Astronomy & Astrophysics

Starlight-polarization-based tomography of the magnetized ISM: PASIPHAE's line-of-sight inversion method

V. Pelgrims[1,2] , G. V. Panopoulou[3] , K. Tassis[1,2] , V. Pavlidou[1,2] , A. Basyrov[4] , D. Blinov[1,2] , E. Gjerløw[4],
S. Kiehlmann[1,2] , N. Mandarakas[1,2] , A. Papadaki[1,2,5] , R. Skalidis[1,2] , A. Tsouros[1,2] , R. M. Anche[6,7] ,
H. K. Eriksen[4], T. Ghosh[8] , J. A. Kypriotakis[1,2], S. Maharana[1,2,7] , E. Ntormousi[1,2,9] , T. J. Pearson[10] ,
S. B. Potter[11,12], A. N. Ramaprakash[1,7,10] , A. C. S. Readhead[10] , and I. K. Wehus[4]

[1] Institute of Astrophysics, Foundation for Research and Technology-Hellas, N. Plastira 100, Vassilika Vouton, 71110 Heraklion, Greece
[2] Department of Physics, and Institute for Theoretical and Computational Physics, University of Crete, Voutes University campus, 70013 Heraklion, Greece
e-mail: pelgrims@physics.uoc.gr
[3] Hubble Fellow, California Institute of Technology, MC350-17, 1200 East California Boulevard, Pasadena, CA 91125, USA
[4] Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, 0315 Oslo, Norway
[5] Institute of Computer Science, Foundation for Research and Technology-Hellas, 71110 Heraklion, Greece
[6] Department of Astronomy/Steward Observatory, Tucson, AZ 85721-0065, USA
[7] Inter-University Centre for Astronomy and Astrophysics, Post bag 4, Ganeshkhind, Pune 411007, India
[8] School of Physical Sciences, National Institute of Science Education and Research, HBNI, Jatni 752050, Odisha, India
[9] Scuola Normale Superiore di Pisa, piazza dei Cavalieri 7, 56126 Pisa, Italy
[10] Cahill Center for Astronomy and Astrophysics, California Institute of Technology, 1216 E California Blvd, Pasadena, CA 91125, USA
[11] Department of Physics, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa
[12] South African Astronomical Observatory, PO Box 9, Observatory, 7935 Cape Town, South Africa

ABSTRACT

We present the first Bayesian method for tomographic decomposition of the plane-of-sky orientation of the magnetic field with the use of stellar polarimetry and distance. This standalone tomographic inversion method presents an important step forward in reconstructing the magnetized interstellar medium (ISM) in three dimensions within dusty regions. We develop a model in which the polarization signal from the magnetized and dusty ISM is described by thin layers at various distances, a working assumption which should be satisfied in small-angular circular apertures. Our modeling makes it possible to infer the mean polarization (amplitude and orientation) induced by individual dusty clouds and to account for the turbulence-induced scatter in a generic way. We present a likelihood function that explicitly accounts for uncertainties in polarization and parallax. We develop a framework for reconstructing the magnetized ISM through the maximization of the log-likelihood using a nested sampling method. We test our Bayesian inversion method on mock data, representative of the high Galactic latitude sky, taking into account realistic uncertainties from *Gaia* and as expected for the optical polarization survey PASIPHAE according to the currently planned observing strategy. We demonstrate that our method is effective at recovering the cloud properties as soon as the polarization induced by a cloud to its background stars is higher than ~0.1% for the adopted survey exposure time and level of systematic uncertainty. The larger the induced polarization is, the better the method's performance, and the lower the number of required stars. Our method makes it possible to recover not only the mean polarization properties but also to characterize the intrinsic scatter, thus creating new ways to characterize ISM turbulence and the magnetic field strength. Finally, we apply our method to an existing data set of starlight polarization with known line-of-sight decomposition, demonstrating agreement with previous results and an improved quantification of uncertainties in cloud properties.

Key words. dust, extinction – ISM: magnetic fields – ISM: structure – polarization – methods: statistical

1. Introduction

Studies of the interstellar medium (ISM) have relied on two-dimensional (2D) projections on the sky until recently. With the advent of sophisticated techniques and state-of-the-art facilities, astronomy has entered a new realm in which the third dimension can finally be accessed with accuracy, enabling the mapping of the ISM in three dimensions (3D). Astronomers – and the public – will soon be able to experience the Universe in 3D flying through real-world data sets loaded in virtual environments.

Gaia data on stellar distances in particular (e.g., Gaia Collaboration 2016; Gaia Collaboration 2021; Bailer-Jones et al. 2021) have allowed for the precise localization in 3D space of more than one billion stars in our Galaxy through the accurate determination of stellar parallaxes. Coupling measurements of stellar parallaxes to reddening, Bayesian inversion methods have already been successful at reconstructing 3D maps of the dust density distribution (e.g., Green et al. 2019; Lallement et al. 2019; Leike & Enßlin 2019; Leike et al. 2020), leading to stunning 3D images mapping the dust content of the ISM, in the

Solar neighborhood from within the first tens of parsec and up to much larger distances within the Galactic disk, already covering a substantial fraction of the Galaxy ($6000 \times 6000 \times 800\,\mathrm{pc}^3$ for Lallement et al. 2019, 2022). Such 3D mappings of the ISM content are of great interest for several areas in astrophysics. They shed new light on the dynamics shaping the Galaxy, breaking degeneracies caused by 2D mapping on the 3D shapes of ISM clouds and cloud complexes, their formation mechanisms, and their history (e.g., Ivanova et al. 2021; Bialy et al. 2021; Rezaei Kh. & Kainulainen 2022; Zucker et al. 2022; Lallement et al. 2022; Konstantinou et al. 2022). Ultimately, 3D images of the dust content of the Galaxy could also help in the characterization of Galactic foregrounds for observational cosmology and extra-galactic astrophysics (e.g., Martínez-Solaeche et al. 2018).

Impressive as they may be, 3D reconstructions of the ISM dust distribution are leaving out an important component of the Galaxy: magnetic fields, which are ubiquitous in the ISM. Magnetic fields are relevant in a variety of processes, from regulating star formation (e.g., Mouschovias et al. 2006; Hennebelle & Inutsuka 2019; Li 2021) to shaping large-scale structures in the disk and the halo of the Galaxy (e.g., Beck 2015). Magnetic fields in the Galaxy also affect our ability to study the Universe's structure and history, from its first moments to its later ages. Aspherical dust grains line up their shortest axis with the ambient magnetic field (e.g., Andersson et al. 2015). As a result, the thermal radiation emitted by those grains is polarized. This emission constitutes the major limitation in cosmologists' search for primordial B modes, the clear proof of primordial gravitational waves from inflation, and cosmic birefringence in the polarization of the cosmic microwave background (CMB; e.g., BICEP2/Keck Collaboration & Planck Collaboration 2015; Planck Collaboration XXX 2016; Planck Collaboration XI 2020; Diego-Palazuelos et al. 2022). This emission also represents a foreground in polarization studies of individual extra-galactic objects (e.g., Pelgrims 2019).

Significant effort has been invested in the last two decades to characterize the dust-polarized emission in order to disentangle it from the cosmological signal. However, this task has been proven to be very convoluted. Variations in dust spectral emission distribution, either in the plane of the sky (POS) or along the line of sight (LOS; Tassis & Pavlidou 2015; Planck Collaboration Int. L 2017; Pelgrims et al. 2021; Ritacco et al. 2023), and unexpected signatures of the dust signal in polarization power spectra (e.g., Planck Collaboration Int. XXXVIII 2016) – all rooted in the tight connection between dust clouds and the magnetic field – add many layers of complexity. Various sophisticated techniques are being developed to address these problems. The most direct way of attacking them and of providing confident and accurate solutions requires 3D mapping of the Galactic magnetic field in dusty regions (e.g., Hervías-Caimapo & Huffenberger 2022; Pelgrims et al. 2022; Konstantinou et al. 2022; Huang 2022).

Accessing the LOS structure of the magnetic field from dust emission alone is not feasible. Three-dimensional maps of the dust distribution can help identify LOSs with several clouds and place constraints on their respective significance; however, those maps alone provide no information about the magnetic fields permeating those clouds. While they can be combined with maps of dust-polarized emission in a coherent analysis to model the Galactic magnetic field (GMF) on large scales (Pelgrims et al. 2020), they cannot provide significant information on cloud scales, with perhaps some exceptions (e.g., Rezaei Kh. et al. 2020).

Fortunately, there are other probes that make it possible to infer the structure of the magnetized ISM in 3D. Among those, the linear polarization of stars, measured from the near-infrared (NIR) to the near-ultraviolet, is of particular interest, and can be used to study and model the dusty and magnetized ISM, from the smallest to the largest scales. While starlight usually starts out unpolarized, the same aspherical dust grains that are responsible for the polarized thermal emission induce a polarization to it when partially absorbing it, due to dichroic extinction (e.g., Andersson et al. 2015). Starlight polarization has been related to the magnetic field and the ISM in the Galaxy since its early observation (e.g., Hiltner 1949, 1951; Davis & Greenstein 1951; Heiles 2000). In comparison to other probes of the magnetized ISM, starlight polarization has the significant advantage that it can provide direct 3D information as soon as stellar distances are known. The potential of such 3D magnetic tomography to recover information on the LOS structure of the magnetic field has been demonstrated recently by Panopoulou et al. (2019b) using data collected from the RoboPol polarimeter (Ramaprakash et al. 2019), while correlation analysis of dust-polarized emission at sub-millimeter wavelengths and starlight polarization data has proven useful to locate the distance to the dominant polarized dust emission component seen at high Galactic latitude (Skalidis & Pelgrims 2019).

In recent years, several regions of the sky have been mapped with a high density of stellar polarization measurements (> 100 stars per sq.degree), including a significant portion of the inner Galaxy (in the NIR Clemens et al. 2020), as well as more diffuse regions of the ISM (e.g., Panopoulou et al. 2019b; Skalidis et al. 2022). These data sets have paved the way to 3D mapping magnetic fields in the general ISM of the Galaxy (e.g., Pavel et al. 2012), far from the dense regions of star formation that had been traditionally studied with large stellar samples (e.g., Pereyra & Magalhães 2004; Sugitani et al. 2011; Marchwinski et al. 2012; Santos et al. 2014; Franco & Alves 2015; Kwon et al. 2015; Eswaraiah et al. 2017). In the near future, the PASIPHAE survey (Tassis et al. 2018) and the SOUTH POL survey (Magalhães et al. 2005) will enable a leap forward by generating stellar polarization data for millions of stars, covering a large fraction of the sky. In conjunction with measurements of stellar distances obtained by *Gaia*, those data sets will pave the way for the characterization of the dusty magnetized ISM in 3D. Since stellar polarization traces the very same medium (magnetized dust) that produces the dominant CMB polarization foreground, starlight polarization data may offer a unique independent means to model out the polarization signal of the Galaxy, allowing the study of the very first moments of the Universe.

The observed polarization of each single star is the integrated effect of dichroic absorption from all interstellar clouds lying between us and the star. For this reason, in order to derive the complex 3D structure of the magnetized ISM from starlight polarization data and stellar distances, we need to develop methods that invert this LOS integration. So far, no standard method has been established in the literature to accomplish such a task in an automated, Bayesian way. On the one hand, different ad hoc methods have been considered (e.g., Andersson & Potter 2006; Panopoulou et al. 2019b; Doi et al. 2021), but they are not easily scalable to large data sets since they are not well adapted for automation and they do not allow for a straightforward, robust estimation of the credible interval of the reconstruction. On the other hand, methods developed for extinction data cannot be used unaltered on starlight polarization data. The main reason is that polarization is a pseudo-vector quantity. This implies that

it cannot be described by a single scalar quantity – two are needed: either the degree of polarization and polarization position angle, or its linear Stokes parameters. Additionally, even if contributions from individual clouds are additive (as is the case for linear Stokes parameters in the case of low amounts of extinction), polarization increments can be either positive, negative, or null. Because of these fundamental differences, dedicated, specialized methods need to be developed for the problem of starlight polarization tomography.

In this paper, we present such a specialized Bayesian method, developed for the PASIPHAE survey, implemented in Python, and now made publicly available for use by the community[1]. The inversion method developed here works on a per line-of-sight basis. We defer to future work for information on the extension of the method, which must take the correlation of the solutions in the plane of the sky into account.

In Sect. 2, we present our model for the distance dependence of starlight polarization along sightlines of the diffuse ISM, and we explain how we built our data equation and derived the corresponding likelihood. In Sect. 3, we provide details on the implementation of our Bayesian method and validate its performance by applying it to two simple examples of mock data. In Sect. 4, we present extensive testing of the performance of the method in the low signal-to-noise (S/N) regime and identify the method's limitations. We apply the method to currently available data in Sect. 5 and compare the results from our method to the literature. We finally summarize and discuss our work in Sect. 6. This paper contains two appendices. In Appendix A, we explain the creation of the mock observations used for performance testing, which are based on actual star samples, realistic estimates of the uncertainties on stellar parallax and polarization, and that rely on a complete toy model for the 3D structure of the ISM along sightlines. In Appendix B, we explore our toy model of the magnetic field geometry to gain intuition on the effects of turbulence-induced fluctuations in the ISM on the polarization observables.

2. Model, data equation, and likelihood

In this section, we lay the foundation for a model that describes the distance-dependence of stellar polarization toward a sightline of the diffuse ISM. We construct a generic data equation and build the corresponding likelihood that relates model parameters and star data. We first discuss the case of a single cloud along the LOS and then proceed to the generalization to cases with multiple-clouds.

2.1. Model: Thin-layer magnetized clouds

We model the LOS polarization induced by an individual cloud to background stars as being dominated by a single thin polarizing dust layer at the cloud distance (d_C). As already described by many authors (e.g., Andersson et al. 2015; Hensley & Draine 2021; Draine & Hensley 2021), the polarization induced by a dust cloud to the light of background stars depends on the dust opacity at the frequency of observation, the polarizing efficiency (which relates dust reddening $E(B-V)$ to a maximum polarization fraction) and on the apparent 3D orientation of the magnetic field (**B**) that permeates the cloud. The latter is described through the inclination angle of the magnetic field lines with respect to the POS ($\gamma_{\mathbf{B}}$) as well as the position angle of the POS component of the magnetic field ($\psi_{\mathbf{B}}$). For a single star behind a cloud,

with starlight intensity I_V, the vector of its relative Stokes parameters in the visible $(q_V, u_V) = (Q_V/I_V, U_V/I_V)$ equals the cloud polarization vector (q_C, u_C) given by

$$\begin{pmatrix} q_C \\ u_C \end{pmatrix} = P_{max} \cos^2 \gamma_{\mathbf{B}} \begin{pmatrix} \cos[2\psi_{\mathbf{B}}] \\ \sin[2\psi_{\mathbf{B}}] \end{pmatrix}. \tag{1}$$

In this equation, in which we have neglected any possible source of noise, $P_{max} \approx 13\% \, E(B-V)$, (Panopoulou et al. 2019a; Planck Collaboration XII 2020) where the reddening $E(B-V)$ generally depends on the dust grain physical properties and on the column density. Using single-frequency starlight polarization only, we have access to the position angle of the POS component of the magnetic field (related to the electric vector position angle, EVPA, of the stellar polarization) and to the magnitude of the induced polarization, that is the degree of polarization: $p_C = (q_C^2 + u_C^2)^{1/2}$ (related to the degree of stellar polarization). The latter is affected by the dust extinction, the dust polarizing efficiency, the inclination of the magnetic field with respect to the POS, and by possible LOS depolarization caused by turbulence within the cloud.

If the distribution of dust and the magnetic field properties were spatially homogeneous within a cloud, a single stellar measurement would suffice to describe the polarization properties it induces. Considering an ensemble of stars to constrain the cloud polarization properties makes it possible to take advantage of the number statistics and to sample the distance axis to provide constraints on the cloud distance. The former is critical for the S/N regime that is expected at high Galactic latitudes (Skalidis et al. 2018). Furthermore, observations of interstellar dust reveal fluctuations in the dust distribution on a range of scales (e.g., Miville-Deschênes et al. 2016) and fluctuations in the density and the magnetic field are also expected as a result of magneto-hydrodynamic (MHD) turbulence (e.g., Goldreich & Sridhar 1995; Cho & Lazarian 2003; Heiles & Crutcher 2005). Therefore, in order to obtain a realistic description of the polarization properties of a cloud within a region, we consider an ensemble of stellar measurements from LOSs within a finite circular aperture (called "beam" in the remainder of the paper) toward the cloud. We describe the Stokes parameters induced by the cloud to the ensemble of stars with a well-defined mean and a measure of dispersion about that mean. We refer to this dispersion as intrinsic scatter, to distinguish it from other sources of dispersion in the measurements, such as noise.

The intrinsic scatter has effects on polarization observables. We explore these effects in Appendix B where we characterize the variations produced by the intrinsic scatter on p_C and on $\psi_{\mathbf{B}}$; or more generally in the (q_V, u_V) plane. In particular, we show that 3D variations of the magnetic field can generate biases and a nonzero cross term in the polarization plane. The biases are irrelevant in our case since we are interested in recovering the mean values. The cross term, however, needs to be accounted for given that it might reach a non-negligible fraction of the variance of the Stokes parameters $(q_V, u_V;$ Montier et al. 2015). Other sources of variance in the polarization properties, such as fluctuations of the dust extinction across the sky, may reduce the importance of the off-diagonal element compared to the diagonal elements. Nevertheless, we retain the off-diagonal terms in our analysis for completeness.

As a first approximation, we assume that the turbulence-induced variations generate a bivariate normal distribution about the mean in the (q_V, u_V) plane. As a result, and in absence of observational noise, our stochastic model for the vector of Stokes

parameters of a star i in the background of a cloud is

$$\mathbf{m}_i = \begin{pmatrix} q_C \\ u_C \end{pmatrix} + G_2(0, C_{int})_i, \tag{2}$$

where q_C and u_C now denote the mean polarization values induced by the cloud. We denote the mean Stokes vector of the cloud as $\bar{\mathbf{m}} = (q_C\, u_C)^\dagger$. $G_2(0, C_{int})_i$ is a random realization of a 2D bivariate normal distribution centered on $(0, 0)$ with the 2-by-2 covariance matrix, C_{int}. The latter encodes the variances and covariances induced by sources of intrinsic fluctuations (e.g., turbulence) on the Stokes parameters. According to this generic description, six parameters are necessary to describe the polarization data of stars toward such a cloud: the distance of the cloud (d_C), the mean Stokes parameters (q_V, u_V) and three numbers to characterize the intrinsic-scatter covariance matrix C_{int}.

2.2. Data equation

To write the data equation, we need to account for the fact that a star at distance d_i may either be in the foreground ($d_i < d_C$) or in the background ($d_i > d_C$) of the cloud. In the former case no polarization is induced by the cloud and in the latter case the star polarization will be given by the mean polarization of the cloud plus one random realization of the intrinsic scatter. This piecewise-constant behavior is implemented through the use of the Heaviside function ($H(x) = 1$ if $x > 0$, 0 otherwise).

We further add a noise term ($\mathbf{n}_i$) to our stochastic model for the Stokes parameters. We consider that the observational noise which results from photon noise, instrumental polarization, and on-sky instrumental calibration, is described by a bivariate normal distribution $G_2(0, C_{obs})$ with covariance matrix, C_{obs}, where the off-diagonal terms can be nonzero. Unlike the intrinsic scatter, the covariance matrix corresponding to the observational uncertainties is generally source dependent; it might depend on the source's brightness, for example. The variance and covariance in the (q_V, u_V) plane result from both the intrinsic scatter and the observational uncertainties. We consider them as independent sources of Gaussian scatter in the polarization plane. Therefore, for stars in the background of a cloud the covariance matrices (C_{int} and C_{obs}) are summed. For stars that are not background to a cloud, only the observational uncertainties are relevant. The total covariance matrix thus takes the form:

$$\Sigma_i = C_{obs,i} + C_{int}\, H(d_i - d_C). \tag{3}$$

As a result, the data equation for the case of a single cloud along the LOS is:

$$\begin{aligned} \mathbf{s}_i &= \mathbf{m}_i\, H(d_i - d_C) + \mathbf{n}_i \\ &= \begin{cases} \bar{\mathbf{m}} + G_2(0, C_{int} + C_{obs,i})_i & \text{if } d_i > d_C \\ \mathbf{0} + G_2(0, C_{obs,i})_i & \text{otherwise,} \end{cases} \end{aligned} \tag{4}$$

where $\mathbf{s}_i$ is the vector of the measured Stokes parameters, d_i is the distance of the star, and $\mathbf{n}_i$ is the source-dependent noise term. Denoting $\bar{\mathbf{m}}_i$ the mean polarization induced by any dust cloud between us and the star and using Eq. (3), we thus write the vector of measured Stokes parameters as a random draw of a bivariate normal distribution with mean $\bar{\mathbf{m}}_i$ (with value either $\mathbf{0}$ or $\bar{\mathbf{m}}$) and 2-by-2 covariance matrix Σ_i:

$$\mathbf{s}_i \leftarrow G_2(\bar{\mathbf{m}}_i, \Sigma_i) = \frac{1}{2\pi\, |\Sigma_i|^{1/2}}\, \exp\left(-\frac{1}{2}\, \eta_i^\dagger\, \Sigma^{-1}\, \eta_i\right) \tag{5}$$

in which $|\Sigma_i| = \det(\Sigma_i)$ is the determinant of Σ_i and where we have introduced $\eta_i = \mathbf{s}_i - \bar{\mathbf{m}}_i$ with $\bar{\mathbf{m}}_i = \mathbf{0}$ or $\bar{\mathbf{m}}$ depending on whether the star is in the foreground of background to the cloud. In Eqs. (3) to (5), both the modeled mean and the modeled covariance terms for the stellar polarization depend on whether the source is in the foreground or the background of the cloud.

2.3. Likelihood

The data equation (Eq. (4)) concerns only the Stokes parameters of the stars. As explicitly written, our model for the Stokes parameters relates to the distance of the stars which, in turn, is also a measured quantity that comes with an uncertainty. This source of uncertainty adds complexity to the problem that we wish to solve, as distance uncertainties might impact the model prediction for $\mathbf{m}_i$, in particular for those stars that are near a cloud. Star distances are nowadays known with great accuracy (e.g., Bailer-Jones et al. 2021) through their parallaxes (ϖ) which, to good approximation, have Gaussian uncertainties (σ_ϖ). For this reason we choose to work in terms of parallaxes rather than distances, the two being related through the inverse relation $\varpi = 1/d$ where parallaxes are measured in arcseconds and distances in parsec. To account for this extra source of complexity, we notice that the star distance entering the data equation above should be the true distance of the star, and therefore we use the true parallax of the star in the following. We modify the argument of the Heaviside function from $(d_i - d_C)$ to $(\varpi_C - \varpi_i^0)$, in which ϖ_i^0 denotes the true parallax of a star. We consider that the measured parallax (ϖ_i) is a random realization of a Gaussian distribution centered on the true parallax with uncertainty σ_{ϖ_i}:

$$\varpi_i \leftarrow G(\varpi_i^0, \sigma_{\varpi_i}) = \frac{1}{\sqrt{2\pi}\, \sigma_{\varpi_i}}\, \exp\left(-\frac{(\varpi_i - \varpi_i^0)^2}{2\, \sigma_{\varpi_i}^2}\right). \tag{6}$$

In this work, we assume that the measurements of the parallax and of the optical polarization of stars are independent and uncorrelated. Further, we assume that the Stokes parameters for star polarization are functions of the true parallax of the star through the generic data equation built in the previous section (Eq. (4)). With these notations, the likelihood of the observational data point for star i with measured parallax ϖ_i and Stokes' vector $\mathbf{s}_i$ takes the form:

$$\begin{aligned} &P\left(\varpi_i, \mathbf{s}_i \,|\, \mathbf{m}_i,\, C_{int},\, \varpi_i^0,\, \sigma_{\varpi_i},\, C_{obs,i}\right) \\ &= P\left(\varpi_i \,|\, \varpi_i^0,\, \sigma_{\varpi_i}\right) P\left(\mathbf{s}_i \,|\, \mathbf{m}_i,\, C_{int},\, C_{obs,i}\right) \\ &= \frac{1}{\sqrt{2\pi}\sigma_{\varpi_i}}\, e^{-\frac{(\varpi_i - \varpi_i^0)^2}{2\sigma_{\varpi_i}^2}}\, \frac{1}{2\pi\, |\Sigma_i|^{1/2}}\, e^{-\frac{1}{2}\eta_i^\dagger \Sigma_i^{-1} \eta_i} \end{aligned} \tag{7}$$

where, in the last line, we explicitly write the parallax likelihood as a 1D Gaussian with standard deviation equal to the observational uncertainty and the polarization likelihood as a 2D Gaussian with the total covariance matrix that accounts for both the observational uncertainties and the contribution to intrinsic scatter from the crossed cloud (Eq. (3)).

We are not interested in modeling the true parallax of the star (ϖ_i^0). Instead we wish to marginalize over it to define the likelihood of the cloud parameters given an observation for star i. This marginalization allows us to separate the likelihood into two parts, one corresponding to the background of the cloud and

the other to its foreground as follows:

$$
\begin{aligned}
&\mathcal{L}_i \left(\varpi_{\mathrm{C}}, \bar{\mathbf{m}}, C_{\mathrm{int}} \right) \\
&= \int_0^\infty \mathrm{d}\varpi_i^0 \, P\left(\varpi_i, \mathbf{s}_i \mid \varpi_i^0, \bar{\mathbf{m}}_i, C_{\mathrm{int}}, \sigma_{\varpi_i}, C_{\mathrm{obs},i} \right) \\
&= \underbrace{\int_0^{\varpi_{\mathrm{C}}} \mathrm{d}\varpi_i^0 \, P\left(\varpi_i, \mathbf{s}_i \mid \varpi_i^0, \bar{\mathbf{m}}, C_{\mathrm{int}}, \sigma_{\varpi_i}, C_{\mathrm{obs},i} \right)}_{\text{background}} \\
&\quad + \underbrace{\int_{\varpi_{\mathrm{C}}}^\infty \mathrm{d}\varpi_i^0 \, P\left(\varpi_i, \mathbf{s}_i \mid \varpi_i^0, \mathbf{0}, \sigma_{\varpi_i}, C_{\mathrm{obs},i} \right)}_{\text{foreground}} \\
&= P\left(\mathbf{s}_i \mid \bar{\mathbf{m}}, C_{\mathrm{int}}, C_{\mathrm{obs},i} \right) \int_0^{\varpi_{\mathrm{C}}} \mathrm{d}\varpi_i^0 \, P\left(\varpi_i \mid \varpi_i^0, \sigma_{\varpi_i} \right) \\
&\quad + P\left(\mathbf{s}_i \mid \mathbf{0}, C_{\mathrm{obs},i} \right) \int_{\varpi_{\mathrm{C}}}^\infty \mathrm{d}\varpi_i^0 \, P\left(\varpi_i \mid \varpi_i^0, \sigma_{\varpi_i} \right) \\
&= P\left(\mathbf{s}_i \mid \bar{\mathbf{m}}, C_{\mathrm{int}}, C_{\mathrm{obs},i} \right) \frac{1}{2} \left(1 + \mathrm{erf}\left(\frac{\varpi_{\mathrm{C}} - \varpi_i}{\sqrt{2}\sigma_{\varpi_i}} \right) \right) \\
&\quad + P\left(\mathbf{s}_i \mid \mathbf{0}, C_{\mathrm{obs},i} \right) \frac{1}{2} \left(1 - \mathrm{erf}\left(\frac{\varpi_{\mathrm{C}} - \varpi_i}{\sqrt{2}\sigma_{\varpi_i}} \right) \right).
\end{aligned}
\tag{8}
$$

For a given sample of stars with polarization measurements and with known parallaxes and uncertainties, and under the assumption that the data are independent, the likelihood of the cloud parameters for a given LOS is given by the product of the likelihoods of the data points:

$$
\mathcal{L}\left(\varpi_{\mathrm{C}}, \bar{\mathbf{m}}, C_{\mathrm{int}} \right) = \prod_{i=1}^{N_{\mathrm{star}}} \mathcal{L}_i \left(\varpi_{\mathrm{C}}, \bar{\mathbf{m}}, C_{\mathrm{int}} \right).
\tag{9}
$$

This is the total likelihood function that we need to maximize to constrain our model parameters given the data.

2.4. Multicloud case

The generalization of the single-cloud model to the case with multiple independent clouds along the LOS is straightforward (we take N_{c} to be the number of clouds along the LOS). We consider that the Stokes parameters of a star in the background of multiple clouds result from the addition of the contributions from individual clouds. This approximation, which is correct in the low polarization regime (e.g., Appendix B of Patat et al. 2010), is well motivated for translucent LOSs through the diffuse ISM given that dust clouds can be considered as weak polarizers. The validity of this approximation may need to be revised for denser regions of the ISM, such as the Galactic plane or other LOSs through very dense molecular clouds.

In this work, we assume the linearity of the polarization signal and defer to future work the addition of more complex cases. Hence, in the low-polarization regime, the mean polarization vectors ($\bar{\mathbf{m}}_i^{[k]}$) of clouds along the LOS are additive. The same is true for the covariance matrices from the intrinsic scatter. Here, we have introduced the superscript $[k]$ to label clouds from $[1]$ to $[N_{\mathrm{c}}]$, for the nearest and the farthest cloud, respectively. As for the case of a single cloud, the total covariance (Σ_i) on the Stokes parameters for a star i depends on the observational uncertainties and on the sum of all sources of intrinsic scatter intervening along the LOS, from the star to the observer. Thus, assuming

N_{c} independent clouds along the LOS, the data equation for a star i is:

$$
\begin{aligned}
\mathbf{s}_i &= \mathbf{m}_i + \mathbf{n}_i \\
&= \bar{\mathbf{m}}_i + \mathrm{G}_2\left(0, \Sigma_i \right),
\end{aligned}
\tag{10}
$$

where

$$
\bar{\mathbf{m}}_i = \sum_{j \leq k} \bar{\mathbf{m}}^{[j]} \quad \text{and} \quad \Sigma_i = C_{\mathrm{obs},i} + \sum_{j \leq k} C_{\mathrm{int}}^{[j]},
\tag{11}
$$

in which we implicitly assume that star i lies behind cloud $[k]$ and in front of cloud $[k+1]$, if $k < N_{\mathrm{c}}$ and behind all clouds if $k = N_{\mathrm{c}}$.

To determine the likelihood for the set of cloud parameters given star data we marginalize over the true parallax of the star. This step allows us to separate the likelihood into $N_{\mathrm{c}} + 1$ terms, building the likelihood of a mixture model where the different terms correspond to a different (increasing) number of foreground clouds:

$$
\begin{aligned}
&\mathcal{L}_i \left(\{ \varpi_{\mathrm{C}}^{[k]}, \bar{\mathbf{m}}^{[k]}, C_{\mathrm{int}}^{[k]} \} \right) \\
&= P_{0,i} \, P\left(\mathbf{s}_i = \mathbf{0} \mid C_{\mathrm{obs},i} \right) \\
&\quad + P_{1,i} \, P\left(\mathbf{s}_i = \bar{\mathbf{m}}^{[1]} \mid C_{\mathrm{obs},i} + C_{\mathrm{int}}^{[1]} \right) \\
&\quad + \ldots \\
&\quad + P_{N_{\mathrm{c}},i} \, P\left(\mathbf{s}_i = \sum_{k=1}^{N_{\mathrm{c}}} \bar{\mathbf{m}}^{[k]} \mid C_{\mathrm{obs},i} + \sum_{k=1}^{N_{\mathrm{c}}} C_{\mathrm{int}}^{[k]} \right).
\end{aligned}
\tag{12}
$$

The coefficients $P_{k,i}$ result from the integration of the probability density function of the star parallax in the inter-cloud ranges of parallax and take the form

$$
P_{k,i} = \frac{1}{2} \left(\mathrm{erf}\left(\frac{\varpi_{\mathrm{C}_k} - \varpi_i}{\sqrt{2}\sigma_{\varpi_i}} \right) - \mathrm{erf}\left(\frac{\varpi_{\mathrm{C}_{k+1}} - \varpi_i}{\sqrt{2}\sigma_{\varpi_i}} \right) \right).
\tag{13}
$$

Finally, from an ensemble of stars, the likelihood of the cloud parameters for a given LOS is given by the product of the likelihoods of the data points, as in Eq. (9).

3. The Bayesian inversion method: Implementation and validation

In this section we first describe the implementation of our maximum-likelihood method to decompose the polarization properties of clouds along the LOS using stellar polarization and distance. We validate our method using mock stellar data sets based on a toy model of discrete clouds along the LOS and explore the sensitivity of the log-likelihood with respect to the different model parameters. We then provide a solution on how, facing real observations, we can select the correct model, that is, to choose the correct number of clouds that exist along a sightline.

3.1. Implementation

To maximize the likelihood function and estimate the posterior distributions of our model parameters we rely on a numerical method. We choose to use the code dynesty (Speagle 2020) to sample the parameter space using the nested sampling method (Skilling 2004). This code has already proven to be powerful and reliable in solving astrophysical problems similar to ours (e.g.,

Zucker et al. 2019, 2022; Alves et al. 2020). The algorithm uses sampling points (named 'live points' in `dynesty`'s definition) to explore the parameter space in a dynamic nested sampling scheme and estimate the posterior distributions on model parameters. It has two main advantages compared to other sampling methods: first, it returns an estimate of the model evidence and second, it includes a well-defined stopping criterion, suitable for automation of the fitting process.

The code `dynesty` takes as input the function of the log-likelihood that has to be maximized and the definition of functions that implement our prior knowledge on the model parameters. We implemented both uniform and Gaussian priors for the cloud parallaxes and cloud mean polarization. The Gaussian priors are defined through their means and standard deviations while uniform priors are defined by their lower and upper limits. We strongly recommend using flat priors for the element of the covariance matrix encoding the effects from turbulence. This is because the diagonal elements of the matrix must be positive and can approach zero depending on the (unknown) orientation of the magnetic field in 3D (see Appendix B). In this case, the diagonal elements ($C_{\mathrm{int},qq}$ and $C_{\mathrm{int},uu}$) are first proposed independently in their respective ranges and then the off-diagonal elements are drawn such that the semi-positive-definiteness of the covariance matrix is guaranteed, that is, $C_{\mathrm{int},qu}$ is drawn from a uniform distribution in the range $(-\sqrt{C_{\mathrm{int},qq}\,C_{\mathrm{int},uu}},\ \sqrt{C_{\mathrm{int},qq}\,C_{\mathrm{int},uu}})$; excluding the limits. In the case of multiple clouds, the prior function makes internally the distinction between clouds, ranked by their distances. We have hard-coded a lower limit on the number of stars that can exist between two clouds. We fix this limit at five.

Given a set of data points $\{\varpi_i, \mathbf{s}_i = (q_{\mathrm{V}}, u_{\mathrm{V}})_i\}$ and corresponding uncertainties (σ_{ϖ_i}, $C_{\mathrm{obs},i}$), the log-likelihood function that `dynesty` has to maximize is given by

$$\log\left[\mathcal{L}\left(\{\varpi_{\mathrm{C}}^{[k]}, \bar{\mathbf{m}}^{[k]}, C_{\mathrm{int}}^{[k]}\}\right)\right] = \log\left[\prod_{i=1}^{N_{\mathrm{star}}} \mathcal{L}_i\left(\{\varpi_{\mathrm{C}}^{[k]}, \bar{\mathbf{m}}^{[k]}, C_{\mathrm{int}}^{[k]}\}\right)\right]$$
$$= \sum_{i=1}^{N_{\mathrm{star}}} \log\left[\mathcal{L}_i\left(\{\varpi_{\mathrm{C}}^{[k]}, \bar{\mathbf{m}}^{[k]}, C_{\mathrm{int}}^{[k]}\}\right)\right] \quad (14)$$

which requires the number of clouds as an entry and where the $\mathcal{L}_i$'s are given by Eq. (12). The number of clouds populating the LOS is a priori unknown. We discuss in Sect. 3.5 how we intend to determine it. We implemented the likelihood functions for up to five clouds along the LOS. Even though the generalization of the implementation to higher number of clouds is trivial we do not deem it necessary given that only few LOSs at intermediate and high Galactic latitudes are expected to show a large number of components with significant contribution to the polarization signal (Panopoulou & Lenz 2020).

3.2. Mock data for two example LOSs

To test our method we developed a simple but complete implementation of our layer model to generate mock data sets with realistic number of stars, stellar distance and brightness distributions, and uncertainties both on parallax and polarization. This implementation, which includes a self-consistent prescription for the intrinsic scatter, is presented in detail in Appendix A.3. Our toy model for generating stellar polarization has five free parameters per cloud: the cloud parallax ($\varpi_{\mathrm{C}} = 1/d_{\mathrm{C}}$), the maximum degree of polarization (P_{max}), the inclination ($\gamma_{\mathbf{B}_{\mathrm{reg}}}$) and position ($\psi_{\mathbf{B}_{\mathrm{reg}}}$) angles of $\mathbf{B}_{\mathrm{reg}}$ and, finally, the relative amplitude of

fluctuations in magnetic field orientation (A_{turb}). We draw the reader's attention to the fact that, apart from the cloud parallax, these parameters are not the same as the model parameters entering our data equation. Additionally, our toy-model is stochastic due to the presence of the intrinsic scatter. Therefore, as demonstrated in Appendix B, the mean values of the Stokes parameters of a cloud and the values characterizing the covariance induced by the intrinsic scatter must be read from the simulated data before observational noise in parallax and polarization are introduced. To do so, we segment the 'clean' data sets at the input cloud distance(s) and we estimate the mean and covariance of the polarization induced by the cloud to the polarization of stars behind the cloud (but in front of the more distant cloud, if any). We refer to these estimates as the "true" values in the remainder of this paper.

We show in Fig. 1 two examples of simulated data for a single-cloud case (left) and a two-cloud case (right) applied to a sample of stars typical to intermediate to high Galactic latitudes for a circular sky area with a diameter of 0.5° (see Appendix A.1). We show the relative Stokes parameters as a function of distance modulus ($\mu_i = 5\log(d_i) - 5$, where d_i is the star distance in parsec). The top row shows the simulated data before noise in parallax and polarization is added and the bottom row shows the same simulated data sets when noise is added. The parameter values used to generate both mock data from our toy model (see Appendix A.3) are reported in Table 1. The two-cloud LOS is chosen such that (i) the far-away cloud alone induces about half the polarization of the nearby cloud alone, (ii) that the presence of the far-away cloud is only clear in one of the two Stokes parameters due to the POS orientation of $\mathbf{B}$ permeating the second cloud, and (iii) that there are at least 20 stars in the background of each cloud. The simulated data sets include uncertainties on stellar parallaxes based on *Gaia* performance (see Appendix A), realistic uncertainties on individual Stokes parameters as expected for WALOP-N (the northern instrument that will be used for the PASIPHAE survey) with 5 min of exposure time (see Appendix A.4 and Fig. A.3), and include a prescription for the intrinsic scatter. For completeness, we report in Table 1 the true values of the mean Stokes parameters and of the elements of the covariance matrix corresponding to the intrinsic scatter corresponding to the two examples shown in Fig. 1. These values correspond to the parameters entering our data equation and, ultimately, should be retrieved from the application of an inversion method to the data.

For the remainder of this section, we use these two examples of mock starlight polarization data to explore the sensitivity of our log-likelihood with respect to the different sampled parameters and to validate our implementation.

3.3. 1D conditional posterior distributions

As a first test of implementation, we perform 1D likelihood scans through the parameter space. This exercise also allows us to study the sensitivity of the likelihood, and therefore of the observables, to each model parameter individually. We present in Fig. 2 the conditional log-likelihood curves corresponding to the scans using the one-cloud LOS mock data shown in Fig. 1 (bottom left). We show in Fig. 3 the conditional probability distribution function (PDF) corresponding to these scans. They are not estimates of the 1D marginalized posterior distributions obtained from a fit since all parameters are not varied during the scans but are fixed to their true values. An actual fit to this particular simulated data set is performed in Sect. 3.4. The validation of our method and its implementation for several realistic cases
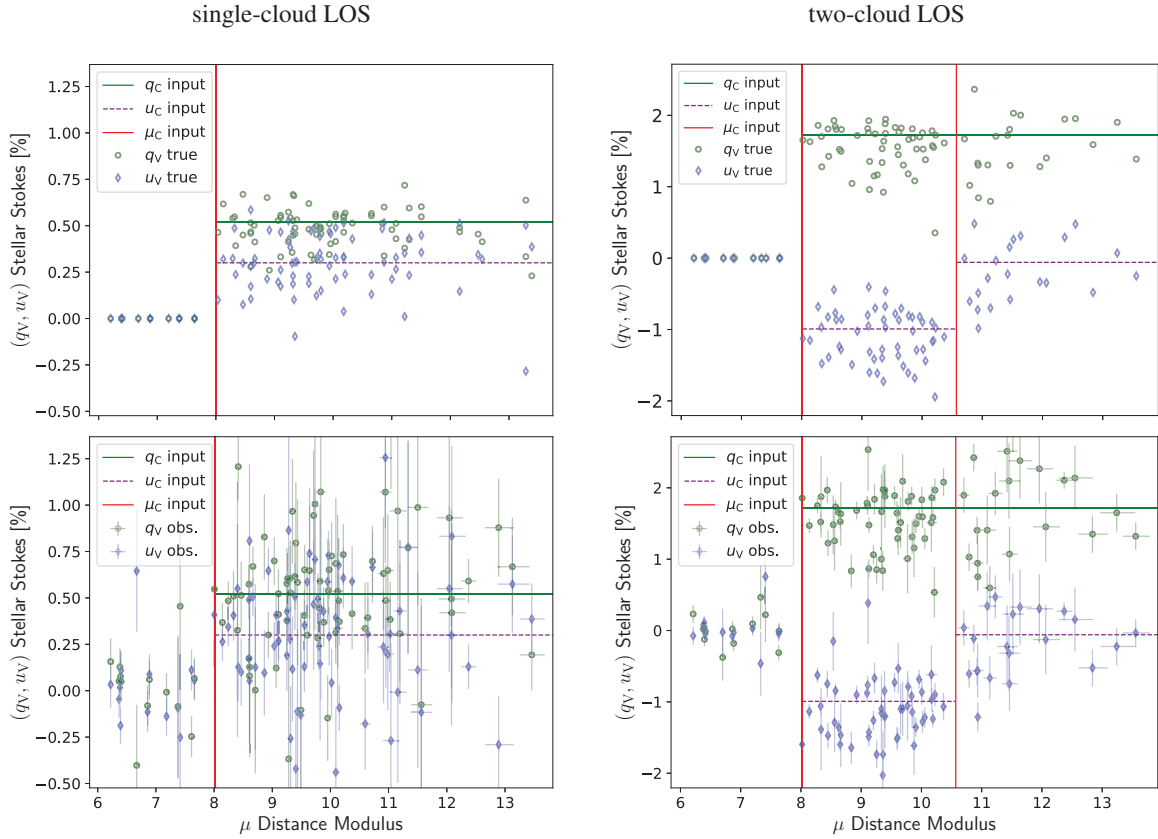
Fig. 1. Example of a single-cloud (left) and two-cloud (right) simulated data set. We show the stellar q_V (green circles) and u_V (blue diamonds) Stokes parameters as a function of distance modulus (μ_i). Top and bottom panels show the same data set. Top panels do not include observational noise (they are the "true" data points) while bottom panels do include uncertainties in both parallax and polarization (shown with errorbars). The vertical red lines indicate the input distance modulus of the clouds. The horizontal green and dashed-purple lines indicate the input (cumulative) mean Stokes parameters (q_C and u_C, respectively) before the inclusion of intrinsic scatter and observational noise, i.e., $\bar{\mathbf{m}}_i$ in Eq. (1).

Table 1. Setups for the simulated data set used in Sect. 3 and illustrated in Fig. 1.

Data set	Cloud parameters					"True" values				
cloud #	d_C (pc)	P_{max} (%)	$\gamma_{\mathbf{B}}$ (°)	$\psi_{\mathbf{B}}$ (°)	A_{turb} (−)	q_C (%)	u_C (%)	$C_{int,qq}$ (%)^2	$C_{int,uu}$ (%)^2	$C_{int,qu}$ (%)^2
Single-cloud										
cloud 1	400	0.8	30	15	0.2	0.49	0.30	0.01	0.02	−0.01
Two-cloud										
cloud 1	400	2.0	5	−15	0.2	1.56	−1.01	0.10	0.12	0.10
cloud 2	1300	1.0	15	45	0.2	0.00	0.91	0.17	0.18	0.16

Notes. The labels of the parameters follow the notations given in the text.

will be presented in the Sect. 4 where the performance and the limitations of the method will be assessed.

First, we note that the input parameter values always fall in the interval where $\log(\mathcal{L}) - \max(\log(\mathcal{L})) > -1$ meaning that the input values fall inside the approximated 68% credible interval. Second, we see that the shapes of the conditional log-likelihood curve, and of its corresponding conditional PDF, from the exploration of the cloud distance shown in Figs. 2 and 3 are somewhat surprising while the curves obtained for the other parameters look quite conventional. The very reasons for the unconventional shapes of the cloud-distance curves come from the unevenly distributed constraints (the stars) on the parallax (distance) space, their unequal uncertainties along that axis, the smearing in the

foreground and the background that the latter can generate, and, last but not least, the unequal constraining power of each star in the fit since polarization uncertainties are unequal. A star with large polarization uncertainties will constrain the fit less, and thus the position of the cloud along the LOS, as compared to a star with small polarization uncertainties. We illustrate part of this complexity in Fig. 4 in which we repeat the top left panel of Fig. 3 but where the true and observed parallaxes are indicated by vertical segments. It is clear that the likelihood of having a cloud with any distance between two distant constraints is constant and that the steepness of the variations depends on the parallax uncertainties. In general, standard statistics are not appropriate for characterizing the posterior distributions on cloud distances
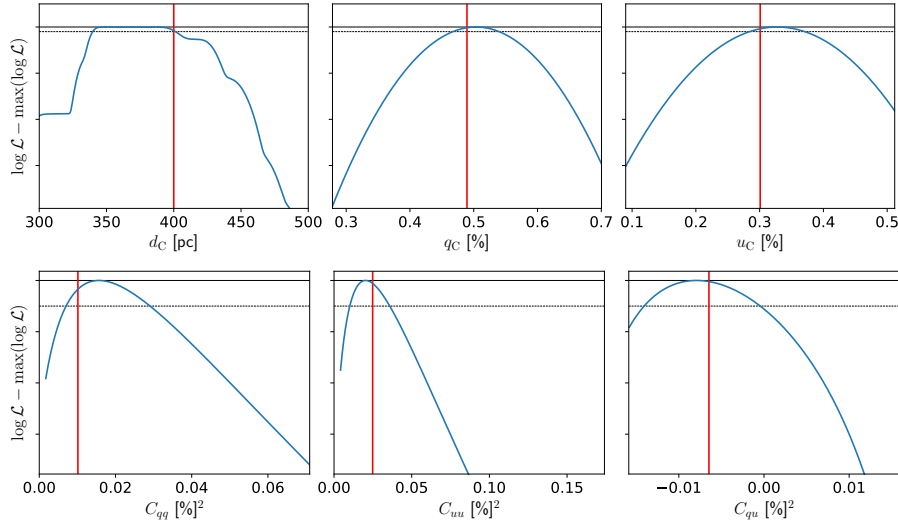
Fig. 2. Curves of $(\log(\mathcal{L}) - \max(\log(\mathcal{L})))$ corresponding to 1D likelihood scans through the parameter space for the mock data set with a single cloud along the LOS. For each scan only the explored parameter varies, while all other parameters are kept fixed to their true values. The log-likelihood $(\log(\mathcal{L}))$ is estimated at each point. The horizontal solid and dashed lines show the values of 0 and −1, respectively, providing an approximate estimate of the location of the 68 per cent credible interval. In the top (bottom) row the vertical axis ranges from −40 to 5 (−7.5 to 0.3). The red vertical line on each panel indicates the so-called true value reported in Table 1.
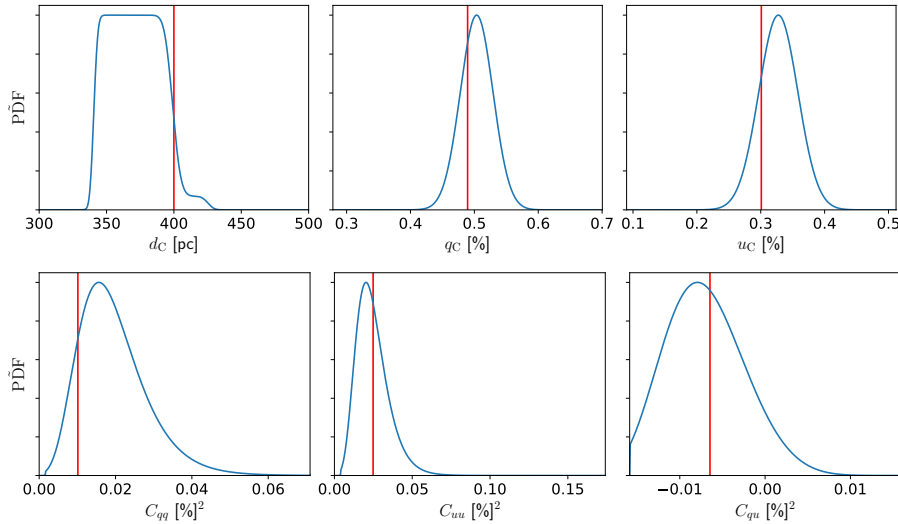
Fig. 3. Conditional probability distributions (P̃DF) corresponding to the $(\log(\mathcal{L}) - \max(\log(\mathcal{L})))$ curves in Fig. 2.

and dedicated metrics have to be considered to quantify success and goodness of fit. We address this point in Sect. 4.2.

3.4. Sanity checks

We check that our implementation of the model and of the maximization of the log-likelihood (Eq. (14)) through the nested sampling method is effective, by first applying our inversion method to the single-cloud LOS data used above and using a model with a single layer. We use 1000 live points, start with loose uniform priors on all parameters (the same as used in Sect. 4.3 and reported in Table 2) and sample the parameter space until an uncertainty of about 0.1 is achieved on the log of the model evidence $\mathcal{Z}$ (see Eq. (15)). Then the sampling is stopped and the samples are post-processed to generate 1D and 2D marginalized posterior distributions of the model parameters. The resulting histograms are shown on a corner plot format (Foreman-Mackey 2016) in Fig. 5. In this example, the

obtained maximum likelihood value is $\log \hat{\mathcal{L}} = 784.55$ and the evidence is $\hat{\mathcal{Z}} = 765.54 \pm 0.18$, all the input polarization-related parameters are found within the 68% credible interval of the estimated posterior distributions and 97% of the posterior on cloud parallax is contained between the true parallaxes of the two stars that directly bracket the input cloud parallax value. The method, therefore, demonstrates high accuracy in this example. We emphasize that not only do we recover the cloud distance and the mean values of the Stokes parameters but also an accurate estimate of the polarization covariance from the intrinsic scatter.

Similar conclusions are reached from the application of our fitting method to the two-cloud LOS example presented above using a model with two layers. In this case the priors set for each cloud are the same as used for the single-cloud LOS above and we use the same setup to analyze the data. We show in Fig. 6 the posterior distributions reconstructed by our fit for the parallaxes and Stokes parameters of the two clouds. The Stokes parameters are found to be well within the 68%
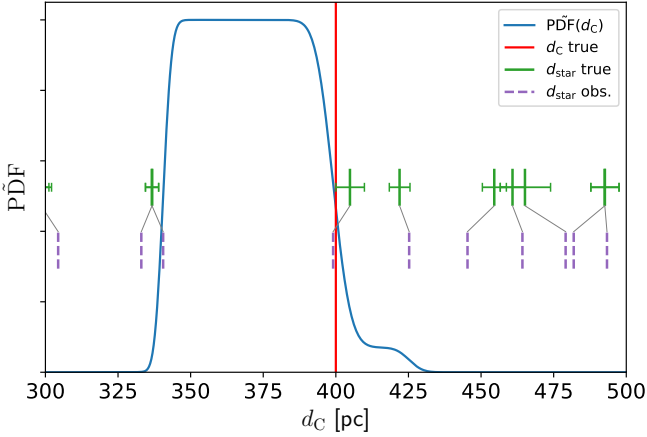
Fig. 4. Conditional probability distribution (P̃DF) corresponding to the 1D likelihood scan through ϖ_C for the one-cloud mock data set using the one-layer model. The vertical (continuous) green and (dashed) purple segments indicate the true and observed distances of the stars, respectively. The gray oblique lines link the two. The green horizontal errorbars indicate the 68% confidence level on star distances obtained from σ_{ϖ_i}. The vertical red line indicates the input cloud parallax. Due to randomization on parallaxes, some stars with very similar (true) parallaxes (green) are dispersed.

Table 2. Model parameters and range limits of uniform priors.

Cases	Model parameters				
cloud #	ϖ_C (mas)	q_C (%)	u_C (%)	$C_{int,qq}$ (%)^2	$C_{int,uu}$ (%)^2
Sect. 4.3					
cloud 1	[0.286, 10]	[–2, 2]	[–2, 2]	[0, 1]	[0, 1]
Sect. 4.4, 4.5.1, and 4.5.2					
cloud 1	[1.667, 10]	[–2, 2]	[–2, 2]	[0, 1]	[0, 1]
cloud 2	[0.286, 3.334]	[–2, 2]	[–2, 2]	[0, 1]	[0, 1]

Notes. Labels of the model parameters follow the notations given in the text: ϖ_C is the cloud parallax, q_C is the cloud's mean Stokes parameter q, u_C is the cloud's mean Stokes parameter u, $C_{int,qq}$ and $C_{int,uu}$ are the diagonal elements of the intrinsic-scatter covariance matrix. In all cases, the flat prior on $C_{int,qu}$, the off-diagonal element of the same matrix, is defined by $(-\sqrt{C_{int,qq}\,C_{int,uu}}, \sqrt{C_{int,qq}\,C_{int,uu}})$ such that the semi-positive-definiteness of the covariance matrix is guaranteed (thus, excluding the limits). The parallax values of 0.286, 1.667, 3.334, and 10 correspond to distance values of 3500, 600, 300, and 100 pc, respectively.

confidence interval of the estimated posteriors and 98% and 83% of the posteriors on clouds' parallaxes are contained between the true parallaxes of the two stars that directly bracket the values of their respective input-cloud parallaxes. We do not show the full corner plot, including the constraints on the intrinsic scatter parts, only because visualization of 12 parameters lead to poor insights; the posterior distributions are similarly very good, all true values falling within the 68% confidence interval of the estimated posteriors. The very fact that the posterior on the parallax of the second cloud appears less tight than that of the first cloud makes sense given that the second cloud is farther away from the observer where stars have larger uncertainties on their parallax and are generally fainter, thus showing larger uncertainties on Stokes parameters, than closer stars. Additionally, the presence of foreground clouds (stars) add noise to the reconstruction of the background cloud.

Figure 7 presents the estimated posterior distributions obtained for cloud parallaxes. The shapes of the posterior distributions are well understood when parallaxes of surrounding stars are considered, in a way similar to Fig. 4, illustrating the inhomogeneous distribution of constraints along that dimension.

In Fig. 8, we reproduce the bottom panels of Fig. 1 where we also represent the model evidences obtained from our maximum-likelihood analysis of the data. Namely, for any distance in the range spanned by the data stars, the (q_V, u_V) are estimated by re-sampling the posterior distributions of the model parameters. The modeled (q_V, u_V) are computed taking into account the intrinsic scatter and the correlation that it induces between the Stokes parameters. Given that our model is stochastic, we generate ten random draws for each of the 1000 sets of the model parameters randomly extracted from the posteriors. This leads to a bivariate distribution of (q_V, u_V) for any distance modulus. The shaded areas span the ranges of [2.5, 97.5] (light) and [16, 84] (dark) percentiles of the distributions for each μ. The medians of the distributions are given by the continuous line. The figure shows that that the data points and the model agree very well.

Another way of examining the agreement between data and model is to look at the residuals of the data points (in the polarization plane) as compared to the model predictions. Given that our model is stochastic and that (i) the polarization must be regarded as a bivariate quantity and (ii) measurement uncertainties are heteroscedastic; we are interested in the distribution of the Mahalanobis distances of the individual measurements in the beam with respect to their respective modeled means and computed with their respective total covariance matrices. The distribution of the Mahalanobis distances squared is expected to follow a χ^2 distribution with two degrees of freedom. We checked (not shown) that this is the case. Our distributions of Mahalanobis distances have a median of about 1.4 and no outliers are found.

3.5. Model selection

When dealing with real data, we generally do not know how many dust clouds exist along the LOS. In our approach, the number of clouds is an implicit parameter of the model. We first choose the number of clouds with which to model the data and then perform the maximum-likelihood analysis of the data. The latter provides an estimate of the posterior distributions on model parameters, an estimate of the evidence and an estimate of the maximum log-likelihood value. In itself, this approach cannot assess the validity of the chosen model. However, if we fit different models (i.e., with different input number of clouds) we can compare the results and decide, based on statistical arguments, which model is preferred by the data. This provides us with a standalone method to infer the number of clouds along the LOS and to proceed to its Bayesian decomposition in terms of dust layers, using polarization data of stars with measured distances only.

We consider two criteria to decide on the model that should be preferred given the data. The first criterion is based on the evidence returned by the nested sampling method (Skilling 2004) and the second is the Akaike information criterion (AIC; Akaike 1974).

The evidence, directly estimated through the use of `dynesty`, results from the integration over the full parameter space (Ω_Θ) of the likelihood of the model parameters (Θ) given the data (multiplied by the prior on the model parameters $\pi(\Theta)$):

$$\mathcal{Z} \equiv \int_{\Omega_\Theta} d\Theta \, \mathcal{L}(\Theta) \, \pi(\Theta). \qquad (15)$$
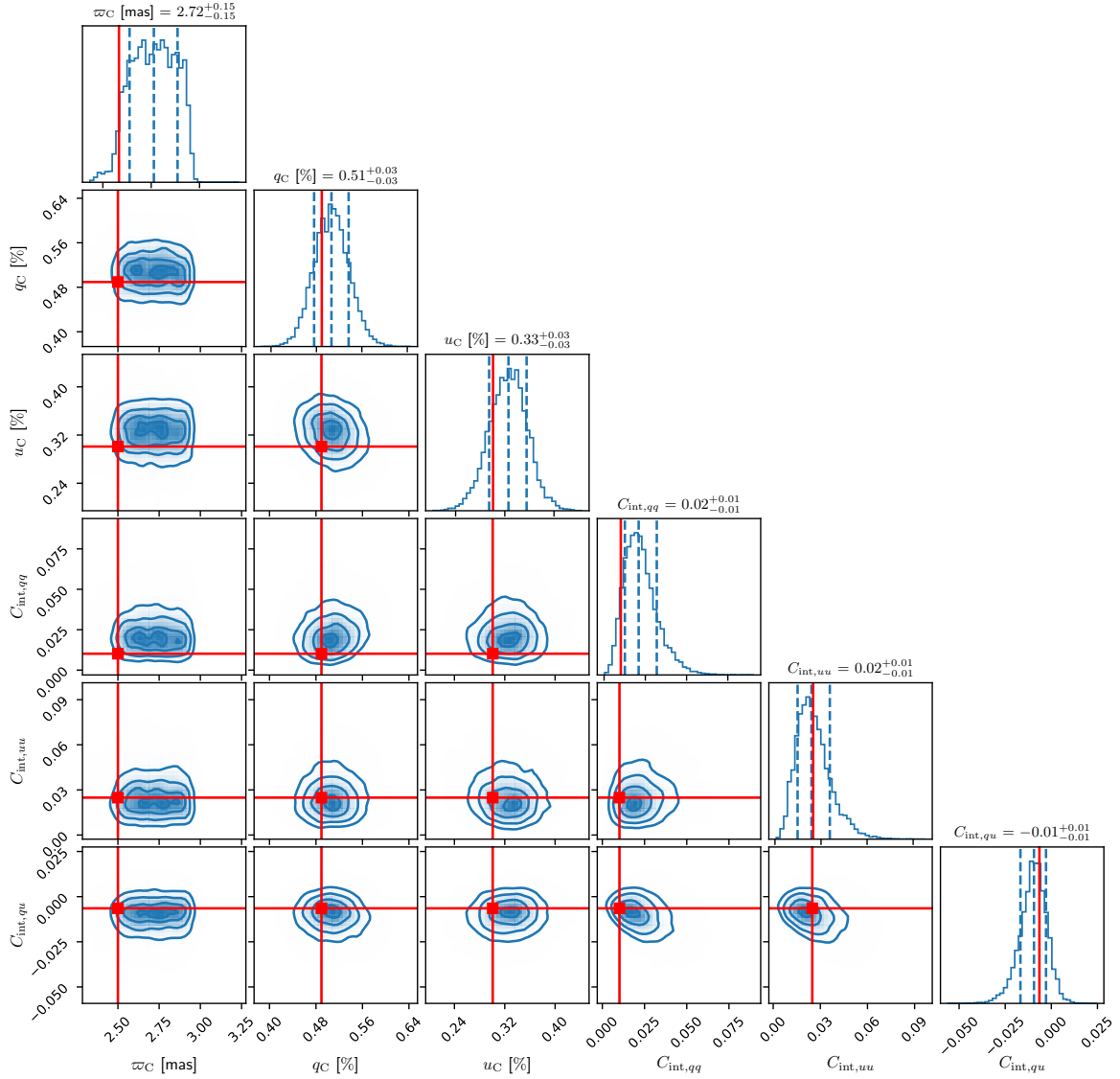
Fig. 5. Performance of the one-layer model in fitting the single-cloud mock dataset. 1D and 2D marginalized posterior distributions for the sampled model parameters obtained by log-likelihood maximization. The red lines indicate the true parameter values. Cloud parallax (ϖ_C) is given in milli-arcseconds (mas), mean Stokes parameters (q_C, u_C) in per cent and the elements of the covariance matrix encoding the effect of the turbulence-induced intrinsic scatter are given in per cent to the square (i.e., multiplied by 10,000). The dashed vertical lines indicate the 16, 50, and 84 percentiles of the 1D marginalized distributions and the values for the 68% confidence interval can be read from the title on each of the diagonal panels.

For an ensemble of models, the "best" model is the one that maximizes the evidence.

The AIC originates from information theory and is a measure of the amount of information that is lost by representing the data by a given model. It is based on the maximum likelihood value and includes a penalty for the number of model parameters. For a model j with M parameters, if the maximum likelihood is denoted by $\hat{\mathcal{L}}_j$, the AIC is given by:

$$\mathrm{AIC}_j = 2M - 2\log(\hat{\mathcal{L}}_j). \tag{16}$$

For an ensemble of models, the 'best' model is the one that minimizes the loss of information. The use of the AIC in model comparison is also attractive as it makes it possible to quantify the probability that a given model may minimize the information loss in comparison to the model that actually minimizes it in our data analysis experiment. Given a set of models $\{m\}$, the probability that model j minimizes the information loss is given by

Boisbunon et al. (2014):

$$P_{j|\{m\}} = \exp\left((\min_m\{\mathrm{AIC}_m\} - \mathrm{AIC}_j)/2\right). \tag{17}$$

In a conservative approach only models with small probability ($\lesssim 1\%$) should be disregarded. We use the above criteria in Sect. 5.

4. Performance

In the previous sections, we have introduced a Bayesian method to decompose the starlight polarization data in terms of independent clouds along the LOS by maximizing a dedicated log-likelihood function. We have demonstrated on two examples that the method is effective in performing the decomposition and in recovering the true values of the modeled data. In this section
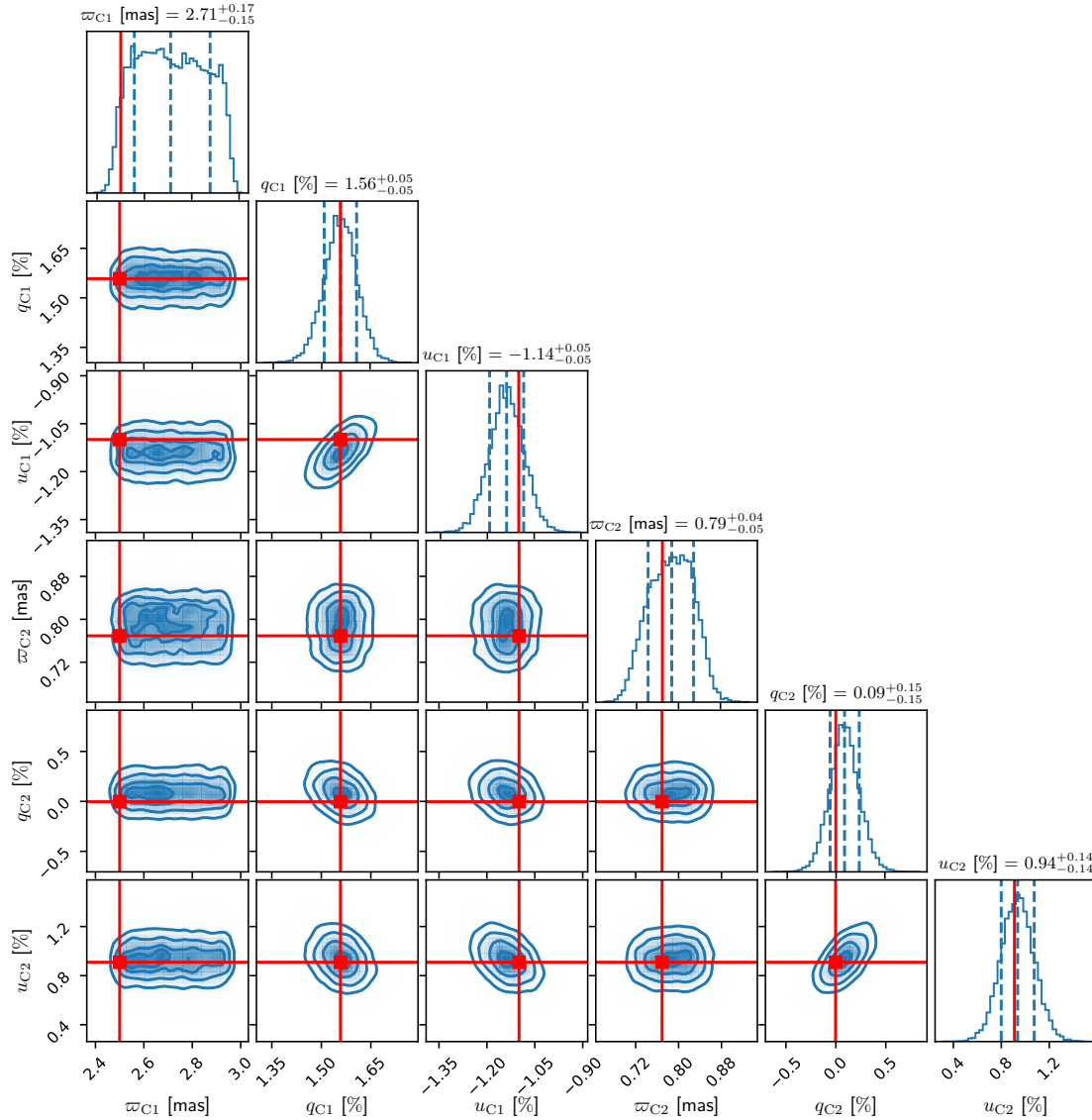
Fig. 6. Same as in Fig. 5 but showing a restricted sample of the model parameters of a two-layer-model fit to the two-cloud LOS shown in example in Fig. 1. We show the estimated 1D and 2D marginalized posterior distributions only for the parallax and Stokes parameters of the 2 modeled clouds.

we aim to investigate the performance of the method and identify the limits of its applicability for the case of PASIPHAE-like observations.

The performance of the method at recovering cloud parameters is expected to depend on: (i) the amplitude of the polarization signal that a given cloud induces to the light of background stars, (ii) the number of stars effectively sampling a given cloud (i.e., stars in the background of that cloud but in the foreground of any potential farther one), (iii) the noise level of the stellar polarization measurements, (iv) the precision on the star parallaxes (which is also dependent on the star distances), (v) the degree of intrinsic scatter, and (vi) on the number of clouds along the LOS.

To determine the performance of the method we need to rely on a metric that quantifies the goodness of the fit in both distance and polarization. Our metrics are introduced below, after we introduce the simulated starlight polarization data which rely on actual stellar magnitudes and parallaxes from *Gaia*. Then, we study the performance of our method by applying it to several single-cloud-LOS and two-cloud-LOS cases that

could be typical of LOSs at intermediate and high Galactic latitudes that will be targeted by PASIPHAE. Finally we explore the use of the criteria to select the most likely model to test their efficiency.

We expect, and have checked (not shown), that the method performs best when the polarization induced by the ISM is large compared to the uncertainties in the stellar polarization (but still within the weak polarization limit assumed for the diffuse ISM). We therefore wish to test our method in conditions of low S/N in polarization to determine its limitations. At intermediate and high Galactic latitude ($|b| > 30°$) the 80th percentile of the distribution of stellar polarization in Berdyugin et al. (2014) is at about 0.3%. This value is also in broad agreement with the distribution obtained from the extrapolation of the H I column density of low-velocity clouds in those sky areas (calculated based on the data by Panopoulou & Lenz 2020). Hence, to generate the mock starlight polarization samples on which to apply our inversion method, we explore the parameter space of our toy model (see below and Appendix A.3) so that p_C ranges from 0% to about 0.3%.
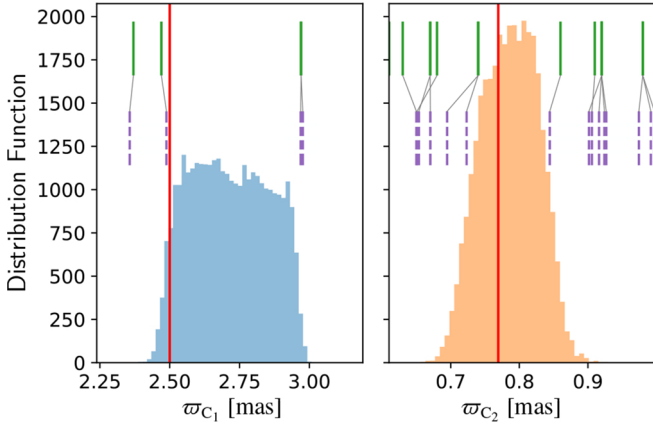
Fig. 7. Estimated posterior distributions for ϖ_{C1} (left) and ϖ_{C2} (right) in [mas] with marked observed and input parallaxes of surrounding stars with (dashed) purple and (continuous) green vertical segments, respectively, similar to Fig. 4. Vertical red lines indicate the input parallaxes of the clouds. As before, some stars with very similar (true) parallaxes (green) are dispersed after randomization.

4.1. Simulated data

To study the performance of our inversion method we rely on simulations of starlight polarization data. We construct realistic stellar samples from the Gaia Universe Model Snapshot[2] (GUMS) database. As fully described in Appendix A.1, we obtain stellar distance and photometry data at high Galactic latitudes within circular beams having two different diameters: 0.5° and 1°. We apply two selection criteria based on stellar brightness and S/N in parallax. We retain stars with SDSS r-band magnitude $r < 16$ mag, which is the expected limiting magnitude of the PASIPHAE survey (Tassis et al. 2018). To ensure high precision in stellar distance information, we keep stars with S/N in parallax larger than five ($\varpi/\sigma_\varpi \geq 5$). The main characteristics of the stellar samples are shown in Fig. A.1, after the application of these selection criteria. We show the distributions of star distance (modulus) and SDSS r-band magnitude for a 1° beam and a 0.5° beam.

To proceed further in the modeling of starlight polarization data for our performance tests, we select one sample per beam size. We choose the samples with the mean number of stars, sampling actual data at high Galactic latitudes. For each star in the two samples we have estimates of the actual parallax, parallax uncertainty, and magnitude in the SDSS-r band. Then, applying our toy model (Appendix A.3) to these star samples we can generate simulated starlight polarization data with realistic uncertainties for any desired setup of the magnetized ISM. The latter is determined by the number of dust clouds along the considered LOS and by five parameters per cloud (see Appendix A.3). Two examples of simulated starlight polarization data obtained for a one-cloud and a two-cloud ISM setup applied to the mean star sample within a beam of 0.5° in diameter are shown in Fig. 1 with, however, a larger induced polarization signal than for the performance tests carried out below.

4.2. Metric to assess the goodness of LOS decomposition

Our primary focus in the LOS decomposition of the starlight polarization data is to retrieve the distance of the ISM clouds along the LOS and to infer their mean polarization properties.

The quality of any reconstruction should therefore focus on these two aspects.

First, to quantify the accuracy with which we recover the cloud parallaxes (distances), we have to deal with the nonregular – often multi modal – shapes of the posterior distributions. In some cases, the cloud parallax at maximum-likelihood may not correspond to the mode of the posterior distribution. This becomes pathological once the cloud parallax at maximum log-likelihood belongs to a peak of the posterior distribution that has a small amplitude as compared to the dominant peak or when the bulk of the posterior distribution is squeezed onto one of the limits imposed by the prior. This may happen when the constraining power of the stars is not sufficient for the log-likelihood hyper-surface to show a strong global maximum at maximum-likelihood value; cases generally corresponding to weak ISM polarization signal, or low star density around the dust cloud. In such an occurrence, the fit fails and should be disregarded.

We found that (i) the relative difference between the cloud parallax at maximum log-likelihood with the parallax of the closest star to the input cloud, and (ii) the fact that the cloud parallax at maximum log-likelihood appears in (one of) the main modes of its posterior distribution are two criteria that jointly allow us to assess the quality of our fit on ϖ_C in addition to the rejection of fits with posterior distribution squeezed on one of the prior limits. If the maximum-likelihood parallax value belongs to (one of) the main modes, then the fit is valid and the relative difference tells us about the accuracy of the recovered cloud parallax. To decide whether the maximum-likelihood parallax belongs to a significant peak, we analyze the posterior distribution using the peak-finder algorithm `find_peaks` of the Python library `Scipy` which identifies all local maxima through simple comparison of neighboring values. Applying it to the marginalized parallax posterior distribution we identify the peaks and their boundaries (i.e., we find local maxima of the marginalized PDF and the range between the two adjacent local minima of each of those). We compute the fraction of the PDF corresponding to each peak. If the fraction associated to the peak to which the maximum-likelihood belongs is higher than a given threshold, then we consider the peak as (one of) the dominant one(s) and the fit as valid. We refer to this, and the detection of squeezed posterior distribution on one of the prior limit, as the criterion on $\hat{\varpi}_C$. We use a threshold of 0.5 in what follows and make sure that our results do not depend sensitively on this choice.

Second, to quantify the accuracy with which we recover the mean Stokes parameters of the clouds, we rely on the computation of the Mahalanobis distance of the true values with respect to the bivariate posterior distributions of the cloud Stokes parameters. If $\mathbf{c}_\star = (q_C\, u_C)^\dagger$ is the vector of true mean polarization of the cloud, $\hat{\mathbf{c}} = (\hat{q}_C\, \hat{u}_C)^\dagger$ is the value at maximum-likelihood, and $\hat{\Sigma}$ is the covariance matrix computed from the re-sampling of the estimated posteriors, the Mahalanobis distance is given by:

$$d_{\text{Maha}}(\mathbf{c}_\star|\hat{\mathbf{c}}, \hat{\Sigma}) = \sqrt{(\mathbf{c}_\star - \hat{\mathbf{c}})^\dagger\, \hat{\Sigma}^{-1}\, (\mathbf{c}_\star - \hat{\mathbf{c}})}. \tag{18}$$

It is also useful to introduce the Mahalanobis distance between the bivariate posterior distribution on the cloud parallax and the point (0, 0) of the polarization plane. This distance, labeled $d_{\text{Maha}}(\mathbf{0}|\hat{\mathbf{c}}, \hat{\Sigma})$ obtained by substituting the true polarization of a cloud by zero, is a measure of the significance with which a cloud polarization is detected given the data and the sensitivity of the inversion method.

The distance $d_{\text{Maha}}(\mathbf{c}_\star|\hat{\mathbf{c}}, \hat{\Sigma})$ determines whether the "true" Stokes parameters are located, and centered, within the respective bivariate distribution. However, if the reconstruction method
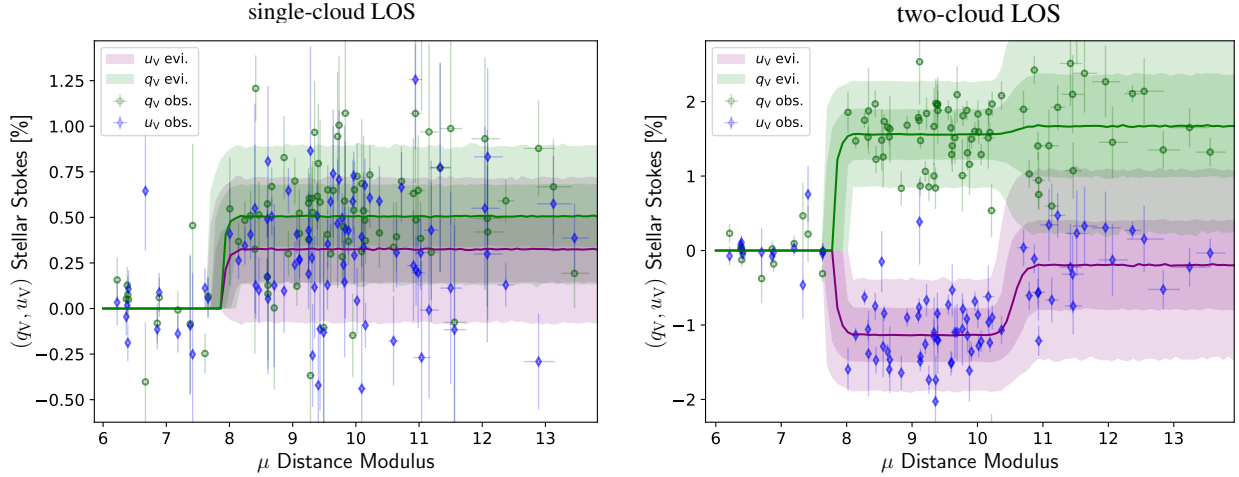
Fig. 8. Representation of the model evidences obtained for the single-cloud LOS (left) and two-cloud LOS (right). The data points are the same as in Fig. 1. The shaded areas illustrate the distributions of (q_V, u_V) obtained at every μ value through re-sampling of the posterior distributions obtained from the maximum-likelihood analysis of the data, as explained in the text.

fails at recovering the cloud distance for example, the posterior distributions estimated from the sampling method can become arbitrarily large. In such a scenario, it is useful to consider both the absolute Euclidean distance between the pair of Stokes parameters at maximum-likelihood value and the corresponding true values and the effective size of the estimated posteriors. We use the Euclidean distance defined as

$$L_2 = \sqrt{(q_C - \hat{q}_C)^2 + (u_C - \hat{u}_C)^2}, \qquad (19)$$

and we infer the sizes of the posterior distributions on q_C and u_C by considering the ratio (ξ) of the posterior sizes that are expected in ideal cases to the measured sizes from the estimated posteriors. If the cloud Stokes parameters were constrained only from N_{bg} stars all with uncertainties of 0.1% (minimal value in our simulated data), the expected size of the posterior is $(0.1\,[\%]/\sqrt{N_{bg}})$. We call this estimate ideal in the sense that, in addition to considering the smallest possible polarization uncertainties, it neglects possible correlation between q_C and u_C, the presence of intrinsic scatter, the presence of foreground stars - which add noise to the reconstruction-, and scrambling along LOS distance from parallax uncertainties. Within the low S/N regime explored in this section, we found that the parameter

$$\xi = 2 \frac{0.1\,[\%]}{\sqrt{N_{bg}}\ \sqrt{\hat{\sigma}_{q_C}\,\hat{\sigma}_{u_C}}}, \qquad (20)$$

where $\hat{\sigma}_{q_C}$ and $\hat{\sigma}_{u_C}$ denote the standard deviation (in %) of the respective estimated posteriors, generally takes values around unity for well-behaved fits and takes small values when the cloud polarization is not well-constrained.

In summary, a reliable reconstruction of the ISM structure along the LOS is expected if, simultaneously, we have a well behaved posterior on cloud parallax with small relative difference (between the cloud parallax at maximum-likelihood and the parallax of the star closest to the input), a low L_2 distance and ξ above a threshold which we choose to be 0.5.

4.3. One-layer cases

In this subsection we explore the behavior and the performance of our inversion method for several cases of a single cloud along

the LOS. We made sure that our inversion method works equally well for any position angle of the magnetic field and will not make any distinction in what follows even though we let it vary to generate our simulated data. We want to infer the power of our method at recovering the cloud distance and at recovering the Stokes parameters when the amplitude of the polarization signal varies, the distance of the cloud varies, and the amplitude of the intrinsic scatter varies. For a given sample of stars, the choice of cloud distance affects the number of stars in the foreground and the background that are available to constrain the model. The amplitude of the (mean) polarization signal depends on both P_{max} and $\gamma_{B_{reg}}$ and both should therefore be explored. The amplitude of the intrinsic scatter directly affects the amount of scatter in the (q_V, u_V) plane for background stars only.

Given a sample of stars, we choose the cloud distances such that 90%, 70%, 50%, 30%, and 10% of the stars are in the background of the cloud and hence, useful to constrain the polarization properties of the cloud. The lower the fraction of stars in the background (f_{bg}), the larger the distance of the cloud, and the larger the parallax uncertainties of stars in the distance neighborhood of the cloud. Thus, we expect larger relative differences in cloud parallax, and consequently larger L_2 and smaller ξ, at small f_{bg} than at large f_{bg}. In practice, however, this picture may be changed due to the imposed prior on cloud parallax and, in particular, due to the lower limit on which the cloud parallax posterior may be squeezed. Those cases, that anyway do not pass the criterion on $\hat{\varpi}_C$, may show small relative differences on cloud parallaxes simply because of the limit of the prior on ϖ_C is close to the input cloud parallax value. In such cases, though, the posterior on the mean polarization Stokes parameter should be broader than expected (i.e., showing low ξ values).

For a fixed level of intrinsic scatter with $A_{turb} = 0.2$ and the mean star sample with a 1° diameter beam, and for all values of f_{bg}, we first study the impact of p_C, the degree of polarization induced by the cloud to background stars, on the quality of the reconstruction of the ISM structure along the LOS. For this sample made of 345 stars the f_{bg} cuts correspond to cloud distances of about 270, 565, 790, 1050, and 1712 pc. To avoid mixing possible dependence on P_{max} and $\gamma_{B_{reg}}$, we impose the regular component of the magnetic field to be in the POS ($\gamma_{B_{reg}} = 0°$) and only vary P_{max} from 0.05% to 0.3% in steps of 0.05%. We further consider 10 realizations for each pair of (f_{bg}, P_{max})

varying the POS position angle of $\mathbf{B}_{\text{reg}}$. This generates a set of 300 simulated samples of starlight polarization data to which we apply our inversion method. Each sample is analyzed through our maximum-likelihood method using the one-layer model. We consider relatively loose flat priors on the six parameters with $\{q_{\text{C}}, u_{\text{C}}\} \in [-2\%, 2\%]$, $\{C_{\text{int},qq}, C_{\text{int},uu}\} \in [0, 10^{-4}]$, $|C_{\text{int},qu}| \leq 10^{-4}$, and ϖ_{C} in the range corresponding to distances between 100 and 3500 pc. A value of $\{C_{\text{int},qq} = 10^{-4}$ corresponds to an intrinsic scatter of 1% on Stokes q. For convenience, the definitions of these priors are repeated in Table 2. We consider 1000 live points and sample the parameter space until we reach a tolerance on the estimated log-evidence below 0.1. Generally in our cases, this corresponds to an uncertainty on the log-evidence on the order of one part per 10 000.

We subsequently analyze the posterior distributions, flag reconstructions with odd posterior on ϖ_{C} (i.e., those not passing the criterion on $\hat{\varpi}_{\text{C}}$), and evaluate the three quantities with which we intend to qualify the goodness of the reconstruction: the relative difference between the $\hat{\varpi}_{\text{C}}$ and the parallax of the star that is nearest to the input cloud, the L_2 distance between $(\hat{q}_{\text{C}}, \hat{u}_{\text{C}})$ and the true $(q_{\text{C}}, u_{\text{C}})$, and the parameter ξ which compares the sizes of the posteriors on $(q_{\text{C}}, u_{\text{C}})$ with the ideal ones (see above). For the reconstructions that pass the criterion on $\hat{\varpi}_{\text{C}}$, we checked that there is no bias in the recovered mean polarization properties of the cloud; $(\hat{q}_{\text{C}}, \hat{u}_{\text{C}})$ is always found close to the input $(q_{\text{C}}, u_{\text{C}})$ within uncertainties.

We present in Fig. 9 our results on parallax relative differences (top), L_2 (middle) and ξ (bottom) as a function of the true p_{C} for the five cloud distances corresponding to the five f_{bg} thresholds. The different colors indicate different f_{bg} values, the solid lines connect the medians obtained for each f_{bg} at the different P_{max} values (all fits). The scatters and offsets in $p_{\text{C}}^{\text{true}}$ about multiples of 0.05% observed in the figure originate from the presence of the intrinsic scatter. Filled (empty) symbols indicate fits that pass (do not pass) the criterion on $\hat{\varpi}_{\text{C}}$.

As a general trend, we see that the performance on both cloud distance and cloud polarization properties is generally good for the cases with $p_{\text{C}}^{\text{true}} \gtrsim 0.1\%$. Below this threshold a significant fraction of reconstructions lead to pathological posterior distributions on cloud parallax. In practice those reconstructions should be disregarded. Pathological reconstructions also extend to larger $p_{\text{C}}^{\text{true}}$ for cases corresponding to $f_{\text{bg}} = 10\%$. This is due to both a lower number of stars to constrain the model parameters and larger parallax uncertainties. Running the same performance test using starlight polarization samples in the 0.5° diameter beam, we find that it is the number of stars in the background of the cloud that is the dominant factor, as pathological cases are found at larger f_{bg} on the order of 30% in tested cases (not shown). It is difficult to constrain the distance of a (relatively) far away cloud beyond which only about 30 stars provide constraints.

When $p_{\text{C}}^{\text{true}} \gtrsim 0.1\%$, the relative differences in cloud parallax obtained for most of the reconstructions with well-behaved ϖ_{C} posteriors are found below the level of 5% regardless of the cloud distance. The Euclidean distances L_2 indicate a good accuracy at recovering the cloud mean polarization, with about 90% of the valid reconstructions leading to an L_2 below 0.05%. The larger the f_{bg}, the smaller L_2, thus the more accurate the reconstruction. This is an expected feature since for lower f_{bg}, the reconstructions can be noisier due to both the presence of (an increasing fraction of) foreground stars and a smaller amount of stars constraining the cloud polarization properties. This is also seen in the ξ panel. For the valid reconstructions, the posterior distributions get larger (ξ decreases) when f_{bg} decreases, thus lowering
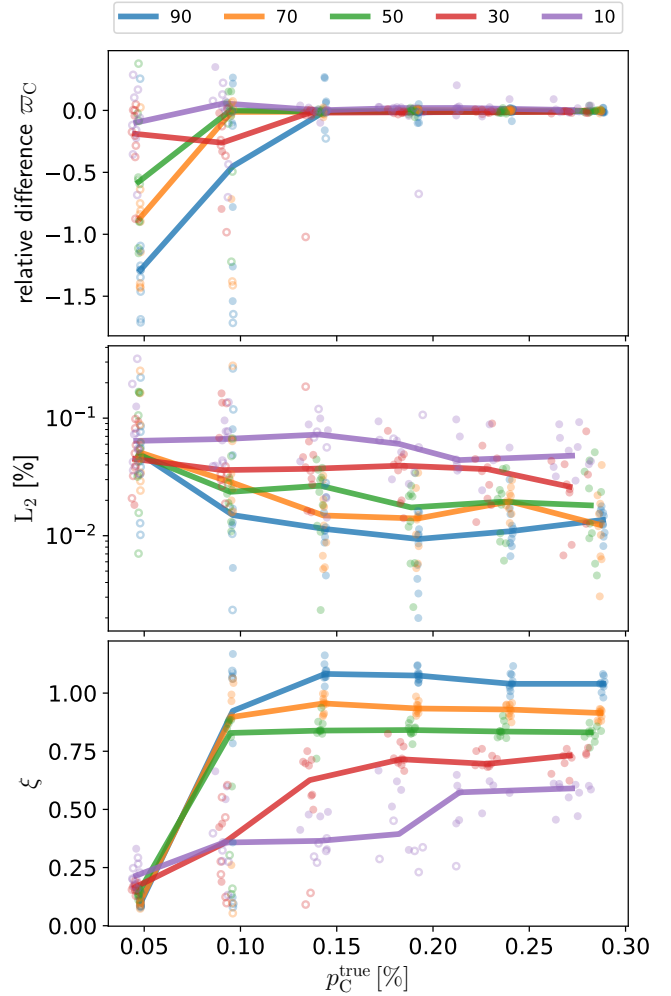
Fig. 9. Performance of the inversion method as a function of the true polarization signal for different cloud distances, probed by the values of f_{bg}, shown in the legend and given in per cent. The method is applied to simulated data from the 1° aperture star sample (345 stars). Top: relative difference between the cloud parallax at maximum log-likelihood with the parallax of the closest star to the input cloud ($2\,(\hat{\varpi}_{\text{C}} - \varpi_{\text{closest}})/(\hat{\varpi}_{\text{C}} + \varpi_{\text{closest}})$). Middle: L_2, distance between the true and estimated mean polarization vector. Bottom: ξ, the ratio between ideal and estimated posterior size on clouds' Stokes parameters. P_{max} varies in multiples of 0.05%, $\gamma_{\mathbf{B}_{\text{reg}}} = 0°$ and $A_{\text{turb}} = 0.2$ and 10 simulated samples are obtained by varying $\psi_{\mathbf{B}_{\text{reg}}}$. For the same f_{bg}, the solid lines connect the median of the 10 reconstructions on simulated samples with the same P_{max}. Filled (empty) symbols correspond to fits that pass (do not pass) the $\hat{\varpi}_{\text{C}}$ criterion explained in the text.

the significance with which a cloud can be detected given the data. The sizes of the posteriors on the Stokes parameters seem to be well determined irrespective of $p_{\text{C}}^{\text{true}}$ as soon as the recovery of the cloud distance is successful.

We show in Fig. 10, the significance of the detection of the cloud defined as the Mahalanobis distance of the zero polarization with respect to the estimated posterior distributions on the cloud polarization as a function of $p_{\text{C}}^{\text{true}}$. As expected, the significance with which a cloud can be detected depends both on the amplitude of the polarization signal that it induces to the background stars and on the fraction of stars in the background. The larger the $p_{\text{C}}^{\text{true}}$ and the larger the f_{bg}, the larger the detection significance.
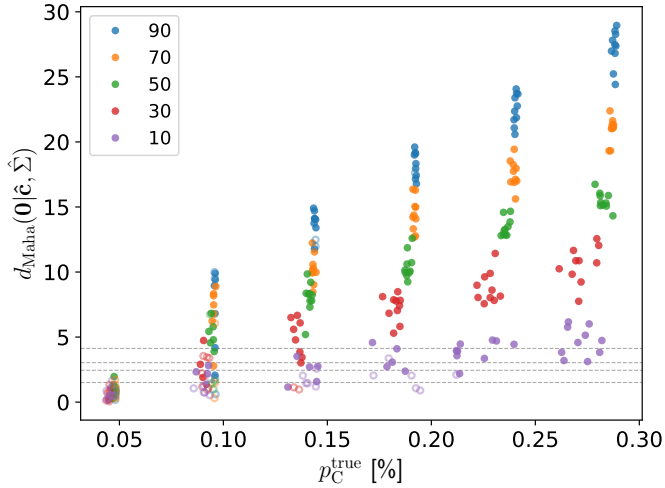
Fig. 10. Significance of the cloud detection ($d_{\mathrm{Maha}}(\mathbf{0}|\hat{\mathbf{c}}, \hat{\Sigma})$) as a function of the input cloud polarization and for the several f_{bg} values corresponding to the same reconstructions characterized through Fig. 9. The horizontal lines indicate threshold values corresponding to detection levels with 68%, 95%, 99%, and 99.98% probabilities of finding a distance lower than threshold. Symbol and color conventions are the same as in Fig. 9.

We repeat the performance analysis presented above with a level of intrinsic scatter given by $A_{\mathrm{turb}} = 0.1$. Owing to the fact that the intrinsic scatter is explicitly accounted for in the maximized log-likelihood, the performance on cloud parallax and mean polarization is only slightly better in this case than in the case of $A_{\mathrm{turb}} = 0.2$.

We finally repeat our performance analysis using starlight polarization data simulated from the mean star sample in a 0.5° diameter beam. The sample is made of 85 stars. The general picture depicted above remains similar. However, the decreased number of stars (a factor of about four) naturally implies a less accurate and precise reconstruction of the ISM along the LOS. The cloud parallaxes can be recovered with a relative difference lower than 15% for a majority of good reconstructions but we found more reconstructions that do not pass the criterion on $\hat{\varpi}_{\mathrm{C}}$. The drop in precision is due to the loss in density of stars that effectively sample the distance axis along the LOS. In what concerns the polarization properties, only 55% (80%) of the reconstructions that pass the criterion on ϖ_{C} and with $p_{\mathrm{C}}^{\mathrm{true}} > 0.1\%$ show an L_2 distance lower than 0.05% (0.1%). No bias is observed and the sizes of the posteriors are well defined. Due to the reduced number of stars (and the low polarization considered here) several reconstructions lead to cloud detections that do not exceed the level of 2σ ($d_{\mathrm{Maha}}(\mathbf{0}|\hat{\mathbf{c}}, \hat{\Sigma}) \gtrsim 2.45$). However, for $f_{\mathrm{bg}} \geq 0.5$ (corresponding to cloud distances lower than ≈ 750 pc) and $p_{\mathrm{C}}^{\mathrm{true}} \gtrsim 0.1\%$, most of the reconstructions are successful and the cloud is detected at high significance ($d_{\mathrm{Maha}}(\mathbf{0}|\hat{\mathbf{c}}, \hat{\Sigma}) > 4.12$).

We now investigate the possible effects of the inclination angle of the magnetic field with the POS on the performance of our inversion method. For this purpose we create a new set of simulated starlight polarization samples for the 1° diameter beam. We fix $A_{\mathrm{turb}} = 0.2$, explore the five thresholds in f_{bg}, sample the inclination angle $\gamma_{\mathbf{B}_{\mathrm{reg}}}$ from 0° to 75° in steps of 15° and adapt the P_{max} values so that the observed p_{C} values are similar and approximately about 0.2% using $P_{\mathrm{max}} = 0.2/\cos^2(\gamma_{\mathbf{B}_{\mathrm{reg}}})$ [%]. We generate 10 simulated samples for each pair of $(\gamma_{\mathbf{B}_{\mathrm{reg}}}, P_{\mathrm{max}})$ by varying $\psi_{\mathbf{B}_{\mathrm{reg}}}$. As before, we then apply our inversion method and study its performance. The results are shown in Fig. 11. For
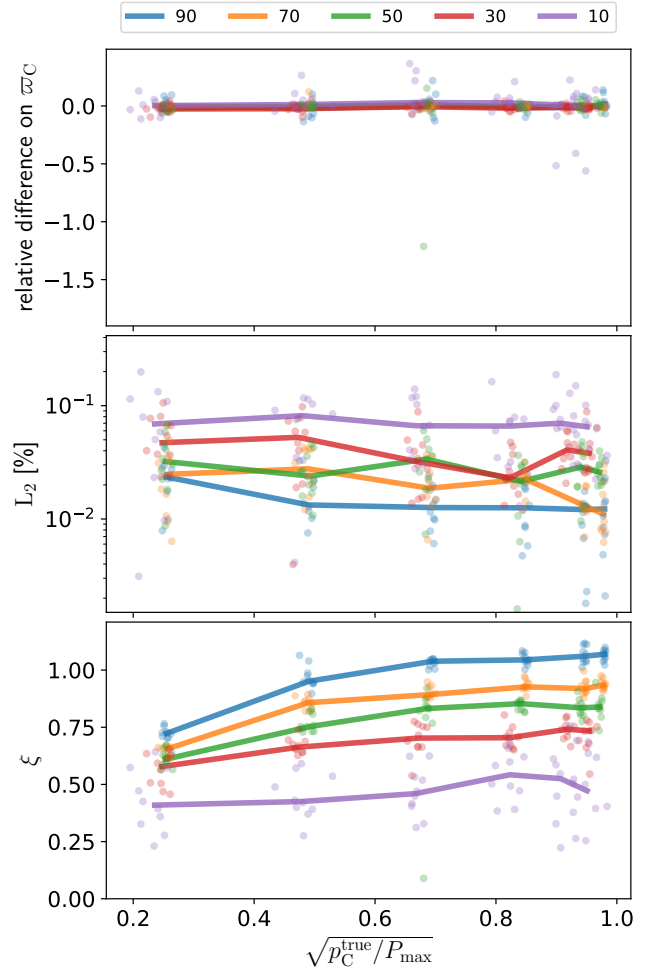
Fig. 11. Performance of the inversion method as a function of $\sim \gamma_{\mathbf{B}}$, the inclination angle of the magnetic field with respect to the POS. Conventions are the same as in Fig. 9. In absence of intrinsic scatter the abscissa would correspond to $\cos(\gamma_{\mathbf{B}_{\mathrm{reg}}})$.

fixed f_{bg}, the relative differences in parallax and the L_2 distances show very little dependence on the inclination angle. A mild dependence is observed for ξ: the larger the inclination angle the larger the size of the posteriors. However, the panels in L_2 and ξ clearly demonstrate that it is the number of stars in the background of the cloud that is decisive for the performance of the inversion method. For a similar degree of polarization, a low number of stars in the background of a far away cloud prevents us from recovering the cloud parallax with great accuracy and leads to loose constraints on its polarization properties. The latter is also seen in Fig. 12 where we show the significance with which the polarization of the cloud is detected compared to zero. A mild dependence of the significance as a function of the inclination angle is observed although we made sure that the mean degree of polarization is similar for all $\gamma_{\mathbf{B}_{\mathrm{reg}}}$. The more the magnetic field is inclined on the LOS the less the significance of the detection in the polarization plane. This is due to the fact that the scatter in starlight polarization is larger for higher inclination angles.

To summarize, for the case of a single cloud along the LOS we find that our inversion method is effective in recovering the cloud properties (parallax and mean polarization) when the polarization signal induced by the cloud to background stars is at least at the level of $\approx 0.1\%$. This threshold corresponds to the
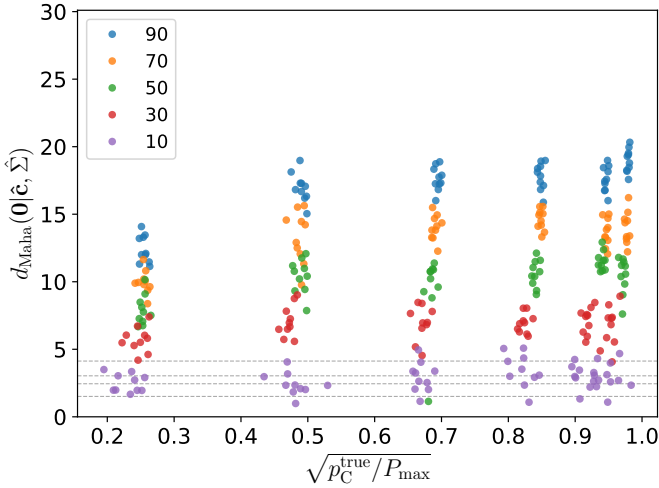
Fig. 12. Significance of the cloud detection in polarization as a function of $\sim \cos(\gamma_{\mathbf{B}})$. Conventions are the same as in Fig. 10.

minimum uncertainty in Stokes parameters of individual stellar measurements that we introduce in our mock observations to account for systematic uncertainties. Below this threshold the cloud parallax is not well recovered and we only recover loose constraints on the cloud polarization properties. Above this threshold the cloud parallax is recovered with high accuracy which increases with both the amplitude of the polarization signal and the fraction of stars in the background of the cloud. The latter also implies that the distance of far away clouds is generally recovered with less accuracy than for nearby clouds. Meanwhile, useful distance (parallax) constraints can be placed for all clouds nearer than about 750 pc even when considering star samples in a 0.5° diameter beam. We find that as soon as the cloud parallax is well recovered, the cloud polarization is recovered without bias and detected with a significance that depends on both the amplitude of the input polarization signal and the number of stars in the background of the cloud. The precision on cloud mean polarization can be as low as 0.05% for realistic cases, owing to the number of constraining stars. We find that, for fixed amplitude of induced polarization signal, the inclination of the magnetic field with respect to the POS (within the investigated range) does not affect the precision on the recovered cloud mean polarization degree but slightly changes the significance with which the polarization is detected. For realistic settings we additionally do not find a strong dependence of the performance of our method on the level of intrinsic scatter.

In conclusion, as soon as the polarization signal induced by a cloud to background stars is larger than the minimum uncertainties on individual polarization measurements and that at least about 30 stars are in the background of the cloud, our inversion method is effective in constraining the cloud parallax and its mean polarization properties. When the polarization signal is weaker or the number of background stars is smaller than that, the data are too limited to allow for the presence of the cloud to be identified with confidence. In these cases, the parameters are poorly constrained leading to unreliable reconstructions. These correspond to the pathological cases which we have shown how to identify. For most of these cases, the AIC values corresponding to a zero-layer model (i.e., no cloud along the LOS) are found to be lower than the AIC obtained with the one-layer model, suggesting no evidence for any cloud along the LOS given the data.

4.4. Two-layer cases

In this subsection, we test the performance of the method at recovering the structure of the magnetized ISM along the LOS when there are two clouds. Specifically, we are interested in exploring the limiting conditions that could lead to an imprecise reconstruction. In this respect, for a fixed level of stellar polarization uncertainties, we may generally expect that the overall performance of our inversion method will depend on the respective distance of the two clouds (affecting the number of stars in the foreground and the background of each cloud but also in between the clouds), the relative amplitude and weakness of the polarization signals as compared to noise and intrinsic scatter, and possibly the difference in magnetic field orientations.

Given the results presented in the previous subsection we might expect to be unable to recover the cloud properties (distance and polarization) when the far-away cloud has low polarization and or large distance. The detection of this second cloud may also be more difficult due to the additional scatter induced by the presence of a foreground cloud that not only modifies the zero-mean but also adds dispersion due to the intrinsic scatter. The contributions from the intrinsic scatter in the two clouds will also add for the stars in the background of the second cloud. Additionally, in cases where there are too few stars in between the two clouds or too few stars in the background of the far-away cloud, we expect that cloud properties of the nearby cloud might not be well retrieved because most of the polarization signal will then be attributed to only one of the two clouds. Generally, if the properties of one of the two clouds are not well recovered, this will automatically affect the reconstruction of the second cloud and therefore might compromise the overall reconstruction. We attempt to assess and quantify any such dependencies in this subsection by applying our inversion method to simulated starlight polarization data obtained from our toy model.

We narrow the range of possible ISM configurations by focusing on most realistic sky conditions. We fix the distance of the nearby cloud at 350 pc from the Sun. At high Galactic latitudes, this corresponds roughly to the maximum distance of the dusty and magnetized shell of the Local Bubble that surrounds the Sun (e.g., Skalidis et al. 2018; Pelgrims et al. 2020) and therefore the maximum distance at which we expect a LOS to have intersected the nearest cloud. In our star sample with 1° diameter beam, this leads to about 55 and 290 stars in the foreground and background of the nearby cloud, respectively. Additionally, we fix the inclination angle of the magnetic field permeating the nearby cloud to 30°, fix its P_{max} so that $p_C = 0.2\%$ and adopt a position angle of 22.5° to distribute equally the polarization signal over both Stokes parameters. According to the results obtained in the previous subsection, the properties of such a cloud are expected to be well recovered by our inversion method if this cloud is alone along the LOS.

We add a second cloud along the LOS and vary its distance, degree of polarization, and magnetic field position angle. We choose to fix the distance of the second cloud so that 90%, 70%, 50%, 30%, and 10% of the stars in the background of the nearby cloud are also background to the far-away cloud. The fraction of stars in the background of the nearby cloud that are also background to the more distant cloud is denoted f_{bg2}. Namely, the second cloud is placed at distance of about 520, 685, 900, 1150, and 1800 pc. The number of stars in between the two clouds are respectively $\sim$30, 90, 145, 200, and 260. For each cloud distance, we then vary the degree of polarization that the second cloud induces to its background stars and the relative position angle of the magnetic field in the far-away cloud with respect to the one
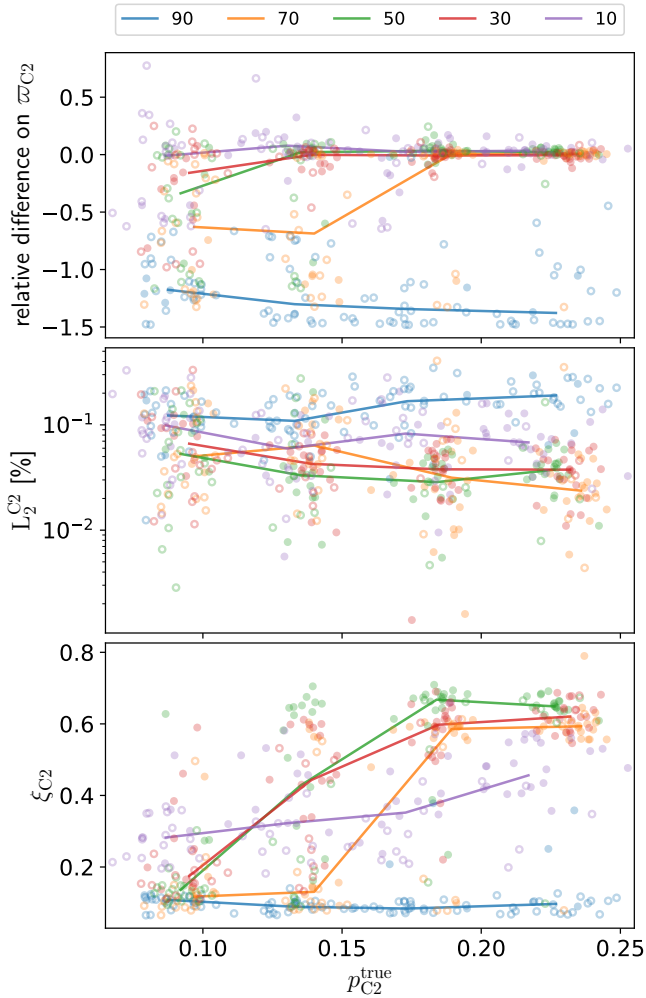
Fig. 13. Performance of the inversion method to retrieve properties of the distant cloud for two-cloud cases, as a function of p_{C2} and for all f_{bg2} values. Conventions are the same as in Fig. 9 except that filled symbols correspond to reconstruction in which both ϖ_{C1} and ϖ_{C2} pass the selection criterion.

in the nearby cloud ($\Delta\psi_{1,2}$) as follows: p_{C2} ranges from 0.1% to 0.25% in steps of 0.05%, and $\Delta\psi_{1,2}$ takes values from 0° to 150° in steps of 30°. We fix the inclination angle of the magnetic field in the far-away cloud to be zero. The level of intrinsic scatter in both clouds is set to $A_{turb} = 0.2$. Lowering this value in one or both of the clouds might result in slightly better performance.

We generate four random realizations of the simulation settings just described. In total, we consider 480 configurations for the magnetized ISM structure along the LOS, to each of which corresponds a mock starlight polarization sample of 345 stars. We apply our inversion method to each of these samples using a two-layer model. For all cloud parameters, except for the parallax, we consider flat priors as used before. For the parallaxes of the clouds, we define a flat prior so that the distance of the nearby cloud can range from 100 to 600 pc and the distance of the far-away cloud can range from 300 to 3500 pc. As before, we use 1000 live points and we sample the parameter space until we reach a tolerance on the estimated log-evidence below 0.1. Then, we analyze the resulting posterior distributions and characterize the results as before to infer the performance of the method.

In Fig. 13, we show the relative differences between the parallax of the second cloud obtained at maximum-likelihood value as compared to the parallax of the closest star to the input

cloud parallax, the L_2 distances on second cloud mean polarization (L_2^{C2}) and the ratio ξ of ideal-to-actual size of the posterior distributions on mean polarization for the second cloud (ξ_{C2}). Similarly to Fig. 9, we show the results for all values of f_{bg2}, the fraction of stars in the background of the nearby cloud that are also in the background of the far-away cloud, and present the data as a function of the true degree of polarization input to the far-away cloud. In this case, the filled (empty) symbols correspond to fits in which the posterior distributions of both cloud parallaxes pass (do not pass) the $\hat{\varpi}_C$ criterion.

We first notice that a large fraction of fits do not pass the joint selection criterion on cloud parallaxes. In fact, while most of the posteriors on ϖ_{C1} pass the criterion (more below), about 45% of the reconstructions lead to posterior distributions on ϖ_{C2} that do not pass it. Invalid posterior distributions generally appear for large f_{bg2} values and for low p_{C2}^{true} values, irrespective of f_{bg2}. As soon as $p_{C2} \gtrsim 0.1\%$ and $f_{bg2} \lesssim 70\%$, the posteriors on cloud parallaxes are valid and lead to reasonable constraints for more than 75% of the cases (this is for the cases that we have investigated, probing the limits of the method capabilities).

Generally, for both valid and invalid posterior distributions on ϖ_{C2}, the parallax of the nearby cloud is well constrained close to the input values. This is true except for a few cases which correspond to large values of f_{bg2}, $p_{C2} \gtrsim p_{C1}$, and particular $\Delta\psi_{1,2}$ which make it impossible to differentiate the bivariate distribution of (q_V, u_V) in the inter-cloud region from the bivariate distribution in the background of the more distant cloud. In such cases, because of the low number of stars sampling only the nearby cloud (≈ 30), because of the high level of uncertainties in individual polarization measurements, and because of the presence of the intrinsic scatter, the fit attributes all the polarization to only one of the two clouds which turns out to be the nearby one in order to account for the dispersion seen in the (q_V, u_V) plane. When $p_{C2} \approx p_{C1}$, the cloud is placed much closer than the input value when $\Delta\psi = 90°$ or much farther than the input value when $\Delta\psi = 0°$. The reason is that when $\Delta\psi \approx 90°$, the mean polarization of all the stars in the background of the second cloud approximate zero while in the case $\Delta\psi \approx 0°$ the majority of the those stars have a degree of polarization around 0.4%. The cloud distance is adapted to ensure the best compromise to minimize the residuals of individual star polarization and thus, to maximize the likelihood. When the polarization of the second cloud is lower, this dependence on $\Delta\psi$ tends to vanish simply because the induced polarization is somewhat negligible as compared to the polarization (and scatter) induced by the foreground cloud which are loosely constrained by only 30 stars. The tendency to drive the cloud parallax far away from the input value is thus not significant and the parallax of the nearby cloud is found close to its input value. In any case, when the second cloud is too close to the nearby cloud, the second-cloud parallax is not well recovered and it exhibits a pathological posterior. Such a fit should therefore be disregarded.

For reconstructions that satisfy the $\hat{\varpi}_C$ criterion for both clouds, we find that the relative differences on ϖ_{C1} are below 15% for 98% of the studied cases (or below 5% in 58% of the cases) and that of ϖ_{C2} are below 15% for 80% of the studied cases (or below 5% in 60% of the cases). For the nearby cloud, the Euclidean distances (L_2^{C1}) between the true Stokes parameters and those at maximum-likelihood differ by less than 0.05% in degree of polarization in 90% of the cases (or below 0.1% in 99% of the cases) and the ξ values corresponding to these reconstructions are generally at about one or greater. Thus we find very good accuracy and precision on the retrieved mean polarization of the foreground cloud. For the far-away cloud, the L_2^{C2}

are below 0.05% in 59% of the cases (or below 0.1% in 84% of the cases) and the ξ values are lower than one but above 0.5 for 64% of the valid reconstructions (regarding the $\hat{\varpi}_C$ criterion). The fact that ξ_{C2} is generally lower than ξ_{C1} is related to the presence of the intrinsic scatter already induced by the nearby cloud in addition to the one in the second cloud. We note that none of these is accounted for in the ideal estimate of the posterior size while computing the ξ ratio. Therefore we find that the mean polarization properties of the far-away cloud are also generally well retrieved as soon as $p_{C2} \gtrsim 0.1\%$ and if $f_{\text{bg2}} \in [70\%, 30\%]$, that is, if the number of stars in the inter-cloud region is large enough and if there are not too few stars in the background of the second cloud to constrain its properties.

In summary, we find that the inversion method works efficiently for a large range of possible configurations of the magnetized ISM with two clouds along the LOS. It fails at recovering the ISM structure if the mean polarization of the second cloud is comparable to the systematic uncertainty in the starlight polarization (i.e., if $p_C \approx 0.1\%$), if the number of stars in the inter-cloud region is smaller than ≈ 30, or if the number of stars in the background of the far-away cloud is lower than ≈ 30. In those cases, the method generally leads to satisfying constraints on the nearby cloud unless the magnetic fields in both clouds are mostly parallel or perpendicular to each other and there are too few stars ($\lesssim 30$) constraining the nearby cloud polarization properties.

4.5. Test on model selection

Finally, for our tests of method performance, we want to explore the use and relevance of our model selection criteria, using statistical evidence, the AIC, or both, to choose among the different trial models (i.e., to choose how many clouds are on the LOS), as described in Sect. 3.5. We limit our investigations to single- and two-layer cases in the low polarization regime and defer the exploration of more complex cases that might occur toward denser ISM regions than those anticipated for the PASIPHAE survey to further dedicated work.

We generate mock starlight polarization catalogs from the star sample within the $1°$ beam, placing either one or two clouds along the LOS as described below. We then apply our inversion method testing the one-layer and the two-layer models one after the other. We check the results of each reconstruction and flag those with pathological posterior distributions on cloud parallax, we record the evidence of the model given the data and estimate the value of the AIC from the obtained maximum log-likelihood, and finally proceed to model comparison.

4.5.1. One-cloud mock samples

We first consider the case when one dust cloud is introduced along the LOS. We use our toy model to produce the set of mock starlight polarization samples. We fix the distance of the cloud at 350 pc, impose a $p_C^{\text{true}} \simeq 0.2\%$ with $\gamma_{\mathbf{B}_{\text{reg}}} = 30°$, and a level of intrinsic scatter given by $A_{\text{turb}} = 0.2$. We then vary the position angle of the magnetic field from $0°$ to $162°$ in steps of $18°$ and finally generate 10 random realizations of each setup. In total, this makes 100 mock samples.

When using the one-layer model, we find that all reconstructions are valid in terms of the shape of the posterior distribution. Only 12 reconstructions from the two-layer model are valid in this respect, i.e., that posteriors on both ϖ_{C1} and ϖ_{C2} pass the selection criterion, implying that the one-layer model should already be favored against the two-layer model in 88% of cases.

We find that the evidence is higher for the one-layer model than for the two-layer model in all cases. 20% show an AIC lower for the two-layer model than for the one-layer model but the probability that, among the two tested models, the one-layer model actually minimizes the loss of information never goes below 36%. However, when limiting ourselves to the 12 samples for which we obtain valid reconstructions with the two-layer model, the AIC values are always lower for the one-layer model than for the two-layer model. The probability that, instead, the two-layer model actually minimizes the loss of information ($P_{2L|\{1L,2L\}}$) reaches a value as high as 20% for one case but more than half (a quarter) of the samples show this probability below the 5% (1%) threshold. We check that for the cases with $P_{2L|\{1L,2L\}} \gtrsim 1\%$ the significance of the far away cloud does not reach the 2σ-detection threshold (i.e., $d_{\text{Maha}}(\mathbf{0}|\hat{\mathbf{c}}, \hat{\Sigma}) < 2.45$ in all cases), meaning that the contribution to polarization from the model-retrieved far-way cloud is compatible with zero.

4.5.2. Two-cloud mock samples

We consider mock samples with two clouds along the LOS. The first cloud is fixed at 350 pc distance from the Sun and has $p_{C1} \simeq 0.2\%$ with an inclination angle of $30°$ and $A_{\text{turb}} = 0.2$, as for the 1-cloud mock samples used above. The POS position angle of the magnetic field in the nearby cloud ($\psi_{\mathbf{B}_{\text{reg}}}$) is fixed at $22.5°$. The second cloud is placed at a distance such that $f_{\text{bg2}} = 70\%$. This corresponds to a distance of about 685 pc which implies that only about 90 stars are located in the inter-cloud region, thus directly constraining the nearby cloud's polarization properties. We fix the inclination angle to $0°$, vary $\Delta\psi_{1,2}$ from $0°$ to $150°$ in steps of $30°$ as before, and consider two values for the far-away cloud polarization: $p_{C2} \simeq 0.15\%$ and 0.2%. We consider 10 random realizations of each setup making a total of 120 mock samples; 60 for each p_{C2} value. According to previous tests, we know that a significant fraction of runs with $p_{C2} \simeq 0.15\%$ will end with pathological posteriors cloud parallaxes but that this fraction should decrease, but not vanish, for $p_{C2} \simeq 0.2\%$ as inferred from Fig. 13.

In the cases with $p_{C2} \simeq 0.15\%$, we find that only 24 out of 60 trials lead to nonpathological posterior distributions of the cloud parallaxes. This implies that due to the low polarization induced by the second cloud (with respect to the level of noise and intrinsic scatter), we could detect the presence of the cloud in only about 40% of the cases. However, out of the 24, 23 have evidence higher for the one-layer model than for the two-layer model. According to the evidence, the two-layer model is to be favored in only one case out of 60. On the other hand, the AIC values are lower for the two-layer model than for the one-layer model (thus favoring the two-layer model) for 20 cases out of the 24 and the probability that instead the two-layer model should actually minimize the loss of information is higher than 1%, and as high as 10% for the four remaining cases. The comparison of the AIC values thus suggests that the two-layer model is to be favored in more than 80% of the cases that pass the selection criterion on ϖ_C. The hypothesis of a two-layer model is never rejected in those cases.

In the case with $p_{C2} \simeq 0.2\%$, we find that 44 out of 60 trials lead to nonpathological ϖ_C posterior distributions. Out of 44, 14 exhibit an evidence larger for the one-layer model than for the two layer-model, thus suggesting that the one-layer model is to be favored in about 30% of the valid reconstructions. However, the AIC values indicate that the one-layer model has to be preferred in only one case out of 44 and yet, for this case, $P_{2L|\{1L,2L\}} > 5\%$ which implies that the possibility for a two-layer model cannot be

disregarded. Therefore, while the evidence favors the two-layer model in only about 70% of the cases that pass the criterion on ϖ_C, the AIC criterion does it in more than 97% of the cases and never rejects the hypothesis of a two-layer model.

4.5.3. Conclusion

When testing the one-layer and two-layer models on mock samples generated with a 1-cloud model, we find that both the evidence and the AIC criterion favor the one-layer model. Hints for the presence of a spurious second cloud (corresponding to case where $P_{2L|\{1L,2L\}} \gtrsim 1\%$) may be encountered in less than 10% of the cases but none exhibit a significant detection in the polarization plane.

When testing the one-layer and two-layer models on mock samples generated with 2-clouds in both regimes where pathological reconstructions are dominant and subdominant, we find that comparison of AIC values is more sensitive to the presence of a second cloud than is the comparison of the evidences as it never leads to the rejection of the correct model but instead favors the correct model for a very large fraction of the cases that pass the $\hat{\varpi}_C$ selection criterion.

As a result, to select between models (number of clouds along the LOS), and after we disregard reconstructions with pathological posterior distributions, we prefer the use of the AIC values rather than the evidences. The comparison of the AIC values never suggests rejecting the correct model. It might, however, suggest spurious detection in some cases which, however, do not lead to an erroneously significant detection of a cloud.

We performed those tests in the low-polarization regime with a somewhat high level of scatter. We expect our conclusions to hold and the performance to improve when the polarization induced by intervening clouds increases with respect to the level of scatter in the polarization data (for example, a deeper survey than assumed here, or for regions of the sky with higher polarization).

5. Application to observational data

In this section we apply our Bayesian inversion method to existing observations of stellar polarization toward the two LOSs with known number of clouds presented by Panopoulou et al. (2019b). We first describe the data, then apply our method, and finally discuss our results in comparison with the previously obtained decomposition.

5.1. Archival polarization data in two LOSs of the diffuse ISM

Panopoulou et al. (2019b) demonstrated that it is possible to perform a tomographic decomposition of the POS magnetic field by combining a large number of stellar polarization data obtained with the RoboPol polarimeter (Ramaprakash et al. 2019) and *Gaia* distances. They selected two neighboring LOSs in the diffuse ISM which have different number of clouds along the LOS, as determined by their H I spectra. One sightline exhibits two distinct peaks (velocity components) corresponding to two clouds overlapping on the POS (2-cloud region). The other sightline exhibits a single prominent peak (1-cloud region). They obtained stellar polarimetry for ∼100 stars in each of these 0.32°-wide circular regions, and were able to recover the polarization properties of each cloud separately.

We choose to work with this data set as it is the only one, to the best of our knowledge, that approaches the expected spatial density of stellar measurements of the PASIPHAE survey in the
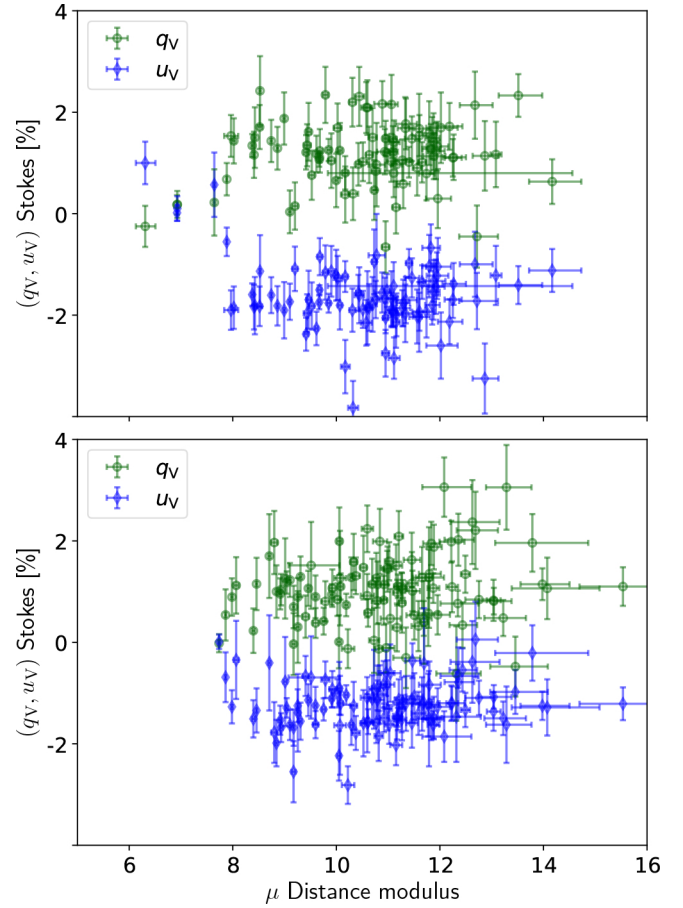
Fig. 14. $(q_V, u_V) - \mu$ plane for the polarization data in Panopoulou et al. (2019b). Top and bottom panels correspond to their 1-cloud and 2-cloud regions, respectively. The relative Stokes parameters q_V's are shown by dark green circles and the u_V's are shown by blue diamonds. Vertical and horizontal errorbars show the 68% confidence level on measurements on polarization and distance modulus, respectively. For visualization purposes we omit to show two stars at distance modulus 1.6 and 18.6 in the 1-cloud region.

optical, while targeting the diffuse ISM. We caution that despite this similarity, the aforementioned data set is not directly representative of the kind of data expected from PASIPHAE, for the following reasons. First, the selected regions in Panopoulou et al. (2019b) are at intermediate latitude, with mean stellar polarization of ∼2%, thus a factor of several higher than the average polarization expected at high Galactic latitudes. Second, their observing strategy did not mimic a survey with fixed exposure time, so that the uncertainties do not reflect exactly those expected by PASIPHAE. Nevertheless, the aforementioned data set remains the most appropriate one to test our Bayesian method and compare directly to the results of the existing tomographic decomposition presented in Panopoulou et al. (2019b).

In Fig. 14, we show their data in the $(q_V, u_V) - \mu$ plane, also showing uncertainties in both polarization and parallaxes (propagated to distance moduli). As noted in this previous work, the stellar data show systematic differences in the stellar polarization properties between the two regions, as a result of differences in the ISM across the sky (e.g., dust column density).

5.2. Results and comparison to previous work

We apply our inversion method to those two regions separately, trying models with one to three layers. We use flat priors on our

Table 3. Statistics for model comparison from inference of data in Panopoulou et al. (2019b).

Model	1-Cloud LOS				2-Cloud LOS			
	$\hat{\mathcal{Z}}$	$\log \hat{\mathcal{L}}$	AIC	$P_{j|\{m\}}$ (%)	$\hat{\mathcal{Z}}$	$\log \hat{\mathcal{L}}$	AIC	$P_{j|\{m\}}$ (%)
1-layer	671.1 ± 0.2	688.7	-1365.4	100	757.6 ± 0.2	775.0	-1538.0	100
2-layer	661.1 ± 0.3	690.6	-1357.1	1.5	750.8 ± 0.2	780.1	-1536.3	42.5
3-layer	655.6 ± 0.3	686.9	-1337.8	1.0×10^{-4}	743.9 ± 0.6	777.4	-1518.9	6.9×10^{-3}

six parameters. In the case of multiple clouds, our implementation choice guarantees that these are at least five stars in between two successive clouds. In the initialization of the flat priors, we further impose that all cloud parallaxes must lie in the parallax ranges defined so that the cloud's minimum distance is $100\,\text{pc}$ and a cloud's maximum distance is such that there are at least ten stars in the background. The limits on cloud parallax thus depend on the studied sample. The limits defining the parameter space of the polarization are rather arbitrary. We set a large range for the mean Stokes parameters of the clouds, from -5% to 5%, and consider that the elements of the intrinsic-scatter covariance matrix can go as high as 0.8×10^{-4} for the diagonal elements (this corresponds to a standard deviation on individual Stokes at the value of almost 1%). Again, internally, the prior function ensures that $C_{\text{int,qu}}^2 < C_{\text{int,qq}}\, C_{\text{int,uu}}$ for the covariance matrix to be invertible. The resulting statistics are given in Table 3 and we discuss the results for each region next.

5.2.1. 1-cloud region

The values of the AICs obtained for the set of tested models inform us that the one-layer model is preferred by the data in the sense that it minimizes the loss of information. The probability $P_{j|\{m\}}$ that the two-layer model is actually the model that should minimize the loss of information takes the value of 1.5% and that probability for the three-layer model is negligible. That is, there might be an indication for a second cloud in the data of the 1-cloud region but it is marginal. Assuming there is a second cloud along that LOS, the data is not sufficient to make the significance of the second cloud to show up. However, we have seen in Sect. 4.5 that such a low value might point to spurious detection. It is therefore safe to conclude that there is most likely only one cloud along this LOS.

The analysis of the posterior distributions of the best model leads to the following results for the cloud. The distance of the cloud is $382.7^{+10.9}_{-23.8}\,\text{pc}$ and the Stokes parameters are $(q_C, u_C) = ((1.21 \pm 0.05)\%, (-1.70 \pm 0.05)\%)$. Here, to summarize the posterior distributions, we report the medians of the distributions and their 16 and 84 percentiles. By re-sampling the posterior distributions of the Stokes parameters we obtain the distributions of degree of polarization and position angle characterized by: $p_C = (2.09 \pm 0.05)\%$ and $\psi_B = (-27.2 \pm 0.7)°$. These values are fully consistent with those obtained by Panopoulou et al. (2019b) for their 1-cloud region ($d_C \in [346, 393]\,\text{pc}$, $p_C = (2.04 \pm 0.04)\%$ and $\psi_B = (-27.5 \pm 0.6)°$).

For this fit, the parameter values corresponding to the maximum-likelihood are well centered in the distributions, with $d_C = 380.9\,\text{pc}$ and $(q_C, u_C) = (1.21–1.71\%)$. As we have seen in the previous section, situations with significant offsets between parameter values at maximum-likelihood and posterior distributions may happen when the data are not sufficient enough to

Table 4. Summary of posteriors on cloud parameters obtained for the 2-cloud region.

| Parameter	1-Layer	2-Layer	
	cloud 1	cloud 1	cloud 2
d_C (pc)	$370.1^{+15.7}_{-12.1}$	$371.2^{+15.5}_{-12.7}$	$2329.0^{+358.1}_{-426.2}$
q_C (%)	1.00 ± 0.05	0.96 ± 0.06	0.21 ± 0.18
u_C (%)	-1.28 ± 0.04	-1.33 ± 0.05	0.19 ± 0.11
p_C (%)	1.62 ± 0.04	1.64 ± 0.05	$0.32^{+0.15}_{-0.12}$
ψ_C (°)	-26.0 ± 0.9	-27.1 ± 1	$20.6^{+18.3}_{-12.6}$

Notes. Values indicates the median of the posterior distributions and the uncertainties are computed from the posteriors and indicate the offset from the median to the 16 and 84 percentiles. p_C and ψ_C are built from posteriors on q_C and u_C. We show results obtained with the 1-layer and 2-layer models only.

provide strong constraints on the cloud distance. This is not the case here.

5.2.2. 2-Cloud region

According to our Bayesian inversion method, the one-layer model is the model preferred by the data available for the 2-cloud region. Both the evidence and the AIC criteria agree, as reported in Table 3. However, the probability that the two-layer model actually minimizes the loss of information is very large ($\gtrsim 40\%$) and therefore we cannot reject the hypothesis that a second cloud exists along this LOS. The probability that the three-layer model minimizes the loss of information is negligible. The fact that both the evidence is smaller and the AIC is larger for the two-layer model than for the one-layer model, likely results from the fact that (i) fewer stars sample the second cloud and (ii) the signal of the second cloud is weak as compared to the signal of the dominant cloud and close to the noise level.

Summary statistics on the cloud model parameters are given in Table 4 for the case of the one-layer and two-layer model. For the fits obtained with the one- and two-layer models, the parameter values corresponding to the maximum-likelihood are well centered in their respective posterior distributions. We have $d_{C1} = 369.2\,\text{pc}$ and $(q_{C1}, u_{C1}) = (1.00–1.28\%)$ for the one-layer model and $d_{C1} = 360.2\,\text{pc}$ and $(q_{C1}, u_{C1}) = (0.93–1.33\%)$ and $d_{C2} = 2380.0\,\text{pc}$ and $(q_{C2}, u_{C2}) = (0.30–0.14\%)$ for the two-layer model. The fits and the results are robust and the data are sufficient to generate relatively stable and significant extrema in the log-likelihood hyper-surface. Therefore, we conclude that it is somewhat likely that there are two clouds in the 2-cloud region. The fact that the best-fit parameters of the nearby cloud from the two-layer model and those from the one-layer model are
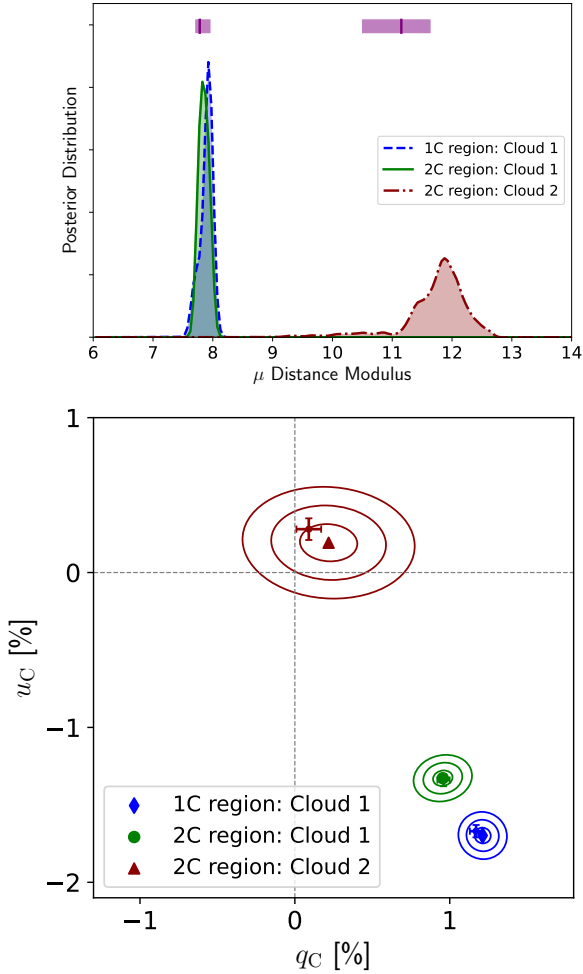
Fig. 15. Posterior distributions for the cloud distance modulus (top) and cloud mean polarization (bottom) obtained for the 1-cloud region while fitted with the one-layer model (blue) and for the 2-cloud region while fitted with the two-layer model with posteriors in green and dark red for the nearby and faraway cloud, respectively. The contours indicate the 1, 2, and 3σ confidence levels. The values obtained by Panopoulou et al. (2019b) are reported, using purple horizontal bands for the distance modulus on the top panel and errorbars to report polarization in the bottom panel using the same colors as for the confidence contours we obtain. Values for the 2-cloud region correspond to their distance cut that maximizes the detection of the faraway cloud.

very similar demonstrates that the polarization signal is dominated by the nearby cloud. This explains why there is no absolute strong evidence for a second cloud and that the one-layer model is ranked first according to the AIC and evidence criteria.

Our method returns results that are consistent with those obtained by Panopoulou et al. (2019b). To study the 2-cloud region, they fixed the nearby cloud distance at 360 pc, found that the far-away cloud distance that maximizes the detection is $d_{C2} \approx 1700$ pc to which they assigned an uncertainty of ± 440 pc. Using $d_{C2} \approx 1700$ pc, they found that the nearby and faraway cloud polarization properties were $(p_{C1}, \psi_{\mathbf{B}_{C1}}) = ((1.65 \pm 0.04)\%, (-27.3 \pm 0.8)°)$ and $(p_{C2}, \psi_{\mathbf{B}_{C2}}) = ((0.28 \pm 0.07)\%, (36 \pm 8)°)$, respectively.

5.2.3. Results comparison

In Fig. 15, we present and compare the posterior distributions for the cloud distance modulus and the cloud mean

polarization that we obtain with the one-layer model applied to the 1-cloud region and the two-layer model applied to the 2-cloud region. For comparison we also report the values obtained by Panopoulou et al. (2019b). It is seen that their results are compatible with our posterior distributions. The confidence intervals that we derive from our marginalized posterior distributions are larger than their uncertainties. This is true in particular for the polarization parameters of the faraway cloud. The reason is that our estimates include any possible correlation between parameters of the model while their estimates are conditional to the choice of both cloud distances.

We thus conclude that even applied to actual data, the method that we have developed makes it possible to perform a tomographic decomposition of the magnetized ISM in dusty regions. We emphasize that at no time did we rely on the a priori assumption that there are one or two clouds along the given LOSs. This shows that our inversion method is a standalone method in the sense that it can be used blindly on stellar polarization and distance data, and independently of any other observable, and lead to a reliable decomposition of the ISM signal along the LOS.

6. Conclusions and discussion

Starlight polarization is a direct tracer of the orientation of the POS component of the magnetic field in the dusty ISM. When combined with measurements of stellar distances, starlight polarization has the potential to allow for a 3D reconstruction of the magnetized ISM in dusty regions. In this paper, and motivated by the forthcoming leap forward in available stellar polarization data from the PASIPHAE survey, we have aimed to develop a robust method to perform such a 3D reconstruction.

We have developed a Bayesian method to reconstruct the POS magnetized ISM structure along the LOS through maximum-likelihood analysis of the stellar data alone. Our method, which relies on a generic model in which dust clouds have a thickness along the LOS smaller than the typical separation between stars, accounts for uncertainties in stellar parallaxes and in the Stokes parameters. It further accounts, in a model-independent way, for the intrinsic scatter that is expected from turbulence within individual clouds. We obtained a likelihood that accounts for all sources of noise and scatter and implemented it in Python. Our code depends on the nested sampling code `dynesty` to maximize the log-likelihood function, construct the posterior distributions of the model parameters, and estimate the evidence. The code, named `BISP-1`, for Bayesian Inference of Starlight Polarization in one dimension (along distance), is made public.

We tested our Bayesian inversion method on mock starlight polarization data obtained from a self-consistent toy model. We have demonstrated that our method is effective at recovering the cloud properties as soon as the polarization signal induced by a cloud to its background stars is higher than ~0.1%. When the minimum (systematic) uncertainty on observed stellar polarization is assumed to be at the level of 0.1% (in the degree of polarization) and the induced polarization is at a similar level, we found that ≈30 stars in the background of a cloud are required to place useful constraints on the cloud properties. The larger the induced polarization signal is, the better the method's performance, and the lower the number of required stars. In addition, to accurately recover the distances and the mean polarization properties of clouds, we found that our method also makes it possible to constrain the parameters characterizing the turbulence-induced intrinsic scatter. This might open new

avenues in the characterization of the ISM turbulence and estimation of the magnetic field strength. We will explore those avenues in future works.

We have further demonstrated that our Bayesian inversion method efficiently recovers cloud properties (distance and polarization) when applied to the actual data sets that were first used to demonstrate that starlight polarization, coupled to distance measurements, can be used to decompose the magnetized ISM signal as a function of distance (Panopoulou et al. 2019b). We obtained results that are fully consistent with those from the original study but within a robust Bayesian framework which allowed us to build proper posterior distributions on our model parameters of our reconstruction and, therefore, to put our results on a more solid footing. With this application we have shown that our method can work independently and blindly on star data to reconstruct the structure of the magnetized ISM. For example, we need not rely on external data to inform the method on the number of components along the LOS. This demonstrates the strength of our method, as well as the great potential of starlight polarization as a direct and fully independent probe to the 3D structure of the magnetized ISM.

Modifications of the sampling method implemented in this work are possible and could lead to speeding up the calculations. An example is a profiled-likelihood method that would estimate on the fly, and from the data, a certain number of the six free parameters per cloud. For example, the parameters related to the intrinsic scatter could be estimated from the scatter of the data points and their observational uncertainties once the cloud distances are fixed. The main advantage of such an implementation is the reduction of computing time owing to the reduction of the parameter-space dimensions. The main disadvantages are that the log-likelihood hyper-surface is no longer homogeneously sampled, preventing reliable estimates of the statistical evidence of a given model (number of clouds along the LOS), and that no proper posterior distributions of the parameters estimated on the fly can be safely reconstructed. We postpone the exploration of such alternative implementations to future work.

In its current implementation, our Bayesian inversion method relies on the dust-layer model that we have introduced. We expect for our model assumptions to hold for small beam sizes, so that any variations in the magnetic field and dust density do not vary appreciably. However, since a minimum number of stars is required to constrain the dust-cloud characteristics effectively, a minimum beam size is determined by both the true distribution of stars in space and the distance and multiplicity of clouds in the beam. While the analysis presented here focused on two beam sizes, we note that the minimum beam size allowed likely varies with sky position. One likely must evaluate, on a case-by-case basis, the trade-off between increasing the number of stars (widening the beam) and minimizing the variations in the ISM properties on the sky.

Furthermore, although the conceptual dust-layer model is supported by current observations of the high latitude sky, we must keep in mind that it might fail to account for all dust along a sightline and that careful analyses of the results and residuals will be mandatory, as it is generally the case in Bayesian modeling (Romero-Shaw et al. 2022). If, in the future, our simple layer-model prescription fails to fit the data, the underlying model will have to be changed, perhaps to be replaced by a smoother function of the LOS distance. In such a case, the generalization of the formalism that we have laid out will be straightforward and somewhat simpler than for the step model that we have assumed in this paper.

Finally, we must note that we have implicitly assumed in this paper that the polarization of stars is due to the magnetized ISM, only. However, stars of certain types may show intrinsic polarization, likely related to the existence of a circum-stellar disk where planets form or other asymmetries in the object (e.g., Cotton et al. 2016; Gontcharov & Mosenkov 2019). These stars usually show a higher degree of polarization than neighboring stars with unrelated position angles. In its current implementation, our method does not account for these intrinsically polarized stars. Therefore, to apply our decomposition method, the input stellar sample must have first been cleaned from potentially intrinsically polarized stars. Various techniques can be used for this purpose such as sigma clipping. Alternatively, in a Bayesian framework, we could consider constructing a likelihood of a mixture model in which the polarization of a star would have a probability of being due to the ISM or of being intrinsic given some of the star's properties. Such an approach to deal with outliers has been implemented in Zucker et al. (2019) to estimate the distance of molecular clouds based on the measured stellar extinction. However, we must defer such efforts to future work, given that basic statistical information on these sources (e.g. number density of intrinsically polarized stars per sky location) is currently lacking, and will only be known with the data from future unbiased surveys.

References

Akaike, H. 1974, IEEE Trans. Automatic Control, 19, 716
Alves, J., Zucker, C., Goodman, A. A., et al. 2020, Nature, 578, 237
Anche, R. M., Maharana, S., Ramaprakash, A. N., et al. 2022, SPIE Conf. Ser., 12188, 121882C
Andersson, B. G., & Potter, S. B. 2006, ApJ, 640, L51
Andersson, B. G., Lazarian, A., & Vaillancourt, J. E. 2015, ARA&A, 53, 501
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
Beck, R. 2015, A&ARv, 24, 4
Berdyugin, A., Piirola, V., & Teerikorpi, P. 2014, A&A, 561, A24
Bialy, S., Zucker, C., Goodman, A., et al. 2021, ApJ, 919, L5
BICEP2/Keck Collaboration, & Planck Collaboration 2015, Phys. Rev. Lett., 114, 101301
Boisbunon, A., Canu, S., Fourdrinier, D., Strawderman, W., & Wells, M. T. 2014, Int. Stat. Rev., 82, 422

Cho, J., & Lazarian, A. 2003, MNRAS, 345, 325
Clemens, D. P., Cashman, L. R., Cerny, C., et al. 2020, ApJS, 249, 23
Cotton, D. V., Bailey, J., Kedziora-Chudczer, L., et al. 2016, MNRAS, 455, 1607
Davis, Leverett, J., & Greenstein, J. L. 1951, ApJ, 114, 206
Diego-Palazuelos, P., Eskilt, J. R., Minami, Y., et al. 2022, Phys. Rev. Lett., 128, 091302
Doi, Y., Hasegawa, T., Bastien, P., et al. 2021, ApJ, 914, 122
Draine, B. T., & Hensley, B. S. 2021, ApJ, 919, 65
Eswaraiah, C., Lai, S.-P., Chen, W.-P., et al. 2017, ApJ, 850, 195
Foreman-Mackey, D. 2016, J. Open Source Softw., 1, 24
Franco, G. A. P., & Alves, F. O. 2015, ApJ, 807, 5
Gaia Collaboration (Prusti, T., et al.) 2016, A&A, 595, A1
Gaia Collaboration (Brown, A. G. A., et al.) 2021, A&A, 649, A1
Goldreich, P., & Sridhar, S. 1995, ApJ, 438, 763
Gontcharov, G. A., & Mosenkov, A. V. 2019, MNRAS, 483, 299
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S., & Finkbeiner, D. 2019, ApJ, 887, 93
Heiles, C. 2000, AJ, 119, 923
Heiles, C., & Crutcher, R. 2005, in Cosmic Magnetic Fields, 664, eds. R. Wielebinski, & R. Beck, 137
Hennebelle, P., & Inutsuka, S.-i. 2019, Front. Astron. Space Sci., 6, 5
Hensley, B. S., & Draine, B. T. 2021, ApJ, 906, 73
Hervías-Caimapo, C., & Huffenberger, K. M. 2022, ApJ, 928, 65
Hiltner, W. A. 1949, Science, 109, 165
Hiltner, W. A. 1951, ApJ, 114, 241
Hu, Y., & Lazarian, A. 2023, MNRAS, 519, 3736
Huang, Z. 2022, Universe, 8, 423
Ivanova, A., Lallement, R., Vergely, J. L., & Hottier, C. 2021, A&A, 652, A22
Jordi, K., Grebel, E. K., & Ammon, K. 2006, A&A, 460, 339
Jordi, C., Gebran, M., Carrasco, J. M., et al. 2010, A&A, 523, A48
Konstantinou, A., Pelgrims, V., Fuchs, F., & Tassis, K. 2022, A&A, 663, A175
Kwon, J., Tamura, M., Hough, J. H., et al. 2015, ApJS, 220, 17
Lallement, R., Babusiaux, C., Vergely, J. L., et al. 2019, A&A, 625, A135
Lallement, R., Vergely, J. L., Babusiaux, C., & Cox, N. L. J. 2022, A&A, 661, A147
Leike, R. H., & Enßlin, T. A. 2019, A&A, 631, A32
Leike, R. H., Glatzle, M., & Enßlin, T. A. 2020, A&A, 639, A138
Li, H.-B. 2021, Galaxies, 9, 41
Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021, A&A, 649, A2
Magalhães, A. M., Pereyra, A., Melgarejo, R., et al. 2005, ASP, Conf. Ser., 343, 305
Maharana, S., Kypriotakis, J. A., Ramaprakash, A. N., et al. 2020, SPIE Conf. Ser., 11447, 114475E
Maharana, S., Kypriotakis, J. A., Ramaprakash, A. N., et al. 2021, J. Astron. Telescopes Instrum. Syst., 7, 014004
Maharana, S., Anche, R. M., Ramaprakash, A. N., et al. 2022, J. Astron. Telescopes Instrum. Syst., 8, 038004
Marchwinski, R. C., Pavel, M. D., & Clemens, D. P. 2012, ApJ, 755, 130

Martínez-Solaeche, G., Karakci, A., & Delabrouille, J. 2018, MNRAS, 476, 1310
Miville-Deschênes, M. A., Duc, P. A., Marleau, F., et al. 2016, A&A, 593, A4
Montier, L., Plaszczynski, S., Levrier, F., et al. 2015, A&A, 574, A135
Mouschovias, T. C., Tassis, K., & Kunz, M. W. 2006, ApJ, 646, 1043
Panopoulou, G. V., & Lenz, D. 2020, ApJ, 902, 120
Panopoulou, G. V., Psaradaki, I., & Tassis, K. 2016, MNRAS, 462, 1517
Panopoulou, G. V., Hensley, B. S., Skalidis, R., Blinov, D., & Tassis, K. 2019a, A&A, 624, A8
Panopoulou, G. V., Tassis, K., Skalidis, R., et al. 2019b, ApJ, 872, 56
Patat, F., Maund, J. R., Benetti, S., et al. 2010, A&A, 510, A108
Pavel, M. D., Clemens, D. P., & Pinnick, A. F. 2012, ApJ, 749, 71
Pelgrims, V. 2019, A&A, 622, A145
Pelgrims, V., Ferrière, K., Boulanger, F., Lallement, R., & Montier, L. 2020, A&A, 636, A17
Pelgrims, V., Clark, S. E., Hensley, B. S., et al. 2021, A&A, 647, A16
Pelgrims, V., Ntormousi, E., & Tassis, K. 2022, A&A, 658, A134
Pereyra, A., & Magalhães, A. M. 2004, ApJ, 603, 584
Planck Collaboration Int. XXXV. 2016, A&A, 586, A138
Planck Collaboration XXX. 2016, A&A, 586, A133
Planck Collaboration XI. 2020, A&A, 641, A11
Planck Collaboration XII. 2020, A&A, 641, A12
Planck Collaboration Int. XXXVIII. 2016, A&A, 586, A141
Planck Collaboration Int. L. 2017, A&A, 599, A51
Ramaprakash, A. N., Rajarshi, C. V., Das, H. K., et al. 2019, MNRAS, 485, 2355
Rezaei Kh., S., & Kainulainen, J. 2022, ApJ, 930, L22
Rezaei Kh., S., Bailer-Jones, C. A. L., Soler, J. D., & Zari, E. 2020, A&A, 643, A151
Ritacco, A., Boulanger, F., Guillet, V., et al. 2023, A&A, in press, https://doi.org/10.1051/0004-6361/202244269
Robin, A. C., Luri, X., Reylé, C., et al. 2012, A&A, 543, A100
Romero-Shaw, I. M., Thrane, E., & Lasky, P. D. 2022, PASA, 39, e025
Santos, F. P., Franco, G. A. P., Roman-Lopes, A., Reis, W., & Román-Zúñiga, C. G. 2014, ApJ, 783, 1
Skalidis, R., & Pelgrims, V. 2019, A&A, 631, L11
Skalidis, R., & Tassis, K. 2021, A&A, 647, A186
Skalidis, R., Panopoulou, G. V., Tassis, K., et al. 2018, A&A, 616, A52
Skalidis, R., Sternberg, J., Beattie, J. R., Pavlidou, V., & Tassis, K. 2021, A&A, 656, A118
Skalidis, R., Tassis, K., Panopoulou, G. V., et al. 2022, A&A, 665, A77
Skilling, J. 2004, in American Institute of Physics Conference Series, 735, 395
Soler, J. D., Alves, F., Boulanger, F., et al. 2016, A&A, 596, A93
Speagle, J. S. 2020, MNRAS, 493, 3132
Sugitani, K., Nakamura, F., Watanabe, M., et al. 2011, ApJ, 734, 63
Tassis, K., & Pavlidou, V. 2015, MNRAS, 451, L90
Tassis, K., Ramaprakash, A. N., Readhead, A. C. S., et al. 2018, ArXiv e-prints, [arXiv:1810.05652]
Zucker, C., Speagle, J. S., Schlafly, E. F., et al. 2019, ApJ, 879, 125
Zucker, C., Goodman, A. A., Alves, J., et al. 2022, Nature, 601, 334

Appendix A: Mock starlight polarization data

To test and validate the inversion method presented in this paper we developed a pipeline to generate mock samples of starlight polarization data. Our simulation pipeline has the main following characteristics: (*i*) it makes use of star samples extracted from the *Gaia* data to guarantee realistic distributions of star distances, number density, brightness, etc., (*ii*) it relies on a self-consistent implementation of the thin-layer dust-cloud model discussed in Sect. 2 and presented below which includes a prescription for the intrinsic scatter, and finally (*iii*) it computes and propagates to the simulated stellar polarization data realistic uncertainties as expected for the PASIPHAE survey. These three parts are described in the rest of this Appendix. We further explore our toy-model to simulate the polarization signal from the magnetized ISM in Appendix B.

Our simulation pipeline works as follows. First, we choose a sample of stars in a small-aperture beam and extract parallaxes, parallax uncertainties and apparent magnitudes. Second, we simulate the stellar polarization measurements for any chosen setup of the magnetized and dusty ISM and estimate the polarization uncertainties. We 'observe' the stars by randomly drawing the Stokes parameters and parallaxes within their respective uncertainties.

A.1. Star samples from the Gaia Universe Model Snapshot

We seek to create realistic samples of stellar photometry and distance, based on the expected sky footprint of the PASIPHAE survey. In this section we discuss how we generate and choose representative samples of stars from the GUMS database.

A.1.1. Mock stellar catalog generation

To obtain realistic samples of stellar parameters (photometry, distances) for the high Galactic latitude sky, we use the GUMS database (Robin et al. 2012) associated with Gaia Early Data Release 3 (EDR3, Lindegren et al. 2021), which provides some photometric information as well as parallaxes. We begin by selecting sightlines toward each of the HEALPix $N_{side} = 4$ pixels at high Galactic latitude (with $|b| > 60°$). We then query the *Gaia* archive within a circular region centered on each of the aforementioned pixels, with two different radii of 0.5° and 0.25°. For ease of computation, we exclude very faint stars with G > 18.5 within the query. This brightness cut is later superseded by a stricter cut based on our derived *r* band photometry, as explained next. We thus obtain 24 samples of stars for each selected beam size.

Next, we wish to compute realistic uncertainties for the parallax and polarization of each star. The GUMS query returns the $V - I$ color, the *V*-band absolute magnitude M_V, the extinction A_V, the apparent magnitude in the *G*-band, the *Gaia* broadband color G_{BP} and the barycentric distance, d, in parsec. We determine the parallax errors as a function of *G*-band magnitude based on EDR3 performance (Lindegren et al. 2021). We use the quoted median parallax error per bin of *G* of their table 4 (5-parameter solution sources) and create an interpolating function. Each star is thus assigned a true parallax by inverting the distance and converting to units of milli-arcseconds (mas) and an uncertainty based on its *G* band magnitude. The estimation of uncertainties in the polarization measurements from PASIPHAE relies on (a) SDSS *r*-band photometry and (b) the expected performance of the survey. We derive *r*-band magnitudes in the following section and briefly discuss the dependence

of the uncertainty in the stellar Stokes parameters on *r*-band magnitude in Appendix A.4.

The stellar properties of the samples cover a broad range of apparent magnitudes, parallaxes and S/N in parallax (Fig. A.2, left, shows a representative sample of size 1°). We tailor the stellar samples to the expected limiting brightness of the PASIPHAE survey, by applying a cut of $r < 16$ mag. The total number of stars per sample (post-cut) varies in the range [225, 414] and [64, 118] for the 1° and 0.5° beams, respectively. For ease of computation, we choose one representative sample for each beam size on which to apply our analysis. The representative sample for each beam is that which has a total number of stars (post-brightness-cut) close to the average number of stars of the ensemble of sightlines. We further limit the samples to stars with a S/N in parallax higher than 5.

As shown in Fig. A.1 (left), the distribution of stellar parallaxes (distances) peaks at ~1 mas (1000 pc), but this is brightness-dependent (see Fig. A.2, right). As one would expect, our brightness cut limits the number of stars at small parallaxes (large distances). We also note there is a less intuitive selection bias that results from our choice to use stars with $\varpi/\sigma_\varpi > 5$. Fig. A.2 shows the line of $\varpi/\sigma_\varpi = 5$ in two different planes (SNR$_\varpi$ vs. ϖ and *r* vs. ϖ). As seen in the right panel, imposing the S/N cut results in a parallax-dependent brightness selection: stars to the left and top of the dashed line are excluded from our analysis. Our samples miss some faint and nearby stars. This is the result of the dependence of the parallax error on stellar brightness. In other words, the limiting magnitude of star samples analyzed in this work is parallax-dependent.

A.2. Derivation of stellar *r*-band photometry given GUMS outputs

We must determine the apparent magnitude in the SDSS-*r* band. This requires some manipulation of different color transformations of the photometric data provided by GUMS. We begin by computing the apparent magnitude in the *V* band:

$$m_V = M_V + 5\log(d) - 5 + A_V. \tag{A.1}$$

We can connect the SDSS photometry with *Gaia G* band photometry using the relations from Jordi et al. (2010), as initially referenced by the GUMS paper (Robin et al. 2012):

$$G - G_{BP} = -0.1703 - 1.0813\,(r - i) \\ - 0.1424\,(r - i)^2 + 0.0271\,(r - i)^3, \tag{A.2}$$

where the uncertainty of the coefficients is $\sigma = 0.1$. We solve this equation for the color $r - i$. We can then use the color transformations by Jordi et al. (2006):

$$(i - I) = (0.247 \pm 0.003)\,(R - I) + (0.329 \pm 0.002) \tag{A.3}$$
$$(r - i) = (1.007 \pm 0.005)\,(R - I) - (0.236 \pm 0.003)\,. \tag{A.4}$$

Solving the second expression for $(R - I)$ and substituting for $(R - I)$ in the first expression, we obtain:

$$(i - I) = 0.0578 + 0.245\,(r - i) + 0.329\,. \tag{A.5}$$

Using the color $(r - I)$, we substitute $(i - I)$ from the previous expression and we solve for m_r:

$$m_r - m_I = (r - I) = (r - i) + (i - I) \tag{A.6}$$
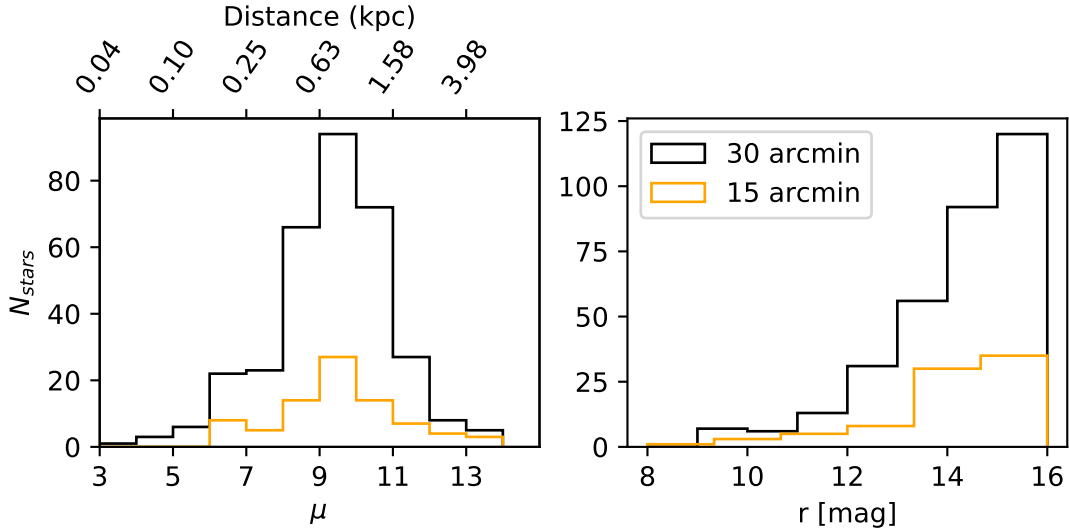$$m_r = m_I + (r - i) + 0.0578 + 0.245\,(r - i) + 0.329. \tag{A.7}$$

Fig. A.1: Distribution of stellar distances (left) and apparent magnitudes (right) after the selection criteria have been applied. The distributions for the large (0.5° radius) and small (0.25° radius) beams are shown with black and orange lines, respectively.
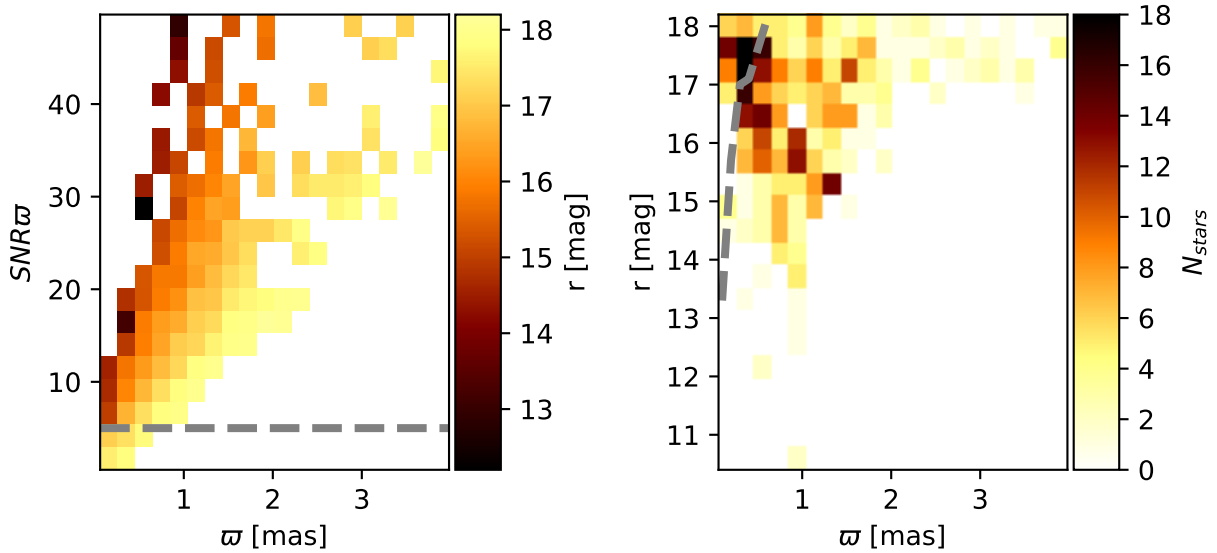

Fig. A.2: Stellar properties and selection criteria. Left: Mean SDSS r-band magnitude in bins defined on the parallax S/N vs. parallax plane. Bins with no stars are shown in white. Right: 2D distribution of the number of stars in the r vs. ϖ plane. The cut at $SNR_\varpi = 5$ is shown with a dashed gray line in both panels. We note that the introduction of the parallax S/N cut imposes a biased selection in the $r - \varpi$ plane. The data shown are for the large (30-arcmin radius) beam. We do not show the full dynamic range of ϖ and SNR_ϖ for better visualization of the key features of the dataset.

We finally obtain an expression for SDSS r-band apparent magnitude:

$$m_r = m_I + 1.245\,(r - i) + 0.3868, \qquad (A.8)$$

where $m_I = m_V - (V - I)$ can be computed from the $V - I$ color provided in GUMS.

A.3. ISM polarization signal: A toy model

Both the 3D orientation of the magnetic field and the dust density distribution are expected to vary within a cloud, at least to some extent, due to compressible MHD turbulence. These two effects induce a spread in the polarization signal of stars in the background of a cloud (the intrinsic scatter). This is implemented by considering the signal part to be made of two components: a regular component and a stochastic component. To build intuition on the effect of intrinsic scatter on the Stokes parameters, we simplify the problem and only implement the scatter that arises from fluctuations in the 3D geometry of the magnetic field. We defer a treatment of fluctuations in the dust distribution to future work. Hence, for a given cloud, a single value of the maximum degree of polarization (P_{max}) is used to model the polarization signal of all background stars (Eq. 1). On the contrary, both the inclination angle and the position angle of the magnetic field are considered to vary from star to star because of fluctuations in magnetic field geometry.

We model the total 3D magnetic field as the sum of a regular component ($\mathbf{B}_{reg}$) and a stochastic one ($\mathbf{B}_{sto}$). We consider $\mathbf{B}_{reg}$ to be uniform within a cloud and to have a norm of unity.

We model $\mathbf{B}_{sto}$ through 3D Gaussian realizations of white noise. That is, the stochastic component is built from realizations of 3D isotropic random vectors obtained from sampling an independent normal distribution for each of its three components. Assuming a large sample of random realizations for $\mathbf{B}_{sto}$ we estimate the sample *rms* and use it to normalized all the random draws. The distributions of the norms of $\mathbf{B}_{sto}$ have a standard deviation of one. To reach a statistically stable normalization of $\mathbf{B}_{sto}$, the *rms* evaluation should be performed for a sufficiently large sample of realizations. We use at least 1000 realizations. Finally, we model the 3D total magnetic field as the sum of a regular ($\mathbf{B}_{reg}$) and a stochastic component:

$$\mathbf{B}_{tot} \propto \mathbf{B}_{reg} + A_{turb}\,\mathbf{B}_{sto}\,, \qquad (A.9)$$

where A_{turb} quantifies the amplitude of the stochastic component with respect to the regular one. A different realization of $\mathbf{B}_{sto}$ is attributed to each star in the sample. As a result, to each star corresponds a different inclination and position angle. Therefore, by virtue of Eq. (1), different values of the Stokes parameters are obtained only from fluctuations in the 3D geometry of the magnetic field. We emphasize that the parameter A_{turb} used above should not be confused with the turbulent-to-mean magnetic field ratio used in methods to estimate the strength of the magnetic field (e.g., Skalidis & Tassis 2021). A_{turb} is a metric of the statistical fluctuations of the magnetic field geometry in 3D and is model dependent. We find in Appendix B that values in the range 0.1 to 0.3 may be representative of clouds at intermediate and high Galactic latitudes.

To summarize, our toy model has five free parameters per cloud: the cloud parallax ($\varpi_C = 1/d_C$), the maximum degree of polarization (P_{max}), the inclination ($\gamma_{\mathbf{B}_{reg}}$) and position ($\psi_{\mathbf{B}_{reg}}$) angles of $\mathbf{B}_{reg}$ and, finally, the relative amplitude of fluctuations in magnetic field orientation (A_{turb}). We notice that, apart from the cloud parallax, these parameters are not the same as the model parameters entering our data equation (Eq. 10) and, therefore, are not the parameters being sampled in the maximization process.

Due to projection effects, and since the inclination angle is positively defined, the mean inclination angle of the total magnetic field may generally be larger than the inclination angle of the regular magnetic field given as input. In general this results in a depolarization of the cloud signal which may be stronger in one of the two polarization channels (Stokes q_V or u_V) depending on the position angle $\psi_{\mathbf{B}}$. Therefore, in addition to producing a scatter in the polarization plane, the stochastic component in the 3D magnetic field geometry induces a bias in the mean Stokes parameters. We further explore our toy model in Appendix B with a particular emphasis on the bias and the covariance that magnetic-field fluctuations produce. Here, we note that the bias is not physical and simply originates from the construction of our model. What matters, and what should be modeled from real observations, are the mean values of the Stokes parameters of a cloud and the dispersion around these means.

A.4. Realistic noise

Once the polarization signal from the input ISM setup is attributed to each star in the sample, we add noise in both polarization and parallax. We take the values of parallax uncertainties from our samples of stars derived from the *Gaia* data (Appendix A.1) and the star parallaxes are simply randomized within their uncertainty range following Gaussian distributions. To give realistic uncertainties on the Stokes parameters of
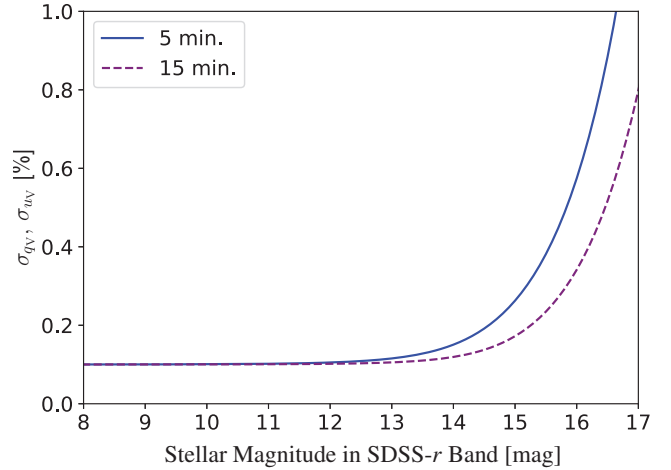

Fig. A.3: Uncertainties on individual measurements of the relative Stokes parameters in SDSS-r band as a function of star's magnitude as expected from PASIPHAE's northern instruments (WALOP-N) for 5 minutes and 15 minutes exposure times shown with solid blue and dashed purple lines, respectively.

starlight polarization, we rely on current expectations of the performance of the WALOP-N instrument to be used for the PASIPHAE survey in the northern hemisphere. The observational uncertainties, which can be estimated from the optical modeling of the WALOP-N instrument (Maharana et al. *in prep.*), depend primarily on the magnitude of the stars in the observation band (SDSS-r band) and on the exposure time. To take into account instrumental systematics, we add to our observational uncertainty budget a contribution of 0.1% to the two relative Stokes values q_V and u_V (see Maharana et al. 2020; Maharana et al. 2021; Maharana et al. 2022; Anche et al. 2022). Therefore, for each star, an estimate of the observational uncertainty is obtained and used to randomize the Stokes parameters obtained from the ISM configuration. The randomization of q_V's and u_V's is performed independently. The uncertainty in the Stokes parameters as a function of star magnitude is shown in Fig. A.3 for two typical exposure times. Examples of mock starlight polarization data obtained for a five minutes exposure time, for a beam of 0.5° circular diameter, and a median number density are shown in Fig. 1 for the cases of one and two clouds along the LOS.

Appendix B: Polarization observables and intrinsic scatter

For the purpose of this work we have developed and implemented a toy model, described in Sect. A.3, that allows us to generate mock observations based on the multilayer magnetized ISM paradigm described in Sect. 2. In particular, our implementation makes it possible to account for a source of intrinsic scatter in the simulated data. We expect such scatter to exist because of small-scale variations in the ISM, that generically result from turbulence. Generally fluctuations are expected in both density distribution and magnetic field geometry but dust grain properties, dust temperature, etc. could also vary on small scales. In our modeling, and as described in Sect. A.3, we only consider fluctuations in 3D geometry of the magnetic field and thus ignore fluctuations in dust density distribution or other properties that may affect the value of the maximum degree of polarization P_{max}

(see Eq. 1). In this appendix we explore our toy-model implementation to gain intuition on the effects of the intrinsic scatter on the Stokes parameters for different levels of intrinsic scatter (controlled through the value of A_{turb} – see Eq. A.9) and for several combinations of the inclination angle ($\gamma_{\mathbf{B}_{reg}}$) and position angle ($\psi_{\mathbf{B}_{reg}}$) of the regular component of the magnetic field. In the adopted polarization convention (IAU), q_V is maximum for $\psi_{\mathbf{B}_{reg}} = 0°$ and zero at $\psi_{\mathbf{B}_{reg}} = 45°$ where u_V is maximum.

We begin by considering the relative Stokes parameters normalized by the value of P_{max}. Therefore, in the absence of a stochastic component, the degree of polarization varies from 0 to 1 for the case of magnetic field pointing toward the observer ($\gamma_{\mathbf{B}_{reg}} = 90°$) to the case where the magnetic field lies in the POS ($\gamma_{\mathbf{B}_{reg}} = 0°$). From geometrical considerations, we expect that the effect of the addition of a stochastic component on the values of the Stokes parameters q_V, u_V will depend on the chosen value of A_{turb}, on $\gamma_{\mathbf{B}_{reg}}$, and possibly on $\psi_{\mathbf{B}_{reg}}$. To infer such dependence, we want to consider different orientations of the regular magnetic field as seen from the observer.

We consider the following setup. We place an observer above the pole of the northern hemisphere ($b \geq 0°$) of a HEALPix map with $N_{side} = 2$ (Górski et al. 2005). This map contains 28 pixels to which correspond 3D orientations with colatitude and longitude coordinates. The former corresponds to the inclination of the magnetic field line with respect to the POS and latter to the position angle in the POS. By construction, these orientations are symmetric in longitude and therefore only half of the points do show different Stokes parameters given the definition of the polarization; $\psi_{\mathbf{B}}$ is a two-circular quantity defined in the range $[0°, 180°]$. Hence, we drop half of the northern hemisphere and thus consider 14 combinations of ($\gamma_{\mathbf{B}_{reg}}$, $\psi_{\mathbf{B}_{reg}}$). To each, corresponds a pair of (q_V, u_V) at position $\cos^2 \gamma_{\mathbf{B}_{reg}}$ ($\cos[2\psi_{\mathbf{B}_{reg}}]$, $\sin[2\psi_{\mathbf{B}_{reg}}]$) in the polarization plane (see Eq. 1). Those positions are shown as red crosses in Fig. B.1. These crosses correspond to the Stokes parameters that would be observed for all background stars if the magnetic field was made of the regular component only. Then, we add the stochastic component in magnetic field geometry according to Eq. A.9. This creates variations about those values. Example of variations obtained in the polarization plane for $A_{trub} = 0.1$ is shown in Fig. B.1.

For most of the 14 configurations, the 2D scatter produced in the polarization plane cannot be considered as resulting from two independent normal distributions; the correlation coefficient is not zero in general. Instead, configurations with the regular magnetic field close to the POS are significantly asymmetric. Interestingly, depending on the position angle, the correlation coefficient can take large values. For example, in the case of $\gamma = 0°$ and $\psi = 22.5°$, the correlation coefficient is close to -1. The more the regular magnetic field points to the observer, the more the scatter becomes symmetric in q_V and u_V.

At this stage, we notice that fluctuations in dust density distribution, or any other ISM parameters affecting the value of P_{max}, would add radial scatter on this plot. These would then reduce the possible asymmetry and therefore would reduce the value of the correlation coefficient. From the shape of the 2D scatter seen in Fig. B.1 we may conclude that modeling the effect of the intrinsic scatter through the characterization of a bivariate normal distribution with given covariance matrix is a fair approach which should be general enough to cover a large class of models of sources of fluctuations in the magnetized ISM. This motivates our choice while writing the model equation in the main text (see Sect. 2). From this point, we therefore consider that the polarization signal in the (q_V, u_V) plane, mean and intrinsic scatter, can
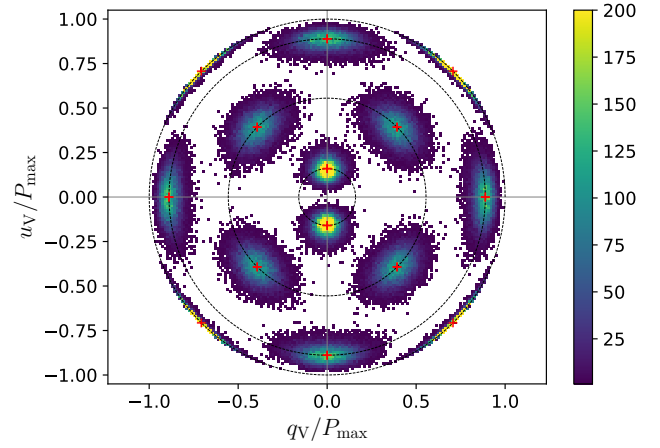

Fig. B.1: Examples of scatter produced in the (q_V, u_V) plane by the addition of a stochastic component to a regular magnetic field for different inclination and position angle of the latter and for $A_{turb} = 0.1$. Fourteen configurations are shown, as explained in the text. The red crosses indicate the (q_V, u_V) obtained for $A_{turb} = 0$ (regular field only). The scatter around each cross shows a 2D histogram of 10,000 realizations of (q_V, u_V) values obtained with the addition of the stochastic component. We note that the (q_V, u_V) are normalized by P_{max}.

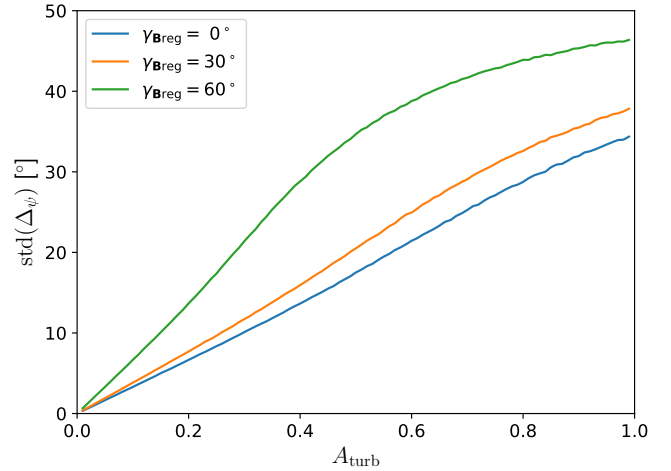Fig. B.2: Scatter in EVPA as a function of A_{turb}. For each value of A_{turb}, 10^5 realizations of $\mathbf{B}_{tot}$ are computed. The Stokes parameters are computed and polarization position angles determined and compared to the polarization vector corresponding to the regular component only. The standard deviation of the polarization angle difference is then computed and plotted against A_{turb}. We carry the analysis for $\gamma_{\mathbf{B}_{reg}} = 0°$, $30°$ and $60°$.

be self-consistently described by a bivariate normal distribution centered on the mean polarization and with a covariance matrix potentially having a nonzero off-diagonal element.

A scatter in the (q_V, u_V) plane corresponds to scatter in the degree of polarization and in the EVPA. The scatter in EVPA is thought to inform on the degree of turbulence in clouds and to be related to the amplitude of the magnetic field (e.g., Skalidis & Tassis 2021), at least when the magnetic field lies mostly in the POS. We explore the dependence of the scatter
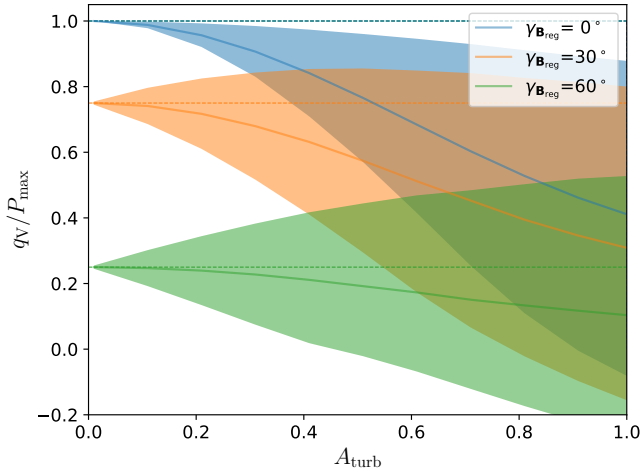
Fig. B.3: Distributions of q_V/P_{max} corresponding to various inclination angle and amplitude of the intrinsic scatter. We carry the analysis out for $\gamma_{B_{reg}} = 0°$, $30°$, and $60°$ and $A_{turb} \in [0, 1]$. For each value pair, 10^5 realizations of $\mathbf{B}_{tot}$ are computed and q_V is evaluated. The continuous lines show the mean of the distribution for $\gamma_{B_{reg}} = 0°$, $30°$, and $60°$ in blue, orange and green, respectively. The shaded areas indicate the range span between percentiles 16 and 84 of the distribution. The dashed lines indicate the q_V/P_{max} values obtained when no stochastic term is added in the magnetic field.

in EVPA as a function of the parameter A_{turb}. Because of projection effects, we carry this analysis for three values of $\gamma_{B_{reg}}$ ($0°$, $30°$, $60°$). Our results are reported in Fig. B.2. The EVPA scatter increases linearly at low values of A_{turb}, then the increase slows down until the EVPA scatter reaches saturation. The larger the angle of inclination, the greater the rate of increase and thus, breakup and saturation occur at lower values of A_{turb}. Given that there is currently a lack of starlight polarization data for diffuse sightlines, there is not much observational constraints on the scatter in EVPA and, hence, on the value that A_{turb} may take. In small angular regions toward denser clouds or denser sightlines, authors report values of the EVPA scatter to range from few degrees up to about fifteen degrees (Soler et al. 2016; Planck Collaboration Int. XXXV 2016; Panopoulou et al. 2016; Skalidis et al. 2022), also consistent with numerical simulations of the ISM in sub- and trans-Alvénic regimes (e.g., Skalidis et al. 2021). We assume that those values are representative to our case. Therefore, we choose values of A_{turb} in the range 0.1 to 0.3 to be representative of degree of intrinsic scatter we may find at intermediate and high Galactic latitude, in general, from our future survey.

In Fig. B.3, we show distributions of the Stokes parameter q_V (normalized to P_{max}) as a function of A_{turb} and for the same three inclination angle values studied before. We show the mean (continuous lines) and the interval between percentiles 16 and 84 (shaded areas) of the distributions. We see that depending on the inclination angle there might be a nonzero difference between the mean of the distributions and the values corresponding to the regular field only (dashed lines). This is a bias that purely results from projection effects. This bias is larger for small inclination angles (i.e., magnetic field close to the POS) than for large inclination angles and vanishes when the field lines point toward the observer. The same figure is obtained for Stokes u_V but with a position angle rotated by $45°$. This result is consistent with what
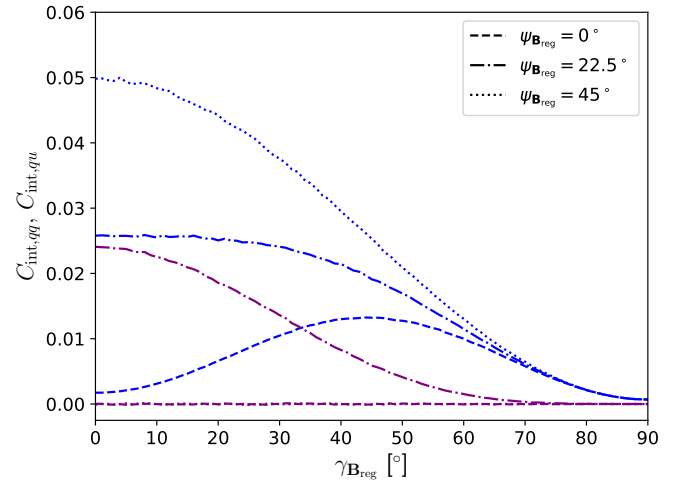

Fig. B.4: Dependence of $C_{int,qq}$ (blue) and $C_{int,qu}$ (purple) as a function of $\gamma_{B_{reg}}$ and for three values of the position angles: $\psi_{B_{reg}} = 0°$ (dashed lines), $\psi_{B_{reg}} = 22.5°$ (dashed-dotted lines), and $\psi_{B_{reg}} = 45°$ (dotted lines). A value of $A_{turb} = 0.2$ is chosen here. The elements of the covariance matrix are given in units of P_{max}^2. A value of 0.01 corresponds to a polarization uncertainty (σ_q or σ_q) of 10 per cent of P_{max}.

Hu & Lazarian (2023) obtained from numerical simulations of incompressible MHD turbulence in molecular cloud.

In the following we explore the effects of the intrinsic scatter on the element of the covariance matrix of the (q_V, u_V) pairs that it generates. In particular, we are interested in exploring the dependence on the inclination and position angle of the regular magnetic field and on the dependence with the amplitude of the stochastic component relative to the regular one.

In Fig. B.4, we present the evolution of $C_{int,qq}$ and $C_{int,qu}$ as a function of $\gamma_{B_{reg}}$, for the three position angles $\psi_{B_{reg}} = 0°$, $22.5°$, and $45°$ and for $A_{turb} = 0.2$. We do not show $C_{int,uu}$, as it is identical to $C_{int,qq}$ but for a position angle rotated by $45°$. We see that the covariance-matrix elements are nontrivial but certainly trigonometric functions of the inclination angle. The cross term vanishes when $\psi_{B_{reg}} = 0°$ or any multiple of $45°$ (i.e., when either q_V or u_V is zero). If fluctuations in magnetic field geometry were the only sources of intrinsic scatter, a detailed analysis of the covariance-matrix elements from observational data points could potentially lead to constraints on inclination angle of the magnetic field with respect to the POS. This information is generally not accessible from dust-related polarization observables and providing a direct access to it would constitute a breakthrough in dust studies of the magnetized ISM. However, as we have already noticed, other sources of fluctuations are expected in the turbulent ISM and they directly affect the values of P_{max}, adding a certainly non-negligible contribution to $C_{int,qq}$ and $C_{int,uu}$ and reducing the relative importance of $C_{int,qu}$, thus bringing some complexity to the situation depicted in Fig. B.4. While future work will be necessary to clarify whether those sources of scatter can be disentangled, we note that other sources of complication will come from both observational noise and from the a-priori unknown value of A_{turb}.

Finally, we address the question of the significance of the intrinsic scatter and how it compares to observational uncertainties for a large range of A_{turb} values and for different inclination and position angles of the regular component of the magnetic field. For this investigation we consider a sample of 100 stars in

the background of a cloud. We fix our observational uncertainty on the polarization as $\sigma_p = \sigma_q = \sigma_u = 0.2\%$ and we fix a degree of polarization of the cloud to be $p_C = 0.6\%$ in absence of scatter. We thus change the value of P_{max} to compensate for change of $\gamma_{\mathbf{B}_{reg}}$.

For each set of $(A_{turb}, \gamma_{\mathbf{B}_{reg}}, \psi_{\mathbf{B}_{reg}})$ values we can generate simulated Stokes parameters (i) computed only from $\mathbf{B}_{reg}$, (ii) computed with $\mathbf{B}_{tot}$ for specified A_{turb} and (iii) with the addition of the observational noise. For each simulated data set we estimate the total covariance matrix of the Stokes parameters and the contributions from the observational and from the intrinsic scatter taken separately. Denoting $C_{tot,xy} = C_{obs,xy} + C_{int,xy}$ with x, y being either q or u, we can measure the different terms in our simulations and thus compare the relative amplitude of the different contributions. We perform this exercise for 10,000 simulated data sets to build distributions of the elements of the covariance matrix. The result is shown in Fig. B.5 in which, for each matrix element taken separately and for each of their contributions (noise and intrinsic scatter), we show how the distributions change as a function of A_{turb} by reporting the median and the interval between 16 and 84 percentiles as shaded areas.

The relative importance of the contribution from the intrinsic scatter with respect to the contribution from the noise to the total scatter in the data points is well demonstrated by the different panels of Fig. B.5. Depending on the exact setup of the ISM and amplitude of the stochastic component in the magnetic field, the data scatter can be dominated either by the intrinsic scatter or by the noise. Keeping in mind that this also depends on the S/N level which we have fixed to ≈ 3 for individual measurements, we see that the contribution from the intrinsic scatter exceeds the noise contribution in at least one of the polarization channels when $A_{turb} \gtrsim 0.3$ and that this threshold reduces when $\gamma_{\mathbf{B}_{reg}}$ increases. When either or both the amplitude of the cloud polarization gets lower or the overall uncertainty on individual star measurements gets larger, this threshold goes to lower A_{turb} values (not shown in the figure).
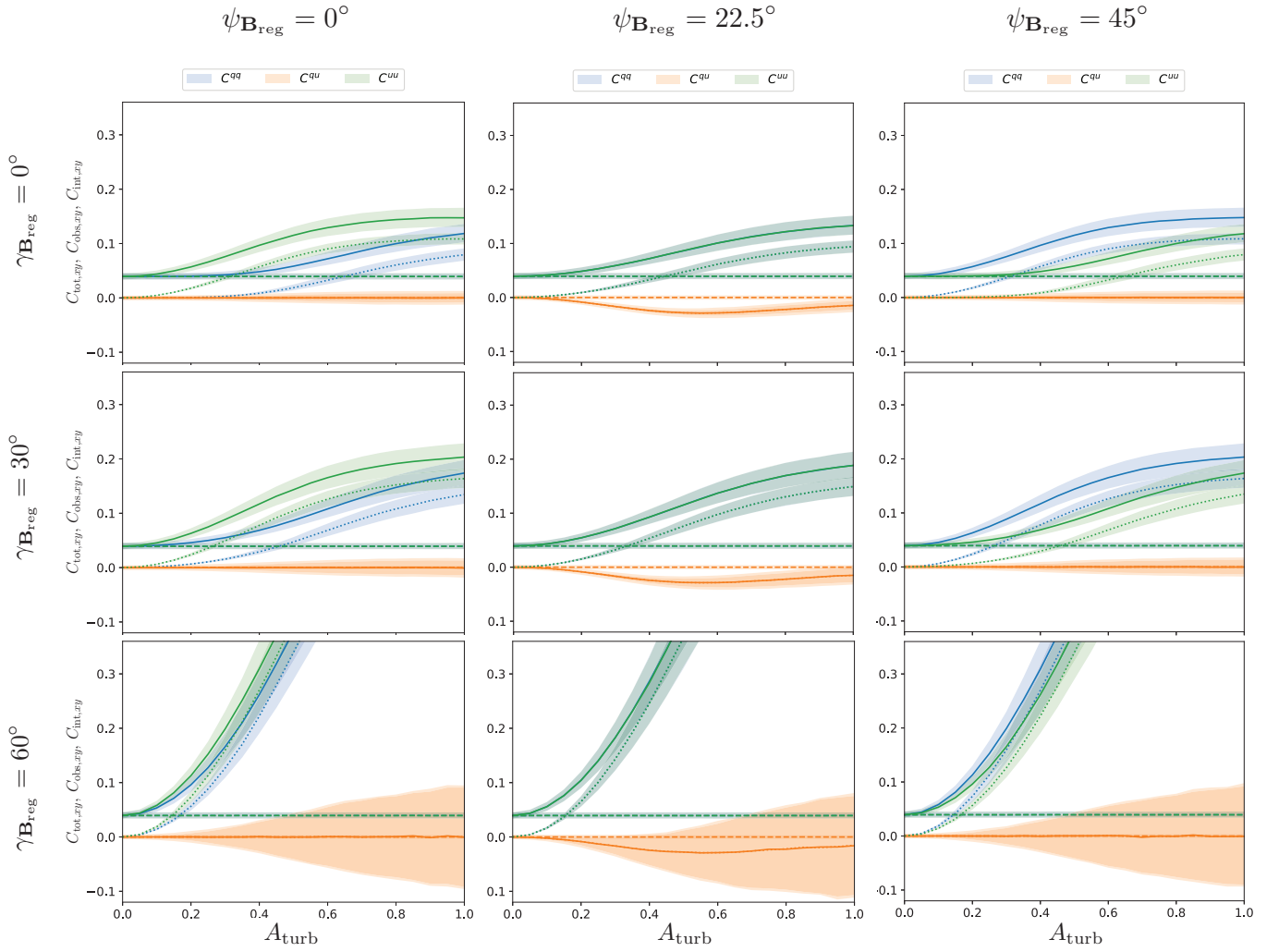
Fig. B.5: Comparison of the contributions from the observational noise and from the intrinsic scatter to the different elements of the covariance matrix as a function of the amplitude of the stochastic component and for different configurations of $\mathbf{B}_{\mathrm{reg}}$. We show the total elements of the covariance matrix (solid lines) along with the contribution from the noise (dashed lines) and from the intrinsic scatter (dotted lines). Blue, green and orange correspond to $C_{\mathrm{int},qq}$, $C_{\mathrm{int},uu}$ and $C_{\mathrm{int},qu}$ respectively. The lines and shaded areas show the median and 16 and 84 percentile of the values obtained while repeating the analysis for 10,000 simulations in which both the noise and the intrinsic scatter vary. In this simulation we have set $p_{\mathrm{reg}} = 0.6\%$ and $\sigma_p = \sigma_q = \sigma_u = 0.2\%$ corresponding roughly to a S/N of 3 for the individual star polarization measurement.