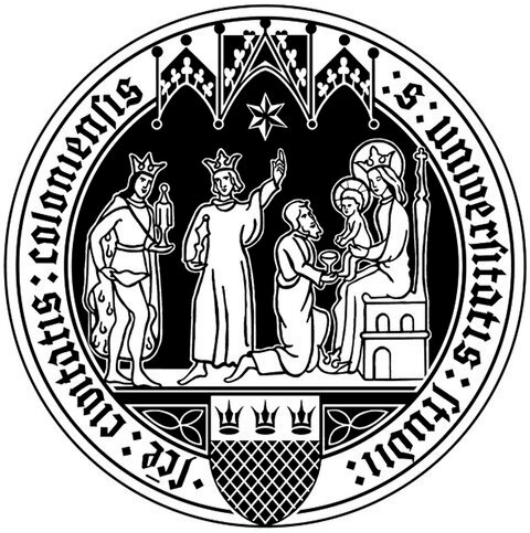
The Milky Way with `kosmatau3d`:

Modelling the Galactic cooling lines using clumpy PDRs



Inaugural-Dissertation
zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Craig Nicholas Yanitski
aus Edmonton

Köln 2023

Berichterstatter/in:	Priv.-Doz. Dr. Volker Ossenkopf-Okada
(Gutachter/in)	Prof. Dr. Stefanie Walch-Gassner
Vorsitzender:	Prof. Dr. Dirk Witthaut
Beisitzer:	Priv.-Doz. Dr. Markus Röllig
Tag der mündlichen Prüfung:	**02.11.2023**

For my family,
both in Canada and Europe.

Foreword

> *Aristotle said a bunch of stuff that was wrong. Galileo and Newton fixed things up. Then Einstein broke everything again. Now, we've basically got it all worked out, except small stuff, big stuff, hot stuff, cold stuff, fast stuff, heavy stuff, dark stuff, turbulence, and the concept of time.*
>
> Zach Weinersmith,
> Science: Abridged Beyond
> the Point of Usefulness

One of the most beautiful aspects of astronomy is the relation between scales. Giant interstellar clouds of gas and dust, with masses several orders of magnitude larger than our Sun, fragment and collapse into new stars. Some of these bright, young stars emit enough radiation to push against and ionise their natal clouds, eventually ceasing their epochs of fusion in violent explosions called supernovae. The impact of these supernovae enrich the surrounding gas and trigger a new generation of stars to be born, and thus the cycle repeats.

This relation of scales goes even further, though, as the structures seen within these gaseous clouds are also related. Inhomogeneity, which is the word we use here to describe a gas without any discernible density or temperature structure, is ubiquitous throughout the interstellar gas. Gaseous clouds and filaments contain small, denser regions called cores, and these themselves contain many smaller unresolved structures called clumps. It is in the research of these clumps that we must begin developing statistical models. The development of these models is thus of great interest to the scientific community, and it is the backbone of the ensuing thesis. By applying this statistical model to the Milky Way, the galaxy in which we reside, we are able to conclude whether these models are accurate on a the scale of a galaxy, or more correctly what are the caveats in such a comparison.

Craig Nicholas Yanitski
21.12.2023

Abstract

Given the number of large-scale Milky Way surveys of the dust continuum and spectroscopic line transitions, we are in a favourable position to test different models of the interstellar medium (ISM) in order to better understand and constrain its physics. While the numerous features of the Galactic ISM observation cannot be matched simultaneously by a test model, the large-scale structure and the typical line ratios and intensities should be reproduced in a statistically significant way. To that end, previous research has utilised the KOSMA-τ photon-dissociation region (PDR) code to fit the line emission as observed with COBE-FIRAS with very few assumptions and constraints. The basis of this work was the fractal structure of the ISM: a theory that accounts for the high surface-to-volume ratio as inferred from observations allowing the sophisticated post-processing of self-consistent model results. I have now extended this approach to arbitrary geometries in the novel PDR code `kosmatau3d`, which also accounts for the dust pumping, continuum emission, and absorption effects. The code uniquely accounts for the clumpy structure of the inhomogeneous ISM. I have used this code to model the molecular and atomic cold neutral medium (CNM) in the Milky Way using axisymmetric distributions of the dense and diffuse gas mass (traced with atomic and molecular hydrogen), average density, cosmic ray ionisation rate, and far-UV radiation. I find that much of the observed large-scale structure in the spectroscopic maps can be replicated in the synthetic observations, though there remains difficulty to fully model the self-absorption features toward the Galactic centre. I also find these `kosmatau3d` synthetic observations are able to replicate the molecular ring and certain Galactic arm features even without an explicit implementation of the spiral structure of the Milky Way. This suggests the structure seen in the galactic position-velocity diagram arises to a large extent from the rotation of the galaxy and the molecular ring, and thus offers insight into the interpretation of galactic longitude-velocity diagrams. I also summarise the benefit of using multiple transition lines in the constraining of the Milky Way parameters, how these constraints improve on and differ from previous work, and finally conclude with an outlook on this method of modelling and how well the Galactic cooling lines can be modelled using PDRs.

Contents

List of Figures

List of Tables

CHAPTER I

Introduction

After all, the universe required ten billion years of evolution before life was even possible; the evolution of the stars and the evolving of new chemical elements in the nuclear furnaces of the stars were indispensable prerequisites for the generation of life.

John Polkinghorne

The interstellar medium (ISM) is a complex mixture of gas, dust and other particles, which span several orders of magnitude in energy, scale, density, temperature, ionisation rate, etc. The physics of the ISM affect several components of the galaxy, from stellar properties (such as composition and formation) to molecular cloud and cluster properties (such as fragmentation and velocity dispersion) to energy transport (such as the various forms of stellar feedback). Due to its ubiquity, the ISM is typically studied in terms of the interstellar medium feedback cycle. The feedback cycle, alluded to in the Foreword, involves the cycle and processing of material from molecular clouds to star-forming regions to stars and back to molecular clouds. Understanding any one of these aspects is certainly of interest, but it is insufficient to fully understand the physics of the ISM. Not only that, but the ISM exists in a wide range of conditions, yielding the ISM phases discussed in §I-1. Still, the best method we know how to develop a more general understanding of the ISM is to first investigate "special cases", that is, understanding the ISM when dominated by certain processes. Examples of these include giant molecular cloud surveys, where the ISM is assumed to be relatively isolated and the primary heating mechanism of the gas and dust is collisional; shock front analyses, where the ISM is assumed to be moving faster than the speed of sound compared with some density interface, and the energy input is due to the physics of crossing the shock front; and PDR analyses, where the dominant heating mechanism is the radiation from O/B stars and stellar clusters.

What is commonly calculated for these cases is the star-formation rate (SFR), $\Sigma_{\rm SFR}$, the surface density of the rate at which ISM matter is lost to stellar matter

Figure I–1: The Rosette H II region as observed with *Herschel*. Credit: ESA and the PACS, SPIRE & HSC consortia, F. Motte (AIM Saclay, CEA/IRFU - CNRS/INSU - Université Paris Didedrot) for the HOBYS key programme.

(see the reviews by McKee & Ostriker 2007; Kennicutt & Evans 2012, and references therin). This is commonly calculated in extra-galactic surveys of giant molecular clouds as a measure of the ISM conditions. Conditions such as turbulence, magnetic field, and gravity affect the SFR, making $\Sigma_{\rm SFR}$ indicative of the types of regions contained in the ISM. For example, a high star formation rate ($\Sigma_{\rm SFR} \gtrsim 0.1\,{\rm M}\odot\,{\rm yr}^{-1}{\rm pc}^{-2}$; see for example Figure 1 in Leroy et al. 2021) is suggests there is a high probability that a given molecular cloud hosts an O/B stellar cluster.

Within the context of this work, we focus on environments of hot stars and stellar clusters. These stellar environments produce most of their radiation in the extreme ultraviolet (EUV) regime, and thus ionise the surrounding hydrogen (see Figure I–1). The resulting object is referred to as either a H II region or Stömgren region (in the ideal case this is a sphere; Strömgren 1939). These regions in turn produce a variety of different types of pressure that contribute to expanding their radii, whether it is internal gas pressure, radiative pressure, or stellar wind ram pressure (Barnes et al. 2021a,b). These pressures cause much of the dust in the ISM to be swept to the edges of the H II region, sometimes creating a slightly-shocked region depending on the magnitude of the pressures. While most studies in H II regions make an assumption of perfect spherical bubble and homogeneous densities (see for example Lopez et al. 2011, 2014; McLeod et al. 2021; Geen & de Koter 2022), there have been studies on the inhomogeneity of the H II region morphology (Lancaster et al. 2021) and even the stellar wind (Driessen et al. 2022). The existence of these inhomogeneities in turn propagate into the edge of the H II region.

This interface between the H II region and neighbouring molecular cloud is known as a photo-dissociation region (PDR; alternatively photon-dominated region; see Figure I–2 and Figure I–1). PDRs are predominantly neutral regions of the

Figure I–2: The Carina nebula as observed with the Near-Infrared Camera (NIRCAM) and Mid-Infrared Instrument (MIRI) on the *James Webb Space Telescope* (*JWST*). Credit: NASA, ESA, CSA, and STScI.

ISM in the vicinity of O/B stars and stellar clusters, where the dominant heating mechanism is far-ultraviolet (far-UV) radiation, while gas is cooled primarily through line-transitions of the gas. These regions have been modelled for decades, however observations from the 1990s (Stutzki & Guesten 1990; Meixner & Tielens 1993) indicated the necessity for a higher surface to volume ratio than that considered in the models. This motivated the fractal approach to PDR-modelling, where the PDR is approximated by an ensemble of several self-similar clumps (introduced in §I-3).

This thesis is concerned with the development of the clumpy PDR code `kosma–tau3d` and its application to the galactic cooling lines of the Milky Way. The supporting theory is discussed in Chapter II, and Chapter III covers its implementation into the PDR code. From that point in the thesis it should be clear how one can use the code to create a three-dimensional PDR model, which information is necessary to know *a priori*, and which quantities can be constrained by the model. In Chapter IV the galactic models are introduced, including the prior information we use to first run the model and which model parameters we attempt to constrain. The observational surveys used to verify our models are presented in Chapter V. The model grids are constrained in Chapter VI, where we are careful to describe clearly how we combine the model fits between different surveys. This is put in context in Chapter VII, and finally summarised in Chapter VIII. By the end of this work, we want to determine, or at least constrain, the following questions:

1. How successfully can a clumpy PDR model be used to simulate the galactic cooling line of a Milky Way-type galaxy?

2. How many, if any, of the galactic cooling line large-scale structures can be reasonably modelled with PDRs?

3. How well can the clumpy models simulate the atomic and molecular mass distribution in the Milky Way?

4. How do our constraints vary depending on fitting dense gas tracers compared to diffuse gas tracers?

5. Which artefacts, if any, exist in our synthetic observations? Is this seen both in the spectroscopic data as well as the integrated intensity data?

6. What effect does the assumption of PDR clumps filling the galactic disk have on the $X_{\rm CO}$ factor throughout the galaxy?

7. Which transitions/wavelengths are ideal to constrain the galactic parameters?

1 Interstellar medium

There is a wide range of conditions in the ISM, and we generally separate it into the ionised, atomic, and molecular phases. This is determined by which phase of hydrogen has the highest abundance, which in turn depends on the conditions of the gas and dust. In general, it is common to differentiate between the following phases (Draine 2011):

- hot ionised medium (HIM; $T > 10^6$ K, $n_{\rm H} \sim 4 \times 10^{-3}$ cm^{-3}),
- warm ionised medium (WIM; $T \sim 10^4$ K, $n_{\rm H} \sim 10^{-1} - 10^4$ cm^{-3}),
- warm neutral medium (WNM; $T > 10^3$ K, $n_{\rm H} \sim 10^{-1}$ cm^{-3}),
- cold neutral medium (CNM; $T \approx 10^2$ K, $n_{\rm H} > 10$ cm^{-3}), and
- the dense and diffuse molecular gas ($T \approx 10$ K, $n_{\rm H} \sim 10^2, 10^3 - 10^6$ cm^{-3}).

The HIM is the hot coronal plasma in stellar environments and accounts for a small fraction of the Milky Way mass, while the rest of the ionised material is in the WIM. The atomic medium is approximately split between the WNM and CNM, which accounts for about half of the mass inside of the orbit of the Sun in the Milky Way (henceforth known as the *solar circle*).

1.1 Observations

It is difficult to observe stars in the Milky Way through the Galactic plane in optical wavelengths since all of this material–the ISM–obscures our view. Indeed most, if not all, optical observations will need to account for some amount a attenuation (the scattering and absorption of light) due to the gas and dust that permeates the Galaxy. In this sense, the ISM is mainly seen as a hindrance in optical stellar observations, another source of error. How then are we able to observe and study the ISM?

For this we need to understand the physics of radiation and spectroscopy: in which wavelengths can we observe the ISM and how does it change as it propagates through different media. Unlike shorter wavelength bands, far-infrared radiation is only slightly attenuated by the dust in the ISM. What is more important for infrared observations is radiative transfer, which accounts for the emission and absorption of the intervening material. For a given path length s along a line-of-sight, the observed intensity is calculated using the radiative transfer equation:

$$\mathrm{d}I_\nu = -I_\nu \, \mathrm{d}\tau_\nu + S_\nu \, \mathrm{d}\tau_\nu \ . \tag{I-1}$$

Here we have the source function S_ν, which is the ratio of how many photons are created to how many are destroyed, the optical depth τ_ν, which a dimensionless quantity related to how many photons are absorbed along a distance s, and I_ν is the observed intensity. Both I_ν and S_ν have units power per unit area per unit frequency per solid angle. We can further define the optical depth in terms of the path length s and the absorption κ_ν:

$$\kappa_\nu \equiv \frac{\mathrm{d}\tau_\nu}{\mathrm{d}s} \ . \tag{I-2}$$

In §I-1.2 we will derive precisely how to calculate the absorption[1]. We can also identify a related term for emissivity j_ν, in units power per volume per frequency per solid angle, and rewrite the radiative transfer equation as a function of path length:

$$\mathrm{d}I_\nu = -I_\nu \, \kappa_\nu \, \mathrm{d}s + j_\nu \, \mathrm{d}s \ . \tag{I-3}$$

From this we can identify the formal definition for the source function:

$$S_\nu \equiv \frac{j_\nu}{\kappa_\nu} \ . \tag{I-4}$$

In local thermal equilibrium (LTE), this is equivalent to the radiation from a blackbody $B_\nu(T)$:

$$B_\nu(T) \equiv \frac{2\,h\,\nu^3}{c^2} \left(\mathrm{e}^{\frac{h\,\nu}{k_\mathrm{B}\,T}} - 1 \right)^{-1} \ , \tag{I-5}$$

with energy density u_ν:

$$\begin{aligned} u_\nu &= \frac{4\pi}{c} \, B_\nu(t) \ , \\ u_\nu &= \frac{8\pi\,h\,\nu^3}{c^3} \left(\mathrm{e}^{\frac{h\,\nu}{k_\mathrm{B}\,T}} - 1 \right)^{-1} \ . \end{aligned} \tag{I-6}$$

Following Draine (2011), it useful to define the *photon occupation number* n_γ:

$$\overline{n}_\gamma \equiv \frac{c^3}{8\pi\,h\,\nu^3} \, u_\nu \ , \tag{I-7}$$

where n_γ is averaged over all directions to express in terms of the energy density.

[1] The absorption is sometimes referred to as the absorption coefficient due to how it enters into the radiative transfer equation (Equation I-3).

Rayleigh-Jeans approximation

In practice, at least for long wavelengths, it is often easier to work instead with the *Rayleigh-Jeans approximation*. In this approximation we define the *brightness temperature* T_B as the temperature of a blackbody that would give the corresponding intensity. We can then rewrite Equation I–5 in terms of the brightness temperature using two Taylor expansions in the limit $h\,\nu \ll k_\mathrm{B}\,T_\mathrm{B}$:

$$\begin{aligned} I_\nu &= \frac{2\,h\,\nu^3}{c^2}\left(\sum_{n=1}^{\infty}\frac{1}{n!}\left(\frac{h\,\nu}{k_\mathrm{B}\,T_\mathrm{B}}\right)^n\right)^{-1} , \\ I_\nu &= \frac{2\,k_\mathrm{B}\,\nu^2}{c^2}\,\frac{T_\mathrm{B}}{1-\sum_{n=2}^{\infty}\frac{1}{n!}\left(\frac{h\,\nu}{k_\mathrm{B}\,T_\mathrm{B}}\right)^{n-1}} , \\ I_\nu &= \frac{2\,k_\mathrm{B}\,\nu^2}{c^2}\left(T_\mathrm{B}-\Delta T_{\mathrm{RJ},\nu}\right) , \end{aligned} \tag{I–8}$$

where $\Delta T_{\mathrm{RJ},\nu}$ is called the *Rayleigh-Jeans correction*. For long wavelengths, the approximation is enough to just use the $n=1$ term of Equation I–8:

$$I_\nu \approx \frac{2\,k_\mathrm{B}\,\nu^2}{c^2}\,T_\mathrm{B} \ . \tag{I–9}$$

The error in this approximation will increase as $h\,\nu \to k_\mathrm{B}T_\mathrm{B}$, in which case the $n>1$ terms should be considered for the Rayleigh-Jeans correction.

1.2 Microphysics of emission

The interaction of atomic and molecular species with radiation is quantised into transitions between energy levels. We study this using the number density of species in a specific energy level, and simplify the discussion by considering species with two energy levels. Thus n_l is the number density in the lower energy level and n_u is the number density in the upper energy level. In order to examine the rate of change of these energy levels, it is necessary know the corresponding Einstein coefficients:

- A_ul: spontaneous emission, yielding the rate $(\mathrm{d}n_\mathrm{u}/\mathrm{d}t)_\mathrm{em,spon} = -n_\mathrm{u}A_\mathrm{ul}$.
- B_ul: stimulated emission, yielding the rate $(\mathrm{d}n_\mathrm{u}/\mathrm{d}t)_\mathrm{em,stim} = -n_\mathrm{u}B_\mathrm{ul}u_\nu$.
- B_lu: absorption, yielding the rate $(\mathrm{d}n_\mathrm{u}/\mathrm{d}t)_\mathrm{abs} = n_\mathrm{l}B_\mathrm{lu}u_\nu$.

The density at which radiative de-excitation is equivalent to collisional de-excitation is referred to as the *critical density*. It is a useful metric to understand roughly the minimum conditions for the upper energy level to be sufficiently populated to sustain the u$\rightarrow$l transition. For a two-level system with collisional de-excitation rate $k_{\rm ul}$ (in cm^3 s^{-1}) and photon occupation number $n_{\gamma,{\rm ul}}$ (with wavelength $\lambda_{\rm ul}$), the critical density is (Draine 2011):

$$n_{\rm crit,ul} = \frac{(1 + n_{\gamma,{\rm ul}})\, A_{\rm ul}}{k_{\rm ul}} \ . \tag{I–10}$$

Restricting the analysis with the two-level approximation simplifies the theoretical treatment of electron transitions in a species, but it should be noted that this typically overestimates quantities such as the critical density. The full multi-level should consider transitions from other energy levels (giving sums in the numerator and denominator of Equation I–10) and account for effects like *radiative-trapping*, where a given transition u$\rightarrow$l is also optically-thick in a molecular cloud and thus the radiation cannot escape the cloud. Still, it is assumed that the two-level approximation is reasonable for many regions in the ISM and we account for radiative trapping for a small subset of molecular transitions (for example CO and H_2).

It is also possible to relate the population in the upper and lower energy levels by defining the *excitation temperature*:

$$\frac{n_{\rm u}}{n_{\rm l}} = \frac{g_{\rm u}}{g_{\rm l}} {\rm e}^{-\frac{E_{ul}}{k_{\rm B}\, T_{\rm ex}}} \ . \tag{I–11}$$

Here g_i refers to the quantum degeneracy of an energy level. The total change in the upper population is then,

$$\frac{{\rm d}n_{\rm u}}{{\rm d}t} = \left(\frac{{\rm d}n_{\rm u}}{{\rm d}t}\right)_{\rm em,spon} + \left(\frac{{\rm d}n_{\rm u}}{{\rm d}t}\right)_{\rm em,stim} + \left(\frac{{\rm d}n_{\rm u}}{{\rm d}t}\right)_{\rm abs} \ , \tag{I–12}$$

$$\frac{{\rm d}n_{\rm u}}{{\rm d}t} = -n_{\rm u}\, A_{\rm ul} - n_{\rm u}\, B_{\rm ul}\, u_\nu + n_{\rm l}\, B_{\rm lu} u_\nu \ . \tag{I–13}$$

If we consider the incident radiation energy density spectrum follows a blackbody spectrum (see Equation I–6) and the species are in thermal equilibrium with the radiation, the levels will be populated according to Equation I–11 with $T_{\rm exc} = T_{\rm gas}$. By assuming ${\rm d}n_{\rm u}/{\rm d}t = 0$ and taking the limits $T \rightarrow 0$ and $T \rightarrow \infty$, it can be shown,

$$B_{\rm ul} = \frac{g_{\rm l}}{g_{\rm u}}\, B_{\rm lu} \ , \tag{I–14}$$

$$A_{\rm ul} = \frac{8\,\pi\, h\, \nu^3}{c^3}\, B_{\rm ul} \ , \tag{I–15}$$

in which we use the principle of detailed balance: absorption events and emission events are balanced in thermal equilibrium.

Emissivity

The emissivity j_ν can be calculated directly from the Einstein $A_{\rm ul}$ coefficient:

$$j_\nu = n_{\rm u} \frac{h\,\nu_{\rm ul}}{4\pi} A_{\rm ul}\,\phi_\nu \ , \tag{I–16}$$

where ϕ_ν is the line profile of the emissivity, explained in depth later. We can also apply the Rayleigh-Jeans approximation to the emissivity to derive it in units of $\mathrm{K\,s^{-1}}$:

$$\epsilon_\nu = n_{\rm u} \frac{h\,c^2}{8\pi\,k_{\rm B}\,\nu_{\rm ul}} A_{\rm ul}\,\phi_\nu \ . \tag{I–17}$$

We would like to consider the interaction of light with the species in the ISM. For this, we can think of the interaction (read: *collision*) of a photon and some particle (molecule, ion, etc.). Thus we necessarily must think in terms of the scattering and absorption cross sections of species sp, $\sigma_{\rm sca,sp}$ and $\sigma_{\rm abs,sp}$, where the extinction cross section is $\sigma_{{\rm ext},s} = \sigma_{{\rm sca},s} + \sigma_{{\rm abs},s}$. These terms are necessarily functions of wavelength λ. When applying this to the ISM in general, which is a mixture of many different species, we identify the cross section per hydrogen nucleus, σ'. If we want to consider the propagation of radiation through a region of the ISM with hydrogen column density $N_{\rm H}$ (parallel with the radiation), the optical depth of this region is calculated by $\tau_\nu = N_{\rm H}\sigma'_\nu$. We note that since column density has units of $\mathrm{cm^{-2}}$ and cross section has units of $\mathrm{cm^2}$, the optical depth is dimensionless. If the region is homogeneous over length ℓ, the attenuation coefficient is by definition:

$$\kappa_\nu \equiv \tau_\nu \ell^{-1} = n_{\rm H}\sigma'_\nu \ . \tag{I–18}$$

Obtaining the absorption coefficient then depends on the cross-section of the transition:

$$\sigma_{\nu,{\rm lu}} = \frac{h\,\nu_{\rm lu}}{c} B_{\rm lu}\,\phi_\nu \ , \tag{I–19}$$

$$\sigma_{\nu,{\rm lu}} = \frac{g_{\rm u}}{g_{\rm l}} \frac{c^2}{8\pi\,\nu_{\rm lu}^2} A_{\rm ul}\,\phi_\nu \ ,$$

$$\sigma_{\nu,{\rm lu}} = \frac{g_{\rm u}}{g_{\rm l}}\,\sigma_{\nu,{\rm ul}} \ . \tag{I–20}$$

Absorption

The absorption κ_ν can be calculated using the cross-section:

$$\begin{aligned}
\kappa_\nu &= n_\mathrm{l}\,\sigma_\mathrm{lu} - n_\mathrm{u}\,\sigma_\mathrm{ul}, \\
\kappa_\nu &= n_\mathrm{l}\,\sigma_\mathrm{lu}\left(1 - \frac{n_\mathrm{u}\,g_\mathrm{l}}{n_\mathrm{l}\,g_\mathrm{u}}\right) , \\
\kappa_\nu &= n_\mathrm{l}\,\frac{g_\mathrm{u}}{g_\mathrm{l}}\,\frac{c^2}{8\pi\,\nu_\mathrm{lu}^2}\,A_\mathrm{ul}\,\phi_\nu\left(1 - \frac{n_\mathrm{u}\,g_\mathrm{l}}{n_\mathrm{l}\,g_\mathrm{u}}\right) , \\
\kappa_\nu &= n_\mathrm{l}\,\frac{g_\mathrm{u}}{g_\mathrm{l}}\,\frac{c^2}{8\pi\,\nu_\mathrm{lu}^2}\,A_\mathrm{ul}\,\phi_\nu\left(1 - \mathrm{e}^{-\frac{h\,\nu_\mathrm{ul}}{k_\mathrm{B}\,T_\mathrm{ex}}}\right) .
\end{aligned} \tag{I–21}$$

1.2.1 [^{12}C II] 158 μm transition

An important species for the ISM is $^{12}C^+$, which, due to the low threshold for the second ionisation state of C ($E_{\mathrm{I,C\,I}\rightarrow\mathrm{II}} \approx 11.3\,\mathrm{eV}$), exists in multiple phases of the ISM. The *fine-structure splitting* of the $1s^2 2s^2 2p^1$ ground state between $^2P^o_{3/2}$ and $^2P^o_{1/2}$ (see Figure I–3) permits the transition energy of 158 μm to be used to cool the gas. Along with the [O I] 63 μm transition, the [^{12}C II] 158 μm is one of the main coolants of hot and dense PDRs (see eg. Pabst et al. 2022; Schneider et al. 2018, 2021). Due to the ubiquity of $^{12}C^+$ in the ISM, there is also a non-negligible foreground component for these PDRs such that the dense ISM is embedded in an extended cold, diffuse atomic gas (see the analysis in eg. Guevara et al. 2020; Kabanovic et al. 2022). Evidently, the [^{12}C II] cooling line is an important line diagnostic for disentangling the WIM, WNM, and CNM. Some recent work even suggests that the [^{12}C II] line emission can be a useful tracer of molecular hydrogen in dwarf galaxies (Ramambason et al. 2023, see the discussion in §I-4).

.

1.2.2 H I 21 cm spin transition

Atomic hydrogen in the ground state $1s^1\ S_{1/2}$ has an electronic and nuclear angular momentum $J = I = 1/2$. Since the electron and proton spins can either be parallel or anti-parallel, the ground state has two energy levels and the total angular momentum can either be $F = 0$ or $F = 1$, resulting in *hyperfine-structure splitting*. This hyperfine splitting of the $1s$ ground state is primarily used to study the atomic content of the ISM. The electron's spin-flip has an energy difference of $\Delta E_\mathrm{s}/k_\mathrm{B} = 0.06816\,\mathrm{K}$, corresponding to a wavelength of 21.106 cm. For this transition, the cosmic microwave background (2.73 K) is sufficient to populate the excited state, giving,

$$\frac{n_\mathrm{u}}{n_\mathrm{l}} = \frac{g_\mathrm{u}}{g_\mathrm{l}}\,\mathrm{e}^{-\frac{h\,\nu_\mathrm{ul}}{k_\mathrm{B}\,T_\mathrm{CMB}}} \approx 3 , \tag{I–22}$$

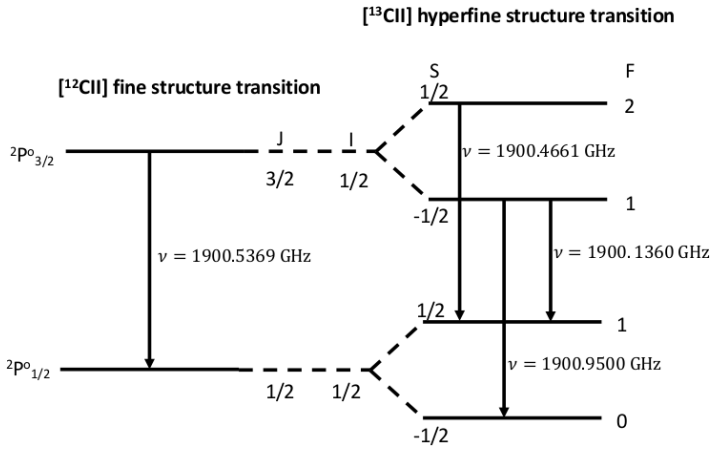


Figure I–3: The fine-structure splitting of the ground-state energy level of $^{12}C^{+}$ and $^{13}C^{+}$.

$$n_{\rm u} \approx \frac{3}{4}\, n({\rm H\,I})\ , n_{\rm l} \approx \frac{1}{4}\, n({\rm H\,I})\ , \tag{I–23}$$

and thus we can assume that H^0 is in thermal equilibrium and the excitation temperature is equal to the spin temperature $T_{\rm s}$. We can then write the emissivity coefficient as,

$$\begin{aligned} j_\nu =& n_{\rm u}\, A_{\rm ul}\, \frac{h\,\nu_{\rm ul}}{4\pi}\, \phi_\nu, \\ j_\nu =& \frac{3}{16\pi}\, n({\rm H\,I})\, A_{\rm ul}\, h\,\nu_{\rm ul}\, \phi_\nu\ . \end{aligned} \tag{I–24}$$

Applying the thermal equilibrium of the H I spin transition, we can write the absorption coefficient as,

$$\begin{aligned} \kappa_\nu &\approx n_{\rm l}\, \frac{g_{\rm u}}{g_{\rm l}}\, \frac{c^2}{8\pi\nu_{\rm lu}^2}\, A_{\rm ul}\, \phi_\nu\ \left(\frac{h\nu_{\rm ul}}{k_{\rm B}T_{\rm s}}\right), \\ \kappa_\nu &\approx \frac{3}{32\pi}\, A_{\rm ul}\, \frac{hc^2/\nu_{\rm ul}}{k_{\rm B}T_{\rm s}}\, n({\rm H\,I})\, \phi_\nu\ . \end{aligned} \tag{I–25}$$

Assuming a background intensity of 0, we calculate the optical depth and brightness temperature as,

$$\tau_\nu = \int_0^\ell {\rm d}s\, \kappa_\nu \approx \frac{3}{32\pi}\, A_{\rm ul}\, \frac{hc^2/\nu_{\rm ul}}{k_{\rm B}T_{\rm s}}\, N({\rm HI})\, \phi_\nu\ , \tag{I–26}$$

$$T_{{\rm B},\nu} \approx \begin{cases} \dfrac{c^2}{2k_{\rm B}\nu^2} \displaystyle\int_0^\ell {\rm d}s\, j_\nu = \dfrac{c^2}{2k_{\rm B}\nu^2}\, j_\nu\, \ell & N({\rm HI}) \leq 10^{20}\,{\rm cm}^{-2} \\ \dfrac{c^2}{2k_{\rm B}\nu^2}\, \dfrac{j_\nu}{\tau_\nu}\, \left(1-{\rm e}^{-\tau_\nu}\right) & \text{else} \end{cases}\ , \tag{I–27}$$

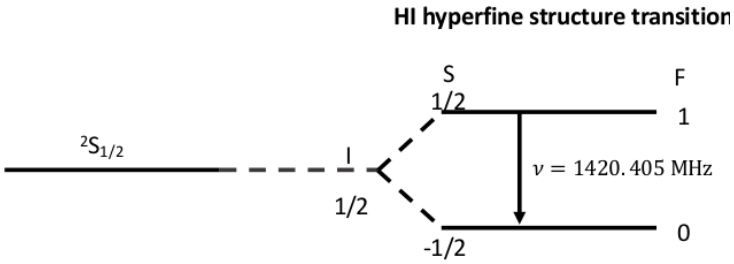


Figure I–4: The hyperfine splitting of the ground-state energy level of H^0.

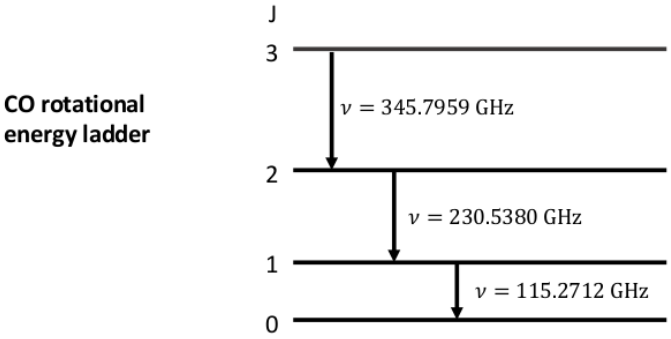


Figure I–5: The rotational energy levels of CO.

where the emissivity and absorption are assumed to be constant over the path length ℓ. The line profile ϕ_ν we want in terms of observing velocity $v_{\rm obs}$, so we use the following conversion:

$$\phi_\nu(v_{\rm obs}) = \frac{1}{\sqrt{2\pi\sigma_{\rm V}^2}} \frac{c}{\nu} \mathrm{e}^{-\frac{v_{\rm obs}^2}{2\sigma_{\rm V}^2}}. \tag{I–28}$$

1.2.3 CO transitions

One of the most important molecules for studying the molecular ISM is CO and its isotopologues, which are observed primarily by their ground electronic state rotational (J) transitions (see Figure I–5). Due to the relative lightness and abundance of its constituents as well as its critical density of $n_{\mathrm{crit},J=1\to0} \approx 10^3\,\mathrm{cm}^{-3}$, it is a convenient tracer of the physical properties and nucleosynthesis in star-forming regions (Dame et al. 2001; Bolatto et al. 2013; Leroy et al. 2021; den Brok et al. 2022). The energy difference for the lower transitions is $E_{\rm ul}/k_{\rm B} \lesssim 20\,\mathrm{K}$, so the low-$J$ levels should be sufficiently populated in molecular regions that these transitions are observable. From the abundance of CO and its low critical density, the $J = 1 \to 0$ is typically optically thick and thus radiative-trapping keeps the $J = 1$ energy level populated even at densities $n \sim 10^2\,\mathrm{cm}^{-3}$ (Heyer & Dame 2015). This means CO should be observable in both the dense and diffuse regions of the cold ISM.

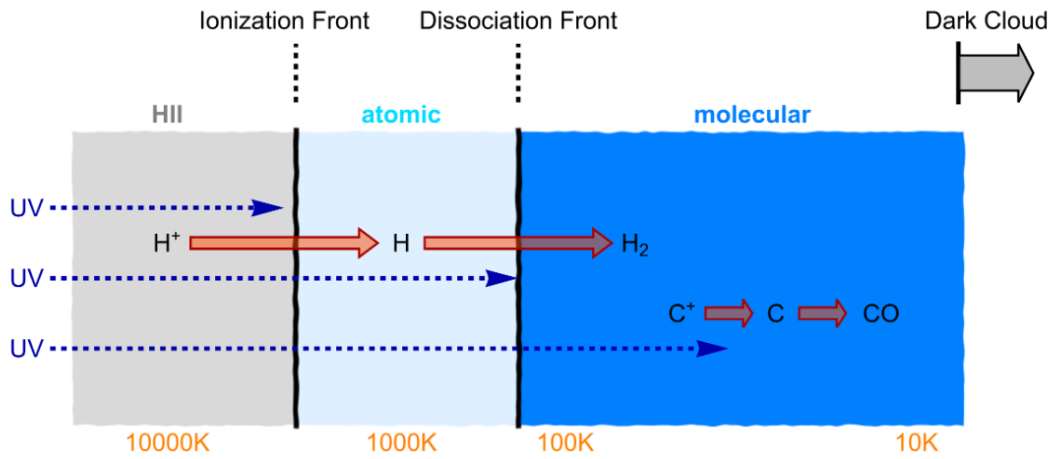


Figure I–6: The PDR structure plot from Röllig & Ossenkopf-Okada (2022).

1.2.4 Dust continuum

Interstellar dust is a catalyst for chemistry in molecular regions, as well as playing a crucial role in energy and momentum balance in the ISM. The grain-size distribution is commonly taken to be $\mathrm{d}n_a/\mathrm{d}a \propto a^{-3.5}$, where $a_\mathrm{d} \in [5-250]$nm (Mathis et al. 1977; Weingartner & Draine 2001c). Most of the incident photons are absorbed by the dust grains due to their large size compared to individual atom/molecules, and thus dust grains are important for the dynamic mechanism of radiative pressure (Weingartner & Draine 2001b). The absorption of radiation by the grain, especially the smaller grains and polycyclic aromatic hydrocarbons (PAHs), leads to photoelectric emission and contributes to the heating of the gas (Weingartner & Draine 2001a). This absorption also heats the dust grains, which is commonly though to be in thermal equilibrium with the incident radiation field. Equilibrium is a reasonable assumption since the dust grains efficiently cool by radiating as a blackbody spectrum (consider Equation I–6 with the dust grain temperature $T_{\mathrm{d,a}}$) rather than the transition lines of gaseous species. While the emission from each dust grain follows a blackbody, we observe the cumulative emission of all dust grain sizes, which is in fact broadens the SED to make a *modified-blackbody* or *greybody*:

$$I(\nu) = \tau_0 \left(\frac{\nu}{\nu_0}\right)^{\beta} B_\nu(T_\mathrm{d}) \ , \tag{I–29}$$

where τ_0 is the optical depth of the observed dust column (proportional to the dust column density), ν_0 is the reference frequency, β is the emissivity index, and T_d is the average dust temperature. At long wavelengths ($\lambda \gtrsim 250\,\mu$m), the emission should be dominated by the opacity of the dust so $\beta \approx 2$ (Draine & Lee 1984; Ossenkopf & Henning 1994). The observations from the FIRAS instrument on *COBE* indicate a lower value of $\beta \propto \lambda^{1.6}$, so Li & Draine (2001) had modified the dielectric wavelength dependence accordingly.

2 Photo-dissociation regions

This photoionisation front is the surface of the photo-dissociation region (PDR) nearest the ionising source, which is defined as the neutral region between an ionised H II region and a denser molecular cloud, where matter is irradiated with far-ultraviolet (far-UV) radiation and molecular hydrogen is disassociated. In particular it lies between the photoionisation front, where H^+ and H^0 are in equal abundance, and the photodissociation front, where H^0 and H_2 are in equal abundance by mass. These regions are ideal to test our understanding of the chemistry of the atomic medium, since their identifying feature is a stratification of species in different ionisation states (for example the transition from $^{12}C^+$ to ^{12}C to CO).

2.1 Heating

The heating of the gas and dust in PDRs is dominated by the incident far-UV radiation. Like every other PDR code, the gas heating in the ISM is provided primarily from the far-UV radiation from nearby stars ($6 - 13.8\,\mathrm{eV}$; $2066 - 912\,\text{Å}$). Commonly this interstellar radiation field (ISRF) is described by the often-cited Habing field (Habing 1968). While the original paper provided an approximate spectral density for three wavelengths (1000 Å, 1400 Å, and 2200 Å), Draine & Bertoldi (1996) provided a fit which covers the far-UV range[2].

$$\lambda u_\lambda \approx 4 \times 10^{-14}\,\mathrm{erg\,cm^{-3}} \left(-\tfrac{25}{24} \left(\tfrac{\lambda}{10^3\text{Å}}\right)^3 + \tfrac{25}{8} \left(\tfrac{\lambda}{10^3\text{Å}}\right)^2 - \tfrac{13}{12} \left(\tfrac{\lambda}{10^3\text{Å}}\right)^1 \right)$$

The energy density at 1000 Å is typically used to compare the performance of the various SEDs available in the literature, where the value stated in Habing (1968) is $u_{\mathrm{H},1000\text{Å}} = 4 \times 10^{-14}\,\mathrm{erg\,cm^{-3}}$. Thus to maintain a direct comparison with Habing (1968), it is characterised by the introduction of the parameter $\chi_{1000\text{Å}} \equiv \left(\lambda u_\lambda(1000\text{ Å})/u_{\mathrm{H},1000\text{Å}}\right)$. A more-robust calculation of the local ISRF was provided by Draine (1978). Besides being *observationally-based*, it is comparable to theoretical determinations of the ISRF. Note that λu_λ has been re-calculated from Draine & Bertoldi (1996).

$$\begin{aligned} F(E) &\approx 1.658 \times 10^6 \left(\tfrac{E}{\mathrm{eV}}\right) - 2.152 \times 10^5 \left(\tfrac{E}{\mathrm{eV}}\right)^2 \\ &\quad + 6.919 \times 10^3 \left(\tfrac{E}{\mathrm{eV}}\right)^3 \mathrm{photons\,cm^{-2}\,s^{-1}\,sr^{-1}\,eV} \\ \lambda u_\lambda &\approx 4 \times 10^{-14}\,\mathrm{erg\,cm^{-3}}\ \chi \\ &\quad \times \left(31.02609 \left(\tfrac{\lambda}{10^3\text{Å}}\right)^{-3} - 49.92879 \left(\tfrac{\lambda}{10^3\text{Å}}\right)^{-4} + 19.90300 \left(\tfrac{\lambda}{10^3\text{Å}}\right)^{-5} \right) \end{aligned} \qquad \text{(I–30)}$$

[2] In this work we adopt the formality of Draine & Bertoldi (1996) of displaying the equation as a multiple of $4 \times 10^{-14}\,\mathrm{erg\,cm^{-3}}$ and using wavelength units of 10^3 Å. The multiple is used since this is the wavelength at which Habing (1968) evaluated the spectral energy density. The wavelength units can be used to conveniently convert between wavelength and energy in eV.

$$E \approx 12.398\,\mathrm{eV} \left(\frac{\lambda}{10^3\text{ Å}}\right)^{-1}$$

Table I–1: A comparison of the various far-UV spectra discussed. Although KOSMA-τ uses a modified Draine field, $\chi_{1000\,\text{Å}}$ and $u\,u_{\text{Habing}}^{-1}$ are remain approximately the same in the 6-13.6 eV range.

spectrum	$\chi_{1000\,\text{Å}}$	$\frac{u}{u_{\text{Habing}}}$	ref
Habing	1	1	1,3
Draine	1.71	1.69	2,3
Mezger et al.	1.23	1.14	4,5
Zucconi et al.	1.05	0.92	6
KOSMA-τ	1.71	1.69	7

(1) Habing (1968); (2) Draine (1978); (3) Draine & Bertoldi (1996); (4) Mezger et al. (1982); (5) Mathis et al. (1983); (6) Zucconi et al. (2001); (7) Röllig et al. (2013).

Another observationally-based characterisation of the ISRF was made by smoothing and fitting the observations in Gondhalekar et al. (1980) with a broken-power law (Mezger et al. 1982; Mathis et al. 1983).

$$\lambda u_\lambda \approx 4 \times 10^{-14}\,\text{erg}\,\text{cm}^{-3} \begin{cases} 0.5932\left(\frac{\lambda}{10^3\text{Å}}\right)^{-0.6678} & 1340\text{Å}<\lambda<2460\text{Å} \\ 17.06\left(\frac{\lambda}{10^3\text{Å}}\right) & 1100\text{Å}<\lambda<1340\text{Å} \\ 32180\left(\frac{\lambda}{10^3\text{Å}}\right)^{4.4172} & 912\text{Å}<\lambda<1100\text{Å} \end{cases} \tag{I–31}$$

One could also describe the ISRF as a superposition of blackbody spectra. This was calculated in Zucconi et al. (2001) for different regimes of the ISRF. A summary of the χ and total energy density ($u = \int u_\lambda d\lambda$) ratio is given in Table I–1.

For KOSMA-τ, however, a modified Draine spectrum is used. Equation I–30 is used for $\lambda < 2000$ Å, but the relation in van Dishoeck & Black (1982) is used for $\lambda \geq 2000$ Å (Röllig et al. 2013).

$$\lambda u_\lambda \approx 4.61438 \times 10^{-14} \left(\frac{\lambda}{10^3\text{Å}}\right)^{0.7} \quad \lambda \geq 2000\text{Å} \tag{I–32}$$

A comparison of all of these spectra are shown in Figure I–7.

2.2 Cooling

As mentioned previously, the most efficient coolant of the ISM is the interstellar dust since it gains and releases energy via the blackbody spectrum (Klessen & Glover 2016). A typical PDR (ie. one with high incident radiation χ) releases most of its energy via line cooling with the CO and ^{13}CO rotational transitions, the [^{12}C II] 158 μm transition, and the [O I] 63 μm and 146 μm transitions. Other mechanisms for cooling include recombination of electrons with charged dust grains, ^{12}C I line emission, and H_2 photo-dissociation.

Even though it is an efficient channel for cooling, the continuum emission is not the most intense observable in star forming regions since the energy is distributed throughout the entire spectrum. Since they are optically thin in PDR conditions,

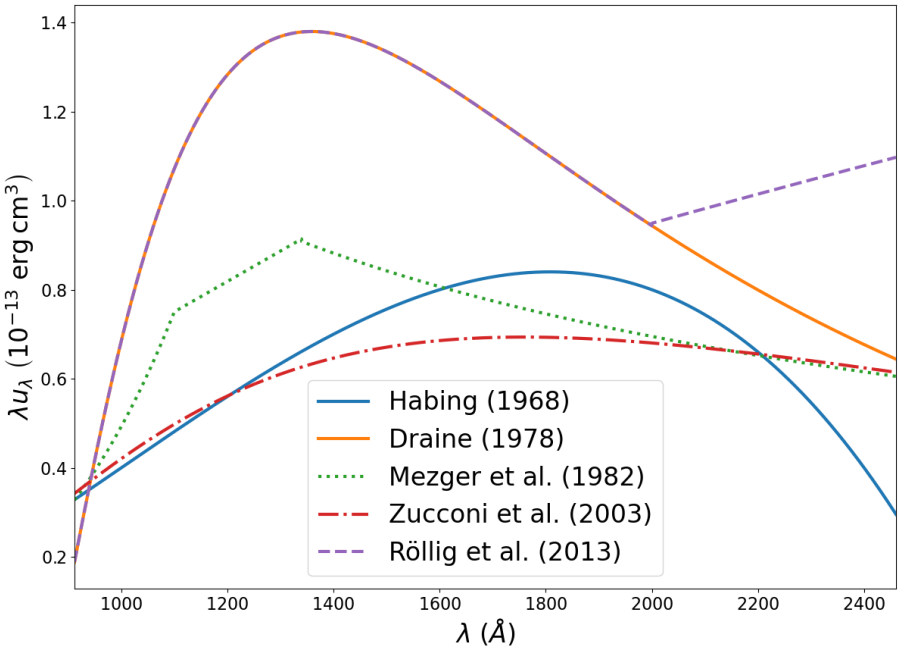


Figure I–7: The various SEDs considered in the literature compared to KOSMA-τ. See Table I–1 for the χ and relative energy density.

the [^{12}C II] and [O I] lines are typically the brightest transitions in diffuse regions. The intense rotational transitions of CO and ^{13}CO make them efficient coolants of the high density molecular region of PDRs ($n_\mathrm{H} \gtrsim 10^4\,\mathrm{cm}^{-3}$). Radiative trapping in the quickly optically-thick low-J CO lines reduces the efficiency of these coolants at densities $n_\mathrm{H} \gtrsim 10^3\,\mathrm{cm}^{-3}$, but increases their efficiency down to $n_\mathrm{H} \sim 10^2\,\mathrm{cm}^{-3}$.

2.3 Structure

It is common (since Hollenbach & Tielens 1997) to view the structure of a PDR along one dimension, with the ionising radiation on the left and the molecular cloud on the right. PDRs are typically the neutral ISM between the ionisation front[3] (where the number of protons in ionised and atomic hydrogen are equal) and the dissociation front (where the number of protons in atomic and molecular hydrogen are equal). Since the far-UV radiation dominates the heating and chemistry, its attenuation as it propagates deeper into the cloud results in a stratification of different chemical chains (such as for carbon or the aforementioned hydrogen). As we identified in §I-1.2.1, the lower ionisation energy of carbon compared to hydrogen results in most ^{12}C existing in the form of ^{12}C$^+$ in the molecular region. This has significant implications for derivations of the H_2-column density, since this is primarily traced via its collisions with ^{12}CO(see the discussion of tracer

[3]The ionised gas in the H II region is optically-thick to photons with $h\nu \geq 13.6\,\mathrm{eV}$, so photo-ionisation events balance recombination events and there is a clear ionisation front.

particles in §I-4).

2.4 PDR modelling

Models have been developed for decades to balance the heating and cooling rates of PDRs. Most have been developed to model the aforementioned PDR structure, that is, in a one-dimensional, plane-parallel configuration irradiated by one or both sides of the PDR such as `Cloudy` (Ferland et al. 2017), `UCL_PDR` Bell et al. (2006b,a), and the Meudon PDR code (Goicoechea & Bourlot 2007; Le Petit et al. 2006), while others have opted for spherical symmetry with isotropic irradiation such as the KOSMA-τ model (Röllig & Ossenkopf-Okada 2022; Röllig & Ossenkopf 2013; Röllig et al. 2013, 2006; Gierens et al. 1992). The advent of these configurations have allowed for the development of self-consistent PDR models with sophisticated chemical networks while being computationally feasible. Although the precise purpose of each code may differ (for example `Cloudy` is a general spectral synthesis and plasma simulation code), the community has gathered in the past to ensure comparable results, although there remains a non-negligible difference in the atomic fine-structure line intensities (see a comparison between PDR models in Röllig et al. 2007).

3 Inhomogeneity

For decades one-dimensional PDR models have served to model the PDR cooling lines and constrain the ISM properties. It has been known for quite some time, however, that these models are inadequate to explain the observed ratio of atomic to molecular regions (traced with [^{12}C II] / ^{12}CO $J = 1 \rightarrow 0$; Stutzki & Guesten 1990). To that end, it was suggested that it is useful to model the ISM structure as a fractal distribution of self-consistent ISM *clumps*. This inhomogeneity is thought to originate from the dissipation of turbulent energy in a cascade from galactic-scale dynamics down to the size of clumps and eventually to the diffuse interclump medium (see further discussion of this in §I-4 Klessen & Glover 2016; Ossenkopf & Mac Low 2002; Ossenkopf 2004). The modelling of this inhomogeneities has proceeded in two general directions: self-consistent three-dimensional modelling and fractal post-processing. The former direction is the

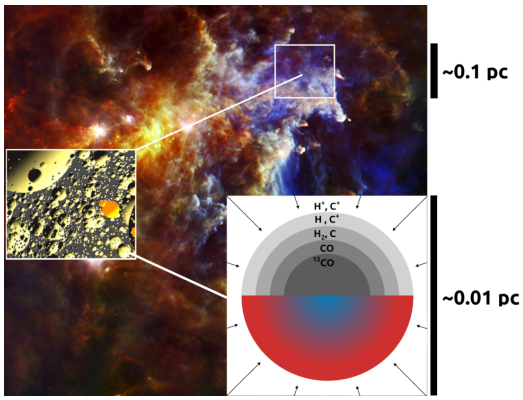


Figure I–8: The Rosette H II region with the PDR and relevant scales identified. Image adopted from Figure I–1 and a graphic created of a clumpy ensemble (from `https://markusroellig.github.io/research/Clumpy-PDR/`). We try to mimic the low-resolution image available at `https://astro.uni-koeln.de/riechers/research/kosma-tau` while also including very approximate values of the relative scale of the clumps in the ISM.

most-accurate, of course, since it ensures thermal and chemical balance throughout the inhomogeneous medium. While this has been successfully been developed from the UCL_PDR model for one, two, and three dimensions, this does come at the cost of chemical complexity (Bisbas et al. 2012). Modelling of the fractal ISM, however, has proceeded in two and three dimensions (§7 in Röllig & Ossenkopf-Okada 2022; Andree-Labsch et al. 2017). These models are able to keep the complexity of the chemical networks of the one-dimensional models by assuming there is self-similarity between the PDR clumps, allowing the modelling of unresolved inhomogeneity. This is the method used in the remainder of this thesis.

4 Galaxies

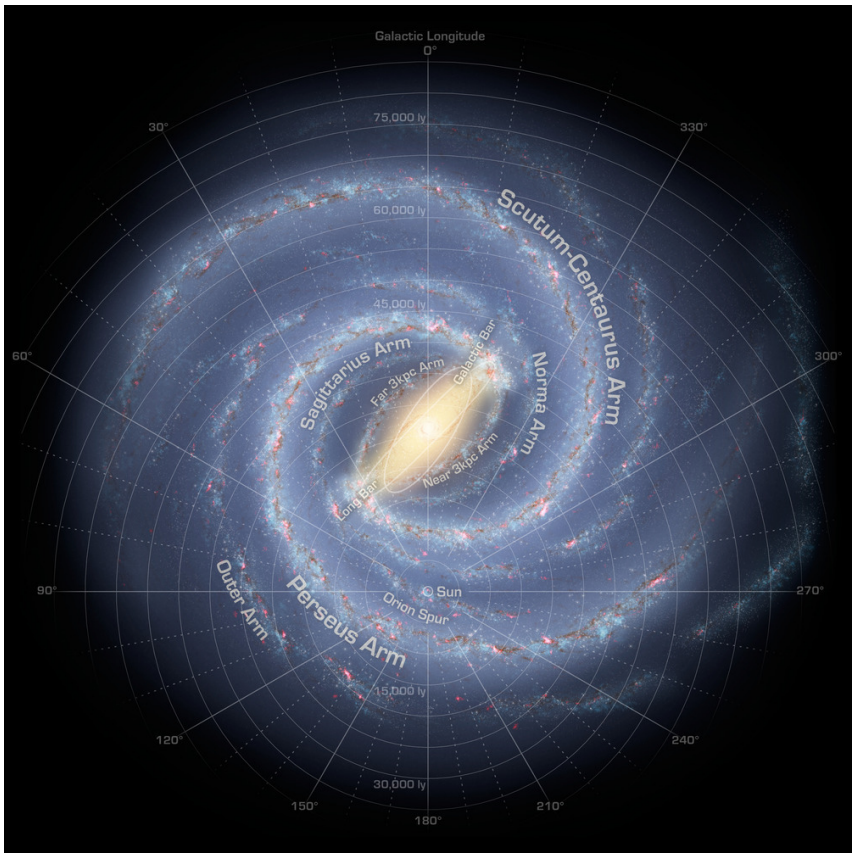


Figure I–9: A representation of the spiral arms of the Milky Way (credit: NASA/JPL-Caltech/R. Hunt).

While galaxy-scale surveys of atomic hydrogen line transitions have existed for several decades (van de Hulst et al. 1954; Kerr et al. 1954), molecular hydrogen was not observed until 20 years later (Carruthers 1970). Due to the difficulty in observing a symmetric diatomic molecule, many observations of the molecular ISM on a galactic scale rely on the rotational transitions of abundant asymmetric *tracer* molecules (den Brok et al. 2021, 2022; Schuller et al. 2021, 2017; Barnes et al. 2015;

Braiding et al. 2015; Dame et al. 2001, 1987). These have proven to be exceptionally powerful in terms of understanding the large-scale structure of the galactic ISM since these lines are not scattered as much as the ultraviolet line transitions of H_2. Indeed we typically use the intensity of collisionally-excited transitions of these tracers as a metric to quantify the column-density of molecular hydrogen. We quantify this relation between $\varpi_{\rm tr}$ and $\mathcal{N}(H_2)$ using the conversion factor $X_{\rm tr}$:

$$X_{\rm tr} = \frac{\mathcal{N}(H_2)}{\varpi_{\rm tr}} \ . \tag{I–33}$$

The most common tracer for H_2 is ^{12}CO due to its ubiquity in the ISM. The conversion factor in the disk of the Milky Way seems to be relatively constant with galactocentric radius, however the precise value changes by a factor of 4 depending on how it is derived (for example using a viral analysis, extinction, dust emission, CO isotopologues, etc.). the average ^{12}CO conversion factor is $X_{^{12}\mathrm{CO,MW}} \approx 2\times10^{20}\,\mathrm{cm^{-2}}\,(\mathrm{K\,km\,s^{-1}})^{-1}$ (Narayanan et al. 2012; Bolatto et al. 2013).

Another useful quantity for quantifying the conditions across a galaxy is the CO line ratio, $R_{\rm ul}$. This is defined as the ratio between the line-integrated intensity from an upper (u) and lower (l) transition:

$$R_{\rm ul} = \frac{\varpi_{\rm u}}{\varpi_{\rm l}} \ . \tag{I–34}$$

The most common ratio to calculate is R_{21}, the ratio between the ^{12}CO $J = 1 \rightarrow 0$ and $J = 2 \rightarrow 1$ integrated intensities den Brok et al. (2021). Since analyse spectral lines using the brightness temperature (see Equation I–9), the line ratio $R_{\rm ul} = 1$ when in local thermal equilibrium. It is thus a useful metric for the ISM thermal properties.

When discussing the clumpy, inhomogeneous nature of a galaxy, it is important to consider how this inhomogeneity is initially created. Plenty of theoretical research has been devoted to the composition of and pressure balance in the ISM in a galactic context (Barnes et al. 2020; McKee & Ostriker 1977; Field et al. 1969) and the existence of spiral arms in grand design galaxies (such as the spiral arms in Figure I–9; Lin & Shu 1964, 1966; Yuan 1969; Elmegreen et al. 1989). This has been complemented by modelling of the dynamic nature of a galaxy in star-forming regions (Elmegreen & Falgarone 1996; Elmegreen 1999b,a; Ossenkopf & Mac Low 2002; Ossenkopf 2004). Generally, a theory has developed that kinetic energy due to large-scale dynamics (originating for example from galactic rotation) cascades down to smaller scales where it is dissipated as heat (Mac Low & Klessen 2004; Klessen & Glover 2016). It is in this cascading regime that there is self-similarity between different scales, motivating the fractal approach to modelling the ISM that is explained in Chapter II.

CHAPTER II

Theory

Science means constantly walking a tightrope between blind faith and curiosity; between expertise and creativity; between bias and openness; between experience and epiphany; between ambition and passion; and between arrogance and conviction - in short, between an old today and a new tomorrow.

Heinrich Rohrer

The discrepancies in line ratios between observations and models reported in papers such as Stutzki & Guesten (1990); Meixner & Tielens (1993); Schneider et al. (2021) indicate PDR modelling in its basic, homogeneous approach should be developed to account for the inhomogeneities mentioned in §I-3. The method adopted in this work is to represent the inhomogeneous ISM by a discrete fractal distribution of self-similar PDR *clumps*. The structure of the KOSMA-τ PDR model is ideal for implementation into the fractal distribution. The theoretical backbone of `kosmatau3d` has been developed to some extent by Cubick et al. (2008) and Andree-Labsch et al. (2017), though there were a number of assumptions that needed to be corrected. This chapter thus is part review and part novel research. In the progression of the theory, one must keep in mind that each ensemble of clumps is contained within a *voxel*: a cubic element with three spatial dimensions[4]. Although the clumps themselves are unresolved within each voxel, they contribute to the emissivity and absorption coefficients of their host voxels.

We will first introduce the some basic properties of the KOSMA-τ PDR model and how to calculate the extrinsic clump-averaged properties. Then we will discuss fractal ensembles in §II-2. All of the advancements in confining them within a voxel are described in §II-3.

[4]This is the definition of a voxel in terms of three-dimensional computer graphics. It is easy to think of voxels as VOlume-piXELs. Be aware that this is different from *spaxels*, which have two spatial dimensions and one spectral dimension. In this sense three-dimensional models are made of several voxels, while synthetic maps are composed of many spaxels.

1 KOSMA-τ clumps

`kosmatau3d` is essentially a framework in which the clump-averaged results from KOSMA-τ are post-processed to compute their superposition in a clumpy medium. Thus it is necessary to review some of the relevant information pertaining to these models. One of the properties affecting further theoretical development is the density profile of the clump, since this affects properties such as emissivity and absorption (see §I-1.2). Assuming a surface total hydrogen density of $n_{\mathrm{s,H}}$, the profile is defined as,

$$n_{\mathrm{cl}}(r) = n_{\mathrm{s,H}} \begin{cases} \left(\frac{r}{r_{\mathrm{cl}}}\right)^{-1.5} & r_{\mathrm{core}} < r < r_{\mathrm{cl}} \ , \\ \left(\frac{r_{\mathrm{core}}}{r_{\mathrm{cl}}}\right)^{-1.5} & r < r_{\mathrm{core}} \ , \end{cases} \tag{II–1}$$

where we use a clump core of $r_{\mathrm{core}} = 0.2 r_{\mathrm{cl}}$.

Clump radius

Since the density profile of each clump is given by $\rho_{\mathrm{cl}} = m_{\mathrm{H}} n_{\mathrm{cl,H}}$, we can derive the clump-averaged number density as,

$$\begin{aligned}
\overline{n_{\mathrm{cl}}} &= \frac{\int_V \mathrm{d}V \, n_{\mathrm{cl,H}}(r)}{\int_V \mathrm{d}V} \\
\overline{n_{\mathrm{cl}}} &= \frac{3\, n_{\mathrm{s,H}}}{r_{\mathrm{cl}}^3} \left(\int_0^{r_{\mathrm{core}}} \mathrm{d}r \, r^2 \left(\frac{r_{\mathrm{core}}}{r_{\mathrm{cl}}}\right)^{-1.5} \right. \\
&\qquad \left. + \int_{r_{\mathrm{core}}}^{r_{\mathrm{cl}}} \mathrm{d}r \, r^2 \left(\frac{r}{r_{\mathrm{cl}}}\right)^{-1.5} \right) \\
\overline{n_{\mathrm{cl}}} &= \frac{3\, n_{\mathrm{s,H}}}{r_{\mathrm{cl}}^3} \left(\frac{r_{\mathrm{core}}^3}{3} \left(\frac{r_{\mathrm{core}}}{r_{\mathrm{cl}}}\right)^{-1.5} + \int_{r_{\mathrm{core}}}^{r_{\mathrm{cl}}} \mathrm{d}r \, \frac{r^{0.5}}{r_{\mathrm{cl}}^{-1.5}} \right) \\
\overline{n_{\mathrm{cl}}} &= \frac{n_{\mathrm{s,H}}}{r_{\mathrm{cl}}^{1.5}} \left(r_{\mathrm{core}}^{1.5} + 2\, r_{\mathrm{cl}}^{1.5} - 2\, r_{\mathrm{core}}^{1.5} \right) \\
\overline{n_{\mathrm{cl}}} &= n_{\mathrm{s,H}} \left(2 - \frac{r_{\mathrm{core}}^{1.5}}{r_{\mathrm{cl}}^{1.5}} \right) \qquad \text{(II–2)} \\
\overline{n_{\mathrm{cl}}} &\approx 1.911\, n_{\mathrm{s,H}} \qquad \text{(II–3)}
\end{aligned}$$

where we have used the $r_{\mathrm{core}} = 0.2 r_{\mathrm{cl}}$ for the approximation in Equation II–3. From there we can derive the clump radius r_{cl} as,

$$r_{\rm cl}^3 = \frac{m_{\rm cl}}{\overline{\rho_{\rm cl}}\, V} \ ,$$
$$r_{\rm cl} \approx \sqrt[3]{\frac{3 m_{\rm cl}}{1.911 \times 4\pi\, m_{\rm H}\, n_{\rm s,H}}} \ . \qquad \text{(II–4)}$$

The intrinsic clump properties (mass, density, radius) are not the only properties of interest; for the intended use of `kosmatau3d` we really require clump-averaged extrinsic properties. Such properties include the column density $\mathcal{N}(\rm sp)$ for a given species sp, line intensity I_ν and optical depth τ_ν for a given wavelength ν, and visual optical depth $\tau_{\rm V}$, which are calculated as a function of impact parameter p. In these cases, the clump-averaged properties are,

$$\langle \mathcal{E} \rangle_{\rm cl} = \frac{\int_A {\rm d}A \mathcal{E}(p)}{\int_A {\rm d}A} \qquad \text{(II–5)}$$
$$\langle \mathcal{E} \rangle_{\rm cl} = \frac{2\pi}{\pi r_{\rm cl}^2} \int_0^{r_{\rm cl}} {\rm d}p\, p\, \mathcal{E}(p), \qquad \text{(II–6)}$$

for some extrinsic property $\mathcal{E}(p)$ that can be integrated over the impact parameter. Since Equation II–6 is essentially a weighted average of the extrinsic property, we must carefully consider how this weighting is applied to the property. The optical depth, for example, is a logarithmic term, therefore it must be weighted as an exponential function. We can thus write the four aforementioned properties as,

$$\langle \mathcal{N}({\rm sp}) \rangle_{\rm cl} = \frac{2}{r_{\rm cl}^2} \int {\rm d}p\, p\, \mathcal{N}({\rm sp}, p) \ , \qquad \text{(II–7)}$$
$$\langle I_\nu \rangle_{\rm cl} = \frac{2}{r_{\rm cl}^2} \int {\rm d}p\, p\, I_\nu(p) \ , \qquad \text{(II–8)}$$
$$\langle \tau_\nu \rangle_{\rm cl} = -\ln\left(\frac{2}{r_{\rm cl}^2} \int {\rm d}p\, p\, {\rm e}^{-\tau_\nu(p)} \right) \ , \qquad \text{(II–9)}$$
$$\langle \tau_{\rm V} \rangle_{\rm cl} = -\ln\left(\frac{2}{r_{\rm cl}^2} \int {\rm d}p\, p\, {\rm e}^{-\tau_{\rm V}(p)} \right) \ . \qquad \text{(II–10)}$$

For the remainder of this thesis, extrinsic properties with the subscript "cl" will refer to the clump-averaged property.

1.1 Additional properties

There are still some properties we wish to use in `kosmatau3d` which must be derived from the KOSMA-τ output[5,6]. In these outputs we have local properties (gas

[5] These outputs are available at `https://ismdb.obspm.fr`. It is planned to implement a function in `kosmatau3d` to query directly from ISMDB, but this is not yet possible. For assistance in querying ISMDB using `python`, contact the author

[6] The KOSMA-τ data can be loaded into `python` by using the scripts from Aleena Baby: `https://github.com/KOSMAsubmm/HDF4_reader_for_KOSMA_tau`

temperature T_g, dust temperature T_d, fractional abundances $X(\mathrm{sp})$ for species sp, etc.) as a function of depth (either visual extinction A_V or distance r). We want to derive some extrinsic properties in order to calculate the clump-averaged emissivity and absorption coefficients, since they are not provided consistently in the model. We can use the quantities we have already defined to calculate the emissivity using Equation I–27 and Equation I–3. In order to calculate the H^0 21 cm spin transition absorption, however, we assume thermal equilibrium (so $T_s = T_{kin}$) and thus require the gas temperature T_{gas}. For this we calculate the clump-averaged gas temperature by performing a weighting function:

$$T_{gas} = \frac{\int_0^{R_{cl}} \mathrm{d}r\, r^2\, T_{gas}(r)}{\int_0^{R_{cl}} \mathrm{d}r\, r^2}. \tag{II–11}$$

Equation II–11 can be used with Equation I–26 and Equation II–4 in order to calculate the clumpy-averaged optical depth.

It is also informative for us to calculate the intrinsic property of the mass of a particular species contained in each clump. In practice this is possible for every species contained in KOSMA-τ, but in this work we focus on H^0 and H_2 to examine the atomic and molecular ISM in the Galaxy[7]. The mass is determined using Equation II–7 and the clump radius:

$$M_{cl,sp} = N_{cl}(\mathrm{sp})\, \pi\, r_{cl}^2 m_{sp}, \tag{II–12}$$

where m_{sp} is the mass of the species.

2 Fractal approach

The fractal approach makes the fundamental assumption that the ISM can be represented by an ensemble of self-similar non-interacting[8] spherically-symmetric clumps. This ensures there is a higher surface-to-volume ratio in the model in order to better-agree with observations. As mentioned in §I-3, we assume a mass-spectrum of,

$$\frac{\mathrm{d}N_{cl}}{\mathrm{d}m_{cl}} \propto m_{cl}^{-\alpha}, \tag{II–13}$$

and a clump-radius relation of,

$$m_{cl} \propto r_{cl}^{\varpi}. \tag{II–14}$$

By assuming some total mass of the ensemble M_{ens} and limiting individual clump masses to be $m_l \leq m_{cl} \leq m_u$, we can write the total number of clumps as,

[7] This information has only been implemented for these two species. The rest of the available species will be implemented at a later date.

[8] This means we do not consider any merging of clumps. This is not an issue anyways since we have no time dependence. Rather the emissivity and absorption coefficient of each clump contribute to the voxel-averaged emissivity and absorption coefficient.

$$N_{\rm ens} = M_{\rm ens} \frac{\int_{m_{\rm l}}^{m_{\rm u}} {\rm d}m\, m_{\rm cl}^{-\alpha}}{\int_{m_{\rm l}}^{m_{\rm u}} {\rm d}m\, m^{-\alpha+1}}, \tag{II–15}$$

$$N_{\rm ens} = M_{\rm ens} \frac{2-\alpha}{1-\alpha} \frac{m_{\rm cl,u}^{-\alpha+1} - m_{\rm cl,l}^{-\alpha+1}}{m_{\rm cl,u}^{-\alpha+2} - m_{\rm cl,l}^{-\alpha+2}}. \tag{II–16}$$

For our purposes, since we do not have infinite computational resources, we want to derive a *discrete approximation* of these relations. We simplify the calculation by replacing the integral by a sum By assuming a total mass for the ensemble, we can directly calculate the number N_j of each clump mass:

$$N_j = M_{\rm ens} \frac{m_j^{-\alpha+1}}{\sum_n m_n^{-\alpha+2}}. \tag{II–17}$$

It naturally follows that the mass represented by the discrete clump mass m_j is,

$$M_j \equiv N_j\, m_j = M_{\rm ens} \frac{m_j^{-\alpha+2}}{\sum_n m_n^{-\alpha+2}}, \tag{II–18}$$

which ensures that,

$$M_{\rm ens} \approx \sum_j M_j. \tag{II–19}$$

We note that Equation II–19 is an approximate relation due to the discretisation of mass into clumps, primarily that N_j is an integer. Thus the difference between the discretised ensemble mass and the true ensemble mass is the *discretisation error*.

The previous section covered how to derive each clump's radius r_j, so we can then calculate the volume occupied by the clumps by,

$$V_{\rm ens} = \sum_j N_j\, v_j, \tag{II–20}$$

$$V_{\rm ens} = \sum_j N_j\, \frac{4\,\pi}{3}\, r_j^3,$$

$$V_{\rm ens} \propto \frac{4\,\pi}{3}\, M_{\rm ens} \frac{\sum_j m_j^{3/\gamma-\alpha+1}}{\sum_n m_n^{-\alpha+2}}, \tag{II–21}$$

where we have used Equation II–14 and Equation II–17[9]. We can write the average hydrogen proton density ($n_{\rm H} \equiv n_{\rm H^0} + n_{\rm H_2}$; we write $\overline{n_{\rm ens}}$ for the ensemble and $\overline{n_j}$ for individual clumps) by,

[9]Practically we use definition Equation II–19 to ensure the modelled mass is consistent and avoid using extrinsic values.

$$\frac{V_{\rm ens}}{v_j} = M_{\rm ens}\,\frac{\sum_n m_n^{3/\gamma-\alpha+1}}{m_j^{3/\gamma}\sum_n m_n^{-\alpha+2}},$$
$$\overline{n_j} = \overline{n_{\rm ens}}\,m_j^{1-3/\gamma}\,\frac{\sum_n m_n^{3/\gamma-\alpha+1}}{\sum_n m_n^{-\alpha+2}}, \tag{II–22}$$

where we have assumed the volume-size relation is valid for both the ensemble and the clump. We can convert to the surface density using Equation II–4. We henceforth refer to the surface hydrogen proton density as n_j and the ensemble-averaged hydrogen proton density as $n_{\rm ens}$.

3 Voxel approach

One advantage of the application of clumpyness in `kosmatau3d` is that we consider the most-probable number of clumps in a line-of-sight. In order to do so, we must stipulate that the ensemble(s) are contained with cubic regions of the ISM, visualised inside as voxel. Each voxel has a side-length of $\ell_{\rm vox}$, and we can thus define the volume-filling-factor, $f_{\rm V}$, for each voxel as,

$$f_{\rm V} = \frac{V_{\rm ens}}{\ell_{\rm vox}^3}. \tag{II–23}$$

Now we must address the advantage of `kosmatau3d` over the clumpy prescription used in previous work such as Cubick et al. (2008) and chapter 7 of Röllig & Ossenkopf-Okada (2022): the inclusion of velocity dispersion as a parameter. Besides `kosmatau3d`, the clumpy PDRs using KOSMA-τ in the literature have used the integrated intensities. In that sense one can successfully compare to a distribution of clumps filling, but not overlapping, the intensity map.

3.1 Velocity dispersion

The clumps contained in each ensemble are moving with respect to the observer. Formally we model this using three types of velocity dispersion that are relevant for the clumps contained within a voxel:

- σ_j — velocity dispersion intrinsic to the gas contained within a clump ($\sigma_{\rm cl}$) with mass index j
- $\sigma_{\rm sys}$ — systematic velocity dispersion between clumps in an ensemble
- $\sigma_{\rm ens}$ — velocity dispersion of the gas contained in an ensemble

These are assumed to describe the superposition of Gaussian profiles. The line profile of each line transition from the voxel will have a full-width at half-maximum (FWHM) corresponding to $\sigma_{\rm ens}$, which is a linear combination of σ_j and $\sigma_{\rm sys}$:

$$\sigma_{\rm ens}^2 \equiv \sigma_{\rm sys}^2 + \sigma_j^2. \tag{II–24}$$

For navigating the velocity-dependence of our calculations, it is useful to define the relevant velocities considered in our formalisation:

- $v_{\rm obs}$ — the observing velocity of the synthetic observation; this corresponds to the radial velocity of the emitting species with respect to the observer.
- $v_{\rm ens}$ — the mean radial velocity of the ensemble. This is equivalent to the observing velocity where the emissivity and absorption are maximum.
- v_i — the *internal* systematic velocity spanning $v_{\rm ens} \pm 3\,\sigma_{\rm sys}$; this is the velocity bin used to discretise the calculation in velocity-space. Each v_i is assumed to be the mean radial velocity for a subset of the clumps with respect to the observer.
- Δv_i — the spacing in internal velocities; in practice this is calculated as $\min\left(\sigma_j/3, \sigma_{\rm ens}/3\right)$ in order to optimise accuracy versus computational efficiency.
- v_i' — a velocity bin like v_i, though it is used to but used to account for the contribution of all velocity bins for each v_i.

Clumps of different masses may have different velocity dispersion, so the quantities which are linear and velocity-dependant along the line-of-sight from each clump (such as column density, emissivity, and absorption) will follow the profile $E_{j,i,i'}$,

$$E_{j,i,i'} = \exp\left(-\frac{(v_i - v_i')^2}{2\,\sigma_j^2}\right), \tag{II–25}$$

where for a particular velocity bin v_i we can see a contribution to all other velocity bins v_i'. One of the benefits of `kosmatau3d` is that we can have a velocity-based calculation of the intensity and optical depth, which can be used to increase the velocity dispersion of the ISM modelled in the voxel compared to single clumps. Since none of the clumps in `kosmatau3d` interact except through radiative transfer, we can discretise the ensemble in velocity-space (assuming $\Delta v_i \lesssim \sigma_{\rm ens}$) by creating separate ensembles covering the velocity dispersion we want modelled. We thus have the total number of a specific clump mass in the ensemble following $N_{j,\rm vox} = \sum_i N_{j,i}$, where $N_{j,i}$ is the number of clumps with mass M_j at mean velocity v_i:

$$\Delta N_{j,i} = \frac{N_j}{\sqrt{2\,\pi\,\sigma_{\rm sys}^2}} \exp\left(-\frac{\left(v_{\rm ens} - v_i\right)^2}{2\,\sigma_{\rm sys}^2}\right) \Delta v_i, \tag{II–26}$$

which, as mentioned above, is valid within $3\sigma_{\rm sys}$ to account for 99% of the ensemble mass. This is the velocity discretisation as it appears in Andree-Labsch et al. (2017). Since we want to calculate a superposition of clumps, we want to simplify the calculation since given the theory thus-far in we would need to account the contribution from Equation II–25 from each internal velocity at each observing velocity.

We can simplify this calculation by performing an integral over v_i. First we need Equation II–26 to be independent of the internal velocity spacing Δv_i,

we can notice that Equation II–26 can be rearranged to provide a discrete change of clump number with velocity. This value depends on the step size of the internal velocities Δv_i. Taking the limit of this change as Δv_i approaches 0 will remove this dependence:

$$\begin{aligned}
\frac{\mathrm{d}N_{j,i}}{\mathrm{d}v_i} &= \lim_{\Delta v_i \to 0} \left(\frac{\Delta N_{j,i}}{\Delta v_i} \right), \\
\frac{\mathrm{d}N_{j,i}}{\mathrm{d}v_i} &= \frac{N_j}{\sqrt{2\,\pi\,\sigma_{\mathrm{sys}}^2}} \exp\left(-\frac{(v_{\mathrm{ens}} - v_i)^2}{2\,\sigma_{\mathrm{sys}}^2} \right), \\
\frac{\mathrm{d}N_{j,i}}{\mathrm{d}v_i} &\equiv A_{j,i},
\end{aligned} \tag{II–27}$$

then we need to integrate over the domain of v_i, which we now take as $(-\infty, \infty)$. If we have a quantity $\mathcal{Q}_{j,i}$ with associated clump-averaged values q_j that is linear along the line-of-sight, we can write it as,

$$\mathcal{Q}_{j,i} = k_{j,i} \times q_j \times \int \mathrm{d}v_{i'}\, E_{j,i,i'}, \tag{II–28}$$

where $k_{j,i}$ is the *j*th mass index and *i*th velocity index of some combination of clumps based on $A_{j,i}$ (explained in §II-3.2) and the integral over $v_{i'}$ of Equation II–25 accounts for the contribution of clumps at different internal velocities (thus accounting for the superposition in velocity space). We now encode this superposition in velocity-space in the calculation of the number of clumps $N_{j,i}$. What makes this difficult to calculate efficiently is that it is a superposition of two velocity distributions: σ_{j} from Equation II–25 and σ_{sys} from Equation II–26. Then the final calculation must result in the desired velocity dispersion of σ_{ens}. Since we want the model to be $> 99\%$ accurate, we want to consider clump ensembles spanning a minimum interval of $[v_{\mathrm{vox}} - 3\sigma_{\mathrm{sys}}, v_{\mathrm{vox}} + 3\sigma_{\mathrm{sys}}]$. Numerically, this results the number of clumps at a particular systematic velocity effectively becoming,

$$\begin{aligned}
N_{j,i,\mathrm{eff}} &= \sum_{i'} \Delta N_{j,i} \exp\left(-\frac{(v_i - v_i')^2}{2\,\sigma_j^2} \right), \\
N_{j,i,\mathrm{eff}} &= \int_{-\infty}^{\infty} \mathrm{d}v_i' \frac{\mathrm{d}N_{j,i}}{\mathrm{d}v_i} \exp\left(-\frac{(v_i - v_i')^2}{2\,\sigma_j^2} \right), \\
N_{j,i,\mathrm{eff}} &= N_j \frac{\sigma_j}{\sigma_{\mathrm{ens}}} \exp\left(-\frac{(v_{\mathrm{vox}} - v_i)^2}{2\,\sigma_{\mathrm{ens}}^2} \right),
\end{aligned} \tag{II–29}$$

where there is now no more coupling between velocity bins (we already accounted for the contribution from other velocity bins), and we set $v_{\mathrm{ens}} = v_{\mathrm{vox}}$ for simplicity in the calculations[10]. Since Equation II–29 is now the effective number of clumps

[10] Practically it is possible to have two ensembles with different v_{ens} in a voxel, but this requires further justification that the ensembles are not collisional. `kosmatau3d` by design does not consider any mechanical interaction between clumps in a voxel.

at velocity v_i regardless of bin size (the superposition is already accounted for), in order to reproduce the line profile we must give it the profile velocity dispersion σ_{ens}.

The integration performed in the final equation means we no longer have to account for the line profile in the superposition (there is an effective velocity dispersion of $\sigma_{j,\text{eff}} = 0$ for the clumps in a particular velocity bin), making $\sigma_{\text{sys,eff}} = \sigma_{\text{ens}}$. Since we removed the need to sum over the internal velocities, we can write $N_{j,\,i,\text{eff}} \equiv N_{j,\text{eff}}\left(v_{\text{obs}}\right)$ for $v_{\text{obs}} = v_i$.

3.2 Probabilistic calculation

In order to calculate voxel-averaged quantities, `kosmatau3d` uses a probabilistic approach (Andree-Labsch et al. 2017). Essentially it utilises the probability of having a certain number of clumps in a line of sight to calculate the average of a given quantity. This is valid for any linear quantity that is additive. The relevant clump-averaged quantities for this work which are linear along the line of sight (and thus we can substitute for q_j in Equation II–28) are the emissivity coefficient $\epsilon_{j,\nu}$ and absorption coefficient $\kappa_{j,\nu}$[11] . Since we must obtain these quantities from the clump-averaged intensity and optical depth obtained from KOSMA-τ, we derive them as,

$$\kappa_{j,\nu} \equiv \frac{\tau_{j,\nu}}{\ell_{j,\text{eff}}}, \tag{II–30}$$

$$\begin{aligned} T_{\text{b},j,\nu} =& \frac{\epsilon_{j,\nu}}{\kappa_{j,\nu}}\left(1 - \mathrm{e}^{-\tau_{j,\nu}}\right), \\ \Rightarrow \epsilon_{j,\nu} =& \frac{T_{\text{b},j,\nu}}{\ell_{j,\text{eff}}} \frac{\tau_{j,\nu}}{\left(1 - \mathrm{e}^{-\tau_{j,\nu}}\right)}, \end{aligned} \tag{II–31}$$

where the effective length scale of clump j is,

$$\ell_{j,\text{eff}} \approx \frac{\frac{4}{3}\pi R_j^3}{\pi R_j^2} = \frac{4}{3} R_j. \tag{II–32}$$

Since we have pre-computed the superposition of clumps in velocity-space, we have the effective number of clumps at each systematic velocity in order to calculate the voxel-averaged quantity. When using multiple clumps sizes, we must consider all combinations of clumps. For example, a voxel containing 10 clumps of the same size will have the possible combinations $\left[0, 1, \ldots, 10\right]$. Likewise if the voxel contains 10 clumps of one size and 20 clumps of another, the combinations would be $\left[(0,0), (0,1), \ldots, (0,20), (1,0), \ldots, (10,20)\right]$. Let refer to a particular combination using the parameter k, whereas the jth index in that combination is k_j. The probabilistic calculation considers the probability of a number of clumps

[11] Note that we perform these calculation separately for both the line transitions and the dust continuum. We can add these afterwards.

$k_j \in [0, N_{j,\,i,\,\mathrm{eff}}]$ in a line-of-sight through the voxel. For the probabilistic approach, we define the probability of observing a particular clump as the area-filling factor of each clump,

$$\mathcal{P}_j \equiv f_{\mathrm{A},j} = \frac{\pi\, R_j^2}{\ell_{\mathrm{vox}}^2}, \tag{II–33}$$

which relates the size of the clump R_j with the size of the voxel[12] ℓ_{vox}. From this it is clear to see there is a higher probability for the larger clump to be in your line-of-sight. This is simply the probability of seeing a clump in a voxel if the voxel only contains one clump. Theoretically, however, a voxel may contain many clumps (the *clumpy* description of the ISM). Now for a given probability density function $f(k)$, we have the probability of each combination is $\mathcal{P}_k = \prod_j f(k_j)$, and ultimately we should arrive at the property,

$$\sum_k \mathcal{P}_k = \sum_k \prod_j f(k_j) = 1. \tag{II–34}$$

There are different options for determining the probability density function $f(k)$ to evaluate the probabilities. One of the simplifications made in Andree-Labsch et al. (2017) was to normalise the number of clumps to a smaller number, so the binomial probability mass function was suitable to explain the probabilities.

$$\mathcal{P}_{k,j,i} = \binom{N_{j,i,\mathrm{eff}}}{k_{j,i}} f_{\mathrm{A},j}^{k_j} \left(1 - f_{\mathrm{A},j}\right)^{N_{j,i,\mathrm{eff}} - k_{j,i}} \tag{II–35}$$

Alternatively there is an option to use a Poisson distribution. A key difference between these distributions is that Equation II–34 will not be true when using a Poisson distribution and the normalisation of the probabilities will create a numerical artefact. For that reason we rely on the binomial distribution. For a given combination of clumps k_i, we have the combination intensity and optical depth as,

$$\tau_{k,i,\nu} = \sum_j k_{j,i} \tau_{j,\nu} \tag{II–36}$$

$$\begin{aligned} \epsilon'_{k,i,\nu} &= \sum_j k_{j,i} \epsilon_{j,\nu} \ell_{j,\mathrm{eff}} \\ I_{k,i,\nu} &= \frac{\epsilon'_{k,i,\nu}}{\tau_{k,i,\nu}} \left(1 - \mathrm{e}^{-\tau_{k,i,\nu}}\right). \end{aligned} \tag{II–37}$$

Finally, it remains to explain our method of calculating voxel-averaged properties. As can be surmised from the preceding explanation, the voxel-averaged properties are a summation of the contribution from all of the combinations. The

[12] Note that in this sense we are comparing the projected area of the clump with the area of one face of the voxel. While in practice you cannot assume this geometry is satisfied, it is enough to approximate the voxel-averaged properties. In practice we then use many small voxels so we can use an average depth to make a synthetic observation.

clump intensity and optical depth, calculated in Equation II–37 and Equation II–36, must be scaled by the probability of the combination. In this calculation, the probability acts on the intensity and attenuation, so the combination optical depth must be processed accordingly. Thus we can calculate the contribution of each clump combination to the final synthetic voxel-averaged intensity and optical depth:

$$\left\langle T_{\mathrm{B},\nu}\right\rangle_{\mathrm{vox},i} = \left(\sum_k \left(\prod_j \mathcal{P}_{k,j,i}\right) T_{\mathrm{B},k,i,\nu}\right) \ , \tag{II–38}$$

$$\left\langle \tau_\nu\right\rangle_{\mathrm{vox},i} = -\ln\left(\sum_k \left(\prod_j \mathcal{P}_{k,j,i}\right) \mathrm{e}^{-\tau_{k,i,\nu}}\right) \ , \tag{II–39}$$

$$\left\langle \epsilon_\nu\right\rangle_{\mathrm{vox},i} = \frac{\left\langle T_{\mathrm{B},\nu}\right\rangle_{\mathrm{vox},i}}{\ell_{\mathrm{vox}}}\left(1-\mathrm{e}^{-\left\langle\tau_\nu\right\rangle_{\mathrm{vox},i}}\right) \ , \tag{II–40}$$

$$\left\langle \kappa_\nu\right\rangle_{\mathrm{vox},i} = \frac{\left\langle \tau_\nu\right\rangle_{\mathrm{vox},i}}{\ell_{\mathrm{vox}}} \ . \tag{II–41}$$

This is valid for each independent ensemble. For a model with multiple ensembles in each voxel (such as the two-ensemble model used in this work), the emissivity and absorption coefficients are summed. The calculation we have covered thus-far is for the internal velocities v_i. To obtain the voxel-averaged quantities as a function of observing velocity, we simply need to perform an interpolation:

$$\widetilde{\left\langle \epsilon_\nu\right\rangle}_{\mathrm{vox}}\left(v_{\mathrm{obs}}\right) = \mathrm{interp}\left(\left\langle \epsilon_\nu\right\rangle_{\mathrm{vox},i}\,;\, v_{\mathrm{obs}}\right), \tag{II–42}$$

$$\widetilde{\left\langle \kappa_\nu\right\rangle}_{\mathrm{vox}}\left(v_{\mathrm{obs}}\right) = \mathrm{interp}\left(\left\langle \kappa_\nu\right\rangle_{\mathrm{vox},i}\,;\, v_{\mathrm{obs}}\right), \tag{II–43}$$

$$\widetilde{\left\langle \tau_\nu\right\rangle}_{\mathrm{vox}}\left(v_{\mathrm{obs}}\right) = \mathrm{interp}\left(\left\langle \tau_\nu\right\rangle_{\mathrm{vox},i}\,;\, v_{\mathrm{obs}}\right), \tag{II–44}$$

$$\widetilde{\left\langle T_{\mathrm{B},\nu}\right\rangle}_{\mathrm{vox}}\left(v_{\mathrm{obs}}\right) = \frac{\widetilde{\left\langle \epsilon_\nu\right\rangle}_{\mathrm{vox},i}\left(v_{\mathrm{obs}}\right)}{\widetilde{\left\langle \tau_\nu\right\rangle}_{\mathrm{vox}}\left(v_{\mathrm{obs}}\right)}\left(1-\mathrm{e}^{-\widetilde{\left\langle\tau_\nu\right\rangle}_{\mathrm{vox}}\left(v_{\mathrm{obs}}\right)}\right), \tag{II–45}$$

where the values are simply linearly-interpolated[13]. To obtain a synthetic observation from our three-dimensional model, one then needs to add the dust contribution to the line emissivity and absorption coefficients before integrate the radiative transfer equation.

3.3 Multiple ensembles

It is useful for us to consider voxels that contain multiple independent ensembles. In this case, all ensembles would contribute to the emissivity and absorption, but

[13] Other kinds of interpolation are possible, such as quadratic and cubic, but these have a tendency to give non-physical results (such as negative absorption or intensity). There is also a possibility to fit a Gaussian to the emissivity and absorption coefficients, but this does not work well in the case of 0 emissivity.

we do not need to use the probabilistic approach to differentiate clumps from different ensembles. The total voxel-averaged emissivity and absorption is thus,

$$\widetilde{\langle \epsilon_\nu \rangle}_{\text{vox}} (v_{\text{obs}}) = \sum_{\text{ens}} \widetilde{\langle \epsilon_\nu \rangle}_{\text{vox,ens}} (v_{\text{obs}}) \ , \tag{II–46}$$

$$\widetilde{\langle \kappa_\nu \rangle}_{\text{vox}} (v_{\text{obs}}) = \sum_{\text{ens}} \widetilde{\langle \kappa_\nu \rangle}_{\text{vox,ens}} (v_{\text{obs}}) \ , \tag{II–47}$$

where $\langle \cdots \rangle_{\text{vox,ens}}$ represents the voxel-averaged value from an ensemble, calculated by Equation II–42 and Equation II–43. A reasonable question might arise at this point: why don't we simply separate the discrete clumps in a given ensemble into different ensembles, one for each clump size? The advantage of doing so would decrease the time it takes to compute the voxel-averaged extrinsic properties, since the probabilistic approach using combinatorics is the most time-consuming process in the calculation. If we consider an ensemble consisting of i clump sizes with n_i in the line-of-sight, the number of combinations we need to consider reduces from $\mathcal{O}\left(\prod_i (n_i + 1)\right)$ to $\mathcal{O}\left(\sum_i (n_i + 1)\right)$. The reason we do not use this method in practice is primarily due to the independence of ensembles.

Firstly, we must know the mass of each individual ensemble in order to initialise the calculation, which is not necessarily true. We are able to make certain assumptions to differentiate between two ensembles (see §III-3) for the dense and diffuse ISM, but we do not know *a priori* how this is split further into different clump sizes. We may use the fractal assumption to split the mass between clumps within an ensemble, but these clumps are not independent.

The probabilistic calculation anyways couples the available combinations between clumps in an ensemble. Effectively what this does for a given set of combinations $(0, 1, \ldots, n_i)$ is flatten the probability mass function, thus boosting the contribution of the smaller clumps.

CHAPTER III

kosmatau3d

So many good ideas are never heard from again once they embark in a voyage on the semantic gulf.

Alan J. Perlis

This chapter is solely about the `kosmatau3d` code itself: what are the parameters that can be altered, what are the types of models one can create, and how well the code performs. The idea here is to provide some reference on how the code works as well (beyond what was mentioned in Chapter II) as how it can be used. We describe the various submodules and classes in §III-1. There is then a brief description of model parameters in §III-2. The evaluation of and output given by each voxel is described in §III-3. We cover the integration of the radiative transfer equation using multiple voxels in §III-4. The files created for a given model are described in §III-5. This chapter should also cover issues native to the language of the code, `python`, since this usually raises a few flags for more experienced programmers. Therefore it will go into detail about the timing and efficiency of the code's performance (§III-6). Finally the chapter will end with a brief example showing how one would create a PDR structure plot using `kosmatau3d` in §III-7, showcasing both the clumpy and non-clumpy results. This chapter will necessarily contain more technical details (bear with me) than the rest of the thesis.

1 Code structure

First, we must download the code from GitHub and install it in a python environment. Any command-line interface in this chapter will be for the Bourne Again Shell (bash) on a Linux workstation.

~/projects/

```
$ git clone https://github.com/CraigYanitski/kosmatau3d.git
$ pip install kosmatau3d
```

This will install all dependencies. There should now be a base directory located at /home/projects/kosmatau3d/ which contains some files defining the python package (.py, .toml, .cfg, etc.), the directory kosmatau3d which contains the python code itself, the directory notebooks which will contain worked examples[14], the directory docs contains various documents covering the development of kosmatau3d, the directory doc will contain the primary documentation utilising sphinx[15], and the directory scripts contains a python script that can be called from the command line to run full grids (at the moment it is just for the Milky Way). The focus of this chapter is the contents of kosmatau3d/kosmatau3d/, hereafter called the *code directory*, which contains all of the input files and scripts required to run the model.

There are two folders in the code directory needed for the basic computation of kosmatau3d: grid and molecular_data. The latter folder contains information from the ONION routine (Gierens et al. 1992) in KOSMA-τ and is used for calculating the wavelength of each transition, while the former contains multiple grids of KOSMA-τ clump-averaged results (intensities, optical depths, fuv extinctions, column densities). The folders containing python code are models, notably containing the classes Voxel() and VoxelGrid() and all routines for the clumpy calculation; properties, which contains functions to plot properties assumed in KOSMA-τ (such as the density profile or the far-UV spectrum); and comparison, which contains code used to resample and regrid observations and compare to the model grids. These folders are independent of each other, and therefore each is considered a module in the kosmatau3d python package.

Importing the clumpy module

```
>>> from kosmatau3d import models
```

The structure of the models module has gone through a couple iterations, but the most efficient version utilises the Voxel() class facilitate most of the required calculations using other submodules and perform the final calculations for the voxel-averaged quantities. The advantage of doing this is that each submodule can

[14] The only worked example currently available is of the functionality of the Voxel() class. More will be added in the future.

[15] https://www.sphinx-doc.org/en/master/

own their variables for the calculations and debugging rather that each `Voxel()` instance. considering how many variables are required for the calculation, this largely reduces the memory footprint of running a model composed of many voxels. Each submodule of course has a specific purpose and was optimised in slightly different ways:

- Setup submodules
  - `constants`: This contains not only physical constants, but constants within the model such as which clump masses are used, the fractal parameters, the observing velocities, the directories to use for loading the input data or saving the model, the wavelengths at which the dust continuum is calculated, etc. This is initialised by passing keyword arguments to either the `Voxel()` class or the `Model()` class.
  - `observations`: A bit of a misnomer, this submodule opens all of the input model files such as the KOSMA-τ output grids and the files defining whichever three-dimensional model will be created. The advantage here is that there are separate methods for initializing the module, so one opens the data required to evaluate a voxel while the other opens the data files for the full three-dinemsional model.
  - `interpolations`: Here we pre-compute all of the interpolation functions from the input files. As with `observations`, the initialisation is split between interpolations required for the each voxel and interpolations required for a model containing many voxels. An additional advantage that is mainly helpful to reduce overhead during setup is that one can save the interpolations to files.
  - `species`: All of the species transition information (such as wavelength and frequency) is calculated and saved in this submodule. The advantage this has is that it reads this information directly from the input files for the ONION (located in `/molecular_data`) routine in KOSMA-τ, and thus remains consistent.
- Calculation submodules
  - `masspoints`: This is where the clump properties are calculated (such as the clump radius or H_2 mass) or interpolated(such as the emissivity and column densities). The name of this submodule follows the work of Andree-Labsch et al. (2017), but will soon be changed to `clumps` in a future release. The methods in this submodule are the only methods in `kosmatau3d` to interpolate from the KOSMA-τ grid.
  - `combinations`: Arrays of all the possible combinations of clumps in a line-of-sight, or at least as many that are considered in the code, are stored in this submodule.
  - `ensemble`: The clump combinations are calculated here, somewhat unintuitively, as well as the probability for each combination.

  - `shape`: This submodule is used for the full three-dimensional model to define the voxel positions based on the shape and dimensions of the desired model with a given orientation. The primary motivator for isolating this code into a submodule is to allow for different model shapes accessible altering the model parameters.

  - `radiatiative_transfer`: All of the code to define how the model is observed is contained within this submodule. For the case of the Milky Way, the observer is embedded in the model and it is necessary to use spherical coordinates. External sources such as PDRs and other galaxies are easier to compute in the sense that one can construct the model in such a way that the voxels are always viewed face-on, simplifying the calculation.

- Classes

  - `Voxel`: This is the most fundamental class. One is able to construct a voxel of given size containing one or more ensembles (to account for the clumpy medium and interclump medium). After evaluating the voxel (here meaning to calculate the voxel-averaged properties), it is possible to extract the velocity-dependant, voxel-averaged emissivity, absorption, or intensity (assuming no background radiation) with or without the continuum contribution. It is also possible to get the integrated intensity, or plot the results for each ensemble. The only values stored as class properties are the emissivity and absorption for the dust and each transition.

  - `VoxelGrid`: This facilitates the initialisation and evaluation of each voxel in the model. It is responsible for calling the appropriate methods from the `interpolations` routine, evaluating the emissivity and absorption coefficients, and exporting the voxel properties to `.fits` files. The technique used to reduce the memory footprint of running the full model is to stream the voxel properties to the corresponding files, then deleting the voxel instance. In this way there is only ever on voxel instance in memory. This of course will increase if and when multiprocessing is implemented.

  - `Model`: The main class interface for running the full three-dimensional model, the idea for this class is to facilitate the evaluation of the voxel grid, after which it would run the radiative transfer portion of the code. Currently this is not the pipeline for creating a model, but it is at least created in the anticipation of future development.

  - `SyntheticModel`: This class is used to open all of the saved data and plot in different views (currently there is a radial plot or scatter plot).

2 Parameters

Practically all of the parameters involved in the clumpy calculation are available as parameters in `kosmatau3d`, with the exception of some KOSMA-τ parameters[16]. These parameters are specified here not only because they affect the output of the model, but because many of them are required to be constants in the model. For example, if this code was used to model a single PDR, we likely do not want to change the fractal parameters or the different clump masses considered in each ensemble. Simulating each voxel thus requires these values to be kept constant. In an effort to minimise the number of redundant calculations, all of these are contained in the submodule `constants` and the calculations are only performed when the submodule is initialised.

Table III–1: Parameters that may be adjusted for each `kosmatau3d` voxel. Many of these are assumed to be constant in this thesis (for example the fractal parameters α and γ), but can be important parameters in future studies in the fractal ISM.

Parameter	Units	Keyword argument	Description
ℓ_{vox}	pc	`voxel_size`	The size (in parsecs) of each voxel in the model. This is necessarily a constant for each model, and is vital to the ensemble calculation (see Chapter II).
n_m	N/A	`clump_mass_number`	The number of different clump mass bins composing the ensemble. This parameter can be a list of values, each corresponding to an independent ensemble.
$M_{\mathrm{min}}, M_{\mathrm{max}}$	$\log_{10}(M_\odot)$	`clump_mass_range`	The lower and upper limits of the clump masses composing the ensemble. This parameter must be a list of the same length of `clump_mass_number`. It is possible to give just one limit, in which case the ensemble contains just that mass.

[16] In order to vary these parameters, one would need to include it in the underlying KOSMA-τ grid.

Table III–1 continued...

Parameter	Units	Keyword argument	Description
N_{max}	N/A	`clump_n_max`	The number of the largest clumps to use during the calculation. This is used to simplify the calculation when there are many clumps in the calculation.
α	N/A	`alpha`	The index for the clump mass number relation. It is assumed to follow a power-law, and has been well-researched (Heithausen et al. 1998; Stutzki & Guesten 1990). We assume a standard value of 1.84.
γ	N/A	`gamma`	The index for the clump mass-size relation (*not the cosmic ray ionisation rate*). We assume a standard local value of 2.31 (Heithausen et al. 1998).
n_v	N/A	`velocity_resolution`	The number of internal velocities v_i at which to perform the calculations.
n_v	N/A	`velocity_number`	The number of observing velocities desired in the output.
$v_{\text{min}}, v_{\text{max}}$	$^{\text{m}}/_{\text{s}}$	`velocity_range`	The range of the velocity axis desired in the output.
M_{ens}	$M_\odot$	`ensemble_mass`	The mass of the ensemble. This is used to determine the total number of clumps contained in each voxel. This can be a single value for one ensemble or a list for multiple ensembles.

Table III–1 continued...

Parameter	Units	Keyword argument	Description
$n_{\rm ens}$	$\rm cm^{-3}$	`ensemble_density`	The ensemble-averaged density, used to calculate the surface density of each clump. This can be a single value for one ensemble or a list for multiple ensembles.
$\sigma_{\rm ens}$	$^{\rm m}/_{\rm s}$	`ensemble_dispersion`	The velocity dispersion of the ensemble. This can be a single value for one ensemble or a list for multiple ensembles.
χ	$\chi_{\rm D}$	`fuv`	The far-UV radiation incident on the ensemble. This is assumed to be isotropically radiating on each clump. This can be a single value for one ensemble or a list for multiple ensembles.
$\zeta_{\rm H}$	$\rm s^{-1}$	`crir`	The primary cosmic ray ionisation rate. We assume a value of 2×10^{-16} $\rm s^{-1}$, which is generally agreed upon as the local value in the literature (Bisbas et al. 2021; Padovani et al. 2020; Hollenbach et al. 2012; Indriolo et al. 2009; Padovani et al. 2009). This can be a single value for one ensemble or a list for multiple ensembles.

Table III–2: Parameters that may be adjusted for each kosmatau3d galactic model. Some here are listed due to their importance, but are not adjusted for this fit (for example α and γ; these are important to alter for models of other galactic environments).

Parameter	Units	Description
$\zeta_{\mathrm{H,CMZ}}$	s^{-1}	The cosmic rate ionisation rate in the central molecular zone. It is assumed to be constant in the CMZ at 10^{-14} s^{-1}.
R_{CMZ}	pc	The radius of the central molecular zone. This only affects ζ_H as this value should be higher in the CMZ (Padovani et al. 2020).
$f_{\mathrm{H^0}}$	N/A	The fraction of atomic hydrogen mass that is represented by the interclump medium.
m_{CM}	$\mathrm{M}_\odot$	The clump mass range considered in the model. These do not affect the clump mass number or mass-size relations.
m_{ICM}	$\mathrm{M}_\odot$	The interclump mass range considered in the model. These do not affect the clump mass number or mass-size relations.
$f_{\mathrm{m,H_2}}$	N/A	A scaling factor for the mass of molecular hydrogen in the Milky Way. This does not alter the profile (inferred from Figure IV–4).
$f_{\mathrm{m,H^0}}$	N/A	A scaling factor for the mass of atomic hydrogen in the Milky Way. This does not alter the profile (inferred from Figure IV–4).
f_{FUV}	N/A	A scaling factor for the far-UV radiation in the Milky Way. This does not alter the profile (seen in Figure IV–5).

3 Single-voxel model

The best way to understand the advantage of kosmatau3d over one-dimensional PDR models is to examine what can be provided with a single voxel. The voxel itself can have any physical size, irrespective of the size of the individual clumps inside. One must just be careful to ensure the total volume of clumps inside the voxel does not exceed the volume of the voxel.The voxel properties, as implied in §II-3.2, are calculated assuming the voxel is viewed face-on and all of the clumps inside the voxel are isotropically irradiated with the same intensity of far-UV radiation. The three primary parameters that are used to create a unique ensemble are `ensemble mass`, `ensemble density`, and `fuv`. The local cosmic ray ionisation rate `crir` may also be specified as an input parameter, but this assumes different initial species abundances in the clumps. In practice we use two ensembles in each voxel: one for the dense, clumpy PDR and one for the diffuse interclump ensemble (as it was used in Andree-Labsch et al. 2017). From these ensembles we aim to model the

dense molecular gas and the atomic CNM, respectively. The base assumption for this approach is that the hydrogen in the dense clumps are primarily molecular while the hydrogen in the diffuse ensemble is primarily atomic (verified in §III-3.2).

3.1 KOSMA-τ grids

As explained in Chapter II, the clump properties are interpolated from a pre-computed grid. While we have the grid used in the Andree-Labsch et al. (2017) application of `kosmatau3d` (hereafter referred to as `OrionBar2013`), we have supplemented these results with two other grids: `CMZ2015` (used in García et al. 2016) and `PDRTB2020` (used in `PhotoDissociation Region Toolbox`; Pound & Wolfire 2023).

Parameter	Grid	Range	Spacing
$n_{\mathrm{H,s}}$	`OrionBar2013`	$10^3, \ldots, 10^7\,\mathrm{cm}^{-3}$	1 dex
	`CMZ2015`	$10^3, \ldots, 10^7\,\mathrm{cm}^{-3}$	1 dex
	`PDRTB2020`	$10, \ldots, 10^7\,\mathrm{cm}^{-3}$	0.5 dex
m_{cl}	`OrionBar2013`	$10^{-3}, \ldots, 10^3\,\mathrm{M}_\odot$	1 dex
	`CMZ2015`	$10^{-3}, \ldots, 10^3\,\mathrm{M}_\odot$	1 dex
	`PDRTB2020`	$10^{-3}, \ldots, 10^3\,\mathrm{M}_\odot$	0.5 dex
χ	`OrionBar2013`	$1, \ldots, 10^6\,\mathrm{cm}^{-3}$	1 dex
	`CMZ2015`	$1, \ldots, 10^6\,\mathrm{cm}^{-3}$	1 dex
	`PDRTB2020`	$1, \ldots, 10^6\,\mathrm{cm}^{-3}$	0.5 dex
ζ_{H}	`OrionBar2013`	N/A	N/A
	`CMZ2015`	$10^{-15}, \ldots, 10^{-13}\,\mathrm{s}^{-1}$	1 dex
	`PDRTB2020`	N/A	N/A

Table III–3: The parameter space and model density for each of the KOSMA-τ grids used in this work.

While the initial abundances for `OrionBar2013` and `PDRTB2020` are quite similar, `CMZ2015` was created for the conditions in the Galactic centre and thus has very different initial abundances. Our base assumption is that the `OrionBar2013` grid is valid for the entire galaxy, and when we use cosmic ray ionisation rate $\zeta_{\mathrm{H}} \geq 10^{-15}\,\mathrm{s}^{-1}$ (the local rate), we use results from `CMZ2015`. The `PDRTB2020` grid reaches lower densities with its parameter space, which is why we use it to model the interclump medium.

While the chemistry in KOSMA-τ keeps track of more than 100 species, a small subset of these have transitions that are used in the radiative transfer calculation. Since it is possible for multiple transitions to exist for a given species, we identify them in ascending order of energy in `kosmatau3d`. For example, the ^{12}CO $J = 1 \rightarrow 0$ at 115 GHz is specified as `CO 1`, the ^{12}CO $J = 2 \rightarrow 1$ at 330 GHz is specified as `CO 2`, and so on. Some of the transitions used in this thesis are given in Table III–4.

[17] This wavelength is obtained from the input to the radiative transfer module of KOSMA-τ (the

Table III–4: The transitions of various species and how they are specified in KOSMA-τ. Not all of these are used in the models presented in Chapter IV

species	ground state	transition	wavelength	KOSMA-τ ID
^{12}C$^+$	$1s^2 2s^2 2p^1$	$^2P^o_{3/2} \rightarrow ^2 P^o_{1/2}$	157.0 μm[17]	C+ 1
^{12}C	$2s^2 2s^2 2p^1$	$^3P_1 \rightarrow ^3 P_0$	610.1 μm	C 1
		$^3P_2 \rightarrow ^3 P_1$	369.1 μm	C 2
		$^3P_2 \rightarrow ^3 P_0$	230.0 μm	C 3
^{12}CO	$J=0$	$J=1\rightarrow 0$	2.601 mm	CO 1
		$J=2\rightarrow 1$	1.300 mm	CO 2
		$\vdots$		
		$J=49\rightarrow 48$	53.90 μm	CO 49
^{13}C$^+$	$1s^2 2s^2 2p^1$	$^2P^o_{3/2} \rightarrow ^2 P^o_{1/2}$	157.7 μm	13C+ 1
^{13}C	$2s^2 2s^2 2p^1$	$^3P_1 \rightarrow ^3 P_0$	609.0 μm	13C 1
		$^3P_2 \rightarrow ^3 P_1$	370.4 μm	13C 2
		$^3P_2 \rightarrow ^3 P_0$	230.3 μm	13C 3
^{13}CO	$J=0$	$J=1\rightarrow 0$	2.731 mm	13CO 1
		$J=2\rightarrow 1$	1.365 mm	13CO 2
		$\vdots$		
		$J=49\rightarrow 48$	56.55 μm	13CO 49
H^0	$1s$	$F=1\rightarrow 0$	21.11 cm	HI

3.2 Atomic fraction

For our application of the two-ensemble model to work we need to ensure that the clumps used in the interclump ensemble is primarily atomic. Since the diffuse interclump medium in the model should be ubiquitous in the ISM, we must ensure that it is atomic at the local far-UV intensity (1 χ_D). We therefore plot the atomic and molecular hydrogen masses at varying clump surface densities in Figure III–1 to verify our grid selection. Here we can see that the clump is indeed primarily atomic at the lower densites provided by the `PDRTB2020` grid, and thus we may use it to model the interclump medium in the remaining analysis and assume an interclump surface density of $n_H = 10^{-3}$ cm^{-3}.

3.3 Single voxel example

For now we will simply create of voxel of size 1 pc containing 100 $M_\odot$ of ISM split into clumps of size 1 $M_\odot$, 10 $M_\odot$, and 100 $M_\odot$. Since this example is purely for demonstration, we do not specify how the change the KOSMA-τ grid from which we

so-called ONION routine). While this discrepancy from the accepted [^{12}C II] wavelength of 158 μm does affect the output of KOSMA-τ, this is a very minor effect in the analysis presented in this thesis.

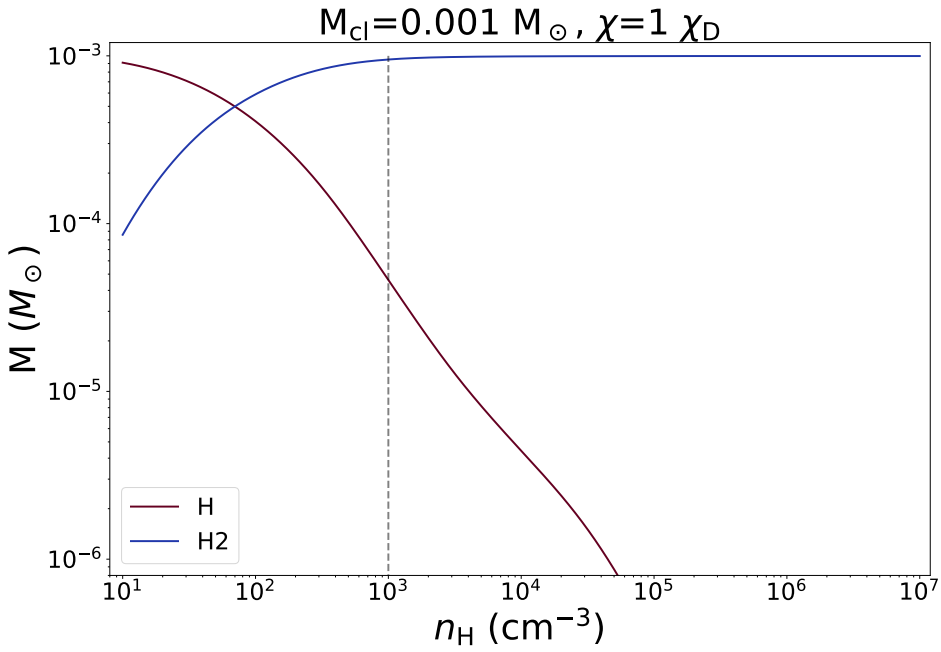


Figure III–1: The hydrogen mass in a 10^{-3} clump with impinging far-UV intensity of $1\,\chi_{\rm D}$ for the `OrionBar2013` (from $n_{\rm H} = 10^3\,{\rm cm}^{-3}$) and `PDRTB2020` (from $n_{\rm H} = 10^1\,{\rm cm}^{-3}$) grids used in this work.

interpolate[18]. We also do not specify which transitions to include, so by default all transitions are included in the model. Note that since the voxel contains the same mass as the largest clump and we force a clumpy distribution, this will contain just a fraction of the largest clump. This is possible due to the statistical approach in §II-3.2.

Voxel setup

```
>>> parameters = {
...     # Model parameters
...     "voxel_size": 1,
...     "molecules" : 'all',
...     "dust" : 'PAH',
...     "clump_mass_range" : [[0, 2]],
...     "clump_mass_number" : [3],
...     "clump_n_max" : [1],
...     "velocity_range" : [-5, 5],
...     "velocity_number" : 201,
...     # Voxel properties
...     "velocity" : 0,
...     "ensemble_dispersion" : 2,
...     "ensemble_mass" : [100],
```

[18] by default we use the `OrionBar2013`

```
...     "ensemble_density" : [1e5],
...     "fuv" : [1e2],
...     # Calculation
...     "velocity_resolution" : 10,
...     }
>>> from kosmatau3d import models
>>> vox = models.Voxel()
>>> vox.set_properties(**parameters)
```

Here we have also set some options to modify our output. `velocity_range` and `velocity_number` define the velocities in the output (the synthetic observing velocities) while `velocity` is the ensemble-averaged velocity of the ensemble. We also specified `molecules` as "all" to include all transitions, and `dust` is set to "PAH" so the continuum is computed to include the PAH features. There are two parameters affecting the calculations from Chapter II: `velocity_resolution` affects the resolution $r_{\rm vel}$ of the internal velocity grid (where the velocity spacing is $\sigma_{\rm cl}/r_{\rm vel}$; here it is 10 to ensure a smooth curve[19]), and `clump_n_max` affects how the calculation is normalised (in this case we ignore the normalisation step).

Now there are various properties we can extract from the voxel, but we are most interested in the line transition and continuum emission: absorption coefficient, emissivity coefficient, intensity, and optical depth. As these are not only interpolated from the limited spectral range of what is given in the ensemble calculations, but there may be some use-cases for including the continuum contribution and some cases for not, all of these spectral properties have their own accessing method. An example for the most sophisticated method, `get_species_intensity`, is given below.

Accessing intensity (no background)

```
>>> vox.get_species_intensity(integrated=False,
...     kind='linear', include_dust=False,
...     total=True, hi=False)
```

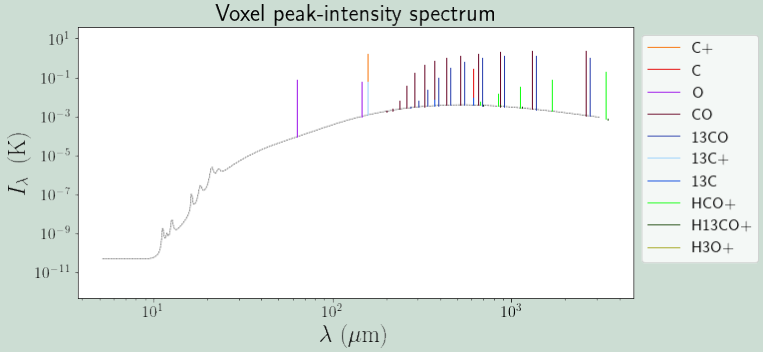
For this method to function, it utilises the voxel-averaged emissivity and absorption coefficients to integrate the radiative transfer equation without any background intensity over voxel length $\ell_{\rm vox}$. Thus it measures the intensity produced by the voxel itself. Similarly, there is an option `integrate` to integrate over the velocity axis. `total` is an option to sum the contributions from all ensembles in the voxel, if there are more than one. The flag to specify the return of the H I intensity is `hi`, since these values are kept separate in the code (this is because the H I data is post-processed rather than an actual result of KOSMA-τ). Finally `kind` is the option specifying the method of interpolation.

One can then compare the relative integrated intensities (or even peak emission values) over the entire spectrum using the method `plot_spectrum()`,

[19] This is not necessary in practice unless you require high spectral resolution.

Plotting the spectrum

```
>>> vox.plot_spectrum(quantity='intensity',
...     title='Voxel peak-intensity spectrum')
```

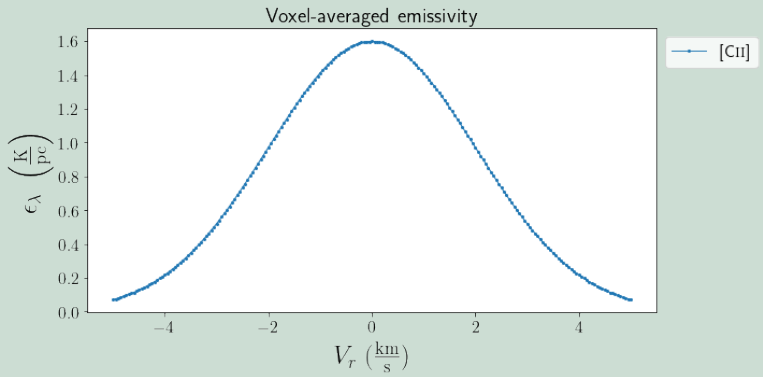


and thus one can clearly see the relative intensities of different transitions, as well as the heating of the ensemble in the CO transition ladder.

If one is more interested in the absorption effects of different transitions, one can examine specific transitions with respect to observing velocity. Here we can produce one plot each for emissivity, absorption, and intensity. Let us consider simply the [^{12}C II] 158 μm transition line (the naming convension in `kosmatau3d` is discussed in §III-3.1).

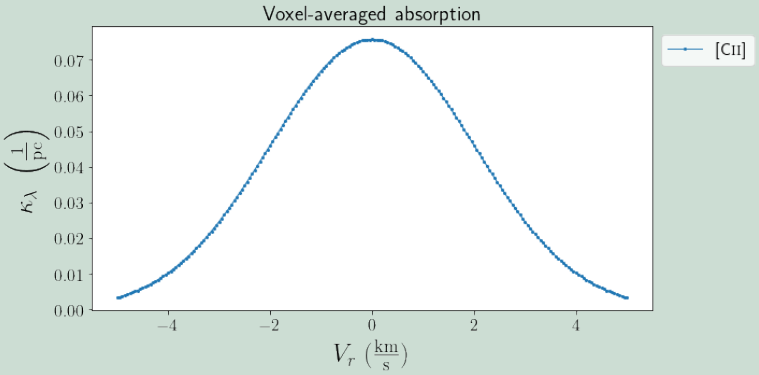
[^{12}C II] emissivity

```
>>> vox.plot_transition(quantity='emissivity',
...     title='Voxel-averaged emissivity',
...     transition='C+ 1',
...     label=r'[C\textsc{ii}]')
```



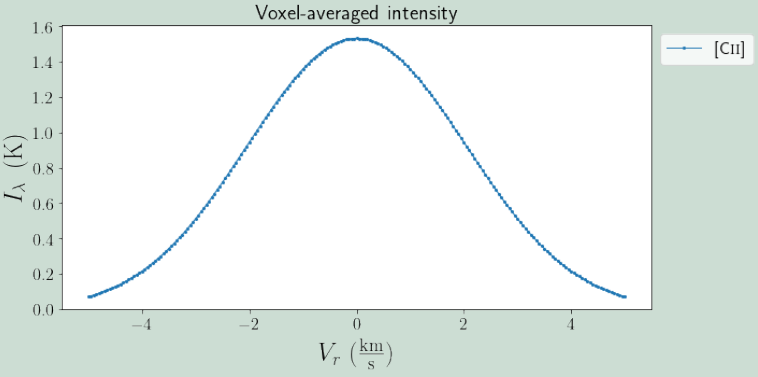
[^{12}C II] absorption

```
>>> vox.plot_transition(quantity='absorption',
...     title='Voxel-averaged absorption',
...     transition='C+ 1',
...     label=r'[C\textsc{ii}]')
```



[^{12}C II] intensity

```
>>> vox.plot_transition(quantity='intensity',
...     title='Voxel-averaged intensity',
...     transition='C+ 1',
...     label=r'[C\textsc{ii}]')
```



What these plots clearly show is that the spectra of the emissivity and absorption coefficients of individual transitions have a Gaussian shape, whereas the intensity spectrum does not. This is meant to be a confirmation that what ultimately is calculated by the voxel object is in line with the underlying theory.

4 Radiative transfer

For the integration of the radiative transfer equation over a line-of-sight passing through several voxels, we must make a couple assumptions. First, we choose to orientate the voxels face-on in a row. Next we must make linear approximations to the voxel-averaged properties along the line-of-sight to improve the accuracy of our calculation. To simplify the equations in the rest of this thesis, any discussion of radiative transfer using multiple voxels will refer to voxel-averaged properties such as $\langle \epsilon_\nu \rangle_{\rm vox}$ as $\epsilon_{i,\nu}$, where i refers to the position of the voxel with respect to the background voxel (where $i = 0$). The linear approximations to the emissivity and absorption coefficients are thus,

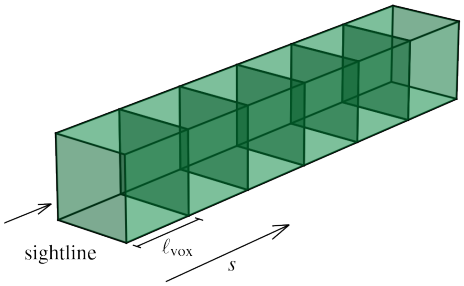


Figure III–2: An illustration of how the voxels are oriented for integrating the radiative transfer equation.

$$\epsilon_{i,\nu}(s) = \epsilon_{i,\nu} + \frac{\Delta\epsilon_{i,\nu}}{\ell_{\rm vox}} s \ , \tag{III–1}$$

$$\kappa_{i,\nu}(s) = \kappa_{i,\nu} + \frac{\Delta\kappa_{i,\nu}}{\ell_{\rm vox}} s \ , \tag{III–2}$$

where the Δ terms refer to the difference between the ith voxel with the voxel just behind. From this assumption, it is now possible to fully integrate Equation I–3.

RT integration

We begin with the radiative transfer equation in the Rayleigh-Jeans limit (so we substitute $T_{{\rm B},\nu}$ and ϵ_ν for I_ν and j_ν),

$$\mathrm{d}T_{{\rm B},\nu} = -\mathrm{d}\tau_\nu \, T_{{\rm B},\nu,{\rm bg}} + \mathrm{d}\tau_\nu \, \frac{\epsilon_\nu}{\kappa_\nu} \ , \tag{III–3}$$

keeping in mind that $\mathrm{d}\tau_\nu = \kappa_\nu \, \mathrm{d}s$ for constant κ_ν over path length s. Suppressing the i notation, we can substitute our linear approximation into Equation III–3:

$$\begin{aligned} \mathrm{d}T_{{\rm B},\nu} = &- \mathrm{d}s \left(\kappa_\nu + s \, \frac{\Delta\kappa_\nu}{\ell_{\rm vox}} \right) T_{{\rm B},\nu,{\rm bg}} \\ &+ \mathrm{d}s \left(\epsilon_\nu + s \, \frac{\Delta\epsilon_\nu}{\ell_{\rm vox}} \right) \ . \end{aligned} \tag{III–4}$$

Integrating with respect to $\ell_{\rm vox}$, we get,

$$T_{\mathrm{B},\nu} = \mathrm{e}^{-(\kappa_\nu + \frac{1}{2}\Delta\kappa_\nu)\ell_{\mathrm{vox}}} \left[T_{\mathrm{B},\nu,\mathrm{bg}} + \int_0^{\ell_{\mathrm{vox}}} \mathrm{d}s \left(\epsilon_\nu + s \frac{\Delta\epsilon_\nu}{\ell_{\mathrm{vox}}} \right) \mathrm{e}^{\int_0^s \mathrm{d}s' \left(\kappa_\nu + s' \frac{\Delta\kappa_\nu}{\ell_{\mathrm{vox}}} \right)} \right] , \tag{III–5}$$

$$\begin{aligned} T_{\mathrm{B},\nu} = {} & \frac{\Delta\epsilon_\nu}{\Delta\kappa_\nu} \left(1 - \mathrm{e}^{-\left(\kappa_\nu + \frac{1}{2}\Delta\kappa_\nu\right)\ell_{\mathrm{vox}}} \right) \\ & - \frac{\epsilon_\nu \Delta\kappa_\nu - \kappa_\nu \Delta\epsilon_\nu}{\sqrt{|\Delta\kappa_\nu|^3}} \sqrt{\frac{\pi}{2}} \\ & \left(\mathrm{e}^{a_\nu^2 - b_\nu^2} \mathrm{erfi}(a_\nu) - \mathrm{erfi}(b_\nu) \right) \\ & + T_{B,\nu}(0)\, \mathrm{e}^{-\left(\kappa_\nu + \frac{1}{2}\Delta\kappa_\nu\right)\ell_{\mathrm{vox}}} , \end{aligned} \tag{III–6}$$

where $\mathrm{erfi}(x)$ is the imaginary error function.

5 Output structure

Each time a model is run there is plenty of data that is saved. A directory for the data is created and all properties are saved in `.fits` format in order to include a header describing the model parameters. In general enough data is saved to rerun each voxel individually without running the entire model, so this includes,

- `clump_number.fits`
- `clump_radius.fits`
- `voxel_density.fits`
- `voxel_ensemble_dispersion.fits`
- `voxel_ensemble_mass.fits`
- `voxel-filling_factor.fits`
- `voxel_FUVabsorption.fits`
- `voxel_fuv.fits`
- `voxel_h2_mass.fits`
- `voxel_hi_mass.fits`
- `voxel_position.fits`
- `voxel_velocity.fits`

The observed intensities of course also need to be saved. We split the saved files between the transition data and the continuum as well as separating the H I 21 cm line data, which is post-processed.

- `dust_absorption.fits`
- `dust_emissivity.fits`
- `hi_absorption.fits`
- `hi_emissivity.fits`
- `species_absorption.fits`
- `species_emissivity.fits`
- `synthetic_hi_intensity.fits`
- `synthetic_hi_optical_depth.fits`
- `synthetic_intensity.fits`
- `synthetic_optical_depth.fits`

6 Performance

Performance is a major factor for any code. It is useful if code developers abstain from making assumptions or simplifications, but most of the time the accuracy of the code is correlated with the computation time. Thus optimising the performance of astrophysical code requires a balance of the code sophistication and the computation time.

While much of the development of `kosmatau3d` has been devoted to the improvement and optimisation of how the underlying theory is evaluated, reducing computation time, most of this section will concern the assumptions used and accuracy of how the code operates.

6.1 Interpolating from the KOSMA-τ grid

As is always the case when performing an interpolation, there is a certain degree of error between the expected output and the expected output. In terms of what we discuss within the context of `kosmatau3d`, the error arises from how we interpolate the clump-averaged brightness temperature and optical depth grids using the input parameters (which is refered to as our *parameter space*). In our case, the grid is defined for three input parameters: $n_{\mathrm{cl,s}}$, M_{cl}, and χ. With such a high dimensionality, at least in `python`, we basically only have linear interpolation or the radial basis function for direct methods of interpolation. We will attempt to measure the accuracy of the linear interpolation method.

Measuring the accuracy of any interpolation typically requires us to understand the underlying function, or at least a trend, in order to know how extraneous is the performance of the interpolation. Since there is too much computation required to evaluate KOSMA-τ for each point in the interpolation we test, we use the existing original grid parameters from which we interpolate. This should not work using the entire parameter space since linear interpolation will exactly fit the input, so our procedure is thus as follows:

1. Remove the interpolation point from the grid.
2. Use the reduced grid to create the interpolation function.
3. Perform the desired interpolation.
4. Use the interpolated quantity and the original quantity to calculate the relative residual.

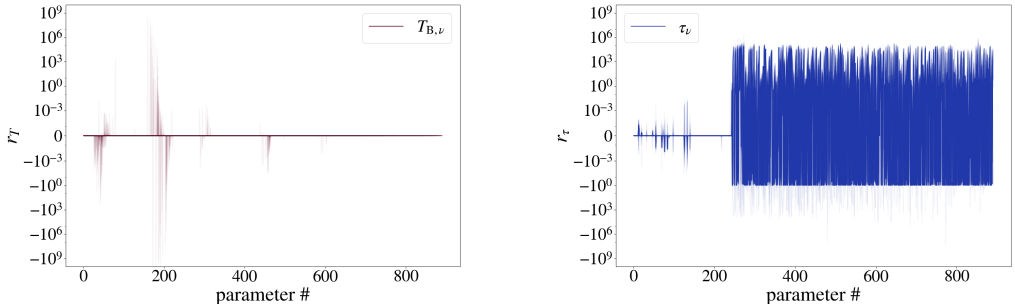
This must be performed for all transitions/wavelengths in the grid. To estimate the interpolation error, we examine the relative residual,

$$r_q \equiv \frac{\tilde{q} - q}{q} \ , \tag{III–7}$$

of the interpolated quantity $\tilde{q}$ compared to the original quantity q, where q is either the clump-averaged brightness temperature $T_{\mathrm{B,cl}}$ or clump-averaged optical depth τ_{cl}. The error of the interpolation is thus $E(q) \equiv |r_q| \times 100\%$.

In this way we enumerate the points in our gridded parameter space, then plot the relative residuals (see Figure III–3). This is done using linear interpolation for the brightness temperature (r_{T}; subfigure **(a)**) and optical depth (r_τ; subfigure **(b)**). There are certain parameters that have have a large amount of error when interpolated, but the majority of the parameters $\lesssim 200$ have interpolation error $Eq \ll 100\%$. The interpolation error for parameters $\gg 200$ is still reasonable for the brightness temperature grid, but there is significantly more error in the optical depth grid. From the definition in Equation III–7, we can also determine whether the interpolated quantities are predominately over- or under-estimated. In subfigure **(a)**, there are isolated locations where the interpolation error is high, and it is evenly distributed between positive and negative residual. The large, mostly positive relative residuals in subfigure **(b)** indicate a tendency of the interpolation of this grid to overestimate the optical depth.

Another method that has been explored for a non-parametric interpolation of the KOSMA-τ grid is an *extremely-randomised trees* (Geurts et al. 2006). A form of supervised machine-learning regression, this ensemble technique builds on the decision tree method (which is basically a nearest-neighbour interpolation) to give a more-accurate representation of the underlying population. Given an input of training data of k attributes (the input variables) and one output variable, the extremely-randomised trees method creates several decision trees taking several samples of the attributes and taking a random sample of where it splits each node. For a given set of input variables, an average is taken over the outputs of the decision trees to obtain a single prediction. A distinct difference in this method



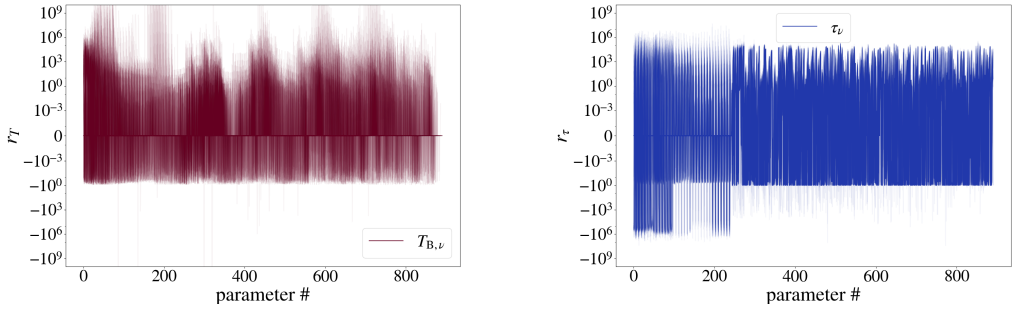
(a) $T_{\mathrm{B,cl}}$ **(b)** τ_{cl}

Figure III–3: The relative error in the interpolated clump-averaged quantities using linear interpolation.

when compared to linear interpolation is that extremely-randomised trees will not fit explicitly the input, but rather attempt to find the most likely output given the training data. It is also a regression method that minimises the variance of the output given a over-sampled parameter space.

From Figure III–4, we see predicted r_T (subplot **(a)** and r_τ (subplot **(b)** not only exhibit much more variance than in Figure III–3, but the peak relative residuals are also much greater. A possible reason for the large residuals when using extremely-randomised trees is that it is not robust to a sampled regular grid, so our method of quantifying the accuracy of these interpolation techniques introduces too many artefacts. We need our method for deriving the interpolation error since it is too computationally expensive to run KOSMA-τ independently of the grid.

This is enough to conclude that linear interpolation is much more accurate for the purpose of interpolating the KOSMA-τ grids.



(a) $T_{\mathrm{B,cl}}$ **(b)** τ_{cl}

Figure III–4: The relative error in the interpolated clump-averaged quantities using extremely-randomised trees.

7 Clumpy PDR profile

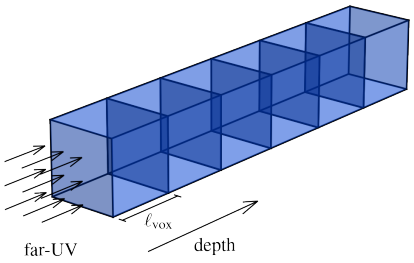


Figure III–5: An illustration of how the voxels are oriented for integrating the radiative transfer equation.

To fully understand how clumpyness affects the PDR model results, it is best to examine local properties as a function of depth into a PDR (for example abundance, temperature, or emissivity). As shown in Figure III–5, we consider a column of voxels, each containing identical ensembles with mass $M_{\rm ens}$ and density $n_{\rm ens}$, illuminated from one end with far-UV radiation (for our purposes this radiation has intensity $\chi = 10^6\,\chi_{\rm D}$–the maximum radiation in our grid). Each voxel in turn attenuates the far-UV radiation for the next, until there is insufficient far-UV radiation for the gas in the voxel to be modelled as a PDR (when $\chi > 1\,\chi_{\rm D}$). There are of course some parameters we are ignoring in such a model such as depth and velocity dispersion, but this rather gives a reasonable indication on how a PDR structure is affected by the clumpy approximation description.

Figure III–6 shows some of these profiles as a function of depth using voxels with size $\ell_{\rm vox} = 2\,$pc. Each voxel contains one ensemble with average density $n_{\rm ens} = 10^4\,{\rm cm}^{-3}$ and mass $M_{\rm ens} = 10^2\,{\rm M}_\odot$, but we differentiate between the profiles we get using one clump mass (solid line) in the ensemble compared to using three clump masses (dashed line). The structure profiles appear quite similar as a function of far-UV intensity, but there are slight differences between them due to the different clump masses. This effect is most noticeable at the PDR surface (where depth $s = 0$). While the gas temperature throughout most of the profiles indicates that the ensembles are part of the CNM, the temperatures with far-UV radiation $\chi \lesssim 10\,\chi_{\rm D}$ (the ambient far-UV radiation in the vicinity of the Sun) indicate the ensemble is part of the molecular medium. Since the temperature is weighted by mass (see Equation II–11), these temperature profiles are unique to the choice of ensemble-averaged density $n_{\rm cl}$ and clump masses $m_{\rm cl}$ in the ensemble. Changing the mass $M_{\rm ens}$ contained in each voxel changes the total number of clumps (see Equation II–17) and thus primarily affects the far-UV absorption and the total depth of the profile.

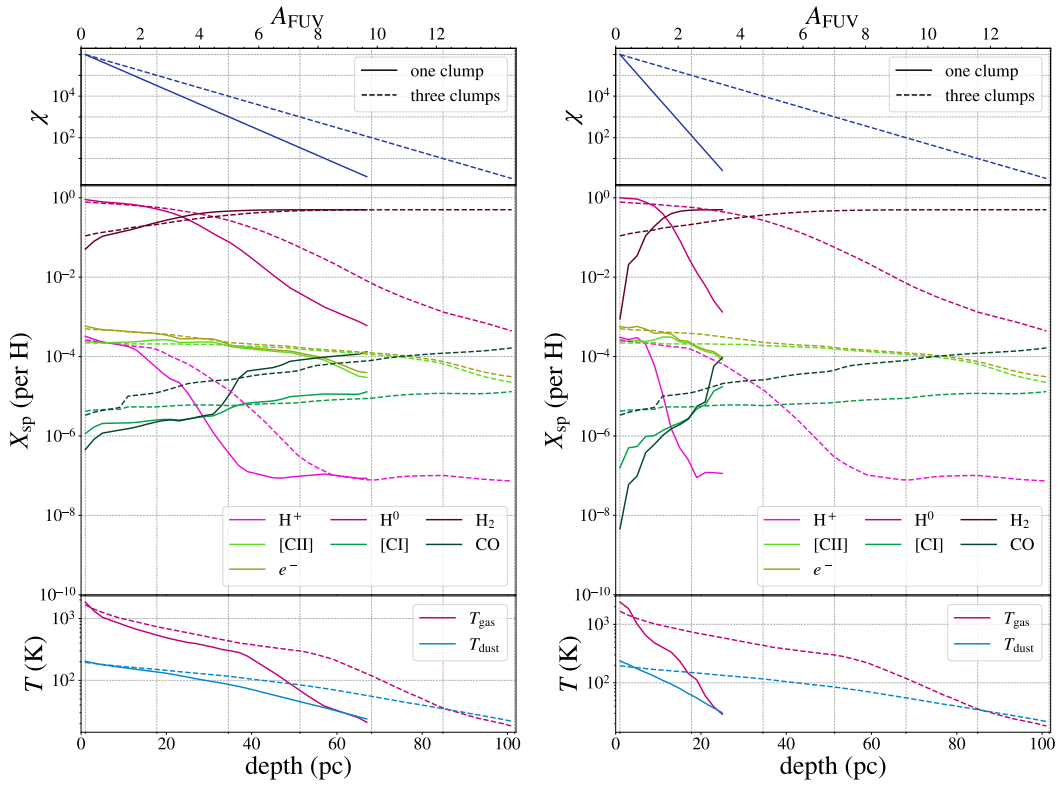


Figure III–6: The PDR structure profile as created with `kosmatau3d`. This profile is created using voxels of size $\ell_{\rm vox} = 2\,$pc containing an ISM mass $M_{\rm ens} = 10^2\,{\rm M}_\odot$ and average density $n_{\rm ens} = 10^4\,{\rm cm}^{-3}$. The top panel shows the far-UV intensity, the middle panel shows the fractional abundance for some important species, and the bottom plot shows the gas and dust temperature as a function of depth. The solid lines correspond to voxels containing either clump mass $m_{\rm cl} = 10\,{\rm M}_\odot$ (subfigure **(a)**) or $m_{\rm cl} = 1\,{\rm M}_\odot$ (subfigure **(b)**), and the dashed lines correspond to voxels containing clump masses $m_{\rm cl} = [1, 10, 100]\,{\rm M}_\odot$. The vertical grid is placed where the far-UV radiation decreases by a factor of 10 (following the *three clumps* profile).

While the clumpy ensemble in the voxels used for Figure III–6, as explained, is representative of the CNM and molecular media, we want to see if the diffuse interclump ensemble can be used to also model the WNM. We therefore compare ensembles composed of the two least-massive clumps available in our KOSMA-τ grid: $m_{\rm cl} = 10^{-2}\,{\rm M}_\odot$ and $m_{\rm cl} = 10^{-3}\,{\rm M}_\odot$ (see Figure III–7). Both ensembles seem to have approximately the same profiles except for the abundance of H_2 and ^{12}CO at at far-UV radiation $\chi \lesssim 10\,\chi_{\rm D}$, where the profile containing clumps with mass $m_{\rm cl} = 10^{-2}\,{\rm M}_\odot$ have a higher abundance. There also seems to be some unexplained effect where the H^0 and e^- abundances increase from the surface and peak at $\chi \sim 10^4\,\chi_{\rm D}$. Since the gas temperature is calculated tracing H^0, this effect is seen as well in $T_{\rm g}$.

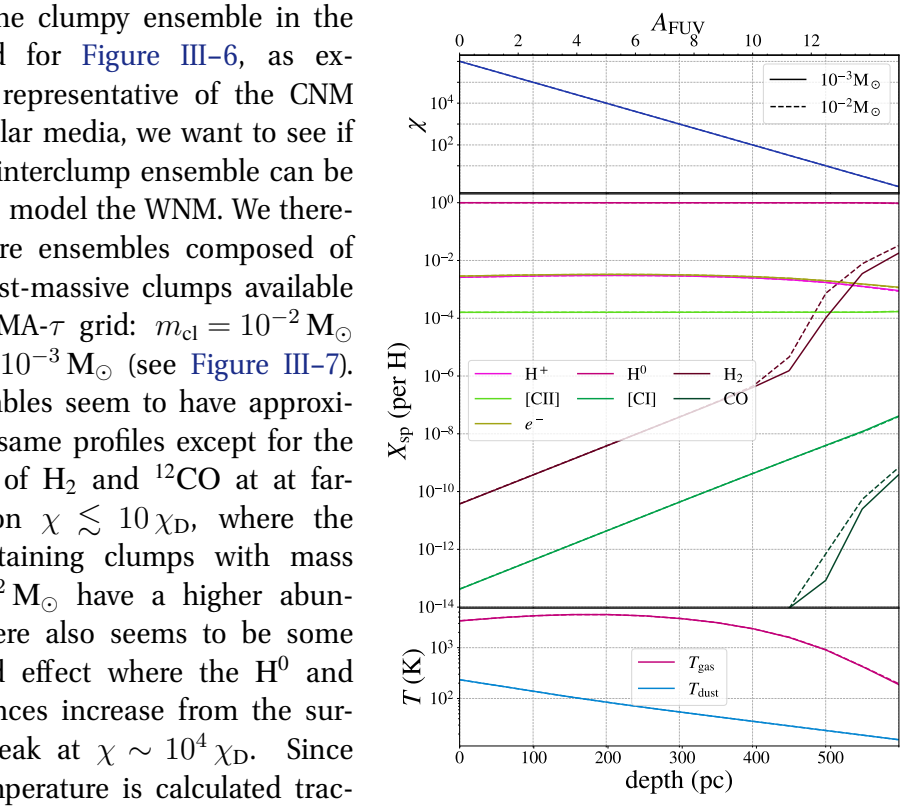


Figure III–7: Like Figure III–6, but using voxels of size $\ell_{\rm vox} = 1\,{\rm pc}$ containing ensembles with $M_{\rm ens} = 10^{-1}\,{\rm M}_\odot$ and $n_{\rm ens} = 19.11\,{\rm cm}^{-3}$ (ensuring the surface density of each clump is $n_{\rm s,cl} = 10\,{\rm cm}^{-3}$). The solid lines correspond to voxels containing clump mass $m_{\rm cl} = 10^{-3}\,{\rm M}_\odot$, and the dashed lines correspond to voxels containing clump mass $m_{\rm cl} = 10^{-2}\,{\rm M}_\odot$.

As we will see in Chapter IV, the ambient far-UV radiation intensity remains at $\chi \approx 1\,\chi_{\rm D}$, and thus the physical conditions of our clumpy and interclump ensembles mean they will suitably model the molecular medium and CNM, respectively. The interclump ensemble gets nearly hot enough to model the WNM at $\chi \sim 10^4\,\chi_{\rm D}$, where the gas temperature $T_{\rm g} \lesssim 8000\,{\rm K}$. These will have to be the conditions we need to use if we are to include the WNM in our galactic models (this is explained further in §IV-5). To ensure we keep the interclump ensemble composed mainly of H^0, we will use clumpy with mass $m_{\rm cl} = 10^{-3}\,{\rm M}_\odot$.

CHAPTER IV

Galactic models

Fools ignore complexity. Pragmatists suffer it. Some can avoid it. Geniuses remove it.

Alan J. Perlis

Having a thorough understanding on how our PDR model functions, we must determine how best to simulate the cooling lines in our Galaxy using `kosmatau-3d`. As seen in Chapter III, the advantage of representing the ISM as a clumpy PDR contained within voxels is not only that one can account for inhomogeneity in the ISM, but one can account for self-absorption in the line spectra. For this reason we want to focus on properly modelling the self-absorption in a galactic context rather than directly resolving the spatial distribution of the galactic ISM. Thus we use axisymmetric relations for the galactic properties and low spatial resolution in relation to the finer structure of the galaxy (meaning the voxels are sometimes larger than the typical scale of a giant molecular cloud and even the galactic scale height) in order to save on computational power. In order to test our PDR model in a galactic context, we perform a detailed model of the Milky Way (the galactic properties are discussed in §IV-1). Since the properties vary as a function of galactocentric radius, we discuss how these trends can be extracted from the data files for each model in §IV-2. One possible source of error using voxels larger than the scale height of the Milky Way is that the voxels are only partially-filled, affecting how we can use and analyse the model results (this is explained thoroughly in §IV-3). Even if we run models with lower spatial resolution (in this case resulting in larger voxels), this limits the lower-limit of our synthetic maps (see §IV-4 for a discussion of how these are made).

1 Milky Way

The lower resolution we find useful is $\ell_{\mathrm{vox}} = 400\,\mathrm{pc}$. While this appears arbitrary, it is high-enough resolution to model the large-scale structure of the Milky Way

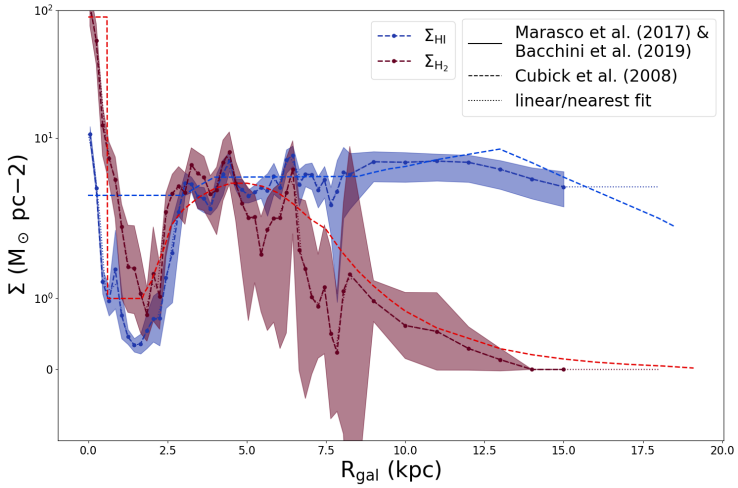


Figure IV–1: The surface mass density of the Milky Way as a function of galactocentric radius. Outside the solar circle we use the results of Bacchini et al. (2019b) while inside the solar circle we use the higher-resolution results of Marasco et al. (2017).

observation, but not so large as to require a supercomputer (each simulation takes around 30 minutes and produces 6.1 GB of data; models with resolution scale 100 pc requires 120 minutes and produces >20 GB of data). We have opted to use a lower resolution to create more model grids to constrain the various parameters. This ensures all development and analysis can be performed on a personal computer (the specifications of the computer used can be found in Appendix A). The resolution is still an improvement upon the 1 kpc clumpy analysis used in Cubick et al. (2008).

1.1 Mass profiles

Simple axisymmetric relations are used for the properties of the Milky Way. These are based on both observations (in the case of surface mass density and clump-averaged surface density) and simulations (for the far-UV radiation). The voxel parameters from Table III–1 necessary to specify a KOSMA-τ model are the primary cosmic ray ionisation rate ($\zeta_{\rm H}$), hydrogen number density ($n_{\rm H}$), dense and diffuse mass ($M_{\rm H_2}$ & $M_{\rm H^0}$), and far-UV radiation (χ).

The mass distributions must be derived from axisymmetric profiles of surface mass density and galaxy scale height. These have been derived by Bacchini et al. (2019b) for both molecular hydrogen and atomic hydrogen for $1\,{\rm kpc} \leq R_{\rm gal} \leq 15\,{\rm kpc}$. Figure IV–2 shows that the H_2 (H^0) scale height is a minimum of ~ 25 pc (~ 50 pc) in the galactic centre, and increases to a maximum of ~ 300 pc (~ 500 pc) at the edge of the galactic disk. We extrapolate the scale height linearly for $R_{\rm gal} < 1\,{\rm kpc}$ and $R_{\rm gal} > 15\,{\rm kpc}$. Since we do not need to resolve structure in latitude, we do not consider a Gaussian distribution of gas in the $\hat{z}$ direction and use the scale height as the thickness of the galactic disk. The profiles of surface mass density are a combination of Marasco et al. (2017) in the inner galaxy

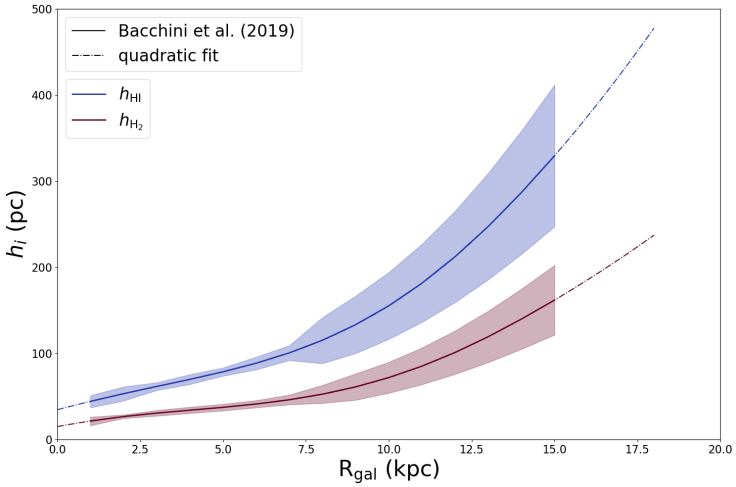


Figure IV–2: The scale height of the Milky Way disk as a function of galactocentric radius. This is the same derived by Bacchini et al. (2019b), extrapolated to the size of model used in this work.

($R_{\rm gal} < R_{\rm gal,\odot}$) and an average of various previous determinations for the outer galaxy (see Figure IV–1). In order to properly model the higher mass in the galactic centre, we rather directly use the higher resolution Marasco et al. (2017) solution for the inner galaxy. We simply convert the surface mass densities to a volume mass density by dividing by twice the scale height of the galaxy ($2h_{\rm i}$, see Figure IV–3).

A major difference between the results we use from Bacchini et al. (2019b) and the previous work by Wolfire et al. (2003) is that the latter assumed relatively constant values for the scale height of the Milky Way (with a maximum difference of $<$ 1dex) while the former fitted the scale height. As seen in Figure IV–2, the results from Bacchini et al. (2019b) indicate a flaring of $\lesssim$ 2dex over the plane of the Milky Way. Among other things, this means there is a much lower ISM density in the outer disk, and thus a lower column density (which is proportional to the integrated intensity of a given line transition).

The ensemble mass of each voxel depends on the volume mass density distribution in the galaxy,

$$M_{\rm ens} = \bar{\rho_{\rm MW}}\, \ell^3_{\rm vox}, \tag{IV–1}$$

where $\rho_{\rm MW}$ is the volume mass density in the Milky Way. While it is important to consider first the volume $V_{\rm MW}$ of the Milky Way that is contained withing each voxel, we use Equation IV–1 to correctly model the column density for the Milky Way plane (see argument in §IV-3). Thus all of the voxels in our galactic models are uniformly-filled $V_{\rm MW}$ is simply the volume of the voxel ($V_{\rm MW} = \ell^3_{\rm vox}$). An initial assumption is that the mass distribution of molecular hydrogen traces the dense gas, while the mass distribution of neutral hydrogen traces more the diffuse gas. Ideally these are modelled with KOSMA-τ using larger, dense models for the dense gas and smaller, diffuse models for the interclump region (the masses of these clumps span $10^{-3}\,{\rm M}_\odot$ to $10^{2}\,{\rm M}_\odot$ according to Heithausen et al. 1998), though the two are not necessarily uncoupled (by definition a PDR contains H^0). For this

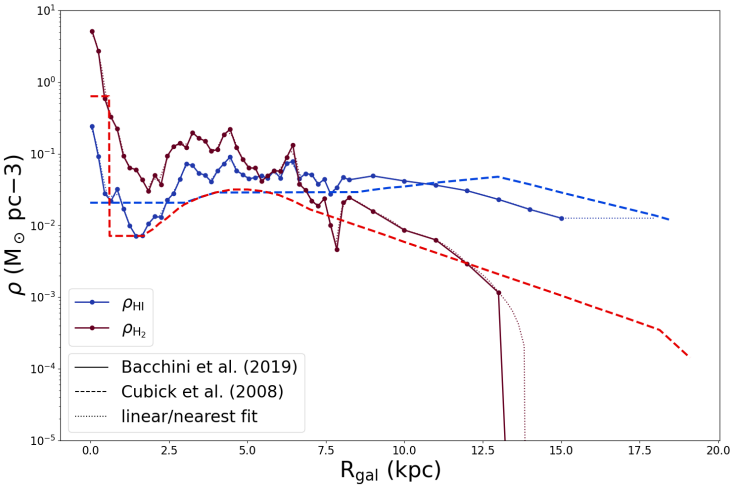


Figure IV–3: Volume mass density model used for the distribution of neutral and molecular mass in the Milky Way as a function of galactocentric radius. These distributions are calculated using the surface mass density profiles from Marasco et al. (2017) and Bacchini et al. (2019b) and the scale height of the Milky Way.

reason, the mass modelled with clumpy ensemble follows the molecular hydrogen mass (M_{H_2}; modelled with clump masses $10^0 - 10^2\,M_\odot$), while the mass of the interclump ensemble we use the atomic hydrogen mass (M_{H^0}; using clumps with mass $10^{-3}\,M_\odot$ as determined in §III-3.2).

1.2 Ensemble-averaged density profile

The ensemble-averaged density we use is the same as that of Cubick et al. (2008) (see Figure IV–4). The profile is that of the CNM (n_{CNM}) and it was derived by balancing the gravitational potential of H^0 and H_2 with thermal pressure (Wolfire et al. 2003). One of the results in Cubick et al. (2008) was that the ensemble-averaged density profile scaled this profile so the density at the solar circle is $n_{CNM}(R_\odot) = 10^{3.8}\,cm^{-3}$. For the two ensembles we use in our work, we will use a radial estimation of the ensemble-averaged density for the dense clumpy ensemble and a constant value for the diffuse interclump ensemble ($19.11\,cm^{-3}$, as determined in §III-3.2)[20]. Since the mass in the clumpy medium is determined by the density of H_2 (see Figure IV–3), it is technically possible to scale this mass density in a similar way to derive the corresponding ensemble-averaged density for the clumpy medium. The issue with this is that the outer galaxy has densities $\sim 10^{-10}$ that of the inner 10 kpc and this region cannot be modelled using KOSMA-τ (the density range is too large). We therefore choose to keep the density profile of Cubick et al. (2008) to use for the clumpy ensemble.

[20] Technically the interclump medium should have a number density extending down to $0.1\ cm^{-3}$ in order to account for the WNM, though the KOSMA-τ models cannot converge on a solution for such a low density. We try to mitigate this by using very small clumps, and assume the difference is negligible. What is anyways important in the result is the column density in the synthetic observation.

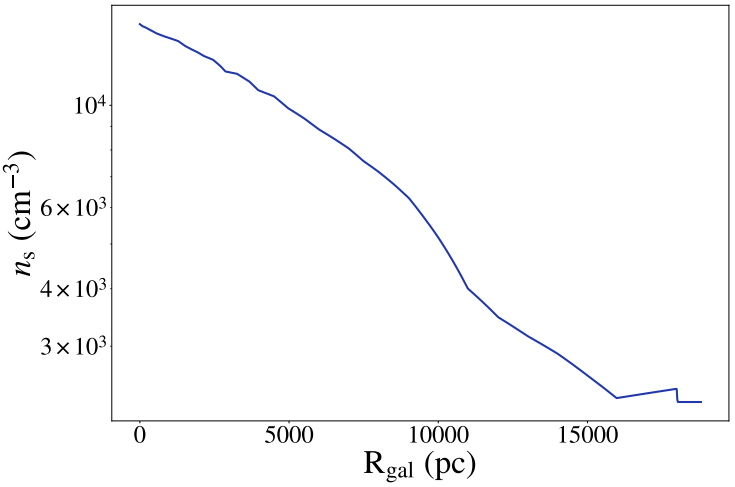


Figure IV–4: The total clump hydrogen density in the Milky Way as a function of galactocentric radius. This is the same as what was used in Cubick et al. (2008).

1.3 Far-UV profile

Since most of the gas in a galaxy is distributed in the ISM, one might think the FUV radiation is the same as the Draine unit (Draine 1978, ; $\chi_{\rm D} \approx 1.71\,G_0 \approx 8.95{\times}10^{-14}$ erg cm^{-3}). It is more accurate to rather simulate the distribution of O-B stars in the Milky Way (Popescu et al. 2011), which are the largest contributors of UV radiation. Such a simulation was done to derive the axisymmetric distribution of the far-UV radiation density, as seen in Figure IV–5, for 9 wavelengths. This is enough to approximate the far-UV SED and calculate our required profile from,

$$\chi(r) = \int_{912\,\text{Å}}^{2066\,\text{Å}} \mathrm{d}\lambda\, u_\lambda(r) \times \frac{\chi_{\rm D}}{2.63\times10^{42}\,\text{erg pc}^{-3}}\ , \tag{IV–2}$$

where we integrate over the wavelength range considered by KOSMA-τ (912 Å to 2066 Å) and multiply by the conversion factor from energy density units (erg per cubic parsec) to the energy density of the Draine far-UV field. It is immediately evident that most of the ISM in the Galactic disk is embedded in a far-UV field less than the Draine field, reaching a minimum of $\sim 10^{-2}\,\chi_{\rm D}$ at the edge of the Galactic disk at $R_{\rm gal} = 18$ kpc. From the constraints of our PDR model, we force the minimum of the profile to be $1\,\chi_{\rm D}$. Our motivation of this is that the PDR gas is located closer to the ionising sources than the large-scale galactic average.

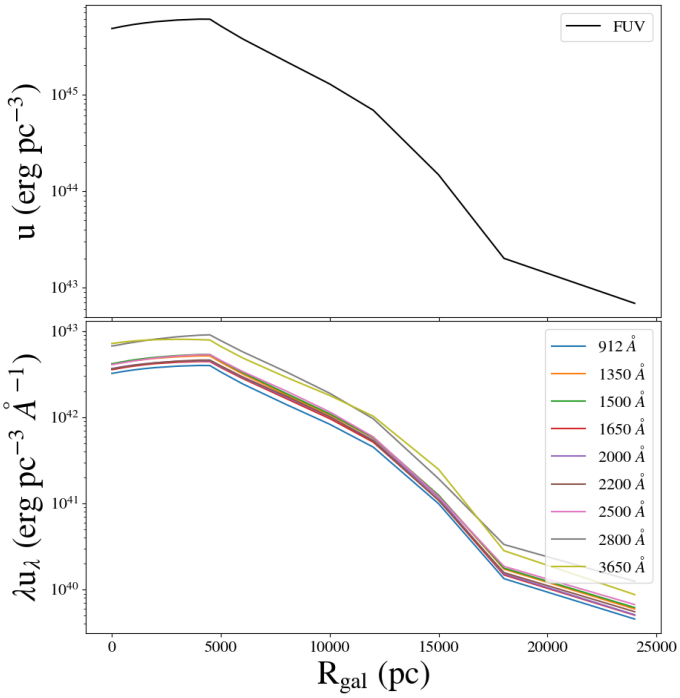


Figure IV–5: The distribution of far-UV radiation as modelled by Popescu et al. (2011). The two results available are either the total far-UV radiation intensity (integrated over the spectrum), or the specific far-UV intensity at the specified wavelengths. The specific intensity values are used in this work to integrate over the far-UV range of KOSMA-τ (912 Å to 2066 Å).

1.4 Rotational velocity profile

Finally, one must adopt a specific rotation curve for the Milky Way. The rotation curve we adopt is an amalgamation of two separate studies. Bhattacharjee et al. (2014) examined the kinematics of a variety of sources in the Milky Way and derived a rotation curve out to $R_{\text{gal}} \approx 200$ kpc. There is also more recent and precise astrometric data for Milky Way objects from *2MASS*, *Gaia*, and *WISE*. This has been supplemented by the spectroscopic data of *Apogee* and analysed by Eilers et al. (2019) to obtain the rotation curve in the interval $(5, 25)$ kpc. For the Milky Way models we created for the following analysis, we combined the inner 5 kpc from Bhattacharjee et al. (2014) with the entire model of Eilers et al. (2019) (see Figure IV–6).

This is sufficient to approximate both the rotational velocity and velocity dispersion of each voxel, since the difference in rotational velocity dominates the velocity dispersion. To put this in terms that can be used by `kosmatau3d`, however, we must derive the local standard of rest (LSR) velocity v_{LSR}. Since all of the voxels in the model necessarily have their positions recorded in galactocentric Cartesian coordinates, we need to record this with respect to a reference position. The reference point is the point defining our LSR (ie. at $R_{\text{gal},\odot} \equiv 8.178$ kpc; GRAVITY Collaboration et al. 2019). From this we can derive the relative longitude and latitude of each voxel, which we combine with the galactocentric longitude and latitude to recover v_{LSR}. This is summarised in the face-on representation of the galactic disk in Figure IV–6, where the primed quantities are relative to Earth and the un-primed quantities are galactocentric.

Using the law of cosines, we derive the angle σ' between voxel's galactocentric radial vector and relative radial vector as,

$$\sigma' = \cos^{-1}\left(\frac{r^2 + r'^2 - R_{\text{gal},\odot}}{2rr'}\right). \tag{IV–3}$$

Since both the voxel and the LSR have some rotational velocity around the Milky Way, one must correct for this to derive the voxel's rotational velocity in the co-rotating frame:

$$v'_{\text{gal}} = v_{\text{gal}} - v_{\text{rot},\odot}\left(\frac{r}{R_{\text{gal},\odot}}\right). \tag{IV–4}$$

Finally these can be derived quantities can be used to calculate the voxel's v_{LSR}:

$$v'_{\text{LSR}} = v'_{\text{gal}} \sin(\sigma') \cos(\beta'). \tag{IV–5}$$

1.5 Constant parameters

We try to vary the parameters we expect to have the greatest effect, or at least those which are not well-constrained. There are some parameters affecting the model which we do not vary since the average Milky Way is sufficient. An example of this is the aforementioned (Chapter III) mass spectrum parameters derived by

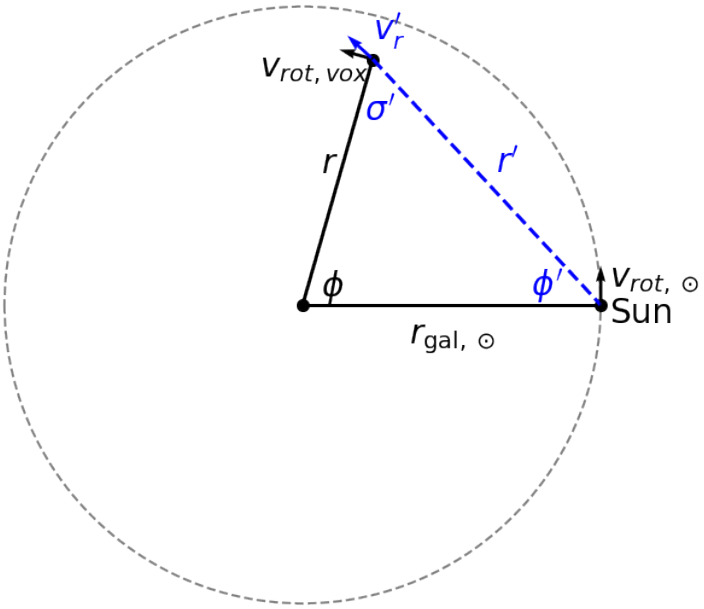


Figure IV–6: The geometry needed for calculating the velocity in a differentially-rotating disk. These variables can be used in Equation IV–3 to Equation IV–5 to calculate each voxel's v_{LSR}.

Heithausen et al. (1998). Their derived values $\alpha = 1.84$ and $\varpi = 2.4$ were thus kept constant for all of the models we ran.

An approximation we make which fails is the average Milky Way cosmic ray ionisation rate, ζ_H. The typical value of the total ionisation of molecular hydrogen is approximated to be $\zeta_{\mathrm{H_2}} \approx 2 \times 10^{-16}$ s^{-1}, though this value is highly variable depending on the galactic environment. Cosmic rays are high-energy particles emitted from various energetic events in the galaxy; fusion, stellar winds, and active galactic nuclei (AGN) are all sources of cosmic rays (Albertsson et al. 2018). The primary sources of cosmic rays, however, are supernovae. Supernovae are very prominent forms of stellar feedback in star-forming environments. The increased cosmic ray ionisation rate can have a significant effect on gas temperature and ionisation fraction (Bisbas et al. 2021). However, Kabanovic et al. (2022) showed that for cold C+, increasing the cosmic ionization rate by an order of magnitude increases the C+ excitation temperature from 15K to 17K. For this reason one should consider galactic ecology and at least derive a function of cosmic ray ionisation with respect to galactocentric radius. According to Padovani et al. (2020), the cosmic ray ionisation rate is roughly constant at $\zeta_H \sim 10^{-14}$ s^{-1} throughout the CMZ, and the rest of the galactic disk has a lower ζ_{H_1}. Adopting their approximation and approximating the disk $\zeta_{\mathrm{H_2}}$ to be the local value of 2×10^{-16} s^{-1} (Hollenbach et al. 2012), we fit the radius of the CMZ, R_{CMZ}, as a parameter in our models.

2 Parameter profiles

We have implemented the possibility to create radial profiles of various parameters directly from the saved data for each `kosmatau3d` model. Since the models assume axial symmetry, these are particularly effective to compare for example line ratios across the galactic disk. This in turn enables us to quantify a variety of trends, from classifying star formation throughout the galaxy to estimating the amount of CO-dark gas and even providing an estimate of how this might be traced. In Figure IV–7 we show the azimuthally-averaged far-UV radiation, which should recover the profile in Figure IV–5. As we have mentioned in §IV-1.3, we forced the profile minimum to be $\chi = 1\,\chi_{\rm D}$. This modification is of course applied after any scaling of the far-UV profile when we examine the parameters in our model grids.

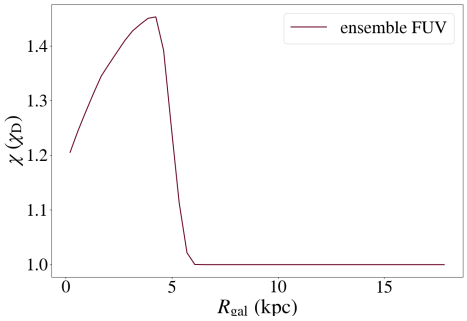


Figure IV–7: A radial profile of the far-UV radiation as obtained from the fiducial `kosmatau3d` model.

We can also examine the voxel-filling factor throughout the Galactic disk, since this affects the modifications we are able to make in §IV-3. As seen in Figure IV–8, the voxels are filled less than 50% within $\sim 7.5\,$kpc for the atomic medium and within $\sim 12\,$kpc for the molecular medium. It is explained in the next chapter why this is important for the calculation of various trends in our models.

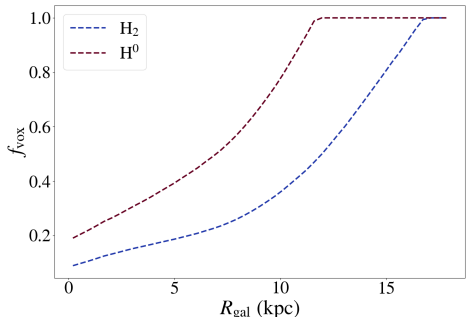


Figure IV–8: A radial profile of the voxel-filling factor $f_{\rm vox}$ as obtained from the fiducial `kosmatau3d` model.

3 Partially-filled voxels

An important concept to understand from the setup of our models is that of a partially-filled voxel. From §II-3.2 we outlined the formalism for calculating voxel-averaged versions of the emissivity and absorption, which by definition do not assume any confinement of the clumps in the voxel. It is therefore advantageous for us to investigate how one might alter the theory behind the voxel-averaged properties, both intrinsic and extrinsic, when we must further confine the clumps in one of the dimensions. This is required in the Milky Way models, for example,

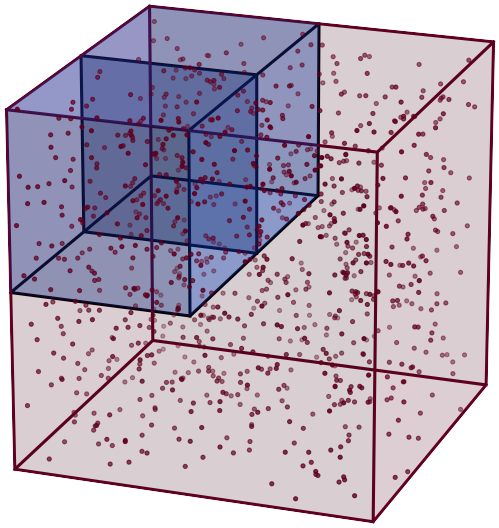


Figure IV–9: A cubic region of ISM (red) modelled using voxels (blue). The scatter points illustrate the distribution of clumps in the voxel. The voxels are fully-occupied with the ISM.

when the scale height on the galactic centre is smaller than the scale of the voxel used to model it. Thus the clumps, which should be confined within $\pm h_{\mathrm{H}}$, are uniformly distributed in $\pm \ell_{\mathrm{vox}}/2$.

Our approach to quantify the effect of this confinement is to model the emission of two hypothetical configurations of ISM: one occupying a cubic region of $1\,\mathrm{pc} \times 1\,\mathrm{pc} \times 1\,\mathrm{pc}$ and one occupying a slab region of $1\,\mathrm{pc} \times 1\,\mathrm{pc} \times 0.25\,\mathrm{pc}$ (Compare Figure IV–9 and Figure IV–10). For simplicity and to make a direct comparison, we consider both regions to contain a mass of $100\,\mathrm{M}_{\odot}$, both embedded in an isotropic far-UV radiation field of $1\,\chi_{\mathrm{D}}$. While the volume density of these regions are much greater than the densities in the Milky Way models presented in this thesis, this is deliberate to ensure there are absorption effects after integrating the radiative transfer equation. If one is to test the integration of the radiative transfer equation along a pencil beam through the centre of the region using the voxel sizes $\ell \in (1, 0.5, 0.25, 0.125)\,\mathrm{pc}$. In the case of the uniform voxel, the calculated intensity of all lines should be very close. For the partially-filled voxel, however, the calculated intensities should increase from the $1\,\mathrm{pc}$ voxel to the $0.25\,\mathrm{pc}$ voxel, at which point it should be uniformly-filled and the calculated intensities will converge on the correct value. This is a result of the mass volume-density in the partially-filled voxels scaling by the area ℓ_{vox}^2 rather than the volume ℓ_{vox}^3 for voxels with $\ell_{\mathrm{vox}} <$ the thickness of the slab.

Considering the two conceptual regions in the context of the Galactic disk (where the voxel), since they both have the same mass and fractal dimensions

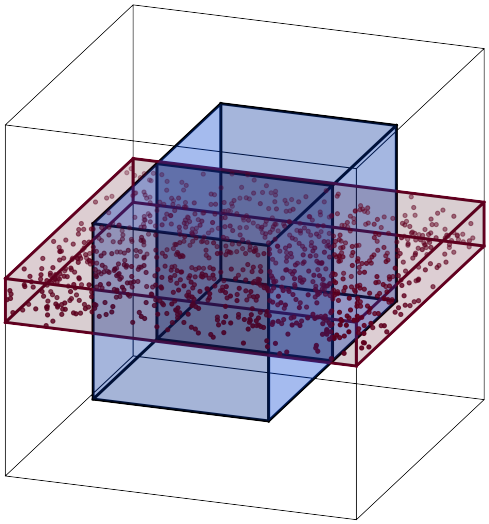


Figure IV–10: The same as Figure IV–9 except the ISM fills just a portion of the voxel, while the black frame outlines the original shape of the ISM

they will have the same number of clumps, N_j. The surface densities and far-UV radiation fields are not affected by the geometry of the hypothetical regions, thus what really changes when observing a pencil beam through each region is what we can call the *clump column-density*:

$$\mathcal{N}_j(\mathrm{cl}) \equiv \frac{N_j}{\ell^2_{\mathrm{vox}}} \frac{\ell^2_{\mathrm{vox}}}{A_{\mathrm{ISM}}}, \tag{IV–6}$$

where A_{ISM} is the projected area of the ISM contained in the voxel and j is the mass index in our discrete ensemble.

We can quantify the difference generally by defining another parameter in the voxel model called the *voxel-filling factor*, f_{vox}. Similar to how the volume-filling factor is defined, in the context of a clumpy ISM, as the ratio between the volume of a region and the volume occupied by clumps, the voxel-filling factor is defined as the ratio between the volume of the voxel and the volume of the ISM it models. Where this new factor enters into the theoretical framework of §II-3.2 is the number of clumps in a line-of-sight, N_j. Our proposed alterations to accommodate for clump confinement would ideally occur at this step in the calculation, so the recorded emissivity and absorption of each voxel is consistent with how they are used in integrating the radiative transfer equation. Thus we can perform these altered calculations for situation where one knows *a priori* that a voxel is partially-filled, but this makes it too time-consuming to recompute the voxel-averaged properties then the synthetic observation for different models to constrain the implementation (which we need to do for the comparison of models

to observations in Chapter VI of this thesis).

If one wants to change the viewing angle of the voxel and/or if the models have already been run, we must test and verify a approximate solution applying the factor to the coefficients in Equation I–3 (in order to account for absorption effects). We do this not only for the aforementioned configurations, but also for modifications to them where there is more mass in either the foreground or the background (similar to the two layer models of Guevara et al. 2020; Kabanovic et al. 2022). By doing so, we should be able to verify that a proposed solution still captures the optical depth effects that are an advantage of `kosmatau3d` over one-dimensional codes such as KOSMA-τ.

Since both emissivity and absorption are additive along a line-of-sight, we propose to use the voxel-filling factor to modify Equation II–42 and Equation II–43 by,

$$\langle \epsilon_\nu \rangle'_{\rm vox} = \frac{\langle \epsilon_\nu \rangle_{\rm vox}}{f_{\rm vox}} \ , \tag{IV–7}$$

$$\langle \kappa_\nu \rangle'_{\rm vox} = \frac{\langle \kappa_\nu \rangle_{\rm vox}}{f_{\rm vox}} \ , \tag{IV–8}$$

where the scaled values are specified as primed ($'$). This approximation should scale the emission values appropriately to account for higher column density from the voxel being partially-filled, however it neglects the probabilistic part of the calculations in Equation II–40 and Equation II–41. Therefore Equation IV–7 and Equation IV–8 should be accurate when $f_{\rm V} \lesssim 1$.

Since the amount by which the voxels are partially-filled in our galactic models is substantial (consider Figure IV–8) and we focus on the Galactic plane ($l = 0$), we use uniformly-filled voxels in our models to ensure the correct column density. This means intrinsic properties such as mass require a correction:

$$M'_{{\rm vox},i} = f_{\rm vox}\, M_{{\rm vox},i} \ , \tag{IV–9}$$

where the primed mass is corrected for the partially-filled voxel, and i may refer to either H^0 or H_2.

4 Radiative transfer

Having come to a solution to modify the partially-filled voxels within our models, we must now utilise the emissivity and absorption in all voxels to calculate our integrated intensity maps. Even though we simplify the problem by considering just the sightlines at $b = 0$, we will still have issues of the pathlengths of the sightlines passing through each voxel are not $\ell_{\rm vox}$ as in §III-4. This results from configurations where the voxels are not viewed face-on[21]. Not only will this affect

[21]This is largely a result of the observations being taken *within* the Milky Way. Other objects (PDRs, galaxies, AGNs, etc.) modelled with `kosmatau3d` can be constructed so the voxels are basically assumed to be face-on.

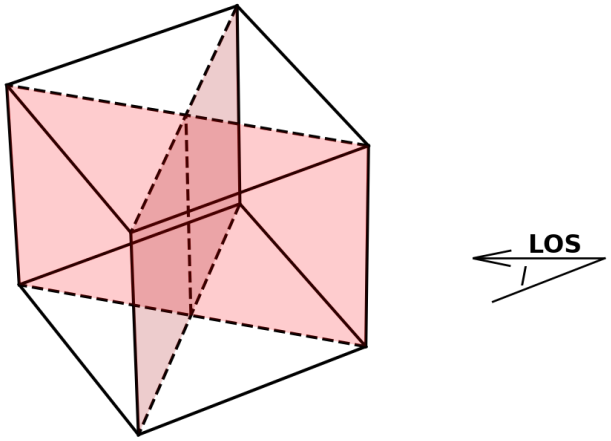


Figure IV–11: The structure used to define the voxel in the `kosmatau3d` radiative transfer module at $b = 0$. The red surfaces denoting diagonals in the x, y-plane and their intersection in $\hat{\mathbf{z}}$ are depicted with dashed black lines.

how we integrate Equation I–3, but this will make the determination of which voxel is in the line-of-sight more complicated.

We will need to begin with how we determine if a voxel exists in a line-of-sight. Since we confine our analysis to the plane of the galaxy, we only need to concern ourselves with rotations about the $\hat{\mathbf{z}}$ axis. In this sense, we only need to consider the diagonals in the x, y-plane in order to determine if a voxel crosses the line-of-sight (see Figure IV–11). For a voxel length of $\ell_{\rm vox}$, the apparent physical width $\Delta w_{\rm vox}$ of each voxel will be within $\Delta w_{\rm vox} \in [1, \sqrt{2}]\,\ell_{\rm vox}$. We thus use the distance to, longitude, and angular width of each voxel ($r_{\rm vox}$, $l_{\rm vox}$, and $\Delta l_{\rm vox}$, respectively) to compare to the line-of-sight $l_{\rm LOS}$. The voxel positions are tracked using their centre points, so the necessary calculations with respect to the reference position for the LSR at $R_{\rm gal,\odot}$ are:

$$r_{\rm vox} = \sqrt{(x_{\rm vox} - R_{\rm gal,\odot})^2 + y_{\rm vox}^2}\ , \tag{IV–10}$$

$$l_{\rm vox} = \arctan\left(\frac{y_{\rm vox}}{x_{\rm vox} - R_{\rm gal,\odot}}\right)\ , \tag{IV–11}$$

$$\Delta l_{\rm vox} = \frac{\sqrt{2}\,\ell_{\rm vox}}{r_{\rm vox}} \times \max\left(\sin\left(l_{\rm vox} \bmod \pi + \pi/4\right), \sin\left(l_{\rm vox} \bmod \pi - \pi/4\right)\right)\ , \tag{IV–12}$$

where the criteria for a voxel to be in a line-of-sight $i_{\rm vox,LOS}$ is,

$$i_{\rm vox,LOS} = \begin{cases} \text{True} & |l_{\rm LOS} - l_{\rm vox}| \leq \Delta l_{\rm vox}/2,\ r_{\rm vox} > \ell_{\rm vox}/2 \\ \text{True} & |l_{\rm LOS} - l_{\rm vox}| \leq \pi, r_{\rm vox} \leq \ell_{\rm vox}/2 \\ \text{False} & \text{else} \end{cases} \tag{IV–13}$$

Here the two criteria for a voxel to be in the line-of-sight are split between $r_{\rm vox} = {}^{\ell_{\rm vox}}/_{2}$ due to the modelling constraint: below this limit it is possible for the reference position to exist within a voxel where $l_{\rm vox} \in [-\pi, \pi]$ (in this case the synthetic observation is taken from a position inside of a voxel, but the centre of the voxel is opposite the sightline). For $r_{\rm vox} > {}^{\ell_{\rm vox}}/_{2}$ it is sufficient to ensure $l_{\rm LOS}$ and $l_{\rm vox}$ align within ${}^{\Delta l}/_{2}$ to minimise the artefacts[22]

Having determined which voxels are in a given line-of-sight, we then arrange them in order of decreasing $r_{\rm vox}$ in order to integrate the radiative transfer equation (since we must integrate starting at the point furthest from our reference). Given our method, the number of voxels in our calculation $N_{\rm LOS}$ satisfy the equation,

$$N_{\rm LOS}\ell_{\rm vox} \geq D_{\rm LOS}, \tag{IV–15}$$

where $D_{\rm LOS}$ is the modelling distance along our LOS (for our Milky Way model, it is the distance to the edge of the Milky Way). Qualitatively, Equation IV–15 is satisfied since the LOS passes through portions of the voxel for voxels that are not viewed face-on (ie. when $l_{\rm vox} \neq \frac{i\pi}{2} \,\forall\, i \in \mathbb{Z}$). Rather than go through the tedious process of defining the path length for each individual voxel $s'_{{\rm vox},i} \,\forall\, i \in [0, N_{\rm LOS} - 1)$, which anyways gets much more complicated when one considers a beam width (and therefore the path length is not a single value for each voxel), we define an average normalised voxel size:

$$\ell'_{\rm vox} \equiv \frac{D_{\rm LOS}}{N_{\rm LOS}} \ . \tag{IV–16}$$

One clear assumption inherent in this approach is that the voxel properties along the LOS are smoothly-varying, which is valid for the Milky Way models presented in this thesis. A more-sophisticated method of integration is required if we need to account for voxels with very different properties.

5 Galactic model grids

In order to constrain the parameter space in Table III–2, we evaluated a few grids of `kosmatau3d` Milky Way models, which are summarised in Table IV–2. The first model that is necessary is the `convergence` model; it is used to ensure the synthetic observation is quite similar for models with different voxel size. Due to some of the features of the [^{12}C II] observations that will be mentioned in §V-2.1, there is also the grid `FUV-Galactic-centre` varying the far-UV radiation and size of the Galactic centre (shown by the subscript GC). For this we vary $f_{\rm FUV}$ along with the radius $R_{\rm GC}$ of the scaled region. Another parameter affecting the Galactic centre is the primary cosmic ray ionisation rate $\zeta_{\rm H}$, so we also compute a model grid to explore the effect of this. Following the description in Padovani et al.

[22] One could alternatively get the beamsize $\Delta l_{\rm beam}$ from the desired resolution of the map, in which case the criterion is,

$$|l_{\rm LOS} - l_{\rm vox}| \leq \frac{\Delta l_{\rm beam} + \Delta l_{\rm vox}}{2}. \tag{IV–14}$$

.

(2020), we adopt a galactic model where the galactic disk has an average cosmic ray ionisation rate equivalent to the local value of $2 \times 10^{-16}\,\mathrm{s}^{-1}$ and the CMZ has an average rate of $10^{-14}\,\mathrm{s}^{-1}$. Thus the model grid `CMZ-radius-radiation` varies the CMZ radius[23] R_{CMZ} as well as the far-UV radiation f_{FUV} throughout the model. We also wanted to highlight the traits affecting the [^{12}C II] 158 μm line observations, which we examine via the `atomic-ISM-radiation`. This varies the fraction of the atomic ISM represented by the interclump medium rather than the clumpy medium, as well as the far-UV radiation. We also evaluated a model grid `mass-factors` varying the mass distributions used to create the models: $f_{\mathrm{m,H_2}}$, $f_{\mathrm{m,H^0}}$. Finally, we explored the effectiveness of using PDRs to model the WNM in the `three-media` model grid, where we vary the fraction f_{WNM} of the interclump ensemble in the WNM and the far-UV radiation χ_{WNM} in which it is embedded. From these models, it should be possible to not only constrain the underlying galactic profiles but also reach a conclusion on the applicability of PDRs to modelling the galactic cooling lines.

The fiducial model with parameters in Table IV–1, which we compare to the `kosmatau3d` model grids listed in Table IV–2.

Table IV–1: The default values of the parameters in Table IV–2.

Parameter	Default value
ℓ_{vox}	400 pc
f_{FUV}	1
$f_{\mathrm{m,H^0}}$	1
$f_{\mathrm{m,h2}}$	1
$f_{\mathrm{H^0}}$	1
f_{WNM}	1
$f_{\mathrm{FUV,GC}}$	1
R_{CMZ}	0 pc
R_{GC}	0 pc
σ_{GMC}	4.47 kms
χ_{cl}	None
χ_{WNM}	None
f_{FUV}	1

[23] It should be noted that R_{CMZ}, is different than R_{GC}, even though they seem similar in Table IV–2. We use the size of the CMZ to determine the voxel for which we can use a higher cosmic ray ionisation rate. The size of the Galactic core is used to scale parameters such as the far-UV intensity in the Galactic core.

Table IV–2: A list of all grids, which parameters are varied and what values, and how they are referred in this work.

Parameters	Values	Name
ℓ_{vox}	$(400, 200, 100)$ pc	convergence
σ_{ens}	$(0.001, 1.10, 1.86, 2.63$ $4.47, 6.33, 10.72)$pc	voxel-dispersion
$\log_{10}(\chi_{\mathrm{cl}})$ $f_{\mathrm{FUV,icl}}$	$(0.1, 0.3, \ldots, 2.5)$ $(10^{0.25}, 10^{0.5}, \ldots, 10^{3})$	ensemble-FUV
$f_{\mathrm{H^0}}$ $\log_{10}(f_{\mathrm{FUV}})$	$(0.3, 0.4, 0.5, 0.6, 0.7)$ $(0.000, 0.125, \ldots, 2.000)$	atomic-ISM-radiation
R_{GC} $\log_{10}(f_{\mathrm{FUV,GC}})$	$(200, 600, 1000, 1400)$ pc $(0.5, 1.0, 1.5, 2.0, 2.5, 3.0)$	FUV-galactic-centre
$\log_{10}(f_{\mathrm{FUV}})$ R_{CMZ}	$(0.000, 0.125, \ldots, 2.000)$ $(0, 50, \ldots, 1000)$ pc	CMZ-radius-radiation
$\log_{10}(m_{\mathrm{cl}})$ $\log_{10}(f_{\mathrm{FUV}})$	$([0, 2], [-1, 2], [-1, 3], [0, 3])$ $(1, 1.5, 2, 2.5, 3, 3.5, 4)$	clump-masses
$f_{\mathrm{m,H_2}}$ $f_{\mathrm{m,H^0}}$	$(0.25, 0.5, 1, 2, 4)$ $(0.25, 0.5, 1, 2, 4)$	mass-factors
f_{WNM} $\log_{10}(\chi_{\mathrm{WNM}})$	$(0.2, 0.3, 0.4, 0.5)$ $(2, 2.5, 3, 3.5, 4)$	three-media

5.1 Convergence

When considering the agreement between galactic models with different voxel sizes, at least in our approach in using axisymmetric profiles, the difference between the *convergence* models will be in the total number clumps for each ensemble, the ensemble velocity dispersion, and the voxel-filling factor. The total number of clumps scales with the voxel volume, so this should not greatly affect our results when observing just the galactic plane. Similarly, the velocity dispersion should affect the resolution of certain features in the overall intensity map, but on average should not cause major differences between our *convergence* models. We define a framework in which the voxel-filling factor does not affect the intensity datacubes, but rather it will affect the interpretation of properties such as H^0 mass and H_2 mass (see §IV-3). Therefore we focus our analysis of this model comparison on the agreement between the synthetic observation maps.

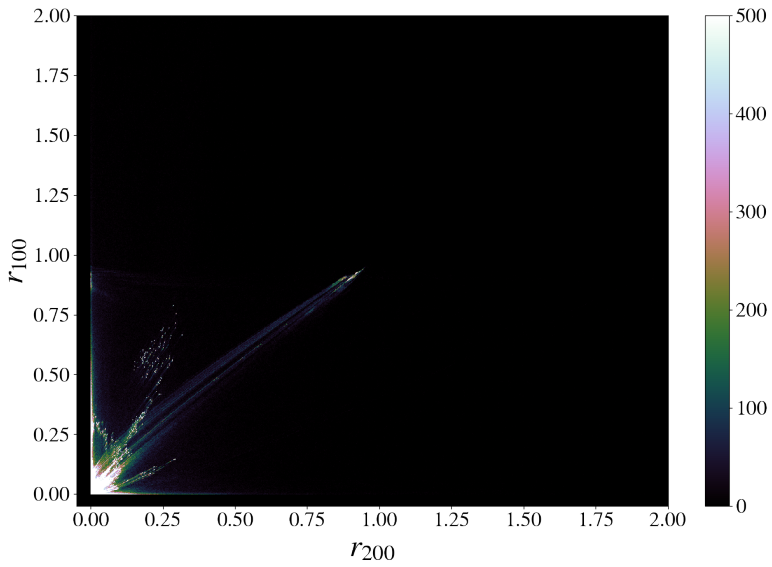


Figure IV–12: A histogram of the square residual of the 200 pc and 100 pc models compared to the 400 pc model, computed with Equation IV–17.

$$r_{\text{model}} = \frac{\left(T_{\text{B},400} - T_{\text{B,model}}\right)^2}{T_{\text{B},400}^2} \tag{IV–17}$$

In order to quantify how much the models differ from each other, we examine the square residual r_{model} which should be < 1 for models are similar to each other. We calculate r_{model} for the datacube of all transitions included in our models (see IV–12), using as a reference the 400 pc. IV–12 shows that $r_{200} < 1$ and $r_{100} < 1$ for most of the datacubes, indicating that there is no clear artefact introduced by changing the voxel size.

CHAPTER V

Milky Way observations

What we observe is not nature itself, but nature exposed to our method of questioning.

Werner Heisenberg

From the discussion in §IV-3, the ensuing analysis will focus on observations with latitude $b = 0$. This will allow us to utilise the velocity resolution of our models and fit trends on a galactic scale rather than constraining galactic arm morphology. Luckily there are are a few different Milky Way surveys focusing on the galactic plane. Although each trace a different wavelength or region of the galaxy, we basically have low-J CO line transitions, ^{12}C I, or dust observations to constrain our galactic models. The spectroscopic surveys in particular are interesting to use as a constraint since they have clear signs of self-absorption which are modelled in `kosmatau3d`.

1 Regridding

The majority of the surveys used to constrain the models of Chapter IV have spatial and spectral higher than the synthetic observations. For this reason we first resample the velocity (spectroscopic) axis, if applicable, to the resolution of the synthetic data (1 km s^{-1}) using the python package `SpectRes` (Carnall 2017). While this might remove some of the finer velocity features, it should help to minimise the noise. We then use `cygrid` (Winkel et al. 2016b) to regrid the data to according to the resolution in longitude and latitude of the synthetic maps.

Table V–1: The axes dimensions to which we regrid the observational surveys. The analysis discussed in this thesis only utilise the subset with $b = 0°$

Axis	Range	Spacing	N_{axis}
l_{gal}	$[-180, 180]°$	$0.5°$	721
b_{gal}	$[-0.5, 0.5]°$	$0.5°$	361
v_{obs}	$[-350, 350]\,\text{km s}^{-1}$	$1\,\text{km s}^{-1}$	701

2 Line emission surveys

Our focus for line transition comparisons are on a few of the carbonaceous species transitions, namely ^{12}CO, ^{13}CO, ^{12}C, and ^{12}C^{+} (with ^{12}C being the only line without spectroscopic data). The fundamental reference for many spectroscopic surveys which we also use is the CfA survey (^{12}CO $J = 1 \rightarrow 0$; Dame et al. 2001). Subsequent surveys have drastically improved both in angular and spectroscopic resolution. We use ThrUMMS (^{12}CO $J = 1 \rightarrow 0$, ^{13}CO $J = 1 \rightarrow 0$; Barnes et al. 2015), Mopra (^{12}CO $J = 1 \rightarrow 0$, ^{13}CO $J = 1 \rightarrow 0$; Braiding et al. 2015), and SEDIGISM (^{13}CO $J = 2 \rightarrow 1$; Schuller et al. 2021). We also use the older data from *COBE*-FIRAS (Fixsen et al. 1999), which provides integrated intensities for various transitions in [^{12}C II], ^{12}C I, and ^{12}CO.

2.1 GOT C+

Spectroscopic [^{12}C II] 158 μm (1.9 THz) data for the galactic plane was taken for the first time in the Galactic Observations of Terahertz C+ (GOT C+) project from Pineda et al. (2013) (see Figure V–1). They utilised the HIFI instrument onboard *Herschel* to investigate [^{12}C II] as a tracer of the various phases of the ISM.

At the time of publication, they chose have the data available as tables for each of the 452 sightlines. These tables were then consolidated and resampled to the according to Table V–1. Even though HIFI had a resolution of ~ 15 arcsec, they still needed to sparsely-sample the observations over galactic longitudes away from the galactic centre. Subsequent publications have focused on narrower windows to examine the fine detail of the CMZ and the data was released in the standard `FITS` format (Langer et al. 2017), but we focus our analysis on the GOT C+ strips. This dataset thus does not have a regular grid.

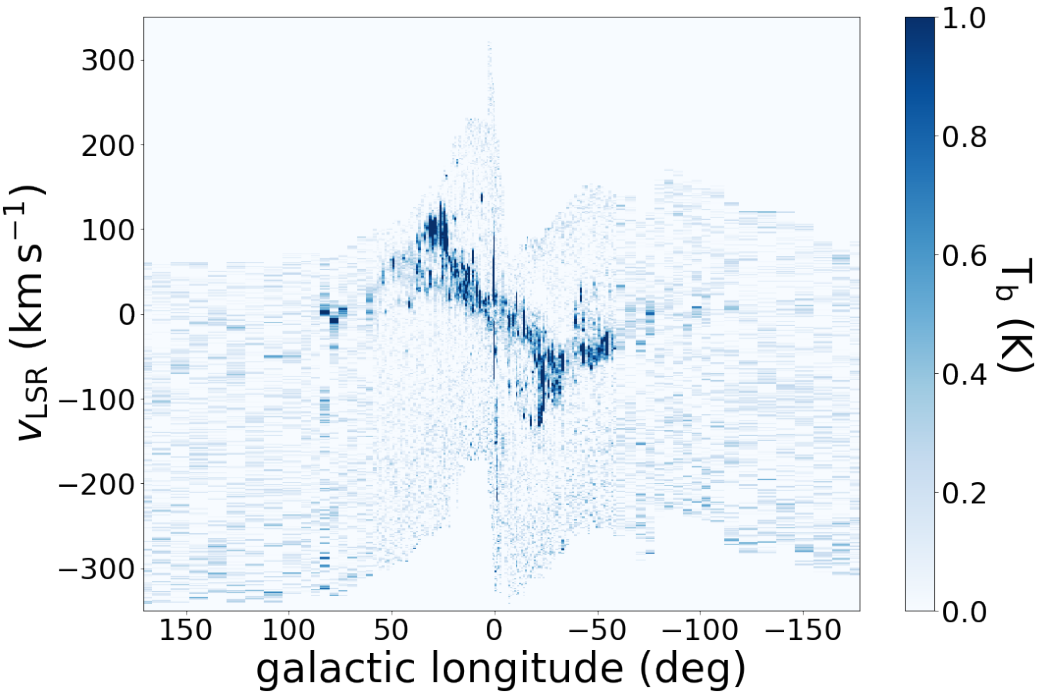


Figure V–1: The position-velocity diagram of the GOT C+ data, resampled according to Table V–1. Note the irregular spacing in longitude results in larger bins away from the galactic centre.

2.2 *COBE*-FIRAS

While it is the oldest survey used in this paper, it is arguably one of the most important. The *Far-Infrared Absolute Spectrometer* (FIRAS) instrument on the *Cosmic Background Explorer* (*COBE*) observed a spectrum with $104\,\mu\text{m} < \nu < 4400\,\mu\text{m}$, from which they could determine a few line transitions. Using a 7° beam, [^{12}C II]and the first 8 rotational transitions of ^{12}CO were observed [24](see Figure V–2). The full-sky maps are provided in Bennett et al. (1994), but the analysis in Fixsen et al. (1999) assumed the intensity in the 7° extent in latitude originates from the galactic plane with $\Delta\beta \approx 1°$ (see Figure V–2)[25] to acquire the emission from the galactic plane. This recovered higher intensities and more lines, and has been used in subsequent model comparisons (eg. Cubick et al. 2008).

We focus on the Fixsen et al. (1999) results in order to perform a direct comparison to our models. This reduces the error associated to the partially-filled voxels in §IV-3, since the original Bennett et al. (1994) would require us to regrid the synthetic observations to a resolution of 7° as well as the partially-filled voxel.

[24]The ^{12}CO $J = 7 \rightarrow 6$ line was unable to be separated from the ^{12}C I 370 μm line, so they provide the sum of the two.

[25]Acquired through personal contact.

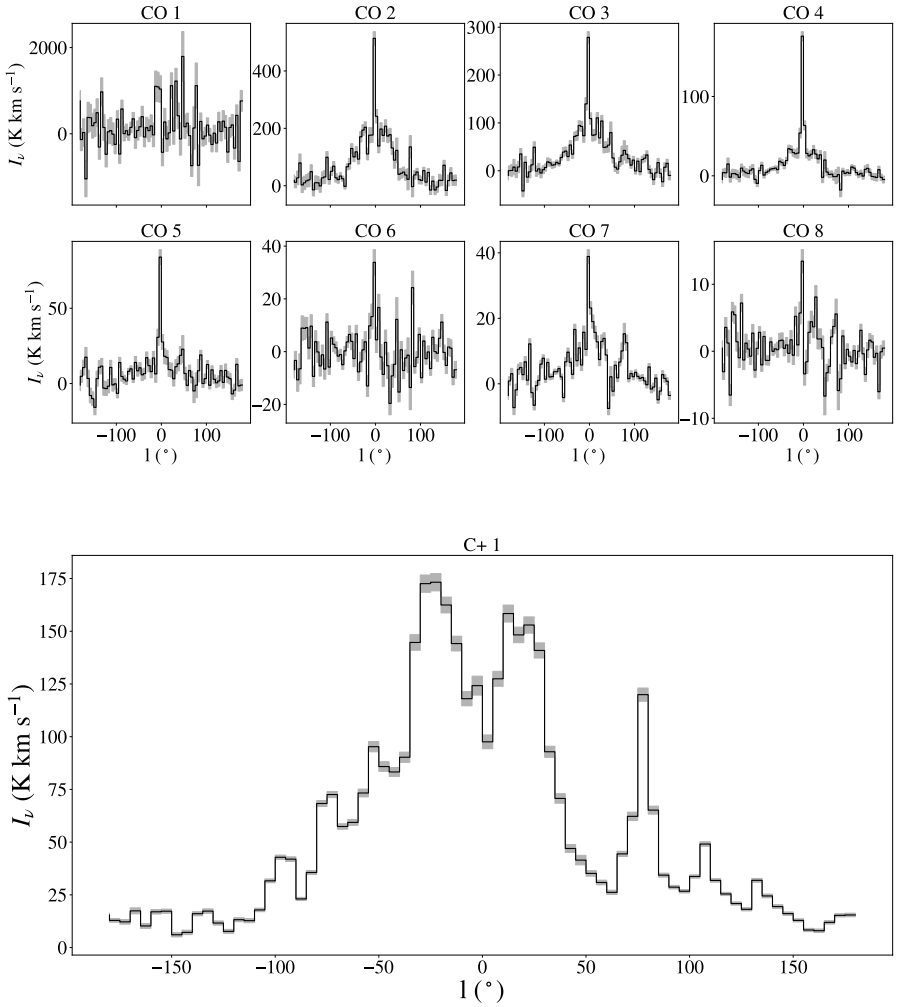


Figure V–2: The *COBE*-FIRAS galactic plane observation Fixsen et al. (1999) calculated from the publicly available data. The labels correspond to the transitions in Table III–4.

2.3 CfA

The seminal ^{12}CO $J = 1 \rightarrow 0$ survey and galactic structure analysis, this amalgamation of observations opened a new era for far-IR astronomy (see Figure V–3). It proved that galactic spectroscopic observations are critical methods of analysing the spiral arm structure and gas heating. The dataset composited several galactic surveys using 1.2m antennae in Harvard-Smithsonian Center for Astrophysics, USA and the Cerro Tololo Interamerican Observatory, Chile (Dame et al. 2001). The angular resolution of these ranged from $0.5°$ to $0.125°$, and the spectral resolution varied depending on the source of each survey. Higher resolution was used for the majority of the map, and was supplemented with lower-resolution data from a previous survey (Dame et al. 1987). As a result of the various surveys used to

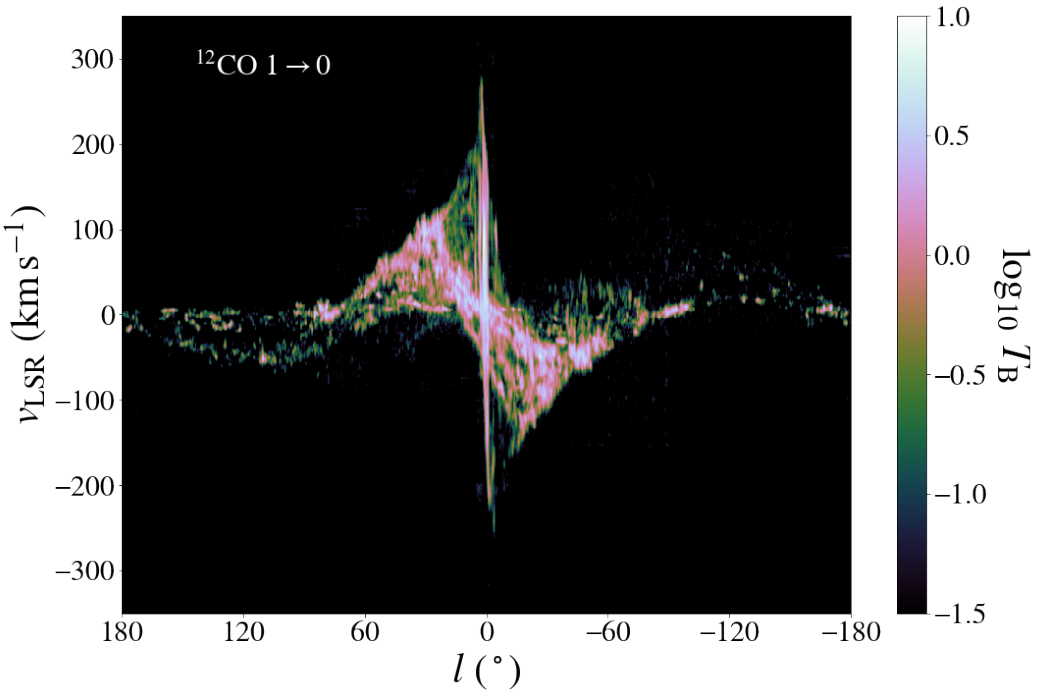


Figure V–3: The CfA ^{12}CO $J = 1 \rightarrow 0$ data resampled and regridded according to Table V–1.

create this dataset, it is contained in a partially-filled datacube[26] covering 360 in longitude, 10 in latitude, and 360 km s^{-1} in V_{LSR}.

2.4 Mopra

The Mopra southern Galactic plane CO Survey maps the $J = 1 \rightarrow 0$ transition for ^{12}CO, ^{13}CO, C^{18}O, and C^{17}O in the fourth galactic quadrant. It utilised the eponymous instrument to cover $320 \leq l \leq 327$, $|b| \leq 0.5$ as well as $327 \leq l \leq 330$, $-0.5 \leq b \leq 1.0$. We utilise the observations of ^{12}CO and ^{13}CO (see Figure V–4 and Figure V–5) to constrain our galactic models in Chapter VI.

[26] A datacube requires a regular three-dimensional grid. Appending several surveys of varying extents in longitude, latitude, and LSR velocity means that there can be gaps between them. These are filled in with undefined (`NaN`) values.

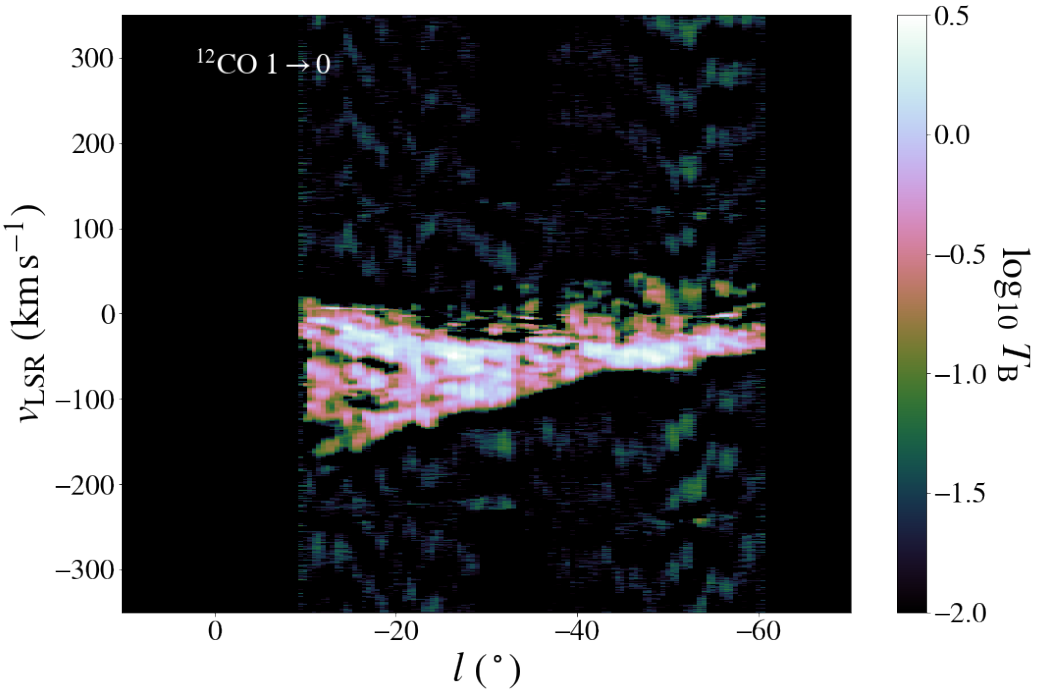


Figure V–4: The Mopra ^{12}CO $J = 1 \rightarrow 0$ data resampled and regridded according to Table V–1.

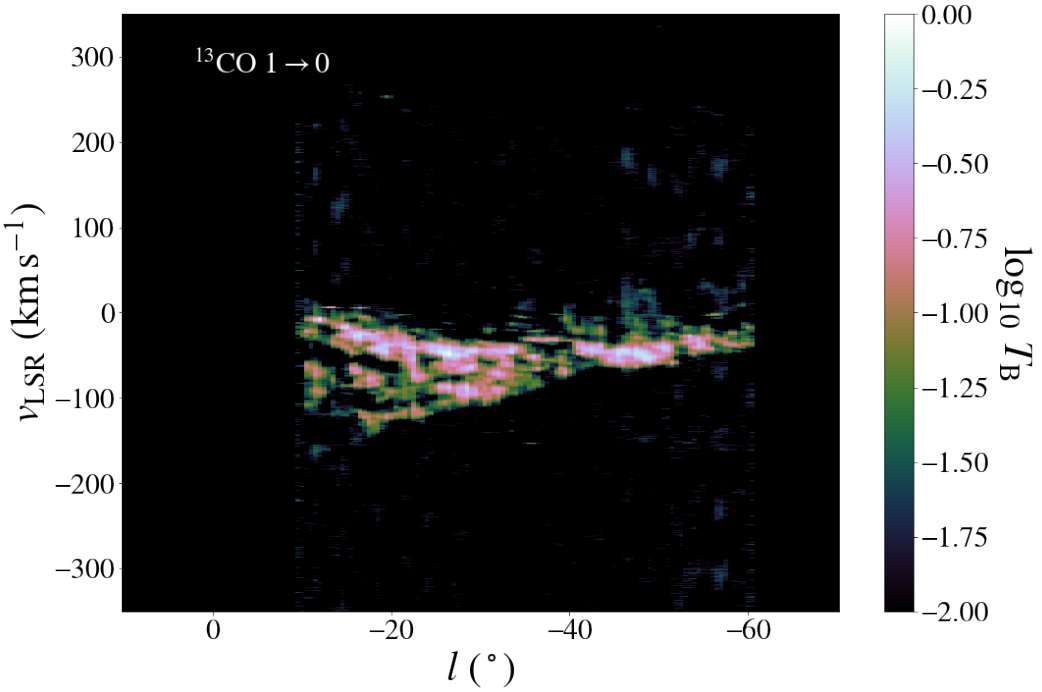


Figure V–5: The Mopra ^{13}CO $J = 1 \rightarrow 0$ data resampled and regridded according to Table V–1.

2.5 ThrUMMS

The Three-mm Ultimate Mopra Milky Way Survey (ThrUMMS) is another independent Mopra survey of the fourth galactic quandrant of the $J = 1 \rightarrow 0$ of ^{12}CO, ^{13}CO, C^{18}O, and CN. The hallmark of this survey is the improved observational techniques explained in Barnes et al. (2015). We utilise the observations of ^{12}CO and ^{13}CO (see Figure V–6 and Figure V–7) to constrain our model grids.

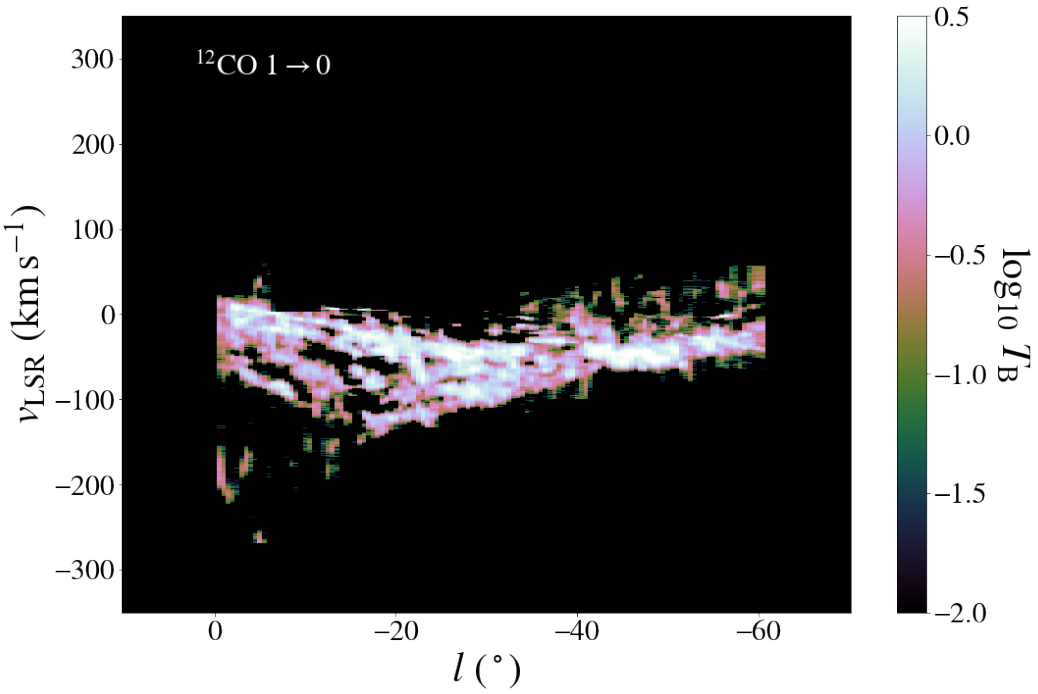


Figure V–6: The ThrUMMS ^{12}CO $J = 1 \rightarrow 0$ data resampled and regridded according to Table V–1.

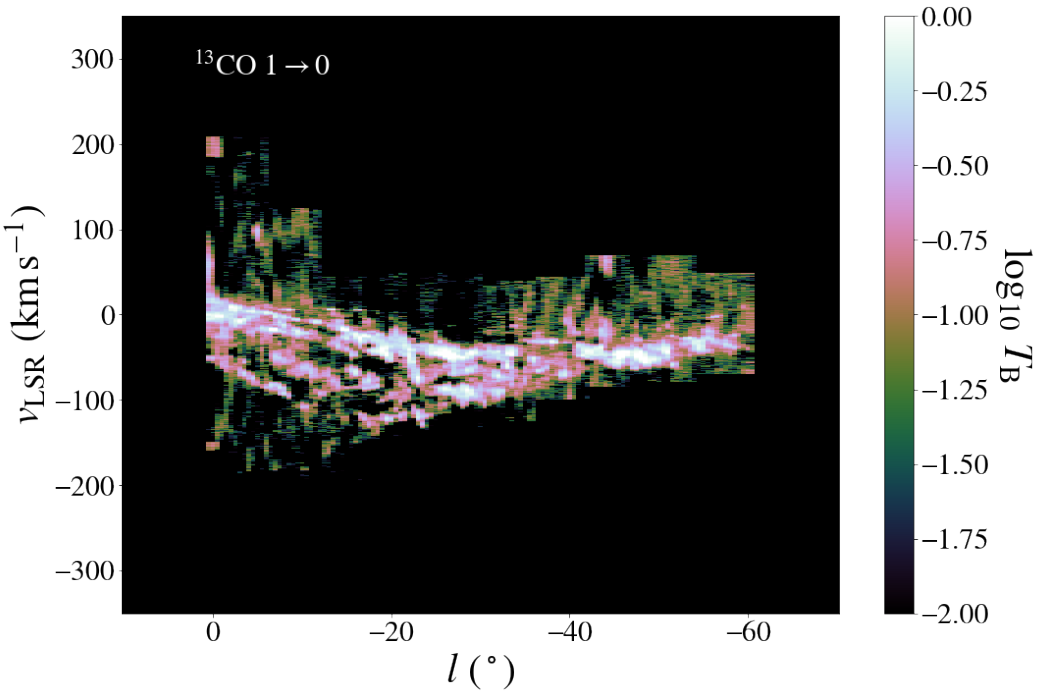


Figure V–7: The ThrUMMS ^{13}CO $J = 1 \rightarrow 0$ data resampled and regridded according to Table V–1.

2.6 SEDIGISM

The Structure, Excitation, and Dynamics of the Inner Galactic Interstellar Medium (SEDIGISM) survey maps the $J = 2 \rightarrow 1$ rotational transition line of ^{13}CO and C^{18}O near the galactic center (Schuller et al. 2017, 2021) (see Figure V–8). Being more optically-thin than the typical CO tracer, these transitions should be more sensitive to the denser spiral structure of the Milky Way. The observations were taken with the $12m$ diameter Atacama Pathfinder Experiment Telescope (APEX) in Chile, covering regions with $-60 \leq l \leq 18$ and $|b| \leq 0.5$ with an angular resolution of $30''$. The data is split into several maps covering 2 in longitude and 1 in latitude. This is the only spectroscopic survey of COwe use in the analysis with a rotational transition $J_{\mathrm{u}} \neq 1$.

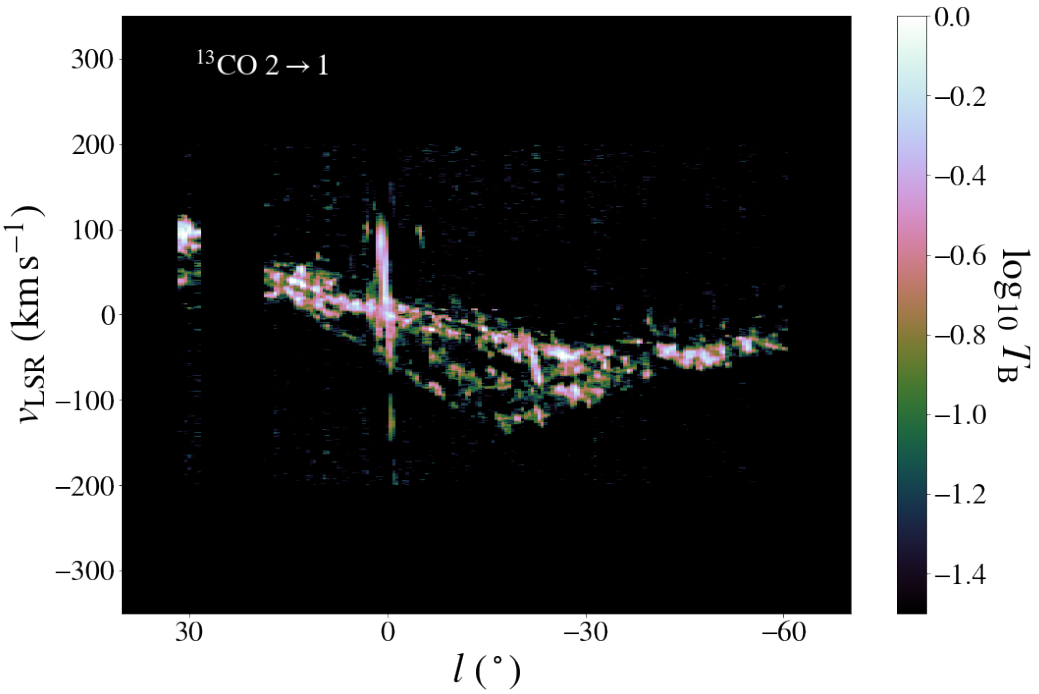


Figure V–8: The SEDIGISM ^{13}CO $J = 2 \rightarrow 1$ data resampled and regridded according to Table V–1.

2.7 H I data

We utilise H I data from the HI4PI[27] (HI4PI Collaboration et al. 2016), which is the name given to the combination of the Effelsberg-Bonn H I Survey (EBHIS; Kerp et al. 2011; Winkel et al. 2016a) and Galactic All-Sky Survey (GASS; McClure-Griffiths et al. 2009; Kalberla & Haud 2015). The combined data cover the galactic plane (see Figure V–9).

[27] http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/594/A116#/browse

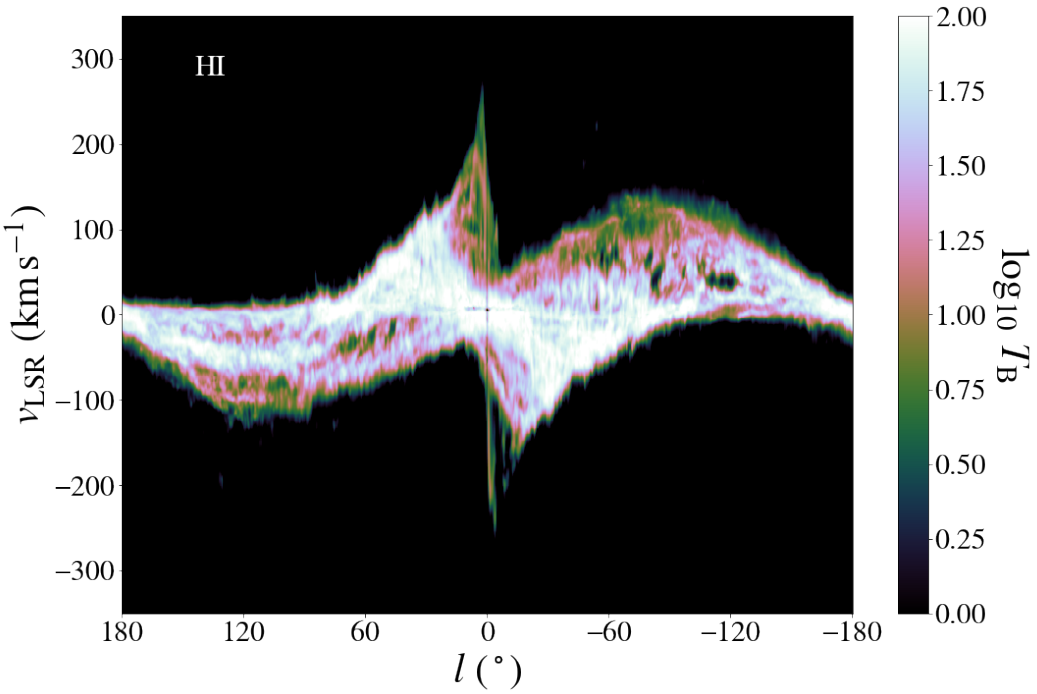


Figure V–9: The combined H I data resampled and regridded according to Table V–1.

3 Continuum emission

For the observed dust intensity, we utilise the data products from the *Planck* collaboration. Our options for the warmer dust emission ideally include `Herschel`, but the data was unavailable during the analysis featured in this thesis. We must therefore resort to the *COBE*-DIRBE results, specifically the 240 μm dataset. We are anyways comparing the data at such a low resolution that the *COBE*-DIRBE results should sufficiently constrain our galactic models.

3.1 *COBE*-DIRBE

With *COBE* we can also utilise data from the *Diffuse InfraRed Background Experiment* (DIRBE) for the dust emission. Its observation of the intensity at $\lambda = 240\,\mu$m is useful to constrain the dust in the optically-thick regime. Since the observational data is in units of Jansky, we converted to brightness temperature using Equation I–9 (see Figure V–10).

3.2 *Planck*

The *Planck* mission mapped frequencies spanning 30 GHz to 857 GHz in order to determine the cosmic microwave background (Planck Collaboration et al. 2014b). There is an extensive set of publications describing the data and the relevant analysis. This data was combined with both the *Wilkinson* Microwave Anisotropy Probe (WMAP) and the MPIfR 408 MHz survey (Haslam et al. 1982) in order to

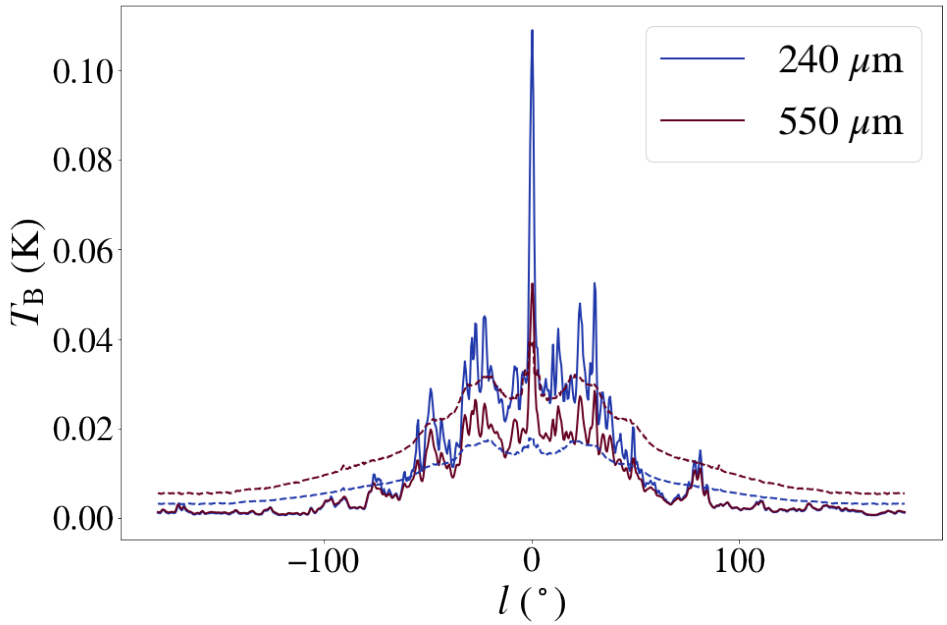


Figure V–10: The 240 μm data from *COBE*-DIRBE (blue) and the 545 μm data from *Planck* (red) regridded according to Table V–1. The identified wavelengths are the wavelengths contained in `kosmatau3d`.

derive also the synchrotron, free-free, spinning dust, and thermal dust emission. The Bayesian analysis code `Commander` is used to fit a parametric model and derive the relevant properties. The available data is the statistics (maximum, mean, and rms) of the posteriors for the three fitted parameters A_d, β_d, and T_d. From these parameters, we use the dust temperature T_d at frequency $\nu = 545\,\mu$m to constrain our galactic models, which is in the optically-thin regime of the dust continuum (see discussion in §I-1.2.4; plot in Figure V–10).

Table V–2: The surveys used with their associated regions in the galactic sphere.

Survey	transition/wavelength	l (°)	b (°)	beam	Δv	ref
GOT C+	[^{12}C II] 158 μm	$[-180, 180]$	$[0]$	$12''$	$0.17\,\mathrm{km\,s^{-1}}$	1, 2
COBE-FIRAS	[^{12}C II] 158 μm, CO up to $J=8\rightarrow7$	$[-180, 180]$	$[0]$	$30'$	N/A	3, 4
CfA	CO $J=1\rightarrow0$	$[-180, 180]$	$[0]$	$15'$	$0.65\,\mathrm{km\,s^{-1}}$	5
Mopra	^{12}CO, ^{13}CO $J=1\rightarrow0$	$[-55, -20]$	$[0]$	$35''$	$0.1\,\mathrm{km\,s^{-1}}$	6
ThrUMMS	^{12}CO, ^{13}CO $J=1\rightarrow0$	$[-60, 0]$	$[0]$	$72''$	$0.3\,\mathrm{km\,s^{-1}}$	7
SEDIGISM	^{13}CO $J=2\rightarrow1$	$[-600, 18]$	$[0]$	$30''$	$0.25\,\mathrm{km\,s^{-1}}$	8, 9
HI4PI	H I	$[-180, 180]$	$[0]$	$16.2'$	$1.29\,\mathrm{km\,s^{-1}}$	10
COBE-DIRBE	continuum (240 μm)	$[-180, 180]$	$[0]$	$39.5'$	N/A	11
Planck	continuum (550 μm)	$[-180, 180]$	$[0]$	$7.5'$	N/A	12

(1) Pineda et al. (2013); (2) Langer et al. (2014); (3) Bennett et al. (1994); (4) Fixsen et al. (1999); (5) Dame et al. (2001); (6) Braiding et al. (2015); (7) Barnes et al. (2015); (8) Schuller et al. (2017); (9) Schuller et al. (2021); (10) HI4PI Collaboration et al. (2016); (11) Hauser et al. (1998); (12) Planck Collaboration et al. (2016);

4 Error Analysis

Determination of the error associated with the aforementioned surveys is crucial for any attempt at model selection. When a map of the survey root-mean-square (RMS) error σ_{RMS} is available, we use it. This sometimes also requires the error to be regridded, for example for the Mopra data, or it is otherwise provided in the source paper. There are particular spectroscopic surveys, such as the CfA survey and GOT C+, where the error must be determined by approximating the standard deviation of the noise. These approaches are summarised in Table V–3.

While σ_{RMS} is important for the error directly associated with the observation, we must define another error due to the configuration of our model used in the comparison. In Chapter IV we saw that our models are axisymmetric, which means we will always have symmetry in the resulting synthetic observation. We therefore define as well a *configuration* error σ_{conf} to account for this (Natale et al. 2022). We must define it separately for the two-dimensional maps (for the integrated maps of *COBE*-FIRAS and the dust maps of *COBE*-DIRBE and *Planck*) and the three-dimensional spectral cubes.

For the integrated intensity and dust maps, this error will consider symmetry in both longitude and latitude:

$$T_{\mathrm{B}}(|l|, |b|) = \text{const} \ . \tag{V–1}$$

Table V–3: A summary of the RMS error determination for the surveys in §V-2 and §V-3.

Survey	$\overline{\sigma_{\mathrm{RMS}}}$	Regridded	Calculated	Notes
COBE-FIRAS	$< 17.7\,\mathrm{K\,km\,s^{-1}}$	No	No	
CfA	$5.38\times10^{-3}\,\mathrm{K}$	Yes	Yes	
GOT C+	$0.109\,\mathrm{K}$	No	Yes	
Mopra	$1.75\times10^{-2}\,\mathrm{K}$	Yes	No	^{12}CO
	$0.625\times10^{-2}\,\mathrm{K}$	Yes	No	^{13}CO
ThrUMMS	$5.69\times10^{-2}\,\mathrm{K}$	Yes	No	^{12}CO
	$1.98\times10^{-2}\,\mathrm{K}$	Yes	No	^{13}CO
SEDIGISM	$1.15\times10^{-2}\,\mathrm{K}$	Yes	No	
H I	$2.20\times10^{-2}\,\mathrm{K}$	Yes	No	
COBE-DIRBE	$3.03\times10^{-5}\,\mathrm{K}$	Yes	No	
Planck	$6.45\times10^{-7}\,\mathrm{K}$	Yes	No	

For the spectroscopic datacubes, the configuration error symmetry in longitude and observing velocity is antisymmetric and in latitude it is symmetric:

$$T_{\mathrm{B}}(v_{\mathrm{obs}}, l, |b|) = T_{\mathrm{B}}(-v_{\mathrm{obs}}, -l, |b|) \ . \tag{V–2}$$

If we consider a subset of pixels in the datacube that should be equivalent due to symmetry[28], $\mathcal{T} \equiv [T_{\mathrm{B},0}, \ldots, T_{\mathrm{B},m}]$, we first find the average brightness temperature,

$$\overline{T_{\mathrm{B}}} = \frac{1}{m} \sum_{i=0}^{m} T_{\mathrm{B},i} \ , \tag{V–3}$$

then use this to compute sample-independent standard deviation:

$$\sigma_{\mathrm{conf}} = \sqrt{\frac{1}{m-1} \sum_{i=0}^{m} \left(T_{\mathrm{B},i} - \overline{T_{\mathrm{B}}}\right)^2} \ . \tag{V–4}$$

Here we used the Bessel correction to ensure we calculate the unbiased standard deviation of the subset.

[28] By the symmetry of our model defined by the relations Equation V–1 and Equation V–2, this subset has two elements for $l = 0$ and four elements for $l \neq 0$

CHAPTER VI

Model selection

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.

Richard P. Feynman

The resolution of this comparison is set by the resolution of the synthetic observation, which is in turn limited by the available computational resources. While one might think the real observation should provide the main constraint, the current observational surveys have such high precision that one cannot simulate anything close without a supercomputing cluster and hundreds of terabytes worth of storage for each model. In the absence of this, we decrease the spatial and spectral resolution of the observations to agree with the resolution of our models.

The benefit of comparing to many surveys is the variety of line transitions which can be used to constrain our Milky Way model. This has the price of comparing to files with a unique definitions for longitude, latitude, and observing velocity ($l_{\rm gal}$, $b_{\rm gal}$, and $v_{\rm LSR}$, respectively). These in turn differ from the standard adopted in the synthetic observations.

1 Method

We ultimately employ two basic methods for comparing real and synthetic observations due to the nature of the observational surveys. For the integrated intensity maps in *COBE*-FIRAS and the dust continuum maps from *COBE*-DIRBE and *Planck*, we must work with the corresponding integrated intensity and dust continuum maps given by the models. We further simplify our comparison by focusing on the galactic plane at $\beta_{\rm gal} = 0$. For the spectroscopic observations, however, we should exploit the full velocity information. Therefore we consider the position-velocity slice from the galactic disk ($\beta_{gal} = 0$). This allows us to fit the primary features of the CMZ and the overall velocity structure of the Milky Way. Comparing the inte-

grated velocity maps from the spectroscopic surveys will at best provide a sanity check for the aforementioned integrated maps.

That explains *what* will be compared between models and observations, but there is still the question of *how* these will be compared. We utilise the χ^2 test for each transition (or continuum wavelength) in each survey, using Equation VI–1.

$$\chi^2 = \sum_i \frac{(x_{\mathrm{obs},i} - x_{\mathrm{model},i})^2}{\sigma_i^2} \ , \tag{VI–1}$$

where x_i is a observable we want to fit and σ_i is the observational error. Here the index i refers to an element in the subset of the data used in the comparison, which in our case is a function of longitude and, for the spectroscopic data, velocity. While this is sufficient to constrain the model for each transition, we can include multiple transitions in a particular likelihood analysis to also fit the relative intensities of different transitions.

The error associated with a given observation enters in two ways, that is, multiplicative or additive (see Equation VI–2). The additive error is random and spurious, occurring from instrument noise or small glitches from cosmic rays (Fixsen et al. 1994). The calibration error manifests itself as a multiplicative quantity, depending on the intensity incident on the receiver, T_{sky}. In that case, we have that the calibration error dominates in the parts of the spectrum where we observe emission, whereas the additive error dominates where we do not have emission.

$$T_{\mathrm{B,obs}} = T_{\mathrm{B,sky}} \, (1 + \sigma_{\mathrm{cal}}) + \sigma_{\mathrm{rms}} \ , \tag{VI–2}$$

giving use a total observational error,

$$\sigma_{\mathrm{T}} = T_{\mathrm{B,sky}} \, \sigma_{\mathrm{cal}} + \sigma_{\mathrm{rms}} \ . \tag{VI–3}$$

χ^2-test statistic for $\log_{10} T_{\mathrm{B}}$

In our application of this equation, we treat reciprocal multiplicative factors of the brightness temperature as equally likely. We do this by comparing the logarithm of brightness temperature for a given transition/wavelength, $\log_{10} T_{\mathrm{B}}$. The error of $\log_{10} T_{\mathrm{B}}$ is then,

$$\begin{aligned} \sigma_{\log T} &= \delta \left(\log_{10} T_{\mathrm{B},\nu} \right) \ , \\ \sigma_{\log T} &= \delta \left(\frac{\ln T_{\mathrm{B},\nu}}{\ln 10} \right) \ , \\ \sigma_{\log T} &= \frac{1}{\ln 10} \frac{\sigma_T}{T_{\mathrm{B}}} \ , \end{aligned} \tag{VI–4}$$

where δ represents the error of a function, and σ_{T} is the error of the observation. The associated χ^2 statistic for each transition t thus becomes,

$$\chi^2_{\mathrm{t}} = \sum_i \frac{\left(\log_{10} T_{\mathrm{B,t,obs},i} - \log_{10} T_{\mathrm{B,t,syn},i}\right)^2}{\sigma^2_{\log T,\mathrm{t}}} \,. \tag{VI–5}$$

For each survey, the χ^2 statistic is the sum of its observed transitions:

$$\chi^2 = \sum_{\mathrm{t}} \chi^2_{\mathrm{t}} \,. \tag{VI–6}$$

One caveat that results from Equation VI–5 is that it effectively removes the additive error contribution, leaving only the multiplicative error. By invoking the χ^2 test, we make the inherent assumption that the error is normally-distributed.

2 Direct comparison

Now we would like to compare directly our models with the observations using Equation VI–6. What this presents for the spectroscopic data is an opportunity to study the accuracy of the fit over the entire position-velocity plane. Due to the axisymmetric configuration of our models, the emission is much more extended in the PV-plane compared to the observations and we are thus fitting synthetic signal to observational noise dominated by the additive error, resulting in some arbitrarily-large $\chi^2_{\nu,i}$ statistic that introduce computational difficulties when constraining the fit of the model to the observation.

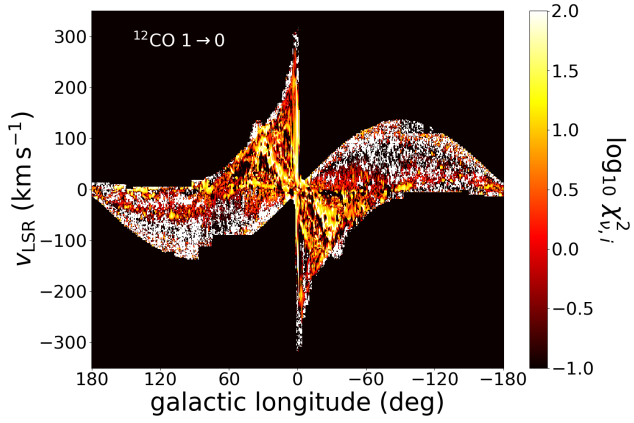


Figure VI–1: The χ^2 statistic in the PV plane comparing the CfA ^{12}CO $J = 1 \rightarrow 0$ observations to the `r400_convergence` model. The white pixels have $\chi^2_{\nu,i} > 10^{300}$.

To illustrate this, we can examine the distribution of the $\chi^2_{\nu,i}$ statistic throughout the PV plane in one of the comparisons. Due to its completeness, we use the example of the CfA ^{12}CO $J = 1 \rightarrow 0$ observations (see Figure VI–1). We see here that most of the molecular ring structure is replicated with $\chi^2_{\nu,i} < 10^2$, but there are some clear areas where there is no signal and $\chi^2_{\nu,i} \gg 10^2$. In fact it is these pixels in the plot where $\chi^2_{\nu,i} > 10^{300}$, which is clearly not fitting the observations. Since these pixels will dominate Equation VI–5 and severely affect the model fit, we filter out pixels with $\chi^2_{\nu,i} > 10^{10}$. While this

limit is arbitrary, it should be larger than the statistic for the meaningful pixels which, according to Figure VI–1, should reach a maximum of $\sim 10^2$.

$$\Delta\,\chi^2_{\rm min} = \left(\frac{\chi^2 - \chi^2_{\rm min}}{\chi^2_{\rm min}}\right) \times 100\% \ . \tag{VI–7}$$

Due to the very large values of χ^2 in our parameter space and the fact that we compare surveys of different sample sizes, we define an alternative metric, $\Delta\chi^2_{\rm min}$, for determining how well a model constrains a parameter in a given parameter space. This method of depicting the χ^2 statistic will be used throughout this chapter, since it more effectively shows how much the statistic varies in the parameter space even when the minimum χ^2 is very large or different amounts of data were used to calculate χ^2, and it will thus be crucial to determine the final constraints on our parameter space. As a result, we refer to $\Delta\chi^2_{\rm min}$ as the test statistic for the remainder of this section[29].

We have three species to consider for our comparison: CO, $^{12}C^+$, and dust[30]. As seen in §IV-5, we also have six model grids exploring the parameter space. The presentation of our results will begin with two grids (`voxel-dispersion` and `ensemble-FUV`) that are important to address first before the discussion of the rest.

[29] Typically in model selection, the reduced-χ^2 is used as the test statistic to account for the number of degrees of freedom in the χ^2 test. While this would account for both the number of parameters varied for each grid as well as the number of independent comparisons with the observations, in our case this is still insufficient to adequately compare different surveys/transition lines. That is why we define this alternative method to constrain our parameters.

[30] Technically we can also use ^{12}C to constrain our models, but the line we include is blended with CO $J = 7 \rightarrow 6$.

3 Voxel velocity dispersion

One question that arises from the construction of the galactic models is how can one compare our models to observations when there is no molecular cloud scale to approximate the velocity dispersion? From §II-3.1, the ensemble dispersion in the model is resulting solely from the variation in the relative rotational velocity inherent in each voxel, but the intrinsic ensemble dispersion in each voxel should not be a relative quantity. In order to judge how much this affects our results, we set the minimal ensemble dispersion σ_{ens} in each voxel by using the Larson relation (Larson 1981),

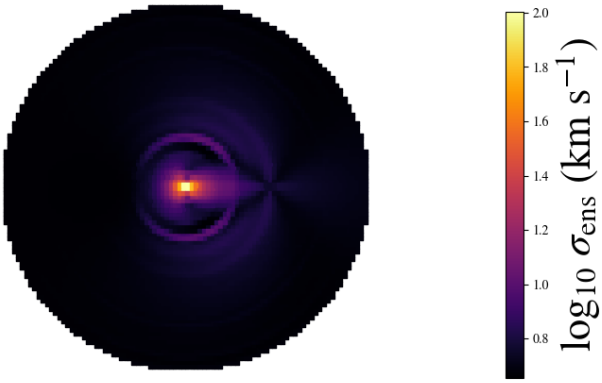


Figure VI–2: The ensemble dispersion just from the uncertainty in calculating the relative velocity according to §II-3.1, where the minimum is set by the clump velocity dispersion σ_{cl}.

$$\sigma_{\mathrm{mc}} \approx 1.1 \left(\frac{L_{\mathrm{mc}}}{\mathrm{pc}} \right)^{0.38} \mathrm{km\,s^{-1}} \,, \tag{VI–8}$$

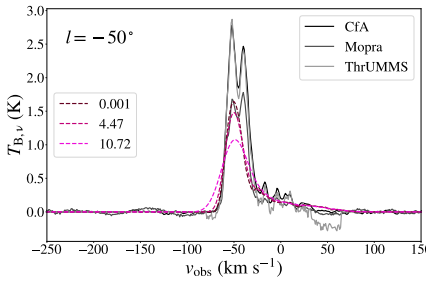
where L_{mc} is the length-scale of the molecular cloud for which the dispersion is being calculated. In our case, since we do not explicitly model molecular clouds, we assume the length-scale is some fraction $f_{\mathrm{mc}} < 1$ of the voxel size:

$$L_{\mathrm{mc}} = f_{\mathrm{mc}}\, \ell_{\mathrm{vox}} \,. \tag{VI–9}$$

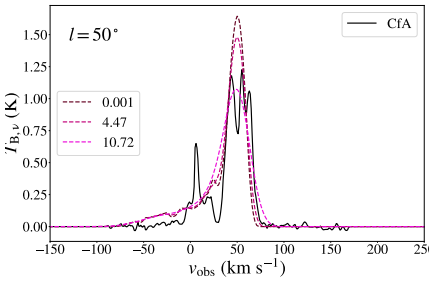
We have computed the `voxel-dispersion` model grid in order to constrain the ideal value of the velocity dispersion σ_{mc}, which we can use to constrain the typical molecular cloud length scale L_{mc} by Equation VI–8 in our models. Since the comparisons to the *COBE*-FIRAS CO lines and the dust continuum are not affected by changes in the velocity dispersion and do not exhibit much variation in the test statistic χ^2_{min}, we do not show these comparisons.

The ensemble dispersion should rather affect the spectra, so it is interesting to compare some of the synthetic spectra with the spectroscopic observations to qualitatively compare the fit (see Figure VI–3 to Figure VI–5). There is also the issue that there is far too much emission from the ^{13}CO transitions except at $l = -50°$. We can plainly see in the $l = 0$ subfigures in these figures that the velocity dispersion and absorption in the synthetic spectrum is not at all matching the observed spectrum. The spectra from the models with lower velocity dispersion seem to have an approximately Lorentzian profile. While the models with increased velocity dispersion have a broader Gaussian shape, the line profile is nowhere near

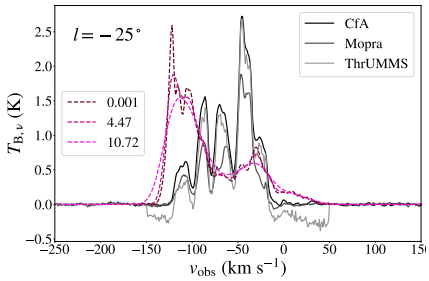
as broad as the observed spectrum in CO $J=1\rightarrow0$. This is rather important, since we will need to determine from Chapter VI whether this is due to a poor selection of parameters, or if it is a defect of how we have chosen to create the model.



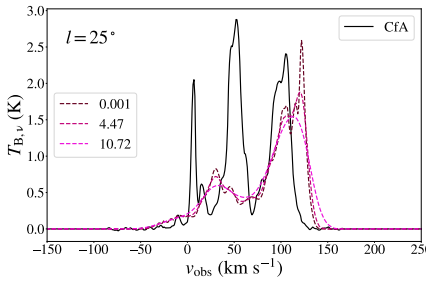
(a) $l=-50°$



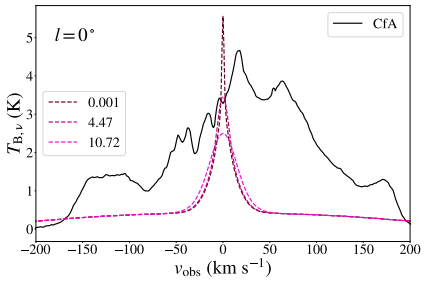
(b) $l=50°$



(c) $l=-25°$



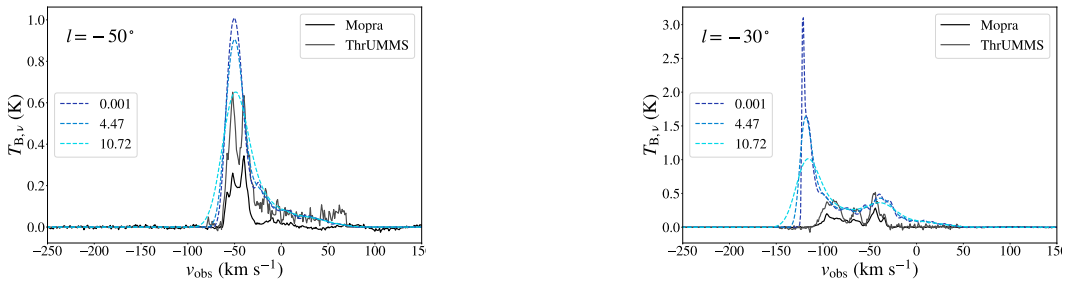
(d) $l=25°$



(e) $l=0°$

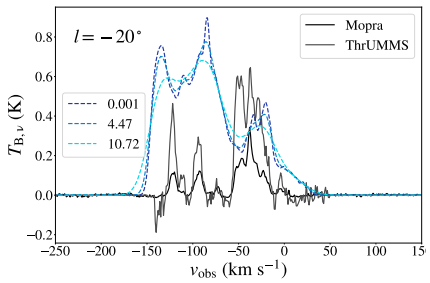
Figure VI–3: The ^{12}CO $J=1\rightarrow0$ spectra at $l\in[-50,-25,0,25,50]°$ from the `voxel-dispersion` model grid. The observed spectra are depicted as shades of solid black lines, while the synthetic spectra are depicted as shades of dashed red lines.

Subfigures **(a)** and **(c)**, meanwhile, depict a much closer fit for the three figures. While it appears that the spectrum with the lower velocity dispersion fits better to the shape of the observed spectrum (away from the bulk emission at $\pm50\,\mathrm{km\,s^{-1}}$), the intensity of the synthetic spectrum is much too high at the bulk emission. All of the synthetic spectra roughly follow a Gaussian profile, but the observed spectra are much more complicated. This difference is addressed in Chapter VII.

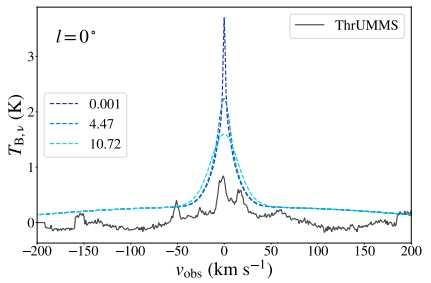


(a) $l = -50°$

(b) $l = -30°$

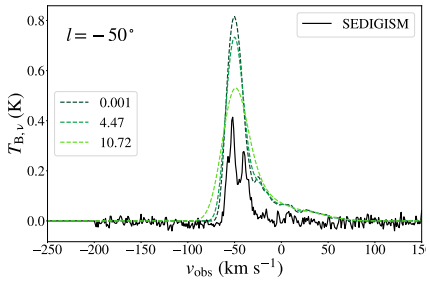


(c) $l = -20°$

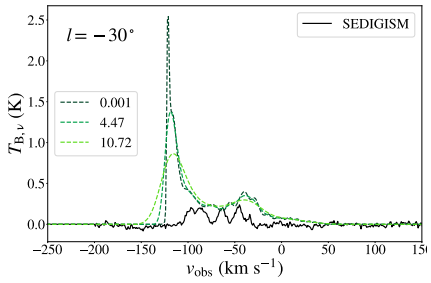


(d) $l = 0°$

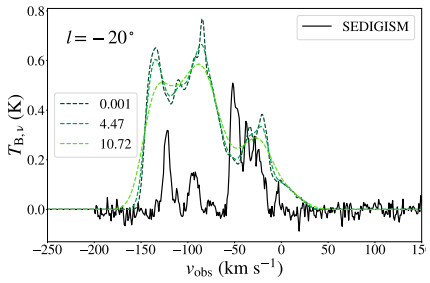
Figure VI–4: The ^{13}CO $J = 1 \rightarrow 0$ spectra at $l \in [-50, -30, -20, 0]°$ from the `voxel-dispersion` model grid. The observed spectra are depicted as shades of solid black lines, while the synthetic spectra are depicted as shades of dashed blue lines.



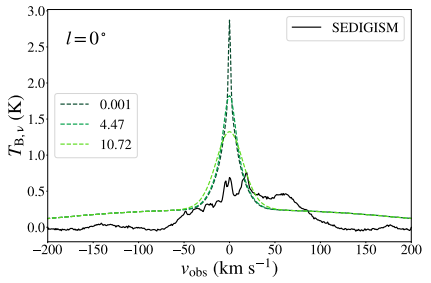
(a) $l = -50°$



(b) $l = -30°$



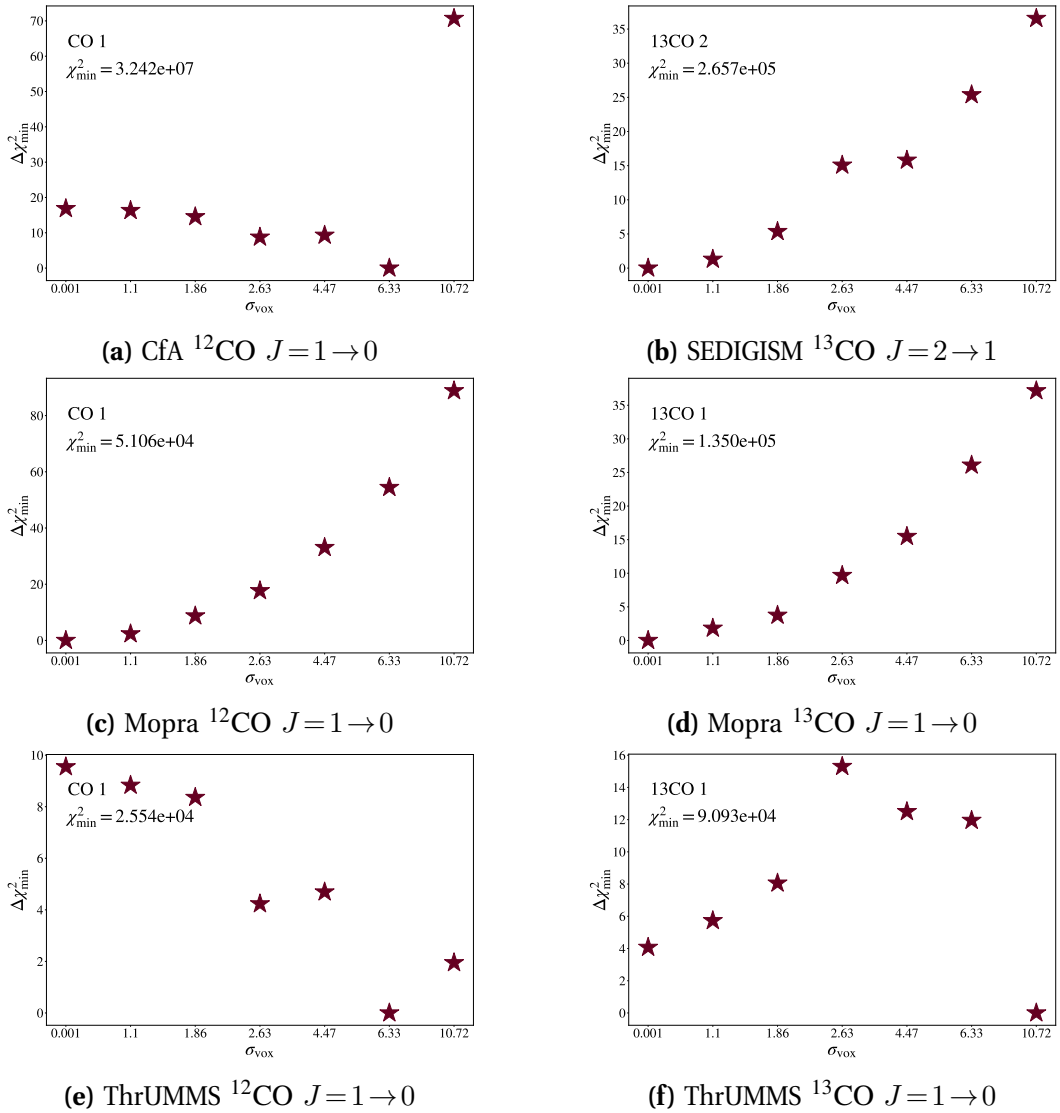
(c) $l = -20°$



(d) $l = 0°$

Figure VI–5: The ^{13}CO $J = 2 \rightarrow 1$ spectra at $l \in [-50, -30, -20, 0]°$ from the `voxel-dispersion` model grid. The observed spectrum is depicted as a solid black line, while the synthetic spectra are depicted as shades of dashed green lines.

In order to provide a quantitative comparison, we also examine the test statistic according to the procedure in §VI-2 for the spectroscopic surveys (see Figure VI–6). For these comparisons χ^2 varies between the different surveys due to differences in the fit and the number of individual points in the position-velocity plane used in the comparison, but about half of the subplots indicate that $\sigma_{\mathrm{ens}} = 0.001\,\mathrm{km\,s^{-1}}$ (meaning we don't consider a typical molecular cloud size). The other half of the surveys indicate $0.25 < f_{\mathrm{mc}} < 1$. The constraints are stronger than what we can interpret from the *COBE*-FIRAS comparisons, though we are still unable to identify a trend in the parameter space.



(a) CfA ^{12}CO $J = 1 \rightarrow 0$

(b) SEDIGISM ^{13}CO $J = 2 \rightarrow 1$

(c) Mopra ^{12}CO $J = 1 \rightarrow 0$

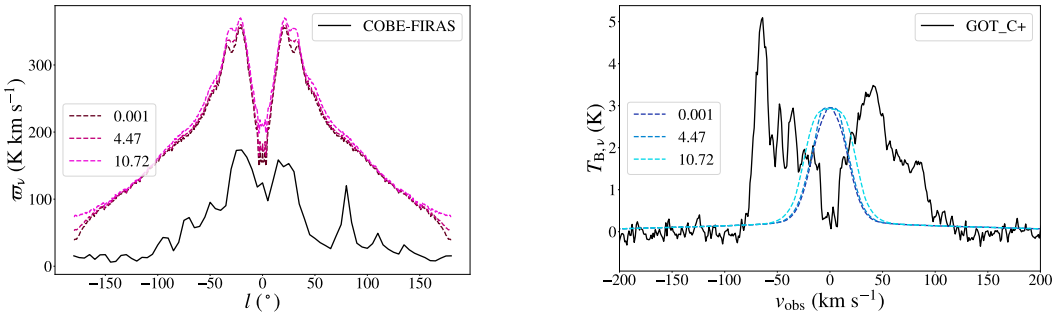
(d) Mopra ^{13}CO $J = 1 \rightarrow 0$

(e) ThrUMMS ^{12}CO $J = 1 \rightarrow 0$

(f) ThrUMMS ^{13}CO $J = 1 \rightarrow 0$

Figure VI–6: The comparison between the spectroscopic CO observations and the synthetic observations of the `voxel-dispersion` model grid for the transitions ^{12}CO $J = 1 \rightarrow 0$, ^{13}CO $J = 1 \rightarrow 0$, and ^{13}CO $J = 2 \rightarrow 1$.

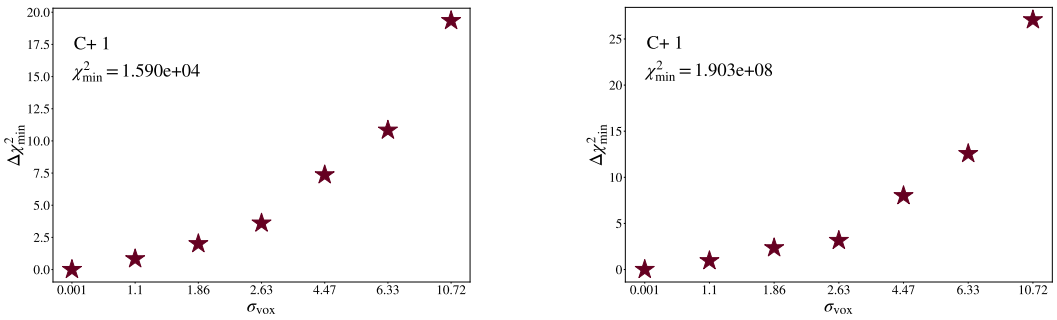
We can plot as well the integrated-intensity profile and the spectrum towards the Galactic centre for models spanning the parameter space, which we plot along with the respective observations (see Figure VI–7). We can plainly see in subfigure



(a) *COBE*-FIRAS **(b)** GOT C+

Figure VI–7: The [^{12}C II] 158 μm intensity profile from *COBE*-FIRAS and spectrum at $l = 0$ from GOT C+ plotted with the best-fit models from the `voxel-dispersion` model grid. The observed profiles are depicted as black solid lines, while the synthetic intensities are depicted as shades of red lines for the profile subplot and shades of dashed blue lines for the spectrum subplot.

(a) that although we correctly get more emission from the Galactic ring rather than the Galactic centre, overall there is too much [^{12}C II] emission in the synthetic profiles compared to *COBE*-FIRAS. The decrease in emission towards the Galactic centre is more evident in the model with higher velocity dispersion, though the fit is not close enough to conclude whether this is significant. in subfigure **(b)**, the synthetic spectra are unable to replicate the self-absorption dip at $v_{obs} = 0\,\mathrm{km\,s^{-1}}$. There are optical depth effects, toward the Galactic centre, evidenced by the small plateau that is visible, but overall there is not enough velocity dispersion in order to make these models fit.

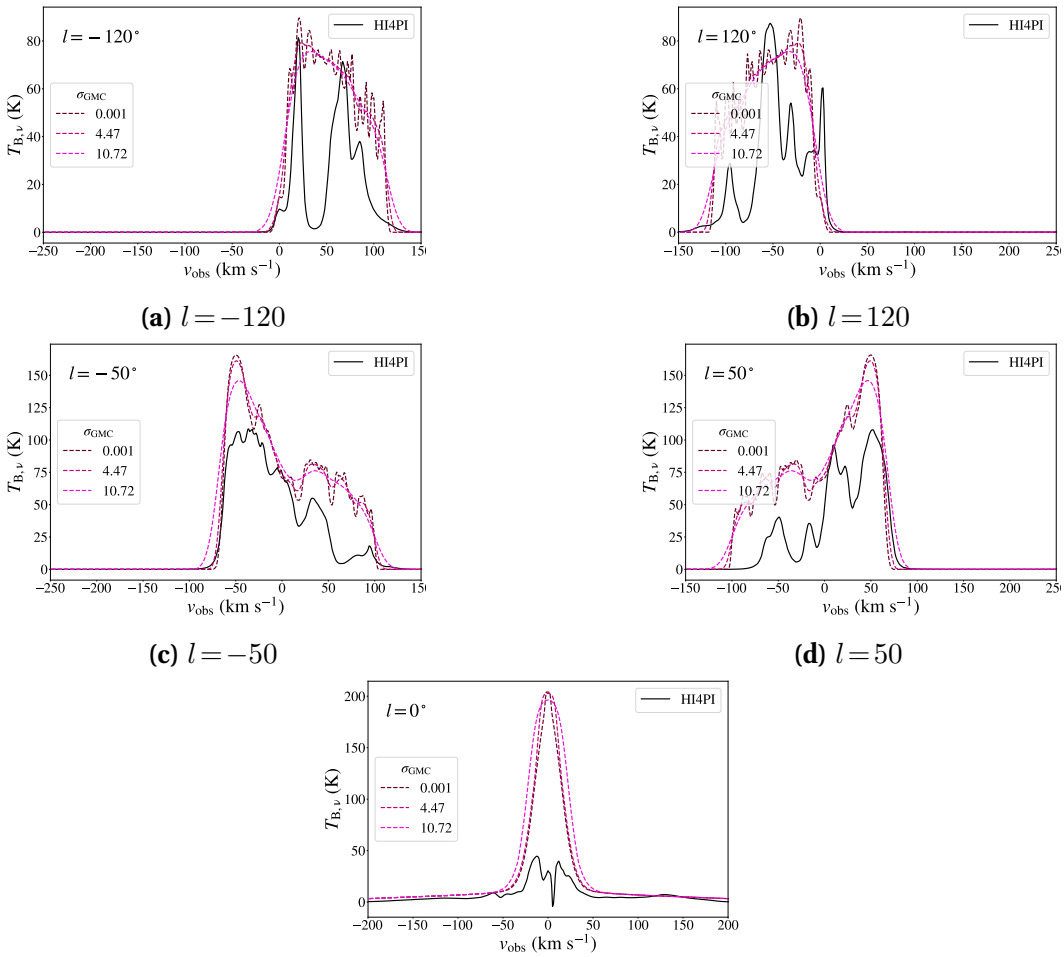


(a) *COBE*-FIRAS **(b)** GOT C+

Figure VI–8: The comparison between the observations to the `voxel-dispersion kosmatau3d` grid for the transition [^{12}C II] 158 μm.

When comparing the synthetic [^{12}C II] 158 μm transition to the observations of the *COBE*-FIRAS and GOT C+ surveys, the variation in the test statistic is $\sim 20\%$ with respect to $\chi^2_{min} = 1.590\times10^4, 1.903\times10^8$ and we have stronger constraints for our parameters (see Figure VI–8). There is a clear constraint of $\sigma_{ens} \lesssim 1.1\,\mathrm{km\,s^{-1}}$.

The H I 21 cm transition has been post-processed (see §I-1.2.2), so the galactic model can be further constrained. Some of the spectra are compared qualitatively in Figure VI–9 for models spanning the parameter space. These synthetic spectra



(a) $l = -120$ **(b)** $l = 120$

(c) $l = -50$ **(d)** $l = 50$

(e) $l = 0$

Figure VI–9: The H I 21 cm spin transition spectra at $l \in [-120, -50, 0, 50, 120]^\circ$ from the `voxel-dispersion` model grid. The observed spectra are depicted as shades of solid black lines, while the synthetic spectra are depicted as shades of dashed red lines.

show a bit of variation in how well they match the observations, with lower values of velocity dispersion exhibiting more spurious features in the spectra. Generally there is too much emission predicted by the models. The synthetic spectra in subfigure **(e)** do not closely match the observed spectrum partially because there is a large HISA feature at $\sim 5\,\mathrm{km\,s^{-1}}$ making the observation difficult, and partially due to the excess of emission from the galactic models (similar to what is seen in CO). At least the position in velocity-space is correct, which can be seen by comparing the emission wings.

The $\Delta\chi^2_{\mathrm{min}}$ test statistic features a large amount of variation in the `voxel-dispersion` parameter space (see Figure VI–10), with $\chi^2_{\mathrm{min}} = 2.955 \times 10^8$, effectively resulting in a moderate constraint of $\sigma_{\mathrm{GMC}} \sim 10.72$.

The $\Delta\chi^2_{\mathrm{min}}$ results are combined for the various comparisons (except for the *Planck* comparison; see §VII-7 for the discussion of why it is neglected) and summarised in VI–1.

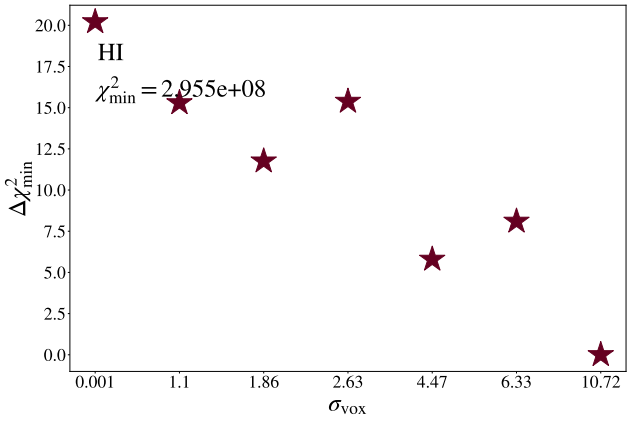


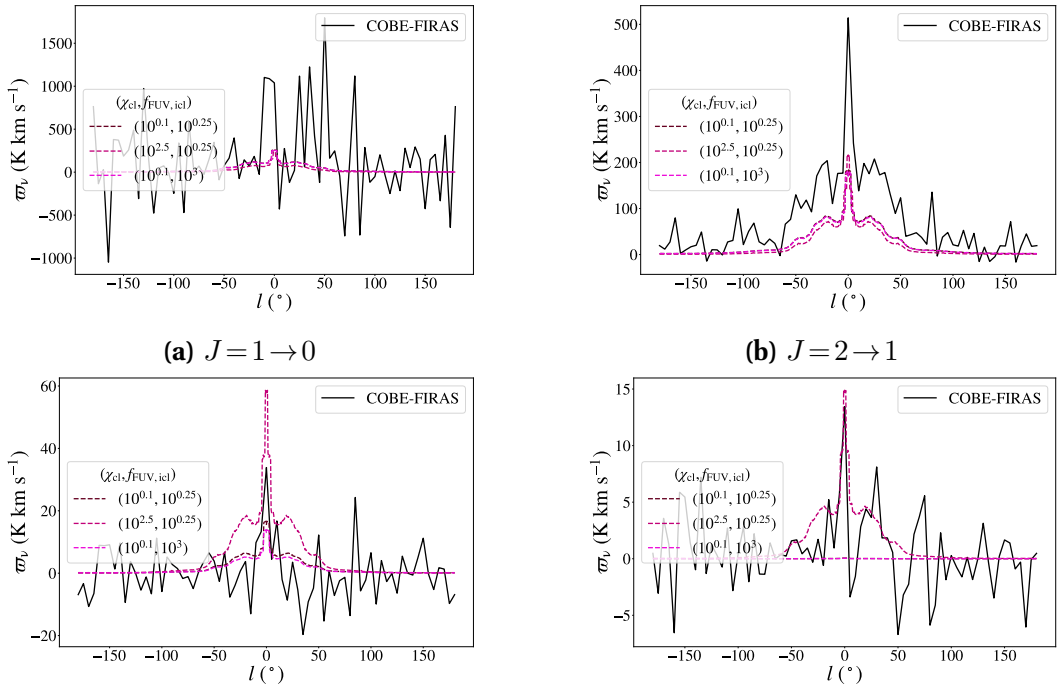
Figure VI–10: The comparison between the H I observations to the `voxel-dispersion kosmatau3d` grid for the transition H I 21 cm.

Table VI–1: The $\Delta\chi^2_{\rm min}$ results for the `voxel-dispersion kosmatau3d` model grid, which varies $\sigma_{\rm GMC}$ The minimum χ^2 is 5.286×10^8.

$\sigma_{\rm GMC}$	$\Delta\chi^2_{\rm min}$
0.001	**51.4**
1.1	54.1
1.86	65.6
2.63	93.3
4.47	112.8
6.33	149.8
10.72	282.5

4 Clumpy far-UV

Before we show the full analysis we must constrain the far-UV radiation intensity for both the dense, clumpy medium and the diffuse interclump medium. The work by Cubick et al. (2008), using one ensemble of dense clumps, had constrained the required far-UV intensity to $\chi_{\rm cl} \approx 10^{1.8}\,\chi_{\rm D}$ when fitting their clumpy model to the *COBE*-FIRAS reduction of Fixsen et al. (1999). We constrain this using the `ensemble-FUV` grid that was computed varying the far-UV radiation intensity $\chi_{\rm cl}$ for the clumpy ensemble (constant throughout the galaxy) and the scaling factor $f_{\rm FUV}$ for the galactic far-UV profile in Figure IV–5 (which is then used for the interclump ensemble).

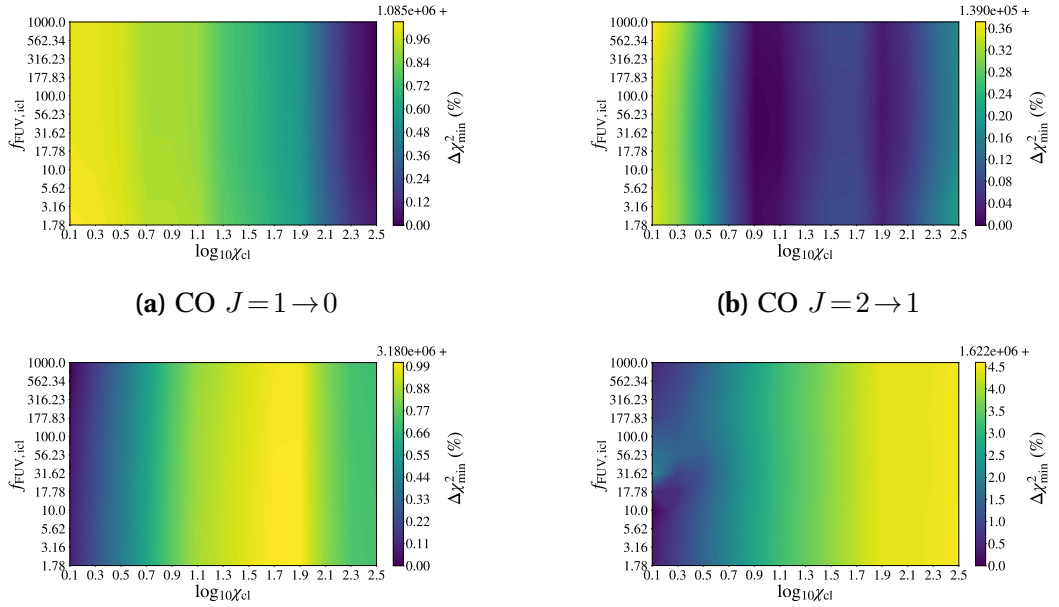


(a) $J=1\rightarrow0$ **(b)** $J=2\rightarrow1$

(c) $J=6\rightarrow5$ **(d)** $J=8\rightarrow7$

Figure VI–11: The ^{12}CO $J=1\rightarrow0$, $J=2\rightarrow1$, $J=6\rightarrow5$, and $J=8\rightarrow7$ integrated-intensity from *COBE*-FIRAS and the `ensemble-FUV` model grid. The observational profile is depicted as a solid black line, while the synthetic profile is depicted as shades of dashed red lines.

While the CO might exist in either ensemble (see abundance profiles in §III-7), we do not expect the rotational transitions to trace very well the diffuse interclump ensemble (its fractional abundance is only $X_{CO,icl}\sim10^{-9}$). A qualitative verification of how the galactic models match the *COBE*-FIRAS CO observations can be obtained by examining the transitions $J=1\rightarrow0$, $J=2\rightarrow1$, $J=6\rightarrow5$, and $J=8\rightarrow7$ for the parameter space $(\log_{10}\chi_{cl}, f_{FUV,icl})\in[(0.1, 10^{0.25}), (2.5, 10^{0.25}), (0.1, 10^{3})]$ in Figure VI–11. While it can be difficult to judge by eye from these plots which of these models is fitting better, subfigure **(b)** seems to indicate a slightly better fit for the model with $\chi_{cl}=10^{2.5}\,\chi_D$, while subfigure **(c)** shows a better fit for the model with $\chi_{cl}=10^{0.1}\,\chi_D$.



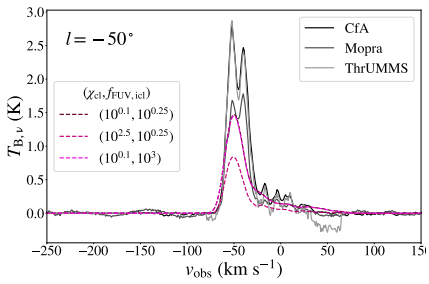
(a) CO $J=1\rightarrow0$ **(b)** CO $J=2\rightarrow1$

(c) CO $J=6\rightarrow5$ **(d)** CO $J=8\rightarrow7$

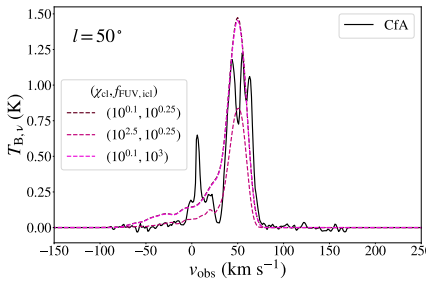
Figure VI–12: The comparison between the CO transitions observed by *COBE*-FIRAS and the `ensemble-FUV` model grid, neglecting extraneous values of χ^2.

By inspecting the χ^2 statistic in Figure VI–12, we can see that the test statistic indicates a better match for the lower transition lines when the clumpy medium is embedded in a higher radiation (as determined by Cubick et al. 2008), while the test statistic for the higher transitions indicate a better match to lower far-UV values. The variation of the test statistic in the parameter space is $\max(\Delta\chi^2_{\mathrm{min}}) < 1\%$ for the all transitions except the $J=8\rightarrow7$ parameter space. Since $\Delta\chi^2_{\mathrm{min}}$ is used as a measure of how well a given transition can constrain the parameter space, Figure VI–12 indicates that there is a better fit to low χ_{cl}. The lower-limit of the far-UV parameter for each clump is $1\,\chi_{\mathrm{D}}$ (see §III-3.1), forcing the far-UV radiation used throughout the galactic disk to be $\gtrsim 1\,\chi_{\mathrm{D}}$. This indicates that the un-scaled far-UV profile discussed in §IV-1.3 may be used for the clumpy ensemble.

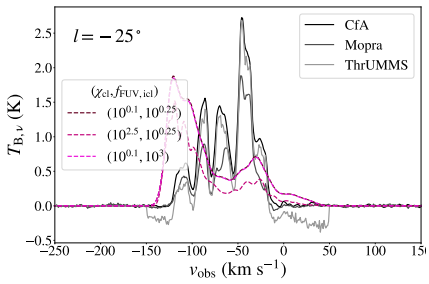
We can qualitatively examine spectra in the models spanning the parameter space at $l\in[-50,0,50]^\circ$ (see Figure VI–13 to Figure VI–15). These plots show that there is very little difference between the low and high far-UV radiation models. The largest difference is seen in subfigures **(a)** and **(c)** of Figure VI–14, where the models with $\log_{10}\chi_{\mathrm{cl}}=0.1$ appears to fit better to the spectra of the observations. The spectra at $l=0$, however, is still fitting poorly to the observations with too little velocity dispersion.



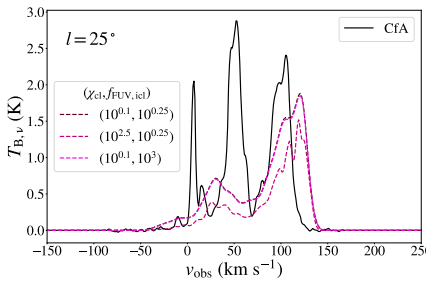
(a) $l = -50°$



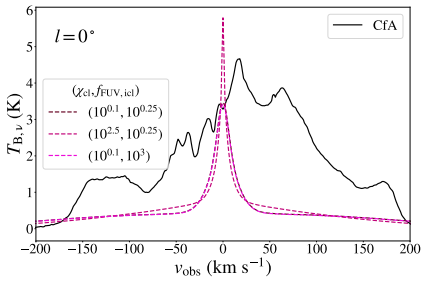
(b) $l = 50°$



(c) $l = -25°$

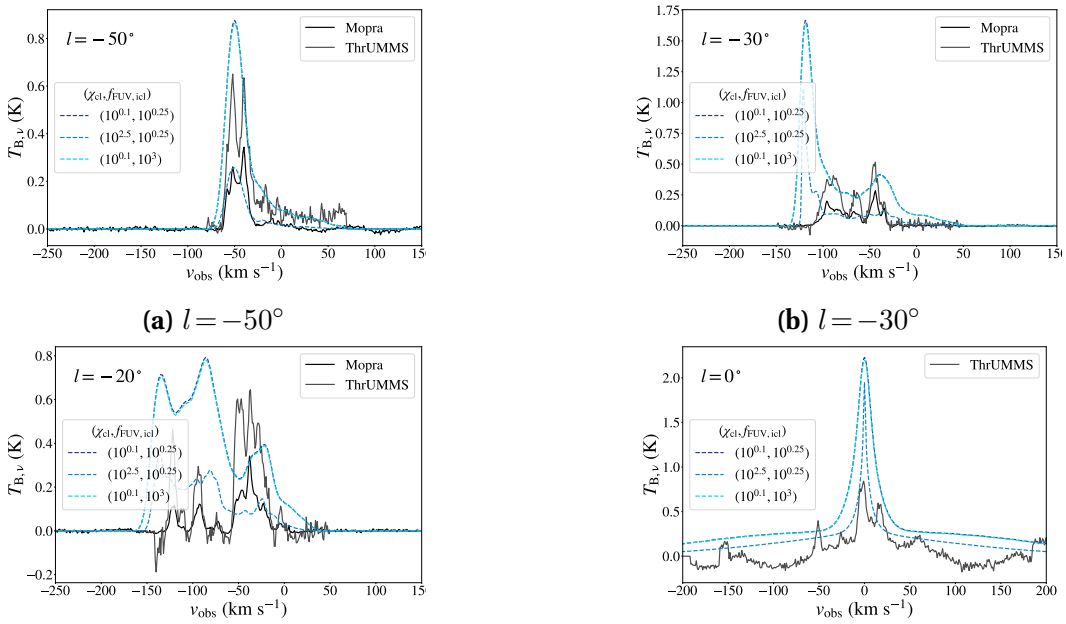


(d) $l = 25°$



(e) $l = 0°$

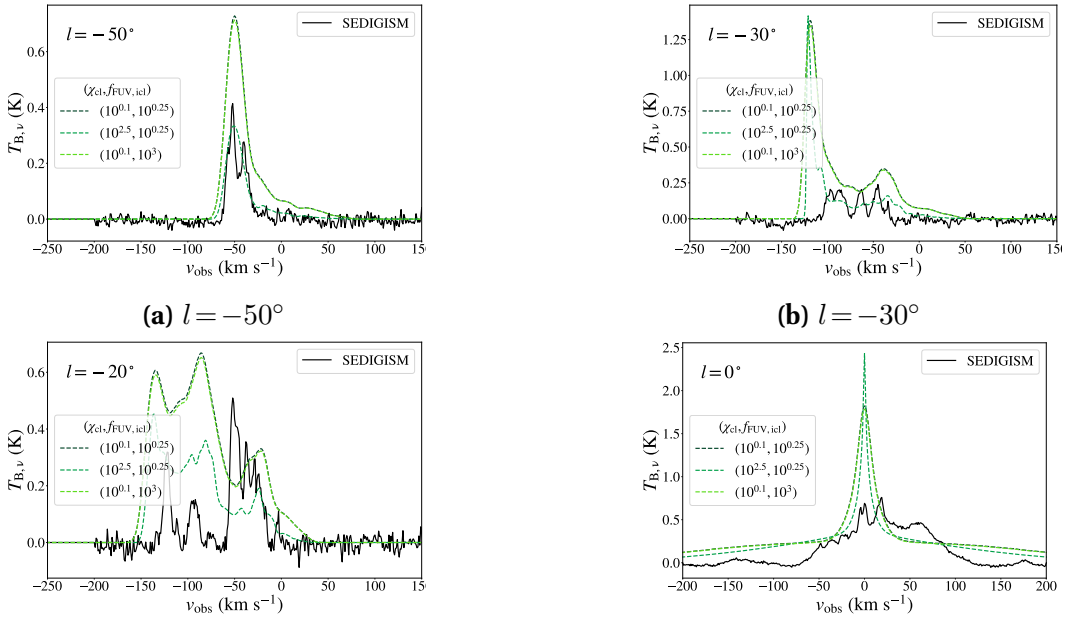
Figure VI–13: The ^{12}CO $J = 1 \rightarrow 0$ spectra at $l \in [-50, -25, 0, 25, 50]°$ from the `ensemble-FUV` model grid. The observed spectra are depicted as shades of solid black lines, while the synthetic spectra are depicted as shades of dashed red lines.



(a) $l = -50°$ **(b)** $l = -30°$

(c) $l = -20°$ **(d)** $l = 0°$

Figure VI–14: The ^{13}CO $J = 1 \rightarrow 0$ spectra at $l \in [-50, -30, -20, 0]°$ from the `ensemble-FUV` model grid. The observed spectra are depicted as shades of solid black lines, while the synthetic spectra are depicted as shades of dashed blue lines.

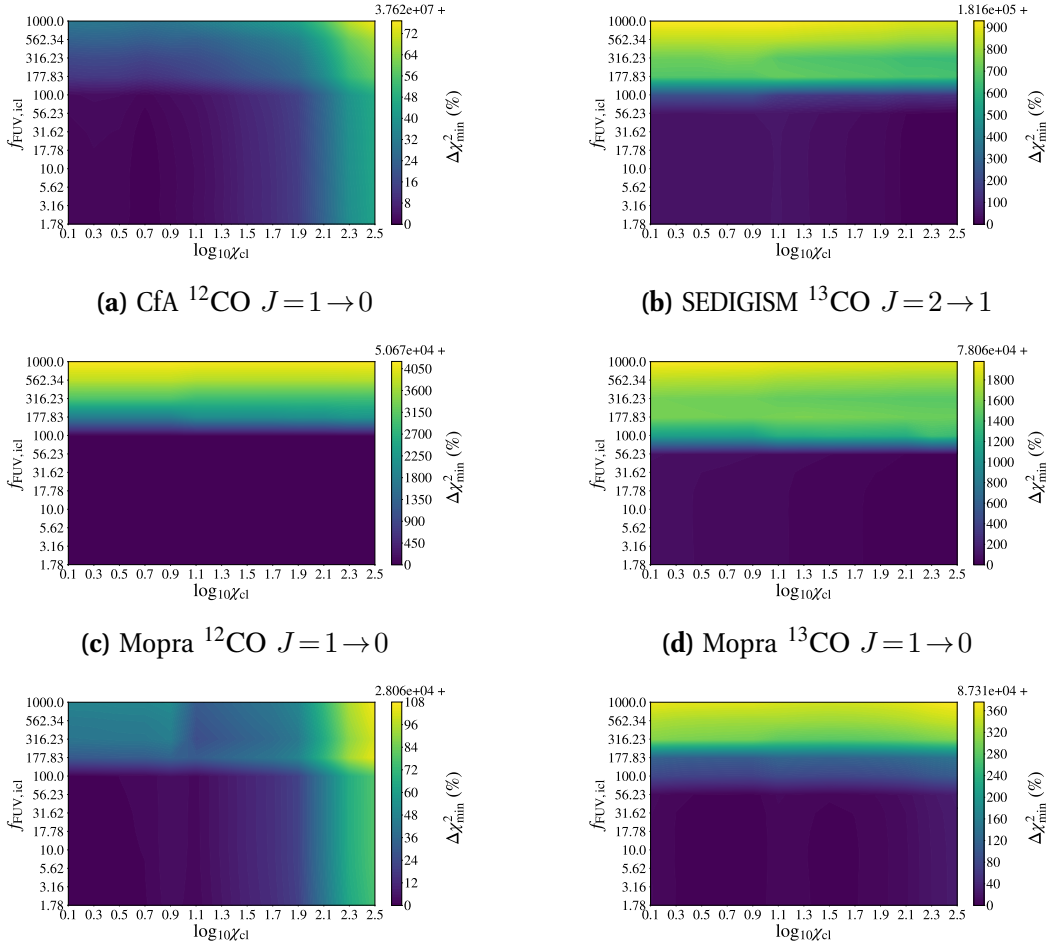


(a) $l = -50°$ **(b)** $l = -30°$

(c) $l = -20°$ **(d)** $l = 0°$

Figure VI–15: The ^{13}CO $J = 2 \rightarrow 1$ spectra at $l \in [-50, -30, -20, 0]°$ from the `ensemble-FUV` model grid. The observed spectrum is depicted as a solid black line, while the synthetic spectra are depicted as shades of dashed green lines.

Contrary to the fit of the low-J transitions in Figure VI–12, the spectroscopic comparison in Figure VI–16 indicates a different trend in the parameter space. Not only do all of the comparison to the spectroscopic surveys show better agreement to low $\chi_{\rm cl}$ and low $\chi_{\rm icl}$, but there is strong agreement since the variance of χ^2 in the parameter space is rather large ($\max(\Delta\chi^2_{\rm min}) > 100\%$). Although we expect there is a contribution of both ensembles to the lower ^{12}CO transitions, it is useful to know that these comparisons are in agreement with our previous conclusion.



(a) CfA ^{12}CO $J=1\rightarrow0$

(b) SEDIGISM ^{13}CO $J=2\rightarrow1$

(c) Mopra ^{12}CO $J=1\rightarrow0$

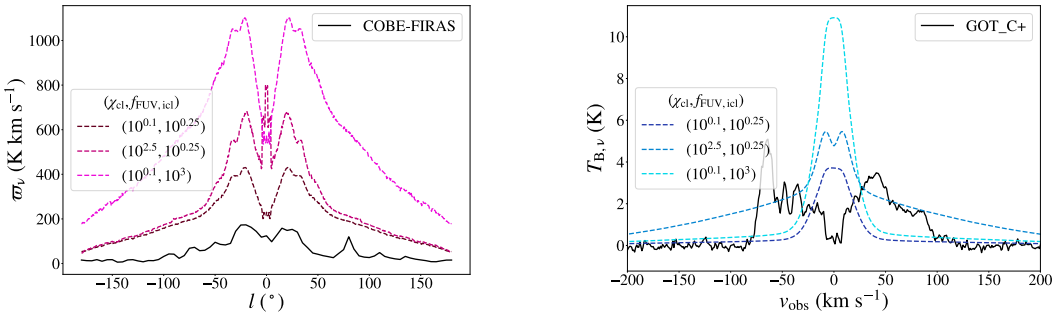
(d) Mopra ^{13}CO $J=1\rightarrow0$

(e) ThrUMMS ^{12}CO $J=1\rightarrow0$

(f) ThrUMMS ^{13}CO $J=1\rightarrow0$

Figure VI–16: The comparison between the spectroscopic CO observations and the synthetic observations of the `ensemble-FUV` model grid for the transitions ^{12}CO $J=1\rightarrow0$, ^{13}CO $J=1\rightarrow0$, and ^{13}CO $J=2\rightarrow1$.

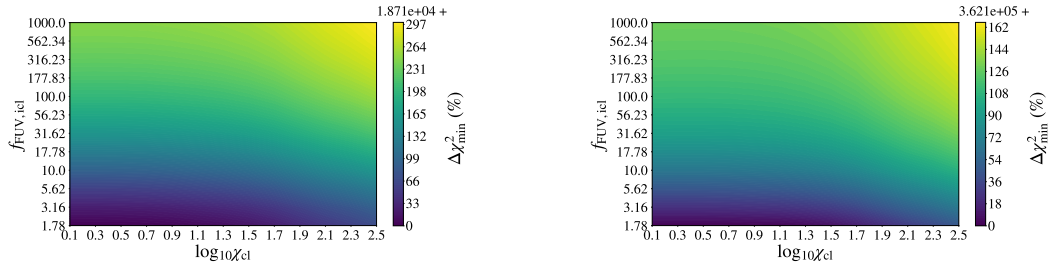
We can plot as well the integrated-intensity profile and the spectrum towards the Galactic centre for models spanning the parameter space, which we plot along with the respective observations (see Figure VI–17). We can plainly see in subfigure **(a)** that although we correctly get more emission from the Galactic ring rather than the Galactic centre, overall there is too much [^{12}C II] emission in both synthetic profiles compared to *COBE*-FIRAS. The decrease in emission towards the Galactic centre is more evident in the model with $\log_{10}\chi_{\rm cl} = 0.1$, though the fit is not close enough to conclude whether this is significant. in subfigure **(b)**, the



(a) *COBE*-FIRAS **(b)** GOT C+

Figure VI–17: The [^{12}C II] 158 μm intensity profile from *COBE*-FIRAS and spectrum at $l = 0$ from GOT C+ plotted with the best-fit models from the `ensemble-FUV` model grid. The observed profiles are depicted as black solid lines, while the synthetic intensities are depicted as shades of red lines for the profile subplot and shades of dashed blue lines for the spectrum subplot.

synthetic spectra are unable to replicate the self-absorption dip at $v_{obs} = 0\,\mathrm{km\,s^{-1}}$. There is self-absorption, toward the Galactic centre, evidenced by the small plateau that is visible, but overall there is not enough velocity dispersion in order to make these models fit. Away from the bulk emission, the model with parameters $(\log_{10}\chi_{cl}, f_{FUV,icl}) = (0.1, 10^{0.25})$ appears to give a closer fit to the GOT C+ observations.

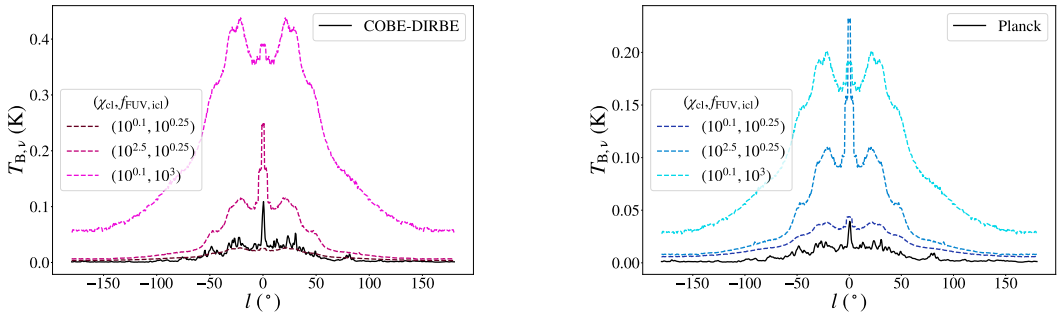


(a) *COBE*-FIRAS **(b)** GOT C+

Figure VI–18: The comparison between the observations to the `ensemble-FUV` `kosmatau3d` grid for the transition [^{12}C II] 158 μm.

When comparing the synthetic [^{12}C II] 158 μm transition to the observations of the *COBE*-FIRAS and GOT C+ surveys, the variation in the test statistic is $> 100\%$ with respect to $\chi^2_{min} = 1.871 \times 10^4, 3.621 \times 10^5$ and we have stronger constraints for our parameters (see Figure VI–18). There is a clear constraint of $f_{FUV} < 10$, and a slight constraint of $\chi_{cl} \lesssim 10^{2.1}\,\chi_D$. This supports our prediction that χ_{cl} is not well-constrained in this comparison.

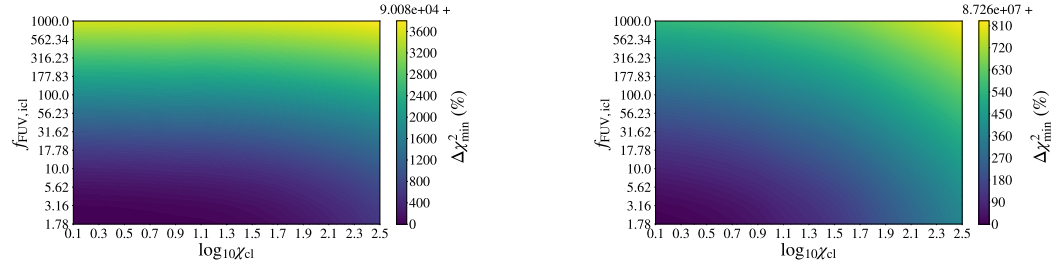
The dust in the PDRs is an important tracer of the far-UV since it should remain in approximate thermal equilibrium (see the discussion in §V-3). We explore the parameter space in Figure VI–19. While the model with $f_{FUV} = 10^{0.25}$ has a synthetic continuum that matches slightly better to the observed continuum at 240 μm rather than the other models, and it definitely fits better to the observed continuum at



(a) *COBE*-DIRBE 240 μm **(b)** *Planck* 550 μm

Figure VI–19: The comparison between the dust continuum observations to the `ensemble-FUV kosmatau3d` grid at the specified wavelengths.

550 μm (although none of the models have a continuum that matches well to the 550 μm profile). The closeness of the comparisons in subfigure **(a)** indicate that the dust intensity is more susceptible to changes in the conditions of the dense, clumpy ensemble rather than the diffuse interclump ensemble. None of the models can correctly replicate the intensity peak of the profile at $l = 0$. From Figure VI–20 there is no clear fitting of the dense, clumpy medium to higher far-UV radiation intensities. Contrarily, the comparison to the *Planck* 550 μm observations fits better to lower intensities. From this comparison we can also see indications that $\chi_{cl} \approx \chi_{icl}$ in subfigure **(b)**[31]. The maximum variation in both comparisons in $\gg 100\%$, confirming that the dust continuum is a good tracer of the far-UV radiation for PDRs. Nonetheless, this is a strong indication that the dense, clumpy medium and the diffuse interstellar medium in our models can use the same far-UV radiation intensity.

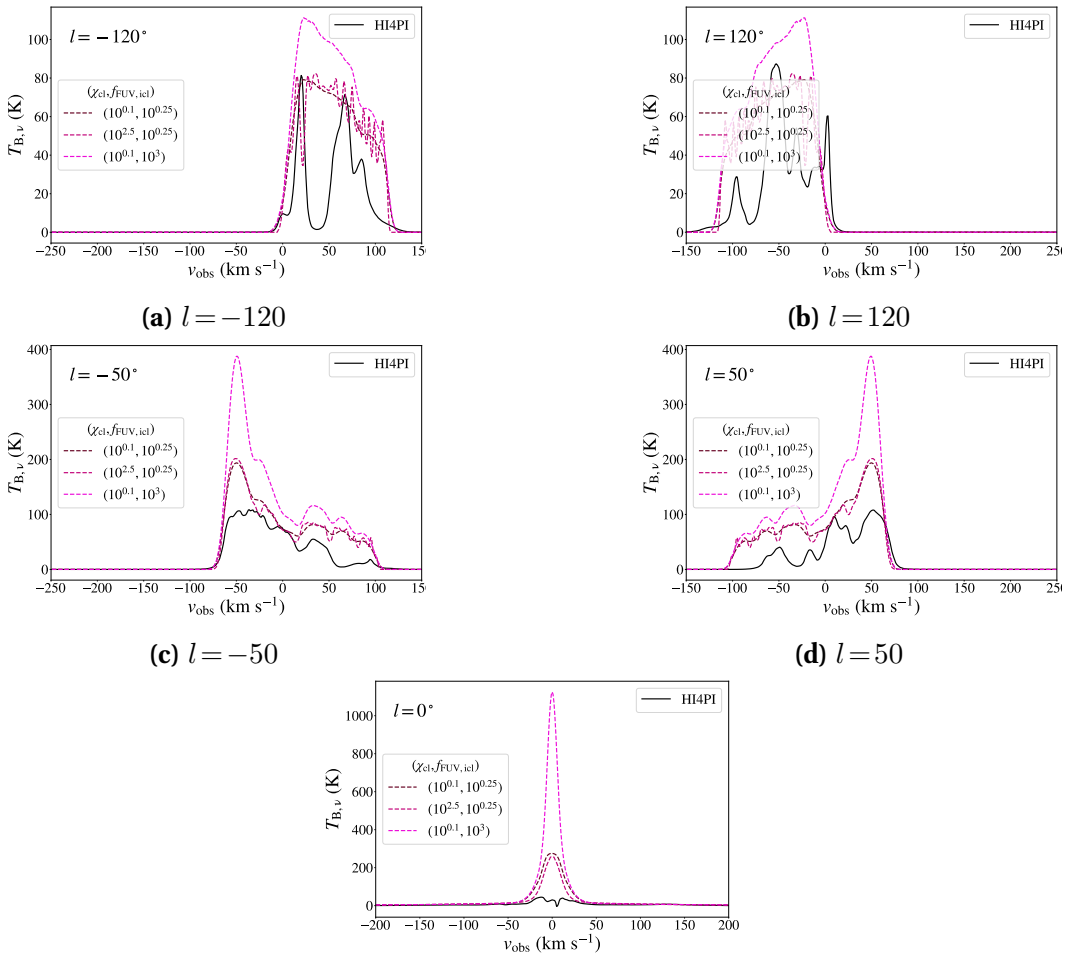


(a) 240 μm **(b)** 550 μm

Figure VI–20: The comparison between the *COBE*-DIRBE and *Planck* observations to the `ensemble-FUV kosmatau3d` grid for the wavelength 240 μm.

The H I 21 cm transition has been post-processed (see §I-1.2.2), so the galactic model can be further constrained. Some of the spectra are compared qualitatively in Figure VI–21 for the models spanning the parameter space. These synthetic spectra confirm our assumption of the H I emission being dominated by the in-

[31] This can be understood since the un-scaled far-UV distribution (see Figure IV–7) is $\gtrsim 1\,\chi_D$ throughout the disk, and thus having the constraint $\chi_{cl} \gtrsim 1$ and $f_{FUV,icl} \gtrsim 1$ implies $\chi_{cl} \approx \chi_{icl}$.



(a) $l = -120$ **(b)** $l = 120$

(c) $l = -50$ **(d)** $l = 50$

(e) $l = 0$

Figure VI–21: The H I 21 cm spin transition spectra at $l \in [-120, -50, 0, 50, 120]°$ from the `ensemble-FUV` model grid. The observed spectra are depicted as shades of solid black lines, while the synthetic spectra are depicted as shades of dashed red lines.

terclump ensemble since they mostly overlap. There is just a bit more structure visible in the models with higher χ_{cl}. The bulk emission in subfigures **(c)** and **(d)** seem to be approximately correctly orientated in v_{obs}, but there is not enough self-absorption at $v_{obs} \approx \mp 50\,\mathrm{km\,s^{-1}}$, respectively. Similarly, there is too much emission at $v_{obs} \approx \pm 25\,\mathrm{km\,s^{-1}}$ to $v_{obs} \approx \pm 100\,\mathrm{km\,s^{-1}}$. This indicates that there is poor optimisation in the mass distribution in the interclump medium. The synthetic spectra in subfigure **(e)** do not closely match the observed spectrum partially because there is a large synchrotron source at $5\,\mathrm{km\,s^{-1}}$ making the observation difficult, and partially due to the excess of emission from the galactic models (similar to what is seen in CO). At least the position in velocity-space is correct, which can be seen by comparing the emission wings.

The $\Delta\chi^2_{min}$ test statistic features a large amount of variation in the `ensemble-FUV` parameter space (see Figure VI–22), with $\chi^2_{min} = 4.485 \times 10^8$, effectively resulting in a strong constraint of $f_{icl} \gtrsim 1$ while nearly no constraint on the χ_{cl} parameter.

This agrees with our assertion that the H^0 is mainly in the interclump ensemble, and thus the H I intensity is primarily affected by parameters affecting the interclump ensemble in our models.

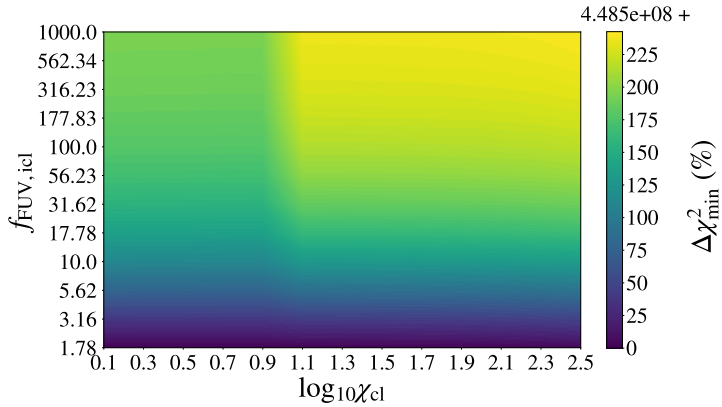


Figure VI–22: The comparison between the H I observations to the `ensemble-FUV kosmatau3d` grid for the transition H I 21 cm.

The $\Delta\chi^2_{\mathrm{min}}$ results are combined for the various comparisons (except for the *Planck* comparison; see §VII-7 for the discussion of why it is neglected) and summarised in VI–2.

Table VI–2: The $\Delta\chi^2_{\mathrm{min}}$ results for the `ensemble-FUV kosmatau3d` model grid, which varies χ_{cl} and f_{FUV} The minimum χ^2 is 4.968×10^8.

χ_{cl}	f_{FUV}	$\Delta\chi^2_{\mathrm{min}}$
$10^{0.1}$	$10^{0.25}$	210.8
$10^{0.1}$	$10^{0.5}$	329.5
$10^{0.1}$	$10^{0.75}$	551.3
$10^{0.1}$	10^{1}	858.4
$10^{0.1}$	$10^{1.25}$	1218.5
$10^{0.1}$	$10^{1.5}$	1621.2
$10^{0.1}$	$10^{1.75}$	2067.4
$10^{0.1}$	10^{2}	4873.2
$10^{0.1}$	$10^{2.25}$	6802.9
$10^{0.1}$	$10^{2.5}$	7661.8
$10^{0.1}$	$10^{2.75}$	8689.8
$10^{0.1}$	10^{3}	9762.0
$10^{0.3}$	$10^{0.25}$	186.4
$10^{0.3}$	$10^{0.5}$	309.2
$10^{0.3}$	$10^{0.75}$	532.9
$10^{0.3}$	10^{1}	840.8

Continued on next page

Table VI–2: The $\Delta\chi^2_{\rm min}$ results for the `ensemble-FUV kosmatau3d` model grid, which varies $\chi_{\rm cl}$ and $f_{\rm FUV}$ The minimum χ^2 is 4.968×10^8.

$\boldsymbol{\chi_{\rm cl}}$	$\boldsymbol{f_{\rm FUV}}$	$\boldsymbol{\Delta\chi^2_{\rm min}}$
$10^{0.3}$	$10^{1.25}$	1200.0
$10^{0.3}$	$10^{1.5}$	1601.2
$10^{0.3}$	$10^{1.75}$	2048.0
$10^{0.3}$	10^{2}	4853.6
$10^{0.3}$	$10^{2.25}$	6781.5
$10^{0.3}$	$10^{2.5}$	7641.1
$10^{0.3}$	$10^{2.75}$	8668.2
$10^{0.3}$	10^{3}	9740.4
$10^{0.5}$	$10^{0.25}$	171.4
$10^{0.5}$	$10^{0.5}$	299.6
$10^{0.5}$	$10^{0.75}$	526.1
$10^{0.5}$	10^{1}	834.9
$10^{0.5}$	$10^{1.25}$	1194.0
$10^{0.5}$	$10^{1.5}$	1594.4
$10^{0.5}$	$10^{1.75}$	2040.2
$10^{0.5}$	10^{2}	4845.7
$10^{0.5}$	$10^{2.25}$	6771.4
$10^{0.5}$	$10^{2.5}$	7631.7
$10^{0.5}$	$10^{2.75}$	8657.5
$10^{0.5}$	10^{3}	9730.4
$10^{0.7}$	$10^{0.25}$	165.7
$10^{0.7}$	$10^{0.5}$	300.9
$10^{0.7}$	$10^{0.75}$	531.1
$10^{0.7}$	10^{1}	841.0
$10^{0.7}$	$10^{1.25}$	1199.8
$10^{0.7}$	$10^{1.5}$	1599.3
$10^{0.7}$	$10^{1.75}$	2044.1
$10^{0.7}$	10^{2}	4848.5
$10^{0.7}$	$10^{2.25}$	6770.1
$10^{0.7}$	$10^{2.5}$	7647.4
$10^{0.7}$	$10^{2.75}$	8655.3
$10^{0.7}$	10^{3}	9727.9
$\mathbf{10^{0.9}}$	$\mathbf{10^{0.25}}$	**165.0**
$10^{0.9}$	$10^{0.5}$	308.8
$10^{0.9}$	$10^{0.75}$	543.5
$10^{0.9}$	10^{1}	854.6
$10^{0.9}$	$10^{1.25}$	1212.8
$10^{0.9}$	$10^{1.5}$	1611.0
$10^{0.9}$	$10^{1.75}$	2054.3

Continued on next page

Table VI–2: The $\Delta\chi^2_{\rm min}$ results for the `ensemble-FUV kosmatau3d` model grid, which varies $\chi_{\rm cl}$ and $f_{\rm FUV}$ The minimum χ^2 is 4.968×10^8.

$\chi_{\rm cl}$	$f_{\rm FUV}$	$\Delta\chi^2_{\rm min}$
$10^{0.9}$	10^{2}	4857.1
$10^{0.9}$	$10^{2.25}$	6772.8
$10^{0.9}$	$10^{2.5}$	7646.8
$10^{0.9}$	$10^{2.75}$	8657.8
$10^{0.9}$	10^{3}	9729.7
$10^{1.1}$	$10^{0.25}$	189.4
$10^{1.1}$	$10^{0.5}$	343.3
$10^{1.1}$	$10^{0.75}$	585.4
$10^{1.1}$	10^{1}	900.5
$10^{1.1}$	$10^{1.25}$	1259.9
$10^{1.1}$	$10^{1.5}$	1657.3
$10^{1.1}$	$10^{1.75}$	2097.5
$10^{1.1}$	10^{2}	5154.8
$10^{1.1}$	$10^{2.25}$	6897.5
$10^{1.1}$	$10^{2.5}$	7528.4
$10^{1.1}$	$10^{2.75}$	8613.8
$10^{1.1}$	10^{3}	9690.0
$10^{1.3}$	$10^{0.25}$	184.9
$10^{1.3}$	$10^{0.5}$	346.4
$10^{1.3}$	$10^{0.75}$	592.5
$10^{1.3}$	10^{1}	908.3
$10^{1.3}$	$10^{1.25}$	1266.3
$10^{1.3}$	$10^{1.5}$	1661.6
$10^{1.3}$	$10^{1.75}$	2099.5
$10^{1.3}$	10^{2}	5154.6
$10^{1.3}$	$10^{2.25}$	6935.0
$10^{1.3}$	$10^{2.5}$	7520.7
$10^{1.3}$	$10^{2.75}$	8606.4
$10^{1.3}$	10^{3}	9680.9
$10^{1.5}$	$10^{0.25}$	205.7
$10^{1.5}$	$10^{0.5}$	374.4
$10^{1.5}$	$10^{0.75}$	623.9
$10^{1.5}$	10^{1}	939.5
$10^{1.5}$	$10^{1.25}$	1295.3
$10^{1.5}$	$10^{1.5}$	1687.2
$10^{1.5}$	$10^{1.75}$	2121.8
$10^{1.5}$	10^{2}	5237.1
$10^{1.5}$	$10^{2.25}$	6938.8
$10^{1.5}$	$10^{2.5}$	7531.0

Continued on next page

Table VI–2: The $\Delta\chi^2_{\rm min}$ results for the `ensemble-FUV kosmatau3d` model grid, which varies $\chi_{\rm cl}$ and $f_{\rm FUV}$ The minimum χ^2 is 4.968×10^8.

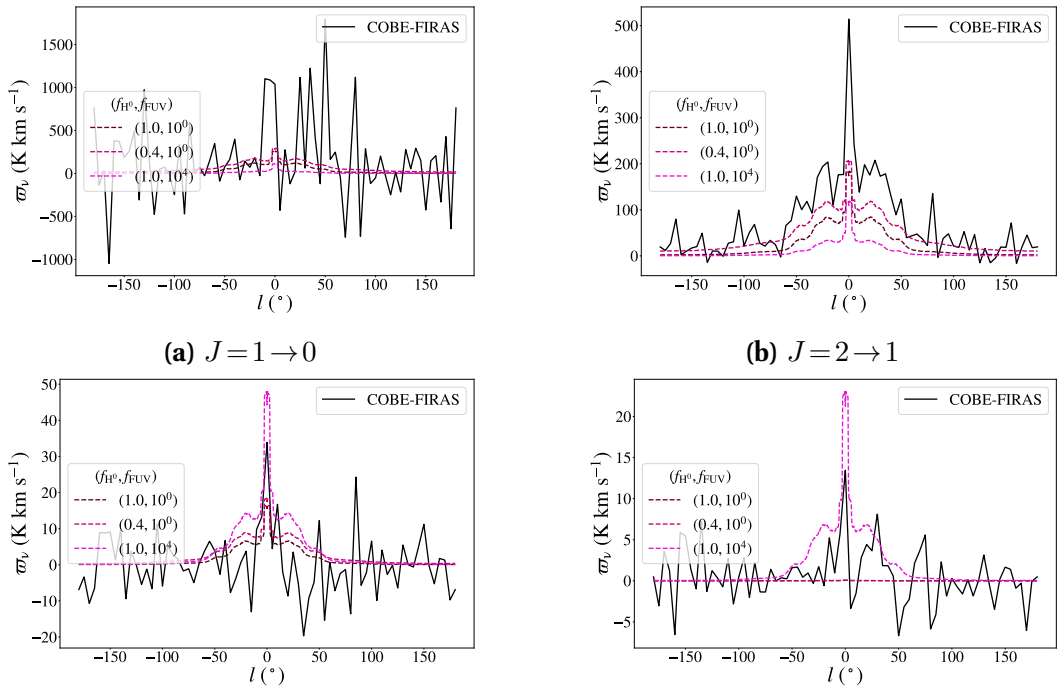
$\chi_{\rm cl}$	$f_{\rm FUV}$	$\Delta\chi^2_{\rm min}$
$10^{1.5}$	$10^{2.75}$	8615.6
$10^{1.5}$	10^{3}	9687.8
$10^{1.7}$	$10^{0.25}$	264.2
$10^{1.7}$	$10^{0.5}$	438.6
$10^{1.7}$	$10^{0.75}$	690.1
$10^{1.7}$	10^{1}	1004.2
$10^{1.7}$	$10^{1.25}$	1356.1
$10^{1.7}$	$10^{1.5}$	1743.3
$10^{1.7}$	$10^{1.75}$	2172.8
$10^{1.7}$	10^{2}	5268.8
$10^{1.7}$	$10^{2.25}$	6970.6
$10^{1.7}$	$10^{2.5}$	7566.7
$10^{1.7}$	$10^{2.75}$	8648.4
$10^{1.7}$	10^{3}	9717.0
$10^{1.9}$	$10^{0.25}$	361.5
$10^{1.9}$	$10^{0.5}$	539.0
$10^{1.9}$	$10^{0.75}$	790.4
$10^{1.9}$	10^{1}	1100.9
$10^{1.9}$	$10^{1.25}$	1447.1
$10^{1.9}$	$10^{1.5}$	1827.5
$10^{1.9}$	$10^{1.75}$	2250.1
$10^{1.9}$	10^{2}	5345.7
$10^{1.9}$	$10^{2.25}$	7026.1
$10^{1.9}$	$10^{2.5}$	7622.9
$10^{1.9}$	$10^{2.75}$	8701.1
$10^{1.9}$	10^{3}	9764.5
$10^{2.1}$	$10^{0.25}$	468.3
$10^{2.1}$	$10^{0.5}$	646.4
$10^{2.1}$	$10^{0.75}$	895.4
$10^{2.1}$	10^{1}	1200.6
$10^{2.1}$	$10^{1.25}$	1539.6
$10^{2.1}$	$10^{1.5}$	1912.0
$10^{2.1}$	$10^{1.75}$	2326.8
$10^{2.1}$	10^{2}	5406.2
$10^{2.1}$	$10^{2.25}$	7076.8
$10^{2.1}$	$10^{2.5}$	7675.3
$10^{2.1}$	$10^{2.75}$	8749.1
$10^{2.1}$	10^{3}	9806.7
$10^{2.3}$	$10^{0.25}$	634.4

Continued on next page

Table VI–2: The $\Delta\chi^2_{\rm min}$ results for the `ensemble-FUV kosmatau3d` model grid, which varies $\chi_{\rm cl}$ and $f_{\rm FUV}$ The minimum χ^2 is 4.968×10^8.

$\chi_{\rm cl}$	$f_{\rm FUV}$	$\Delta\chi^2_{\rm min}$
$10^{2.3}$	$10^{0.5}$	811.0
$10^{2.3}$	$10^{0.75}$	1056.1
$10^{2.3}$	10^{1}	1354.7
$10^{2.3}$	$10^{1.25}$	1685.6
$10^{2.3}$	$10^{1.5}$	2049.2
$10^{2.3}$	$10^{1.75}$	2455.4
$10^{2.3}$	10^{2}	5841.6
$10^{2.3}$	$10^{2.25}$	7181.5
$10^{2.3}$	$10^{2.5}$	7777.0
$10^{2.3}$	$10^{2.75}$	8845.8
$10^{2.3}$	10^{3}	9895.9
$10^{2.5}$	$10^{0.25}$	829.2
$10^{2.5}$	$10^{0.5}$	1002.6
$10^{2.5}$	$10^{0.75}$	1241.6
$10^{2.5}$	10^{1}	1531.6
$10^{2.5}$	$10^{1.25}$	1852.3
$10^{2.5}$	$10^{1.5}$	2205.0
$10^{2.5}$	$10^{1.75}$	2600.4
$10^{2.5}$	10^{2}	5903.1
$10^{2.5}$	$10^{2.25}$	7293.4
$10^{2.5}$	$10^{2.5}$	7890.4
$10^{2.5}$	$10^{2.75}$	8951.9
$10^{2.5}$	10^{3}	9993.6

5 Atomic fraction

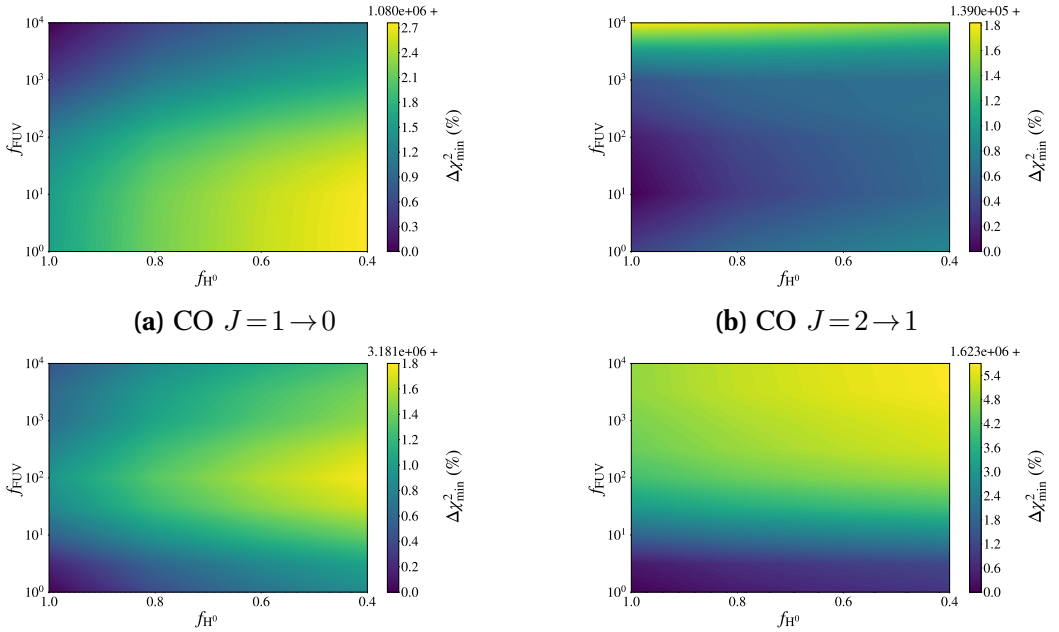


(a) $J = 1 \rightarrow 0$ **(b)** $J = 2 \rightarrow 1$

(c) $J = 6 \rightarrow 5$ **(d)** $J = 8 \rightarrow 7$

Figure VI–23: The same as Figure VI–11 for the `atomic-ISM-radiation` model grid.

By varying the fraction of the atomic hydrogen in the interclump medium to the far-UV radiation in which the ensembles are embedded, as in the `atomic-ISM-radiation` grid, it is possible to constrain the validity of our default assumption that molecular hydrogen is represented by the clumpy ensemble and atomic hydrogen by the interclump ensemble. The comparison of the synthetic intensity profiles to the *COBE*-FIRAS survey indicates that the upper limit of far-UV in this grid does not match well (see Figure VI–23). It results in too high intensity in the upper-J lines and too low indensity in the low-J lines. By eye it seems that the $J = 2 \rightarrow 1$ and $J = 6 \rightarrow 5$ provide the strongest constraint, both of which indicate that the models with $f_{H^0} = 0.4$ more closely match the observations. However, the quantitative comparison of CO transitions in Figure VI–24 slightly indicates the opposite trend (with variation of the test statistic remaining $< 10\%$ compared to $\Delta\chi^2_{\text{min}} \sim 10^6$). While there is no clear trend regarding the far-UV intensity, a slight conclusion that can be drawn by separately considering the constraints of each transition: it is sufficient for all of the atomic hydrogen to be bound in the interclump ensemble (subfigures **(a)**-**(c)**) and there is no need to drastically increase the far-UV radiation galactic profile (subfigure **(d)**). The χ^2 statistic seems to be lower when all of the atomic hydrogen is in the interclump medium, in agreement with the result for all transitions.

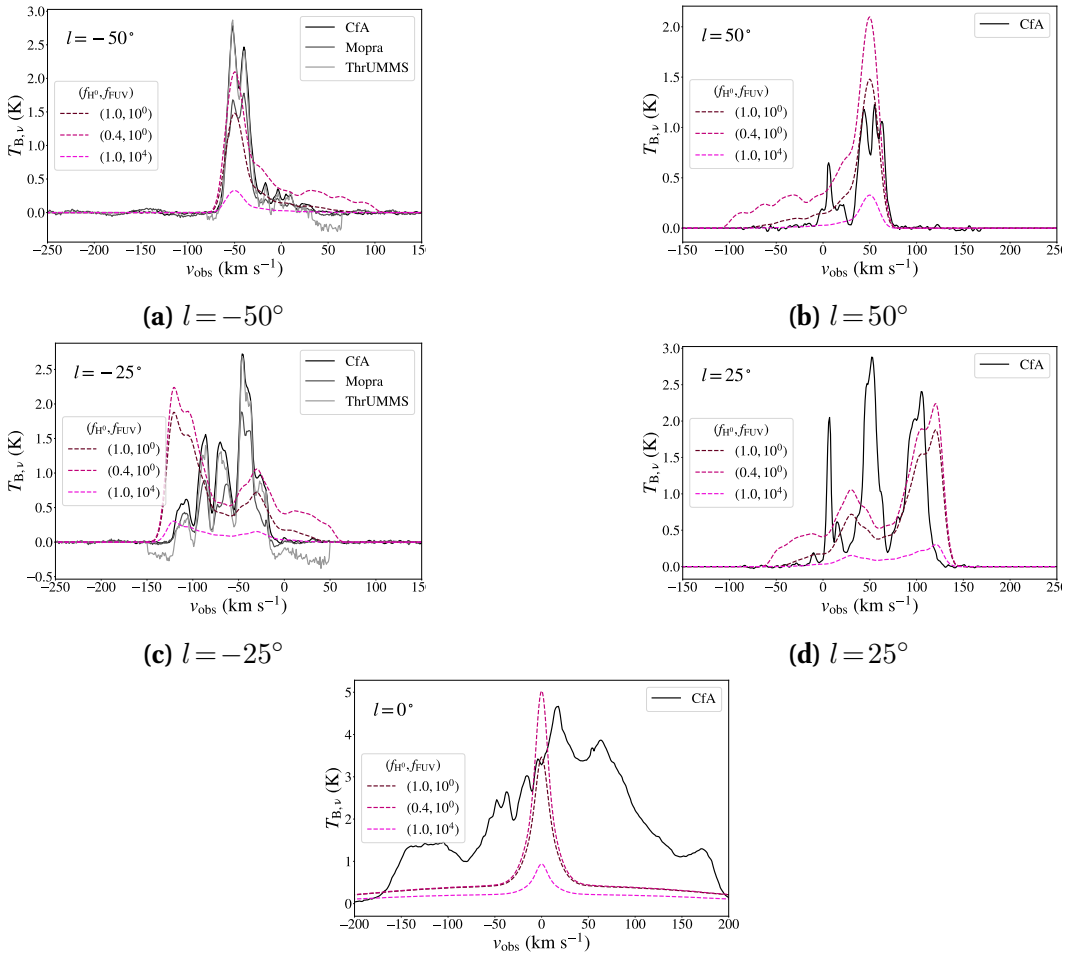


(a) CO $J=1\rightarrow0$ **(b)** CO $J=2\rightarrow1$

(c) CO $J=6\rightarrow5$ **(d)** CO $J=8\rightarrow7$

Figure VI–24: The comparison between the CO transitions observed by *COBE*-FIRAS and the `atomic-ISM-radiation` model grid, neglecting extraneous values of χ^2.

By plotting the synthetic observations with the spectroscopic observational surveys for various sightlines, we can get a qualitative impression of how well the models match (see Figure VI–25 to Figure VI–25). What is interesting to see in these direct comparisons is that decreasing $f_{\mathrm{H^0}}$ also increases the CO emission away from the bulk emission (seen in Figure VI–25 subfigure **(a)** and **(b)**, Figure VI–26 subfigure **(a)** and **(c)**, and Figure VI–27 subfigure **(a)** and **(b)**). This is due to the increased mass of the clumpy ensemble, which is primarily molecular gas containing CO. The intensity spectra from the model with $(f_{\mathrm{H^0}}, f_{\mathrm{FUV}}) = (0.4, 1.0)$ therefore is too large for most of the spectral domain compared to the observations. Conversely, the f_{FUV} parameter matches better for lower values, since the synthetic spectra from model $(f_{\mathrm{H^0}}, f_{\mathrm{FUV}}) = (1.0, 10^4)$ typically too-low in intensity compared to the observations (except for subfigure **(d)** in Figure VI–26 and Figure VI–27). The lower values of f_{FUV} appear to underestimate the CO emission and overestimate the ^{13}CO emission, though. A possible explanation of this is that CO has higher optical depth than ^{13}CO, and by construction the `kosmatau3d` galactic models overestimate the optical depth by having clumpy gas in each voxel.

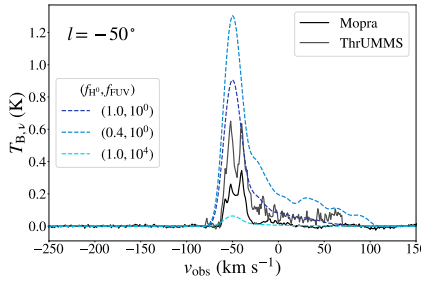


(a) $l = -50°$ **(b)** $l = 50°$

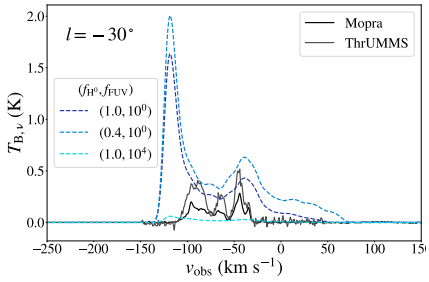
(c) $l = -25°$ **(d)** $l = 25°$

(e) $l = 0°$

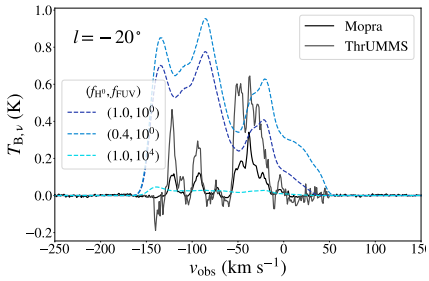
Figure VI–25: The same as Figure VI–13 for the `atomic-ISM-radiation` model grid.



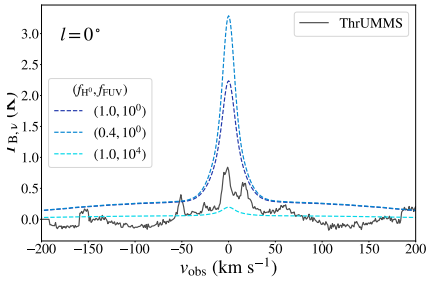
(a) $l = -50°$



(b) $l = -30°$

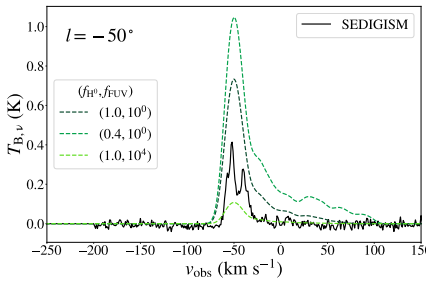


(c) $l = -20°$

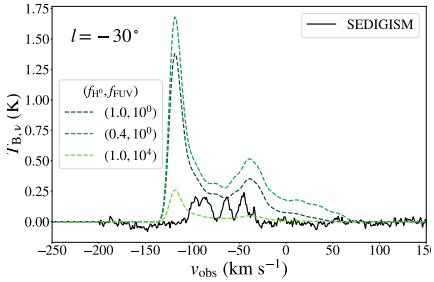


(d) $l = 0°$

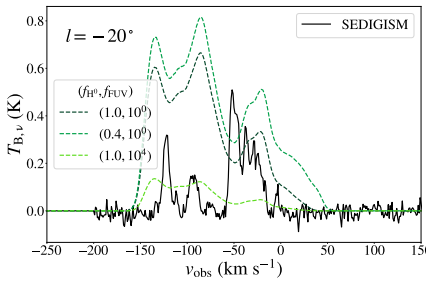
Figure VI–26: The same as Figure VI–14 for the `atomic-ISM-radiation` model grid.



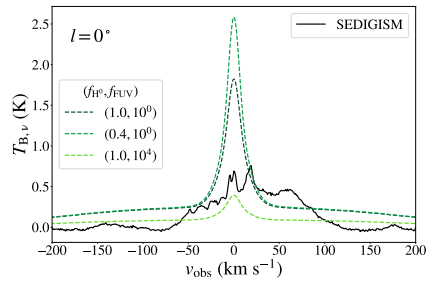
(a) $l = -50°$



(b) $l = -30°$



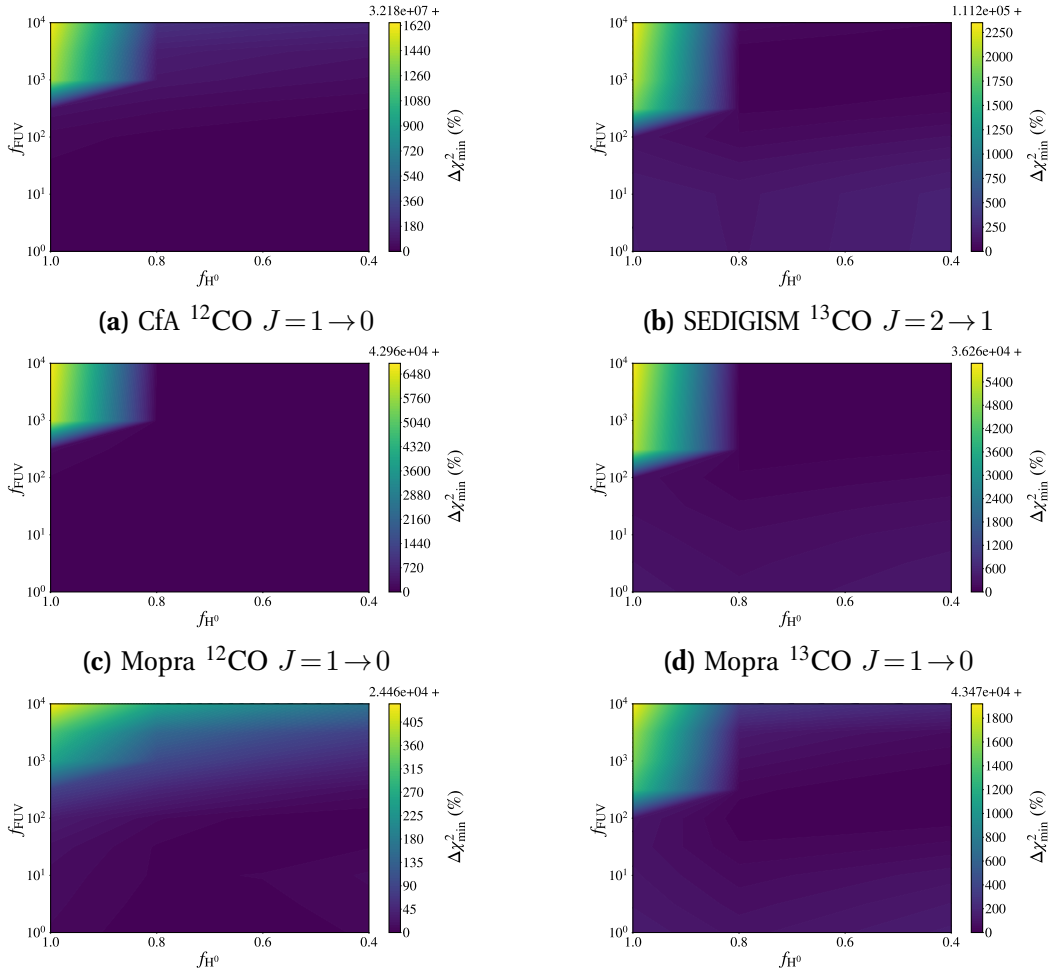
(c) $l = -20°$



(d) $l = 0°$

Figure VI–27: The same as Figure VI–15 for the `atomic-ISM-radiation` model grid.

The comparison of the `atomic-ISM-radiation` grid to the spectroscopic observation survey again yields larger variations of the test statistic in the parameter space ($\chi^2_{\mathrm{min}} \gg 100\%$ for $\chi^2_{\mathrm{min}} \sim 10^4$; the CfA comparison has $\chi^2_{\mathrm{min}} \sim 10^7$). While it is possible that the χ^2_{min} is higher for the comparison to the CfA CO $J = 1 \rightarrow 0$ line since there are more degrees of freedom, the difference is not so great as to cause a discrepancy of the order 3. The $\Delta\chi^2_{\mathrm{min}}$ test statistic should still be an accurate metric by which we can constrain the parameter space. There is a clear indication that we can constrain the far-UV radiation to $\lesssim 10\,\chi_{\mathrm{D}}$ and a slight constraint that $f_{\mathrm{H^0}} < 1$. The comparison of the SEDIGISM CO $J = 2 \rightarrow 1$ transition seems to be the only comparison to constrain $10^1 \lesssim 10^2$.



(a) CfA ^{12}CO $J = 1 \rightarrow 0$

(b) SEDIGISM ^{13}CO $J = 2 \rightarrow 1$

(c) Mopra ^{12}CO $J = 1 \rightarrow 0$

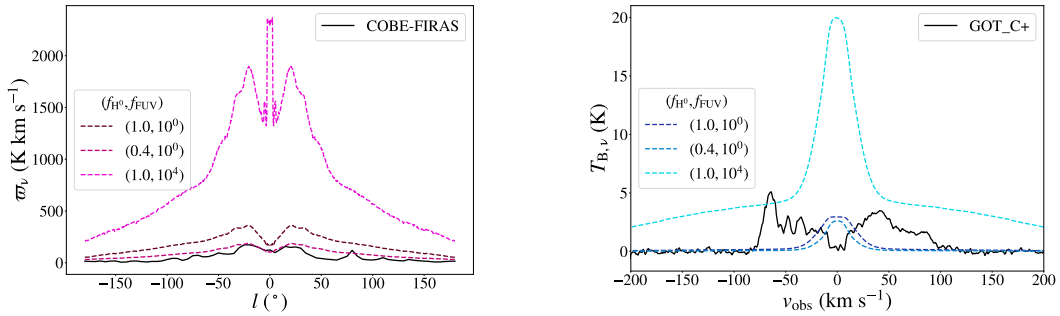
(d) Mopra ^{13}CO $J = 1 \rightarrow 0$

(e) ThrUMMS ^{12}CO $J = 1 \rightarrow 0$

(f) ThrUMMS ^{13}CO $J = 1 \rightarrow 0$

Figure VI–28: The comparison between the spectroscopic CO observations and the synthetic observations of the `atomic-ISM-radiation` model grid for the transitions ^{12}CO $J = 1 \rightarrow 0$, ^{13}CO $J = 1 \rightarrow 0$, and ^{13}CO $J = 2 \rightarrow 1$.

The [^{12}C II] 158 μm line emission traces better the atomic medium and should thus provide a stronger constraint on the interclump ensemble properties. We qualitatively compare our model parameter space to the integrated-intensity map and spectrum towards the Galactic centre to gain an impression of how well our models fit the observations (see Figure VI–29). Here it can be seen that there is a much

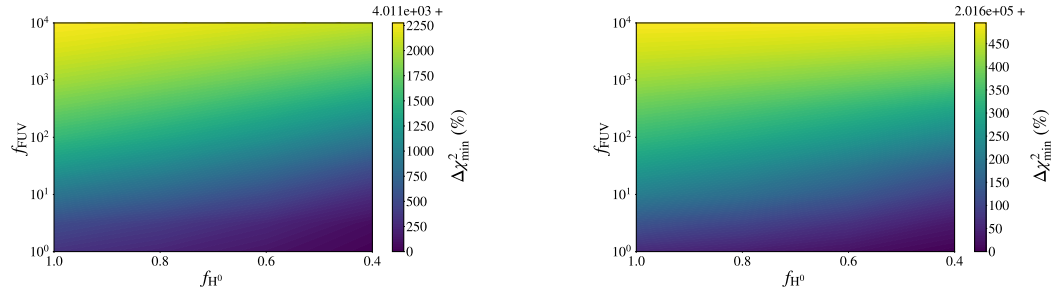


(a) *COBE*-FIRAS **(b)** GOT C+

Figure VI–29: The same as Figure VI–17 for the `atomic-ISM-radiation` model grid.

closer fit to the Galactic ring and Galactic centre features in subfigure **(a)** (seen in the peaks of the profile), although the emission away from the Galactic centre is quite a bit larger than the *COBE*-FIRAS observations. The synthetic spectrum in subfigure **(b)** does not have the characteristic self-absorption of the [^{12}C II] 158 μm transition like the models in Figure VI–17, and still lacks the velocity dispersion of the observations.

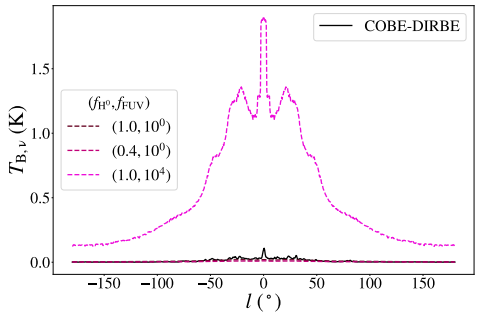
The `atomic-ISM-radiation` parameter space features a very large variation of $\Delta\chi^2_{\text{min}}$ compared to $\chi^2_{\text{min}} = 4.011 \times 10^3,\ 2.016 \times 10^5$ (see Figure VI–30). This indicates that there exists a strong constraint in the parameter space. It can thus be concluded that $f_{\text{H}^0} \approx 0.4$ and $f_{\text{FUV}} \sim 1$.



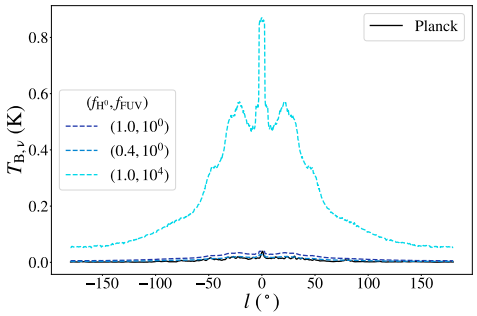
(a) *COBE*-FIRAS **(b)** GOT C+

Figure VI–30: The comparison between the observations to the `atomic-ISM-radiation kosmatau3d` grid for the transition [^{12}C II] 158 μm.

It is expected that the dust comparison is sensitive to the far-UV radiation (see §I-1.2.4), but it is important to know how the comparison is affected by f_{H^0}. The comparison of the synthetic and observed continuum wavelengths is shown in Figure VI–31 for three model approximately spanning the parameter space. None of the models can replicate the dust intensity of the molecular ring at 240 μm, however the model with $f_{\text{H}^0} = 1$ can approximately replicate the intensity away from the Galactic centre. Conversely the model with $f_{\text{H}^0} = 0.4$ better represents the Galactic ring at 550 μm, but cannot match the intensity away from the Galactic centre. Neither model can fit both wavelengths well. Quantitatively, there is a very large variation of the test statistic in the parameter space, with $\Delta\chi^2_{\text{min}} \gg 100\%$



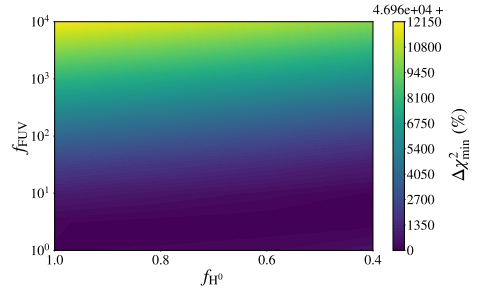
(a) *COBE*-DIRBE 240 μm



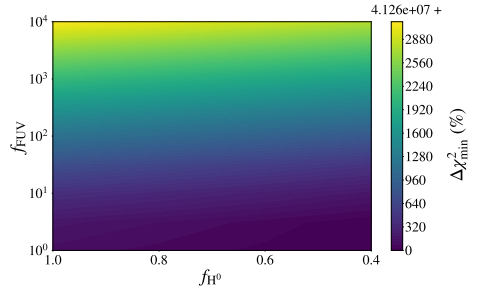
(b) *Planck* 550 μm

Figure VI–31: The same as Figure VI–19 for the `atomic-ISM-radiation` model grid.

compared to $\chi^2_{\rm min} = 4.696 \times 10^4,\ 4.126 \times 10^7$ (see Figure VI–32). While there is a strong constraint that $f_{\rm FUV} \lesssim 3.2$ for the comparison to both wavelengths, there is also a slight (barely visible) constraint to $f_{\rm H^0} \leq 0.8$.

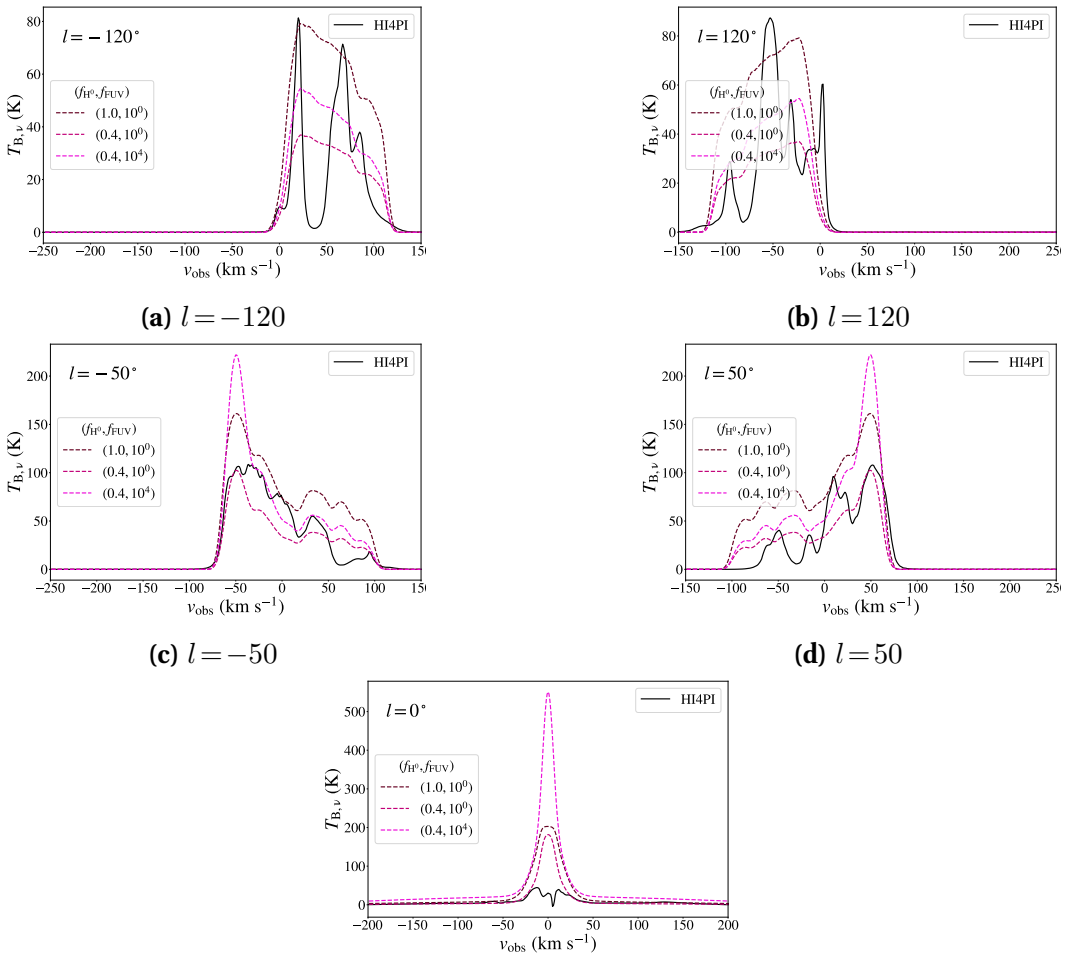


(a) *COBE*-DIRBE 240 μm



(b) *Planck* 550 μm

Figure VI–32: The comparison between the dust continuum observations to the `atomic-ISM-radiation kosmatau3d` grid at the specified wavelengths.



(a) $l = -120$ **(b)** $l = 120$

(c) $l = -50$ **(d)** $l = 50$

(e) $l = 0$

Figure VI–33: The same as Figure VI–21 for the `atomic-ISM-radiation` model grid.

The H I 21 cm transition mainly traces the interclump medium, and thus it should be very sensitive to parameters such as f_{H^0}. The synthetic spectra from the models spanning the parameter space are compared to the observed spectra in Figure VI–33. Here we see the model with $f_{H^0} = 0.4$ overestimates the intensity in all spectra. The synthetic spectra at $l = \pm 120°$ (subfigures **(a)** and **(b)**) do not show as much structure as the observed spectra (seen with the multiple intensity peaks). The model with $f_{H^0} = 1$ has a peak intensity comparable to the observed spectra, but the model with $f_{H^0} = 1$ has too low intensity. The spectra at $l = \pm 50°$ (subfigures **(c)** and **(d)**) are fitting a bit closer to the observations than the other spectra for $f_{FUV} = 1$, but the model with $f_{H^0} = 1$ overestimates the intensity and the model with $f_{H^0} = 0.4$ underestimates the intensity. Subfigure **(e)** shows that the spectra towards the Galactic centre is fitting well in the wings, but there is definitely too high intensity compared to the observed spectrum.

The parameter space of the `atomic-ISM-radiation` model grid varies two parameters directly affecting the interclump ensemble which, as seen in the test statistic (Figure VI–34), are both strongly constrained. The minimum χ^2 of $\sim$

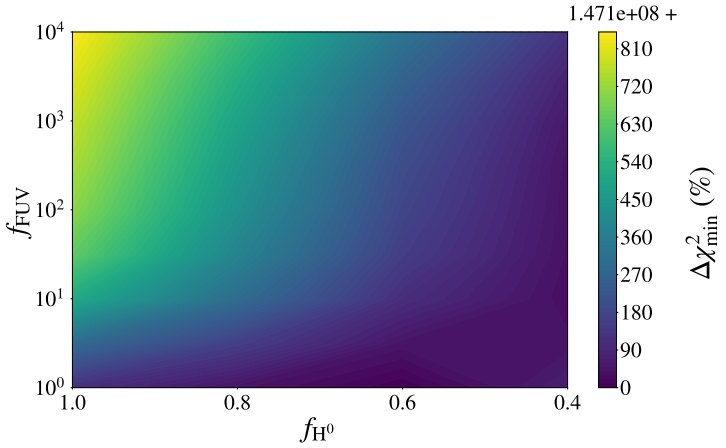


Figure VI–34: The comparison between the H I observations to the `atomic-ISM-radiation kosmatau3d` grid for the transition H I 21 cm.

1.471×10^8 occurs for $f_{\rm H^0} \gtrsim 0.4$ and $f_{\rm FUV} \gtrsim 1$. Overall, this test statistic is lower than and remains in agreement with the best-fit `ensemble-FUV`, giving further support to lowering the amount of atomic hydrogen we assume is represented by our interclump ensemble.

The $\Delta\chi^2_{\rm min}$ results are combined for the various comparisons (except for the *Planck* comparison; see §VII-7 for the discussion of why it is neglected) and summarised in VI–3.

Table VI–3: The $\Delta\chi^2_{\rm min}$ results for the `atomic-ISM-radiation kosmatau3d` model grid, which varies $f_{\rm H^0}$ and $f_{\rm FUV}$ The minimum χ^2 is 1.895×10^8.

$f_{\rm H^0}$	$f_{\rm FUV}$	$\Delta\chi^2_{\rm min}$
0.4	1	1833.2
0.4	$10^{0.5}$	1290.0
0.4	10^{1}	1471.3
0.4	$10^{1.5}$	2332.9
0.4	10^{2}	3790.0
0.4	$10^{2.5}$	5649.7
0.4	10^{3}	7915.8
0.4	$10^{3.5}$	10387.7
0.4	10^{4}	13287.0
0.6	1	1452.2
0.6	$10^{0.5}$	1288.8
0.6	10^{1}	1927.1
0.6	$10^{1.5}$	3119.0
0.6	10^{2}	4766.7
0.6	$10^{2.5}$	6747.4
0.6	10^{3}	9071.0

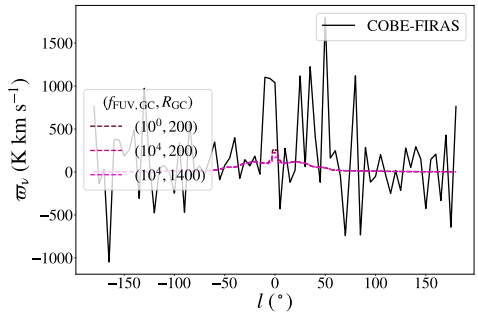
Continued on next page

Table VI–3: The $\Delta\chi^2_{\rm min}$ results for the `atomic-ISM-radiation kosmatau3d` model grid, which varies $f_{\rm H^0}$ and $f_{\rm FUV}$ The minimum χ^2 is 1.895×10^8.

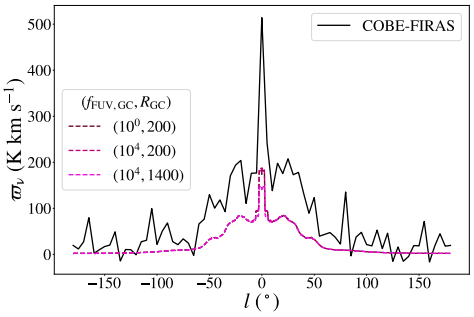
$f_{\rm H^0}$	$f_{\rm FUV}$	$\Delta\chi^2_{\rm min}$
0.6	$10^{3.5}$	11557.1
0.6	10^{4}	14511.7
0.8	**1**	**1331.9**
0.8	$10^{0.5}$	1461.1
0.8	10^{1}	2462.0
0.8	$10^{1.5}$	3929.4
0.8	10^{2}	5751.1
0.8	$10^{2.5}$	7857.7
0.8	10^{3}	10247.3
0.8	$10^{3.5}$	12754.6
0.8	10^{4}	15780.4
1	1	1756.5
1	$10^{0.5}$	2129.7
1	10^{1}	3445.4
1	$10^{1.5}$	5185.0
1	10^{2}	7301.2
1	$10^{2.5}$	23736.7
1	10^{3}	28463.8
1	$10^{3.5}$	31603.8
1	10^{4}	35112.8

6 Galactic core far-UV

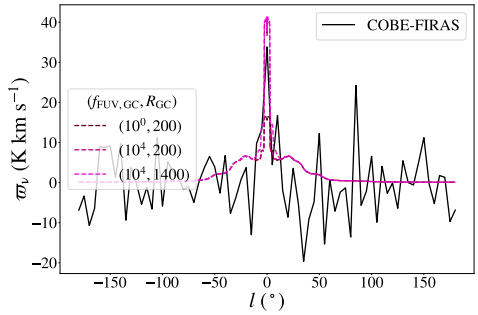
In the `FUV-Galactic-centre` model grid, we examine the effect of varying the Galactic centre properties $(f_{\rm FUV,GC}, R_{\rm GC})$. This is attempt to increase the amount of H^0 and $^{12}C^+$ in the Galactic centre, and thus recreate the observed absorption features. Since we alter the parameters within 1400 pc of the Galactic centre, we should only see an effect at longitudes $|l| \lesssim 10^\circ$ (see the comparison in Figure VI–35 for models spanning the parameter space). Here we can see that there is little effect to the low-J CO lines, but the increased far-UV radiation creates a peak in the high-J CO lines. Moreover, the variation $\max(\Delta\chi^2) \ll 1\%$, so we cannot gather much information from this comparison Figure VI–36.



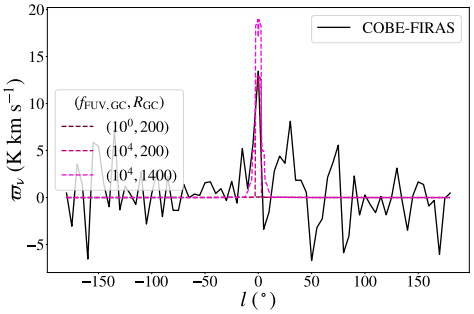
(a) $J = 1 \rightarrow 0$



(b) $J = 2 \rightarrow 1$

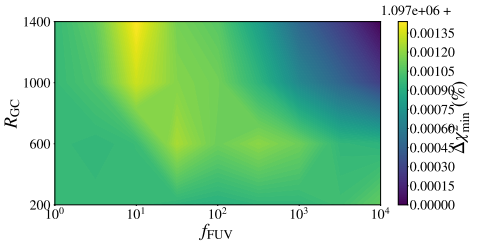


(c) $J = 6 \rightarrow 5$

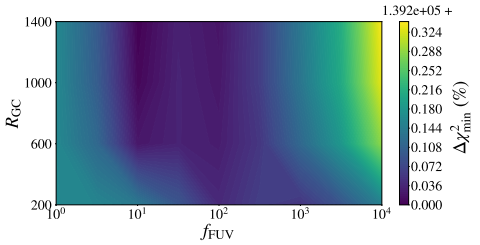


(d) $J = 8 \rightarrow 7$

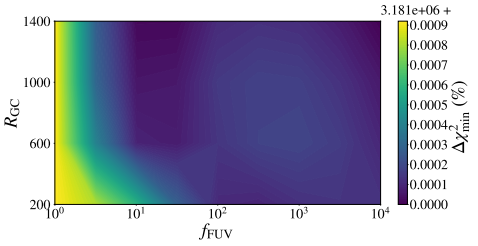
Figure VI–35: The same as Figure VI–11 for the `FUV-Galactic-centre` model grid.



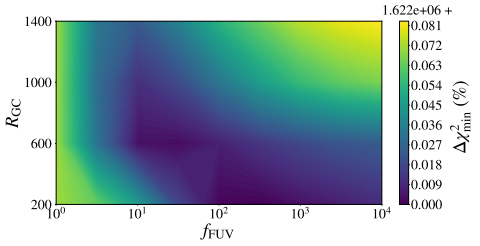
(a) CO $J = 1 \rightarrow 0$



(b) CO $J = 2 \rightarrow 1$

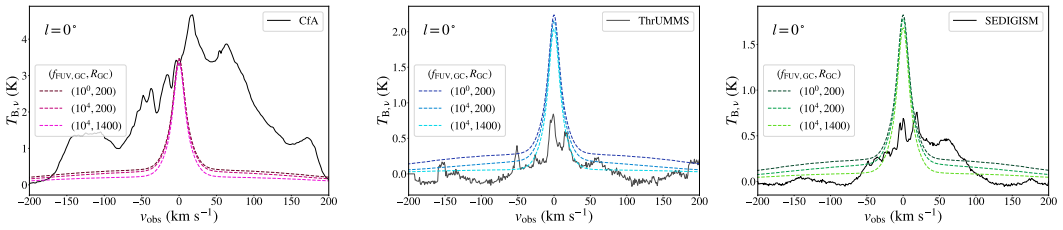


(c) CO $J = 6 \rightarrow 5$



(d) CO $J = 8 \rightarrow 7$

Figure VI–36: The comparison between the CO transitions observed by *COBE*-FIRAS and the `FUV-Galactic-centre` model grid, neglecting extraneous values of χ^2.

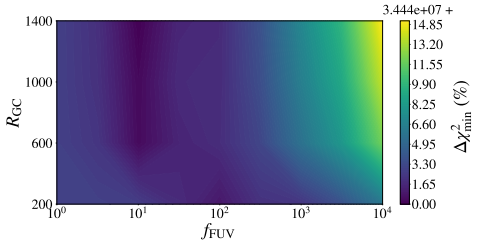


(a) $l = 0°$, CO $J = 1 \rightarrow 0$ **(b)** $l = 0°$, ^{13}CO $J = 1 \rightarrow 0$ **(c)** $l = 0°$, ^{13}CO $J = 2 \rightarrow 1$

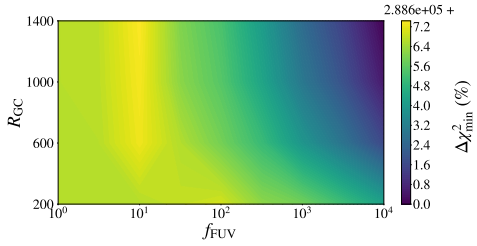
Figure VI–37: The same as Figure VI–13 to Figure VI–15 for the `FUV-Galactic-centre` model grid.

We can qualitatively examine spectra in the models with parameters $(f_{\mathrm{FUV}}, r_{\mathrm{GC}}) \in [(0.4, 1.0), (1.0, 1.0), (1.0, 31.6)]$ at $l \in [-50, 0, 50]°$ (see Figure VI–37). Since this grid varies the Galactic centre parameters and our spectrum at $l = 0$ does not fit at all, we cannot conclude anything from the qualitative analysis. Very little variation is seen between the reference models, but in general the CO line is underestimated (still not matching the observed velocity dispersion) and the ^{13}CO lines are overestimated compared to the observations.

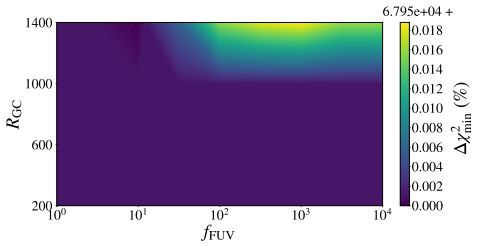
This lack of variation in the test statistic is also seen in the comparison to the spectroscopic surveys (see Figure VI–38). The only very minor constraints come from the SEDIGISM ^{13}CO $J = 2 \rightarrow 1$ comparison that the far-UV radiation intensity is weakly constrained to $f_{\mathrm{FUV}} \lesssim 10^3$, and from the CfA CO $J = 1 \rightarrow 0$ comparison constrains the far-UV radiation intensity to $10^1 \lesssim f_{\mathrm{FUV}} \lesssim 10^2$. These constraints are not very binding, however.



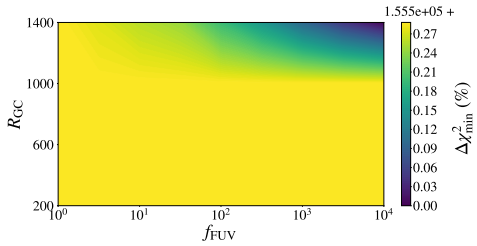
(a) CfA ^{12}CO $J=1\rightarrow0$



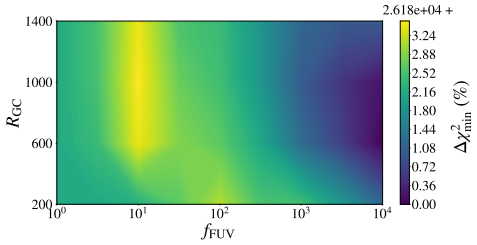
(b) SEDIGISM ^{13}CO $J=2\rightarrow1$



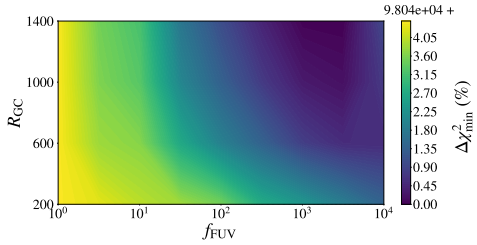
(c) Mopra ^{12}CO $J=1\rightarrow0$



(d) Mopra ^{13}CO $J=1\rightarrow0$



(e) ThrUMMS ^{12}CO $J=1\rightarrow0$

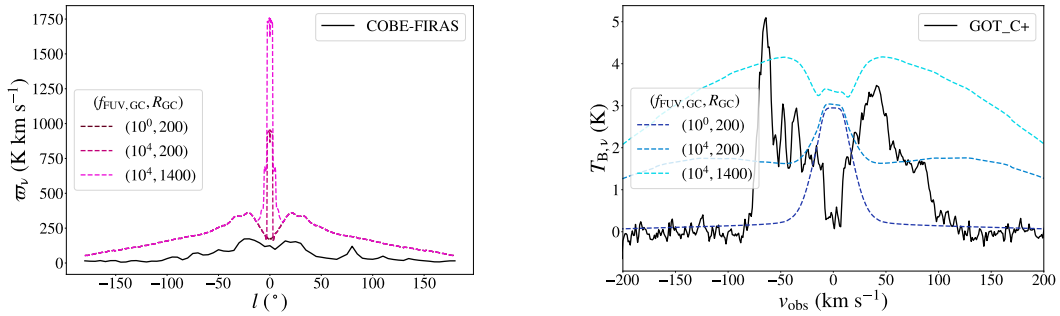


(f) ThrUMMS ^{13}CO $J=1\rightarrow0$

Figure VI–38: The comparison between the spectroscopic CO observations and the synthetic observations of the `FUV-Galactic-centre` model grid for the transitions ^{12}CO $J=1\rightarrow0$, ^{13}CO $J=1\rightarrow0$, and ^{13}CO $J=2\rightarrow1$.

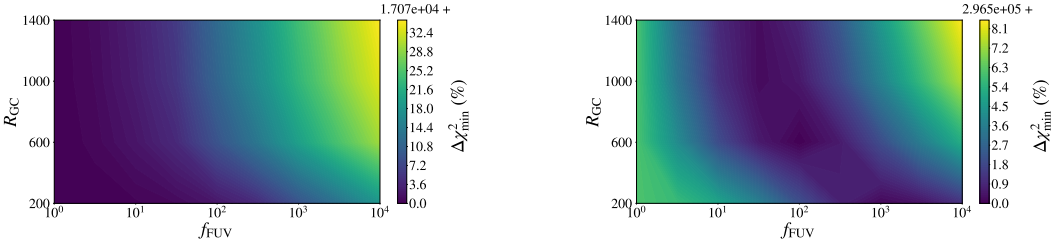
The [^{12}C II] 158 μm line exhibits more variation in this model grid than the CO lines, but we cannot replicate the self-absorption feature of the observations (see the synthetic and observed integrated-intensity profile and spectra in Figure VI–39). We see in subfigure **(a)** that there is no longer self-absorption in synthetic integrated-intensity profile when $R_{\rm GC}=200$ pc and $f_{\rm FUV,GC}=10^4$, nor is there self absorption in the synthetic spectra in subfigure **(b)**. The synthetic profile when $R_{\rm GC}$ model we have included also is missing the absorption feature in subfigure **(a)**, but it does have an absorption feature in subfigure **(b)**. Besides the existence of this absorption feature, the rest of the synthetic spectrum is too intense compared to the observations.

There is slight variation also in test statistic, with $\chi^2_{\rm min}\sim10\%$ compared to $\chi^2_{\rm min}=1.707\times10^4,\ 2.965\times10^5$ (see Figure VI–40). The test for the comparisons to *COBE*-FIRAS and GOT C+ appear to be fitting opposite areas of the parameter space, though, with an approximate best-fit region with $f_{\rm FUV}\approx10$ and $r_{\rm GC}\in(600,1200)$ pc.



(a) *COBE*-FIRAS **(b)** GOT C+

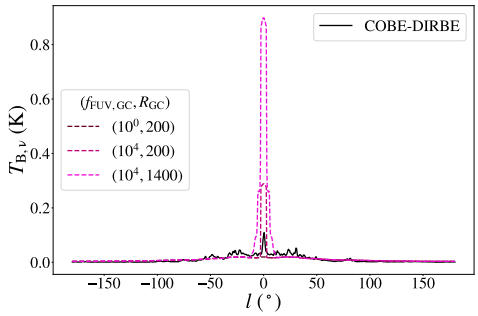
Figure VI–39: The same as Figure VI–17 for the `FUV-Galactic-centre` model grid.



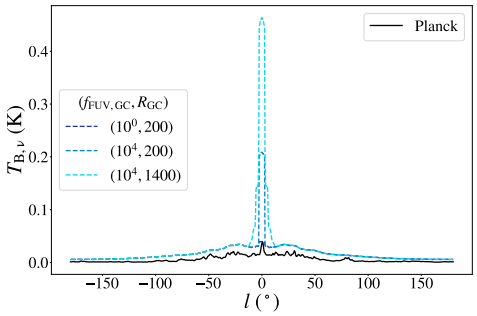
(a) *COBE*-FIRAS **(b)** GOT C+

Figure VI–40: The comparison between the observations to the `FUV-Galactic-centre` `kosmatau3d` grid for the transition [^{12}C II] 158 μm.

The dust continuum is sensitive to far-UV radiation, thus the comparison of the synthetic continuum to the observed continuum should give the strongest constraint for the `FUV-Galactic-centre` model grid. The intensity profiles shown in Figure VI–41 for models spanning the parameter space confirm that it is possible to fit better the peak continuum emission at $l = 0°$. The intensity profile at 240 μm is smaller than the observed profile while the profile at 550 μm is larger than the observed profile for all models.



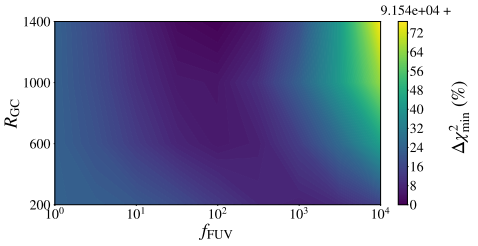
(a) *COBE*-DIRBE 240 μm



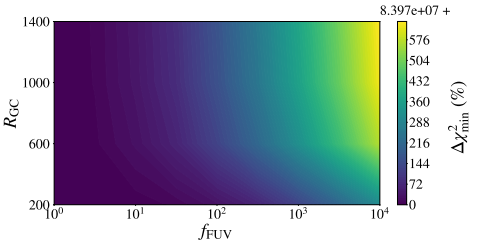
(b) *Planck* 550 μm

Figure VI–41: The same as Figure VI–19 for the `FUV-Galactic-centre` model grid.

If we are to consider the minimum χ^2, our best-fitting results for both wavelengths are similar to the results of §VI-5: $(f_{\mathrm{H^0}}, f_{\mathrm{FUV}}) = (0.4, 1.78)$. In the quantitative comparison to the model grid, the variation of the test statistic is $\chi^2_{\mathrm{min}} \sim 10^2\%$ compared to $\chi^2_{\mathrm{min}} = 9.154 \times 10^4,\ 8.397 \times 10^7$ (see Figure VI–42). By approximately combining the the results of subfigures **(a)** and **(b)**, there is a constraint on the size of the galactic core ($r_{\mathrm{GC}} \gtrsim 200$ pc) and the radiation in the core ($f_{\mathrm{FUV}} \lesssim 10$).

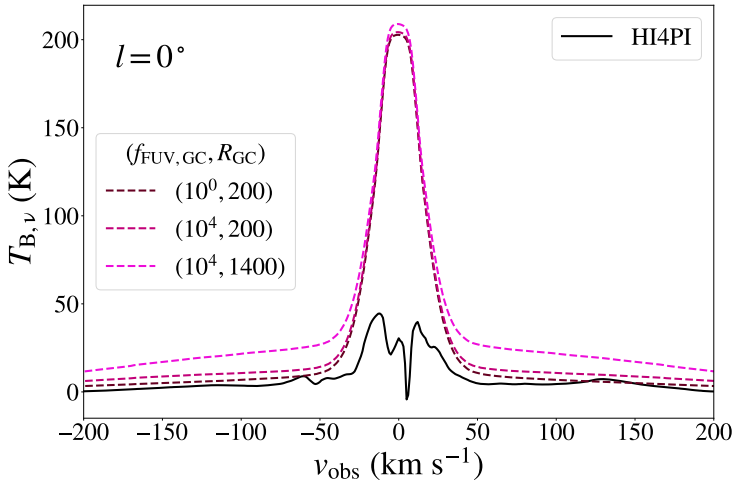


(a) *COBE*-DIRBE 240 μm



(b) *Planck* 550 μm

Figure VI–42: The comparison between the dust continuum observations to the `FUV-Galactic-centre` `kosmatau3d` grid at the specified wavelengths.



(a) $l = 0$

Figure VI–43: The same as Figure VI–21 for the `FUV-Galactic-centre` model grid.

We still compare the spectra in Figure VI–43, though this time only at $l = 0$ since only the variables in the Galactic centre are affected. This plot clearly shows a lack of velocity dispersion in these models, and none of the synthetic models feature self-absorption effects. For the `FUV-Galactic-centre` model grid there is not significant variation of the test statistic (see Figure VI–44). There is a minor constraint of $(f_{\mathrm{FUV}}, r_{\mathrm{GC}}) = (10^3, 600\,\mathrm{pc})$, but the minimum χ^2 is 3.122×10^8.

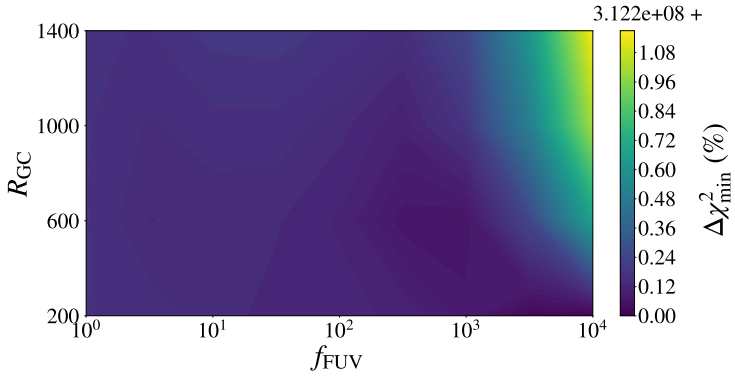
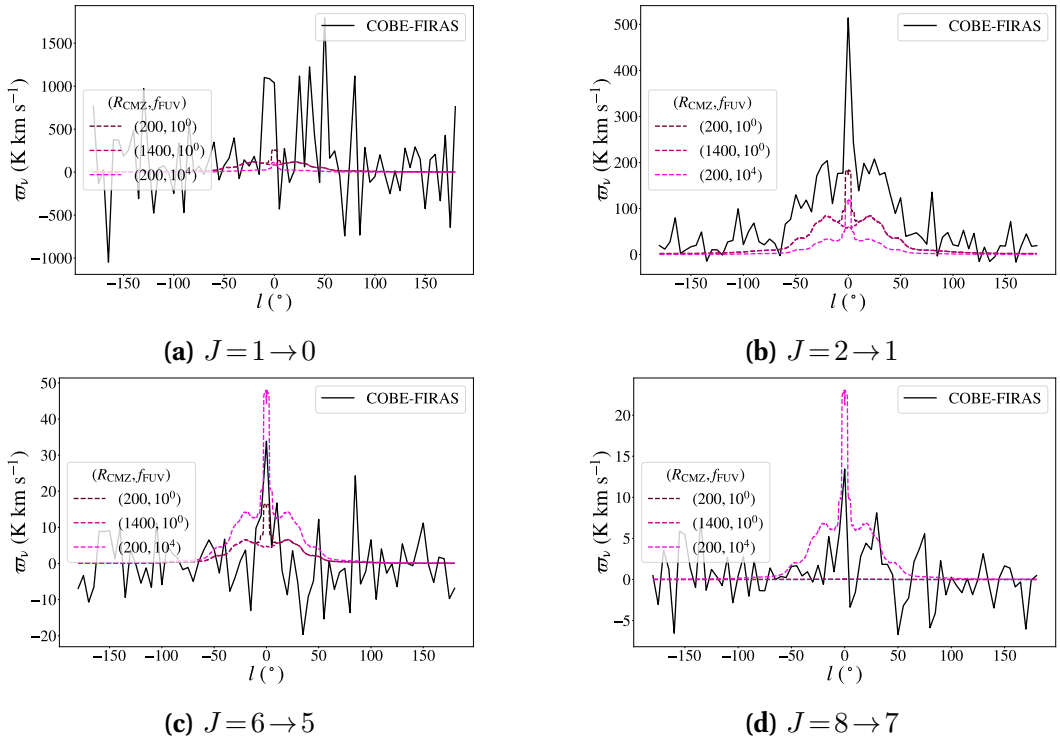


Figure VI–44: The comparison between the H I observations to the `FUV-Galactic-centre` `kosmatau3d` grid for the transition H I 21 cm.

The $\Delta\chi^2_{\mathrm{min}}$ results are combined for the various comparisons (except for the *Planck* comparison; see §VII-7 for the discussion of why it is neglected) and summarised in VI–4.

Table VI–4: The $\Delta\chi^2_{\rm min}$ results for the `FUV-Galactic-centre` `kosmatau3d` model grid, which varies $f_{\rm FUV,GC}$ and $R_{\rm GC}$ The minimum χ^2 is 3.575×10^8.

$f_{\rm FUV,GC}$	$R_{\rm GC}$	$\Delta\chi^2_{\rm min}$
1	200	49.2
1	600	49.2
1	1000	49.2
1	1400	49.2
$10^{0.5}$	200	47.9
$10^{0.5}$	600	41.7
$10^{0.5}$	1000	38.8
$10^{0.5}$	1400	37.5
10^{1}	200	44.9
10^{1}	600	32.8
10^{1}	1000	29.1
10^{1}	1400	26.6
$10^{1.5}$	200	39.5
$10^{1.5}$	600	26.8
$10^{1.5}$	1000	23.7
$10^{1.5}$	**1400**	**19.8**
10^{2}	200	32.7
10^{2}	600	26.4
10^{2}	1000	27.1
10^{2}	1400	22.6
$10^{2.5}$	200	28.3
$10^{2.5}$	600	32.8
$10^{2.5}$	1000	39.2
$10^{2.5}$	1400	34.1
10^{3}	200	29.6
10^{3}	600	47.3
10^{3}	1000	60.8
10^{3}	1400	57.2
$10^{3.5}$	200	36.2
$10^{3.5}$	600	67.4
$10^{3.5}$	1000	88.1
$10^{3.5}$	1400	90.3
10^{4}	200	48.1
10^{4}	600	94.4
10^{4}	1000	123.2
10^{4}	1400	138.6



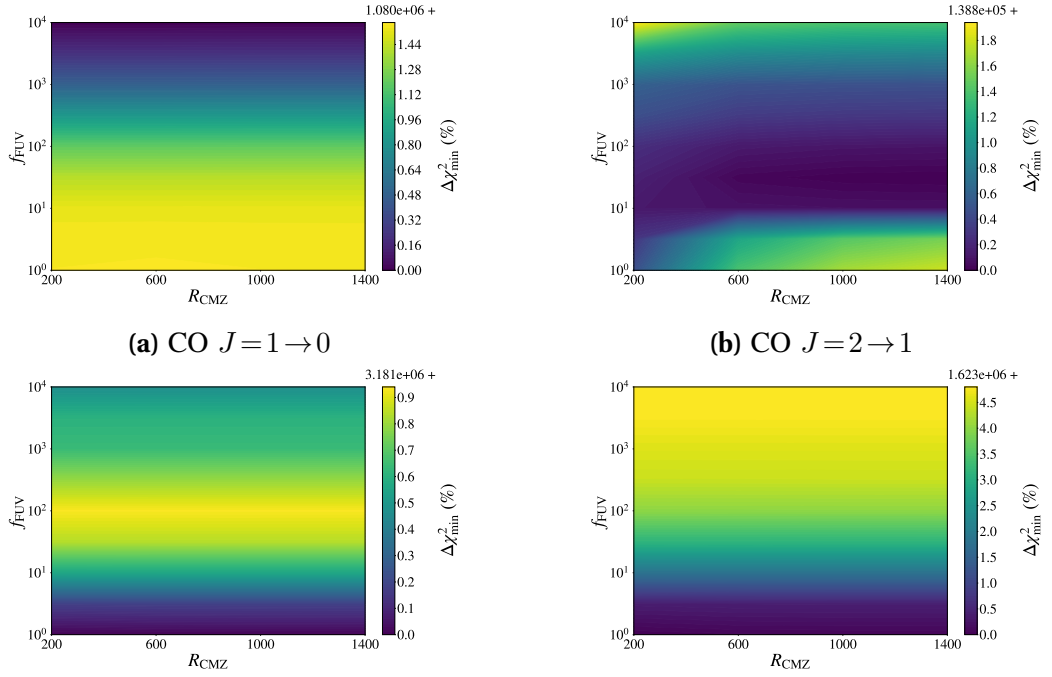
(a) $J=1\rightarrow0$ **(b)** $J=2\rightarrow1$

(c) $J=6\rightarrow5$ **(d)** $J=8\rightarrow7$

Figure VI–45: The same as Figure VI–11 for the `CMZ-radius-radiation` model grid.

7 Cosmic ray ionisation rate

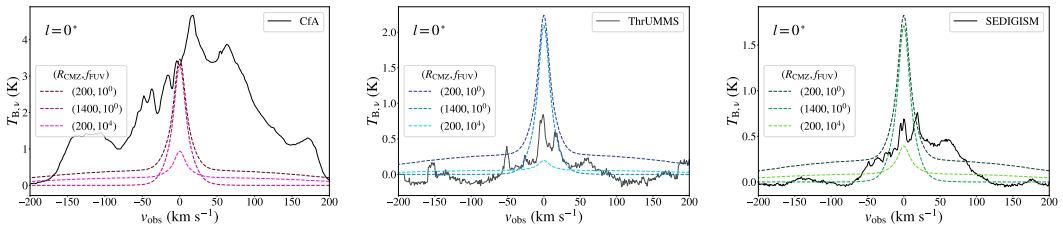
In the grid varying the cosmic ray ionisation rate in the CMZ, `CMZ-radius-radiation`, it is expected that the increased cosmic ray ionisation rate ζ_H in the CMZ decreases the amount of CO, increasing the abundance of ^{12}C and $^{12}C^+$. The increased CRIR also heats the gas, so there are two opposing processes affecting the CO emission. In the qualitative comparison of the synthetic observations to the *COBE*-FIRAS observations (see Figure VI–45), it is evident that the increased CRIR actually removes the intensity peak at $l=0°$. This is in disagreement with the observed profiles. The far-UV radiation has an effect similar to that in §VI-4, where the low-J CO lines are underestimated and the high-J CO lines are overestimated. The variation of these profiles in the parameter space is not large, though, since we do not see much variation in the test statistic when comparing to the *COBE*-FIRAS observations (see Figure VI–46). The best we can constrain is that $1\lesssim f_{FUV}\lesssim10$.



(a) CO $J=1\rightarrow0$ **(b)** CO $J=2\rightarrow1$

(c) CO $J=6\rightarrow5$ **(d)** CO $J=8\rightarrow7$

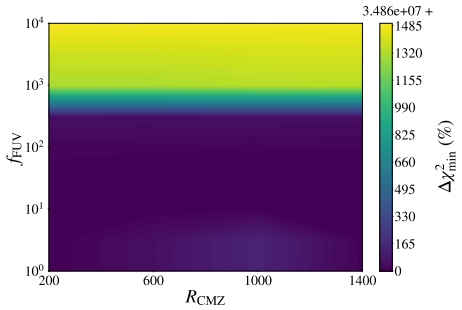
Figure VI–46: The comparison between the CO transitions observed by *COBE*-FIRAS and the `CMZ-radius-radiation` model grid, neglecting extraneous values of χ^2.



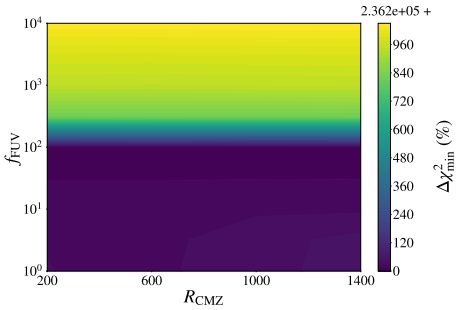
(a) $l=0°$ **(b)** $l=0°$ **(c)** $l=0°$

Figure VI–47: The same as Figure VI–13 to Figure VI–15 for the `CMZ-radius-radiation` model grid.

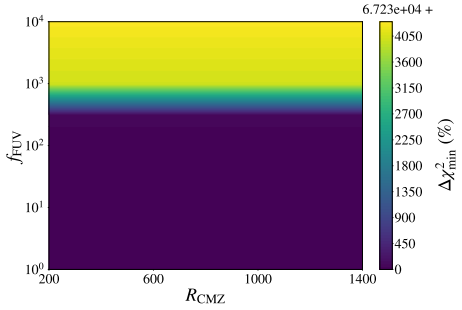
We can qualitatively examine the comparison of the CO spectroscopic surveys to models spanning the parameter space at $l=0°$ (see Figure VI–47). While all of the models underestimate the intensity in subfigure **(a)**, the model with increased CRIR overestimates the wing emission (at $|v_{\rm obs}|>50\,{\rm km\,s^{-1}}$) in subfigures **(b)** and **(c)**. The models with $f_{\rm FUV}=1$ overestimates the bulk ^{13}CO intensity, compared to the model with higher $f_{\rm FUV}$ which is fitting better to the observed spectra. The quantitative comparisons in Figure VI–48 constrain the far-UV radiation factor to $f_{\rm FUV}\lesssim10^{1.5}$, and there is a weak constraint to $R_{\rm CMZ}\approx200\,{\rm pc}$.



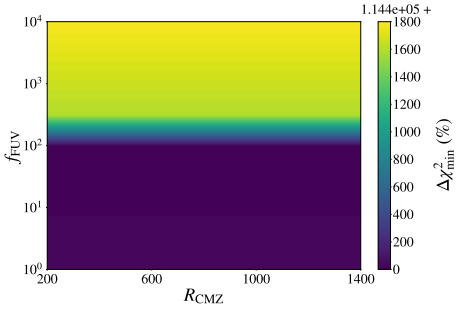
(a) CfA ^{12}CO $J=1\rightarrow0$



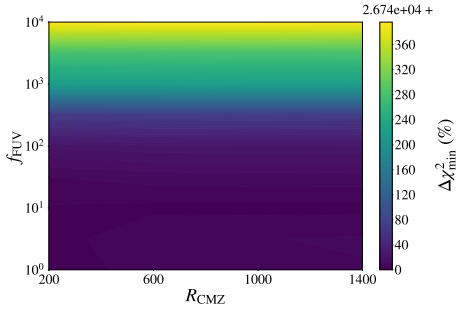
(b) SEDIGISM ^{13}CO $J=1\rightarrow0$



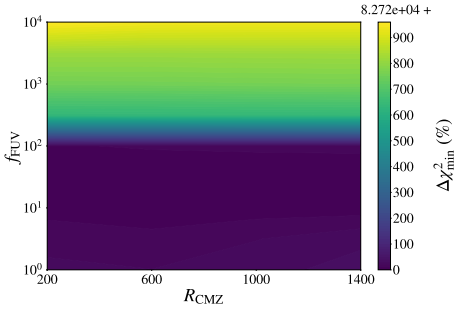
(c) Mopra ^{12}CO $J=1\rightarrow0$



(d) Mopra ^{13}CO $J=1\rightarrow0$

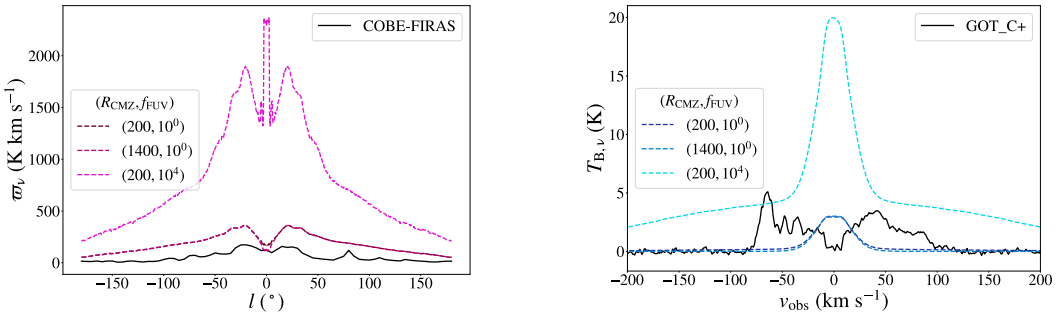


(e) ThrUMMS ^{12}CO $J=1\rightarrow0$



(f) ThrUMMS ^{13}CO $J=1\rightarrow0$

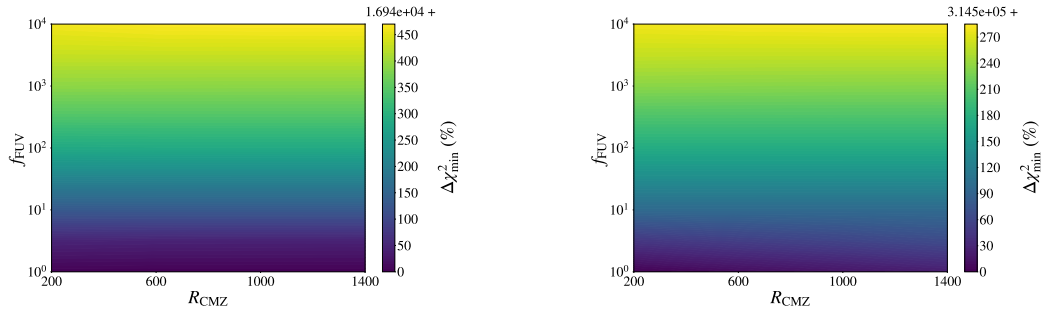
Figure VI–48: The comparison between the spectroscopic CO observations and the synthetic observations of the `CMZ-radius-radiation` model grid for the transitions ^{12}CO $J=1\rightarrow0$, ^{13}CO $J=1\rightarrow0$, and ^{13}CO $J=2\rightarrow1$.



(a) *COBE*-FIRAS **(b)** GOT C+

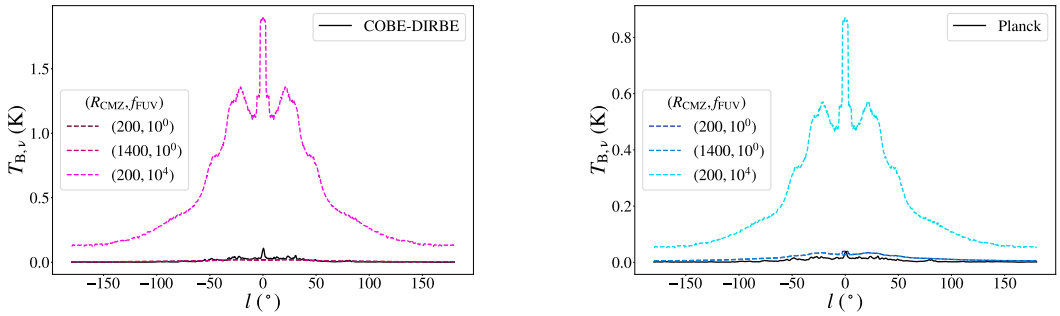
Figure VI–49: The same as Figure VI–17 for the `CMZ-radius-radiation` model grid.

Since the CRIR should create more $^{12}C^+$, the comparison to the *COBE*-FIRAS and GOT C+ observations should give a strong constraint to r_{CMZ}. By examining the qualitative comparison in Figure VI–49, however, the spectra are mainly affected by the far-UV radiation. We can at least see in subfigure **(a)** that the increased CRIR makes the self-absorption at $l = 0$ more pronounced, but the intensity profile id overestimated. In subfigure **(b)** there are clear signs of absorption effects in the synthetic spectra from models with $f_{FUV} = 1$, but increased far-UV radiation does not make it more pronounced. The comparison of the [^{12}C II] 158 μm emission to the synthetic observations of the `CMZ-radius-radiation` grid show similar results to that of the CO comparison. There is a large variance in the test statistic with $\Delta\chi^2_{min} > 10^2$ compared to $\chi^2_{min} = 1.694 \times 10^4,\ 3.145 \times 10^5$ (see Figure VI–50), so there is a strong constraint that $f_{FUV} \sim 1$ and $R_{CMZ} \lesssim 600$ pc.



(a) *COBE*-FIRAS **(b)** GOT C+

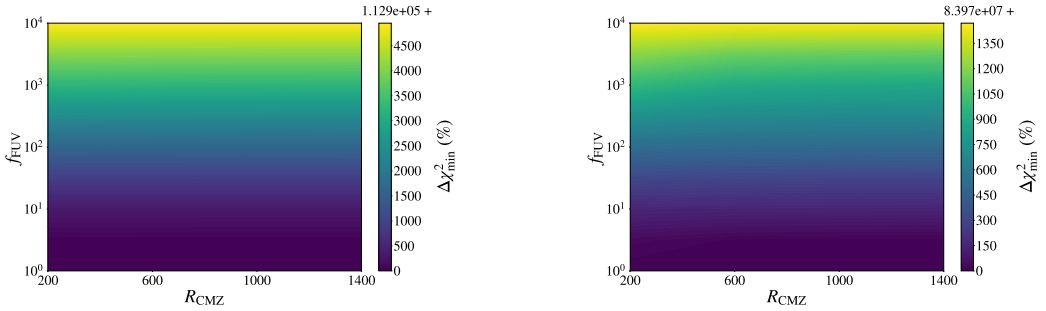
Figure VI–50: The comparison between the observations to the `CMZ-radius-radiation kosmatau3d` grid for the transition [^{12}C II] 158 μm.



(a) *COBE*-DIRBE 240 μm **(b)** *Planck* 550 μm

Figure VI–51: The same as Figure VI–19 for the `CMZ-radius-radiation` model grid.

Dust does not have much CRIR-induced chemistry, so it will mainly be heated. Since the dust continuum is sensitive to far-UV radiation, it is not expected that we can constrain the CMZ radius R_{CMZ} with this comparison. This is seen in Figure VI–51, where the variation in the continuum profiles is only affected by the far-UV factor f_{FUV}. This is confirmed in the quantitative comparison (see Figure VI–52), where there is a strong constraint that $1 < f_{\mathrm{FUV}} \lesssim 10^{0.5}$ mostly independent of R_{CMZ}.



(a) *COBE*-DIRBE 240 μm **(b)** *Planck* 550 μm

Figure VI–52: The comparison between the dust continuum observations to the `CMZ-radius-radiation kosmatau3d` grid at the specified wavelengths.

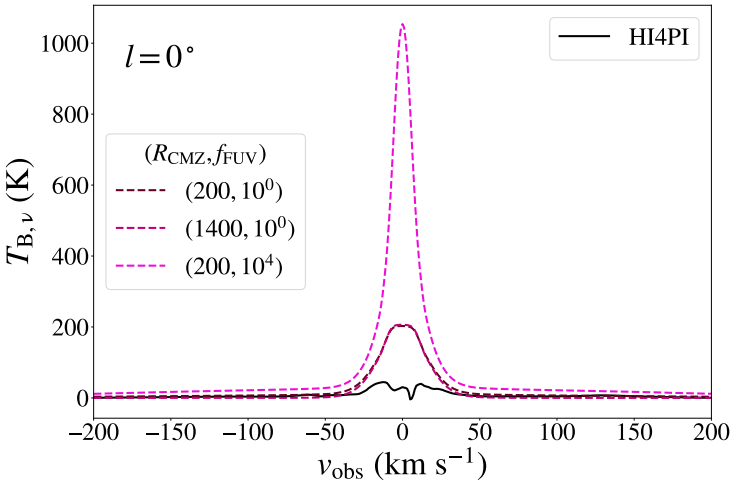


Figure VI–53: The same as Figure VI–21 for the `CMZ-radius-radiation` model grid.

Similar to the continuum comparison, the comparison to the H I 21 cm observations do not show any variation with CRIR (see Figure VI–49 and Figure VI–50). all variation in the plot is due to the far-UV factor f_{FUV}, so the results are similar to that of §VI-4.

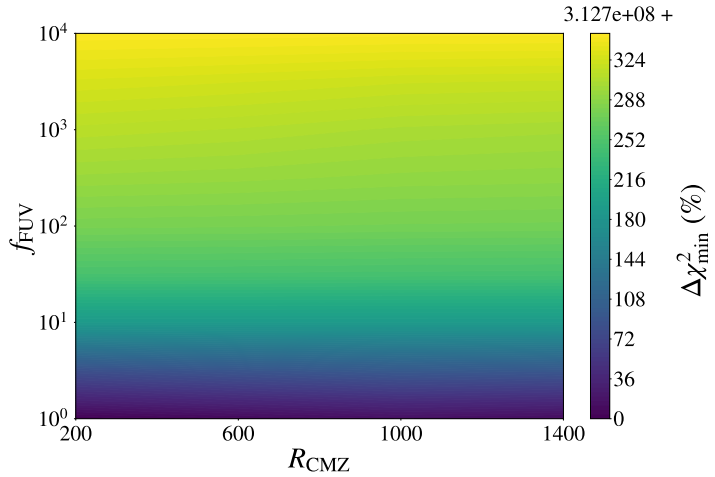
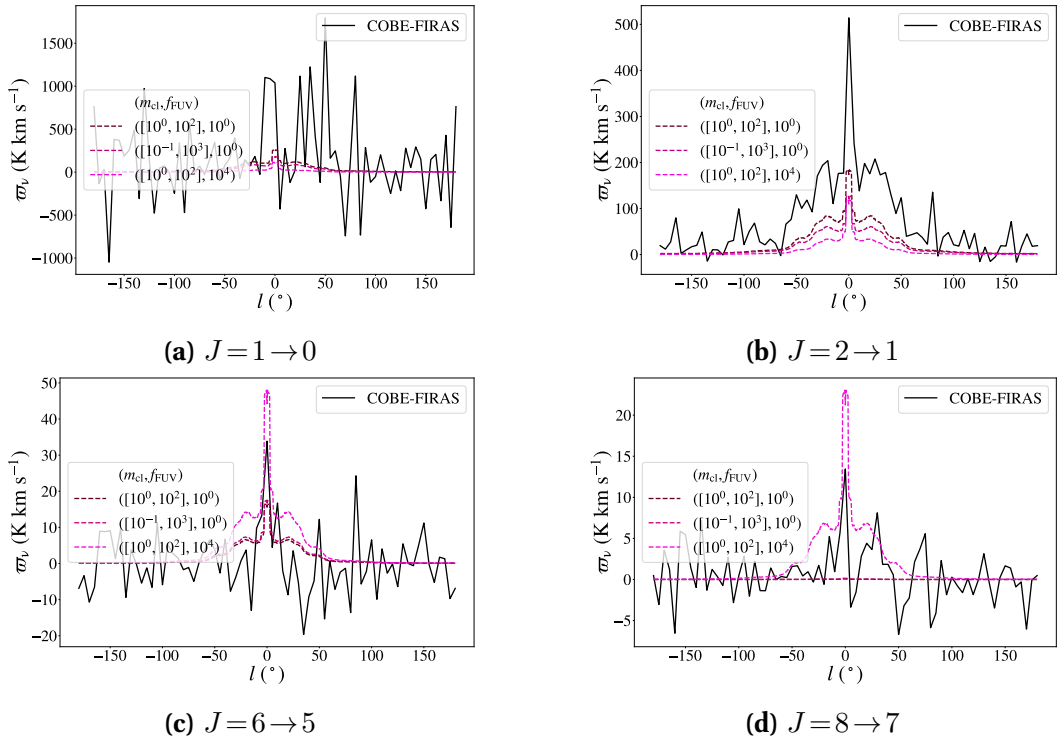


Figure VI–54: The comparison between the H I observations to the `CMZ-radius-radiation kosmatau3d` grid for the transition H I 21 cm.

The $\Delta\chi^2_{min}$ results are combined for the various comparisons (except for the *Planck* comparison; see §VII-7 for the discussion of why it is neglected) and summarised in VI–5.

Table VI–5: The $\Delta\chi^2_{\rm min}$ results for the `CMZ-radius-radiation kosmatau3d` model grid, which varies $R_{\rm CMZ}$ and $f_{\rm FUV}$ The minimum χ^2 is 3.583×10^8.

$R_{\rm CMZ}$	$f_{\rm FUV}$	$\Delta\chi^2_{\rm min}$
200	**1**	**157.2**
200	$10^{0.5}$	305.7
200	10^{1}	823.6
200	$10^{1.5}$	1516.5
200	10^{2}	2369.4
200	$10^{2.5}$	8561.3
200	10^{3}	11217.8
200	$10^{3.5}$	12533.3
200	10^{4}	14059.9
600	1	222.1
600	$10^{0.5}$	368.1
600	10^{1}	820.0
600	$10^{1.5}$	1509.8
600	10^{2}	2372.9
600	$10^{2.5}$	8556.1
600	10^{3}	11205.8
600	$10^{3.5}$	12519.4
600	10^{4}	14022.4
1000	1	347.2
1000	$10^{0.5}$	492.3
1000	10^{1}	818.6
1000	$10^{1.5}$	1507.8
1000	10^{2}	2374.5
1000	$10^{2.5}$	8555.1
1000	10^{3}	11223.7
1000	$10^{3.5}$	12530.8
1000	10^{4}	14019.2
1400	1	244.0
1400	$10^{0.5}$	391.6
1400	10^{1}	818.1
1400	$10^{1.5}$	1506.7
1400	10^{2}	2374.9
1400	$10^{2.5}$	8555.3
1400	10^{3}	11228.6
1400	$10^{3.5}$	12533.0
1400	10^{4}	14016.3



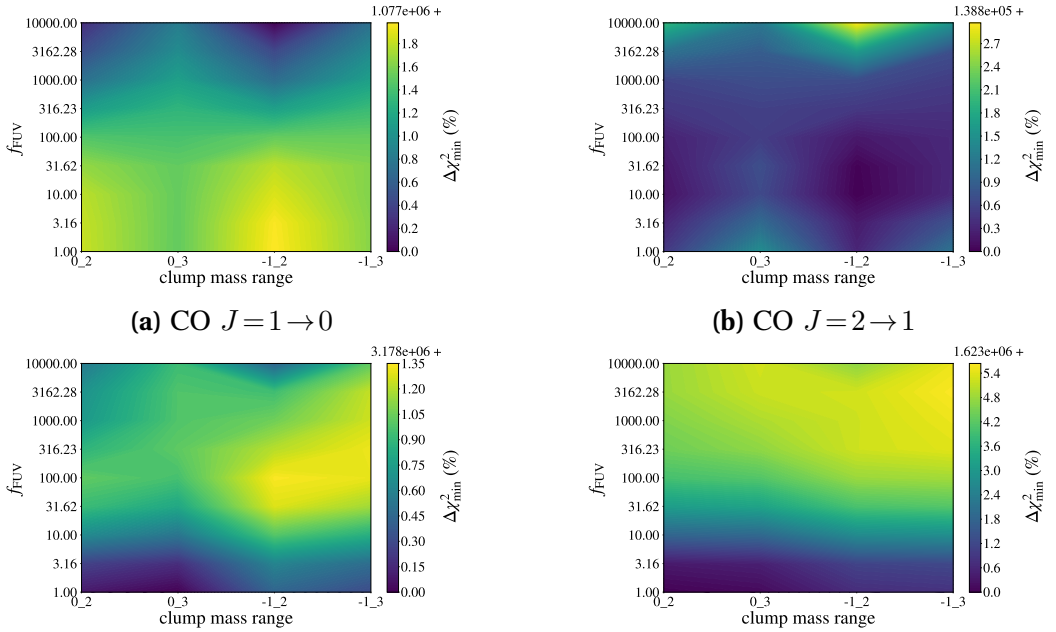
(a) $J = 1 \rightarrow 0$ **(b)** $J = 2 \rightarrow 1$

(c) $J = 6 \rightarrow 5$ **(d)** $J = 8 \rightarrow 7$

Figure VI–55: The same as Figure VI–11 for the H^0 mass in the `clump-masses` model grid.

8 Clump masses

The parameters varied in the `clump-masses` model grid, $m_{\rm cl}$ and f_subFUV, will change the optical depth of a given voxel (see the discussion in §III-7). In the qualitative comparison of the CO $J = 1 \rightarrow 0$ from models spanning the parameter space to the *COBE*-FIRAS observations, it appears that using a wider range of clump masses slightly decreases the intensity for the low-J CO lines (see Figure VI–55). Subfigure **(c)** indicates the reverse trend for CO $J = 6 \rightarrow 5$, and the far-UV factor has a similar constraint to previous grids. The quantitative comparison of the grid to the observations does not show much variation in the test statistic, since the maximum variation is $\max(\Delta\chi^2) \sim 1\%$ Thus the results of this test are inconclusive.



(a) CO $J = 1 \rightarrow 0$

(b) CO $J = 2 \rightarrow 1$

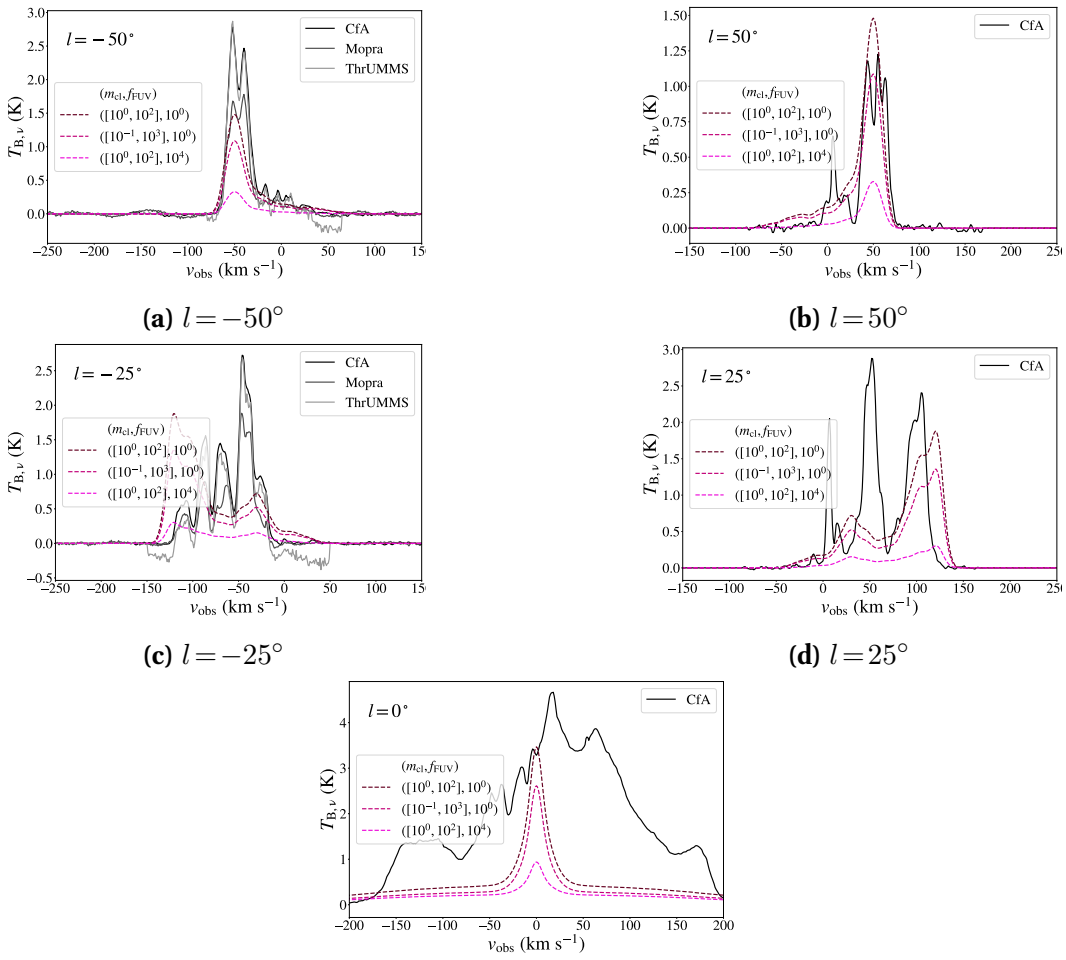
(c) CO $J = 6 \rightarrow 5$

(d) CO $J = 8 \rightarrow 7$

Figure VI–56: The comparison between the CO transitions observed by *COBE*-FIRAS and the `clump-masses` model grid, neglecting extraneous values of χ^2.

We can qualitatively examine spectra for the same models (see Figure VI–57 to Figure VI–59). Beyond the trend of f_{FUV} velocity dispersion that was discussed in previous sections, increasing the mass clump range to $m_{\mathrm{cl}} = [10^{-1}, 10^{3}]\,\mathrm{M}_{\odot}$ reduces the intensity in the spectra. This makes it fit better to Figure VI–57 subfigure **(b)**, Figure VI–58 subfigure **(a)**, and Figure VI–59 subfigure **(a)**, but the model with mass clump range $m_{\mathrm{cl}} = [10^{0}, 10^{2}]\,\mathrm{M}_{\odot}$ fits better to the spectrum in Figure VI–57 subfigure **(a)**. It is difficult to gain much information from the other spectra (with $20^{\circ} < |l| < 30^{\circ}$) since there is not much agreement between the synthetic and observed intensities.

The lack of variation is not seen in the comparison to the spectroscopic surveys (see Figure VI–60), where the maximum variation of the test statistic is $\Delta\chi^2_{\mathrm{min}} \sim 100\%$ for low f_{FUV}. The trend seems to indicate that the models with $m_{\mathrm{cl}} = ([10^{0}, 10^{2}]\,\mathrm{M}_{\odot}$ and $m_{\mathrm{cl}} = ([10^{-1}, 10^{2}]\,\mathrm{M}_{\odot}$ fit better than the models with $10^{3}\,\mathrm{M}_{\odot}$ as the most-massive clump.



(a) $l = -50°$

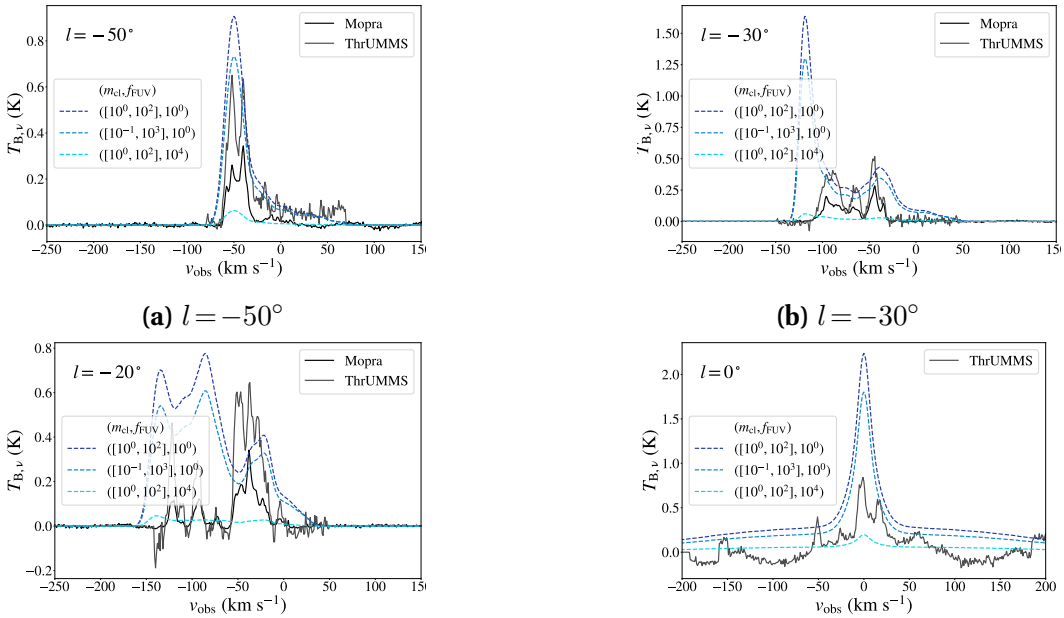
(b) $l = 50°$

(c) $l = -25°$

(d) $l = 25°$

(e) $l = 0°$

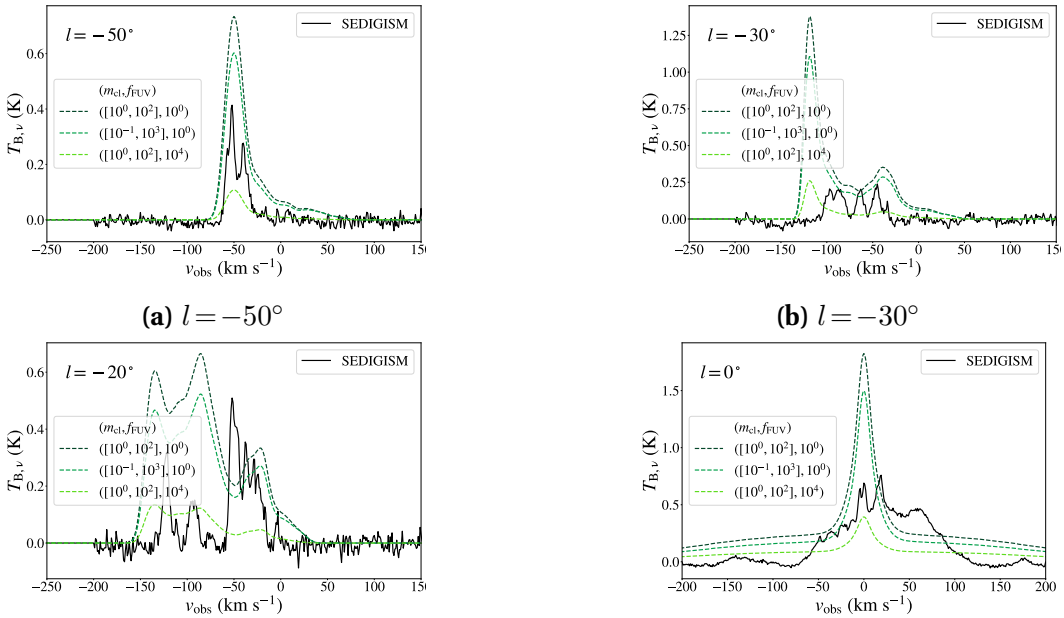
Figure VI–57: The same as Figure VI–13 for the `clump-masses` model grid.



(a) $l = -50°$ **(b)** $l = -30°$

(c) $l = -20°$ **(d)** $l = 0°$

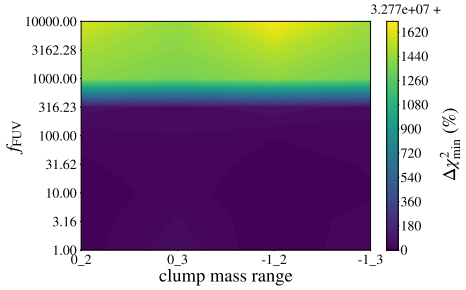
Figure VI–58: The same as Figure VI–14 for the `clump-masses` model grid.



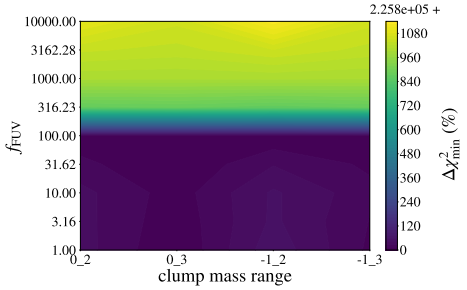
(a) $l = -50°$ **(b)** $l = -30°$

(c) $l = -20°$ **(d)** $l = 0°$

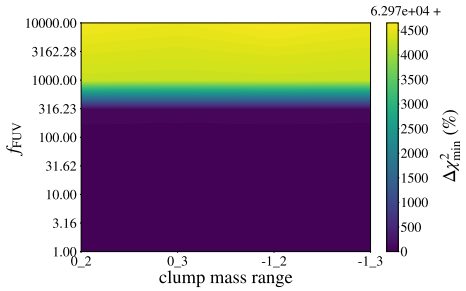
Figure VI–59: The same as Figure VI–13 to Figure VI–15 for the `clump-masses` model grid.



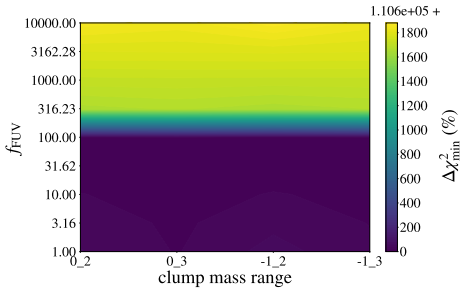
(a) CfA ^{12}CO $J=1\rightarrow0$



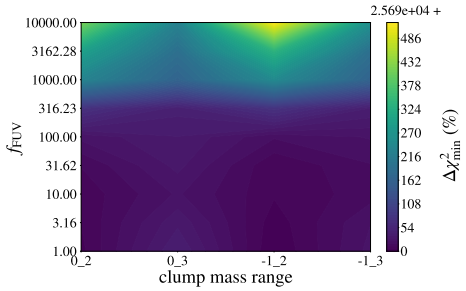
(b) SEDIGISM ^{13}CO $J=2\rightarrow1$



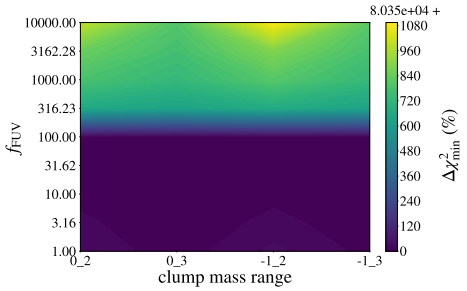
(c) Mopra ^{12}CO $J=1\rightarrow0$



(d) Mopra ^{13}CO $J=1\rightarrow0$



(e) ThrUMMS ^{12}CO $J=1\rightarrow0$



(f) ThrUMMS ^{13}CO $J=1\rightarrow0$

Figure VI–60: The comparison between the spectroscopic CO observations and the synthetic observations of the `clump-masses` model grid for the transitions ^{12}CO $J=1\rightarrow0$, ^{13}CO $J=1\rightarrow0$, and ^{13}CO $J=2\rightarrow1$.

Since the `clump-masses` model grid does not show any variation variation in the qualitative comparison to the [^{12}C II], continuum, or H I observations with respect to the clump masses $m_{\rm cl}$, we just summarise the quantitative comparison in VI–6.

The $\Delta\chi^2_{\rm min}$ results are combined for the various comparisons (except for the *Planck* comparison; see §VII-7 for the discussion of why it is neglected) and summarised in VI–6.

Table VI–6: The $\Delta\chi^2_{\rm min}$ results for the `clump-masses kosmatau3d` model grid, which varies $m_{\rm cl}$ and $f_{\rm FUV}$ The minimum χ^2 is 3.561×10^8.

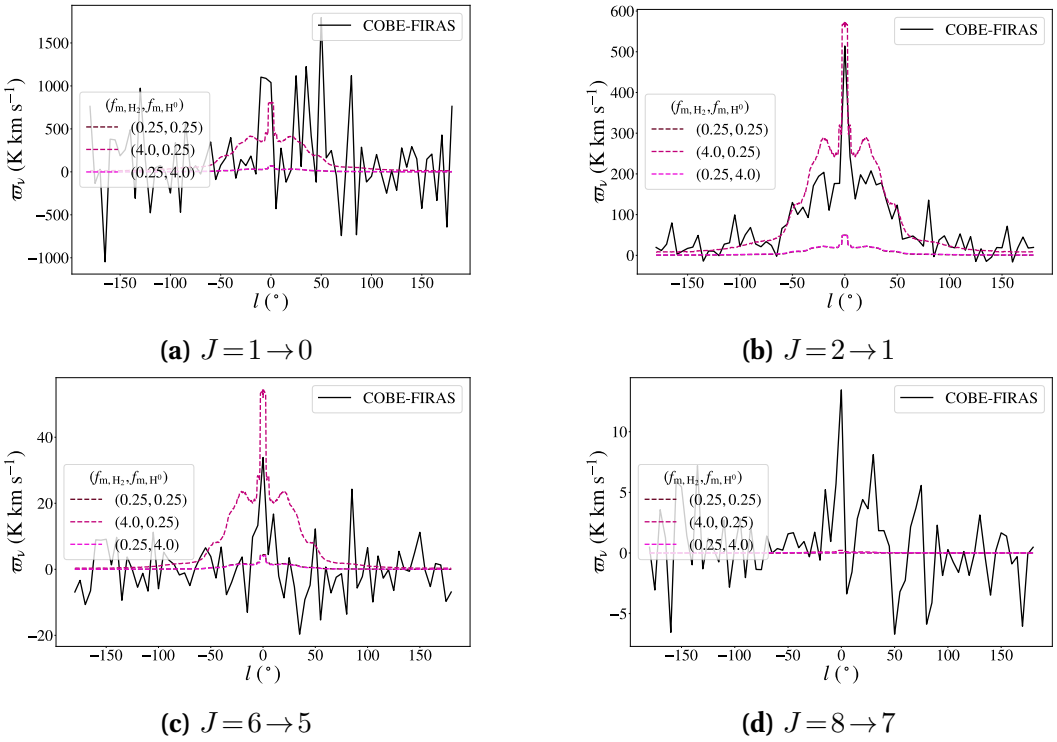
$m_{\rm cl}$	$f_{\rm FUV}$	$\Delta\chi^2_{\rm min}$
$[10^0, 10^2]$	1	188.0
$[10^0, 10^2]$	$10^{0.5}$	335.3
$[10^0, 10^2]$	10^1	856.5
$[10^0, 10^2]$	$10^{1.5}$	1556.5
$[10^0, 10^2]$	10^2	2420.6
$[10^0, 10^2]$	$10^{2.5}$	8811.9
$[10^0, 10^2]$	10^3	11578.3
$[10^0, 10^2]$	$10^{3.5}$	12922.3
$[10^0, 10^2]$	10^4	14485.1
$[10^0, 10^3]$	**1**	**116.4**
$[10^0, 10^3]$	$10^{0.5}$	276.3
$[10^0, 10^3]$	10^1	794.0
$[10^0, 10^3]$	$10^{1.5}$	1495.8
$[10^0, 10^3]$	10^2	2346.3
$[10^0, 10^3]$	$10^{2.5}$	8647.9
$[10^0, 10^3]$	10^3	11264.8
$[10^0, 10^3]$	$10^{3.5}$	12487.0
$[10^0, 10^3]$	10^4	13660.1
$[10^{-1}, 10^2]$	1	224.9
$[10^{-1}, 10^2]$	$10^{0.5}$	354.0
$[10^{-1}, 10^2]$	10^1	878.0
$[10^{-1}, 10^2]$	$10^{1.5}$	1578.1
$[10^{-1}, 10^2]$	10^2	2439.8
$[10^{-1}, 10^2]$	$10^{2.5}$	8867.3
$[10^{-1}, 10^2]$	10^3	11691.8
$[10^{-1}, 10^2]$	$10^{3.5}$	13133.9
$[10^{-1}, 10^2]$	10^4	14921.7
$[10^{-1}, 10^3]$	1	130.6
$[10^{-1}, 10^3]$	$10^{0.5}$	288.0
$[10^{-1}, 10^3]$	10^1	811.2
$[10^{-1}, 10^3]$	$10^{1.5}$	1511.0

Continued on next page

Table VI–6: The $\Delta\chi^2_{\text{min}}$ results for the `clump-masses kosmatau3d` model grid, which varies m_{cl} and f_{FUV} The minimum χ^2 is 3.561×10^8.

m_{cl}	f_{FUV}	$\Delta\chi^2_{\text{min}}$
$[10^{-1}, 10^3]$	10^2	2358.0
$[10^{-1}, 10^3]$	$10^{2.5}$	8669.5
$[10^{-1}, 10^3]$	10^3	11304.0
$[10^{-1}, 10^3]$	$10^{3.5}$	12544.1
$[10^{-1}, 10^3]$	10^4	13772.1

9 Mass factor



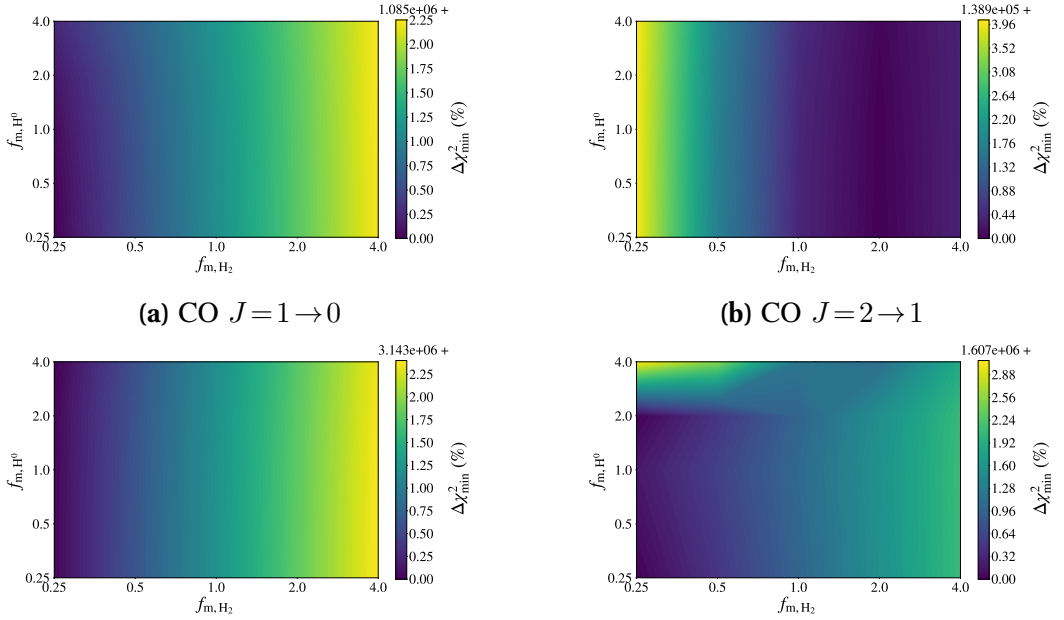
(a) $J = 1 \rightarrow 0$ **(b)** $J = 2 \rightarrow 1$

(c) $J = 6 \rightarrow 5$ **(d)** $J = 8 \rightarrow 7$

Figure VI–61: The same as Figure VI–11 for the `mass-factors` model grid.

The `mass-factors` model grid scales the galactic molecular and atomic mass profiles (see §IV-1.1), which is used for the mass in the clumpy and interclump ensembles, respectively. Thus the comparison to observations should be able to constrain either factor depending on whether the observation traces the atomic or molecular ISM. The qualitative comparison of the `mass-factors` grid with the *COBE*-FIRAS observations shows that the CO integrated intensity profiles are most sensitive to the molecular hydrogen mass M_{H_2}, but not the atomic hydrogen mass M_{H^0}. Even though the comparison is sensitive to changes in the molecular mass,

there is not much variation in $\Delta\chi^2_{\text{min}}$ to provide a strong constraint (see Figure VI–62). Most of the transitions show a better fit to $f_{\text{cm}} \lesssim 1$, with no clear constraint on f_{icm}. There is slightly more variation in the test statistic for the $J = 8 \rightarrow 7$ transition, giving the slight constraint of $f_{\text{icm}} \lesssim 2$.

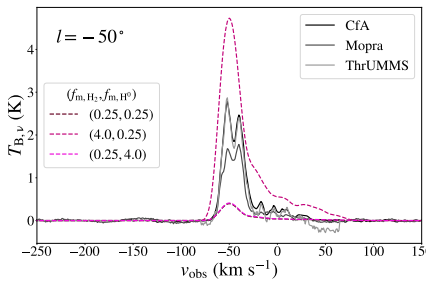


(a) CO $J = 1 \rightarrow 0$ **(b)** CO $J = 2 \rightarrow 1$

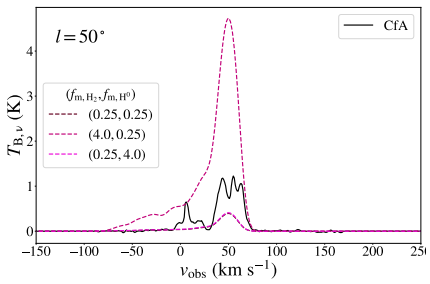
(c) CO $J = 6 \rightarrow 5$ **(d)** CO $J = 8 \rightarrow 7$

Figure VI–62: The comparison between the CO transitions observed by *COBE*-FIRAS and the `mass-factors` model grid, neglecting extraneous values of χ^2.

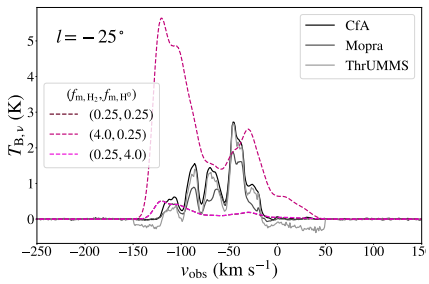
The qualitative comparison of the grid to the spectroscopic CO surveys (Figure VI–63 to Figure VI–63) again shows that the intensity of the spectra scale proportionally with $f_{\text{m,H}_2}$. Now the quantitative comparison exhibits more variation in the test statistic since it reaches a maximum of $\Delta\chi^2_{\text{min}} > 100\%$, and thus stronger constraints on $f_{\text{m,H}_2}$. None of the transitions can be used to constrain f_{icm}, but there is a strong constraint that $1 < f_{\text{cm}} < 2$. Ultimately this indicates that most of the CO emission is, understandably, coming from the denser regions of the clumpy medium.



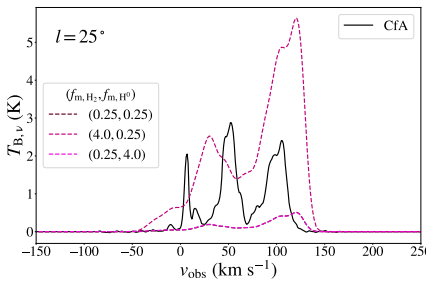
(a) $l = -50°$



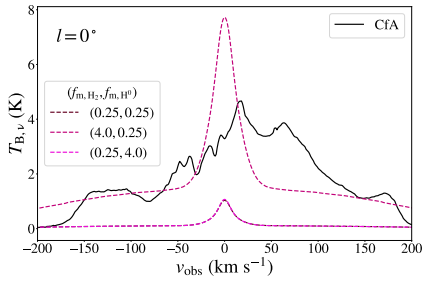
(b) $l = 50°$



(c) $l = -25°$

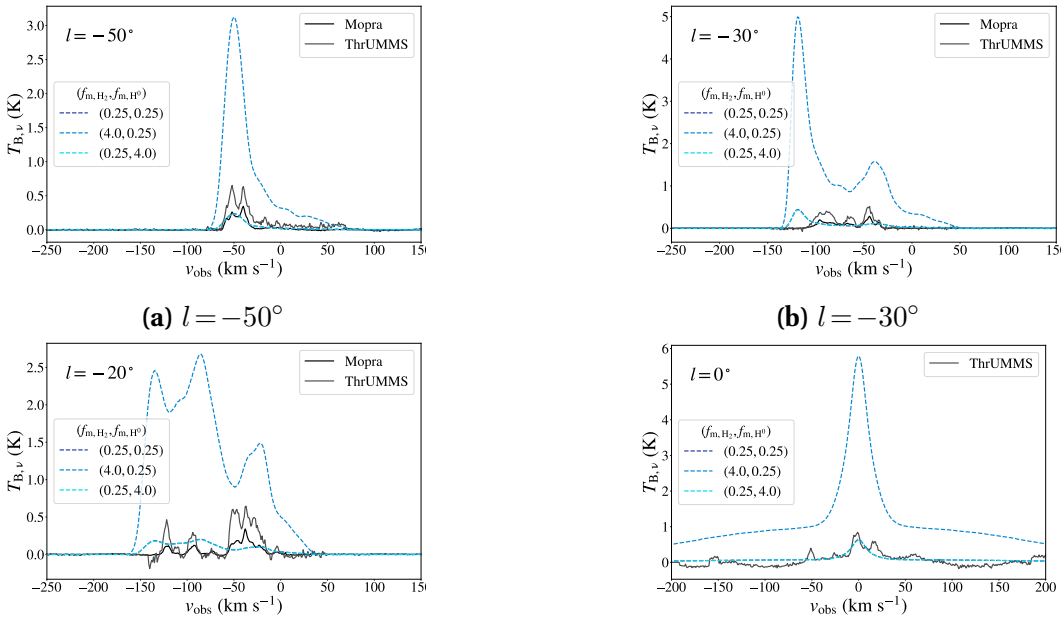


(d) $l = 25°$



(e) $l = 0°$

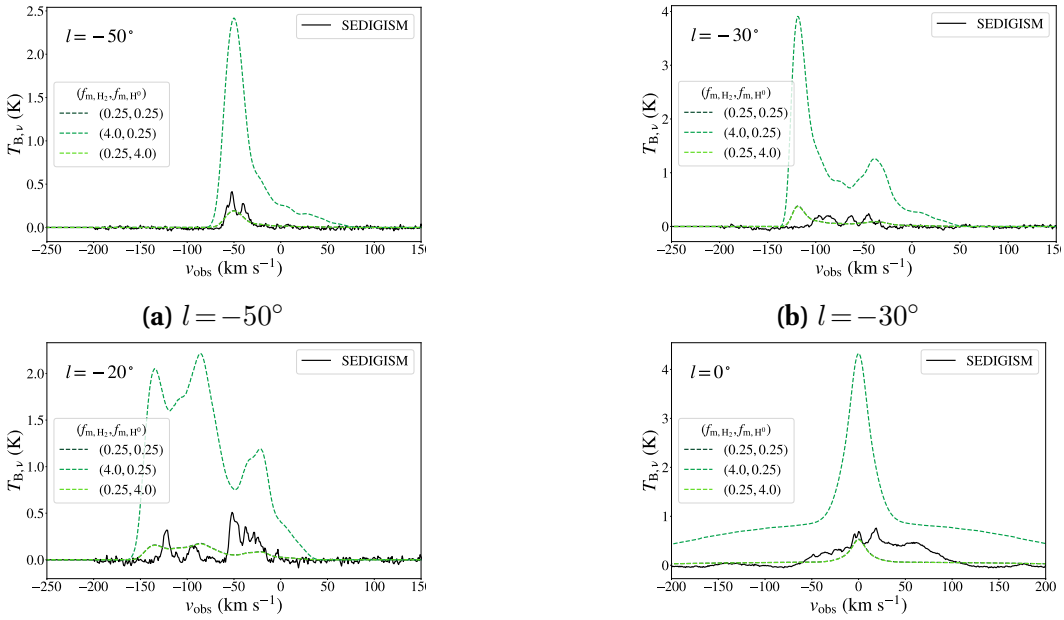
Figure VI–63: The same as Figure VI–13 for the `mass-factors` model grid.



(a) $l = -50°$ **(b)** $l = -30°$

(c) $l = -20°$ **(d)** $l = 0°$

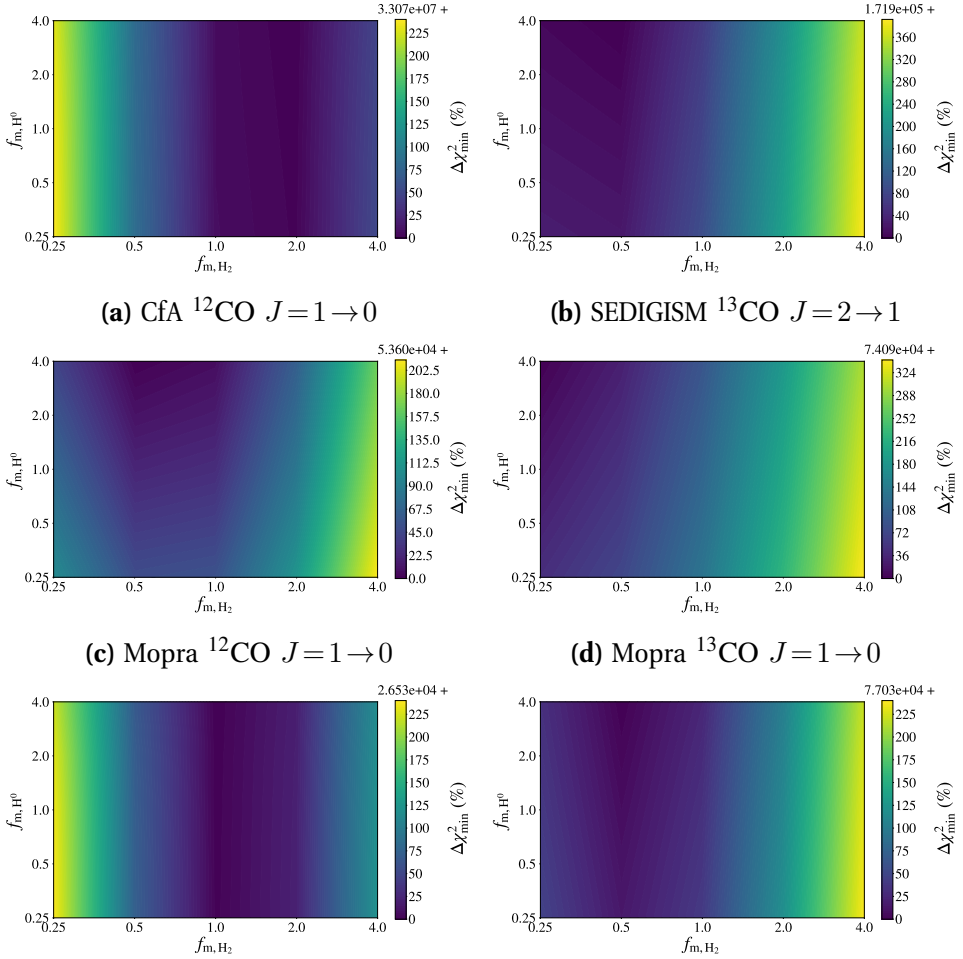
Figure VI–64: The same as Figure VI–14 for the `mass-factors` model grid.



(a) $l = -50°$ **(b)** $l = -30°$

(c) $l = -20°$ **(d)** $l = 0°$

Figure VI–65: The same as Figure VI–15 for the `mass-factors` model grid.

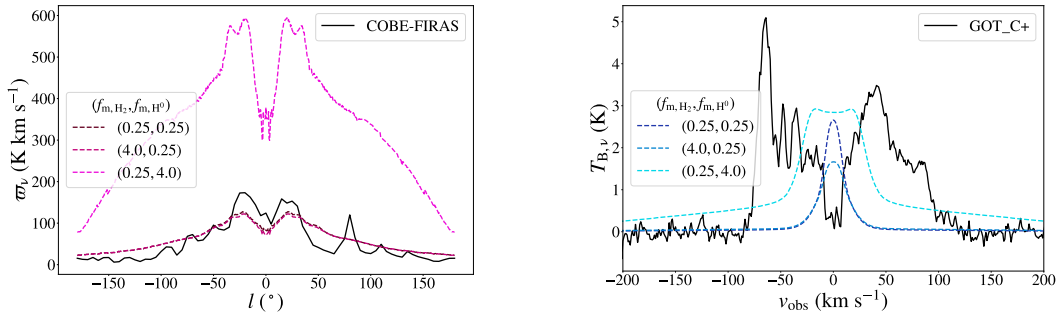


(a) CfA ^{12}CO $J=1\rightarrow0$ **(b)** SEDIGISM ^{13}CO $J=2\rightarrow1$

(c) Mopra ^{12}CO $J=1\rightarrow0$ **(d)** Mopra ^{13}CO $J=1\rightarrow0$

(e) ThrUMMS ^{12}CO $J=1\rightarrow0$ **(f)** ThrUMMS ^{13}CO $J=1\rightarrow0$

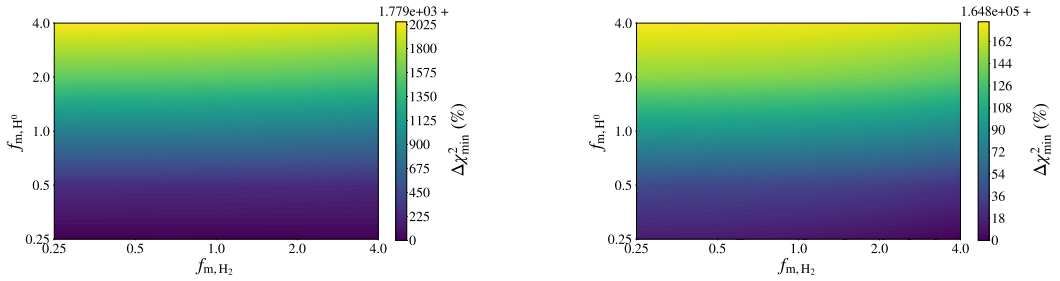
Figure VI–66: The comparison between the spectroscopic CO observations and the synthetic observations of the `mass-factors` model grid for the transitions ^{12}CO $J=1\rightarrow0$, ^{13}CO $J=1\rightarrow0$, and ^{13}CO $J=2\rightarrow1$.



(a) *COBE*-FIRAS **(b)** GOT C+

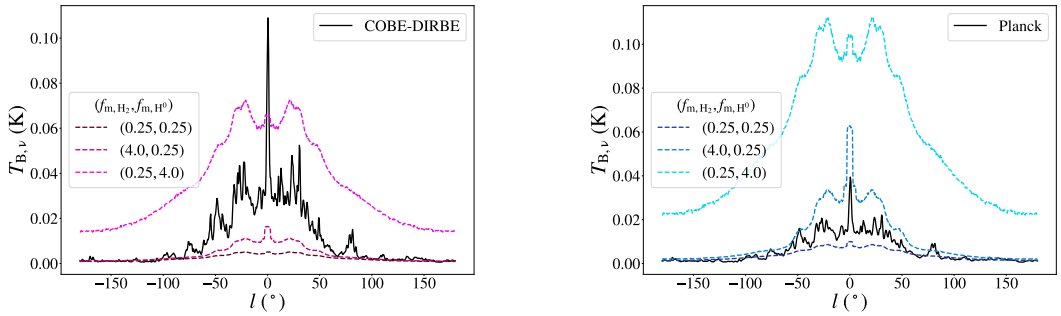
Figure VI–67: The same as Figure VI–17 for the `mass-factors` model grid.

The comparison of the [^{12}C II] observations to models spanning the parameter space show that [^{12}C II] is most sensitive to the atomic mass M_{H^0}. (see Figure VI–67). Here we can see that changing the mass of the interclump ensemble has a significant effect on the closeness of the fit in subfigure **(a)**, although increasing it too much overestimates the synthetic intensity profile. Subfigure **(b)**, however, indicates that the increased atomic mass also increases the amount of self-absorption at $l = 0°$ in the synthetic spectra. None of the models show as much velocity dispersion as the observation, though. The comparison of the `mass-factors` model grid to *COBE*-FIRAS and GOT C+ yields a strong constraint to the parameter space, with the maximum variation of the test statistic $\Delta\chi^2_{\rm min} > 100\%$ with $\chi^2_{\rm min} = 1.779\times10^3,\ 1.648\times10^5$ (see Figure VI–68). This suggests there is a strong constraint for $f_{\rm icm} \approx 0.25$ and a moderate constraint for $f_{\rm cm} \sim 4$.



(a) *COBE*-FIRAS **(b)** GOT C+

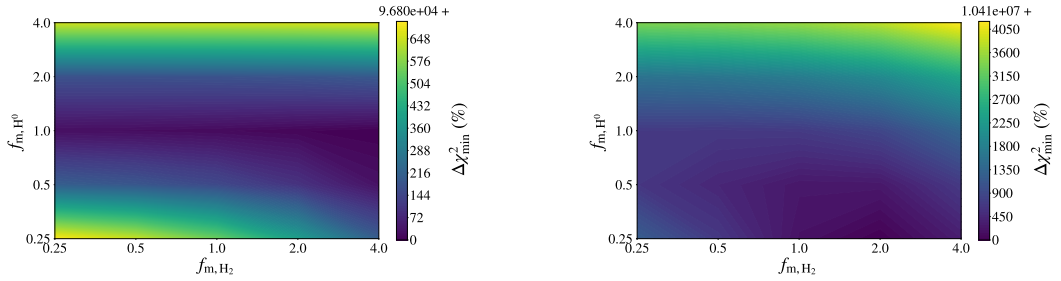
Figure VI–68: The comparison between the observations to the `mass-factors` `kosmatau3d` grid for the transition [^{12}C II] 158 μm.



(a) *COBE*-DIRBE 240 μm **(b)** *Planck* 550 μm

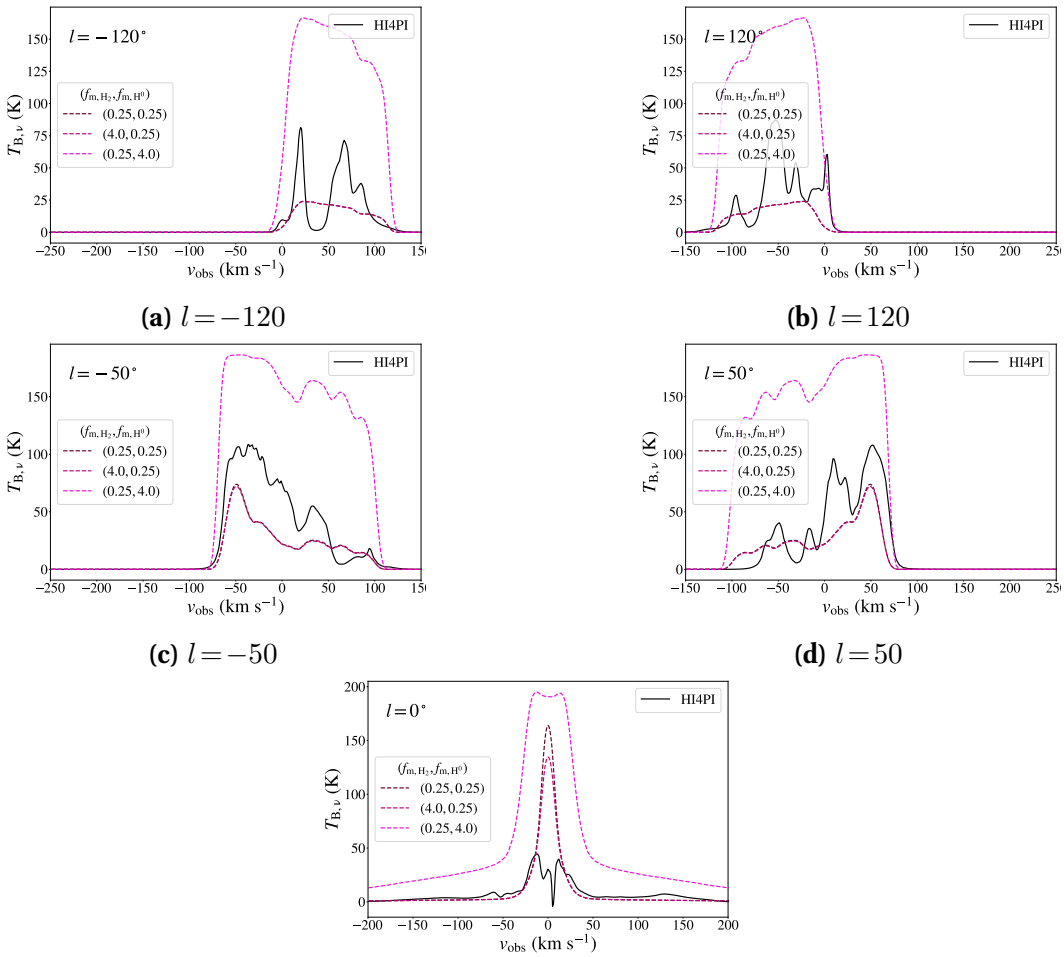
Figure VI–69: The same as Figure VI–19 for the `mass-factors` model grid.

The effect of the ensemble masses on the dust continuum is explored in the intensity profiles in Figure VI–69 for models spanning the parameter space. Both the molecular and atomic masses affect the dust continuum, but it appears there is not a model that can match the profiles for both wavelengths. The test statistics for the comparison of the dust continuum with the model grid exhibits a large amount of variation in the parameter space, since the maximum variation of the test statistic $\Delta\chi^2_{\rm min} \sim 10^3\%$ compared to $\chi^2_{\rm min} = 9.680\times10^4,\ 1.041\times10^7$ (see Figure VI–70). While the *COBE*-DIRBE comparison enforces a strong constraint of $f_{\rm icm} = 1$ and a moderate constraint of $f_{\rm cm} \approx 4$, the *Planck* comparison shows a much stronger constraint of $f_{\rm icm} < 1$ and $f_{\rm cm} \approx 2$. This is rather interesting since strong constraints for the two wavelengths act on different parameters. This can be understood since wavelengths $\lambda \sim 240\,\mu$m are more energetic than wavelengths $\lambda \sim 550\,\mu$m, and thus 240 μm is more sensitive to the warmer interclump ensemble rather than the cold, dense clumpy ensemble. There is anyways an issue with the comparison to the *Planck* data (see §VII-7).



(a) *COBE*-DIRBE 240 μm **(b)** *Planck* 550 μm

Figure VI–70: The comparison between the dust continuum observations to the `mass-factors kosmatau3d` grid at the specified wavelengths.



(a) $l = -120$ **(b)** $l = 120$

(c) $l = -50$ **(d)** $l = 50$

(e) $l = 0$

Figure VI–71: The same as Figure VI–21 for the `mass-factors` model grid.

The H I emission is sensitive to the properties affecting the interclump ensemble, thus we can constrain better the atomic mass factor f_{m,H^0} from the comparison to the `mass-factors` model grid. This is clearly shown in Figure VI–71, where we plot the spectra for the models spanning the parameter space. The synthetic spectra reinforce the assumption that the interclump ensemble properties affect the H I emission, and the and there is a minor effect from the molecular mass factor f_{m,H_2} at $l = 0$. All synthetic spectra are missing the self-absorption features present in the observations. The quantitative comparison of the H I observations to the synthetic observations in the model grid give a strong constraint since the maximum variation of the test statistic is $\Delta\chi^2_{min} > 100\%$ with $\chi^2_{min} = 1.624 \times 10^8$ (see Figure VI–72). This provides the strong constraint $0.5 \lesssim f_{icm} < 1$ for all f_{cm}.

The $\Delta\chi^2_{min}$ results are combined for the various comparisons (except for the *Planck* comparison; see §VII-7 for the discussion of why it is neglected) and summarised in VI–7.

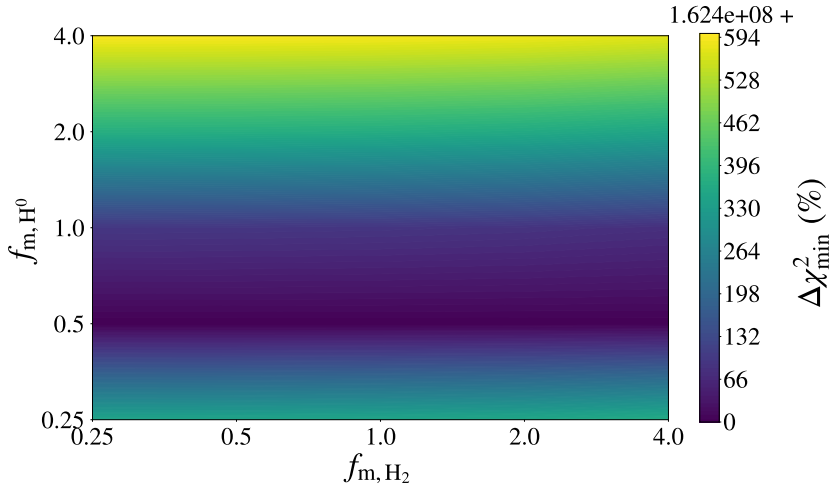


Figure VI–72: The comparison between the H I observations to the `mass-factors kosmatau3d` grid for the transition H I 21 cm.

Table VI–7: The $\Delta\chi^2_{\rm min}$ results for the `mass-factors kosmatau3d` model grid, which varies $f_{\rm m,H_2}$ and $f_{\rm m,H^0}$ The minimum χ^2 is 2.058×10^8.

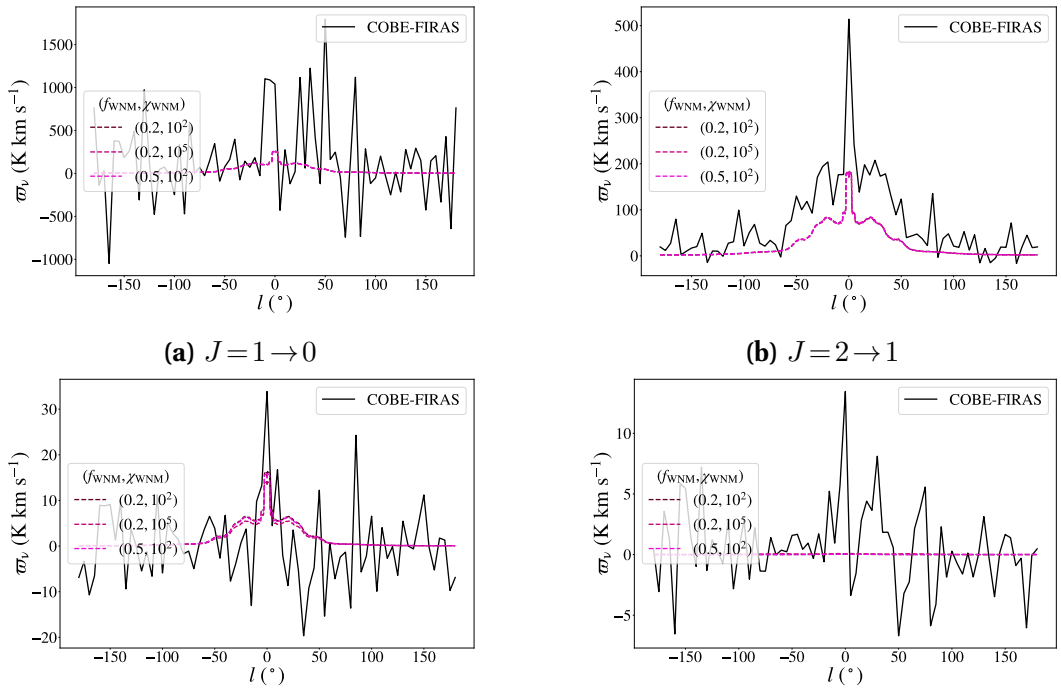
$f_{\rm m,H_2}$	$f_{\rm m,H^0}$	$\Delta\chi^2_{\rm min}$
0.25	0.25	1518.2
0.25	0.5	931.8
0.25	1	1473.1
0.25	2	2589.3
0.25	4	3766.8
0.5	0.25	1285.8
0.5	**0.5**	**738.4**
0.5	1	1293.2
0.5	2	2413.7
0.5	4	3592.9
1	0.25	1350.9
1	0.5	864.0
1	1	1438.2
1	2	2561.5
1	4	3738.9
2	0.25	1670.0
2	0.5	1266.9
2	1	1869.3
2	2	2995.8
2	4	4171.6
4	0.25	2195.3
4	0.5	1880.0

Continued on next page

Table VI–7: The $\Delta\chi^2_{\rm min}$ results for the `mass-factors kosmatau3d` model grid, which varies $f_{\rm m,H_2}$ and $f_{\rm m,H^0}$ The minimum χ^2 is 2.058×10^8.

$f_{\rm m,H_2}$	$f_{\rm m,H^0}$	$\Delta\chi^2_{\rm min}$
4	1	2510.9
4	2	3640.5
4	4	4813.3

10 Warm neutral medium

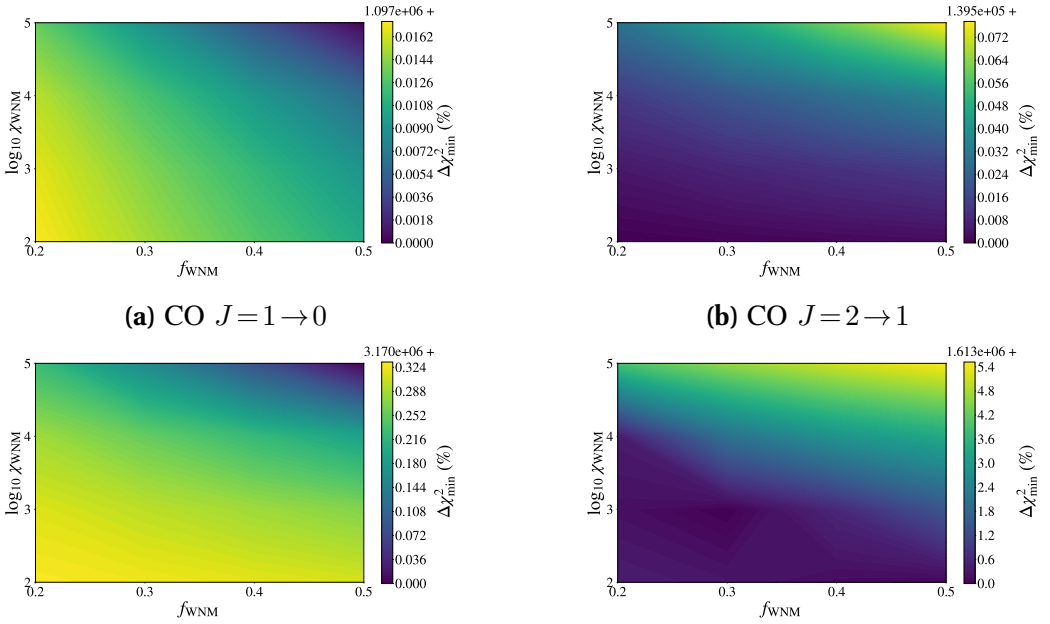


(a) $J = 1 \rightarrow 0$ **(b)** $J = 2 \rightarrow 1$

(c) $J = 6 \rightarrow 5$ **(d)** $J = 8 \rightarrow 7$

Figure VI–73: The same as Figure VI–11 for the `three-media` model grid.

The `three-media kosmatau3d` model grid should be able to better-replicate some of the atomic ISM features, particularly the [^{12}C II] 158 μm and H I 21 cm self-absorption features in the spectrum towards the Galactic Centre since there should be two temperatures of atomic gas in each voxel. Exploring the effect of this parameter space on the comparison to the *COBE*-FIRAS CO profiles (see Figure VI–73), it is difficult to find any constraint. It should come as no surprise since CO traces rather the molecular than the atomic medium. The test statistic of the comparison, shown in Figure VI–74, confirms that the CO integrated intensity profiles do not constrain well the WNM.

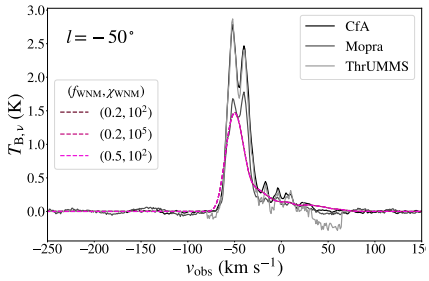


(a) CO $J=1\to0$ **(b)** CO $J=2\to1$

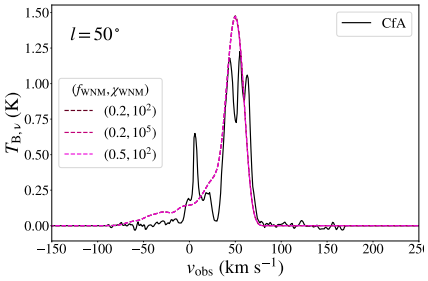
(c) CO $J=6\to5$ **(d)** CO $J=8\to7$

Figure VI–74: The comparison between the CO transitions observed by *COBE*-FIRAS and the `three-media` model grid, neglecting extraneous values of χ^2.

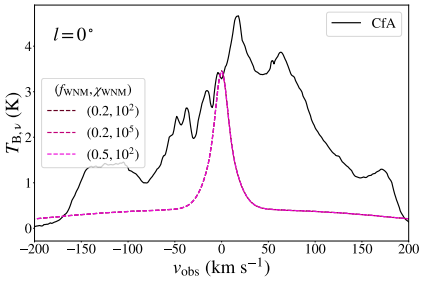
The comparison of the `three-media` model grid to the CO spectroscopic surveys, both qualitatively (Figure VI–75 to Figure VI–77) and quantitatively (Figure VI–78), likewise do not show a clear any clear variation of the spectra when varying the WNM parameters.



(a) $l = -50°$

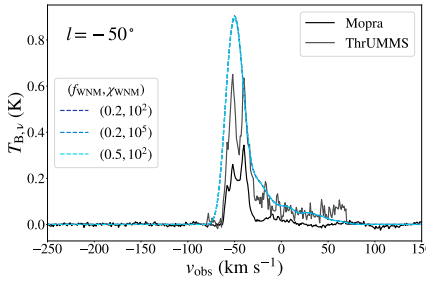


(b) $l = 50°$

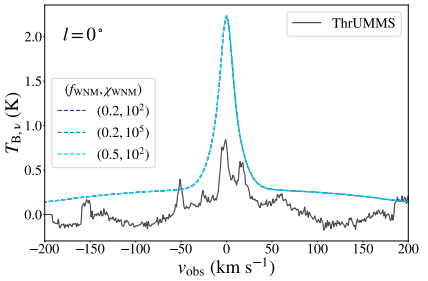


(c) $l = 0°$

Figure VI–75: The same as Figure VI–13 for the `three-media` model grid.

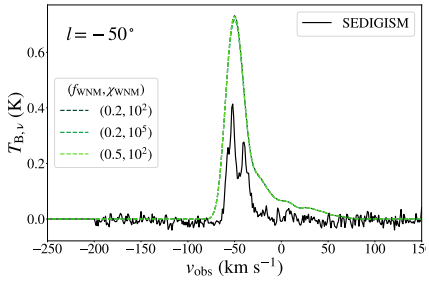


(a) $l = -50°$

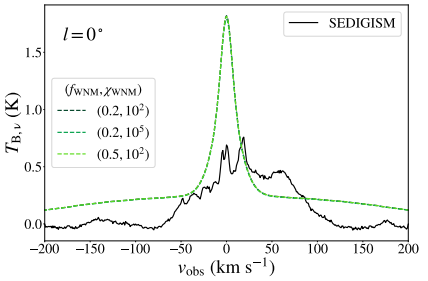


(b) $l = 0°$

Figure VI–76: The same as Figure VI–14 for the `three-media` model grid.

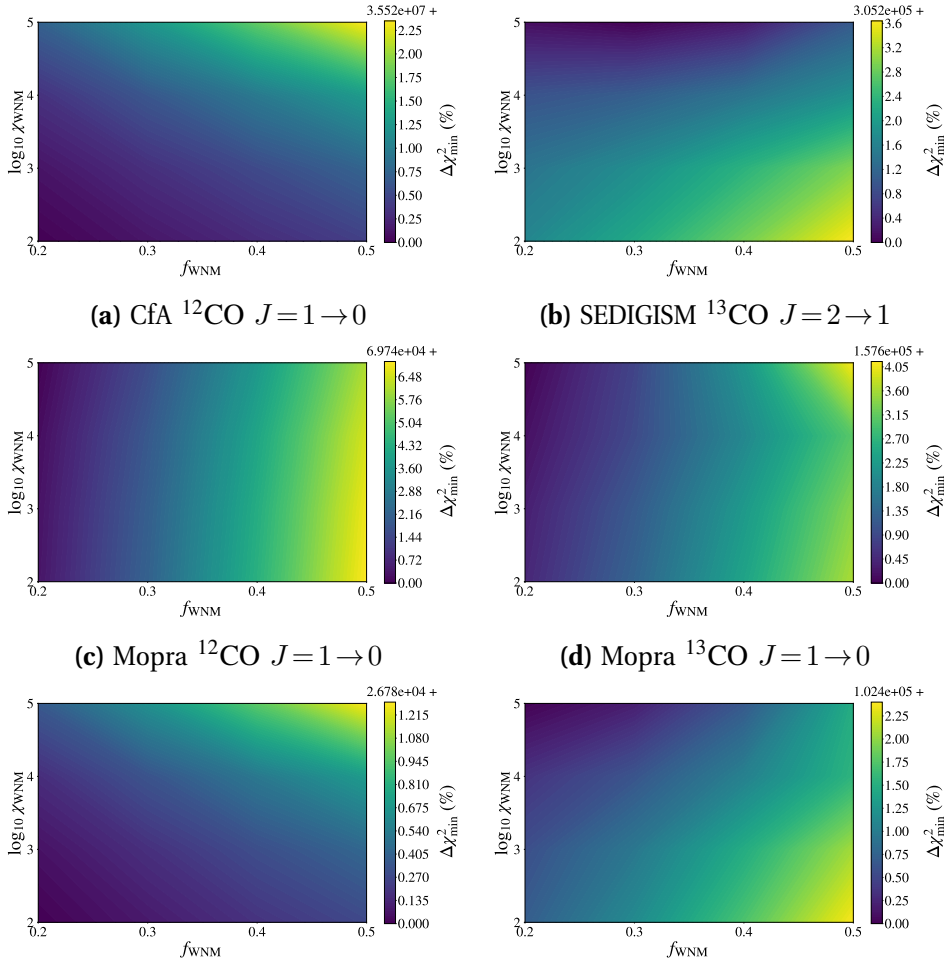


(a) $l = -50°$



(b) $l = 0°$

Figure VI–77: The same as Figure VI–15 for the `three-media` model grid.



(a) CfA ^{12}CO $J=1\rightarrow0$

(b) SEDIGISM ^{13}CO $J=2\rightarrow1$

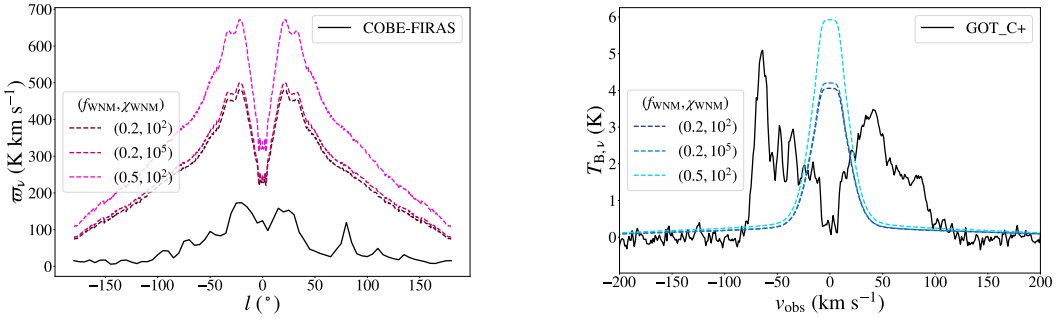
(c) Mopra ^{12}CO $J=1\rightarrow0$

(d) Mopra ^{13}CO $J=1\rightarrow0$

(e) ThrUMMS ^{12}CO $J=1\rightarrow0$

(f) ThrUMMS ^{13}CO $J=1\rightarrow0$

Figure VI–78: The comparison between the spectroscopic CO observations and the synthetic observations of the `three-media` model grid for the transitions ^{12}CO $J=1\rightarrow0$, ^{13}CO $J=1\rightarrow0$, and ^{13}CO $J=2\rightarrow1$.

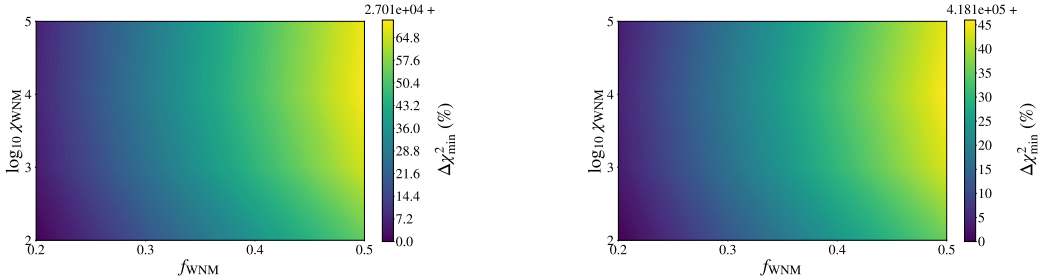


(a) *COBE*-FIRAS **(b)** GOT C+

Figure VI–79: The same as Figure VI–17 for the `three-media` model grid.

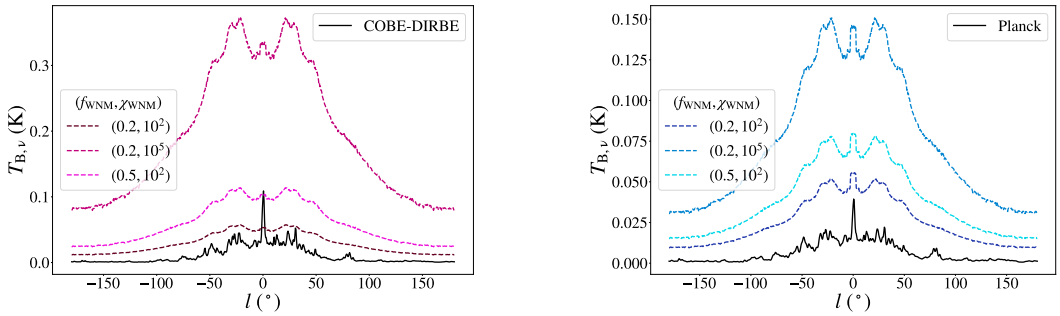
Unlike the comparison to the CO observations, the qualitative comparison of the `three-media` models to the observational surveys of *COBE*-FIRAS and GOT C+ do indicate some variation throughout the parameter space (see Figure VI–79). Subfigure **(a)** shows rather broad emission over the Galactic plane with clearly defined Galactic ring features without an intensity peak at $l = 0°$. Compared to subfigure **(a)** in Figure VI–29, however, it is clear that the models with WNM have much too high intensity. The reason for this is likely that the increased far-UV radiation creating the WNM is too much to apply globally in the Galactic disk. The spectra through the Galactic centre, seen in subfigure **(b)**, all exhibit signs of self-absorption, although the emission is very localised to the window $v_{\rm obs} \in [-50, 50]\,{\rm km\,s^{-1}}$. The self-absorption feature in the synthetic spectra is also a saturated feature rather than the self-absorption dip seen in the GOT C+ observations.

The test statistic for the comparison of the [^{12}C II] line shows considerable variation, with $\Delta\chi^2_{\rm min} \lesssim 50\%$ of $\chi^2_{\rm min} = 2.701 \times 10^4, 4.181 \times 10^5$ (see Figure VI–80). This produces a strong constraint of $f_{\rm WNM} \approx 0.2$ and a minor constraint $\chi_{\rm WNM} \approx 100\,\chi_{\rm D}$.



(a) *COBE*-FIRAS **(b)** GOT C+

Figure VI–80: The comparison between the observations to the `three-media` `kosmatau3d` grid for the transition [^{12}C II] 158 μm.

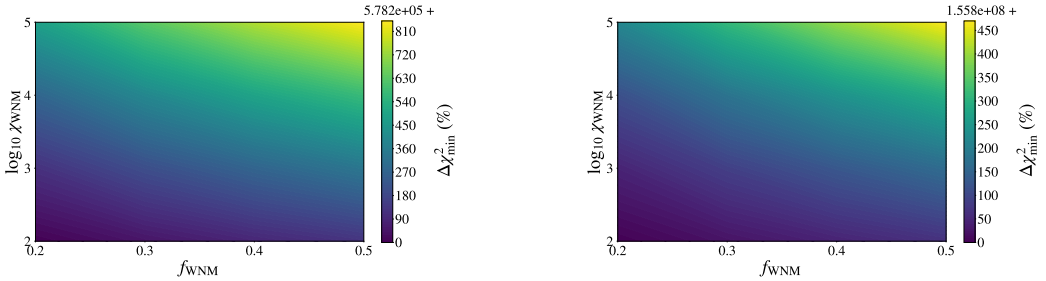


(a) *COBE*-DIRBE 240 μm **(b)** *Planck* 550 μm

Figure VI–81: The same as Figure VI–19 for the `three-media` model grid.

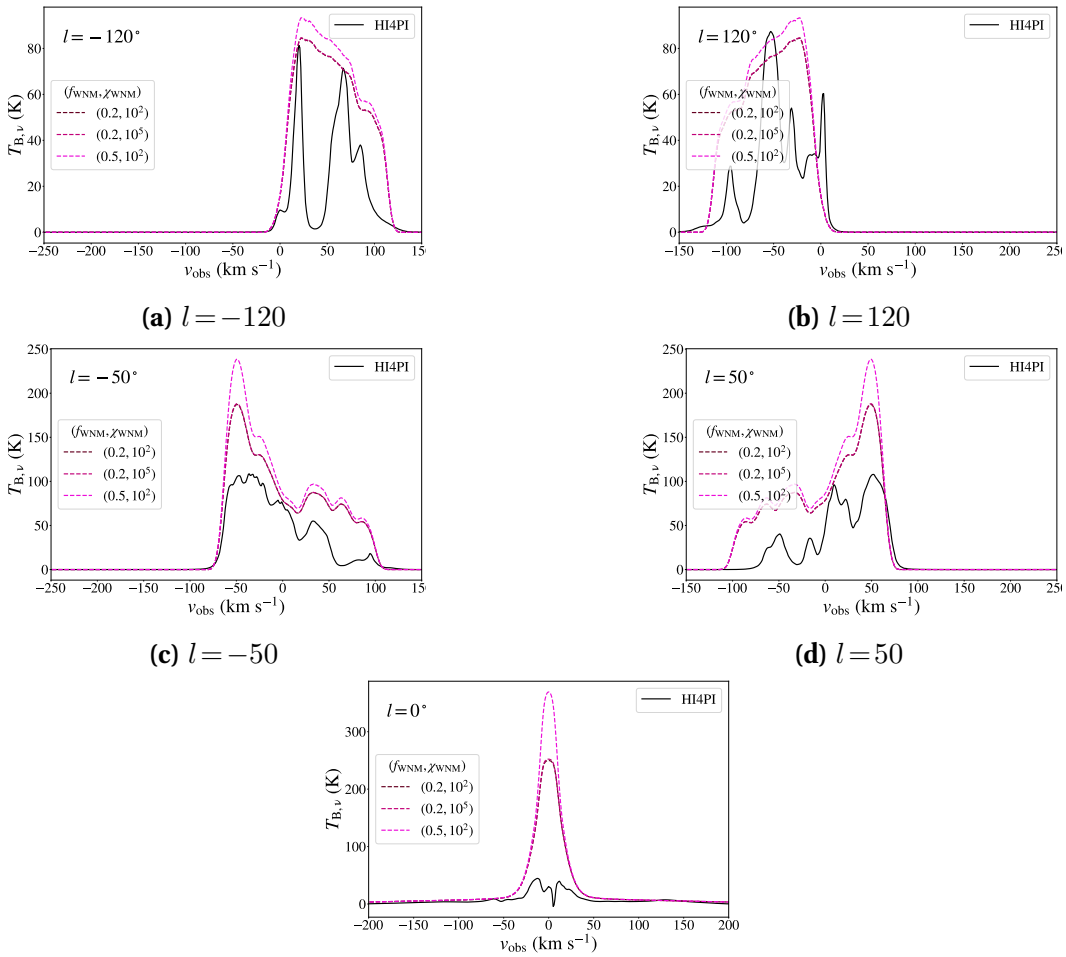
The comparison of the synthetic dust continuum to the *COBE*-DIRBE and *Planck* observations also indicates that the far-UV radiation is too high in the `three-media` model grid (see Figure VI–81). While it can be seen in both subfigures that the intensity of the synthetic continuum is too high, subfigure **(a)** shows that continuum from the model with $(f_{\rm WNM}, \chi_{\rm WNM}) = (0.2, 100\,\chi_{\rm D})$ is the most appropriate match to the observations.

The variation in the quantitative comparison provides the constraint $(f_{\rm WNM}, \chi_{\rm WNM}) \approx (0.2, 100\,\chi_{\rm D})$ (see Figure VI–82).



(a) *COBE*-DIRBE 240 μm **(b)** *Planck* 550 μm

Figure VI–82: The comparison between the dust continuum observations to the `three-media` `kosmatau3d` grid at the specified wavelengths.



(a) $l = -120$ **(b)** $l = 120$

(c) $l = -50$ **(d)** $l = 50$

(e) $l = 0$

Figure VI–83: The same as Figure VI–21 for the `three-media` model grid.

The qualitative comparison of the `three-media` grid the H I 21 cm transition indicates a closer fit of the low far-UV models to the HI4PI observations (see Figure VI–83). Unlike the previous atomic ISM tracers, there is little indication that the far-UV radiation is too intense. Subfigure **(e)** still does not exhibit any self-absorption features, but it cannot fully replicate this spectrum anyways due to the strong synchrotron source. Still, it seems this spectrum is too intense compared to the observation.

Even with such high intensities, the constraint given by the test statistic is $f_{\mathrm{WNM}} = 0.2$ (see Figure VI–84).

The $\Delta\chi^2_{\mathrm{min}}$ results are combined for the various comparisons (except for the *Planck* comparison; see §VII-7 for the discussion of why it is neglected) and summarised in VI–8.

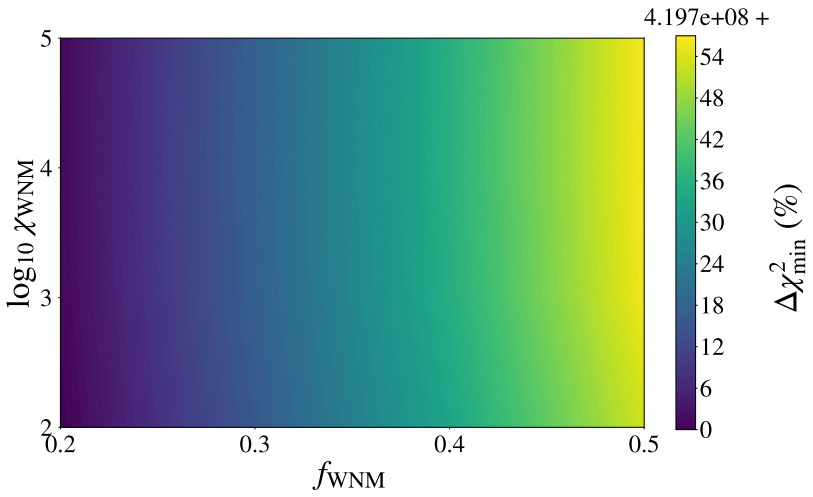


Figure VI–84: The comparison between the H I observations to the `three-media kosmatau3d` grid for the transition H I 21 cm.

Table VI–8: The $\Delta\chi^2_{\rm min}$ results for the `three-media kosmatau3d` model grid, which varies $f_{\rm WNM}$ and $\chi_{\rm WNM}$ and $m_{\rm WNM}$ The minimum χ^2 is 4.667×10^8.

$f_{\rm WNM}$	$\chi_{\rm WNM}$	$m_{\rm WNM}$	$\Delta\chi^2_{\rm min}$
0.2	$\mathbf{10^2}$	$\mathbf{10^{-3}}$	**4.8**
0.2	10^3	10^{-3}	133.6
0.2	10^4	10^{-3}	294.1
0.2	1	10^{-3}	496.6
0.3	10^2	10^{-3}	100.0
0.3	10^3	10^{-3}	270.3
0.3	10^4	10^{-3}	468.4
0.3	1	10^{-3}	702.8
0.4	10^2	10^{-3}	194.6
0.4	10^3	10^{-3}	397.6
0.4	10^4	10^{-3}	622.2
0.4	1	10^{-3}	880.4
0.5	10^2	10^{-3}	287.1
0.5	10^3	10^{-3}	518.1
0.5	10^4	10^{-3}	763.3
0.5	1	10^{-3}	1042.2

CHAPTER VII

Galactic model properties and shortcomings

> *One can only display complex information in the mind. Like seeing, movement or flow or alteration of view is more important than the static picture, no matter how lovely.*
>
> Alan J. Perlis

Given the long presentation of the results in Chapter VI, the aim in this chapter is to discuss some of the features and implications from our clumpy-PDR galactic models. Currently, `kosmatau3d` mainly evaluates the synthetic spectra throughout a model assuming just a few intrinsic quantities (see Chapter III). As such I will be focusing primarily on trends in the models relating to either the emission or mass, as well as discussing some of the assumptions or issues that remain in our analysis. To begin discussing the results of Chapter VI, I will first explain how we have combined the model comparison to constrain the parameters in §VII-1. For the galactic properties, the I will discuss the constraints on the velocity dispersion in §VII-2, the galactic profiles in §VII-3, galactic centre properties in §VII-4, and global properties (those constant throughout the model) in §VII-5. I will then discuss the attempt to include the WNM in §VII-6, in which I also discuss some of the observed trends in the model. I will explain some of the difficulties fitting the dust continuum in §VII-7, which is the motication to neglect the comparison to the *Planck* 550 μm observations. Finally, I will end with a discussion on possible improvements that can be made to the model (§VII-8).

1 Model selection

Every test statistic I list in Chapter VI have $\chi^2_{\rm min} \gg 10$, which greatly exceeds the number of degrees of freedom. For simpler theoretical models, this can be adjusted using the *reduced* $\chi^2_{\rm red}$, though this is not effective for models as complex as `kosma-tau3d`. This is due in part because of how I construct the comparison (I compare up to 10^5 intensities), and also due to the complexity of determining number of

free parameters in our models. While there are many possible parameters to adjust for each `kosmatau3d` galactic model, I focus on 15 of them. This is not at all sufficient to bring our χ^2 analysis to the limit of a close fit,

$$\chi^2 \lesssim f \tag{VII–1}$$

for the degrees of freedom f. It is neither useful for us to calculate the *information criterion* for constraining models since these are more accurate when there are varying degrees of freedom between the models.

Since the test statistic varies so much not only between models in a grid, but also in the comparison to different surveys/transitions, I have opted to essentially use the difference of χ^2 relative to the minimum as a test statistic ($\Delta\chi^2_{\text{min}}$; defined in Equation VI–7). This test statistic should therefore be a measure of how well a parameter can be constrained in a grid of models. Thus for a given grid, I use this new test statistic to constrain the parameters for each comparison (for example for the comparison to *COBE*-FIRAS [^{12}C II], HI4PI H I, etc.) and sum the results to determine the constraint in the grid parameter space.

2 Ensemble dispersion and scale height

The ensemble dispersion σ_{ens} as I had specified it in §IV-1.4 neglects any minimum velocity dispersion of giant molecular clouds, for which I had defined the parameter σ_{GMC}. Without setting a lower limit for the ensemble dispersion, the minimum ensemble dispersion for each voxel is equal to the clump velocity dispersion σ_{cl} ($0.71\,\text{km}\,\text{s}^{-1}$). In the code, I identify such models as having $\sigma_{\text{GMC}} = 10^{-2}\,\text{km}\,\text{s}^{-1}$. I computed the `voxel-dispersion` model grid for the purpose of constraining this parameter (see §VI-3). While the comparison of the grid to the *COBE*-FIRAS CO integrated intensity profiles did not yield significant results, the comparison of the [^{12}C II] 158 μm profile provided a strong constraint (see Figure VI–7 subfigure **(a)**). The main effect we notice from increasing the lower limit of ensemble dispersion is the reduced optical depth effects at the Galactic centre ($l = 0$) and the Galactic ring ($l = \pm 40\,\text{km}\,\text{s}^{-1}$). This follows logically from number of clumps being split into a wider range of velocity bins (see Equation II–26). Although in general increasing the velocity dispersion does not change the integrated intensity (see Figure VII–1), the increased absorption effects can be seen in $^{12}\text{C}^+$ since it is optically thin in the ISM.

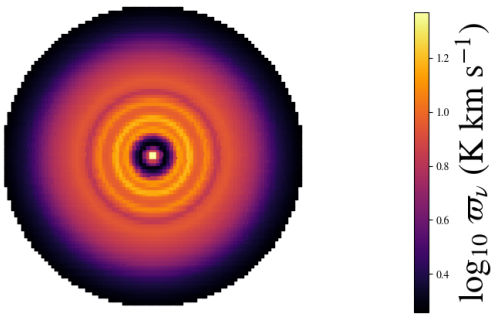


Figure VII–1: The integrated-intensity of the [^{12}C II] 158 μm transition for each voxel in the model.

While molecular clouds can exist anywhere in the Galactic disk scale height, which GMCs can have any size up to twice the scale height. Although the scale

height increases in size proportionally with $R_{\rm gal}$ both for molecular and atomic hydrogen (the flared Galactic disk advocated by Bacchini et al. 2019b, 2023), the velocity dispersion actually decreases towards the edge of the galactic disk (based on the H I velocity dispersion $\sigma_{\rm H\,I}$ in nearby galaxies, Bacchini et al. 2019a; Marasco et al. 2017). This is not present in my implementation, as I set a global value for $\sigma_{\rm GMC}$. It would be interesting to examine how the fit changes when the voxel-averaged dispersion is decreased with Galactocentric radius. Since an increased velocity dispersion does not affect the line-integrated intensity or dust continuum (which depend on the column density), this type of analysis would be useful for comparison to spectroscopic surveys.

3 Galactic profiles

The mass distributions used for the clumpy and interclump ensembles is constrained in the `mass-factors` model grid, and the results of which are listed in VI–7. The best-fitting model has the test statistic $\Delta\chi^2_{\rm min} = 997.3$ with molecular mass factor $f_{\rm m,H_2} = 1$ and atomic mass factor $f_{\rm m,H^0} = 0.5$, which refers to the clump ensemble mass and interclump ensemble mass, respectively. The constraint, with $\Delta\chi^2_{\rm min} < 2000$, is $0.5 < f_{\rm m,H_2} = 2$ and $0.25 < f_{\rm m,H^0} = 1$. The atomic mass profile is constrained to lower values because the [^{12}C II] and H I lines are overestimated by the fiducial model (see Figure VI–67 and Figure VI–71). In the continuum comparison (see Figure VI–69), the intensity profile is most sensitive to the atomic mass since there is a higher column density in the interclump ensemble rather than the clumpy ensemble (as implied by the density profiles in Figure IV–3, where there is primarily atomic mass in the outer Milky Way). Although the fiducial model overestimates the ^{13}CO lines, the ^{12}CO lines are overestimated and thus the molecular mass profile does not need to be scaled. It was not explored in this thesis, but the constraint acquired from the analysis in §VI-5 will change if we also vary the fraction of atomic hydrogen in the interclump medium (see the discussion in §VII-5)

While the mass distributions are empirically estimated (a list of the observational papers can be found in Bacchini et al. 2019b), the ensemble-averaged density and far-UV profiles are obtained rather from fitting models. It is useful and physically-motivated to use ISM properties that vary as a function of Galactocentric radius, but there is always some degree of uncertainty in using these models. For instance, the profile in Figure IV–4 is that of the CNM as determined by Wolfire et al. (2003). However, this assumes a two-medium ISM (the other one being the WNM), with the molecular gas being part of the CNM. Our galactic models similarly use two ensembles, but considering their gas temperatures (see Figure VII–2) and how I determine their relative mass, it is inconsistent to assume they correspond to the CNM and WNM. I rather model the diffuse atomic CNM and the molecular CNM in the galactic models (the attempt at including the WNM is discussed in §VII-6). With the assumptions made for the Milky Way, it is important to the models that fit well to the observations were constrained without the WNM. An accurate constraint should therefore be made using the CO lines rather than the [^{12}C II] or H I lines.

It is also inconsistent that I use the outdated solutions of Wolfire et al. (2003)

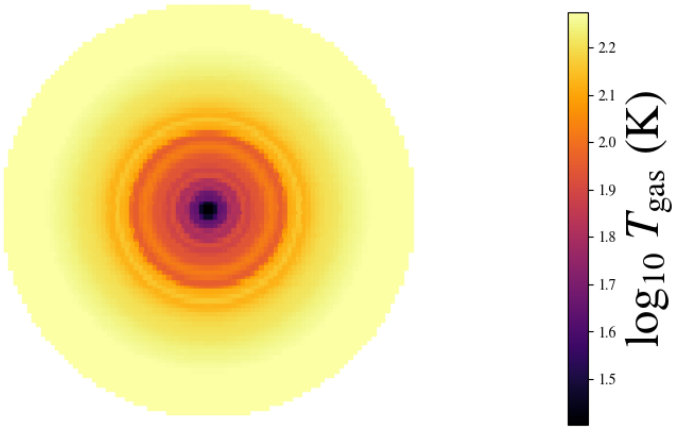


Figure VII–2: The mass-weighted gas temperature in each voxel of the model.

considering that their models do not account for the flaring of the Galactic disk as required (Bacchini et al. 2019b, 2023). Practically, it would be better to use our calculated volume densities for H^0 and H_2 and scale them for the ensemble-averaged density profile. There are two reasons I did/could not do this: the H_2 mass drops by a factor of $\sim 10^{10}$ near the disk edge, which makes modelling the molecular gas near the edge of the Galaxy impossible without further assumptions, and I enforced that the interclump ensemble is primarily atomic, which was only possible when it is composed of $10^{-3}\,M_\odot$ clumps. Therefore the interclump ensemble-averaged density could not exceed $\sim 19\times10^{-3}\,cm^{-3}$ (see the definition of the clump-averaged density Equation II–2), and it cannot vary by the two orders of magnitude as in the H^0 profile. It is anyways more accurate to perform a dynamic analysis such as in Wolfire et al. (2003) in order to make the model self-consistent. Considering the setup of my models and the mass distributions we employ, the error in using their density results should be small in comparison to the error associated with the other assumptions (for example the hydrogen distribution mentioned in §VII-6).

One concerning aspect of Figure IV–7 is that it implies the local far-UV radiation density is below the generally accepted value of $1\,\chi_D$. What the Popescu et al. (2011) profile nicely captures is the increased energy density due to the molecular ring, and according to the limitations of our model the unscaled far-UV energy density profile changes by a factor < 1.5. Of course this is not seen in the original distribution (see the first plot of Figure IV–5), but rather a lower limit of our PDR model. For this reason I scale the profile over a large range in a few of the grids (up to a factor 10^4), and the physical motivation for this is that the ISM existing in PDRs are located closer to the ionising sources than the volume-averaged far-UV radiation energy density. Such an argument was used also by Cubick et al. (2008) to motivate their Galactic disk solution of $10^{1.8}\,\chi_D$, which was taken as a constant. Even in the method in which I apply the far-UV radiation profile, I still obtain a better fit to the unscaled profile. What was interesting to note is that I obtain different results depending on whether or not I compare line-integrated intensities; the comparisons to *COBE*-FIRAS seemed to favour higher values of the χ_{cl}, in accordance with the results of Cubick et al. (2008), and no constraint on the interclump medium far-UV

radiation. The comparisons to both the dust continuum and the spectroscopic surveys instead indicate the far-UV radiation for both the clumpy and interclump ensembles should not only be similar, but that the best-fitting model has a far-UV profile scaled by a factor $f_{\rm FUV} \lesssim 10^{0.5}$. The dust continuum was critical to form such a definitive constraint. Part of the reason for this strong constraint to lower values of $f_{\rm FUV}$ is the excess of emission from the outer disk (seen at longitudes $|l| > 90°$). It is possible that an improvement to the dust treatment in KOSMA-τ (see §VII-7 for this discussion) might improve this synthetic observation in this longitude range, but this need to be tested in future research into these models. With this in mind, it would be better for the future work to use more than one wavelength to constrain the continuum.

Given the previously-mentioned inconsistencies, it would be a useful exercise to solve for a self-consistent density and far-UV radiation profile. A possible way forward is to perform an analysis similar to Wolfire et al. (2003) using the updated mass and scale height profiles, only this time allow for the molecular, cold neutral, and warm neutral media. Using the self-consistent density profiles (which do not make the vague assumption of a molecular clumpy ensemble and an atomic interclump ensemble) from this analysis, one could constrain the far-UV radiation of the voxels along each line-of-sight, using the Popescu et al. (2011) estimation as an initial guess. This could be useful for the lines-of-sight where there are many voxels with different voxel-averaged velocities (basically for voxels between the observer and the Galactic centre; see Figure VI–2) or towards the edge of the Galaxy.

4 Galactic centre properties

I have created the `CMZ-radius-radiation` and `FUV-Galactic-centre` model grids for the purpose of fitting better the line-of-sight at towards the Galactic centre. Both of these grids had the far-UV factor as a parameter, either the global factor $f_{\rm FUV}$ or just in the galactic core $f_{\rm FUV,GC}$. As such, it exhibited similar effects to some of the other model grids that varied the far-UV factor, namely that the dust typically forces the constraint to values $f_{\rm FUV} \lesssim 10^{0.5}$. The test statistic in the `FUV-Galactic-centre` grid was $\sim 10\%$ for most models, and the best-fitting model had parameters $f_{\rm FUV,GC} = 10^{1.5}$ and $R_{\rm GC} = 1400\,$pc. The increased radiation in the Galactic centre was mainly constrained by the [^{12}C II] line and dust continuum. For the `CMZ-radius-radiation` grid, the test statistic was a factor had more variation and was constrained mainly by the dust continuum due to the far-UV. When comparing spectra, the largest effect from the cosmic ray ionisation rate (which is higher within $R_{\rm CMZ}$) was in the CO spectra. The best-fitting model had $R_{\rm CMZ} = 200\,$pc and $f_{\rm FUV} = 1$ mainly due to the large effect the far-UV factor has on the comparison to the dust continuum. By comparison of the profiles in Figure VI–31 (for example), the comparison to the dust continuum yields a stronger constraint due to the low intensity at longitudes $|l| \gtrsim 50°$. As such, it would be more enlightening to combine these grids, increasing the far-UV just in the Galactic CMZ, which also has an increased cosmic-ray ionisation rate. Such a model grid should be better at constraining the radius of the CMZ while allowing for higher

far-UV radiation intensity.

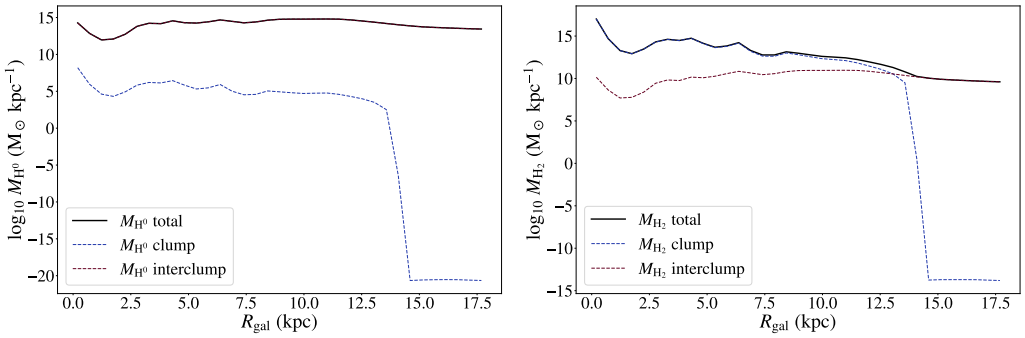
5 Global properties

Some of the global properties that were changed include the clump mass range $m_{\rm cl}$ chosen for the clumpy ensemble in the `clump-masses`model grid and the fraction of atomic hydrogen in the interclump ensemble $f_{\rm H^0}$ in the `atomic-ISM-radiation` model grid. While these grids also changed the far-UV factor $f_{\rm FUV}$, to which the dust continuum is very sensitive, there were some noticeable effects from the global parameters. For instance, decreasing the atomic fraction also decreased the intensity of the H I spectra, which fit better to the HI4PI observations. Likewise, the [^{12}C II] integrated intensity was decreased and fits better to the *COBE*-FIRAS profile. However, the CO spectra were increased due to the added mass. This improved the fit to the ^{12}CO $J=1\rightarrow0$ observations, but the fit to the ^{13}CO lines worsened. The $\Delta\chi^2_{\rm min}$ analysis ultimately constrained $f_{\rm H^0}\gtrsim0.4$ for $f_{\rm FUV}\lesssim10^{0.5}$. The variation of the test statistic in the `clump-masses`model grid was mainly dominated by the far-UV factor, but there is a slight constraint favouring lower clump mass. Since the $m_{\rm cl}$ parameter is categorical rather than continuous, it is best to rank the options of the fit. In order from best fit to worst, the clump mass ranges $m_{\rm cl}=[0,2],[-1,2],[-1,3],[0,3]\,{\rm M}_\odot$ fit comparatively well for $f_{\rm FUV}\approx1$.

6 Hydrogen in different media

The atomic medium in the Milky Way, which I trace with H^0, can be obtained both in terms of mass and emission. For our clumpy models, as discussed in §III-3.2 and verified by the comparison to of the H^0 21 cm line in Chapter VI, the atomic medium is contained within the interclump ensemble. I can verify that in Figure VII–3, where I can track the atomic and molecular masses in the Milky Way models as they are split between the ensembles. It clearly shows that the atomic mass is primarily in the interclump medium throughout the Milky Way model.

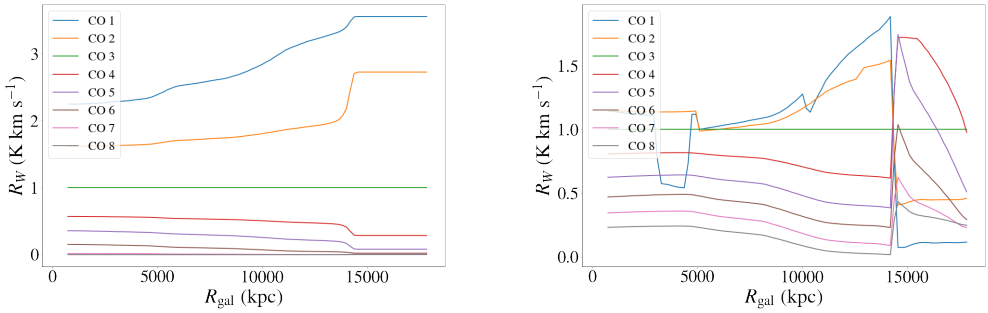
For our `three-media` model grid, I test the use of three ensembles: one dense clumpy ensemble for the molecular gas, one diffuse atomic interclump ensemble for the CNM, and one diffuse atomic ensemble for the WNM (see the ISM media in §I-1). This was an unsuccessful attempt to increase the HISA in the synthetic maps and produce a better fit to the observations. One issue is that the temperature of the WNM in these models is not quite high enough for the theoretical estimates of the WNM, so there is less emissivity and more absorption from these regions, but this at first glance improve the HISA. Another (major) explanation that explains the lack of HISA in the synthetic observations is that the calculations introduced in §I-1.2.2 assumes thermal equilibrium, which will never feature self-absorption. The it was assumed that the variety of environments in the Galactic centre (as seen in the voxel parameters) would introduce enough large-scale variation to reproduce the HISA. This same explanation was used to replicate the self-absorption in the [^{12}C II] 158 μm transition. While the [^{12}C II] self-absorption feature could be replicated



(a) H^0 **(b)** H_2

Figure VII–3: The neutral and molecular hydrogen profiles as they appear in our 400 pc model from the `convergence` grid. The solid black, dashed blue, and dashed red profiles correspond to the total, clump, and interclump masses, respectively.

with increased radiation in the Galactic centre, HISA was not present in any of the models.



(a) $f_{FUV} = 1$ **(b)** $f_{FUV} = 10^4$

Figure VII–4: The line ratio of the ^{12}CO lines as a function of Galactocentric radius. Each line is compared with the $J = 3 \rightarrow 2$ transition.

From this mass profile, I can also compare now to the intensity profiles of various transitions throughout the Milky Way. In Figure VII–4 subfigure (a), the line integrated-intensity ratios are plotted as a function of Galactocentric radius. All lines are sub-thermally excited (since $R_W \neq 1$), and thus there are non-LTE conditions. The increase in R_{13} and R_{23} as a function of r_{gal} paired with the decrease in the ratio for higher transitions is consistent with a decrease in gas temperature in the clumpy ensemble[32]. Subfigure **(b)** shows in general that increasing the far-UV radiation moves the CO line ratios closer to LTE, but there exist some significant artefacts in the data at around 4 pc and 10 kpc, but the discontinuity at 15 kpc is due to the decrease in clumpy ensemble mass. This appears as a discontinuity since the diffuse clumps are more sensitive to changes in gas temperature (compare this with the PDR profile of the diffuse clumpy in Figure III–7).

[32]This is not visible in Figure VII–2 since T_{gas} normalised to the atomic hydrogen abundance.

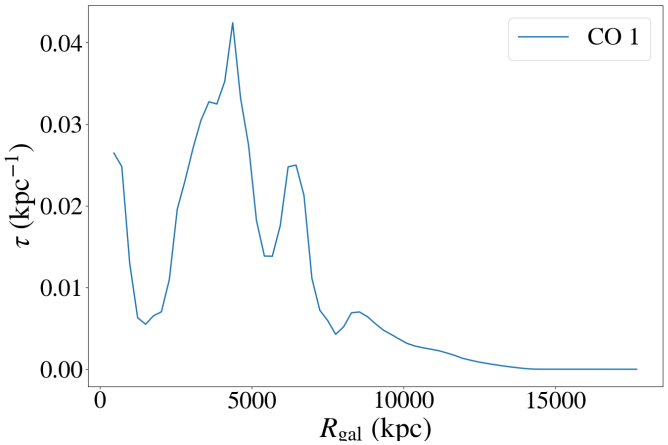


Figure VII–6: The maximum optical depth of the ^{12}CO $J = 1 \rightarrow 0$ line as a function of Galactocentric radius. The peak of the profile corresponds to an optical depth effect of $\sim 4\%$.

I can also estimate the amount of CO-dark H_2, or rather, how does the X_{co} factor change throughout the Galactic disk compared to the Milky Way average? Recalling Equation I–33 for CO:

$$X_{CO} = \frac{\mathcal{N}(H_2)}{\varpi_{CO}} \ ,$$

where $\mathcal{N}(H_2)$ is the column density of molecular hydrogen and ϖ_{CO} is the integrated intensity of ^{12}CO $J = 1 \rightarrow 0$. I calculate these values using the length-scale of each voxel, ℓ_{vox}. Figure VII–5 shows the azimuthally-averaged profile of X_{CO} in from our fiducial Galactic model, normalised to the Milky Way average of $X_{CO,MW} = 2 \times 10^{22}\,\mathrm{cm^{-2}(K\,km\,s^{-1})}$. When $X_{CO} < 1$, it means there is more ^{12}CO $J = 1 \rightarrow 0$ emission than the Milky Way average (consequently overestimating the amount of H_2), and $X_{CO} > 1$ means there is more H_2 compared to what I can calculate using Equation I–33 (meaning there exists CO-dark H_2). What I can infer directly from this figure is that, for the CO emission produced solely from PDRs, I will overestimate the amount mass in H_2 using $X_{CO,MW}$, and all of the CO-dark H_2 exists in the outer 3 kpc. As it has already been shown in Figure VII–3, it is an artefact in our models that there is so much molecular mass in the outer galaxy.

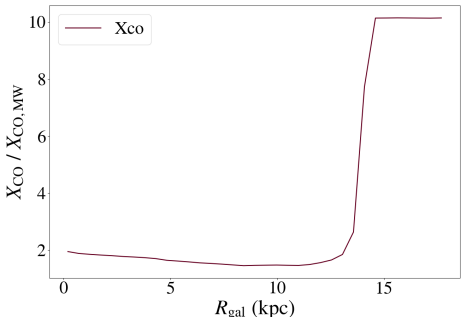


Figure VII–5: The X_{CO} factor as determined by the ^{12}CO $J = 1 \rightarrow 0$ transition.

The result for the inner 11 pc of the galactic disk is not an artefact, but instead reflects a property inherent in our galactic models: there is more CO-dark H_2 in

PDRs. Based on the comparisons of the spectra in Chapter VI, it is clear that the column density of CO is too low to replicate the observations (that is partially why lower values of the atomic hydrogen fraction f_{H^0} fit better in the `atomic-ISM-radiation` comparison).

7 Continuum long-wavelength limit

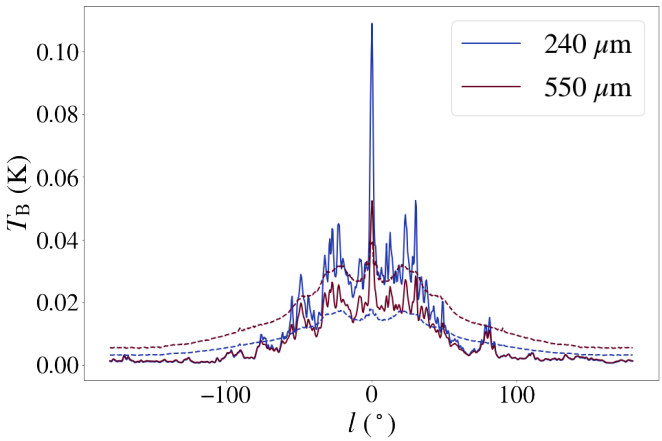


Figure VII–7: The 240 μm and 550 μm observations of the Milky Way galactic disk. The synthetic observations of the fiducial `kosmatau3d` galactic model is plotted with dashed curves.

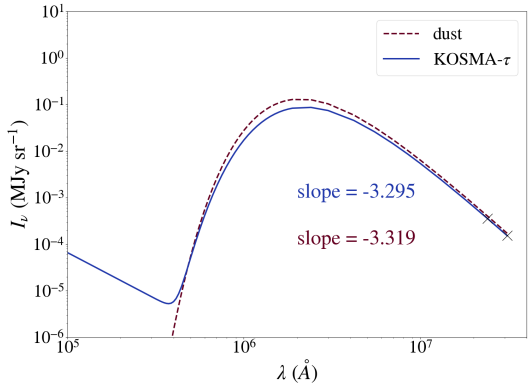


Figure VII–8: The continuum provided by KOSMA-τ for a model with $m_{cl} = 10^{-3}\,M_\odot$, $n_s = 10\,cm^{-3}$, $\chi = 1\,\chi_D$. This model uses the MRN dust model. The red dashed curve is an independent calculation of the dust SED from the extinction curve and dust temperature results of the KOSMA-τ model. The slopes result from a linear fit to the curves at 3.2 mm and 1 mm, shown by the black markers

As I have seen in the continuum comparison in Chapter VI, it was not possible to fit both the 240 μm and the 550 μm wavelengths simultaneously. Figure VII–7 shows the issue as I have first discovered it. From this it is clear that the optical depth at 550 μm is too large, and I am thus calculating a synthetic intensity higher than that at 240 μm, and any correction should decrease the emission at 550 μm. One possible reason for this, in so far as the properties assumed in our clumpy model, is the relative temperatures of our clumpy and interclump ensembles. If I consider the dust in a region of the ISM at some temperature T_d and its grain-size distribution, I can effectively model the intensity SED by a *greybody* (Désert 2022):

$$I_\nu = \tau_0 \left(\frac{\lambda}{\lambda_0} \right)^\beta B_\nu(T_\mathrm{d}) \ , \tag{VII–2}$$

where $B_\nu(T_\mathrm{d})$ is the Planck function and β is the free emissivity index, which at long wavelengths should be $\beta = -2$. For the brightness temperature I use in this work, this gives us the long-wavelength dependency of $I_\nu \propto \lambda^{-4}$; the Rayleigh limit and the emissivity index each give a dependency of λ^{-2} Ossenkopf & Henning (1994). A modification of this relation was introduced by Li & Draine (2001) in an attempt to better model the *COBE*-FIRAS observations, where the imaginary part of the silicate dielectric constant ϵ_DE has been altered for wavelengths $\lambda > 250\,\mu$m to effectively increase the long-wavelength dependence by $\lambda^{0.4}$ (so $\beta = -1.6$). The results from the *Planck* survey indicate a similarly-shallow emissivity index of the observed continuum at latitudes away from the Galactic plane, but $\beta \approx 1.8$ (Planck Collaboration et al. 2014a,c, 2020).

In KOSMA-τ, the long-wavelength dependence of the brightness temperature is a slightly more-shallow $I_\nu \propto \lambda^{3.3}$ (see Figure VII–8). The reasons for this shallow dependency are due in part to the distribution of dust grain sizes and temperatures, which slightly broaden the modified blackbody, as well as the assumed extinction dependency (which is proportional to the dielectric constant). The extinction wavelength-dependency for the KOSMA-τ model has been $A(\lambda) \propto \lambda^{-1.5}$ since its implementation (see for example Figure 25 in Röllig et al. 2013). Practically this has the effect that the dust temperatures in the clump are about a factor of 20% too low[33], which decreases the formation of H_2 on dust grains and reduces the size of the molecular region of the clumps. A robust analysis of how much this affects the KOSMA-τ results over the past 10 years is still required, but regarding the results of this thesis it is mainly important to stay consistent; the error should be highest at longer wavelengths, which is why I neglected the model-fit to the *Planck* 550 μm observations. It is left for future research to constrain the Milky Way models with longer wavelengths and determine how much this dust issue has affected the results, and a possible way forward is discussed in Appendix C.

8 Future development

Although there are still a few limitations to the approach of modelling galaxies outlined in this thesis, I have a few ideas for future developments that would benefit this type of analysis:

1. The modifications for partially-filled voxels need to be correctly implemented. The method we used for the Milky Way is valid only for observations where the ISM fills the beam (in the sense that we can treat the disk as an extended object with our lines-of-sight, not that the beam-filling factor is $f_\mathrm{beam} \approx 1$).

[33] This estimate was acquired through personal communication with the developer of KOSMA-τ, Markus Röllig.

2. Dynamic processes such as shock can be included, at least statistically. It is likely that this requires the inclusion of plane-parallel results, for example using irradiated shocks from the Paris-Durham Shock code (Godard et al. 2019).

3. There should really be a way to include a warmer ensemble to properly model the WNM. This requires development in KOSMA-τ.

4. Similar development necessary in KOSMA-τ is diffusion, advection, and time-dependent chemistry. This would remove the current assumption of static clumps.

There are also a few general developments that would benefit `kosmatau3d` to make it more accessible for modelling the observations of PDRs:

1. It is possible to use `kosmatau3d` voxels to individually model the spaxels in an observational datacube (this will be presented in Okada et al. in prep.). This requires one to create a grid of voxel results from which on can interpolate using the parameters of each spaxel. This will should be implemented as a function in `kosmatau3d`. It should be optionally extended to include the two-component models of Guevara et al. (2020) and Kabanovic et al. (2022).

2. There should exist some functionality in `kosmatau3d` to constrain PDR properties based on a set of lines and line ratios similar to `PDRToolbox`. This will make it easier for observational astronomers to utilise `kosmatau3d` to compare their observations assuming a clumpy PDR.

3. The clumpy modelling shown in Röllig & Ossenkopf-Okada (2022) should be included as an additional function. This would provide the base for comparing the two analyses and constraining the associated error in deriving the PDR properties. This comparison is motivated by the difference between the results of Cubick et al. (2008) and the results of my comparison in §VI-5. It is certainly useful to understand why these results differ by such a large factor.

CHAPTER VIII

Summary

A conclusion is the place where you got tired thinking.

Martin H. Fischer

The Milky Way is a very complex object in terms of the ISM, with several phases each with different physical conditions. There are processes such as stellar accretion and galactic outflows removing material from the ISM, and processes such as YSO jets, stellar winds, supernovae, and galactic accretion adding material to the ISM. For these reasons it is somewhat surprising to find that the clumpy PDR models created for this thesis, which build upon the stationary time-steady KOSMA-τ model, are able to replicate several trends and much of the structure of the observations of galactic plane. Some of the features we can reproduce seem forced; the molecular and atomic hydrogen mass was separated into different ensembles (except for the `atomic-ISM-radiation` model grid), which we in turn forced to be primarily molecular or primarily atomic. Thus it was no surprise that the modelled molecular and atomic distributions closely follow the input. Since even the diffuse interclump ensemble still contained some molecular hydrogen ($\sim 10\%$ by mass), we are overestimating the amount of molecular hydrogen in the outer Galactic disk. That is if we assume our ensembles are modelling the CNM and molecular media of the ISM.

Even though I could not completely replicate the self-absorption features of the diffuse gas tracers, it is still remarkable to see the imprint of the molecular ring in the position-velocity diagrams of the Galactic plane. Since the structure seen in the synthetic intensity datacubes is mainly due to the velocity structure of the gas, I did not attempt any structure-finding analysis to constrain the model grids. By constraining the grids using the intensity, there are a few main conclusions I can draw from the analysis outlined in this thesis[34]:

1. The large-scale structure of spectra away from the Galactic centre can be

[34]Note that the enumeration of the conclusions do not match 1:1 with the enumeration of the questions in Chapter I.

reproduced reasonable well using CO, while the [^{12}C II] 158 μm and H I 21 cm emission typically overestimated.

2. The parameter I defined as the GMC dispersion $\sigma_{\rm GMC}$ was constrained to $2.63\,{\rm km\,s^{-1}}$, corresponding to a cloud scale $L_{\rm GMC} = 10\,{\rm pc}$. Although the higher values ($6.33\,{\rm km\,s^{-1}}$, corresponding to a cloud scale $L_{\rm GMC} = 100\,{\rm pc}$) had a lower test statistic, this was determined to be unreasonable due to the limited scale height.

3. I constrained the clumpy ensemble far-UV radiation intensity $\chi_{\rm cl} \lesssim 10^{0.7}\,\chi_{\rm D}$ and the interclump ensemble far-UV scaling factor $f_{\rm FUV,icl} \approx 10^{0.25}$. This is about an order of magnitude lower than the results of Cubick et al. (2008) using solely the *COBE*-FIRAS observations. From the comparison the the different survey it was concluded that it is sufficient to use the same far-UV intensity for the clump and interclump ensembles.

4. The far-UV scaling factor $f_{\rm FUV}$, which scaled the far-UV radiation intensity for both the clump and interclump ensembles, was included in many on the model grids since there can be a large amount of variation in the far-UV radiation in star-forming regions (Xia et al. 2022).

5. The fraction of atomic fraction in the interclump ensemble compared to the clumpy ensemble $f_{\rm H^0}$ has been constrained to $\gtrsim 0.4$. The decrease in mass of the interclump ensemble made the [^{12}C II] and H I comparison fit better, since they were typically overestimated when $f_{\rm H^0} = 1$. The increased mass in the clumpy medium made the ^{12}CO spectral lines fit better as they were underestimated in the fiducial model, but the ^{13}CO lines were overestimated for the acceptable values of $f_{\rm FUV}$.

6. I have constrained the atomic mass factor $f_{\rm m,H^0} \gtrsim 0.5$ and the molecular mass factor $f_{\rm m,H_2} \lesssim 1$ in the `mass-factors` model grid. This was achieved by using the molecular mas for the clumpy ensemble and the atomic mass for the interclump ensemble. The constraint of this grid had effects similar to that of the `atomic-ISM-radiation` grid. These results will change if one considers an atomic mass factor $f_{\rm H^0} < 1$.

7. The Galactic centre far-UV factor $f_{\rm FUV,GC}$ and size $R_{\rm GC}$ were constrained to $\lesssim 10^{1.5}$ and $\lesssim 1400\,{\rm pc}$. The increased far-UV radiation increased the size of the absorption feature at $v_{\rm obs} = 0\,{\rm km\,s^{-1}}$ of the [^{12}C II] spectrum at $l = 0$, but increased too much the intensity of the spectrum at $|v_{\rm obs}| > 100\,{\rm km\,s^{-1}}$.

8. I determined the radius of the CMZ $R_{\rm CMZ} = 200\,{\rm pc}$, in agreement with the accepted size of the CMZ (García et al. 2016).

9. It was determined that the clump mass range $m_{\rm cl}$ used in the fiducial model, $[10^0, 10^2]\,{\rm M_\odot}$, fitted best to the observations. The models using larger masses fitted worst, likely due to increased optical depth effects.

10. I managed to get an approximate constraint for the global WNM parameters, which was acquired in the comparison to the `three-media` model grid. The factor of the atomic mass in the WNM ensemble was constrained to $f_{\mathrm{WNM}} \gtrsim 0.2$ and the far-UV radiation intensity of the WNM ensemble was constrained to $\chi_{\mathrm{WNM}} \lesssim 100\,\chi_{\mathrm{D}}$.

Despite the successes of the galactic models and the trends that I found, there are also some features of the comparison that did not fit well:

1. The spectrum through the galactic centre ($l = 0°$) cannot be replicated for any transition. In the case of the CO transitions, the observed spectra are quite broad compared to the synthetic intensities, with a FWHM of $\gtrsim 100\,\mathrm{km\,s^{-1}}$. The synthetic spectra, conversely, appear to be more localised and almost Lorentzian in shape. The observed ^{12}CO intensity exceeds the synthetic intensity, however the synthetic ^{13}CO intensities are much larger than the observed intensities.

2. The self-absorption features in the spectra of diffuse tracers such as the [^{12}C II] 158 μm and H I 21 cm transitions cannot be fully modelled. Models with increased far-UV radiation in the Galactic Centre are able to show signs of self-absorption in the [^{12}C II] 158 μm at $l = 0°$, but these models do not show the same feature in H I.

3. A possible issue with the H I post-processed clump emission is the assumption made to calculate it: thermal equilibrium. We do not consider any H I shielding within the clump, but rather assume self-absorption occurs within the ensemble and to a large degree between different ISM environmental conditions. It is possible that fully calculating the H I transition clump-averaged intensity in the same way as the other line transitions will increase the self-absorption effects since there will be more absorption from the cooler temperature atomic hydrogen.

4. The H I 21 cm absorption feature due to the strong synchrotron source cannot be replicated using PDRs.

5. The X_{CO} factor is always greater than the Milky Way average by at least a factor of 2, with the outer Milky Way disk increasing to 13 since the gas is mainly atomic. This implies that, if the CO $J = 1 \rightarrow 0$ emission in the Galaxy is due to PDRs, we are underestimating the molecular gas mass by a factor of at least 2.

Regarding future developments, it must first be noted which developments will be useful from the available KOSMA-τ model grids:

1. The continuum issue mentioned in §VII-7 must be addressed as this affects the dust temperature by $\sim 20\%$, which decreases the amount of H_2formation on dust grains.

2. It would be useful to have a portion of the model grid that can be used for the WNM, where the gas temperature is $8\times10^3 \lesssim T_{\rm g} \lesssim 10^4$ K.

3. Certain effects such as diffusion, advection, and time dependence are still missing from the available model grids, but are certainly interesting to include as options in `kosmatau3d`. Some of these should become available soon (see for example Aleena B et al. 2023).

4. All of these model grids should be made available on `ISMDB`, from where we can query and download them using `kosmatau3d`.

The development of the Milky Way galactic models based on `kosmatau3d` has been used as a proof-of-concept of modelling galactic cooling lines using PDRs. It has yielded many interesting trends and promoted the improvement of the model from (Andree-Labsch et al. 2017). The knowledge gained from this analysis will serve as a benchmark for the modelling of nearby star-forming galaxies such as M81. With all of the features I have implemented during the development of `kosmatau3d` and the future plans for continued improvement, it is poised to have a lasting impact on the modelling of both PDRs and galaxies.

Bibliography

Albertsson, T., Kauffmann, J., & Menten, K. M. 2018, *The Astrophysical Journal*, 868, 40

Aleena B, Ossenkopf-Okada, V., & Röllig, M. 2023, in Physics and Chemistry of Star Formation: The Dynamical ISM Across Time and Spatial Scales, 139

Andree-Labsch, S., Ossenkopf-Okada, V., & Röllig, M. 2017, *Astronomy & Astrophysics*, 598, A2

Bacchini, C., Fraternali, F., Iorio, G., & Pezzulli, G. 2019a, *Astronomy & Astrophysics*, 622, A64

Bacchini, C., Fraternali, F., Pezzulli, G., et al. 2023, IAU Symposium, 373, 310

Bacchini, C., Fraternali, F., Pezzulli, G., et al. 2019b, *Astronomy & Astrophysics*, 632, A127

Barnes, A. T., Glover, S. C. O., Kreckel, K., et al. 2021a, *Monthly Notices from the Royal Astronomical Society*, 508, 5362

Barnes, A. T., Henshaw, J. D., Fontani, F., et al. 2021b, *Monthly Notices from the Royal Astronomical Society*, 503, 4601

Barnes, A. T., Longmore, S. N., Dale, J. E., et al. 2020, *Monthly Notices from the Royal Astronomical Society*, 498, 4906

Barnes, P. J., Muller, E., Indermuehle, B., et al. 2015, *The Astrophysical Journal*, 812, 6

Bell, T. A., Hartquist, T. W., Viti, S., & Williams, D. A. 2006a, *Astronomy & Astrophysics*, 459, 805

Bell, T. A., Roueff, E., Viti, S., & Williams, D. A. 2006b, *Monthly Notices from the Royal Astronomical Society*, 371, 1865

Bennett, C. L., Fixsen, D. J., Hinshaw, G., et al. 1994, *The Astrophysical Journal*, 434, 587

Bhattacharjee, P., Chaudhury, S., & Kundu, S. 2014, *The Astrophysical Journal*, 785, 63

Bisbas, T. G., Bell, T. A., Viti, S., Yates, J., & Barlow, M. J. 2012, *Monthly Notices from the Royal Astronomical Society*, 427, 2100

Bisbas, T. G., Tan, J. C., & Tanaka, K. E. I. 2021, *Monthly Notices from the Royal Astronomical Society*, 502, 2701

Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, *Annual Reviews in Astronomy and Astrophysics*, 51, 207

Braiding, C., Burton, M. G., Blackwell, R., et al. 2015, *Publications of the Astronomical Society of Australia*, 32, e020

Carnall, A. C. 2017, arXiv e-prints, arXiv:1705.05165

Carruthers, G. R. 1970, *The Astrophysical Journal Letters*, 161, L81

Cubick, M., Stutzki, J., Ossenkopf, V., Kramer, C., & Röllig, M. 2008, *Astronomy & Astrophysics*, 488, 623

Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, *The Astrophysical Journal*, 547, 792

Dame, T. M., Ungerechts, H., Cohen, R. S., et al. 1987, *The Astrophysical Journal*, 322, 706

den Brok, J. S., Bigiel, F., Sliwa, K., et al. 2022, å, 662

den Brok, J. S., Chatzigiannakis, D., Bigiel, F., et al. 2021, *Monthly Notices from the Royal Astronomical Society*, 504, 3221

Désert, F.-X. 2022, *Astronomy & Astrophysics*, 659, A70

Draine, B. T. 1978, *The Astrophysical Journal Supplementary Series*, 36, 595

Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium (Princeton: Princeton University Press)

Draine, B. T. & Bertoldi, F. 1996, *The Astrophysical Journal*, 468, 269

Draine, B. T. & Lee, H. M. 1984, *The Astrophysical Journal*, 285, 89

Driessen, F. A., Sundqvist, J. O., & Dagore, A. 2022, *Astronomy & Astrophysics*, 663, A40

Eilers, A.-C., Hogg, D. W., Rix, H.-W., & Ness, M. K. 2019, *The Astrophysical Journal*, 871, 120

Elmegreen, B. 1999a, in The Physics and Chemistry of the Interstellar Medium, ed. V. Ossenkopf, J. Stutzki, & G. Winnewisser, 77

Elmegreen, B. G. 1999b, *The Astrophysical Journal*, 527, 266

Elmegreen, B. G., Elmegreen, D. M., & Seiden, P. E. 1989, *The Astrophysical Journal*, 343, 602

Elmegreen, B. G. & Falgarone, E. 1996, *The Astrophysical Journal*, 471, 816

Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, *Revista Mexicana de Astronomía y Astrofísica*, 53, 385

Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, *The Astrophysical Journal Letters*, 155, L149

Fixsen, D. J., Bennett, C. L., & Mather, J. C. 1999, *The Astrophysical Journal*, 526, 207

Fixsen, D. J., Cheng, E. S., Cottingham, D. A., et al. 1994, *The Astrophysical Journal*, 420, 457

García, P., Simon, R., Stutzki, J., et al. 2016, *Astronomy & Astrophysics*, 588, A131

Geen, S. & de Koter, A. 2022, *Monthly Notices from the Royal Astronomical Society*, 509, 4498

Geurts, P., Ernst, D., & Wehenkel, L. 2006, Machine Learning, 63, 3

Gierens, K. M., Stutzki, J., & Winnewisser, G. 1992, *Astronomy & Astrophysics*, 259, 271

Godard, B., Pineau des Forêts, G., Lesaffre, P., et al. 2019, *Astronomy & Astrophysics*, 622, A100

Goicoechea, J. R. & Bourlot, J. L. 2007, *Astronomy & Astrophysics*, 467, 1

Gondhalekar, P. M., Phillips, A. P., & Wilson, R. 1980, *Astronomy & Astrophysics*, 85, 272

GRAVITY Collaboration, Abuter, R., Amorim, A., et al. 2019, *Astronomy & Astrophysics*, 625, L10

Guevara, C., Stutzki, J., Ossenkopf-Okada, V., et al. 2020, *Astronomy & Astrophysics*, 636, A16

Habing, H. J. 1968, *Bulletin of the Astronomical Institutes of the Netherlands*, 19, 421

Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, *Astronomy & Astrophysics Supplement Series*, 47, 1

Hauser, M. G., Arendt, R. G., Kelsall, T., et al. 1998, *The Astrophysical Journal*, 508, 25

Heithausen, A., Bensch, F., Stutzki, J., Falgarone, E., & Panis, J. F. 1998, *Astronomy & Astrophysics*, 331, L65

Heyer, M. & Dame, T. M. 2015, *Annual Reviews in Astronomy and Astrophysics*, 53, 583

HI4PI Collaboration, Ben Bekhti, N., Flöer, L., et al. 2016, *Astronomy & Astrophysics*, 594, A116

Hollenbach, D., Kaufman, M. J., Neufeld, D., Wolfire, M., & Goicoechea, J. R. 2012, *The Astrophysical Journal*, 754, 105

Hollenbach, D. J. & Tielens, A. G. 1997, *Annual Reviews in Astronomy and Astrophysics*, 35, 179

Indriolo, N., Fields, B. D., & McCall, B. J. 2009, *The Astrophysical Journal*, 694, 257

Kabanovic, S., Schneider, N., Ossenkopf-Okada, V., et al. 2022, *Astronomy & Astrophysics*, 659, A36

Kalberla, P. M. W. & Haud, U. 2015, *Astronomy & Astrophysics*, 578, A78

Kennicutt, R. C. & Evans, N. J. 2012, *Annual Reviews in Astronomy and Astrophysics*, 50, 531

Kerp, J., Winkel, B., Ben Bekhti, N., Flöer, L., & Kalberla, P. M. W. 2011, Astronomische Nachrichten, 332, 637

Kerr, F. J., Hindman, J. F., & Robinson, B. J. 1954, Australian Journal of Physics, 7, 297

Klessen, R. S. & Glover, S. C. O. 2016, in Saas-Fee Advanced Course, Vol. 43, Saas-Fee Advanced Course, ed. Y. Revaz, P. Jablonka, R. Teyssier, & L. Mayer, 85

Lancaster, L., Ostriker, E. C., Kim, J.-G., & Kim, C.-G. 2021, *The Astrophysical Journal*, 914, 89

Langer, W. D., Velusamy, T., Morris, M. R., Goldsmith, P. F., & Pineda, J. L. 2017, *Astronomy & Astrophysics*, 599, A136

Langer, W. D., Velusamy, T., Pineda, J. L., Willacy, K., & Goldsmith, P. F. 2014, *Astronomy & Astrophysics*, 561, A122

Larson, R. B. 1981, *Monthly Notices from the Royal Astronomical Society*, 194, 809

Le Petit, F., Nehmé, C., Le Bourlot, J., & Roueff, E. 2006, *The Astrophysical Journal Supplementary Series*, 164, 506

Leroy, A. K., Schinnerer, E., Hughes, A., et al. 2021, *The Astrophysical Journal Supplementary Series*, 257, 43

Li, A. & Draine, B. T. 2001, *The Astrophysical Journal*, 554, 778

Lin, C. C. & Shu, F. H. 1964, *The Astrophysical Journal*, 140, 646

Lin, C. C. & Shu, F. H. 1966, Proceedings of the National Academy of Science, 55, 229

Lopez, L. A., Krumholz, M. R., Bolatto, A. D., Prochaska, J. X., & Ramirez-Ruiz, E. 2011, *The Astrophysical Journal*, 731, 91

Lopez, L. A., Krumholz, M. R., Bolatto, A. D., et al. 2014, *The Astrophysical Journal*, 795, 121

Mac Low, M.-M. & Klessen, R. S. 2004, Reviews of Modern Physics, 76, 125

Marasco, A., Fraternali, F., van der Hulst, J. M., & Oosterloo, T. 2017, *Astronomy & Astrophysics*, 607, A106

Mathis, J. S., Mezger, P. G., & Panagia, N. 1983, *Astronomy & Astrophysics*, 128, 212

Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, *The Astrophysical Journal*, 217, 425

McClure-Griffiths, N. M., Pisano, D. J., Calabretta, M. R., et al. 2009, *The Astrophysical Journal Supplementary Series*, 181, 398

McKee, C. F. & Ostriker, E. C. 2007, *Annual Reviews in Astronomy and Astrophysics*, 45, 565

McKee, C. F. & Ostriker, J. P. 1977, *The Astrophysical Journal*, 218, 148

McLeod, A. F., Ali, A. A., Chevance, M., et al. 2021, *Monthly Notices from the Royal Astronomical Society*, 508, 5425

Meixner, M. & Tielens, A. G. G. M. 1993, *The Astrophysical Journal*, 405, 216

Mezger, P. G., Mathis, J. S., & Panagia, N. 1982, *Astronomy & Astrophysics*, 105, 372

Narayanan, D., Krumholz, M. R., Ostriker, E. C., & Hernquist, L. 2012, *Monthly Notices from the Royal Astronomical Society*, 421, 3127

Natale, G., Popescu, C. C., Rushton, M., et al. 2022, *Monthly Notices from the Royal Astronomical Society*, 509, 2339

Okada, Y., Yanitski, C. N., Kabanovic, S., et al. in prep., *Astronomy & Astrophysics*

Ossenkopf, V. 2004, in The Dense Interstellar Medium in Galaxies, ed. S. Pfalzner, C. Kramer, C. Staubmeier, & A. Heithausen, Vol. 91, 387

Ossenkopf, V. & Henning, T. 1994, *Astronomy & Astrophysics*, 291, 943

Ossenkopf, V. & Mac Low, M. M. 2002, *Astronomy & Astrophysics*, 390, 307

Pabst, C. H. M., Goicoechea, J. R., Hacar, A., et al. 2022, *Astronomy & Astrophysics*, 658, A98

Padovani, M., Galli, D., & Glassgold, A. E. 2009, *Astronomy & Astrophysics*, 501, 619

Padovani, M., Ivlev, A. V., Galli, D., et al. 2020, *Space Science Reviews*, 216, 29

Pineda, J. L., Langer, W. D., Velusamy, T., & Goldsmith, P. F. 2013, *Astronomy & Astrophysics*, 554, A103

Planck Collaboration, Abergel, A., Ade, P. A. R., et al. 2014a, *Astronomy & Astrophysics*, 571, A11

Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2016, *Astronomy & Astrophysics*, 594, A10

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014b, *Astronomy & Astrophysics*, 571, A1

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014c, *Astronomy & Astrophysics*, 571, A12

Planck Collaboration, Akrami, Y., Ashdown, M., et al. 2020, *Astronomy & Astrophysics*, 641, A4

Popescu, C. C., Tuffs, R. J., Dopita, M. A., et al. 2011, *Astronomy & Astrophysics*, 527, A109

Pound, M. W. & Wolfire, M. G. 2023, *The Astronomical Journal*, 165, 25

Ramambason, L., Lebouteiller, V., Madden, S. C., et al. 2023, arXiv e-prints, arXiv:2306.14881

Röllig, M., Abel, N. P., Bell, T., et al. 2007, *Astronomy & Astrophysics*, 467, 187

Röllig, M. & Ossenkopf, V. 2013, *Astronomy & Astrophysics*, 550, A56

Röllig, M., Ossenkopf, V., Jeyakumar, S., Stutzki, J., & Sternberg, A. 2006, *Astronomy & Astrophysics*, 451, 917

Röllig, M. & Ossenkopf-Okada, V. 2022, *Astronomy & Astrophysics*, 664, A67

Röllig, M., Szczerba, R., Ossenkopf, V., & Glück, C. 2013, *Astronomy & Astrophysics*, 549, A85

Schneider, N., Röllig, M., Polehampton, E. T., et al. 2021, *Astronomy & Astrophysics*, 653, A108

Schneider, N., Röllig, M., Simon, R., et al. 2018, *Astronomy & Astrophysics*, 617, A45

Schuller, F., Csengeri, T., Urquhart, J. S., et al. 2017, *Astronomy & Astrophysics*, 601, A124

Schuller, F., Urquhart, J. S., Csengeri, T., et al. 2021, *Monthly Notices from the Royal Astronomical Society*, 500, 3064

Strömgren, B. 1939, *The Astrophysical Journal*, 89, 526

Stutzki, J. & Guesten, R. 1990, *The Astrophysical Journal*, 356, 513

van de Hulst, H. C., Muller, C. A., & Oort, J. H. 1954, *Bulletin of the Astronomical Institutes of the Netherlands*, 12, 117

van Dishoeck, E. F. & Black, J. H. 1982, *The Astrophysical Journal*, 258, 533

Weingartner, J. C. & Draine, B. T. 2001a, *The Astrophysical Journal*, 548, 296

Weingartner, J. C. & Draine, B. T. 2001b, *The Astrophysical Journal*, 553, 581

Weingartner, J. C. & Draine, B. T. 2001c, *The Astrophysical Journal Supplementary Series*, 134, 263

Winkel, B., Kerp, J., Flöer, L., et al. 2016a, *Astronomy & Astrophysics*, 585, A41

Winkel, B., Lenz, D., & Flöer, L. 2016b, *Astronomy & Astrophysics*, 591, A12

Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, *The Astrophysical Journal*, 587, 278

Xia, J., Tang, N., Zhi, Q., et al. 2022, Research in Astronomy and Astrophysics, 22, 085017

Yuan, C. 1969, *The Astrophysical Journal*, 158, 871

Zucconi, A., Walmsley, C. M., & Galli, D. 2001, *Astronomy & Astrophysics*, 376, 650

APPENDIX A

Work station

All of the analysis for this thesis was completed at a standard computer in the institute, though with a bit more computational power[35]. The development of `kosmatau3d` in its current form began during the course of my doctoral project, and I have yet to implement funtionality to utilise distributed computing. Future plans for `kosmatau3d` can be found in Chapter VIII or at the bottom of `https://github.com/CraigYanitski/kosmatau3d`.

Table A–1: Computer specifications

Name	`pdr_hpc`
Operating System	Ubuntu 20.04
Architecture	64-bit
RAM	64 GB
Processor	Intel Xeon CPU 3.7 GHz
Threads	12

Future work using `kosmatau3d` will utilise as well computing clusters, but until then it has been incredibly useful to use the cores available on `pdr_hpc`.

[35]This was very generously lent to the office by Nicola Schneider.

APPENDIX B

KOSMA-τ grid assumptions

For the development of the current thesis, the same underlying grid from Andree-Labsch et al. (2017), `OrionBar2013`, was used for testing the models (the elemental abundances are given in Röllig & Ossenkopf-Okada 2022). It was assumed that all of the assumptions required for the `kosmatau3d` approach were satisfied. Namely this means that the larger, partially molecular clumps are embedded in a medium of smaller, atomic interclump medium.

This grid was later supplemented with another grid, `CMZ2015`, exploring the cosmic ray ionisation rate (used for the analysis in García et al. 2016). I have implemented this only when using the increased primary cosmic ray ionisation in §VI-7. It is worth noting that since this grid was computed to model the CMZ, it has higher metallicity than the `OrionBar2013`.

After the clump-averaged column-densities where implemented in `kosmatau3d` (however late in the project), we were able to approximate and examine the atomic and molecular gas mass in each voxel. From this it was determined that the mass in the smallest, most-diffuse clump in the grid is actually primarily molecular at low far-UV radiation ($1\,\chi_{\mathrm{D}}$). As shown in §VI-5, the KOSMA-τ clumpy we use for the interclump ensemble from `OrionBar2013` were molecular. By using instead the lower-density clumps from the `PDRTB2020`, the grid used by `PDRToolbox`, I was able to at least model the interclump ensemble as primarily atomic gas.

APPENDIX C

KOSMA-τ continuum

It is also theoretically possible to apply a "correction" to the dust continuum used in this work, though for the wavelengths used this has the opposite effect on our results. One can effectively steepen the slope by removing the modification in Li & Draine (2001). The correction factor I remove, according to the corrected silicate dielectric function, is,

$$f_{\rm DE} = \begin{cases} \left(1 + \beta \dfrac{\ln(\frac{\lambda}{250\,\mu{\rm m}})\ln(\frac{\lambda}{850\,\mu{\rm m}})}{\ln(\frac{850\,\mu{\rm m}}{250\,\mu{\rm m}})}\right) , & 250 \leq \lambda \leq 850\,\mu{\rm m} , \\ \left(\dfrac{\lambda}{850\,\mu{\rm m}}\right)^{\beta} , & 850 \leq \lambda \leq 10^4\,\mu{\rm m} , \\ \left(\dfrac{10^4\,\mu{\rm m}}{850\,\mu{\rm m}}\right)^{\beta} , & \lambda \geq 10^4\,\mu{\rm m} , \end{cases} \tag{C–1}$$

where β is the index by which I modify the long-wavelength dependence of the absorption cross section (while a value of $\beta = 0.4$ was used in Li & Draine 2001, to fit the *COBE*-FIRAS observations, for our purpose I would need to use $\beta = 0.5$). Figure C–1 shows $f_{\rm DE}$ follows the desired long-wavelength trend, however the correction at 550 μm does not improve the issue I note in Figure VII–7 (that is why I neglect the comparison to the *Planck* observations). Future

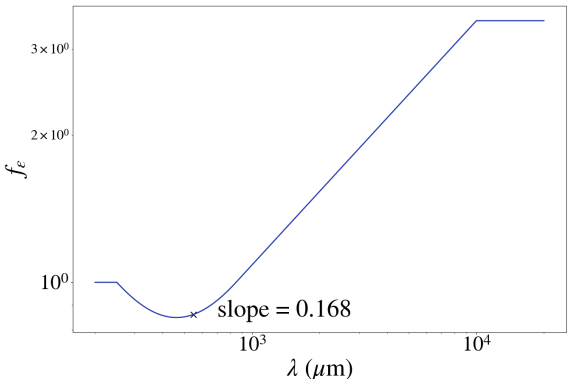


Figure C–1: $f_{\rm DE}$ over the relevant wavelength range.

modelling of the Milky Way using `kosmatau3d` should take this correction into account when planning the models and use wavelengths with $\lambda \gtrsim 10^3\,\mu$m for the dust continuum.

List of publications

Contributed references:

Yanitski, C. N., Ossenkopf-Okada, V., & Röllig, M. (2023). Constraining the Galactic far-UV field using a clumpy PDR model. *Physics and Chemistry of Star Formation: The Dynamical ISM Across Time and Spatial Scales. Proceedings of the 7th Chile-Cologne-Bonn Symposium*, ed. V. Ossenkopf-Okada, R. Schaaf, I. Breloy, & J. Stutzki, 265.

Okada, Y., Yanitski, C. N., et al. (2024), in prep. Dynamics of bright-rimmed clouds in IC 1396 and application of KOSMA-τ 3D.

First-author peer-reviewed references:

Yanitski, C. N., Ossenkopf-Okada, V., and Röllig, M. (2024), in prep.

Acknowledgements

The past four years at the Universität zu Köln have presented many new challenges. I arrived in a new country with no knowledge of the language and began a new project unrelated to my Master's degree, and there were many people who help me feel like this was home.

I would like to begin by thanking my supervisor, Volker Ossenkopf-Okada, for granting me this opportunity to join this working group and giving me freedom to develop `kosmatau3d` as I saw fit. It may have taken a long time to produce any papers (I think I recall him saying the Milky Way paper should take a couple of months...), but at least what we have developed should endure and produce much more science for the PDR group. I hope we manage to maintain a collaboration going forward, especially since I have a few ideas and collaborators interested in this model. His advice throughout my PhD, whether it was related to the science discussions or my project, has had a profound impact on how I approach each new problem I encounter.

I would also like to thank my Beisitzer, Markus Röllig, for all of his assistance helping me understand how does KOSMA-τ work and how may I interface with it from `kosmatau3d`. I apologise that it is taking so long for me to write his `Mathematica` routines in `python`; I swear I will start on it soon :-) I have fond memories of my three iterations as the teaching assistant for his Data Analysis course, even if two thirds of that needed to be given virtually.

I am grateful to Steffi Walch-Gassner not only for her guidance during my TAC meetings, but for agreeing to be a Gutachterin in my defence committee. Her expertise is definitely appreciated. I know my project was supposed to bridge the gap between our PDR modelling group and her MHD modelling group, so I am sorry if I failed on that front. Perhaps it will be easier to implement `kosmatau3d` as a subgrid model when time-dependence is implemented and tested in KOSMA-τ.

My Thesis Advisory Committee (TAC) has guided me and ensured that I did not waste too much time on unproductive areas of research. I thank the TAC members Volker, Steffi, and Frank Bigiel.

Someone who was instrumental in helping me settle when I first arrived in Germany was Bettina Krause. All of the logistics of handling contract extensions and travel notifications were processed so much easily because of her help.

I was fortunate to take join the student council during SFB 956 for two years, during which time I was able to organise workshops for coding and soft skills. I thank Shashwata, Mitra, Aleena, Stefano, Bettina, Ekaterina, Gerrit, and Parit for all of the support.

I had the privilege of founding and managing Astronomy on Tap with Aleena Baby in 2020. Although it was a rough start to launch this initiative during the pandemic (thanks again to Volker for being our first presenter!), I think we still managed to have a positive influence for our audience. There were times during the virtual events where it felt like we were the only ones watching, but we still managed to entice our audience to get engaged and ask questions. Of course this was much more obvious once we could begin hosting events in person (when ExPat Cafe was still open..). I thank Aleena, Roya, Parit, Shashwata, Daniel, Jonathan, Aaron, Kateryna, Isabelle Breloy, and everyone else who helped plan these events.

I have been given an immense amount of help from Slawa Kabanovic. We have had many fruitful discussions on science and life over the past three years. He is the only reason I was able to navigate the bureaucracy of the Arbeitsamt. He also came up with increasingly innovative ways to sabotage my thesis as I was trying to finish. I will never forget our encounter on the bus in Houston.

Cristian Guevara has given me many useful ideas of ways to test my model and compare with observations. I cherish all of the times he invited me for dinner with him and his wife, Carolina, and the conversations that lasted until well into the next morning. He was always gracious enough to allow me to sleep on the couch.

I spent many evenings discussing science and life in general with Juan Luis Verbena. Whether we are attending the open mic night at a local bar or he invites me to celebrate the holidays with his family, I always felt enriched by the experience. I look forward to our future collaboration with Jakob den Brok and everything we

I am grateful for the many discussions I had with Daniel Vieira regarding our ideas and projects outside of the doctoral work. It has been most illuminating to discuss these things over coffee.

Lastly I want to thank Aleena Baby. We started our projects around the same time, submitted our theses on the same day, and entered unemployment at the same time. Throughout the doctoral project it has been incredibly useful to discuss the problems we face with our projects or otherwise. I thank her for always being there for our various struggles during the doctoral time.

To Jürgen Stutzki, Dominik Riechers, Yoko Okada, Nicola Schneider, Robert Simon, Christof Buchbender, Ronan Higgins, and anybody else I forgot in our group, I am grateful for the all of the useful discussions we had about science, life, and German bureaucracy.

Erklärung

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäSS aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.“

Teilpublikationen:

Yanitski, C. N., Ossenkopf-Okada, V., & Röllig, M. (2023). Constraining the Galactic far-UV field using a clumpy PDR model. Physics and Chemistry of Star Formation: The Dynamical ISM Across Time and Spatial Scales. Proceedings of the 7th Chile-Cologne-Bonn Symposium, 265.

Craig Nicholas Yanitski
14.08.2023, Köln

Amendments

Between the initial submission of this thesis and its final publication, various aspects have been updated. Beyond the correction of general typographical and formatting errors, I have updated the RMS error I use for the following surveys,

- CfA
- Mopra
- ThruMMS
- SEDIGISM
- HI4PI

In my initial submission of this work, some of these surveys had a single error for a given transition line. In my recalculation of the RMS error, I calculate an error map (as a function of both longitude and latitude). As a result, some of the comparisons in Chapter VI have changed by more than scaling χ^2.

The figures that exhibit a different dependence in the parameter space compared to the initial version of this thesis are Figure VI–6, Figure VI–10, Figure VI–16, Figure VI–22, Figure VI–28, Figure VI–34, Figure VI–38, Figure VI–44, Figure VI–48, Figure VI–54, Figure VI–60, Figure VI–66, Figure VI–72, Figure VI–78, and Figure VI–84. The difference is most visible in Figure VI–38 and Figure VI–44 due to the small variation seen in these parameter spaces. Even though some of the trends have changed, the discussion and conclusions in Chapter VII and Chapter VIII are still valid.

Craig Nicholas Yanitski
21.12.2023