

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

L. Capitanio<sup>1</sup>, R. Lallement<sup>1</sup>, J.L. Vergely<sup>2</sup>, M. Elyajouri<sup>1</sup>, and A. Monreal-Ibero<sup>1</sup>

<sup>1</sup> 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

<sup>2</sup> 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  $\approx 5,000$  color excess values estimated from the strength of the  $\lambda 15273$  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 & Magor-

rian 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  $\leq 500$ pc), (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  $\approx 22,500$  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 ( $d \leq 500$ pc), 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  $\approx 400$  targets with photometric distances that had no catalog identifications. The coordinates, brightnesses, estimated distances, and reddening reasonably of these  $\approx 400$  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 sight-lines 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;  $1.51 - 1.70 \mu\text{m}$ ) at high spectral resolution ( $R \sim 22,500$ ; Wilson et al. 2010). Specifically, we downloaded spectra from the SDSS data release<sup>1</sup> (DR12; Alam et al. 2015), which provides all the data taken between April 2011 and July 2014. In addition to the  $\sim 160,000$  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  $\sim 17,000$  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  $5.5 \leq H \leq 11$  mag and are therefore hot and bright stars with spectra that are most often featureless.

The APOGEE products contain the decontaminated spectra

<sup>1</sup> <http://www.sdss.org/dr12/>

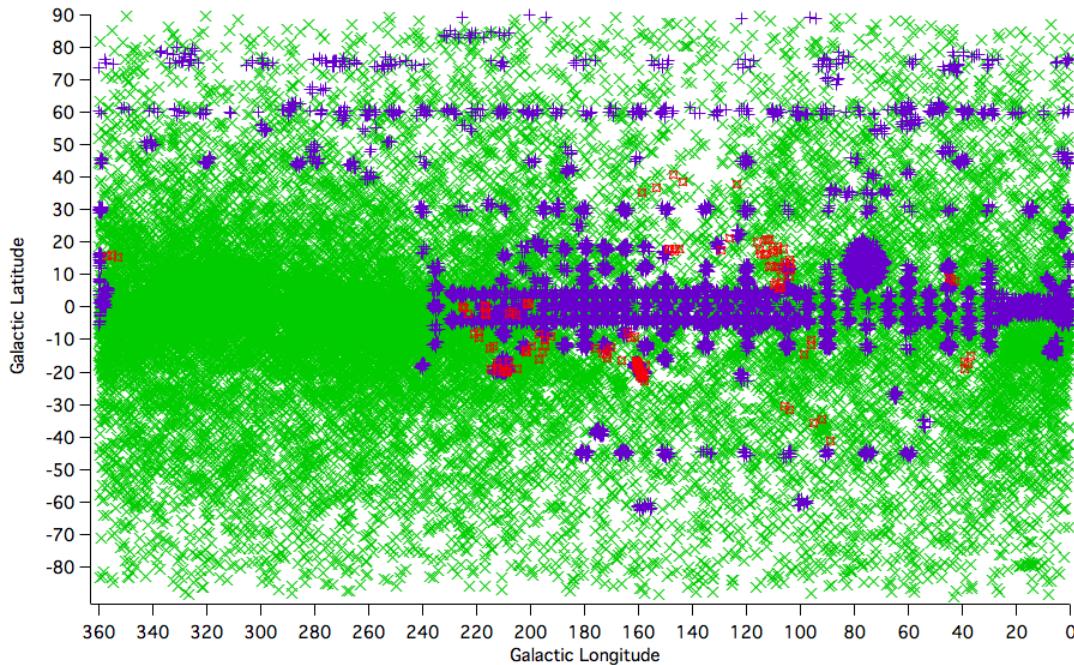


Fig. 1: Location of the targets from the LVV catalog compilation (green crosses) and of the APOGEE target stars retained for the DIB extraction and the inversion (violet signs). Coordinates are Galactic. Also shown are the directions of the molecular clouds of the Schlafly et al. (2014) catalog mentioned 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  $\lambda 15273$  Å 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<sup>2</sup>.

## 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 Brackett 19-4 stellar line at 15264.7 Å. We used for the fit the Levenberg-Marquardt algorithm implemented in IGOR PRO<sup>2</sup>.

We fit the data to a combination of an adjusted synthetic spectrum, DIB model, and smooth continuum as shown in equa-

tion 1. Below we provide information specific to each of the equation variables as follows:

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

- $[S_\lambda]^\alpha$ , an adjusted stellar spectrum, where  $S_\lambda$  is the stellar model provided by APOGEE. We included the scaling factor  $\alpha$  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 ( $\sigma$ ), central wavelength ( $\lambda_c$ ), 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  $A = [\lambda_c - 10 \text{ Å}, \lambda_c + 10 \text{ Å}]$  close to the DIB.

- RB was calculated over a spectral range  $B = [15890-15960] \text{ Å}$ , which is free of DIB absorption, strong stellar lines, and telluric residuals.

- RC was obtained from region  $C = [15200-15250] \text{ Å}$  relatively close to the DIB, and is also free of DIB and potentially

<sup>2</sup> <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  $\sigma$  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 re-

markable 3D dust reddening maps with a very high angular resolution (5-15 arc-minutes) covering distances from  $\approx 300$  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  $\psi$  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  $\xi_1$  and  $\xi_2$ , 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  $\xi_1=15$  pc and  $\xi_2=30$  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  $\sigma_1=0.8$  for the first kernel and  $\sigma_2=1.0$  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	$\lambda_c$ 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]	$\leq 1.4$	$\leq 2 \times RA$

Table 1: Fitting constraints of  $\lambda 15273$  NIR DIB for the selected APOGEE spectra. RA, RB, RC are the standard deviation in three spectral ranges A = [ $\lambda_c - 10\text{Å}$ ,  $\lambda_c + 10\text{Å}$ ], 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 dis-

tances 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  $\rho_0$  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).

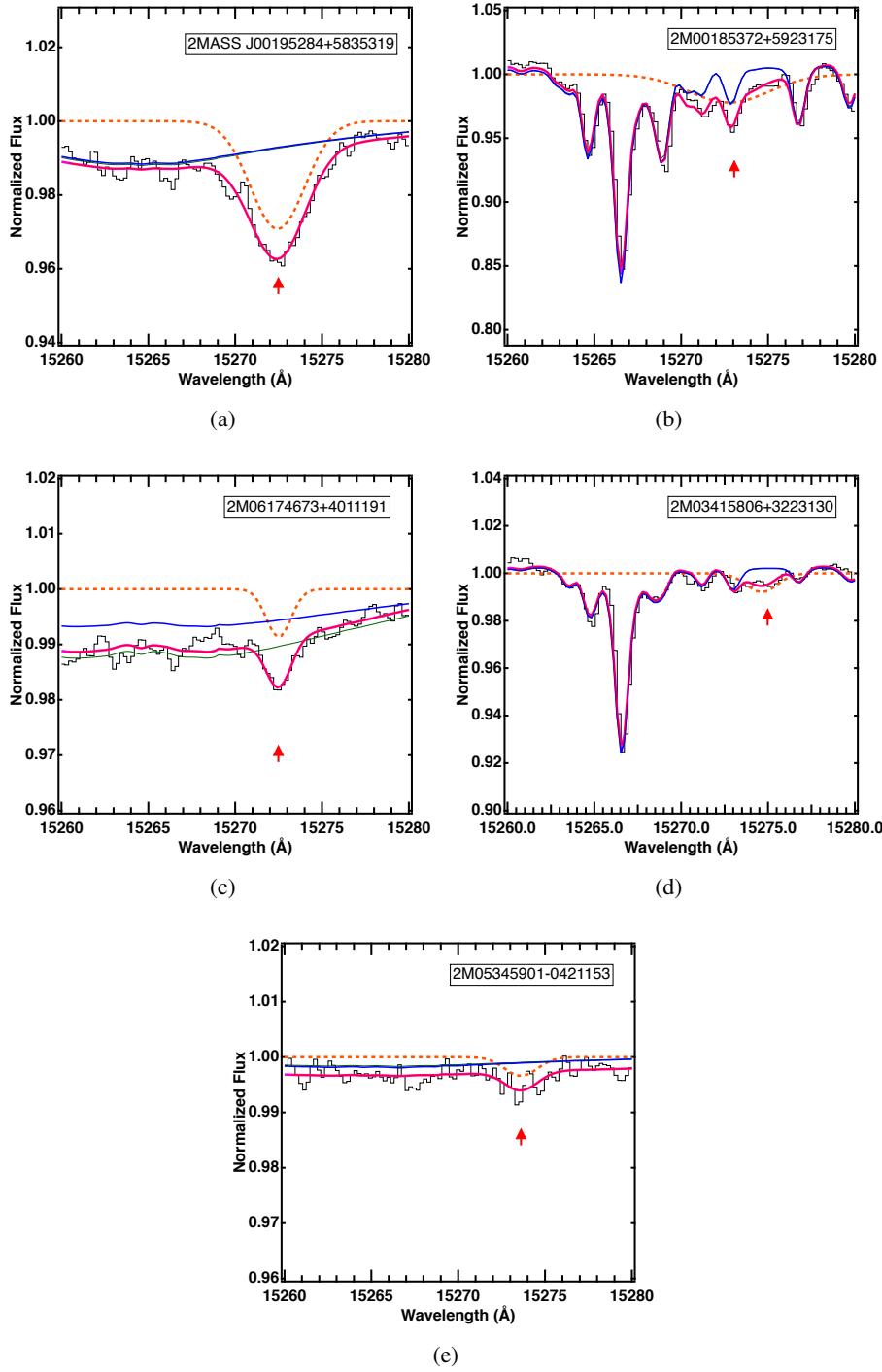


Fig. 2: Illustration of the various categories of DIB extraction: a) Detected; b) Recovered; c) Narrow; d) Recovered and Narrow; e) Upper limit. APOGEE spectra are shown with solid black curves. The initial stellar model provided by APOGEE and the model obtained after application of the scaling factor 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 ini-

tial 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 par-

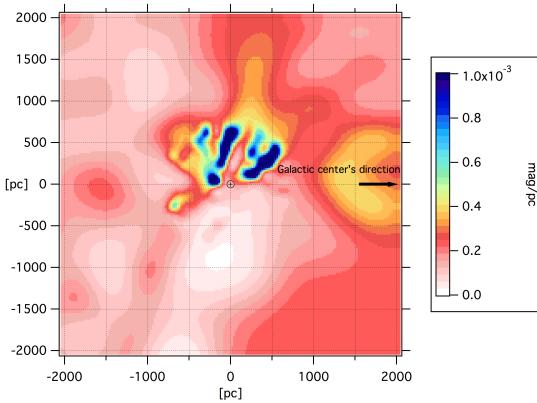


Fig. 3: Galactic plane cut in the very low-resolution 3D opacity distribution to be used as a prior in the inversion. It is based on Pan-STARRS reddening measurements of Green et al. (2015). The Sun is at (0,0) and the Galactic center to the right. Longitudes increase counter-clockwise. Units are  $\text{mag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

allax uncertainties. 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  $\chi^2$  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  $\chi^2$  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  $\chi^2$  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 particu-

lar in the Cygnus Rift region. At  $l \approx 70^\circ$  the Rift is located significantly farther away in the new inversion, at about 800 pc instead of  $\approx 600$  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  $\chi^2$  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 than 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

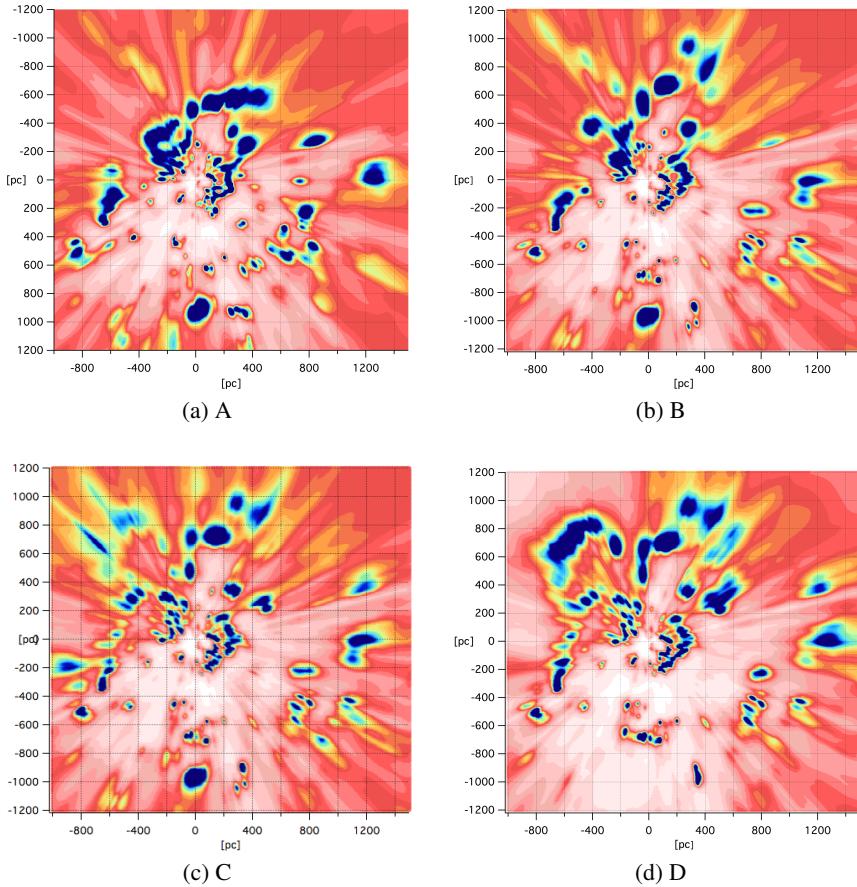


Fig. 4: Galactic plane cut in the 3D opacity distribution resulting from the four successive inversions: a) A distribution: color excess data alone and Hipparcos or photometric distances. b) B distribution: same as A except for the replacement of 80% of the initial distances with Gaia-DR1 (TGAS) values. 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 $\chi^2$	FINAL $\chi^2$
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  $\chi^2$  per line of sight in the four inversions.

the transformation of wide structures into smaller clumps and/or 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  $\approx 750$  lines of sight from the photometric catalog and  $\approx 375$  from the APOGEE catalog within the longitude interval  $l=(208:218^\circ)$ . 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  $\approx 450$  pc 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

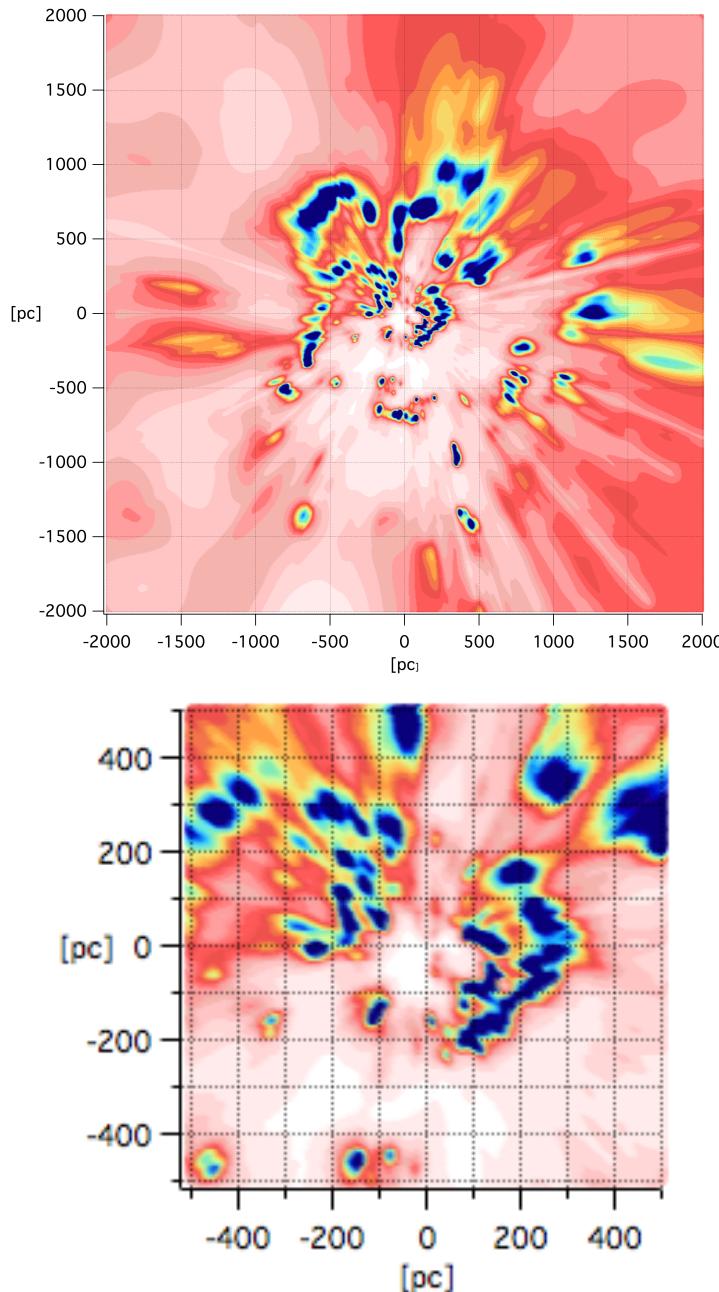


Fig. 5: Planar cut along the Galactic plane in the inverted 3D distribution D shown in Fig 4, panel d. The color scale is same as in Fig 3. Top: the entire 4000pc  $\times$  4000pc computed area. Bottom: Zoom in a 1000 pc  $\times$  1000 pc area around the Sun.

mapping, the increased compactness of the foreground structure 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  $l=(105^\circ;115^\circ)$  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  $\approx 300$  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.

#### 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

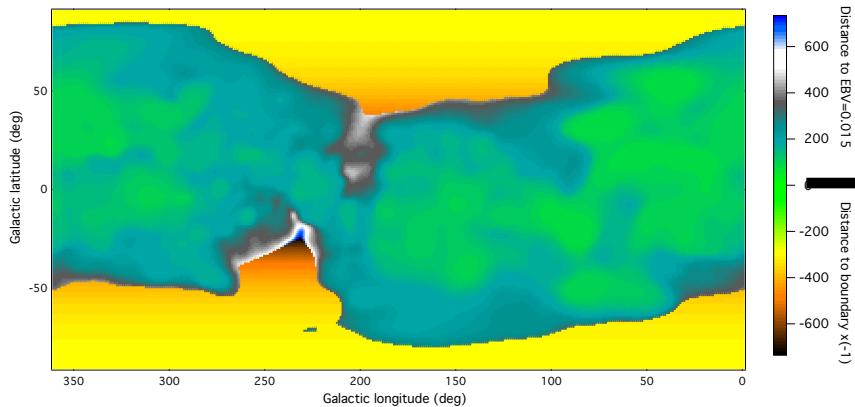


Fig. 6: Local cavity boundaries: the distance at which a color excess of 0.015 mag is reached is plotted as a function of rectangular Galactic coordinates. In very weakly reddened directions for which 0.015 mag is not reached within our computation domain, the distance to the domain boundary is plotted 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.

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  $\chi^2$  per point significantly and decreasing the final  $\chi^2$  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., Lalle-

ment (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 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

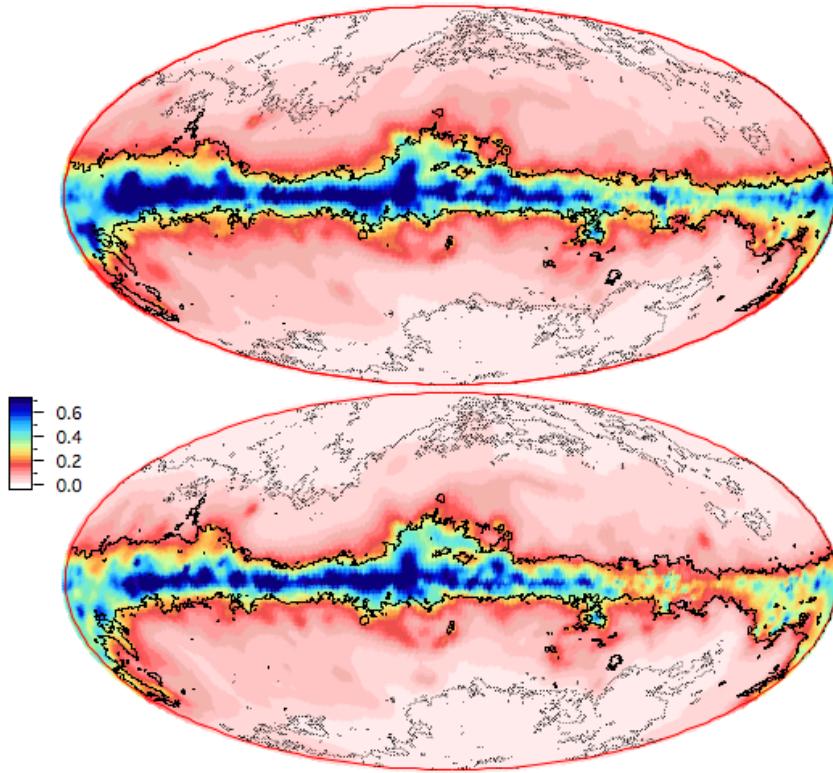


Fig. 7: Full-sky 2D map of opacity distribution integrated from the Sun to the boundaries of our computational domain ( $\text{abs}(Z)=300\text{pc}$ ,  $\text{abs}(X)=2\text{kpc}$ ,  $\text{abs}(Y)=2\text{kpc}$ ). Top: LVV results are shown (distribution A). Bottom: the final inversion with APOGEE data 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  $+180^\circ$  to  $-180^\circ$  from left to right. The color scale refers to  $E(B-V)$ , linear, in units of mag.

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  $\rho$  (in  $\text{pc}^{-3}$ ) 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 rednings 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  $50 \times 50 \times 50 \text{ pc}^3$ ; in case the cube contains less than 10 targets we use a volume of  $200 \times 200 \times 200 \text{ pc}^3$ . Uncertainties on distances and color excess will be available on our website and displayed interactively.

## 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).

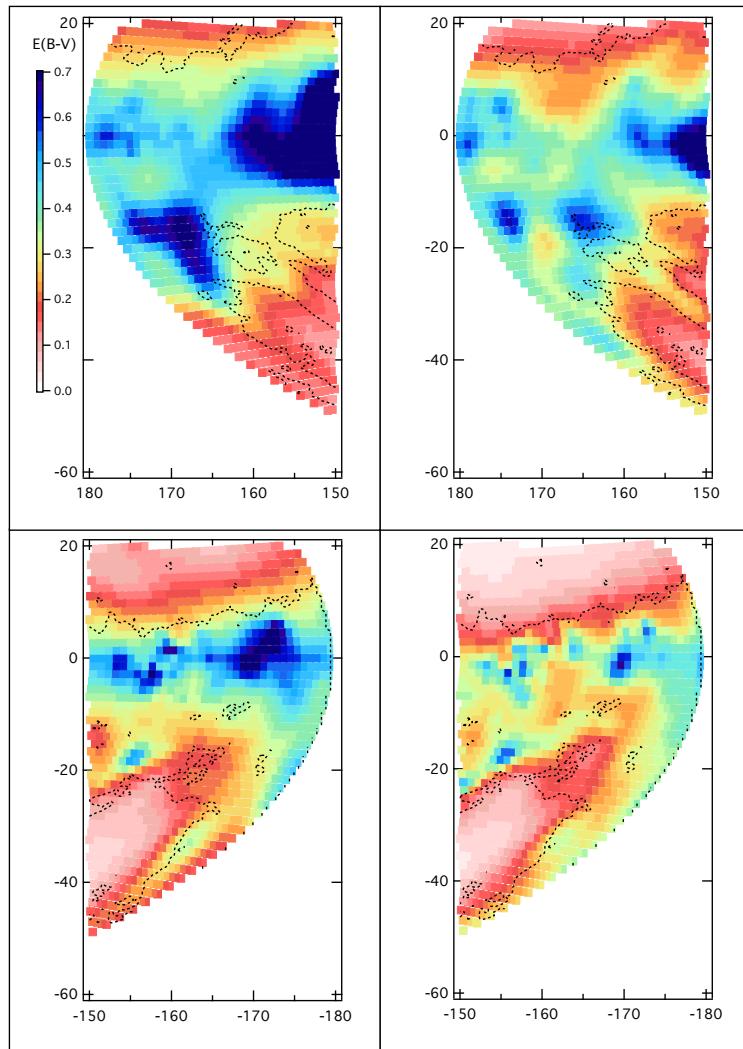


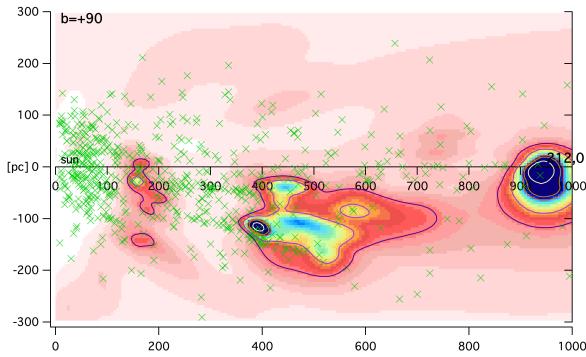
Fig. 8: Same as Fig. 7 for limited regions in the anticenter area. Left: LVV data are shown. Right: final inversion is shown. Isocontours from Schlegel et al. (1998) for  $E(B-V) = 0.025$  are superimposed as dashed black lines.

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.

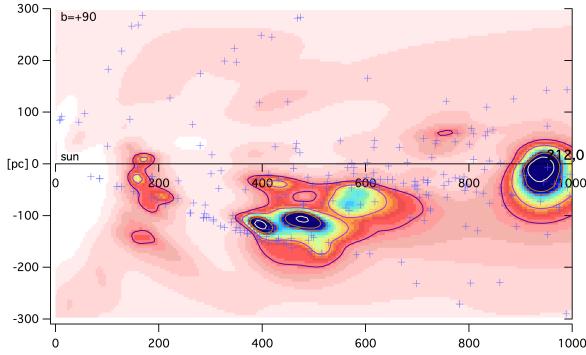
In Fig. 9, 12, and the bottom panel of Fig. 11, we superimposed the opacity distribution resulting from our D inversion and the position of the cloud cores identified by Schlafly et al. (2014) with associated error bars on their positions. For the Orion region ( galactic longitude =  $212^\circ$ ; Fig. 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 =  $158^\circ$ ; Fig 12), our map also indicates the same distances. In the case of the Cepheus region (galactic longitude =  $110^\circ$ ; 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

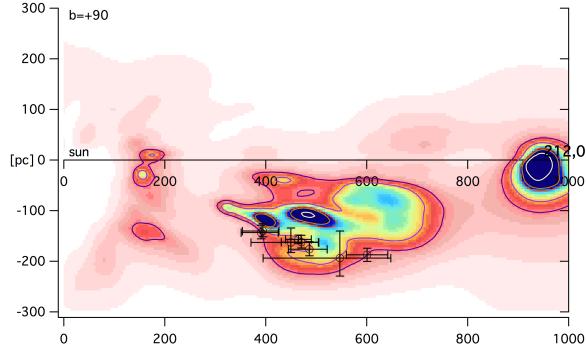
Following the work of Puspitarini et al. (2014) we searched for cavities in our 3D maps, which can be counterparts of supernova remnants (SNRs) or massive stellar winds. Such cavities are detected through their soft X-ray emission in ROSAT maps. Fig. 15 shows four vertical cuts in the 3D distribution 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  $l,b = (-30^\circ, +14^\circ)$ ,  $(-18^\circ, +18^\circ)$ ,  $(-5^\circ, +9^\circ)$ , and  $(-8^\circ, -10^\circ)$ . 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 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  $l,b,d=(-18^\circ, +18^\circ, 125$  pc) and  $(-30^\circ, +14^\circ, 125$  pc) are within  $3^\circ$  and  $7^\circ$ , respectively, from the directions of the estimated sites of the two most recent



(a) B opacity distribution



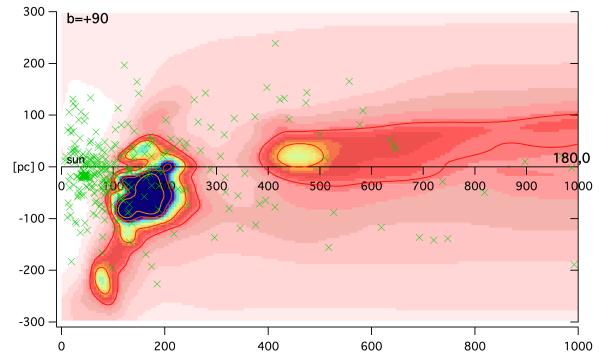
(b) C opacity distribution



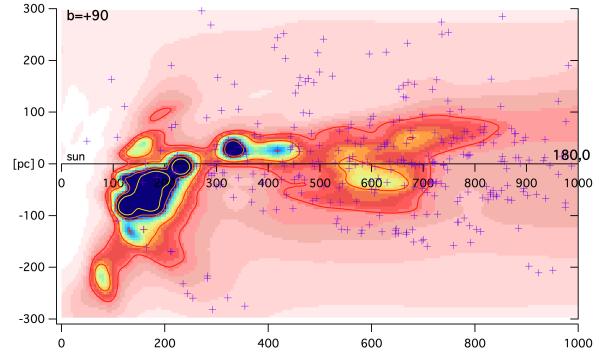
(c) D opacity distribution

Fig. 9: Comparison between the three B-C-D distributions (see Sect. 4 and Fig. 4) for a vertical plane along  $l=212^\circ$ . 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  $\rho_0 = 0.0002, 0.00035, 0.00076, 0.001$ , and  $0.002 \text{ mag.pc}^{-1}$ . 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).

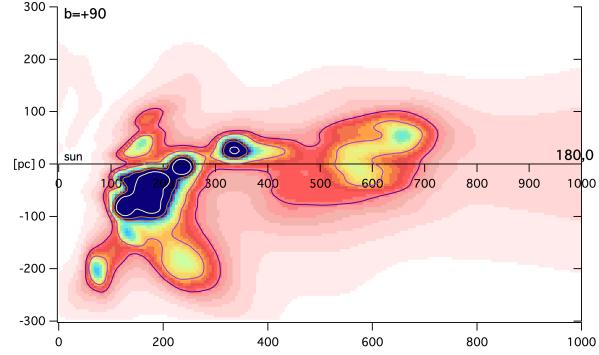
SNRs ( $l,b = (-33^\circ, +11^\circ)$  and  $(-17^\circ, +25^\circ)$ , Breitschwerdt et al. (2016)), and at similar distances. These authors have used the signature of  $^{60}\text{Fe}$  in deep-sea crusts as a sign of recent (about 2.2 million years ago) explosions of supernovae in the solar neighborhood, and, based on proper motions and ages of the young



(a) B opacity distribution



(b) C opacity distribution



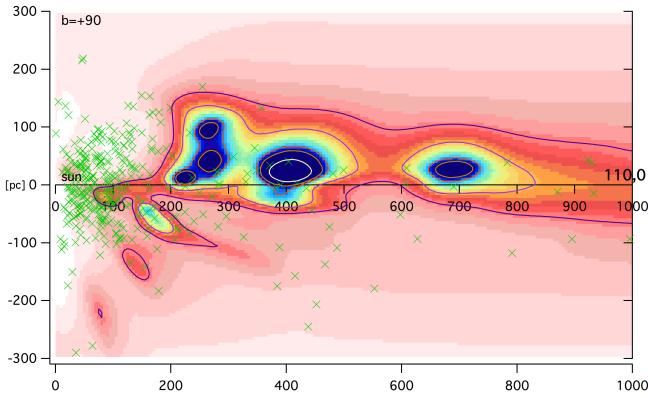
(c) D opacity distribution

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

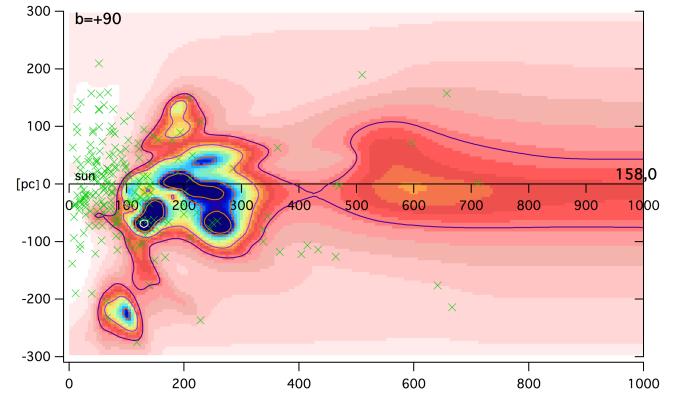
stellar groups, they computed the most probable present locations of the latest explosions. Given the uncertainties associated with their statistical modeling, the small differences between our cavities 1 and 2 and their results is negligible, and these identifications reinforce their model and the  $^{60}\text{Fe}$  measurements and interpretations.

## 6. Conclusions and perspectives

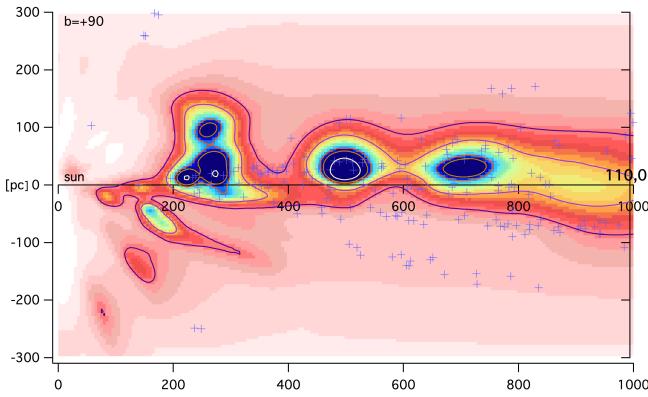
We updated and improved our previous 3D maps of the local interstellar dust, based on the Bayesian inversion of photometry-based color excess measurements, in several ways. Our goal is to analyze the influence of the use of new Gaia parallaxes and to test two additional techniques aimed at providing more accurate 3D maps. First, the merging of photometric color excesses



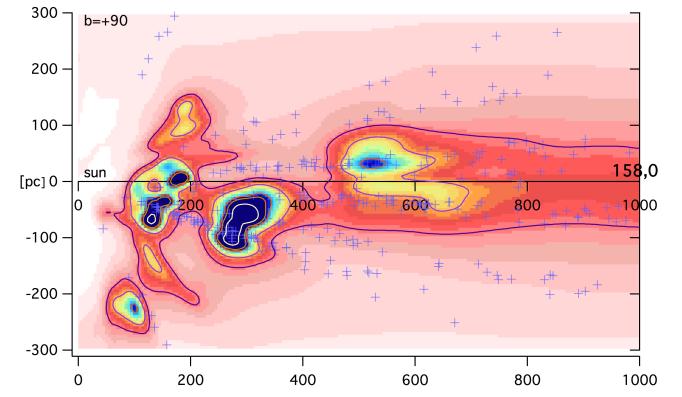
(a) B opacity distribution



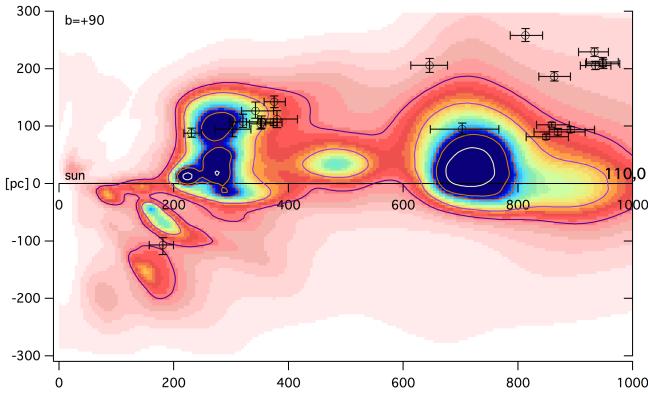
(a) B opacity distribution



(b) C opacity distribution



(b) C opacity distribution

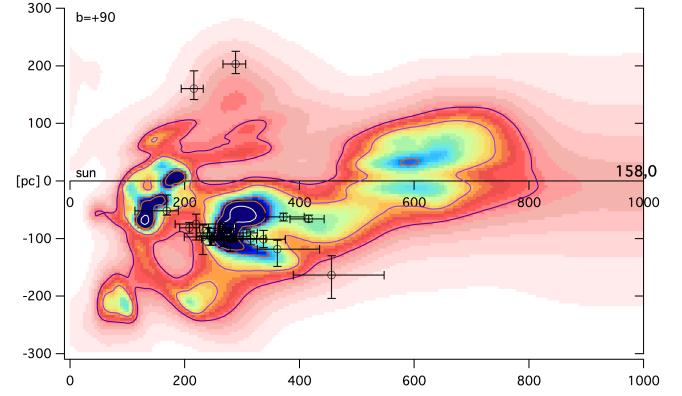


(c) D opacity distribution

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

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  $\simeq 5,100$  Hipparcos and  $\simeq 17,400$  photometric distance measurements with  $\simeq 18,700$  new Gaia-DR1 available parallactic distances and  $\simeq$  only 4,700 remaining photometric distances. In the resulting maps some elongated struc-



(c) D opacity distribution

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

tures have been suppressed, which demonstrates that such features were artefacts owing to the lower accuracy of former 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  $\simeq 5,000$  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

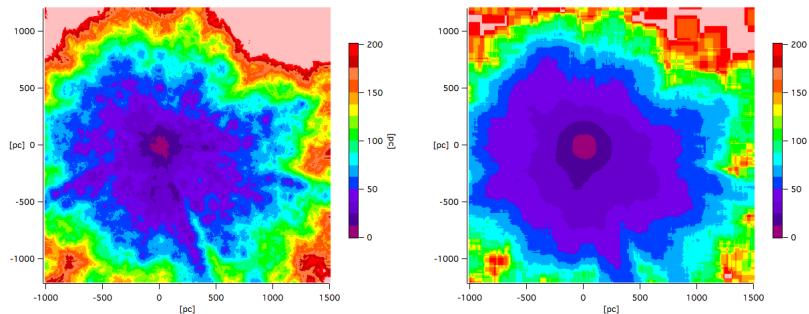


Fig. 13: Error in the distance due to the limited target space density, in the Galactic plane, as estimated from the radii of the smallest spheres that contain at least 10 targets (left figure) or from the target density in a cube of side 200 pc (right). The pale pink color corresponds to errors larger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

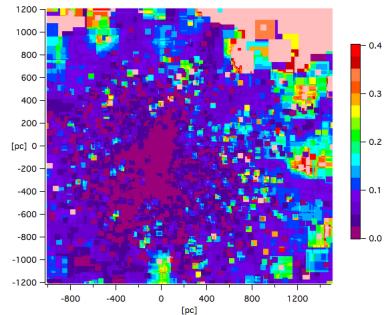


Fig. 14: Estimated error on the integrated color excess for the Galactic plane computed as the mean difference between data and model in a restricted volume (see Sect. 5.1). The pale pink color corresponds to errors larger than 0.4 mag.

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 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  $30^\circ$ . Our final map is available at [stilism.obspm.fr](http://stilism.obspm.fr) in its entirety and online tools allow one to

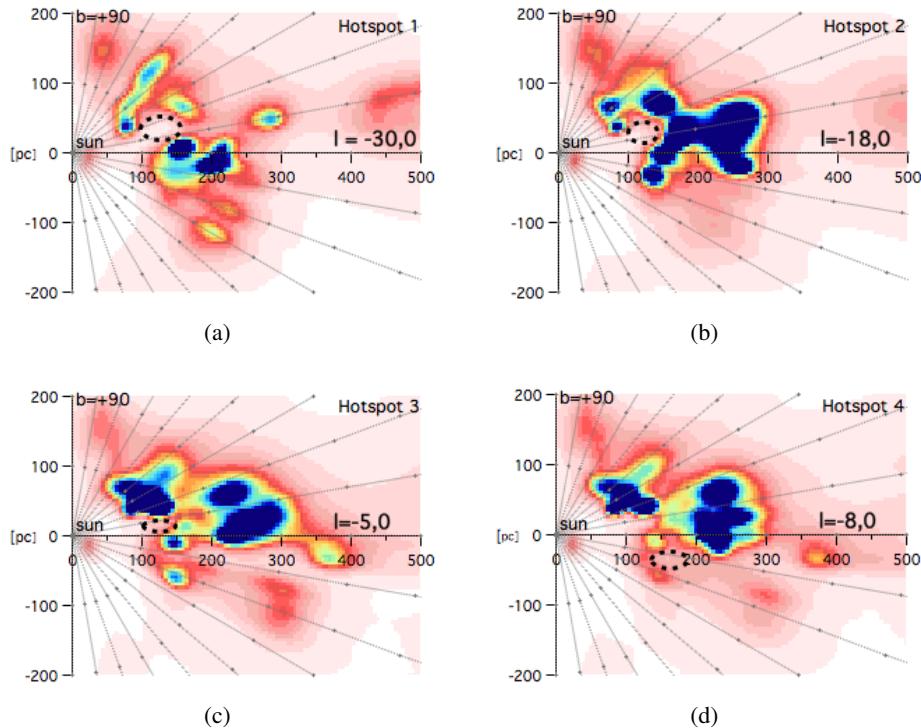


Fig. 15: Four vertical planar cuts in the 3D opacity maps in planes containing the ROSAT 3/4 keV X-ray bright spots labeled 1 to 4 in Puspitarini et al. (2014) and here as a) to d); the black dashed ellipses show the likely locations of the corresponding SNRs; the latitudinal grid is every 10 deg.

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 Ferrière 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.

- Monreal-Ibero, A. & Lallement, R. 2015, in Memorie della Società 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](https://arxiv.org/abs/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

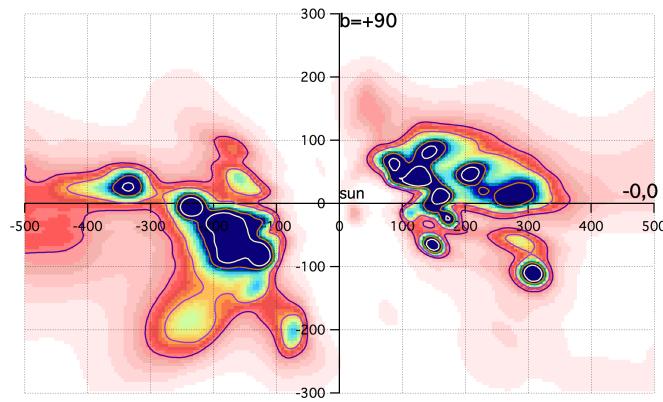


Fig. 16: Vertical cut in the 3D opacity distribution for D inversion along the meridian plane  $l=0-180^\circ$ . Color scale and iso-contours as in Fig 9.

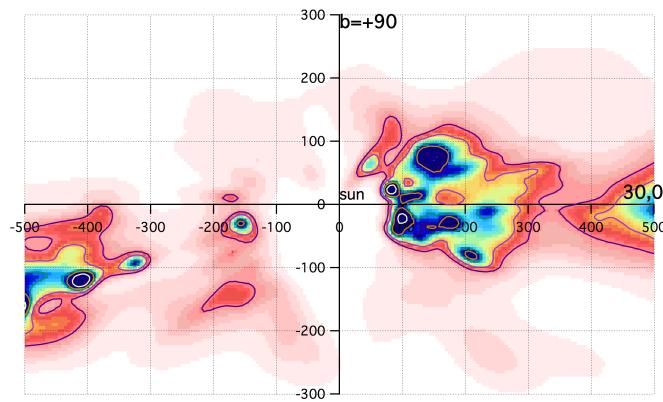


Fig. 17: Same as Fig. 16 for the vertical plane along  $l=30-210^\circ$

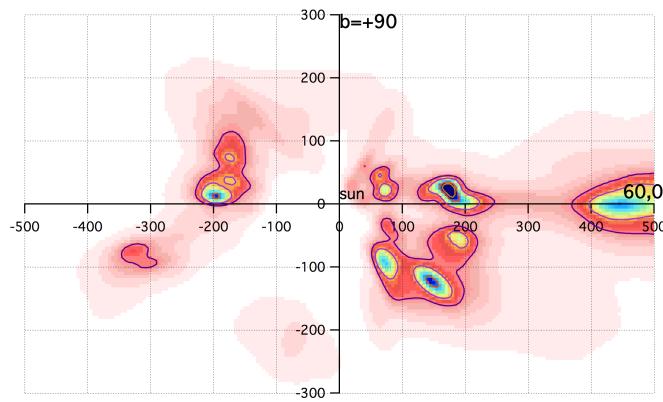


Fig. 18: Same as Fig. 16 for the vertical plane  $l=60-240^\circ$

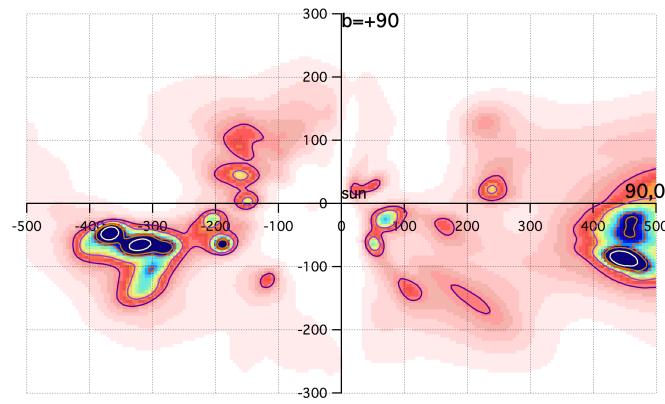


Fig. 19: Same as Fig. 16 for the vertical plane  $l=90-270^\circ$  (the rotation plane)

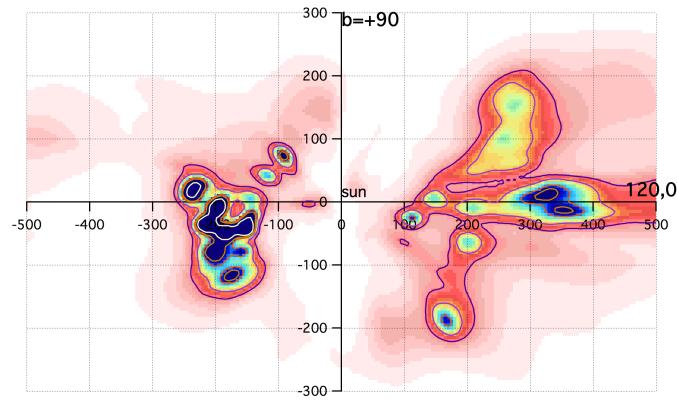


Fig. 20: Same as Fig. 16 for the vertical plane  $l=120-300^\circ$ .

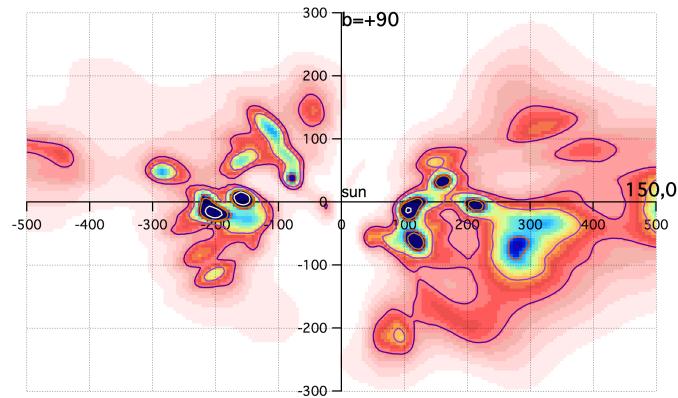


Fig. 21: Same as Fig. 16 for the vertical plane  $l=150-330^\circ$ .