

Bayesian inference of three-dimensional gas maps: Galactic CO

Philipp Mertsch^{a,*} and Andrea Vittino^a

^a*Institute for Theoretical Particle Physics and Cosmology (TTK),
RWTH Aachen University, 52056 Aachen, Germany*
E-mail: pmertsch@physik.rwth-aachen.de

The three-dimensional distribution of both atomic and molecular gas in the Galaxy is a crucial modelling input, both for the generation of diffuse emission in gamma-rays and the transport of cosmic rays. Here, we present new 3D maps of molecular hydrogen based on the Dame et al. (2001) CO survey compilation. We consider the deprojection as a Bayesian variational inference problem. The posterior distribution of the gas densities allows us to estimate both the mean and uncertainty of the reconstructed density. Unlike most of the previous attempts, we take into account the correlations of gas on a variety of scales which allows curing some of the well-known pathologies, like fingers-of-god effects. Both gas flow models that we adopt incorporate a Galactic bar which induces radial motions in the inner few kiloparsecs and thus offers spectral resolution towards the Galactic centre. We compare our gas maps with those of earlier studies and characterise their statistical properties, e.g. the radial profile of the average surface mass density.

37th International Cosmic Ray Conference (ICRC 2021)
July 12th – 23rd, 2021
Online – Berlin, Germany

*Presenter

1. Introduction

About nine out of ten gamma-ray photons observed at GeV energies are diffuse, that is they have not been produced inside gamma-ray sources but in the interstellar medium (ISM) of the Galaxy. Together with charged particles misidentified as gamma-rays, Galactic diffuse emission constitutes the dominant background. Oftentimes, conventional background subtraction is not efficient in mitigating this background when studying sources, for instance for sources with a soft spectrum or for extended sources. The only way of dealing with this is through careful modelling and subsequent subtraction, but diffuse emission is of course interesting in its own right as it contains information on the spatial and spectral distribution of cosmic rays.

Such modelling requires detailed information on the spectral and spatial distribution of the gamma-ray emissivity. For galactic diffuse emission at GeV energies, the dominant contribution for large regions of the sky is π^0 production and subsequent decay into gamma-rays. The π^0 's in turn, are produced by inelastic collision of cosmic ray nuclei with the gas in the ISM. This gas is predominantly hydrogen, either in the atomic phase (HI) or molecular phase (H_2), with only a small admixture of ionised hydrogen (HII). Modelling of Galactic diffuse emission therefore requires detailed information on HI and H_2 in the Galaxy.

While HI is famously emitting at 1.4 GHz through the hyperfine transition, H_2 does not. Instead, it is believed that carbon monoxide (CO) is well-mixed with H_2 and can therefore be used as a tracer [1]. ^{12}CO emits in various lines, with the $J = 1 \rightarrow 0$ transition at 115 GHz the best-studied one. Thanks to galactic rotation, the emission gets Doppler-shifted by varying amounts along a line of sight. What is observed are therefore spectra that encode the distribution along various lines of sight. If the velocity v_{LSR} along a line of sight and hence the Doppler shift was known as a function of position, it could be used to deproject an observed spectrum and gain 3D information on the distribution of emission.

However, a number of issues remain:

1. Kinematic distance ambiguity: For the inner Galaxy there can be two solutions for the distance from the observer for a given v_{LSR} , even under the assumption of perfectly circular rotation.
2. Lack of kinematic resolution: Again for circular rotation, there is no component of velocity along the line of sight for directions towards the Galactic centre (longitude $\ell \sim 0^\circ$) and anti-centre ($\ell \sim 180^\circ$).
3. Small-scale variations of velocities: Winds and supernova explosions lead to perturbations of the regular velocity field of (circular) rotation, thus affecting the distance estimates.

Traditionally, data from gas line surveys has been deprojected into galacto-centric rings. For axisymmetric distributions of cosmic rays (the other ingredient in producing the gamma-ray emissivity), the kinematic distance ambiguity does not matter. The rings are also rather wide, thus the issues with small-scale variations of velocities get mitigated. However, comparison of the modelled diffuse emission with observations by *Fermi*-LAT have shown that significant residuals remain [2]. Alternatively, forward modelling has been proposed [3], but currently these models provide only upper limits on the gas density. There are a number of deprojections of the gas spectra into 3D distributions of gas [4, 5], but even studies that build on particular gas flow model show artefacts,

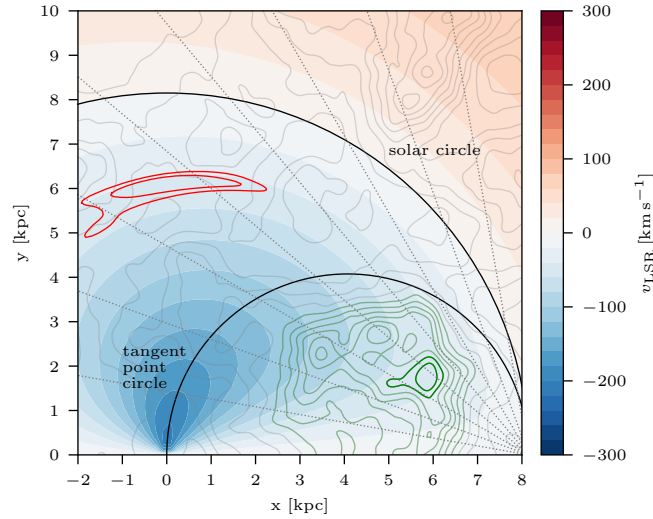


Figure 1: A part of the Galactic plane. The shaded areas indicate the line-of-sight velocity as seen by an observer at the position of the Sun, $(x, y) = (8.15, 0)$ kpc. The lines are equicontours of a simulated log-normal gas density. Consider the structure shown by the dark green contours, which happens to be the highest-density region of the simulated density. Based on the observations of a gas line survey one would not be able to tell if this emission was at the near distance (i.e. inside the tangent point circle) or at the far distance (i.e. outside the tangent point circle). In fact, the structure shown in red would look very similar in a gas line survey. However, this red structure is significantly distorted, e.g. very anisotropic, and would thus have a lower probability to be reconstructed in a Bayesian inference algorithm.

e.g. finger-of-god effects. What the latter studies largely ignore is that fact that the gas density in the Galaxy must be correlated, due to various physical processes: On the largest scales, the spiral structure is likely due to density waves of interstellar matter. On smaller scales, turbulence determines the correlation of the gas density.

Here, we propose using a Bayesian inference method that is able to take into account such correlations to deproject a gas line survey of CO to produce 3D maps of Galactic H_2 . This method has the potential to cure some of the above mentioned deficiencies of the kinematic deprojection. Note that we do not need to impose the exact correlation structure, e.g. the power spectrum, but have the data determine the power spectrum under some simplifying assumptions, e.g. log-normal statistics, isotropy and homogeneity. In addition, our reconstructed 2D gas maps readily come with uncertainty information, something that is difficult to estimate for the other reconstruction methods. An example for the power of correlations to resolve issues of the kinematic deprojection is provided in Fig. 1.

2. Methodology

2.1 Data

We have used the spectral cubes from the CfA compilation of CO surveys [6]. These data cover the entire galactic plane and virtually all other parts of the sky for which significant emission has been reported. The angular resolution of the constituent surveys varies between $1/16$ to $1/2$ of

a degree and we conservatively fix the noise level to 0.3 K per channel. We have updated the data for the now better-known velocity of the local standard of rest [7, 8]. Throughout, we assume an X_{CO} factor of 2×10^{20} molecules $\text{cm}^{-2} (\text{K km/s})^{-1}$.

2.2 Gas flow models

While assuming only circular rotation would be the most conservative assumption for a kinematic deprojection, its validity will be limited even on the largest galactic scales: We have overwhelming evidence for the presence of a galactic bar which will induce non-circular motion. While the details are everything but certain, we have adopted two models of the gas flow. The first is based on a smoothed particle hydrodynamics simulation [9, hereafter BEG03], and has already been used for deprojecting line survey data in Ref. [5]. In addition, we have generated our own gas flow model, based on the semi-analytical model of Ref. [10, hereafter SBM15] for gas-carrying orbits in a potential dominated by a Galactic bar. (See Ref. [11] for details).

2.3 Bayesian inference

Here, we have employed Metric Gaussian Variational Inference (MGVI) [12] for the reconstruction of the H_2 density. MGVI forms part of the larger set of methods known in the framework of information field theory (IFT). As a probabilistic method, MGVI does not provide one “best-fit” reconstruction, but quantifies the posterior of the reconstructed field, in this case of the gas density.

In order to derive the posterior in the adopted Bayesian framework, we need to specify a likelihood $p(T_{\ell b v} | \varepsilon_{xyz})$ and a prior $p(\varepsilon_{xyz})$. Here, $T_{\ell b v}$ denotes the data field, that is the observed brightness temperature, and ε_{xyz} the signal field, that is the gas density that is to be reconstructed. For the prior, we make the assumption that the gas density is a realisation of homogeneous and isotropic random field with log-normal statistics. In the harmonic domain, this is readily reflected by a 1D power spectrum and within MGVI, there is a rather flexible parametrisation that treats the power spectrum itself as a random field. We note that we do not need to enforce the detailed structure of the power spectrum; instead the data essentially determine the appropriate power spectrum.

As for the likelihood $p(T_{\ell b v} | \varepsilon_{xyz})$, we assume the following data model,

$$T_{\ell b v} = R[\varepsilon_{xyz}] + n_{\ell b v}, \quad (1)$$

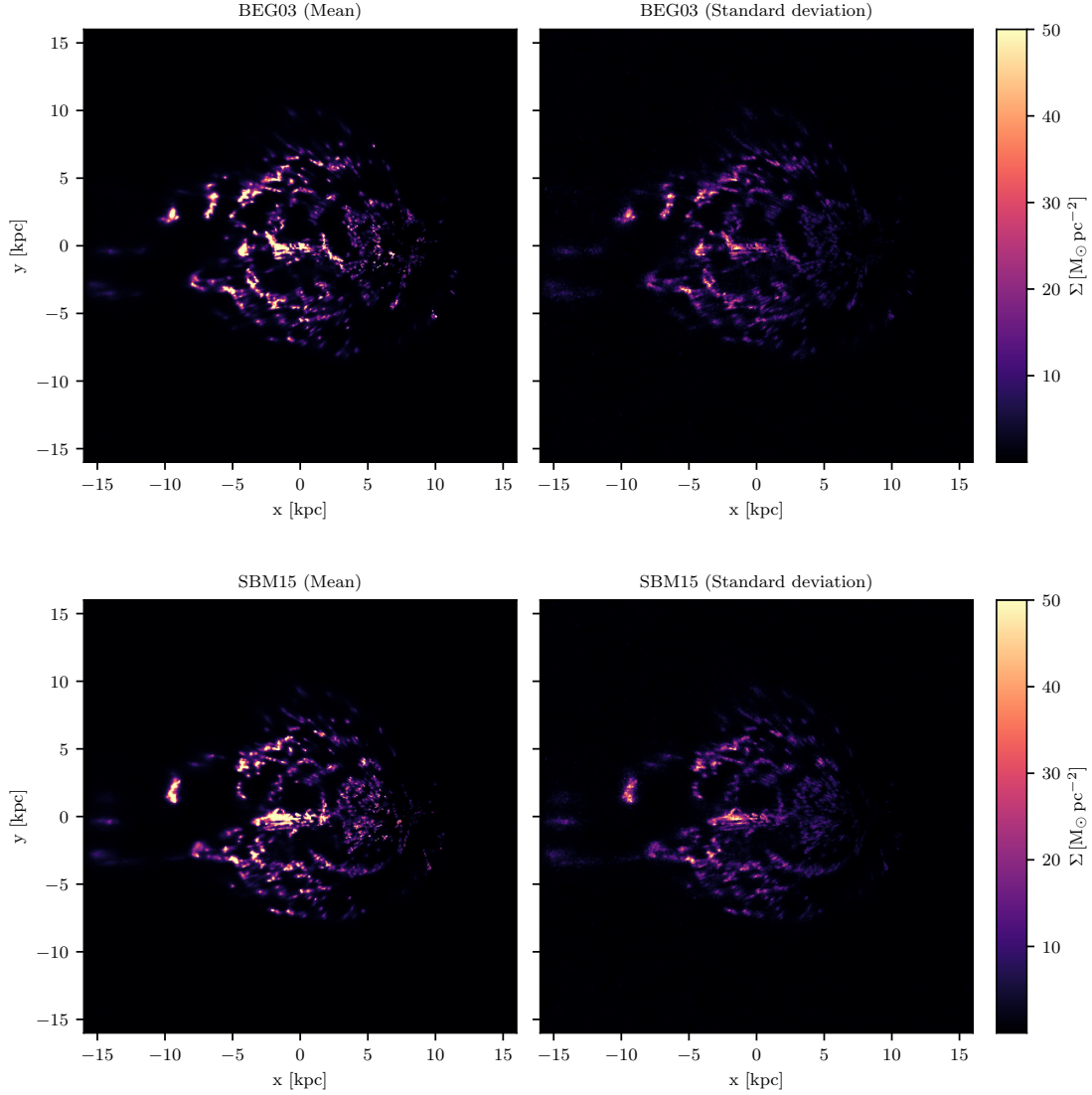
relating the observed data $T_{\ell b v}$ to the gas density ε_{xyz} in the presence of noise $n_{\ell b v}$. Here, R is a linear map from signal space (x, y, z) to data space (ℓ, b, v)

$$R[\varepsilon](\ell, b, v) = \int_0^\infty ds \varepsilon(\vec{r}) \delta(v - v_{\text{LSR}}(\vec{r})) \Big|_{\vec{r}=\vec{r}(\ell, b, s)}. \quad (2)$$

Given that the noise $n_{\ell b v}$ is Gaussian with covariance N , we can formulate a Gaussian likelihood, $\tilde{p}(T_{\ell b v} | \varepsilon_{xyz}) = \mathcal{G}(T_{\ell b v} - R[\varepsilon_{xyz}], N)$. However, we take into account the fact that the measured velocity can deviate from the one prescribed by the gas flow model due to thermal line width or turbulence, by smearing of the linear map with a velocity distribution, $p(T_{\ell b v} | \varepsilon_{xyz}) = \mathcal{G}(T_{\ell b v} - R'[\varepsilon_{xyz}], N)$ where

$$R'[\varepsilon_{xyz}] \equiv \int d\hat{v} \mathcal{G}(v - \hat{v}, \sigma_v^2) R[\varepsilon_{xyz}] = \int_0^\infty ds \varepsilon(\vec{r}) \mathcal{G}(v - v_{\text{LSR}}(\vec{r}), \sigma_v^2) \Big|_{\vec{r}=\vec{r}(\ell, b, s)}. \quad (3)$$

We adopt $\sigma_v = 5$ km/s throughout. For technical details, we refer the reader to Ref. [11]



POS (ICRG2021) 682

Figure 2: Surface mass density reconstructed for the two gas flow models adopted: **BEG03** (upper panels) and **SBM15** (lower panels). In the left column, we show the mean reconstructed density and in the right column its standard deviation.

3. Results

In Fig. 2, we show the surface mass density, that is the 3D gas density projected onto the Galactic plane for both gas flow models. In addition to the mean of the posterior, we also show the standard deviation which can be used as a measure of the uncertainty. The reconstruction shows structure on a range of scales as expected. Some structures align with spiral arms while others are likely artefacts. Total gas mass is $1.4 \times 10^9 M_\odot$ and $1.3 \times 10^9 M_\odot$ for the **BEG03** and **SBM15** gas flow models, respectively. Our maps are available at <https://dx.doi.org/10.5281/zenodo.4405437>.

In order to investigate this further, in Fig. 3 we provide maps of the signal-to-noise ratio (S/N)

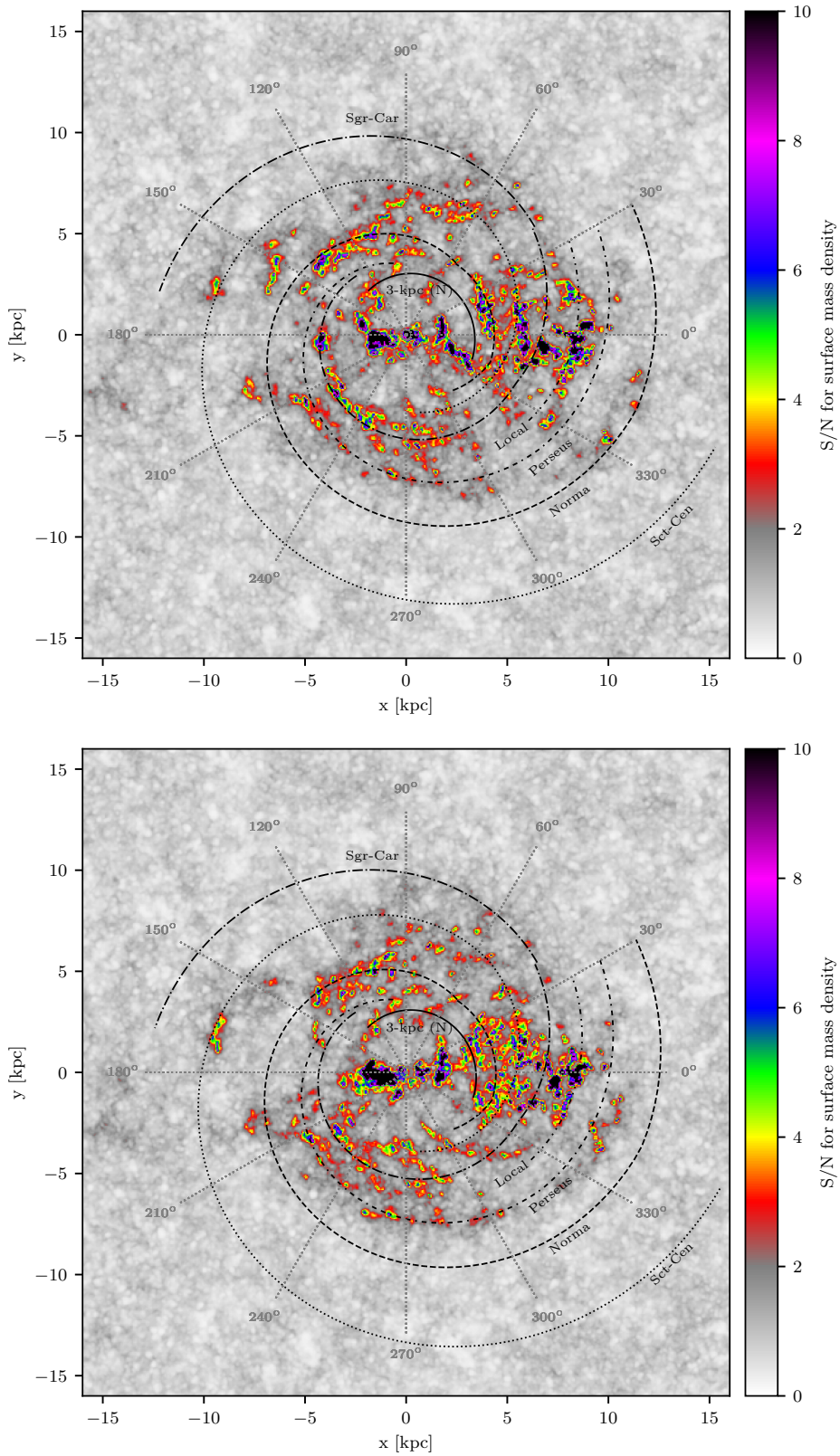


Figure 3: Maps of the signal-to-noise ratio (S/N) for the surface mass density in the two gas flow models adopted: **BEG03** (upper panel) and **SBM15** (lower panel). We also indicate a number of spiral arms and the galacto-centric azimuth φ .

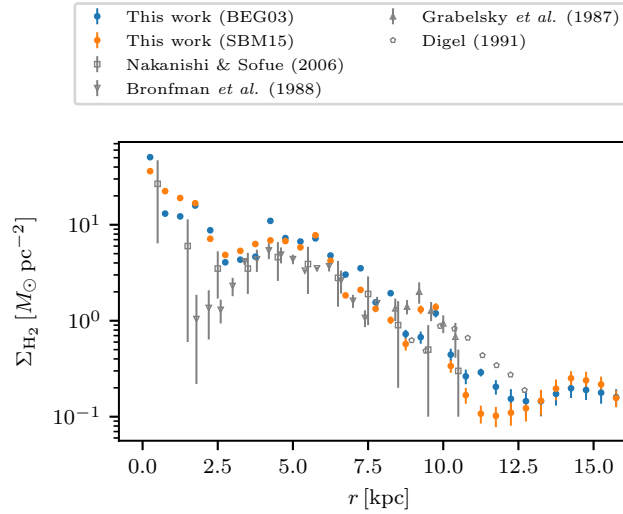


Figure 4: Surface mass density for the two gas-flow models (BEG03 and SBM15) and compared with earlier studies ([13], [14], [15] and [4]).

defined as the mean divided by the standard deviation. Much of the emission has been resolved into regions with high S/N and much of it along spiral arms. This is most clearly seen along the Perseus arm for galacto-centric azimuths $\varphi = (210. \dots 300)^\circ$ and for the Norma arm for $\varphi = (60. \dots 150)^\circ$. Finally, we show the radial profiles of the mass surface density and compare with other studies in Fig. 4. See Ref. [11] for a more detailed discussion of the results and comparison with other studies.

4. Conclusion

We have presented a new methods for kinematic deprojection of gas line surveys. Our methods builds on a Bayesian inference method [12] that allows taking into account the correlation structure of Galactic gas. We have applied our method to the Dame *et al.* survey compilation [6] to reconstruct the 3D density of H₂ in the Galaxy. Our reconstruction exhibits structure on a variety of scales, some associated with spiral arms. We have also provided the radial profile of the reconstructed gas densities. More details can be found in Ref. [11]. The reconstructed densities are available at <https://dx.doi.org/10.5281/zenodo.4405437>. An application to HI data is in preparation.

References

- [1] M. Heyer and T.M. Dame, *Molecular Clouds in the Milky Way*, *ARA&A* **53** (2015) 583.
- [2] M. Ackermann, M. Ajello, W.B. Atwood, L. Baldini, J. Ballet, G. Barbiellini et al., *Fermi-LAT Observations of the Diffuse γ -Ray Emission: Implications for Cosmic Rays and the Interstellar Medium*, *ApJ* **750** (2012) 3 [1202.4039].
- [3] G. Jóhannesson, T.A. Porter and I.V. Moskalenko, *The Three-dimensional Spatial Distribution of Interstellar Gas in the Milky Way: Implications for Cosmic Rays and High-energy Gamma-ray Emissions*, *ApJ* **856** (2018) 45 [1802.08646].

- [4] H. Nakanishi and Y. Sofue, *Three-Dimensional Distribution of the ISM in the Milky Way Galaxy: II. The Molecular Gas Disk*, *PASJ* **58** (2006) 847 [[astro-ph/0610769](#)].
- [5] M. Pohl, P. Englmaier and N. Bissantz, *Three-Dimensional Distribution of Molecular Gas in the Barred Milky Way*, *ApJ* **677** (2008) 283 [[0712.4264](#)].
- [6] T.M. Dame, D. Hartmann and P. Thaddeus, *The Milky Way in Molecular Clouds: A New Complete CO Survey*, *ApJ* **547** (2001) 792 [[astro-ph/0009217](#)].
- [7] T.V. Wenger, D.S. Balsaer, L.D. Anderson and T.M. Bania, *Kinematic Distances: A Monte Carlo Method*, *ApJ* **856** (2018) 52 [[1802.04203](#)].
- [8] M.J. Reid, K.M. Menten, A. Brunthaler, X.W. Zheng, T.M. Dame, Y. Xu et al., *Trigonometric Parallaxes of High-mass Star-forming Regions: Our View of the Milky Way*, *ApJ* **885** (2019) 131 [[1910.03357](#)].
- [9] N. Bissantz, P. Englmaier and O. Gerhard, *Gas dynamics in the Milky Way: second pattern speed and large-scale morphology*, *MNRAS* **340** (2003) 949 [[astro-ph/0212516](#)].
- [10] M.C. Sormani, J. Binney and J. Magorrian, *Gas flow in barred potentials*, *MNRAS* **449** (2015) 2421 [[1502.02740](#)].
- [11] P. Mertsch and A. Vittino, *Bayesian inference of three-dimensional gas maps: I. Galactic CO, in print*, arXiv:2012.15770 [[2012.15770](#)].
- [12] J. Knollmüller and T.A. Enßlin, *Metric Gaussian Variational Inference*, arXiv e-prints (2019) arXiv:1901.11033 [[1901.11033](#)].
- [13] D.A. Grabelsky, R.S. Cohen, L. Bronfman, P. Thaddeus and J. May, *Molecular Clouds in the Carina Arm: Large-Scale Properties of Molecular Gas and Comparison with H I*, *ApJ* **315** (1987) 122.
- [14] L. Bronfman, R.S. Cohen, H. Alvarez, J. May and P. Thaddeus, *A CO Survey of the Southern Milky Way: The Mean Radial Distribution of Molecular Clouds within the Solar Circle*, *ApJ* **324** (1988) 248.
- [15] S.W. Digel, *Molecular Clouds in the Distant Outer Galaxy.*, Ph.D. thesis, Harvard University, Cambridge, MA., Jan., 1991.