

# Resolving Our Dusty Neighborhood in the Milky Way

Gordian Edenhofer



Munich 2024



# Resolving Our Dusty Neighborhood in the Milky Way

Gordian Edenhofer

Dissertation an der Fakultät für Physik der  
Ludwig-Maximilians-Universität München

*Dissertation at the Faculty of Physics at the  
Ludwig Maximilian University of Munich*

erstellt am  
Max-Planck-Institut für Astrophysik

*created at the  
Max Planck Institute for Astrophysics*

vorgelegt von  
*submitted by*

Gordian Edenhofer  
aus Frankfurt am Main  
*born in Frankfurt am Main*

München, 8. Mai 2024  
*Munich, 8<sup>th</sup> May 2024*

1<sup>st</sup> Referee: Assoc. Prof. Dr. Torsten Enßlin

2<sup>nd</sup> Referee: Prof. Dr. Tilman Birnstiel

Date of Doctoral Examination: 26<sup>th</sup> September 2024

# Abstract

For centuries, our view of the night sky was merely a two-dimensional image of the vast cosmos surrounding us. With new observational data from *Gaia* and significant methodological advancements, we have begun to study the structures closest to us in the Milky Way in three spatial dimensions (3D). The goal of this thesis is to reconstruct the spatial 3D distribution of clouds of matter between stars, important for the formation of stars and galactic dynamics, at high resolution in a much larger volume than ever before. To probe these clouds in 3D, we rely on measurements of interstellar dust as a tracer.

Reconstructing the density of interstellar dust in 3D in ever larger volumes requires new statistical tools to model spatially correlated densities in large volumes and new computational tools to handle the degrees of freedom encountered in the modeling. First, we develop a fast algorithm utilizing a multi-resolution view of the modeled volume. The algorithm achieves a linear time and memory complexity with the number of modeled volume elements and allows representing spatially correlated structures in volumes with hundreds of millions of elements. Next, we create a Bayesian modeling framework designed to efficiently handle hundreds of millions to billions of degrees of freedom. This framework is a rewrite of the Numerical Information Field Theory package (NIFTy) in JAX. The rewrite accelerates the runtime of typical models coded in NIFTy by two orders of magnitude.

Combining our statistical and computational tools with the most recent data from *Gaia*, we create the largest high-resolution 3D interstellar dust map to date. The map has an angular resolution of  $14'$ , a parsec-scale distance resolution, and extends out to 1.25 kpc from the Sun, more than one tenth of the distance to the Galactic center. The map represents a significant improvement over previous maps in terms of dynamic range and volume covered at high resolution.

Using the new 3D interstellar dust map, we study three prominent molecular clouds in the vicinity of the Sun: Musca, Chameleon, and Coalsack. Our 3D analysis reveals that these three famous nearby clouds, previously thought to be distinct, are part of a single coherent cloud structure, only revealed in 3D. The map and the methodologies developed offer new opportunities and a fresh perspective for future studies of our neighborhood in the Milky Way and the medium between stars.



# Zusammenfassung

Jahrhundertlang war unser Blick in den Nachthimmel lediglich ein zweidimensionales Bild des Kosmos. Neue Beobachtungsdaten von *Gaia* und bedeutende methodische Fortschritte haben es uns ermöglicht, die uns am nächsten liegenden Strukturen in der Milchstraße in drei räumlichen Dimensionen (3D) zu untersuchen. Ziel dieser Arbeit ist es, die räumliche 3D-Verteilung von Materiewolken zwischen Sternen, die für die Sternentstehung und die galaktische Dynamik wichtig sind, in einem viel größeren Volumen als je zuvor hochaufgelöst zu rekonstruieren. Um diese Wolken in 3D zu untersuchen, verwenden wir Messungen von interstellarem Staub.

Die Rekonstruktion der 3D-Dichteverteilung von interstellarem Staub in immer größeren Volumina erfordert neue Methoden zur Auflösung korrelierter Dichten und neue Modellierungsverfahren zur Handhabung der Freiheitsgrade. Zuerst entwickeln wir einen Algorithmus mit linearer Zeit- und Speicherkomplexität, der in der Lage ist, räumlich korrelierte Strukturen in Volumina mit Hunderten von Millionen Volumenelementen darzustellen. Als Nächstes entwickeln wir eine bayessche Modellierungsbibliothek, die es erlaubt, effizient mit Hunderten von Millionen bis Milliarden von Freiheitsgraden zu arbeiten. Die Bibliothek ist eine grundlegende Überarbeitung und Übersetzung des Softwarepakets “Numerical Information Field Theory” (NIFTy) nach JAX und beschleunigt typische Modelle um zwei Größenordnungen.

Mit unseren statistischen und rechnerischen Methoden und den neuesten Daten von *Gaia* erstellen wir die bisher größte hochauflösende 3D-Karte des interstellaren Staubs. Die Karte hat eine Winkelauflösung von  $14'$ , eine Entfernungsauflösung von Parsecs und reicht bis zu  $1,25$  kpc in jede Richtung von der Sonne. Die Karte stellt eine erhebliche Vergrößerung des hochaufgelösten Volumens und Verbesserung des Dynamikbereichs gegenüber früheren Karten dar.

Wir untersuchen drei bekannte Molekülwolken in der Nähe der Sonne — Musca, Chameleon und Coalsack — mit der neuen 3D-Karte. Unsere 3D-Analyse zeigt, dass diese drei berühmten sonnennahen Wolken, die bisher als voneinander unabhängig galten, Teil einer einzigen zusammenhängenden Wolkenstruktur sind, die nur in 3D sichtbar ist. Unsere 3D-Karte und die entwickelten Methoden eröffnen neue Perspektiven für zukünftige Studien unserer Nachbarschaft in der Milchstraße und des Mediums zwischen den Sternen.



# Acknowledgements

I would like to express my deepest gratitude to my advisor, *Torsten Enßlin*, for giving me incredible freedom in my pursuit of my doctoral research. I am very grateful to have been given the freedom to follow my interests in astrophysics, statistics, and computer science without constraints, at the Max Planck Institute for Astrophysics and on distant and not so distant shores. His rigorous statistical lessons, out-of-the-box thinking, and unwavering support have greatly enriched my research. I am also very grateful to have been part of Torsten’s vibrant group at the Max Planck Institute for Astrophysics. I have had the privilege of engaging in countless inspiring conversations with both current and former students. To name just a few, I would like to thank *Philipp Frank*, *Reimar Leike*, *Philipp Arras*, *Sebastian Hutschenreuter*, *Vincent Eberle*, *Matteo Guardiani*, and *Margret Westerkamp*. I would like to extend a very special thank you to *Jakob Roth*, my office mate, sparring partner for scientific ideas, and dear friend for nearly a decade. I am also thankful to *Martin Reinecke* for insightful discussions about C++, performance optimization, and numerics.

I am profoundly grateful to *João Alves* for his advice on non-scientific matters such as choosing a postdoc position and for the truly joyful discussions about science during my research stay at the University of Vienna. I would like to extend my great appreciation to *Cameron Swiggum* and *Laura Posch* for discussions on reference frames and velocity measurements, and overall for being a joy to be around, and I thank all the other members of João’s group for taking me into their midst during my stay.

My time at the Harvard–Smithsonian Center for Astrophysics was deeply influenced by *Douglas Finkbeiner*, whose breadth of knowledge in astrophysics, statistics, and computer science was truly inspiring. This thesis would not be the same without our very productive meetings. I am grateful for the welcoming atmosphere in his group and felt very much at home. I would also like to deeply thank *Catherine Zucker* for her support in all scientific and non-scientific matters throughout my doctoral research. Thanks to our long and captivating discussions, I learned a great deal about the medium in between stars, why molecular clouds are fascinating, and about astrophysics in general. *Andrew Saydjari* deserves a big thank you as well for our in-depth discussions about statistics and computer science, even though I was unsuccessful in converting him to the good side of the Python-versus-Julia debate.

I am thankful to the German Academic Scholarship Foundation (“Studiens-tiftung”) for their support in my pursuit of independent research projects through a PhD stipend and two international travel scholarships for my studies abroad. I would also like to thank *Volker Springel* for his support for my research stay at Harvard against a multitude of administrative hurdles. A big thank you also to the IT team at the Max Planck Institute for Astrophysics for their incredibly fast response times in all hardware and software matters.

A heartfelt thank you goes to my group of friends — *Jakob Roth, Haochun Ma, Lukas Nakamura, Paul Ockenfuß, and Johannes Bürger*. We grew together during intense study sessions more than 8 years ago and since then have been upholding the Bavarian culture through cooking non-Bavarian dishes, frequent visits to classical concerts, and many movie nights. I am glad to call you my close friends.

*Mia Lu* deserves a special mention for her emotional support throughout the most turbulent phase of my doctoral studies. You have truly brightened my life. I would also like to thank *my brother and sister* for their unwavering support throughout my studies and beyond, and *my parents*, whose love and support have been the foundation for where I am today.

---

I acknowledge the use of ChatGPT, DeepL, and Grammarly in the preparation of this thesis. For parts of this thesis, I used ChatGPT to translate detailed bullet points into first rough paragraphs which I then extensively reworked. I used DeepL to improve the wording throughout my thesis. Furthermore, I used Grammarly to enhance the writing and grammar of parts of my thesis.

# Contents

<b>Abstract</b>	<b>v</b>
<b>Zusammenfassung</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The interstellar medium and interstellar dust . . . . .	2
1.2 Probing the Milky Way in 3D with interstellar dust . . . . .	5
1.2.1 Measuring interstellar dust . . . . .	5
1.2.2 Measuring distances with <i>Gaia</i> . . . . .	7
1.3 Overview of the work presented in this thesis . . . . .	8
1.3.1 Statistical challenge . . . . .	9
1.3.2 Computational challenge . . . . .	10
1.3.3 Largest high-resolution 3D map of interstellar dust . . . . .	11
<b>2 Sparse kernel Gaussian processes through iterative charted refinement (ICR)</b>	<b>13</b>
2.1 Introduction . . . . .	13
2.2 Related work . . . . .	15
2.3 Background . . . . .	15
2.3.1 Gaussian processes . . . . .	15
2.3.2 Standardization . . . . .	16
2.4 Iterative Charted Refinement . . . . .	17
2.4.1 Refinement of a three pixel grid . . . . .	17
2.4.2 Larger grids . . . . .	19
2.4.3 Arbitrary points . . . . .	20
2.4.4 Larger refinement matrices and arbitrary dimensions . . . . .	21
2.4.5 Sources of error . . . . .	22
2.5 Experiments . . . . .	22
2.5.1 Accuracy . . . . .	23
2.5.2 Covariance matrix representation . . . . .	23
2.5.3 Computational speed . . . . .	24
2.6 Conclusion . . . . .	26

CONTENTS

2.7	Supplementary material . . . . .	27
2.7.1	Proof of positive definiteness of $K_{\text{ICR}}$ . . . . .	27
2.7.2	Proof of linear scaling . . . . .	28
2.7.3	Automatic charting . . . . .	29
2.7.4	Optimal sizes for the refinement matrices . . . . .	30
2.7.5	Accounting for accumulating errors . . . . .	30
2.7.6	Multivariate ICR with non-stationary kernel on irregular grids . . . . .	30
<b>3</b>	<b>Re-envisioning numerical information field theory (NIFTy.re):</b>	
	<b>A library for Gaussian processes and variational inference</b>	<b>33</b>
3.1	Statement of need . . . . .	34
3.2	Core components . . . . .	35
3.2.1	Gaussian processes . . . . .	35
3.2.2	Building up complex models . . . . .	36
3.2.3	Variational inference . . . . .	37
3.2.4	Performance of NIFTy.re compared to NIFTy . . . . .	39
3.3	Conclusion . . . . .	41
<b>4</b>	<b>A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun</b>	<b>43</b>
4.1	Introduction . . . . .	44
4.2	Stellar distance and extinction data . . . . .	45
4.3	Priors . . . . .	48
4.4	Likelihood . . . . .	49
4.4.1	Response . . . . .	50
4.4.2	Likelihood and joint probability density . . . . .	51
4.5	Posterior inference . . . . .	53
4.6	Caveats . . . . .	54
4.7	Results . . . . .	55
4.7.1	Comparison to existing 3D dust maps . . . . .	61
4.8	Conclusions . . . . .	68
4.9	Supplementary material . . . . .	69
4.9.1	ZGR23 in dust-free regions . . . . .	69
4.9.2	Metric Gaussian variational inference . . . . .	71
4.9.3	Extinction within the innermost 69 pc . . . . .	72
4.9.4	Extinction catalog . . . . .	73
4.9.5	2 kpc reconstruction . . . . .	76
4.9.6	Using the reconstruction . . . . .	80
4.9.7	Molecular clouds by distance . . . . .	80
<b>5</b>	<b>The “C”: The large Chameleon-Musca-Coalsack cloud</b>	<b>87</b>
5.1	Introduction . . . . .	87
5.2	Methods and results . . . . .	89

5.2.1	Topology . . . . .	89
5.2.2	Mass . . . . .	90
5.2.3	Dynamics . . . . .	93
5.2.4	Age . . . . .	94
5.2.5	Caveats . . . . .	94
5.3	Discussion and conclusion . . . . .	96
<b>6</b>	<b>Additional contributions</b>	<b>97</b>
6.1	JAXbind: Bind any function to JAX . . . . .	97
6.2	Galactic 3D large-scale distribution of interstellar dust . . . . .	98
6.3	Studies of the local ISM . . . . .	98
6.4	Applications outside of astrophysics . . . . .	101
<b>7</b>	<b>Conclusion</b>	<b>103</b>
7.1	Summary . . . . .	103
7.2	Outlook . . . . .	104



# List of figures

1.1	Reddening of starlight in the Bernard 68 molecular cloud . . . . .	4
1.2	Basic measurement principle underpinning distance measurements by <i>Gaia</i> . . . . .	7
2.1	1D Grid refinement . . . . .	18
2.2	Multiple grid refinements of a large grid . . . . .	21
2.3	MAE comparison of ICR and KISS-GP . . . . .	24
2.4	Comparison of ICR' and KISS-GP' covariance . . . . .	25
2.5	Performance of KISS-GP versus ICR . . . . .	26
3.1	Data and reconstruction for a synthetic example . . . . .	39
3.2	Benchmark . . . . .	40
4.1	2D histograms of the density of stars in heliocentric Galactic Cartesian (X, Y, Z) projections, as well as the density of stars as a function of distance . . . . .	47
4.2	POS distribution of the ZGR23 stars . . . . .	48
4.3	Mollweide projection of the POS integrated $A_V$ extinction . . . . .	57
4.4	Same as Figure 4.3 but showing the difference between the inte- grated extinctions in between distance slices . . . . .	57
4.5	Heliocentric Galactic Cartesian (X, Y, Z) projections of the pos- terior mean . . . . .	58
4.6	Same as Figure 4.5 but with a catalog of clusters of YSOs . . . . .	59
4.7	Heliocentric Galactic Cartesian (X, Y, Z) projections of the rela- tive uncertainty . . . . .	60
4.8	Side by side views of the 3D dust maps from Bayestar19, VLC22, L+22, and this work . . . . .	62
4.9	Zoomed-in version of Figure 4.8 for the volume reconstructed in Leike et al. (2020) . . . . .	63
4.10	Histogram of the mean posterior extinction of our map versus LGE20, VLC22, and Bayestar19 for 58 million test points . . . . .	64
4.11	Mollweide projections of total integrated extinction and 3D extinc- tion maps integrated out the maximum distance of the respective map . . . . .	66

*LIST OF FIGURES*

4.12	Zoomed-in views toward the individual molecular clouds . . . . .	67
4.13	Absolute and relative ZGR23 extinction in dust-free regions . . .	70
4.14	Mollweide projection of the integrated $A_V$ extinction in the inner- most 69 pc . . . . .	73
4.15	ZGR23's extinction versus our predicted extinction for Gaia BP/RP stars . . . . .	74
4.16	Mean standardized extinctions . . . . .	75
4.17	Similar to Figure 4.15 but for the posterior standard deviation of our extinctions . . . . .	76
4.18	Axis parallel projections of the reconstructed dust extinction . .	78
4.19	Same as Figure 4.18 but with a catalog of clusters of YSOs . . .	79
4.20	Comparison of different dust maps for Perseus . . . . .	82
4.21	Same as Figure 4.20 but for Orion A . . . . .	83
4.22	Same as Figure 4.20 but for Taurus . . . . .	84
4.23	Same as Figure 4.20 but for CrA . . . . .	85
4.24	Same as Figure 4.20 but for Chameleon . . . . .	86
5.1	Plane-of-sky region toward the “C” . . . . .	91
5.2	3D view of the isodensity surface . . . . .	92
5.3	Cartesian X-Z-projections of the “C” . . . . .	95
6.1	POS and radial data constraints for Musca . . . . .	99
6.2	3D view of the Local Bubble . . . . .	100
6.3	3D magnetic field reconstruction from synthetic data . . . . .	101

## List of tables

4.1	Parameters of the prior distributions . . . . .	52
5.1	Key properties of the “C” . . . . .	92



# 1 Introduction

The main focus of my thesis is the exploration of the Milky Way in three dimensions (3D). I am fascinated by how little we currently know about the true 3D structure of the universe surrounding us, and it is this curiosity that has driven me to delve deeply into studying our Galaxy in 3D.

The 3D structure of the Milky Way and especially the 3D structure of the medium in between stars in the Milky Way — the interstellar medium (ISM) — is crucial for our understanding of star formation and the dynamics of our Galaxy. Stars are born from molecular clouds that make up most of the non-stellar, visible matter in galaxies. Understanding the topology of these clouds in 3D is vital, as the physical processes within them are inherently three-dimensional.

Traditional methods of observing these structures in two-dimensional plane-of-sky (POS) projections often present a misleading picture. The analysis of POS projections can distort our perception of the ISM, making it challenging to discern the actual spatial relationships and boundaries of molecular clouds.

Our local neighborhood in the Milky Way provides a unique opportunity to resolve astrophysical structures in 3D. The three-dimensional reconstruction of our Galactic neighborhood, also known as Galactic tomography, hinges on measurements of interstellar dust. Despite its mundane name, interstellar dust plays a critical role as a proxy for the gaseous molecular clouds in the ISM and is currently the only quantity in the ISM we can precisely measure in 3D.

My research during my doctoral studies has been encapsulated in four works. The first focuses on the required statistical algorithm for modeling the encountered degrees of freedom for representing interstellar dust in 3D. The second discusses the computational challenges of modeling hundreds of millions to billions of parameters and proposes a new probabilistic programming framework to work with extremely high dimensional statistical structures efficiently. The third work presents a new 3D map of interstellar dust using the statistical and computational methods developed in the previous two works. In the fourth work, we analyze the 3D topology of three famous molecular clouds in our new 3D map of interstellar dust and demonstrate the importance of a 3D perspective. Finally, we discuss research projects to which I contributed that utilized the 3D dust map or parts of the developed methodology.

The introductory section of my thesis opens with a discussion on why interstellar dust is of interest. Afterwards, we discuss what constitutes interstellar

dust, and then explain the principles behind its measurement in 3D. Finally, we motivate the need for both a new statistical and a new computational framework to advance our understanding of the distribution of interstellar dust in 3D.

## 1.1 The interstellar medium and interstellar dust

The summary provided here is based on Klessen and Glover (2016) and Draine (2011) which discuss the physical processes in the ISM.

The ISM is a tenuous, gaseous mixture of mostly hydrogen and helium that fills the space in between stars. It is the medium from which stars form and to which they expel their outflows. It hosts many vastly different pressure regimes, encompassing faint gaseous clouds off the disk as well as dense molecular, star forming clouds.

The ISM is often assumed to be in thermal equilibrium (c.f. Field et al., 1969). Neglecting complicating factors such as supernova feedback and ionized matter far from the Milky Way’s midplane, the ISM has two thermally stable states: a cold, dense phase around 100 K known as the Cold Neutral Medium (CNM) and a warm, diffuse phase around  $10^4$  K known as the Warm Neutral Medium (WNM). The CNM is host to collapsing clouds and star formation, while the WNM plays an important role as a transition medium.

By mass, the ISM is approximately 70% hydrogen (H), 28% helium (He), and 2% heavier elements — often referred to as “metals” in astrophysics. Ionized gas occupies the largest volume, however, it comprises only about 25% of the mass. The majority of the mass in the ISM resides in clouds of molecular and atomic hydrogen. These dense pockets make up about 1% to 2% of the overall volume. The total mass of all the gas in the Milky Way is on the order of  $10^{10} M_{\odot}$  (Kalberla and Kerp, 2009), significantly less than the mass found in the form of stars. Yet as the seed for star formation and their primary interacting medium, the ISM is critically important for the Milky Way.

Stars form in dense molecular clouds within the CNM of the ISM. These clouds are composed of hydrogen and interstellar dust and host virtually all star formation. Interstellar dust is negligible in mass, contributing only approximately 1% of mass to the total mass of the ISM, however, it is crucial for the ISM. The formation and dissociation of molecular clouds and the triggers for star formation are dictated by many physical processes, several of which involve or are traced by interstellar dust. These processes couple many vastly different scales ranging from the chemical processes occurring on the surface of interstellar dust grains, stellar feedback, the propagation of shock waves through molecular clouds, and the dynamics of these clouds on galactic scales.

Stars form from small overdensities within molecular clouds which become dense and cold enough to gravitationally collapse. The precise mechanisms driving the collapse are yet to be fully understood, but turbulence is thought to play an important role. The timescale of collapse is roughly the same as the

timescale of dissipation, making the process highly unpredictable. Turbulence is believed to drive both star formation within dense molecular clouds on small scales while simultaneously counterbalancing gravity on larger scales hindering the whole cloud from collapsing.

The processes driving the formation and destruction of molecular clouds are similarly uncertain. Molecular clouds likely form through a combination of large-scale turbulence, leading to compression, and localized heating and cooling processes counterbalancing respectively enhancing further compression. These turbulent instabilities likely trigger the transition from warm, tenuous, predominantly atomic gas to dense, cold, molecular gas.

The sizes of molecular clouds range widely, from 2 pc to 20 pc (Cernicharo, 1991; Bergin and Tafalla, 2007). Typically, they have mean molecular hydrogen densities between  $10^2 \text{ cm}^{-3}$  to  $10^3 \text{ cm}^{-3}$ , and masses ranging from a few hundred to millions of solar masses (Cernicharo, 1991; Bergin and Tafalla, 2007). The exact processes keeping them from collapsing are yet to be fully understood but pure thermal pressure or stellar feedback are likely insufficient. Strong turbulence within the clouds, characterized by supersonic gas motions, appears to be the likely mechanism preventing rapid collapse.

Interstellar dust is a necessary ingredient in the creation of molecular clouds. It plays a vital role in facilitating the formation of molecular hydrogen  $\text{H}_2$  from atomic hydrogen by shielding its surroundings from interstellar ultraviolet (UV) radiation. Interstellar dust is the tracer of choice for molecular clouds as  $\text{H}_2$  itself radiates weakly due to the absence of a permanent dipole moment. Other common tracers of molecular clouds such as  $^{12}\text{CO}$  are easily saturated and quickly become optically thick, tracing only the cloud surfaces.

The shielding of interstellar dust is key to observing it. The absorption of interstellar dust in the UV makes the light of stars, as observed from Earth, appear less or more reddish, depending on the amount of dust along the line of sight (LOS) from us to the star. This process is called reddening. An example observation through a molecular cloud is shown in Figure 1.1. The change in color of the stars observed in this image is due to interstellar dust between us and the stars. Observationally, the reddening due to interstellar dust is closely linked with the hydrogen density; while hydrogen makes up almost all the mass of a molecular cloud, dust is crucial for allowing it to become dense and cold enough to form a cloud in the first place. The strong correlation between the amount of reddening of starlight with the hydrogen column density led to the discovery of interstellar dust.

Dust acts as an essential coolant, not only allowing for molecular hydrogen to form but also facilitating gravitational collapse within molecular clouds. This cooling occurs through collisions with particles at higher temperatures and subsequent re-emission of the absorbed kinetic energy as thermal radiation. The exact cooling rate per unit volume remains a topic of active research, but dust is known to dominate cooling processes at high densities (Hollenbach and McKee, 1989; Goldsmith, 2001).

1. Introduction

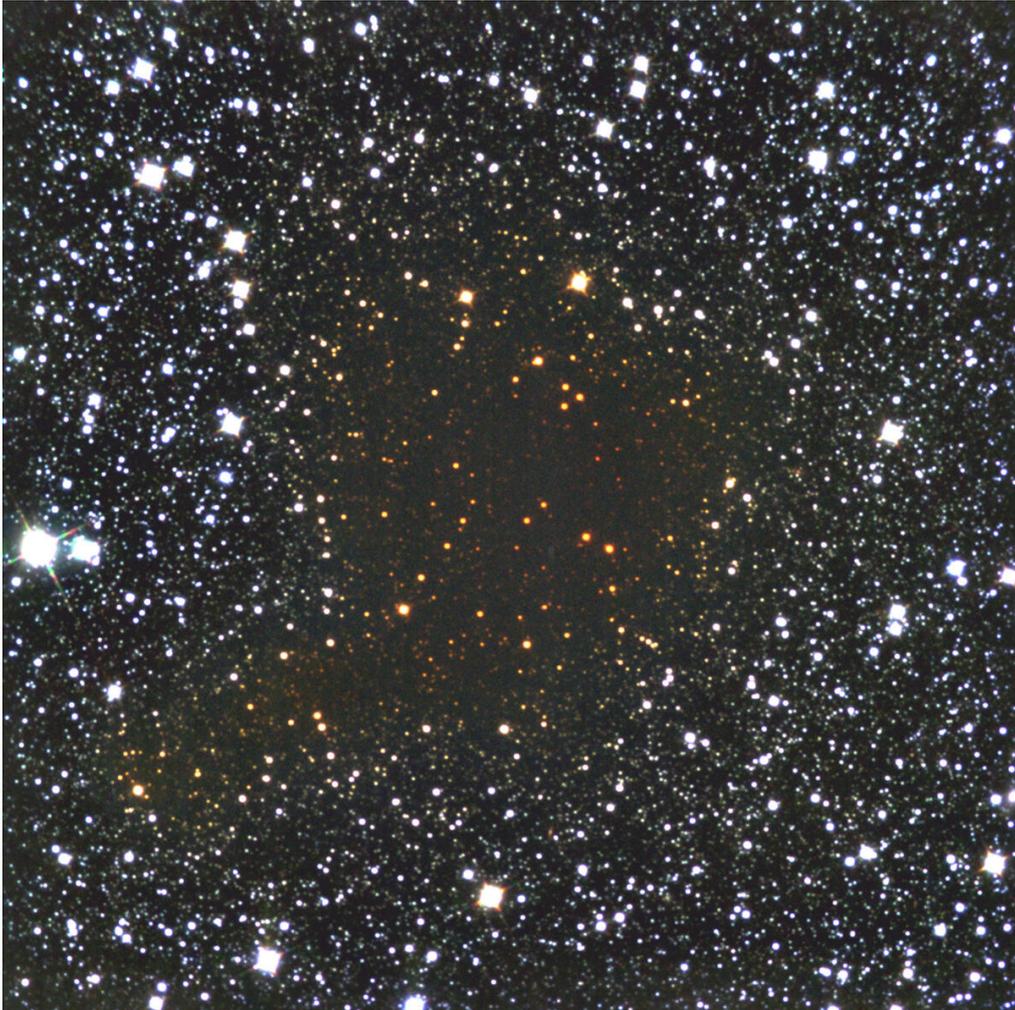


Figure 1.1: Reddening of starlight in the Bernard 68 molecular cloud in a  $4.9' \times 4.9'$  patch of the sky. The image obtained on March 8<sup>th</sup> and March 9<sup>th</sup> 1999 is a false-color, infrared composite photo of observations in three wavelength bands. It was taken with the SOFI instrument at the ESO 3.5-m New Technology Telescope (NTT) at La Silla. Figure taken from ESO (1999).

The strong absorption of interstellar dust in the UV spectrum indicates that its size is typically on the order of UV wavelengths, with less absorption in the optical and even less in the infrared. The absorption suggests a scarcity of larger grains. The size distribution of dust grains (ranging from about 50 nm to 0.25  $\mu\text{m}$ ) is characteristic of the properties of dust within the ISM. Interstellar dust is believed to be made up of graphite and silicate because of notable features in the extinction curve associated with these materials. Following Draine (2011), the total mass of interstellar dust is dominated by large grains while the total surface area, important for astrochemistry, is dominated by the sum of many small grains.

Dust primarily emits in the range from 5  $\mu\text{m}$  to 600  $\mu\text{m}$ , with two-thirds of its total power radiated in the mid and far infrared. Its emissivity roughly corresponds to that of a gray-body with a mean temperature of approximately 20 K (Planck Collaboration, 2014b)

$$B_\nu(T) \left( \frac{\nu}{\nu_0} \right)^\beta, \quad (1.1)$$

with  $B_\nu(T)$  the spectral energy density of a Planck black-body at temperature  $T$ ,  $\beta$  the spectral index, and  $\nu_0$  some reference frequency.

The simplicity of measuring dust via reddening — a process requiring minimal physical interpretation and subject to little saturation — makes it an invaluable tool in ISM studies, readily detectable even at low densities.

## 1.2 Probing the Milky Way in 3D with interstellar dust

Reconstructing the 3D distribution of interstellar dust is a complex task that hinges on utilizing tens of millions of data points, specifically measurements of the extinctions along lines of sight to stars and measurements of the distances to those stars. Inverting these measurements to retrieve a 3D map of the distribution of interstellar dust is a statistically challenging and computationally expensive tomographic inversion problem. The precise measurement processes and the uncertainties encountered therein are crucial for the final model. Let us thus start with a discussion of how extinctions and distances are measured.

### 1.2.1 Measuring interstellar dust

The study of 3D interstellar dust is intricately linked to the observation of stars. Whenever we observe a star, the light that we observe first travels through the dusty medium between us and the star. This changes the observed light of the star making it become more red than what the star originally sent out. With precise models of how a star's spectrum is supposed to look like, we can infer the amount of dust from us to a star.

## 1. Introduction

Stellar modeling can be used to propagate intrinsic stellar parameters to a star’s intrinsic spectrum. Using careful statistical modeling (Speagle et al., 2021; Paxton et al., 2011, 2013, 2015, 2018, 2019; Choi et al., 2016) we can probabilistically invert the star’s intrinsic spectrum to intrinsic stellar parameters. Loosely speaking, the absolute brightness in a specific frequency band, and the difference in brightness between two frequency bands of a star, are a good proxy for the star’s effective black-body temperature, metallicity, and surface gravity. These parameters roughly correspond to the star’s temperature, age, and mass, respectively.

However, what we observe on Earth are not the star’s intrinsic brightness values, but rather dust-extincted brightness values. The star’s intrinsic brightness is modulated by extrinsic parameters, namely the distance toward the star and the amount of dust from us to the star. The distance uniformly reduces the observed brightness across all frequencies while interstellar dust preferentially extincts blue light and the amount of total extinction depends on the amount of dust along the LOS. The precise frequency behavior of interstellar dust is described by the extinction law of interstellar dust. Typically, during stellar modeling, a fixed extinction law is assumed.

Putting both stellar models and models of the extrinsic processes modulating stellar brightness together enables us to construct models for both intrinsic and extrinsic stellar parameters. However, the extrinsic and intrinsic parameters are highly degenerate. For example, a star’s distance estimate depends strongly on the assumed intrinsic brightness while in turn a star’s temperature and brightness is strongly degenerate with the extinction.

The *Gaia* space observatory has been revolutionary in the field of stellar modeling by breaking this degeneracy. *Gaia* provides an independent measurement of a star’s distance, breaking the degeneracy at the heart of modeling extrinsic stellar parameters together with intrinsic stellar parameters. The tight constraint on a star’s distance drastically simplifies the inference of its intrinsic parameters which in turn allows us to estimate its extinction along the LOS much more precisely.

There are two primary approaches for inferring the intrinsic and extrinsic parameters of a star. The first approach utilizes analytical models of stars and probabilistically solves the inversion problem using rigorous statistical models and a lot of computing power. The second approach utilizes black-box surrogate models for the mapping from intrinsic and extrinsic stellar parameters to observed brightness values. Analytical stellar models need no training data and give accurate and well-understood uncertainties, yet they are usually computationally expensive and can have peculiar failure modes that are obscure to the average user. Surrogate models require a pristine training set (e.g. derived not only from photometric measurements but spectroscopic ones), usually provide only approximate uncertainties at best but are oftentimes computationally inexpensive and promise to exploit undiscovered or hard-to-exploit details in the data (Zhang et al., 2023). Both approaches benefit tremendously from *Gaia* dis-

tance measurements, significantly simplifying the complex disentangling of parameters influencing stellar observations. The catalog (Zhang et al., 2023) that we will use for our 3D interstellar dust map relies on the second approach.

### 1.2.2 Measuring distances with *Gaia*

Launched in 2013 by the European Space Agency (ESA), the space observatory *Gaia* is designed to cartograph our neighborhood in the Milky Way. *Gaia* aims to measure the distance to every object brighter than 20 mag, covering approximately 1% of all stars in the Milky Way. By 2020, *Gaia* had observed approximately 2 billion stars, extending down to 20.7 mag (ESA, 2022). This substantial observational reach and precision has been revolutionary for the study of our Milky Way.

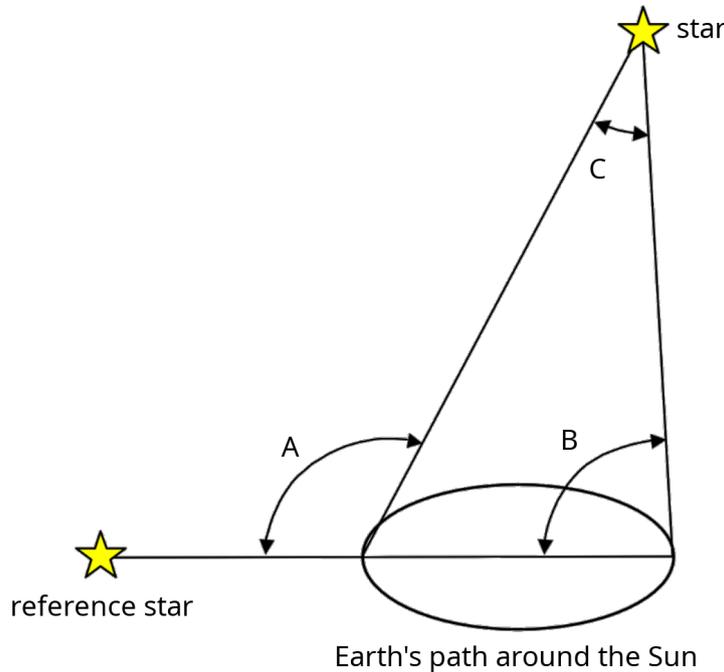


Figure 1.2: Basic measurement principle underpinning distance measurements by *Gaia*, adapted from Lindegren (2005). The angle  $C$  can be computed from the angles  $A$  and  $B$ , and the distance to the star directly follows from the distance from Earth to the Sun.

The main scientific instrument on the *Gaia* space observatory is the astrometric instrument. The astrometric instrument measures the separation of thousands of stars simultaneously. These measurements are repeated over and over while Earth and hence *Gaia* travels around the Sun. Over time, the stars will move relative to each other due to Earth's changing position in the solar system. While these differences are minute, they are yet large enough to infer the dis-

## 1. Introduction

tance to stars using geometric arguments (Lindgren, 2005). The relative shift in the angle under which we observe a star at opposite sides of the Sun relative to some fixed reference point is called parallax. A simplified sketch of the geometric measurement principle used to infer the distance is shown in Figure 1.2.

In addition to the astrometric instrument, *Gaia* hosts a blue photometer (BP) with a wavelength coverage from 330 nm to 680 nm and a red photometer (RP) with a coverage from 640 nm to 1000 nm. Both have a relatively low spectroscopic resolution ( $R = \lambda/\Delta\lambda$  30–100). The two photometers are used for various science results including for correcting the astrometric instrument as well as for stellar modeling pipelines. See de Bruijne (2009) for further details on the photometers.

*Gaia* has been transformative for the study of our Milky Way, providing a much larger catalog of stellar positions, three to four orders of magnitude larger than the largest previous catalog by *Hipparcos* (Gaia Collaboration, 2023b). In addition to measuring distances and providing valuable data for stellar modeling pipelines, *Gaia* also tracks the motion of stars on the plane of the sky, allowing the study of the full phase-space information of stars if combined with radial velocity measurements. For a small subset of approximately 10 million objects, *Gaia* also takes high-resolution spectra with its Radial Velocity Spectrometer (RVS), capable of reaching down to about 16 mag. All of this makes *Gaia* the premier observatory for studying the Milky Way.

In the early data releases, the *Gaia* collaboration did not release any RP and BP spectra. With *Gaia* EDR3, the RP and BP spectroscopy was released for a subset of stars. Although this spectroscopy is of low resolution, it has allowed for very accurate estimations of stellar properties and has significantly enhanced the quantification of extinction, making *Gaia* exceptionally effective for mapping the 3D distribution of interstellar dust. Compared to spectroscopic surveys *Gaia* has a much better coverage of the sky while achieving a comparable precision in extinction. Compared to photometric surveys *Gaia* is orders of magnitude more precise.

### 1.3 Overview of the work presented in this thesis

Reconstructing the three-dimensional distribution of interstellar dust from extinction and distance measurements to stars requires devising a statistical model of the measurement process and quantifying our prior knowledge about interstellar dust. Our model is built around two key components: the likelihood and the prior. The likelihood is defined by the measurement process, the data, and its uncertainties and the prior is defined by our knowledge about the distribution of dust before having measured any data.

Representing the likelihood is computationally extremely challenging as it combines distance and extinction data for many tens of millions of stars and needs to rigorously account for their respective uncertainties. While the extinction uncertainties are somewhat straightforward to incorporate, the distance un-

certainties significantly complicate the model. The distance uncertainties point toward one recurring challenge in the modeling of interstellar dust, namely that we have a much better POS resolution than LOS resolution. How the additional LOS uncertainty is incorporated is a key distinguishing factor between models of 3D interstellar dust.

The prior in the modeling of interstellar dust is deliberately physics agnostic, merely encoding the assumption that the interstellar dust distribution is gas-like, specifically that its 3D distribution is spatially smooth. The assumption of spatial smoothness, while innocuous sounding is essential to reconstruct interstellar dust in 3D. The importance of the spatial smoothness prior is tightly connected to the distance uncertainties in the measurements and part of the role of the prior is to counterbalance this distance uncertainty.

### 1.3.1 Statistical challenge

One key challenge in the reconstruction of interstellar dust in 3D is the so-called fingers-of-god effect. The effect denotes the preference of Galactic 3D reconstructions to align structures with the LOS. This alignment is a direct consequence of the much superior POS resolution compared to the distance resolution.

A crucial insight for the modeling of interstellar dust is to realize that dust is a physical field with a smoothly varying density. Information Field Theory (IFT) offers a rigorous mathematical framework for working with infinite-dimensional structures encountered by thinking of interstellar dust as a field. IFT enables the statistical description of smooth fields and the representation of probability spaces in infinite dimensions. Thus, IFT provides a comprehensive approach to phenomenologically model interstellar dust. We refer to Enßlin (2019, 2022, 2023) for a detailed introduction to IFT.

Modeling infinite-dimensional fields on a computer is infeasible, necessitating the discretization of our modeled space. We discretize the space into voxels and instead of reconstructing an infinite-dimensional field we content with reconstructing the amount of interstellar dust extinction in each voxel. Specifically, we model the smooth interstellar dust density field with a Gaussian Process (GP). By working in the framework of IFT, regardless of the specific choice of discretization, we can write down a continuous limit for our GP.

Modeling interstellar dust at parsec-scale resolution while covering volumes spanning kiloparsecs poses significant challenges. Naively modeling GPs on those scales is infeasible. One key trick from previous 3D interstellar dust maps (Leike and Enßlin, 2019; Leike et al., 2020) was to employ regularly spaced voxels and a stationary kernel for added efficiency (c.f. Wiener-Khinchin theorem). However, a regular grid is highly inefficient to probe volumes much beyond a couple of hundred parsecs at high resolution as a lot of computing power is wasted at the edges of the grid which can never be resolved as well as the most close by interstellar dust voxels.

Only with an uneven spacing of voxels can we probe large volumes while

## 1. Introduction

maintaining high resolution close by. Ideally, we would like to align voxels such that we have a higher resolution close by, where we have the most data and the smallest uncertainties, and a lower resolution farther out, where the data gets sparser and features higher uncertainties. In other words, we want to radially thin out the density of voxels to achieve a high resolution locally and a lower resolution farther away. The first step toward developing our new dust map, therefore, is to develop a new GP model for coupling millions to billions of unequally spaced voxels.

Chapter 2 discusses a new, highly efficient GP model. This model forms the basis of our new 3D interstellar dust map. The model leverages the geometric structure of the modeled space, adding details iteratively to a low-resolution view of the modeled volume. In each iteration, it adds detail locally. It does so at multiple resolutions, coupling even far away voxels. This approach is particularly effective for kernels whose strength diminishes at least logarithmically with distance, providing a robust foundation for accurately modeling the distribution of interstellar dust in three dimensions.

### 1.3.2 Computational challenge

Working with millions to billions of parameters such as voxels of dust densities on a computer and invoking them in statistical models is computationally expensive. Furthermore, these hundreds of millions of parameters are connected to tens of millions of data points with a statistical model. Modeling hundreds of millions of dust voxels and propagating the data constraints to these parameters makes the reconstruction of interstellar dust in 3D very computationally expensive.

Chapter 3 introduces a new Bayesian modeling and inference framework called `NIFTy.re`. This framework is a significant revision of the previous Numerical Information Field Theory package (`NIFTy`). `NIFTy.re` puts forward a much simpler modeling approach, a new concept for optimization, and makes the inference machinery much more flexible. Most importantly, the rewrite features a new backend for `NIFTy`, written in JAX (Bradbury et al., 2018). Through JAX, `NIFTy.re` can now utilize arbitrary order automatic differentiation, just-in-time compilation, automatic vectorization of the model, and seamlessly use accelerator hardware such as GPUs. The rewrite accelerates the runtime of typical models by a factor of 100 compared to the previous `NIFTy` backend.

`NIFTy.re` represents the second step toward creating a comprehensive 3D dust map. All components required for modeling 3D interstellar dust are well aligned with GPU programming paradigms, and we are able to significantly speed up the reconstruction of the 3D distribution of interstellar dust using `NIFTy.re` and JAX. The rewrite of `NIFTy`, and especially the ability to run the reconstruction on the GPU, has been critical to probing larger volumes at high resolution.

### 1.3.3 Largest high-resolution 3D map of interstellar dust

Equipped with the necessary statistical and computational tooling, we describe the reconstruction of the largest high-resolution 3D dust map to date in Chapter 4. We employ the statistical methodology outlined in Chapter 2 to model the interstellar dust efficiently in 3D space, and we use the Bayesian modeling and inference framework described in Chapter 3 to represent the more than half a billion degrees of freedom. The reconstruction relies on the most recent distance and extinction measurements from *Gaia*.

We model the positive definite density of interstellar dust with a log-normal GP. We discretize the modeled volume on a grid defined on the outer product of two subspaces, one spherical HEALPix space (Gorski et al., 1999, 2005) and one logarithmic distance space. This strategy of placing voxels results in many volume elements and thus a high resolution close-by and fewer volume elements and thus a lower resolution at larger distances where the data quality diminishes. This spacing strategy is highly uneven but can be efficiently modeled with the methodology discussed in Chapter 2.

Our reconstruction of interstellar dust in 3D is constrained by tens of millions of measurements of stars. Specifically, we use the processed *Gaia* BP/RP catalog described in Zhang et al. (2023). We carefully incorporate the distance uncertainties into the model, necessitating a highly efficient LOS integration strategy.

All computational components are written in `NIFTy.re`. We extensively use the new modeling concept put forward by `NIFTy.re` and code our model with GPU programming paradigms in mind. Our final model fully embraces JAX and `NIFTy.re` and runs highly efficiently on the GPU.

Our new map resolves faint molecular clouds at high resolution, providing a new, much larger, high-resolution picture of the ISM surrounding us. Thanks to the new *Gaia* data, the reconstruction also achieves a higher dynamic range compared to previous reconstructions in the literature. The larger volume and the higher dynamic range are instrumental for the study of molecular clouds in our vicinity.

In Chapter 5, we showcase the new map and analyze three famous molecular clouds in 3D: Musca, Chameleon, and Coalsack. We find that the three molecular clouds are part of a large faint cloud. We analyze the mass and geometry of the large cloud using our 3D dust map. By further integrating additional information on Young Stellar Object (YSO) clusters into our analysis, we are able to determine the motion of gas within the cloud. The 3D topology and motion suggest a single origin story for all three clouds. This analysis hinges on a 3D perspective, underscoring the vital importance of 3D for understanding the complexities of structures in the ISM.

Chapter 6 discusses further contributions made during my doctoral studies. This includes the co-development of `JAXbind`, a software package for bridging code written in languages like C, C++, or Julia to JAX and by extension to `NIFTy.re`. Furthermore, we discuss research projects to which I contributed

## *1. Introduction*

which use 3D interstellar dust to study molecular clouds, as well as research projects showcasing the broader application of our modeling methods within and beyond astrophysics.

Finally, we summarize the discussed work and provide an outlook for future work in Chapter 7.

## 2 Sparse kernel Gaussian processes through iterative charted refinement (ICR)

*The following chapter is an extended version of an unpublished article led by me (Edenhofer et al., 2022). The project emerged from a close collaboration between Reimar Leike and me. Reimar Leike had the initial idea and wrote parts of the manuscript. I developed the idea, conducted the experiments, and wrote parts of the manuscript. Philipp Frank and Torsten Enßlin read, commented, and approved the manuscript. The text was adapted to fit the thesis.*

### Abstract

Gaussian processes are highly expressive, probabilistic models. A major limitation is their computational complexity. Naively, exact GP inference requires  $\mathcal{O}(N^3)$  computations with  $N$  denoting the number of data-points (modeled points). Current approaches to overcome this limitation either rely on sparse, structured, or stochastic representations of data or kernel, respectively, and usually involve nested optimizations to evaluate a GP. We present a new, generative method named Iterative Charted Refinement (ICR) to perform inference with GPs on nearly arbitrarily spaced points in  $\mathcal{O}(N)$  time without nested optimizations. ICR represents long- as well as short-range correlations by combining views of the modeled points at multiple levels of coarsening. In our experiment with points whose spacing varies over three orders of magnitude, ICR’s accuracy is comparable to state-of-the-art GP methods. For points spanning more than three orders of magnitude in their spacing, ICR is significantly more accurate than competing GP methods. ICR outperforms existing methods in terms of computational speed by two orders of magnitude on the CPU and GPU.

### 2.1 Introduction

GPs are flexible function approximators. Their capacity to learn rich statistical representations scales with the amount of data they are provided with. The

## 2. Sparse kernel Gaussian processes through iterative charted refinement (ICR)

statistical structure in the data is learned via the kernel. It relates any two points in data-space and can be used to inter- and extrapolate. Through their kernel, GPs are capable of representing intricate structures in large datasets and yield reliable uncertainty estimates for their predictions.

The applicability of GPs, however, is limited by their scaling. Naively evaluating a GP requires  $\mathcal{O}(N^3)$  computations where  $N$  is the number of data-points (modeled points). The classical approach to GP inference requires applying the inverse of the kernel matrix (matrix representation of the kernel) and computing its log-determinant. These computations are often carried out via nested optimizations within each evaluation of the GP. Commonly, Krylov subspace methods are used for this purpose (Gardner et al., 2018; Pleiss et al., 2018, 2020; Wang et al., 2019). These methods are usually terminated after relatively few iterations, well before their theoretically guaranteed convergence. In practice, this reduces the computational complexity of evaluating a GP to  $\mathcal{O}(n_{\text{solve}}N^2)$ , with  $n_{\text{solve}}$  the number of matrix-vector-multiplications (MVMs) of the kernel matrix, respectively, the number of iterations of the Krylov subspace method.

Several approaches exist to reduce the quadratic computational complexity for applying the kernel matrix. They require either sparse kernels, structured kernels, regularly spaced modeled points, a set of inducing points (Liu et al., 2020), or a mixture of these. Inducing point methods are popular because they do not require a special structure in the data or kernel. The number of inducing points defines the method’s ability to resolve structures. Applying a kernel matrix with  $M$  inducing points naively scales with  $\mathcal{O}(M^2+NM)$ . It is desirable to choose  $M \propto N$  which, however, renders their application impractical for large datasets with many modeled points. Structured Kernel Interpolation (SKI), described in Wilson and Nickisch (2015), combines the advantages of regular grids and inducing points. The authors use SKI to achieve  $\mathcal{O}(N \log N)$  computational complexity for applying a kernel matrix with  $M = N$  inducing points.

The application of the inverse as well as computation of the log-determinant of the kernel matrix remains computationally expensive within all of these approaches. Furthermore, inducing point methods often yield singular kernel matrices, becoming non-singular only under projections or in combination with a non-singular matrix. With the goal of resolving the computational constraints and guaranteeing a positive definite kernel matrix, we reformulate GP inference as generative process. Without loss of generality, we shift the complexity of applying the inverse and evaluating the log-determinant of the kernel matrix to applying the “square-root” of it.

In this paper we devise an efficient algorithm to apply an approximate square-root of the kernel matrix. We provide: (I) A generative approach to GPs which avoids inverting and taking the log-determinant of the kernel matrix. (II) Iterative Charted Refinement (ICR), an algorithm with  $\mathcal{O}(N)$  computational complexity and a guaranteed positive definite kernel matrix that requires no nested optimization to evaluate a GP on nearly arbitrarily spaced points. Our open-source implementation (BSD 2-Clause license) of ICR is available as part of

NIFTy.re, described in Chapter 3. It is written in Python and uses JAX (Bradbury et al., 2018) to just-in-time compile code for the CPU and GPU.

## 2.2 Related work

There are a wide variety of approaches to modeling GPs among which are full (Wang et al., 2019), sparse, structured sparse (Wilson et al., 2014; Wilson and Nickisch, 2015; Wilson et al., 2015; Snelson and Ghahramani, 2005; Rossi et al., 2021; Gardner et al., 2018; Pleiss et al., 2018), and stochastic (Hensman et al., 2013; Wilson et al., 2016) approaches. We refer to the review in Liu et al. (2020) for a comprehensive discussion. In the following, we would like to highlight a few notable developments which focus on the non-stochastic application of GPs to big data without strong grid or kernel constraints.

Without additional constraints on the kernel or the points, GPs require at least  $\mathcal{O}(n_{\text{solve}}N^2)$  computations as they invoke  $n_{\text{solve}}$  MVMs of the kernel matrix to apply its inverse. The required application of the inverse kernel matrix for evaluating a GP can be done efficiently with Krylov subspace methods, in particular preconditioned conjugate gradient and Lanczos methods (Gardner et al., 2018; Pleiss et al., 2018, 2020). Furthermore, the computation can be distributed with minimal communication overhead (Wang et al., 2019). However, neither gain in efficiency resolves the quadratic computational scaling inherent from applying and representing the kernel matrix.

Inducing point methods bypass the quadratic scaling with the number of modeled points by approximating the true kernel matrix  $K_{XX}$  with the kernel matrix of the inducing points  $K_{UU}$ . Either much fewer inducing points can be used than modeled points or their spacing can be chosen such that applying  $K_{UU}$  becomes very efficient. With KISS-GP (Wilson and Nickisch, 2015), the authors propose a method with regularly spaced inducing points and Toeplitz and Kronecker structures in  $K_{UU}$  (Wilson and Nickisch, 2015). The kernel matrix of the inducing points is mapped to the kernel matrix of the modeled points via a sparse interpolation matrix. In general, the kernel matrix is singular and is not a proper covariance matrix. It only becomes non-singular in combination with projections (e.g. via a preconditioner) or additive corrections (e.g. diagonal jitter). Applying the KISS-GP kernel matrix requires  $\mathcal{O}(N \log N)$  computations for points with a single dimension,  $M = N$  inducing points as suggested by the authors and a Toeplitz structure in the kernel matrix of the inducing points.

## 2.3 Background

### 2.3.1 Gaussian processes

GPs are stochastic processes for which every finite marginal is a multivariate normal distribution. A GP is completely determined by its mean function  $\mu(x)$

## 2. Sparse kernel Gaussian processes through iterative charted refinement (ICR)

acting on points  $x$  and kernel function  $k(x, x')$  correlating any two points  $x$  and  $x'$ . We denote the space of all possible  $x$  with  $\Omega_x$ . For a finite set of points  $X$  we denote the kernel matrix by  $K_{XX}$ , the mean at  $X$  by  $\mu(X)$  and the corresponding multivariate normal distribution by  $\mathcal{N}(\mu(X), K_{XX})$ . We denote vectors drawn from  $\mathcal{N}(\mu(X), K_{XX})$  by  $s$ . Without loss of generality we assume the mean of our GP to be zero, i.e.  $\mu(x) = 0$ , respectively,  $\mu(X) = 0$ .

The kernel is commonly chosen from a set of well known kernel functions. Popular kernels include the Radial Basis Function (RBF) and the Matérn kernel. These kernels have various parameters  $\theta$  that are to be optimized during inference with a GP. The kernel matrix  $K_{XX}$  depends on the parameters of the kernel  $\theta$ . For brevity of notation, we do not make the dependence explicit.

### 2.3.2 Standardization

Inference with a GP is often formulated in terms of a likelihood and a Gaussian process prior

$$\log p((s, \theta)^T, y) = \log p(y|s) - \frac{1}{2} [\log |2\pi K_{XX}| + s^T K_{XX}^{-1} s] + \log p(\theta) \quad (2.1)$$

with the likelihood  $p(y|s)$  penalizing the level of agreement between some noisy data  $y$  and  $s$  a realization of the GP prior. The quantity of interest is the posterior  $p((s, \theta)^T | y) \propto p((s, \theta)^T, y)$ . While closed form solutions exist for the case with a Gaussian likelihood, none exist for the general inference problem with a generic likelihood. Thus, to perform an inference, the posterior is approximated. A popular choice is to use variational inference (VI; Morningstar et al., 2021; Knollmüller and Enßlin, 2019; Frank et al., 2021).

Naively, to evaluate Equation (2.1), it is necessary to apply the inverse of the kernel as well as compute its log-determinant. By change of variables, these computations can be avoided. A generative approach provides one such way. It allows us to express the random variable  $s$  in terms of a random variable  $\xi$  which follows a “simpler” distribution. A common choice for the distribution of  $\xi$  is the standard normal distribution. The act of expressing  $s$  in terms of a standard normally distributed variable  $\xi$  is called standardization and is possible under very mild regularity conditions. We refer to Rezende and Mohamed (2015) for a more detailed discussion on this subject.

Without loss of generality, we can absorb  $K_{XX}$  into the mapping  $s : \xi, \theta \mapsto s(\xi, \theta)$  and rewrite Equation (2.1) to

$$\log p((\xi, \theta)^T, y) = \log p(y|s(\xi, \theta)) - \frac{1}{2} [\log |2\pi \mathbb{1}| + \xi^T \xi] + \log p(\theta) . \quad (2.2)$$

The inference problems from Equations (2.1) and (2.2) are equivalent, but Equation (2.2) involves neither an inversion nor a log-determinant of the kernel if  $s(\xi, \theta)$  involves none. Loosely speaking, the mapping  $s(\xi, \theta)$  applies the square-root of the kernel matrix  $\sqrt{K_{XX}}$  to correlate the standard Gaussian random vari-

ables  $\xi$  via  $s(\xi, \theta) = \sqrt{K_{XX}} \cdot \xi$ . The term  $\sqrt{K_{XX}}$  denotes an implicit linear operator that correlates the values of  $\xi$  such that  $\mathbb{E}_{p(\xi)} \left[ (\sqrt{K_{XX}} \cdot \xi) (\sqrt{K_{XX}} \cdot \xi)^T \right] = \sqrt{K_{XX}} \mathbb{E}_{p(\xi)} [\xi \xi^T] \sqrt{K_{XX}}^T = K_{XX}$  with  $p(\xi)$  a standard normal distribution over  $\xi$ .

The square-root of the kernel is not uniquely defined. Any (potentially non-square) matrix  $\sqrt{K_{XX}}$  which fulfills  $\sqrt{K_{XX}} \sqrt{K_{XX}}^T = K_{XX}$  would suffice. Note, naively applying some form of square-root of the matrix, e.g. via the Cholesky decomposition, is computationally as expensive as applying its inverse.

## 2.4 Iterative Charted Refinement

In the following, we devise an algorithm with  $\mathcal{O}(N)$  complexity to efficiently apply an approximate  $\sqrt{K_{XX}}$  for smoothly varying kernels. We denote our approximation by  $\sqrt{K_{ICR}}$ . Applied to standard normally distributed variables  $\xi$ , our approximation yields  $\mathbb{E}_{p(\xi)} \left[ (\sqrt{K_{ICR}}(\xi)) (\sqrt{K_{ICR}}(\xi))^T \right] \approx K_{XX}$ . By construction, our kernel matrix defined by

$$K_{ICR} = \sqrt{K_{ICR}} \sqrt{K_{ICR}}^T \quad (2.3)$$

is positive definite and the resulting process is a GP with covariance  $K_{ICR} \approx K_{XX}$ , see Section 2.7.1.

$\sqrt{K_{ICR}}$  can be used for both kernel learning and sampling. Our algorithm can be used for the former by optimizing Equation (2.2) for  $\theta$  and for the latter by applying it – for a given  $\theta$  – to standard normally distributed numbers.

The idea at the core of our algorithm is to model the desired GP at various coarsening levels of the modeled points simultaneously. Our algorithm starts off at a coarse view of our modeled points and iteratively adds finer structures until we arrive at the desired modeled points. Each view of the GP at a given coarsening level is also a view on correlations at a different length scale.

### 2.4.1 Refinement of a three pixel grid

To derive our method, we constrain ourselves to a 3 pixel sub-grid of a regular grid. In this example, we would like to refine the central pixel of this sub-grid. See Figure 2.1 for an illustration. We interpret the pixel values as realizations of a GP at their respective centers (lower row of Figure 2.1). The central property of a GP states that the 3+2 coarse and fine pixels are jointly multivariate normally distributed:

$$\begin{pmatrix} s^{(1)} \\ s^{(0)} \end{pmatrix} \sim \mathcal{N} \left( 0, \begin{pmatrix} K_{ff} & K_{fc} \\ K_{cf} & K_{cc} \end{pmatrix} \right), \quad (2.4)$$

with  $s^{(1)}$  the realization of the GP on the two fine grid coordinates  $x^{(1)}$  and  $s^{(0)}$  the realization of the GP at the three coarse grid positions  $x^{(0)}$ . The entries of

## 2. Sparse kernel Gaussian processes through iterative charted refinement (ICR)

the covariance are given by

$$\begin{aligned} K_{\text{ff}} &= \begin{pmatrix} k(x_0^{(1)}, x_0^{(1)}) & k(x_0^{(1)}, x_1^{(1)}) \\ k(x_1^{(1)}, x_0^{(1)}) & k(x_1^{(1)}, x_1^{(1)}) \end{pmatrix}, \\ K_{\text{fc}} &= \begin{pmatrix} k(x_0^{(1)}, x_0^{(0)}) & k(x_0^{(1)}, x_1^{(0)}) & k(x_0^{(1)}, x_2^{(0)}) \\ k(x_1^{(1)}, x_0^{(0)}) & k(x_1^{(1)}, x_1^{(0)}) & k(x_1^{(1)}, x_2^{(0)}) \end{pmatrix}, \\ K_{\text{cc}} &= \begin{pmatrix} k(x_0^{(0)}, x_0^{(0)}) & k(x_0^{(0)}, x_1^{(0)}) & k(x_0^{(0)}, x_2^{(0)}) \\ k(x_1^{(0)}, x_0^{(0)}) & k(x_1^{(0)}, x_1^{(0)}) & k(x_1^{(0)}, x_2^{(0)}) \\ k(x_2^{(0)}, x_0^{(0)}) & k(x_2^{(0)}, x_1^{(0)}) & k(x_2^{(0)}, x_2^{(0)}) \end{pmatrix}, \end{aligned} \quad (2.5)$$

with  $K_{\text{fc}} = K_{\text{cf}}^T$ .

Given the realization of the GP at the coarse grid points  $x^{(0)}$ , the conditional distribution of the fine grid points  $x^{(1)}$  is given by

$$p(s^{(1)}|s^{(0)}) = \mathcal{N}(s^{(1)}|Rs^{(0)}, D), \quad (2.6)$$

$$\text{with } R = K_{\text{fc}} \cdot K_{\text{cc}}^{-1} \quad (2.7)$$

$$\text{and } D = K_{\text{ff}} - K_{\text{fc}} \cdot K_{\text{cc}}^{-1} \cdot K_{\text{cf}}. \quad (2.8)$$

One can reformulate this into a generative process of the fine grid realizations  $s^{(1)}$  given the coarse grid realizations  $s^{(0)}$ :

$$s^{(1)} = R \cdot s^{(0)} + \sqrt{D} \cdot \xi^{(1)} \quad (2.9)$$

where  $\sqrt{D}$  is the Cholesky decomposition of  $D$  and  $\xi^{(1)}$  is a two-dimensional independent standard normally distributed vector. We call  $R$  and  $\sqrt{D}$  the refinement matrices. The generative formulation for  $s^{(1)}$  using the refinement matrices is linear in  $(s^{(0)}, \xi^{(1)})$ .

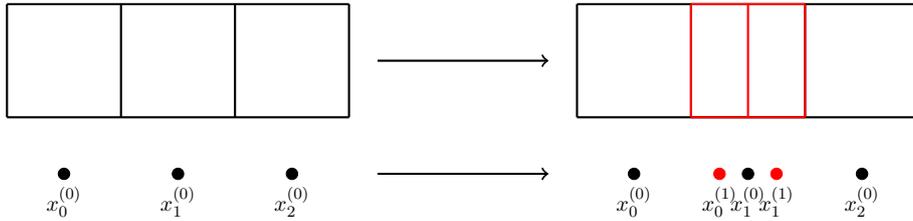


Figure 2.1: Refinement of a 3-pixel grid. The upper row shows the grid refinement in terms of pixels, the lower row shows the grid refinement in terms of pixel centers. For GPs defined on these coordinates, one can analytically calculate the conditional distribution of the fine grid (red dots) given the coarse grid (black dots).

### 2.4.2 Larger grids

For arbitrarily large grids, we take each pixel of the coarse grid together with its two neighbors and apply the procedure outlined in Section 2.4.1 to refine it. Thus, for a coarse grid of size  $N^{(0)}$  we refine pixels in overlapping subsets of 3 coarse pixels to 2 fine pixels and retrieve a new fine grid of size  $N^{(1)} = 2(N^{(0)} - 2)$ . The grid shrinks by two coarse pixels because the outermost two pixels of the coarse grid do not have sufficiently many neighbors.

Our scheme is an approximation because instead of conditioning the fine pixels on all coarse pixels we factorize the fine pixels into pairs of 2 pixels and condition only on the nearest 3 coarse pixels. The approximation can be understood in terms of probability distributions

$$\begin{aligned}
 p((s_i^{(1)})_{i=0\dots N^{(1)}-1} | (s_i^{(0)})_{i=0\dots N^{(1)}/2+1}) &\approx p_{\text{approx}}(s^{(1)} | s^{(0)}) \quad \text{with} \\
 p_{\text{approx}} &= p((s_i^{(1)})_{i=0,1} | (s_i^{(0)})_{i=0,1,2}) \\
 &\cdot \dots \\
 &\cdot p((s_i^{(1)})_{i=N^{(1)}-1, N^{(1)}} | (s_i^{(0)})_{i=N^{(1)}/2-1, N^{(1)}/2, N^{(1)}/2+1}).
 \end{aligned} \tag{2.10}$$

There are two ways in which we loose information. First, factoring

$$p((s_i^{(1)})_{i=0\dots N^{(1)}} | (s_i^{(0)})_{i=0\dots N^{(1)}/2+1})$$

with respect to pairs of points  $(s_i^{(1)})_{i=2j, 2j+1}$  neglects correlations between points on the fine grid. Second, information about the fine grid points' dependence on coarse grid points at larger distances is neglected. Both approximations are expected to be less severe if there are no systematic long range correlations when conditioning on a local subset of the coarse grid. This is the case for kernels that correlate closer points more strongly than points that are further away.

Our approximative way to calculate the fine grid points has computational advantages. It enables calculating all pairs of fine grid points independently in the same way in parallel. On a regular grid, if the kernel  $k$  is stationary, i.e.  $k(x, x') = k(x - x')$ , then we can interpret the refinement for multiple pixels following Equation (2.9) as a convolution, with  $R_{oj}$  being the convolution kernel of size  $2 \times 3$ , such that

$$r_{2i+o} = \sum_{j=0}^2 R_{oj} s_{i+j}^{(0)} \tag{2.11}$$

$$s_{2i+o}^{(1)} = r_{2i+o} + \sum_{j=0}^1 \sqrt{D_{oj}} \xi_{2i+j}^{(1)} \tag{2.12}$$

with  $r$  being the filtered coarse grid,  $o \in \{0, 1\}$  and  $\xi^{(1)}$  an independent standard normally distributed vector. Equations (2.11) and (2.12) show similarities

## 2. Sparse kernel Gaussian processes through iterative charted refinement (ICR)

to multigrid approaches, which are used for simulations (Heath, 2018; Wesseling, 2004; Bakhvalov, 1966). The values of the fine grid are constructed through smoothing the values of the coarse grid with  $R$  (Equation (2.11)) and then applying a correction (Equation (2.12)).

The approximate probability distribution  $p_{\text{approx}}(s^{(1)}|s^{(0)})$  allows us to draw a sample of  $s^{(1)}$  given the realization  $s^{(0)}$  on the coarse grid. One can iterate this procedure, using the fine grid as a coarse grid to an even finer grid. This gives rise to a generative process, depicted in Figure 2.2a.

The overall cost of generating a sample when performing iterative grid refinement is  $\mathcal{O}(N)$ , where  $N$  is the number of pixels of the final refinement. The crucial ingredient to achieve a linear scaling is the doubling of pixels within a refinement. Let  $n_{\text{lvl}}$  denote the number of levels it takes our algorithm to go from  $N^{(0)}$  points at the coarse-most grid to  $N$  points at the final grid, then the total number of computations in terms of  $N^{(0)}$  reads

$$N^{(0)} + 2(N^{(0)} - 2) + 2(2(N^{(0)} - 2) - 2) + \underbrace{\quad}_{n_{\text{lvl}} - 2 \text{ additional terms}} \cdot \quad (2.13)$$

For  $N^{(0)} > 4$  this is in  $\mathcal{O}(2^{n_{\text{lvl}}} \cdot N^{(0)}) = \mathcal{O}(N)$ . One can start this process from an arbitrarily coarse grid for which the covariance matrix can be factorized explicitly at negligible computational cost. For a full proof of the linear scaling, see Section 2.7.2.

### 2.4.3 Arbitrary points

To generalize the refinement to nearly arbitrarily spaced points, we need to define what a coarse, respectively, fine point in the space of modeled points is. To retrieve such a definition, it is convenient to employ the concept of a coordinate chart from topology. A coordinate chart is a homeomorphism from an open subset  $\Omega_x$  of a manifold, i.e. the space on which our GP is defined (see Section 2.3.1), to an open subset of a Euclidean space. Every point in  $\Omega_x$  is associated with a unique point in Euclidean space via a coordinate chart.

ICR requires a coordinate chart together with the modeled points. For many data analysis problems, the modeled points are spaced in a very specific way given by the setup of the experiment and retrieving a chart is straightforward, e.g. a pixel detector measuring energies might have a regular, spatial pixel axis and a logarithmic, spectral energy axis. For problems in which the choice of chart is non-obvious even with expert knowledge, it can be constructed solely from the modeled points, see Section 2.7.3.

We can split off the modeled points from the actual refinement via the coordinate chart. The modeled points enter the refinement only via  $R$  and  $\sqrt{D}$ , respectively via Equation (2.5). Conceptually ICR carries out the refinement on a regular grid with points  $\tilde{x}$  but for the purpose of computing  $R$  and  $\sqrt{D}$ , each regularly spaced point is associated with a point in the space of the modeled points  $\Omega_x$ , see Figures 2.2a and 2.2b. The mapping to the space of the modeled

points is performed right before evaluating Equation (2.5). This amendment to the refinement algorithm can be interpreted as an update to the kernel. Let  $\phi^{-1}$  denote a coordinate chart from a Euclidean grid to  $\Omega_x$ , then our new kernel reads  $\tilde{k} : (\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}$  with  $\tilde{k} : (\tilde{x}, \tilde{x}') \mapsto k(\phi^{-1}(\tilde{x}), \phi^{-1}(\tilde{x}'))$ . With this minimally invasive update, our discussion in Sections 2.4.1 and 2.4.2 adapts to nearly arbitrarily spaced points.

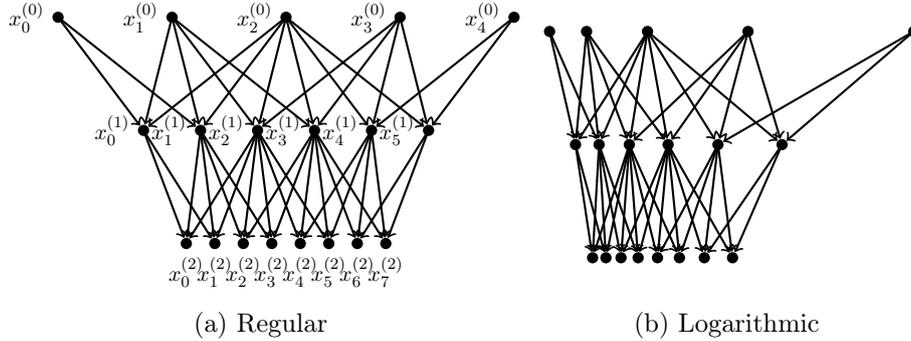


Figure 2.2: Multiple grid refinements of a large grid. Panel (a): Pixel centers on a regular grid within two refinement steps. The fine grid realizations are calculated on the basis of the three nearest, coarse grid realizations. Arrows indicate which coarse grid values inform which fine grid values. Panel (b): Refinement from Panel (a) translated with a logarithmic coordinate chart.

The refinement matrices  $R$  and  $\sqrt{D}$  need to be computed separately for every point on the irregularly spaced grid. Furthermore, if the parameters of the kernel change, the refinement matrices must be recomputed. The computational cost associated with constructing the refinement matrices for all pixels scales with  $\mathcal{O}(N)$ .

#### 2.4.4 Larger refinement matrices and arbitrary dimensions

How much information we neglect at a refinement level depends on how many adjacent pixels we use to construct the refinement matrices. In Sections 2.4.1 and 2.4.2 we refined 3 coarse pixels to 2 fine pixels. The number of coarse pixels, the number of fine pixels as well as the position of the fine pixels on our Euclidean grid can be tuned. The optimal setting for these parameters depends on the kernel and the coordinate chart. In our open-source implementation which is available as part of `NIFTy.re`, described in Chapter 3, we provide helper utilities to retrieve the information theoretical optimal settings for a given kernel and chart, see Section 2.7.4.

With more coarse and fine grid points in a refinement, the construction of the refinement matrices gets more expensive. It requires  $\mathcal{O}\left(\max\{n_{\text{csz}}, n_{\text{fsz}}\}^3 \cdot N\right)$  computations where  $n_{\text{csz}}$  denotes the number of coarse pixels (first axis of  $R$  from Equation (2.7)) which are refined to  $n_{\text{fsz}}$  fine pixels (second axis of  $R$  from

## 2. Sparse kernel Gaussian processes through iterative charted refinement (ICR)

Equation (2.7)). The power of 3 stems from the explicit inversion and Cholesky decomposition in Equations (2.7) to (2.9).

In  $d$  dimensions ICR uses  $n_{\text{csz}}^d$  neighbors and refines the center-most pixel to  $n_{\text{fsz}}^d$  pixels. It requires  $\mathcal{O}\left(\max\{n_{\text{csz}}, n_{\text{fsz}}\}^{3d} \cdot N\right)$  computations. Section 2.7.6 summarizes the most general algorithm with arbitrary  $n_{\text{csz}}$  and  $n_{\text{fsz}}$  in  $d$  dimensions.

If the kernel factorizes along certain dimensions, the computational complexity can be significantly reduced. Likewise, for rotationally or translationally invariant axes within  $\Omega_x$  and stationary kernels, the refinement matrices need only be computed once and can be broadcasted along these axes akin to Equations (2.11) and (2.12). For commonly used coordinate systems like polar or spherical coordinates, significant performance improvements can be obtained by utilizing these symmetries.

### 2.4.5 Sources of error

ICR’s approximation of the desired covariance incurs errors in two ways. The first source of error lies within the refinement step itself. Within a step, we disregard correlations by interpolating the coarse pixels with the refinement matrix  $R$ , which uses only  $n_{\text{csz}}^d$  neighboring coarse pixels instead of all coarse pixels. Furthermore, we add small corrections to the previous level but only correlate these with the refinement matrix  $\sqrt{D}$  in blocks of  $n_{\text{fsz}}^d$  fine pixels instead of correlating all fine pixels jointly. Both approximations strictly decrease the strength of correlations between pixels.

The second source of error lies in the iterative nature of our algorithm. In each refinement step we assume the previous refinement level to have modeled our desired GP without error. As outlined above, this assumption does not hold after the first iteration. The interpolation matrix  $R$  in our refinement mixes values of which the variance may be overestimated and the correlation underestimated. Already after the second refinement level, the incurred errors due to our approximations within the scheme are smeared out and potentially amplified (c.f. Section 2.7.5).

## 2.5 Experiments

We evaluate the accuracy and speed of our algorithm. In particular, we compare our implicit representation of the covariance against the true one for a low number of modeled points for which explicitly instantiating the kernel matrix is possible. Furthermore, we compare against KISS-GP. We choose KISS-GP because it has a similar computational complexity and promises to be applicable to millions or even billions of modeled points, outperformed previous popular inducing point methods in terms of speed and accuracy (Wilson and Nickisch, 2015; Wilson et al., 2015), and it highlights the conceptual differences in modeling GPs compared to our approach.

### 2.5.1 Accuracy

In contrast to our method, KISS-GP must not necessarily produce a proper GP prior because their approximate representation of  $K_{XX}$  is not always full rank<sup>1</sup>. This makes it impossible to quantify the loss of information from an information theoretical perspective (Leike and Ensslin, 2017). Instead, we quantify the error in terms of the absolute difference to the true covariance.

We test the accuracy of ICR’s representation of the covariance against the true one in terms of the mean absolute error (MAE) for  $N = 10.000$  points while varying the spacing between points. Analogously, we evaluate KISS-GP’s error in representing the covariance. For KISS-GP we use  $M = N = 10.000$  inducing points. As kernel, we use the homogeneous and isotropic Matérn covariance with degree-of-freedom parameter  $3/2$ :

$$k(x, x') = \left(1 + \frac{\sqrt{3}|x - x'|}{\rho}\right) \cdot \exp\left(-\frac{\sqrt{3}|x - x'|}{\rho}\right) \Big|_{\rho=\rho_0}, \quad (2.14)$$

where  $\rho$  denotes the characteristic length scale. For the experiment we space points logarithmically to probe orders of magnitude in spanned spacing. We set the minimal distance between points to be  $10^{-3}\rho_0$  and successively increase the maximal distance between points from  $10^{-2}\rho_0$  to  $10^2\rho_0$ .

Figure 2.3 depicts the accuracy of ICR and KISS-GP for one to five orders of magnitude spanned in the distance between points. The parametrizations of ICR are denoted by different colors. They all share  $n_{\text{lvl}} = 5$ . The fine pixels within each parametrization of ICR are located around the coarse center pixel and spaced such that they take up half the volume on the Euclidean grid of a coarse pixel, see Section 2.7.6.

Irrespective of the parametrization, ICR’s MAE changes minimally with the order of magnitude spanned in distance between points. ICR’s accuracy in terms of MAE improves with larger  $n_{\text{csz}}$ . The MAE of KISS-GP increases strongly with the spanned spacing. For every order of magnitude increase in distances spanned, the MAE increases by about 1.5 orders of magnitude. At three orders of magnitude in distances spanned, ICR and KISS-GP are comparable in accuracy. Below this threshold, KISS-GP is significantly more accurate and above this threshold ICR is significantly more accurate.

### 2.5.2 Covariance matrix representation

We compare the covariance matrix of ICR to the true one and the one of KISS-GP at the tipping point found in Section 2.5.1. Our modeled points are spaced

<sup>1</sup>KISS-GP yields a singular kernel matrix for  $M < N$  but also for  $M \geq N$  and points with vastly different spacing between modeled points such that the interpolation does not use at least  $M - N + 1$  of the regularly spaced inducing points. To apply the inverse of the kernel matrix, it is necessary to add some small diagonal correction or use an appropriate preconditioner respectively projection operator.

## 2. Sparse kernel Gaussian processes through iterative charted refinement (ICR)

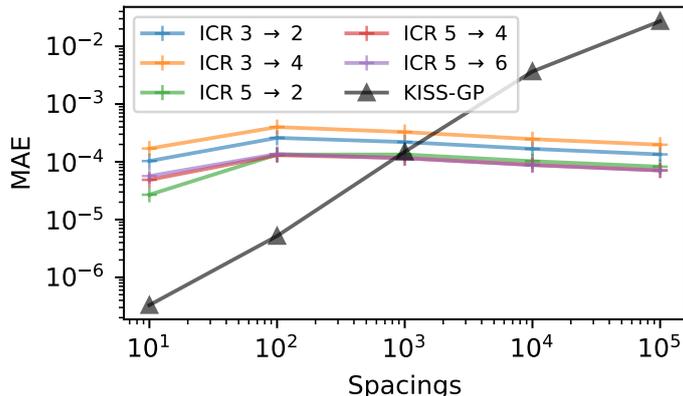


Figure 2.3: MAE between the true covariance and the approximate one by ICR (pluses) respectively KISS-GP (triangles). Each parametrization of ICR is denoted by its own color. The leftmost number in the legend for entries on ICR denotes  $n_{\text{CSZ}}$  and the rightmost number  $n_{\text{fSZ}}$ .

such that the distances to nearest neighbors range from  $10^{-3}\rho_0$  up to  $\rho_0$ . We again choose  $n_{|\text{V}|} = 5$  and within  $(n_{\text{CSZ}}, n_{\text{fSZ}}) \in \{(3, 2), (3, 4), (5, 2), (5, 4), (5, 6)\}$  we choose  $n_{\text{CSZ}} = 5$  and  $n_{\text{fSZ}} = 4$ .

The right column of Figure 2.4 compares the implicit covariance matrix of ICR to the true one. Both covariance matrices are overall in agreement. The diagonal is approximated comparatively well with errors up to  $6.6 \cdot 10^{-2}$ . On the off diagonal at intermediate spacing between points, there are significant differences in the strength of correlations with absolute errors up to 0.13 (13% of the variance). The mean absolute error (MAE) is  $1.2 \cdot 10^{-4}$ .

The left column of Figure 2.4 compares the covariance of KISS-GP to the true one. The maximum absolute error is  $1.9 \cdot 10^{-2}$  (14% of ICR's) and occurs on the diagonal (28% of ICR's maximum error on the diagonal). The MAE with respect to the true covariance is  $1.5 \cdot 10^{-4}$ , which is 1.3 times the MAE of ICR. ICR is more accurate in terms of the MAE but yields higher absolute errors at few diagonal and off-diagonal locations.

### 2.5.3 Computational speed

Let us now compare the computational speed of ICR versus KISS-GP. For each method we time the execution of a single forward pass of the model. In the case of our algorithm, we time the application of  $\sqrt{K_{\text{ICR}}}$ . For KISS-GP a forward pass involves applying the inverse as well as computing the log-determinant of the kernel matrix. We compare their respective performance on both the CPU and GPU. All experiments are carried out with double precision using 8 cores of an Intel Xeon IceLake-SP 8360Y CPU with 62 GB of RAM and a single NVIDIA A100 GPU with 40 GB of HBM2.

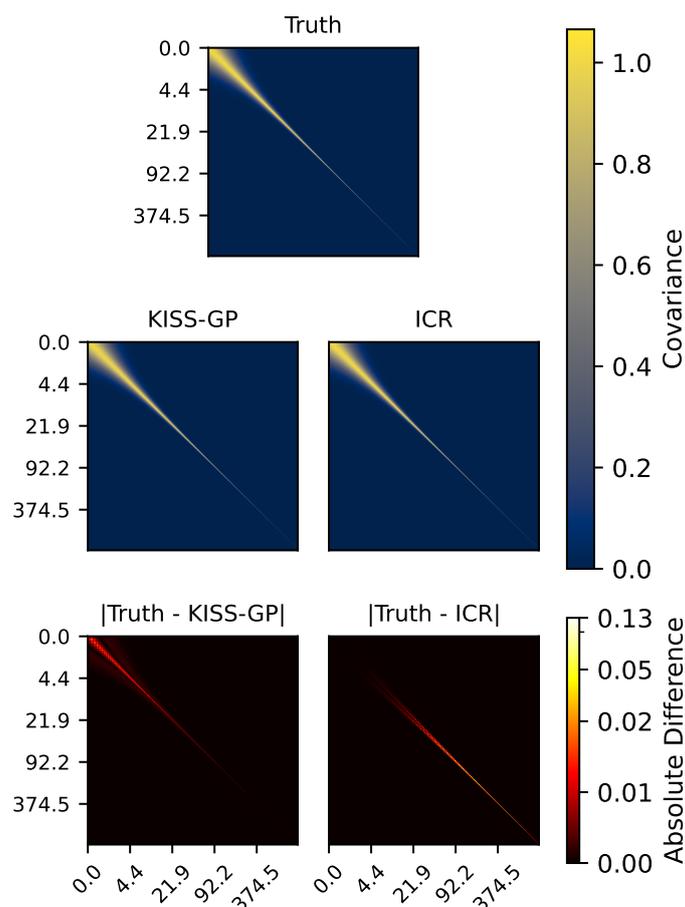


Figure 2.4: The true covariance, the implicit covariance of ICR and KISS-GP as well as their respective absolute difference to the true one. Abscissas and ordinates are logarithmic. Their labels are given in multiples of  $\rho_0$ . The colorbar is linear for the upper three plots of the covariances. For the difference plots, the colorbar is linear up to 0.02 and logarithmic from there on.

For the experiment, we use relatively few Krylov subspace iterations to create a favorable setting for KISS-GP. We use 40 CG iterations to apply the inverse of the kernel matrix, and 10 samples each optimized for 15 Lanczos iterations to stochastically estimate the log-determinant. KISS-GP’s speed crucially depends on the configuration of the Krylov subspace methods. Larger inference problems would likely require more CG and Lanczos iterations. For all probed numbers of modeled points, we set the number of inducing points to the number of modeled points  $M = N$ .

Figure 2.5 depicts the execution time of KISS-GP versus ICR for a varying number of modeled points for the CPU and GPU on double logarithmic scales. Our algorithm is consistently about two orders of magnitude faster than KISS-GP on both the CPU and GPU for more than  $10^5$  modeled points. Its computational

## 2. Sparse kernel Gaussian processes through iterative charted refinement (ICR)

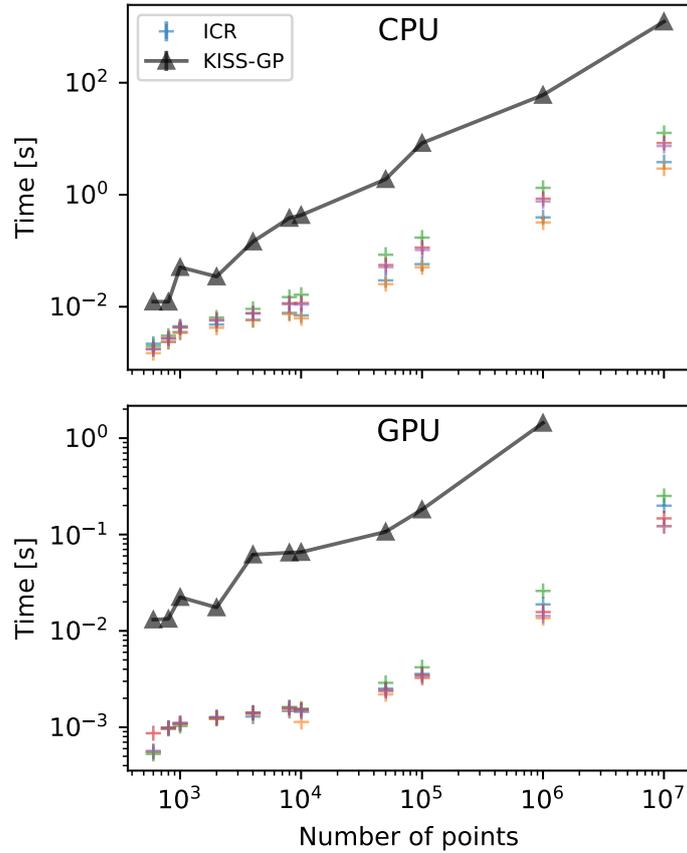


Figure 2.5: Performance benchmark on the CPU (top) and GPU (bottom) of KISS-GP versus ICR on double logarithmic axes. KISS-GP measurements are shown as triangles and ICR measurements as plus signs. The markers are placed at the median time it took the model to perform a forward pass. Minimum and maximum timings are shown as vertical bars but are fully absorbed by the marker size. Each parametrization of ICR is denoted by its own color.

advantage changes minimally with the choice of parameters for ICR.

## 2.6 Conclusion

We introduce a new algorithm called Iterative Charted Refinement (ICR) for evaluating GPs on nearly arbitrarily spaced points. We reformulate the problem of applying the inverse and computing the log-determinant of the kernel matrix to applying the “square-root” of it. To efficiently approximate the square-root, our algorithm transforms the modeled points with a coordinate chart which must either be provided by the user or generated from the modeled points. ICR has a computational complexity of  $\mathcal{O}(N)$ .

Its major advantages are its guaranteed positive definite kernel matrix, its computational speed and its accuracy in representing covariance matrices for highly unevenly spaced points. ICR is consistently faster than KISS-GP by about two orders of magnitude on both the CPU and GPU for more than  $10^5$  modeled points. Our algorithm gets more accurate relative to KISS-GP the more orders of magnitude are spanned in the spacing of points. For points whose spacing varies over several order of magnitudes, ICR can be more than 100 times more accurate than KISS-GP in terms of MAE in representing the covariance matrix while being 100 times faster.

## Acknowledgements

Gordian Edenhofer acknowledges that support for this work was provided by the German Academic Scholarship Foundation in the form of a PhD scholarship (“Promotionsstipendium der Studienstiftung des Deutschen Volkes”).

## 2.7 Supplementary material

### 2.7.1 Proof of positive definiteness of $K_{\text{ICR}}$

We want to show that ICR models a proper GP. A process is called Gaussian iff every discrete realization of it is a multivariate normal distribution. A multivariate normal distribution is a distribution that is fully determined by its mean and its positive definite covariance matrix. ICR models a GP iff  $K_{\text{ICR}} = \sqrt{K_{\text{ICR}}}\sqrt{K_{\text{ICR}}}^T$  is a positive definite matrix.

By virtue of our algorithm applying  $\sqrt{K_{\text{ICR}}}$  in a generative model, the covariance matrix is  $K_{\text{ICR}} = \sqrt{K_{\text{ICR}}}\sqrt{K_{\text{ICR}}}^T$ , see Section 2.4. Due to the existence of the decomposition of  $K_{\text{ICR}}$  into  $\sqrt{K_{\text{ICR}}}$ , we know that the covariance must be at least positive semidefinite. To prove positive definiteness, let us look at the positions of the non-zero entries of the matrix representation of  $\sqrt{K_{\text{ICR}}}$ . To retrieve a unique matrix representation, assume the standard normally distributed variable  $\xi$  to be a one-dimensional vector starting with the flattened entries at the coarsest grid going to the flattened entries at the finest grid. ICR adds part of  $\xi$  correlated in blocks of  $n_{\text{fsz}}$  to the previous level, see Equation (2.12), thus  $\sqrt{K_{\text{ICR}}}$  can be split as follows

$$\sqrt{K_{\text{ICR}}} = \left( \begin{array}{c|c|c} & \boxed{\sqrt{D}'} & \\ \hline B & & \ddots \\ \hline & & \boxed{\sqrt{D}'\dots'} \end{array} \right) \quad (2.15)$$

with the first block  $B$  denoting the refinement combining all of  $\xi$  up to the final level as well as the mapping to the final level and  $\sqrt{D}', \dots, \sqrt{D}'\dots'$  the Cholesky decomposition of Equation (2.8) for each block of  $n_{\text{fsz}}$  at the final level.

## 2. Sparse kernel Gaussian processes through iterative charted refinement (ICR)

Thus,  $K_{\text{ICR}}$  splits into<sup>2</sup>

$$\begin{aligned}
 K_{\text{ICR}} &= \sqrt{K_{\text{ICR}}} \sqrt{K_{\text{ICR}}}^T \\
 &= BB^T + \begin{pmatrix} \boxed{\sqrt{D}'} & & \\ & \ddots & \\ & & \boxed{\sqrt{D}' \dots'} \end{pmatrix} \begin{pmatrix} \boxed{\sqrt{D}'^T} & & \\ & \ddots & \\ & & \boxed{\sqrt{D}' \dots'^T} \end{pmatrix}.
 \end{aligned} \tag{2.16}$$

The first summand is at least positive semidefinite by the existence of  $B$ . This holds irrespective of the concrete value of  $B$ . The latter summand is positive definite by virtue of being a block-diagonal matrix of Cholesky decompositions. As the sum of a matrix which is at least positive semidefinite and a matrix which is positive definite,  $K_{\text{ICR}}$  is itself positive definite.

### 2.7.2 Proof of linear scaling

We want to show that ICR's computational complexity scales linearly with the number of modeled points. For simplicity, we assume a 1D space,  $n_{\text{csz}} = 3$ ,  $n_{\text{fsz}} = 2$  and  $N^{(0)} > 4$  number of initial, coarse-most points. The proof holds analogously for any  $n_{\text{fsz}}$  and  $n_{\text{csz}}$  in arbitrary dimensions for which the refinement produces more points in the next level than in the previous one.

Let us start by first relating the number of modeled points  $N$  to the number of initial points  $N^{(0)}$ . In terms of the number of pixels at the previous level,  $N$  can be expressed as  $N = N^{(n_{\text{lvl}})} = 2(N^{(n_{\text{lvl}}-1)} - 2)$ . We can recurse this expression to connect  $N$  to  $N^{(0)}$ :

$$N = 2(N^{(n_{\text{lvl}}-1)} - 2) \tag{2.17}$$

$$= 2(2(N^{(n_{\text{lvl}}-2)} - 2) - 2) \tag{2.18}$$

$$\vdots \tag{2.19}$$

$$= 2(2(2(\dots(N^{(0)} - 2)\dots) - 2) - 2) \tag{2.20}$$

$$= 2^{n_{\text{lvl}}} N^{(0)} - 2^2 - 2^3 - \dots - 2^{n_{\text{lvl}}+1} \tag{2.21}$$

$$= 2^{n_{\text{lvl}}} N^{(0)} - 2^{n_{\text{lvl}}+1} \cdot \left(1 + \frac{1}{2^{-1}} + \dots + \frac{1}{2^{-(n_{\text{lvl}}-1)}}\right) \tag{2.22}$$

$$\stackrel{\text{geometric series}}{>} 2^{n_{\text{lvl}}} N^{(0)} - 2^{n_{\text{lvl}}+1} \cdot 2 \tag{2.23}$$

$$= 2^{n_{\text{lvl}}} (N^{(0)} - 4) . \tag{2.24}$$

---

<sup>2</sup>One can think of this step as implicitly rewriting the matrix product as a sum of column-by-row outer products. For the product of two matrices  $A$  and  $B$  it holds that  $AB = \sum_{i=1}^k a_i b_i^T$  with  $a_i$  the  $i$ th column of  $A$ ,  $b_i^T$  the  $i$ th row of  $B$  and  $k$  the length of the second axis of  $A$ , respectively, the first axis of  $B$ .

The complexity of ICR is linear in the total number of points  $N^{\text{tot}}$  from all coarsening levels that appear in ICR as it only touches a point at most  $n_{\text{csz}} + n_{\text{fsz}}$  times. To show that ICR's computational complexity is also linear in  $N$ , we need to show that the total number of points  $N^{\text{tot}}$  is a multiple of the number of modeled points. Analogously to Equation (2.20), we get for the total number of points

$$N^{\text{tot}} = N^{(0)} + 2(N^{(0)} - 2) + \dots + 2(2(2(\dots(N^{(0)} - 2)\dots) - 2) - 2) \quad (2.25)$$

$$< N^{(0)} + 2(N^{(0)} - 2) + \dots + 2^{n_{\text{ivl}}}(N^{(0)} - 2) \quad (2.26)$$

$$= N^{(0)} + 2^{n_{\text{ivl}}}(N^{(0)} - 2) \cdot \left(1 + \frac{1}{2^{-1}} + \dots + \frac{1}{2^{-(n_{\text{ivl}}-1)}}\right) \quad (2.27)$$

$$\stackrel{\text{geometric series}}{<} N^{(0)} + 2^{n_{\text{ivl}}}(N^{(0)} - 2) \cdot 2 \quad (2.28)$$

$$= N^{(0)} + 2^{n_{\text{ivl}}}(N^{(0)} - 4) \cdot 2 + 2^{n_{\text{ivl}}} \cdot 4 \quad (2.29)$$

$$= N^{(0)} + 2^{n_{\text{ivl}}}(N^{(0)} - 4) \cdot \left(2 + \frac{4}{N^{(0)} - 4}\right) \quad (2.30)$$

$$\stackrel{\text{Equation (2.24)}}{<} N^{(0)} + \left(2 + \frac{4}{N^{(0)} - 4}\right) N. \quad (2.31)$$

Since  $N^{(0)} > 4$ , the total number of points is bound by  $6N$  and the computational complexity is linear in  $N$ .

### 2.7.3 Automatic charting

In the absence of a user-provided coordinate chart, one can be constructed from the positions of the modeled points. Any invertible map from the modeled points to a Euclidean grid is sufficient and there are infinitely many such maps. One possible choice of an invertible map is given in Algorithm 1 which linearly interpolates within the sorted set of all modeled points to construct a mapping to a one-dimensional regularly spaced Euclidean grid position.

**Data:** modeled points  $X$ , to-be-charted point  $x'$  in  $\Omega_x$

**Result:**  $\tilde{x}'$ : charted point in Euclidean space

$M \leftarrow \text{sorted}(X)$ ;

**for**  $i$  **from** 1 **to**  $\text{len}(X) - 1$  **do**

| **if**  $M_i \leq x' \leq M_{i+1}$  **then**

|  $\tilde{x}' \leftarrow i + \frac{M_i - x'}{M_{i+1} - M_i}$ ;

| **break**;

**end**

**Algorithm 1:** Coordinate chart  $\phi$  constructed from the modeled points using a (inefficient) linear search algorithm to find the suitable Euclidean grid position. For multidimensional data  $\leq$  denotes a comparison along all dimensions and is true iff all components of the vector are less or equal.

### 2.7.4 Optimal sizes for the refinement matrices

ICR has two tunable parameters  $n_{\text{csz}}$  and  $n_{\text{fsz}}$  which determine the sizes of the refinement matrices  $R$  and  $\sqrt{D}$ . Neglecting computational constraints, the optimal sizes for  $R$  and  $\sqrt{D}$  are such that they encompass the full grid. In this limit ICR is exact (cf. Section 2.4.1). For reasons of computational efficiency (see Section 2.4.4), the size of the refinement matrices must be chosen within a set of small integer values.

Assume we have a set  $\Omega_{\text{csz}}$  of allowed  $n_{\text{csz}}$  respectively  $\Omega_{\text{fsz}}$  for  $n_{\text{fsz}}$ . The optimal choice in terms of the least amount of information lost (Knuth and Skilling, 2012; Leike and Ensslin, 2017) by representing the GP using  $K_{\text{ICR}}$  instead of  $K_{XX}$  is the tuple of  $n_{\text{csz}}$  and  $n_{\text{fsz}}$  whose Kullback-Leibler divergence of the GP with the desired covariance  $K_{XX}$  to the GP with covariance  $K_{\text{ICR}}$  is the lowest

$$\min_{n_{\text{csz}} \in \Omega_{\text{csz}}, n_{\text{fsz}} \in \Omega_{\text{fsz}}} D_{\text{KL}}(\mathcal{N}(0, K_{XX}) \parallel \mathcal{N}(0, K_{\text{ICR}})) \quad (2.32)$$

$$= \min_{n_{\text{csz}} \in \Omega_{\text{csz}}, n_{\text{fsz}} \in \Omega_{\text{fsz}}} \frac{1}{2} \left[ \text{tr}(K_{\text{ICR}}^{-1} K_{XX}) - d + \ln \left( \frac{\det K_{\text{ICR}}}{\det K_{XX}} \right) \right]. \quad (2.33)$$

In the implementation available as part of `NIFTy.re`, described in Chapter 3, we compare the true GP to our approximate one on a small subset of the modeled points. We only use a small subset because for a large set of modeled points, it is intractable to explicitly compute the matrices  $K_{\text{ICR}}$  and  $K_{XX}$ . We return the optimal  $n_{\text{csz}}$  and  $n_{\text{fsz}}$  by computing Equation (2.33) explicitly for two predefined sets  $\Omega_{\text{csz}}$  and  $\Omega_{\text{fsz}}$ .

### 2.7.5 Accounting for accumulating errors

In theory, it is possible to account for the accumulating errors by not using the desired kernel but the actual one for the coarse pixels. The actual kernel is implicitly available which would be sufficient. However, to evaluate the implicit kernel we would need to pass through all previous layers for every fine pixel thus yielding a quadratic scaling.

### 2.7.6 Multivariate ICR with non-stationary kernel on irregular grids

Algorithm 2 summarizes the most general form of ICR introduced in Section 2.4. It extends Section 2.4.2 to  $d$ -dimensional irregular grids with non-stationary kernels using the methodology described in Sections 2.4.3 and 2.4.4. The algorithm takes standard normally distributed  $\xi$ , kernel  $k$ , chart  $\phi^{-1}$  and inverse chart  $\phi$ , as well as the modeled points  $X$  as input and produces a realization of a multivariate normally distributed  $s$  with  $K_{\text{ICR}}$  as covariance.

The full algorithm needs to adjust the size of the Euclidean grid to which the modeled points are mapped to with the refinement depth. The extent

of the grid shrinks with each refinement level due to the loss of the outermost  $n_{\text{csz}} - 1$  coarse pixels along each dimension. Thus, the boundaries of the grid shrink by  $\frac{1}{2}(n_{\text{csz}} - 1) - \frac{1}{4}(n_{\text{fsz}} - 1)$ . After  $l$  refinement steps, the grid is smaller by  $2(1 - 2^{-l})\left(\frac{1}{2}(n_{\text{csz}} - 1) - \frac{1}{4}(n_{\text{fsz}} - 1)\right)$  zero-level pixels respectively  $2^{n_{\text{lvl}}}(1 - 2^{-l})\left(\frac{1}{2}(n_{\text{csz}} - 1) - \frac{1}{4}(n_{\text{fsz}} - 1)\right)$  pixels at refinement level  $n_{\text{lvl}}$ . From the refinement level  $n_{\text{lvl}}$  to the  $l$ th refinement level, the edges of the grid move out by

$$\left[(1 - 2^{-n_{\text{lvl}}}) - (1 - 2^{-l})\right] 2^{n_{\text{lvl}}} \left(\frac{1}{2}(n_{\text{csz}} - 1) - \frac{1}{4}(n_{\text{fsz}} - 1)\right) \quad (2.34)$$

$$= \left[-2^{-n_{\text{lvl}}} + 2^{-l}\right] 2^{n_{\text{lvl}}} \left(\frac{1}{2}(n_{\text{csz}} - 1) - \frac{1}{4}(n_{\text{fsz}} - 1)\right) \quad (2.35)$$

$$= \left[2^{n_{\text{lvl}}-l} - 1\right] \left(\frac{1}{2}(n_{\text{csz}} - 1) - \frac{1}{4}(n_{\text{fsz}} - 1)\right) \quad (2.36)$$

in terms of pixels at the refinement level  $n_{\text{lvl}}$ .

## 2. Sparse kernel Gaussian processes through iterative charted refinement (ICR)

**Data:**  $\xi^{(0)}, \dots, \xi^{(n_{lv})}$ , kernel  $k$ , chart  $\phi^{-1}$  and inverse chart  $\phi$ , set of modeled points  $X$   
**Result:**  $s^{n_{lv}}$ : A sample with approximate covariance  $K_{XX}$   
 $\tilde{X}^{(n_{lv})} \leftarrow \phi(X)$ ;  
**for**  $m$  **from** 1 **to**  $d$  **do**  
     $N_m^{(n_{lv})} \leftarrow \text{size}(x_{1, \dots, m, \dots, 1_d})$ ;  
     $p_m \leftarrow \tilde{X}_{1, \dots, 2m, \dots, 1_d}^{(n_{lv})} - \tilde{X}_{1, \dots, 1m, \dots, 1_d}^{(n_{lv})}$ ;  
**end**  
**for**  $l$  **from**  $n_{lv} - 1$  **to** 1 **do**  
    **for**  $m$  **from** 1 **to**  $d$  **do**  
         $N_m^{(l)} \leftarrow N_m^{(l+1)} / 2 + (n_{csz} - 1)$ ;  
        // Treat non-integer results for  $N_m^{(l)}$  in a slightly more elaborate way  
        :  
    **end**  
    // Grid gets slightly smaller with every refinement  
     $o \leftarrow (2^{-(n_{lv}-l)} - 1) (\frac{1}{2}(n_{csz} - 1) - \frac{1}{4}(n_{fsz} - 1))$ ;  
    **for**  $i_1, \dots, i_d$  **from** 1,  $\dots$ , 1 **to**  $N_1^{(l)}, \dots, N_d^{(l)}$  **do**  
         $\tilde{X}_{i_1, \dots, i_d}^{(l)} \leftarrow (p_1 o + i_1 p_1 2^{n_{lv}-l}, \dots, p_d o + i_d p_d 2^{n_{lv}-l})^T$ ;  
    **end**  
**end**  
 $D \leftarrow k(\phi^{-1}(\tilde{X}_{1:N_1^{(0)}, \dots, 1:N_d^{(0)}}^{(0)}), \phi^{-1}(\tilde{X}_{1:N_1^{(0)}, \dots, 1:N_d^{(0)}}^{(0)}))$ ;  
 $\sqrt{D} \leftarrow \text{cholesky}(D)$ ;  
 $\sqrt{D} \leftarrow \text{reshape}(\sqrt{D}, (N_1^{(0)}, \dots, N_d^{(0)}, N_1^{(0)}, \dots, N_d^{(0)}))$ ;  
 $s_{i_1, \dots, i_d}^{(0)} \leftarrow \sum_{j_1, \dots, j_d} \sqrt{D}_{i_1, \dots, i_d, j_1, \dots, j_d}^{(0)} (\xi^{(0)})_{j_1, \dots, j_d}$ ;  
**for**  $l$  **from** 1 **to**  $n_{lv}$  **do**  
    **for**  $i_1, \dots, i_d$  **from** 1,  $\dots$ , 1 **to**  $N_1^{(l-1)} - n_{csz}, \dots, N_d^{(l-1)} - n_{csz}$  **by**  $n_{fsz}/2, \dots, n_{fsz}/2$   
        **do**  
            **for**  $m$  **from** 1 **to**  $d$  **do**  
                 $i_m^{\text{next}} \leftarrow 2i_m - 1$ ;  
            **end**  
            **for**  $j_1, \dots, j_d$  **from** 1,  $\dots$ , 1 **to**  $n_{fsz}, \dots, n_{fsz}$  **do**  
                 $K_{cc} \leftarrow k(\phi^{-1}(\tilde{X}_{i_1:i_1+n_{csz}, \dots, i_d:i_d+n_{csz}}^{(l-1)}), \phi^{-1}(\tilde{X}_{i_1:i_1+n_{csz}, \dots, i_d:i_d+n_{csz}}^{(l-1)}))$ ;  
                 $K_{ff} \leftarrow k(\phi^{-1}(\tilde{X}_{i_1^{\text{next}}:i_1^{\text{next}}+n_{fsz}, \dots, i_d^{\text{next}}:i_d^{\text{next}}+n_{fsz}}^{(l)}), \phi^{-1}(\tilde{X}_{i_1^{\text{next}}:i_1^{\text{next}}+n_{fsz}, \dots, i_d^{\text{next}}:i_d^{\text{next}}+n_{fsz}}^{(l)}))$ ;  
                 $K_{cf} \leftarrow k(\phi^{-1}(\tilde{X}_{i_1:i_1+n_{csz}, \dots, i_d:i_d+n_{csz}}^{(l-1)}), \phi^{-1}(\tilde{X}_{i_1^{\text{next}}:i_1^{\text{next}}+n_{fsz}, \dots, i_d^{\text{next}}:i_d^{\text{next}}+n_{fsz}}^{(l)}))$ ;  
                 $R \leftarrow K_{cf}^T \cdot K_{cc}^{-1}$ ;  
                 $D \leftarrow K_{ff} - K_{cf}^T \cdot K_{cc}^{-1} \cdot K_{cf}$ ;  
                 $\sqrt{D} \leftarrow \text{cholesky}(D)$ ;  
                 $R \leftarrow \text{reshape}(R, (n_{fsz}, \dots, n_{fsz}, n_{csz}, \dots, n_{csz}))$ ;  
                 $\sqrt{D} \leftarrow \text{reshape}(\sqrt{D}, (n_{fsz}, \dots, n_{fsz}, n_{fsz}, \dots, n_{fsz}))$ ;  
                 $r \leftarrow \sum_{k_1, \dots, k_d}^{n_{csz}, \dots, n_{csz}} R_{j_1, \dots, j_d, k_1, \dots, k_d} s_{i_1+k_1, \dots, i_d+k_d}^{(l-1)}$ ;  
                 $s_{i_1^{\text{next}}+j_1, \dots, i_d^{\text{next}}+j_d}^{(l)} \leftarrow r + \sum_{k_1, \dots, k_d}^{n_{fsz}, \dots, n_{fsz}} \sqrt{D}_{j_1, \dots, j_d, k_1, \dots, k_d} \xi_{i_1^{\text{next}}+k_1, \dots, i_d^{\text{next}}+k_d}^{(l)}$ ;  
            **end**  
        **end**  
    **end**  
**end**

**Algorithm 2:** Iterative Charted Refinement of a  $d$ -dimensional irregularly charted grid with non-stationary kernel. The algorithm uses NumPy's indexing notation (Harris et al., 2020a) and the centered dot symbol  $\cdot$  denotes a matrix-matrix product.

### 3 Re-envisioning numerical information field theory (NIFTy.re): A library for Gaussian processes and variational inference

*The following chapter is an article currently under review at the Journal of Open Source Software (Edenhofer et al., 2024b). This article emerged from a collaboration in the group of Torsten Enßlin with me leading the development effort and the writing of the paper. Philipp Frank and Jakob Roth contributed toward a high-level interface for the implemented variational inference algorithms. Reimar Leike and Massin Guerdi contributed code toward the early prototype of the framework while Lukas Scheel-Platz, Matteo Guardiani, Vincent Eberle, and Margret Westerkamp contributed code and tests toward the final version of the framework. Torsten Enßlin contributed to discussions about the direction for the framework and the design of user-friendly interfaces. All authors read, commented, and approved the manuscript. The text was adapted to fit the thesis.*

#### **Abstract**

Imaging is the process of transforming noisy, incomplete data into a space that humans can interpret. NIFTy is a Bayesian framework for imaging and has already successfully been applied to many fields in astrophysics. Previous design decisions held the performance and the development of methods in NIFTy back. We present a rewrite of NIFTy, coined NIFTy.re, which reworks the modeling principle, extends the inference strategies, and outsources much of the heavy lifting to JAX. The rewrite dramatically accelerates models written in NIFTy, lays the foundation for new types of inference machineries, improves maintainability, and enables interoperability between NIFTy and the JAX machine learning ecosystem.

### 3.1 Statement of need

Imaging commonly involves millions to billions of pixels. Each pixel usually corresponds to one or more correlated degrees of freedom in the model space. Modeling this many degrees of freedom is computationally demanding. However, imaging is not only computationally demanding but also statistically challenging. The noise in the data requires a statistical treatment and needs to be accurately propagated from the data to the uncertainties in the final image. To do this, we require an inference machinery that not only handles extremely high-dimensional spaces, but one that does so in a statistically rigorous way.

NIFTy is a Bayesian imaging library (Arras et al., 2019; Selig et al., 2013; Steininger et al., 2019). It is designed to infer the million- to billion-dimensional posterior distribution in the image space from noisy input data. At the core of NIFTy lies a set of powerful GP models and accurate VI algorithms.

NIFTy.re is a rewrite of NIFTy in JAX (Bradbury et al., 2018) with all relevant previous GP models, new, more flexible GP models, and a more flexible machinery for approximating posterior distributions. Being written in JAX, NIFTy.re effortlessly runs on accelerator hardware such as the GPU and TPU, vectorizes models whenever possible, and just-in-time compiles code for additional performance. NIFTy.re switches from a home-grown automatic differentiation engine that was used in NIFTy to JAX’s automatic differentiation engine. This lays the foundation for new types of inference machineries that make use of the higher order derivatives provided by JAX. Through these changes, we envision to harness significant gains in maintainability of NIFTy.re compared to NIFTy and a faster development cycle for new features.

We expect NIFTy.re to be highly useful for many imaging applications and envision many applications within and outside of astrophysics (Arras et al., 2019, 2022; Eberle et al., 2023; Eberle et al., 2022; Frank et al., 2021; Frank, 2022; Hutschenreuter et al., 2022, 2023; Leike and Enßlin, 2019; Leike et al., 2020; Mertsch and Phan, 2023; Roth et al., 2023b,a; Scheel-Platz et al., 2023; Tsouros et al., 2024b; Welling et al., 2021; Westerkamp et al., 2023). NIFTy.re underpins the reconstruction described in Chapter 4. We use NIFTy.re to infer a 500-million-dimensional posterior distribution using VI (Knollmüller and Enßlin, 2019). We make extensive use of NIFTy.re’s GPU support to reduce the runtime by two orders of magnitude compared to the CPU. With NIFTy.re bridging ideas from NIFTy to JAX, we envision many new possibilities for inferring classical machine learning models with NIFTy’s inference methods and a plethora of opportunities to use NIFTy-components such as the GP models in classical neural network frameworks.

NIFTy.re competes with other GP libraries as well as with probabilistic programming languages and frameworks. Compared to GPyTorch (Hensman et al., 2015), GPflow (De G. Matthews et al., 2017), george (Ambikasaran et al., 2015), or TinyGP (Foreman-Mackey et al., 2024), NIFTy.re focuses on GP models for structured spaces and does not assume the posterior to be analytically accessi-

ble. Instead, `NIFTy.re` tries to approximate the true posterior using VI. Compared to classical probabilistic programming languages such as Stan (Carpenter et al., 2017) and frameworks such as Pyro (Bingham et al., 2019), NumPyro (Phan et al., 2019), pyMC3 (Salvatier et al., 2016), emcee (Foreman-Mackey et al., 2013), dynesty (Koposov et al., 2023; Speagle, 2020), or BlackJAX (Cabezas et al., 2024), `NIFTy.re` focuses on inference in extremely high-dimensional spaces. `NIFTy.re` exploits the structure of probabilistic models in its VI techniques (Frank et al., 2021). With `NIFTy.re`, the GP models and the VI machinery are now fully accessible in the JAX ecosystem and `NIFTy.re` components interact seamlessly with other JAX packages such as BlackJAX and JAXopt/Optax (Blondel et al., 2022; DeepMind et al., 2020).

## 3.2 Core components

`NIFTy.re` brings tried and tested structured GP models and VI algorithms to JAX. GP models are highly useful for imaging problems, and VI algorithms are essential to probing high-dimensional posteriors, which are often encountered in imaging problems. `NIFTy.re` infers the parameters of interest from noisy data via a stochastic mapping that goes in the opposite direction: from the parameters of interest to the data.

`NIFTy` and `NIFTy.re` build up hierarchical models for the posterior. The log-posterior function reads  $\ln p(\theta|d) := \ell(d, f(\theta)) + \ln p(\theta) + \text{const}$  with log-likelihood  $\ell$ , forward model  $f$  mapping the parameters of interest  $\theta$  to the data space, and log-prior  $\ln p(\theta)$ . The goal of the inference is to draw samples from the posterior  $p(\theta|d)$ .

What is considered part of the likelihood versus part of the prior is ill-defined. Without loss of generality, `NIFTy` and `NIFTy.re` re-formulate models such that the prior is always standard Gaussian. They implicitly define a mapping from a new latent space with a priori standard Gaussian parameters  $\xi$  to the parameters of interest  $\theta$ . The mapping  $\theta(\xi)$  is incorporated into the forward model  $f(\theta(\xi))$  in such a way that all relevant details of the prior model are encoded in the forward model. This choice of re-parameterization (Rezende and Mohamed, 2015) is called standardization. It is often carried out implicitly in the background without user input.

### 3.2.1 Gaussian processes

One standard tool from the `NIFTy.re` toolbox is the so-called correlated field GP model from `NIFTy`. This model relies on the harmonic domain being easily accessible. For example, for pixels spaced on a regular Cartesian grid, the natural choice to represent a stationary kernel is the Fourier domain. In the generative picture, a realization  $s$  drawn from a GP then reads  $s = \text{FT} \circ \sqrt{P} \circ \xi$  with FT the (fast) Fourier transform,  $\sqrt{P}$  the square-root of the power-spectrum in harmonic space, and  $\xi$  standard Gaussian random variables. In the implementation

### 3. Re-envisioning numerical information field theory (*NIFTy.re*): A library for Gaussian processes and variational inference

in *NIFTy.re* and *NIFTy*, the user can choose between two adaptive kernel models, a non-parametric kernel  $\sqrt{P}$  and a Matérn kernel  $\sqrt{P}$  (for details on their implementation see Arras et al., 2022; Guardiani et al., 2022). A code example that initializes a non-parametric GP prior for a  $128 \times 128$  space with unit volume is shown in the following.

```
from nifty8 import re as jft

dims = (128, 128)
cfm = jft.CorrelatedFieldMaker("cf")
cfm.set_amplitude_total_offset(
    offset_mean=2,
    offset_std=(1e-1, 3e-2)
)
# Parameters for the kernel and the regular 2D Cartesian grid for
# which it is defined
cfm.add_fluctuations(
    dims,
    distances=tuple(1.0 / d for d in dims),
    fluctuations=(1.0, 5e-1),
    loglogavgslope=(-3.0, 2e-1),
    flexibility=(1e0, 2e-1),
    asperity=(5e-1, 5e-2),
    prefix="ax1",
    non_parametric_kind="power",
)
# Get the forward model for the GP prior
correlated_field = cfm.finalize()
```

Not all problems are well described by regularly spaced pixels. For more complicated pixel spacings, *NIFTy.re* features Iterative Charted Refinement (see Chapter 2), a GP model for arbitrarily deformed spaces. This model exploits nearest neighbor relations on various coarsenings of the discretized modeled space and runs very efficiently on GPUs. For one-dimensional problems with arbitrarily spaced pixels, *NIFTy.re* also implements multiple flavors of Gauss-Markov processes.

#### 3.2.2 Building up complex models

Models are rarely just a GP prior. Commonly, a model contains at least a few non-linearities that transform the GP prior or combine it with other random variables. For building more complex models, *NIFTy.re* provides a `Model` class that offers a somewhat familiar object-oriented design yet is fully JAX compatible and functional under the hood. The following code shows how to build a slightly more complex model using the objects from the previous example.

```
from jax import numpy as jnp

class Forward(jft.Model):
    def __init__(self, correlated_field):
```

```

self._cf = correlated_field
# Tracks a callable with which the model can be
# initialized. This is not strictly required, but comes in
# handy when building deep models. Note, the init method
# (short for "initialization" method) is not to be confused
# with the prior, which is always standard Gaussian.
super().__init__(init=correlated_field.init)

def __call__(self, x):
    # NOTE, any kind of masking of the output, non-linear and
    # linear transformation could be carried out here. Models
    # can also be combined and nested in any way and form.
    return jnp.exp(self._cf(x))

forward = Forward(correlated_field)

data = jnp.load("data.npy")
lh = jft.Poissonian(data).amend(forward)

```

All GP models in NIFTy.re as well as all likelihoods behave like instances of `jft.Model`, meaning that JAX understands what it means if a computation involves `self`, other `jft.Model` instances, or their attributes. In other words, `correlated_field`, `forward`, and `lh` from the code snippets shown here are all so-called pytrees in JAX, and, for example, the following is valid code `jax.jit(lambda l, x: l(x))(lh, x0)` with `x0` some arbitrarily chosen valid input to `lh`. Inspired by equinox (Kidger and Garcia, 2021), individual attributes of the class can be marked as non-static or static via `dataclass.field(metadata=dict(static=...))` for the purpose of compiling. Depending on the value, JAX will either treat the attribute as an unknown placeholder or as a known concrete attribute and potentially inline it during compilation. This mechanism is extensively used in likelihoods to avoid inlining large constants such as the data and to avoid expensive re-compilations whenever possible.

### 3.2.3 Variational inference

NIFTy.re is built for models with millions to billions of degrees of freedom. To probe the posterior efficiently and accurately, NIFTy.re relies on VI. Specifically, NIFTy.re implements Metric Gaussian Variational Inference (MGVI; Knollmüller and Enßlin, 2019) and its successor geometric Variational Inference (geoVI; Frank et al., 2021; Frank, 2022). At the core of both MGVI and geoVI lies an alternating procedure in which one switches between optimizing the Kullback–Leibler divergence for a specific shape of the variational posterior and updating the shape of the variational posterior. MGVI and geoVI define the variational posterior via samples, specifically, via samples drawn around an expansion point. The samples in MGVI and geoVI exploit model-intrinsic knowledge of the posterior’s approximate shape, encoded in the Fisher information metric and the prior curvature (Frank et al., 2021).

### 3. Re-envisioning numerical information field theory (*NIFTy.re*): A library for Gaussian processes and variational inference

*NIFTy.re* implements both MGVI and geoVI and allows for much finer control over the way samples are drawn and updated compared to *NIFTy*. Furthermore, *NIFTy.re* exposes stand-alone functions for drawing MGVI and geoVI samples from any arbitrary model with a likelihood from *NIFTy.re* and a forward model that is differentiable by JAX. In addition to stand-alone sampling functions, *NIFTy.re* also provides tools to configure and execute the alternating Kullback–Leibler divergence optimization and sample adaption at a lower abstraction level. These tools are provided in a JAXopt/Optax-style optimizer class (Blondel et al., 2022; DeepMind et al., 2020).

A typical minimization with *NIFTy.re* is shown in the following. It retrieves six independent, antithetically mirrored samples from the approximate posterior via 25 iterations of alternating between optimization and sample adaption. The final result is stored in the `samples` variable. A convenient one-shot wrapper for the code below is `jft.optimize_kl`. By virtue of all modeling tools in *NIFTy.re* being written in JAX, it is also possible to combine *NIFTy.re* tools with BlackJAX (Cabezas et al., 2024) or any other posterior sampler in the JAX ecosystem.

```
from jax import random

key = random.PRNGKey(42)
key, sk = random.split(key, 2)
# NIFTy is agnostic w.r.t. the type of inputs it gets as long as
# they support core arithmetic properties. Tell NIFTy to treat our
# parameter dictionary as a vector.
samples = jft.Samples(pos=jft.Vector(lh.init(sk)), samples=None)
delta = 1e-4
absdelta = delta * jft.size(samples.pos)
opt_vi = jft.OptimizeVI(lh, n_total_iterations=25)
opt_vi_st = opt_vi.init_state(
    key,
    # Implicit definition for the accuracy of the KL-divergence
    # approximation; typically on the order of 2-12
    n_samples=lambda i: 1 if i < 2 else (2 if i < 4 else 6),
    # Parametrize the conjugate gradient method at the heart of the
    # sample-drawing
    draw_linear_kwargs=dict(
        cg_name="SL", cg_kwargs=dict(
            absdelta=absdelta / 10.0, maxiter=100
        )
    ),
    # Parametrize the minimizer in the nonlinear update of the
    # samples
    nonlinearly_update_kwargs=dict(
        minimize_kwargs=dict(
            name="SN",
            xtol=delta,
            cg_kwargs=dict(name=None),
            maxiter=5,
        )
    )
)
```

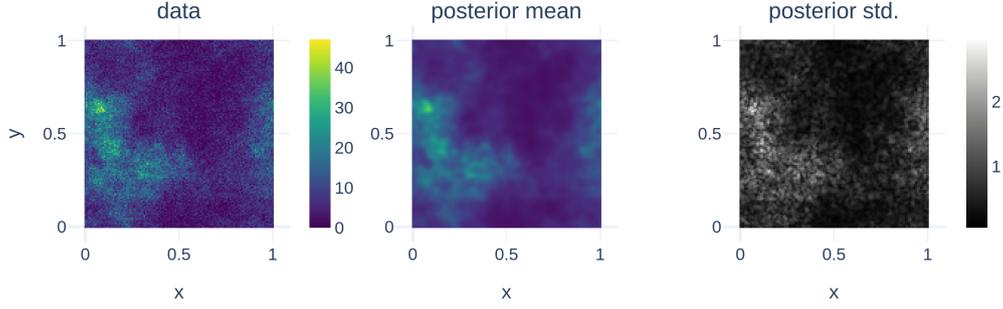


Figure 3.1: Data (left), posterior mean (middle), and posterior uncertainty (right) for a simple toy example.

```

    ),
    # Parametrize the minimization of the KL-divergence cost
    # potential
    kl_kwargs=dict(minimize_kwargs=dict(
        name="M", xtol=delta, maxiter=35
    )),
    sample_mode="nonlinear_resample",
)
for i in range(opt_vi.n_total_iterations):
    print(f"Iteration {i+1:04d}")
    # Continuously update the samples of the approximate posterior
    # distribution
    samples, opt_vi_st = opt_vi.update(samples, opt_vi_st)
    print(opt_vi.get_status_message(samples, opt_vi_st))

```

Figure 3.1 shows an exemplary posterior reconstruction employing the above model. The posterior mean agrees with the data but removes noisy structures. The posterior standard deviation is approximately equal to typical differences between the posterior mean and the data.

### 3.2.4 Performance of NIFTy.re compared to NIFTy

We test the performance of NIFTy.re against NIFTy for the simple yet representative model from above. To assess the performance, we compare the time required to apply  $M_p := F_p + \mathbb{1}$  to random input with  $F_p$  denoting the Fisher metric of the overall likelihood at position  $p$  and  $\mathbb{1}$  the identity matrix. Within NIFTy.re, the Fisher metric of the overall likelihood is decomposed into  $J_{f,p}^\dagger N^{-1} J_{f,p}$  with  $J_{f,p}$  the implicit Jacobian of the forward model  $f$  at  $p$  and  $N^{-1}$  the Fisher-metric of the Poisson likelihood. We choose to benchmark  $M_p$  as a typical VI minimization in NIFTy.re and NIFTy is dominated by calls to this function.

Figure 3.2 shows the median evaluation time in NIFTy of applying  $M_p$  to a new, random tangent position and the evaluation time in NIFTy.re of building  $M_p$  and applying it to a new, random tangent position for exponentially larger models. The 16%-quantiles and the 84%-quantiles of the timings are ob-

### 3. Re-envisioning numerical information field theory (*NIFTy.re*): A library for Gaussian processes and variational inference

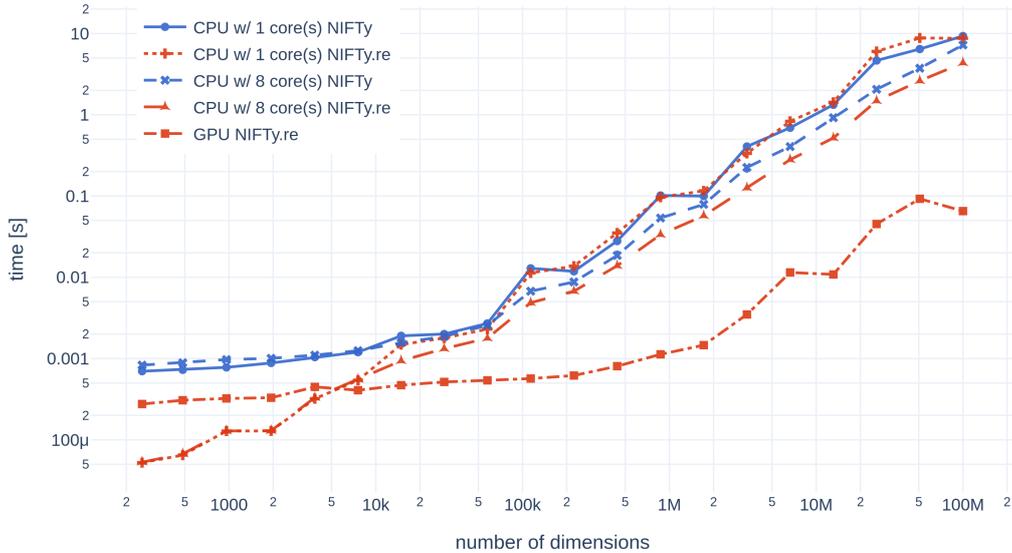


Figure 3.2: Median evaluation time of applying the Fisher metric plus the identity metric to random input for *NIFTy.re* and *NIFTy* on the CPU (one and eight core(s) of an Intel Xeon Platinum 8358 CPU clocked at 2.60G Hz) and the GPU (A100 SXM4 80 GB HBM2). The quantile range from the 16%- to the 84%-quantile is obscured by the marker symbols.

sured by the marker symbols. We choose to exclude the build time of  $M_p$  in *NIFTy* from the comparison, putting *NIFTy* at an advantage, as its automatic differentiation is built around calls to  $M_p$  with  $p$  rarely varying. We ran the benchmark on one CPU core, eight CPU cores, and on a GPU on a compute-node with an Intel Xeon Platinum 8358 CPU clocked at 2.60G Hz and an NVIDIA A100 SXM4 80 GB HBM2 GPU. The benchmark used `jax==0.4.23` and `jaxlib==0.4.23+cuda12.cudnn89`. We vary the size of the model by increasing the size of the two-dimensional square image grid.

For small image sizes, *NIFTy.re* on the CPU is about one order of magnitude faster than *NIFTy*. Both reach about the same performance at an image size of roughly 15,000 pixels and continue to perform roughly the same for larger image sizes. The performance increases by a factor of three to four with eight cores for *NIFTy.re* and *NIFTy*, although *NIFTy.re* is slightly better at using the additional cores. On the GPU, *NIFTy.re* is consistently about one to two orders of magnitude faster than *NIFTy* for images larger than 100,000 pixels.

We believe the performance benefits of *NIFTy.re* on the CPU for small models stem from the reduced Python overhead by just-in-time compiling computations. At image sizes larger than roughly 15,000 pixels, both evaluation times are dominated by the fast Fourier transform and are hence roughly the same as both use the same underlying implementation (Reinecke, 2024). Models in *NIFTy.re* and *NIFTy* are often well aligned with GPU programming models and thus con-

sistently perform well on the GPU. Modeling components such as the new GP models implemented in `NIFTy.re` are even better aligned with GPU programming paradigms and yield even higher performance gains (c.f. Chapter 2).

### 3.3 Conclusion

We implemented the core GP and VI machinery of the Bayesian imaging package `NIFTy` in `JAX`. The rewrite moves much of the heavy-lifting from home-grown solutions to `JAX`, and we envision significant gains in maintainability of `NIFTy.re` and a faster development cycle moving forward. The rewrite accelerates typical models written in `NIFTy` by one to two orders of magnitude, lays the foundation for new types of inference machineries by enabling higher order derivatives via `JAX`, and enables the interoperability of `NIFTy`'s VI and GP methods with the `JAX` machine learning ecosystem.

### Acknowledgements

Gordian Edenhofer acknowledges that support for this work was provided by the German Academic Scholarship Foundation in the form of a PhD scholarship (“Promotionsstipendium der Studienstiftung des Deutschen Volkes”). Philipp Frank acknowledges funding through the German Federal Ministry of Education and Research for the project “ErUM-IFT: Informationsfeldtheorie für Experimente an Großforschungsanlagen” (Förderkennzeichen: 05D23EO1). Jakob Roth acknowledges financial support by the German Federal Ministry of Education and Research (BMBF) under grant 05A20W01 (Verbundprojekt D-MeerKAT). Matteo Guardiani, Vincent Eberle, and Margret Westerkamp acknowledge financial support from the “Deutsches Zentrum für Luft- und Raumfahrt e.V.” (DLR) through the project Universal Bayesian Imaging Kit (UBIK, Förderkennzeichen 500O2103). Lukas Scheel-Platz acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon Europe research and innovation programme under grant agreement No 101041936 (EchoLux).



## 4 A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

*The following chapter is an article in *Astronomy & Astrophysics* led by me (Edenhofer et al., 2023b). This work would not have been possible without the early guidance of Torsten Enßlin, whose rigorous training in statistical methods and discussions throughout my doctoral research contributed in uncountable ways to making this work possible. The final reconstruction profited immensely from regular interactions and discussions with Douglas Finkbeiner and Catherine Zucker. Additionally, Catherine Zucker contributed text toward the introduction. This work also benefited from extensive conversations and feedback from Andrew Saydjari and Philipp Frank. All authors, Catherine Zucker, Philipp Frank, Andrew K. Saydjari, Joshua S. Speagle, Douglas Finkbeiner, and Torsten A. Enßlin, read, commented, and approved the manuscript. The text was adapted to fit the thesis.*

### Abstract

High-resolution 3D maps of interstellar dust are critical for probing the underlying physics shaping the structure of the interstellar medium, and for foreground correction of astrophysical observations affected by dust. We aim to construct a new 3D map of the spatial distribution of interstellar dust extinction out to a distance of 1.25 kpc from the Sun. We leveraged distance and extinction estimates to 54 million nearby stars derived from the *Gaia* BP/RP spectra. Using the stellar distance and extinction information, we inferred the spatial distribution of dust extinction. We modeled the logarithmic dust extinction with a Gaussian process in a spherical coordinate system via iterative charted refinement and a correlation kernel inferred in previous work. In total, our posterior has over 661 million degrees of freedom. We probed the posterior distribution using the VI method MGVI. Our 3D dust map has an angular resolution of up to  $14'$  ( $N_{\text{side}} = 256$ ), and we achieve parsec-scale distance resolution, sampling the dust in 516 logarithmically spaced distance bins spanning 69 pc to 1250 pc. We generated 12 samples from the variational posterior of the 3D dust distribution and release the samples alongside the mean 3D dust map and its corresponding uncertainty. Our map resolves the internal structure of hundreds of molecular

#### 4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

clouds in the solar neighborhood and will be broadly useful for studies of star formation, Galactic structure, and young stellar populations. It is available for download in a variety of coordinate systems online and can also be queried via the publicly available `dustmaps` Python package.

### 4.1 Introduction

Interstellar dust comprises only 1% of the ISM by mass but absorbs and re-radiates  $> 30\%$  of starlight at infrared wavelengths (Popescu and Tuffs, 2002). As such, dust plays an outsized role in the evolution of galaxies, catalyzing the formation of molecular hydrogen, shielding complex molecules from the UV radiation field, coupling the magnetic field to interstellar gas, and regulating the overall heating and cooling of the ISM, see Section 1.1.

Dust’s ability to scatter and absorb starlight is precisely the reason why we can probe it in three spatial dimensions. It preferentially absorbs shorter wavelengths of a stellar spectrum, thus leading to stars behind dense dust clouds appearing reddened relative to their intrinsic colors. The amount by which stars behind dust clouds appear reddened allows us to infer the amount of dust extinction between us and the reddened star. In combination with distance measurements to reddened stars, we can de-project the integrated extinction measurements into a 3D map of differential dust extinction.

*Gaia* has been transformative for the field by providing accurate distance information to more than 2 billion stars, primarily within a few kiloparsecs of the Sun. Precise distances not only improve our knowledge about a star’s position, they also break degeneracies inherent in the modeling of extinction and significantly reduce the extinction uncertainties (Zucker et al., 2019). Thanks to the large quantity of extinction and distance measurements available in the era of large photometric, astrometric, and spectroscopic surveys, we can now probe the 3D distribution of dust in the Milky Way on parsec scales, see Section 1.2.

A number of 3D dust maps that combine *Gaia* and vast photometric and spectroscopic surveys already exist. These maps primarily differ in the way they account for the so-called fingers-of-god effect, or the tendency of dust structures to be smeared out along the LOS. The effect stems from superior constraints on stars’ POS positions relative to their LOS distance uncertainties.

Three-dimensional dust maps predominantly fall into two categories, each representing a trade-off between angular resolution and distance resolution: reconstructions on a Cartesian grid and reconstructions on a spherical grid. Cartesian reconstructions commonly feature less pronounced fingers-of-god but scale poorly with the size of the reconstructed volume. They either encompass a limited volume of the Galaxy (Leike et al., 2020; Leike and Enßlin, 2019) at high resolution or cover a larger volume of the Galaxy at low resolution (Vergely et al., 2022; Lallement et al., 2022, 2019, 2018; Capitanio et al., 2017). Spherical reconstructions often have a much higher resolution and probe larger volumes of the

Galaxy but come with more strongly pronounced fingers-of-god artifacts (Green et al., 2019, 2018; Chen et al., 2019). Alternative approaches using many small reconstructions (Leike et al., 2022), an analytical approach (Rezaei Kh. and Kainulainen, 2022; Rezaei Kh. et al., 2020, 2018, 2017), or inducing point methods (Dharmawardena et al., 2022) have so far been unsuccessful in reconstructing dust at high resolution over large volumes without artifacts.

Physical smoothness priors counterbalance the fingers-of-god effect as finger-like structures are a priori unlikely. In a Cartesian coordinate system it is comparatively easy to incorporate physical priors into the model, such as the distribution of dust being spatially smooth. Smoothness priors are often incorporated using GP priors. Sparsities and symmetries in the prior can be exploited to efficiently apply a GP to a regular Cartesian coordinate system.

Spherical coordinate systems break these sparsities and symmetries in the prior but are much better aligned with the desired spacing of voxels along the LOS. Nearby, voxels can be spaced densely, while at greater distances voxels can be spaced further apart. Naively using a GP prior is infeasible, and approximations either trade fingers-of-god artifacts for other artifacts (Leike et al., 2022) or are too weak to regularize the reconstructions (Green et al., 2019).

In this work, we present a 3D dust map that achieves high distance and angular resolution and probes a large volume of the Galaxy, all at a feasible computational cost. The map uses a new GP prior methodology to incorporate smoothness in a spherical coordinate system, mitigating fingers-of-god artifacts. With a spherical coordinate system we were able to probe dust beyond 1 kpc while still resolving nearby dust clouds at parsec-scale resolution. In Section 4.2 we present the stellar distance and extinction estimates upon which our map is based. In Section 4.3 we present our GP prior methodology for incorporating smoothness in a spherical coordinate system. Section 4.4 describes how we combine the data with our prior model and how we incorporate the distance uncertainties of stars. In Section 4.5 we describe our inference before recapitulating all approximations of the model and their implications in Section 4.6. Finally, in Section 4.7 we present the final map and compare it to existing 3D dust maps and 2D observations.

## 4.2 Stellar distance and extinction data

To construct a 3D dust map, we used the stellar distance and extinction estimates from Zhang et al. (2023), which are primarily based on the *Gaia* BP/RP spectra (spectral resolution  $R \sim 30 - 100$ ). Zhang et al. (2023) adopted a data-driven approach to forward-model the extinction, distance, and intrinsic parameters of each star given the combination of the *Gaia* BP/RP spectra and infrared photometry from the two micron all sky survey (2MASS) and unWISE, a processed catalog based on the wide-field infrared survey explorer (WISE; Carrasco et al., 2021; De Angeli et al., 2023; Gaia Collaboration, 2023a; Montegriffo et al., 2022;

#### 4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

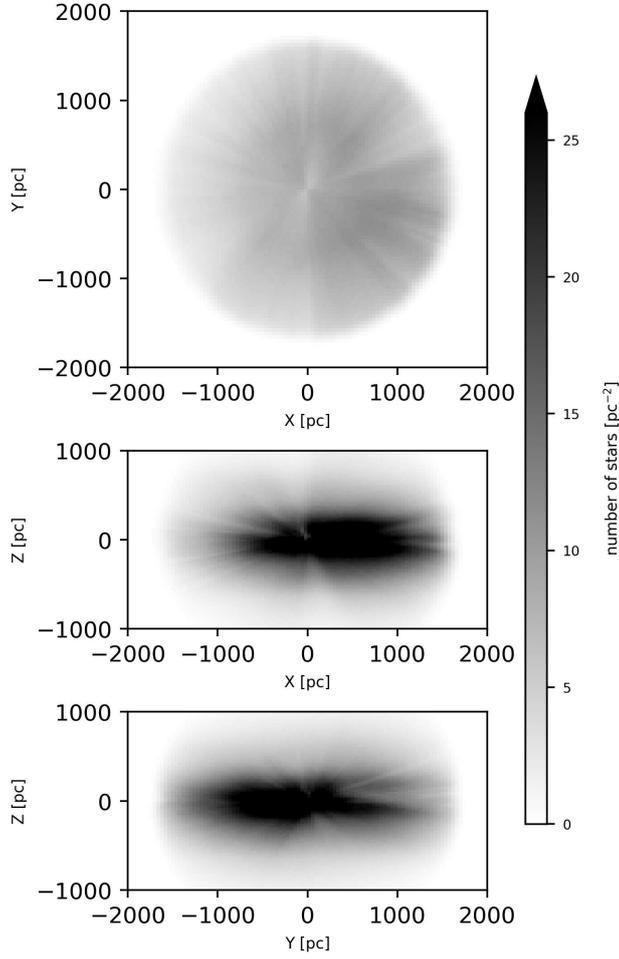
Schlafly et al., 2019; Wright et al., 2010; Skrutskie et al., 2006). The model is trained using a subset of stars with higher-resolution spectra ( $R \sim 1800$ ) available from the large sky area multi-object fibre spectroscopic telescope (LAMOST; Wang et al., 2022; Xiang et al., 2022). The resulting catalog contains distance, extinction, and stellar type ( $T_{\text{eff}}$ ,  $[\text{Fe}/\text{H}]$ ,  $\log g$ ) information for 220 million stars. Throughout this work, we denote the Zhang et al. (2023) catalog by ZGR23.

Compared to other stellar distance and extinction catalogs, the ZGR23 catalog features smaller uncertainties on the extinction estimates while still targeting a significant number of stars. Approximately 87 million ZGR23 stars have an  $A_V$  uncertainty below 60 mmag. Thus, ZGR23 achieves similar extinction uncertainties compared to the subset of 39,538 stars in the StarHorse catalog (Queiroz et al., 2023) that have both higher-resolution spectra from the Apache point observatory galactic evolution experiment (APOGEE) and *grizy* photometry from the panoramic survey telescope and rapid response system (Pan-STARRS), specifically Pan-STARRS1 (PS1; Chambers et al., 2019) (typical  $A_V$  extinction uncertainty of 60 mmag). While the ZGR23 catalog is limited to stars with *Gaia* BP/RP measurements, the quality of the data makes the inference from the ZGR23 catalog competitive with models based on catalogs with larger numbers of stars — 799 million stars in Bayestar19 (Green et al., 2019), 265 million in StarHorse DR2 (Anders et al., 2019), and 362 million in StarHorse EDR3 (Anders et al., 2022). We further find the ZGR23 catalog to have fewer systematic shifts in the extinction and reliable extinction uncertainties based on an analysis in dust-free regions; further details are given in Section 4.9.1.

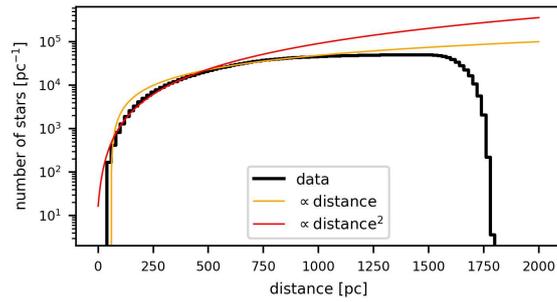
For our reconstruction, we restricted our analysis to ZGR23 stars that have `quality_flags` < 8, as recommended by the authors. We further sub-selected the stars based on their distance. We required  $1/(\omega - \sigma_\omega) < 1.8$  kpc and  $1/(\omega + \sigma_\omega) > 40$  pc with  $\omega$  the parallax of a star and  $\sigma_\omega$  the parallax uncertainty to enforce that all stars are likely within our reconstructed volume. In total, we selected 53,880,655 stars.

The reliability of our reconstruction is predominantly limited by the quality and quantity of the data. Both strongly depend on the POS position and distance. Figure 4.1 shows 2D histograms of stellar density in heliocentric Galactic Cartesian (X, Y, Z) projections, as well as the number of stars as a function of distance. The densities of stars per distance bin first increases approximately quadratically with distance before falling off to a linear increase. At approximately 1.5 kpc the number of stars per distance bin levels off due to our requirement that stars have a >1 sigma chance of being within 1.8 kpc in distance. Figure 4.2 shows a POS histogram of the stars. A clear imprint of the *Gaia* BP/RP selection function is visible (cf. Cantat-Gaudin et al. 2023). A systematic under-sampling of stars behind dense dust clouds is also apparent. We expect our reconstruction to be more trustworthy in regions of higher stellar density. Due to the obscuring effect of dust, regions within and behind dense dust clouds should be treated with more caution.

## 4.2 Stellar distance and extinction data



(a)



(b)

Figure 4.1: 2D histograms of the density of stars in heliocentric Galactic Cartesian ( $X$ ,  $Y$ ,  $Z$ ) projections, as well as the density of stars as a function of distance, for the subset of the ZGR23 catalog used in the reconstruction of our 3D dust map. Panel (a): Heliocentric Galactic Cartesian-projected histograms. Panel (b): Number of stars as a function of distance. This panel also shows a linear growth and a quadratic growth with distance for comparison.

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

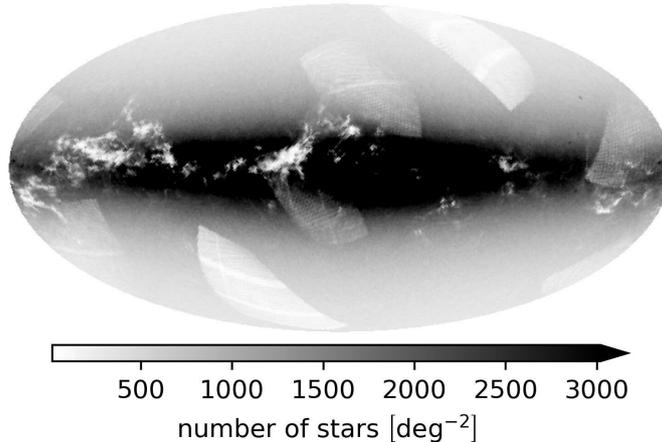


Figure 4.2: POS distribution of the subset of ZGR23 stars used in the reconstruction of our 3D dust map.

### 4.3 Priors

Our quantity of interest is the 3D distribution of differential ZGR23 extinction  $\rho$ .<sup>1</sup> By definition, the differential extinction is positive. Furthermore, we assume it to be spatially smooth. A priori we assume the level of smoothness to be spatially stationary and isotropic.

To reconstruct the 3D volume efficiently, we discretize it in spherical coordinates. Specifically, we discretize our reconstructed volume into HEALPix spheres at logarithmically spaced distances. We adopt an  $N_{\text{side}}$  of 256, which corresponds to 786,432 POS bins. This  $N_{\text{side}}$  corresponds to an angular size of our voxels of  $14'$ . For the LOS direction, we adopt 772 logarithmically spaced distance bins, of which 256 are used for padding. Our highest distance discretization is 0.4 pc and our lowest distance discretization is 7 pc. In contrast to reconstructions with linearly spaced voxels in distance, we are able to probe much larger volumes while maintaining a high sampling at nearby distances. The discretization provides a lower bound on the minimum separation between dust structures that we are able to resolve. In practice the resolvable separation depends on the quantity and quality of the data and varies with the POS and LOS position.

We encode both positivity and smoothness in our model by assuming the differential extinction to be log-normally distributed:

$$\rho = \exp s, \quad (4.1)$$

<sup>1</sup>The ZGR23 extinction is in arbitrary units but can be translated to an extinction at any given wavelength by using the extinction curve published at <https://doi.org/10.5281/zenodo.7692680>. Furthermore, dust extinction can be translated to a rough hydrogen volume density by assuming a constant extinction to hydrogen column density ratio (see, e.g., Zucker et al., 2021).

with normally distributed  $s$ , where  $s$  is drawn from a GP with a homogeneous and isotropic correlation kernel,  $k$ . From previous reconstructions of the differential extinction for the *Gaia* DR2 G-band  $A_G$  (Leike et al., 2020), we have constraints on the correlation kernel of the logarithm of the differential extinction in a volume around the Sun ( $|X| < 370$  pc,  $|Y| < 370$  pc,  $|Z| < 270$  pc). As part of our prior model, we use the inferred  $A_G$  extinction kernel from Leike et al. (2020). To account for the conversion between the ZGR23 extinction and  $A_G$  extinction, we add a global multiplicative factor to  $s$  in our model.<sup>2</sup> Furthermore, we add an additive offset in the differential extinction. We place a log-normal prior on the multiplicative parameter and a normal prior on the additive one.

We enforce the correlation kernel  $k$  using ICR (see Chapter 2). ICR enables us to enforce a kernel on arbitrarily spaced voxels by representing the modeled volume at multiple discretizations. It starts from a very coarse view of our modeled volume. On this coarsest scale, ICR models the GP with learned voxel excitations  $\xi_e^{(0)}$  and an explicit full kernel covariance matrix. A priori the parameters  $\xi_e^{(0)}$  are standard normally distributed and coupled according to  $k$  via ICR. It then iteratively refines  $n_{\text{lvl}}$  times its coarse view of the space with local, fine, a priori standard normally distributed corrections  $\xi_e^{(1)}, \dots, \xi_e^{(n_{\text{lvl}})}$  until reaching the desired discretization. In each refinement, it uses  $n_{\text{csz}}$  neighbors from the previous refinement to refine one coarse pixel into  $n_{\text{fsz}}$  fine pixels.

Iterative charted refinement uses local corrections at varying discretizations and within a refinement assumes the previous iteration to have modeled the GP without error. Both lead to slight errors in representing the kernel. For our use case, we encountered errors in representing the kernel of a few percent. We accepted these errors as a trade-off that enables the reconstruction to probe larger volumes. We refer to Chapter 2 for a detailed discussion of the kernel approximation errors.

Overall, our model for the prior reads

$$\rho = \exp \left[ \text{scl}(\xi_{\text{scl}}) \cdot s \left( \xi_e^{(0)}, \dots, \xi_e^{(n_{\text{lvl}})} \right) + \text{off}(\xi_{\text{off}}) \right], \quad (4.2)$$

where we denote the learned multiplicative scaling of  $s$  by  $\text{scl}$ , the learned additive offset by  $\text{off}$ , and re-expressed both in terms of a priori standard normally distributed parameters  $\xi_{\text{scl}}$  and  $\xi_{\text{off}}$ , respectively. The act of expressing  $\text{scl}$ ,  $\text{off}$  and  $s$  via parameters with an a priori simpler distribution, here a standard normal distribution, is called re-parameterization. A detailed discussion on this subject is given in Rezende and Mohamed (2015).

## 4.4 Likelihood

To construct the likelihood we first need to define how the differential extinction  $\rho$  — our quantity of interest — connects to the measured data  $\mathcal{D}$ . Our data

<sup>2</sup>By doing so (and by using ZGR23) we implicitly assume a spatially stationary reddening law for dust.

#### 4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

comprise POS position, extinction  $\mathcal{D}_A$ , and parallax  $\mathcal{D}_\omega$  data. The POS position is in essence without error. The extinction data  $\mathcal{D}_A = \{A, \sigma_A\}$  are in the form of integrated LOS extinctions to stars  $A$  and associated uncertainties  $\sigma_A$ . The parallax data  $\mathcal{D}_\omega = \{\omega, \sigma_\omega\}$  similarly are in the form of parallax estimates  $\omega$  and uncertainties  $\sigma_\omega$ .

Our model focuses on the measured extinction,  $A$ , and does not predict parallaxes to stars. Instead, we conditioned our model on the parallax data,  $\mathcal{D}_\omega$ , and split the likelihood into the probability of the measured extinction given the true extinction,  $a$ , and the probability of the true extinction given uncertain parallax information:

$$P(A | \rho, \mathcal{D}_\omega) = \int da P(A, a | \rho, \mathcal{D}_\omega) \quad (4.3)$$

$$= \int da P(A | a) \cdot P(a | \rho, \mathcal{D}_\omega) . \quad (4.4)$$

The first term of the integrand is constrained by the quality of the extinction measurements and the second by the quality of the parallax measurements.

##### 4.4.1 Response

The second term in Equation (4.4),  $P(a | \rho, \mathcal{D}_\omega)$ , can be expressed as the joint probability of extinction and true distance,  $d$ , marginalized over the true distance:

$$P(a | \rho, \mathcal{D}_\omega) = \int dd P(a, d | \rho, \mathcal{D}_\omega) \quad (4.5)$$

$$= \int dd P(a | \rho, \mathcal{D}_\omega, d) \cdot P(d | \rho, \mathcal{D}_\omega) . \quad (4.6)$$

We neglect data selection effects (i.e.,  $a$ 's dependence on  $\mathcal{D}_\omega$  given  $d$  and  $d$ 's dependence on  $\rho$  given  $\mathcal{D}_\omega$ ) and use the fact that the true extinction,  $a$ , at known distance  $d$  is simply the LOS integral of  $\rho$  along the LOS to the star from zero to  $d$ :

$$P(a | \rho, \mathcal{D}_\omega) = \int dd P(a | \rho, d) \cdot P(d | \mathcal{D}_\omega) \quad (4.7)$$

$$= \int dd \delta \left( a - \underbrace{\int_0^d d\tilde{d} \rho[\text{POS}](\tilde{d})}_{:=R^d(\rho)} \right) \cdot P(d | \mathcal{D}_\omega), \quad (4.8)$$

with  $\rho[\text{POS}]$  the slice of  $\rho$  at the POS positions of the stars,  $\delta$  the Dirac delta distribution defined by  $\int_{-\infty}^{\infty} dx f(x)\delta(x) = f(0)$  for any continuous  $f$  with compact support, and  $R$  the response that maps from  $\rho$  to the domain of the measured extinction.

We approximated  $P(a|\rho, \mathcal{D}_\omega)$  with a normal distribution,

$$P(a|\rho, \mathcal{D}_\omega) \approx \mathcal{G}(a|\bar{a}, \sigma_a^2), \quad (4.9)$$

with mean  $\bar{a}$  and standard deviation  $\sigma_a$  to obtain a tractable expression for Equation (4.4). The mean extinction,  $\bar{a}$ , is

$$\bar{a} := \langle a \rangle_{P(a|\rho, \mathcal{D}_\omega)} \quad (4.10)$$

$$= \int da a \int dd \delta(a - R^d(\rho)) \cdot P(d|\mathcal{D}_\omega) \quad (4.11)$$

$$= \int \int da dd a \cdot \delta(a - R^d(\rho)) \cdot P(d|\mathcal{D}_\omega) \quad (4.12)$$

$$= \int dd R^d(\rho) \cdot P(d|\mathcal{D}_\omega) \quad (4.13)$$

$$= \langle R^d(\rho) \rangle_{P(d|\mathcal{D}_\omega)}. \quad (4.14)$$

Assuming the parallax  $1/d$  is normally distributed (i.e.,  $P(d|\mathcal{D}_\omega) = \mathcal{G}(1/d|\omega, \sigma_\omega^2)$  with mean  $\omega$  and standard deviation  $\sigma_\omega$ ), then

$$\begin{aligned} \langle a \rangle_{P(a|\rho, \mathcal{D}_\omega)} &= \langle R^d(\rho) \rangle_{\mathcal{G}(1/d|\omega, \sigma_\omega^2)} \\ &= \int_0^\infty d\tilde{d} \rho[\text{POS}](\tilde{d}) \cdot \text{sf}_{\mathcal{G}}(1/\tilde{d}|\omega, \sigma_\omega^2), \end{aligned} \quad (4.15)$$

with  $\text{sf}_{\mathcal{G}}(1/d|\omega, \sigma_\omega^2) := 1 - \int_{-\infty}^{1/d} d\omega' \mathcal{G}(\omega'|\omega, \sigma_\omega^2)$  the survival function of the normal distributed parallax.

The standard deviation  $\sigma_a$  can be understood as an additional error contribution for marginalizing over the distance. The error depends on the distance uncertainty and the dust along the full LOS:

$$\sigma_a^2 := \left\langle \left( R^d(\rho) \right)^2 \right\rangle_{\mathcal{G}(1/d|\omega, \sigma_\omega^2)} - \langle R^d(\rho) \rangle_{\mathcal{G}(1/d|\omega, \sigma_\omega^2)}^2. \quad (4.16)$$

Evaluating both  $\bar{a}$  and  $\sigma_a^2$  is comparatively cheap in a spherical coordinate system since for a discretized sphere  $R^d(\rho)$  is simply the cumulative sum of  $\rho$  along the distance axis weighted by the radial extent of each voxel.

#### 4.4.2 Likelihood and joint probability density

We assume the measured extinction to be normally distributed around the true extinction  $a$ . We take the inferred extinction,  $A$ , from the catalog to be the mean of the normal distribution. The accompanying uncertainty  $\sigma_A$  in the catalog is assumed to be the standard deviation of  $P(A|a)$ .

Some stars will have underestimated uncertainties due to either mismodeled intrinsic stellar properties in the inference or bad photometric measurements that

#### 4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

were not flagged. We want our model to be able to detect and deselect stars that are in strong disagreement with the rest of the reconstruction. We achieve this by inferring an additional multiplicative factor per star,  $n_\sigma$ , which scales  $\sigma_A$ . A priori, we assume  $n_\sigma$  to be drawn from a heavy-tailed distribution. Specially, we assume  $n_\sigma$  to follow an inverse gamma distributed. We again express  $n_\sigma$  in terms of standard normally distributed parameters  $n_\sigma(\xi_\sigma)$  in the inference.

To summarize, our approximate likelihood first introduced in Equation (4.4), reads

$$P(A | \rho, n_\sigma, \mathcal{D}_\omega) \approx \int da \mathcal{G} \left( A \mid a, (n_\sigma \cdot \sigma_A)^2 \right) \cdot \mathcal{G} \left( a \mid \bar{a}(\rho), \sigma_a^2(\rho) \right) \quad (4.17)$$

$$= \mathcal{G} \left( A \mid \bar{a}(\rho), [n_\sigma \cdot \sigma_A]^2 + \sigma_a^2(\rho) \right). \quad (4.18)$$

The uncertainty in the extinction  $\sigma_A$  is scaled by  $n_\sigma$  to deselect outliers and increased by  $\sigma_a^2$  due to marginalizing over the distance uncertainty.

The joint probability density function of data and parameters reads

$$P(A, \rho(\xi), n_\sigma(\xi) | \mathcal{D}_\omega) = \mathcal{G} \left( A \mid \bar{a}(\rho(\xi)), [n_\sigma(\xi) \cdot \sigma_A]^2 + \sigma_a^2(\rho(\xi)) \right) \cdot \mathcal{G}(\xi | 0, 1), \quad (4.19)$$

with  $\xi$  the vector of all parameters of the model  $\{\xi_e^{(0)}, \dots, \xi_e^{(n_{\text{lvl}})}, \xi_{\text{scl}}, \xi_{\text{off}}, \xi_\sigma\}$ . The complexity of the prior distributions has been fully absorbed into the transformations  $s(\xi)$ ,  $\text{scl}(\xi)$ ,  $\text{off}(\xi)$ , and  $n_\sigma(\xi)$  from the a priori standard normally distributed parameters  $\xi$ .

Table 4.1: Parameters of the prior distributions. The parameters  $s$ ,  $\text{scl}$ , and  $\text{off}$  fully determine  $\rho$ . They are jointly chosen to a priori yield the kernel reconstructed in Leike et al. (2020).

Name	Distribution	Mean	Standard Deviation	Degrees of Freedom
$s$	Normal	0.0	Kernel from Leike et al. (2020)	$786,432 \times 772$
$\text{scl}$	Log-Normal	1.0	0.5	1
$\text{off}$	Normal	$-6.91 (\approx \ln 10^{-3})$ prior median extinction from Leike et al. (2020)	1.0	1
		Shape Parameter	Scale Parameter	
$n_\sigma$	Inverse Gamma	3.0	4.0	# Stars = 53,880,655

Our priors in terms of nonstandard-normal parameters are summarized in Table 4.1. The priors for  $s$ ,  $\text{scl}$ , and  $\text{off}$  are chosen to a priori yield the kernel reconstructed in Leike et al. (2020). In contrast to Leike et al. (2020), we do not learn a full non-parametric kernel. However, we do infer  $\text{scl}$  and  $\text{off}$ , the scale, and zero-mode of the kernel. The prior for  $n_\sigma$  was chosen such that the inverse gamma distribution has mode 1 and standard deviation 2.

## 4.5 Posterior inference

In the previous section we took special care to express our model not only in terms of physical parameters, like the differential extinction density  $\rho$ , but also in terms of simpler parameters  $\xi$ . The act of expressing the parameters of the model  $s_{cl}$ ,  $off$ ,  $s$ , and  $n_\sigma$  in terms of a priori standard normal distributed variables  $\xi$  is called standardization, a special form of re-parameterization (see Rezende and Mohamed, 2015). Effectively, we are shifting complexity from the prior to the likelihood. However, both the nonstandardized and the standardized formulation of the joint model are equivalent. Standardizing models can lead to better conditioned inference problems as the parameters all vary on the same scales — if the prior is not in conflict with the likelihood. We use an inference scheme that relies on the standardized formulation.

We want to infer the posterior for our standardized model from Equation (4.19). Directly probing the posterior via sampling methods such as Hamiltonian Monte Carlo Hoffman and Gelman (2014) is computationally infeasible. Instead, we used VI to approximate the true posterior. Specifically, we used MGVI (Knollmüller and Enßlin, 2019). We summarize the main idea behind MGVI in Section 4.9.2. We did not approximate the posterior of the noise inference parameter  $n_\sigma(\xi_\sigma)$  via VI and instead used only the maximum of the posterior for  $\xi_\sigma$ .

To speed up the inference, we started the reconstruction at a lower resolution (196, 608 POS bins at  $N_{side} = 128$  and 388 LOS distance bins) and restricted the inference to a subset of stars with a  $\geq 2$  sigma chance of being within 600 pc and a  $\geq 2$  sigma chance of being farther than 40 pc. We successively increased the distance range of the map up to which stars are incorporated in steps of 300 pc from 600 pc to 1.8 kpc. Every time we increase the distance range, we reset the parameters for  $n_\sigma$ . Then, after all data was incorporated, we increased the angular and distance resolution of the reconstruction to the final resolution.

Our data selection deselects stars close to the maximum distance probed (c.f. Figure 4.1). This effect leads to the outer regions of the map being informed by relatively few stars compared to the inner regions. We observe that these regions are prone to producing spurious features. For our final data products, we removed the outermost 550 pc from the data-constrained volume as we observed artifacts aligned with our data incrementation strategy within these regions. We believe 550 pc to be a conservative cut but we advise caution when finding structures perfectly aligned with a sphere around the Sun at 600 pc, 900 pc, or 1200 pc.

ZGR23 assumes all extinctions to be strictly positive. We neglected this constraint by assuming Gaussian errors, which led to an artificial spike in extinction in the first few voxels in each direction. As we know those regions to be effectively free of dust from previous reconstructions c.f. Leike et al. (2020), we removed the innermost 69 pc (see Section 4.9.3). We release an additional HEALPix map of integrated extinction out to 69 pc from the Sun and suggest using it to correct integrated LOS predictions for the removed extinction.

#### 4. *A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun*

Our inference heavily utilizes derivatives of various components of our model. Derivatives are used for the minimization as well as for the variational approximation of the posterior. Previous models such as those described in Leike and Enßlin (2019) and Leike et al. (2020) relied on the Numerical Information Field Theory (NIFTy) package (Selig et al., 2013; Steininger et al., 2019; Arras et al., 2019) and were limited to running on CPUs.

We employed a new framework called NIFTy.re (see Chapter 3) for deploying NIFTy models to GPUs. NIFTy.re is part of the NIFTy Python package and internally uses JAX (Bradbury et al., 2018) to run models on the GPU. We were able to speed up the evaluation of the value and gradient of Equation (4.19) by two orders of magnitude by transitioning from CPUs to GPUs. Our reconstruction ran on a single NVIDIA A100 GPU with 80 GB of memory for about four weeks.

## 4.6 Caveats

We believe statistical uncertainties are the dominant source of uncertainty for our reconstruction. However, it is important to also consider sources of systematic uncertainties. Depending on the application, the systematic uncertainties may be more important than the statistical uncertainties. The data that informed the reconstruction, the model with which we inferred it, and the inference procedure all contribute to the model’s systematic uncertainties.

Naturally, the data themselves are a source of systematic uncertainties (spatially stationary reddening law, mismodeling of binaries, etc.; see Zhang et al., 2023) and additionally is known to be incomplete, c.f. Figure 4.2. A lower stellar density, for example in heavily obscured regions, limits the map’s resolution and results in volumes of the map behind dense dust clouds being poorly constrained. Thus, we believe our dust reconstruction to be an underestimation of the true extinction toward dense dust clouds. Zucker et al. (2021) also note this effect when comparing the Leike et al. (2020) map with 2D integrated extinction maps based on infrared photometry, finding that the Leike et al. (2020) is not sensitive to regions with  $A_V \gtrsim 2$  mag.

We advise visualizing the stellar density in the region of interest to assess the magnitude of the systematic uncertainties due to data incompleteness. We release all stars used in the reconstruction as an additional data product. This data product can be used to visualize the stellar density. In regions with a significant underdensity of stars, we expect the systematic lack of stars to dominate the statistical uncertainties in the reconstruction. The reconstruction produced plausible infills in those regions based on adjacent stars. However, the resulting uncertainties do not capture the cause for the systematic lack of stars as the model implicitly assumes that underdensities of stars are not systematic but purely random.

Our model includes a number of approximations. First, we assumed a GP

prior on the logarithmic dust extinction using the kernel from Leike et al. (2020) and additionally only applied it approximately via ICR. Second, we assumed  $\mathcal{D}_\omega$  to be independent of  $\mathcal{D}_A$ . Third, we assumed the parallax error to be Gaussian, and fourth, we assumed the extinction error to be Gaussian.

For extremely low extinctions, the assumption of  $A$  being Gaussian is poor due to the positivity prior in the ZGR23 catalog. We corrected for this bias toward higher estimated extinction in regions with assumed extremely low true extinctions post-hoc by cutting away the innermost 69 pc as described in Section 4.5. We publish an auxiliary map of integrated extinction out to 69 pc from the Sun to correct integrated LOS predictions for the removed extinction. We suggest adding the removed local extinction back to the map when comparing integrated extinctions. By default, the removed local extinction is added back to the map when querying integrated extinctions via `dustmaps`.

We further release a catalog of the predicted extinctions of our model to all stars that we used for the reconstruction. In Section 4.9.4, we perform a non-exhaustive consistency test comparing our predictions to the ZGR23 ones. We find that both predictions for the extinction to stars disagree below 50 mmag and above 4 mag; 34% more stars than expected have larger (respectively smaller) extinction predictions compared to ZGR23. We expect the range from 50 mmag to 4 mag to correspond to the range within which our map is reliable. More details are provided in Section 4.9.4.

Furthermore, our posterior inference is an approximation. We assume our approximation of the true posterior accurately captures the intrinsic model uncertainties (c.f. Galan et al., 2024; Arras et al., 2022; Leike and Enßlin, 2019; Leike et al., 2020; Mertsch and Phan, 2023; Roth et al., 2023a; Hutschenreuter et al., 2023; Tsouros et al., 2024b; Roth et al., 2023b; Hutschenreuter et al., 2022). However, we worried about structures getting burned in when we increase the maximum distance probed during the inference from 600 pc to 1800 pc in steps of 300 pc as described in Section 4.5. We checked the final reconstruction for this effect by comparing it against a larger reconstruction that does not sub-select the stars based on their distance during the inference but uses only a small subsample of ZGR23 stars with more stringent quality flags. The larger reconstruction, which extends out to 2 kpc in distance, is released as an additional data product. We find no significant differences between both runs. Details on the larger reconstruction are provided in Section 4.9.5.

## 4.7 Results

We reconstructed 12 samples (6 antithetically drawn samples) of the 3D dust extinction distribution, each of which encompasses 607,125,504 differential extinction voxels. The voxels are arranged on 772 HEALPix spheres with  $N_{\text{side}} = 256$  spaced at logarithmically increasing distances. After removing the innermost <69 pc and outermost >1250 pc HEALPix spheres, we are left with 516 HEALPix

#### 4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

spheres. The samples, the posterior mean, and the posterior standard deviation for the reconstruction are publicly available online<sup>3</sup>. We strongly advise using the samples for any quantitative analysis. For convenience, we also provide the posterior mean and standard deviation of the reconstruction interpolated to heliocentric Galactic Cartesian Coordinates ( $X, Y, Z$ ) and Galactic spherical Coordinates ( $l, b, d$ ) as well as the scripts for the interpolation. Furthermore, the map can be queried via the `dustmaps` Python package (Green, 2018). Further details on using the reconstruction are given in Section 4.9.6.

The distance discretization in our reconstruction is highest for close-by voxels and decreases further out. Our highest distance discretization is 0.4 pc and our lowest distance discretization is 7 pc while our angular discretization is  $14'$  and is independent of the distance. The stated discretizations specify the lower bound on our resolution. The minimum separation that we are able to resolve depends on the position and is encoded in the posterior samples. For small regions, we suggest additionally analyzing the stellar density (see Section 4.6) to assess the strength of the systematic uncertainties due to the density of stars.

The reconstruction is in terms of the unitless ZGR23 extinction as defined in Zhang et al. (2023). For visualization purposes, we translated the ZGR23 extinction to Johnson’s V-band  $\lambda = 540.0$  nm, that is to say,  $A_V := A(V = 540.0 \text{ nm})$ . To perform the conversion, we adopt the extinction curve published in ZGR23 and multiplied the unitless ZGR23 extinction by a factor of 2.8. We refer readers to the full extinction curve<sup>4</sup> from Zhang et al. (2023) for the coefficients needed to translate the extinction to other bands.

Figure 4.3 depicts the POS projection of the posterior mean reconstruction integrated out to 250 pc, 500 pc, 750 pc, and up to the end of our sphere. The  $A_V$  values are in units of magnitudes. We see that higher-latitude features like the Aquila Rift are comparatively close-by while structures in the Galactic plane appear only gradually. Figure 4.4 shows the difference between the integrated POS projections. We recover well-known features of integrated dust but are now able to de-project them.

Figure 4.5 show a bird’s eye ( $X, Y$ ), side-on ( $X, Z$ ), and ( $Y, Z$ ) projection of the posterior mean of our reconstruction in heliocentric Galactic Cartesian coordinates. The image depicts the innermost  $2.5 \text{ kpc} \times 2.5 \text{ kpc} \times 0.8 \text{ kpc}$  around the Sun in  $A_V$  extinction integrated over  $z$  from  $-400 \text{ pc}$  to  $400 \text{ pc}$ ,  $y$  in  $-1.25 \text{ kpc}$  to  $1.25 \text{ kpc}$ , and  $x$  in  $-1.25 \text{ kpc}$  to  $1.25 \text{ kpc}$ , respectively. In Figure 4.6 we overlay a catalog of clusters of young stellar objects (YSOs; Kuhn, 2023) based on Kuhn et al. (2021), Winston et al. (2020), and Marton et al. (2023), which are shown as blue dots. The positions of the YSO clusters visually agree with the positions of dust clouds within the YSO clusters’ reported distance uncertainties.

The posterior standard deviation divided by the posterior mean of the reconstruction is shown in Figure 4.7. The map features a faint speckle pattern. This is likely due to the low number of samples relative to the number of degrees of

<sup>3</sup><https://doi.org/10.5281/zenodo.8187942>

<sup>4</sup><https://doi.org/10.5281/zenodo.7692680>

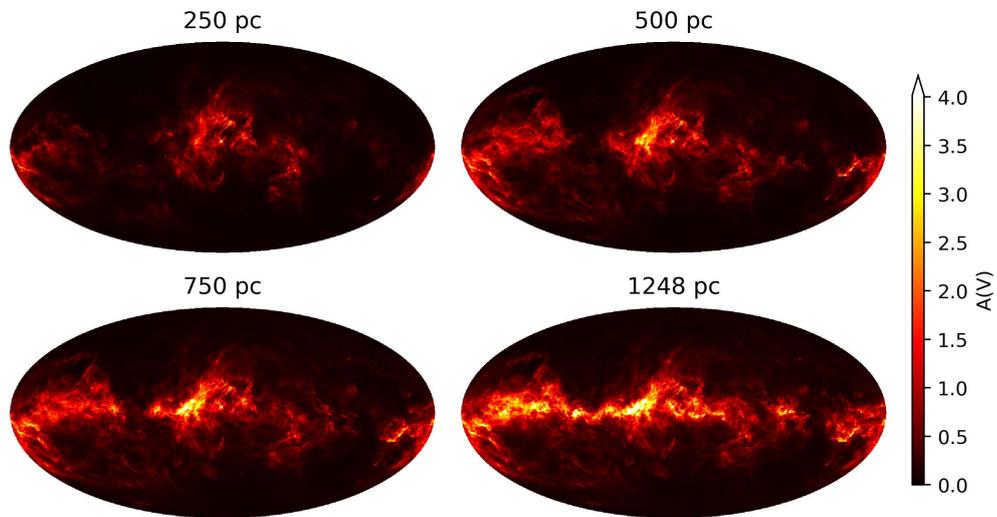


Figure 4.3: Mollweide projection of the POS integrated  $A_V$  extinction out to 250 pc, 500 pc, 750 pc, and up to the maximum distance of our map. The colorbar saturates at the 99.9% quantile.

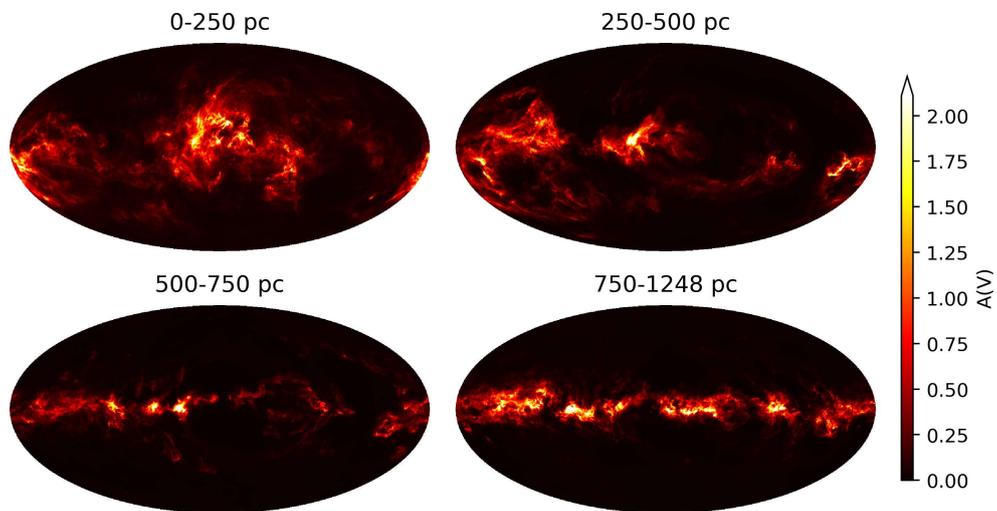


Figure 4.4: Same as Figure 4.3 but showing the difference between the integrated extinctions in between distance slices projected on the POS. The colorbar saturates at the 99.9% quantile.

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

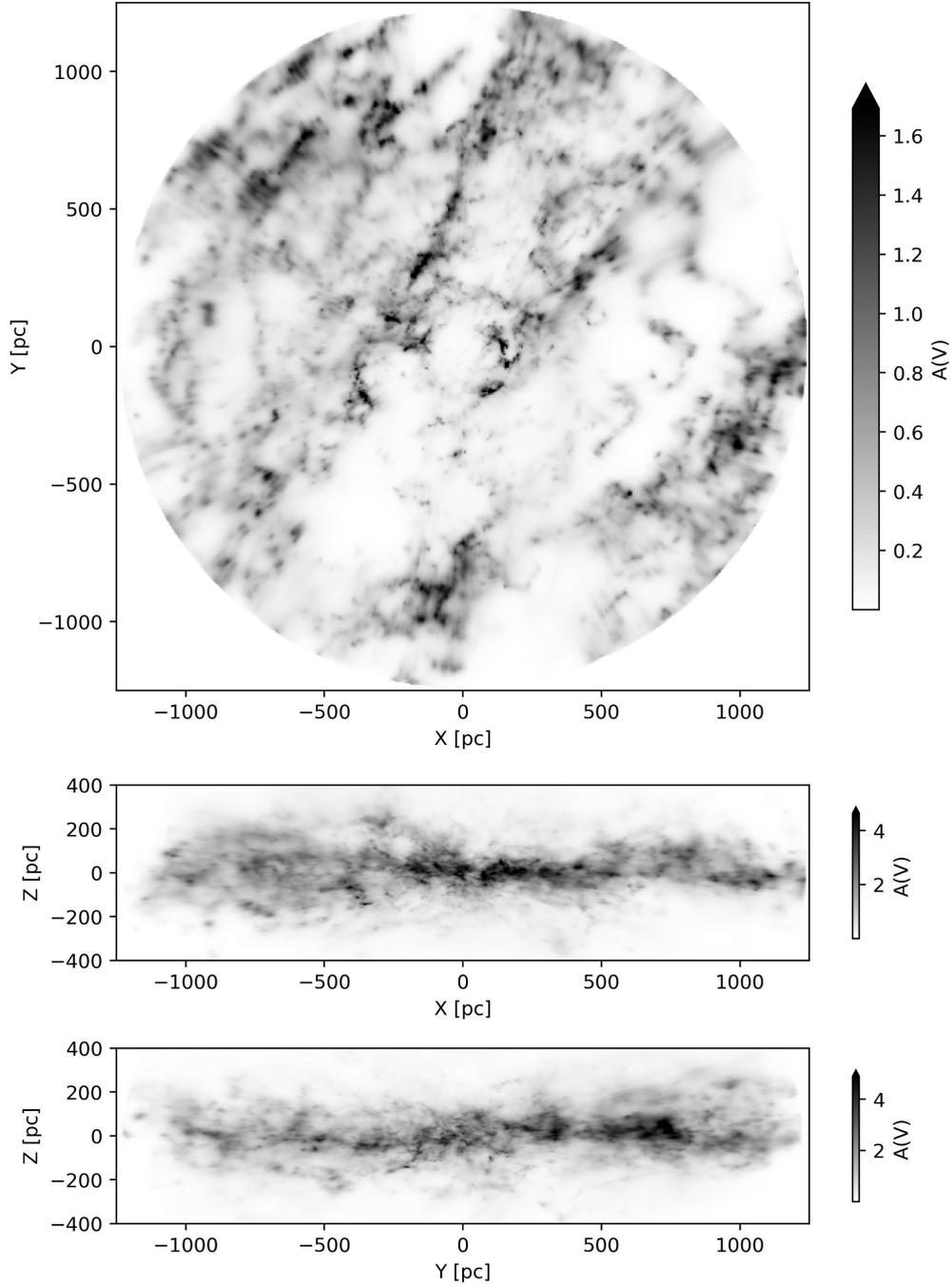


Figure 4.5: X-, Y-, Z-projections of the posterior mean of our 3D dust map in a box of  $2.5 \text{ kpc} \times 2.5 \text{ kpc} \times 0.8 \text{ kpc}$  centered on the Sun. The colorbar is linear and saturates at the 99.9% quantile. A GIF of the posterior samples is shown at [https://faun.rc.fas.harvard.edu/gedenhofer/perm/E+23/21b9\\_final.gif](https://faun.rc.fas.harvard.edu/gedenhofer/perm/E+23/21b9_final.gif). A low-resolution 3D interactive figure is available at [https://faun.rc.fas.harvard.edu/czucker/Paper\\_Figures/3D\\_Dust\\_Edenhofer2023.html](https://faun.rc.fas.harvard.edu/czucker/Paper_Figures/3D_Dust_Edenhofer2023.html).

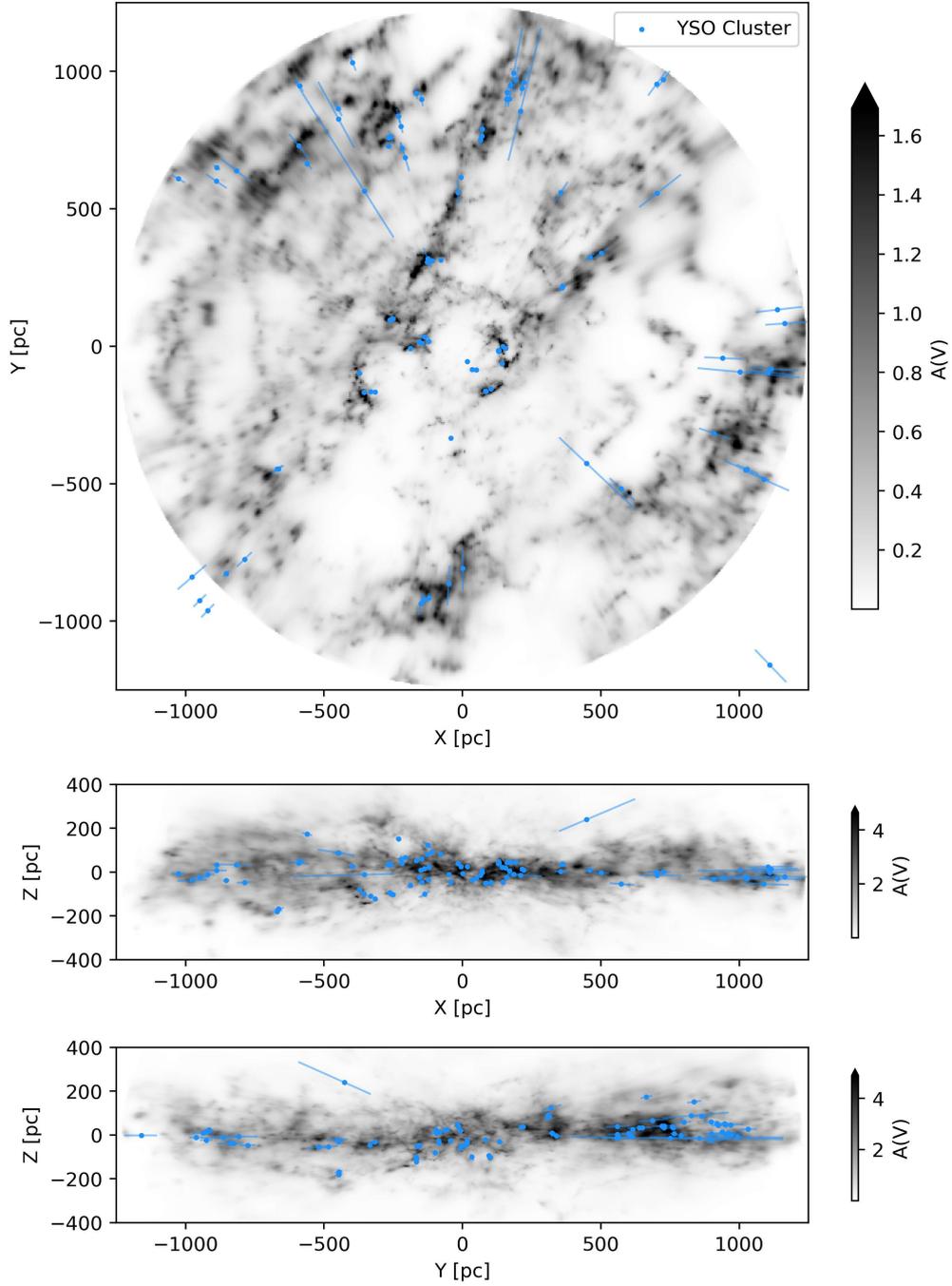


Figure 4.6: Same as Figure 4.5 but with a catalog of clusters of YSOs (Kuhn, 2023) based on Kuhn et al. (2021), Winston et al. (2020), and Marton et al. (2023) shown as blue dots on top of the reconstruction; their distance uncertainties are shown as extended lines.

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

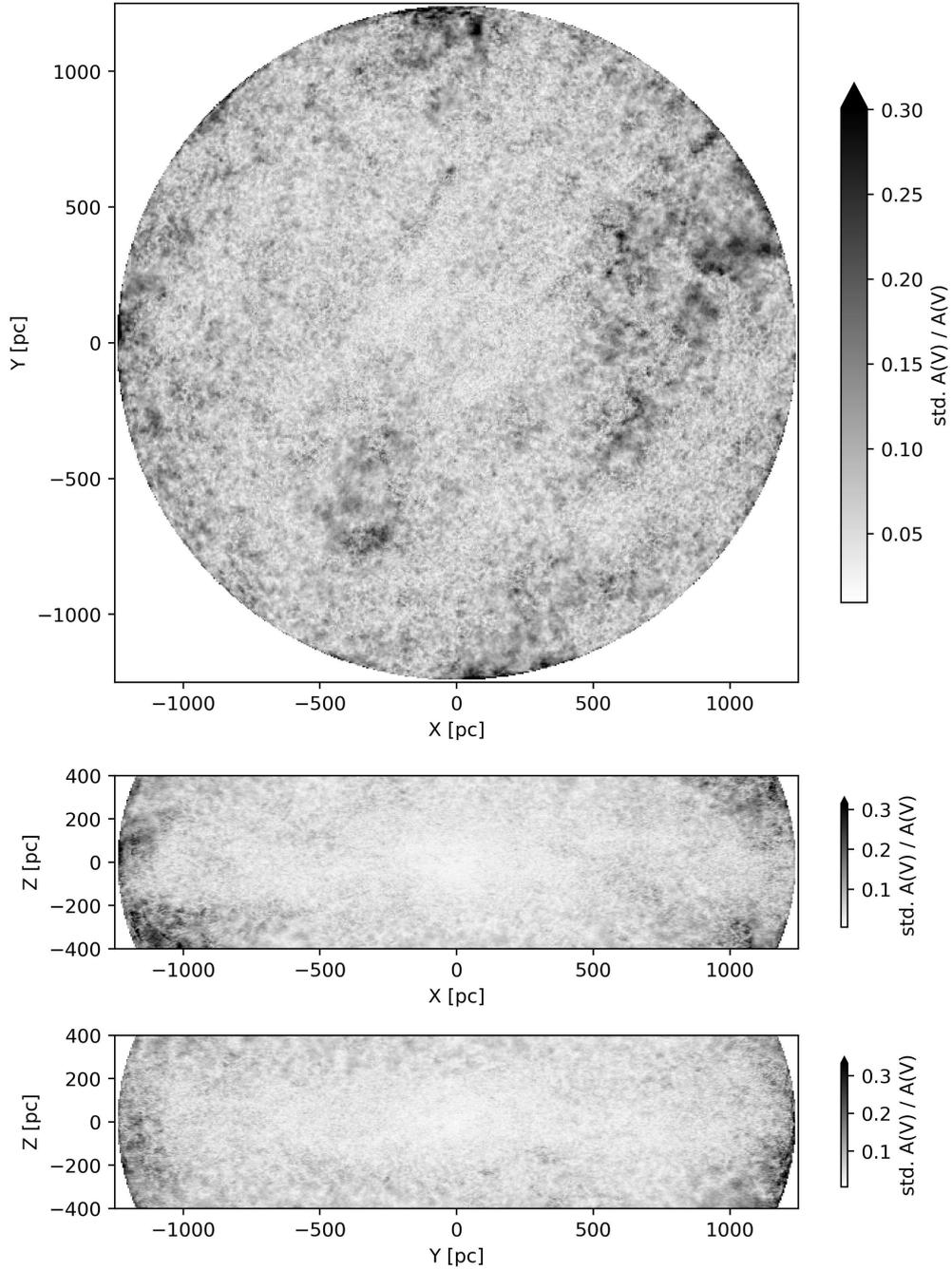


Figure 4.7: Heliocentric Galactic Cartesian (X, Y, Z) projections of the relative uncertainty of the reconstructed dust extinction integrated within a box of  $2.5 \text{ kpc} \times 2.5 \text{ kpc} \times 0.8 \text{ kpc}$  centered on the Sun. The colorbar is linear and saturates at the 99.9% quantile.

freedom. The standard deviation is on the order of 10% of the posterior mean and slightly increases with distance. Toward the galactic center behind the dust in the immediate vicinity of roughly 300 pc, the relative uncertainty is noticeably higher.

The reconstruction has a high dynamic range and reveals faint dust lanes in the reconstructed volume. Small approximately spherical cavities are evident throughout the map. The dust clouds in the reconstruction are compact and only weakly elongated radially. Prominent large-scale features, such as the Radcliffe Wave (Alves et al., 2020) and the Split (Lallement et al., 2019), have been resolved at an unprecedented level of detail, previously only accessible for the most nearby dust clouds.

### 4.7.1 Comparison to existing 3D dust maps

In this section, we compare our map to other 3D dust maps in the literature. We denote the dust map described in Leike et al. (2020) by LGE20, Vergely et al. (2022) by VLC22, Green et al. (2019) by Bayestar19, and Leike et al. (2022) by L+22 in this section. For the purposes of comparison, we show the posterior mean. We release the statistical uncertainties as additional data products, and we strongly advise taking into account the released statistical uncertainties for any quantitative analysis. However, the differences between the various 3D dust reconstructions discussed here are systematic differences and are not captured by the reconstructed statistical uncertainties.

In Figure 4.8, we show 3D (X, Y, Z) projections of the maps, comparing Bayestar19, VLC22, L+22, and this work side by side. All four maps agree on the general structure of the distribution of dust.

This work, L+22, and VLC22 have comparable distance resolutions, while Bayestar19 features comparatively few distance bins and more strongly pronounced fingers-of-god. Compared to L+22, we feature more homogeneously extended dust clouds and significantly fewer wiggles in the distances to dust clouds. Compared to VLC22, we feature more compact dust clouds, less grainy structures, and a higher dynamic range. Both this work and VLC22 feature dust clouds in a comparable volume around the Sun despite the VLC22 map technically extending out farther in Galactic heliocentric X and Y.

Figure 4.9 shows the same projections for the volume reconstructed in the LGE20 map and includes the LGE20 map for comparison. The zoom-in highlights the close similarity between this work and the LGE20 map. All larger structures have direct correspondences in the other map, yet the distances to the structures are slightly different. Furthermore, the LGE20 map appears slightly sharper. The model in LGE20 is very similar to ours but uses fewer approximations. LGE20 also uses compiled data (StarHorse DR2; see Anders et al., 2019). More work is needed to assess the validity of the sharper features in LGE20 not present in this work. The VLC22 map is in good agreement as well but lower resolution. Bayestar19 poorly resolves distances at the scale of the LGE20 map.

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

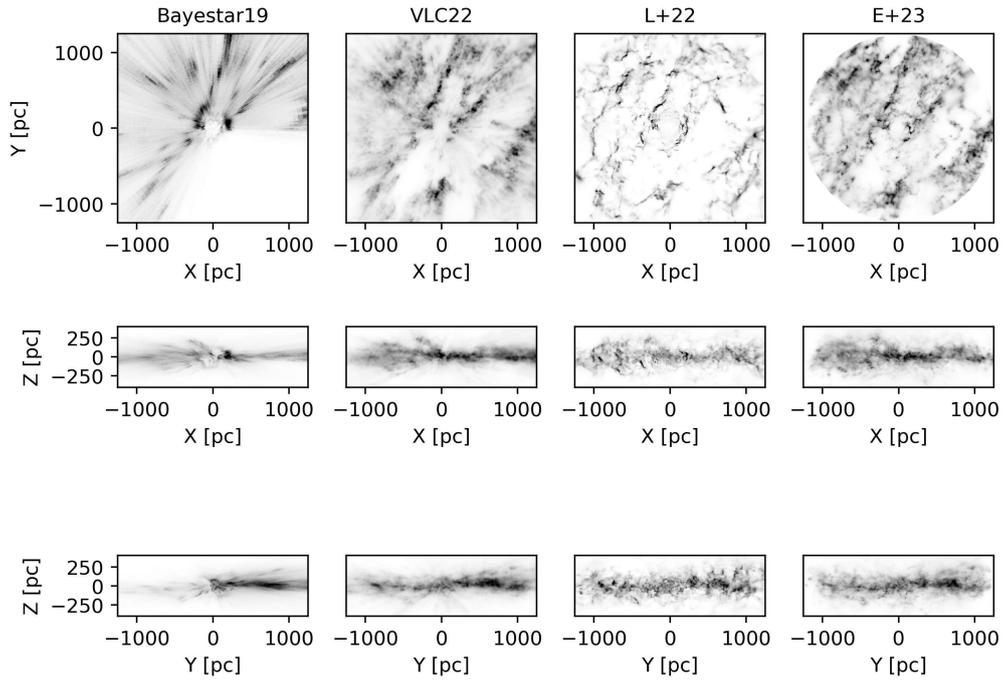


Figure 4.8: Side by side views of the 3D dust maps from Bayestar19, VLC22, L+22, and this work, shown in heliocentric Galactic Cartesian ( $X$ ,  $Y$ ,  $Z$ ) projections. The colorbars are saturated at the 99.9% quantile of the respective reconstruction.

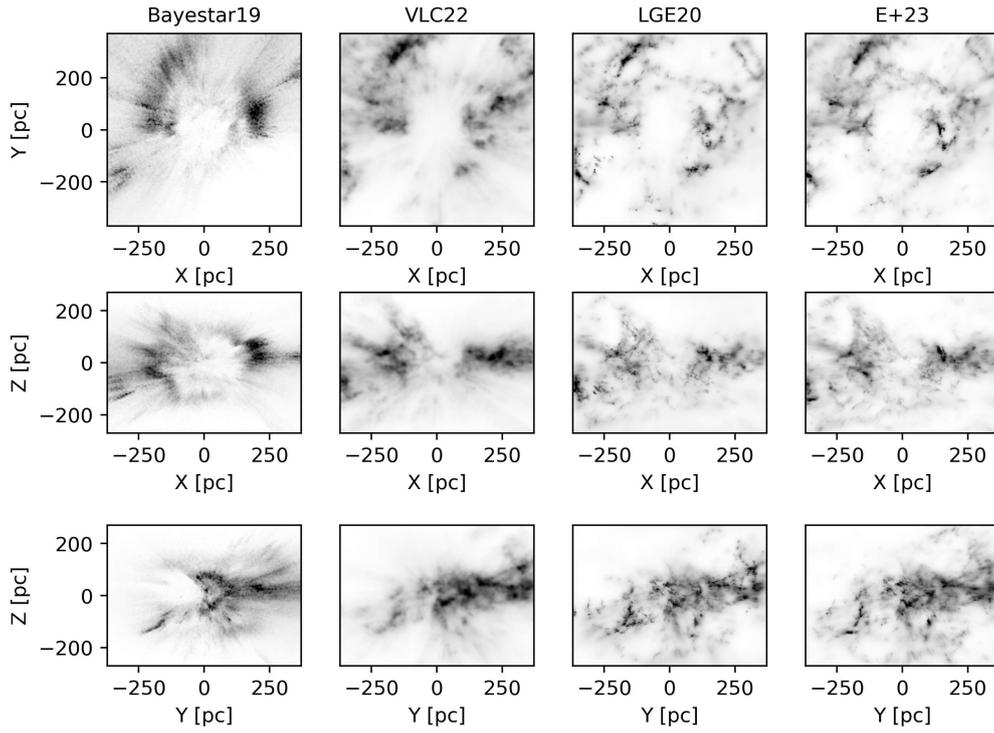


Figure 4.9: Zoomed-in version of Figure 4.8 for the volume reconstructed in Leike et al. (2020), now also showing the LGE20 reconstruction for comparison. We omit L+22 from the comparison because the authors explicitly focus on larger volumes and trade strongly pronounced artifacts in the inner couple hundred parsecs for a larger probed volume. The colorbars are again saturated at the 99.9% quantile of the respective reconstruction.

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

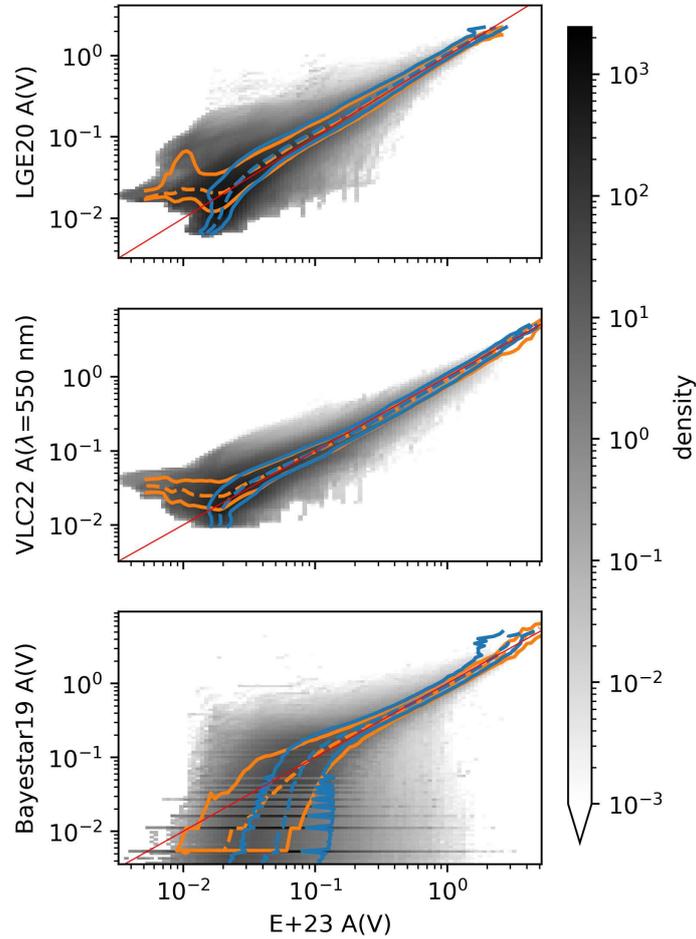


Figure 4.10: Histogram of the mean posterior extinction of our map versus LGE20, VLC22, and Bayestar19 for 58 million test points. For each pixel center of a HEALPix sphere with  $N_{\text{side}} = 64$ , 1182 test points are placed at 1 pc intervals in distance starting at 69 pc. The orange lines show the 16th, 50th, and 84th quantiles of the predicted extinction by LGE20, VLC22, and Bayestar19, respectively, for each bin of our mean extinction. The respective quantiles of our predictions in bins of the other reconstruction are shown as blue lines. The bisectors are shown in red. The colorbars are logarithmic and truncated at the lower end at 1 mmag.

Figure 4.10 quantitatively compares our map with LGE20, VLC22, and Bayestar19. We compare the integrated extinction for each pixel center of a HEALPix sphere with  $N_{\text{side}} = 64$  at 1182 test points that are placed at 1 pc intervals in distance starting at 69 pc. We compared integrated extinctions because slight shifts in the distances to extincted regions can be better distinguished from discrepant predictions with integrated extinctions than by directly comparing differential extinctions. Above 15 mmag, VLC22 and LGE20 are in good agreement with our map. Below an  $A_V$  of 40 mmag and above an  $A_V$  of approximately 3.5 mag, Bayestar19 significantly deviates from our predictions.

In Figure 4.11 we compare the POS view of Bayestar19, L+22, VLC22, LGE20, and this work. The respective POS views are integrated out to the maximum distance probed by each map —  $< 63$  kpc in distance for Bayestar19 (maximum reliable distance 10 kpc),  $< 4$  kpc in distance for L+22 (authors trust structures up to 4 kpc though the map extends to 16 kpc), a heliocentric box of size  $3 \text{ kpc} \times 3 \text{ kpc} \times 800 \text{ pc}$  with  $10^3 \text{ pc}^3$  voxels for VLC22 with at most 2.16 kpc in distance, and a heliocentric box of size  $|X|, |Y| < 370 \text{ pc}$ ,  $|Z| < 270 \text{ pc}$  and up to 590 pc in distance for LGE20. In addition, we show the *Planck* 2013 extragalactic dust map (Planck Collaboration, 2014a) and the *Gaia* total Galactic extinction (TGE) 2022 map (Delchambre et al., 2023).

All maps agree on fine structures at high Galactic latitudes but differ in the Galactic plane due to the difference in distance up to which the respective reconstruction extends. 3D dust reconstructions do not probe deep enough into the Galactic plane to fully recover the *Planck* 2013 extragalactic dust map. Bayestar19 and L+22 probe much deeper than VLC22, LGE20, and this work, yet they do not probe the full column of dust seen in *Planck* 2013 and *Gaia* TGE 2022. Both VLC22 and our map probe up to a similar depth while LGE20 only probes dust at much closer distances.

Figure 4.12 shows a zoomed-in comparison of the Perseus, Orion A, Taurus, Corona Australis (CrA), and Chameleon molecular clouds, integrated out to the maximum distance of each map (4 kpc for L+22). Among the 3D dust reconstructions, Bayestar19 and L+22 have arguably the highest angular resolution (angular discretization of  $N_{\text{side}} = 1024$  or  $3.4'$  and  $1.9'$ , respectively). They resolve the high latitude dust clouds with great detail although L+22 suffers from localized artifacts in patches of the sky. Both LGE20 ( $1 \text{ pc}^3$  boxes) and this work ( $N_{\text{side}} = 256$ ) achieve a comparable angular resolution. The VLC22 reconstruction ( $10^3 \text{ pc}^3$  voxels) is noticeably lower in resolution and does not resolve the cloud substructure on the POS. A comparison between the same molecular clouds in different distance slices in VLC22, LGE20, and our map is provided in Section 4.9.7.

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

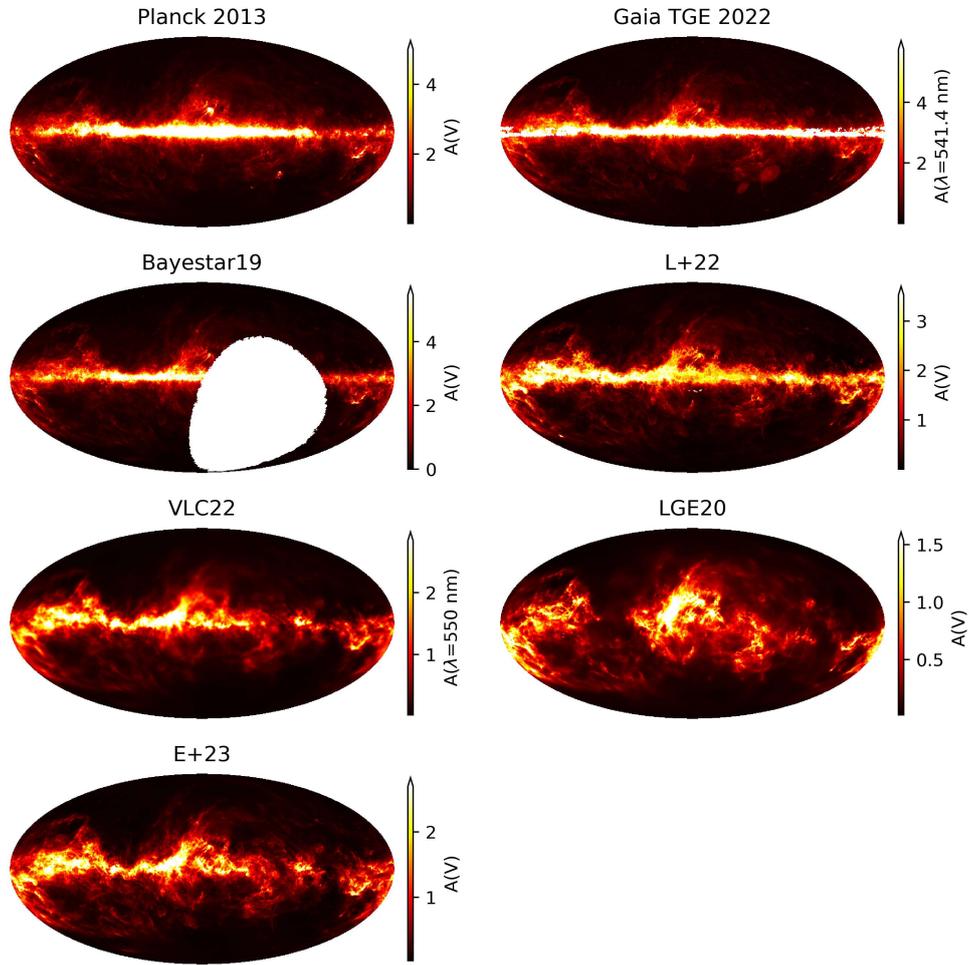


Figure 4.11: Mollweide projections of total integrated extinction and 3D extinction maps integrated out the maximum distance of the respective map. Bayestar19 reconstructs up to a maximum distance of 63 kpc (maximum reliable distance 10 kpc) and is integrated out to that volume. L+22 reconstructs up to a maximum distance of 16 kpc but the authors trust their map only out to 4 kpc and we integrate their map only to 4 kpc. VLC22 reconstructs a heliocentric box of size  $3 \text{ kpc} \times 3 \text{ kpc} \times 800 \text{ pc}$  with  $10^3 \text{ pc}^3$  voxels with at most  $< 2.16 \text{ kpc}$  in distance and is integrated out to the end of the box. Likewise, LGE20 reconstructs a heliocentric box of size  $|X| < 370 \text{ pc}$ ,  $|Y| < 370 \text{ pc}$ ,  $|Z| < 270 \text{ pc}$  covering at most  $< 590 \text{ pc}$  in distance and is integrated out to the end of the box. The colorbars saturate at the respective 99% quantile of the map except for the colorbar of *Planck* 2013, which saturates at 5 mag for better comparability.

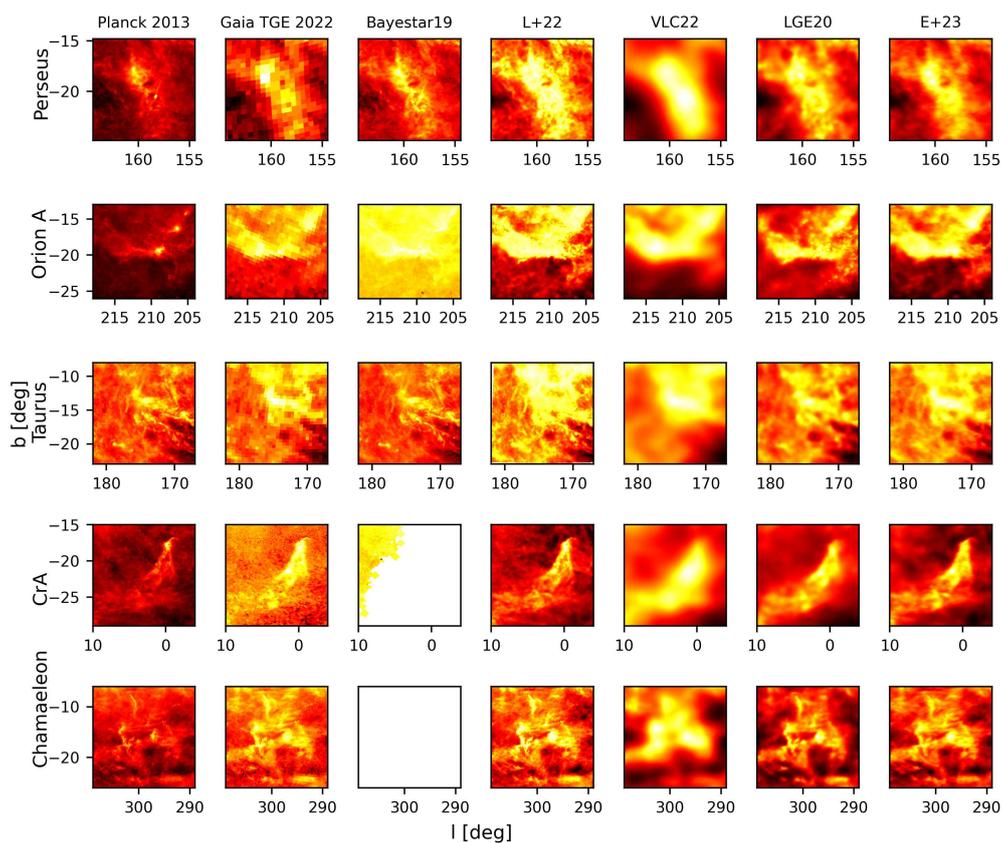


Figure 4.12: Zoomed-in views toward the individual molecular clouds (Perseus, Orion, Taurus, Corona Australis, and Chamaeleon) seen in Figure 4.11. The colorbars are logarithmic and span the full dynamic range of the selected POS slice in every image. Each row is a separate region and each column a separate reconstruction.

## 4.8 Conclusions

We present a 3D dust map with a POS and LOS resolution comparable to Leike et al. (2020) that extends out to 1.25 kpc. We used the distance and extinction estimates of Zhang et al. (2023), which have much lower extinction uncertainties than competing catalogs while probing a similar number of stars. Our reconstruction has a resolution comparable to the 2 pc resolution of Leike et al. (2020). Specifically, it has an angular resolution of up to 14' and parsec-scale distance resolution. Our map is in good agreement with existing 3D dust maps and improves upon them in terms of volume covered at a high spatial resolution. The map is publicly available online<sup>5</sup> and can be queried via the `dustmaps` Python package. We anticipate that the map will be useful for a wide range of applications in studying the distribution of dust and the ISM more broadly.

## Acknowledgements

We thank Joao Alves for many fruitful discussions at the “Self-Organization Across Scales: From nm to parsec (SOcraSCALES)” workshop at the Munich Institute for Astro-, Particle and BioPhysics, an institute of the Excellence Cluster ORIGINS in 2022 and afterwards. Furthermore, we thank Jakob Roth for many invaluable discussions about the model and for providing feedback on the early versions of the reconstruction. We thank Alyssa Goodman for providing invaluable feedback on the late versions of the reconstructions. We also thank Michael A. Kuhn for providing us with a unified catalog of YSOs. Gordian Edenhofer acknowledges that support for this work was provided by the German Academic Scholarship Foundation in the form of a PhD scholarship (“Promotionsstipendium der Studienstiftung des Deutschen Volkes”). Catherine Zucker acknowledges that support for this work was provided by NASA through the NASA Hubble Fellowship grant #HST-HF2-51498.001 awarded by the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. Philipp Frank acknowledges funding through the German Federal Ministry of Education and Research for the project ErUM-IFT: Informationsfeldtheorie für Experimente an Großforschungsanlagen (Förderkennzeichen: 05D23EO1). Andrew K. Saydjari acknowledges support by a National Science Foundation Graduate Research Fellowship (DGE-1745303). Andrew K. Saydjari and Douglas Finkbeiner acknowledge support by NASA ADAP grant 80NSSC21K0634 “Knitting Together the Milky Way: An Integrated Model of the Galaxy’s Stars, Gas, and Dust”. This work was supported by the National Science Foundation under Cooperative Agreement PHY-2019786 (The NSF AI Institute for Artificial Intelligence and Fundamental Interactions). A portion of this work was enabled by the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Har-

---

<sup>5</sup><https://doi.org/10.5281/zenodo.8187942>

vard University. We acknowledge support by the Max-Planck Computing and Data Facility (MPCDF). This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

## 4.9 Supplementary material

### 4.9.1 ZGR23 in dust-free regions

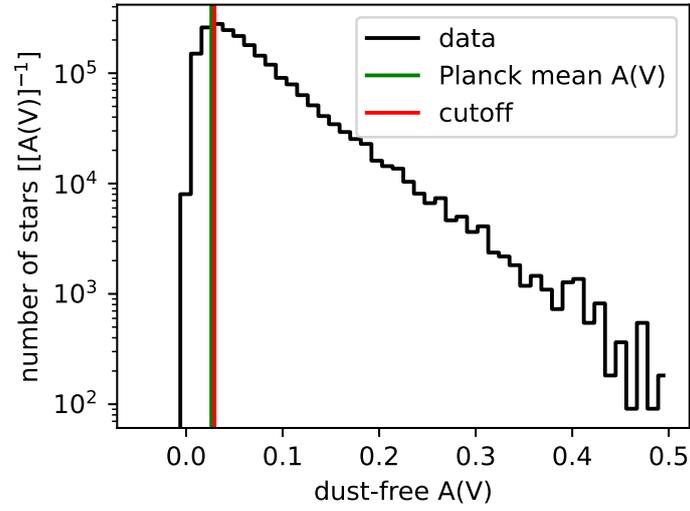
To gauge the reliability of the ZGR23 catalog, we analyzed the extinction to stars in dust-free regions (cf. Leike and Enßlin 2019; Leike et al. 2020). In dust-free regions, we would expect the extinction to be zero within the uncertainties of the catalog. To classify a region as dust-free, we use the *Planck* dust emission map (Planck Collaboration, 2014a). A region is said to be dust-free if the *Planck*  $E(B-V)$  map is below or equal to 9.5 mmag or approximately 29 mmag in terms of  $A_V$ .

The first panel of Figure 4.13 shows the histogram of ZGR23 extinction to stars with `quality_flags` < 8 in dust-free regions translated to  $A_V$ . We see that the histogram of the extinction peaks at the cutoff value and coincides with the mean total extinction as measured by *Planck* in those regions translated to  $A_V$ . The density of extinction values falls off exponentially after the cutoff value. Overall, the ZGR23 extinction seems to be in good agreement with Planck Collaboration (2014a) for dust-free regions.

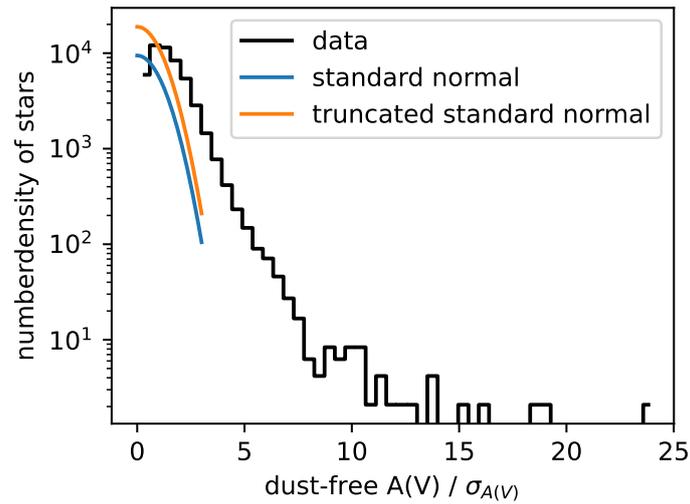
The second panel of Figure 4.13 shows the extinction divided by their uncertainties for the stars from Figure 4.13. The standardized extinctions are centered around unity, indicating that the ZGR23 extinction indeed are offset from zero by about one standard deviation in dust-free regions. This is in agreement with the previous finding that the extinctions are centered around the cutoff value instead of clustering around zero. The width around the center is comparable to a truncated standard normal distribution or a normal distribution. In total, about 1% of the probability mass lies outside the possible range of all ZGR23 extinction with `quality_flags` < 8 if we assume a normal distribution for the extinctions.

Except for outliers far from the center, which can be captured by an outlier model, the ZGR23 catalog seems to be in agreement with POS measurements in dust-free regions and the spread around the cutoff value approximately follows a (truncated) normal distribution. We deem the ZGR23 catalog to be reliable for our purposes and approximate the uncertainties using a normal distribution. We accepted the mismodeling of a small fraction of probability mass for a simpler model (see Sections 4.5 and 4.6).

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun



(a)



(b)

Figure 4.13: Absolute and relative ZGR23 extinction in dust-free regions. Panel (a): Histogram of the ZGR23 extinctions in dust-free regions translated to  $A_V$ . The mean extinction in dust-free regions based on *Planck* is shown as a vertical green line, and the cutoff value translated to  $A_V$  for our definition of dust-free is shown in red. Panel (b): Same as Panel (a) but the extinctions are scaled by their accompanying uncertainties. A truncated standard normal distribution and a standard normal distribution are plotted on top. Both ordinates are logarithmic.

### 4.9.2 Metric Gaussian variational inference

The VI method MGVI approximates the true posterior  $P(\xi | d)$  with a standard normal distribution in a linearly transformed space in which the posterior more closely resembles a standard normal. Let  $Q_{\bar{\xi}}(y | d) = \mathcal{G}(y(\xi) | y(\bar{\xi}), \mathbb{1}) \left| \frac{dy}{d\xi} \right|$  be the approximate posterior and  $y(\xi) : \xi \mapsto y(\xi)$  the coordinate transformation. In this space the transformed posterior reads  $P(\xi(y) | d) \left| \frac{d\xi}{dy} \right|$ . We denote the metric of the space in which  $P(\xi(y) | d) \left| \frac{d\xi}{dy} \right|$  is “more” standard normal by  $M := \frac{dy}{d\xi} \left( \frac{dy}{d\xi} \right)^\dagger$ . Assuming  $y(\xi) : \xi \mapsto y(\xi)$  is known, the difficulty lies solely in finding the optimal  $\bar{\xi}$  for  $Q_{\bar{\xi}}$ .

Based on the Fisher information metric and Frequentist statistics, Knollmüller and Enßlin derive a coordinate transformation  $y_{\bar{\xi}}(\xi)$  centered on  $\bar{\xi}$  that is linear in  $\xi$ . In Frank et al. (2021), the authors find that a set of Riemannian normal coordinates  $y_{\bar{\xi}}(\xi)$  centered on  $\bar{\xi}$  are an improved, nonlinear estimate of the coordinate transform  $y(\xi) \approx y_{\bar{\xi}}(\xi)$ . The improvements, though, come at slightly higher computational costs. We refer the reader to Frank et al. (2021) and Frank (2022) for further details on geoVI and its relation to MGVI, the choice of metric, and an analysis of its failure modes. For computational reasons, we used MGVI for our inference.

MGVI and geoVI start at a random initial position for  $\bar{\xi}$  and draw  $n_{\text{samples}}$  standard normal samples in the space of  $y$ . Next, they transform the samples to the space of  $\xi$  via  $y_{\bar{\xi}}$ , the local, linear (respectively, nonlinear for geoVI) approximation to  $y(\xi)$  at  $\bar{\xi}$ . We denote the samples in parameter space by  $\{\xi_1, \dots, \xi_{n_{\text{samples}}}\}$ . Relative to the expansion point  $\bar{\xi}$  the samples read  $\{\Delta\xi_1 := \xi_1 - \bar{\xi}, \dots, \Delta\xi_{n_{\text{samples}}} := \xi_{n_{\text{samples}}} - \bar{\xi}\}$ . The samples  $\{\Delta\xi_1, \dots, \Delta\xi_{n_{\text{samples}}}\}$  around  $\bar{\xi}$  provide an empirical, sampled approximation to  $Q_{\bar{\xi}}$ , which we denote by  $\tilde{Q}_{\bar{\xi}}$ . MGVI and geoVI then optimize  $\bar{\xi}$  of the sampled distribution,  $\tilde{Q}_{\bar{\xi}}$ , by minimizing the variational Kullback-Leibler (KL) divergence between  $\tilde{Q}_{\bar{\xi}}$  and the true distribution  $P$

$$\bar{\xi}' = \arg \min_{\bar{\xi}} \text{KL} \left( \tilde{Q}_{\bar{\xi}}, P(\xi | d) \right) \quad (4.20)$$

$$= \arg \min_{\bar{\xi}} \left\langle \ln \frac{\tilde{Q}_{\bar{\xi}}}{P(\xi | d)} \right\rangle_{\tilde{Q}_{\bar{\xi}}} \quad (4.21)$$

$$= \arg \min_{\bar{\xi}} \langle -\ln P(\xi | d) \rangle_{\tilde{Q}_{\bar{\xi}}} \quad (4.22)$$

$$= \arg \min_{\bar{\xi}} \frac{-1}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}} \ln P(\Delta\xi_i - \bar{\xi} | d) . \quad (4.23)$$

They keep the relative samples  $\{\Delta\xi_1 := \xi_1 - \bar{\xi}, \dots, \Delta\xi_{n_{\text{samples}}} := \xi_{n_{\text{samples}}} - \bar{\xi}\}$  fixed during the optimization and only vary  $\bar{\xi}$ . Finally, they update the expansion point  $\bar{\xi}$  to the new found optimum  $\bar{\xi}'$ .

#### 4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

After the minimization, MGVI and geoVI draw a new set of samples, transform them via a local (linear) expansion of  $y$ , and then minimize again. The drawing of samples and minimization is repeated until a fixed point for  $\bar{\xi}$  is reached. Algorithm 3 summarizes the algorithmic steps of the variational approximation to the true posterior.

```

def sample( $\xi$ ,  $n_{\text{samples}}$ ):
    ...
    return  $\{\xi_1, \dots, \xi_{n_{\text{samples}}}\}$ 
 $\bar{\xi} \leftarrow \dots$  // random init

while  $\bar{\xi}$  not converged do
     $\{\xi_1, \dots, \xi_{n_{\text{samples}}}\} \leftarrow \text{sample}(\bar{\xi}, n_{\text{samples}})$ 
     $\{\Delta\xi_1, \dots, \Delta\xi_{n_{\text{samples}}}\} \leftarrow \{\xi_1 - \bar{\xi}, \dots, \xi_{n_{\text{samples}}} - \bar{\xi}\}$ 
     $\bar{\xi} \leftarrow \arg \min_{\bar{\xi}} \frac{-1}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}} \ln P(\Delta\xi_i + \bar{\xi} | d)$ 
end

```

**Algorithm 3:** Pseudocode for the MGVI and geoVI expansion point VI scheme.

#### 4.9.3 Extinction within the innermost 69 pc

By construction of our likelihood, we know that our model is biased high for low extinction values because we neglected the positivity prior of the ZGR23 catalog. This imprints on to the map in the form of a thin layer of dust extinction right at the beginning of the modeled volume. As voxels in our reconstruction are correlated, the extinguished first layer of voxels pulls the next layer of voxels to slightly higher extinctions too. At 69 pc the differential extinction as a function of distance reaches a local minimum, and we expect little to now influence of the innermost layers of voxels. Thus, we determined 69 pc to be our cutoff.

Figure 4.14 shows the integrated extinction that is cut out from the final reconstruction. Most of the extinction is likely spurious. Overall, no structure contributes a significant amount of extinction. However, to be consistent with ZGR23, we suggest adding the removed extinction back to the map when comparing integrated extinction.

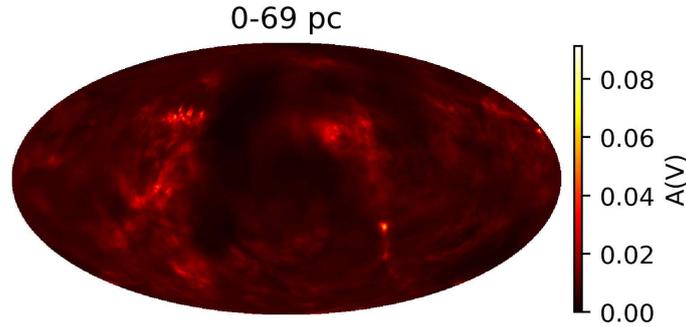


Figure 4.14: Mollweide projection of the integrated  $A_V$  extinction in the innermost 69 pc, which is likely dominated by spurious effects and is therefore excluded from the reconstruction. The colorbar is linear and covers the full range of the extinction that is cut out of the reconstruction.

#### 4.9.4 Extinction catalog

We release a catalog of expected extinction for all stars within the subset of the ZGR23 catalog that we used for our reconstruction (see Section 4.2). We predict the expected extinction conditional on the known parallax including parallax uncertainties. Our prediction is the best guess of our model for the extinction toward a star but is not necessarily the best guess for the extinction at the mean parallax of the star.

Our extinction predictions (see Sections 4.3 and 4.4.1) differ from the extinctions in the ZGR23 catalog by coupling the individual stars via the 3D dust extinction density. By virtue of every star depending on all nearby stars via the prior, our extinction predictions come in the form of joint predictions for all stars. In regions where the 3D dust extinction density is well constrained, the joint predictions to first order factorize into predictions for individual stars, and we can compute expected extinctions for individual stars and their uncertainties.

Our catalog of extinction includes the innermost 69 pc from the beginning of our grid and the outer 550 pc beyond 1.25 kpc that we cut away in the 3D map. We advise caution when analyzing the stars of our catalog within those regions as they might carry additional biases. Details on why these regions were removed from the final map are given in Sections 4.5 and 4.6.

The top panel of Figure 4.15 compares the ZGR23 extinctions to our mean extinction predictions for Gaia BP/RP stars. Overall, our mean extinctions are in very good agreement with the extinctions in the ZGR23 catalog for the vast majority of stars. However, below 50 mmag and above 4 mag, our extinction predictions deviate from the predictions in ZGR23. In any given extinction bin from ZGR23 respectively our work, we would expect half of the respective other extinctions to be below the bisector and the other half to be above. At 50 mmag, 34% more stars than expected have higher extinctions than the corresponding

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

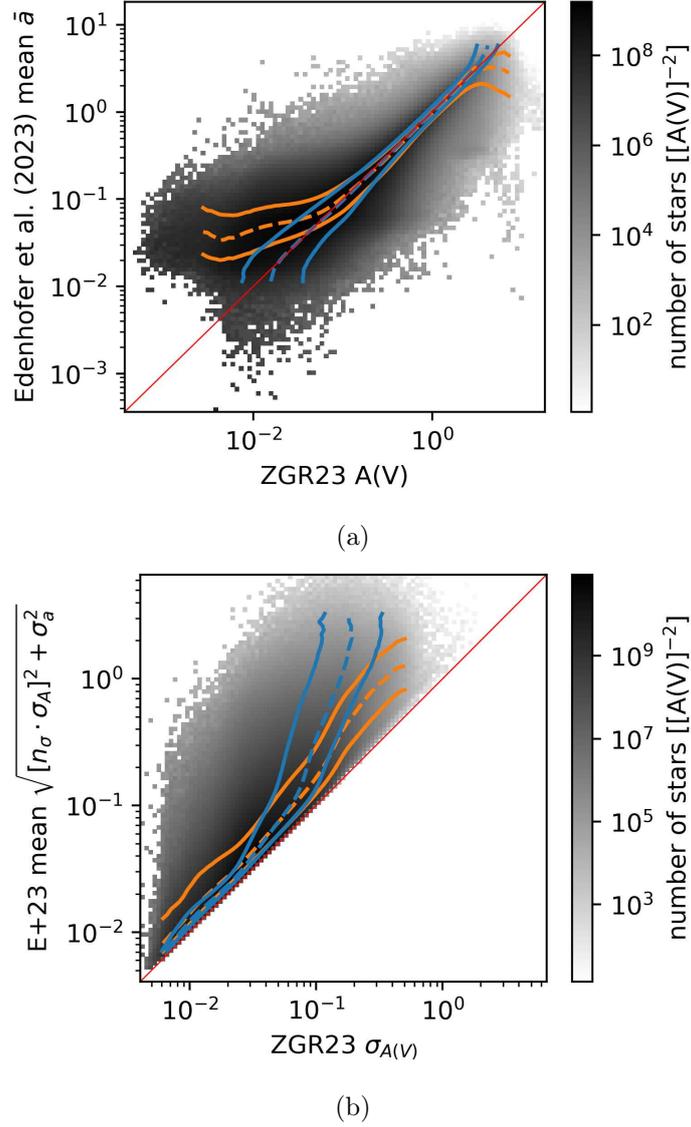


Figure 4.15: ZGR23’s extinction versus our predicted extinction for Gaia BP/RP stars. Panel (a): Our mean posterior extinctions versus the ZGR23 extinctions to stars as 2D histogram. The 16<sup>th</sup>, 50<sup>th</sup>, and 84<sup>th</sup> quantiles of the ZGR23 extinctions for each bin of our mean extinction are shown as blue lines. The respective quantiles of our predictions in bins of the ZGR23 extinctions are shown as orange lines. Panel (b): Same comparison but for our posterior mean predictions for the ZGR23 measurement uncertainties,  $\sqrt{[n_\sigma(\xi) \cdot \sigma_A]^2 + \sigma_a^2}$ , versus the ZGR23 uncertainties. Note that the predictions for the ZGR23 measurement uncertainties are not the uncertainties of our extinction predictions. See Section 4.4 and in particular Equation (4.18) for further details on the quantities shown here. The bisectors are shown in red. The colorbars are logarithmic.

extinctions in ZGR23. The difference further widens for lower ZGR23 extinctions. At 4 mag, 34% more stars than expected have lower extinctions than the corresponding extinctions in ZGR23.

The bottom panel of Figure 4.15 shows our and the ZGR23 extinction uncertainties. We note that our extinction uncertainties are predictions for the measured uncertainties of the ZGR23 catalog  $[n_\sigma(\xi) \cdot \sigma_A]^2 + \sigma_a^2(\rho(\xi))$  and not the uncertainties of our extinction predictions  $\text{std}(\bar{a})$  (see Section 4.4). Overall, both uncertainties agree well for the vast majority of stars. At low extinctions our uncertainties only marginally inflate the ZGR23 uncertainties. However, at high extinction uncertainties, our predictions cover a larger range, and we find that the ZGR23 significantly underpredicts our extinction uncertainties of stars.

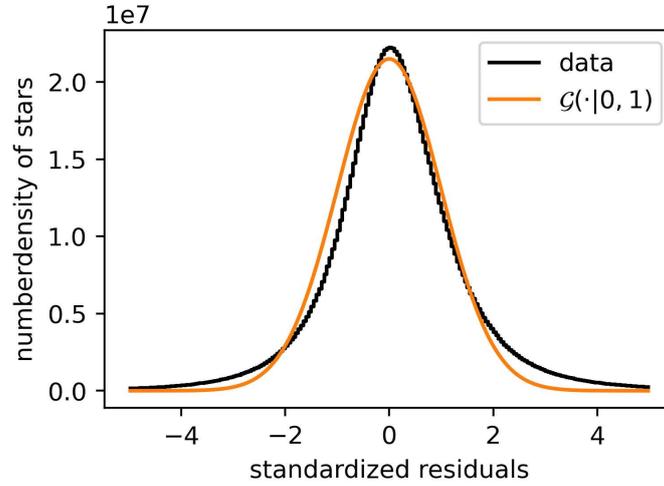


Figure 4.16: Mean standardized extinctions:  $(A - \bar{a}) / \sqrt{(n_\sigma \cdot \sigma_A)^2 + \sigma_a^2}$  (see Section 4.4 and in particular Equation (4.18)) within the range  $-5$  to  $5$ .

Figure 4.16 summarizes the extinctions and the extinction uncertainties of both ZGR23 and our predictions into a single histogram of the mean standardized extinction. A standard Gaussian is shown as reference. The mean standardized residuals have two slight overdensities at each tail compared to the Gaussian.

Figure 4.17 shows the posterior standard deviation of our extinction predictions versus the ZGR23 uncertainties. Our model yields approximately one order in magnitude lower extinction uncertainties than the ZGR23 uncertainties for the vast majority of stars. The effect is less pronounced for low ZGR23 extinction uncertainties.

Our predictions for the extinction to stars theoretically contain more information since we allow for the cross-talk of nearby stars via the 3D distribution of dust and thus might be more accurate. However, the ZGR23 catalog might yield better results in practice because it does not discretize the 3D volume within

#### 4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

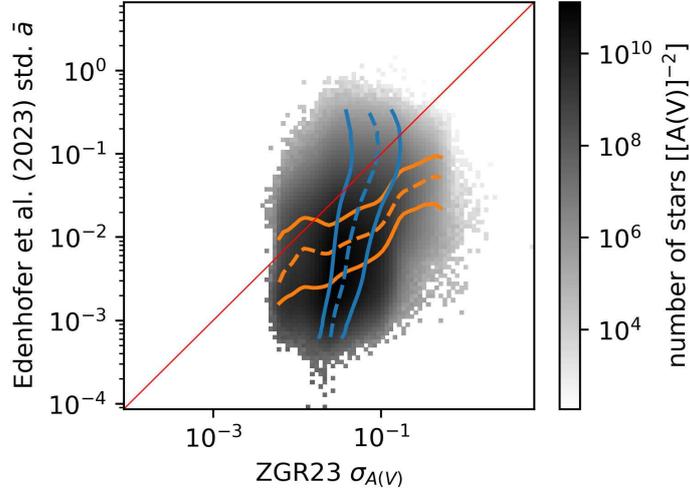


Figure 4.17: Similar to Figure 4.15 but for the posterior standard deviation of our extinctions versus the ZGR23 uncertainties. The 16<sup>th</sup>, 50<sup>th</sup>, and 84<sup>th</sup> quantiles of the ZGR23 uncertainties for each bin of our standard deviation are shown as blue lines. The respective quantiles of our standard deviation in bins of the ZGR23 uncertainties are shown as orange lines. The bisectors are shown in red. The colorbars is logarithmic.

which the stars reside. By discretizing the modeled volume we can produce contradicting data that in a continuous space is non-contradicting, for example by putting highly extinguished stars that lie in a dust cloud into the same voxel as less extinguished stars that are adjacent to the dust cloud. Overall, both predictions agree very well for stars below between 50 mmag and 4 mag. More work is needed to validate the discrepant predictions at very low and very high extinctions.

#### 4.9.5 2 kpc reconstruction

In Section 4.5 we describe how we iteratively increase the distance out to our maximum reconstructed distance. We did so to improve the convergence of the reconstruction. We also tried naively reconstructing the full volume at once. Using all the available data is computationally prohibitive, so we limited the reconstruction to high quality data using `quality_flags == 0`,  $\sigma_A \leq 0.04$ , and  $\sigma_\omega/\omega < 0.33$ .

We used  $1/(\omega - \sigma_\omega) < 3 \text{ kpc}$  and  $1/(\omega + \sigma_\omega) > 40 \text{ pc}$  to select the stars within a 3 kpc sphere. To further speed up the inference, we started the inference using at first only a sample of 10%, then 20%, 45%, 67%, and finally 100% of the stars. In total, we selected 59,334,214 stars. After the inference, we cut away the outermost 1 kpc of the sphere of the data-constrained region to avoid degradation effects due to the thinning out of stars at the edge. The overall reconstructed

volume after removing the outermost HEALPix spheres extends out to 2 kpc in distance.

The reconstruction is shown in Figure 4.18 and again in Figure 4.19 with a catalog of YSO clusters (Kuhn, 2023) overlaid on top. It shows the same large-scale features as the smaller reconstruction discussed in the main text. The distribution of dense dust clouds is in agreement with the positions of YSO clusters within the distance uncertainties of the YSO clusters. Compared to Figure 4.5 the reconstruction is less detailed and features more pronounced artifacts.

We used the larger reconstruction to validate the inference of the smaller one. Specifically, we used the larger reconstruction to ensure that structures aligned with or close to the radial boundaries at which we increased the distance of the main reconstruction are independent of the locations at which we increased the distance covered.

We release the larger reconstruction as an additional data product together with the main reconstruction. We advise using the main reconstruction for all regions that fall within its volume. Care should be taken when interpreting small-scale features or structures at high distances in the larger reconstruction.

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

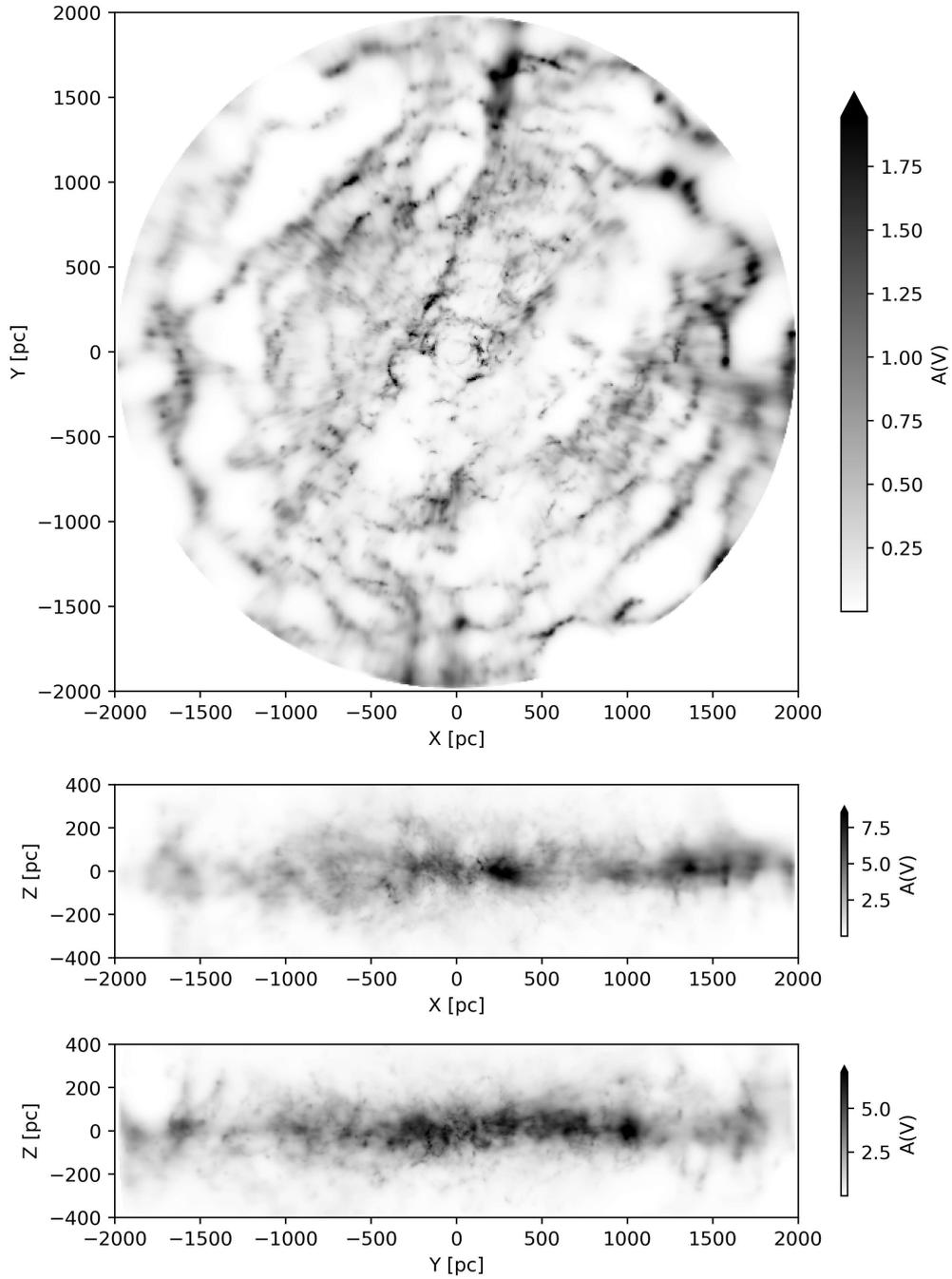


Figure 4.18: Axis parallel projections of the reconstructed dust extinction in a box of dimensions  $4 \text{ kpc} \times 4 \text{ kpc} \times 0.8 \text{ kpc}$  centered on the Sun. The colorbar is linear and saturates at the 99.9% quantile.

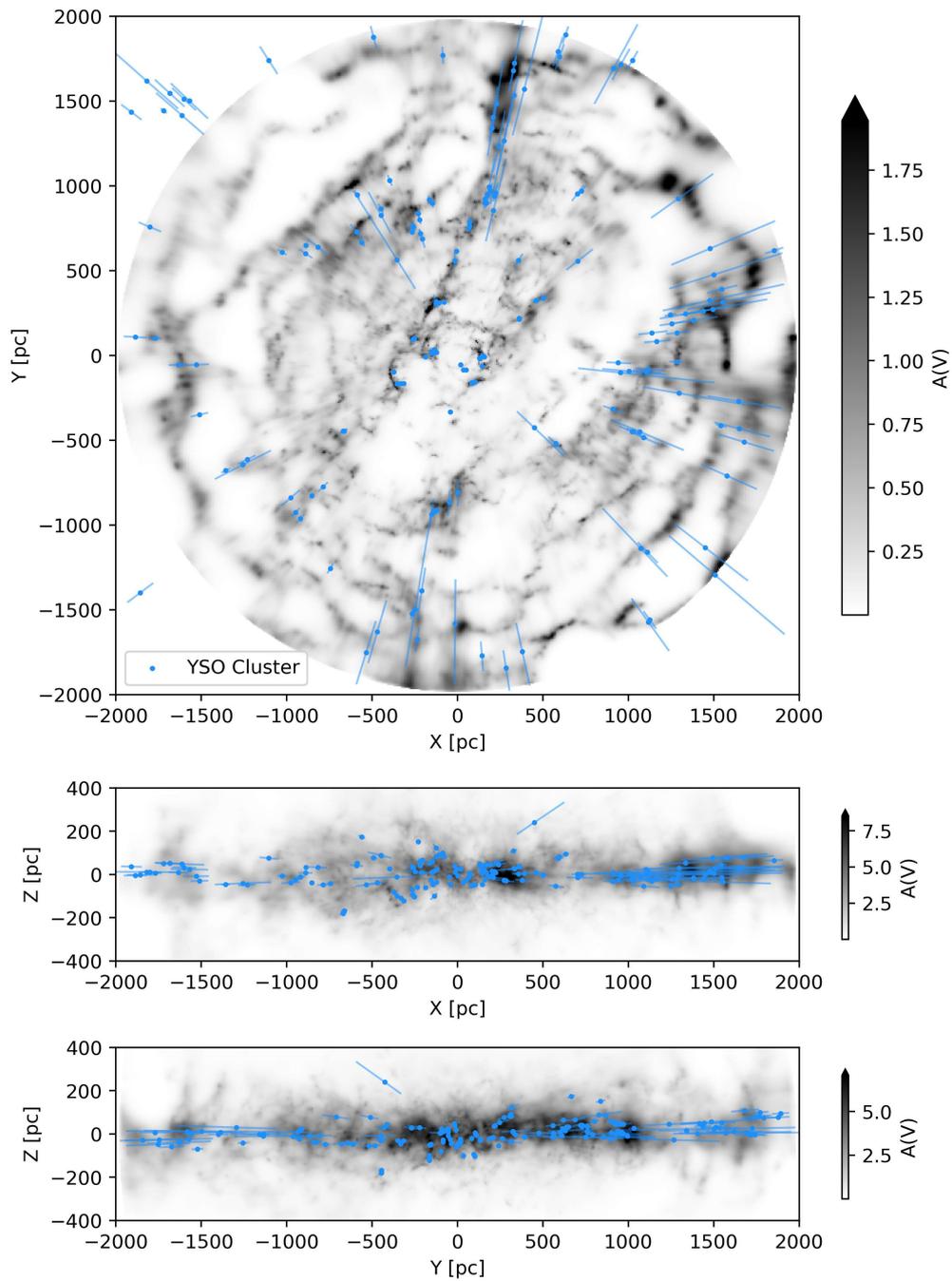


Figure 4.19: Same as Figure 4.18 but with a catalog of clusters of YSOs (Kuhn, 2023) based on Kuhn et al. (2021), Winston et al. (2020), and Marton et al. (2023) shown as blue dots on top of the reconstruction; their distance uncertainties are shown as extended lines.

### 4.9.6 Using the reconstruction

All data products are made publicly available online<sup>6</sup>. The data products are stored in the FITS file format. The main data products are the posterior samples of the spatial 3D distribution of dust extinction discretized to HEALPix spheres at logarithmically spaced distances. For convenience, we also provide the posterior mean and standard deviation of the samples of the HEALPix spheres at logarithmically spaced distances.

We additionally interpolated the posterior mean and standard deviation to a Cartesian grid. The interpolation was carried out at a lower discretization using  $2^3 \text{ pc}^3$  voxels to keep the file size reasonably small. We recommend re-interpolating the map at a higher discretization for the study of individual regions within the map.

We release the interpolation script as part of the data release. Its signature reads `interp2box.py [-h] [-o OUTPUT_DIRECTORY] [-b BOX] healpix_path .` A box is a string of two tuples separated by two colons. The first tuple specifies the number of voxels along each axis of the box and the second tuple specifies the corners of the box in parsecs in heliocentric coordinates. To interpolate the map to a box with  $1051 \times 1051 \times 351$  voxels of size  $|X|, |Y| \leq 2100 \text{ pc}$  and  $|Z| \leq 700 \text{ pc}$ , use `interp2box.py -b '(1051,1051,351)::((-2100,2100),(-2100,2100),(-700,700))' -- mean_and_std_healpix.fits.`

In addition, we interpolated the posterior mean and standard deviation to Galactic longitude, latitude, and distance. The signature of the interpolation script reads `interp21bd.py [-h] [-o OUTPUT_DIRECTORY] [-b BOX] healpix_path.` Its behavior is similar to `interp2box.py` but the box is specified in terms of Galactic longitude, latitude, and distance in units of degrees, degrees, and parsecs, respectively.

Both scripts require the Python packages `numpy` (Harris et al., 2020b), `astropy` (Astropy Collaboration et al., 2013; Astropy Collaboration and Astropy Contributors, 2018; Astropy Collaboration and Astropy Project Contributors, 2022), and `healpy` (Gorski et al., 2005; Zonca et al., 2019). Depending on the number of output voxels, the interpolation can be very memory intensive and computationally expensive.

### 4.9.7 Molecular clouds by distance

Figures 4.20 to 4.24 depict zoomed-in views of Figure 4.11 at different distant slices, alongside histograms comparing the extinction of VLC22, LGE20, and our mean map toward Perseus, Orion A, Taurus, Corona Australis, and Chameleon. The top panels depict the same POS views for the clouds as in Figure 4.12, but now showing the dust extinction in finite distance bins, rather than integrated over the full distance range. A low-resolution 3D interactive figure of the

---

<sup>6</sup><https://doi.org/10.5281/zenodo.8187942>

reconstruction including all the above molecular clouds is available online<sup>7</sup>.

The bottom panels of the figures are akin to Figure 4.10 but compare the extinction within the selected distance and POS area of the respective molecular cloud only. The probing points to which we integrated are spaced in the distance range of the top panel. We started the integration at the lowermost distance of the top panel and successively integrated out in steps of 0.5 pc. The probing points are spaced 15' apart along the POS.

LGE20 and our map agree well for Perseus, Orion A, Taurus, and Chameleon. Some structures appear at slightly larger distances in our map as indicated by the arches above the diagonal in the second panels. LGE20 saturates sooner than our map as it flattens off at high extinctions below the diagonal. The flattening off is less pronounced for Chameleon.

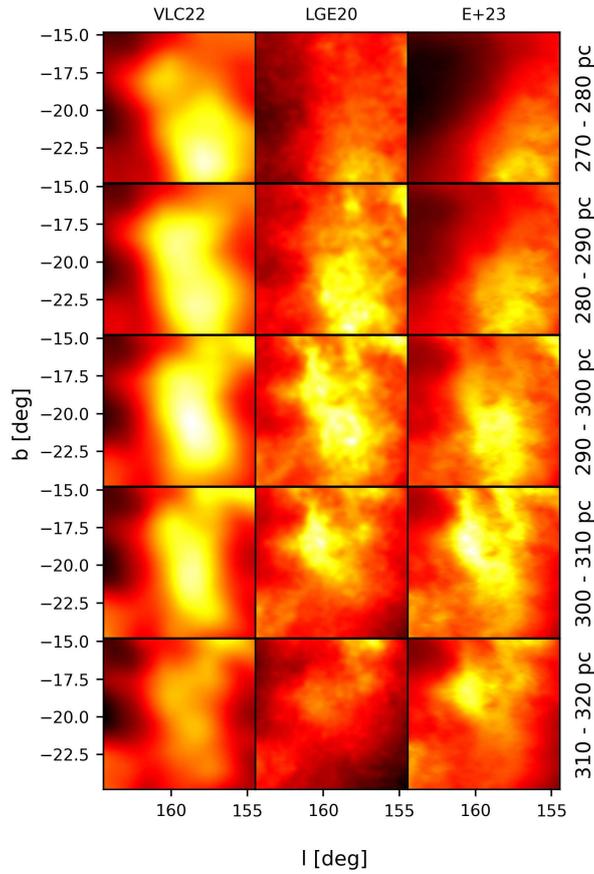
Corona Australis is an outlier among the generally good level of agreement. In the lower panel, there is noticeably more density off the diagonal. We see arches both above and below the diagonal indicating that some structures are closer while others are farther in our map compared with LGE20. While LGE20 puts the head of Corona Australis in the distance slice ranging from 140 pc to 160 pc, our map puts the head of Corona Australis dozens of parsecs farther in distance. Possible reasons for this shift include an insufficient number of stars to constrain the distance to the head of Corona Australis or a failure mode in our posterior approximation.

VLC22 is much lower in resolution and does not resolve degree-scale high extinction regions. The extinction in VLC22 within the selected distance slice is much lower than in our map and by extension in LGE20. The missing extinction partially lies outside the selected box as it is strongly smeared out radially.

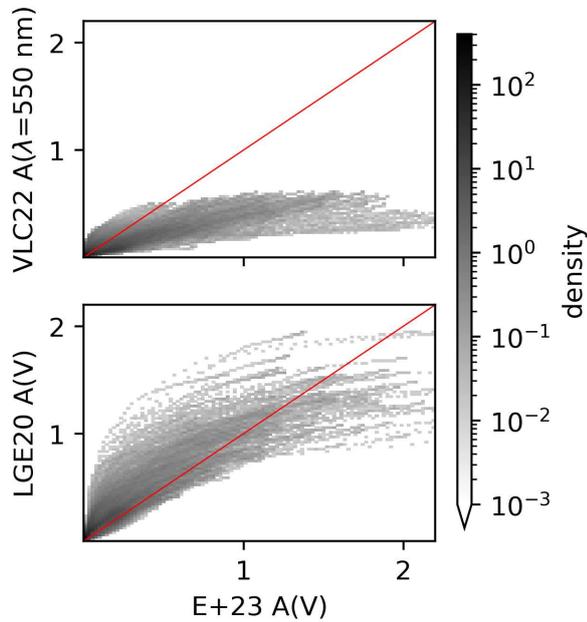
---

<sup>7</sup>[https://faun.rc.fas.harvard.edu/czucker/Paper\\_Figures/3D\\_Dust\\_Edenhofer2023.html](https://faun.rc.fas.harvard.edu/czucker/Paper_Figures/3D_Dust_Edenhofer2023.html)

4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun



(a)



(b)

Figure 4.20: Comparison of different dust maps for Perseus. Panel (a): Zoomed-in view of Figure 4.11 toward Perseus akin to Figure 4.12, in different distant slices. Columns depict dust reconstructions, while rows depict distance slices. The logarithmic colorbars are separate for the reconstructions but shared for different distance slices. Panel (b): Comparison of the mean posterior extinction integrated from the lowest distance of Panel (a) to regularly spaced points in the distance range of Panel (a). The extinction predictions for our map versus LGE20 and VLC22 are shown as histograms. The binning is linear and the colorbar logarithmic. The bisectors are shown in red.

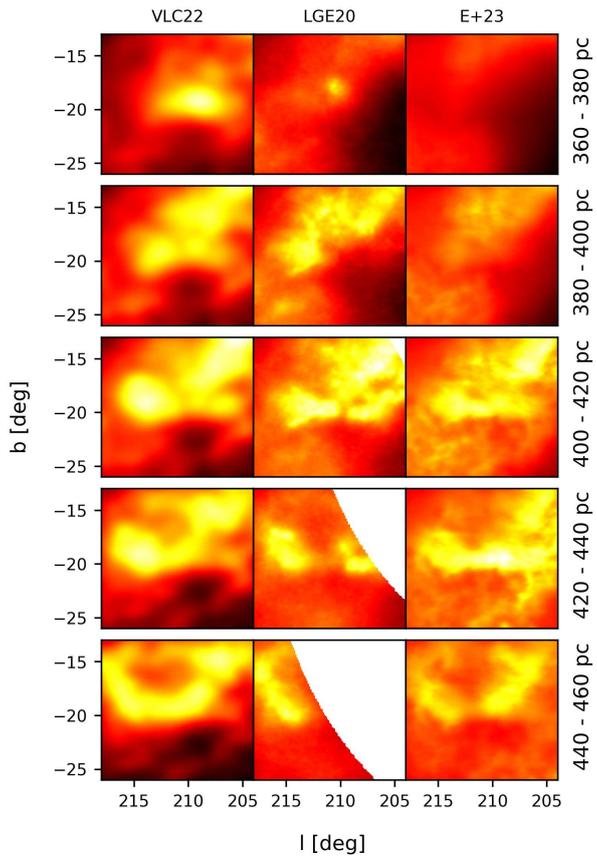
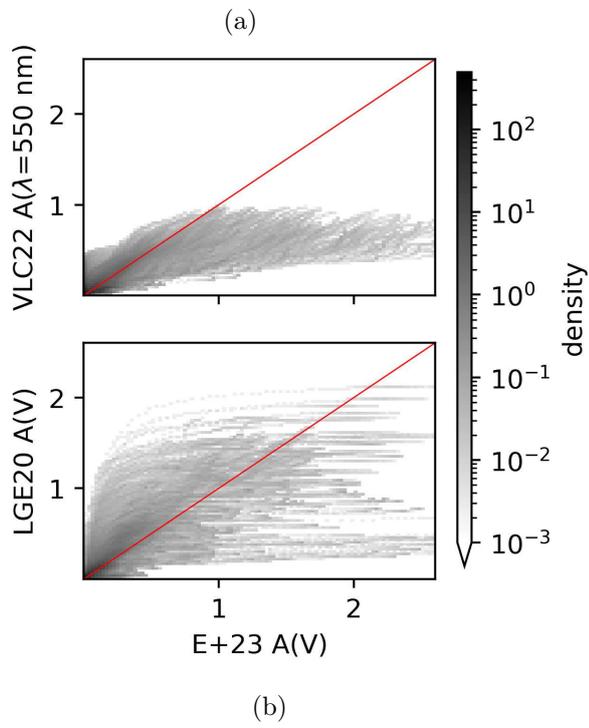


Figure 4.21: Same as Figure 4.20 but for Orion A.



4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

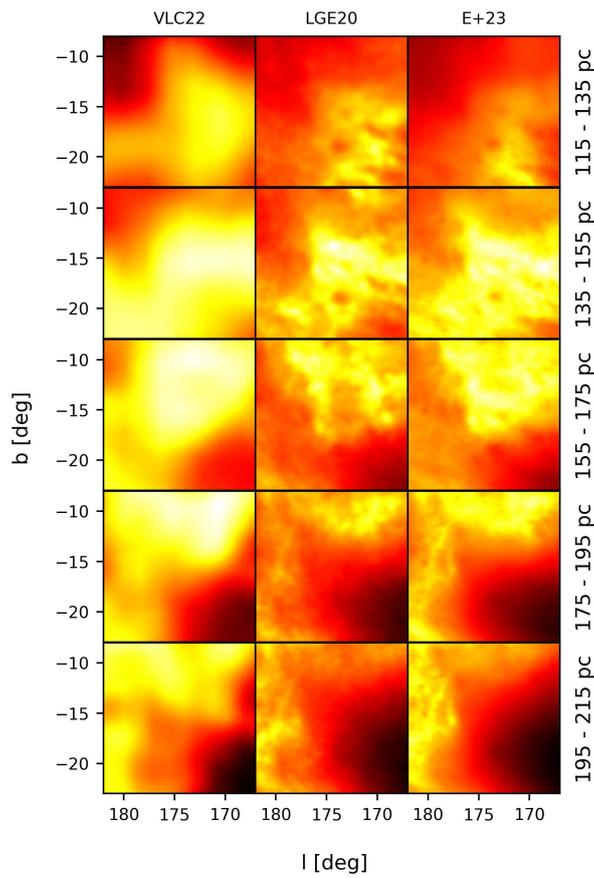
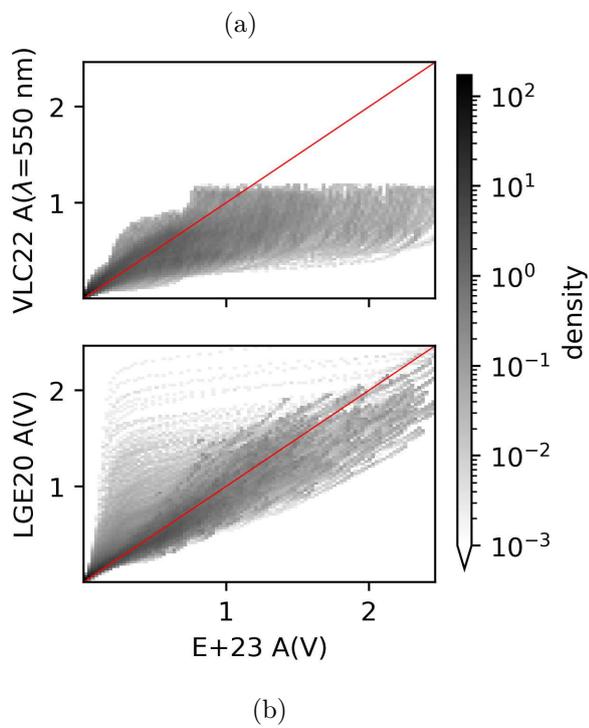


Figure 4.22: Same as Figure 4.20 but for Taurus.



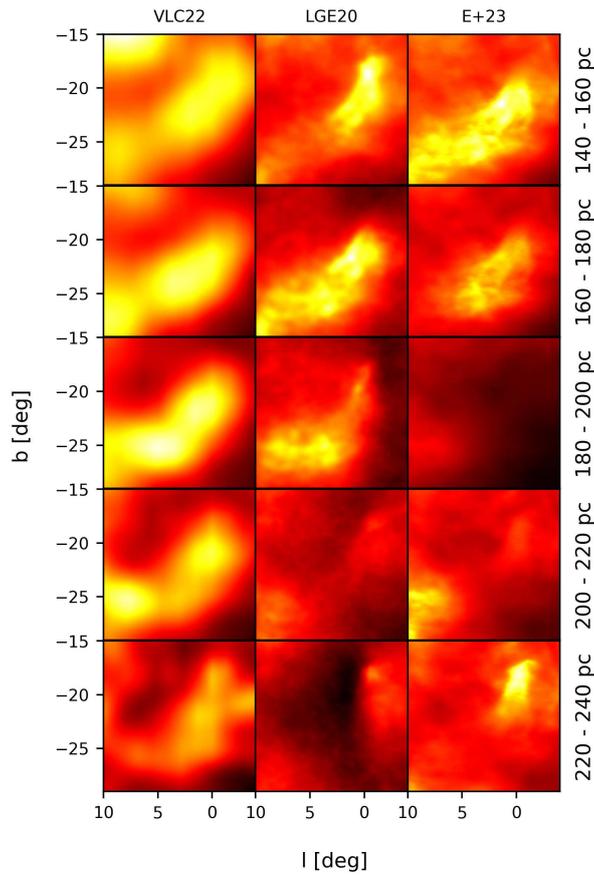
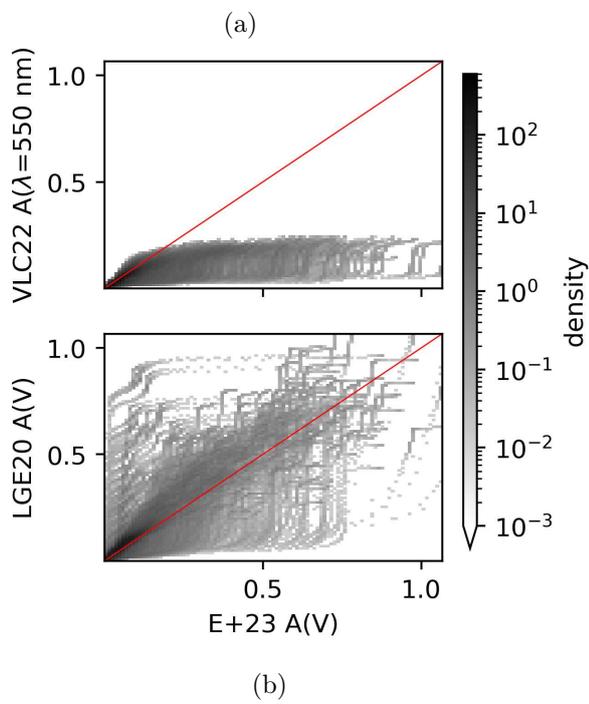


Figure 4.23: Same as Figure 4.20 but for Corona Australis (CrA).



4. A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun

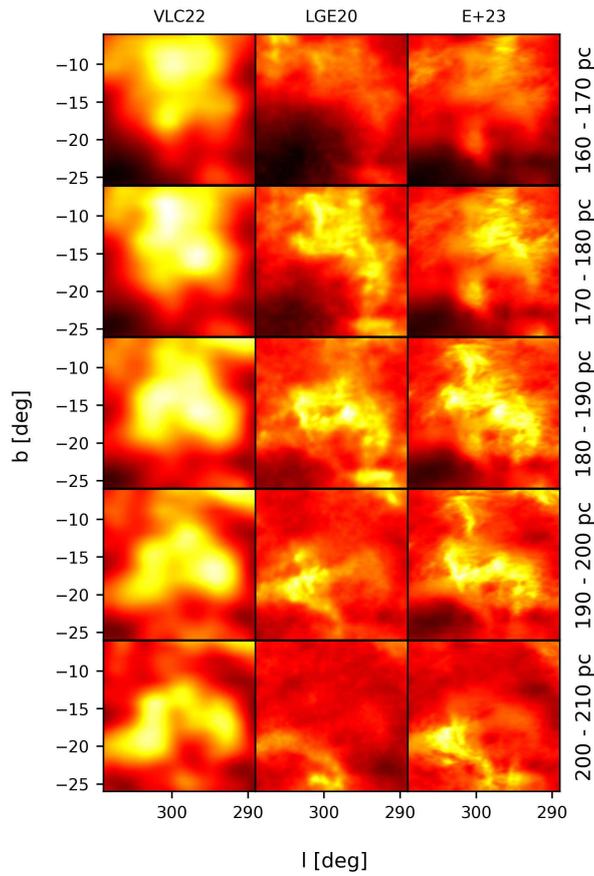
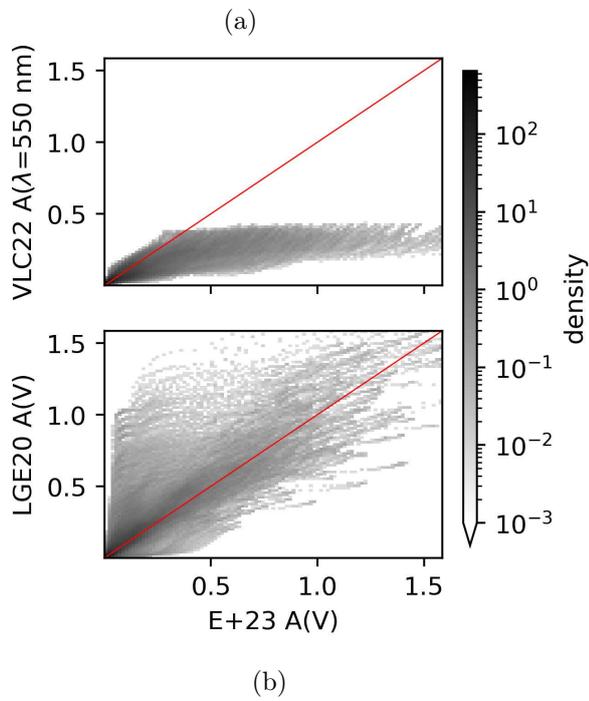


Figure 4.24: Same as Figure 4.20 but for Chameleon.



## 5 The “C”: The large Chameleon-Musca-Coalsack cloud

*The following chapter is an article currently under review at Astronomy & Astrophysics (Edenhofer et al., 2024a). This article emerged from a close collaboration with João Alves who contributed to this work via countless discussions. Furthermore, João Alves contributed significant portions of the text for the abstract and introduction. Catherine Zucker contributed text and very valuable feedback throughout the article. Torsten Enßlin contributed comments on later versions of the manuscript. All authors read, commented, and approved the manuscript. The text was adapted to fit the thesis.*

### Abstract

Recent advancements in 3D dust mapping have transformed our understanding of the Milky Way’s local interstellar medium, enabling us to explore its structure in three spatial dimensions for the first time. In this letter, we use the Edenhofer et al. (2023) 3D dust map to study the well-known Chameleon, Musca, and Coalsack cloud complexes, located about 200 pc from the Sun. We find that these three complexes are not isolated but rather connect to form a surprisingly well-defined half-ring, constituting a single “C”-shaped cloud with a radius of about 50 pc, a thickness of about 45 pc, and a total mass of about  $5 \times 10^4 M_{\odot}$  respectively  $9 \times 10^4 M_{\odot}$  for everything within its vicinity. Despite the absence of an evident feedback source at its center, the dynamics of young stellar clusters associated with the “C” structure suggest that a single supernova explosion about 4 Myr to 10 Myr ago likely shaped this structure. Our findings support a single origin story for these cloud complexes, suggesting that they were formed by feedback-driven gas compression, offering new insights into the processes that govern the birth of star-forming clouds in feedback-dominated regions, such as Sco-Cen.

### 5.1 Introduction

For over a hundred years, our understanding of the ISM has been limited to 2D projections. Analysis of two-dimensional projections, while a mainstay in tradi-

## 5. The “C”: The large Chameleon-Musca-Coalsack cloud

tional studies of the ISM, can create a deceptive picture of the true and complex three-dimensional structures of the ISM. For example, atomic and molecular clouds can appear as distinct or overlapping entities, obscuring their true spatial relationships. This illusion is especially problematic when clouds exhibit similar radial velocities, making it nearly impossible to distinguish their boundaries and depths using traditional observational methods.

The nearby molecular clouds of Chameleon, Musca, and Coalsack are a potential example of this type of confusion, having long been suspected to be physically associated (Corradi et al., 1997, 2004). However, their seemingly distinct appearances in the 2D plane of the sky has complicated any possible physical association. This projected view is potentially hiding diffuse connections that could shed new light on their origin and evolution.

In recent years, the proliferation of 3D dust maps has begun a revolution in the field (e.g., Chapter 4; Leike and Enßlin, 2019; Vergely et al., 2022; Lallement et al., 2022, 2019, 2018; Capitanio et al., 2017; Babusiaux et al., 2020; Hottier et al., 2020; Green et al., 2019, 2018; Leike et al., 2022; Chen et al., 2019; Rezaei Kh. and Kainulainen, 2022; Rezaei Kh. et al., 2020, 2018, 2017; Dharmawardena et al., 2022). These new maps capitalize on the astrometric data from the *Gaia* satellite (Gaia Collaboration, 2023a) and emerge as powerful tools to unravel the complex volume density distribution of gas within the local Milky Way, providing unprecedented insights into the structure and dynamics of molecular clouds from kiloparsec down to parsec scales (Zucker et al., 2023; Zucker et al., 2018; Zucker et al., 2019; Zucker et al., 2020; Alves et al., 2020; Zucker et al., 2021; Kuhn et al., 2022; Posch et al., 2023).

Chameleon, Musca, and Coalsack are among the closest molecular clouds to Earth at a distance of about 200 pc (see e.g. Zucker et al., 2021; Corradi et al., 2004). Chameleon has been extensively studied for its active star formation (e.g., Luhman, 2008) and Musca for its simple filamentary structure and subsonic nature (e.g., Mizuno et al., 1998; Hacar et al., 2016; Kainulainen et al., 2016). The Coalsack, despite being one of the few dark clouds visible to the naked eye, is one of the least studied nearby clouds, probably because it is not star-forming and is seen against the complicated background of the Galactic plane (e.g., Nyman, 2008; Beuther et al., 2011). So far, all three clouds have been treated as separate entities with conflicting origin stories. For example, Musca, is hypothesized to have been shaped by magnetic fields, by dissipation of supersonic turbulence, by a cloud-cloud collision, or a combination of these processes (Cox et al., 2016; Tritsis and Tassis, 2018; Tritsis et al., 2022; Bonne et al., 2020a,b; Yahia et al., 2021; Kaminsky et al., 2023).

In this letter, we study the region of Chameleon, Musca, and Coalsack from a new perspective, namely in full spatial 3D. Using the 3D dust map by Edenhofer et al., we characterize the topology and mass of this region and analyze the relationship between the three molecular clouds. In addition, we analyze the dynamics of young stellar clusters (YSO clusters) embedded in Chameleon and Coalsack as a proxy for the large-scale dynamics of the gas.

## 5.2 Methods and results

We use the 3D reconstruction of the interstellar dust distribution presented in Chapter 4 throughout our analysis. The reconstruction is based on the catalog described in Zhang et al. (2023) which in turn is based on the *Gaia* DR3 BP/RP spectra (Gaia Collaboration, 2023a). Using the stellar extinction and distances in the catalog, we reconstruct the 3D distribution of interstellar dust. The map extends to 1.25 kpc in distance from the Sun and achieves  $14'$  angular resolution and parsec-scale distance resolution.

The map presented in Chapter 4 is one of the highest resolution 3D dust maps of the local Milky Way. In comparison to other 3D dust maps in the literature, it focuses on a relatively small volume (c.f. Green et al., 2019), but resolves the structures within this volume at high resolution (parsec-scale) comparable to Leike et al. (2020). Compared to Leike et al. (2020), it reconstructs a much larger volume and yields a higher dynamic range thanks to the new Gaia data. The methodology extends the one from Leike et al. (2020) and employs its non-parametrically inferred correlation kernel. Furthermore, the model strictly enforces physical constraints such as positive definiteness of differential interstellar dust densities and rigorously incorporates distance uncertainties to stars.

### 5.2.1 Topology

Using the 3D map of the distribution of interstellar dust described in Chapter 4, we analyze the region around Chameleon, Musca, and Coalsack. For convenience we interpolate the 3D map with irregularly spaced voxels to a regular Cartesian and a regular spherical grid using the scripts provided by the authors online<sup>1</sup> and distributed as part of the Python package `dustmaps` (Green, 2018). We find that all three molecular clouds are embedded in a large 3D structure forming a “C”-shaped half-ring. The structure lies at the edge of the Local Bubble (Zucker et al., 2022; O’Neill et al., 2024; Pelgrims et al., 2020), close to the Scorpius-Centaurus association. Due to obscuration and projection effects, the peculiar “C”-shape is only revealed using 3D reconstructions of interstellar dust. In an ordinary 2D plane-of-sky projection, the diffuse bridges connecting the molecular clouds become indistinguishable from faint dust extinction at larger distances.

Figure 5.1 shows the extinction within the region over the range  $335^\circ$  to  $275^\circ$  in Galactic longitude and  $-40^\circ$  to  $15^\circ$  in Galactic latitude. The first panel shows the *Planck* 2013 extragalactic  $E(B - V)$  extinction of interstellar dust (Planck Collaboration, 2014a) converted to  $A_V$  via  $A_V = 3.1 \cdot E(B - V)$ . The “C”-shaped structure is almost completely obscured. The second panel shows the 3D interstellar dust map integrated over a 55 pc distance range from 165 pc to 220 pc. Here, the “C”-shaped structure appears as a single coherent structure hosting Chameleon, Musca, and Coalsack. The individual dense molecular clouds are connected through faint lanes of interstellar dust extinction. Thanks to the

<sup>1</sup><https://zenodo.org/doi/10.5281/zenodo.8187942>

## 5. The “C”: The large Chameleon-Musca-Coalsack cloud

3D distance selection, the figure shows this region free of confusion stemming from extinction at farther distances.

With 3D interstellar dust maps, we are not restricted to plane-of-sky projections and can instead explore the distribution of interstellar dust in full spatial 3D. In Figure 5.2 we show the isodensity surface in full spatial 3D at a total hydrogen nuclei density of approximately  $n_{\text{H}} = 4.5 \text{ cm}^{-3}$  (0.98-quantile in the selected volume; see O’Neill et al. (2024) for converting the units of the 3D dust map to  $n_{\text{H}}$ ). The outline traces the “C”-shaped spine observed in Figure 5.1. The isodensity surface fully envelops the half-ring including the Chameleons, Musca, and Coalsack molecular clouds. We term the “C”-shaped structure the “C”.

Based on its suggestive half-ring structure, we defined a center for the “C”. We defined its center to lie below the Coalsack in Galactic X and Y and slightly above the Galactic Z of the two Chameleons. Specifically, we set the center to be at Galactic X=138 pc, Y=−140 pc, and Z=−33 pc, respectively, at Galactic  $l = 314.6^\circ$ ,  $b = -9.5^\circ$ , and  $d = 199.3$  pc (see interactive Figure 5.2). We tested random shifts of  $\pm 10$  pc in Galactic X, Y, and Z for the center and validated that the qualitative results presented in the following are insensitive with respect to the precise choice of center.

Relative to the center location, we defined a radius using the volume filling fraction of interstellar dust. We find that the volume filling fraction reaches a maximum at a distance of 48 pc and we assume this to be the radius of the “C”. This estimate is robust with respect to different posterior samples of the dust distribution described in Chapter 4. The full-width-half-maximum (FWHM) of the peak is  $(47 \pm 2)$  pc and most of the dust lies within a distance of 80 pc from the center (ceiling of the more distant edge of the FWHM).

We find that the “C” is almost perfectly parallel to the Galactic Z axis. Given its proximity to the Galactic disk, this makes it perpendicular to the line of sight. To quantify the inclination, we selected the 10,000 highest density points of a 1 pc interpolation of the 3D dust map within  $20 \leq X \leq 230$  pc,  $-240 \leq Y \leq -90$  pc, and  $-185 \leq Z \leq 50$  pc and decomposed them with a singular value decomposition into a plane. We find that the Galactic Z axis and the plane’s normal vector form a  $86^\circ$  angle (2/3 of samples within  $\pm 1^\circ$ ) and the inclination between the LOS toward the center and the fitted plane to be  $72^\circ$  (2/3 of samples within  $\pm 1^\circ$ ). Table 5.1 summarizes key properties of the “C”.

### 5.2.2 Mass

To estimate the mass, we again utilize the 3D interstellar dust density. We adopt the conversion ratio  $n_{\text{H}} = 1653 \text{ cm}^{-3} \rho$  as derived in O’Neill et al. (2024) using Draine (2003, 2009) to convert our interstellar dust density  $\rho$  to a total hydrogen nuclei density  $n_{\text{H}}$ . Analogously to O’Neill et al. (2024) we then convert the hydrogen nuclei density to a mass via  $M = 1.37 \cdot m_p \cdot \sum_i n_{\text{H},i} \cdot dv_i$ , adopting a mean molecular weight of hydrogen ( $\mu$ ) of 1.37,  $m_p$  the proton mass, and  $dv_i$  the volume of the  $i$ th voxel  $n_{\text{H},i}$ . We integrate the mass from the center out

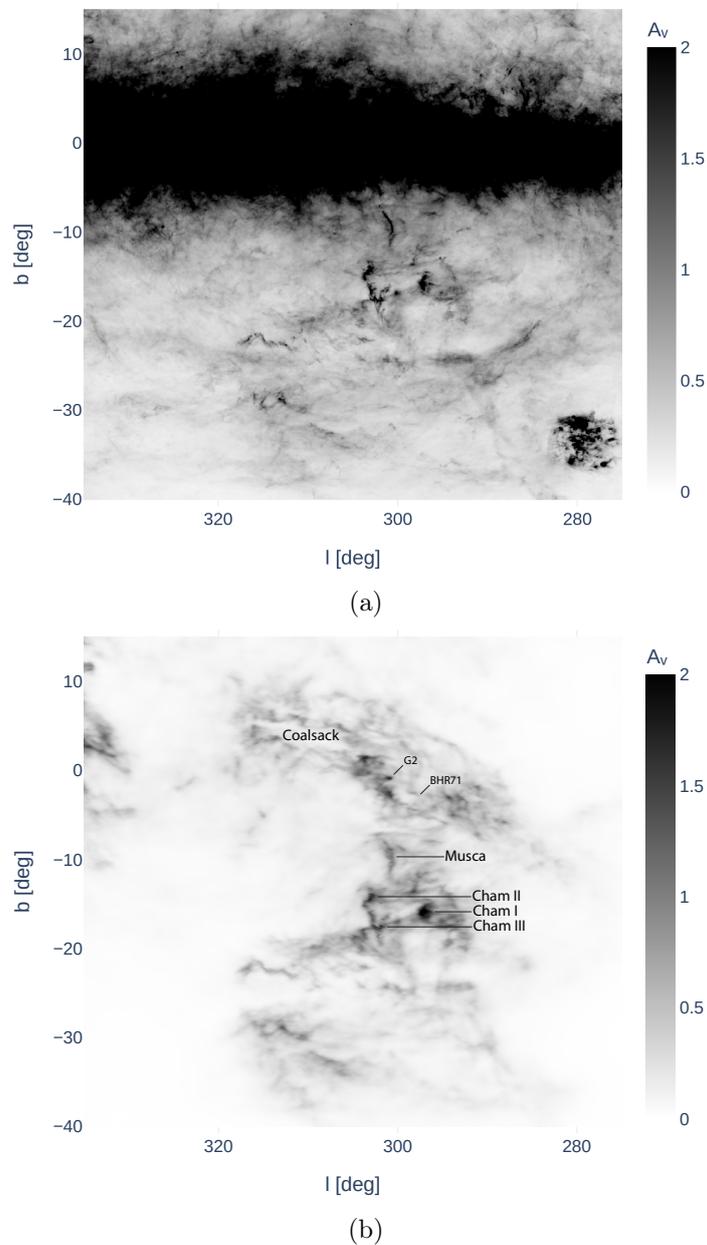


Figure 5.1: Plane-of-sky region toward the “C” of the posterior mean of the map described in Chapter 4. Panel (a): *Planck* 2013 extragalactic  $E(B-V)$  extinction toward the “C” converted to  $A_V$  via  $A_V = 3.1 \cdot E(B - V)$ . Panel (b): Visual extinction  $A_V$  between 165 pc to 220 pc toward the “C”-shaped structure in the 3D interstellar dust map. Both figures display the extinction in units of magnitude and the colorbars are linear but clipped at an extinction of  $A_V = 2$  mag.

### 5. The “C”: The large Chameleon-Musca-Coalsack cloud

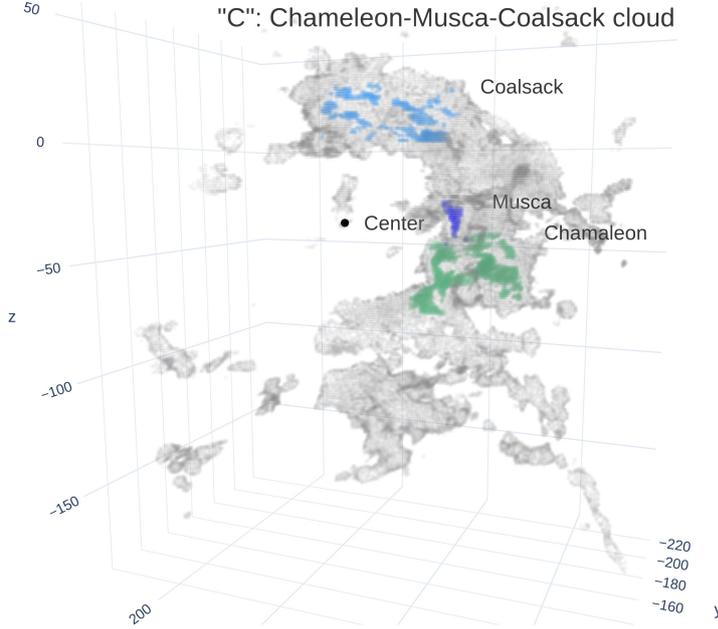


Figure 5.2: 3D view of the isodensity surface of the posterior mean of the 3D dust map of Edenhofer et al. within the Cartesian selection given in Table 5.1 showing the “C”. High density structures ( $n_{\text{H}} \geq 23.3 \text{ cm}^{-3}$ ; 10,000 highest density points of a 1 pc Cartesian interpolation of the “C”) toward Chameleon ( $290^\circ \leq l \leq 306^\circ$  and  $-21^\circ \leq b \leq -12^\circ$ ) are color-coded in green, toward Musca ( $300^\circ \leq l \leq 302.5^\circ$  and  $-13^\circ \leq b \leq -7.25^\circ$ ) in violet, and toward Coalsack ( $300^\circ \leq l \leq 318^\circ$  and  $0.5^\circ \leq b \leq 7.5^\circ$ ) in blue. An interactive version of this figure is available at [https://faun.rc.fas.harvard.edu/gedenhofer/perm/C/C\\_Chameleon\\_Musca\\_Coalsack\\_cloud.html](https://faun.rc.fas.harvard.edu/gedenhofer/perm/C/C_Chameleon_Musca_Coalsack_cloud.html).

Table 5.1: Key properties of the “C”. The specified uncertainties encompass at least  $2/3$  of the 3D dust map posterior samples. The specified positions capture the core features of the “C” while including as little as possible from other molecular clouds. The spherical and Cartesian cuts are not equivalent as for each of them we tried to minimize projection and obscuration effects from molecular clouds not affiliated with the “C” (c.f. Figures 5.1 and 5.3).

	Cartesian Position	Spherical Position	Radius	FWHM	Inclination	Mass
“C”	$20 \leq X/\text{pc} \leq 230$ $-240 \leq Y/\text{pc} \leq -90$ $-185 \leq Z/\text{pc} \leq 50$	$275 \leq l^\circ \leq 335$ $-40 \leq b^\circ \leq 15$ $165 \leq d/\text{pc} \leq 220$	48 pc (all samples)	$(29.5 \pm 1.0) \text{ pc}$ – $(76.0 \pm 0.5) \text{ pc}$	$(72 \pm 1)^\circ$	$(50.5 \pm 0.3) \times 10^3 M_\odot$ in isodensity surface
Center	$X=137 \text{ pc}$ $Y=-140 \text{ pc}$ $Z=-33 \text{ pc}$	$l=314.6^\circ$ $b=-9.5^\circ$ $d=199.3 \text{ pc}$				$(93.4 \pm 0.3) \times 10^3 M_\odot$ in vicinity

to a distance of 80 pc or the edge of the Cartesian selection given in Table 5.1, whichever is lower, and find the mass of everything within the vicinity of the “C” to be  $(9.34 \pm 0.03) \times 10^4 M_\odot$  (statistical uncertainty covering the 0.16- to 0.84-quantile of the interstellar 3D dust map). This mass estimate is roughly in between the total mass of everything above a density of  $n_H = 1 \text{ cm}^{-3}$  and  $n_H = 2 \text{ cm}^{-3}$  within the Cartesian selection given in Table 5.1 ( $(1.12 \pm 0.01) \times 10^5 M_\odot$  and  $(8.00 \pm 0.08) \times 10^4 M_\odot$ , respectively). The mass of everything within the isosurface shown in Figure 5.2 (i.e. above approximately  $n_H = 4.5 \text{ cm}^{-3}$ ) is  $(5.05 \pm 0.03) \times 10^4 M_\odot$ .

### 5.2.3 Dynamics

To investigate the dynamics of the peculiar “C”-shaped structure, we study YSO clusters embedded in the “C”. We use the catalogs of Ratzenböck et al. (2023) and Hunt and Reffert (2023), applying the following cuts to the catalog by Hunt and Reffert:  $\log \text{Age} \leq \log_{10}(30 \text{ Myr})$ , Astrometric SNR  $\geq 5$ , and 50th percentile of Color Magnitude Diagram (CMD) class  $\geq 0.5$  (c.f. Hunt and Reffert, 2023). Nested inside the “C”, we find three YSOs clusters: Centaurus-Far (HSC 2630 in Hunt and Reffert, 2023), Chameleon I (Hunt and Reffert, 2023; Ratzenböck et al., 2023), and Chameleon II (Ratzenböck et al., 2023). In the following, we use the velocities given in (Ratzenböck et al., 2023).

Relative to the local standard of rest (Schönrich et al., 2010), we find that the two Chameleon clusters are moving toward negative Galactic Z more strongly than Centaurus-Far. This implies that the “C” is expanding in Galactic Z. If we assign our center point a hypothetical velocity of  $1/2 \cdot (v_{\text{CenFar}} + 1/2 \cdot (v_{\text{ChamI}} + v_{\text{ChamII}}))$  with  $v_\square$  the velocities of Centaurus-Far and the two Chameleons respectively, we find that Centaurus-Far moves up and the two Chameleons move down relative to the center. Their velocity relative to the center is 3 km/s for Centaurus-Far and 4 km/s and 2 km/s for the two Chameleons.

As the YSOs are embedded in the “C”, we propose that not only the cluster move away at an average velocity of 3 km/s but the whole “C” expands at an average velocity of 3 km/s. Moving the mass in the “C” at this speed requires a significant amount of energy. Specifically, moving  $9 \times 10^4 M_\odot$  at 3 km/s requires an energy input of  $8 \times 10^{48}$  erg. This energy input is approximately 1% of the total energy release of a supernova feedback event and on the order of the expected kinetic energy input of a supernova (Kim and Ostriker, 2015).

The observation is robust with respect to the assumed center velocity and cluster velocities. To test the dependence on the center velocity, we assumed that the center velocity is the average velocity of all YSO clusters in the vicinity (clusters with 0.84-age-quantile  $\leq 30 \text{ Myr}$ ) in the catalog of Hunt and Reffert (2023) within  $-20 \leq X \leq 260 \text{ pc}$ ,  $-300 \leq Y \leq -10 \text{ pc}$ , and  $-240 \leq Z \leq 80 \text{ pc}$  and find a qualitatively equivalent result. We tested the sensitivity to the cluster velocities by studying the systematic and statistical uncertainties in the clusters.

## 5. The “C”: The large Chameleon-Musca-Coalsack cloud

The uncertainty in the bulk motion is dominated by the spread in the individual YSO velocities (order of magnitude of the cluster velocity itself). The true cluster velocity uncertainty is likely in between both uncertainties and leads to errors on the order of the relative velocity that we found above. As a proxy for the systematic uncertainties in the clustering in the catalog used here, we validated that the qualitative result holds equivalently for the cluster velocities from Hunt and Reffert (2023) for Centaurus-Far and Chameleon I. Note, other traces for the velocity such as HI or CO are unavailable as in addition to obscuration and confusion, the inclination of the “C” results in its expansion being perpendicular to the line of sight.

### 5.2.4 Age

We analyzed the relative position of the clusters in the past by propagating their current position and velocity back in time using the software package `galpy` and the `MWPotential2014` Galactic potential (Bovy, 2015). In doing so, we neglect other gravitational forces such as the cluster’s own gravitational potential, gravitational forces due to nearby molecular clouds, respectively any other acceleration not due to the Milk Way’s potential. We find that they have been approaching each other in the past 4 Myr to 10 Myr relative to the center with the center velocity described above. Centaurus-Far gets closest to the center with a minimum separation of 25 pc while the two Chameleons get as close as 46 pc respectively 40 pc to the center in the last 10 Myr. The observation is robust with respect to the precise choice of center velocity and the minimum separations vary by 10 pc or less if we assume the velocity of the center to be the average velocity of all YSO clusters in the vicinity (see Section 5.2.3). Given the large uncertainty in the velocities and ages and the simplistic model, these discrepancies appear modest, and the velocities and ages seem in good agreement with the hypothesis of an expanding half-ring.

Figure 5.3 shows the “C” in Cartesian X-Z-projections. The three clusters and their tracebacks relative to the center are indicated by colored lines.

### 5.2.5 Caveats

We have determined the shape and mass of the “C” with considerable accuracy using 3D interstellar dust maps. However, the dynamics and energy calculations remain uncertain. The “C”’s overall shape and the velocities of YSO clusters suggest it may be an expanding half-ring, but our analysis is constrained by the limited number of YSO clusters observed. Additionally, the uncertainties in the velocity measurements are comparable to the relative uncertainties, further complicating the analysis. Finally, although the “C” surrounds a largely empty space, there is evidence of faint interstellar dust density at its center, as shown in Figure 5.3.

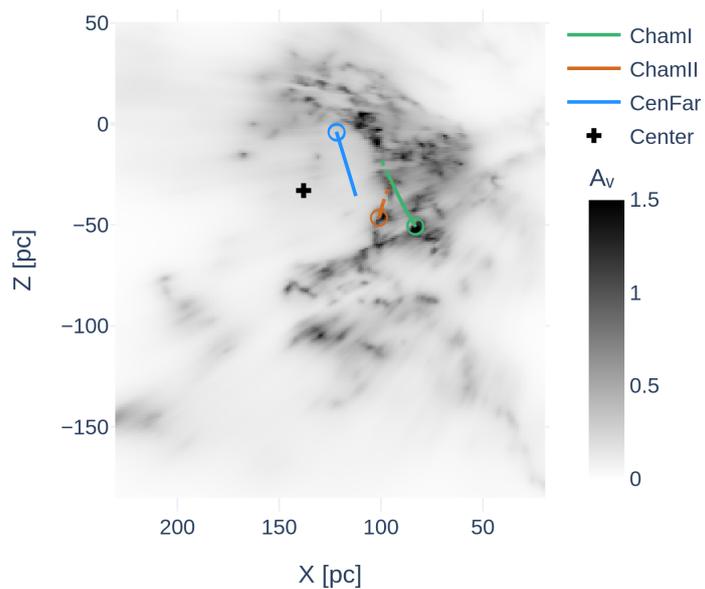


Figure 5.3: Cartesian X-Z-projections of the “C”. The colorbar is linear and clipped at  $A_V = 1.5$  mag. The traced back clusters relative to the center are shown as colored lines on top. The solid lines show the traces of the cluster up to their 0.84-age-quantile while the dash-lines show the traces back to  $-10$  Myr whenever their 0.84-age-quantile is lower. See the interactive version of Figure 5.2 and toggle on the clusters to see them in full spatial 3D.

### 5.3 Discussion and conclusion

3D maps of interstellar dust have unlocked new ways to explore the intricate structures within the ISM in comprehensive 3D detail. Using these maps, we have established that the Chameleon, Musca, and Coalsack molecular clouds are spatially connected, forming a “C”-shaped structure that encompasses a striking, well-defined cavity. This structure, which we refer to as the “C”, is a half-ring structure with a radius of approximately 50 pc and a mass of about  $5 \times 10^4 M_{\odot}$  respectively  $9 \times 10^4 M_{\odot}$  for everything within its vicinity. The “C” extends across roughly  $50^{\circ}$  in both longitude and latitude, and has a depth of about 55 pc. The structure is discernible only through 3D reconstructions of interstellar dust, as obscuration and projection effects obscure it in conventional 2D views.

We propose a single origin for all three molecular clouds comprising the “C”. The orbits of clusters in the “C” are indicative of an expansion that when combined with the mass of the “C” requires the energy released by a single supernova explosion. We suggest that a singular supernova explosion shaped an existing dense cloud into the “C” between 4 Myr to 10 Myr ago.

The realization that the Chameleon, Musca, and Coalsack molecular clouds are interconnected in 3D space introduces a novel perspective for their study. These previously distinct clouds are now understood to be parts of a single large cloud, the “C”. This broader context is especially significant for understanding Musca, where a large body of observational work has been acquired over the last decade to explain its formation. Viewing Musca within the scope of the “C” suggests a relatively simple formation mechanism: cloud formation on an expanding ring, likely driven by stellar feedback. This insight suggests that feedback driven expansion is the primary cause for the shape of Musca (c.f. Inutsuka et al., 2015, for role of magnetic fields), and calls for a new analysis of the existing data on this cloud. This new perspective on Musca’s formation and shape is a reminder that the environment in which a molecular cloud resides can significantly influence its formation mechanism and morphology. This work highlights the importance of considering the broader context offered by the new 3D dust maps when interpreting molecular cloud formation and evolution.

### Acknowledgements

The authors would like to thank Laura Posch and Cameren Swiggum for many helpful comments and discussions throughout the work. Gordian Edenhofer acknowledges that support for this work was provided by the German Academic Scholarship Foundation in the form of a PhD scholarship (“Promotionsstipendium der Studienstiftung des Deutschen Volkes”). This work was co-funded by the European Union (ERC, ISM-FLOW, 101055318). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

## 6 Additional contributions

This chapter outlines work to which I contributed but where I was not the lead author. These projects demonstrate the versatility and usefulness of the developed methods and the importance of 3D dust in studying the ISM.

### 6.1 JAXbind: Bind any function to JAX

JAX has proven to be an incredibly powerful tool for handling large statistical models. Of particular interest when working with JAX is its transformation system for taking arbitrary high derivatives of functions, batching computations, and just-in-time compiling code for additional performance. JAX has become widely adopted in the scientific community and underpins `NIFTy.re`. However, JAX's transformation system requires that all parts of the computation be written in JAX. To make `NIFTy.re` a truly generic modeling toolkit, we need to interact with functions (and their derivatives) that cannot be efficiently expressed in JAX due to JAX's control flow constraints, its focus on regular-shaped tensors, or simply due to time constraints. One such function that we would like to use in `NIFTy.re` is the highly optimized radio-interferometric gridder implemented in C++ and described in Arras et al. (2021). Implementing the gridder natively in JAX is infeasible due to JAX's focus on regular-shaped tensors.

Motivated by our desire to do radio-interferometric imaging with `NIFTy.re`, Jakob Roth, Martin Reinecke, and I developed a generic software package called `JAXbind`, which bridges foreign functions and their derivative to JAX. The package complements the focus of `NIFTy.re` on JAX. To the best of our knowledge, no other code currently exists for connecting generic functions with full derivative support to JAX. We describe the software in an equal-author article (Roth et al., 2024) which is currently under review at the Journal of Open Source Software. Jakob Roth and I wrote most of the Python code and extended the C++ code of which Martin Reinecke wrote the initial version and helped in its later development.

Internally, `JAXbind` registers a so-called primitive with JAX. To JAX, this primitive appears like a generic JAX code that supports all its transformation rules. Users are required to hand-code their derivatives and transposition rules, but they only need to expose these functions through a standard Python interface. Whenever JAX calls any transformation of our primitive, we call back

## 6. Additional contributions

to the appropriate code where the user implemented these transformations in plain Python. `JAXbind` calls these functions at the right time such that codes connected to JAX via `JAXbind` can be utilized as if they were JAX-native.

`JAXbind` has proven highly useful in situations where new code is written in JAX but where it is still desirable to use previous codes not written in JAX. It enabled Jakob Roth to connect the radio-interferometric gridded described in Arras et al. (2021) to `NIFTy.re` and allowed for a much larger reconstruction by combining the previous gridded with JAX (Roth et al., 2024). `JAXbind` also enabled Andres Ramirez to connect a cosmic ray propagation code (PICARD; Kissmann, 2014) to JAX for inferring the cosmic ray density distribution in an ongoing research project.

## 6.2 Galactic 3D large-scale distribution of interstellar dust

In an unpublished article led by Reimar Leike (Leike et al., 2022), a predecessor of the GP methodology described in Chapter 2 is applied to data from PANSTARRS, 2MASS, Gaia DR2, and ALLWISE to reconstruct the 3D distribution of interstellar dust out to 4kpc. To accommodate the large number of degrees of freedom, the reconstruction is split into small cones that were reconstructed separately. Only in a final post-processing step are the cones aligned with one another. Compared to the methodology described in Chapter 4, the work uses a simpler voxelization scheme that does not allow for reconstructing the full sky all at once and approximates the posterior much more crudely. The reconstruction successfully demonstrated the idea behind Chapter 2, but the voxelization scheme, the splitting scheme, and the posterior approximation approach led to strong artifacts throughout the map. See Chapter 4 for a comparison of this map with our new map.

Reimar Leike wrote most of the paper, led the development, and managed the reconstruction. My contribution entailed an extensive validation of the statistical methodology for the spatial correlation prior. I demonstrated that, for the specific kernel used, the model incurs errors on the order of 10% in representing the variance between voxels at small distances.

## 6.3 Studies of the local ISM

I have contributed to several works that have utilized 3D dust maps for the study of the ISM. Here, I would like to highlight a few notable articles and summarize my contributions.

In an article led by Bouzelou Tritsis, the shape of the Musca molecular cloud (Tritsis et al., 2022) is studied. Tritsis et al. present evidence suggesting that Musca is radially elongated. The evidence is primarily based on theoretical considerations and POS-projected polarization measurements tracing the magnetic

field. My contribution to the article involved an extensive analysis of the 3D dust map described in Leike et al. (2020) and the data informing it in the region around Musca. The key data constraints for Musca’s POS and radial extent are shown in Figure 6.1. My findings indicate that the data in the 3D dust map is insufficient to determine whether Musca is radially elongated on scales relevant to the study.

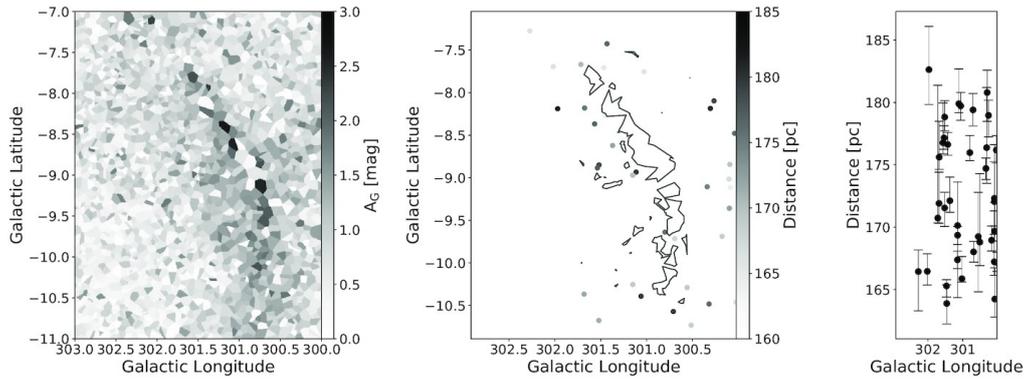


Figure 6.1: POS and radial data constraints for Musca. Left-hand panel: Voronoi tessellation of  $A_G$  extinction from 2187 stars within 500 pc from the Sun. Musca is visible as a more extinct thin strip. Middle panel: Stars within the POS and radial vicinity of Musca (distance range 165 pc to 180 pc), color coded by their distance. The contour shows the region with  $A_G \geq 1.5$  mag from the first panel and roughly traces Musca. Right-hand panel: Median distances of stars and their uncertainties over the Galactic longitude of the stars. In the POS and radial vicinity of Musca, only 37 stars directly constrain the cloud’s 3D morphology. Figure created by me and taken from Tritsis et al. (2022).

In an article led by Kristina Monsch, the discovery and initial observations of a giant edge-on protoplanetary disk (Monsch et al., 2024) is presented. The protoplanetary disk is particularly noteworthy because it does not appear to be connected to any notable star-forming region. My contribution to this research involved analyzing the 3D distribution of interstellar dust in the vicinity of the disk. I identified dense molecular clouds along the LOS, which might be potential birthplaces of the star. This analysis provided valuable constraints on the distance to the protoplanetary disk and helped constrain its size.

In O’Neill et al. (2024), an article led by Theo O’Neill, the precise shape of the Local Bubble is studied. The newly inferred shape of the Local Bubble is shown in Figure 6.2. The analysis suggests that the Local Bubble more closely resembles a Local Chimney with an open top. The new analysis is enabled by the 3D dust map described in Chapter 4, which extends to much larger distances, including at high Galactic latitudes. I contributed to the study with early access to the 3D dust map and an extensive analysis of peculiarities in 3D dust maps for the study of faint structures at low distances.

## 6. Additional contributions

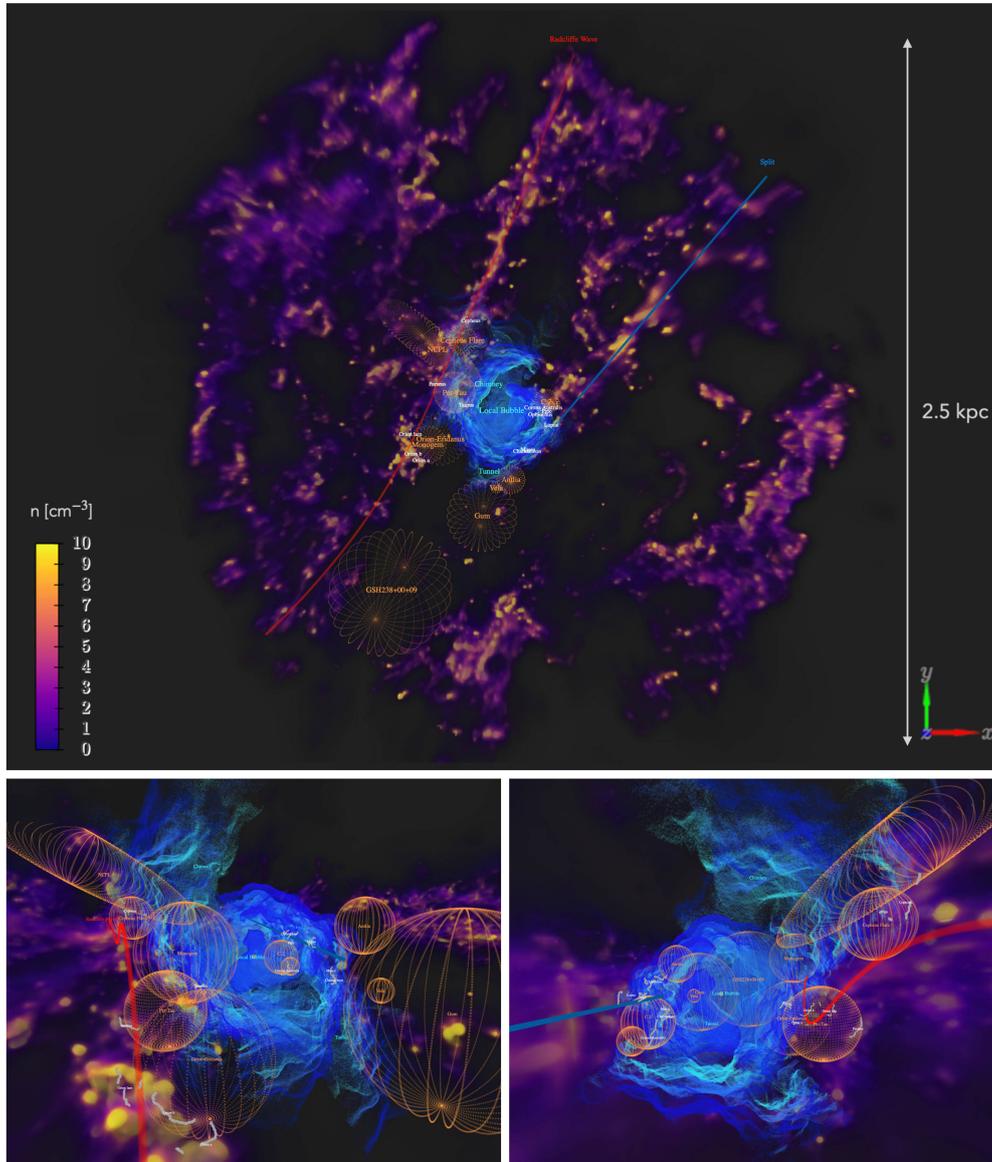


Figure 6.2: A 3D view of the Local Bubble respectively the Local Chimney inferred by O’Neill et al. from the map described in Chapter 4. The inner, peak, and outer edges of the Local Chimney’s shell are shown in light blue, medium blue, and dark blue, respectively (see O’Neill et al., 2024, for further details). The structure is overlaid on top of our 3D map of interstellar dust. An interactive figure is available at <https://theo-oneill.github.io/localbubble/neighborhood/>. Figure created by Theo O’Neill and adapted from O’Neill et al. (2024).

Beyond the analysis of interstellar dust, I contributed to exploring the magnetic field in the ISM. In two articles led by Alexandros Tsouros (Tsouros et al., 2024b,a), a method for probing the 3D distribution of the magnetic field as a means to trace back ultra-high-energy cosmic rays is developed. The method is validated on synthetic data. See Figure 6.3 for a reconstruction of the magnetic field in 3D from synthetic data. I contributed to the articles through extensive discussions and pair-programming sessions focused on developing a divergence-free model for the magnetic field.

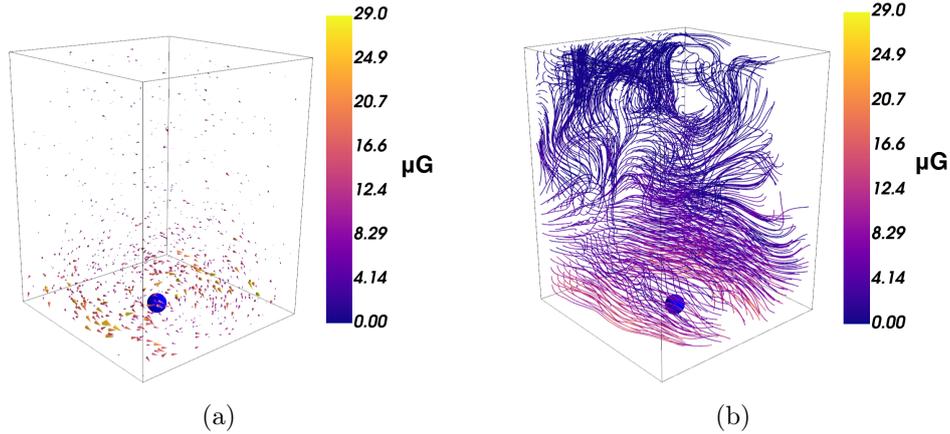


Figure 6.3: A 3D view of a magnetic field reconstruction from synthetic data. Panel (a): Local and sparse POS synthetic magnetic field measurements. The observer location is highlighted in blue. Panel (b): Posterior mean of the magnetic field reconstruction from the local and sparse POS synthetic magnetic field measurements. Figures created by Alexandros Tsouros and adapted from Tsouros et al. (2024a).

## 6.4 Applications outside of astrophysics

In addition to exploring the physics of the ISM, I contributed to a project applying similar statistical and computational methods outside of astrophysics. In a publication led by Matteo Guardiani (Guardiani et al., 2022), an IFT-based method is developed to quantify the causal relationship between the age and virus load in patients diagnosed with COVID-19. My role in this project involved writing code for the causal inference model. Specifically, I helped write the final algorithm that models probability density functions as smooth normalized functions. This model is a key component for the quantification of the causal relationship. The algorithm for inferring smooth probability densities from noisy count data which was developed as part of this article, was published as a separate software publication (Edenhofer et al., 2023a).



# 7 Conclusion

## 7.1 Summary

The thesis discusses the creation of the largest high-resolution 3D map of interstellar dust to date. We begin by examining the required statistical GP methodology for arbitrarily spaced volume elements, an essential ingredient for modeling interstellar dust in 3D. Our new method represents the modeled volume at multiple resolutions, coupling voxels only locally but across resolutions. Thus, all voxels are correlated, though correlations become coarser with increasing distance between voxels. The multi-resolution view results in an algorithm that is both accurate in representing the desired GP and highly efficient. Compared to the current state-of-the-art, the algorithm is about two orders of magnitude faster. The developed approach scales linearly with the number of volume elements, enabling the modeling of larger volumes at high resolutions.

Next, we discuss `NIFTy.re`, a novel Bayesian probabilistic modeling framework for handling the hundreds of millions to billions of degrees of freedom encountered in reconstructions of the 3D interstellar dust distribution. The framework, a rewrite of `NIFTy` in `JAX`, introduces a new modeling approach and a new optimization concept, and offers much more flexible posterior inference strategies compared to the original `NIFTy`. On average, it speeds up models by two orders of magnitude.

The main result of the thesis is our new 3D dust map. Utilizing processed Gaia data (Zhang et al., 2023) and the methods described in Chapters 2 and 3, we reconstruct the largest high-resolution 3D dust map to date. This map models the density of dust via a log-normal GP and carefully accounts for the distance uncertainties. The reconstruction extends out to 1.25 kpc in distance, features an angular resolution of 14' and a parsec-scale distance resolution.

The map encompasses more than half a billion voxels. As the reconstruction of the map is not an algebraic inversion problem but a statistical one with uncertainties in the result, we provide not just one true reconstruction but 12 posterior samples drawn via the MGVI algorithm (Knollmüller and Enßlin, 2019). For better accessibility of the map, we interpolate the map to spherical and regular Cartesian coordinates and provide a set of interpolation scripts to create custom zoom-ins in either coordinate frame. Additionally, we integrated the map into the software package `dustmaps` (Green, 2018) for easy access.

## 7. Conclusion

Our new map offers a fresh perspective for studying the ISM. It provides a detailed parsec-scale 3D view of molecular clouds and their internal structure. The map improves upon previous maps in terms of dynamic range and volume covered at high resolution.

Using our new 3D map of interstellar dust, we study three famous molecular clouds in the vicinity of the Sun — Musca, Chameleon, and Coalsack. We find these clouds to be connected in 3D. We characterize the region and find the molecular cloud which encompasses Musca, Chameleon, and Coalsack to form a “C”-shaped structure. We term the molecular cloud the “C”. It has a radius of about 50 pc and encompasses about  $5 \times 10^4 M_{\odot}$  respectively  $9 \times 10^4 M_{\odot}$  for everything within its vicinity.

We extend our spatial 3D perspective with 3D velocity data from YSO cluster data. Using the velocity information from these clusters, we propose a feedback-driven origin story for the entire region. Although we only have limited velocity information from clusters, the observed shape and velocity patterns suggest that a supernova feedback event shaped the “C”, providing a simple explanation for the topology of the whole region. The true shape of the structure is completely obscured in POS projections, underscoring the limitations of traditional two-dimensional observations. Our study emphasizes the critical importance of examining the ISM in 3D.

Lastly, we discuss miscellaneous projects inside and outside of astrophysics to which I contributed. I contributed to the development of a library for bridging scientific software to JAX. This library has already been utilized for radio-interferometry applications with `NIFTy.re` (Roth et al., 2024). Furthermore, I contributed to several articles that utilized the 3D dust map and helped in applying the developed methods within and beyond astrophysics.

## 7.2 Outlook

This thesis sets the stage for studying a significant fraction of the Milky Way’s ISM in 3D. Our 3D map of interstellar dust provides the backbone for probing more components of the ISM in 3D and for probing much larger volumes in the future.

We envision that our reconstruction of the 3D distribution of interstellar dust to be used to help constrain the distribution of HI and CO in 3D. By combining interstellar dust with HI and CO measurements, we can leverage dust to disentangle POS observations of HI and CO while at the same time assigning radial velocities to interstellar dust. Reconstructing HI and CO using interstellar dust will enable us to quantify the spatial variation in the dust-to-gas ratio, a vital ingredient for understanding star formation.

We not only envision radial velocities to be added to interstellar dust maps but full 3D velocity information from *Gaia*’s proper motion measurements. YSO clusters provide an excellent data source to constrain the 3D velocity vector

space. With a 3D velocity vector field reconstruction, we will be able to study the impact of feedback-driven expansion as well as galactic shear on molecular clouds.

Looking further ahead, many more components could be combined with 3D interstellar dust maps. With rigorous physical models, we envision all-sky maps of the 3D distribution of the temperature of dust by combining 3D maps of dust extinction with POS emission measurements at multiple frequencies. Furthermore, we envision 3D maps of the cosmic ray density from gamma-ray data as well as 3D magnetic field reconstructions based on stellar polarization measurements (e.g. with PASIPHAE data in the near future).

At the same time that we are probing more components of the ISM, we should also reach out further in distance. We need to increase our sample size of molecular clouds to quantify potential local biases and to study the dependence of molecular clouds on their galactic environment. Furthermore, we have yet to see galactic-scale structures like spiral arms in the ISM. Studying their formation and interaction with medium-scale structures would teach us a lot about the dynamics of our Milky Way. Future research will be able to answer what drives molecular cloud formation on scales of hundreds of parsecs, how medium- to large-scale structures such as the Split (Lallement et al., 2019) or the Radcliffe wave (Alves et al., 2020) form, and what impact the galactic environment has on molecular cloud properties.

Interstellar dust will continue to be the quantity of choice for probing the Milky Way in 3D at high resolution and *Gaia* will continue to be the backbone of 3D interstellar dust maps. The upcoming fourth data release (DR4) from *Gaia* will need to be included in all future 3D dust maps. It is expected to be another major step forward in quantity and quality, releasing a much larger volume of low resolution spectra. *Gaia* DR4 will provide pristine extinction measurements for up to ten times more stars compared to the current third data release. The data will provide an almost complete sample of stars in the solar neighborhood.

To push out to larger distances, we need to turn to ground-based measurements. While no ground-based catalog achieves a similar coverage and fidelity as *Gaia*, they can make up for that with quality or quantity. Spectroscopic measurements can provide precise anchor points at large distances to constrain the next generation of 3D dust maps while photometric measurements can provide the necessary data quantity for a high angular resolution. For reconstructions focusing on the Galactic disk, data from the Dark Energy Camera Plane Survey (DECaPS; Saydjari et al., 2023) and the VISTA Variables in the Via Lactea (VVV; Minniti et al., 2010) survey is already readily available and will enable probing much higher extinctions. Off the Galactic plane, the SDSS-V Milky Way Mapper mission, a spectroscopic survey of the Milky Way, is set to probe regions much beyond *Gaia*, and will be a perfect complement for larger distances, promising to expand our observational horizon substantially.

To incorporate these new datasets and significantly increase the reconstructed volume further, new and better algorithms will remain a high priority. For prob-

## *7. Conclusion*

ing larger volumes, we will need to incorporate more distant stars. This will require handling complex distance uncertainties, in particular coupled distance and extinction uncertainties. A completely new approach to represent and sample from likelihoods that incorporate these complex, coupled uncertainties is required. Furthermore, the GP methodology needs to be extended with a highly distributed computing infrastructure in mind to probe ever larger volumes with soon-to-be tens of billions of voxels.

# Bibliography

- Alves, J., Zucker, C., Goodman, A. A., Speagle, J. S., Meingast, S., Robitaille, T., Finkbeiner, D. P., Schlafly, E. F., and Green, G. M. A Galactic-scale gas wave in the solar neighbourhood. *Nature*, 578(7794):237–239, February 2020. doi: 10.1038/s41586-019-1874-z.
- Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., and O’Neil, M. Fast Direct Methods for Gaussian Processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38:252, June 2015. doi: 10.1109/T-PAMI.2015.2448083.
- Anders, F., Khalatyan, A., Chiappini, C., Queiroz, A. B., Santiago, B. X., Jordi, C., Girardi, L., Brown, A. G. A., Matijević, G., Monari, G. et al. Photo-astrometric distances, extinctions, and astrophysical parameters for Gaia DR2 stars brighter than  $G = 18$ . *A&A*, 628:A94, August 2019. doi: 10.1051/0004-6361/201935765.
- Anders, F., Khalatyan, A., Queiroz, A. B. A., Chiappini, C., Ardèvol, J., Casamiquela, L., Figueras, F., Jiménez-Arranz, Ó., Jordi, C., Monguió, M. et al. Photo-astrometric distances, extinctions, and astrophysical parameters for Gaia EDR3 stars brighter than  $G = 18.5$ . *A&A*, 658:A91, February 2022. doi: 10.1051/0004-6361/202142369.
- Arras, P., Baltac, M., Ensslin, T. A., Frank, P., Hutschenreuter, S., Knollmueller, J., Leike, R. H., Newrzella, M.-N., Platz, L., Reinecke, M. et al. NIFTy5: Numerical Information Field Theory v5. Astrophysics Source Code Library, record ascl:1903.008, 03 2019.
- Arras, P., Reinecke, M., Westermann, R., and Enßlin, T. A. Efficient wide-field radio interferometry response. *A&A*, 646:A58, February 2021. doi: 10.1051/0004-6361/202039723.
- Arras, P., Frank, P., Haim, P., Knollmüller, J., Leike, R. H., Reinecke, M., and Enßlin, T. A. Variable structures in M87\* from space, time and frequency resolved interferometry. *Nature Astronomy*, 6:259–269, January 2022. doi: 10.1038/s41550-021-01548-0.
- Astropy Collaboration and Astropy Contributors. The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *AJ*, 156(3):123, September 2018. doi: 10.3847/1538-3881/aabc4f.
- Astropy Collaboration and Astropy Project Contributors. The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the

## BIBLIOGRAPHY

- Latest Major Release (v5.0) of the Core Package. *ApJ*, 935(2):167, August 2022. doi: 10.3847/1538-4357/ac7c74.
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M. et al. Astropy: A community Python package for astronomy. *A&A*, 558:A33, October 2013. doi: 10.1051/0004-6361/201322068.
- Babusiaux, C., Fournelle-Ravard, C., Hottier, C., Arenou, F., and Gómez, A. FEDReD. I. 3D extinction and stellar maps by Bayesian deconvolution. *A&A*, 641:A78, September 2020. doi: 10.1051/0004-6361/202037466.
- Bakhvalov, N. S. On the convergence of a relaxation method with natural constraints on the elliptic operator. *USSR Computational Mathematics and Mathematical Physics*, 6(5):101–135, 1966.
- Bergin, E. A. and Tafalla, M. Cold Dark Clouds: The Initial Conditions for Star Formation. *ARA&A*, 45(1):339–396, September 2007. doi: 10.1146/annurev.astro.45.071206.100404.
- Beuther, H., Kainulainen, J., Henning, T., Plume, R., and Heitsch, F. The coalsack near and far. *arXiv [astro-ph.SR]*, July 2011.
- Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletos, T., Singh, R., Szerlip, P. A., Horsfall, P., and Goodman, N. D. Pyro: Deep universal probabilistic programming. *Journal of Machine Learning Research*, 20:28:1–28:6, 2019. URL <http://jmlr.org/papers/v20/18-403.html>.
- Blondel, M., Berthet, Q., Cuturi, M., Frostig, R., Hoyer, S., Llinares-López, F., Pedregosa, F., and Vert, J.-P. Efficient and Modular Implicit Differentiation. *Advances in Neural Information Processing Systems*, 35:5230–5242, 2022. URL [https://proceedings.neurips.cc/paper\\_files/paper/2022/file/228b9279ecf9bbafe582406850c57115-Paper-Conference.pdf](https://proceedings.neurips.cc/paper_files/paper/2022/file/228b9279ecf9bbafe582406850c57115-Paper-Conference.pdf).
- Bonne, L., Bontemps, S., Schneider, N., Clarke, S. D., Arzoumanian, D., Fukui, Y., Tachihara, K., Csengeri, T., Guesten, R., Ohama, A. et al. Formation of the musca filament: evidence for asymmetries in the accretion flow due to a cloud-cloud collision. *A&A*, 644:A27, December 2020a.
- Bonne, L., Schneider, N., Bontemps, S., Clarke, S. D., Gusdorf, A., Lehmann, A., Steinke, M., Csengeri, T., Kabanovic, S., Simon, R. et al. Dense gas formation in the musca filament due to the dissipation of a supersonic converging flow. *A&A*, 641:A17, September 2020b.
- Bovy, J. galpy: A python Library for Galactic Dynamics. *ApJS*, 216(2):29, February 2015. doi: 10.1088/0067-0049/216/2/29.
- Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S. et al. JAX: composable transformations of Python+NumPy programs, 2018. URL <http://github.com/google/jax>.
- Cabezas, A., Corenflos, A., Lao, J., and Louf, R. Blackjax: Composable Bayesian inference in JAX, 2024.
- Cantat-Gaudin, T., Fouesneau, M., Rix, H.-W., Brown, A. G. A., Castro-Ginard, A., Kostrzewa-Rutkowska, Z., Drimmel, R., Hogg, D. W., Casey, A. R.,

- Khanna, S. et al. An empirical model of the *Gaia* DR3 selection function. *A&A*, 669:A55, jan 2023. doi: 10.1051/0004-6361/202244784.
- Capitanio, L., Lallement, R., Vergely, J. L., Elyajouri, M., and Monreal-Ibero, A. Three-dimensional mapping of the local interstellar medium with composite data. *A&A*, 606:A65, October 2017. doi: 10.1051/0004-6361/201730831.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1):1–32, 2017. doi: 10.18637/jss.v076.i01. URL <https://www.jstatsoft.org/index.php/jss/article/view/v076i01>.
- Carrasco, J. M., Weiler, M., Jordi, C., Fabricius, C., De Angeli, F., Evans, D. W., van Leeuwen, F., Riello, M., and Montegriffo, P. Internal calibration of Gaia BP/RP low-resolution spectra. *A&A*, 652:A86, August 2021. doi: 10.1051/0004-6361/202141249.
- Cernicharo, J. The Physical Conditions of Low Mass Star Forming Regions. In Lada, C. J. and Kylafis, N. D., editors, *The Physics of Star Formation and Early Stellar Evolution*, volume 342 of *NATO Advanced Study Institute (ASI) Series C*, page 287, January 1991. doi: 10.1007/978-94-011-3642-6\_8.
- Chambers, K. C., Magnier, E. A., Metcalfe, N., Flewelling, H. A., Huber, M. E., Waters, C. Z., Denneau, L., Draper, P. W., Farrow, D., Finkbeiner, D. P. et al. The pan-starrs1 surveys, 2019.
- Chen, B. Q., Huang, Y., Yuan, H. B., Wang, C., Fan, D. W., Xiang, M. S., Zhang, H. W., Tian, Z. J., and Liu, X. W. Three-dimensional interstellar dust reddening maps of the Galactic plane. *MNRAS*, 483(4):4277–4289, March 2019. doi: 10.1093/mnras/sty3341.
- Choi, J., Dotter, A., Conroy, C., Cantiello, M., Paxton, B., and Johnson, B. D. Mesa Isochrones and Stellar Tracks (MIST). I. Solar-scaled Models. *ApJ*, 823(2):102, June 2016. doi: 10.3847/0004-637X/823/2/102.
- Corradi, W. J. B., Franco, G. A. P., and Knude, J. Physical association between the Southern Coalsack and the Chamaeleon-Musca dark clouds. *A&A*, 326:1215–1227, October 1997.
- Corradi, W. J. B., Franco, G. A. P., and Knude, J. Local interstellar medium kinematics towards the Southern Coalsack and Chamaeleon-Musca dark clouds. *MNRAS*, 347(4):1065–1083, February 2004. doi: 10.1111/j.1365-2966.2004.07219.x.
- Cox, N. L. J., Arzoumanian, D., André, P., Rygl, K. L. J., Prusti, T., Men’shchikov, A., Royer, P., Kóspál, A., Palmeirim, P., Ribas, A. et al. Filamentary structure and magnetic field orientation in musca. *A&A*, 590:A110, June 2016.
- De Angeli, F., Weiler, M., Montegriffo, P., Evans, D. W., Riello, M., Andrae, R., Carrasco, J. M., Busso, G., Burgess, P. W., Cacciari, C. et al. Gaia Data Release 3. Processing and validation of BP/RP low-resolution spectral data. *A&A*, 674:A2, June 2023. doi: 10.1051/0004-6361/202243680.

## BIBLIOGRAPHY

- de Bruijne, J. Gaia - taking the galactic census: Photometric instrument, 08 2009. URL [https://www.cosmos.esa.int/documents/29201/302420/IN\\_photometric\\_instrument.pdf/d7dd9dbb-5726-4810-90db-6b2e8c750fcd](https://www.cosmos.esa.int/documents/29201/302420/IN_photometric_instrument.pdf/d7dd9dbb-5726-4810-90db-6b2e8c750fcd).
- De G. Matthews, A. G., Van Der Wilk, M., Nickson, T., Fujii, K., Boukouvalas, A., León-Villagrà, P., Ghahramani, Z., and Hensman, J. Gpflow: A gaussian process library using tensorflow. *Journal of Machine Learning Research*, 18(1): 1299–1304, jan 2017. ISSN 1532-4435.
- DeepMind, Babuschkin, I., Baumli, K., Bell, A., Bhupatiraju, S., Bruce, J., Buchlovsky, P., Budden, D., Cai, T., Clark, A. et al. The DeepMind JAX Ecosystem, 2020. URL <http://github.com/google-deepmind>.
- Delchambre, L., Bailer-Jones, C. A. L., Bellas-Velidis, I., Drimmel, R., Garabato, D., Carballo, R., Hatzidimitriou, D., Marshall, D. J., Andrae, R., Dafonte, C. et al. Gaia Data Release 3. Apsis. III. Non-stellar content and source classification. *A&A*, 674:A31, June 2023. doi: 10.1051/0004-6361/202243423.
- Dharmawardena, T. E., Bailer-Jones, C. A. L., Foesneau, M., and Foreman-Mackey, D. Three-dimensional dust density structure of the Orion, Cygnus X, Taurus, and Perseus star-forming regions. *A&A*, 658:A166, February 2022. doi: 10.1051/0004-6361/202141298.
- Draine, B. T. Interstellar Dust Grains. *ARA&A*, 41:241–289, January 2003. doi: 10.1146/annurev.astro.41.011802.094840.
- Draine, B. T. Interstellar Dust Models and Evolutionary Implications. In Henning, T., Grün, E., and Steinacker, J., editors, *Cosmic Dust - Near and Far*, volume 414 of *Astronomical Society of the Pacific Conference Series*, page 453, December 2009. doi: 10.48550/arXiv.0903.1658.
- Draine, B. T. *Physics of the Interstellar and Intergalactic Medium*. 2011.
- Eberle, V., Frank, P., Stadler, J., Streit, S., and Enßlin, T. Efficient representations of spatially variant point spread functions with butterfly transforms in bayesian imaging algorithms. *Physical Sciences Forum*, 5(1), 2022. ISSN 2673-9984. doi: 10.3390/psf2022005033. URL <https://www.mdpi.com/2673-9984/5/1/33>.
- Eberle, V., Frank, P., Stadler, J., Streit, S., and Enßlin, T. Butterfly Transforms for Efficient Representation of Spatially Variant Point Spread Functions in Bayesian Imaging. *Entropy*, 25(4):652, April 2023. doi: 10.3390/e25040652.
- Edenhofer, G., Leike, R. H., Frank, P., and Enßlin, T. A. Sparse kernel gaussian processes through iterative charted refinement (icr), 2022. URL <https://arxiv.org/abs/2206.10634>.
- Edenhofer, G., Ensslin, T. A., Frank, P., Guardiani, M., and Roth, J. DENSe: Bayesian density estimation for Poisson data. *Astrophysics Source Code Library*, record ascl:2312.004, December 2023a.
- Edenhofer, G., Zucker, C., Frank, P., Saydjari, A. K., Speagle, J. S., Finkbeiner, D., and Enßlin, T. A Parsec-Scale Galactic 3D Dust Map out to 1.25 kpc from the Sun. *arXiv e-prints*, art. arXiv:2308.01295, August 2023b. doi: 10.48550/arXiv.2308.01295.

- Edenhofer, G., Alves, J., Zucker, C., and Enßlin, T. A. The “C”: The large Chameleon-Musca-Coalsack cloud. *arXiv e-prints*, art. arXiv:2404.09592, April 2024a. doi: 10.48550/arXiv.2404.09592.
- Edenhofer, G., Frank, P., Roth, J., Leike, R. H., Guerdi, M., Scheel-Platz, L. I., Guardiani, M., Eberle, V., Westerkamp, M., and Enßlin, T. A. Re-Envisioning Numerical Information Field Theory (NIFTy.re): A Library for Gaussian Processes and Variational Inference, 2024b.
- Enßlin, T. A. Information Theory for Fields. *Annalen der Physik*, 531(3):1800127, March 2019. doi: 10.1002/andp.201800127.
- Enßlin, T. A. Information Field Theory and Artificial Intelligence. *Entropy*, 24(3):374, March 2022. doi: 10.3390/e24030374.
- Enßlin, T. A. Information theory & information field theory, 03 2023. URL <https://wwwmpa.mpa-garching.mpg.de/~enssln/lectures/Files/ScriptIT&IFT.pdf>.
- ESA. Expected science performance for the nominal and extended mission based on gaia (e)dr3, 2022. URL <https://www.cosmos.esa.int/web/gaia/science-performance>.
- ESO. Secrets of a dark cloud: Unique infrared sofi images of barnard 68 probe the very first stages of star formation, 07 1999. URL <https://elt.eso.org/public/news/eso9934/>.
- Field, G. B., Goldsmith, D. W., and Habing, H. J. Cosmic-Ray Heating of the Interstellar Gas. *ApJ*, 155:L149, March 1969. doi: 10.1086/180324.
- Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306, March 2013. doi: 10.1086/670067.
- Foreman-Mackey, D., Yu, W., Yadav, S., Becker, M. R., Caplar, N., Huppenkothen, D., Killestein, T., Tronsgaard, R., Rashid, T., and Schmerler, S. dfm/tinygp: The tiniest of Gaussian Process libraries, January 2024.
- Frank, P. Geometric variational inference and its application to bayesian imaging. *Physical Sciences Forum*, 5(1), 2022. ISSN 2673-9984. doi: 10.3390/psf2022005006. URL <https://www.mdpi.com/2673-9984/5/1/6>.
- Frank, P., Leike, R., and Enßlin, T. A. Geometric variational inference. *Entropy*, 23(7):853, 07 2021. doi: 10.3390/e23070853.
- Gaia Collaboration. Gaia Data Release 3. Summary of the content and survey properties. *A&A*, 674:A1, June 2023a. doi: 10.1051/0004-6361/202243940.
- Gaia Collaboration. Gaia Data Release 3. Summary of the content and survey properties. *A&A*, 674:A1, June 2023b. doi: 10.1051/0004-6361/202243940.
- Galan, A., Caminha, G. B., Knollmüller, J., Roth, J., and Suyu, S. H. El Gordo needs El Anzuelo: Probing the structure of cluster members with multi-band extended arcs in JWST data. *arXiv e-prints*, art. arXiv:2402.18636, February 2024. doi: 10.48550/arXiv.2402.18636.
- Gardner, J., Pleiss, G., Weinberger, K. Q., Bindel, D., and Wilson, A. G. Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. In *Advances in Neural Information Processing Systems*, volume 31. Cur-

## BIBLIOGRAPHY

- ran Associates, Inc., 2018. URL <https://proceedings.neurips.cc/paper/2018/file/27e8e17134dd7083b050476733207ea1-Paper.pdf>.
- Goldsmith, P. F. Molecular Depletion and Thermal Balance in Dark Cloud Cores. *ApJ*, 557(2):736–746, August 2001. doi: 10.1086/322255.
- Gorski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., and Bartelmann, M. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *ApJ*, 622: 759–771, April 2005. doi: 10.1086/427976.
- Gorski, K. M., Wandelt, B. D., Hansen, F. K., Hivon, E., and Banday, A. J. The HEALPix Primer. *arXiv e-prints*, art. astro-ph/9905275, May 1999. doi: 10.48550/arXiv.astro-ph/9905275.
- Green, G. dustmaps: A Python interface for maps of interstellar dust. *The Journal of Open Source Software*, 3(26):695, Jun 2018. doi: 10.21105/joss.00695.
- Green, G. M., Schlafly, E. F., Finkbeiner, D., Rix, H.-W., Martin, N., Burgett, W., Draper, P. W., Flewelling, H., Hodapp, K., Kaiser, N. et al. Galactic reddening in 3D from stellar photometry - an improved map. *MNRAS*, 478 (1):651–666, July 2018. doi: 10.1093/mnras/sty1008.
- Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S., and Finkbeiner, D. A 3D Dust Map Based on Gaia, Pan-STARRS 1, and 2MASS. *ApJ*, 887(1):93, December 2019. doi: 10.3847/1538-4357/ab5362.
- Guardiani, M., Frank, P., Kostić, A., Edenhofer, G., Roth, J., Uhlmann, B., and Enßlin, T. Causal, Bayesian, & non-parametric modeling of the SARS-CoV-2 viral load distribution vs. patient’s age. *PLOS ONE*, 17(10):1–21, 10 2022. doi: 10.1371/journal.pone.0275011.
- Hacar, A., Kainulainen, J., Tafalla, M., Beuther, H., and Alves, J. The musca cloud: A 6 pc-long velocity-coherent, sonic filament. *A&A*, 587:A97, March 2016.
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J. et al. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020a. doi: 10.1038/s41586-020-2649-2.
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J. et al. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020b. doi: 10.1038/s41586-020-2649-2.
- Heath, M. T. *Scientific Computing: An Introductory Survey, Revised Second Edition*. SIAM, 2018. ISBN 1611975573. doi: 10.2514/1.J060261.
- Hensman, J., Fusi, N., and Lawrence, N. D. Gaussian processes for big data. In *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence*, page 282–290. AUAI Press, 2013. URL <https://proceedings.mlr.press/v37/wilson15.html>.
- Hensman, J., de G. Matthews, A. G., and Ghahramani, Z. Scalable variational gaussian process classification. In Lebanon, G. and Vishwanathan, S. V. N.,

- editors, *AISTATS*, volume 38 of *JMLR Workshop and Conference Proceedings*. JMLR.org, 2015. URL <http://dblp.uni-trier.de/db/conf/aistats/aistats2015.html#HensmanMG15>.
- Hoffman, M. D. and Gelman, A. The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15(1):1593–1623, jan 2014. ISSN 1532-4435.
- Hollenbach, D. and McKee, C. F. Molecule Formation and Infrared Emission in Fast Interstellar Shocks. III. Results for J Shocks in Molecular Clouds. *ApJ*, 342:306, July 1989. doi: 10.1086/167595.
- Hottier, C., Babusiaux, C., and Arenou, F. FEDReD. II. 3D extinction map with 2MASS and Gaia DR2 data. *A&A*, 641:A79, September 2020. doi: 10.1051/0004-6361/202037573.
- Hunt, E. L. and Reffert, S. Improving the open cluster census. II. An all-sky cluster catalogue with Gaia DR3. *A&A*, 673:A114, May 2023. doi: 10.1051/0004-6361/202346285.
- Hutschenreuter, S., Anderson, C. S., Betti, S., Bower, G. C., Brown, J. A., Brüggem, M., Carretti, E., Clarke, T., Clegg, A., Costa, A. et al. The Galactic Faraday rotation sky 2020. *A&A*, 657:A43, January 2022. doi: 10.1051/0004-6361/202140486.
- Hutschenreuter, S., Haverkorn, M., Frank, P., Raycheva, N. C., and Enßlin, T. A. Disentangling the Faraday rotation sky. *arXiv e-prints*, art. arXiv:2304.12350, April 2023. doi: 10.48550/arXiv.2304.12350.
- Inutsuka, S.-i., Inoue, T., Iwasaki, K., and Hosokawa, T. The formation and destruction of molecular clouds and galactic star formation. An origin for the cloud mass function and star formation efficiency. *A&A*, 580:A49, August 2015. doi: 10.1051/0004-6361/201425584.
- Kainulainen, J., Hacar, A., Alves, J., Beuther, H., Bouy, H., and Tafalla, M. Gravitational fragmentation caught in the act: the filamentary musca molecular cloud. *A&A*, 586:A27, February 2016.
- Kalberla, P. M. W. and Kerp, J. The Hi Distribution of the Milky Way. *ARA&A*, 47(1):27–61, September 2009. doi: 10.1146/annurev-astro-082708-101823.
- Kaminsky, A., Bonne, L., Arzoumanian, D., and Coudé, S. On the 3D curvature and dynamics of the musca filament. *ApJ*, 948(2):109, May 2023.
- Kidger, P. and Garcia, C. Equinox: neural networks in JAX via callable PyTrees and filtered transformations. *Differentiable Programming workshop at Neural Information Processing Systems 2021*, 2021.
- Kim, C.-G. and Ostriker, E. C. Momentum injection by supernovae in the interstellar medium. *ApJ*, 802(2):99, March 2015.
- Kissmann, R. PICARD: A novel code for the Galactic Cosmic Ray propagation problem. *Astroparticle Physics*, 55:37–50, March 2014. doi: 10.1016/j.astropartphys.2014.02.002.
- Klessen, R. S. and Glover, S. C. O. Physical Processes in the Interstellar Medium. *Saas-Fee Advanced Course*, 43:85, January 2016. doi: 10.1007/978-3-662-47890-5\_2.

## BIBLIOGRAPHY

- Knollmüller, J. and Enßlin, T. A. Metric Gaussian Variational Inference, January 2019.
- Knuth, K. H. and Skilling, J. Foundations of inference. *Axioms*, 1(1):38–73, 06 2012. ISSN 2075-1680. doi: 10.3390/axioms1010038.
- Koposov, S., Speagle, J., Barbary, K., Ashton, G., Bennett, E., Buchner, J., Scheffler, C., Cook, B., Talbot, C., Guillochon, J. et al. joshspeagle/dynesty: v2.1.3, October 2023.
- Kuhn, M. A., de Souza, R. S., Krone-Martins, A., Castro-Ginard, A., Ishida, E. E. O., Povich, M. S., Hillenbrand, L. A., and COIN Collaboration. SPICY: The Spitzer/IRAC Candidate YSO Catalog for the Inner Galactic Midplane. *ApJS*, 254(2):33, June 2021. doi: 10.3847/1538-4365/abe465.
- Kuhn, M. A., Saber, R., Povich, M. S., de Souza, R. S., Krone-Martins, A., Ishida, E. E. O., Zucker, C., Benjamin, R. A., Hillenbrand, L. A., Castro-Ginard, A. et al. Spectroscopic confirmation of a population of isolated, intermediate-mass YSOs. *arXiv [astro-ph.GA]*, June 2022.
- Kuhn, M. personal communication, 05 2023.
- Lallement, R., Capitanio, L., Ruiz-Dern, L., Danielski, C., Babusiaux, C., Vergely, L., Elyajouri, M., Arenou, F., and Leclerc, N. Three-dimensional maps of interstellar dust in the Local Arm: using Gaia, 2MASS, and APOGEE-DR14. *A&A*, 616:A132, August 2018. doi: 10.1051/0004-6361/201832832.
- Lallement, R., Babusiaux, C., Vergely, J. L., Katz, D., Arenou, F., Valette, B., Hottier, C., and Capitanio, L. Gaia-2MASS 3D maps of Galactic interstellar dust within 3 kpc. *A&A*, 625:A135, May 2019. doi: 10.1051/0004-6361/201834695.
- Lallement, R., Vergely, J. L., Babusiaux, C., and Cox, N. L. J. Updated Gaia-2MASS 3D maps of Galactic interstellar dust. *A&A*, 661:A147, May 2022. doi: 10.1051/0004-6361/202142846.
- Leike, R. H. and Enßlin, T. A. Charting nearby dust clouds using Gaia data only. *A&A*, 631:A32, November 2019. doi: 10.1051/0004-6361/201935093.
- Leike, R. H., Glatzle, M., and Enßlin, T. A. Resolving nearby dust clouds. *A&A*, 639:A138, July 2020. doi: 10.1051/0004-6361/202038169.
- Leike, R. H., Edenhofer, G., Knollmüller, J., Alig, C., Frank, P., and Enßlin, T. A. The Galactic 3D large-scale dust distribution via Gaussian process regression on spherical coordinates. *arXiv e-prints*, art. arXiv:2204.11715, April 2022. doi: 10.48550/arXiv.2204.11715.
- Leike, R. H. and Ensslin, T. A. Optimal belief approximation. *Entropy*, 19(8), 2017. ISSN 1099-4300. doi: 10.3390/e19080402.
- Lindegren, L. The Astrometric Instrument of Gaia: Principles. In Turon, C., O’Flaherty, K. S., and Perryman, M. A. C., editors, *The Three-Dimensional Universe with Gaia*, volume 576 of *ESA Special Publication*, page 29, January 2005.
- Liu, H., Ong, Y., Shen, X., and Cai, J. When gaussian process meets big data: A review of scalable gps. *IEEE Transactions on Neural Networks and Learning Systems*, 31:4405–4423, 2020. doi: 10.1109/TNNLS.2019.2957109.

- Luhman, K. L. Chamaeleon. *Handbook of Star Forming Regions*, 64(5):169–370, December 2008.
- Marton, G., Ábrahám, P., Rimoldini, L., Audard, M., Kun, M., Nagy, Z., Kóspál, Á., Szabados, L., Holl, B., Gavras, P. et al. Gaia Data Release 3. Validating the classification of variable young stellar object candidates. *A&A*, 674:A21, June 2023. doi: 10.1051/0004-6361/202244101.
- Mertsch, P. and Phan, V. H. M. Bayesian inference of three-dimensional gas maps. II. Galactic HI. *A&A*, 671:A54, March 2023. doi: 10.1051/0004-6361/202243326.
- Minniti, D., Lucas, P. W., Emerson, J. P., Saito, R. K., Hempel, M., Pietrukowicz, P., Ahumada, A. V., Alonso, M. V., Alonso-Garcia, J., Arias, J. I. et al. VISTA Variables in the Via Lactea (VVV): The public ESO near-IR variability survey of the Milky Way. *New A*, 15(5):433–443, July 2010. doi: 10.1016/j.newast.2009.12.002.
- Mizuno, A., Hayakawa, T., Yamaguchi, N., Kato, S., Hara, A., Mizuno, N., Yonekura, Y., Onishi, T., Kawamura, A., Tachihara, K. et al. A  $^{13}\text{CO}$  (J = 1-0) Search for Small Molecular Clouds toward the Chamaeleon-Musca Dark Cloud Complex. *ApJ*, 507(1):L83–L87, November 1998. doi: 10.1086/311662.
- Monsch, K., Lovell, J. B., Berghea, C. T., Edenhofer, G., Keating, G. K., Andrews, S. M., Bayyari, A., Drake, J. J., and Wilner, D. J. High-resolution Pan-STARRS and SMA observations of IRAS 23077+6707: A giant edge-on protoplanetary disk. *arXiv e-prints*, art. arXiv:2402.01941, February 2024. doi: 10.48550/arXiv.2402.01941.
- Montegriffo, P., De Angeli, F., Andrae, R., Riello, M., Pancino, E., Sanna, N., Bellazzini, M., Evans, D. W., Carrasco, J. M., Sordo, R. et al. Gaia Data Release 3: External calibration of BP/RP low-resolution spectroscopic data. *arXiv e-prints*, 674:arXiv:2206.06205, June 2022. doi: 10.48550/arXiv.2206.06205.
- Morningstar, W., Vikram, S., Ham, C., Gallagher, A., and Dillon, J. Automatic differentiation variational inference with mixtures. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pages 3250–3258. PMLR, 04 2021. URL <https://proceedings.mlr.press/v130/morningstar21b.html>.
- Nyman, L.-a. The southern coalsack. *Handbook of Star Forming Regions*, 64(5): 222–370, December 2008.
- O’Neill, T. J., Zucker, C., Goodman, A. A., and Edenhofer, G. The Local Bubble is a Local Chimney: A New Model from 3D Dust Mapping. *arXiv e-prints*, art. arXiv:2403.04961, March 2024. doi: 10.48550/arXiv.2403.04961.
- Paxton, B., Bildsten, L., Dotter, A., Herwig, F., Lesaffre, P., and Timmes, F. Modules for Experiments in Stellar Astrophysics (MESA). *ApJS*, 192(1):3, January 2011. doi: 10.1088/0067-0049/192/1/3.
- Paxton, B., Cantiello, M., Arras, P., Bildsten, L., Brown, E. F., Dotter, A., Mankovich, C., Montgomery, M. H., Stello, D., Timmes, F. X. et al. Modules for Experiments in Stellar Astrophysics (MESA): Planets, Oscillations, Rota-

## BIBLIOGRAPHY

- tion, and Massive Stars. *ApJS*, 208(1):4, September 2013. doi: 10.1088/0067-0049/208/1/4.
- Paxton, B., Marchant, P., Schwab, J., Bauer, E. B., Bildsten, L., Cantiello, M., Dessart, L., Farmer, R., Hu, H., Langer, N. et al. Modules for Experiments in Stellar Astrophysics (MESA): Binaries, Pulsations, and Explosions. *ApJS*, 220(1):15, September 2015. doi: 10.1088/0067-0049/220/1/15.
- Paxton, B., Schwab, J., Bauer, E. B., Bildsten, L., Blinnikov, S., Duffell, P., Farmer, R., Goldberg, J. A., Marchant, P., Sorokina, E. et al. Modules for Experiments in Stellar Astrophysics (MESA): Convective Boundaries, Element Diffusion, and Massive Star Explosions. *ApJS*, 234(2):34, February 2018. doi: 10.3847/1538-4365/aaa5a8.
- Paxton, B., Smolec, R., Schwab, J., Gaultschi, A., Bildsten, L., Cantiello, M., Dotter, A., Farmer, R., Goldberg, J. A., Jermyn, A. S. et al. Modules for Experiments in Stellar Astrophysics (MESA): Pulsating Variable Stars, Rotation, Convective Boundaries, and Energy Conservation. *ApJS*, 243(1):10, July 2019. doi: 10.3847/1538-4365/ab2241.
- Pelgrims, V., Ferrière, K., Boulanger, F., Lallement, R., and Montier, L. Modeling the magnetized Local Bubble from dust data. *A&A*, 636:A17, April 2020. doi: 10.1051/0004-6361/201937157.
- Phan, D., Pradhan, N., and Jankowiak, M. Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro. *arXiv e-prints*, art. arXiv:1912.11554, December 2019. doi: 10.48550/arXiv.1912.11554.
- Planck Collaboration. Planck 2013 results. XI. All-sky model of thermal dust emission. *A&A*, 571:A11, November 2014a. doi: 10.1051/0004-6361/201323195.
- Planck Collaboration. Planck intermediate results. XVII. Emission of dust in the diffuse interstellar medium from the far-infrared to microwave frequencies. *A&A*, 566:A55, June 2014b. doi: 10.1051/0004-6361/201323270.
- Pleiss, G., Gardner, J., Weinberger, K., and Wilson, A. G. Constant-time predictive distributions for Gaussian processes. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 4114–4123. PMLR, 07 2018. URL <https://proceedings.mlr.press/v80/pleiss18a.html>.
- Pleiss, G., Jankowiak, M., Eriksson, D., Damle, A., and Gardner, J. Fast matrix square roots with applications to gaussian processes and bayesian optimization. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H., editors, *Advances in Neural Information Processing Systems*, volume 33, pages 22268–22281. Curran Associates, Inc., 2020. URL <https://proceedings.neurips.cc/paper/2020/file/fcf55a303b71b84d326fb1d06e332a26-Paper.pdf>.
- Popescu, C. C. and Tuffs, R. J. The percentage of stellar light re-radiated by dust in late-type Virgo Cluster galaxies. *MNRAS*, 335(2):L41–L44, September 2002. doi: 10.1046/j.1365-8711.2002.05881.x.
- Posch, L., Miret-Roig, N., Alves, J., Rattenböck, S., Großschedl, J., Meingast,

- S., Zucker, C., and Burkert, A. The corona australis star formation complex is accelerating away from the galactic plane. *A&A*, 679:L10, November 2023.
- Queiroz, A. B. A., Anders, F., Chiappini, C., Khalatyan, A., Santiago, B. X., Nepal, S., Steinmetz, M., Gallart, C., Valentini, M., Dal Ponte, M. et al. StarHorse results for spectroscopic surveys + Gaia DR3: Chronochemical populations in the solar vicinity, the genuine thick disk, and young-alpha rich stars. *arXiv e-prints*, 673:arXiv:2303.09926, March 2023. doi: 10.48550/arXiv.2303.09926.
- Ratzenböck, S., Großschedl, J. E., Möller, T., Alves, J., Bomze, I., and Meingast, S. Significance mode analysis (SigMA) for hierarchical structures. An application to the Sco-Cen OB association. *A&A*, 677:A59, September 2023. doi: 10.1051/0004-6361/202243690.
- Reinecke, M. DUCC: Distinctly useful code collection, 2024. URL <https://gitlab.mpcdf.mpg.de/mtr/ducc>.
- Rezaei Kh., S., Bailer-Jones, C. A. L., Hanson, R. J., and Fouesneau, M. Inferring the three-dimensional distribution of dust in the Galaxy with a non-parametric method . Preparing for Gaia. *A&A*, 598:A125, February 2017. doi: 10.1051/0004-6361/201628885.
- Rezaei Kh., S. and Kainulainen, J. Three-dimensional Shape Explains Star Formation Mystery of California and Orion A. *ApJ*, 930(2):L22, May 2022. doi: 10.3847/2041-8213/ac67db.
- Rezaei Kh., S., Bailer-Jones, C. A. L., Hogg, D. W., and Schultheis, M. Detection of the Milky Way spiral arms in dust from 3D mapping. *A&A*, 618:A168, October 2018. doi: 10.1051/0004-6361/201833284.
- Rezaei Kh., S., Bailer-Jones, C. A. L., Soler, J. D., and Zari, E. Detailed 3D structure of Orion A in dust with Gaia DR2. *A&A*, 643:A151, November 2020. doi: 10.1051/0004-6361/202038708.
- Rezende, D. J. and Mohamed, S. Variational Inference with Normalizing Flows. In *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*, ICML'15, pages 1530–1538. JMLR.org, 2015. URL <http://proceedings.mlr.press/v37/rezende15.html>.
- Rossi, S., Heinonen, M., Bonilla, E., Shen, Z., and Filippone, M. Sparse gaussian processes revisited: Bayesian approaches to inducing-variable approximations. In Banerjee, A. and Fukumizu, K., editors, *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pages 1837–1845. PMLR, 04 2021. URL <https://proceedings.mlr.press/v130/rossi21a.html>.
- Roth, J., Arras, P., Reinecke, M., Perley, R. A., Westermann, R., and Enßlin, T. A. Bayesian radio interferometric imaging with direction-dependent calibration. *A&A*, 678:A177, October 2023a. doi: 10.1051/0004-6361/202346851.
- Roth, J., Li Causi, G., Testa, V., Arras, P., and Enßlin, T. A. Fast-cadence High-contrast Imaging with Information Field Theory. *The Astronomical Journal*, 165(3):86, March 2023b. doi: 10.3847/1538-3881/acabc1.

## BIBLIOGRAPHY

- Roth, J., Reinecke, M., and Edenhofer, G. JAXbind: Bind any function to JAX. *arXiv e-prints*, art. arXiv:2403.08847, March 2024. doi: 10.48550/arXiv.2403.08847.
- Roth, J., Frank, P., Bester, H. L., Smirnov, O., Westermann, R., and Enßlin, T. A. fast-resolve, a fast bayesian radio interferometric imaging framework. in preparation, 2024.
- Salvatier, J., Wiecki, T. V., and Fonnesbeck, C. Probabilistic programming in python using pymc3. *PeerJ Computer Science*, 2:e55, April 2016. ISSN 2376-5992. doi: 10.7717/peerj-cs.55.
- Saydjari, A. K., Schlafly, E. F., Lang, D., Meisner, A. M., Green, G. M., Zucker, C., Zelko, I., Speagle, J. S., Daylan, T., Lee, A. et al. The Dark Energy Camera Plane Survey 2 (DECaPS2): More Sky, Less Bias, and Better Uncertainties. *ApJS*, 264(2):28, February 2023. doi: 10.3847/1538-4365/aca594.
- Scheel-Platz, L. I., Knollmüller, J., Arras, P., Frank, P., Reinecke, M., Jüstel, D., and Enßlin, T. A. Multicomponent imaging of the Fermi gamma-ray sky in the spatio-spectral domain. *A&A*, 680:A2, December 2023. doi: 10.1051/0004-6361/202243819.
- Schlafly, E. F., Meisner, A. M., and Green, G. M. The unWISE Catalog: Two Billion Infrared Sources from Five Years of WISE Imaging. *ApJS*, 240(2):30, February 2019. doi: 10.3847/1538-4365/aafbea.
- Schönrich, R., Binney, J., and Dehnen, W. Local kinematics and the local standard of rest. *MNRAS*, 403(4):1829–1833, April 2010. doi: 10.1111/j.1365-2966.2010.16253.x.
- Selig, M., Bell, M. R., Junklewitz, H., Oppermann, N., Reinecke, M., Greiner, M., Pachajoa, C., and Ensslin, T. A. NIFTY: A versatile Python library for signal inference. *Astrophysics Source Code Library*, record ascl:1302.013, 02 2013.
- Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, S., Carpenter, J. M., Beichman, C., Capps, R., Chester, T., Elias, J. et al. The Two Micron All Sky Survey (2MASS). *AJ*, 131(2):1163–1183, February 2006. doi: 10.1086/498708.
- Snelson, E. and Ghahramani, Z. Sparse gaussian processes using pseudo-inputs. In Weiss, Y., Schölkopf, B., and Platt, J., editors, *Advances in Neural Information Processing Systems*, volume 18. MIT Press, 2005. URL <https://proceedings.neurips.cc/paper/2005/file/4491777b1aa8b5b32c2e8666dbe1a495-Paper.pdf>.
- Speagle, J. S. DYNesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. *MNRAS*, 493(3):3132–3158, April 2020. doi: 10.1093/mnras/staa278.
- Speagle, J. S., Zucker, C., Beane, A., Cargile, P. A., Dotter, A., Finkbeiner, D. P., Green, G. M., Johnson, B. D., Schlafly, E. F., Bonaca, A. et al. Deriving Stellar Properties, Distances, and Reddenings using Photometry and Astrometry with BRUTUS. in preparation, 2021.

- Steininger, T., Dixit, J., Frank, P., Greiner, M., Hutschenreuter, S., Knollmüller, J., Leike, R. H., Porqueres, N., Pumpe, D., Reinecke, M. et al. NIFTy 3 - Numerical Information Field Theory: A Python Framework for Multicomponent Signal Inference on HPC Clusters. *Annalen der Physik*, 531(3):1800290, March 2019. doi: 10.1002/andp.201800290.
- Tritsis, A., Bouzelou, F., Skalidis, R., Tassis, K., Enßlin, T., and Edenhofer, G. The Musca molecular cloud: The perfect ‘filament’ is still a sheet. *Monthly Notices of the Royal Astronomical Society*, 514(3):3593–3603, 06 2022. doi: 10.1093/mnras/stac1572.
- Tritsis, A. and Tassis, K. Magnetic seismology of interstellar gas clouds: Unveiling a hidden dimension. *Science*, 360(6389):635–638, May 2018.
- Tsouros, A., Bendre, A. B., Edenhofer, G., Enßlin, T., Frank, P., Mastorakis, M., and Pavlidou, V. Non-parametric Bayesian reconstruction of Galactic magnetic fields using Information Field Theory: The inclusion of line-of-sight information in ultra-high energy cosmic ray backtracking. *arXiv e-prints*, art. arXiv:2403.05531, March 2024a. doi: 10.48550/arXiv.2403.05531.
- Tsouros, A., Edenhofer, G., Enßlin, T., Mastorakis, M., and Pavlidou, V. Reconstructing Galactic magnetic fields from local measurements for backtracking ultra-high-energy cosmic rays. *A&A*, 681:A111, January 2024b. doi: 10.1051/0004-6361/202346423.
- Vergely, J. L., Lallement, R., and Cox, N. L. J. Three-dimensional extinction maps: Inverting inter-calibrated extinction catalogues. *A&A*, 664:A174, August 2022. doi: 10.1051/0004-6361/202243319.
- Wang, C., Huang, Y., Yuan, H., Zhang, H., Xiang, M., and Liu, X. The Value-added Catalog for LAMOST DR8 Low-resolution Spectra. *ApJS*, 259(2):51, April 2022. doi: 10.3847/1538-4365/ac4df7.
- Wang, K., Pleiss, G., Gardner, J., Tyree, S., Weinberger, K. Q., and Wilson, A. G. Exact gaussian processes on a million data points. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL <https://proceedings.neurips.cc/paper/2019/file/01ce84968c6969bdd5d51c5eeaa3946a-Paper.pdf>.
- Welling, C., Frank, P., Enßlin, T., and Nelles, A. Reconstructing non-repeating radio pulses with Information Field Theory. *Journal of Cosmology and Astroparticle Physics*, 2021(4):071, April 2021. doi: 10.1088/1475-7516/2021/04/071.
- Wesseling, P. *An Introduction to Multigrid Methods*. An Introduction to Multigrid Methods. R.T. Edwards, 2004. ISBN 9781930217089. URL <https://books.google.de/books?id=RoRGAAAAAYAAJ>.
- Westerkamp, M., Eberle, V., Guardiani, M., Frank, P., Platz, L., Arras, P., Knollmüller, J., Stadler, J., and Enßlin, T. First spatio-spectral Bayesian imaging of SN1006 in X-ray. *arXiv e-prints*, art. arXiv:2308.09176, August 2023. doi: 10.48550/arXiv.2308.09176.
- Wilson, A. G., Gilboa, E., Nehorai, A., and Cunningham, J. P. Fast kernel learning for multidimensional pattern extrapolation. In *Advances*

## BIBLIOGRAPHY

- in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014. URL <https://proceedings.neurips.cc/paper/2014/file/77369e37b2aa1404f416275183ab055f-Paper.pdf>.
- Wilson, A. G., Hu, Z., Salakhutdinov, R. R., and Xing, E. P. Stochastic variational deep kernel learning. In *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. URL <https://proceedings.neurips.cc/paper/2016/file/bcc0d400288793e8bdcd7c19a8ac0c2b-Paper.pdf>.
- Wilson, A. G. and Nickisch, H. Kernel interpolation for scalable structured gaussian processes (kiss-gp). In *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37, ICML'15*, page 1775–1784. JMLR.org, 2015.
- Wilson, A. G., Dann, C., and Nickisch, H. Thoughts on massively scalable gaussian processes. *arXiv*, abs/1511.01870, 2015. doi: 10.48550/ARXIV.1511.01870.
- Winston, E., Hora, J. L., and Tolls, V. A Census of Star Formation in the Outer Galaxy. II. The GLIMPSE360 Field. *AJ*, 160(2):68, August 2020. doi: 10.3847/1538-3881/ab99c8.
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., Ressler, M. E., Cutri, R. M., Jarrett, T., Kirkpatrick, J. D., Padgett, D., McMillan, R. S., Skrutskie, M. et al. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. *AJ*, 140(6):1868–1881, December 2010. doi: 10.1088/0004-6256/140/6/1868.
- Xiang, M., Rix, H.-W., Ting, Y.-S., Kudritzki, R.-P., Conroy, C., Zari, E., Shi, J.-R., Przybilla, N., Ramirez-Tannus, M., Tkachenko, A. et al. Stellar labels for hot stars from low-resolution spectra. I. The HotPayne method and results for 330 000 stars from LAMOST DR6. *A&A*, 662:A66, June 2022. doi: 10.1051/0004-6361/202141570.
- Yahia, H., Schneider, N., Bontemps, S., Bonne, L., Attuel, G., Dib, S., Ossenkopf, V., Turiel, A., Zebadua, A., Elia, D. et al. Description of turbulent dynamics in the interstellar medium: Multifractal-microcanonical analysis I. application to herchel observations of the musca filament. *A&A*, 649:A33, March 2021.
- Zhang, X., Green, G. M., and Rix, H.-W. Parameters of 220 million stars from Gaia BP/RP spectra. *arXiv e-prints*, 524(2):arXiv:2303.03420, March 2023. doi: 10.48550/arXiv.2303.03420.
- Zonca, A., Singer, L., Lenz, D., Reinecke, M., Rosset, C., Hivon, E., and Gorski, K. healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python. *The Journal of Open Source Software*, 4(35):1298, March 2019. doi: 10.21105/joss.01298.
- Zucker, C., Alves, J., Goodman, A., Meingast, S., and Galli, P. The Solar Neighborhood in the Age of Gaia. In Inutsuka, S., Aikawa, Y., Muto, T., Tomida, K., and Tamura, M., editors, *Protostars and Planets VII*, volume 534 of *Astronomical Society of the Pacific Conference Series*, page 43, July 2023. doi: 10.48550/arXiv.2212.00067.

- Zucker, C., Schlafly, E. F., Speagle, J. S., Green, G. M., Portillo, S. K. N., Finkbeiner, D. P., and Goodman, A. A. Mapping distances across the perseus molecular cloud using CO observations, stellar photometry, and gaia DR2 parallax measurements. *ApJ*, 869(1):83, December 2018.
- Zucker, C., Speagle, J. S., Schlafly, E. F., Green, G. M., Finkbeiner, D. P., Goodman, A. A., and Alves, J. A Large Catalog of Accurate Distances to Local Molecular Clouds: The Gaia DR2 Edition. *ApJ*, 879(2):125, July 2019. doi: 10.3847/1538-4357/ab2388.
- Zucker, C., Speagle, J. S., Schlafly, E. F., Green, G. M., Finkbeiner, D. P., Goodman, A., and Alves, J. A compendium of distances to molecular clouds in the star formation handbook, 2020.
- Zucker, C., Goodman, A., Alves, J., Bialy, S., Koch, E. W., Speagle, J. S., Foley, M. M., Finkbeiner, D., Leike, R., Enßlin, T. et al. On the Three-dimensional Structure of Local Molecular Clouds. *ApJ*, 919(1):35, September 2021. doi: 10.3847/1538-4357/ac1f96.
- Zucker, C., Goodman, A. A., Alves, J., Bialy, S., Foley, M., Speagle, J. S., Großschedl, J., Finkbeiner, D. P., Burkert, A., Khimey, D. et al. Star formation near the Sun is driven by expansion of the Local Bubble. *Nature*, 601(7893): 334–337, January 2022. doi: 10.1038/s41586-021-04286-5.