm^{-3}, equation (12) predicts $f_{[C\,II],mol} = 75.84 \pm 5.97$ per cent for the Milky Way. As a comparison, the alternative prescriptions evoking fewer input parameters, equations (16), (15), (14) and (13) predict values of 77.6 $\pm$ 6.3 per cent, 74.42 $\pm$ 6.94 per cent, 74.42 $\pm$ 6.94 per cent and 73.80 $\pm$ 6.10 per cent, respectively. All of these predictions are in excellent agreement with the measured value of 75 per cent (Pineda et al. 2013).

Extragalactic observations have been done, however, which accurately measure the fraction of [C II] emerging from *ionized* gas regions, using the [C II]/[N II]205 μm and [N II]122 μm/[N II]205 μm ratios (Oberst et al. 2006). This fraction has been measured to be between 15 per cent and 65 per cent in NGC 891 (Hughes et al. 2015) and between 20 per cent and 30 per cent in the star-forming region BCLMP 302 of M33 (Mookerjea et al. 2011). Our ISM model here is unable to measure exactly the [C II] fraction arising from similar ionized regions; it would have to be modified to produce the emissivity profiles across the ionized and neutral phases to discriminate the origin of [C II] between these two phases of the ISM. We can however provide an upper limit for the fraction of [C II] emerging from the ionized regions as $1 - f_{[C\,II],mol} \simeq$ 20–40 per cent, in agreement with these observations.

We also apply our prescription to the HRS galaxies that have measurements available for all four physical parameters going into equation (12), and find that the typical value of $f_{[C\,II],mol}$ for these representative local galaxies is 60–80 per cent, shown in Fig. 7. This agrees well with Olsen et al. (2016), who also found [C II] emission to be dominated by the molecular gas. Furthermore, Fig. 8 shows how these values of $f_{[C\,II],mol}$ depend on key parameters. In

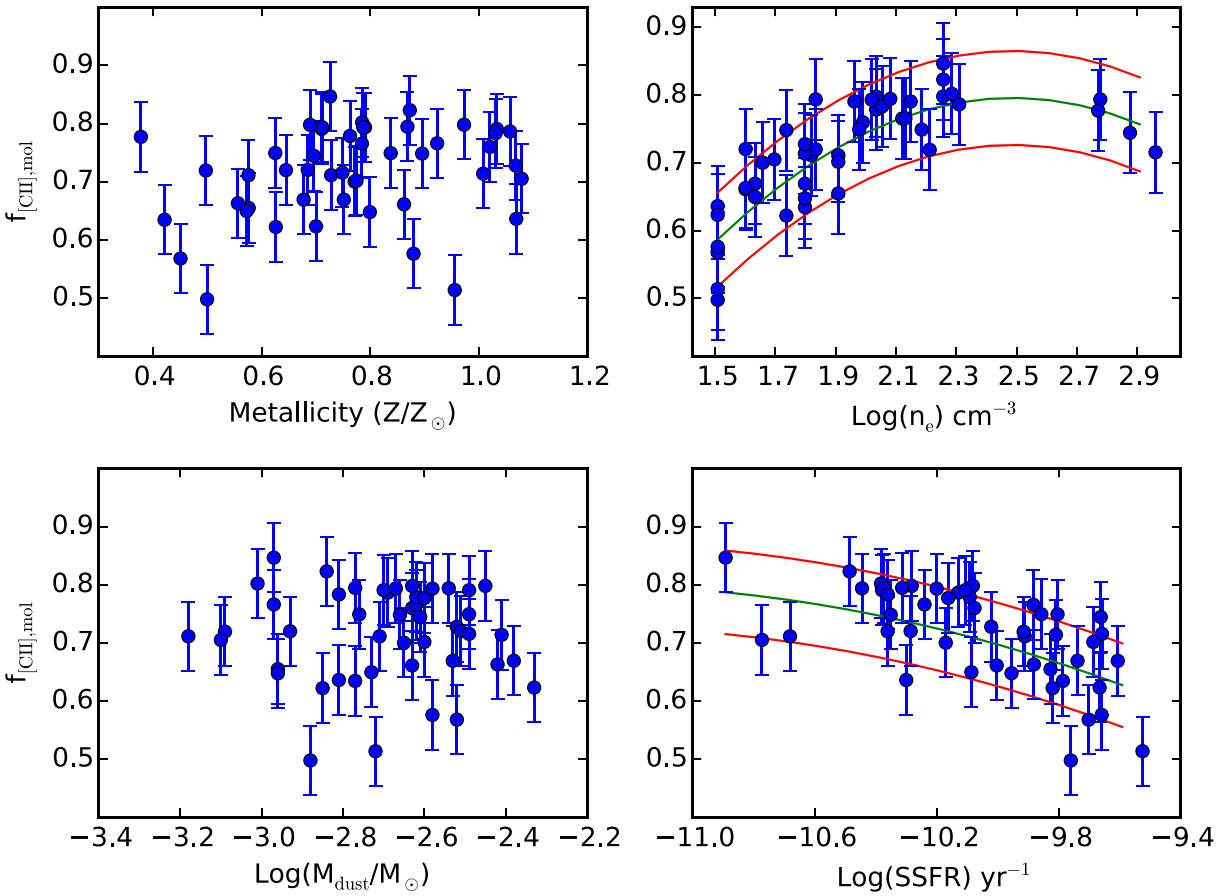

Figure 8. Relation between $f_{[C\,II],mol}$ as calculated from equation (12) for galaxies from the HRS sample that have four key integrated properties (metallicity, density, dust mass fraction and SSFR). We also overlay the single prescriptions for density and SSFR, equations (15) and (16), respectively, in the upper-right and bottom-right panels by the green lines. The red lines represent the 1σ errors on both prescriptions.

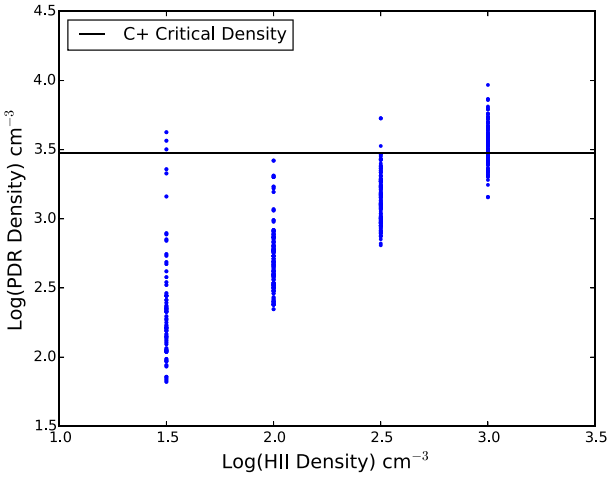

Figure 9. We plot the input hydrogen number density (equivalent to n_e) used in ionized regions against the calculated PDR number densities. As can be seen for an H II density of 10^3 cm^{-3} the PDR density starts to exceed the critical density of [C II].

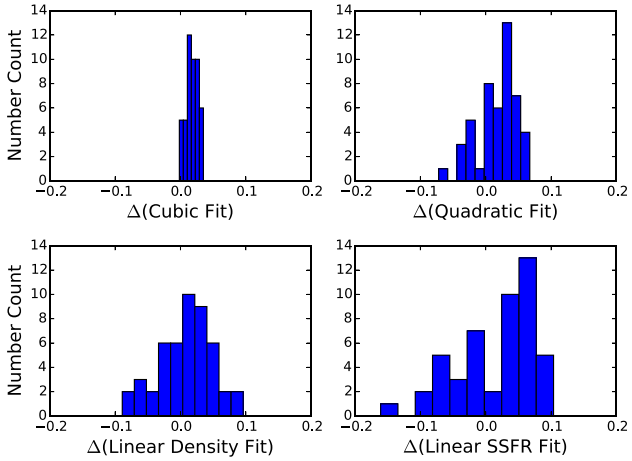

Figure 10. Distribution of the offset between the best prediction of $f_{[C\,II],mol}$ from equation (12) for the HRS galaxies, and the values produced by the four alternative prescriptions (equations 13–16).

the model grid (Table 2), all the parameters were varied independently, without enforcing any correlations between each of them. However, observations of star-forming regions and local galaxies make it clear that many of these physical properties are highly correlated. The scaling relations of Fig. 8 therefore implicitly contain these physical correlations, and interestingly show no trend with metallicity and dust mass fraction. Correlations are seen however with SSFR and n_e with explanations similar to those in Sections 5.1 and 5.4. The initial increase in $f_{[C\,II],mol}$ with n_e occurs as we have already reached the critical density of [C II] in the H II region; hence, emission from the H II region saturates while continuing to increase in the PDR region. There is then a plateau of the relation, which occurs at $\sim 10^{2.5}$ cm^{-3} because the densities are starting to approach the critical density of [C II] in the PDR regions, as shown in Fig. 9. At these densities, the [C II] emission from the PDR also saturates, leading to an overall constant value of $f_{[C\,II],mol}$. Finally, we test the consistency of the values of $f_{[C\,II],mol}$ obtained from equations (12)–(16). The value of $f_{[C\,II],mol}$ is calculated for the HRS galaxies using all five of these equations. Assuming that the most accurate estimate is given by the four-parameter equation (12), the offset between the other sets of measurements and this reference are shown in Fig. 10. As expected, the dispersion increases as the number of parameters used to calculate $f_{[C\,II],mol}$ decreases, and the uncertainty on $f_{[C\,II],mol}$ increases when using the one-parameter equation (15) or (16) compared to the four-parameter equation (12). This increase in uncertainty is accounted for by the larger $\sigma_{f_{[C\,II],mol}}$ values of equations (15) and (16) versus that of equation (12). As these relations were derived using the HRS to determine weighting factors, they are mostly applicable over the parameter space probed by the HRS galaxies, which can be seen in Fig. 8. While the full parameter space covered by our simulated clouds was very large (see Table 2), some regions of this space were ignored via the weighting factors if found to be not representative of physical conditions in local galaxies. Throughout the analysis presented in Section 6, a Galactic cosmic ray ionization rate was used, and so our prescriptions should only be used for low-redshift, normal star-forming galaxies. High-redshift galaxies, local ULIRGS, and other intensely star-forming objects are very likely to have higher cosmic ray ionization rates (100–1000× the Milky Way value). An analysis of the simulated clouds with high ionization rates would require a representative sample of galaxies at high redshift, similar to the HRS at $z \sim 0$, which is beyond the scope of this paper.

Finally, throughout the modelling presented here, we held constant the N/O abundance ratio, which is known to vary as a function of metallicity (Pettini et al. 2008; Guseva et al. 2011). In Appendix B, we explore how variations of this abundance ratio change our results, and find an uncertainty on $f_{[C\,II],mol}$ of less than 3 per cent, less than the reported errors in equations 12–16); ergo, this does not affect our results or the conclusions of this paper.

7 WHERE DOES A GALAXY'S [C II] EMISSION COME FROM?

Here, we summarize the main prescriptions detailed in this paper, and provide a cookbook to help decide which prescription is appropriate for a user's specific needs given their available data. There are five equations (equations 12, 13, 14, 15 and 16) that accurately quantify the fraction of [C II] emission emerging from molecular regions; however, which one should be used? To answer this, we present a flowchart in Fig. 11, which can be used to make this decision. The main decisions lie in determining which physical parameters of the galaxy have been observed and, therefore, what data are available.

7.1 SUMMARY AND CONCLUSIONS

We built a new 3D multiphase radiative transfer interface through the combination of STARBURST99, MOCASSIN and 3D-PDR, which can simulate all phases of the ISM, from ionized to molecular, where photoionization and photochemistry dominates. We assume pressure equilibrium between the ionized and neutral phases of the ISM, solving the thermal balance equations between the two regions to ensure self-consistency. This interface was used to simulate a broad family of spherically symmetric star-forming regions, with the aim of understanding how much of the total [C II] emission originates

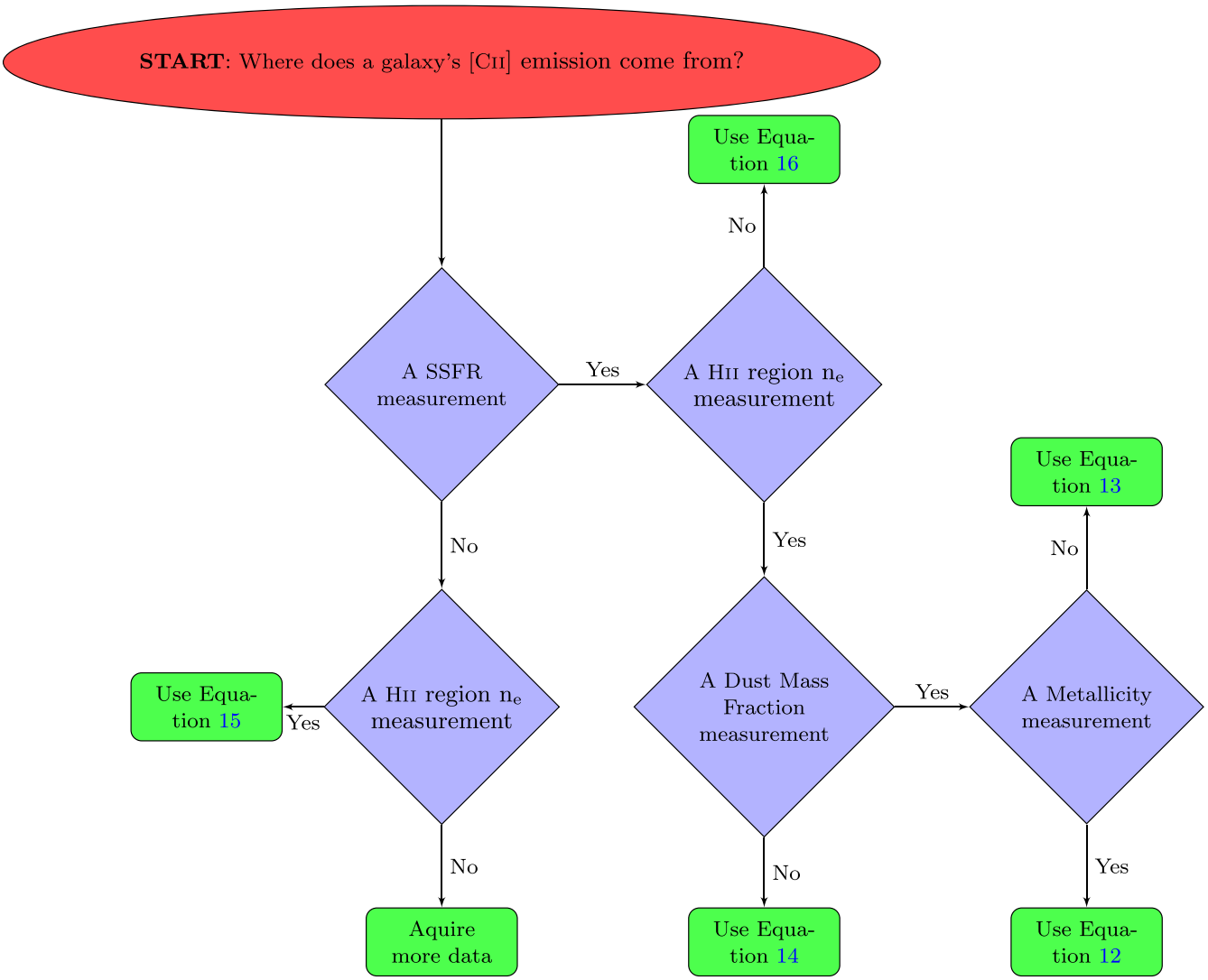

Figure 11. We here present a flowchart that can be used to constrain the fraction of [C II] emission from molecular regions from a galaxy. This will help to decide which equation should be used depending on which physical parameters of the galaxy have been observed and, therefore, what data are available for an individual galaxy.

from the cold molecular ISM under varying conditions. This is of importance, for example, to correctly interpret integrated [C II] measurements for distant galaxies.

An analysis of the relations obtained between $f_{[\mathrm{C\,II}],\mathrm{mol}}$ and the key input parameters of the multiphase code for these star-forming regions shows that an increase in the strength of the UV radiation field (whether by increasing the current SFR or having a recent burst of star formation) leads to a decrease of $f_{[\mathrm{C\,II}],\mathrm{mol}}$. Metallicity variations can lead to both an increase and a reduction of $f_{[\mathrm{C\,II}],\mathrm{mol}}$ depending on other global parameters, as it can both affect the cooling rate and the level of photodissociation of the CO molecule.

To extend the analysis to the integrated [C II] emission from extragalactic objects, we employed a hierarchical Bayesian inference method to identify the simulated clouds that are representative of the physical conditions in local star-forming galaxies, as found in the HRS. This is possible under the assumption that the physical conditions found in a simulated star-forming cloud can represent the average conditions found on galaxy-wide scales for objects with similar physical properties such as metallicity, SSFR and density.

We find that $f_{[\mathrm{C\,II}],\mathrm{mol}}$ is best predicted using four key parameters: n_{e}, SSFR, dust mass fraction and metallicity (equation 12). We tested this prescription on the Milky Way and obtained an estimate that 75.9 ± 5.9 per cent of its total [C II] emission arises from molecular regions, which is in very good agreement with observations placing this number at 75 per cent (Pineda et al. 2013). Given that it is relatively rare for measurements of all four of these parameters to be available for large samples of galaxies, we provide alternative prescriptions that invoke fewer parameters. These other prescriptions (equations 13–16) also produce estimates for the Milky Way consistent with direct observations, although the uncertainty on $f_{[\mathrm{C\,II}],\mathrm{mol}}$ increases slightly as the number of parameters involved in the prescription decreases. Of most practical use for many extragalactic studies is equation (16), which relates $f_{[\mathrm{C\,II}],\mathrm{mol}}$ to SSFR.

Applying the prescription to a sample of galaxies from the HRS, we find that typical galaxies in the local universe have 60–80 per cent of their [C II] emission arising from molecular regions. Within this sample, $f_{[\mathrm{C\,II}],\mathrm{mol}}$ increases with density, and decreases with SSFR. Combining the relations obtained through the Bayesian analysis,

we propose a decision tree in Fig. 11 to help determine which equation to use depending on the type of galaxy and the data products available. Using this, it is possible to estimate the relative fraction of [C II] emerging from the molecular phase of the ISM when only galaxy-wide observations are available.

ACKNOWLEDGEMENTS

GA would like to thank the UK Science and Technologies Facilities Council (STFC) for their support via a postgraduate Studentship, as well as Patrick J. Owen for informative discussions regarding the use of MOCASSIN. AS acknowledges the support of the Royal Society through the award of a University Research Fellowship and of a Research Grant.

REFERENCES

Abel N. P., Ferland G. J., Shaw G., van Hoof P. A. M., 2005, ApJS, 161, 65
Akaike H., 1981, J. Econometrics, 16, 3
Baldwin J. A., Ferland G. J., Martin P. G., Corbin M. R., Cota S. A., Peterson B. M., Slettebak A., 1991, ApJ, 374, 580
Bell T. A., Roueff E., Viti S., Williams D. A., 2006, MNRAS, 371, 1865
Bisbas T. G., Bell T. A., Viti S., Yates J., Barlow M. J., 2012, MNRAS, 427, 2100
Bisbas T. G., Bell T. A., Viti S., Barlow M. J., Yates J., Vasta M., 2014, MNRAS, 443, 111
Bisbas T. G., Haworth T. J., Barlow M. J., Viti S., Harries T. J., Bell T., Yates J. A., 2015, MNRAS, 454, 2828
Boselli A. et al., 2010, PASP, 122, 261
Boselli A., Hughes T. M., Cortese L., Gavazzi G., Buat V., 2013, A&A, 550, A114
Boselli A., Fossati M., Gavazzi G., Ciesla L., Buat V., Boissier S., Hughes T. M., 2015, A&A, 579, A102
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Carral P., Hollenbach D. J., Lord S. D., Colgan S. W. J., Haas M. R., Rubin R. H., Erickson E. F., 1994, ApJ, 423, 223
Cazaux S., Tielens A. G. G. M., 2004, ApJ, 604, 222
Chabrier G., 2003, PASP, 115, 763
Colombo D. et al., 2014, ApJ, 784, 3
Conroy C., 2013, ARA&A, 51, 393
Cormier D. et al., 2012, A&A, 548, A20
Cortese L. et al., 2012, A&A, 540, A52
Cortese L. et al., 2014, MNRAS, 440, 942
Cummings A. et al., 2015, Am. Astron. Soc. Meeting Abstr #SH52B-02
Daddi E. et al., 2007, ApJ, 670, 156
Dalgarno A., 2006, Proc. Natl. Acad. Sci., 103, 12269
Davé R., Oppenheimer B. D., Finlator K., 2011, MNRAS, 415, 11
Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98
de Jong T., 1977, A&A, 55, 137
de Jong T., Boland W., Dalgarno A., 1980, A&A, 91, 68
de Looze I., Baes M., Bendo G. J., Cortese L., Fritz J., 2011, MNRAS, 416, 2712
Draine B. T., 1978, ApJS, 36, 595
Draine B. T., Lee H. M., 1984, ApJ, 285, 89
Elbaz D. et al., 2007, A&A, 468, 33
Ercolano B., Barlow M. J., Storey P. J., Liu X.-W., 2003, MNRAS, 340, 1136
Ercolano B., Barlow M. J., Storey P. J., 2005, MNRAS, 362, 1038
Ercolano B., Dale J. E., Gritschneder M., Westmoquette M., 2012, MNRAS, 420, 141
Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994, A&AS, 105, 29
Ferland G. J. et al., 2013, Rev. Mex. Astron. Astrofis., 49, 137
Flower D. R., 1968, in Osterbrock D. E., O'dell C. R., eds, Proc. IAU Symp. 34, Planetary Nebulae. Reidel, Dordrecht, p. 205
Foster J. B., Mandel K. S., Pineda J. E., Covey K. R., Arce H. G., Goodman A. A., 2013, MNRAS, 428, 1606
Freedman D., Diaconis P., 1981, Z. Wahrscheinlichkeitstheor. Verwandte Geb., 57, 453
Goldsmith P. F., Langer W. D., Pineda J. L., Velusamy T., 2012, ApJS, 203, 13
Goodman J., Weare J., 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Gusev A. S. et al., 2016, MNRAS, 457, 3334
Guseva N. G., Izotov Y. I., Stasińska G., Fricke K. J., Henkel C., Papaderos P., 2011, A&A, 529, A149
Harries T. J., Howarth I. D., 1997, A&AS, 121, 15
Heiles C. et al., 1994, ApJ, 436, 720
Henney W. J., Arthur S. J., Williams R. J. R., Ferland G. J., 2005, ApJ, 621, 328
Ho I.-T. et al., 2014, MNRAS, 444, 3894
Hollenbach D., McKee C. F., 1979, ApJS, 41, 555
Hollenbach D. J., Takahashi T., Tielens A. G. G. M., 1991, ApJ, 377, 192
Hughes A. et al., 2013, ApJ, 779, 46
Hughes T. M. et al., 2015, A&A, 575, A17
Hunt L. K., Hirashita H., 2009, A&A, 507, 1327
Kamp I., Bertoldi F., 2000, A&A, 353, 276
Karczewski O. Ł. et al., 2013, MNRAS, 431, 2493
Kauffmann G. et al., 2003, MNRAS, 346, 1055
Kaufman M. J., Wolfire M. G., Hollenbach D. J., Luhman M. L., 1999, ApJ, 527, 795
Kelly B. C., 2007, ApJ, 665, 1489
Kelly G., Viti S., Bayet E., Aladro R., Yates J., 2015, A&A, 578, A70
Kramer C. et al., 2013, A&A, 553, A114
Kroupa P., 2001, MNRAS, 322, 231
Langer W. D., Velusamy T., Pineda J. L., Goldsmith P. F., Li D., Yorke H. W., 2010, A&A, 521, L17
Langer W. D., Velusamy T., Pineda J. L., Willacy K., Goldsmith P. F., 2014, A&A, 561, A122
Lee H.-H., Herbst E., Pineau des Forets G., Roueff E., Le Bourlot J., 1996, A&A, 311, 690
Leitherer C. et al., 1999, ApJS, 123, 3
Leitherer C., Ortiz Otálvaro P. A., Bresolin F., Kudritzki R.-P., Lo Faro B., Pauldrach A. W. A., Pettini M., Rix S. A., 2010, ApJS, 189, 309
Leroy A. K. et al., 2011, ApJ, 737, 12
Licquia T., Newman J., 2014, Am. Astron. Soc. Meeting Abstracts, 223, 336.04
Lucy L. B., 1999, A&A, 344, 282
McElroy D., Walsh C., Markwick A. J., Cordiner M. A., Smith K., Millar T. J., 2013, A&A, 550, A36
Maiolino R. et al., 2015, MNRAS, 452, 54
Makrymallis A., Viti S., 2014, ApJ, 794, 45
Mandel K. S., Narayan G., Kirshner R. P., 2011, ApJ, 731, 120
Mookerjea B. et al., 2011, A&A, 532, A152
Mookerjea B. et al., 2016, A&A, 586, A37
Noeske K. G., Weiner B. J., Faber S. M., Papovich C., Koo D. C., Somerville R. S., Bundy K., 2007, ApJ, 660, L43
Oberst T. E. et al., 2006, ApJ, 652, L125
Offner S. S. R., Bisbas T. G., Bell T. A., Viti S., 2014, MNRAS, 440, L81
Olsen K. P., Greve T. R., Brinch C., Sommer-Larsen J., Rasmussen J., Toft S., Zirm A., 2016, MNRAS, 457, 3306
Ota K. et al., 2014, ApJ, 792, 34
Papadopoulos P. P., 2010, ApJ, 720, 226
Peng Y.-j. et al., 2010, ApJ, 721, 193
Pettini M., Zych B. J., Steidel C. C., Chaffee F. H., 2008, MNRAS, 385, 2011
Pierce-Price D. et al., 2000, ApJ, 545, L121
Pineda J. L., Langer W. D., Velusamy T., Goldsmith P. F., 2013, A&A, 554, A103
Reddy N. A., Pettini M., Steidel C. C., Shapley A. E., Erb D. K., Law D. R., 2012, ApJ, 754, 25
Riechers D. A. et al., 2014, ApJ, 796, 84

Röllig M., Ossenkopf V., Jeyakumar S., Stutzki J., Sternberg A., 2006, A&A, 451, 917
Sanders R. L. et al., 2016, ApJ, 816, 23
Schwarz G., 1978, Ann. Stat., 6, 461
Shetty R., Kelly B. C., Bigiel F., 2013, MNRAS, 430, 288
Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15
Stacey G. J., Hailey-Dunsheath S., Ferkinhoff C., Nikola T., Parshley S. C., Benford D. J., Staguhn J. G., Fiolet N., 2010, ApJ, 724, 957
Tremonti C. A. et al., 2004, ApJ, 613, 898
van Dishoeck E. F., Black J. H., 1988, ApJ, 334, 771
Vasta M., 2010, Doctoral thesis, Univ. College London
Vasta M., Barlow M. J., Viti S., Yates J. A., Bell T. A., 2010, Highlights Astron., 15, 408
Vázquez G. A., Leitherer C., 2005, ApJ, 621, 695
Velusamy T., Langer W. D., Pineda J. L., Goldsmith P. F., Li D., Yorke H. W., 2010, A&A, 521, L18
Viti S., 2013, Astrophysics Source Code Library, record ascl:1303.004
Walch S., 2016, in EAS Publ. Ser., 75, 359
Wong T. et al., 2011, ApJS, 197, 16
Wood K., Mathis J. S., Ercolano B., 2004, MNRAS, 348, 1337
Wright N. J., Drake J. J., Drew J. E., Vink J. S., 2010, ApJ, 713, 871

APPENDIX A: VARYING THE COSMIC RAY IONIZATION RATE

We performed a similar analysis as in Section 6.1 for clouds with a cosmic ray ionization rate 10× the average Milky Way value (10^{-16} s^{-1}). We find that the same four parameters emerge as necessary to provide the a prescription for $f_{[\mathrm{C\,II}],\mathrm{mol}}$. We use this higher cosmic ray ionization rate and produce identical plots to Figs 6–8. We show in Fig. A1 a plot for the AIC and BIC, similar to Fig. 6, and find three to four parameters are needed, similar to that in Section 6.1. Once we obtained our prescription, we applied it to the HRS objects, which have observed values for the four important parameters. We present these results in Fig. A2 and, qualitatively, it is clear that the results are very similar and almost identical to those shown in Fig. 8. Finally, we bin the HRS results, similar to Fig. 7, and again find that, even for clouds with a cosmic ray ionization rate 10× the average Milky Way value, the majority of the galaxies have 60–80 per cent of their total integrated [C II] emission arising from molecular regions, shown in Fig. A3.

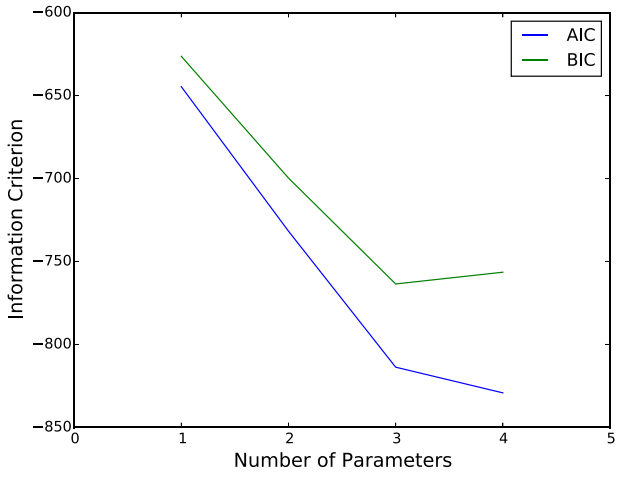

Figure A1. We provide a plot for the variation of the Akaike Information Criterion and the Bayesian Information Criteria in blue and green, respectively, but now for clouds with a cosmic ray ionization rate 10× the average Milky Way value. It can be seen how they reach a minimum of three to four necessary parameters, similar to Fig. 6.

Overall, we claim that, even if a galaxy is thought to have cosmic ray ionization rates 10 times larger than the Milky Way value, our prescriptions detailed in Section 6 are still robust and accurate as the cosmic ray ionization rate value does not affect the results at these levels. Higher redshift objects and ULIRGS will have cosmic ray ionization rates more than $\sim 10^3$ times that of the Milky Way, and for those cases our prescription, in Section 6, starts to break down. Furthermore, our prescription would not be valid at high redshift because the complete HRS sample is only complete for the low-redshift universe. Therefore, our prescription is accurate only for low-redshift, star-forming and quiescent galaxies regardless of their cosmic ray ionization rate.

Obtaining a similar prescription for high-redshift objects is possible, using the above method; however, a statistically complete sample of galaxies at high redshift would be needed to provide the weightings necessary for the Bayesian inference method. This could be done via a machine learning technique to generate a predictive sample of galaxies at high redshift; however, this is beyond the scope of this paper.

APPENDIX B: VARYING CHEMICAL ABUNDANCES

When varying the metallicity parameter, we scaled all the abundances in Table 1 equally, except for hydrogen and helium. This means that the relative abundances between non-hydrogen and helium elements is constant. While this is generally correct, it is not true of the N/O ratio, which varies as a function of metallicity (Pettini et al. 2008; Guseva et al. 2011). When $\log(\mathrm{O/H}) + 12 > 8.2$, nitrogen is a secondary element and the N/O ratio decreases with metallicity. However, when $\log(\mathrm{O/H}) + 12 < 8.2$, nitrogen is a primary element and the N/O ratio is of constant value $10^{-1.5}$. Therefore, by assuming a constant N/O ratio, we have oversupplied the low-metallicity clouds with nitrogen, which could lead to an erroneous [C II] emission calculation in the MOCASSIN simulations. Given that nitrogen is a coolant in ionized regions only, this would manifest itself as an underestimation of the [C II] emission from the ionized regions.

To test for the effect of the varying N/O ratio with metallicity, we adopt the prescription from Pettini et al. (2008) and re-run MOCASSIN for three of the low-metallicity clouds ($Z = 0.2\,Z_{\odot}$) with different SSFR and n_{e} (the age of the secondary burst is kept constant). We find that assuming a constant N/O could lead to an underestimation of the [C II] emission from the ionized of only 3.7–7.7 per cent. Since ionized regions contribute between 20 and 40 per cent of the total [C II] emission (Section 6.5), this corresponds to an uncertainty on $f_{[\mathrm{C\,II}],\mathrm{mol}}$ of less than 3 per cent, less than the reported errors in equations (12)–(16). This is therefore not a dominant source of uncertainty; the results for these runs are shown in Table A1.

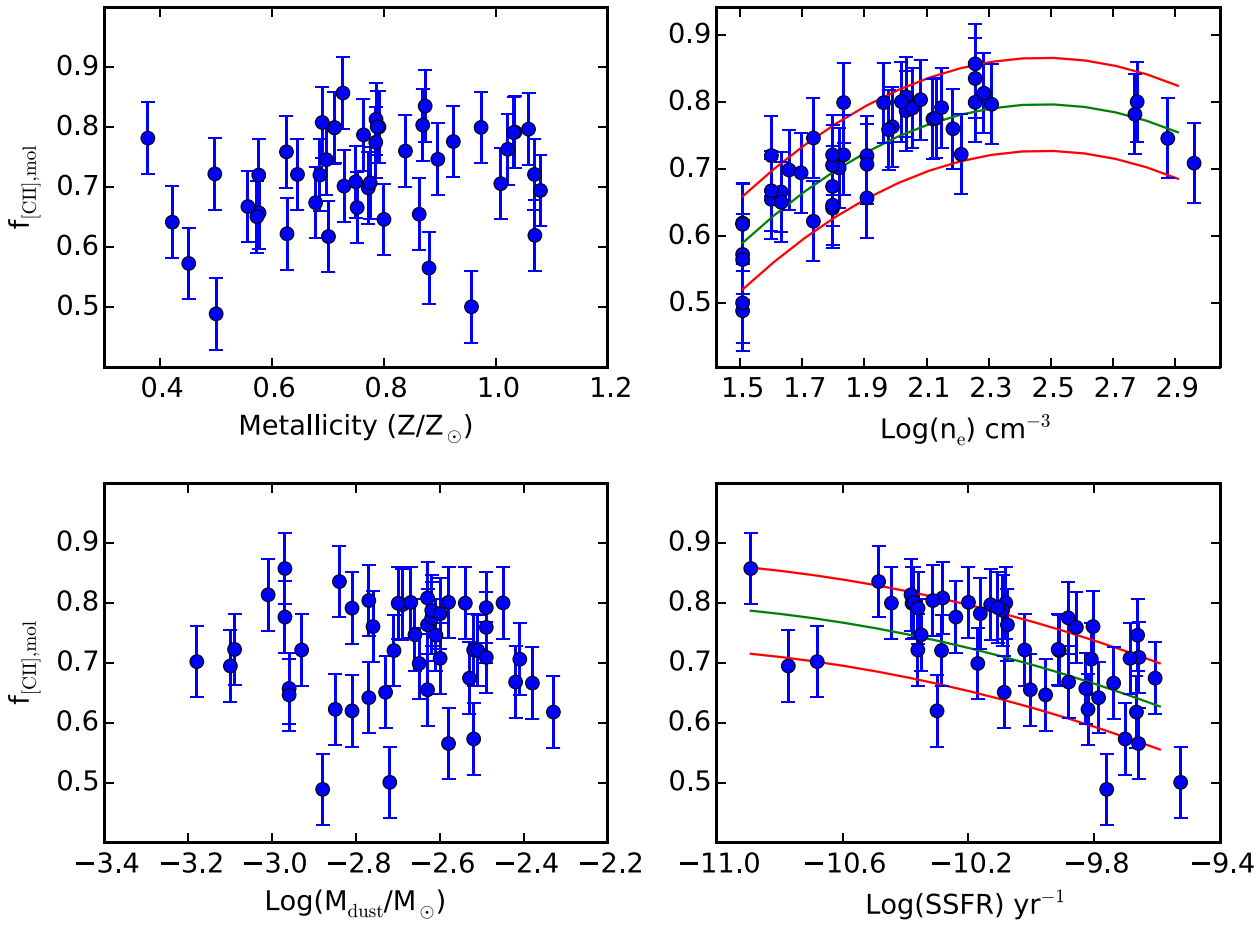

Figure A2. Relation between $f_{[\mathrm{C\,II}],\mathrm{mol}}$ for galaxies from the HRS sample for a cosmic ray ionization rate 10× the average Milky Way value, a comparable plot to Fig. 8. The different coloured lines here represent the same as those in Fig. 8.

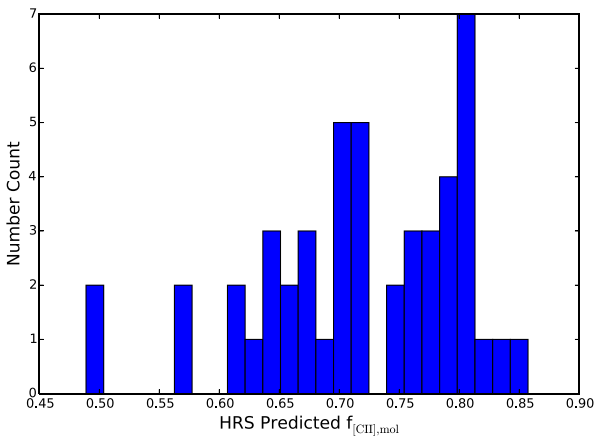

Figure A3. From the HRS sample we find that the majority of the galaxies have 60–80 per cent of their total integrated [C II] emission arising from molecular regions, even for a cosmic ray ionization rate 10× the average Milky Way value, a similar result to that in Fig. 7.

Table A1. To test for the effect of varying the N/O ratio as a function of metallicity we re-run three of our MOCASSIN runs. We find that, even in the most extreme cases, our calculations could underestimate the [C II] emission from the ionized regions by 3.7–7.7 per cent.

Metallicity ($Z/Z_\odot$)	log(SSFR) (yr^{-1})	n_e (cm^{-3})	[C II] with N/O constant ($L/L_\odot$)	[C II] with N/O varying ($L/L_\odot$)	Fractional difference (%)
0.2	−11.5	$10^{1.5}$	5.15×10^{-3}	5.19×10^{-3}	7.7
0.2	−10.5	$10^{2.0}$	11.32×10^{-3}	11.92×10^{-3}	5.3
0.2	−10.5	$10^{3.0}$	9.55×10^{-3}	9.91×10^{-3}	3.77

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