The Diffuse Interstellar Band "GAIA" : optimisation of its extraction and use with double stellar and interstellar goal

Letizia Capitanio

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'Observatoire de Paris

La bande interstellaire diffuse Gaia: optimisation de son extraction et de son utilisation

Soutenue par
Letizia CAPITANIO
Le 29/11/2018

École doctorale n° 127
Astronomie et Astrophysique de l'Ile de France

Spécialité
Astrophysique

Composition du jury :

Jacques, LEBOURLOT Titre, Observatoire de Paris	*Président*
Dieter, BREITSCHWERDT Professeur, Berlin Institute of Technology	*Rapporteur*
Giovanni, VLADILO Director INAF, Trieste Astronomical Observatory	*Rapporteur*
Annie, ROBIN Astronome, Institut Utinam	*Examinateur*
Arnaud, SIEBERT Astronome, Observatoire astronomique de Strasbourg	*Examinateur*
Rosine, LALLEMENT Titre, Observatoire de Paris	*Directeur de thèse*

l'Observatoire de Paris | PSL

Abstract

The Interstellar Medium (ISM) represents a small fraction of the Milky Way baryonic matter (typically 10%), but it plays a fundamental role because it governs the galactic evolution and its structure. Understanding this evolution requires the knowledge of its various phases, their physical parameters, temperatures, pressures, their compositions in ionised, atomic, and molecular gas as well as in dust grains, to better characterise their dynamical, physical and chemical interactions. Most of the ISM gaseous constituents have been identified through their emission or absorption characteristics, however, despite decades of efforts a category of interstellar absorptions, the "diffuse interstellar bands" (or DIBs), still has no counterpart in absorbing species. More specifically, an increasing number of these hundreds of weak optical and infrared bands is found to have internal structures that correspond to large organic molecules, but at the exception of the C60+ fullerene cation none of these molecular carriers has been identified yet. Moreover, nothing is clearly established about the life cycle of the DIB carriers, neither their formation, nor evolution or destruction. Because DIBs may represent a large fraction of the organic matter, it is mandatory to understand their role during star and planet formation, in particular their interactions with other species and grains, and this requires measurements of their distributions and preferential sites. Large surveys from ground or space have started to provide such statistical information on their spatial distribution and links with the other constituents. In particular, a serendipitously favourable situation is the presence of a DIB (with central wavelength 862 nm) within the spectral interval of the *Radial Velocity Spectrometer* (RVS), one of the instruments of the ESA Gaia astrometric mission. Thanks to the numerous stellar spectra that are currently (and will be) recorded with the RVS, one may expect the construction of a very large catalog of measurements of this DIB, and the first part of the work presented in this document is directly related to the construction of this future catalog and its potential use. More specifically, I have attempted to answer the following questions:

-How many DIB measurements can be potentially reliably extracted from the RVS spectra of individual stars?

-What will be the characteristics of this sample of targets?

-What kind of limitations will have this DIB survey?

-What range of distances will be potentially explored?

To answer these questions, I have developed a profile fitting code adapted to the *Gaia* DIB and to the RVS instrument. I have applied this adjustment code to a large series of synthetic RVS spectra obtained from the simulator GOG developed for the RVS, for different stellar types, star magnitudes and extinction

ranges, and I have merged the results with the statistics derived from the Gaia Universe Model Snapshot (GUMS), a Milky Way model adapted to the Gaia instrumental performances. The resulting information on the characteristics of the targets that are adapted to DIB measurements as well as their numbers is presented.

The second part of this thesis has been related to the three-dimensional (3D) mapping of the ISM, a goal motivated by the increasing number of large photometric stellar surveys as well as the availability of the Gaia parallax distances. In particular, I have focused on the potential use of DIBs for the mapping, here motivated by the advent of large spectroscopic surveys and the Gaia DIB catalog. Can we use these absorptions to chart the ISM? What are the advantages and disadvantages of such a choice? How to optimise the use of the catalogs of spectra, in particular the catalog of RVS spectra which will be available with the Gaia DR3 (or DR4)? To answer these questions I have followed several paths. In ran parallel inversions of two extinction and DIB datasets for the same target stars, both data derived from the SDSS/APOGEE survey, and compared the distributions. I made the first tomography based on composite data, both extinctions and DIBs and showed this is a viable way. The resulting map has been used by GEPI colleagues during a photometric calibration of the G band Gaia, a calibration they in turn used for the production of additional extinction estimates. I ran the inversion of all data to produce improved 3D maps. In both cases of mapping from extinction or DIB measurements, one encounters the same difficulty associated with the use of massive datasets. The full 3D tomographic inversion technique in this case requires an unrealistically large computing time. I have developed a technique called "hierarchical tomographic technique", an iterative inversion using the previous step 3D distribution as a prior for the next inversion, and where each new step corresponds to an increased spatial resolution in reduced volumes that depend on the target distribution density. I have applied this technique to a new catalog of 30 millions of extinction measurements based on Gaia and 2MASS. Preliminarily to this development I have tested the used of an arbitrary, complex 3D prior distribution for the inversion based on Pan-STARRS extinction results. Finally, I used high resolution ground-based stellar spectra to extracted through profile-fitting radial velocities and columns of IS absorbing species in several specific regions and participated to the first step of a "kinematic" 3D map of the ISM based on these data, radio emission and 3D dust maps.

Abstract

Le milieu interstellaire (MIS) représente une fraction mineure de la matière baryonique de la Voie Lactée (environ 10%), mais joue un rôle fondamental parce qu'il gouverne l'evolution de la Galaxie en contenu et structure. Comprendre cette évolution nécessite la connaissance de ses diverses phases, ses paramètres physiques, température, pression, sa composition dans le gas ionisé, atomique et moléculaire, et de la même façon dans les grains de poussière, pour mieux caractériser ses interactions dynamiques, physiques et chimiques. La majorité des constituants du MIS ont été identifiés grace à leurs émissions ou absorptions, mais une catégorie d'absorptions interstellaires, les "bandes interstellaires diffuses" (DIBs), n'ont pas encore de particules identifiées comme responsables des absorptions, malgré des dizaines d'années dédiées à leur étude. Un nombre de plus en plus élevé de bandes parmi les centaines détectés des en optique et infrarouge présente des profils spectraux qui correspondent à des larges molecules organiques, mais aucune molécule nà été encore identifiée, hormis le cation du fullerene C60+. Rien n'est vraiment établi sur le cycle de vie des porteurs de DIBs, ni leur formation ni leur évolution et destruction. Comme les DIBs pourraient représenter une large fraction de la matière organique, il est fondamental de comprendre leur rôle dans la formation stellaire et planétaire, en particulier leur interaction avec les autres espèces et les grains; pour atteindre cet objectif il est nécessaire de mesurer leur distribution et leurs sites préférentiels. Les grands relevés sol soit spatiaux ont commencé à apporter des informations statistiques sur la distribution spatiale des porteurs de DIBs et aussi sur leurs liens avec d'autres constituants. Par un heureux hasard, une DIB dont la longueur d'onde centrale est 862 nm est présente dans l'intervalle spectroscopique du Radial Velocity Spectrometer (RVS), un des instruments de la mission astrométrique Gaia de l'ESA. Grace aux spectres stellaires en cours de mesure avec le RVS, on s'attend à obtenir un large catalogue de mesures de DIBs et la première partie de ma thèse est directement relié à la construction de ce futur catalogue et son usage potentiel. Plus précisément, j'ai exploré les questions suivantes:

-combien des mesures de DIB seront potentiellement extraites des spectres RVS individuelles?

-quelles caractéristiques aura cet échantillonnage de DIB?

-quelles limitations aura cette survey des DIB?

-quel range des distances pourra être exploré?

Pour répondre à ces questions, j'ai développé un code de profile fitting adapté à la DIB "Gaia" et à l'instrument RVS. J'ai appliqué ce code à une large série de spectres stellaires synthétiques RVS obtenus avec le simulateur GOG développé expressément pour le RVS, pour différents types stellaires,

magnitudes et intervalles d'extinction. J'ai comparé les résultats avec des statistiques dérivées du Gaia Universe Model Snapshot (GUMS), un modèle de la Voie Lactée adapté aux performances instrumentales GAIA. Les informations sur les caractéristiques des étoiles cibles ont été adaptées aux mesures de la DIB. Je présente dans la thèse une estimation des DIB mesurables pour chaque ensemble de paramètres stellaires et observationels simulés, en fonction de la qualité de l'ajustement. La second partie de la thèse concerne les cartes en trois dimensions du MIS, un objectif motivé par l'augmentation des grands relevés photométriques et spectroscopiques stellaires, et par la disponibilité des parallaxes Gaia. En particulier, j'ai commencé à étudier l'usage potentiel des DIBs pour la cartographie. Peut-on utiliser ces absorptions pour cartographier l'ISM? Quels sont les avantages et désavantages de ce choix? Comment optimiser l'usage des catalogues des spectres, en particulier le catalogue des spectres du RVS qui seront disponibles avec la DR3 (ou DR4) Gaia? Pour répondre à ces questions j'ai suivi différents chemins. J'ai fait tourner en parallèle des programmes d'inversions des données des extinctions et des DIBs pour les mêmes cibles stellaires, les deux dérivés du relevé SDSS/APOGEE et comparés les distributions inversées obtenues. J'ai fait la première tomographie basée sur des données composites, extinctions et DIBs et j'ai montré que c'est praticable. La carte résultante a été utilisé par des collègues du GEPI pour aider à la calibration photométrique de la bande G de Gaia, une calibration qu'ils ont utilisée pour la production de l'estimation de nouvelles extinctions supplémentaires. Ces extinctions ont été utilisés ensuite pour produire et améliorer les cartes 3D. Dans tous les cas, cartes d'extinction ou mesures des DIBs, on rencontre les mêmes difficultés associées aux des dataset massives. La technique de l'inversion du 3D tomographic dans ce cas demande un temps de calcul non realiste. J'ai développé une technique appelée "technique tomographique hiérarchique", une inversion itérative qui utilise la distribution 3D de l'étape précédente comme prior pour l'inversion suivante, et où chaque étape nouvelle correspond à une résolution spatiale plus grande. J'ai appliqué cette technique aux nouveau catalogue de 30 millions des mesures d'extinction basé sur Gaia et 2MASS. Comme préliminaire à ce développement, j'ai testé un prior complexe dérivé des profils d'extinction obtenus à partir de Pan-STARRS. Finalement, j'ai utilisé des spectres à haute résolution stellaire obtenus au sol pour extraire soit les vitesses soit la colonne des matériaux interstellaires absorbants avec un programme de profile-fitting et j'ai participé aux premiers étapes d'une carte cinématique 3D du MIS basée sur ces données, sur les émissions radio CO et les cartes 3D des poussières.

In memory of Giulio Regeni
TRUTH FOR GIULIO REGENI.

Contents

Part I

Thesis context

Chapter 1

Interstellar Medium general properties

1.1 Interstellar Medium environments

The interstellar medium is an astronomical object with variable characteristics, but it can be defined as non-homogeneous mixture of gas-phase baryonic matter, dust, cosmic rays (interstellar matter) and magnetic field and radiation (electromagnetic field). It is not simply an environment where stars float, but the physical processes that occur create, with stars, the galactic ecology.

The mass of the Milky Way is 10-15% composed of interstellar matter - atomic gas, molecular gas, dust -, concentrated on the disk, with small structures highly fractal and non-homogeneous; 50% of the interstellar matter is organised into cloudy structures, which occupy ∼1-2% of the interstellar volume (Karttunen et al., 2003; Ferrière, 2001). Clouds structures could be classified as dark, cold and translucent clouds and matter in more diffuse interstellar medium is in warm gas, warm ionised gas or ionised gas. Different clouds propriety are summarised in Tab 1.1.

I present here a general description of the interstellar medium components and in Chapter 1.2 I am going to describe the measurement techniques.

Component	Temperature [K]	n $[cm^{-3}]$
Molecular	10-20	$10^2 - 10^6$
Cold Atomic	50-100	20-50
Warm Atomic	6000-10000	0.2-0.5
Warm Ionized	∼8000	0.2-0.5
Hot Ionized	10^6	∼0.0065

Table 1.1 – Mean temperatures and particles density in various interstellar medium environments (Ferrière, 2001).

Dust

Dust is the solid component in interstellar matter, composed in amorphous or regular crystals of carbon, silicon, ice and others impurities - depending on the environment and the dynamical time of formation. There are some space missions which have collected directly interstellar dust grains in our Solar System, most recent Cassini Mission (Altobelli et al., 2016), but only non-charged or larger particles can enter in the heliosphere. Dust is very important for stellar cycle, because dust coagulation is the beginning for the gravitational contraction in dark clouds. Dust origin is not completely understood, but there are many identified formation environments: outer atmospheres of red giants and supergiants, and planetary nebula, but also in the ISM itself there are favourable condition for formation, as in the high density clouds.

From the observations, astronomers have deduced the distribution of the size of the interstellar dust grains. Usually dust grains are described as ranging in size from nanometers to micrometrics, but the steps of coagulation from micrometers and planetary objects (kilometre) are still unknown. Macromolecules like polycyclic aromatic hydrocarbons (PAHs), which have nano-metric dimensions, are considered as components of the dust, because they could represent the pre-coagulation steps in ISM. In the Milky Way, the dust is concentrated in the galactic disc, with a scale height of ~ 120 pc (Kos et al., 2014).

Physical processes involving dust and light are:

- absorption and re-emission: the photon energy is transformed into heat, which is then re-radiated at the infrared wavelengths corresponding to the temperature of the grains particles
- scattering: light trajectory is changed, therefore the intensity on the original direction is reduced

Gas phases

Gas is the main component of interstellar medium, but it is more elusive because it can be measure only by spectrometric observations, see Chapter 1.2. Hydrogen in different states - H, H^+ and H_2 - is the main component (70%), followed by He (28%) and the other elements (2%). Heavy elements are measured as significantly less abundant than in the Sun and other stars of population I; since the global composition in the interstellar medium would be regular, the interpretation is that heavy elements are embedded in the dust grains, so they can no longer be measured in the absorption lines.

Many atoms are ionised in the interstellar space, because the density of interstellar matter is very low and free electrons cannot be easily captured; so ionisation states due to the UV radiation from stars or cosmic rays do not come back easily at the neutral state.

Atomic gas Atomic gas structures have wide variety depending on the physical conditions. Atomic hydrogen is observed in UV band with Lymanα absorption (n=2, 121.6 nm) and in radio band with the hyperfine suppressed transition at 21 cm. The H^+ regions usually are ionised by the UV radiation, and we measure them in the star formation regions or in OB complexes.

The most common metals C, O, N are depleted by a factor 1.2-3, while other as Mg, Si and Fe have a 10 - 100 depletion factor in comparison to the stars, so they are supposed to be in dust grains, but it is difficult to measure directly the dust composition.

Molecular gas Molecular gas has different states, from simply bi-atomic gas (i.e: H_2, CO, CH) to more complex poly-atomic molecules. Bi-atomic hydrogen forms by recombination of atomic hydrogen, having dust grains as chemical catalysts. They can be stable only isolated from UV and in sufficiently cold environment, because of the their molecular inner energy, to avoid collisional dissociation; for these reasons their environment are the cores of darks clouds.

Thanks to the ease in measurement, it is the CO emission - in equilibrium with the cosmic rays heating - which provides us the observed temperature, estimated 10-20 K, but probably lower in denser cores.

Molecular gas, measured in dark clouds, is confined in Milky Way disk as inferred from different studies. The total mass is estimated as 1.3-2.5 10^9 solar mass (Ferrière, 2001).

Molecular gas structures are hierarchically organised in interstellar medium, with different scales from tens to few tenths of parsec. Larger structures have lower density as $\sim 100 - 1000\text{cm}^{-3}$, smaller ones can contain billions of molecules per cubic centimetre.

Diffuse Interstellar Bands

Special absorptions from interstellar medium are Diffuse Interstellar Bands (DIB) - an example is shown in Fig. 1.1. There are a few hundreds of large absorptions in optical bands and some tens in the infrared band (Hobbs et al., 2008), but new DIBs are going to be discovered thanks to the ESO Diffuse Interstellar Bands Large Exploration Survey (EDIBLES) (Cox et al., 2017). Their carriers are not identified, but for the most solid theories they are macro-organic molecules (Jones et al., 2013; Jones, 2016). Sometimes we use the acronym DIBs both for the absorptions and for the unknown carriers: in this manuscript I will refer to DIB only for absorptions.

In the last few decades, some near infrared DIB carriers have been partially identified, such as the fullerene $C_{60}{}^+$ (Foing and Ehrenfreund, 1994), and these have been confirmed in the laboratory experiments in gas phase (Campbell et al., 2015). This detection suggests that PAHs (Poli Aromatic Hydrocarbons) could be good candidates for DIBs carriers and in general for interstellar medium environmental (van der Zwet and Allamandola, 1985; Salama et al., 1996).

DIB carriers have different regimes; for example there is a basic categorisation that distinguishes DIBs in irradiated regions an (σ -type) and others (ζ -type). The discriminator factor is the ratio between the Equivalent Width (EW) of the DIB at 5797 Åand 5780 Å. Lines of sight with lower ratio values are classified as σ -type, while those with a high ratio are considered ζ -type, (Cami et al., 1997; Maíz Apellániz et al., 2015).

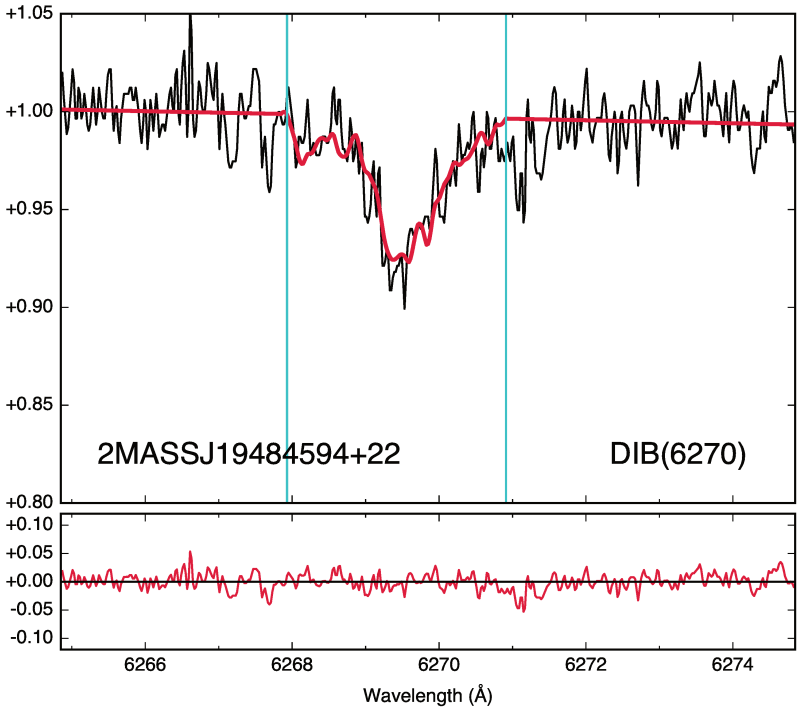


Figure 1.1 – Example of fitted DIBs absorption on normalised spectrum.

Other components

Magnetic field

Electromagnetic fields and cosmic rays permeate the interstellar medium. Magnetic fields very likely originated before and/or the formation of the Galaxy, and were magnified during collapse and later due to differential rotation and turbulence (Ferrière, 2015). The direction of the magnetic field direction can be measured by the polarisation of the stellar light in optics, a polarisation due to the anisotropy of the dust grains and their rotation, because during this rotation there is a preferential rotation around the sort axis and an alignment of the short/rotation axis along the field, affecting the stellar light absorption. Due to this alignment, dust grains also emit a polarised light. Recently Planck made full-sky observations of the dust emission at 353 GHz and provided us with a global map of the Galactic magnetic fields.

Three physical effects can be used to measure the strength of the magnetic field:

- the Zeeman splitting $\Delta\nu$ in the atomic emissions is proportional at the magnetic-field strength B. On the HI 21 cm line, $\Delta\nu$ usually is too small to be measured, and what is measured is the difference between the two circularly polarised components, that is $B_{||}$. However such measurements are difficult and scarce.

- the Faraday rotation of the linearly polarised radio light. In this case the

measured rotation is additionally depending on the column of electrons.

- the radio synchrotron emission from relativistic electrons. In this case the measured rotation is additionally depending on the column of cosmic-ray electrons. To obtain the magnetic field from this emission it is mandatory to break the degeneracy with the spectrum of cosmic-ray electrons (Rockstroh and Webber, 1978).

Cosmic rays

The cosmic rays are energetic particles, composed by protons, 10% alpha particles, 1% heavier nuclei, 2% electrons, and smaller amounts of positrons and antiprotons (Blandford and Eichler, 1987). The supernova explosions are supposed to be the main source of cosmic rays, as recently confirmed by measurements of the Fermi satellite, but other mechanisms are also proposed because the only supernovae are not enough to account for all energy ranges, as for example ejections from active galactic nuclei. Measurements of these particles were very important for the development of the Standard Model of particle physics, for example muons, pions and positrons were identified using cosmic rays showers. The low energy part of the cosmic rays (on the order of hundred of MeV or lower) is very important since it is responsible for most of the interstellar gas ionization in the dense molecular cores. Since the ionised fraction of the gas is the only one to be coupled to the magnetic field, and the field plays an essential role during the cloud collapse, low energy cosmic rays are an important actor in star formation. Along the Galactic plane there is an approximate equipartition between the gas thermal pressure, the magnetic pressure and the total cosmic rays pressure. The fraction of pressure associated with the cosmic rays increases with the distance from the Galactic mid-plane due to their escape to the halo.

1.2 Distance - limited interstellar measurements

The interstellar medium is widely observed thanks to the lack of light, because, with the exception of some emissions in the infrared and radio bands, the main effect of dust and atomic/molecular gas is the disappearance of photons emitted by stellar sources. It is fundamental to identify and quantify the matter that - laying between the source and the observer - affects the light in the path to the observer. This is the basic characteristic of the measurements used for the interstellar medium maps, because knowing the spatial segment where the light has been modified it is possible to obtain the position of the agents responsible for the modification, that is the ISM.

Extinction and reddening

Considering the space from a bright source to the observer completely empty, the difference from the real brightness of a source to the observed one is due only by distance from the source and the observer:

$$M - m = 5 \log \frac{d}{10\text{pc}} \tag{1.1}$$

But when the space is not empty, the interstellar medium modifies the light flux. The processes that take place between dust grains and light are essentially absorption and diffusion; these two processes subtract photons from the lines of sight. Absorption is followed by temperature-dependent re-emission, but since the photon is re-emitted in a random direction, there is a net light loss on the line of sight. The dispersion is very efficient from grains with dimension similar to the optical wavelength and it is the main cause of photon loss. In theory, gas diffusion could also contribute to light-subtraction, but it is not an efficient process, since the density is too low to have a high impact (Karttunen et al., 2003).

As consequence, taking in account the presence of dust, the difference between emitted and observed magnitude is increased by a factor $A > 0$, that is the *extinction* due to the medium.

$$M - m = 5 \cdot \log \frac{d}{10\text{pc}} + A(d) \tag{1.2}$$

Scattering efficiency is size-grain depending, thus A is not a constant, but extinction is stronger for shorter wavelength, producing in optical bands the so called *reddening*. For this reason in optical band more distant stars seem to have a (B - V) index higher than the intrinsic for their own stellar type.

$$E(B-V) = (B-V)_{observed} - (B-V)_{intrinsic} \tag{1.3}$$

or more general for two bands:

$$E(B_1 - B_2) = (B_1 - B_2)_{observed} - (B_1 - B_2)_{intrinsic} \tag{1.4}$$

The empirical law connects the E(B - V) - called *excess of colour in V band* - with extinction in optical band, Av, and it is usually expressed as a proportion:

$$R_V = \frac{Av}{E(B-V)} \approx 3.1 \tag{1.5}$$

Important studies of this relationship were done for different bands, to describe R_λ - the milestone is the Cardelli et al. (1989) empirical extinction law. Next studies modified this law, and in the last years there are was important improvement in this branch thanks to Planck Collaboration et al. (2014) dust map results, which show important variability of the relationship at different latitude.

Measurements of the extinction curve show some features (Fig 1.2), the most famous one is the UV - bump, whose origin is still under investigation (Gavilan et al., 2017; Berruto et al., 2018) and probably depends on peculiar components or grains dimension.

The proportion between extinction (or excess of colour) and dust column is possible only supposing strong hypothesis about the dust nature, shape and dimension. It is presumed a general uniformity in dust composition and reaction at the scattering, but it is always true that extinction us always correlated with dust presence.

Extinction changes for directions and distances, depending on the dust concentration in the line of sight: for example dense dust clouds in the Galactic centre direction produce an optical extinction index of 30 mag, which makes impossible direct observations in the optical band.

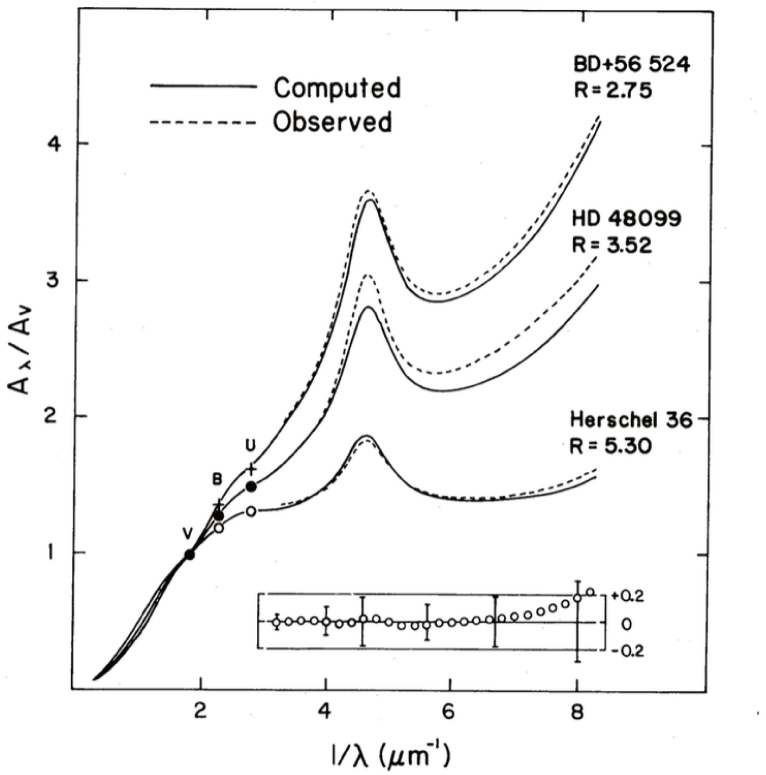


Figure 1.2 – Observed extinction curves and empirical extinction law as in Cardelli et al. (1989). The main peak is the famous UV-bump.

Extinction representation In schematic standard model for extinction, light intensity decreases exponentially during the optical path from **r1** to **r2** depending on the medium opacity λ.

$$I = I_0 \exp^{-\tau_\lambda} \tag{1.6}$$

I is the intensity after the absorption, I_0 the non absorbed intensity, τ_λ is the optical depth: this value takes into account the whole absorbing medium - its absorbing coefficient k_λ, its absorbing particles density ρ and its distribution along the optical path. In the case of interstellar medium absorption, the absorption is due to the dust, and k_λ depends on the dust composition, dimension, shape and polarisation, which are not homogeneous along a line of sight:

$$\tau_\lambda = \int_0^R k_\lambda(r)\rho(r)dr \tag{1.7}$$

The flux at the point **r** depends on the distance from the emitting source in **r0** and absorption on the optical path. Absorption adds a degeneracy in measurements if we use the brightness as distance proxy, but we can break the degeneration using the excess of colour instead of the generic absorption. Integrated flux in a specific band degrades uniformly for each wavelength with

distances, but the absorption law depends on the photon energy and the cross section of the dust grains. Measuring the degradation for two different wavelengths and knowing the emitted spectral curve of the star, we can resolve the degeneracy:

$$E(\lambda_1, \lambda_2) = C \int_0^R (k_{\lambda_1}(r) - k_{\lambda_2}(r))\rho(r)dr \tag{1.8}$$

As far as we cannot resolve the dust density and proprieties degeneracy in Eq 1.7, the excess of colour becomes:

$$E(\lambda_1, \lambda_2) = C \int_0^R \theta_{\lambda_1,\lambda_2}(r)dr \tag{1.9}$$

with $\theta_{\lambda_1,\lambda_2}(r)$ a global variable for differential opacity, expressed in magnitude per pc units. Observed $E(\lambda_1, \lambda_2)$ is a variable which depends only on the distance from the source and the opacity, it means that it contains the information about the global presence of matter on the line of sight, but it not depends on the matter distribution along the line of sight.

Atomic and molecular absorptions

The gas does not produce a generic light extinction on one band, but we can observe it by their absorption lines in UV and optical range. The central position of the absorption line (frequency ν or wavelength λ) depends on the type of particles present in the gas; the shape and depth of a line depends on the physical characteristics of the global medium, i.e. temperature and optical depth. If we schematise the interstellar medium as a gas composed by homogeneous clouds, the cross-section between atoms and light is constant and the optical depth is proportional to the cross-section, σ_ν, and the number of absorbers:

$$\tau_\nu = \sigma_\nu N \tag{1.10}$$

So, it could be sufficient to measure the absorption curve:

$$\phi(\nu) = \sigma_\nu \frac{m_e c}{\pi e^2} f \tag{1.11}$$

where m_e is the electron mass, c is the speed of light and f is the oscillator strength specific for each absorption.

An absorption line can be schematised as three different regimes, depending on the number of encountered particles: optical thin, saturated regime, damped regime. An empirical quantity used as physical units is the Equivalent Widht (EW), the width of a rectangle below the normalised spectrum having the same surface as the one comprised between the line profile and the continuum. The curve that describes the variation of EW depending on the optical depth is called the curve of growth. For various regimes, the curve follows various functions:

- **Optical thin regime:** the Equivalent Width is proportional to the number of absorber particles

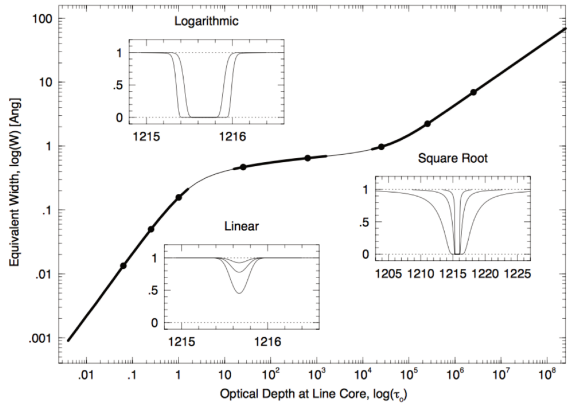


Figure 1.3 – Framework of curve of growth. For small τ, the curve approachs a straight line.

- **Saturated regime:** the Equivalent Width follows the logarithmic of number of absorber particles
- **Damped regime:** the Equivalent Width follows the square root of number of absorber particles

Some atomic lines can provide information about the total gas - quantity: for example NaI usually lies in neutral interstellar regions and knowing the column of NaI carriers it is possible to estimate the HI column (Hobbs, 1978; Lallement et al., 2003). To make a good fit profile for the interstellar lines, it is mandatory to have good stellar models, because the stellar lines affect the quality of the profile of the IS line. In hot stars the stellar lines are less important and packed, the stellar continuum can be sometimes modelled as a polynome for the interstellar goals, and it makes simpler the profile.

Since interstellar absorptions contain information about metallicity, kinematic, temperature and optical depth of the medium, an interstellar medium study using the spectral informations is very substantial and interesting, but gas absorption in interstellar medium goes rapidly in the saturated and damped regimes, therefore it is more complicated get quantitative informations; atomic lines have been used for accurate solar neighbour maps, see Chapter 2.2.

DIBs

DIBs absorptions follow the same mechanism than the gas, but they are not strong absorptions, and it means that DIBs do not reach the saturated and damped regime. For this reason, these absorptions are particularly interesting to derive information on interstellar matter from spectral information for distant objects, because atomic lines usually saturate beyond one dense clouds.

Because of their broad absorptions, for long time DIBs profile were mainly derived in the spectra of hot stars, whose smoothed continuum does not need high sophisticated or precise models, but recently more and more complete stellar catalogs were used for DIBs extractions, both in optical and infrared

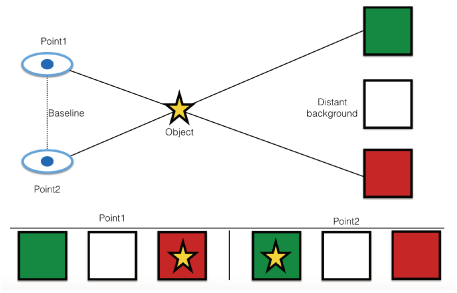


Figure 1.4 – Parallax principle, simplified.

bands (Chen et al., 2013; Puspitarini et al., 2013; Zasowski et al., 2015). In the case of early type stars, it is very important a precise determination of stellar parameter to distinguish stellar from interstellar features.

As their carriers are unknown, we are not able to connect the absorption to the number of absorbing particles, but starting from theoretical shapes we can describe DIBs variations; it is largely demonstrated the existence of a correlation between DIBs absorption and excess of colour (Puspitarini et al., 2015). DIBs shapes are empirically derived by high resolution observations (Chen et al., 2013); where an empirical shape is not available, a gaussian shape is usually attributed, because of its breadth (Elyajouri et al., 2016).

1.3 Distances measurements

Measuring distances is a fundamental goal in astronomy, and there many techniques adapted to the different objects. For galactic objects there are three main techniques: parallaxes, photometric distances and standard candles, as maser for the interstellar medium.

Trigonometric parallaxes

If we observe an object from different points, we see it in different direction. This angular difference is called *parallax* and it is a natural mechanism for measuring distances because the parallax depends on the distance. For astronomical objects, we use as the orbital movement of the Earth around the Sun, and in particular the *annual parallax* is the angle subtended by the radius of our orbit, or better a Astronomical Units - since the observer (us) orbit is elliptical. Since the parallaxes are completely independent on stellar population models or other stellar hypothesis, they can be used as distance probe, but their application is distance-limited. Sure enough, our instruments can variate the observational base-line in the order of $\sim$1 AU and for more distant objects the angular difference needs the use of very sensible instruments to be detected.

Estimate distances from parallaxes If the parallax is given in arc seconds, the distance in parsec is:

$$1pc = \frac{1}{\pi} =\sim 3.26 \text{ light-years} \tag{1.12}$$

The equation 1.12 needs special attentions, because it is valid only in case of parallaxes measured without uncertainty. Since the measured parallaxes in the astronomic catalogs have associated errors, the distance determination is more complicated (Bailer-Jones, 2015). In particular, using the simpler transformation:

$$d_{min}(pc) = \frac{1}{\pi + \sigma_\pi} \quad \text{and} \quad d_{max}(pc) = \frac{1}{\pi - \sigma_\pi}$$

we pass from a gaussian and symmetric probability distribution in the parallaxes errors space to an asimmetric distribution in the real distance space. Moreover, some negative parallaxes belong to the catalogs, but they are not outlier, but they contain statistical information about distances, but if we apply the transformation, we get a nonsense negative distance. Rejecting all the negative parallaxes add some bias in the used sample. Since these problems are more and more important for distant objects with parallaxes errors larger then 20%, and we use for our analysis only objects with distances with small errors, in this thesis I widely applied the naif transformation.

Available catalogs for parallaxes

For parallaxes, main catalogs come from *Hipparcos* and Gaia astrometric missions - see Chapter 3.1 for more detailed description of the Gaia mission.

Hipparcos This European mission was launched in August 1989, and it observed sky for 3.5 years, until March 1993. Hipparcos catalog contains 118 218 stars, but comparing the Hipparcos measurement to others previous less precise catalog, the Tycho Catalogue of 1 058 332 stars and Tycho 2 Catalogue of 2 539 913 stars were finally compiled. Last catalog was released in 2007, Hipparcos 2 (van Leeuwen, 2007).

Gaia DR1 Gaia DR1 was available in 2016 September and it is based on the TGAS catalog (Arenou et al., 2017a). Parallaxes for ~ 2 millions objects were available, therefore it was not a complete survey and it contains some bias. This release missed some objects in very dense regions - above 400 000 sources per square degree - where only brighter objects were observed; many bright stars with magnitude G < 7 were lost in DR1 and also objects near brighter objects could be missing. Colour dependent and spatially correlated systematic errors occurs at the level of ± 0.2 mas.

Gaia DR2 Since 2017 April, second data release is available, improving the first one in several critical points. The completeness in this case is respected for sources in G = (12;7) interval, but brighter stars and objects near bright sources are still partially lost. Denser regions have always some selections in magnitude, in particular the magnitude limit of Gaia DR2 is as bright as G = 18. More than 1 billion of parallaxes are now published, with associated uncertainties. These errors are known to be underestimated for G>16 source by a factor of 10% and for G<12 by a factor of 30%. Some systematic are present depending on celestial position, magnitude and colour, and Arenou et al. (2018) estimated them to be lower than 0.1 mas.

Photometric distances

The photometric distances are based on some presupposition about the stellar type and the extinction. Basically they are based on the Eq 1.2, and there are many strategy to determine the absolute magnitude M and extinction A. A calibration for M can be done by standard candles as RR Lyrae, Cepheids or Red Clump stars. Extinction can be estimated at the same time that distance or $A(d)$ could be assumed as know, for example extracted from the 3D maps.

Photometric estimations are all model depent, at different levels, and they benefit from independent distances measurements at nearer distances, as trigonometric parallaxes.

Red Clump technique - APOGEE/Gaia calibration In the post-main sequence area on HR diagram, the Giant Branch, the Red Clump stars have temperature between 4500 K and 5300 K; they are characterised by a low mass He-burning core and usually higher metallicity than the Horizontal Branch. Since all stars have the helium flash at the same mass of the degenerated helium core (0.5 solar masses), their luminosity does not vary a lot with metallicity, temperature or radius: they can be used as standard candles in the Milky Way (Girardi, 2016).

During my thesis we collaborated with Ruiz-Dern and Danielski to improve the Gaia photometric calibrations, with a virtuous iterative exchange between extinction maps and distance - extinction couples estimated from stellar catalogs (Danielski et al., 2018; Ruiz-Dern et al., 2018).

Standard Candles - Maser

Standard Candles are in general emitting points with regular luminosity, thus the observed luminosity depends only by the distance and the interstellar medium medium. For the ISM maps, we use point-like emissions which are directly related to the dense clouds, the Maser sources.

Maser (Microwave Amplification by Stimulated Emission of Radiation) is a particular kind of emitted, coherent and monochromatic light; it occurs in population inversion, when en excited state contains more atoms then the ground state. Typically the excited state is *metastable*, with a long average life, so spontaneous emissions are negligible and this gives rise to coherence.

In the interstellar medium there are some molecular clouds where it is measured a maser emissions from OH molecules, H_2O and SiO; since they are very small regions (5-10 AU), we can consider point-like sources. They are generally in HII regions, or near Mira variables or some supergiant stars. The positions of the Maser sources can be used directly to identify the spiral arms, star formation regions and molecular clouds, because they are associated with short-lived objects and this means that they do not move away from the formation site.

Fig 1.5 representes the maser positions for the Local Arm maser from Lépine et al. (2017), based by the Bar and Spiral Structure Legacy Survey2 key science project, the European VLBI Network3, and the Japanese VLBI Exploration of Radio Astrometry project.

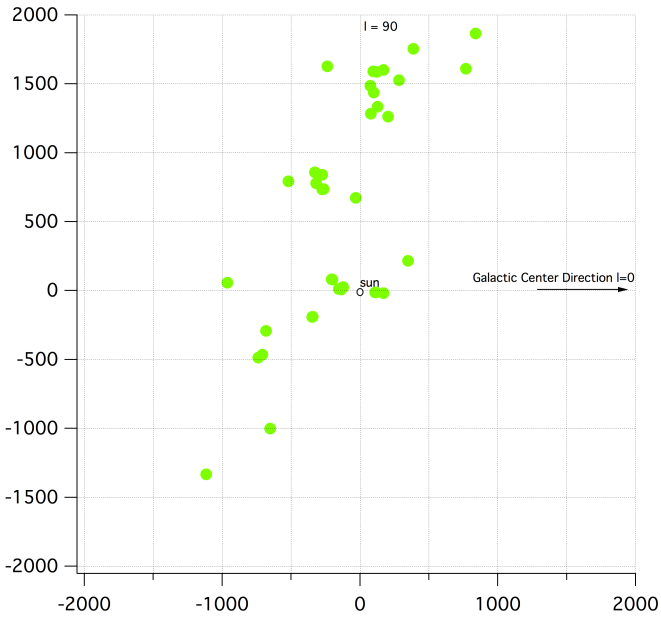


Figure 1.5 – Maser source emission position on the galactic plane for Local Arm (Lépine et al., 2017). The Sun is is (0,0), the x axis is the Galactic Center direction, distance units are in parsec.

Chapter 2

Mapping the Interstellar Medium

Mapping the spatial distribution of objects in the Universe is the oldest and main objective of astronomy. The ISM is diffuse matter and its volumes are not well defined, there are different types of plasma that require different techniques of observation. The mapping of diffuse matter is complicated, and often involves the measurement of starlight.

Three-dimensional (3D) maps of the Galactic interstellar matter (ISM) are useful in several ways. First, they provide observations to be compared with evolutionary models of the interstellar matter under the action of stellar winds and supernovae. Such models (see e.g. de Avillez et al. (2015)) will help explaining star-ISM feedback processes that are not yet well understood.

They are also useful for evolutionary models of Milky Way stellar populations. Evolutionary models require measurements of the spatial distribution of massive amounts of stars as well as their dynamical, physical and chemical properties, all quantities currently provided by the ESA satellite Gaia (Gaia Collaboration et al., 2016a,b) and ground-based surveys. In this context, three-dimensional (3D) maps of the Galactic interstellar dust are an additional and mandatory tool since they allow de-reddening of stellar spectra and suppression of degeneracies between star temperature and dust reddening on the one hand, and, on the other hand, since they may shed important additional light on the star formation through the inter-play between star and interstellar matter. Fortunately, Gaia parallaxes and Gaia photometric data complemented by ground-based photometric data allow the construction of the required IS dust maps in parallel with the stellar studies.

Finally, 3D ISM maps are a general tool for various purposes: studies of foreground, background or environment for specific Galactic objects, modelling of photon or high energy particle propagation, estimates of integrated Galactic foregrounds, and finally, as also mentioned above, of particular importance today in the context of Gaia, calibration for the reddening of the Milky Way objects.

2.1 Existing interstellar medium mapping techniques

There are two main categories of maps: 2D maps, which describe the distribution on the projection of the celestial sphere, and 3D maps, which try to assign

to each of them not only the position on the sky, but also the distance from Earth.

2D maps

2D maps could be based on absorption or emission, in particular from the thermal emission of dust particles. The advantage for this kind of maps is that for each line of sight the emission directly measures the total amount of dust, and therefore the total extinction, without a distance - limitation.

Schlegel et al. (1998) published an extinction map based on DIRBE and IRAS/ISSA, with an accurate remotion of noising light from zodiacal radiation and point sources. It is a far-infrared map, with data from 100 to 240 μm. To make the extinction map, they suppose a standard reddening law, and they use the elliptical galaxies as calibration for the reddening corresponding to the 100 μm emission. Their observational range is dominated by the emission from the biggest grains that are in thermal equilibrium with the local radiation field.

More recently, Planck Collaboration et al. (2014) provides emission 2D maps, with a similar extraction technique. For the interstellar medium science, this map discovered large structures in the emissivity spectral index of the dust β without relationship with known stellar structures.

3D maps

In this section, several 3D mapping methods are described with the categorisation of Sale (2015). More or less every map is based on a stellar catalog, with stellar positions, on the sky and distance, and probe for interstellar matter. As said in Chapter 1.2, there are essentially two distance - limited measurements: extinction, linked with dust grains, and absorptions, that measure the gas components. A 3D map could also be based both on extinctions (at different bands) and the absorbed lines. Generally, the ion, atomic or bi-atomic lines were used for the local maps, because they reach rapidly the saturation regime, where the Equivalent Width (EW) is not proportional to the matter quantity, therefore it is useless for mapping. These maps evolved a lot passing from an handful of line of sight to some thousands (Vladilo et al., 1985; van Loon et al., 2013; Vergely et al., 2010a). The DIB absorptions are more interesting for the larger scale mapping because they never reach the saturated regime, and their EW is always proportional to the absorbers column, and the DIBs were used both for local ISM (van Loon et al., 2015) and larger galactic environment (Zasowski et al., 2015).

Independently from the matter probe, there are three main techniques to chart interstellar medium in Milky Way: voxelised methods, model-based maps and gridless methods.

- **Voxelised methods** are for example used by Green et al. (2015): they divide the 2D sky in "pixels", for each pixel they assign bins at various distances. Globally, bins at the same direction on the sky form a voxel. In each voxel, they evaluate the extinction curve in different bins, to match individual observations or statistical variability along the voxel. These methods produce maps with high coherence in single bins on 2D sky, but each bins is independent and the results are maps with unnaturally

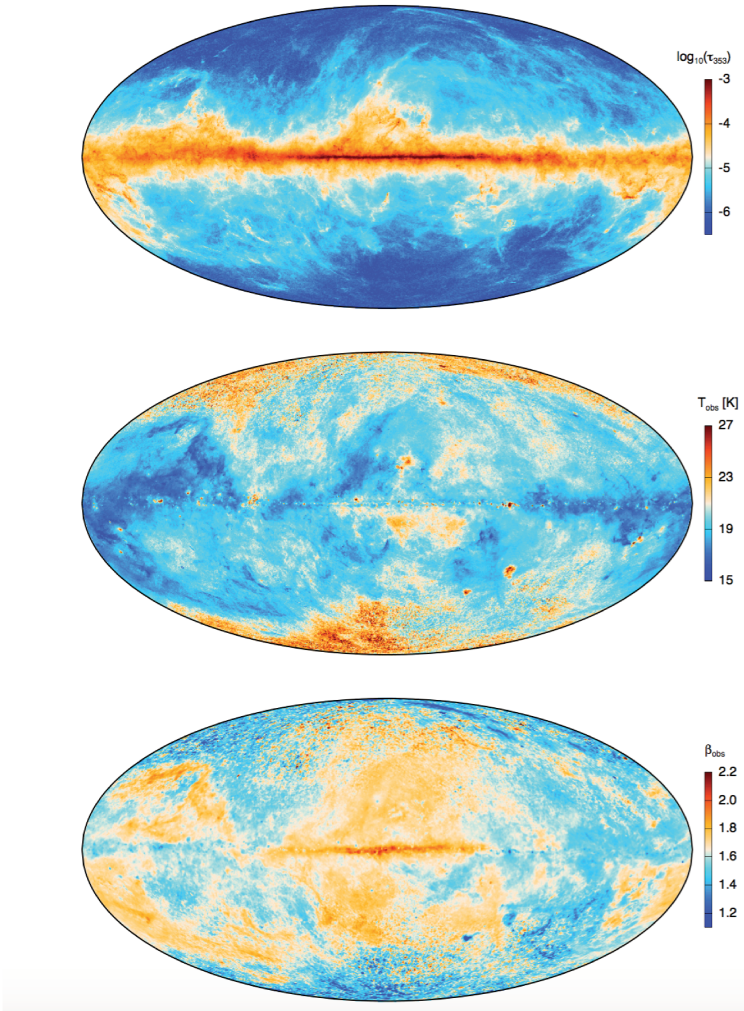


Figure 2.1 – All-sky 2D maps of the modified blackbody fit of *Planck*. Top) Optical depth at 353 GHz. Middle) Observed dust temperature, in kelvin. Bottom) Observed dust spectral index $\beta_{observed}$. Credits: Planck Collaboration et al. (2014).

elongated pencils. An example of this method applied using a different probe is the Zasowski et al. (2015), where the chosen feature to describe interstellar medium is a DIB absorption in APOGEE spectra.

- **Model-based methods** produces a synthetic stellar population in the modelled map of interstellar medium, to reproduce the observed population (Marshall et al., 2006). The similarity with the voxelised method is the use of an tessellation approach. They are actually pseudo-three dimensional maps, because they evaluate the extinction at different lines of sight, but there is not a correlation between to adjacent directions. The consequence of a non-correlation between adjacent directions is a fragmentary map, with strong gradients which are not realistic for the interstellar medium physics, as it is remarkable in Fig 2.2. In the top) panel there is a zoom in the Green map for the solar neighbour, where the interstellar medium seems to have non-natural segmented shapes; in the other hand, their map has a high precision on the sky - bottom panel.

- **Gridless methods** do not use a first guess of specific grid-distribution for the interstellar matter. They use information on stellar position and distances, and extinction for single lines of sight, but also the correlation in orthogonal directions. Light measured from two neighbour points in the space must pass through similar environments, and it means that the inferred map has to take in account not only the information for each line of sight, but also the reciprocal position between the sources. It is un under-estimated problem, because it guesses to determined extinction value for all points (infinite) in a given volume, using finite extinction estimations.

  A possible technique is to describe the correlation between the different points in the volume as an analytical function. For these techniques, the advantages are more physical structures; unfortunately these methods are very slow and they cannot be used for very large volumes, because of the calculation time. First maps with this technique were done by Vergely et al. (2001), but recently other groups start to use similar approach. In particular promising mapping technique are developed by Sale and Magorrian (2014a) and Rezaei Kh. et al. (2016).

In this thesis, I apply the gridless method from Vergely and I adapted it at new data and larger stellar catalogs, in preparation at the final Gaia catalog.

2.2 Known 3D Structures

Thank to previous maps, we yet know some important ISM structure in Milky Way. The interstellar medium is concentred on the galactic disk, it reproduces the same large stellar structures. In particular, depending on the charted volumes in the maps, large structures or local ISM are well investigated.

Large structures

From measured distances of emitting regions and the 21 cm radio emissions, Milky Way is schematised as a spiral double barred galaxy (Georgelin and

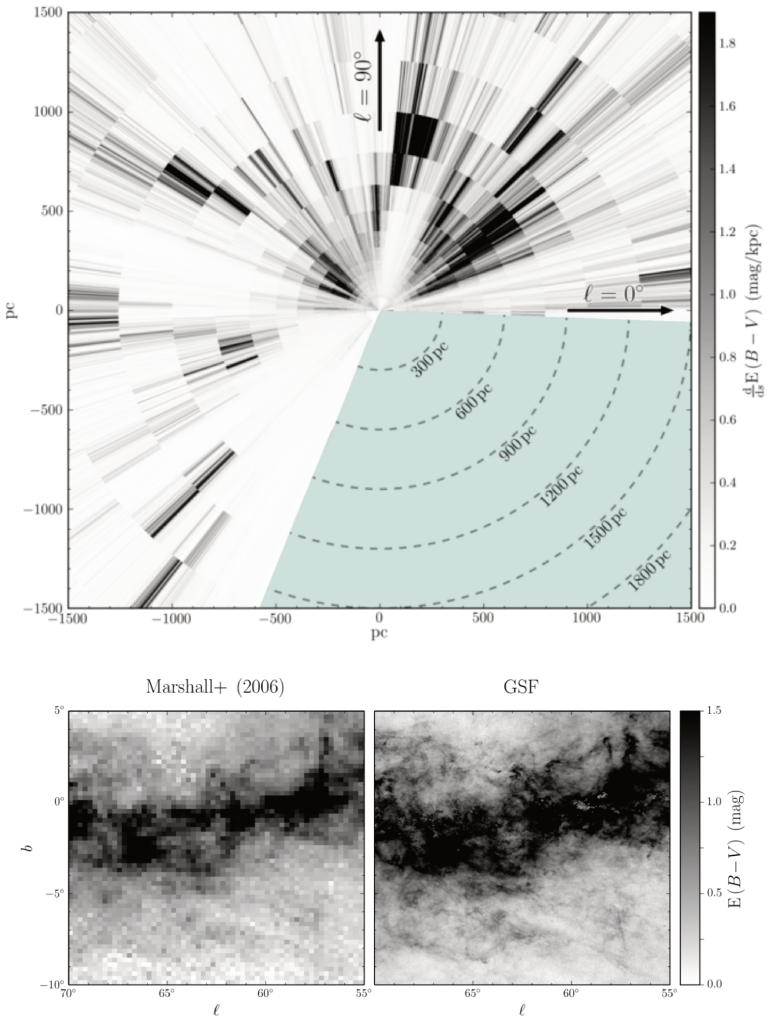


Figure 2.2 – Top) Example of galactic plane as in voxelised maps. Bottom) Comparison of precision on the sky between a model-depending map and a voxelised map. Credits: Green et al. (2015).

Component	Scale height
CO and H_2	$\sim$ 80 pc
Atomic gas	115 pc
Ionised gas	>200 pc
Dust	117 pc
862 nm DIB	209 pc

Table 2.1 – Different scale heights for interstellar medium component, Ferrière (2001); Kos et al. (2014).

Georgelin, 1976; Dame et al., 2001; Reid et al., 2014). Also maser sources are largely used to locate arms, since they are young objects related to high mass star formation (Hou and Han, 2014; Lépine et al., 2017). An artistic Milky Way representation is in Fig. 2.3 describe our knowledge about the large structures. Sun is located on the Orion Spur, a small arm between Sagittarius and Perseus, at 8 kpc from the Galactic Centre.

Despite all these accurate studies, the arms structure is not yet well determined, and in particular an accurate dust distribution map is not available; Gaia DR2 boosted extinction-based maps, and some examples are now in publication (Chen et al., 2018; Rezaei Kh. et al., 2018).

Considering the galactic plane as a unlimited plane with uniform gravitational potential, the distribution of matter in a cylinder perpendicular to the plane is:

$$\rho = \rho_0 \exp\left(\frac{-z}{H}\right) \tag{2.1}$$

where ρ_0 is the density on the plane, z the distance to the plane and H is the so-called "scale height".

Moving away from the disk, the scale height varies for different interstellar medium components, as reported in Tab 2.1. These variabilities are due to their origins environments, their preferable locations and their resistibility in different environments.

Local Bubble, Local ISM

The Sun sets in a hot and diffuse interstellar medium cavity, called Local Bubble, located in the Orion Spur and at ∼8.5 kpc from the galactic centre. In the Local Bubble plasma has a kinematic temperature of million of degrees, and it is delimitated by a ring of denser environment. The total size depends on the explaination, but in general it occurs for ∼100 pc on the galactic plane and hundreds of parsec in a vertical chimney through the disk. Using atomic absorptions, three dimensional maps have been done for fifteen years (Vergely et al., 2001; Welsh et al., 2010; Vergely et al., 2010b), then also through light extinction (Lallement et al., 2014) and different optical DIBs (van Loon et al., 2015). The Bubble is surrounded from neutral gas clouds, with some tunnels in direction l = 70°and l = 250°, in direction of the radio emission source GSH238+00+09 (Lallement et al., 2014). In Fig 2.4 there is the Local Bubble as charted in 2014 from individual excess of colours for some tens thousand of stellar targets. Another local structure is the Gould belt, a ring of young stars

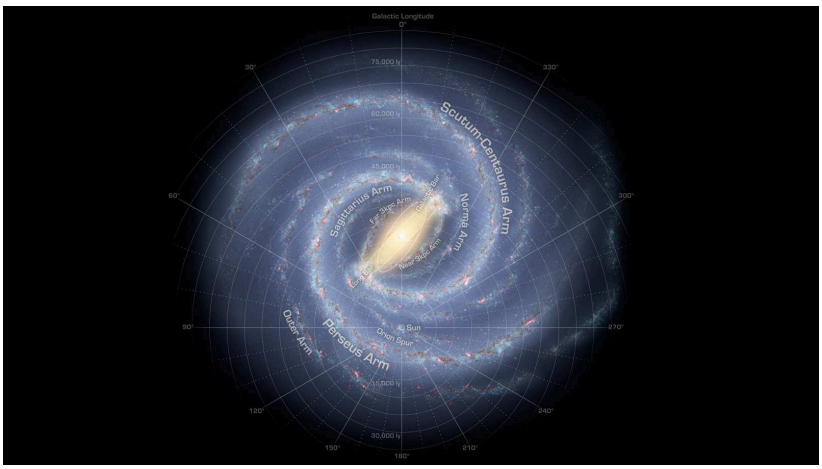


Figure 2.3 – Artistic Milky Way representation, based on observational consideration. Credits: NASA/JPL-Caltech/R. Hurt (SSC/Caltech).

tilted by ∼20°from the galactic plane. Since the Local Bubble and the Gould belt has a similar inclination, it could mean a common origin mechanism. The local environment origins are largely unknown, many possible mechanisms were been evaluated - from supernovae series to interaction between a super cloud and spiral arm - and a accurate 3D mapping can strengthen or exclude some hypothesis.

For the local astronomy this environment is objects of investigations, because its small dense structures and their kinematics impact on the interstellar dust and the few comets which arrives in the Solar System and penetrate heliosphere.

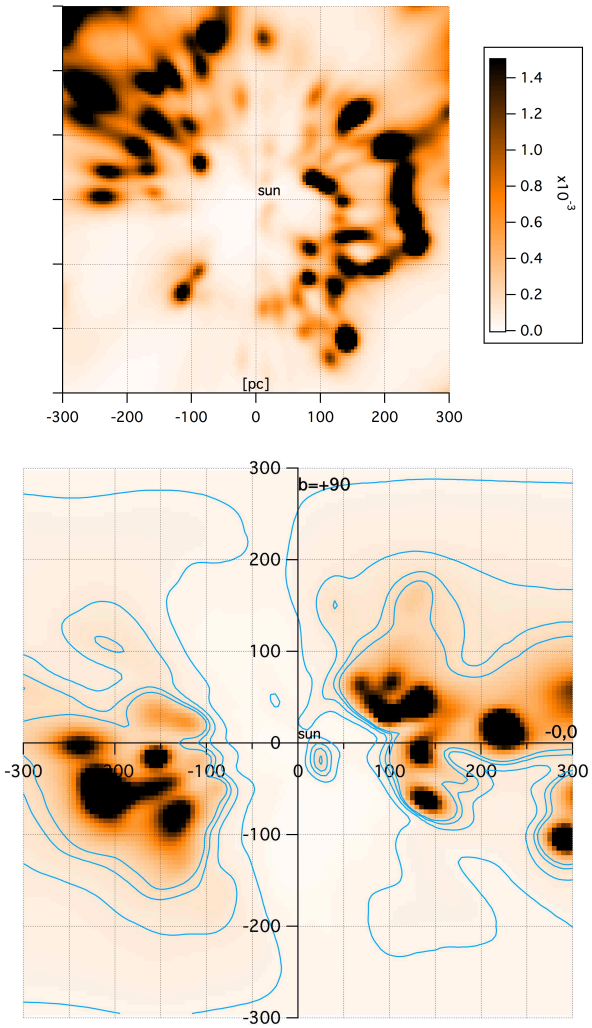


Figure 2.4 – Top) Excess of colour per pc map of the Local Bubble, slice on the Galactic plane of the map Lallement et al. (2014). The Sun is at (0,0) and the Galactic centre to the right. Longitudes increase counter-clockwise. Bottom) Local Bubble in a vertical plane long the direction of galactic longitude 0/180°. Colour scale is the same, colour units are mag/pc^{-1} and distance scale are parsec. Blue contours are for excess of colours per pc (1 e-05, 6 -05, 11 E-05, 16 E-05, 21 E-15).

Chapter 3

Used observational surveys

3.1 The Gaia Mission

Gaia is a ESA space mission, launched on 2013 December 19, arrived in L2 one month later and at the moment in operative state. Scanning the full sky for a minimum of 5 years, it is a global astrometric mission, whose main goal is to measure position, parallax and proper motion for at least 10^9 sources, to describe Milky Way three dimensional structures. It is measuring some tens million stars with distance accuracy better than 1%, and hundreds million around 10%.

To achieve this scientific objective, Gaia has three instruments: astrometric, photometric and spectrometric.

Data realises are scheduled[1] and they provide (short list):

- First release: 14 September 2016. Positions (α, δ) and G magnitudes only for sources with acceptable formal standard errors on positions; the five-parameter astrometric solution - positions, parallaxes, and proper motions - for stars in common between the Tycho-2 and Gaia is contained in Gaia DR1; positions and G magnitudes for some thousands of quasars;
- Second release: 25 April 2018. The five-parameters astrometric solution - positions on the sky (α, δ), parallaxes, and proper motions - for more than 1.3 billion sources, in the magnitude limit G = (3, 21); median radial velocities for $\sim$>7.2 million stars with a mean G magnitude between about 4 and 13;
- Third release: 2021. Improved astrometry and photometry, object classification and astrophysical parameters, together with Blu Photometer/Red Photometer spectra and/or RVS spectra they are based on, for spectroscopically and (spectro-)photometrically well-behaved objects; mean radial velocities for those stars not showing variability and with available atmospheric-parameter estimates; variable-star classifications together with the epoch photometry used for the stars; solar-system results with preliminary orbital solutions and individual epoch observations; non-single star catalogues;
- Final release for the nominal mission: end 2022

[1] https://www.cosmos.esa.int/web/gaia/release

Gaia instruments

There are three scientific instruments on the Gaia satellite; the most important instrument for the goal mission is the one dedicated at the astrometric measurements, and it is based on the global astrometry concept than in Hipparcos intruments. The photometric instrument is composed by two low resolution fused-silica prisms - called Blu and Red Photometer- , which disperse all the light entering the field of view. The Blu Photometer and Red Photometer operate respectively in the wavelength range 330-680 nm and 640-1050 nm (Carrasco et al., 2016). It permits to roughly determinate stellar proprieties as temperature, mass and composition for all stars in catalogue. Third instrument was coinceved to be dedicated at the stellar velocity measurements:

Radial Velocity Spectrometer (RVS) is an integral spectrograph that disperses all the light entering in its field of view ($\sim 2.5 \times 1.6\text{deg}^2$). Its resolving power is $R = \frac{\lambda}{\delta\lambda} = 11500$. As its name suggests, the main aim is the measurement of radial velocities, to describe the kinematics in Milky Way. Radial velocities are crucial to study the kinematical and dynamical evolution of the Milky Way and they are needed to correct the astrometry from the perspective accelerations (Katz et al., 2004). Moreover, thanks to the multi-epoch measurements, it will be possible to identify and characterise not-single systems. Because of scattered light and other non-conformity, the precision is actually 15 km s^{-1} at the V limiting magnitude of 15.2 for GV2 stars and of 11.3 for B1V stars.

The RVS range between [848-874] nm was chosen because of the presence of the Ca II doublet; all the stellar lines will be used to determinate stellar intrinsic parameters as temperature, gravity, metallicity and obviously the observational parameters as radial velocity and rotation velocity.

Broad scientific yield from the RVS measurements, here I list some examples: kinematics of the globular cluster could constrain the MW accretion history, general kinematics is a fondamental measure to determine the Dark Matter distribution, as it is inferred principally from kinematical considerations; also the disk kinematic needs to be studied to separate thin and thick disk behaviours, for example velocity structures with peculiar velocities could be old small merges debris; the metallicity determination is fundamental to built a realist evolution model. RVS shall discover and help the classification of the poor stars, the fossil stars which could represent the Population III and which are the most important constrain for the Big Bang nucleogenesis rate, in particular the helium abundances (Wilkinson et al., 2005).

Gaia for the interstellar medium astronomy

Gaia catalogues provide 1.5 and a half billion precise parallaxes, and extinction for a subset of targets; basically, extinctions will be estimated from photometry and spectrometry, using excess of colour as a proxy for extinction. For the subsample of the whole catalogue also measured by the RVS, stars will associate magnitude, parallax, temperature and metallicity; in those cases the distance-extinction pairs will have high precision, because it will possible to disentangle the reddening due to the temperature, the metallicity and the dust. In DR2 there is a first extinction catalogue, but it is strongly conditioned by the fact

that Gaia spectra are not yet available and reddening measurements are based only on Red/Blue photometric filters.

In addiction to extinction estimations, from RVS spectra, it will possible directly measure an interstellar medium absorption - the Gaia Diffuse Interstellar Band (Gaia DIB).

Gaia DIB Gaia DIB is a small DIB absorption, centred at 862.0 nm; the DIB is rather in the middle of the spectral range of Gaia-RVS. Only a small part of the RVS spectra will be usable for the extraction of the DIB, because of the high SNR required for the extraction. In addition, the closest objects, with apparently higher magnitudes, have less interstellar absorption, while the more distant objects for which the DIB is deeper and easier to measure, have lower apparent magnitudes and therefore a lower SNR. One of the objectives of my thesis is to develop a fit technique adapted to RVS spectra and predict the order of magnitude of the measurable DIB; these issues are investigated in Part II.

The Gaia DIB was studied and characterised by Munari (2000). This study is based on 37 observations at the Asiago Echelle spectrograph and describes its empirical law for correlation with excess of colour:

$$E_{B-V} = 2.69 EW \text{Å} \tag{3.1}$$

but Munari him-self proposed that this law should be:

$$E_{B-V} = \alpha(l, b, D) EW \tag{3.2}$$

because the exact slope depends on the interstellar medium proprieties, and it is a good ansatz its variation for different Galactic structures. Puspitarini et al. (2015) did an accurate fit for the 862 nm DIB for 64 stellar targets in CoRot field l = 212.9° and b = -2° of Gaia-ESO Survey (GES), founding an α proportion value is 2.3, compatible with the Munari value of 2.69. Since the spectra in Puspitarini are all concentrated in the same line of sight, the Munari estimation for the proportion is more safe. They empirically extracted the DIB shape, using the technique for DIBs shapes profiling described in Chen et al. (2014).

Variations of the DIB shape or the extinction/EW proportion are a strong constraint for the identification of DIB carriers, as well as understanding the different positions for the different carriers of DIB and grains of dust. A map of the Gaia DIB was still drawn from Kos et al. (2014) up to 3 kpc from the Sun; based on 500,000 spectra from the Radial Velocity Experiment (RAVE), they mapped a pseudo 3D map of the DIBs and independently of extinction. They estimated the distance-extinction pairs using Bayesian techniques, while the 862 nm DIBs were detected after combining different RAVE spectra to obtain an SNR ratio of ∼300. The maximum distance of their points from the galactic plane is 1.5 kpc and 40° at galactic latitude, for operational reasons. The map, in Fig 3.1, is composed of bins at different distances, with variable altitude on the plane, so all information on the z directions are compressed into 2 points, for the North and South hemispheres. Despite its distance and direction limits, this first map shows very similar behaviour when comparing the extinction and equivalent widths of the DIB, but the height of the DIB

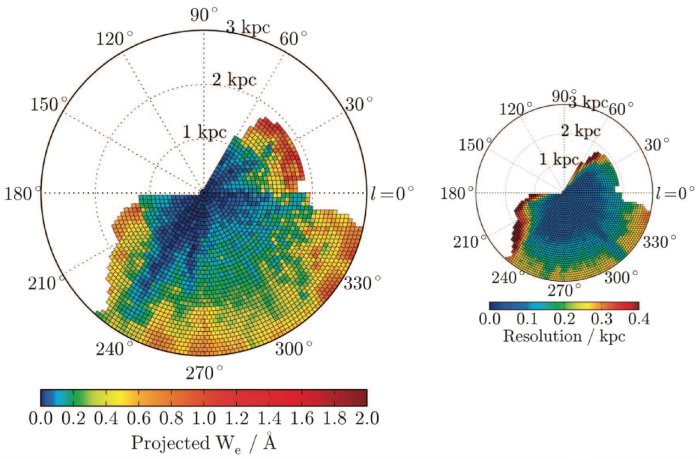


Figure 3.1 – On left, the projection on the galactic plane of the equivalent width value, coded by colours. The small map on the right is the spatial resolution. Credits: Kos et al. (2014)

scale results very high compared to that of the dust: 209.0 ± 11.9 pc for the 862 nm DIB and 117.7 ± 4.7 pc for the extinction.

Gaia spectra are not yet published, but three examples are shown explanatory in Fig 3.2. These normalised spectra show clearly the DIB, for three stars of different stellar types; spectra are sorted with growing extinction from top to bottom.

3.2 Photometric and spectrometric surveys

Astronomy is enjoying its first large survey era, waiting for the ELT, which probably will provide another astronomical revolution. In my thesis I used Gaia results, but also the catalogues published by large on-ground survey. In particular, two spectrometric surveys get us massive dataset to study the effect of big data on the interstellar medium mapping.

Pan-STARRS

The Panoramic Survey Telescope and Rapid Response System (Pan-STARRS), located at Haleakala Observatory, Hawaii, USA, is a survey based on two 1.8 m Ritchey Chretien telescopes observations, Pan-STARRS1 and Pan-STARRS2 (Chambers et al., 2016). The survey covers the northern sky down to $\delta \approx 30$ and the first realise provides a database with more than 10 billion objects (Flewelling et al., 2016).

From the Pan-STARRS survey I widely used in my thesis the derived extinction curves for line of sight (Green et al., 2015). They are based on ∼800 million stars, covering three quarter of the sky, and they have an hight angular resolution from 3.40' to 13.70'. Each line of sight is sampled in distance modulus, but the extinction curve does not start from the Sun position (distance zero) because they have severe constraints in the reliability and at the closer

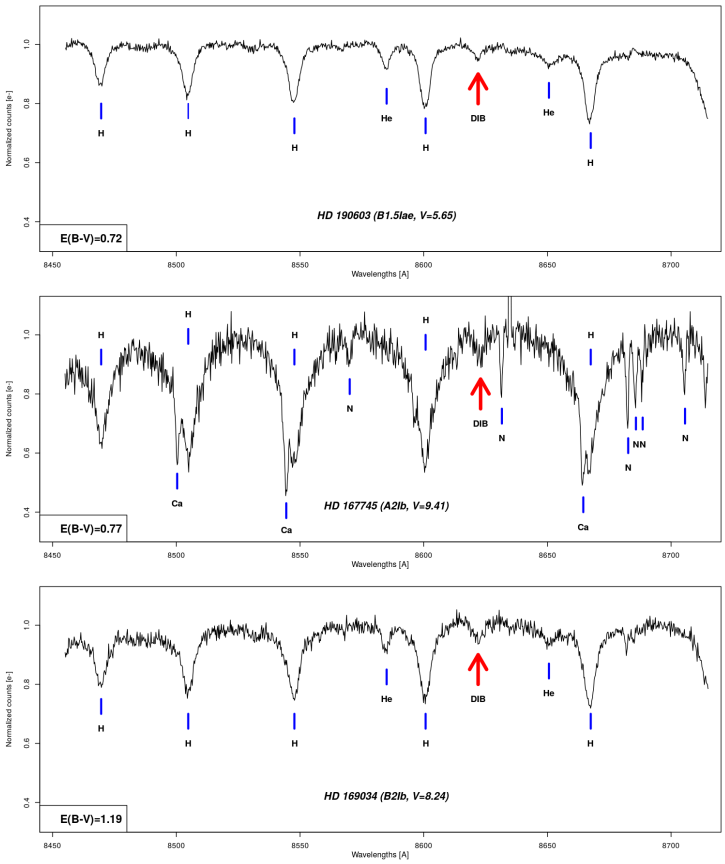


Figure 3.2 – Only yet published Gaia DIB from RVS spectra. Credits: ESA/ Gaia/ DPAC/ CU6/ Observatoire de Paris-Meudon/Olivier Marchal, Carine Babusiaux and David Katz

distances they have not enough stellar targets for their criteria. In effect, the resulted extinction map is very interesting for the large structures but not for the very local mapping. It is completely available [2], in Fig 3.4 an example of extinction curve is shown: different curves are from different stellar samples and the best result in bold line. Distances indicated in the graph with dashed background have not associated a reliable extinction curve. For the closer distances, as said, it is due at the low number of stars, for the more distant bins the decreasing of stellar targets is due at the observational selection in magnitude.

SDSS/APOGEE

The Sloan Digital Sky Survey is a large survey using a 2.5-m wide-angle optical telescope at Apache Point Observatory in New Mexico, United States. The Apache Point Observatory Galactic Evolution Experiment (APOGEE) is a high-resolution and high signal-to-noise infrared spectroscopy survey in the

[2] http://argonaut.skymaps.info

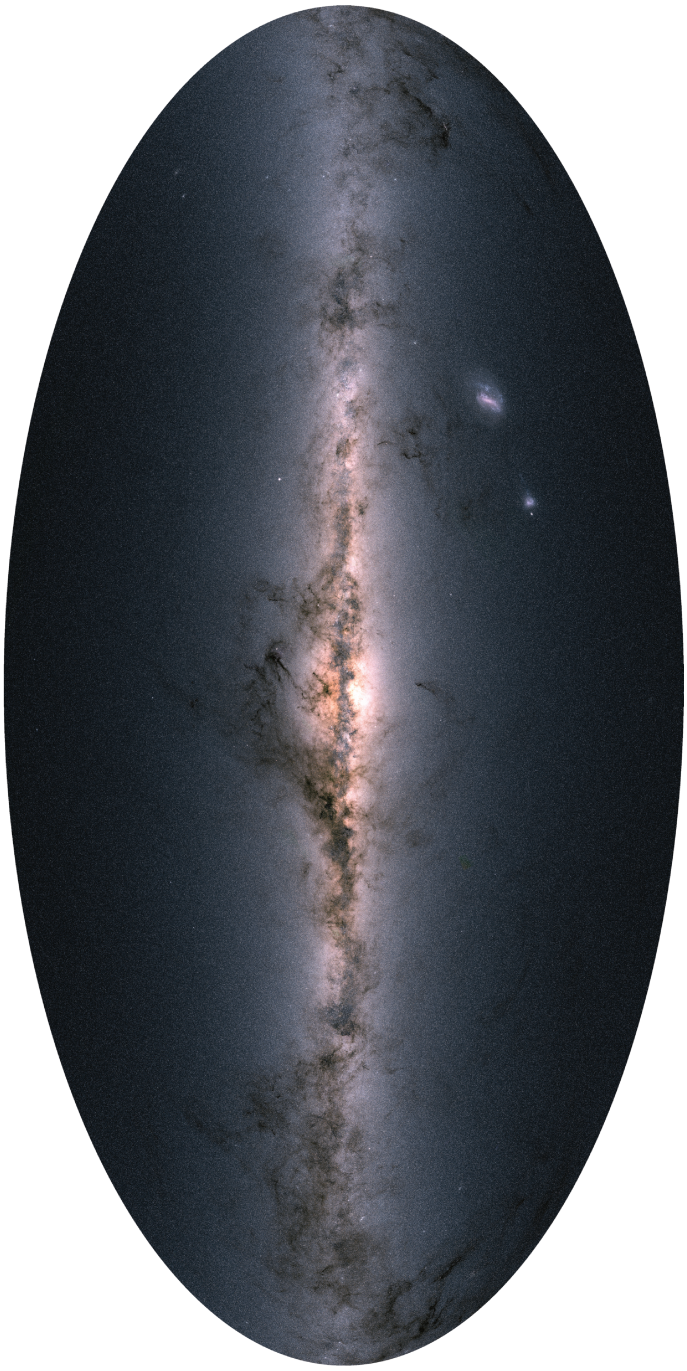


Figure 3.3 – Map of the sky from Gaia DR2. Credits: ESA/Gaia/DPAC, A. Moitinho / A. F. Silva / M. Barros / C. Barata, University of Lisbon, Portugal; H. Savietto, Fork Research, Portugal.

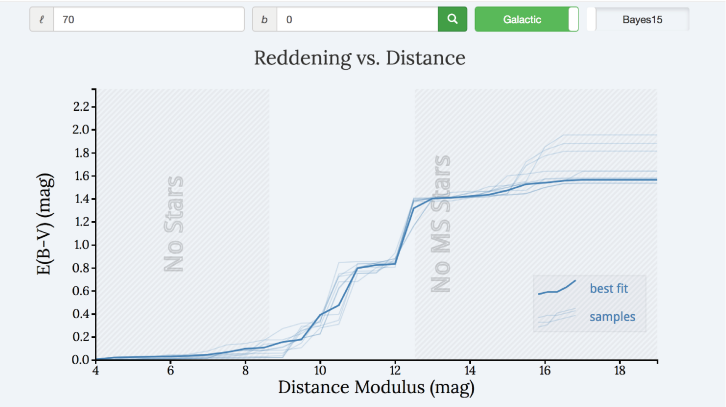


Figure 3.4 – Example of extinction curve from the Green et al. (2015) map, for galactic coordinates l,b=(70, 0).

SDSS context; it has provided a large catalog of stellar spectra and several stellar estimations: stellar parameters, radial velocities, distances and extinctions. Its main scientific goal is constraining dynamical and chemical evolution models. Its targets scheme is quite complex and based on precise fields which do not cover all the sky (Zasowski et al., 2013). APOGEE spectra database was fundamental in my thesis, because I tested the use of DIBs for a full-3D map, see Chapter 16, using the so-called APOGEE DIB. In the APOGEE spectra band there is a DIB absorption at 1527.3 nm, which is the most intense in the H band. Zasowski et al. (2015) made a large catalogue of the 1527.3 DIB and used it to map the DIB distribution; another accurate DIB catalogue was done and used also to determine some physical constraints on the carrier (Elyajouri et al., 2016, 2017b), but they analysed especially the calibration targets.

The APOGEE survey contributed in more accurate calibrations of distance-extinction couple, achieved using different photometric and spectrometer considerations. During my thesis I used very accurate distance-extinction couples, derived using calibration based on Gaia and APOGEE information (Ruiz-Dern et al., 2018; Danielski et al., 2018). For the mapping purposes, a great problem of the APOGEE survey is that it observes in specific fields, which do not cover homogeneously the sky, as it is shown in Fig. 3.5.

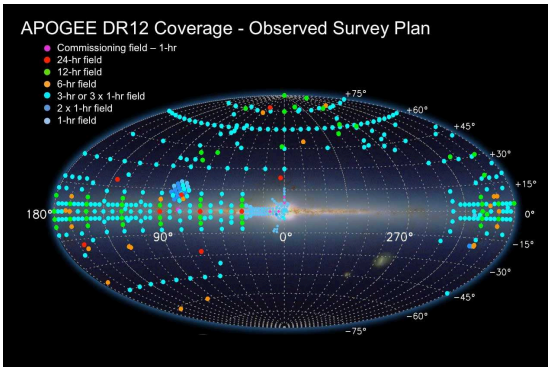


Figure 3.5 – APOGEE DR12 distribution on the sky of the stellar targets.

Chapter 4

This thesis project

My thesis is developed in a period of large surveys and strives to prepare tools to map in the most efficient way the huge amount of data Gaia will provide. In particular, I am interested to use the excess of colour and DIB absorptions from the Gaia catalogs to improve the 3D maps. In practice my project follows two main branches:

Gaia DIB: tools for efficient Gaia-DIB profile-fitting and estimate of expected results The first issue is to estimate how many DIBs will be measurable with the Gaia spectra. To make this prevision, I simulate the DIB extraction in synthetic spectra in the RVS band, with wide variation in stellar types, magnitudes and extinctions; then I code a specific fit profile for the Gaia DIB, adapted to the simulated Gaia RVS spectra. For each synthetic spectrum, I estimate how many objects of that type will be observed by the RVS, using a Milky Way simulation. In this way, it is possible to estimate how many objects will have a good DIB fit. This work is presented in Part II.

Three dimensional maps: improvement of techniques adapted to future massive datasets The gridless techniques are not adapted to chart large volumes or to manage massive target catalogs. In preparation to the large surveys, it is fundamental to understand how to use in the most efficient way all the available data. I test for the first time a tomographic inversion using different types of input data, that is individual DIBs absorptions, individual photometric extinctions and statistical interstellar medium 3D maps based on extinction in groups of stars. The maps were used to help the Gaia photometry to calibrate the Gaia photometry and, in return, new additional extinction measurements were obtained from the new calibrations, that were subsequently inverted to produce better 3D maps. I propose also a way to estimate uncertainties of quantities in the maps (see Part III). I developed a new method to invert data into a 3D map with reduction of the computational time, called "hierarchical technique", which reduces the calculation time without actually changing the map precision, because the map resolution is now directly the best resolution achievable at each point in space, and depends on the volume density of the stellar targets (Part IV). This technique was applied to 30 million of stellar targets from the DR2 Gaia, with positive results - Part V.

New developments using 3D maps Moreover, other astronomical informations are available from the 3D maps: in Part VI I investigated the comparison between DIBs carriers and dust distribution on the one hand, and developed a first 3D-kinematic map on the other hand.
As the organic matter in interstellar medium seems to be largely in DIBs carriers Snow (2014), it is important to identify these molecules, their origins and distributions; the APOGEE DR14 provides a large database of distance - extinction couples and spectra from which the $\lambda \sim 1.527\mu$m DIBs were extracted by my colleague Meriem Elyajouri. I produced two maps from the same stellar target using only photometrical extinctions and then only extinctions estimated from the DIBs absorptions. As the future Gaia DIBs will be measured with high precision, similar maps will be done and comparisons between the distributions of different DIBs will constrain the DIB preferred environments and as consequence the life cycle of the carriers.

To understand the evolution and the reciprocal relationships between the clouds, it is important to assign specific velocities to the interstellar medium clouds. I did this exercise in the Taurus region, based on various observational constraints: potassium absorptions in optical bands, CO emissions in radio bands and extinction 3D maps.

To resume, the scientific project was both how to extract the DIBs and how to use them, to be prepared to the challenge of Gaia data exploitation and to be ready and armed with the appropriate technical tools.

Part II

Gaia DIB: previsions and tools

Chapter 5

Simulated analyses of Gaia RVS spectra

In 2021, 150 000 millions of RVS spectra should be available to the community. This number applies to combined spectra from all visits to the same target (about 35 visits per star in average) and it is mainly due to the fact that only targets with G(RVS) <16 are retained on board for RVS spectroscopy[1] (G(RVS) is the target magnitude in the RVS spectral domain).

As already mentioned in the introduction, the Gaia RVS wavelength domain contains a weak diffuse interstellar band (DIB) centered at ∼862 nm, commonly called the *Gaia DIB*. Its response to the extinction was first studied by Munari (2000) thanks to the spectra from the RAVE survey, and later by Chen et al. (2014) and Puspitarini et al. (2015) based on ESO/VLT spectra with the FLAMES-GIRAFFE spectrograph. The DIB spectral profile (shape) has been first derived by Chen et al. (2014) and later confirmed by Puspitarini et al. (2015) using GES data (see Chapter 3.1 for an introduction). The normalised shape used in the following analysis is represented in Fig 5.3.

What we want to estimate is how many spectra will be useful for the extraction of the DIB and what kind of precision will be achieved for these spectra. Preliminary and very crude estimates had been done prior to the commissioning of the instrument, but much more precise predictions can be done after this phase and after the first results. Unfortunately, the unexpected straylight, discovered after the launch, affects primarily the RVS, and expectations in terms of signal and DIB extraction are lower than initially hoped.

The two main difficulties of the DIB measurements are associated with the relative weakness of the DIB and the subsequent requirement for a high signal to noise, and the need to disentangle stellar features and DIB absorption. Examples of stellar spectra recorded by the RVS are shown in Fig. 5.1 for stellar types varying from M to early B stars. As can be see, stellar lines strongly vary with the star temperature and one needs to possess accurate models for all types in order to obtain a good adjustment. The target stars of Fig. 5.1 are all very bright and close. Note that for such targets, the extinction is very low and no detectable absorption is expected, as it is indeed the case.

An example of 862 nm DIB observed from the ground for a reddened star is displayed in Fig. 5.2. The spectrum is the one of the cool targets from

[1] https://www.cosmos.esa.int/web/gaia/news-2018#DataReleaseUpdate2018]

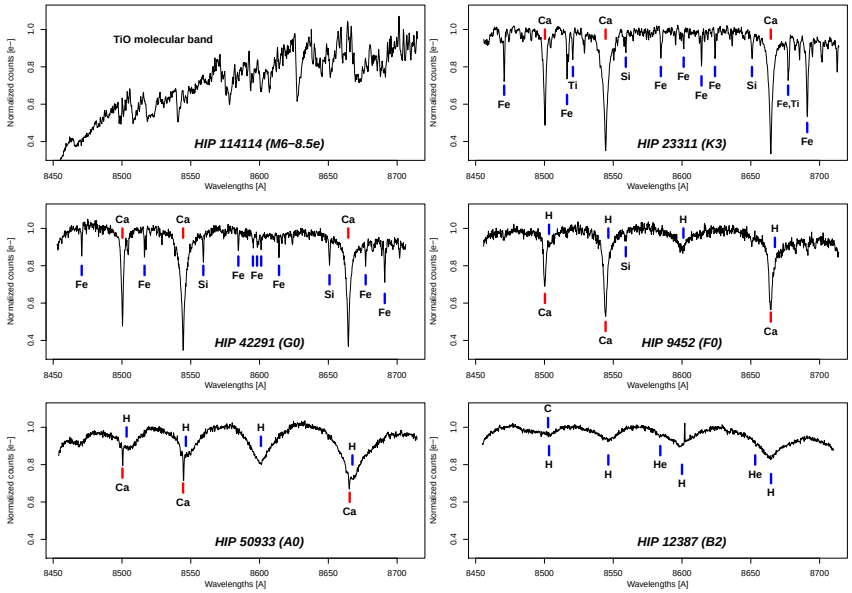


Figure 5.1 – Examples of RVS spectra for stars with increasing temperature. The stellar lines must be very precisely modelled and adapted to all types if we wants to extract the diffuse interstellar band from the data.

the Gaia-ESO Survey at the VLT and was recorded with FLAMES/GIRAFFE with a resolution R=17,000, i.e. on the same order than the RVS resolution of 11,500. The DIB depth amounts to several percents and corresponds to the target color excess E(B-V) on the order of 0.7 ($A_V \sim 2$, note in all along our study we assume that $A_V \propto$ E(B-V).). This shows how accurate must be stellar models if we want to extract such DIBs. Fortunately, no telluric emissions or absorptions will contaminate the spectra in the case of the RVS, at variance with the ground spectrum (see caption).

In the frame of the preparation of the Gaia data analysis and catalogues, the Gaia Data Processing and Analysis Consortium (DPAC) has prepared a set of simulators, and I used two of them to perform my estimations: the Gaia Object Generator (GOG) and the Gaia Universe Model (GUM).

First, I simulated a very large set of RVS spectra, using the Gaia Object Generator (GOG) (Luri et al., 2014) . GOG can simulate RVS spectra as they are expected at the end of the mission (final mission spectra), and takes into account a large number of observational properties for the simulated object, in addition to the star properties themselves. It provides also an assigned uncertienties for each simulated pixel, based on the target's light and all instrumental effects. In the simulation it is possible to vary the position on the sky, the distance, and the radial velocity.

Second, I developed a pipeline for the DIB profile fitting, adapted to varying stellar types, extinction and magnitude.

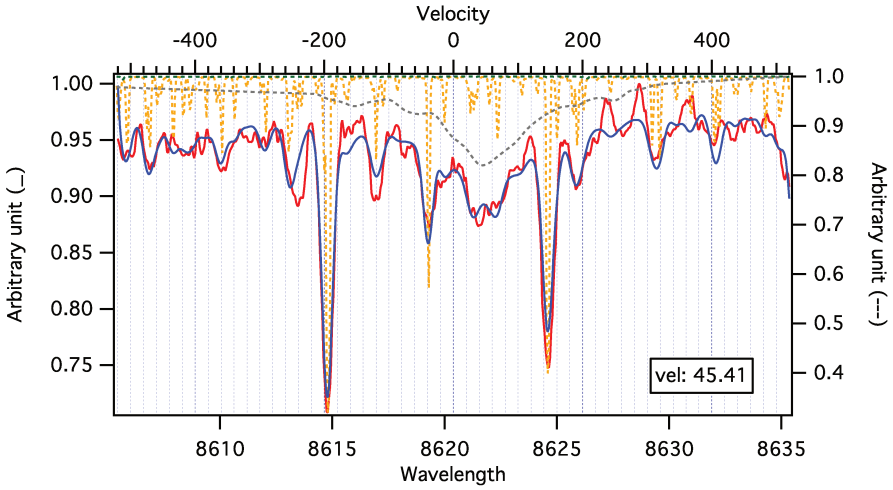


Figure 5.2 – Example of ground-based spectrum of a reddened cool star (red solid line) and a corresponding modelling (solid dark blue line). The model is a combination of a stellar theoretical model (dashed orange line) and a DIB model (dashed blue line), adjusted to the data. The spectrum is from the GES survey with the ESO/VLT/GIRAFFE spectrograph. There are sky background emission residuals in the red wing of the DIB (from Puspitarini et al. (2015)).

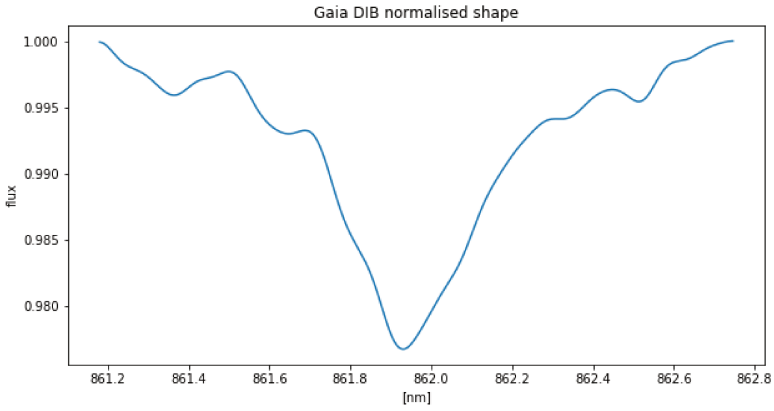


Figure 5.3 – Gaia DIB shape, empirically estimated for $A_V = 1.0$.

5.1 Preliminary estimate of upper limits on DIB measurements

If we assume that the stellar spectrum is completely flat, i.e. there are no stellar features at all (equivalently, if we assume that each stellar spectrum is perfectly known), and in addition that the DIB shape is constant, it is possible to estimate what is the achievable maximum G (RVS) magnitude for which the DIB can be extracted; we can estimate that for any DIB depth or equivalently any extinction A_V if we assume the DIB-A_V proportionally from the above formula. The key point for this estimate is the relationship between the stellar magnitude in the RVS wavelength range and the S/N. This SNR-G(RVS) relationship has been well calibrated and is available in numerical form from the Gaia website; it is shown in Fig. 5.4. It corresponds to end-of-nominal mission performances, and the average total number of visits to a given star.

We did a simulation using simply a Gaussian noise (for a series of S/N values) and a superimposed DIB, whose Equivalent Width (EW) corresponds to three extinction levels A_V=1, 3, and 6 mag. For the DIB we used the normalised DIB spectral profile shown in Fig 5.3. We assume an exact proportionality between the DIB EW and the reddening, according to the Munari (2000) law $E_{(B-V)} = 2.69 \times$ Equivalent Width(Å). The fitting function is made of a flat continuum (equal to unity) and a DIB profile, whose position and depth are free parameters. The error on the DIB EW is estimated in two ways:

1) with a bootstrap-type method: 100 different noise patterns are generated and for each one a fit is performed. The average solution and its standard deviation are taken as EW and σ(EW)

2) using an approximate simple formula based on the S/N (SNR), the number of pixels defining the DIB (N) and the wavelength interval for each pixel $\delta(\lambda)$ (assumed to be constant over the spectrum) $\frac{\delta(EW)}{EW} \simeq \delta(\lambda) \cdot \sqrt{(N)} \cdot 1/SNRr$.

Fig. 5.5 shows the resulting relative error on the equivalent width $\sigma(EW)/EW$ as a function of G(RVS) for the three A_V values. The two methods give slightly different results, as expected, however they can be considered as consistent for only an estimation. The figure can be used to guess which upper limit on the G(RVS) magnitude is compatible with a DIB detection for different quality criteria: e.g. a relative error of 50% is achievable for G(RVS) $<$ 12, 13.2, 14 for A_V=1, 3 and 6 mag respectively. Or, a relative error of 20% is achievable for G(RVS) $<$ 11, 12.2, 13.5 for the same A_V values of 1, 3 and 6 mag.

The actual DIB extractions will be evidently more difficult due to the stellar features, and will depend on the quality of the stellar models. As already found for DIBs observed from ground, the most difficult stellar types are not the same for all DIBs since the detection difficulty depends on the DIB width: schematically, stellar lines either much narrow or on the contrary much wider that the DIBs allow a better detection than stellar features with widths on the same order as the DIB width. This is the reason why nothing can replace a simulation for each type, and this is done in the next section.

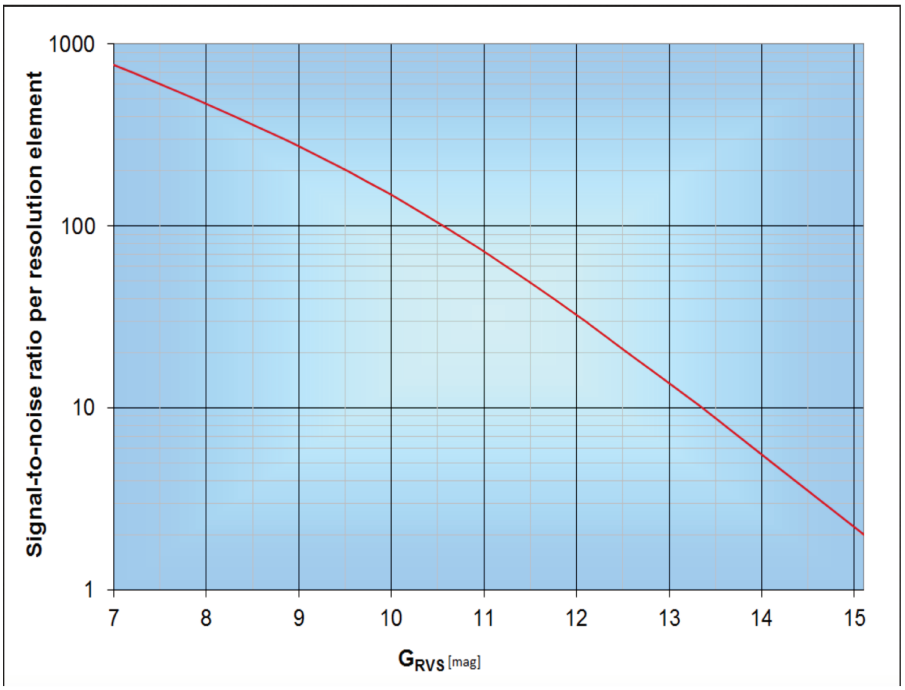


Figure 5.4 – End of mission signal to noise ratio of RVS spectra as function of the G (RVS).

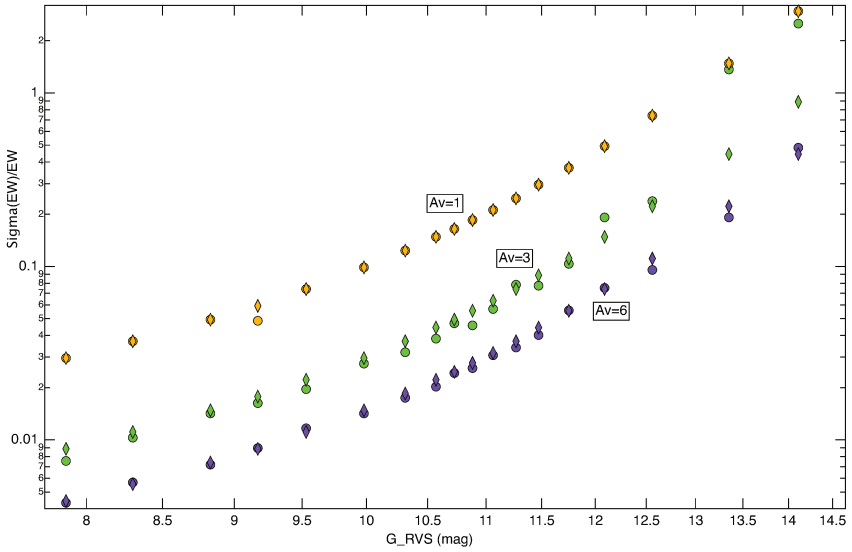


Figure 5.5 – Simulation of DIB extraction in the idealised case of a flat continuum. Two different estimations of the relative error on EW are shown for each value of A_V and each value of the S/N ratio (see text).

5.2 Simulation of RVS spectra with the GOG tool

The RVS spectra are not publicly available yet and are expected for the third data release (DR3). The Gaia DPAC members work on the data, but they are not allowed to communicate any data before the release. For this reason, I used only simulated data.

The GOG simulates realistic RVS spectra, in particular it simulates their associated noise, that is modelled using all the knowledge about its various sources: errors linked with the instrument characteristics and errors due to the stellar input flux.

The GOG simulator was not initially conceived with Gaia DIB absorptions, however one of the free parameters in a GOG simulation is the extinction which affects the spectra. This allowed C. Babusiaux from GEPI (with help of P. Sartoretti an N. Leclerc) to briefly test the inclusion of an additional absorption proportional to the extinction in one of the GOG versions. Thanks to the collaboration with Carine Babusiaux, I could make use of such a simulation.

The simulation uses the normalised DIB spectral profile shown in Fig 5.3. This initial profile corresponds to a reddening of reference. In all the study we assume that the DIB shape is unchanged and that there is an exact proportionality between the DIB equivalent width (EW) and the reddening, according to the Munari (2000) law:

$$E_{(B-V)} = 2.69 \times \text{Equivalent Width(Å)}$$

Because the DIB is shallow, we are always far from the saturation for the absorption and in this case it is well known that multiplying the EW by a coefficient α is equivalent to elevate the DIB normalised profile D(λ) at the power α. Therefore, the amount of reddening is very simply proportional to the coefficient α.

$$- \quad \mathbf{E}_{(B-V)} \rightarrow \alpha \ \mathbf{x} \ \mathbf{E}_{(B-V)} \tag{5.1}$$

$$- \quad \mathbf{EW} \rightarrow \alpha \ \mathbf{x} \ \mathbf{EW} \tag{5.2}$$

$$- \quad \mathbf{D}(\lambda) \rightarrow \mathbf{D}(\lambda)^{\alpha} \tag{5.3}$$

5.3 Grid of parameters for the simulated spectra

Each simulated spectrum is specific of a set of parameters Π.
-Four stellar atmosphere parameters - effective temperature, gravity, metallicity and rotational velocity
-The extinction (or equivalently the reddening).
-The star absolute magnitude

$$\Pi = [\text{T, log(g), [Fe/H], vsini; visual extinction, absolute magnitude V}]$$

It may seem surprising to vary the absolute magnitude of the star and not the distance from the star (and equivalently the visual magnitude and the flux). However this is motivated by practical reasons. Keeping the distance fixed

and varying the absolute magnitude (even if the magnitudes are unrealistic) produces variable fluxes in an equivalent way. There are some small changes in the RVS response to the same stellar light depending on the location of the star in the sky, however their effects are much smaller than the uncertainties in our final estimates and I did not vary the position on the sky. For practical reasons we used a fixed distance d= 91.69 pc and a fixed direction - α= 249.68° and δ= -16.2° -, however these values are not important in our study.

The simulation table is a squared grid. It contains non-existing stellar types, however those will be eliminated later. The grid is the following:

- **Temperature [K]**: 3500, 4500, 5500, 6500, 7500, 8500, 10000, 15000, 20000, 25000, 30000;
- **Log(g)**: 1.01, 1.61, 2.21, 2.81, 3.40, 4.01, 4.61;
- **[Fe/H]**: -3.0, -2.5, -2.0, -1.5, -1.0, -0.5, 0.0, 0.5, 1.0;
- **vsini [km/s]**: 0.0, 5.0, 10.0, 50.0, 99.0;
- **mean absolute V [mag]**: -1.0, 0.0, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0;
- $\mathbf{A}_V$ **[mag]**: 0.0 , 0.6, 1.2, 1.8, 2.4, 3.0 , 3.6, 4.2, 4.8, 5.4;

As we will see in the Chapter 7, the above range of parameters corresponds to $G_{RVS} > 9$ mag, but magnitudes below 9 correspond to high SNR and for such spectra the DIB will always be accurately measured. The number of spectra for stars with $G_{RVS} > 9$ are estimated as ∼700 000, to be add to our following results for fainter targets.

It is also possible with GOG to simulate the RVS spectra as they would be in the absence of photon noise, and we will also use this capability (see next section). Fig. 5.6 shows an example of GOG simulated spectra for the same stellar type, including the GOG simulated noise and in absence of any noise.

The inclusion of the DIB is illustrated in Fig 5.7. The figure displays a series of GOG simulated spectra that corresponds to the same stellar parameters and mean absolute V magnitude, but a variable extinction, and therefore a variable corresponding DIB. The extinction varies from A_V=0 to A_V=5.4 and allows to estimate the S/N required to extract DIBs of various strengths. It shows that the detection will crucially depend on both the signal and the amount of extinction. An advantage is that faint stars are generally more distant and more reddened for a given direction close to the galactic plane, i.e. a weak signal corresponds to a stronger DIB (in average and, again, for the same direction).

GOG is not remodelling a new theoretical spectrum for each input, but it is based on a very extended library of pre-computed stellar models for grids of parameters. For computational reasons, it does not interpolate between them, instead it searches for the stellar type (equivalently the set of parameters) that corresponds the most closely to the required parameter values. As a consequence, GOG does not exactly provide us with a an output spectrum that corresponds exactly to the set of parameters entered in input, but to the closest set. However, in some cases, the combination entered in input is too far from all the parameter sets of the database and there is no solution, a situation that corresponds to very rare stellar types. For this reason, the output from GOG consisted of 381 900 spectra from the 450 450 required.

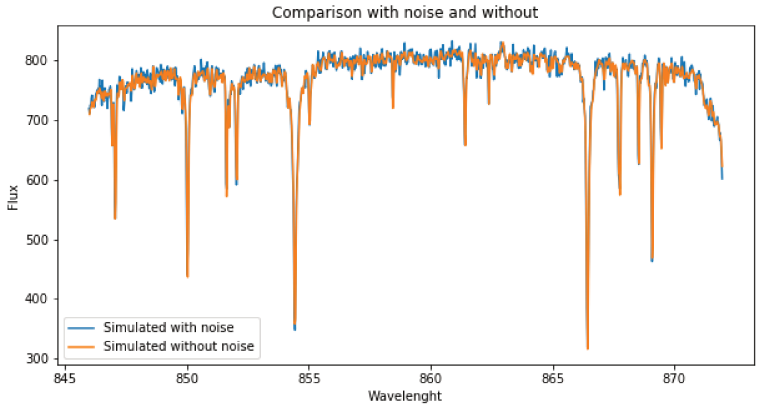


Figure 5.6 – Comparison between two output of GOG simulation, with noise (blu line) or without (orange line), for stellar parameter T = 3500 K, log(g) = 4.01 and [Fe/H] = -0.5.

5.4 Fitting of RVS spectra simulated with GOG

Our adjustment made to derive the DIB equivalent width (EW) is quite simple: it is based on the principle that the stellar continuum can be perfectly well fitted and that the DIB shape is similar to our empirical shape. Because we use here a stellar model perfectly adapted to the simulation (it comes from the same library), and because the simulated DIB and the one used in our adjustment are identical, this assumption is evidently totally valid here.

All simulated RVS data were fitted to a model that is a combination of a normalised stellar theoretical spectrum, a DIB absorption and a linear continuum.

$$\text{Flux} = (\text{D}(\lambda))^{\alpha} \times (Y_{StellarModel}(vsini, \epsilon, V_{rad}))^{\beta} \times (m \cdot \lambda + q) \tag{5.4}$$

-The DIB model D(λ) and its free parameter α were presented in the previous section. α is proportional to the extinction as we discussed above.

-We chose as stellar models simulated spectra as they come out from GOG when removing the fluctuations associated with the instrumental effects and with the photon noise, and without any DIB. In this case they correspond to state-of-the-art model theoretical spectra (Sordo et al., 2011). Note that we initially performed all the fits using completely different stellar models: these previously used models were interpolated in a grid of stellar spectra built for ground-based DIB analyses. However, we found that multiple stellar features were absent from these models, while they were present in GOG spectra. Therefore we switched to the GOG models. Our justification is that we can expect that at the time of future RVS data analyses excellent models will be available and will be as accurate or better than the GOG models (e.g. data-driven models as done for the APOGEE survey).

-The coefficient β introduces the possibility to vary the depth of the stellar features. This allows to partially take into account the metallicity and/or the

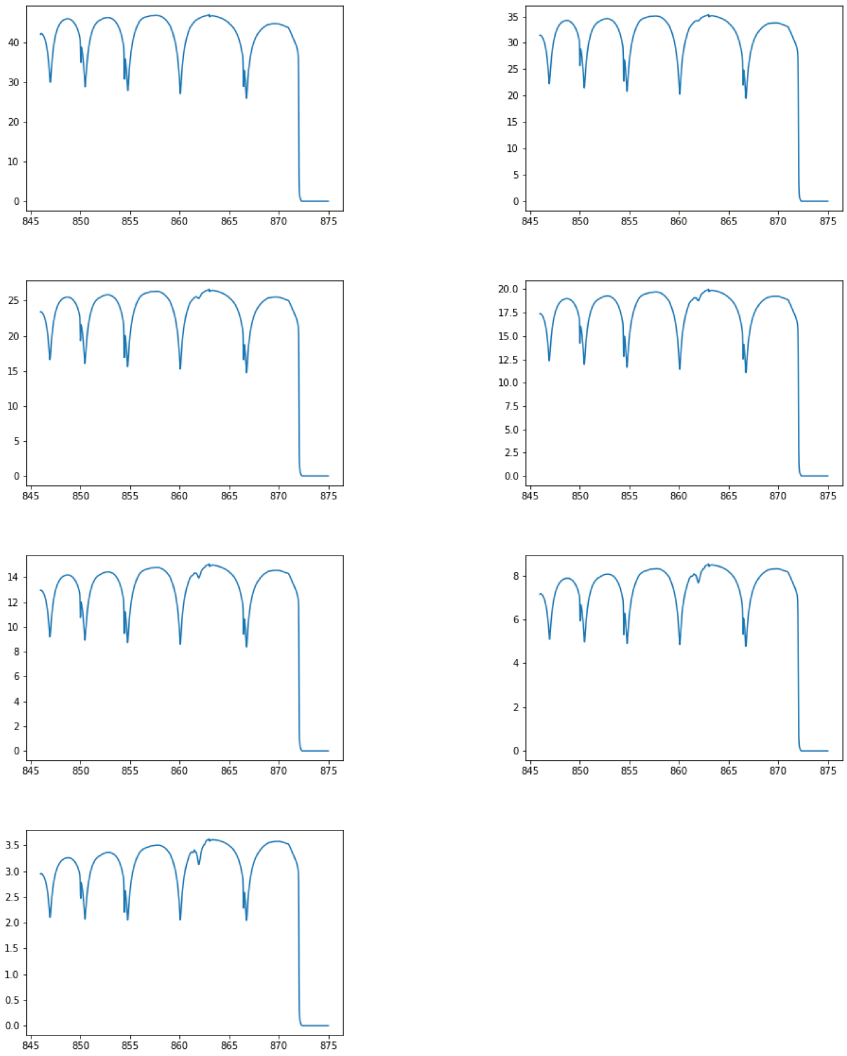


Figure 5.7 – Example of GOG models, for the same stellar parameters T = 7500 K, log(g) = 1.01, [Fe/H] = -3 and for increased extinction (and 862.0 nm DIB), for A_V= 0, 0.6, 1.2, 1.8, 2.4, 3.6, 5.4 from top to bottom and left to right.

temperature.

-Vrad is the star radial velocity

-The parameter vsini (stellar rotation velocity) controls the stellar line broadening and the parameter ϵ the atmospheric limb darkening.

-The use of a linear continuum has proven to be adequate. A flat continuum is not fully adapted since the extinction reduces the flux differently as a function of the wavelength (the reddening). Indeed, allowing for a slope gives slightly better results than the absence of slope.

The fit parameters (ensemble ϕ) are finally: α , β, vsini, ϵ, the radial velocity, the continuum slope m and its intercept q. Since there are no doppler shifts associated with the star radial velocity in the simulated data, and no variability of the location of the DIB absorption there should be not be any variability of this parameter. However Doppler shifts will evidently exist in real Gaia spectra. For this reason I kept the radial velocity as a variable parameter, but as expected the adjusted value was always close to zero. In future our simulations will be extended to the case of variable DIB Doppler shifts and variable stellar radial velocity.

I coded the adjustment program in Python language. The convergence to the best-fit solution is based on a classical non-linear least squares criterion, using the open - packages scipy.optimize.curve_fit and PyAstronomy.

Choice of initial parameters We must fit enter initial guesses for the parameters, a set ϕ_0.

-Our input α_0 value is the input A_V, used to simulate the spectra. We justify this choice by the fact that an estimate of the extinction will be available based on photometry (as is already the case, work in progress or published work of Chen et al. (2018)). We impose lower and upper limits to this parameter and used the values -0.1 and 20; I have chosen in purpose an unrealistic negative value to avoid a bias around zero.

-The parameters m and q are very important because they define the continuum, that must be realistic in order to avoid secondary minimum of the χ^2. I imposed for the initial gradient m_0=0 and for the *y-intercept*a_0 as the maximum value of the simulated flux in the spectral interval used for the fit.

-β_0 is equal to 1 and corresponds to the case of a perfect stellar model

-vsini$_0$ is the input vsini, used to simulated the spectra. This is justified by the fact that vsini will be deduced from the width of the stellar lines.

-ϵ_0 is taken equal to 0.5, a median value between the limits for this parameter (0,1).

-There are not Doppler effects associated to a radial velocity in these simulated spectra, since Vrad=0, but they will exist in real Gaia spectra. For this reason I inserted its influence in the fit function as already discussed, but fitted values were compatible with zero.

α parameter is the most important because is directly connected with the DIB absorption and the fitted A_V.

Spectral interval used for the adjustment The adjustment was not done along all the whole RVS spectral interval, but in a smaller region, between 860.6 and 863.6 nm. I have chosen these limits after some tests because it provides equivalent results for a shorter computation time, as illustrated in one typical

case shown in Fig 5.8. The figure displays adjustments to the same spectrum but using two different spectral intervals; in the top panel the fit is done on the entire RVS band and in the bottom panel on the shorter interval. It can be seen from the top panel that the model reproduces very well the spectra over the whole RVS interval, and that our assumption of a linear continuum is acceptable, even for the entire interval. On the other hand, here, as in other cases, the fit results for the fit in the complete interval are generally not closer to the input values than for the fit in the limited region. E.g. $\alpha = \mathrm{A}_{V\ fit}$ is 0.93±0.10 mag for the entire interval and $\alpha = \mathrm{A}_{V\ fit}$ is 1.06±0.17 mag in the smaller interval, while the input extinction is 1.2 mag. The fit on the complete interval being not significantly better but slower, I decided to use the smaller interval.

In Fig 5.9,5.11 and 5.12 are shown some example of fits. Fig 5.9 shows two stellar spectra corresponding to the same stellar parameters and the same absolute V magnitude, but with different extinctions Av - respectively 5.6 mag and 0 mag. The extinction changes both the flux level and the DIB depth. The fitted α value is indicated on the figure.

Fig 5.11 and 5.12 are examples of fit for one cold and one hot star. In the second case the stellar model is almost featureless in the fit region, this is the case for most stellar types above $\sim$ 8000 K.

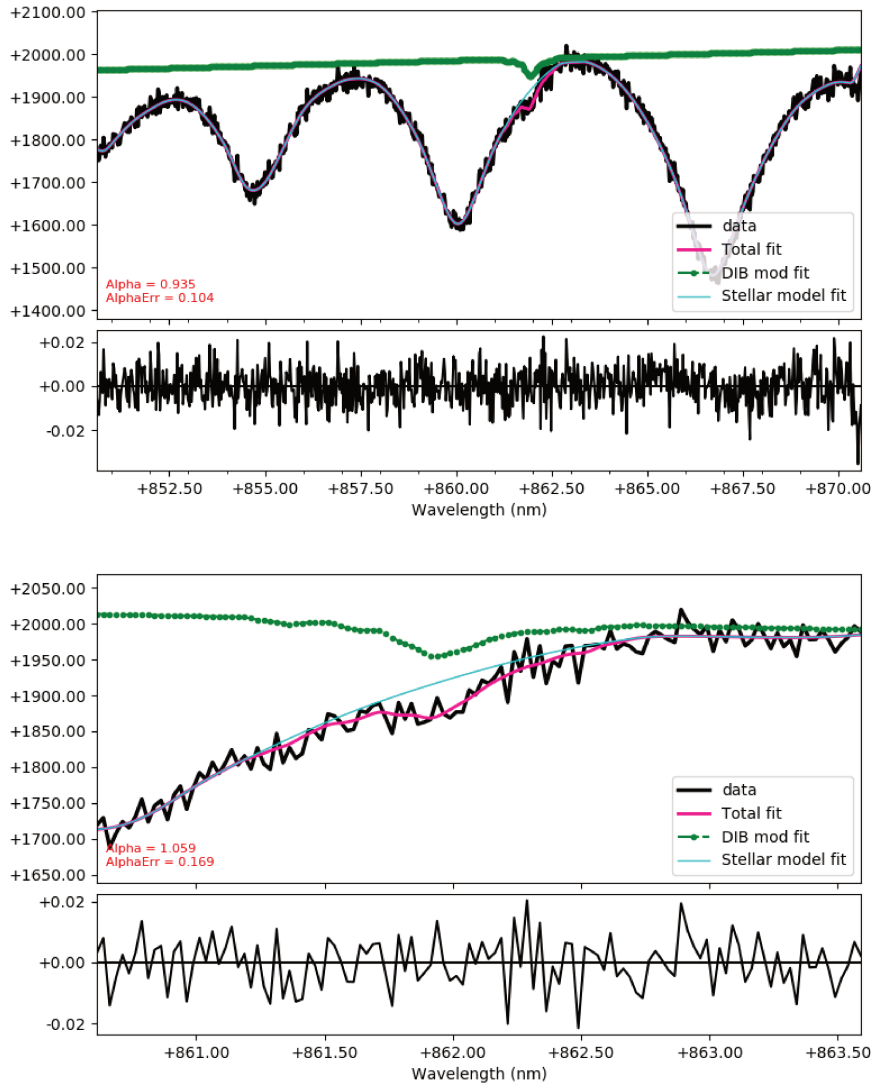


Figure 5.8 – Comparison between adjustments to the same spectrum (effective temperature = 10000 K; log(g) = 4.61 ; [Fe/H] = -3.0; input Av is 1.2 mag.) Top) Fit done over the full RVS spectral interval. Bottom) Fit done over a smaller interval around the DIB (860.6, 863.6) nm.

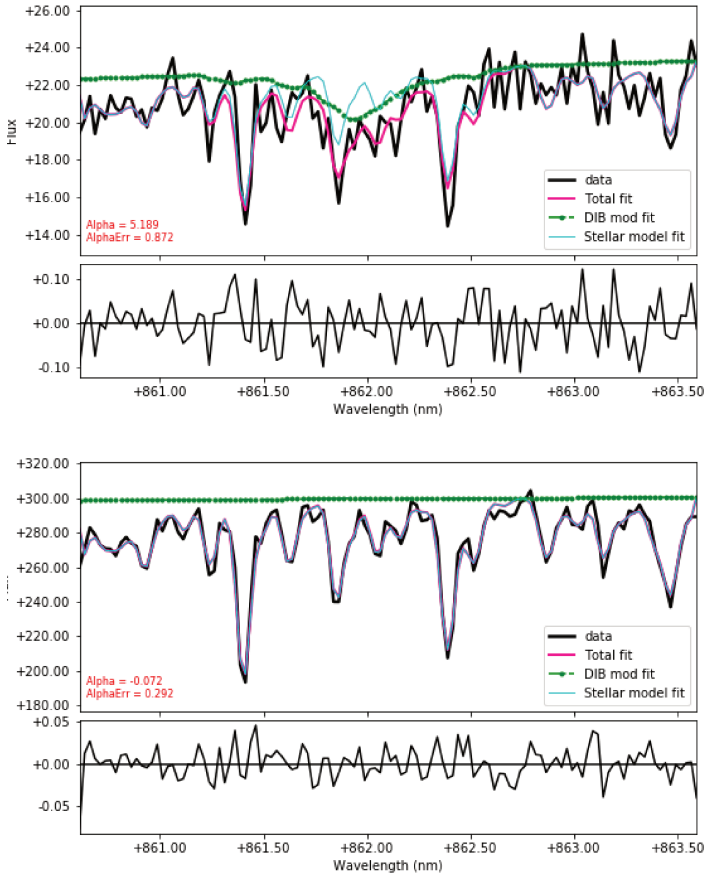


Figure 5.9 – Example of adjustments for two stars with identical stellar parameters, T = 4500, log(g) = 1.01 and [Fe/H] = 0.0, but different input extinctions for the simulation. -Top) A_V input= 5.4. -Bottom) A_V input= 0. The green line is the fitted DIB profile, the light blue line is the adjusted stellar model and the pink line is global fit. Note that in the bottom panel, the two blue and pink lines are superimposed, producing the violet color.

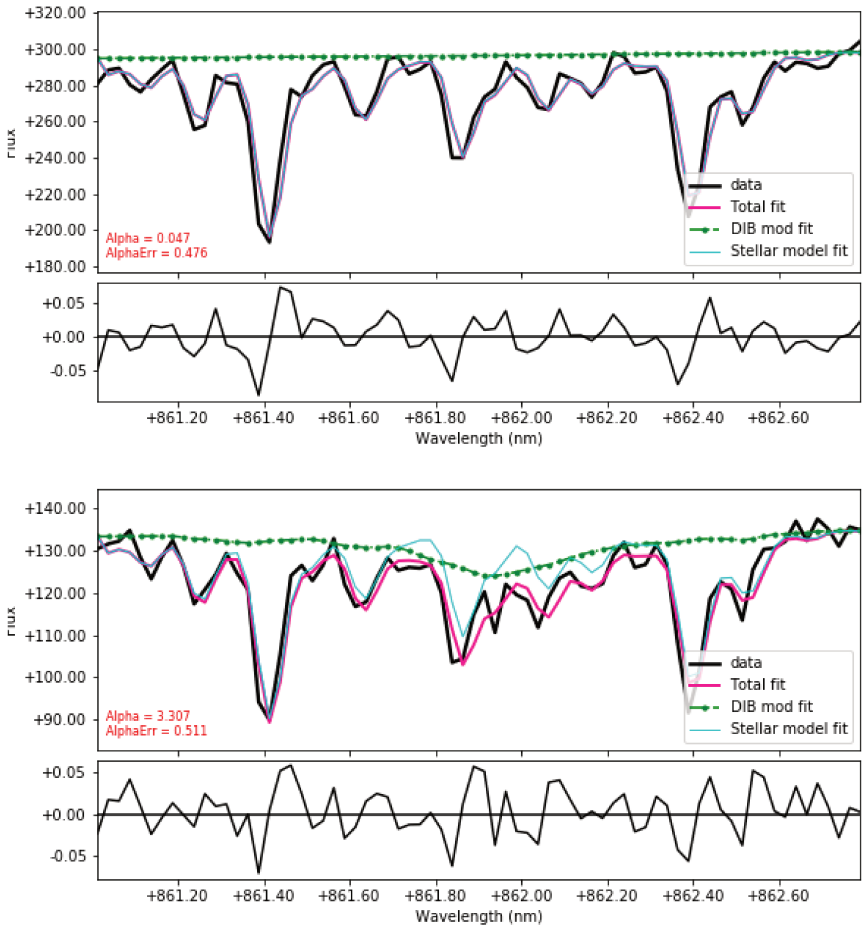


Figure 5.10 – Same as fig. 5.9 for a brighter star and A_V input = 0 (top), A_V input = 3.0 (bottom). The spectral interval is narrower to better display the stellar lines and the DIB.

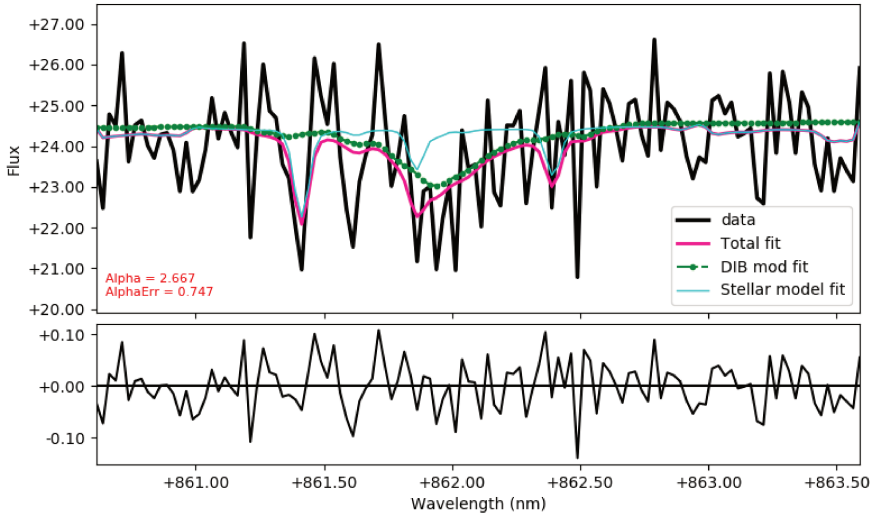


Figure 5.11 – Exemple of fit in case of a low SNR and a cool star, T = 5500 K, log(g) = 2.21, [Fe/H] = -1.0, A_V input = 2.4 mag.

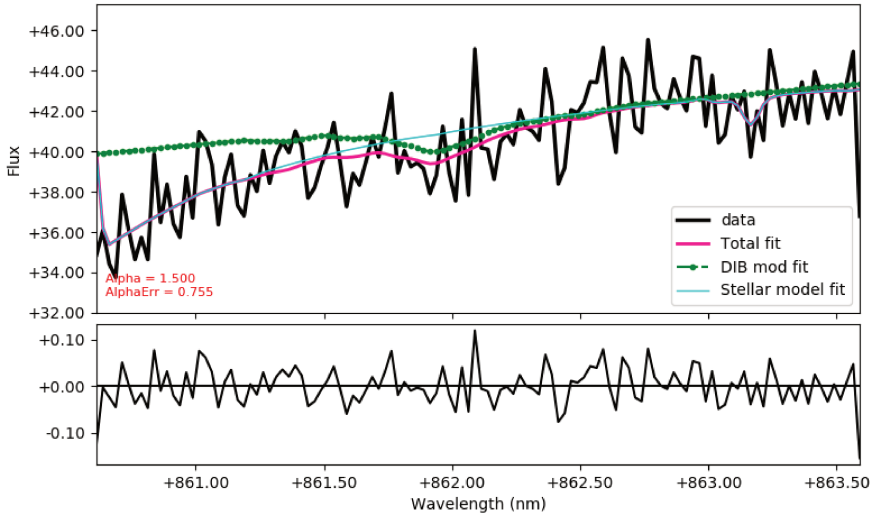


Figure 5.12 – Example of fit in case of a low SNR and a hotter star T = 10000 K, log(g) = 2.21, [Fe/H] = 0.5, A_V input = 1.8mag.

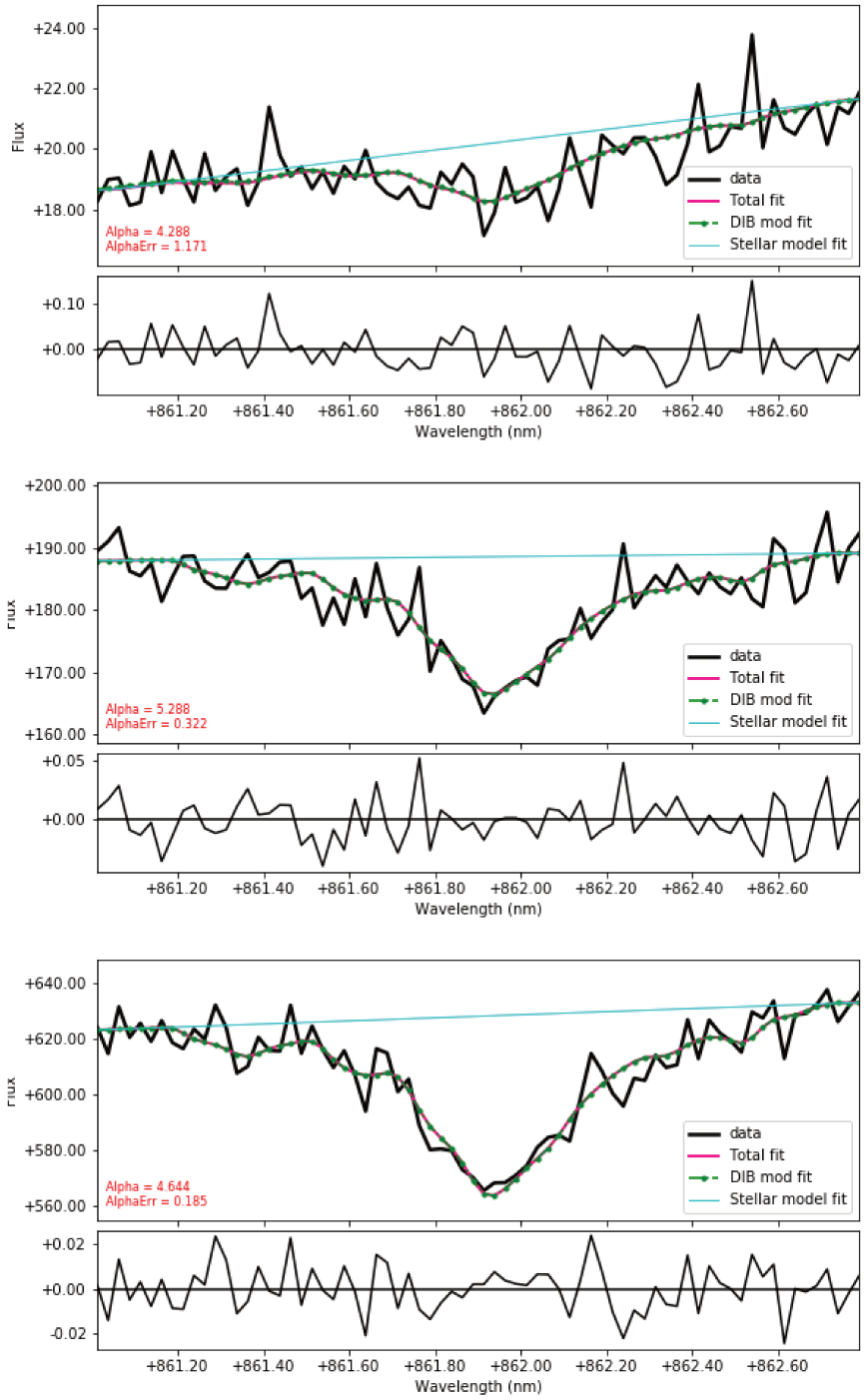


Figure 5.13 – Example of spectra and fitted models, for three stars with different parameters: Top: T = 8500 K, log(g) = 2.21, [Fe/H] = 0.5, Vsini = 10.0, A_V input = 5.4; Middle: T = 30000 K, log(g) = 2.21, [Fe/H] = 0.5, Vsini = 50.0, A_V input = 5.4; Bottom: T = 30000 K, log(g) = 2.21, [Fe/H] = 0.5, Vsini = 10.0, A_V input = 4.8

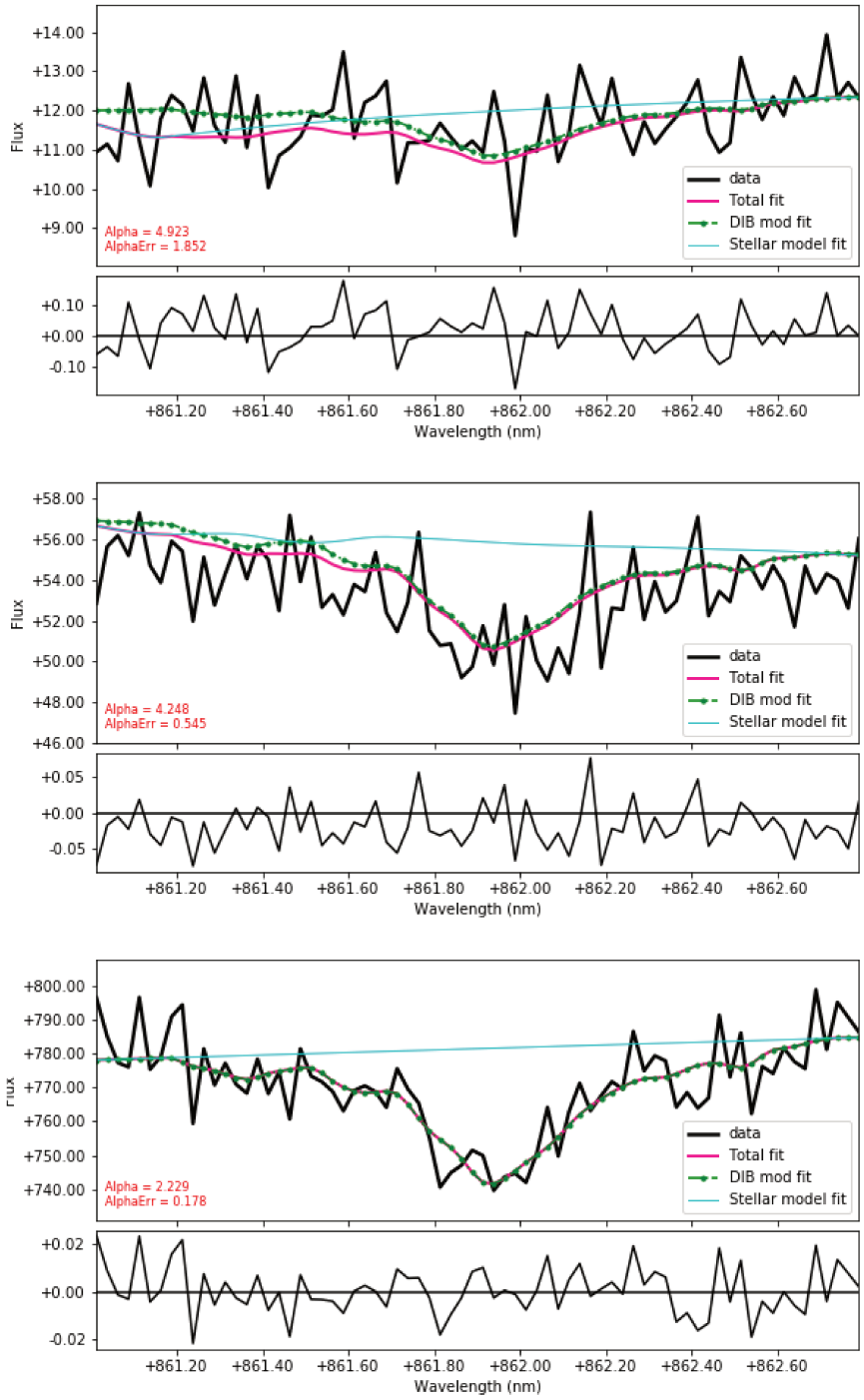


Figure 5.14 – Example of spectra and fitted models, for three stars with different parameters: Top: T = 10000 K, log(g) = 2.81, [Fe/H] = 0.5, Vsini = 50.0, A_V input = 4.2; Middle: T = 25000 K, log(g) = 2.81, [Fe/H] = 0.5, Vsini = 50.0, A_V input = 4.2; Bottom: T = 30000 K, log(g) = 2.21, [Fe/H] = 0.5, Vsini = 50.0, A_V input = 2.4

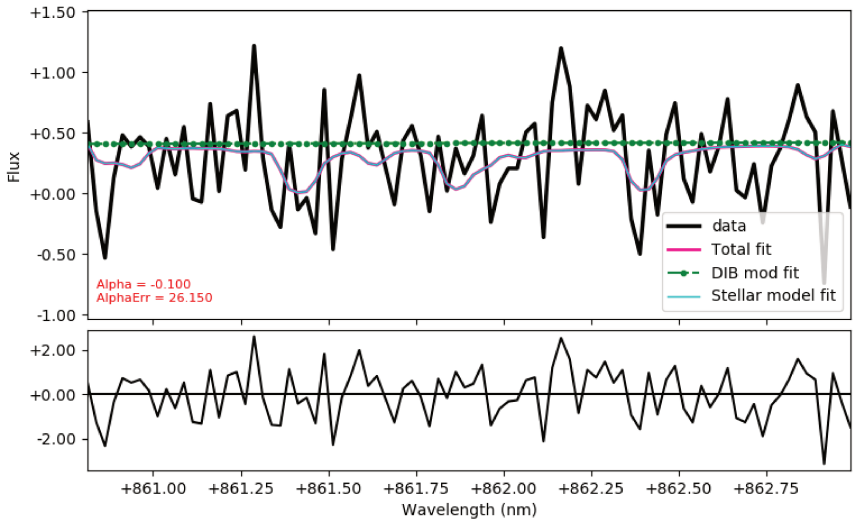


Figure 5.15 – Example of fit for a very low SNR, T = 5500 K, log(g) = 2.21, [Fe/H] = -0.5, A_V input = 1.8 mag. Such spectra will not be useful.

Chapter 6

Gaia Universe Model Snapshot

6.1 Gaia Universe Model Snapshot: generalities

The Gaia Universe Model Snapshot (GUM) is a simulation of the Milky Way objects, developed at the Besancon observatory, see Robin et al. (2012). This simulation was done to produce a useful tool to make predictions and to prepare the production and exploitation of the Gaia catalogues. It is based on the fully developed Besancon model of Milky Way stellar populations, which among other qualities has the advantage of being self-consistent from the dynamical point of view.

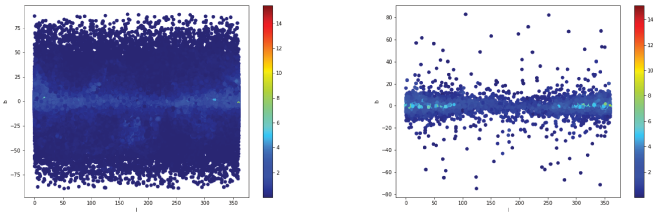
The model distinguishes four different stellar populations: the thick disk, the thin disk, the stellar halo and the bulge. White dwarfs are computed separately. All the objects are seen through a three-dimensional extinction model, based on Drimmel and Marshall dust maps - at very large and large distances respectively.

Using a very complex modelling (Robin et al., 2003) that I will briefly present in the following paragraph, they computed the number of stars for any given age, stellar type and parameters, absolute magnitude, and how they are distributed all around the Galaxy. In their model the Sun is assumed to be at 8.5 kpc from the Galactic center and 15 pc above the plane. A series of Monte Carlo drawings are performed, knowing the expected density for each stellar type.

Apparent magnitudes and star colours as seen at the Sun are computed star by star following the luminosity in wavelength intervals along the path from the star to the Sun, in repsonse to the reddening due to the encountered dust. Astrometry and kinematics are calculated for the different stellar populations and for different ages using several initial velocity dispersions and velocity dispersion gradients. At the end of the entire process a simulated catalog is produced, including both intrinsic and observable stellar parameters, the star positions and each interstellar extinction.

GUM simulations are publicly available. I used one of the GUM simulations available under the form of twelve files, corresponding to Healpix decompositions. These files are in the HDF5 format and huge (up to 350 MO), therefore I did not concatenate them and worked on them always individually.

I used in this study the last version of November 2017. Galactic objects are simulated using the BGM model, which provides stellar parameters, distribu-



(a) T = 7500 (b) T = 25000

Figure 6.1 – Distribution on the sky of objects extracted from the GOG simulation for two different temperature intervals, centred respectively at 7500 K and 25000 K. Colour scale is the excess of colour.

tions and kinematics for single stars. Not only single stars, but also binaries are part of GUM, but their distributions are calculated in a different way, because the BGM does not contain them. I am not considering the binaries in this study, because I developed the fitting method only for single stars. The binaries will be automatically eliminated from my selections in GUM, because they have no defined stellar parameters, therefore our GOG-GUM links will not be affected by the binaries.

GUM is a huge, complete catalog of **3 542 213 066** emitting points, and contains **136 963 318** stars for which the magnitude G RVS is defined and lower than 15 mag. I selected only objects with magnitude lower then 15 mag for the RVS, because the RVS SNR law shows that for objects weaker than 15 mag there will not be a sufficient signal to perform any spectral analysis (see our estimates in the previous chapter) [1].

GUMs parameters useful for my thesis

- effective temperature teff
- surface gravity log(g)
- metallicity [Fe/H]
- rotation velocity vsini
- extinction A_V
- magnitude G for the RVS instrument mag g rvs

6.2 Using GUM for the predictions of Gaia DIB extractions

As we mentioned, we used very wide intervals for the stellar parameters in our GOG simulations. Some are not corresponding to realistic objects and we must first calculate how many stars simulated with GOG have parameters

[1] (https://www.cosmos.esa.int/web/gaia/sn-rvs)

	All objects	Objects with magnitude lower than 15
UM00	131617189	9759873
UM01	28025830	3151767
UM02	23979085	2488923
UM03	369245499	18044246
UM04	22112069	2491353
UM05	131942564	10002968
UM06	22277852	2371755
UM07	1410836067	35755946
UM08	22065865	2514418
UM09	296022539	15659298
UM10	897667889	27506142
UM11	186420618	7216629

Table 6.1 – For the different parts of Universe Model (UM files), how many stars with magnitude G RVS lower than 15 are contained in the file.

that are actually in the GUMs. This includes both stellar and observational parameters, since I want to quantify the fits results. To do so, I looked in each GUM HDF5 file (UM) parts for the parameters I varied in GOG. For example I'm interested in variation in extinction but not in radial velocity because I did not vary this value in GOG simulations. Hereafter are listed the intervals for the varied parameters (the sign ... means regularly spaced intervals similar to the previous one), and, for the twelve HDF5 files, the numbers of objects in the file that have the corresponding characteristics.

Temperature:

- 3500 K= (3000 : 4000)
- 4500 K= (4000 : 5000)
- ...
- 25000 K= (22500 : 27500)
- 30000 K= (27500 : 33500)

Gravity:

- 1.01 = (0.7 : 1.3)
- 1.61 = (1.3 : 1.9)
- ...
- 4.01 = (3.7 : 4.3)
- 4.61 = (4.3 : 4.9)

[Fe/H]:

- -3.0= (-3.2 : -2.8)
- -2.5= (-2.7 : -2.3)

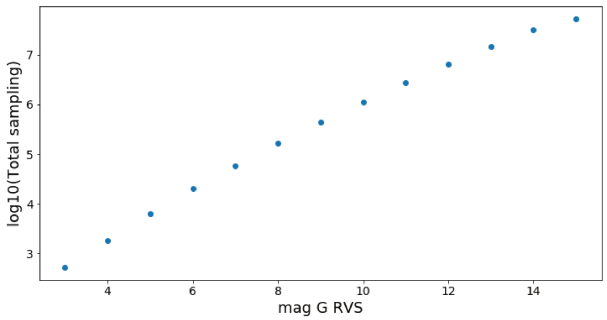


Figure 6.2 – How many objects meeting our criteria are parts of GUM for a series of magnitudes.

- ...
- 0.5= (0.3 : 0.7)
- 1.0= (0.8 : 1.2)

Vsini:

- 0 Km/s= (0 : 2.5) Km/s
- 5 Km/s= (2.5 : 7.5) Km/s
- 10 Km/s= (7.5 : 25) Km/s
- 50 Km/s= (25 : 75) Km/s
- 100 Km/s= (75 : 125)Km/s

Extinction A_V:

- 0 mag = (-0.3 : 0.3) mag
- 0.6 mag = (0.3 : 0.9) mag
- ...
- 4.8 mag = (4.5 : 5.1) mag
- 5.4 mag = (5.1 : 5.7) mag

Mag G RVS:

- 3 mag = (2.5 : 3.5) mag
- ...
- 14 mag = (13.5 : 14.5) mag
- 15 mag = (14.5 : 15) mag

	All objects	With all conditions applied
UM00	131617189	810530
UM01	28025830	287086
UM02	23979085	226168
UM03	369245499	1548238
UM04	22112069	230983
UM05	131942564	824090
UM06	22277852	215627
UM07	1410836067	4196926
UM08	22065865	234613
UM09	296022539	1327605
UM10	897667889	2749157
UM11	186420618	4699191

Table 6.2 – For different parts of GUM, how many stars meet the conditions.

Condition on A_V	How many in GUM
0.0	23225311
0.6	29145529
1.2	13094491
1.8	6824809
2.4	4493995
3.0	2924968
3.6	1777819
4.2	1059270
4.8	616042
5.4	340754

Table 6.3 – How many stars in our bins of different extinctions, A_V.

Condition on Temperature	How many in GUM
3500	4386362
4500	31772611
5500	23551613
6500	19391691
7500	2427791
8500	964486
10000	873526
15000	108164
20000	16222
25000	6799
30000	3723

Table 6.4 – How many stars in our bins of effective temperature.

Condition on magnitude	How many in GUM
$2.5< 3 < 3.5$	522
$3.5 < 4 < 4.5$	1853
$4.5 < 5 < 5.5$	6239
$5.5 < 6 < 6.5$	20094
$6.5 < 7 < 7.5$	58955
$7.5 < 8 < 8.5$	166722
$8.5 < 9 < 9.5$	437047
$9.5 < 10 < 10.5$	1102757
$10.5 < 11 < 11.5$	2702765
$11.5 < 12 < 12.5$	6428939
$12.5 < 13 < 13.5$	14549599
$13.5 < 14 < 14.5$	31270207
$14.5 <15$	26757289

Table 6.5 – How many stars in our bins of magnitude.

Chapter 7

DIB measurements with the RVS: a global estimate

My goal is to establish an estimate of the total number of DIB measurements of reasonable quality we can expect from the RVS instrument. To do so we will use our results for the GOG simulated and fitted spectra for each parameter set:

$$\Pi = [\text{T, log(g), [Fe/H], vsini; visual extinction, absolute magnitude V}]$$

$$= [\pi, \text{absolute magnitude V}]$$

and consider that they are representative of what would be obtained for all the GUM stars with the same parameter set.

First of all, we recall that varying the absolute magnitude V, as we did, does not correspond to real variable physical conditions, because we know that the absolute value is intrinsic for each stellar type: we used the variability of the number entered in GOG for the absolute magnitude V as a way to vary the apparent magnitude instead of varying the distance. The extinction also contributes at the apparent magnitude variation, because for the same simulated absolute magnitude V, a stronger extinction results in a lower luminosity (higher apparent magnitude). In Chapter 5 the Figure 5.9 shows an example of the influence of the extinction for two simulations with the same absolute magnitude, i.e. the mean absolute V is 4.0 for the two spectra, but the extinction changes the total flux, and as a consequence the magnitude in the RVS spectra interval.

After all the calibration phases, the fluxes measured with the RVS instrument are associated with a specific magnitude, called G (RVS). G (RVS) depends on the incoming stellar flux but also from the instrument it-self, because it is an operational magnitude (Recio-Blanco et al., 2016).

From the GUMs, I extracted the total number of objects in 6D boxes, using as Σ parameters, effective temperature, gravity, metallicity, rotational velocity, extinction and magnitude G (RVS). It means that we have disentangled for the GUMs counts the influence of the extinction and the magnitude.

The idea is to understand for each spectrum with parameter set Π_i simulated with GOG, how many similar objects are in GUMs in a corresponding Σ_j interval. The GOG interval is centred in temperature, gravity, metallicity, vsini and extinction as for the Π GOG simulations, but when converting into

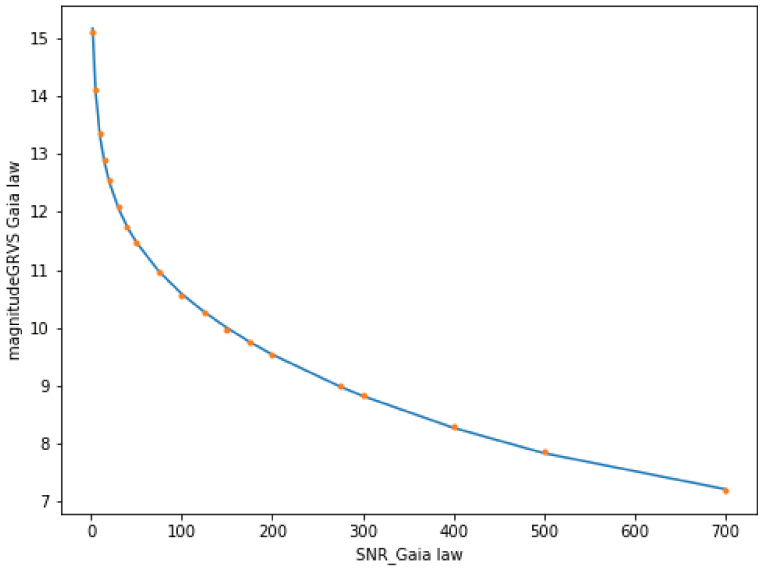


Figure 7.1 – Law to connect the SNR and the magnitude G (RVS). Orange dots are the point of the law available on the Gaia website, the continuum line is the fitted function.

intervals for objects in GUMs we must know at which magnitude G (RVS) corresponds the simulated GOG spectrum, i.e., to go from the i-spectra in GOG to the number of objects in GUM, I need to calculate the magnitude G (RVS) for the GOG spectra.

7.1 Estimating the magnitude G (RVS) corresponding to each simulated GOG spectrum

There is a well calibrated relationship SNR=f(G (RVS)) between the magnitude G (RVS) and the SNR of the RVS spectrum[1]. It will be central in all what follows. For our GOG simulations we do not have the magnitude G (RVS), but we can estimate the SNR in our simulated spectra, and then use the inverse relationship G (RVS)= f^{-1}(SNR) to compute the magnitude G (RVS). The inverse relationship is shown in Fig 7.1, and we modelled it as:

$$magnitudeG\ (RVS) = (a * SNR * *b) * (c * SNR * *2 + d * SNR + e) \quad (7.1)$$

,

Using this relationship, we can convert the SNR of our GOG synthetic spectra into a magnitude G (RVS), the parameter present and sampled in GUM for all the Σ intervals.

[1]https://www.cosmos.esa.int/web/gaia/sn-rvs

I tested two ways to derive the G (RVS) magnitude for each i-spectrum. Both are based at some level on the simple idea to derive the noise as the difference between the GOG simulated spectrum and the fitted model, and find the standard deviation (or SNR) in the residuals. To do so we have to assume that the model is excellent, but this is the case in our simulations.

The first method uses each fit and computes the standard deviation in the fit residuals (i.e., in the spectral interval used for the fit), then transforms this estimate of noise into a SNR by dividing by the median flux in the fit interval

$$SNR = \frac{median_{model\_flux}}{standard\_deviation\_of\_noise}$$

and in a last step converts the obtained SNR into a G (RVS) magnitude, by applying the formula 7.1 derived from the Gaia calibration. In this way each spectrum has a G (RVS) magnitude calculated in the fit interval, with its-own SNR, depending on the fit.

For the second technique, we use a series of test spectra to build a law connecting the flux level in the GOG simulation to the magnitude G (RVS). Seven spectra were selected as test - shown in Fig 7.2; they have same π parameters T = 10 000 K, log(g) = 4.01, [Fe/H] = -0.5, vsini = 10 km/s, A_V = 0.0, and different mean absolute magnitudes V = (4, 5, 6, 7, 8, 9, 10). As they have not extinction, I used directly the synthetic spectra without noise and with the simulated noise to calculated the standard deviation of the noise on the whole spectra.

This time, I derived the SNR on the entire spectrum:

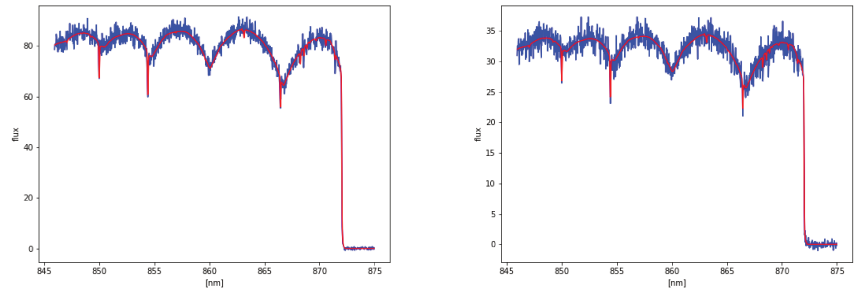
$$SNR = \frac{max_{model\_flux}}{standard\_deviation\_of\_noise}$$

For the seven spectra I calculated their SNRs and then the corresponding magnitudes G (RVS) (again from the general law described above). In used this list of seven flux values and magnitudes G (RVS) to establish a flux - magnitude G (RVS) relationship, a relationship that can be applied to all median flux values in the i-spectra of GOG. In short, instead of using fit results, and independently of the fits, we transform the flux in the GOG spectrum into a SNR, then into a G (RVS).

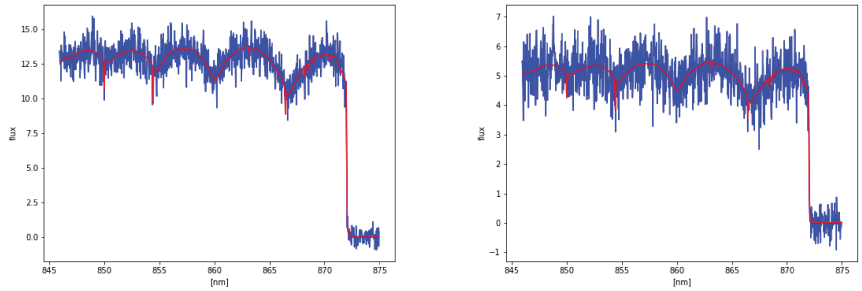
Comparing the results for the two approaches, as shown in Fig 7.3, it is evident that the first method has a large dispersion at small magnitudes, but globally they give comparable results. Our simulated spectra cover very well the range between G (RVS) (9 : 16), with only some spectra at lower and higher magnitude. The RVS catalogue will be published with also $\sim$ 700 000 objects with magnitude G (RVS) lower than 9, which are not part of our simulation; as we have already said, for those objects the DIB will be accurately measured in all cases. On the other side, as the SNR is very low for objects with magnitude G (RVS) higher than 15, the loss of objects in that range does not change our estimate.

7.2 Estimates of DIB measurements from RVS spectra

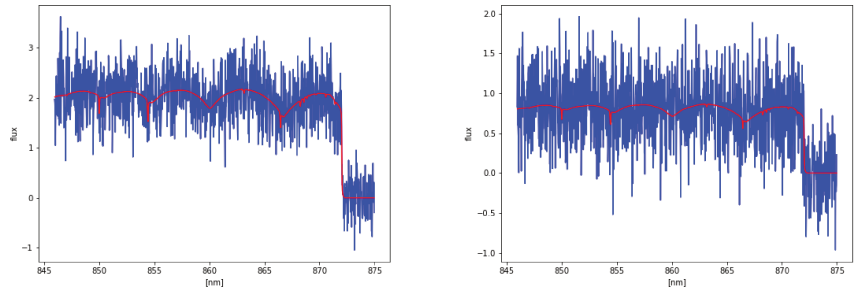
Having the magnitude G (RVS) for each i-spectrum from the GOG simulations, it is possible to connect to each simulated spectrum the number of correspond-



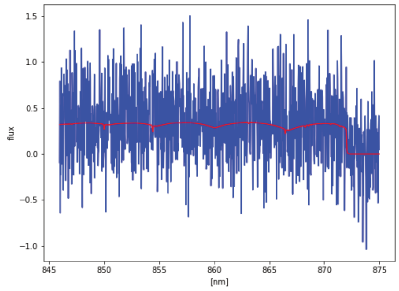
(a)



(b)



(c)



(d)

Figure 7.2 – Spectra simulated with noise (blu line) and their stellar model (red line) used for the estimation of magnitude G (RVS) - first method.

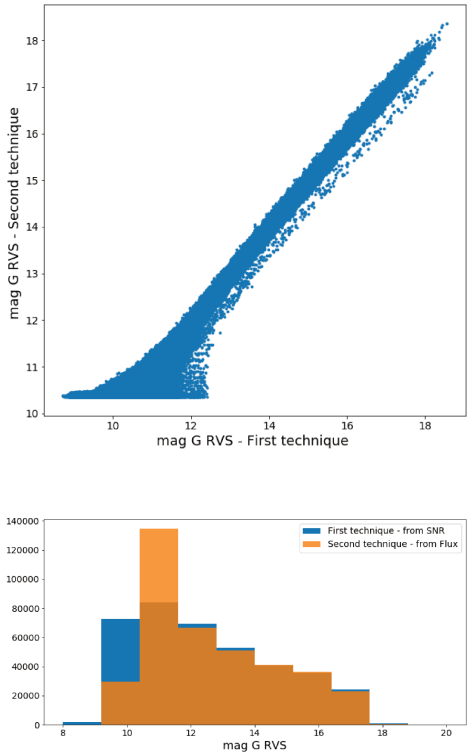


Figure 7.3 – Comparison between the two techniques to estimate the magnitude G (RVS).

ing objects in GUM, using all computed numbers of objects in the Σ intervals of our GUM sampling.

On one side, we have all GUM object counts, with magnitude sampling by steps of 1 magnitude. The total number of objects in GUM for the Σ parameters is 83 502 988 objects distributed in the 27 012 Σ parameters.

On the other side, the GOG simulation results contain 381 900 Π parameter combinations (381 900 simulated spectra). As we knew, a large fraction of the parameter sets are physically un-realistic and have no counterpart as a measurable object, even if they have a good fit. However, this will not impact our final statistics. Note that the number of failed fits is very small, there was no convergence for only 198 simulated spectra.

We can now start to find for each π group of parameters [T, log(g), [Fe/H], vsini, A_V] the total number of objects for which we can have DIB measurements. However we have to define a criterion for the quality (or possibility) of the DIB extraction from the data.

I selected in GOG the list of spectra with the same combination of π parameters. I removed the failed fit cases, then extracted from the GOG simulated spectra and their fits the list λ of magnitude G (RVS) and the list of fit results, especially the values $A_{V\,fitted}$ - or the α parameter in my fit - and their associ-

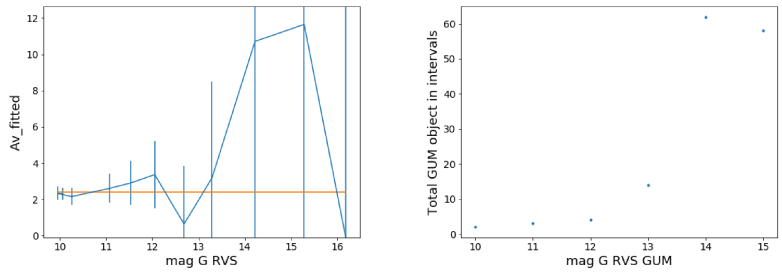


Figure 7.4 – For a π = [T = 20000.0, log(g)= 3.4, [Fe/H] = 0.5, vsini = 10.0, A_V (input) = 2.4] selection of parameters Left) For the GOG spectra with π parameters, the blue line connects the fitted A_V results depending on the magnitude G (RVS) and the error bar is the double of the calculated error for the fit; the orange line corresponds to the input value for A_V - the "true" results guessed from the fit. Magnitude G (RVS) for the GOG simulation is calculated in this example with the first technique.Right) How many corresponding objects contains GUM for each G (RVS) magnitude interval.

ated error bars. Fig 7.4, 7.5, 7.6, 7.7 show some example of λ lists, and displays how the fitted value for Av, Av_{fitted} and its associated error (blue curve) vary as a function of the magnitude G (RVS). The comparison between the blue curve and its errors with the orange line that represents the "true" value for our fit (the input value of $A_V = 2.4$ mag) shows that for this set of parameters the DIB equivalent width will be extracted with reasonable accuracy (at better than 30%) for G(RVS) smaller or on the order of 12. Beyond this magnitude limit, the fitted value diverges very rapidly from the input value and error bars become extremely large.

In the following we will consider four different criteria of quality for the DIB extraction and for each of this criteria estimate the number of objects for which this criterion is achieved. We call the criterion COND. For each criterion we follow the same procedure, explained below:
Given COND, I select in each λ list as above (corresponding to a set of parameters π) only the magnitude intervals which satisfy COND and define a minimal and a maximal value for the G (RVS) magnitude interval - [mag_GOG_cond_MIN ; mag_GOG_cond_MAX]- that is acceptable with respect to the condition COND. I.e., a star with stellar parameters and extinction corresponding to the π selection and within this range of magnitude may provide a fit with an *interesting* result.

In order to derive how many objects correspond to these *interesting* results, I use the function $N_{object} = f(mag\_G\_RVS)$ that relates the magnitude G (RVS) and the total number objects with the π parameters in GUM - see some example of this function in Fig 7.4, 7.5, 7.6, 7.7 . At the end, the integral of N_{object} in the interval [mag_GOG_cond_MIN ; mag_GOG_cond_MAX] gives the total number of objects for which RVS observations will satisfy the condition COND.

A first condition COND0 is whether the fitted value A_V (fitted) from a

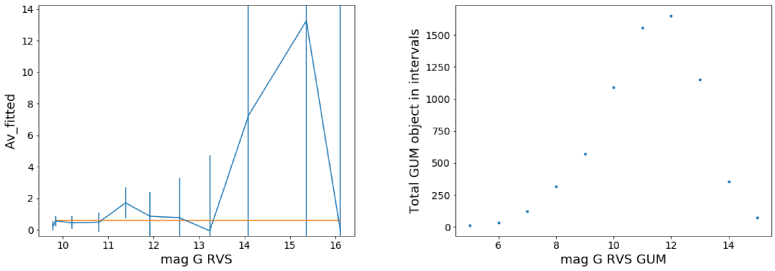


Figure 7.5 – Same than Fig 7.4, for a π = [T = 5500.0, log(g)= 2.21, [Fe/H] = 0.0, vsini = 10.0, A_V (input) = 0.6]selection of parameters.

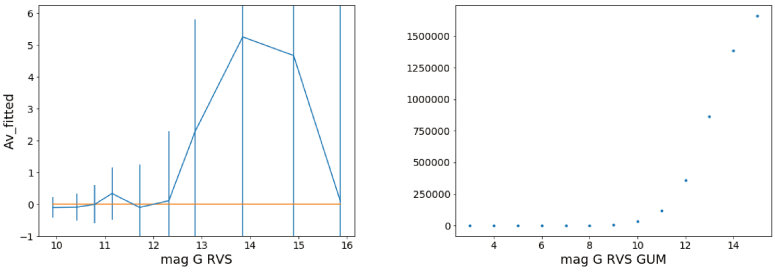


Figure 7.6 – Same than Fig 7.4, for a π = [T = 5500.0, log(g)= 4.61, [Fe/H] = 0.0, vsini = 0.0, A_V (input) = 0.0] selection of parameters.

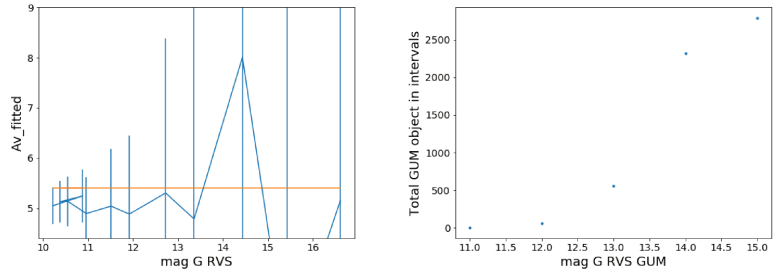


Figure 7.7 – Same than Fig 7.4, for a π =[T = 3500.0, log(g)= 1.01, [Fe/H] = -0.5, vsini = 0.0, A_V (input) = 5.4] selection of parameters.

Table 7.1 – Numbers of stars satisfying to the different criteria associated to the fit: the two estimates correspond to the two methods used to derive the magnitude G (RVS) for each GOG simulated spectrum.

	Total objects with estimated magnitude by the SNR for each spectrum in small range	Total objects with estimated magnitude by the SNR through 7 spectra in all RVS band
No condition on the fit	67 375 215	67 377 777
Reproducibility condition - COND0	56 241 030	56 231 415
Associated error 0.5/ 50% - COND1	3 942 016	3 738 784
Associated error 0.3/30% - COND2	1 339 523	1 368 433
Associated error 0.15/15% - COND3	81 478	81 533

data-model adjustment such as those we simulated is able to reproduce the input value A_V. I translate this requirement into our first condition COND:

$$Av_{input} \cap (Av_{fitted} \pm 2error_{fit})$$

a condition that must be apply to the λ intervals for each π set. Testing for this condition is interesting since it gives us some insight on the conditions of reliability of the fitting method, however it will not be reproduced at the time of real data analysis, because we will not know the *real* extinction; on the other hand, extinction measurements will be available, based on photometry and distances, and testing this condition provides us with some limits on future comparisons between photometric extinctions and DIB absorptions.

Other requirements correspond to a limitation on the uncertainty on A_v(fitted) (COND1,COND2,COND3 respectively:

- for spectra with a $Av_{fit} < 1$ only fits with an associated absolute error on the α parameter in the fit respectively < 0.5 / 0.3 / 0.15 mag will be acceptable

- for spectra with a $Av_{fit} > 1$ only fits with associated relative error on the α parameter in the fit respectively $< 50\%$ / 30% / 15% are acceptable

These criteria are the most useful since they predict the numbers of potential fits with limited uncertenties, i.e. the numbers of objects in a final DIB catalog.

Tab 7.1 contains the total number of objects estimated in this way for the whole set of GOG simulated spectra, using the four different criteria COND 1,2,3,4. Since we can associate to the GOG spectra two different estimated magnitudes G (RVS), we end up with two different estimates of the numbers of objects, because different limiting magnitudes correspond to different limits for the integration. However, it can be seen that the order of magnitude of the two estimates is very similar.

Table 7.1 shows that DIB extractions using good stellar models and with relative errors better than 30% should be achievable for an unprecedented number of objects, on the order of 1.3 millions. Since the $\sim$ 700 000 objects with magnitude lower than 9 mag G (RVS) will have a very high SNR, they will be used for the DIB measures, despite their absorption will be probably very small, the number of objects with a good DIB measure associate will be on the order of 2 millions.

Table 7.2 – Statistics on extinctions for targets stars with DIB detections at conditions COND1 and COND2.

A_V [mag]	COND1	COND2
0.0	1 168 245	553 220
0.6	813 278	211 713
1.2	535 581	54 523
1.8	392 596	148 227
2.4	246 939	116 238
3.0	366 214	90 297
3.6	175 257	69 534
4.2	112 861	42 464
4.8	65 344	38 208
5.4	62 157	12 501

Table 7.3 – Statistics on distances of targets stars with DIB detections at condition COND1 and COND2

	COND1	COND2
0 < dist < 1000 [pc]	1 439 928	631 250
1000 < dist < 2000 [pc]	602 581	159 744
2000 < dist < 3000 [pc]	561528	139 838
3000 < dist < 4000 [pc]	397 939	142 233
4000 < dist < 5000 [pc]	283 772	100 618
5000 < dist [pc]	652 724	163 242

Statistics on the extinction I also binned the results for extinctions intervals, in order to understand to which range of extinctions our fit applies. In Fig 7.8 is shown the estimated number of detections for different extinctions ranges; higher extinctions correspond objects because the magnitudes become too high. Our global count in this case is similar for conditions COND1 and COND2, we see that we will measure accurately the DIBs in those rare case of both sufficient SNR and high extinction. Results are summarised in Tab 7.2.

Statistics on the distances I have binned the results in intervals of distances. The results are shown in Tab 7.3. It is clear that at least 40% - 50% will be closer than 1 kpc for COND1 - COND2 which is not surprising. On the other hand, interestingly, hundreds of targets beyond 1 kpc will participated at the detection and even than 1 hundred thousand more distant than 5 kpc with DIB detected at better than 30% (COND2).

Possible influence of GOG simulation on our count

Luri et al. (2014) investigated the difference between the input and output parameter sets for the GOG simulator, using for comparison the Gaia Universe Model (GUM) as a source of all stellar types and magnitudes. They found a good agreement in most cases, as can be seen in Fig 7.9. E.g., in the case of temperature, the GOG library contains temperatures between 850 and 102 000 K, and the maximum difference between input and output values is 640K, with a mean difference of 390K. Therefore, since our temperature grid has intervals of 1000K, we can consider that this parameter is well simulated. On the other hand, there is a significant spread of metallicity output values for a given input value, especially around [Fe/H]=0. Our input metallicity grid has bins of 0.5 dex for a total range -6.5:+4.6, and the average input-output difference is 0.57 dex. Since we have used a perfect "stellar model" for the fit, the fit results for the extinction have not trends for the metallicity: for this reason the possible

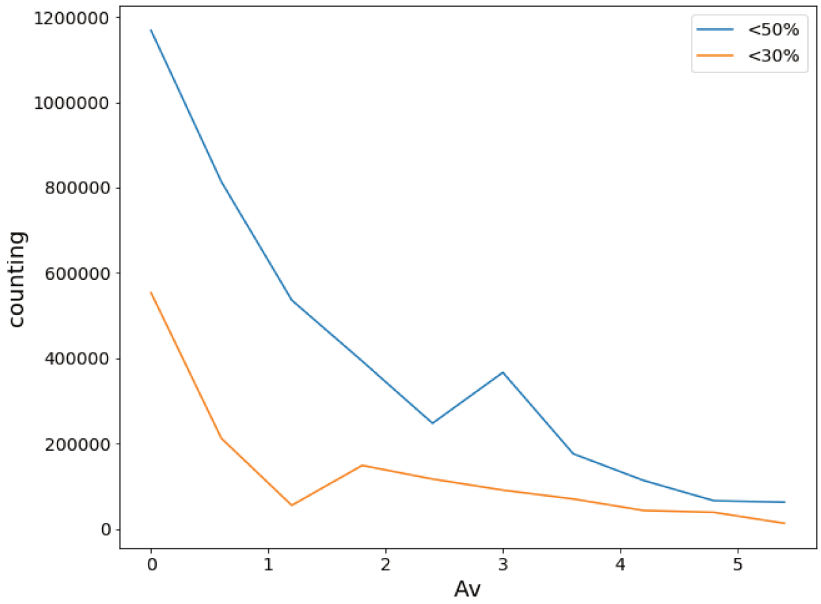


Figure 7.8 – Counting for different extinctions, with two different conditions - COND1 and COND2.

deviation have not influence in our estimate. In Fig 7.10 there is shown the count of objects in GUM and the count for the good fit from the GOG spectra, using the COND1: the number of objects are proportional at every metallicity, ant it means that the metallicity variation has not an high impact on our fit quality. We verified similar trends also for the other COND.

7.3 Conclusion: numbers and characteristics of future Gaia DIB detections

Our results allow us to predict the size of the future Gaia DIB catalogs from the RVS spectra. About ∼4 millions of detections with uncertainties smaller than 50%, ∼1.3 millions of detections with uncertainties smaller than 30%, ∼81 000 of detections with uncertainties smaller than 15% are expected for all stars with RVS magnitude higher than 9, based on our estimation using simulated spectra. Moreover, there will be ∼700 000 additional spectra corresponding at RVS magnitude lower than 9, with signal largely sufficient to measure the DIB than 15%.

Today, the largest catalog for a single DIB is for the APOGEE 1527.3 nm DIB (submitted paper El Yajouri & Lallement, 2018), which contains ∼120 000 DIB detections or upper limits, for 120 000 lines of sight; since it is based from a ground telescope, the measures do not cover all the sky. Moreover APOGEE observations are concentrated in specific fields in the sky observable from the

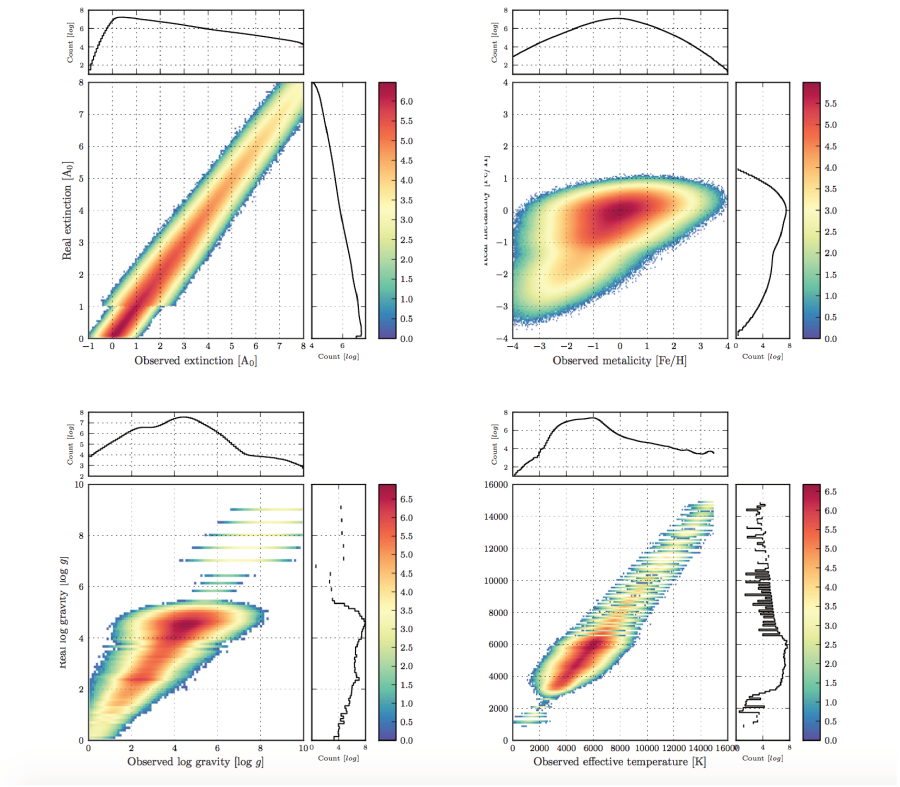


Figure 7.9 – Comparison on the "true" values of physical parameters with the GOG simulated values, with colour scale representing the log density: top left) extinction - bins of 50 mmag, top right) metallicity - bins of 0.4 dex, bottom left) surface gravity - bins 0.5 dex, bottom right) effective temperature - bins 100 K. Credits: Luri et al. (2014).

North Hemisphere. The Gaia DIB catalogs will be tens of times larger, and all the sky will be uniformly covered. It will be the most complete one-DIB catalog available, and it will be fundamental for the future DIB studies.

Having a large DIB catalog, based on uniform on the sky lines of sight and at various distances, we will approach a complete statistical study to constrain the physical environments where the molecular DIB carrier does exist; more numerous measurements of Equivalent Widths and extinctions will allow us probably to measure variation of the DIB to extinction ratio in different ISM regions, and it is particular important to constrain those relationships between the dust and DIB carrier molecule. It is particularly interesting in the context of my thesis that is focused in maps of the ISM based on the DIBs Equivalent Width.

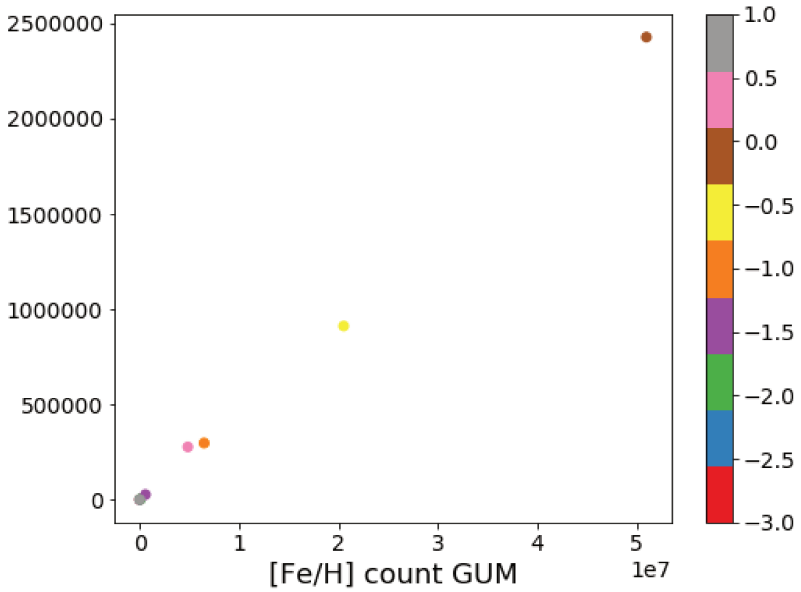


Figure 7.10 – Comparison between the count at different metallicity (colour) in the GUM sampling and for the GOG spectra with the magnitude calculated with the first technique, applying the COND1. As they are proportional, the variation in metallicity does not change the fit quality.

Part III

Mapping improvements in the context of the full 3D tomographic technique

Chapter 8

Description of the full 3D tomographic inversion technique

8.1 Inversion: basic principles

In a milestone work, Tarantola and Valette (1982) formulated a general solution for generalised non-linear inverse problems which, among many projects in other fields, inspired earlier work on inversion of distance-limited columns of interstellar matter (Vergely et al., 2001, 2010b; Lallement et al., 2014) as well as the inversions presented in this thesis.

In direct (or forward) modelling, one uses a theoretical relationship $\mathbf{g}$ and a set of model parameters $\mathbf{p}$ to compute the modelled values $\mathbf{d}^f$ of some measurable quantities. In most cases, the modelled data $\mathbf{d}^f$ are subsequently compared to the actual measurements $\mathbf{d}$ and some method is used to tune the parameters $\mathbf{p}$ in order to minimise the discrepancies between model and data.

$$- \quad \mathbf{d}^f = \mathbf{g} \cdot (\mathbf{p}), \tag{8.1}$$

$$- \quad \mathbf{d}^f \longleftrightarrow \mathbf{d}, \tag{8.2}$$

$$- \quad \textbf{Tune parameters (p) back to (1)} \tag{8.3}$$

where $\mathbf{d}$ are a set of measurements, $\mathbf{g}$ is the operator which expresses the mathematical model of the physical system (usually non-linear) and $\mathbf{d}^f$ are the modelled data. Forward modelling has the advantage of some control on the parameters and estimates of the source of uncertainties: since the predicted values from a forward model usually are not exactly like the measures and this deviation is the combination of measurements errors and imperfections in model, such imperfections can be quantified. Forward modelling is the most common way to model measurements and is widely used.

In inverse problems (as opposed to direct ones) one uses values of measured quantities $\mathbf{d}$ and a theoretical relationship $\mathbf{g}$ to derive information on the values of a set of parameters $\mathbf{p}$.

$$- \quad \mathbf{(d)}, \tag{8.4}$$

$$- \quad +\textbf{theoretical relationship } \mathbf{d} = \mathbf{g} \cdot \mathbf{(p)} \tag{8.5}$$

$$- \quad \mapsto \mathbf{(p)} \tag{8.6}$$

Due to the lack of direct control on the parameters, non-uniqueness of the solution (i.e. different sets of parameters correspond to identical measurements) is more difficult to resolve in the case of inverse problems. More generally, it is crucial for inverse problems to be *well posed*, with existing, unique and stable solutions. In many inverse problems one wants to reconstruct a analytical function in the parameter space with only a finite number of measures and the solution is not unique; in order to prevent an *ill posed* problem one has to impose some regularisation. This is the case for ISM tomography in the actual state of measurements, a situation characterise by a very strong underdetermination. The regularisation may have diverse forms, generally a priori information on the parameters (p_0) and smoothness of the solution through a covariance function linking neighbouring points C (both will be used in the ISM tomography).

Tarantola and Valette have shown that when all informations can be described by Gaussian probabilities the solution of the inversion can always be obtained trough a least squares formalism. In the linear case, the solution can be computed explicitly (as in the very simple example given below) and in the non linear case it can be reached in an iterative way. Here we give only the general form of the iterative solution as taken from Tarantola formulation:

$$p_{k+1} = p_k + (G_k^T \cdot C_{d_0 d_0}^{-1} \cdot G_k + C_{p_0 p_0}^{-1})^{-1} \cdot (G_k^T \cdot C_{d_0 d_0}^{-1} \cdot (d_0 - g(p_k)) - C_{p_0 p_0}^{-1} \cdot (p_k - p_0)) \quad (8.7)$$

A least squares criterion and a threshold for the χ_2 decrease rate rule the end of the interations. In the 3D tomography of Vergely et al. (2001) and subsequent work, the relationship is a an integral in the 3D distribution of interstellar matter and this iterative solution must be used.

8.2 Illustration in (very) simple cases

To illustrate the difficulties encountered in the reversal, let's start with a very simple situation: 5 girls return from a day of shopping during a period of discount. They all bought clothes but spent different amounts of money, respectively 100, 200, 300, 1000 and 2000 euros. The problem is to estimate how many clothes each bought (and then the prices of the various clothes). The problem is highly indeterminate if you only have this information, however there are several additional statistical information:

- in the same shopping area the prices of the dresses remain very similar, but can strongly depart from one shopping area to the other.
- the first three girls shopped all three in the same area, the two others in another area. (These first two conditions can be seen as a regularisation based on spatial smoothing)
  -
- in average a girl buys two dresses during the discount period, sometimes one or three, rarely four (and rarely zero). This third condition corresponds to some "a priori", favoured value.

With this additional information, you can exclude high prices in the first area and low prices in the second, otherwise the number of dresses would be

too high or too low and lower than one. In the first area, since prices must be similar, you can quickly converge to 100 euros per dress and to a solution of 1, 2 and 3 dresses purchased by the first three girls. In the second area, the most likely solution is one and two dresses, with a price of 1000 euros per dress. There are other solutions, although much less likely.

The following second example is used to clarify how the choices of the additional information may influence the inversion. Here one wants to estimate the values of a function **f(x)** for a continuous set of **x** values within a given interval for **x** (between 0 and 1). The data are the measured values $\mathbf{d}^i$ for 10 values of $\mathbf{x}^i$ **(i=1,10)** and the associated measurement uncertainties (blue circles and bars in the figure). The problem is again strongly underdetermined but there is some additional information: first, one assumes that the solution can not depart strongly from a *prior* value $\mathbf{y}_0$ and that the dispersion around the *prior* is on the order of σ. Second, there is a regularization under the form of smoothing, i.e., for two values of $\mathbf{x}$, $\mathbf{x}_a$ **and** $\mathbf{x}_b$ separated by a distance $\mathbf{x}_b$-$\mathbf{x}_a$ on the order or smaller than a given smoothing length ξ, then $\mathbf{y}_a$= **f($\mathbf{x}_a$)** and $\mathbf{y}_b$=**f($\mathbf{x}_b$)** must be very similar. More specifically, there is a non-null covariance between such points and there is a correlation length ξ along the **x** axis that defines the smoothing scale. Following these two conditions, let's assume that the covariance function **C** has the following form:

$$C(x_a, x_b) = \sigma^2 \exp(-\frac{1}{2}\frac{(x_a - x_b)^2}{\xi^2}) \tag{8.8}$$

which means a variance σ and a correlation length ξ. For this one-dimensional, linear case with an extremely simple relationship $\mathbf{G}^i$**(x)**=δ**(x**^{i}**-x) (i=1,10)** (where δ is the delta function) and a covariance function **C** the solution for the inversion becomes:

$$y(x) = y_0(x) + \Sigma_i \Sigma_j C_{y_0 y_0}(x, x^i) \cdot (S^{-1})^{ij} \cdot (y_0^j - d_0(x^j)) \tag{8.9}$$

where

$$S^{ij} = C^{ij}_{d_0 d_0} + C_{y_0 y_0}(x^i, x^j) \tag{8.10}$$

Figure 8.1 (top panel) shows the computed solution for the Gaussian kernel of equation 8.8 with ξ=0.05 and σ=1, a constant prior y_0 equal to 2 and the 10 shown data points. For clarity σ and ξ are shown by vertical and horizontal segments. The standard error be computed in this case from the a posteriori covariance function (see Tarantola and Valette for the formulation). One can see that far from the data the standard error tends to reach the standard error on the prior value.

The role of the correlation length in the covariance function is illustrated in the middle and bottom panels of figure 8.1. For unrealistically large ξ values (0.3 in the figure) the smoothing is so spatially extended that the solution always departs from the data confidence intervals. On the contrary, in case of a too narrow correlation length (here 0.01), i.e. a length narrower than the mean interval between two consecutive data point abscissae x, the solution is unrealistically close to the exact data values and too rapidly forced to the prior value even at short distance from the data points (the inversion is over-fitted). Therefore, the choice of the kernel is closely associated with the distribution of the data points, and governs the spatial variation of the distribution. This is true here in one-D, but also in 3D, and as we will see in the inversion of

columns of interstellar matter the kernel ξ must be chosen according to the mean distance between two target stars, and will govern the minimum size of the recovered structures. One can easily deduce that, ideally, the size of the kernel should correspond approximately to the dimensions of the interstellar medium structures in 3D space. A wide kernel will permit only large structures and miss smaller ones, a too narrow kernel will produce unrealistic, small structures. As we will see, the choice of a double kernel will allow to better reconstruct both diffuse and compact structures, but it remains that a fractal structure is extremely difficult to reconstruct by inversion.

The role of the variance σ on the prior is illustrated in the middle and bottom panels of figure 8.2. For unrealistically large (respectively low) σ values (2.5 then 0.05 in the figure) the errors on the the solutions outside the data points are over- (respectively under-) estimated. One can see σ as a factor governing the global dynamical range, with small (respectively high) σ values associated with a low (respectively high) dynamical range around the prior. Finally, the choice of the prior level value itself (not illustrated here) is also important and obvious. Here a value of 2 is simply chosen close to the average of the data for the ten data points.

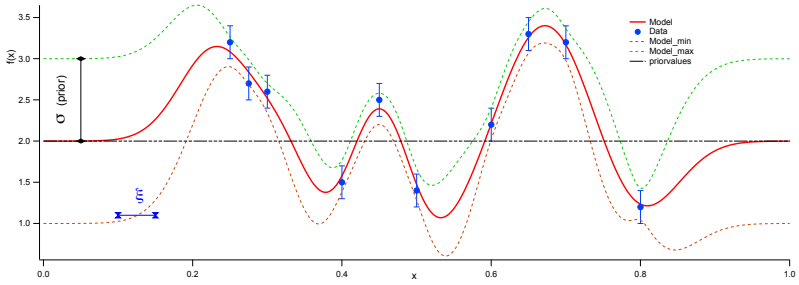
8.3 Interstellar medium tomographic inversion

In the case of the ISM full 3D tomography:

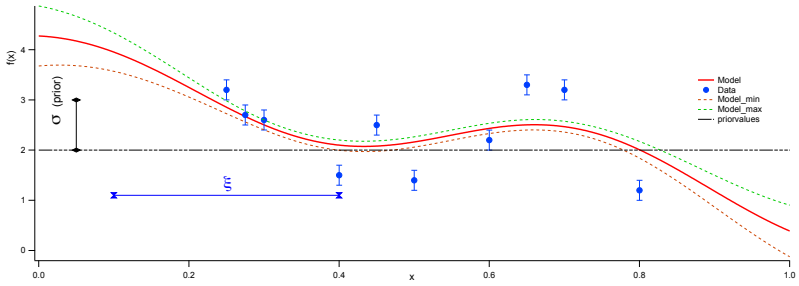
- Instead of a one-dimensional problem, data and parameters are in 3D and the parameters **p** one wants to derive are the values of the volume density of matter (grains, atoms or molecules) at series of points in 3D space.
- The data **d** are the integrated columns of interstellar matter, dust grains or molecules that are measured in the lines of sights toward as many as possible target stars. These columns are derived from extinction (for the dust grains) or absorption lines or bands (for gaseous species). Therefore, instead of the simple inverse function y^i of the illustrative example, where $f(x)$ must approach the measured value y^i for N points, here the theoretical relationship $\mathbf{d}^f = \mathbf{g} \cdot \mathbf{(p)}$ is the equality between these measurements **d** and the integrals within the 3D distribution of grains or molecules along the corresponding sightlines (the parameters **p** one wants to derive.
- Following previous work, a double kernel is preferred to the unique Gaussian kernel.
- Instead of a constant the *prior* is an analytically computed 3D distribution fro the parameters, or, as described in Chapter 10, a numerically computed distribution.

8.4 Inversion Algorithm

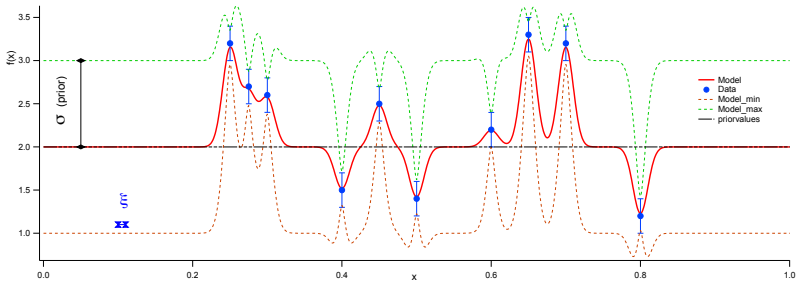
Except for cases presented as DIB inversion, see Chapter 16, the tomography described here uses measured values of excess colour in a series of lines of sight to reconstruct a three-dimensional distribution of dust opacity, i.e. a quantity proportional to the dust grain volume density.The transposition to gaseous



(a) Kernel: ξ= 0.05

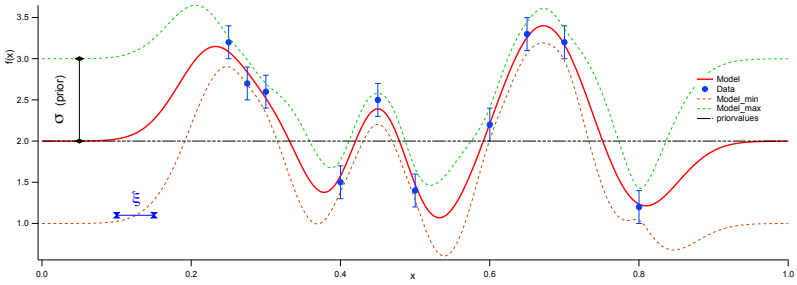


(b) Kernel: ξ= 0.30

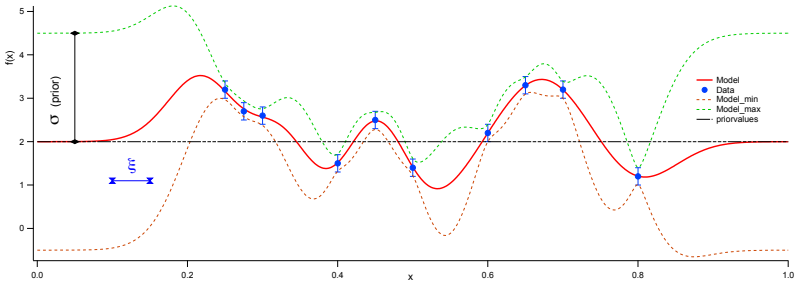


(c) Kernel: ξ= 0.01

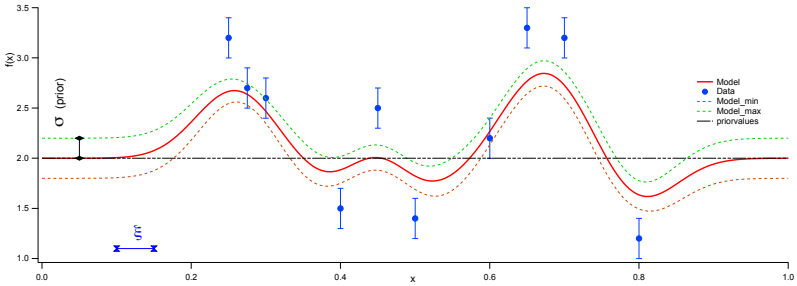
Figure 8.1 – Three exemple of one-dimensional inversion with same data and different kernel parameters.



(a) sigma: σ= 1.0



(b) sigma: σ= 2.5



(c) sigma: σ= 0.05

Figure 8.2 – Three exemple of one-dimensional inversion with same data and different dynamic parameters.

absorption line or band inversion is straightforward since it is equivalent to replacing dust grains by atoms or molecules. Note that the actual computed physical quantity at a point P is the reddening per unit distance subjected by the stellar light when it crosses P (e.g. a quantity expressed in mag pc^{-1}). This quantity is proportional to the number of nano-micrometric dust grains encountered by the photons per unit traveled distance, and, therefore, to the local volume density of these dust grains.

For each star i at distance R_i and Galactic coordinates l_i, b_i, its excess of colour can be expressed as the integral along the line of sight between the Sun and the star of the optical thickness per unit distance, expressed as the product of a constant and a quantity θ proportional to the volume density of grains and an absorption coefficient:

$$E_i = \int_0^R \frac{d\tau}{dr} dr \quad = \quad C \int_0^R \theta(r) dV(r) = C \int_0^R \theta(r) sin(b) cos(b) dl db dr \quad (8.11)$$

As in the 1D example above, the inversion is regularised by additional assumptions. The main one is that two points in space that are very close to each other can not have extremely different densities of matter (the smoothing assumption). The second is that the solution is as close to a prior distribution as permitted by the constraints from the data.

The inversion algorithm was developed from Jean-Luc Vergely then further developed and maintained thanks to the STILISM project. This software is a FORTRAN programme. The main programme input is the set of measurements along a series of lines of sight, i.e. their directions in the sky - in galactic coordinates (l, b), the distances to the targets (in parsec) and associated relative errors, the excesses of colour between B and V standard bands (in in magnitude) and absolute errors on the excess of colour.

Algorithm is fed by different parameters:

Prior distribution : parameters in the case of an analytical prior distribution The initial inversion code uses as a prior a quasi-homogeneous distribution that depends only on the distance to the galactic plane, i.e. it is independent of the location of the projection onto the plane. The prior interstellar medium density is assumed to decrease exponentially as a function of the height above/below the plane z with a height scale $\mathcal{H}$.

$$\rho = \rho_0 \exp\left(\frac{-(z - z_0)}{\mathcal{H}}\right) \quad (8.12)$$

In practice the input $\mathcal{H}$ parameter is deliberately over-estimated in comparison with the measured height scale. There are operational reasons for that, as explained in Vergely et al. (2010a). An higher height scale allows a better inversion of the small structures at hight latitude/distance from the plane and better results were obtained when charting the Local Bubble, which has a very low opacity. As a result, the two parameters are:

- ρ_0, density on the Galactic plane
- $\mathcal{H}$, the prior height scale

Prior distribution: case of a non-homogeneous prior distribution defined everywhere in the 3D space For the first time during my thesis I ran new inversions using a non-analytical prior distribution. The main goal is to allow the mapping of regions with low numbers of targets, and the basic principle is to build a preliminary map at very low spatial resolution that covers all the space (including target-poor regions and that is subsequently used as a prior for subsequent inversions at better resolution and restricted to regions that possess more targets. This will be detailed in the next sections and Parts III and IV.

Two types of such prior distributions were used:

- prior based on completely different data, i.e. not the data to invert
- prior based on an inversion made in a previous step, during an iterative process

Parameters controlling the regularization: spatial smoothing and dynamics: The covariance function C controls the relationship between the densities in two neighbouring points, therefore the size of the structures through the kernel parameter ξ, and the allowed dispersion around the prior for all data points, therefore the dynamical range of the structures, through the parameter σ. It may take various forms.

Two different autocorrelation function f(ξ) are accepted in the code:

$$C_1(x) = \frac{{\sigma_{\xi_0}}^2}{\cosh(-\frac{x}{\xi_0})} + {\sigma_{\xi_1}}^2 \exp^{\frac{-x^2}{\xi_1}} \tag{8.13}$$

and:

$$C_2(x) = {\sigma_{\xi_0}}^2 \exp\left(\frac{-|x|}{\xi_0}\right) + {\sigma_{\xi_1}}^2 \exp\left(\frac{-|x|}{\xi_1}\right) \tag{8.14}$$

The choice of an appropriate kernel is very difficult, especially in 3D, since it is related to the interplay between the distribution of the targets and the characteristic sizes of the structures (the dust clouds). It is much more difficult for an irregular dataset, which is the case here: the distribution of the targets in space is far from homogeneous, instead the number of targets per unit volume decreases with distance from the Sun (observational selection for fainter sources), and with distance from the galactic plane (the stellar distribution is concentrated close to the plane). On the other hand, the sizes of the dust clouds vary from fractions of parsecs to tens, even hundreds of parsecs.

For all the presented inversions I used always the C_1 function. It has the advantage of allowing the representation of relatively compact and small clouds (through the Gaussian term) and more extended and much less contrasted clouds (through the $\cosh^{-1}$ term. The first type (resp. the second) corresponds to dense (resp more diffuse) entities.

Whatever the quality of the kernel, the decrease of target density at increasing distances from the Sun implies that a unique kernel can not describe correctly nearby and distant regions. This is certainly one of the most difficult problem from the point of view of the algorithm. A part of this thesis work has been devoted to the search for a solution to this problem. This is discussed in Chapter 11.

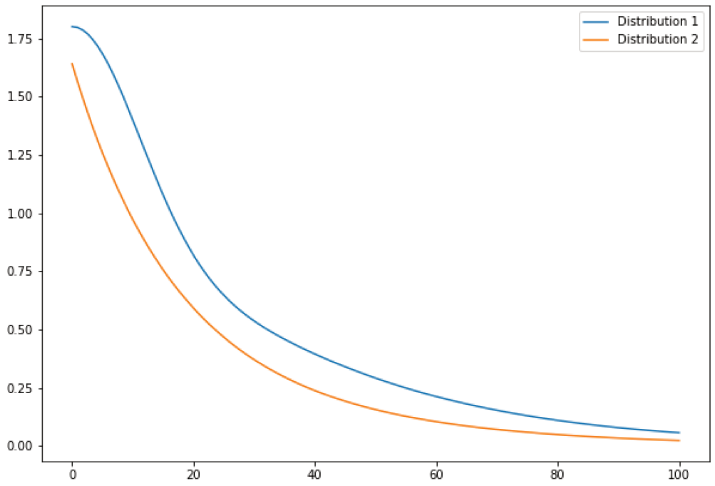


Figure 8.3 – C_1 and C_2 for the parameters $\sigma_{\xi_0} = 0.8$, $\xi_0 = 30$, $\sigma_{\xi_1} = 1.0$,$\xi_1 = 15$.

Parameters defining the volume where the dust density will be computed and the selection of input sightlines for the inversion One can select the computation area through:

- maximum and minimum Galactic longitude
- maximum and minimum Galactic latitude
- maximum distance along the Sun-Galactic centre axis X, imposed to be equal to the maximum distance along the perpendicular axis Y in the Plane
- size of the voxel for the final distribution

This defines a parallelepiped-shape volume for the 3D map. It can be a large volume containing the Sun or any other volume. The size of the voxel is chosen about three to five times smaller than the map resolution (defined hereafter). The output distribution is the series of excess of colour per pc in discrete voxels in cartesian space.

Similarly, the set of targets used for the inversion is defined by the same list of parameters but not the same values. Indeed, in order to avoid artefacts at the boundary of the computational volume, the range for longitude, latitude, distance from the Sun and to the galactic plane of the input targets must be chosen wider than those of the computational volume. Note that targets that are outside the final computational volume influence the computation inside the volume for geometrical reasons (their sighline crosses the volume).

Input data and data quality criteria The data are the list of individual lines of sight and their measured distances and colour excesses:

- target star longitude and latitude
- target star distance and associated relative error on this distance
- target star color excess and associated uncertainty

Stars with very uncertain distance or very uncertain colour excess do not bring useful constraints and may increase the computation time in a useless manner. A selection is made on the distance relative error.

Moreover, in order to be conservative in case the smaller error considered for the excess of colour is of 0.01 mag. For input points with associated excess of colour error lower than 0.01, this the value used in the inversion.

Parameters controlling the computation and treatment of outliers The central part of the code is the computation (by blocks) of a very big matrix representing the mutual influences of all points in space that are constrained by the data. To do so, all sightlines with a known colour excess are discretised into distance bins that will enter the matrix. One parameter, called resolution, defines the size of the distance bins. It is chosen to be close to the integration step for the sightline reddening in Eq 8.11. For points in space that are very distant one can obviously neglect their mutual influence. For this purpose, two parameters control the thresholds for the distance between two points beyond which their mutual influence is not computed. One parameter applies to the computation of the matrix linking the observed lines of sight and the second to the computation of the final density in each point in space.

Finally, a special treatment of the outliers has been implemented in the latest version of the program. Starting with the third iteration, errors bars on the colour excess for lines of sights that are highly discrepant from the solution are multiplied by a factor of three, and this is done foe each subsequent iteration. Tests made by J.L. Vergely have shown that this eliminates data points with unrealistic measurements.

The convergence criterion is based on the relative decrease of the χ^2. When it becomes lower than 0.01 the iterations are stopped. This happens in general for between 5 and 10 iterations. The maximum number of iteration is modifiable, but in this thesis I always use 10.

Parameter choices

As we have more than twenty parameters, it is not trivial to tune them in a proper way. In general, for the prior parameters I used the same than in Lallement et al. (2014), i.e. $\rho = 0.00025$ mag/pc and height scale = 200 pc. The final distribution was calculated with voxels of 5pc^3, in a global MW region of $4000\times4000\times600$ pc^3, otherwise it is specified. The other parameters will be specified for each inversion.

8.5 Specific difficulties of full 3D inversions

3D mapping is complex, and each technique has its own advantages and shortcomings. The full 3D inversion has several critical points. They are intrinsic to the technique and the used algorithm. In the following chapters I will present

several improvements that I have made using different strategies with the aim of limiting the effects of two main problems discussed below.

Non-uniformity of target distribution in 3D space

Ideal inversions are those based on a homogeneous spatial distribution of sight-line constaints, which would here imply a uniform volume density of target stars, with similar errors in distance and reddening. Since we invert distance-limited quantities, we already said that it is (by far) not the case, due to dimming of starlight with distance and rarefaction of targets out of the plane. In addition, every survey has its observational selection biases. For all these reasons, it is obvious that the target distribution is non homogeneous. As explained above the targets' spatial distribution in the inversion is fundamental regarding the choice of an appropriate kernel. A reasonable choice of kernels takes in account the Nyquist-Shannon criterion, and it affects the maps reliability: in not well-constrained areas (usually far from the Sun) the inversion tends to "invent" structures that are based on too few data and not reliable. Our available inversion code don't allows us to use different kernels at different distances, which implies that a kernel appropriate in the inner regions will no longer be appropriate at large distances and vice versa.

Computational time

In the new era of big surveys, the main practical problem of full 3D tomographic inversions becomes the computational time. This is illustrated by the following test. I made use of the CPU time of different inversions based on increasing number of targets, and I fitted how this number influences the time. Our fit shows that the general trend is :

$$CPU \quad Time[s] = a(target \quad number)^2 + b(target \quad number) + c \qquad (8.15)$$

with a, b, c = [6e-03, 5e-17, 1e-20].

The CPU time is related not only with the number of targets, but also with the integration step and the pixel resolution for the final distribution. The pixel resolution for the final 3D distribution influences only the very last step of the inversion after the convergence is reached, and is not so important. The integration steps influences every iteration, but for a given target it does not increase if the target distance does not increase. Finally, the number of targets is the main factor governing the CPU. One can see from Fig. 8.4 that it increases roughly as the square of the target number.

One of the main goals of my thesis has been to test a new method allowing to avoid the unsurmountable increase of computational time linked to the existing code. This need for a new technique is especially motivated by the availability of massive Gaia data. With the current techniques, even 10^5 objects are feasible.

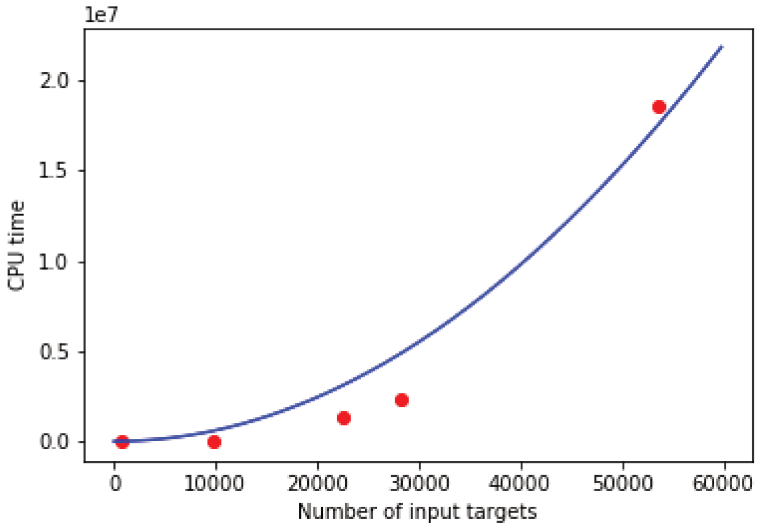


Figure 8.4 – CPU time for inversions with different input targets number. Time is measured in seconds.

Chapter 9

Mapping improvements: using APOGEE DIBs as additional data

The improvements and tests on full 3D mapping in this thesis work have as start point the map performed in 2014 (Lallement et al., 2014). This map is very accurate in the solar vicinity, but it is much less precise beyond a few hundreds parsecs (say, $\sim$ 400 pc). This map was based on $\sim$ 24000 lines of sight, whose distances were determined for $\sim$ 20% from the Hipparcos parallaxes and the rest based on photometry. Stellar targets for this map were concentrated in the first 200 pc from the Sun, because of the difficulty to have good estimations beyond the first matter wall (targets are too faint and thus absent from catalogues).

With the goal of extending and improving the map at longer distances, I started using for the mapping the absorptions derived from APOGEE DIBs. APOGEE DIB measurements have a strong advantage: being measured in the H-band, where extinction is less important than in the optical, they allow to probe denser, more opaque clouds with more distant targets.

9.1 Description of APOGEE DIB data

We used results of APOGEE DR12 (Alam et al., 2015, DR12); in addition to the $\sim$ 160000 science targets, APOGEE also recorded a sample of $\sim$ 17000 spectra of hot calibration stars, used for telluric line elimination (called Telluric Standard Stars, TSSs). These are bright sources with H-band magnitudes ranging from 5.5 to 11(Zasowski et al., 2013). For such hot stars the DIB extraction could be much simpler and precise than for cool stars, due to the absence of stellar features.

APOGEE data are very high quality stellar spectra, from which precise atmospheric stellar parameters were obtained; consequently target distances and estimated colour excess have become very accurate, because they are based on both spectroscopy and photometry as well as on carefully adjusted theoretical stellar spectra, especially since the latest Data Release and the use of the Cannon software. However, this is true only for the cooler scientific targets and not for the calibration stars (the TSSs). For the TSSs, stellar theoretical spectra are not appropriate and there are no distance-extinction estimates based on the spectro-photometry. This is why we had to make use of parallax

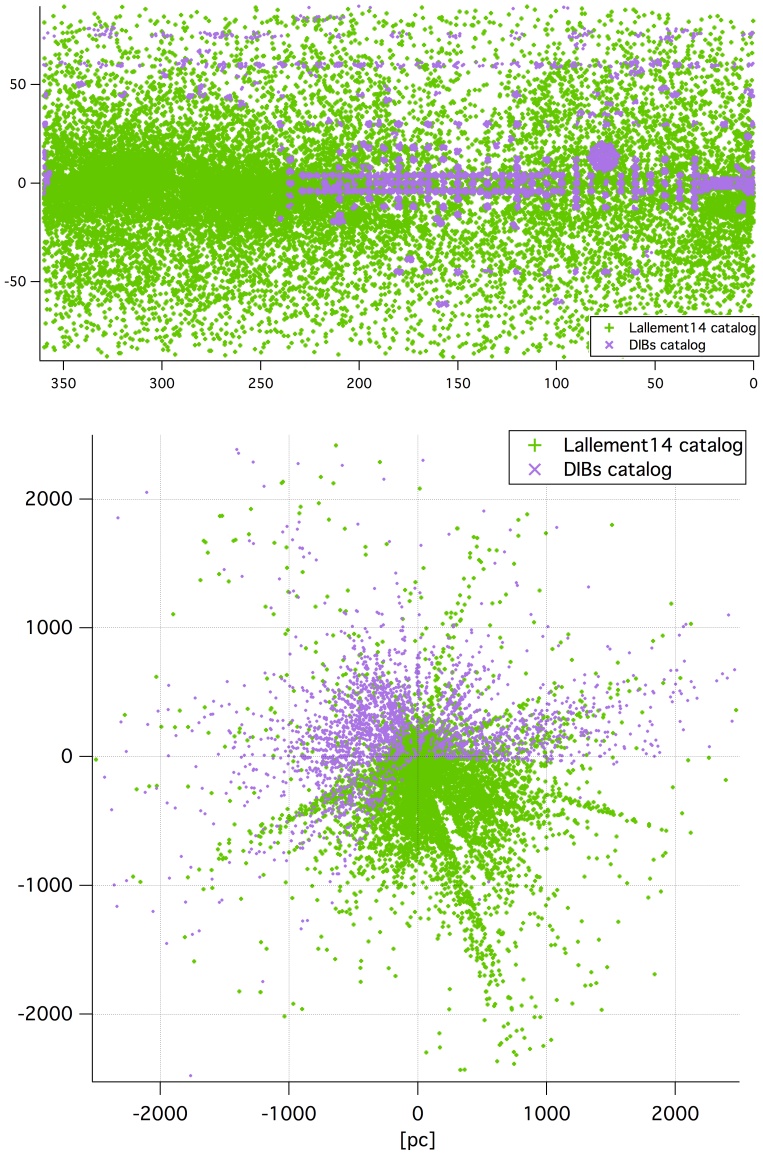


Figure 9.1 – Top) Target distributions for the Lallement14 catalog (green crosses) and the added DIBs absorptions (violet crosses) on the sky. Bottom) Target distributions for the Lallement14 catalog (green crosses) and the added DIBs absorptions (violet crosses) on the galactic plane, only targets with distances to the plane in the range (-50,+50) pc are represented.

Flags	λ_c Å	Width Å	Depth Å
Detected (4861)	[15 268 − 15 280]	[1.4 − 5.5]	[1.5× max(RA,RB,RC) −0.1]
Narrow DIBs (364)	[15 260 − 15 280]	[1 − 1.4]	[1.5× max(RA,RB,RC) −0.1]
Recovered DIBs (2515)	[15 268 − 15 280]	[1.4 − 5.5]	[2× RA −1.5× max(RA,RB,RC)]
Narrow recovered DIBs (420)	[15 268 − 15 280]	[1 − 1.4[	]2× RA −1.5× max(RA,RB,RC)]
Upper limit (1753)	[15 268 − 15 280]	≤1.4	≤ 2×RA

Table 9.1 – Fitting constraints of λ15273 DIB for the selected APOGEE spectra. RA, RB, RC are the standard deviation in three spectral ranges A = [λ_c - 10Å, λ_c+ 10Å], B= [15890-15960] Å and C= [15200-15250] Å, respectively.

distances and retain targets possessing a parallaxes. At the time of this work, only Gaia DR1 data were available, i.e., only targets belonging to the Tycho catalog could have a Gaia parallax. The Gaia DR1 was published with data collected in the first year of observation, and to assign a parallax estimation they used the previous Tycho measures.

I explain briefly the spectra fit and selection of targets, but for details, I refer at the accompanying paper Capitanio et al. (2017).

DIBs profile fitting

The series of the 25 409 APOGEE DR12 spectra of targets belonging to the Tycho catalog were analysed to determine the presence or absence of a detectable DIB at 1527.3 nm and measure its equivalent width. The methodology and the technique applied to the TSSs spectra is described in Elyajouri et al. (2016).

Among the 25 000 star spectra analysed, only the best spectra were selected, excluding target stars for which the stellar models were not suitable, for which S/N was too low or if there were too many telluric line residuals, following the criteria exposed in Tab 9.1; this filtering selected 9913 spectra.

9.2 Inclusion of DIB data addiction to colour excess data

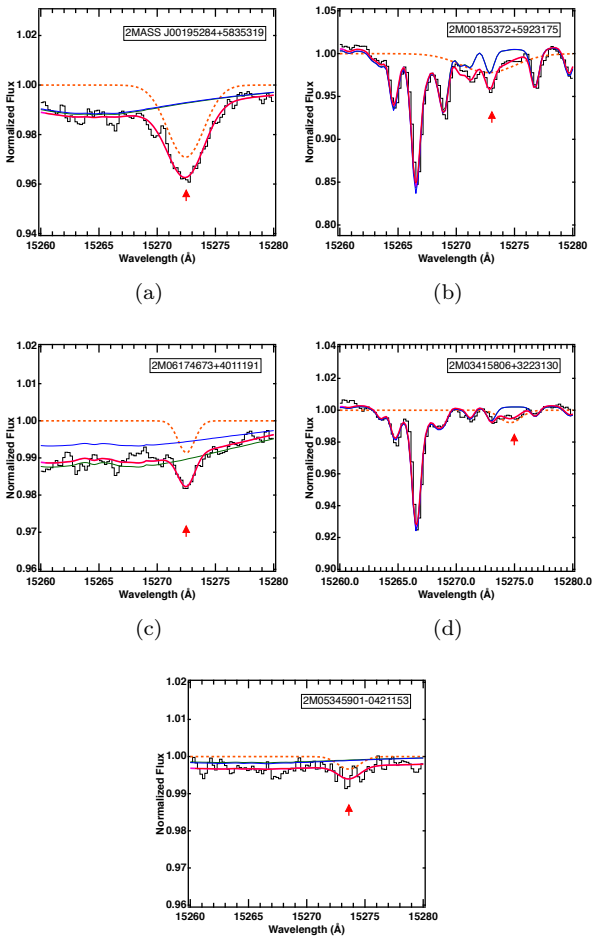
After crossmatch with the targets of the Gaia DR1 catalogue (Arenou et al., 2017b) I could obtain a catalogue of 7800 stars with Gaia distances and well measured DIBs. I subsequently filtered to eliminate stellar targets with fractal error on distances larger than 33%: 4886 targets were finally used for the inversion.

DIB equivalent widths were converted into extinction then into excess of colour using the relationship established by Zasowski et al. (2015):

$$EW = 0.102\ A_V^{1.01} \text{Å} \tag{9.1}$$

and the ratio $R_V = A_V/E(B-V) = 3.1$ (Savage and Mathis, 1979).

DIB absorptions are not produced by dust grains, as does the extinction, but the carriers are probably large molecules in the gas phase. Therefore, they trace IS matter but not exactly the same kind and they are not exactly proportional to the extinction. To take into account the dispersion around the average relationship, I assigned to the excess of colour derived from DIBs an error of 50% that comes in addition to the uncertainty associated with the fit.



(a) (b) (c) (d) (e)

Figure 9.2 – Illustration of the various categories of DIB extraction: a) Detected; b) Recovered; c) Narrow; d) Recovered and Narrow; e) Upper limit. APOGEE spectra are shown with solid black curves. The initial stellar model provided by APOGEE and the model obtained after application of the scaling factor are shown in solid blue and green respectively. The DIB absorption alone is represented in orange. The solid magenta curves represent the final stellar+DIB modelled spectrum.

9.3 New inversions

Using Gaia DR1 parallaxes

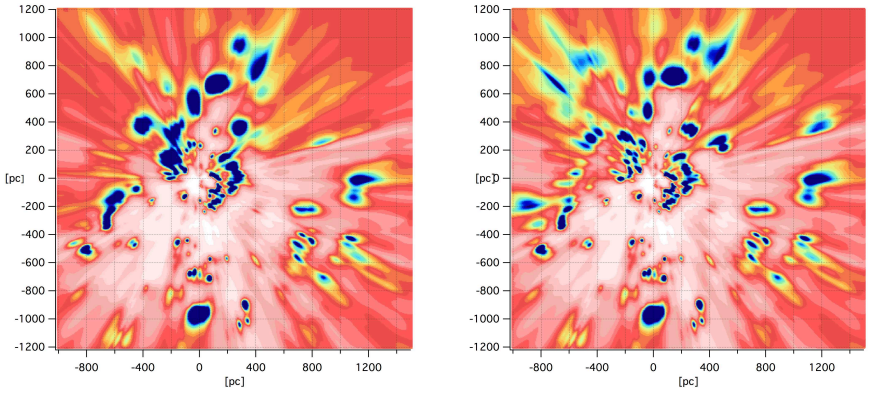
Before merging DIBs and colour excess data, I ran a preliminary inversion without DIBs, but using the newly published Gaia DR1 (TGAS) parallaxes. Starting with the Lallement catalog of 23444 targets available for inversion, I searched for targets possessing a a Gaia DR1 parallax (Arenou et al., 2017a) and found $\sim 80\%$ of them in this case; I replaced the photometric or Hipparcos parallaxes-derived distances by these Gaia DR1 distances, when available. Because the Gaia parallaxes have smaller errors compared to the Hipparcos or photometry distances, and because for our inversion only stars with distance errors smaller than 33% are finally retained, the new catalog available for an inversion is slightly different from the one used by Lallement et al. (2014). The result of the new inversion with the same technique as in Lallement et al. (2014) is illustrated in Fig 9.3. The figure shows the distribution of dust in the Galactic plane. We note that the use of Gaia DR1 distances introduces several modifications, as expected, and they are explained in the paper Capitanio et al. (2017); the main change on the galactic plane is seen in the first quadrant at longitudes on the order of 70-80° where the absorbing matter is now found at larger distances, beyond 600 pc. We knew that in this region the inversion was potentially incorrect due to strong oscillations during the iterations. The new distance is in better agreement with the radial profiles found by Green et al. (2015).

Merging color excess and DIB measurements

The second inversion is an attempt, the first of its kind, to run inversions based on composite data. The DIB catalog converted into colour excess described above corresponds to significantly different characteristics of the lines of sight. For the colour excess catalog the stellar median distance was 219 pc, while for APOGEE stars the average distance is 724 pc. This means that we augmented the previous catalog with 20% of new targets with much larger distances.

As shown in Fig. 9.3, after the combination of photometric and DIB-based excess of colour values, the new inverted map reveals in the Galactic Plane several important changes in the distant regions in the anti-center directions. This is due to the fact that 14% of the APOGEE targets are closer than 20 pc from the galactic plane and to their large distances, as said above. The absence of changes in the fourth quadrant is simply explained by the absence of new APOGEE stellar targets in this area (the APOGEE observing site is in the Northern hemisphere). Local Bubble has not changed much, because there are a lot of old data in the first 200 pc, and the added data are consistent with the previous ones: for this reason the inversion leads to stable results in the nearest regions.

More details can be found in Capitanio et al. (2017) (published article inserted in this document).



(a) (b) C

Figure 9.3 – Left) Galactic plane for the 3D map based on the Lallement et al. (2014) catalog, with Gaia DR1 distances. Right) Galactic plane for the 3D map based on the Lallement et al. (2014) catalog with new Gaia DR1 distances and the addition of the selection of APOGEE DIB-based extinctions.

.

Chapter 10

Using the Pan-STARRS - based extinction radial profiles for 3D dust map and a prior 3D distribution

In this chapter we describe two different ways of using the highly detailed extinction mapping recently performed by Green et al. (2015). Both exercises correspond to attempts of improving our 3D dust maps. The first attempt is the extraction of several extinction-distance pairs from the Green et al. (2015) extinction radial profiles, in a large number of directions. After combination with the Galactic coordinates of the direction, these pairs are considered as *virtual* targets to be added to our database and treated in the same way. The second attempt is the construction of a 3D prior distribution based on actual measurements, here Pan-STARRS, at variance with the previously used, analytical prior. This second prior will also be based on the Pan-STARRS reddening radial profiles.

10.1 Introducing to Pan-STARRS based map of reddening radial profiles

The Green et al. (2015) map is based on Pan-STARRS and 2MASS photometry, and the derived classifications, stellar distances and reddenings for $\sim$ 800 millions of Pan-STARRS targets; Green et al. (2014) developed a statistical method to derive from this dataset the extinction radial profile along all lines of sight accessible with the survey at high angular resolution. To do so, they work solid angle by solid angle, i.e. for a given central direction they select all the targets that are contained within a small solid angle around this direction. Based on their photometric distance and reddening they construct the best fit, monotonously increasing reddening profile along the radial direction. The resulting map of radial profiles is available at the website http://argonaut.skymaps.info/. It is possible to download the complete map or to access at the information choosing the line of sight. There are some fully-prepared query forms in Python to download large amounts of data. The map has a variable but always very high angular resolution, ranging from 3.40 to

13.70 arcmin, however the bins of distance moduli (i.e. a logarithmic scale for the distance) correspond to a distance-dependent distance resolution of 25%, i.e. very broad radial distance bins, especially at large distance.

Because of their statistical technique, there are some well explained limitations in excess of colour reliability at different lines of sight. In pixels corresponding to a very low number of objects - as in the case of shorter distances - their statistics fail because they do not have enough information. For the more distant pixels the problem is still the decrease in the number of objects that leads to insufficient information, but also the observational distortion, because the more distant stars observed are usually close to the limit of magnitude of the investigation and are favoured targets weakly reddened. They use a very refined technique to determine where their determination of the reddening is reliable, and the distance without a good reliability are well indicated for each line of sight.

The Pan-STARRS radial profiles, as said above, have low accuracy and reliability at small distances, and the reliability area starts at distances on the order of 300 pc. This fact entails that there is a complementarity between our 3D maps that are the most precise at small distances and the Pan-STARRS profiles and this complementarity has motivated the attempts described in the next sections.

10.2 Pan-STARRS extinction radial profiles as sources of additional color excess data

A 3D inversion is alimented by a large series of distance - excess of colour pairs, and such pairs usually come from individual stars for which distances and reddenings are simultaneously estimated. The Pan-STARRS extinction radial profiles are also initially based on a large number of targets used in conjunction, but for each direction the resulting profile is a statistical result. Still, it is a series of distance modulus-extinction pairs, and our idea has been to consider each point of the profile, i.e. each one of these pairs as a potential *virtual* target star having the same distance and the same excess of colour value. As mentioned above, it provides a possibility to extend and improve our mapping because it has information at further distances and because the spatial distribution of the additional targets is uniform.

Green et al. (2015) provides the extinction, with its standard deviation, for all distance moduli between 4 and 19 with a distance modulus (DM) step of 0.5. after transformation of DM in parsecs using the relationship:

$$d[pc] = 10^{(0.2 \cdot DM + 1)} \tag{10.1}$$

we can realise the non-uniformity in distances and the bin widths, with 5 bins over 600 pc between 400 and 1000 pc (DM 8-10) and with 5 bins over 1500 pc between 1000 pc and 2512 pc (DM 10-12). The set of reliability flags shows that the most reliable data start at about 400 pc. Since we need additional targets over 500 pc, in this sense, the Pan-STARRS may enormously help.

Data selection:

- For each line of sight several extinction curves are given, based on different star subsamples. I did not use the samples but only the profile marked "Best Fit", with its corresponding error.
- Lines of sight were selected with one degree steps in Galactic longitude and latitude.
- We retained only the parts of the profiles in the reliability zone, using the given flag.
- To be conservative in our choice, we selected only data with relative error on the excess of colour less than 0.3. This criterion strongly reduces the imported data.
- For our 3D inversion, we only use data with distances to the galactic plane lower than 600 pc. We retained Pan-STARRS data that fulfil this criterion.
- For each line of sight I have selected a maximum of three points chosen to be in the flat areas of the reddening profile, because they are the regions in which the different samples are in good agreement, and as consequence also the "Best Fit" curve is more reliable.
- My selection has only excess of colour that grows monotonically.
- The most distant point is taken into account only if the difference between the second and third point is less than 20% because for the inversion code, a too important gradient is a source of errors and elongated structures that we want to avoid.

After this selection, I retained only virtual stellar targets closer than 2500 pc.

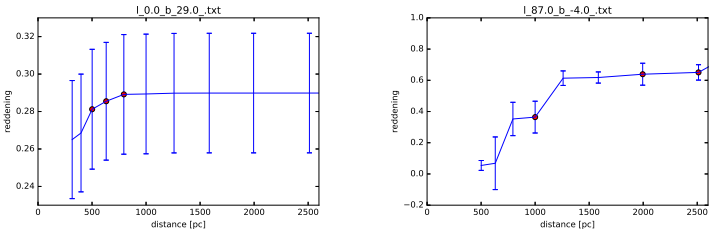
These selection criteria could seem too strict, but anyway the inversion cannot manage more than $\sim$ 70000 points, and it takes a lot of computational time. The strategy was to select only the more useful virtual stars.

In Fig 10.1 are shown some examples for different lines of sight, the extinction curve from Green et al. (2015) and the selected points. In Fig 10.2 is shown the sky distribution of the virtual stars. A fraction of the sky is not covered by the Pan-STARRS survey (norther hemisphere data), unlike the initial catalog.

Inversion based on real and *virtual* stars

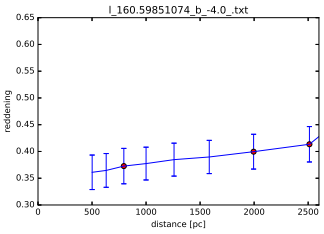
A new inversion has been attempted for a total number of 69 707 distance - excess of colour couples, with only 22 900 points from the initial catalogue, the rest being Pan-STARRS *virtual* stars. The spatial distribution of all the targets is shown in Fig 10.3 with different colours for the Lallement et al. (2014) and my Pan-STARRS selection. Most of the improvement is in the sampling of distances and increase of distance range. An arbitrary error of 5% was assigned to the distance of the *virtual* Pan-STARRS targets. As prescribed by Green et al. (2015), we corrected the reddening taken from the Pan-STARRS website by a factor 0.88.

The resulted map is damaged by an artefact in the anticenter direction (see Fig. 10.4), and we still have not completely understood the origin. Given the

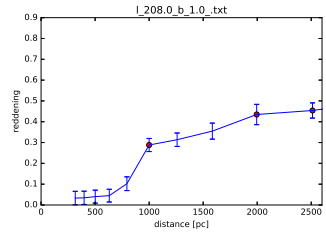


(a) Linesight l = 0, b = 29.

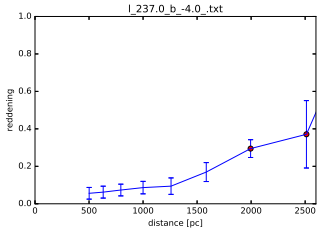
(b) Linesight l = 87 b = -4.



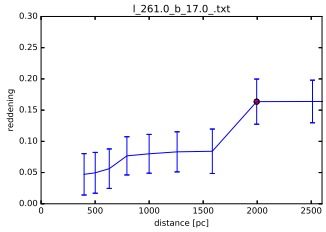
(c) Linesight l = 160.59 b = -4.



(d) Linesight l = 208 b = 1.



(e) Linesight l = 237 b = -4.



(f) Linesight l = 261 b = 17.

Figure 10.1 – Example of lines of sight and targets. For each line of sight the reliable excess of colour values as in Green et al. (2015) and red points are the selected virtual stellar targets.

distribution of the spurious voxels along a sphere centered on a point in space without any particular characteristics, very likely one of the elements of the matrix of mutual influences has a false value, maybe due to a problem during the transmission of one of the files during its computation. Since the artefact affects a limited region only, and the computational time for such a direct inversion is very large (1 month and 26 days), we did not perform a second run.

Fig 10.4 shows the Galactic plane distribution from the new inversion and how it compares with the inverted distribution without the *virtual* stars. There are some new structures in the new map, especially at large distances, as expected, for example a second rank of massive clouds is mapped. On the other hand, some of the structures in the third quadrant are more compact, a sign of improved realism.

When comparing some vertical cuts through the distribution, - Fig 10.5 - we can see that in the direction l = 80° the new Pan-STARRS points add a lot

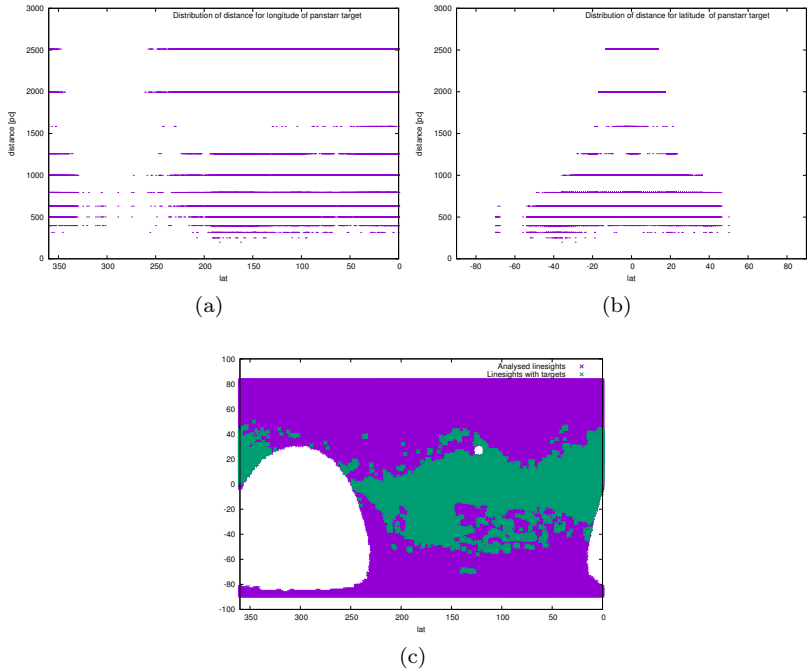


Figure 10.2 – Distribution on the sky of Pan-STARRS selected targets. a) x axis: galactic longitude, y axis: distance [pc]. b) x axis: galactic latitude, y axis: distance [pc]. c) Comparison on the sky between the total Green et al. (2015) distribution (violet points) and the selected points (green points).

of information at large distance, while towards l = 160° the most important changes are in the nearest 800 pc and then the exponential prior dominates.

Again, we emphasise that this inversion has a much lower angular resolution than the original Green et al. (2015) map, due to our selection of directions and because tour 3D inversion can not reproduce very small structures. Also, the sizes of the distant structures are large, since the initial distance bins are wide and have increasing spacing. However, we derive from these comparison that this is a potential technique of combining different datasets.

10.3 Prior based on Pan-STARRS extinction map

Another possible use of the Pan-STARRS extinction map is the derivation of a large scale 3D prior distribution for our 3D inversions. The prior distribution is very important in inversion processes as it controls how the medium should be described where there is no information from the data, and it is especially fundamental in 3D inversions that are generally intrinsically under-constrained. For those reasons, it is important to choose the most realistic one. Moreover, a realistic prior facilitates the convergence towards the best-fit solution. I prepared and tested the evolution from an analytically computed prior, namely an exponentially decreasing function of the distance from the galactic plane (Chen et al., 1998), to a more realistic prior. For our 3D mapping, we need a

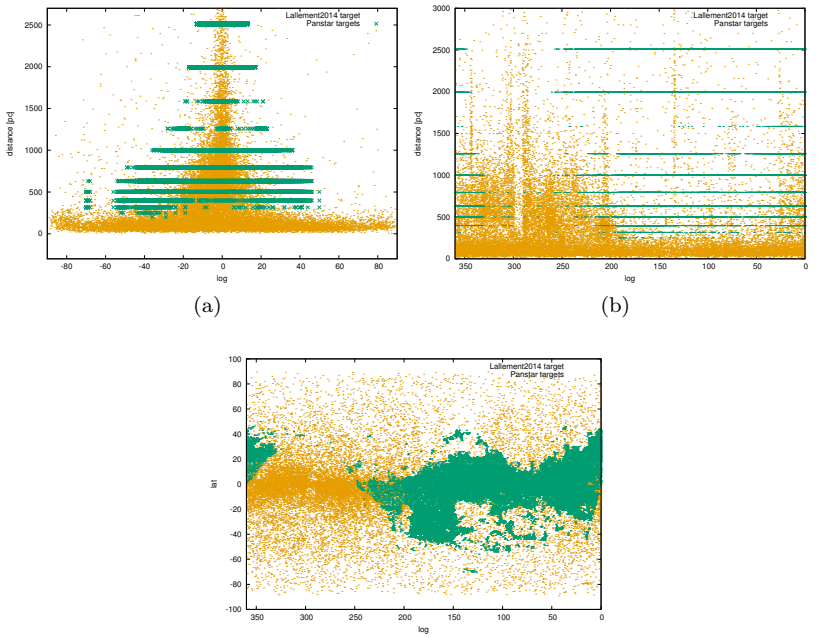


Figure 10.3 – General comparison of spatial distribution between Lallement14 catalog (orange points) and new selected points (green points). a) x axis: galactic longitude, y axis: distance [pc]. b) x axis: galactic latitude, y axis: distance [pc]. c) Comparison on the sky.

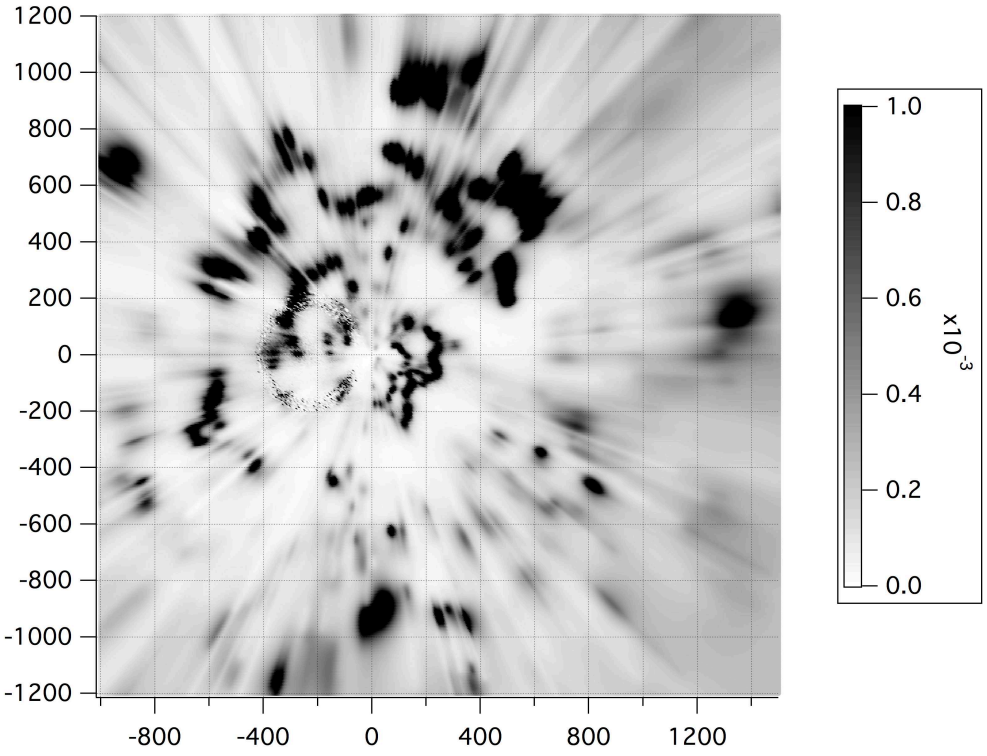
3D prior, therefore the very detailed 2D maps, e.g. from Schlegel et al. (1998) and Planck Collaboration et al. (2014) are not useful for our purpose.

Using the Pan-STARRS map as a prior does not mean that we make use its high angular resolution, because the inversion can not reproduce structures with a very small size (due to the use of the kernels), but we can better recover the large structures. Moreover, since the Pan-STARRS radial profile distance bins are increasing rapidly with distance, the prior is expected to generate structures at different scales, from the smaller and better defined near the Sun to the wider and largely smoothed for more distant areas.

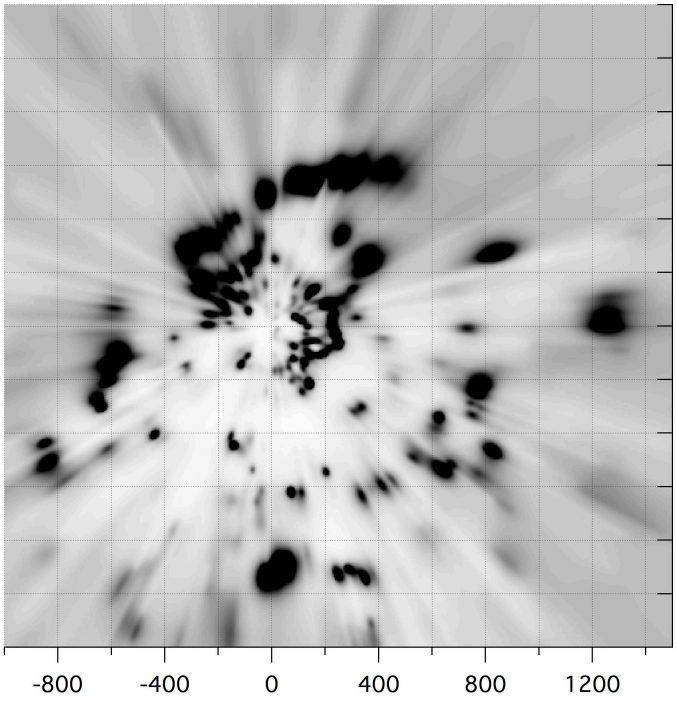
10.4 Building the prior

To build the prior we used all the points for the extinction curves with reliability flag, sampling the sky each 0.5° in longitude and latitude. I obtained a total sampling of 3 071 230 distance-excess of colour value pairs. As done previously and recommended by Green et al. (2015), I corrected the reddening as they come from the Pan-STARRS website by a factor 0.88.

We used for the first time a hierarchical method to build the new prior. This new technique, devised by J.L. Vergely, allows the users to reconcile reasonable computation time and large number of targets as well as a distance-dependent structure size (from very large structures at large distance to increasingly small

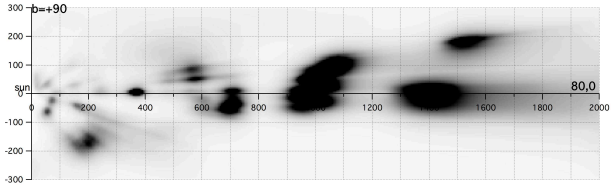


(a) *Galactic plane for the inversion Lallement14 + our selection.*

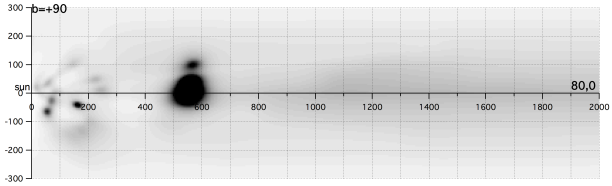


(b) *Galactic plane for the inversion Lallement14.*

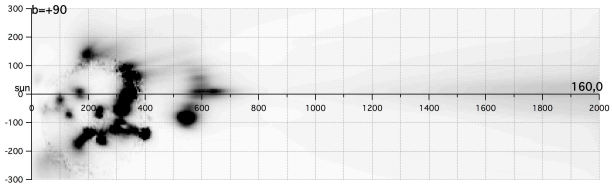
Figure 10.4 – a) Galactic plane for my inversion based on the Lallement14 catalog and my selection from Green15 map, Sun in (0,0) point, x axis [-1000,1500] pc, y axis [-1200,1200] b) Galactic plane for inversion based on the Lallement14 catalog, Sun in (0,0) point, x axis [-1000,1500] pc, y axis [-1200,1200] b). Color scale in magnitude per parsec.



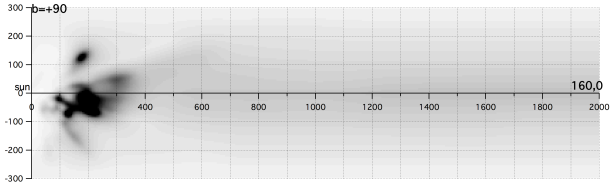
(a) *Cut l = 80° for the inversion Lallement14 + our selection.*



(b) *Cut l = 80° for the inversion Lallement14.*



(c) *Cut l= 80° for the inversion Lallement14 + our selection.*



(d) *Cut l= 160° for the inversion Lallement14.*

Figure 10.5 – Vertical cut of different planes in the 3D map. Comparison between old and new map, based on only Lallement14 catalog and the combined catalog. Colour scale as in Fig 10.4.

structures at decreasing distance). It starts with a very low frequency map (structures of 800 pc and a 5 kpc volume), using a dataset of values obtained by averaging data in cells of a size that allows to fulfil the Nyquist-Shannon criterion. This technique will be largely explained in Part IV.

1. The Pan-STARRS reddening data subsample is converted into local opacities by computing the local gradients of the reddening profile ρ_0: $\rho_0 = \delta_{E(B-V)}/\delta_R$.

2. A spatial scale S of restitution of the opacity is fixed, starting with the (large) scale that can possibly be restored on the whole map. The average opacity is computed in a 3D grid of boxes of size SxSxS by averaging local values of locations contained within each grid cell. Thus, the resulting average opacity is dependent on the scale.

3. Based on these average opacities, maps of smoothed opacities are computed by Bayesian interpolation with a smoothing kernel twice as large as S.

4. We then move on to a scale twice smaller. S/2 x S/2 x S/2 boxes are built and the average opacity is calculated for each box. The dynamics of the average opacity for this set of boxes is stronger. We suppress the boxes that do not fulfil the Nyquist-Shannon criterion for S/2, and usually they are the more distant boxes.

5. A Bayesian interpolation is carried out again with a smaller smoothing kernel and using the previous map as prior.

6. Again, the scale is divided by two and steps 2) and 3) are performed again for this new scale, etc.

New prior

Final Pan-STARRS prior map is a 3D map, with dimension [-3000:3000]x[-3000:3000]x[-600:600] pc, calculated with voxels of 20 pc. With respect to the Fig 10.4, the solution has no longer radial features (the *fingers of God* that are due to uncertainties on the radial position of targets, because the hierarchical technique locates structures only if they are really constrained by data. Also in this low resolution maps some features, as the Local Bubble, are recognisable.

10.5 Mapping results

I performed the inversion of the same dataset as in Chapter 9, this time with the Pan-STARRS prior. Fig. 10.6 shows the resulting distribution in the Galactic plane, to be compared with the previous maps. More details can be found in the published article included in this chapter. It is shown that with a more realistic prior, the Bayesian inversion has an improved accuracy and assigns more reliable distances to the interstellar medium structures. The improvement is shown using the 2D Schlegel et al. (1998) map and soft X-ray emission maps. Moreover, we compared the bubbles found in our 3D map and locations of recent supernovae, modelled by Breitschwerdt et al. (2016). The

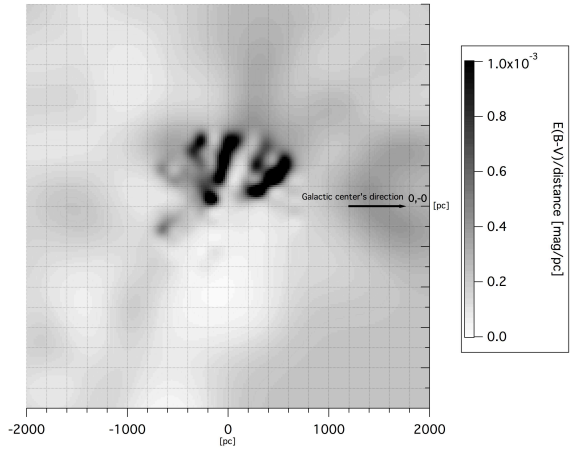
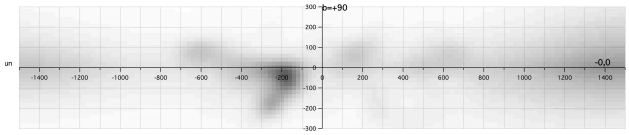
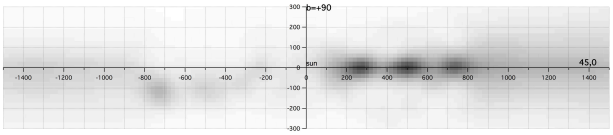


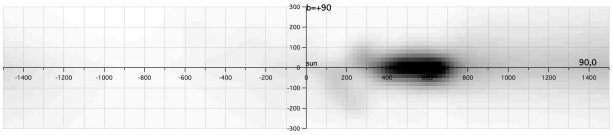
Figure 10.6 – Galactic plane cut for the built prior based on Pan-STARRS map. The Sun is at (0,0) and the Galactic centre to the right. Longitudes increase counter-clockwise. Color scale in magnitude per parsec.



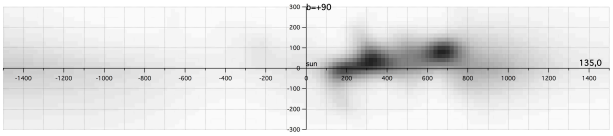
(a) Cut l= 0°.



(b) Cut l= 45°.



(c) Cut l= 90°.



(d) Cut l= 135°.

Figure 10.7 – Vertical cut of different longitude planes in the 3D map; c) and d) vertical cuts have the negative parts in the part of the prior which remains as the analytical one, because in that directions there are not Pan-STARRS data. Colour scale as in 10.6.

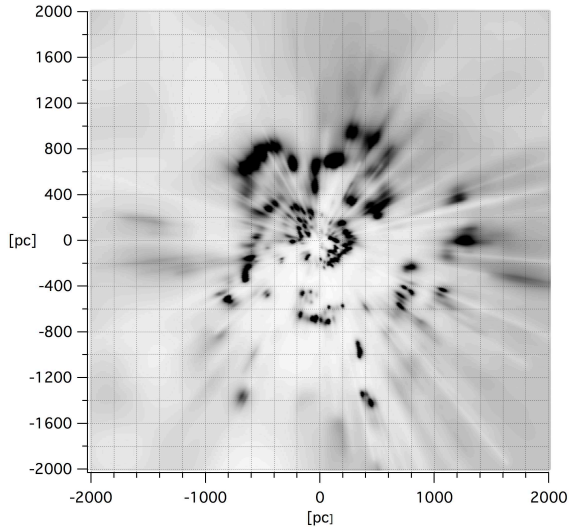


Figure 10.8 – Galactic plane cut for map based on Lallement14 + APOGEE DIB and with Pan-STARRS prior. The Sun is at (0,0) and the Galactic center to the right. Longitudes increase counter-clockwise. Colour scale as in 10.6.

structures' locations are validated in Capitanio et al. (2017) using molecular clouds catalog. Close to the Sun, the prior does not play a strong role due to the large number of individual data, and Local Bubble keeps the same shape, as well as the small structures in the second quadrant closer than 200 pc. In the second quadrant at 600 pc there is a more important change, because some elongated structures are now compacted into a single dense filament. Note that this inversion combines three different kinds of data: extinction based on photometry for individual stars, extinctions from spectroscopic observations (from DIBs), and extinctions based on statistical methods.

10.6 Maps representation

In this thesis and in different papers, we must represent a 3D map in 2D figures, in particular we are not interested at represent projections, but slices in the 3D, interpolated in the 3D matrix. I usually show for illustration the Galactic plane and some vertical planes, perpendicular at the galactic plane. Actually, for me the Galactic plane is the plane at galactic latitude equal zero. In Fig 10.9 there is the same map with different colour scale; a) panel is a representation in logarithmic scale and it is particularly interesting if we are interesting in the variability in the low values, b) panel represents the galactic plane in linear colour scale without saturation, but it is interesting only if you want to look at the very dense clouds. To represent the low redded regions it is more useful make a saturation in the colour scale, but also in this case there are many possibility. Panel c) ad d) are two examples in monochrome and polychrome palettes. The first one is particularly good because there are not "special region", in comparison at the second one where the yellow regions seems to surround special area. The second one is on the other hand better to

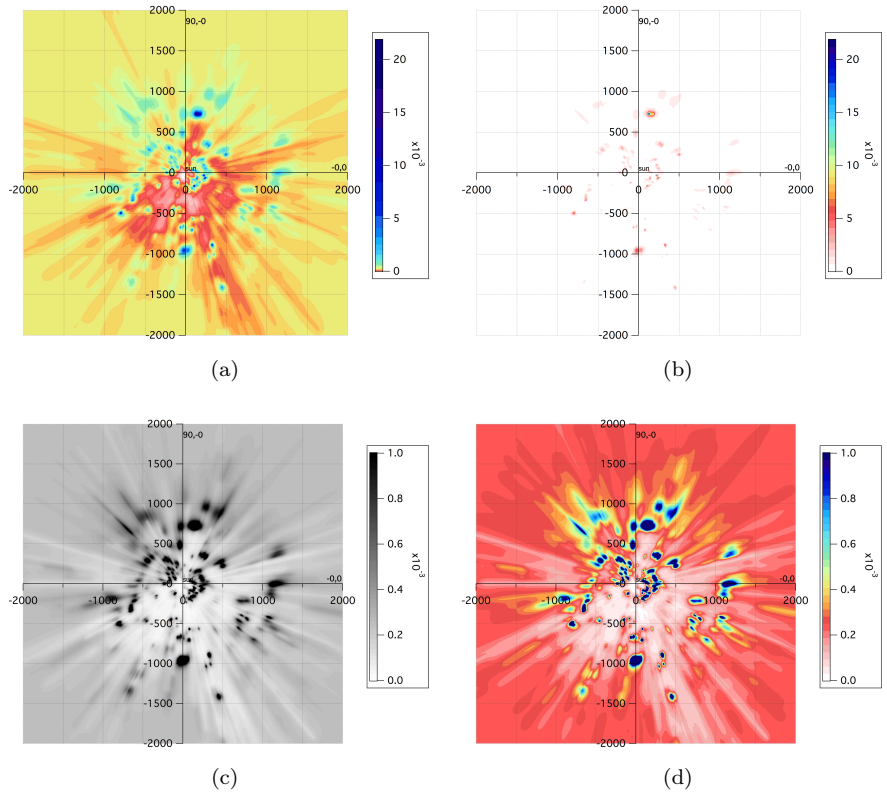


Figure 10.9 – Same plane for the same map but with three different colorscale, to compare the difference. a) Logaritmic color scale with the real value mag/pc in the map. b) Linear color scale with the real value mag/pc in the map. c) Colourblind friendly linear color scale, with color in range [0:0.001]. d) Coloured linear color scale, with color in range [0:0.001]

identify the dynamics in the low values regions.

For this reason, the 3D map is available and for different uses different representations could be more or less appropriated. In this thesis, I try to note everywhere the colour scale, but I prefer the c) and d) representations, because I can identify easily the Local Bubble, cavities and regions where the analytical prior is dominant.

10.7 Maps errors

Our excess of colour/pc map contains some errors, in distances positions and fro the reddening/pc values. Uncertainties depend from the error on data - on distances and on reddening values, the positions on the sky are assumed as without errors. Uncertanties' propagation is very complicated by the relationship between prior and data, the reciprocal covariances and additional uncertainties vary because of the bias in observations. Sale and Magorrian (2014b) proposed a technique for calculating errors during the Bayesian inver-

sion, but this evaluation are too expensive in computational time to be really applied. Without an assigned error it is very difficult to interpret the map, so it is necessary to develop an approach to obtain some estimates of errors. In the strategy I used, I processed separately the errors on distances and excess of colour, even if they are connected in inversion. I did not calculate the errors in the excess of colour/pc value but on the integrated value on the line of sight.

Uncertainty on the clouds' positions

I have assumed that distance errors are directly related to the spatial distribution of the volume of the star targets, and not to the errors assigned to them. Structures cannot be constrained to a better position than is acceptable by the Nquist-Shannon theorem, otherwise the inversion produces variations in the distribution of opacity with a frequency too high in relation to the distribution of data, which should be considered noise. Translated in a very simplified way, if in a certain volume of $10pc^3$ I have only 9 star targets, the reversal can not have an accuracy of more than 10 pc at that volume.

I calculate for each point on the map the smallest sphere centred in that point that contains 10 targets - because I don't know the spatial distribution of the sphere, nor the uncertainties of the targets. As our map for the Capitanio et al. (2017), it contains 800x800x120 points, and it is based on $\sim$ 28000 stellar targets, this approach is too expensive in computational time. Another simpler technique is to calculated for each map's point the mean targets' volume density N/pc^3 around this point, using a fixed volume, and the error is the reciprocal value of the mean volume density.

Uncertainty on the excess of colour value in the maps

Because of the fixed kernels and the σ, our reversal tends to smooth the structures and is not flexible enough to recognise the denser clouds. We estimated the error on the total excess of colour on the line of sight simply looking at the difference between excess of colour in input and excess of colour in output. For each point in the map $p = (x, y, z)$, there are considered the N stellar targets closer to p. The difference between the input excess of colour and the output excess of colour is averaged. We will obtain a total cube of error on excess of colour. As the emission maps give a maximum value for the excess of colour, I used the Schlegel et al. (1998) map as maximum value for excess of colour. It is not a *statistically true* incertitude on the map, but it helps to understand where our maps are more reliable.

10.8 STILISM website

In the STILISM project frame, they developed a new website[1] where they published the 3D maps inverted during my thesis, and the incertitudes, and it is possible to visualise extinction curves (Fig 10.10) and slices at different inclinations.

[1] stilism.obspm.fr

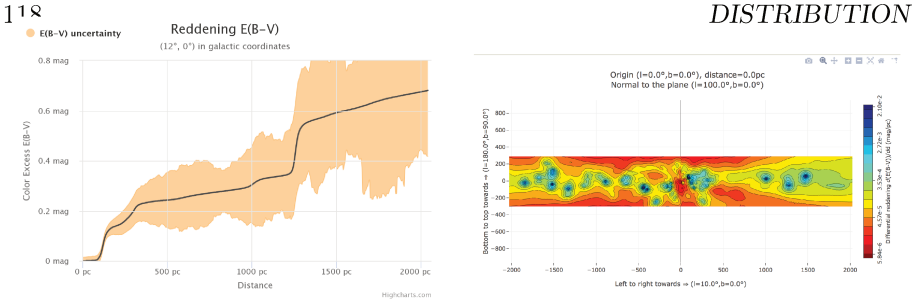


Figure 10.10 – Top) Example of extinction curve at l = 12° available on the STILISM website, from the Lallement et al. (2018) map. Bottom) Example of visualisable image on the STILISM website, a vertical slice at longitude 10°.

10.9 Piecewise-defined inversion

Also with this improved prior, in case of large dataset to invert, the direct inversion is too much time-consuming. It was mandatory to find an appropriate strategy to menage data from 10^4 to 10^5. A possibility - investigated and used in this work - is to separate the input data for their position on the sky, specifically for their longitude.

This challenge was presented in the case of the inversion for the re-calibrated distance - excess of colour pairs, calibrated through APOGEE and Gaia photometry; this inversion was published in Lallement et al. (2018), where are widely detailed the dataset characteristic and the resulted map.

We separate our data in two subsets, depending on the longitude. In Fig 10.11 there are the steps for the Lallement et al. (2018) map. Note that separation in two parts permits us to strongly reduce the calculation time, because it is reduces by a factor four, and the two inversions have run simultaneously. Anti-centre and centre map are the result of the inversion of respectively 38 806 and 35 748 stellar targets; the computational time was 24 days and 17 days. The merging matrix has the same dimension than the 3D map, and it is a piecewise-defined function:

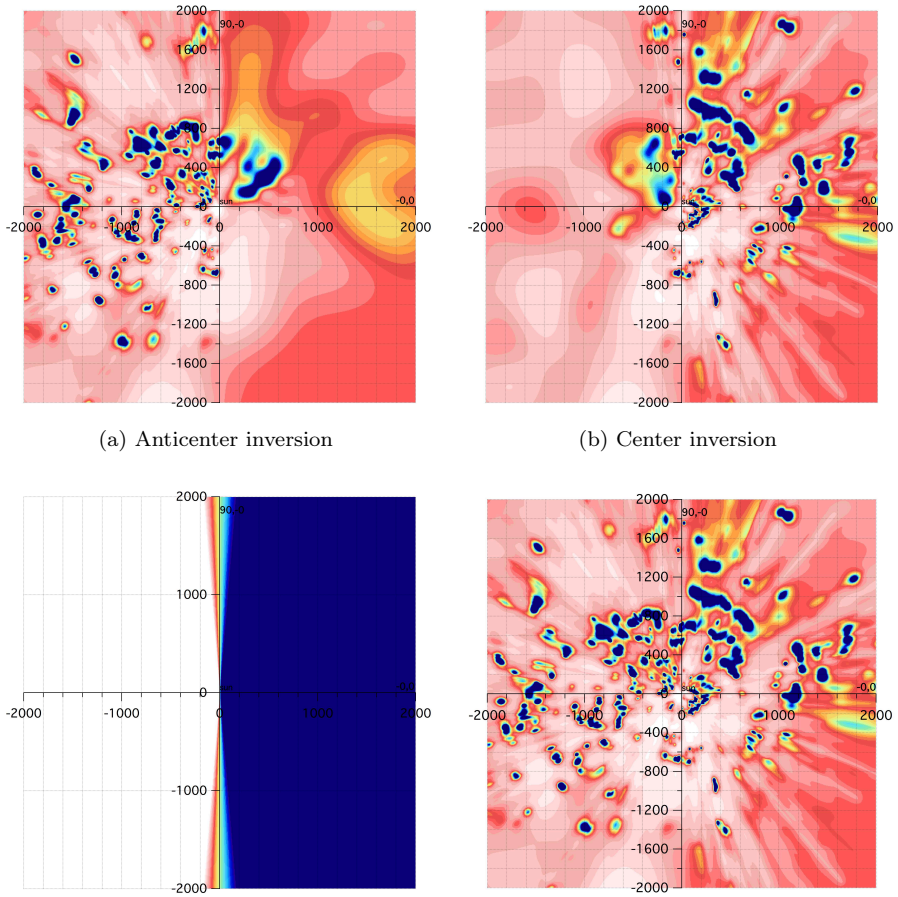
- for (x, y, z) at a galactic longitude from -85 to 85 $f(x, y, z) = 1$
- for (x, y, z) corresponding at a galactic longitude from 95 to -95 $f(x, y, z) = 0$
- for (x, y, z) corresponding at a galactic longitude from 85 to 95 $f(x, y, z) = (95 - atan2(y/x) * 180/pi)/10$
- for (x, y, z) corresponding at a galactic longitude from -95 to 85 $f(x, y, z) = -(-95 - atan2((y/x) * 180/pi)/10$

Then I used as total matrix:

$$M_{total} = M_{center} \times M_{merging} + M_{anticenter} \times (1 - M_{merging}) \tag{10.2}$$

The colour scale for the c) panel in Fig 10.11 illustrates the smoothed passage 0-1 in the superposition's interval. This technique is not optimised,

but it is a operational choice that I applied also in other cases, as illustrated in Chapter 15 and 16. Since the borders of two original maps do not show any jumps or discontinuities, the total merged map is reliable.



(a) Anticenter inversion

(b) Center inversion

(c) Matrix for merging

(d) Total resutls

Figure 10.11 – Galactic planes for the inversions used to reach at the global map Lallement et al. (2018) a) the Galactic plane for the anti-centre inversion, b) the Galactic plane for the centre inversion, c) the Galactic plane for the merging matrix, d) the final Galactic plane. Colour scale for panel a), b), d), it is like 10.9 - panel d). The colour scale of panel c) is the interval (0,1).

Astronomy & Astrophysics manuscript no. Capitanio2017
October 24, 2018

Three-dimensional mapping of the local interstellar medium with composite data

L. Capitanio[1], R. Lallement[1], J.L. Vergely[2], M. Elyajouri[1], and A. Monreal-Ibero[1]

[1] GEPI, Observatoire de Paris, PSL Research University, CNRS, Université Paris-Diderot, Sorbonne Paris Cité, Place Jules Janssen, 92195 Meudon, France
e-mail: letizia.capitanio@obspm.fr
[2] ACRI-ST, Route du Pin Montard, Sofia-Antipolis, France

Received xx, 2017; accepted xx, 2017

ABSTRACT

Context. Three-dimensional maps of the Galactic interstellar medium are general astrophysical tools. Reddening maps may be based on the inversion of color excess measurements for individual target stars or on statistical methods using stellar surveys. Three-dimensional maps based on diffuse interstellar bands (DIBs) have also been produced. All methods benefit from the advent of massive surveys and may benefit from Gaia data.
Aims. All of the various methods and databases have their own advantages and limitations. Here we present a first attempt to combine different datasets and methods to improve the local maps.
Methods. We first updated our previous local dust maps based on a regularized Bayesian inversion of individual color excess data by replacing Hipparcos or photometric distances with Gaia Data Release 1 values when available. Secondly, we complemented this database with a series of $\simeq$ 5,000 color excess values estimated from the strength of the λ15273 DIB toward stars possessing a Gaia parallax. The DIB strengths were extracted from SDSS/APOGEE spectra. Third, we computed a low-resolution map based on a grid of Pan-STARRS reddening measurements by means of a new hierarchical technique and used this map as the prior distribution during the inversion of the two other datasets.
Results. The use of Gaia parallaxes introduces significant changes in some areas and globally increases the compactness of the structures. Additional DIB-based data make it possible to assign distances to clouds located behind closer opaque structures and do not introduce contradictory information for the close structures. A more realistic prior distribution instead of a plane-parallel homogeneous distribution helps better define the structures. We validated the results through comparisons with other maps and with soft X-ray data.
Conclusions. Our study demonstrates that the combination of various tracers is a potential tool for more accurate maps. An online tool makes it possible to retrieve maps and reddening estimations (http://stilism.obspm.fr).

Key words. – ISM: lines and bands – ISM: dust, extinction – Line: profiles

1. Introduction

Three-dimensional (3D) maps of the nearby and distant Milky Way interstellar medium (ISM) are useful multipurpose tools have been only recently developed at variance with the extremely detailed, multiwavelength 2D emission maps or spectral cubes that have been accumulated over decades. As a matter of fact, a revolution is underway that dramatically changes the situation: photometric, spectrophotometric, and spectroscopic massive stellar surveys have started to provide extinction and absorption data that, coupled with parallax or photometric star distances, bring the third dimension for the ISM, i.e., the distance to the clouds. Thanks to 3D maps of the interstellar matter, structures overlapping on the sky are now converted into distinct clouds distributed at different distances, which will progressively allow emission components to be assigned a well-located production site.

Three-dimensional maps or pseudo-3D maps have already been produced based on various kinds of stellar data; that is mainly photometric extinction but also diffuse interstellar bands (DIBs) and spectral lines of gaseous species (Marshall et al. 2006; Vergely et al. 2010; van Loon et al. 2013; Welsh et al. 2010; Lallement et al. 2014; Schlafly et al. 2014; Sale & Magorrian 2014a; Kos et al. 2014; Schultheis et al. 2014; Zasowski et al. 2015; Green et al. 2015; Schlafly et al. 2015; Rezaei Kh. et al. 2016). Various methods have been used to synthesize the distance-limited data, including full 3D tomographic inversion of individual sightline extinctions (Vergely et al. 2001; Sale & Magorrian 2014b), statistical methods based on photometry, and stellar population synthesis or color-color diagrams; see the Monreal-Ibero & Lallement (2015) proceedings for recent developments in this field. In the coming years, maps of increasing quality are expected thanks to continuing or new stellar surveys, on the one hand, and precise Gaia parallaxes, on the other hand.

Each mapping technique has its advantages and limitations. Statistical methods require large amounts of targets and are appropriate at large distances, while the synthesis of individual line-of-sight data may work closer to the Sun. Near-infrared (NIR) or infrared (IR) surveys allow one to go deeper in cloud opacities compared to optical data but are more difficult to correct for telluric emissions and absorptions. Maps based on DIBs may slightly suffer from the fact that DIB carriers tend to disappear in dense cores of clouds (Lan et al. 2015), thereby potentially reducing the resolution. Nevertheless, there are numerous DIBs and their combinations may become powerful tools

to probe the physical state of the encountered media. Whatever the technique, a major difficulty is associated with the decrease with distance of the achievable spatial resolution in the radial direction, due to increasing uncertainties on target distances. Full 3D inversions that combine the radial and ortho-radial information may partly overcome this difficulty provided they are based on massive datasets, however at the price of strongly increased computing time (see, e.g., Sale & Magorrian (2014b)).

We present a first attempt to address these limitations by means of a combination of three different datasets as follows: (i) individual color excess measurements from broad- and narrowband photometry of nearby stars (mostly ≤500pc), (ii) NIR DIB equivalent widths (EWs) measured in individual spectra of more distant stars (500-1500pc), and (iii) color excess measurements at larger distance (up to 3kpc) based on a statistical analysis of multiband photometric data. Our purpose is to test such an association and its potential for improved 3D mapping of nearby or distant ISM. The first two datasets are merged by means of a conversion from DIB strengths to E(B-V) color excesses. A Bayesian inversion is applied to the resulting composite database. The large distance color excess data are used to build a very low-resolution 3D distribution that is subsequently used as the prior distribution in the Bayesian inversion to replace a plane-parallel homogeneous distribution.

Prior to the combination of these different sources, we analyzed the changes that are induced in the inverted 3D maps based on the first color excess dataset alone when Hipparcos or photometric distances are replaced with new Gaia Data Release 1 (DR1) parallax distances (the TGAS catalog), when available.

In section 2 we describe the three datasets used in the inversions. In section 3 we briefly recall the Bayesian technique developed in Vergely et al. (2010) and Lallement et al. (2014) (hereafter LVV) and the changes introduced in the method. We also describe the hierarchical method developed to build the large-scale prior distribution. Section 4 describes the 3D map evolution introduced successively using Gaia distances, the addition of the DIB-based data, and finally the large-scale prior map. Section 5 discusses uncertainties and shows comparisons between the resulting final map and other measurements of cloud distances or cavities. Section 6 discusses the results and perspectives of such composite maps.

2. Data

2.1. Individual color excess data

We started with the reddening dataset of ≃22,500 sightlines compiled and inverted by LVV. It comprises four different catalogs that were cross-matched and homogenized; for more details about the stellar types and characteristics of the calibration method for each separate dataset, see Vergely et al. (2010); Cramer (1999); Nordström et al. (2004); Dias et al. (2012). Most target stars are nearby (d≤500pc), cluster data only correspond to larger distances. The four catalogs are complementary in terms of sky coverage. Of the target distances used for the inversion presented in LVV, 75% were photometric. For one-quarter of these distances the Hipparcos distance was preferred. In the four inversions presented here we kept the same error bars on the reddening as those used by LVV. The only difference in the LVV series of targets is the removal of ≃400 targets with photometric distances that had no catalog identifications. The coordinates, brightnesses, estimated distances, and reddening reasonably of these ≃400 targets correspond to those of nearby stars. Therefore, these targets were not excluded for the inversion presented in LVV. However, as we discuss in section 4, after the addition of new, more distant APOGEE targets and the subsequent mapping improvement we realized that for these particular targets the color excesses were significantly above other data in the same area. For this reason, we conservatively excluded the targets. Fig. 1 shows the distribution of the remaining targets over the sky.

As shown in LVV, the distribution of the targets is highly biased toward low reddening regions in the sky. This had strong consequences on the resulting maps; cavities surrounding the Sun were well defined along with the boundaries of the first encountered dense clouds. On the other hand, clouds located behind foreground dense structures were either not mapped or were very poorly mapped.

2.2. Measurements of DIBs and DIB-based color excess estimates

Except for the recent identification of ionized fullerenes C_{60}^+, the carriers of the DIBs are still unknown, and it may seem strange to use absorptions due to unknown species for mapping purposes. However the link between DIB strengths and the columns of interstellar matter is strong and successful DIB-based 3D mapping has now been carried out (see Sect. 1). Up to now such mapping used solely DIBs. At variance with these works, here we attempted to merge color excess data and DIB absorption data and perform a unique inversion of a composite dataset. This requires a conversion of DIB equivalent widths into reddening (or the inverse); because none of the DIBs have been shown to be perfectly correlated with the reddening, such a mixing of color excesses computed in so many different ways— for example, with photometric and spectrometric techniques—is potentially hazardous. On the other hand, initial comparisons between DIB equivalent widths and reddening radial profiles in the same sightlines have demonstrated that their similarity (Puspitarini et al. 2015) and global maps also have strong similarities (Kos et al. 2014). Additionally, it is possible to account for the variable E(B-V)/DIB ratio by increasing error bars on the estimated E(B-V), which we have carried out in this work.

2.2.1. Data

We used the results of the SDSS-III Apache Point Observatory Galactic Evolution Experiment (APOGEE), which is one of the four Sloan Digital Sky Survey III (SDSS-III; Eisenstein et al. 2011; Aihara et al. 2011) experiments. The APOGEE uses a 300-fiber multiobject spectrograph working in the near-infrared (H-band; 1.51- 1.70 μm) at high spectral resolution (R∼22 500; Wilson et al. 2010). Specifically, we downloaded spectra from the SDSS data release[1] (DR12; Alam et al. 2015), which provides all the data taken between April 2011 and July 2014.
In addition to the ∼160 000 scientific targets selected from the source catalog of the Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006) and distributed across all Galactic environments, APOGEE contains a sample of ∼17 000 spectra of hot stars, called telluric standard stars (TSSs). The TSSs form the basis for the correction of the telluric absorption lines across all field types (see Zasowski et al. (2013), for a detailed description of the different target classes). These stars are the bluest stars on a given APOGEE plate with a magnitude in the range 5.5 ≤ H ≤ 11 mag and are therefore hot and bright stars with spectra that are most often featureless.
The APOGEE products contain the decontaminated spectra

[1] http://www.sdss.org/dr12/

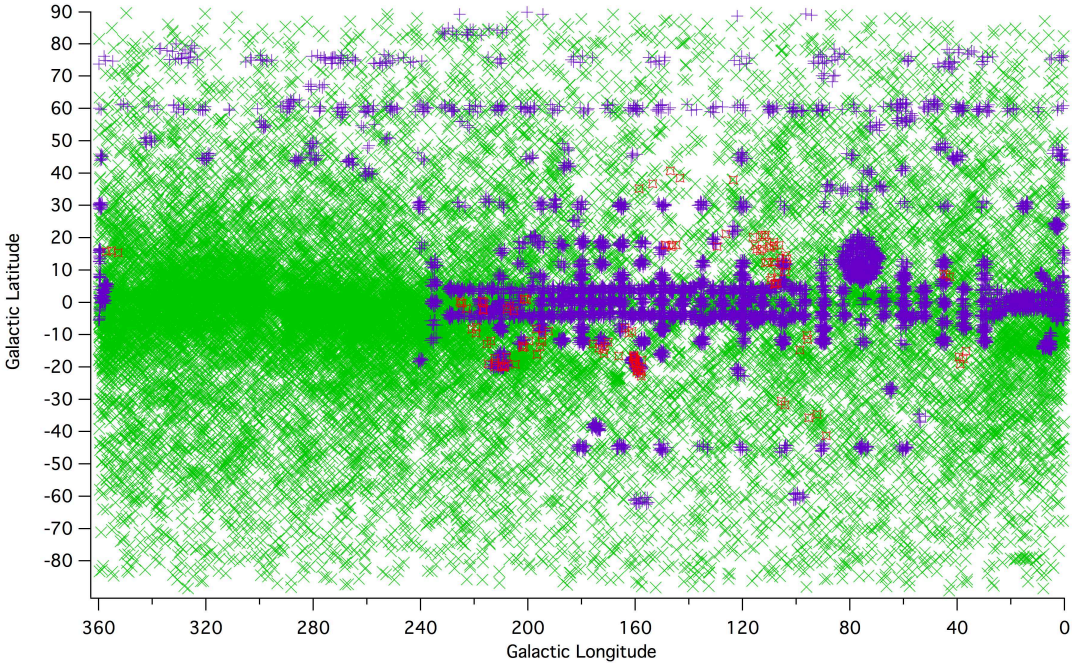


Fig. 1: Location of the targets from the LVV catalog compilation (green crosses) and of the APOGEE target stars retained for the DIB extraction and the inversion (violet signs). Coordinates are Galactic. Also shown are the directions of the molecular clouds of the Schlafly et al. (2014) catalog mentioned in section 5.2 (red signs).

and synthetic stellar spectra that provide the main stellar line locations and relative depths and widths. Both were used by Elyajouri et al. (2016) to extract a catalog of λ15273 DIB measurements from 6700 TSS spectra. Further details on the sample can be found in Elyajouri et al. (2016, and references therein). Weaker DIBs have been further studied based on the same TSS spectra (Elyajouri et al. 2017). Here we focus on the strongest 15273 Å DIB with the aim of extracting as many accurate measurements as possible from both scientific targets and TSSs possessing a Gaia parallax.

We analyzed the series of 25 409 DR12 spectra of targets belonging to the Tycho catalog to determine the presence or absence of a detectable DIB at 15273 Å and measure its equivalent width (resp. an upper limit) in case of positive (resp. negative) detection. We took as a starting point the methodology presented by Elyajouri et al. (2016) for TSSs spectra. We kept the same fully automated fitting technique and method of classification but we adapted the fitting constraints to take into account the inclusion of targets cooler than the TSSs. The code was developed using the IGOR PRO environment 2.

2.2.2. Fitting method

We restricted the fit to the spectral range [15 260 − 15 280]Å in the vicinity of the DIB. This range is wide enough to ensure an adequate sampling of the neighboring Bracket 19-4 stellar line at 15264.7 Å. We used for the fit the Levenberg-Marquardt algorithm implemented in IGOR PRO[2].

We fit the data to a combination of an adjusted synthetic spectrum, DIB model, and smooth continuum as shown in equation 1. Below we provide information specific to each of the equation variables as follows:

$$M_\lambda = [S_\lambda]^\alpha \times DIB[\sigma, \lambda, D] \times (1 + [A] \times \lambda), \quad (1)$$

- $[S_\lambda]^\alpha$, an adjusted stellar spectrum, where S_λ is the stellar model provided by APOGEE. We included the scaling factor α to adjust the model stellar line depths to the data.
- $DIB[\sigma, \lambda_c, D]$, the DIB shape was modeled as a Gaussian function allowed to vary in strength, depth, and velocity shift with three free parameters associated with its Gaussian RMS width (σ), central wavelength (λ_c), and depth (D).
- $(1 + [A] \times \lambda)$, a 1-degree polynomial function introduced to reproduce the continuum around the DIB as close possible.

An example illustrating our fitting procedure is shown in Figure 2.

2.2.3. Selection criteria and error estimates

Residuals, computed as differences between data and model, were used to obtain three measurements of the standard deviations RA, RB, and RC. To do so, we performed the local fit and a second fit over the whole spectral range covered by APOGEE using the same fit function as that presented in equation 1.

We selected the standard deviations RA, RB, and RC of the fit residuals, as follows:

- RA from the local fit was calculated in a region A = [λ_c - 10 Å , λ_c+ 10 Å] close to the DIB.
- RB was calculated over a spectral range B= [15890-15960] Å, which is free of DIB absorption, strong stellar lines, and telluric residuals.
- RC was obtained from region C= [15200-15250] Å relatively close to the DIB, and is also free of DIB and potentially

[2] http://www.wavemetrics.com

contaminated by stellar residuals.

These are used for the selection criteria shown in Table 1 and also used below to estimate the uncertainty on the DIB equivalent width. The uncertainty is conservatively estimated based on the following formula:

$$\delta_{EW} = \sqrt{2\pi}\,\sigma\, max(RA, RB, RC), \quad (2)$$

where σ is the Gaussian RMS width that results from the fit. The quantity max (RA,RB,RC) is the maximum of the three standard deviations.

Table 1 shows the fitting constraints that were used to select our final sample. They are somewhat different from the criteria of Elyajouri et al. (2016). In particular, we no longer used HI 21 cm data to limit the width of the DIB, but we used limits based on the FWHM histograms built in this previous work. Representative examples illustrating our fitting procedure and the various categories are shown in Fig. 2.

2.2.4. Final data selection for the inversion

Equivalent widths and stellar rest-frame wavelengths are determined from the best-fit parameters in each sightline for the 25,000 attempted spectra of targets belonging to the Tycho catalog. We rejected those cases where the fit failed due to low S/N, an inadequate stellar model, or very strong telluric contamination most of the time, and we retained 9913 spectra following the criteria in Table 1. The corresponding targets were crossmatched with the targets of the Gaia DR1 catalog (Arenou et al. 2017) to obtain a catalog of 7800 stars with Gaia distances. We added to the DIB measurement an additional 50% uncertainty to account for the E(B-V)/EW variability associated with the influence of the physical properties of the encountered media. As a matter of fact, from the most neutral to the most ionized media, this ratio may reach a variability factor of three (see, e.g., Elyajouri et al. (2017)). After the additional filtering for our criteria on target distances and distance uncertainties (see Sect 4.1), only 4886 targets were retained for the present inversion. Other targets will become useful after the next Gaia data releases. The number of retained targets is small with respect to the initial number of APOGEE spectra of Tycho stars (less than 20%) and targets from the reddening catalog. Despite the limited number of these retained objects, they are precious since they allow us to probe more reddened regions, in particular to reveal a second or third rank of clouds, something often precluded using the bright targets measurements in the optical.
We converted the DIB EW into a color excess using the Zasowski et al. (2015) formula

$$EW = 0.102\, A_V^{1.01}\,\text{Å} \quad (3)$$

and the classical average total-to-selective ratio A_V/E(B-V)= 3.1 (Savage & Mathis 1979). The relationship established by Zasowski et al. (2015) has been confirmed, by Elyajouri et al. (2016), to be valid locally based on the closer TSSs.

2.3. Subsample of Pan-STARRS color excess measurements

Green et al. (2015) have used Pan-STARRS and 2MASS photometric measurements toward 800 millions targets to build remarkable 3D dust reddening maps with a very high angular resolution (5-15 arc-minutes) covering distances from ≃300 pc to 4-5 kpc. The publicly available maps contain a quality flag that allows to exclude results at small or large distances that are uncertain because of limitations on the number of targets for geometrical or sensitivity reasons. Both surveys were conducted from the Northern Hemisphere. We used those data to build a low-resolution 3D *prior* distribution to be used in our Bayesian inversion in replacement of the previous analytical solution corresponding to a homogeneous plane-parallel opacity.

To do this, we downloaded a grid of Pan-STARRS E(B-V) estimates for lines of sight equally spaced every 0.5 degree in Galactic latitude and longitude and for all distances available from the Pan-STARRS online tool, in this way obtaining 3,071,230 pairs of distance and extinction after exclusion of all data points flagged as uncertain. As result of this process, we do not benefit from the very high angular resolution of the maps, however this information would be lost in our approach. In many directions the first 200 pc are flagged because of the limited number of targets available to define the reddening in the very small solid angle chosen for this Green et al. (2015) study. However this is where the present inversion of individual data may be advantageous. The method applied to these Pan-STARRS results is described in Section 3.2.

3. Inversion techniques

3.1. Bayesian inversion of individual sightlines

The Bayesian inversion technique is based on the pioneering work of Tarantola & Valette (1982) and its first developments and applications to the ISM are described in detail in Vergely et al. (2001, 2010); Welsh et al. (2010); Lallement et al. (2014). Briefly, the inversion optimizes an analytical solution for the 3D opacity distribution through the adjustment of all integrated opacities along the target sightlines to the observed color excess data. Because the solution is strongly under-constrained, it is regularized based on the assumption that the 3D solution is smooth. Opacities ψ at two points (x), (x') in space follow correlation kernels: here $\psi_1(x,x') = \exp(\frac{-\|x-x'\|^2}{2\xi_1^2})$ and $\psi_2(x,x') = \frac{1}{\cosh(-\|x-x'\|/\xi_2)}$, where the first (resp. second) term represents the compact (resp. diffuse) structures. The correlation distances ξ_1 and ξ_2, which roughly correspond to the minimum size of the computed structures, are evidently limited by the average distance between the target stars. Here we have chosen ξ_1=15 pc and ξ_2=30 pc. As we discuss in the next sections, such kernels are valid in internal regions where the target distribution is the densest. The model variance, which controls the departures from the prior distribution, is σ_1=0.8 for the first kernel and σ_2=1.0 for the second (see LVV for its definition). The Tarantola & Valette (1982) formalism allows one to compute the least-squares solution iteratively. For the present maps the convergence is reached in about six iterations. Uncertainties on the reddening and target distance are treated in a combined way (see Vergely et al. (2001)). The distribution is computed in a volume of 4 kpc by 4 kpc along the **Galactic plane** and 600 pc along the perpendicular direction.

Compared to the method used in Lallement et al. (2014), an additional treatment of outliers was introduced. During each iteration except the first two, in case one measured color excess value is found to be incompatible with the kernels and the surrounding data, its error bar is multiplied by 2. This method has proven to solve for data points that have underestimated er-

Flags	λ_c Å	Width Å	Depth Å
Detected (4861)	[15 268 − 15 280]	[1.4 − 5.5]	[1.5× max(RA,RB,RC) −0.1]
Narrow DIBs (364)	[15 260 − 15 280]	[1 − 1.4]	[1.5× max(RA,RB,RC) −0.1]
Recovered DIBs (2515)	[15 268 − 15 280]	[1.4 − 5.5]	[2× RA −1.5× max(RA,RB,RC)]
Narrow recovered DIBs (420)	[15 268 − 15 280]	[1 − 1.4[	]2× RA −1.5× max(RA,RB,RC)]
Upper limit (1753)	[15 268 − 15 280]	≤1.4	≤ 2×RA

Table 1: Fitting constraints of λ15273 NIR DIB for the selected APOGEE spectra. RA, RB, RC are the standard deviation in three spectral ranges A = [λ_c - 10Å, λ_c+ 10Å], B= [15890-15960] Å and C= [15200-15250] Å, respectively.

rors and eliminates oscillations. The Bayesian approach implemented in Vergely et al. (2001) and subsequent works allows us to include an *a priori* knowledge of the ISM 3D distribution. This prior distribution can in principle be provided by a model that represents the density variations in the Galactic plane by an analytic law or alternatively by a distribution based on external data. The latter case is fully appropriate to our objective of combining different sources of extinction. As a matter of fact, it is possible to produce density maps by successive refinements following integration of data from different sources in a sequential and non-simultaneous manner. The advantage of such an approach is twofold: first, the amount of data to be integrated at each step is reduced, which makes it possible to reduce computation times and, second, it is easier to see the contribution of each set of data.

In this perspective, a new development was introduced in our inversion code. Now, it is possible to use any 3D prior distribution, i.e., we can replace the homogeneous plane-parallel prior that corresponded to an analytical representation (we used an exponential decrease from the Galactic plane) with any arbitrary precomputed distribution. We used this possibility to test a two-step inversion based on two different datasets as described in section 4.3. The construction of the prior distribution is described in the next section.

3.2. Large-scale prior distribution based on Pan-STARRS

We used the Green et al. (2015) color excess subsample described in 2.3 to develop a low-resolution opacity model extended to large distances, to be used as a prior for the inversion of individual lines of sight as described above. Because this reddening model is based on photometric measurements of hundreds of millions of objects, the Green et al. (2015) model has a very high angular resolution, whereas the radial resolution is limited by uncertainties on photometric distances. In principle our prior distribution is losing all the details contained in the angular variations, first because we start with a coarse angular sampling of 0.5 deg, and also because our full 3D inversion has a spatial resolution limited by the radial distance uncertainties.

Ideal inversions are based on a homogeneous spatial sampling of targets, i.e., a homogeneous volume density of targets and similar uncertainties on their locations at all distances, which allows us to retrieve structures of the same minimal spatial scale everywhere in space. But for most surveys, including Pan-STARRS, the achievable resolution is dependent on the distance to the Sun, especially owing to distance-dependent uncertainties on the star locations. The sampling of the data in this case does not allow the Nyquist-Shannon criterion to be fulfilled for the same minimal autocorrelation length everywhere; also, some of the energy attributable to high spatial frequencies is likely to be redistributed in low spatial frequencies, especially at large distances or in the radial directions when the sampling is weak. Therefore, we need an approach that allows us to retrieve only the large scales at large distances and all large and small scales at short distances. We implemented a hierarchical approach that allows us to achieve this goal in an iterative way by progressively updating the opacity model. Schematically, we first constructed a low spatial frequency map in a sphere with a large radius (here 5 kpc). In a subsequent step, we estimated spatial frequencies that are higher by a factor of 2, however this time in a smaller sphere (for instance 3 kpc) by removing data at large distances for which the available information does not make it possible to reach these frequencies using the Nyquist-Shannon criterion. The information at large distances is kept under the form of the prior, which is simply the distribution derived from the previous step. This procedure is repeated for increasingly high frequencies until their maxima compatible with the radial sampling (reached at small distances) are obtained. In practice we proceeded as follows:

1. The Pan-STARRS reddening data subsample (see section 2.3) is converted into local opacities ρ_0 through derivation w.r.t. the radial distance R,
   $\rho_0 = \delta_{E(B-V)}/\delta_R$.
2. A spatial scale S of restitution of the opacity is fixed, starting with the (large) scale that can possibly be restored on the whole map. The average opacity is computed in a 3D grid of boxes of size SxSxS. The resulting average opacity is thus dependent on the scale.
3. Based on these average opacities, maps of smoothed opacities are computed by Bayesian interpolation with a smoothing kernel twice as large as S.
4. We then move on to a scale twice smaller. S/2 x S/2 x S/2 boxes are built and the average opacity is calculated for each box. The dynamics of the average opacity for this set of boxes is stronger. We suppress the boxes that do not fulfill the Nyquist-Shannon criterion for S/2.
5. A Bayesian interpolation is carried out again with a smaller smoothing kernel and using the previous map as prior.
6. Again, the scale is divided by two and steps 2) and 3) are performed again for this new scale, etc.

A planar cut along the Galactic plane in the resulting 3D prior is shown in Fig. 3. The Sun is at center and Galactic longitude increases counter-clockwise. As expected, far from the Sun we can only restore the very large scales. On the contrary, one can reach scales on the order of 100 pc within the first few hundreds parsecs. Because Pan-STARRS data are recorded from the Northern Hemisphere, the prior in the fourth (bottom-right) quadrant appears homogeneous at large distances from the Sun. This map can be used as a prior in the inversion of the individual lines of sight for which accurate distances are available, with the goal of improving estimates of opacity fluctuations and refining the distance assignment of the structures (see Sect 4.3).

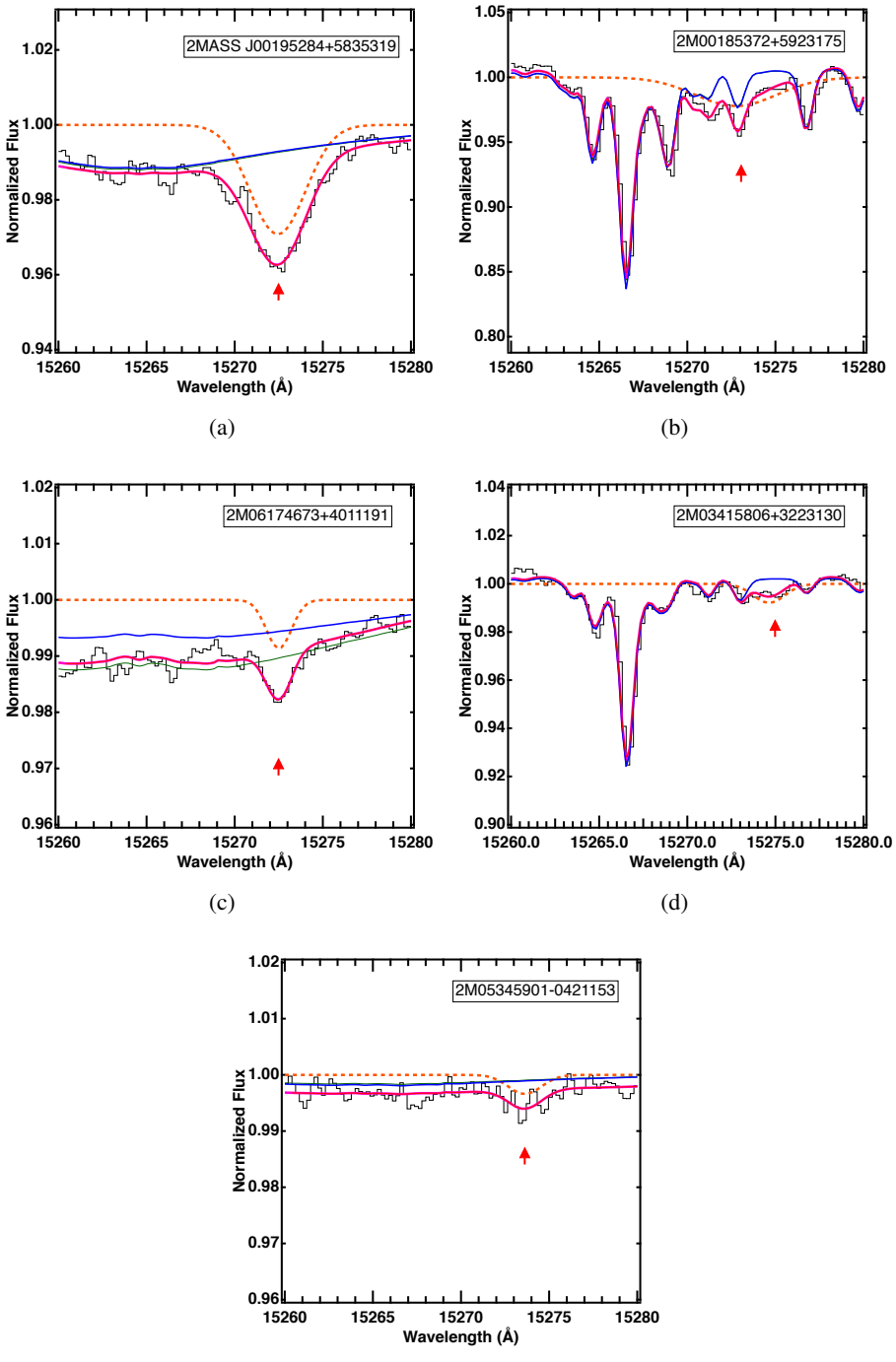


Fig. 2: Illustration of the various categories of DIB extraction: a) Detected; b) Recovered; c) Narrow; d) Recovered and Narrow; e) Upper limit. APOGEE spectra are shown with solid black curves. The initial stellar model provided by APOGEE and the model obtained after application of the scaling factor are shown in solid blue and green respectively (note that in many cases they are very similar). The DIB absorption alone is represented in orange. The solid magenta curves represent the final stellar+DIB modeled spectrum.

4. Inversion of datasets

4.1. Inverting pre- and post-Gaia individual color excess data

In order to test the influence of the new Gaia parallax distances we inverted the reddening dataset described in section 2.1 twice with exactly the same method, criteria, and parameters, but with different target distances. The first inversion (A) used the initial Hipparcos+photometric distances and errors as in LVV and the second (B) used Gaia-TGAS distances and associated errors when available, or unchanged values in the absence of Gaia result. Gaia DR1 parallaxes are about 20 times more numerous than Hipparcos parallaxes and about three to four times more accurate. These characteristics will evolve with the next releases, which will contain many more stars and reduced parallax uncer-

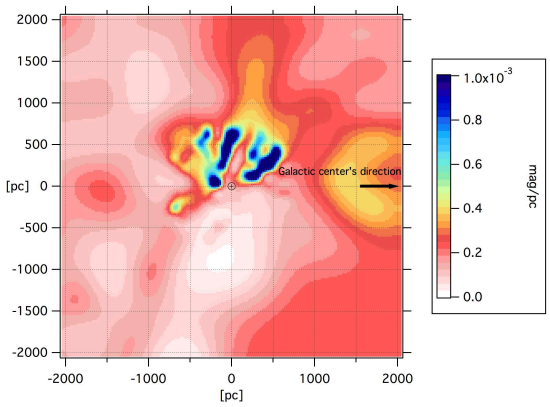


Fig. 3: Galactic plane cut in the very low-resolution 3D opacity distribution to be used as a prior in the inversion. It is based on Pan-STARRS reddening measurements of Green et al. (2015). The Sun is at (0,0) and the Galactic center to the right. Longitudes increase counter-clockwise. Units are mag pc^{-1}.

tainties. As in LVV the prior dust opacity is a plane-parallel homogenous distribution decreasing exponentially from the Galactic plane with a scale height of 200 pc. We also similarly retained for the inversion targets closer to the Galactic plane than 1000 pc and less distant than 2601 pc with relative errors on the distance smaller than 33%. The number of targets retained for the first inversion was 22467 and among these 5106 were assigned their Hipparcos distance, and the rest were photometric. In the second inversion 23444 targets were retained, with about 80% having a Gaia parallax. As the Gaia-DR1 uncertainties were 3-4 times smaller than Hipparcos uncertainties, the number of targets excluded for uncertain distances was significantly reduced. Importantly, the fraction of photometric distances decreased by a factor of four and Gaia distances started to be largely dominant.

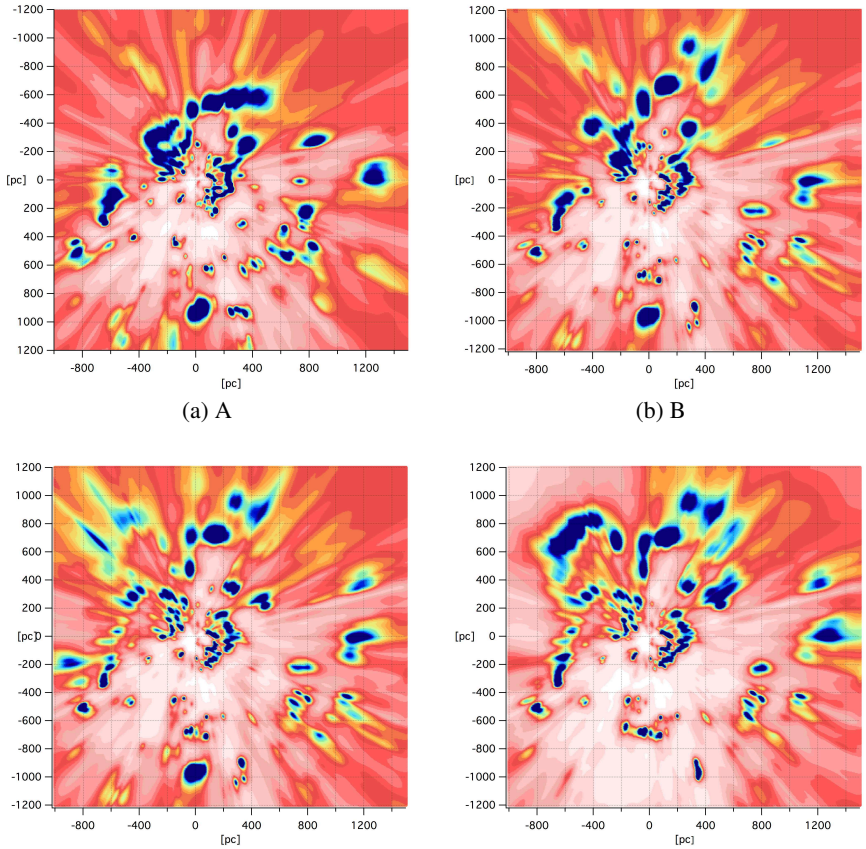
The values of the χ^2 per target are computed from the differences between the observed color excesses and the color excesses computed by integration of the model 3D distribution along the target lines of sight. These values are reported in Table 2 for both the prior and final distribution. The value of χ^2 for the prior distribution is significantly higher for the second inversion, reflecting the fact that error bars entering its computation are on average smaller because of the reduced uncertainties on Gaia-DR1 distances, especially by comparison with previous uncertainties on the photometric distances. On the contrary, the final χ^2 is lower despite reduced error bars, demonstrating that the found distribution is more compatible with the new distances than with the previous distances.

Panels a and b (top left and right) in Figure 4 show planar cuts in the two inverted opacity distributions A and B. The plane is parallel to the Galactic plane and contains the Sun. The pre-Gaia inverted distribution is not exactly the same as in LVV because of the removal of some targets and the introduction of the new method to treat outliers (see section 3.1). The difference between the two pre- and post-Gaia distributions is visible: on average the structures are more compact. Some elongated structures diminished owing to the reduction of distance uncertainties, as in the third and fourth quadrants (bottom left and right); some others appeared, which may be due to contradictions between new Gaia distances and the remaining photometric distances. The latter regions will have to be carefully studied once all targets benefit from Gaia parallaxes. The Local Bubble keeps its global shape, but looks slightly more complex owing to the higher compactness of structures. At larger distance the most conspicuous difference appears in the first quadrant, in particular in the Cygnus Rift region. At l≃70°the Rift is located significantly farther away in the new inversion, at about 800 pc instead of ≃600 pc, interestingly now in agreement with the Green et al. (2015) distance assignment. We knew that this was a region of strong oscillations in the first inversion; these oscillations have now disappeared. In the second quadrant there are substantial changes and we start to sketch a ring of clouds, more specifically the very large structure located from 100 to 400 pc is decomposed into two smaller structures.

4.2. Inversion of combined photometric and DIB-based color excess data

The same inversion code was applied to the merged catalog of photometric and DIB-based reddening estimates (inversion C). As can be seen in Table 2, the introduction of the APOGEE targets reduces both the initial and final χ^2 per point. The former decrease means that error bars on the DIB-based estimated E(B-V)s are not underestimated. The latter decrease strongly suggests that the APOGEE targets do not introduce contradictory data and their inclusion produces a reasonable solution. In comparing the new Galactic plane cut in Fig 4 and the distribution C at the bottom left with the previous inversion with photometric data alone (distribution B, top right), we can clearly see some changes; these changes are particularly apparent in the second quadrant, where some decoupling or fragmentation of large structures into distinct cloud complexes along with new structures at large distances are visible. Since 14 % of the APOGEE targets are closer than 20 pc from the Galactic plane, are highly concentrated in this quadrant, and are more distant that the LVV targets, it is not surprising to find most changes there. The fourth quadrant does not change because the APOGEE catalog adds targets observed from the Northern Hemisphere.

Since it is the first time that such measurements of different types are used, which could trigger incoherence in the inversion, it is mandatory to investigate the observed changes in all areas in order to validate this new technique. Four vertical planar cuts in the 3D opacity map are shown in Fig 9 (Orion region), Fig 10 (anticenter direction), Fig 11 (Cepheus region), and Fig 12 (Perseus region). The Perseus region is discussed in detail in Sect 4.3. The images show, in the top panel, the B opacity distribution results and, in the middle panel, the C opacity distribution after inclusion of APOGEE data. In principle, the effect of adding coherent data in poorly covered areas should result in the transformation of wide structures into smaller clumps and/or



(a) A (b) B

(c) C (d) D

Fig. 4: Galactic plane cut in the 3D opacity distribution resulting from the four successive inversions: a) A distribution: color excess data alone and Hipparcos or photometric distances. b) B distribution: same as A except for the replacement of 80% of the initial distances with Gaia-DR1 (TGAS) values. For the 20% of stars without Gaia distance we kept the previous value. c) C distribution: same as B except for the addition of DIB-based color excess estimates for stars with a Gaia-DR1 parallax. d) D distribution: same as C except for the use of a Pan-STARRS based prior distribution instead of a homogeneous distribution. Color scale, direction of Galactic center, and Sun position are the same as in Fig 3

.

DISTRIBUTION	DATA	INITIAL χ^2	FINAL χ^2
A	LVV catalog, Hipparcos+phot	5.361	1.100
B	LVV catalog, Gaia-DR1+phot	5.810	1.085
C	LVV catalog + APOGEE, Gaia-DR1+phot	5.080	1.023
D	LVV catalog + APOGEE, Gaia-DR1+phot , Pan-STARRS prior	3.521	1.006

Table 2: Evolution of the prior and final χ^2 per line of sight in the four inversions.

in the appearance of new structures, especially at large distances in locations where there was initially no targets.

The Orion region is well suited for comparisons owing to the existence in similar proportions of targets of the two types beyond 150 pc. There are ≃750 lines of sight from the photometric catalog and ≃375 from the APOGEE catalog within the longitude interval l=(208:218°). As shown in Fig 9, the two types of targets are spatially well mixed. If the data were not compatible we would observe radially elongated and/or wide structures in the middle map that were absent in the top. Instead, there are no significant changes, which demonstrates that the two datasets trace the same structures. Indeed, the figure shows that the APOGEE data reinforce the dense structures at 500 and 600 pc, at distances poorly covered by targets with photometric measurements.

Another region suited to a further check is the anticenter region. At variance with the Orion area, the LVV targets are concentrated in the first 100 pc, while APOGEE introduces many more distant targets. Fig 10 shows clearly that this addition allows us to map new clouds beyond 300 pc, in particular to transform a unique wide structure at ≃ 450pc into at least three structures located between 300 and 700 pc, confirming that infrared data such as APOGEE add valuable information in areas where stars are already reddened by close clouds. The structures at small distances already seen in the first map (inversion B) remain in the second map but appear more compact with their relative opacities somewhat changed. Because the sightlines to the distant targets cross the nearby clouds, contributing to their mapping, the increased compactness of the foreground structure

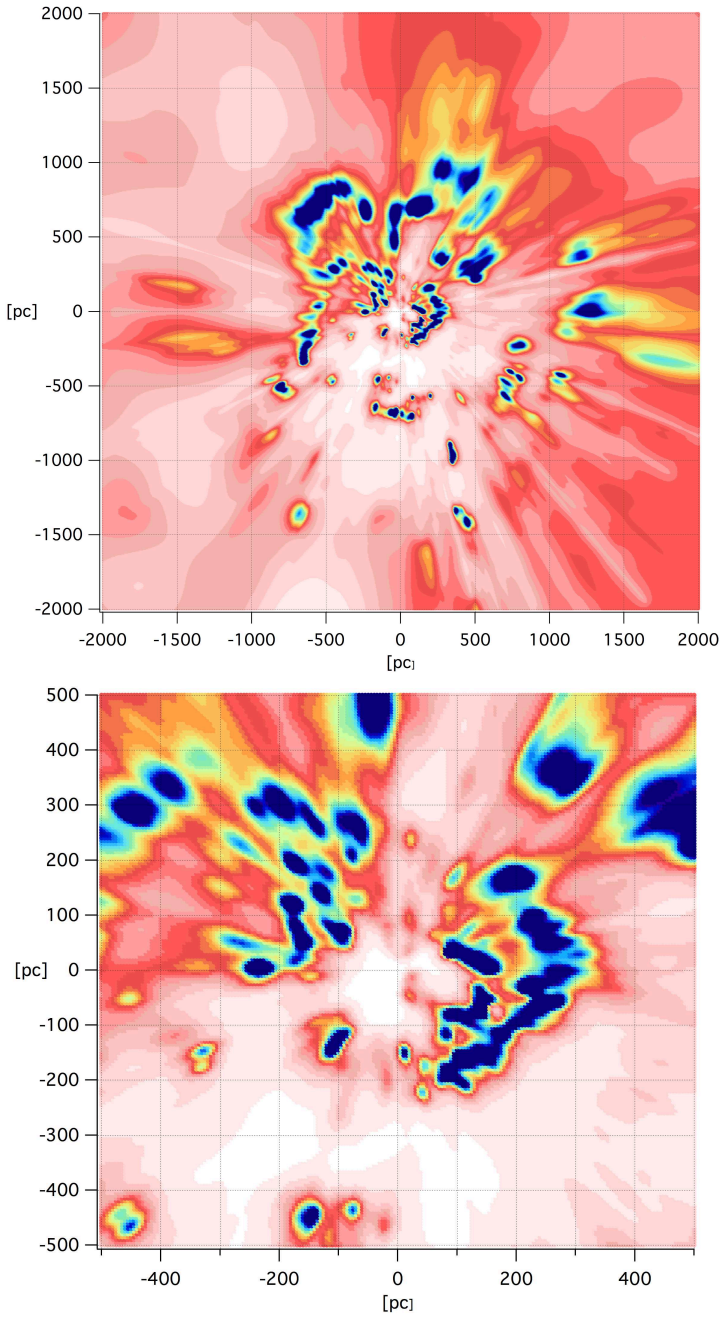


Fig. 5: Planar cut along the Galactic plane in the inverted 3D distribution D shown in Fig 4, panel d. The color scale is same as in Fig 3. Top: the entire 4000pc × 4000pc computed area. Bottom: Zoom in a 1000 pc × 1000 pc area around the Sun.

shows that the new data do not contradict the old data and instead help shape the close clouds.

We finally consider the Cepheus region in the second quadrant (Fig 11). There are several clouds distributed in the same region of the sky, which is a very good test for the inversions. Here, the number of targets in the longitude interval l=(105°;115°) increases from 390 to 664. It can be seen that the large structure at distances between 200 and 300 pc and at heights above the Galactic plane between 0 and 150 pc is now more clearly decomposed into three cores and an additional fourth cloud just below the Galactic plane at ≃ 300 pc. This last structure below the Galactic plane had apparently been erroneously associated spatially with a more distant and wide structure at 500 pc. As a result of Gaia distances and additional targets, its location has now been better defined. Indeed, the final distribution D (bottom panel of Fig 11 and section 4.3) confirms this new location. Behind those structures a hole as well as two different more distant clouds are revealed near the Galactic plane. The large structure located at about 400 pc in distribution A, which was derived from only a few targets, is now much better defined and significantly displaced farther away at about 500 pc. This again shows that using infrared data enables the mapping of distributed clouds and holes between them.

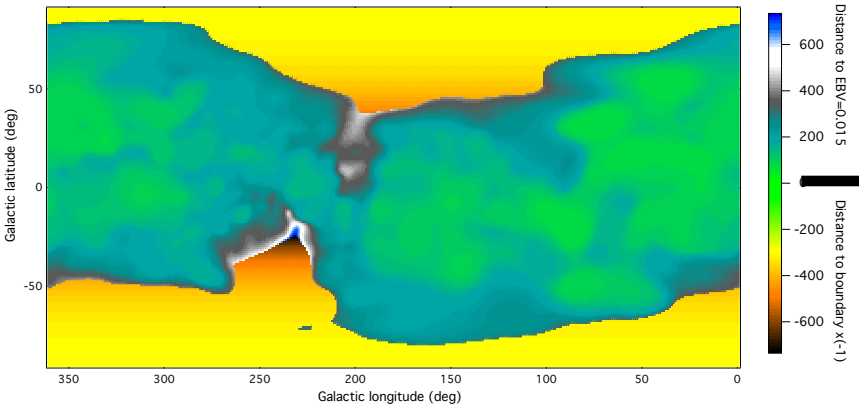


Fig. 6: Local cavity boundaries: the distance at which a color excess of 0.015 mag is reached is plotted as a function of rectangular Galactic coordinates. In very weakly reddened directions for which 0.015 mag is not reached within our computation domain, the distance to the domain boundary is plotted after conversion to negative values. This allows us to distinguish between the two cases well. The figure illustrates well the local warp of gas associated with the Gould belt, i.e., more matter at positive latitudes in the Galactic center hemisphere and at negative latitudes in the anticenter hemisphere. It also reveals very clearly the *local bubble* openings to the halo at about 210-240° longitude.

4.3. Inversion of merged data using the Pan-STARRS-based prior 3D distribution

We performed an additional inversion (inversion D) of the set of LVV+ APOGEE targets, this time replacing the plane-parallel homogeneous prior with the large-scale distribution described in Sect. 3.2, based on Green et al. (2015) results. This allowed us to test our new capacity of using an arbitrary 3D prior distribution. In the new prior, the smallest structures allowed are 100 pc wide and the kernel sizes for the largest structures are 800 pc.

Table 2 shows that using the new prior has the effect of reducing the initial χ^2 per point significantly and decreasing the final χ^2 too. This means that the data for individual sightlines are, not surprisingly, in better agreement with this more realistic prior. A comparison of panels c and d in Fig. 4 shows the differences introduced by using the new prior in the Galactic plane. As expected, the main variations are at large distance, especially beyond 1 kpc, while closer the changes are very small. The same effects are seen out of the Galactic plane, as shown in Fig 9 (Orion), Fig 10 (anticenter), Fig 11 (Cepheus), and Fig 12 (Perseus), where the bottom panels show the D opacity distribution. There are significant changes far from the Sun and negligible changes closer to the Sun, where the target density is high enough to suppress the influence of the prior. In all cases, it is interesting to see how the nearby structures connect with the distant very low-resolution Pan-STARRS based distribution. This two-step process allows large-scale and smaller-scale structures to coexist in a map, while avoiding an overinterpretation of the data. Other vertical planes drawn every 30°Galactic longitude are shown in Appendix. These vertical planes give us a general idea of the nearby ISM distribution out of the Galactic plane.

Fig 5 (upper panel) shows the dust opacity in the Galactic plane for the entire computational area 4000pc x 4000 pc. As can be seen in the Figure, at large distances there are co-existing compact structures with the minimum size allowed by the kernels and large-scale structures that mainly result from the Pan-STARRS prior. The small structures however should be considered with caution, as is discussed in the next section.

Bottom panel of Fig 5 shows a zoom in the 1000pc x 1000pc central area. It shows numerous clouds and cavities and that the shape of the so-called Local Bubble is far from simple. It is well known that the interstellar matter within the first hundreds of parsecs is very complex and characterized by the series of clouds forming the Gould belt, a wide structure tilted w.r.t. the Galactic plane by about 20° (e.g., Perrot & Grenier (2003)). The origin of this structure is still a matter of debate (see, e.g., Lallement (2015) for a recent discussion and suggested origin). Two *chimneys* connect the Local Bubble to the northern and southern halos forming a hole whose axis is nearly perpendicular to the Gould belt plane; this peculiar structure is illustrated in Fig 6. The figure shows for all directions the line-of-sight length that corresponds to an integrated color excess E(B-V)=0.015 magnitude. When the distance required to reach this color excess is larger than the distance to the boundary of the computational volume, such as toward high Galactic latitudes, we used this distance multiplied by -1. The figure shows very clearly the local warp of IS matter as well as the *chimneys* connecting to the halo.

Fig. 7 shows 2D color excesses from inversions A (top map) and D (bottom map), which are computed by integrating local opacities from the Sun to the boundaries of our computational domain; the distance to the Galactic plane is smaller than 300 pc and distances along the X and Y axes are smaller than 2,000 pc. Iso-contours from Schlegel et al. (1998) for E(B-V) = 0.025 mag and E(B-V)=0.32 mag are drawn for comparison. Compared to our previous inversion (A, at top) the structures are more compact for inversion D (at bottom), showing that the data are more consistent and the resolution is improved thanks to Gaia distances and additional data. Fig. 8 is a zoom in on these 2D maps of A and D integrated color excesses, this time in a limited region of the sky close to the anticenter direction. Compared to our previous inversion, D leads to better agreement with the Schlegel et al. (1998) contours at low reddening, demonstrating that the inclusion of DIB-based data has improved the detection and location of the structures.

5. Uncertainties and comparison with other 3D measurements

5.1. Resolution and errors

Uncertainties on cloud distances and dust columns are strongly dependent on the distance to the Sun due to the combined effect

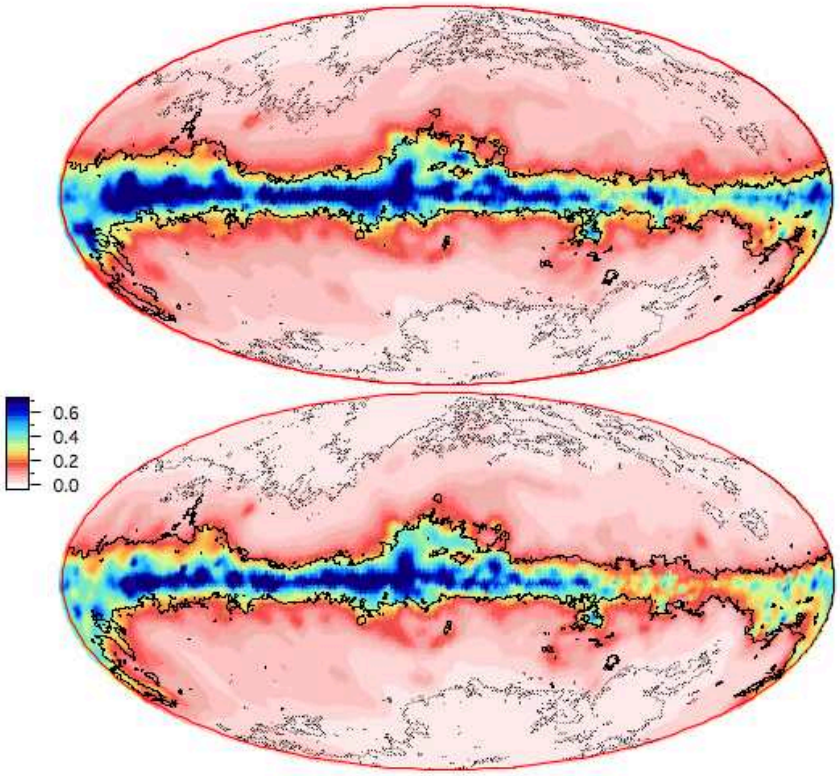


Fig. 7: Full-sky 2D map of opacity distribution integrated from the Sun to the boundaries of our computational domain (abs(Z)=300pc, abs(X)=2kpc, abs(Y)=2kpc). Top: LVV results are shown (distribution A). Bottom: the final inversion with APOGEE data and the Pan-STARRS based prior distribution are shown (distribution D). Isocontours from Schlegel et al. (1998) for E(B-V) = 0.025 and E(B-V)=0.32 are superimposed as dotted and continuous black lines. The map is in Aitoff Galactic coordinates. Longitudes decrease from +180° to -180° from left to right. The color scale refers to E(B-V), linear, in units of mag.

of target space density decrease and target distance uncertainty increase. Additionally, uncertainties vary from one region to the other since databases have strong biases. Theoretically the errors may be calculated for each point in space along with the Bayesian inversion, but this kind of evaluation is very expensive in computation time (Sale & Magorrian (2014b)). On the other hand less precise error estimates for the distance to the structures and the amount of integrated reddening may be obtained by means of simpler approaches. We present these approaches here.

In the frame of a first simple approach we assumed that the error in distance varies as the achievable resolution, i.e., depends on the target space density. We choose as an estimate of the error at each point the size of the smallest sphere around this point that contains at least 10 targets. Another simpler and computationally faster approach is to calculate again in each point the target density ρ (in pc^{-3}) in a cube of a given size centered on this point and use $(1/\rho)^{1/3}$ as the error estimate. This second technique is less precise because we keep the same volume everywhere in an arbitrary way. The results for the Galactic plane are shown in Fig 13. The results from both methods are not markedly different and as expected there is a strong increase of the distance uncertainty from the center to the peripheral regions. The error can be compared with a similar estimate by LVV: it can be seen that the addition of the APOGEE targets improves the resolution in a number of regions from the first to third quadrant. The errors on cloud localization owing to uncertainties on the target distances are limited by the use of our threshold (see Sect 4.1).

As already mentioned in LVV, estimating the error on the local dust opacity is meaningless because by principle the clouds have a minimum size and as a consequence the dust is distributed in larger volumes in the model compared to the volumes of the actual, clumpy clouds. This spreading of the opacity in volumes that are wider than the volumes of the actual clouds introduces systematic underestimates of the opacities and integrated reddenings in the cloud cores. As a result, this systematic error is expected to be small for sightlines that do not intercept small, dense cores, but may reach high values in highly reddened regions in case the reddening is generated by one or a few small structures. We come back to this point in Section 6. On the other hand, it is possible to give rough estimates of the errors on the reddening (integrated opacity) that are linked to the dataset, independently of the errors due to the assumed cloud minimum size. To do this, for every point in our 3D map we selected all targets contained in a cube of specific size around the considered point, and for those targets we calculated the mean difference between the data and model color excess. Fig 14 shows the results for the Galactic plane with a cube of 50x50x50 pc^3; in case the cube contains less than 10 targets we use a volume of 200x200x200 pc^3. Uncertainties on distances and color excess will be available on our website and displayed interactively.

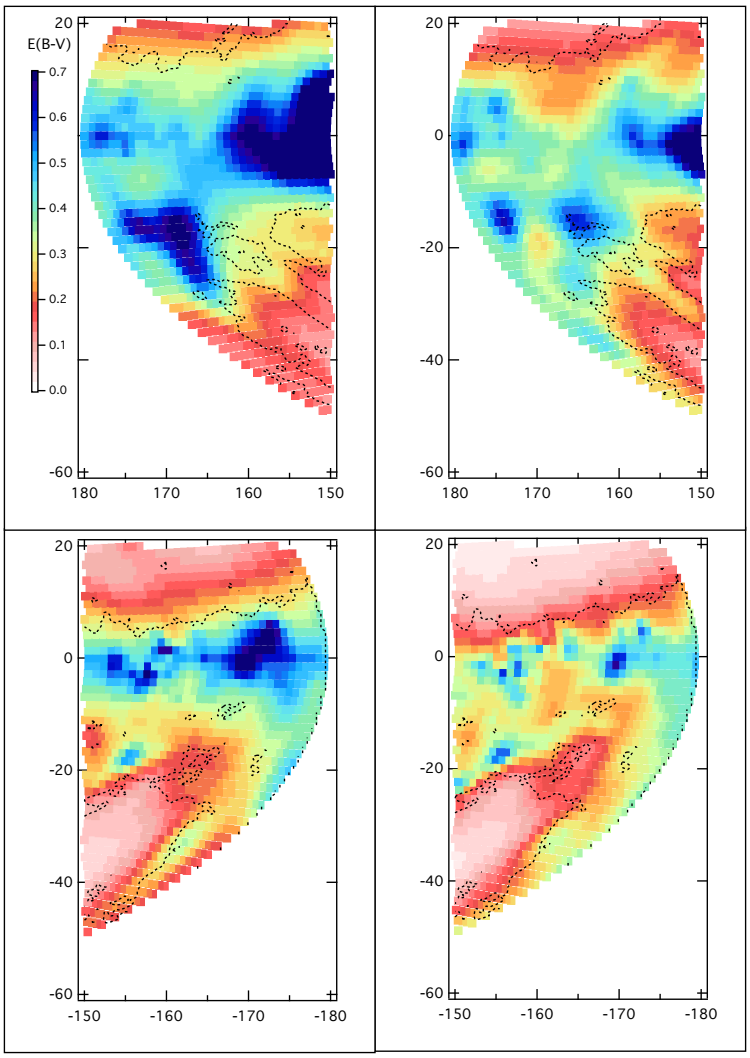


Fig. 8: Same as Fig. 7 for limited regions in the anticenter area. Left: LVV data are shown. Right: final inversion is shown. Isocontours from Schlegel et al. (1998) for E(B-V) = 0.025 are superimposed as dashed black lines.

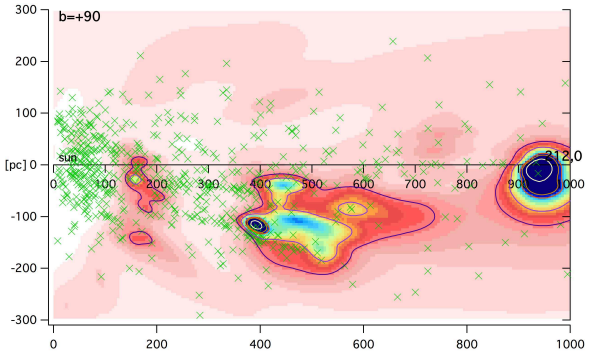
5.2. Molecular clouds catalogs

Schlafly et al. (2014) produced a catalog of molecular cloud distances based on Pan-STARRS photometric data. Most of the clouds are out of the Galactic plane, where the number of structures distributed along the sightline is relatively small (see Fig 1). These authors have mapped the clouds cores with a very high angular resolution, while our inversion produces much wider structures, however it is interesting to use this catalog for a comparison of cloud locations.

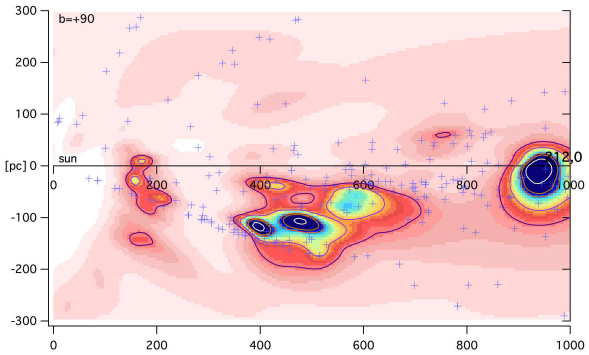
In Fig. 9, 12, and the bottom panel of Fig. 11, we superimposed the opacity distribution resulting from our D inversion and the position of the cloud cores identified by Schlafly et al. (2014) with associated error bars on their positions. For the Orion region (galactic longitude = 212°; Fig. 9), these authors found a sequence of clouds distributed from 400 pc to 600 pc, which are distances similar to the structures appearing after the inversion. In the case of Perseus (galactic longitude = 158°; Fig 12), our map also indicates the same distances. In the case of the Cepheus region (galactic longitude =110°; Fig. 11) the inversion produces two large, opaque structures above the Galactic plane, the first between 250-300 pc and the second between 700-800 pc. Two groups of high latitude molecular clouds have also been mapped in this region by Schlafly et al. (2014). The closest group is in reasonable agreement with our mapped structures, while the second is at somewhat larger distances. However, in this case our targets are very scarce at high latitude (see upper and middle panels in Fig 11) and more data are needed to allow a comparison.

5.3. Soft X-ray emitting cavities as signatures of nearby recent supernovae

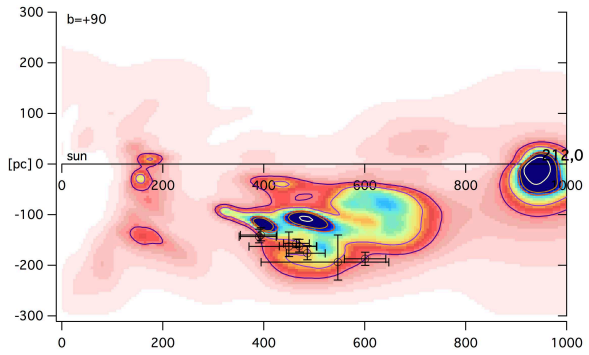
Following the work of Puspitarini et al. (2014) we searched for cavities in our 3D maps, which can be counterparts of supernova remnants (SNRs) or massive stellar winds. Such cavities are detected through their soft X-ray emission in ROSAT maps. Fig. 15 shows four vertical cuts in the 3D distribution along meridian lines crossing the four ROSAT 3/4 keV bright spots labeled 1 to 4 by Puspitarini et al. (2014). The coordinates of those X-ray spots are l,b = (-30°, +14°), (-18°, +18°), (-5°, +9°), and (-8°, -10°). In the four cases, we see that at the latitude of the bright spot centers there are cavities surrounded by dense clouds, which is the typical configuration for moderately old SNRs. Moreover, these cavities are not separated from the Sun by opaque clouds, which is a condition for their detection in soft X-rays, since million K



(a) B opacity distribution

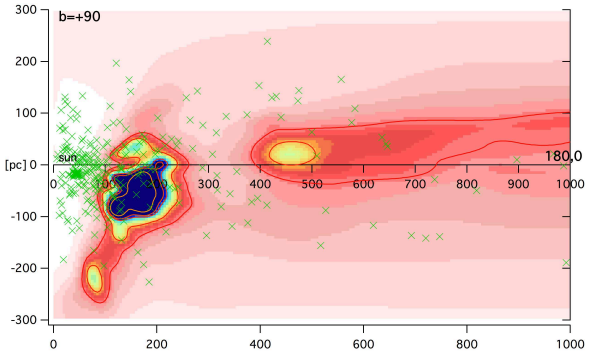


(b) C opacity distribution

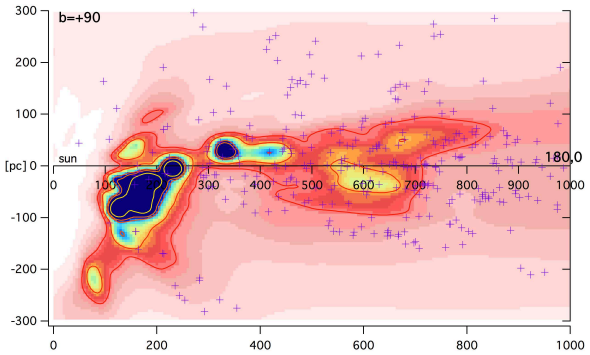


(c) D opacity distribution

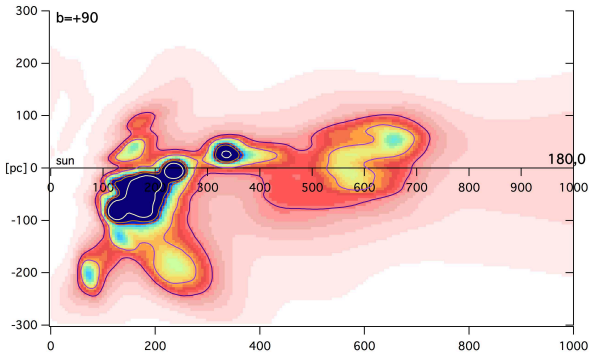
Fig. 9: Comparison between the three B-C-D distributions (see Sect. 4 and Fig. 4) for a vertical plane along l= 212°. The X-axis is within the Galactic plane, the Y-axis is toward the North Galactic pole direction. The color scale is the same as in Fig. 3. Iso-contours are superimposed for ρ_0 = = 0.0002, 0.00035, 0.00076, 0.001, and 0.002 mag.pc^{-1}. The green crosses in top panel are the positions of the targets from the LVV catalog in a slice of 10 degrees around the considered vertical plane, violet crosses in middle panel are the targets from the APOGEE catalog, and the black circles in bottom panel are the molecular clouds positions and associated uncertainties based on the catalog of Schlafly et al. (2014).



(a) B opacity distribution



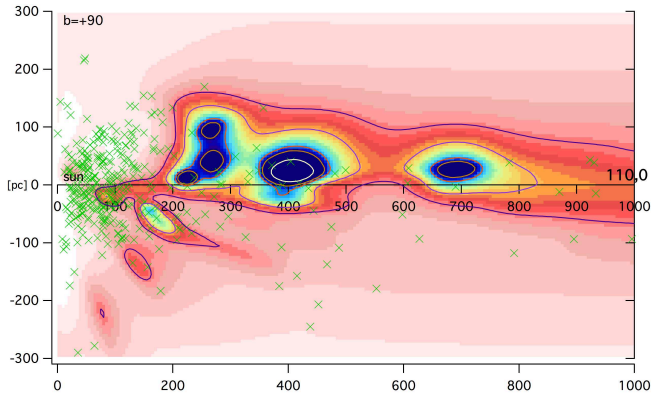
(b) C opacity distribution



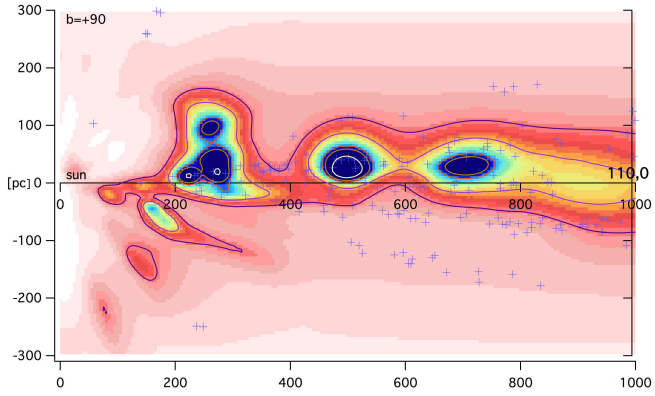
(c) D opacity distribution

Fig. 10: Same as fig 9 for the vertical plane along l= 180°(anticenter region).

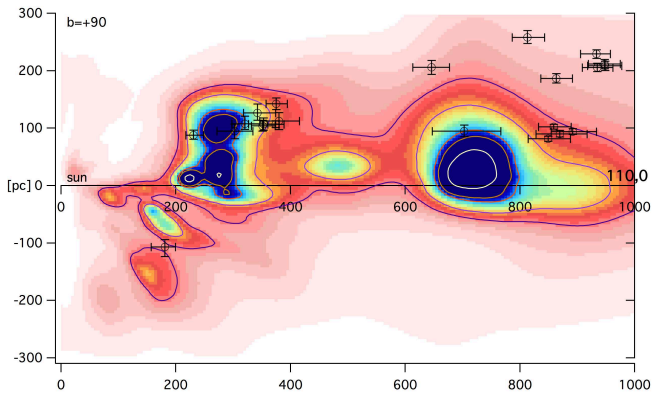
gas soft X-rays photons are easily absorbed; more precisely, the cavities must have foreground extinction that is lower than E(B-V)=0.1 to be easily detected. Compared to the previous search using LVV maps, the cavities are now better defined. Interestingly, the two cavities seen in the maps at l,b,d=(-18°,+18°,125 pc) and (-30°,+14°,125 pc) are within 3° and 7°, respectively, from the directions of the estimated sites of the two most recent SNRs (l,b = (-33°,+11°) and (-17°, +25°), Breitschwerdt et al. (2016)), and at similar distances. These authors have used the signature of ^{60}Fe in deep-sea crusts as a sign of recent (about 2.2 million years ago) explosions of supernovae in the solar neighborhood, and, based on proper motions and ages of the young stellar groups, they computed the most probable present locations of the latest explosions. Given the uncertainties associated with their statistical modeling, the small differences between our cavities 1 and 2 and their results is negligible, and these identifications reinforce their model and the ^{60}Fe measurements and interpretations.



(a) B opacity distribution

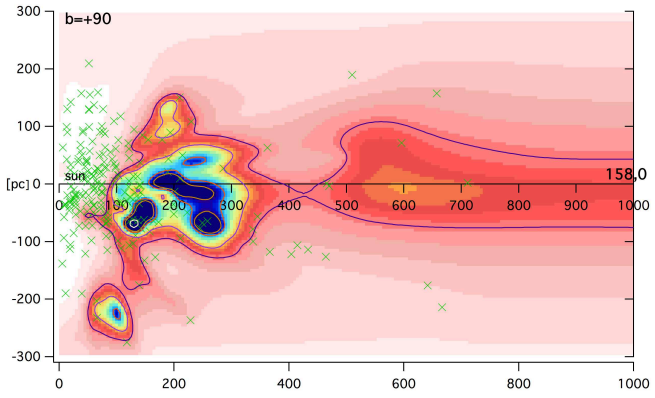


(b) C opacity distribution

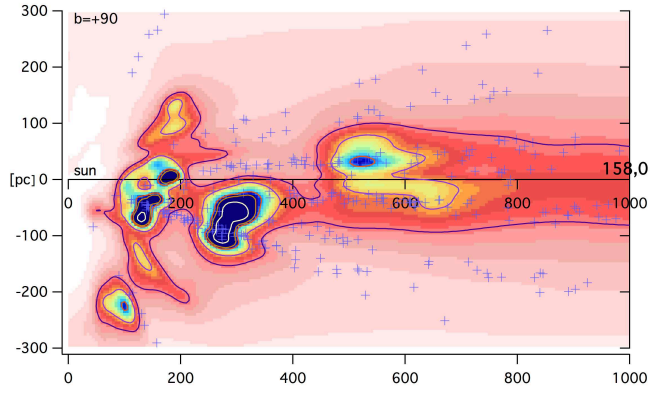


(c) D opacity distribution

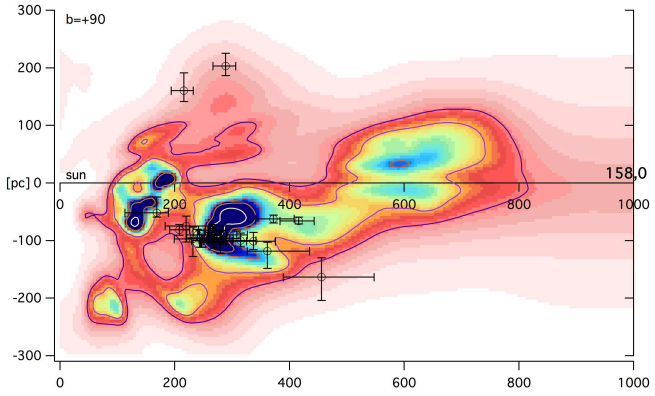
Fig. 11: Same as fig 9 for the vertical plane along l= 110°(Cepheus region).



(a) B opacity distribution



(b) C opacity distribution



(c) D opacity distribution

Fig. 12: Same as fig 9 for the vertical plane along l= 158°(Perseus region).

6. Conclusions and perspectives

We updated and improved our previous 3D maps of the local interstellar dust, based on the Bayesian inversion of photometry-based color excess measurements, in several ways. Our goal is to analyze the influence of the use of new Gaia parallaxes and to test two additional techniques aimed at providing more accurate 3D maps. First, the merging of photometric color excesses with color excesses estimated from diffuse interstellar band absorptions and, second, the use of a non-homogeneous large-scale prior distribution deduced from massive photometric surveys. The first technique increases the number of sightlines available for the inversion and, in turn, the resolution. The second work aims at testing hierarchical techniques to be used for Gaia photometric data in the future .

– We replaced ≃ 5,100 Hipparcos and ≃ 17,400 photometric distance measurements with ≃ 18,700 new Gaia-DR1 available parallactic distances and ≃ only 4,700 remaining photometric distances. In the resulting maps some elongated structures have been suppressed, which demonstrates that such features were artefacts owing to the lower accuracy of for-

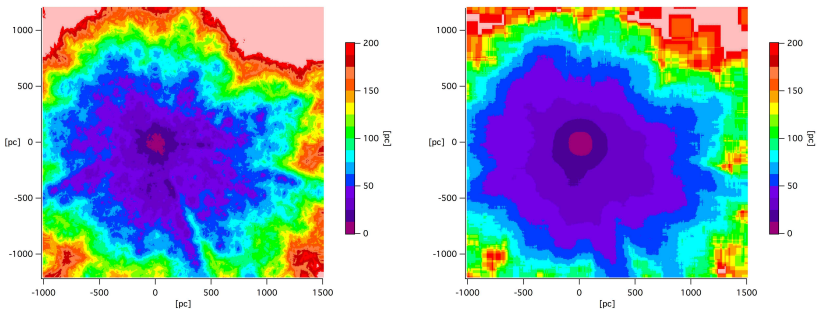


Fig. 13: Error in the distance due to the limited target space density, in the Galactic plane, as estimated from the radii of the smallest spheres that contain at least 10 targets (left figure) or from the target density in a cube of side 200 pc (right). The pale pink color corresponds to errors larger than 200 pc.

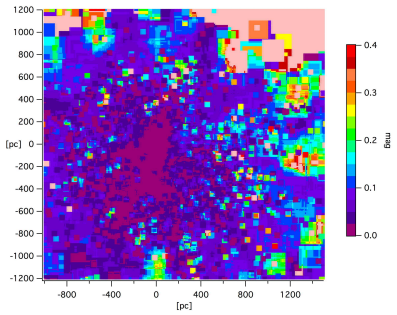


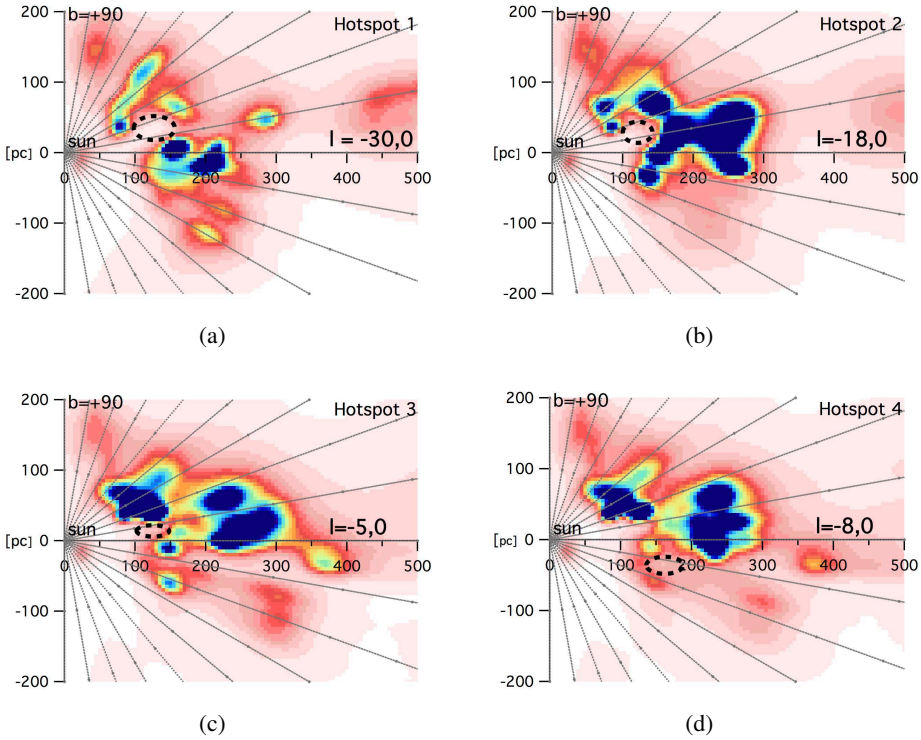
Fig. 14: Estimated error on the integrated color excess for the Galactic plane computed as the mean difference between data and model in a restricted volume (see Sect. 5.1). The pale pink color corresponds to errors larger than 0.4 mag.

mer distances. Significant changes are also seen in specific areas, in particular the Cygnus Rift.

– We extracted by means of a dedicated profile-fitting technique a series of ≃ 5,000 equivalent widths of 15273 Å diffuse interstellar bands from SDSS/APOGEE high-resolution spectra of targets possessing a Gaia parallax and converted their equivalent widths into E(B-V) color excesses. We allowed for an additional 50% uncertainty to take into account the absence of a strict proportionality between the DIB EW and the color excess E(B-V). We then performed a new inversion using the combination of the previous reddening data and the DIB-based reddening estimates. In general the APOGEE targets are more distant than the previous stars, which results in more extended maps. We do not find any contradiction between previous and new maps in regions where both types of targets are present in similar numbers. Instead, a number of structures that were broad and had low opacities are assigned more complex shapes and denser cores, which suggests that the DIB-based color excess estimates can enter the Bayesian inversion and be used to get better constraints on the cloud locations. This has important consequences because all spectroscopic surveys provide DIB measurements that may help locate the interstellar clouds, especially if all targets have precise parallax measurements.
– We used a subsample of Green et al. (2015) Pan-STARRS reddening estimates to build a low-resolution 3D map by means of a new hierarchical Bayesian interpolation method. This map is used as a prior during the Bayesian inversion of the individual data instead of a homogeneous distribution. We see that such a prior complements well the low distance map by adding spatially averaged large structures farther away. This means that data from statistical methods applied to massive photometric surveys and also future Gaia data may be used in conjunction with the databases of individual sightline data.

The resulting 3D maps are significantly improved with respect to the previous maps thanks to the use of Gaia parallaxes, the additional DIB database, and the use of a prior. This is demonstrated by the residuals, and especially by the increased compactness of the structures and the disappearance or decrease of some of the elongated artifacts due to contradictory distances (the *fingers of god*) or lack of targets. Many of these artifacts remain, however, and are expected to disappear when all targets are assigned an accurate distance. In addition, the comparisons with other cloud location assignments and X-ray emitting cavities are encouraging. This shows that future Gaia parallax catalogs should considerably help the 3D mapping by allowing the use of a much larger number of targets. Merging photometric and DIB-based color excess estimates does not reveal any contradictory information and instead improves the maps, especially beyond foreground dense clouds. This opens the perspective of using the massive amounts of DIB data that are expected to be provided from high-resolution spectroscopic surveys. Finally, the use of a Pan-STARRS based prior demonstrates the feasibility of hierarchical methods for the inversion based on different datasets or massive amounts of data.

Our distributions are particularly useful at short distances, as they show how cloud complexes and cavities are spatially organized. In this respect, they are also appropriate tools for comparing with multiwavelength emission data, understanding environments of specific objects, computing particle or photon propagation modeled distributions, and searching for unreddened areas in case of stellar photometric model calibrations. In terms of angular resolution, they cannot compete with the distributions based on massive surveys. Especially, owing to the imposed cloud minimum size, opacities, and integrated reddenings



(a) (b) (c) (d)

Fig. 15: Four vertical planar cuts in the 3D opacity maps in planes containing the ROSAT 3/4 keV X-ray bright spots labeled 1 to 4 in Puspitarini et al. (2014) and here as a) to d); the black dashed ellipses show the likely locations of the corresponding SNRs; the latitudinal grid is every 10 deg.

toward the cores of the dense clouds can be strongly underestimated. In such nearby, opaque regions a valid approach may be the combination of both types of information, from high angular resolution maps, on the one hand, and from our 3D structure, on the other.

Vertical planar cuts in the 3D distribution are given in Appendix at longitude intervals of 30°. Our final map is available at stilism.obspm.fr in its entirety and online tools allow one to obtain reddening curves estimated through integration in the 3D opacity distribution along with rough estimates of the uncertainties on distances of the structures and reddening. We plan to update the maps and improve online tools when additional data are included in the inversion.

Acknowledgements. We deeply thank our referee Katia Ferriére for careful reading of the manuscript and numerous constructive comments. Her corrections and suggestions resulted in significant improvement of the article. L.C. acknowledges doctoral grant funding from the Centre National d'Etudes Spatiales (CNES). R.L. and A.M.-I. acknowledge support from "Agence Nationale de la Recherche" through the STILISM project (ANR-12-BS05-0016-02) and the CNRS PCMI national program. M.E. acknowledges funding from the "Region Ile-de-France" through the DIM-ACAV project. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France.

References

Aihara, H., Allende Prieto, C., An, D., et al. 2011, ApJS, 193, 29
Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12
Arenou, F., Luri, X., Babusiaux, C., et al. 2017, A&A, 599, A50
Breitschwerdt, D., Feige, J., Schulreich, M. M., et al. 2016, Nature, 532, 73
Cramer, N. 1999, New A Rev., 43, 343
Dias, W. S., Alessi, B. S., Moitinho, A., & Lepine, J. R. D. 2012, VizieR Online Data Catalog, 1
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72
Elyajouri, M., Lallement, R., Monreal-Ibero, A., Capitanio, L., & Cox, N. L. J. 2017, A&A, 600, A129
Elyajouri, M., Monreal-Ibero, A., Remy, Q., & Lallement, R. 2016, ApJS, 225, 19
Green, G. M., Schlafly, E. F., Finkbeiner, D. P., et al. 2015, ApJ, 810, 25
Kos, J., Zwitter, T., Wyse, R., et al. 2014, Science, 345, 791
Lallement, R. 2015, in Journal of Physics Conference Series, Vol. 577, Journal of Physics Conference Series, 012016
Lallement, R., Vergely, J.-L., Valette, B., et al. 2014, A&A, 561, A91
Lan, T.-W., Ménard, B., & Zhu, G. 2015, MNRAS, 452, 3629
Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., & Picaud, S. 2006, A&A, 453, 635
Monreal-Ibero, A. & Lallement, R. 2015, in Memorie della Societa Italiana Astronomica, Vol. 86, Geeting ready for Gaia:3D structure of the ISM, 515–645
Nordström, B., Mayor, M., Andersen, J., et al. 2004, A&A, 418, 989
Perrot, C. A. & Grenier, I. A. 2003, A&A, 404, 519
Puspitarini, L., Lallement, R., Babusiaux, C., et al. 2015, A&A, 573, A35
Puspitarini, L., Lallement, R., Vergely, J.-L., & Snowden, S. L. 2014, A&A, 566, A13
Rezaei Kh., S., Bailer-Jones, C. A. L., Hanson, R. J., & Fouesneau, M. 2016, ArXiv e-prints [arXiv:1609.08917]
Sale, S. E. & Magorrian, J. 2014a, MNRAS, 445, 256
Sale, S. E. & Magorrian, J. 2014b, MNRAS, 445, 256
Savage, B. D. & Mathis, J. S. 1979, ARA&A, 17, 73
Schlafly, E. F., Green, G., Finkbeiner, D. P., et al. 2014, ApJ, 786, 29
Schlafly, E. F., Green, G., Finkbeiner, D. P., et al. 2015, ApJ, 799, 116
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schultheis, M., Chen, B. Q., Jiang, B. W., et al. 2014, A&A, 566, A120
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Tarantola, A. & Valette, B. 1982, Reviews of Geophysics and Space Physics, 20, 219
van Loon, J. T., Bailey, M., Tatton, B. L., et al. 2013, A&A, 550, A108
Vergely, J.-L., Freire Ferrero, R., Siebert, A., & Valette, B. 2001, A&A, 366, 1016
Vergely, J.-L., Valette, B., Lallement, R., & Raimond, S. 2010, A&A, 518, A31
Welsh, B. Y., Lallement, R., Vergely, J.-L., & Raimond, S. 2010, A&A, 510, A54
Wilson, J. C., Hearty, F., Skrutskie, M. F., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1
Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, AJ, 146, 81
Zasowski, G., Ménard, B., Bizyaev, D., et al. 2015, ApJ, 798, 35

7. Appendix

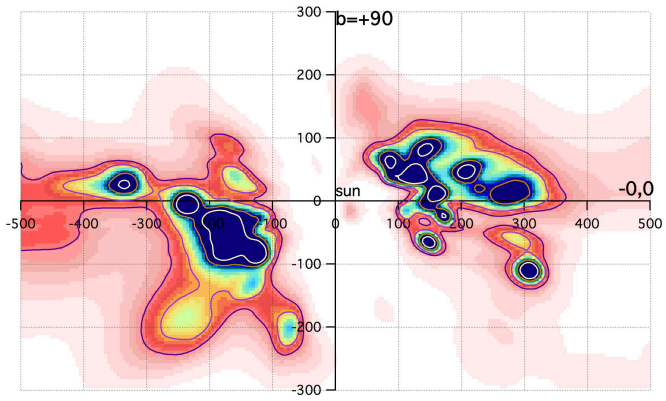


Fig. 16: Vertical cut in the 3D opacity distribution for D inversion along the meridian plane l=0-180°. Color scale and iso-contours as in Fig 9.

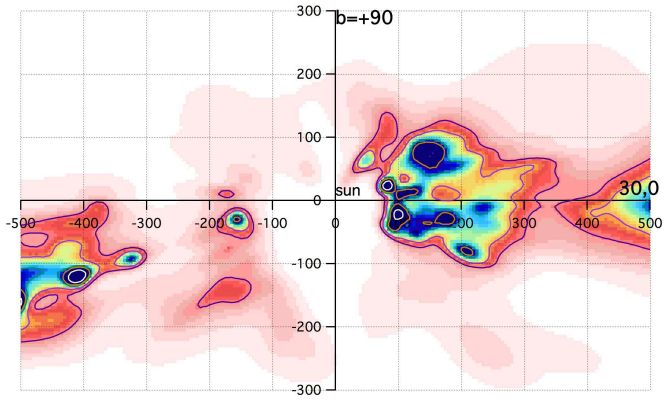


Fig. 17: Same as Fig. 16 for the vertical plane along l=30-210°

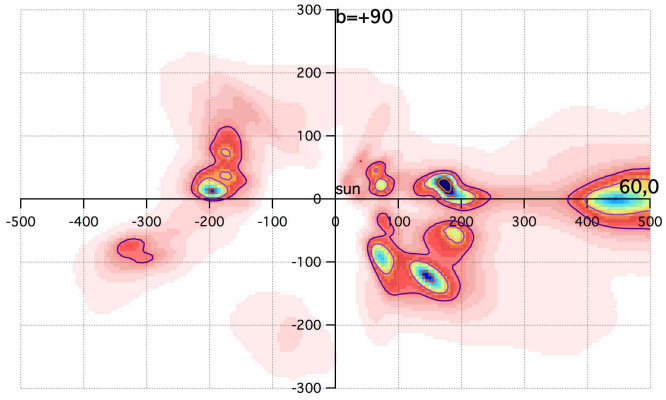


Fig. 18: Same as Fig. 16 for the vertical plane l=60-240°

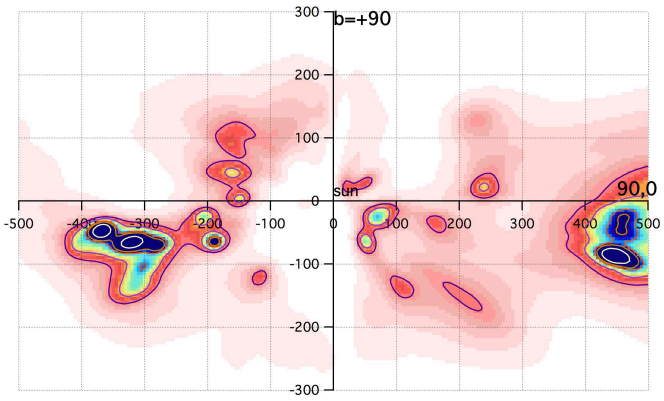


Fig. 19: Same as Fig. 16 for the vertical plane l=90-270°(the rotation plane)

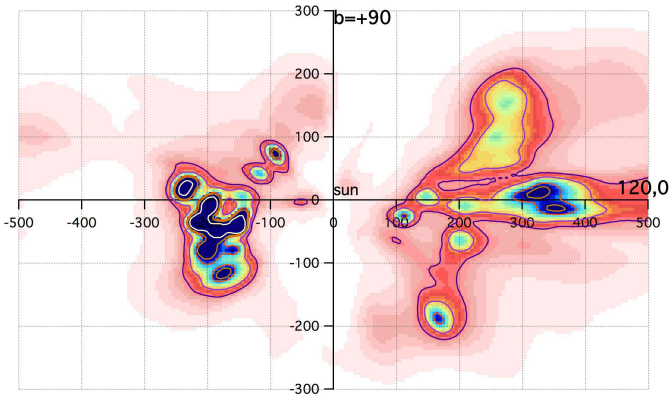


Fig. 20: Same as Fig. 16 for the vertical plane l=120-300°.

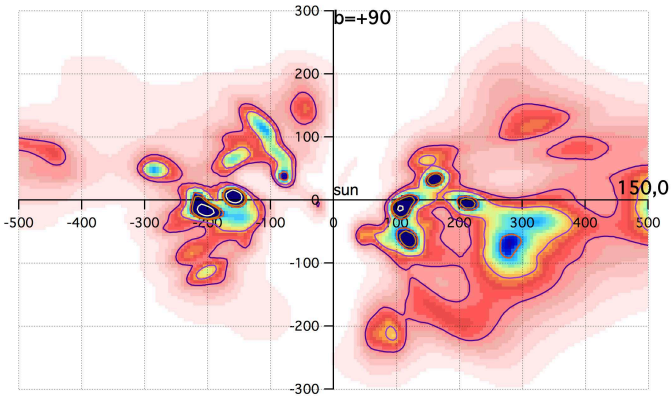


Fig. 21: Same as Fig. 16 for the vertical plane l=150-330°.

Part IV

Mapping improvements: building and testing a full 3D hierarchical method

Chapter 11

Principle of the hierarchical method

11.1 Theory

As already mentioned, the distribution of targets available for 3D inversion is far from ideally homogeneous due to several limitations and biases, essentially observational constraints inherent to each survey and to stellar types, and, most importantly, magnitude limits. As a consequence of limitations on magnitude, 3D ISM mapping is increasingly under-constrained at increasing distances, and, in turn, a given covariance function can not be appropriate simultaneously at small and large distances: e.g., a distance of correlation of, say, $\sim$20 pc may be convenient close to the Sun where bright targets are numerous and distant from each other by a distance of this order, but may become totally inappropriate 2 or 3 kpc from the Sun if distances between target stars become significantly larger because only the brightest ones can be observed, as it is the case for the surveys we are currently using. Reciprocally, a distance of correlation on the order of 100 pc will not correspond to an optimal use of the nearby targets and will smooth the nearby structures. The hierarchical method presented below is a possible approach to solve this problem, while keeping the full-3D approach.

The basic idea of the hierarchical method is to perform inversions in an iterative way, and to decrease progressively both the distance of correlation (or structure size) and the volumes containing the targets to be used in the inversion. In the course of this process, at each step with a distance of correlation ξ, all stars located in regions where the mean distance between them is larger than ξ are eliminated from the input target list and as 3D prior, the 3D the result of the previous inversion, i.e. the one performed for a larger ξ and larger volume. Doing so, constraints from the eliminated targets are used only indirectly by means of this prior distribution. In other words, little by little data, which do not add a useful constraint for the attempted structures' dimensions, are deleted.

In parallel, a second important change has been introduced in the frame of the hierarchical method, namely the use of metadata, instead of individual targets. Metadata are obtained by averaging extinctions towards targets contained within a given fractional volume (cell). The size of the cell at each step will be governed by the correlation length used for this step, i.e., first iterations

will use large cells and last iterations small cells, with sizes ultimately converging to the mean distances between targets in the Sun vicinity. An enormous gain is associated to the use of metadata, in the sense that it will considerably reduce the number of lines of sight at each iteration and consequently the computation time. The first iterations will be performed over a large global volume but a small number of large averaging cells, and the latest will use smaller cells but will be done only over a small volume close to the Sun. Such a reduction of the computational time is mandatory at the time of massive surveys and especially Gaia data.

Fig 11.1 shows the target stars within 20 pc from the Plane: the target stars circled in the top-left corner are isolated, and obviously they cannot provide any useful constraint for structures of size 15 pc^3, as required by the inversion program for a kernel parameter $\xi = 15$ pc. The circled targets are only an example, it can be seen that a large fraction of the targets represented (mostly those more than 250 pc away) cannot constrain such small structures.

In this case, we will start inverting only large structures of, say, $\sim$500 pc characteristic size, and, step by step, we will use the previous inversion results as prior distribution for the following inversion and eliminate the targets in distant regions of space. It means, for a m-step inversion:

- First inversion for structures of size ξ_0 pc using an analytical prior (or any 3D prior with structures larger than ξ_0)
- For i=1,(m-1)
  ...
  Inversion for structures of size ξ_i pc, with $\xi_i < \xi_{i-1}$
  ...
- Final inversion for the smallest scales ξ_m pc

The core of the inversion program itself is not different from the one used for the direct method, but for the hierarchical method, we change the input data, the parameters and the 3D prior at each step. At the end, we get a 3D map with a spatially-variable resolution.

Following this procedure, the input dataset that is used at each step is produced from the same initial catalog but with a specific distribution, and it is important to construct metadata with a well-determined information. In case of structure size ξ, we need data distributed within a volume of size at least $\xi/2$, we need a number of targets sufficient in the cell to obtain a significant average, and we need realistic errors associated with the average distance to the cell and with the averaged extinction in the cell.

11.2 Implementation of the hierarchical method

Preparation of intermediate datasets (metadata)

From the total dataset of distance-excess of colour pairs we now have for each inversion step to build a data subset adapted to the chosen covariance length, with all cells containing a minimum number of targets (e.g. 12 in the first example below). These metadata are based on averages of actual data, and their use implies new estimates of the associated uncertainties. As we will see this is the most complex aspect of the whole procedure.

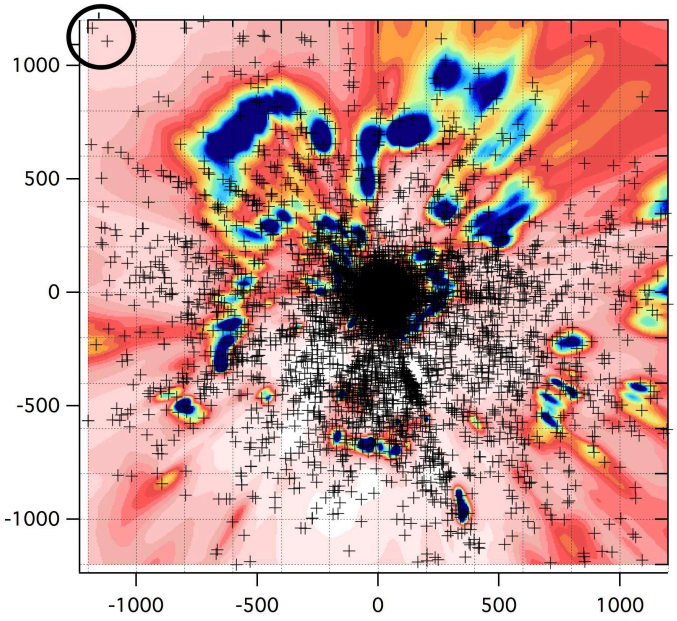


Figure 11.1 – Galactic plane resulting from inversion D of Capitanio et al. (2017): input targets within 20 pc from the Plane are superimposed (black plus signs). Targets like those circled at top left will be eliminated along the hierarchical process, since the target density at their location is too small to contribute to the definition of small structures. They will be taken into account at the first steps only. Colour scale is like 10.9 - panel d)

The problem of uncertainty estimates for metadata We start with a given cell containing N targets. We replace the N targets by a metastar at an average location l_C, b_C, D_C (generally close to the coordinates and distance of the cell barycentre) and characterised by an average colour excess excess of colour $_C$ computed as the mean of the N excess of colour values. It comes to mind to use the dispersion σ(excess of colour) of excess of colour values among targets within the cell as the uncertainty associated with the metastar mean colour excess. However this is not correct and overestimates the uncertainty. This comes from the fact that in the same way the colour excess excess of colour for any sightline to a target contained in the cell is built over the whole path, i.e. including foreground cells, the information contained in the dispersion corresponds to the entire pathlengths from 0 to D. For this reason, the dispersion on the reddening measurements is the sum of the dispersion generated in the cell and the dispersion generated in front of the cell. As a result, in order to avoid overestimating the errors, we cannot simply use the dispersion among the data σ(excess of colour), instead we must subtract from this dispersion the variance in the foreground cell that is crossed by the N sightlines. This is the source of complexity, since the variance computed in this way does not only depend on measurement errors, but also on measurements in the region

in front of the cell. The general formulation is:

$$\sigma = \sqrt{{\sigma_C}^2 - {\sigma_{FC}}^2 + \frac{{\sigma_C}^2}{Nobjectincell}} \tag{11.1}$$

where C (cell) refers to the considered cell and FC (ForeGround) refers to the cell just in front of C. N_C is the number of objects in the cell. The first (and usually dominant) term is linked to the gradient of variance and the second is the simple statistical error associated with the average over the N targets. The way the uncertainty estimate is made has strong consequences in terms of cell geometries. As a matter of fact, at each step and for each cell one must use the colour excess dispersion in a front cell that is crossed by the same sightlines. As we will see, it implies two geometries that are used in parallel, the geometry of the cells used for the inversion, unique (we we call it the inversion geometry), and the *geometries* computed at each step for the error estimates only and that vary all along the computation (we will call them the intermediate geometry).

Cells shape An important choice in the building of metadata (or metastars) is the choice of the cell shapes for the inversion geometry. The main goal is to obtain metadata cells as regularly spaced as possible, and with similar volumes (for a given step), however the structure must be also compatible with the computation scheme. A simple choice could be cubic cells, since most computations within the inversion use cartesian coordinates, and it would to be easy to define cells of identical volumes and to compute the mean distance between our meta-stars. However it becomes impractical when we come to the point of calculating the errors as discussed above. Another choice could be to simply divide the space in cells defined by constant intervals in distance, longitude, and latitude. This choice is perfect to calculate errors, because all lines of sight to the stars in a given cell cross the same foreground cell, but evidently having constant distance, longitude and latitude intervals implies increasing distances along the ortho-radial direction and increasing volumes of the cells with increasing distance, which is opposite to our main goal. Keeping constant intervals in longitude and latitude and decreasing the radial dimension to keep a constant volume is also a unappropriated solution since the cells will become increasingly flattened.

At the end, I chose a solution that is based on constant radial distance intervals and variable longitude and latitude intervals in such a way the volume is roughly conserved in all cells. The geometry is shown in Fig 11.2 and in the left part of Fig. 11.3 for the first, second, and third ranks around the Sun. As discussed in the previous section, such a geometry does not allow to estimate the variance in the cell, and and additional decomposition in cells with a different and adapted geometry must be recomputed and used at each step.

For the inversion geometry I start by selecting the mean distance U that I want between the meta-stars, and divides the 3D space in series of spherical shells, with widths U. If the maximum accepted distance is DMAX I divide DMAX by U and retain the integer Ns the closest to the ratio. Ns is the number of spherical shells. Then, each shell is divided in angular bins in longitude and latitude in order to obtain cells with volumes as close as possible to U^3. This is obtained by dividing the perimeter of the shell in sections of equal length l as close as possible to U (again, since the number of cells is an integer, one can

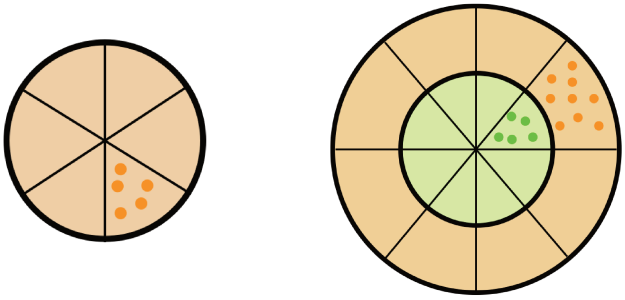


Figure 11.2 – Illustration in 2D of the cell geometry for the inner cells (left) and the first rank.

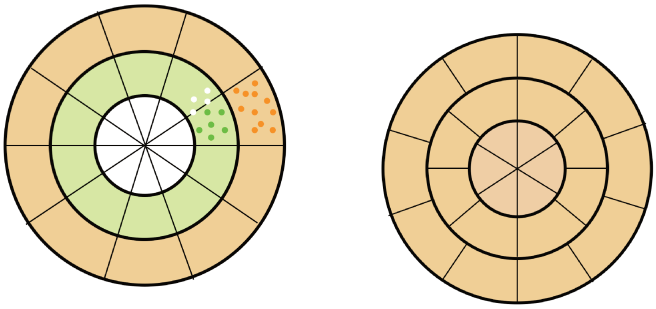


Figure 11.3 – Illustration in 2D of the two different cell geometries in case of the third rank. (Left): the geometry used for the error estimate at the third rank stage. In this case radial boundaries are the same for current (third rank) and foreground (second rank) cells. I.e. light paths to stars in a cell are all crossing the foreground one, a condition required for estimating the variance (see text) (Right): geometry of the three first ranks of cells actually used for the inversion. The volumes of the cells are approximately constant but the conditions for error estimates are not fulfilled.

not use a simple mathematical division but choose the solution for the closest integer).

Fig 11.2 illustrates the cell geometries for the first and second ranks. Fig 11.3 illustrates the geometry of the cells used for the inversion in case of the third rank, and the associated, and different, cell geometry that is necessary to estimate the uncertainties (note the differences between the two panels of Fig 11.3). As a matter of fact, as mentioned above, since the foreground variance has to be subtracted, it is important, at each step, to use a foreground cell with a completely corresponding optical path. Therefore the adapted geometry is re-computed at each step and used only for the variance.

Metadata calculation in the cell After having chosen the cell shape [l, b, d], for each cell C the selection code extracts from the initial database the stars contained within the cell and simultaneously it extracts the foreground cell FC, using the same constraints for angular position but using a distance smaller by the quantity U. Note that, doing so, we do not use any uncertainty on target longitudes, latitudes, and distances. The nest step is to calculate for the subsample the mean distance and its associated error, the mean excess of colour and the variance on excess of colour. The following additional constraints are imposed:

- Every cell retained for the inversion must contain a minimum number of objects, sufficient to get obtain a statistically significant average.
- To be part of the inversion at a given step, a cell in a specific direction must have cells filled with targets in the same direction and closer to the Sun, i.e. ensuring a foreground without any observational *hole*

These constraints reduce significantly the useful volume in case of inhomogeneous databases. The reason for the second requirement is not only the need for homogeneous constraints, but the need to calculate errors associated to the metastar excess of colour, as explained above. The way these constraints are implemented is as follows (for an imposed minimum number of targets N in one cell): if the foreground cell FC does not contain N stars, then the cell C is removed from the inversion. This is not an optimal implementation of the second constraint because I could test at this step that each foreground cell contains the minimum stars number before accepting the background cell, but since most selection biases are functions of distance, I assume that if there are enough targets at a given distance, this is also true at smaller distances.

Final selection A final selection takes in account all the calculated metadata and verifies that each metadata has at least one other metadata closer than U. This selection is important because it suppresses possible far isolated cells. In Fig 11.4 is illustrated why it is possible to have isolated cells; if there is a large concentration of stars located after a hole, it can provide some calculated metadata, because the foreground cell is not empty. As measured are limited in distances in the observational survey, these cases are quite rare: the stellar catalogs usually contains number of objects that decrease with the distances.

11.3 Example of metadata preparation and hierarchical inversion

In what follows is presented one of the several tests I did to guide and test the metadata calculations and the choices of the inversion parameters.

I used the large dataset of extinction-distance pairs estimated based on spectroscopic observations with the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) in China. The data I used are from the data release DR4 (Cui et al., 2012). The distance-extinction pairs are based on the work of Wang et al, 2016. I selected targets stars located in the interesting region of the nearby Taurus clouds. From the initial LAMOST published dataset I extracted 29 448 objects within the interval l = (155°, 180°) and b = (-30°,

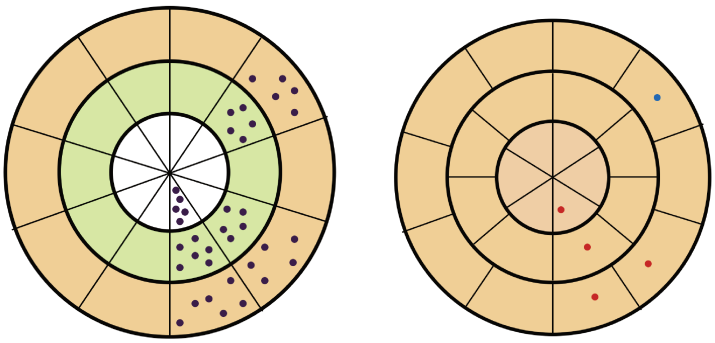


Figure 11.4 – Example for the last selection in metadata. Left) Black rounds are indicative stellar position in the original catalog, with the geometry geometry used for the error estimate at the third rank stage Right) Rounds are the calculated metadata, located in different cells at different rank. The blue metadata is an isolated point and it is rejected thanks to the last selection.

-5°) and with relative errors on the extinction less than 30%. This first hierarchical inversion was performed in 5 steps. The mean sizes of the metadata cells were 100, 80, 50, 35, 25 pc for the 5 steps respectively.

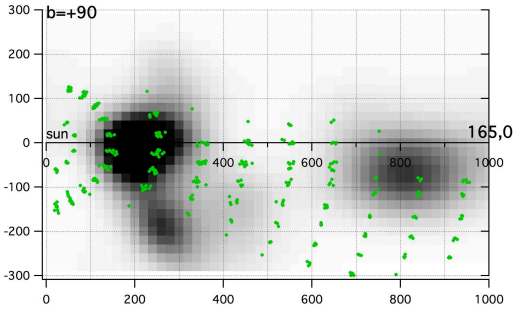
At each iteration correspond different parameters for the inversion. It is important to note that, if the cells are restricted to the above Taurus area coordinates, close to the Sun they contain too few targets and one gets border effects and unreliable results. For this reason, it is necessary to add targets taken in a larger volume, i.e. to extend to larger intervals in longitude and latitude: for targets closer than 200 pc we extended the latitude and longitude interval by 25 degrees on each side, and for targets closer than 100 pc the extension by 50 degrees.

A summary of the metadata and inversion choices are in Tab 11.1. In step 5 I decreased the minimum number of targets per cell in order to avoid a too large number of empty cells. For this test I did an unweighted colour excess average. I used as prior distribution the Pan-STARRS prior, as in Capitanio et al. (2017). The evolution of the hierarchical inversion and of the metadata can be seen in the Figures 11.5, 11.6, 11.7, 11.8 and 11.9 . The size of the voxels used for the inverted distribution at each step is a compromise: it is not useful to use a size much smaller than the achieved resolution so we choose the maximum size at each step to save computation time.

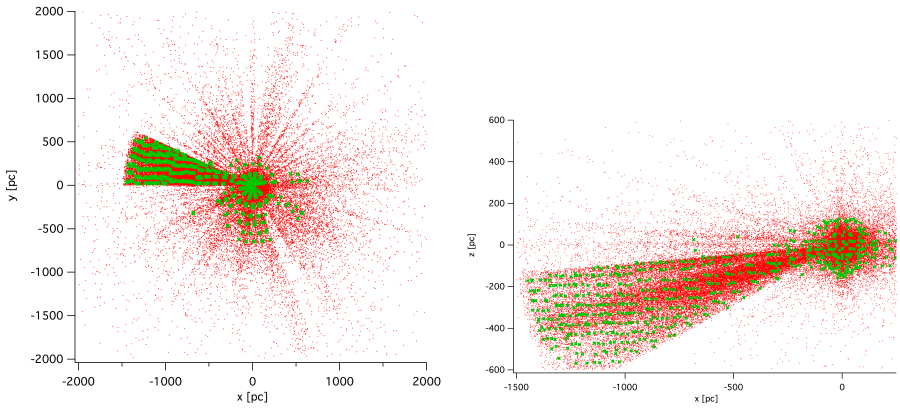
The figures illustrate the progressive decrease of metadata at large distance from the Sun, this is due to the constraints given above that are very strict. While the first step - u = 100 pc - has metadata until 1500 pc, for the third step they do not exist beyond 1000 pc and for the fourth one, they are closer than 800 pc. One can see the decrease in the sizes of the inverted structures. In regions without metadata the inversion does not introduce any significant change, as expected. Here I show only one vertical cut. The global 3D distribution is complex and the distribution varies with the longitude in this area, as we will see in the next sections. Note that the final resolution for this test is not high, with 50 pc for the smallest kernel. We will see that the data quality does not allow to reach a much better resolution.

	Mean Distance D Metadata	Minimum number N of objects	Objects Ns	ξ_0	ξ_1	σ_{ξ_0}	σ_{ξ_1}
STEP 1	100 pc	12	666	220 pc	200 pc	0.4	0.8
STEP 2	80 pc	12	753	180 pc	160 pc	0.5	0.9
STEP 3	50 pc	12	751	110 pc	100 pc	0.6	1.0
STEP 4	35 pc	12	443	80 pc	70 pc	0.7	1.1
STEP 5	25 pc	5	686	60 pc	50 pc	0.8	1.2

Table 11.1 – Parameters used for the LAMOST-Taurus hierarchical inversion. D is the distance between cells. N is the minimum numbers of targets required to keep the cell in the inversion process. Ns is the number of cells. ξ_i and σ_i are the correlation length and variance for the two kernels (Gaussian and 1/cosh)

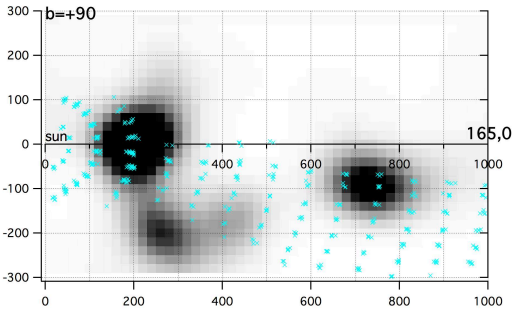


(a) step 1.

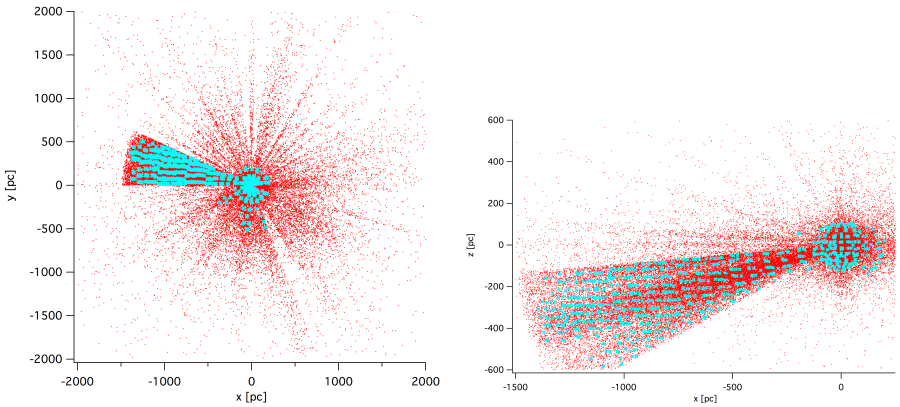


(b) Target distribution step 1.

Figure 11.5 – Hierarchical step 1: a) Vertical cut at longitude 165° in the inverted distribution after the first step, with input metadata overlapped (green crosses). The colour scale represents the differential reddening excess of colour/pc in mag.pc^{-1}, as in 10.6. b) Metadata (green crosses) and initial stellar targets (red dots) along the Galactic plane and in the vertical plane of a).

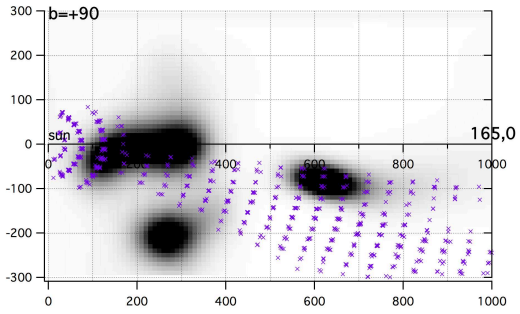


(a) step 2.

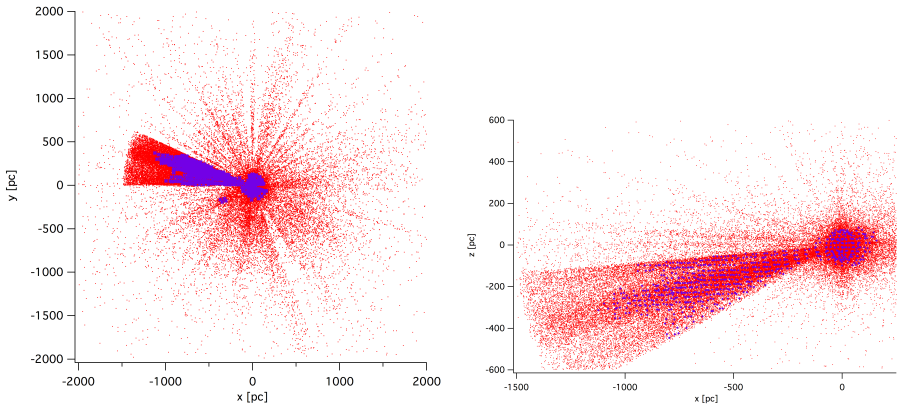


(b) Target distribution step 2.

Figure 11.6 – Same as Fig. 11.5 for the hierarchical step 2

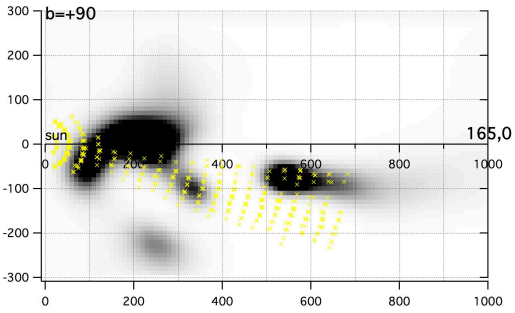


(a) step 3.

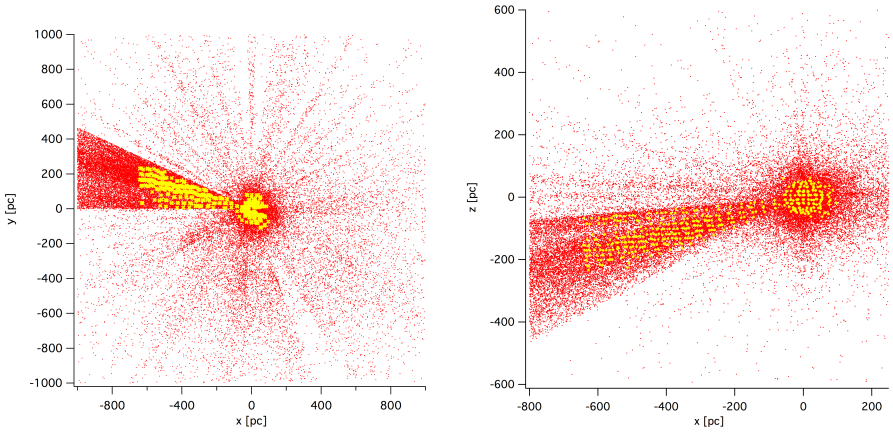


(b) Target distribution step 3.

Figure 11.7 – Same as Fig. 11.5 for the hierarchical step 3.

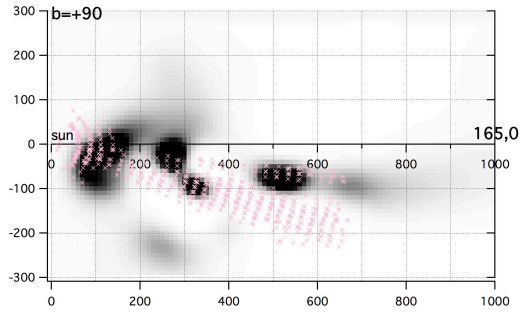


(a) step 4.

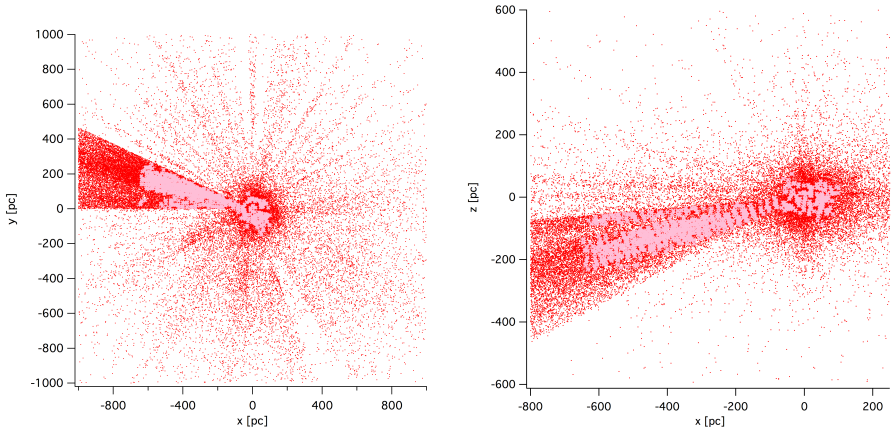


(b) Target distribution step 4.

Figure 11.8 – Same as Fig. 11.5 for the hierarchical step 4.



(a) *step 5.*



(b) *Target distribution step 5.*

Figure 11.9 – Same as Fig. 11.5 for the hierarchical step 5.

Chapter 12

Hierarchical method and data quality

As in other inversion techniques it is difficult to estimate the influence of uncertainties associated with the data on the resulting distribution, globally and at each point in space. As a matter of fact, there are complex interdependencies between the error bars, the volume density of targets and the cloud clumpiness along the sightlines. This is true also for the hierarchical technique and, since the metadata are calculated differently at each step, it is important to understand how errors on input distances and reddening values influence the inversion at each different step. On the other hand, the step-by-step evolution brings additional information and sheds additional light on these influences.

I used two distinct subsamples of the LAMOST DR4 data, the first one being the one presented in the last section (called ODD) and another set of data, completely independent and with the same number of data points (called EVEN). As their names indicate, the ODD-EVEN selection was very simply made from the initial catalog based on the line number in the file (ODD or EVEN number). Since the ordering was based on the location in the sky, the two datasets cover the same regions and correspond to random selections everywhere.

The goal here is to compare the results obtained with the two different datasets, from both direct and hierarchical inversions, and to study how the data uncertainties influence the quality of the results at the various steps of the hierarchical inversion.

12.1 Direct Inversion

I inverted the two ODD and EVEN datasets using the direct code, and made some tests. Inversions here were performed locally on our Mac server and took $\sim$ 20 days. Data with l = (152°, 183°) were retained, with no limitations in latitude other than those of the survey. The inversion parameters were σ_0 0.9, ξ_0 50.0, σ_1 1.0, ξ_1 25.0. The EVEN inversion used 29359 stellar targets and the ODD one 29448.

In the input catalog, there are not threshold for the errors, but mostly associated relative errors on distances and on excess of colour have values lower than 50%. In Fig 12.1 the histograms illustrate the relative errors distribution

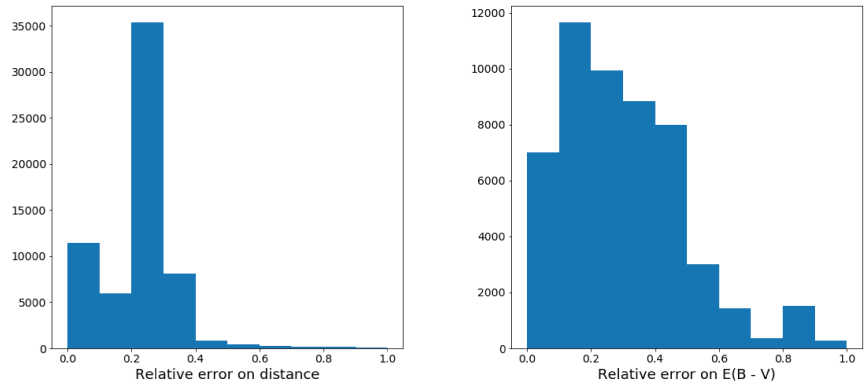


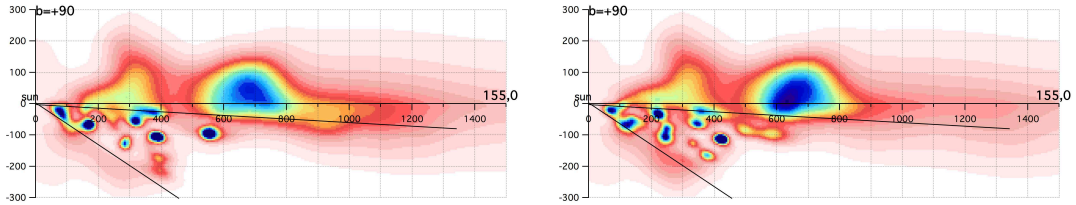
Figure 12.1 – Histogram distributions for the relative errors on distances and excess of colour for the EVEN input catalog. The ODD input catalog has similar distribution.

in the interval [0:1] for the EVEN input catalog; there are some rare outliers with high relative errors, but as their values are averaged in the cells, we decided to not remove them.

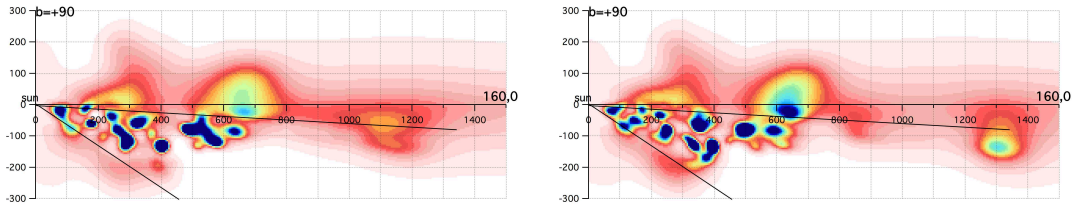
Looking at Figures 12.2 and 12.3 (remember to restrict comparisons to the area between the black lines that delimitate the latitudes of effective inversion) it is evident that we have significant differences between the ODD and EVEN maps, which may be surprising since the datasets are two separate subsamples from the same data release, with the same observational selections and biases, and with a very good data coverage in both cases. This clearly points to an inadequacy between the inversion parameters, and especially the 25 pc resolution, and the uncertainties in distances, reddenings or both, or point to under-estimated errors. In the inner part of the data, within about 250 pc from the Sun, the two inversions seem to be roughly compatible, at least in the shown vertical cuts at longitudes $l = 165°$ and $l = 170°$. When we say compatible, we mean the cloud locations, not their opacities. We tested different values of the parameters ξ and σ, and the tests show that the choices for the ξ and σ pairs made in Capitanio et al. (2017) were appropriate for this region. As a conclusion, the tests suggest that the combination of errors on distances - reddenings and the number of targets precludes the use of a kernel as small as 25 pc beyond 300 pc. This implies that direct inversion with this resolution and with this dataset are inadequate for the whole Taurus area.

12.2 Hierarchical inversion for EVEN and ODD sub sets

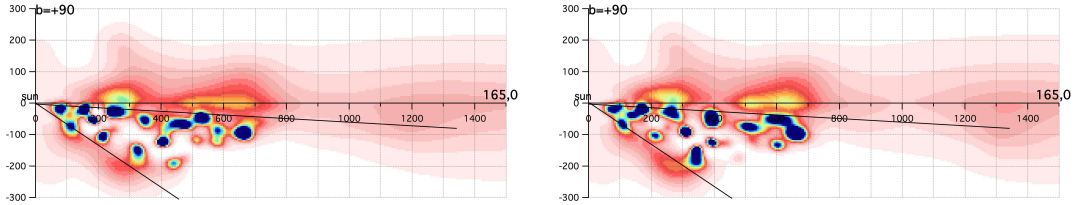
From the direct inversion it has already appeared clearly that the combination of data uncertainties/data coverage does not allow to obtain reliable cloud structures beyond 300 pc for a resolution of 25 pc. As we will see, the hierarchical inversion will be very useful here: in fact it is possible to use the successive steps of the hierarchical ODD and EVEN inversions to check for differences at



(a) $l = 155°$.



(b) $l = 160°$.



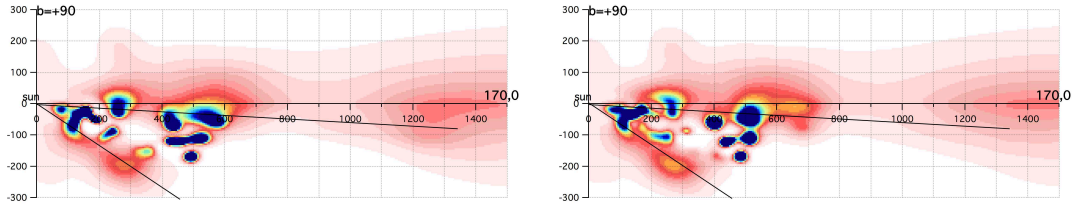
(c) $l = 165°$.

Figure 12.2 – Direct inversion: comparison between the EVEN (left) and ODD (right) databases in three different vertical planes. Colour scale as 10.9 - panel d).

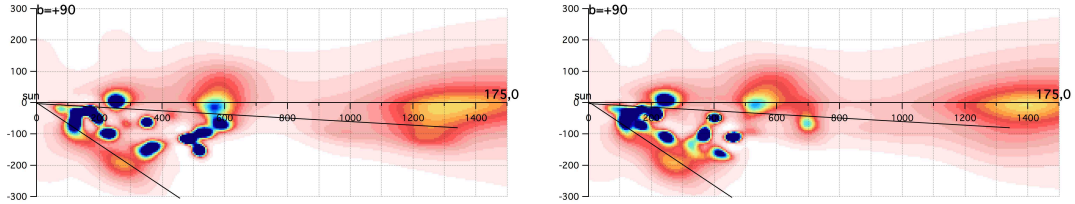
each step and understand the actual minimum size of the structures that can be inverted with this dataset at each distance.

To do so, I inverted the two EVEN and ODD datasets in parallel, using the same hierarchical technique for both. The parameters for the different steps are listed in Tab 11.1. In Fig 12.4 and 12.5 we can observe the evolution of the two distributions and infer until which step the two results remain compatible (note again that the comparison has to be done only for the region well covered by data, i.e. between latitude -35° and 0°). In particular we can see that the two separate inversions start to diverge at step 5, i.e for 60 and 50 pc kernels. The full computation of the 5 steps takes less then half an hour, i.e. it is considerably faster than the direct inversion.

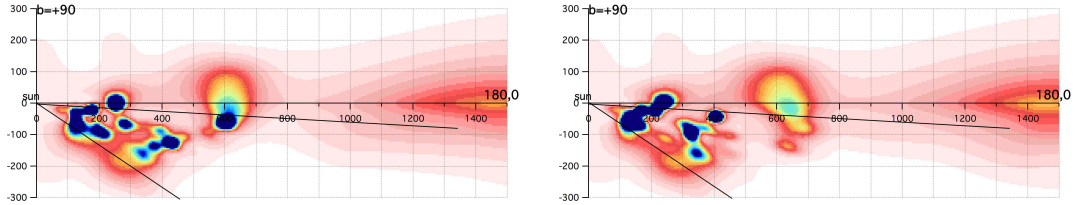
As a conclusion, the tests with the hierarchical inversion show that it will be impossible to reach a resolution much better than 50 pc with this LAMOST dataset, even if one can expect a better performance for the ODD+EVEN combined catalog. However, one advantage of the hierarchical method is that the step 5 inversion should be reliable at ALL distances, thanks to the basic principles, something impossible to guarantee with the direct method (for this latter method, there may be artefacts at large distance).



(a) l = 170°.



(b) l = 175°.

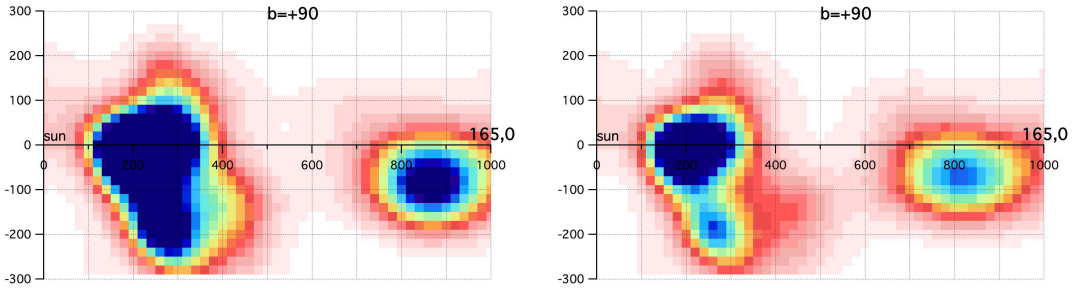


(c) l = 180°.

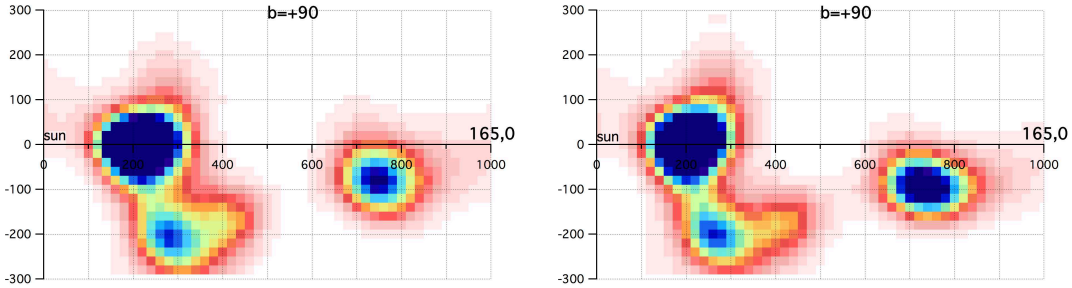
Figure 12.3 – Same as figure 12.3 for three more vertical planes.

The difficulty of Local Cavity (or Local Bubble) mapping with the LAMOST data only The maps derived from direct or hierarchical inversions with the LAMOST DR4 dataset show Taurus clouds very close to the Sun, in contradiction with previous maps that all well reveal the cavity around the Sun, the Local Bubble. We have investigated this problem and believe it is due to the absence of target stars closer than 30 pc and also to large uncertainties on the distances (estimated from photometry). In the absence of constraints at small distances the inversion is distributing the absorbing matter measured towards the distant stars all along the lines of sight, including at very small distances. To test this hypothesis we made an additional inversion, this time including target stars from our own catalog located within 150 pc in the Taurus direction.

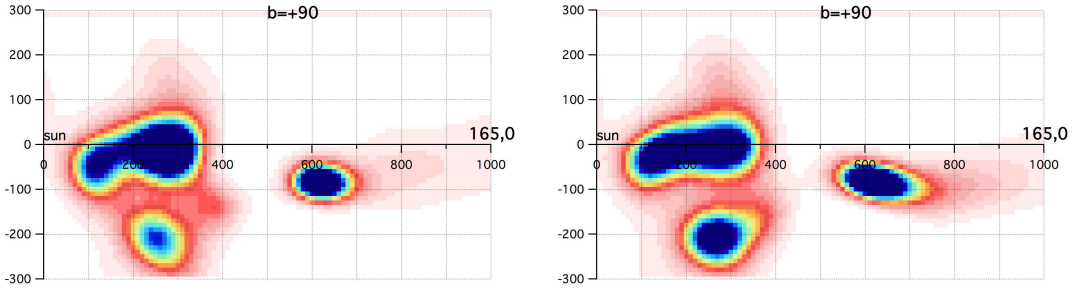
Fig 12.6 shows three vertical cuts in the 3D inverted distribution for -a) the direct inversion with LAMOST data only and for the 25 pc resolution, b) the hierarchical inversion for LAMOST data (ODD) only and for the step 5 (structures of 50 pc) and c) the direct inversion with the distribution b) as prior and both LAMOST DR4 ODD data and stars from the previous catalogue closer than 150 pc. I.e., the last inversion c) is combining the hierarchical and direct techniques. It takes 3 hours, using 9724 stellar targets. It can be seen that there is some decrease of the density close to the Sun - the Local Bubble - for the last inversion (although, due to the coarse resolution, the improvement



(a) step 1.



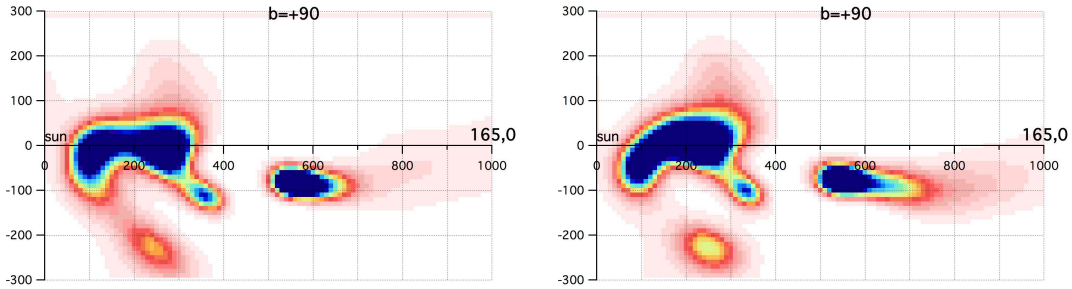
(b) step 2 .



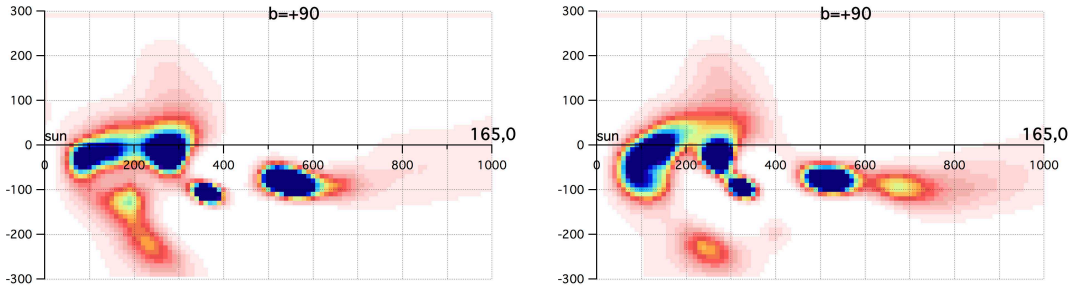
(c) step 3.

Figure 12.4 – Hierarchical inversion steps 1-2-3, for EVEN (left) and ODD (right) datasets, vertical plane at longitude 165°. Colour scale as 10.9 - panel d).

is limited and the cavity is not well defined).

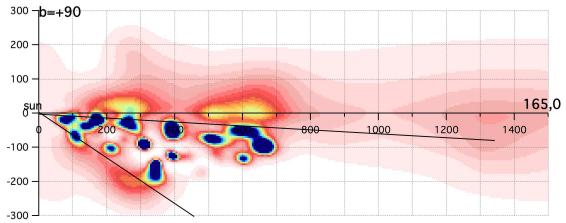


(a) *step 4.*

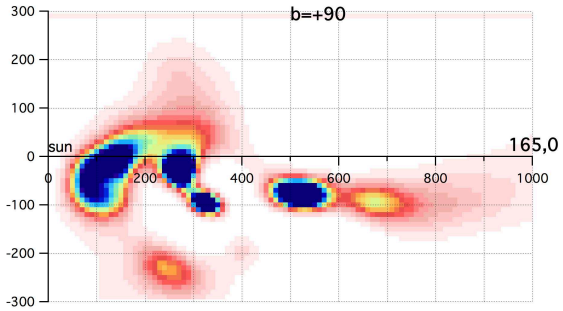


(b) *step 5.*

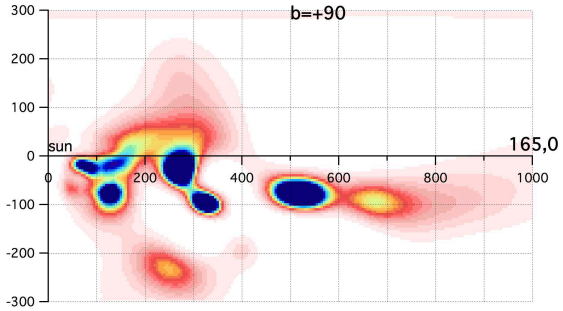
Figure 12.5 – Hierarchical inversion steps 4-5, for EVEN (left) and ODD (right) datasets, vertical plane at longitude 165°. Colour scale as 10.9 - panel d).



(a) Direct inversion, 25 pc structures ODD dataset



(b) Hierarchical inversion step 5 - ODD dataset.



(c) Direct inversion with prior from hierarchical step 5 and adding data within 150 pc to the LAMOST ODD dataset.

Figure 12.6 – Tests of direct-hierarchical combination, adding nearby stars (see text). Colour scale as 10.9 - panel d).

Chapter 13

Hierarchical method and data characteristics

In this chapter I present a hierarchical inversion based on a combination of extinctions from LAMOST DR3 (190 858 targets) and APOGEE DR14 (7 146 targets). Everything was done before the Gaia Data Release 2 (DR2), so we used either Gaia DR1 (TGAS) distances or photometric distances estimated with the Gaia DR1 calibration. We also used in addition our *old* catalog, the one used in Capitanio et al. (2017), which contains a combination of data with photometric or Gaia DR1 parallax distances. The inclusion of this catalog is important because of the numerous stellar targets at short distance and within the Local Bubble: other catalogs have only a few targets at short distance. Moreover, extinctions deduced from Bayesian statistics tend to have increased biases for small extinctions, while the Stromgren or Geneva photometric calibrations that form the majority of the *old* catalog provide accurate data for small extinctions and individual targets. I analysed again the Taurus region, well covered by the LAMOST survey, and in order to be able to compare with the other inversions.

Metadata generation

Every metadata generation must be adapted to the set of stellar data. Target distances vary from one dataset to the other, therefore limits in distances for the cells must change accordingly to adapt to each dataset distribution. Here we study a small region, because our goal is to identify a convenient strategy for the following inversions with more data and future Gaia accurate parallaxes. Another important aspect is the choice of the quality criteria requested to enter the database to be inverted. In the context of our hierarchical method, we do not use errors on the distance in the calculation of metadata, i.e. only mean distances are used to locate a star in a given cell. Therefore, if we want a minimum accuracy on the distances we must impose a criterion at the beginning during the choice of the targets to consider for the inversion, and reject data in catalogs that do not possess the requested quality. For the tests shown in this section, we did not add additional constraints on distance or extinction and considered that the catalogs were of sufficient quality. A test on LAMOST data shows that except for a few percent of the targets the relative error on the estimated distance is smaller than 30% (note that for the

posterior inversions with new LAMOST DR4 we will impose additional quality criteria). For APOGEE data, we rely on the distance estimate made by Carine Babusiaux, who made series of tests for their quality. For the *old* catalog, we also rely on the quality of the catalogs.

I select stars from the APOGEE and LAMOST catalog in the Taurus direction, using the following geometrical criteria:

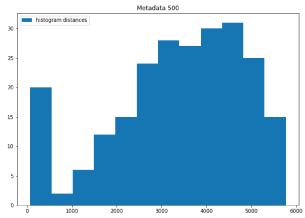
- for stars farer then 200 pc, longitudes from 150° to 180° and latitudes for -30° and -5°
- for stars between 100 and 200 pc, longitudes from 110° and 225° and latitudes for -60° and +25°
- for stars closer then 100 pc, longitudes from 65° to 270° and latitudes for -90° and +55°

The strong increase of longitude and latitude interval widths at short distance ensures a minimum number of stars around the Sun to be part of the closest cells. As a matter of fact, using pure limitations in angles reduces too strongly the allowed volumes (see, e.g. Fig. 11.8). The total number of stellar targets was 201 079 objects, and I made metadata with mean excess of colour values in cells having mean distances of 200, 100, 80, 70, 60, 50, 40, 30, 20, 15 pc. These metadata were maintained in the inversion process only if they contained at least 12 targets, excepted for metadata with cells of 15 pc, for which the limit was decreased to 5 stellar targets. In Fig 13.1 and 13.2 are shown the histograms of the distances of the metadata at the various steps. The double peaks are due the existence of changes in the stellar selection criteria as a function of distance. The second bump, at larger distances, is due to intrinsic limitations in the stellar catalog. In Fig 13.3 we can observe the metadata projected on the sky, in the Taurus region. The colour scale indicates the reddening. There are some irregular patterns in the sky projection, they are consequences of our cell shapes and of the survey's observational fields. For the large cells of the first steps, the projections extend to all the Taurus clouds, but for the small cells of last steps only those close to the Sun contain enough targets to remain selected. As a consequence only the closest clouds are represented in the metadata.

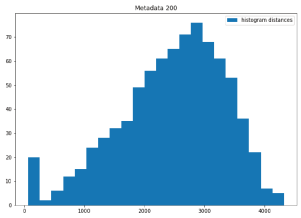
Inversion and results

I made the hierarchical inversion in 10 steps, corresponding to structures of sizes decreasing from 400 pc to 30 pc, and using the exponentially decreasing analytical prior at the beginning (the length scale is still 200 pc, that corresponds to a value above the dust scale height). The steps for the inversion are listed in Tab 13.1.

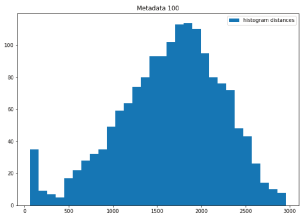
Some examples of the evolution of the 3D distribution are shown in Fig 13.4, for steps 5, 6, 8 and 10. In these figures it is possible to observe how structures are increasingly detailed during the progression of the inversion steps. The results are encouraging, since there is apparently a coherent evolution at each step, and the inversions have good χ^2, which confirms a general agreement between our statistical data and the modelled distribution after convergence.



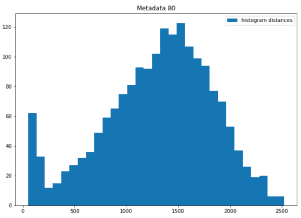
(a) Cells 500 pc.



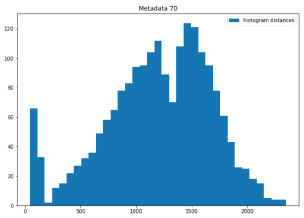
(b) Cells 200 pc.



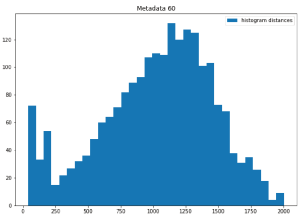
(c) Cells 100 pc .



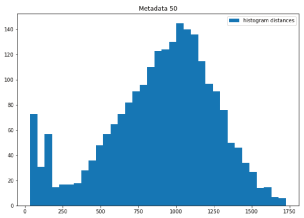
(d) Cells 80 pc.



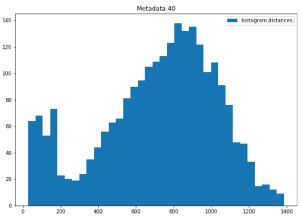
(e) Cells 70 pc.



(f) Cells 60 pc.

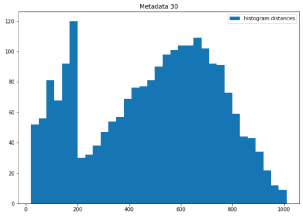


(g) Cells 50.

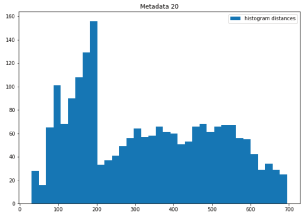


(h) *Cells 40 pc.*

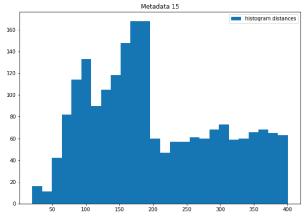
Figure 13.1 – Histograms of the distances to the target stars for the various steps (or types of metadata).



(a) Cells 30 pc.



(b) Cells 20 pc.

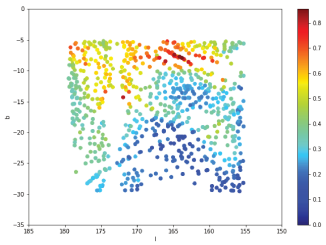


(c) Cells 15 pc.

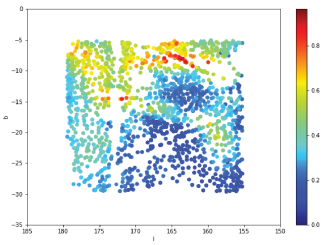
Figure 13.2 – Same as fig 13.1 for further steps.

	Mean Distance Metadata	Minimum number of objects	Total objects	ξ_0	σ_{ξ_0}	ξ_1	σ_{ξ_1}
STEP 1	400 pc	12	390	450	0.3	400	0.7
STEP 2	200 pc	12	1213	250	0.4	200	0.8
STEP 3	160 pc	12	1606	200	0.4	160	0.8
STEP 4	140 pc	12	1846	160	0.5	140	0.9
STEP 5	120 pc	12	2080	140	0.6	120	0.95
STEP 6	100 pc	12	2278	120	0.7	100	1.0
STEP 7	80 pc	12	2407	100	0.75	80	1.1
STEP 8	60 pc	12	2209	80	0.75	60	1.1
STEP 9	40 pc	12	1331	60	0.8	40	1.2
STEP 10	30 pc	5	2059	40	0.8	30	1.2
STEP 10 bis	30 pc	5	4149	40	0.8	30	1.2

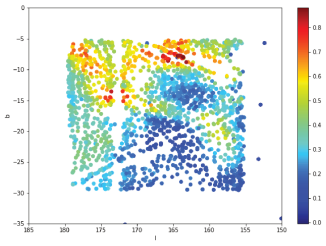
Table 13.1 – Inversion parameters for hierarchical inversion.



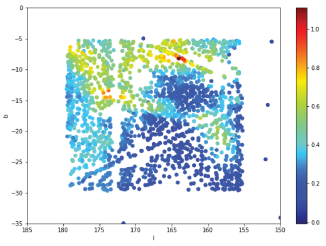
(a) Cells 200 pc.



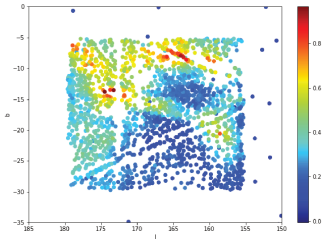
(b) Cells 100 pc.



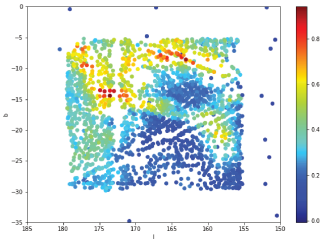
(c) Cells 80 pc .



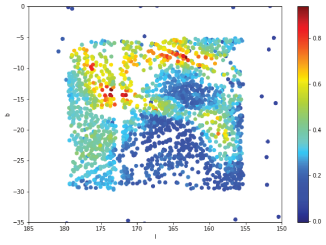
(d) Cells 70 pc .



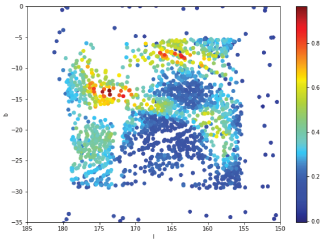
(e) Cells 60 pc.



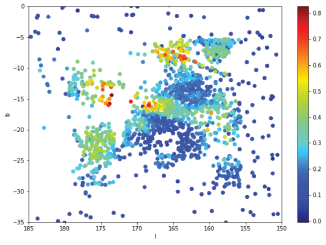
(f) Cells 50 pc .



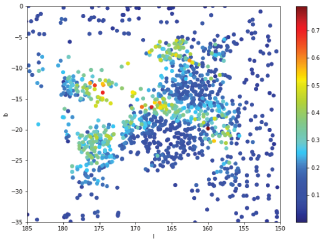
(g) Cells 40 pc .



(h) Cells 30 pc.

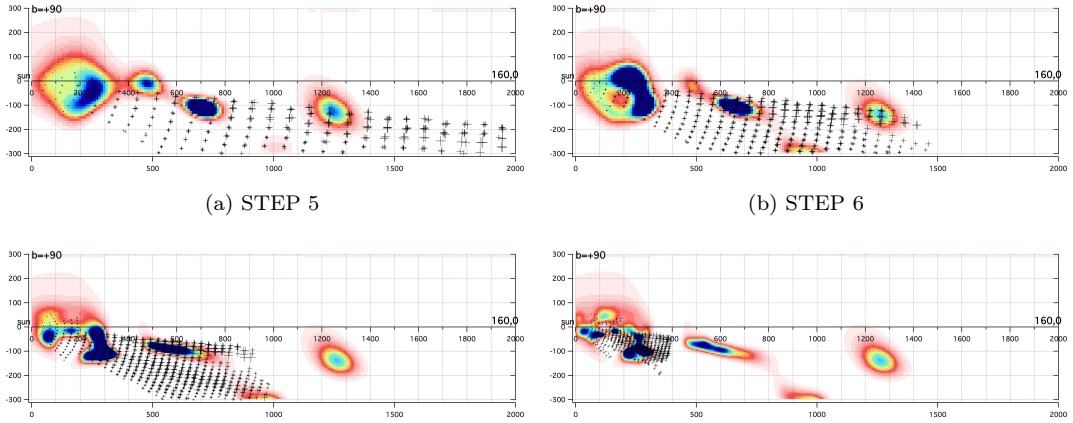


(i) Cells 20 pc.



(j) Cells 15 pc.

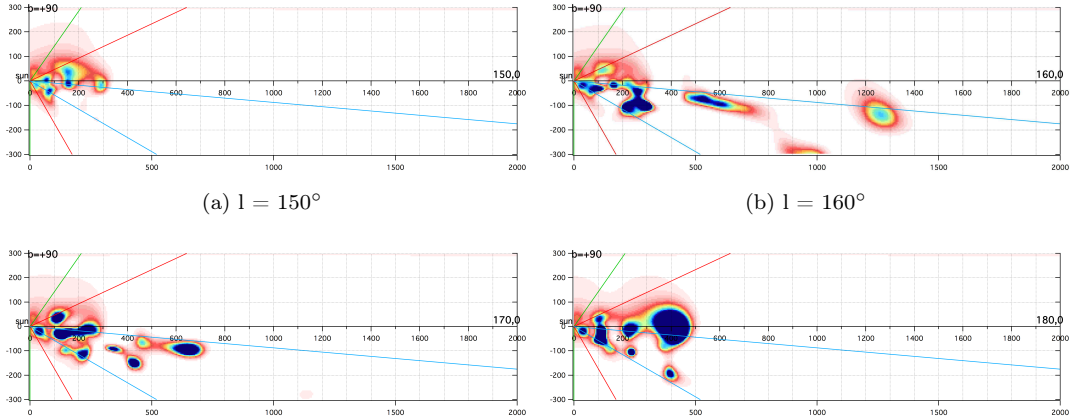
Figure 13.3 – Sky locations of meta-targets used for the inversion, in the Taurus region. The colour scale refers to the reddening excess of colour.



(a) STEP 5 (b) STEP 6

(c) STEP 8 (d) STEP 10

Figure 13.4 – Hierarchical inversion: some steps for the vertical plane at longitude 160°. Black crosses are the metadata used in the inversion, with sizes increasing with the mean excess of colour.



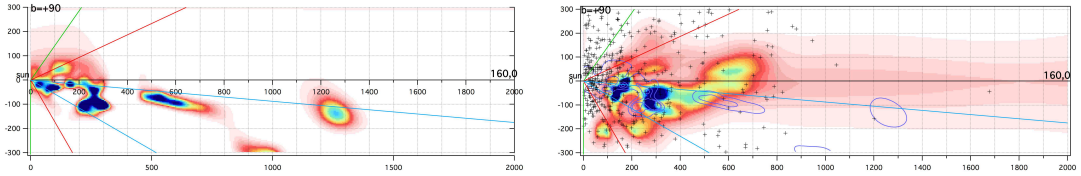
(a) l = 150° (b) l = 160°

(c) l = 170° (d) l = 180°

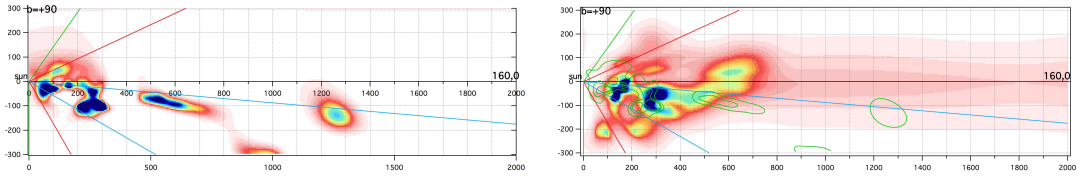
Figure 13.5 – Vertical cuts for the final step of the hierarchical inversion. Longitude: 150°, 160°, 170°, 180°. The green lines delineate the volume allowed for stars closer than 100 pc, the red lines delineate the volume allowed for stars between 100 and 200 pc, the light blue lines the volume containing all stars farther than 200 pc.

Fig 13.5 shows for the step 10 several different vertical planes, allowing to explore how the Taurus clouds are distributed when approaching the boundaries of the volume of our selected data.

The Local Bubble is apparently very strongly shrunk in these Taurus maps: indeed, at step 10, the inversion predicts some matter closer then 50 pc. This is closer than for all previous maps from direct inversions. One explanation is



(a) Hierarchical inversion - step 10 and inversion from Capitanio et al. (2017)



(b) Hierarchical inversion - step 10bis and inversion from Capitanio et al. (2017)

Figure 13.6 – Last inversion steps and comparisons: left) Inversion step 10 (top) and step 10 bis (bottom). Note the difference at short distance (see text). Right top: last direct inversion of Capitanio et al. (2017) (called the *old* map) with superposed isocontours from hierarchical inversion step 10. Also shown are the targets used in the old map inversion. Note that comparisons are valid only for negative latitudes between -5° and -30°. Right bottom: last direct inversion of Capitanio et al. (2017) with superposed isocontours from hierarchical inversion step 10 bis. Clouds beyond 500 pc could not be present in the old map due to lack of targets at large distance.

the dominant role here of extinctions deduced from Bayesian techniques and we know that they may over - estimate excess of colour for the smallest values close to the Sun, or have large error bars compared to other methods like Stromgren or Geneva photometry. To test this hypothesis and reinforce the weight of individual measurements with good accuracy for nearby stars, I added to the metadata of the last step all individual data from the old catalog corresponding to distances lower than 50 pc, well adapted for weaker extinctions. This new inversion is in 13.1 and referenced as STEP 10 bis. It can be seen in Fig 13.6 (left top, and left bottom) that the step 10bis distribution has effectively a wider cavity close to the Sun, as if matter had been pushed away at larger distance. This shows that our hypothesis is correct. However, it is also probable that using metadata is not a well adapted technique at short distance, unless one decreases even more number of objects per cell (one target per cell), which ultimately is equivalent to the direct inversion.

In the same figure is a comparison with the Capitanio et al. (2017) map in the same plane (right panel): also show at top are the targets used to generate this previous map. As can be seen, there are considerable differences (however, comparisons are valid only for negative latitudes between -5° and -30°): first, all in the old map, in this direction there are two concentred clouds 200 pc and 300 pc. Any clouds could be present in the Capitanio et al. (2017) map because there are not targets behind 1000 pc.

Chapter 14

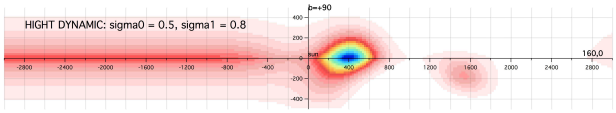
Influence of dynamics parameters

The hierarchical method requires to build efficiently metadata from the stellar dataset, but also to choose adequate inversion parameters. This latter aspect is quite challenging, because the optimal choice depends on a complex way on the dataset characteristics and what we want to extract from the data in each region. As explained in the introduction on the inversion, the way the absorbing matter is distributed in 3D space is mainly governed by the four parameters ξ_0, ξ_1, σ_0, σ_1. ξ_0, ξ_1 act on the minimum sizes of the structures, as illustrated for a very simple case in Fig 8.1, and σ_0, σ_1 control the relative influences of the prior and the data, as illustrated in Fig 8.2.

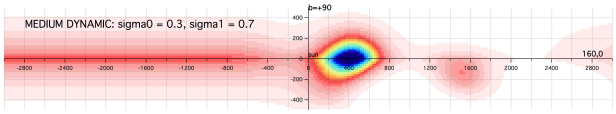
The previous inversion for example is not optimised, because apart from concentration of matter in the clouds there are completely empty spaces in the regions where we have no metadata. This is not realistic, because without data, the inversion should keep the prior values, i.e. a dust density decreasing with the distance from the Plane.

In Fig 14.1 is represented the reason of this anomaly: for this inversion we have used relatively high σs, as we had previously done for the direct inversions. However, during the first steps of the hierarchical inversion with very wide cells and accordingly very large ξ values, the large size of the modeled clouds and their influences at very large distance from their cores, combined with the allowance for very strong departures from the prior (high σ) has permitted solutions with empty regions along the Plane where there are no data. Since these regions become the new prior and there are no data to influence them at the next steps, they are kept empty until the last steps, whatever the final ξ value. Clearly, there must be some better adaptation of the σ parameter to the fist steps.

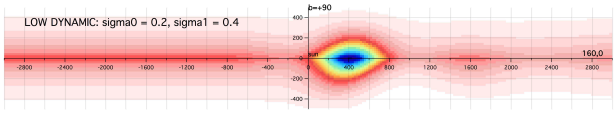
Following this idea, I tested exactly the same inversion as previously except for a change of the values of the σ parameters. The results are compared to the previous results in Fig 14.1. The result of the inversion is also shown towards the directions opposite to Taurus, where there are no data, in order to show clearly the shape of the prior. The comparison indicates that a low value for the dynamic parameters is more appropriate for the first step in a hierarchical inversion.



(a) High values for dynamic parameters



(b) Medium values for dynamic parameters



(c) Low values for dynamic parameters

Figure 14.1 – Vertical cut at l = 160°, comparison between different *dynamical* parameters.

14.1 LAMOST after GAIA DR2

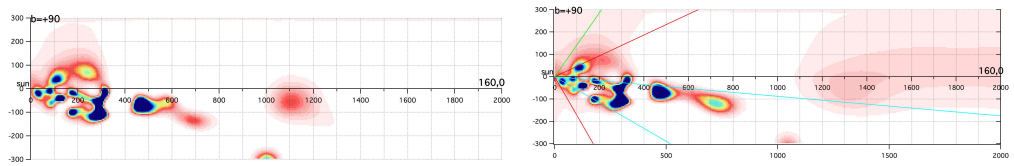
In April, the Data Release 2 of Gaia was available, and we crossmatched the most recent LAMOST dataset available, i.e. LAMOST DR4, with the Gaia DR2 catalog, to substitute the photometric distances with distances derived from the Gaia parallaxes. In this section I show the changes due to the new distances and the new dynamical parameters, using the new LAMOST DR4 dataset.

This inversion was done in ten steps similar to those in previous chapter. In Tab 14.1 are listed the used parameters. The inversion was now done both with the former σs and σs lower than in the previous tests. In Fig 14.2 is shown a comparison of the results based on the new LAMOST catalog for the same parameters as in Tab 13.1 and the new ones. The difference is very significant in regions where there are no data, showing the influence of the dynamical parameters. It is clearly not realistic to *empty* regions along the Plane in regions without data. In areas well covered by data, the inversions reproduce about the same structures, since the influence of data is dominant.

Looking at differences between Fig 14.2 and panel a) in Fig 13.4 there are some differences, these are mainly due to more accurate distances. For step 10 at l=160°, the structure previously located at 600 pc is now clearly separated in two parts, a dominant one at 400 pc and a less dense, more distant one. As we will see in Chapter 22, the elongated cloud at 400 pc is the well-known California complex. The cavity between ∼100 and ∼200 pc is confirmed, although smaller. More generally, all nearby structures are now smaller or more fragmented. This is a very positive sign, demonstrating that the data are more coherent. This is clearly due to the more accurate parallaxes.

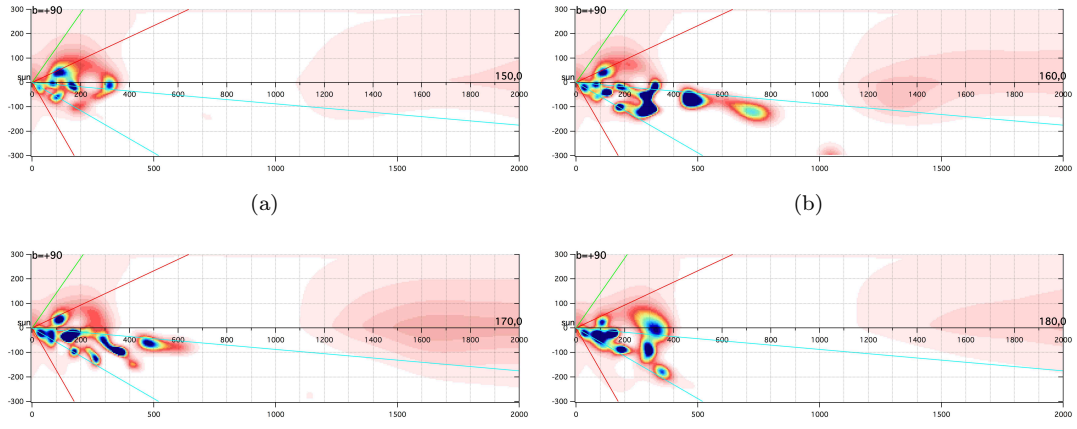
	Mean Distance Metadata	Minimum number of objects	Total objects	ξ_0	σ_{ξ_0}	ξ_1	σ_{ξ_1}
STEP 1	400 pc	12	344	450	0.2	400	0.3
STEP 2	200 pc	12	993	250	0.25	200	0.35
STEP 3	160 pc	12	1335	200	0.3	160	0.4
STEP 4	140 pc	12	1552	160	0.35	140	0.45
STEP 5	120 pc	12	1797	140	0.4	120	0.5
STEP 6	100 pc	12	1999	120	0.45	100	0.55
STEP 7	80 pc	12	2158	100	0.5	80	0.6
STEP 8	60 pc	12	2294	80	0.6	60	0.8
STEP 9	40 pc	12	1667	60	0.7	40	0.9
STEP 10	30 pc	5	2683	40	0.8	30	1.0

Table 14.1 – Inversion parameters for hierarchical inversion.



(a) STEP 10 - With Medium Dynamic (b) STEP 10 - With Low Dynamic

Figure 14.2 – Vertical cut at longitude 160°, for the steps 10 in hierarchical inversion after Gais DR2, using the same set of metadata based on LAMOST catalog, with two different dynamic: the smaller dynamic parameters drive to more realistic results.



(a) (b)

(c) (d)

Figure 14.3 – Distribution in several vertical planes, after the last step of the hierarchical inversion LAMOST DR4- Gaia DR2. Coloured lines have the same meaning than in Fig 13.5.

Part V

Full-sky hierarchical method applied to Gaia DR2 extinctions

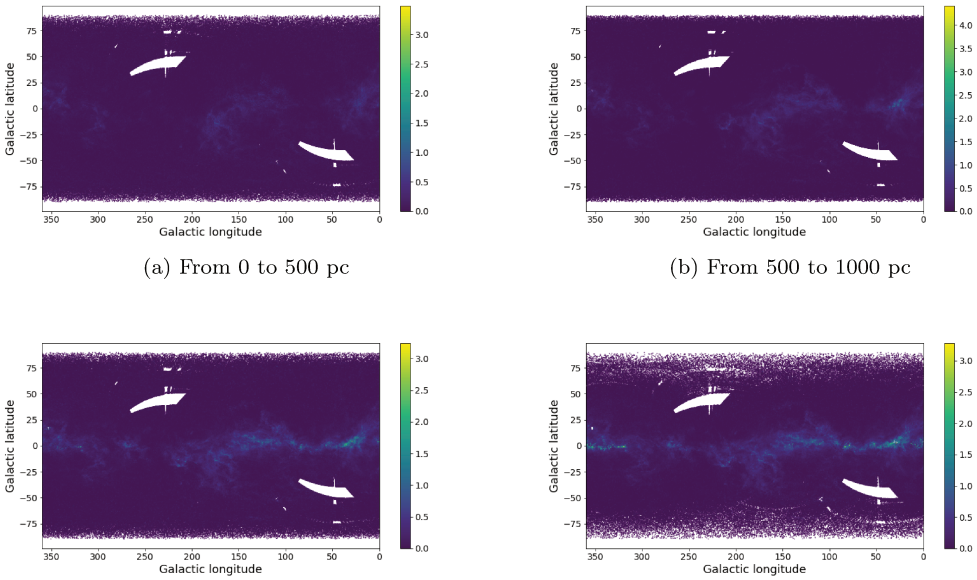
Chapter 15

Application of the hierarchical inversion to full-sky data

The main interests of the hierarchical method are its adaptability to massive, full-sky datasets and to inhomogeneous target distributions (in practical, mostly distance-dependent target densities). These are the characteristics of Gaia-based data, namely present and future extinction estimates and future DIB measurements. The series of tests presented in previous sections and limited to a specific region has been done in this perspective of extending the inversion to full-sky data.

The final exploitation of the future Gaia DIB catalog, in the context of ISM studies, is a DIB carrier 3D map based on full-sky measurements of the DIB equivalent width. As Gaia RVS spectra are not yet available, here I tested a full-sky hierarchical inversion of dust colour excess excess of colour derived from Gaia DR2 data. We did not use extinctions made publicly available from the Gaia website since they were determined from the broad-band photometric measurements of Gaia only; that caused a strong degeneracy between stellar temperature and extinction determinations (Arenou et al., 2018). Instead, we used a new catalog of extinction estimates computed by C. Babusiaux (paper in preparation). The extinction A_0 is converted into colour excess excess of colour using a simple proportionality A_0=3.1 excess of colour. The extinction is deduced from the combination of Gaia and 2MASS photometry. The inclusion of 2MASS eliminates the degeneracy.

The main characteristics of the catalog we have used are listed in Tab 15.1. Most important for our purpose, it is limited to targets with relative uncertainties on the parallax lower than 20%, a point that is essential for mapping. This limitation, in turn, implies relatively bright targets, and explains why 94% of stars are closer to the Sun then 3000 pc, even if the maximal distance reaches 17 kpc. A small fraction of the sky is not covered - see Fig 15.1 - but in theory with the hierarchical inversion this problem is less important, especially because the resolution will remain limited in those directions and no spurious features can be created where there is no information. Moreover, these small holes are at high galactic latitudes, while the Galactic plane is extremely well covered, and, because the Plane is the richest region in dust, these holes should not impact the results.



(a) From 0 to 500 pc

(b) From 500 to 1000 pc

(c) From 1000 to 1500 pc

(d) From 1500 to 2000 pc

Figure 15.1 – Position on the sky for all stellar data at differente distances, colour is excess of colour.

Table 15.1 – .

Total catalog	30630398
Total catalog closer then 600 pc to the plane	26182597
Mean sigma data	0.098
Median sigma data	0.099
Distance max	16792.35 pc
Distance min	14.37 pc
excess of colour max	4.41
excess of colour min	-8.62E-5

15.1 Metadata selection

For the first inversion of all sky GAIA extinction measurements, I proceeded with the following step characteristics:

- 800 pc, with metadata calculated in distances slices of 400 pc, max distance 3000 pc
- 400 pc, with metadata calculated in distances slices of 200 pc, max distance 3000 pc
- 200 pc, with metadata calculated in distances slices of 100 pc, max distance 2500 pc

- 160 pc, with metadata calculated in distances slices of 80 pc, max distance 2500 pc
- 100 pc, with metadata calculated in distances slices of 50 pc, max distance 1750 pc
- 60 pc, with metadata calculated in distances slices of 30 pc, max distance 1400 pc

Further steps are in progress, however will not be entirely computed before the end of preparation of this document. As we have a huge quantity of input data, the production of metadata was slow, so I imposed some specific limitations. The first limitation is on the altitude above/below the galactic plane, because during the inversion I do not use data that are too far from the disk, is it not useful to produce them. The second limitation was on the extinction uncertainty. Input targets were selected in the DR2 catalog with a threshold of 10% on the relative parallax error, so I have very precise data in distances. I impose then for excess of colour lower then 0.3, I accept only data with absolute error lower then 0.1; for excess of colour higher then 0.3, the relative error must be lower then 20%. To get only the best metadata I selected only cells with at least 50 objects par cells.

15.2 Inversion

To make this inversion I started with an analytical prior distribution, with an opacity decreasing exponentially at increasing distance from the Plane. Our motivation was to obtain results with only Gaia data. I used only the Gaussian correlation function:

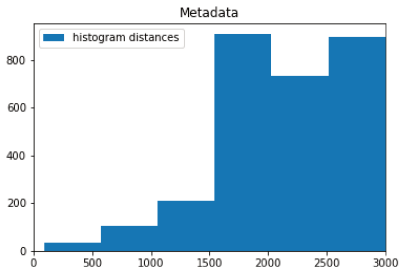
$$C(x) = {\sigma_{\xi_1}}^2 \exp^{\frac{-x^2}{\xi_1}} \tag{15.1}$$

and not the two-term function C_1 8.13 used in all the previous inversions that contains a hyperbolic cosine. The tails in the C_1 cosh term are very important for the mapping of small structures, i.e. when we use small correlation distances ($\sim$10pc), because it allows to represent two phases of the clouds, a compact phase for the dense central part and a more diffuse external part. Here however, for all the first steps of the hierarchical inversion with large correlation distances, the cosh term introduces very extended tails and imposes a correlation between very distant points that has any physical meaning.

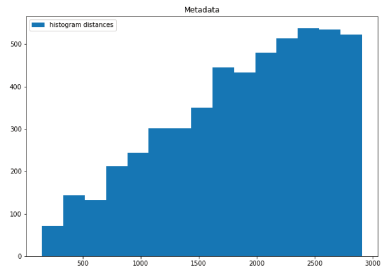
As our errors on excess of colour are really small, I kept the additional error of 0.01 mag, as in others inversions.

First step - 800 pc For the first step 879 cells were selected.

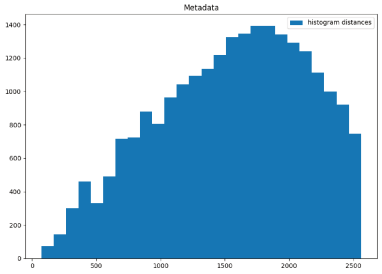
- $\sigma_1 = 0.3$ and $\chi_1 = 400$ pc
- Beginning $\chi^2 = 41.18$
- End $\chi^2 = 2.14$
- CPU time = 210505
- Threads number = 6



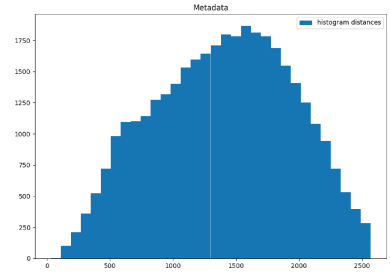
(a) Metadata with mean distance 400 pc



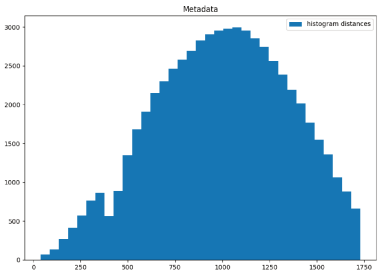
(b) Metadata with mean distance 200 pc



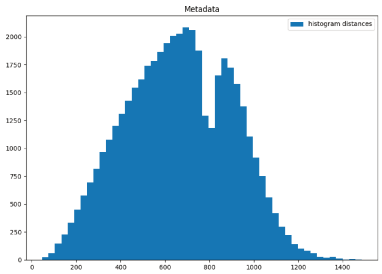
(c) Metadata with mean distance 100 pc



(d) Metadata with mean distance 80 pc



(e) Metadata with mean distance 50 pc



(f) Metadata with mean distance 30 pc

Figure 15.2 – Distance histograms for metadata with different mean distances.

- Real time = 8h57m

Second step - 400 pc For the second step 5223 cells were selected.

- $\sigma_1 = 0.4$ and $\chi_1 = 400$ pc
- Beginning $\chi^2 = 35.05$
- End $\chi^2 = 2.20$
- CPU time = 198959
- Threads number = 6
- Real time = 11h 48m

Third step - 200 pc For the third step 23488 cells were selected.

- $\sigma_1 = 0.4$ and $\chi_1 = 200$ pc
- Beginning $\chi^2 = 17.59$
- End $\chi^2 = 1.50$
- CPU time = 1741599
- Threads number = 6
- Real time = 4d 14h 35m

Fourth step - 160 pc For the first step 35610 cells were selected.

- $\sigma_1 = 0.5$ and $\chi_1 = 160$ pc
- Beginning $\chi^2 = 24.70$
- End $\chi^2 = 1.28$
- CPU time = 5911498
- Threads number = 6
- Real time = 15d 15h

Fifth step - 100 pc To diminish the computational time, I separated the metadata in two parts, centre and anti-centre, and I merged the results as explained in Section 10.9. Centre inversion was done with data from l = (-95 ; 95) and 35492 targets were selected; anticentre was inverted with data in the longitude interval l = (85; 275), and 26140 targets were used.

- $\sigma_1 = 0.6$ and $\chi_1 = 100$ pc
- Beginning χ^2 centre = 9.75, anticentre = 13.52
- End χ^2 centre = 1.08, anticentre = 1.09
- CPU time centre = 4026913 ; CPU time anticentre = 1880766
- Threads number = 6
- Real time centre = 11j 4h, real time anticentre = 6j 10h

Sixth step - 60 pc Also in this case I spitted the inversion in four parts, for the four quadrants on the galactic plane. First, second, third and fourth quadrants contains respectively data in longitude intervals [-5°: +95°], [85°: 185°], [175°, 275°] and [-95°: 5°]. The targets belonging at these intervals and used for the inversion are respectively 27416, 17787, 19320, 29789.For these inversions the calculation time is higher than expected, and also too high by comparison with the computational time in the previous step at 100 pc. We believe that the way the work-space in our computer was configured was not adapted. In the future, the same inversions with other computers and the same algorithm will be performed to discern the origin of this problem.

- $\sigma_1 = 0.7$ and $\chi_1 = 60$ pc
- Beginning χ^2 first quadrant = 4.43, second quadrant = 4.88, third quadrant = 1.97 , fourth quadrant = 2.78
- End χ^2 first quadrant = 1.08, second quadrant = 0.90, third quadrant = 0.60 , fourth quadrant = 0.82
- CPU time first quadrant = 18 310 103; CPU time second quadrant = 6 177 830; CPU time third quadrant = 6094402 ; CPU time fourth quadrant = 19 595 894
- Threads number = 6
- Real time first quadrant = ∼37j, real time second quadrant = 18j; real time second quadrant = 17j; real time fourth quadrant = 40j

15.3 Results

The results are really encouraging, because the strong point in this technique is the suppression of non-realistic features (given the resolution achived at each step), and it seems that the objective is reached. Looking at the Galactic plane, at first steps the matter is located as as a semi-ring around the Sun, with a bloc in the Galactic centre direction - Fig 15.3. In the third quadrant appears a large cavity without matter. Step by step, the structures are effectively located, and prior values are kept at distances without meta-stars. No *fingers of God* are present on the Galactic plane, even at smaller scales, see Fig 15.4. Such features are generally present in other inversion techniques that are not full-3D inversions. At 60 pc resolution, the Galactic plane map shows two non-radial thick chains of clouds in the second and third quadrants, and a more distant and parallel filament in the Galactic centre direction.

Fig 15.5 illustrates the evolution away from the Galactic plane, showing the Taurus clouds. While the first steps show extended blobs centred below the Plane, in further steps there are important fragmentations into smaller structures, e.g. between step 3 (200 pc) and step 4 (160 pc); starting with steps 5 and 6, the locations of the cloud cores are stabilised, but they become thinner: the cavity between 200 and 400 pc near the plane is emptier, the structure at 100 pc is separated in two smaller cores. Note that for this figure the final resolution is 60 pc, still above the one achieved for the same area after the previous inversions that were spatially limited to Taurus (25 pc).

In Fig 15.6, there are some examples of distributions in vertical planes for the step 6. The original analytical prior is not discernible, and this at all longitudes. This reflects the fact that the initial dataset is really homogeneously distributed and the metadata fill entirely the space where we are computing the extinction. This is the strenght of Gaia data over ground-based surveys.

Errors estimation

For the hierarchical method it is simple to estimate errors on distances to the structures, because it is intrinsically related to the metadata size and sampling. Based on the histograms' *bumps* in Fig 15.2, I can assume that after step 6 all structures closer to 1000 pc have distance errors on the order of 50 pc, those between 1000 pc and 1500 pc an error of 80 pc, and between 1500 and 2000 pc an error of $\sim$ 100 pc.

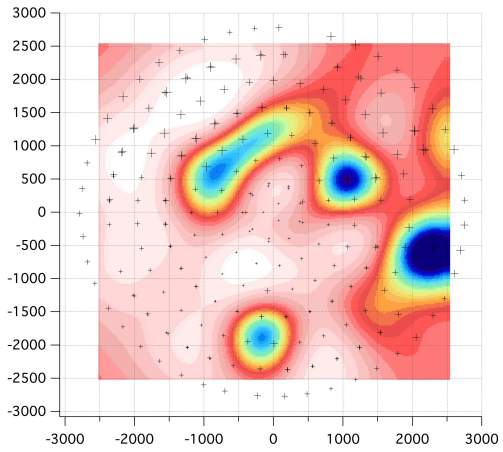
At variance with the case of distances, any estimate of errors on the reddening excess of colour is challenging because in this case we have different input data set. A possibility could be to calculated the mean difference from excess of colour of metadata and the results. Another possibility is to implement a test as in Chapter 12, selecting subsample and inverting step by step to understand if the dataset contains some bias. Moreover, I could change the geometry of our metadata, for example shifting the angles, and observe potential changes. All this possibility are interesting, but this investigation it is too slow for my 3 years thesis.

Comparison with other maps

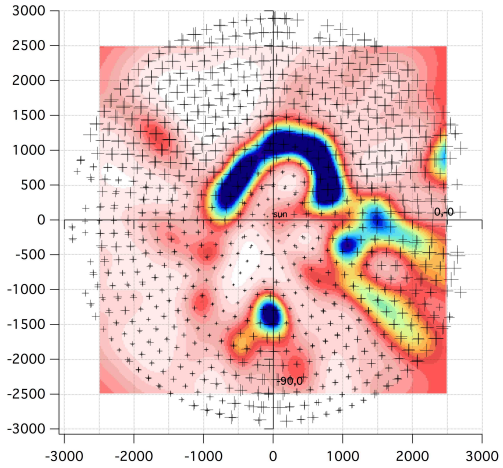
The same dataset with exactly the same targets and extinction values has been used by our collaborator Jean-Luc Vergely to enter a different type of inversion. This newly developed mapping technique (here after called technique-2) is not a full 3D like the one I handled: the inversion is hierarchical but with two different computations at each step of resolution: the first part considers each line of sight individually and builds series of reddening profile, then during the second part an azimuthal average is performed to smooth the structures. His map is computed directly in extinction A_0 and not in excess of colour but as we said the difference is simply a scaling factor. This technique is completely independent, and it is interesting to compare a different map build with the same dataset. In Fig 15.7 are shown the Galactic planes distributions for the Vergely map with final resolution 25 pc and for step 6 of our inversion.

The figure shows that globally the maps are in very good agreement, taking into account that our 60 pc resolution is poorer than the 25 pc resolution of the second inversion. Most importantly, some of the features looking like *fingers of God* in the technique-2 map are not present in our map. This is very important, since it demonstrates that our full-3D hierarchical method may provide a *safety-check* for other non full-3D inversions where radial artefacts remain.

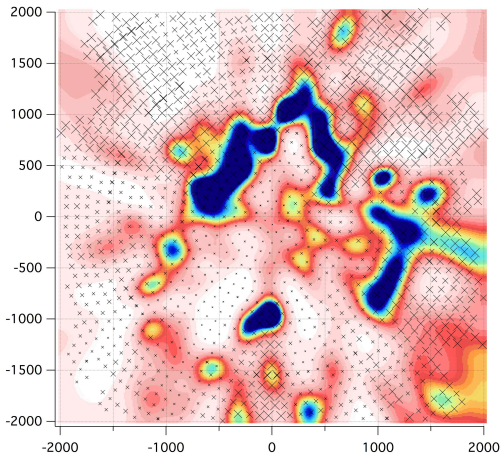
While, as said above, our step 6 hierarchical map and the technique-2 map are similar, the main difference is close to the Sun in the so-called Local Bubble area. This region does not appear empty in our step 6 map, while it is a visible cavity in the other map. There are several potential reasons for this difference and their investigation is still in progress. Maybe metadata boxes are too wide, or it is due to my strict condition for the excess of colour relative error.



(a)

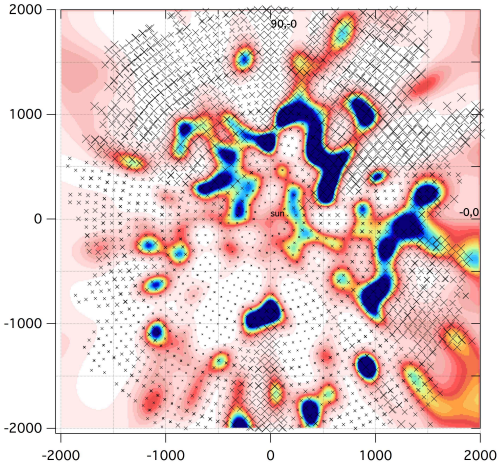


(b)

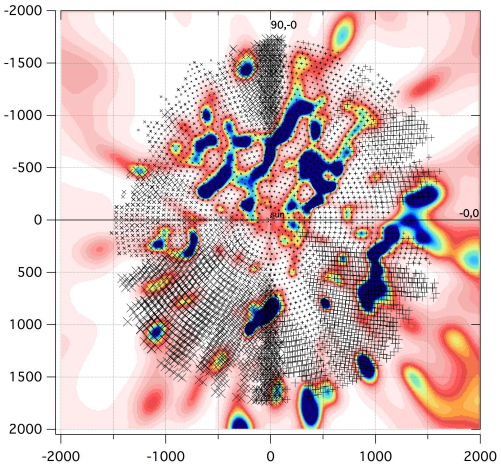


(c)

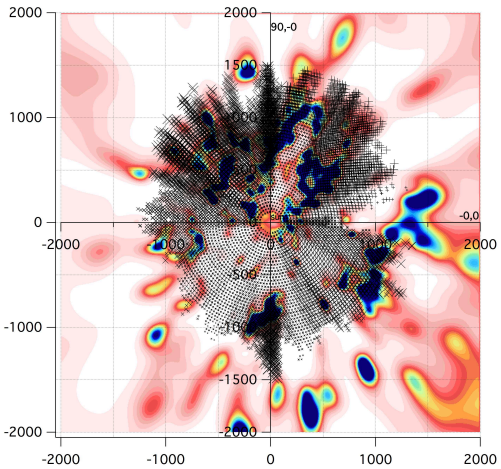
Figure 15.3 – Galactic plane differential extinction distribution after step 1, step 2 and step 3. Targets with distances to the Plane smaller than 200 pc are represented with a cross whose dimension represents the reddening excess of colour. Colour scale as in Fig 10.9 - panel d).



(a)

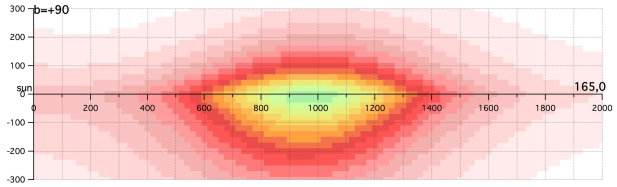


(b)

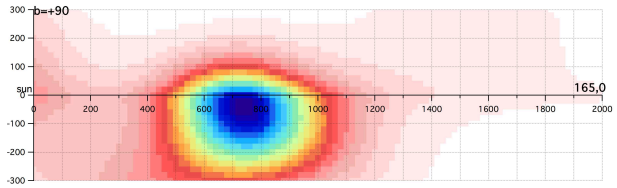


(c)

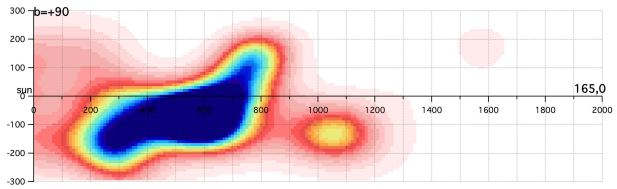
Figure 15.4 – Galactic plane for inversion for step 4, step 5 and step 6. Colour scale as in Fig 10.9 - panel d).



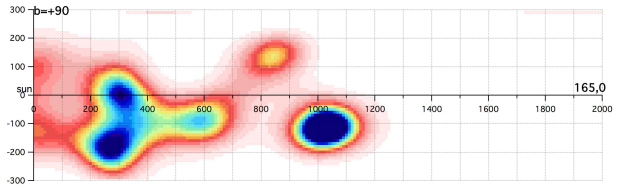
(a)



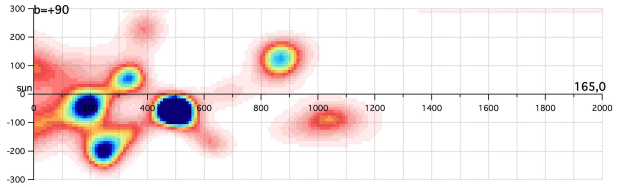
(b)



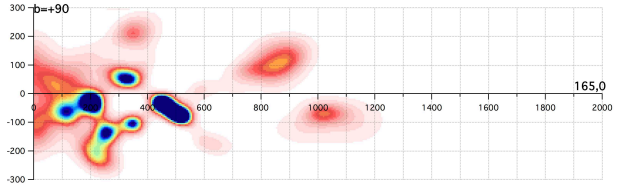
(c)



(d)

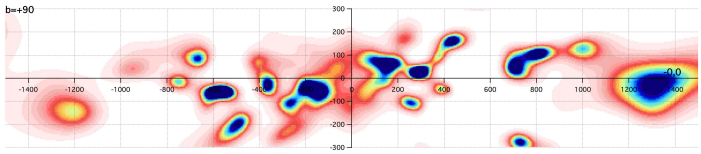


(e)

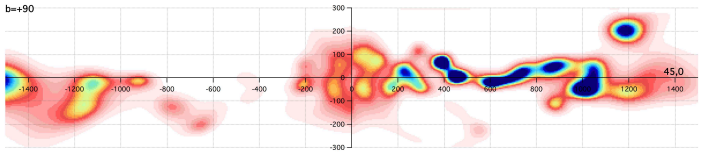


(f)

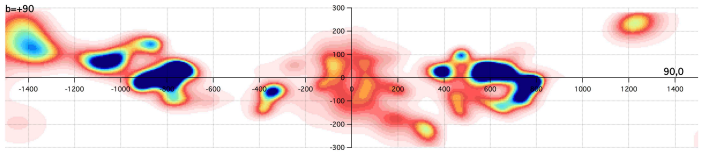
Figure 15.5 – Vertical plane in Taurus direction - galactic longitude = 165° - for different steps. a) Step 1 b) Step 2 c) Step 3 d) Step 4 e) Step 5 f) Step 6. Colour scale as in Fig 10.9 - panel d).



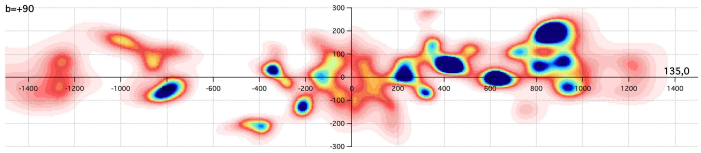
(a) l = 0°



(b) l = 45°



(c) l = 90°

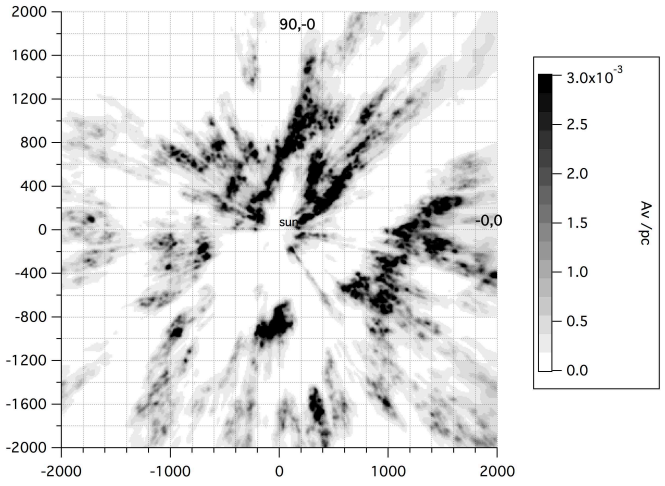


(d) l = 135°

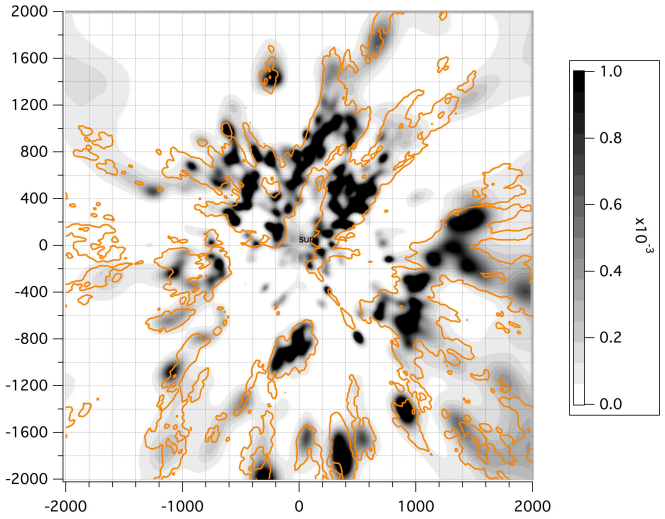
Figure 15.6 – Distribution in vertical planes for the step 6 inversion, at different longitudes.

Another (likely) possibility is that it is due to the four separate inversions for four complementary regions I did, since I made the separation based on the coordinates only. Boxes close to the Sun do not contain enough targets to really constrain the small opacity and the prior distribution forces the filling near the Sun. For each inversion I should have included all targets close to the Sun until a given distance whatever the coordinates, as done for the Taurus region.

Rezaei Kh. et al. (2018) have published a large scale map based on APOGEE DR14 measurements; they use a non-parametric method to infer the dust distribution in 3D space (Rezaei Kh. et al., 2016); as it is based only on APOGEE DR14 they map a small part of the Galactic plane. Fig 15.8 displays their map with our last step map superimposed. Our map is dedicated at the local interstellar medium, on the contrary their objective is to study the spiral arm structure in the galaxy, but in the zoom (bottom of the figure) it is possible to compare our results: the filaments elongated near the Sun are coherent in two maps, and in third quadrant our map find a cavity and a new group of clouds at 1000 pc which have a counterpart in their map. For the filament in the direction of the Galactic centre, we find it located at 1000 pc, but their map



(a)



(b)

Figure 15.7 – Top) Galactic plane for the extinction map performed by J.L. Vergely with technique 2. Bottom) Galactic plane for step 6 with superimposed in orange iso-contours measured in the technique 2 map at $A_V = 0.005$.

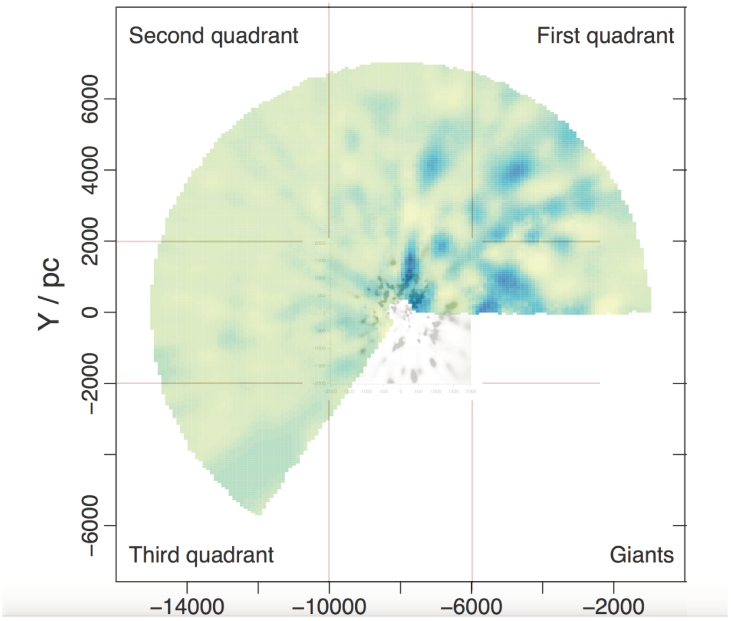
suggest a more distant wall of matter. This discrepancy could be caused by the fact that this direction is at the boundary of the APOGEE target distribution over the sky, or there is a bias in this direction for the Gaia parallaxes. Others maps with different datasets or using the new Gaia calibration shall clarify this point.

15.4 Conclusion and perspective

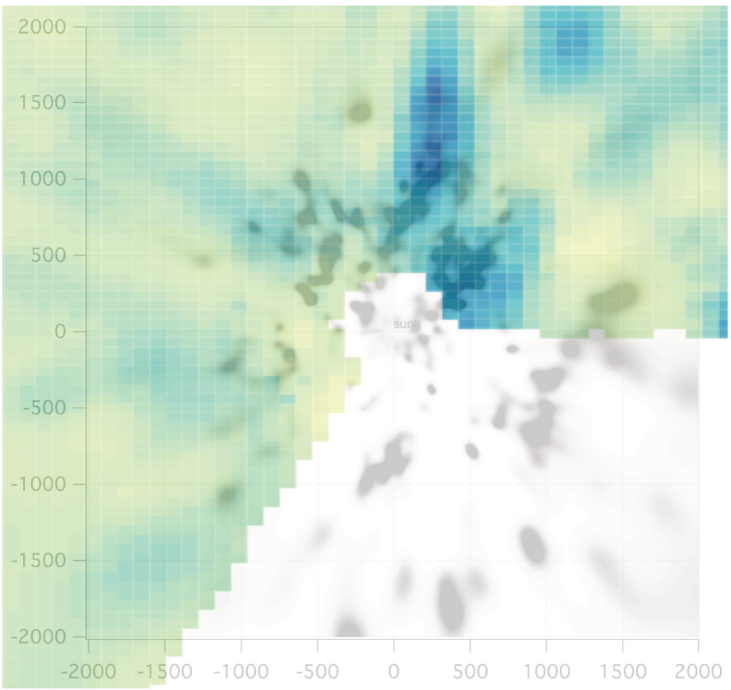
The hierarchical approach for full 3D mapping seems to be a satisfying solution adapted to massive datasets and works well, as shown by comparison with maps made with other techniques. It is also particularly well adapted to catalogs of targets uniformly distributed over the sky: the comparison with the direct inversion from Lallement et al. (2018) that was largely based on APOGEE DR14 photometric data with very limited and inhomogeneous sky coverage demonstrates well that the spatial uniformity provides better reliability of the structures.

On the other hand, the metadata calculation is still too slow and not optimised. Using smarter algorithms with faster metadata calculation, it should be possible to create different metadata set and inverse them, to better understand the sources of biases in the maps due to the metadata and to better constrain the structures locations. Another important study to be done is varying the choices for the steps and resolutions, and check whether the iterations converge toward the same structures, to better assess the hierarchical method reliability. This will be possible with faster computers.

As we have seen, it can be used to detect artefacts in maps based on other techniques, especially radial features that are otherwise very difficult to dismiss.



(a)



(b)

Figure 15.8 – a) Galactic plane in Rezaei Kh. et al. (2018) with superposed our map for the step 6. b) Zoom on the local interstellar medium of a) map.

Part VI

Astrophysical information in maps

Chapter 16

First 3D map completely based on the inversion of DIB measurements

In the perspective of a large Gaia DIB dataset at Data Release 3, it is important to understand how such data can be used for 3D mapping and what can be learned from these maps. Evidently, an inversion based on a given DIB only will give us the 3D distribution of the interstellar species responsible for the absorption. On the other hand, since all DIBs are at least partially correlated with the amount of IS matter, DIB distributions should be very similar to dust or gas distributions. This is why DIBs may also potentially be used a an additional tracer to be used e.g. in conjunction with extinction (this has been already presented in Chapter 9.

We did a first test with data from the SDSS/APOGEE data release 14 (DR14), because this data release has been accompanied by large excess of colour catalogs, based on different statistical techniques and because the DIB located at 15 273 Å has been studied extensively by my colleague M. ElYajouri. She has fitted the spectra and extracted the DIB with high precision, and these absorptions, after conversion into extinctions, have been used to improve extinction maps, as I have shown in Chapter 9. The idea of the present chapter is different: we have performed an inversion for the same stellar targets, using the excess of colour provided by photometric calibrations on the one hand, and the DIB absorption (converted into extinction with a simple scaling factor) on the other hand.

Our goal is to compare the maps to understand the difference between the location of the dust (through the reddening) and the location of the macromolecules causing the DIB absorptions. At the same time such comparisons are additional tests of our inversions and we infer whether our inversion technique needs particular adaptations in case of DIB-absorption inversions.

16.1 Description of the dataset

Source of extinction values There are several extinction catalogs associated with the APOGEE DR14 dataset; the APOGEE team istelf proposes different extinction estimates. I used the so-called BPG estimate based on the

Starhorse technique(Santiago et al., 2016; Queiroz et al., 2018). [1]

Target distances The BPG catalog was crossmatched with Gaia DR1 to extract stellar targets with Gaia parallaxes; I compared the relative error on distance based on parallaxes with those based on photometry and in used in each case the one with the smallest relative error.

Conversion of DIB data I used the full catalog of 120 362 stars with a DIB measurement and fitted the relationship between the DIB equivalent width and the extinction with a simple linear relation - Fig 16.1:

$$E_{(B-V)} = \frac{EW}{0.063 * 3.1} \tag{16.1}$$

This conversion factor was used for all DIBs.

Inversion Because the total number of targets is large, in order to reduce the computation time I made the inversion in two parts as in Chapter 10.9: the centre-inversion was done with all targets in the galactic longitude interval (0°: 145°) and the anticentre-inversion used targets in the galactic longitude interval (135°: 360°).

Sky coverage of DIB-extinction targets Fig. 16.2 shows the location of the APOGEE target stars used for the inversion. It can be seen that the sky coverage is very limited and very far from uniform. This is extremely important here: it implies that the derived 3D maps will be very limited and very far from complete, and they will not be appropriate to any astrophysical use: on the other hand, since the DIB and extinction inversions are made exactly for the same targets and suffer from the same limitations, we can use the two maps for our own comparisons.

16.2 Comparison between the map based on DIBs and the map based on extinctions

We warn again here that the inverted maps are not reliable interstellar medium distributions. A 3D inversion needs a dataset as uniform as possible and here the target distribution is far from reaching this requirement: the data are grouped into regions of several degrees width with large empty regions between and almost no data far from the galactic plane. The real interesting result here is the comparison between the reddening per parsec distribution derived from DIB absorption and the one derived from extinction.

Fig 16.3 shows the galactic Plane distribution from the inversion based on DIB absorption (top panel) and on photometric extinction (medium panel).

Both maps reveal structures in the Plane that are very roughly distributed in the same way, however there are marked differences. E.g. there are some similar structures such like the filament at 1000 pc toward the Galactic centre, or the horseshoe structure in third quadrant between x=[-800, -400] pc and y

[1] [http://www.sdss.org/dr14/data_access/value-added-catalogs/?vac_id= apogee-dr14-based-distance-estimations/]

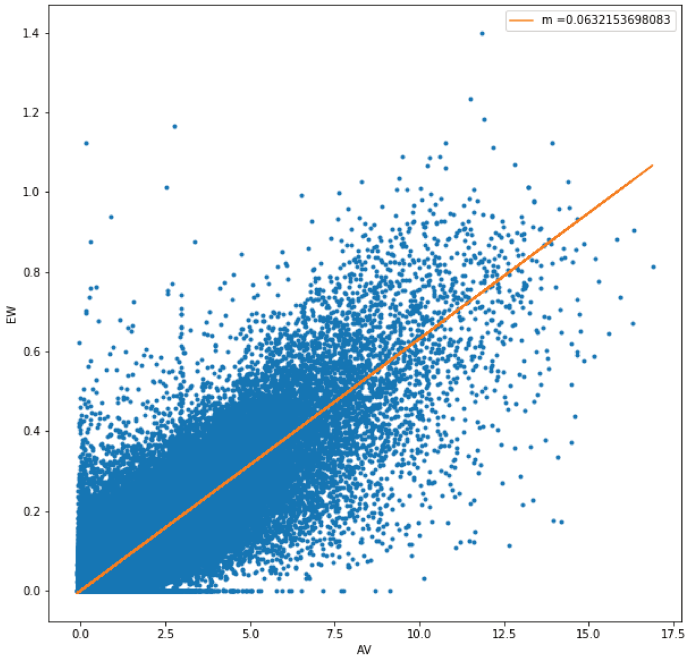


Figure 16.1 – Relationship between AV BPG and equivalent widths.

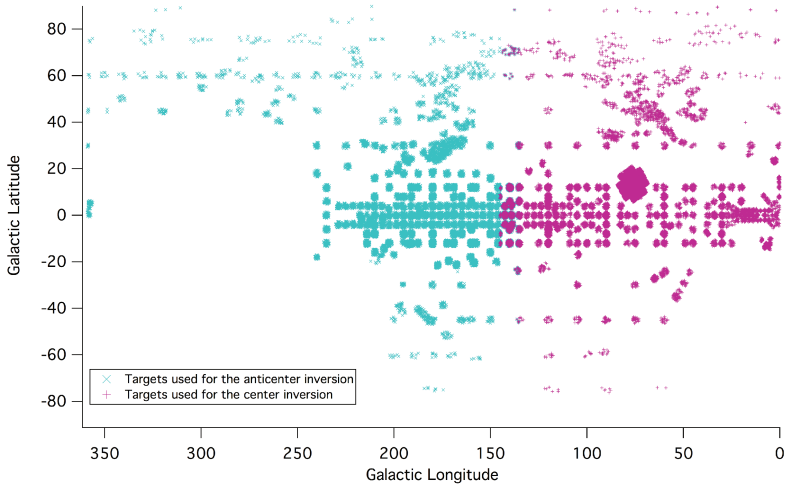


Figure 16.2 – Targets position on the sky.

[200, 800] pc, but there are very different cloud distributions in the longitude interval of [60°, 90°]. Clearly the data sampling and the uncertainties on the two quantities are such that the cloud locations are not very accurate.

To help the comparisons isocontours were derived from the DIB-based distribution and they are superimposed on the extinction-based distribution in Fig 16.4. The figure shows that in general isocontours based on DIBs absorptions correspond to wider structures, however with smaller dense cores. This suggests that DIB carriers do not follow exactly the dust grains but more numerous in cloud envelopes and less numerous in cloud cores. This , in fact, corresponds to the results of studies of specific regions (e.g. DIBs in the Barnard 68 cloud, see Elyajouri et al. (2017a)).

Fig. 16.5 shows the dust and DIB matter distributions in vertical planes containing the Sun. For some of these planes the two distributions are very different (e.g. the l=90- 270 °Plane) , for others the main structures look similar (e.g. the l=135- 315 °Plane). This suggests that the combination of data distribution and data quality is not sufficient to perform maps with good significance. On the other hand, a striking result is the presence of DIB carriers in the Local Bubble at high altitude above or below the Plane, while the map based on the extinction reveals a completely empty area. This is probably not related to the target distribution very surprising and needs further investigations with other DIBs.

Comparison of input extinction and DIB data for series of lines of sight

On order to understand the source of the marked discrepancies between the two maps and especially to determine whether the results for the Local Bubble are really constrained by the input data or are due to the inversion process, we have directly compared the input data in several regions. In Fig 16.6 is shown how excess of colour varies with distance at various different latitudes and for the longitude l=90°. The general trend is that that on the plane DIB carriers and dust follow each other, but at higher latitude there are a number of high values for the DIB carriers. This would be compatible with the hypothesis that absorbing macromolecules are more closely linked with the gas distributions than with the dust grains. My colleague M. ElYajouri is currently reaching similar conclusions about the DIB APOGEE carriers with others analyses.

16.3 Conclusion

This is only a first test, but it returns some important information. First, to convert a DIB EW into an extinction we shall be very careful about possible variations of the ratio between the two quantities that may vary from one direction to another, in longitude and latitude. The future map based on Gaia DIBs will be very interesting because the target distribution will be more regular and the map itself will be more reliable.

We have used in Chapter 9 APOGEE DIB absorptions converted into extinction and, for the inversion, we have allowed for a very large associated uncertainty. This uncertainty can account for the DIB to extinction ratio variability. The addition of this dataset has added information about the matter

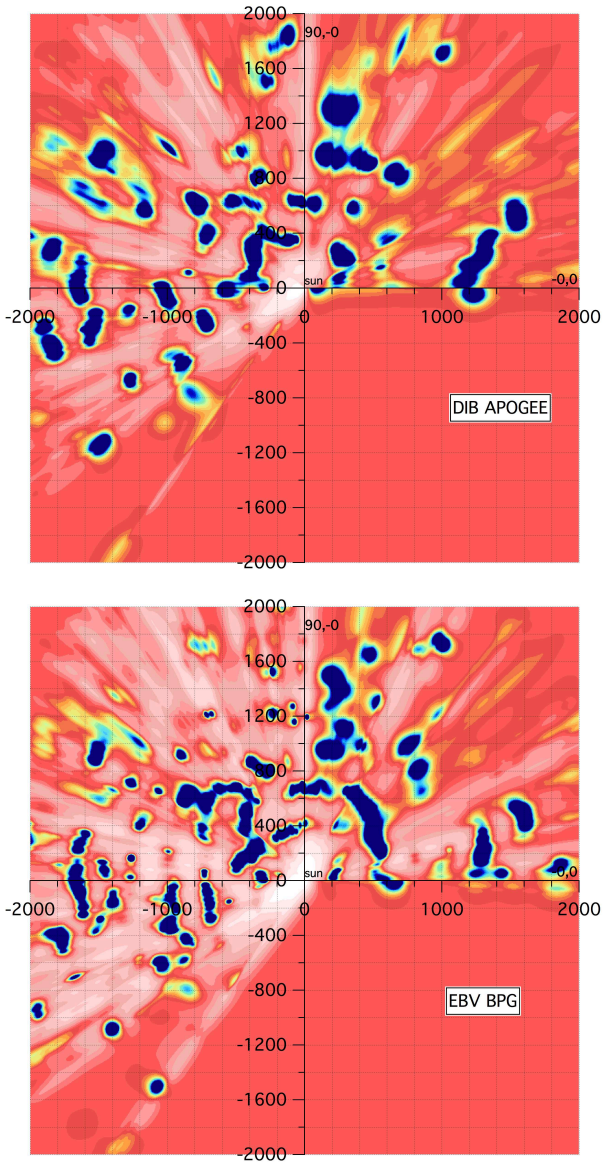


Figure 16.3 – Top) Galactic plane for the inversion based on excess of colour derived from the photometric extinction. Medium) Galactic plane for the inversion based on excess of colour derived from the APOGEE DIB absorption.

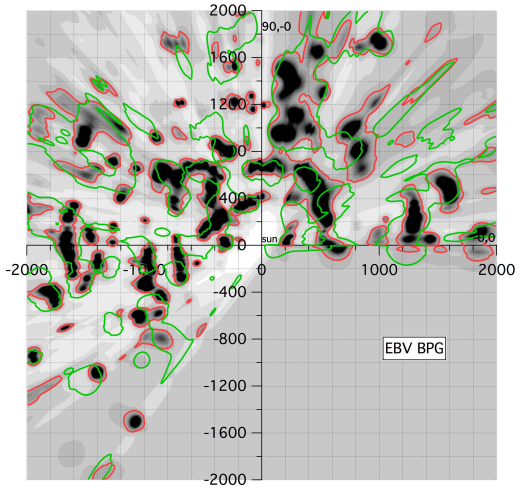
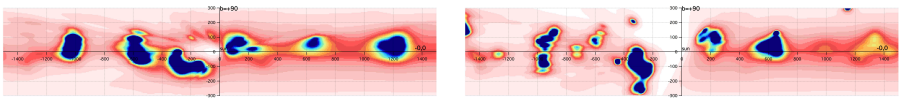
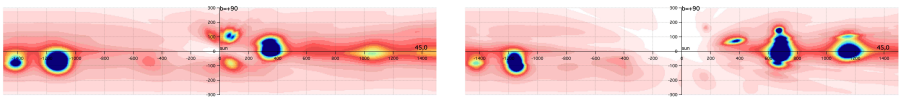


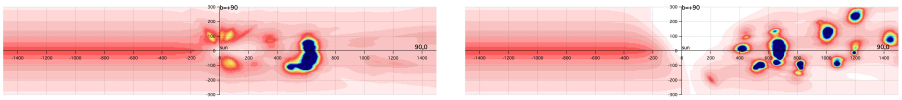
Figure 16.4 – The galactic plane for the inversion based on the extinctions, with superposed the isocontours for the map itself (red) and the map based on DIB absorptions (green). Isocontours at 0.0003 mag/pc



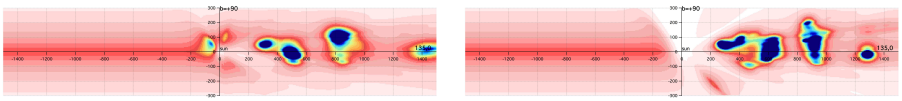
(a) l = 0°



(b) l = 45°

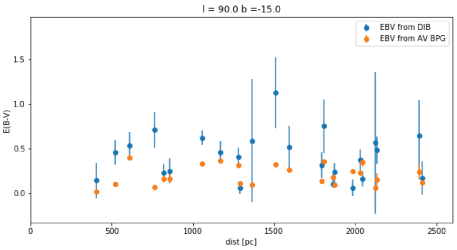


(c) l = 90°

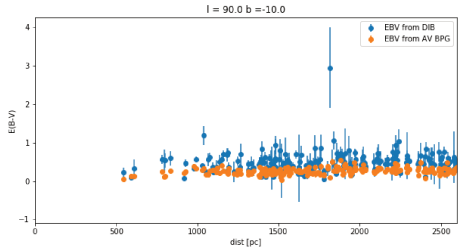


(d) l = 135°

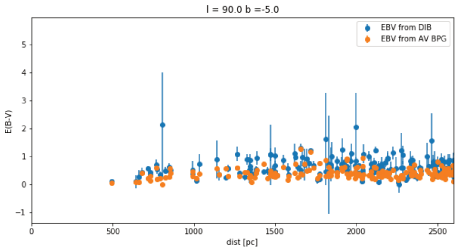
Figure 16.5 – Some vertical cut for inversion based on DIB (left column) and on extinction (right column). Different rows are respectevely: 0°, 45°, 90° and 135°.



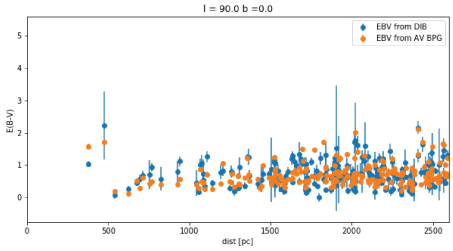
(a) b = - 15



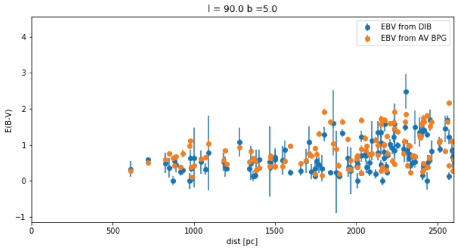
(b) b = - 10



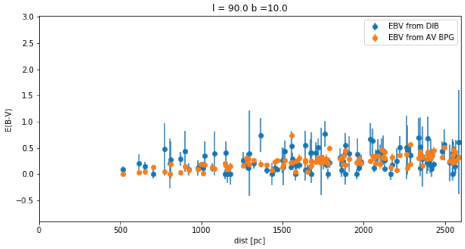
(c) b = - 5



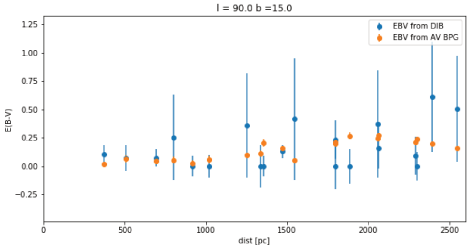
(d) b = 0



(e) b = 5



(f) b = 10



(g) b = 15

Figure 16.6 – For longitude 90°, the variation of excess of colour for distance at different latitude a) b = -15° b) b = -10° c) b = -5° d) b = 0° e) b = 5° f) b = 10° g) b = 15°.

distribution, but in this case it was a small fraction of the full dataset. If we want a pure DIB carriers distribution we need both good DIB measurements - as it is the case for the ElYajouri catalog - and better targets distribution. In the case of the Gaia DIB and the DR3 catalog, both requirements will be reached.

To improve the APOGEE DIB map in the context of the current catalog, it could be interesting to study the EW/ Av ratio for different ISM environments. Moreover, for the DIB-based map a new analytical prior could be introduced, with a different scale height and/or a different opacity on the galactic plane. Similar tests shall be done in preparation of the pure Gaia DIB.

Chapter 17

To the 3D kinematic

The position of clouds on the sky and their distances are not sufficient to have a detailed knowledge of the mutual relations between clouds. In fact, in order to understand the origin and evolution of a system of clouds, it is important to know also the velocities of the structures. The kinematics of the cloud, that is the set of the global velocities of the particles, is an information that we can obtain from the deviation from the central position of the interstellar absorptions, due to the Doppler effect. I have not done a study of the visibility of Gaia's DIB as a function of its relative velocity with the star in the background, because with GOG simulations it has not been possible to modify the relative velocity between star and interstellar absorption; however we can predict that knowing the theoretical profile of the DIB it will be possible to encode a fit profile with "multi-clouds" analysis as we did for gas lines. DIBs in general are not optimal for studying kinematics because they are very large absorptions, so it is difficult to distinguish the difference between various velocities (Puspitarini et al., 2015): however, they can be used effectively for very different velocities, such as for the different arms of the spiral.

I coded a fit profile for the sodium and potassium gas lines with a Python code, adapted for hot and cold stars, so as to have star parameters for all Gaia DR3 spectra; I studied the Taurus nebulae, using stellar spectra recorded with NARVAL and ESPaDOnS in the frame of dedicated observing programs, already performed before the beginning of my thesis and/or continued during the last years. Then, we compared the results of the resulting radial speeds with the radial speeds corresponding to the CO emission lines in the same directions; I also used the series of observations in a synthetic way and assigned to different clouds at different distances their own speeds.

The Taurus clouds complex is located in direction of the Taurus constellation, below the Galactic anticentre, between galactic longitudes 150° and 180° and galactic latitudes -5° and -50°; it is very well known as a conspicuous emissive structure in CO radio bands, as it is shown in Fig 17.1, but also in dust maps, as in the Planck maps - see at Fig 17.2. As expected, the dust and CO distributions are very similar, but not completely congruent.

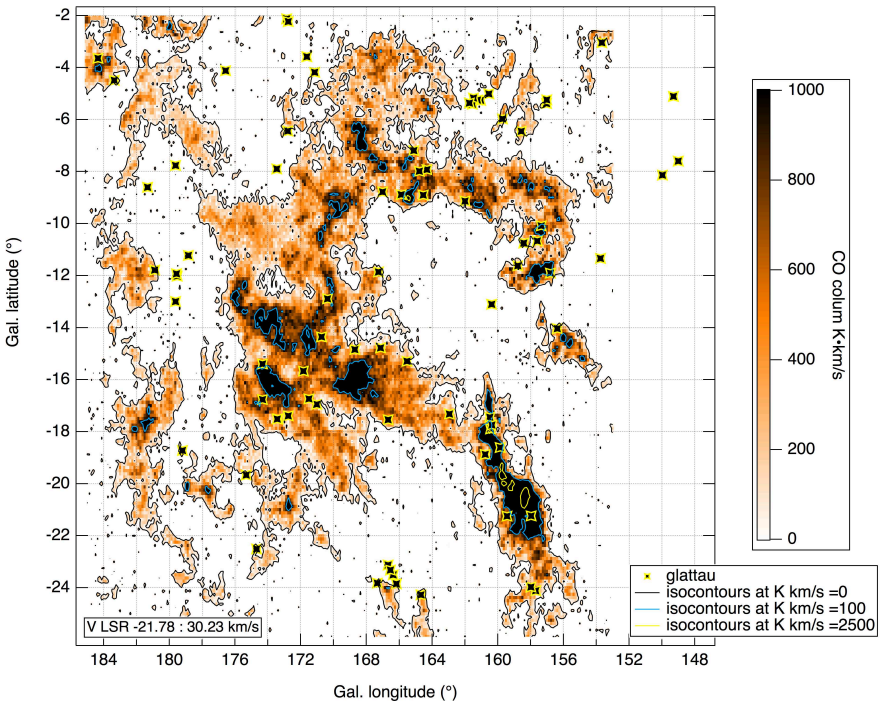


Figure 17.1 – CO map of the Taurus nebula with superposed all targets in our sample - black/yellow cross symbols. Colour scale depends on the CO column measured in radio units.

17.1 NARVAL and ESPaDOnS spectra

Stellar spectra used for the atomic lines profile fitting come from two twin instruments: ESPaDOnS (Echelle SpectroPolarimetric Devicefor the Observation of Stars) and NARVAL, respectively available at the Canada-France-Hawaii Telescope (CFHT, Mouna Kea Observatory) and the Telescope Bernard Lyot (TBL, Pic du Midi Observatory). They are two high resolution ($\sim$ 76000) spectro-polarimeters, whose first scientific motivation is the measure of stellar magnetic fields. Here, we used them in their spectrometric mode.

Spectra were measured partially on purpose for this research and others were available on the database Polarbase; PolarBase spectra are downloadable[1] in ASCII format. I selected only hot stars to avoid complex stellar spectra and faciliate the extration of IS lines. All spectra are in the heliocentric frame.

17.2 Profile-fitting and measurements of radial velocities

The implemented profile-fitting code is based on the scipy optimisation packages from the Python software. The absorption lines are modelled as Voigt functions, with for each cloud three free parameters: temperature [K], line center position (or cloud radial velocity) [km/s] and optical thickness of the absorber. Each

[1] http://polarbase-dev.irap.omp.eu/Espadons_Narval/index.php

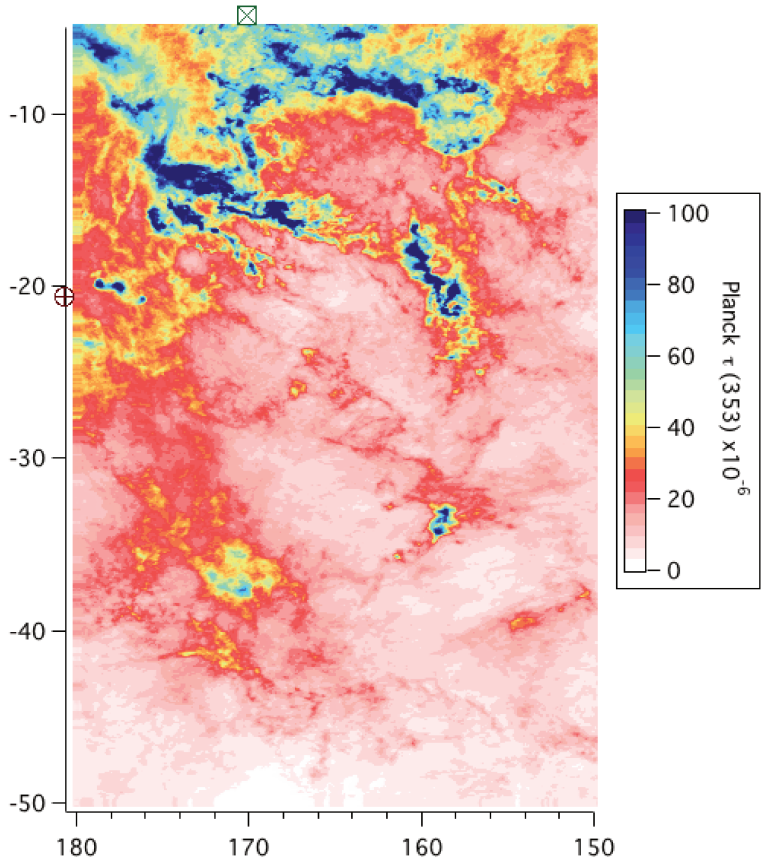


Figure 17.2 – Taurus map in galactic units of dust emissions from Planck.

spectrum is normalized. I used mainly the 769.8974 nm absorption of neutral potassium KI. This potassium absorption is part of a doublet, however I used only one of the two lines, the second (at 766.4911 nm) being coincidentally very strongly polluted by atmospheric O2 lines. Using the two lines is particularly relevant when one is interested in precise values of the optical thickness, but here I was essentially interested in velocity, and using one transition is sufficient.

We derive the Doppler shift associated with the radial velocity of a cloud from the measured difference between the absorption central velocity and the KI absorption in the rest frame (i.e. 769.8974 nm): for this reason I transformed all wavelengths to radial velocities (in km/s), with 0 km/s corresponding to 769.8974 nm. The code is adapted to fit from zero to a maximum of 4 clouds. I used a the velocity range -50 to +50 km/s, since all nearby clouds have motions smaller than 50 km/s, but in the graph is represented a wider range, to include the telluric oxygen lines; their centres vary in the heliocentric reference frame from one date to another due to the earth motion. Due to these variations, sometimes the reddest one of the telluric lines falls within our fitting range, and this could cause some misinterpretations. However, by looking at a larger wavelength interval I can easily detect these "false positives".

Particular cases among the various fits For V * AA Tau and TYC 2397-44-1 spectra, which show clearly additional stellar absorptions, I did not calculate a stellar synthetic profile, but only added "by hand" a Gaussian component to the fit. The results of these fits are uncertain. For TYC 2864-514-1, HD 279953 and HD 27923 I allowed for 5 absorbing clouds, but only in the TYC 2864-514-1 spectra the fitting code converges with five components. For 2MASS J04282605+4129297, I constrained the fit within -10 and +50 km/s to avoid the stellar components in the bluer part of the spectra.

I show some examples of fit results in Fig 17.3. The whole set of figures showing the fits are in Appendix B with the Table of the fit results. The spectra are in the heliocentric reference frame (more precisely the BSR, see below), and fits are done in this reference. Note that, after the fit, all velocities are transformed to be in the same system as CO observations, i.e. the LSR frame in its kinematic definition (see also below).

Velocity reference frames Ground-based optical observations use the so-called barycentric standard-of-rest (BSR) that considers the barycentre of the Solar System as the reference point. Since the difference between barycentric and heliocentric velocities is extremely small and negligible in most cases, the BSR frame is also often referred to as the heliocentric standard-of-rest (HSR) frame.

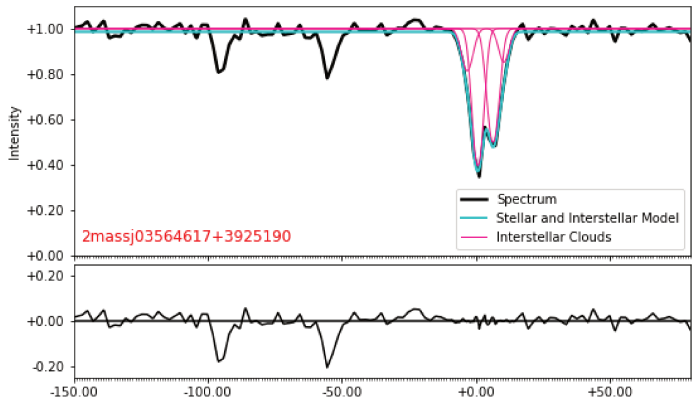
For objects very far from the Sun, the peculiar solar motion with respect to the general quasi-circular motion of Milky Way stars and gas in our Galaxy corner is taken into account and removed to obtain the radial velocity in the Local Standard Rest (LSR) frame. This is the commonly used frame for radio velocities, in particular the CO velocities used in our investigation. The difference between the LSR and BSR velocities depends on the position on the sky defined by Galactic coordinates glon, glat, and it is obtained from the following calculation:

```
w1 = cos(glon) cos(glat)
w2 = sin(glon) cos(glat)
w3 = sin(glat)
glon_sun = 56.1602 π/180
glat_sun = 22.762 π/180
sun1 = cos(glon_sun) cos(glat_sun)
sun2 = sin(glon_sun) cos(glat_sun)
sun3 = sin(glat_sun)
δ v = (w1 sun1 + w2 sun2 + w3 sun3)·20
```

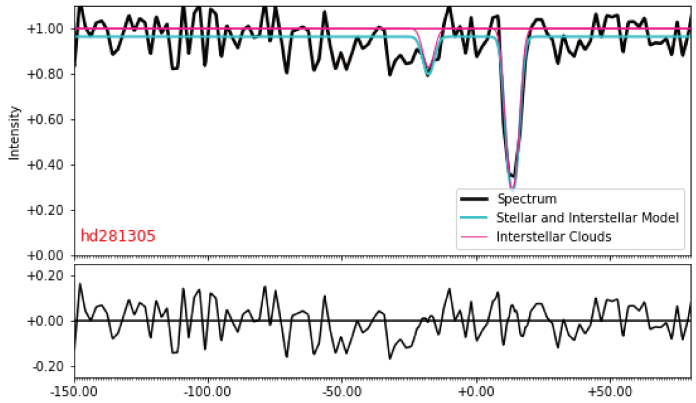
i.e., it assumes that Sun moves at 20 km/s towards (glon, glat)= (56.1602,22.762). I converted all interstellar medium absorption velocities into LSR velocities instead of converting CO velocities into BSR ones, because it was simpler to change several individual values than a full spectral cube.

17.3 3D map in the Taurus region

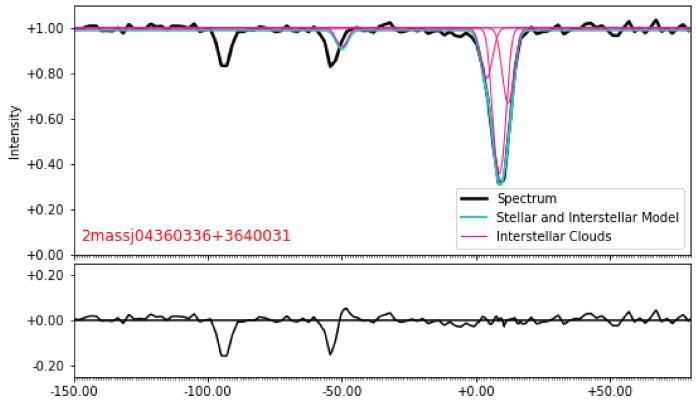
In this section, I show some vertical cuts in the opacity map produced by Jean Luc Vergely, that I presented in Chapter 15. Superimposed are the lines-of-sight to the analysed stars, and the observed velocities. The colour scale, mag per pc, is the same for all figures and green, yellow, pale orange, and reds



(a) HD 275941



(b) HD 281305



(c) HD 280006

Figure 17.3 – Example of fit for three spectra. The two absorptions in the bluer part are always telluric lines of oxygen. Black line: observation. Pink lines: different velocities components. Light blue: total model. Top) Good fit. Middle) Example of fit with a stellar component at -17 km/s. Bottom) Example of fit with a telluric component at the border of the fit interval.

isocontours lines are respectively (0.005, 0.002, 0.001, 0.00043, 0.0002) mag pc-1. We draw the attention to the fact that, here, the colour scale does not refer to excess of colour/pc, but A_V/pc - different from the other maps in the thesis that I produced by inversion. Since we consider that extinction and reddening are simply proportional, the maps are not different.

Following there are a series of figures with annotated velocities: these are the velocities assigned to the structures as suggested by the whole set of stellar velocities and the geometry of the lines-of-sight. They were derived from logical deductions, and not from automatic computation. We do not have a uniform distribution of stellar targets for the fit profile and more numerous and accurate assignations will be given in future thanks to new absorption data.

17.4 CO velocities maps

We used 12 CO J= 1:0 observations at 115 GHz from the CfA survey of the Galactic plane (Dame et al., 2001; Dame and Thaddeus, 2004). For Taurus the survey is based on a 0.125° x 0.125° Cartesian grid and the LSR velocity step is 0.65 km/s.

In this section I thoroughly analyse the region between 155° and 180° using our 3D extinction maps, the CO emission spectral maps and the spectral information on interstellar potassium absorption to assign position and kinematics to the interstellar clouds. My decomposition of velocities for CO is based on a work contained in the thesis of Quentin Remy, in particular his Fig 12 (private communication), pag 18.

Comparing our 3D map, the atomic absorptions and the CO emissions it is possible to identify molecular clouds positions and surrounding atomic gas velocities.

Main Taurus

Main Taurus is a complex cloud from 178° to 160° in longitude and from -3° to -20° in latitude. It is characterised by velocities from 4 km/s to 10 km/s. It is closer than California, and I tried to identify both the beginning and the end of the complex thanks to the spectra of stars at different distances and the 3D map. In particular Fig 17.15, 17.14, 17.12, 17.11, 17.10, 17.9, 17.8 are the vertical slices in the map, and I can use them to understand the evolution in distances. It is evident that in 3D there are various structures in these directions.

In Main Taurus direction we have a lot of targets but all closer than 500 pc. In Fig 17.24 and 17.24 are shown velocities and distances for stellar targets in this region, globally and separated for the different velocities. In the central core of the cloud (between 175° and 165° longitude) I found several absorption components; for stars at distance lower than 176 pc only a component is present, but for more distant stars than this value I found two components, coherent with the range. In direction from 167° and 162°, there are four stellar targets and they have wide range of distances (from 176 to 500 pc), but only one component.

From Fig 17.18 and 17.19, there is a coherence with the higher velocities around 10 km/s for these longitude as in Fig 17.25 bottom. Moving in the

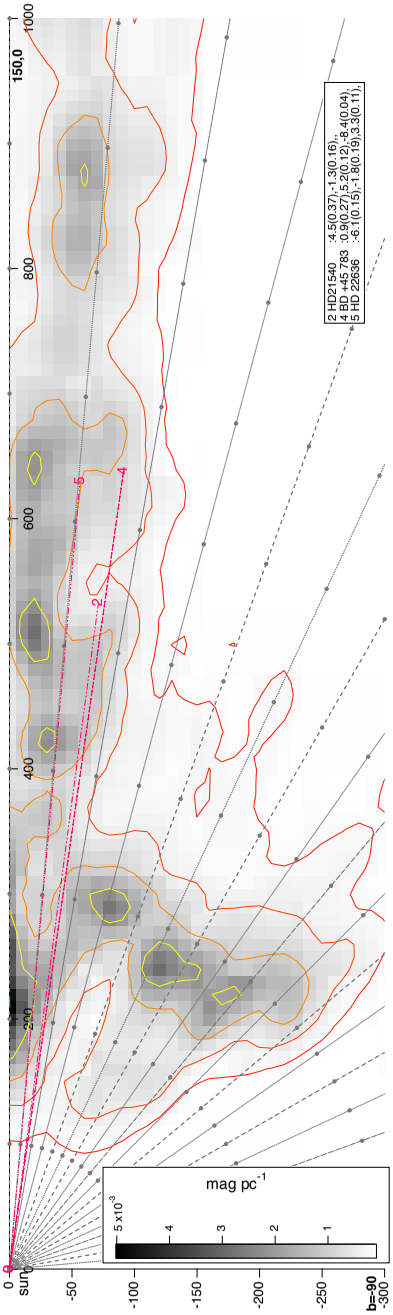


Figure 17.4 – Slice at longitudes 150° from the 3D map, with stellar targets closer then 2° at the plane.

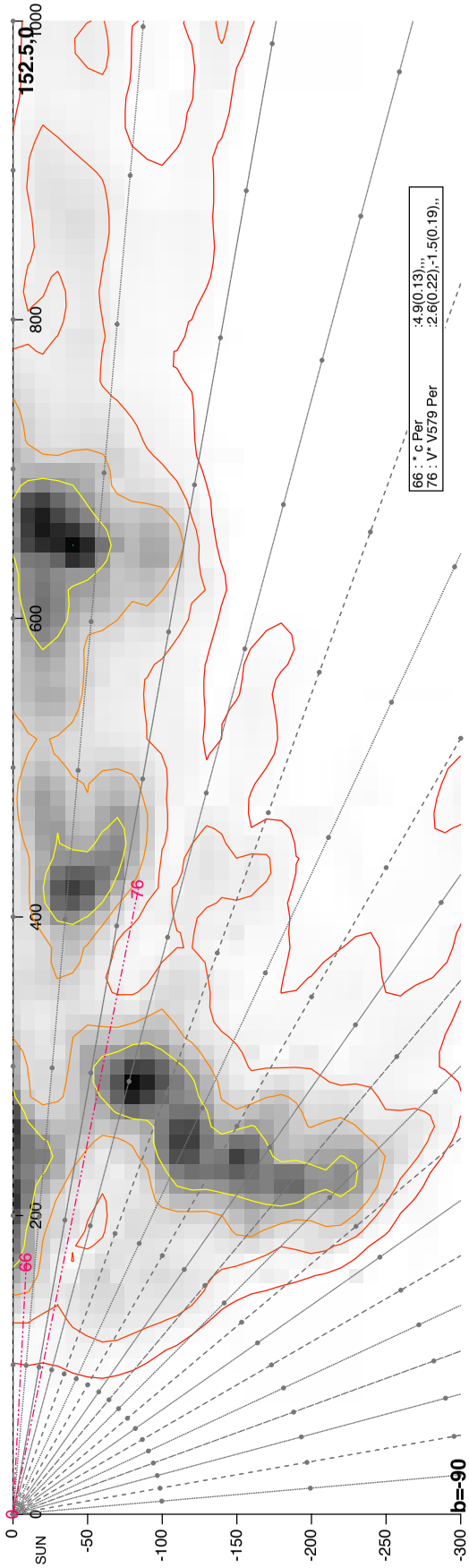


Figure 17.5 – Slice at longitudes 152.5° from the 3D map.

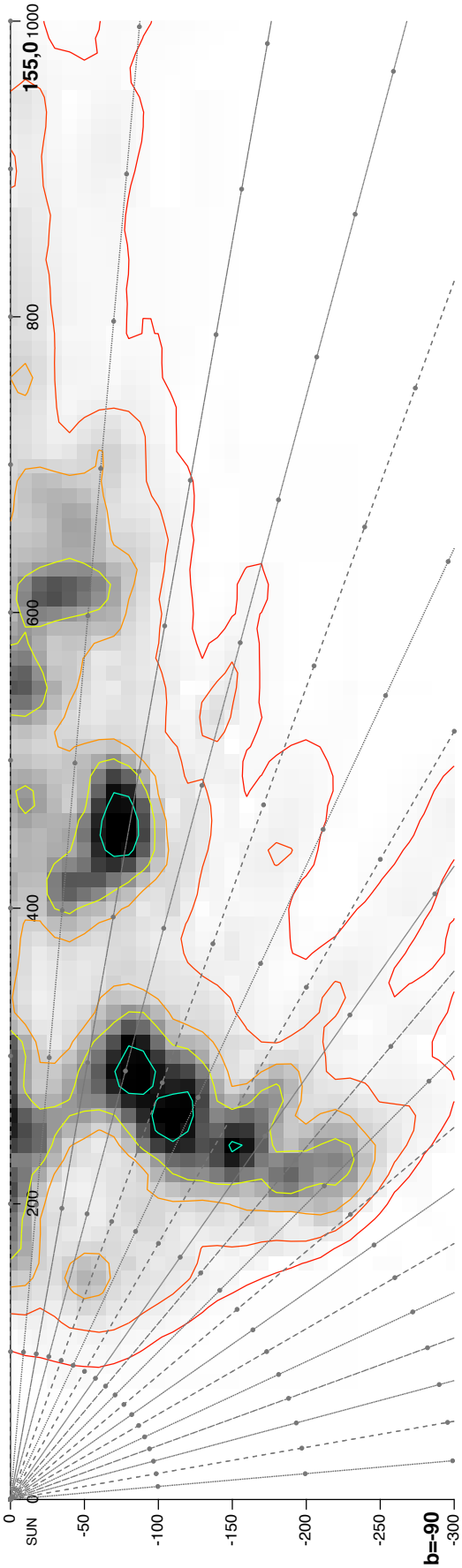


Figure 17.6 – Slice at longitudes 155° from the 3D map.

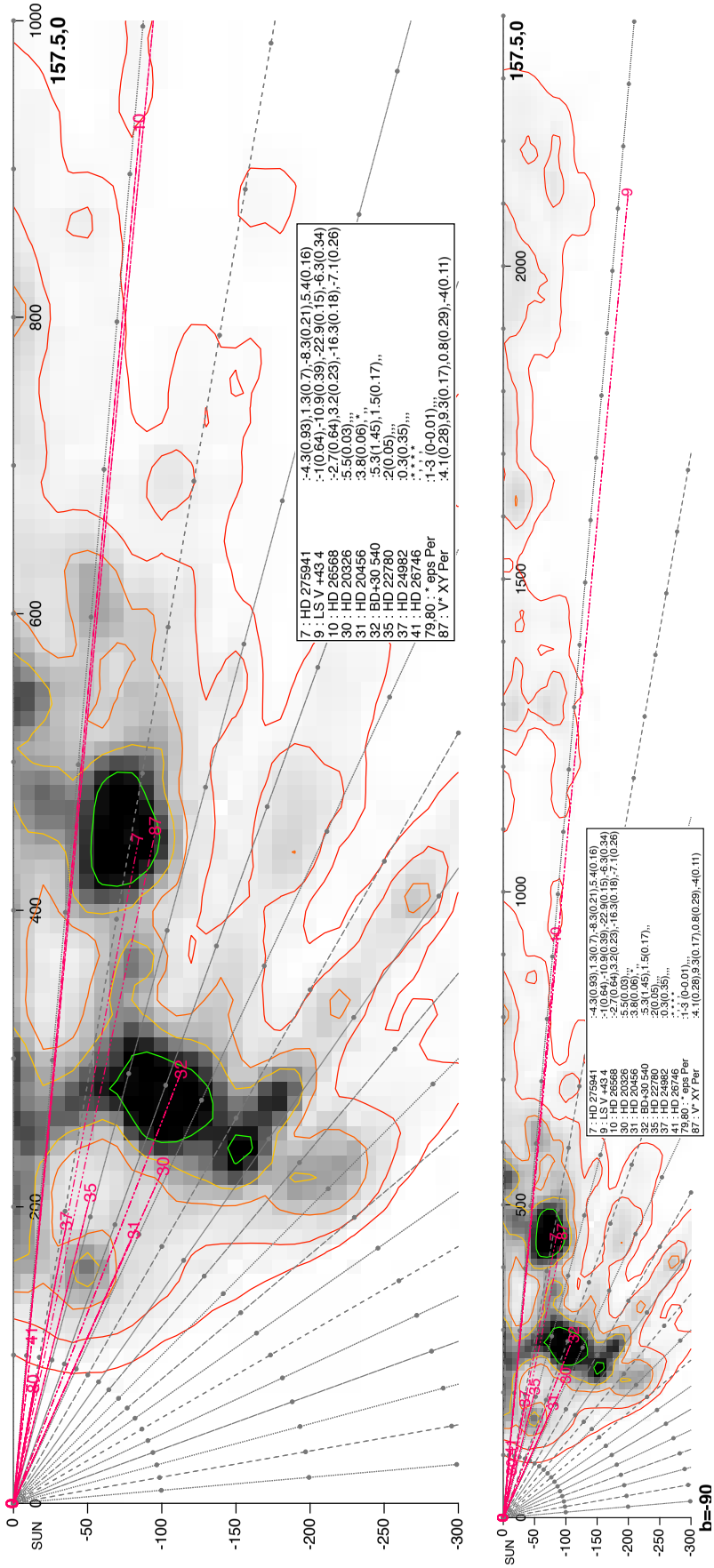


Figure 17.7 – Slice at longitudes 157° from the 3D map.

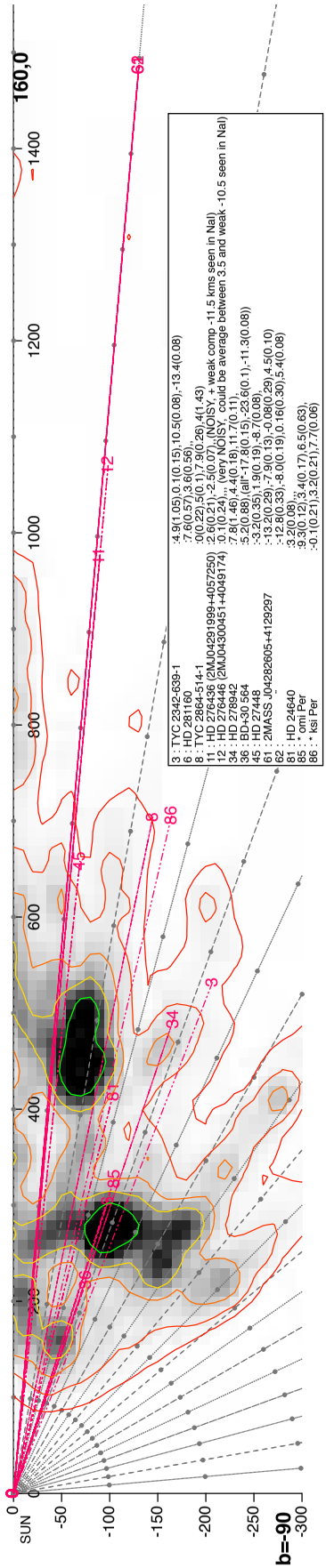


Figure 17.8 – Slice at longitudes 160° from the 3D map.

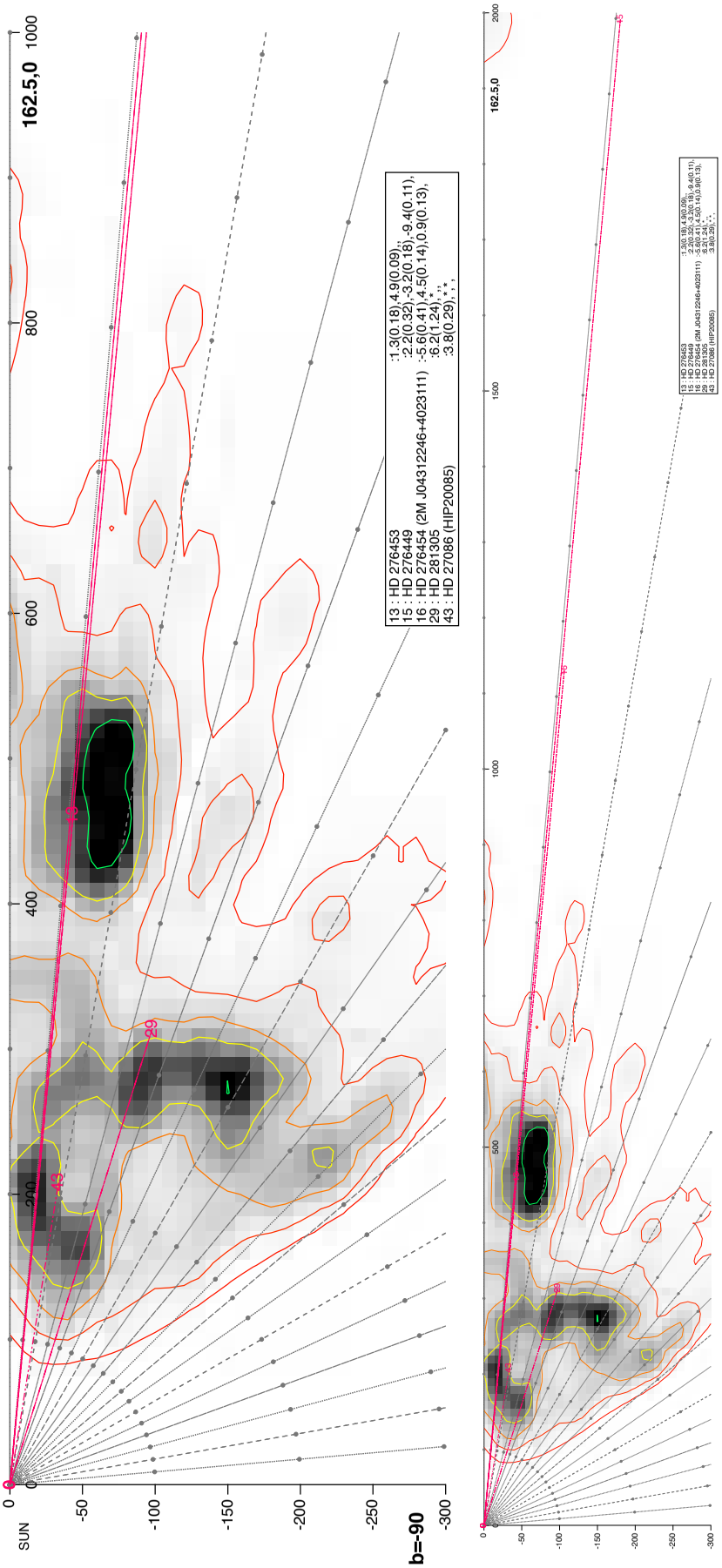


Figure 17.9 – Slice at longitudes 162.5° from the 3D map.

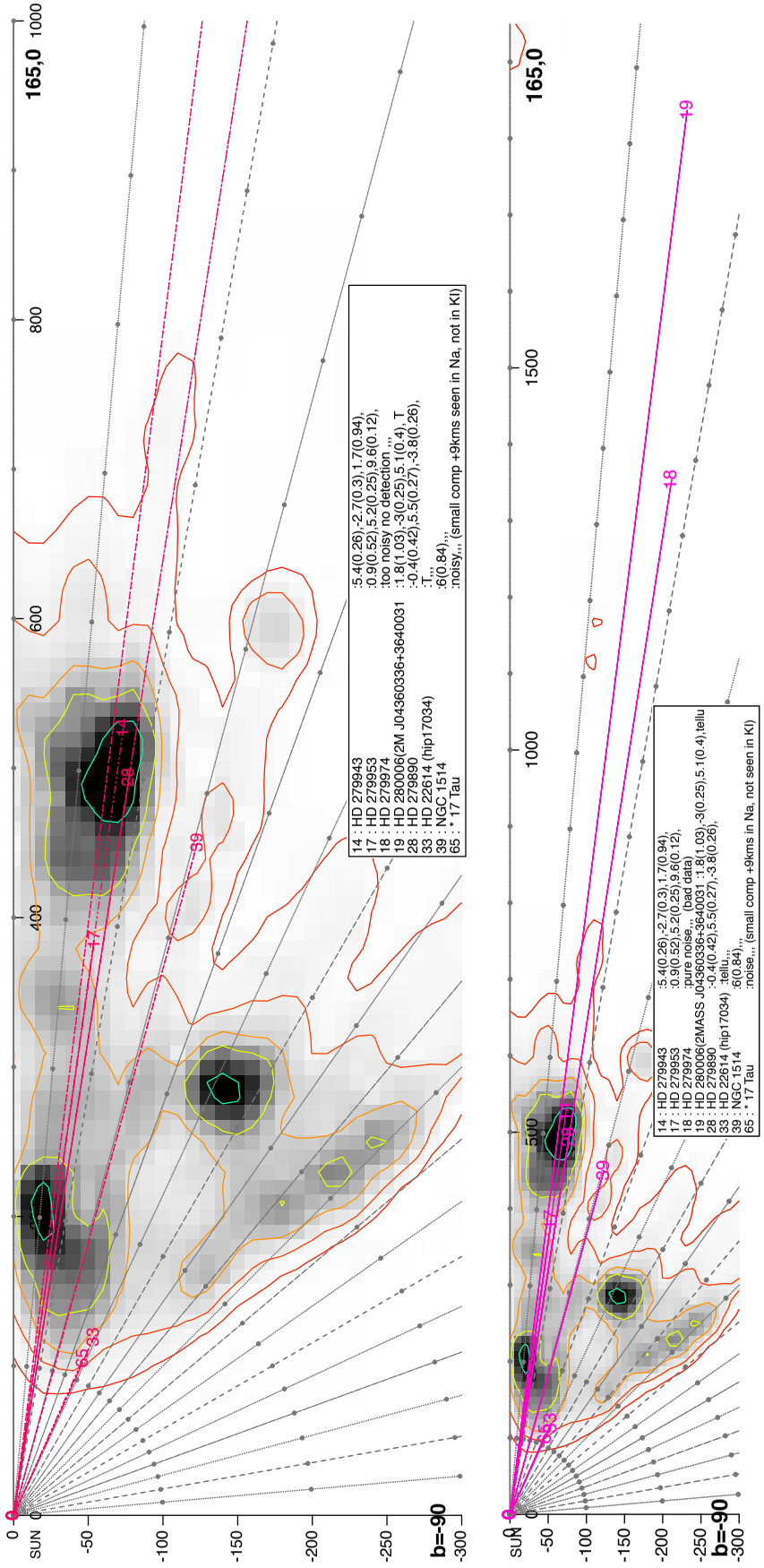


Figure 17.10 – Slice at longitudes 165° from the 3D map.

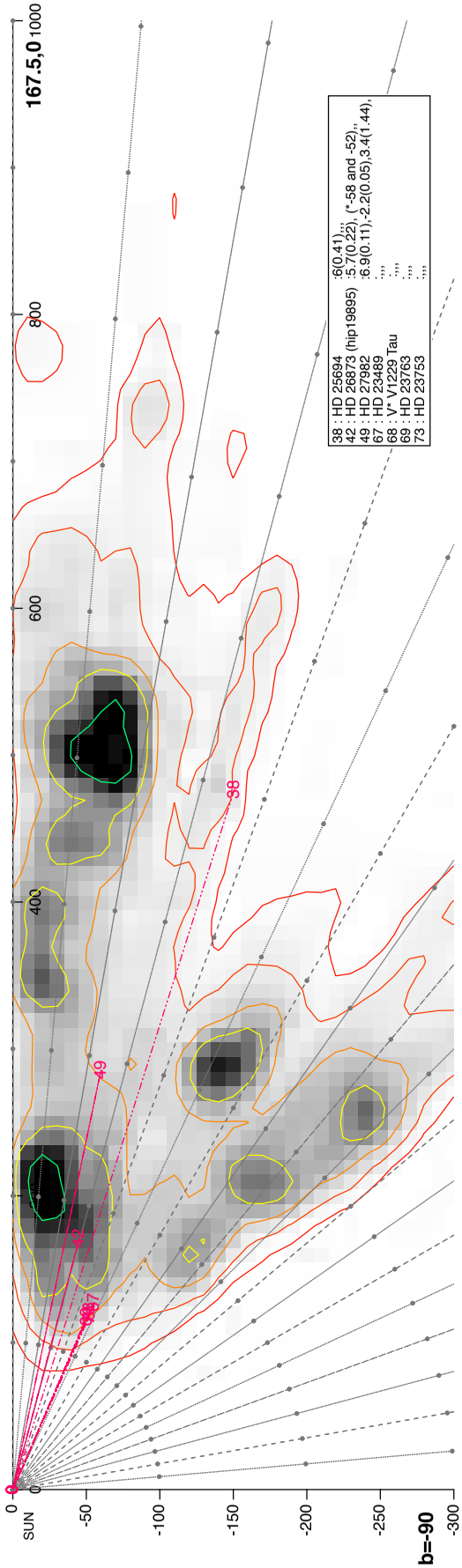


Figure 17.11 – Slice at longitudes 167.5° from the 3D map.

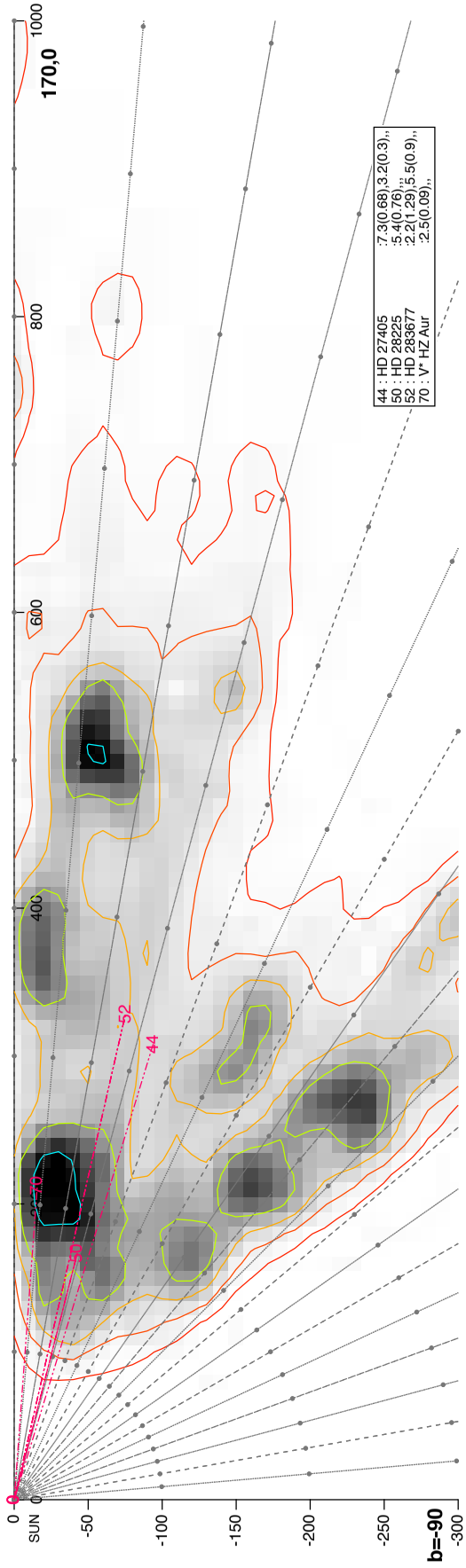


Figure 17.12 – Slice at longitudes 170° from the 3D map.

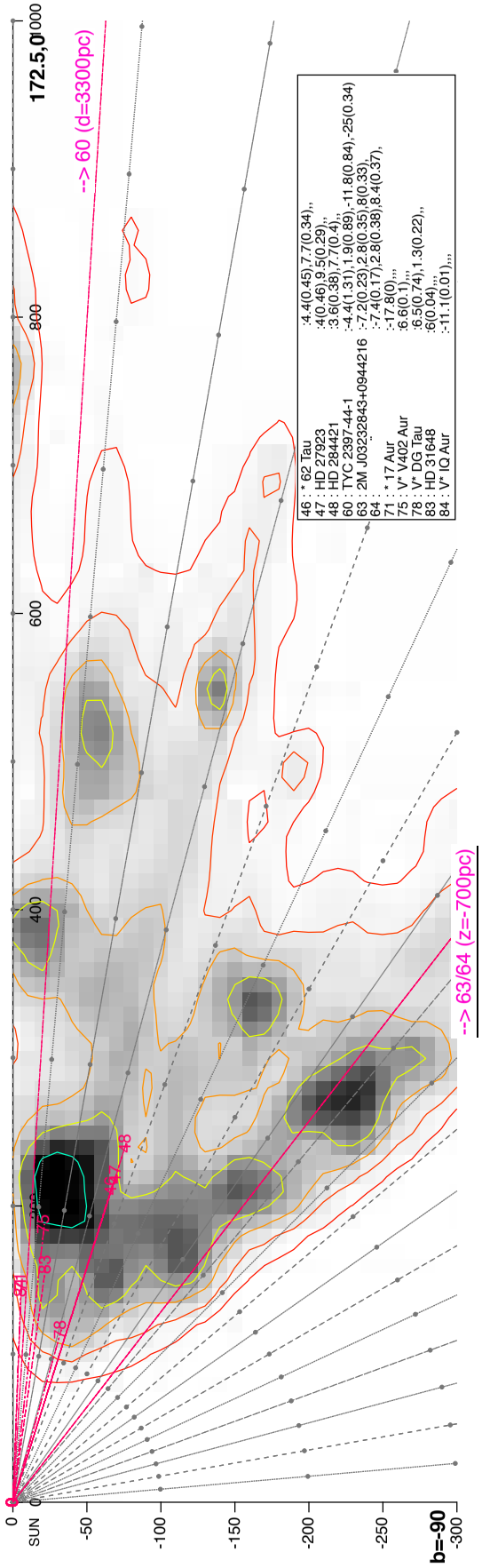


Figure 17.13 – Slice at longitudes 172.5° from the 3D map.

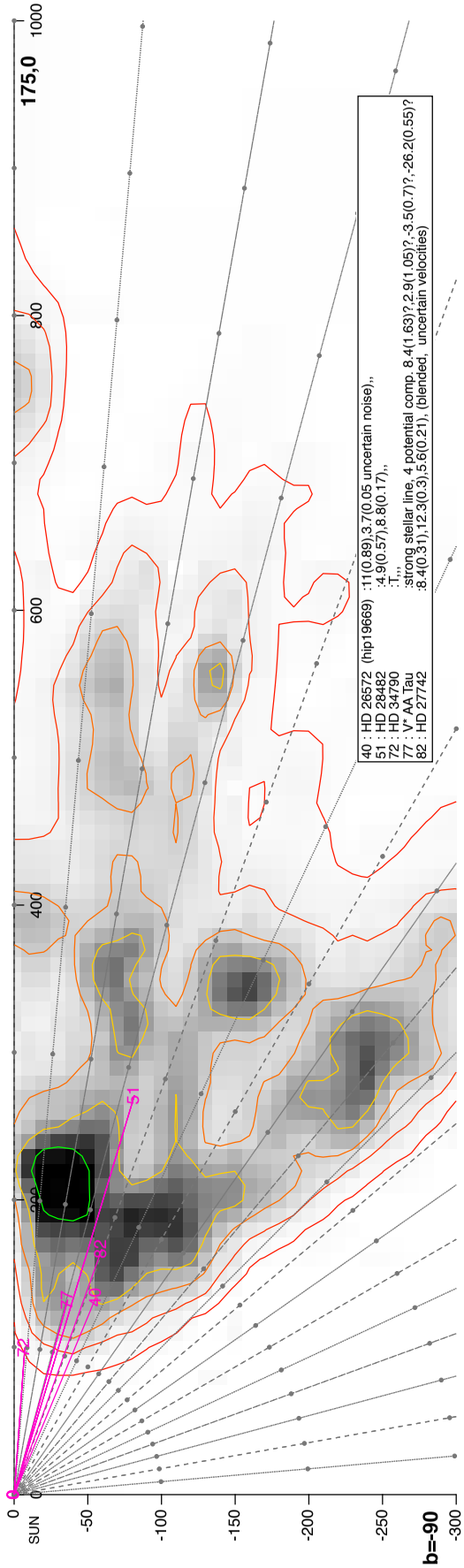


Figure 17.14 – Slice at longitudes 175° from the 3D map.

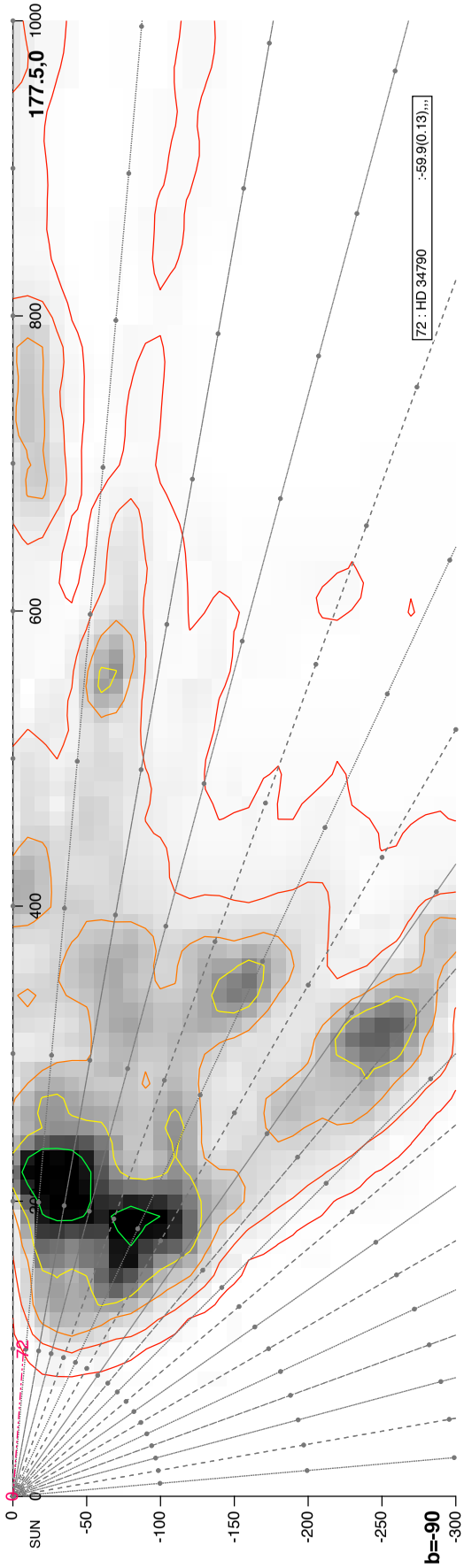


Figure 17.15 – Slice at longitudes 177.5° from the 3D map.

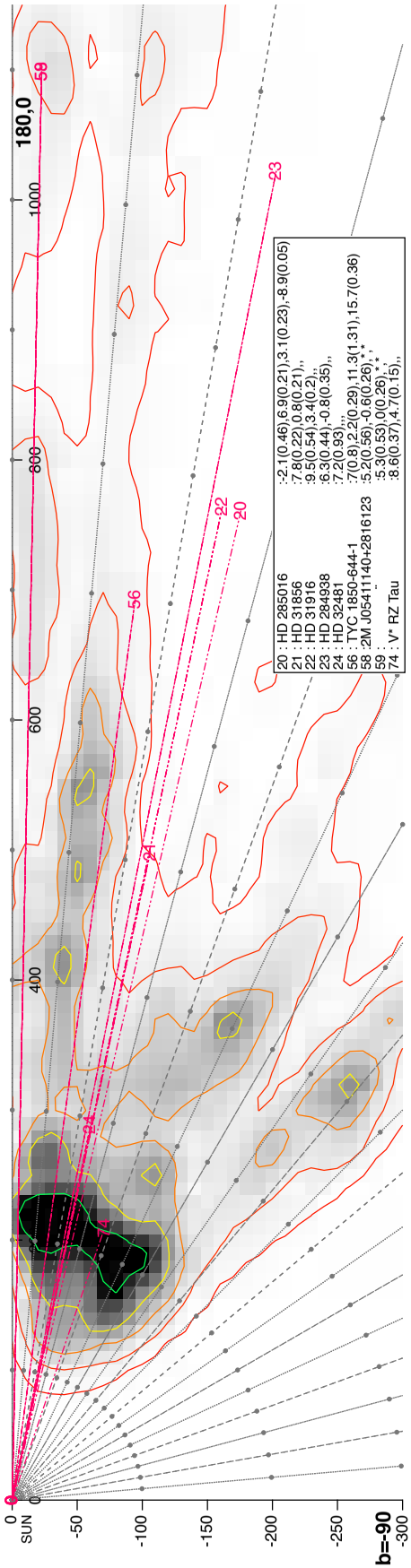


Figure 17.16 – Slice at longitudes 180° from the 3D map.

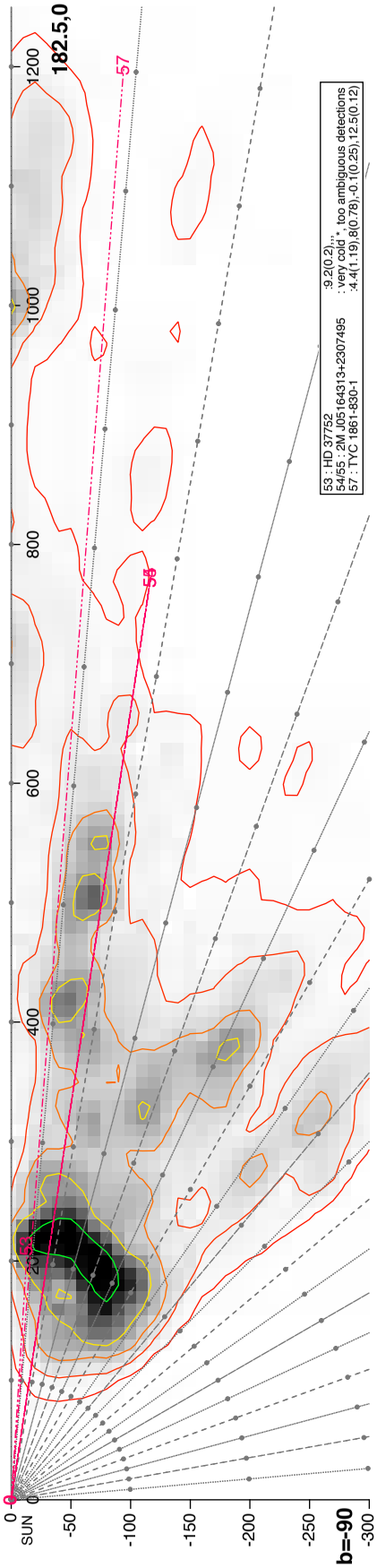


Figure 17.17 – Slice at longitudes 185° from the 3D map.

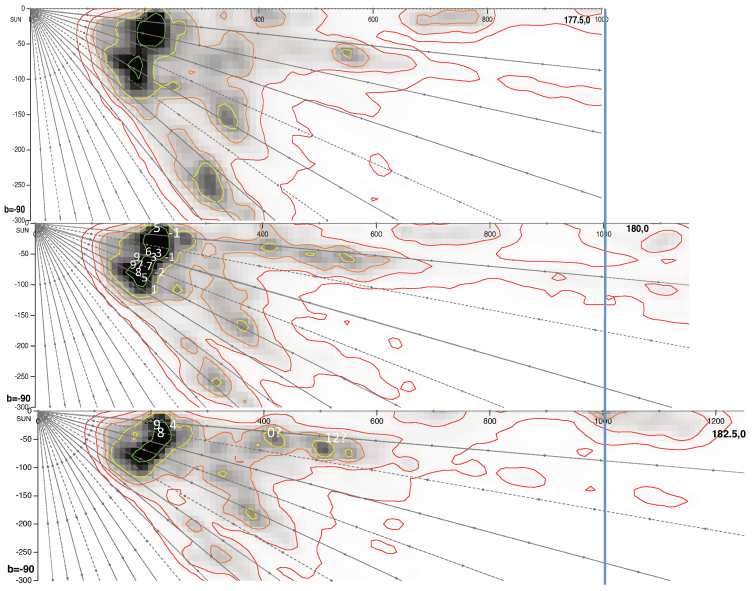


Figure 17.18 – Slice at longitudes 177.5°, 180° and 182.5° from the 3D map, with the velocities assigned at 3D clouds, as suggested by the interstellar K absorptions.

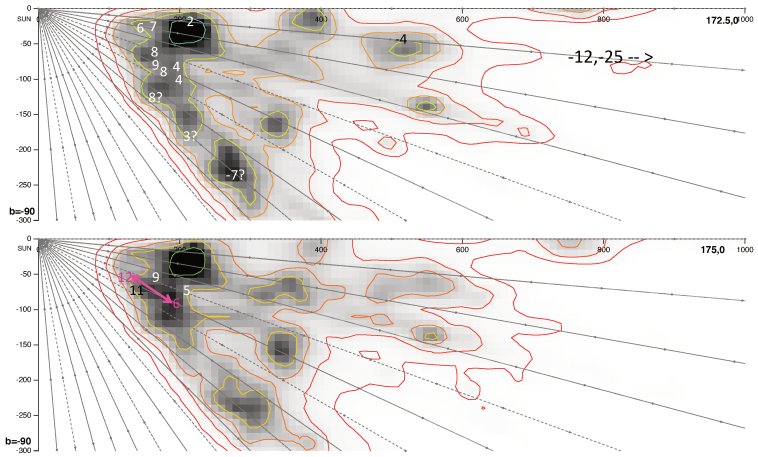


Figure 17.19 – Slice at longitudes 172.5° and 175° from the 3D map, with the velocities assigned at 3D clouds, as suggested by the interstellar K absorptions.

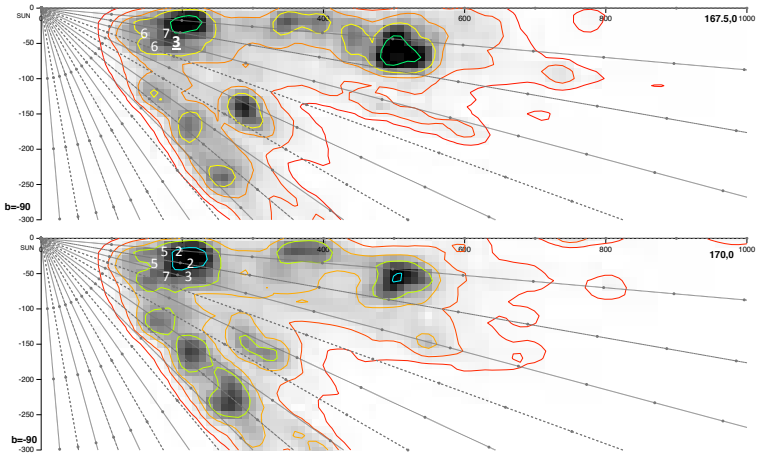


Figure 17.20 – Slice at longitudes 167.5° and 170° from the 3D map, with the velocities assigned at 3D clouds, as suggested by the interstellar K absorptions.

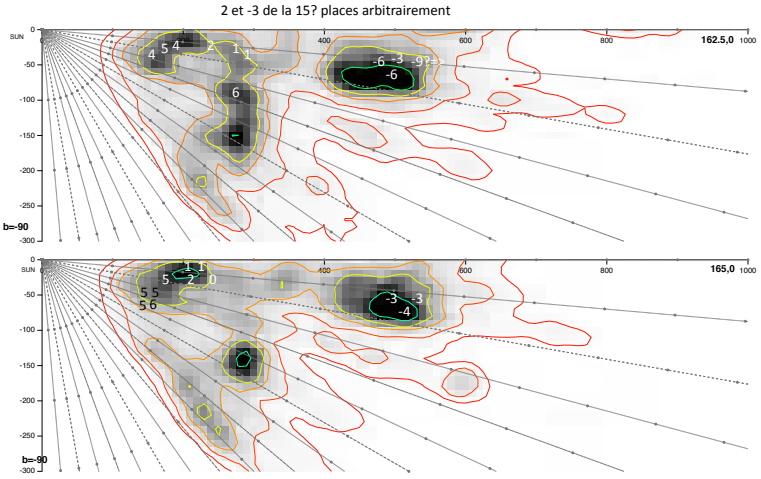


Figure 17.21 – Slice at longitudes 162.5° and 165° from the 3D map, with the velocities assigned at 3D clouds, as suggested by the interstellar K absorptions.

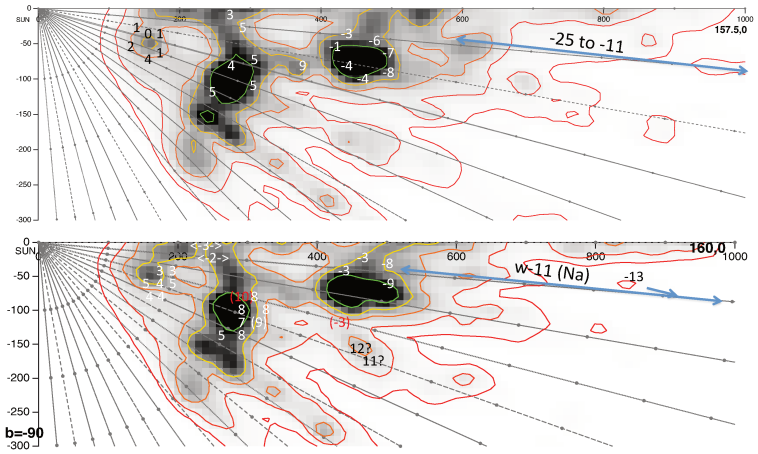


Figure 17.22 – Slice at longitudes 157.5° and 160.0° from the 3D map, with the velocities assigned at 3D clouds, as suggested by the interstellar K absorptions.

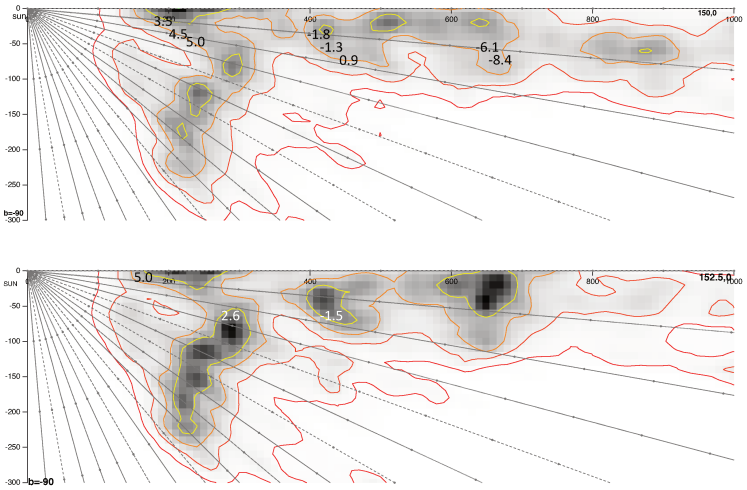


Figure 17.23 – Slice at longitudes 150° and 152.5° from the 3D map, with the velocities assigned at 3D clouds, as suggested by the interstellar K absorptions.

lower longitudes, velocities from atomic line are lower. In Fig 17.22, we assign at the closer structure at ∼200 pc velocities around 1 km/s, but there are higher velocities associated at a more distant structure at 300 pc, at same latitude. Fig 17.24 b) shows the CO emission for velocities in the interval (3-4) km/s and from our measures, the emission molecules are in a molecular cloud at 150 pc.

South Taurus

Properly called South Taurus Clouds have CO velocities from 5 to 15 km/s, and are located between l = (175°, 170°) and b = (-20°, -25°). In that direction, we have only 8 stellar targets, and the velocity puzzle seems to be quite simple to understand. The part with velocities higher than 10 km/s are in the lower latitudes, as clear in Fig 17.26. In particular HD 26572 arranges the higher velocity part behind 144 pc, while HD 27742 positions the slower part behind 177 pc. It is very interesting because in Fig 17.14 b) there is a core in 175° with latitude lower then 20°, that I can identify as the source of this absorptions. Using the vertical cut, Fig 17.13 and 17.12, I can identify the cloud South Taurus between 150 and 220 pc.

I include in this paragraph also a small cloud with same velocities but longitude (167°, 165°). Here we have 5 stars near each others in position and distance and only the furthest stars HD 23489 at 139 pc has a very small absorption: I cannot propose it as the border of the cloud and only suggest that this fragment is behind 140 pc, Fig 17.11 and 17.10.

Perseus

Perseus Cloud is characterised by a velocity interval in CO from 0 to 8 km/s, in l = (162, 157) and b = (-25, -20) and we have four stellar targets in that region.

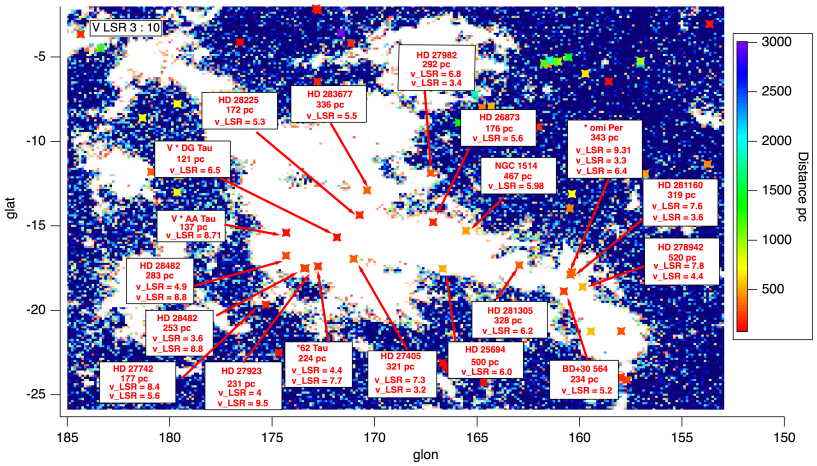
In Fig 17.21 there are slices to understand: the so-called Perseus cloud in CO is actually two clouds, first one between 150 and 200 pc ("Near Perseus") and the second one from 260 and 310 pc ("Far Perseus"). The more distant cloud show its core in Fig 17.9, 17.8 and 17.7. This cloud in not exactly between b = (-25, -20), but (-25, -15). From -20 and -15, for the CO classification the emission from 0 to 8 km/s belongs to the Main Taurus, but is is more coherent to associate it with "Far Perseus".

Auriga

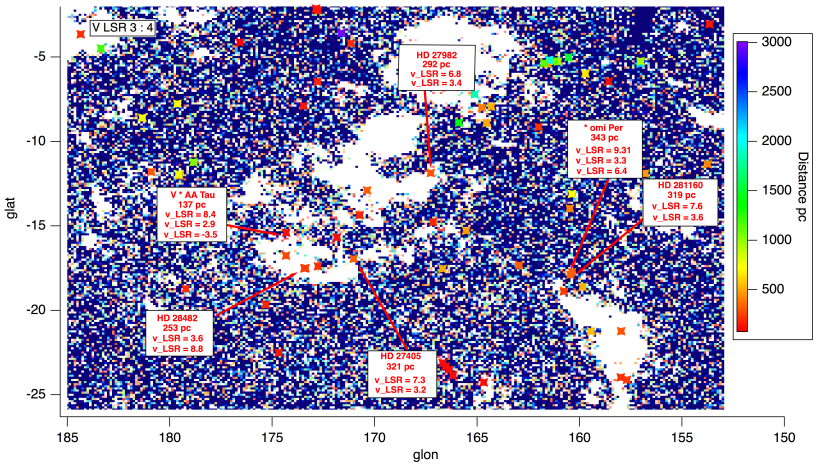
Auriga Cloud is characterised by a velocity interval from 0 to 6 km/s, near the Galactic Plane. Its position on the sky is longitude (170°, 163°) and latitude from -5° to -7.5°. We have not stars well centred on Auriga but in the direction with the border Auriga/California. Far star HD 279974 at 1370 pc has not a good spectra, so I did not use it.

HD 27086 has a absorption in the Auriga range but not in the California (see later) velocity range, so Auriga clouds should begin closer then 210 pc and only beyond there is California.

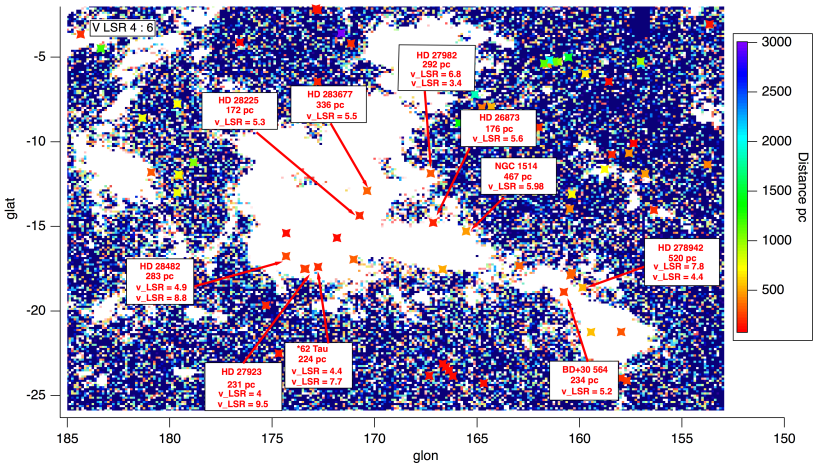
Using HD 280006 and HD 279943, it seems that the three components are yet present closer than 531 pc, and beyond we have not important dense clouds.



(a) Main Taurus

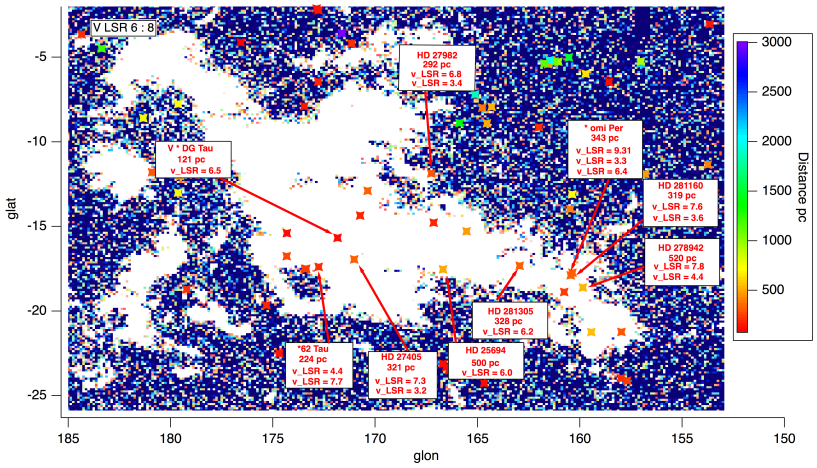


(b) Main Taurus, velocities 3-4 km/s

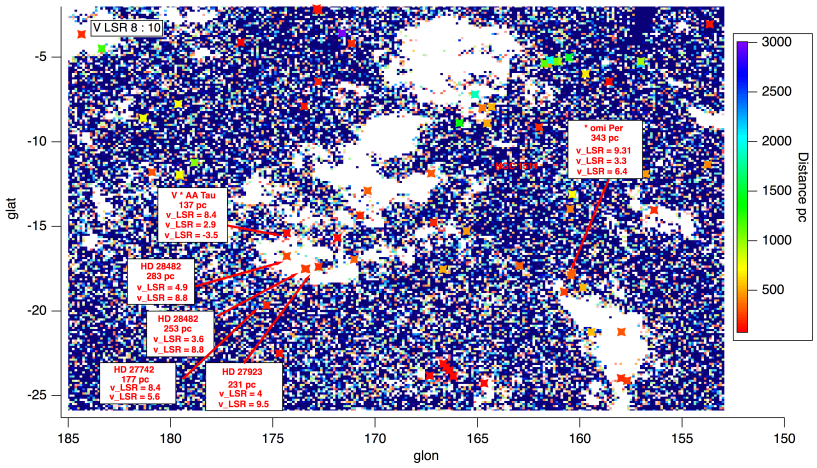


(c) Main Taurus, velocities 4-6 km/s

Figure 17.24 – Structure in CO maps at for the velocity range of Main Taurus, total on the top and separated by velocities. Colour scale of the stellar targets represents the distance.

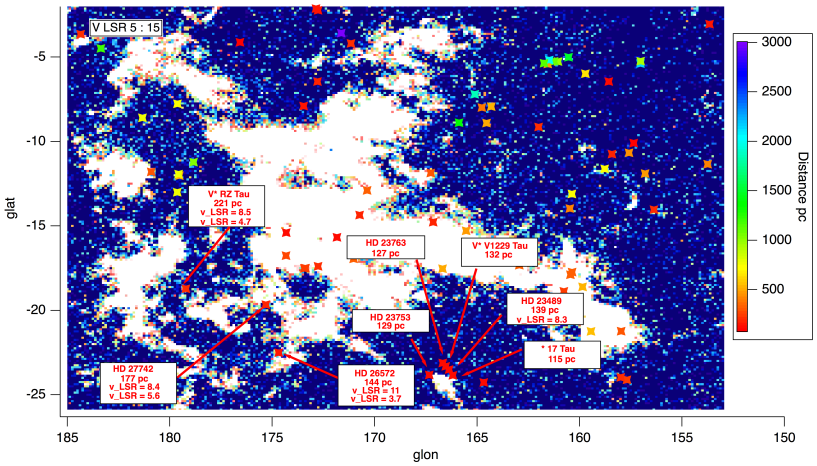


(a) Main Taurus, velocities 6-8 km/s

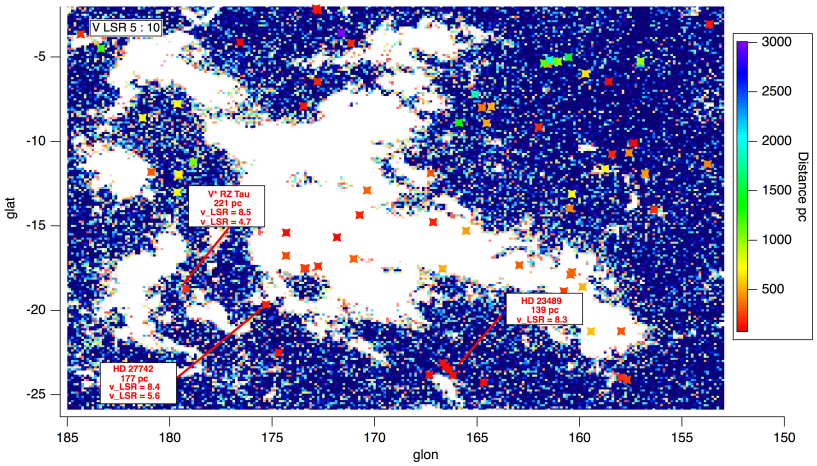


(b) Main Taurus, velocities 8-10 km/s

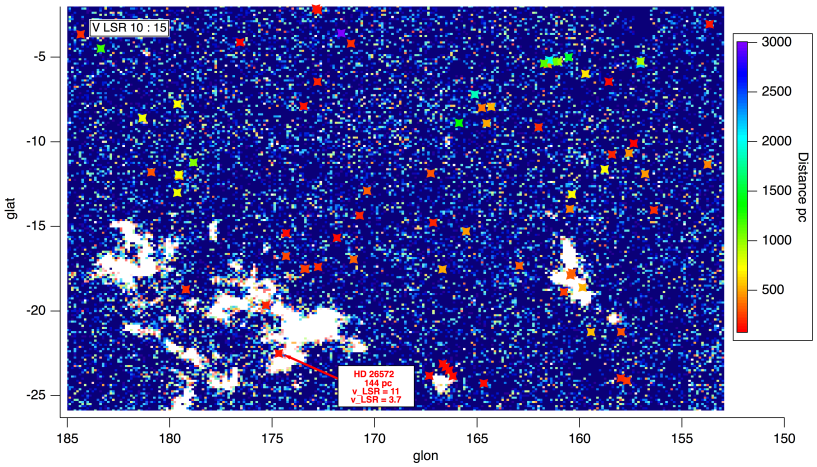
Figure 17.25 – Structure in CO maps at for the velocity range of Main Taurus, separated by velocities. Colour scale of the stellar targets represents the distance.



(a) South Taurus

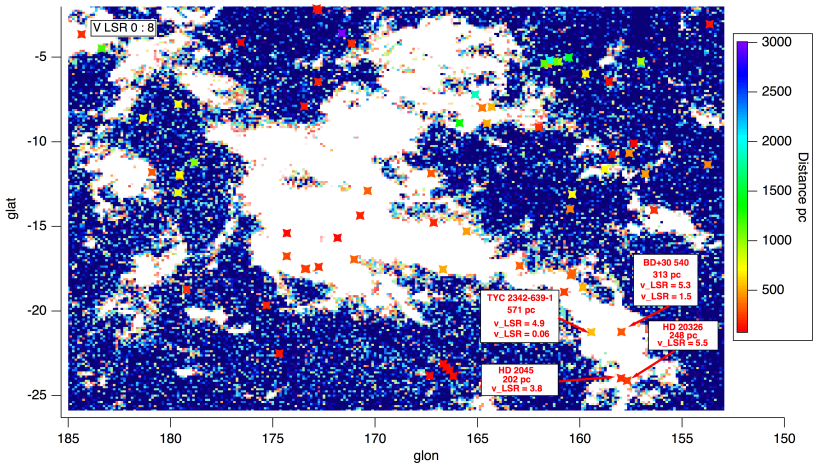


(b) South Taurus, velocities 5-10 km/s

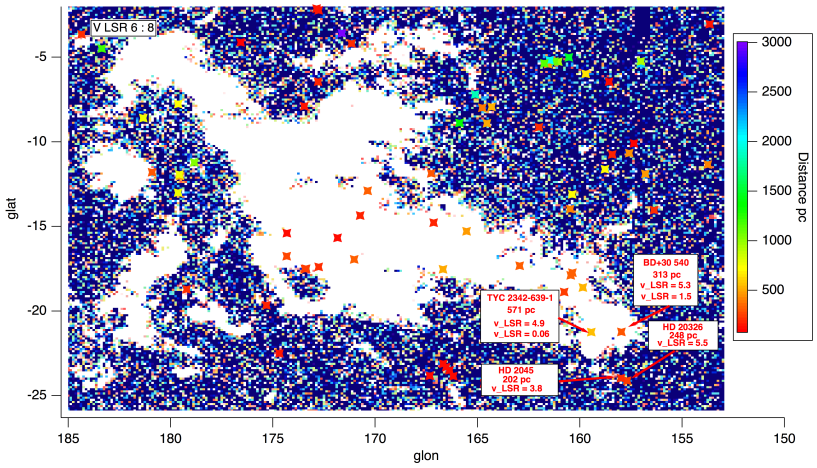


(c) South Taurus, velocities 10-15 km/s

Figure 17.26 – Structure in CO maps at for the velocities range of South Taurus, total on the top and separated by velocities. Colour scale of the stellar targets represents the distance.

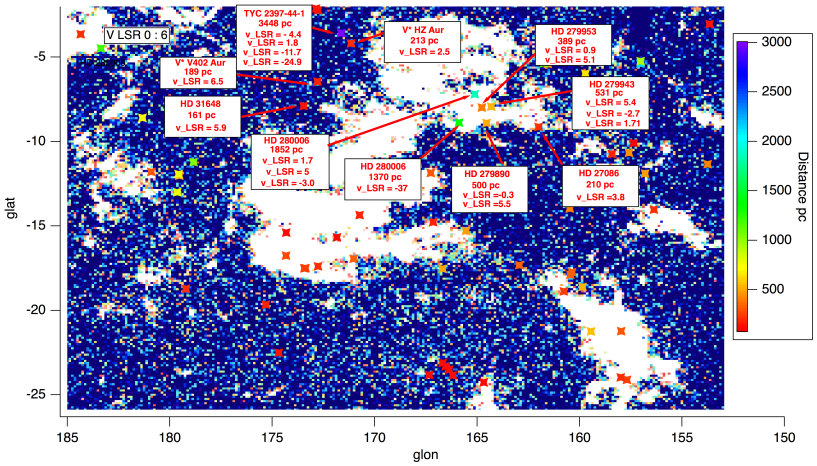


(a) Perseus

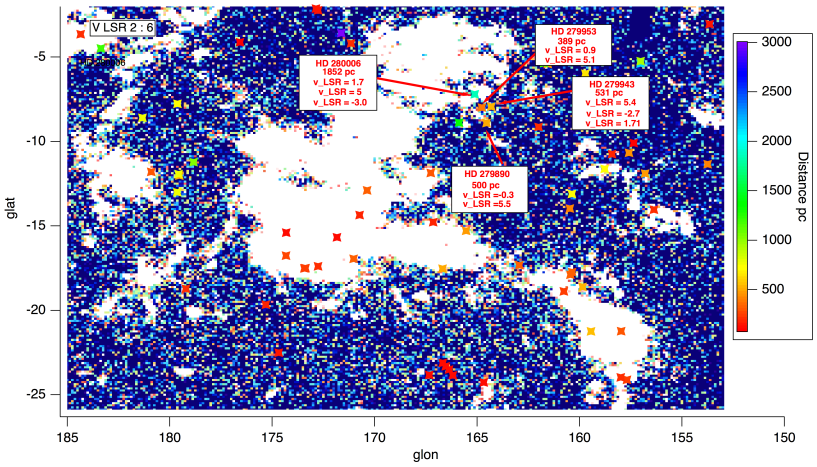


(b) Perseus, velocities 6-8 km/s

Figure 17.27 – Structure in CO maps at for the velocity range of Perseus region, total on the top and separated by velocities. Colour scale of the stellar targets represents the distance.



(a) Auriga

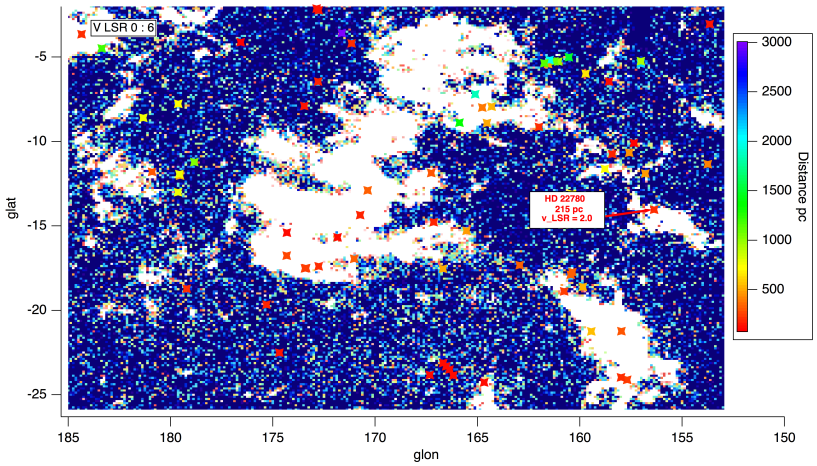


(b) Auriga, velocities 2-6 km/s

Figure 17.28 – Structure in CO maps at for the velocity range of Auriga region, total on the top and separated by velocities. Colour scale of the stellar targets represents the distance.

Looking at the Fig 17.20 and 17.21, the only structure with compatible in longitude and latitude is located at 200 pc; it is a small elongated over-density in 5° direction from the core to 130 pc, and it could cause the weak absorption for stars HD 31648, V* V402 Au and HD 3164. This structure has not a CO corresponding part, but those stars has absorptions at gas velocities corresponding with Auriga's ones.

From the 3D map and the interstellar absorptions available, it seems that Auriga is a part of a close cloud which correspond largely with Main Taurus, simply closer to the Galactic Plane.



(a) North Taurus

Figure 17.29 – Structure in CO maps at for the velocity range of North Taurus region. Colour scale of the stellar targets represents the distance.

North Taurus

In North Taurus region, Fig 17.29 we have only stellar target, HD 22780 . Properly called North Taurus Clouds has CO velocity from 0 to 6 km/s, and it is between l = (157°, 153°) and b = -15°. In Fig 17.7 d) and 17.22 there is an important structure from 200 and 300 pc and in Fig 17.6 it has one of its core at latitude 15°. I can assign North Taurus at this structure. The North Taurus and Perseus are connected in the 3D: they have different position on the sky, but in 3D and for the velocities they are coherent.

California

California is a molecular cloud in the Taurus complex, in the interval l = (155°, 170°) and b = (-12.5°, -7.5°). Its velocities are between -10 and 0 km/s and it could be decomposed in four parts, which respectively velocities from -8 to 6.5 km/s, from -6.5 to -4.0 km/s, and from -4.0 to 0 km/s and a bloom of little structures with variable velocities.

Fig 17.30 shows at the top the structure at velocity from -4 km/s to 0 km/s in CO emission and our stellar targets. It is indicated for the stars in the direction of the CO structure the stellar identifier, the distance and the velocity corresponding to the structure, where it is measured. Our spectra allows us to chart two structures; first one in the longitude from 160° and 170°, positioned the beginning between 390 pc and 500 pc, and it is properly the California cloud. Second one, it is closer and from longitude 162° to 150°.

In the bottom) Fig 17.30, it is shown the structure with velocities from -6 km/s to -4 km/s. It is evident the same cloud from longitude 162° to 150°. Thanks to the stellar target positions, I can chart its beginning definitely between 210 and 456 pc.

At centre, it is possible to constrain also the structure at velocity from -10 km/s to -6 km/s, as the previous one. It is interesting that the star V * XY

Per is far, but it doesn't show any absorption in this range, probably because it is on the border of the structure.

Using the 17.12, 17.11, 17.10, 17.9, 17.8, 17.7, 17.6 it is possible to name as California at structure from 400 to 600 pc, which starts at longitude 175° and ends at 155°.

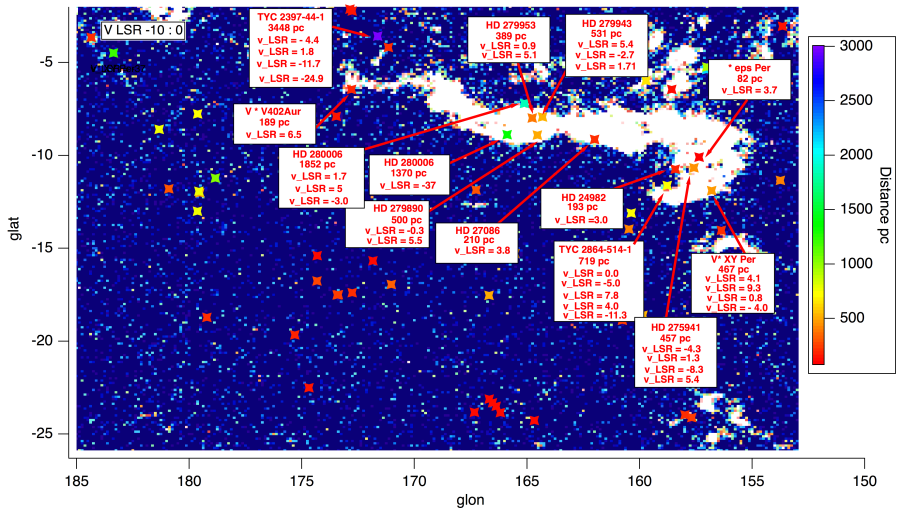
Between 149°and 152°

In Fig 17.31 the vertical cuts in JLcube1 Gaia for 149°, 150°and 152°. I can identify the velocity of cloud part in figure as a pale pink like a cloud with velocity around ∼-7 km/s behind 537 pc. Using *c Per, it is evident that the component at ∼5 km/s is near the plane: this component is also absent for V* V 579 Per, whose latitude is 11°. Since V* V 579 Per has two components comparable in absorption, I suppose they are the same clouds with poli - component, as for example the beginning of North Taurus (from -1 to 5 km/s).

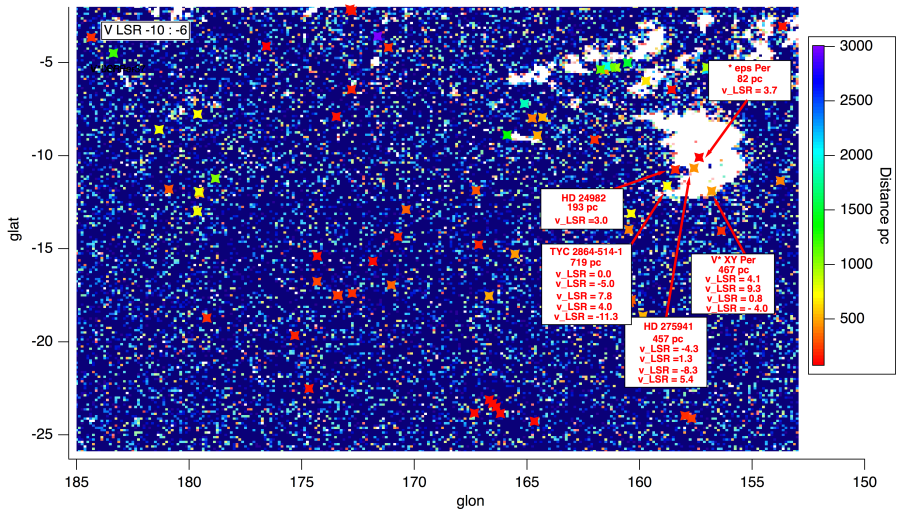
- ∼ 5 km/s, near the plane, closer then 150 pc,pale light blue
- ∼ 2 km/s, from

17.5 Conclusion

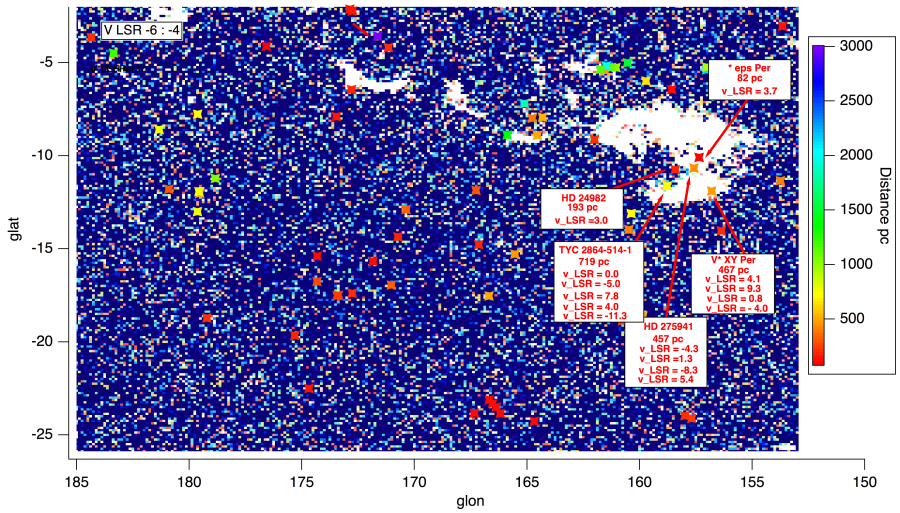
Using CO emission, potassium absorption and the 3D map, it is possible to draw a coherent kinematic picture for the Taurus complex. Molecular and atomic gas and dust are not in the same ISM positions, but using various ISM components, the global puzzle of Taurus nebulae is assembled. Each component adds some information about the nebula. The velocity analyses derived from potassium absorption are already ready also for the Cepheus Nebula region, and a similar reconstruction work will surely be carried out.



(a) California

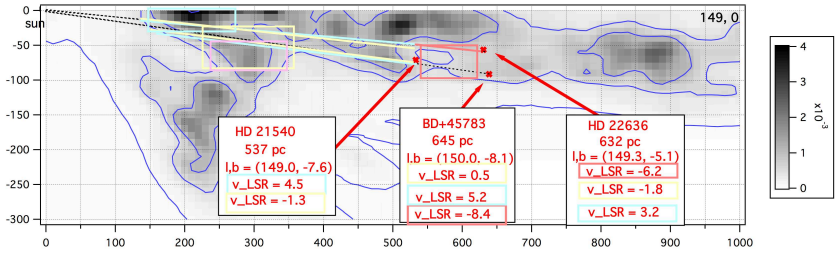


(b) From -10 km/s to -6 km/s

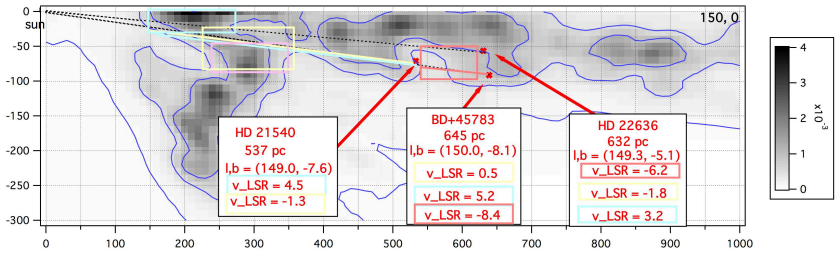


(c) From -6 km/s to -4 km/s

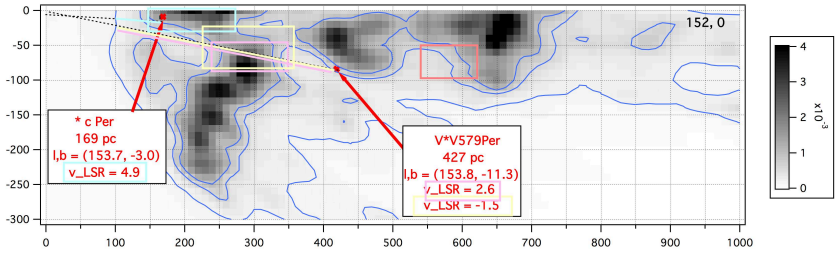
Figure 17.30 – Structure in CO maps at different velocities and stellar targets for California region. Colour scale of the stellar targets represents the distance.



(a)



(b)



(c)

Figure 17.31 – Slice in the 3D map, with stars with galactic longitude between 148.75° and 151.25° for a) and b) and between 151.25° and 153.75°.

Part VII

Conclusion

Chapter 18

Conclusion and perspectives

18.1 Conclusion

This thesis was developed as a test work, to be prepared to Gaia data analyses. My main goal was to develop and test essential tools for the analysis and the exploitation of the Gaia DIB in the RVS spectra.

For the DIB analysis, I coded a profile fitting program in Python and I applied it to a series of synthetic spectra to estimate how many good DIB profiles we may expect to have from the RVS spectra. I used two simulations: the Gaia Universe Model Snapshot (Robin et al., 2012) and the Gaia Object Generator (Luri et al., 2014). To obtain the results I tested several profile fitting functions, and I retained the function which returns the best results in reasonably short time. Finally, I found that a reliable DIB measurement- with uncertainty on the fitted DIB strength (or equivalently extinction) lower than 50%- should be obtained for $\sim$3 million objects.

To improve the full 3D inversions, I tested new techniques to be ready to big data and the Gaia catalog. I tested the inversion with combined data, in particular with reddening derived from photometry of individual stars, spectrometry (APOGEE DIB) again for individual stars, and from statistical processes (Pan-STARRS) applied to groups of stars. We are able now to estimate uncertainties on cloud distances and on integrated extinctions in our maps, despite the fact that full error calculations require too long computation times to be applied to our maps. I developed also a new technique, following the intuition of J. L. Vergely, called "hierarchical technique", which has the two advantages of a shorter calculation time and a more realistic representation of structures, with a new adaptation to the spatial density of stellar targets. I made several tests, partially discussed in this thesis, to understand how to control the inversion code in this context and I applied this technique to a very large sample of extinctions estimated from Gaia photometry. Hierarchical maps suppress artefacts, i.e. they allow structures to emerge from the prior's values only when they are strictly constrained by data, and for this reason this kind of maps can be used as validation tools for other techniques. The map is build using the lines of sight, but for it-own inversion we have not radial effect. Some improvements for this technique are suggested: testing different geometrical configurations, i.e. translating the angles for the determination of the cells, it will be possible to recognise eventual artefacts or uncertain fea-

tures. I was very conservative in the cells' selection, but the results can reach fine resolution for larger volume using lighter constraints.

I also show, benefit from the my colleague M. ElYajouri competence on the profile fitting of the APOGEE DIB, how it is possible, with large DIB absorption datasets, to map the DIB carriers distribution, and I illustrated the results; this kind of studies is still embryonic, but it gives us important indication about the future maps that will be based on Gaia DIBs. We should inverse the DIB taking in account in the prior a different height scale respect to the extinction inversion, perhaps doing several test to understand for each DIB the most appropriate parameter; we must remember also the importance in this frame of the not-detection measurements or upper-limit detections, because they impose the Local Bubble bounds.

In the small Taurus region, a kinematical map was produced, based on profile fitting of interstellar potassium absorptions, on CO radio emission and on 3D maps. These totally new types of maps will be useful because they will help constraining the dynamical evolution of clouds and their interactions. These maps will be designed for various regions, we have yet the data for the Cepheus region.

18.2 Perspectives

Using future Gaia DIB measurements

From my estimations, RVS will provide millions of interesting 862.0 nm DIB measurements, well distributed over the sky. After the publication of the spectra, it is trivial to say that the priority will be the extraction of the maximum number of DIBs, for all stars with a RVS magnitude brighter than 15 mag. Additionally, it will be also possible to extract DIBs from lower SNR data using stacked spectra, (Elyajouri et al., 2017a).

Respect to previous catalog, Gaia data will have various and strong advantages: numerous, clean of telluric lines, full sky coverage. They should provide very valable constraints.

In general, massive DIB data may provide several types of information. First, some information on the DIB carrier properties (location of formation, destruction) thanks to their distribution on the space and the proportion with dust and other gas component. Since the DIBs carriers are essentially matter, we can use the DIBs information to make a ISM tomography.

At the end, the DIBs can be add information on ISM kinematic because it will be possible to extract the DIB radial velocities for the strong DIBs and the most distant objects, even if the large DIB width hinders a very simple extraction (Puspitarini et al., 2015). As a matter of fact, since the DIB is broad, only clouds with velocities differing by at least 5-10 km.s^{-1} can be extracted and it will mainly concern clouds belonging to different arms.

Tomographic techniques and future 3D maps (dust or DIB carriers)

The hierarchical technique that I developed guarantees the possibility to inverse a large database; new DIB maps will be charted and a more precise scale height will be provided, allowing accurate comparisons between excess of

colour- and DIB absorption- based maps. These maps will bring constraints on the environments that are favourable for the formation and survival of the GAIA DIB carriers. At the moment there is no specific study of the sites of formation of the Gaia DIB, but according to Kos et al. (2014) the DIB is linked with the gas, and it means that we will be able to make a 3D map of the gas distribution, to compare with the 3D dust maps provided by extinctions. Our hierarchical methods are unique and adapted to future datasets. One can foresee applications to most ground based surveys for both extinctions or DIB measurements. Others large surveys can provide new large DIB catalogs in optical or infrared bands, that we will be able to use and compare with the Gaia DIB. WEAVE is a new multi-object spectrograph for the 4.2-m William Herschel Telescope (WHT) at the Observatorio del Roque de los Muchachos. The first light is scheduled for May 2019. The instrument is designed to make large stellar surveys in the spectral interval from 370 to 960 nm, and in this range there are a lot of observable optical DIBs. MOONS (Multi Object Optical and Near-infrared Spectrograph) is another interesting instrument in the infrared (range 0.65 - 1.8 micron), mounted on the VLT; it will start scientific observations at the end of 2020. Measurements with this instrument will allow to explore the infrared DIBs (Cirasuolo et al., 2014). In addition to these data, other spectra at high resolution will be potentially available using the new SPIRou (SPectropolarimètre InfraROUge) facility, (Moutou et al., 2015). The spectrograph is mounted on the 3.6m Canada-France-Hawaii telescope located on top of Mauna Kea, and it provides YJHK spectra from 0.95 to 2.35 microns in a single shot, at a spectral resolution of 75,000. With this very precise instrument it is possible to measure the Gaia DIB at high resolution, allowing to extract a new and more detailed shape of the DIB.

Our methods have their own advantages (they take benefit from omni directional links, from full hierarchy of sizes), and also their own shortcomings (difficulties for error estimates, limited angular resolution): ideally they should be combined with other Bayesian techniques to take benefit of all positive aspects.

In the frame of Gaia, derivations of stellar populations properties for stars too faint to be observed by the RVS will benefit from dust maps and reciprocally additional extinction estimates for individual objects will serve to improve the maps, this can be done iteratively. Still in the Gaia frame, comparisons between the locations of stellar populations of different ages and the dust clouds will give clues to the history of star formation.

A completely different contest, detailed dust maps and associated properties linked to dust (reddening law, beta parameter) are crucial for foreground removal in CMB measurements. The galactic dust emits in microwave, and it pollutes the cosmological measurements, but more and more precise maps will permit to remove this spurious component. The *Planck* measures opened in the last years one important question at large scale about the extinction laws, in particular the the Planck Collaboration et al. (2014) β coefficient map shows a variability in β which is not linkable with any known dust, gas or stellar structure. To understand this variability, Schlafly et al. (2017) tried a comparison with their extinction profiles based on Pan-STARRS photometry, looking at the variability over the sky of the Rv value and they found some links with Planck β: the explanation may be that since Rv varies in response to the grain size distribution, this distribution is not spatially uniform, and it is reflected in

β variations. Such informations on the grain sizes and their favoured locations on the large scale will help constraining the Galactic evolution.

Future astrophysical studies of the galactic mechanism will pass through detailed interstellar medium maps for the different components - gas species, dust species and macro molecules species. In this sense techniques as the 3D inversions, that is replicable for various sources, will permit us direct comparisons.

Kinetic tomography

Our study of the Taurus clouds is a precursor of what could be done on the large scale by means of the combination of emission data (HI, CO), absorption measurements (gaseous lines, DIBs) and dust maps, i.e. the construction of a dynamical picture of the ISM on the large scale. Such a tool would be particularly useful in the frame of Gaia, allowing to build and test dynamical models of the Milky Way including stars and gas, and predict star formation. In particular comparing the dust distribution, the populations distribution, the clouds and the stellar velocities, a complete analysis of the motions and history of our galaxy. These maps can also be used at much smaller scales - as the Taurus one - in star-forming regions to better understand cloud collapse and feedback from newly born stars. We have yet compared our maps with X-ray observations, but we can better understand the bubble generated by supernovae using a kinetic tomography.

And for me, that's all.

Appendix A

Inversion parameter example

Exemple of configuration file for the inversion.

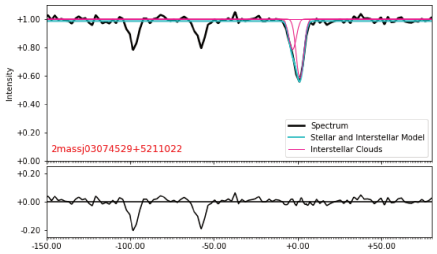
<?xml version="1.0"?>
<tomo>
<CONF>
<l>25000</l> ! maximum lines of sight number
<isec>0</isec> ! if 1, code uses a sec norm, otherwise a gaussian norm
<ml>10</ml>
<switchbloc>1</switchbloc> ! if 1, it divided in blocks the dataset
<groupb>500</groupb> ! physical dimension of blocks, here 500x500 pc
<switchcarteprior>0</switchcarteprior> ! if 0, analytical prior; if 1 external prior
<switchautocorr>1</switchautocorr> ! choice of autocorrelation function
<ro0>0.00025</ro0> ! density on the prior in the Galactic plane
<h0>200</h0> ! scale height
<sr0>0.8</sr0> ! σ_0 for the autocorrelation function
<sr1>1.0</sr1> ! ξ_0 for the autocorrelation function
<xi0>30.0</xi0> ! σ_1 for the autocorrelation function
<xi1>15.0</xi1>! ξ_1 for the autocorrelation function
<errmin>0.01</errmin> ! minimum error on the excess of colour
<lmax>2500.0</lmax> ! selected target closer to the Sun then lmax [pc]
<hmax>500.0</hmax> ! selected target closer to the Galactic plane then hmax [pc]
<resolution>10</resolution> ! resulution for the integration on the line of sight
<longmin>0.</longmin> ! selected target with longitude highr then longmin
<longmax>2.</longmax> ! selected target with longitude lower then longmax
<hsig>10000.0</hsig>
<spri>1.</spri> ! typical error on the prior
<ndlimi>20.0</ndlimi> ! reciprocal influence distance ndlimi*max(xi0,xi1)) for the density extimation
<ndd0>12.0</ndd0> ! maximal distance between two target to be taken in account during the integration ndd0*max(xi0,xi1)
<dmax>2000.0</dmax> ! half-base cube dimension [pc]
<dmaxz>300.0</dmaxz> ! half - altitude cube dimension [pc]
<pas>5.0</pas> ! sampling in cube
<nsigdata>3.0</nsigdata> ! how many sigma difference between datapoint and model to be considered outlier
<repres>resultscramer</repres> ! directory for the results

<reptemp>tmp</reptemp>
<ficdata>listatarget.txt</ficdata> ! input data
<ficprior>cartepri0</ficprior> ! in case, external prior adress
<ficpriorout>cartepri0</ficpriorout> ! output name for the results, ready to be used as prior
<ficcube>cube</ficcube>
<ficautocorr>fileautocorr</ficautocorr>
</CONF>
</tomo>

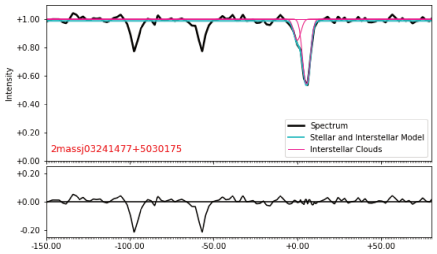
Appendix B

Fit NARVAL and ESPaDOnS spectra for Taurus

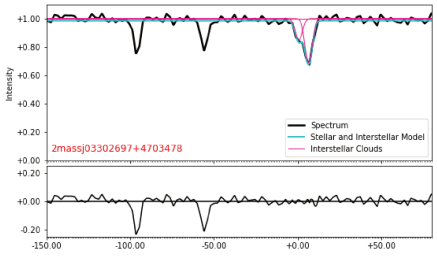
In this appendix I present all image for fit not yet in Chapter 17 and the fit results tables.



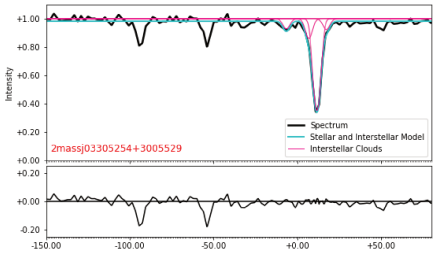
(a) HD19231



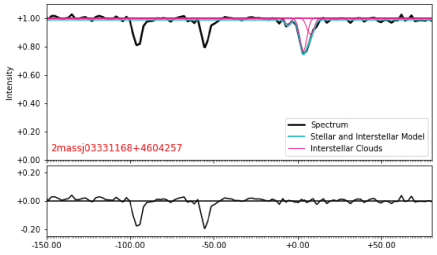
(b) BD+49915



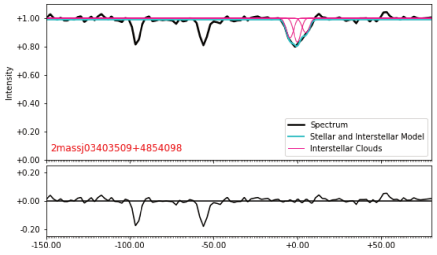
(c) HD21540



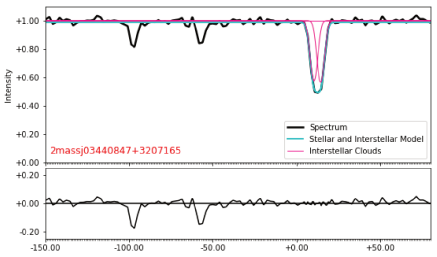
(d) TYC2342-639-1



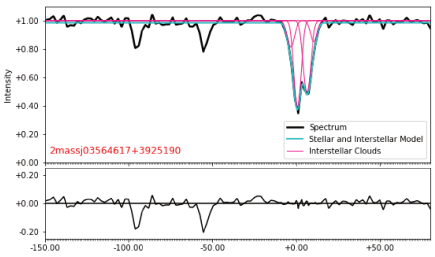
(e) BD+45783



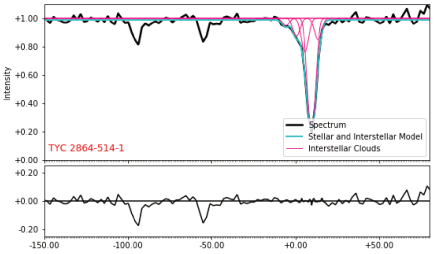
(f) HD22636



(g) HD281160

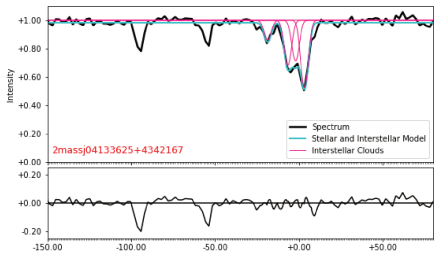


(h) HD275941

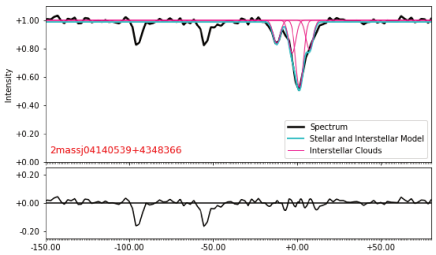


(i) TYC2864-514-1

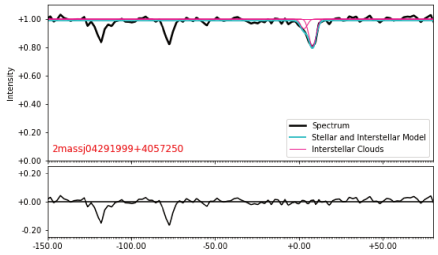
Figure B.1 – .



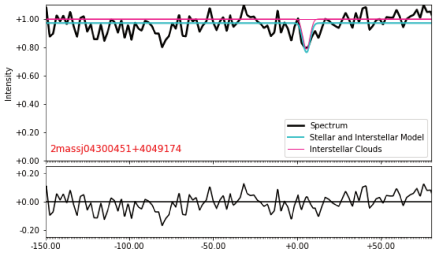
(a) LSV+434



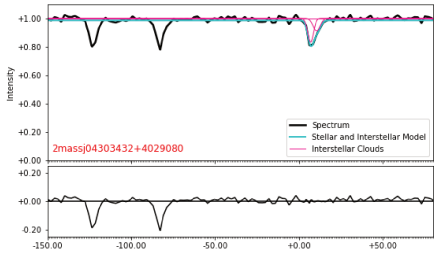
(b) HD26568



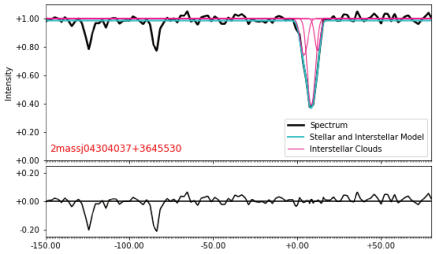
(c) HD276436



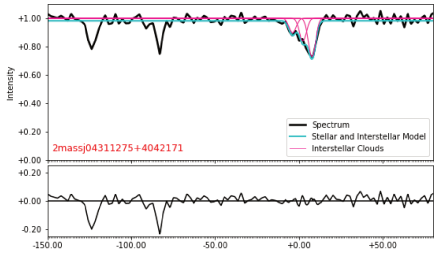
(d) HD276446



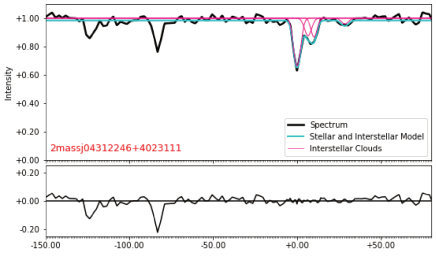
(e) HD276453



(f) HD279943

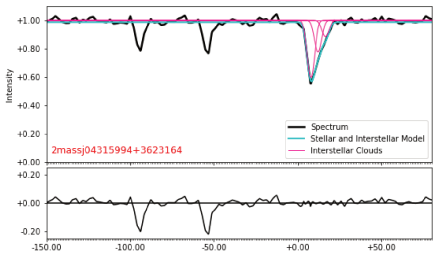


(g) HD276449

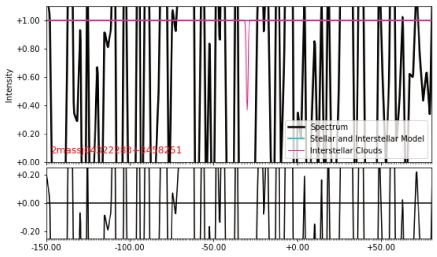


(h) HD276454

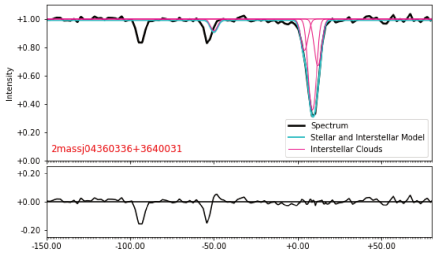
Figure B.2 – .



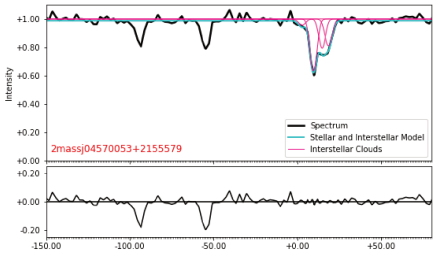
(a) HD279953



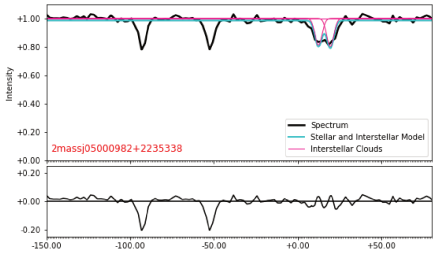
(b) HD279974



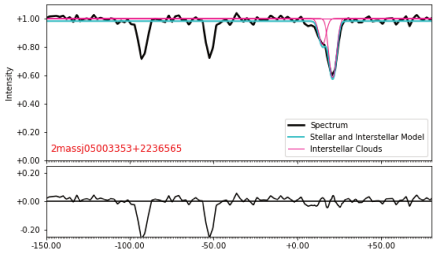
(c) HD280006



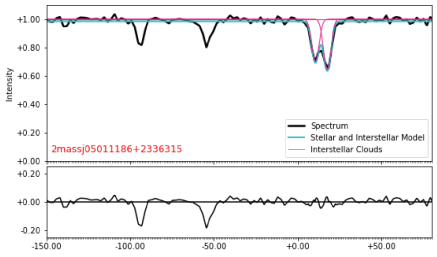
(d) HD285016



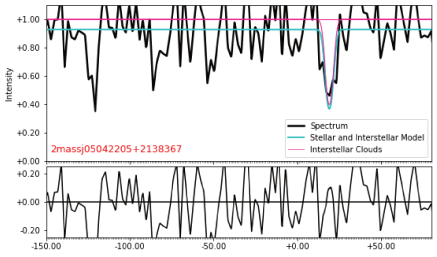
(e) HD31856



(f) HD31916

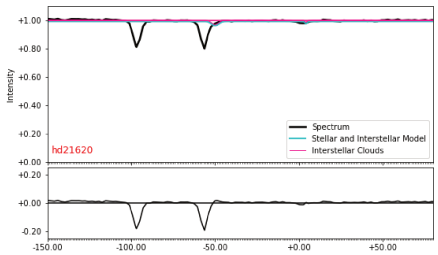


(g) HD284938

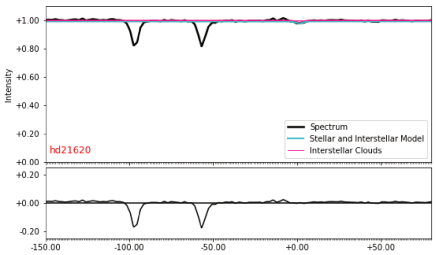


(h) HD32481

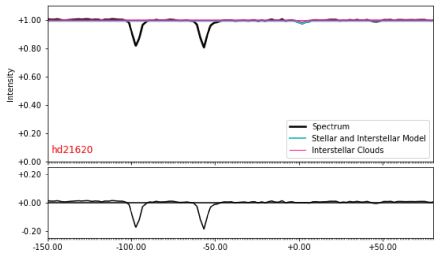
Figure B.3 – .



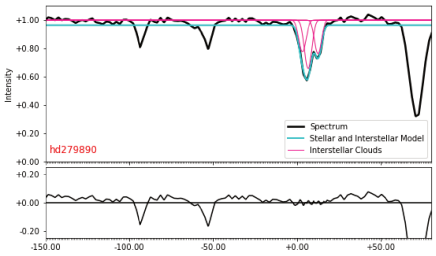
(a) HD21620



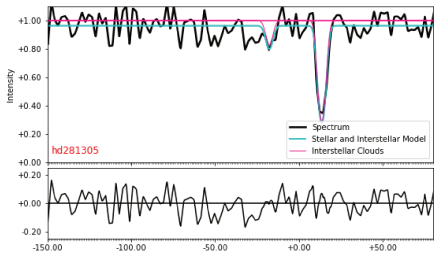
(b) HD21620



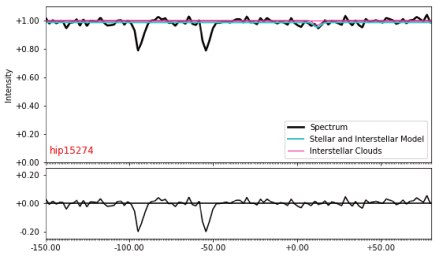
(c) HD21620



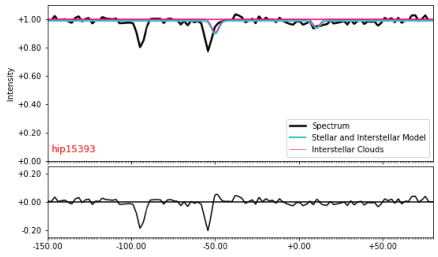
(d) HD279890



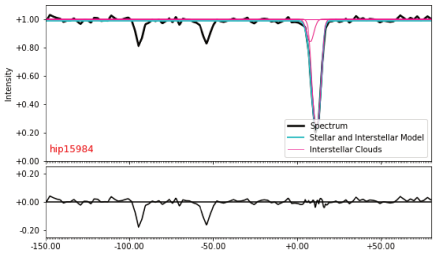
(e) HD281305



(f) HD20326

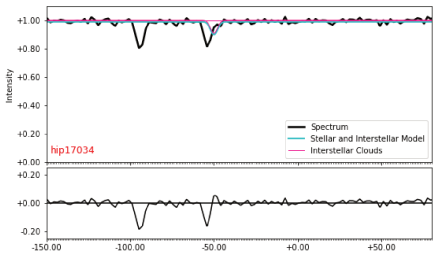


(g) HD20456

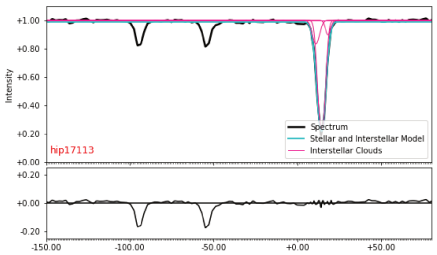


(h) BD+30540

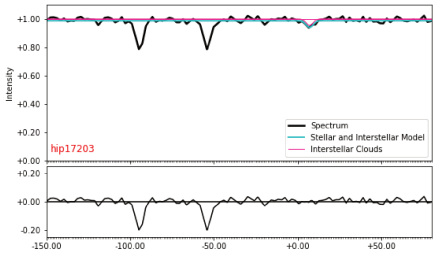
Figure B.4 – .



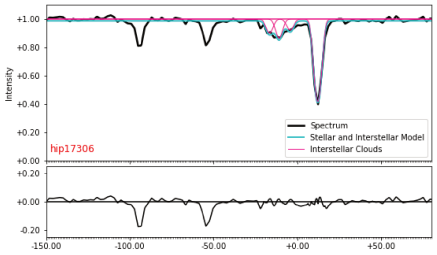
(a) HD22614



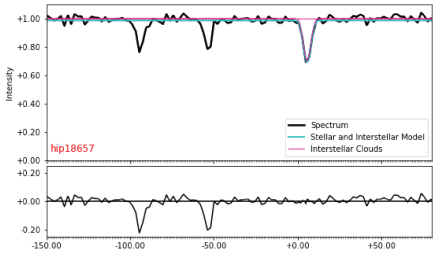
(b) HD278942



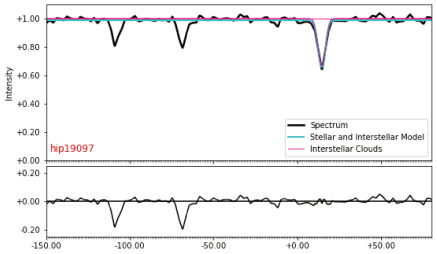
(c) HD22780



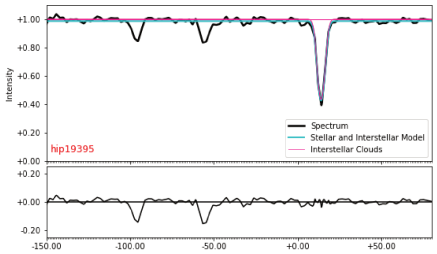
(d) BD+30564



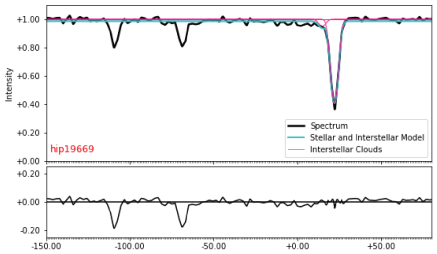
(e) HD24982



(f) HD25694

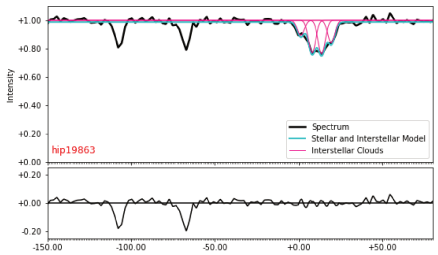


(g) NGC1514

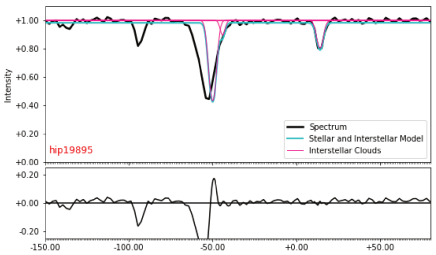


(h) HD26572

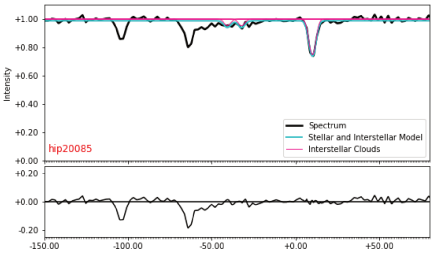
Figure B.5 – .



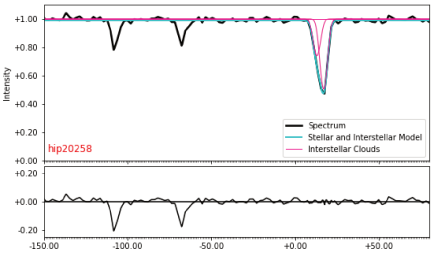
(a) HD26746



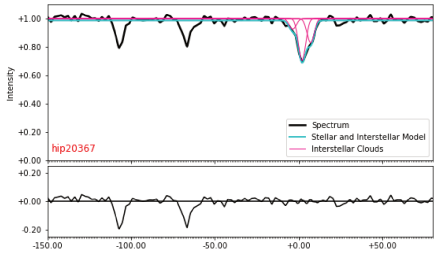
(b) HD26873



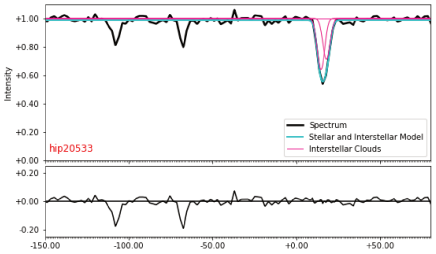
(c) HD27086



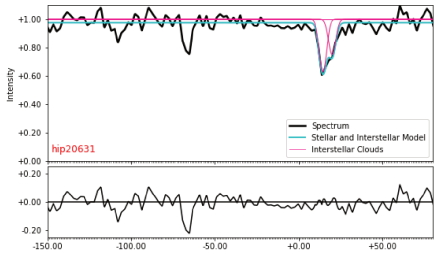
(d) HD27405



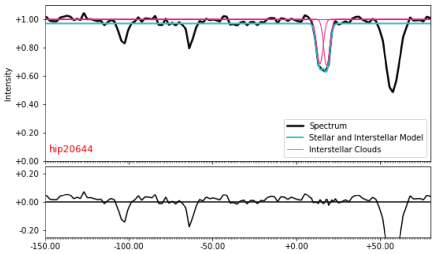
(e) HD27448



(f) *62Tau

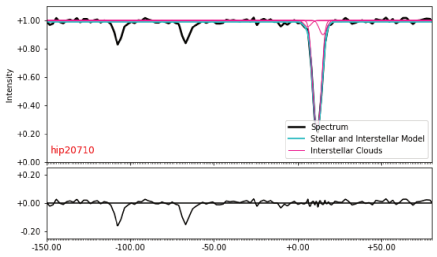


(g) HD27923

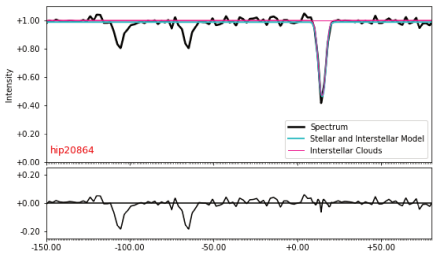


(h) HD284421

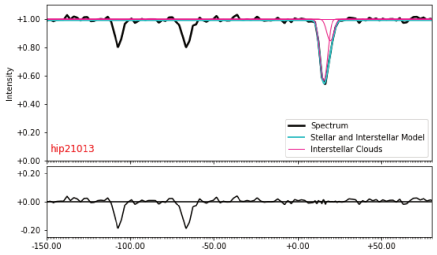
Figure B.6 – .



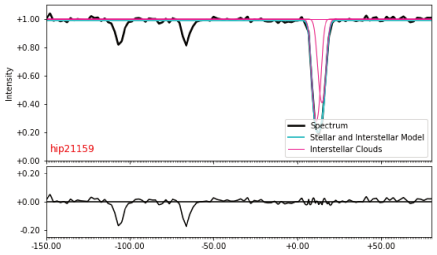
(a) HD27982



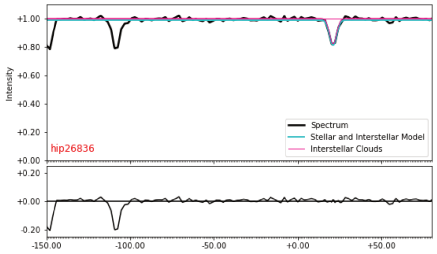
(b) HD28225



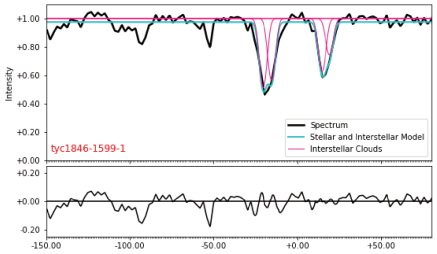
(c) HD28482



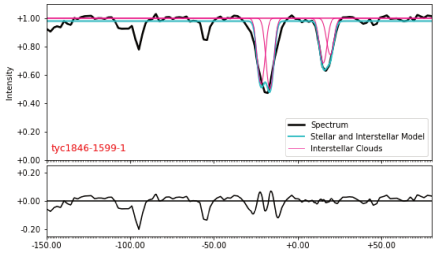
(d) HD283677



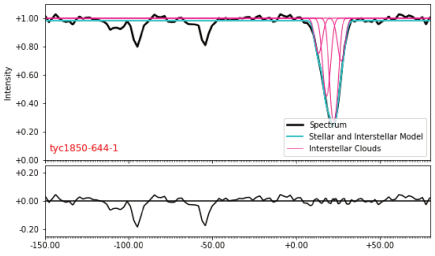
(e) HD37752



(f) 2MASSJ05164313+2307495

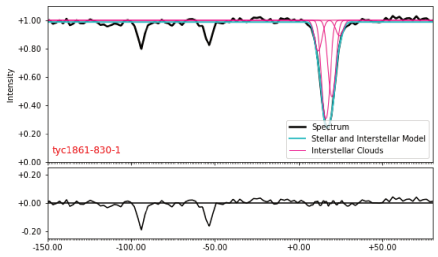


(g) 2MASSJ05164313+2307495

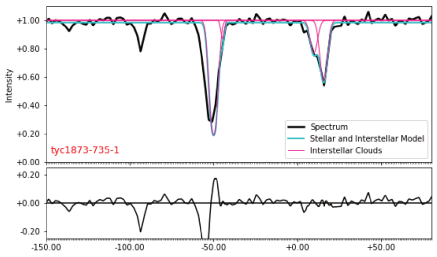


(h) TYC1850-644-1

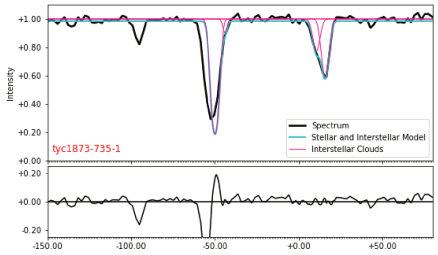
Figure B.7 – .



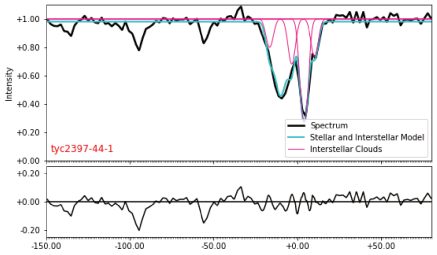
(a) TYC1861-830-1



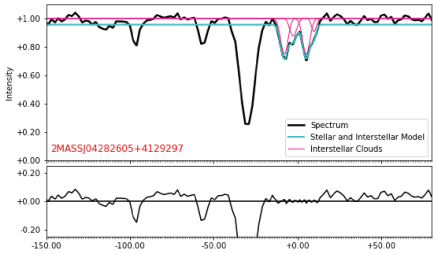
(b) 2MASSJ05411140+2816123



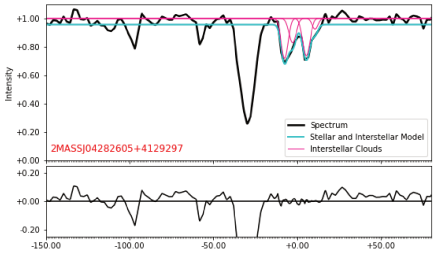
(c) 2MASSJ05411140+2816123



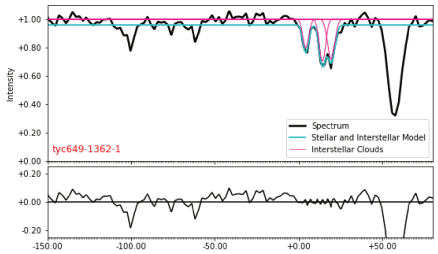
(d) TYC2397-44-1



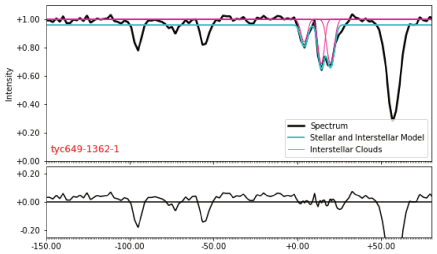
(e) 2MASSJ04282605+4129297



(f) 2MASSJ04282605+4129297

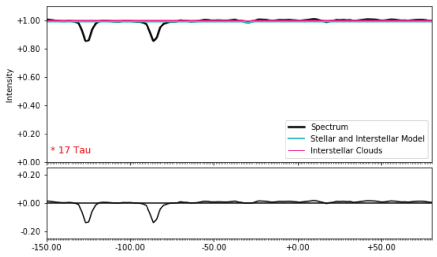


(g) 2MASSJ03232843+0944216

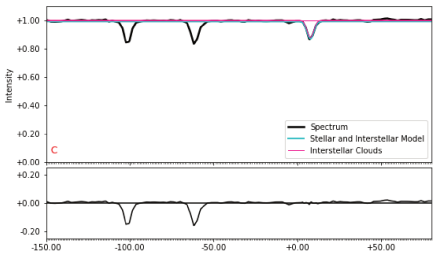


(h) 2MASSJ03232843+0944216

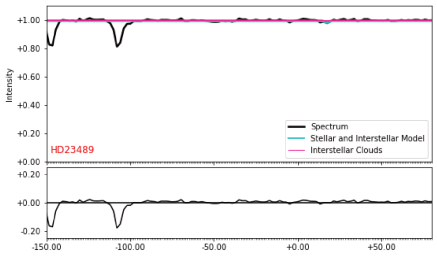
Figure B.8 – .



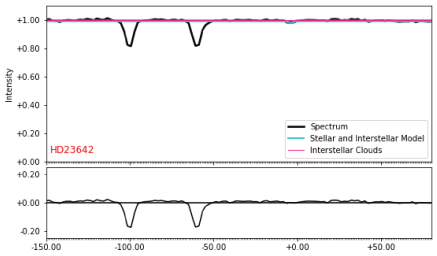
(a) *17Tau



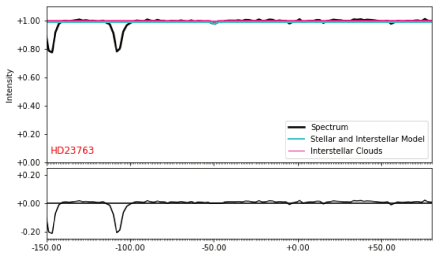
(b) *cPer



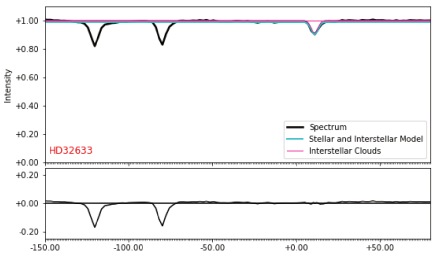
(c) HD23489



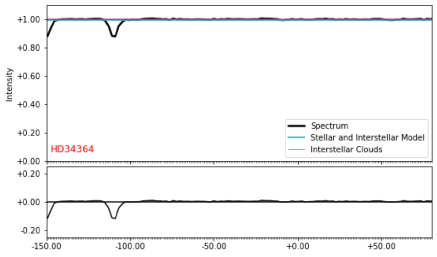
(d) V*V1229Tau



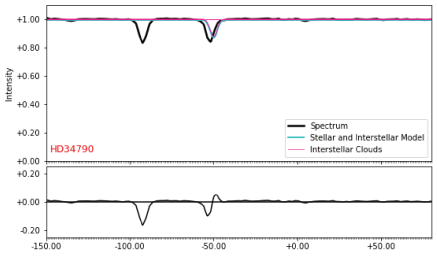
(e) HD23763



(f) V*HZAur

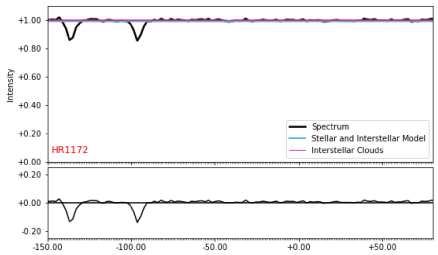


(g) *17Aur

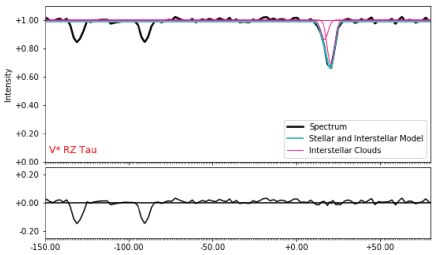


(h) HD34790

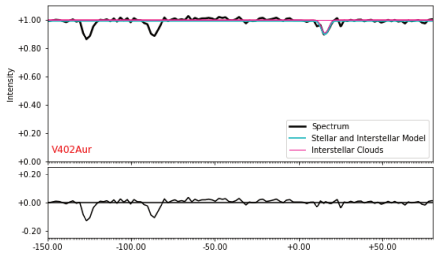
Figure B.9 – .



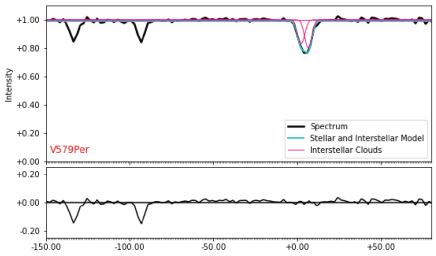
(a) HD23753



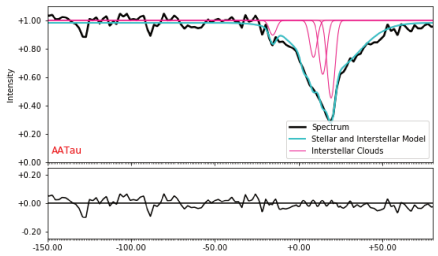
(b) V*RZTau



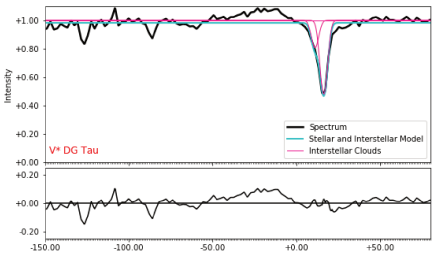
(c) V*V402Aur



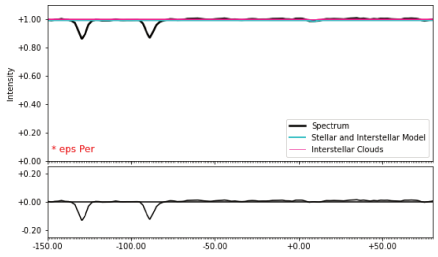
(d) V*V579Per



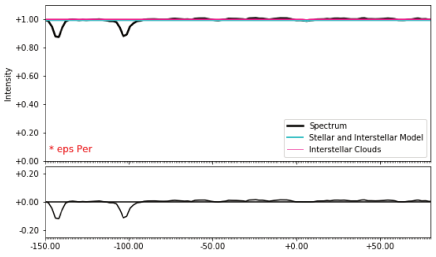
(e) V*AATau



(f) V*DGTau

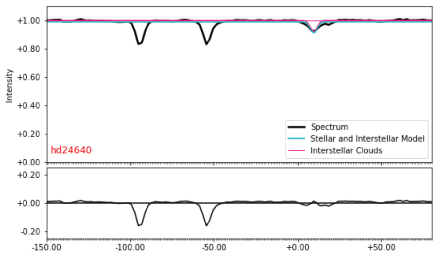


(g) *epsPer

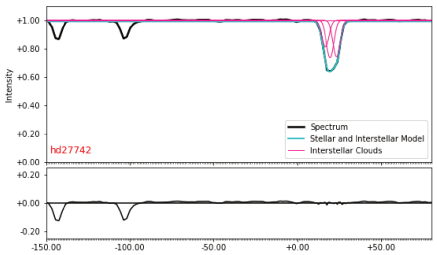


(h) *epsPer

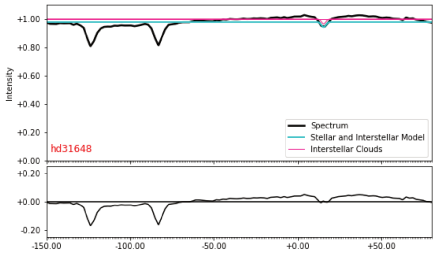
Figure B.10 – .



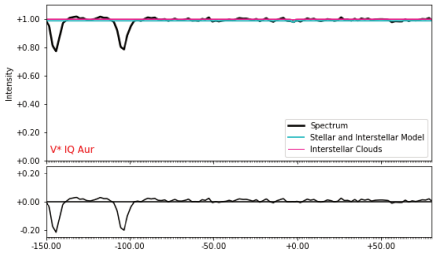
(a) HD24640



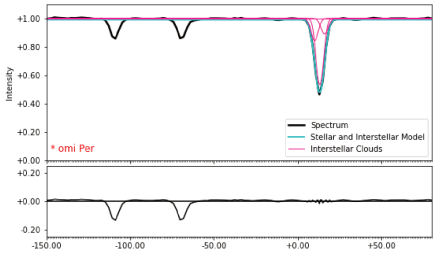
(b) HD27742



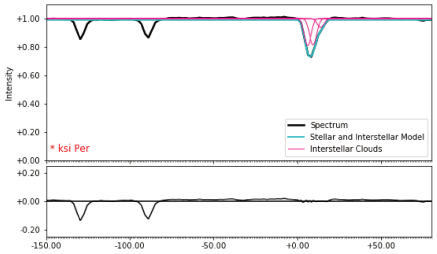
(c) HD31648



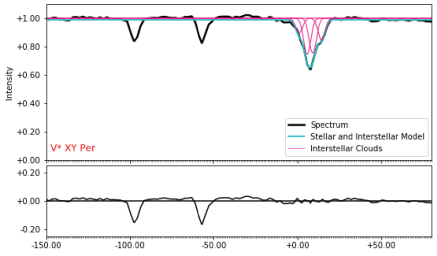
(d) V*IQAur



(e) *omiPer



(f) *ksiPer



(g) V*XYPer

Figure B.11 – .

Table B.1 – Fit NARVAL and ESPaDOnS spectra for Taurus, * symbols are for velocities from stellar lines, T symbols are for velocities from telluric lines.

identifier	glon	glat	pi	v0 HBR	v1 HBR	v2 HBR	v3 HBR	v0 LSR	v1 LSR	v2 LSR	v3 LSR	τ0	τ1	τ2	τ3
HD 19231	143.115	-5.264	3.04	1.294	-3.094	9999.0	9999.0	1.56	-2.828	9999.266	9999.266	0.541	0.243	9999.0	9999.0
BD+49 915	146.201	-5.333	4.06	5.366	-0.031	9999.0	9999.0	4.634	-0.764	9998.268	9998.268	0.613	0.161	9999.0	9999.0
HD 21540	149.002	-7.6	1.86	6.425	0.626	9999.0	9999.0	4.495	-1.303	9997.07	9997.07	0.372	0.164	9999.0	9999.0
TYC 2342-639-1	159.44	-21.247	1.75	11.603	6.816	17.254	-6.605	4.851	0.064	10.501	-13.358	1.05	0.153	0.081	0.075
BD+45 783	149.968	-8.134	1.55	3.252	7.516	-6.045	9999.0	0.945	5.209	-8.352	9996.693	0.267	0.121	0.043	9999.0
HD 22636	149.295	-5.116	1.58	-4.443	-0.058	4.964	9999.0	-6.138	-1.753	3.27	9997.306	0.15	0.185	0.109	9999.0
HD 281160	160.446	-17.892	3.13	14.341	10.301	9999.0	9999.0	7.633	3.593	9992.292	9992.292	0.571	0.556	9999.0	9999.0
HD 275941	157.586	-10.669	2.19	0.703	6.362	-3.25	10.381	-4.32	1.339	-8.273	5.358	0.932	0.698	0.206	0.161
TYC 2864-514-1	158.771	-11.62	1.39	5.538	0.493	13.361	9.465	0.035	-5.009	7.858	3.962	0.218	0.102	0.264	1.429
LS V +43 4	157.031	-5.384	0.47	3.236	-6.682	-18.698	-2.091	-0.953	-10.871	-22.887	-6.279	0.643	0.386	0.15	0.34
HD 26568	157.021	-5.247	1.07	1.44	7.344	-12.157	-2.899	-2.728	3.176	-16.325	-7.067	0.639	0.234	0.177	0.264
HD 276436	161.045	-5.259	1.02	7.986	2.938	9999.0	9999.0	2.559	-2.489	9993.573	9993.573	0.215	0.072	9999.0	9999.0
HD 276446	161.241	-5.25	0.93	5.586	9999.0	9999.0	9999.0	0.1	9993.514	9993.514	9993.514	0.24	9999.0	9999.0	9999.0
HD 276453	161.552	-5.411	2.16	6.89	10.47	9999.0	9999.0	1.288	4.867	9993.397	9993.397	0.182	0.093	9999.0	9999.0
HD 279943	164.309	-7.93	1.88	12.165	4.103	8.47	9999.0	5.408	-2.655	1.713	9992.243	0.257	0.298	0.944	9999.0
HD 276449	161.474	-5.172	0.5	7.7	2.311	-3.854	9999.0	2.151	-3.238	-9.403	9993.451	0.317	0.176	0.109	9999.0
HD 276454	161.729	-5.367	0.88	0.026	10.111	6.54	28.523	-5.626	4.459	0.888	22.871	0.412	0.143	0.133	0.042
HD 279953	164.771	-7.99	2.57	7.8	12.065	16.473	9999.0	0.896	5.161	9.569	9992.096	0.518	0.253	0.122	9999.0
HD 279974	165.88	-8.886	0.73	-30.0	9999.0	9999.0	9999.0	-37.343	9991.657	9991.657	9991.657	1.0	9999.0	9999.0	9999.0
HD 280006	165.113	-7.198	0.54	8.669	3.959	11.989	-50.0(T)	1.757	-2.953	5.077	-56.912(T)	1.03	0.251	0.403	0.088
HD 285016	179.632	-13.001	1.28	9.591	18.522	14.714	2.794	-2.061	6.87	3.063	-8.858	0.458	0.211	0.232	0.047
HD 31856	179.534	-12.022	1.96	19.359	12.345	9999.0	9999.0	7.825	0.811	9987.466	9987.466	0.221	0.207	9999.0	9999.0
HD 31916	179.57	-11.935	1.28	21.025	14.961	9999.0	9999.0	9.49	3.426	9987.464	9987.464	0.54	0.2	9999.0	9999.0
HD 284938	178.843	-11.227	0.96	17.542	10.53	9999.0	9999.0	6.267	-0.745	9987.725	9987.725	0.439	0.354	9999.0	9999.0
HD 32481	180.904	-11.794	3.39	19.052	9999.0	9999.0	9999.0	7.182	9987.13	9987.13	9987.13	0.927	9999.0	9999.0	9999.0

Table B.2 – Fit NARVAL and ESPaDOnS spectra for Taurus, * symbols are for velocities from stellar lines, T symbols are for velocities from telluric lines.

identifier	glon	glat	pi	v0 HBR	v1 HBR	v2 HBR	v3 HBR	v0 LSR	v1 LSR	v2 LSR	v3 LSR	τ0	τ1	τ2	τ3
HD 21620	147.903	-5.744	7.41	-50.0	-50.0	4.22	9999.0	-51.332	-51.332	2.887	9997.668	0.031	0.0	0.014	9999.0
HD 21620	147.903	-5.744	7.41	1.278	9999.0	9999.0	9999.0	-0.054	9997.668	9997.668	9997.668	0.013	9999.0	9999.0	9999.0
HD 21620	147.903	-5.744	7.41	1.812	9999.0	9999.0	9999.0	0.479	9997.668	9997.668	9997.668	0.016	9999.0	9999.0	9999.0
HD 279890	164.537	-8.903	2.0	6.558	12.452	3.176	9999.0	-0.384	5.51	-3.766	9992.058	0.423	0.273	0.256	9999.0
HD 281305	162.941	-17.32	3.04	13.578	-17.683(*)	9999.0	9999.0	6.191	-25.07(*)	9991.613	9991.613	1.239	0.19	9999.0	9999.0
HD 20326	157.685	-24.114	4.02	12.015	9999.0	9999.0	9999.0	5.491	9992.476	9992.476	9992.476	0.034	9999.0	9999.0	9999.0
HD 20456	157.984	-23.975	4.94	10.43	-50.0(*)	9999.0	9999.0	3.833	-56.597(*)	9992.403	9992.403	0.056	0.096	9999.0	9999.0
BD+30 540	157.971	-21.238	3.19	11.577	7.858	9999.0	9999.0	5.256	1.537	9992.678	9992.678	1.448	0.174	9999.0	9999.0
HD 22614	164.675	-24.272	7.66	-50.0(T)	9999.0	9999.0	9999.0	-58.52(T)	9990.48	9990.48	9990.48	0.096	9999.0	9999.0	9999.0
HD 278942	159.85	-18.61	1.92	14.4	11.029	18.286	9999.0	7.794	4.423	11.68	9992.394	1.462	0.184	0.109	9999.0
HD 22780	156.373	-14.038	4.64	7.053	9999.0	9999.0	9999.0	2.004	9993.951	9993.951	9993.951	0.048	9999.0	9999.0	9999.0
BD+30 564	160.769	-18.877	4.26	12.11	-10.937	-16.651(*)	-4.358(*)	5.205(*)	-17.842(*)	-23.555(*)	-11.263(*)	0.88	0.146	0.104	0.082
HD 24982	158.42	-10.751	5.16	5.586	9999.0	9999.0	9999.0	0.295	9993.709	9993.709	9993.709	0.354	9999.0	9999.0	9999.0
HD 25694	166.685	-17.54	2.0	14.527	9999.0	9999.0	9999.0	6.03	9990.503	9990.503	9990.503	0.408	9999.0	9999.0	9999.0
NGC 1514	165.534	-15.288	2.14	13.924	9999.0	9999.0	9999.0	5.982	9991.058	9991.058	9991.058	0.838	9999.0	9999.0	9999.0
HD 26572	174.687	-22.504	6.91	22.08	14.769	9999.0	9999.0	10.982	3.671	9987.901	9987.901	0.888	0.052	9999.0	9999.0
HD 26746	158.575	-6.452	8.81	13.634(*)	7.936(*)	19.392(*)	1.219(*)	8.825(*)	3.126(*)	14.582(*)	-3.591(*)	0.268	0.26	0.175	0.114
HD 26873	167.14	-14.778	5.66	-50.0(*)	14.037	-44.025(*)	9999.0	-58.358(*)	5.678	-52.384(*)	9990.642	0.841	0.218	0.111	9999.0
HD 27086	161.998	-9.136	4.74	10.008	-32.252(*)	-40.731(*)	9999.0	3.809	-38.451(*)	-46.929(*)	9992.802	0.293	0.05	0.049	9999.0
HD 27405	171.024	-16.953	3.11	16.948	12.889	9999.0	9999.0	7.274	3.216	9989.326	9989.326	0.684	0.302	9999.0	9999.0
HD 27448	159.717	-5.979	1.51	1.929	7.032	-3.612	9999.0	-3.176	1.927	-8.717	9993.895	0.351	0.186	0.077	9999.0
* 62 Tau	172.763	-17.393	4.46	14.622	17.922	9999.0	9999.0	4.428	7.728	9988.806	9988.806	0.447	0.339	9999.0	9999.0
HD 27923	173.375	-17.507	4.32	14.377	19.856	9999.0	9999.0	4.006	9.485	9988.629	9988.629	0.463	0.288	9999.0	9999.0
HD 284421	173.435	-17.518	3.94	13.996	18.084	9999.0	9999.0	3.607	7.696	9988.611	9988.611	0.379	0.396	9999.0	9999.0
HD 27982	167.263	-11.861	3.42	14.987	5.858	11.471	9999.0	6.898	-2.23	3.383	9990.911	0.107	0.046	1.442	9999.0
HD 28225	170.735	-14.353	5.8	14.699	9999.0	9999.0	9999.0	5.35	9989.651	9989.651	9989.651	0.759	9999.0	9999.0	9999.0
HD 28482	174.327	-16.765	3.53	15.507	19.373	9999.0	9999.0	4.94	8.805	9988.432	9988.432	0.568	0.169	9999.0	9999.0
HD 283677	170.369	-12.895	2.97	11.256	14.628	9999.0	9999.0	2.157	5.53	9989.901	9989.901	1.288	0.898	9999.0	9999.0
HD 37752	184.347	-3.645	4.69	21.024	9999.0	9999.0	9999.0	9.153	9987.129	9987.129	9987.129	0.196	9999.0	9999.0	9999.0
2MASS J05164313+2307495	181.337	-8.603	1.28	-20.409	14.778	-15.453	19.06	-32.072	3.115	-27.116	7.397	0.672	0.465	0.573	0.296
2MASS J05164313+2307495	181.337	-8.603	1.28	-16.644	15.182	-21.91	19.318	-28.307	3.52	-33.572	7.655	0.694	0.383	0.624	0.298
TYC 1850-644-1	179.625	-7.772	1.43	18.068	13.284	22.404	26.792	6.945	2.161	11.282	15.669	0.8	0.286	1.307	0.361
TYC 1861-830-1	183.363	-4.494	0.83	16.107	19.716	11.636	24.259	4.384	7.993	-0.087	12.536	1.19	0.78	0.245	0.117

Table B.3 – Fit NARVAL and ESPaDOnS spectra for Taurus, * symbols are for velocities from stellar lines, T symbols are for velocities from telluric lines.

identifier	glon	glat	pi	v0 HBR	v1 HBR	v2 HBR	v3 HBR	v0 LSR	v1 LSR	v2 LSR	v3 LSR	τ0	τ1	τ2	τ3
2MASS J05411140+2816123	180.056	-1.177	0.91	-50.0(*)	15.673	9.839	-44.639	-60.442(*)	5.231	-0.603	-55.081	1.67	0.561	0.262	0.14
2MASS J05411140+2816123	180.056	-1.177	0.91	-50.0(*)	15.722	10.484	-44.08	-60.442(*)	5.28	0.042	-54.522	1.674	0.525	0.262	0.101
TYC 2397-44-1	171.634	-3.585	0.29	4.0	10.248	-3.381	-16.561	-4.4	1.848	-11.781	-24.961	1.306	0.895	0.836	0.341
2MASS J04282605+4129297	160.539	-5.014	0.67	5.153	-7.965	-2.622	9.705	-0.086	-13.204	-7.861	4.467	0.289	0.29	0.133	0.098
2MASS J04282605+4129297	160.539	-5.014	0.67	-7.548	5.395	-2.802	10.601	-12.786	0.156	-8.041	5.363	0.33	0.305	0.187	0.078
2MASS J03232843+0944216	173.188	-37.879	0.89	14.121	4.203	19.32	9999.0	2.755	-7.163	7.954	9987.634	0.354	0.226	0.328	9999.0
2MASS J03232843+0944216	173.188	-37.879	0.89	19.743	14.193	4.003	9999.0	8.377	2.827	-7.363	9987.634	0.375	0.381	0.175	9999.0
* 17 Tau	166.18	-23.849	8.7	-29.761	9999.0	9999.0	9999.0	-38.664	9990.097	9990.097	9990.097	0.007	9999.0	9999.0	9999.0
* c Per	153.654	-3.045	5.93	7.733	9999.0	9999.0	9999.0	4.92	9996.187	9996.187	9996.187	0.132	9999.0	9999.0	9999.0
HD 23489	166.369	-23.502	7.22	17.281	9999.0	9999.0	9999.0	8.353	9990.072	9990.072	9990.072	0.012	9999.0	9999.0	9999.0
V* V1229 Tau	166.535	-23.319	7.57	-4.138	9999.0	9999.0	9999.0	-13.097	9990.04	9990.04	9990.04	0.011	9999.0	9999.0	9999.0
HD 23763	166.678	-23.121	7.84	-49.899	9999.0	9999.0	9999.0	-58.883	9990.017	9990.017	9990.017	0.011	9999.0	9999.0	9999.0
V* HZ Aur	171.157	-4.19	4.69	10.865	9999.0	9999.0	9999.0	2.527	9990.662	9990.662	9990.662	0.091	9999.0	9999.0	9999.0
* 17 Aur	172.767	-2.233	6.72	-9.276	9999.0	9999.0	9999.0	-17.831	9990.445	9990.445	9990.445	0.002	9999.0	9999.0	9999.0
HD 34790	176.572	-4.124	10.05	-50.0	9999.0	9999.0	9999.0	-59.868	9989.132	9989.132	9989.132	0.132	9999.0	9999.0	9999.0
HD 23753	167.332	-23.826	7.72	-41.615	9999.0	9999.0	9999.0	-50.834	9989.781	9989.781	9989.781	0.004	9999.0	9999.0	9999.0
V* RZ Tau	179.214	-18.724	4.52	20.582	16.715	9999.0	9999.0	8.571	4.704	9986.99	9986.99	0.372	0.149	9999.0	9999.0
V* V402 Aur	172.788	-6.445	5.29	15.641	9999.0	9999.0	9999.0	6.559	9989.918	9989.918	9989.918	0.098	9999.0	9999.0	9999.0
V* V579 Per	153.746	-11.339	2.34	6.53	2.367	9999.0	9999.0	2.622	-1.542	9995.091	9995.091	0.221	0.187	9999.0	9999.0
V* AA Tau	174.321	-15.395	7.29	18.869	13.312	6.961	-15.742	8.423	2.866	-3.484	-26.187	1.633	1.049	0.696	0.55
V* DG Tau	171.842	-15.672	8.25	16.321	11.09	9999.0	9999.0	6.536	1.305	9989.214	9989.214	0.74	0.215	9999.0	9999.0
* eps Per	157.354	-10.088	12.16	8.589	9999.0	9999.0	9999.0	3.709	9994.12	9994.12	9994.12	0.007	9999.0	9999.0	9999.0
* eps Per	157.354	-10.088	12.16	5.995	9999.0	9999.0	9999.0	1.115	9994.12	9994.12	9994.12	0.005	9999.0	9999.0	9999.0
HD 24640	160.469	-13.974	2.32	9.532	9999.0	9999.0	9999.0	3.241	9992.708	9992.708	9992.708	0.082	9999.0	9999.0	9999.0
HD 27742	175.317	-19.665	5.64	19.477	23.345	16.673	9999.0	8.411	12.28	5.608	9987.935	0.309	0.299	0.21	9999.0
HD 31648	173.466	-7.902	6.18	15.406	9999.0	9999.0	9999.0								
V* IQ Aur	172.865	-2.128	6.88	-2.551	9999.0	9999.0	9999.0	-11.121	9990.431	9990.431	9990.431	0.009	9999.0	9999.0	9999.0
* omi Per	160.364	-17.74	2.91	15.987	10.026	13.145	9999.0	9.32	3.358	6.477	9992.332	0.116	0.17	0.629	9999.0
* ksi Per	160.372	-13.107	1.38	6.104	9.326	13.818	9999.0	-0.061	3.161	7.654	9992.835	0.211	0.205	0.064	9999.0
V* XY Per	156.796	-11.902	2.14	9.024	14.24	5.697	0.935	4.097	9.313	0.77	-3.992	0.282	0.166	0.294	0.109

Appendix C

Others NARVAL and ESPaDOnS fits

As I cannot use all the NARVAL and ESPaDOnS spectra available and fitted in potassium to make a kinematic analysis like for Taurus, I present here the results for future employments.

Table C.1 – Fit NARVAL and ESPaDOnS spectra.

name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
19cep	104.871	5.391	1.16	-12.412	-20.234	-16.4	-8.765	20000.0	20000.0	20000.0	20000.0	0.345	0.112	0.252	0.123
19cep	104.871	5.391	1.16	-20.755	-16.743	-12.749	-9.191	20000.0	20000.0	20000.0	20000.0	0.091	0.244	0.338	0.144
19cep	104.871	5.391	1.16	-9.14	-12.649	-16.672	-20.719	20000.0	20000.0	20000.0	20000.0	0.135	0.341	0.254	0.095
19cep	104.871	5.391	1.16	-12.587	-20.46	-16.577	-8.976	20000.0	20000.0	20000.0	20000.0	0.344	0.096	0.249	0.133
19cep	104.871	5.391	1.16	-20.185	-16.329	-12.309	-8.719	20000.0	20000.0	20000.0	20000.0	0.108	0.257	0.344	0.112
19cep	104.871	5.391	1.16	-12.555	-20.4	-16.594	-8.936	20000.0	20000.0	20000.0	20000.0	0.347	0.087	0.251	0.13
19cep	104.871	5.391	1.16	-16.462	-20.203	-12.491	-8.956	20000.0	20000.0	20000.0	20000.0	0.253	0.096	0.342	0.124
2massj00052394+6347207	117.874	1.37	0.9	-20.806	-3.643	-15.35	-11.506	20000.0	20000.0	20000.0	20000.0	0.146	0.095	0.387	0.337
2massj00103725+5844536	117.647	-3.704	2.32	-11.679	-1.609	-7.058	9999.0	20000.0	5965.747	20000.0	9999.0	0.1	0.019	0.125	9999.0
2massj00104636+5846102	117.67	-3.686	2.57	-8.04	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.214	9999.0	9999.0	9999.0
2massj00215670+6330397	119.65	0.827	0.26	-21.817	-15.663	-6.198	-10.69	20000.0	20000.0	20000.0	20000.0	0.152	0.331	0.406	1.136
2massj00330087+5856180	120.552	-3.848	1.29	7.686	2.104	-4.485	9999.0	20000.0	20000.0	20000.0	9999.0	0.419	0.528	0.211	9999.0
2massj00362292+6313162	121.237	0.399	1.73	-16.105	-19.743	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.13	0.462	9999.0	9999.0
2massj00380988+6251111	121.419	0.02	1.33	-12.155	-16.621	-20.128	-6.792	20000.0	20000.0	20000.0	20000.0	0.216	0.18	0.599	0.041
2massj00401172+5854561	121.478	-3.924	1.18	-10.164	-14.446	-19.741	9999.0	20000.0	20000.0	17038.043	9999.0	0.065	0.08	0.314	9999.0
2massj00535454+5945115	123.244	-3.117	1.76	-30.347	-38.645	-35.137	-43.026	20000.0	20000.0	20000.0	20000.0	0.165	1.034	0.769	0.236
2massj01022616+5822498	124.377	-4.46	0.99	-21.644	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.286	9999.0	9999.0	9999.0
2massj02102704+4846405	136.096	-12.064	1.3	-3.545	50.0	9999.0	9999.0	19999.386	19905.0	9999.0	9999.0	0.079	0.0	9999.0	9999.0
2massj02205086+5519394	135.482	-5.35	0.23	-11.075	-4.274	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.084	0.609	9999.0	9999.0
2massj02255659+5500312	136.279	-5.395	0.3	-11.546	0.745	-6.067	9999.0	20000.0	20000.0	20000.0	9999.0	0.094	0.037	0.067	9999.0
2massj11073631-1036344	265.733	44.633	1.23	32.966	37.45	41.647	27.006	20000.0	20000.0	20000.0	20000.0	0.287	0.302	0.173	0.127
2massj11260029+0357048	257.853	59.128	5.42	-15.926	-21.047	-11.361	-5.877	20000.0	20000.0	20000.0	20000.0	0.349	0.132	0.347	0.28
2massj11522069-1107291	280.442	49.149	2.73	-2.656	9999.0	9999.0	9999.0	3698.441	9999.0	9999.0	9999.0	0.024	9999.0	9999.0	9999.0

Table C.2 – Fit NARVAL and ESPaDOnS spectra.

name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
2massj11572921-1701023	284.884	43.956	2.67	-1.376	-2.89	9999.0	9999.0	2677.895	1000.019	9999.0	9999.0	0.015	0.011	9999.0	9999.0
2massj12001800+1619268	252.679	73.966	5.55	-14.804	-10.224	-5.388	9999.0	20000.0	20000.0	20000.0	9999.0	0.448	0.79	0.319	9999.0
2massj12033383-1728547	286.998	43.904	5.23	-11.635	-16.667	-22.224	-5.164	20000.0	20000.0	20000.0	20000.0	0.369	0.428	0.253	0.181
2massj14300097+0928247	0.404	60.802	2.27	-50.0	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.564	9999.0	9999.0	9999.0
2massj14300097+0928247	0.404	60.802	2.27	-35.439	-44.395	-40.57	-50.0	20000.0	20000.0	20000.0	20000.0	0.095	0.571	0.417	0.098
2massj15145975+0005386	0.817	46.201	6.74	-2.83	0.719	6.592	-6.238	20000.0	20000.0	20000.0	20000.0	0.684	0.226	0.025	0.277
2massj15183903-0010040	1.39	45.339	6.01	-4.346	1.571	-9.391	-15.483	20000.0	20000.0	20000.0	20000.0	0.407	0.218	0.359	0.249
2massj15191467-0210021	359.419	43.906	5.61	-16.115	-13.95	9999.0	9999.0	2084.718	15568.986	9999.0	9999.0	0.048	0.327	9999.0	9999.0
2massj15564545-1112464	358.747	30.94	1.87	-10.991	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.445	9999.0	9999.0	9999.0
2massj16024746-1037567	0.382	30.198	6.78	-13.983	-9.312	-4.024	-18.888	20000.0	20000.0	20000.0	20000.0	0.391	0.402	0.161	0.139
2massj16032651-1122030	359.855	29.601	4.08	-12.224	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	1.211	9999.0	9999.0	9999.0
2massj16032651-1122030	359.855	29.601	4.08	-5.756	-11.876	-0.223	-15.905	20000.0	20000.0	20000.0	20000.0	0.383	0.833	0.203	0.303
2massj16044607-1213451	359.347	28.794	1.23	29.068	32.927	20.263	25.463	20000.0	20000.0	20000.0	20000.0	1.211	0.327	0.048	0.583
2massj16044607-1213451	359.347	28.794	1.23	29.306	33.483	21.221	25.736	20000.0	20000.0	20000.0	20000.0	1.148	0.312	0.105	0.625
2massj16050283-1113179	0.275	29.395	4.77	5.279	9999.0	9999.0	9999.0	14740.837	9999.0	9999.0	9999.0	0.051	9999.0	9999.0	9999.0
2massj17593838+0351569	30.462	13.281	0.92	-3.925	-19.823	-17.804	-11.566	20000.0	7977.146	13350.149	20000.0	0.083	0.1	0.153	0.231
2massj17593838+0351569	30.462	13.281	0.92	-13.193	-8.954	-2.745	-18.387	20000.0	20000.0	16488.327	20000.0	0.115	0.099	0.055	0.224
2massj18004177+0308574	29.93	12.721	1.86	-0.424	9999.0	9999.0	9999.0	731.052	9999.0	9999.0	9999.0	0.021	9999.0	9999.0	9999.0
2massj18024973+0136076	28.772	11.539	2.69	-9.552	-13.805	-18.34	-21.337	2304.469	14966.948	19999.728	10362.851	0.041	0.109	0.213	0.1
2massj18024973+0136076	28.772	11.539	2.69	-13.658	-18.873	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.116	0.227	9999.0	9999.0
2massj18030727+0057194	28.221	11.176	1.77	-18.053	-13.825	-12.072	-10.278	20000.0	14658.004	2340.882	9639.047	0.12	0.35	0.072	0.078
2massj18033168+0246138	29.915	11.919	1.75	-3.077	40.499	9999.0	9999.0	3042.867	4651.876	9999.0	9999.0	0.026	0.025	9999.0	9999.0
2massj18041166+0152399	29.182	11.362	0.68	-15.048	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.128	9999.0	9999.0	9999.0
2massj18043166+0251232	30.109	11.736	1.83	-2.968	4.752	1.171	9999.0	2880.329	3635.387	11292.897	9999.0	0.027	0.051	0.038	9999.0
2massj18051031+0237092	29.969	11.484	1.59	-1.5	-1.648	9999.0	9999.0	1000.0	1000.0	9999.0	9999.0	0.0	0.017	9999.0	9999.0
2massj18061938+0133273	29.141	10.742	1.85	-36.413	12.298	9999.0	9999.0	1003.158	1000.0	9999.0	9999.0	0.0	0.029	9999.0	9999.0
2massj18062780+0212510	29.752	11.012	2.17	-36.543	-32.406	9999.0	9999.0	20000.0	4138.52	9999.0	9999.0	0.071	0.171	9999.0	9999.0
2massj18073558+0154124	29.602	10.619	0.3	-13.063	-16.521	9999.0	9999.0	20000.0	15217.06	9999.0	9999.0	0.249	0.196	9999.0	9999.0
2massj18074980+0205065	29.794	10.65	1.4	-15.197	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.239	9999.0	9999.0	9999.0
2massj18075408+0305240	30.713	11.092	0.84	-11.185	-7.684	-15.197	-3.055	20000.0	20000.0	20000.0	20000.0	0.671	1.331	0.232	0.408
2massj18075746+0313237	30.84	11.14	1.91	-0.157	-1.07	9999.0	9999.0	1000.0	1347.547	9999.0	9999.0	0.038	0.088	9999.0	9999.0
2massj18101306-0000277	28.186	9.157	2.49	-5.088	-15.148	-10.746	-19.83	20000.0	20000.0	20000.0	18515.381	0.108	0.695	0.303	0.271
2massj18123623+0131238	29.841	9.332	1.73	-16.403	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	7.805	9999.0	9999.0	9999.0
2massj18134458+0103348	29.556	8.866	1.46	-15.816	-20.139	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	1.969	0.749	9999.0	9999.0
2massj18140097+0035338	29.168	8.591	0.96	-14.848	-21.358	-9.188	9999.0	20000.0	20000.0	20000.0	9999.0	1.352	0.391	0.271	9999.0
2massj18155782-0021313	28.541	7.721	2.22	-7.287	-12.254	-4.43	9999.0	20000.0	20000.0	20000.0	9999.0	0.371	0.796	0.161	9999.0
2massj18163758+0203474	30.792	8.684	0.92	-12.818	-16.513	-21.177	-7.701	20000.0	20000.0	9060.101	20000.0	0.599	1.476	0.118	0.615
2massj18164452-0023490	28.597	7.531	1.27	-12.153	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	1.351	9999.0	9999.0	9999.0
2massj18170252-0000435	28.977	7.641	1.79	-12.927	-8.171	-17.392	-4.674	20000.0	20000.0	20000.0	16815.618	0.921	0.208	0.375	0.069
2massj18180936+0135298	30.544	8.129	1.58	-16.606	-19.45	-12.088	-7.196	20000.0	20000.0	20000.0	20000.0	1.172	0.507	0.825	0.284
2massj18184381+0019470	29.479	7.424	0.75	-13.301	-16.734	-5.559	-9.605	20000.0	20000.0	20000.0	20000.0	1.188	0.617	0.165	0.398
2massj18194204-0003478	29.239	7.028	0.94	-37.23	-30.936	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.569	0.504	9999.0	9999.0
2massj18274097-0225358	28.053	4.169	2.17	-20.953	-10.324	-6.555	-14.616	4884.964	20000.0	20000.0	20000.0	0.03	1.386	0.241	0.336
2massj18474175+1326322	44.534	6.9	2.18	-6.519	-10.679	-16.343	-3.657	8346.29	20000.0	20000.0	3495.266	0.112	0.226	0.253	0.017

Table C.3 – Fit NARVAL and ESPaDOnS spectra.

name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
2massj18534172+1412156	45.873	5.935	1.52	1.418	9999.0	9999.0	9999.0	2508.357	9999.0	9999.0	9999.0	0.023	9999.0	9999.0	9999.0
2massj18544633+1428396	46.235	5.823	0.95	-2.627	-18.255	-7.702	-12.238	20000.0	20000.0	20000.0	20000.0	0.734	0.098	0.255	0.143
2massj19225969+2557167	59.521	5.148	1.46	-12.053	-8.917	-5.333	9999.0	13267.102	12041.58	4707.169	9999.0	0.401	0.364	0.073	9999.0
2massj19241521+1803224	52.674	1.18	2.03	-11.893	-9.532	-20.204	9999.0	20000.0	20000.0	20000.0	9999.0	0.743	0.45	0.024	9999.0
2massj19272539+2455174	59.077	3.78	1.62	-15.507	-12.232	-13.7	-8.452	6200.408	17984.351	1996.004	19999.778	0.119	0.277	0.046	0.132
2massj19272539+2455174	59.077	3.78	1.62	-13.09	-9.146	-5.256	9999.0	20000.0	18493.141	3541.125	9999.0	0.289	0.23	0.066	9999.0
2massj19285670+2546085	59.988	3.879	1.7	-9.631	-14.559	-5.382	9999.0	19813.51	17974.9	9124.524	9999.0	0.812	0.102	0.085	9999.0
2massj19294113+1811575	53.414	0.113	0.92	-1.489	-11.075	-7.658	-14.114	20000.0	18128.537	19999.112	8571.958	0.095	0.721	0.219	0.074
2massj19421487+2415408	60.134	0.513	0.98	-9.827	-14.196	-4.698	9999.0	20000.0	15087.822	20000.0	9999.0	1.601	0.135	0.049	9999.0
2massj19484594+2256137	59.736	-1.443	0.66	-8.438	-5.158	-12.521	9999.0	20000.0	8045.096	14618.832	9999.0	1.037	0.094	0.292	9999.0
2massj19594179+3054499	67.826	0.584	2.22	0.336	9999.0	9999.0	9999.0	928.623	9999.0	9999.0	9999.0	0.011	9999.0	9999.0	9999.0
2massj20012170+2217258	60.681	-4.267	0.16	-14.06	-4.251	0.032	-6.965	20000.0	20000.0	20000.0	20000.0	0.29	0.404	0.128	0.591
2massj20014398+4418389	79.486	7.253	0.99	-3.845	-7.564	-12.208	0.359	20000.0	20000.0	20000.0	20000.0	0.991	0.739	0.144	0.24
2massj20025554+4559129	81.042	7.941	0.38	-9.54	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.298	9999.0	9999.0	9999.0
2massj20025554+4559129	81.042	7.941	0.38	-9.975	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.272	9999.0	9999.0	9999.0
2massj20025554+4559129	81.042	7.941	0.38	-19.904	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.265	9999.0	9999.0	9999.0
2massj20135903+3632379	74.168	1.118	0.43	-1.19	-6.907	-15.017	5.936	20000.0	20000.0	20000.0	20000.0	0.341	0.22	0.213	0.052
2massj20141795+3709286	74.715	1.405	1.08	-15.839	-12.127	-8.656	-10.182	14800.445	5314.836	20000.0	3232.708	0.122	0.188	0.501	0.147
2massj20145498+3722420	74.967	1.425	0.38	-0.062	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
2massj20250713+3638161	75.509	-0.673	0.81	-0.843	-12.484	-6.31	-16.751	20000.0	20000.0	20000.0	20000.0	0.07	0.233	0.146	0.245
2massj20270260+4444576	82.34	3.723	0.99	-9.17	4.202	-0.699	-4.542	20000.0	20000.0	20000.0	20000.0	0.39	0.338	1.325	0.864
2massj20270260+4444576	82.34	3.723	0.99	-0.311	-9.434	-4.267	4.376	20000.0	20000.0	20000.0	20000.0	1.162	0.354	1.007	0.284
2massj20444908+3157167	74.175	-6.73	0.62	-9.716	-5.522	-17.726	9999.0	20000.0	20000.0	20000.0	9999.0	0.493	0.624	0.233	9999.0
2massj20451060+5112379	89.35	5.16	0.39	-11.88	-7.998	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.759	0.072	9999.0	9999.0
2massj20501250+5139580	90.199	4.832	1.05	-9.886	-14.618	-20.059	9999.0	20000.0	20000.0	20000.0	9999.0	0.127	0.377	0.796	9999.0
2massj20510469+5025102	89.317	3.938	1.15	-8.35	-16.979	-12.851	-21.307	20000.0	20000.0	20000.0	9793.522	0.38	0.193	0.253	0.119
2massj20513653+5157071	90.559	4.846	0.86	-15.252	-20.265	-10.023	-3.102	20000.0	20000.0	20000.0	20000.0	0.234	0.373	0.266	0.028
2massj20523532+5020555	89.416	3.707	1.06	-18.303	-5.279	-9.596	-13.501	20000.0	20000.0	20000.0	20000.0	0.261	0.272	1.487	0.765
2massj20550326+3928488	81.359	-3.606	0.95	-15.824	-11.024	-6.121	-0.959	20000.0	20000.0	20000.0	8412.65	0.294	0.163	0.616	0.148
2massj20564108+3957218	81.926	-3.538	0.78	-14.895	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.915	9999.0	9999.0	9999.0
2massj20565852+5013052	89.768	3.087	1.27	-22.154	-19.059	-13.997	-9.23	20000.0	20000.0	20000.0	20000.0	0.301	1.022	0.435	0.662
2massj20595186+3858384	81.581	-4.641	0.57	-5.638	-12.032	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.355	0.866	9999.0	9999.0
2massj20595186+3858384	81.581	-4.641	0.57	-11.935	-24.192	-5.439	-19.152	20000.0	20000.0	20000.0	20000.0	0.923	0.181	0.363	0.236
2massj21100235+4913175	90.428	0.847	1.5	-1.117	-5.818	-11.121	-14.397	8847.809	20000.0	19999.907	5793.101	0.069	0.467	0.285	0.076
2massj21122845+4703145	89.125	-0.931	0.85	-1.089	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.754	9999.0	9999.0	9999.0
2massj21161964+4901093	90.99	-0.038	0.63	-1.643	-6.432	-11.271	9999.0	5649.35	18496.145	20000.0	9999.0	0.05	0.881	0.263	9999.0
2massj21163884+4747154	90.143	-0.934	1.37	-28.799	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	2.433	9999.0	9999.0	9999.0
2massj21183302+6644202	104.062	12.034	0.86	-22.801	-18.396	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	1.686	0.569	9999.0	9999.0
2massj21282592+4647232	90.861	-3.049	0.68	4.393	-1.568	9999.0	9999.0	4106.964	3739.494	9999.0	9999.0	0.023	0.034	9999.0	9999.0
2massj21282648+4655259	90.955	-2.953	0.9	-8.272	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.201	9999.0	9999.0	9999.0
2massj21301511+5626264	97.704	3.766	0.94	-13.01	-20.551	-15.867	-9.508	20000.0	9734.608	20000.0	19039.835	0.426	0.063	0.328	0.703
2massj21301511+5626264	97.704	3.766	0.94	-11.204	-20.607	-8.567	-15.794	20000.0	3607.712	20000.0	20000.0	0.663	0.026	0.373	0.451
2massj21304593+5712001	98.277	4.271	1.02	-15.733	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.681	9999.0	9999.0	9999.0
2massj21304593+5712001	98.277	4.271	1.02	-15.674	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.664	9999.0	9999.0	9999.0
2massj21344455+4432322	90.128	-5.443	1.04	-0.127	-6.05	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.469	0.461	9999.0	9999.0

Table C.4 – Fit NARVAL and ESPaDOnS spectra.

name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
2massj21363278+4303344	89.363	-6.757	0.63	-7.725	-1.867	-11.205	-15.686	20000.0	20000.0	20000.0	20000.0	0.156	0.129	0.347	0.144
2massj21373102+5259450	96.146	0.52	0.66	-23.573	-13.017	-7.134	-16.997	20000.0	20000.0	20000.0	20000.0	0.056	0.175	0.244	0.314
2massj21375836+4152509	88.76	-7.806	2.36	-6.863	-6.194	-12.771	-2.412	19971.87	1707.151	20000.0	4138.66	0.178	0.04	0.086	0.017
2massj21432261+5850422	100.625	4.375	0.76	-8.566	-24.318	-13.766	-18.255	20000.0	20000.0	20000.0	20000.0	0.175	0.062	1.144	0.289
2massj21434429+4323427	90.574	-7.377	0.36	-22.666	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.562	9999.0	9999.0	9999.0
2massj21434429+4323427	90.574	-7.377	0.36	-14.399	-9.723	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.378	0.265	9999.0	9999.0
2massj21444465+4345462	90.955	-7.218	1.15	-14.732	-20.747	-12.659	-10.471	10984.411	20000.0	4538.787	8915.491	0.242	0.083	0.114	0.093
2massj21444465+4345462	90.955	-7.218	1.15	0.109	-30.482	9999.0	9999.0	16096.936	20000.0	9999.0	9999.0	0.084	0.119	9999.0	9999.0
2massj21455910+4337591	91.042	-7.462	1.77	2.151	9999.0	9999.0	9999.0	2613.568	9999.0	9999.0	9999.0	0.015	9999.0	9999.0	9999.0
2massj21462326+5212411	96.65	-0.955	0.48	2.776	2.74	-2.474	9999.0	2134.695	1911.237	4018.67	9999.0	0.0	0.009	0.011	9999.0
2massj21502003+3856054	88.589	-11.572	0.75	-34.294	-38.472	9999.0	9999.0	20000.0	10790.722	9999.0	9999.0	0.101	0.277	9999.0	9999.0
2massj21502178+5357414	98.226	0.012	2.27	-22.033	-17.417	-9.301	-12.424	20000.0	20000.0	9761.917	20000.0	0.044	0.099	0.068	0.172
2massj21502178+5357414	98.226	0.012	2.27	0.666	9999.0	9999.0	9999.0	15306.446	9999.0	9999.0	9999.0	0.05	9999.0	9999.0	9999.0
2massj21502178+5357414	98.226	0.012	2.27	1.033	9999.0	9999.0	9999.0	500.005	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
2massj21502178+5357414	98.226	0.012	2.27	-10.025	-13.53	-16.392	-4.461	19999.999	15788.985	19999.163	11867.293	0.089	0.121	0.211	0.045
2massj21534939+3951119	89.718	-11.298	0.99	-5.982	-10.166	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.184	0.166	9999.0	9999.0
2massj21541026+3952378	89.786	-11.322	0.08	-5.898	-14.275	-10.271	-17.157	15666.683	20000.0	19999.997	19999.998	0.068	0.268	0.426	0.098
2massj21551055+5326166	98.454	-0.842	0.4	-11.993	-8.318	-19.127	-4.688	20000.0	20000.0	14440.444	20000.0	0.123	0.131	0.059	0.233
2massj22032307+5129046	98.26	-3.153	0.43	-0.045	9999.0	9999.0	9999.0	2494.917	9999.0	9999.0	9999.0	0.006	9999.0	9999.0	9999.0
2massj22205423+5325463	101.565	-3.08	1.02	-10.763	-17.782	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.197	0.067	9999.0	9999.0
2massj22292816+5258587	102.416	-4.139	1.11	-31.52	-37.064	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	1.559	1.413	9999.0	9999.0
2massj22465612+5305205	104.764	-5.334	2.46	-14.993	-19.641	-7.565	-11.378	19999.991	18087.488	5311.218	16632.07	0.214	0.038	0.045	0.288
2massj23280614+5825229	112.186	-2.699	0.98	-5.603	-4.804	-7.628	-12.215	1473.424	16964.302	20000.0	20000.0	0.055	0.175	0.353	0.69
HD200775	104.062	14.193	2.77	26.505	9999.0	9999.0	9999.0	4349.043	9999.0	9999.0	9999.0	0.013	9999.0	9999.0	9999.0
HD200775	104.062	14.193	2.77	-1.5	9999.0	9999.0	9999.0	500.004	9999.0	9999.0	9999.0	0.001	9999.0	9999.0	9999.0
HD200775	104.062	14.193	2.77	-2.879	9999.0	9999.0	9999.0	1934.352	9999.0	9999.0	9999.0	0.002	9999.0	9999.0	9999.0
HD200775	104.062	14.193	2.77	47.632	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
HD200775	104.062	14.193	2.77	-5.924	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
HD200775	104.062	14.193	2.77	-13.902	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.114	9999.0	9999.0	9999.0
HD200775	104.062	14.193	2.77	-0.903	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
HD200775	104.062	14.193	2.77	-47.238	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
HD20336	137.457	7.061	6.07	-5.248	-2.671	9999.0	9999.0	6661.327	6936.148	9999.0	9999.0	0.011	0.015	9999.0	9999.0
HD20336	137.457	7.061	6.07	-3.756	9999.0	9999.0	9999.0	12492.669	9999.0	9999.0	9999.0	0.018	9999.0	9999.0	9999.0
HD20336	137.457	7.061	6.07	-3.813	9999.0	9999.0	9999.0	15484.273	9999.0	9999.0	9999.0	0.02	9999.0	9999.0	9999.0
HD20336	137.457	7.061	6.07	1.574	-3.827	9999.0	9999.0	1000.0	17893.517	9999.0	9999.0	0.0	0.016	9999.0	9999.0
HD20336	137.457	7.061	6.07	-25.859	-5.516	-3.407	9999.0	1076.93	2445.164	7716.336	9999.0	0.0	0.007	0.015	9999.0
HD20336	137.457	7.061	6.07	-3.916	22.871	9999.0	9999.0	5676.882	1001.074	9999.0	9999.0	0.017	0.0	9999.0	9999.0
HD20336	137.457	7.061	6.07	-3.561	9999.0	9999.0	9999.0	6858.797	9999.0	9999.0	9999.0	0.01	9999.0	9999.0	9999.0
HD20336	137.457	7.061	6.07	-3.627	4.283	9999.0	9999.0	7291.274	1017.165	9999.0	9999.0	0.017	0.0	9999.0	9999.0
HD209339	104.578	5.869	1.36	-18.004	-22.53	-10.337	-14.246	20000.0	20000.0	20000.0	20000.0	0.446	0.2	0.083	0.222
HD209339	104.578	5.869	1.36	-10.1	-22.376	-17.753	-13.809	20000.0	20000.0	20000.0	20000.0	0.069	0.212	0.153	0.203
bd+453341	86.982	0.692	0.36	-11.606	-16.845	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	1.953	0.819	9999.0	9999.0
bd+523122	99.807	-2.435	9999.0	-39.43	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.471	9999.0	9999.0	9999.0
bd+523122	99.807	-2.435	9999.0	-39.22	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.489	9999.0	9999.0	9999.0
bd+523135	100.637	-2.133	0.16	-11.998	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.405	9999.0	9999.0	9999.0

Table C.5 – Fit NARVAL and ESPaDOnS spectra.

name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
bd+552770	104.764	-1.403	9999.0	-20.005	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.47	9999.0	9999.0	9999.0
bd+602380	105.367	3.203	1.07	-17.99	-25.6	-12.643	-8.784	20000.0	20000.0	20000.0	20000.0	0.166	0.076	1.361	0.14
bd+622142	110.622	3.233	1.13	-5.573	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.475	9999.0	9999.0	9999.0
bd+622154	110.938	3.257	1.19	-30.72	-25.535	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.752	1.161	9999.0	9999.0
bd+65121	123.94	3.065	1.22	-21.672	-11.385	-50.0	-16.536	20000.0	20000.0	3686.163	20000.0	0.846	1.455	0.005	1.87
bd+65122	123.931	3.585	1.49	-19.438	-13.144	-9.162	-4.492	20000.0	20000.0	20000.0	20000.0	0.273	0.325	0.888	0.383
bd+65131	124.671	3.399	0.91	-17.218	-22.376	-11.158	-6.278	20000.0	20000.0	20000.0	20000.0	1.476	0.136	1.091	0.282
bd+65131	124.671	3.399	0.91	-21.249	-6.379	-17.143	-11.18	20000.0	20000.0	20000.0	20000.0	0.138	0.306	1.403	1.071
bd+65142	125.079	3.38	1.24	-9.268	-13.17	-17.293	-4.819	20000.0	20000.0	20000.0	20000.0	1.099	1.482	0.902	0.116
bd+65154	126.026	3.282	1.23	-11.785	-15.125	1.468	-6.671	20000.0	20000.0	20000.0	20000.0	1.548	0.389	0.038	0.636
bd+65154	126.026	3.282	1.23	-16.621	-12.266	-5.699	-8.06	20000.0	20000.0	20000.0	20000.0	0.21	1.649	0.281	0.502
bd+651637	105.391	9.892	1.12	-18.504	-23.145	-28.485	-13.279	20000.0	20000.0	20000.0	20000.0	0.38	0.748	0.129	0.134
bd+651637	105.391	9.892	1.12	-23.023	-17.88	-28.483	-12.245	20000.0	20000.0	20000.0	20000.0	0.8	0.357	0.149	0.093
bd+651637	105.391	9.892	1.12	-18.314	-23.095	-28.412	-13.14	20000.0	20000.0	20000.0	20000.0	0.383	0.769	0.145	0.109
bd+651637	105.391	9.892	1.12	-18.231	-23.109	-28.571	-12.513	20000.0	20000.0	20000.0	20000.0	0.355	0.755	0.131	0.112
bd+66100	125.358	4.749	1.09	-12.286	-7.274	-16.108	-2.189	20000.0	20000.0	20000.0	19487.0	0.366	0.208	1.467	0.074
bd+66100	125.358	4.749	1.09	-15.92	-20.004	-11.402	-5.747	20000.0	7017.142	20000.0	20000.0	1.531	0.092	0.33	0.176
bd+66104	125.717	4.732	1.15	-19.453	-15.59	-12.091	-6.314	13153.848	20000.0	20000.0	20000.0	0.082	1.088	0.257	0.15
bd+66104	125.717	4.732	1.15	-15.264	-11.039	-6.142	-18.339	20000.0	20000.0	20000.0	20000.0	1.106	0.154	0.158	0.153
bd+66111	126.06	4.776	0.93	-9.124	-14.659	-16.633	-3.532	20000.0	20000.0	20000.0	20000.0	0.243	0.806	0.304	0.061
bd+66112	126.098	4.476	1.05	-6.197	-0.813	-12.713	-17.397	20000.0	20000.0	20000.0	20000.0	1.872	0.809	1.437	1.276
bd+66116	126.417	4.344	1.41	-15.378	-12.851	-6.633	-3.167	20000.0	20000.0	20000.0	20000.0	0.869	0.33	0.188	0.091
bd+661673	118.196	5.093	1.02	-15.921	-21.782	-10.131	-27.379	20000.0	20000.0	20000.0	20000.0	1.963	2.46	1.181	0.369
bd+661673	118.196	5.093	1.02	-21.862	-27.601	-10.038	-15.93	20000.0	20000.0	20000.0	20000.0	2.39	0.331	1.197	2.005
bd+661673	118.196	5.093	1.02	-27.704	-21.909	-15.99	-10.131	20000.0	20000.0	20000.0	20000.0	0.336	2.43	1.991	1.199
bd+6663	122.73	4.634	2.24	-10.55	-14.419	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.311	1.116	9999.0	9999.0
bd+6663	122.73	4.634	2.24	-14.504	-10.969	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	1.064	0.339	9999.0	9999.0
bd+6671	123.379	4.51	1.32	-18.804	-14.766	-10.257	-6.154	12751.383	20000.0	20000.0	20000.0	0.085	1.506	0.291	0.082
bd+6689	124.082	4.098	0.97	-13.585	-17.049	-9.773	-21.45	20000.0	20000.0	20000.0	20000.0	1.069	0.998	0.435	0.382
hd168963	30.546	6.95	2.6	-11.298	-14.801	-18.12	-8.547	7233.377	14460.09	10345.152	3087.356	0.308	1.868	0.296	0.063
hd185418	53.602	-2.171	1.32	-6.668	-10.761	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.223	0.65	9999.0	9999.0
hd185418	53.602	-2.171	1.32	-10.623	-12.982	-7.31	-2.237	16604.189	11025.645	19998.351	8482.387	0.538	0.132	0.277	0.019
hd190864	72.467	2.018	0.49	0.404	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.286	9999.0	9999.0	9999.0
hd190864	72.467	2.018	0.49	0.443	-7.159	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.258	0.102	9999.0	9999.0
hd191396	74.882	2.928	0.58	-6.507	-10.697	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.338	0.132	9999.0	9999.0
hd191396	74.882	2.928	0.58	-4.426	-11.823	-7.489	9999.0	12430.243	20000.0	20000.0	9999.0	0.098	0.073	0.32	9999.0
hd191611	73.621	1.852	0.37	-20.63	-16.031	-5.778	-11.425	20000.0	20000.0	20000.0	20000.0	0.25	0.384	0.169	0.237
hd211880	106.668	5.289	1.04	-8.143	-15.454	-11.235	-19.327	20000.0	20000.0	20000.0	20000.0	0.099	0.913	0.42	0.228
hd216248	107.726	-0.589	0.26	-23.28	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.433	9999.0	9999.0	9999.0
hd217782	102.517	-15.759	7.74	-48.958	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
hd217782	102.517	-15.759	7.74	-49.807	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
hd217782	102.517	-15.759	7.74	-49.385	-45.487	9999.0	9999.0	1000.0	1000.0	9999.0	9999.0	0.0	0.0	9999.0	9999.0
hd217919	111.266	3.309	1.16	-11.942	-3.843	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.349	0.323	9999.0	9999.0
hd217919	111.266	3.309	1.16	-3.874	-11.855	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.314	0.335	9999.0	9999.0

Table C.6 – Fit NARVAL and ESPaDOnS spectra.

name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
hd218323	111.799	3.728	1.07	-23.406	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	1.233	9999.0	9999.0	9999.0
hd218323	111.799	3.728	1.07	-23.431	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	1.261	9999.0	9999.0	9999.0
hd224424	116.21	-2.45	0.27	-9.912	-16.811	-23.797	-6.6	20000.0	20000.0	20000.0	20000.0	0.418	0.053	0.082	0.131
hd224424	116.21	-2.45	0.27	-16.608	-24.178	-6.522	-9.949	20000.0	20000.0	20000.0	20000.0	0.068	0.083	0.128	0.42
hd229049	76.995	1.351	0.43	-9.502	-25.614	-4.529	-14.664	20000.0	12430.872	20000.0	20000.0	0.053	0.075	0.053	0.315
hd2619	120.74	2.49	0.96	-14.899	-19.829	-11.039	-7.218	20000.0	20000.0	20000.0	20000.0	1.02	0.117	1.869	0.338
hd2619	120.74	2.49	0.96	-19.0	-14.685	-11.149	-7.854	20000.0	20000.0	20000.0	20000.0	0.143	0.936	1.75	0.41
hd344784	59.402	-0.153	1.74	-10.608	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.946	9999.0	9999.0	9999.0
hd344880	60.287	-0.309	0.47	-2.149	-7.931	-11.81	9999.0	20000.0	20000.0	20000.0	9999.0	0.549	0.606	1.392	9999.0
hd344894	59.616	-0.724	0.48	-10.606	-13.342	-7.092	-5.217	16557.288	20000.0	6867.562	16667.665	0.976	0.427	0.335	0.189
hd345214	60.042	-2.602	3.25	-5.762	-10.549	1.795	-1.237	20000.0	20000.0	20000.0	7753.595	0.514	1.114	0.021	0.102
hd71155	227.779	18.892	27.48	-49.883	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
hd71155	227.779	18.892	27.48	22.548	49.888	9999.0	9999.0	1000.0	1000.0	9999.0	9999.0	0.0	0.0	9999.0	9999.0
hd71155	227.779	18.892	27.48	-1.919	5.91	9999.0	9999.0	1000.0	1000.0	9999.0	9999.0	0.0	0.0	9999.0	9999.0
hd71155	227.779	18.892	27.48	7.375	29.505	9999.0	9999.0	1216.679	6585.99	9999.0	9999.0	0.0	0.008	9999.0	9999.0
hip100314	100.483	16.563	3.14	1.082	9999.0	9999.0	9999.0	5469.395	9999.0	9999.0	9999.0	0.018	9999.0	9999.0	9999.0
hip100936	101.917	16.584	3.22	-27.708	-23.49	-15.357	-19.887	20000.0	20000.0	20000.0	20000.0	0.12	0.699	0.234	1.191
hip101770	94.964	10.729	2.25	-24.327	-19.664	-15.477	-11.455	20000.0	20000.0	20000.0	20000.0	0.253	1.056	1.415	0.677
hip101903	92.779	8.906	2.82	11.625	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.147	9999.0	9999.0	9999.0
hip102187	94.323	9.487	3.29	-24.247	-27.604	-31.74	-20.224	20000.0	20000.0	20000.0	20000.0	0.932	0.758	0.212	0.245
hip102449	85.574	2.09	0.02	-10.861	-20.201	-5.495	-16.697	20000.0	20000.0	20000.0	20000.0	0.519	0.261	0.1	0.47
hip102518	102.434	14.779	2.32	-11.386	-15.386	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.159	0.098	9999.0	9999.0
hip102699	108.664	18.844	3.28	-22.735	-31.997	-15.656	-27.089	20000.0	20000.0	20000.0	20000.0	0.69	0.49	0.883	1.614
hip102926	93.274	7.131	1.15	-7.223	-12.659	-16.723	-20.079	20000.0	20000.0	20000.0	20000.0	0.12	0.724	1.427	0.241
hip102926	93.274	7.131	1.15	-19.275	-12.736	-16.493	-7.201	20000.0	20000.0	20000.0	20000.0	0.367	0.696	1.276	0.101
hip102969	107.589	17.846	5.72	-0.872	-1.873	9999.0	9999.0	1000.035	2327.414	9999.0	9999.0	0.031	0.022	9999.0	9999.0
hip103188	92.23	5.681	1.94	-0.094	5.016	-5.015	-11.119	20000.0	20000.0	20000.0	20000.0	1.521	0.459	1.087	1.097
hip103245	92.267	5.584	1.54	-11.691	-17.212	-8.804	-20.347	20000.0	20000.0	20000.0	20000.0	0.73	0.714	0.289	0.722
hip103292	97.201	9.538	1.78	-22.037	-17.062	-11.369	-27.337	20000.0	20000.0	20000.0	20000.0	1.255	1.527	0.589	0.399
hip103474	103.511	14.217	3.03	-19.473	-11.397	-15.117	-6.95	20000.0	20000.0	20000.0	20000.0	0.279	1.21	1.072	0.288
hip103530	90.111	3.061	3.02	4.503	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
hip103530	90.111	3.061	3.02	21.029	33.154	9999.0	9999.0	3077.347	1070.662	9999.0	9999.0	0.058	0.0	9999.0	9999.0
hip103530	90.111	3.061	3.02	48.161	9999.0	9999.0	9999.0	14616.406	9999.0	9999.0	9999.0	0.024	9999.0	9999.0	9999.0
hip103543	107.286	16.982	5.33	-2.629	-0.67	9999.0	9999.0	1712.85	6227.412	9999.0	9999.0	0.008	0.024	9999.0	9999.0
hip103549	97.827	9.541	3.27	8.028	-9.235	2.352	-2.979	20000.0	20000.0	20000.0	20000.0	0.206	0.136	0.311	0.26
hip103606	94.378	6.519	3.25	-9.062	-17.042	-13.361	-21.549	20000.0	20000.0	20000.0	20000.0	0.308	1.041	1.116	0.282
hip103619	108.956	18.12	3.9	-10.177	-2.751	-6.11	1.097	20000.0	20000.0	20000.0	20000.0	0.164	1.012	0.773	0.229
hip103763	104.062	14.193	2.77	-26.927	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.032	9999.0	9999.0	9999.0
hip103763	104.062	14.193	2.77	-2.097	3.736	9999.0	9999.0	1000.0	1014.506	9999.0	9999.0	0.002	0.006	9999.0	9999.0
hip103772	99.319	10.315	2.3	-11.883	-7.875	9999.0	9999.0	20000.0	7085.26	9999.0	9999.0	0.626	0.061	9999.0	9999.0
hip103772	99.319	10.315	2.3	-15.757	-13.089	-11.039	-11.127	8422.305	6743.493	20000.0	1525.435	0.069	0.183	0.478	0.14
hip103875	106.829	16.22	3.07	-17.63	-22.295	-14.102	-9.892	20000.0	20000.0	20000.0	20000.0	1.271	0.211	1.126	0.256

Table C.7 – Fit NARVAL and ESPaDOnS spectra.

name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
hip104211	92.133	3.106	2.04	-11.236	-7.362	-4.443	9999.0	20000.0	20000.0	20000.0	9999.0	0.32	0.957	0.287	9999.0
hip104366	99.855	9.71	1.76	-22.021	-12.582	-5.894	-16.797	20000.0	20000.0	20000.0	20000.0	0.021	0.455	0.028	0.352
hip104573	95.012	4.899	1.85	-16.495	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.298	9999.0	9999.0	9999.0
hip104573	95.012	4.899	1.85	-16.008	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.311	9999.0	9999.0	9999.0
hip104849	108.94	16.789	6.73	16.784	49.871	9999.0	9999.0	1000.0	19999.909	9999.0	9999.0	0.0	0.0	9999.0	9999.0
hip105137	105.389	13.323	3.13	-16.874	-21.287	-8.977	-13.244	20000.0	20000.0	20000.0	20000.0	0.712	0.168	0.293	1.139
hip105490	108.272	15.451	3.66	-1.455	-0.69	-1.523	9999.0	1000.027	1000.0	1000.0	9999.0	0.0	0.019	0.005	9999.0
hip105490	108.272	15.451	3.66	22.702	9999.0	9999.0	9999.0	2265.483	9999.0	9999.0	9999.0	0.005	9999.0	9999.0	9999.0
hip105635	106.482	13.609	1.85	-10.766	-15.858	-18.198	-17.587	8854.747	10753.758	15452.996	1000.0	0.018	0.161	0.369	0.097
hip105635	106.482	13.609	1.85	-17.546	-15.32	-17.666	-9.652	1000.0	2760.785	17545.803	20000.0	0.075	0.04	0.412	0.027
hip105690	97.063	4.272	1.01	-11.731	-22.98	-17.183	-7.54	20000.0	20000.0	20000.0	20000.0	0.309	0.046	1.371	0.103
hip105891	94.891	1.551	6.06	-0.134	9999.0	9999.0	9999.0	6492.165	9999.0	9999.0	9999.0	0.021	9999.0	9999.0	9999.0
hip106585	104.449	10.029	2.72	1.904	9999.0	9999.0	9999.0	1491.28	9999.0	9999.0	9999.0	0.013	9999.0	9999.0	9999.0
hip106733	104.936	10.243	1.54	-10.835	-15.086	-6.107	-20.9	20000.0	20000.0	20000.0	20000.0	0.871	1.0	0.339	0.217
hip106830	100.339	5.054	1.28	-20.772	-11.512	-8.321	-16.454	20000.0	20000.0	20000.0	20000.0	0.207	1.161	0.212	1.977
hip106948	108.82	14.027	5.07	3.727	9999.0	9999.0	9999.0	13489.869	9999.0	9999.0	9999.0	0.015	9999.0	9999.0	9999.0
hip107025	108.429	13.507	3.01	-30.605	-25.928	-11.238	-20.818	20000.0	20000.0	20000.0	20000.0	0.559	1.452	0.153	0.45
hip107025	108.429	13.507	3.01	-30.555	-25.929	-20.877	-11.115	20000.0	20000.0	20000.0	20000.0	0.558	1.443	0.47	0.141
hip107123	100.26	4.234	0.93	-12.298	-15.961	-5.003	9999.0	20000.0	20000.0	20000.0	9999.0	0.778	0.787	0.113	9999.0
hip107420	109.682	14.266	3.0	-10.424	-13.647	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.112	0.266	9999.0	9999.0
hip107449	111.372	16.045	2.55	-16.63	-14.322	-10.79	9999.0	19999.989	20000.0	20000.0	9999.0	0.697	1.178	0.686	9999.0
hip107480	100.222	3.335	1.04	-14.564	-18.253	-11.406	-4.627	20000.0	20000.0	20000.0	20000.0	0.963	0.079	0.649	0.04
hip107955	110.304	14.171	3.03	-15.557	-11.889	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.229	0.121	9999.0	9999.0
hip107966	109.726	13.492	2.83	-47.249	9999.0	9999.0	9999.0	3602.972	9999.0	9999.0	9999.0	0.017	9999.0	9999.0	9999.0
hip108653	111.078	14.103	1.34	-1.723	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	8.843	9999.0	9999.0	9999.0
hip110662	104.432	0.326	0.4	-10.854	-14.994	-6.426	-20.931	20000.0	20000.0	20000.0	20000.0	0.242	0.197	0.035	0.093
hip110662	104.432	0.326	0.4	-18.937	-23.023	-10.069	-13.494	20000.0	20000.0	20000.0	20000.0	0.166	0.034	0.218	0.23
hip111161	109.046	6.448	1.38	-18.953	-24.382	-13.374	-9.252	20000.0	20000.0	20000.0	20000.0	0.63	0.255	0.577	0.741
hip112718	109.303	2.729	1.15	-9.282	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	1.098	9999.0	9999.0	9999.0
hip113065	109.865	2.745	1.17	-26.455	-7.377	-15.601	-22.127	20000.0	20000.0	20000.0	20000.0	1.608	0.441	0.516	0.723
hip113853	109.987	0.318	1.96	-12.222	-16.3	-7.657	9999.0	20000.0	9961.076	18916.935	9999.0	0.364	0.058	0.026	9999.0
hip11401	132.372	5.072	1.42	-11.849	-15.538	-7.067	-2.265	17728.33	1868.302	20000.0	20000.0	0.813	0.041	0.789	0.081
hip11401	132.372	5.072	1.42	-6.903	-2.138	-11.762	9999.0	20000.0	20000.0	19097.366	9999.0	0.778	0.064	0.833	9999.0
hip114836	107.85	-9.309	2.99	-7.793	-4.489	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.119	0.994	9999.0	9999.0
hip114892	107.854	-9.579	5.97	-2.839	-4.22	-0.825	3.978	3421.98	2740.557	1587.721	1721.02	0.005	0.012	0.002	0.02
hip114900	107.936	-9.413	2.58	-23.444	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	10.377	9999.0	9999.0	9999.0
hip114904	112.019	1.121	0.93	-18.497	-12.539	-3.354	-8.032	19732.948	20000.0	20000.0	20000.0	0.342	0.653	0.092	0.165
hip114904	112.019	1.121	0.93	-8.766	-12.669	-4.318	-18.601	20000.0	20000.0	20000.0	20000.0	0.176	0.633	0.11	0.307
hip114958	112.737	2.801	1.03	-10.928	-15.596	-6.197	-20.025	20000.0	20000.0	20000.0	20000.0	0.331	1.656	0.299	0.178
hip115169	108.783	-8.648	2.25	-11.355	-7.516	-5.152	9999.0	18928.221	10948.336	20000.0	9999.0	0.048	0.184	0.404	9999.0
hip118113	117.807	5.22	0.81	-6.743	-12.273	-23.536	-18.729	20000.0	20000.0	20000.0	20000.0	0.214	0.618	0.996	1.061
hip118113	117.807	5.22	0.81	-23.56	-18.713	-12.437	-7.334	20000.0	20000.0	20000.0	20000.0	0.996	1.061	0.577	0.241

Table C.8 – Fit NARVAL and ESPaDOnS spectra.

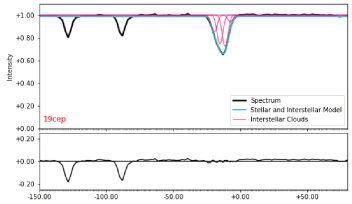
name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
hip13003	134.458	5.474	5.38	5.054	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.035	9999.0	9999.0	9999.0
hip13219	132.876	9.327	2.2	-8.304	-5.714	-12.662	-6.528	13080.914	20000.0	19119.535	1000.0	0.218	0.263	0.045	0.04
hip15647	138.663	5.487	1.53	3.99	-15.435	-6.296	-1.074	20000.0	20000.0	20000.0	20000.0	0.248	0.063	0.154	0.313
hip172	118.214	4.99	0.93	-13.86	-18.96	-22.787	-9.279	20000.0	20000.0	20000.0	20000.0	1.394	2.069	1.301	0.693
hip234	118.821	7.754	1.33	-13.634	-11.389	-19.149	-6.56	20000.0	20000.0	20000.0	20000.0	0.106	0.32	0.192	0.039
hip2390	120.74	2.49	0.96	-19.937	-14.912	-11.309	-7.941	20000.0	20000.0	20000.0	20000.0	0.121	0.945	1.777	0.503
hip26304	211.016	-20.097	4.06	1.39	9999.0	9999.0	9999.0	3573.711	9999.0	9999.0	9999.0	0.014	9999.0	9999.0	9999.0
hip2679	121.12	2.72	1.23	-8.624	-13.003	-18.34	-4.763	20000.0	20000.0	20000.0	20000.0	1.404	1.396	0.099	0.139
hip270	119.419	10.631	4.87	48.362	9999.0	9999.0	9999.0	7507.083	9999.0	9999.0	9999.0	0.008	9999.0	9999.0	9999.0
hip27545	205.65	-13.569	5.62	1.777	9999.0	9999.0	9999.0	4378.031	9999.0	9999.0	9999.0	0.013	9999.0	9999.0	9999.0
hip27854	204.649	-12.022	5.11	-6.054	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.047	9999.0	9999.0	9999.0
hip29282	205.616	-7.772	2.64	-0.908	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
hip29433	190.921	0.635	3.03	7.205	9999.0	9999.0	9999.0	500.037	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
hip29433	190.921	0.635	3.03	0.154	-0.976	9999.0	9999.0	1000.0	1000.0	9999.0	9999.0	0.002	0.003	9999.0	9999.0
hip32839	216.649	-2.103	4.12	-30.839	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
hip3349	122.114	5.306	2.83	-14.729	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	3.541	9999.0	9999.0	9999.0
hip338	118.218	3.91	0.92	-18.979	-22.997	-13.886	-9.347	20000.0	20000.0	20000.0	20000.0	1.349	1.013	1.698	0.63
hip338	118.218	3.91	0.92	-22.864	-18.749	-9.294	-13.688	20000.0	20000.0	20000.0	20000.0	1.073	1.341	0.606	1.67
hip338	118.218	3.91	0.92	-22.935	-18.86	-13.828	-9.382	20000.0	20000.0	20000.0	20000.0	1.062	1.323	1.682	0.64
hip3544	122.047	-7.64	9.07	1.735	9999.0	9999.0	9999.0	1070.952	9999.0	9999.0	9999.0	0.021	9999.0	9999.0	9999.0
hip3611	122.187	-7.559	5.92	-3.637	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.099	9999.0	9999.0	9999.0
hip3942	122.816	-6.259	1.67	2.403	9999.0	9999.0	9999.0	19999.997	9999.0	9999.0	9999.0	0.021	9999.0	9999.0	9999.0
hip4121	123.123	-6.244	9999.0	-23.293	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.24	9999.0	9999.0	9999.0
hip4121	123.123	-6.244	9999.0	-23.603	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.259	9999.0	9999.0	9999.0
hip422	118.573	5.361	2.74	-29.832	-22.89	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.896	0.719	9999.0	9999.0
hip440	118.778	6.385	5.42	0.754	9999.0	9999.0	9999.0	4206.49	9999.0	9999.0	9999.0	0.018	9999.0	9999.0	9999.0
hip4679	124.181	-7.56	2.08	-11.302	-7.206	-16.688	1.296	20000.0	20000.0	4217.292	16399.275	0.029	0.174	0.016	0.018
hip4681	124.265	-9.869	0.9	-14.129	-19.137	-6.333	-8.91	20000.0	20000.0	19117.74	20000.0	0.04	0.042	0.185	0.154
hip4897	124.355	-2.713	1.15	-9.549	-6.118	-17.246	-13.579	20000.0	9115.055	19999.986	20000.0	0.347	0.096	0.068	0.156
hip4955	124.684	-7.801	5.78	-18.854	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	1.592	9999.0	9999.0	9999.0
hip5029	124.657	-4.847	2.95	-9.491	-7.36	-10.113	9999.0	2109.873	10314.579	20000.0	9999.0	0.048	0.082	0.178	9999.0
hip5147	124.404	3.141	2.52	-8.627	-13.338	-3.904	-17.544	20000.0	17045.932	20000.0	18483.706	0.124	0.178	0.076	0.291
hip5147	124.404	3.141	2.52	-8.219	-13.207	-4.387	-17.644	20000.0	20000.0	13386.413	17955.624	0.117	0.185	0.089	0.309
hip5161	124.89	-5.06	0.97	-0.593	9999.0	9999.0	9999.0	500.002	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
hip5232	124.874	-2.955	7.53	1.227	9999.0	9999.0	9999.0	1282.062	9999.0	9999.0	9999.0	0.017	9999.0	9999.0	9999.0
hip5361	125.189	-4.533	4.1	0.919	9999.0	9999.0	9999.0	2360.49	9999.0	9999.0	9999.0	0.028	9999.0	9999.0	9999.0
hip5536	125.621	-5.995	3.38	-1.051	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.001	9999.0	9999.0	9999.0
hip5569	125.619	-5.171	4.6	-12.823	-17.729	-22.967	-8.032	20000.0	20000.0	20000.0	20000.0	0.201	0.19	0.076	0.135
hip5609	125.632	-4.513	1.54	-11.147	-13.408	-4.922	-8.197	4053.781	11239.757	12125.741	12691.111	0.068	0.124	0.133	0.272
hip5657	125.079	3.38	1.24	-17.443	-13.384	-9.55	-5.542	20000.0	20000.0	20000.0	20000.0	0.879	1.407	1.178	0.166
hip5670	125.95	-6.807	3.56	-14.778	-9.687	-2.531	9999.0	20000.0	20000.0	20000.0	9999.0	0.04	0.222	0.047	9999.0
hip56780	231.861	71.892	5.95	4.708	9999.0	9999.0	9999.0	2781.263	9999.0	9999.0	9999.0	0.004	9999.0	9999.0	9999.0
hip56780	231.861	71.892	5.95	9.354	-6.906	2.925	9999.0	11981.587	1952.095	5871.394	9999.0	0.008	0.008	0.01	9999.0
hip56920	220.981	73.561	14.62	1.834	-6.193	9999.0	9999.0	7208.718	20000.0	9999.0	9999.0	0.031	0.014	9999.0	9999.0

Table C.9 – Fit NARVAL and ESPaDOnS spectra.

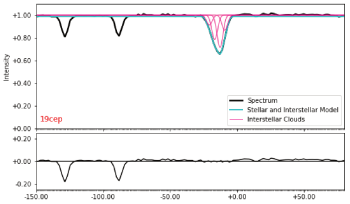
name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
hip57234	199.579	75.003	8.63	1.888	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.012	9999.0	9999.0	9999.0
hip57234	199.579	75.003	8.63	-0.057	9999.0	9999.0	9999.0	1182.339	9999.0	9999.0	9999.0	0.007	9999.0	9999.0	9999.0
hip57279	203.078	75.194	6.42	1.537	9999.0	9999.0	9999.0	19285.604	9999.0	9999.0	9999.0	0.025	9999.0	9999.0	9999.0
hip575	118.348	3.063	1.02	-13.758	-17.52	-8.98	9999.0	20000.0	20000.0	20000.0	9999.0	0.83	1.003	0.236	9999.0
hip575	118.348	3.063	1.02	-9.718	-18.068	-14.746	9999.0	20000.0	20000.0	20000.0	9999.0	0.359	0.731	1.045	9999.0
hip5755	125.819	-3.691	1.58	-10.322	-15.481	-6.765	9999.0	20000.0	20000.0	20000.0	9999.0	0.218	0.175	0.332	9999.0
hip575	118.348	3.063	1.02	-8.736	-12.987	-16.656	-19.764	20000.0	20000.0	20000.0	20000.0	0.197	0.582	1.095	0.199
hip87633	24.818	11.977	5.32	-14.676	-12.71	-15.931	-9.656	1000.017	14728.148	16334.902	1279.632	0.059	0.085	0.416	0.008
hip88148	27.589	11.659	8.48	0.339	1.332	9999.0	9999.0	1003.101	2232.41	9999.0	9999.0	0.004	0.019	9999.0	9999.0
hip88376	26.117	10.184	4.29	-13.559	-15.992	-8.233	9999.0	20000.0	13153.746	19999.936	9999.0	0.358	0.176	0.041	9999.0
hip88671	27.311	9.862	4.36	-7.178	-12.103	-17.032	-1.796	20000.0	20000.0	20000.0	20000.0	0.607	0.66	0.401	0.252
hip88698	26.715	9.461	3.59	0.37	-14.962	-32.776	9999.0	20000.0	20000.0	20000.0	9999.0	0.029	0.498	0.127	9999.0
hip88740	24.91	8.381	12.18	15.094	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
hip88753	26.238	9.029	2.74	-21.829	-16.879	-12.801	-7.75	20000.0	20000.0	20000.0	20000.0	0.086	0.113	0.909	0.156
hip88841	29.948	10.646	8.02	0.552	-0.451	9999.0	9999.0	1040.593	2394.207	9999.0	9999.0	0.006	0.004	9999.0	9999.0
hip88870	28.649	9.847	3.66	-36.457	-39.653	-32.521	-43.48	20000.0	20000.0	20000.0	20000.0	1.079	0.695	0.303	0.174
hip88878	25.405	8.154	5.61	-24.531	-28.835	-20.524	-14.584	20000.0	20000.0	20000.0	20000.0	1.092	0.318	0.58	0.141
hip89148	25.924	7.57	5.5	-5.279	9999.0	9999.0	9999.0	10616.409	9999.0	9999.0	9999.0	0.05	9999.0	9999.0	9999.0
hip89151	29.096	9.205	4.52	-0.693	-1.151	9999.0	9999.0	1000.0	1000.0	9999.0	9999.0	0.004	0.007	9999.0	9999.0
hip89472	29.003	8.082	2.88	-17.229	-12.715	9999.0	9999.0	20000.0	19602.237	9999.0	9999.0	0.06	0.447	9999.0	9999.0
hip89680	27.92	6.801	4.0	-12.5	-10.812	-10.597	-13.559	2868.543	1210.445	20000.0	20000.0	0.187	0.288	0.943	0.301
hip9278	129.488	5.039	9.09	-0.026	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.031	9999.0	9999.0	9999.0
hip99726	100.49	17.311	2.95	-17.049	-12.226	-7.853	9999.0	20000.0	20000.0	20000.0	9999.0	0.462	1.551	0.55	9999.0

Table C.10 – Fit NARVAL and ESPaDOnS spectra.

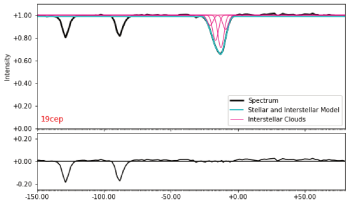
name	glon	glat	pi	v1	v2	v3	v4	T1	T2	T3	T4	tau1	tau2	tau3	tau4
TYC 132-864-1	208.985	-7.195	0.82	-22.071	9999.0	9999.0	9999.0	500.0	9999.0	9999.0	9999.0	0.0	9999.0	9999.0	9999.0
TYC 3152-864-1	77.274	-0.188	0.61	-19.065	-5.968	-1.357	-13.634	19903.442	20000.0	20000.0	20000.0	0.93	0.26	0.073	0.39
TYC 3152-864-1	77.274	-0.188	0.61	-13.817	-19.175	-6.034	-1.653	20000.0	20000.0	20000.0	20000.0	0.418	0.924	0.267	0.076
TYC 3185-1390-1	88.856	-12.655	0.92	-18.928	-11.697	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.672	0.076	9999.0	9999.0
TYC 3590-2213-1	90.339	-3.212	1.36	-18.608	-22.768	-14.982	-10.941	20000.0	20000.0	20000.0	20000.0	0.972	0.228	1.005	0.344
TYC 3682-1132-1	128.28	-3.959	0.81	-15.815	-3.213	-20.658	-9.147	20000.0	20000.0	20000.0	20000.0	0.168	0.368	0.167	0.124
TYC 3698-2525-1	134.622	-1.405	0.68	-15.782	-8.74	-4.124	0.508	20000.0	20000.0	20000.0	20000.0	0.763	0.811	1.459	0.919
TYC 3698-2525-1	134.622	-1.405	0.68	-9.134	-15.968	0.576	-4.323	20000.0	20000.0	20000.0	20000.0	0.785	0.706	0.874	1.566
TYC 3990-1799-1	103.694	-0.511	0.61	-14.867	-10.628	-22.173	-8.159	20000.0	11077.741	20000.0	1510.133	0.133	0.357	0.222	0.041
TYC 3990-1799-1	103.694	-0.511	0.61	-21.702	-11.647	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.179	0.311	9999.0	9999.0
TYC 4014-3145-1	117.478	-1.798	0.36	-27.578	-18.77	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.154	0.338	9999.0	9999.0
TYC 4014-3145-1	117.478	-1.798	0.36	-18.728	-27.006	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.33	0.16	9999.0	9999.0
TYC 4018-914-1	117.379	0.559	0.13	-11.352	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.821	9999.0	9999.0	9999.0
TYC 4018-914-1	117.379	0.559	0.13	-11.235	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	0.841	9999.0	9999.0	9999.0
TYC 4287-238-1	113.503	2.914	1.09	-0.055	9999.0	9999.0	9999.0	20000.0	9999.0	9999.0	9999.0	1.745	9999.0	9999.0	9999.0
TYC 3045-332-1	62.412	59.364	3.73	-39.405	-34.749	-25.194	-30.913	20000.0	20000.0	16037.143	20000.0	0.176	0.634	0.091	0.44
TYC 430-1922-1	28.957	10.605	1.42	-14.311	-12.206	9999.0	9999.0	19999.994	16635.024	9999.0	9999.0	0.275	0.145	9999.0	9999.0
TYC 431-1605-1	28.121	9.491	2.84	-9.653	-17.795	-11.665	-14.162	13517.203	20000.0	3381.555	13481.353	0.131	0.316	0.083	0.545
TYC 431-394-1	29.805	9.917	1.63	-26.066	-18.667	-14.658	-10.516	4667.427	19999.999	20000.0	20000.0	0.027	0.1	1.056	0.385
TYC 431-812-1	28.749	9.323	2.22	-19.691	-14.317	9999.0	9999.0	20000.0	20000.0	9999.0	9999.0	0.421	1.341	9999.0	9999.0
TYC 431-835-1	28.933	9.481	0.84	-20.762	-17.351	-13.797	9999.0	20000.0	20000.0	20000.0	9999.0	0.331	1.89	0.267	9999.0
TYC 432-1126-1	29.282	7.232	2.05	-8.419	-25.575	-17.901	-13.226	20000.0	20000.0	20000.0	20000.0	0.267	0.059	0.577	1.362
TYC 434-4617-1	29.621	10.893	2.08	9999.0	9999.0	9999.0	9999.0	9999.0	9999.0	9999.0	9999.0	9999.0	9999.0	9999.0	9999.0



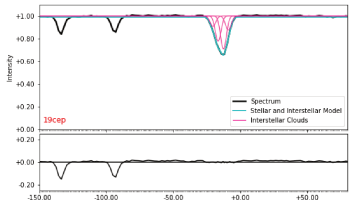
(a) 19cep



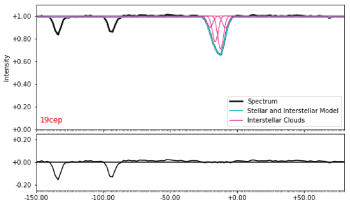
(b) 19cep



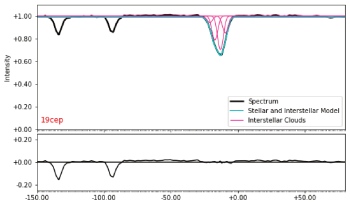
(c) 19cep



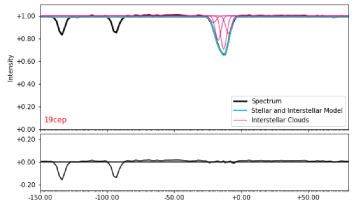
(d) 19cep



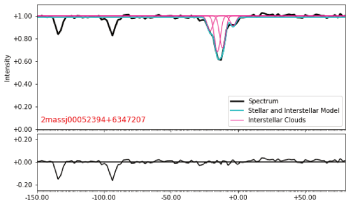
(e) 19cep



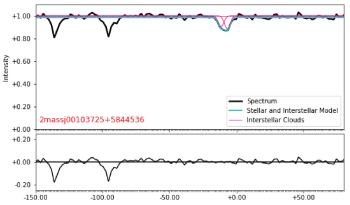
(f) 19cep



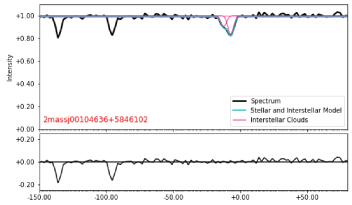
(g) 19cep



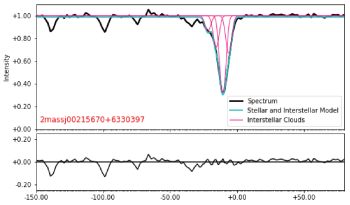
(h) 2massj00052394+6347207



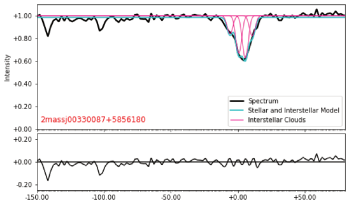
(i) 2massj00103725+5844536



(j) 2massj00104636+5846102

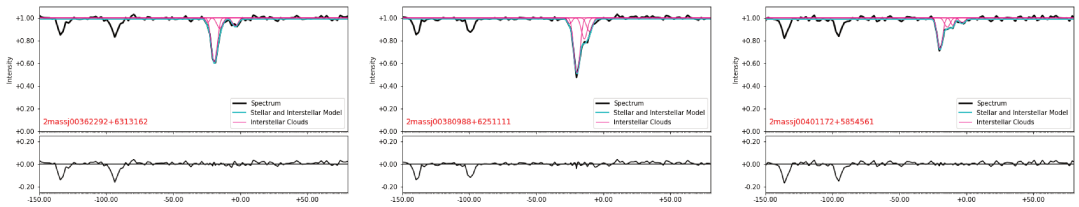


(k) 2massj00215670+6330397



(l) 2massj00330087+5856180

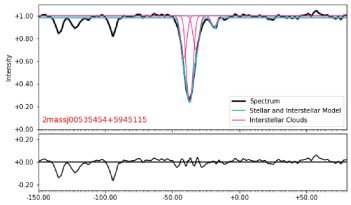
Figure C.1 – .



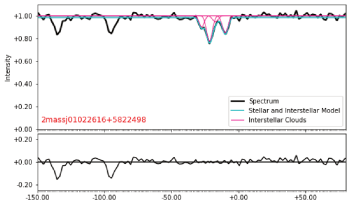
(a) 2massj00362292+6313162

(b) 2massj00380988+6251111

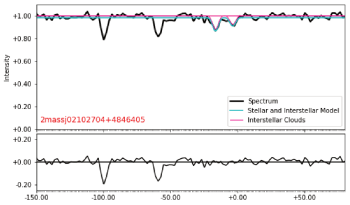
(c) 2massj00401172+5854561



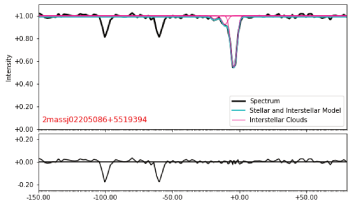
(d) 2massj00535454+5945115



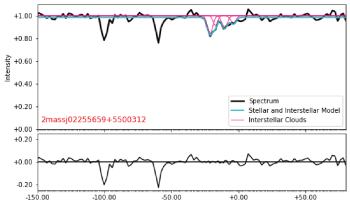
(e) 2massj01022616+5822498



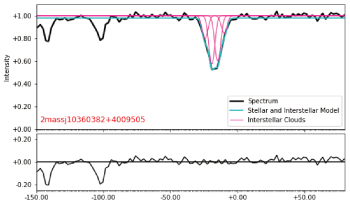
(f) 2massj02102704+4846405



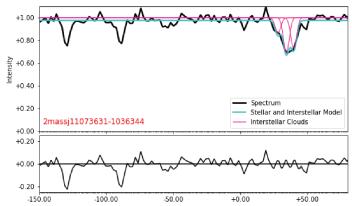
(g) 2massj02205086+5519394



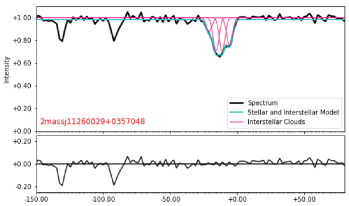
(h) 2massj02255659+5500312



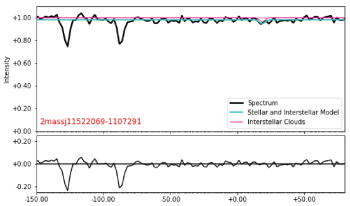
(i) 2massj10360382+4009505



(j) 2massj11073631-1036344

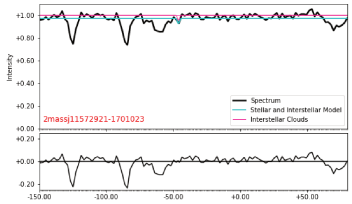


(k) 2massj11260029+0357048

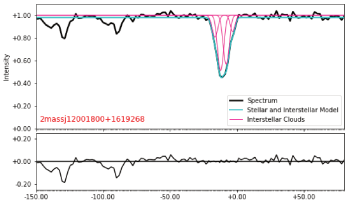


(l) 2massj11522069-1107291

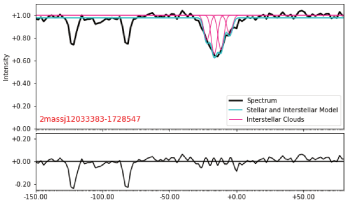
Figure C.2 – .



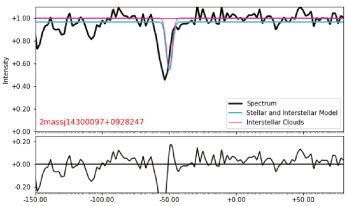
(a) 2massj11572921-1701023



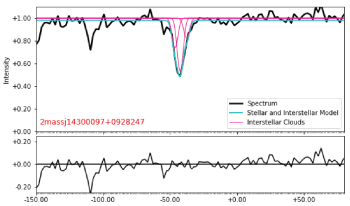
(b) 2massj12001800+1619268



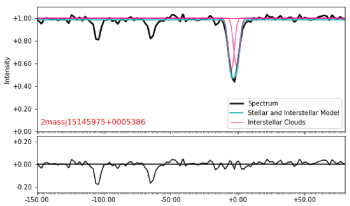
(c) 2massj12033383-1728547



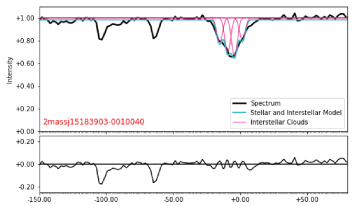
(d) 2massj14300097+0928247



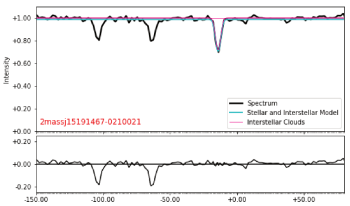
(e) 2massj14300097+0928247



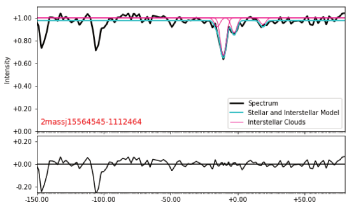
(f) 2massj15145975+0005386



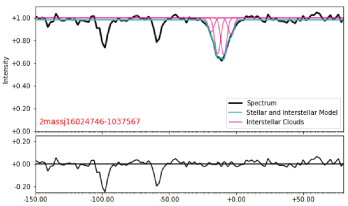
(g) 2massj15183903-0010040



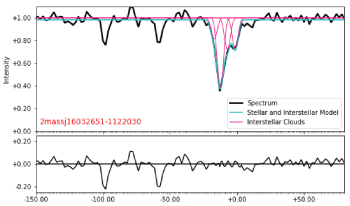
(h) 2massj15191467-0210021



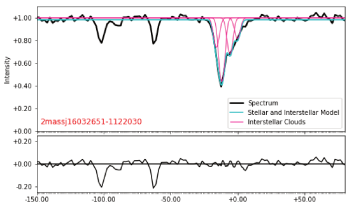
(i) 2massj15564545-1112464



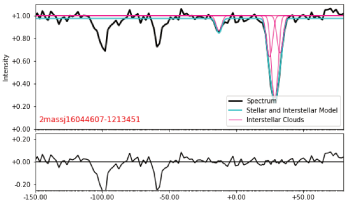
(j) 2massj16024746-1037567



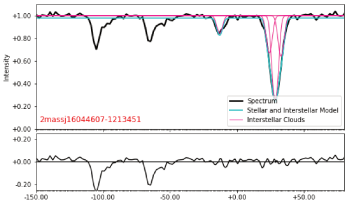
(k) 2massj16032651-1122030



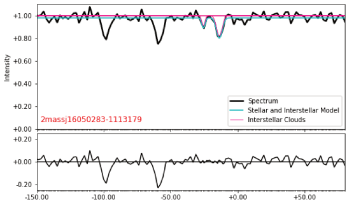
(l) 2massj16032651-1122030



(m) 2massj16044607-1213451

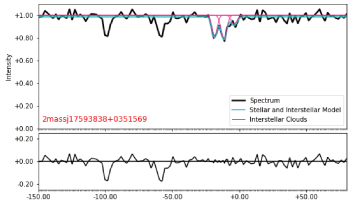


(n) 2massj16044607-1213451

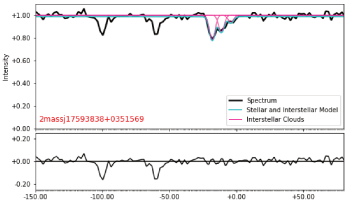


(o) 2massj16050283-1113179

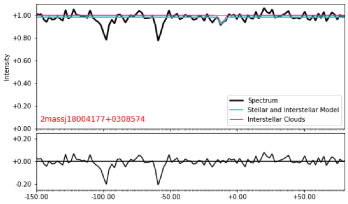
Figure C.3 – .



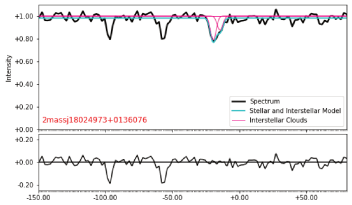
(a) 2massj17593838+0351569



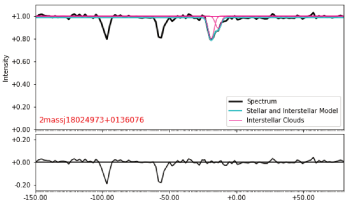
(b) 2massj17593838+0351569



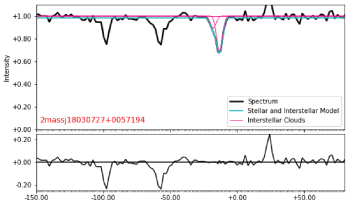
(c) 2massj18004177+0308574



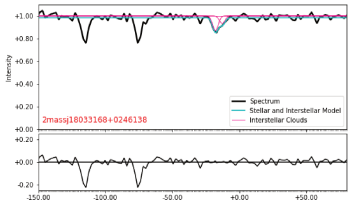
(d) 2massj18024973+0136076



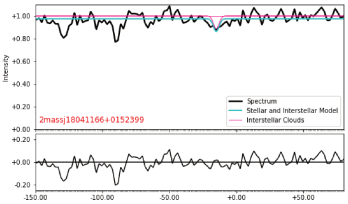
(e) 2massj18024973+0136076



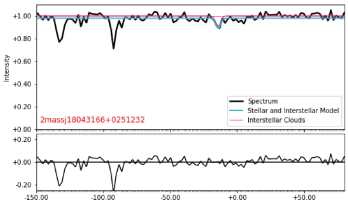
(f) 2massj18030727+0057194



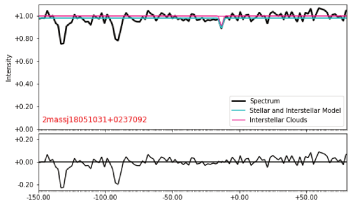
(g) 2massj18033168+0246138



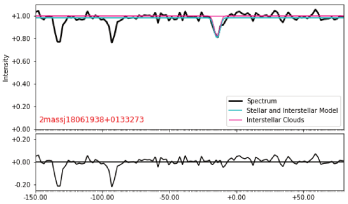
(h) 2massj18041166+0152399



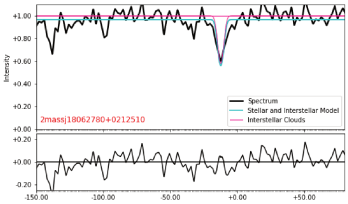
(i) 2massj18043166+0251232



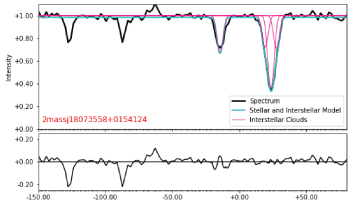
(j) 2massj18051031+0237092



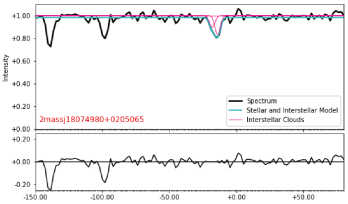
(k) 2massj18061938+0133273



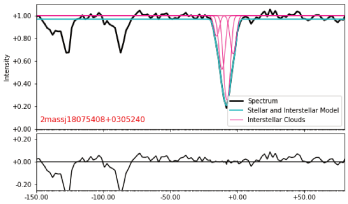
(l) 2massj18062780+0212510



(m) 2massj18073558+0154124

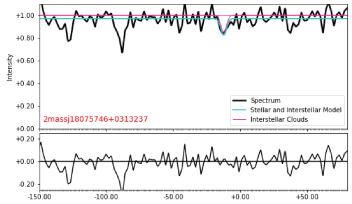


(n) 2massj18074980+0205065

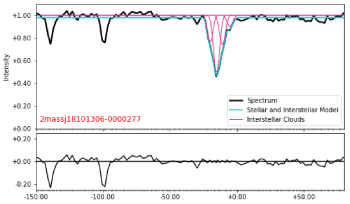


(o) 2massj18075408+0305240

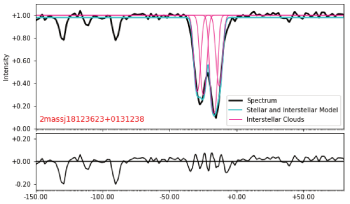
Figure C.4 – .



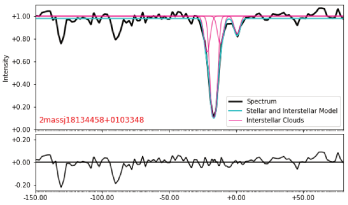
(a) 2massj18075746+0313237



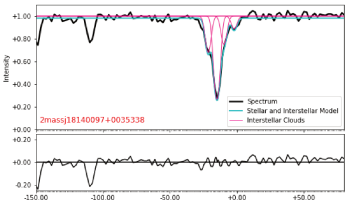
(b) 2massj18101306-0000277



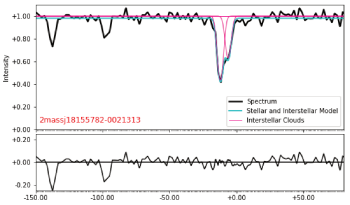
(c) 2massj18123623+0131238



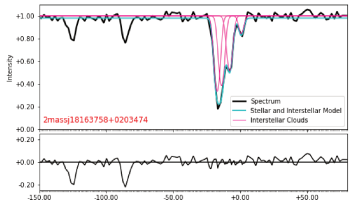
(d) 2massj18134458+0103348



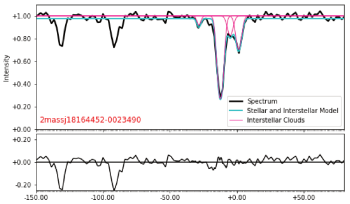
(e) 2massj18140097+0035338



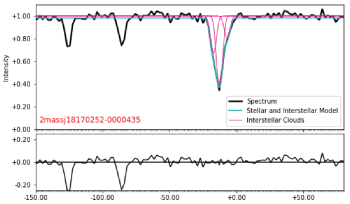
(f) 2massj18155782-0021313



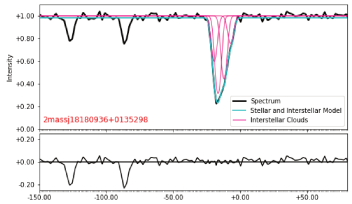
(g) 2massj18163758+0203474



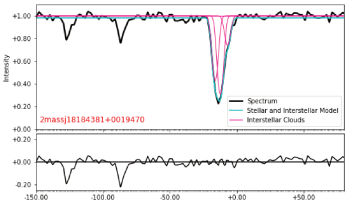
(h) 2massj18164452-0023490



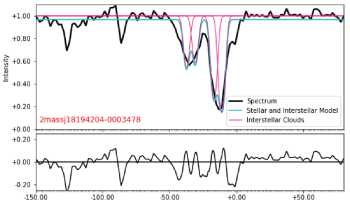
(i) 2massj18170252-0000435



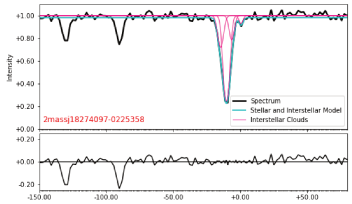
(j) 2massj18180936+0135298



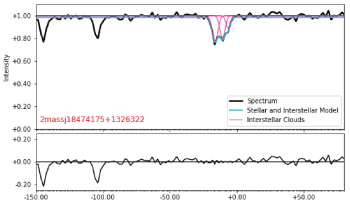
(k) 2massj18184381+0019470



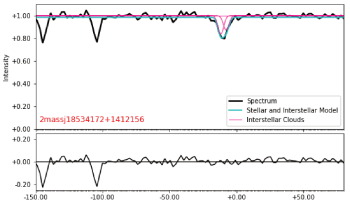
(l) 2massj18194204-0003478



(m) 2massj18274097-0225358

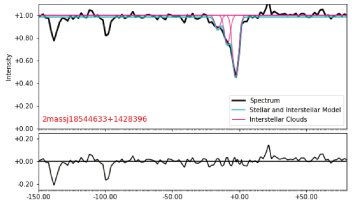


(n) 2massj18474175+1326322

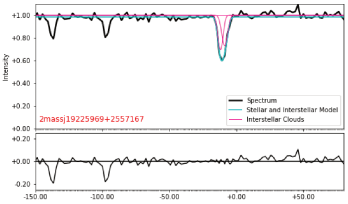


(o) 2massj18534172+1412156

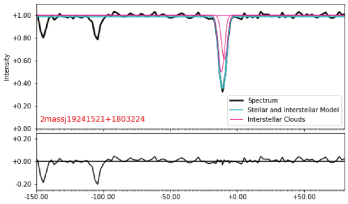
Figure C.5 – .



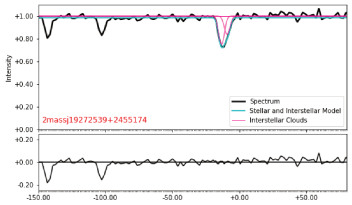
(a) 2massj18544633+1428396



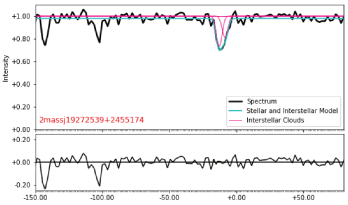
(b) 2massj19225969+2557167



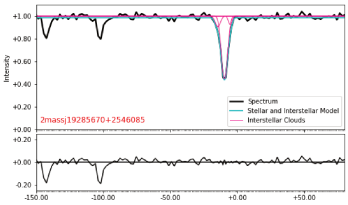
(c) 2massj19241521+1803224



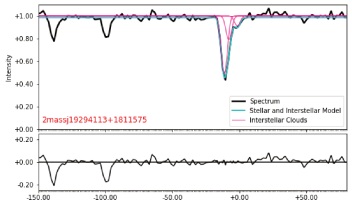
(d) 2massj19272539+2455174



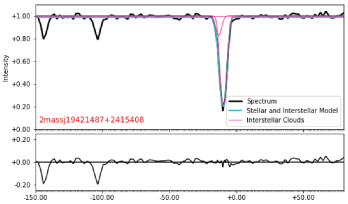
(e) 2massj19272539+2455174



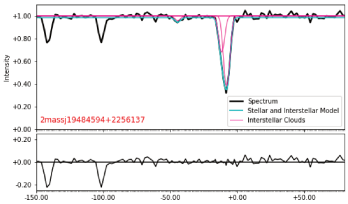
(f) 2massj19285670+2546085



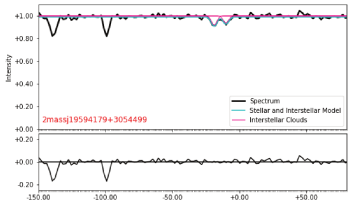
(g) 2massj19294113+1811575



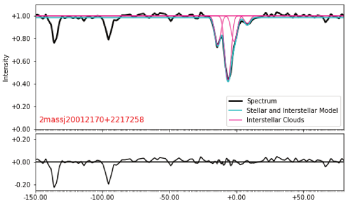
(h) 2massj19421487+2415408



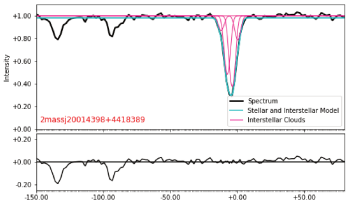
(i) 2massj19484594+2256137



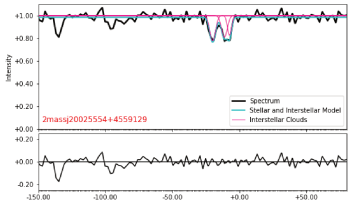
(j) 2massj19594179+3054499



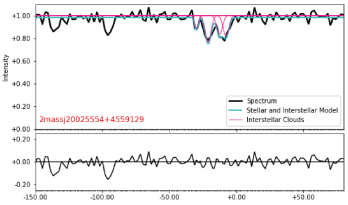
(k) 2massj20012170+2217258



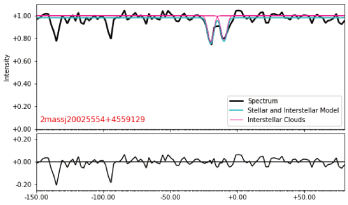
(l) 2massj20014398+4418389



(m) 2massj20025554+4559129

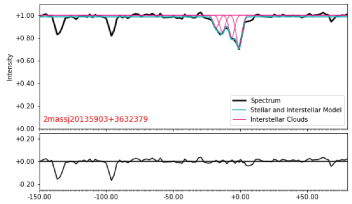


(n) 2massj20025554+4559129

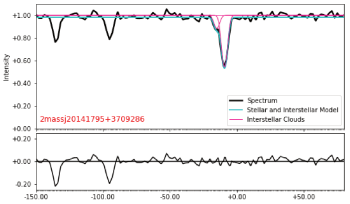


(o) 2massj20025554+4559129

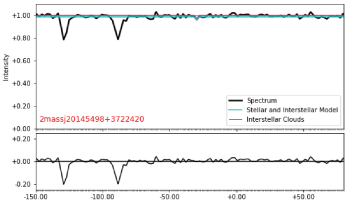
Figure C.6 – .



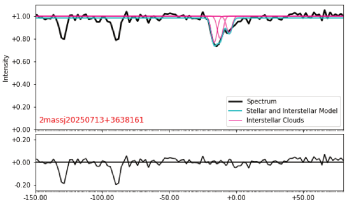
(a) 2massj20135903+3632379



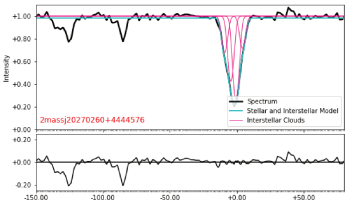
(b) 2massj20141795+3709286



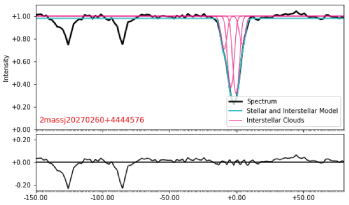
(c) 2massj20145498+3722420



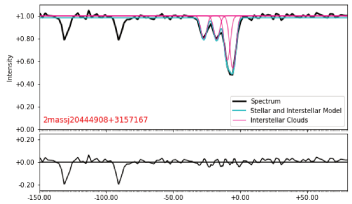
(d) 2massj20250713+3638161



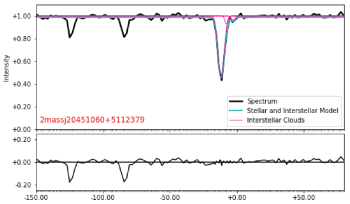
(e) 2massj20270260+4444576



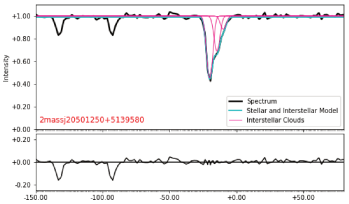
(f) 2massj20270260+4444576



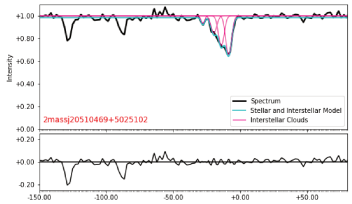
(g) 2massj20444908+3157167



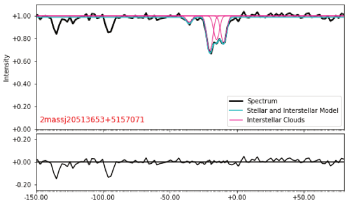
(h) 2massj20451060+5112379



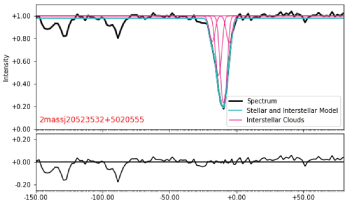
(i) 2massj20501250+5139580



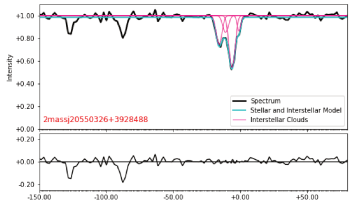
(j) 2massj20510469+5025102



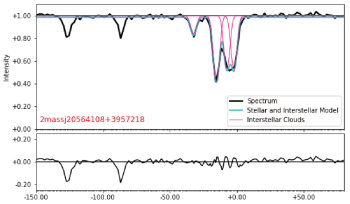
(k) 2massj20513653+5157071



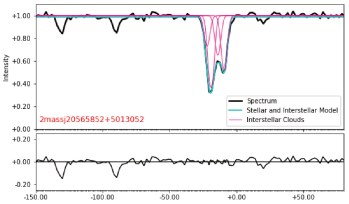
(l) 2massj20523532+5020555



(m) 2massj20550326+3928488

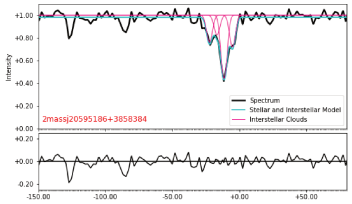


(n) 2massj20564108+3957218

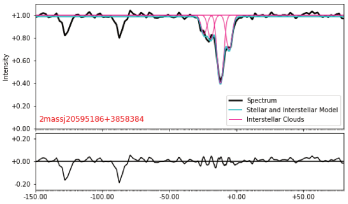


(o) 2massj20565852+5013052

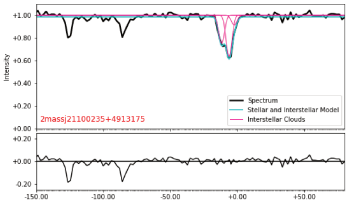
Figure C.7 – .



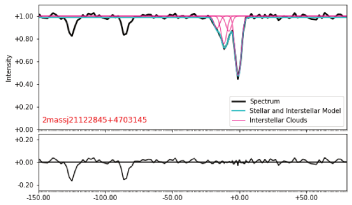
(a) 2massj20595186+3858384



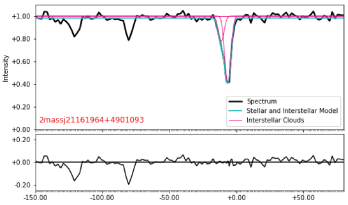
(b) 2massj20595186+3858384



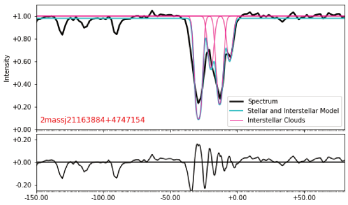
(c) 2massj21100235+4913175



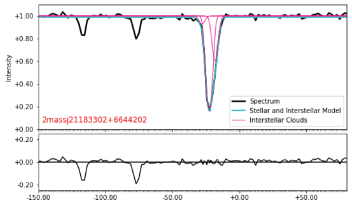
(d) 2massj21122845+4703145



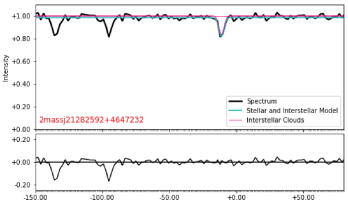
(e) 2massj21161964+4901093



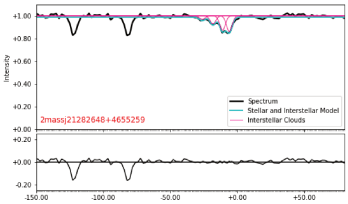
(f) 2massj21163884+4747154



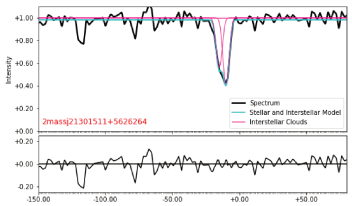
(g) 2massj21183302+6644202



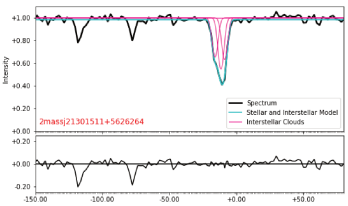
(h) 2massj21282592+4647232



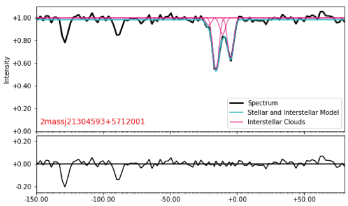
(i) 2massj21282648+4655259



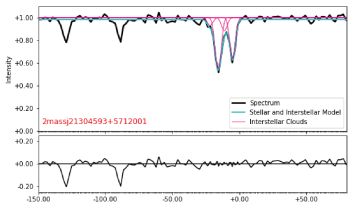
(j) 2massj21301511+5626264



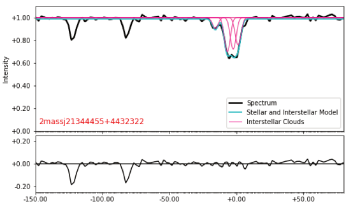
(k) 2massj21301511+5626264



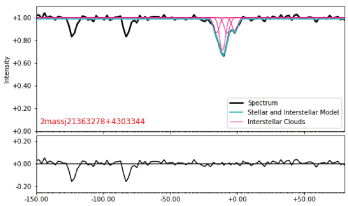
(l) 2massj21304593+5712001



(m) 2massj21304593+5712001

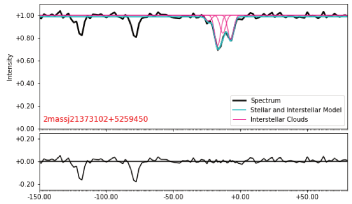


(n) 2massj21344455+4432322

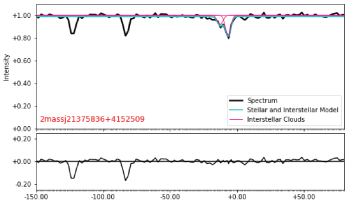


(o) 2massj21363278+4303344

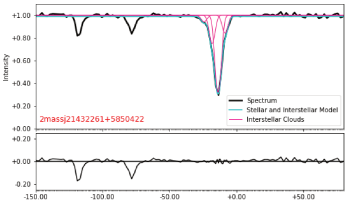
Figure C.8 – .



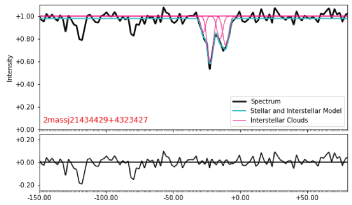
(a) 2massj21373102+5259450



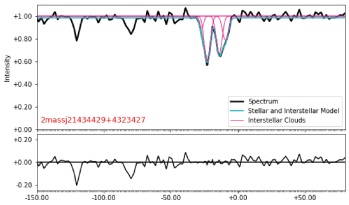
(b) 2massj21375836+4152509



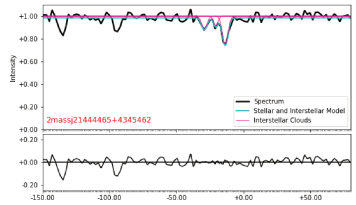
(c) 2massj21432261+5850422



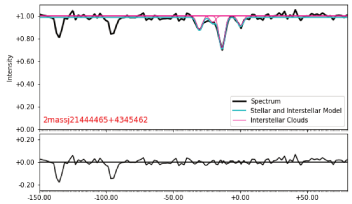
(d) 2massj21434429+4323427



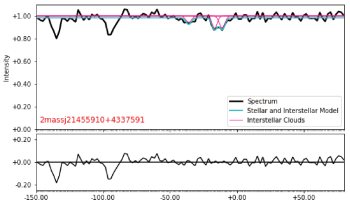
(e) 2massj21434429+4323427



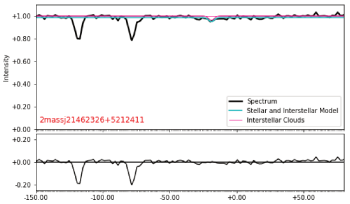
(f) 2massj21444465+4345462



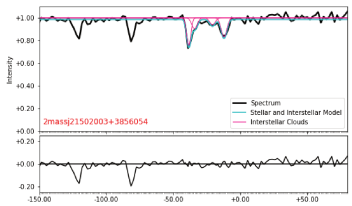
(g) 2massj21444465+4345462



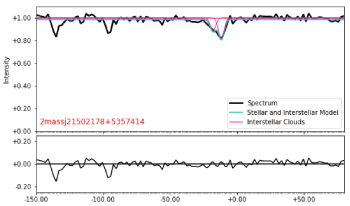
(h) 2massj21455910+4337591



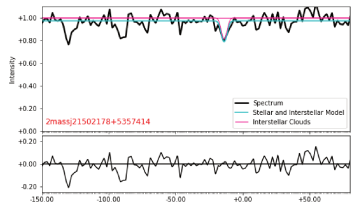
(i) 2massj21462326+5212411



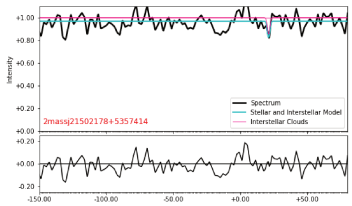
(j) 2massj21502003+3856054



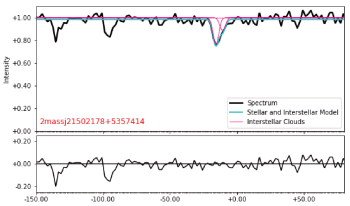
(k) 2massj21502178+5357414



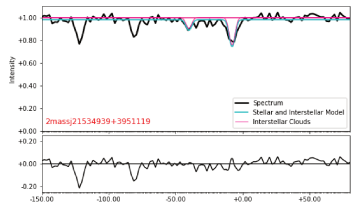
(l) 2massj21502178+5357414



(m) 2massj21502178+5357414

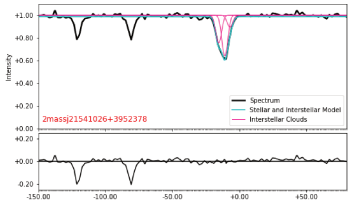


(n) 2massj21502178+5357414

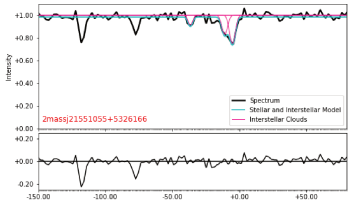


(o) 2massj21534939+3951119

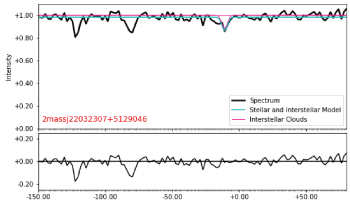
Figure C.9 – .



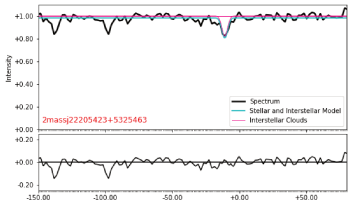
(a) 2massj21541026+3952378



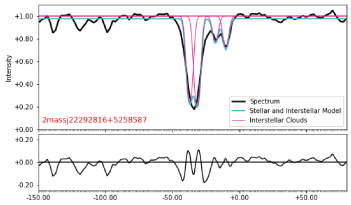
(b) 2massj21551055+5326166



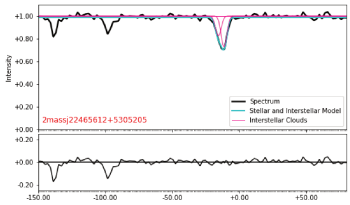
(c) 2massj22032307+5129046



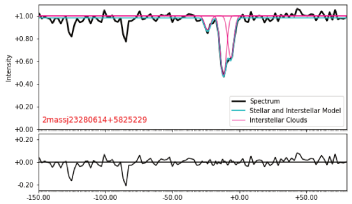
(d) 2massj22205423+5325463



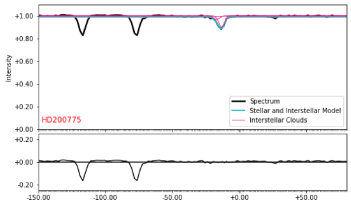
(e) 2massj22292816+5258587



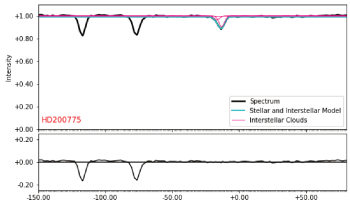
(f) 2massj22465612+5305205



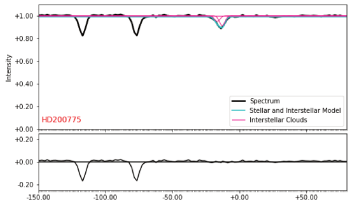
(g) 2massj23280614+5825229



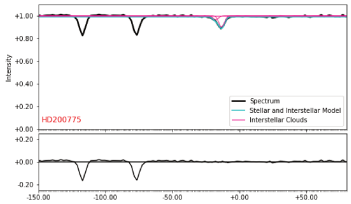
(h) HD 200775



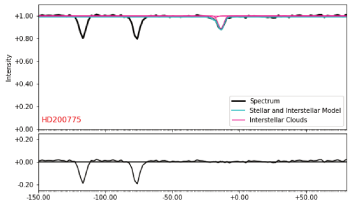
(i) HD 200775



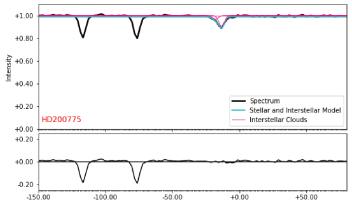
(j) HD 200775



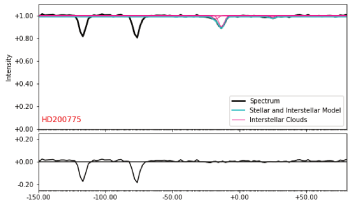
(k) HD 200775



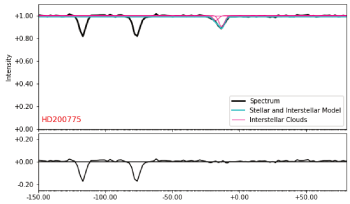
(l) HD 200775



(m) HD 200775

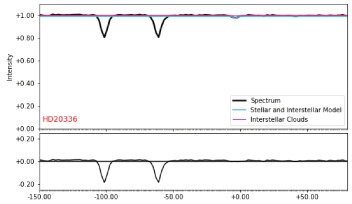


(n) HD 200775

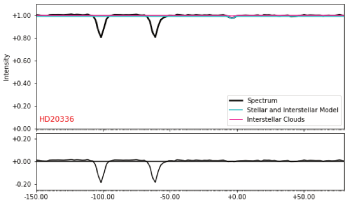


(o) HD 200775

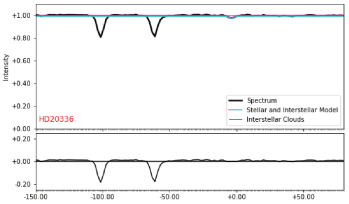
Figure C.10 – .



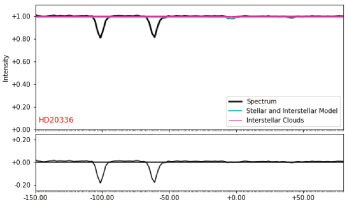
(a) HD 20336



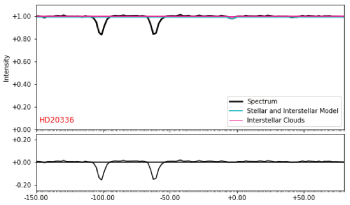
(b) HD 20336



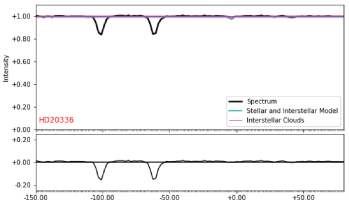
(c) HD 20336



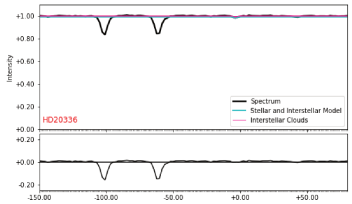
(d) HD 20336



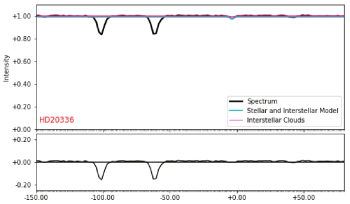
(e) HD 20336



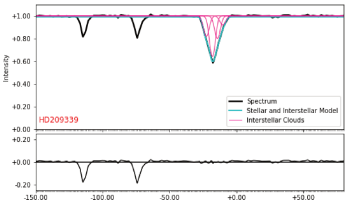
(f) HD 20336



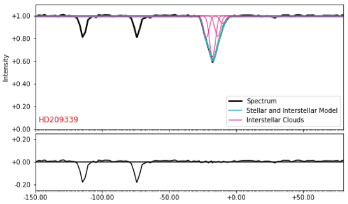
(g) HD 20336



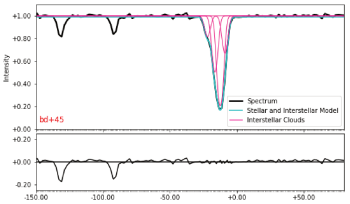
(h) HD 20336



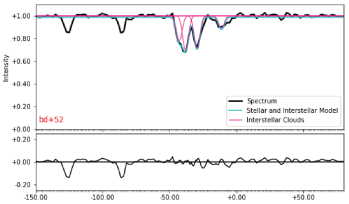
(i) HD 209339



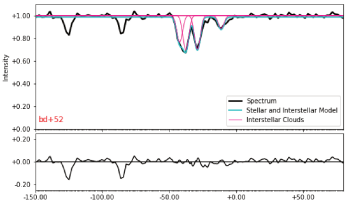
(j) HD 209339



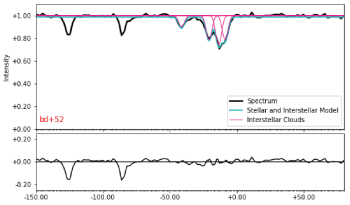
(k) BD+45 3341



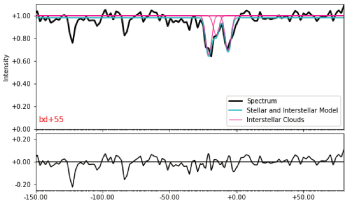
(l) BD+52 3122



(m) BD+52 3122

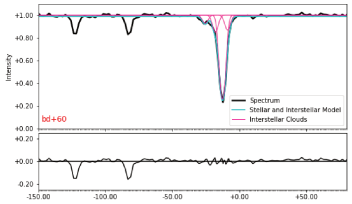


(n) BD+52 3135

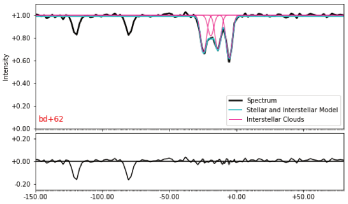


(o) BD+55 2770

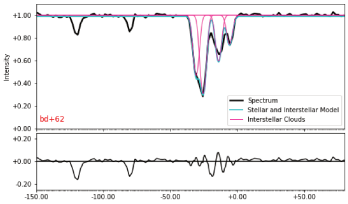
Figure C.11 – .



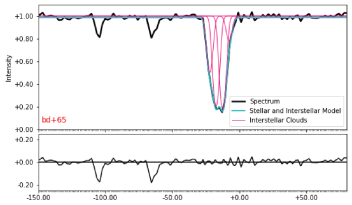
(a) BD+60 2380



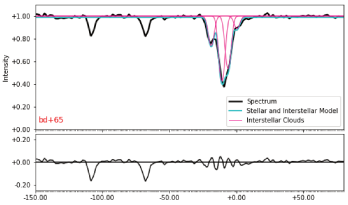
(b) BD+62 2142



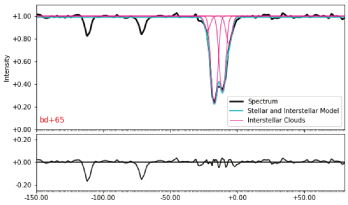
(c) BD+62 2154



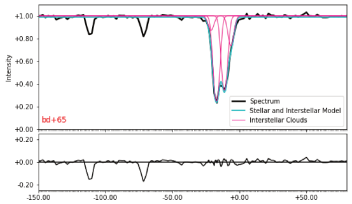
(d) BD+65 121



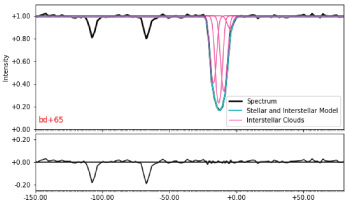
(e) BD+65 122



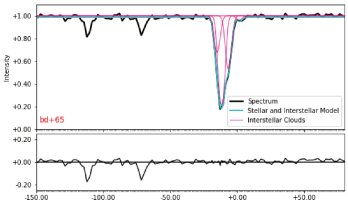
(f) BD+65 131



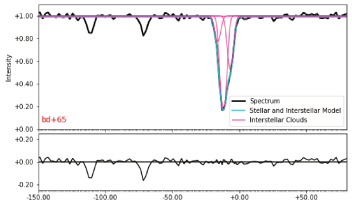
(g) BD+65 131



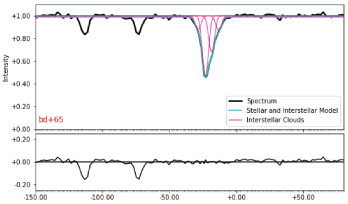
(h) BD+65 142



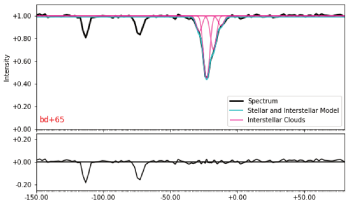
(i) BD+65 154



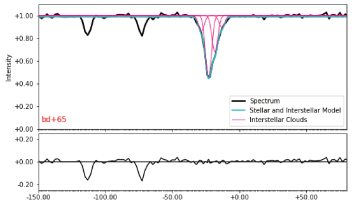
(j) BD+65 154



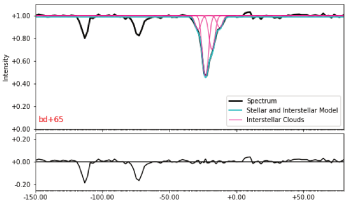
(k) BD+65 1637



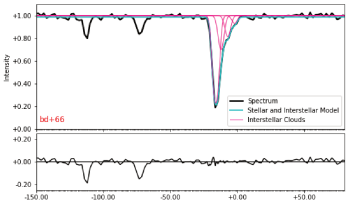
(l) BD+65 1637



(m) BD+65 1637

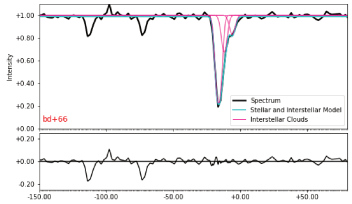


(n) BD+65 1637

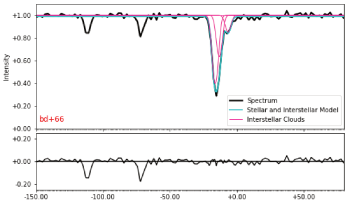


(o) BD+66 100

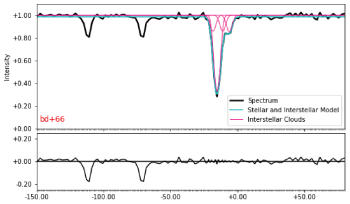
Figure C.12 – .



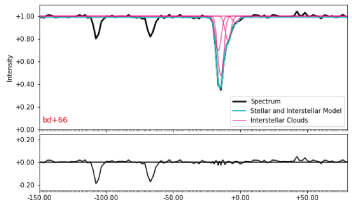
(a) BD+66 100



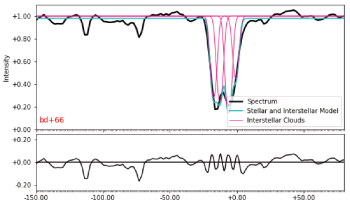
(b) BD+66 104



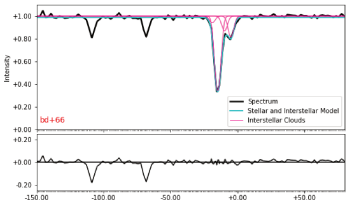
(c) BD+66 104



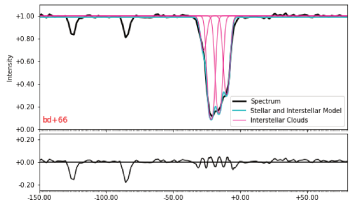
(d) BD+66 111



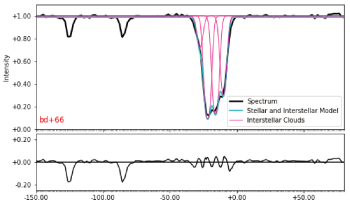
(e) BD+66 112



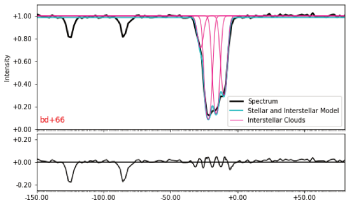
(f) BD+66 116



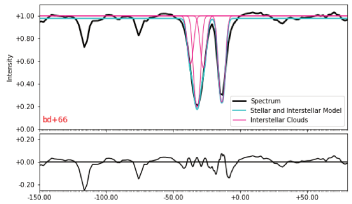
(g) BD+66 1673



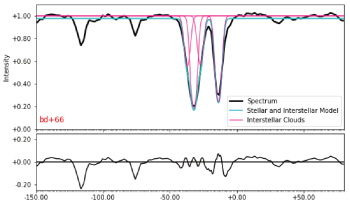
(h) BD+66 1673



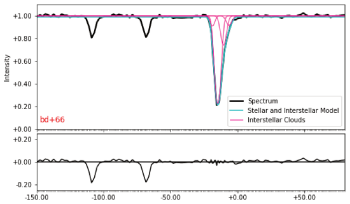
(i) BD+66 1673



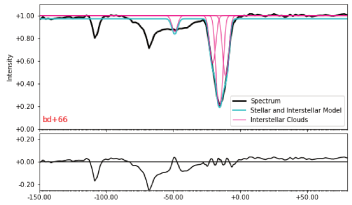
(j) BD+66 63



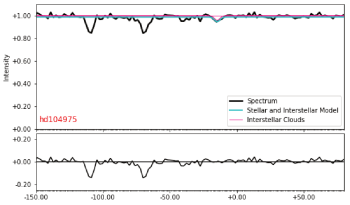
(k) BD+66 63



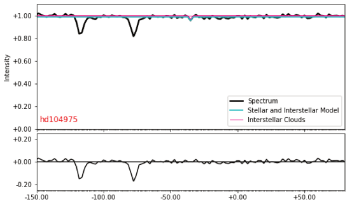
(l) BD+66 71



(m) BD+66 89

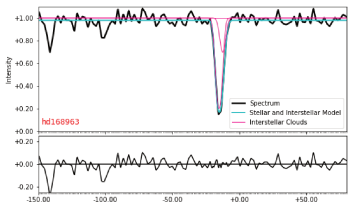


(n) HD 104975

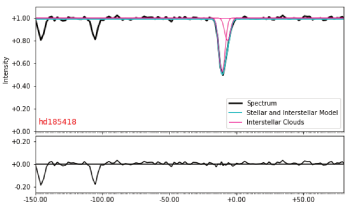


(o) HD 104975

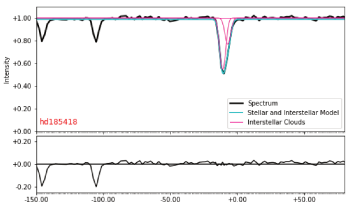
Figure C.13 – .



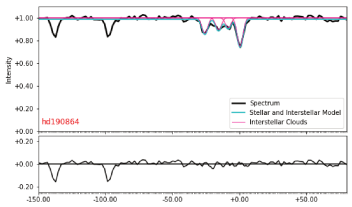
(a) HD 168963



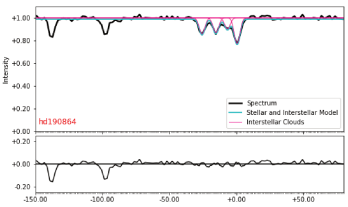
(b) HD 185418



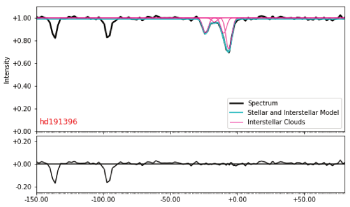
(c) HD 185418



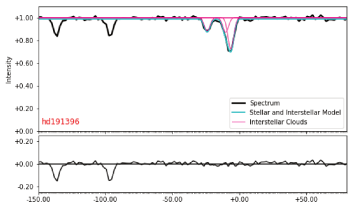
(d) HD 190864



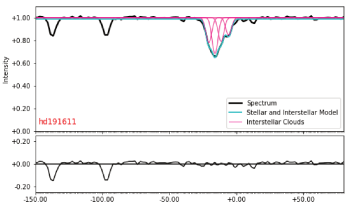
(e) HD 190864



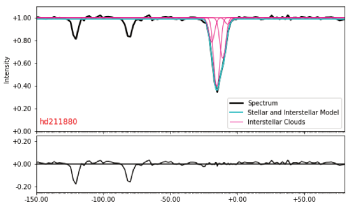
(f) HD 191396



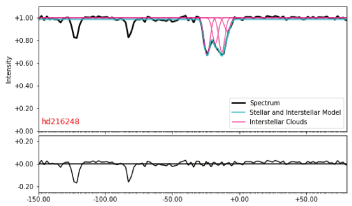
(g) HD 191396



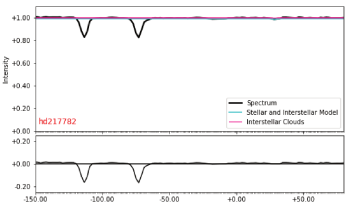
(h) HD 191611



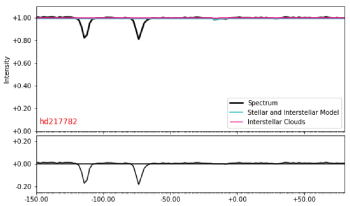
(i) HD 211880



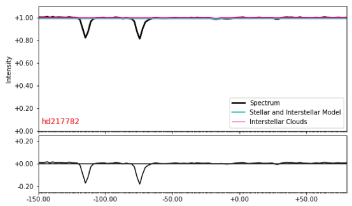
(j) HD 216248



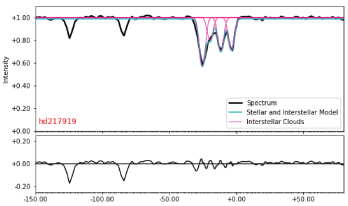
(k) HD 217782



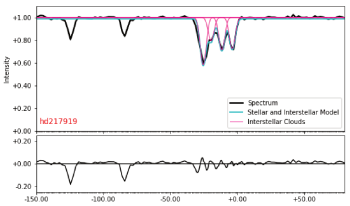
(l) HD 217782



(m) HD 217782

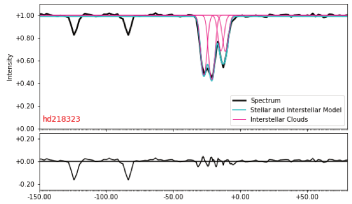


(n) HD 217919

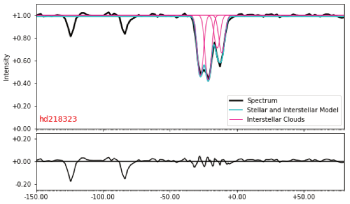


(o) HD 217919

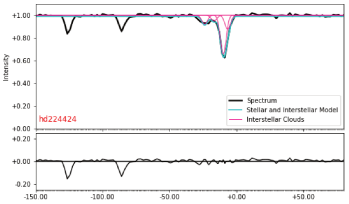
Figure C.14 – .



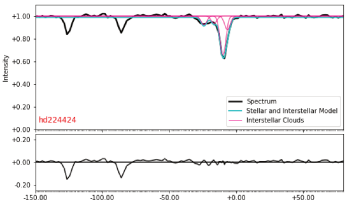
(a) HD 218323



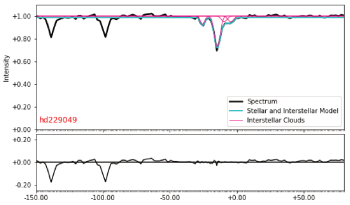
(b) HD 218323



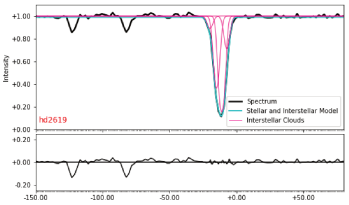
(c) HD 224424



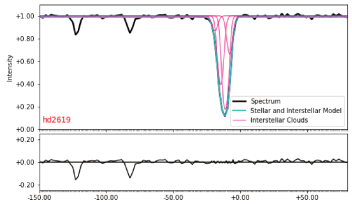
(d) HD 224424



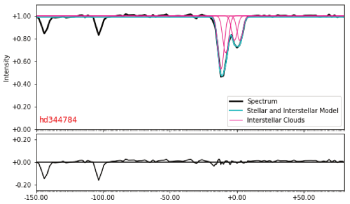
(e) HD 229049



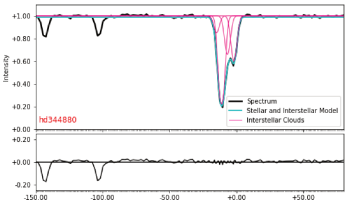
(f) HD 2619



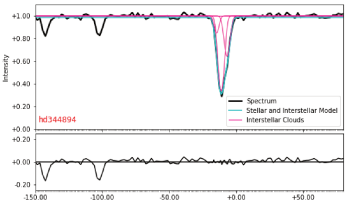
(g) HD 2619



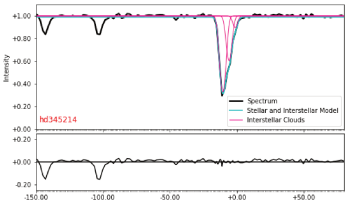
(h) HD 344784



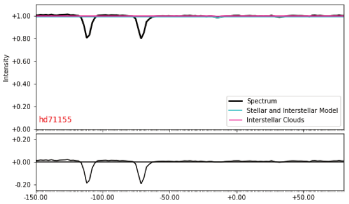
(i) HD 344880



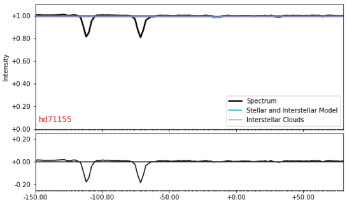
(j) HD 344894



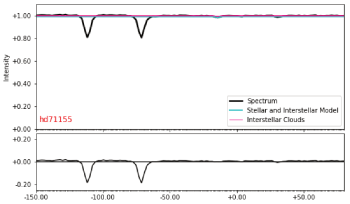
(k) HD 345214



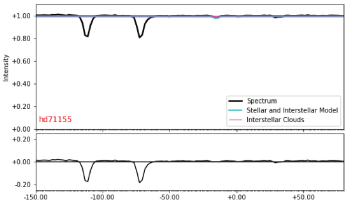
(l) HD 71155



(m) HD 71155

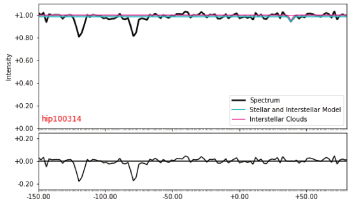


(n) HD 71155

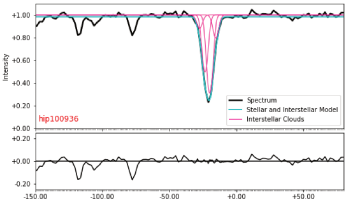


(o) HD 71155

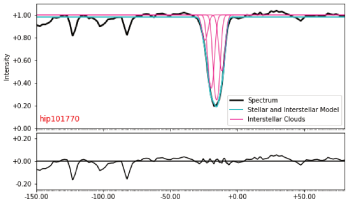
Figure C.15 – .



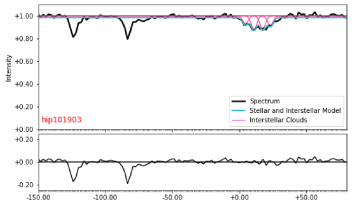
(a) hip100314



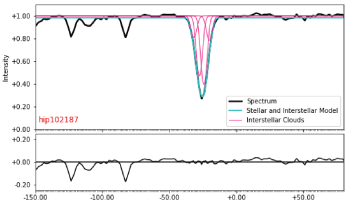
(b) hip100936



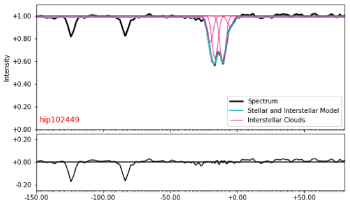
(c) hip101770



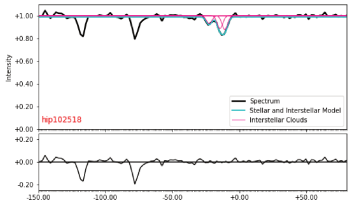
(d) hip101903



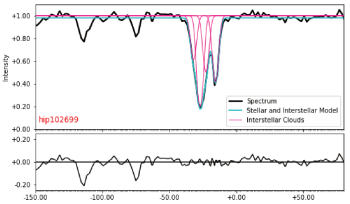
(e) hip102187



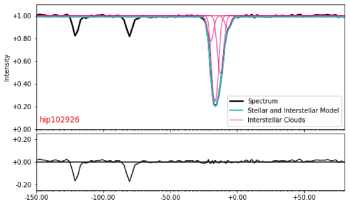
(f) hip102449



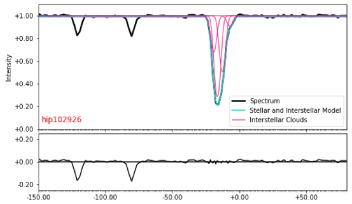
(g) hip102518



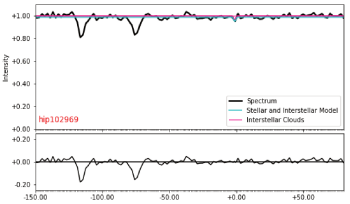
(h) hip102699



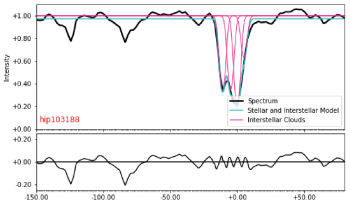
(i) hip102926



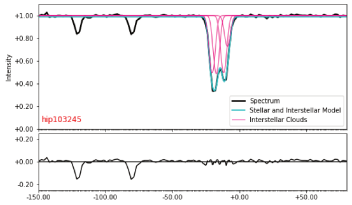
(j) hip102926



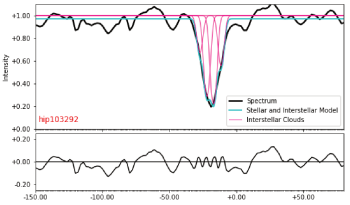
(k) hip102969



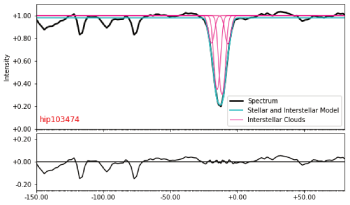
(l) hip103188



(m) hip103245

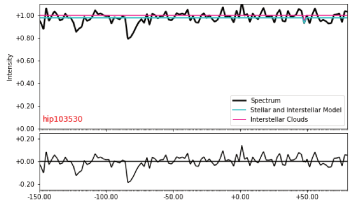


(n) hip103292

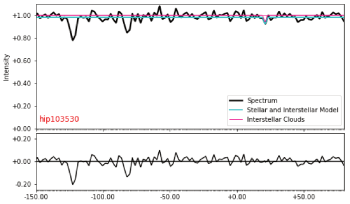


(o) hip103474

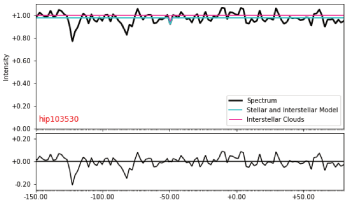
Figure C.16 – .



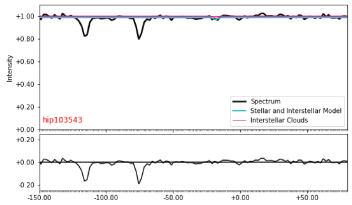
(a) hip103530



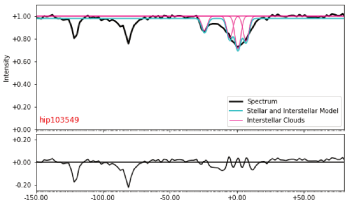
(b) hip103530



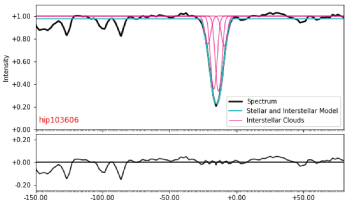
(c) hip103530



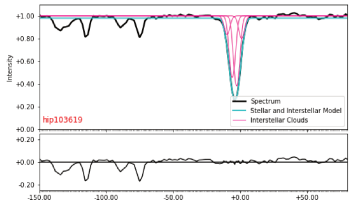
(d) hip103543



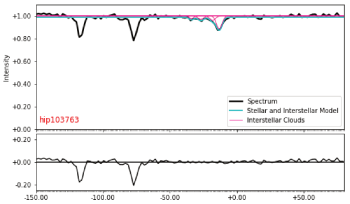
(e) hip103549



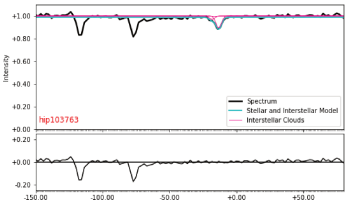
(f) hip103606



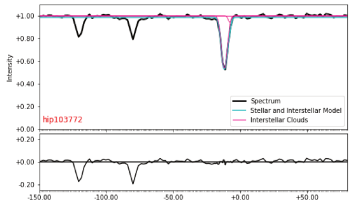
(g) hip103619



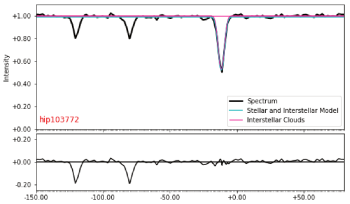
(h) hip103763



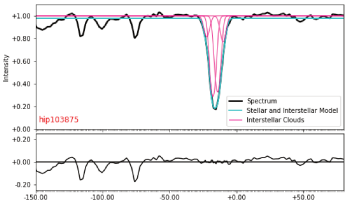
(i) hip103763



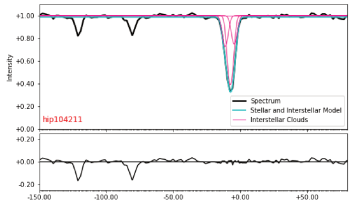
(j) hip103772



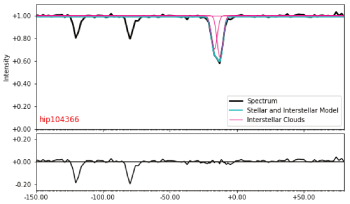
(k) hip103772



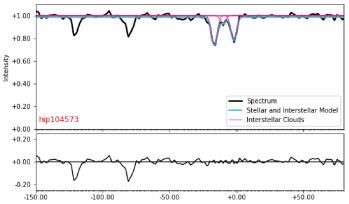
(l) hip103875



(m) hip104211

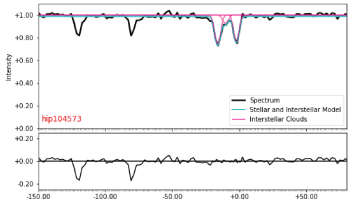


(n) hip104366

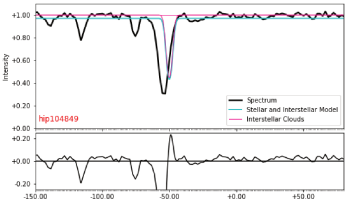


(o) hip104573

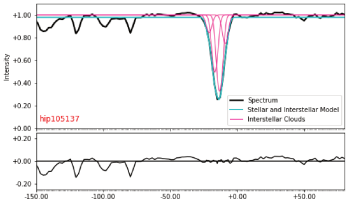
Figure C.17 – .



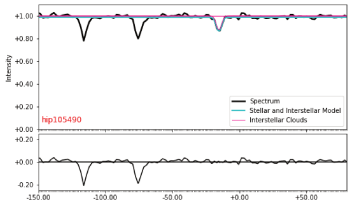
(a) hip104573



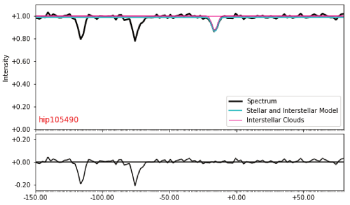
(b) hip104849



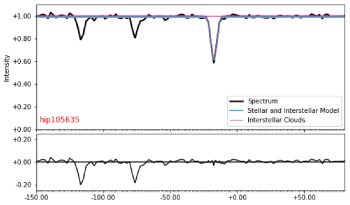
(c) hip105137



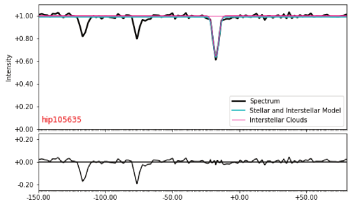
(d) hip105490



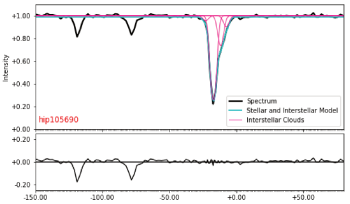
(e) hip105490



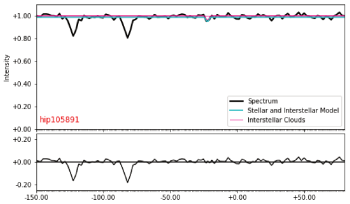
(f) hip105635



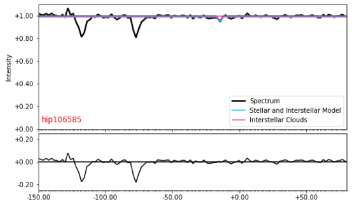
(g) hip105635



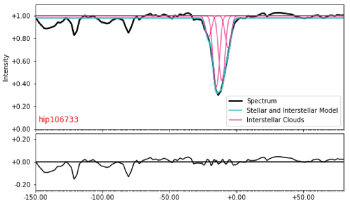
(h) hip105690



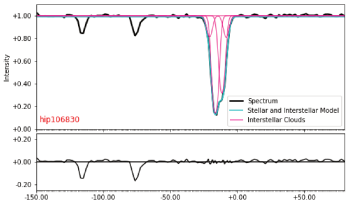
(i) hip105891



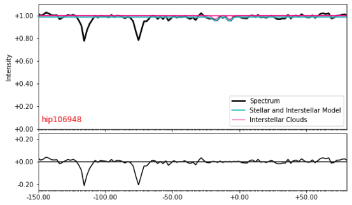
(j) hip106585



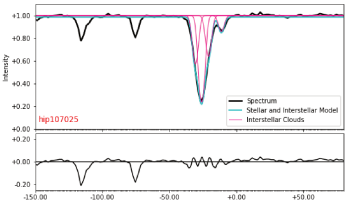
(k) hip106733



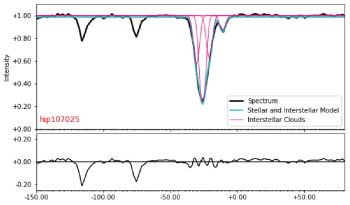
(l) hip106830



(m) hip106948

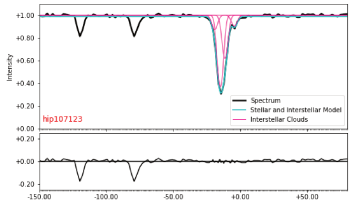


(n) hip107025

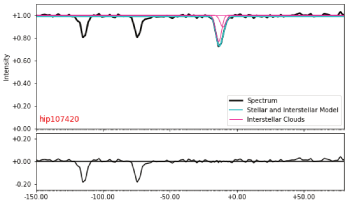


(o) hip107025

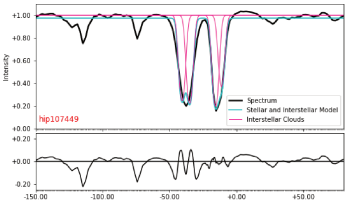
Figure C.18 – .



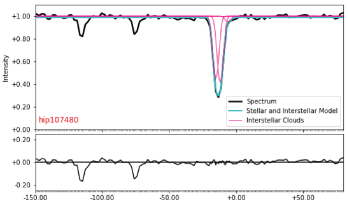
(a) hip107123



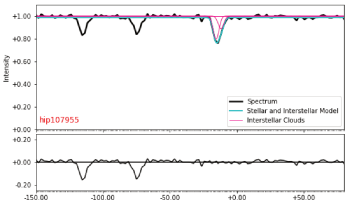
(b) hip107420



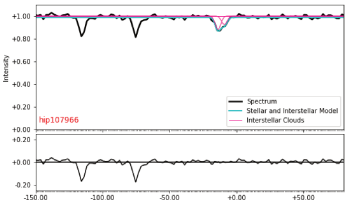
(c) hip107449



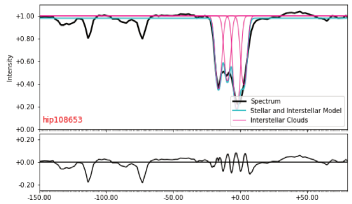
(d) hip107480



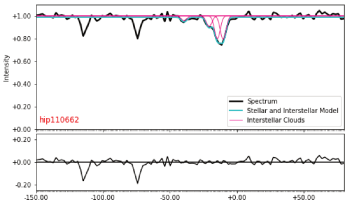
(e) hip107955



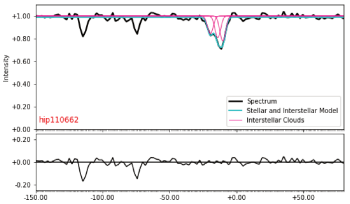
(f) hip107966



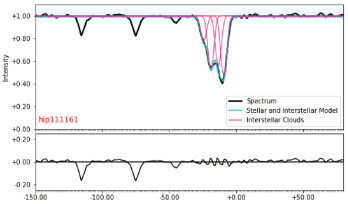
(g) hip108653



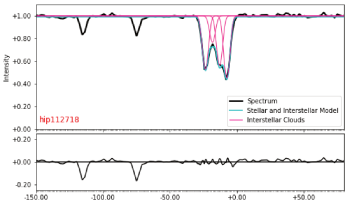
(h) hip110662



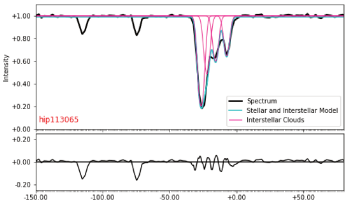
(i) hip110662



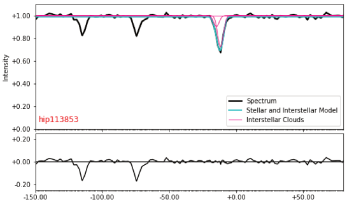
(j) hip111161



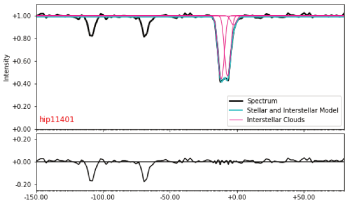
(k) hip112718



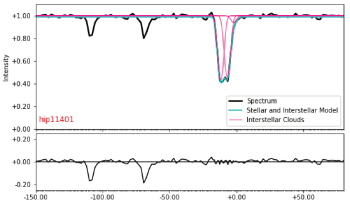
(l) hip113065



(m) hip113853

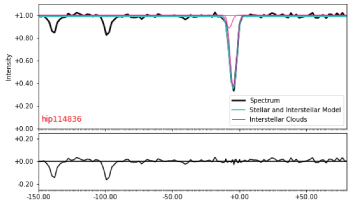


(n) hip11401

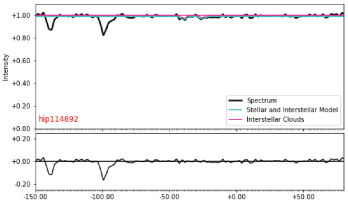


(o) hip11401

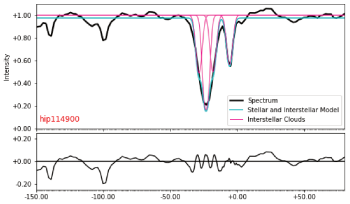
Figure C.19 – .



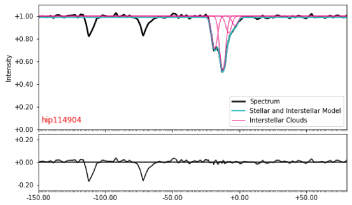
(a) hip114836



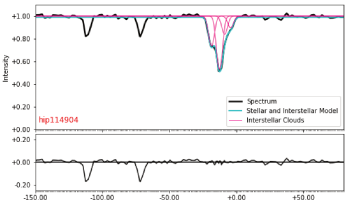
(b) hip114892



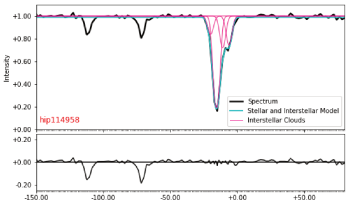
(c) hip114900



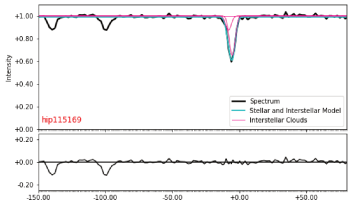
(d) hip114904



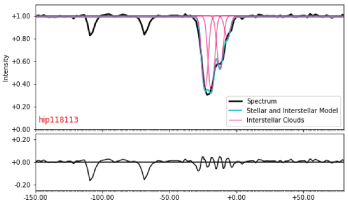
(e) hip114904



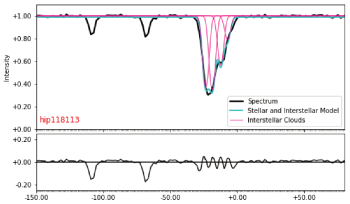
(f) hip114958



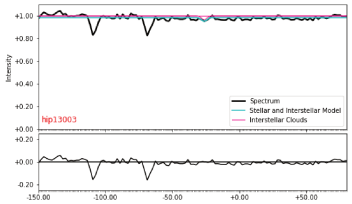
(g) hip115169



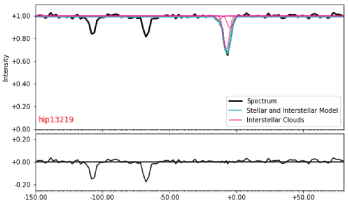
(h) hip118113



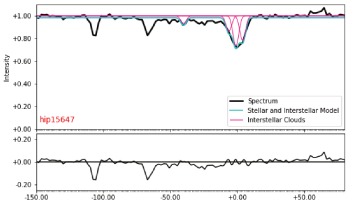
(i) hip118113



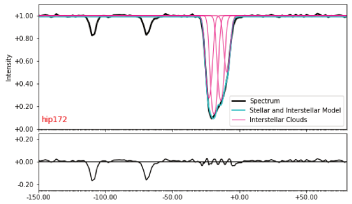
(j) hip13003



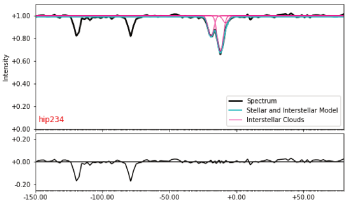
(k) hip13219



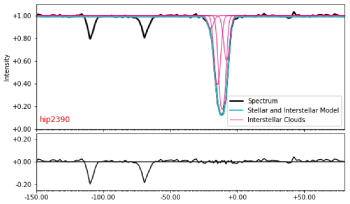
(l) hip15647



(m) hip172

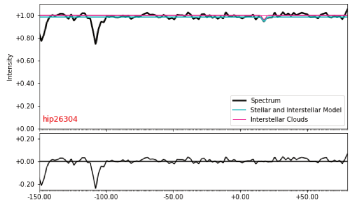


(n) hip234

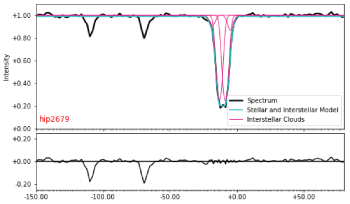


(o) hip2390

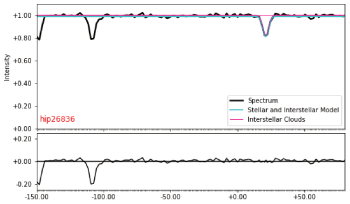
Figure C.20 – .



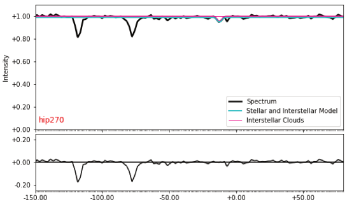
(a) hip26304



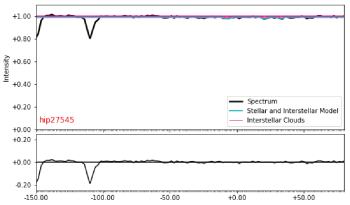
(b) hip2679



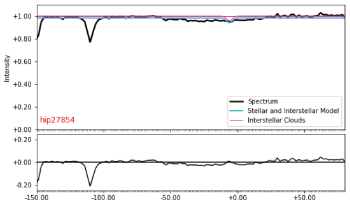
(c) hip26836



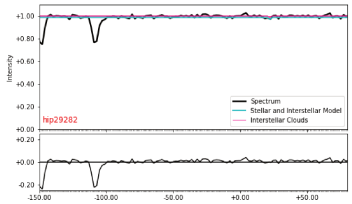
(d) hip270



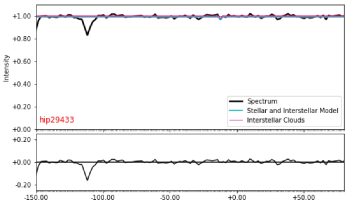
(e) hip27545



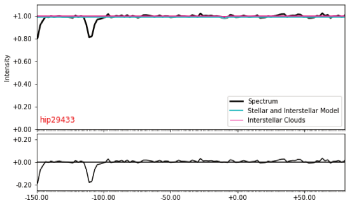
(f) hip27854



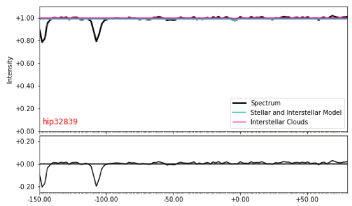
(g) hip29282



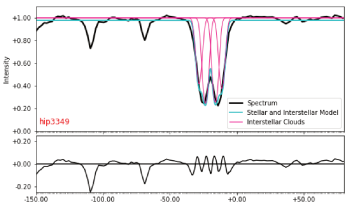
(h) hip29433



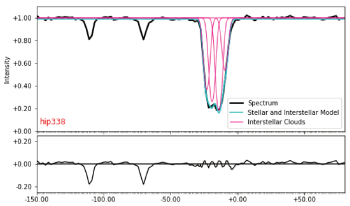
(i) hip29433



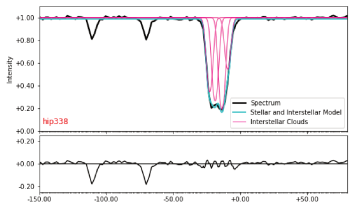
(j) hip32839



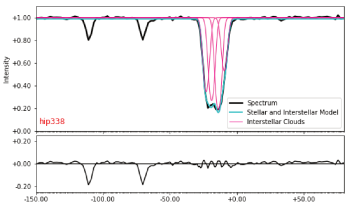
(k) hip3349



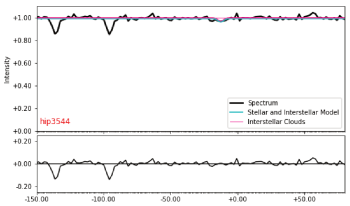
(l) hip338



(m) hip338

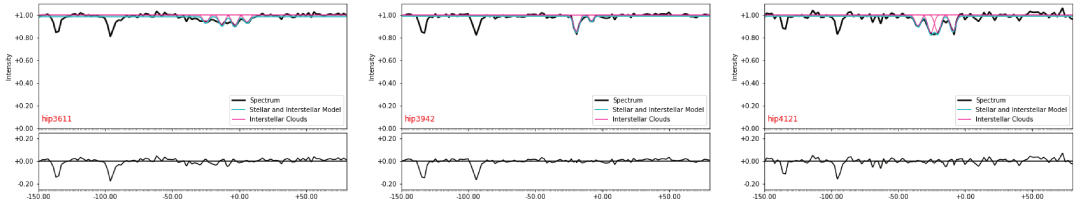


(n) hip338



(o) hip3544

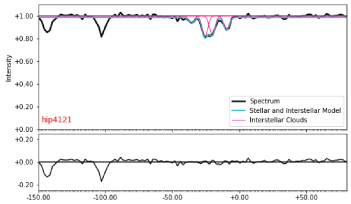
Figure C.21 – .



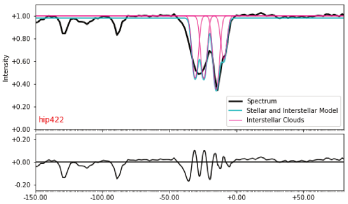
(a) hip3611

(b) hip3942

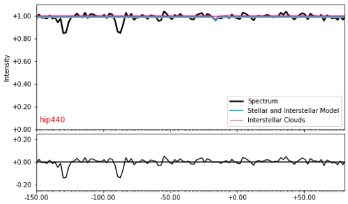
(c) hip4121



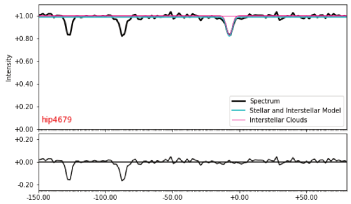
(d) hip4121



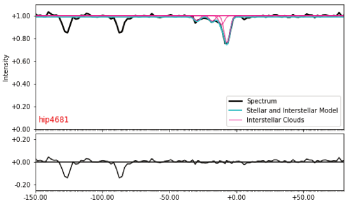
(e) hip422



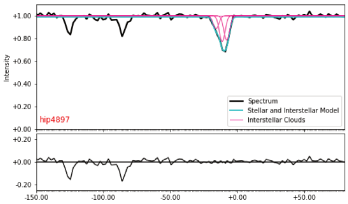
(f) hip440



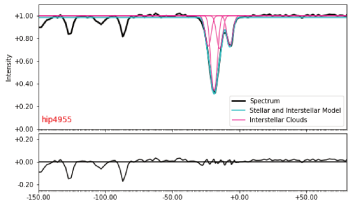
(g) hip4679



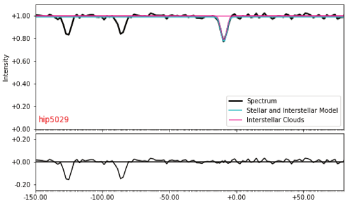
(h) hip4681



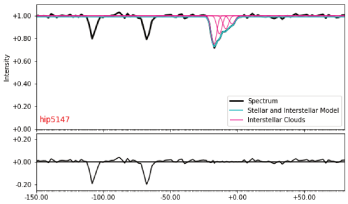
(i) hip4897



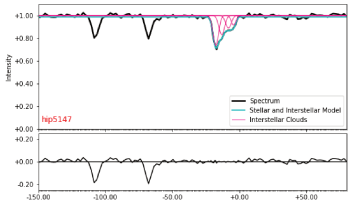
(j) hip4955



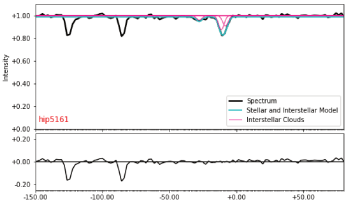
(k) hip5029



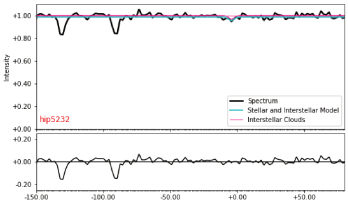
(l) hip5147



(m) hip5147

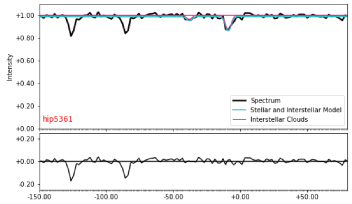


(n) hip5161

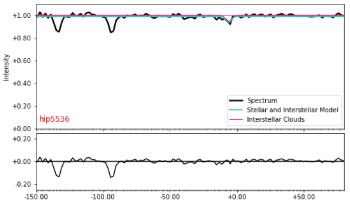


(o) hip5232

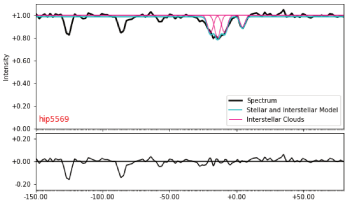
Figure C.22 – .



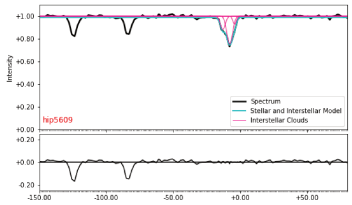
(a) hip5361



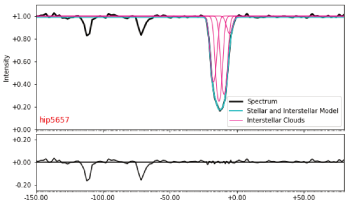
(b) hip5536



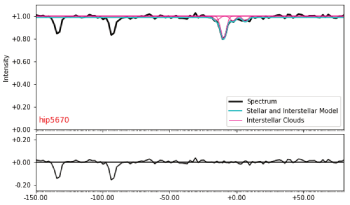
(c) hip5569



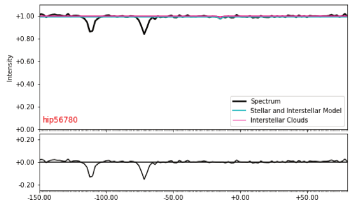
(d) hip5609



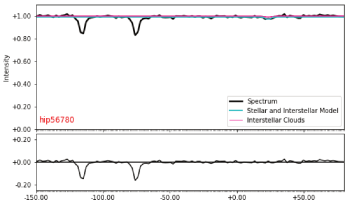
(e) hip5657



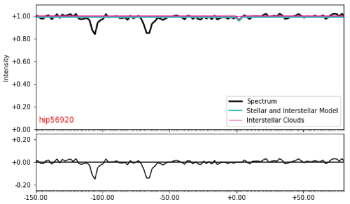
(f) hip5670



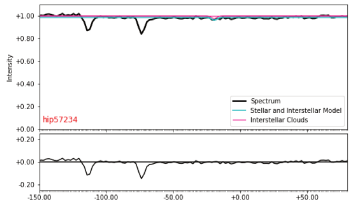
(g) hip56780



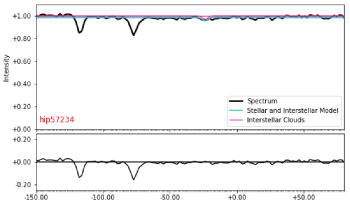
(h) hip56780



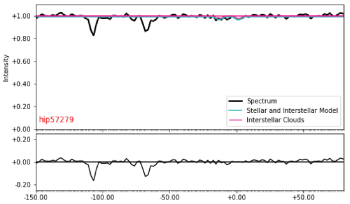
(i) hip56920



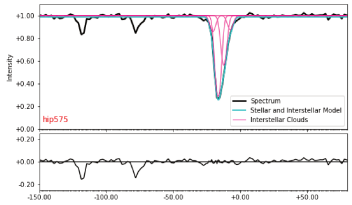
(j) hip57234



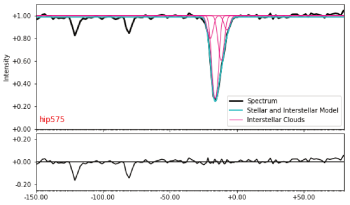
(k) hip57234



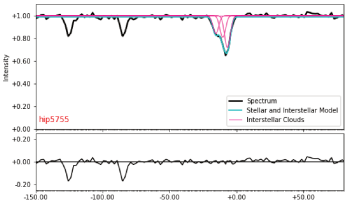
(l) hip57279



(m) hip575

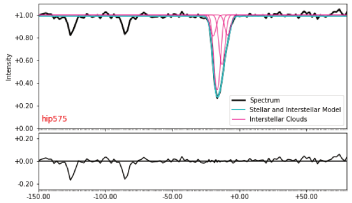


(n) hip575

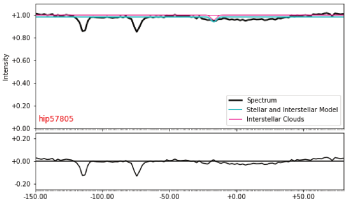


(o) hip5755

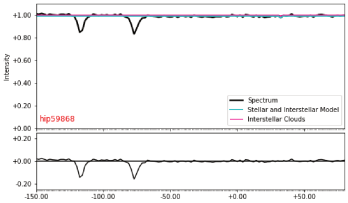
Figure C.23 – .



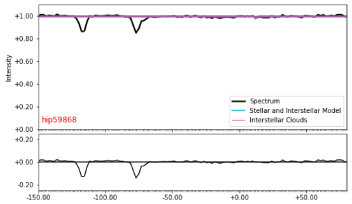
(a) hip575



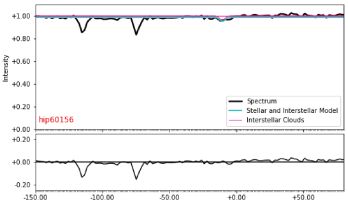
(b) hip57805



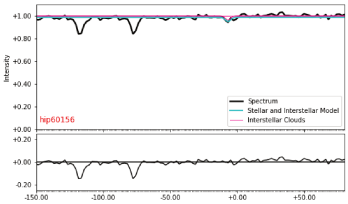
(c) hip59868



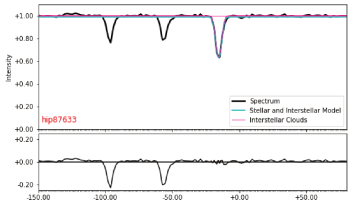
(d) hip59868



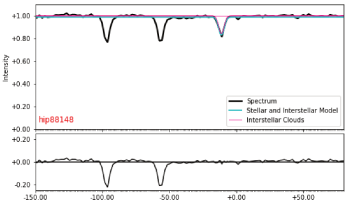
(e) hip60156



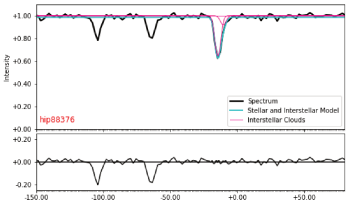
(f) hip60156



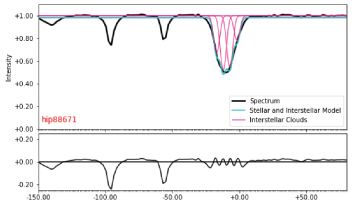
(g) hip87633



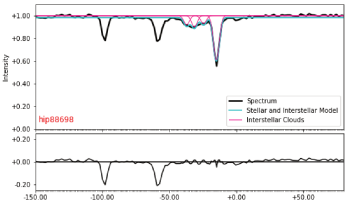
(h) hip88148



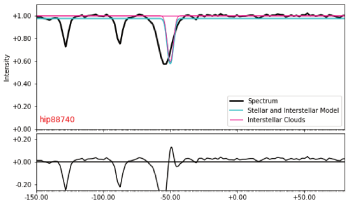
(i) hip88376



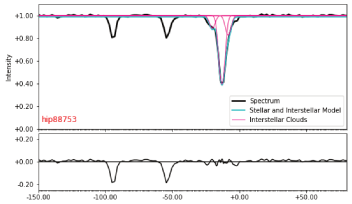
(j) hip88671



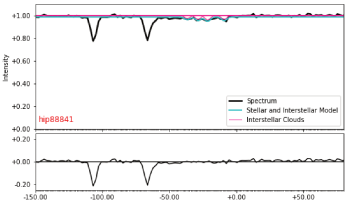
(k) hip88698



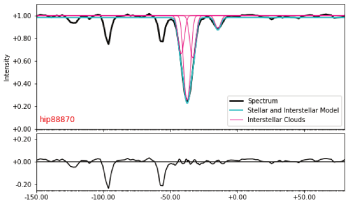
(l) hip88740



(m) hip88753

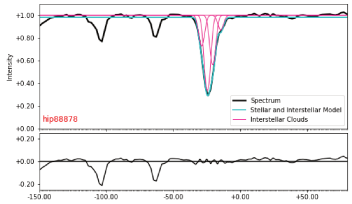


(n) hip88841

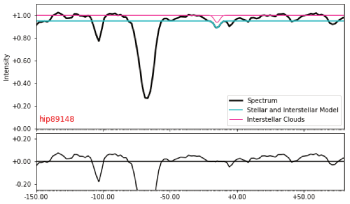


(o) hip88870

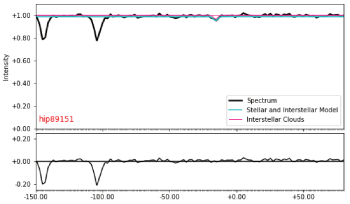
Figure C.24 – .



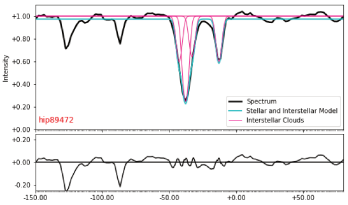
(a) hip88878



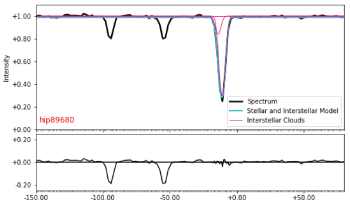
(b) hip89148



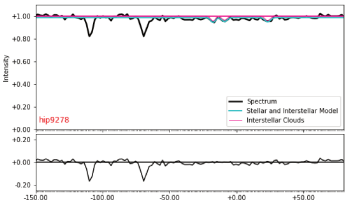
(c) hip89151



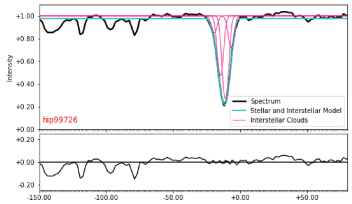
(d) hip89472



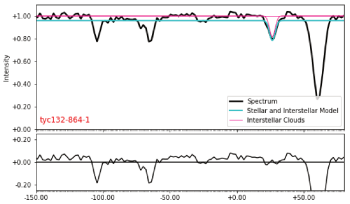
(e) hip89680



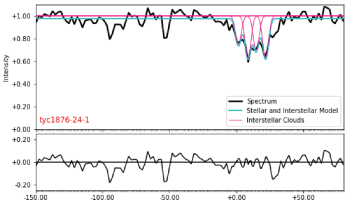
(f) hip9278



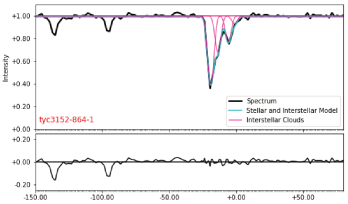
(g) hip99726



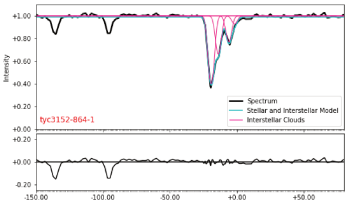
(h) TYC 132-864-1



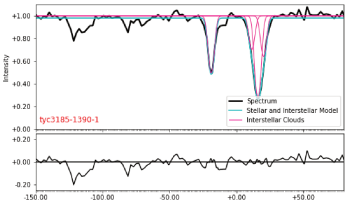
(i) TYC 1876-24-1



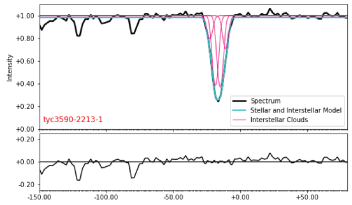
(j) TYC 3152-864-1



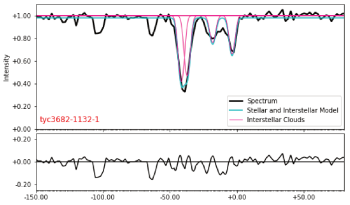
(k) TYC 3152-864-1



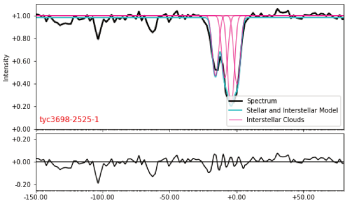
(l) TYC 3185-1390-1



(m) TYC 3590-2213-1

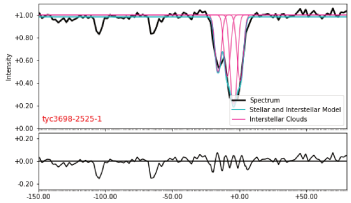


(n) TYC 3682-1132-1

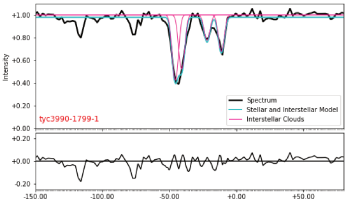


(o) TYC 3698-2525-1

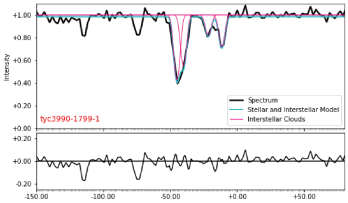
Figure C.25 – .



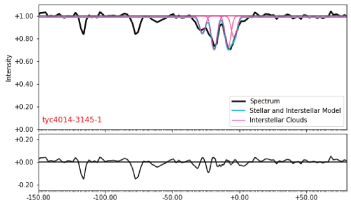
(a) TYC 3698-2525-1



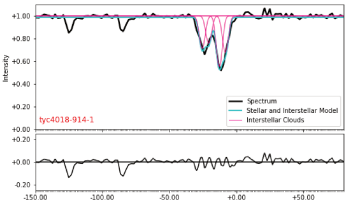
(b) TYC 3990-1799-1



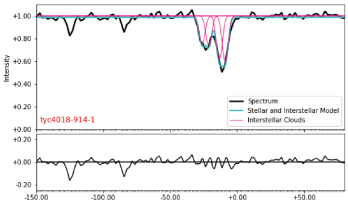
(c) TYC 3990-1799-1



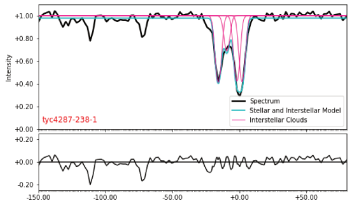
(d) TYC 4014-3145-1



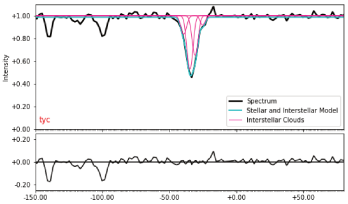
(e) TYC 4018-914-1



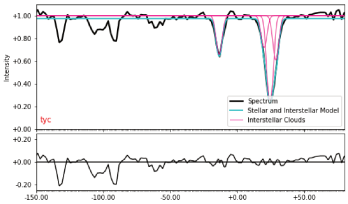
(f) TYC 4018-914-1



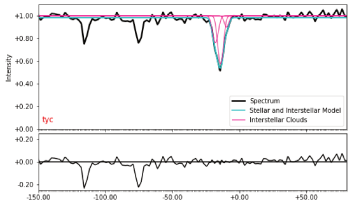
(g) TYC 4287-238-1



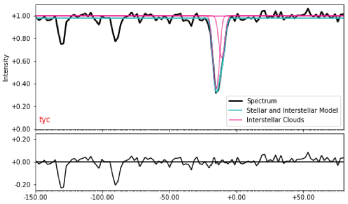
(h) TYC 3045-332-1



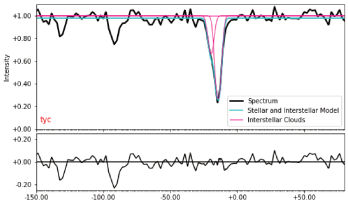
(i) TYC 430-1922-1



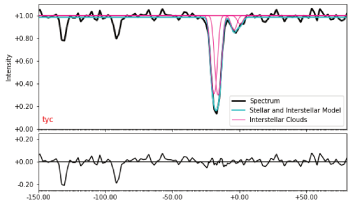
(j) TYC 431-1605-1



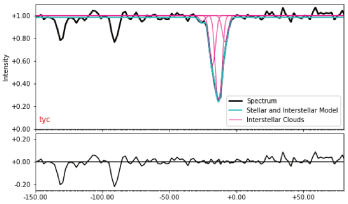
(k) TYC 431-394-1



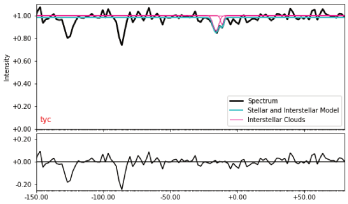
(l) TYC 431-812-1



(m) TYC 431-835-1



(n) TYC 432-1126-1



(o) TYC 434-4617-1

Figure C.26 – .

Bibliography

Alam, S., Albareti, F. D., Allende Prieto, C., Anders, F., Anderson, S. F., Anderton, T., Andrews, B. H., Armengaud, E., Aubourg, É., Bailey, S., and et al. (2015). The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. *ApJS*, 219:12.

Altobelli, N., Postberg, F., Fiege, K., Trieloff, M., Kimura, H., Sterken, V. J., Hsu, H.-W., Hillier, J., Khawaja, N., Moragas-Klostermeyer, G., Blum, J., Burton, M., Srama, R., Kempf, S., and Gruen, E. (2016). Flux and composition of interstellar dust at Saturn from Cassinis Cosmic Dust Analyzer. *Science*, 352:312–318.

Arenou, F., Luri, X., Babusiaux, C., Fabricius, C., Helmi, A., Muraveva, T., Robin, A. C., Spoto, F., Vallenari, A., Antoja, T., Cantat-Gaudin, T., Jordi, C., Leclerc, N., Reylé, C., Romero-Gómez, M., Shih, I., Soria, S., Barache, C., Bossini, D., Bragaglia, A., Breddels, M. A., Fabrizio, M., Lambert, S., Marrese, P. M., Massari, D., Moitinho, A., Robichon, N., Ruiz-Dern, L., Sordo, R., Veljanoski, J., Di Matteo, P., Eyer, L., Jasniewicz, G., Pancino, E., Soubiran, C., Spagna, A., Tanga, P., Turon, C., and Zurbach, C. (2018). Gaia Data Release 2: Catalogue validation. *ArXiv e-prints*.

Arenou, F., Luri, X., Babusiaux, C., Fabricius, C., Helmi, A., Robin, A. C., Vallenari, A., Blanco-Cuaresma, S., Cantat-Gaudin, T., Findeisen, K., Reylé, C., Ruiz-Dern, L., Sordo, R., Turon, C., Walton, N. A., Shih, I.-C., Antiche, E., Barache, C., Barros, M., Breddels, M., Carrasco, J. M., Costigan, G., Diakité, S., Eyer, L., Figueras, F., Galluccio, L., Heu, J., Jordi, C., Krone-Martins, A., Lallement, R., Lambert, S., Leclerc, N., Marrese, P. M., Moitinho, A., Mor, R., Romero-Gómez, M., Sartoretti, P., Soria, S., Soubiran, C., Souchay, J., Veljanoski, J., Ziaeepour, H., Giuffrida, G., Pancino, E., and Bragaglia, A. (2017a). Gaia Data Release 1. Catalogue validation. *Astronomy&Astrophysics*, 599:A50.

Arenou, F., Luri, X., Babusiaux, C., Fabricius, C., Helmi, A., Robin, A. C., Vallenari, A., Blanco-Cuaresma, S., Cantat-Gaudin, T., Findeisen, K., Reylé, C., Ruiz-Dern, L., Sordo, R., Turon, C., Walton, N. A., Shih, I.-C., Antiche, E., Barache, C., Barros, M., Breddels, M., Carrasco, J. M., Costigan, G., Diakité, S., Eyer, L., Figueras, F., Galluccio, L., Heu, J., Jordi, C., Krone-Martins, A., Lallement, R., Lambert, S., Leclerc, N., Marrese, P. M., Moitinho, A., Mor, R., Romero-Gómez, M., Sartoretti, P., Soria, S., Soubiran, C., Souchay, J., Veljanoski, J., Ziaeepour, H., Giuffrida, G., Pancino, E., and Bragaglia, A. (2017b). Gaia Data Release 1. Catalogue validation. *Astronomy&Astrophysics*, 599:A50.

Bailer-Jones, C. A. L. (2015). Estimating Distances from Parallaxes. *Publication of the Astronomical Society of the Pacific*, 127:994.

Berruto, G., Madan, I., Murooka, Y., Vanacore, G. M., Pomarico, E., Rajeswari, J., Lamb, R., Huang, P., Kruchkov, A. J., Togawa, Y., LaGrange, T., McGrouther, D., Rønnow, H. M., and Carbone, F. (2018). Laser-Induced Skyrmion Writing and Erasing in an Ultrafast Cryo-Lorentz Transmission Electron Microscope. *Physical Review Letters*, 120(11):117201.

Blandford, R. and Eichler, D. (1987). Particle acceleration at astrophysical shocks: A theory of cosmic ray origin. *Physics Reports*, 154:1–75.

Breitschwerdt, D., Feige, J., Schulreich, M. M., Avillez, M. A. D., Dettbarn, C., and Fuchs, B. (2016). The locations of recent supernovae near the Sun from modelling ^{60}Fe transport. *Nature*, 532:73–76.

Cami, J., Sonnentrucker, P., Ehrenfreund, P., and Foing, B. H. (1997). Diffuse Interstellar Bands in single clouds: new families and constraints on the carriers. *Astronomy&Astrophysics*, 326:822–830.

Campbell, E. K., Holz, M., Gerlich, D., and Maier, J. P. (2015). Laboratory confirmation of C_{60}^+ as the carrier of two diffuse interstellar bands. *Nature*, 523:322–323.

Capitanio, L., Lallement, R., Vergely, J. L., Elyajouri, M., and Monreal-Ibero, A. (2017). Three-dimensional mapping of the local interstellar medium with composite data. *Astronomy&Astrophysics*, 606:A65.

Cardelli, J. A., Clayton, G. C., and Mathis, J. S. (1989). The relationship between infrared, optical, and ultraviolet extinction. *Astrophysical Journal*, 345:245–256.

Carrasco, J. M., Evans, D. W., Montegriffo, P., Jordi, C., van Leeuwen, F., Riello, M., Voss, H., De Angeli, F., Busso, G., Fabricius, C., Cacciari, C., Weiler, M., Pancino, E., Brown, A. G. A., Holland, G., Burgess, P., Osborne, P., Altavilla, G., Gebran, M., Ragaini, S., Galleti, S., Cocozza, G., Marinoni, S., Bellazzini, M., Bragaglia, A., Federici, L., and Balaguer-Núñez, L. (2016). Gaia Data Release 1. Principles of the photometric calibration of the G band. *Astronomy&Astrophysics*, 595:A7.

Chambers, K. C., Magnier, E. A., Metcalfe, N., Flewelling, H. A., Huber, M. E., Waters, C. Z., Denneau, L., Draper, P. W., Farrow, D., Finkbeiner, D. P., Holmberg, C., Koppenhoefer, J., Price, P. A., Saglia, R. P., Schlafly, E. F., Smartt, S. J., Sweeney, W., Wainscoat, R. J., Burgett, W. S., Grav, T., Heasley, J. N., Hodapp, K. W., Jedicke, R., Kaiser, N., Kudritzki, R.-P., Luppino, G. A., Lupton, R. H., Monet, D. G., Morgan, J. S., Onaka, P. M., Stubbs, C. W., Tonry, J. L., Banados, E., Bell, E. F., Bender, R., Bernard, E. J., Botticella, M. T., Casertano, S., Chastel, S., Chen, W.-P., Chen, X., Cole, S., Deacon, N., Frenk, C., Fitzsimmons, A., Gezari, S., Goessl, C., Goggia, T., Goldman, B., Grebel, E. K., Hambly, N. C., Hasinger, G., Heavens, A. F., Heckman, T. M., Henderson, R., Henning, T., Holman, M., Hopp, U., Ip, W.-H., Isani, S., Keyes, C. D., Koekemoer, A., Kotak, R., Long, K. S., Lucey, J. R., Liu, M., Martin, N. F., McLean, B., Morganson,

E., Murphy, D. N. A., Nieto-Santisteban, M. A., Norberg, P., Peacock, J. A., Pier, E. A., Postman, M., Primak, N., Rae, C., Rest, A., Riess, A., Riffeser, A., Rix, H. W., Roser, S., Schilbach, E., Schultz, A. S. B., Scolnic, D., Szalay, A., Seitz, S., Shiao, B., Small, E., Smith, K. W., Soderblom, D., Taylor, A. N., Thakar, A. R., Thiel, J., Thilker, D., Urata, Y., Valenti, J., Walter, F., Watters, S. P., Werner, S., White, R., Wood-Vasey, W. M., and Wyse, R. (2016). The Pan-STARRS1 Surveys. *ArXiv e-prints*.

Chen, B., Vergely, J. L., Valette, B., and Carraro, G. (1998). Comparison of two different extinction laws with HIPPARCOS observations. *Astronomy&Astrophysics*, 336:137–149.

Chen, B.-Q., Huang, Y., Liu, X.-W., Yuan, H.-B., Wang, C., Fan, D.-W., Xiang, M.-S., Zhang, H.-W., and Tian, Z.-J. (2018). Three-dimensional interstellar dust reddening maps of the Galactic plane. *ArXiv e-prints*.

Chen, H.-C., Lallement, R., Babusiaux, C., Puspitarini, L., Bonifacio, P., and Hill, V. (2013). Extracting interstellar diffuse absorption bands from cool star spectra. Application to bulge clump giants in Baade's window. *Astronomy&Astrophysics*, 550:A62.

Chen, H.-C., Lallement, R., Puspitarini, L., Bonifacio, P., Babusiaux, C., and Hill, V. (2014). Automated Extraction of DIBs from Cool Star Spectra. In Cami, J. and Cox, N. L. J., editors, *The Diffuse Interstellar Bands*, volume 297 of *IAU Symposium*, pages 110–112.

Cirasuolo, M., Afonso, J., Carollo, M., Flores, H., Maiolino, R., Oliva, E., Paltani, S., Vanzi, L., Evans, C., Abreu, M., Atkinson, D., Babusiaux, C., Beard, S., Bauer, F., Bellazzini, M., Bender, R., Best, P., Bezawada, N., Bonifacio, P., Bragaglia, A., Bryson, I., Busher, D., Cabral, A., Caputi, K., Centrone, M., Chemla, F., Cimatti, A., Cioni, M.-R., Clementini, G., Coelho, J., Crnojevic, D., Daddi, E., Dunlop, J., Eales, S., Feltzing, S., Ferguson, A., Fisher, M., Fontana, A., Fynbo, J., Garilli, B., Gilmore, G., Glauser, A., Guinouard, I., Hammer, F., Hastings, P., Hess, A., Ivison, R., Jagourel, P., Jarvis, M., Kaper, L., Kauffman, G., Kitching, A. T., Lawrence, A., Lee, D., Lemasle, B., Licausi, G., Lilly, S., Lorenzetti, D., Lunney, D., Maiolino, R., Mannucci, F., McLure, R., Minniti, D., Montgomery, D., Muschielok, B., Nandra, K., Navarro, R., Norberg, P., Oliver, S., Origlia, L., Padilla, N., Peacock, J., Pedichini, F., Peng, J., Pentericci, L., Pragt, J., Puech, M., Randich, S., Rees, P., Renzini, A., Ryde, N., Rodrigues, M., Roseboom, I., Royer, F., Saglia, R., Sanchez, A., Schiavon, R., Schnetler, H., Sobral, D., Speziali, R., Sun, D., Stuik, R., Taylor, A., Taylor, W., Todd, S., Tolstoy, E., Torres, M., Tosi, M., Vanzella, E., Venema, L., Vitali, F., Wegner, M., Wells, M., Wild, V., Wright, G., Zamorani, G., and Zoccali, M. (2014). MOONS: the Multi-Object Optical and Near-infrared Spectrograph for the VLT. In *Ground-based and Airborne Instrumentation for Astronomy V*, volume 9147 of *Proceedings of the SPIE*, page 91470N.

Cox, N. L. J., Cami, J., Farhang, A., Smoker, J., Monreal-Ibero, A., Lallement, R., Sarre, P. J., Marshall, C. C. M., Smith, K. T., Evans, C. J., Royer, P., Linnartz, H., Cordiner, M. A., Joblin, C., van Loon, J. T., Foing, B. H., Bhatt, N. H., Bron, E., Elyajouri, M., de Koter, A., Ehrenfreund, P., Javadi,

A., Kaper, L., Khosroshadi, H. G., Laverick, M., Le Petit, F., Mulas, G., Roueff, E., Salama, F., and Spaans, M. (2017). The ESO Diffuse Interstellar Bands Large Exploration Survey (EDIBLES) . I. Project description, survey sample, and quality assessment. *Astronomy&Astrophysics*, 606:A76.

Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., Li, G.-P., Li, Q., Zhang, L.-P., Su, H.-J., Yao, Z.-Q., Wang, Y.-N., Xing, X.-Z., Li, X.-N., Zhu, Y.-T., Wang, G., Gu, B.-Z., Luo, A.-L., Xu, X.-Q., Zhang, Z.-C., Liu, G.-R., Zhang, H.-T., Yang, D.-H., Cao, S.-Y., Chen, H.-Y., Chen, J.-J., Chen, K.-X., Chen, Y., Chu, J.-R., Feng, L., Gong, X.-F., Hou, Y.-H., Hu, H.-Z., Hu, N.-S., Hu, Z.-W., Jia, L., Jiang, F.-H., Jiang, X., Jiang, Z.-B., Jin, G., Li, A.-H., Li, Y., Li, Y.-P., Liu, G.-Q., Liu, Z.-G., Lu, W.-Z., Mao, Y.-D., Men, L., Qi, Y.-J., Qi, Z.-X., Shi, H.-M., Tang, Z.-H., Tao, Q.-S., Wang, D.-Q., Wang, D., Wang, G.-M., Wang, H., Wang, J.-N., Wang, J., Wang, J.-L., Wang, J.-P., Wang, L., Wang, S.-Q., Wang, Y., Wang, Y.-F., Xu, L.-Z., Xu, Y., Yang, S.-H., Yu, Y., Yuan, H., Yuan, X.-Y., Zhai, C., Zhang, J., Zhang, Y.-X., Zhang, Y., Zhao, M., Zhou, F., Zhou, G.-H., Zhu, J., and Zou, S.-C. (2012). The Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST). *Research in Astronomy and Astrophysics*, 12:1197–1242.

Dame, T. M., Hartmann, D., and Thaddeus, P. (2001). The Milky Way in Molecular Clouds: A New Complete CO Survey. *Astrophysical Journal*, 547:792–813.

Dame, T. M. and Thaddeus, P. (2004). A Large Extension of the CfA Galactic CO Survey. In Clemens, D., Shah, R., and Brainerd, T., editors, *Milky Way Surveys: The Structure and Evolution of our Galaxy*, volume 317 of *Astronomical Society of the Pacific Conference Series*, page 66.

Danielski, C., Babusiaux, C., Ruiz-Dern, L., Sartoretti, P., and Arenou, F. (2018). The empirical Gaia G-band extinction coefficient. *Astronomy&Astrophysics*, 614:A19.

de Avillez, M. A., Breitschwerdt, D., Asgekar, A., and Spitoni, E. (2015). ISM simulations: an overview of models. *Highlights of Astronomy*, 16:606–608.

Elyajouri, M., Cox, N. L. J., and Lallement, R. (2017a). The 15 273 Å diffuse interstellar band in the dark cloud Barnard 68. *Astronomy&Astrophysics*, 605:L10.

Elyajouri, M., Lallement, R., Monreal-Ibero, A., Capitanio, L., and Cox, N. L. J. (2017b). Near-infrared diffuse interstellar bands in APOGEE telluric standard star spectra . Weak bands and comparisons with optical counterparts. *Astronomy&Astrophysics*, 600:A129.

Elyajouri, M., Monreal-Ibero, A., Remy, Q., and Lallement, R. (2016). A Catalog of 1.5273 um Diffuse Interstellar Bands Based on APOGEE Hot Telluric Calibrators. *Astrophysical Journal Supplement*, 225:19.

Ferrière, K. M. (2001). The interstellar environment of our galaxy. *Reviews of Modern Physics*, 73:1031–1066.

Ferrière, K. M. (2015). Theoretical understanding of Galactic magnetic fields. *Highlights of Astronomy*, 16:396–396.

Flewelling, H. A., Magnier, E. A., Chambers, K. C., Heasley, J. N., Holmberg, C., Huber, M. E., Sweeney, W., Waters, C. Z., Chen, X., Farrow, D., Hasinger, G., Henderson, R., Long, K. S., Metcalfe, N., Nieto-Santisteban, M. A., Norberg, P., Saglia, R. P., Szalay, A., Rest, A., Thakar, A. R., Tonry, J. L., Valenti, J., Werner, S., White, R., Denneau, L., Draper, P. W., Hodapp, K. W., Jedicke, R., Kaiser, N., Kudritzki, R.-P., Price, P. A., Wainscoat, R. J., Chastel, S., McClean, B., Postman, M., and Shiao, B. (2016). The Pan-STARRS1 Database and Data Products. *ArXiv e-prints*.

Foing, B. H. and Ehrenfreund, P. (1994). Detection of two interstellar absorption bands coincident with spectral features of C_{60}^+. *Nature*, 369:296–298.

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Mignard, F., Drimmel, R., Babusiaux, C., Bailer-Jones, C. A. L., Bastian, U., and et al. (2016a). Gaia Data Release 1. Summary of the astrometric, photometric, and survey properties. *Astronomy&Astrophysics*, 595:A2.

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., Vallenari, A., Babusiaux, C., Bailer-Jones, C. A. L., Bastian, U., Biermann, M., Evans, D. W., and et al. (2016b). The Gaia mission. *Astronomy&Astrophysics*, 595:A1.

Gavilan, L., Le, K. C., Pino, T., Alata, I., Giuliani, A., and Dartois, E. (2017). Polyaromatic disordered carbon grains as carriers of the UV bump: Far-UV to mid-IR spectroscopy of laboratory analogs. *Astronomy&Astrophysics*, 607:A73.

Georgelin, Y. M. and Georgelin, Y. P. (1976). The spiral structure of our Galaxy determined from H II regions. *Astronomy&Astrophysics*, 49:57–79.

Girardi, L. (2016). Red Clump Stars. *Annual Review of Astronomy and Astrophysics*, 54:95–133.

Green, G. M., Schlafly, E. F., Finkbeiner, D. P., Jurić, M., Rix, H.-W., Burgett, W., Chambers, K. C., Draper, P. W., Flewelling, H., Kudritzki, R. P., Magnier, E., Martin, N., Metcalfe, N., Tonry, J., Wainscoat, R., and Waters, C. (2014). Measuring Distances and Reddenings for a Billion Stars: Toward a 3D Dust Map from Pan-STARRS 1. *Astrophysical Journal*, 783:114.

Green, G. M., Schlafly, E. F., Finkbeiner, D. P., Rix, H.-W., Martin, N., Burgett, W., Draper, P. W., Flewelling, H., Hodapp, K., Kaiser, N., Kudritzki, R. P., Magnier, E., Metcalfe, N., Price, P., Tonry, J., and Wainscoat, R. (2015). A Three-dimensional Map of Milky Way Dust. *Astrophysical Journal*, 810:25.

Hobbs, L. M. (1978). Optical interstellar lines toward 18 stars of low reddening. III. *Astrophysical Journal*, 222:491–507.

Hobbs, L. M., York, D. G., Snow, T. P., Oka, T., Thorburn, J. A., Bishof, M., Friedman, S. D., McCall, B. J., Rachford, B., Sonnentrucker, P., and Welty, D. E. (2008). A Catalog of Diffuse Interstellar Bands in the Spectrum of HD 204827. *Astrophysical Journal*, 680:1256–1270.

Hou, L. G. and Han, J. L. (2014). The observed spiral structure of the Milky Way. *Astronomy&Astrophysics*, 569:A125.

Jones, A. P. (2016). Dust evolution, a global view: III. Core/mantle grains, organic nano-globules, comets and surface chemistry. *Royal Society Open Science*, 3:160224.

Jones, A. P., Fanciullo, L., Köhler, M., Verstraete, L., Guillet, V., Bocchio, M., and Ysard, N. (2013). The evolution of amorphous hydrocarbons in the ISM: dust modelling from a new vantage point. *Astronomy&Astrophysics*, 558:A62.

Karttunen, H., Kroeger, P., Oja, H., Poutanen, M., and Donner, K. J. (2003). *Fundamental astronomy*.

Katz, D., Munari, U., Cropper, M., Zwitter, T., Thévenin, F., David, M., Viala, Y., Crifo, F., Gomboc, A., Royer, F., Arenou, F., Marrese, P., Sordo, R., Wilkinson, M., Vallenari, A., Turon, C., Helmi, A., Bono, G., Perryman, M., Gómez, A., Tomasella, L., Boschi, F., Morin, D., Haywood, M., Soubiran, C., Castelli, F., Bijaoui, A., Bertelli, G., Prsa, A., Mignot, S., Sellier, A., Baylac, M.-O., Lebreton, Y., Jauregi, U., Siviero, A., Bingham, R., Chemla, F., Coker, J., Dibbens, T., Hancock, B., Holland, A., Horville, D., Huet, J.-M., Laporte, P., Melse, T., Sayède, F., Stevenson, T.-J., Vola, P., Walton, D., and Winter, B. (2004). Spectroscopic survey of the Galaxy with Gaia-I. Design and performance of the Radial Velocity Spectrometer. *Monthly Notices of the Royal Astronomical Society*, 354:1223–1238.

Kos, J., Zwitter, T., Wyse, R., Bienaymé, O., Binney, J., Bland-Hawthorn, J., Freeman, K., Gibson, B. K., Gilmore, G., Grebel, E. K., Helmi, A., Kordopatis, G., Munari, U., Navarro, J., Parker, Q., Reid, W. A., Seabroke, G., Sharma, S., Siebert, A., Siviero, A., Steinmetz, M., Watson, F. G., and Williams, M. E. K. (2014). Pseudo-three-dimensional maps of the diffuse interstellar band at 862 nm. *Science*, 345:791–795.

Lallement, R., Capitanio, L., Ruiz-Dern, L., Danielski, C., Babusiaux, C., Vergely, L., Elyajouri, M., Arenou, F., and Leclerc, N. (2018). Three-dimensional maps of interstellar dust in the Local Arm: using Gaia, 2MASS, and APOGEE-DR14. *Astronomy&Astrophysics*, 616:A132.

Lallement, R., Vergely, J.-L., Valette, B., Puspitarini, L., Eyer, L., and Casagrande, L. (2014). 3D maps of the local ISM from inversion of individual color excess measurements. *A&A*, 561:A91.

Lallement, R., Welsh, B. Y., Vergely, J. L., Crifo, F., and Sfeir, D. (2003). 3D mapping of the dense interstellar gas around the Local Bubble. *Astronomy&Astrophysics*, 411:447–464.

Lépine, J. R. D., Michtchenko, T. A., Barros, D. A., and Vieira, R. S. S. (2017). The Dynamical Origin of the Local Arm and the Sun's Trapped Orbit. *Astrophysical Journal*, 843:48.

Luri, X., Palmer, M., Arenou, F., Masana, E., de Bruijne, J., Antiche, E., Babusiaux, C., Borrachero, R., Sartoretti, P., Julbe, F., Isasi, Y., Martinez,

O., Robin, A. C., Reylé, C., Jordi, C., and Carrasco, J. M. (2014). Overview and stellar statistics of the expected Gaia Catalogue using the Gaia Object Generator. *Astronomy&Astrophysics*, 566:A119.

Maíz Apellániz, J., Barbá, R. H., Sota, A., and Simón-Díaz, S. (2015). The little-studied cluster Berkeley 90. II. The foreground ISM. *Astronomy&Astrophysics*, 583:A132.

Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., and Picaud, S. (2006). Modelling the Galactic interstellar extinction distribution in three dimensions. *Astronomy&Astrophysics*, 453:635–651.

Moutou, C., Boisse, I., Hébrard, G., Hébrard, E., Donati, J.-F., Delfosse, X., and Kouach, D. (2015). SPIRou: a spectropolarimeter for the CFHT. In Martins, F., Boissier, S., Buat, V., Cambrésy, L., and Petit, P., editors, *SF2A-2015: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, pages 205–212.

Munari, U. (2000). The diffuse interstellar band at 8620 Å: A good reddening tracer for GAIA. In Porceddu, I. and Aiello, S., editors, *Molecules in Space and in the Laboratory*, page 179.

Planck Collaboration, Abergel, A., Ade, P. A. R., Aghanim, N., Alves, M. I. R., Aniano, G., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., and et al. (2014). Planck 2013 results. XI. All-sky model of thermal dust emission. *Astronomy&Astrophysics*, 571:A11.

Puspitarini, L., Lallement, R., Babusiaux, C., Chen, H.-C., Bonifacio, P., Sbordone, L., Caffau, E., Duffau, S., Hill, V., Monreal-Ibero, A., Royer, F., Arenou, F., Peralta, R., Drew, J. E., Bonito, R., Lopez-Santiago, J., Alfaro, E. J., Bensby, T., Bragaglia, A., Flaccomio, E., Lanzafame, A. C., Pancino, E., Recio-Blanco, A., Smiljanic, R., Costado, M. T., Lardo, C., de Laverny, P., and Zwitter, T. (2015). The Gaia-ESO Survey: Extracting diffuse interstellar bands from cool star spectra. DIB-based interstellar medium line-of-sight structures at the kpc scale. *A&A*, 573:A35.

Puspitarini, L., Lallement, R., and Chen, H.-C. (2013). Automated measurements of diffuse interstellar bands in early-type star spectra. Correlations with the color excess. *A&A*, 555:A25.

Queiroz, A. B. A., Anders, F., Santiago, B. X., Chiappini, C., Steinmetz, M., Dal Ponte, M., Stassun, K. G., da Costa, L. N., Maia, M. A. G., Crestani, J., Beers, T. C., Fernández-Trincado, J. G., García-Hernández, D. A., Roman-Lopes, A., and Zamora, O. (2018). StarHorse: a Bayesian tool for determining stellar masses, ages, distances, and extinctions for field stars. *Monthly Notices of the Royal Astronomical Society*, 476:2556–2583.

Recio-Blanco, A., de Laverny, P., Allende Prieto, C., Fustes, D., Manteiga, M., Arcay, B., Bijaoui, A., Dafonte, C., Ordenovic, C., and Ordoñez Blanco, D. (2016). Stellar parametrization from Gaia RVS spectra. *Astronomy&Astrophysics*, 585:A93.

Reid, M. J., Menten, K. M., Brunthaler, A., Zheng, X. W., Dame, T. M., Xu, Y., Wu, Y., Zhang, B., Sanna, A., Sato, M., Hachisuka, K., Choi, Y. K., Immer, K., Moscadelli, L., Rygl, K. L. J., and Bartkiewicz, A. (2014). Trigonometric Parallaxes of High Mass Star Forming Regions: The Structure and Kinematics of the Milky Way. *Astrophysical Journal*, 783:130.

Rezaei Kh., S., Bailer-Jones, C. A. L., Hanson, R. J., and Fouesneau, M. (2016). Inferring the three-dimensional distribution of dust in the Galaxy with a non-parametric method: Preparing for Gaia. *ArXiv e-prints*.

Rezaei Kh., S., Bailer-Jones, C. A. L., Hogg, D. W., and Schultheis, M. (2018). Detection of the Milky Way spiral arms in dust from 3D mapping. *ArXiv e-prints*.

Robin, A. C., Luri, X., Reylé, C., Isasi, Y., Grux, E., Blanco-Cuaresma, S., Arenou, F., Babusiaux, C., Belcheva, M., Drimmel, R., Jordi, C., Krone-Martins, A., Masana, E., Mauduit, J. C., Mignard, F., Mowlavi, N., Rocca-Volmerange, B., Sartoretti, P., Slezak, E., and Sozzetti, A. (2012). Gaia Universe model snapshot. A statistical analysis of the expected contents of the Gaia catalogue. *Astronomy&Astrophysics*, 543:A100.

Robin, A. C., Reylé, C., Derrière, S., and Picaud, S. (2003). A synthetic view on structure and evolution of the Milky Way. *Astronomy&Astrophysics*, 409:523–540.

Rockstroh, J. M. and Webber, W. R. (1978). A new determination of the local interstellar electron spectrum from the radio background. *Astrophysical Journal*, 224:677–690.

Ruiz-Dern, L., Babusiaux, C., Arenou, F., Turon, C., and Lallement, R. (2018). Empirical photometric calibration of the Gaia red clump: Colours, effective temperature, and absolute magnitude. *Astronomy&Astrophysics*, 609:A116.

Salama, F., Bakes, E. L. O., Allamandola, L. J., and Tielens, A. G. G. M. (1996). Assessment of the Polycyclic Aromatic Hydrocarbon–Diffuse Interstellar Band Proposal. *Astrophysical Journal*, 458:621.

Sale, S. E. (2015). Galactic 3D extinction maps. *Memorie della Societá Astronomica Italiana*, 86:568.

Sale, S. E. and Magorrian, J. (2014a). Three-dimensional extinction mapping using Gaussian random fields. *Monthly Notices of the Royal Astronomical Society*, 445:256–269.

Sale, S. E. and Magorrian, J. (2014b). Three-dimensional extinction mapping using Gaussian random fields. *MNRAS*, 445:256–269.

Santiago, B. X., Brauer, D. E., Anders, F., Chiappini, C., Queiroz, A. B., Girardi, L., Rocha-Pinto, H. J., Balbinot, E., da Costa, L. N., Maia, M. A. G., Schultheis, M., Steinmetz, M., Miglio, A., Montalbán, J., Schneider, D. P., Beers, T. C., Frinchaboy, P. M., Lee, Y. S., and Zasowski, G. (2016). Spectrophotometric distances to stars: A general purpose Bayesian approach. *Astronomy&Astrophysics*, 585:A42.

Savage, B. D. and Mathis, J. S. (1979). Observed properties of interstellar dust. *Annual Review of Astronomy and Astrophysics*, 17:73–111.

Schlafly, E. F., Peek, J. E. G., Finkbeiner, D. P., and Green, G. M. (2017). Mapping the Extinction Curve in 3D: Structure on Kiloparsec Scales. *Astrophysical Journal*, 838:36.

Schlegel, D. J., Finkbeiner, D. P., and Davis, M. (1998). Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds. *Astrophysical Journal*, 500:525–553.

Snow, T. P. (2014). Diffuse Interstellar Bands: Past and Present. In Cami, J. and Cox, N. L. J., editors, *The Diffuse Interstellar Bands*, volume 297 of *IAU Symposium*, pages 3–12.

Sordo, R., Vallenari, A., Tantalo, R., Liu, C., Smith, K., Allard, F., Blomme, R., Bouret, J.-C., Brott, I., de Laverny, P., Edvardsson, B., Frémat, Y., Heber, U., Josselin, E., Kochukhov, O., Korn, A., Lanzafame, A., Martayan, C., Martins, F., Plez, B., Schweitzer, A., Thévenin, F., and Zorec, J. (2011). Stellar libraries for Gaia. In *Journal of Physics Conference Series*, volume 328 of *Journal of Physics Conference Series*, page 012006.

Tarantola, A. and Valette, B. (1982). Generalized Nonlinear Inverse Problems Solved Using the Least Squares Criterion (Paper 1R1855). *Reviews of Geophysics and Space Physics*, 20:219.

van der Zwet, G. P. and Allamandola, L. J. (1985). Polycyclic aromatic hydrocarbons and the diffuse interstellar bands. *Astronomy&Astrophysics*, 146:76–80.

van Leeuwen, F. (2007). Validation of the new Hipparcos reduction. *Astronomy&Astrophysics*, 474:653–664.

van Loon, J. T., Bailey, M., Tatton, B. L., Maíz Apellániz, J., Crowther, P. A., de Koter, A., Evans, C. J., Hénault-Brunet, V., Howarth, I. D., Richter, P., Sana, H., Simón-Díaz, S., Taylor, W., and Walborn, N. R. (2013). The VLT-FLAMES Tarantula Survey. IX. The interstellar medium seen through diffuse interstellar bands and neutral sodium. *AA*, 550:A108.

van Loon, J. T., Farhang, A., Javadi, A., Bailey, M., and Khosroshahi, H. G. (2015). Probing the Local Bubble with Diffuse Interstellar Bands (DIBs). *Memorie della Societá Astronomica Italiana*, 86:534.

Vergely, J.-L., Freire Ferrero, R., Siebert, A., and Valette, B. (2001). NaI and HI 3-D density distribution in the solar neighbourhood. *Astronomy&Astrophysics*, 366:1016–1034.

Vergely, J.-L., Valette, B., Lallement, R., and Raimond, S. (2010a). Spatial distribution of interstellar dust in the Sun's vicinity. Comparison with neutral sodium-bearing gas. *Astronomy&Astrophysics*, 518:A31.

Vergely, J.-L., Valette, B., Lallement, R., and Raimond, S. (2010b). Spatial distribution of interstellar dust in the Sun's vicinity. Comparison with neutral sodium-bearing gas. *A&A*, 518:A31.

Vladilo, G., Beckman, J. E., Crivellari, L., Franco, M. L., and Molaro, P. (1985). The distribution of the local interstellar medium derived from MG II column densities towards seven cool stars. *Astronomy&Astrophysics*, 144:81–86.

Welsh, B. Y., Lallement, R., Vergely, J.-L., and Raimond, S. (2010). New 3D gas density maps of NaI and CaII interstellar absorption within 300 pc. *Astronomy&Astrophysics*, 510:A54.

Wilkinson, M. I., Vallenari, A., Turon, C., Munari, U., Katz, D., Bono, G., Cropper, M., Helmi, A., Robichon, N., Thévenin, F., Vidrih, S., Zwitter, T., Arenou, F., Baylac, M.-O., Bertelli, G., Bijaoui, A., Boschi, F., Castelli, F., Crifo, F., David, M., Gomboc, A., Gómez, A., Haywood, M., Jauregi, U., de Laverny, P., Lebreton, Y., Marrese, P., Marsh, T., Mignot, S., Morin, D., Pasetto, S., Perryman, M., Prša, A., Recio-Blanco, A., Royer, F., Sellier, A., Siviero, A., Sordo, R., Soubiran, C., Tomasella, L., and Viala, Y. (2005). Spectroscopic survey of the Galaxy with Gaia- II. The expected science yield from the Radial Velocity Spectrometer. *Monthly Notices of the Royal Astronomical Society*, 359:1306–1335.

Zasowski, G., Johnson, J. A., Frinchaboy, P. M., Majewski, S. R., Nidever, D. L., Rocha Pinto, H. J., Girardi, L., Andrews, B., Chojnowski, S. D., Cudworth, K. M., Jackson, K., Munn, J., Skrutskie, M. F., Beaton, R. L., Blake, C. H., Covey, K., Deshpande, R., Epstein, C., Fabbian, D., Fleming, S. W., Garcia Hernandez, D. A., Herrero, A., Mahadevan, S., Mészáros, S., Schultheis, M., Sellgren, K., Terrien, R., van Saders, J., Allende Prieto, C., Bizyaev, D., Burton, A., Cunha, K., da Costa, L. N., Hasselquist, S., Hearty, F., Holtzman, J., García Pérez, A. E., Maia, M. A. G., O'Connell, R. W., O'Donnell, C., Pinsonneault, M., Santiago, B. X., Schiavon, R. P., Shetrone, M., Smith, V., and Wilson, J. C. (2013). Target Selection for the Apache Point Observatory Galactic Evolution Experiment (APOGEE). *AJ*, 146:81.

Zasowski, G., Ménard, B., Bizyaev, D., García-Hernández, D. A., García Pérez, A. E., Hayden, M. R., Holtzman, J., Johnson, J. A., Kinemuchi, K., Majewski, S. R., Nidever, D. L., Shetrone, M., and Wilson, J. C. (2015). Mapping the Interstellar Medium with Near-infrared Diffuse Interstellar Bands. *ApJ*, 798:35.

RÉSUMÉ

Le milieu interstellaire (MIS) représente une fraction mineure de la matière baryonique de la Voie Lactée (environ 10%), mais joue un rôle dans l'evolution et la structure galactique. La majorité des constituants du MIS ont été identifiés grace à leur emissions or absorptions, une catégorie d'absorption,les "bandes interstellaires diffuses"(DIBs), ont pas des particules absorbants identifiées. En particulier, il y a une DIB centrée `a 862 nm dans e intervalle spectroscopique du Radial Velocity Spectrometer (RVS), un des instruments des la mission ESA Gaia. La première partie de ma thèse est relié à la construction de ce future catalogue et son usage potentiel. J'ai développé un code de profile fitting adapté à la DIB Gaia et à l'instrument RVS. Je pr ´esente dans ma thèse l'estimation de DIB mesurable. La second partie de ma thèse est en relation aux cartes 3D du MIS. J'ai focalisé sur l'usage potentiel des DIBs pour le cartographie. J'ai fait la première tomography basée sur des données composée, extinction et DIB. J'ai tournée des inversions parallels des extinctions et DIB pour les mêmes lignes des visée, et comparés les distributions. J'ai développé une technique exprès pour les données massives, appelée "technique hiérarchique" et j'ai appliqué cet technique aux nouveau catalog de 30 millions des mesures d'extinction base sur Gaia. Finalement, j'ai utilisé des spectres au sol pour extraire soit les vitesses soit la colonne des matériaux interstellaire avec du profile-fitting et j'ai participé aux premiers steps d'une carte cinématique 3D du MIS.

MOTS CLÉS

Milieu interstellaire - Mission Gaia - Tomographie - Spectroscopie

ABSTRACT

The Interstellar Medium (ISM) represents a small fraction of the Milky Way matter, but it plays a fundamental role because it governs the galactic evolution and its structure. Most of the ISM gaseous constituents have been identified through their emission or absorption characteristics, however, the "diffuse interstellar bands" (or DIBs), still has no counterpart in absorbing species. A favorable situation is the presence of a DIB (with central wavelength 862 nm) within the spectral interval of the Radial Velocity Spectrometer (RVS), one of the instruments of the ESA Gaia mission. The first part of the work presented in this document is directly related to the construction of this future catalog and its potential use. I have developed a profile fitting code adapted to the Gaia DIB and to the RVS instrument. I make some estimations in my thesis how many DIBs will be measured. The second part of this thesis is related to the three-dimensional (3D) mapping of the ISM. I made the first tomography based on composite data, both extinctions and DIBs and showed this is a viable way. I ran parallel inversions of two extinction and DIB datasets for the same target stars and compared the distributions. The full 3D tomographic inversion technique in this case requires an unrealistically large computing time. I have developed a technique and I have applied it to a new catalog of 30 millions of extinction measurements based on Gaia. Finally, I used ground-based stellar spectra to extracted radial velocities and columns of IS absorbing species and participated to the first step of a "kinematic" 3D map of the ISM.

KEYWORDS

Interstellar medium - Gaia mission - Tomography - Spectroscopy