Dissertation

submitted to the

Combined Faculty of

Mathematics, Engineering and Natural Sciences

of Heidelberg University, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

Xiangyu Zhang

born in: Baotou, China

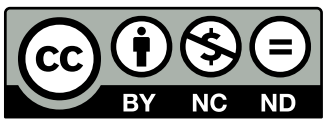
Oral examination: 03-07-2025

On the Variation of Interstellar Dust Extinction

Referees: Dr. Gregory M. Green

Prof. Dr. Ralf S. Klessen

This work is licensed under a Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.



Zusammenfassung

Staub erschwert astronomische Beobachtungen und spielt eine wichtige Rolle in der Entwicklung von Galaxien. Die dreidimensionale Kartierung von Staub ermöglicht nicht nur eine präzise Korrektur von Beobachtungen, sondern liefert auch Einschränkungen für die Entwicklung des interstellaren Mediums (ISM). Extinktion ist die zentrale beobachtbare Größe zur Bestimmung der Dichte und Eigenschaften von Staub und wird häufig in Staubkarten verwendet. Die meisten bestehenden Staubkarten gehen von einer universellen Extinktionskurve aus und behandeln die Staubeigenschaften als einheitlich. Diese Arbeit führt Variationen in der Steilheit von Extinktionskurven ein, parametrisiert durch R(V). Ich entwickle ein datengetriebenes Vorwärtsmodell, das niederaufgelöste Spektren als Funktion von Sternparametern (effektive Temperatur, Oberflächengravitation und Metallizität), Parallaxe und Extinktionseigenschaften vorhersagt. Dieses Modell wird auf alle 220 Millionen Gaia-XP-Spektren angewendet und liefert präzise Extinktionskurven für 130 Millionen Sterne in der Milchstraße, der LMC und der SMC. Daraus erstelle ich die erste dreidimensionale Himmelskarte von Extinktionskurven-Variationen. Unerwarteterweise finde ich, dass die Extinktionskurven in durchscheinenden Staubwolken mit zunehmender Staubdichte steiler und nicht flacher werden. Ich schlage eine theoretische Erklärung vor, die die Steilheit auf das Wachstum polyzyklischer aromatischer Kohlenwasserstoffe (PAHs) zurückführt. Die Ergebnisse liefern zudem wichtige Implikationen für zukünftige Beobachtungen im Zeitalter des JWST.

Abstract in English

Dust obscures astronomical observations and plays a important role in the evolution of galaxies. Mapping dust in three dimensions not only allows accurate correction of observations but also constrains the evolution of the interstellar medium (ISM). Extinction is the central observable to trace the density and properties of dust, and is widely employed in dust mapping. Most existing dust maps assume a universal extinction curve, treating the properties of dust as uniform. This thesis introduces variations in the slope of extinction curves, parameterized by R(V). I construct a data-driven forward model that predicts low-resolution spectra as a function of stellar parameters (effective temperature, surface gravity, and metallicity), parallax, and extinction properties. This model is applied to all 220 million Gaia XP spectra, yielding precise extinction curves for 130 million stars in the Milky Way, LMC, and SMC. From these, I construct the first 3D all-sky map of extinction curve variations. Unexpectedly, I find extinction curves steepen, rather than flatten, with increasing dust density in translucent dust clouds. I propose a theoretical explanation, attributing the steepening to the growth of polycyclic aromatic hydrocarbons (PAHs). The results also provide implications for future observations in the era of the JWST.

Contents

Zusammenfassung **vii**

Abstract in English **ix**

Contents **x**

List of Figures **xiv**

List of Tables **xvii**

1 Dust properties and models **3**

1.1 Abstract 3

1.2 Physics of Extinction 3

1.3 Cross sections and grain size distribution 5

1.3.1 Mie theory and its extensions 5

1.3.2 Size distribution of large grains 8

1.3.3 Extinction cross-sections from $\pi \rightarrow \pi^*$ transitions: small grains 10

1.3.4 Size distribution of small grains (PAHs) 11

1.4 Methodologies for determining extinction curves 13

1.4.1 Pair method 13

1.4.2 Forward modelling 15

1.4.3 Data-Driven forward modeling 16
1.5 Extinction curves 19
1.5.1 IR features 20
1.5.2 Optical-NIR 23
1.5.3 UV features 26
1.6 Evolution of dust 28
1.6.1 Gas-dust interaction 28
1.6.2 Dust-dust interaction 30
1.7 Dust Mapping 31
1.7.1 Dust maps from emission 31
1.7.2 Dust maps from extinction 32
1.8 On this thesis 33

2 Determining extinction with data-driven forward model 35
2.1 Abstract 35
2.2 Data 36
2.2.1 *Gaia* XP spectra 36
2.2.2 LAMOST 39
2.2.3 2MASS 42
2.2.4 unWISE 43
2.2.5 Training and validation datasets 45
2.3 Method 46
2.3.1 Bayesian formulation of model 48
2.3.2 Training the model 52
2.4 Determination of stellar parameters 56
2.4.1 Formulation and priors 57
2.4.2 Optimization of stellar parameters 60
2.4.3 Estimation of uncertainties 61

2.4.4 Validation of stellar parameters 64
2.5 Results 73
2.5.1 Trained model of stellar spectra and extinction 73
2.5.2 220 million stellar parameter estimates 78
2.5.3 External validation of stellar parameters 83
2.5.4 Preliminary map of extinction in 3D 90
2.5.5 Flux residuals 95
2.6 Discussion 98
2.7 Conclusions 102
2.8 Data availability 104
2.9 Gaia Archive Queries 104
2.10 Stellar atmospheric parameter confidence estimates 107

3 A 3D map of extinction-curve variations 111
3.1 Abstract 111
3.2 Adopted datasets 112
3.3 Stellar model and quality cuts 112
3.4 Mapping the distribution of $R(V)$ 114
3.5 Implications for dust properties 119
3.6 Utility for extinction corrections 123

4 A hypothesis on $R(V)$ variations 125
4.1 abstract 125
4.2 An Anti-correlation between $R(V)$ and PAH mass fraction 126
4.3 Theoretical explanation of $R(V)$ variation in the translucent ISM . . 130
4.4 Discussion 135
4.4.1 Constraints on the PAH evolution 135
4.4.2 Implications for future JWST observations 137

4.5 Summary 138
4.6 2D $R(V)$ maps of clouds in the Milky Way 138
4.7 2D $R(V)$ maps for LMC and SMC 140
4.8 Estimation of time scale for PAH accretion 141

5 Conclusion **143**
5.1 Summary 143
5.2 Future work 145

6 List of Publications **149**

Bibliography **150**

7 Acknowledgement **167**

List of Figures

1.1 Sky distribution of polarization fraction 8

1.2 Extinction cross section of a spheroid dust grain 9

1.3 Initial size distribution of two components 12

1.4 Initial extinction curves 13

1.5 Measured $R(V)$ from OB stars. 20

1.6 Measured $R(V)$ from S16. 20

1.7 Measured $R(V)$ from ZYC23. 21

1.8 Measured $R(V)$ from ZG24. 21

1.9 CCM extinction curves 22

1.10 SFD dust map 32

1.11 SFD dust map 33

1.12 Bayestar dust map 34

2.1 Kiel diagram of matched stars. 40

2.2 Neighboring stars in photometric data 44

2.3 The structure of our stellar flux model. 47

2.4 "Self-cleaning" process of our approach 51

2.5 The prior assumption about the distribution of stellar types 59

2.6 Observed parallaxes reduce the correlations among the stellar parameter estimates 62

2.7 Difference between inferred stellar parameters ($T_{\rm eff}$, [Fe/H], $\log g$, E, ϖ) and the values determined by LAMOST, Bayestar19 and GDR3 in our validation dataset. 65

2.8 Residuals normalized by the uncertainties reported by LAMOST, Bayestar19 and GDR3. 66

2.9 Normalized flux residuals ($\triangle f/\sigma \equiv (f_{\rm pred} - f_{\rm obs})/\sigma_f$) in the validation set, as a function of $T_{\rm eff}$, [Fe/H], $\log g$ and E. 67

2.10 Residuals (model – LAMOST) in $T_{\rm eff}$ (top panels), [Fe/H] (middle panels) and $\log g$ for sources flagged as reliable (`good`, left panels) vs. unreliable (`bad`, right panels), as a function of the LAMOST atmospheric parameter estimates. 69

2.11 The distribution of normalized $T_{\rm eff}$ residuals (model minus LAMOST, divided by combined model and LAMOST uncertainties, as a function of LAMOST $T_{\rm eff}$. 70

2.12 Intrinsic flux as a function of $T_{\rm eff}$ 74

2.13 Intrinsic flux as a function of [Fe/H] 75

2.14 Intrinsic flux as a function of $\log g$ 75

2.15 The extinction curve 77

2.16 Residual vs. APOGEE 79

2.17 Residual vs. XGBoost 80

2.18 Fraction of sources that pass our standard reliability cut 81

2.19 Validation of our parallax estimates in the low-SNR regime 86

2.20 Comparison of our reddening estimates with SFD 87

2.21 Comparison of our reddening estimates with a recent `StarHorse` catalog 88

2.22 Sky maps of differential extinction in different distance ranges . . . 91

2.23 Sky maps of differential extinction. 92

2.24 A bird's-eye view of the extinction density across the Galactic plane 93
2.25 Maps of average model – data flux residuals, normalized by observed flux uncertainties. 96
2.26 Maps of average model – data flux residuals, normalized by observed flux uncertainties. 97

3.1 Map of R(V) in the Galactic plane. 115
3.2 Map of $R(V)$ across the sky, in discrete distance bins. 117
3.3 Sky maps of $R(V)$ around five dense clouds. 118
3.4 The dependence of R(V) on extinction in four clouds. 120
3.5 Maps of R(V) in the Magellanic Clouds. 121

4.1 Comparison of 2D distribution of $A(V)$, $R(V)$, $q_{\rm PAH}$ for dust clouds. 128
4.2 Continuation of Figure 4.1 129
4.3 Correlation between $R(V)$ and $q_{\rm PAH}$ 129
4.4 $R(V)$ and $q_{\rm PAH}$ distribution of LMC and SMC 131
4.5 Quantitative comparison on the per-star basis between $R(V)$ and $q_{\rm PAH}$ in LMC (left panel) and SMC (right panel). 132
4.6 $R(V)$ variation as a function of the mass accreted onto astrodust or PAHs relative to initial mass 136

List of Tables

2.1 Training procedure 50

2.2 The columns of our stellar parameter catalog. 82

2.3 Numbers of stars that pass different quality cuts. 83

4.1 Coefficients to select stars into the 2D $R(V)$ maps. 140

4.2 Time scales for typical types of PAH to double their mass by accretion. 142

Chapter 1

Dust properties and models

1.1 Abstract

In this chapter, I review the study of interstellar dust extinction. Beginning with its definition, I derive the connection between extinction and the physical properties of dust, and discuss observational and experimental constraints on these properties. I then review the development of methodologies for determining extinction curves. Observed features of extinction curves are reviewed, including the infrared power law, $R(V)$, broad structures in the optical, diffuse interstellar bands, and the UV bump, as these observations are relevant to this thesis. Finally, I review dust evolution processes and their influence on extinction curves, which provides the basis for explaining our discovery.

1.2 Physics of Extinction

As starlight passes through the interstellar medium (ISM), it is attenuated by scattering and absorption by ”dust,” the solid grains present in the ISM. These two processes—scattering and absorption—are collectively referred to as extinction. This attenuation can be described by the *radiative transfer equation*:

$$\frac{dI_\lambda(\mathbf{n}, s)}{ds} = j_\lambda(\mathbf{n}, s) - \alpha_\lambda(s) I_\lambda(\mathbf{n}, s), \tag{1.1}$$

where $I_\lambda(\mathbf{n}, s)$ is the specific intensity of light at wavelength λ, as a function of direction **n** and distance s. $j_\lambda(s)$ is the *emissivity* of the ISM, and α_λ is the *absorption coefficient* per unit length, which equals the total extinction cross-section of dust:

$$\alpha_\lambda = \int_0^{+\infty} \frac{dn(a, s)}{da} \sigma_\lambda(a) da, \tag{1.2}$$

where $n(s)$ is the number density of grains at wavelength λ with effective grain size $\leq a$. The cross-sections ($\sigma_\lambda(a)$) from both scattering and absorption are functions of the effective size of the grain.

At typical ISM temperatures ($T \sim 20$–200, K), emission is concentrated at mid- to far-infrared (mid-IR to far-IR) wavelengths and is negligible from the ultraviolet (UV) through the near-infrared; that is, $j_\lambda(\mathbf{n}, s) = 0$. Equation 1.1 yields:

$$I_\lambda(\mathbf{n}, s) = I_\lambda(\mathbf{n}, 0) \exp(-\tau_\lambda), \tag{1.3}$$

where $\tau_\lambda \equiv \int_0^s \alpha_\lambda ds = \int_0^{+\infty} da \int_0^{+\infty} ds \, \frac{dn(\sigma_\lambda, s)}{da} \sigma_\lambda(a)$, is the total extinction cross-section of all dust grains at wavelength λ along the line of sight. The quantity τ_λ is referred to as the *optical depth*. The commonly used observable extinction, $A(\lambda)$, is proportional to the optical depth, as defined by:

$$A(\lambda) \equiv -2.5 \log(f_{\rm obs}/f_{\rm intrinsic}) = 1.08 \, \tau_\lambda, \tag{1.4}$$

where $f_{\rm obs}$ is the observed flux at wavelength λ, and $f_{\rm intrinsic}$ is the "intrinsic" flux if there were no dust extinction.

Extinction as a function of wavelength, which is proportional to the optical depth as a function of wavelength, is referred to as the *extinction curve*.

If the distribution of grain size is uniform along the line of sight, that is, if the number density can be separated into a a normalized size distribution and an overall number density term:

$$n(a, s) = n_1(a)n_2(s), \tag{1.5}$$

extinction will be an indicator of the dust column density:

$$A(\lambda) = 1.086\tau_\lambda \propto \int_0^s ds\, n_2(s) \tag{1.6}$$

This idea is widely used in studies of the ISM. For example, dust extinction is often used to trace the column density of gas under the assumption that dust and gas are well mixed. However, this relies on the assumption that the dust size distribution remains the same along the line of sight. As an example that violates this assumption, if dust grows through accretion, the cross-sections increase and thus extinction increases, even though the dust column density remains unchanged.

As the ISM density increases, the evolution of dust can affect the cross section significantly. Changes in extinction can arise from grain size distribution, chemical composition, alignment of dust grains, etc, making extinction a key observable for studying dust physics and chemistry. In the following sections, I review the physics of extinction cross sections and the efforts of modelling on size distribution of dust components.

1.3 Cross sections and grain size distribution

1.3.1 Mie theory and its extensions

To study the physical properties of dust through extinction, it is essential to link the cross sections to the dust's physical morphology and chemical composition. As-

sumptions must be made when relating macroscopic extinction to the microscopic morphology of dust. The zeroth-order approximation assumes that all dust grains are spherical, with a single degree of freedom, the radius "a". In this case, the scattering and absorption cross sections (C_{sca}, C_{abs}) can be calculated using Mie theory, which provides the solution to Maxwell's equations. The total extinction cross section, C_{ext}, is the sum of both cross sections:

$$C_{\text{ext}} = C_{\text{sca}} + C_{\text{abs}} \tag{1.7}$$

In Mie theory, the scattering cross section is

$$C_{\text{sca}} = \frac{\lambda^2}{2\pi} \sum_{n=1}^{\infty} (2n+1) \left(|a_n|^2 + |b_n|^2 \right), \tag{1.8}$$

while the absorption cross section is

$$C_{\text{abs}} = \frac{\lambda^2}{2\pi} \sum_{n=1}^{\infty} (2n+1) \left[\text{Re}\left(a_n + b_n\right) - \left(|a_n|^2 + |b_n|^2 \right) \right], \tag{1.9}$$

where Re() stands for the real part. a_n and b_n are Mie coefficients determined by the material of the dust grain, wavelength λ and the dust grain radius a.

When the grain size a is much smaller than the wavelength λ, such as in the case of dust extinction in the near- or mid-infrared, the extinction predicted by Mie theory approaches the Rayleigh limit:

$$C_{\text{sca}} = \frac{8\pi}{3} \left(\frac{2\pi}{\lambda} \right)^4 \left(\frac{n^2 - 1}{n^2 + 2} \right)^2 a^6. \tag{1.10}$$

Under the Rayleigh limit, the cross section scales as λ^{-4}, meaning that **small grains contribute negligibly to extinction at longer wavelengths.** We refer to this as Conclusion 1.

In the other extreme case, when $a \gg \lambda$, such as for mid- and far-UV photons extincted by large dust grains, Mie theory approaches the "extinction paradox":

$$C_{\text{ext}} = 2\pi a^2, \tag{1.11}$$

which is twice the geometric area. The factor of 2 is known as the "extinction paradox," highlighting the difference between classical and quantum mechanics. In this limit, the extinction cross section becomes independent of the wavelength, so **large dust grains contribute a "grey term" to the extinction.** We refer to this as Conclusion 2.

However, in reality, the assumption of spherical dust grains is over simplistic. Spherical grains cannot account for polarized extinction or emission, both of which require alignment of non-spherical grains. Polarized emission has been observed ubiquitously in all-sky surveys. For example, the Planck survey has revealed large-scale variations in the thermal dust emission (Planck Collaboration et al., 2020a,b,c), as shown in Figure 1.1. Polarized extinction has also been widely detected. Serkowski et al. (1975) observed 180 stars and and derived an empirical relation between the polarization fraction of extincted starlight and wavelength:

$$p(\lambda)/p_{\text{max}} = \exp\left[-1.15 \ln^2\left(\lambda_{\text{max}}/\lambda\right)\right], \tag{1.12}$$

where λ_{max} is the wavelength where the maximum interstellar linear polarization p_{max} occurs.

To account for dust grain alignment without over-complicating the model, researchers usually adopt a first-order approximation: the dust grains are spheroidal, with one axis shorter then another two equal-length axes: $a_1 < a_2 = a_3$. Computation of cross-sections of spheroids is computationally expensive, so I adopt a precomputed library (Draine & Hensley, 2021a) in this thesis, which provides average

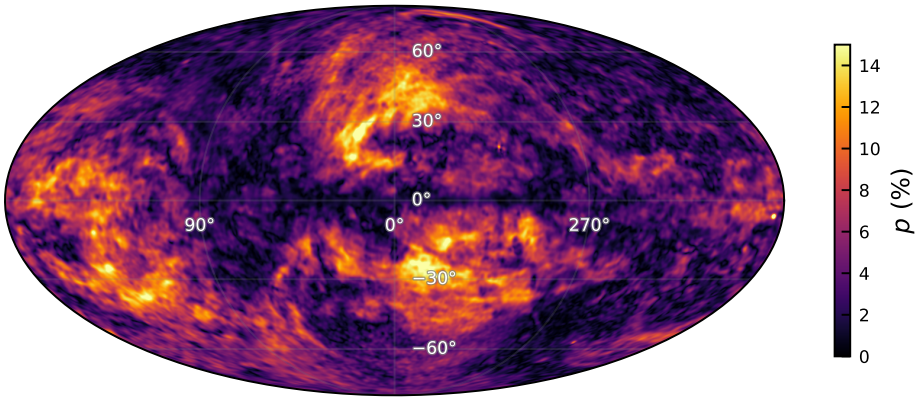


Figure 1.1. Polarization fraction ($p \equiv \sqrt{Q^2 + U^2}/I$) of dust thermal emission from Planck 2018 shown in HEALPIX. In the presence of a magnetic field, dust grains align with their long axes perpendicular to the magnetic filed. Large-scale variations in polarization are evident both above and below the Galactic plane, indicating the distribution of interstellar magnetic filed.

cross sections as a function of the "equivalent size" $a = (a_1 a_2 a_3)^{1/3}$. Despite this added complexity, **Conclusions 1 and 2 still hold**, as demonstrated in Figure 1.2. Therefore, to effectively change reddening ratios, such as $A(V)/E(B-V)$, through extensions of Mie theory, the grain size distribution near the corresponding wavelengths must be modified.

1.3.2 Size distribution of large grains

As discussed in section 1.3.1, at a given wavelength, dust grains much smaller than the wavelength contribute negligibly to extinction, while grains much larger produce a nearly wavelength-independent, or "grey," extinction. As a result, features in extinction curves are effective tracers of dust grains of which sizes are comparable to the corresponding wavelength of light we are interested in ($a = 10 \sim 100$ nm). These grains are referred to as "large grains". Previous studies suggest that large grains have a different chemical composition compared to small grains.

Mathis et al. (1977, "MRN" model) fitted the observed extinction curves with combination of several dust components, and found that the dust grain size distribu-

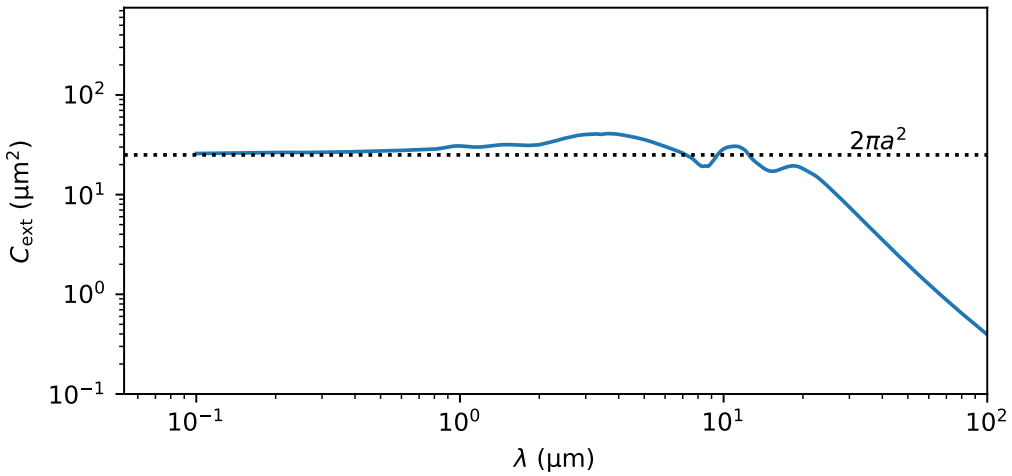


Figure 1.2. Extinction cross section $C_{\rm ext}$ as a function of wavelength for spheroid dust grains with equivalent size $a = 2.0$ μm from the precomputed library by Draine & Hensley (2021a). Dotted line marks the extinction cross-section in the small-wavelength limit, or the "extinction paradox", which equals *twice* of the physical area. Both conclusion 1 and 2 from Mie theory hold in the spheroid case: When $\lambda << a$, $C_{\rm ext} \to 2\pi a^2$. The extinction contributed by large grains become a grey term – independent of wavelength. When $\lambda >> a$, extinction cross section becomes negligible.

tions follow roughly power laws, with an exponent of about -3.3 to -3.6. Building upon the MRN model, Draine & Lee (1984) proposed a model with graphite and silicates. With the grain size model adopted from MRN, the reproduced extinction curves agree well with the observed results (Draine & Lee, 1984). To account the variation of extinction curves, Weingartner & Draine (2001) modeled the extinction curves in the Milky Way, LMC and SMC with two dust populations, a silicate population to account for large grains, and a carbonaceous population for small grains. Zubko et al. (2004) introduced elemental abundance constraints into the dust grain size model, and found that acceptable solutions include composite grains made of silicate, organic material, water ice, and voids. Compiègne et al. (2011) fit the grain size distribution model with the full emission spectra of dust, from mid-IR to the mm range. Guillet et al. (2018) developed the model of Compiègne et al. (2011) to account for polarized emission and extinction.

In this thesis, I adopt the model by Hensley & Draine (2023), in which large grains are represented by a single composite material, "astrodust" (Draine & Hensley, 2021b). The grain size distribution is determined by fitting of extinction, emission and polarization, as shown in Figure 1.3. Combining size distributions and cross-sections, I present the initial extinction curve contributed by astrodust in Figure 1.4.

1.3.3 Extinction cross-sections from $\pi \rightarrow \pi^*$ transitions: small grains

In addition to Mie theory and its extensions, transitions of elections also contribute to extinction. Electrons can be excited by photons, absorbing light across a wide range of wavelengths, from far-UV to optical. Various types of electronic transitions can absorb UV/optical photons, including transitions from non-bonding electron to anti-bonding sigma orbital ($n \rightarrow \pi^*$) and from π bonding electron to a π anti-bonding molecular orbital ($\pi \rightarrow \pi^*$). The $n \rightarrow \pi^*$ needs lower energy, but its probability is much lower than that of the $\pi \rightarrow \pi^*$. Therefore, the $\pi \rightarrow \pi^*$ dominates extinction. This arises from molecules with abundant π bonds, such as polycyclic aromatic hydrocarbons (PAHs). Previous studies indicate that these extinction features come from small grains ($\lesssim 1$ nm).

$\pi \rightarrow \pi^*$ transition have been proposed as the source of the 2175 Å-bump, the most prominent extinction feature that spans near- and far-UV. Joblin et al. (1992) compared observed 2175 Å-feature with laboratory-measured absorption features from PAH mixtures, and found that the mixtures produce broad, instead of sharp features, in UV. In addition, the red wing of the lab-measured features from the PAH mixtures extend to as large as ~ 400 nm, and the extinction of PAH mixtures do not vanish until $\lambda \approx 500$ nm (see FIG. 1 in Joblin et al. (1992)), making $\pi \rightarrow \pi^*$ transition also an important source of extinction in optical wavelengths as well.

$\pi \rightarrow \pi^*$ transition enables PAH molecules to exhibit extinction cross sections at the wavelengths that are much larger than their physical sizes. In those range, the extinction cross sections Mie theory and its extension are negligible.

Numerous studies have modeled or measured the extinction cross sections of PAHs to understand their role in interstellar extinction. Li & Draine (2001) modeled cross sections of PAHs based on laboratory studies and observed PAH emissions in mid- and far-infrared. Malloci et al. (2004) computed the extinction cross sections of 20 types of PAHs as functions of wavelength, and reproduced the laboratory-measured PAH features by Joblin et al. (1992). Malloci et al. (2008) explored the effect of dehydrogenation of PAHs with theoretical calculation. Cecchi-Pestellini et al. (2008) proposed that the charge state of PAHs could produce the variation of the UV-bump and the far-UV rise. Lin et al. (2023) have quantitatively modeled the 2175 feature with combination of pericondensed PAH molecules using quantum chemical computations.

In this work, I adopt the library of PAH cross-sections from Hensley & Draine (2023), which is developed from Draine & Li (2007) and Draine (2016). In this model, the cross sections of PAHs are simplified to depend only on grain effective radius a and the grain charge.

1.3.4 Size distribution of small grains (PAHs)

Observations and experiments have placed constraints on the size distribution of PAHs. By modeling the states of hydrogenation and charge of PAHs in the ISM, Le Page et al. (2003) have found that tiny PAHs (fewer than 15-20 carbon atoms per molecule) are rapidly destroyed. This is supported by the absence of sharp UV absorption features, which are characteristic of tiny PAHs, in observations made using the Hubble Space Telescope (Clayton et al., 2003). Steglich et al. (2010) made mixtures of PAHs in lab that reproduced UV bumps, with its position mainly deter-

mined by the mean size of PAHs. Steglich et al. (2010) inferred that the observed 2175-Å bump can be produced by PAH mixtures with mean size of 50-60 carbon atoms. Steglich et al. (2012) showed that PAH mixtures can reproduce the smooth extinction curves (without sharp features) between $300 < \lambda < 400$ nm.

In this work, I adopt the PAH size distribution of Hensley & Draine (2023) as the initial distribution, which is developed from Draine et al. (2021):

$$\frac{1}{n_{\rm H}}\frac{dn_{\rm PAH}}{da} = \sum_{j=1}^{2}\frac{B_j}{a}\exp\left\{-\frac{\left[\ln\left(a/a_{0,j}\right)\right]^2}{2\sigma^2}\right\}, \tag{1.13}$$

where $n_{\rm H}$ is the number density of PAHs. a is the effective radius. $a_{0,1} = 4.0$, $a_{0,2} = 30$, and $\sigma = 0.40$. I adopt $B_1 = 7.52\times10^{-7}$ H^{-1} and $B_2 = 8.09\times10^{-10}$ H^{-1} from Hensley & Draine (2023). The size distribution of PAHs is shown in Figure 1.3.

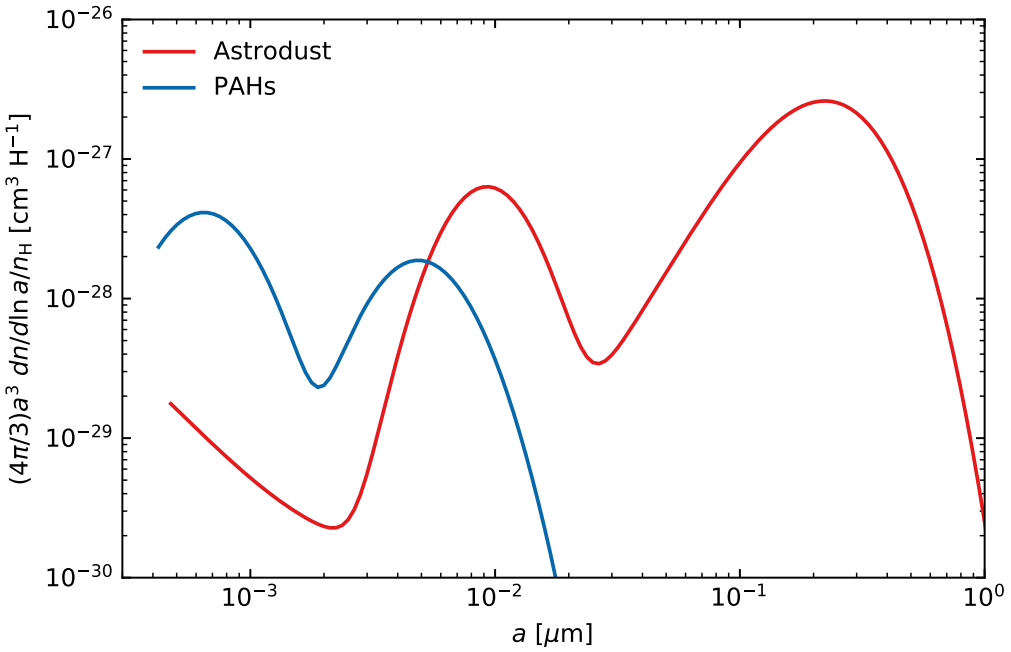


Figure 1.3. Initial size distribution of two components in the model adopted by this thesis (Hensley & Draine, 2023). x axis is the effective radius of dust grains, and y axis is the equivalent volume of dust grains per unit $\ln a$ per Hydrogen atom in the ISM.

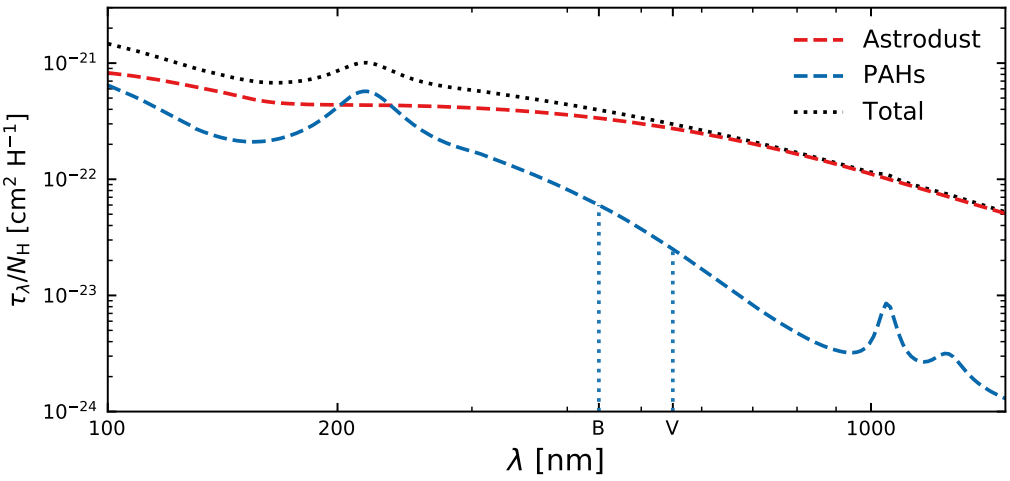


Figure 1.4. Extinction curves contributed by PAHs (red dashed line) and astrodust (blue dashed line). y-axis is the optical depth per Hygrogen atom in the ISM. x-axis is the wavelength.

1.4 Methodologies for determining extinction curves

By definition, extinction is calculated by comparing the observed spectra with the inferred intrinsic spectra, i.e., the spectra of the object if it were placed at the same distance without any intervening dust. There are two primary methods for inferring intrinsic spectra: "pair method" and "forward modeling". In this section, I discuss the methods used to measure extinction curves and review significant discoveries made through these approaches. Finally, I introduce the concept of data-driven forward modeling, which combines the advantages of both approaches.

1.4.1 Pair method

The pair method compares the spectra of a reddened target star with those of a "standard reference star" of similar spectral type using the same instrument (Massa et al., 1983; Stecher, 1965). The key advantage of this method is its direct cancellation of instrumental systematics - particularly important when combining multi-instrument datasets that cover different wavelength ranges. Additionally, the method requires

minimal sample sizes - In principle, a single pair of reddened and standard reference stars should suffice. However, this method faces three main limitations:

- Finding reference stars with identical types can be challenging;
- The spectral features of the reference stars may contaminate the derived extinction curves;
- Stellar extinction becomes negligible only for stars in close proximity to the Sun, which significantly limits both the number of suitable reference stars and their spectral types.

Many significant discoveries have been made using the pair method. Fitzpatrick & Massa (1986, 1988) compared 45 reddened stars with 10 standard stars and derived their extinction curves in UV using the International Ultraviolet Explorer (IUE). Fitzpatrick & Massa (1986) found that the 2175-Å bump position is stable, but the bump width (FWHM) can vary from 360 Å to 600 Å. The width of the bump correlates positively with the density of the dust (Fitzpatrick & Massa, 1986). Fitzpatrick & Massa (1988) found that all 45 of their UV extinction curves can be represented by a six-parameter model, known as the FM parameterization of the UV extinction. Valencic et al. (2004) derived 417 UV extinction curves in the Milky Way using pair method, comparing their IUE spectra with the reference stars from Cardelli et al. (1992). Despite 93% of the sightlines can be described by CCM parameterization, Valencic et al. (2004) reported that 4 sightlines in dense molecular clouds significantly deviate from CCM. Using pair method, Gordon et al. (2003) derived 5 sightlines in the SMC, and 19 sightlines in the LMC. Gordon et al. (2003) found that the majority of the LMC/SMC extinction curves do not follow the CCM relationship. Gordon et al. (2009) used pair method for 75 stars and found strong correlation between $a(\lambda)/A(V)$ and $1/R(V)$ in the far-UV wavelength range.

Large spectroscopic surveys have significantly facilitated the derivation of extinction curves using the pair method by providing precise measurements of stellar

spectral types. Yuan et al. (2013) built a low-extinction reference samples of 50,053 stars using stellar parameters derived from SDSS DR7 spectra (Abazajian et al., 2009). With these reference stars, Yuan et al. (2013) determined reddening of 9,202 higher-extinction stars using photometric surveys. Zhang et al. (2024, no relation) used common stars of LAMOST DR7 (Wu et al., 2011) and Gaia XP spectra (De Angeli et al., 2022; Montegriffo et al., 2022). Zhang et al. (2024) built the reference sample of 67,650 stars, and used these stars to derive extinction curves of 370,000 high-quality sources.

1.4.2 Forward modelling

The forward modeling method derives intrinsic spectra from theoretical or empirical models, enabling simultaneous determination of stellar types and extinction parameters while providing self-consistent estimates for both properties. Unlike the pair method, this approach can analyze any star *within the parameter space covered by the training set* without requiring specific low-extinction counterparts. The main challenges of forward modeling include:

- Flux calibration is necessary;
- Empirical models require larger and more representative training sets, compared to the pair method;
- Potential systematic errors when using theoretical models that may not fully reproduce observed spectral features;
- Difficulties in dealing with systematic errors when combining data from multiple instruments.

This method has been developed over a long period. Whiteoak (1966a) determined the intrinsic spectra of their samples by fitting the observed Balmer discontinuities with theoretical models. Fitzpatrick & Massa (2005) developed a method named “extinction without standards”, in which they formulate the observed flux

($f_{\rm obs}(\lambda)$) as:

$$f_{\rm obs}(\lambda) = f_{\rm abs}(\lambda, T_{\rm eff}, \log g, [{\rm M/H}], v_t)\theta_R^2 10^{-0.4E(B-V)[k(\lambda-V)+R(V)]}, \quad (1.14)$$

where $[{\rm M/H}]$ is the metallicity, v_t is the microturbulence velocity, and θ_R is the angular size of the star. $f_{\rm abs}$ is predicted by the theoretical model "Kurucz ATLAS9" by Kurucz (1991). Fitzpatrick & Massa (2005) determined both stellar types and the extinction curves simultaneously by fitting the observed spectra with Equation 1.14.

Fitzpatrick & Massa (2007) applied this method to 328 OB stars, and found that the width of the 2175-Å feature varies significantly, while the variation of the central wavelength is mild. Fitzpatrick & Massa (2007) also found that no clear correlation between $R(V)$ and the shape of the UV bump in the majority of their sightlines. However, for sightlines with highest $R(V)$, Fitzpatrick & Massa (2007) found that the strength of the bump is anti-correlated with $R(V)$. In the following work, Fitzpatrick & Massa (2009) studied the near-IR extinction curves and confirmed that the exponent in the power law of IR extinction, β, is anti-correlated with $R(V)$. Fitzpatrick et al. (2019) derived 72 extinction curves with this method, and found various structures on the extinction curves that are independent of $R(V)$.

1.4.3 Data-Driven forward modeling

Studies using forward modeling have been limited to O/B stars due to their high UV brightness. OB stars also exhibit relatively simple spectra that are easier to model with theory. However, O/B stars are sparsely distributed in the Milky Way, making it difficult to trace the 3D variation of dust properties. Figure 1.5 shows the distribution of measured extinction curves towards O/B stars by Valencic et al. (2004) and Gordon et al. (2023). In contrast, modern photometric and spectroscopic surveys include numerous non-O/B stars that have proven valuable for extinction

studies through the pair method (see Section 1.4.1).

Over the past decade, significant progress has been made through the data-driven forward modeling, which combines the advantages of both method. The general idea is to simultaneously learn intrinsic spectra and extinction curves directly from observational data while minimizing assumptions on their mathematical forms. This method has two key advantages:

- **Self-consistency**: Both intrinsic spectra and extinction curves are learned directly from the data, with instrumental systematics accounted.
- **Population-based learning**: Both extinction and spectral properties are inferred from the large number of spectra in the training set, which avoids bias from individual reference stars.

Combining APOGEE (Majewski et al., 2017) stellar parameters and photometric surveys Pan-STARRS1 (Hodapp et al., 2004), 2MASS (Skrutskie et al., 2006) and WISE (Wright et al., 2010), Schlafly et al. (2016) measured variation of extinction curves for 37,000 stars. Like forward modeling, they formulate the observed photometric data as:

$$\boldsymbol{m}_i = \boldsymbol{f}\left(T_{\text{eff,i}}, [\text{Fe/H}]_i\right) + \mu_i + \boldsymbol{R}_0 E_i', \tag{1.15}$$

where $\boldsymbol{m}_i$ is the vector of the 10 photometric bands of the star. i is the index of stars. $\mu_i \equiv 5\log(d/10\text{ pc})$ is the distance modulus. $\boldsymbol{R}_0$ is the extinction curve, and E_i' is a scalar that represents the amount of extinction. $\boldsymbol{f}\left(T_{\text{eff,i}}, [\text{Fe/H}]_i\right)$ represents the "intrinsic" spectra (unreddened, observed at 10 pc) as a function of effective temperature and metallicity. Instead of using theoretical models, Schlafly et al. (2016) assumed $\boldsymbol{f}$ is unknown, and parameterized it using polynomial representation, with the coefficients learned from the data.

This approach is equivalent to modeling the intrinsic spectra observed through

the photometric instruments, which avoids the complications from instrumental biases as in the traditional forward modeling. Unlike the pair method, this method allows extinction curves to be constrained for stars that do not have unreddened counterparts.

Schlafly et al. (2016) found $R(V)$ variations that have significant larger scales than the typical size of molecular clouds. Schlafly et al. (2016) also found anti-correlation between $R(V)$ and spectral index (β)[1] of the dust emissivity, which set constraints on the dust evolution models as both parameters are tracers of grain size distribution.

The key limitation of the work by Schlafly et al. (2016) is the cross-matching between photometric surveys and the APOGEE spectroscopic survey for precise stellar parameters. Their selected 37,000 stars represents $\sim 1/4$ of all APOGEE DR12 sources, but are only $\sim 10^{-4}$ of the stars in the photometric surveys.

Addressing this limitation, Green et al. (2021) developed a data-driven model that expands the applicability of data-driven forward modeling by predicting intrinsic spectra directly from photometric data. Ideally, when there are more photometric bands than unknown physical parameters for each star, all parameters can be fully constrained. In practice, while the photometric bands of a star are correlated, and therefore more bands are required for the parameter determination, the available number of photometric bands often exceeds the necessary number of parameters. This is especially true with the increasing depth and coverage of modern photometric surveys. Green et al. (2021) trained their model with stars cross-matched with spectroscopic surveys (LAMOST, APOGEE and GALAH), and used the model to infer parameters from photometry, without relying on higher-resolution spectroscopic data. Incorporating *Gaia* parallaxes enables them to predict the absolute flux, as well as absolute extinction curves instead of reddening. Green et al. (2021) showed that

[1]Emissivity is also denoted as β, same as infrared power law index of extinction curves. I will specify the physical meaning when either is being discussed.

the mean extinction curve learned from data is consistent with Schlafly et al. (2016), if converted to reddening.

The idea of this thesis is developed from the data-driven forward model by Green et al. (2021). We have published the work as Zhang et al. (2023b) and Zhang & Green (2025), which are discussed in Chapter 2 and Chapter 3.

To conclude this section, I summaries the measurements of extinction curves, parameterized by $R(V)$, in Figure 1.5, Figure 1.6, Figure 1.7 and Figure 1.8 for Valencic et al. (2004), Gordon et al. (2023), Schlafly et al. (2016), Zhang et al. (2024) separately.

1.5 Extinction curves

In this section, I summarize key features of extinction curves and their empirical parameterizations, followed by a brief review of discovered correlations among these features. These correlations reveal physical or chemical interactions between different dust components, making them central to the observational study of dust extinction. Although this thesis aims at measuring extinction curves via data-driven model, previous works provide important implications to our work.

The most widely recognized extinction parameter is $R(V) \equiv A(V)/E(B-V)$, which correlates with dust properties across the far-UV to near-IR spectrum. Cardelli et al. (1989) proposed the canonical parameterization of extinction curves by $R(V)$ from UV to near-infrared range (the "CCM" model; see Figure 1.9). Cardelli et al. (1989) parameterize extinction curves as

$$A(\lambda)/A(V) = a(x) + b(x)/R(V), \tag{1.16}$$

where $x = 1/\lambda$. $a(x)$ and $b(x)$ have different forms in different wavelength ranges. Given the demonstrated success of the CCM model, I adopt it as our framework for

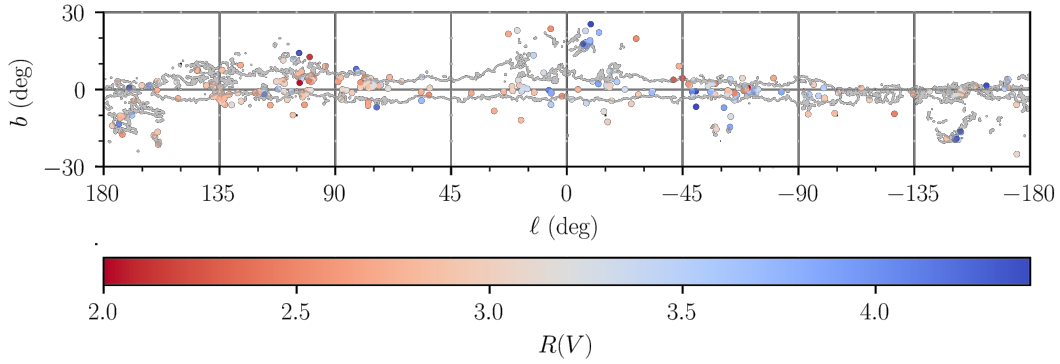


Figure 1.5. Measured $R(V)$ from O/B stars by Valencic et al. (2004) and Gordon et al. (2023), plotted in galactic coordinates. The grey contour shows the contour of high extinction regions (with the SFD extinction $E = 0.5$ mag). Extinction curves measured with high-resolution spectroscopic data of OB stars provides detailed structures, but are limited by the sparse distribution of OB stars.

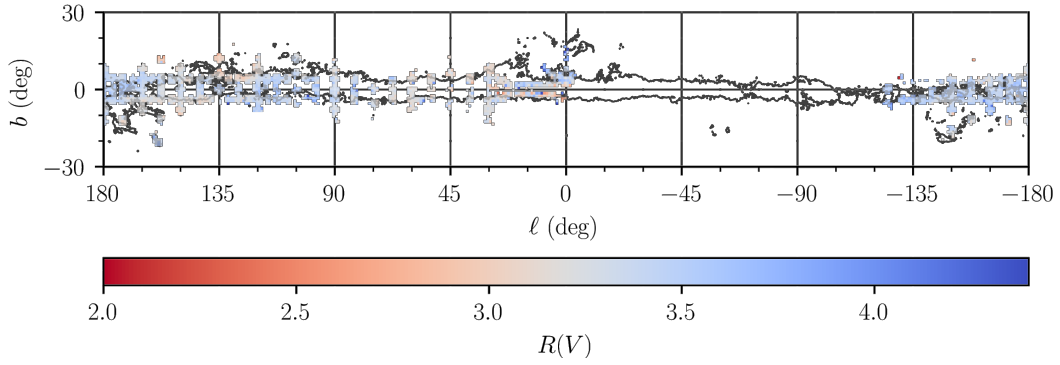


Figure 1.6. Measured $R(V)$ by Schlafly et al. (2016) using cross match results of APOGEE and photometric data. Utilizing forward modelling, they measured $R(V)$ of 37,000 stars. While the coverage of the sky is limited by APOGEE footprint, this result shows large-scale variation on $R(V)$.

discussion. In the following sections, I first discuss the CCM parameterization of different wavelength ranges, and then review the development based on CCM.

1.5.1 IR features

Extinction is generally weak in the infrared range, as the wavelength ($> 1\mu$m) is larger than most of the interstellar dust (typically $\leq 0.1\mu$m). According to our discussion in Section 1.2, extinction cross-sections follow a power law (λ^{-4}) in the Rayleigh limit, so it is not surprising that IR extinction curves are usually modeled

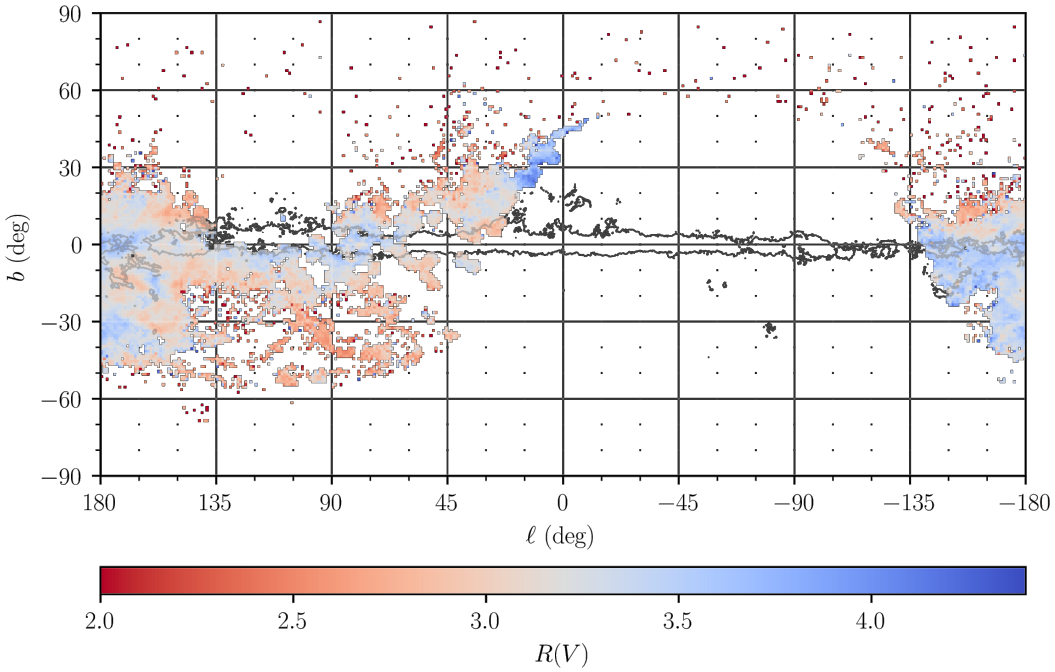


Figure 1.7. Measured $R(V)$ by Zhang et al. (2023a) using forward modeling. Despite being limited by LAMOST footprint, more structures have been revealed.

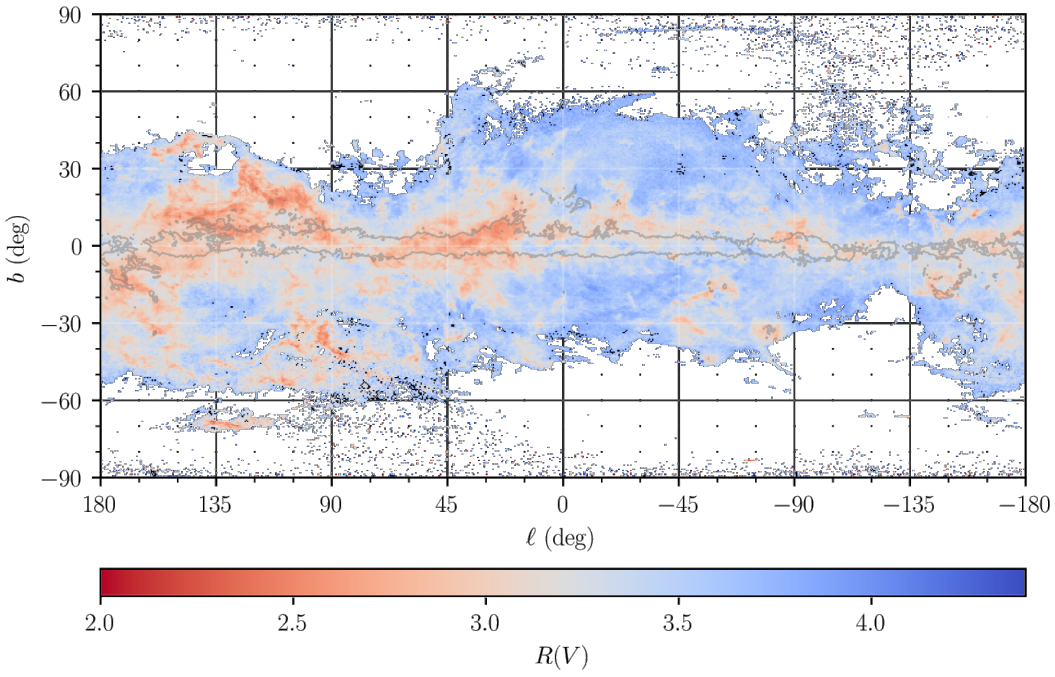


Figure 1.8. We measure $R(V)$ using data-driven forward modeling (Zhang & Green, 2025), where intrinsic spectra and extinction curves are inferred simultaneously from observed *Gaia* XP spectra, enabling all-sky measurement of $R(V)$.

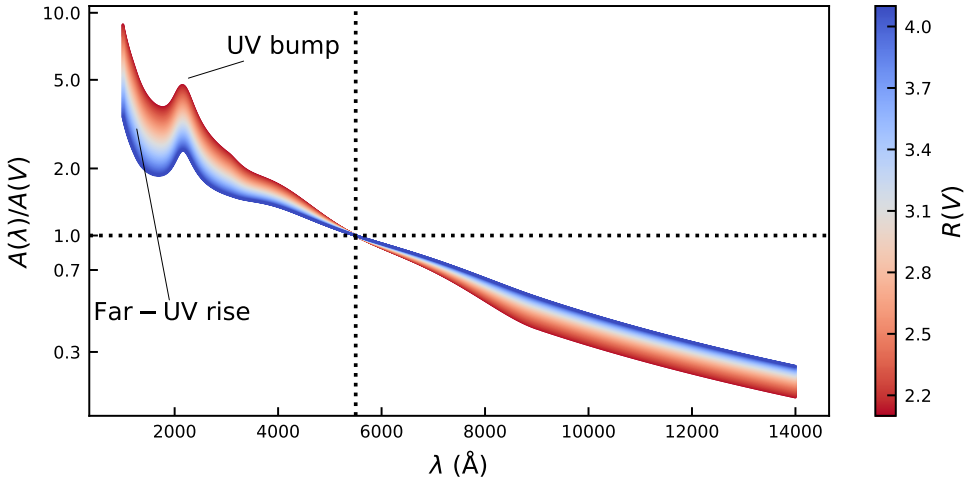


Figure 1.9. Extinction curves from Cardelli et al. (1989), color-coded by $R(V)$

as a power law:

$$A(\lambda) \propto \lambda^{-\beta} \tag{1.17}$$

In the CCM model, both $a(x)$ and $b(x)$ are proportional to $\lambda^{-1.61}$ for near-infrared extinction curves, which is adopted from Rieke & Lebofsky (1985). However, many different values have been reported. To list a few, Draine (1989) found that $\beta = 1.75 \pm 0.2$ applies for their observation data both in the Milky Way and the Magellenic Clouds. Martin & Whittet (1990) reported the value of $\beta = 1.8$ using sightlines of diffuse regions and towards the dense region of ρ-Oph. Various studies have shown higher β towards the Galactic center: Schödel et al. (2010) derived $\beta = 2.21 \pm 0.24$, and Fritz et al. (2011) reported the value of $\beta = 2.11 \pm 0.06$ using hydrogen emission lines. Wang & Chen (2019) found the index $\beta = 2.07 \pm 0.03$ using 2MASS and WISE photometry.

Many of these values do not agree with one another even within 2 or 3 σ. Addressing this problem, Martin & Whittet (1990) suggested that the power-law index might depend systematically on the range of passbands, which is confirmed by Fitzpatrick & Massa (2009). Using data from the Hubble Space Telescope, Fitzpatrick

& Massa (2009) found that β varies significantly among different sightlines, though extinction curves follow the power law. Fitzpatrick & Massa (2009) also reported anti-correlation between β and $R(V)$, suggesting the connection of IR extinction and the grain size distribution. Fitzpatrick & Massa (2009) proposed an alternative expression of the IR extinction curve:

$$k(\lambda - V) = [0.349 + 2.087R(V)]\frac{1}{1 + (\lambda/0.507)^{\alpha}} - R(V), \tag{1.18}$$

where α is their alternative power-law index, similar to β.

Mid- and far-UV features of extinction are complicated: They are even weaker than near-IR, and are mixed with dust emission features (e.g. PAH emission features at 3.3, 6.2, 7.7, 8.6, 11.2, 12.7, and 16.4 μm (Tielens, 2008)). This is beyond the scope of this thesis.

Informed by the studies of IR extinction curves, I construct our data-driven forward models assuming near-IR extinction are functions of $R(V)$ (Zhang & Green, 2025). However, I do not impose prior knowledge to the shapes of the extinction curves from literature, which are expected to be learned from the data.

1.5.2 Optical-NIR

The optical-NIR extinction curves are more complicated than Mid-IR, as extinction in optical-NIR is the combined effect of various absorption features. Carriers of many features remain unknown, but observations have set constraints on the physics behind those features.

CCM parameterized $a(x)$ and $b(x)$ in Equation 1.16 by 7th-order polynomials of $x \equiv 1/\lambda$. However, CCM only had 5 photometric bands (UBVRI) in the optical-NIR range, which are not enough to derive fine structures in this range. The CCM model has been improved by O'Donnell (1994) in the wavelength range of $1.1\ \mu\text{m}^{-1} <$

$1/\lambda < 3.3\ \mu\mathrm{m}^{-1}$, with more photometric bands. The most prominent feature in the optical wavelengths is $R(V)$ itself, but according to its definition, $R(V)$ only describes the slope of the extinction curve. Multiple structures have been found in the optical-NIR range, and I briefly review them in the following sections.

1.5.2.1 Very broad structure (VBS) and Intermediate scale structure (ISS)

The "very broad structure" was found and named by Whiteoak (1966b). They discovered "depression" between 500 nm and 625 nm towards extincted stars in Perseus, Orion, Monoceros, Cygnus, and Cepheus dust clouds. VBS has been found in various sightlines (Krelowski et al., 1986; Schild, 1977; Walker et al., 1980; Whittet et al., 1976). Massa et al. (2020) identified three "intermediate scale structures" at 4370, 4870, and 6300 Å, and modeled the with three Drude profiles:

$$D\left(x, x_i, \gamma_i\right) = \frac{x^2}{\left(x^2 - x_i^2\right)^2 + x^2\gamma_i^2}, \tag{1.19}$$

where i is the index of profile. x_i and γ_i are free parameters to determine the shape of the profiles. Maíz Apellániz et al. (2021) found a broad feature near 7700 Å. Zhang et al. (2024) found bump structures on their extinction curves at 487, 540, and 769 nm [2] with LAMOST and Gaia XP spectra. Utilizing decomposition technique in magnitude space and extinction curves derived from Gaia XP spectra, Green et al. (2024) found two bump structures at 770 and 850 nm, in addition to the ISS and the VBS.

The carriers of the VBS and ISS remain unknown by the time of writing this thesis (2025). Observations, especially their correlation, have set constraints on the potential carriers. Massa et al. (2020) found that the amplitudes of the features centered at 4370 and 4870 Å positively correlates with the strength of the 2175-Å fea-

[2] This is the same structure as the 7700-Å structure mentioned above, but it is named differently in different works.

ture, indicating their potential connection to carbonaceous grains. Maíz Apellániz et al. (2021) found that the 7700-AA feature strength is weaker towards sightlines rich in molecular carbon, indicating chemical connection between the carrier with carbon. Gordon et al. (2023) found positive slopes (with respect to $\frac{1}{R(V)} - \frac{1}{3.1}$) for the 4370 and 4870 feature, and slightly negative slope for the 6800 faeture. Zhang et al. (2024) found no correlation between $R(V)$ and their 487- or 540-nm features. [3] Green et al. (2024) found that the strength of the 770- and 850-nm features positively correlation between $R(V)$ and that the VBS anti-correlates with $R(V)$.

1.5.2.2 Diffuse Interstellar bands

Diffuse interstellar bands (DIBs) are absorption features that are broader than atomic absorption lines in stellar spectra, but narrower than broad extinction features such as the VBS. The first DIB was discovered a century ago (Heger, 1922), and since then more than 500 DIBs have been reported (Snow, 2014). Most of the confirmed DIBs are at optical-NIR wavelength range, though discovery of DIBs in UV and infrared have been made.

The carriers of most DIBs remain unknown. The ubiquity of DIBs indicate that they are from common elements, and their correlation with various physical parameters suggests their carriers being reactive, which make organic molecules the most popularly believed carriers. There is only one confirmed carrier, C_{60}^{+} confirmed by Campbell et al. (2015) using laboratory spectroscopy.

In addition to chemical composition tracers, DIBs can be used to measure the radial velocity of the ISM, and set constraints on ISM dynamics (e.g. Gaia Collaboration et al. (2023a)).

[3] This contradicts with the conclusion of Gordon et al. (2023) and needs further investigation.

1.5.3 UV features

Extinction is strong in the UV range. In addition to extinction from transition of electron, larger grains contribute a grey term of extinction ($\sim 2\pi a^2$) that is not negligible (see Figure 1.2). Fitzpatrick & Massa (1988) presented the canonical form of the parameterization of UV extinction, adopted by various study:

$$A(\lambda)/A(V) \text{ or } E(\lambda-V)/E(B-V) = c_1+c_2\lambda^{-1}+c_3 D\left(\lambda^{-1},\gamma,\lambda_0^{-1}\right)+c_4 F\left(\lambda^{-1}\right), \tag{1.20}$$

where the first two terms are referred to as the "background". c_1 is the grey term that is independent of the wavelength. The second term is inversely proportional to wavelength. The term with c_3 is a Drude bump profile that describes the 2175-Å bump:

$$D\left(\lambda^{-1},\gamma,\lambda_0^{-1}\right) = \frac{\lambda^{-2}}{\left(\lambda^{-2}-\lambda_0^{-2}\right)^2+\gamma^2\lambda^{-2}}, \tag{1.21}$$

where γ and λ_0^{-1} are the parameters that determines the width and the position of the bump. The term with c_4 describes the far-UV rise:

$$F\left(\lambda^{-1}\right) = \begin{cases} k_2\left(\lambda^{-1}-5.9\right)^2+k_3\left(\lambda^{-1}-5.9\right)^3 & \text{for } \lambda^{-1} > 5.9\ \mu\text{m}^{-1}, \\ 0.0 & \text{for } \lambda^{-1} < 5.9\ \mu\text{m}^{-1}, \end{cases} \tag{1.22}$$

where k_2 and k_3 are coefficients that vary among different models. The far-UV rise vanishes when $\lambda^{-1} < 5.9\ \mu\text{m}^{-1}$, or $\lambda > 1695$ Å.

Cardelli et al. (1989) has found that the $R(V)$ correlates negatively with the bump strength of the 2175-Å feature normalized by $A(V)$, which is adopted by Fitzpatrick (1999) in their model. Gordon et al. (2023) also found correlation between the bump strength and $R(V)$ - In their parameterization, the slope of $A(\lambda)/A(V)$ versus $1/R(V)$ is positive. This correlation is consistent with the extinction curves

observed in lab that the extinction cross-sections of PAH mixtures do not vanish even at $400 \sim 500$ nm Joblin et al. (1992, e.g.). In Chapter
refchapter4, I propose an explanation that the red wing of the UV bump extends to B band and significantly affects $R(V)$. This part is also published as a paper (Zhang et al., 2025).

1.5.3.1 UV extinction features outside of the Galaxy

While this thesis primarily focuses on dust extinction variations within the Milky Way, our results provide important constraints on the dust evolution processes and the interpretation of extinction in extragalactic systems. In this section I review the observational progress of UV extinction features of extragalactic sources.

The UV bump has been observed outside of the Milky Way for long. Some sightlines in SMC and LMC exhibit the 2175-bump feature (Gordon et al., 2003, 2024, e.g.). Motta et al. (2002) discovered the UV bump in a galaxy at the redshift $z = 0.83$. More detections have been made in low- and intermediate-redshift galaxies ($z \lesssim 3$) (Buat et al., 2011; Conroy, 2010; Elíasdóttir et al., 2009; Noll et al., 2009; Shivaei et al., 2022; Wild et al., 2011). UV bumps are also observed in many high-redshift ($z \gtrsim 6$) galaxies. Witstok et al. (2023) detected the UV bump at a Galaxy of $z = 6.7$. Markov et al. (2023) found a galaxy with the UV bump at $z = 6.9$. Markov et al. (2025) reported the detection of a galaxy at $z = 7.55$. Fisher et al. (2025) reported detection of UV bump at three galaxies with $z \sim 7$. Ormerod et al. (2025) reported the detection in a galaxy at $z = 7.11$.

The detection of UV bumps in high-redshift galaxies challenges traditional theories that PAH formation is dominated by low-mass asymptotic giant branch (AGB) stars. These galaxies are in such early stage (~ 700 Myr after the big bang) that there is not enough time for low-mass stars to evolve into AGB, which takes more than 1 billion years (Yang & Li, 2023). To account for this paradox, Witstok et al. (2023)

suggested that there are rapid channels of PAH formation, such as Wolf–Rayet stars or supernova ejecta. However, questions remain in the formation of PAHs. Wolf-Rayet stars are too rare to account for strength of the feature, and reverse shock from supernova destroys tiny dust grains, from which UV bump arises. As Yang & Li (2023) argued, supernova-condensed graphite grains found in meteorites are too large, and there should be ways of tiny PAH grains forming from those graphite if they are from SNe.

1.6 Evolution of dust

Dust is not static, but is actively interacting with gas and other dust grains in the ISM. Dust evolution is concluded as four processes:

- **Accretion**: Dust grains gather metal elements from the gas;
- **Coagulation**: Small dust grains collide and merge into larger dust grains;
- **Sputtering**: Dust grains collide with hot gas and lose mass;
- **Shattering**: Dust grains collide with other grains and fragment.

In the following sections, I discuss the properties of these processes, with a focus on factors that affect their efficiency.

1.6.1 Gas-dust interaction

Accretion and sputtering are both interactions between gas and dust grains, but with opposite effect. Both effects happen at the surface of the dust grains, and therefore the speed of gaining (or losing) mass ($d\,m/dt$) is proportional to the surface area of the dust:

$$\frac{dm}{dt} = \frac{4\pi}{3}\rho_{\text{dust}}a^2\frac{da}{dt} \propto 4\pi a^2, \tag{1.23}$$

where ρ_{dust} is the density of the dust, and a is the effective radius of the dust grains. a^2 can be canceled from both sides, and therefore $d\,a/dt$ is independent of a. This

means that dust grains will gain (or lose) the same amount in radius, regardless of their initial size. The following discussion will be based on $d\,a/dt$.

In the following analysis, I adopt the formulation of McKinnon et al. (2018). For accretion,

$$\frac{\mathrm{d}a}{\mathrm{d}t} \approx \left(\frac{Z}{Z_\odot}\right)\left(\frac{n_\mathrm{H}}{10^3\ \mathrm{cm}^{-3}}\right)\left(\frac{T}{10\ \mathrm{K}}\right)^{1/2}\left(\frac{S_\mathrm{acc}}{0.3}\right)\ \mathrm{nm\ Myr}^{-1} \tag{1.24}$$

where the square root of gas temperature $\sqrt{T}$ accounts for the number of gas particles that collide with the dust surface, assuming gas velocity follows Maxwellian distribution. $Z_\odot = 0.0127$ is the solar metallicity. Z is the local metallicity of the ISM. n_H is the local number density of Hydrogen. S_acc is the probability of dust capturing the particle, the "sticking" probability. The sticking probability depends on the material of the dust, gas temperature, velocity of particles, incident angles, etc. Despite its complexity, the sticking probability is not expected to change over orders of magnitudes. McKinnon et al. (2018) adopts 0.3 from Hirashita & Kuo (2011).

For intermediate-density/translucent ISM, I assume metallicity $Z = Z_\odot$, number density $n_H = 10^2$ cm^{-3} and temperature of 50 k. Equation 4.2 yields an accretion rate of 0.22 nm Myr^{-1}. During the lifespan of a molecular cloud (typically ~ 10 Myr), grains grow larger by 2 nm through accretion. [4] In denser regions, accretion can be accelerated by higher gas density, but the lower temperature can slow it down. Metallicity can be enriched, e.g. by supernova ejecta (Silvia et al., 2012). This amount of variation is negligible for large grains (such as large silicate grains with ~ 100 nm), but could significantly change the distribution and the total mass of small grains. For example, PAHs with ~ 50 carbon have the typical size of 0.5 nm, which is the average size of the carrier of the 2175-Å feature (Steglich et al., 2010).

[4] The actual amount may be lower due to depletion of metal elements, but this estimation generally agree with literature values in order of magnitude, such as Hirashita (2012) and de Bennassuti et al. (2014).

For sputtering,

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\left(1.02 \times 10^3 \text{ cm}^3\right)\left(\frac{\rho_\mathrm{g}}{m_\mathrm{p}}\right)\left[\left(\frac{T_\mathrm{sput}}{T}\right)^{2.5} + 1\right]^{-1} \text{ nm Myr}^{-1}, \qquad (1.25)$$

where ρ_g is the density of the gas. m_p is the mass of proton. $T_\mathrm{sput} \equiv 2 \times 10^6$ K is the scale temperature.

In the translucent ISM, following our assumption above, Equation 1.25 yields an sputtering rate of -3×10^{-7} nm Myr^{-1}. The negative sign represents decrease of the grain size. The absolute value is smaller than accretion by 10^6 in the translucent regions. Denser regions have lower temperature, making it even slower. Sputtering is prominent only when the temperature dependent term approaches 1, or $T \gtrsim 10^6$ K, which is limited to plasma (McGee & Balogh, 2010) or SN shocks (Silvia et al., 2012). Therefore I conclude that **in diffuse and translucent (intermediate-density) ISM, the interaction between gas and dust grains is dominated by accretion.**

1.6.2 Dust-dust interaction

Dust-dust interactions contains coagulation and shattering. I discuss these effects qualitatively. The rate of coagulation depends on the velocity of dust. If the velocity is too high, energies of the grains will be too high to capture. The critical velocity below which coagluation can happen is referred to as "coagulation threshold" (Chokshi et al., 1993). The threshold velocity depends on grain size, grain morphology, grain surface property, etc. Generally, larger grains have smaller threshold velocity. Meanwhile, larger grains also tend to have larger velocity dispersions. Both effects prevent dust from growing to infinity large. Hirashita & Yan (2009) estimated that coagulation becomes negligible for dust grains with $a > 200$ nm. Hirashita (2012) estimated that the time scale[5] of coagulation ($\lesssim 10$ Myr) is comparable with the

[5]Coagulation time scale is defined as the expected time for a dust grain to merge with another.

lifespan of molecular clouds.

Shattering, in contrast, can only happen with high-velocity grains. The threshold velocity (lower limit) for shattering to happen is estimated to be $1\ \mathrm{km/s}$ (Jones et al., 1996), which is difficult to reach by thermal motion of dust. In translucent cloud, $T \sim 50$ K, and I assume the typical dust grain mass[6] is 1.3×10^{-14} g. The velocity dispersion from thermal motion is $\sqrt{kT/m} = 7.4\ \mathrm{mm/s}$, which is far smaller than the threshold velocity of shattering. However, large grains can be accelerated to the threshold due to turbulence (Hirashita & Yan, 2009, e.g.) and SN shocks (Dwek & Scalo, 1980, e.g.).

1.7 Dust Mapping

This thesis aims to measure the variation of extinction curves across the Milky Way, LMC and SMC. One of the most important (and direct) applications of the result is dust mapping.

1.7.1 Dust maps from emission

Dust dominates the mid- to far-infrared emission of galaxies. In this wavelength regime, dust extinction is negligible, making dust emission a reliable tracer of the dust distribution integrated along the line of sight. Dust emission maps of the Milky Way are important for the study of cosmology and the cosmic microwave background (CMB), as they enable the subtraction of foreground contamination. Schlegel et al. (1998, "SFD") utilized all-sky emission data at $100\ \mu$m and $240\ \mu$m to determine the dust temperature. After removing zodiacal light and point sources, Schlegel et al. (1998) converted the dust emission to dust column density, as shown in Figure 1.10. Planck Collaboration et al. (2016a) determined dust temperature and pro-

[6] I assume a spherical dust grain with $a = 100$ nm and that its density is $3\ \mathrm{g/cm^3}$

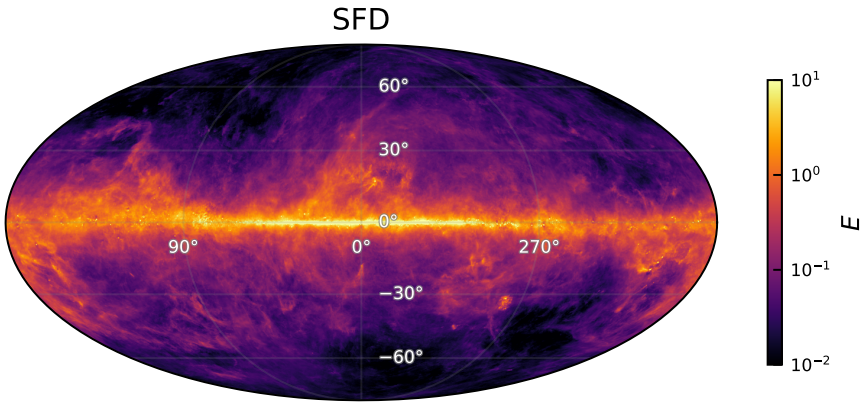


Figure 1.10. 2D mapping of dust extinction by Schlegel et al. (1998).

duced dust maps using emission data at 353, 545, and 857 GHz (corresponding to 850, 550 and 350 μm).

Emission data is the tracer of the dust components, as different components dominate the emission at different wavelengths. Both Schlegel et al. (1998) and Planck Collaboration et al. (2016a) utilized far-infrared emission, which is dominated by cold large grains. Utilizing 12-μm emission data (Meisner & Finkbeiner, 2014) dominated by PAHs, Planck Collaboration et al. (2016b) determined PAH fraction across the sky along with Planck and IRAS data, as shown in Figure 1.11. However, many artifacts can be seen on the picture, especially towards the Galactic plane. I discuss the limitation in detail in Chapter 4.

1.7.2 Dust maps from extinction

Dust extinction is strong at UV and optical wavelengths. A key advantage of mapping dust distribution through extinction is the ability to resolve distances by using stars as tracers, especially with the precise measurements of stellar parallax from *Gaia*.

Stellar extinction data provide only the integrated extinction along the line of sight to each star; inferring the full 3D distribution of dust requires probabilistic mod-

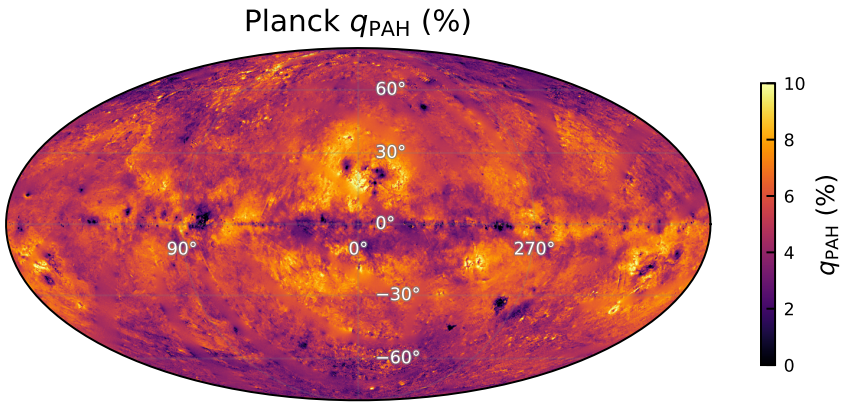


Figure 1.11. 2D PAH mass fraction map by Planck Collaboration et al. (2016b) using Planck, IRAS and WISE emission data. Artifacts from the three datasets can be seen. This is due to the limitation of 2D data and model fits. I discuss the limitation in detail in Chapter 4.

eling. Green et al. (2019) developed a 3D map using *Gaia* parallax, Pan-STARRS 1 and 2MASS. Green et al. (2019) first determine the posterior distribution of stellar type, distance and extinction of individual stars, and then they infer the distribution of dust along sightlines using the probability distribution of stellar distance and extinction. Green et al. (2019) applied a prior distribution in space to account for the continuity of dust clouds. This map is shown in Figure 1.12.

3D dust extinction maps are limited by the depth of photometric and spectroscopic surveys, as distant stars are often too faint to detect. However, with the advent of deeper surveys—particularly those targeting the Galactic plane—it will become possible to construct dust maps that cover a much broader range of distances.

1.8 On this thesis

In this thesis, I will measure the extinction curves of 220 million stars in the Milky Way, LMC and SMC. Inverting these data, I will present the distribution of the variation of extinction curves (parameterized by $R(V)$), and investigate the underlying physics. Chapter 2 describes the detailed method. Chapter 3 presents the phenomena

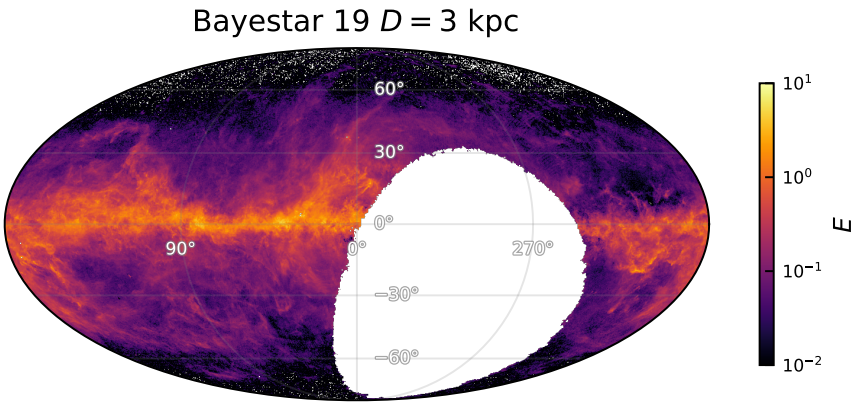


Figure 1.12. Bayestar19 dust map at distance of 3 kpc. The white area is not covered by Pan-STARRS 1 survey.

I discovered from our results, and the physical explanation. Contents in Chapter 2 and Chapter 3 are published in Zhang & Green (2025); Zhang et al. (2023b). Chapter 4 describes our theory that the anomaly I found can be explained by the growth of PAHs.

Chapter 2

Determining extinction with data-driven forward model

Disclaimer: This chapter has been published by the Monthly Notices of the Royal Astronomical Society as Zhang et al. (2023b).

2.1 Abstract

We develop, validate and apply a forward model to estimate stellar atmospheric parameters ($T_{\rm eff}$, $\log g$ and $[{\rm Fe/H}]$), revised distances and extinctions for 220 million stars with XP spectra from *Gaia* DR3. Instead of using *ab initio* stellar models, we develop a data-driven model of *Gaia* XP spectra as a function of the stellar parameters, with a few straightforward built-in physical assumptions. We train our model on stellar atmospheric parameters from the LAMOST survey, which provides broad coverage of different spectral types. We model the *Gaia* XP spectra with all of their covariances, augmented by 2MASS and WISE photometry that greatly reduces degeneracies between stellar parameters, yielding more precise determinations of temperature and dust reddening. Taken together, our approach overcomes a number of important limitations that the astrophysical parameters released in *Gaia* DR3 faced,

and exploits the full information content of the data. We provide the resulting catalog of stellar atmospheric parameters, revised parallaxes and extinction estimates, with all their uncertainties. The modeling procedure also produces an estimate of the optical extinction curve at the spectral resolution of the XP spectra ($R \sim 20 - 100$), which agrees reasonably well with the $R(V) = 3.1$ CCM model. Remaining limitations that will be addressed in future work are that the model assumes a universal extinction law, ignores binary stars and does not cover all parts of the Hertzsprung-Russell Diagram (*e.g.*, white dwarfs).

2.2 Data

We infer stellar parameters using three sources of spectro-photometric data (*Gaia* XP, 2MASS and unWISE), as well as *Gaia* parallaxes. The inclusion of near-infrared photometry from 2MASS and WISE helps us disentangle stellar temperature and extinction, by anchoring our model at longer wavelengths, where extinction has a far smaller effect. In order to train our stellar model, we additionally make use of stellar atmospheric parameter estimates based on LAMOST spectroscopy. We describe each source of data in greater detail below.

2.2.1 *Gaia* XP spectra

Gaia is a satellite-based observatory launched in 2013 by the European Space Agency (ESA), aimed at providing 6-D (sky position, parallax, proper motion and radial velocity) astrometry measurements for objects in the Milky Way and Local Group (Gaia Collaboration et al., 2016a). *Gaia* Data Release 3 (GDR3) provides positions and parallaxes of 1.46×10^9 sources (Gaia Collaboration et al., 2022). On the *Gaia* satellite, there are two spectrophotometers, with the "Blue Photometer" (BP) covering 330-680 nm and the "Blue Photometer" (RP) covering 640-1050 nm (De Angeli

et al., 2022). As of GDR3, BP/RP spectra (hereafter "XP spectra") of ~ 220 million stars have been published (Carrasco et al., 2021; Gaia Collaboration et al., 2022).

Gaia XP spectra are not natively reported as pixelized wavelength-space measurements. Instead, BP and RP spectra are each projected onto 55 orthonormal Hermite functions, and reported as a set of coefficients. Therefore, the XP spectrum of each star is represented by a 110-dimensional coefficient vector. The reason for such a representation is that the observation of each star is the combination of a series of epochs. In different epochs, the instrumental influence varies with time, the focal plane position, the detector used, the field of view, and other factors. These differences are the integral transforms of the PSF, and therefore it is mathematically easier to work with a series of continuous orthogonal basis functions (Carrasco et al., 2021; Montegriffo et al., 2022). However, for our application, it is preferable to transfer these coefficients back to wavelength space for two major reasons:

1. The dust extinction effect is linear in magnitude space, and it would be complicated to parameterize the extinction in the space of Hermite-function coefficients.

2. The spectra are noisy at the edges of BP and RP, including in their overlap region. In Hermite space, the noisy part of the spectrum is encoded by a wide range of coefficients. It is therefore impossible to remove noisy data in the edges of the BP and RP wavelength ranges by excluding certain coefficients.

We use `GaiaXPy` package[1] to convert the XP coefficients to sampled (*i.e.*, wavelength-space) spectra. Cutting out the noisy edges of the XP spectral range results in the loss of degrees of freedom. We sample the spectra from 392-992 nm, in increments of 10 nm. We find that this is the finest sampling that reliably ensures that the covariance matrices of the sampled spectra remain positive-definite. Because the flux

[1]GaiaXPy is described at `https://gaia-dpci.github.io/GaiaXPy-website/`, and version 1.1.4 can be found at `https://doi.org/10.5281/zenodo.6674521`

in our representation is sampled at discrete wavelengths, we refer to it as the "**flux vector**."

In order to accurately model XP spectra, it is also critical to include information about the covariances between the observed fluxes at different wavelengths, which are in general non-negligible. We use `GaiaXPy` to transfer the covariance matrices of the 110 basis-function coefficients to the sampled space.[2] We denote the resulting sample-space covariance matrix as $C_{\vec{f}_{\rm gaia}}$. To prevent individual wavelengths with very small reported uncertainties from dominating our results, we inflate the covariance matrix by introducing a lower limit on the individual flux uncertainties. We additionally add in zero-point uncertainties to the covariance matrices. The final covariance matrix that we use, $C_{\vec{f}_{\rm obs}}$, is given by

$$C_{\vec{f}_{\rm obs}} = C_{\vec{f}_{\rm gaia}} + \text{diag}\left(0.005\vec{f}_{\rm gaia}\right)^2 + 0.005^2\vec{f}_{\rm gaia}\vec{f}_{\rm gaia}^{T} + 0.001^2\vec{f}_{\rm BP}\vec{f}_{\rm BP}^{T} + 0.001^2\vec{f}_{\rm RP}\vec{f}_{\rm RP}^{T}\,, \tag{2.1}$$

where $\vec{f}_{\rm gaia}$ is the sampled flux, and $\vec{f}_{\rm BP}$ and $\vec{f}_{\rm RP}$ are zero-padded vectors containing the sampled BP and RP spectra, respectively. The first term is the sample-space covariance matrix, calculated directly from *Gaia*'s reported coefficient-space covariance matrix. The second term allows the flux at each wavelength to vary independently by 0.5%. The third term allows the zero-point of the overall luminosity of the entire spectrum to vary by 0.5%. The fourth and fifth terms allow the zero-points of the BP and RP sides of the spectrum to independently vary by 0.1%, in order to capture possible errors in the relative zero-point calibration of BP vs. RP.

The inflation of the flux covariance matrix ensures reasonable error bars for each element in the flux vector, but it cannot guarantee that the covariance matrix as a

[2]Note that we do not scale the BP and RP covariance matrices by `bp_standard_deviation` and `rp_standard_deviation`, respectively, due to an oversight during the preparation of the data. These factors typically change the uncertainties by only a few percent, though the corrections can be larger for a small subset of stars.

whole is well behaved (for example, that the condition number is small). We therefore diagonalize each covariance matrix using its eigendecomposition:[3]

$$C_{\vec{f}_{\text{obs}}} = UDU^T \,, \tag{2.2}$$

where U is an orthonormal matrix and D is a diagonal matrix. There are sometimes extremely small (or even more rarely, negative) values in D, which means that the predicted flux vectors are extremely strongly constrained along the corresponding eigenvectors. These "constraints" are far stronger than those one would expect from the typical scale of the flux uncertainties, and can lead to practical difficulties during training. We limit the strength of these problematic constraints by setting the minimum allowable value along the diagonal of the D matrix to 10^{-9}. We denote the modified D matrix as $\hat{D}$, and let $L \equiv \hat{D}^{-1/2}U^T$, where $\hat{D}^{-1/2}$ is well defined because $\hat{D}$ is diagonal. We can then calculate the χ^2 between the predicted flux and observed flux as

$$\chi^2 = \Delta\vec{f}^T C^{-1}_{\vec{f}_{\text{obs}}} \Delta\vec{f} = \left| L\,\Delta\vec{f} \right|^2 \,, \tag{2.3}$$

where $\Delta\vec{f} \equiv \vec{f}_{\text{pred}} - \vec{f}_{\text{obs}}$ is the residual between the predicted and observed flux.

2.2.2 LAMOST

The Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) is a ground-based telescope which has observed over ~ 10 optical stellar spectra in the Northern Hemisphere with a resolution $R \sim 1800$ and with the limiting magnitude of $r < 19$ (Cui et al., 2012; Zhao et al., 2012). In this work, we use the AFGK catalog from LAMOST Data Release 8 (DR8),[4] which contains atmospheric pa-

[3] We use `numpy.linalg.eigh` for the eigendecomposition of the covariance matrices, taking advantage of the fact that they are Hermitian.

[4] LAMOST DR8: `http://www.lamost.org/dr8/v2.0/`

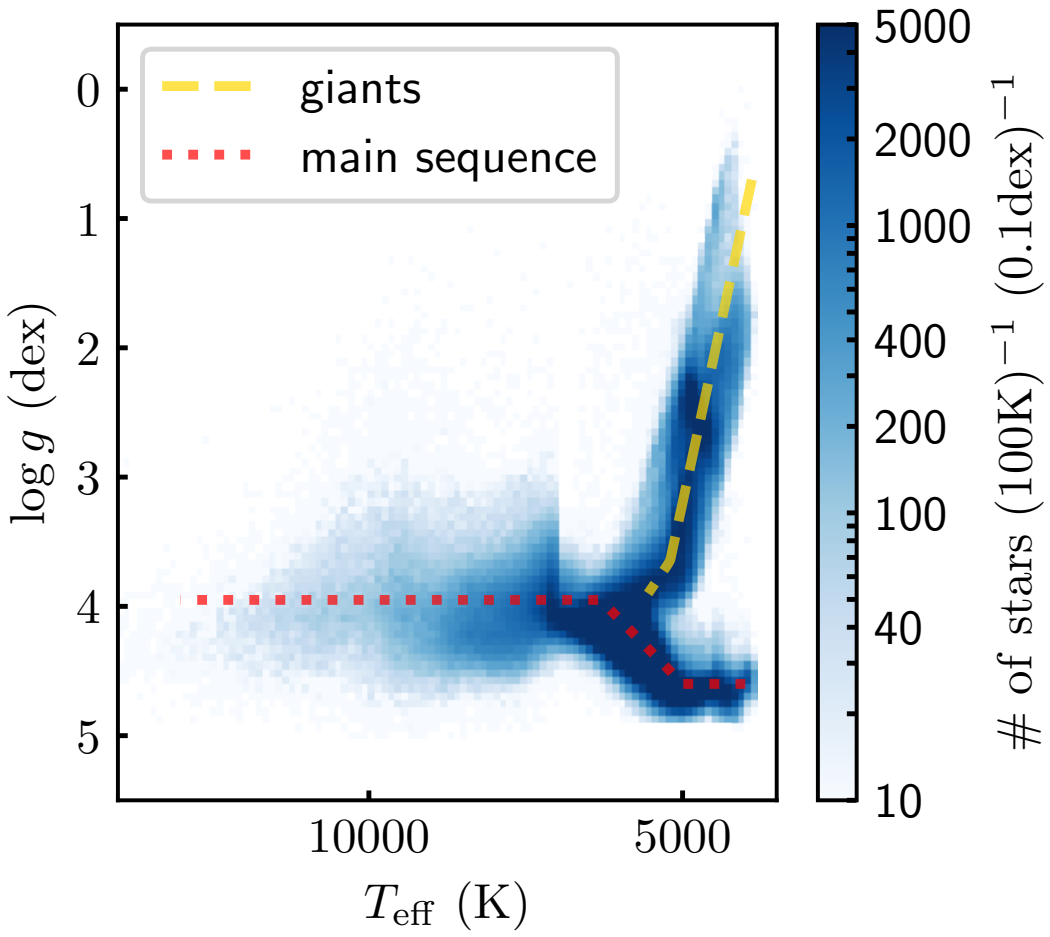


Figure 2.1. Kiel diagram ($\log g$ *vs.* $T_{\rm eff}$) of our matched GDR3 XP and LAMOST catalog, used for training and validating our stellar model. The matched catalog contains 2,575,354 stars, covering much – but not all - of stellar parameter space. The sharp feature at $T_{\rm eff} = 7000\,\mathrm{K}$ is the point at which we transition from the standard DR8 stellar atmospheric parameter estimates to the "Hot Payne" (Xiang et al., 2022), which provides more accurate temperatures for hot stars. We overplot piecewise-linear main-sequence and giant-branch tracks (See Eqs. 2.27 and 2.28), which we later use when visualizing our trained stellar models.

rameters of over 6 million stars. The typical errors of the stellar parameters in this catalog are ~ 40 K for $T_{\rm eff}$, 0.06 dex for $\log g$, and 0.04 dex for $[{\rm Fe/H}]$. For hot stars ($T_{\rm eff} \gtrsim 7500$ K), we obtain stellar atmospheric parameters from the "Hot Payne" (Xiang et al., 2022), which applies the method, "The Payne" (Ting et al., 2019) to determine the atmospheric parameters of $\sim 330,000$ O-, B- and A-type stars in LAMOST Data Release 6. The standard LAMOST stellar parameter pipeline has difficulty modeling the Balmer lines of stars with $8,000\,{\rm K} \lesssim T_{\rm eff} \lesssim 12,000\,{\rm K}$, leading to inaccurate temperature estimates in this range. The Hot Payne produces more accurate temperature estimates for sta. Thrers in this regime. However, as can be seen in Fig. 2.1, the distribution of $\log g$ in our combined catalog changes sharply at 7000 K, where we transition from the standard LAMOST catalog to the Hot Payne catalog. We leave the problem of spanning these two temperature regimes more seamlessly to future work.

We crossmatch LAMOST with GDR3 by searching the closest star, out to a maximum angular separation of 0.25". Fig. 2.1 shows a Kiel diagram of the resulting cross-matched catalog. The matched stars cover a wide area of stellar atmospheric parameter space, including the main sequence from $\sim 4,000 - 12,000\,{\rm K}$, and the giant branch up to $\log g \sim 0.5$. Certain stellar types and stellar remnants, such as white dwarfs and subdwarfs, are not covered by our matched catalog. In addition, some regions of parameter space, such as the horizontal branch and stars with $T_{\rm eff} \gtrsim 15,000\,{\rm K}$, have only sparse coverage. Although the overall number of matched sources (~ 2 million) comprises only $\sim 1\%$ of the *Gaia* XP catalog, it is still sufficient to train a model of XP spectra over a wide range of stellar parameter space.

2.2.3 2MASS

2MASS is a near-infrared, all-sky imaging survey, which obtained photometry for $\sim 4 \times 10^8$ stars in three bands, namely J, H and K_s (Skrutskie et al., 2006). The effective wavelengths of the bands are J $(1.25\mu\text{m}), H$ $(1.65\mu\text{m})$, and $K_s(2.16\mu\text{m})$, respectively. The 10 σ limiting sensitivities for these bands are $J < 15.8$, $H < 15.1$ and $K_s < 15.8$. The typical uncertainty for bright sources ($K_s < 13$) is 0.03 mag, and the calibration offsets are < 0.02 mag.

We use the crossmatched 2MASS data from the *Gaia* Archive (Gaia Collaboration, 2022a,b,c). See section 2.9 for our full ADQL query.

2MASS used three 256×256 pixel arrays, with a pixel scale of $2'' \times 2''$, which is much larger than that of *Gaia* CCD pixels ($\sim 177\,\text{mas} \times 59\,\text{mas}$, Gaia Collaboration et al. 2016b). Therefore, a star clearly resolved by *Gaia* could be severely contaminated by its neighbors in 2MASS, as shown in Fig. 2.2. The two stars at $(\alpha_{J2000}, \delta_{J2000})$= (09:17:13.8, 53:25:09.6) are resolved by GDR3, and fulfill the standard of GDR3 to be spectroscopically analysed, but they cannot be resolved by 2MASS because their angular separation is $\sim 3.2''$. In such cases, we expect a positive bias in the observed J, H and K_s fluxes, compared with the predicted fluxes given by our model. To solve this, we disregard 2MASS observations (by setting their variance to infinity) if `norm_dg` > -5. The parameter `norm_dg` is defined as $max\{\frac{\Delta G_i}{\text{mag}} - \frac{\theta_i}{\text{arcsec}}\}$, where ΔG_i is the difference of G-band magnitude between the target star and its i-th neighbor with an angular separation $< 30''$, and θ_i is the angular separation between them (Rybizki et al., 2022). Therefore, `norm_dg` is a measure of crowding, with larger values indicating the presence of closer, brighter neighbors.

2MASS photometry is natively provided in Vega magnitudes. In order to put 2MASS photometry on a scale that is easily comparable to *Gaia* XP spectral fluxes, we convert it to spectral flux using an assumed central wavelength and AB-Vega

magnitude offset:

$$f_\lambda = (3631\,\mathrm{Jy})\, c\lambda_0^{-2} 10^{-0.4(m+\Delta m)}\,. \tag{2.4}$$

Where m is the reported Vega magnitude; c is the speed of light; the AB-Vega offset $\Delta m = 0.91$, 1.39 and 1.85 mag for J, H and K_s, respectively (Blanton et al., 2005); and the central wavelength $\lambda_0 = 1.235$, 1.662 and 2.159 μm for J, H and K_s, respectively. This conversion is only approximate, as a correct conversion between broad-band photometry and spectral flux requires a knowledge of the shape of the source spectrum. Nevertheless, this conversion is mathematically well defined, can be inverted to recover the broad-band photometry, and allows a rough comparison with XP spectral fluxes.

2.2.4 unWISE

The WISE mission has observed the entire sky in four bands, $W1$ $(3.4\mu m)$, $W2$ $(4.6\mu m)$, $W3$ $(12\mu m)$ and $W4$ $(22\mu m)$, with a space-based 40 cm telescope ("WISE cryogenic", Wright et al. 2010). Because the HgCdTe detector arrays used by $W1$ and $W2$ remain functional without solid hydrogen as coolant, more data in these two bands was accumulated in search of near-Earth objects before the hibernation in 2011 ("NEOWISE", Mainzer et al. 2011). Yet more $W1$ and $W2$ observations have been collected since the reactivation of NEOWISE in 2013 ("NEOWISER", Mainzer et al. 2014). Schlafly et al. (2019) co-added "WISE cryogenic", "NEOWISE" and "NEOWISER" to build the "unWISE" catalog, which is deeper than each individual catalog it used. We make use of the unWISE "neo6" catalog, which is based on WISE exposures captured through 13 December 2019 (Meisner et al., 2021). The unWISE catalog behaves better in the Galactic disk, as it uses the "crowdsource" pipeline, which was developed to handle the extremely crowded fields observed by

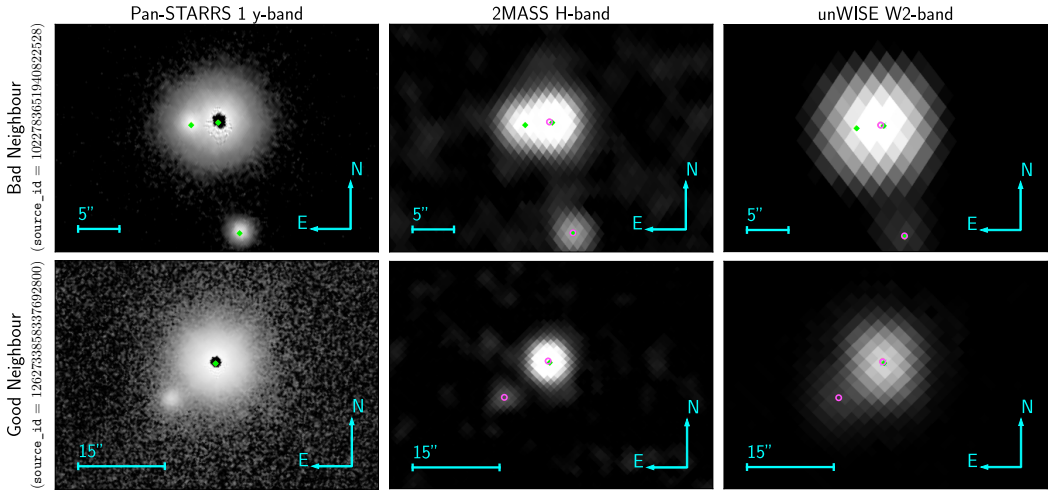


Figure 2.2. Example of a source (centered in each image) with a problematic (top panels) and an unproblematic neighbor (bottom panels). The left panels show PS1 y-band images, the center panels show 2MASS H-band images, and the right panels show unWISE $W2$-band images. The green diamonds (♦) mark the locations of sources with XP spectra in GDR3. In the middle and right panels, the violet circles (○) mark the locations of point sources included in the 2MASS and unWISE catalogs, respectively. 2MASS and WISE do not detect the problematic neighbor, due to their lower resolution (approximately 2" and 6", respectively) compared to *Gaia*. The source with the problematic neighbor has a relatively high `norm_dg` value of -5.98, indicating that it has a close, bright neighbor. This source fails our cut for including unWISE measurements, but barely passes our cut for using 2MASS measurements. The bottom panels show a source with an unproblematic neighbor, which is detected in both 2MASS and unWISE. As this source has a `norm_dg` value of -19.4, indicating that it is relatively isolated, we use both its 2MASS and unWISE measurements.

the Dark Energy Camera Plane Survey (Schlafly et al., 2018).

The FWHM of the point spread function (PSF) of WISE is $6''.1$ for $W1$ and $6''.4$ for $W2$, which is several times larger than GDR3. For this reason, we disregard unWISE photometry if `norm_dg` > -10, indicating the presence of nearby, bright neighbors.

As with 2MASS, we convert unWISE magnitudes from their native Vega system to spectra flux using Eq. (2.4). We assume central wavelengths of $\lambda_0 = 3.3526$ and 4.6028 μm, and and AB offsets of m_0 2.699 and 3.339 mag for $W1$ and $W2$, respectively.[5]

2.2.5 Training and validation datasets

We train our model on our cross-matched *Gaia* XP–LAMOST catalog, with *Gaia* parallaxes and photometry from 2MASS and unWISE. We impose a number of additional quality cuts on our cross-matched catalog:

- LAMOST SNR of greater than 20 in g-, r-, and i-bands.
- LAMOST uncertainties of less than 500 K in $T_{\rm eff}$, and less than 0.5 dex in [Fe/H] and $\log g$.
- Well constrained parallaxes: `parallax_over_error` > 3.
- Reliable *Gaia* astrometry: `fidelity_v2` > 0.5.
- Low BP/RP flux excess: `bp_rp_flux_excess` < 1.3.

After these quality cuts, our cross-matched catalog contains a total of 2,575,354 sources. We split our dataset into a training set (80% of sources) and a validation set (20% of sources). The training dataset is used to train a model of stellar flux and dust extinction, as a function of wavelength, while the validation set is used to evaluate the performance of the trained model.

[5] WISE Vega-AB magnitude offsets are given by `https://wise2.ipac.caltech.edu/docs/release/allsky/expsup/sec4_4h.html#conv2ab`.

2.3 Method

We build a forward model that maps from stellar type, extinction and parallax to the expected XP spectrum and near-infrared 2MASS and WISE photometry. We build our model in the auto-differentiable TensorFlow 2 framework (Abadi et al., 2015), which allows us to optimize both the model's internal parameters and to fit individual stellar parameters using gradient descent methods. We train our forward model using atmospheric parameters from LAMOST, parallaxes from *Gaia* DR3 (Gaia Collaboration et al., 2022), and reddenings from the 3D dust map Bayestar19 (Green et al., 2019). We then apply our model to all stars with *Gaia* XP spectra, to determine their types, distances and extinctions.

We map stellar atmospheric parameters, $\Theta = (T_{\rm eff}, \log g, [{\rm Fe/H}])$, to the unreddened spectral profile at 1 kpc, which we call the "absolute flux" and denote by $\vec{f}_{\rm abs}(\Theta)$:

$$\Theta \rightarrow \vec{f}_{\rm abs}(\Theta)\,. \tag{2.5}$$

We implement this model as a simple feed-forward neural network, which takes Θ as an input and outputs a vector, with each entry representing the natural log of the absolute flux at a different wavelength. To obtain the unreddened flux at the distance of the star, we multiply by ϖ^2, the square of the parallax (note that this is a model parameter, and not the observed noisy parallax). We assume that the effect of extinction is to multiply the absolute flux by

$$\exp(-E\vec{R})\,, \tag{2.6}$$

where E is a scalar measurement of the amount of extinction in front of the star along the line of sight. The vector $\vec{R}$ is shared by all stars, and represents the relative amount of extinction at each wavelength. We ignore extinction effects that would

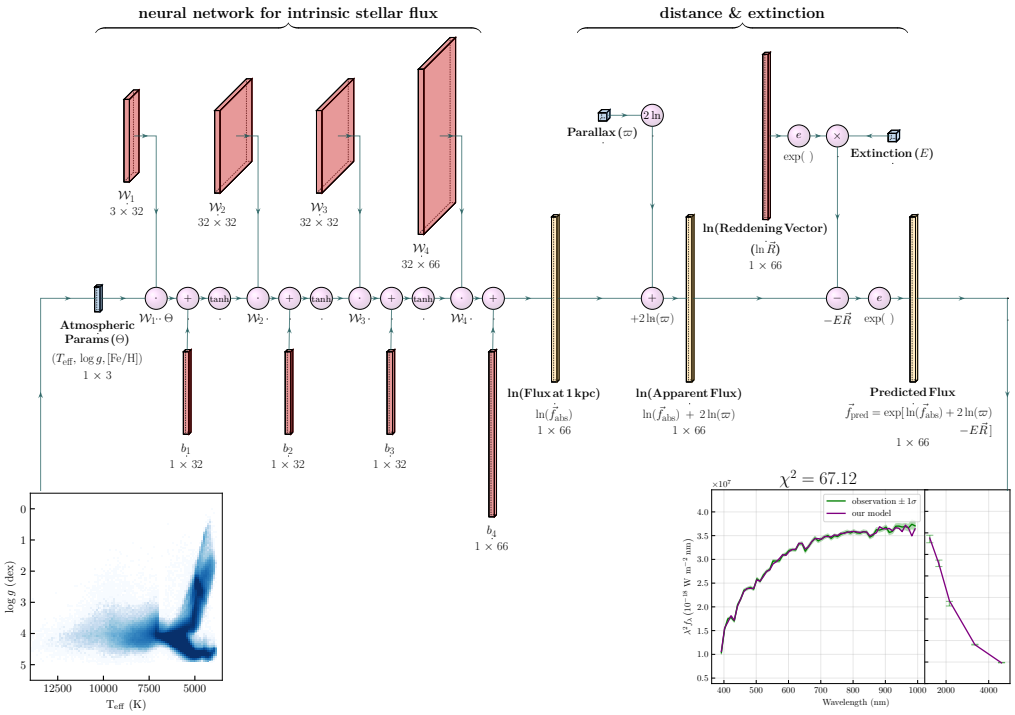


Figure 2.3. The structure of our stellar flux model. Individual stellar parameters are represented in the light blue blocks, $\Theta \equiv (T_{\rm eff}, \log g, [\mathrm{Fe/H}])$, the parallax ϖ and the extinctions E. Global model parameters are represented as red blocks, and intermediate calculations and outputs are represented in yellow. The overall structure of the model is chosen to only use neural networks in the aspects where simple *ab initio* physical models fall short. The stellar atmospheric parameters (Θ) are fed into a neural network consisting of three 16-neuron dense layers, with weights $\mathcal{W}$ and biases b. The neural network outputs the natural logarithm of the unreddened absolute flux observed at 1 kpc, which is a 66-dimensional vector (61 wavelength samples from XP with an interval of 10 nm, plus J, H, $\mathrm{K_s}$ from 2MASS and $W1$ and $W2$ bands from unWISE). The inverse-square law decay in observed flux is accounted for with the factor ϖ^2, and the extinction is represented by the factor of $\exp(-E\vec{R})$, where $\vec{R}$ is the relative amount of extinction at each wavelength and E is a scalar that represents the overall amount of extinction for the given star. The final output is the predicted flux, $\vec{f}_{\rm pred} = \vec{f}_{\rm abs}\varpi^2 \exp(-E\vec{R})$. This prediction is compared with the observation, $\vec{f}_{\rm obs}$, to calculate χ^2. The loss function consists of the χ^2, the likelihoods of (Θ, ϖ, E) from outside observations (LAMOST, *Gaia* and Bayestar19, respectively), and the L2 regularization of the neural network weights ($\mathcal{W}$). In the training process, we first fix (Θ, ϖ, E), the blue blocks in the picture, and find the optimal global parameters, ($\mathcal{W}$, b, $\vec{R}$), that minimize the loss function. Then we fix ($\mathcal{W}$, b, $\vec{R}$), the red blocks, and optimize the stellar parameters (Θ, ϖ, E). We repeatedly switch between updating the red blocks and optimizing blue blocks, until the loss function stops decreasing.

be caused by the finite width of our bandpasses, because XP spectral elements are much narrower than typical photometric bands. For 2MASS and WISE photometry, this will introduce somewhat larger fractional errors in extinction, but this effect is mitigated by the fact that extinction is smaller in the near-infrared.

Putting all the piece of our model together, the predicted flux is given by

$$\vec{f}_{\text{pred}} = \vec{f}_{\text{abs}}(\Theta, \mathcal{W}, b)\varpi^2 \exp\left(-E\vec{R}\right), \tag{2.7}$$

where $\mathcal{W}$ and b represent all trainable neural-network weights and biases, respectively, in the absolute flux model. The structure of our model is shown in Fig. 2.3. In total, our model has 4484 global parameters (contained in $\mathcal{W}$ and b, the weights and biases of the neural network, as well as the extinction curve $\vec{R}$), and an additional five parameters per star (T_{eff}, [Fe/H], $\log g$, ϖ, and E).

In the following, we provide a detailed description of the workings of our model, and of our training procedure. We invite readers who are primarily interested in the results we obtain using the model to inspect Fig. 2.3, which gives an overview of the model structure, and then to proceed to Section 2.5.

2.3.1 Bayesian formulation of model

Before describing how we learn the parameters in the model, we will write down the posterior density of the model and individual stellar parameters, given our observations. Then, we will show how we use gradient descent to infer both the global model parameters and the individual stellar parameters.

The likelihood of a single star's observed flux $\vec{f}_{\text{obs}}$ (with uncertainties described by the covariance matrix C_f), given our global model parameters $\mathcal{W}$, b and $\vec{R}$, as well as our modeled type Θ, true stellar parallax ϖ, and reddening E is given (up to

a constant) by

$$\ln p\left(\vec{f}_{\text{obs}} \mid \mathcal{W}, b, \vec{R}, \Theta, \varpi, E\right) = -\frac{1}{2}\underbrace{\Delta\vec{f}^{T} C_f^{-1} \Delta\vec{f}}_{\equiv\chi^2}, \tag{2.8}$$

where $\Delta\vec{f} \equiv \vec{f}_{\text{obs}} - \vec{f}_{\text{pred}}\left(\Theta, \varpi, E \mid \mathcal{W}, b, \vec{R}\right)$. In practice, we calculate χ^2 using our matrix decomposition of C_f^{-1}, as shown in Equation (2.3).

For each star, we also have an observed *Gaia* parallax $\hat{\varpi}$ (with Gaussian uncertainty σ_{ϖ}), a LAMOST estimate of stellar type $\hat{\Theta}$ (with diagonal covariance matrix described by $\vec{\sigma}_{\Theta}$), and a Gaussian estimate of reddening which we derive from Bayestar19 (with uncertainty σ_E). We treat these as observed quantities, so that each contributes an independent Gaussian likelihood term. We additionally place priors on the individual stellar parameters. We place improper flat priors on ϖ, E, and Θ, with the additional condition that ϖ and E must be positive (we achieve this by using $\ln\varpi$ and $\ln E$ as our variables, and including the appropriate Jacobian terms to transform our flat priors in ϖ and E).

We also impose simple priors on the global model parameters, as follows. We place a flat prior on the logarithm of each component of $\vec{R}$, a Gaussian prior on the neural network weights $\mathcal{W}$ (which is equivalent to adding an L2 regularization penalty to the network), and a flat prior on the neural network biases b. As will be seen later, we set the standard deviation of the prior on $\mathcal{W}$ so that it makes a similar contribution to the posterior as one extra degree of freedom per star would be expected to make.

Phase	Only HQ data	Initial LR of ($\mathcal{W}$, b, $\vec{R}$)	Initial LR of (Θ, ϖ, E)	Optimizer
1	✓	10^{-3}	–	SGD
2	✓	10^{-5}	10^{-4}	SGD
3		–	10^{-4}	SGD
4	Self-cleaning			
5		10^{-7}	10^{-5}	SGD
6		–	10^{-2}	Adam

Table 2.1. Training procedure described in Section 2.3.2. "LR" stands for the learning rate. "HQ" stands for high-quality sources, as defined in Equation (2.14). When a given set of parameters are not updated in a given phase, the initial learning rate is marked "–". In phase 1, we first train the model parameters ($\mathcal{W}$, b, $\vec{R}$) on HQ data while holding the stellar parameters (Θ, ϖ, E) fixed. This produces an initial, rough model. We then train both the model and individual stellar parameters in phase 2, again using only HQ data. In phase 3, we apply the model trained with HQ data to all stars in the training set to refine their stellar parameters. In phase 4, we use a self-cleaning procedure to remove unresolved binary systems and other outliers. In phase 5, we use the remaining stars to train the model, while simultaneously updating the stellar parameters. At this point, the training of our model is complete. Finally, in phase 6, we use our trained model to update the parameters of all stars in the training set.

In full, the posterior of our model, assuming one observed star, can be written as

$$\begin{aligned} & p\left(\mathcal{W}, b, \vec{R}, \Theta, \varpi, E \mid f_{\text{obs}}, \hat{\Theta}, \sigma_{\Theta}, \hat{\varpi}, \sigma_{\varpi}, \hat{E}, \sigma_{E}, C_{f}\right) \\ & \quad \propto p\left(\vec{f}_{\text{obs}} \mid \mathcal{W}, b, \vec{R}, \Theta, \varpi, E, C_{f}\right) \\ & \quad \times p\left(\hat{\varpi} \mid \varpi, \sigma_{\varpi}\right) p\left(\hat{E} \mid E, \sigma_{E}\right) p\left(\hat{\Theta} \mid \Theta, \sigma_{\Theta}\right) \\ & \quad \times p(\varpi)\, p(E)\, p(\Theta)\, p(\mathcal{W})\, p(b)\, p\left(\vec{R}\right), \end{aligned} \tag{2.9}$$

with a normalizing constant that is independent of the model parameters. The generalization to multiple stars is trivial: every term except for the priors on $\mathcal{W}$, b and $\vec{R}$ is repeated for each star.

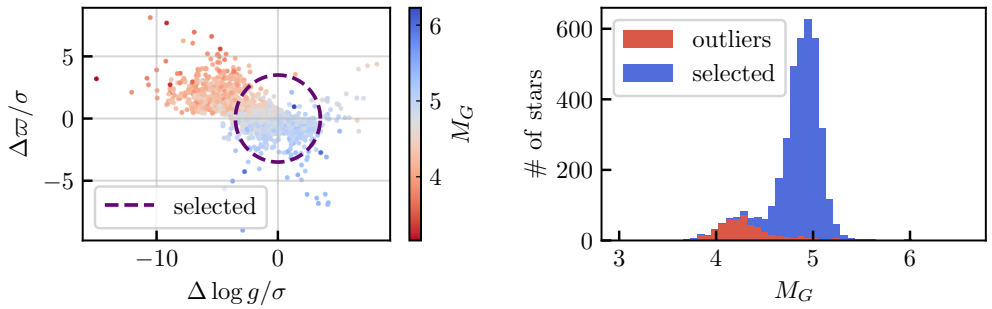


Figure 2.4. "Self-cleaning" process of our approach, which generates outlier rejection flags. In essence, the approach identifies instances where the model's fundamental assumption – stellar parameters uniquely predict absolute spectral fluxes – create significant tension with the data (e.g. the parallax). For the illustrative regime of low-extinction Solar-type stars (defined by Eq. 2.15), the left panel shows the deviation of our optimized ϖ and $\log g$ from the observed values. Each star is colored by the inferred absolute magnitude, $M_G = m_G - \mu$, where m_G is the apparent G-band magnitude, and the distance modulus is given by $\mu = 10 - 5\log(\varpi/1\text{ mas})$. We define $\Delta\varpi/\sigma \equiv (\varpi_{\text{opt}} - \varpi_{\text{obs}})/\sigma(\varpi_{\text{obs}})$ and $\Delta \log g/\sigma \equiv (\log g_{\text{opt}} - \log g_{\text{obs}})/\sigma(\log g_{\text{obs}})$. We remove stars with large residuals, $(\Delta \log g/\sigma)^2 + (\Delta\varpi/\sigma)^2 > 3.5^2$. The right panel shows a histogram of the absolute G-band magnitude of same population of stars, with the outliers identified in the left panel colored in red. The outliers form a distinct population with absolute magnitudes that are ~ 0.75 mag brighter (corresponding to twice the luminosity) than the typical star with the same LAMOST atmospheric parameters. The outliers are thus likely dominated by near-equal-mass binaries.

2.3.2 Training the model

Our goal is to find the set of both global model parameters and individual stellar parameters that maximize the above posterior density. In an ideal world, we would simultaneously infer all parameters – both the global model parameters and the individual stellar parameters. However, we find an iterative approach to be more manageable. We implement our model in Tensorflow 2, and use gradient descent to iteratively update the global model parameters (holding the individual stellar parameters fixed) and the individual stellar parameters (holding the global model parameters fixed). Not all terms in the posterior need to be calculated for each type of gradient descent step – we only need to calculate the terms that depend on the set of parameters being updated. Thus, when updating the global model parameters, we do not calculate the stellar priors or the likelihoods of $\hat{\varpi}$, $\hat{E}$ and $\hat{\Theta}$.

To train the global model parameters $\mathcal{W}$, b and $\vec{R}$, holding the stellar parameters Θ, ϖ and E fixed, we seek to maximize

$$p\left(\vec{f}_{\text{obs}} \mid \mathcal{W}, b, \vec{R}, \Theta, \varpi, E, C_f\right) p\left(\mathcal{W}\right) p\left(b\right) p\left(\vec{R}\right) . \tag{2.10}$$

This is equivalent to minimizing χ^2_f plus an L2-norm penalty on $\mathcal{W}$:

$$\mathcal{L}_{\text{model}}\left(\mathcal{W}, b, \vec{R}\right) = \chi^2_f + \frac{1}{n_W}\|\mathcal{W}\|_2 \,, \tag{2.11}$$

with respect to the global model parameters $\mathcal{W}$, b, and $\vec{R}$. Here, χ^2_f is the sum of the flux likelihood for a single random star, calculated according to Equation 2.3, and n_W is the total number of weights in $\mathcal{W}$. The L2-norm penalty is equivalent to a Gaussian prior on the neural network weights, and causes our loss function to prefer smaller weights in $\mathcal{W}$, and thus simpler models. For unit weights, this prior would be the equivalent of adding one to χ^2_f. In practice, during each gradient descent step,

we calculate the average of $\mathcal{L}_{\text{model}}$ for batches of 512 stars, in order to reduce the noise of the gradient.

To learn the parameters describing an individual star, holding the global model fixed, we seek to maximize all of the terms on the right-hand side of Equation (2.9), except for the last three prior terms, which do not depend on the individual stellar parameters. This is equivalent to minimizing the stellar loss function

$$\mathcal{L}_{\text{star}}\left(\Theta, \ln\varpi, \ln E\right) = \chi_f^2 + \left(\frac{\varpi - \hat{\varpi}}{\sigma_\varpi}\right)^2 + \left(\frac{E - \hat{E}}{\sigma_E}\right)^2 \tag{2.12}$$

$$+ \sum_{i=0}^{3}\left(\frac{\Theta_i - \hat{\Theta}_i}{\sigma_{\Theta,i}}\right)^2 - 2\ln\varpi - 2\ln E, \tag{2.13}$$

with respect to the individual stellar parameters, Θ, ϖ and E. The 1st term corresponds to the likelihood of the observed fluxes (calculated using Equation 2.3), while the 2nd, 3rd and 4th terms correspond to the likelihoods of ϖ, E and Θ. The last two terms correspond to the Jacobian terms $\frac{\partial\varpi}{\partial \ln\varpi}$ and $\frac{\partial \mathrm{E}}{\partial \ln \mathrm{E}}$, respectively.

We train our model and update the stellar parameters in stages. In the first stage, we fix Θ, ϖ and E at their mean observed values, as determined by LAMOST, *Gaia* and Bayestar19.

We use a stochastic gradient descent (SGD) optimizer with momentum of 0.5 to update the global model parameters. In order to update individual stellar parameters, we again use SGD, but with zero momentum. When alternating between different batches of stars, this prevents the gradient descent direction of one batch of stellar parameters from contaminating the gradient descent direction of the following batch of stellar parameters.

We train the model in several phases. In the first phase, we use a subset of stars with high-quality measurements to train our stellar flux and extinction model. In this phase, we update the global model parameters, while holding the stellar parameters fixed at their measured values (as determined by LAMOST, *Gaia* and Bayestar19).

We define the "high-quality subset" as the stars (in the training set) that pass all of the following cuts:

$$\begin{aligned} &(i) \quad \sigma_{T_{\rm eff}} < 200\,{\rm K} \\ &(ii) \quad \sigma_{\rm [Fe/H]} < 0.2\,{\rm dex} \\ &(iii) \quad \sigma_{\log g} < 0.2\,{\rm dex} \\ &(iv) \quad \hat{\varpi}/\sigma_{\varpi} > 10 \\ &(v) \quad \sigma_E < 0.1 \end{aligned} \tag{2.14}$$

The high-quality subset contains 1,861,666 sources (90.1% of the training set). Using this high-quality subset, we minimize $\mathcal{L}_{\rm model}$ by performing stochastic gradient descent on $\mathcal{W}$, b, and $\vec{R}$, or the "red blocks" in Fig. 2.3. We use batches of 512 stars to calculate $\mathcal{L}_{\rm model}$ at each gradient descent step. In the following, we will use the concept of a training "epoch," which is the number of training batches (or equivalently, training steps) required to run through the training dataset. For our training dataset, one epoch contains approximately 5,000 batches. In the first phase, we train for 128 epochs. We begin with a learning rate of 10^{-3}, and reduce the learning rate by a factor of 10 at epochs 32, 64 and 96.

After this initial training phase, which produces a reasonable first guess of our stellar flux and extinction model, we proceed to a second training phase, in which we simultaneously refine both the global model parameters and the individual parameters of the stars in the high-quality subset. We again train for 128 epochs. In each training step, we first update the global model parameters, and then update the parameters of the current batch of stars. We begin with a learning rate of 10^{-5} for the global model parameters, and 10^{-4} for the individual stellar parameters, and reduce the learning rates by a factor of 10 at epochs 32, 64 and 96.

We would next like to train using the entire training set, including stars with larger measurement uncertainties. However, before using these stars to update the model, we would like to refine their individual parameters (Θ, ϖ, E). We thus

take 128 gradient descent steps for each star (including sources in the high-quality subset), holding the global model parameters fixed. We begin with a learning rate of 10^{-4}, and again reduce the learning rates by a factor of 10 at epochs 32, 64 and 96.

Before proceeding to again update the global model parameters, we first carry out a self-cleaning step, in which we remove outlier stars based on large discrepancies in their measured vs. predicted fluxes or stellar parameters. We identify outliers in two ways. The first method is to remove any star for which our estimate of Θ or ϖ is more than 4σ removed from the measured value. Binary systems, in particular, are likely a major contaminant in our dataset. A near-equal-mass binary will appear approximately twice as bright as would be predicted from the stellar type and parallax of the system. Our model will attempt to accommodate such systems by decreasing $\log g$ or increasing ϖ, both of which have the effect of increasing predicted flux. However, because LAMOST measures stellar atmospheric parameters using individual line shapes, rather than the overall luminosity of the source, it should not respond as drastically to the presence of a binary companion. We thus expect this self-cleaning method to identify (and remove) a large number of potential binary systems. This self-cleaning procedure is illustrated for low-extinction Solar-type stars in Fig. 2.4. We define "low-extinction Solar-type" as:

$$\begin{aligned}
&|T_{\rm eff} - 5700| < 50 \text{ K}, \quad \sigma_{T_{\rm eff}} < 100 \text{ K}, \\
&|\log g - 4.5| < 0.1, \qquad \sigma_{\log g} < 0.1, \\
&|[\mathrm{Fe/H}] - 0.0| < 0.1, \quad \sigma_{[\mathrm{Fe/H}]} < 0.1, \\
&|\sigma_{\varpi,\mathrm{obs}}/\varpi_{\mathrm{obs}}| < 0.1, \\
&E < 0.05, \qquad\qquad\qquad \sigma_E < 0.05
\end{aligned} \tag{2.15}$$

The second method of identifying outliers is to flag stars for which the residual between the observed and predicted flux at any wavelength is discrepant by more than

4σ. Our parameter-based outlier rejection flags 21.4% of sources, while our flux-based outlier rejection flags 6.9% of sources. Because there is some overlap between these two populations, in total, we remove 24.7% of sources through self-cleaning.

Next, we train both the model and refine the individual stellar parameters. We include all sources that pass our self-cleaning cut (1,556,666 sources, or 75.3% of the training set) in this phase, and train for 128 epochs. In each training step, we first update the global model parameters, and then the parameters of the stars in the batch. We begin with a learning rate of 10^{-7} for the global model parameters, and 10^{-5} for the individual stellar parameters, and again reduce the learning rates by a factor of 10 at epochs 32, 64 and 96. This is the final phase in which we update the global model parameters.

Finally, we refine the individual stellar parameters, holding the global model parameters fixed. In this phase, use the Adam optimizer (Kingma & Ba, 2014), with an initial learning rate of 0.01, which we halve every 512 steps. We take a total of 4096 gradient descent steps for each star.

We summarize the key properties of each training phase in Table 2.1. At the end of this process, we have a trained model of stellar flux and extinction as a function of wavelength, as well as an updated estimate of the type, parallax (or equivalently, distance) and extinction of each star.

2.4 Determination of stellar parameters

We now wish to determine the parameters of all stars with XP spectra. Our approach is to use the model trained in the previous section to fit the stellar type, extinction and parallax, based on the observed flux and parallax. Although our optimization algorithm will be similar to the algorithm used to refine the stellar parameters during training in Section 2.3, 99% of the BP/RP sources do not have observational con-

straints on stellar atmospheric parameters from LAMOST, and constraints on E from Bayestar19 are unavailable for stars with $\delta < -30°$. However, prior constraints on these parameters are still necessary, in order to prevent the optimizer entering regions of parameter space which are not physical, or which are not covered by the training data. We therefore replace the LAMOST constraints on stellar type by a Gaussian mixture model (GMM) prior. We additionally enforce positivity of extinction and distance by using $\ln E$ and $\ln \varpi$ as our parameters.

2.4.1 Formulation and priors

We determine the stellar parameters by maximizing the posterior probability of (Θ, ϖ, E), given only the observed flux ($\vec{f}_{\text{obs}}$, C_f) and observed parallax ($\hat{\varpi}$, σ_ϖ), under necessary constraints and priors. The global model parameters ($\mathcal{W}$, b, $\vec{R}$) have been fixed by the training process described in Section 2.3.2. If we omit constant factors, the posterior of a single star is given by

$$\begin{aligned} p\left(\Theta, \varpi, E \mid \vec{f}_{\text{obs}}, C_f, \hat{\varpi}, \sigma_\varpi, \mathcal{W}, b, \vec{R}\right) \\ \propto p\left(\vec{f}_{\text{obs}} \mid \Theta, \varpi, E, C_f, \mathcal{W}, b, \vec{R}\right) \\ \times p\left(\hat{\varpi} \mid \varpi, \sigma_\varpi\right) \\ \times p\left(\varpi\right) p\left(E\right) p\left(\Theta\right) \ . \end{aligned} \tag{2.16}$$

The first term, $p\left(\vec{f}_{\text{obs}} \mid \mathcal{W}, b, \vec{R}, \Theta, \varpi, E, C_f\right)$, is the likelihood of the observed flux, given by Equation 2.8. The second term is the likelihood of the observed parallax, which is a Gaussian with mean and standard deviation given by the GDR3 observation. We place a flat prior on positive parallaxes. As true parallax is strictly positive, we use $\ln \varpi$ as our variable (and include a corresponding Jacobian term in the prior). As Green et al. (2019) does not cover $\delta < -30°$, and as we want our determination of stellar extinction to be independent of Green et al. (2019), we

impose a flat prior distribution of E on $(0, +\infty)$. We use $\ln E$ as our variable, and include a corresponding Jacobian term $(-2 \ln E)$ in the prior. We use a Gaussian mixture model as the prior on atmospheric parameters ($T_{\rm eff}$, $[{\rm Fe/H}]$, $\log g$), in order not to place them in regions of parameter space that are not covered by the training set (and for which we therefore do not have reliable models). The GMM is a probabilistic model describing an overall population (i.e. stellar type distribution in the training set) by the combination sub-populations (e.g., main sequence stars or the red clump), each with a Gaussian distribution in stellar-type-space:

$$p_{\rm gmm}(\Theta) = \sum_{i=1}^{N} k_i \mathcal{N}\left(\Theta \mid \boldsymbol{\Theta}_i, \boldsymbol{C}_i\right) , \tag{2.17}$$

where N is the number of sub-populations, and $\boldsymbol{\Theta}_i$, $\boldsymbol{\Sigma}_i$ and k_i are the mean, covariance matrix and weight, respectively, of the $i^{\rm th}$ sub-population. The weights fulfill $\sum_{i=1}^{N} k_i = 1$. We use the `scikit-learn` package (Pedregosa et al., 2011) to fit our GMM on the training set, and then implement the resulting model in TensorFlow 2, so that it is auto-differentiable. We use $N = 16$ components to represent the stellar types in our training set. Our result is shown in Fig. 2.5. The main sequence and the red clump stars are well represented, and the regions sparsely covered by the training set are assigned low probabilities. By applying this GMM prior to the entire BP/RP spectral dataset, we implicitly assume that the stellar population of all BP/RP sources are well represented by the stars with LAMOST counterparts. Although this assumption is not true in detail, the purpose of our prior is primarily to prevent Θ from straying into regions of parameter space in which our model is unconstrained by data. In general, we expect the likelihood terms from the observed BP/RP spectra and parallax to dominate, and for the prior to play a subdominant role.

Combining all of our constraints, the loss function we finally use in the determi-

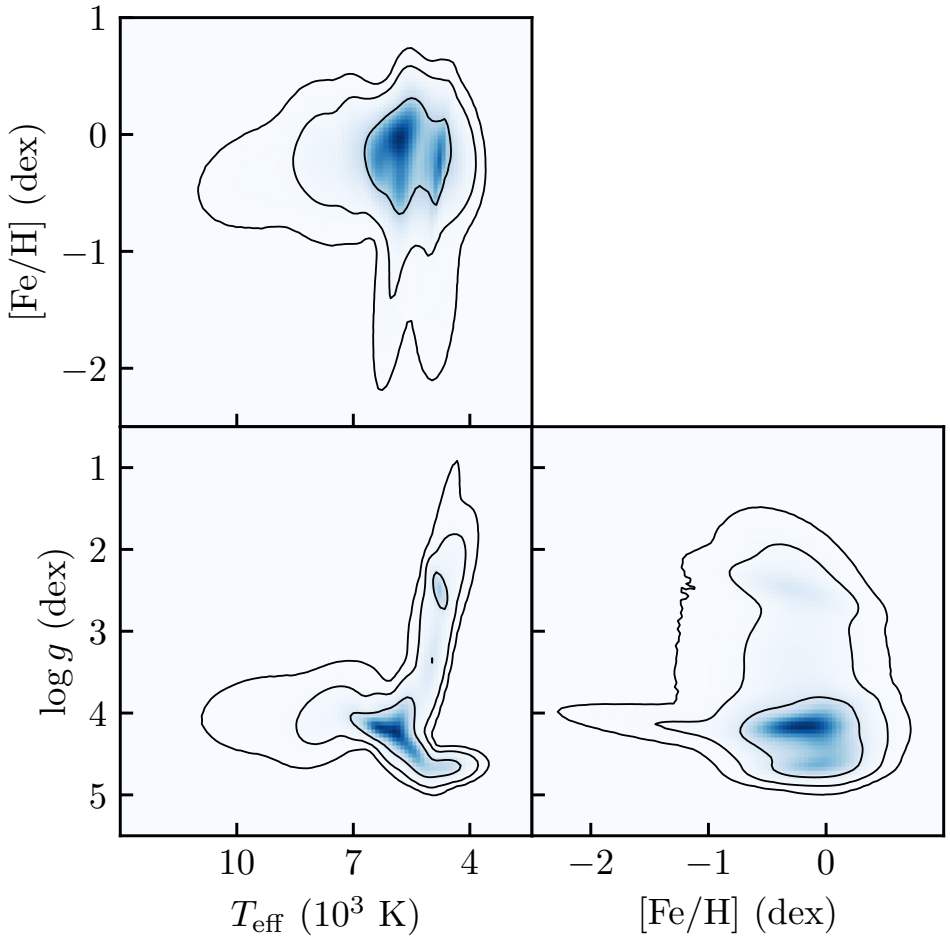


Figure 2.5. The prior assumption about the distribution of stellar types, represented by a Gaussian mixture model with N=16, with contours enclosing the 68%, 95%, 99% percentiles. The Gaussian mixture model (GMM) is a linear combination of 16 independent normal distributions. As shown on the picture, the main sequence and the red clump stars are well represented, and regions are ruled out by the GMM, where the training set does not cover. This prior distribution prevents unrealistic solutions when optimizing stellar types, such as meaningless interpolations of parameters at locations where there are not enough stars in the training set.

nation of stellar parameters is:

$$\begin{aligned}\mathcal{L}_{\text{inference}}(\Theta, \ln\varpi, \ln E) &= -2\ln(\text{Posterior}) \\ &= \chi_f^2 + \left(\frac{\varpi - \hat{\varpi}}{\sigma_\varpi}\right)^2 - 2\ln p_{\text{gmm}}(\Theta) \\ &\quad - 2\ln\varpi - 2\ln E \end{aligned} \tag{2.18}$$

The first three terms correspond to the constraints from BP/RP flux, parallax and the prior of Gaussian mixture model, respectively. The last two terms are the Jacobian terms to convert variables from ϖ and E to $\ln\varpi$ and $\ln E$.

2.4.2 Optimization of stellar parameters

Before optimizing the parameters of each star, we first conduct a rough grid search to locate a reasonable starting point. To construct our grid of possible starting points, we draw 512 samples from the Gaussian mixture model as the initial guess of stellar types, which is learned from the training set in section 2.4.1. For each stellar type, we check a geometric series of extinction values, from 10^{-2} to 10, with a common ratio of $10^{0.05}$. For each combination of stellar type and extinction, we calculate a parallax guess for each star by comparing the ratio of observed flux and predicted flux at 1 kpc. During calculation of this parallax guess, we only use fluxes between 592 nm and 782 nm, in order to avoid the edges of the BP/RP spectrum, which are typically noisy. Finally, for each star, we select the starting point (Θ, ϖ, E) that minimizes $\chi^2 - 2p_{\text{GMM}}(\Theta)$.

After the starting points are determined, we use gradient descent to find the parameters that maximize the stellar posterior, which is equivalent to minimizing the loss function Equation (2.18). We separate our stars into batches of 32,768 (10^{15}), conducting several rounds of optimization of stellar parameters $\vartheta = (T_{\text{eff}}, [\text{Fe/H}], \log g, E, \varpi)$ on each batch. In the first round of optimization, we use the Adam optimizer, with an

initial learning rate of 0.01 and 8,192 steps, halving the learning rate every 768 steps. For each round of optimization, we optimize the stellar parameters for 12288 steps to minimize the loss function of Equation 2.18, in which we set the initial learning rate as 0.01, and reduce it by 8 every 512 steps. After each round of optimization, we compute the Hessian matrix of the loss for each star:

$$H[\mathcal{L}_{\text{inference}}, \vartheta]_{i,j} = \frac{\partial^2 \mathcal{L}_{\text{inference}}}{\partial \vartheta_i \partial \vartheta_j} \,. \tag{2.19}$$

We select the stars that have non-positive-semidefinite Hessian matrices for further optimization. In subsequent rounds of optimization, we use pure Stochastic Gradient Descent (SGD), with 12,288 steps per round. In the n^{th} round of SGD optimization, we use an initial learning rate of $10^{-5}/2^{n-1}$, halving the learning rate every 4,096 steps. If fewer than 0.01% of the stars in a batch have non-positive-semidefinite Hessian matrices, we terminate optimization early and proceed to the next batch. Otherwise, we conduct at most 8 rounds of optimization for a given batch. Here we use (the eigenvalues of the) Hessian matrices as the criterion of convergence, because non-positive-semidefinite Hessian matrices can indicate that a minimum loss has not been reached.

2.4.3 Estimation of uncertainties

We calculate uncertainties for our inferred stellar parameters using the shape of the likelihood function in the neighborhood of the best fit. In detail, we make use of the Fisher information matrix of the likelihood, which is related to the Hessian matrix, but which is guaranteed to be positive-semidefinite.

For our model, the Fisher information matrix of the combined flux and parallax

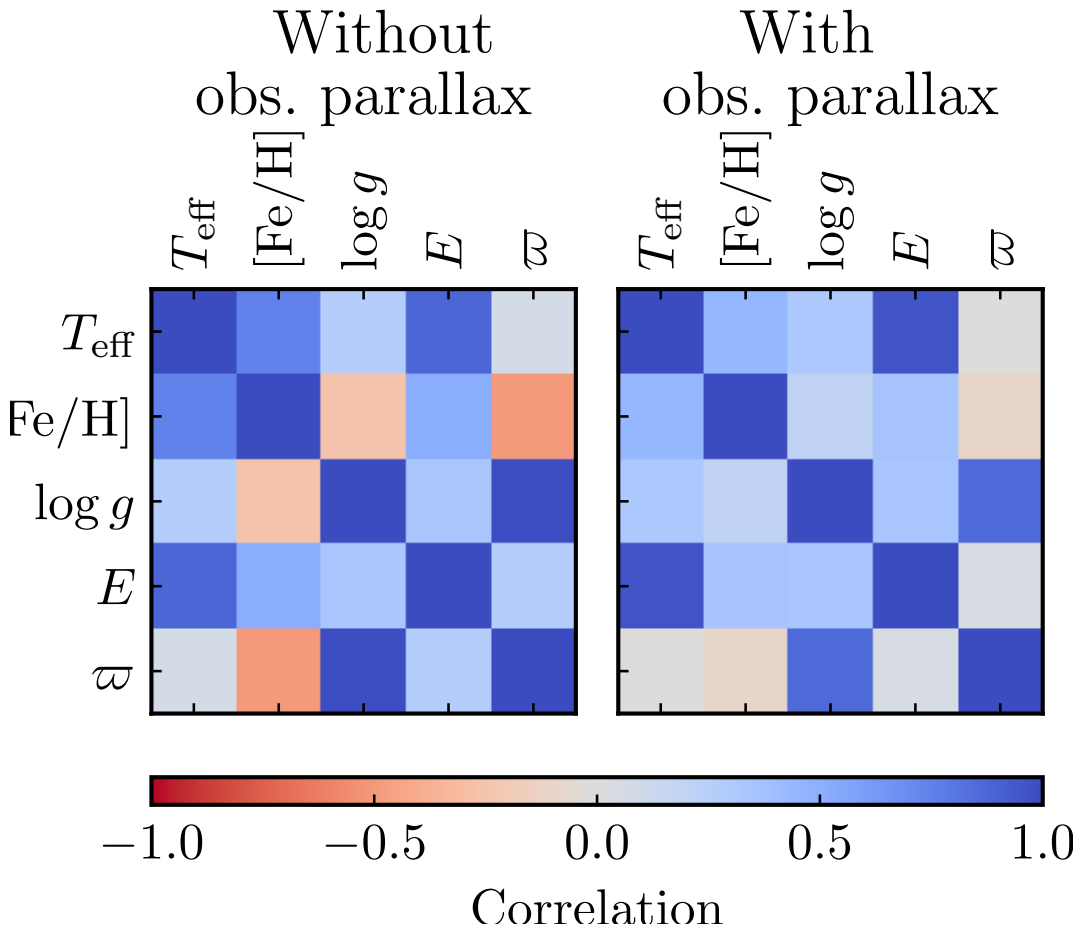


Figure 2.6. Observed parallaxes reduce the correlations among the stellar parameter estimates. The panels show the median correlations in the uncertainties of inferred stellar parameters before (left) and after (right) taking observed *Gaia* parallaxes into account. The left panel only shows the correlations given by flux (*i.e.* using only the first term in Eq. 2.21). The right panel contains the constraint from *Gaia*'s observed parallaxes (*i.e.*, the second term in Eq. 2.21). Correlations between several pairs of parameters are significantly reduced by inclusion of a parallax observation: $(\varpi, \log g)$, $(\varpi, [\mathrm{Fe/H}])$, (ϖ, E), $(T_{\rm eff}, [\mathrm{Fe/H}])$, and $([\mathrm{Fe/H}], E)$. The strongest remaining correlation in the right panel is between $T_{\rm eff}$ and E, which both affect the slope of the spectra similarly.

likelihood is given by

$$\mathcal{I}_{\alpha,\beta}\left(\vartheta\right) = \sum_{i,j} \frac{\partial f_{\text{pred},i}(\vartheta)}{\partial \vartheta_\alpha} [C_f^{-1}]_{i,j} \frac{\partial f_{\text{pred},j}(\vartheta)}{\partial \vartheta_\beta} \tag{2.20}$$

$$+ \text{diag}\left(0, 0, 0, 0, \sigma_\varpi^2\right)_{\alpha,\beta} , \tag{2.21}$$

where $\vartheta \equiv (\Theta, E, \varpi)$ is the full set of stellar parameters we fit, C_f^{-1} is the inverse covariance matrix of the observed fluxes (see Eqs. 2.1–2.2), and $\frac{f_{\text{pred},i}(\vartheta)}{\partial \vartheta_\alpha}$ is the partial derivative of the i^{th} component of the predicted flux by the α^{th} stellar parameter. As we implement our model in TensorFlow, these derivatives can be calculated quickly and accurately with automatic differentiation. The final term, which depends on σ_ϖ^2, represents the the additional information contributed by *Gaia*'s parallax measurement.

We evaluate the Fisher information matrix at the best-fit value of ϑ for each source. If our model $\vec{f}_{\text{pred}}$ depended linearly on ϑ, and if all of our measurements had perfectly Gaussian uncertainties, then the Fisher information matrix would be equal to the inverse covariance matrix of ϑ. In general, the Fisher information matrix produces a *lower bound* (the "Cramér-Rao bound") on the parameter uncertainties. We use the Fisher information matrix as an estimate of our inverse covariance matrix, with the caveat that the true covariance matrix may encode somewhat larger uncertainties:

$$C_\vartheta \equiv \left[\mathcal{I}\left(\vartheta\right)\right]^{-1} . \tag{2.22}$$

Note that we do not take our GMM prior (see Eq. 2.17) on stellar type Θ into account when estimating our covariance matrix. This prior is intended mainly to prevent our inference from straying towards stellar types that are unphysical, or for which our model is unconstrained by training data. We therefore estimate our uncertainties

solely on the basis of the flux and parallax likelihoods.

Fig. 2.6 shows the median (over the entire XP catalog) Pearson correlation between each pair of stellar variables (where the correlation ρ of variables α and β is related to the covariance matrix C by $\rho_{\alpha,\beta} = C_{\alpha,\beta}/\sqrt{C_{\alpha,\alpha}C_{\beta,\beta}}$), both before (left) and after (right) the constraints from *Gaia*'s observed parallaxes are incorporated into the uncertainty estimate (*i.e.*, with and without the last term in Eq. 2.21).

If we rely on the XP spectra alone, there is a strong degeneracy between our inferred $\log g$ and ϖ. This is because the *Gaia* XP spectral resolution is not fine enough to precisely measure linewidths. Holding other parameters constant, lower surface gravity corresponds to stars with larger radii, meaning that the major observable effect of $\log g$ on *Gaia* XP spectra is to increase or decrease the overall luminosity of the star. Parallax, ϖ, also causes an overall scaling of apparent flux, leading to a strong degeneracy between ϖ and $\log g$. Inclusion of the constraint from *Gaia*'s observed parallax lessens this degeneracy, decreasing the correlation coefficient between ϖ and $\log g$ from 0.987 to 0.861. There are a number of other moderately strong degeneracies that are reduced by taking *Gaia*'s parallax measurements into account. Correlations are reduced between ϖ and [Fe/H] (from -0.497 to -0.098), ϖ and E (from 0.266 to 0.047), $T_{\rm eff}$ and [Fe/H] (from 0.748 to 0.448), and [Fe/H] and E (from 0.515 to 0.352). The strongest remaining degeneracy in our model is between $T_{\rm eff}$ and E. This is because to first order, both parameters change the slope of the observed spectra. An increase in inferred temperature can be balanced by an increase in inferred reddening, rendering the color of the star nearly unchanged. This degeneracy is not lessened by inclusion of observed parallaxes.

2.4.4 Validation of stellar parameters

We validate our method by applying it to our validation set. We conduct the same optimization procedure discussed in Section 2.4.2, which does not take into account

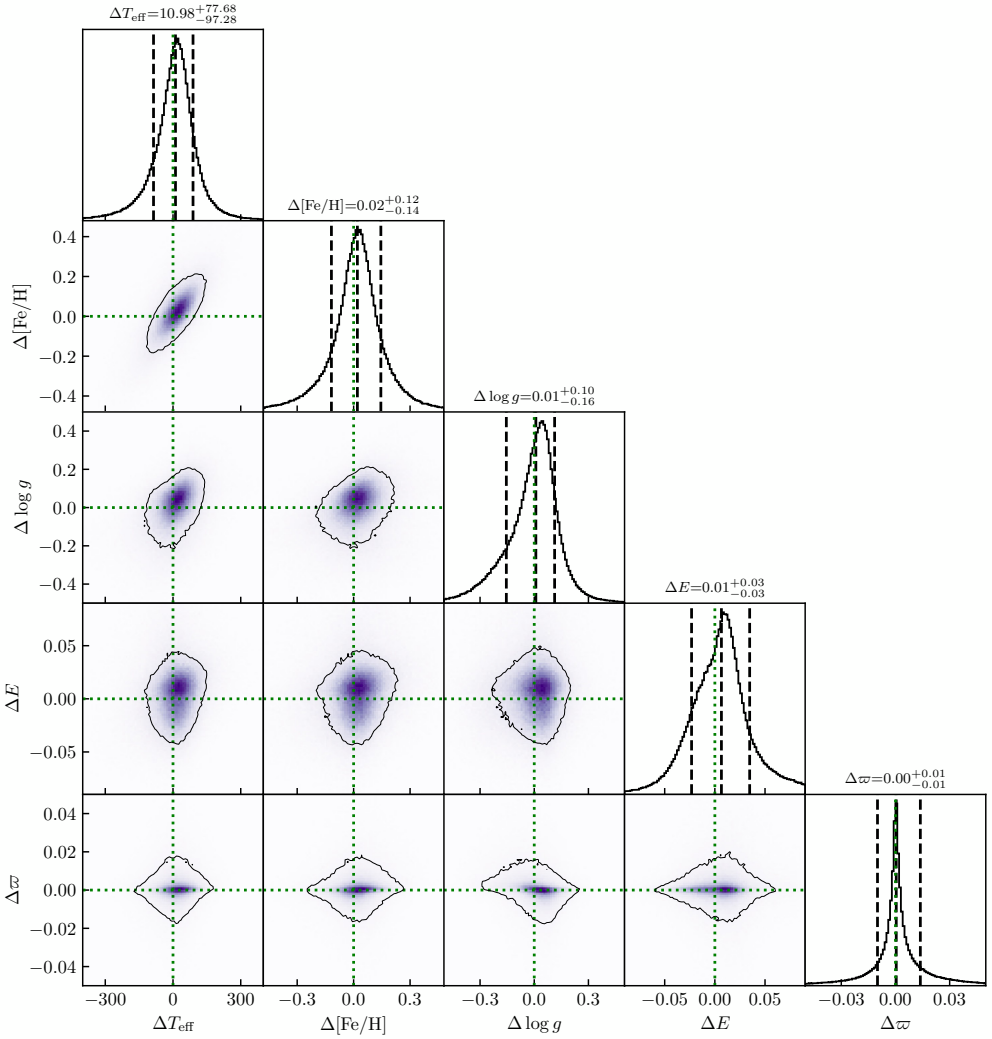


Figure 2.7. Difference between inferred stellar parameters $(T_{\rm eff}, [{\rm Fe/H}], \log g, E, \varpi)$ and the values determined by LAMOST, Bayestar19 and GDR3 in our validation dataset. We remove stars with bad parallax measurement or bad fitting results, which take $\sim$ 3% of all the stars. The contours enclose 68% of the probability. These parameter estimates are determined using the loss function Eq. (2.18), which encodes a likelihood of the observed XP spectra and GDR3 parallax, with flat priors on E and ϖ and a simple Gaussian mixture model prior on $(T_{\rm eff}, [{\rm Fe/H}], \log g)$. Our model performs well on the validation dataset, obtaining nearly unbiased estimates of the parameters, with typical uncertainties of $\sim 90\,{\rm K}$ in $T_{\rm eff}$, $\sim 0.15\,{\rm dex}$ in $[{\rm Fe/H}]$ and $\log g$, and $0.03\,{\rm mag}$ in E, although these uncertainties depend on the SNR of the observed XP spectra and parallaxes.

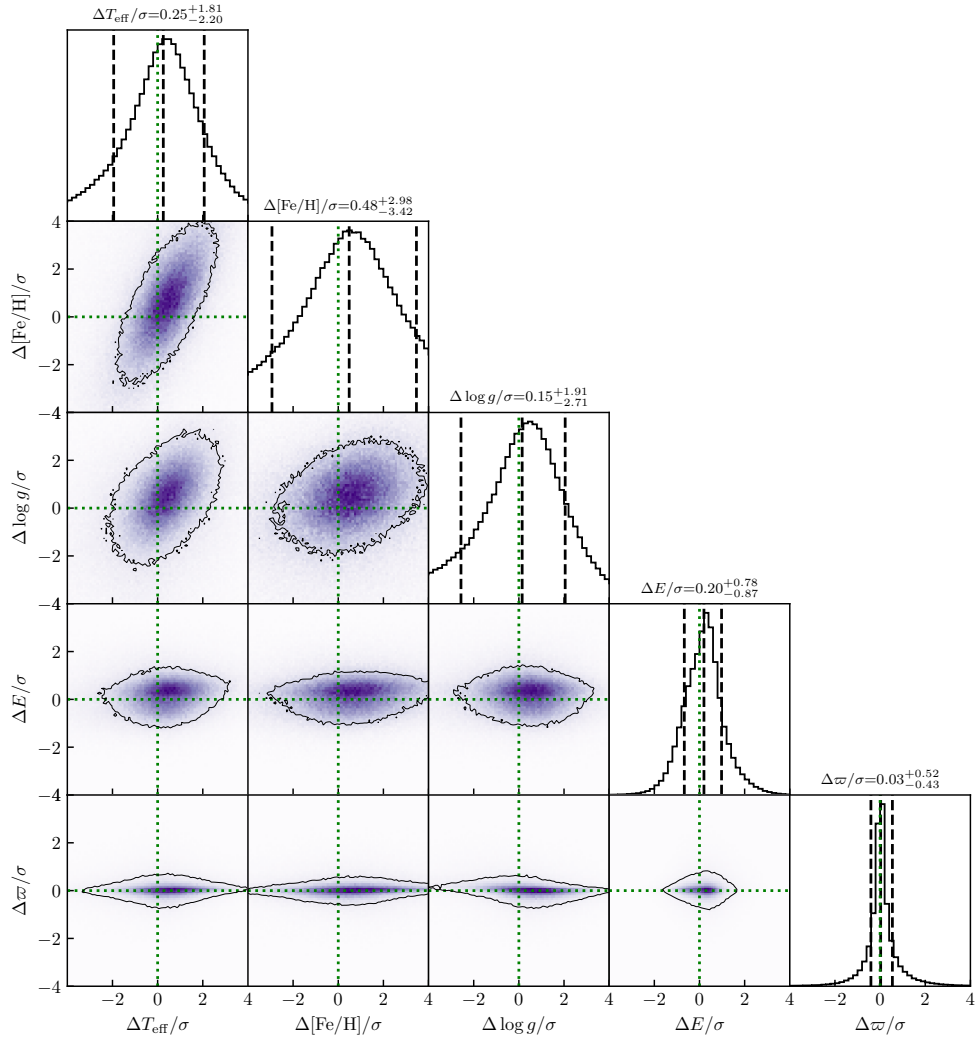


Figure 2.8. As Fig. 2.7, but with residuals normalized by the uncertainties reported by LAMOST, Bayestar19 and GDR3. In our validation dataset, we find that our stellar parameters typically agree with those reported by LAMOST to within $1.5-3\sigma$. Specifically, We find that the deviations are of the same order of magnitude as those from LAMOST, Bayestar19 and GDR3: $\sigma_{\rm inference,\ T_{eff}} \simeq 2\sigma_{\rm LAMOST,\ T_{eff}}$, $\sigma_{\rm inference,\ [Fe/H]} \simeq 3\sigma_{\rm LAMOST,\ [Fe/H]}$, $\sigma_{\rm inference,\ log\,g} \simeq 2\sigma_{\rm LAMOST,\ log\,g}$, $\sigma_{\rm inference,\ E} \simeq 0.8\sigma_{\rm Bayestar19,\ E}$, and $\sigma_{\rm inference,\ \varpi} \simeq 0.5\sigma_{\rm GDR3,\ \varpi}$. Our inferred uncertainties of ϖ are even smaller than those reported by GDR3, because our parallax estimates are based on both the observed GDR3 parallaxes and the observed XP spectra.

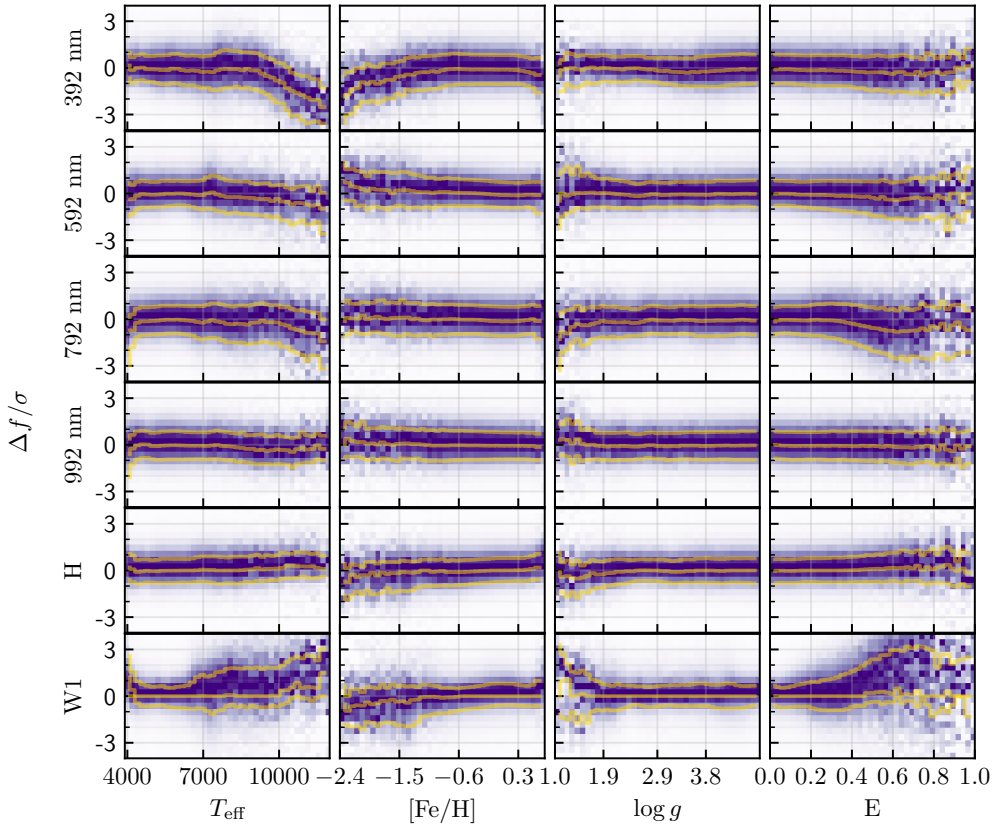


Figure 2.9. Normalized flux residuals ($\triangle f/\sigma \equiv (f_{\rm pred} - f_{\rm obs})/\sigma_f$) in the validation set, as a function of $T_{\rm eff}$, [Fe/H], $\log g$ and E. The densities are normalized by the maximum value at each parameter value (*i.e.*, in each pixel column). The yellow lines mark the position of the $15^{\rm th}$, $50^{\rm th}$ and $84^{\rm th}$ percentiles of the normalized residuals. During the optimization of parameters in the validation set, the model is unaware of the LAMOST and Bayestar19 estimates of stellar atmospheric parameters and extinction, respectively. The fluxes predicted by the model match the observed values to within $1 \sim 2\sigma$ over a wide range of parameter values, with the largest residuals occurring in $W1$ (particularly for $T_{\rm eff} \gtrsim 7000\,{\rm K}$ or $E \gtrsim 0.4$), and at 392 nm for $T_{\rm eff} \gtrsim 9000\,{\rm K}$.

the LAMOST estimates of stellar type or the Bayestar19 reddening estimates. We then compare our inferred stellar parameters with those determined by LAMOST (Θ), Bayestar19 (E), and GDR3 (ϖ). In order to estimate the typical errors with proper samples, we also remove outliers. Since the high-resolution spectroscopic constraints on $\log g$ are not available for most of the 220 million XP sources, we cannot apply the self-cleaning method used during training to remove suspected binary stars (see Section 2.3.2 and Fig. 2.4). We only remove stars with poorly constrained GDR3 parallax measurements ($\hat{\varpi}/\sigma_{\varpi} < 5$) and the stars with bad fit quality ($\mathcal{L}_{\text{inference}}/\text{DOF} > 5$), which altogether make up $\sim$ 3% of stars in the validation set.

Fig. 2.7 shows the residuals between our inferred stellar parameters and the parameters reported by LAMOST, Bayestar19 and GDR3 (hereafter, the "observed" parameters), for stars in the validation set. The inferred and observed parameters match with typical errors of $\sigma_{T_{\text{eff}}} \approx 90\,\text{K}$, $\sigma_{[\text{Fe/H}]} \approx 0.15$, $\sigma_{\log g} \approx 0.15$, $\sigma_E \approx 0.03\,\text{mag}$, and $\sigma_{\varpi} \approx 0.01\,\text{mas}$ (with σ_{ϖ} depending strongly on distance). We find that the errors are of the same order of magnitude as the observed uncertainties, as shown in Fig. 2.8. We can therefore conclude that the parameters inferred using our model match the observations well, with no prominent biases, and that the typical residuals are comparable to the uncertainties of the observations.

We additionally estimate the quality of our predicted fluxes (evaluated at the parameter values inferred by our model) by comparison with the observed fluxes. Fig. 2.9 shows the normalized flux residuals (prediction minus observation, divided by the uncertainty in the observation) in the validation set at 392, 592, 792 and 992 nm, and in 2MASS H- and unWISE W1-band, as a function of T_{eff}, [Fe/H], $\log g$ and E (as estimated by LAMOST and Bayestar19). Over most of the parameter space, the model prediction matches the observation within $\pm 1\sigma$.

Our stellar model does not cover all possible stellar types, and we therefore do not obtain reliable parameter estimates for all types of stars. In general, we recommend

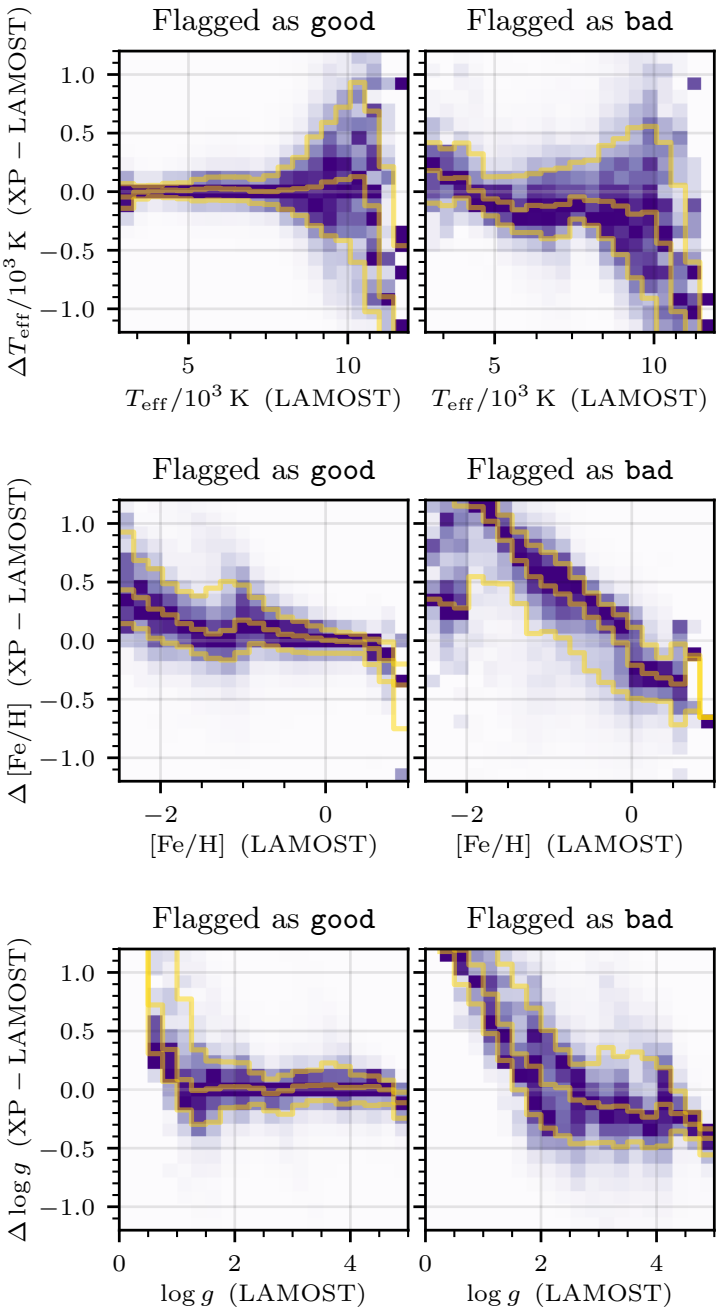


Figure 2.10. Residuals (model – LAMOST) in $T_{\rm eff}$ (top panels), [Fe/H] (middle panels) and $\log g$ for sources flagged as reliable (`good`, left panels) vs. unreliable (`bad`, right panels), as a function of the LAMOST atmospheric parameter estimates. The densities are normalized by the maximum value at each parameter value (*i.e.*, in each pixel column). The yellow lines mark the positions of the $15^{\rm th}$, $50^{\rm th}$ and $84^{\rm th}$ percentiles of the residuals. The reliability classifiers are unaware of the LAMOST estimate, working instead with features that are available for all sources (including those without LAMOST observations). For each parameter, there is a stark difference between stars flagged as `good` and `bad`. There is a nearly flat trend in the median [Fe/H] residuals of stars labeled `good` over a wide range of metallicities. $\log g$ estimates agree well for $1 \lesssim \log g \lesssim 5$, but agreement degrades for $\log g \lesssim 1$, where there is little LAMOST training data. $T_{\rm eff}$ residuals are approximately flat for $T_{\rm eff} \lesssim 11,000\,{\rm K}$, though the the scatter in the residuals increases dramatically for $T_{\rm eff} \gtrsim 7,500\,{\rm K}$.

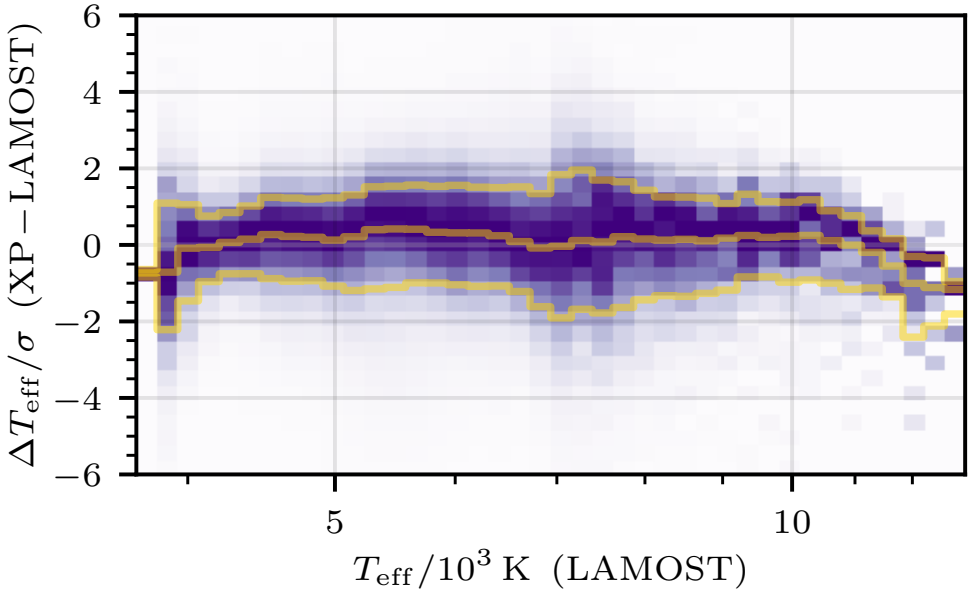


Figure 2.11. The distribution of normalized $T_{\rm eff}$ residuals (model minus LAMOST, divided by combined model and LAMOST uncertainties, as a function of LAMOST $T_{\rm eff}$. We include only estimates flagged as `good` by our $T_{\rm eff}$ reliability classifier. By comparison with the top-left panel of Fig. 2.10, we see that although the scatter in absolute $T_{\rm eff}$ residuals grows past $T_{\rm eff} \gtrsim 7,500$ K, the reported $T_{\rm eff}$ uncertainties also grow proportionally, leading to a well-behaved distribution of normalized residuals.

applying the following set of reliability cuts on our catalog of stellar parameters:

$$\texttt{chi2_opt}/61 < 2 \tag{2.23}$$

$$\texttt{ln_prior} > -7.43 \tag{2.24}$$

$$\frac{|\texttt{gaia_parallax} - \texttt{parallax_est}|}{\texttt{gaia_parallax_error}} < 10 \tag{2.25}$$

For convenience of users of the catalog, we combine these cuts together into one flag, which we call `reliability`. The first cut corresponds to a cut on χ^2/dof, which captures whether the predicted spectrum matches the observed spectrum. The second cut ensures that the inferred stellar atmospheric parameters lie in a region that was covered by our LAMOST training dataset. The threshold on this cut is chosen so that 99.9% of the probability mass of our prior on the stellar atmospheric parameters is enclosed in the allowed region. The third cut removes sources for which our parallax estimate strongly disagrees with that of *Gaia*.

In addition to the above set of cuts, we develop a separate "confidence" measurement for each stellar atmospheric parameter, which provides the possibility of further ruling out spurious parameter estimates. We train a neural-network classifier for each atmospheric parameter (T_{eff}, $[\mathrm{Fe/H}]$, $\log g$), which predicts whether or not our estimate agrees well with LAMOST. We define agreement based on the metric

$$\chi^2_\epsilon \equiv \frac{(\Delta x)^2}{\sigma^2_{\mathrm{opt}} + \sigma^2_{\mathrm{LAMOST}} + \epsilon^2}\,, \tag{2.26}$$

where Δx is the residual between our optimized parameter and LAMOST; σ_{opt} and σ_{LAMOST} are the uncertainties in our and LAMOST's parameter estimates, respectively; and ϵ is an uncertainty floor that we add in quadrature, so that sufficiently small residuals are considered to be in agreement. We use $\epsilon = 100\,\mathrm{K}$ for T_{eff}, and 0.1 dex for $[\mathrm{Fe/H}]$ and $\log g$. For each atmospheric parameter, we construct a train-

ing set by defining stars as `bad` when $\chi^2_\epsilon > 9$, and as `good` when $\chi^2_\epsilon < 4$. Our parameter estimates for this training set are derived without knowledge of LAMOST. During training, we withhold 20% of the LAMOST sources to evaluate the performance of our flags. We only train this model on sources that pass our basic reliability cut (see Eq. 2.25). Our parameter estimates for sources which fail the basic reliability cut can be disregarded, regardless of our neural network "confidence" estimate.

As we would like to eventually apply these classifiers to the entire XP dataset, we can only take into account features that are available for every source (including sources not observed by LAMOST). These features include the reduced χ^2 of our fit, the inferred $T_{\rm eff}$, $[{\rm Fe/H}]$ and $\log g$, the *Gaia* `parallax_over_error`, and the normalized flux residuals (model minus observed, divided by observational uncertainty) at every wavelength. More details about our classifier, including complete list of features used, are given in section 2.10.

Each "confidence" classifier assigns a number between 0 and 1 to each source, which is the probability that our parameter estimate ($T_{\rm eff}$, $[{\rm Fe/H}]$ or $\log g$) is reliable. We report these probabilities for every XP source. Fig. 2.10 shows the performance of our confidence measurements on the 20% of data withheld during training of the classifiers, considering only sources that pass our basic reliability cuts (Eq. 2.25). The distribution of $T_{\rm eff}$, $[{\rm Fe/H}]$, and $\log g$ residuals (model – LAMOST) is plotted on the y-axis for stars labeled `good` (left panels) vs. `bad` (right panels), using a threshold of 0.5 to distinguish the two classes. Along the x-axis, we separate stars into bins based their LAMOST parameter estimates. For all three atmospheric parameters, stars labeled `good` have much better behaved residuals than stars labeled `bad`. Fig. 2.11 focuses on the $T_{\rm eff}$ residuals, *normalized* by the reported uncertainties. Although the absolute scale of the residuals are large for hot stars, the normalized residuals continue to be well behaved. Our model thus delivers highly uncertain $T_{\rm eff}$

estimates for hot stars, which are nevertheless statistically well behaved with respect to LAMOST. In Table 2.3, we show the number of stars passing each quality flag.

2.5 Results

This results section consists of three parts. First, we present our trained model, which maps stellar parameters ($T_{\rm eff}$, [Fe/H], $\log g$, E, ϖ) to predicted *Gaia* XP, 2MASS and WISE fluxes. Second, we apply our model to determine the stellar parameters of all 220 million XP sources in GDR3. Third, we discuss possible uses of our inferred stellar parameters. We show a preliminary three-dimensional dust map, based on our distance and extinction estimates, which shows evidence of unmodeled extinction curve variations. We also explore the metallicity distribution as a function of position in the Milky Way, based on our inferred stellar [Fe/H] values.

2.5.1 Trained model of stellar spectra and extinction

We obtain a model that predicts *Gaia* XP spectra (from 392-992 nm, with 10 nm resolution), as well as 2MASS and WISE ($W1$ and $W2$) bands, for any given stellar atmospheric parameters ($T_{\rm eff}$, [Fe/H] and $\log g$), extinction (E) and parallax (ϖ). This requires our model to predict not only stellar absolute flux, but also to capture the wavelength-dependence of dust extinction.

Fig. 2.12 shows our model spectra for Solar-metallicity main-sequence stars from 4,000-11,000 K. In this figure, we set $\log g$ using a piecewise-linear function of $T_{\rm eff}$:

$$\log g_{\rm MS} = \begin{cases} 4.6\,, & T_{\rm eff} < 5000 \\ 4.6 - 5.0\times10^{-4}(T_{\rm eff} - 5000)\,, & 5000 \leq T_{\rm eff} < 6300 \\ 3.95\,, & T_{\rm eff} \geq 6300 \end{cases} \tag{2.27}$$

In the right panel of the figure, we plot the continuum-normalized model XP spectra,

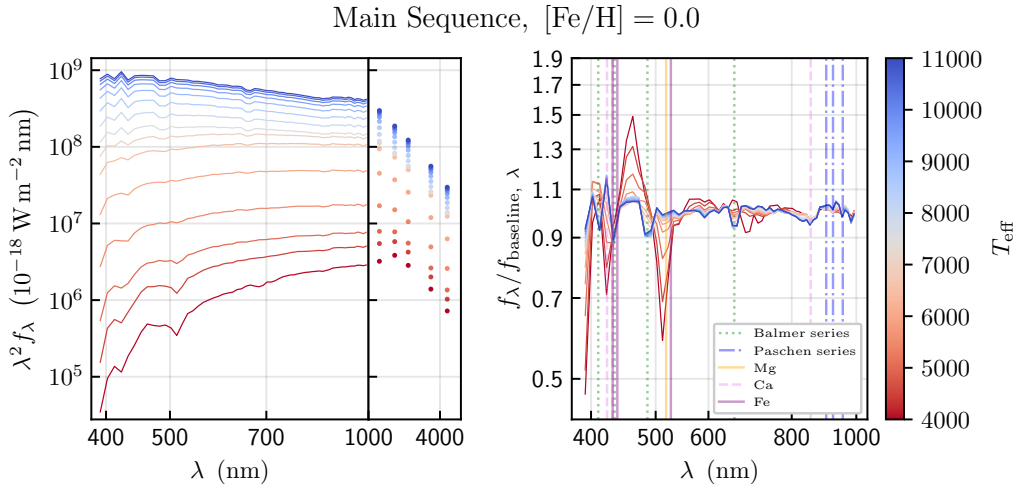


Figure 2.12. Model-predicted, zero-extinction fluxes (at 1 kpc) for the XP spectra and the NIR photometry (left panel), illustrated here for Solar-metallicity main sequence stars as a function of effective temperature. We set $\log g$ as a function of T_{eff} (see Eq. 2.27). The right panel shows continuum-normalized XP model spectra to highlight the spectral line features. Balmer and Paschen series are shown as dotted and dash-dotted lines. We also show the position of Ca lines and Mg b_1 line. As effective temperature increases, the flux increases, and Balmer lines become stronger. We also show the position of the Paschen series, and find that there is a weak trend of strengthening with the increase of T_{eff}. It is clear that Ca lines around ~ 430 nm and the Mg line attenuate as T_{eff} increase, but the trend for other metal lines are unclear, due to saturation and the low resolution of the XP spectra. Our full model, along with example code to evaluate it for arbitrary stellar types, extinctions and distances, is available at `https://doi.org/10.5281/zenodo.7692680`.

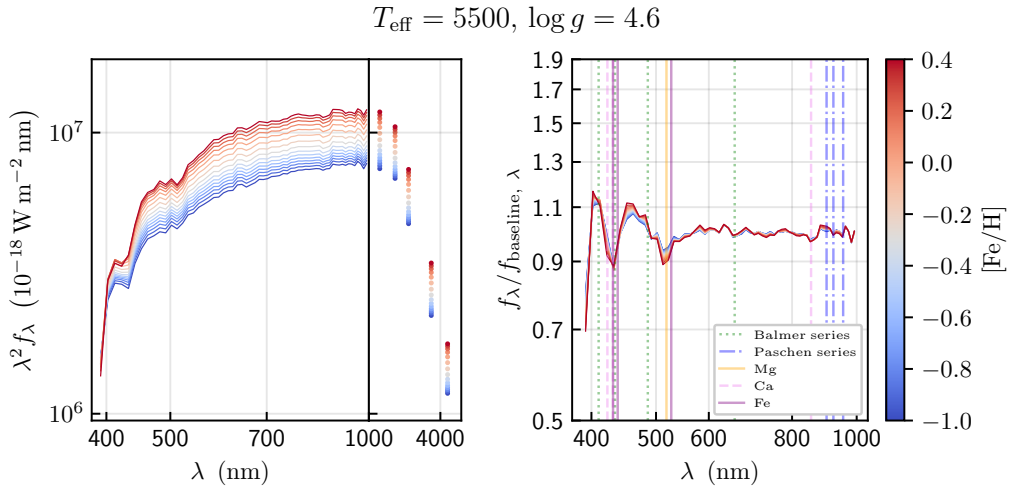


Figure 2.13. Predicted model fluxes as in Fig. 2.12, here shown as a function of [Fe/H] while fixing $T_{\rm eff} = 5500\,{\rm K}$ and $\log g = 4.6$. As in Fig. 2.12, the left panel shows spectral flux density for XP spectra and five NIR bands, while the right panel shows continuum-normalized XP spectra. As metallicity increases, Mg and Fe lines become deeper, while hydrogen lines do not vary significantly. There is also an overall increase of the flux with increasing [Fe/H].

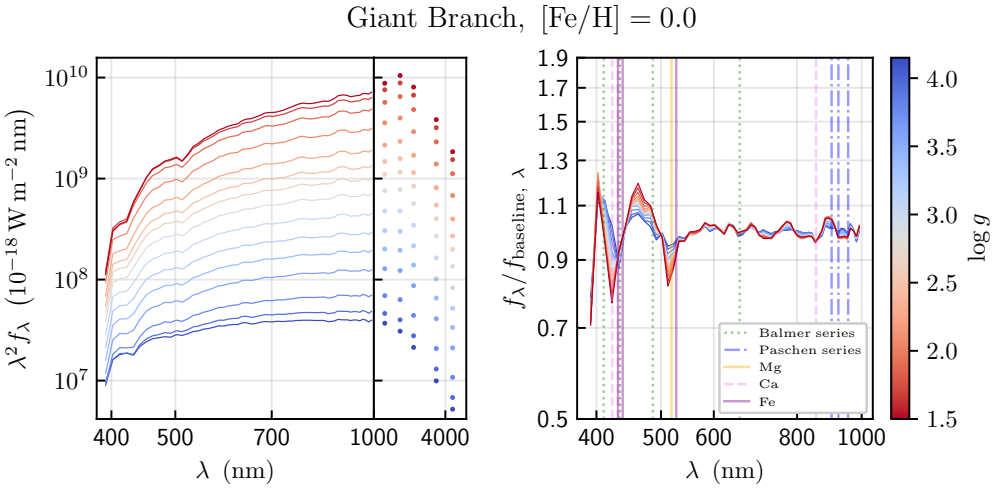


Figure 2.14. Predicted model fluxes as in Fig. 2.12, here shown as a function of $\log g$ along the giant branch. For this figure, we use a piecewise function to set $T_{\rm eff}$ as a function of $\log g$ along the giant branch (see Eq. 2.28). As in Fig. 2.12, the left panel shows spectral flux density for XP spectra and five NIR bands, while the right panel shows continuum-normalized XP spectra. Because the absorption lines are not resolved, the primary effect of $\log g$ is to change the overall luminosity of the star. The variation of the continuum slope and line depths is due to the dependence of $T_{\rm eff}$ on $\log g$ along the giant branch.

using a 5$^{\text{th}}$-order polynomial to represent each spectrum's continuum. Our model captures the temperature-dependence of both the slope of the flux continuum and of the depth of different classes of absorption lines. Although the low resolution of XP spectra makes it difficult to resolve individual absorption lines, a few classes of lines are nevertheless visible in our models. Both the Balmer and Paschen series appear in our models, with line strength increasing past $T_{\text{eff}} \gtrsim 6000\,\text{K}$, as expected, while features likely associated with Ca (at 423 nm) and Mg (at 523 nm) become stronger at lower temperatures. The Ca II triplet (from 850-867 nm) is present as a single blended line across a wide range of temperatures.

Fig. 2.13 shows the dependence of our model spectra on $[\text{Fe/H}]$ for 5,500 K main-sequence stars. Increasing $[\text{Fe/H}]$ not only changes the overall scale of the flux, but also strengthens metal lines, as expected. As expected, the strength of hydrogen lines is insensitive to $[\text{Fe/H}]$. A feature likely associated with Mg is not cleanly centered at the expected location of the absorption line, at 523 nm, possibly because of the low resolution of our model (10 nm), and because this feature may be a combination of different lines.

Fig. 2.14 shows our model spectra as a function of $\log g$ along the giant branch. For plotting purposes, we define effective temperature along the giant branch as a piecewise-linear function of $\log g$:

$$T_{\text{eff,GB}} = \begin{cases} 5200 - 441.86(3.65 - \log g)\,, & \log g < 3.65 \\ 5900 - 1400(4.15 - \log g)\,, & \log g \geq 3.65 \end{cases} \tag{2.28}$$

As individual absorption lines are not fully resolved, the primary observable effect of $\log g$ in XP spectra is to change the overall luminosity of a star. The variation of the spectral shape in Fig. 2.14 is due to the relation between T_{eff} and $\log g$ along the giant branch.

Our model contains a universal extinction curve, which describes the relative

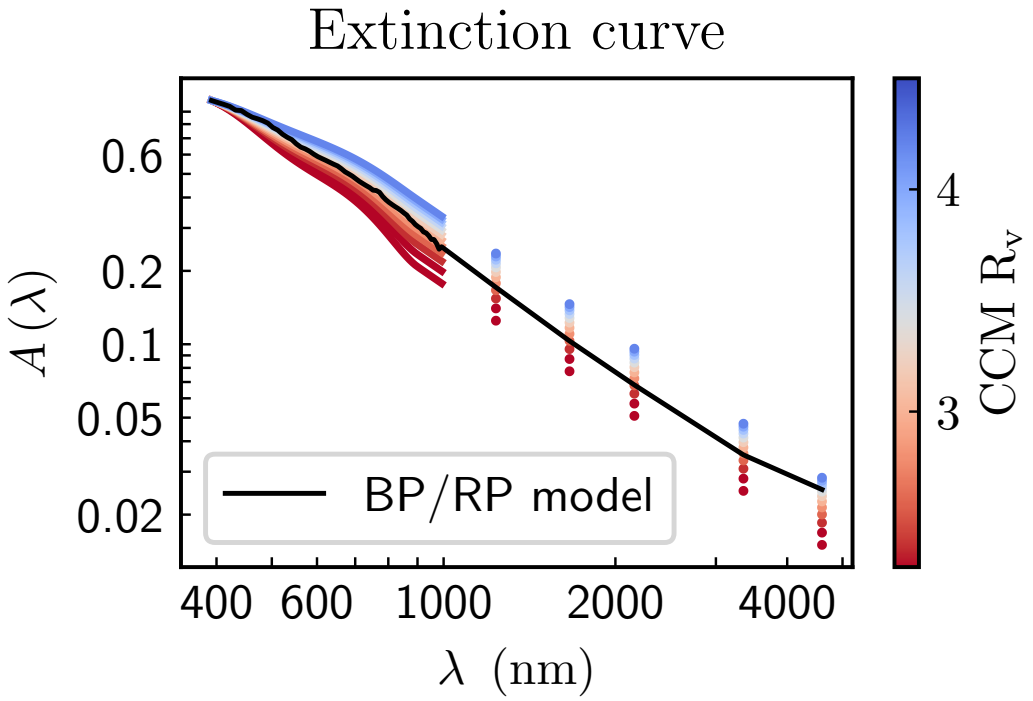


Figure 2.15. The extinction curve, assumed here to be universal across the sky, learned by our model (in black) from the data, compared with the CCM model (colored curves; Cardelli et al., 1989). All models are normalized at $\lambda = 392$ nm Although we apply no constraints to ensure smoothness in the extinction curve, our model learns a curve that varies smoothly with wavelength. We find that our extinction model roughly aligns with CCM extinction curves with $R_V \sim 2.8 - 3.2$ in the XP spectral region, but deviates towards slightly larger values of R_V in the NIR. Note, however, that the broad bandpasses of 2MASS and WISE cause extinction to vary by $\sim 5\%$, compared to the mono-wavelength extinction. The full extinction curve is available as an electronic table at `https://doi.org/10.5281/zenodo.7692680`.

extinction as a function of wavelength. Our model learns this extinction curve from the data alone, without reference to any previous extinction curve models. Each wavelength is modeled separately, without any prior on smoothness, meaning that the smoothness of the resulting extinction curve is a consequence purely of the data preferring such a curve. Fig. 2.15 shows our extinction curve (in black), and compares it with the family of R_V-dependent models of Cardelli et al. (1989, hereafter "CCM"). Our extinction model is smooth in the XP wavelength range, and roughly follows CCM models with $R_V \sim 2.8 - 3.2$. In the WISE W2 band, our extinction curve is slightly higher than the mean CCM curves, which corresponds to a deviation towards slightly higher values of R_V. This may be a reflection of the fact that our model predicts the extinction in the W2 band, which has a non-negligible spectral width. The extinction in the band is dependent on the slope of the stellar spectrum, which is generally sharply falling across W2 band. The band-weighted extinction is therefore expected to be larger than the monochromatic extinction at the center of the band.

2.5.2 220 million stellar parameter estimates

We apply our model to the entire XP dataset, using the method described in Section 2.4, to obtain stellar parameter estimates for 220 million sources, along with corresponding uncertainty estimates.

Our entire stellar parameter catalog is available for download at `https://doi.org/10.5281/zenodo.7692680`. We additionally plan to release the catalog through the Virtual Observatory in the near future. The columns of our catalog are described in Table 2.2.

Table 2.2 describes the contents of our catalog. 82% of our parameter estimates pass our basic reliability cut (see Eq. 2.25), which can be obtained from the catalog by requiring `quality_flags` < 8. Fig. 2.18 shows the fraction of sources that are

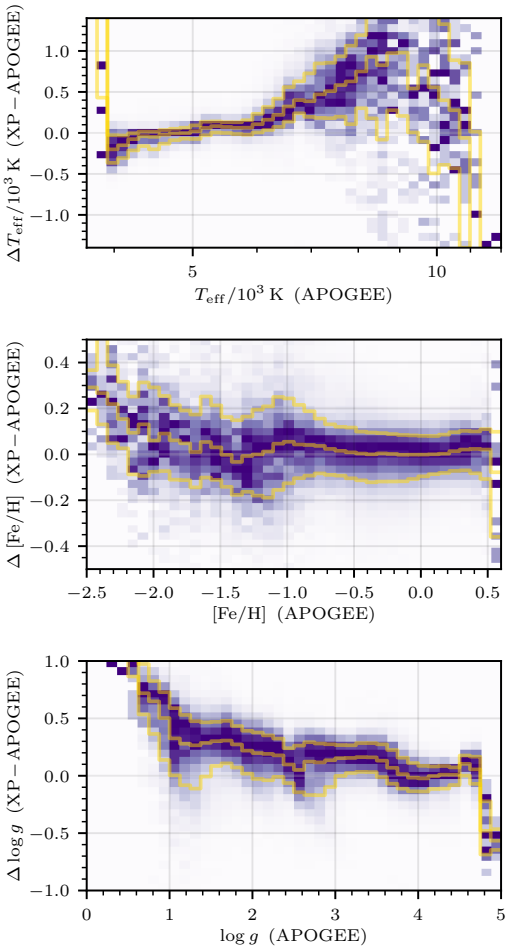


Figure 2.16. The distributions of residuals (model – APOGEE) in $T_{\rm eff}$ (top panel), [Fe/H] (middle panel), and $\log g$ (bottom panel), as a function of the respective APOGEE parameter estimates. The densities are normalized by the maximum value at each parameter value (*i.e.*, in each pixel column). The yellow lines mark the positions of the $15^{\rm th}$, $50^{\rm th}$ and $84^{\rm th}$ percentiles of the residuals. In each panel, we plot only sources labeled as `good` by the respective reliability classifier for the given parameter (See Section 2.4.4). We observe good agreement between our [Fe/H] and $\log g$ estimates and those of APOGEE. However, we observe a trend in our $T_{\rm eff}$ residuals vs. APOGEE. This is due to two factors: a trend in LAMOST vs. APOGEE $T_{\rm eff}$ estimates, and large uncertainties in our $T_{\rm eff}$ estimates for hot stars.

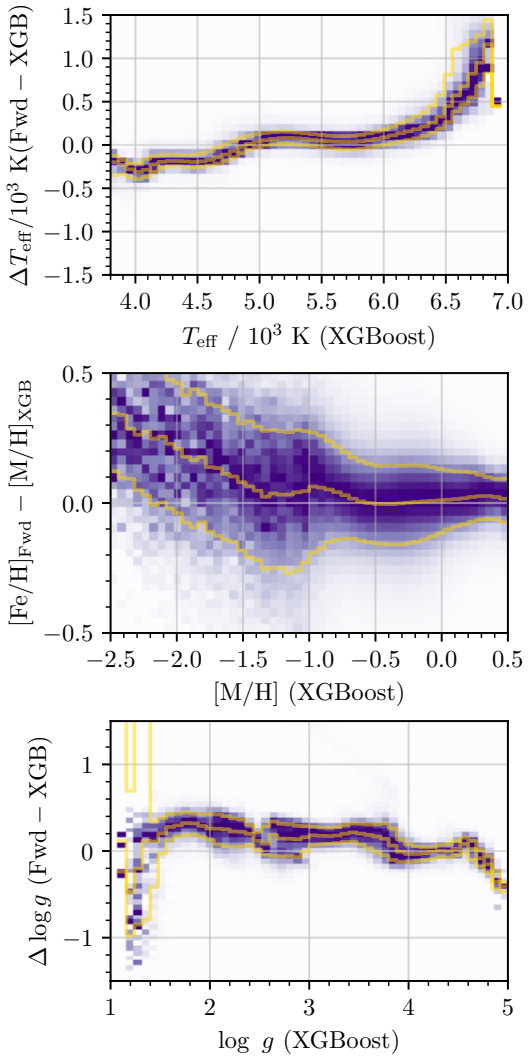


Figure 2.17. The distributions of residuals (model – XGBoost) in the same layout as Fig. 2.16. "Fwd" represents the forward model we develop in this work, while "XGB" represents the XGBoost results from Andrae et al. (2023). We only compare the parameters of stars in which both catalogs have confidence. In our catalog, we select the stars passing the basic cut. In each panel, we additionally require that the corresponding paramter have confidence > 0.5 (see Section 2.4.4). In the XGBoost catalog, we select stars with parallax $\varpi > 1$ mas, because the estimate of [M/H] in XGBoost is much less reliable for small parallax, according to Fig. 11 in Andrae et al. (2023). We find similar systematic trend in the $T_{\rm eff}$ residuals as we find in our comparison with APOGEE. This is due to the fact that XGBoost uses APOGEE as a training set, while we use LAMOST. A similar trend can be seen in a direct comparison of the underlying training datasets, LAMOST and APOGEE. Our [Fe/H] and $\log g$ estimates have better agreement (with a scatter of approximately 0.2 dex), though we observe a systematic trend in the [Fe/H] residuals for [Fe/H] $\lesssim -1.3$ and a systematic offset in $\log g$ of order 0.2 dex. As with the $T_{\rm eff}$ comparison, most of the systematic differences are due to disagreements between the underlying training datasets, LAMOST and APOGEE.

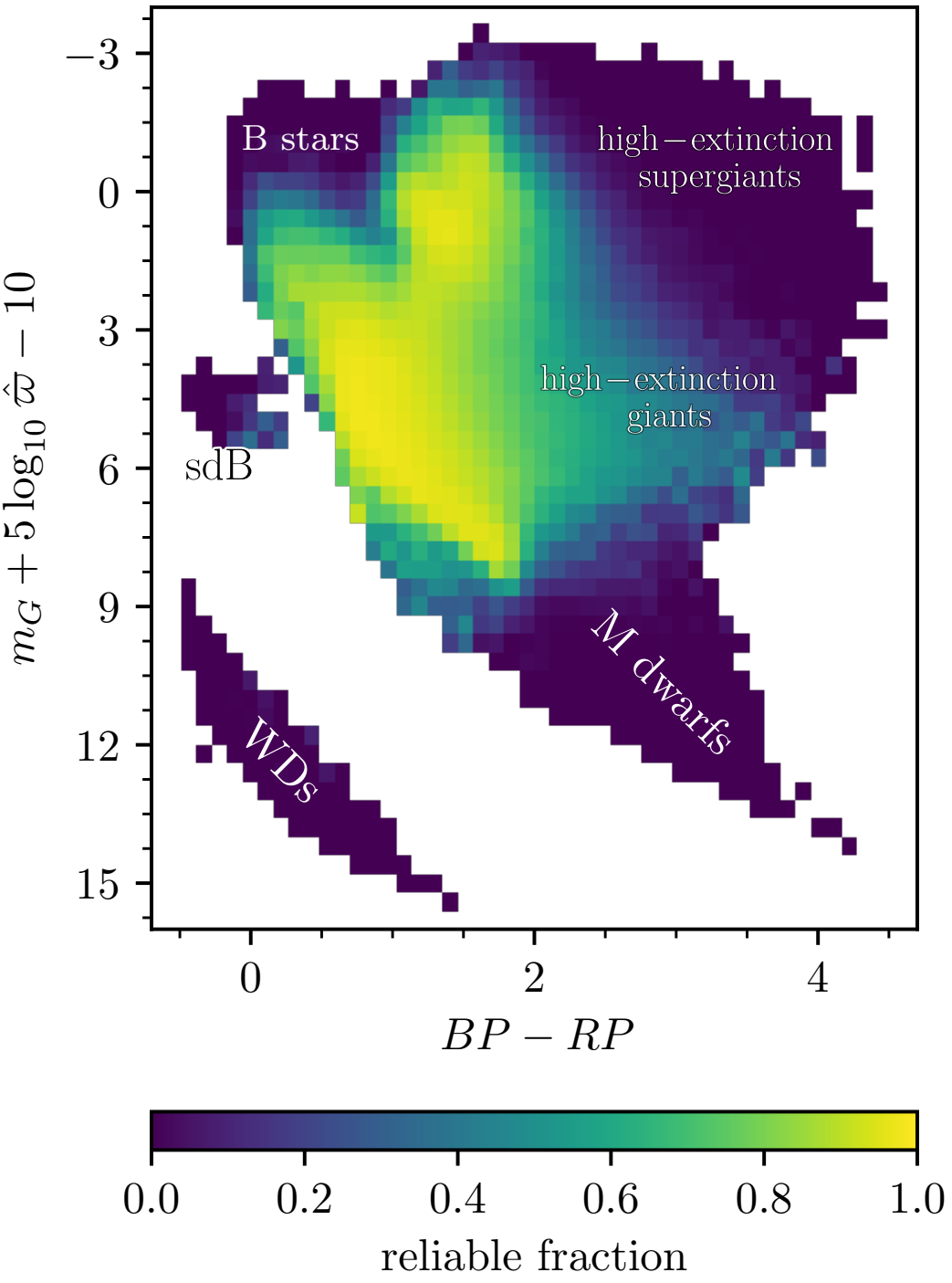


Figure 2.18. Fraction of sources that pass our standard reliability cut, as a function of position on the *Gaia* CAMD. While we obtain reliable stellar parameters across the parameter range covered by our LAMOST training dataset, a number of classes of stellar objects lie outside this range: M-dwarfs, main-sequence B-stars, white dwarfs, B-type subwarfs (sdB), and supergiants. In addition, due to our assumption of a universal extinction curve, the reliability of our parameter estimates drops at large extinctions. This can be seen above for high-extinction giants. Use of additional sources of training data beyond LAMOST would expand the range of stellar types covered by our model.

Name	Data type	Description
gdr3_source_id	integer	*Gaia* DR3 source_id.
ra	float	Right Ascension (in deg), as measured by *Gaia* DR3.
dec	float	Declination (in deg), as measured by *Gaia* DR3.
stellar_params_est	$5 \times$ float	Estimates of stellar parameters ($T_{\rm eff}$ in kiloKelvin, [Fe/H] in dex, $\log g$ in dex, E in mag, ϖ in mas).
stellar_params_err	$5 \times$ float	Uncertainties in the stellar parameters.
chi2_opt	float	χ^2 of the best-fit solution.
ln_prior	float	Natural log of the GMM prior on stellar type, at the location of the optimal solution.
teff_confidence	float	A neural-network-based estimate of the confidence in the $T_{\rm eff}$ estimate, on a scale of 0 (no confidence) to 1 (high confidence).
feh_confidence	float	As teff_confidence, but for [Fe/H].
logg_confidence	float	As teff_confidence, but for $\log g$.
quality_flags	8-bit uint	The three least significant bits represent whether the confidence in $T_{\rm eff}$, [Fe/H] and $\log g$ is less than 0.5, respectively. The $4^{\rm th}$ bit is set if chi2_opt$/61 > 2$. The $5^{\rm th}$ bit is set if ln_prior < -7.43. The $6^{\rm th}$ bit is set if our parallax estimate is more than 10σ from the GDR3 measurement (using reported parallax uncertainties from GDR3). The two most significant bits are always unset. We recommend a cut of quality_flags < 8 (the "basic reliability cut"), although a stricter cut of quality_flags $== 0$ ensures higher reliability at the cost of lower completeness.
stellar_params_icov_triu	$15 \times$ float	Upper triangle of the inverse covariance matrix of our stellar parameters.
stellar_params_cov_triu	$15 \times$ float	Upper triangle of the covariance matrix of our stellar parameters, obtained from the inverse covariance matrix in a numerically stable manner that ensures positive semi-definiteness.

Table 2.2. The columns of our stellar parameter catalog. The full catalog is available at https://doi.org/10.5281/zenodo.7692680.

Quality cuts	# of stars passing the cuts	Percentage
all stars	219,197,643	100%
Basic cuts	180,344,401	82.3%
Basic cut + Confident in $T_{\rm eff}$	126,501,649	57.7%
Basic cut + Confident in $\log g$	138,890,532	63.4%
Basic cut + Confident in $[{\rm Fe/H}]$	132,573,594	60.5%
Basic cut + Confident in $(T_{\rm eff}, [{\rm Fe/H}], \log g)$	90,390,241	41.2%

Table 2.3. Numbers of stars that pass different quality cuts. The "basic cut" is defined as quality_flags < 8, as described in Section 2.4.4 and Table 2.2. "Confident in param" is defined as confidence_param > 0.5.

judged reliable by this cut, as a function of position on the CAMD. We obtain reliable parameter estimates for stars in the parameter range covered by our LAMOST training set. We do not obtain reliable parameter estimates for M-dwarfs, white dwarfs, B-type subdwarfs, or supergiants. Additionally, our reliability fraction drops at very large extinctions, likely due to variation in the extinction curve. Our model learns a universal extinction curve, and at very large extinctions, slight variations in the slope of the extinction curve can have a large effect on the ability of our model to ability to accurately model the stellar spectrum.

We additionally use the classifiers described in Section 2.4.4 to assign a "confidence" estimate to each estimate of $T_{\rm eff}$, $[{\rm Fe/H}]$, and $\log g$. These classifiers assign greater than 0.5 confidence in our $T_{\rm eff}$, $[{\rm Fe/H}]$ and $\log g$ estimates for 61%, 68% and 71% of sources, respectively. A cut of `quality_flags` $== 0$ selects sources for which we are confident in $T_{\rm eff}$, $[{\rm Fe/H}]$ and $\log g$, and which additionally pass our basic reliability cut (Eq. 2.25).

2.5.3 External validation of stellar parameters

We validate our inferred stellar atmospheric parameters by comparison against an external catalog: APOGEE DR17. We crossmatch the XP sources with APOGEE

DR17 stars, rejecting stars that have any of the following APOGEE flags set: `METALS_BAD`, `SNR_BAD`, `CHI2_BAD`. When comparing each atmospheric parameter, we additionally require that our XP reliability probability be greater than 0.5 (See Section 2.4.4). Fig. 2.16 shows the distributions of residuals in $T_{\rm eff}$, $[{\rm Fe/H}]$ and $\log g$, as a function of the APOGEE measurement of each respective parameter. Our $[{\rm Fe/H}]$ measurements agree to within 0.1–0.2 dex above $[{\rm Fe/H}] \approx -2$, with somewhat larger scatter at lower metallicities. Our $\log g$ measurements agree to within a few dex for $\log g \gtrsim 1$, but with a slight trend for XP to overestimate $\log g$ relative to APOGEE. Temperatures agree well for $4,000\,{\rm K} \lesssim T_{\rm eff} \lesssim 6,000\,{\rm K}$, but show significant trends at higher temperatures. These trends are due to differences in LAMOST and APOGEE $T_{\rm eff}$ estimates. As shown in Figs. 2.10 and 2.11 our temperature estimates match those of LAMOST, though with large uncertainties for stars hotter than $T_{\rm eff} \approx 7,500\,{\rm K}$.

We also compare our results with XGBoost (Andrae et al., 2023), in which a discriminative model is applied to XP spectra to estimate stellar parameters, using APOGEE labels as training data. We only compare the parameters of stars in which both catalogs have confidence. In our catalog, we select the stars passing the basic cut and with confidence > 0.5 in the parameter under consideration (see Section 2.4.4). In the XGBoost catalog, we select stars with parallax $\varpi > 1\,{\rm mas}$, because the estimation of $[{\rm M/H}]$ in XGBoost is less reliable for small parallaxes, according to Fig. 11 in Andrae et al. (2023). Moreover, for stars with $\varpi < 1\,{\rm mas}$ in the XGBoost catalog, we find a long-range distribution in $\log g$ for Red Clump stars, which we believe is because of the rapid degradation of discriminative models in low signal-to-noise regime. Fig. 2.17 shows our comparison with XGBoost. Our estimates of $[{\rm Fe/H}]$ and $\log g$ generally agree well, with a typical scatter of 0.2 dex in most regimes. However, we find a systematic trend in the $[{\rm Fe/H}]$ residuals for $[{\rm Fe/H}] \lesssim -1.3$, and systematic offsets in the $\log g$ estimates of order 0.2 dex. We additionally find a sharp feature in the $\log g$ residuals near $\log g \approx 2.6$. We note

that we find very similar features in our comparison with APOGEE, leading us to conclude that these systematic differences are primarily due to differences between the underlying training sets, LAMOST and APOGEE.

Our estimated parallaxes are constrained by both GDR3 parallax measurements and the observed stellar spectral energy densities, which are proportional to $1/\varpi_{\rm est}^2$. We therefore expect our estimated parallaxes to be closer to the true parallaxes, particularly in the regime in which GDR3 provides only weak constraints on parallax. Globular clusters provide a means of testing the validity of our parallax estimates, as many globular clusters lie at well known distances, and as relatively pure samples of cluster members can be straightforwardly obtained using simple cuts on sky location and proper motion. Our estimated parallaxes of the stars in a globular cluster should be more tightly concentrated around the central parallax, compared with the values observed by Gaia, particularly for stars with large GDR3 parallax uncertainties. At the same time, due to crowding, globular clusters are likely to have more contaminated XP spectra (particularly in the low-SNR regime that we are interested in), and thus present a relatively difficult test case for our method. In Fig. 2.19, we show the three most populous globular clusters within 7 kpc: NGC 5139 (ω Cen), NGC 104 (47 Tuc) and NGC 6752 (Vasiliev & Baumgardt, 2021). We restrict our comparison to stars with reliable parameter estimates, as identified by our basic reliability cut (Eq. 2.25) and good confidence (> 0.5) on ($T_{\rm eff}$, [Fe/H], $\log g$), as described in Section 2.4.4. Because we wish to investigate the low-parallax-SNR regime, where XP spectra contribute proportionally more information about distance, we select stars with $\sigma(\varpi_{\rm obs}) > 0.2\varpi_{\rm gc}$, where $\varpi_{\rm gc}$ is the central parallax value of the given globular cluster, as reported by Vasiliev & Baumgardt (2021). As can be seen in the right panels of Fig. 2.19, our estimated parallaxes are slightly more concentrated around the central values in all three globular clusters.

We validate our stellar reddening estimates by comparison with the Schlegel et al.

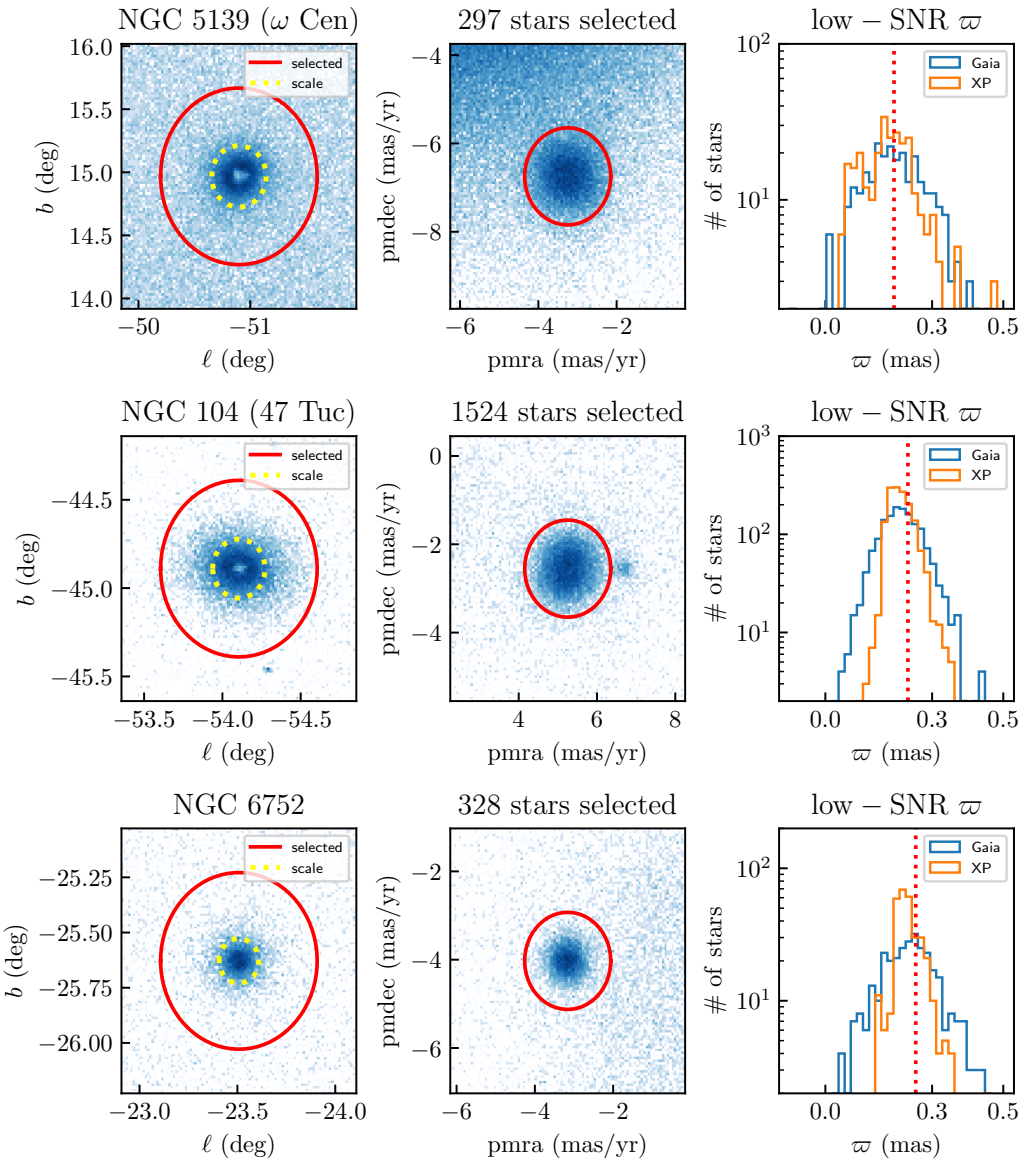


Figure 2.19. Validation of our parallax estimates in the low-SNR regime, using globular clusters of known distance as a test-bed. We use the three richest globular clusters, NGC 5139 (top row), NGC 104 (middle row) and NGC 6752 (bottom row) within 7 kpc. We identify stars in a globular cluster based on their positions in the sky (left panels) and in proper-motion space (middle panels). The solid red circles are the outer radius of our selection. In order to lessen the impact of crowding, which we expect to impact data quality, we exclude stars that are less than one "scale separation" (determined by Vasiliev & Baumgardt 2021, and represented here by dotten yellow circles) from the center of the given cluster. We further require that stars have reliable parameter estimates (as defined by Eq. 2.25), and good confidence (> 0.5) on $(T_{\rm eff}, [{\rm Fe/H}], \log g)$, which are estimated in Section 2.4.4. In the right panels, we show the distribution of the parallaxes, both as estimated by Gaia and as estimated with our model (using XP spectra). We conduct this comparison only for the low-SNR regime, defined as $\sigma(\varpi_{\rm obs}) > 0.2\varpi_{\rm gc}$, where $\varpi_{\rm gc}$ is the central parallax value of the given globular cluster, as reported by Vasiliev & Baumgardt (2021). In all three cases, our parallax estimates are more tightly concentrated about the central parallaxes (marked by red, dotted, vertical lines) of the respective globular clusters.

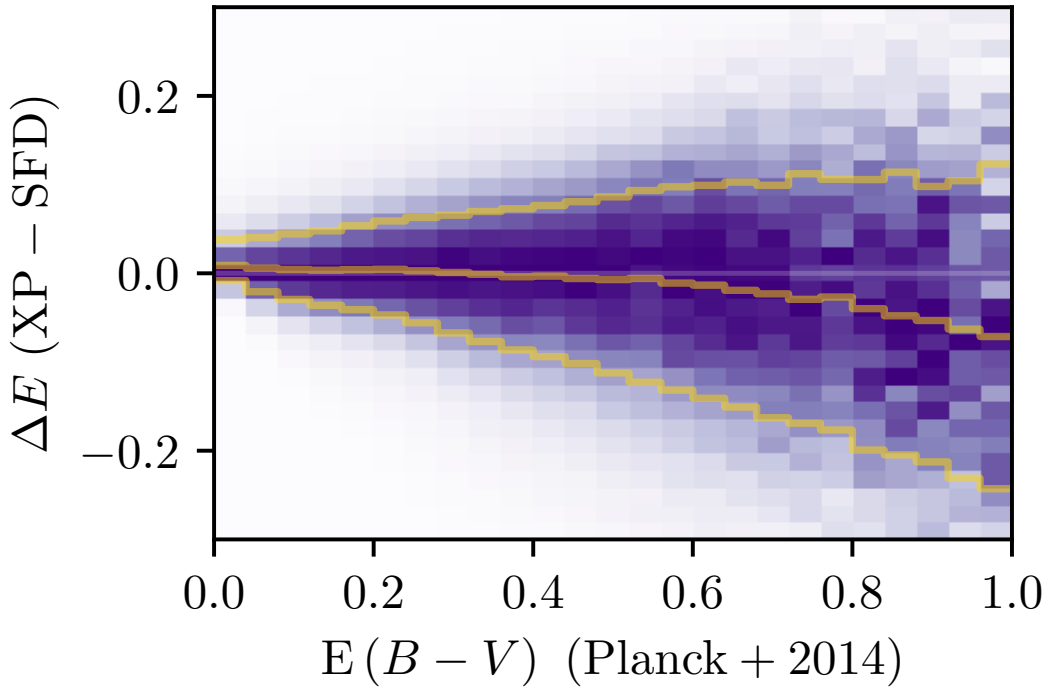


Figure 2.20. Comparison of our reddening estimates with SFD (Schlegel et al., 1998a), for stars that lie outside of the Galactic midplane ($|z| > 400$ pc, $|b| > 10$ deg) and outside the vicinity of the LMC and SMC, as a function of an independent measure of reddening, based on the Planck mission (Planck Collaboration et al., 2014a). The yellow envelopes mark the 16$^{\text{th}}$, 50$^{\text{th}}$, and 84$^{\text{th}}$ percentiles of the reddening residuals. Our stellar reddening estimates agree well with SFD, with a median residual of approximately zero out to E $(B-V) \approx 0.5$, and a scatter of 15-20%.

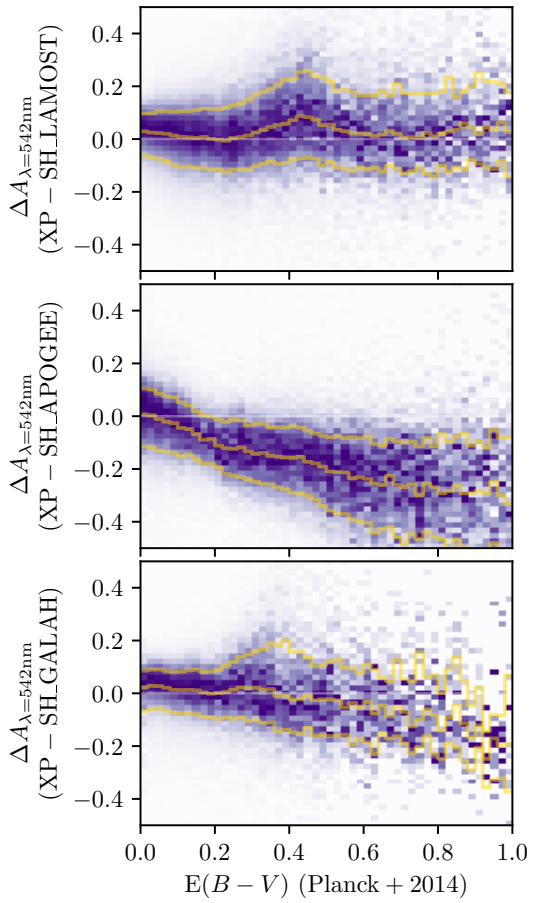


Figure 2.21. Comparison of our reddening estimates with a recent `StarHorse` catalog (Queiroz et al., 2023). The layout is the same as in Fig 2.20. We compare the extinction estimates of stars in which both catalogs have confidence. In particular, we require that stars which pass the basic cut (see Section 2.4.4) in our catalog, and that they have input flags of "panstarrs", "2mass" and "PARALLAX" but no output flags in the `StarHorse` catalogs. `StarHorse` makes use of different spectroscopic catalogs. We find good consistency between our catalog and the `StarHorse` catalog using LAMOST LRS (top panel) and GALAH (bottom panel). However, there is a systematic trend when we compare our extinctions with those of `StarHorse` using APOGEE. We find a similar trend when we compares `StarHorse` estimates using LAMOST and APOGEE against one another. We believe that this is likely due to systematic differences in the T_{eff} estimates given by LAMOST LRS and APOGEE.

(1998a, hereafter, "SFD") dust map. Fig. 2.20 shows the difference between our stellar reddening estimates (our parameter E) and the E $(B-V)$ estimate from SFD, as a function of reddening, as measured independently by the Planck mission, using dust optical depth at 353 GHz (Planck Collaboration et al., 2014a). As SFD only measures integrated reddening along the entire sightline, we restrict our comparison to stars which we expect to lie behind all of (or nearly all of) the dust. We therefore use stars that are more than 400 pc above or below the Galactic midplane (using our updated parallax estimate to calculate distance). We additionally exclude stars with $|b| < 10$ deg, which lie within 8 deg of the Large Magellanic Cloud or 6 deg of the Small Magellanic Cloud, or which fail our basic reliability cut (Eq. 2.25). Our stellar reddening estimates agree well with the reddening estimates provided by SFD, with only a slight trend in the median residuals past E $(B-V) \gtrsim 0.5$, and a scatter of 15-20%.

We also compare our extinction estimates with those given by the `StarHorse` catalogs (Queiroz et al., 2023). `StarHorse` infers stellar parameters from photometric, spectroscopic and astrometric data, using *ab initio* stellar models. In Fig 2.21, we compare our extinction estimates (at $\lambda = 542$ nm, denoted as "XP") to those of `StarHorse`. `StarHorse` makes use of several different catalogs of spectroscopically determined atmosopheric parameters. We treat `StarHorse` estimates based on different spectroscopic surveys (LAMOST LRS, APOGEE and GALAH) separately (denoting the `StarHorse` estimates as `SH_survey`). We find good agreement between our extinction estimates and those of `SH_LAMOST` (top panel) and `SH_GALAH` (bottom panel). However, there is a trend in the extinction residuals with `SH_APOGEE`. We find a similar trend when comparing `SH_APOGEE` with `SH_LAMOST-LRS`. These differences may be related to differences between the temperature estimates produced by APOGEE and LAMOST LRS.

2.5.4 Preliminary map of extinction in 3D

Our final catalog of stellar parameters contains estimates of extinction (E) and parallax (ϖ) of all 220 million GDR3 XP sources, covering the entire sky. This is precisely the information required to produce a large-scale three-dimensional dust map. Compared to broad-band photometry, the XP spectra allow a much more precise determination of stellar extinction. While higher-resolution spectra would enable even more precise extinction estimates (by pinning down $T_{\rm eff}$, which is highly covariant with extinction), high-resolution spectral surveys do not presently observe enough stars to densely cover the sky and enable high-resolution, all-sky dust maps. With 220 million sources, the GDR3 XP catalog provides a unique combination of high sky density and precision extinction estimates. We therefore believe that *Gaia* XP spectra will enable the next generation of three-dimensional Milky Way extinction maps. Here, we demonstrate a preliminary extinction map, based on naive spatial binning of stellar extinction estimates. We leave fuller exploitation of our stellar extinction estimates to future work.

Figs. 2.22 and 2.23 show preliminary sky maps of differential extinction in a number of distance ranges. In each distance bin, we calculate a HEALPix map of the inverse-variance-weighted mean extinction of stars falling in the bin (based on the inferred distance, $1/\varpi$, reddening E and reddening uncertainty $1/\sigma_E$ of each star). In each distance bin, we calculate mean extinction maps at HEALPix resolutions of $\texttt{nside} = 256$, 128 and 64 (equivalent to angular resolutions of $13.74'$, $27.48'$ and $54.97'$, respectively), throwing out pixels that are based on fewer than 10 stars. We then combine these maps, using the highest-resolution map that covers any given part of the sky. After calculating a map of cumulative extinction in each distance bin, we calculate differential extinction by subtracting off the previous distance bin. It is this differential extinction that we display in Fig. 2.22. We build this naive extinction map with the highest-quality stars: We remove stars with $\mathcal{L}_{\rm inference}/{\rm DOF} > 5$. We

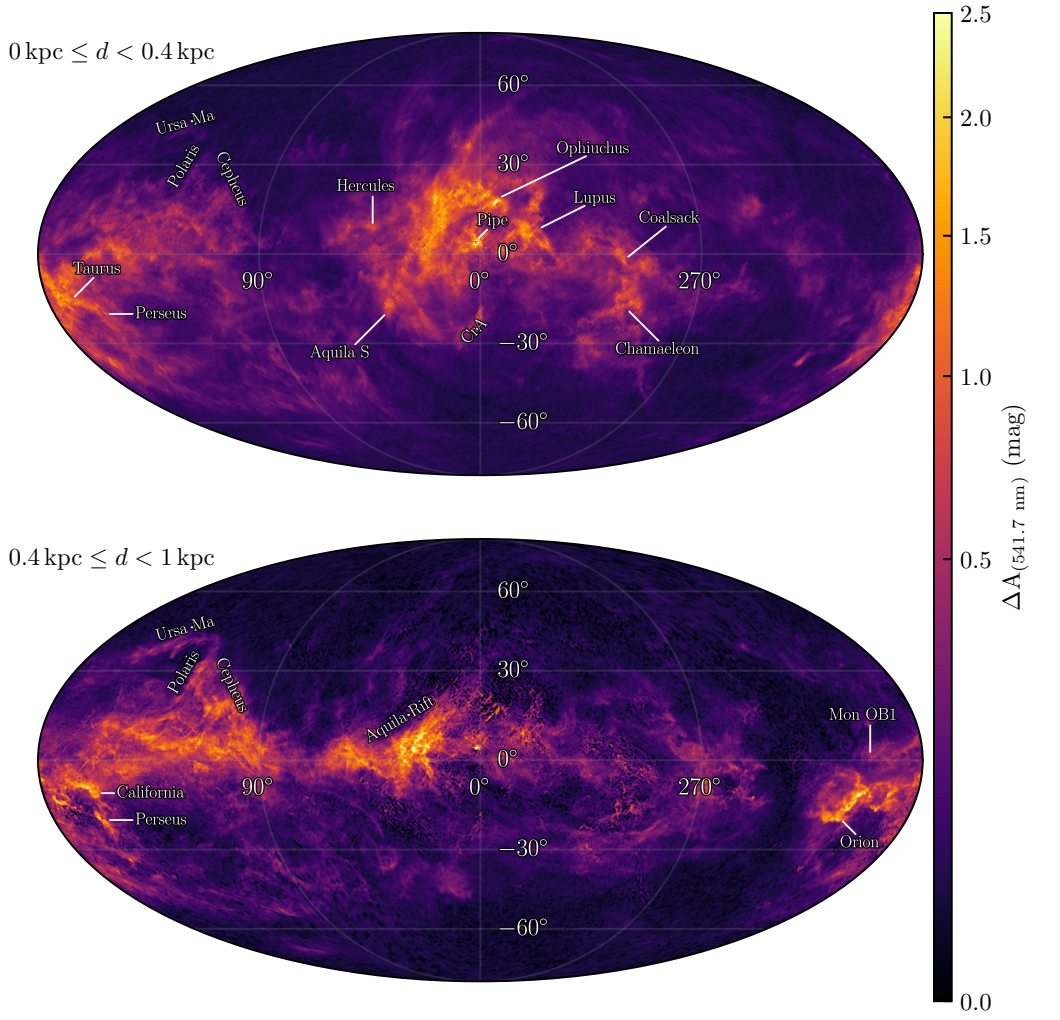


Figure 2.22. Sky maps of differential extinction in different distance ranges. We construct adaptive-resolution HEALPix maps of inverse-variance-weighted mean stellar extinction (using our inferred E and σ_E) in each distance range, and subtract off the result for the previous distance range. We omit stars with poor fit quality ($\chi^2/\mathrm{DOF} > 5$) or uncertain inferred parallaxes (defined as $\sigma_\varpi/\varpi > 0.1$). We also require that $E < 10$, $\sigma_E < 0.04$, to remove outliers in extinction. This simple method of determining differential extinction is only intended to display the information present in our stellar reddening and distance inferences, and is not meant to replace (or reproduce) more sophisticated three-dimensional dust mapping methods. However, even this naive dust-mapping method recovers rich information on the three-dimensional structure of the interstellar medium, demonstrating the power of XP spectra to map dust. We recover the now-familiar structure of local dust clouds, several of which are labeled above.

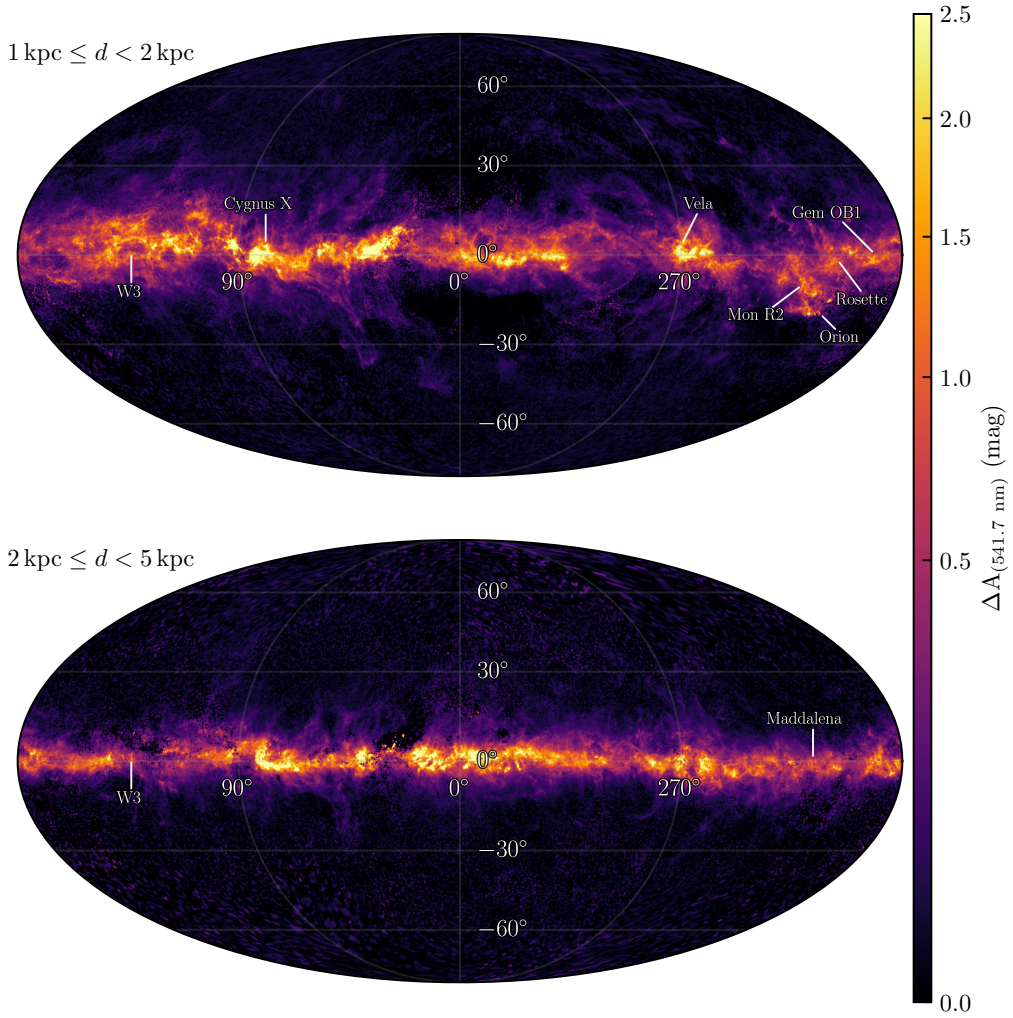


Figure 2.23. Sky maps of differential extinction. As in Fig 2.22, but for more distant slices.

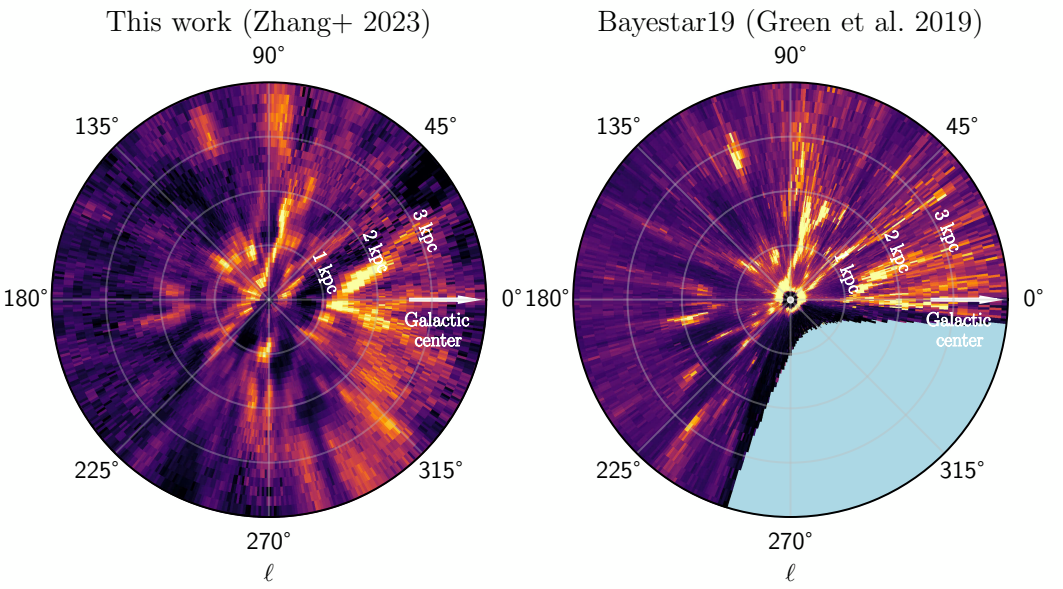


Figure 2.24. A bird's-eye view of the extinction density across the Galactic plane in our preliminary dust map (left) in Bayestar19 (right; Green et al., 2019). Both panels display average extinction density, integrated along the z-axis from $z = -400\,\mathrm{pc}$ to $+400\,\mathrm{pc}$). The Sun is located at $(0, 0, 0)$, with the Galactic Center off the plot to the right. Positions where Bayestar19 does not cover are marked as light blue. Using *Gaia* XP sources alone, we are able to identify most of the dust structures in Bayestar19, which makes use of photometric extinction estimates for four times the number of stars. Because *Gaia* surveys the entire sky, we also cover the Southern Galactic plane, which is not included in Bayestar19, due to the latter's reliance on Pan-STARRS 1, a photometric survey of the Northern Hemisphere. Even with extremely simple mapping techniques based on naive spatial binning, our XP-spectral-based extinction and distance estimates reveal a detailed three-dimensional dust distribution. More sophisticated mapping techniques hold the potential to extract much more detailed information from our stellar extinction and distance estimates.

also remove stars with extremely large extinctions ($E > 10$), which are likely to be outliers, or with imprecisely determined extinctions ($\sigma_E > 0.04$). Finally we remove stars with large distance uncertainties ($\sigma_\varpi/\varpi > 0.1$, using our inferred parallax and its corresponding uncertainty). Although we do not impose spatial continuity on the dust, the rich structure of the ISM is already apparent, out to a distance of ~ 5 kpc. We mark the positions of several prominent dust clouds, including ρ Ophiuchus, Aquila South, Hercules, Cepheus, Perseus, California, Ursa Major, Polaris, Taurus, Orion, the Pipe Nebula, Lupus, Chamaeleon, the Coalsack, Monoceros OB1, Cygnus X, Gemini OB1, Maddalena, W3, Rosette and Monoceros R2.

In Fig. 2.24, we compare bird's-eye views of Galactic extinction density in our new catalog and in Bayestar19 (Green et al., 2019). In detail, both panels of Fig. 2.24 show

$$A_{z=400\,\mathrm{pc}}\,(x,y) \equiv \frac{1}{800\mathrm{pc}} \int_{-400\ \mathrm{pc}}^{400\ \mathrm{pc}} \frac{\mathrm{d}E(x,y,z)}{\mathrm{d}r}\,\mathrm{d}z, \tag{2.29}$$

where $r = (x^2 + y^2 + z^2)^{1/2}$ is distance from the origin in Sun-centered Galactic Cartesian coordinates, and $\mathrm{d}E/\mathrm{d}r$ is the extinction density (with units of mag pc^{-1}). In order to calculate this integral, we need to construct an approximation of $\mathrm{d}E/\mathrm{d}r$ from our stellar distance and extinction estimates. To do so, we construct multi-resolution HEALPix maps of differential extinction in the same manner as described above for our sky maps of extinction, but with much finer distance bins. Since this is only a naive average of stellar extinction (E), the precision of the naive dust map cannot exceed the uncertainties of distance measurements in GDR3, the lower limit of which increase linearly with distance. We use 40 bins to cover 0-5 kpc. The length of the n$^{\mathrm{th}}$ bin is $\sinh[n \cdot \mathrm{arcsinh}(5)/40] - \sinh[(n-1)\mathrm{arcsinh}(5)/40]$. We remove outlier stars with large uncertainties on E ($\sigma_E > 0.04$), and stars with large uncertainties on parallax ($\sigma_\varpi/\varpi > 0.33$). We further require that the stars have $\mathcal{L}_{\mathrm{inference}}/\mathrm{DOF} < 5$,

to remove poorly fitting results. We thus obtain a rough three-dimensional extinction map, $E(\ell, b, r)$. The differential extinction is then approximated by taking the difference in cumulative extinction between successive distance bins. We integrate the differential extinction through the Galactic plane vertically, using Eq. 2.29, to obtain our bird's-eye view of Galactic extinction.

While Bayestar19 makes use of a much larger stellar catalog than we use here, its stellar parameters are determined using broad-band photometry, rather than spectra. As can be seen in Figs. 2.22 and 2.24, using *Gaia* XP sources alone, we are able to identify most of the prominent dust structures in Bayestar19, despite the difference in catalog size. Our stellar distance and extinction estimates build the foundation of a next-generation dust map.

2.5.5 Flux residuals

Fig. 2.25 and Fig. 2.26 maps average flux residuals at 402 nm and 652 nm, and in 2MASS J and unWISE $W1$ bands, normalized by the uncertainty of observations: $\chi \equiv (f_{\rm pred} - f_{\rm obs})/\sigma(f_{\rm obs})$. *Gaia* scanning-law patterns are faintly visible in the average XP spectral flux residuals, while the 2MASS observing pattern is visible in the J-band residual map. However, the most prominent residuals are in the WISE bands, and clearly trace dust density. These patterns have two features which strongly suggest that they arise from unmodeled variations in the dust extinction curve. First, while the residuals clearly follow dust density, different clouds have residuals of opposite sign. For example, our predicted $W1$ fluxes are higher than observed fluxes in the Aquila Rift, Cepheus, and the Perseus-Taurus-Auriga complex, but are lower than observed fluxes in the ρ Ophiuchi cloud complex. Such sign differences would arise if these two sets of clouds were to diverge from the mean R_V in different directions. Second, the same residual patterns are visible to a lesser extent at shorter wavelengths (*i.e.*, at 652 nm), but with the opposite sign. Changing

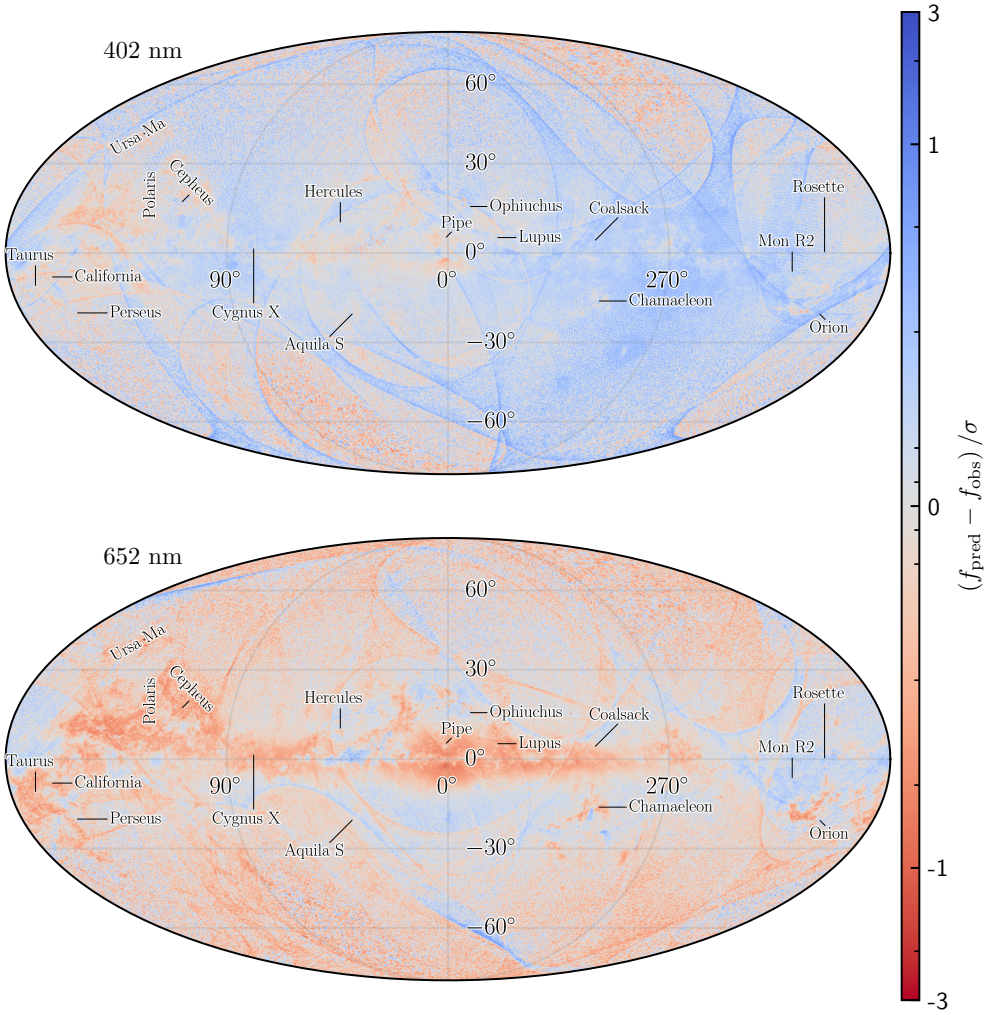


Figure 2.25. Maps of average model – data flux residuals, normalized by observed flux uncertainties. In detail, we plot $\chi = (f_{\rm pred} - f_{\rm obs})/\sigma(f_{\rm obs})$ at 402 nm and 652 nm, and in 2MASS J and WISE $W1$ band (see Fig 2.26). We exclude the stars with bad parallax agreement ($|\varpi_{\rm pred} - \varpi_{\rm obs}|/\sigma_\varpi > 10$), poor fits ($\chi^2/{\rm DOF} > 5$), or which are not well represented in the training set (`ln_prior` < -7.43). In each panel, we also remove stars with extreme residuals ($|\chi| > 10$) – which are rare – in the wavelength or band in question. With the exception of the WISE bands, average flux residuals are far smaller than typical flux uncertainties. However, several spatial patterns are visible in the residuals. *Gaia* scanning-law patterns, which appear as arcs across the sky, can be seen at high Galactic latitudes in the 402 and 652 nm residual maps. There are also faint dust-related residuals in the 652 nm map, visible both in the inner Galactic plane and in the vicinity of several well-known clouds (such as Orion, Cepheus, California and Taurus).

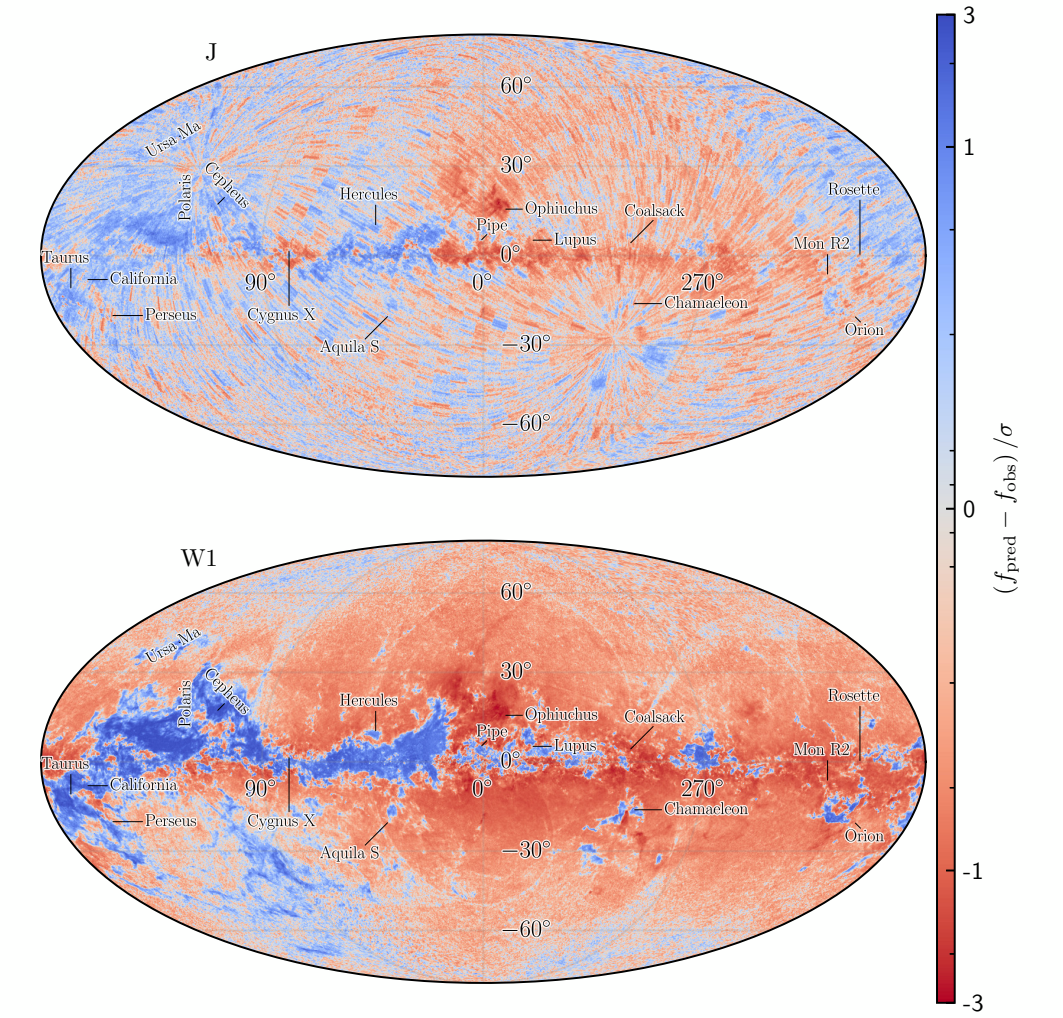


Figure 2.26. Maps of average model – data flux residuals, normalized by observed flux uncertainties. As in Fig 2.25, but for 2MASS J and unWISE $W1$ bands. 2MASS exposure tiling patterns, which are aligned with equatorial coordinates, are visible in the J-band residuals. There are clear dust-related residuals in $W1$. These dust-related patterns tend to have opposite signs at short vs. long wavelengths (*e.g.*, compare $W1$ here with 652 nm in Fig. 2.25), and have different signs in different clouds, suggesting that they are caused by variations in the extinction curve.

R_V alters the ratio of extinction at short and long wavelengths, and thus should lead to flux residuals of opposite sign on different ends of the observed spectrum. Given these factors (flux residuals that are proportional to dust density, but with different signs in different clouds; opposite flux residuals at short and long wavelengths), we hypothesize that these patterns are caused by extinction curve variations.

While it may appear counter-intuitive that dust-related residuals would appear most strongly at the longest wavelengths, this effect can arise if most of the likelihood constraints are from observations at shorter wavelengths, as in our case, where 61 of 66 observed wavelengths are in the optical range. Our model can compensate for unmodeled variations in the extinction curve by adjusting both E and ϖ in a way that nearly matches the observed fluxes at optical wavelengths, at the cost of introducing dust-dependent flux residuals in the NIR. In order to correctly model the spectrum across the entire observed spectrum, it is necessary to incorporate extinction curve variations into our model. We leave this to future work.

2.6 Discussion

Our model makes a number of assumptions:

1. A star's intrinsic spectrum is a function of just three parameters: $T_{\rm eff}$, [Fe/H], and $\log g$. At the coarse resolution of XP spectra, other parameters, such as $[\alpha/\mathrm{Fe}]$, rotation, or detailed chemical abundances can be ignored.

2. The extinction curve is universal. That is, the extinction of any given star i is a scalar multiple, E_i, of a universal function of wavelength, $R(\lambda)$.

3. All observed spectra are of single stars. Our model does not account for binary systems.

4. All observed spectra are stars in the range of stellar types covered by our

model. Our model roughly covers the range $4000 \lesssim T_{\text{eff}} \lesssim 10000$, and does not include subdwarfs and white dwarfs.

We will discuss each of these assumptions in turn.

First, we model intrinsic stellar spectra as a function of T_{eff}, $[\text{Fe}/\text{H}]$, and $\log g$. However, in our formalism, it is simple to include additional intrinsic stellar parameters, as long as training data is available in which these parameters have been measured. Given its use in measuring star formation history and the history of metal enrichment in the interstellar medium, $[\alpha/\text{Fe}]$ would be a logical parameter to include in our model. However, studies using synthetic XP spectra suggest that for most stellar types, it is not possible to constrain $[\alpha/\text{Fe}]$ precisely enough to be of interest (Gavel et al., 2021; Witten et al., 2022). Nevertheless, inclusion of $[\alpha/\text{Fe}]$ as a fourth intrinsic stellar parameter would be a relatively straightforward addition to our model, requiring only a trivial change to the structure of our model (i.e., the addition of one dimension to Θ; see Fig. 2.3).

Second, this work is based on the assumption that the extinction curve is "universal" for all stars. That is, the extinction of any given star is proportional to a single, universal function of wavelength: $A(\lambda) \propto R(\lambda)$. We represent this function as a common 66-dimensional vector ($\vec{R}$), with each component representing a different wavelength. However, the Milky Way extinction curve is known to have non-negligible variation (*e.g.*, Fitzpatrick & Massa 1986, 1988, 1990, 2005, 2007, 2009). As shown in Fig. 2.15, the ratio of extinction at 900 and 400 nm ($R(\lambda \simeq 900\text{nm})/R(\lambda \simeq 400\text{nm})$) can vary by as much as a factor of ~ 2. As discussed in Section 2.5.5 (see Fig. 2.25), we see evidence of extinction-curve variation in sky maps of flux residuals.

In the optical-through-near-infrared wavelengths, variation in the extinction curve can be effectively parameterized by a scalar, $R_V \equiv A_V/E(B-V)$ (Cardelli et al., 1989). In the ultraviolet, which is not covered by our dataset, additional degrees of

freedom may be essential to parameterize variation in the strength of the 2175 Å bump and the slope of the far-UV rise in extinction (Peek & Schiminovich, 2013). Therefore, the most practical way forward may be to add an additional degree of freedom to our extinction model, which allows the direction of $\vec{R}$ to vary. We believe that the XP spectra, in combination with 2MASS and WISE photometry, contain enough information to "learn" the variation of the extinction curve with minimal priors about the exact form of the variation. One possible way of implementing extinction curve variation in our model would be to expand the "universal" $\vec{R}$ in Equation (2.7) to $\vec{R} + \xi\vec{R}_1$, where $\vec{R}_1$ is a vector orthogonal to $\vec{R}$, and ξ is a scalar that is determined separately for each star. We would then expect ξ to be related to R_V. This method is similar to that of Schlafly et al. (2016), which decomposes the reddening vector into a series of orthogonal basis vectors. Schlafly et al. (2016) finds two statistically significant components, one of which represents a "mean" reddening curve, and the second of which is related linearly to R_V (for small excursions from $R_V = 3.3$).

Although the specific physics behind R_V variation are still not fully understood (Draine, 2003), the slope of the extinction curve is thought to be related to the dust grain-size distribution, which determines whether Rayleigh or Mie scattering applies. Larger dust grains lead to flatter curves, corresponding to larger R_V. A precise determination of R_V as a function of position in the Milky Way would allow a study of the dependence of R_V on environmental conditions, and would be of importance for the study of the dust properties as well as the evolution of the Galaxy. R_V (or in our proposed model extension above, ξ) should vary with the physical properties of the dust, and should therefore be spatially continuous. This suggests the utility of using continuous function of position, $R_v(r, l, b)$, to model extinction curve variation. One way to implement this in our model would be to represent R_V (or, equivalently, ξ) as a basis-function expansion, with the expansion coefficients being learned during training.

A third assumption made by our model is that all sources are single stars. However, a significant fraction of stars reside in binary systems. For large luminosity ratios, the observed flux will be dominated by the brighter star in the pair, and our model should reasonably recover the parameters of the brighter star. However, in binary systems in which the luminosity ratio is near order unity, we expect our model to break down. In the case of an unresolved, equal-mass binary system, in which both stars formed from the same cloud and thus have the same age and metallicity, both stars will be of equal temperature. The system will thus appear as a single star of twice the luminosity of a single star of the same $(T_{\text{eff}}, [\text{Fe}/\text{H}], \log g)$. In attempting to fit this system as a single star, our model could decrease the inferred distance by a factor of $\sqrt{2}$ (or equivalently, increase inferred parallax by a factor of $\sqrt{2}$), obtaining a solution matching the observed spectrum. The model also has an additional parameter, $\log g$, which it can decrease in order to increase the predicted luminosity of the observed spectrum. During training, we address the issue of binary stars using self-cleaning, as described in Section 2.3.2. This self-cleaning method makes use of the residuals between our inferred values of $\log g$ and ϖ, and those determined by LAMOST and Gaia. However, independently measured stellar atmospheric parameters are unavailable for the entire XP catalog (and indeed, if they were, they would render the need for stellar parameters determined from XP spectra moot), so this self-cleaning method cannot be applied to our entire catalog of inferred stellar parameters. A possible way forward here is to explicitly model each XP source as a binary system. This would not require a full doubling of stellar parameters, as binary systems can be reasonably assumed to share the same $[\text{Fe}/\text{H}]$, and as stellar evolution places constraints on the region of $(T_{\text{eff}}, \log g)$-space that the two stars can simultaneously inhabit. Comparison of the goodness-of-fit (e.g., reduced χ^2) of single-star and binary solutions for each source would then allow identification of possible binary systems.

Finally, while our method is formulated in a way that is agnostic towards the origin of the stellar parameters used during training, our stellar model can only be as good as the higher-resolution spectroscopic data used to train it. In this work, we have used LAMOST spectra, because of LAMOST's good coverage of the Hertzsprung-Russell Diagram, and of the main sequence, in particular. However, there are two drawbacks to this choice. First, LAMOST does not cover the giant branch as fully as APOGEE. Second, as discussed in Section 2.2.2, we use a combination stellar parameters from the standard LAMOST catalog and from the "Hot Payne", due to the former's difficulty in modeling hot stars. This leads to a discontinuity in our training data at $T_{\rm eff} \approx 7000\,{\rm K}$.

In our future work, we plan to combine more spectroscopic surveys for better coverage of stellar parameter space. Possible options may include LAMOST MRS, APOGEE and GALAH. However, methods must be developed to explain and reconcile the systematic difference between catalogs (such as the discontinuity seen at $\sim$7000 K in our training dataset), which are caused by their use of different reduction pipelines and spectroscopic instruments. Another particularly promising dataset for future work is the SDSS-V Milky Way Mapper (MWM, Kollmeier et al. 2017), which is gathering high-resolution ($R \sim 22,000$) NIR and medium-resolution ($R \sim 2,000$) optical spectra of over 6 million stars. MWM targets are primarily at low Galactic latitudes (*i.e.*, in the Milky Way disk). Compared to LAMOST, MWM will contain a relatively large number of stars at high extinction, making it a powerful dataset for studying variations in the extinction curve.

2.7 Conclusions

In this paper, we determine stellar parameters ($T_{\rm eff}$, [Fe/H], $\log g$, E, ϖ) for all 220 million *Gaia* XP sources.

These stellar parameter determinations are based on an empirical forward model of *Gaia* XP spectra and infrared photometric bands J, H, K_s from 2MASS and $W1$, $W2$ from unWISE. This model maps stellar atmospheric parameters ($T_{\rm eff}$, $[{\rm Fe/H}]$, $\log g$), extinction (E) and parallax (ϖ) to predicted XP spectra and infrared photometry. The model is trained using the subset of the XP sources which have higher-resolution spectroscopy from LAMOST, and which thus have well determined atmospheric parameters. We additionally use extinctions from Bayestar19 and parallaxes from GDR3. The matched XP–LAMOST data is separated into training (80%) and validation (20%) sets. We then simultaneously learn a stellar model and refine the parameters of the stars in the training set. Our model generally recovers the flux of the stars in the validation set within 3σ in the regions of the parameter space where we have good data coverage. We implement our stellar model in an auto-differentiable framework, TensorFlow 2, which allows us to efficiently infer parameters of observed stars and propagate observational uncertainties into uncertainties in stellar parameters.

We apply this model to all 220 million published GDR3 XP spectra, 99% of which do not have corresponding LAMOST spectra. When inferring the stellar parameters of all 220 million stars, we impose a weak prior on the stellar atmospheric parameters, based on the distribution of parameters in our training data. This is intended to prevent the optimizer from reaching regions of parameter space not covered by the training set. As our model is auto-differentiable, we determine the stellar parameters using simple gradient descent methods, and estimate the corresponding uncertainties using Fisher information matrices.

Our entire catalog of stellar parameters, along with our trained stellar model and dust extinction curve, are available for download at `https://doi.org/10.5281/zenodo.7692680`, and can also be queried using ADQL/TAP from the German Astrophysical Virtual Observatory (GAVO; for details, see `https://dc.zah.`

uni-heidelberg.de/tableinfo/xpparams.main). In order to obtain reliable stellar parameter estimates, we strongly urge users to apply our "basic reliability" cut: quality_flags < 8. We include additional "confidence" flags for each stellar atmospheric parameter, which provide even stricter cuts on the quality of our parameter inferences.

Alongside the stellar atmospheric parameters, we obtain a large catalog of precisely determined stellar distances and extinctions, which will be the foundation of a next-generation 3D dust map in the Milky Way. We additionally find evidence that the *Gaia* XP spectra contain information on variation in the dust extinction curve. We therefore expect "the next generation" of dust maps to benefit now only from the increased precision allowed by *Gaia* XP spectra, but also to include variation of the extinction curve. We leave this avenue of investigation to follow-up work.

Finally, the stellar parameters presented in this paper are a promising resource for studying stellar populations in the Milky Way. Based on our validation dataset, we achieve typical errors of 90 K in $T_{\rm eff}$, 0.15 dex in $[{\rm Fe/H}]$ and $\log g$, and 0.03 mag in E (a parameter that is roughly equivalent to E $(B-V)$).

2.8 Data availability

All underlying data used in this work is in the public domain, as detailed in Section 2.2. Our results can be acquired at https://doi.org/10.5281/zenodo.7692680, which contains all contents in Table 2.2.

2.9 Gaia Archive Queries

In order to crossmatch 2MASS to GDR3 sources with XP spectra, we run the following ADQL query on the Gaia Archive:

```
SELECT
  g.random_index, g.source_id,
  g.ra, g.dec,

  tmass.designation,
  tmass.ra as tm_ra, tmass.dec as tm_dec,
  tmass.j_m, tmass.j_msigcom,
  tmass.h_m, tmass.h_msigcom,
  tmass.ks_m, tmass.ks_msigcom,
  tmass.ph_qual,

  xmatch.angular_distance
    as gaia_tmass_angular_distance

FROM gaiadr3.gaia_source AS g
JOIN gaiadr3.tmass_psc_xsc_best_neighbour
AS xmatch
  USING (source_id)
JOIN gaiadr3.tmass_psc_xsc_join AS xjoin
  USING (clean_tmass_psc_xsc_oid)
JOIN gaiadr1.tmass_original_valid AS tmass
  ON xjoin.original_psc_source_id
     = tmass.designation

WHERE
  (g.has_xp_continuous = 'true')
  AND (
       (tmass.ph_qual LIKE 'A__')
    OR (tmass.ph_qual LIKE '_A_')
    OR (tmass.ph_qual LIKE '__A')
  )
  AND (tmass.ext_key is NULL)
  AND (g.source_id BETWEEN sid0 AND sid1)
```

```
34
35 ORDER BY g.source_id
```

Above, `sid0` and `sid1` are integers used to divide our query into manageable chunks.

We use the following query to fetch information from the GDR3 `gaia_source` catalog and from the astrometric fidelity catalog (Rybizki et al., 2022) on each source with an XP spectrum:

```
1  SELECT
2    g.random_index, g.source_id, g.ref_epoch,
3
4    g.ra, g.ra_error, g.dec, g.dec_error,
5    g.parallax, g.parallax_error,
6    g.pmra, g.pmra_error, g.pmdec, g.pmdec_error
7
8    g.ruwe, g.astrometric_excess_noise,
9
10   spur.fidelity_v2, spur.norm_dg,
11
12   g.phot_g_mean_mag,
13   g.phot_g_mean_flux, g.phot_g_mean_flux_error,
14   g.phot_bp_mean_mag,
15   g.phot_bp_mean_flux, g.phot_rp_mean_flux_error,
16   g.phot_rp_mean_mag,
17   g.phot_rp_mean_flux, g.phot_bp_mean_flux_error,
18
19   g.has_xp_continuous,
20   g.phot_bp_n_obs, g.phot_rp_n_obs,
21   g.phot_bp_rp_excess_factor,
22   g.visibility_periods_used
23
24 FROM gaiadr3.gaia_source as g
25   JOIN gaiadr3.astrophysical_parameters as p
```

```
26      ON g.source_id = p.source_id
27    LEFT OUTER JOIN external.gaiaedr3_spurious as spur
28      ON g.source_id = spur.source_id
29
30  WHERE
31    (g.has_xp_continuous = 'True')
32    AND (g.source_id BETWEEN sid0 AND sid1)
33
34  ORDER BY g.source_id
```

2.10 Stellar atmospheric parameter confidence estimates

As described in Section 2.4.4, we train a classifier for each stellar atmospheric parameter ($T_{\rm eff}$, [Fe/H], $\log g$), which assigns a "confidence" between 0 and 1 to each parameter estimate. Here, we list the input features used by the classifiers, and describe the neural-network structure of the classifiers, both of which are uniform across all three classifiers.

We use the following features:

- `ln_rchi2_opt`
- `ln_dplx2`
- `ln_prior`
- `teff_est`
- `logg_est`
- `feh_est`
- `teff_est_err`
- `logg_est_err`
- `feh_est_err`

- `asinh_plx_snr`
- `asinh_g_snr`
- `asinh_bp_snr`
- `asinh_rp_snr`
- `ln_phot_bp_rp_excess_factor`
- `phot_g_mean_mag`
- `ln_ruwe`
- `fidelity_v2`
- `norm_dg`
- `ln_bp_chi_squared`
- `ln_rp_chi_squared`

Features of the form `asinh_X_snr` are computed using $\text{asinh}\left(X/\sigma_X\right)$, using the GDR3 parallax and flux measurements. Features of the form `ln_X` are calculated as $\ln\left[\max\left(X, 10^{-7}\right)\right]$. The `ln_dplx2` feature is computed from $\left|\varpi_{\text{GDR3}} - \varpi_{\text{est}}\right| / \sigma_{\varpi,\text{GDR3}}$. The fields `fidelity_v2` and `norm_dg` are indicators of the quality of the GDR3 astrometric measurements and of crowding, respectively (Rybizki et al., 2022). Features of the form `X_est` and `X_est_err` are based on our stellar parameter estimates and their associated uncertainties. In addition to the above features, we additionally provide the network with information about the flux residuals between our best-fit models and the observed XP spectra. In detail, we provide the network with $\text{asinh}\left[\left(f_{\text{pred}} - f_{\text{obs}}\right)/\sigma_f\right]$ at each sampled wavelength.

We concatenate all of our input features into a single vector for each source, and feed it into a feed-forward neural network with three densely-connected hidden layers, each with 64 neurons and ReLU activation, and a single output neuron with a sigmoid activation. In order to regularize the network during training, we use a dropout fraction of 0.1 before each hidden layer (Srivastava et al., 2014), and impose an L2 penalty of 10^{-3} on the weights of each hidden layer. The output neuron

represents the probability of the given stellar atmospheric parameter estimate being reliable. We use the binary cross-entropy between our output neuron and the training label as our loss function. We train each classifier using the Adam optimizer (Kingma & Ba, 2014), with a batch size of 4096, and an initial learning rate of 10^{-3}, which we decrease by a factor of 10 whenever the validation loss fails to decrease over a span of 256 training epochs. We terminate the training procedure when the validation loss fails to decrease over a span of 1024 epochs.

Chapter 3

A 3D map of extinction-curve variations

Disclaimer: This chapter has been published by Science as Zhang & Green (2025).

3.1 Abstract

Interstellar dust grains cause extinction (absorption and scattering) of light from background astronomical sources. The spectral shape of the extinction curve depends on the dust composition. We used low-resolution optical spectra to measure the extinction curve of 130 million stars. By inverting these data, we mapped the extinction curve parameter $R(V)$ within the Milky Way in three dimensions and within the Magellanic Clouds in two dimensions. These maps provide improved extinction corrections for astronomical observations. We find that $R(V)$ varies with extinction, consistent with dust grains growing by accretion in low-extinction regions and by coagulation in higher-extinction regions. Star-forming regions have high $R(V)$ values, indicating either preferential destruction of small dust grains or additional supply of large dust grains in those regions.

3.2 Adopted datasets

We investigated how $R(V)$ varies using archival low-resolution, flux-calibrated blue photometer and red photometer (BP and RP, collectively referred to as XP) spectra collected by the Gaia space telescope (De Angeli et al., 2023; Gaia Collaboration et al., 2016a, 2023b; Montegriffo et al., 2023). The XP spectra provide flux measurements at 61 wavelengths, many more than broadband photometric surveys, which is sufficient spectral resolution to determine the shape of the extinction curve (Andrae et al., 2023; Li et al., 2024; Zhang et al., 2023b). Most stars with XP spectra also have Gaia parallax measurements, which can be combined with photometry to determine the distance to the star. To supplement the XP spectra, we adopted additional photometry from NIR (Skrutskie et al., 2006) and mid-infrared (MIR) (Schlafly et al., 2019) surveys, extending our wavelength coverage out to $4.6 \sim \mu$m.

We used machine learning to determine both a stellar model and a dust extinction curve along each sightline from the Gaia data. The machine learning model was trained on a small subset ($\sim 1\%$) of the Gaia XP catalog that has matching higher-resolution spectroscopy from the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) surveys (Cui et al., 2012). The LAMOST data allowed the stellar parameters of the training set to be determined independently of the dust extinction parameters. The training set provides sufficient coverage for stars with effective temperatures of ~ 4000 to 8000 K.

3.3 Stellar model and quality cuts

We constructed a forward model that predicts the observed flux as a function of stellar and dust parameters. The intrinsic stellar spectrum is described by a neural network, which maps atmospheric parameters (effective temperature, metallicity, and surface gravity, which we collectively denote as θ) to the intrinsic flux (F_λ) at

each wavelength. The stellar flux is dimmed according to its distance, which we represent using stellar parallax (ϖ). In contrast to our previous work (Zhang et al., 2023b), we allowed the shape of the extinction curve, $R(\xi)$, to vary within a family of curves (Eq. 2) controlled by a single fitting parameter ξ, which is allowed to vary between stars. ξ functions similarly to $R(V)$ in previous dust models (Cardelli et al., 1989; Fitzpatrick, 1999; Gordon et al., 2023). The total amount of extinction is quantified by a fitting parameter E. We implemented the model in the machine learning programming package TensorFlow 2 (Abadi et al., 2016). Our forward model maps the stellar and dust parameters to the observed stellar spectra using the equation

$$\mathbf{f}_\lambda(\theta, \varpi, E, \xi) = \mathbf{F}_\lambda(\theta)\varpi^2 \mathrm{e}^{-E\mathbf{R}(\xi)}, \tag{3.1}$$

where the mapping from ξ to the extinction curve, $R(\xi)$, is given by

$$\ln \mathbf{R}(\xi) = \ln(\mathbf{R}_0) + \tanh(\xi)\Delta \ln(\mathbf{R}), \tag{3.2}$$

where $\mathbf{R}_0$ and $\Delta \ln(\mathbf{R})$ are model parameters shared by all stars that were learned while training the model. Applying the tanh function to ξ limits the size of its effect on the extinction curve.

We trained this model using 2.4 million stellar-type parameter estimates from LAMOST data release 8 (Cui et al., 2012; Zhao et al., 2012). We used a catalog of hot stars (Xiang et al., 2022) as priors on θ, reddening estimates from a three-dimensional (3D) dust map (Green et al., 2019) as priors on E, parallax measurements from Gaia data release 3 (Gaia Collaboration et al., 2023b) as priors on ϖ, and a unit normal distribution prior on ξ. For stars that are not in the prior reddening map (Green et al., 2019), we adopted values from a 2D dust map (Schlegel et al., 1998). We sampled the XP spectral flux at 10-nm increments from 392 to 992 nm, similar to previous work (Zhang et al., 2023b).

We applied our trained model to all 220 million Gaia XP spectra [augmented with the NIR and MIR photometry (Schlafly et al., 2019; Skrutskie et al., 2006) and Gaia parallaxes] to infer the parameters θ, ϖ, E, and ξ for each star. Because 99% of the sources with XP spectra do not have LAMOST observations, we imposed weak priors on the stellar parameters, on the basis of their distribution in our training set; this restricted the model to regions of stellar parameter space that were covered by the training set. We used the Fisher information matrix to propagate observational uncertainties into the stellar parameter estimates. We then converted the estimates of ξ into $R(V)$ for each star.

We then imposed quality cuts on the model output, to ensure that the extinction curve was sufficiently well constrained. We built three classifiers to set thresholds on the reliability of our estimates on effective temperature ($T_{\rm eff}$), metallicity ([Fe/H]), and surface gravity ($\log g$). For the 130 million stars that passed our quality cuts, we obtained median 1σ uncertainties of ± 130 K in $T_{\rm eff}$, ± 0.03 mag in E, and ± 0.18 in $R(V)$. The learned family of extinction curves is consistent with those derived using higher-resolution spectra. Our $R(V)$ measurements have a monotonic correspondence with previous work.

3.4 Mapping the distribution of $R(V)$

We generated 3D maps of $R(V)$ from our extinction estimates of 130 million stars. We first calculated sky maps of the average extinction at 440 and 550 nm (the approximate central wavelengths of B and V bands) at different distances. We then calculated the value of $R(V)$ in the 3D space (voxel) between adjacent distance bins and angular positions, using $R(V) = \Delta A(V)/[\Delta A(B) - \Delta A(V)]$, where $\Delta A(B)$ and $\Delta A(V)$ are the difference of extinction of adjacent distance bins in B and V bands.

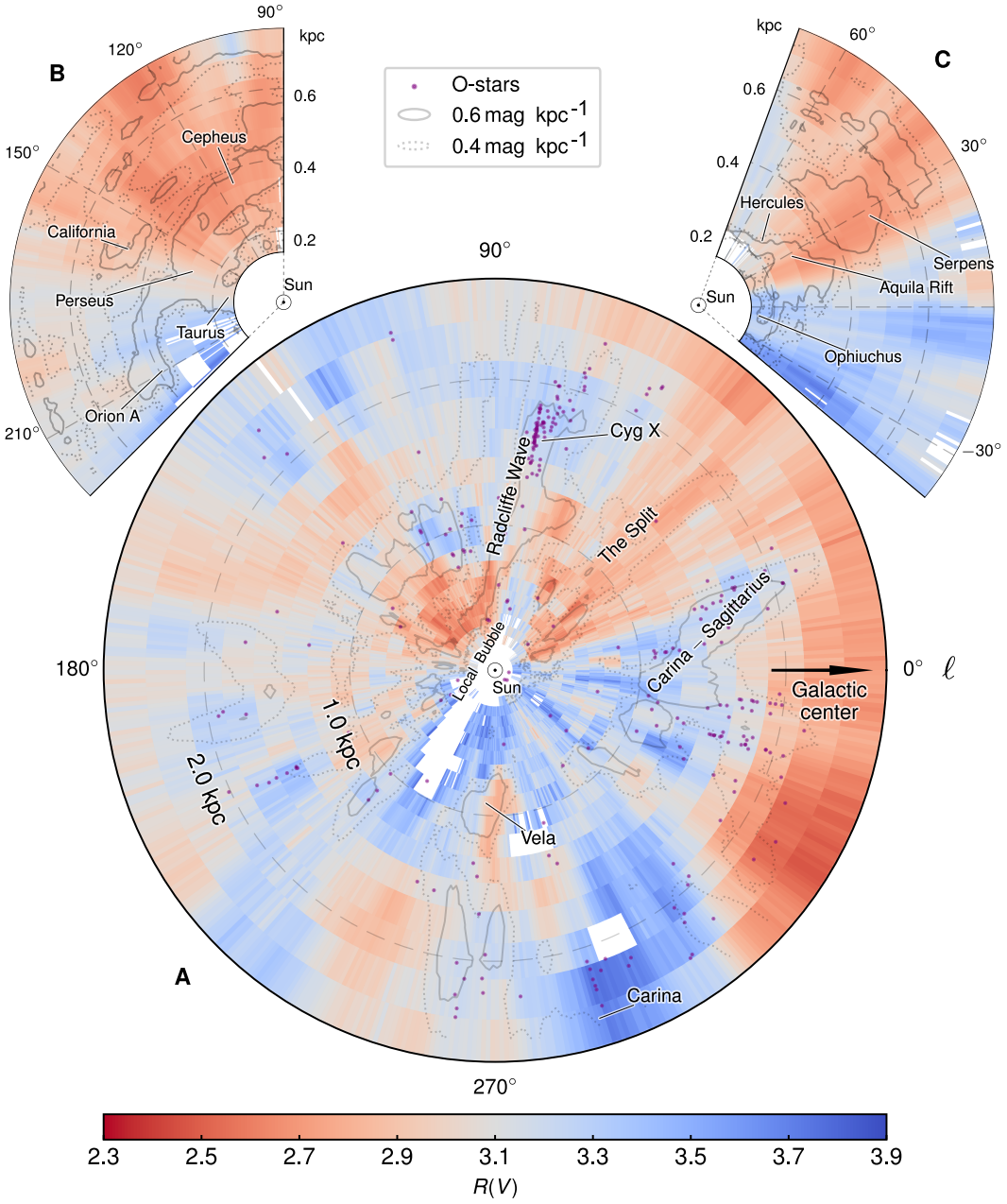


Figure 3.1. (A) Polar map for targets within 2.6 kpc of the Sun (marked with the ⊙ symbol) and within 400 pc above or below the plane. ℓ is the Galactic longitude. R(V) (color bar) was calculated by averaging our maps of differential extinction at 440 and 550 nm. Contours enclose regions of high differential extinction at 550 nm (see legend). White regions were excluded owing to insufficient extinction [$\Delta A(550\text{ nm})/\Delta D < 0.1$ mag kpc^{-1}] to measure R(V), including within the Local Bubble. Purple dots indicate the locations of O-type stars (Xu et al., 2018). Specific regions are labeled. (B and C) Zoomed views of the same map in two ranges of Galactic longitude, out to a distance of 0.8 kpc.

Figure 3.1 shows the projection of our $R(V)$ map onto the Galactic plane, for locations within 2.6 kpc of the Sun and <400 pc above or below the plane. We find spatial correspondence between the value of $R(V)$ and several known local structures (Alves et al., 2020; Lallement et al., 2019), which are labeled in the figure. The Carina-Sagittarius spiral arm has higher $R(V)$ than the average, while the Split feature has lower $R(V)$; both regions contain numerous molecular clouds. We ascribe this difference to the high density of OB stars and star-forming regions in Carina-Sagittarius, which are absent in the Split (Lallement et al., 2019). Averaging over the entire 3D map within 4 kpc of the Sun, we find a median $R(V)$ in each spatial voxel of 3.11 (1σ range from 2.60 to 3.88). Integrating along each line of sight, the median $R(V)$ is 3.08 (1σ range from 2.86 to 3.28). These values are consistent with previous studies that found a mean $R(V)$ within the Milky Way of about 3.1 (Cardelli et al., 1989; Schlafly et al., 2016).

Figure 3.2 shows maps of $R(V)$ across the sky in four distance bins. We find that regions containing molecular clouds have low $R(V)$ values. In the most distant bin, 2 to 5 kpc, we find low $R(V)$ values toward the Galactic Center ($\ell = 0°$), which is consistent with previous photometric results (Udalski, 2003).

Figure 3.3 shows sky maps of $R(V)$ zoomed into the positions of five dense clouds. We find low $R(V)$ in the outer, intermediate-density regions of three of these clouds (Orion A, Taurus, and part of Mon OB1). However, in the lines of sight toward the densest regions of the clouds, such as the Orion Nebula, we find higher $R(V)$ values than in the outskirts. For reasons that are unclear, the Perseus cloud shows a more complex pattern. In ρ Ophiuchus, both the central and outer regions have higher $R(V)$ than the average value in the Milky Way (~ 3.1), which could be due to the recent and ongoing star formation in this region (Briceño-Morales & Chanamé, 2023; Miret-Roig et al., 2022; Ratzenböck et al., 2023). Figure S9 shows the $R(V)$ maps of ρ Ophiuchus and the adjacent region of Upper Scorpius compared

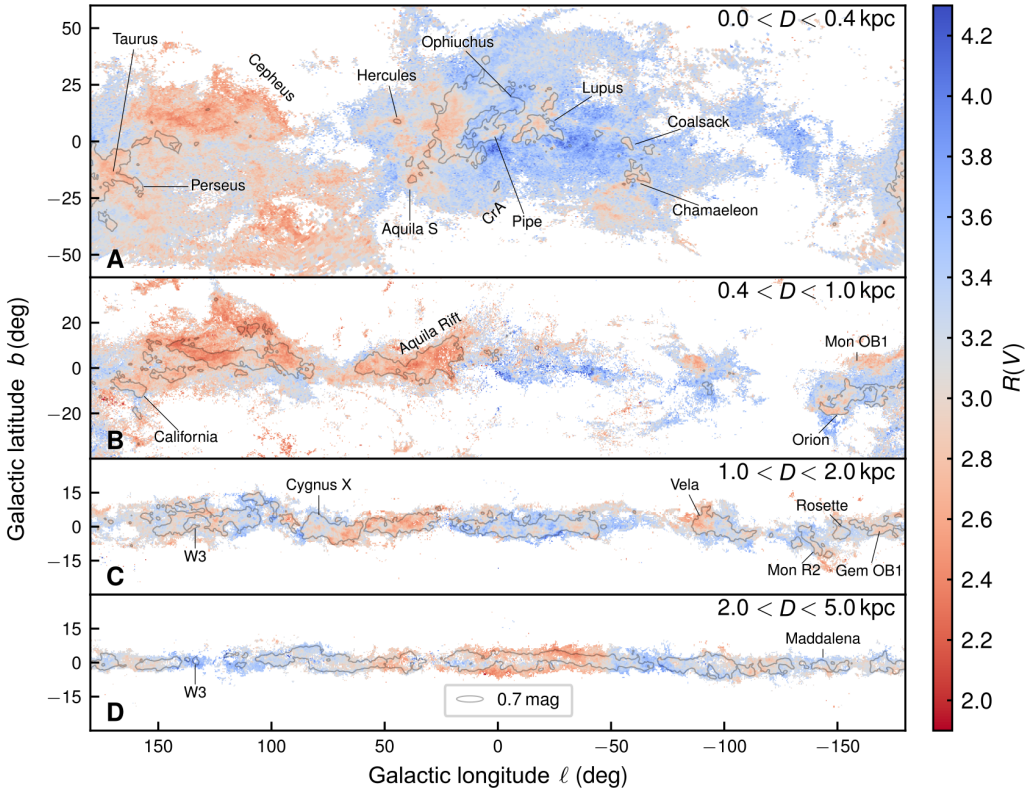


Figure 3.2. $R(V)$ (color bar) was calculated equivalently to Fig. 1 and projected onto the plane of the sky in four distance bins. $R(V)$ maps are shown in the distance slices (A) 0.0 to 0.4 kpc, (B) 0.4 to 1.0 kpc, (C) 1.0 to 2.0 kpc, and (D) 2.0 to 5.0 kpc. White areas are excluded owing to having $\Delta E/\Delta D < 0.1$ mag kpc^{-1} or $\Delta E < 0.1$ mag, where ΔE is the difference of extinction (E) in adjacent distance bins. Contours enclose high-extinction regions with $\Delta A(550\text{ nm}) > 0.7$ mag. Labels indicate the positions of dense molecular cloud complexes (Zucker et al., 2020).

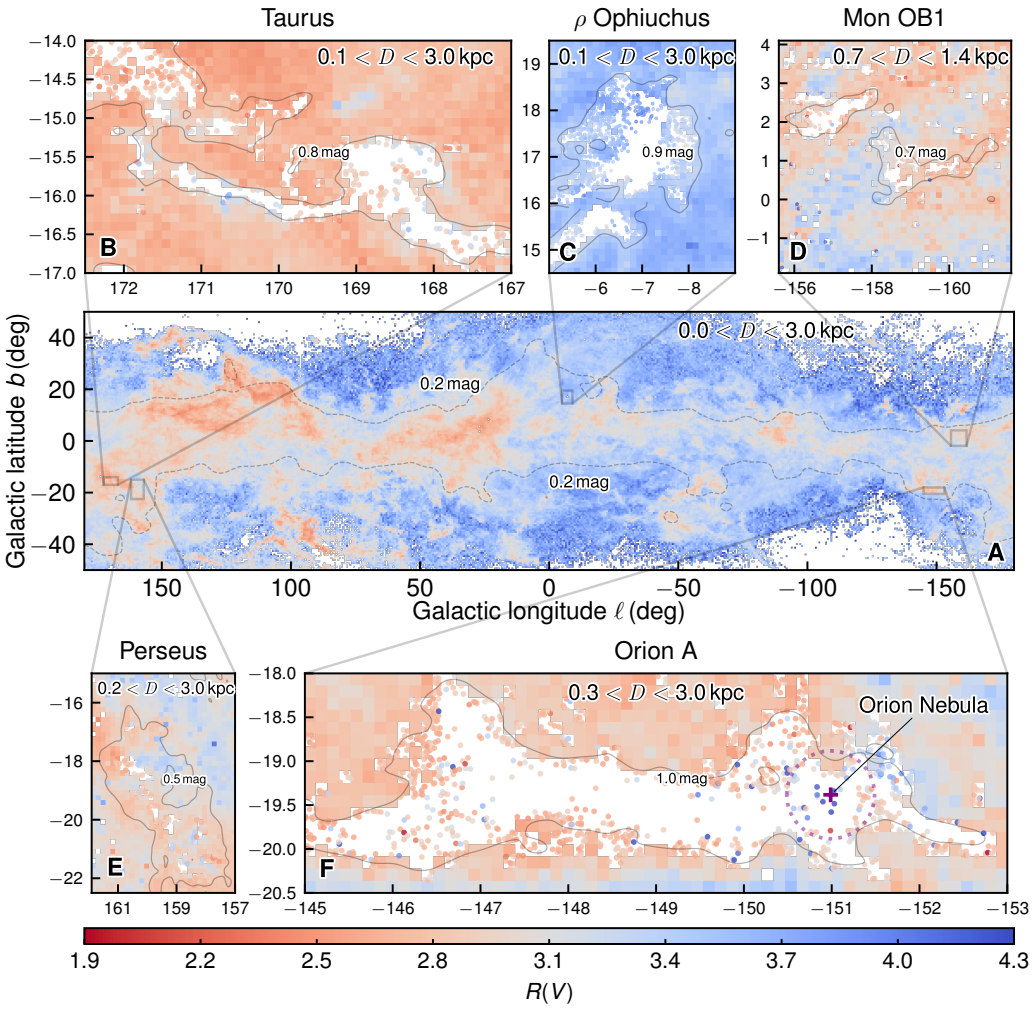


Figure 3.3. (A) $R(V)$ map using stars with $E > 0.1$ mag at distances between 0 and 3 kpc. (B to F) Zoomed views of five dense clouds, with $R(V)$ calculated using stars in distance ranges (labeled in each panel) that contain the selected cloud, determined from a previous 3D dust map (Green et al., 2019). Foreground extinction has been removed from each panel by subtracting maps of the average extinction of stars in front of each cloud. Gray contours enclose regions of high extinction (labels indicate the values of E), taken from the previous map (Green et al., 2019). In highly extinguished regions with few observed stars, $R(V)$ values of individual stars are plotted as dots. In (F), the center and approximate radius of the Orion Nebula are denoted by the purple cross and circle, respectively.

with the distribution of OB stars, which are tracers of star formation.

We investigated how $R(V)$ depends on extinction in four clouds (Fig. 4). For Orion A, Perseus, and Taurus, we find a U-shaped relation: $R(V)$ decreases with increasing E at low extinctions and then shows the opposite trend at high extinction ($E \gtrsim 0.75$ mag). No such minimum is seen in the ρ Ophiuchus data.

We also mapped $R(V)$ in the Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC), two satellite galaxies of the Milky Way. The LMC and SMC both have lower metallicity than does the Milky Way and so could potentially have different dust properties. After removing foreground extinction of the Milky Way, we find that the SMC has lower $R(V)$ than the Milky Way average, which is consistent with previous studies (Udalski, 2003). The LMC has a wider range of $R(V)$, with higher values in some star-forming regions (Fig. 5A).

3.5 Implications for dust properties

We have identified a general trend wherein $R(V)$ is usually higher than average in regions of the ISM with low column density of dust, as traced by E (Fig. 3A), and $R(V)$ typically decreases as E increases. In the highest-extinction regions of some clouds, such as Orion A, Taurus, and Perseus (Fig. 4, A to C), we find that this trend reverses, with $R(V)$ increasing with greater extinction. Previous work has shown that $R(V)$ tends to increase in the cold, dense cores of molecular clouds (Chapman et al., 2008; Steinacker et al., 2010). We generally observe decreasing $R(V)$ with increasing extinction in the translucent (moderate-density) ISM, before $R(V)$ increases again in the densest inner regions of clouds. Next, we considered possible physical explanations for this trend. The two dominant mechanisms of dust grain growth in the ISM are (i) accretion of elements from the gas phase as they condense onto the surface of grains and (ii) coagulation of grains that collide and stick to-

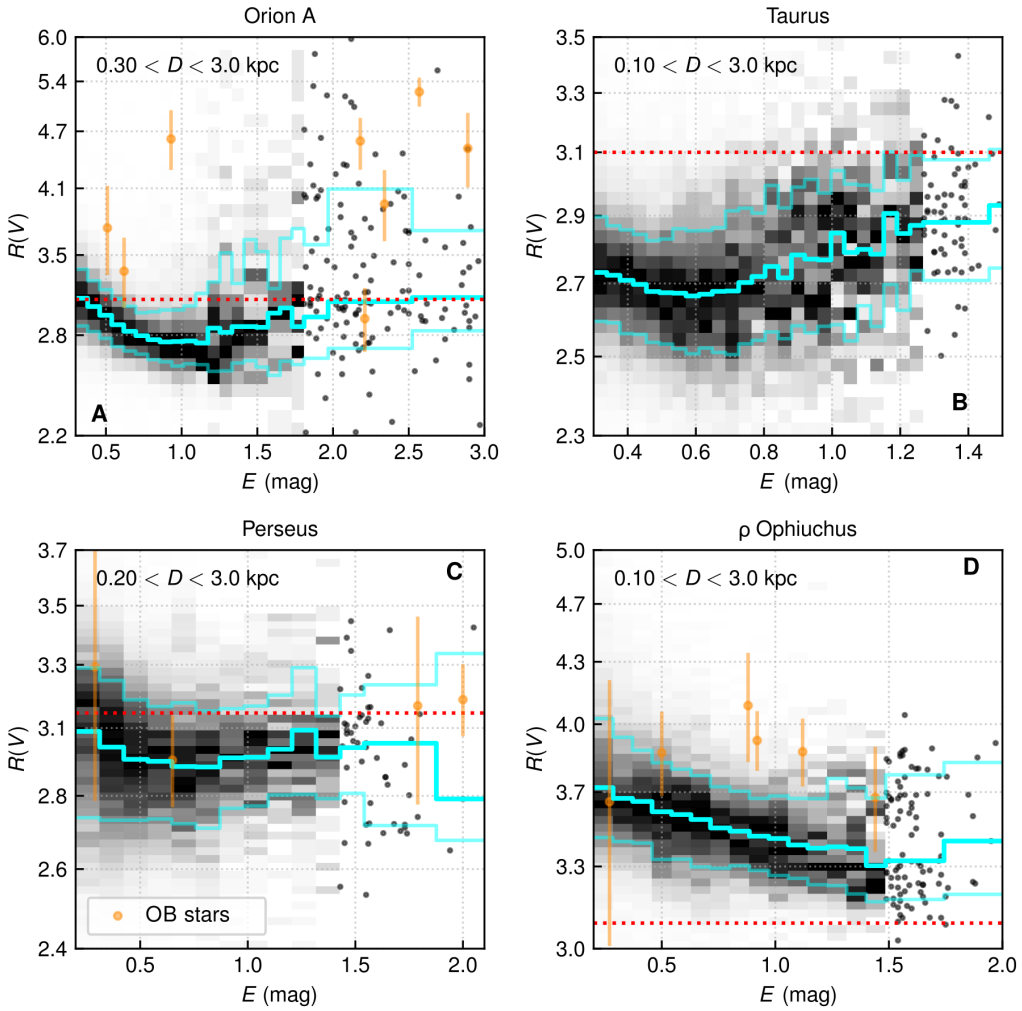


Figure 3.4. Shading in each panel shows the conditional distribution of $R(V)$ for each value of dust extinction E at a distance of 3 kpc, as predicted independently using a 3D dust extinction map (Green et al., 2019). White indicates zero probability density and black corresponds to the maximum probability density in each extinction bin. Data are shown for four clouds: (A) 15,147 stars in Orion A, (B) 19,996 stars in Taurus, (C) 50,120 stars in Perseus, and (D) 44,830 stars in ρ Ophiuchus. For each cloud, we selected stars within a distance range (labeled in each panel) chosen to select stars that are either embedded in or behind the cloud. At high extinctions, where there are fewer than 20 stars in each bin, gray dots indicate individual stars. Cyan lines mark the 16th, 50th, and 84th percentiles of the $R(V)$ distribution in each extinction bin. Red dotted lines show the average Milky Way $R(V)$ of ~ 3.1. Yellow circles indicate $R(V)$ measurements (error bars show 1σ uncertainties) of individual hot stars from previous studies (Gordon et al., 2023; Valencic et al., 2004)

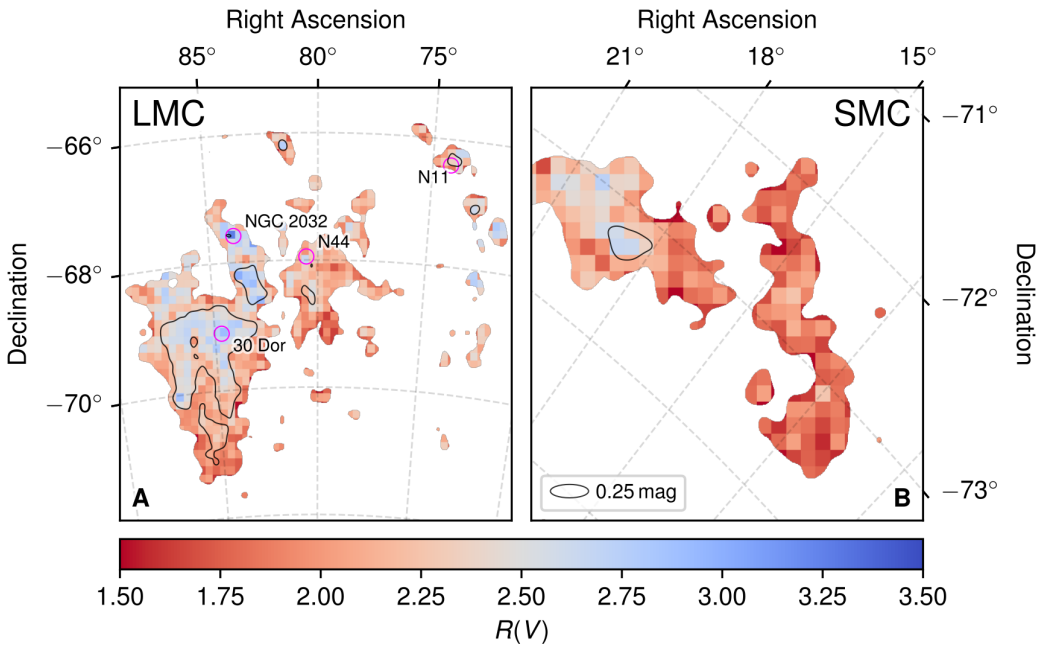


Figure 3.5. $R(V)$ (color bar) mapped in (A) the LMC and (B) the SMC. Regions with $A(550\ \mathrm{nm}) < 0.15$ mag have been masked out. Foreground extinction in the Milky Way has been subtracted (see text). Magenta circles in (A) indicate star-forming regions in the LMC, most of which have higher $R(V)$ than the LMC average. Dashed gray lines are a coordinate grid.

gether. Although both processes increase the average grain size, their effects on the grain-size distribution, and thus $R(V)$, are different (4). Accretion tends to deposit a layer of approximately equal thickness onto each grain, so smaller grains experience a larger fractional increase in surface area than do the larger grains. The resulting cross section–weighted average grain size decreases, so the extinction curve steepens and $R(V)$ decreases. In contrast, coagulation converts pairs of small grains into single larger grains, preferentially depleting the population of small grains. This drives the grain-size distribution toward larger grains, so the extinction curve becomes flatter and $R(V)$ increases. Because accretion involves interactions between dust grains and the surrounding gas, it occurs at lower densities than does coagulation, which requires direct collisions between dust grains. As more-refractory (low condensation temperature) elements accrete onto grains, they become depleted from the gas phase, so the accretion rate decreases (2). We interpret our results as indicating that accretion is the dominant mechanism of grain growth in the moderately dense translucent regions of interstellar clouds, which is consistent with the observed decrease in $R(V)$ with dust density. We suggest that, at higher densities, coagulation becomes the dominant grain growth mechanism, so $R(V)$ increases. In the maps, we identify regions of high $R(V)$ close to locations with recent (or ongoing) star formation. We observe high $R(V)$ across the entire Upper Scorpius and ρ Ophiuchus region (Fig. 3C), which hosts an association of OB stars and contains both ongoing and recent star formation (Briceño-Morales & Chanamé, 2023; Miret-Roig et al., 2022; Ratzenböck et al., 2023). Two-dimensional sky projections show that this region has high $R(V)$ in both low- and high-extinction regions (Fig. 4D), which might be typical of star-forming regions. In Fig. 1, concentrations of OB stars are often surrounded by regions of high $R(V)$. For example, although much of the Radcliffe Wave region has low $R(V)$ and contains few OB stars, Cygnus X (located at one end of the wave) contains a large concentration of OB stars and has higher $R(V)$. The Split region,

which contains few OB stars, has low $R(V)$. The Carina-Sagittarius spiral arm has a higher density of OB stars and also has high $R(V)$. This behavior is also apparent in the LMC, where a few star-forming regions have higher $R(V)$ than the rest of that galaxy (Fig. 5A). There are several physical mechanisms that could produce this association between star formation and high $R(V)$, including preferential destruction of small grains by supernova shocks (Asano et al., 2013), or the formation of large grains in dense molecular cores (Hirashita & Yan, 2009; Ormel et al., 2009), which are dispersed into the translucent ISM by stellar feedback (Murray et al., 2009). Previous work (Schlafly et al., 2016) found an anticorrelation between $R(V)$ and the dust emissivity spectral index β (Abergel et al., 2014), which quantifies the spectral shape of dust thermal emission in the far infrared. We find a similar trend for low-Galactic-latitude sightlines. Those sightlines are complex, with multiple interstellar clouds along each one, and so cannot be used to determine whether β and $R(V)$ are anticorrelated in individual clouds. We therefore investigated these parameters in clouds at higher Galactic latitudes, finding a weaker anticorrelation between β and $R(V)$ than in previous work.

3.6 Utility for extinction corrections

Many astronomical observations require correction for foreground dust extinction. Previous dust extinction maps have assumed a constant value of $R(V)$, which we estimate introduces a systematic uncertainty of $\sim 10\%$ in extinction [assuming the typical range of $R(V)$ we observe in the Milky Way]. Our 3D map provides the $R(V)$ value for foreground dust as a function of distance along any given sightline in the Milky Way, which can be incorporated into dust correction calculations to avoid this systematic uncertainty.

Chapter 4

A hypothesis on $R(V)$ variations

Disclaimer: This chapter has been published by the Astrophysical Journal Letters as Zhang et al. (2025).

4.1 abstract

The first all-sky, high-resolution, 3D map of the optical extinction curve of the Milky Way (Zhang & Green, 2025, "ZG24") revealed an unexpected steepening of the extinction curve in the moderate-density, "translucent" interstellar medium (ISM). We argue that this trend is driven by growth of the total mass of polycyclic aromatic hydrocarbons (PAHs) through gas-phase accretion. We find a strong anti-correlation between the slope of the optical extinction curve – parameterized by $R(V)$ – and maps of the PAH mass fraction (relative to the total dust mass) – parameterized by $q_{\rm PAH}$ – derived from infrared emission. The range of observed $q_{\rm PAH}$ indicates PAH growth by a factor of $\sim$2 between $A_V \simeq 1$ and 3. This implies a factor-of-two stronger 2175 Å feature, which is sufficient to lower $R(V)$ by the observed amount. This level of PAH growth is possible given rapid accretion timescales and the depletion of carbon in the translucent ISM. Spectral observations by JWST would provide a definitive test of this proposed explanation of $R(V)$ variation.

4.2 An Anti-correlation between $R(V)$ and PAH mass fraction

In the regions where $R(V)$ decreases with $A(V)$ in ZG24, we generally find an anti-correlation between $q_{\rm PAH}$ and $R(V)$ from the maps of Planck Collaboration et al. (2016b) and Zhang & Green (2025). This relation persists among the translucent regions in the vicinity of various molecular clouds (Aquila S, California, Orion A, Perseus, Ursa MA, etc) in the Milky Way and the Large Magellanic Cloud (LMC).

An ideal comparison between $R(V)$ and PAH abundance would be conducted in 3D. However, PAH maps are based on emission features of PAHs integrated to infinite distance, and are thus 2D in nature. To measure the same column of dust, we build 2D $R(V)$ maps by a binning method, using only stars that are behind most of the dust cloud. We describe the details of the method, as well as the selection of stars, in section 4.6.

We use the PAH mass fraction ($q_{\rm PAH}$) map from the Planck intermediate results (Planck Collaboration et al., 2016b), which fits the dust model of Draine & Li (2007) to dust emission data from the Planck satellite (Planck Collaboration et al., 2014b,c), the Infrared Astronomical Satellite ("IRAS"; Miville-Deschênes & Lagache, 2005), and the Wide-Field Infrared Survey Explorer ("WISE"; Meisner & Finkbeiner, 2014) observations. The fitting of 2D emission data assumes that the radiation field and the temperature of the dust along the line of sight are the same, which does not necessarily hold for lines of sight that cross multiple dust clouds or extended dust structures. Therefore, we select lines of sight where the majority of the extinction (over 70%) is concentrated within a 500-pc distance range:

$$\max_{0<d_0<4.5\rm kpc}\{\int_{d_0}^{d_0+500\,\rm pc} dE\}/E_{5\,\rm kpc} > 0.7. \tag{4.1}$$

We present $A(V)$, $R(V)$ and $q_{\rm PAH}$ of five clouds in the Milky Way in Figures 4.1 and 4.2, in which most sightlines fulfill our requirements. Despite finer structures in both $R(V)$ and $q_{\rm PAH}$ maps, we generally observe lower $R(V)$ in higher-$q_{\rm PAH}$ regions in all five clouds, which is shown quantitatively in Figure 4.3. An increase in from 4% to 8% is observed to result in a decrease in $R(V)$ by $\sim 0.2 - 0.5$. The median $R(V)$ at low $q_{\rm PAH}$ varies among the clouds, which may be attributed to their different initial dust grain size distribution, as well as magnetic field variations (Hensley, 2024; Voshchinnikov, 1989). In all five clouds, $R(V)$ approaches a value of ~ 2.9 at $q_{\rm PAH} \simeq 8\%$.

The scatter of $R(V)$ at given $q_{\rm PAH}$ is large (~ 0.3) in some clouds shown in Figures 4.1 and 4.2. The following factors may contribute to the scatter:

- Finer structures of clouds exist but the $q_{\rm PAH}$ map has limited angular resolution.
- Variation of both $q_{\rm PAH}$ and $R(V)$ exist along the line of sight, which 2D maps cannot represent.
- Uncertainty of measured $R(V)$ of individual stars.
- The WISE $W3$ (centered at $12\,\mu$m, Wright et al., 2010) photometric band has an underlying continuum that is not from PAHs, and the decomposition provided by the dust model (Draine & Li, 2007) may not be perfect. Further, a number of gas emission lines can contribute to the total $W3$ flux.
- Magnetic fields may vary in the dust clouds, which can lead to $R(V)$ variation of order ~ 0.2 (Hensley, 2024).

In addition to the Milky Way, the anti-correlation between $q_{\rm PAH}$ and $R(V)$ is also observed in the LMC. Figure 4.4 shows the 2D $R(V)$ distribution from ZG24 and the map from Chastenet et al. (2019). When generating our $R(V)$ maps for the LMC and SMC, we subtract foreground extinction from Milky Way dust, using the method described in section 4.7. We label H II regions by magenta circles (Lopez

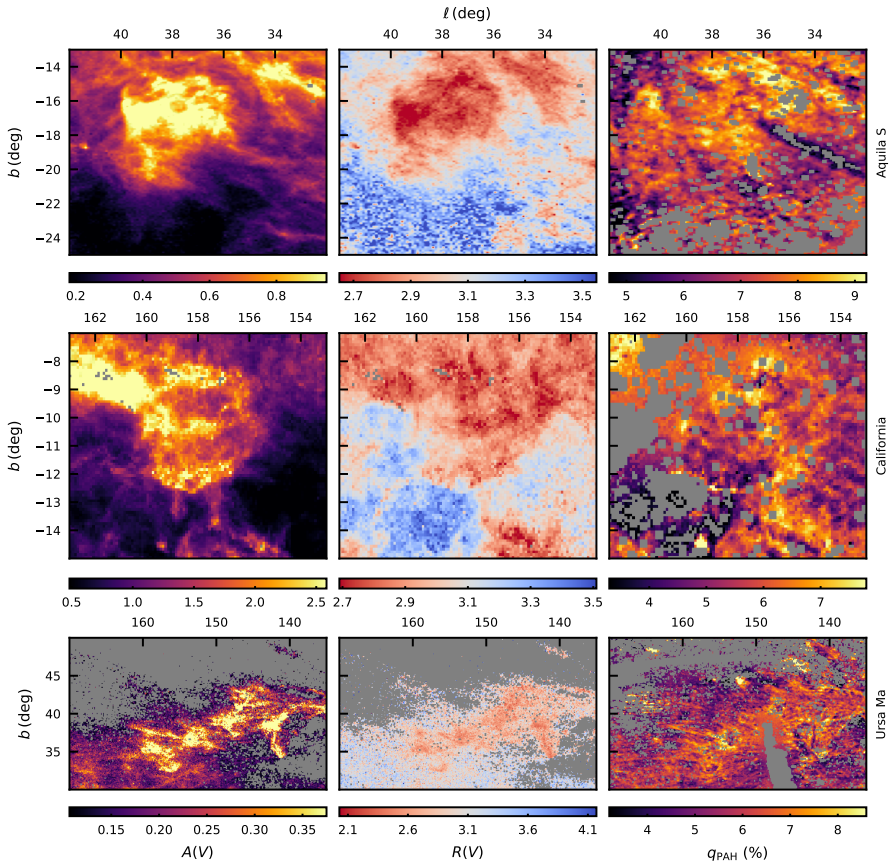


Figure 4.1. Comparison of 2D distribution of $A(V)$, $R(V)$, $q_{\rm PAH}$ for dust clouds. For $A(V)$ and $R(V)$ maps, we select high-quality stars from Zhang & Green (2025), and remove lines of sight where only partial extinction is covered by XP. For $q_{\rm PAH}$ map, we remove lines of sight where the $q_{\rm PAH}$ map is of low quality ($\chi^2/{\rm DOF} > 3$ or $q_{\rm PAH}/\sigma(q_{\rm PAH}) < 3$), or that cross multiple clouds (Equation 4.1). The binning method and selection of stars in the 2D maps are described in section 4.6. Despite different baselines and finer structures in different maps, all clouds tend to have lower $R(V)$ at high-$q_{\rm PAH}$ regions, as indicated quantitatively on the per-star basis in Figure 4.3.

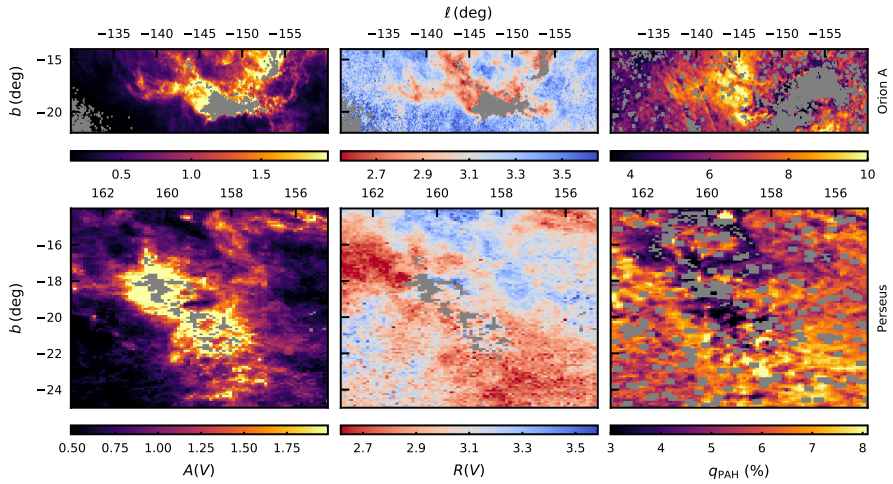


Figure 4.2. Continuation of Figure 4.1

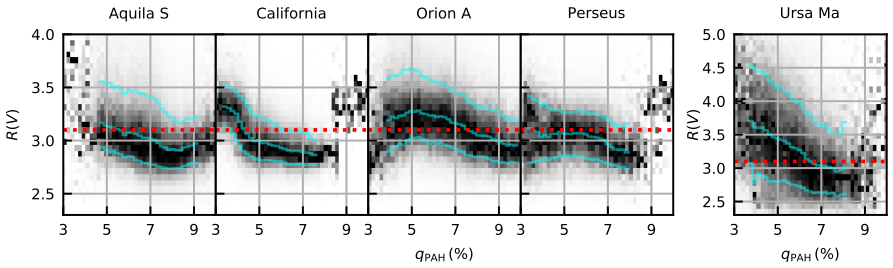


Figure 4.3. Correlation between $R(V)$ and $q_{\rm PAH}$ using stars that pass all the cuts in Figure 4.1. 16th, 50th and 84th percentiles are in cyan lines. We also label $R(V) = 3.1$, the Milky Way average value, in red dotted lines. Although different clouds have different baselines of $R(V)$ due to different initial distribution of dust grain size, all dust clouds show anti-correlation between $R(V)$ and $q_{\rm PAH}$.

et al., 2014). In the LMC, most H II regions have lower and higher $R(V)$ than the surrounding regions, which is consistent with the trend we find in the Milky Way. We also find an anti-correlation between $R(V)$ and on a per-star basis in the LMC, as illustrated in Figure 4.5.

In the SMC, we do not see any trend between $R(V)$ and $q_{\rm PAH}$. However, it is notable that the PAH mass fraction, as well as the range of its variation, is low in the SMC. The size distribution of larger grains could also be significantly different from that of the Milky Way (Weingartner & Draine, 2001).

4.3 Theoretical explanation of $R(V)$ variation in the translucent ISM

In this section, we demonstrate theoretically that the variation of is capable of changing $R(V)$ as we observe in Section 4.2, and rule out the possibility of accretion onto larger dust grains causing the observed decrease in $R(V)$.

Within a dust cloud, we assume that dust grains are well mixed when the cloud initially forms, such that the initial grain size distribution (of each dust component) is homogeneous. As a cloud evolves, growth and destruction of dust grains shape the size distribution differently at different densities within the cloud, which gives rise to $A(V)$ and $R(V)$ variations.

We adopt the dust grain model by Hensley & Draine (2023), which consists of two components, "astrodust" (Draine & Hensley, 2021b), and PAHs. Astrodust is a composite material to account for large grains ($\gtrsim 500$ Å) that contains the majority of the elements depleted from the gas phase, such as C, O, Mg, Si and Fe. Small grains are accounted for by PAHs.

By fitting the grain-size distribution of astrodust and PAHs, this model reproduces the observed extinction, polarized extinction, emission, and polarized emis-

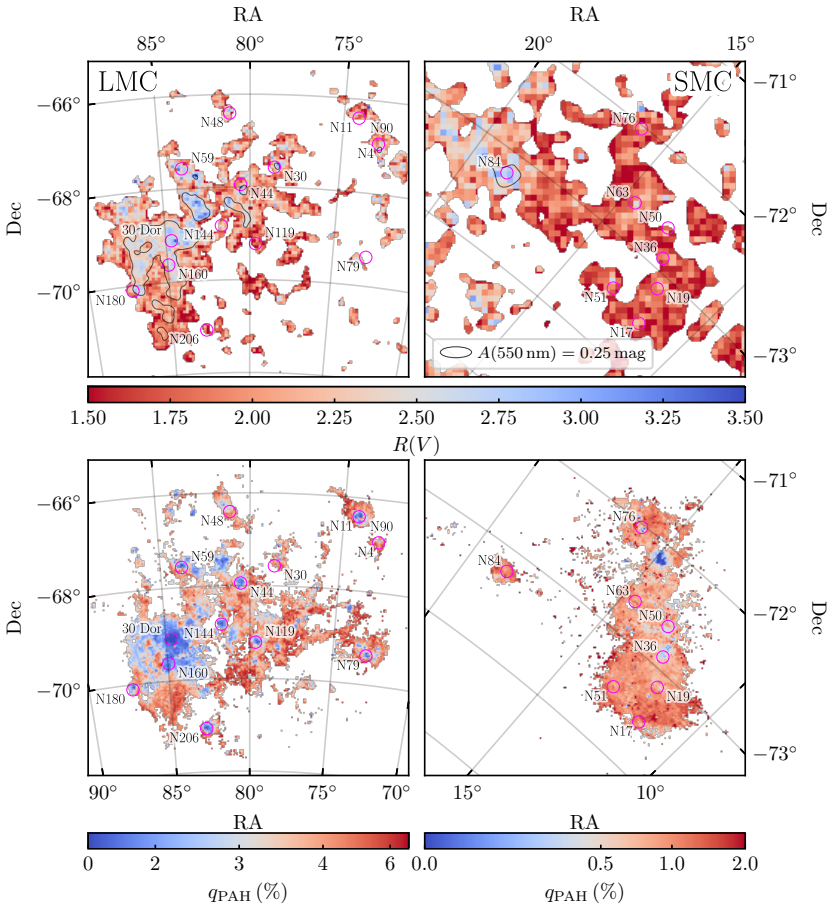


Figure 4.4. $R(V)$ distribution of LMC and SMC is shown in the top panels. We select stars by distance and proper motion and remove foreground extinction, described in section 4.7, We label a few dense dust clouds in magenta circles indicated by H II signal from Lopez et al. (2014), which tend to have higher $R(V)$ than their neighborhood. We also show the $q_{\rm PAH}$ map from Chastenet et al. (2019) in the bottom panels. The H II regions generally have lower $q_{\rm PAH}$. The quantitative comparison on the per-star basis is shown in Figure 4.5.

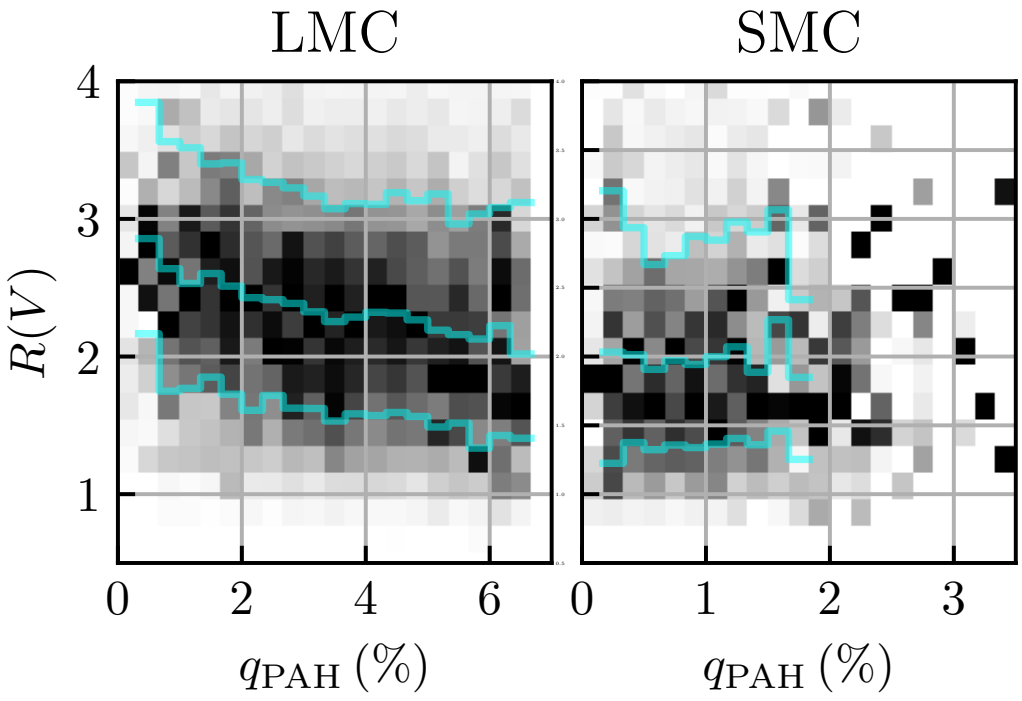


Figure 4.5. Quantitative comparison on the per-star basis between $R(V)$ and q_{PAH} in LMC (left panel) and SMC (right panel). The median value of $R(V)$ decreases by 0.4 as q_{PAH} increases from 1% to 6% in the LMC. In the SMC, the trend is not clear because of small variation range of q_{PAH}.

sion from dust in the diffuse ISM. We use the best-fit grain-size distribution of astrodust and PAHs in Hensley & Draine (2023) as the initial distribution, and then model the growth of dust through accretion.

In the translucent ISM, where ZG24 observe a decrease in $R(V)$ with increasing $A(V)$, the evolution of dust is dominated by accretion (Hirashita, 2012), whereby metals from the gas phase are accumulated on the surface of dust grains. As a result, the growth rate of grain mass (m) is proportional to the grain surface area:

$$\frac{dm}{dt} \propto 4\pi\rho a^2, \tag{4.2}$$

where ρ is the density of dust grains and a is the "effective radius," defined as $a \equiv (3m/4\pi\rho)^{1/3}$, a measurement of grain size. Factors affecting the accretion rate include the dust temperature, the metal abundance in the gas phase and the sticking probability (Hirashita, 2012), all of which are nearly independent of a. If we further assume the mass density of a dust grain, ρ, to be independent of its effective radius a, then $\frac{da}{dt}$ is independent of effective radius a. Therefore, the net growth of effective radius, Δa, over some timespan is roughly the same for all dust grains despite their different initial sizes.

Optical-NIR extinction of the astrodust component is dominated by the grains with effective radius of $\gtrsim 0.1\,\mu$m (Draine & Hensley, 2021b; Hensley & Draine, 2023). As illustrated in Figure 4.6, in order for accretion onto astrodust to effectively change $R(V)$ from 3.1 to 2.5, Δa has to reach $\simeq 0.3\,\mu$m. This would require $\simeq 291$ times the initial mass of astrodust to be accreted onto the astrodust grains, greatly exceeding what is available in the gas phase. The abundance of major elements of astrodust, Mg, Fe and Si, available in gas phase is 7.1, 0.88 and 6.6 parts per million hydrogen atoms (ppm) (Hensley & Draine, 2021; Jenkins, 2009), respectively, which only enables growth of dust mass by $\sim 16\%$, 2% and 17%. This

rules out the possibility that accretion onto large grains could change $R(V)$ by the observed amount.

On the other hand, the red wing of the 2175 Å feature extends to optical-NIR wavelengths and has a profile that is steep enough to change the slope of the extinction curve in this regime (Hensley & Draine, 2023, Figure 3). Arising from $\pi^* \leftarrow \pi$ electronic transitions in carbon rings, the strength of the feature is proportional to the total mass of PAHs, since their physical size (nanometers) is much smaller than 2175 Å (i.e., absorption is in the electric dipole limit). We can estimate $R(V)$ variation as a function of PAH mass fraction, represented by the total dust mass fraction of PAHs (). In Figure 4.6, the orange curve shows $R(V)$ variation with PAH growth, holding the astrodust component fixed. In order to change $R(V)$ by ~ 0.6, PAH mass only needs to grow by 91%, corresponding of an increase in $q_{\rm PAH}$ from 6% to 11%. This amount of growth requires another 54 ppm of C to be added to PAHs via enhanced depletion, while there is ~ 200 ppm of carbon available in the gas phase, (Hensley & Draine, 2021; Jenkins, 2009). Enhanced depletion of carbon has been observed in denser regions of the Milky Way (e.g. Roman-Duval et al., 2021; Zuo et al., 2021). However, large uncertainty remains in the measurements of solid-phase carbon abundance (e.g. Hensley & Draine, 2021; Jenkins, 2009; Ysard et al., 2024). Consequently, this scenario is plausible within the current framework of our theory, but needs to be verified with improved measurements of depletion.

Although allowed by the mass budget, the question then arises as to whether PAHs can double in mass within a reasonable timescale, which we take the lifespan of a typical molecular cloud as a lower limit. Unfortunately, the growth pattern of PAHs in molecular clouds remains unclear. Inspired by Krasnokutski et al. (2017), we estimate the time scale by a simple accretion model. We assume carbon atoms in the gas that collide with the rim (width ≈ 1.4 Å) of a PAH molecule are incorporated into the molecule. We assume atoms in gas follow a Maxwellian velocity distribution

with $T = 20\,\mathrm{K}$, and calculate the time scale for a few particular PAH molecules to double in size, tabulated in Table 4.2.

The resulting timescales of only $\sim$2 Myr are much shorter than the lifespan of a molecular cloud (typically 10–20 Myr). Further, the accretion rate of carbon ions onto neutral PAHs could be enhanced by more than an order of magnitude due to electromagnetic effects (Weingartner & Draine, 1999). We describe this estimate in detail in section 4.8. We therefore conclude that doubling the PAH mass is plausible.

Although PAHs have a significant impact on $R(V)$ through the 2175 Å feature, other effects also influence $R(V)$. The size distribution of larger grains may vary considerably among clouds, which can substantially alter $R(V)$ (as indicated in Figure 4.3). However, in a single cloud, assuming that the grains are initially well mixed and the initial grain-size distribution therefore does not vary within the cloud, we find that accretion onto large grains cannot explain the observed variations in $R(V)$ between the diffuse outskirts of the cloud and the translucent, intermediate-density regions of the cloud. In contrast, in situ growth of PAHs can explain the observed trends in $R(V)$ with density.

Finally, we emphasize that our result is not specific to the astrodust model. Although the precise range of $R(V)$ variation may change with different dust models, the trend of decreasing $R(V)$ with PAH mass fraction holds, because it only relies on PAHs being the carrier of the 2175 Å feature.

4.4 Discussion

4.4.1 Constraints on the PAH evolution

Our findings underscore the importance of the growth of PAHs in the ISM, which is still poorly understood. To list a few questions:

- Do PAHs form in-situ or migrate from denser regions?

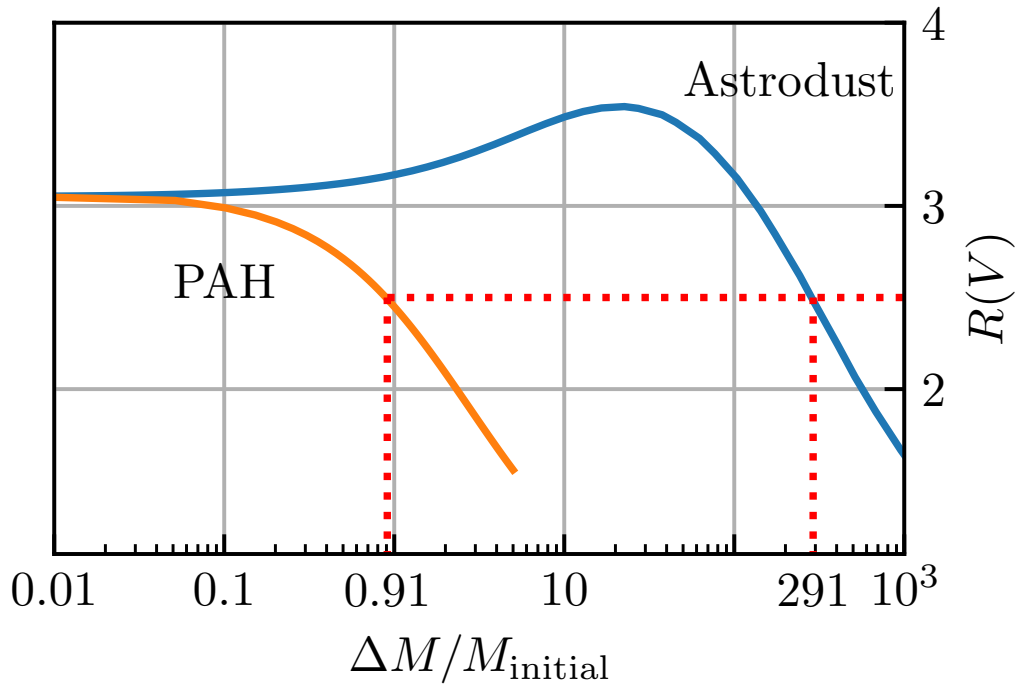


Figure 4.6. $R(V)$ variation as a function of the mass accreted onto astrodust or PAHs relative to initial mass, holding the other component (PAHs or astrodust) fixed. For $R(V)$ to decrease from 3.1 to 2.5, the astrodust mass must grow by a factor of ~ 291. This greatly exceeds the available mass budget of Mg, Fe, and Si (the major components of astrodust) in the gas. However, PAHs only have to grow by 91% in mass, or 54 ppm of C, which is a viable amount given the mass budget of carbon in the gas phase (Hensley & Draine, 2021; Jenkins, 2009).

- How are carbon atoms “assembled” onto PAHs?
- What affects the growth or destruction rate of PAHs?

PAH growth in the ISM has been suggested by various recent observations. Whitcomb et al. (2024) found that with decreasing metallicity, the flux of long-wavelength PAH emission features tends to decrease, while the flux of short-wavelength features increases. Whitcomb et al. (2024) attribute this to inhibited grain growth at low metallicity. Type Ia supernovae are observed to have anomalously low $R(V)$ (Popovic et al., 2023), which could be connected to PAHs: The 2175 Å features observed in galaxies from cosmic dawn indicate the possibility that supernovae trigger PAH formation through shattering of graphite or amorphous carbon grains (Witstok et al., 2023; Yang & Li, 2023). Aromatic molecules can also be formed in a “bottom-up” way from precursors (e.g., Jones et al. 2011; McGuire et al. 2021). Experiments and computations indicate that PAHs can acquire carbon from the gas over a wide temperature range (Krasnokutski et al., 2017).

Taken together with the results of the present study, these observations suggest that PAH growth in the ISM is a major component of the PAH life cycle in galaxies and is responsible for much of the observed PAH mass.

4.4.2 Implications for future JWST observations

JWST/MIRI covers PAH emission features at 6.2, 7.7, 8.6, 11.2, 12.7 and 16.4 μm and is powerful in determining PAH composition (e.g. Chastenet et al. 2023; García-Bernete et al. 2022; Sturm et al. 2024). JWST/MIRI MRS measurements of a series of targets in the transluscent ISM with similar $A(V)$ but a gradient in $R(V)$ could trace the different evolutionary phases of PAHs. In addition to testing the hypothesis put forward here, this could provide valuable constraints on the carbon cycle in the ISM.

$R(V)$ variation caused by the 2175 Å feature relies primarily on the total mass

fraction of PAHs, and is not sensitive to the PAH size distribution. Nevertheless, $R(V)$ maps indicate locations of significant variation in PAH mass fraction, which are potential sites of growth, destruction, or migration of PAHs. Therefore, these maps can be used to identify optimal targets for observing PAH evolution with JWST.

4.5 Summary

In this work, we find an anti-correlation between $R(V)$ and in the translucent ISM. We observe this phenomenon both in the vicinity of various molecular clouds in the Milky Way and in the LMC. These regions are also where $R(V)$ decreases with $A(V)$ in ZG24.

We propose a theory to explain both anti-correlations: $R(V)$ variation is driven by the growth of PAHs. The 2175 Å feature enables PAHs to change the slope of extinction curves at optical-NIR wavelengths even though the sizes of the molecules are far smaller than the wavelength of the B or V band. This theory is consistent both with the mass budget of carbon atoms available in gas phase and with the estimated time scale for PAHs to gain the needed mass.

We discuss potential JWST observations of the translucent ISM to resolve PAH absorption and emission features. Such observations would not only test the PAH-$R(V)$ correlation but also impose important constraints on the evolution of PAHs and the carbon cycle in the ISM.

4.6 2D $R(V)$ maps of clouds in the Milky Way

In order to compare the $R(V)$ results with the 2D $q_{\rm PAH}$ map, we build $A(V)$ and $R(V)$ maps using binning method. We select stars that pass the `basic cut` and the $T_{\rm eff}$ `confidence` cut in Zhang & Green (2025). Because the clouds we select are the main source of extinction along the lines of sight, we can simply remove the

foreground contamination by only using stars beyond the minimal distances of dust clouds, as tabulated in Table 4.1, determined by Schlafly et al. (2014).

The Gaia XP catalog has a magnitude limit of 17.65 mag (De Angeli et al., 2023; Montegriffo et al., 2023), which excludes heavily-extincted stars behind some dust clouds so that binning method could only measure partial extinction. To make sure our $R(V)$ map describe the majority of dust along the line of sight, we select stars by comparing extinction measure by Gaia XP, $A(V)_{\rm XP}$, with that of 2D extinction map "SFD" (Schlegel et al., 1998b), $E_{\rm SFD}$:

$$A(V)_{\rm XP} - kE_{\rm SFD} - \texttt{bias} > -0.5\,\text{mag}. \tag{4.3}$$

Coefficients k and `bias` are fitted using low-extinction regions of the cloud with $0.1 < E_{\rm SFD} < 0.8\,$mag, where most of the XP stars are behind clouds. We tabulate the coefficient values in Table 4.1 .

We allocate stars into bins of $\Delta\ell \times \Delta b = 0.1^\circ \times 0.1^\circ$, and calculate the average $A(B)$, $A(V)$ and $R(V)$ by the following equation in each bin:

$$\begin{aligned}
\bar{A}(B) &= \sum_i w_{B,i} A(B)_i / \sum_i w_{B,i}, \\
\bar{A}(V) &= \sum_i w_{V,i} A(V)_i / \sum_i w_{V,i}, \\
\bar{R}(V) &= \bar{A}(V)/(\bar{A}(B) - \bar{A}(V))
\end{aligned} \tag{4.4}$$

where i is the index of stars in each bin, and the weight of each star is given by $w_i = (\sigma(A(B))^2 + \sigma(A(V))^2 + 0.01^2)^{-1}$.

Cloud name	Aquila S	California	Orion A	Perseus	Ursa Ma
Minimal distance (kpc)	0.05	0.3	0.3	0.2	0.2
k	2.14	1.78	1.92	1.87	2.19
bias	0.17	0.24	0.12	0.21	0.14

Table 4.1. Coefficients to select stars into the 2D $R(V)$ maps.

4.7 2D $R(V)$ maps for LMC and SMC

In order to compare with the q_{PAH} maps from Chastenet et al. (2019), which remove foreground emission, we also remove the foreground extinction of LMC and SMC. We still select stars by the quality cuts described in section 4.6, but additionally define two group of stars: "foreground" and "target" stars. The foreground stars are the stars in the Milky Way halo that fulfill **both** of the following condition:

$$\begin{aligned} &\varpi_{\text{est}} - 2\sigma(\varpi_{\text{est}}) > 1/(5\text{kpc}) \\ &|\sin(b)|\,(\varpi_{\text{est}} + 2\sigma(\varpi_{\text{est}})) < 1/(0.4\text{kpc}) \end{aligned} \tag{4.5}$$

We use binning method (Equation 4.4) calculate extinction contributed by the dust in the Milky Way disk and halo.

We select "target" stars in the LMC or SMC by choosing stars that fulfill **either**:

$$\varpi_{\text{est}} + 2\sigma(\varpi_{\text{est}}) < 1/(30\text{kpc}), \tag{4.6}$$

or

$$\begin{aligned} &\varpi_{\text{est}} - \sigma(\varpi_{\text{est}}) < 1/(30\text{kpc}); \\ &(\mu_\alpha - \bar{\mu}_\alpha)^2 + (\mu_\delta - \bar{\mu}_\delta)^2 < (0.8\text{mas/yr})^2. \end{aligned} \tag{4.7}$$

$\bar{\mu}_\alpha$ and $\bar{\mu}_\delta$ are the central proper motion. We use $\bar{\mu}_\alpha = 1.91\,\text{mas/yr}$ and $\bar{\mu}_\delta = 0.229\,\text{mas/yr}$ for LMC and $\bar{\mu}_\alpha = 0.772\,\text{mas/yr}$ and $\bar{\mu}_\delta = -1.117\,\text{mas/yr}$ for SMC (Kallivayalil et al., 2013). Proper motion of stars are from Gaia DR3 pmra and

`pmdec`. We calculate $R(V)$ map of LMC or SMC by the "target" stars using Equation 4.4, but using the residual $A(V)$ and $A(B)$ subtracting the foreground values.

4.8 Estimation of time scale for PAH accretion

In this section, we estimate the time scale of PAH growth by accretion. Based on Krasnokutski et al. (2017), we assume that carbon atoms that collide with the rim of PAH molecules are incorporated into the molecule. We assume the width of the rim is the spacing between carbon atoms in PAHs, $\Delta R \approx 1.4\,\text{Å}$. The number of atoms collected per unit time by a PAH molecule (assumed to be circular) is:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \frac{\int_0^\pi |\cos\theta| \sin\theta \,\mathrm{d}\theta}{\int_0^\pi \sin\theta \,\mathrm{d}\theta} \int_0^\infty \mathrm{d}v \; p(v) \; 2\pi R \Delta R n_C v = \pi R \,\Delta R \, n_C \langle v \rangle, \tag{4.8}$$

where N is the number of carbon atoms in the PAH molecule, θ is the inclination angle, R is the radius of the PAH molecule, n_C is the number density of carbon atoms in gas, v is the speed of the carbon atoms, and $p(v)$ is the probability of speed v. If the 3D velocity follows Maxwellian distribution, then we have

$$\langle v \rangle = \sqrt{\frac{8 k_B T}{\pi m}}, \tag{4.9}$$

where T is the temperature of the carbon atoms, and $m = 12\,\text{Da}$ is the atomic mass of carbon.

The radius of a compact pericondensed PAH molecule is related to number of carbon atoms by:

$$\pi R^2 \approx N \left(\Delta R\right)^2 . \tag{4.10}$$

We can use the above to formulate Equation 4.8 in terms of N (instead of R),

Typical types of PAH	Time scale of accretion to double the mass (Myr)
Naphthalene ($C_{10}H_8$)	1.28
Phenalene ($C_{13}H_{10}$)	1.46
Pyrene ($C_{16}H_{10}$)	1.61
Corannulene ($C_{20}H_{10}$)	1.80
Coronene ($C_{24}H_{12}$)	1.98

Table 4.2. Time scales for typical types of PAH to double their mass by accretion. These time scales are much shorter than the lifespan of molecular clouds ($\simeq 20\,\mathrm{Myr}$), such that growth by accretion is a viable means of explaining the variations of PAH mass fraction that give rise to the $R(V)$ variation observed in ZG24.

resulting in

$$\frac{\mathrm{d}N}{\mathrm{d}t} = (\Delta R)^2 n_C \sqrt{\frac{8 k_B T N}{m}}\,. \tag{4.11}$$

The solution to this differential equation is given by

$$N = \left((\Delta R)^2 n_C t \sqrt{\frac{2 k_B T}{m}} + \sqrt{N(t=0)} \right)^2 \tag{4.12}$$

In translucent ISM clouds, the typical gas density is $10^2\,\mathrm{cm}^{-3}$. If we assume there are 100 ppm of carbon atoms available in gas, we have $n_\mathrm{C} = 10^{-2}\,\mathrm{cm}^{-3}$. If $T = 20\,\mathrm{K}$, we find time scale for $\sqrt{N}$ to increase by 1 is:

$$\left((\Delta R)^2 n_\mathrm{C} \sqrt{\frac{2 k_B T}{m}} \right) = 1.0\,\mathrm{Myr}^{-1} \tag{4.13}$$

Various factors could affect the time scale. To name a few, sticking probability is assumed to be 100% in our estimate, but in reality should depend on factors such as binding energy, temperature, inclination, velocity and angle. Accretion efficiency also decreases with the depletion of carbon atoms in gas. Because a fraction of PAHs is ionized, Coulomb focusing could accelerate accretion (Weingartner & Draine, 1999).

Chapter 5

Conclusion

5.1 Summary

In this thesis, I began by reviewing the fundamental concepts of extinction. I discussed two important mechanisms responsible for extinction: (generalized) Mie theory and $\pi \rightarrow \pi^*$ transitions of electron. I discuss the factors that affects cross sections associated with each mechanism, which shape the observed extinction curves. Existing dust maps typically assume universal extinction curves across the Milky Way, which is equivalent to assume universal physical and chemical properties of dust. This thesis aims to address this oversimplification. I review the methodologies for determining extinction curves as well as important discoveries made by these methods, and then discuss the idea of data-driven forward modeling adopted by this thesis. I discuss the prominent features on the extinction curves from UV to IR, and their correlation with the parameter $R(V)$. Finally, I introduce the evolution of dust grains, which provides basis for the explanation of the phenomenon we have discovered.

In Chapter 2, I developed a data-driven forward model that predicts the low-resolution spectra observed by *Gaia* XP with physical parameters. I trained this

model with the stars that have both stellar parameters determined by LAMOST, and spectra observed by *Gaia* XP. We validate the result of stellar parameters with higher-resolution spectroscopic surveys. This model was then used to fit all 220 million *Gaia* XP spectra, determining their stellar types, distances and extinction properties simultaneously. As variation of extinction, parameterized by $R(V)$, is a second-order effect, we hold $R(V)$ constant in this stage. Nevertheless, we find residuals of flux (the difference between our best fitting spectra and the observed spectra) exhibit spatial variation patters that correlate closely with dust clouds, which indicate underlying variation of extinction curves.

In Chapter 3, I address the variation of extinction curves by adding an additional degree of freedom that varies the slope of extinction curves, equivalent to the effect of $R(V)$. I apply the generalized model to all XP spectra again, and precisely determined 130 million extinction curves. $R(V)$ is only meaningful for stars with significant extinction, so these results are limited to that subset. Using a binning method, I construct a 3D map of extinction curve variation across the Milky Way, and 2D maps for the Large and Small Magellanic Clouds. I find star-forming regions tend to have higher $R(V)$, or flatter extinction curves—consistent with theoretical predictions that dust grains grow larger in dense environments. Counterintuitively, $R(V)$ decreases, rather than increase, with dust density in the translucent regions of dust clouds.

In Chapter 4, I propose theoretical explanations for the anomaly identified in Chapter 3. I compared the mass needed to decrease $R(V)$ by the growth of large (silicate) grains, and found it greatly exceed what is available in the gas phase ISM. The contribution to $R(V)$ variation by changing the grain size distribution (modeled with generalized Mie theory) is negligible. On the other hand, the red wing of the 2175-Å feature contributes to B band extinction but is negligible to V band, effectively decreasing $R(V)$. As the 2175-Å[] feature arises from electronic transition, of

which the strength is proportional to the total PAH mass (independent of size distribution of PAHs), there is enough carbon in the gas phase ISM to support the growth of PAHs that decreases $R(V)$ by the observed amount. I present observation correlations of PAH mass fraction and $R(V)$ in the Milky Way and the LMC, which are consistent with the theory prediction.

This thesis reveals the variation of extinction curves from small-scale molecular clouds to large-scale Galactic structures. The observed connection between $R(V)$ and PAHs has important implications for future observations: if $R(V)$ is indeed driven by PAHs, then its spatial variation traces the growth and destruction of these molecules, placing important constraints on the carbon cycle in the ISM. Regions exhibiting significant changes in $R(V)$ are especially promising targets for follow-up observations, as they may be sites of active chemical and physical processes involving carbonaceous dust.

5.2 Future work

Various projects are in process to follow up the observed and theoretical results of this thesis.

First, the theory I have proposed is based on observational correlations between $R(V)$ and $q_{\rm PAH}$. To directly verify the connection between $R(V)$ and PAHs, I have been proposing observations in UV to resolve the 2175-Å feature, especially its red wing, with the Hubble Space Telescope (HST). Combined with optical-NIR extinction curves derived from *Gaia* XP and NIR photometric data, these observations will yield UV-through-NIR extinction curves. This allows direct measurements of the UV bump and $R(V)$ arisen from the same column of dust towards the target stars, which could potentially reduce the scattering of the trend we have found with 2D PAH mass fraction maps by emission data. Regardless of the outcome, these

observations will greatly expand the sample of high-spectral-resolution extinction curves, providing important in constraining the chemical composition of dust.

The strength of the 2175-Å feature is proportional to the total mass of PAHs, but does not constrain their size distribution. Key questions remain unresolved: Do PAHs grow via accretion of gas-phase carbon in the ISM, or through the shattering of larger carbonaceous grains? Addressing these questions requires modeling the mid-infrared PAH emission features. Therefore I am pursuing observations by JWST that resolve the PAH emission features, which provides information on the size distributions of PAHs, and constrains the in-situ growth and destruction patterns of PAHs. These processes are suggested by recent observations within molecular clouds in the Milky Way (McGuire et al., 2021) and are one of the potential PAH formation channels in high-redshift galaxies (Yang & Li, 2023).

The carriers of the broad features on extinction curves remain unclear (e.g. Green et al. (2024)). To constrain their origins, it is important to search for correlations between these features and DIBs, $R(V)$, and other structures on the extinction curve. I am pursuing observations with HST and JWST to investigate these correlations. The behavior of potential carriers in the magnetic fields may also provides key constraints. I have been conducting observation on the polarization of extinction (as a function of wavelength) towards sources with different $R(V)$ and equivalent width of those features.

Currently, three-dimensional reconstructions of gas structures in the Milky Way often rely on 3D dust extinction maps, under the assumption that extinction is proportional to dust column density. This thesis highlights the need to re-examine that relationship and to incorporate dust grain size distribution into estimates of the gas-to-dust ratio. I am proposing a high-resolution HI survey using the Five-hundred-meter Aperture Spherical Telescope (FAST). By combining HI data with a 3D dust map that accounts for extinction curve variations, we will be able to place new con-

straints on dust evolution and investigate its interactions with both the interstellar medium and stellar populations.

In this work, we have provided a preliminary map of $R(V)$ using a simple binning method, with space for improvement in many aspects. For example, spatial continuity could be imposed to better reflect the coherent structure of dust clouds. Additionally, the current mapping could be biased by selection effects: stars with higher extinction tend to have less precise parallaxes, and thus are more likely to be excluded by simple quality cuts. Incorporating corrections for such biases via Poisson point process could yield a more accurate map of $R(V)$ (Sale, 2015).

Last but not least, this work is limited by the depth of the *Gaia* XP catalog-only stars with G-band magnitudes brighter than 17.65 mag have XP spectra available. To address this limitation, I have been extending the data-driven forward models to photometric surveys, such as Pan-STARRS 1 (Hodapp et al., 2004), DECaPS (Schlafly et al., 2018) and future surveys (e.g. LSST (Ivezić et al., 2019)). In principle, these surveys have large number of bands that can constrain both stellar parameters and extinction properties. Extending the model to include photometric data will allow us to probe deeper into the Galactic plane and unveil the rich structure of the dense regions within molecular clouds.

Chapter 6

List of Publications

Here I list the publications of which I am the first author. These three publications are used in this thesis:

- **Zhang, X.**, Hensley, B., & Green, G., (2025) Dust extinction-curve variation in the translucent interstellar medium is driven by PAHs. The Astronomical Journal Letters,. 979(1), L17, 9 pp.
- **Zhang, X.** & Green, G., (2025) Three-dimensional maps of the interstellar dust extinction curve within the Milky Way galaxy. Science 387, 1209-1214(2025). DOI:10.1126/science.ado9787
- **Zhang, X.**, Green, G., & Rix. H.-W. (2023) Parameters of 220 million stars from Gaia BP/RP spectra. Monthly Notices of the Royal Astronomical Society 524, no. 2 (2023): 1855-1884.

I have also contributed to other works, but these are *not* used in this thesis:

- Green, G., **Zhang, X.**, & Zhang, R. , (2024) The Dust Extinction Curve: Beyond R(V). submitted. arXiv:2410.22537
- **Zhang, X.**, et al. (2020). OGLE-2015-BLG-1771Lb: A Microlens Planet Orbiting an Ultracool Dwarf?" The Astronomical Journal, 159(3), 116.
- Yang, H., **Zhang, X.**, et al. (2020). KMT-2016-BLG-1836Lb: A Super-

Jovian Planet from a High-cadence Microlensing Field. The Astronomical Journal, 159(3), 98.

Bibliography

Abadi, M., Agarwal, A., Barham, P., et al. 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. https://www.tensorflow.org/

Abadi, M., Barham, P., Chen, J., et al. 2016, in 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16) (Savannah, GA: USENIX Association), 265–283. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543, doi: 10.1088/0067-0049/182/2/543

Abergel, A., Ade, P. A., Aghanim, N., et al. 2014, Astronomy & Astrophysics, 571, A11

Alves, J., Zucker, C., Goodman, A. A., et al. 2020, Nature, 578, 237, doi: 10.1038/s41586-019-1874-z

Andrae, R., Rix, H.-W., & Chandra, V. 2023, arXiv e-prints, arXiv:2302.02611, doi: 10.48550/arXiv.2302.02611

Asano, R. S., Takeuchi, T. T., Hirashita, H., & Nozawa, T. 2013, Monthly Notices of the Royal Astronomical Society, 432, 637

Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, AJ, 129, 2562, doi: 10.1086/429803

Briceño-Morales, G., & Chanamé, J. 2023, Monthly Notices of the Royal Astronomical Society, 522, 1288

Buat, V., Giovannoli, E., Heinis, S., et al. 2011, A&A, 533, A93, doi: 10.1051/0004-6361/201117264

Campbell, E. K., Holz, M., Gerlich, D., & Maier, J. P. 2015, Nature, 523, 322, doi: 10.1038/nature14566

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245, doi: 10.1086/167900

Cardelli, J. A., Sembach, K. R., & Mathis, J. S. 1992, AJ, 104, 1916, doi: 10.1086/116367

Carrasco, J. M., Weiler, M., Jordi, C., et al. 2021, A&A, 652, A86, doi: 10.1051/0004-6361/202141249

Cecchi-Pestellini, C., Malloci, G., Mulas, G., Joblin, C., & Williams, D. A. 2008, A&A, 486, L25, doi: 10.1051/0004-6361:200810015

Chapman, N. L., Mundy, L. G., Lai, S.-P., & Evans, N. J. 2008, The Astrophysical Journal, 690, 496

Chastenet, J., Sandstrom, K., Chiang, I.-D., et al. 2019, ApJ, 876, 62, doi: 10.3847/1538-4357/ab16cf

Chastenet, J., Sutter, J., Sandstrom, K., et al. 2023, ApJ, 944, L12, doi: 10.3847/2041-8213/acac94

Chokshi, A., Tielens, A. G. G. M., & Hollenbach, D. 1993, ApJ, 407, 806, doi: 10.1086/172562

Clayton, G. C., Gordon, K. D., Salama, F., et al. 2003, ApJ, 592, 947, doi: 10.1086/375771

Compiègne, M., Verstraete, L., Jones, A., et al. 2011, A&A, 525, A103, doi: 10.1051/0004-6361/201015292

Conroy, C. 2010, MNRAS, 404, 247, doi: 10.1111/j.1365-2966.2010.16266.x

Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, Research in Astronomy and Astrophysics, 12, 1197, doi: 10.1088/1674-4527/12/9/003

De Angeli, F., Weiler, M., Montegriffo, P., et al. 2022, arXiv e-prints, arXiv:2206.06143. https://arxiv.org/abs/2206.06143

—. 2023, A&A, 674, A2, doi: 10.1051/0004-6361/202243680

de Bennassuti, M., Schneider, R., Valiante, R., & Salvadori, S. 2014, MNRAS, 445, 3039, doi: 10.1093/mnras/stu1962

Draine, B. T. 1989, in Infrared Spectroscopy in Astronomy, ed. E. Böhm-Vitense, 93

Draine, B. T. 2003, ARA&A, 41, 241, doi: 10.1146/annurev.astro.41.011802.094840

—. 2016, ApJ, 831, 109, doi: 10.3847/0004-637X/831/1/109

Draine, B. T., & Hensley, B. S. 2021a, ApJ, 919, 65, doi: 10.3847/1538-4357/ac0050

—. 2021b, ApJ, 909, 94, doi: 10.3847/1538-4357/abd6c6

Draine, B. T., & Lee, H. M. 1984, ApJ, 285, 89, doi: 10.1086/162480

Draine, B. T., & Li, A. 2007, ApJ, 657, 810, doi: 10.1086/511055

Draine, B. T., Li, A., Hensley, B. S., et al. 2021, ApJ, 917, 3, doi: 10.3847/1538-4357/abff51

Dwek, E., & Scalo, J. M. 1980, ApJ, 239, 193, doi: 10.1086/158100

Elíasdóttir, Á., Fynbo, J. P. U., Hjorth, J., et al. 2009, ApJ, 697, 1725, doi: 10.1088/0004-637X/697/2/1725

Fisher, R., Bowler, R. A. A., Stefanon, M., et al. 2025, MNRAS, 539, 109, doi: 10.1093/mnras/staf485

Fitzpatrick, E. L. 1999, PASP, 111, 63, doi: 10.1086/316293

Fitzpatrick, E. L., & Massa, D. 1986, ApJ, 307, 286, doi: 10.1086/164415

—. 1988, ApJ, 328, 734, doi: 10.1086/166332

—. 1990, ApJS, 72, 163, doi: 10.1086/191413

—. 2005, AJ, 130, 1127, doi: 10.1086/431900

—. 2007, ApJ, 663, 320, doi: 10.1086/518158

—. 2009, ApJ, 699, 1209, doi: 10.1088/0004-637X/699/2/1209

Fitzpatrick, E. L., Massa, D., Gordon, K. D., Bohlin, R., & Clayton, G. C. 2019, ApJ, 886, 108, doi: 10.3847/1538-4357/ab4c3a

Fritz, T. K., Gillessen, S., Dodds-Eden, K., et al. 2011, ApJ, 737, 73, doi: 10.1088/0004-637X/737/2/73

Gaia Collaboration. 2022a, 2MASS PSC XSC Best Neighbour Table,. https://doi.org/10.17876/gaia/dr.3/40

—. 2022b, 2MASS PSC XSC Join,. https://doi.org/10.17876/gaia/dr.3/42

—. 2022c, Gaia archive website,. https://gea.esac.esa.int/archive/

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016a, A&A, 595, A1, doi: 10.1051/0004-6361/201629272

—. 2016b, A&A, 595, A1, doi: 10.1051/0004-6361/201629272

Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2022, arXiv e-prints, arXiv:2208.00211. https://arxiv.org/abs/2208.00211

Gaia Collaboration, Schultheis, M., Zhao, H., et al. 2023a, A&A, 674, A40, doi: 10.1051/0004-6361/202243283

Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023b, A&A, 674, A1, doi: 10.1051/0004-6361/202243940

García-Bernete, I., Rigopoulou, D., Alonso-Herrero, A., et al. 2022, A&A, 666, L5, doi: 10.1051/0004-6361/202244806

Gavel, A., Andrae, R., Fouesneau, M., Korn, A. J., & Sordo, R. 2021, A&A, 656, A93, doi: 10.1051/0004-6361/202141589

Gordon, K. D., Cartledge, S., & Clayton, G. C. 2009, ApJ, 705, 1320, doi: 10.1088/0004-637X/705/2/1320

Gordon, K. D., Clayton, G. C., Decleir, M., et al. 2023, ApJ, 950, 86, doi: 10.3847/1538-4357/accb59

Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U., & Wolff, M. J. 2003, ApJ, 594, 279, doi: 10.1086/376774

Gordon, K. D., Fitzpatrick, E. L., Massa, D., et al. 2024, ApJ, 970, 51, doi: 10.3847/1538-4357/ad4be1

Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S., & Finkbeiner, D. 2019, ApJ, 887, 93, doi: 10.3847/1538-4357/ab5362

Green, G. M., Zhang, X., & Zhang, R. 2024, arXiv e-prints, arXiv:2410.22537, doi: 10.48550/arXiv.2410.22537

Green, G. M., Rix, H.-W., Tschesche, L., et al. 2021, ApJ, 907, 57, doi: 10.3847/1538-4357/abd1dd

Guillet, V., Fanciullo, L., Verstraete, L., et al. 2018, A&A, 610, A16, doi: 10.1051/0004-6361/201630271

Heger, M. L. 1922, Lick Observatory Bulletin, 10, 141, doi: 10.5479/ADS/bib/1922LicOB.10.141H

Hensley, B. S. 2024, arXiv e-prints, arXiv:2411.16839, doi: 10.48550/arXiv.2411.16839

Hensley, B. S., & Draine, B. T. 2021, ApJ, 906, 73, doi: 10.3847/1538-4357/abc8f1

—. 2023, ApJ, 948, 55, doi: 10.3847/1538-4357/acc4c2

Hirashita, H. 2012, Monthly Notices of the Royal Astronomical Society, 422, 1263, doi: 10.1111/j.1365-2966.2012.20702.x

Hirashita, H. 2012, MNRAS, 422, 1263, doi: 10.1111/j.1365-2966.2012.20702.x

Hirashita, H., & Kuo, T.-M. 2011, Monthly Notices of the Royal Astronomical Society, 416, 1340, doi: 10.1111/j.1365-2966.2011.19131.x

Hirashita, H., & Yan, H. 2009, Monthly Notices of the Royal Astronomical Society, 394, 1061, doi: 10.1111/j.1365-2966.2009.14405.x

Hodapp, K. W., Kaiser, N., Aussel, H., et al. 2004, Astronomische Nachrichten, 325, 636, doi: 10.1002/asna.200410300

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: 10.3847/1538-4357/ab042c

Jenkins, E. B. 2009, ApJ, 700, 1299, doi: 10.1088/0004-637X/700/2/1299

Joblin, C., Leger, A., & Martin, P. 1992, ApJ, 393, L79, doi: 10.1086/186456

Jones, A. P., Tielens, A. G. G. M., & Hollenbach, D. J. 1996, ApJ, 469, 740, doi: 10.1086/177823

Jones, B. M., Zhang, F., Kaiser, R. I., et al. 2011, Proceedings of the National Academy of Science, 108, 452, doi: 10.1073/pnas.1012468108

Kallivayalil, N., van der Marel, R. P., Besla, G., Anderson, J., & Alcock, C. 2013, The Astrophysical Journal, 764, 161, doi: 10.1088/0004-637X/764/2/161

Kingma, D. P., & Ba, J. 2014, arXiv e-prints, arXiv:1412.6980, doi: 10.48550/arXiv.1412.6980

Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv e-prints, arXiv:1711.03234. https://arxiv.org/abs/1711.03234

Krasnokutski, S. A., Huisken, F., Jäger, C., & Henning, T. 2017, ApJ, 836, 32, doi: 10.3847/1538-4357/836/1/32

Krelowski, J., Maszkowski, R., & Strobel, A. 1986, A&A, 166, 271

Kurucz, R. L. 1991, in NATO Advanced Study Institute (ASI) Series C, Vol. 341, Stellar Atmospheres - Beyond Classical Models, ed. L. Crivellari, I. Hubeny, & D. G. Hummer, 441, doi: 10.1007/978-94-011-3554-2_42

Lallement, R., Babusiaux, C., Vergely, J., et al. 2019, Astronomy & Astrophysics, 625, A135

Le Page, V., Snow, T. P., & Bierbaum, V. M. 2003, ApJ, 584, 316, doi: 10.1086/345595

Li, A., & Draine, B. T. 2001, ApJ, 554, 778, doi: 10.1086/323147

Li, J., Wong, K. W. K., Hogg, D. W., Rix, H.-W., & Chandra, V. 2024, ApJS, 272, 2, doi: 10.3847/1538-4365/ad2b4d

Lin, Q., Yang, X. J., & Li, A. 2023, MNRAS, 525, 2380, doi: 10.1093/mnras/stad2405

Lopez, L. A., Krumholz, M. R., Bolatto, A. D., et al. 2014, ApJ, 795, 121, doi: 10.1088/0004-637X/795/2/121

Mainzer, A., Grav, T., Bauer, J., et al. 2011, ApJ, 743, 156, doi: 10.1088/0004-637X/743/2/156

Mainzer, A., Bauer, J., Cutri, R. M., et al. 2014, ApJ, 792, 30, doi: 10.1088/0004-637X/792/1/30

Maíz Apellániz, J., Barbá, R. H., Caballero, J. A., Bohlin, R. C., & Fariña, C. 2021, MNRAS, 501, 2487, doi: 10.1093/mnras/staa2371

Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94, doi: 10.3847/1538-3881/aa784d

Malloci, G., Mulas, G., Cecchi-Pestellini, C., & Joblin, C. 2008, A&A, 489, 1183, doi: 10.1051/0004-6361:200810177

Malloci, G., Mulas, G., & Joblin, C. 2004, A&A, 426, 105, doi: 10.1051/0004-6361:20040541

Markov, V., Gallerani, S., Ferrara, A., et al. 2025, Nature Astronomy, 9, 458, doi: 10.1038/s41550-024-02426-1

Markov, V., Gallerani, S., Pallottini, A., et al. 2023, A&A, 679, A12, doi: 10.1051/0004-6361/202346723

Martin, P. G., & Whittet, D. C. B. 1990, ApJ, 357, 113, doi: 10.1086/168896

Massa, D., Fitzpatrick, E. L., & Gordon, K. D. 2020, ApJ, 891, 67, doi: 10.3847/1538-4357/ab6f01

Massa, D., Savage, B. D., & Fitzpatrick, E. L. 1983, ApJ, 266, 662, doi: 10.1086/160813

Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, ApJ, 217, 425, doi: 10.1086/155591

McGee, S. L., & Balogh, M. L. 2010, Monthly Notices of the Royal Astronomical Society, 405, 2069, doi: 10.1111/j.1365-2966.2010.16616.x

McGuire, B. A., Loomis, R. A., Burkhardt, A. M., et al. 2021, Science, 371, 1265, doi: 10.1126/science.abb7535

McKinnon, R., Vogelsberger, M., Torrey, P., Marinacci, F., & Kannan, R. 2018, MNRAS, 478, 2851, doi: 10.1093/mnras/sty1248

Meisner, A. M., & Finkbeiner, D. P. 2014, ApJ, 781, 5, doi: 10.1088/0004-637X/781/1/5

Meisner, A. M., Lang, D., Schlafly, E. F., & Schlegel, D. J. 2021, Research Notes of the American Astronomical Society, 5, 200, doi: 10.3847/2515-5172/ac21ca

Miret-Roig, N., Galli, P., Olivares, J., et al. 2022, Astronomy & Astrophysics, 667, A163

Miville-Deschênes, M.-A., & Lagache, G. 2005, ApJS, 157, 302, doi: 10.1086/427938

Montegriffo, P., De Angeli, F., Andrae, R., et al. 2022, arXiv e-prints, arXiv:2206.06205. https://arxiv.org/abs/2206.06205

—. 2023, A&A, 674, A3, doi: 10.1051/0004-6361/202243880

Motta, V., Mediavilla, E., Muñoz, J. A., et al. 2002, ApJ, 574, 719, doi: 10.1086/341118

Murray, N., Quataert, E., & Thompson, T. A. 2009, The Astrophysical Journal, 709, 191

Noll, S., Pierini, D., Cimatti, A., et al. 2009, A&A, 499, 69, doi: 10.1051/0004-6361/200811526

O'Donnell, J. E. 1994, ApJ, 422, 158, doi: 10.1086/173713

Ormel, C., Paszun, D., Dominik, C., & Tielens, A. 2009, Astronomy & Astrophysics, 502, 845

Ormerod, K., Witstok, J., Smit, R., et al. 2025, arXiv e-prints, arXiv:2502.21119, doi: 10.48550/arXiv.2502.21119

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Peek, J. E. G., & Schiminovich, D. 2013, ApJ, 771, 68, doi: 10.1088/0004-637X/771/1/68

Planck Collaboration, Abergel, A., Ade, P. A. R., et al. 2014a, A&A, 571, A11, doi: 10.1051/0004-6361/201323195

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014b, A&A, 571, A1, doi: 10.1051/0004-6361/201321529

—. 2014c, A&A, 571, A8, doi: 10.1051/0004-6361/201321538

Planck Collaboration, Aghanim, N., Ashdown, M., et al. 2016a, A&A, 596, A109, doi: 10.1051/0004-6361/201629022

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, A&A, 586, A132, doi: 10.1051/0004-6361/201424945

Planck Collaboration, Akrami, Y., Ashdown, M., et al. 2020a, A&A, 641, A4, doi: 10.1051/0004-6361/201833881

—. 2020b, A&A, 641, A11, doi: 10.1051/0004-6361/201832618

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020c, A&A, 641, A12, doi: 10.1051/0004-6361/201833885

Popovic, B., Brout, D., Kessler, R., & Scolnic, D. 2023, The Astrophysical Journal, 945, 84, doi: 10.3847/1538-4357/aca273

Queiroz, A. B. A., Anders, F., Chiappini, C., et al. 2023, A&A, 673, A155, doi: 10.1051/0004-6361/202245399

Ratzenböck, S., Großschedl, J. E., Alves, J., et al. 2023, Astronomy & Astrophysics, 678, A71

Rieke, G. H., & Lebofsky, M. J. 1985, ApJ, 288, 618, doi: 10.1086/162827

Roman-Duval, J., Jenkins, E. B., Tchernyshyov, K., et al. 2021, ApJ, 910, 95, doi: 10.3847/1538-4357/abdeb6

Rybizki, J., Green, G. M., Rix, H.-W., et al. 2022, MNRAS, 510, 2597, doi: 10.1093/mnras/stab3588

Sale, S. E. 2015, MNRAS, 452, 2960, doi: 10.1093/mnras/stv1459

Schild, R. E. 1977, AJ, 82, 337, doi: 10.1086/112055

Schlafly, E. F., Meisner, A. M., & Green, G. M. 2019, ApJS, 240, 30, doi: 10.3847/1538-4365/aafbea

Schlafly, E. F., Green, G., Finkbeiner, D. P., et al. 2014, ApJ, 786, 29, doi: 10.1088/0004-637X/786/1/29

Schlafly, E. F., Meisner, A. M., Stutz, A. M., et al. 2016, ApJ, 821, 78, doi: 10.3847/0004-637X/821/2/78

Schlafly, E. F., Green, G. M., Lang, D., et al. 2018, ApJS, 234, 39, doi: 10.3847/1538-4365/aaa3e2

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, The Astrophysical Journal, 500, 525, doi: 10.1086/305772

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998a, ApJ, 500, 525, doi: 10.1086/305772

—. 1998b, ApJ, 500, 525, doi: 10.1086/305772

Schödel, R., Najarro, F., Muzic, K., & Eckart, A. 2010, A&A, 511, A18, doi: 10.1051/0004-6361/200913183

Serkowski, K., Mathewson, D. S., & Ford, V. L. 1975, ApJ, 196, 261, doi: 10.1086/153410

Shivaei, I., Boogaard, L., Díaz-Santos, T., et al. 2022, Monthly Notices of the Royal Astronomical Society, 514, 1886, doi: 10.1093/mnras/stac1313

Silvia, D. W., Smith, B. D., & Shull, J. M. 2012, ApJ, 748, 12, doi: 10.1088/0004-637X/748/1/12

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163, doi: 10.1086/498708

Snow, T. P. 2014, in IAU Symposium, Vol. 297, The Diffuse Interstellar Bands, ed. J. Cami & N. L. J. Cox, 3–12, doi: 10.1017/S1743921313015512

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. 2014, Journal of Machine Learning Research, 15, 1929. http://jmlr.org/papers/v15/srivastava14a.html

Stecher, T. P. 1965, ApJ, 142, 1683, doi: 10.1086/148462

Steglich, M., Carpentier, Y., Jäger, C., et al. 2012, A&A, 540, A110, doi: 10.1051/0004-6361/201118618

Steglich, M., Jäger, C., Rouillé, G., et al. 2010, ApJ, 712, L16, doi: 10.1088/2041-8205/712/1/L16

Steinacker, J., Pagani, L., Bacmann, A., & Guieu, S. 2010, Astronomy & Astrophysics, 511, A9

Sturm, J. A., McClure, M. K., Harsono, D., et al. 2024, A&A, 689, A92, doi: 10.1051/0004-6361/202450865

Tielens, A. G. G. M. 2008, ARA&A, 46, 289, doi: 10.1146/annurev.astro.46.060407.145211

Ting, Y.-S., Conroy, C., Rix, H.-W., & Cargile, P. 2019, ApJ, 879, 69, doi: 10.3847/1538-4357/ab2331

Udalski, A. 2003, The Astrophysical Journal, 590, 284

Valencic, L. A., Clayton, G. C., & Gordon, K. D. 2004, ApJ, 616, 912, doi: 10.1086/424922

Vasiliev, E., & Baumgardt, H. 2021, MNRAS, 505, 5978, doi: 10.1093/mnras/stab1475

Voshchinnikov, N. V. 1989, Astronomische Nachrichten, 310, 265, doi: 10.1002/asna.2113100405

Walker, G. A. H., Yang, S., Fahlman, G. C., & Witt, A. N. 1980, PASP, 92, 411, doi: 10.1086/130687

Wang, S., & Chen, X. 2019, ApJ, 877, 116, doi: 10.3847/1538-4357/ab1c61

Weingartner, J. C., & Draine, B. T. 1999, ApJ, 517, 292, doi: 10.1086/307197

—. 2001, ApJ, 548, 296, doi: 10.1086/318651

Whitcomb, C. M., Smith, J.-D. T., Sandstrom, K., et al. 2024, The Astrophysical Journal, 974, 20, doi: 10.3847/1538-4357/ad66c8

Whiteoak, J. B. 1966a, ApJ, 144, 305, doi: 10.1086/148605

—. 1966b, ApJ, 144, 305, doi: 10.1086/148605

Whittet, D. C. B., van Breda, I. G., & Glass, I. S. 1976, MNRAS, 177, 625, doi: 10.1093/mnras/177.3.625

Wild, V., Charlot, S., Brinchmann, J., et al. 2011, MNRAS, 417, 1760, doi: 10.1111/j.1365-2966.2011.19367.x

Witstok, J., Shivaei, I., Smit, R., et al. 2023, Nature, 621, 267, doi: 10.1038/s41586-023-06413-w

Witten, C. E. C., Aguado, D. S., Sanders, J. L., et al. 2022, MNRAS, 516, 3254, doi: 10.1093/mnras/stac2273

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868, doi: 10.1088/0004-6256/140/6/1868

Wu, Y., Luo, A. L., Li, H.-N., et al. 2011, Research in Astronomy and Astrophysics, 11, 924, doi: 10.1088/1674-4527/11/8/006

Xiang, M., Rix, H.-W., Ting, Y.-S., et al. 2022, A&A, 662, A66, doi: 10.1051/0004-6361/202141570

Xu, Y., Bian, S. B., Reid, M. J., et al. 2018, A&A, 616, L15, doi: 10.1051/0004-6361/201833407

Yang, X., & Li, A. 2023, Nature, 621, 260, doi: 10.1038/d41586-023-02766-4

Ysard, N., Jones, A. P., Guillet, V., et al. 2024, A&A, 684, A34, doi: 10.1051/0004-6361/202348391

Yuan, H. B., Liu, X. W., & Xiang, M. S. 2013, MNRAS, 430, 2188, doi: 10.1093/mnras/stt039

Zhang, R., Yuan, H., & Chen, B. 2023a, ApJS, 269, 6, doi: 10.3847/1538-4365/acf764

Zhang, R., Yuan, H., Huang, B., et al. 2024, ApJ, 971, 127, doi: 10.3847/1538-4357/ad613e

Zhang, X., & Green, G. M. 2025, Science, 387, 1209, doi: 10.1126/science.ado9787

Zhang, X., Green, G. M., & Rix, H.-W. 2023b, MNRAS, 524, 1855, doi: 10.1093/mnras/stad1941

Zhang, X., Hensley, B. S., & Green, G. M. 2025, ApJ, 979, L17, doi: 10.3847/2041-8213/ada28f

Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, Research in Astronomy and Astrophysics, 12, 723, doi: 10.1088/1674-4527/12/7/002

Zubko, V., Dwek, E., & Arendt, R. G. 2004, ApJS, 152, 211, doi: 10.1086/382351

Zucker, C., Speagle, J. S., Schlafly, E. F., et al. 2020, Astronomy & Astrophysics, 633, A51

Zuo, W., Li, A., & Zhao, G. 2021, ApJS, 257, 63, doi: 10.3847/1538-4365/ac2cc3

Chapter 7

Acknowledgement

I would like to express my deepest gratitude to my supervisor, Dr. Gregory Green, for his guidance and support throughout the four years of my Ph.D. study. He has always been reachable, whether online or in person, generously sharing his insights — from the grand design of the universe to the smallest coding techniques. In addition to academic discussions, I enjoy our Friday hikes and the spontaneous conversations on whatever we found interesting — experiences I have come to see as another form of learning. I could not have come this far without his help. I wish him all the best in his future career, and especially, I wish his son, Euclid, grow up in good health and happiness.

I am profoundly thankful to my collaborators — Dr. Brandon Hensley and Prof. Dr. Hans-Walter Rix — for their invaluable contributions, feedback, and support throughout my research. Visiting Brandon in the summer of 2024 remains one of the best traveling memories. I also greatly enjoyed all the activities organized by Hans-Walter, including — but not limited to — the group meetings and the retreats.

I thank my friends who overlapped with me at MPIA for part or all of my time there: Taavet Kalda, Soumavo Ghosh, and Jiadong Li. Special thanks go to Maosheng Xiang for offering valuable insights into both my research and career development.

I am also grateful to Johanna Müller-Horn for her kind help with the German portion of this thesis.

I thank Prof. Ralf Klessen and Dr. Gregory Green for reviewing this thesis.

Finally, I owe my heartfelt thanks to my parents and grandparents. Despite the 6- or 7-hour time difference, their support has been my greatest source of motivation. Our video conversations have been one of the most comforting and relaxing parts of my routine.