Autonomous Gaussian decomposition of the Galactic Ring Survey

II. The Galactic distribution of ^{13}CO[*]

M. Riener[1],[**], J. Kainulainen[2],[1], J. D. Henshaw[1], and H. Beuther[1]

[1] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
e-mail: riener@mpia.de
[2] Chalmers University of Technology, Department of Space, Earth and Environment, 412 93 Gothenburg, Sweden

ABSTRACT

Knowledge about the distribution of CO emission in the Milky Way is essential to understanding the impact of the Galactic environment on the formation and evolution of structures in the interstellar medium. However, our current insight as to the fraction of CO in the spiral arm and interarm regions is still limited by large uncertainties in assumed rotation curve models or distance determination techniques. In this work we use the Bayesian approach from Reid et al. (2016, ApJ, 823, 77; 2019, ApJ, 885, 131), which is based on our most precise knowledge at present about the structure and kinematics of the Milky Way, to obtain the current best assessment of the Galactic distribution of ^{13}CO from the Galactic Ring Survey. We performed two different distance estimates that either included (Run A) or excluded (Run B) a model for Galactic features, such as spiral arms or spurs. We also included a prior for the solution of the kinematic distance ambiguity that was determined from a compilation of literature distances and an assumed size-linewidth relationship. Even though the two distance runs show strong differences due to the prior for Galactic features for Run A and larger uncertainties due to kinematic distances in Run B, the majority of their distance results are consistent with each other within the uncertainties. We find that the fraction of ^{13}CO emission associated with spiral arm features ranges from 76 to 84% between the two distance runs. The vertical distribution of the gas is concentrated around the Galactic midplane, showing full-width at half-maximum values of ∼75 pc. We do not find any significant difference between gas emission properties associated with spiral arm and interarm features. In particular, the distribution of velocity dispersion values of gas emission in spurs and spiral arms is very similar. We detect a trend of higher velocity dispersion values with increasing heliocentric distance, which we, however, attribute to beam averaging effects caused by differences in spatial resolution. We argue that the true distribution of the gas emission is likely more similar to a combination of the two distance results discussed, and we highlight the importance of using complementary distance estimations to safeguard against the pitfalls of any single approach. We conclude that the methodology presented in this work is a promising way to determine distances to gas emission features in Galactic plane surveys.

Key words. methods: data analysis – radio lines: ISM – ISM: kinematics and dynamics – ISM: lines and bands – Galaxy: structure – Galaxy: kinematics and dynamics

1. Introduction

A long-standing problem in astrophysics is how molecular gas, in particular the isotopologues of carbon monoxide (CO), is distributed in the Milky Way (for reviews see e.g. Combes 1991; Heyer & Dame 2015). Knowledge about the location of the molecular gas in our Galaxy is essential to answer important open questions in interstellar medium (ISM) research, such as the impact and importance of different Galactic environments (e.g. spiral arm and interarm regions) on star formation and the origin and evolution of ISM structures.

Addressing these scientific questions in an unbiased and systematic way requires the detailed analysis of CO emission line surveys of the Galactic plane, which usually consist of hundreds of thousands to millions of spectra (e.g. Dame et al. 2001;

Jackson et al. 2006; Umemoto et al. 2017; Su et al. 2019). Many studies have focussed on extracting structures from these surveys, which have been compiled into catalogues of physical objects such as molecular clouds and clumps (e.g. Solomon et al. 1987; Rathborne et al. 2009; Rice et al. 2016; Miville-Deschênes et al. 2017; Rigby et al. 2019). Alternative approaches (e.g. Sawada et al. 2012; Roman-Duval et al. 2016; Riener et al. 2020) have focussed on an analysis of these data sets without a segmentation into pre-defined physical objects, which bypasses the step of classifying the fundamentally continuous nature of the ISM into discrete objects. However, both approaches require the determination of distances to the gas emission to permit a homogeneous analysis and comparison across different Galactic environments that accounts for differences in spatial resolution introduced by our vantage point inside the Galactic disc.

Molecular gas observations entail additional information about the radial velocity of the gas emission along the line of sight – the velocity difference to the local standard of rest or $v_{\rm LSR}$, which many Galactic plane studies have used in conjunction with an assumed model for the rotation curve of our Galaxy to estimate distances via the kinematic distance (KD) method

(e.g. Dame et al. 1986; Roman-Duval et al. 2009, 2016; Elia et al. 2017; Miville-Deschênes et al. 2017). However, the KD method is based on a model for the Galactic rotation curve and thus assumes the gas to be in rotational equilibrium, whereas the Milky Way is characterised by streaming motions (e.g. Combes 1991; Reid et al. 2009, 2019; López-Corredoira & Sylos Labini 2019). Especially around spiral arms, we expect strong deviations from purely circular rotation that can reach values of up to $10\,km\,s^{-1}$ and can lead to large kinematic distance uncertainties of up to $2-3$ kpc (e.g Burton 1971; Liszt & Burton 1981; Stark & Brand 1989; Gómez 2006; Reid et al. 2009; Ramón-Fox & Bonnell 2018). Moreover, the non-axisymmetric potential introduced by the Galactic bar causes large non-circular motions in the gas within Galactocentric distances of $\sim$5 kpc (e.g. Reid et al. 2019). Towards the Galactic centre and close to the Sun, the observed gas velocity also has almost no radial component, which yields large distance uncertainties (see e.g. the kinematic distance avoidance zones in Fig. 1 of Ellsworth-Bowers et al. 2015).

Another big problem of the KD method is that it always yields two possible distance solutions in the inner Galaxy (i.e. for emission within the solar orbit), which has been termed the kinematic distance ambiguity (KDA). Additional information is needed to resolve the KDA and previous studies have utilised an abundance of methods to solve for it by using, for example, H I self absorption (e.g. Jackson et al. 2002; Anderson & Bania 2009; Roman-Duval et al. 2009; Wienen et al. 2012; Urquhart et al. 2018), H I absorption against ultracompact H II regions (Fish et al. 2003), H I emission and absorption (e.g. Anderson & Bania 2009), association with infrared dark clouds (e.g. Simon et al. 2006a; Duarte-Cabral, in prep.), or the use of scaling relationships (e.g. Rice et al. 2016; Miville-Deschênes et al. 2017).

Notwithstanding all these issues in establishing reliable distances, many studies of molecular clouds extracted from $^{12}CO\,(1-0)$ or $^{13}CO\,(1-0)$ surveys tried to identify their position within the Galaxy (Combes 1991; Heyer & Dame 2015; Rice et al. 2016; Miville-Deschênes et al. 2017) and found large variations in how well the clouds trace the gaseous spiral arm structure and the fraction of clouds located in interarm regions. In terms of star formation, we expect an enhancement in spiral arms due to effects of gravitational instabilities, cloud collisions, and orbit crowding (e.g. Elmegreen 2009). Even though sites of massive star formation seem to be predominantly associated with spiral arms (e.g. Urquhart et al. 2018), recent studies have found no significant impact of Galactic structure on the clump or star formation efficiency of dense clumps (e.g. Moore et al. 2012; Eden et al. 2013, 2015; Ragan et al. 2016, 2018), or the physical properties of filaments (Schisano et al. 2020) and molecular clumps (Rigby et al. 2019). However, the last study reported differences in the linewidths between clumps located in interarm and spiral arm structures. A recent study by Wang et al. (2020) also found clear differences in the ratio of atomic to molecular gas in arm and interarm regions.

We note that many of these studies used different Galactic rotation curve models and rotation parameters (e.g. Clemens 1985; Brand & Blitz 1993; Reid et al. 2014) in their distance estimation; also different spiral arm models (e.g. Taylor & Cordes 1993; Vallee 1995; Reid et al. 2014) have been used as a comparison. The exact number and precise locations of the spiral arms in our Milky Way is debated, even though recent years have seen huge progress in our understanding of Galactic structure (see e.g. the recent review from Xu et al. 2018). In particular, advances have been made by precise parallax measurements

of masers associated with high-mass star-forming regions (e.g. Reid et al. 2009, 2014, 2019; VERA Collaboration 2020). New distance estimation approaches have emerged that use a Bayesian approach to combine these parallax measurements with additional information from CO and H I surveys (Reid et al. 2016, 2019), which has already been used in the distance estimation to molecular clouds and clumps (e.g. Rice et al. 2016; Urquhart et al. 2018; Rigby et al. 2019; Duarte-Cabral, in prep.).

Our main motivation with this work is to use the currently most precise model for the structure and rotation curve of the Milky Way from Reid et al. (2019) in conjunction with the Bayesian approach presented in Reid et al. (2016, 2019) to analyse the distribution of molecular gas within the Galactic disc. With the distance results we further can discuss variations of the gas emission properties with Galactic environment or Galactocentric distance. By using additional priors based on literature resolutions of the KDA for molecular clouds and clumps and considerations based on a size-linewidth relationship, we derive distance estimates to all Gaussian components we fitted to the data set of a large $^{13}CO\,(1-0)$ Galactic plane survey in the first quadrant (Riener et al. 2020). In this work we present the results of two distance runs, one including and one excluding a prior for Galactic features. This approach allows us to determine lower and upper limits for the fraction of emission within spiral arm and interarm locations, and enables us to discuss the robustness of our results in terms of how much the gas emission varies with Galactocentric distance and Galactic features.

2. Data and methods

2.1. Gaussian decomposition of the GRS

In this work we use the Gaussian decomposition results of the entire Boston University-Five College Radio Astronomy Observatory Galactic Ring Survey (GRS; Jackson et al. 2006) data set as presented in Riener et al. (2020)[1]. The GRS is a $^{13}CO\,(1-0)$ emission line survey that spans ranges[2] in Galactic longitude, Galactic latitude, and velocity of $14° < \ell < 55.7°$, $|b| < 1.1°$, and $-5 < v_{LSR} < 135\,km\,s^{-1}$. The GRS consists of about 2.28 million spectra; the data has an angular resolution of $46''$, a pixel sampling of $22''$, and a spectral resolution of $0.21\,km\,s^{-1}$. Riener et al. (2020) used the fully automated Gaussian decomposition package GAUSSPY+ (Riener et al. 2019)[3] to fit all GRS spectra in their native spatial and spectral resolution, which resulted in about 4.65 million velocity fit components. They estimate that the decomposition was able to recover about 87.5% of the flux from the GRS data set, with the remaining fraction of flux being due to diffuse emission or spectra with elevated noise levels that made the extraction of signal very challenging. Riener et al. (2020) made the entire decomposition results available, and also provide quality metrics (such as the number of strongly blended components) for the fit results[1].

2.2. Bayesian distance calculator

For the distance estimation we used the Bayesian distance calculator (BDC) tool (Reid et al. 2016, 2019) that was designed for the distance calculation of spiral arm sources. For a given (ℓ, b, v_{LSR}) coordinate, the BDC calculates a distance probability

---

[1] http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/633/A14
[2] For $\ell < 18°$ the latitude coverage is incomplete and for $\ell > 40°$ the velocity range is limited to -5 to $85\,km\,s^{-1}$.
[3] https://ascl.net/1907.020

density function (PDF) based on multiple priors that can be selected by the user. In the current version of the BDC (v2.4, Reid et al. 2019) this includes the following priors:

KD: the kinematic distance;

GL: the Galactic latitude value or displacement from the Galactic midplane;

PS: the proximity to parallax sources; these are high mass star-forming regions, whose trigonometric parallaxes have been determined as part of the Bar and Spiral Structure Legacy (BeSSeL) Survey[4] and the Japanese VLBI Exploration of Radio Astrometry (VERA)[5];

SA: the proximity to features from an assumed spiral arm model; these features (such as spiral arms and spurs) have been inferred from combining information from the parallax sources with archival CO and H I Galactic plane surveys;

PM: the proper motion of the source.

The BDC allows users to set weights for these priors (P_{KD}, P_{GL}, P_{PS}, P_{SA}, P_{PM}) that can range from 0 to 1. If the weight of a prior is set to 0 it is neglected in the distance estimation. In the default settings of the BDC, all prior weights are set to 0.85, with the exception of P_{PS}, which receives a lower weight of 0.15. In addition, users can also supply a prior for the resolution of the KDA, that means they can provide information on whether the source location is expected to be on the near or far side of the Galactic disc. The weight P_{far} for this prior is by default set to 0.5, so that the near and far solutions of the KD prior receive equal weight. In this work we introduce two additional priors based on literature solutions of the KDA (Sect. 3.2) and a size-linewidth relationship (Sect. 3.3) that inform the P_{far} value for individual sources.

3. Distance estimation

Here we describe our method for the distance estimation. Since the BDC was designed as a distance estimator for spiral arm sources, its default settings have an inherent bias of associating (ℓ, b, v_{LSR}) coordinates with the assumed spiral arm model (see Fig. 6 in Reid et al. 2016). To better characterise the impact of this bias, we decided to perform and compare distance calculations with and without the SA prior, which we refer to further on as Run A and Run B, respectively. The two distance results represent two very useful extremes in the parameter space of the distance estimation. In one case we intentionally bias the emission towards our currently best knowledge of spiral arm features or overdensities of H I and CO, which we would expect to also coincide with overdensities in ^{13}CO. In the other case we obtain a picture that is unbiased by an assumed spiral arm model, but is much more dominated by the chosen Galactic rotation curve and suffers more from kinematic distance uncertainties and errors introduced by streaming motions. In the following, we present our settings for these two BDC runs, detail how we incorporated additional prior information based on literature KDA information and the fitted linewidths, and discuss how we choose the final distance results.

### 3.1. Modification of the BDC and setting of prior weights

For the distance calculation we use the most recent version of the BDC tool (Sect. 2.2) with the default Galactic rotation curve parameters as determined by Reid et al. (2019); Table 1 lists the

---

[4] http://bessel.vlbi-astrometry.org
[5] http://veraserver.mtk.nao.ac.jp

Table 1. Galactic rotation curve parameters used in the BDC runs.

Parameter	Value	Parameter	Value
R_0 [kpc]	8.15 ± 0.15	$U_\odot$ [km s^{-1}]	10.6
Θ_0 (a$_1$) [km s^{-1}]	236 ± 7	$V_\odot$ [km s^{-1}]	10.7
a$_2$	0.96	$W_\odot$ [km s^{-1}]	7.6
a$_3$	1.62		
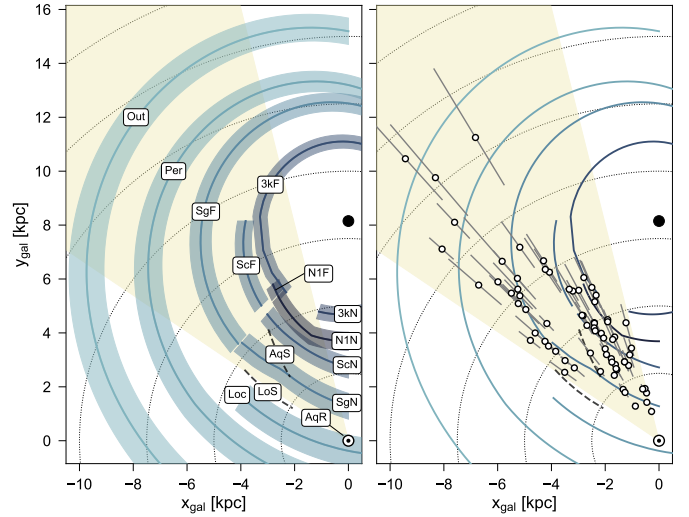
Fig. 1. Face-on view of the 1st Galactic quadrant showing the GRS coverage (beige shaded area), and the positions of the Sun (Sun symbol) and Galactic centre (black dot). Dotted black lines indicate distances to the Sun in 2.5 kpc intervals. *Left panel*: positions of Galactic features as determined by Reid et al. (2019). Spiral arms are shown with green solid lines and the shaded green areas indicate ~3.3σ widths of the arms. Spurs are shown with dashed black lines. The features are labelled as follows: 3 kpc far arm (3kF), Aquila Rift (AqR), Aquila Spur (AqS), Local arm (Loc), Local spur (LoS), Norma 1st quadrant near and far portions (N1N, N1F), Outer (Out), Perseus (Per), Scutum near and far portions (ScN, ScF), Sagittarius near and far portions (SgN, SgF). *Right panel*: position and uncertainties in distance of 71 maser sources overlapping with the GRS coverage. Spiral arm and spur positions are the same as in the *left panel*.

most important parameters. R_0 denotes the distance to the Galactic centre and Θ_0 (or a$_1$) is the estimated circular rotation speed at the position of the Sun; both values are in very good agreement with independent observations and measurements (Gravity Collaboration 2019; Kawata et al. 2019). The a$_1$, a$_2$, and a$_3$ values are parameters used in the 'universal' form of the rotation curve from Persic et al. (1996) that was adopted in the BDC (Reid et al. 2014, 2016, 2019). $U_\odot$, $V_\odot$, and $W_\odot$ denote the solar peculiar motions towards the Galactic centre, in the direction of Galactic rotation, and towards the north Galactic pole, respectively.

The BDC results are strongly influenced by the choice of the spiral arm model and the included parallax measurements to maser sources. It is therefore instructive to discuss and illustrate how many spiral arm features and maser sources overlap with the GRS coverage as these are decisive factors in the distance estimation. The left panel in Fig. 1 shows Galactic features, such as spiral arms and spurs, that were inferred from distance measurements to maser parallax sources and archival CO and H I surveys (Reid et al. 2016, 2019) and are used as spiral arm model for the

SA prior. The width of the spiral arm features shows the approximate extent for associations of data points with these features. In the right panel we show the position and distance uncertainties of 71 maser sources from Reid et al. (2019) that are overlapping with the spatial and spectral coverage of the GRS. These maser sources all have parallax uncertainties <20%, which is the BDC default requirement for the inclusion of parallax sources for the PS prior. The PS prior for GRS sources is determined by association with one or more of these parallax sources (see Reid et al. 2016 for how this association is performed). Most of the measured maser sources are associated with the Scutum and Sagittarius spiral arms, thus leading to an additional emphasis of these features in the distance determination.

Since we only have access to the radial velocity component of the gas, we do not use the PM prior that would require knowledge about the proper motion of the gas. In the following, we motivate and explain the chosen settings for our two BDC runs.

For Run A we used all priors (KD, GL, PS, SA). We used the default weights for P_{KD} and P_{PS}. In test runs of the BDC, we found that the default weight of 0.85 for P_{SA} led to a strong domination of the spiral arm model (see Appendix C.2) compared to the remaining priors. We thus opted to reduce P_{SA} to 0.5, which led to a more balanced ratio between the priors in our tests. Since in the default settings of the BDC the priors for the spiral arm model and the Galactic latitude are combined, we also set P_{GL} to 0.5 to keep the ratio between these priors intact.

For Run B we did not use the priors for the proper motion and the spiral arm model. By default, the BDC combines the SA and GL priors, which means that setting $P_{SA} = 0$ has the effect of also setting $P_{GL} = 0$. As the Galactic latitude information contains important prior information for the distances, we slightly modified the BDC source code so that we could use the P_{GL} prior without using the P_{SA} prior. However, we found that in this case the default settings of $P_{GL} = 0.85$ could yield a strong bias towards the far KD solution. To reduce this bias, we opted to decrease P_{GL} to a value of 0.5, which yielded a more balanced ratio between the priors in our tests.

In addition to these settings, we include priors that incorporate literature KDA resolutions and fold in information from the fitted linewidth in both BDC runs. These additional priors are described in more detail in the next two sections.

3.2. Prior for the kinematic distance ambiguity

For all sources located within the solar circle the KD prior yields two possible distance solutions (called the near and far distances). However, over recent years many works have already solved the KDA for many objects such as molecular clouds and clumps that overlap with the GRS coverage. Many of these studies even used the GRS data set directly in their distance estimation. To take advantage of these previous works, we implemented a new scheme that uses these literature KDA solutions to inform the P_{far} prior of the BDC, which results in a preference for the near or far distance solution. In Appendix A we list all literature KDA solutions that we incorporated in our method and describe in detail how we use this information to determine the P_{far} weight for individual sources. In total, this prior was used in the distance estimation of about 30% of the ^{13}CO fit components (see Appendix C.3 for more details). In Appendix A we also discuss the performance of this prior; we found that its inclusion leads to a significant increase in consistency of the BDC distance results with the reported literature distances.

We illustrate the effect of the KDA prior on the distance estimation with an example in Fig. 2, which shows the resulting distance PDFs for the individual priors. In panel a the spiral arm prior (P_{SA}, in red) was switched off and no KDA prior was supplied (i.e. $P_{far} = 0.5$), so the only remaining contributions to the combined distance PDF come from the priors for the kinematic distances (P_{KD}, in blue), the association with parallax sources (P_{PS}, in green), and the Galactic latitude (P_{GL}, in orange). Since the source is located close to the centre of the Galaxy, the two peaks of the kinematic distance PDF are not Gaussian-shaped, but were down-weighted to reflect expected large peculiar motions near the Galactic bar (Reid et al. 2019). Distances are estimated by fitting Gaussians to the peaks of the combined distance PDF (in black); the most likely distance value corresponds to the Gaussian component with the highest integrated probability density, so the highest peak of the combined distance PDF need not result in the most likely distance estimate. The distance uncertainty is given by the standard deviation of the Gaussian fit component.

With no associated parallax sources or conclusive latitude information the two distance solutions would have corresponded to the two peaks of the kinematic distance PDF and would have received the same probability (50%). In our case, the prior incorporating the Galactic latitude position favours the far distance, but associated parallax sources shift the balance towards the near kinematic distance solution, yielding a most likely distance value for the source of ~3.4 kpc. We note that even though the SA prior is switched off the BDC still gives the information of whether the distance results do overlap with locations of spiral arm and interarm features; the extent for such associations is indicated with the red-shaded areas in Fig. 2. For the example depicted in panel a, D_1 is associated with the near portion of the Scutum spiral arm, whereas D_2 corresponds to an interarm position.

If the spiral arm prior is included (panel b), the most likely distance shifts to a higher value of about 4.2 kpc[6]. Also the distance estimate with the second highest probability corresponds to a near distance solution, which illustrates the strength of the spiral arm prior.

Finally, panel c shows the effect of adding a prior for the KDA, which in our case favours the far kinematic distance solution (for this example we assume $P_{far} = 0.875$; see Appendix A for how exactly P_{far} is determined from literature KDA solutions). Setting the KDA prior has the effect of rescaling the kinematic distance PDF, which in this example shifts the most likely distance value to a far distance solution.

This example illustrated that the KDA prior can be a decisive factor for the distance estimation. However, while the P_{far} prior can give a strong preference for one of the kinematic distance solutions, we note that the combination with the other priors can still result in a different choice for the most likely distance.

3.3. Prior for the fitted linewidth

In our tests of the BDC, we noticed that sources with low v_{LSR} velocities are preferentially placed at larger distances (see Fig. 3). This effect is strongest for sources with $v_{LSR} \lesssim 5 \ km \ s^{-1}$; for sources with $v_{LSR} \lesssim 0 \ km \ s^{-1}$ the KD prior permits essentially only the far distance solution. This effect can be mitigated by the inclusion of the SA prior as sources can receive a strong association with the nearby Aquila Rift cloud complex. However, since the association with Aquila Rift is only performed over a very limited distance range, this leads to narrow high peaks in the distance PDF, which in turn yield associated Gaussian fit

---

[6] In this case, the spiral arm and Galactic latitude probabilities are by default combined.
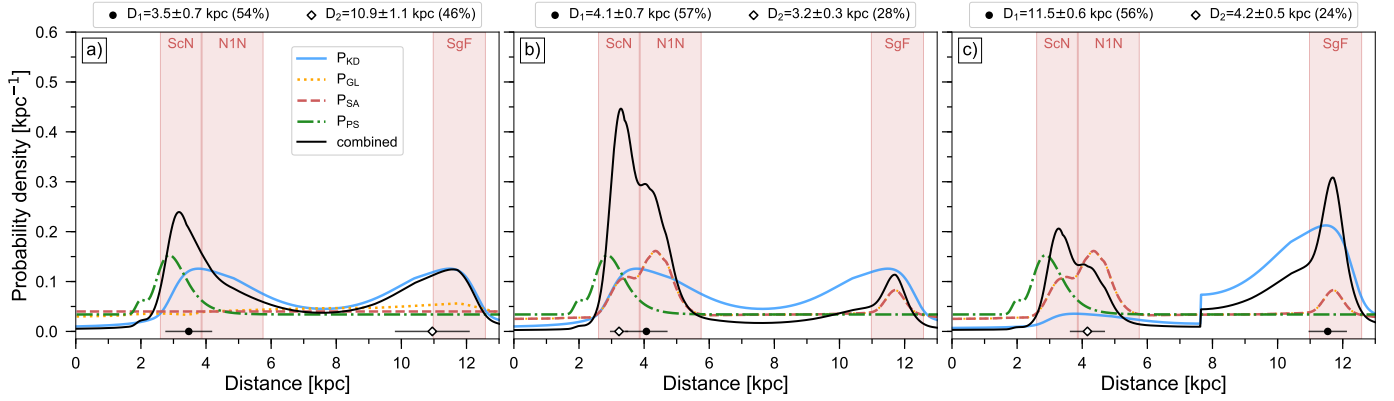
Fig. 2. Effect of BDC priors on the distance estimation for a source located at $\ell = 20.66°$, $b = -0.14°$, and $v_{LSR} = 56.2$ km s^{-1}. For *panel a* we set $P_{SA} = 0$ and $P_{far} = 0.5$, for *panel b* we set $P_{SA} = 0.5$ and $P_{far} = 0.5$, and for *panel c* we set $P_{SA} = 0.5$ and $P_{far} = 0.875$. Coloured lines show the probabilities of the priors for kinematic distances (blue solid line), latitude probability (orange dotted line), spiral arms (red dashed line), and associations with parallax sources (green dash-dotted line). The black solid line shows the combined distance PDF. The red-shaded areas show the distance ranges for which the source would be associated with individual Galactic features. Black circles and white diamonds indicate the first and second choice for the distance estimates, with the horizontal bars showing their corresponding uncertainties. The boxes above the panels list the resulting distances, their uncertainties, and estimated probabilities.
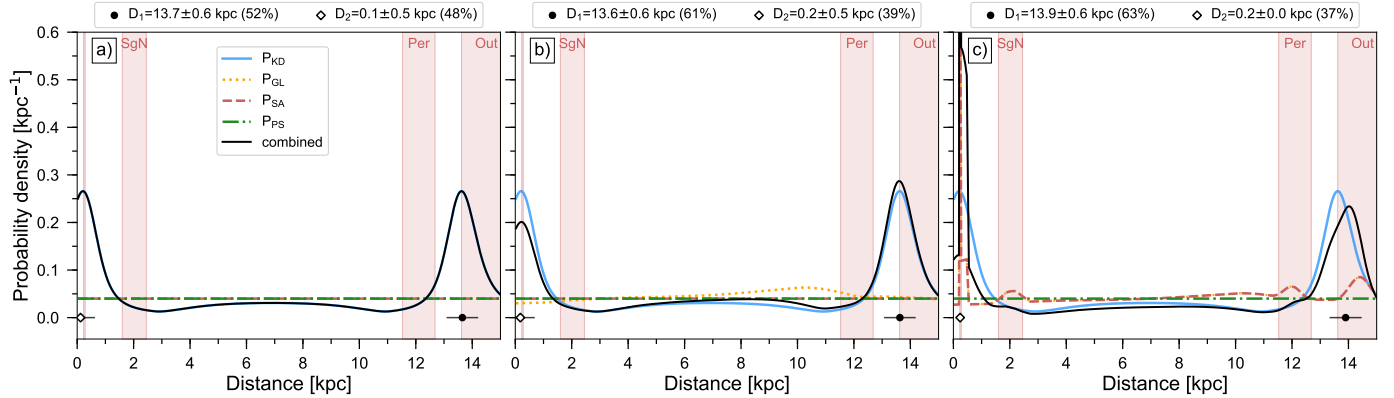
Fig. 3. BDC examples for a source located at $\ell = 32°$, $b = 0°$, and $v_{LSR} = 5$ km s^{-1}, illustrating how sources with low v_{LSR} velocities are biased towards the far distance solution. The panels show BDC results using the KD and PS priors (*a*), in addition to the GL (*b*) as well as the SA priors (*c*). The meaning of the lines and symbols is the same as in Fig. 2.

components with a lower integrated area than for the far distance solution (Fig. 3c). This effect thus has a large impact on our distance results, since we expect strong confusion between local emission from the solar neighbourhood and the far side of the Galactic disc at $-5 < v_{LSR} \lesssim 20$ km s^{-1} (Riener et al. 2020).

However, as suggested in Riener et al. (2020), we can try to use the velocity dispersion values of the fit components as an additional prior information for the distance calculation. Figure 4 recaps the argument put forth in Riener et al. (2020): due to averaging of bigger spatial areas at larger distances, we expect broadened lines due to, for example, sub-beam structure and velocity crowding, velocity gradients of the line centroids (either along the line of sight or in the plane of the sky), or fluctuations in the non-thermal contribution to the linewidth (e.g. due to regions with higher turbulence). The example shown in Fig. 4 highlights the effect of sub-beam structure and velocity crowding. If a region with two strongly blended velocity components is located at close distances, the individual emission peaks can be well resolved and fitted with two narrow Gaussian components (bottom centre and right panels in Fig. 4). However, if the same region is located at far distances, the individual velocity components might not be resolved, leading
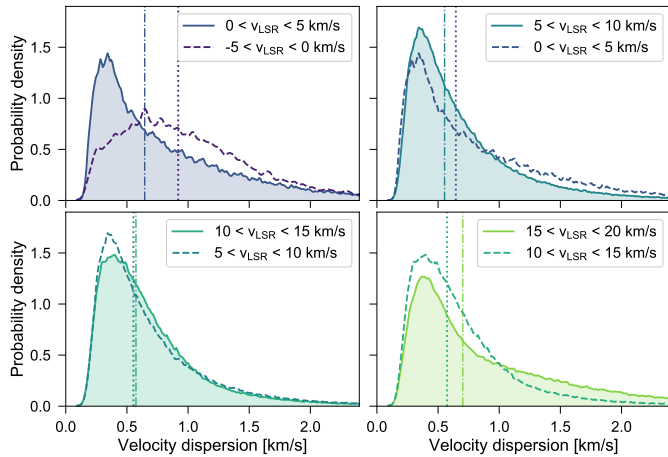
to a decomposition with a single broad Gaussian component (top centre and right panels in Fig. 4).

Given these expected differences due to beam averaging effects, it is unlikely that very narrow fitted linewidths are associated with emission at large distances. For most of the molecular gas in the GRS, the molecular gas temperatures is about 10 to 20 K, which is the typical temperature of gas at intermediate density ($\sim 10^3$ cm^{-3}) in molecular clouds. The thermal broadening of the spectral lines for these temperatures is about 0.2 to 0.3 km s^{-1}, so effectively the spectral resolution of the GRS. The physical extent of the GRS beam is ~ 0.1 pc at the distance of the Aquila Rift complex and increases to $\gtrsim 2$ pc at distances beyond the solar radius. Therefore the physical areas covered by the beam at the nearby distances of the Aquila Rift and the far distances of the Perseus and Outer arm are different by a factor of >400. Even in the case of no sub-beam structure and velocity crowding and no significant non-thermal contributions to the linewidth, expected variations in the line centroids across the beam-averaged area are enough to broaden the lines significantly (see Appendix B).

The effect of broader linewidths for emission originating at larger heliocentric distances is already noticeable in the
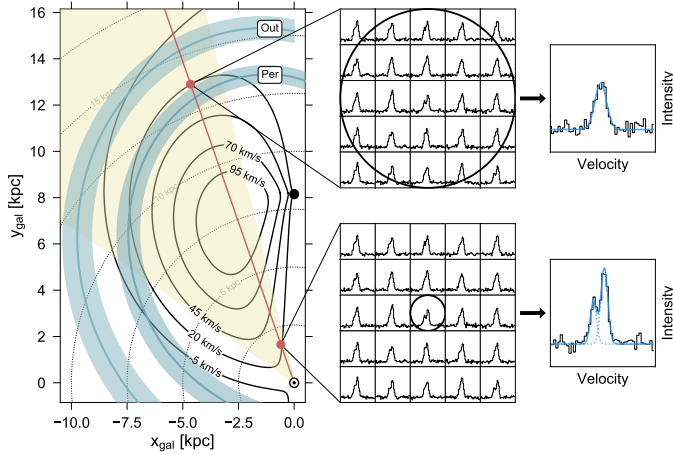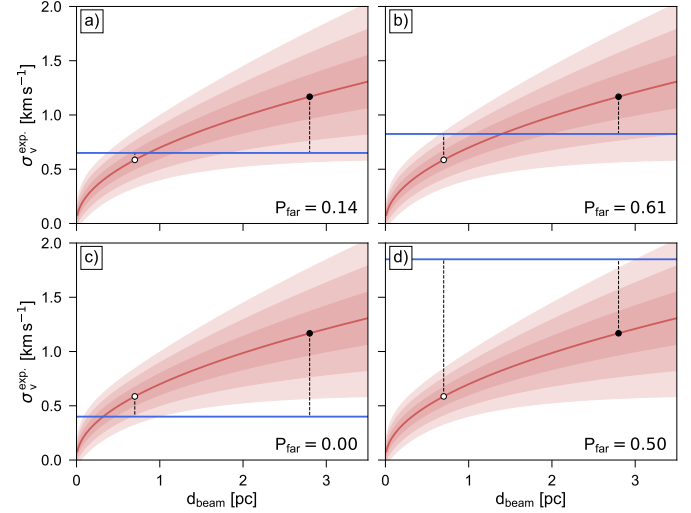
Fig. 4. Illustration of linewidth broadening caused by beam averaging. *Left panel*: same as the left panel of Fig. 1, but showing only the positions and estimated widths of the Perseus (Per) and Outer (Out) spiral arms. Black solid lines show curves of constant projected v_{LSR} values. The red line shows a random line of sight with the corresponding intersections with the $v_{LSR} = 20$ km s^{-1} curve indicated with red dots. The *centre panels* illustrate the change in spatial extent of the beam (black circle) for a region with two blended velocity components embedded at the near (bottom centre) and far (top centre) distance. The *right panels* illustrate the resulting observed spectra (black line) and Gaussian fit components (blue lines).

Fig. 6. Illustration of the distance prior based on the fitted linewidth. The red line shows the size-linewidth relation from Solomon et al. (1987) corrected for the most recent distance estimates to the Galactic centre. Red-shaded areas show the $1\tilde{\sigma}$, $2\tilde{\sigma}$, and $3\tilde{\sigma}$ intervals determined from variations of the parameters of the size-linewidth relation. The two dots show expected velocity dispersion values for corresponding physical extents of the beam for the near (white dot) and far (black dot) kinematic distance solutions. Blue horizontal lines show velocity dispersion values of fit components and dashed vertical lines indicate the relevant distances to the expected values, from which the P_{far} prior is determined. See Sect. 3.3 for more details.

to inform our decision about whether a fitted σ_v value is more likely associated with a region at near or far distances. This size-linewidth relationship has the form of:

$$\sigma_v^{exp.} = \sigma_{v,0} \cdot \left(\frac{L}{1\,\text{pc}} \right)^{\gamma}, \tag{1}$$

with $\gamma = 0.5$ and $\sigma_{v,0} = 0.7$ (corrected for the most recent distance estimates to the Galactic centre; Gravity Collaboration 2019; Reid et al. 2019). In Fig. 6 we show the expected velocity dispersion values based on this relation as a function of physical extents of the beam (d_{beam}) with the solid red line. The shaded red areas indicate $1\tilde{\sigma}$, $2\tilde{\sigma}$, and $3\tilde{\sigma}$ intervals for the size-linewidth relation assuming variations in γ and $\sigma_{v,0}$ of ± 0.1. The magnitude of these variations was motivated for consistency with results obtained from more local molecular clouds (Larson 1981; Shetty et al. 2012).

We use this size-linewidth relationship to inform the KDA prior as follows. We first calculate the physical extent of the beam (d_{beam}) for the two kinematic distance solutions that are always obtained for positive v_{LSR} values in the inner Galaxy. We then use the size-linewidth relationship to calculate the expected velocity dispersions for both d_{beam} values. Subsequently, we compare the actual fitted velocity dispersion with these expected velocity dispersion values to decide whether it is more consistent with the near or far distance value. This decision is driven by how close the fitted σ_v value is to the expected values from the near and far distances. We calculate for both distances the difference between the fitted and expected σ_v values; if the difference is within the $3\tilde{\sigma}$ interval indicated in Fig. 6 we give it the corresponding weight from a normalised Gaussian function:

Fig. 5. PDFs of fitted velocity dispersion values in different v_{LSR} ranges. Each panel compares PDFs of two different v_{LSR} ranges; the median values of the PDFs with the solid and dashed lines are indicated with the dash-dotted and dotted vertical lines, respectively.

fitted linewidths (Fig. 5). We would expect fit components in the interval of $-5 < v_{LSR} < 0$ km s^{-1} (left upper panel of Fig. 5) to predominantly originate from large distances and indeed the distribution of σ_v values is shifted towards larger values compared to similar v_{LSR} ranges between 0 and 20 km s^{-1} (remaining panels of Fig. 5). The distribution of these other ranges has a strong peak at $\sigma_v < 0.5$ km s^{-1}, consistent with the assumption that this corresponds to emission lines originating from nearby spatially resolved regions.

Having established that the fitted velocity dispersion values can contain information about the distance to the gas emission, we explain in the following how we implemented this as prior information for our distance calculation. Similar to Riener et al. (2020), we used the size-linewidth relationship established by Solomon et al. (1987) for molecular clouds in the Galactic disc

$$w_\sigma = \exp\left(-0.5 \cdot \left(\frac{\sigma_v - \sigma_v^{exp}}{\tilde{\sigma}} \right)^2 \right), \tag{2}$$

where $\tilde{\sigma}$ is the standard deviation for σ_v^{exp}. From this we calculate values for the P_{far} prior as:

$$P_{\mathrm{far}} = \tfrac{1}{2} + \tfrac{1}{2}\left(w_\sigma^{\mathrm{far}} - w_\sigma^{\mathrm{near}}\right), \tag{3}$$

with w_σ^{near} and w_σ^{far} indicating the weights for the near and far distance value. If the fitted σ_v value falls above the $3\tilde{\sigma}$ interval, we set the corresponding weight w_σ to zero. If the fitted σ_v value falls below the $3\tilde{\sigma}$ interval for the far distance but is not above the $3\tilde{\sigma}$ interval for the near distance, we automatically assume $P_{\mathrm{far}} = 0$.

We illustrate this procedure for four different cases in Fig. 6; for all these cases the kinematic distance solution and the corresponding d_{beam} values are the same but the values of the fitted σ_v values (horizontal blue lines) vary. For the first case (panel a) the σ_v value is more consistent with the near distance; we obtain $w_\sigma^{\mathrm{near}} = 0.76$ and $w_\sigma^{\mathrm{far}} = 0.04$, yielding $P_{\mathrm{far}} = 0.14$ and thus strongly favouring the near distance. In the second case (panel b) the far distance is favoured, as $w_\sigma^{\mathrm{near}} = 0.17$ and $w_\sigma^{\mathrm{far}} = 0.24$. In the third case, the fitted σ_v value is much lower than the expected σ_v value and falls below the $3\tilde{\sigma}$ range ($w_\sigma^{\mathrm{far}} = 0$); in such cases we always assume $P_{\mathrm{far}} = 0$ unless the σ_v value is above the $3\tilde{\sigma}$ interval for the near distance (in which case P_{far} would be 0.5). Finally, the last case (panel d) yields no P_{far} prior as the fitted σ_v value is much higher than the expected σ_v values for both the near ($w_\sigma^{\mathrm{near}} = 0$) and far ($w_\sigma^{\mathrm{far}} = 0$) distance. This ensures that we do not exclude the possibility that a source with high σ_v value can come from a nearby region with high non-thermal contributions to the linewidth.

Recent studies have found large dispersions of the size-linewidth relation across the Galactic disc (e.g. Heyer et al. 2009; Miville-Deschênes et al. 2017) and advocate a scaling relation that also takes the surface density into account. Moreover, especially in the inner part of the Galaxy, linewidths can be systematically higher than predicted by the size-linewidth relation, indicating that σ_v is at least partly set by Galactic environment (Shetty et al. 2012; Henshaw et al. 2016, 2019; Rice et al. 2016). We want to emphasise here that we do not use the size-linewidth relation to make conclusive decisions about the distance to a gas emission peak, but only use it as additional prior KDA information for sources with $v_{\mathrm{LSR}} < 20\ \mathrm{km\,s^{-1}}$. For sources with larger v_{LSR} values the difference between the σ_v^{exp} values for the two KD solutions gets smaller and the size-linewidth prior might bias components with narrower fitted linewidths to be preferentially placed at the near distance solution. We also do not use the σ_v prior in case the literature solutions for the KDA (Sect. 3.2) already yielded a P_{far} value $\neq 0.5$.

3.4. Choice of distance solution

The distance calculation with the BDC yields multiple alternative distance solutions with corresponding estimates of their probability. These probabilities are obtained from Gaussian fits to the combined distance PDF (Reid et al. 2016). By default, the Gaussian distance component with the highest integrated area is chosen as the most likely distance value. So even if the distance PDF shows a clear peak, this need not correspond to the selected most likely distance value. Our tests showed that this could be problematic, as very broad Gaussian components with low peak values can be selected as the most probable distance component, resulting in unlikely distance solutions (Fig. 7). For our BDC runs we found that such broad components with low peak values would be chosen as the preferred distance value in ~2.5 (Run A) and ~9% (Run B) of the distance assignments.
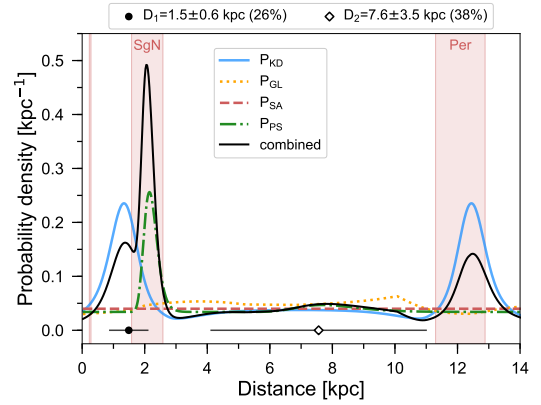
Fig. 7. Example of distance choice in case one of the distance components has a high integrated area but low peak amplitude value. The meaning of the lines and symbols is the same as in Fig. 2.

To avoid the selection of such broad components with low peak values, we adapted the choice for the most likely distance as follows. In case of two reported distance solutions (as is the default in v2.4 of the BDC), we first check whether the peaks of the individual Gaussian fit components exceed a pre-defined limit. We set this limit to 0.12, which corresponds to three times the value of a flat distance PDF[7]. If one of the distance components does not satisfy this criterion, we choose the remaining distance solution, regardless of whether its integrated area was less (see Fig. 7). If both of the distance components exceed or fail the amplitude limit, we choose the distance component with the highest assigned probability (i.e. the Gaussian fit component having the highest integrated area). In case both distance components have the same assigned probability, we choose the distance solution with the lower absolute distance error. If both components are also tied in the distance errors (as can happen if the combined distance PDF is dominated strongly by the KD prior), we choose the distance component with the lower distance value. The last two conditions were only used in ~1% of the distance choices for the two BDC runs (see Appendix C.3 for more details).

4. Galactic distribution of the gas emission

In this section we report the distance results obtained for the BDC runs including (Run A) and excluding (Run B) the prior for the spiral arm model (Sect. 3.1). In the subsections discussing the results, we always show and compare both BDC runs; if not indicated otherwise, the left- and right-hand panels depict the results of Run A and B, respectively. We first present an overview of the results and then discuss the differences in terms of the face-on and vertical distribution of the gas emission and its variation with heliocentric and Galactocentric distance. Finally, we discuss problems and biases of the two distance runs and compare our results with previous studies.

4.1. Catalogue description

With this work, we also make a catalogue of all our distance results for the GRS available. In this section we describe the

---

[7] The distance PDF is evaluated from 0 to 25 kpc. Requiring that the integrated area of a flat distance PDF is equal to unity yields a value of 0.04 for the PDF at all distances.

Table 2. Distance results.

ℓ [°] (1)	b [°] (2)	v_{LSR} [km s^{-1}] (3)	Run A: with SA prior								Run B: without SA prior							
			$d_{\odot, A}$ [kpc] (4)	$\Delta d_{\odot, A}$ [kpc] (5)	$R_{gal, A}$ [kpc] (6)	P_A (7)	Arm$_A$ (8)	$P_{far, A}$ (9)	Ref$_A$ (10)	F_A (11)	$d_{\odot, B}$ [kpc] (12)	$\Delta d_{\odot, B}$ [kpc] (13)	$R_{gal, B}$ [kpc] (14)	P_B (15)	Arm$_B$ (16)	$P_{far, B}$ (17)	Ref$_B$ (18)	F_B (19)
54.241	−1.088	45.187	4.27	0.86	6.63	1.0	LoS	0.5	–	2	4.01	0.79	6.66	1.0	LoS	0.5	–	2
54.241	−1.088	24.809	1.59	0.64	7.33	0.54	...	0.5	–	2	1.68	0.58	7.30	0.78	...	0.5	–	2
54.235	−1.088	44.955	4.27	0.87	6.63	1.0	LoS	0.5	–	2	4.0	0.8	6.66	1.0	LoS	0.5	–	2
54.235	−1.088	24.849	1.6	0.64	7.33	0.54	...	0.5	–	2	1.69	0.58	7.29	0.78	...	0.5	–	2
54.229	−1.088	24.878	1.6	0.64	7.33	0.54	...	0.5	–	2	1.69	0.58	7.29	0.78	...	0.5	–	2
54.217	−1.088	24.836	1.6	0.64	7.33	0.54	...	0.5	–	2	1.69	0.58	7.29	0.78	...	0.5	–	2
54.210	−1.088	24.755	1.59	0.64	7.33	0.54	...	0.5	–	2	1.68	0.58	7.30	0.78	...	0.5	–	2
54.204	−1.088	24.842	1.6	0.64	7.33	0.54	...	0.5	–	2	1.69	0.58	7.29	0.78	...	0.5	–	2
54.192	−1.088	24.794	1.59	0.64	7.33	0.54	...	0.5	–	2	1.68	0.58	7.30	0.78	...	0.5	–	2
54.186	−1.088	24.863	1.6	0.64	7.33	0.54	...	0.5	–	2	1.69	0.58	7.29	0.78	...	0.5	–	2
54.180	−1.088	24.721	1.59	0.64	7.33	0.54	...	0.5	–	2	1.68	0.58	7.29	0.78	...	0.5	–	2

Notes. This table is available in its entirety in electronic form at the CDS. A portion is shown here for guidance regarding its form and content.

entries of the catalogue, which includes useful parameters that help gauge the performance of the distance results.

We show a subset of the distance results in Table 2. Each row corresponds to a single Gaussian fit component; a spectrum fitted with eight Gaussian components thus occupies eight consecutive rows in the table.

Columns (1) and (2) show the Galactic coordinate values and Col. (3) gives the mean position v_{LSR} of the fit component. Columns (4–11) list the parameters of the distance results for Run A. Columns (4) and (5) give the heliocentric distance $d_{\odot}$ and its associated uncertainty $\Delta d_{\odot}$, and Col. (6) gives the Galactocentric distance R_{gal}. Columns (7) and (8) give the estimated probabilities P and the associated Galactic features (Arm) for the distance results. Columns (9) and (10) list the probability P_{far} that was used for the KDA prior and the corresponding reference for an associated literature distance ($P_{far} = 0.5$ corresponds to the default value, in case no literature sources could be associated; see Sect. 3.2 and Appendix A) for more details. Column (11) gives the flag that indicates which criterion was used for the choice of the final distance solutions (see Sect. 3.4 and Appendix C.3 for more details). Columns (12–19) list the same parameters as Cols. (4–11), but for the distance results of Run B.

4.2. Face-on view of the ^{13}CO emission

We show face-on view maps of the integrated ^{13}CO emission, the number of Gaussian fit components, and the median σ_v value in Figs. 8–10. Comparing the maps of the ^{13}CO emission (Fig. 8), we can clearly see the effect of the SA prior in the left panel, which tends to concentrate most of the emission close to the Galactic features as they are defined in the spiral arm model (Fig. 1). By neglecting the SA prior we get a distribution of the ^{13}CO emission that is much more spread out and extends over a much larger area in between the arms, which can also be clearly observed in Fig. 9. This spreading of the emission to interarm locations is to a large part due to our use of archival KDA solutions to inform the P_{far} prior. We present a comparison of the face-on map of ^{13}CO emission with and without the use of archival KDA solutions in Appendix C.2. While we find only moderate differences in the fraction of emission assigned to interarm locations, the distribution of the gas emission itself changes significantly.
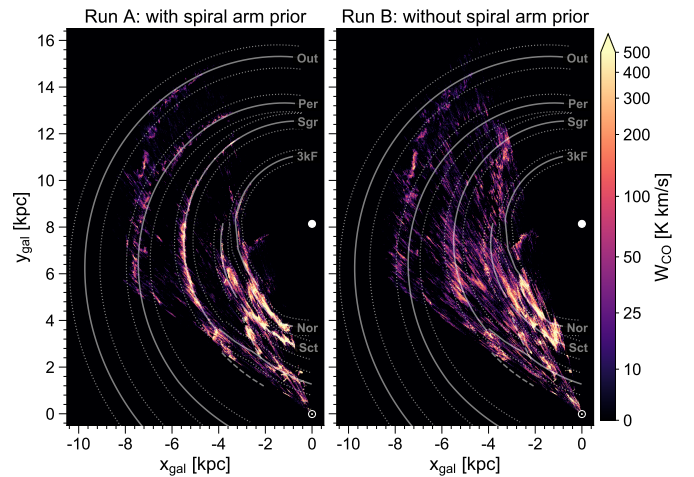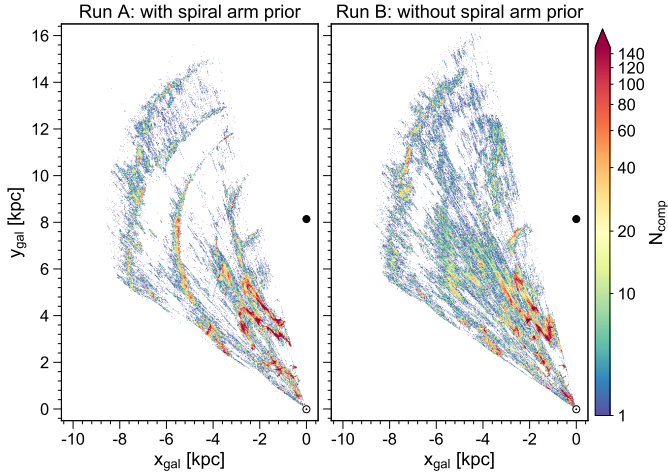

Fig. 8. Face-on view of the integrated ^{13}CO emission for the BDC results obtained with (*left*) and without (*right*) the spiral arm prior. The values are binned in 10×10 pc cells and are summed up along the z_{gal} axis. The position of the Sun and Galactic centre are indicated by the Sun symbol and black dot, respectively. When displayed in Adobe Acrobat, it is possible to hide the spiral arm positions and show the grid.

Even though Run B shows a larger spreading of emission into interarm regions, we can still identify ^{13}CO overdensities at the positions of the Galactic features of the SA model (right panels of Figs. 8 and 9). This is not surprising, as Run B has still a contribution from the maser parallax sources, which tend to be concentrated at spiral arms and spurs as well (cf. right panel in Fig. 1). Moreover, the Galactic features for the spiral arm model are also based on overdensities in archival H I and ^{12}CO Galactic plane surveys, so we would expect that the ^{13}CO emission is also present at these same locations.

Looking at the maps of the median σ_v values (Fig. 10), we qualitatively observe that spiral arm features seem to be associated with ^{13}CO components with larger linewidths. In general, we can see increased median σ_v values within Galactocentric distances $\lesssim 6$ kpc; as already speculated in Riener et al. (2020), these increasing σ_v values towards the inner Galaxy could be due to the presence of the Galactic bar and the observed overdensity

Fig. 9. Face-on view of the number of Gaussian fit components for the BDC results obtained with (*left*) and without (*right*) the spiral arm prior. The values are binned in 10×10 pc cells and are summed up along the z_{gal} axis. The position of the Sun and Galactic centre are indicated by the Sun symbol and black dot, respectively. When displayed in Adobe Acrobat, it is possible to show the spiral arm positions and the grid.
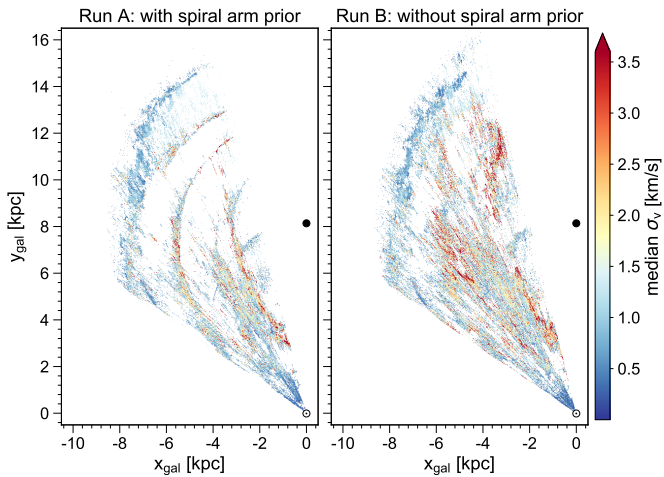

Fig. 10. Face-on view of the median velocity dispersion values of Gaussian fit components for the BDC results obtained with (*left*) and without (*right*) the spiral arm prior. The values are binned in 10×10 pc cells and the median was calculated along the z_{gal} axis. The position of the Sun and Galactic centre are indicated by the Sun symbol and black dot, respectively. When displayed in Adobe Acrobat, it is possible to show the spiral arm positions and the grid.

of star-forming regions (Anderson et al. 2017; Ragan et al. 2018), but could also partly result from our inability to correctly decompose strongly blended emission lines. We can further see an increase of the median σ_v values with heliocentric distance; for emission lines with $v_{\mathrm{LSR}} < 20$ km s^{-1} this is partly due to our use of the size-linewidth prior (Sect. 3.3). However, this effect is also present if we do not use this prior (see Appendix C.2 for a comparison between the maps of median σ_v values obtained with and without the size-linewidth prior).

Figures 8–10 also show a persistent feature at a Galactic longitude range of $29° \lesssim \ell \lesssim 38°$ that seemingly connects the Perseus and Outer arm[8]. This is very likely emission originating

---

[8] The position of this feature is indicated with a yellow dashed ellipse in Fig. 21.
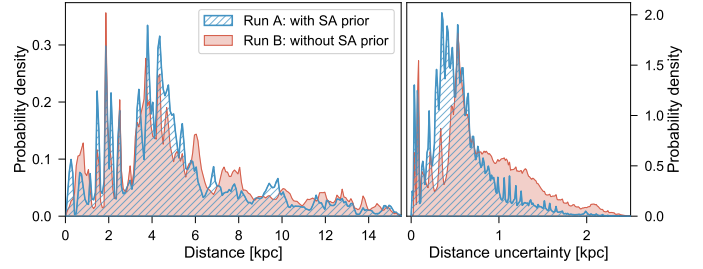
Fig. 11. PDFs for the estimated heliocentric distances (*left*) and corresponding uncertainties (*right*).

close to the Sun that has been erroneously placed at far distances. We can find evidence for this in Fig. 10, where the median σ_v value shows significantly lower values ($\lesssim 0.5$ km s^{-1}) at these locations than for most of the other parts of the Perseus arm. This erroneously placed local emission is also clearly identifiable in the offset positions from the Galactic midplane, which we discuss in Sect. 4.7.

4.3. Comparison of the distance results

Comparing the resulting distances of the two BDC runs, we find that 68.2% are compatible with each other within their estimated uncertainties[9]. The distance uncertainties are given by the standard deviation of the chosen Gaussian component fit to the combined distance PDF (see Sect. 3.4). In terms of differences in absolute distance uncertainty values, 67.2, 78.8, and 83.8% of the distance results are compatible within ± 0.5, ± 1.0, and ± 1.5 kpc, respectively. We thus conclude that for the majority of the GRS fit components the two distance runs yielded similar results. We can use the PDFs of the estimated heliocentric distances and corresponding distance uncertainties (Fig. 11) to identify where the distance estimates deviated. For example, Run B yielded more distances above 8 kpc (21.2%) but produced fewer distance assignments <0.5 kpc (1.2%) compared to Run A (17.1 and 2.4%, respectively).

The difference between the BDC runs is even more pronounced in the distance uncertainties. Half of the distance assignments of Run A have distance uncertainties <0.5 kpc, but only a quarter of the distance assignments for Run B are below this distance uncertainty threshold. This difference is also reflected in the estimated probabilities of the distances: about 44% of the results from Run A have high-confidence probabilities >0.75; for Run B only ~34% of the distance results exceed this probability threshold (see Appendix C.3 for more details). We caution that the estimated uncertainties and probabilities do not allow for a straightforward comparison of the quality of the distance results. Strongly favouring the distance assignments towards a particular prior may yield small uncertainties and high probabilities but the prior itself may lead to biased distance results. We discuss these issues further in Sect. 4.8.

In the top panels of Fig. 12 we show how the intensity and velocity dispersion values of the Gaussian fit components vary with heliocentric distance for both BDC runs. While the intensity values cover a large range, their median values stay flat over all considered distances.

The bottom panels in Fig. 12 show how the σ_v values of the fit components vary with their estimated distances. We can see a

---

[9] We note that each percentage point corresponds to about 46 500 independent distance assignments.
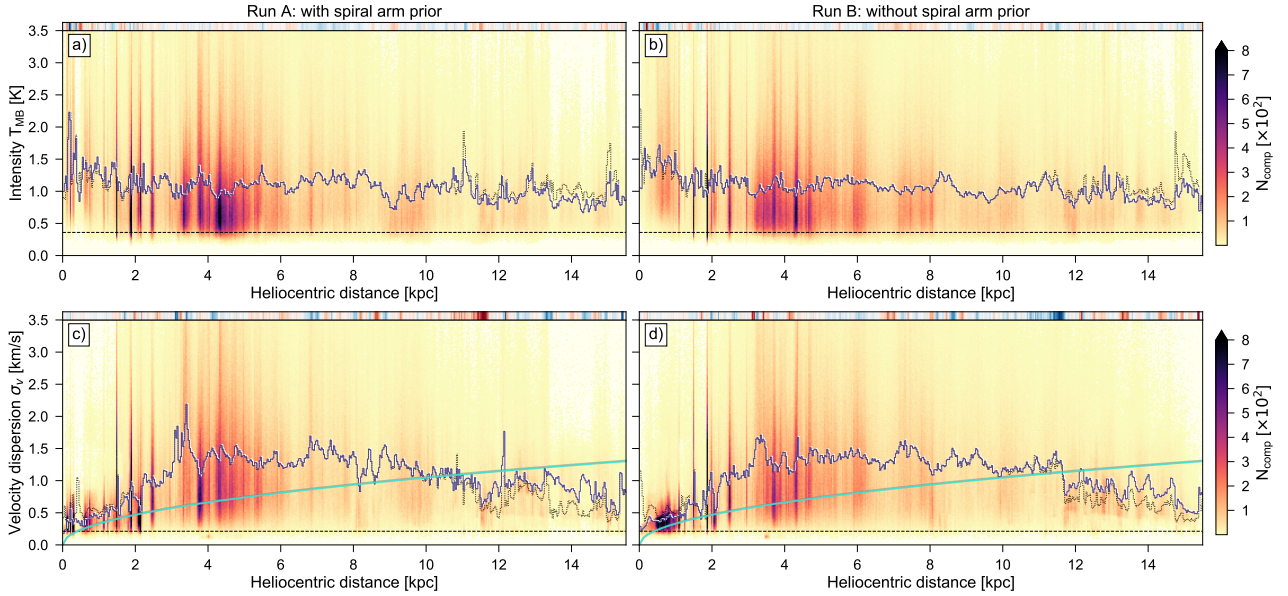
Fig. 12. 2D histograms of estimated distance values and intensity (*top*) and velocity dispersion (*bottom*) values for the BDC run with (*left column*) and without (*right column*) the spiral arm prior. The blue lines show the respective median values per distance bin and dotted black lines give the corresponding values for distances obtained without the size-linewidth prior. The small strips at the top of the individual panels show where the median value is higher (blue) or lower (red) compared to the opposite BDC run, with the strength of the colour corresponding to the magnitude of the difference. The turquoise line in the bottom panels indicates the expected values from the size-linewidth relationship (Eq. (1)). The dashed horizontal line in the top panels at $T_{MB} = 0.36$ K corresponds to the 3× S/N limit for the 0.1st percentile of the GRS noise distribution (see Riener et al. 2020). The dashed horizontal line in the bottom panels indicates the velocity resolution of the GRS (0.21 km s^{-1}).

clear increase in the median σ_v values up until heliocentric distances of about 3.5 kpc, after which it stays at increased values of >1 km s^{-1}, until it drops again at distances $\gtrsim 11.5$ kpc. This drop at the largest distances is due to a bias in the distance calculation that erroneously puts emission from nearby regions at large distances from the Sun (see Sect. 4.2). We also show the median σ_v values we would have gotten if we had not used the size-linewidth prior (Sect. 3.3), which shows an even bigger drop at these large distances. However, for $d < 4$ kpc we recover a similar trend of increased linewidths with larger heliocentric distances, indicating that beam averaging effects play a crucial role in producing these increased linewidths. Another explanation could be a larger non-thermal contribution to the linewidth for emission located in the inner part of the Galaxy. The comparison of the median σ_v curve with the size-linewidth relationship from Eq. (1) shows that most of the fit components have linewidths that are significantly larger than those expected values. In Fig. 12 we also indicate the median intensity and velocity dispersion values without the use of the size-linewidth prior (black dotted lines). Since we restricted the use of the size-linewidth prior to v_{LSR} values <20 km s^{-1}, the distribution does not change between $2 \lesssim d_\odot \lesssim 10$ kpc. However, we can see that in absence of the size-linewidth prior the distribution of the velocity dispersion values (bottom panels) contain much more gas emission with $\sigma_v < 1$ km s^{-1} at heliocentric distances $d_\odot > 11$ kpc, which based on our considerations in Sect. 3.3 is likely not correct. We thus conclude that while we need to exercise caution in the use of the size-linewidth prior its restricted use for v_{LSR} values <20 km s^{-1} led to significant improvements.

4.4. Gas fraction in spiral arm and interarm regions

In this section we discuss the fraction of ^{13}CO residing in spiral arm and interarm environments, which also serves to give a more quantitative overview of the distance results. In Table 3 we split our distance results into different sub-samples

that correspond to the determined association with Galactic features (left panel in Fig. 1) by the BDC. This association is based on the (ℓ, b, v_{LSR}) coordinates and the position and extents of the spiral arm and interarm features (see Sect. 2.1 in Reid et al. 2016 for more details about this association). For each sub-sample we report the fraction of the total integrated ^{13}CO intensity (W_{CO}), the fraction of the total number of fit components (N_{comp}), and the median velocity dispersion value ($\sigma_{v, med.}$) with the corresponding interquartile range (IQR) in brackets. We also list the combined values for all spiral arm (3kF, N1F, N1N, Out, Per, ScF, ScN, SgF, and SgN) and interarm (AqR, AqS, LoS, N/A) features as *Spiral arm* and *Interarm*, respectively.

In the two BDC runs, about 76−84% of the integrated ^{13}CO emission and 66−76% of the ^{13}CO fit components were associated with spiral arm features, mostly with the Norma, Scutum, and Sagittarius arms. Run B placed about 1.5 times more ^{13}CO emission in interarm regions not associated with any of the Galactic features shown in Fig. 1. To put these numbers into perspective and check whether also the gas distribution in Run B shows a significant concentration towards spiral arm features, we determined the fraction of ^{13}CO gas in spiral arms based on only kinematic distances. We calculate the kinematic distances using methods contained in the BDC v2.4 and solve for the KDA by using the Monte Carlo approach outlined in Sect. 3.1 of Roman-Duval et al. (2016), assuming a Gaussian vertical density profile of the molecular gas with a FWHM of 110 pc as was done in that study. For these pure kinematic distance solutions we find that ~58% of the integrated ^{13}CO emission and ~52% of the fit components overlap with the positions of spiral arms from our assumed model. These results demonstrate that compared to pure kinematic distances both our BDC runs contain a significant enhancement of ^{13}CO emission at the position of spiral arm features.

To further check the robustness of our results we also looked at the distance results of only the ~75% of fit components that

Table 3. Distance results for the two BDC runs.

|	Run A: with SA prior			Run B: without SA prior		
Feature [a]	W_{CO} [%]	N_{comp} [%]	$\sigma_{v,med.}$ [b] [km s^{-1}]	W_{CO} [%]	N_{comp} [%]	$\sigma_{v,med.}$ [b] [km s^{-1}]
3kF	2.6	2.4	1.4 $\begin{bmatrix}0.9\\2.2\end{bmatrix}$	1.4	1.6	1.3 $\begin{bmatrix}0.8\\2.0\end{bmatrix}$
AqR [c]	1.1	2.8	0.4 $\begin{bmatrix}0.3\\0.5\end{bmatrix}$	1.0	2.5	0.4 $\begin{bmatrix}0.3\\0.6\end{bmatrix}$
AqS	7.0	7.2	1.4 $\begin{bmatrix}0.9\\2.2\end{bmatrix}$	4.3	4.5	1.3 $\begin{bmatrix}0.8\\2.1\end{bmatrix}$
LoS	1.8	2.5	1.0 $\begin{bmatrix}0.7\\1.6\end{bmatrix}$	1.5	2.1	1.0 $\begin{bmatrix}0.7\\1.5\end{bmatrix}$
N1F	1.3	1.0	1.3 $\begin{bmatrix}0.9\\2.0\end{bmatrix}$	3.2	2.3	1.5 $\begin{bmatrix}1.0\\2.3\end{bmatrix}$
N1N	20.6	14.9	1.5 $\begin{bmatrix}1.0\\2.5\end{bmatrix}$	16.8	11.7	1.5 $\begin{bmatrix}1.0\\2.5\end{bmatrix}$
Out [c]	0.6	1.3	0.7 $\begin{bmatrix}0.5\\1.0\end{bmatrix}$	0.6	1.4	0.7 $\begin{bmatrix}0.5\\1.0\end{bmatrix}$
Per [c]	3.9	6.0	0.9 $\begin{bmatrix}0.6\\1.4\end{bmatrix}$	4.0	6.0	1.0 $\begin{bmatrix}0.6\\1.5\end{bmatrix}$
ScF	5.7	4.3	1.4 $\begin{bmatrix}0.9\\2.2\end{bmatrix}$	6.2	5.1	1.4 $\begin{bmatrix}0.9\\2.2\end{bmatrix}$
ScN	24.6	20.1	1.4 $\begin{bmatrix}0.9\\2.3\end{bmatrix}$	23.6	19.2	1.4 $\begin{bmatrix}0.9\\2.3\end{bmatrix}$
SgF	12.2	10.3	1.4 $\begin{bmatrix}0.9\\2.1\end{bmatrix}$	11.9	9.7	1.4 $\begin{bmatrix}0.9\\2.3\end{bmatrix}$
SgN	12.1	15.8	0.8 $\begin{bmatrix}0.5\\1.4\end{bmatrix}$	8.0	9.4	1.0 $\begin{bmatrix}0.6\\1.5\end{bmatrix}$
N/A	6.5	11.2	0.7 $\begin{bmatrix}0.5\\1.1\end{bmatrix}$	17.5	24.5	0.8 $\begin{bmatrix}0.5\\1.4\end{bmatrix}$
Spiral arms	83.6	76.2	1.2 $\begin{bmatrix}0.8\\2.0\end{bmatrix}$	75.6	66.3	1.3 $\begin{bmatrix}0.8\\2.1\end{bmatrix}$
Interarm	16.4	23.8	0.8 $\begin{bmatrix}0.5\\1.4\end{bmatrix}$	24.4	33.7	0.8 $\begin{bmatrix}0.5\\1.4\end{bmatrix}$

Notes. [a]3 kpc far arm (3kF), Aquila Rift (AqR), Aquila Spur (AqS), Local spur (LoS), Norma 1st quadrant near and far portions (N1N, N1F), Outer (Out), Perseus (Per), Scutum near and far portions (ScN, ScF), Sagittarius near and far portions (SgN, SgF), unassociated (N/A). [b]The two values in the brackets give the corresponding IQR. [c]Values are likely severely impacted by confusion between emission from the solar neighbourhood and far distances; see Sects. 4.7 and 4.8.

had a signal-to-noise ratio (S/N) > 3. We do not find significant deviations from the trends presented in Table 3. In particular, we recover the same difference in $\sigma_{v,med.}$ between the Galactic features, which we discuss in the next section.

4.5. Velocity dispersion in spiral arm and interarm regions

One interesting exercise is to look for possible variations of the gas velocity dispersion between spiral arm and interarm regions, which has been observed for the nearby spiral galaxy M 51 (Colombo et al. 2014). To split our data points into spiral arm and interarm features, we again use the BDC assignment with Galactic features from the previous section. Figure 13 shows σ_v-PDFs for these Galactic features and Table 3 gives the corresponding median values and interquartile ranges for these distributions. Generally speaking, spiral arm structures are associated with larger σ_v values than interarm structures, with the spiral arm PDF peaking at larger σ_v values. We note that the PDF labelled *Interarm* contains also associations with the spur features (AqS, LoS) and the nearby Aquila Rift complex (AqR). To check how this might skew the results, we also show PDFs for interarm emission not associated with any of the Galactic features from the SA model (labelled *Interarm (N/A)*) and emission only associated with spur features (*Spurs*). Interestingly, the PDF for the spurs is almost indistinguishable from the PDF of the spiral arms.

We make a more detailed comparison between emission associated with spiral arm and spur structures in Figs. 13c–f. The emission associated with the two major spiral structures covered by the GRS, the Scutum and Sagittarius arms, essentially has
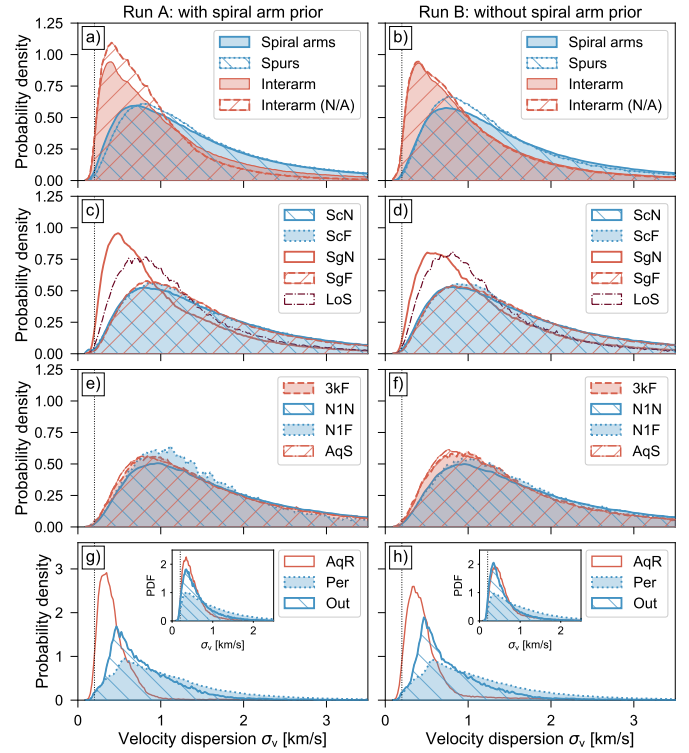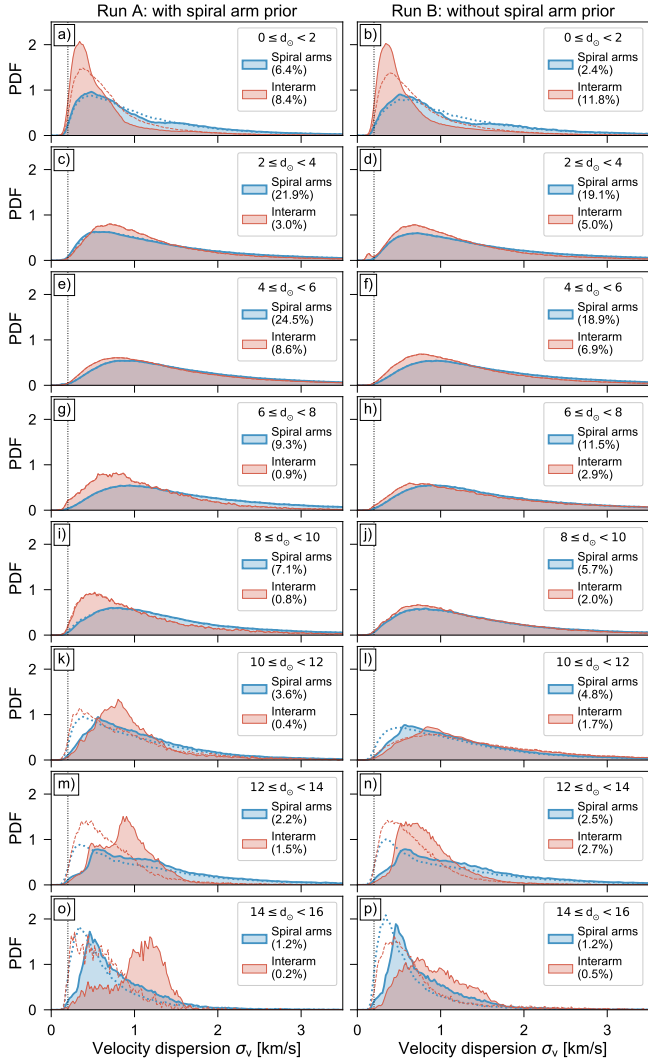

Fig. 13. PDFs of velocity dispersion values associated with Galactic features for the distance results with (*left columns*) and without (*right columns*) the SA prior. The dotted vertical line indicates the GRS velocity resolution (0.21 km s^{-1}). The insets in the bottom panels show the corresponding PDFs without the use of the size-linewidth prior.

identical σ_v-PDFs apart from the near portion of the Sagittarius arm (SgN), whose distribution peaks at much lower σ_v values and is more similar to the PDF of the Local Spur (LoS) and the interarm PDFs in panels a and b. Other structures in the inner Galaxy – the Norma arm (N1N, N1F), the far portion of the 3-kpc-arm (3kF), and the Aquila spur (AqS) – all show a very similar σ_v distribution that is essentially identical to the PDFs of the Scutum arm and the far portion of the Sagittarius arm. Since the near portion of the Sagittarius arm and the Local Spur are located at the highest longitude ranges covered by the GRS, this might point to real differences in terms of the linewidth distribution in the innermost and more outer parts of the GRS coverage. However, since parts of the SgN are also located close to the Sun ($d_\odot < 3$ kpc), its emission lines might simply be better resolved spatially, leading to narrower linewidths (see also discussion in Sect. 3.3). The difference in the σ_v-PDFs might also be explained by difficulties in the decomposition of strongly blended emission lines in the inner Galaxy, which could have led to higher fitted σ_v values.

The bottom panels (g, h) show PDFs for the Aquila Rift (AqR) complex and the Perseus (Per) and Outer (Out) arms. As already mentioned, we are not able to fully separate the near and far contribution of this emission with low v_{LSR} values. This problem is reflected in the shape of the PDFs, which are moreover impacted by our use of the size-linewidth prior. For comparison, we also show how the PDFs would look like if we did not use the size-linewidth prior (small insets in panels g and h). In this case their σ_v-PDFs become more similar, which is in contrast to expectations based on beam averaging effects (see Sect. 3.3 and Appendix B) and the other spiral arm PDFs, which show much higher σ_v values. We currently also have no reason to suspect

Fig. 14. PDFs of velocity dispersion values for 2 kpc heliocentric distance bins for the distance results with (*left columns*) and without (*right columns*) the SA prior. The dotted vertical line indicates the GRS velocity resolution (0.21 km s^{-1}). Dotted and dashed histograms show the distribution for distance results obtained without the size-linewidth prior. Percentages in the legend indicate the respective fraction of Gaussian fit components associated with spiral arm and interarm structures.

that the Perseus and Outer arms should be peculiar in terms of their linewidth distribution compared to other spiral arms.

To further check the significance of the difference in the σ_v-PDFs of spiral arm and interarm structures, we looked at the σ_v-PDFs in 2 kpc heliocentric bins (Fig. 14). About one third of the fit components associated with interarm structures have distances <2 kpc (panels a, b), compared to a much lower fraction of fit components associated with spiral arms in this distance range. This difference seems to be the major cause for the difference in the total σ_v-PDFs in Figs. 13a and b. The remaining interarm distributions in Fig. 14 show much closer resemblance to the spiral arm PDFs, and indicate no consistent or considerable trend towards lower linewidths.

Figure 14 once more highlights the problem of confusion between emission originating from the near and far side of the Galactic disc. For most of the PDFs in Figs. 14a–j we do see a shift towards higher linewidths with increasing distance ranges, which would match our expectations based on beam averaging

effects. The interarm PDFs in panels g and i show a deviation from this trend, which could be indicative of an increased confusion between near and far emission at these distance bins. The second bump at low σ_v values ($<0.2 \text{ km s}^{-1}$) in panel d is due to an instrumental artefact in the GRS data set that led to the fitting of very narrow components (see Appendix A.4 in Riener et al. 2020).

The strong confusion for emission at low v_{LSR} values ($<20 \text{ km s}^{-1}$) also becomes apparent again in panels a, b, and k–p of Fig. 14, which also highlight the effect of the size-linewidth prior. While we do find artefacts induced by the prior in these σ_v-PDFs, the distributions are nonetheless more consistent with the trend of higher σ_v values with increasing heliocentric distance. So even though Figs. 13 and 14 show that we need to be careful in interpreting the distance results for the emission features with low v_{LSR} values, we conclude that the use of the size-linewidth prior was justified and successful in disentangling part of the confusion between near and far emission.

4.6. Galactocentric variation of the gas properties

We now focus on the distribution of intensity and velocity dispersion values of the ^{13}CO fit components with Galactocentric distance (Fig. 15). These distributions also reveal some intriguing differences between the BDC runs. For example, for Run A we can identify an accumulation of data points at the approximate R_{gal} extent of the far portion of the 3-kpc arm (3kF), which however is almost entirely missing in Run B. Indeed, a comparison with Fig. 8 confirms that Run B puts significantly less emission at the location of the 3kF arm than Run A. This is most likely due to very large non-circular motions near the Galactic bar that introduces errors and large uncertainties for Run B, which depends mostly on the KD assumption of circular motions (see Sect. 4.8). In addition, there is large uncertainty in the rotation velocity at small Galactocentric radii, which also contributes to increased uncertainty for KD estimates. Another striking difference occurs at an R_{gal} value of ~8 kpc, where Run A shows large peaks that are missing in Run B. This emission corresponds to the position of the nearby Aquila Rift complex, but in Run B most of its emission is allocated to R_{gal} distances of ~7.5 kpc. We can confirm this in the top panels, where the accumulation of data points <0.5 kpc for Run A is shifted to higher distances (between 0.5 and 1 kpc) in Run B.

The intensity distribution (top panels in Fig. 15) shows large variation but an almost constant median value with no significant trends, similar to Fig. 12. The σ_v distributions (bottom panels in Fig. 15) show a more interesting behaviour; the median σ_v value stays at a large value of ~1.5 km s^{-1} from $3 \lesssim R_{\text{gal}} \lesssim 6$ kpc, after which it drops significantly to a value of ~0.5 km s^{-1}. As mentioned before, this could indicate that in the inner Galaxy the ^{13}CO components have higher non-thermal contributions or that there are increased problems in the decomposition of strongly blended emission in the inner parts of the GRS. We can however also interpret this trend as yet another indication that most of the emission at $R_{\text{gal}} \gtrsim 6.5$ kpc is associated with regions close to the Sun and thus has better resolved emission lines (Sect. 3.3).

4.7. Vertical distribution of the ^{13}CO emission

The Galactic plane has long been known to show a warp towards positive z_{gal} values in the first quadrant at Galactocentric distances $R_{\text{gal}} \gtrsim 7$ kpc (Gum et al. 1960)[10]. In Fig. 16 we show

---

[10] The BDC takes into account the effects of this warping in its calculation for the GL prior.
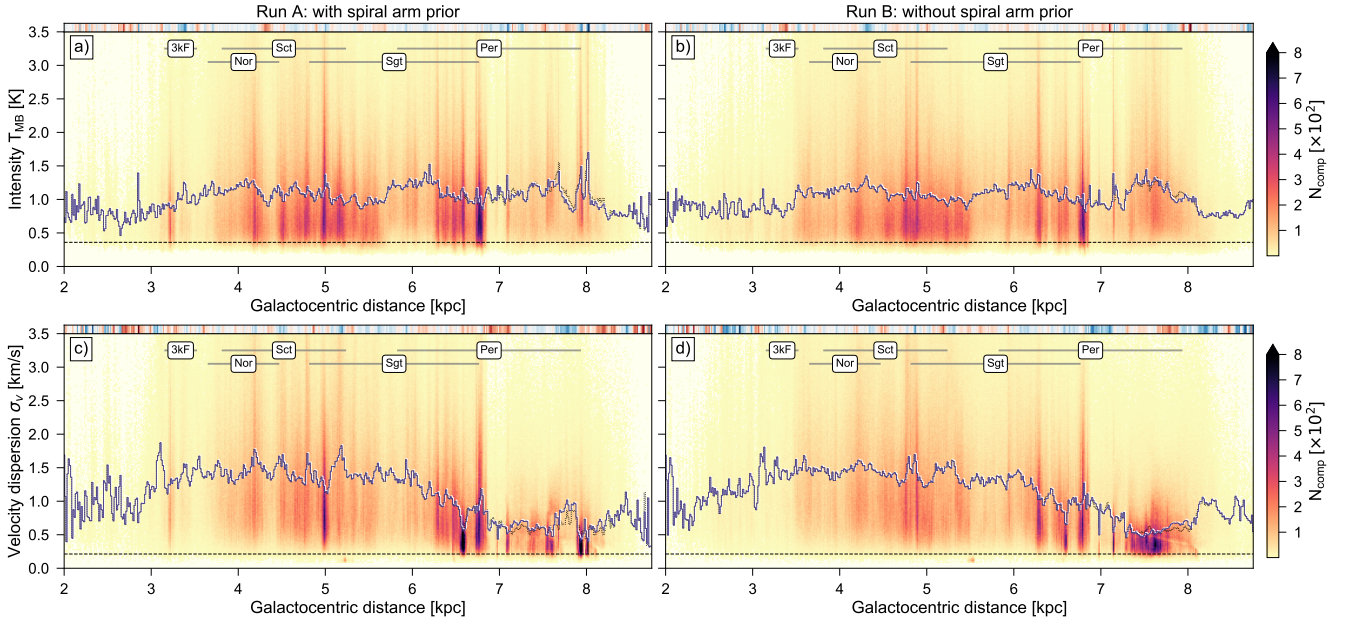
Fig. 15. 2D histograms of estimated Galactocentric distance values and intensity (*top*) and velocity dispersion (*bottom*) values for the BDC run with (*left column*) and without (*right column*) the spiral arm prior. The blue lines show the respective median values per distance bin and dotted black lines give the corresponding values for distances obtained without the size-linewidth prior. The small strips at the top of the individual panels show where the median value is higher (blue) or lower (red) compared to the opposite BDC run, with the strength of the colour corresponding to the magnitude of the difference. The grey horizontal lines in all panels show the approximate R_{gal} extent of five spiral arms overlapping with the GRS coverage. The dashed horizontal line in the top panels at $T_{MB} = 0.36$ K corresponds to the $3\times$ S/N limit for the 0.1st percentile of the GRS noise distribution (see Riener et al. 2020). The dashed horizontal line in the bottom panels indicates the velocity resolution of the GRS (0.21 km s^{-1}).
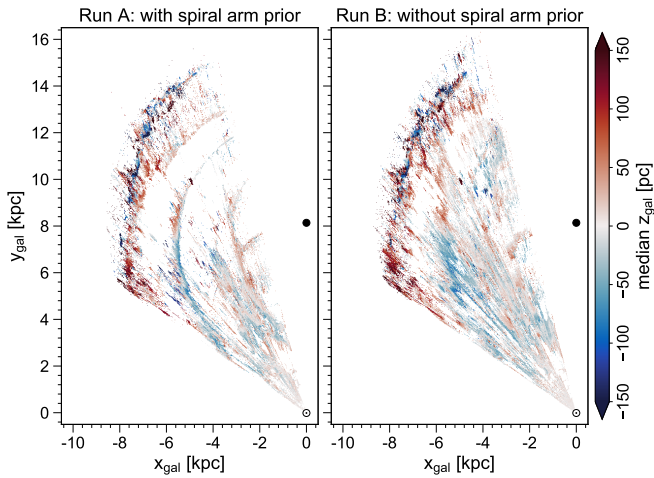
Fig. 16. Face-on view of the median z_{gal} values from the BDC results obtained with (*left*) and without (*right*) the spiral arm prior. The values are binned in 10×10 pc cells and the median was calculated along the z_{gal} axis. The position of the Sun and Galactic centre are indicated by the Sun symbol and black dot, respectively. When displayed in Adobe Acrobat, it is possible to show only the negative median z_{gal} positions, show spiral arm positions and show the grid.

a face-on view of the median z_{gal} values of our estimated distances, which clearly shows this warp of the molecular gas disc at $R_{gal} \gtrsim 7$ kpc. However, we can also see patches of negative z_{gal} values (< -100 pc) for regions that coincide with the Perseus and Outer arms. A comparison with Fig. 10 shows that these patches also correspond to the anomalously low σ_v values we already pointed out in Sect. 4.2. This confirms our suspicion that these patches most likely correspond to gas emission that originates

from very nearby regions that were erroneously assigned to large distances[11].

Another conspicuous feature is the presence of substantial negative z_{gal} values at the location of the Sagittarius arm at Galactic longitude values of $35° \lesssim \ell \lesssim 50°$ and 5 kpc $\lesssim R_{gal} \lesssim$ 7 kpc. More quantitatively, the estimated vertical heights for gas emission for these ℓ and R_{gal} ranges associated with the far portion of the Sagittarius arm have a median value of $z_{gal} = -34$ pc and span an IQR of -58 to -8 pc for both distance runs. This bend towards negative z_{gal} values at this longitude range is already clearly visible in the zeroth moment maps of the GRS data set (cf. Fig. 2 in Riener et al. 2020) and has also been observed in the *Herschel* Hi-GAL survey (Molinari et al. 2016). Since this distortion seems to be mainly present in the diffuse ISM component of the Milky Way, Molinari et al. (2016) speculated that it might be due to interaction with gas flows that originate from the Galactic halo or the Galactic fountain. However, instead of a global phenomenon these negative z_{gal} values could also simply indicate substructure of the Sagittarius arm.

In Fig. 17 we present PDFs for the estimated z_{gal} values, which have very similar shapes in both BDC runs. The most notable difference is that Run A shows a higher concentration at $z_{gal} = 0$, whereas Run B shows a dip at this position. This difference is mostly due to the association of sources with the Aquila Rift complex in Run A.

In our calculations we assumed that the Sun is located in the Galactic midplane, which is consistent with results from the most recent studies (Anderson et al. 2019; Reid et al. 2019). However, previous studies and observations found that the Sun has a vertical offset of $z_{offset} \sim 25$ pc from the IAU definition of the

---

[11] We note that the presence of these incorrect distance assignments do not change our general conclusion about the warp of the Galactic disc towards positive z_{gal} values.
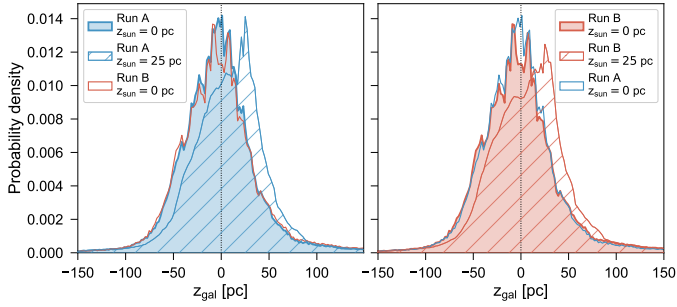
**Fig. 17.** PDFs of the vertical distribution of the ^{13}CO emission for the entire GRS data set. Shaded PDFs are for the BDC runs with (*left*) and without (*right*) the SA prior, and unfilled PDFs show the distribution of the opposite panel for reference. Hatched PDFs show the z_{gal} distribution assuming an offset of the Sun above the midplane of $z_{offset} = 25$ pc.

Galactic midplane (Goodman et al. 2014; Bland-Hawthorn & Gerhard 2016). Figure 17 shows how the PDFs would change if we correct for this assumed vertical offset of the Sun using Eq. (C3) from Ellsworth-Bowers et al. (2013). Accounting for such an offset leads to a shift of the distribution towards positive z_{gal} values, with an asymmetric peak at $z_{gal} \sim 25$ pc introduced by emission originating close to the Sun ($\lesssim 1$ kpc).

A Gaussian fit to the PDFs in Fig. 17 yields full-width at half-maximum (FWHM) values of 72 and 77 pc with corresponding mean or peak positions at $z_{gal} = -5$ and -6 pc for Run A and B, respectively. If a z_{offset} value of 25 pc is factored in, the FWHM values increase slightly to values of 78 and 83 pc and the centroid position changes to $z_{gal} = 10$ and 8 pc, respectively. Our FWHM estimate is lower by about one third than the value of 110 pc Roman-Duval et al. (2016) found for the dense gas (corresponding to H_2 surface densities $\gtrsim 25\ M_\odot\ pc^{-2}$) in the inner Milky Way. However, our results correspond very well with scale heights of ~ 30 to 40 pc and peak values of -4 to -10 pc that have been determined from high mass star forming regions, H II regions, and dust emission surveys in the far-infrared (see Table 1 in Anderson et al. 2019 for a compilation of literature results). A peak position at $z_{gal} = -5$ pc also agrees well with $z_{offset} \sim 5$ pc as found by Anderson et al. (2019) and Reid et al. (2019). We also note that the scale height of the H I cold neutral medium (~ 150 pc; Kalberla 2003) significantly exceeds our determined scale height for the ^{13}CO gas by about a factor of five. To check whether our results are impacted by the inclusion of both near and far emission, we also estimated the FWHM estimates for individual 1 kpc bins in the R_{gal} range of 3−6 kpc. We find a maximum FWHM extent of ~ 90 pc for $5 < R_{gal} < 6$ kpc and FWHM values of 70−75 pc at lower R_{gal} bins.

Figure 18 shows the distribution of σ_v values with vertical height z_{gal}. For both distance results we can see a clear concentration of data points towards the midplane. The decrease in the median σ_v value around a z_{gal} value of 0 is due to very nearby emission located <1 kpc from the Sun, which has very narrow linewidths. We also note the presence of an asymmetry, especially striking in the curve of median values, with a larger fraction of components with broader linewidths located at negative z_{gal} values. Riener et al. (2020) already found a similar asymmetry in the distribution of σ_v values with Galactic latitude. As argued in Riener et al. (2020), such an asymmetry could be explained by an offset position of the Sun above the Galactic midplane. However, as mentioned, recent results have found that the vertical position of the Sun agrees well with the
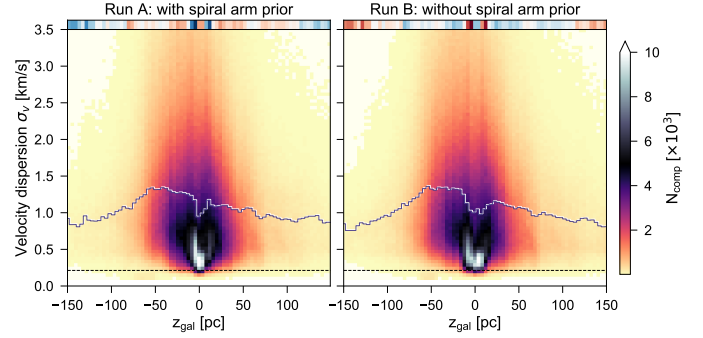
Fig. 18. 2D histograms of velocity dispersion and estimated vertical distances z_{gal} for the BDC run with (*left*) and without (*right*) the spiral arm prior. The blue line shows the median σ_v value per z_{gal} bin. The small strips at the top of the individual panels show where the median value is higher (blue) or lower (red) compared to the opposite BDC run, with the strength of the colour corresponding to the magnitude of the difference. The dashed horizontal line indicates the GRS velocity resolution (0.21 km s^{-1}).

location of the Galactic midplane (Anderson et al. 2019; Reid et al. 2019).

4.8. Potential problems, artefacts, and biases

The BDC tool was designed to estimate distances for spiral arm sources, which means that its default settings have an inherent bias of associating sources with Galactic features from its spiral arm model. Since we use the BDC in assigning distances to the gas emission of an entire Galactic plane survey, we need to be careful in interpreting its results and should be aware of the biases present in the distance calculation.

It is a priori not clear which of our BDC runs yields more trustworthy or better distance solutions. Run A has the obvious problem that the gas emission is preferentially located closer to the Galactic features included in the spiral arm model. For this run we expect biased results in terms of the distribution of emission in spiral arm and interarm regions, with the latter likely severely underestimated. Run B gives more unbiased results with regards to the allocation of the gas to arm and interarm regions. However, we note that for the distance results from Run B an association with maser parallax sources can be a decisive factor for the choice of the most likely distance (cf. left panel of Fig. 2). Since these maser sources do mostly overlap with the Galactic features of the spiral arm model (Fig. 1), the distance results thus still contain an implicit, albeit moderate, association with these Galactic features. Moreover, since Run B is dominated by the KD prior, it is also more strongly affected by the ambiguities and uncertainties of the KD method.

We can identify an accumulation of emission features around the locus of tangent points for both distance estimates. The problem with determining kinematic distances near tangent points is that small changes in the v_{LSR} value result in large changes in the estimated KD value. Thus often a threshold is used for the tangent point distance allocation, where for example all sources with v_{LSR} values within 10 km s^{-1} of the tangent point velocity are assigned the tangent point distance (e.g. Urquhart et al. 2018). We indicate the corresponding region where v_{LSR} values are within 10 km s^{-1} of the tangent point velocity in Fig. 21. This threshold of 10 km s^{-1} corresponds to expected velocity deviations introduced by streaming motions (e.g. Burton 1971; Ramón-Fox & Bonnell 2018). We thus speculate that some of the empty voids within this region might be at least partly due
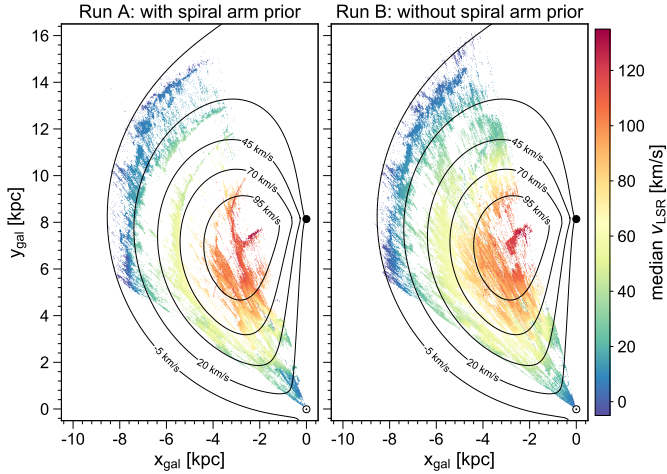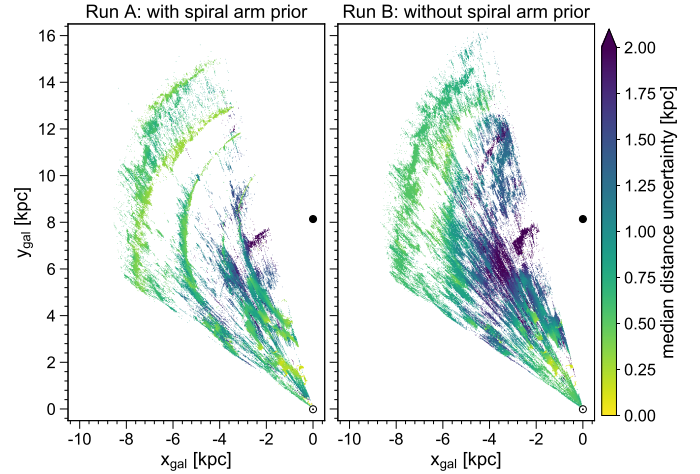
Fig. 19. Face-on view of the median v_{LSR} values from the BDC results obtained with (*left*) and without (*right*) the spiral arm prior. The values are binned in 10×10 pc cells and the median was calculated along the z_{gal} axis. The position of the Sun and Galactic centre are indicated by the Sun symbol and black dot, respectively. When displayed in Adobe Acrobat, it is possible to show the spiral arm positions, hide the curves of constant projected v_{LSR}, and show the grid.

Fig. 20. Face-on view of the median distance uncertainty for the BDC results obtained with (*left*) and without (*right*) the spiral arm prior. The values are binned in 10×10 pc cells and are summed up along the z_{gal} axis. The position of the Sun and Galactic centre are indicated by the Sun symbol and black dot, respectively. When displayed in Adobe Acrobat, it is possible to show the spiral arm positions and show the grid.

to this confusion around the tangent point velocity, at least for Run B that is dominated by the KD prior. However, a comparison with the default BDC runs (Fig. C.2) shows that our use of literature distance solutions helped to substantially decrease artefacts around the locus of tangent points.

Figure 19 shows a comparison of the median associated v_{LSR} values for the distance results. We overplot this figure with curves of constant projected v_{LSR} values that were calculated with the methods included in the BDC. In general, the BDC runs produced distance results that are well in agreement with the assumed Galactic rotation curve model. This good correspondence is not surprising, given that the KD solutions are calculated using the rotation curve model. However, an anticipated problem is that peculiar gas motions, for example introduced by streaming motions within spiral arms, might cause a significant deviation from the expected v_{LSR} velocities of the assumed rotation curve model (e.g. Ramón-Fox & Bonnell 2018). This strongly affects regions with R_{gal} values $\lesssim 5$ kpc, for which we expect large peculiar motions due to the influence of the Galactic bar (Reid et al. 2019). The BDC takes this into account by down-weighting the KD prior for regions closer to the Galactic centre (see Fig. 2); this however has a significant impact on the estimated distance uncertainties, which show a substantial increase with decreasing R_{gal} values (Fig. 20). As already noticed in Fig. 11, this effect is much stronger for Run B, since in this case the combined PDF shows broader peaks and corresponding Gaussian fits to these peaks result in higher estimated distance uncertainties. We give a more detailed discussion about the deviations from the rotation curve velocities and the regions where they occur in Appendix C.4. We also note that the v_{LSR} uncertainties of the fit components can have a large impact on the distance calculation routine, as larger uncertainty values can lead to an association with more parallax sources or Galactic features. We illustrate and discuss this effect with an example in Appendix C.1. Moreover, the BDC is expected to have problems at lower Galactic longitudes for both Run A and Run B, since for low longitude ranges emission in the various arm segments
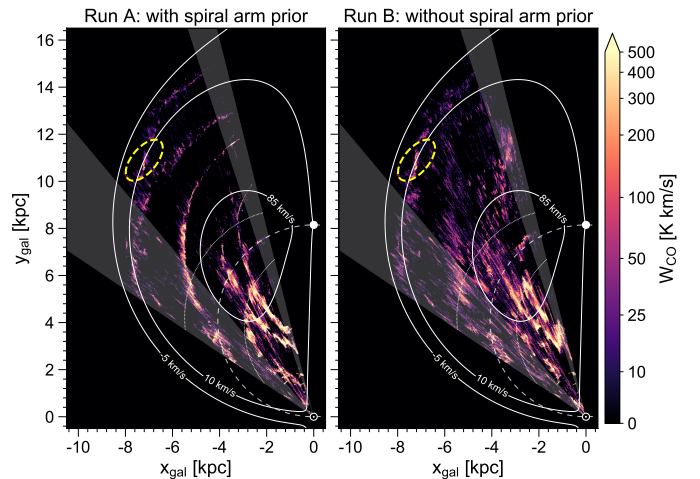
Fig. 21. Same as Fig. 8, but overplotted with diagnostics to illustrate potential issues in the distance assignment. The position of the Sun and the Galactic centre are indicated with a Sun symbol and a white dot, respectively. Solid white curves show constant projected v_{LSR} velocities. The dashed white curve marks the locus of tangent points and dotted lines indicate the area where v_{LSR} values are within 10 km s^{-1} of the tangent point velocity. Shaded areas show GRS regions that either had limited latitude or velocity coverage. The yellow dashed ellipse indicates a likely artefact of nearby emission that was erroneously placed at far distances. See Sect. 4.8 for more details.

along the lines of sight has similar Galactic latitude and v_{LSR} values, thus complicating the distance allocation.

Since in this work we do not explicitly correlate the distance results of neighbouring lines of sight, it is possible that assigned distance values can show strong variation between neighbouring lines of sight. We can see this effect as emission features that are spread out along the line of sight (reminiscent of the 'Fingers of God effect'; see e.g. the right panel of Fig. 21).

The limited spatial and spectral coverage of the GRS also introduces some artefacts in the distance estimation (Fig. 21).

The limited latitude coverage from $14° \lesssim \ell \lesssim 18°$ results in missing patches of CO emission at this longitude range. Moreover, the face-on view is restricted by the lower limit of the velocity coverage ($v_{LSR} = -5$ km s^{-1}) and thus contains no emission past $d \sim 13$ kpc at $\ell > 40°$. However, the reduced velocity coverage of $-5 \lesssim v_{LSR} \lesssim 85$ km s^{-1} for $40° \lesssim \ell \lesssim 56°$ should not impact our distance results as we would not expect emission peaks with $v_{LSR} > 85$ km s^{-1} at these larger longitude values, as demonstrated in Fig. 21.

We note that the GRS decomposition also has uncertainties that might cause problems for the distance estimation. Especially in the inner part of the Galaxy emission lines can be strongly blended, which could have led to difficulties in the decomposition (see the discussion about flagged components in Sect. 3.1 of Riener et al. 2020). We tried to fold these considerations into the uncertainties of the v_{LSR} position supplied to the BDC (see discussion in Appendix C.1). Moreover, we tested the effects of a quality cut based on the S/N ratio of the fit components and found that this does not change our overall conclusions (see Sect. 4.4).

As already mentioned in previous sections, the feature that seemingly bridges the Perseus and Outer arm and corresponds to emission with v_{LSR} values around 10 km s^{-1} (yellow dashed ellipse in Fig. 21) is most likely an artefact introduced by the KD prior. As discussed in Riener et al. (2020) and Sect. 3.3, for this v_{LSR} regime there is strong confusion between emission from the solar neighbourhood and the far disc. Moreover, the BDC biases emission lines with $v_{LSR} \lesssim 5$ km s^{-1} towards far distances (cf. Fig. 3). We can estimate the magnitude of this error by counting all components with an unlikely combination of distance, z_{gal}, and σ_v values. Choosing $d > 8$ kpc, $z_{gal} < -50$ pc, and $\sigma_v < 0.5$ km s^{-1} as unlikely combination of parameters, we find that for both distance results about 0.3% of the fit components satisfy these parameters[12]. Since the W_{CO} values of these components account for only about 0.1% of the total emission, we conclude that this problem has a very minor impact on our overall conclusions concerning the fraction of emission located in arm and interarm regions. However, this issue has strong impacts on our conclusions about the σ_v distribution in the Perseus and Outer arm (see Sect. 4.5).

4.9. Comparison with previous results

The BDC has already been used in the distance estimation to clouds and clumps extracted from other Galactic plane surveys (Urquhart et al. 2018; Rigby et al. 2019). However, since these works used an older version of the BDC (v1.0), it is not straightforward to compare their distance results with the distances obtained in this work. Moreover, our distance estimations are not independent from these previous results, since we use them as input for our KDA prior. Notwithstanding these issues, in Appendix A.4 we discuss how well we are able to match these previous distance results and conclude that we recover the vast majority of literature distance results within the degree of expected uncertainties introduced by the updated rotation curve parameters from BDC v1 to v2.4.

Unfortunately, it is also extremely challenging to compare our results in terms of the spiral arm to interarm fraction and variation of physical properties with previous studies, given that other works used a combination of different tracers, different

spiral arm models, or different assumptions about the Galactic rotation curve and the distance to the Galactic centre. It would be necessary to homogenise all data sets before any attempted comparison, which potentially requires recalculating and updating the literature distance results with our assumed Galactic parameters. Since such a homogenisation exceeds the scope of this work, we decided on a strictly qualitative comparison with some of the previous results and do not attempt to account for any of these systematic differences.

Previous works analysing the GRS found a similar overdensity of ^{13}CO with spiral arm features. Roman-Duval et al. (2009) found for their sample of GRS molecular clouds that the ^{13}CO surface brightness is strongly enhanced at the location of spiral arms from the model of Vallee (1995). Sawada et al. (2012) also found that the GRS emission shows bright and compact concentrations along spiral arm features, whereas more diffuse and extended emission dominates the interarm regions.

However, recent results from other Galactic plane surveys in the first quadrant found a weaker correspondence of molecular clouds with spiral arms. Colombo et al. (2019) analysed a large ^{12}CO (3–2) survey overlapping with the GRS and could only attribute about 35% of the flux to molecular clouds associated with spiral arms[13]. Colombo et al. (2019) attribute this low fraction of flux in spiral arm clouds to difficulties in the distance assignments and optical depth effects of the ^{12}CO (3–2) emission.

Recently, Rigby et al. (2019) found that clumps from a distance limited ($6 < d_\odot < 9$ kpc) sample associated with spiral arms have significantly higher σ_v values than clumps at the same distances that are located in interarm regions. They further note that this difference in linewidth is comparable to what has been found in extragalactic work (Colombo et al. 2014) and smoothed particle-hydrodynamics simulations (Duarte-Cabral & Dobbs 2016). We do find a trend for lower σ_v values at around the same distances for interarm regions for Run A, but not for Run B (Fig. 12). This is somewhat surprising, given that Rigby et al. (2019) used the BDC without the SA prior for their distance calculation, which should have yielded a better agreement with our Run B. However, we note that Rigby et al. (2019) used a higher-density tracer (^{13}CO (3–2)) and v1 of the BDC[14], which could both account for any differences compared to our results.

5. Conclusions

In this work we present distance estimates for the Gaussian decomposition results of the Galactic Ring Survey presented in Riener et al. (2020). Using the most recent version of the Bayesian Distance Calculator tool Reid et al. (2016, 2019), we perform two separate distance calculations for the ~4.6 million individual Gaussian fit components, for which we vary the settings so as to either incorporate or neglect a prior for an association with spiral arm structure (labelled Run A and Run B,

---

[12] If we had not used the size-linewidth prior this fraction would increase to about 0.6 and 0.9% of the fit components for the runs with and without the SA prior, respectively.

[13] Colombo et al. (2019) used the spiral arm model by Vallée (2017) as a comparison, in which positions for the Scutum and Sagittarius arms deviate by up to ~1 kpc compared to the corresponding arms defined by Reid et al. (2019).

[14] BDC v2.4 includes new maser parallax sources, updated models for the Galactic rotation curve and spiral arm features, and contains significant changes in the distance estimation, such as a down-weighting of the KD prior in the inner Galaxy to accommodate expected large streaming motions introduced by the Galactic bar. See Reid et al. (2019) for more information.

respectively). In addition, we include literature distance information of objects overlapping with the GRS coverage as prior information for solving the kinematic distance ambiguity. We also incorporate a size-linewidth prior to solve for the confusion between emission from the solar neighbourhood and the far Galactic disc for emission peaks with line centroids of $v_{LSR} < 20 \, \text{km s}^{-1}$.

We find that most of the distance results of the two BDC runs are consistent with each other within their uncertainties, with most of the differences either due to the strong influence of the spiral arm prior for Run A or larger uncertainties introduced by the stronger effect of the kinematic distance prior for Run B. The two distance runs complement each other and show opposing strengths and weaknesses, thus suggesting that the true distribution of the gas emission is closer to a combination of the two results than to each of the individual distance runs.

In the following we present our main findings based on these two distance results:

i) The majority of the ^{13}CO emission is associated with spiral arm features as defined in the model by Reid et al. (2019). The fraction of ^{13}CO emission located in interarm regions varies from 16 to 24% in terms of the total ^{13}CO integrated emission and 24 to 34% in terms of the total number of ^{13}CO velocity components.

ii) The vertical distribution of the gas emission has a FWHM extent of ~75 pc. We recover a significant warp of the molecular disc towards positive z_{gal} values of more than 100 pc for the far side of the disc at $R_{gal} > 7$ kpc and the entire covered longitude range of $14° < \ell < 56°$. The gas disc shows a significant bend towards negative z_{gal} values at the position of the Sagittarius arm at Galactic longitude values of $35° \lesssim \ell \lesssim 50°$ and $5 \, \text{kpc} \lesssim R_{gal} \lesssim 7 \, \text{kpc}$.

iii) We find a trend of higher velocity dispersion values with increasing heliocentric distance, which we attribute mostly to beam averaging effects. Most of the velocity dispersion values also significantly exceed expected values based on an assumed size-linewidth relationship.

iv) The ^{13}CO emission associated with spiral arms and spur features has a similar distribution of velocity dispersion values, which is shifted to higher values compared to the distribution of velocity dispersion values in interarm structures. However, we find that most of this difference is due to the location of a significant fraction of interarm gas at close distances to the Sun, which resulted in spatially better resolved lines and narrower linewidths. While we cannot exclude variations in the linewidth between spiral arm and interarm gas, we conclude that our present results do not support strong differences in σ_v between these environments.

v) There is strong confusion between ^{13}CO emission coming from the local solar neighbourhood and regions associated with the Perseus and Outer arm. By using the velocity dispersion values of the fit components as an additional prior we could significantly reduce the confusion between near and far emission for low v_{LSR} velocities ($-5 < v_{LSR} < 20 \, \text{km s}^{-1}$).

While we use the currently best knowledge about the structure of our Galaxy for our distance results, we anticipate that these are subject to change, in particular due to updates on the BDC method, the Galactic rotation model, and the position of Galactic features, with additional and more precise maser parallax measurements, and new KDA solutions for sources overlapping with the GRS coverage. The BDC tool and its enhancements discussed in this work are designed to be versatile enough to incorporate these changes. We thus conclude that the approach presented herein should be a helpful contribution to the problem of estimating distances to gas emission features from Galactic plane surveys.

References

Anderson, L. D., & Bania, T. M. 2009, ApJ, 690, 706
Anderson, L. D., Bania, T. M., Jackson, J. M., et al. 2009, ApJS, 181, 255
Anderson, L. D., Bania, T. M., Balser, D. S., et al. 2014, ApJS, 212, 1
Anderson, L. D., Wang, Y., Bihr, S., et al. 2017, A&A, 605, A58
Anderson, L. D., Wenger, T. V., Armentrout, W. P., Balser, D. S., & Bania, T. M. 2019, ApJ, 871, 145
Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Battisti, A. J., & Heyer, M. H. 2014, ApJ, 780, 173
Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529
Brand, J., & Blitz, L. 1993, A&A, 275, 67
Burton, W. B. 1971, A&A, 10, 76
Clemens, D. P. 1985, ApJ, 295, 422
Colombo, D., Hughes, A., Schinnerer, E., et al. 2014, ApJ, 784, 3
Colombo, D., Rosolowsky, E., Duarte-Cabral, A., et al. 2019, MNRAS, 483, 4291
Combes, F. 1991, ARA&A, 29, 195
Dame, T. M., Elmegreen, B. G., Cohen, R. S., & Thaddeus, P. 1986, ApJ, 305, 892
Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792
Dempsey, J. T., Thomas, H. S., & Currie, M. J. 2013, ApJS, 209, 8
Duarte-Cabral, A., & Dobbs, C. L. 2016, MNRAS, 458, 3667
Eden, D. J., Moore, T. J. T., Plume, R., & Morgan, L. K. 2012, MNRAS, 422, 3178
Eden, D. J., Moore, T. J. T., Morgan, L. K., Thompson, M. A., & Urquhart, J. S. 2013, MNRAS, 431, 1587
Eden, D. J., Moore, T. J. T., Urquhart, J. S., et al. 2015, MNRAS, 452, 289
Elia, D., Molinari, S., Schisano, E., et al. 2017, MNRAS, 471, 100
Ellsworth-Bowers, T. P., Glenn, J., Rosolowsky, E., et al. 2013, ApJ, 770, 39
Ellsworth-Bowers, T. P., Rosolowsky, E., Glenn, J., et al. 2015, ApJ, 799, 29
Elmegreen, B. G. 2009, IAU Symp., 254, 289
Fish, V. L., Reid, M. J., Wilner, D. J., & Churchwell, E. 2003, ApJ, 587, 701
Ginsburg, A., Glenn, J., Rosolowsky, E., et al. 2013, ApJS, 208, 14
Gómez, G. C. 2006, AJ, 132, 2376
Goodman, A. A., Alves, J., Beaumont, C. N., et al. 2014, ApJ, 797, 53
Gravity Collaboration (Abuter, R., et al.) 2019, A&A, 625, L10
Green, D. A. 2019, JApA, 40, 36
Gum, C. S., Kerr, F. J., & Westerhout, G. 1960, MNRAS, 121, 132
Henshaw, J. D., Longmore, S. N., Kruijssen, J. M. D., et al. 2016, MNRAS, 457, 2675
Henshaw, J. D., Ginsburg, A., Haworth, T. J., et al. 2019, MNRAS, 485, 2457
Henshaw, J. D., Kruijssen, J. M. D., Longmore, S. N., et al. 2020, Nat. Astron., in press [arXiv:2007.01877]
Heyer, M., & Dame, T. M. 2015, ARA&A, 53, 583
Heyer, M., Krawczyk, C., Duval, J., & Jackson, J. M. 2009, ApJ, 699, 1092
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Jackson, J. M., Bania, T. M., Simon, R., et al. 2002, ApJ, 566, L81
Jackson, J. M., Rathborne, J. M., Shah, R. Y., et al. 2006, ApJS, 163, 145
Kalberla, P. M. W. 2003, ApJ, 588, 805

Kawata, D., Bovy, J., Matsunaga, N., & Baba, J. 2019, MNRAS, 482, 40
Larson, R. B. 1981, MNRAS, 194, 809
Leahy, D. A., & Ranasinghe, S. 2018, ApJ, 866, 9
Liszt, H. S., & Burton, W. B. 1981, ApJ, 243, 778
López-Corredoira, M., & Sylos Labini, F. 2019, A&A, 621, A48
Marshall, D. J., Joncas, G., & Jones, A. P. 2009, ApJ, 706, 727
Miville-Deschênes, M.-A., Murray, N., & Lee, E. J. 2017, ApJ, 834, 57
Molinari, S., Noriega-Crespo, A., Bally, J., et al. 2016, A&A, 588, A75
Moore, T. J. T., Urquhart, J. S., Morgan, L. K., & Thompson, M. A. 2012, MNRAS, 426, 701
Persic, M., Salucci, P., & Stel, F. 1996, MNRAS, 281, 27
Ragan, S. E., Moore, T. J. T., Eden, D. J., et al. 2016, MNRAS, 462, 3123
Ragan, S. E., Moore, T. J. T., Eden, D. J., et al. 2018, MNRAS, 479, 2361
Ramón-Fox, F. G., & Bonnell, I. A. 2018, MNRAS, 474, 2028
Ranasinghe, S., & Leahy, D. A. 2018a, AJ, 155, 204
Ranasinghe, S., & Leahy, D. A. 2018b, MNRAS, 477, 2243
Ranasinghe, S., Leahy, D. A., & Tian, W. 2018, Open Phys. J., 4, 1
Rathborne, J. M., Johnson, A. M., Jackson, J. M., Shah, R. Y., & Simon, R. 2009, ApJS, 182, 131
Reid, M. J., Menten, K. M., Zheng, X. W., et al. 2009, ApJ, 700, 137
Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2014, ApJ, 783, 130
Reid, M. J., Dame, T. M., Menten, K. M., & Brunthaler, A. 2016, ApJ, 823, 77
Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, ApJ, 885, 131
Rice, T. S., Goodman, A. A., Bergin, E. A., Beaumont, C., & Dame, T. M. 2016, ApJ, 822, 52
Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, A&A, 628, A78
Riener, M., Kainulainen, J., Beuther, H., et al. 2020, A&A, 633, A14
Rigby, A. J., Moore, T. J. T., Plume, R., et al. 2016, MNRAS, 456, 2885
Rigby, A. J., Moore, T. J. T., Eden, D. J., et al. 2019, A&A, 632, A58
Robitaille, T., & Bressert, E. 2012, Astrophysics Source Code Library [record ascl:1208.017]
Roman-Duval, J., Jackson, J. M., Heyer, M., et al. 2009, ApJ, 699, 1153
Roman-Duval, J., Heyer, M., Brunt, C. M., et al. 2016, ApJ, 818, 144
Rosolowsky, E., Dunham, M. K., Ginsburg, A., et al. 2010, ApJS, 188, 123
Sawada, T., Hasegawa, T., & Koda, J. 2012, ApJ, 759, L26
Schisano, E., Molinari, S., Elia, D., et al. 2020, MNRAS, 492, 5420
Shetty, R., Beaumont, C. N., Burton, M. G., Kelly, B. C., & Klessen, R. S. 2012, MNRAS, 425, 720
Simon, R., Rathborne, J. M., Shah, R. Y., Jackson, J. M., & Chambers, E. T. 2006a, ApJ, 653, 1325
Simon, R., Jackson, J. M., Rathborne, J. M., & Chambers, E. T. 2006b, ApJ, 639, 227
Solomon, P. M., Rivolo, A. R., Barrett, J., & Yahil, A. 1987, ApJ, 319, 730
Stark, A. A., & Brand, J. 1989, ApJ, 339, 763
Su, Y., Yang, J., Zhang, S., et al. 2019, ApJS, 240, 9
Svoboda, B. E., Shirley, Y. L., Battersby, C., et al. 2016, ApJ, 822, 59
Taylor, J. H., & Cordes, J. M. 1993, ApJ, 411, 674
Umemoto, T., Minamidani, T., Kuno, N., et al. 2017, PASJ, 69, 78
Urquhart, J. S., Csengeri, T., Wyrowski, F., et al. 2014, A&A, 568, A41
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059
Vallee, J. P. 1995, ApJ, 454, 119
Vallée, J. P. 2017, Astron. Rev., 13, 113
VERA Collaboration (Hirota, T., et al.) 2020, PASJ, 72, 50
Wang, Y., Beuther, H., Rugel, M. R., et al. 2020, A&A, 634, A83
Waskom, M., Botvinnik, O., O'Kane, D., et al. 2018, mwaskom/seaborn: v0.9.0 (July 2018)
Wenger, T. V., Balser, D. S., Anderson, L. D., & Bania, T. M. 2018, ApJ, 856, 52
Wienen, M., Wyrowski, F., Schuller, F., et al. 2012, A&A, 544, A146
Xu, Y., Hou, L.-G., & Wu, Y.-W. 2018, Res. Astron. Astrophys., 18, 146

Appendix A: Further details about the KDA prior
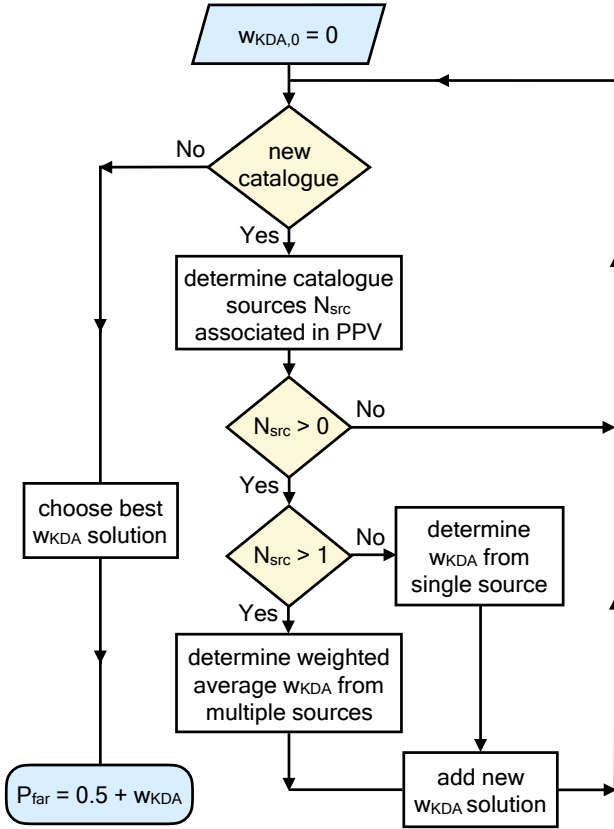
Fig. A.1. Flowchart outlining how literature distance estimates are used to determine a prior for the KDA solution.

Here we give more details about the KDA prior (Sect. 3.2). We describe the method we used to calculate the prior, present the literature results we considered for the prior, and discuss its effect on the distance calculation.

A.1. Calculation of priors

In the following we describe the iterative loop that is used to determine priors for the KDA solution (see also Fig. A.1). For a given $(\ell, b, v_{\mathrm{LSR}})$ coordinate, we first determine how many sources N_{src} from a catalogue are associated with this coordinate. For this we define a Gaussian weight using the definition of a Gaussian function:

$$f(x) = a \cdot \exp\left(-4 \cdot \ln(2) \cdot \frac{(x - \mu)^2}{\Theta^2}\right), \tag{A.1}$$

where a, μ, and Θ denote the amplitude, mean position, and FWHM values, respectively. From this definition we construct a Gaussian weight w_{g} that evaluates to unity at $x = \Theta/2$:

$$w_{\mathrm{g}}(x) = \tilde{a} \cdot \exp\left(-4 \cdot \ln(2) \cdot x^2\right), \tag{A.2}$$

with a normalisation factor of $\tilde{a} = \exp(\ln(2))$ and x being the distance to the mean position (in our case $\mu = 0$) in fractions of the FWHM (f_Θ). We can apply this Gaussian weighting straightforwardly along the spectral axis, where $x \equiv f_\Theta = |v_{\mathrm{LSR}}^{\mathrm{src}} - v_{\mathrm{LSR}}|/\Theta_{\mathrm{src}}$, with $v_{\mathrm{LSR}}^{\mathrm{src}}$ and Θ_{src} being the measured centroid velocity and

linewidth of the catalogue sources, respectively. We define the weight along the velocity axis w_{V} as:

$$w_{\mathrm{V}} = \begin{cases} 1, & \text{for } w_{\mathrm{g}}(f_\Theta) \geq 1 \\ w_{\mathrm{g}}(f_\Theta), & \text{for } 1 > w_{\mathrm{g}}(f_\Theta) \geq w_{\mathrm{V}}^{\mathrm{min}} \\ 0, & \text{otherwise,} \end{cases} \tag{A.3}$$

where $w_{\mathrm{V}}^{\mathrm{min}}$ is a user-defined threshold. For example, with $w_{\mathrm{V}}^{\mathrm{min}} = 0.125$ all points along the spectral axis for which $|v_{\mathrm{LSR}}^{\mathrm{src}} - v_{\mathrm{LSR}}| > \Theta_{\mathrm{src}}$ receive a weight of $w_{\mathrm{V}} = 0$.

For the spatial association, we use the elliptical or circular FWHM extent of the sources as defined in the respective catalogue. With the following equation we can check whether a point is located within a rotated ellipse:

$$\epsilon = \frac{\left[\cos(\alpha)(x_{\mathrm{p}} - x_0) + \sin(\alpha)(y_{\mathrm{p}} - y_0)\right]^2}{a^2}$$
$$+ \frac{\left[\sin(\alpha)(x_{\mathrm{p}} - x_0) + \cos(\alpha)(y_{\mathrm{p}} - y_0)\right]^2}{b^2}, \tag{A.4}$$

where in our case x_{p} and y_{p} refer to the ℓ and b coordinates of our PPV data point and x_0, y_0, a, b, and α are the central ℓ and b coordinates, the semi-major and semi-minor axis, and the rotation of the ellipse, respectively. If $\epsilon \leq 1$, the point is located within or on the ellipse. Given that $\sqrt{\epsilon}$ corresponds to the distance to the centre of the source expressed in fractions of its spatial FWHM extent, we can define a weight w_{PP} for the association of Galactic or position-position coordinates with a catalogue source as follows:

$$w_{\mathrm{PP}} = \begin{cases} 1, & \text{for } w_{\mathrm{g}}(\sqrt{\epsilon}) \geq 1 \\ w_{\mathrm{g}}(\sqrt{\epsilon}), & \text{for } 1 > w_{\mathrm{g}}(\sqrt{\epsilon}) \geq w_{\mathrm{PP}}^{\mathrm{min}} \\ 0, & \text{otherwise,} \end{cases} \tag{A.5}$$

where $w_{\mathrm{PP}}^{\mathrm{min}}$ is a user-defined threshold. For example, with $w_{\mathrm{PP}}^{\mathrm{min}} = 0.125$ all points located beyond twice the extent of the source receive a weight of $w_{\mathrm{PP}} = 0$.

We then combine the spectral and spatial weights to a total PPV weight of $w_{\mathrm{PPV}} = w_{\mathrm{PP}} \cdot w_{\mathrm{V}}$. We only retain catalogue sources that have a weight $w_{\mathrm{PPV}} > 0$. If there are no sources from the catalogue that have $w_{\mathrm{PPV}} > 0$ we proceed to the next catalogue and repeat the source association.

If there was at least one catalogue source associated with the coordinate, we calculate w_{KDA}, which gives the weight that the source is located on the near or far distance. The w_{KDA} weight can range from -0.5 (which puts all the weight on the near distance) to 0.5 (which puts all the weight on the far distance). In case of only a single associated catalogue source $w_{\mathrm{KDA}} = w_{\mathrm{PPV}} \cdot f_{\mathrm{N/F}} \cdot w_{\mathrm{CAT}}$, where $f_{\mathrm{N/F}}$ is -0.5 or 0.5 if the catalogue source is associated with the near or far KD solution, respectively, and w_{CAT} is a user-defined weight for the catalogue (see Appendix A.3). If multiple sources are associated, we calculate w_{KDA} as a weighted average:

$$w_{\mathrm{KDA}} = \frac{\sum_{i=1}^{N_{\mathrm{src}}} w_{\mathrm{PPV,i}} \cdot f_{\mathrm{N/F,i}}}{\sum_{i=1}^{N_{\mathrm{src}}} w_{\mathrm{PPV,i}}} \cdot w_{\mathrm{CAT}}. \tag{A.6}$$

After we searched all catalogues, we use the w_{KDA} with the highest absolute value as our final value to determine our prior for the kinematic distance solution with $P_{\mathrm{far}} = 0.5 + w_{\mathrm{KDA}}$. We did not make the final calculation of the weights cumulative or additive as many of the catalogues we use are not fully independent

from each other. For example, the CHIMPS catalogue used many of the other catalogues in resolving the KDA for their clumps. In case our routine yields multiple w_{KDA} values with the same absolute value, we calculate the average w_{KDA} value from these solutions. Thus, if for the same clump two catalogues determined different KDA solutions, the resulting P_{far} value is 0.5, which means that no prior for the KDA is supplied.

A.2. Inferring literature KDA solutions

We use the literature distance information only for prior information on whether the source is located on the near or far side of the Galactic disc. In case one of the catalogues already supplied an information about the KDA resolution (that means if the near or far distance was chosen) we adopt these KDA results. If the KDA resolution was not stated explicitly, we use one of the following two methods to infer it.

Method 1. If the corresponding catalogue only gives the information about the heliocentric distance $d_\odot$ without any further information on the KDA, we use the relation

$$d_\odot = R_0 \cos \ell \pm \sqrt{R_{gal}^2 - (R_0 \sin \ell)^2} \qquad (A.7)$$

to obtain the Galactocentric radius R_{gal}. With R_{gal}, we then establish the near and far kinematic distances, from which we determine the chosen KDA solution. We only attempt to resolve the KDA if the near and far kinematic distances differ by more than 1 kpc; otherwise we remove the source from the catalogue.

Method 2. In case the KDA cannot be inferred via Method 1, we use functions contained in the BDC tool to calculate kinematic distance solutions and expected tangent point velocities v_{LSR}^{TP} for the catalogue sources. If the v_{LSR} velocity of a catalogue source is within 10 km s^{-1} or is higher than v_{LSR}^{TP}, we assume that the source is too close to the tangent point to resolve the KDA properly. These sources are subsequently removed from the catalogue.

Next we check for clear KDA solutions. If the v_{LSR} velocity of the source is lower than the near KD solution or higher than the far KD solution we resolve the KDA as "N" (near) or "F" (far), respectively.

For the remaining sources we compare the given catalogue distance to the kinematic distance solutions and the tangent point distance. In case the difference of the literature distance to the tangent point is lower than the differences to both of the kinematic distance solutions, we do not attempt to resolve the KDA and remove the source from the catalogue. Otherwise, we choose the kinematic distance solution that has the smallest difference to the literature distance.

A.3. Literature results used in the KDA prior

We now discuss the catalogues that were used in this work to infer priors for the KDA. We required that the catalogues contain information about molecular gas v_{LSR} velocities, either reported directly or reported indirectly via a given kinematic distance. We only retain the catalogue sources for which we could infer whether the near or far distance solution was chosen; we exclude all catalogue entries which were assigned tangent point distances or which had uncertain KDA resolutions. For this work we did not attempt to incorporate distances obtained from alternative distance estimation methods, such as dust extinction mapping,

if there was no information about the association with molecular gas. We also chose not to include catalogues based on ^{12}CO (1–0) observations.

Where available, we incorporated measured ^{13}CO (1–0) linewidths for the catalogue sources. If these did not exist, we either settled on linewidth measurements of a higher-density tracer or used median linewidth measurements obtained for comparable sources (i.e. clumps or clouds). We use a w_V^{min} threshold of 0.125 for all catalogues, which means that we make a spectral association with a catalogue source if $|v_{LSR}^{src} - v_{LSR}| < \Theta_{src}$. We make a spatial association with sources from the clump catalogues if the (ℓ, b) coordinate of the PPV point is contained within twice the extent of the clump (i.e. $w_{PP}^{min} = 0.125$). For the more extended sources of the remaining catalogues, we only make a spatial association if the (ℓ, b) coordinate of the PPV point is located within or close to the elliptical or circular extent of the source (with $w_{PP}^{min} = 0.9$)

We give the clump catalogues a higher weight ($w_{CAT} = 0.75$) than the catalogues of more extended objects, such as molecular clouds and IRDCs ($w_{CAT} = 0.5$). This was done to favour the KDA information from the clumps on small scales, as the clumps are usually embedded within these more extended objects.

In Table A.1 we give an overview about the catalogues we used as KDA solutions in this work. N_{src}^{GRS} gives the number of sources overlapping with the GRS coverage, and N_Θ, N_N, and N_F give the corresponding numbers of sources with measured linewidths, near, and far KD solutions, respectively. The columns w_{PP}^{min}, w_V^{min}, and w_{CAT} list the weights we used for the association of catalogue sources (Appendix A.1); the KDAR column specifies if the KDA resolution was given in the respective catalogue or how we calculated it otherwise. In the last two columns we give the abbreviation for the catalogues we refer to further on in the text and list the main references used to obtain information about the location, size, velocity, and distance information of the catalogue sources. In the following we discuss the individual catalogues in more detail.

ATLASGAL clumps. Urquhart et al. (2018) presented distance results for clumps from the ATLASGAL survey in the inner Galactic plane ($|l| < 60°$, $|b| < 1.5°$). The catalogue does not contain explicit information on how the KDA was resolved but lists the kinematic distance solutions, the distance estimated with the BDC (v1), and the chosen distance. With that information we could infer the KDA information "N", "F") for 6317 clumps. Of these, 4457 clumps have linewidth measurements, of which 3139 were obtained from the SEDIGISM survey (Urquhart et al., in prep.), 668 measurements were taken from Urquhart et al. (2018), and 292 measurements were taken from Wienen et al. (2012). For the remaining clumps with resolved KDAs but missing linewidth information we assume a FWHM linewidth of 3.367 km s^{-1}, which corresponds to the median linewidth computed from the clumps with measurements. We took the size information for the ATLASGAL clumps from Urquhart et al. (2014). In total, 1745 ATLASGAL clumps with resolved KDAs overlap with the GRS coverage.

BGPS clumps (v1). Eden et al. (2012, 2013) presented KD determinations for clumps of the Bolocam Galactic Plane Survey (BGPS; Rosolowsky et al. 2010). We only use the sample of 165 BGPS sources whose distances were not inferred via an association with molecular clouds from Roman-Duval et al. (2009). We established the chosen KDA solution via Method 1, using $R_0 = 8.5$ kpc as assumed by Eden et al. (2012, 2013). We take the corresponding position and size information of these clumps

Table A.1. Overview of the catalogues used as KDA solutions in this work.

| Info	$N_{\mathrm{src}}^{\mathrm{GRS}}$	N_Θ	N_{N}	N_{F}	$w_{\mathrm{PP}}^{\min}$	$w_{\mathrm{V}}^{\min}$	w_{CAT}	KDAR	Abb.	Refs.
ATLASGAL clumps	1745	426	1221	524	0.125	0.125	0.75	See Appendix A.3	U+18	1, 2, 3
BGPS v1 clumps	146	0	105	41	0.125	0.125	0.75	Method 1	E+12	4, 5, 6
BGPS v2.1 clumps	1046	455	754	292	0.125	0.125	0.75	Given	EB+15	7, 8, 9
CHIMPS clumps	3294	3294	2318	976	0.125	0.125	0.75	Given	R+19	10
Hi-GAL clumps	4021	0	3536	485	0.125	0.125	0.75	Given	E+17	11
COHRS clouds	396	396	262	134	0.9	0.125	0.5	Method 2	C+19	12
GRS clouds	652	652	453	199	0.9	0.125	0.5	Method 1	RD+09	13, 14
GRS clouds (BGPS)	381	381	203	178	0.9	0.125	0.5	Method 1	BH14	15
MSX IRDCs	263	263	261	2	0.9	0.125	0.5	Method 2	S+06	16, 17, 18
GRS H II regions	169	169	49	120	0.9	0.125	0.5	Given	A+09	19, 20
WISE H II regions	351	0	72	279	0.9	0.125	0.5	Given	A+14	21
SNRs	23	0	17	6	0.9	0.125	0.5	Given	R+18	22, 23, 24, 25, 26

References. (1) Wienen et al. (2012); (2) Urquhart et al. (2014); (3) Urquhart et al. (2018); (4) Rosolowsky et al. (2010); (5) Eden et al. (2012); (6) Eden et al. (2013); (7) Ginsburg et al. (2013); (8) Ellsworth-Bowers et al. (2015); (9) Svoboda et al. (2016); (10) Rigby et al. (2019); (11) Elia et al. (2017); (12) Colombo et al. (2019); (13) Rathborne et al. (2009); (14) Roman-Duval et al. (2009); (15) Battisti & Heyer (2014); (16) Simon et al. (2006b); (17) Simon et al. (2006a); (18) Marshall et al. (2009); (19) Anderson et al. (2009); (20) Anderson & Bania (2009); (21) Anderson et al. (2014); (22) Leahy & Ranasinghe (2018); (23) Ranasinghe & Leahy (2018a); (24) Ranasinghe & Leahy (2018b); (25) Ranasinghe et al. (2018); (26) Green (2019).

from v1 of the BGPS catalogue (Rosolowsky et al. 2010). For the linewidth of the clumps we assume a value of $\Theta = 3.316$ km s^{-1}, which corresponds to the median linewidth of the BGPS v2.1 sample (see next paragraph).

BGPS clumps (v2.1). Ellsworth-Bowers et al. (2015) resolved the KDA for 1320 BGPS clumps. We took the corresponding position and size information from v2.1 of the BGPS catalogue (Ginsburg et al. 2013). Svoboda et al. (2016) determined linewidths for 610 of these clumps from NH_3 observations. For clumps the associated ^{13}CO emission has in general broader linewidths than the NH_3 emission (Wienen et al. 2012). We thus decided to multiply the measured NH_3 linewidths of Svoboda et al. (2016) by a factor of two, which is based on the difference found by Wienen et al. (2012). For the remaining clumps without measured linewidths we assume the median NH_3 linewidth value from the Svoboda et al. (2016) sample corrected by a factor two, which corresponds to 3.316 km s^{-1} and compares very well to the median value from the ATLASGAL sample.

CHIMPS clumps. Rigby et al. (2019) used literature information to resolve the KDA for their sample of clumps compiled from the $^{13}CO/C^{18}O$ ($J = 3$–2) Heterodyne Inner Milky Way Plane Survey (CHIMPS; Rigby et al. 2016). We use 3294 clumps that have the highest reliability flag and a resolved KDA. We take position, size, and spectral information of the clumps from Rigby et al. (2019).

Hi-GAL clumps. Elia et al. (2017) presented a compact source catalogue for the *Herschel* InfraRed Galactic Plane Survey (Hi-GAL), for which KDs were determined via the Brand & Blitz (1993) rotation curve. We only retain the sources for which an external indicator was used to solve the KD (flag "G"). The Elia et al. (2017) catalogue does not contain information about the associated v_{LSR} velocities of the sources. We used the function `brand_rotcurve.calc_vlsr` from the kinematic distance package presented in Wenger et al. (2018)[15] to infer the v_{LSR}

---

[15] https://ascl.net/1712.001
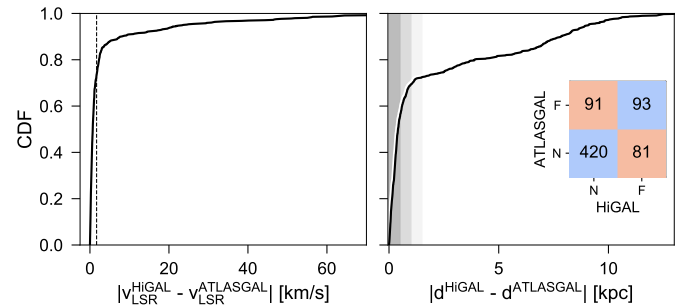
Fig. A.2. Comparison between 685 associated Hi-GAL and ATLASGAL sources (Urquhart et al. 2018). *Left*: CDF of the absolute difference between the inferred and measured v_{LSR} velocities for the Hi-GAL and ATLASGAL catalogues, respectively. The dashed vertical line indicates the assumed linewidth for the Hi-GAL sources. *Right*: CDF of the absolute difference of reported distances for Hi-GAL and associated ATLASGAL sources. The vertical grey areas indicate distance bins of 0.5 kpc. The inset shows the corresponding confusion matrix for the KDA resolution.

velocities for the longitude and distance values of the sources in the Elia et al. (2017) catalogue. In Fig. A.2 we benchmark these estimated Hi-GAL v_{LSR} velocities for 685 sources that we could associate with ATLASGAL clumps from Urquhart et al. (2018). We associate a Hi-GAL source with an ATLASGAL clump if their central positions are less than 19.2" apart, which corresponds to the ATLASGAL beam size. The left panel in Fig. A.2 shows that the majority of estimated Hi-GAL v_{LSR} velocities match very well with the associated ATLASGAL clump velocities. The fraction of sources for which the v_{LSR} values are not consistent is likely due to wrong associations of spectral lines with the dust features, as there can be multiple molecular gas emission features along the line of sight that the dust feature could be associated with. For this work, we make no attempt to resolve these inconsistent v_{LSR} values. The right panel in Fig. A.2 shows the differences in the estimated distances for the 685 associated sources. Again, for the majority of associated sources the distance values are similar. More important for our context, for

75% of the associated sources the KDA resolution is identical. We take the location and size information for the Hi-GAL clumps from Elia et al. (2017). For the linewidth, we assume for each clump the median linewidth estimated from the measured values from the ATLASGAL survey, which corresponds to $\Theta = 3.367$ km s^{-1}.

COHRS molecular clouds. Colombo et al. (2019) presented a cloud catalogue for the JCMT CO(3–2) High Resolution Survey (COHRS; Dempsey et al. 2013). We use their fiducial sample of 540 molecular clouds with well-defined distance estimations. We infer best-matching KDA solutions via Method 2, which yielded 396 KDA ("N", "F") solutions.

GRS molecular clouds. Roman-Duval et al. (2009) determined distances to 750 clouds from the catalogue of Rathborne et al. (2009) but do not explicitly specify the KDA resolution. Using Method 1 (with $R_0 = 8.5$ kpc), we were able to infer the KDA solutions for 652 clouds. We took the position, size, and spectral information of the clouds from Rathborne et al. (2009).

GRS molecular clouds crossmatched with BGPS clumps. Battisti & Heyer (2014) compiled a catalogue of 437 molecular clouds from the GRS survey, which they associated with BGPS sources from the v1 catalogue (Rosolowsky et al. 2010). They resolve the KDA for their sources but do not explicitly list whether they chose the near or far solution. Using Method 1 (with $R_0 = 8.3$ kpc), we could infer the chosen KDA solution for 389 clouds.

MSX IRDCs with velocities from GRS. Simon et al. (2006a) determined KDs to 313 IRDCs identified from observations of the Midcourse Space Experiment (Simon et al. 2006b), by morphologically matching the IRDCs to ^{13}CO (1–0) emission from the GRS. Simon et al. (2006a) argued that since the IRDCs are seen as extinction features they probably are located in the foreground and thus always resolve the KDA in favour of the near solution. Marshall et al. (2009) used an alternative approach based on modelling the three-dimensional distribution of interstellar extinction towards 115 of these IRDCs, which allowed them to obtain distances that do not suffer from near/far ambiguities. Whenever these were available, we chose the Marshall et al. (2009) distance determinations over the ones obtained by Simon et al. (2006a). Using Method 2 (Appendix A.2), we could resolve the KDA for 272 IRDCs. We take the v_{LSR} and FWHM linewidth information for this sample of IRDCs from Simon et al. (2006a) and take their position and size information from Simon et al. (2006b).

H II Regions associated with GRS emission. Anderson et al. (2009) associated 301 Galactic H II regions located within the GRS coverage with the corresponding ^{13}CO (1–0) properties. Anderson & Bania (2009) resolved the KDA for 266 of these H II regions using the H I emission/absorption and H I self-absorption methods. We include H II region sources as prior information if one of the following two conditions was fulfilled: the two methods yielded the same KDA resolution; or one of the methods received a high confidence label, in which case we use its KDA solution. We thus retained 169 sources, which had measured associated ^{13}CO (1–0) properties and a resolved KDA. We take the position, size, and spectral information of the sources from Anderson et al. (2009).

WISE H II Regions. Anderson et al. (2014) compiled an H II region catalogue from observations of the Wide-Field Infrared Survey Explorer (WISE). Using v2.0 of the catalogue, we select all H II regions with resolved KDAs; we also take the position, size, and v_{LSR} information from this version of the catalogue. For the spectral extent, we use the median linewidth value from the sample of H II regions of Anderson et al. (2009) (see previous paragraph), which corresponds to a value of $\Theta = 4.05$ km s^{-1}.

Supernova remnants. We include KDA solutions for 23 supernova remnants (SNRs) that have been obtained from H I 21 cm and GRS ^{13}CO line emission (Ranasinghe & Leahy 2018a,b; Ranasinghe et al. 2018). In case no information on the spatial extent was given in these works, we adopted the information given in Leahy & Ranasinghe (2018) and Green (2019). For the spectral extent, we assume the average linewidth of 3.6 km s^{-1} that Rathborne et al. (2009) find for their catalogue of molecular clouds of the GRS.

A.4. Effect of the KDA prior

In this section we discuss the effect of the KDA prior on the BDC results. For our tests we use the catalogues as detailed in Table A.1, so we include only sources that overlap with the GRS coverage and for which we could infer near or far KDA solutions.

First, we quantify the effect of the w_{CAT} weight on the distance estimation, for the two cases where the spiral arm priors are included ($P_{SA} = 0.5$) or switched off ($P_{SA} = 0$). For this test we use only the KDA information from the ATLASGAL sample and do not consider any of the other catalogues. We perform different distance runs with the BDC for the ATLASGAL sample; for each run, we supply the BDC with the correct KDA solutions for the sources and just vary the weight w_{CAT}, which determines the strength of the resulting P_{far} prior. For example, for $w_{CAT} = 0.5$ a far KDA solution yields $P_{far} = 0.75$, whereas with $w_{CAT} = 0.75$ this increases to $P_{far} = 0.875$. For $w_{CAT} = 1$ we would thus expect the highest correspondence between our calculated distances and the distances given in Urquhart et al. (2018)[16]. To test how robust the BDC results are against wrong KDA solutions, we also perform distance calculations for which we intentionally supply the incorrect KDA solutions for the ATLASGAL sample.

Table A.2 lists the performance of the BDC results for 20 runs, for which we vary w_{CAT} between the values $0, 0.25, 0.5, 0.75$, and 1, use or switch off the P_{SA} prior, and supply either the correct or incorrect KDA solutions. The $\delta(x)$ parameter gives the percentage of calculated distance values whose absolute error is within x kpc of the literature distances; we report $\delta(x)$ for x intervals of ± 0.5, ± 1.0, and ± 1.5 kpc. The runs with $w_{CAT} = 0$ correspond to the default BDC distance estimations that do not consider any prior information on the KDA and thus serve as our benchmarks.

The runs for which we supplied the correct KDA solution show a clear increase in the fraction of matching distances with increasing w_{CAT} value. This is expected, as KDA solutions are more enforced with increasing w_{CAT} values. In the runs where we use the prior for spiral arms ($P_{SA} = 0.5$), the percentage of matching distance values is less; however, the vast majority of the estimated distances is still close to the literature values.

Unsurprisingly, the supply of incorrect KDA solutions yields wrong distance estimates, especially the higher the w_{CAT} value, that means the more we enforce these KDA solutions. This effect is very strong for the runs where the spiral arm prior was switched off and the prior for the kinematic distances was

---

[16] However, since Urquhart et al. (2018) used an older version of the BDC (v1), we would not expect a perfect correspondence of the distance results even in this best case scenario.
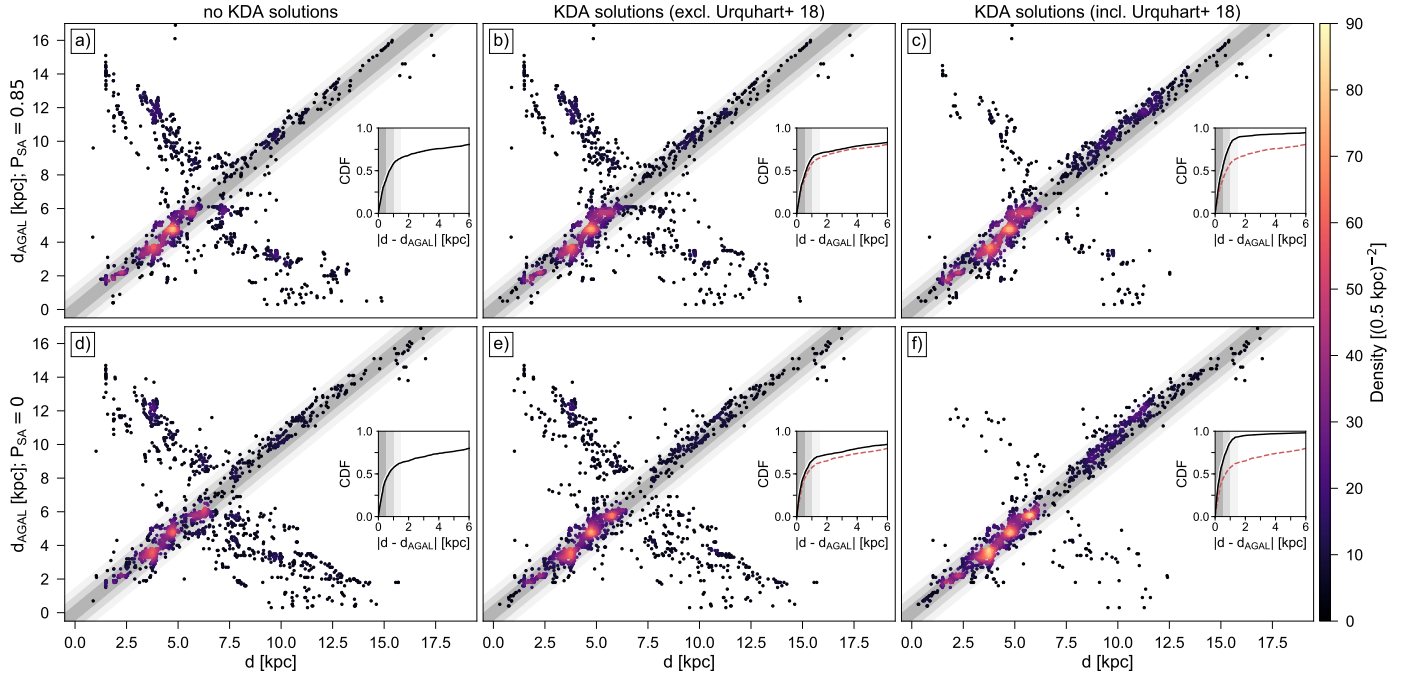
Fig. A.3. BDC results for the ATLASGAL clump sample plotted against their literature values (d_{AGAL}). The six panels correspond to different settings for the spiral arm and KDA priors. The points are colour-coded by their density. The insets show the cumulative distribution functions (CDFs) for the difference between the BDC results and the literature distances of the respective panels. The red dashed line in the insets of *panels b* and *c* and *panels e* and *f* correspond to the CDFs of *panel a* and *d*, respectively. The grey-shaded areas in the main panels and insets correspond to $|d - d_{AGAL}|$ intervals of 0.5, 1.0, and 1.5 kpc. See Appendix A.4 for more details.

Table A.2. BDC results for the ATLASGAL clump sample for different w_{CAT} and P_{SA} values and correct or incorrect KDA priors.

P_{SA}	w_{CAT}	Correct KDA			Incorrect KDA		
		$\delta(0.5)$ [%]	$\delta(1.0)$ [%]	$\delta(1.5)$ [%]	$\delta(0.5)$ [%]	$\delta(1.0)$ [%]	$\delta(1.5)$ [%]
0.5	0	36.2	56.3	62.7	36.2	56.3	62.7
0.5	0.25	40.9	63.7	70.1	33.4	49.7	55.0
0.5	0.5	46.6	70.8	77.5	28.1	43.6	47.6
0.5	0.75	50.1	75.6	83.3	21.0	32.0	35.9
0.5	1	55.1	83.7	92.6	7.6	10.6	14.2
0	0	37.4	52.3	58.9	37.4	52.3	58.9
0	0.25	44.5	62.6	72.3	23.8	32.3	36.7
0	0.5	53.5	77.0	84.8	16.6	20.7	23.7
0	0.75	59.5	82.7	91.1	8.9	10.7	12.8
0	1	58.3	81.9	91.6	2.4	3.0	5.4

Table A.3. BDC results for the ATLASGAL clump sample for different spiral arm and KDA priors.

KDA	$P_{SA} = 0.5$			$P_{SA} = 0$		
	$\delta(0.5)$ [%]	$\delta(1.0)$ [%]	$\delta(1.5)$ [%]	$\delta(0.5)$ [%]	$\delta(1.0)$ [%]	$\delta(1.5)$ [%]
No	39.1	58.5	65.1	43.6	57.4	63.0
excl. U+18	43.7	64.8	70.1	48.9	65.3	70.6
incl. U+18	53.7	81.0	89.1	63.7	86.8	93.4

Next, we want to quantify the effect of using all available KDA information with their corresponding weights (Table A.1). Using the ATLASGAL sample, we again perform different BDC runs with $P_{SA} = 0.5$ and $P_{SA} = 0$. For the KDA prior we either use none of the catalogues (giving us a benchmark for the default BDC performance), all KDA solutions from Table A.1 excluding the ATLASGAL catalogue itself, and all KDA solutions including the ATLASGAL catalogue. Figure A.3 shows the BDC results plotted against the literature values for these six BDC runs. Table A.3 gives the corresponding percentages of matching distance values within ranges of $\delta(x)$ (with $x = 0.5, 1$, and 1.5 kpc) that are highlighted with the grey-shaded areas in Fig. A.3. The runs which use no KDA solutions (panels a and d) show a large dichotomy between near and far distance solutions. Including all KDA solutions apart from the Urquhart et al. (2018) catalogue itself already manages to improve the correspondence between the distances and indicates that the sources of the remaining catalogues overlap with many of the ATLASGAL clumps. Finally, as expected, the inclusion of the Urquhart et al. (2018) catalogue leads to the best correspondence of the distance results. However, for many sources the spiral arm

the dominating factor for the BDC results. Using the prior for the spiral arms can mitigate the negative impact of the incorrect KDA solutions; for $w_{CAT} = 0.5$ almost half the sources have distance results within 1.5 kpc to the literature distances even though we intentionally forced the BDC to prioritise the wrong KDA solution.

This test demonstrated that the BDC run with $P_{SA} = 0.5$ is more robust against priors using wrong KDA solutions than the run with $P_{SA} = 0$. Moreover, we find that w_{CAT} values of 0.5 to 0.75 are preferable values for the catalogue weights, as they offer a good balance between recovery of correct distances with the right KDA solutions for $P_{SA} = 0$ and robustness against incorrect distances with the wrong KDA solutions for $P_{SA} = 0.5$.

Table A.4. BDC results for the remaining catalogues from Table A.1 for different spiral arm and KDA priors.

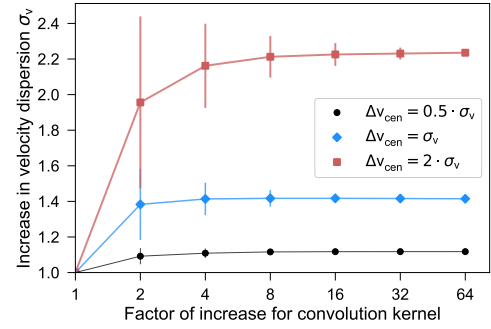|	$P_{SA} = 0.5$			$P_{SA} = 0$		
Abb.	$\delta(1.5)$ [%]	$\delta^-(1.5)$ [%]	$\delta^+(1.5)$ [%]	$\delta(1.5)$ [%]	$\delta^-(1.5)$ [%]	$\delta^+(1.5)$ [%]
E+12	21.2	45.9	80.8	33.6	67.8	91.1
EB+15	56.6	80.2	88.5	59.1	87.1	92.4
R+19	50.6	79.8	91.3	63.5	90.6	95.7
E+17	51.3	80.6	91.4	62.6	89.3	95.9
C+19	47.0	69.4	74.2	52.3	78.8	84.6
RD+09	38.8	69.3	80.1	52.5	82.5	92.0
BH14	29.8	54.0	65.0	36.0	66.1	76.1
S+06	44.5	74.1	87.5	51.0	76.0	89.4
A+09	25.4	50.3	63.3	30.2	56.8	69.2
A+14	32.2	58.7	68.7	41.9	72.1	83.2
R+18	43.5	78.3	78.3	60.9	91.3	91.3

Fig. B.1. Increase in observed velocity dispersion with decreasing spatial resolution. The points and errorbars show the results of 100 different realisations of a mock PPV cube containing emission lines with identical σ_v values; errorbars indicate 1σ intervals. For each realisation, the standard deviation for the variation of emission line centroids (Δv_{cen}) was varied between 0.5, 1, or 2 times σ_v (indicated in black, blue, and red, respectively). See Appendix B for more details.

priors lead to a preference of different KDA resolutions (panel c). By switching the spiral arm prior off, we manage to drastically reduce the instances for which a different KDA solution was favoured. The remaining fraction of clumps for which a different KDA solution was chosen was due to mismatching KDA solutions from different catalogues. For example, Urquhart et al. (2018) give a near distance solution of 3.5 kpc for the clump AGAL020.662-00.139, whereas spatially and spectrally overlapping sources in five other catalogues (A+09, RD+09, BH14, EB+15, E+17) favour a far distance solution, leading to a BDC value of 10.1 kpc.

Finally, we list the performance of the BDC results for the remaining catalogues used for KDA information in Table A.4. We again perform different BDC runs with $P_{SA} = 0.5$ and $P_{SA} = 0$ and give the percentage of sources for which the distance was within a range of 1.5 kpc to the literature distance for the cases where no literature KDA solution is used (labelled δ), and KDA solutions from all catalogues excluding and including the one for which the distances are calculated (labelled δ^- and δ^+, respectively). We see already an improvement in matching distances for the δ^- runs, which indicates that there is a good overlap between sources from all catalogues. As expected, we see the highest correspondence between the BDC and literature distance results for the runs in which the spiral arm prior is switched off. In conclusion, our tests showed that the BDC runs with supplied literature KDA solutions are able to match the vast majority of distance results from each of the individual catalogues used to infer KDA priors. This result is not self-evident, given that many of these catalogues use different assumptions about the rotation curve parameters. We thus infer that our obtained distance results are consistent with the vast majority of the literature results.

Appendix B: Effects of beam averaging on the observed linewidth

We designed the following simplified experiment to test how fluctuations of the line centroids can broaden the linewidth via beam averaging effects. We perform different runs for which we vary the spatial resolution of a given PPV cube to simulate observations of regions at different heliocentric distances. For each run we construct PPV cubes with dimension (100, 100, 30)

and populate each of the spectra with a single Gaussian component, whose velocity dispersion is two spectral channels. For simplicity we do not assume any noise. The mean position of the Gaussian is centred along the spectral axis with the centroid of the spectrum allowed to vary for each component; the standard deviation of this variation (Δv_{cen}) is set either to 0.5, 1, or 2 times the velocity dispersion. Assuming the pixel size to be equal to the FWHM of the resolution element or beam, we convolve this cube with a 2D Gaussian kernel whose FWHM is set to either 2, 4, 8, 16, 32, or 64 times the pixel size, thus simulating observations of regions at 2 to 64 times the distance of our original cube, which corresponds approximately to the variation present in the GRS. We then determine the velocity dispersion of a Gaussian fit to the central spectrum of the spatially smoothed cube.

Figure B.1 presents the results for 100 different realisations of the PPV cube and shows that the measured σ_v in the spatially smoothed cubes increases significantly with increasing Δv_{cen}. Variations of Δv_{cen} of the order of the velocity dispersion of the emission line in the resolved cube lead to increases in σ_v in the spatially unresolved cubes by a factor of ~1.4.

In real observations, Δv_{cen} will not be distributed randomly. Rather, the distribution of line centroids is observed to be highly structured (Henshaw et al. 2020; Riener et al. 2020), with coherent gradients, which will result in similar effects as in our simplified case. Moreover, variation in the non-thermal contribution to the linewidth can lead to an additional broadening of the lines at coarser spatial resolution. We thus conclude that due to beam averaging effects it is very unlikely that we observe the same population of linewidths in regions located at far distances as in nearby regions, whose emission lines are spatially better resolved.

Appendix C: Characterisation of the BDC performance

Here we give some further details about the performance of the BDC. We also discuss the effect of the KDA and size-linewidth priors (Sects. 3.2 and 3.3) on our final distance results.

C.1. Effect of v_{LSR} uncertainties

Version 2.4 of the BDC allows to supply uncertainties for the v_{LSR} measurement, which can have strong effects on the distance
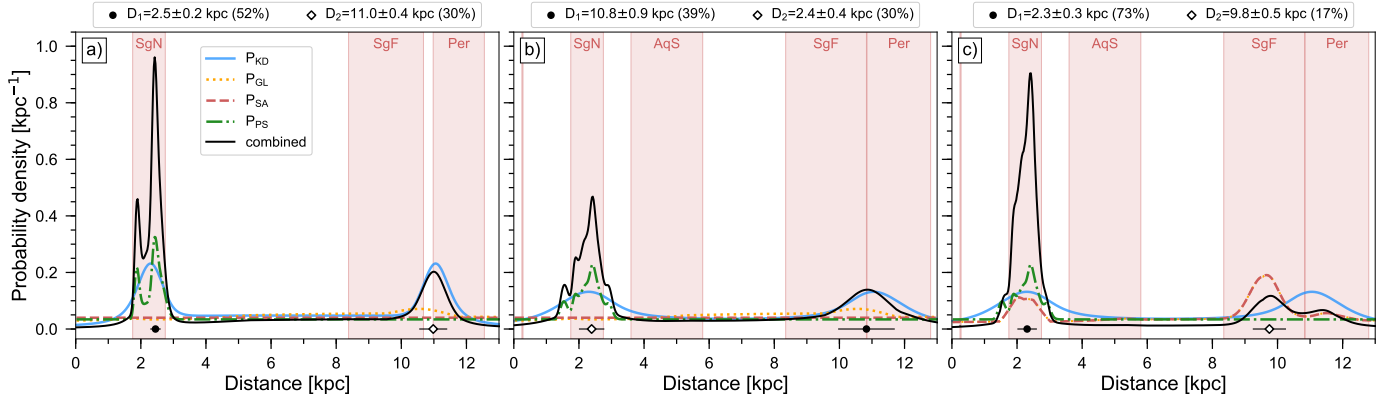
Fig. C.1. Effect of v_{LSR} uncertainties on BDC results for a source located at $\ell = 35°$, $b = 0.1°$, and $v_{LSR} = 40$ km s^{-1}. The BDC settings are the same for *panels a* and *b*; for *panel c* we included the prior for spiral arms. We increase Δv_{LSR} from 1 km s^{-1} (*panel a*) to 10 km s^{-1} (*panels b* and *c*). The meaning of the lines and symbols is the same as in Fig. 2. See Appendix C.1 for more details.

estimation (Fig. C.1). For each fit component, we chose either its estimated v_{LSR} uncertainty or its σ_v value for Δv_{LSR}, whichever was the higher value. The median uncertainty of the Δv_{LSR} values for all fit components was 1.1 km s^{-1}, with an IQR of 0.7 to 1.9 km s^{-1}.

Figure C.1 illustrates the effect of Δv_{LSR} on the distance results. The first two panels (a, b) show distance results obtained with identical BDC settings, where only the supplied uncertainty on the v_{LSR} coordinate was different. Increasing the v_{LSR} uncertainty has multiple effects: the KD peaks get broadened and the association with parallax sources is increased. In our example this causes a shift of the estimated most likely distance from the near to the far KD solution. Finally, panel c illustrates the effect of Δv_{LSR} on the association with spiral arms. Larger Δv_{LSR} values lead to an increase in associations with Galactic features. In our example it led to the consideration of the Aquila Spur and Aquila Rift as possible candidates for an association; however this had only a very limited effect on the combined distance PDF.

C.2. Distance results without P_{far} priors

Here we discuss the impact of the KDA and size-linewidth priors (see Sects. 3.2 and 3.3) on our distance results. For this purpose we created four more distance runs with the BDC:

– Run C: uses the default settings of the BDC, that means $P_{SA} = 0.85$, but does not include the KDA and size-linewidth priors.
– Run D: same as Run C, but with the SA prior switched off ($P_{SA} = 0$), which also sets $P_{GL} = 0$ as these two priors are combined in the default version of the BDC.
– Run E: same as Run A, but without the size-linewidth prior.
– Run F: same as Run B, but without the size-linewidth prior.

Figure C.2 shows the map of W_{CO} values for Run C and D. A comparison with Fig. 8 reveals substantial differences to Run A and B, respectively. The default BDC settings in Run C lead to a much stronger association with the SA model and the results contain much less emission at close distances ($d < 2$ kpc). However, compared to Run A, Run C does not put emission in between the Perseus and Outer arm. Run D shows that without the KDA and GL priors the distance results contain a higher fraction of emission-free areas, which is especially notable around the far portion of the Sagittarius arm. Without the size-linewidth prior, Run D also puts significantly more emission from close distances
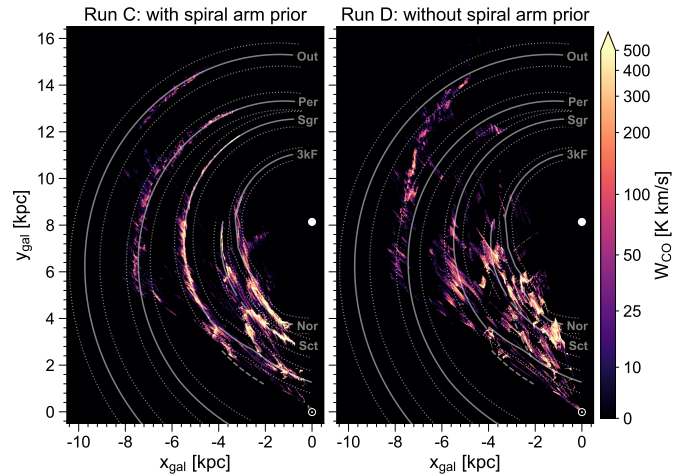

Fig. C.2. Same as Fig. 8, but for the distance results obtained with the default settings of the BDC. When displayed in Adobe Acrobat, it is possible to switch to the original map (Fig. 8), hide the spiral arm positions and show the grid.

towards the Perseus and Outer arm regions. In terms of association with Galactic features: for Run C 13.8% of the emission (and 18.6% of the fit components) were associated with interarm regions, which increased to 25.6% (34.6% of N_{comp}) for Run D.

Figure C.3 shows the resulting σ_v map if we had not used the size-linewidth prior. Since we use this prior only for emission with $v_{LSR} < 20$ km s^{-1}, only positions at the largest R_{gal} values are affected by it. A comparison with Run A and B shows that the size-linewidth prior helps in decreasing the confusion between near and far emission that causes the large fraction of emission lines with narrow linewidths at large distances in Run E and F.

C.3. Estimated probabilities, P_{far} weights, and distance choices

Figure C.4 shows a face-on view of the median estimated probability values and the left panel in Fig. C.5 gives the cumulative distribution of all estimated probabilities. These probabilities were estimated from the integrated areas of Gaussian fit components to the combined distance PDF (Sect. 3.4). We thus get higher probabilities for regions where the combined distance PDF produced a dominant peak, which however could be caused
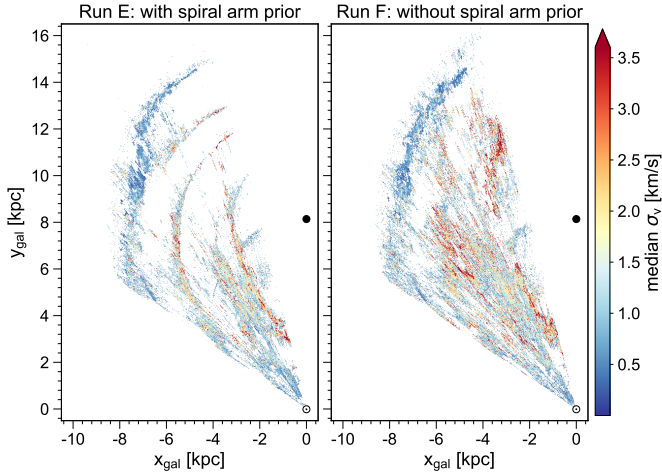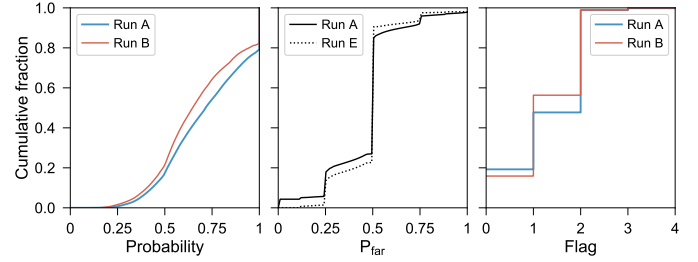
Fig. C.3. Face-on view of the median velocity dispersion values of Gaussian fit components for the BDC results obtained with (*left*) and without (*right*) the spiral arm prior. The values are binned in 10×10 pc cells and the median was calculated along the z_{gal} axis. The position of the Sun and Galactic centre are indicated by the Sun symbol and black dot, respectively. When displayed in Adobe Acrobat, it is possible to switch to the original map (Fig. 10), show the spiral arm positions and the grid.
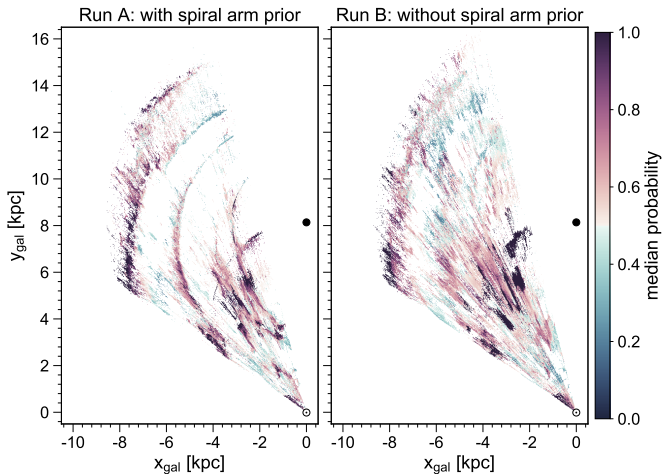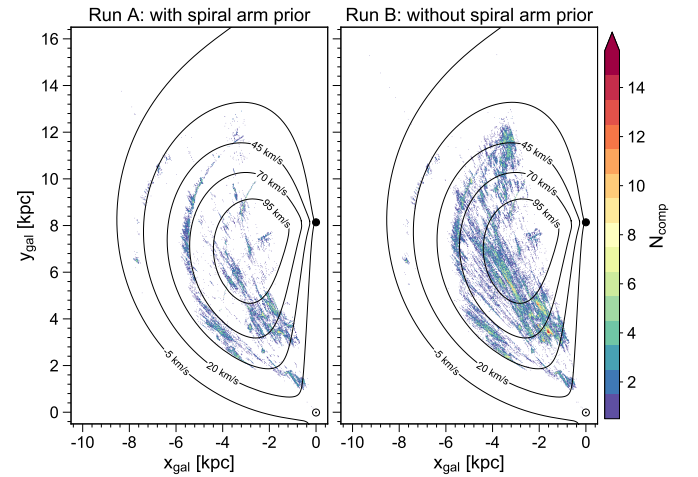

Fig. C.4. Face-on view of the median probability values from the BDC results obtained with (*left*) and without (*right*) the spiral arm prior. The values are binned in 10×10 pc cells and the median was calculated along the z_{gal} axis. The position of the Sun and Galactic centre are indicated by the Sun symbol and black dot, respectively. When displayed in Adobe Acrobat, it is possible to show the spiral arm positions and the grid.

by negative effects. For example, the near distance solution of the KD prior is cut off for low v_{LSR} values (cf. Fig. 3), thus yielding far distance solutions with high estimated probabilities. The KD prior is also down-weighted for lower R_{gal} values, which could lead to strongly blended KD peaks. This could result in broad Gaussian fits over both of these peaks, with high distance uncertainties as well as high estimated probability, which seems to occur at the lowest R_{gal} values (cf. Fig. 20).

The cumulative distribution of the assigned probabilities for the distance values (left panel of Fig. C.5) shows that the chosen distance values from the BDC run with the SA prior have

Fig. C.5. Cumulative distributions for the estimated distance probabilities (*left*), values for the P_{far} prior (*middle*), and choice of the distance values (*right*). See Sect. 4.4 for more details.

Fig. C.6. Face-on view of the number of components with distance results that cause an absolute v_{LSR} deviation of more than 10 km s^{-1} compared to the Galactic rotation curve model for the BDC results obtained with (*left*) and without (*right*) the spiral arm prior. The values are binned in 10×10 pc cells and the values were summed up along the z_{gal} axis. The position of the Sun and Galactic centre are indicated by the Sun symbol and black dot, respectively. When displayed in Adobe Acrobat, it is possible to show the spiral arm positions, hide the curves of constant projected v_{LSR}, and show the grid.

higher associated probabilities. For Run A, 43.8% of the chosen distance values had probabilities >0.75 and 16.7% of the distance values had probabilities <0.5; for Run B, these fractions change to 33.6 and 21.0%, respectively. Thus for Run A the Gaussian fits to the combined distance PDF had higher integrated areas, confirming that the addition of the SA prior leads to more well-defined peaks.

The middle panel of Fig. C.5 shows the cumulative distribution of chosen weights for the P_{far} prior for the case where the KDA prior was used (dotted line) and the case where the size-linewidth prior (Sect. 3.3) was used in addition to the KDA (solid line). For the case in which P_{far} was only informed by the KDA prior, 23.1 and 7.6% of the fit components received a preference for the near and far KD solution, respectively. If the size-linewidth prior is used in addition to the KDA prior, these percentages increase to 27.5 and 13.3%, respectively.

Finally, the right panel of Fig. C.5 shows the cumulative distribution for the choice of distance values (Sect. 3.4). The numbers indicate the following conditions: the distance assignment yielded only one distance solution (0); the associated Gaussian fit of one distance solution did not satisfy the criterion

for the amplitude threshold (1); the distance solution with the highest probability (i.e. the highest integrated intensity of the associated Gaussian fit) was chosen (2); the distance solution with the lowest absolute distance error was chosen (3); and the near KD solution was picked randomly (4). The cumulative distribution shows that conditions (0), (1), and (2) were responsible for the vast majority of final distance choices (contributing 19.2, 28.5, and 51.2% for Run A and 15.9, 40.4, and 42.6% for Run B), whereas conditions (3) and (4) only contributed minimally.

C.4. Deviation from Galactic rotation curve

Here we quantify the deviation of our distance results from the expected values from the Galactic rotation curve model. These deviations are interesting as they can identify regions with high peculiar velocities. Figure C.6 shows the number of fit components whose estimated distance values caused a deviation of more than 10 km s^{-1} from the expected v_{LSR} value based on the Galactic rotation curve. In Run A and B, 2.2 and 6.2% of the components showed such a large v_{LSR} deviation. Since both panels show similar deviations occurring around the positions of the Sagittarius, Scutum, and Norma arms, it is likely that these differences from pure rotation curve velocities are to a large part due to the effect of the maser parallax sources (cf. Fig. 1). For Run B we can also identify an increase in the number of components with deviating v_{LSR} values inside $R_{gal} \lesssim 5$ kpc. For this region the BDC downweighted the KD prior, so these deviating components might to a large part simply be due to large associated distance uncertainties. A comparison with Fig. 20 shows that these regions are indeed associated with increased uncertainty values.