Dissertation
submitted to the
Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Richard James Hanson, M.Sc.
born in Huddersfield, United Kingdom

Oral examination: 19.05.2017

Mapping 3D extinction and structures in the Milky Way

Referees: PD. Dr. Coryn Bailer-Jones
Prof. Dr. Hans-Walter Rix

Zusammenfassung

In dieser Arbeit präsentieren wir eine Bayes'sche Methode um gleichzeitig die interstellare Extinktion, die effektive stellare Temperatur und den Entfernungsmodul zu einzelnen Sternen zu messen. Wir erstellen dreidimensionale Extinktionskarten in Richtung der galaktischen Pole, indem wir photometrische Daten von der Sloan Digital Sky Survey und UKIRT Infrared Deep Sky Survey kombinieren, sowie Karten in der galaktischen Ebene mit Daten von Pan-STARRS1 und der Spitzer GLIMPSE Durchmusterung.

Durch das Zusammenwirken von optischer und nah-infrarot Photometrie definieren wir ein Model, welches in der Lage ist, die Farbänderung aufgrund physikalischer Eigenschaften der Sterne und des interstellaren Mediums effektiv nachzubilden. Wir berechnen die Wahrscheinlichkeitsdichten der astrophysikalischen Parameter für Millionen Sterne einzeln und bestimmen anschließend gewichtete Entfernungsprofile für die Extinktion A_0 und den Extinktionsparameter R_0.

Weiterhin führen wir eine neuartige, nicht-parametrische Methode ein, die als gaußscher Prozess in sich selbst folgerichtig die dreidimensionale Staubdichte in der Milchstraße bestimmen kann. Wir verwenden individuelle Extinktions- und Entfernungsmessungen, sowie einfache Annahmen über die Entferungskorrelation der Staubdichte, um die wahrscheinlichste Dichte an einem beliebigen Punkt zu berechnen, sogar dann, wenn sich dieser Punkt nicht auf einer beobachteten Sichtlinie befindet. Wir demonstrieren anhand simulierter Daten, dass die Methode zuverlässig bekannte Strukturen rekonstruieren kann.

Abstract

We present a parametric Bayesian method to simultaneously infer interstellar extinction, stellar effective temperature and distance modulus to stars. We create three dimensional maps of extinction towards the Galactic poles using multiband photometry from Sloan Digital Sky Survey and UKIRT Infrared Deep Sky Survey, and maps in the Galactic plane using data from Pan-STARRS1 and Spitzer GLIMPSE surveys.

Using optical and near-infrared photometry we train a forward model to emulate the colour change due to properties of stars and the interstellar medium. We predict the probability density function of astrophysical parameters for millions of stars individually and then construct weighted distance profiles in extinction A_0 and extinction parameter R_0.

Furthermore we present a non-parametric model to self-consistently predict the three-dimensional dust density in the Milky Way using a Gaussian process. Using individual extinction and distance measurements to stars and basic assumptions about the spatial correlation of the dust density we infer the most probable density at any point, even if no observations are present along that line of sight. We demonstrate the method's ability to reliably reconstruct known dust structures with mock data.

Acknowledgements

A man of few words, I feel like I have already said enough. I would, however, like to thank my collaborators, colleagues, friends and family, who have, one way or another, contributed to this work and my professional development. It goes without saying that many of them were part of this journey, but there are two people I must thank in particular. The first is Coryn for giving me this opportunity and keeping up the friendly pressure to finish. The second is Jana for believing I would *actually* finish at some point and pushing me to finally get my act together. *Thank you!*

These projects were funded by the Sonderforschungsbereich SFB881 'The Milky Way System' (subproject B5) of the German Research Foundation (DFG). I was a member of the International Max-Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg (IMPRS-HD) and the Heidelberg Graduate School of Fundamental Physics (HGSFP).

This work makes use of the ninth data release of UKIDSS. The UKIDSS project is defined in Lawrence et al. (2007). UKIDSS uses the UKIRT Wide Field Camera (WFCAM; Casali, 2007).

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is www.sdss3.org.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck In-

stitute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

The Pan-STARRS1 Surveys (PS1) have been made possible through contributions of the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation under grant No. AST-1238877, the University of Maryland, and Eotvos Lorand University (ELTE).

Contents

1

Introduction

1.1 Dust

The distribution of dust varies strongly throughout our Galaxy, the Milky Way. However, an accurate picture of the three dimensional distribution is still not known. This has significant implications, considering dust attenuates and reddens light from Galactic and extra-galactic sources. This restricts the precision of many observations, in particular within the Galaxy. Considering the Milky Way is the only laboratory we have to measure many individual stars and test theories on, it is worthwhile understanding and measuring the dust distribution within in.

The first systematic insight into this resulted from the heliocentric star map of Kapteyn (1922), who realised that the star count dropped off in all directions from the sun. A few years later, Trumpler (1930) was able to demonstrate that an interstellar mechanism was responsible for this drop-off.

In the Milky Way approximately 1% of the mass is in solid form – dust (e.g. Draine, 2003). Despite its small mass fraction, dust plays an outsized role in observations. It absorbes nearly 50% of all stellar light in the Universe, processing it and re-emitting it at longer wavelengths (e.g. Kennicutt & Evans, 2012). Different dust molecules have a wide range of energy levels to interact with light over a range of frequencies. Perhaps it is therefore unsurprising that up to 30% of light emitted by the Galaxy was last emitted by dust grains, rather than from the original source (Li & Greenberg, 2003).

Not only does dust influence how we perceive the Universe, interstellar dust grains are the key ingredient in the formation of molecules. Their surfaces are catalysts for molecular reactions. The most abundant element, hydrogen, primar-

ily forms as molecules on the surface of dust grains. The other core constituents of the interstellar medium (ISM), traces of heavier elements, are present in multiple temperature and density phases. They are produced in outflows of evolved, high mass-loss stars that shed their atmospheres, and supernovae. This hot gas cools and condenses into particles that enrich the ISM (e.g. Speck et al., 2009; Kochanek, 2011).

Furthermore, dust is a primary source of cooling in hot gas clouds and enables these to collapse and initiate star formation in the first place. At the same time, the diffuse ISM is heated by dust absorbing energetic photons.

From observations we have learnt that dust forms as the metal-enriched stellar ejecta cool. Dust is destroyed on timescales of roughly $5 \cdot 10^8$ yr (Seab, 1987), for example through collisions with high velocity ions in dense environments, like those found in supernovae shocks. At the same time, timescales of stellar ejecta are much longer, at around $2.5 \cdot 10^9$ yr (Tielens et al., 1994). Dust formation, therefore, must be dominated by other channels (e.g. Tielens et al., 1994; Draine, 2003).

1.2 Interstellar Extinction

The combined effect of absorption and scattering of light is called extinction. Following Cardelli, Clayton & Mathis (1988), extinction in a narrow band at wavelength λ can be parameterized as

$$A_\lambda = A_0(a_\lambda + b_\lambda/R_0) \,, \tag{1.1}$$

where A_0 is the extinction and R_0 is the total-to-selective extinction ratio, often referred to as R_V. The two parameters, a_λ and b_λ are fixed polynomials that are fit to measurements in Johnson photometry. We use A_0, the monochromatic extinction at 5495 Å (following Cardelli, Clayton & Mathis, 1989), to characterise extinction as a property of the ISM. Other parameterizations of extinction that tie its value to a particular broad photometric band (such as A_V), implicitly are functions of the spectral energy distribution (SED) of a star and thereby depend strongly on the intrinsic physical properties of the star itself, rather than solely on the characteristics of the ISM.

Extinction is calculated as

$$A_0 = -2.5 \log \left(\frac{\int F_\lambda h_{\lambda,0} 10^{-0.4 A_\lambda} \mathrm{d}\lambda}{\int F_\lambda h_{\lambda,0} \mathrm{d}\lambda} \right) \,, \tag{1.2}$$

by integrating over an SED $F_\lambda \equiv F_\lambda(T_{\rm eff})$ with fixed stellar effective temperature $T_{\rm eff}$, whilst using the pass band function $h_{\lambda,0}$ of the filter over all wavelengths.

Most extinction laws assume that reddening over each band corresponds to reddening at the band's effective wavelength. However, due to the wavelength-dependent nature of extinction the actual effective reddening wavelength is shorter than that of the photometric filter. As a solution to this problem, Fitzpatrick (1999) uses the effective reddening wavelength to fit splines to extinction values. Through this approach, these extinction laws are more accurate for narrow-band photometry, as well as spectroscopic data.

The extinction parameter R_0 is an indicator of dust grain size distribution and composition. The parameter typically has values in the range $2.1 - 5.5$ (Cardelli, Clayton & Mathis, 1988). Higher R_0 tend to be associated with dense regions of dust (Roman-Zuniga et al., 2007), whereas lower values prevail in the diffuse interstellar medium.

The physical properties of dust are typically inferred from extinction curves, which are constructed empirically by comparing obscured and unobscured stellar sources at various wavelengths. They describe the relationship between total extinction and reddening and are often fit by polynomials between colour excesses in different photometric bands. Different regions in the curves characterise different dust components. A set of example curves for different values of $R_{\rm V}$ are shown in Figure 1.1. In the wavelength regime from $5 - 20$ μm dust emission is dominated by the molecular bands of Polycyclic Aromatic Hydrocarbons (PAH, compounds containing only carbon and hydrogen). At longer wavelengths the main effects come from thermal continuum emission from the main dust population. Here, smaller dust grains that are heated by radiation fields in star forming regions emit out to approximately 60 μm. Larger dust grains emit at even longer wavelengths (Draine, 2003).

Variations in the ultra-violet and optical regimes of extinction laws are well studied (e.g. Massa & Savage, 1989; Cardelli, Clayton & Mathis, 1989; Mathis, 1990). More recently the infrared regime has become accessible through near- and mid-infrared surveys such as Galactic Legacy Infrared Mid-Plane Survey Extraordinaire (GLIMPSE, Benjamin et al., 2003; Churchwell et al., 2009) and UKIRT Infrared Deep Sky Survey (UKIDSS, Lawrence et al., 2007). Despite a family of parameterizations being valid, typically the shape of extinction curves in optical and near-infrared regimes depends only on R_0, with minor variations for different wavelengths (Rieke & Lebofsky, 1985; O'Donnell, 1994; Fitzpatrick, 1999). In the infrared regime, the shape depends less on R_0, but for ultra-violet (and shorter) wavelengths the dependency is strongest. Using multiband photometry to measure extinction is therefore

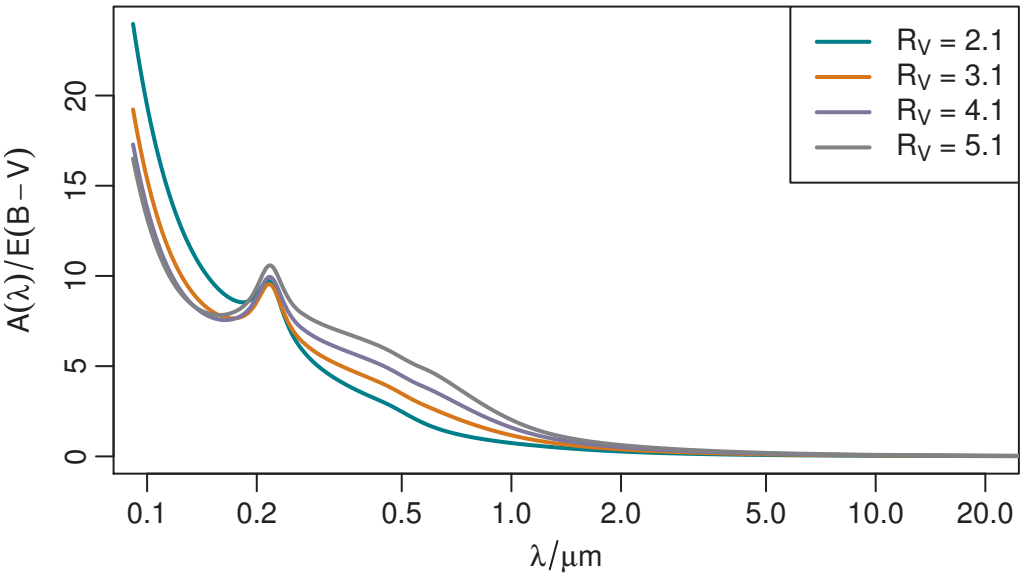

Figure 1.1: Extinction curves for R_V values of 2.1, 3.1, 4.1 and 5.1. using the parameterization of Fitzpatrick (1999). Extinction at a wavelength is normalised to the corresponding $B - V$ reddening.

a good approach to characterising the dust properties in the ISM.

1.3 Extinction Models

There are two main categories of large-scale extinction models, or dust maps. The first attempts to map dust in two dimensions, effectively characterising line of sight extinction. The second, more recent, group focusses on improving the three dimensional understanding of the dust distribution in our Galaxy.

The first two dimensional dust maps used the correlation between the dust column density and the distribution of neutral hydrogen (Burstein & Heiles, 1978). The most widely used, large-scale dust study is that of Schlegel, Finkbeiner & Davis (1998, hereafter SFD), who created a two dimensional map of total line of sight dust column density in the Milky Way by analysing dust emission at 100 and 240 μm. Assuming a standard reddening law and using calibrated dust temperatures, SFD were able to produce maps that were at least twice as accurate, and more detailed, as earlier ones.

However, the SFD maps are known to overestimate reddening by a factor of up to 1.5 in regions where V-band extinction exceeds 0.5 mag, and underestimate it in regions with steep extinction gradients (e.g. Arce & Goodman, 1999; Schlafly & Finkbeiner, 2011).

The main shortcoming of these kind of maps is the fact that they are only two di-

mensional. Sources inside the Galaxy lie within the dust and are likely to be affected by only a portion of the total dust column. To overcome this, several smaller-scale and then large scale three dimensional dust maps were created.

Hakkila et al. (1997) created an all-sky empirical extinction model by combining results from several other surveys (Fitzgerald, 1968; Neckel, Klare & Sarcander, 1980; Berdnikov & Pavlovskaya, 1991; Arenou, Grenon & Gómez, 1992). Using an extinction law parameterization from Cardelli, Clayton & Mathis (1989) with fixed $R_V = 3.1$ with intrinsic and observed B- and V-band magnitudes, they were able to reconstruct line of sight extinctions to $78,000$ individual stars. Due to the assumption of a fixed R_V these maps had relatively large measurement errors, as well as limited distance and angular resolution. Although many studies find $R_V = 3.1$ to be a common value for many sightlines through the diffuse interstellar medium, Schlafly et al. (2010), Hanson & Bailer-Jones (2014) and others have found a broad range of possible values across the sky.

In the last decade or so, significant advancements on these early maps have been made. Marshall et al. (2006) built an extinction map using data from Two Micron All-Sky Survey (2MASS) for Galactic latitudes below 10 degrees and longitudes lower than 100 degrees towards the Galactic centre. The maps have a high resolution component towards the Galactic plane. The method involves comparing observed $J - K_S$ colours of K and M giants to predicted intrinsic colours from stellar population models (Robin et al., 2003) to determine extinction in the inner Galaxy. The advantage of this approach is that infrared light is less sensitive to variations in R_V, thereby being more robust than similar approaches in optical bands. They are able to detect traces of spiral arms, dust in the Galactic bar as well as local features.

Sale et al. (2009) developed an algorithm to determine extinction from A to early K-type stars by comparing intrinsic $r - i$ band photometry (with lower R_V dependence) from INT/WFC Photometric Hα Survey (IPHAS, Drew et al., 2005) and achieve high resolution maps at low latitudes.

Schlafly et al. (2010) and Schlafly & Finkbeiner (2011) use photometric and spectroscopic data, respectively, from Sloan Digital Sky Survey (SDSS, York et al., 2000; Abazajian et al., 2009) to model extinction at higher latitudes using the colours of main-sequence turn-off stars. They test reddening from SFD and provide large-scale two dimensional map of R_V in the Galaxy.

A Bayesian approach is introduced by Bailer-Jones (2011) using the Hertzsprung-Russell diagram to place a prior on available stellar astrophysical parameter space and simultaneously infer extinction, effective temperature and distances to individual stars using broadband photometry and Hipparcos parallaxes.

Berry et al. (2012) use SDSS and 2MASS photometry and compare observations to the spectral energy distribution from stellar templates by applying a χ^2 fit to the data. This way they not only derive line of sight extinction estimates to millions of stars towards the Galactic poles, they also infer R_V to reasonable precision.

Gonzalez et al. (2011, 2012, 2013) compare red clump stars to reference measurements in Baade's window to produce high resolution extinction maps of the Galactic bulge that are insensitive to differential extinction.

Chen et al. (2013) and Schultheis et al. (2014) infer extinction in the Galactic bulge with high resolution by comparing colour-magnitude diagrams to a Galaxy model. Chen et al. (2014) build a map towards the Galactic anticentre using Xuyi Schmidt Telescope Photometric Survey of the Galactic Anticentre (XSTPS-GAC), Wide-field Infrared Survey Explorer (WISE) and 2MASS data, by fitting SEDs to photometric observations.

Sale et al. (2014) expands on a hierarchical Bayesian system developed in Sale (2012) and applies it to IPHAS data to map extinction and differential extinction in the Northern Galactic plane with fine distance resolution.

Green et al. (2014, 2015) and Schlafly et al. (2014a) combine Galactic priors to obtain a fully probabilistic three dimensional extinction map for most of the Galaxy above a declination of -30° using Panoramic Survey Telescope & Rapid Response System (PanSTARRS 1 or PS1) photometry. They achieve a maximum distance resolution of roughly 25% and find good agreement with the Planck-based reddening map.

In Hanson & Bailer-Jones (2014) and Hanson et al. (2016) we expand on the work by Bailer-Jones (2011), employing a Bayesian approach to simultaneously infer several stellar parameters for stars towards the Galactic poles and in the Galactic plane, respectively.

A common shortcoming of most of the aforementioned maps is the fact that angular correlations are not accounted for, meaning neighbouring lines of sight are solved for independently. Extinction of stars that are close in space should be affected by the same dust structures, whereas stars with large angular separation are less correlated. Significant progress on this topic has started to be made in recent years.

Lallement et al. (2014) create 3D maps of the local ISM by combining colour excess data and distance information of individual sightlines and then applying an inversion method (Vergely et al., 2010), which is based on a regularised Bayesian approach, thereby being able to map nearby dense clouds with a spatial resolution between 10 pc and 100 pc, depending on the distance to the clouds.

Sale & Magorrian (2014) introduce a survey-agnostic method based on Gaussian random fields and a model of interstellar turbulence to address the discontinuities seen in most extinction maps. They demonstrate the value of being able to utilise the spatial resolution set naturally by the data, rather than imposing any binning.

In Rezaei Kh. et al. (2017), we develop a non-parametric model for inferring dust density. Using a Gaussian Process model and assuming simple spatial correlations, a fully self-consistent map of dust can be produced, even predicting the dust density and line of sight extinction in regions not directly observed by the data.

1.4 Thesis Outline

This thesis is split into two conceptually distinct parts. In Part 1, I present work based on Hanson & Bailer-Jones (2014) and Hanson et al. (2016), where we use multiband photometry in conjunction with a Bayesian probability model trained on synthetic spectra to simultaneously predict stellar effective temperatures, distance moduli, extinction and relative extinction to millions of individual stars towards the Galactic poles and in the Galactic plane.

In Part 2, I present a non-parametric model to infer the three dimensional distribution of dust in the Galaxy using line of sight extinction measurements to individual stars. A Gaussian process is used to compute full dust density probability distributions for mock data examples. Further results from this method, along with applications of current extinction datasets, are published in Rezaei Kh. et al. (2017).

Part 1

Photometric Extinction Mapping

2

Extinction Model

This chapter covers the theoretical background and algorithmic approaches common to chapters 3 and 4. These, in turn, contain the astrophysical results published in Hanson & Bailer-Jones (2014) and Hanson et al. (2016), respectively.

The aim is to determine the full probability density function (PDF) over a set of astrophysical parameters for a star given its photometric colours $\boldsymbol{p}$. We infer interstellar extinction A_0, effective temperature T_{eff} and distance modulus μ. In an additional step we extend the method to also infer the extinction parameter R_0 to fully understand the three dimensional properties of dust.

2.1 Theory

We use a Bayesian approach to determine the aforementioned probability density function. This allows us to include prior knowledge on the components of our model. We build on the method introduced by Bailer-Jones (2011) by generalising it to include distance modulus.

In the first stage, we want to infer monochromatic extinction, A_0, the effective temperature, T_{eff}, and the distance modulus, μ. In the second stage, we expand the method to also estimate the relative extinction R_0. The measurement data are the set of stellar colours, the vector $\boldsymbol{p}$, and the apparent magnitude of the star in one band, m. By separating out the spectral and distance information in this way, instead of using the individual magnitudes, we can construct a cleaner posterior distribution function.

The Hertzsprung–Russell diagram, H (or HRD), introduces our prior knowledge of stellar structure and evolution. It is a two-dimensional probability distribution over

M, the absolute magnitude in a band, and $T_{\rm eff}$. The specific implementations of H will be detailed sections 3.1 and 4.1.

Using Bayes' theorem, the posterior distribution is proportional to the product of a likelihood and a prior

$$\underbrace{P(A_0, T_{\rm eff}, \Delta|\boldsymbol{p}, m, H)}_{\text{posterior}} = \frac{1}{Z}\underbrace{P(\boldsymbol{p}, m|A_0, T_{\rm eff}, \Delta, H)}_{\text{likelihood}}\underbrace{P(A_0, T_{\rm eff}, \Delta|H)}_{\text{prior(s)}}, \tag{2.1}$$

where the first term on the right-hand side is the normalisation factor $Z^{-1} = P(\boldsymbol{p}, m)^{-1}$. This is a constant for every star, and therefore does not need to be explicitly computed. The second term is the likelihood function, the third is the prior.

Note that we use a proxy variable, $\Delta = m_{\rm r} - M_{\rm r}$, instead of distance modulus in our formulation. We make this choice, as it prevents us from having to explicitly model the dependency of distance modulus on extinction. However, when reporting actual parameters for each star we compute distance modulus as $\mu = \Delta - A_{\rm r} = m_{\rm r} - M_{\rm r} - A_{\rm r}$. The r-band extinction is computed as a function of A_0 and $T_{\rm eff}$. In principle we could implement this step using any band, but for practical reasons we use r-band magnitudes and extinction whenever they are used, such as for the HRD prior H.

Considering that $\boldsymbol{p}$ is independent of m, Δ, and H once conditioned on A_0 and $T_{\rm eff}$, we can write the likelihood function as

$$P(\boldsymbol{p}, m|A_0, T_{\rm eff}, \Delta, H) = P(\boldsymbol{p}|A_0, T_{\rm eff})P(m|A_0, T_{\rm eff}, \Delta, H)\ . \tag{2.2}$$

The second term can be written as a marginalisation over M

$$\begin{aligned} P(m|A_0, T_{\rm eff}, \Delta, H) &= \int_M P(m|M, A_0, T_{\rm eff}, \Delta, H)P(M|A_0, T_{\rm eff}, \Delta, H)\,{\rm d}M \\ &= \int_M P(m|M, \Delta)\frac{P(M, T_{\rm eff}|H)}{P(T_{\rm eff}|H)}\,{\rm d}M \end{aligned} \tag{2.3}$$

where, due to conditional independence, we can remove A_0, $T_{\rm eff}$ and H from the first term under the integral. This is because $m = \Delta + M$, by definition. Note that because m and Δ are measured – and therefore noisy – quantities, $P(m|M, \Delta)$ is not a delta function. We also remove A_0 and Δ from the second term, because given the HRD and $T_{\rm eff}$, the distribution over M is fully defined. Note that the right-hand-side no longer has any dependence on A_0. The apparent magnitude, m, is independent of A_0 because H and $T_{\rm eff}$ specify a distribution over M, which together with Δ specifies a distribution over m.

Assuming the prior is separable such that we can write

$$P(A_0, T_{\rm eff}, \Delta|H) = P(A_0, \Delta)P(T_{\rm eff}|H) \ , \tag{2.4}$$

then substituting equations 2.2, 2.3 and 2.4 into 2.1 yields

$$\begin{aligned} &P(A_0, T_{\rm eff}, \Delta|\boldsymbol{p}, m, H) = \\ &\frac{1}{Z}P(\boldsymbol{p}|A_0, T_{\rm eff})P(A_0, \Delta)\int_M P(m|M, \Delta)P(M, T_{\rm eff}|H)\,{\rm d}M \ . \end{aligned} \tag{2.5}$$

This expression is the product of three terms. The first is the probability of measuring the colours given the astrophysical parameters A_0 and $T_{\rm eff}$. The second is the prior over extinction and the distance modulus proxy, Δ. The third is an integral over the unknown absolute magnitude, constrained by the HRD and the relationship between m, M, and Δ.

To obtain the corresponding relation including R_0, we simply replace A_0 with (A_0, R_0):

$$\begin{aligned} &P(A_0, R_0, T_{\rm eff}, \Delta|\boldsymbol{p}, m, H) = \\ &\frac{1}{Z'}P(\boldsymbol{p}|A_0, R_0, T_{\rm eff})P(A_0, R_0, \Delta)\int_M P(m|M, \Delta)P(M, T_{\rm eff}|H)\,{\rm d}M \ , \end{aligned} \tag{2.6}$$

knowing that the previous arguments of conditional independence also hold for R_0.

2.2 Modelling

With the posterior fully defined for our purposes we construct a model that utilises it to infer the astrophysical parameters for a star. For this we build a forward model that predicts the star's colours given a set of parameters. It is calculated by fitting a thin plate spline as a function of A_0 and $T_{\rm eff}$ (and where relevant, R_0) to each colour in $\boldsymbol{p}$, such that $\boldsymbol{f}(A_0, T_{\rm eff}, \{R_0\}) = \boldsymbol{p}'$.

Likelihood

Given the forward model predictions of the colours $\boldsymbol{f}$ and assuming normally distributed errors on the measurements, the likelihood (equation 2.2) of the colours $\boldsymbol{p}$ is

$$P(\boldsymbol{p}|A_0, T_{\rm eff}) = k \cdot \exp\left(-0.5\,(\boldsymbol{p}-\boldsymbol{f})^T \boldsymbol{C}^{-1}(\boldsymbol{p}-\boldsymbol{f})\right)\ , \tag{2.7}$$

where $k = (2\pi)^{-n/2}|\boldsymbol{C}|^{-1/2}$ is the normalisation factor, which depends on the length n of the colour vector $\boldsymbol{p}$. The covariance matrix $\boldsymbol{C}$ for each star is determined from its magnitude measurements and takes into account that errors on consecutive colours are not independent.

Generally, for two random variables A and B, the variance Var() of their sum is

$$\mathrm{Var}(A+B) = \mathrm{Var}(A) + \mathrm{Var}(B) + 2\mathrm{Cov}(A,B)\,, \tag{2.8}$$

where Cov() is the covariance. Furthermore, for the random variables A, B, C and D and constants a, b, c, d the covariance of

$$\begin{aligned}&\mathrm{Cov}(aA+bB, cC+dD) = \\ &ac\mathrm{Cov}(A,C) + ad\mathrm{Cov}(A,D) + bc\mathrm{Cov}(B,C) + bd\mathrm{Cov}(B,D)\ .\end{aligned} \tag{2.9}$$

Assuming that magnitudes are measured independently (their covariance is zero), then for two colours with one common band, e.g. $p_1 = m_1 - m_2$ and $p_2 = m_2 - m_3$, the covariance of these colours is

$$\mathrm{Cov}(m_1 - m_2, m_2 - m_3) = -\mathrm{Var}(m_2). \tag{2.10}$$

For the example magnitude bands m_1, m_2, m_3, m_4, m_5, the covariance matrix for the resulting colours $\boldsymbol{p} = (m_1 - m_2, m_2 - m_3, m_3 - m_4, m_4 - m_5)$ has the following form

$$\boldsymbol{C} = \begin{pmatrix} \sigma_1^2+\sigma_2^2 & -\sigma_2^2 & 0 & 0 \\ -\sigma_2^2 & \sigma_2^2+\sigma_3^2 & -\sigma_3^2 & 0 \\ 0 & -\sigma_3^2 & \sigma_3^2+\sigma_4^2 & -\sigma_4^2 \\ 0 & 0 & -\sigma_4^2 & \sigma_4^2+\sigma_5^2 \end{pmatrix}\ , \tag{2.11}$$

where σ_i^2 is the variance in band i. If we were to use more colours, the matrix would expand analogously. Standard χ^2 approaches tend to overlook the fact that colours are not independent, and therefore ignore the off-diagonal entries of the covariance matrix.

HRD prior

We construct an HRD prior, $P(M, T_{\rm eff}|H)$ (see Equation 2.5), to characterise the relative probability of a star's parameters being located in some part of the HRD parameter space. It is well understood from stellar evolution observations and models that the HRD is inhomogeneously populated. This implies that, within certain boundaries, fixing one parameter constrains the other. This property allows us to extract valuable information and constrain the inferred parameters of our model.

Setting aside the explicit technical implementation of an HRD prior, which is covered separately in chapters 3 and 4, we must construct an equation the links apparent and absolute magnitudes and distance. If we independently estimate extinction and effective temperature from photometric colours and use the measured apparent magnitude $m_{\rm r}$, together with the HRD we can in principle determine the absolute magnitude of a star and from that estimate the distance modulus.

By introducing a noise model we determine a probability distribution from noisy measurements and the finite width of the HRD. For this, we define the random variable

$$\kappa = m_{\rm r} - M_{\rm r} - \Delta \ , \tag{2.12}$$

where Δ is the difference between the true (but unknown) apparent and absolute magnitudes, as used previously. Were the measurements to be error-free, then κ would be zero. Δ can be considered a proxy for distance modulus, which is the actual parameter we aim to infer.

For the noise model $P(\kappa|M_{\rm r})$ we select a one-dimensional Gaussian in κ, $\mathcal{N}(0, \sigma_{m_{\rm r}})$ which has an expectation value of zero and a standard deviation $\sigma_\Delta = \sigma_{m_{\rm r}}$, the uncertainty in the *r*-band magnitude measurement for the observed star. The absolute magnitude is not measured and as such has no error, Δ is sampled. For a particular measurement of $m_{\rm r}$ the probability is

$$P(\kappa|M_{\rm r}) = \mathcal{N}(m_{\rm r} - M_{\rm r} - \Delta,\, \sigma_{m_{\rm r}}) \ . \tag{2.13}$$

Figure 2.1 illustrates this situation for a sampled $X = 10$ mag, which can correspond to various combinations of absolute and apparent magnitudes, as long as $M_{\rm X}$ is in the physical range of the HRD.

As $m_{\rm r}$ is the only noisy quantity in κ, this term can also be written as $P(m_{\rm r}|M_{\rm r})$. This represents the first term of the integral in equations 2.5 and 2.6. So, for any

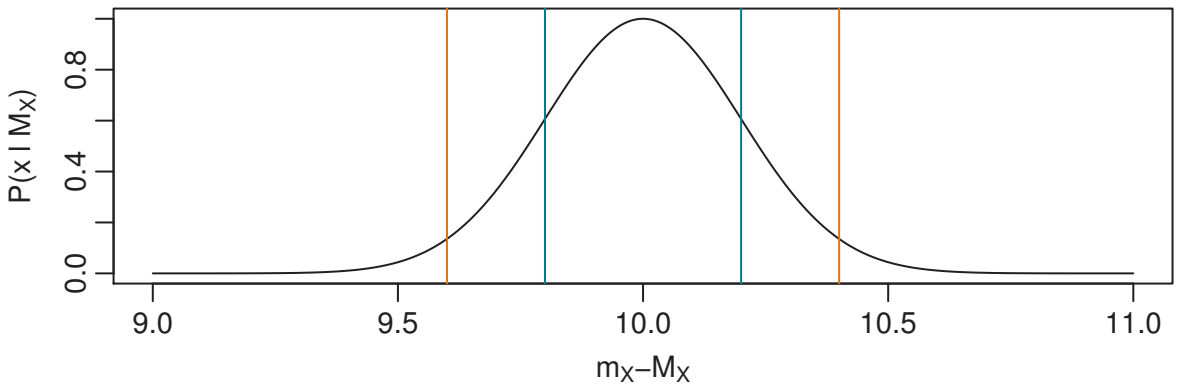

Figure 2.1: Illustration of the way the HRD prior is sampled in practice. For any sample in $\Delta_X = m_X - M_X$, a Gaussian with width corresponding to the uncertainty in the magnitude measurement (here 0.2 mag for a mean $X = 10$ mag) is used to integrate over the probability distribution of the HRD. The vertical turquoise and orange lines indicate the positions of 1 and 2 standard deviations, respectively.

measure of Δ, obeying $\kappa = 0$, the Gaussian describes the scatter around this value. The full HRD prior probability is then calculated by integrating the Gaussian for a Δ over the probability distribution of the HRD at a given $T_{\rm eff}$, as formalised in equation 2.3.

Using the HRD gives us a self-consistent means to estimate distances from photometric information. A stronger constraint on APs can be constructed if distance information is available from independent measurements.

2.3 Mapping

To summarise and visualise these results we bin stars with a fixed angular resolution in l and b. To compute the variation in extinction $A_{\rm r}$ (and R_0) along the line-of-sight at any value of distance modulus μ_j we calculate the weighted mean extinction $\langle A_{\rm r} \rangle_j$ and standard deviation Σ_j (analogously for R_0) for all stars in a single bin which have a distance modulus estimate within one magnitude of our selected position. These are

$$\begin{aligned} \langle A_{\rm r} \rangle_j &= \frac{\sum A_{{\rm r},i} w_{i,j}}{\sum w_{i,j}} \quad , \\ \Sigma_j &= \sqrt{\frac{\sum w_{i,j} (A_{{\rm r},i} - \langle A_{\rm r} \rangle_j)^2}{\frac{N-1}{N} \sum w_{i,j}}} \quad , \end{aligned} \tag{2.14}$$

where the sums are over i. The weight $w_{i,j}$ is a measure of the difference between the inferred stellar distance modulus μ_i and the cell centre distance μ_j. The confi-

dence intervals about the mode are not symmetric, so we use a split Gaussian to approximate the distribution they describe. For each star we compute the weight using the asymmetric Gaussian (or split normal distribution), parameterized by the mode and the standard deviations, σ_1 and σ_2, of each half of the Gaussian:

$$w_{i,j} = \frac{2}{\sqrt{2\pi}(\sigma_1 + \sigma_2)} \cdot \exp\left(-\frac{(\mu_i - \mu_j)^2}{2\sigma_k^2}\right) . \tag{2.15}$$

In the case when μ_i is smaller than μ_j then $\sigma_k = \sigma_1$, otherwise $\sigma_k = \sigma_2$. This is a convenient and fast substitute for summing over all the 2D PDFs we obtain from the inference. Stars with small confidence intervals are weighted more strongly than those with large ones. This procedure can be applied to any arbitrary distance modulus step μ_j. This is repeated for every angular bin to construct a full three dimensional representation of the cumulative line-of-sight extinction. Analogously we use the same procedure with the extinction parameter R_0, allowing us not only to follow the extinction variation along the line of sight, but also to look at the properties of the dust. Due to the selection process, it is in principle possible that individual stars appear in two consecutive distance bins, indicating that this measure is similar to a running (weighted) mean. For each cell we require at least 10 stars to compute the result.

We use distance modulus as the distance variable because it straightforwardly captures the uncertainty which increases with distance. For example, the relative error in distance for a distance modulus error of $\delta\mu = 1$ mag at $d = 1$ kpc is $\delta d = 0.46$ kpc, whereas at $d = 5$ kpc it increases to $\delta d = 2.3$ kpc. It is important to note that although the uncertainty in μ may be symmetric, it will not be in d.

3

Dust at the Poles

This chapter is adapted from Hanson & Bailer-Jones (2014). We use photometric data from *UKIRT Infrared Deep Sky Survey*, (UKIDSS, Lawrence et al., 2007) DR9 Plus and *Sloan Digital Sky Survey* (SDSS, Aihara et al., 2011) DR8 to infer extinction and distance modulus to stars towards the Galactic poles. We estimate the effects of unknown metallicity in the models and demonstrate the benefit of including near-infrared data.

This chapter is organised as follows. First, we introduce the data products utilised in the analysis and then detail how they are used to construct and test the forward model as described in Chapter 2. We then validate the performance of the method by measuring the accuracy of the model and precision of individual estimates of astrophysical parameters. We quantify the effects of excluding data and inferring additional APs (R_0 in particular), and present maps and profiles for the regions covered by the data. We use the SFD results as a baseline to compare the outcome to. We conclude with a summary.

3.1 Data Products

To perform the analysis described in Chapter 1, a set of photometric data is required, for which astrophysical parameters have been independently determined. We use stars selected from UKIDSS and SDSS. We use data from the UKIDSS *Large Area Survey* (LAS) which is designed to overlap with the footprint of SDSS. By using real data, we are able to use the proper photometric errors and intrinsic scatter to build the forward model, instead of depending on synthetic estimates. We perform a crossmatch to generate a catalogue of stars with photometry in nine bands. The

bands *ugriz* are from SDSS and *YJHK* are provided by UKIDSS LAS. In principle we can also use (or combine) other surveys, as our method can easily be trained on further photometric data.

Survey data

We need to simulate the effects of extinction to build an extended catalogue of stars with known APs. We therefore require a catalogue of (ideally) zero extinction stars. To build our training data we select only stars at high Galactic latitudes, $b > 70^\circ$, which avoids the Galactic plane and most of the interstellar dust. Additionally these stars are required to be dwarfs (surface gravity $\log g > 4$). This ensures that the catalogue stars only have minimal variations in $\log g$, which improves the temperature estimates due to the tight correlation with colour for fixed extinction (as in Figure 3.3). Furthermore, effective temperatures determined with the *SEGUE Stellar Parameter Pipeline* (SSPP, Lee et al., 2008; Yanny et al., 2009) must be available, with errors on these of less than 200 K.

We use a weighted average of several spectroscopic and photometric temperature estimates (called `ADOP` in the SSPP papers) for this work. When instead using the purely spectroscopic estimate, e.g. the `ANNRR` routine (Re Fiorentin et al., 2007), we see no significant differences in our inferred extinction and distance estimates. The temperature estimates do differ, with a standard deviation of about 80 K. But as we are not primarily interested in inferring temperatures, we can in principle use either of the estimates, though using the spectroscopic estimator avoids circular arguments, due to nature of how they are generated.

We require the *r*-band magnitudes to be $r < 19$, which typically results in photometric errors approaching the systematic limit of approximately 0.02 mag. This also generally forces the photometry in the other bands to be complete and of good quality. To further ensure selection of sources with high quality data, we impose the following selection criteria on the SDSS flags: we select only unresolved sources (`sdss.type` = 6) with clean photometry (`sdss.clean` = 1) and general quality flags set (`sdss.'4295229440'` = 0, `sdss.mode` = 1). Similarly, we select objects with clean photometry from UKIDSS (`ukidss.*ppErrBits` $< 65\,536$, one for each band) that are classified as stellar objects (`ukidss.mergedClass` $= -1$).

The data (SDSS and UKIDSS) are queried and cross-matched using the *WFCAM Science Archive*, requiring complete photometry in all bands. In total, we obtain roughly 3 200 stars that fulfil all the above criteria. Compared to the hundreds of thousands of stars analysed in SEGUE, we retrieve comparatively few. Aside from

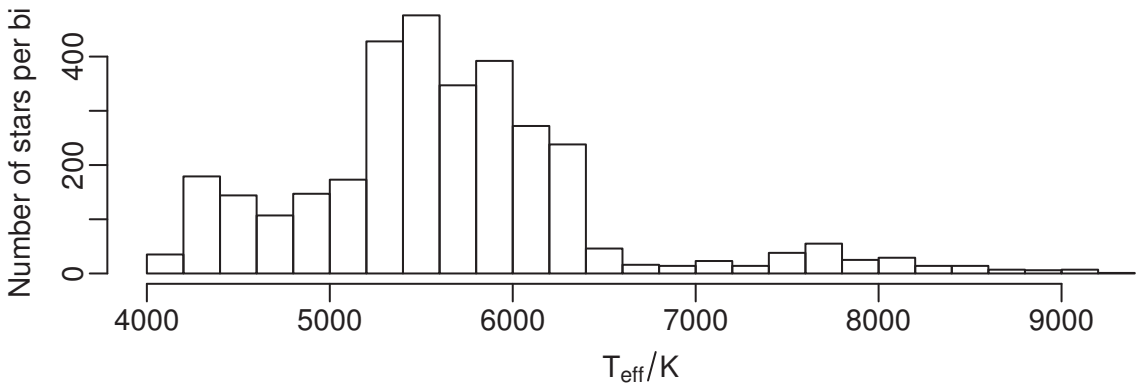

Figure 3.1: Distribution of effective temperatures of approximately 3 200 crossmatched stars. Temperatures are estimated by SSPP with mean uncertainties of 58.5 K. Of these stars, one thousand are randomly chosen to build the extended dataset.

completeness and crossmatching reasons, this is mainly influenced by the magnitude, surface gravity and temperature limits we set. In spite of this, this does not introduce systematic biases. Our selection is representative of the full SEGUE dataset (see below). The photometry is then corrected for the small estimated extinction from SFD, where the assumption holds that at high Galactic latitudes the stars are behind all the layers of dust.

Of the selected data, a randomly sampled set of one thousand stars is chosen, covering a temperature range of approximately $4000 - 9000$ K. A histogram of the temperature distribution for the full training set is shown in Figure 3.1, showing a peak around 5600 K. Only 9 % of the stars have effective temperatures above 6500 K. This is representative of the SEGUE dataset as a whole.

When later applying the method to data with unknown stellar parameters, we still create a selection with the quality flags used above, though we naturally impose no requirement on the existence of temperature estimates or the position in the Galaxy (low extinction).

HRD prior

The HRD prior is constructed using stellar isochrones from Sordo et al. (2010). The chosen stellar population comprises 100 000 stars drawn from a Salpeter initial mass function with masses ranging from $0.2 - 107$ $M_\odot$ (though we only use stars with $T_{\rm eff} \leq 9\,000$ K and the vast majority of stars have masses below 1.6 $M_\odot$). Assuming a constant star formation rate over the age of the Universe (13.7 Gyr), all stars evolve independently with solar metallicity, i.e. there is no chemical enrichment. By using the forward model, however, we are implicitly assuming that the metallicity of stars we encounter have the same mean metallicity of the SSPP stars, which are used to

fit the model. This mismatch between the metallicities of the forward model and the HRD does not strongly affect the estimation of A_0 and $T_{\rm eff}$ directly, though it does limit the ability of achieving good distance estimates. As we are not estimating metallicity, this mismatch will naturally occur for most stars, independently of the choice of HRD metallicity. An example of this limitation is given in Section 3.4.

The resulting temperatures and absolute magnitudes then are used to place the stars in the HRD plane. To use the HRD as a prior, we need to convert the data representation to a probabilistic description of the distribution. We achieve this by applying a two-dimensional binned kernel density estimate with a bandwidth of 12.5 K and 0.0625 mag in effective temperature and r-band absolute magnitude, respectively. The resulting grid has the pixel dimensions of 600^2, we limit the temperature range to $3010 - 9000$ K as to extend slightly beyond the limits of the forward model parameters. This allows for the main sequence to be fully represented down to low temperatures without an artificial cutoff. The resulting magnitude range is $M_{\rm r} \in (-4, 12)$ mag. Before normalisation, a small, but non-zero, offset is added to each point to account for the regions which nominally have zero probability, but in reality may not be completely empty. These regions include white dwarfs in the lower left part of the figure or (post) AGB stars higher up the branch to the right. Statistically, these regions will be very thinly populated and we don't expect to find many (if any) of these stars in our samples, so we don't model these in a more sophisticated manner. A representation of the final distribution is presented in Figure 3.2.

3.2 Forward Model

Having presented the data input to the model, we are now ready to build the forward model itself that will be used to compute the most likely set of colours given an input set of astrophysical parameters.

Adding artificial extinction

Using the filter response functions for all filters and an extinction law (Fitzpatrick, 1999), the change in magnitude for a given extinction A_0 in each band is computed for a range synthetic spectral energy distributions (SED). Owing to the fact that we are using broad band filters, the extinction in a band varies smoothly with temperature and we do not need to do this for every unique temperature present in our data sample. Instead we use eleven *PHOENIX* model SEDs (Hauschildt, Allard & Baron, 1999) evenly spaced in the temperature range from 4000 K to 9000 K. This

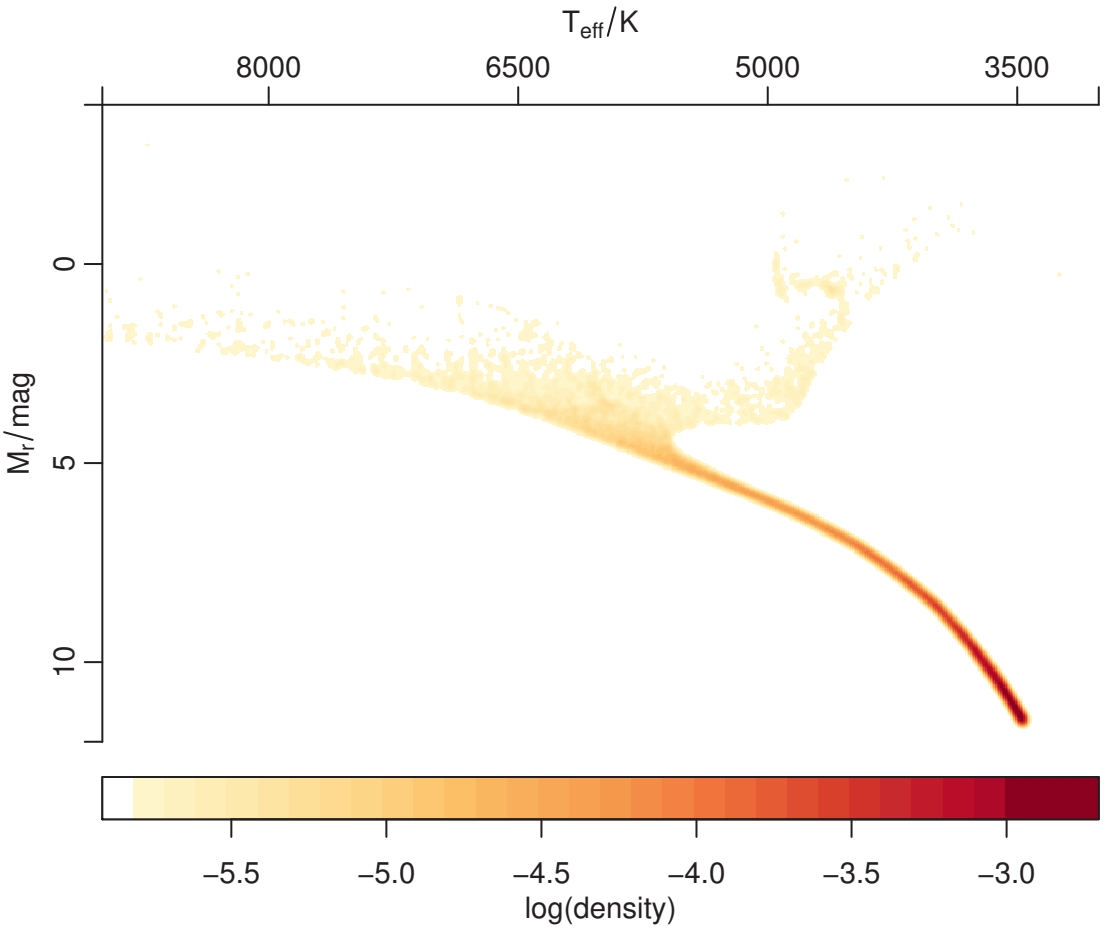

Figure 3.2: Density representation of the HRD used in this analysis. The integrated probability of the HRD is normalised to one, red is high number density and yellow is low. The colour scale shows base 10 logarithm of the density. White areas denote regions of the parameter space with initially zero probability. A small offset is added to each point before normalisation to avoid this in the actual computation. In this case the offset is approximately 10^{-3} times the maximum density.

covers the temperature range in our catalogue. For each band we fit a set of one-dimensional quadratic functions in $T_{\rm eff}$ for fixed A_0, covering the whole temperature range noted above and extinction in steps of 0.25 mag from 0 to 5 magnitudes.

Knowing how the magnitudes vary with A_0 and $T_{\rm eff}$, we expand our initial dataset of zero extinction stars by adding artificial extinction in 21 discrete steps over the range noted above for each of the 1 000 stars. The photometric errors are adapted to take into account the change in magnitudes by following the average magnitude-error relation in the data. Our expanded dataset (*training set*) now comprises 21 000 stars with known artificial extinction and effective temperatures.

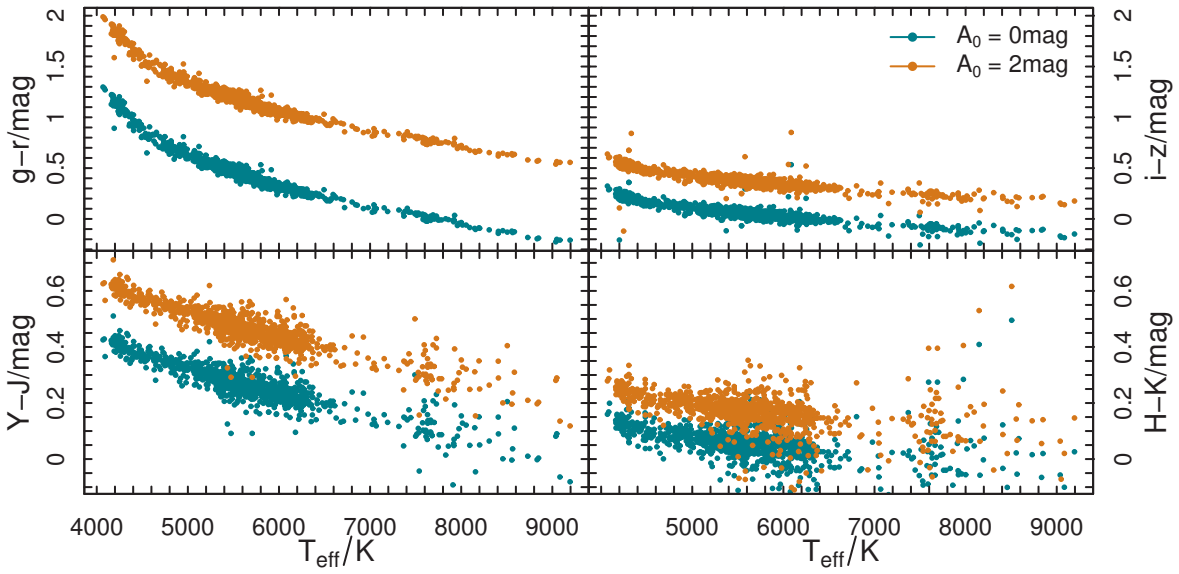

Figure 3.3: Effects of effective temperature and extinction on four selected colours. Turquoise points show a subset of the training data with no extinction as a function of effective temperature. Orange dots show the effects for $A_0 = 2$ mag. Each sample consists of 1 000 stars.

Building the forward model

In Figure 3.3 we show the effect of temperature and extinction on four different colours. We do not convert the magnitudes in both surveys to a common system (UKIDSS uses Vega magnitudes, for example), because the colours are merely affected by a constant offset with respect to that of the other survey, and we are only interested in changes in colours. As long as the training data and the final datasets are handled consistently, we need not take into account this offset.

From the training set we select a random subset of 4 000 stars to create the forward model itself, with the aim of being able to predict colours given temperature and extinction. For each of the eight reduced colours we fit a two-dimensional thin-plate spline to both APs. A good fit for all colours is achieved when giving the splines 20 degrees of freedom.

In Figure 3.4 we show the direct performance of the splines predicting colours given some input APs. For four example colours we plot the residuals of colour $(p_{X,true} - p_{X,spline})$ for the ranges of temperature and extinction present in the model as contours. This is not the final result of the full parameter inference, instead it illustrates the intrinsic inaccuracies of the model which derive from the scatter in the real data. We do not see any systematic variation, and measure a typical scatter in the order of $0.03 - 0.06$ mag for all colours, which is roughly $\sqrt{2}$ larger than for the magnitudes in a band. The forward model captures the intrinsic variance, as this is

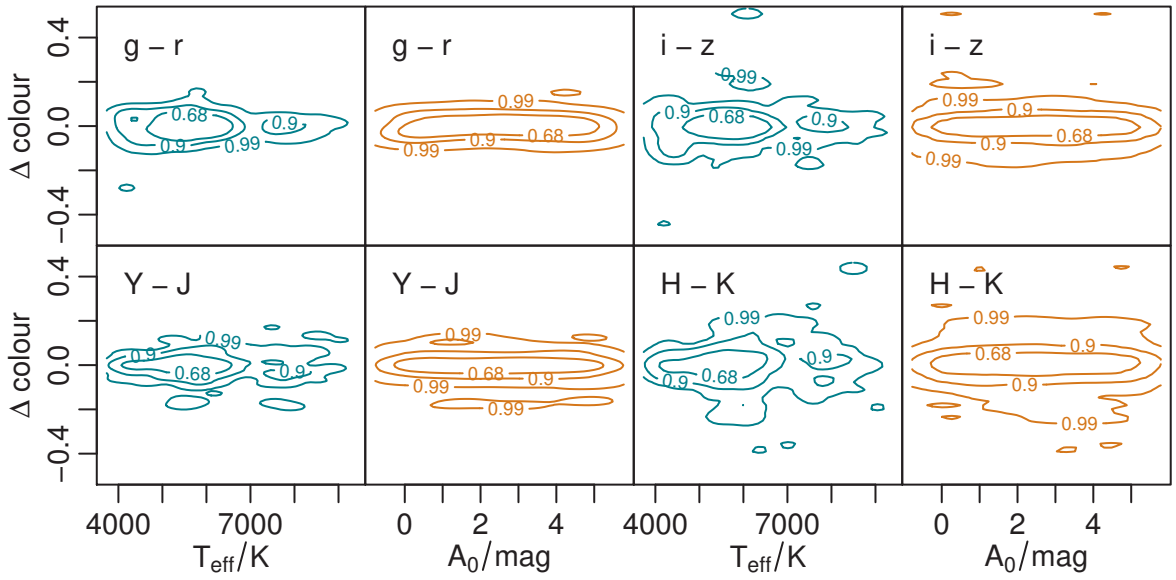

Figure 3.4: Residuals in the forward model. This shows the difference between the true colours and the colours predicted from the forward model thin plate spline as functions of true effective temperature and extinction, for four colours. The contour levels shown incorporate 68%, 90% and 99% of the data points. In the top left of each panel the corresponding colour is named, with the x-axis depicting the respective parameter ranges. All panels are scaled to show the same range on the y-axis. Typical scatter is in the order of 0.03 – 0.06 mag for all colours.

the same order of magnitude or within a small factor of the reported uncertainties of the survey data magnitudes. In Section 3.4 we look in detail at how assuming larger photometric errors affects the AP estimation.

We characterise extinction using A_0 to indicate that this is a parameter that solely depends on the physics of the interstellar medium, rather than on the type of star. When presenting results and maps later on we will use extinction in SDSS r-band to allow more convenient comparisons with other work. Note that the conversion from A_0 to A_r has a small dependence on temperature. For $R_0 = 3.1$ they roughly convert as $\langle A_r/A_0 \rangle = 0.837$. The standard deviation from this value over the parameter range $T_{\rm eff} = 4000 - 9000$ K and $A_0 = 0 - 5$ mag is only 0.005. The model values for extinction in other bands, e.g. the relative extinctions A_X/A_r are listed in Table 3.1 for three different values of R_0. Values reported are averages over the above $T_{\rm eff}$ and A_0 ranges. In particular, changes occur for the short wavelength optical bands. When varying R_0, we see that the relative extinction in the other, redder, bands changes only slightly. In Figure 3.5 these changes are illustrated for all the bands used and for two different effective temperatures.

For the APs used to build the forward model we implicitly implement priors that are uniform over the APs' training ranges, and have zero probability outside them. We will demonstrate the use of a prior on extinction based on SFD estimates later

R_0	u	g	i	z	Y	J	H	K
2.1	2.393	1.698	0.707	0.542	0.439	0.337	0.230	0.147
3.1	1.848	1.445	0.744	0.555	0.427	0.308	0.204	0.133
4.1	1.610	1.329	0.761	0.561	0.422	0.296	0.193	0.126

Table 3.1: Model values for extinction in all bands, compared to SDSS r-band extinction, A_X/A_r for three values of R_0. Reported values are averages over the full temperature range of the data.

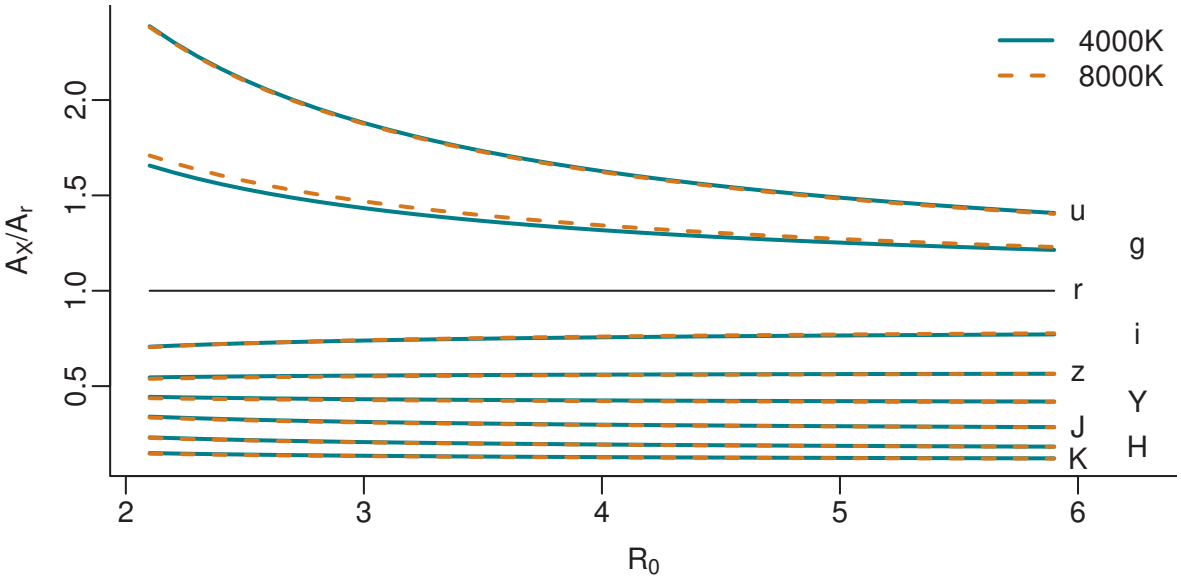

Figure 3.5: Model values for extinction relative to r-band, A_X/A_r, for two temperatures over a range of R_0 and the photometric bands used in this work. Small temperature dependencies can be seen for the optical bands u and g in particular at low R_0.

in Section 3.5.

3.3 Sampling and Computation

The forward model, priors and data are combined to compute the posterior function, as detailed in Section 2.1. As we are interested in inferring multiple APs, we use a standard Metropolis-Hastings Markov-Chain Monte-Carlo (MCMC) routine to efficiently sample the parameter space and to compute the posterior probability distribution. The parameter space is explored in logarithmic units of the respective quantity, forcing the estimates to remain positive and physical. The sampling matrix is diagonal with variances of 0.1 dex in each variable (AP). We assume no correlation between the variables. Convergence is sufficient with a chain length and burn in of 10 000 steps each. This does not need to be adjusted if we also infer R_0. We only expand the sampling matrix to account for the additional parameter.

To compute the likelihoods we can either use the actual spline function or build a

lookup table. Considering the numerical speed-up we achieve by tabulating results in advance, we generally use this approach. For this, we build a grid in extinction and effective temperature (and later in R_0 too) and calculate the predicted colours. We use the same parameter ranges as for training (avoiding extrapolation into regions not covered by the data), creating a grid with A_0 from $0 - 5$ mag in steps of 0.001 mag and T_{eff} from $3000 - 9000$ K in steps of 10 K (effectively the same as for the HRD). For R_0 we use a coarser stepsize of 0.2 in the range from $2.1 - 5.9$. The stepsizes are chosen as to be significantly smaller than the mean absolute error (MAE) of the residuals when using the full spline function directly (see Section 3.4.)

When assessing the performance in the APs and computing further quantities, we only use those stars whose inferred parameters are within the boundaries of the lookup table (i.e. within our training space). As indicated in section 3.1, at lower temperatures and smaller wavelengths the linear relation between colour and temperature for fixed extinction no longer holds, therefore extrapolation is not advisable. This avoids skewing the results due to boundary effects, like MCMC chains sticking to the edge of the grid in one or more parameters.

We apply this filter after parameter inference as we cannot rule out including stars that have true APs outside of our training range (we do not apply any colour-cuts to the data selection). If we were to instead force a prior over the same ranges during sampling we would artificially force those stars to have incorrect parameters.

Typically this filter throws out roughly 5 % of stars when estimating extinction, effective temperature and distance modulus. When including the R_0 parameter, this increases to close to 30 %, due to the additional variation in the model.

3.4 Validation

In this section we first demonstrate the viability of the model and then analyse the precision and accuracy of the method, as well as the colours. The following analysis is performed on a different set of 4 000 stars (*validation set*), of which none where used to train the model.

Simulated photometry

In this first step we aim to show the expected performance of the method and illustrate its limitations with respect to metallicity (see Figure 3.6). For this purpose we set up two extinction clouds at 500 pc ($\mu \approx 8.5$ mag) and 1500 pc ($\mu \approx 10.9$ mag) with extensions of 100 pc. This true extinction distribution with μ is shown as the

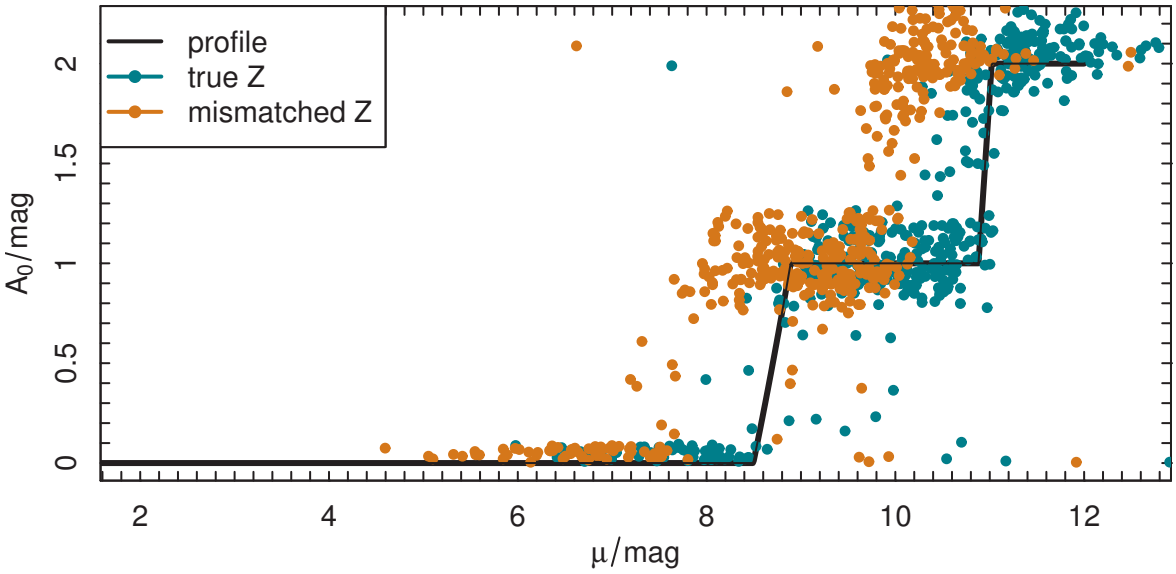

Figure 3.6: Inferred distance moduli μ and extinction A_0 for 500 stars. Turquoise dots indicate results for which matching HRD metallicity ($Z = 0.019$) is used for sampling and inference. Orange dots show results for mismatched HRD ($Z' = 0.0019$).

solid black line. We then place 500 stars along the line of sight, generating effective temperatures and absolute magnitudes from the HRD prior (solar metallicity). According to their effective temperatures, extinctions and distances, we simulate appropriate photometry and photometric errors, which are characteristic of the SDSS/UKIDSS sample we use. Finally we use our method to infer these parameters. These are plotted as points in the figure.

We infer the APs for two cases. In (a) we use the same solar metallicity HRD prior ($Z = 0.019$) as for selecting the stars (turquoise dots). In this case, we can see that our method recovers the distances and extinctions very well, with points closely tracing the extinction profile. The few points that lie furthest from the true profile typically have bad temperature estimates, resulting in erroneous distance moduli. In (b) we use an HRD with a tenth of that metallicity ($Z' = 0.0019$, orange dots) to infer parameters. Although the individual estimates for extinction and temperature are compatible, the mismatched stars are systematically shifted by $\Delta\mu = -1$ to closer distances. Analogously, the opposite shift occurs when reversing the mismatch. Therefore the accuracy of distance estimates is strongly limited by the unknown metallicity of each individual star.

It is difficult to estimate metallicity using only broad band photometry, as used here, when neither temperature nor extinction are fixed. As shown in Bailer-Jones et al. (2013) for simulated Gaia data, given spectroscopic information it is possible to estimate metallicity simultaneously, resulting in more accurate distance estimates.

Effects of colours and NIR data

Considering that we are using photometry from two surveys covering optical and near-infrared bands, it is instructive to compare how inclusion (or exclusion) of certain bands affects the parameter estimation. To examine this we train the forward model using (a) all nine bands, resulting in the eight colours *u-g*, *g-r*, *r-i*, *i-z*, *z-Y*, *Y-J*, *J-H* and *H-K*, (b) using only the five SDSS bands and thus only four colours and (c) using all but the SDSS *u* band magnitudes (seven colours). Generally, one would expect the use of more colours (more information) to yield better results, particularly when taking into account that NIR wavelengths are less affected by dust attenuation than shorter ones.

As expected, the performance of the method is significantly better when using the NIR data (using only NIR data is also much worse). Table 3.2 summarises the differences between the three cases. We characterise the performance with three key statistics. The first is the bias, which is purely the mean value of the difference between predicted and true parameter for all stars. The second is the mean absolute error (MAE), which similarly is the mean of the absolute difference and the third is the root mean square (RMS). For temperatures below 6500 K the performance is generally better than for higher temperatures. This results from the combination of having fewer stars in the higher temperature range and the scatter in the data being larger there (see Figures 3.1 and 3.3).

	A_0/mag			$T_{\rm eff}$/K		
	(a)	(b)	(c)	(a)	(b)	(c)
bias	-0.08	-0.41	-0.29	-67	-295	-269
MAE	0.23	0.59	0.41	300	628	440
RMS	0.45	1.05	0.68	586	1028	712

Table 3.2: Performance using all colours (a), just the four SDSS colours (b) and all bands except *u* (c). Table shows performance characterised by bias, mean absolute error (MAE) and root mean square (RMS) of the residuals, i.e. the difference between estimated and true parameters. $R_0 = 3.1$ is fixed in these models.

We note that the bias is negative in all three cases and parameters, though much closer to zero in case (a). Intrinsic scatter in the data also precludes higher residual accuracy. In Figure 3.7 we plot the results for the validation set of stars. Towards negative values in both parameters we see two distinct tails. Stars in the tails tend to have hotter true temperatures and larger intrinsic scatter. Only a few percent of stars are effected by such a large error in the parameter estimation. Without these tails, the performance would improve only minimally in MAE and RMS.

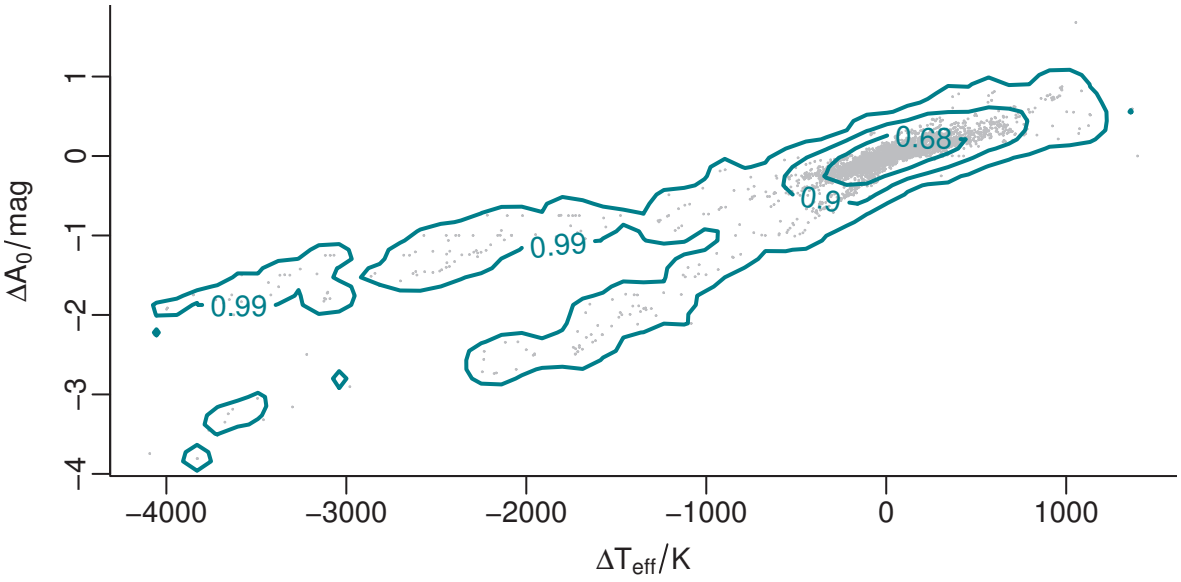

Figure 3.7: Residuals in effective temperature and extinction for case (a) using all eight colours. Plotted are the 68%, 90% and 99% density contours of the distribution of points in the plot. The differences are computed as $\Delta X = X_{estimated} - X_{true}$ for A_0 and $T_{\rm eff}$.

When excluding only the *u* band, case (c), we find that the performance is still better than case (b) although significantly worse than (a). The bias, MAE and RMS for A_0 are -0.29 mag, 0.41 mag and 0.68 mag, respectively. For $T_{\rm eff}$ these are -269 K, 440 K and 712 K.

In temperature and extinction we see no significant systematic effect due to the metallicity of the HRD. As indicated in the previous section, this mainly affects the distance estimator.

Figure 3.8 shows the residuals for both parameters as a function of the true parameters. We see no systematic trends, though we see a general bias towards negative values, which is evident in the values reported in Table 3.2.

As we do not have any true distances to compare the inferred distances to, we can not reference any results for distance modulus here. We must rely on the other parameters to be accurate in order to believe the distance results we obtain. We can however roughly estimate distances using the true extinctions and temperatures and finding the corresponding $\Delta = m_{\rm r} - M_{\rm r}$ via the HRD. Doing this we compute a bias of -0.23 mag, MAE of 2.29 mag and RMS of 3.31 mag when using all bands. Approximately 15 % of the estimates are better than 0.2 mag. In this case, the distance is poorly constrained by the combination of photometric data and our broad HRD prior, when $T_{\rm eff}$ and A_0 also have to be estimated from the data. The bias we encounter is primarily due to the unspecified mismatch of metallicities of the HRD and training data.

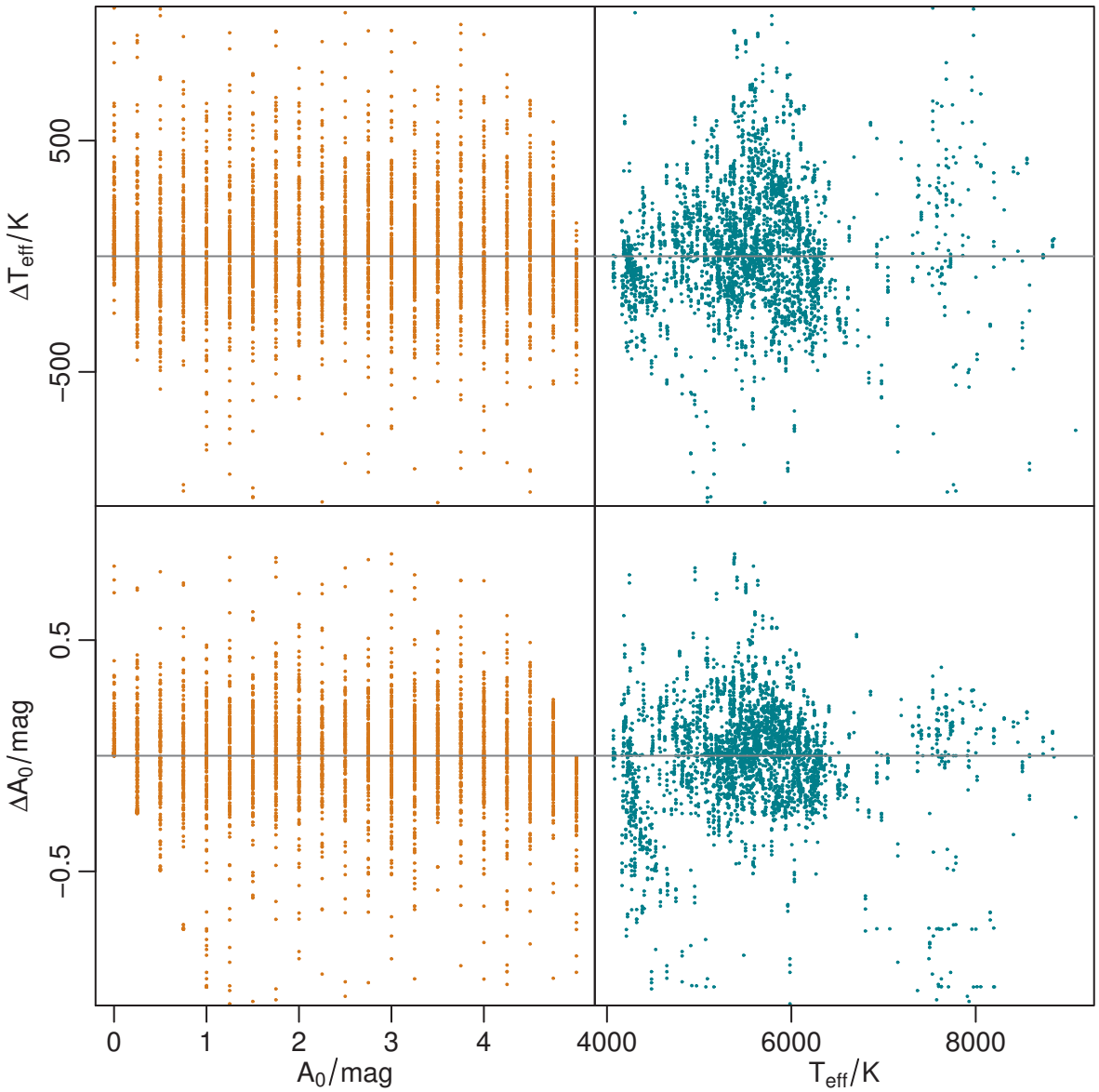

Figure 3.8: Residuals in effective temperature and extinction as functions of the true values in both parameters, as applied to the validation set of data. All eight colours are used. The true variation is shown for the same data as in Figure 3.7. A few dozen points are cut off by the selection of the axis range.

Precision of AP estimation

Beyond AP estimation accuracy (of the method), we are also interested in the precision, i.e. what are the individual parameter uncertainties for any given star. We quantify this using the width of the 68% confidence interval (CI_{68}) of the posterior PDF (which is equivalent to the 1σ variation), using a Gaussian kernel density estimation with the mean as the central point. The distribution is not generally symmetric, so we obtain marginally different values for the left and right confidence bounds. In Figure 3.9 we show the histograms of the average CI_{68} bounds for both extinction and effective temperature, as well as Δ. Mean values of the confidence intervals are $\langle CI_{68}(A_0)\rangle$ = 0.04 mag, $\langle CI_{68}(T_{\rm eff})\rangle$ = 58 K and $\langle CI_{68}(\Delta)\rangle$ = 2.0 mag. We can therefore be confident that our parameter estimates are generally precise in a statistical sense.

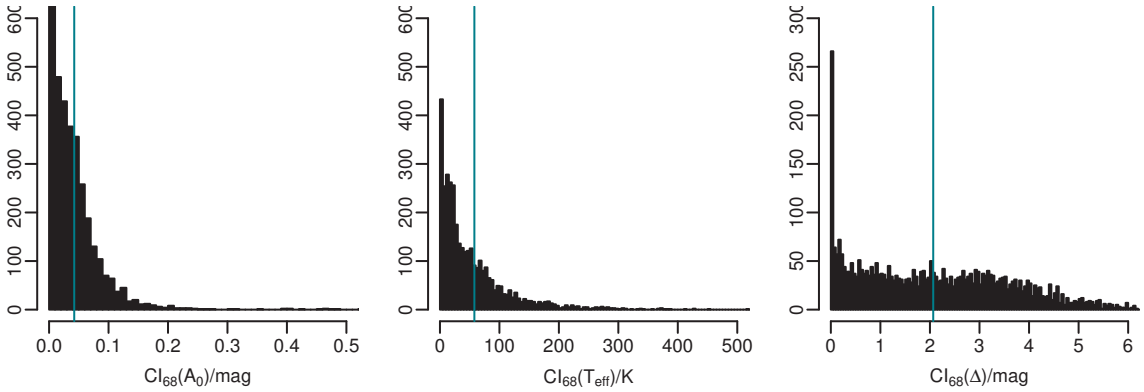

Figure 3.9: Histograms of 68% confidence intervals for extinction, temperature and Δ using photometric errors in the covariance matrix. The turquoise vertical lines indicate the mean values in each case. For A_0 this corresponds to 0.04 mag, for $T_{\rm eff}$ it is 58 K and for Δ it is 2.0 mag.

These values of precision are significantly smaller than the accuracy of the model (Table 3.2), which could be taken to mean that the parameter inference is overconfident compared to the information content of the data. The intrinsic scatter of the spline fit is of the same order of magnitude as the photometric errors of the data (see Section 3.2 and Figure 3.4), therefore properly representing the variation and uncertainties of the model.

Average photometric errors in all nine bands vary from 0.01 to 0.03 mag, the mean temperature uncertainty in the training data is 60 K, uncertainties in extinction can only arise from an erroneous correction of the SFD estimates, unless the reddening law is substantially wrong. These, though, are typically only $A_{\rm SFD,r}$ = 0.03 mag. Our precision therefore is on the order of the parameter errors. If we fix the errors in all colours to σ_{px} = 0.1 mag, which is roughly five times larger

than the photometric errors, we obtain precisions of $\langle \mathrm{CI}_{68}(A_0) \rangle = 0.36$ mag and $\langle \mathrm{CI}_{68}(T_{\mathrm{eff}}) \rangle = 402$ K. These values are in the order of 25 % larger than the MAE reported in Table 3.2. The corresponding biases, mean absolute errors and root mean squares only change by a few percent in respect to the tabulated results. Therefore, a more accurate account for model limitations could imply using significantly larger errors on photometry, although on average the statistical performance remains compatible in either case.

To illustrate the output from the Monte Carlo sampling, Figure 3.10 shows the contours and one-dimensional PDFs of the three parameters extinction, effective temperature and Δ for a single example star. In the extinction-temperature plane (bottom left panel) we clearly see the degeneracy between the two parameters. In the one-dimensional PDF for Δ (middle panel) we see the typical bimodal distribution, which is the result of sampling over the parts in the HRD that represent the main sequence and the giant branch. One peak is usually higher, which is why we use the mode to quantify this parameter rather than the mean. In most cases the higher peak relates to the main sequence (though not in this particular example), which is a reasonable justification to use it, as we have trained our models using photometry from dwarf stars.

It must be noted that the confidence intervals for Δ are large, though the inner 68% are more tightly constrained. We can reduce this spread by either running a longer Monte-Carlo chain, which adds more weight to the non-background regions in the HRD, or by directly decreasing the background level by a few orders of magnitude when computing it in the first place. Both options result in nearly unchanged parameter estimates but improved precision in Δ. The second option is clearly more appealing, as it does not increase the computation time.

Effects of relative extinction

The results shown so far were for a fixed relative extinction of $R_0 = 3.1$. In reality, this value is only the conventional mean value for the diffuse interstellar medium, whereas denser regions will typically be characterised by a higher value of R_0, arising from a change in size distribution and composition of the dust grains responsible for extinction in optical bands. This does not, however, equate high line-of-sight extinction with large R_0; it simply describes properties of the local dust.

We train a new forward model in which we simulate the change of colours due to extinction and R_0 for stars with known temperatures. We expand our parameter grid and again use 4 000 stars to train and validate the model. The general setup

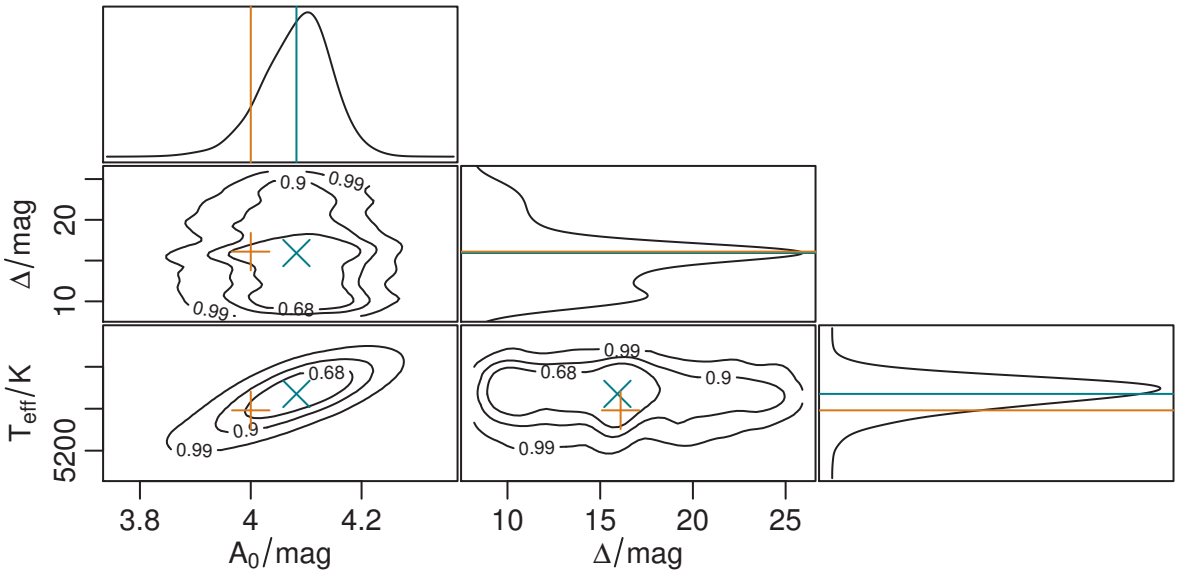

Figure 3.10: Contours of the posterior PDF for a star with true APs of $A_{0,true}$ = 4 mag and $T_{eff,true}$ = 5392 K. In each panel the turquoise cross or line shows the position of the estimated parameter value (A_0 = 4.06 mag, $T_{\rm eff}$ = 5446 K, Δ = 15.9 mag), whereas the orange plus or line is the true (constructed) value. Contours show 68%, 90% and 99% confidence intervals. The side panels show the marginalised one-dimensional PDFs for the single parameters, as aligned with the respective axes. The peak densities are scaled to 1 in each case. The indicated true value of distance modulus is estimated using the true temperature and the HRD from section 3.1. The best estimate values for temperature and extinction are computed using the mean of the distribution, whereas the mode is used for distance modulus.

remains the same as before. We use R_0 in the range from 2.1 to 5.9. The model performance we achieve like this is presented in Table 3.3.

	A_0/mag	$T_{\rm eff}$/K	R_0
bias	−0.22	−206	0.32
MAE	0.39	438	0.73
RMS	0.69	771	0.95

Table 3.3: Performance using all colours and variable R_0, characterised by bias, mean absolute error (MAE) and root mean square (RMS) of the residuals, i.e. the difference between estimated and true parameters, as detailed previously.

Comparing case (a) in Table 3.2 with Table 3.3, we see that modelling R_0 slightly deteriorates the performance in the other two parameters, the standard deviation and MAE in the residuals are increased by roughly 50% for temperature and extinction. This comes as no surprise, as we are now extracting more parameters, but with the same initial data, noting that the observed change in colours are quite small for larger differences in this parameter. In general we are able to roughly infer R_0, though the accuracy is not sufficient to confidently differentiate between, say, 3.1

and 3.8. We therefore build the lookup table using somewhat larger steps of 0.2 in R_0 for the range noted above, and the other parameters remaining the same.

In Figure 3.11 we present the precision of the predicted parameters as histograms, where the turquoise lines again depict the mean values. We obtain $\langle \mathrm{CI}_{68}(A_0) \rangle$ = 0.05 mag, $\langle \mathrm{CI}_{68}(T_{\mathrm{eff}}) \rangle$ = 66 K, $\langle \mathrm{CI}_{68}(R_0) \rangle$ = 0.07 and $\langle \mathrm{CI}_{68}(\Delta) \rangle$ = 1.6 mag. The first two values are minimally larger than in the previous case, though it is to be expected considering that we have added an additional degree of variation to the model. Again, the confidence intervals are significantly smaller than the model accuracy.

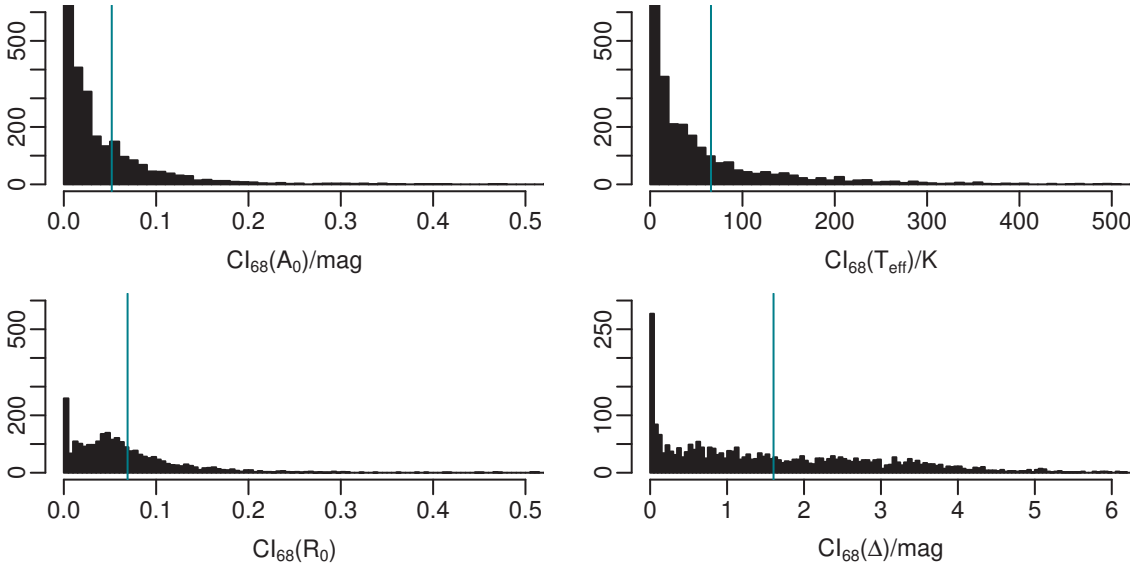

Figure 3.11: Histogram of 68% confidence intervals of extinction, temperature, extinction factor and Δ, which are estimates of the uncertainty on each parameter we infer. The turquoise vertical lines indicate the mean values in each case. For A_0 this corresponds to 0.05 mag, for T_{eff} it is 66 K, for R_0 it is 0.07 and for Δ it is 1.6 mag.

3.5 Maps

We have so far summarised the accuracy and precision of the model using data with known stellar parameters. We now apply the model to estimate APs for large areas of the sky where individual estimates are not available. Querying the *WFCAM Science Archive* with the flags detailed in section 3.1, but imposing no constraints on SSPP information, we obtain a total of 4 191 659 unique stellar objects, of which 3 055 954 and 1 135 705 are north and south of the Galactic equator, respectively. We also pick out a section of the southern sky with $l > 180^{\circ}$, which we use to illustrate the variations of the model, namely using only photometry from SDSS, using all photometry, and also including R_0 in the parameter estimation.

We first show the full projected maps of the northern and southern part of the surveys in Figure 3.12. Here we use a Mollweide projection in Galactic longitude and latitude to obtain equal area pixels, using a resolution of 11.48 arcmin/pixel in both coordinates. This corresponds to an average of 29 stars per pixel, depending on the region observed, with extremes of 3 and 250 stars per pixel.

The interstellar medium towards the Galactic poles typically hosts diffuse gas and lines of sight with low extinction. Nevertheless, the area jointly probed by the two surveys does cover some regions with higher average extinction. Generally we can pick out larger structures that are also visible in SFD maps of the same resolution. Towards the right of the southern part of the sky one can start to make out parts of a larger structure, with extinction estimates in the range of 1 to 1.5 mag in the *r*-band.

To more quantitatively assess the credibility of the presented map, we compare the computed extinction values with those from SFD. For this we project the data in an identical manner and calculate the difference between the SFD extinction values and our own estimates. This is plotted in Figure 3.13 for the southern part of the sky. We refrain from showing the equivalent image for the northern part, as it qualitatively shows the same behaviour. In Figure 3.14 we show the histograms for the differences between SFD extinction estimates and our own for both the southern and northern part of the sky, based on the binned data. In the south the mean difference and standard deviation are -0.02 mag and 0.13 mag, respectively. In the north we obtain -0.06 mag and 0.11 mag.

The general trend favours very small differences, with the mean ΔA_r being close to zero. Some regions indicate that SFD predicts higher extinction than we do. This can have several reasons, one being the fact that we compute extinction along a line of sight to individual stars, rather all the way to the edge of the Galaxy. Then again, we are looking at high (absolute) latitudes, where we assume extinction in general to be small, allowing us to cover most of the stars. Selection effects and limitations set by our data retrieval could affect this assumption. The few pixels that have a red hue, indicating that our model predicts higher extinction, are mostly artificial effects that arise from binning the data, leaving only very few stars in those bins. If these stars happen to have (erroneously) high extinction estimates, they may not be averaged out.

This is visualised in Figure 3.15, where we show the standard deviation of A_r for each pixel over the southern part of the sky. We see a large range of estimates in each pixel, independent of the actual mean extinction in them and the position on the sky. There is a noticeable correlation between the standard deviation and the (absolute) difference between our extinction estimate and that of SFD, $\Delta_{A_\mathrm{r}} =$

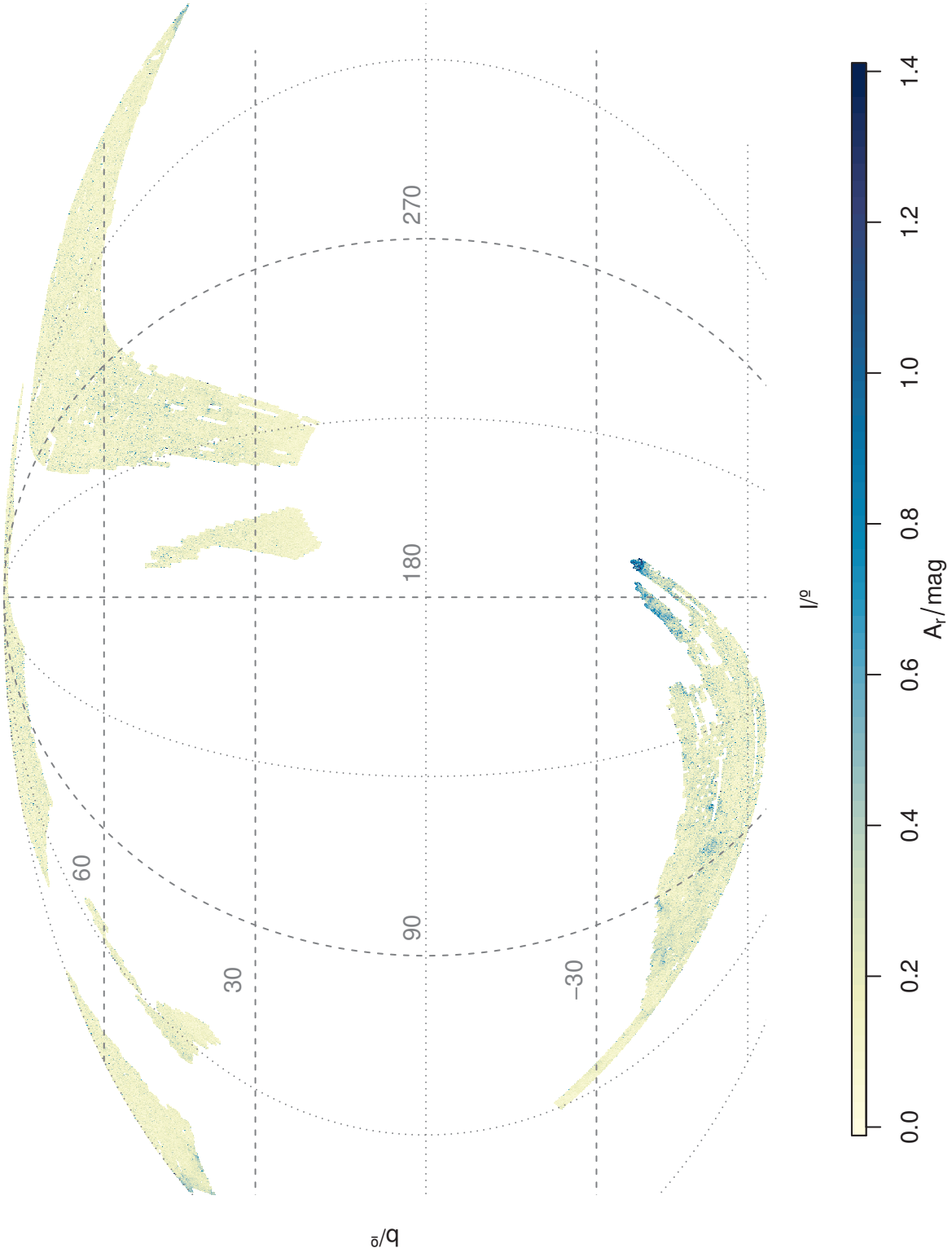

Figure 3.12: A Mollweide equal area projection of the mean extinction in the full cross-matched portion of the sky, centred on the Galactic anti-centre. The stars are binned using resolution of 11.48 arcmin/pixel in l and b. The non-Cartesian grid in latitude and longitude is overplotted. The colour scale shows the mean extinction A_r in any given pixel. White areas are not jointly covered in the surveys.

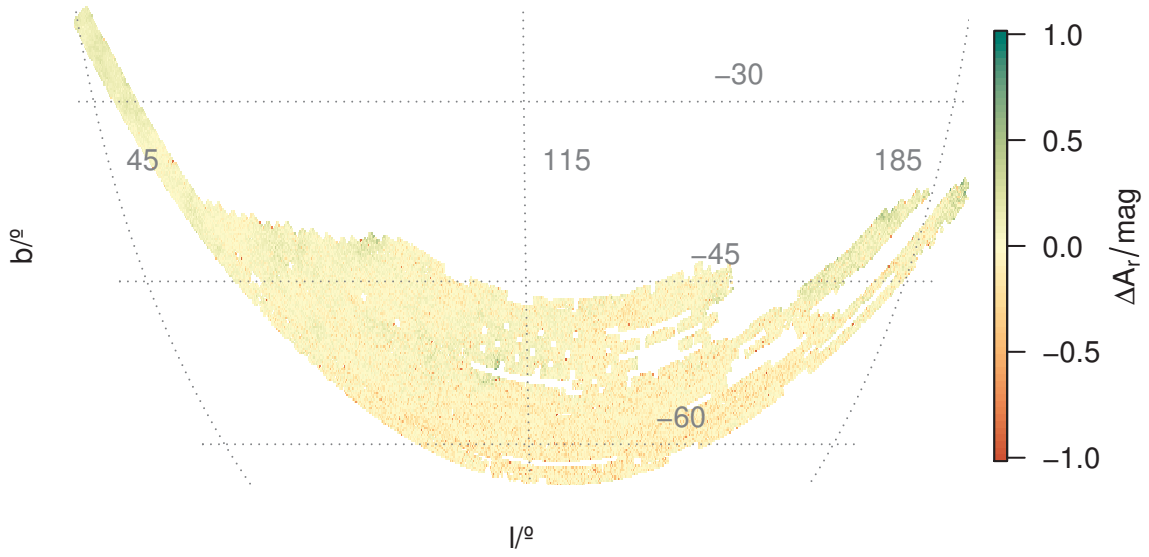

Figure 3.13: Difference between SFD extinction estimates and our own for stars in the southern part of the sky, using a Mollweide projection with 11.48 arcmin/pixel resolution. The colour coding in the range of $\Delta A_r = -1$ to 1 mag shows regions where SFD predicts a higher extinction in green/blue (positive), and where our model has a higher estimate in yellow/red (negative).

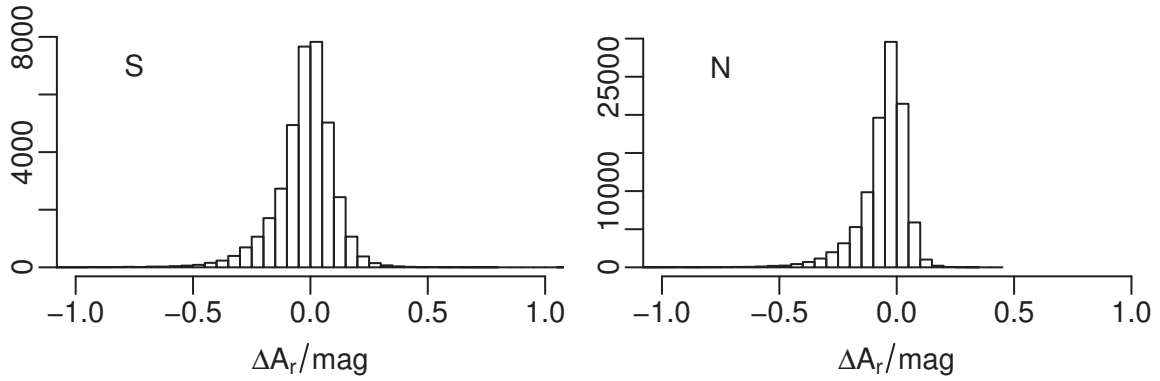

Figure 3.14: Histograms of differences between SFD extinction estimates and our own for stars ($\Delta A_r = A_{r,SFD} - A_r$) in the southern part of the sky (left) and in the northern (right).

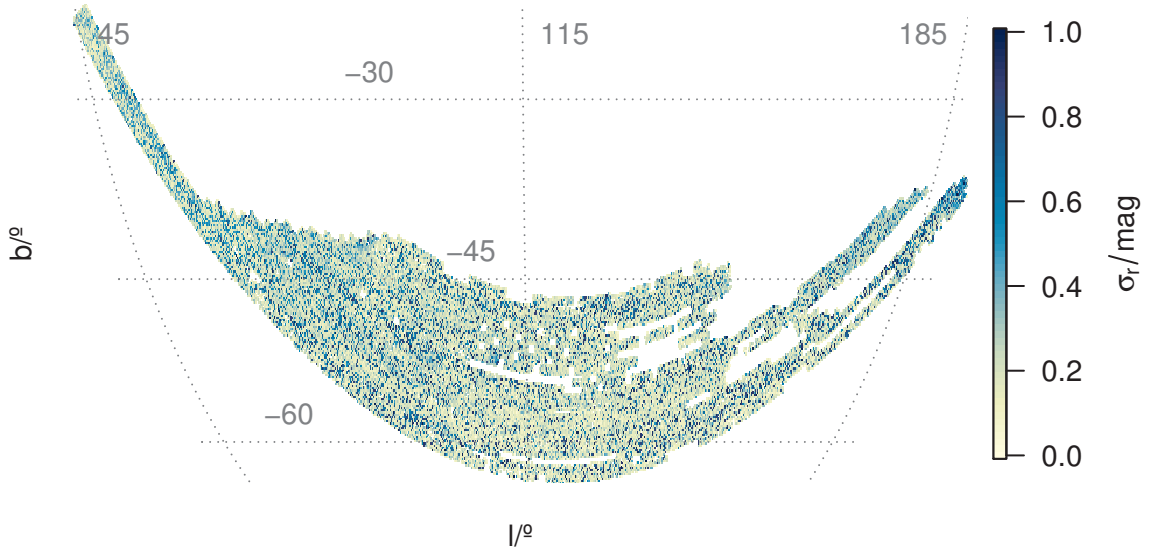

Figure 3.15: Standard deviation σ_{A_r} of r-band extinction over the southern sky using nine photometric bands and fixed R_0. The colour axis is constrained from 0 to 1, only 0.5 % of the cells have higher values.

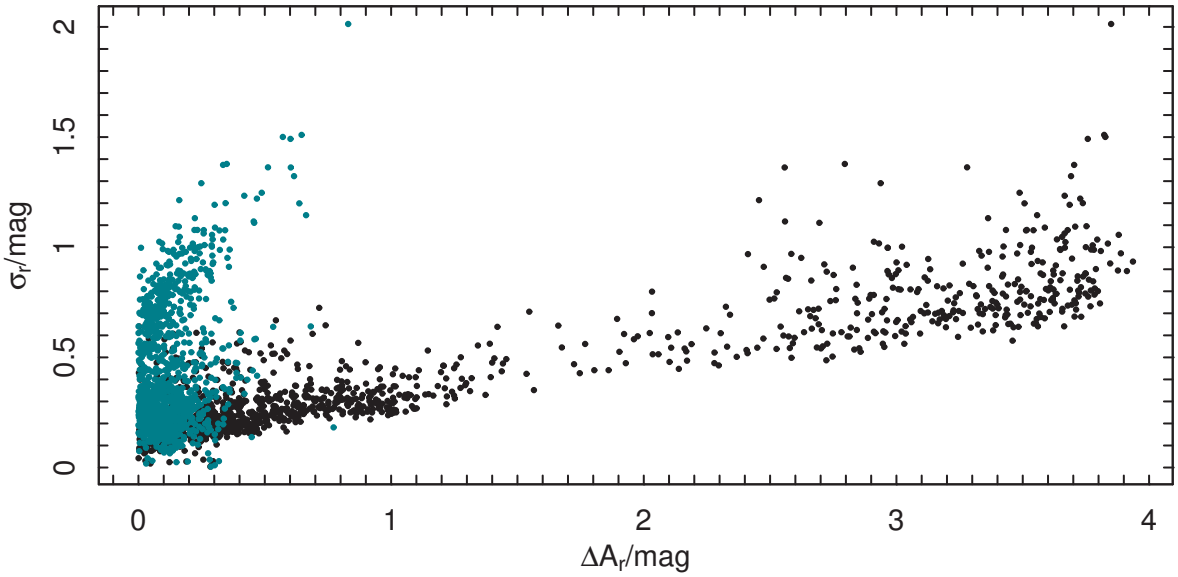

Figure 3.16: Standard deviation σ_{A_r} of r-band extinction as a function of $\Delta A_r = |A_r - A_{r,SFD}|$, the difference between our estimates and SFD values for each pixel in the southern sky using nine photometric bands and fixed R_0. Turquoise dots are used for mean extinction estimates, whereas black dots use the maximum extinction value in each pixel (see Figure 3.18).

$|A_r - A_{r,SFD}|$, see Figure 3.16. Turquoise dots are computed using the mean extinction estimates for A_r in the equation above. Black dots use the maximum extinction value in each pixel, i.e. $A_{r,max}$ (see Section 3.5 and Figure 3.18 for further details). In this case the range of residuals is naturally larger. In both cases, though, the standard deviation in a pixel is a function of the difference between the two extinction estimates.

On the whole, the results are consistent with how we expect the model to function, with results obtained from SFD or other data that deliver the integrated extinction to the edge of the Galaxy. On average, we slightly underestimate the mean extinction in a pixel due to the fact that we are averaging over multiple stars with varying degrees of precision in their individual estimates.

Extinction prior

A simple prior to implement is using SFD estimates to constrain maximum extinction along any line of sight. Though the SFD maps do have several systematic problems, overestimating extinction particularly towards the Galactic plane and in regions of high extinction (see, e.g. Schlafly et al., 2010; Schlafly & Finkbeiner, 2011), at the latitudes probed here the corrections are only on the percent level. By constructing a broader prior we can safely ignore this.

In practice we implement the prior as a step function with the probability dropping from 1 to 0 at $1.3 \cdot A_{\mathrm{r,SFD}}$ with $A_{\mathrm{r,SFD}}$ being the converted SFD r-band extinction. This factor of 1.3 allows for enough range in the extinction estimates to account for errors in the reference value and smoothing of the data, whilst also not being too restrictive. To compare the performance of a model using this prior we use the known, simulated extinction. We see only negligible changes to the values in Table 3.2 when adding the prior. Naturally, more stars remain within the training grid, as maximal values in extinction are constrained. Considering average parameter estimates do not change much, we can be confident that the model can properly characterise the degeneracy intrinsic to the data, independent of stronger priors.

When applying this updated model to data (i.e. with unknown APs), we do notice some changes in respect to average extinction estimates. In Figure 3.17 we show the differences for a region of the southern sky. The data are projected using a Mollweide projection and colour-coded according to the mean extinction in each pixel. The left panel depicts the results without priors (fixed R_0 and eight colours). In the right panel, the equivalent map of SFD estimates is shown, and the centre panel that of our model including the SFD-based prior on extinction. Clearly, inclusion of the prior reconstructs more closely the SFD reference map, whereas the alternative produces a less smooth map, although it still closely follows the depicted distribution on a whole, picking out the high extinction regions towards the top-right, as well as the general trends. Individual pixels are not always accurate though. This is due to the fact that we estimate the APs for every star individually, whereas the smoothed SFD map is a proxy for the average extinction along a line of sight.

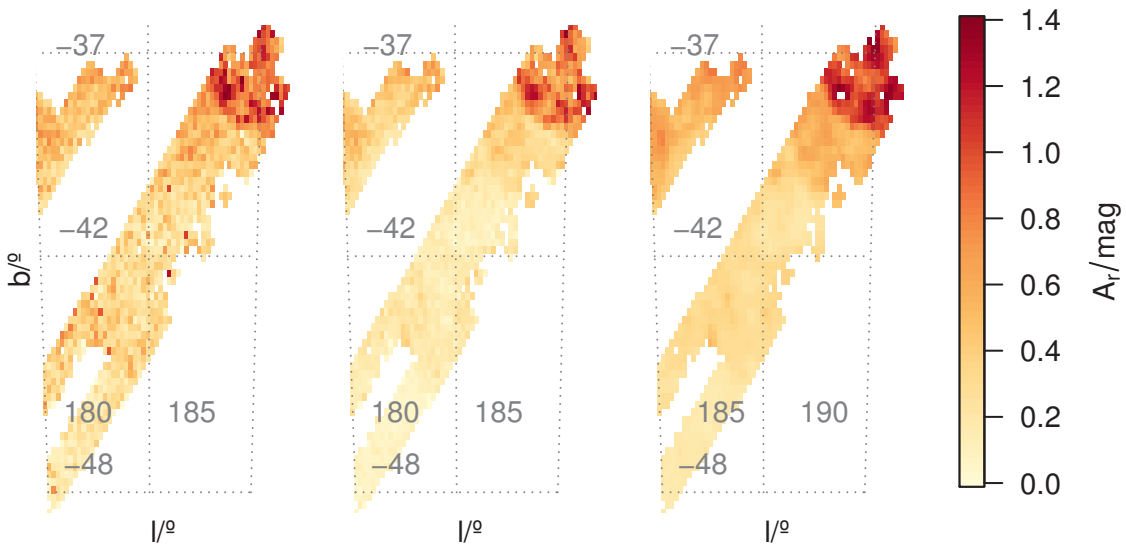

Figure 3.17: Extinction maps for a subset of the data in the southern sky in a Mollweide projection. The non-Cartesian grid in latitude and longitude is overplotted. The data are binned to a resolution of 11.48 arcmin/pixel. The left panel shows the mean r-band extinction in each pixel for the case of using no extinction prior, the middle panel that for a prior based on SFD estimates. In both cases eight colours are used with fixed R_0. The right panel shows for comparison the same region using the SFD estimates for the direction of each pixel directly. White areas are either not covered by the data or are missed due to binning (due to post-processing, the panels do not necessarily use an identical set of stars, therefore some pixels appear white in the left panel but not in the right two.)

This issue becomes more pronounced when we use the maximal value in any bin to create the map. This is shown in Figure 3.18, where we project the maximum values for the case of using the standard model (left) and including the prior (right). The equivalent map for SFD values is identical to the right panel in Figure 3.17 as the reported values are smoothed already. With this prior we limit the maximum value for each line of sight. The resulting map is almost identical to the SFD reference map, with slight differences visible particularly at the edges of the footprints due to strongly varying stellar number densities. Ignoring the prior we compute varying maximum values of extinctions, as we have no low boundary limit. As can be seen in Figure 3.17, when using the mean these extreme values tend to average out, even though individual estimates may still miss the true (or expected) value.

Extinction profiles

In addition to extinction and temperature we have inferred distances to all stars. We also have full probabilistic information available for each parameter which we can use to compute profiles of extinction as a function of distance for individual lines of sight, as detailed in Section 2.3. We repeat this for every (arbitrary) value of distance

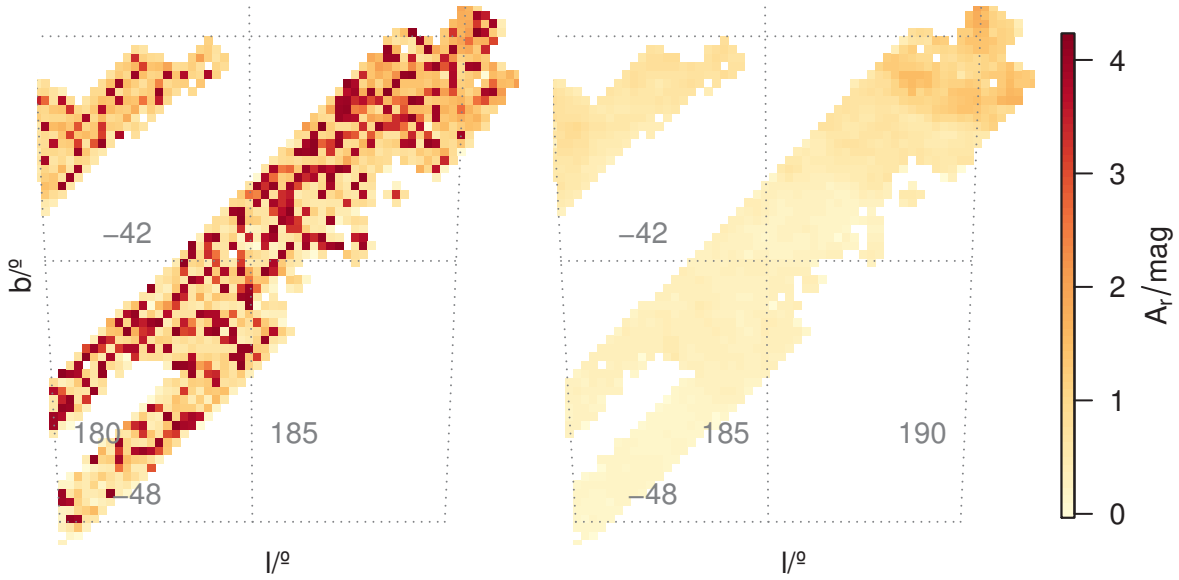

Figure 3.18: Region and resolution as in Figure 3.17. The left panel shows the maximum *r*-band extinction in each pixel for the case of using no extinction prior, the right panel that for a prior based on SFD estimates. In both cases eight colours are used with fixed R_0. Note that the colour scale has changed to cover the adjusted parameter range.

modulus we would like to use. This way we can compute extinction profiles using the full PDF of each star instead of merely averaging the data directly. Similarly this can be done for A_0 or proper distance too.

An example is shown in Figure 3.19 for the region $l \in (189.56, 190.88)^\circ$ and $b \in (-37.35, -36.35)^\circ$, which is in the top right region of Figure 3.17. We have chosen this region because it has higher extinction and as such can demonstrate the application of Equations 2.14 and 2.15. The figure shows two profiles, both using the nine band forward model, one without prior (turquoise) and one including a prior on extinction as in Section 3.5 (orange). The profiles are similar in the sense that they both are essentially flat, possibly with a slight increase towards further distances, though the error bars on extinction are compatible with both interpretations. The errors on distance modulus are implicitly included in the way the mean extinction and its error are computed (see Equations 2.14 and 2.15).

The average precision of distance modulus estimates varies from approximately 0.5 mag to 1.5 mag when going from the smallest to the largest value of μ in Figure 3.19. The SFD extinction for this field is $A_{\mathrm{SFD,r}} = 1.16$ mag, which is covered by the asymptotic error bars. Values in the orange curve tend to be slightly smaller, in line with conclusions from the previous section, where the maximal extinction estimates are constrained by the prior. However, the values are compatible within their uncertainties. Despite imposing no constraints on distance modulus when inferring the APs (apart from fitting into the parameter space of the HRD) and when comput-

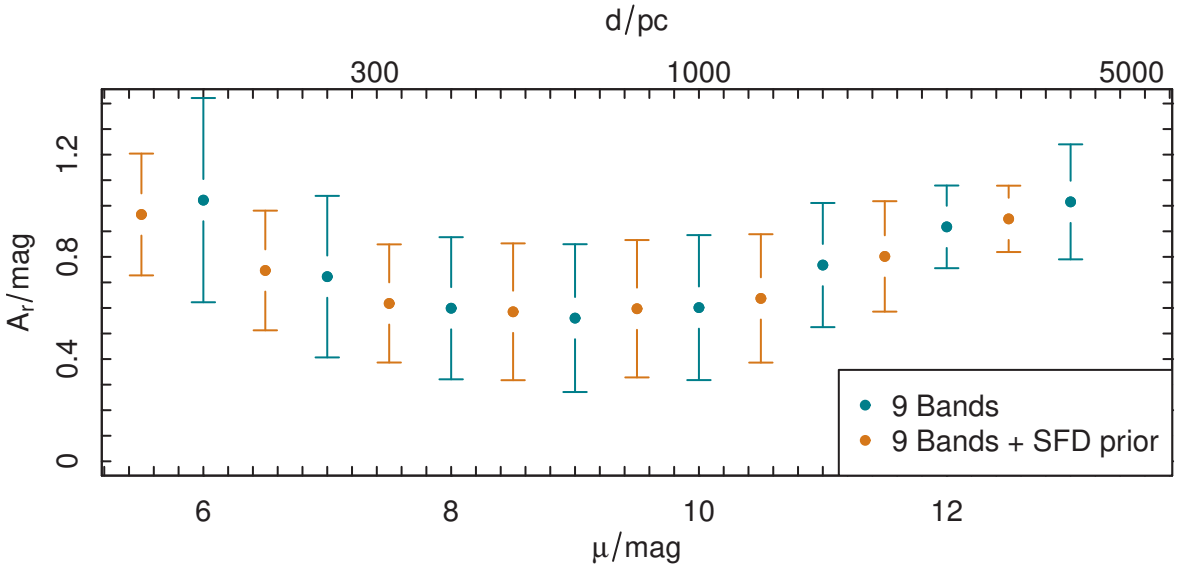

Figure 3.19: Mean *r*-band extinction as function of distance modulus for 520 stars in the region $l \in (189.56, 190.88)^\circ$ and $b \in (-37.35, -36.35)^\circ$. Mean values and error bars are computed as per Equations 2.14 and 2.15. The average SFD extinction for this field is $A_{\mathrm{SFD,r}} = 1.16$ mag. Both profiles use all nine bands, turquoise points show the case of no prior and orange points including a prior on extinction (Section 3.5). Top axis shows the corresponding distances $d = 10^{\mu/5+1}$ in parsec.

ing the profiles, the variation of extinction with distance is physically consistent with extinction not decreasing with increasing distance. Furthermore, the shapes of the profiles are in agreement with the assumption that the stars at high latitudes reside behind the dust layers. Given the bright magnitude limit of the surveys used, we do not expect to find stars with distance moduli less than $\mu = 6$ mag. Indeed, our method estimates very few stars to have smaller values of μ.

Lines of sight in regions with lower extinction behave in a similar manner, though naturally with lower asymptotic values of the extinction. In particular, cells at small distances show larger scatter, which can be attributed to low sampling statistics at these values, with only a few stars being computed to be very close by. In contrast, at larger distances the extinction in each cell is computed using hundreds of stars (owing to large uncertainties in the distance estimation and roughly d^3 growth in the volume of a cell with distance).

Inferring R_0

As detailed in Section 3.4 we are able to infer the R_0 parameter with moderate accuracy, in addition to extinction, temperature and Δ. Using the corresponding forward model, we rerun the method on the same small field of stars as before. We refrain from showing the projected extinction map as it does not differ qualitatively

from that with constant R_0 (Figure 3.17 left and centre panels). We instead show in Figure 3.20 the mean distribution of R_0 in the same projection.

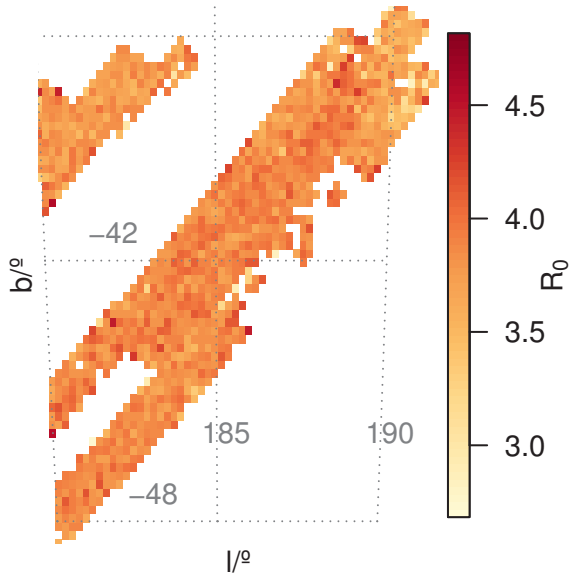

Figure 3.20: Distribution of the parameter R_0 over a subsection of the southern sky with an average over the whole field of $\langle R_0 \rangle = 3.81 \pm 0.20$. The same resolution as in previous maps is used.

Generally we see quite some scatter around the mean $\langle R_0 \rangle = 3.81 \pm 0.20$ in this field. Using the raw data (i.e. not binned) we obtain $\langle R_0 \rangle_{raw} = 3.82 \pm 0.47$, which has a slightly larger standard deviation about the mean. Noting that we average over a large region with varying degrees of extinction and that the model accuracy is only MAE(R_0) = 0.74, this should not be taken as a strong statement concerning the average R_0 of the diffuse ISM.

Nonetheless, this supports results by Gontcharov (2012a,b), who finds R_0 to be quite large and to have significant variations at high latitudes and low extinctions. In comparison, Berry et al. (2012) find $R_0 = 3.0$ with 0.1 random and systematic uncertainties, although they argue that for $A_r < 2$ mag, estimates on R_0 are very unreliable.

Looking at the relation between A_r and R_0 (for raw and binned data) we see in Figure 3.21 that there is a trend ($R_0 = (-0.45 \pm 0.2) A_r[\text{mag}] + 4.0$), again consistent with the previous statement in the limit of low extinctions.

Using the same technique as in Section 3.5 we compute the profiles of R_0 as a function of distance for a small region at the top right of Figure 3.17 (as in Figure 3.19). This is shown in Figure 3.22 in orange (left axis), where the mean extinction has been replaced with mean R_0 in the equations. In turquoise the extinction profile for the same data is plotted (right axis). As in the fixed R_0 case, we see a

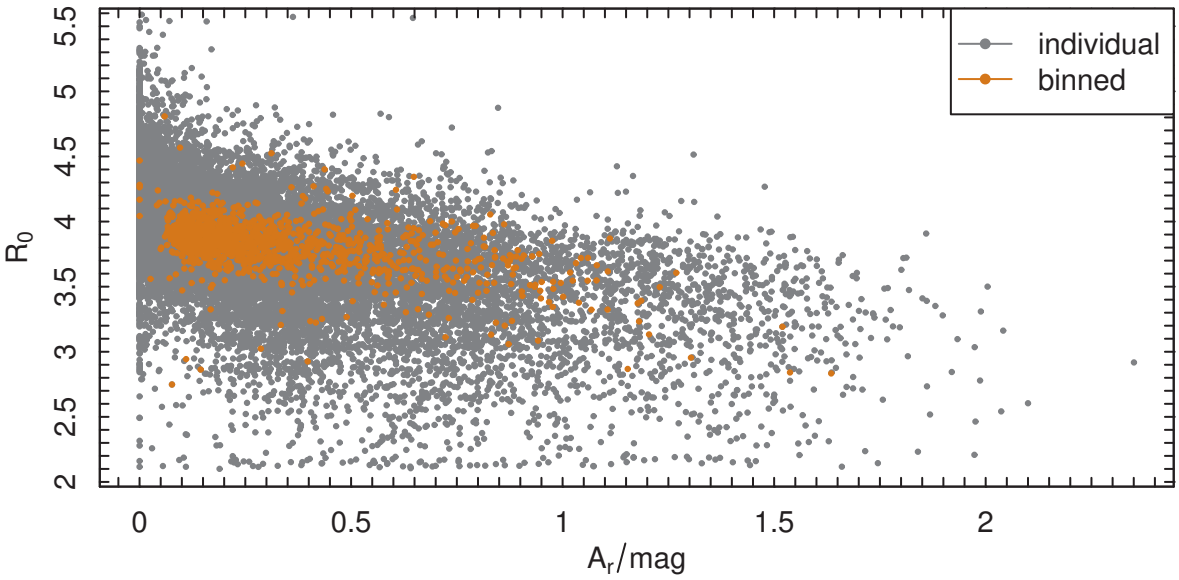

Figure 3.21: For each pixel of the map in Figure 3.20 we plot R_0 vs. A_r. Grey points are the raw, unbinned data, orange are for the binned cells. We see the scatter about the mean $\langle R_0 \rangle = 3.81 \pm 0.20$ (red) and $\langle R_0 \rangle = 3.82 \pm 0.47$ (black).

slight increase in extinction towards larger distances, though it is more pronounced now. At the same time R_0 increases quite quickly, although the amplitude of variations are roughly the size of the error bars and well within the model accuracy. To be able to further analyse this interesting aspect, we need to probe regions with higher values of extinction (not covered by the crossmatched SDSS/UKIDSS data) or obtain more informative data from other sources. As it stands, we can only make the general statements mentioned above.

3.6 Summary

Combining large optical and near infrared surveys of stars, such as SDSS and UKIDSS, is an effective means of constraining interstellar extinction as well as addressing the degeneracy between effective temperature $T_{\rm eff}$ and extinction A_0 for individual stars. To achieve this, we construct a forward model trained on colour changes of real data due to extinction, effective temperature and R_0 in combination with self-consistent use of an HRD prior to infer distance information. Incorporating these physical constraints, we infer astrophysical parameters (APs) in a Bayesian framework using an MCMC algorithm to efficiently sample over the posterior distribution. This way we are able to naturally extract full probability distribution (PDF) information in each AP. Testing with synthetic data has shown that this method produces accurate results, with the caveat of not fully incorporating the systematic effect

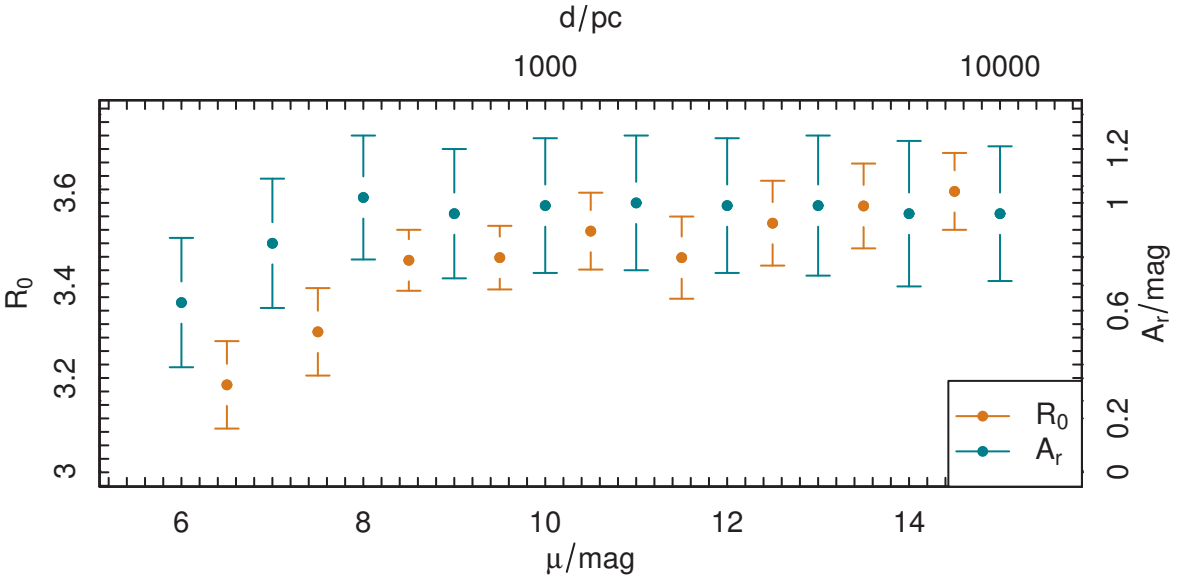

Figure 3.22: Mean R_0 (orange, left axis) as function of distance modulus for a subset of stars from Figure 3.20. The same technique as in Section 3.5 is used to compute the mean and standard deviation. In turquoise the extinction profile is plotted (right axis). Distances in parsec are given in the top axis.

metallicity has on distance estimation.

We have investigated the impact of the data and assumptions on the model, in particular the effect of removing photometric bands, inferring R_0 in addition, and the effect of an incorrectly assumed metallicity on the results. When using nine bands (five from SDSS, four from UKIDSS) we obtain a mean absolute error of the residuals of the model for A_0 of 0.23 mag and 300 K in $T_{\rm eff}$. These values are up to 50% better than just using SDSS photometry. The MAE for $\Delta = m_{\rm r} - M_{\rm r}$ is 2.2 mag. With the same set of data we can also achieve an accuracy of 0.73 in R_0, with acceptable changes to the performance in the other parameters. Using the 68% confidence intervals (CI) to quantify the individual precision of the AP estimates, we obtain the following values: $\langle {\rm CI}_{68}(\Delta)\rangle$ = 1.6 mag, $\langle {\rm CI}_{68}(A_0)\rangle$ = 0.05 mag, $\langle {\rm CI}_{68}(T_{\rm eff})\rangle$ = 66 K, $\langle {\rm CI}_{68}(R_0)\rangle$ = 0.07. When estimating APs for fixed R_0 the precision improves slightly. Accuracies of distance modulus (or Δ) estimates are strongly dependent on the correct assignment of metallicity to the stellar population being analysed. This can either be achieved by matching the metallicity of the HRD prior, or by directly inferring metallicity along with the other parameters. This, however, requires higher quality and more informative data or spectroscopic observations of the stars.

Although we estimate APs for stars individually, our method is able to trace the intermediate-size dust structures visible in e.g. SFD dust maps of the same regions. In addition to line of sight extinction estimates (either averaged or per star) we can use the full PDFs to compute probabilistic profiles of A_0 and R_0 as a function of

distance, indicating how dust is distributed along the line of sight. We generally find the expected result that, at these Galactic latitudes, the observed stars are behind the layers of dust.

We have limited this work to stars at high Galactic latitudes common to SDSS and UKIDSS Large Area Survey, allowing us to obtain photometry in nine bands. In principle this method can use fewer bands, with the decrease in accuracy detailed above, or use data from other surveys, such as Pan-STARRS (Kaiser et al., 2002) or 2MASS (Skrutskie et al., 2006). This would allow us to probe high extinction regions towards the Galactic plane and further address the performance in respect to R_0.

As analysed in Bailer-Jones (2011), the method can be expanded to combine distance estimates, such as parallaxes, with the HRD prior in order to provide a more accurate estimate of APs. Furthermore, metallicity and surface gravity may be estimated as well, although, as noted above, this requires higher quality data. If we could estimate them using spectroscopy or more photometric bands, then this would alleviate the strong assumption made by implementing an HRD with fixed metallicity, and thus potentially improve the accuracy of distance estimates.

Current work for the Gaia (Perryman et al., 2001) data processing pipeline, as summarised in Bailer-Jones et al. (2013), illustrates AP estimation performance using Gaia spectra and photometry. By combining current multiband surveys with Gaia parallaxes we expect to be able to increase the accuracy of the method and precision of the parameter estimates significantly.

4

Dust in the Galactic Plane

This chapter is adapted from Hanson et al. (2016). Whilst conceptually similar to Hanson & Bailer-Jones (2014), we adjust the method to use photometry from Pan-STARRS 1 and GLIMPSE to infer extinction and distances to stars in the Galactic plane. This significantly different environment allows us to map stars with a much larger range of extinction values.

4.1 Data Sources

We crossmatch PS1 and Spitzer IRAC 3.6 μm point source data from the Glimpse surveys using the API of the cross-match service provided by CDS, Strasbourg[1] with a 1 $''$ search radius. This results in a data set with 19 885 031 individual stars. Details on the surveys and data selection are noted below.

Pan-STARRS1

The Pan-STARRS1 survey has observed the entire sky north of declination $-30\,^\circ$ in five filters (Stubbs et al., 2010; Tonry et al., 2012). These cover the wavelength range $400 - 1000$ nm. The resulting global photometric calibration is better than 1 % (Schlafly et al., 2012).

We select all point sources classified as stars that have good observations in the five bands, g_{P1}, r_{P1}, i_{P1}, z_{P1} and y_{P1}, using the epoch-averaged photometry in each band. We use data collected up to February 2013. We do not take into account any variability observed across multiple epochs. About 90 % of stars have g_{P1}-band magnitudes between $16.19 - 21.95$ mag. Only a tiny fraction of the stars have

[1] cdsxmatch.u-strasbg.fr/xmatch

photometric uncertainties worse than 0.1 mag, the median uncertainties in the five bands are in the range $0.01 - 0.02$ mag.

Spitzer GLIMPSE

The Spitzer Space Telescope Legacy program GLIMPSE consists of four separate surveys (I, II, 3D, 360), which together cover most of the Galactic plane within a few degrees in latitude. The Infrared Array Camera (IRAC; Fazio et al., 2004) is used to image at 3.6, 4.5, 5.8, and 8.0 μm. We use only the 3.6 μm data, as the longer wavelength measurements do not improve our parameter estimation. We select point sources that have signal to noise ratios greater than 3 and closed source flags (`csf`) of 0, indicating that no other sources are within 3 $''$ of a source. This is to ensure that sources are extracted reliably. The 90 % quantile for 3.6 μm-band magnitudes is $11.20 - 16.47$ mag.

Hertzsprung-Russell Diagram

As in Chapter 3, we use a HRD prior as a constraint in the $T_{\rm eff} - M_{\rm r}$ -plane. To fully account for the distribution in stellar types expected in the Galactic plane, in particular K and M dwarfs, as well as giants, we use the Dartmouth Stellar Evolution Database (Dotter et al., 2008). For fixed solar metallicity, we smooth the data in the HRD plane using a binned kernel density estimate with bandwidths of 25 K and 0.125 mag in $T_{\rm eff}$ and $M_{\rm r}$, respectively. The temperature range is from $2500 - 10\,000$ K, the absolute magnitudes vary from -4 mag to 12 mag. The resulting grid has the pixel dimensions of 751×600 (as $T_{\rm eff} \times M_{\rm r}$). Before normalisation, a small, non-negative offset is added to all pixels to account for the fact that the regions that are empty in the Dartmouth model HRD in reality may not have exactly zero probability. We show a representation of the HRD in Figure 4.1.

The HRD of course depends on the metallicity, and as demonstrated in Section 3.4 the choice affects the results. Unsurprisingly, it is not possible to also estimate metallicity from our photometric data (due in part to the large - a priori unknown - range of $T_{\rm eff}$ and A_0 in the data). If we fixed the metallicity of the HRD to a single value, we would obtain artificially precise (but not necessarily more accurate) results for the inferred parameters. To avoid this, we took an HRD and then smoothed it (using a kernel density estimation method). This produces a smooth but conservatively broad HRD; it is broader than the one used in Chapter 3. As demonstrated there, the lack of a metallicity determination will be the main limiting factor on the distance accuracy, while the extinction, extinction parameter, and effective temperature are

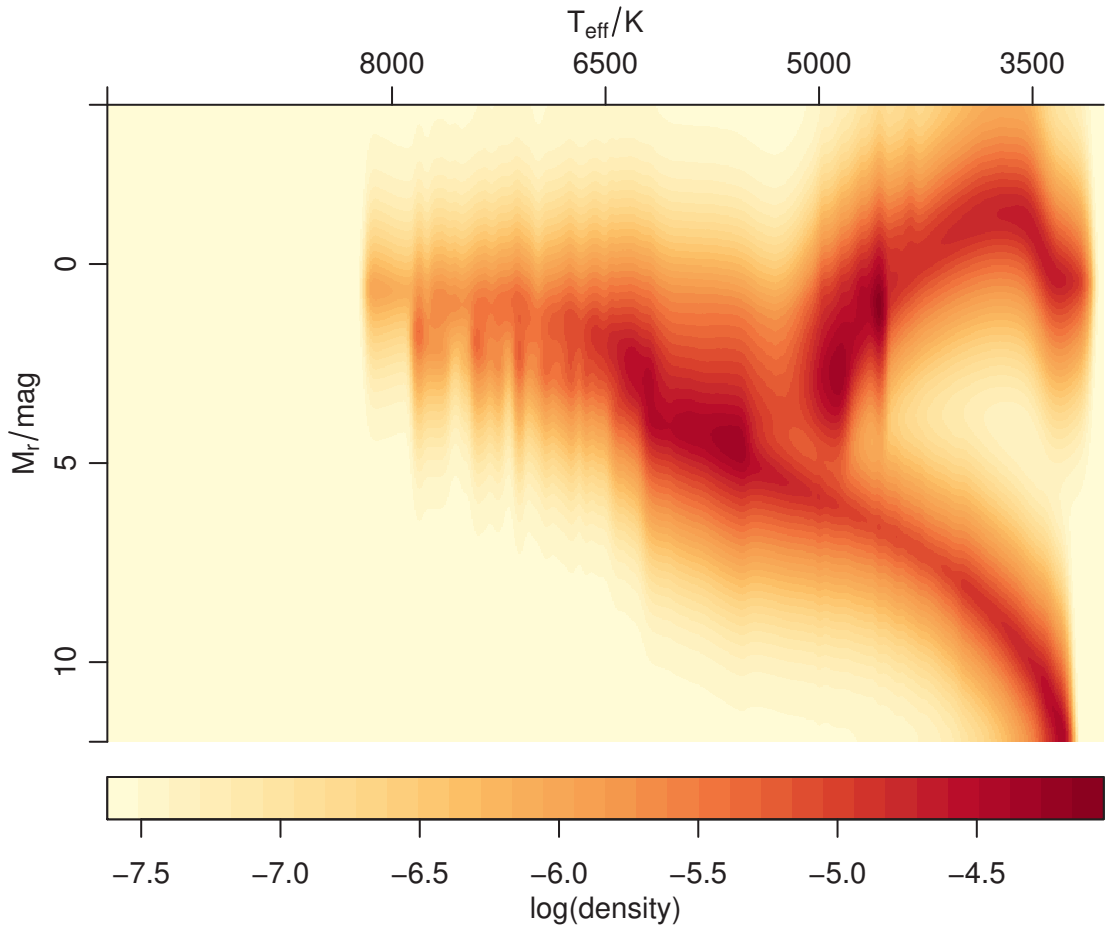

Figure 4.1: Density representation of the HRD, where the integrated probability is normalised to one. The colour scale shows base 10 logarithm of the density, dark red is high number density and yellow is low. Light yellow areas denote regions of the parameter space with initially zero probability. A small offset is added to each point before normalisation to avoid this in the actual computation. In this case the offset is approximately 10^{-4} times the maximum density, resulting in a value of -7.5 in logarithmic density.

less influenced by this. We make this compromise of a simple HRD as we do not wish to introduce yet more dependencies by imposing a complex Galaxy model.

4.2 Forward Model

We build a synthetic forward model based on MARCS model spectra (Gustafsson et al., 2008) in the temperature range $2500 - 10\,000$ K. Based on the bandpass functions of the survey filters, we compute the absolute photometry for stars with simulated extinction. The zero points are computed in the AB system. We convert the Spitzer IRAC data (which is reported in Vega magnitudes) accordingly.

As the synthetic libraries do not model colours of M dwarfs well, we combine

these with empirical stellar loci for the Pan-STARRS1 bands from Tonry et al. (2012) and adapt the synthetic loci at low temperatures ($\approx$ 3000 K). Synthetic and empirical loci match very well for other spectral types.

For each colour we fit a three-dimensional thin-plate spline to its variation over A_0, R_0 and $T_{\rm eff}$. These spline models are used to predict the colour for given trial APs, which are compared with the measured colours via the likelihood in Equation 2.1. To fully model all variations over small parameter changes, we use $\approx$ 8000 synthetic stars and allow the splines to have 1000 degrees of freedom.

Computation

We use a Metropolis-Hastings Markov Chain Monte Carlo (MCMC) routine to sample the parameter space in logarithmic units of the APs. Using the logarithm forces them to remain positive without the use of an explicit prior to this effect. Sufficient convergence is achieved with 10 000 steps each for burn-in and sampling. The sampling steps are of the order of 0.1 dex in all variables. To speed up the computation time, we use a lookup table for all parameters. This has a resolution much better than the model accuracy in order to avoid biasing the results from grid effects.

After inferring parameters for all stars, we remove those with parameters at the grid boundaries, resulting in ranges of $3100 - 9900$ K in $T_{\rm eff}$ and $2.2 - 5.8$ in R_0. This post-processing step removes close to 10 % of the stars. In the available dataset these stars have an indicator flag set to 1 for each affected AP (see Appendix in Hanson et al., 2016, and www.rhanson.de/gpdust).

4.3 Maps

We apply our method to the cross-matched PS1-GLIMPSE photometry to obtain individual AP estimates for all stars individually. To summarise and visualise these results we bin stars with a fixed angular resolution of $7' \times 7'$ in l and b. We present the maps after converting the extinction values to the r-band extinction $A_{\rm r_{PS1}}$ (see Section 3.2 for details of this conversion). To compute the variation in extinction $A_{\rm r}$ (and R_0) along the line-of-sight at any value of distance modulus μ_j we calculate the weighted mean extinction $\langle A_{\rm r} \rangle_j$ and standard deviation Σ_j (and analogously for R_0) for all stars in a single bin which have a distance modulus estimate within one magnitude of our selected position. The recipe for doing so is detailed in Section 2.3.

The mean uncertainties on extinction A_0, extinction parameter R_0, effective temperature $T_{\rm eff}$ and distance modulus μ, based on the widths of the 68% confidence

intervals of the individual stellar parameter estimates, are 0.17 mag, 0.36, 185 K and 2.6 mag, respectively. For each star we obtain an entire PDF over the parameters, from which we compute the confidence intervals. The lower bound of the 68% confidence interval has 16% of the probability below it, whereas the upper bound has 16% of the probability above it. Histograms of the uncertainty distributions are shown in Figure 4.2.

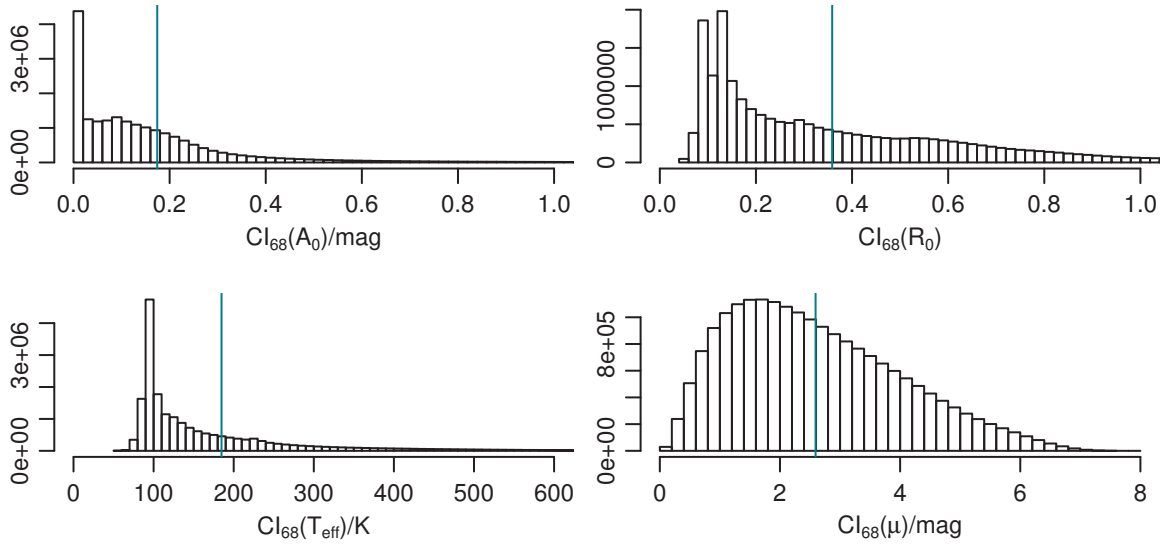

Figure 4.2: Histograms of the widths of the 68% confidence intervals of extinction A_0, extinction parameter R_0, effective temperature $T_{\rm eff}$ and distance modulus μ. The turquoise vertical lines indicate the mean values in each case. For A_0 this corresponds to 0.17 mag, for R_0 it is 0.36 , for $T_{\rm eff}$ it is 185 K and for μ it is 2.6 mag.

In Figure 4.3 we show histograms of the relative uncertainties for the APs for each star (distance modulus is not included, as it is a fractional distance.) These are computed by dividing the width of the 68 per cent confidence intervals by the mean. The mean relative uncertainties are 0.17, 0.09 and 0.04 for extinction, extinction parameter and effective temperature, respectively.

In Figure 4.4 we illustrate the density of stars per pixel for each line of sight. Note that this does not indicate directly how many stars are used at each distance slice. We impose minimal requirements in this case (10 stars per cell). The mean density is nearly 400 stars per $7' \times 7'$ pixel, whereby some pixels have only a few stars (not counting regions not covered by the data set). The maximum is 2 931, the most dense pixels tend to be situated slightly above and below $b = 0$ around the Galactic centre. As expected the density decreases as we move away from the Galactic centre in longitude.

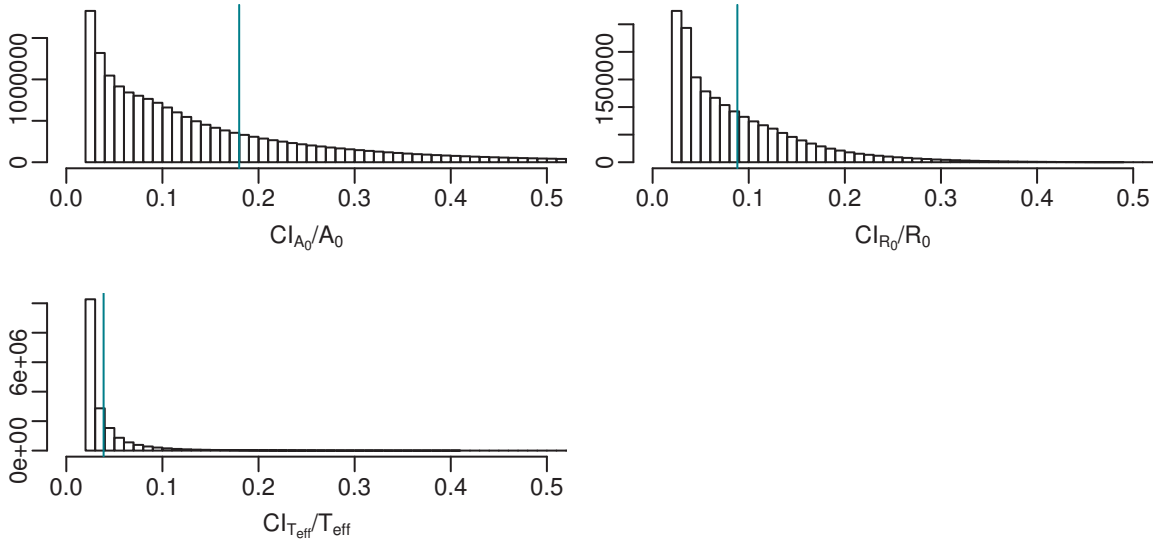

Figure 4.3: Histograms of relative uncertainties as defined by the widths of 68% confidence intervals divided by the mean for A_0, R_0 and $T_{\rm eff}$. The turquoise vertical lines indicate the mean values in each case. For A_0 this corresponds to 0.17, for R_0 it is 0.09 and for $T_{\rm eff}$ it is 0.04.

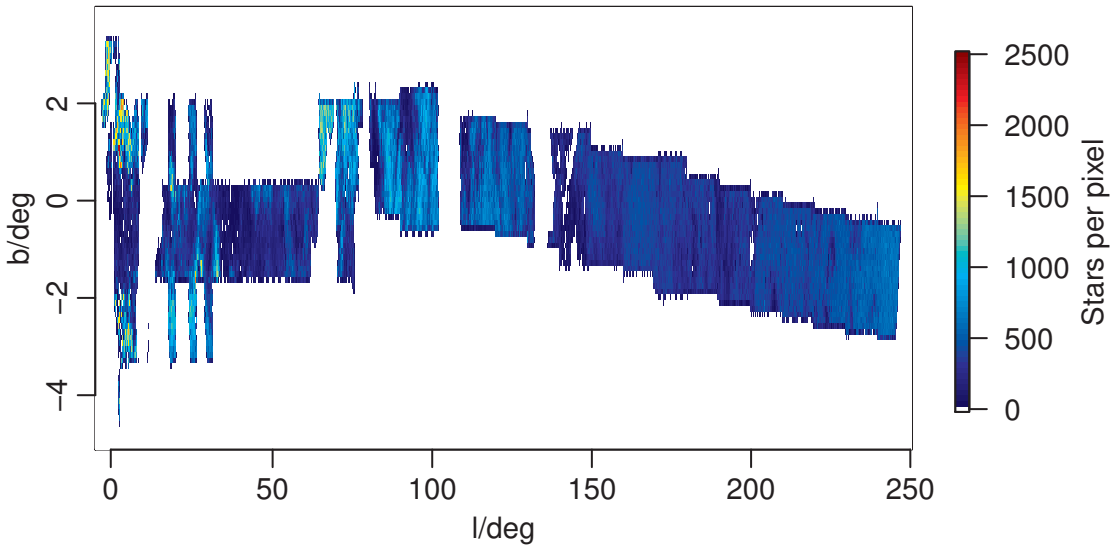

Figure 4.4: Stars per $7' \times 7'$ pixel across the data footprint shown as a colour density scale (white areas are not covered by the data). The mean density is 400 stars per pixel. The scale is limited to 2500 stars per pixel (red), and any pixels with more stars are shown in this colour.

Extinction A_0

Figure 4.5 shows the cumulative line of sight extinction for eight distance slices from $\mu = 6 - 13$ mag in units of r_{PS1} -band extinction as two-dimensional slices of the full map through the Galactic plane. Various structures are visible. In particular the lack of higher extinctions between $l = 100 - 150^{\circ}$ and towards larger distances coincides with the warp in the dust distribution noted by Marshall et al. (2006) and Sale et al. (2014). In Section 4.4 we will analyse in more detail a few particular molecular clouds, which we will also use to validate the overall method.

At closer distances some cells contain insufficient stars to be assigned an extinction estimate and therefore appear white. The colour scale is limited to $A_{\mathrm{r}} \leq 6$ mag; the highest extinction estimate for any pixel is A_{r}= 5.2 mag, although individual stellar estimates may be larger.

Based on the distribution of the standard deviation of individual stellar distances within the three dimensional cells and the standard error of the mean in each cell (per angle and distance, for which a summary is shown in Figure 4.6), we estimate that distances are only reliable from $\mu = 6 - 13$ mag. At closer distances we observe few to no stars due to the bright magnitude limits of the surveys. Beyond the upper limit, distance uncertainties become very large and the distance estimates themselves are no longer useful (recall the relation between distance modulus and distance uncertainties). Those distance slices are not presented here (although the individual stellar distances are available in our published data set).

The predicted uncertainty is illustrated by the distribution of the model-predicted standard errors in the distance modulus and is shown in the left panel of Figure 4.6. For each cell, we compute the standard error of all inferred distance moduli from the fixed cell distance. The average of these is 0.12 mag with a standard deviation of 0.08 mag. The distribution over all cells of the standard deviation of distance moduli within each cell is shown in the right panel of Figure 4.6. The distribution has a mean of 0.56 mag and standard deviation of 0.09 mag. The results indicate that the selected distance slices represent the underlying distance distribution of the stars well.

In Figure 4.7 we show a top-down view of the Galaxy at $b = 0$ in which we average over the five central latitude slices, i.e. from $b = -0.21\,^{\circ}$ to $b = 0.25\,^{\circ}$. As a reference, a distance modulus of 5 mag (10 mag) is equivalent to a distance of 100 pc (1000 pc). Here we can clearly see the 1 mag length scale of the distance modulus slices as well as the expected increase of extinction within a few kpc towards the Galactic centre at the top of the figure. As the measured extinction in neighbouring

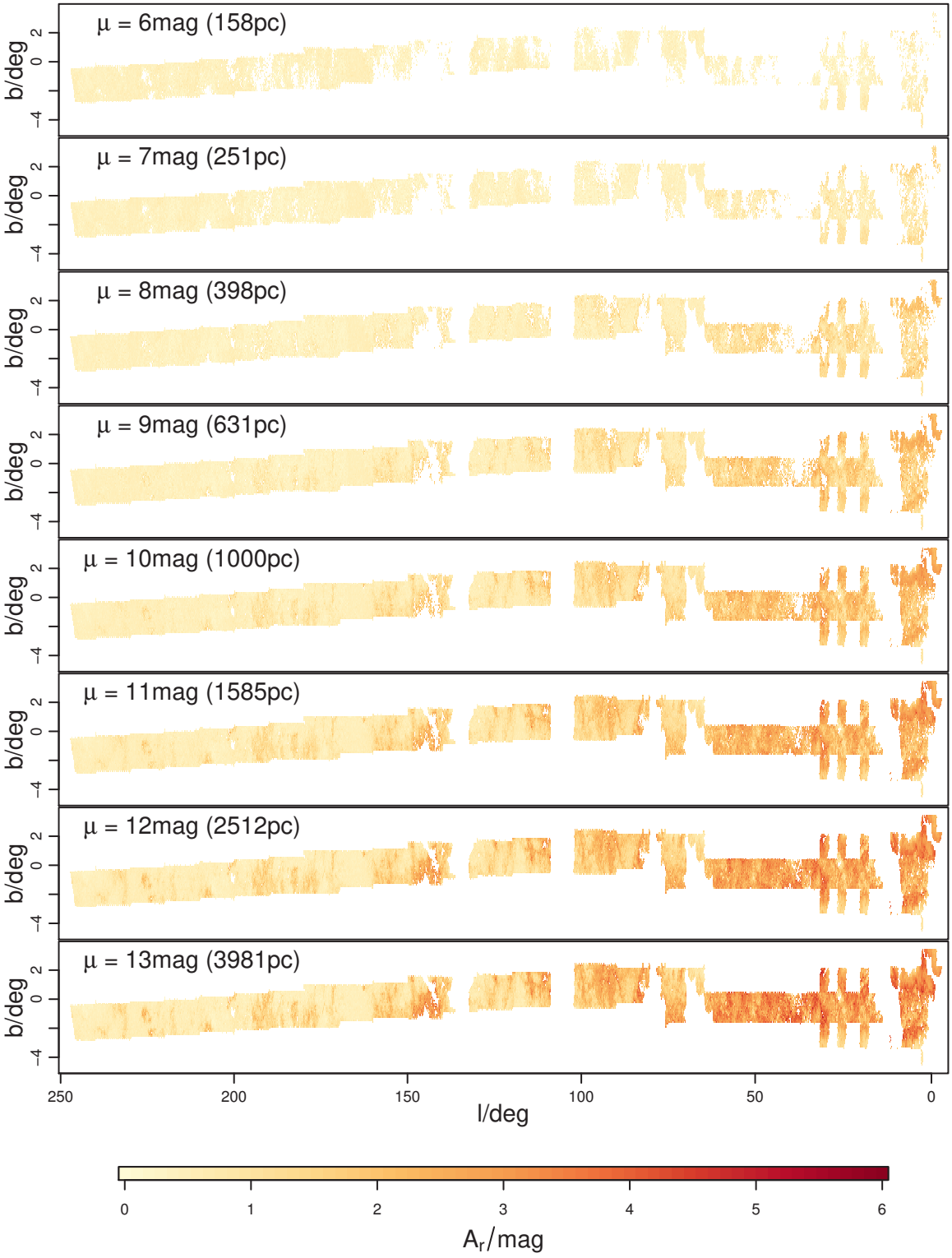

Figure 4.5: Cumulative line of sight extinction at distance moduli from $\mu = 6 - 13$ mag in r_{PS1} -band. White regions are either not covered by the data footprint or (particularly at closer distances) do not contain a sufficient number of stars.

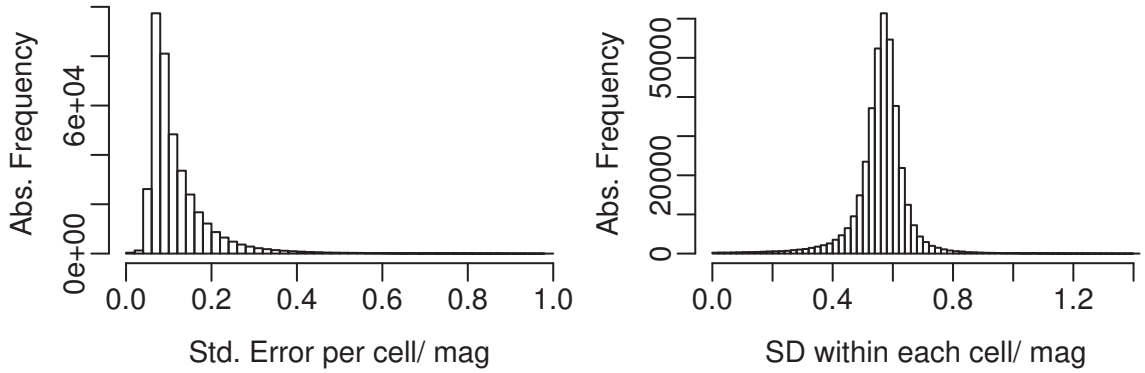

Figure 4.6: Left panel: Histogram of the predicted uncertainty of distance modulus. The standard error is computed for each cell, using the differences of all inferred distance moduli and the fixed cell distance. The mean of this distribution is 0.12 mag. Right panel: Distribution over all cells of the standard deviation of distance moduli within each cell. This distribution has a mean of 0.56 mag. In both panels, cells with distance moduli $\mu = 6 - 13$ mag are included.

cells are only correlated in the radial direction, but not in longitude (or latitude), many discontinuities can be seen.

Extinction parameter R_0

We not only infer extinction A_0 but also the extinction parameter R_0. In Figure 4.8 we show this parameter in slices of distance modulus, analogously to Figure 4.5. It is clear that variations here follow those in extinction. Although there is an indication that in some regions with higher extinction R_0 increases above the mean of $\langle R_0 \rangle = 4.1 \pm 0.27$, we do not detect a global correlation between the two parameters. Only for the two closest distance slices and for low extinctions ($A_r < 0.5$ mag), is there an inkling that R_0 increases with A_0. Whilst we trust the variations of R_0 we measure, we are less certain about the absolute values. This again has to do with model uncertainties and parameter degeneracies that we are unable to remove. Both extinction and the extinction parameter are cumulative along the line of sight to any given distance. All the dust along the line of sight contributes to any individual estimate. Because of this, correlations between these two cumulative parameters are harder to see: at larger distances, the length scale over which the dust properties are averaged increases. For both the A_0 and R_0 estimates we use only stars in a limited distance range around the specified distance.

Our results show that the extinction law is not universal. This has previously been asserted by other authors, such as Gao, Jiang & Li (2009) and Chen et al. (2013) who also look at the variation in large regions of the Galaxy.

The estimates of R_0 for individual stars have, on average, an uncertainty of about

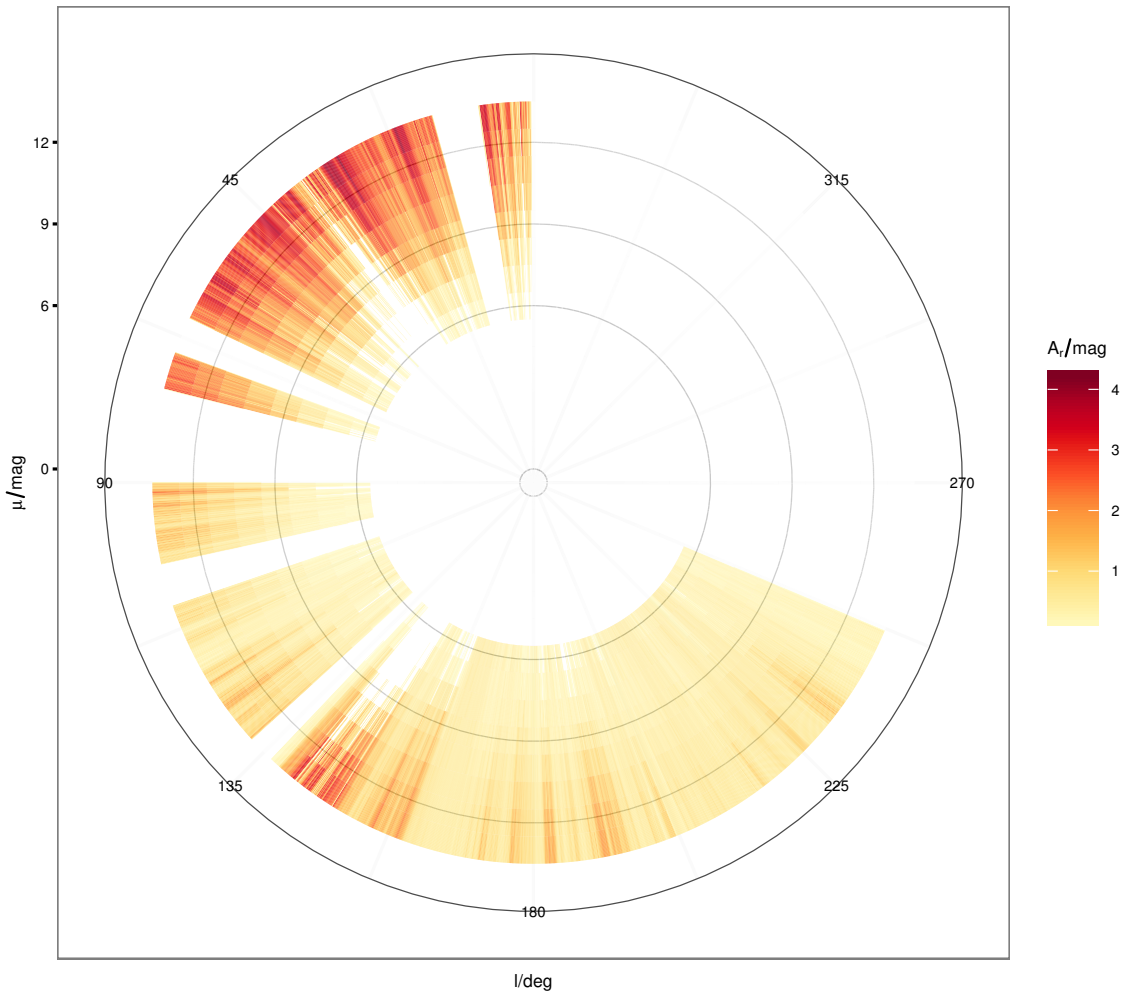

Figure 4.7: Projected extinction map with vertical extent of the Galactic plane from $b = -0.21\,^{\circ} - +0.25\,^{\circ}$ in which the five central latitude slices are averaged. The Sun is at the centre of the diagram. The distance moduli on the left edge refer to the radii of the circles. The physical distances of the indicated distance moduli 6, 9 and 12 mag are 158, 631 and 2512 pc, respectively.

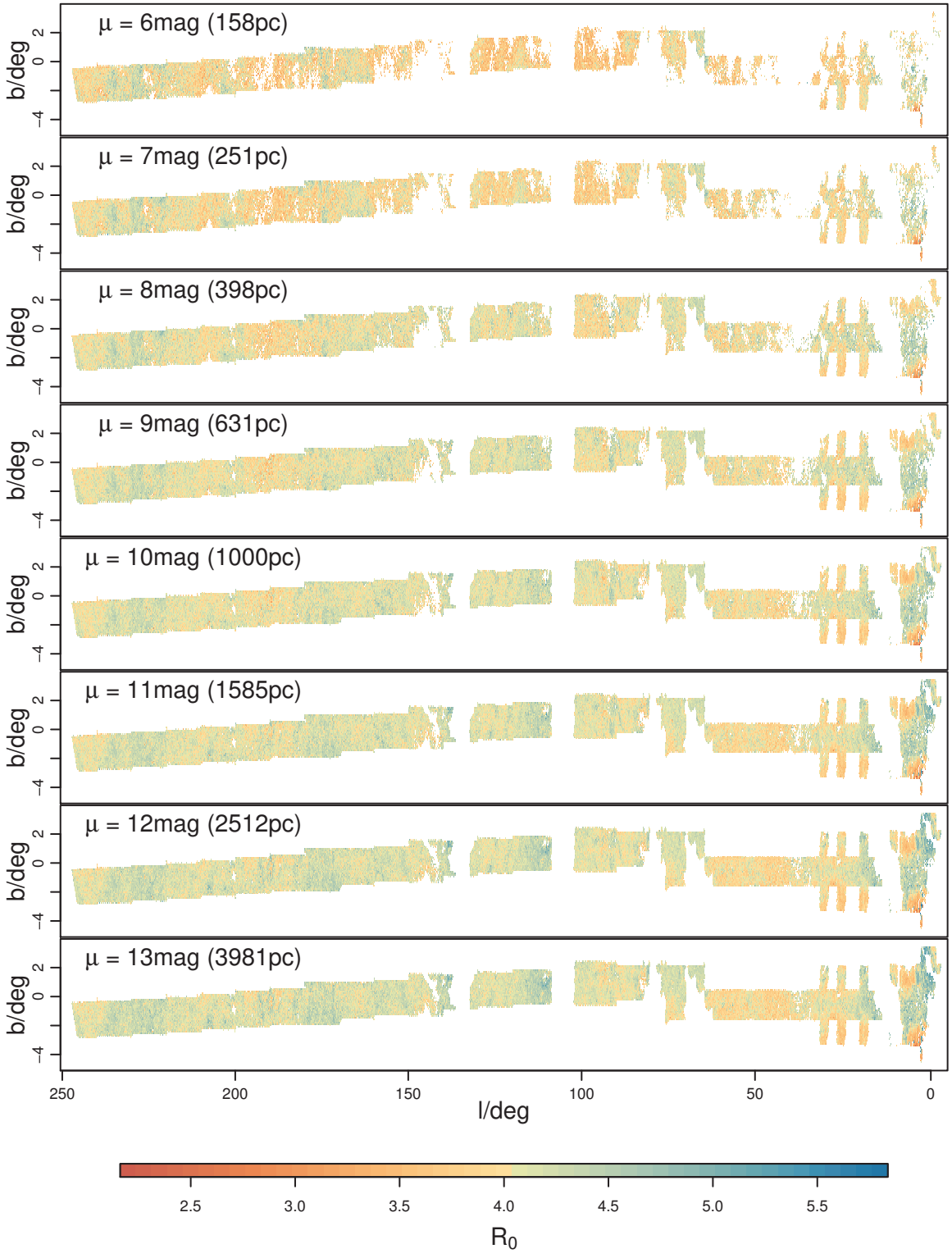

Figure 4.8: Extinction parameter R_0 at distance moduli from $\mu = 6 - 13$ mag, computed according to Equation 2.14 (weighted mean). Again, white regions are either not covered by the data footprint or do not contain enough stars to be assigned a parameter estimate.

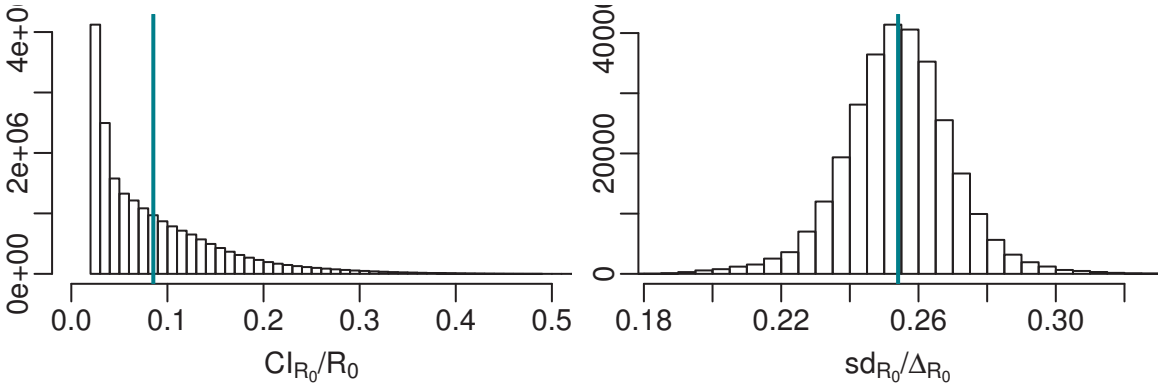

Figure 4.9: Left panel: Relative uncertainty of the R_0 estimates for all individual stars. The value denotes 68 per cent confidence interval over the mean inferred parameter. The blue line indicates the mean of the distribution at 0.085. Right panel: Histogram of the standard deviation relative to the range of R_0 estimates of all cells. The mean of 0.25 is indicated by the blue line.

10 %, as characterised by the ratio of the width of the confidence interval to parameter estimate. This is shown in the left panel of Figure 4.9 as a histogram of all stars. The right panel illustrates the accuracy of the average R_0 estimates from Figure 4.8. We compute the ratio of the standard deviation to the range of R_0 for the stars contained in each cell. This average is 0.25 and indicates that for any individual cell the mean R_0 estimate is well constrained, despite possible variations arising from the fact that APs are inferred for all stars individually.

Zasowski et al. (2009) find that the inner fields of the Galaxy correspond to a larger R_0, whereas outer fields tend to have a lower value. We also find this, as exemplarily shown in Figure 4.10 where we plot the average extinction parameter, $\langle R_0 \rangle$ over several cells as a function of distance modulus for two different lines of sight. The first (left panel) is centred on $l = 0.5^\circ, b = 0$ towards the Galactic centre. The second (right panel) is centred on $l = 47.2^\circ, b = -0.5^\circ$. In both cases we average over approximately half a degree in l and b, corresponding to 5 pixels in each direction at our resolution. We immediately see that the inner profile increases towards the Galactic centre, above the average of 4.1 for our data, an effect that is also seen by Gontcharov (2012b).

The profile for the outer field, which we expect to look through more diffuse dust, remains basically flat at a value below the global average. The mean extinction parameter for this line of sight has a value of 3.9 ± 0.37, very close to the value of 3.8 ± 0.20 we find in Section 3.5 for regions towards the Galactic poles.

Our results for R_0 suggest a higher value for the diffuse interstellar medium than previous studies indicate. Mörtsell (2013) uses quasar data towards the Galactic poles to find $R_V \approx 3$ with a relative uncertainty of 10 %. Savage & Mathis (1979)

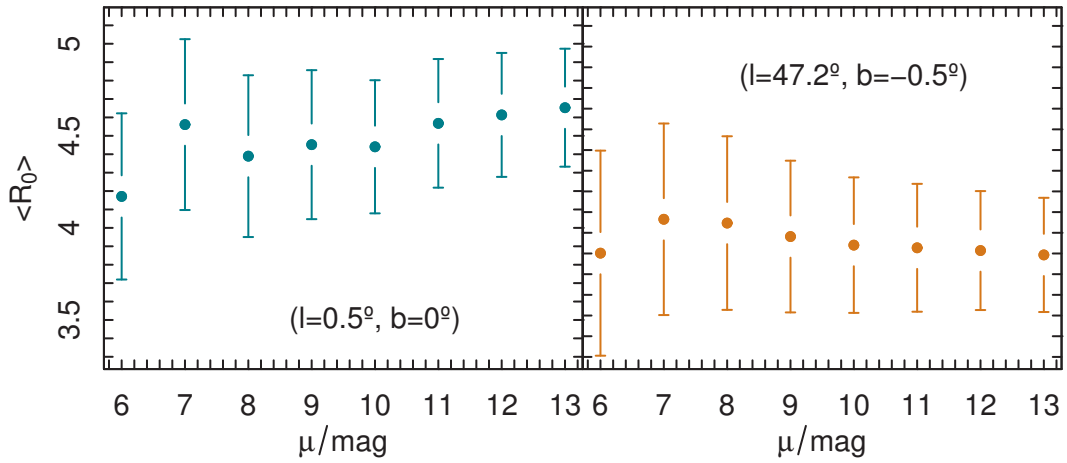

Figure 4.10: Distance modulus versus mean extinction parameter of cells within half a degree, centred at $l = 0.5°, b = 0$ (left panel) and $l = 47.2°, b = -0.5°$ (right panel).

obtain a value of 3.1 with a similar uncertainty. However, Jones, West & Foster (2011) find a median value of 3.38 with at median uncertainty of 0.42 after fitting SDSS spectra of M dwarfs within 1 kpc of the Sun. Their resulting distribution is incompatible with a Gaussian with a width of $\sigma = 0.42$ centered at 3.1.

We have no reason not to believe our results: we find no systematic errors in the data that could, for example, arise from unexpected correlations between R_0 and A_0 and/or $T_{\rm eff}$ and thus affect the parameter inference. This is clear from Chapter 3 where the extinction results for Galactic pole regions are not strongly affected by the inclusion of R_0 as an inferred parameter.

4.4 Validation

To validate our results, in particular the relatively uncertain distances, we compare some of our lines-of-sight with distance estimates to molecular clouds in Schlafly et al. (2014b), who use Pan-STARRS1 photometry to measure and model distances to high statistical accuracy. From Table 1 in that work we select the clouds whose coordinates lie within our survey limits. These are *CMa OB1* with three individual measurements at $(l, b) = (224.5°, -0.2°)$, $(222.9°, -1.9°)$ and $(225.0°, -0.2°)$, as well as *Maddalena* at $(l, b) = (217.1°, 0.4°)$. The reported distances to these clouds are 1369^{+64}_{-56}, 1561^{+79}_{-77}, 1398^{+63}_{-59} and 2280^{+71}_{-66} pc, respectively, which in distance modulus are $10.68^{+0.10}_{-0.09}$, $10.97^{+0.11}_{-0.11}$, $10.73^{+0.10}_{-0.09}$ and $11.79^{+0.07}_{-0.06}$ mag.

In Figure 4.11 we show the extinction $A_{\rm r}$ (turquoise circles) and extinction param-

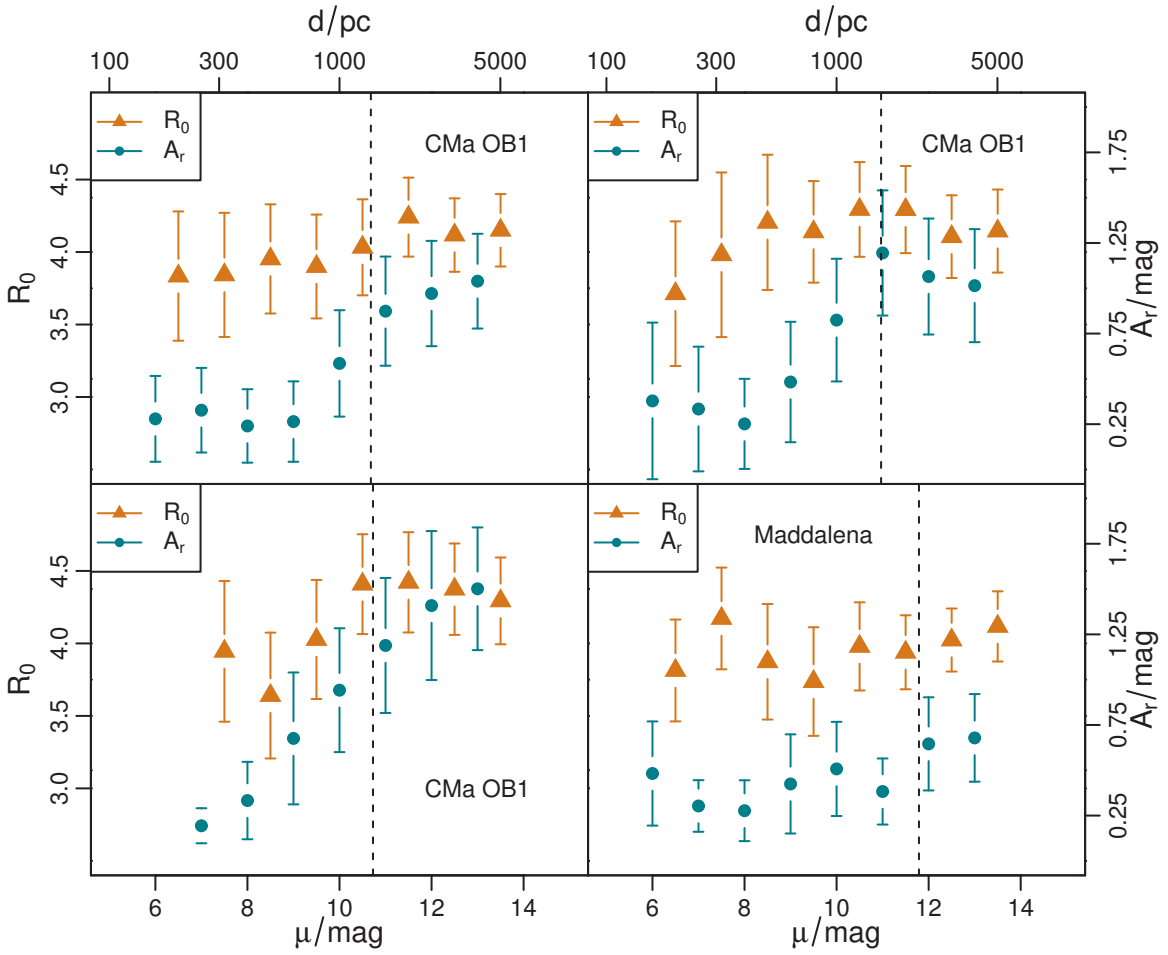

Figure 4.11: Cumulative extinction A_r (turquoise circles) and extinction parameter R_0 (orange triangles) as functions of distance modulus μ towards four molecular clouds. See text for the coordinates of the fields. The error bars are computed using Equation 2.14. The dashed vertical lines indicate the distances reported in Schlafly et al. (2014b).

eter R_0 (orange triangles) as a function of distance modulus for our data using stars within 7 $'$ of the coordinates given above. The dashed lines indicate the Schlafly et al. (2014b) distances of the clouds. The mean and error bars are computed according to Equation 2.14.

Similarly, in Figure 4.12 we show differential profiles of A_r (turquoise circles) and R_0 (orange triangles), where the values quantify the change in both parameters in steps of $\Delta\mu = 1$ mag.

Despite not explicitly measuring distances to individual large-scale objects, it is clear that our method manages to capture real features in the extinction distribution. We see that the total extinction A_r generally increases around the inferred positions of the clouds, indicating an increase of the underlying dust density around that position. This feature is more pronounced in the two top panels, although the clouds could be responsible for the more gradual increase in extinction in the other two

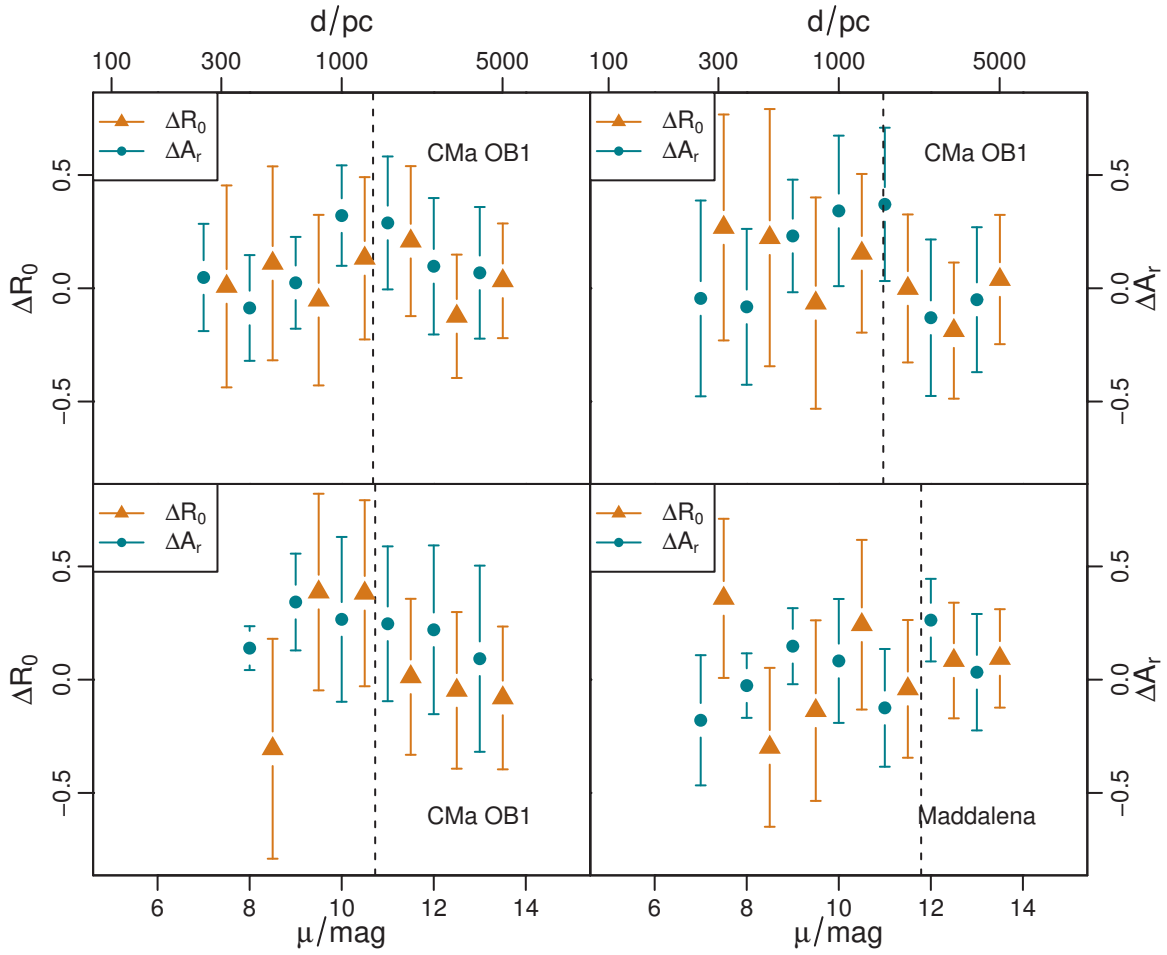

Figure 4.12: Differential extinction A_r (turquoise circles) and extinction parameter R_0 (orange triangles) as functions of distance modulus μ towards four molecular clouds. The differentials are computed between distance modulus steps of 1 mag. See text for the coordinates of the fields. The dashed vertical lines indicate the distances reported in Schlafly et al. (2014b).

panels as well. This is highlighted in Figure 4.12, where the increase in extinction can be seen more clearly in the top two panels. The interpretation of the bottom two panels in both figures is less clear cut, despite there being marginal changes in A_r and R_0 around the literature distances of the clouds. However, the spread in A_r (and R_0) is generally quite large, and the distances are so uncertain that we are not necessarily confident of having detected the clouds. In all four panels the mean extinction decreases slightly again beyond $\mu = 13$ mag. We do not trust values beyond this distance (see Section 4.3 and Figure 4.6 for details), as we do not expect to detect many stars at large distances due to the faint magnitude limits of the input catalogues and the resulting selection effects.

The value of the extinction parameter R_0 also appears to increase in sync with the increase of extinction, although the magnitude of variation tends to be within the

range of uncertainty. Nevertheless, the overall picture is one where there are dense dust clouds which cause the cumulative line of sight extinction to increase above some foreground value. This suggests that the inferred parameters we obtain with our method are trustworthy and physically plausible, at least on a relative scale.

To further probe this, we compare our results with those of Berry et al. (2012, B12) who combine SDSS and 2MASS data to calculate A_r and R_V using a straightforward fit to stellar templates. We take a subset of the common survey area from $l = 49-51^\circ$ and $b = -1^\circ$ to $b = 1^\circ$ and compute a 3D dust map based on their results using Equation 2.14. Due to the different sensitivities and depths of the surveys we use the distance slices at $\mu = 9$ mag and $\mu = 10$ mag for further comparison, as other distances have many empty cells in one or both data sets. Qualitatively we find similar behaviour and features in the A_r extinction map, although the average extinction in our data is $\langle A_r \rangle = 1.63 \pm 0.44$ mag, whereas the B12 data suggests an average of $\langle A_r \rangle_{B12} = 2.41 \pm 0.46$ mag. The standard deviations are similar in both cases. For the extinction parameter R_0 we obtain an average value of $\langle R_0 \rangle = 4.04$ with a standard deviation of 0.20, whereas $\langle R_0 \rangle_{B12} = 3.02$ (standard deviation is also 0.20.) These differences are also reflected when individually cross-matching the stars in the common footprint. The differences (this work minus B12) on average are -0.18 mag for *r*-band extinction and 0.79 for the extinction parameter. As expected, due to B12's work strongly favouring a value of 3.1 for a large fraction of stars, we measure a standard deviation of 1.26 in R_0 between the two datasets and see that the differences increase as our R_0 estimates increase. Our extinction results agree reasonably well with the previous work. However, we seem to have systematically higher values of R_0, which, as discussed previously, may result from fixing the metallicity in the HRD and/or using synthetic spectral templates. Nevertheless, we are much more confident in our relative values of the extinction parameter (and A_0), as our model assumptions have much less effect on our ability to measure these.

To exclude the possibility that requiring NIR data could be a cause for the aforementioned differences, we select a random sample of 10 000 stars in the same region based purely on their presence in the PS1 data-set. We require no counterpart in the GLIMPSE surveys. Comparing the average widths of the 68% confidence intervals and the average relative uncertainties with results that include GLIMPSE data, as shown in Table 4.1, we find that including the 3.6 μm photometry significantly improves the precision of the inferred APs. Especially the R_0 estimates benefit from the additional band, reducing the average width of the confidence intervals from 0.30 to 0.22.

	PS1 + 3.6 μm		PS1 only	
AP	⟨CI/AP⟩	⟨CI⟩	⟨CI/AP⟩	⟨CI⟩
A_0	0.12	0.23 mag	0.14	0.24 mag
$T_{\rm eff}$	0.04	217 K	0.04	206 K
R_0	0.05	0.22	0.07	0.30
μ	-	2.1 mag	-	2.6 mag

Table 4.1: Average widths of 68% confidence intervals ⟨CI⟩ and mean relative uncertainties ⟨CI/AP⟩ for extinction A_0, effective temperature $T_{\rm eff}$, extinction parameter R_0 and distance modulus μ in the cases of including 3.6 μm photometry (left) and using only PS1 bands (right).

As this sample generally lacks GLIMPSE counterparts, we cannot measure differences in the AP estimates for all individual stars. To nevertheless ensure that we have not introduced (or removed) any systematic effects on the inference, we compare the inferred APs for stars from the initial cross-matched sample when including and excluding 3.6 μm photometry. In this situation the mean differences (including minus excluding 3.6 μm data) of the APs for these stars are only 0.07 mag, -45 K, 0.01 and 0.33 mag for A_0, $T_{\rm eff}$, R_0 and μ, respectively. This indicates that the inclusion of the NIR band does not introduce systematic differences, but actually improves the inference.

4.5 Summary

We have presented three dimensional maps in cumulative line of sight extinction A_0 and extinction parameter R_0 which are constructed using a Bayesian method. This method is general and not bound to specific photometric systems. It is based on work by Bailer-Jones (2011) that was expanded in Hanson & Bailer-Jones (2014). We take advantage of the physical understanding of stellar evolution that is encapsulated in the Hertzsprung-Russell Diagram. Using photometric measurements of 19 885 031 stars with data from the cross-matched Pan-STARRS1 and Spitzer GLIMPSE surveys (six bands in total), we infer extinction A_0, extinction parameter R_0, effective temperature $T_{\rm eff}$ and distance modulus μ to all stars individually. We achieve mean relative uncertainties of 0.17, 0.09, 0.04 and 0.18 for extinction, extinction parameter, effective temperature and distance modulus, respectively, whilst obtaining average uncertainties of 0.17 mag, 0.36, 185 K and 2.6 mag for the four parameters. We emphasise that while we believe the R_0 variations we measure, we are less confident in the absolute values, based on comparisons with other work.

Using these inferred parameters we compute the estimated total extinction to ar-

bitrary distances and estimates of the extinction parameter, as formulated in Equation 2.14. The angular stellar density allows us achieve a reliable resolution of $7' \times 7'$ in latitude and longitude. We select steps of 1 mag in distance modulus. From the distribution of distance estimates within all three-dimensional cells, we estimate that the reported extinction map is reliable from $\mu = 6 - 13$ mag. At closer distances we have too few stars for trustworthy estimates due to the bright magnitude limits of both surveys. Beyond that distance range, individual estimates become too uncertain. We do not expect many stars beyond that distance due to the faint magnitude limits, so we do not report values outside this range. We find that the extinction law varies with each line of sight and along the line of sight, supporting previous works which contend that using a single value to parametrize extinction is insufficient to properly model the three dimensional dust distribution in the Galaxy. The data are made available via www.rhanson.de/gpdust.

As previously discussed in Section 3.6, the key limitation at this stage is the distance inference, which is limited by photometric errors and intrinsic model degeneracies. Furthermore, on the account of our use of stellar models to estimate stellar effective temperatures, there are likely to be systematic uncertainties in our estimates of A_0 and R_0. These enter through the assumption of 'true' model temperatures, the use of an HRD prior and lack of metallicity variations (again, see Section 3.4). Furthermore, our extinction estimates for individual lines of sight do not account for correlations in angular dimensions. That is, neighbouring lines of sight are solved for independently. This clearly does not mirror reality, where the extinction estimates for stars that are close in space (and whose photons are affected by the same dust structures) should be strongly correlated, whereas those of stars that have a large separation should be less so. Theoretically, due to the finite cross-sectional area of a line-of-sight, a more distant star could show less extinction. This shortcoming is now starting to be addressed. Sale & Magorrian (2014) introduce a method based on Gaussian random fields and a model of interstellar turbulence, which addresses the discontinuities we currently see in most extinction maps. Lallement et al. (2014) use an inversion method with spatial correlation kernels that attempts to reconstruct structures of the ISM in a more realistic manner. In Chapter 5 I will detail a new method to self-consistently infer the three-dimensional dust distribution in a region using individual stellar AP measurements.

Combining current large area photometric surveys, such as those employed here, with parallax measurements from *Gaia* will enable us to construct accurate 3D maps of stars in the Galaxy. Including stellar parameter estimates from future data releases by the Data Processing and Analysis Consortium (*DPAC*), as summarised in

Bailer-Jones et al. (2013), will significantly increase our capabilities of reconstructing the full three dimensional distribution of dust.

Part 2

Non-parametric Dust Mapping

5

Model for Non-parametric Dust Density

This chapter is adapted from Rezaei Kh. et al. (2017) and presents a non-parametric model for inferring the three-dimensional distribution of dust density in the Galaxy. This model is based on Gaussian processes and provides and analytic solution to the problem.

The chapter is organised as follows. First, I give a brief introduction to Gaussian processes, as they are vital to understanding the model. I then formalise it by deriving its analytical form. I demonstrate the performance using mock data and explain the influence of the model hyperparameters. Finally I illustrate the computational complexity and suggest potential improvements and additions.

5.1 Introduction to Gaussian Processes

At its core, a Gaussian process (GP) is a type of probability distribution of functions. Although GPs can be used in principle for most types of machine learning problems (see, e.g., MacKay, 2003, for a general overview), for our purpose we focus on the regression application. We follow typical approaches, such as those in Gibbs (1997) and Gibbs & MacKay (1997), and refer to Rasmussen & Williams (2006) and Snelson (2007) for more in-depth introductions.

In this situation we have a data set $\mathcal{D}$ with N input vectors $\{\boldsymbol{x}_N\} = \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ and outputs $\boldsymbol{y}_N = y_1, y_2, \ldots, y_N$. The outputs are assumed to be observed with noise from an underlying function $f(\boldsymbol{x})$ that is typically unknown. We therefore intend to estimate $f(\boldsymbol{x})$ from the data $\mathcal{D}$. More accurately, we infer a probability distribution

over likely functions. This then allows us to compute the probability of the quantity in question at a new position, such that we solve $P(y_{N+1}|\mathcal{D}, \boldsymbol{x}_{N+1})$. Here, y_{N+1} is a single datum, whereas $\boldsymbol{y}_{N+1}$ is the vector of the $N+1$ outputs, mapped by the set of input vectors $\{\boldsymbol{x}_{N+1}\}$.

Gaussian process model

The Gaussian process model is a method to solve this equation by assuming that $\{\boldsymbol{x}_N\}$ has a joint Gaussian distribution

$$P(\boldsymbol{y}_N|\{\boldsymbol{x}_N\}, \boldsymbol{C}_N, \boldsymbol{\mu}) = \frac{1}{Z} \exp\left(-\frac{1}{2}(\boldsymbol{y}_N - \boldsymbol{\mu})^\mathsf{T} \boldsymbol{C}_N^{-1} (\boldsymbol{y}_N - \boldsymbol{\mu})\right) , \tag{5.1}$$

where $\boldsymbol{C}_N$ is the $N \times N$ covariance matrix and $\boldsymbol{\mu}$ is the mean vector of the distribution. Z is a normalisation constant which we ignore for now. For simplicity we consider the $\boldsymbol{\mu} = 0$ case, which makes further notation more straight forward. Equation 5.1 then becomes

$$P(\boldsymbol{y}_N|\{\boldsymbol{x}_N\}, \boldsymbol{C}_N) = \frac{1}{Z'} \exp\left(-\frac{1}{2}\boldsymbol{y}_N^\mathsf{T} \boldsymbol{C}_N^{-1} \boldsymbol{y}_N\right) . \tag{5.2}$$

Returning to the original goal of predicting y_{N+1} at a new position $\boldsymbol{x}_{N+1}$, we can write the joint distribution of $\boldsymbol{y}_{N+1}$ as

$$P(\boldsymbol{y}_{N+1}|\{\boldsymbol{x}_N\}, \boldsymbol{x}_{N+1}, \boldsymbol{C}_{N+1}) = \frac{1}{Z''} \exp\left(-\frac{1}{2}\boldsymbol{y}_{N+1}^\mathsf{T} \boldsymbol{C}_{N+1}^{-1} \boldsymbol{y}_{N+1}\right) . \tag{5.3}$$

The full probability distribution for y_{N+1} can thus be written as

$$\begin{aligned} P(y_{N+1}|\boldsymbol{y}_N, \{\boldsymbol{x}_N\}, \boldsymbol{x}_{N+1}, \boldsymbol{C}_{N+1}) &= \frac{P(\boldsymbol{y}_{N+1}|\{\boldsymbol{x}_N\}, \boldsymbol{x}_{N+1}, \boldsymbol{C}_{N+1})}{P(\boldsymbol{y}_N|\{\boldsymbol{x}_N\}, \boldsymbol{C}_N)} \\ &= \frac{Z'}{Z''} \exp\left[-\frac{1}{2}\left(\boldsymbol{y}_{N+1}^\mathsf{T} \boldsymbol{C}_{N+1}^{-1} \boldsymbol{y}_{N+1} - \boldsymbol{y}_N^\mathsf{T} \boldsymbol{C}_N^{-1} \boldsymbol{y}_N\right)\right] . \end{aligned} \tag{5.4}$$

If we now express $\boldsymbol{C}_{N+1}$ in terms of $\boldsymbol{C}_N$ as

$$\boldsymbol{C}_{N+1} = \begin{bmatrix} \begin{bmatrix} & & \\ & \boldsymbol{C}_N & \\ & & \end{bmatrix} & \begin{bmatrix} \\ \boldsymbol{k} \\ \\ \end{bmatrix} \\ \begin{bmatrix} & \boldsymbol{k}^{\mathsf{T}} & \end{bmatrix} & \begin{bmatrix} k \end{bmatrix} \end{bmatrix} , \tag{5.5}$$

where the scalar k is just the covariance of the new position vector with itself $k = C(\boldsymbol{x}_{N+1}, \boldsymbol{x}_{N+1})$ and $\boldsymbol{k}$ is the vector of the covariances of all N original positions with the new position, $\boldsymbol{k} = [C(\boldsymbol{x}_1, \boldsymbol{x}_{N+1}), C(\boldsymbol{x}_2, \boldsymbol{x}_{N+1}), \ldots, C(\boldsymbol{x}_N, \boldsymbol{x}_{N+1})]$. We can rewrite Equation 5.4 as

$$P(y_{N+1}|\boldsymbol{y}_N, \{\boldsymbol{x}_N\}, \boldsymbol{x}_{N+1}, \boldsymbol{C}_{N+1}) = \frac{Z'}{Z''} \exp\left[-\frac{1}{2}\begin{bmatrix}\boldsymbol{y}_N & y_{N+1}\end{bmatrix} \boldsymbol{C}_{N+1}^{-1} \begin{bmatrix}\boldsymbol{y}_N \\ y_{N+1}\end{bmatrix}\right] . \tag{5.6}$$

Now, we express $\boldsymbol{C}_{N+1}^{-1}$ in terms of $\boldsymbol{C}_N$ and $\boldsymbol{C}_N^{-1}$

$$\boldsymbol{C}_{N+1}^{-1} = \begin{bmatrix} \boldsymbol{M} & \boldsymbol{m} \\ \boldsymbol{m}^{\mathsf{T}} & \mu \end{bmatrix} , \tag{5.7}$$

where

$$\begin{aligned} \mu &= \left(k - \boldsymbol{k}^{\mathsf{T}} \boldsymbol{C}_N^{-1} \boldsymbol{k}\right)^{-1} \\ \boldsymbol{m} &= -\mu \boldsymbol{C}_N^{-1} \boldsymbol{k} \\ \boldsymbol{M} &= \boldsymbol{C}_N^{-1} + \frac{1}{\mu} \boldsymbol{m}\boldsymbol{m}^{\mathsf{T}} . \end{aligned} \tag{5.8}$$

Substituting these factors into Equation 5.6 and simplifying, we find

$$P(y_{N+1}|\boldsymbol{y}_N, \{\boldsymbol{x}_N\}, \boldsymbol{x}_{N+1}, \boldsymbol{C}_{N+1}) = \frac{1}{Z^*} \exp\left(-\frac{(y_{N+1} - \hat{y}_{N+1})^2}{2\sigma_{\hat{y}_{N+1}}^2}\right) , \tag{5.9}$$

where $\hat{y}_{N+1} = \boldsymbol{k}^{\mathsf{T}} \boldsymbol{C}_N^{-1} \boldsymbol{y}_N$ and $\sigma_{\hat{y}_{N+1}}^2 = k - \boldsymbol{k}^{\mathsf{T}} \boldsymbol{C}_N^{-1} \boldsymbol{y}_N$, and we have folded all the constants into Z^*. Thus, the maximum probability is given by the mean $\hat{y}_{N+1}$ and variance $\sigma_{\hat{y}_{N+1}}^2$.

Covariance functions

In order to effectively use a Gaussian process to solve a regression problem, a suitable covariance function $\boldsymbol{C}$ must be defined (as well as the mean μ which we have assumed to be zero). In general the elements of the covariance matrix are $\boldsymbol{C}_{ij} = C(\boldsymbol{x}_i, \boldsymbol{x}_j)$, where the function C describes correlations between different points

$$C(x, x') = \mathcal{E}[f(x)f(x')] \,, \tag{5.10}$$

where $\mathcal{E}$ denotes the expectation value of some function. Constructing the covariance matrix $\boldsymbol{C}$ using a covariance function ensures that the matrix is symmetric and positive semidefinite.

There are many different choices for covariance function (see, e.g., Gneiting, 2002; Snelson, 2007; Rasmussen & Williams, 2006, for examples). Depending on the choice, one is able to vary the smoothness, amplitude and effective lengthscales of the Gaussian process prior. Herein lies the appeal of Gaussian processes – instead of trying to define and parametrize the unknown function f explicitly, we can instead define the correlation between different values of it.

5.2 Method

As discussed, the aim is to determine the three-dimensional distribution of interstellar dust given measurements of line of sight extinction towards individual stars. Hence, we wish to find the probability distribution over the dust density at any point in space, irrespective of the positions of the stars used to measure the line of sight extinction. It is important to emphasise that we are not attempting to parametrize a specific dust model – on the contrary, the use of a Gaussian process enables us to avoid having to define a (invariably simplified) parametric model. In effect, we are able to probe a much wider range of functional variations than a parametric model.

To develop and derive the model we define the problem we are trying to solve more explicitly. Defining $\rho(\boldsymbol{r})$ as the dust density at position $\boldsymbol{r}$ measured from the observer, we can model the attenuation of stellar light caused by this dust for a star at position $\boldsymbol{r}_n$ as

$$f_n \propto \int_0^{r_n} \rho(\boldsymbol{r}) \, \mathrm{d}r \,, \tag{5.11}$$

where $r = |\boldsymbol{r}|$ is the length of the distance vector. The central goal is to invert the above to get $\rho(\boldsymbol{r})$ for an arbitrary point in space given measurements of the attenuation towards multiple stars. If we adopt a parametric form for $\rho(\boldsymbol{r})$ then this is straightforward, but the result would be highly limited by the form adopted. Instead, we construct a non-parametric model by dividing the imagined pencil beam along the line of sight towards each star into several dust cells, as illustrated in Figure 5.1

for three stars.

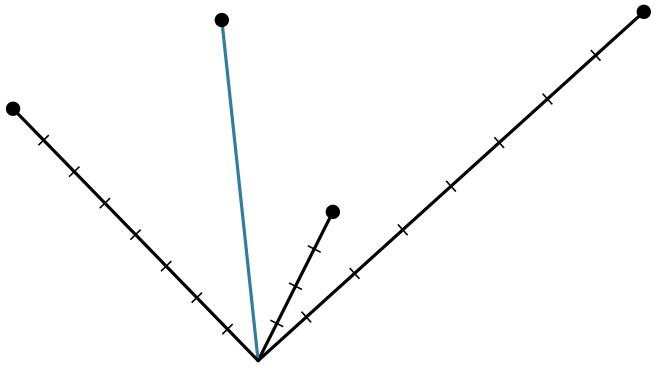

Figure 5.1: Schematic of the principle of splitting each line of sight into multiple dust cells (black lines) for 3 stars. (Here we show cells of constant size.) The total number of dust cells (towards all stars) is J, and the length of dust cell j towards star n is denoted $g_{n,j}$. The blue line illustrates a new point at which we attempt to find the dust density.

Denoting the average dust density in cell j towards star n as $\rho_{n,j}$, the integral in Equation 5.11 can then be replaced by a sum

$$f_n = \sum_j g_{n,j}\rho_{n,j} \ . \tag{5.12}$$

The attenuation itself is unitless. The geometric factor $g_{n,j}$ is the length of the cell along the line of sight (see Figure 5.2 for Cartesian and radial examples). Thus $\rho_{n,j}$, the dust density, has units of attenuation per unit length. Opting for parsecs as our distance unit, $\left[\rho_{n,j}\right] = \mathrm{pc}^{-1}$. Now, let a_n be a measurement of the attenuation towards star n. Adopting a Gaussian noise model with standard deviation σ_n means that the probability of the measurements is

$$P(a_n|\{\rho_{n,j}\}) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left[-\frac{1}{2\sigma_n^2}(a_n - f_n)^2\right] \ , \tag{5.13}$$

where $\{\rho_{n,j}\}$ denotes just those cells along the line of sight towards star n. If the non-attenuated source intensity is I_0, then we assume that the observed intensity due to an attenuation a_n is

$$I = I_0\, e^{-a_n} \ . \tag{5.14}$$

The measured extinction A_n (in magnitudes) is then computed as

$$A_n = -2.5 \log_{10} \left(\frac{I}{I_0} \right) \ , \tag{5.15}$$

which results in extinction and attenuation being related as $A_n \simeq 1.0857 a_n$.

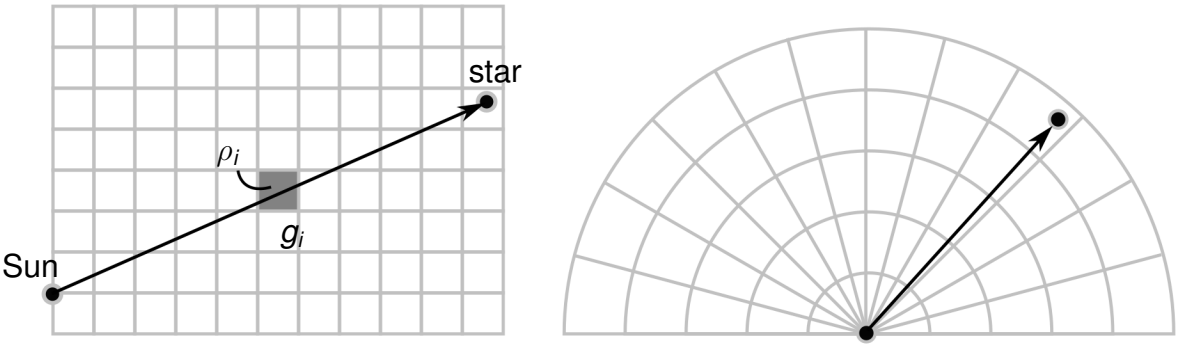

Figure 5.2: Schematic of the geometric factor, left for a Cartesian grid, right for radial grid, centered on the Sun. The radially symmetric grid significantly simplifies the computation of the geometric factor.

If we measure the extinction towards N stars and use a total of J dust cells (not necessarily the same number for each star), then we define $\boldsymbol{G}$ as the $N \times J$ matrix with elements $g_{n,j}$, such that the n^{th} row of $\boldsymbol{G}$ contains the geometric factors just for star n. Assuming a radial geometry (as in the right panel of Figure 5.2), the matrix is very sparse, because most row elements are zero (corresponding to the cells for all other stars). Specifically, each column has just one non-zero element because stars do not share lines of sight. Thus we can write Equation 5.12 as

$$\boldsymbol{f}_N = \boldsymbol{G}\boldsymbol{\rho}_J \ , \tag{5.16}$$

where we define the set of dust densities in all J cells as the J-dimensional vector $\boldsymbol{\rho}_J$. The predicted attenuation towards the N stars is denoted as $\boldsymbol{f}_N$. With the N attenuation measurements (vector $\boldsymbol{a}_N$) and covariance $\boldsymbol{V}_N$, we can generalize Equation 5.13 to be an N-dimensional Gaussian

$$P(\boldsymbol{a}_N|\boldsymbol{\rho}_J) = \frac{1}{(2\pi)^{N/2}|\boldsymbol{V}_N|^{1/2}} \exp\left[-\frac{1}{2} (\boldsymbol{a}_N - \boldsymbol{G}\boldsymbol{\rho}_J)^{\mathsf{T}} \boldsymbol{V}_N^{-1} (\boldsymbol{a}_N - \boldsymbol{G}\boldsymbol{\rho}_J) \right] \ . \tag{5.17}$$

This is a likelihood and defines the probability of observing the data given a set of model parameters. From this we want to find $P(\rho_{J+1}|\boldsymbol{a}_N)$, to be able to estimate the dust density ρ_{J+1} at an arbitrary point $\boldsymbol{r}_{J+1}$ in 3D space. Whereas each element of $\boldsymbol{\rho}_J$ refers to the average dust density in the corresponding cell, ρ_{J+1} is the density at the

point $\boldsymbol{r}_{J+1}$. There is no concept of a cell for points where we want to predict the dust density.

Gaussian process model

A typical problem may involve $N = 10^4$ and $J = 10^5$ (on average 10 cells per line of sight). In order to infer the dust densities, we need to introduce some connection between the lines of sight, otherwise we just have N independent equations like Equation 5.12 with J unknowns.

Therefore lines of sight are connected using a Gaussian process (see Section 5.1) to calculate the joint probability distribution of the dust density at any J different points (or cell centres). This is a J-dimensional Gaussian that is fully characterised by the choice of covariance matrix, $\boldsymbol{C}_J$, which depends on the distance between the points (or cell centres), i.e.

$$P(\boldsymbol{\rho}_J) = \frac{1}{(2\pi)^{J/2}|\boldsymbol{C}_J|^{1/2}} \exp\left[-\frac{1}{2}\boldsymbol{\rho}_J^{\mathsf{T}}\boldsymbol{C}_J^{-1}\boldsymbol{\rho}_J\right] \ . \tag{5.18}$$

In this formulation we explicitly assume a zero mean Gaussian, implying that the dust density is zero in regions that are not constrained by the data. To fully utilise the Gaussian process approach, we exploit the fact that the conditional distribution, $P(\rho_{J+1}|\boldsymbol{\rho}_J)$, is also Gaussian.

As indicated above, the choice of covariance function is an essential factor when using a Gaussian process as it determines the elements, $c_{i,j}$, of the covariance matrix between two points (or cells) i and j at positions $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$, respectively. In practice we use a truncated covariance function from Gneiting (2002)

$$c_{i,j} = \begin{cases} \theta\,(1+t^{\alpha})^{-3}\left[(1-t)\cos(\pi t) + \frac{1}{\pi}\sin(\pi t)\right] & \text{if } 0 \leq t \leq 1 \\ 0 & \text{otherwise ,} \end{cases} \tag{5.19}$$

where $t = |\boldsymbol{r}_i - \boldsymbol{r}_j|/\lambda$ with $\theta > 0$, $\lambda > 0$. Examples of this function are shown in Figure 5.3 for different values of α. The choice of α affects the smoothness of the spatial variation because the gradient at zero separation changes significantly with it (e.g. $\alpha = 2$ has zero gradient at $t = 0$ and a smoother spatial variation than $\alpha = 1$). In the following work we use $\alpha = 1$ (the solid line in the figure).

The covariance decreases monotonically from $t = 0$ to the scale length (λ), where it drops to zero. The FWHM (full width half maximum – the point at which the covariance drops to half the maximum value, θ) is $t = 0.2$. This means the effective

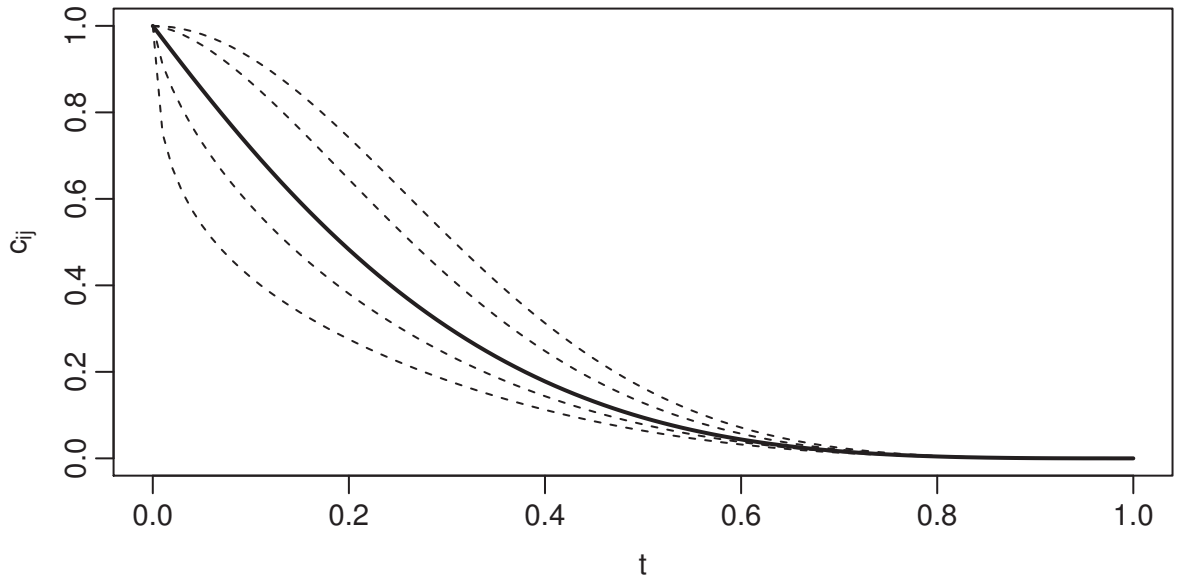

Figure 5.3: Covariance function in equation 5.19 with θ = 1. The solid line is the case of α = 1 (used in this work) and the dashed lines, from bottom to top, are for different values of $\alpha = \{0.5, 0.75, 1.5, 2\}$.

correlation distance is actually lower than the value chosen for λ. The hyperparameter θ determines the overall scale of variations of the dust density. We consider the two hyperparameters, λ and θ, to be fixed (see Section 5.3 for analysis of their influence on the model priors), although in principle they could also be inferred from the data.

Beyond the fact that this covariance function has compact support (it is truncated at λ – points separated by more than this distance will not have any effect on each other), the specific form of the function is not physically motivated. A more practical reason for using this kind of function is that it results in sparse covariance matrices, which have significant computational benefits (both in memory and operations terms) compared to dense ones (such as an exponential covariance function). It is important to note, however, that not every truncated function is necessarily a valid covariance function (being truncated is not a sufficient requirement – see, e.g., Rasmussen & Williams, 2006).

In Figure 5.4 we show samples drawn from the prior for different values of the hyperparameters. This is achieved by setting up a grid of 1000 equally-spaced points from r = 0 to r = 5000 pc. A draw from this 1000-dimensional Gaussian is plotted as a function r. For each fixed pair of λ and θ, we draw three samples from the prior (orange, turquoise and purple lines). The sharpness of the function is due to the shape of our covariance function (Figure 5.3) which influences the smoothness of variations. As indicated earlier, a larger value of α in Equation 5.19 produces

smoother variations. It is clear that higher values of θ result in more marked and larger amplitude variations. Larger values of λ reduce variations at smaller scales. The fact that the samples vary around zero is due to using a zero-mean prior. We will see that in the presence of good data our posterior is mostly determined by the likelihood rather than the prior (Section 5.4).

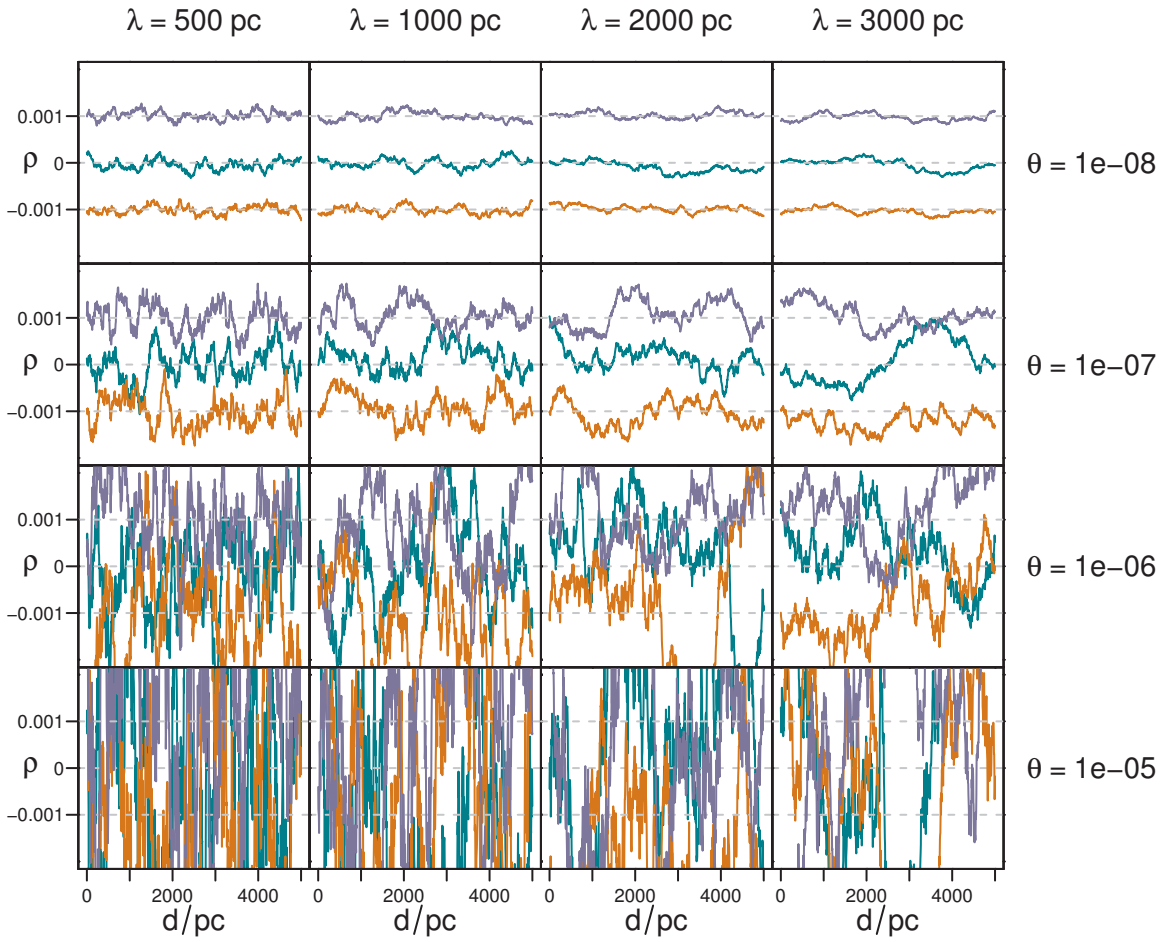

Figure 5.4: Dust density drawn from the covariance function prior (equation 5.19) with different hyperparameters. Each orange, turquoise and purple line shows one draw from a 1000-dimensional Gaussian for a fixed λ and θ. The purple (orange) line is shifted up (down) by $\Delta\rho = 0.001$ in order to more clearly see the variation between separate draws from the prior.

Combining this Gaussian process prior with the likelihood (Equation 5.17), we can now determine $P(\rho_{J+1}|\boldsymbol{a}_N)$, the probability of the dust density ρ_{J+1} at point $\boldsymbol{r}_{J+1}$ given N attenuation measurements $\boldsymbol{a}_N$. This means we are estimating $J+1$ parameters from N measurements. As $J \geq N$, the resulting density estimates will not be independent. However, this is exactly the point of the Gaussian process model: we introduce correlations between the dust cells to make the problem tractable and – more significantly – to allow us to infer a PDF over the dust density at unobserved

points.

Numeric solution

Using the law of marginalization over each line of sight towards observed stars and then applying Bayes theorem, we can write the posterior PDF of the dust density at any point given the data as

$$\begin{aligned} P(\rho_{J+1}|\boldsymbol{a}_N) &= \int P(\rho_{J+1}, \boldsymbol{\rho}_J|\boldsymbol{a}_N)\, \mathrm{d}\boldsymbol{\rho}_J \\ &= \int \frac{P(\rho_{J+1}, \boldsymbol{\rho}_J) P(\boldsymbol{a}_N|\rho_{J+1}, \boldsymbol{\rho}_J)}{P(\boldsymbol{a}_N)}\, \mathrm{d}\boldsymbol{\rho}_J \\ &= \frac{1}{P(\boldsymbol{a}_N)} \int P(\rho_{J+1}, \boldsymbol{\rho}_J) P(\boldsymbol{a}_N|\boldsymbol{\rho}_J)\, \mathrm{d}\boldsymbol{\rho}_J \ , \end{aligned} \tag{5.20}$$

where in the last line we use the fact that $\boldsymbol{a}_N$ is independent of ρ_{J+1} once conditioned on $\boldsymbol{\rho}_J$. This is a J-dimensional integral evaluated over all values of each component of $\boldsymbol{\rho}_J$. The term outside the integral is independent of ρ_{J+1} so is just part of the normalization constant. The first term under the integral is the Gaussian process prior (equation 5.18), now in $J+1$ dimensions. The second term is the likelihood (equation 5.17). Both are Gaussians, but not in ρ_{J+1}. To solve this numerically, we apply the Monte Carlo principle and approximate the integral as

$$P(\rho_{J+1}|\boldsymbol{a}_N) \simeq \frac{1}{K} \sum_k P(\rho_{J+1}|(\boldsymbol{\rho}_J)_k) \ , \tag{5.21}$$

where $(\boldsymbol{\rho}_J)_k$ is a sample of the dust density vector drawn from the posterior PDF $P(\boldsymbol{\rho}_J|\boldsymbol{a}_N)$. Thus the posterior PDF over ρ_{J+1} is a sum of 1D Gaussian process prior PDFs, each with a different value of $\boldsymbol{\rho}_J$ drawn from its J-dimensional posterior. This, in general, is not a Gaussian itself.

Performing this integration in practice is (prohibitively) computationally expensive. The posterior is defined in J dimensions, and for real problems we expect $J > 10^4$. It would be extremely hard to draw reliable samples from this, and/or the sum would have to be over an impossibly large number of samples.

Analytic solution

Luckily, the arguments of the terms in Equation 5.20 are linear functions of ρ_{J+1}, and the integral has an analytic solution. Let

$$\boldsymbol{\rho}_{J+1} = \begin{bmatrix} \boldsymbol{\rho}_J \\ \rho_{J+1} \end{bmatrix} \tag{5.22}$$

be the concatenation of the J dust densities with the dust density at the new point. Its covariance matrix is $\boldsymbol{C}_{J+1}$ and the distribution of $\boldsymbol{\rho}_{J+1}$ follows Equation 5.18 with $J \to J+1$. We now partition the inverse of the covariance matrix in an analogous way as before

$$\boldsymbol{C}_{J+1}^{-1} = \begin{bmatrix} \boldsymbol{M}_J & \boldsymbol{m}_J \\ \boldsymbol{m}_J^{\mathsf{T}} & \mu \end{bmatrix} \tag{5.23}$$

where $\boldsymbol{M}_J$ is a $J \times J$ matrix, $\boldsymbol{m}_J$ is a $J \times 1$ vector, and μ is a scalar. We refer to the steps in Appendix A to show that the result of the integration is a Gaussian with mean $-\beta/\alpha$ and variance $1/\alpha$, such that

$$P(\rho_{J+1}|\boldsymbol{a}_N) = \sqrt{\frac{\alpha}{2\pi}} \exp\left[-\frac{\alpha}{2}\left(\rho_{J+1} + \frac{\beta}{\alpha}\right)^2\right] , \tag{5.24}$$

where

$$\begin{aligned} \alpha &= \mu - \boldsymbol{m}_J^{\mathsf{T}} \boldsymbol{R}_J^{-1} \boldsymbol{m}_J \,, \\ \beta &= \boldsymbol{a}_N^{\mathsf{T}} \boldsymbol{V}_N^{-1} \boldsymbol{G} \boldsymbol{R}_J^{-1} \boldsymbol{m}_J \;\text{ and} \\ \boldsymbol{R}_J &= \boldsymbol{M}_J + \boldsymbol{G}^{\mathsf{T}} \boldsymbol{V}_N^{-1} \boldsymbol{G} \,. \end{aligned} \tag{5.25}$$

The calculation for the one-dimensional PDF over any single ρ_{J+1} at position $\boldsymbol{r}_{J+1}$ involves inverting several matrices of size J, and this takes time $\mathcal{O}(J^n)$, where $n \lesssim 3$ for exact matrix inversion. In Appendix A we illustrate how some of these calculations can be accelerated using matrix identities.

5.3 Simulations

We set up mock observations to explore the model's ability to capture known structures in the dust distribution. For this we test three simplified galaxy models. One has a constant dust density throughout, the second has a dust density profile that decreases exponentially from the galactic centre in all directions. The third adds a dust cloud with increased density along the line of sight. In all cases we place the observer 8.5 kpc from the centre at $l = 0^\circ$ and observe along a pencil beam towards

the galactic centre (l = 0°, b = 0°). For simplicity, all three examples are limited to 1D representations, although the method is defined generally to handle 3D data.

Constant dust density

We set up a simple model that has a constant dust density throughout the galaxy. In Figure 5.5 we show the known dust density profile (left panel) and integrated line of sight attenuation (right panel) along a pencil beam towards the galactic centre. Due to the constant dust density, the line of sight extinction shows a linear increase, as expected.

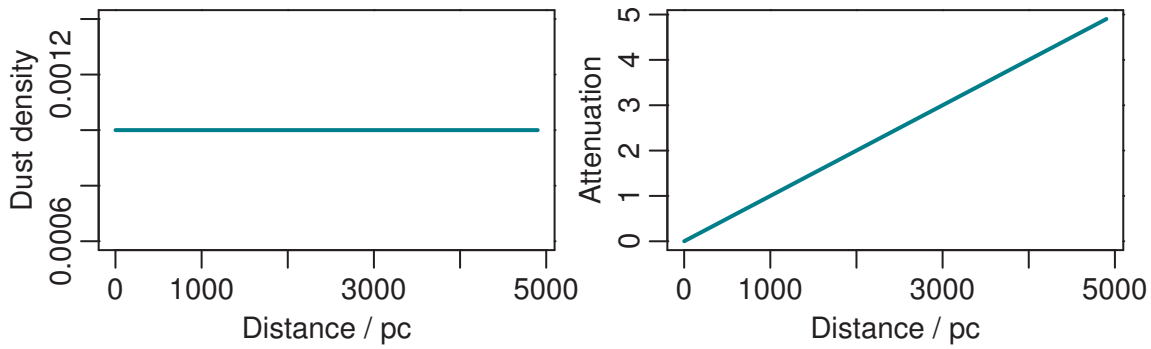

Figure 5.5: Dust density (left panel) and attenuation (right panel) for the constant density galaxy model as a function of distance for a pencil beam towards the galactic centre.

We now randomly select 20 positions along the line of sight and use the Gaussian process model to estimate the dust density at these positions. We also use 20 randomly spaced extinction "observations" along this line of sight as input data. We use two values for measurement noise (σ_n in Equation 5.13), 0.1 and 0.01. We use $\theta = 10^{-7}$ and λ = 2000 pc as the hyperparameters of the covariance function (Equation 5.19; see below for an assessment of their impact). In Figure 5.6 we show the results for both σ_n cases. Black circles are true values (constant), whereas the turquoise circles with errorbars show the inferred densities. Note that the size of the errorbars is exaggerated to emphasise the impact of the measurement uncertainties on the predictions.

We see very good agreement with the model, in particular with a smaller σ. The mean absolute error (MAE) of the predicted dust density is $7.0 \cdot 10^{-5}$ ($4.7 \cdot 10^{-5}$) for σ = 0.1 (σ = 0.01). The relative accuracies ($< |\rho_{true} - \rho_{\mathrm{pred}}|/\rho_{\mathrm{true}} >$) are $4.1 \cdot 10^{-2}$ and $3.2 \cdot 10^{-2}$, respectively (the constant dust density is set to 0.001, as can be seen in Figures 5.5 and 5.6). As we can see, for this very simple case the model is capable of inferring the dust density with a very high accuracy.

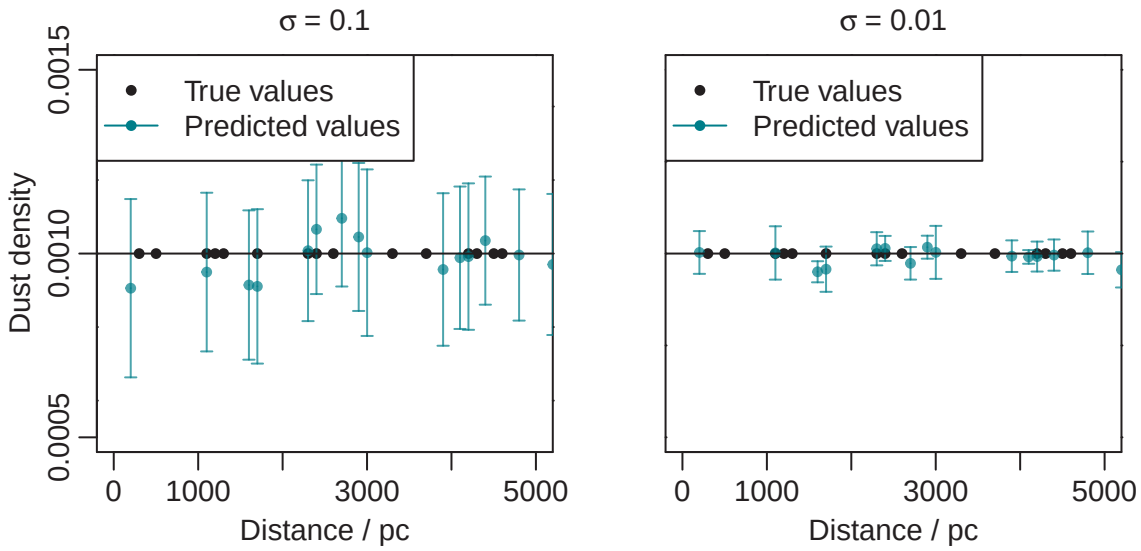

Figure 5.6: Dust density as a function of distance for the constant density model. Left panel uses $\sigma = 0.1$, panel on the right $\sigma = 0.01$. True values at the measurement positions are shown as black circles, predicted values as turquoise circles and errorbars. Errorbars are exaggerated to show the impact of σ on the predictions more clearly.

Exponential dust density

We now repeat the experiment with a slightly more complicated model. The underlying dust density decreases exponentially from the galactic centre (see solid lines in Figure 5.7). Considering attenuation is the integral of the dust density along the line of sight, it should come as no surprise that this also follows an exponential profile (the observations are centred on the galactic centre and thus along the density gradient vector.)

As before, we randomly generate 20 observations along the line of sight and infer the dust density at a further 20 random positions. With fixed θ and λ we illustrate the outcome for the two values of σ in Figure 5.7.

For this model, we also measure a very good performance. The predicted dust densities clearly recover the underlying model. The mean absolute error of the dust density is $1.2 \cdot 10^{-3}$ ($7.8 \cdot 10^{-4}$) for $\sigma = 0.1$ ($\sigma = 0.01$). The relative accuracies are $1.7 \cdot 10^{-1}$ and $1.0 \cdot 10^{-1}$, respectively. Due to the more complex model, the accuracy is slightly lower that in the case of constant density. However, the method is able to reliably predict the dust densities in both cases.

Exponential dust density with cloud

We now modify this setup slightly by adding an artificial dust cloud 2000 pc from the observer with a depth of 500 pc. This dust cloud has a fixed density that is

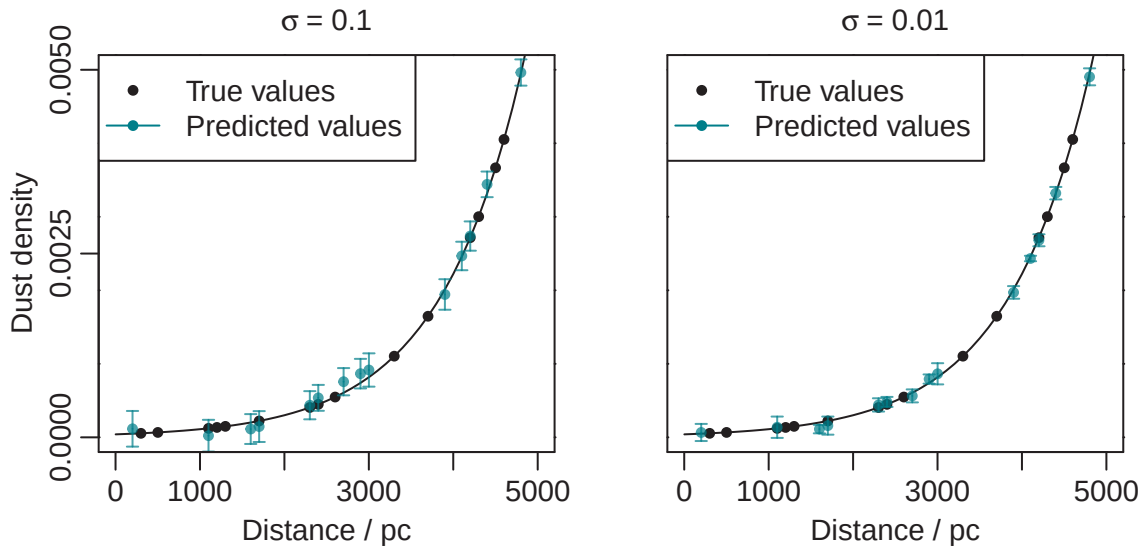

Figure 5.7: Dust density as a function of distance for exponential density model. Left panel uses σ = 0.1, panel on the right σ = 0.01. True values at the measurement positions are shown as black circles, predicted values as turquoise circles and errorbars. Errorbars are exaggerated to show the impact of σ on the predictions.

higher than the underlying exponential distribution at that position (see solid lines in Figure 5.8). This time we generate 40 observations and infer the density at 40 positions. This is done to ensure we have a sufficient number of observations around the position of the dust cloud. In any case, the effect it has on the extinction measurements would be evident. For the same θ, λ and σ values we show the results in Figure 5.8.

We again are able to reliably recover the theoretical density. This time the difference in measurement uncertainty (σ) is more evident. The profile is closer matched for the lower uncertainty (right panel), even though the impact of the cloud is clearly visible in both cases. The mean absolute error of the dust density is $7.5 \cdot 10^{-3}$ ($8.0 \cdot 10^{-3}$) for σ = 0.1 (σ = 0.01). The relative accuracies are $2.8 \cdot 10^{-1}$ and $2.0 \cdot 10^{-1}$, respectively. We again see a slight decrease in accuracy compared to the previous models, although this is to be expected given the added feature in the dust density.

A more complicated model is shown in Rezaei Kh. et al. (2017). It shows that the method is capable of predicting the dust density in the presence of localised dust clouds, as well as returning consistent predictions in regions that have no direct observations.

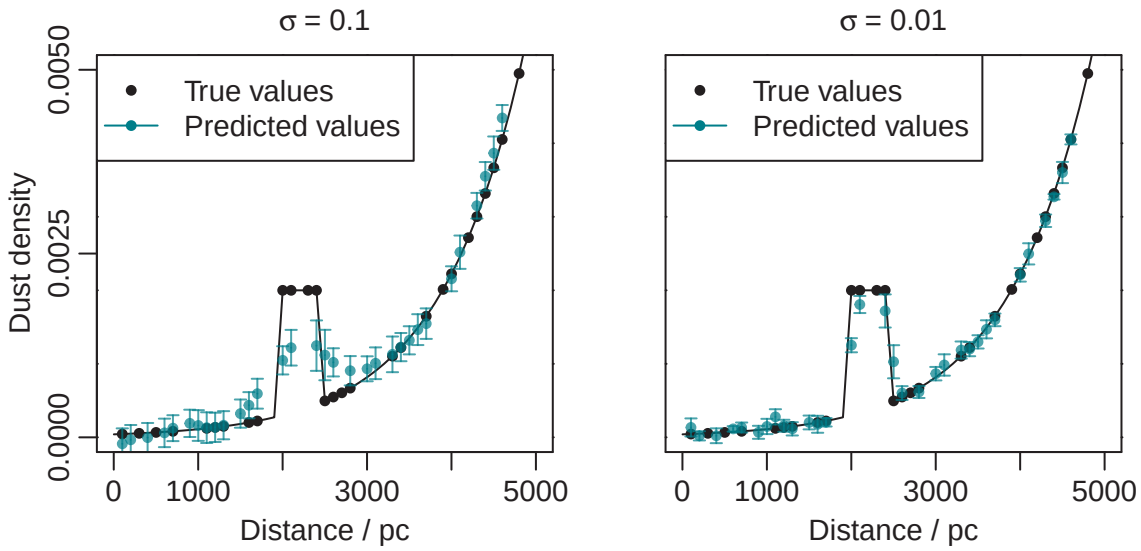

Figure 5.8: Dust density as a function of distance for exponential density model with added dust cloud (solid black lines). Left panel uses σ = 0.1, panel on the right σ = 0.01. True values at the measurement positions are shown as black circles, predicted values as turquoise circles and errorbars. Errorbars are exaggerated to show the impact of σ on the predictions.

Effect of hyperparameters

There are three key parameters that we can set in the model: the hyperparameters of the covariance function, θ and λ, and the size of the cells g (which could also be different for every star, although we use a global value).

The hyperparameter θ sets the absolute scale of the covariance and with it the amplitude of dust variations. For fixed λ and g a larger value of θ means we can capture larger variations. For zero separation, θ is thus the expected variance in the dust density at any point.

The length scale λ defines the maximum correlation length between dust cells (recall, we are using a truncated covariance function). The effective correlation length is actually smaller (see Figure 5.3). Therefore, using a very small value will result in most cells being disconnected (uncorrelated) from one another, meaning that information about dust clouds, etc. will not be propagated effectively. Using a very large correlation length will naturally connect very distant points and thus smooth out local variations in the data.

The cell size is the distance over which we assume the dust density to be constant in the model. It is used to represent the attenuation towards observed stars, however it is not used directly in any further calculations and therefore does not impose a resolution limit for the model. It defines the distance scale for which the model is sensitive to variations. Therefore we aim to use a small value. Unfortunately,

inverting the covariance matrix (which has dimensions $J \times J$) is an $\mathcal{O}(J^3)$ operation and therefore computationally constrains the size we can use in practice.

The actual minimum length scale over which the model can detect dust variations is given by a combination of these three parameters and, importantly, by the data themselves.

In Figure 5.9 we show the impact of varying θ and λ on the model predictions for stars using the exponential dust model. In all panels the black dots denote the true dust densities at the positions of the observations, whereas the turquoise points and errorbars are the predictions. Each column uses a fixed value for θ (see above the top row), each row uses a fixed λ (see right hand side of figure). All simulations are run with an attenuation error of σ = 0.1, although the qualitative results are the same for σ = 0.01.

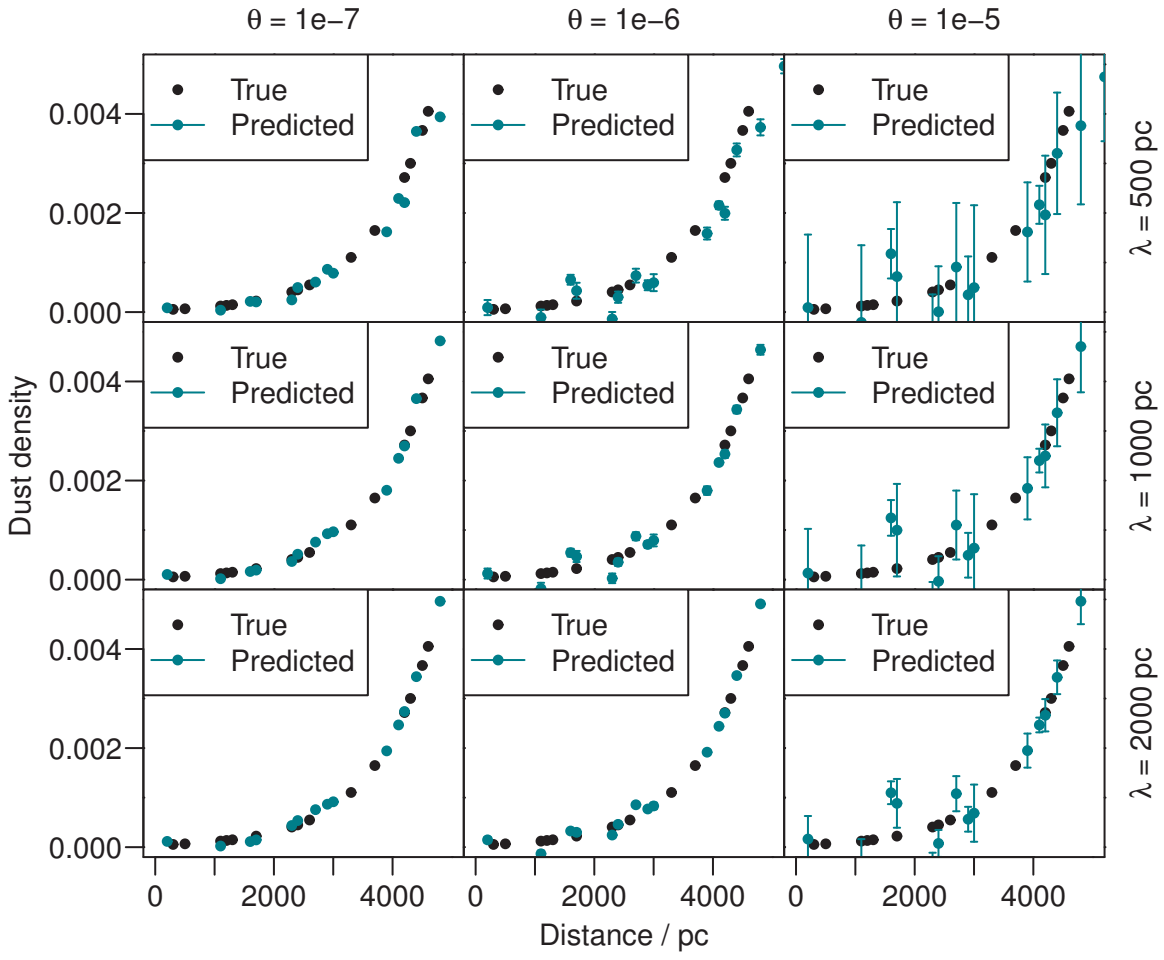

Figure 5.9: Effects of the hyperparameters of the covariance function on the model predictions for an exponential dust model for σ = 0.1. Black points are the true values, turquoise points with error bars are the predictions. Each column uses a different value for θ (see above first row), each row uses a different λ (see to the right of the third column), resulting in nine different (θ, λ) pairs. Errorbars are multiplied by a constant factor to emphasise the differences.

In general we see two main effects. Larger values of the correlation length λ result in smoother variations of the predicted dust density. This is what we expect, considering more cells are connected with each other the larger it becomes and thus increases the number of measurements used to infer dust densities at new points. This also causes the size of the errorbars to decrease with increasing λ.

As indicated above, larger values of θ enable the model capture steeper changes in the dust density. For fixed λ we see smoother changes in the predicted dust density for smaller θ.

Our choice of covariance function implies that points separated by less than the correlation length λ have correlated posterior distributions. Not only are the predicted dust densities correlated, so are their uncertainties. The errorbars of neighbouring points in Figures 5.7 and 5.9 are therefore highly correlated. The choice of where to predict the dust density is not affected by the number of points at which the density is predicted, only by the number of measurements (and their uncertainties), assuming a particular model. Therefore we can trace variations of the dust density at a smaller scale than given by the errorbars.

5.4 Outlook

We have introduced a novel non-parametric method for reconstructing a smooth, three-dimensional map of dust density from line of sight extinction (attenuation) measurements. By using a Gaussian process prior to constrain the variation of dust density, we avoid having to define a specific functional form of this spatial correlation. Instead, we define a simple covariance function that depends on the distance between two points. We are therefore able to predict the dust density in regions where we have no observations and overcome the discontinuities seen in previous 3D dust maps. For each of these new points we predict the posterior probability function over the dust. This function is a Gaussian, and we have derived an analytic solution for the mean and standard deviation.

We have applied the method to three simplified mock galaxy models and showed that we are able to reliably reconstruct dust structures in regions with observed extinctions. We have explored the effect of the covariance hyperparameters on the predicted dust densities, illustrating that a larger correlation length scale λ results in smoother predictions, whereas the amplitude θ affects how well steep changes in density can be captured. We are able to probe dust structures at significantly smaller scales than defined by λ.

Despite having an analytical solution to the problem, and thus making a computational solution feasible in the first place (see Section 5.2), the computation time required to solve a small problem set is still significant. This is due to the matrix algebra involved, in particular inverting the covariance matrix, which scales non-linearly with the number of cells J (generally $\mathcal{O}(J^n)$ with $n \lesssim 3$). Although this particular step only needs to be performed once, there are several other matrix inversions and multiplications which must be computed for every new prediction. For example, inverting the covariance matrix with $N = 1000$ and $J = 10\,000$ takes around one hour with a single core on a modest computer. In principle this step can be parallelised trivially. Predicting the density at a new point with these parameters takes approximately 50 seconds (several points can also be parallelised). However, a significant amount of memory is needed per core, even using sparse matrix techniques, thus limiting the potential problem size on normal machines.

As such, to use this method to probe large regions (requiring thousands of observations) will require some additional optimisation beyond basic parallelisation. One approach could be to use approximate matrix inversion techniques that require $\mathcal{O}(N^2)$ time. Furthermore, intelligently partitioning and overlapping regions that are solved for individually may be possible, as long as discontinuities are taken into account properly.

In its current form, the method only uses the uncertainties in the extinction measurements, but doesn't take into account uncertainties that arise from inferring distances to the stars. In practice these will have to factored in. Even assuming good general accuracy from Gaia, distant and faint stars will still have uncertain distance estimates (see, e.g. Bailer-Jones, 2015). Future work will explore the possibility of including distance errors in the model, as well as incorporating the other suggested modifications.

6

Conclusion

In this thesis we have presented two conceptually different approaches to measuring dust in the Milky Way Galaxy. Dust plays a crucial role in understanding the physical properties of the Galaxy. Therefore, being able to reliably infer its distribution will enable a better understanding of the objects that are affected by it and the processes occurring within it.

In Part 1 we extend a parametric Bayesian method to infer extinction, extinction parameter, effective temperature and distance modulus to millions of stars individually. We are able to predict these parameters with a high precision and good model accuracy. We present maps of the Galactic poles by utilising broadband photometry from SDSS and UKIDSS and of the Galactic plane using Pan-STARRS 1 and Spitzer GLIMPSE data.

In Part 2 we introduce a non-parametric Gaussian process method to infer the local dust density from extinction and distance measurements to individual stars. We demonstrate its ability on a mock galaxy model and are able to reliably recover known features in the dust density.

Both of these methods will benefit immensely from the upcoming *Gaia* data releases. In particular, combining multiband photometry with the vast number of *Gaia* parallaxes will enable an increase in accuracy of the method presented in Part 1, which in its current form only has a relatively weak handle on distance modulus. With this potential improvement on distance estimates we may be able to expand the method to take into account metallicity by using a more detailed, metallicity-dependent, HRD prior to further improve the model's accuracy (as illustrated in Section 3.4).

Similarly, utilising *Gaia* parallaxes for the Gaussian process method, along with

the inferred extinctions (see, e.g. Bailer-Jones et al., 2013), will enable us to infer the dust density for large regions of the Galaxy. Providing we continue to develop the method to overcome some of the computational limitations detailed in Section 5.4, a full, self-consistent 3D dust map of the observable Milky Way might be a viable prospect.

A

Analytic Solution for Gaussian Process

I explicitly show the steps to obtain the analytic solution of the integral in Equation 5.20, for which the final solution is shown in Equation 5.24 (Section A.1). I also highlight the matrix identities that are used (Section A.2).

A.1 Analytic solution of the integral

We rewrite Equation 5.20 in the general form

$$P(\rho_{J+1}|\boldsymbol{a}_N) = \frac{1}{Z}\int e^{-\psi/2}\,\mathrm{d}\boldsymbol{\rho}_J\,, \tag{A.1}$$

where Z is a normalisation constant. From Equations 5.18 and 5.17 (with $J \to J+1$) we see that

$$\begin{aligned}\psi &= \boldsymbol{\rho}_{J+1}^{\mathsf{T}}\boldsymbol{C}_{J+1}^{-1}\boldsymbol{\rho}_{J+1} + (\boldsymbol{a}_N - \boldsymbol{G}\boldsymbol{\rho}_J)^{\mathsf{T}}\boldsymbol{V}_N^{-1}(\boldsymbol{a}_N - \boldsymbol{G}\boldsymbol{\rho}_J) \\ &= \boldsymbol{\rho}_{J+1}^{\mathsf{T}}\boldsymbol{C}_{J+1}^{-1}\boldsymbol{\rho}_{J+1} + \boldsymbol{a}_N^{\mathsf{T}}\boldsymbol{V}_N^{-1}\boldsymbol{a}_N - \boldsymbol{\rho}_J^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{V}_N^{-1}\boldsymbol{a}_N + \boldsymbol{\rho}_J^{\mathsf{T}}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{V}_N^{-1}\boldsymbol{G}\boldsymbol{\rho}_J - \boldsymbol{a}_N^{\mathsf{T}}\boldsymbol{V}_N^{-1}\boldsymbol{G}\boldsymbol{\rho}_J\,,\end{aligned} \tag{A.2}$$

with

$$\boldsymbol{\rho}_{J+1} = \begin{bmatrix} \boldsymbol{\rho}_J \\ \rho_{J+1} \end{bmatrix}\,. \tag{A.3}$$

We can eliminate the third and fifth terms in Equation A.2, because all terms are scalars and transposing them does not change them. To evaluate the integral we

need to separate $\boldsymbol{\rho}_J$ from ρ_{J+1} in $\boldsymbol{\rho}_{J+1}$. We first partition the inverse covariance matrix in the following way

$$\boldsymbol{C}_{J+1}^{-1} = \begin{bmatrix} \boldsymbol{M}_J & \boldsymbol{m}_J \\ \boldsymbol{m}_J^{\mathsf{T}} & \mu \end{bmatrix} \tag{A.4}$$

where $\boldsymbol{M}_J$ is a $J \times J$ matrix, $\boldsymbol{m}_J$ is a $J \times 1$ vector, and μ is a scalar. It is important to note that these are components of the inverted matrix, not components of $\boldsymbol{C}_{J+1}$ which are then inverted. We can then write

$$\begin{aligned} \boldsymbol{\rho}_{J+1}^{\mathsf{T}} \boldsymbol{C}_{J+1}^{-1} \boldsymbol{\rho}_{J+1} &= \begin{bmatrix} \boldsymbol{\rho}_J^{\mathsf{T}} & \rho_{J+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{M}_J & \boldsymbol{m}_J \\ \boldsymbol{m}_J^{\mathsf{T}} & \mu \end{bmatrix} \begin{bmatrix} \boldsymbol{\rho}_J \\ \rho_{J+1} \end{bmatrix} \\ &= \boldsymbol{\rho}_J^{\mathsf{T}} \boldsymbol{M}_J \boldsymbol{\rho}_J + 2\rho_{J+1} \boldsymbol{m}_J^{\mathsf{T}} \boldsymbol{\rho}_J + \mu \rho_{J+1}^2 \ . \end{aligned} \tag{A.5}$$

Substituting this into equation A.2 and gathering together terms gives

$$\psi = \boldsymbol{\rho}_J^{\mathsf{T}} (\boldsymbol{M}_J + \boldsymbol{G}^{\mathsf{T}} \boldsymbol{V}_N^{-1} \boldsymbol{G}) \boldsymbol{\rho}_J + 2(\rho_{J+1} \boldsymbol{m}_J^{\mathsf{T}} - \boldsymbol{a}_N^{\mathsf{T}} \boldsymbol{V}_N^{-1} \boldsymbol{G}) \boldsymbol{\rho}_J + (\boldsymbol{a}_N^{\mathsf{T}} \boldsymbol{V}_N^{-1} \boldsymbol{a}_N + \mu \rho_{J+1}^2) \ , \tag{A.6}$$

which is a quadratic expression in $\boldsymbol{\rho}_J$. The last term is independent of $\boldsymbol{\rho}_J$ so can be taken out of the integral, allowing us to write equation A.1 as

$$\begin{aligned} P(\rho_{J+1} | \boldsymbol{a}_N) = \frac{1}{Z} \exp\left(-\frac{1}{2} \boldsymbol{a}_N^{\mathsf{T}} \boldsymbol{V}_N^{-1} \boldsymbol{a}_N \right) \exp\left(-\frac{1}{2} \mu \rho_{J+1}^2 \right) \\ \times \int \exp\left(-\frac{1}{2} \boldsymbol{\rho}_J^{\mathsf{T}} \boldsymbol{R}_J \boldsymbol{\rho}_J + \boldsymbol{b}_J^{\mathsf{T}} \boldsymbol{\rho}_J \right) \mathrm{d}\boldsymbol{\rho}_J \ , \end{aligned} \tag{A.7}$$

where

$$\begin{aligned} \boldsymbol{R}_J &= \boldsymbol{M}_J + \boldsymbol{G}^{\mathsf{T}} \boldsymbol{V}_N^{-1} \boldsymbol{G} \\ \boldsymbol{b}_J &= \boldsymbol{G}^{\mathsf{T}} \boldsymbol{V}_N^{-1} \boldsymbol{a}_N - \rho_{J+1} \boldsymbol{m}_J \ . \end{aligned} \tag{A.8}$$

This integral again has a standard form with analytical solution

$$\int \exp\left(-\frac{1}{2} \boldsymbol{\rho}_J^{\mathsf{T}} \boldsymbol{R}_J \boldsymbol{\rho}_J + \boldsymbol{b}_J^{\mathsf{T}} \boldsymbol{\rho}_J \right) \mathrm{d}\boldsymbol{\rho}_J = \frac{(2\pi)^{J/2}}{|\boldsymbol{R}_J|^{1/2}} \exp\left(\frac{1}{2} \boldsymbol{b}_J^{\mathsf{T}} \boldsymbol{R}_J^{-1} \boldsymbol{b}_J \right) \ , \tag{A.9}$$

assuming $|\boldsymbol{R}_J| > 0$. Using this we can write equation A.7 as

$$P(\rho_{J+1} | \boldsymbol{a}_N) = \frac{1}{Z} e^{-\phi/2} \qquad \text{with} \qquad \phi = \mu \rho_{J+1}^2 - \boldsymbol{b}_J^{\mathsf{T}} \boldsymbol{R}_J^{-1} \boldsymbol{b}_J \ , \tag{A.10}$$

where we have absorbed all factors which do not depend on ρ_{J+1} into the normalisa-

tion constant. Substituting for $\boldsymbol{b}_J$ from equation A.8 this becomes

$$\begin{aligned}\phi = & -\boldsymbol{a}_N^{\mathsf{T}}\boldsymbol{V}_N^{-1}\boldsymbol{G}\boldsymbol{R}_J^{-1}\boldsymbol{G}^{\mathsf{T}}\boldsymbol{V}_N^{-1}\boldsymbol{a}_N + (\mu - \boldsymbol{m}_J^{\mathsf{T}}\boldsymbol{R}_J^{-1}\boldsymbol{m}_J)\rho_{J+1}^2 + \\ & 2\boldsymbol{a}_N^{\mathsf{T}}\boldsymbol{V}_N^{-1}\boldsymbol{G}\boldsymbol{R}_J^{-1}\boldsymbol{m}_J\rho_{J+1} \ .\end{aligned} \tag{A.11}$$

The first term does not depend on ρ_{J+1} so can also be absorbed into the normalization constant. Inserting this into equation A.10 gives us

$$P(\rho_{J+1}|\boldsymbol{a}_N) = \frac{1}{Z}\exp\left(-\frac{1}{2}\alpha\rho_{J+1}^2 - \beta\rho_{J+1}\right) \ , \tag{A.12}$$

where

$$\begin{aligned}\alpha &= \mu - \boldsymbol{m}_J^{\mathsf{T}}\boldsymbol{R}_J^{-1}\boldsymbol{m}_J \\ \beta &= \boldsymbol{a}_N^{\mathsf{T}}\boldsymbol{V}_N^{-1}\boldsymbol{G}\boldsymbol{R}_J^{-1}\boldsymbol{m}_J \ .\end{aligned} \tag{A.13}$$

By completing the square in the exponent

$$-\frac{1}{2}\alpha\rho_{J+1}^2 - \beta\rho_{J+1} \ = \ -\frac{\alpha}{2}\left(\rho_{J+1} + \frac{\beta}{\alpha}\right)^2 + \frac{\beta^2}{2\alpha} \ , \tag{A.14}$$

and taking the term $\exp(\beta^2/2\alpha)$ outside of the integral (and absorbing it too into normalization constant), we see that $P(\rho_{J+1}|\boldsymbol{a}_N)$ is just a one-dimensional Gaussian with mean $-\beta/\alpha$ and variance $1/\alpha$, i.e.

$$P(\rho_{J+1}|\boldsymbol{a}_N) \ = \ \sqrt{\frac{\alpha}{2\pi}}\exp\left[-\frac{\alpha}{2}\left(\rho_{J+1} + \frac{\beta}{\alpha}\right)^2\right] \ . \tag{A.15}$$

A.2 Accelerating the matrix evaluations via matrix identities

We can simplify the expressions for α and β and thereby reduce the number of matrix multiplications using some standard matrix identities. We first relate the components of the inverted covariance matrix (Equation A.4) to the components of the covariance matrix itself, which we partition as (see also Equation 5.5)

$$\boldsymbol{C}_{J+1} = \begin{bmatrix}\boldsymbol{C}_J & \boldsymbol{k}_J \\ \boldsymbol{k}_J^{\mathsf{T}} & k\end{bmatrix} \ . \tag{A.16}$$

$\boldsymbol{C}_J$ is the $J \times J$ covariance matrix involving only the fixed data. $\boldsymbol{k}_J$ is the $J \times 1$ vector with elements given by $c_{j,J+1}$, i.e. the covariance between the fixed data and the new point. k is the scalar $c_{J+1,J+1}$, the covariance of the new point with itself (i.e. its variance). Recall that each element of the covariance matrix is determined by the covariance function, such as Equation 5.19.

A standard result of matrix algebra (e.g. Press et al., 1992) allows us to write the components of the inverted matrix as

$$\begin{aligned} \boldsymbol{M}_J &= \boldsymbol{C}_J^{-1} + (\boldsymbol{C}_J^{-1}\boldsymbol{k}_J)(k - \boldsymbol{k}_J^{\mathsf{T}}\boldsymbol{C}_J^{-1}\boldsymbol{k}_J)^{-1}(\boldsymbol{k}_J^{\mathsf{T}}\boldsymbol{C}_J^{-1}) \\ \boldsymbol{m}_J &= -(\boldsymbol{C}_J^{-1}\boldsymbol{k}_J)(k - \boldsymbol{k}_J^{\mathsf{T}}\boldsymbol{C}_J^{-1}\boldsymbol{k}_J)^{-1} \\ \mu &= (k - \boldsymbol{k}_J^{\mathsf{T}}\boldsymbol{C}_J^{-1}\boldsymbol{k}_J)^{-1} \ . \end{aligned} \tag{A.17}$$

Temporarily defining

$$\boldsymbol{h}_J = \boldsymbol{C}_J^{-1}\boldsymbol{k}_J \ , \tag{A.18}$$

allows us to write the matrix and vector parts of the inverted covariance matrix as

$$\begin{aligned} \boldsymbol{M}_J &= \boldsymbol{C}_J^{-1} + \mu\boldsymbol{h}_J\boldsymbol{h}_J^{\mathsf{T}} \\ \boldsymbol{m}_J &= -\mu\boldsymbol{h}_J \ . \end{aligned} \tag{A.19}$$

It should be noted that $\boldsymbol{h}_J\boldsymbol{h}_J^{\mathsf{T}}$ is an outer product, and produces a matrix of size $J \times J$. We now use the matrix inversion lemma (also known as the Woodbury matrix identity; e.g. Press et al., 1992) to write, from the first line of Equations A.8

$$\boldsymbol{R}_J^{-1} \ = \ \boldsymbol{M}_J^{-1} - \boldsymbol{M}_J^{-1}\boldsymbol{G}^{\mathsf{T}}(\boldsymbol{V}_N + \boldsymbol{G}\boldsymbol{M}_J^{-1}\boldsymbol{G}^{\mathsf{T}})^{-1}\boldsymbol{G}\boldsymbol{M}_J^{-1} \ . \tag{A.20}$$

The Sherman–Morrison formula (a special case of the matrix inversion lemma) allows us to write

$$\begin{aligned} \boldsymbol{M}_J^{-1} &= \boldsymbol{C}_J - (1 + \mu\boldsymbol{h}_J^{\mathsf{T}}\boldsymbol{C}_J\boldsymbol{h}_J)^{-1}\mu\boldsymbol{C}_J\boldsymbol{h}_J\boldsymbol{h}_J^{\mathsf{T}}\boldsymbol{C}_J \\ &= \boldsymbol{C}_J - (1 + \mu\boldsymbol{k}_J^{\mathsf{T}}\boldsymbol{C}_J^{-1}\boldsymbol{k}_J)^{-1}\mu\boldsymbol{k}_J\boldsymbol{k}_J^{\mathsf{T}} \ , \end{aligned} \tag{A.21}$$

where we have substituted for $\boldsymbol{h}_J$ from Equation A.18.

All of the above allows us to simplify our expressions for α and β in Equation A.13 by using

$$\begin{aligned} \boldsymbol{R}_J^{-1} &= \boldsymbol{M}_J^{-1} - \boldsymbol{M}_J^{-1}\boldsymbol{G}^{\mathsf{T}}(\boldsymbol{V}_N + \boldsymbol{G}\boldsymbol{M}_J^{-1}\boldsymbol{G}^{\mathsf{T}})^{-1}\boldsymbol{G}\boldsymbol{M}_J^{-1} \\ \boldsymbol{M}_J^{-1} &= \boldsymbol{C}_J - (1 + \mu\boldsymbol{k}_J^{\mathsf{T}}\boldsymbol{C}_J^{-1}\boldsymbol{k}_J)^{-1}\mu\boldsymbol{k}_J\boldsymbol{k}_J^{\mathsf{T}} \\ \boldsymbol{m}_J &= -\mu\boldsymbol{C}_J^{-1}\boldsymbol{k}_J \\ \mu &= (k - \boldsymbol{k}_J^{\mathsf{T}}\boldsymbol{C}_J^{-1}\boldsymbol{k}_J)^{-1} \\ k &= c_{J+1,J+1} \\ \boldsymbol{k}_J &= (c_{1,J+1}, c_{2,J+1}, \ldots, c_{J,J+1})^{\mathsf{T}} \ . \end{aligned} \tag{A.22}$$

$\boldsymbol{C}_J^{-1}$ takes time $\mathcal{O}(J^n)$ to compute, where n is typically $\lesssim 3$, but must only be done once. As we use a truncated covariance function, $\boldsymbol{C}_J$ is a sparse matrix, a fact which we exploit when inverting it. To compute the PDF at each new point, various other

matrix inversions and multiplications must be performed, the longest of which takes time $\mathcal{O}(NJ^2)$, because $J \geq N$. It should be noted that $\boldsymbol{G}$ is always sparse (when $J = N$ it is even diagonal, although this would result in a poor resolution map of dust density). Because $\boldsymbol{C}_J$ is sparse, $\boldsymbol{M}_J^{-1}$ is also sparse.

Publications

1. **R. J. Hanson** and C. A. L. Bailer-Jones
   3D Galactic extinction modelling.
   Proceedings of the International Astronomical Union 2013; 9 (S298): 410.
   doi: 10.1017/S1743921313006777

2. **R. J. Hanson**, C. A. L. Bailer-Jones
   3D Galactic dust extinction mapping with multiband photometry.
   Monthly Notices of the Royal Astronomical Society 2014; 438 (4): 2938-2953.
   doi: 10.1093/mnras/stt2401

3. S. Rezaei Kh., C. A. L. Bailer-Jones, **R. J. Hanson**
   Getting ready for Gaia: three-dimensional modeling of dust in the Milky Way.
   Memorie della Societa Astronomica Italiana 2015; 86: 642.
   code: 2015MmSAI..86..642R

4. **R. J. Hanson**, C. A. L. Bailer-Jones, W. S. Burgett, K. C. Chambers, K. W. Hodapp, N. Kaiser, J. L. Tonry, R. J. Wainscoat, C. Waters
   Dust in three dimensions in the Galactic plane.
   Monthly Notices of the Royal Astronomical Society 2016; 463 (4): 3604-3615.
   doi: 10.1093/mnras/stw2240

5. S. Rezaei Kh., C. A. L. Bailer-Jones, **R. J. Hanson**, M. Fouesneau
   Inferring the three-dimensional distribution of dust in the Galaxy with a non-parametric method.
   Astronomy & Astrophysics 2017; 598: A125.
   doi: 10.1051/0004-6361/201628885

Publications 2 and 4 are adapted for use in Part 1. The theoretical model from publication 5 is used in conjunction with elementary simulations and analysis in Part 2. The applications shown in sections 3 and 4 of the publication do not form part of this thesis.

Bibliography

Abazajian K. N. et al., 2009, The Astrophysical Journal Supplement Series, 182, 543 1.3

Aihara H. et al., 2011, The Astrophysical Journal Supplement Series, 193, 29 3

Arce H. G., Goodman A. A., 1999, The Astrophysical Journal, 512, L135 1.3

Arenou F., Grenon M., Gómez A., 1992, Astronomy & Astrophysics, 104 1.3

Bailer-Jones C. A. L., 2011, Monthly Notices of the Royal Astronomical Society, 411, 435 1.3, 2.1, 3.6, 4.5

Bailer-Jones C. A. L., 2015, Publications of the Astronomical Society of the Pacific, 127, 994 5.4

Bailer-Jones C. A. L. et al., 2013, Astronomy & Astrophysics, 559A, 74B 3.4, 3.6, 4.5, 6

Benjamin R. A. et al., 2003, Publications of the Astronomical Society of Pacific, 115, 953 1.2

Berdnikov L., Pavlovskaya E., 1991, Soviet Astronomy Letters, 17, 215 1.3

Berry M. et al., 2012, The Astrophysical Journal, 757, 166 1.3, 3.5, 4.4

Burstein D., Heiles C., 1978, The Astrophysical Journal, 225, 40 1.3

Cardelli J. A., Clayton G. C., Mathis J. S., 1988, The Astrophysical Journal, 329, L33 1.2, 1.2

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, The Astrophysical Journal, 345, 245 1.2, 1.2, 1.3

Casali M., 2007, Astronomy & Astrophysics, 467, 777 (document)

Chen B.-Q. et al., 2014, Monthly Notices of the Royal Astronomical Society, 443, 1192 1.3

Chen B. Q., Schultheis M., Jiang B. W., Gonzalez O. A., Robin A. C., Rejkuba M., Minniti D., 2013, Astronomy & Astrophysics, 550, A42 1.3, 4.3

Churchwell E. et al., 2009, Publications of the Astronomical Society of the Pacific, 121, 213 1.2

Dotter A., Chaboyer B., Jevremović D., Kostov V., Baron E., Ferguson J. W., 2008, The Astrophysical Journal Supplement Series, 178, 89D 4.1

Draine B. T., 2003, Annual Review of Astronomy & Astrophysics, 41, 241 1.1, 1.2

Drew J. et al., 2005, Monthly Notices of the Royal Astronomical Society, 362, 753 1.3

Fazio G. G. et al., 2004, The Astrophysical Journal Supplement Series, 154, 10 4.1

Fitzgerald M. P., 1968, The Astronomical Journal, 73, 983 1.3

Fitzpatrick E. L., 1999, Publications of the Astronomical Society of the Pacific, 111, 63 1.2, 1.2, 1.1, 3.2

Gao J., Jiang B. W., Li A., 2009, The Astrophysical Journal, 707, 89 4.3

Gibbs M. N., 1997, PhD thesis, University of Cambridge 5.1

Gibbs M. N., MacKay D. J. C., 1997, Efficient Implementation of Gaussian Processes. Tech. rep. 5.1

Gneiting T., 2002, Journal of Multivariate Analysis, 83, 493 5.1, 5.2

Gontcharov G. A., 2012a, Astronomy Letters, 38, 87 3.5

Gontcharov G. A., 2012b, Astronomy Letters, 38, 12 3.5, 4.3

Gonzalez O. A., Rejkuba M., Zoccali M., Valent E., Minniti D., Tobar R., 2013, Astronomy & Astrophysics, 552, A110 1.3

Gonzalez O. A., Rejkuba M., Zoccali M., Valenti E., Minniti D., 2011, Astronomy & Astrophysics, 534, A3 1.3

Gonzalez O. A., Rejkuba M., Zoccali M., Valenti E., Minniti D., Schultheis M., Tobar R., Chen B. Q., 2012, Astronomy & Astrophysics, 543, A13 1.3

Green G. M. et al., 2014, The Astrophysical Journal, 783, 114 1.3

Green G. M. et al., 2015, The Astrophysical Journal, 810, 25 1.3

Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, Astronomy & Astrophysics, 486, 951 4.2

Hakkila J., Myers J. M., Stidham B. J., Hartmann D. H., 1997, The Astronomical Journal, 114, 2043 1.3

Hanson R. J., Bailer-Jones C. A. L., 2014, Monthly Notices of the Royal Astronomical Society, 438, 2938 1.3, 1.4, 2, 3, 4, 4.5

Hanson R. J. et al., 2016, Monthly Notices of the Royal Astronomical Society, 463, 3604 1.3, 1.4, 2, 4, 4.2

Hauschildt P. H., Allard F., Baron E., 1999, The Astronomical Journal, 512, 377 3.2

Jones D. O., West A. A., Foster J. B., 2011, The Astronomical Journal, 142, 44 4.3

Kaiser N. et al., 2002, in Proc. SPIE 4836, Survey and Other Telescope Technologies and Discoveries, Tyson J. A., Wolff S., eds., Vol. 4836, p. 154 3.6

Kapteyn J. C., 1922, The Astrophysical Journal, 55, 302 1.1

Kennicutt R. C., Evans N. J., 2012, Annual Review of Astronomy and Astrophysics, 50, 531 1.1

Kochanek C. S., 2011, The Astrophysical Journal, 743, 73 1.1

Lallement R., Vergely J., Valette B., Puspitarini L., Eyer L., Casagrande L., Paris G. O. D., 2014, Astronomy & Astrophysics, 561, A91 1.3, 4.5

Lawrence A. et al., 2007, Monthly Notices of the Royal Astronomical Society, 379, 1599 (document), 1.2, 3

Lee Y. S. et al., 2008, The Astronomical Journal, 136, 2022 3.1

Li A., Greenberg J., 2003, in Solid State Astrochemistry, Pirronello V., Krelowski J., Manicò G., eds., Vol. 120, pp. 37–84 1.1

MacKay D. J. C., 2003, Information Theory, Inference, and Learning Algorithms. Cambridge University Press 5.1

Marshall D. J., Robin A. C., Reylé C., Schultheis M., Picaud S., 2006, Astronomy & Astrophysics, 453, 635 1.3, 4.3

Massa D., Savage B., 1989, in Interstellar Dust, Allamandola L., Tielens A., eds., Vol. 135, p. 3 1.2

Mathis J. S., 1990, Annual Review of Astronomy and Astrophysics, 28, 37 1.2

Mörtsell E., 2013, Astronomy & Astrophysics, 80, A80 4.3

Neckel T., Klare G., Sarcander M., 1980, Astronomy & Astrophysics Supplement Series, 42, 251 1.3

O'Donnell J. E., 1994, The Astrophysical Journal, 422, 158 1.2

Perryman M. A. C., Boer K. S. D., Gilmore G., Høg E., Lattanzi M. G., Lindegren L., Luri X., 2001, Astronomy & Astrophysics, 369, 339 3.6

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes in C: the art of scientific computing, Second Edition A.2, A.2

Rasmussen C. E., Williams C. K. I., 2006, Gaussian Processes for Machine Learning. MIT Press, Cambridge, MA, p. 266 5.1, 5.1, 5.2

Re Fiorentin P., Bailer-Jones C. A. L., Lee Y. S., Beers T. C., Sivarani T., Wilhelm R., Allende Prieto C., Norris J. E., 2007, Astronomy & Astrophysics, 467, 1373 3.1

Rezaei Kh. S., Bailer-Jones C. A. L., Hanson R. J., Fouesneau M., 2017, Astronomy & Astrophysics, 598, A125 1.3, 1.4, 5, 5.3

Rieke G. H., Lebofsky M. J., 1985, The Astrophysical Journal, 288, 618 1.2

Robin A. C., Reylé C., Derriére S., Picaud S., 2003, Astronomy & Astrophysics, 409, 523 1.3

Roman-Zuniga C. G., Lada C. J., Muench A., Alves J. F., 2007, The Astrophysical Journal, 664, 357 1.2

Sale S. E., 2012, Monthly Notices of the Royal Astronomical Society, 427, 2119 1.3

Sale S. E. et al., 2014, Monthly Notices of the Royal Astronomical Society, 443, 2907 1.3, 4.3

Sale S. E. et al., 2009, Monthly Notices of the Royal Astronomical Society, 392, 497 1.3

Sale S. E., Magorrian J., 2014, Monthly Notices of the Royal Astronomical Society, 445, 256 1.3, 4.5

Savage B. D., Mathis J. S., 1979, Annual Review of Astronomy and Astrophysics, 17, 73 4.3

Schlafly E. F., Finkbeiner D. P., 2011, The Astrophysical Journal, 737, 14 1.3, 3.5

Schlafly E. F. et al., 2012, The Astrophysical Journal, 756, 158 4.1

Schlafly E. F., Finkbeiner D. P., Schlegel D. J., Jurić M., Ivezić Ž., Gibson R. R., Knapp G. R., Weaver B. a., 2010, The Astrophysical Journal, 725, 1175 1.3, 3.5

Schlafly E. F. et al., 2014a, The Astrophysical Journal, 789, 15 1.3

Schlafly E. F. et al., 2014b, The Astrophysical Journal, 786, 29 4.4, 4.11, 4.4, 4.12

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, The Astrophysical Journal, 500, 525 1.3, 3

Schultheis M. et al., 2014, Astronomy & Astrophysics, 566, 1 1.3

Seab C., 1987, in Astrophysics and Space Science Library, Vol. 134, Interstellar Processes, Hollenbach D., Thronson Jr. H., eds., pp. 491–512 1.1

Skrutskie M. F. et al., 2006, The Astronomical Journal, 131, 1163 3.6

Snelson E., 2007, PhD thesis, University of London 5.1, 5.1

Sordo R. et al., 2010, Astrophysics and Space Science, 328, 331 3.1

Speck A. K., Corman A. B., Wakeman K., Wheeler C. H., Thompson G., 2009, The Astrophysical Journal, 691, 1202 1.1

Stubbs C. W., Doherty P., Cramer C., Narayan G., Brown Y. J., Lykke K. R., Woodward J. T., Tonry J. L., 2010, The Astrophysical Journal Supplement Series, 191, 376 4.1

Tielens A. G. G. M., Mckee C. F., Seab C. G., Hollenbach D. J., 1994, The Astrophysical Journal, 431, 321 1.1

Tonry J. L. et al., 2012, The Astrophysical Journal, 750, 99 4.1, 4.2

Trumpler R. J., 1930, Publications of the Astronomical Society of the Pacific, 42, 214 1.1

Vergely J.-L., Valette B., Lallement R., Raimond S., 2010, Astronomy & Astrophysics, 518, A31 1.3

Yanny B. et al., 2009, The Astronomical Journal, 137, 4377 3.1

York D. G. et al., 2000, The Astronomical Journal, 120, 1579 1.3

Zasowski G. et al., 2009, The Astrophysical Journal, 707, 510 4.3

Eigenständigkeitserklärung

Hiermit versichere ich, die vorliegende Arbeit selbstständig verfasst zu haben und keine Quellen oder Hilfsmittel außer den angegebenen verwendet zu haben.

London, den 20. Februar 2017, _____ (Richard Hanson)